THE UNIVERSITY OF CALGARY

SOME TOPICS IN ROBUST ESTIMATION AND EXPERIMENTAL DESIGN

by

SHAWN X. LIU

A DISSERTATION
SUBMITTED TO THE FACULTY OF GRADUATE STUDIES
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE
DEGREE OF DOCTOR OF PHILOSOPHY

DEPARTMENT OF MATHEMATICS AND STATISTICS

CALGARY, ALBERTA, CANADA

JANUARY, 1994

ISBN 0-315-93938-9

Name Shawn X. Liu

Dissertation Abstracts International is arranged by broad, general subject categories. Please select the one subject which most nearly describes the content of your dissertation. Enter the corresponding four-digit code in the spaces provided.

Statistics | 0 4 6 3 | U·M·I

SUBJECT TERM | SUBJECT CODE

Subject Categories

THE HUMANITIES AND SOCIAL SCIENCES

COMMUNICATIONS AND THE ARTS
Architecture 0729
Art History 0377
Cinema 0900
Dance 0378
Fine Arts 0357
Information Science 0723
Journalism 0391
Library Science 0399
Mass Communications 0708
Music 0413
Speech Communication 0459
Theater 0465

EDUCATION
General 0515
Administration 0514
Adult and Continuing 0516
Agricultural 0517
Art 0273
Bilingual and Multicultural 0282
Business 0688
Community College 0275
Curriculum and Instruction 0727
Early Childhood 0518
Elementary 0524
Finance 0277
Guidance and Counseling 0519
Health 0680
Higher 0745
History of 0520
Home Economics 0278
Industrial 0521
Language and Literature 0279
Mathematics 0280
Music 0522
Philosophy of 0998
Physical 0523
Psychology 0525
Reading 0535
Religious 0527
Sciences 0714
Secondary 0533
Social Sciences 0534
Sociology of 0340
Special 0529
Teacher Training 0530
Technology 0710
Tests and Measurements 0288
Vocational 0747

LANGUAGE, LITERATURE AND LINGUISTICS
Language
- General 0679
- Ancient 0289
- Linguistics 0290
- Modern 0291

Literature
- General 0401
- Classical 0294
- Comparative 0295
- Medieval 0297
- Modern 0298
- African 0316
- American 0591
- Asian 0305
- Canadian (English) 0352
- Canadian (French) 0355
- English 0593
- Germanic 0311
- Latin American 0312
- Middle Eastern 0315
- Romance 0313
- Slavic and East European 0314

PHILOSOPHY, RELIGION AND THEOLOGY
Philosophy 0422
Religion
- General 0318
- Biblical Studies 0321
- Clergy 0319
- History of 0320
- Philosophy of 0322

Theology 0469

SOCIAL SCIENCES
American Studies 0323
Anthropology
- Archaeology 0324
- Cultural 0326
- Physical 0327

Business Administration
- General 0310
- Accounting 0272
- Banking 0770
- Management 0454
- Marketing 0338

Canadian Studies 0385
Economics
- General 0501
- Agricultural 0503
- Commerce-Business 0505
- Finance 0508
- History 0509
- Labor 0510
- Theory 0511

Folklore 0358
Geography 0366
Gerontology 0351
History
- General 0578
- Ancient 0579
- Medieval 0581
- Modern 0582
- Black 0328
- African 0331
- Asia, Australia and Oceania 0332
- Canadian 0334
- European 0335
- Latin American 0336
- Middle Eastern 0333
- United States 0337

History of Science 0585
Law 0398
Political Science
- General 0615
- International Law and Relations 0616
- Public Administration 0617

Recreation 0814
Social Work 0452
Sociology
- General 0626
- Criminology and Penology 0627
- Demography 0938
- Ethnic and Racial Studies 0631
- Individual and Family Studies 0628
- Industrial and Labor Relations 0629
- Public and Social Welfare 0630
- Social Structure and Development 0700
- Theory and Methods 0344

Transportation 0709
Urban and Regional Planning 0999
Women's Studies 0453

THE SCIENCES AND ENGINEERING

BIOLOGICAL SCIENCES
Agriculture
- General 0473
- Agronomy 0285
- Animal Culture and Nutrition 0475
- Animal Pathology 0476
- Food Science and Technology 0359
- Forestry and Wildlife 0478
- Plant Culture 0479
- Plant Pathology 0480
- Plant Physiology 0817
- Range Management 0777
- Wood Technology 0746

Biology
- General 0306
- Anatomy 0287
- Biostatistics 0308
- Botany 0309
- Cell 0379
- Ecology 0329
- Entomology 0353
- Genetics 0369
- Limnology 0793
- Microbiology 0410
- Molecular 0307
- Neuroscience 0317
- Oceanography 0416
- Physiology 0433
- Radiation 0821
- Veterinary Science 0778
- Zoology 0472

Biophysics
- General 0786
- Medical 0760

EARTH SCIENCES
Biogeochemistry 0425
Geochemistry 0996
Geodesy 0370
Geology 0372
Geophysics 0373
Hydrology 0388
Mineralogy 0411
Paleobotany 0345
Paleoecology 0426
Paleontology 0418
Paleozoology 0985
Palynology 0427
Physical Geography 0368
Physical Oceanography 0415

HEALTH AND ENVIRONMENTAL SCIENCES
Environmental Sciences 0768
Health Sciences
- General 0566
- Audiology 0300
- Chemotherapy 0992
- Dentistry 0567
- Education 0350
- Hospital Management 0769
- Human Development 0758
- Immunology 0982
- Medicine and Surgery 0564
- Mental Health 0347
- Nursing 0569
- Nutrition 0570
- Obstetrics and Gynecology 0380
- Occupational Health and Therapy 0354
- Ophthalmology 0381
- Pathology 0571
- Pharmacology 0419
- Pharmacy 0572
- Physical Therapy 0382
- Public Health 0573
- Radiology 0574
- Recreation 0575
- Speech Pathology 0460
- Toxicology 0383

Home Economics 0386

PHYSICAL SCIENCES
Pure Sciences
Chemistry
- General 0485
- Agricultural 0749
- Analytical 0486
- Biochemistry 0487
- Inorganic 0488
- Nuclear 0738
- Organic 0490
- Pharmaceutical 0491
- Physical 0494
- Polymer 0495
- Radiation 0754

Mathematics 0405
Physics
- General 0605
- Acoustics 0986
- Astronomy and Astrophysics 0606
- Atmospheric Science 0608
- Atomic 0748
- Electronics and Electricity 0607
- Elementary Particles and High Energy 0798
- Fluid and Plasma 0759
- Molecular 0609
- Nuclear 0610
- Optics 0752
- Radiation 0756
- Solid State 0611

Statistics 0463

Applied Sciences
Applied Mechanics 0346
Computer Science 0984
Engineering
- General 0537
- Aerospace 0538
- Agricultural 0539
- Automotive 0540
- Biomedical 0541
- Chemical 0542
- Civil 0543
- Electronics and Electrical 0544
- Heat and Thermodynamics 0348
- Hydraulic 0545
- Industrial 0546
- Marine 0547
- Materials Science 0794
- Mechanical 0548
- Metallurgy 0743
- Mining 0551
- Nuclear 0552
- Packaging 0549
- Petroleum 0765
- Sanitary and Municipal 0554
- System Science 0790

Geotechnology 0428
Operations Research 0796
Plastics Technology 0795
Textile Technology 0994

PSYCHOLOGY
General 0621
Behavioral 0384
Clinical 0622
Developmental 0620
Experimental 0623
Industrial 0624
Personality 0625
Physiological 0989
Psychobiology 0349
Psychometrics 0632
Social 0451

Nom ________________________________

Dissertation Abstracts International est organisé en catégories de sujets. Veuillez s.v.p. choisir le sujet qui décrit le mieux votre thèse et inscrivez le code numérique approprié dans l'espace réservé ci-dessous.

________________________________ ☐☐☐☐ U·M·I

SUJET — CODE DE SUJET

Catégories par sujets

HUMANITÉS ET SCIENCES SOCIALES

COMMUNICATIONS ET LES ARTS
- Architecture 0729
- Beaux-arts 0357
- Bibliothéconomie 0399
- Cinéma 0900
- Communication verbale 0459
- Communications 0708
- Danse 0378
- Histoire de l'art 0377
- Journalisme 0391
- Musique 0413
- Sciences de l'information 0723
- Théâtre 0465

ÉDUCATION
- Généralités 515
- Administration 0514
- Art 0273
- Collèges communautaires 0275
- Commerce 0688
- Économie domestique 0278
- Éducation permanente 0516
- Éducation préscolaire 0518
- Éducation sanitaire 0680
- Enseignement agricole 0517
- Enseignement bilingue et multiculturel 0282
- Enseignement industriel 0521
- Enseignement primaire. 0524
- Enseignement professionnel 0747
- Enseignement religieux 0527
- Enseignement secondaire 0533
- Enseignement spécial 0529
- Enseignement supérieur 0745
- Évaluation 0288
- Finances 0277
- Formation des enseignants 0530
- Histoire de l'éducation 0520
- Langues et littérature 0279
- Lecture 0535
- Mathématiques 0280
- Musique 0522
- Orientation et consultation 0519
- Philosophie de l'éducation 0998
- Physique 0523
- Programmes d'études et enseignement 0727
- Psychologie 0525
- Sciences 0714
- Sciences sociales 0534
- Sociologie de l'éducation 0340
- Technologie 0710

LANGUE, LITTÉRATURE ET LINGUISTIQUE
- Langues
  - Généralités 0679
  - Anciennes 0289
  - Linguistique 0290
  - Modernes 0291
- Littérature
  - Généralités 0401
  - Anciennes 0294
  - Comparée 0295
  - Mediévale 0297
  - Moderne 0298
  - Africaine 0316
  - Américaine 0591
  - Anglaise 0593
  - Asiatique 0305
  - Canadienne (Anglaise) 0352
  - Canadienne (Française) 0355
  - Germanique 0311
  - Latino-américaine 0312
  - Moyen-orientale 0315
  - Romane 0313
  - Slave et est-européenne 0314

PHILOSOPHIE, RELIGION ET THÉOLOGIE
- Philosophie 0422
- Religion
  - Généralités 0318
  - Clergé 0319
  - Études bibliques 0321
  - Histoire des religions 0320
  - Philosophie de la religion 0322
- Théologie 0469

SCIENCES SOCIALES
- Anthropologie
  - Archéologie 0324
  - Culturelle 0326
  - Physique 0327
- Droit 0398
- Économie
  - Généralités 0501
  - Commerce-Affaires 0505
  - Économie agricole 0503
  - Économie du travail 0510
  - Finances 0508
  - Histoire 0509
  - Théorie 0511
- Études américaines 0323
- Études canadiennes 0385
- Études féministes 0453
- Folklore 0358
- Géographie 0366
- Gérontologie 0351
- Gestion des affaires
  - Généralités 0310
  - Administration 0454
  - Banques 0770
  - Comptabilité 0272
  - Marketing 0338
- Histoire
  - Histoire générale 0578
  - Ancienne 0579
  - Médiévale 0581
  - Moderne 0582
  - Histoire des noirs 0328
  - Africaine 0331
  - Canadienne 0334
  - États-Unis 0337
  - Européenne 0335
  - Moyen-orientale 0333
  - Latino-américaine 0336
  - Asie, Australie et Océanie 0332
- Histoire des sciences 0585
- Loisirs 0814
- Planification urbaine et régionale 0999
- Science politique
  - Généralités 0615
  - Administration publique 0617
  - Droit et relations internationales 0616
- Sociologie
  - Généralités 0626
  - Aide et bien-àtre social 0630
  - Criminologie et établissements pénitentiaires 0627
  - Démographie 0938
  - Études de l' individu et de la famille 0628
  - Études des relations interethniques et des relations raciales 0631
  - Structure et développement social 0700
  - Théorie et méthodes. 0344
  - Travail et relations industrielles 0629
- Transports 0709
- Travail social 0452

SCIENCES ET INGÉNIERIE

SCIENCES BIOLOGIQUES
- Agriculture
  - Généralités 0473
  - Agronomie. 0285
  - Alimentation et technologie alimentaire 0359
  - Culture 0479
  - Élevage et alimentation 0475
  - Exploitation des péturages 0777
  - Pathologie animale 0476
  - Pathologie végétale 0480
  - Physiologie végétale 0817
  - Sylviculture et faune 0478
  - Technologie du bois 0746
- Biologie
  - Généralités 0306
  - Anatomie 0287
  - Biologie (Statistiques) 0308
  - Biologie moléculaire 0307
  - Botanique 0309
  - Cellule 0379
  - Écologie 0329
  - Entomologie 0353
  - Génétique 0369
  - Limnologie 0793
  - Microbiologie 0410
  - Neurologie 0317
  - Océanographie 0416
  - Physiologie 0433
  - Radiation 0821
  - Science vétérinaire 0778
  - Zoologie 0472
- Biophysique
  - Généralités 0786
  - Medicale 0760

SCIENCES DE LA TERRE
- Biogéochimie 0425
- Géochimie 0996
- Géodésie 0370
- Géographie physique 0368
- Géologie 0372
- Géophysique 0373
- Hydrologie 0388
- Minéralogie 0411
- Océanographie physique 0415
- Paléobotanique 0345
- Paléoécologie 0426
- Paléontologie 0418
- Paléozoologie 0985
- Palynologie 0427

SCIENCES DE LA SANTÉ ET DE L'ENVIRONNEMENT
- Économie domestique 0386
- Sciences de l'environnement 0768
- Sciences de la santé
  - Généralités 0566
  - Administration des hipitaux 0769
  - Alimentation et nutrition 0570
  - Audiologie 0300
  - Chimiothérapie 0992
  - Dentisterie 0567
  - Développement humain 0758
  - Enseignement 0350
  - Immunologie 0982
  - Loisirs 0575
  - Médecine du travail et thérapie 0354
  - Médecine et chirurgie 0564
  - Obstétrique et gynécologie 0380
  - Ophtalmologie 0381
  - Orthophonie 0460
  - Pathologie 0571
  - Pharmacie 0572
  - Pharmacologie 0419
  - Physiothérapie 0382
  - Radiologie 0574
  - Santé mentale 0347
  - Santé publique 0573
  - Soins infirmiers 0569
  - Toxicologie 0383

SCIENCES PHYSIQUES

Sciences Pures
- Chimie
  - Genéralités 0485
  - Biochimie 487
  - Chimie agricole 0749
  - Chimie analytique 0486
  - Chimie minérale 0488
  - Chimie nucléaire 0738
  - Chimie organique 0490
  - Chimie pharmaceutique 0491
  - Physique 0494
  - PolymÇres 0495
  - Radiation 0754
- Mathématiques 0405
- Physique
  - Genéralités 0605
  - Acoustique 0986
  - Astronomie et astrophysique 0606
  - Electronique et électricité 0607
  - Fluides et plasma 0759
  - Météorologie 0608
  - Optique 0752
  - Particules (Physique nucléaire) 0798
  - Physique atomique 0748
  - Physique de l'état solide 0611
  - Physique moléculaire 0609
  - Physique nucléaire 0610
  - Radiation 0756
- Statistiques 0463

Sciences Appliqués Et Technologie
- Informatique 0984
- Ingénierie
  - Généralités 0537
  - Agricole 0539
  - Automobile 0540
  - Biomédicale 0541
  - Chaleur et ther modynamique 0348
  - Conditionnement (Emballage) 0549
  - Génie aérospatial 0538
  - Génie chimique 0542
  - Génie civil 0543
  - Génie électronique et électrique 0544
  - Génie industriel 0546
  - Génie mécanique 0548
  - Génie nucléaire 0552
  - Ingénierie des systämes 0790
  - Mécanique navale 0547
  - Métallurgie 0743
  - Science des matériaux 0794
  - Technique du pétrole 0765
  - Technique minière 0551
  - Techniques sanitaires et municipales 0554
  - Technologie hydraulique 0545
- Mécanique appliquée 0346
- Géotechnologie 0428
- Matières plastiques (Technologie) 0795
- Recherche opérationnelle 0796
- Textiles et tissus (Technologie) 0794

PSYCHOLOGIE
- Généralités 0621
- Personnalité 0625
- Psychobiologie 0349
- Psychologie clinique 0622
- Psychologie du comportement 0384
- Psychologie du développement 0620
- Psychologie expérimentale 0623
- Psychologie industrielle 0624
- Psychologie physiologique 0989
- Psychologie sociale 0451
- Psychométrie 0632

THE UNIVERSITY OF CALGARY

FACULTY OF GRADUATE STUDIES

The undersigned certify that they have read, and recommended to the Faculty of Graduate Studies for acceptance, a dissertation entitled "Some Topics in Robust Estimation and Experimental Design", submitted by Shawn X. Liu in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

John R Collins
Supervisor, Dr. J.R. Collins
Dept. of Mathematics & Statistics

M. D. Burke
Dr. M.D. Burke
Dept. of Mathematics & Statistics

Dr. E.G. Enns
Dept. of Mathematics & Statistics

Gordon Sick
Dr. G. Sick
Dept. of Management

Stephen L Portnoy
External Examiner, Dr. S.L. Portnoy
University of Illinois

Date: January 10, 1994

ABSTRACT

Tukey (1960) considered the problem of robust estimation of a location parameter θ when the c.d.f. of the error is

$$F(x) = (1-\epsilon)H(x) + \epsilon H(\frac{x}{\sqrt{s}})$$

where ϵ and s could be any fixed number such that $0 < \epsilon < \frac{1}{2}$, and $s > 1$, and $H(x)$ is chosen to be the standard normal distribution. In Part I, we study a further problem of finding the bounds on asymptotic relative efficiencies of some robust estimators while the c.d.f. of the error has the above form with ϵ and s to be random variables.

In Part II, the problem of finding optimal designs against the possible model violation is considered. We confine ourself to the use of the least squares estimator, $\hat{y} = \underset{\sim}{\theta}^T \underset{\sim}{f}(x)$, of the true regression function $y(x)$. When the real regression model is

$$y_i = y(x_i) = \underset{\sim}{\theta}^T \underset{\sim}{f}(x_i) + \psi(x_i) + \epsilon_i,$$

then the mean squared error of $\hat{y}$ is

$$MSE(\hat{y}) = \mathrm{Var}(\xi) + \mathrm{Bias}(\psi, \xi),$$

where ξ is a design measure and ψ is a possible bias term. It was Box and Draper (1959) who first pointed out that the usual optimal design which minimizes Var(ξ) only is no longer optimal when the bias term is present. Several different approaches to the problem are discussed separately. They are summarized in the following topics: 1. Restricted optimal designs; 2. Bounded bias optimal designs; 3. Robust designs for some regression models with random bias.

ACKNOWLEDGEMENTS

I am indebted to Professor John R. Collins for his supervision, excellent guidance, and encouragement as well as the time he spent with me during the preparation of this dissertation.

I express deep gratitude to the members of my examining committee, Dr. S. L. Portnoy, Dr. M. D. Burke, Dr. E. G. Enns and Dr. G. A. Sick for all the help they gave me.

I express my special thanks to Dr. D. P. Wiens for his help and valuable suggestions.

The financial support of NSERC, the University of Calgary through GAT/GAR and that of Professor John Collins are gratefully acknowledged.

I extend my thanks to Mrs. Maria Dourado, Ms. Silvia Knodel, Mr. Jerry Kissinger, and Ms. Joanne Mellard who did an excellent job of typing this manuscript.

Dedicated to

my beloved parents

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

PART I
ROBUST ESTIMATION

Chapter 1

Introduction to Robust Estimation

The idea of robustness and robust methods probably dates back to the prehistory of statistics. Looking at the data and rechecking conspicuous observations is a step towards robustness; excluding highly deviant values is an informal robust procedure. However, the systematic study on the problem of robust estimation is a recent event. It was Tukey (1960) who, in summarizing earlier work of his group in the 1940s and 1950s, demonstrated the drastic nonrobustness of the mean and also investigated some useful robust alternatives. His work made robust estimation a general research area and broke the isolation of the early pioneers. Among a growing flood of papers were the first attempts at a manageable and rather realistic and comprehensive theory of robustness by Huber (1964, 1965, 1968) and Hampel (1968).

Huber's (1964) paper on "Robust estimation of a location parameter" formed the first basis for a theory of robust estimation. In that paper, Huber introduced a flexible class of estimators, called "M-estimators", which became a very useful tool, and he derived properties like consistency and asymptotic normality. Huber then introduced the "gross-error model:" instead of believing in a strict parametric model of the form $H(x-\theta)$ for known H, he assumes that a known fraction ϵ $(0 \le \epsilon < 1)$ of the data may consist of gross errors with an arbitrary unknown distribution $J(x-\theta)$. The distribution underlying the observations is thus $F(x-\theta) = (1-\epsilon)H(x-\theta)+\epsilon J(x-\theta)$.

This is the first time that a rather full kind of "neighborhood" for a strict parametric model is considered. Huber's aim is to optimize the worst that can happen over the neighborhood of the model, as measured by the asymptotic variance of the estimator. He uses the formalism of a two-person zero-sum game. Nature chooses an F from the neighborhood of the model, the statistician chooses an M-estimator via its ψ,and the gain for Nature and loss for the statistician is the asymptotic variance $V(\psi, F)$ which under the mild regularity conditions turns out to be $\int \psi^2 dF/(\psi' \, dF)^2$. Huber shows that under very general conditions there exists a saddle point of the game; in the gross-error model, it consists of what has been called Huber's least favorable distribution, which is normal in the middle and exponential in the tails, and of the famous Huber-estimator with $\psi(x) = \max\{-k, \min\{k, x\}\}$, as the maximum likelihood estimator for the least favorable distribution and the minimax strategy of the statistician.

Another important approach to the robust estimation theory is called the "infinitesimal approach" which was introduced by Hampel (1968, 1974). The infinitesimal approach is based on three central concepts : qualitative robustness, influence function, and breakdown point. Qualitative robustness is defined as equicontinuity of the distributions of the statistic as n changes; it is very closely related to continuity of the statistic viewed as a functional in the weak topology. The quantitative robustness information is provided by the influence function and derived quantities. The breakdown point is a simple quantitative global robustness measure. It is the distance from the model distribution beyond which the statistic becomes totally unreliable.

We are now going to discuss these concepts in detail. Let $X_1, ..., X_n$ be independent and identically distributed (i.i.d.) random variables with distribution function $F((x-\theta)/\sigma)$, where θ is the unknown location parameter and σ is a known scale parameter. We identify the sample $x_1, ..., x_n$ with its empirical distribution F_n. Formally F_n is given by $\left(\frac{1}{n}\right) \sum_{i=1}^{n} \Delta_{x_i}$, where Δ_x is the point mass 1 at x. As estimators of θ we consider real-valued statistics $T_n = T_n(x_1, ..., x_n) = T_n(F_n)$. Moreover, we consider estimators which are functionals (i.e. $T_n(F_n) = T(F_n)$ for all n and F_n) or can asymptotically be replaced by functionals. This means that we assume that there exists a functional T : domain $(T) \rightarrow \mathbf{R}$, such that

$$T_n\,(X_1, ..., X_n) \;\overrightarrow{n \rightarrow \infty}\; T(F)$$

in probability when the observations are i.i.d. according to the true distribution F in domain (T). We say that $T(F)$ is the asymptotic value of $\{T_n;\ n \geq 1\}$ at F. We often assume asymptotic normality, that is,

$$\mathcal{L}_F(\sqrt{n}\,[T_n - T(F)]) \;\underset{n \rightarrow \infty}{\xrightarrow{weakly}}\; N(0,\ V(T, F)),$$

where $\mathcal{L}_F$ means "the distribution of ... under F" and $V(T, F)$ is called the asymptotic variance of $\{T_n; n \geq 1\}$ at F.

We are now going to define an important concept, the so-called influence function, as the following:

Definition 1.1 *The influence function* (IF) *of* T *at* F *is given by*

$$IF(x;\ T, F) = \lim_{t \rightarrow 0} \frac{T((1-t)F + t\Delta_x) - T(F)}{t} \tag{1.1}$$

in those $x \in \mathcal{X}$ *where this limit exists and* $\mathcal{X}$ *is the sample space.*

There is a relation between the IF and the asymptotic variance $V(T, F)$. Under some regularity conditions, we have

$$V(T, F) = \int IF(x; T, F)^2\, dF(x). \tag{1.2}$$

For a detailed discussion, see Hampel (1986), Reeds (1976), Boos and Serfling (1980), and Fernholz (1983).

Apart from the asymptotic variance, there are some other important quantities related to the IF. We define the gross-error sensitivity of T at F by

$$\gamma^* = \gamma^*(T, F) = \sup_x \mid IF(x; T, F) \mid, \tag{1.3}$$

the supremum being taken over all x where $IF(x; T, F)$ exists. The gross-error sensitivity measures the worst influence which a small amount of contamination of fixed size can have on the value of the estimator. Therefore, it may be regarded as an upper bound on the (standardized) asymptotic bias of the estimator. It is a desirable feature that γ^* be finite, in which case we say that T is B-robust.

The local-shift sensitivity is defined by

$$\lambda^* = \lambda^*(T,F) = \sup_{x \neq y} \mid IF(y;T,F) - IF(x;T,F) \mid / \mid y - x \mid, \tag{1.4}$$

the smallest Lipschitz constant the IF obeys.

Moreover, we define the rejection point as the following:

$$\rho^* = \rho^*(T,F) = \inf\{\gamma > 0;\ IF(x;T,F) = 0 \ \text{ when } \ |x| > \gamma\}. \tag{1.5}$$

(If there exists no such γ, then $\rho^* = \infty$ by definition of the infimum). All observations farther away than ρ^* are rejected completely. Therefore, it is desirable that ρ^* is finite.

The gross-error sensitivity γ^* is an important robustness measure. But there is one limitation: it is an entirely local concept. Therefore, it must be complemented by a measure of the global reliability of the estimator, which describes up to what distance from the model distribution the estimator still gives some relevant information. First, we need a metric to measure the distance of two probability distributions. One choice is the Prohorov distance (Prohorov (1956)), of two probability distributions F and G, which is given by

$$\pi(F,G) := \inf\{\epsilon :\ F(A) \leq G(A^\epsilon) + \epsilon \text{ for all events } A\},$$

where A^ϵ is the set of all points whose distance from A is less than ϵ.

The important global robustness measure, the so-called breakdown point can be defined as follows.

Definition 1.2 *The breakdown point ϵ^* of the sequence of estimators $\{T_n;\ n \geq 1\}$ at F is defined by*

$$\begin{aligned} \epsilon^* := &\ \sup\{\epsilon \leq 1; \textit{there is a compact set } K_\epsilon \subsetneq \Theta \textit{ such that } \pi(F,G) < \epsilon \\ &\ \textit{implies } G(\{T_n \in K_\epsilon\}) \xrightarrow{n \to \infty} 1\} . \end{aligned} \tag{1.6}$$

If $\Theta = R$, *we obtain*

$$\begin{aligned} \epsilon^* : \; = \; & \sup_{n \to \infty} \{\epsilon \le 1 : \textit{there exists } \gamma_\epsilon \textit{ such that} \\ & \pi(F, G) < \epsilon \textit{ implies } G(\{\mid T_n \mid \le \gamma_\epsilon\}) \; \to 1\}. \end{aligned} \tag{1.7}$$

Note that one can also consider the gross-error breakdown point where $\pi(F,G) < \epsilon$ is replaced by $G \in \{(1-\epsilon)\ F + \epsilon H$ where H is arbitrary$\}$. Loosely speaking, this is the largest fraction of gross errors that never can carry the estimate over all bounds. This notion often leads to the same value of ϵ^*. Hampel (1971) also introduced some qualitative notions.

Definition 1.3 *We say that a sequence of estimators* $\{T_n;\ n \ge 1\}$ *is qualitatively robust at* F *if for every* $\epsilon > 0$ *there exists* $\delta > 0$ *such that for all* G *in* $\mathcal{F}(\mathcal{X})$ *and for all* n :

$$\pi(F,G) < \delta \Longrightarrow \pi(\mathcal{L}_F(T_n),\ \mathcal{L}_G(T_n)) < \epsilon,$$

where π *is the Prohorov distance,* $\mathcal{L}_F$ *means "the distribution of ... under* F*", and* $\mathcal{F}(\mathcal{X})$ *is the set of all the possible probability distributions on* $\mathcal{X}$.

Thus this definition describes equicontinuity of the distributions of T_n with respect to n.

We now discuss some examples of location estimators. Let $X_1, ..., X_n$ be the i.i.d. random variables with distribution function F, where F has density function f, and f is symmetric. Let us consider the sample mean $\bar{X}$, $\alpha-$trimmed mean $\bar{X}_\alpha$,and median M. Note that the $\alpha-$trimmed mean is obtained by removing both the $[\alpha n]$ smallest and the $[\alpha n]$ largest observations and calculates the mean of the remaining ones. It is easy to see that $\bar{X}$, $\bar{X}_\alpha$, and M can be represented as functionals by $T(F) = \int x dF$, $T(F) = \dfrac{1}{1-2\alpha} \displaystyle\int_\alpha^{1-\alpha} F^{-1}(t)dt$, and $T(F) = F^{-1}\left(\frac{1}{2}\right)$ respectively. According to Definition 1.1, one can find the influence functions of $\bar{X}$, $\bar{X}_\alpha$, and M at F which are the following:

$$
\begin{array}{llcl}
\text{(i)} & IF(x;\,\bar{X},F) & = & x; \\
\text{(ii)} & IF(x;\,\bar{X}_\alpha,F) & = & \begin{cases} \dfrac{1}{1-2\alpha}F^{-1}(\alpha) & x < F^{-1}(\alpha) \\ \dfrac{1}{1-2\alpha}x & F^{-1}(\alpha) \le x \le F^{-1}(1-\alpha); \\ \dfrac{1}{1-2\alpha}F^{-1}(1-\alpha) & x > F^{-1}(1-\alpha) \end{cases} \\
\text{(iii)} & IF(x;\,M,F) & = & \dfrac{sign(x)}{2f(0)} .
\end{array}
$$

Based on the influence functions, one can find the following conclusions:

(i) The arithmetic mean $\bar{X}$ is nowhere qualitatively robust, with
$\epsilon^* = 0,\ \gamma^* = \infty,\ \lambda^* = 1,$ and $\rho^* = \infty.$

(ii) The $\alpha-$ trimmed mean $\bar{X}_\alpha$ is qualitatively robust, with
$\epsilon^* = \alpha,\ \gamma^* = \dfrac{1}{1-2\alpha}F^{-1}(1-\alpha),\ \lambda^* = \dfrac{1}{1-2\alpha},$ and $\rho^* = \infty.$

(iii) The median M is qualitatively robust, with $\epsilon^* = \dfrac{1}{2},\ \gamma^* = \dfrac{1}{2f(0)},\ \lambda^* = \infty,$
and $\rho^* = \infty.$

As we mentioned earlier, Huber (1964) introduced the important concept of M-estimator. We have

Definition 1.4 *Any estimator T_n defined by a minimum problem of the form*

$$\sum \rho(X_i;\ T_n) = \min!, \tag{1.8}$$

or by an implicit equation

$$\sum \psi(X_i;\ T_n) = 0, \tag{1.9}$$

where ρ is an arbitrary function, $\psi(x;\theta) = (\partial/\partial\theta)\ \rho(x;\theta)$, is called an M-estimator.

Note that the choice of $\rho(x;\theta) = -\log\ f(x;\theta)$ gives the Maximum Likelihood estimator. Hence sometimes we call $M-$estimator an ML type estimator or generalized ML estimator.

Another type of estimator often considered in robust estimation theory is the so-called L-estimator.

Definition 1.5 *L-estimators are of the form*

$$T_n(X_1, ..., X_n) = \sum_{i=1}^{n} a_i \, X_{(i)}, \tag{1.10}$$

where $X_{(1)}, ..., X_{(n)}$ are the ordered samples and the a_i's are some coefficients.

The name "L-estimators" comes from "linear combinations of order statistics."

The famous Huber estimator which is defined by $\psi_b(x) = \min\{b, \ \max\{x, -b\}\}$ is an example of $M-$estimator, and $\alpha-$trimmed mean $\bar{X}_\alpha$ is an example of L-estimator. The sample mean $\bar{X}$ and median M may serve as examples for both $M-$ and $L-$estimators.

There are some other types of estimators being considered in robust estimation theory. For details, see Hampel (1986), and Andrews, Bickel, Hampel, Rogers and Tukey (1972).

There is an important topic in robust estimation theory, the problem of finding maximum asymptotic variances of different estimators. Collins (1976, 1977, 1986, and 1991) sufficiently studied the asymptotic variances of many robust estimators under symmetric or asymmetric contaminations. For instance, Collins (1986) considered the following problem arising in robust estimation theory: Find the maximum asymptotic variance of α-trimmed mean used to estimate an unknown location parameter when the error distribution is subject to asymmetric contamination. The model for the error distribution is $F = (1 - \epsilon) \ F_0 + \epsilon \ G$, where F_0 is a known distribution symmetric about 0, ϵ is fixed proportion of contamination, and G is an unknown and asymmetric distribution. Under some assumptions, he found that the maximal asymptotic variance is obtained when G places mass 1 at either $+\infty$ or $-\infty$.

The problems that are relevant to the asymptotic variance have been studied not only for the location estimators but also for the scale estimators as well. Moreover, the problems have also been discussed under some contamination models other than the gross-error neighbourhoods. Some examples are the Kolmogorov neighbourhood which is defined by $K_\epsilon(G) = \{F : \sup_{-\infty<x<\infty} \mid F(x) - G(x) \mid \leq \epsilon\}$ in which ϵ and G are known and fixed, and the Lévy neighbourhood which is defined by $L_{\epsilon,\delta}(G) = \{F : G(x - \delta) - \epsilon \leq F(x) \leq G(x + \delta) + \epsilon \text{ for all } x\}$. Some discussions can be found, for example in Wiens (1986), Collins and Wiens (1989), and Wiens and Wu (1991).

Portnoy (1977 and 1979) studied the problem of robust estimation in dependent situations.

Instead of studying the problem of finding the maximum asymptotic variance for each individual estimator, DasGupta (1990) discussed the bounds of asymptotic relative efficiencies (ARE) for a pair of estimators under some contamination models.

There are many papers about robust estimation theory. Most of the topics can be found in Huber (1981), Hampel (1986), and the references cited therein. Here we only present some basic concepts and the minimum amount of material which is relevant to the problem we are interested in, the problem of finding the bounds on asymptotic relative efficiencies of some robust estimators under random contaminations. The problem is an extension of the work of DasGupta (1990).

Chapter 2

Bounds on Asymptotic Relative Efficiencies of some Robust Estimators under Random Contaminations

2.1 Introduction

Tukey (1960) considered the problem of robust estimation of a location parameter θ where the cumulative distribution function (cdf) of the error is the following

$$F(x) = (1-\epsilon)\Phi(x) + \epsilon\Phi\left(\frac{x}{\sqrt{s}}\right),$$

where ϵ and s could be any fixed numbers such that $0 < \epsilon < \frac{1}{2}$ and $s > 1$, and $\Phi(x)$ is the standard normal distribution function.

The idea has been developed in many different ways by many authors, especially by Huber (1964). DasGupta (1990) proposed the generalized model as follows:

$$F(x) = \int\int \left[(1-\epsilon)\,H(x) + \epsilon H\left(\frac{x}{\sqrt{s}}\right)\right]\, dG_1(\epsilon)\, dG_2(s),$$

where ϵ, s are taken to be random, but with known expectations $E(\epsilon) = \epsilon_0$ and $E(s) = s_0$. Moreover, he assumed that ϵ and s are independent. He discussed the upper and lower bounds of Asymptotic Relative Efficiencies (ARE) of the Hodges-Lehmann estimator W, the median M, and the $\alpha-$trimmed mean $\bar{X}_\alpha$ with respect to the sample mean $\bar{X}$ over the class of distributions

$$\begin{aligned}\mathcal{F}_0 &= \left\{F(x): \ F(x) = \int\int[(1-\epsilon)\ H(x) + \epsilon H\left(\frac{x}{\sqrt{s}}\right)]\ dG_1(\epsilon)\ dG_2(s)\right\} \\ &= \left\{F(x): \ F(x) = (1-\epsilon_0)\ H(x) + \epsilon_0 \int H\left(\frac{x}{\sqrt{s}}\right)\ dG_2(s)\right\},\end{aligned} \tag{2.1.1}$$

where $\epsilon_0 = \int \epsilon dG_1(\epsilon)$ and $G_2[s_1, \infty) = 1$ with known constants $s_1 > 1$ and $s_0 = \int s dG_2(s)$.

In this chapter, we discuss a problem similar to that of DasGupta (1990) but with a different consideration for $\mathcal{F}$. We are looking for the bounds of ARE among the location estimators sample mean $\bar{X}$, $\alpha-$trimmed mean $\bar{X}_\alpha$, and median M over the scale mixing random contaminated class

$$\mathcal{F} = \left\{F(x): \ F(x) = (1-\epsilon)H(x) + \epsilon \int \ H(\frac{x}{\sqrt{s}})\ dG(s)\right\}, \tag{2.1.2}$$

where s is random with distribution function $G(s)$ such that $G[s_1, \ s_2] = 1$ for some fixed numbers s_1, s_2, and $1 \le s_1 < s_2 < \infty$. Here ϵ is any fixed number such that $0 < \epsilon < \frac{1}{2}$, and $H(x)$ is absolutely continuous.

In (2.1.1), DasGupta assumed that ϵ and s are independent and $\int \epsilon dG_1(\epsilon) = \epsilon_0$ is known. In this situation, ϵ plays no role in finding the bounds of ARE among $\bar{X}$, $\bar{X}_\alpha$, and M. The bounds only depend on the value ϵ_0 and $G_2(s)$. Hence, in (2.1.2), we treat ϵ as any fixed number between 0 and $\frac{1}{2}$ rather than a random variable. On the other hand, it seems more reasonable to assume that $G[s_1, s_2] = 1$ for some $1 \le s_1 < s_2 < \infty$ rather than $G_2[s_1, \infty) = 1$ with known constant $s_0 = \int s dG_2(s)$. Therefore, we suggest the scale mixing random contaminated class as we indicated in (2.1.2). From (1.2), we know that $V(T, F) = \int IF(x; \ T, F)^2\ dF(x)$. Hence $e(T, S, F) = V(S, F)/V(T, F)$ is the asymptotic relative efficiency of a pair of estimators $\{T_n : \ n \ge 1\}$ and $\{S_n : n \ge 1\}$.

Let F be the cumulative distribution function (cdf) of the error, where $F \in \mathcal{F}$. The main purpose of this chapter is looking for the bounds on $e(M, \bar{X}, F)$, $e(\bar{X}_\alpha, \bar{X}, F)$, and $e(M, \bar{X}_\alpha, F)$ over the class of distributions $\mathcal{F}$. i.e. we are looking for the following quantities:

$$\begin{aligned} e_*(M,\bar{X}) &= \inf_{F\in\mathcal{F}} e(M,\bar{X},F), & e^*(M,\bar{X}) &= \sup_{F\in\mathcal{F}} e(M,\bar{X},F); \\ e_*(\bar{X}_\alpha,\bar{X}) &= \inf_{F\in\mathcal{F}} e(\bar{X}_\alpha,\bar{X},F), & e^*(\bar{X}_\alpha,\bar{X}) &= \sup_{F\in\mathcal{F}} e(\bar{X}_\alpha,\bar{X},F); \\ e_*(M,\bar{X}_\alpha) &= \inf_{F\in\mathcal{F}} e(M,\bar{X}_\alpha,F), & e^*(M,\bar{X}_\alpha) &= \sup_{F\in\mathcal{F}} e(M,\bar{X}_\alpha,F). \end{aligned}$$

We define

$$\mathcal{G} = \{G(s): \quad G[s_1, s_2] = 1 \qquad 1 \le s_1 < s_2 < \infty\}. \tag{2.1.3}$$

It is clear that the asymptotic relative efficiencies eventually depend on G, since $F(x) = (1-\epsilon)H(x) + \epsilon \int H\left(\frac{x}{\sqrt{s}}\right) dG(s)$ where $H(x)$ is a known function. Therefore, we have

$$\begin{aligned} e_*(M,\bar{X}) &= \inf_{G\in\mathcal{G}} e(M,\bar{X},G), & e^*(M,\bar{X}) &= \sup_{G\in\mathcal{G}} e(M,\bar{X},G); \\ e_*(\bar{X}_\alpha,\bar{X}) &= \inf_{G\in\mathcal{G}} e(\bar{X}_\alpha,\bar{X},G), & e^*(\bar{X}_\alpha,\bar{X}) &= \sup_{G\in\mathcal{G}} e(\bar{X}_\alpha,\bar{X},G); \\ e_*(M,\bar{X}_\alpha) &= \inf_{G\in\mathcal{G}} e(M,\bar{X}_\alpha,G), & e^*(M,\bar{X}_\alpha) &= \sup_{G\in\mathcal{G}} e(M,\bar{X}_\alpha,G). \end{aligned}$$

In section 2.3, we find the explicit solutions to $e_*(M,\bar{X})$ and $e^*(M,\bar{X})$. Let $\mathcal{E} = \{\Delta_s : s_1 \le s \le s_2\}$ and $\mathcal{H} = \{\lambda\Delta s_2 + (1-\lambda)\Delta s_1 : 0 \le \lambda \le 1\}$. We find that

$$e_*(M,\bar{X}) = \inf_{G\in\mathcal{E}} e(M,\bar{X},G) = e(M,\bar{X},\Delta s_1)$$

and

$$e_*(M,\bar{X}) = \sup_{G\in\mathcal{H}} e(M,\bar{X},\mathcal{G}) = e(M,\bar{X},G^*)$$

where $G^* = \Delta s_2$ or $G^* = \lambda\Delta s_2 + (1-\lambda)\Delta s_1$ for some $\lambda \in (0,1)$.

In section 2.4, we study the bounds on $e(\bar{X}_\alpha,\bar{X},G)$ and $e(M,\bar{X}_\alpha,G)$. We indicate that the bounds are located within certain ranges. Let $F^{-1}(1-\alpha) = \gamma$, i.e.,

$$F(\gamma) = (1-\epsilon)H(\gamma) + \epsilon \int H(\frac{\gamma}{\sqrt{s}})dG(s) = 1-\alpha.$$

We define $\mathcal{G}_\gamma = \{G : G \in \mathcal{G}, F^{-1}(1-\alpha) = \gamma\}$, $\mathcal{G}_n = \{G : G \in \mathcal{G}, \mathrm{Card}\sigma(G) \le n\}$, and $\mathcal{T} = \{\lambda\Delta_{t_1} + (1-\lambda)\Delta_{t_2} : t_1, t_2 \in [s_1, s_2], 0 \le \lambda \le 1\}$.

We find that

$$\sup_{G\in\mathcal{G}_n} e(\bar{X}_\alpha, \bar{X}, G) \leq e^*(\bar{X}_\alpha, \bar{X}) = \max_{\gamma_1\leq\gamma\leq\gamma_2} \sup_{G\in\mathcal{G}_\gamma} e(\bar{X}_\alpha, \bar{X}, G) \leq \max_{\gamma_1\leq\gamma\leq\gamma_2} \sup_{G\in\mathcal{T}} e(\bar{X}_\alpha, \bar{X}, G),$$

where γ_1 and γ_2 are the lower and upper bounds of γ.

The situations of $e_*(\bar{X}_\alpha, \bar{X})$, $e^*(M, \bar{X}_\alpha)$, and $e_*(M, \bar{X}_\alpha)$ are similar to the case of $e^*(\bar{X}_\alpha, \bar{X})$. A general result is given by Theorem 2.4.3. some special results are also found when $h(x) = \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$ and $h(x) = \frac{1}{2}e^{-|x|}$. These results are presented through Theorem 2.4.4 to Theorem 2.4.7.

In section 2.2, we proved some preliminary results which will be used to find these bounds. Some numerical results and comments are presented in Section 2.5.

2.2 Preliminaries

In this section, we are going to present some preliminaries which will be used in later sections to find the bounds of ARE among the sample mean $\bar{X}$, $\alpha-$trimmed mean $\bar{X}_\alpha$, and median M.

Let T be the functional defined on $\mathcal{G}$, $T:\ \mathcal{G} \to \mathbb{R}^2$, by

$$T(G) = (\int h_1(s)\, dG(s),\ \int h_2(s)\, dG(s)),$$

where $h_1(s)$ and $h_2(s)$ are bounded and continuous functions on $[s_1, s_2]$.

Note that the extreme points of $\mathcal{G}$ is the set

$$\mathcal{E} = \{\Delta_s :\ s_1 \le s \le s_2\},$$

where Δ_s is the distribution function that put all its mass at s. The image of $\mathcal{E}$ under T is the following:

$$\begin{aligned} T(\mathcal{E}\,) &= \{(\textstyle\int h_1(s) d\Delta_s,\ \int h_2(s) d\Delta_s) : s_1 \le s \le s_2\} \\ &= \{(h_1(s),\ h_2(s)) : s_1 \le s \le s_2\} \\ &:= S. \end{aligned}$$

Let $\hat{S}$ be the convex hull of S, i.e., the smallest convex set containing S or the set of all convex combinations of S. Furthermore, let us define P to be the set of convex combinations of any two points of S, i.e.,

$$P = \{\lambda p_1 + (1-\lambda)p_2\ :\ p_1, p_2 \in S,\ 0 \le \lambda \le 1\}\,, \tag{2.2.1}$$

and let

$$\mathcal{T} = \{\lambda \Delta_{t_1} + (1-\lambda)\Delta_{t_2}\ :\ t_1, t_2 \in [s_1, s_2],\quad 0 \le \lambda \le 1\}\,. \tag{2.2.2}$$

Under the notations and the assumptions we made above, we have

Theorem 2.2.1 $P = T(\mathcal{T}) = T(\mathcal{G}) = \hat{S}$.

Some well known results are needed to prove Theorem 2.2.1. We state them as the following two Lemmas.

Lemma 2.2.2 *For every distribution function G in $\mathcal{G}$, there exists a sequence of step distribution functions G_n in $\mathcal{G}$ such that $G_n \Longrightarrow G$, where " $\Longrightarrow$ " refers to weak convergence.*

Proof: For given $G \in \mathcal{G}$, let

$$G_n = \sum_{i=1}^{n} p_{n_i} \Delta_{S_{n_i}},$$

where $S_{n_i} = s_1 + \dfrac{i(s_2 - s_1)}{n}$ and $p_{n_i} = G(S_{n_i}) - G(S_{n,i-1})$.
It is obvious that $G_n \Longrightarrow G$. □

Lemma 2.2.3 *(Billingsley 1986). The following two conditions are equivalent:*
(i) $G_n \Longrightarrow G$.
(ii) $\int h \, dG_n \to \int h \, dG$ for every bounded, continuous function h.

The proof of Theorem 2.2.1:
We prove the theorem by showing the following four steps:

(i) $P = T(\mathcal{T})$

For any $G \in \mathcal{T}$, we have $G = \lambda \Delta_{t_1} + (1 - \lambda)\Delta_{t_2}$ for some $t_1, t_2 \in [s_1, s_2]$ and $0 \leq \lambda \leq 1$, and

$$\begin{aligned} T(G) &= (\textstyle\int h_1(s)dG, \int h_2(s)dG) \\ &= (\lambda h_1(t_1) + (1-\lambda)h_1(t_2),\ \lambda h_2(t_1) + (1-\lambda)h_2(t_2)) \\ &= \lambda(h_1(t_1), h_2(t_1)) + (1-\lambda)(h_1(t_2),\ h_2(t_2)) \\ &= \lambda p_1 + (1-\lambda)p_2, \end{aligned}$$

where $p_1 = (h_1(t_1),\ h_2(t_1))$ and $p_2 = (h_1(t_2),\ h_2(t_2))$. Hence, we have $p_1, p_2 \in S$ and $T(G) = \lambda p_1 + (1-\lambda)p_2 \in P$. Therefore, we have $T(\mathcal{T}) \subseteq P$.

On the other hand, for any $\lambda p_1 + (1-\lambda)p_2 \in P$, there exist $t_1, t_2 \in [s_1, s_2]$ such that

$$p_1 = (\int h_1(s)d\Delta_{t_1}, \int h_2(s)d\Delta_{t_1}) = (h_1(t_1),\ h_2(t_1))$$

and

$$p_2 = (\int h_1(s)d\Delta_{t_2}, \int h_2(s)d\Delta_{t_2}) = (h_1(t_2),\ h_2(t_2)).$$

Let $G = \lambda\Delta_{t_1} + (1-\lambda)\Delta_{t_2}$. Then $G \in \mathcal{T}$ and $T(G) = \lambda p_1 + (1-\lambda)p_2 \in T(\mathcal{T})$. Hence, we have $P \subseteq T(\mathcal{T})$.

Combining the above two results we get $P = T(\mathcal{T})$.

(ii) $T(\mathcal{T}) \subseteq T(\mathcal{G})$.

This is obvious since $\mathcal{T} \subseteq \mathcal{G}$.

(iii) $T(\mathcal{G}) \subseteq \hat{S}$.

For any $G \in \mathcal{G}$, Lemma 2.2.2 implies that there exists $\{G_n\}$, such that $G_n \in \mathcal{G}$ and $G_n \Longrightarrow G$ where $G_n = \sum_{i=1}^{n} p_{n_i}\Delta_{S_{n_i}}$ with $\sum_{i=1}^{n} p_{n_i} = 1$ and $S_{n_i} \in [s_1, s_2]$. Especially, we can choose $S_{n_i} = s_1 + \dfrac{i(s_2 - s_1)}{n}$ and $p_{n_i} = G(S_{n_i}) - G(S_{n,i-1})$ $i = 1, ..., n$.

Let

$$q = (\int h_1(s)dG(s),\ \int h_2(s)dG(s))$$

and

$$\begin{aligned} q_n &= (\int h_1(s)dG_n(s),\ \int h_2(s)dG_n(s)) \\ &= (\sum_{i=1}^{n} p_{n_i}\ h_1(S_{ni}),\ \sum_{i=1}^{n} p_{n_i}\ h_2(S_{ni})) \\ &= \sum_{i=1}^{n} (p_{n_i}\ h_1(S_{ni}),\ p_{n_i} h_2(S_{n_i}) \\ &= \sum_{i=1}^{n} p_{n_i}\ (h_1(S_{ni}),\ h_2(S_{ni})). \end{aligned}$$

Since q_n is the convex combination of $(h_1(S_{ni}),\ h_2(S_{ni}))$, where $(h_1(S_{ni}), h_2(S_{ni})) \in S$. Therefore, $q_n \in \hat{S}$ for all n.

We assume that $h_1(s)$ and $h_2(s)$ are the bounded and continuous functions on $[s_1, s_2]$. The implication of Lemma 2.2.3 gives us

$$\begin{aligned} q_n &= (\int h_1(s)dG_n(s),\ \int h_2(s)dG_n(s)) \\ &\to (\int h_1(s)dG(s),\ \int h_2(s)dG(s)) = q \text{ as } n \to \infty. \end{aligned}$$

Note that $\hat{S}$ is a closed set. Hence $q_n \in \hat{S}$ and $q_n \to q$ implies $q \in \hat{S}$. We have proved $T(G) = q \in \hat{S}$ for any $G \in \mathcal{G}$. Therefore, we get $T(\mathcal{G}) \subseteq \hat{S}$.

(iv) $\hat{S} \subseteq P$.

The assumption that $h_1(s)$ and $h_2(s)$ are the bounded and continuous functions implies that S is a bounded continuous curve on $\mathbf{R}^2$. Hence, there exist real numbers a, b, c, and d such that the set

$$U = \{(x, y) \ : \ a \leq x \leq b, \ c \leq y \leq d\}$$

is the smallest rectangle containing S as well as $\hat{S}$.

For any $p_0 \in U \backslash S$, let l_p be the straight line that goes through the point p_0 and p, where $p \in S$. Varying p within S, we get a family of straight line denoted as

$$\mathcal{L} = \{l_p : \ p \in S, \ l_p \text{ is the line going through } p \text{ and } p_0\}.$$

If $p_0 \in \hat{S} \backslash S$, we claim that the following statement is true.

$S1$: There exists a point $p_1 \in S$, such that l_{p_1} intersect S at another point $p_2 \in S$ with p_0 in between.

The proof of S1:

Let $p_0 \in \hat{S} \backslash S$ and $p_0 = (x_0, y_0)$. We draw a vertical line l_0 going through p_0, i.e., $l_0 : x = x_0$. If $S1$ does not hold, l_0 must intersect S either above or below p_0. Without loss of generality, we assume that l_0 intersect S above p_0.

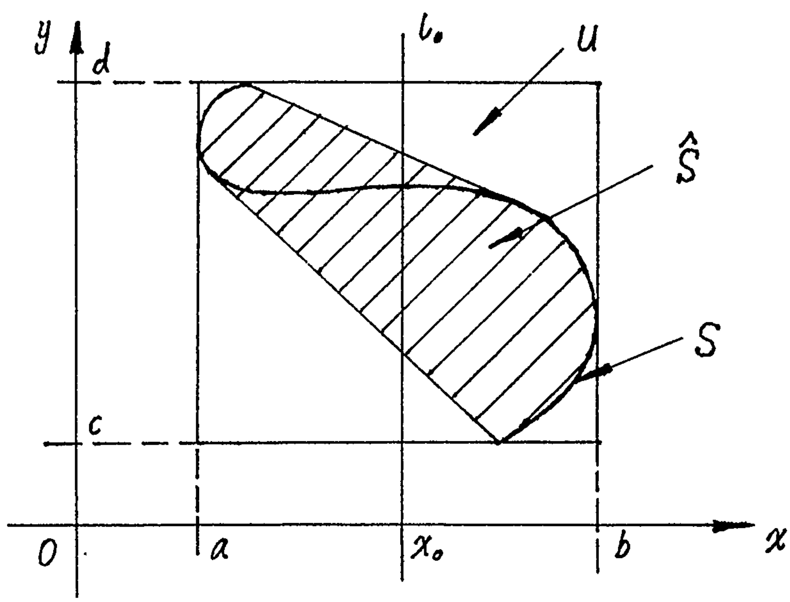

Figure 2.2.1

For the sake of argument, let us color the "upper half" (above p_0) of l_0 to be red and the "lower half" (below p_0) to be blue. We turn around l_0 clockwise with p_0 fixed. Stop turning when the blue half first hit S. Because S1 does not hold, the red half must stay away from S. If we slightly turn back l_0, there must be a position of l_0 such that the whole l_0 stay away from S i.e. $l_0 \cap S = \phi$. It is easy to see that there exists a convex set C containing S such that $C \cap l_0 = \phi$. Hence $\hat{S} \cap l_0 = \phi$, and $p_0 \notin \hat{S}$. This contradicts the fact that $p_0 \in \hat{S} \backslash S$. Therefore S1 is a true statement.

$S1$ indicates that, for any $p_0 \in \hat{S} \backslash S$ there exist $p_1$and $p_2 \in S$ such that p_0 is on the segment of straight line connecting p_1 to p_2. Therefore p_0 is the convex combination of p_1 and p_2, i.e., $p_0 = \lambda\, p_1 + (1-\lambda)p_2$ for some $0 < \lambda < 1$. On the other hand, for any $p \in S$, there exist p_1 and $p_2 \in S$ and with $\lambda = 0$ or $\lambda = 1$ such that $p = \lambda p_1 + (1-\lambda)p_2$ belong to P. This implies $\hat{S} \subseteq P$.

Combining (i), (ii), (iii), and (iv), we have proved Theorem 2.2.1. □

Under the assumption we made before, S is a continuous curve. If S is a convex curve, we will have a result which is simpler than Theorem 2.2.1. Let

$$P^{(1)} = \{\lambda p^{(1)} + (1-\lambda)p \;:\; p^{(1)} = (h_1(s_1), h_2(s_1)),\, p \in S,\; 0 \leq \lambda \leq 1\} \qquad (2.2.3)$$

and

$$\mathcal{T}^{(1)} = \{\lambda \Delta_{s_1} + (1-\lambda)\Delta_s \;:\; s \in [s_1, s_2],\;\; 0 \leq \lambda \leq 1\}. \qquad (2.2.4)$$

We have:

Theorem 2.2.4 *Let us maintain the same notations and assumptions as we made in Theorem 2.2.1. Let $P^{(1)}$ and $\mathcal{T}^{(1)}$ be defined as in (2.2.3) and (2.2.4). If S is a convex curve, then we have*

$$P^{(1)} = T(\mathcal{T}^{(1)}) = T(\mathcal{G}) = \hat{S}.$$

Proof: It is obvious that $T(\mathcal{T}^{(1)}) \subseteq T(\mathcal{G})$. The proof of $P^{(1)} = T(\mathcal{T}^{(1)})$ is similar to the proof of $P = T(\mathcal{T})$ in Theorem 2.2.1. Moreover, in Theorem 2.2.1, we have shown $T(\mathcal{G}) \subseteq \hat{S}$. Hence it is sufficient to show that $\hat{S} \subseteq P^{(1)}$ and this is obviously true by the convexity of S. □

Remark 1. Theorem 2.2.4 remains true if S is a concave curve.

Remark 2. It is obvious that

$$P^{(2)} = T(\mathcal{T}^{(2)}) = T(\mathcal{G}) = \hat{S},$$

when S is either convex or concave, where

$$P^{(2)} = \left\{\lambda p^{(2)} + (1-\lambda)p \ : p^{(2)} = (h_1(s_2),\ h_2(s_2))\,,\ p \in S,\ 0 \le \lambda \le 1\right\} \tag{2.2.5}$$

and

$$\mathcal{T}^{(2)} = \{\lambda\Delta_{s_2} + (1-\lambda)\Delta_s \ : \ s \in [s_1, s_2],\ \ 0 \le \lambda \le 1\}\,. \tag{2.2.6}$$

We are not sure whether the natural extension of Theorem 2.2.1 to the m-dimensional case is true or not. Let

$$T_m(G) = \left(\int h_1(s)dG(s), ..., \int h_m(s)dG(s)\right),$$

where $h_1(s), ..., h_m(s)$ are bounded and continuous functions on $[s_1, s_2]$. Let S_m be the image of $\mathcal{E}$ under T_m, i.e.,

$$\begin{aligned} S_m \ &:= T_m(\mathcal{E}) \\ &= \{(\textstyle\int h_1(s)d\Delta_s, ..., \int h_m(s)d\Delta_s) \ : \ s_1 \le s \le s_2\} \\ &= \{(h_1(s), ..., h_m(s)) \ : \ s_1 \le s \le s_2\}\,. \end{aligned}$$

Let $\hat{S}_m$ be the convex hull of S_m. Let P_m be the set of convex combinations of any m points of S_m, i.e.,

$$P_m = \left\{\sum_{i=1}^{m} \lambda_i p_i \ : \ p_i \in S_m,\ 0 \le \lambda_i \le 1,\ i = 1, ..., m,\ \sum_{i=1}^{m} \lambda_i = 1\right\}$$

and

$$\mathcal{T}_m = \left\{\sum_{i=1}^{m} \lambda_i \Delta_{t_i} \ : \ t_i \in [s_1, s_2],\ 0 \le \lambda_i \le 1,\ i = 1, ..., m,\ \sum_{i=1}^{m} \lambda_i = 1\right\}.$$

We make the following conjecture:

Conjecture 2.2.5 $P_m = T_m(\mathcal{T}_m) = T_m(\mathcal{G}) = \hat{S}_m$.

Some general results similar to Theorem 2.2.1 and Conjecture 2.2.5 can be found for example, in Rockafellar (1970) and some implications and applications to robust estimation can be found in Collins and Portnoy (1981). Comparing with those results, both the conditions and conclusions of Theorem 2.2.1 and Conjecture 2.2.5 are stronger.

2.3 Bounds on $e(M, \bar{X}, F)$

In this section, we are going to find the bounds on $e(M, \bar{X}, F)$, the bounds of the asymptotic relative efficiency of the pair of estimators $\bar{X}$ and M. We start from the concept of influence function. As we mentioned in Chapter 1, the IF of $\bar{X}$ and M are $IF(x; \bar{X}, F) = x$ and $IF(x; M, F) = sign(x)/2f(0)$ respectively. These are well-known results, but for completeness, we provide the detailed proof here. Let $X_1, ..., X_n \sim F(x)$ and X_i's are i.i.d. It is clear that the sample mean can be represented as a functional of F by $T(F) = \int x dF(x)$. The influence function of T at F can be calculated according to (1.1). We have

$$\begin{aligned}
IF(x; T, F) &= \lim_{t \to 0} \frac{T((1-t)F + t\Delta_x) - T(F)}{t} \\
&= \lim_{t \to 0} \frac{\int u d[(1-t)F + t\Delta_x](u) - \int u dF(u)}{t} \\
&= \lim_{t \to 0} \frac{(1-t)\int u dF(u) + t \int u d\Delta_x(u) - \int u dF(u)}{t} \\
&= \lim_{t \to 0} \frac{t \cdot x - t \cdot \int u dF(u)}{t} \\
&= x - \int u dF(u) = x
\end{aligned}$$

if $\int u dF(u) = 0$.

In the case we discussed in Section 2.1, $F(x)$ is symmetric. Hence, we have $IF(x; \bar{X}, F) = x$. Let us denote the asymptotic variance of $\bar{X}$ under F as $V(\bar{X}, F)$. By (1.2), we have

$$\begin{aligned}
V(\bar{X}, F) &= \int IF(x; \bar{X}, F)^2 dF(x) \\
&= \int x^2 dF(x) \\
&= (1-\epsilon) \int x^2 h(x) dx + \epsilon \int [x^2 \int h(\frac{x}{\sqrt{s}}) \cdot \frac{1}{\sqrt{s}} dG(s)] dx \\
&= (1-\epsilon) \int x^2 h(x) dx + \epsilon \int_{s_1}^{s_2} s [\int_{-\infty}^{\infty} y^2 h(y) dy] dG(s) \\
&= (1-\epsilon)\sigma_h^2 + \epsilon \sigma_h^2 \int s dG(s),
\end{aligned}$$

where $F \in \mathcal{F}$ and $\sigma_h^2 = \int x^2 h(x) dx$.

Similarly, the median M can be represented as a functional of F by $T(F) = F^{-1}(\frac{1}{2})$. The influence function of T under F is the following limit:

$$IF(x;T,F) = \lim_{t\to 0} \frac{[(1-t)F + t\Delta_x]^{-1}(\frac{1}{2}) - F^{-1}(\frac{1}{2})}{t}.$$

Case (i) $x < F^{-1}(\frac{1}{2})$

In this case, $[(1-t)F+t\Delta_x]^{-1}(\frac{1}{2})$ is that value of y for which $(1-t)F(y)+t\Delta_x(y) = \frac{1}{2}$. This solution will either be a value of $y < x$ for which $(1-t)F(y) = \frac{1}{2}$, or a value of $y > x$ for which $(1-t)F(y) + t = \frac{1}{2}$. If $y < x$, we have $F(y) \leq F(x) < \frac{1}{2}$. Hence $(1-t)F(y) = \frac{1}{2}$ is impossible for small $t > 0$. Therefore, we have $y > x$, and $(1-t)F(y) + t = \frac{1}{2}$. Take the derivative with respect to t at both sides and evaluate at $t = 0$, we get

$$(1-t)\cdot \frac{dF(y)}{dt} + F(y)\cdot(-1) = -1$$

and

$$(1-t)\cdot f(y)\cdot \frac{dy}{dt}\,|_{t=0} - F(y)\,|_{t=0} = -1 \Longrightarrow f(0)\frac{dy}{dt}\,|_{t=0} - \frac{1}{2} = -1.$$

Hence, we get $IF(x;\ M,F) = \frac{dy}{dt}\,|_{t=0} = -\frac{1}{2f(0)}$.

Case (ii) $x > F^{-1}(\frac{1}{2})$

In this case, we have $IF(x;\ M,F) = \frac{1}{2f(0)}$. The proof is similar to case (i).

We finally get

$$IF(x;\ M,F) = \frac{sign(x)}{2f(0)}.$$

Denote the asymptotic variance of M under F as $V(M,F)$. According to (1.2), we have

$$\begin{aligned} V(M,F) &= \int IF(x;M,F)^2 dF(x) \\ &= \frac{1}{4f^2(0)} \\ &= \frac{1}{4[(1-\epsilon)h(0) + \epsilon\int(h(\frac{0}{\sqrt{s}})\frac{1}{\sqrt{s}})dG(s)]^2} \\ &= \frac{1}{4h^2(0)[(1-\epsilon)+\epsilon\int\frac{1}{\sqrt{s}}dG(s)]^2}. \end{aligned}$$

Hence, the asymptotic relative efficiency of the pair of estimators $\bar{X}$ and M is

$$\begin{aligned} e(M,\bar{X},F) &= \frac{V(\bar{X},F)}{V(M,F)} \\ &= 4h^2(0)\cdot\sigma_h^2\cdot[(1-\epsilon)+\epsilon\textstyle\int sdG(s)]\cdot\left[(1-\epsilon)+\epsilon\textstyle\int\frac{1}{\sqrt{s}}dG(s)\right]^2. \end{aligned} \tag{2.3.1}$$

We are now going to find the bounds on $e(M,\bar{X},F)$ where $F\in\mathcal{F}$ or $G\in\mathcal{G}$. In light of Theorem 2.2.1, we consider the functional T from $\mathcal{G}$ to $\mathbf{R}^2$ as follows:

$$T(G)=\left(\int sdG(s),\quad \int\frac{1}{\sqrt{s}}dG(s)\right),$$

i.e., we choose $h_1(s)=s$ and $h_2(s)=\frac{1}{\sqrt{s}}$. We defne

$$T(\mathcal{E})=\{(s,\frac{1}{\sqrt{s}}):s_1\le s\le s_2\}:=S.$$

Then S is a convex curve on $\mathbf{R}^2$ and the convex hull of $S,\hat{S}$, is the shaded region in Figure 2.3.1.

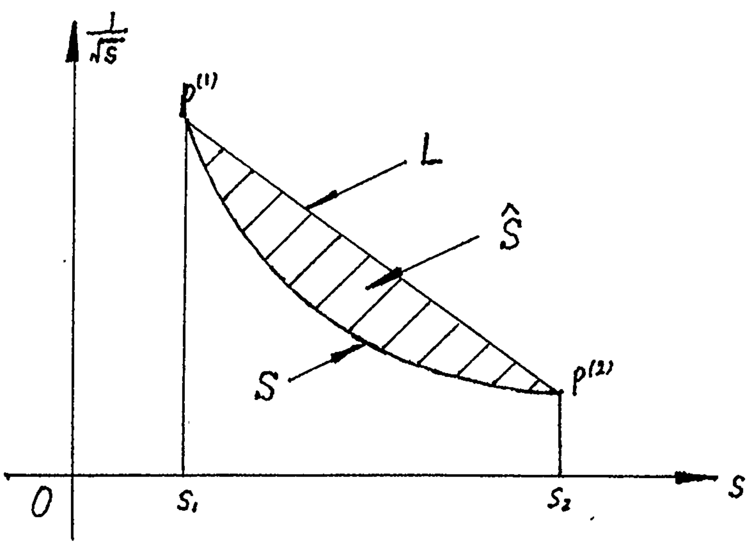

Figure 2.3.1

Because S is a convex curve, Theorem 2.2.4 can be used to solve our problem. Furthermore, it is clear that $e^*(M,\bar{X})$ can be achieved on the upper boundary of $\hat{S}$ and $e_*(M,\bar{X})$ can be achieved on the lower boundary of $\hat{S}$. Let

$$\mathcal{H}=\{G:G=\lambda\Delta_{s_2}+(1-\lambda)\Delta_{s_1},\ 0\le\lambda\le1\}.$$

Then the upper boundary of $\hat{S}$ is the image of $\mathcal{H}$ under T, i.e., $T(\mathcal{H}) = L$, where $L = \{p : p = \lambda p^{(2)} + (1-\lambda)p^{(1)}, 0 \le \lambda \le 1\}$ and the lower boundary of $\hat{S}$ is S which is the image of $\mathcal{E}$ under T. Hence, we have

$$\begin{aligned} e_*(M,\bar{X}) &= \inf_{\mathcal{F}} \ e(M,\bar{X},F) \\ &= \inf_{\mathcal{G}} \ e(M,\bar{X},G) \\ &= \inf_{\mathcal{E}} \ e(M,\bar{X},G) \\ &= \inf_{\mathcal{E}} 4h^2(0)\sigma_h^2[(1-\epsilon)+\epsilon \textstyle\int sdG(s)][(1-\epsilon)+\epsilon \textstyle\int \frac{1}{\sqrt{s}}dG(s)]^2 \\ &= \inf_{s_1 \le s \le s_2} 4h^2(0)\sigma_h^2[(1-\epsilon)+\epsilon\cdot s][(1-\epsilon)+\frac{\epsilon}{\sqrt{s}}]^2 \end{aligned} \tag{2.3.2}$$

and

$$\begin{aligned} e^*(M,\bar{X}) &= \sup_{\mathcal{F}} \ e(M,\bar{X},F) \\ &= \sup_{\mathcal{G}} \ e(M,\bar{X},G) \\ &= \sup_{\mathcal{H}} \ e(M,\bar{X},G) \\ &= \sup_{\mathcal{H}} \ 4h^2(0)\sigma_h^2[(1-\epsilon)+\epsilon \textstyle\int sdG(s)][(1-\epsilon)+\epsilon \textstyle\int \frac{1}{\sqrt{s}}dG(s)]^2 \\ &= \sup_{0\le\lambda\le 1} 4h^2(0)\sigma_h^2[(1-\epsilon)+\epsilon(\lambda s_2+(1-\lambda)s_1)] \\ &\quad \cdot[(1-\epsilon)+\epsilon(\frac{\lambda}{\sqrt{s_2}}+\frac{(1-\lambda)}{\sqrt{s_1}})]^2. \end{aligned} \tag{2.3.3}$$

In order to find $e_*(M,\bar{X})$, we define $L(s) = [(1-\epsilon)+\epsilon s]\cdot[(1-\epsilon)+\frac{\epsilon}{\sqrt{s}}]^2$. Then we get

$$\begin{aligned} \frac{dL(s)}{ds} &= [(1-\epsilon)+\epsilon s]\cdot 2[(1-\epsilon)+\frac{\epsilon}{\sqrt{s}}]\cdot(-\frac{\epsilon}{2}\cdot s^{-\frac{3}{2}}) + [(1-\epsilon)+\frac{\epsilon}{\sqrt{s}}]^2\cdot\epsilon \\ &= \epsilon[(1-\epsilon)+\frac{\epsilon}{\sqrt{s}}]\cdot\{(1-\epsilon)+\frac{\epsilon}{\sqrt{s}} - s^{-\frac{3}{2}}[(1-\epsilon)+\epsilon s]\} \\ &= \epsilon(1-\epsilon)[(1-\epsilon)+\frac{\epsilon}{\sqrt{s}}]\cdot(1-s^{-\frac{3}{2}}) > 0, \end{aligned}$$

since $1\le s_1 \le s \le s_2$. Therefore, $L(s)$ obtains its minimum value at $s = s_1$, and we have

$$e_*(M,\bar{X}) = 4h^2(0)\sigma_h^2[(1-\epsilon)+\epsilon s_1][(1-\epsilon)+\frac{\epsilon}{\sqrt{s_1}}]^2,$$

which is realized by the distribution function $G = \Delta_{s_1}$.

In order to find $e^*(M, \bar{X})$, let

$$\begin{aligned} U(\lambda) &= [(1-\epsilon)+\epsilon(\lambda s_2+(1-\lambda)s_1)]\cdot[(1-\epsilon)+\epsilon(\frac{\lambda}{\sqrt{s_2}}+\frac{(1-\lambda)}{\sqrt{s_1}})]^2 \\ &= [(1-\epsilon)+\epsilon s_1+\epsilon(s_2-s_1)\lambda]\cdot[(1-\epsilon)+\frac{\epsilon}{\sqrt{s_1}}+\epsilon(\frac{1}{\sqrt{s_2}}-\frac{1}{\sqrt{s_1}})\lambda]^2, \end{aligned}$$

and let λ be any real number. We have

$$\begin{aligned} \frac{dU(\lambda)}{d\lambda} &= [(1-\epsilon)+\epsilon s_1+\epsilon(s_2-s_1)\lambda] \\ &\quad \cdot 2[(1-\epsilon)+\frac{\epsilon}{\sqrt{s_1}}+\epsilon(\frac{1}{\sqrt{s_2}}-\frac{1}{\sqrt{s_1}})\lambda]\cdot\epsilon(\frac{1}{\sqrt{s_2}}-\frac{1}{\sqrt{s_1}}) \\ &\quad +[(1-\epsilon)+\frac{\epsilon}{\sqrt{s_1}}+\epsilon(\frac{1}{\sqrt{s_2}}-\frac{1}{\sqrt{s_1}})\lambda]^2\cdot\epsilon(s_2-s_1) \\ &= \epsilon[(1-\epsilon)+\frac{\epsilon}{\sqrt{s_1}}+\epsilon(\frac{1}{\sqrt{s_2}}-\frac{1}{\sqrt{s_1}})\lambda] \\ &\quad \cdot\{2(\frac{1}{\sqrt{s_2}}-\frac{1}{\sqrt{s_1}})[(1-\epsilon)+\epsilon s_1+\epsilon(s_2-s_1)\lambda] \\ &\quad +(s_2-s_1)[(1-\epsilon)+\frac{\epsilon}{\sqrt{s_1}}+\epsilon(\frac{1}{\sqrt{s_2}}-\frac{1}{\sqrt{s_1}})\lambda]\} \\ &= \epsilon[(1-\epsilon)+\frac{\epsilon}{\sqrt{s_1}}+\epsilon(\frac{1}{\sqrt{s_2}}-\frac{1}{\sqrt{s_1}})\lambda] \\ &\quad \cdot[2(\frac{1}{\sqrt{s_2}}-\frac{1}{\sqrt{s_1}})(1-\epsilon+\epsilon s_1) \\ &\quad +(s_2-s_1)(1-\epsilon+\frac{\epsilon}{\sqrt{s_1}})+3\epsilon(s_2-s_1)(\frac{1}{\sqrt{s_2}}-\frac{1}{\sqrt{s_1}})\lambda]. \end{aligned}$$

Setting $\frac{dU(\lambda)}{d\lambda}=0$, we get

$$\lambda_1 = \frac{(1-\epsilon)+\frac{\epsilon}{\sqrt{s_1}}}{\epsilon(\frac{1}{\sqrt{s_1}}-\frac{1}{\sqrt{s_2}})}$$

and

$$\lambda_2 = \frac{2(\frac{1}{\sqrt{s_2}}-\frac{1}{\sqrt{s_1}})(1-\epsilon+\epsilon s_1)+(s_2-s_1)(1-\epsilon+\frac{\epsilon}{\sqrt{s_1}})}{3\epsilon(s_2-s_1)(\frac{1}{\sqrt{s_1}}-\frac{1}{\sqrt{s_2}})}.$$

We are going to show that (i) $\lambda_1 > 1$ and (ii) $\lambda_2 > 0$ by the following calculation:

(i)

$$\begin{aligned}
\lambda_1 &= \frac{(1-\epsilon)+\frac{\epsilon}{\sqrt{s_1}}}{\epsilon(\frac{1}{\sqrt{s_1}}-\frac{1}{\sqrt{s_2}})} \\
&= \frac{\epsilon(\frac{1}{\sqrt{s_1}}-\frac{1}{\sqrt{s_2}})-\epsilon(\frac{1}{\sqrt{s_1}}-\frac{1}{\sqrt{s_2}})+(1-\epsilon)+\frac{\epsilon}{\sqrt{s_1}}}{\epsilon(\frac{1}{\sqrt{s_1}}-\frac{1}{\sqrt{s_2}})} \\
&= 1+\frac{1-\epsilon+\frac{\epsilon}{\sqrt{s_1}}-\epsilon(\frac{1}{\sqrt{s_1}}-\frac{1}{\sqrt{s_2}})}{\epsilon(\frac{1}{\sqrt{s_1}}-\frac{1}{\sqrt{s_2}})} \\
&= 1+\frac{1-\epsilon+\frac{\epsilon}{\sqrt{s_2}}}{\epsilon(\frac{1}{\sqrt{s_1}}-\frac{1}{\sqrt{s_2}})} \\
&> 1,
\end{aligned}$$

since $(1-\epsilon+\dfrac{\epsilon}{\sqrt{s_2}})/\epsilon(\dfrac{1}{\sqrt{s_1}}-\dfrac{1}{\sqrt{s_2}})>0$.

(ii) Note that the denominator of λ_2 is positive and the numerator of λ_2 is linear in ϵ. Let

$$N(\epsilon)=2(\frac{1}{\sqrt{s_2}}-\frac{1}{\sqrt{s_1}})(1-\epsilon+\epsilon s_1)+(s_2-s_1)(1-\epsilon+\frac{\epsilon}{\sqrt{s_1}}).$$

Then we have

$$\begin{aligned}
N(0) &= 2(\frac{1}{\sqrt{s_2}}-\frac{1}{\sqrt{s_1}})+s_2-s_1 \\
&= (\sqrt{s_2}+\sqrt{s_1})(\sqrt{s_2}-\sqrt{s_1})-\frac{2}{\sqrt{s_1}\sqrt{s_2}}(\sqrt{s_2}-\sqrt{s_1}) \\
&= (\sqrt{s_2}-\sqrt{s_1})(\sqrt{s_2}+\sqrt{s_1}-\frac{2}{\sqrt{s_1}\sqrt{s_2}})>0
\end{aligned}$$

and

$$\begin{aligned} N(1) &= 2(\frac{1}{\sqrt{s_2}} - \frac{1}{\sqrt{s_1}}) \cdot s_1 + (s_2 - s_1) \cdot \frac{1}{\sqrt{s_1}} \\ &= \frac{2s_1(\sqrt{s_1} - \sqrt{s_2})}{\sqrt{s_1}\sqrt{s_2}} + \frac{(\sqrt{s_2} + \sqrt{s_1})(\sqrt{s_2} - \sqrt{s_1})}{\sqrt{s_1}} \\ &= \frac{\sqrt{s_2} - \sqrt{s_1}}{\sqrt{s_1}\sqrt{s_2}}[\sqrt{s_2}(\sqrt{s_2} + \sqrt{s_1}) - 2s_1] > 0. \end{aligned}$$

We conclude that $N(\epsilon) > 0$ for $\epsilon \in [0,1]$. Hence we have $\lambda_2 > 0$.

Combining the results (i) and (ii) along with the fact that $U(\lambda)$ is a polynomial in λ of degree three with positive leading coefficient, $U(\lambda)$ will achieve its maximum value on interval $[0,1]$ at $\lambda^* = \lambda_2$ if $0 < \lambda_2 < 1$ and at $\lambda^* = 1$ if $\lambda_2 \geq 1$. Hence we have

$$e^*(M, \bar{X}) = \begin{cases} 4h^2(0)\sigma_h^2\{(1-\epsilon) + \epsilon[\lambda_2 s_2 + (1-\lambda_2)s_1]\} & \text{if } 0 < \lambda_2 < 1 \\ \quad \cdot[(1-\epsilon) + \epsilon(\frac{\lambda_2}{\sqrt{s_2}} + \frac{1-\lambda_2}{\sqrt{s_1}})]^2 & \\ 4h^2(0)\sigma_h^2[(1-\epsilon) + \epsilon s_2][(1-\epsilon) + \frac{\epsilon}{\sqrt{s_2}}]^2 & \text{if } \lambda_2 \geq 1, \end{cases}$$

where

$$\lambda_2 = \frac{2(\frac{1}{\sqrt{s_2}} - \frac{1}{\sqrt{s_1}})(1 - \epsilon + \epsilon s_1) + (s_2 - s_1)(1 - \epsilon + \frac{\epsilon}{\sqrt{s_1}})}{3\epsilon(s_2 - s_1)(\frac{1}{\sqrt{s_1}} - \frac{1}{\sqrt{s_2}})}$$

and $e^*(M, \bar{X})$ can be realized by the distribution functions $G = \lambda_2 \Delta_{s_2} + (1 - \lambda_2)\Delta_{s_1}$ or $G = \Delta_{s_2}$ accordingly.

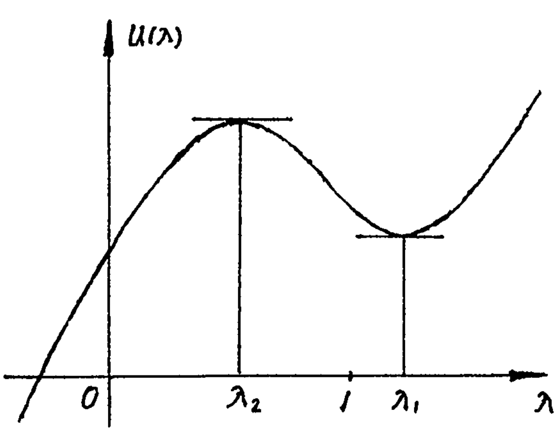

Figure 2.3.2

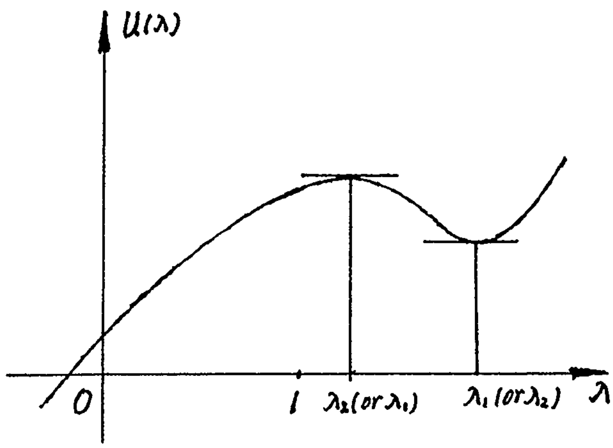

Figure 2.3.3

Note that $\lambda_2 \geq 1$ if and only if

$$\begin{aligned} l(\epsilon) \; &:= (s_2 - s_1)\left(1 - \epsilon + \frac{\epsilon}{\sqrt{s_1}}\right) + 2\left(\frac{1}{\sqrt{s_2}} - \frac{1}{\sqrt{s_1}}\right)(1 - \epsilon + \epsilon s_1) \\ &-3\epsilon(s_2 - s_1)\left(\frac{1}{\sqrt{s_1}} - \frac{1}{\sqrt{s_2}}\right) \geq 0. \end{aligned}$$

Note also $l(\epsilon)$ is linear in ϵ with negative slope. Hence we have

$$\begin{aligned} l(\epsilon) \geq l(\tfrac{1}{2}) \; &= (s_2 - s_1)\left(\frac{1}{2} + \frac{1}{2\sqrt{s_1}}\right) + \left(\frac{1}{\sqrt{s_2}} - \frac{1}{\sqrt{s_1}}\right)(1 + s_1) \\ &-\frac{3}{2}(s_2 - s_1)\left(\frac{1}{\sqrt{s_1}} - \frac{1}{\sqrt{s_2}}\right) \\ &= \frac{1}{2}(s_2 - s_1)\left(1 + \frac{1}{\sqrt{s_1}}\right) \\ &-\left(\frac{1}{\sqrt{s_1}} - \frac{1}{\sqrt{s_2}}\right)\left(1 + \frac{3}{2}s_2 - \frac{1}{2}s_1\right) \\ &= \frac{1}{2}(s_2 - s_1) + \frac{1}{2}(s_2 - s_1)\frac{1}{\sqrt{s_1}} \\ &-\left(\frac{1}{\sqrt{s_1}} - \frac{1}{\sqrt{s_2}}\right)\left[(1 + s_2) + \frac{1}{2}(s_2 - s_1)\right] \\ &= \frac{1}{2}(s_2 - s_1) - \left(\frac{1}{\sqrt{s_1}} - \frac{1}{\sqrt{s_2}}\right)(1 + s_2) \\ &+\frac{1}{2}(s_2 - s_1)\cdot\frac{1}{\sqrt{s_2}} \\ &= \frac{(s_2 - s_1)(\sqrt{s_2} + 1)\sqrt{s_1} - 2(1 + s_2)(\sqrt{s_2} - \sqrt{s_1})}{2\sqrt{s_1}\sqrt{s_2}}. \end{aligned}$$

For $0 \leq \epsilon \leq \frac{1}{2}$, it is clear that $(\sqrt{s_2} + \sqrt{s_1})(\sqrt{s_2} + 1)\sqrt{s_1} - 2(1 + s_2) \geq 0$ is a sufficient condition to guarantee that $\lambda_2 \geq 1$. We have proved the following:

Theorem 2.3.1 *Given* $0 \leq \epsilon \leq \frac{1}{2}$ *and* $1 \leq s_1 < s_2 < \infty$, *we have*

(i) $e_*(M, \bar{X}) = 4h^2(0)\sigma_h^2[(1 - \epsilon) + \epsilon s_1][(1 - \epsilon) + \frac{\epsilon}{\sqrt{s_1}}]^2$ *and* $e_*(M, \bar{X})$ *is achieved by the distribution function* $G = \Delta_{s_1}$.

$$\text{(ii) } e^*(M, \bar{X}) = \begin{cases} 4h^2(0)\sigma_h^2\{(1 - \epsilon) + \epsilon[\lambda_2 s_2 + (1 - \lambda_2)s_1]\} & \text{if } 0 < \lambda_2 < 1 \\ \quad \cdot[(1 - \epsilon) + \epsilon(\frac{\lambda_2}{\sqrt{s_2}} + \frac{1 - \lambda_2}{\sqrt{s_1}})]^2 & \\ 4h^2(0)\sigma_h^2[(1 - \epsilon) + \epsilon s_2][(1 - \epsilon) + \frac{\epsilon}{\sqrt{s_2}}]^2 & \text{if } \lambda_2 \geq 1, \end{cases}$$

where

$$\lambda_2 = \frac{2(\frac{1}{\sqrt{s_2}} - \frac{1}{\sqrt{s_1}})(1 - \epsilon + \epsilon s_1) + (s_2 - s_1)(1 - \epsilon + \frac{\epsilon}{\sqrt{s_1}})}{3\epsilon(s_2 - s_1)(\frac{1}{\sqrt{s_1}} - \frac{1}{\sqrt{s_2}})}$$

and $e^*(M,\bar{X})$ *can be achieved by the distribution functions* $G = \lambda_2\Delta_{s_2} + (1-\lambda_2)\Delta_{s_1}$ *or* $G = \Delta_{s_2}$ *accordingly.*

(iii) A sufficient condition for $\lambda_2 \geq 1$, *hence*

$$e^*(M,\bar{X}) = 4h^2(0)\sigma_h^2[(1-\epsilon)+\epsilon s_2][(1-\epsilon)+\frac{\epsilon}{\sqrt{s_2}}]^2,$$

is $(\sqrt{s_2}+\sqrt{s_1})(\sqrt{s_2}+1)\sqrt{s_1} - 2(1+s_2) \geq 0$.

2.4 Bounds on $e(\bar{X}_\alpha, \bar{X}, F)$ and $e(M, \bar{X}_\alpha, F)$

We start this section by finding the influence function of $\bar{X}_\alpha$. Let $X_1, ..., X_n \sim F(x)$ and $X_i's$ are i.i.d. The α-trimmed mean $\bar{X}_\alpha$ is the estimator that one obtains by removing the $\alpha\%$ largest and $\alpha\%$ smallest observations, and computing the mean of the rest. It is clear that $\alpha = 0$ corresponding to the usual mean $\bar{X}$ and $\alpha = \frac{1}{2}$ corresponding to the median M. $\bar{X}_\alpha$ can be represented as a functional of F by

$$T(F) = \frac{1}{1-2\alpha} \int_\alpha^{1-\alpha} F^{-1}(s) ds,$$

and the influence function of T under F is the following limit:

$$\begin{aligned} IF(x;T,F) &= \lim_{t \to 0} \frac{T[(1-t)F + t\Delta_x] - T(F)}{t} \\ &= \lim_{t \to 0} \frac{\frac{1}{1-2\alpha} \int_\alpha^{1-\alpha} [(1-t)F + t\Delta_x]^{-1}(s) ds - \frac{1}{1-2\alpha} \int_\alpha^{1-\alpha} F^{-1}(s) ds}{t} \\ &= \frac{1}{1-2\alpha} \int_\alpha^{1-\alpha} \{\lim_{t \to 0} \frac{[(1-t)F + t\Delta_x]^{-1}(s) - F^{-1}(s)}{t}\} ds. \end{aligned}$$

Case (i) $x < F^{-1}(\alpha)$ (i.e. $F(x) < \alpha$)

Let $y = [(1-t)F + t\Delta_x]^{-1}(s)$. Then we have $s = (1-t)F(y) + t\Delta_x(y)$.

(1) If $y < x$, then $s = (1-t)F(y)$. But this is impossible, since $s \in [\alpha, 1-\alpha]$ and $F(y) \leq F(x) < \alpha$.

(2) If $y \geq x$, then we have $s = (1-t)F(y) + t$. Take the derivative at both sides with respect to t, we get

$$0 = (1-t)f(y)\frac{dy}{dt} + F(y)(-1) + 1.$$

Evaluate the above at $t = 0$ and note that $t = 0$ if and only if $F(y) = s$. We have

$$0 = f[F^{-1}(s)] \cdot \frac{dy}{dt}|_{t=0} - s + 1.$$

Hence,

$$\frac{dy}{dt}\Big|_{t=0} = \frac{s-1}{f[F^{-1}(s)]} := \lim_{t\to 0} \frac{[(1-t)F + t\Delta_x]^{-1}(s) - F^{-1}(s)}{t}.$$

For $x < F^{-1}(\alpha)$, we conclude

$$\begin{aligned} IF(x;T,F) &= \frac{1}{1-2\alpha}\int_{\alpha}^{1-\alpha} \frac{s-1}{f[F^{-1}(s)]}ds \\ &= \frac{1}{1-2\alpha}\int_{F^{-1}(\alpha)}^{F^{-1}(1-\alpha)} \frac{F(y)-1}{f(y)}\cdot f(y)dy \\ &= yF(y)\Big|_{F^{-1}(\alpha)}^{F^{-1}(1-\alpha)} - \int_{F^{-1}(\alpha)}^{F^{-1}(1-\alpha)} yf(y)dy - y\Big|_{F^{-1}(\alpha)}^{F^{-1}(1-\alpha)} \\ &= \frac{1}{1-2\alpha}\{[(1-\alpha)F^{-1}(1-\alpha) - \alpha F^{-1}(\alpha)] \\ &\quad - \int_{\alpha}^{1-\alpha} F^{-1}(s)ds - [F^{-1}(1-\alpha) - F^{-1}(\alpha)]\} \\ &= \frac{1}{1-2\alpha}[F^{-1}(\alpha) - C], \end{aligned}$$

where

$$C = \int_{\alpha}^{1-\alpha} F^{-1}(s)ds + \alpha F^{-1}(\alpha) + \alpha F^{-1}(1-\alpha).$$

Case (ii) $F^{-1}(\alpha) \leq x \leq F^{-1}(1-\alpha)$ (i.e. $\alpha \leq F(\alpha) \leq 1-\alpha$)

Let $y = [(1-t)F + t\Delta_x]^{-1}(s)$. Then $s = (1-t)F(y) + t\Delta_x(y)$.

(1) For $y < x$, we have $s = (1-t)F(y)$ and $\frac{dy}{dt}\Big|_{t=0} = \frac{s}{f[F^{-1}(s)]}$.

(2) For $y \geq x$, we have $s = (1-t)F(y) + t$ and $\frac{dy}{dt}\Big|_{t=0} = \frac{s-1}{f[F^{-1}(s)]}$.

Hence, for $F^{-1}(\alpha) \le x \le F^{-1}(1-\alpha)$, we have

$$
\begin{aligned}
IF(x;T,F) &= \frac{1}{1-2\alpha}\{\int_{\alpha}^{F(x)} \frac{s}{f\left[F^{-1}(s)\right]}ds + \int_{F(x)}^{1-\alpha} \frac{s-1}{f\left[F^{-1}(s)\right]}ds\} \\
&= \frac{1}{1-2\alpha}\{\int_{F^{-1}(\alpha)}^{x} \frac{F(y)}{f(y)}f(y)dy + \int_{x}^{F^{-1}(1-\alpha)} \frac{F(y)-1}{f(y)}dy\} \\
&= \frac{1}{1-2\alpha}\{yF(y)|_{F^{-1}(\alpha)}^{x} - \textstyle\int_{F^{-1}(\alpha)}^{x} yf(y)dy + yF(y)|_{x}^{F^{-1}(1-\alpha)} \\
&\quad - \int_{x}^{F^{-1}(1-\alpha)} yf(y)dy - [F^{-1}(1-\alpha) - x]\} \\
&= \frac{1}{1-2\alpha}\{xF(x) - \alpha F^{-1}(\alpha) - \int_{\alpha}^{F(x)} F^{-1}(s)ds + (1-\alpha)F^{-1}(1-\alpha) \\
&\quad -xF(x) - \int_{F(x)}^{1-\alpha} F^{-1}(s)ds - F^{-1}(1-\alpha) + x\} \\
&= \frac{1}{1-2\alpha}[x - \int_{\alpha}^{1-\alpha} F^{-1}(s)ds - \alpha F^{-1}(\alpha) - F^{-1}(1-\alpha)] \\
&= \frac{1}{1-2\alpha}[x - C].
\end{aligned}
$$

Case (iii) $x > F^{-1}(1-\alpha)$ (i.e. $F(x) > 1-\alpha$)

Let $y = [(1-t)F + t\Delta_x]^{-1}(s)$. Then $s = (1-t)F(y) + \Delta_x(y)$.

(1) If $y < x$, then we have $s = (1-t)F(y)$ and $\frac{dy}{dt}\,|_{t=0} = \frac{s}{f[F^{-1}(s)]}$.

(2) If $y \ge x$, we have $(1-t)F(y) + t = s$. But this is impossible, since $s \in [\alpha, 1-\alpha]$ and $F(y) \ge F(x) > 1-\alpha$.

For $x > F^{-1}(1-\alpha)$, we get

$$\begin{aligned}
IF(x;T,F) &= \frac{1}{1-2\alpha}\int_{\alpha}^{1-\alpha}\frac{s}{f[F^{-1}(s)]}ds = \frac{1}{1-2\alpha}\int_{F^{-1}(\alpha)}^{F^{-1}(1-\alpha)}F(y)dy \\
&= \frac{1}{1-2\alpha}[yF(y)\mid_{F^{-1}(\alpha)}^{F^{-1}(1-\alpha)} - \int_{F^{-1}(\alpha)}^{F^{-1}(1-\alpha)} yf(y)dy] \\
&= \frac{1}{1-2\alpha}[F^{-1}(1-\alpha) - \int_{\alpha}^{1-\alpha}F^{-1}(s)ds - \alpha F^{-1}(\alpha) - \alpha F^{-1}(1-\alpha)] \\
&= \frac{1}{1-2\alpha}[F^{-1}(1-\alpha) - C].
\end{aligned}$$

Combining the above three cases, we get

$$IF(x;T,F) = \begin{cases} \frac{1}{1-2\alpha}[F^{-1}(\alpha) - C] & x < F^{-1}(\alpha) \\ \frac{1}{1-2\alpha}[x - C] & F^{-1}(\alpha) \le x \le F^{-1}(1-\alpha) \\ \frac{1}{1-2\alpha}[F^{-1}(1-\alpha) - C] & x > F^{-1}(1-\alpha), \end{cases}$$

where $C = \int_{\alpha}^{1-\alpha} F^{-1}(s)ds + \alpha F^{-1}(\alpha) + \alpha F^{-1}(1-\alpha)$.

Note that $C = \int_{\alpha}^{1-\alpha} F^{-1}(s)ds + \alpha F^{-1}(\alpha) + \alpha F^{-1}(1-\alpha) = 0$ if F is symmetric, which is the case we are interested in. Denote the asymptotic variance of $\bar{X}_\alpha$ under F as $V(\bar{X}_\alpha, F)$. By (1.2), we have

$$\begin{aligned}
V(\bar{X}_\alpha, F) &= \int IF(x;T,F)^2 dF \\
&= \frac{2}{(1-2\alpha)^2}\left\{\int_0^{F^{-1}(1-\alpha)} x^2 dF(x) + [F^{-1}(1-\alpha)]^2 \int_{F^{-1}(1-\alpha)}^{\infty} dF(x)\right\} \\
&= \frac{2}{(1-2\alpha)^2}\{\int_0^{\gamma} x^2 dF(x) + \alpha\gamma^2\} \\
&= \frac{2}{(1-2\alpha)^2}\{(1-\epsilon)\int_0^{\gamma} x^2 h(x)dx \\
&\quad + \epsilon \int_0^{\gamma} x^2 [\int h(\frac{x}{\sqrt{s}})\cdot\frac{1}{\sqrt{s}}dG(s)]dx + \alpha\gamma^2\} \\
&= \frac{2}{(1-2\alpha)^2}\{(1-\epsilon)\int_0^{\gamma} x^2 h(x)dx \\
&\quad + \epsilon \int s[\int_0^{\gamma}(\frac{x}{\sqrt{s}})^2 h(\frac{x}{\sqrt{s}})d(\frac{x}{\sqrt{s}})]dG(s) + \alpha\gamma^2\} \\
&= \frac{2}{(1-2\alpha)^2}\{(1-\epsilon)\int_0^{\gamma} x^2 h(x)dx \\
&\quad + \epsilon \int s[\int_0^{\gamma/\sqrt{s}} y^2 h(y)dy]dG(s) + \alpha\gamma^2\} \\
&= \frac{2}{(1-2\alpha)^2}\{(1-\epsilon)\int_0^{\gamma} x^2 h(x)dx + \epsilon\int sk(s)dG(s) + \alpha\gamma^2\},
\end{aligned}$$

where $k(s) = \int_0^{\gamma/\sqrt{s}} y^2 h(y)dy$, and $\gamma = F^{-1}(1-\alpha)$.

Therefore, we get

$$\begin{aligned} e(\bar{X}_\alpha, \bar{X}, F) &= \frac{V(\bar{X}, F)}{V(\bar{X}_\alpha, F)} \\ &= \frac{(1-2\alpha)^2}{2} \cdot \frac{(1-\epsilon)\sigma_h^2 + \epsilon\sigma_h^2 \cdot \int s dG(s)}{(1-\epsilon)\int_0^\gamma x^2 h(x)dx + \alpha\gamma^2 + \epsilon \int sk(s)dG(s)} \end{aligned} \tag{2.4.1}$$

and

$$\begin{aligned} e(M, \bar{X}_\alpha, F) &= \frac{V(\bar{X}_\alpha, F)}{V(M, F)} \\ &= \frac{8h^2(0)}{(1-2\alpha)^2}[(1-\epsilon) + \epsilon \int \frac{1}{\sqrt{s}} dG(s)]^2 \\ &\cdot [(1-\epsilon)\int_0^\gamma x^2 h(x)dx + \alpha\gamma^2 + \epsilon \int sk(s)dG(s)]. \end{aligned} \tag{2.4.2}$$

Note that $\gamma = F^{-1}(1-\alpha)$, i.e., $F(\gamma) = 1-\alpha$. We have $(1-\epsilon)H(\gamma) + \epsilon \int H(\frac{\gamma}{\sqrt{s}})dG(s) = 1-\alpha$. For any given $0 < \alpha < \frac{1}{2}$, and $0 < \epsilon < \frac{1}{2}$, γ is different in F(or G). We need to know the range of γ for F over $\mathcal{F}$(or G over $\mathcal{G}$). For any $G \in \mathcal{G}$, we define

$$\tilde{H}_G(\gamma) = (1-\epsilon)H(\gamma) + \epsilon \int H(\frac{\gamma}{\sqrt{s}})dG(s).$$

Moreover, let

$$\gamma_1 = \min_{G\in\mathcal{G}} \{\gamma : \tilde{H}_G(\gamma) = 1-\alpha\} \text{ and } \gamma_2 = \max_{G\in\mathcal{G}} \{\gamma : \tilde{H}_G(\gamma) = 1-\alpha\}.$$

Then, we have the following:

Lemma 2.4.1 *For given $0 < \alpha < \frac{1}{2}$, $0 < \epsilon < \frac{1}{2}$, and $1 \le s_1 < s_2 < \infty$, we have that γ_i is the solution of $H_i(\gamma) = 1-\alpha$, where $H_i(\gamma) = (1-\epsilon)H(\gamma) + \epsilon H(\frac{\gamma}{\sqrt{s_i}})$, $\quad i = 1, 2$.*

Proof: Note that

$$\frac{d\tilde{H}_G(\gamma)}{d\gamma} = (1-\epsilon)h(\gamma) + \epsilon \int h(\frac{\gamma}{\sqrt{s}}) \cdot (\frac{1}{\sqrt{s}})dG(s) > 0.$$

Hence $\tilde{H}(\gamma)$ is monotone increasing in γ. On the other hand, we have

$$H(\frac{\gamma}{\sqrt{s_2}}) \le \int H(\frac{\gamma}{\sqrt{s}})dG(s) \le H(\frac{\gamma}{\sqrt{s_1}})$$

for any fixed γ. Hence, we have $H_2(\gamma) \le \tilde{H}_G(\gamma) \le H_1(\gamma)$. Consequently, we have $H_1^{-1}(1-\alpha) \le \tilde{H}_G^{-1}(1-\alpha) \le H_2^{-1}(1-\alpha)$, i.e., $\gamma_1 \le \gamma \le \gamma_2$, where γ is the solution of $\tilde{H}_G(\gamma) = 1-\alpha$, and $G \in \mathcal{G}$. $\square$

For any $\gamma \in [\gamma_1, \gamma_2]$, let $\mathcal{G}_\gamma = \{G : G \in \mathcal{G} \text{ and } \tilde{H}_G(\gamma) = 1 - \alpha\}$. Then we have

$$\begin{aligned} e^*(\bar{X}_\alpha, \bar{X}) &= \max_{\gamma_1 \le \gamma \le \gamma_2} \sup_{G \in \mathcal{G}_r} e(\bar{X}_\alpha, \bar{X}, G), \\ e_*(\bar{X}_\alpha, \bar{X}) &= \min_{\gamma_1 \le \gamma \le \gamma_2} \inf_{G \in \mathcal{G}_r} e(\bar{X}_\alpha, \bar{X}, G); \end{aligned}$$

and

$$\begin{aligned} e^*(M, \bar{X}_\alpha) &= \max_{\gamma_1 \le \gamma \le \gamma_2} \sup_{G \in \mathcal{G}_r} e(M, \bar{X}_\alpha, G), \\ e_*(M, \bar{X}_\alpha) &= \min_{\gamma_1 \le \gamma \le \gamma_2} \inf_{G \in \mathcal{G}_r} e(M, \bar{X}_\alpha, G). \end{aligned}$$

Unfortunately, we are unable to find the exact bounds on $e(\bar{X}_\alpha, \bar{X}, G)$ and $e(M, \bar{X}_\alpha, G)$. We can only indicate that the bounds are located within certain ranges. Let

$$\mathcal{G}_n = \{G : G \in \mathcal{G} \quad Card\ \sigma(G) = n\}.$$

Then we have the following:

Theorem 2.4.2 *Let $F \in \mathcal{F}$ as we defined in (2.1.2), where $H(x)$ is known and absolutely continuous. For given $0 < \alpha < \frac{1}{2}$, $0 < \epsilon < \frac{1}{2}$, and $1 \le s_1 < s_2 < \infty$, we have*

(i) $\sup_{G \in \mathcal{G}_n} e(\bar{X}_\alpha, \bar{X}, G) \le e^*(\bar{X}_\alpha, \bar{X}) \le \max_{\gamma_1 \le \gamma \le \gamma_2} \sup_{G \in \mathcal{T}} e(\bar{X}_\alpha, \bar{X}, G)$, *and*

$\min_{\gamma_1 \le \gamma \le \gamma_2} \inf_{G \in \mathcal{T}} e(\bar{X}_\alpha, \bar{X}, G) \le e_*(\bar{X}_\alpha, \bar{X}) \le \inf_{G \in \mathcal{G}_n} e(\bar{X}_\alpha, \bar{X}, G)$;

(ii) $\sup_{G \in \mathcal{G}_n} e(M, \bar{X}_\alpha, G) \le e^*(M, \bar{X}_\alpha, G) \le \max_{\gamma_1 \le \gamma \le \gamma_2} \sup_{G \in \mathcal{T}} e(M, \bar{X}_\alpha, G)$, *and*

$\min_{\gamma_1 \le \gamma \le \gamma_2} \inf_{G \in \mathcal{T}} e(M, \bar{X}_\alpha, G) \le e_*(M, \bar{X}_\alpha) \le \inf_{G \in \mathcal{G}_n} e(M, \bar{X}_\alpha, G)$,

where $\mathcal{T}$ is defined by (2.2.2) and $n \ge 1$.

Proof: We have $\mathcal{G}_n \subseteq \mathcal{G}$, hence $\sup_{G \in \mathcal{G}_n} e(\bar{X}_\alpha, \bar{X}, G) \le \sup_{G \in \mathcal{G}} e(\bar{X}_\alpha, \bar{X}, G) = e^*(\bar{X}_\alpha, \bar{X})$. Similarly, $\mathcal{G}_r \subseteq \mathcal{G}$ implies that

$$e^*(\bar{X}_\alpha, \bar{X}) = \max_{\gamma_1 \le \gamma \le \gamma_2} \sup_{G \in \mathcal{G}_\gamma} e(\bar{X}_\alpha, \bar{X}, G) \le \max_{\gamma_1 \le \gamma \le \gamma_2} \sup_{G \in \mathcal{G}} e(\bar{X}_\alpha, \bar{X}, G).$$

On the other hand, for any fixed γ, we have that $(s, sk(s))$ is a continuous curve where $k(s) = \int_0^{\gamma/\sqrt{s}} y^2 h(y) dy$ depends on γ. The application of Theorem 2.2.1 gives us

$$\sup_{G\in\mathcal{G}} e(\bar{X}_\alpha, \bar{X}, G) = \sup_{G\in\mathcal{T}} e(\bar{X}_\alpha, \bar{X}, G).$$

Hence, we have

$$\sup_{G\in\mathcal{G}_n} e(\bar{X}_\alpha, \bar{X}, G) \le e^*(\bar{X}_\alpha, \bar{X}) \le \max_{\gamma_1\le\gamma\le\gamma_2} \sup_{G\in\mathcal{T}} e(\bar{X}_\alpha, \bar{X}, G).$$

The proof for other cases is similar. □

Let us define

$$C_1 = \left\{(u,v) : u = s, v = sk(s), k(s) = \int_0^{\gamma/\sqrt{s}} y^2 h(y)dy,\ s_1 \le s \le s_2\right\}$$

and

$$C_2 = \left\{(u,v) : u = \frac{1}{\sqrt{s}}, v = sk(s), k(s) = \int_0^{\gamma/\sqrt{s}} y^2 h(y)dy,\ s_1 \le s \le s_2\right\}.$$

Usually, C_1 and C_2 are the continuous curves. If we have some more properties of C_1 and C_2, for example the convexity, we will have some results about the bounds of $e(\bar{X}_\alpha, \bar{X}, G)$ and $e(M, \bar{X}_\alpha, G)$ which are simpler than the results in Theorem 2.4.2 in the sense of numerical calculation. Denote

$$\mathcal{D} = \{G :\ G \in \mathcal{G}, G = \lambda\Delta_{s_1} + (1-\lambda)\Delta_{s_2}, s_1 \le s \le s_2\} \quad \text{and}$$
$$\mathcal{E} = \{G : G \in \mathcal{G}, G = \Delta_s, s_1 \le s \le s_2\}.$$

Then we have the following:

Theorem 2.4.3 *For any fixed $\gamma \in [\gamma_1, \gamma_2]$, let us make the same assumptions as we did in Theorem 2.4.2.*

(i) If C_1, C_2 are the convex or concave curves, then we have

(i,a) $\sup_{G\in\mathcal{T}} e(\bar{X}_\alpha, \bar{X}, G) = \sup_{G\in\mathcal{D}\cup\mathcal{E}} e(\bar{X}_\alpha, \bar{X}, G),$

(i,b) $\inf_{G\in\mathcal{T}} e(\bar{X}_\alpha, \bar{X}, G) = \inf_{G\in\mathcal{D}\cup\mathcal{E}} e(\bar{X}_\alpha, \bar{X}, G),$

(i,c) $\sup_{G\in\mathcal{T}} e(M, \bar{X}_\alpha, G) = \sup_{G\in\mathcal{D}\cup\mathcal{E}} e(M, \bar{X}_\alpha, G),$

(i,d) $\inf_{G\in\mathcal{T}} e(M, \bar{X}_\alpha, G) = \inf_{G\in\mathcal{D}\cup\mathcal{E}} e(M, \bar{X}_\alpha, G).$

(ii) If C_1, C_2 are the convex (concave) and monotone increasing curves, then we have

(ii,a) $\sup_{G\in\mathcal{T}} e(\bar{X}_\alpha, \bar{X}, G) = \sup_{G\in\mathcal{E}} e(\bar{X}_\alpha, \bar{X}, G)$

$\left(\sup_{G\in\mathcal{T}} e(\bar{X}_\alpha, \bar{X}, G) = \sup_{G\in\mathcal{D}} e(\bar{X}_\alpha, \bar{X}, G)\right)$,

(ii,b) $\inf_{G\in\mathcal{T}} e(\bar{X}_\alpha, \bar{X}, G) = \inf_{G\in\mathcal{D}} e(\bar{X}_\alpha, \bar{X}, G)$

$\left(\inf_{G\in\mathcal{T}} e(\bar{X}_\alpha, \bar{X}, G) = \inf_{G\in\mathcal{E}} e(\bar{X}_\alpha, \bar{X}, G)\right)$,

(ii,c) $\sup_{G\in\mathcal{T}} e(M, \bar{X}_\alpha, G) = e(M, \bar{X}_\alpha, \Delta_{s_1})$,

(ii,d) $\inf_{G\in\mathcal{T}} e(M, \bar{X}_\alpha, G) = e(M, \bar{X}_\alpha, \Delta_{s_2})$.

(iii) If C_1, C_2 are the convex (concave) and monotone decreasing curves, then we have

(iii,a) $\sup_{G\in\mathcal{T}} e(\bar{X}_\alpha, \bar{X}, G) = e(\bar{X}_\alpha, \bar{X}, \Delta_{s_2})$,

(iii,b) $\inf_{G\in\mathcal{T}} e(\bar{X}_\alpha, \bar{X}, G) = e(\bar{X}_\alpha, \bar{X}, \Delta_{s_1})$,

(iii,c) $\sup_{G\in\mathcal{T}} e(M, \bar{X}_\alpha, G) = \sup_{G\in\mathcal{D}} e(M, \bar{X}_\alpha, G)$

$\left(\sup_{G\in\mathcal{T}} e(M, \bar{X}_\alpha, G) = \sup_{G\in\mathcal{E}} e(M, \bar{X}_\alpha, G)\right)$,

(iii,d) $\inf_{G\in\mathcal{T}} e(M, \bar{X}_\alpha, G) = \inf_{G\in\mathcal{E}} e(M, \bar{X}_\alpha, G)$

$\left(\inf_{G\in\mathcal{T}} e(M, \bar{X}_\alpha, G) = \inf_{G\in\mathcal{D}} e(M, \bar{X}_\alpha, G)\right)$.

Proof: The results are followed easily by looking at the forms of $e(\bar{X}_\alpha, \bar{X}, G)$ and $e(M, \bar{X}_\alpha, G)$, and the regions of $\hat{S}$ in each case. □

We are now going to find the bounds on $e(\bar{X}_\alpha, \bar{X}, G)$ when $H(x)$ is specified. For $H(x) = \Phi(x)$, the standard normal distribution, we have

Theorem 2.4.4 *Let $H(x) = \Phi(x)$ in (2.1.2), $0 < \epsilon < \frac{1}{2}, 0 < \alpha < \frac{1}{2}$, and $1 \le s_1 < s_2 < \infty$. If $\gamma_2 < \sqrt{s_1}$, then we have*

$$
\begin{aligned}
&(i) \quad \sup_{G\in\mathcal{G}_n} e(\bar{X}_\alpha, \bar{X}, G) \le e^*(\bar{X}_\alpha, \bar{X}) \\
&\qquad \le \frac{(1-2\alpha)^2}{2}\cdot\frac{(1-\epsilon)+\epsilon s_2}{(1-\epsilon)\Psi(1,\gamma_1)+\epsilon\Psi(s_2,\gamma_1)+\alpha\gamma_1^2}, \\
&(ii) \quad \frac{(1-2\alpha)^2}{2}\cdot\frac{(1-\epsilon)+\epsilon s_1}{(1-\epsilon)\Psi(1,\gamma_2)+\epsilon\Psi(s_1,\gamma_2)+\alpha\gamma_2^2} \\
&\qquad \le e_*(\bar{X}_\alpha, \bar{X}) \le \inf_{G\in\mathcal{G}_n} e(\bar{X}_\alpha, \bar{X}, G),
\end{aligned}
$$

where $\Psi(w,\gamma) = w\Phi\left(\frac{\gamma}{\sqrt{w}}\right) - \frac{1}{2}w - \frac{\gamma\sqrt{w}}{\sqrt{2\pi}}e^{-\frac{\gamma^2}{2w}}$, *and* γ_i *is the solution of* $(1-\epsilon)\Phi(\gamma)+\epsilon\Phi\left(\frac{\gamma}{\sqrt{s_i}}\right) = 1-\alpha, \quad i=1,2.$

Moreover, we have that a sufficient condition for $\gamma_2 < \sqrt{s_1}$ *is* $\Phi^{-1}\left(\frac{1-\alpha-\frac{1}{2}\epsilon}{1-\epsilon}\right) \le \sqrt{s_1}$.

Proof: When $H(x) = \Phi(x)$, we have

$$C_1 = \left\{(u,v):\ u=s,\ v=sk(s), k(s)=\int_0^{\gamma/\sqrt{s}} y^2\varphi(y)dy, s_1\le s\le s_2\right\},$$

where $\varphi(y)$ is the probability density function of the standard normal distribution, i.e., $\varphi(y) = \frac{1}{\sqrt{2\pi}}e^{-\frac{y^2}{2}}$.

Let $v(s) = sk(s) = s\int_0^{\gamma/\sqrt{s}} y^2\varphi(y)dy = \frac{1}{\sqrt{2\pi}}s\int_0^{\gamma/\sqrt{s}} y^2 e^{-\frac{y^2}{2}}dy$. Then we get

$$
\begin{aligned}
\frac{dv(s)}{ds} &= \frac{1}{\sqrt{2\pi}}\left[\int_0^{\gamma/\sqrt{s}} y^2 e^{-\frac{y^2}{2}}dy + s\frac{\gamma^2}{s}e^{-\frac{\gamma^2}{2s}}(-\frac{1}{2})\gamma\cdot s^{-\frac{3}{2}}\right] \\
&= \frac{1}{\sqrt{2\pi}}\left[\int_0^{\gamma/\sqrt{s}} y^2 e^{-\frac{y^2}{2}}dy - \frac{1}{2}(\frac{\gamma}{\sqrt{s}})^3 e^{-\frac{\gamma^2}{2s}}\right],
\end{aligned}
$$

and

$$
\begin{aligned}
\frac{d^2v(s)}{ds^2} &= \frac{1}{\sqrt{2\pi}}\{\frac{\gamma^2}{s}e^{-\frac{\gamma^2}{2s}}(-\frac{1}{2}\gamma s^{-\frac{3}{2}}) \\
&\qquad -\frac{1}{2}\gamma^3[s^{-\frac{3}{2}}e^{-\frac{\gamma^2}{2s}}\frac{\gamma^2}{2s^2} + e^{-\frac{\gamma^2}{2s}}(-\frac{3}{2})s^{-\frac{5}{2}}]\} \\
&= \frac{1}{\sqrt{2\pi}}\left(\frac{1}{4}\gamma^3 s^{-\frac{5}{2}}e^{-\frac{\gamma^2}{2s}} - \frac{1}{4}\gamma^5 s^{-\frac{7}{2}}e^{-\frac{\gamma^2}{2s}}\right) \\
&= \frac{\gamma^3}{4\sqrt{2\pi}}s^{-\frac{5}{2}}e^{-\frac{\gamma^2}{2s}}(1-\frac{\gamma^2}{s}) > 0,
\end{aligned}
$$

since $\gamma \le \gamma_2 < \sqrt{s_1}$.

The second derivative of $v(s)$ with respect to s is positive, which implies that $C_1$is a convex curve. On the other hand, if we let

$$L(z) = \frac{1}{\sqrt{2\pi}}\left(\int_0^z y^2 e^{-\frac{y^2}{2}}dy - \frac{1}{2}z^3 e^{-\frac{z^2}{2}}\right),$$

then we have

$$\begin{aligned}\frac{dL(z)}{dz} &= \frac{1}{\sqrt{2\pi}}\left\{y^2 e^{-\frac{z^2}{2}} - \frac{1}{2}[z^3 e^{-\frac{z^2}{2}}(1-z) + e^{-\frac{z^2}{2}}\cdot 3z^2]\right\} \\ &= \frac{1}{2\sqrt{2\pi}}z^2 e^{-\frac{z^2}{2}}(z^2-1) < 0,\end{aligned}$$

if $0 < z < 1$. Note that $L(0) = 0$; hence we have $L(z) < 0, z \in (0,1)$. Note also, $L(\frac{\gamma}{\sqrt{s}}) = \frac{dv(s)}{ds}$, and $0 < \frac{\gamma}{\sqrt{s}} < 1$ (since $\gamma_2 < \sqrt{s_1}$ and $\gamma \geq 0$. $\gamma_1 = 0$ if and only if $\alpha = \frac{1}{2}$). We conclude that $\frac{dv(s)}{ds} < 0$. This implies that C_1 is monotone decreasing.

According to Theorem 2.4.3 (iii,a) and (iii,b) we have

$$\begin{aligned}\sup_{G\in\mathcal{T}} e(\bar{X}_\alpha, \bar{X}, G) &= e(\bar{X}_\alpha, \bar{X}, \Delta_{s_2}) \\ &= \frac{(1-2\alpha)^2}{2}\cdot\frac{(1-\epsilon)+\epsilon s_2}{(1-\epsilon)\int_0^\gamma x^2\varphi(x)dx + \epsilon s_2 k(s_2) + \alpha\gamma^2} \\ &= \frac{(1-2\alpha)^2}{2}\cdot[(1-\epsilon)+\epsilon s_2]/\{(1-\epsilon)[\Phi(\gamma) - \tfrac{1}{2} - \tfrac{1}{\sqrt{2\pi}}\gamma e^{-\frac{\gamma^2}{2}} \\ &\qquad +\epsilon[s_2\Phi(\tfrac{\gamma}{\sqrt{s_2}}) - \tfrac{1}{2}s_2 - \tfrac{1}{\sqrt{2\pi}}\gamma\sqrt{s_2}e^{-\frac{\gamma^2}{2s_2}}] + \alpha\gamma^2\} \\ &= \frac{(1-2\alpha)^2}{2}\cdot\frac{(1-\epsilon)+\epsilon s_2}{(1-\epsilon)\Psi(1,\gamma)+\epsilon\Psi(s_2,\gamma)+\alpha\gamma^2},\end{aligned} \tag{2.4.3}$$

and similarly

$$\begin{aligned}\inf_{G\in\mathcal{T}} e(\bar{X}_\alpha, \bar{X}, G) &= e(\bar{X}_\alpha, \bar{X}, \Delta_{s_1}) \\ &= \frac{(1-2\alpha)^2}{2}\cdot\frac{(1-\epsilon)+\epsilon s_1}{(1-\epsilon)\Psi(1,\gamma)+\epsilon\Psi(s_1,\gamma)+\alpha\gamma^2},\end{aligned} \tag{2.4.4}$$

where $\Psi(w,\gamma) = w\Phi(\frac{\gamma}{\sqrt{w}}) - \frac{1}{2}w - \frac{\gamma\sqrt{w}}{\sqrt{2\pi}}e^{-\frac{\gamma^2}{2w}}$. For any given $\gamma \in [\gamma_1, \gamma_2]$, $\sup_{G\in\mathcal{T}} e(\bar{X}_\alpha, \bar{X}, G)$ and $\inf_{G\in\mathcal{T}} e(\bar{X}_\alpha, \bar{X}, G)$ are given by (2.4.3) and (2.4.4). By Theorem 2.4.2, we know that $e^*(\bar{X}_\alpha, \bar{X}) \leq \max_{\gamma_1\leq\gamma<\gamma_2}\sup_{G\in\mathcal{T}} e(\bar{X}_\alpha, \bar{X}, G)$, and $\min_{\gamma_1\leq\gamma<\gamma_2}\inf_{G\in\mathcal{T}} e(\bar{X}_\alpha, \bar{X}, G) \leq e_*(\bar{X}_\alpha, \bar{X})$. Hence, we choose $\max_{\gamma_1\leq\gamma<\gamma_2}\sup_{G\in\mathcal{T}} e(\bar{X}_\alpha, \bar{X}, G)$ as an upper bound for $e^*(\bar{X}_\alpha, \bar{X})$ and $\min_{\gamma_1\leq\gamma<\gamma_2}\inf_{G\in\mathcal{T}} e(\bar{X}_\alpha, \bar{X}, G)$ as a lower bound for $e_*(\bar{X}_\alpha, \bar{X})$. Let

$$K_1(\gamma) = (1-\epsilon)\Psi(1,\gamma) + \epsilon\Psi(s_1,\gamma) + \alpha\gamma^2$$

and

$$K_2(\gamma) = (1-\epsilon)\Psi(1,\gamma) + \epsilon\Psi(s_2,\gamma) + \alpha\gamma^2.$$

Then we have

$$\begin{aligned}\frac{dK_1(\gamma)}{d\gamma} &= 2\alpha\gamma + (1-\epsilon)\{\frac{1}{\sqrt{2\pi}}e^{-\frac{\gamma^2}{2}} - \frac{1}{\sqrt{2\pi}}[e^{-\frac{\gamma^2}{2}} + \gamma e^{-\frac{\gamma^2}{2}}(-\gamma)]\} \\ &+\epsilon\{\frac{s_2}{\sqrt{2\pi}}e^{-\frac{\gamma^2}{2s_1}} \cdot \frac{1}{\sqrt{s_1}} - \frac{\sqrt{s_1}}{\sqrt{2\pi}}[e^{-\frac{\gamma^2}{2s_1}} + \gamma e^{-\frac{\gamma^2}{2s_1}}(-\frac{\gamma}{s_1})]\} \\ &= 2\alpha\gamma + \frac{(1-\epsilon)}{\sqrt{2\pi}}\gamma^2 e^{-\frac{\gamma^2}{2}} + \frac{\epsilon}{\sqrt{2\pi s_1}}\gamma^2 e^{-\frac{\gamma^2}{2}} > 0,\end{aligned}$$

and similarly

$$\frac{dK_2(\gamma)}{d\gamma} = 2\alpha\gamma + \frac{(1-\epsilon)}{\sqrt{2\pi}}\gamma^2 e^{-\frac{\gamma^2}{2}} + \frac{\epsilon}{\sqrt{2\pi s_2}}\gamma^2 e^{-\frac{\gamma^2}{2}} > 0.$$

Hence, both $K_1(\gamma)$ and $K_2(\gamma)$ are increasing in γ. It is clear that

$$\max_{\gamma_1\le\gamma<\gamma_2} e(\bar{X}_\alpha, \bar{X}, \Delta_{s_2}) = \frac{(1-2\alpha)^2}{2} \cdot \frac{(1-\epsilon)+\epsilon s_2}{(1-\epsilon)\Psi(1,\gamma_1) + \epsilon\Psi(s_2,\gamma_1) + \alpha\gamma_1^2}$$

and

$$\min_{\gamma_1\le\gamma\le\gamma_2} e(\bar{X}_\alpha, \bar{X}, \Delta_{s_1}) = \frac{(1-2\alpha)^2}{2} \cdot \frac{(1-\epsilon)+\epsilon s_1}{(1-\epsilon)\Psi(1,\gamma_2) + \epsilon\Psi(s_1,\gamma_2) + \alpha\gamma_2^2}.$$

We have proved part (i) and (ii) of Theorem 2.4.4.

Finally, note that $(1-\epsilon)\Phi(\gamma) + \epsilon\int\Phi(\frac{\gamma}{\sqrt{s}})dG(s) = 1-\alpha$, and

$$\begin{aligned}\gamma &= \Phi^{-1}\left(\frac{1-\alpha-\epsilon\int\Phi\left(\frac{\gamma}{\sqrt{s}}\right)dG(s)}{1-\epsilon}\right) \\ &< \Phi^{-1}\left(\frac{1-\alpha-\epsilon\int\Phi(0)\,dG(s)}{1-\epsilon}\right) \\ &= \Phi^{-1}\left(\frac{1-\alpha-\frac{1}{2}\epsilon}{1-\epsilon}\right).\end{aligned}$$

Therefore $\Phi^{-1}\left(\frac{1-\alpha-\frac{1}{2}\epsilon}{1-\epsilon}\right) \le \sqrt{s_1}$ implies that $\gamma_2 < \sqrt{s_1}$. □

We can also discuss the problem of finding the bounds of $e(\bar{X}_\alpha, \bar{X}, G)$ for the other choices of $H(x)$. For example, we can choose $H(x) = \int_{-\infty}^{x}\frac{1}{2}e^{-|t|}dt$. In this case, we will get a result which is similar to Theorem 2.4.4, the case when $H(x) = \Phi(x)$. We state the result as follows.

Theorem 2.4.5 *Let* $H(x) = \int_{-\infty}^{x} \frac{1}{2} e^{-|t|} dt$ *in (2.1.2),* $0 < \epsilon < \frac{1}{2}$, $0 < \alpha < \frac{1}{2}$ *and* $1 \le s_1 < s_2 < \infty$. *If* $\gamma_2 < \sqrt{s_1}$, *then we have*

$$\begin{aligned}
(i) \quad \sup_{G \in \mathcal{G}_n} e(\bar{X}_\alpha, \bar{X}, G) &\le e^*(\bar{X}_\alpha, \bar{X}) \\
&\le \frac{(1-2\alpha)^2(1-\epsilon+\epsilon s_2)}{(1-\epsilon)\Psi_0(1,\gamma_1) + \epsilon\Psi_0(s_2,\gamma_1) + \alpha\gamma_1^2}, \\
(ii) \quad \frac{(1-2\alpha)^2(1-\epsilon+\epsilon s_1)}{(1-\epsilon)\Psi_0(1,\gamma_2) + \epsilon\Psi_0(s_1,\gamma_2) + \alpha\gamma_2^2} &\le e_*(\bar{X}_\alpha, \bar{X}) \\
&\le \inf_{G \in \mathcal{G}_n} e(\bar{X}_\alpha, \bar{X}, G),
\end{aligned}$$

where $\Psi_0(w,\gamma) = w - \left(\frac{1}{2}\gamma^2 + \gamma\sqrt{w} + w\right) e^{-\frac{\gamma}{\sqrt{w}}}$ *and* γ_i *is the solution of* $(1-\epsilon)H(\gamma) + \epsilon H\left(\frac{\gamma}{\sqrt{s_i}}\right) = 1-\alpha$, $i = 1, 2$.

Moreover, we have that a sufficient condition for $\gamma_2 < \sqrt{s_1}$ *is* $H^{-1}\left(\frac{1-\alpha-\frac{1}{2}\epsilon}{1-\epsilon}\right) \le \sqrt{s_1}$.

Proof: For $H(x) = \int_{-\infty}^{x} \frac{1}{2} e^{-|t|} dt$, we have $h(x) = \frac{dH(x)}{dx} = \frac{1}{2} e^{-|x|}$ and $\sigma_h^2 = \int x^2 h(x) dx = \int_0^\infty x^2 e^{-x} dx = 2$. We define

$$C_1 = \{(u,v) : u = s, v = sk(s), \quad k(s) = \int_0^{\gamma/\sqrt{s}} \tfrac{1}{2} y^2 e^{-y} dy, s_1 \le s \le s_2\}.$$

Let $v(s) = sk(s) = \frac{1}{2} s \int_0^{\gamma/\sqrt{s}} y^2 e^{-y} dy$. Then we get

$$\begin{aligned}
\frac{dv(s)}{ds} &= \frac{1}{2}\left[\int_0^{\gamma/\sqrt{s}} y^2 e^{-y} dy + s\frac{\gamma^2}{s} e^{-\frac{\gamma}{\sqrt{s}}} \left(-\frac{1}{2} s^{-\frac{3}{2}} \cdot \gamma\right)\right] \\
&= \frac{1}{2}\left(\int_0^{\gamma/\sqrt{s}} y^2 e^{-y} dy - \frac{1}{2}\gamma^3 s^{-\frac{3}{2}} e^{-\gamma/\sqrt{s}}\right),
\end{aligned}$$

and

$$\begin{aligned}\frac{d^2v(s)}{ds^2} &= \frac{1}{2}\{\frac{\gamma^2}{s}e^{-\frac{\gamma}{\sqrt{s}}}\cdot(-\frac{1}{2}\gamma s^{-\frac{3}{2}}) \\ &\quad -\frac{\gamma^3}{2}[s^{-\frac{3}{2}}e^{-\gamma/\sqrt{s}}(\frac{1}{2}\gamma s^{-\frac{3}{2}}) + e^{-\gamma/\sqrt{s}}(-\frac{3}{2}s^{-\frac{5}{2}})]\} \\ &= \frac{1}{2}\left[-\frac{1}{2}\gamma^3 s^{-\frac{5}{2}}\cdot e^{-\frac{\gamma}{\sqrt{s}}} - \frac{1}{4}\gamma^4 s^{-3}e^{-\gamma/\sqrt{s}} + \frac{3}{4}\gamma^3 s^{-\frac{5}{2}}e^{-\gamma/\sqrt{s}}\right] \\ &= \frac{1}{8}\gamma^3 s^{-\frac{5}{2}}e^{-\gamma/\sqrt{s}}\left(1-\frac{\gamma}{\sqrt{s}}\right) > 0,\end{aligned}$$

since $\gamma \le \gamma_2 < \sqrt{s_1}$.

The second derivative $\frac{d^2v(s)}{ds^2} > 0$ implies that C_1 is convex. On the other hand, let

$$L(z) = \frac{1}{2}\left(\int_0^z y^2e^{-y}dy - \frac{1}{2}z^3e^{-z}\right),$$

we have

$$\frac{dL(z)}{dz} = \frac{1}{2}\left\{z^2e^{-z} - \frac{1}{2}[z^3e^{-z}(-1) + e^{-z}3z^2]\right\} = \frac{1}{4}z^2e^{-z}(z-1) < 0$$

if $0 < z < 1$. Note that $L(0) = 0$, hence we have $L(z) < 0,\ z \in (0,1)$. Note also $L(\frac{\gamma}{\sqrt{s}}) = \frac{dv(s)}{ds}$, and $0 < \frac{\gamma}{\sqrt{s}} < 1$. We conclude that $\frac{dv(s)}{ds} < 0$. Hence, C_1 is monotone decreasing.

By Theorem 2.4.3 (iii,a) and (iii,b), we have

$$\begin{aligned}&\sup_{G\in\mathcal{T}} e(\bar{X}_\alpha, \bar{X}, G) = e\left(\bar{X}_\alpha, \bar{X}, \Delta_{s_2}\right) \\ &= \frac{(1-2\alpha)^2(1-\epsilon+\epsilon s_2)}{(1-\epsilon)\int_0^\gamma \frac{1}{2}x^2e^{-x}dx + \epsilon s_2 k(s_2) + \alpha\gamma^2} \\ &= \frac{(1-2\alpha)^2(1-\epsilon+\epsilon s_2)}{(1-\epsilon)[1-(\frac{1}{2}\gamma^2+\gamma+1)e^{-\gamma}] + \epsilon[s_2 - (\frac{1}{2}\gamma^2 + \gamma\sqrt{s_2} + s_2)]e^{-\frac{\gamma}{\sqrt{s_2}}} + \alpha\gamma^2} \\ &= \frac{(1-2\alpha)^2(1-\epsilon+\epsilon s_2)}{(1-\epsilon)\Psi_0(1,\gamma) + \epsilon\Psi_0(s_2,\gamma) + \alpha\gamma^2},\end{aligned}$$

and similarly,

$$\inf_{G \in \mathcal{T}} e(\bar{X}_\alpha, \bar{X}, G) = e(\bar{X}_\alpha, \bar{X}, \Delta_{s_1}) = \frac{(1-2\alpha)^2(1-\epsilon+\epsilon s_1)}{(1-\epsilon)\Psi_0(1,\gamma)+\epsilon\Psi_0(s_1,\gamma)+\alpha\gamma^2},$$

where $\Psi_0(\omega,\gamma) = \omega - (\frac{1}{2}\gamma^2 + \gamma\sqrt{\omega} + \omega)e^{-\frac{\gamma}{\sqrt{\omega}}}$.

Similar to the proof of Theorem 2.4.4, we define

$$K_1(\gamma) \quad = (1-\epsilon)\Psi_0(1,\gamma) + \epsilon\Psi_0(s,\gamma) + \alpha\gamma^2$$

and

$$K_2(\gamma) \quad = (1-\epsilon)\Psi_0(1,\gamma) + \epsilon\Psi_0(s_2,\gamma) + \alpha\gamma^2.$$

Then we have

$$\begin{aligned} \frac{dK_1(\gamma)}{d\gamma} &= (1-\epsilon)[-(\frac{1}{2}\gamma^2+\gamma+1)e^{-\gamma}(-1) + e^{-\gamma}(-\gamma-1)] \\ &\quad +\epsilon[-(\frac{1}{2}\gamma^2+\gamma\sqrt{s_1}+s_1)e^{-\frac{\gamma}{\sqrt{s_1}}}(\frac{-1}{\sqrt{s_1}}) \\ &\quad +e^{-\frac{\gamma}{\sqrt{s_1}}}(-\gamma-\sqrt{s_1})] + 2\alpha\gamma \\ &= \frac{(1-\epsilon)}{2}\gamma^2 e^{-\gamma} + \frac{\epsilon}{2\sqrt{s_1}}\gamma^2 e^{-\frac{\gamma}{\sqrt{s_1}}} + 2\alpha\gamma > 0 \end{aligned}$$

and

$$\frac{dK_2(\gamma)}{d\gamma} = \frac{(1-\epsilon)}{2}\gamma^2 e^{-\gamma} + \frac{\epsilon}{2\sqrt{s_2}}\gamma^2 e^{-\frac{\gamma}{\sqrt{s_2}}} + 2\alpha\gamma > 0.$$

Hence, both $K_1(\gamma)$ and $K_2(\gamma)$ are increasing in γ. This yields the following conclusion:

$$\max_{\gamma_1 \le \gamma \le \gamma_2} e(\bar{X}_\alpha, \bar{X}, \Delta_{s_2}) = \frac{(1-2\alpha)^2(1-\epsilon+\epsilon s_2)}{(1-\epsilon)\Psi_0(1,\gamma_1)+\epsilon\Psi_0(s_2,\gamma_1)+\alpha\gamma_1^2}$$

and

$$\min_{\gamma_1 \le \gamma \le \gamma_2} e(\bar{X}_\alpha, \bar{X}, \Delta_{s_1}) = \frac{(1-2\alpha)^2(1-\epsilon+\epsilon s_1)}{(1-\epsilon)\Psi_0(1,\gamma_2)+\epsilon\Psi_0(s_1,\gamma_2)+\alpha\gamma_2^2}. \qquad \square$$

The problem of finding the bounds of $e(M, \bar{X}_\alpha, G)$ is very similar to the situation of $e(\bar{X}_\alpha, \bar{X}, G)$. When $H(x) = \Phi(x)$, we have the following:

Theorem 2.4.6 *Let $H(x) = \Phi(x)$ in (2.1.2), $0 < \epsilon < \frac{1}{2}$, $0 < \alpha < \frac{1}{2}$, and $1 \le s_1 < s_2 < \infty$. If $\gamma_2 < \sqrt{s_1}$, then we have*

$$\begin{aligned}
(i)\quad \sup_{G\in\mathcal{G}_n} e(M,\bar{X}_\alpha,G) &\le e^*(M,\bar{X}_\alpha)\\
&\le \frac{4(1-\epsilon+\frac{\epsilon}{\sqrt{s_1}})^2}{\pi(1-2\alpha)^2}\\
&\quad\cdot[(1-\epsilon)\Psi(1,\gamma_2)+\epsilon\Psi(s_1,\gamma_2)+\alpha\gamma_2^2],\\
(ii)\quad \frac{4(1-\epsilon+\frac{\epsilon}{\sqrt{s_2}})^2}{\pi(1-2\alpha)^2}[(1-\epsilon)\Psi(1,\gamma_1)+\epsilon\Psi(s_2,\gamma_1)+\alpha\gamma_1^2] &\le e_*(M,\bar{X}_\alpha)\\
&\le \inf_{G\in\mathcal{G}_n} e(M,\bar{X}_\alpha,G),
\end{aligned}$$

where $\Psi(w,\gamma)=w\Phi\left(\frac{\gamma}{\sqrt{w}}\right)-\frac{1}{2}w-\frac{\gamma\sqrt{w}}{\sqrt{2\pi}}e^{-\frac{\gamma^2}{2w}}$, *and* γ_i *is the solution of* $(1-\epsilon)\Phi(\gamma)+\epsilon\Phi\left(\frac{\gamma}{\sqrt{s_i}}\right)=1-\alpha$, $i=1,2$. *Moreover, we have that a sufficient condition for* $\gamma_2<\sqrt{s_1}$ *is* $\Phi^{-1}\left(\frac{1-\alpha-\frac{1}{2}\epsilon}{1-\epsilon}\right)\le\sqrt{s_1}$.

Proof: For $H(x)=\Phi(x)$, we define

$$C_2=\{(u,v): u=\frac{1}{\sqrt{s}}, v=sk(s),\ k(s)=\int_0^{\frac{\gamma}{\sqrt{s}}} y^2\varphi(y)dy,\ s_1\le s\le s_2\}.$$

Consider v as a function of u, $v=v(u)$, and s as a parameter. We find

$$\begin{aligned}
\frac{dv(u)}{du}=\frac{dv(s)/ds}{du(s)/ds} &= \frac{1}{\sqrt{2\pi}}\left[\int_0^{\frac{\gamma}{\sqrt{s}}} y^2e^{-\frac{y^2}{2}}dy+s\frac{\gamma^2}{s}e^{-\frac{\gamma^2}{2s}}\left(-\frac{1}{2}\gamma s^{-\frac{3}{2}}\right)\right](-2s^{\frac{3}{2}})\\
&= \frac{1}{\sqrt{2\pi}}s^{\frac{3}{2}}\left[\left(\frac{\gamma}{\sqrt{s}}\right)^3e^{-\frac{\gamma^2}{2s}}-2\int_0^{\frac{\gamma}{\sqrt{s}}} y^2e^{-\frac{y^2}{2}}dy\right]
\end{aligned}$$

and

$$\begin{aligned}
\frac{d^2v(u)}{du^2} &= \frac{d\left(\frac{dv(w)}{du}\right)/ds}{du(s)/ds}\\
&= \frac{1}{\sqrt{2\pi}}\left\{\gamma^3e^{-\frac{\gamma^2}{2s}}\frac{\gamma^2}{2s^2}-2[\left(\int_0^{\frac{\gamma}{\sqrt{s}}} y^2e^{-\frac{y^2}{2}}dy\right)\frac{3}{2}s^{\frac{1}{2}}+s^{\frac{3}{2}}\frac{\gamma^2}{s}e^{-\frac{\gamma^2}{2s}}(-\frac{1}{2}\gamma s^{-\frac{3}{2}})]\right\}\cdot(-2s^{\frac{3}{2}})\\
&= -\frac{2s^{\frac{3}{2}}}{\sqrt{2\pi}}\left(\frac{\gamma^5}{2s^2}e^{-\frac{\gamma^2}{2s}}-3s^{\frac{1}{2}}\int_0^{\frac{\gamma}{\sqrt{s}}} y^2e^{-\frac{y^2}{2}}dy+\frac{\gamma^3}{s}e^{-\frac{\gamma^2}{2s}}\right)\\
&= \frac{s^2}{\sqrt{2\pi}}\left[6\int_0^{\frac{\gamma}{\sqrt{s}}} y^2e^{-\frac{y^2}{2}}dy-\left(\frac{\gamma}{\sqrt{s}}\right)^5e^{-\frac{\gamma^2}{2s}}-2\left(\frac{\gamma}{\sqrt{s}}\right)^3e^{-\frac{\gamma^2}{2s}}\right].
\end{aligned}$$

We conclude that C_2 is a concave and increasing curve by showing $\frac{d^2v(u)}{du^2}<0$, and

$\dfrac{dv(u)}{du} > 0$. For this purpose, let us define

$$L_1(z) = z^3 e^{-\frac{z^2}{2}} - 2\int_0^z y^2 e^{-\frac{y^2}{2}} dy$$

and

$$L_2(z) = 6\int_0^z y^2 e^{-\frac{y^2}{2}} dy - z^5 e^{-\frac{z^2}{2}} - 2z^3 e^{-\frac{z^2}{2}}.$$

We find

$$\frac{dL_1(z)}{dz} = e^{-\frac{z^2}{2}}\left[z^3(-z) + 3z^2\right] - 2z^2 e^{-\frac{z^2}{2}} = z^2 e^{-\frac{z^2}{2}}(1 - z^2) > 0$$

if $z \in (0,1)$. Note that $L_1(0) = 0$. Hence $L_1(z) > 0$ for $z \in (0,1)$. On the other hand, we know that $L_1\left(\dfrac{\gamma}{\sqrt{s}}\right) = \dfrac{dv(u)}{du}$ and $0 < \dfrac{\gamma}{\sqrt{s}} < 1$, since we have $\gamma_2 < \sqrt{s_1}$. This yields $\dfrac{dv(u)}{du} > 0$. We also find

$$\begin{aligned}\frac{dL_2(z)}{dz} &= 6z^2 e^{-\frac{z^2}{2}} - e^{-\frac{z^2}{2}}\left[z^5(-z) + 5z^4\right] - 2e^{-\frac{z^2}{2}}\left[z^3(-z) + 3z^2\right] \\ &= z^4 e^{-\frac{z^2}{2}}(z^2 - 3) < 0,\end{aligned}$$

at least for $z \in (0,1)$. We have $L_2(z) < 0$ for $z \in (0,1)$, since $L_2(0) = 0$. The fact that $L_2(\frac{\gamma}{\sqrt{s}}) = \dfrac{d^2v(u)}{du^2}$ and $0 < \dfrac{\gamma}{\sqrt{s}} < 1$ implies $\dfrac{d^2v(u)}{du^2} < 0$. Therefore, C_2 is a concave and increasing curve. By Theorem 2.4.3 (ii,c) and (ii,d), we have

$$\begin{aligned}\sup_{G\in T} e(M, \bar{X}_\alpha, G) &= e(M, \bar{X}_\alpha, \Delta_{s_1}) \\ &= \frac{4(1-\epsilon+\frac{\epsilon}{\sqrt{s_1}})^2}{\pi(1-2\alpha)^2} \\ &\quad\cdot\left[(1-\epsilon)\int_0^\gamma y^2 \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy + \epsilon s_1 \int_0^{\frac{\gamma}{\sqrt{s_1}}} y^2 \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy + \alpha\gamma^2\right] \\ &= \frac{4(1-\epsilon+\frac{\epsilon}{\sqrt{s_1}})^2}{\pi(1-2\alpha)^2}\left[(1-\epsilon)\Psi(1,\gamma) + \epsilon\Psi(s_1,\gamma) + \alpha\gamma^2\right]\end{aligned}$$

and

$$\begin{aligned}\inf_{G\in T} e(M, \bar{X}_\alpha, G) &= e(M, \bar{X}_\alpha, \Delta_{s_2}) \\ &= \frac{4(1-\epsilon+\frac{\epsilon}{\sqrt{s_2}})^2}{\pi(1-2\alpha)^2}\left[(1-\epsilon)\Psi(1,\gamma) + \epsilon\Psi(s_2,\gamma) + \alpha\gamma^2\right],\end{aligned}$$

where $\Psi(w,\gamma) = w\Phi(\frac{\gamma}{\sqrt{w}}) - \frac{1}{2}w - \frac{\gamma\sqrt{w}}{\sqrt{2\pi}}e^{-\frac{\gamma^2}{2w}}$. Furthermore, let

$$\begin{aligned} K_1(\gamma) &= (1-\epsilon)\Psi(1,\gamma) + \epsilon\Psi(s_1,\gamma) + \alpha\gamma^2 \\ &= \frac{(1-\epsilon)}{\sqrt{2\pi}} \int_0^\gamma y^2 e^{-\frac{y^2}{2}} dy + \frac{\epsilon s_1}{\sqrt{2\pi}} \int_0^{\frac{\gamma}{\sqrt{s_1}}} y^2 e^{-\frac{y^2}{2}} dy + \alpha\gamma^2 \end{aligned}$$

and

$$\begin{aligned} K_2(\gamma) &= (1-\epsilon)\Psi(1,\gamma) + \epsilon\Psi(s_2,\gamma) + \alpha\gamma^2 \\ &= \frac{(1-\epsilon)}{\sqrt{2\pi}} \int_0^\gamma y^2 e^{-\frac{y^2}{2}} dy + \frac{\epsilon s_2}{\sqrt{2\pi}} \int_0^{\frac{\gamma}{\sqrt{s_2}}} y^2 e^{-\frac{y^2}{2}} dy + \alpha\gamma^2. \end{aligned}$$

Then we have

$$\frac{dK_1(\gamma)}{d\gamma} = \frac{(1-\epsilon)}{\sqrt{2\pi}}\gamma^2 e^{-\frac{\gamma^2}{2}} + \frac{\epsilon s_1}{\sqrt{2\pi}}\frac{\gamma^2}{s_1}e^{-\frac{\gamma^2}{2s_1}}\frac{1}{\sqrt{s_1}} + 2\alpha\gamma > 0$$

and

$$\frac{dK_2(\gamma)}{d\gamma} = \frac{(1-\epsilon)}{\sqrt{2\pi}}\gamma^2 e^{-\frac{\gamma^2}{2}} + \frac{\epsilon s_2}{\sqrt{2\pi}} \cdot \frac{\gamma^2}{s_2}e^{-\frac{\gamma^2}{2s_2}}\frac{1}{\sqrt{s_2}} + 2\alpha\gamma > 0.$$

Since $K_1(\gamma)$ and $K_2(\gamma)$ are increasing in γ, we then have

$$\max_{\gamma_1 \le \gamma \le \gamma_2} e(M, \bar{X}_\alpha, \Delta_{s_1}) = \frac{4(1-\epsilon+\frac{\epsilon}{\sqrt{s_1}})^2}{\pi(1-2\alpha)^2}\left[(1-\epsilon)\Psi(1,\gamma_2) + \epsilon\Psi(s_1,\gamma_2) + \alpha\gamma_2^2\right]$$

and

$$\min_{\gamma_1 \le \gamma \le \gamma_2} e(M, \bar{X}_\alpha, \Delta_{s_2}) = \frac{4(1-\epsilon+\frac{\epsilon}{\sqrt{s_2}})^2}{\pi(1-2\alpha)^2}\left[(1-\epsilon)\Psi(1,\gamma_1) + \epsilon\Psi(s_2,\gamma_1) + \alpha\gamma_1^2\right].$$

Theorem 2.4.6 follows by Theorem 2.4.2 and the above results. □

The situation when $H(x) = \int_{-\infty}^x \frac{1}{2}e^{-|t|}dt$ is very similar to the case of $H(x) = \Phi(x)$. We simply state the result without proof.

Theorem 2.4.7 *Let* $H(x) = \int_{-\infty}^x \frac{1}{2}e^{-|t|}dt$ *in (2.1.2),* $0 < \epsilon < \frac{1}{2}$, $0 < \alpha < \frac{1}{2}$, *and* $1 \le s_1 < s_2 < \infty$. *If* $\gamma_2 < \sqrt{s_1}$, *then we have*

(i)
$$\begin{aligned} \sup_{G\in\mathcal{G}_n} e(M,\bar{X}_\alpha,G) &\le e^*(M,\bar{X}_\alpha) \\ &\le \frac{2\left(1-\epsilon+\frac{\epsilon}{\sqrt{s_1}}\right)^2}{(1-2\alpha)^2} \\ &\quad \cdot\left[(1-\epsilon)\Psi_0(1,\gamma_2) + \epsilon\Psi_0(s_1,\gamma_2) + \alpha\gamma_2^2\right], \end{aligned}$$

$$(ii)\quad \frac{2\left(1-\epsilon+\frac{\epsilon}{\sqrt{s_2}}\right)^2}{(1-2\alpha)^2} \cdot \left[(1-\epsilon)\Psi_0(1,\gamma_1)+\epsilon\Psi_0(s_2,\gamma_1)+\alpha\gamma_1^2\right] \le e_*(M,\bar{X}_\alpha) \le \inf_{G\in\mathcal{G}_n} e(M,\bar{X}_\alpha,G),$$

where $\Psi_0(w,\gamma)=w-\left(\frac{1}{2}\gamma^2+\gamma\sqrt{w}+w\right)e^{-\frac{\gamma}{\sqrt{w}}}$ *and* γ_i *is the solution of* $(1-\epsilon)H(\gamma)+\epsilon H\left(\frac{\gamma}{\sqrt{s_i}}\right)=1-\alpha$, $i=1,2$. *Moreover, we have that a sufficient condition for* $\gamma_2<\sqrt{s_1}$ *is* $H^{-1}\left(\frac{1-\alpha-\frac{1}{2}\epsilon}{1-\epsilon}\right)\le\sqrt{s_1}$.

2.5 Some Numerical Results and Comments

In the last two sections, we discussed the bounds on the asymptotic relative efficiencies among the median M, α-trimmed mean $\bar{X}_\alpha$, and the sample mean $\bar{X}$. We found the explicit solutions for $e_*(M,\bar{X})$ and $e^*(M,\bar{X})$. When $H(x) = \Phi(x)$, we have

$$e_*(M,\bar{X}) = \frac{2}{\pi}(1-\epsilon+\epsilon s_1)(1-\epsilon+\frac{\epsilon}{\sqrt{s_1}})^2,$$

and

$$e^*(M,\bar{X}) = \begin{cases} \frac{2}{\pi}\{(1-\epsilon)+\epsilon[\lambda_2 s_2+(1-\lambda_2)s_1]\} & \\ \quad\cdot\left[(1-\epsilon)+\epsilon\left(\frac{\lambda_2}{\sqrt{s_2}}+\frac{1-\lambda_2}{\sqrt{s_1}}\right)\right]^2 & \text{if } 0<\lambda_2<1 \\ \frac{2}{\pi}(1-\epsilon+\epsilon s_2)\left(1-\epsilon+\frac{\epsilon}{\sqrt{s_2}}\right)^2 & \text{if } \lambda_2\geq 1, \end{cases}$$

where

$$\lambda_2 = \frac{2\left(\frac{1}{\sqrt{s_2}}-\frac{1}{\sqrt{s_1}}\right)(1-\epsilon+\epsilon s_1)+(s_2-s_1)\left(1-\epsilon+\frac{\epsilon}{\sqrt{s_1}}\right)}{3\epsilon(s_2-s_1)\left(\frac{1}{\sqrt{s_1}}-\frac{1}{\sqrt{s_2}}\right)}.$$

Hence, $e_*(M,\bar{X})$ and $e^*(M,\bar{X})$ can be easily calculated. For some different values of s_1, s_2 and ϵ, the corresponding results of $e_*(M,\bar{X})$ and $e^*(M,\bar{X})$ are presented in Table 2.5.1. Note that we have two numbers for each cell inside the table. The top one is $e_*(M,\bar{X})$ and the bottom one is $e^*(M,\bar{X})$. With the results in Table 2.5.1 and some further calculation, we are able to answer a general version of a question raised by Tukey (1960): given s_1 and s_2, how large an $\epsilon > 0$ is required for the infimum of the asymptotic relative efficiency of M with respect to $\bar{X}$ over $\mathcal{F}$ to exceed 1? In this case, one understands that even in the sense of asymptotic relative efficiency, the median M is still preferable than the sample mean $\bar{X}$. There are some other features have been observed from Table 2.5.1. Firstly, we note that, for given s_1 and s_2, $e_*(M,\bar{X})$ and $e^*(M,\bar{X})$ are monotone nondecreasing (monotone increasing in most of the cases) in ϵ. Intuitively, it also makes sense as we know that ϵ is the proportion of contaminations. Secondly, we find that $e_*(M,\bar{X})$ is monotone increasing in s_1 when ϵ and s_2 are given. This is obvious since we have the fact that $L(s) := (1-\epsilon+\epsilon s)(1-\epsilon+\frac{\epsilon}{\sqrt{s}})^2$ is a function increasing in s, when $s \geq 1$. In the cases we considered in Table 2.5.1, we also have $e^*(M, \bar{X})$ is monotone increasing in s_2 when ϵ and s_1 are given. In these cases, we always get $e^*(M, \bar{X}, \Delta_{s_2})$. In general, this may not always happen. For example, when $s_1 = 1$, $s_2 = 36$, and $\epsilon = 0.4$, we have $\lambda_2 = 0.952$ and $e^*(M, \bar{X}) = e(M, \bar{X}, G^*) = 4.2509$ where $G^* = \lambda_2 \Delta_{s_2} + (1-\lambda_2)\Delta s_1$. On the other hand, we have $e(M, \bar{X}, \Delta_{s_2}) = 4.2441$. We also note that the differences between $e^*(M, \bar{X})$ and $e_*(M, \bar{X})$ are small when s_1, s_2, ϵ, or $s_2 - s_1$ are relatively small. This reflects the fact that the asymptotic relative efficiencies are stable over the corresponding class of distribution functions $\mathcal{F}$.

The situations are more complicated for the asymptotic relative efficiencies of $\bar{X}_\alpha$ with respect to $\bar{X}$, and M with respect to $\bar{X}_\alpha$. As we mentioned before, we are unable to find the exact values of $e_*(\bar{X}_\alpha, \bar{X})$, $e^*(\bar{X}_\alpha, \bar{X})$, $e_*(M, \bar{X}_\alpha)$, and $e^*(M, \bar{X}_\alpha)$. Instead, we can only find a range for each of them. For the sake of argument, we define

$$\begin{aligned}
LL(\bar{X}_\alpha, \bar{X}) &:= \frac{(1-2\alpha)}{2} \cdot \frac{1-\epsilon+\epsilon s_1}{(1-\epsilon)\Psi(1,\gamma_2)+\epsilon\Psi(s_1,\gamma_2)+\alpha\gamma_2^2}, \\
UU(\bar{X}_\alpha, \bar{X}) &:= \frac{(1-2\alpha)}{2} \cdot \frac{1-\epsilon+\epsilon s_2}{(1-\epsilon)\Psi(1,\gamma_1)+\epsilon\Psi(s_2,\gamma_1)+\alpha\gamma_1^2}, \\
LL(M, \bar{X}_\alpha) &:= \frac{4}{\pi(1-2\alpha)^2} \cdot (1-\epsilon+\frac{\epsilon}{\sqrt{s_2}})^2 \\
&\quad \cdot [(1-\epsilon)\Psi(1,\gamma_1)+\epsilon\Psi(s_2,\gamma_1)+\alpha\gamma_1^2],
\end{aligned}$$

and

$$\begin{aligned}
UU(M, \bar{X}_\alpha) &:= \frac{4}{\pi(1-2\alpha)^2} \cdot (1-\epsilon+\frac{\epsilon}{\sqrt{s_1}})^2 \\
&\quad \cdot [(1-\epsilon)\Psi(1,\gamma_2)+\epsilon\Psi(s_1,\gamma_2)+\alpha\gamma_2^2],
\end{aligned}$$

where $\Psi(w,\gamma) = w\Phi(\frac{\gamma}{\sqrt{w}}) - \frac{1}{2}w - \frac{\gamma\sqrt{w}}{\sqrt{2\pi}}e^{-\frac{\gamma^2}{2w}}$, and γ_i is the solution of $(1-\epsilon)\Phi(\gamma) + \epsilon\Phi(\frac{\gamma}{\sqrt{s_i}}) = 1-\alpha$, $i = 1, 2$. We also define

$$\begin{aligned}
UL_n(\bar{X}_\alpha, \bar{X}) &:= \inf_{G\in\mathcal{G}_n} e(\bar{X}_\alpha, \bar{X}, G) \\
&= \inf_{G\in\mathcal{G}_n} \frac{(1-2\alpha)^2}{2} \cdot \frac{1-\epsilon+\epsilon\sum_{i=1}^n p_i x_i}{(1-\epsilon)\Psi(1,\gamma)+\epsilon\sum_{i=1}^n p_i\Psi(x_i,\gamma)+\alpha\gamma^2},
\end{aligned}$$

$$\begin{aligned}
LU_n(\bar{X}_\alpha, \bar{X}) &:= \sup_{G \in \mathcal{G}_n} e(\bar{X}_\alpha, \bar{X}, G) \\
&= \sup_{G \in \mathcal{G}_n} \frac{(1-2\alpha)^2}{2} \cdot \frac{1-\epsilon+\epsilon\sum_{i=1}^n p_i x_i}{(1-\epsilon)\Psi(1,\gamma)+\epsilon\sum_{i=1}^n p_i \Psi(x_i,\gamma)+\alpha\gamma^2}, \\
UL_n(M, \bar{X}_\alpha) &:= \inf_{G \in \mathcal{G}_n} e(M, \bar{X}_\alpha, G) \\
&= \inf_{G \in \mathcal{G}_n} \frac{4}{\pi(1-2\alpha)^2} \cdot (1-\epsilon+\epsilon\textstyle\sum_{i=1}^n \frac{p_i}{\sqrt{x_i}})^2 \\
&\quad \cdot [(1-\epsilon)\Psi(1,\gamma)+\epsilon\textstyle\sum_{i=1}^n p_i \Psi(x_i,\gamma)+\alpha\gamma^2],
\end{aligned}$$

and

$$\begin{aligned}
LU_n(M, \bar{X}_\alpha) &:= \sup_{G \in \mathcal{G}_n} e(M, \bar{X}_\alpha, G) \\
&= \sup_{G \in \mathcal{G}_n} \frac{4}{\pi(1-2\alpha)^2} \cdot (1-\epsilon+\epsilon\textstyle\sum_{i=1}^n \frac{p_i}{\sqrt{x_i}})^2 \\
&\quad \cdot [(1-\epsilon)\Psi(1,\gamma)+\epsilon\textstyle\sum_{i=1}^n p_i \Psi(x_i,\gamma)+\alpha\gamma^2],
\end{aligned}$$

where $G = \sum_{i=1}^n p_i \Delta_{x_i}$ with $s_1 \le x_i \le s_2$, $0 \le p_i \le 1$, $\sum_{i=1}^n p_i = 1$, and γ is the solution of $(1-\epsilon)\Phi(\gamma) + \epsilon\sum_{i=1}^n p_i \Phi(\frac{\gamma}{\sqrt{x_i}}) = 1-\alpha$.

According to Theorem 2.4.4 and Theorem 2.4.6, we know that $LL(\bar{X}_\alpha, \bar{X})$ and $UL_n(\bar{X}_\alpha, \bar{X})$ are lower and upper bounds of $e_*(\bar{X}_\alpha, \bar{X})$. Similarly, $LU_n(\bar{X}_\alpha, \bar{X})$ and $UU(\bar{X}_\alpha, \bar{X})$, $LL(M, \bar{X}_\alpha)$ and $UL_n(M, \bar{X}_\alpha)$, $LU_n(M, \bar{X}_\alpha)$ and $UU(M, \bar{X}_\alpha)$ are lower and upper bounds of $e^*(\bar{X}_\alpha, \bar{X})$, $e_*(M, \bar{X}_\alpha)$ and $e^*(M, \bar{X}_\alpha)$ respectively. It is clear that $UL_n(\bar{X}_\alpha, \bar{X})$ and $UL_n(M, \bar{X}_\alpha)$ are monotone nonincreasing and bounded below by $LL(\bar{X}_\alpha, \bar{X})$ and $LL(M, \bar{X}_\alpha)$; $LU_n(\bar{X}_\alpha, \bar{X})$ and $LU_n(M, \bar{X}_\alpha)$ are monotone nondecreasing and bounded above by $UU(\bar{X}_\alpha, \bar{X})$ and $UU(M, \bar{X}_\alpha)$. Hence these four sequences have limits and the limits when n goes to infinity will be the exact values of $e_*(\bar{X}_\alpha, \bar{X})$, $e_*(M, \bar{X})$, $e^*(\bar{X}_\alpha, \bar{X})$, and $e^*(M, \bar{X}_\alpha)$. In practice, we can only do the numerical search when n is small. For $n = 1$, and for some different values of α, ϵ, s_1 and s_2 we present the corresponding bounds of $e_*(\bar{X}_\alpha, \bar{X})$ and $e^*(\bar{X}_\alpha, \bar{X})$ through Table 2.5.2 to Table 2.5.6. Note that each cell of these tables has four numbers. The first one is $LL(\bar{X}_\alpha, \bar{X})$, the second $UL_1(\bar{X}_\alpha, \bar{X})$; the third $LU_1(\bar{X}_\alpha, \bar{X})$ and the fourth $UU(\bar{X}_\alpha, \bar{X})$. Similarly, the bounds of $e_*(M, \bar{X}_\alpha)$ and $e^*(M, \bar{X}_\alpha)$ are presented in Table 2.5.7 through Table 2.5.11. There are some missing values inside the tables. These are the cases when $\gamma_2 > \sqrt{s_1}$; hence, Theorem 2.4.4 and Theorem 2.4.6 do not apply to the calculation of $LL(\bar{X}_\alpha, \bar{X})$, $UU(\bar{X}_\alpha, \bar{X})$, $LL(M, \bar{X}_\alpha)$, and $UU(M, \bar{X}_\alpha)$.

In many cases we note that the differences between $LL(\bar{X}_\alpha, \bar{X})$ and $UL_1(\bar{X}_\alpha, \bar{X})$; $LU_1(\bar{X}_\alpha, \bar{X})$ and $UU(\bar{X}_\alpha, \bar{X})$; $LL(M, \bar{X}_\alpha)$ and $UL_1(M, \bar{X}_\alpha)$; $LU_1(M, \bar{X}_\alpha)$ and $UU(M, \bar{X}_\alpha)$ are very small. Hence, the bounds we provided there are very accurate. In some cases, even the differences between $LL(\bar{X}_\alpha, \bar{X})$ and $UU(\bar{X}_\alpha, \bar{X})$; $LL(M, \bar{X}_\alpha)$ and $UU(M, \bar{X}_\alpha)$ are small. This fact reflects the stability of the asymptotic relative efficiencies, since we always have $e^*(\bar{X}_\alpha, \bar{X}) - e_*(\bar{X}_\alpha, \bar{X}) \leq UU(\bar{X}_\alpha, \bar{X}) - LL(\bar{X}_\alpha, \bar{X})$, and $e^*(M, \bar{X}_\alpha) - e_*(M, \bar{X}_\alpha) \leq UU(M, \bar{X}_\alpha) - LL(M, \bar{X}_\alpha)$. In the cases when the bounds on $e_*(\bar{X}_\alpha, \bar{X})$, $e^*(\bar{X}_\alpha, \bar{X})$, $e_*(M, \bar{X}_\alpha)$, and $e^*(M, \bar{X}_\alpha)$ are not too close, we can always get more precise results by increasing n in $UL_n(\bar{X}_\alpha, \bar{X})$, $LU_n(\bar{X}_\alpha, \bar{X})$, $UL_n(M, \bar{X}_\alpha)$, and $LU_n(M, \bar{X}_\alpha)$. We did some calculations of the above quantities when $n = 2$. The results are either the same or very close to the results when $n = 1$. Hence the exact values of $e_*(\bar{X}_\alpha, \bar{X})$, $e^*(\bar{X}_\alpha, \bar{X})$, $e_*(M, \bar{X}_\alpha)$, and $e^*(M, \bar{X}_\alpha)$ will be closer to $UL_n(\bar{X}_\alpha, \bar{X})$, $LU_n(\bar{X}_\alpha, \bar{X})$, $UL_n(M, \bar{X}_\alpha)$, and $LU_n(M, \bar{X}_\alpha)$ than to $LL(\bar{X}_\alpha, \bar{X})$, $UU(\bar{X}_\alpha, \bar{X})$, $LL(M, \bar{X}_\alpha)$, and $UU(M, \bar{X}_\alpha)$. We know that the sample mean $\bar{X}$ and the median M are the extreme cases of α-trimmed mean $\bar{X}_\alpha$ corresponding to $\alpha = 0$ and $\alpha = \frac{1}{2}$ respectively. Hence the topics we discussed in this chapter and the tables we presented in this section will provide a guideline to choose a suitable value of α, hence $\bar{X}_\alpha$, in the sense of the asymptotic relative efficiency when we have the assessment of the values of s_1, s_2, and ϵ.

Table 2.5.1 Values of $e_*(M, \bar{X})$ and $e^*(M, \bar{X})$ when $H(x) = \Phi(x)$

s_1	s_2	ϵ 0.01	0.05	0.1	0.15	0.25	0.375
1	4	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366
		0.6492	0.6960	0.7469	0.7898	0.8530	0.8931
	9	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366
		0.6784	0.8328	0.9982	1.1345	1.3263	1.4324
	16	0.6366	0.6366	0.6366	0.6366	0.6366	0.6366
		0.7212	1.0321	1.3618	1.6297	1.9963	2.1788
4	9	0.6492	0.6960	0.7469	0.7898	0.8530	0.8931
		0.6784	0.8328	0.9982	1.1345	1.3263	1.4324
	16	0.6492	0.6960	0.7469	0.7898	0.8530	0.8931
		0.7212	1.0321	1.3618	1.6297	1.9963	2.1788
	25	0.6492	0.6960	0.7469	0.7898	0.8530	0.8931
		0.7768	1.2908	1.8320	2.2678	2.8521	3.1194
9	16	0.6784	0.8328	0.9982	1.1345	1.3263	1.4324
		0.7212	1.0321	1.3618	1.6297	1.9963	2.1788
	25	0.6784	0.8328	0.9982	1.1345	1.3263	1.4324
		0.7768	1.2908	1.8320	2.2678	2.8521	3.1194
	36	0.6784	0.8328	0.9982	1.1345	1.3263	1.4324
		0.8452	1.6079	2.4072	3.0463	3.8902	4.2502
16	25	0.7212	1.0321	1.3618	1.6297	1.9963	2.1788
		0.7768	1.2908	1.8320	2.2678	2.8521	3.1194
	36	0.7212	1.0321	1.3618	1.6297	1.9963	2.1788
		0.8452	1.6079	2.4072	3.0463	3.8902	4.2502
	49	0.7212	1.0321	1.3618	1.6297	1.9963	2.1788
		0.9261	1.9830	3.0865	3.9642	5.1092	5.5696

Table 2.5.2 Bounds of $e_*(\bar{X}_\alpha, \bar{X})$ and $e^*(\bar{X}_\alpha, \bar{X})$ when $H(x) = \Phi(x)$

$\epsilon = 0.01$

		α					
s_1	s_2	0.01	0.05	0.1	0.15	0.25	0.375
						0.8267	0.7318
	4	0.9960	0.9744	0.9430	0.9092	0.8367	0.7397
		1.0059	0.9885	0.9586	0.9251	0.8524	0.7541
1						0.8627	0.7622
						0.8231	0.7291
	9	0.9960	0.9744	0.9430	0.9092	0.8367	0.7397
		1.0410	1.0286	0.9993	0.9653	0.8903	0.7880
						0.9050	0.7994
			0.9795	0.9520	0.9199	0.8487	0.7514
	9	1.0059	0.9885	0.9586	0.9251	0.8524	0.7541
		1.0410	1.0286	0.9993	0.9653	0.8903	0.7880
4			1.0380	1.0061	0.9708	0.8942	0.7909
			0.9745	0.9485	0.9172	0.8468	0.7500
	16	1.0059	0.9885	0.9586	0.9251	0.8524	0.7541
		1.0994	1.0907	1.0608	1.0253	0.9460	0.8376
			1.1064	1.0720	1.0342	0.9524	0.8422
		1.0305	1.0233	0.9956	0.9625	0.8883	0.7865
	16	1.0410	1.0286	0.9993	0.9653	0.8903	0.7880
		1.0994	1.0907	1.0608	1.0253	0.9460	0.8376
9		1.1106	1.0964	1.0647	1.0283	0.9482	0.8391
		1.0233	1.0200	0.9934	0.9607	0.8870	0.7857
	25	1.0410	1.0286	0.9993	0.9653	0.8903	0.7880
		1.1789	1.1730	1.1417	1.1038	1.0188	0.9022
		1.1992	1.1829	1.1485	1.1091	1.0225	0.9048
		1.0917	1.0871	1.0584	1.0234	0.9447	0.8366
	25	1.0994	1.0907	1.0608	1.0253	0.9460	0.8376
		1.1789	1.1730	1.1417	1.1038	1.0188	0.9022
16		1.1871	1.1768	1.1443	1.1058	1.0202	0.9032
		1.0862	1.0847	1.0567	1.0222	0.9439	0.8360
	36	1.0994	1.0907	1.0608	1.0253	0.9460	0.8376
		1.2784	1.2747	1.2413	1.2005	1.1083	0.9815
		1.2938	1.2818	1.2461	1.2041	1.1108	0.9833

Table 2.5.3 Bounds of $e_*(\bar{X}_\alpha, \bar{X})$ and $e^*(\bar{X}_\alpha, \bar{X})$ when $H(x) = \Phi(x)$

$\epsilon = 0.05$

		α					
s_1	s_2	0.01	0.05	0.1	0.15	0.25	0.375
1						0.7875	0.7007
	4	0.9960	0.9744	0.9430	0.9092	0.8367	0.7397
		1.0347	1.0373	1.0142	0.9833	0.9105	0.8078
						0.9674	0.8527
						0.7698	0.6877
	9	0.9960	0.9744	0.9430	0.9092	0.8367	0.7397
		1.1516	1.2117	1.1978	1.1670	1.0859	0.9660
						1.1802	1.0389
4			0.9878	0.9786	0.9548	0.8901	0.7928
	9	1.0347	1.0373	1.0142	0.9833	0.9105	0.8078
		1.1517	1.2117	1.1978	1.1670	1.0859	0.9660
			1.2724	1.2414	1.2018	1.1108	0.9843
			0.9589	0.9593	0.9399	0.8797	0.7853
	16	1.0347	1.0373	1.0142	0.9833	0.9105	0.8078
		1.3037	1.4781	1.4725	1.4392	1.3432	1.1966
			1.5986	1.5566	1.5055	1.3901	1.2308
9			1.1764	1.1742	1.1488	1.0733	0.9568
	16	1.1518	1.2117	0.1978	1.1670	1.0859	0.9660
		1.3038	1.4781	1.4725	1.4392	1.3432	1.1966
			1.5224	1.5052	1.4620	1.3590	1.2080
			1.1535	1.1595	1.1377	1.0657	0.9514
	25	1.1518	1.2117	1.1978	1.1670	1.0859	0.9660
		1.4416	1.8287	1.8318	1.7942	1.6777	1.4960
			1.9205	1.8920	1.8405	1.7096	1.5190
16			1.4495	1.4541	1.4253	1.3336	1.1897
	25	1.3038	1.4781	1.4725	1.4392	1.3432	1.1966
		1.4415	1.8287	1.8318	1.7942	1.6777	1.4960
			1.8648	1.8549	1.8118	1.6897	1.5046
			1.4297	1.4417	1.4159	1.3273	1.1852
	36	1.3038	1.4781	1.4725	1.4392	1.3432	1.1968
		1.5612	2.2604	2.2734	2.2302	2.0881	1.8632
			2.3367	2.3219	2.2669	2.1132	1.8811

Table 2.5.4 Bounds of $e_*(\bar{X}_\alpha, \bar{X})$ and $e^*(\bar{X}_\alpha, \bar{X})$ when $H(x) = \Phi(x)$

$\epsilon = 0.1$

s_1	s_2	α 0.01	0.05	0.1	0.15	0.25	0.375
1	4	0.9960 1.0501	0.9744 1.0825	0.9430 1.0702	0.9092 1.0436	0.7400 0.8367 0.9725 1.0994	0.6628 0.7397 0.8660 0.9663
	9	0.9960 1.1434	0.9744 1.3664	0.9430 1.3885	0.9092 1.3698	0.7059 0.8367 1.2904 1.5288	0.6376 0.7397 1.1556 1.3401
4	9	1.0502 1.1434	0.9718 1.0825 1.3664 1.5220	0.9918 1.0702 1.3885 1.4981	0.9812 1.0436 1.3698 1.4567	0.9278 0.9725 1.2904 1.3526	0.8331 0.8660 1.1556 1.2012
	16	1.0502 1.2089	1.0825 1.7852	0.9489 1.0702 1.8579 2.0944	0.9485 1.0436 1.8477 2.0324	0.9052 0.9725 1.7528 1.8830	0.8168 0.8660 1.5750 1.6698
9	16	1.1434 1.2091	1.2706 1.3664 1.7852 1.9190	1.3286 1.3885 1.8579 1.9414	1.3242 1.3698 1.8477 1.9112	1.2590 1.2904 1.7528 1.7965	1.1329 1.1556 1.5750 1.6064
	25	1.1434 1.2575	1.2025 1.3664 2.3226 2.6299	1.2912 1.3885 2.4666 2.6518	1.2963 1.3698 2.4673 2.6068	1.2402 1.2904 2.3516 2.4468	1.1194 1.1556 2.1176 2.1859
16	25	1.2092 1.2575	1.6935 1.7852 2.3225 2.4474	1.8056 1.8579 2.4666 2.5380	1.8088 1.8477 2.4673 2.5203	1.7266 1.7528 2.3516 2.3874	1.5562 1.5750 2.1176 2.1431
	36	1.2090 1.2927	1.6258 1.7852 2.9702 3.2580	1.7699 1.8579 3.2102 3.3693	1.7827 1.8477 3.2248 3.3423	1.7090 1.7528 3.0839 3.1679	1.5437 1.5750 2.7813 2.8375

Table 2.5.5 Bounds of $e_*(\bar{X}_\alpha, \bar{X})$ and $e^*(\bar{X}_\alpha, \bar{X})$ when $H(x) = \Phi(x)$

$\epsilon = 0.15$

s_1	s_2	α					
		0.01	0.05	0.1	0.15	0.25	0.375
1	4	0.9960 1.0512	0.9744 1.1120	0.9430 1.1121	0.9092 1.0911	0.6943 0.8367 1.0234 1.2329	0.6260 0.7397 0.9147 1.0804
	9	0.9960 1.1111	0.9744 1.4424	0.9430 1.5195	0.9092 1.5211	0.6450 0.8367 1.4529 1.8828	0.5894 0.7397 1.3106 1.6433
4	9	1.0512 1.1111	1.1120 1.4424	0.9846 1.1121 1.5195 1.7158	0.9899 1.0911 1.5211 1.6761	0.9509 1.0234 1.4529 1.5635	0.8614 0.9147 1.3106 1.3917
	16	1.0512 1.1477	1.1120 1.8856	0.9140 1.1121 2.1070 2.5607	0.9369 1.0911 2.1425 2.4934	0.9145 1.0234 2.0721 2.3180	0.8351 0.9147 1.8798 2.0587
9	16	1.1111 1.1478	1.2523 1.4424 1.8858 2.1697	1.4108 1.5195 2.1070 2.2687	1.4397 1.5211 2.1425 2.2633	1.3974 1.4529 2.0721 2.1543	1.2706 1.3106 1.8798 1.9388
	25	1.1111 1.1716	1.1036 1.4424 2.3912 3.1107	1.3416 1.5195 2.8574 3.2335	1.3898 1.5211 2.9414 3.2181	1.3641 1.4529 2.8696 3.0561	1.2469 1.3106 2.6132 2.7465
16	25	1.1478 1.1716	1.6633 1.8858 2.3913 2.7073	2.0041 2.1070 2.8574 3.0036	2.0684 2.1425 2.9414 3.0467	2.0227 2.0721 2.8696 2.9395	1.8447 1.8798 2.6132 2.6629
	36	1.1478 1.1872	1.4676 1.8858 2.9084 3.7173	1.9327 2.1070 3.7642 4.1015	2.0184 2.1425 3.9128 4.1527	1.9899 2.0721 3.8416 4.0001	1.8214 1.8798 3.5079 3.6201

Table 2.5.6 Bounds of $e_*(\bar{X}_\alpha, \bar{X})$ and $e^*(\bar{X}_\alpha, \bar{X})$ when $H(x) = \Phi(x)$

$\epsilon = 0.25$

		α					
s_1	s_2	0.01	0.05	0.1	0.15	0.25	0.375
1	4	0.9960 1.0407	0.9744 1.1332	0.9430 1.1588	0.9092 1.1511	0.6081 0.8367 1.0943 1.5044	0.5557 0.7397 0.9857 1.3103
	9	0.9960 1.0700	0.9744 1.4033	0.9430 1.6225	0.9092 1.6850	0.5323 0.8367 1.6631 2.6072	0.4988 0.7397 1.5251 2.2557
4	9	1.0407 1.0700	1.1332 1.4034	0.9193 1.1588 1.6225 2.0449	0.9612 1.1511 1.6850 2.0164	0.9581 1.0943 1.6631 1.8987	0.8852 0.9857 1.5251 1.6978
	16	1.0407 1.0852	1.1332 1.5980	0.7811 1.1588 2.2279 3.2977	0.8611 1.1511 2.4238 3.2330	0.8904 1.0943 2.4649 3.0256	0.8366 0.9857 2.2883 2.6947
9	16	1.0700 1.0852	1.4034 1.5980	1.3780 1.6225 2.2279 2.6204	1.5099 1.6850 2.4238 2.7035	1.5461 1.6631 2.4648 2.6510	1.4415 1.5251 2.2883 2.4209
	25	1.0700 1.0937	1.4034 1.7047	1.2110 1.6225 2.9326 3.9138	1.4010 1.6850 3.3458 4.0182	1.4762 1.6631 3.4837 3.9228	1.3922 1.5251 3.2623 3.5732
16	25	1.0852 1.0937	1.5980 1.7047	1.9596 2.2279 2.9326 3.3310	2.2496 2.4238 3.3458 3.6038	2.3537 2.4649 3.4837 3.6479	2.2101 2.2883 3.2623 3.3775
	36	1.0852 1.0989	1.5980 1.7712	1.7562 2.2279 3.7072 4.6879	2.1302 2.4238 4.4424 5.0506	2.2799 2.4649 4.7143 5.0956	2.1586 2.2883 4.4428 4.7094

Table 2.5.7 Bounds of $e_*(M, \bar{X}_\alpha)$ and $e^*(M, \bar{X}_\alpha)$ when $H(x) = \Phi(x)$

$\epsilon = 0.01$

s_1	s_2	α: 0.01	0.05	0.1	0.15	0.25	0.375
1	4	0.6392 0.6454	0.6533 0.6567	0.6751 0.6772	0.7002 0.7017	0.7525 0.7609 0.7616 0.7701	0.8517 0.8607 0.8608 0.8699
	9	0.6392 0.6517	0.6533 0.6595	0.6751 0.6789	0.7002 0.7028	0.7496 0.7609 0.7620 0.7735	0.8487 0.8607 0.8610 0.8731
4	9	0.6454 0.6517	0.6536 0.6567 0.6595 0.6627	0.6743 0.6772 0.6789 0.6819	0.6988 0.7017 0.7028 0.7057	0.7587 0.7616 0.7620 0.7649	0.8578 0.8608 0.8610 0.8640
	16	0.6454 0.6560	0.6518 0.6567 0.6612 0.6662	0.6727 0.6772 0.6798 0.6844	0.6974 0.7017 0.7034 0.7078	0.7573 0.7616 0.7623 0.7667	0.8563 0.8608 0.8610 0.8656
9	16	0.6494 0.6517 0.6560 0.6583	0.6578 0.6595 0.6612 0.6629	0.6773 0.6789 0.6798 0.6814	0.7013 0.7028 0.7034 0.7049	0.7606 0.7620 0.7623 0.7638	0.8594 0.8610 0.8610 0.8625
	25	0.6478 0.6517 0.6590 0.6630	0.6567 0.6595 0.6623 0.6651	0.6764 0.6789 0.6804 0.6830	0.7004 0.7028 0.7038 0.7062	0.7597 0.7620 0.7625 0.7648	0.8585 0.8610 0.8611 0.8635
16	25	0.6544 0.6560 0.6590 0.6606	0.6601 0.6612 0.6623 0.6634	0.6789 0.6798 0.6804 0.6814	0.7025 0.7034 0.7038 0.7047	0.7614 0.7623 0.7625 0.7634	0.8601 0.8610 0.8611 0.8620
	36	0.6532 0.6560 0.6611 0.6640	0.6594 0.6612 0.6630 0.6649	0.6783 0.6798 0.6809 0.6825	0.7019 0.7034 0.7040 0.7055	0.7609 0.7623 0.7626 0.7640	0.8595 0.8610 0.8611 0.8626

Table 2.5.8 Bounds of $e_*(M, \bar{X}_\alpha)$ and $e^*(M, \bar{X}_\alpha)$ when $H(x) = \Phi(x)$

$\epsilon = 0.05$

		α					
s_1	s_2	0.01	0.05	0.1	0.15	0.25	0.375
1	4	0.6392 0.6726	0.6533 0.6709	0.6751 0.6862	0.7002 0.7078	0.7194 0.7609 0.7644 0.8084	0.8162 0.8607 0.8616 0.9085
	9	0.6392 0.7232	0.6533 0.6873	0.6751 0.6953	0.7002 0.7136	0.7057 0.7609 0.7669 0.8270	0.8016 0.8607 0.8622 0.9258
4	9	0.6726 0.7231	0.6545 0.6709 0.6873 0.7046	0.6709 0.6862 0.6953 0.7112	0.6930 0.7078 0.7136 0.7289	0.7497 0.7644 0.7669 0.7819	0.8462 0.8616 0.8622 0.8779
	16	0.6726 0.7917	0.6456 0.6709 0.6982 0.7258	0.6631 0.6862 0.7009 0.7255	0.6855 0.7078 0.7171 0.7404	0.7425 0.7644 0.7684 0.7911	0.8385 0.8616 0.8626 0.8863
9	16	0.7231 0.7917	0.6779 0.6873 0.6982 0.7080	0.6872 0.6953 0.7009 0.7093	0.7059 0.7136 0.7171 0.7249	0.7595 0.7669 0.7684 0.7760	0.8544 0.8622 0.8626 0.8704
	25	0.7231 0.8954	0.6721 0.6873 0.7058 0.7220	0.6822 0.6953 0.7047 0.7182	0.7013 0.7136 0.7194 0.7320	0.7550 0.7669 0.7693 0.7815	0.8497 0.8622 0.8628 0.8754
16	25	0.7916 0.8955	0.6922 0.6982 0.7058 0.7120	0.6959 0.7009 0.7047 0.7098	0.7124 0.7171 0.7194 0.7241	0.7639 0.7684 0.7693 0.7739	0.8579 0.8626 0.8628 0.8675
	36	0.7916 1.0299	0.6881 0.6982 0.7113 0.7219	0.6925 0.7009 0.7073 0.7159	0.7093 0.7171 0.7209 0.7289	0.7609 0.7684 0.7700 0.7776	0.8547 0.8626 0.8629 0.8708

Table 2.5.9 Bounds of $e_*(M, \bar{X}_\alpha)$ and $e^*(M, \bar{X}_\alpha)$ when $H(x) = \Phi(x)$

$\epsilon = 0.1$

s_1	s_2	α					
		0.01	0.05	0.1	0.15	0.25	0.375
1	4	0.6392 0.7113	0.6533 0.6900	0.6751 0.6979	0.7002 0.7157	0.6794 0.7609 0.7680 0.8603	0.7730 0.8607 0.8625 0.9605
	9	0.6392 0.8731	0.6533 0.7305	0.6751 0.7189	0.7002 0.7288	0.6529 0.7609 0.7736 0.9019	0.7449 0.8607 0.8638 0.9985
4	9	0.7112 0.8731	0.6559 0.6900 0.7306 0.7686	0.6663 0.6979 0.7189 0.7531	0.6852 0.7157 0.7288 0.7612	0.7380 0.7680 0.7736 0.8051	0.8310 0.8625 0.8638 0.8966
	16	0.7112 1.1264	0.6900 0.7628	0.6502 0.6979 0.7329 0.7871	0.6700 0.7157 0.7370 0.7874	0.7232 0.7680 0.7769 0.8252	0.8155 0.8625 0.8646 0.9145
9	16	0.8730 1.1263	0.7096 0.7305 0.7628 0.7856	0.7014 0.7189 0.7329 0.7513	0.7125 0.7288 0.7370 0.7538	0.7580 0.7736 0.7769 0.7928	0.8477 0.8638 0.8646 0.8811
	25	0.8730 1.4568	0.6966 0.7305 0.7888 0.8284	0.6909 0.7189 0.7427 0.7731	0.7028 0.7288 0.7425 0.7701	0.7487 0.7736 0.7791 0.8049	0.8381 0.8638 0.8652 0.8917
16	25	1.1264 1.4569	0.7486 0.7628 0.7888 0.8041	0.7219 0.7329 0.7427 0.7542	0.7269 0.7370 0.7425 0.7528	0.7674 0.7769 0.7791 0.7887	0.8548 0.8646 0.8652 0.8751
	36	1.1264 1.8622	0.7389 0.7628 0.8105 0.8376	0.7145 0.7329 0.7499 0.7694	0.7202 0.7370 0.7465 0.7639	0.7611 0.7769 0.7806 0.7968	0.8484 0.8646 0.8655 0.8821

Table 2.5.10 Bounds of $e_*(M, \bar{X}_\alpha)$ and $e^*(M, \bar{X}_\alpha)$ when $H(x) = \Phi(x)$

$\epsilon = 0.15$

s_1	s_2	α = 0.01	0.05	0.1	0.15	0.25	0.375
1	4	0.6392 0.7513	0.6533 0.7103	0.6751 0.7102	0.7002 0.7239	0.6406 0.7609 0.7717 0.9170	0.7311 0.8607 0.8634 1.0170
	9	0.6392 1.0210	0.6533 0.7865	0.6751 0.7466	0.7002 0.7458	0.6025 0.7609 0.7808 0.9870	0.6903 0.8607 0.8656 1.0802
4	9	0.7514 1.0210	0.7103 0.7865	0.6612 0.7102 0.7466 0.8022	0.6769 0.7239 0.7458 0.7979	0.7256 0.7717 0.7808 0.8306	0.8152 0.8634 0.8656 0.9169
	16	0.7514 1.4199	0.7103 0.8643	0.6364 0.7102 0.7735 0.8641	0.6536 0.7239 0.7606 0.8431	0.7030 0.7717 0.7865 0.8637	0.7916 0.8634 0.8670 0.9458
9	16	1.0210 1.4198	0.7511 0.7865 0.8642 0.9059	0.7183 0.7466 0.7735 0.8041	0.7200 0.7458 0.7606 0.7880	0.7565 0.7808 0.7865 0.8118	0.8406 0.8656 0.8670 0.8928
	25	1.0210 1.9357	0.7290 0.7865 0.9484 1.0280	0.7013 0.7466 0.7936 0.8456	0.7047 0.7458 0.7710 0.8163	0.7421 0.7808 0.7903 0.8317	0.8257 0.8656 0.8678 0.9098
16	25	1.4198 1.9357	0.8377 0.8642 0.9483 0.9798	0.7550 0.7735 0.7936 0.8132	0.7444 0.7606 0.7710 0.7879	0.7715 0.7865 0.7903 0.8057	0.8516 0.8670 0.8678 0.8834
	36	1.4198 2.5657	0.8195 0.8642 1.0474 1.1105	0.7427 0.7735 0.8093 0.8432	0.7336 0.7606 0.7786 0.8074	0.7616 0.7865 0.7930 0.8190	0.8415 0.8670 0.8684 0.8947

Table 2.5.11 Bounds of $e_*(M, \bar{X}_\alpha)$ and $e^*(M, \bar{X}_\alpha)$ when $H(x) = \Phi(x)$

$\epsilon = 0.25$

s_1	s_2	α = 0.01	0.05	0.1	0.15	0.25	0.375
1	4					0.5670	0.6510
		0.6392	0.6533	0.6751	0.7002	0.7609	0.8607
		0.8196	0.7527	0.7361	0.7410	0.7795	0.8654
						1.0470	1.1457
	9					0.5087	0.5880
		0.6392	0.6533	0.6751	0.7002	0.7609	0.8607
		1.2395	0.9451	0.8174	0.7871	0.7975	0.8696
						1.1960	1.2763
4	9			0.6486	0.6577	0.6985	0.7812
		0.8196	0.7527	0.7361	0.7410	0.7795	0.8654
		1.2395	0.9451	0.8174	0.7871	0.7975	0.8696
				0.9278	0.8874	0.8903	0.9636
	16			0.6053	0.6175	0.6598	0.7408
		0.8196	0.7527	0.7361	0.7410	0.7795	0.8654
		1.8396	1.2492	0.8960	0.8236	0.8099	0.8724
				1.0921	0.9906	0.9579	1.0196
9	16			0.7618	0.7384	0.7530	0.8246
		1.2396	0.9451	0.8174	0.7871	0.7975	0.8696
		1.8396	1.2492	0.8960	0.8236	0.8099	0.8724
				0.9625	0.8784	0.8578	0.9201
	25			0.7287	0.7098	0.7270	0.7982
		1.2396	0.9451	0.8174	0.7871	0.7975	0.8696
		2.6077	1.6730	0.9725	0.8524	0.8187	0.8743
				1.0952	0.9467	0.8984	0.9527
16	25			0.8562	0.7914	0.7818	0.8444
		1.8396	1.2492	0.8960	0.8236	0.8099	0.8724
		2.0677	1.6730	0.9725	0.8524	0.8187	0.8743
				1.0187	0.8874	0.8481	0.9032
	36			0.8298	0.7702	0.7634	0.8260
		1.8396	1.2492	0.8960	0.8236	0.8099	0.8724
		3.5400	2.1964	1.0493	0.8757	0.8252	0.8756
				1.1367	0.9371	0.8756	0.9248

PART II

ROBUST EXPERIMENTAL DESIGN

Chapter 3

Introduction to Robust Experimental Design

3.1 Some Basic Concepts of Optimal Design

Before we study the problem of robust experimental design, we would like to present some basic concepts of optimal design as our starting point. These concepts will be used frequently in the later chapters.

Let us consider the following regression model:

$$E(y \mid \underset{\sim}{x}_i) = \underset{\sim}{\theta}^T \underset{\sim}{f}(\underset{\sim}{x}_i), \qquad i = 1, ..., n, \tag{3.1.1}$$

or alternatively

$$y_i = y(\underset{\sim}{x}_i) = \underset{\sim}{\theta}^T \underset{\sim}{f}(\underset{\sim}{x}_i) + \epsilon_i, \qquad i = 1, ..., n, \tag{3.1.2}$$

where $\underset{\sim}{\theta}^T = (\theta_0, ..., \theta_p)$, $\underset{\sim}{f}^T(\underset{\sim}{x}) = (f_0(\underset{\sim}{x}), ..., f_p(\underset{\sim}{x}))$ and $\underset{\sim}{x}_i \in S \subseteq \mathbb{R}^q$, $i = 1, ..., n$. S denotes the design space we are interested in. Particularly, we may choose $S = \{(x_1, ..., x_q) : -1 \leq x_j \leq 1,\ j = 1, ..., q\}$. We assume that $\underset{\sim}{x}_i'$s are subject to no error and ϵ_i's are independent and identically distributed with mean 0 and variance $\sigma^2 > 0$.

Furthermore, we define

$$F(\underset{\sim}{x}) = \begin{pmatrix} f_0(\underset{\sim 1}{x}) & f_1(\underset{\sim 1}{x}) & \dots & f_p(\underset{\sim 1}{x}) \\ f_0(\underset{\sim 2}{x}) & f_1(\underset{\sim 2}{x}) & \dots & f_p(\underset{\sim 2}{x}) \\ \dots & \dots & \dots & \dots \\ f_0(\underset{\sim n}{x}) & f_1(\underset{\sim n}{x}) & \dots & f_p(\underset{\sim n}{x}) \end{pmatrix},$$

$B(\underset{\sim}{x}) = \frac{1}{n}F^T(\underset{\sim}{x})F(\underset{\sim}{x}) = \frac{1}{n}\sum_{i=1}^{n} \underset{\sim}{f}(\underset{\sim i}{x})\cdot \underset{\sim}{f}^T(\underset{\sim i}{x})$, $\underset{\sim}{y}^T = (y_1, ..., y_n)$, and $\underset{\sim}{\epsilon}^T = (\epsilon_1, ..., \epsilon_n)$. Then, we can write (3.1.2) as the following form:

$$\underset{\sim}{y}(\underset{\sim}{x}) = F(\underset{\sim}{x})\,\underset{\sim}{\theta} + \underset{\sim}{\epsilon}\ . \tag{3.1.3}$$

It is well known that the least squares estimator of $\underset{\sim}{\theta}$ under (3.1.3) is

$$\hat{\underset{\sim}{\theta}} = (F^T F)^{-1} F^T \underset{\sim}{y},$$

and the covariance matrix of $\hat{\underset{\sim}{\theta}}$ is

$$\text{cov}(\hat{\underset{\sim}{\theta}}) = E[(\hat{\underset{\sim}{\theta}} - \underset{\sim}{\theta})(\hat{\underset{\sim}{\theta}} - \underset{\sim}{\theta})^T] = \frac{\sigma^2}{n}B^{-1}(\underset{\sim}{x}).$$

We confine ourself to the use of the least squares estimator $\hat{\underset{\sim}{\theta}}$. It is clear that the covariance matrix $\text{cov}(\hat{\underset{\sim}{\theta}})$ depends on the observations $\underset{\sim i}{x}$, $i = 1, ..., n$. The design problem which we shall be concerned with is the following: How should the values $\underset{\sim 1}{x}, ..., \underset{\sim n}{x}$ of the independent variable be chosen in order to give the "best" experiment? The question of best design depends on the meaning of "best". Many optimality criteria have been posed and studied in the past. See for example, Kiefer (1959), Box and Draper (1959, 1963) and also Fedorov (1972). Before we present some of the optimal criteria commonly used, we first give a precise definition of an experimental design.

Definition 3.1.1 *A design of an experiment is the collection of quantities*

$$\begin{pmatrix} \underset{\sim 1}{x}, & \underset{\sim 2}{x}, \dots, & \underset{\sim n}{x} \\ n_1, & n_2, \dots, & n_n \end{pmatrix} \tag{3.1.4}$$

where $\underset{\sim i}{x} \in S$*, and* n_i *is the numbers of repetition at point* $\underset{\sim i}{x}$, $i = 1, ..., n$.

For the theoretical study, the concepts of normalized design and continuous normalized design are more useful.

Definition 3.1.2 *A normalized design is the collection of quantities*

$$\begin{pmatrix} \underset{\sim}{x}_1, & \underset{\sim}{x}_2, \ldots, & \underset{\sim}{x}_n \\ p_1, & p_2, \ldots, & p_n \end{pmatrix} \tag{3.1.5}$$

where $p_i = n_i / \sum_{j=1}^{n} n_j$, $i = 1, \ldots, n$. *Moreover, we call a design to be a continuous normalized experiment, if in (3.1.5) we allow* p_i *to be any real number between* 0 *and* 1.

The concept of experimental design can be further extended by allowing design measure to be any probability measure $\xi(\underset{\sim}{x})$ supported on some design space S. With this extension, we define

$$B(\xi) = \int_S \underset{\sim}{f}(\underset{\sim}{x}) \cdot \underset{\sim}{f}^T(\underset{\sim}{x}) d\xi(\underset{\sim}{x}).$$

Then we have

$$\hat{\underset{\sim}{\theta}} = B^{-1}(\xi) \cdot \int_S \underset{\sim}{f}(\underset{\sim}{x}) \cdot y(\underset{\sim}{x}) d\xi(\underset{\sim}{x}), \tag{3.1.6}$$

and the least squares estimator of $\underset{\sim}{\theta}^T \underset{\sim}{f}(\underset{\sim}{x})$ is $\hat{y} = \hat{\underset{\sim}{\theta}}^T \underset{\sim}{f}(\underset{\sim}{x})$. We use $M(\xi)$ to denote the mean squared error matrix of $\hat{\underset{\sim}{\theta}}$ as estimator of $\underset{\sim}{\theta}$ which depends on the design measure $\xi(\underset{\sim}{x})$. Under (3.1.3), we know that $\hat{\underset{\sim}{\theta}}$ is an unbiased estimator of $\underset{\sim}{\theta}$. Hence $M(\xi)$ is simply the covariance matrix of $\hat{\underset{\sim}{\theta}}$, i.e.,

$$M(\xi) = E[(\hat{\underset{\sim}{\theta}} - \underset{\sim}{\theta})(\hat{\underset{\sim}{\theta}} - \underset{\sim}{\theta})^T] = \mathrm{cov}(\hat{\underset{\sim}{\theta}}) = \frac{\sigma^2}{n} B^{-1}(\xi).$$

Also, we denote

$$d(\underset{\sim}{x}, \xi) = \underset{\sim}{f}^T(\underset{\sim}{x}) M(\xi) \underset{\sim}{f}(\underset{\sim}{x})$$

which is the mean squared error of $\hat{\underset{\sim}{\theta}}^T \underset{\sim}{f}(\underset{\sim}{x})$ as estimator of $\underset{\sim}{\theta}^T \underset{\sim}{f}(\underset{\sim}{x})$.

The goal of optimal design is to seek design measures such that some loss functions of $M(\xi)$ or $d(\underset{\sim}{x}, \xi)$ is minimized. Here are some of the loss functions used by many authors: (i) $\mathcal{L}_D(\xi) = \mid M(\xi) \mid$, (ii) $\mathcal{L}_A(\xi) = tr\ M(\xi)$, (iii) $\mathcal{L}_Q(\xi) = \int_S d(\underset{\sim}{x}, \xi) d\underset{\sim}{x}$, and (iv) $\mathcal{L}_G(\xi) = \max_{\underset{\sim}{x} \in S} d(\underset{\sim}{x}, \xi)$. The minimization of these loss functions over some class of design measures yields different kinds of optimality.

Definition 3.1.3 *Let $\mathcal{F}$ to be a set of design measures we are interested in. We call design measure $\xi_0 \in \mathcal{F}$ to be D−, A−, Q−, or G−optimal over $\mathcal{F}$, if*

$$\begin{aligned} \mid M(\xi_0) \mid &= \min_{\xi\in\mathcal{F}} \mid M(\xi) \mid, \\ trM(\xi_0) &= \min_{\xi\in\mathcal{F}} trM(\xi), \\ \int_S d(\underset{\sim}{x}, \xi_0) d\,\underset{\sim}{x} &= \min_{\xi\in\mathcal{F}} \int_S d(\underset{\sim}{x}, \xi) d\,\underset{\sim}{x}, \end{aligned}$$

or

$$\max_{\underset{\sim}{x}\in S} d(\underset{\sim}{x}, \xi_0) = \min_{\xi\in\mathcal{F}} \max_{\underset{\sim}{x}\in S} d(\underset{\sim}{x}, \xi)$$

respectively.

It is clear that the D-optimal design minimizes the generalized variance of the least squares estimator $\underset{\sim}{\hat{\theta}}$, while A-optimal design minimizes the mean of the normalized dispersion $\frac{1}{n} \cdot trM(\xi)$. These two designs are the optimal designs in the space of parameters. On the other hand, the Q and $G-$optimal designs are the optimal designs in the space of control variables. The $Q-$optimal design minimizes the average of $d(\underset{\sim}{x}, \xi)$ over the design space S, while the $G-$optimal design minimizes the maximum value of $d(\underset{\sim}{x}, \xi)$ over S. Hence, $G-$optimal design is also known as minimax design in the space of control variables.

The problem of optimal design has been studied extensively by many authors, especially by Kiefer. The famous theorem about the equivalence of $D-$optimal and minimax designs is also due to Kiefer and Wolfowitz (1960). Many topics about optimal design theory can also be found in Fedorov (1972). We are not going to discuss the usual theory of optimal design in depth. Instead, we confine ourself to the robust considerations of experimental design which we are going to discuss in the next section and to study some different aspects in the next three chapters.

3.2 Historical Review of Robust Experimental Design

There is a major consideration in robust experimental design problems, namely the possible violation of the assumed regression model.

Consider the regression model (3.1.2). The regression problem is to make inference about $\underset{\sim}{\theta}$ in some "optimal" way. In particular, an optimal estimator of $\underset{\sim}{\theta}$ has to be chosen and in connection with this estimator the design problem is to choose the experimental points, $\underset{\sim i}{x}$'s in an optimal manner. When we choose the least squares method of estimation, then a variety of optimality criteria could be considered in the associated design problem as we have discussed in Section 3.1. Unfortunately, as was noticed by Box and Draper (1959), the strict formulation of the regression function becomes dangerous in the situations when the "true" regression function $y(\underset{\sim}{x})$ is only approximated by $\underset{\sim}{\theta}^T \underset{\sim}{f}\,(\underset{\sim}{x})$ thereby introducing a bias term which may be considerable. The corresponding model can be given now by

$$y_i = y(\underset{\sim i}{x}) = \underset{\sim}{\theta}^T \underset{\sim}{f}\,(\underset{\sim i}{x}) + \psi(\underset{\sim i}{x}) + \epsilon_i, \qquad i = 1, ..., n, \tag{3.2.1}$$

where $\psi(\underset{\sim}{x})$ is an unknown "contamination function" defined on S. $\psi(\underset{\sim}{x})$ belongs to some set Ψ with some specified properties.

Let $\hat{\underset{\sim}{\theta}}$ be the least squares estimator of $\underset{\sim}{\theta}$ as we defined in (3.1.6), and $\hat{y} = \hat{\underset{\sim}{\theta}}^T \underset{\sim}{f}\,(\underset{\sim}{x})$ be the least squares estimator of $\underset{\sim}{\theta}^T \underset{\sim}{f}\,(\underset{\sim}{x})$. We know that $\hat{\underset{\sim}{\theta}}$ is no longer an unbiased estimator of $\underset{\sim}{\theta}$. In fact, we have the following:

Lemma 3.2.1 *Under the regression model (3.2.1), we have*

$$(i)\quad E[\hat{\underset{\sim}{\theta}} - \underset{\sim}{\theta}] = B^{-1}(\xi)\, \underset{\sim}{b}\,(\psi, \xi) \tag{3.2.2}$$

$$\begin{aligned} (ii)\quad M(\psi, \xi) &= E[(\hat{\underset{\sim}{\theta}} - \underset{\sim}{\theta})][(\hat{\underset{\sim}{\theta}} - \underset{\sim}{\theta})^T] \\ &= \frac{\sigma^2}{n} B^{-1}(\xi) + B^{-1}(\xi)\, \underset{\sim}{b}\,(\psi, \xi)\, \underset{\sim}{b}^T\,(\psi, \xi) B^{-1}(\xi) \end{aligned} \tag{3.2.3}$$

$$\text{(iii)}\quad MSE(\hat{y}) = \frac{\sigma^2}{n}\cdot \underset{\sim}{f}^T(\underset{\sim}{x})B^{-1}(\xi)\,\underset{\sim}{f}(\underset{\sim}{x}) + \underset{\sim}{f}^T(\underset{\sim}{x})B^{-1}(\xi)\,\underset{\sim}{b}(\psi,\xi)\,\underset{\sim}{b}^T(\psi,\xi)B^{-1}(\xi)\,\underset{\sim}{f}(\underset{\sim}{x}) \tag{3.2.4}$$

where $B(\xi) = \int_S \underset{\sim}{f}(\underset{\sim}{x})\,\underset{\sim}{f}^T(\underset{\sim}{x})d\xi(\underset{\sim}{x})$ *and* $\underset{\sim}{b}(\psi,\xi) = \int_S \underset{\sim}{f}(\underset{\sim}{x})\psi(\underset{\sim}{x})d\xi(\underset{\sim}{x})$ and ξ is the design measure on S.

Proof:

(i) By (3.1.6) and (3.2.1), we have

$$\begin{aligned}
\hat{\underset{\sim}{\theta}} &= B^{-1}(\xi)\int_S \underset{\sim}{f}(\underset{\sim}{x})y(\underset{\sim}{x})d\xi(\underset{\sim}{x}) \\
&= B^{-1}(\xi)\int_S \underset{\sim}{f}(\underset{\sim}{x})(\underset{\sim}{f}^T(\underset{\sim}{x})\,\underset{\sim}{\theta} + \psi(\underset{\sim}{x}) + \epsilon)d\xi(\underset{\sim}{x}) \\
&= B^{-1}(\xi)(\underset{\sim}{\theta}\int_S \underset{\sim}{f}(\underset{\sim}{x})\,\underset{\sim}{f}^T(\underset{\sim}{x})d\xi(\underset{\sim}{x}) + \int_S \underset{\sim}{f}(\underset{\sim}{x})\psi(\underset{\sim}{x})d\xi(\underset{\sim}{x}) + \epsilon\int_S \underset{\sim}{f}(\underset{\sim}{x})d\xi(\underset{\sim}{x})) \\
&= \underset{\sim}{\theta} + B^{-1}(\xi)\,\underset{\sim}{b}(\psi,\xi) + \epsilon\, B^{-1}(\xi)\int_S \underset{\sim}{f}(\underset{\sim}{x})d\xi(\underset{\sim}{x}).
\end{aligned}$$

Hence, we have $E(\hat{\underset{\sim}{\theta}}) = \underset{\sim}{\theta} + B^{-1}(\xi)\,\underset{\sim}{b}(\psi,\xi)$, i.e., $E(\hat{\underset{\sim}{\theta}} - \underset{\sim}{\theta}) = B^{-1}(\xi)\,\underset{\sim}{b}(\psi,\xi)$.

(ii) A direct calculation yields the following:

$$\begin{aligned}
M(\psi,\xi) &= E[(\hat{\underset{\sim}{\theta}} - \underset{\sim}{\theta})(\hat{\underset{\sim}{\theta}} - \underset{\sim}{\theta})^T] \\
&= E[(\hat{\underset{\sim}{\theta}} - E(\hat{\underset{\sim}{\theta}}) + E(\hat{\underset{\sim}{\theta}}) - \underset{\sim}{\theta})(\hat{\underset{\sim}{\theta}} - E(\hat{\underset{\sim}{\theta}}) + E(\hat{\underset{\sim}{\theta}}) - \underset{\sim}{\theta})^T] \\
&= E[(\hat{\underset{\sim}{\theta}} - E(\hat{\underset{\sim}{\theta}}))(\hat{\underset{\sim}{\theta}} - E(\hat{\underset{\sim}{\theta}}))^T] + E[(\hat{\underset{\sim}{\theta}} - E(\hat{\underset{\sim}{\theta}}))(E(\hat{\underset{\sim}{\theta}}) - \underset{\sim}{\theta})^T] \\
&\qquad + E[(E(\hat{\underset{\sim}{\theta}}) - \underset{\sim}{\theta})(\hat{\underset{\sim}{\theta}} - E(\hat{\underset{\sim}{\theta}}))^T] + E[(E(\hat{\underset{\sim}{\theta}}) - \underset{\sim}{\theta})(E(\hat{\underset{\sim}{\theta}}) - \underset{\sim}{\theta})^T] \\
&= \text{cov}(\hat{\underset{\sim}{\theta}}) + (E(\hat{\underset{\sim}{\theta}} - \underset{\sim}{\theta}))\cdot(E(\hat{\underset{\sim}{\theta}} - \underset{\sim}{\theta}))^T \\
&= \frac{\sigma^2}{n}B^{-1}(\xi) + (B^{-1}(\xi)\,\underset{\sim}{b}(\psi,\xi))\cdot(B^{-1}(\xi)\,\underset{\sim}{b}(\psi,\xi))^T \\
&= \frac{\sigma^2}{n}B^{-1}(\xi) + B^{-1}(\xi)\,\underset{\sim}{b}(\psi,\xi)\,\underset{\sim}{b}^T(\psi,\xi)B^{-1}(\xi).
\end{aligned}$$

Note that the last second equality followed by (i) and the last equality by the fact that $B^{-1}(\xi)$ is a symmetric matrix.

(iii) We know that the mean squared error of $\hat{y}$, $MSE(\hat{y})$, is $d(\underset{\sim}{x},\psi,\xi) = \underset{\sim}{f}^T(\underset{\sim}{x})M(\psi,\xi)\,\underset{\sim}{f}(\underset{\sim}{x})$. Hence (iii) is followed by (ii) immediately. □

It is clear that the usual optimal designs are no longer optimal under the model (3.2.1) since they minimize some scale valued functions of $B^{-1}(\xi)$ only. There are some disadvantages of the usual optimal designs. They are dependent on the assumed model and very sensitive to the possible model violation. They provide no opportunity for a check of the model's adequacy etc. One attempt to meet these objections has been made by Box and Draper (1959). They use $\hat{y}(x) = \hat{\underset{\sim}{\theta}}^T \underset{\sim}{f}(x)$ to be the least squares estimator of the true regression function $y(x)$, calculated under the assumption that $\psi(x) = 0$, while in fact the real model is (3.2.1). They considered the case when $\underset{\sim}{f}^T(x) = (1, x, ..., x^m)$ and suggested that the design should minimize

$$\int_{-1}^{1} E[\hat{\underset{\sim}{\theta}}^T \underset{\sim}{f}(x) - y(x)]^2 dx = \int_{-1}^{1} \mathrm{var}[\hat{y}]dx + \int_{-1}^{1} (E[\hat{y}(x)] - y(x))^2 dx := V + B$$

within the class of symmetric design measures supported on $[-1, 1]$, where they referred to V as variance error and B as bias error. The major difficulty with adopting the criterion "minimize $V + B$" is that the optimal design depends on the function $y(x)$, which is unknown. Even if it is assumed that $y(x)$ is a polynomial of degree $m + 1$, the optimal design still cannot be found, as it will depend on the unknown coefficient of x^{m+1}. To avoid this difficulty, Box and Draper (1959) recommended that one choose the design to minimize B alone. As they noted that "The somewhat unexpected conclusion is reached that, at least in the case considered, the optimal design in typical situations in which both variance and bias occur is very nearly the same as would be obtained if variance were ignored completely and the experiment designed so as to minimize bias alone."

Beginning with Box and Draper (1959), the problem of finding robust design against the model violation has been further studied by many authors in different aspects. Designs for versions of (3.2.1) have been constructed in a series of papers. These differ in the class of Ψ, the design space, the regressors, and in the loss functions used.

Note that the least squares estimator, which disregards the presence of $\psi(x)$, may no longer be optimal among linear estimators for $\underset{\sim}{\theta}$, and therefore the search for new estimators is of special interest. Marcus and Sacks (1977) considered the one-dimensional regression model

$$y(x_i) = a + bx_i + \psi(x_i), \qquad i = 1, ..., n,$$

where $|\ \psi(x)\ | \leq \varphi(x)$, $\varphi(0) = 0$, with φ a given function. They restricted the estimators to be linear, but not necessarily the standard least squares estimator, and look for designs and estimators to minimize the mean squared error

$$\sup_{\Psi} E[(\hat{a} - a)^2 + \lambda^2(\hat{b} - b)^2],$$

where $\hat{a}, \hat{b}$ denote the estimators and λ is specified.

Pesotchinsky (1982), posed the similar problem for the multiple linear regression model

$$y(\underset{\sim}{x}_i) = \underset{\sim}{\theta}^T \underset{\sim}{f}\,(\underset{\sim}{x}_i) + \psi(\underset{\sim}{x}_i) + \epsilon_i, \qquad i = 1, ..., n.$$

He confined himself to the use of the standard least squares estimator because in the case of small deviations, the performance of least squares estimator is nearly the same as of the best linear estimator as was shown by Marcus and Sacks (1977).

Li (1984) also considered the similar problem when $\Psi = \{\psi(x) : |\ \psi(x)\ | \leq \varphi(x)\}$ where $\varphi(x)$ is a known function. But for the class of design measures, he focused on the case that $\sigma(\xi) = \left\{\frac{k}{2N}, -\frac{k}{2N} : k = 1, ..., N\right\}$ for a fixed natural number N.

In a related direction, Huber (1975) formulated a problem that

$$y(x_i) = a + b\, x_i + \psi(x_i) + \epsilon_i, \qquad i = 1, ..., n,$$

where $\psi(x) \in \Psi = \left\{\psi(x) : \int_{-\frac{1}{2}}^{\frac{1}{2}} \psi^2(x)dx \leq \eta^2\right\}$. Huber also confined himself to the use of the standard least squares estimator based on the above model with $\psi(x) = 0$, and found the design which minimizes the loss

$$\sup_{\Psi} E \int_{-\frac{1}{2}}^{\frac{1}{2}} \left(\hat{a} + \hat{b}x - y(x)\right)^2 dx.$$

An unfortunate consequence of this formulation is that it leads to the restriction that the designs must be absolutely continuous, otherwise the above loss is infinite.

This means that no implementable designs can have finite loss. However, it is to be understood that the continuous designs will be approximated by discrete designs in practice. Wiens (1992) mentioned the comment that "Our attitude is that an approximation to a design which is robust against more realistic alternatives is preferable to an exact solution in a neighbourhood which is unrealistically sparse." Wiens (1990 and 1992) extended the Huber's (1975) result to the multiple linear regression model and to some other loss functions.

Stigler (1971) suggested the so-called C-restricted D-optimal design for the one dimensional polynomial regression model

$$P_m : \ y(x_i) = \sum_{i=0}^{m} \theta_i x^i + \epsilon_i, \qquad i = 1, ..., n.$$

Instead of minimizing $|B_m^{-1}(\xi)|$ over the class of whole design measures supported on $[-1, 1]$, which is the usual $D-$optimal design, he suggested the design measure minimizing $|B_m^{-1}(\xi)|$ among all designs ξ satisfying $|B_m(\xi)\ |\leq c\,|B_{m+1}(\xi)|$. The justification for this choice is based on the fact that if $\hat{\theta}_{m+1}$ is the least squares estimator of θ_{m+1} for the model $P_{m+1} : y(x_i) = \sum_{i=0}^{m+1} \theta_i x^i + \epsilon_i$, and ξ corresponds to an experiment run at $x_1, .., x_n$, then $n{\cdot}\mathrm{var}(\hat{\theta}_{m+1}) = |B_m(\xi)| \cdot |B_{m+1}(\xi)|^{-1} \cdot \sigma^2$. Thus this criterion says "minimize the generalized variance of the least squares estimators $(\hat{\theta}_0, ..., \hat{\theta}_m)$ for the model P_m subject to the constraint that $\mathrm{var}(\hat{\theta}_{m+1}) \leq c \cdot \dfrac{\sigma^2}{n}$. Similarly, he introduced the $C-$restricted $G-$optimal design for the model P_m : that is, the design ξ_0 which minimizes $\max\limits_{-1\leq x\leq 1} d_m(x, \xi)$ among all designs ξ satisfying $|B_m(\xi)| \leq c\,|B_{m+1}(\xi)|$, where $d_m(x, \xi) = \underset{\sim}{f}^T(x) M_m(\xi) \underset{\sim}{f}(x)$ and $\underset{\sim}{f}^T(x) = (1, x, ..., x^m)$.

The choice of C reflects a compromise between two conflicting goals: precise inferences about θ_{m+1} and precise inferences about the model P_m. On the one hand, C should be chosen sufficiently small so that it will be possible to detect practically significant departures from the model with a specified precision (this requirement could be phrased in terms of the power of the test: $H : \theta_{m+1} = 0$); while on the other hand, large values of C will yield more efficient designs for the model P_m. Stigler (1971) found the C-restricted $D-$ and $G-$optimal designs for the model P_1, i.e., the one dimentional linear regression model.

Some previous papers, for example Marcus and Sacks (1977), Pesotchinsky (1982), Li (1984), Huber (1975), Wiens (1990 and 1992), have assumed that a class of possible bias functions exists but that all functions in the class are equally likely to be the actual bias presented in the model. In some cases it would seem that certain bias functions would be more likely than the others, and perhaps the experimenter can specify a prior probability distribution on the form of the possible bias in the model. Hence, Notz (1989) attempted to take into account prior information about the possible bias and suggested the following model:

$$y(x_i, w) = \underset{\sim}{\theta}^T \underset{\sim}{f}(x_i) + \psi(x_i, w) + \epsilon_i(w), \qquad i = 1, ..., n, \tag{3.2.5}$$

where w is a random variable on some probability space Ω with probability measure $\prod(w)$.

Let $\underset{\sim}{\hat{\theta}}(w)$ be the least squares estimator of $\underset{\sim}{\theta}$ in model (3.2.5) pretending ψ is 0. Notz (1989) found the optimal design ξ which minimizes

$$\int_{\Omega} \Phi \left\{ E\left[\left(\underset{\sim}{\hat{\theta}}(w) - \underset{\sim}{\theta} \right) \left(\underset{\sim}{\hat{\theta}}(w) - \underset{\sim}{\theta} \right)^T \right] \right\} d\prod(w)$$

where Φ is a scale valued function of $E[(\underset{\sim}{\hat{\theta}}(w) - \underset{\sim}{\theta})(\underset{\sim}{\hat{\theta}}(w) - \underset{\sim}{\theta})^T]$.

Apart from the "model robustness", there is another direction concerning the dependence of random errors. The design problems against the dependence of random errors had its beginning in 1940-50's. Some important work has been done by Ylvisaker (1964), Sacks and Ylvisaker (1966 and 1968), Bickel and Herzberg (1979 and 1981).

Recently, Wiens (1991) studied robust designs against simultaneously the model violation and the contaminated data by using an M-estimator of the parameters instead of the least squares estimator.

In the next three chapters, we focus on the problem of robust experimental design against model violation. We are going to make some extentions of the results of Stigler (1971) and Notz (1989). We are also going to propose a new consideration, called "bounded bias optimal design", which was suggested by Wiens in 1992. Details will be presented in Chapter 4, Chapter 5, and Chapter 6 respectively.

Chapter 4

Restricted Optimal Designs for Approximately Linear and Quadratic Polynomial Regressions

4.1 Introduction

Consider the following regression model:

$$y_i = y(x_i) = \underset{\sim}{\theta}^T \underset{\sim}{f}\,(x_i) + \epsilon_i, \qquad i = 1, \ldots, n, \tag{4.1.1}$$

where $\underset{\sim}{\theta}^T = (\theta_0, \theta_1, \ldots, \theta_m)$, $\underset{\sim}{f}^T\,(x) = (1, x, \ldots x^m)$, ϵ_i's are i.i.d. with mean 0 and variance $\sigma^2 > 0$, and $x_i \in S := [-1, 1]$, $i = 1, \ldots, n$. Let $\underset{\sim}{y}^T = (y_1, \ldots, y_n)$, $\underset{\sim}{\epsilon}^T = (\epsilon_1, \ldots \epsilon_n)$, and

$$F = \begin{pmatrix} 1 & x_1 & \cdots & x_1^m \\ 1 & x_2 & \cdots & x_2^m \\ \cdots & \cdots & \cdots & \cdots \\ 1 & x_n & \cdots & x_n^m \end{pmatrix}.$$

Then we can write (4.1.1) in the following form, which we call model P_m,

$$P_m: \qquad \underset{\sim}{y} = F \cdot \underset{\sim}{\theta} + \underset{\sim}{\epsilon}\ , \tag{4.1.2}$$

where $\underset{\sim}{\epsilon}$ has mean $\underset{\sim}{0}$ and covariance matrix $\sigma^2 I$. Define $B(\xi) = \int_S \underset{\sim}{f}(x) \underset{\sim}{f}^T(x) d\xi(x)$, where $\xi(x)$ is a design measure defined on S. The least squares estimator of $\underset{\sim}{\theta}$ is then $\hat{\underset{\sim}{\theta}} = B^{-1}(\xi) \int_S \underset{\sim}{f}(x) y(x) d\xi(x)$ which has mean $\underset{\sim}{\theta}$ and covariance matrix $\frac{\sigma^2}{n} B^{-1}(\xi)$ which depends on $\xi(x)$.

Some of the most frequently required properties of a comparison of designs are the following: We say that ξ_1 is preferred to ξ_2 if (i) $|M(\xi_1)| \leq |M(\xi_2)|$ or $|M^{-1}(\xi_1)| \geq |M^{-1}(\xi_2)|$, (ii) $trM(\xi_1) \leq trM(\xi_2)$, (iii) $\int_S d(x, \xi_1) dx \leq \int_S d(x, \xi_2) dx$, or (iv) $\max_{x \in S} d(x, \xi_1) \leq \max_{x \in S} d(x, \xi_2)$, where $M(\xi) = \frac{\sigma^2}{n} B^{-1}(\xi)$ and $d(x, \xi) = \underset{\sim}{f}^T(x) M(\xi) \underset{\sim}{f}(x)$. The design measure ξ_0 which minimizes each of the four cases over some class of design measures is called *D-*, *A-*, *Q-*, or *G*-optimal design respectively.

As we mentioned in Chapter 3, Box and Draper (1959) pointed out the danger of assuming the regression model to be exact, since the violation of the regression model is very possible in practice and the usual optimal designs as we mentioned here are very sensitive to the possible model violation. The usual optimal designs also have some other serious shortcomings. For example, *D*-optimal design (as well as some of the others) permit no check of the adequacy of the model etc. It is therefore desirable to find a criterion and designs which meet the following considerations:

(i) The design should allow for a check of whether or not the assumed model provides an adequate fit to the true regression function.

(ii) If it is concluded that the model is adequate, it should be possible to make reasonably efficient inferences concerning that model.

(iii) The optimal design should not depend on unknown parameters.

In order to find designs to meet the above requirements , S. M. Stigler (1971) proposed some new criteria, so-called *C*-restricted *D-* and *G*-optimal designs. In this chapter, we are going to extend the *C*-restricted optimal design to *A-* and *Q*-optimal criteria and also to a general situation which we call $\underset{\sim}{C}^k$-restricted optimal designs, where $\underset{\sim}{C}^k = (c_1, \ldots, c_k)^T$. Here are the precise definitions:

Definition 4.1.1 *We shall call ξ_0 a $\underset{\sim}{C}^k$-restricted D-optimal design for the model P_m if ξ_0 maximizes $|B_m(\xi)|$ (or minimizes $|B_m^{-1}(\xi)|$) among all designs ξ satisfying*

$|B_{m+j-1}(\xi)| \leq c_j \, |B_{m+j}(\xi)|$ *for* $j = 1, \ldots, k$.

The justification for this choice of definition is based on the fact that if $\hat{\theta}_{m+j}$ is the least squares estimator of θ_{m+j} for the model P_{m+j}, and ξ corresponds to an experiment run at $x_1, \ldots, x_n$, then n$\cdot$var$(\hat{\theta}_{m+j}) = |B_{m+j-1}(\xi)| \, |B_{m+j}(\xi)|^{-1} \sigma^2, \quad j = 1, \ldots, k$. Thus this criterion says "minimize the generalized variance of the least squares estimators $\hat{\underset{\sim}{\theta}}^T = (\hat{\theta}_0, \hat{\theta}_1, \ldots, \hat{\theta}_m)$ for the model P_m subject to the constraints that $\text{var}(\hat{\theta}_{m+j}) \leq c_j \cdot \dfrac{\sigma^2}{n}, \quad j = 1, \ldots, k$. Similarly, we have:

Definition 4.1.2 *We shall call ξ_0 a $\underset{\sim}{C}^k$-restricted A-optimal design for the model P_m if ξ_0 minimizes $trB_m^{-1}(\xi)$ among all designs ξ satisfying $|B_{m+j-1}(\xi)| \leq c_j \, |B_{m+j}(\xi)|$ for $j = 1, \ldots, k$.*

The motivation of A-optimality is that the minimization of $trB_m^{-1}(\xi)$ is equivalent to the minimization of the mean dispersion of the estimates of the parameters $T = m^{-1} \sum_{i=1}^{m} \text{var}(\hat{\theta}_i)$, since T is proportional to $trB_m^{-1}(\xi)$. The meaning of Q- and G-optimality are clear. We simply state them as the following two definitions.

Definition 4.1.3 *We shall call ξ_0 a $\underset{\sim}{C}^k$-restricted Q-optimal design for the model P_m if ξ_0 minimizes $\int_S \, d_m(x, \xi) dx$ among all designs ξ satisfying $|B_{m+j-1}(\xi)| \leq c_j \, |B_{m+j}(\xi)|$ for $j = 1, \ldots, k$.*

Definition 4.1.4 *We shall call ξ_0 a $\underset{\sim}{C}^k$-restricted G-optimal design for the model P_m if ξ_0 minimizes $\max_{x \in S} \, d_m(x, \xi)$ among all designs ξ satisfying $|B_{m+j-1}(\xi)| \leq c_j \, |B_{m+j}(\xi)|$ for $j = 1, \ldots, k$.*

According to the definitions of $\underset{\sim}{C}^k$-restricted optimal designs , Stigler (1971) solved the $\underset{\sim}{C}^k$-restricted D- and G- optimal designs for the case of $m = 1$ and $k = 1$. In this chapter, we are going to find the $\underset{\sim}{C}^k$-restricted D-, A-, Q-, and G- optimal designs for (i) $m = 1, \ k = 1$; (ii) $m = 1, \ k = 2$; and (iii) $m = 2, \ k = 1$.

Case (i) is studied in Section 4.3. We use different methods to find the same result as in Stigler(1970). Moreover, we point out that the C-restricted D-, A-, Q-, and G-optimal designs are all the same in this case.

In Section 4.5, we discuss the case when $m = 2$ and $k = 1$. We simplify the problem to a non-linear programming problem with three variables. Numerical searching for the $\underset{\sim}{C}^2$-restricted optimal design is needed.

Some interesting results are found in case (ii). When $m = 1$ and $k = 2$, the problem becomes the following:

$$\max_{\xi \in \mathcal{F}} \mu_2 \qquad \text{subject to } \frac{1}{c_1} \le \mu_4 - \mu_2^2 \text{ and } \mu_2 \le c_2(\mu_2\mu_6 - \mu_4^2),$$

where $\mathcal{F} = \{\xi : \sigma(\xi) \subseteq S\}$.

With the aid of preliminaries in Section 4.2, we have proved that the restricted maximization of μ_2 over $\mathcal{F}$ is equivalent to that over $\mathcal{F}_s$ and furthermore is equivalent to that over $\mathcal{F}_1$, where $\mathcal{F}_s = \{\xi : \xi \in \mathcal{F}, \xi(-x) = \xi(x)\}$ and $\mathcal{F}_1 = \{\xi : \xi \in \mathcal{F}_s,$ $\xi = \frac{\alpha}{2}\Delta_{\pm 1} + \frac{\beta}{2}\Delta_{\pm\sqrt{x}} + (1-\alpha-\beta)\Delta_0,\ 0 \le \alpha \le 1,\ 0 \le \beta \le 1,$ and $0 \le x \le 1\}$.

There are two limiting cases: (1) $c_2 = \infty$ and (2) $c_1 = \infty$. When $c_2 = \infty$, we know that this is the case of $m = 1$ and $k = 1$. For the case when $c_1 = \infty$, we find the explicit solution to the problem. In general, the solution can only be found by numerical search. However, when c_1 and c_2 have some special relationship, we are still able to find the solution explicitly. These are the main results in Section 4.4 which are presented by Theorem 4.4.3 and Theorem 4.4.5.

4.2 Preliminaries

The search for $\mathcal{C}^k$-restricted D-optimal designs with $k = 1$ can be simplified by searching for the optimal design within the class of symmetric design measures. This fact was proved by Stigler (1971). In this section, we first indicate that the fact is also true for restricted A- and Q- optimal designs and for $k > 1$. In order to prove this fact, we need the following two lemmas:

Lemma 4.2.1 *Let $A = (a_{ij})_{n\times n}$ and $B = (b_{ij})_{n\times n}$ be two matrices. If $b_{ij} = (-1)^{i+j}a_{ij}$, then $|A| = |B|$.*

Proof: We know that

$$|A| = \sum_{(j_1,\ldots,j_n)} (-1)^{\tau(j_1,\ldots,j_n)} a_{1j_1},\ldots a_{nj_n}, \text{ and } |B| = \sum_{(j_1,\ldots,j_n)} (-1)^{\tau(j_1,\ldots,j_n)} b_{1j_1},\ldots,b_{nj_n},$$

where $\tau(j_1,\ldots,j_n)$ is the number of inverse order of the permutation $(j_1,\ldots,j_n)$, and the summation is over all the possible permutations of $(1,\ldots,n)$. We claim that $a_{1j_1},\ldots,a_{nj_n} = b_{1j_1},\ldots,b_{nj_n}$ for all $(j_1,\ldots,j_n)$.

Followed by the assumption $b_{ij} = (-1)^{i+j}a_{ij}$, it is sufficient to show that there are even number of pairs (i, j_i) such that $i + j_i$ is odd. This must be true since $\sum_{i=1}^{n}(i + j_i) = \sum_{i=1}^{n} i + \sum_{i=1}^{n} j_i = n(n+1)$ which is even. □

Lemma 4.2.2 *Let $A = (a_{ij})_{n\times n}$ and $B = (b_{ij})_{n\times n}$ be two non-degenerate matrices. Denote $A^{-1} = (a^*_{ij})_{n\times n}$ and $B^{-1} = (b^*_{ij})_{n\times n}$. If $b_{ij} = (-1)^{i+j}a_{ij}$, then $b^*_{ij} = (-1)^{i+j}a^*_{ij}$.*

Proof: Let

$$P = (p_{ij})_{n\times n} = \begin{pmatrix} -1 & & & & 0 \\ & \ddots & & & \\ & & (-1)^i & & \\ & & & \ddots & \\ 0 & & & & (-1)^n \end{pmatrix},$$

$$\text{i.e.,} \quad p_{ij} = \begin{cases} (-1)^i & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}.$$

Then, we have $P^{-1} = P$, and $B = PAP$. Hence, we get $B^{-1} = P^{-1}A^{-1}P^{-1} = PA^{-1}P$. This implies that $b^*_{ij} = (-1)^{i+j}a^*_{ij}$. □

Let $V(c_j)$ be the class of designs ξ satisfying the constraint $|B_{m+j-1}(\xi)| \leq c_j\,|B_{m+j}(\xi)|$, $j = 1, \ldots, k$, and $V(\underset{\sim}{C}^k) = \bigcap_{j=1}^{k} V(c_j)$. We claim that the following lemma is true.

Lemma 4.2.3 *(i)* $V(\underset{\sim}{C}^k)$ *is convex.* *(ii)* $\xi(x) \in V(\underset{\sim}{C}^k)$ *if and only if* $\xi(-x) \in V(\underset{\sim}{C}^k)$.

Proof: (i) The convexity of each $V(c_j)$ was proved by Stigler (1971). Hence $V(\underset{\sim}{C}^k)$ is convex.

(ii) $\xi(x) \in V(\underset{\sim}{C}^k)$ implies that $\xi(x) \in V(c_j)$ for $j = 1, \ldots, k$. Hence we have $|B_{m+j-1}(\xi)| \leq c_j\,|B_{m+j}(\xi)|$ for $j = 1, \ldots, k$. Let b_{il} be the i^{th} row and l^{th} column element of $B(\xi(x))$ and b^*_{il} be the i^{th} row and l^{th} column element of $B(\xi(-x))$. It is clear that $b_{il} = (-1)^{i+l}b^*_{il}$, since we have $B_m(\xi) = \int_S \underset{\sim}{f}\,(x)\,\underset{\sim}{f}^T\,(x)d\xi(x)$ with $\underset{\sim}{f}^T\,(x) = (1, x, \ldots, x^m)$. By Lemma 4.2.1, we have $|B_{m+j-1}(\xi(-x))| \leq c_j\,|B_{m+j}(\xi(-x))|$ for $j = 1, \ldots, k$. This implies that $\xi(-x) \in V(c_j)$ for $j = 1, \ldots, k$. Hence we have $\xi(-x) \in V(\underset{\sim}{C}^k)$.

Similarly, we can show that $\xi(-x) \in V(\underset{\sim}{C}^k)$ implies $\xi(x) \in V(\underset{\sim}{C}^k)$. □

We denote $D(\underset{\sim}{C}^k)$, $A(\underset{\sim}{C}^k)$, $Q(\underset{\sim}{C}^k)$ to be the sets of $\underset{\sim}{C}^k$-restricted *D-, A-, Q-*optimal designs respectively, i.e., $D(\underset{\sim}{C}^k) = \{\xi_0 : |B^{-1}(\xi_0)| = \min_{\xi \in V(\underset{\sim}{C}^k)} |B^{-1}(\xi)|\}$, $A(\underset{\sim}{C}^k) = \{\xi_0 : trB^{-1}(\xi_0) = \min_{\xi \in V(\underset{\sim}{C}^k)} trB^{-1}(\xi)\}$, and $Q(\underset{\sim}{C}^k) = \{\xi_0 : \int_{-1}^{1} d(x, \xi_0)\,dx = \min_{\xi \in V(\underset{\sim}{C}^k)} \int_{-1}^{1} d(x, \xi)\,dx\}$. Then we have:

Lemma 4.2.4 *(i)* $D(\underset{\sim}{C}^k)$, $A(\underset{\sim}{C}^k)$, *and* $Q(\underset{\sim}{C}^k)$ *are the convex subsets of* $V(\underset{\sim}{C}^k)$.

(ii) $\xi(x) \in D(\underset{\sim}{C}^k)$, $A(\underset{\sim}{C}^k)$, *or* $Q(\underset{\sim}{C}^k)$ *if and only if* $\xi(-x) \in D(\underset{\sim}{C}^k)$, $A(\underset{\sim}{C}^k)$, *or* $Q(\underset{\sim}{C}^k)$ *respectively.*

Proof: (i) The following fact was noticed by Kiefer (1959):

" If A and B are any symmetric positive definite matrices, then

$$[\lambda A + (1-\lambda)B]^{-1} \leq \lambda A^{-1} + (1-\lambda)B^{-1},$$

where we write $A \le B$ to mean $B - A$ is semipositive definite."

Let $\xi_1, \xi_2 \in A(\underset{\sim}{C}^k)$ and $\xi^* = \lambda\xi_1 + (1-\lambda)\xi_2$, where $0 < \lambda < 1$. By the convexity of $V(\underset{\sim}{C}^k)$, we have $\xi^* \in V(\underset{\sim}{C}^k)$. It is obvious that $B_m(\xi^*) = \lambda B_m(\xi_1) + (1-\lambda)B_m(\xi_2)$. Taking $A = B_m(\xi_1)$ and $B = B_m(\xi_2)$, we have $B_m^{-1}(\xi^*) \le \lambda B_m^{-1}(\xi_1) + (1-\lambda)\, B_m^{-1}(\xi_2)$. Hence,

$$trB_m^{-1}(\xi^*) \le \lambda trB_m^{-1}(\xi_1) + (1-\lambda)trB_m^{-1}(\xi_2) = trB_m^{-1}(\xi_1).$$

Note that the second equality followed by the A-optimality of ξ_1 and ξ_2. Since $\xi_1 \in A(\underset{\sim}{C}^k)$, $\xi^* \in V(\underset{\sim}{C}^k)$, and $trB_m^{-1}(\xi^*) \le trB_m^{-1}(\xi_1)$, we must have $trB_m^{-1}(\xi^*) = trB_m^{-1}(\xi_1)$, and hence $\xi^* \in A(\underset{\sim}{C}^k)$.

Similarly, let $\xi_1, \xi_2 \in Q(\underset{\sim}{C}^k)$ and $\xi^* = \lambda\xi_1 + (1-\lambda)\xi_2$, where $0 < \lambda < 1$. We know that $B_m^{-1}(\xi^*) \le \lambda B_m^{-1}(\xi_1) + (1-\lambda)B_m^{-1}(\xi_2)$. Hence we have $d_m(x,\xi^*) \le \lambda d_m(x,\xi_1) + (1-\lambda)d_m(x,\xi_2)$ and

$$\int_S d_m(x,\xi^*)\, dx \le \lambda \int_S d_m(x,\xi_1)\, dx + (1-\lambda)\int_S d_m(x,\xi_2)\, dx = \int_S d_m(x,\xi_1)\, dx.$$

Again, the second equality followed by the Q-optimality of ξ_1 and ξ_2. The fact that $\xi_1 \in Q(\underset{\sim}{C}^k)$, $\xi^* \in V(\underset{\sim}{C}^k)$, and $\int_S d_m(x,\xi^*)\, dx \le \int_S d_m(x,\xi_1)\, dx$ implies $\int_S d_m(x,\xi^*)\, dx = \int_S d_m(x,\xi_1)\, dx$ and $\xi^* \in Q(\underset{\sim}{C}^k)$.

The convexity of $D(\underset{\sim}{C}^k)$ has been proved by Stigler (1971).

(ii) Let

$$B_l(\xi(x)) = \int_{-1}^{1} \underset{\sim}{f}(x)\cdot \underset{\sim}{f}^T(x) d\xi(x) = \int_{-1}^{1} \begin{pmatrix} 1 \\ x \\ \vdots \\ x^l \end{pmatrix} (1, x, \ldots, x^l)\, d\xi(x)$$

$$= \begin{pmatrix} 1 & \mu_1 & \cdots & \mu_l \\ \mu_1 & \mu_2 & \cdots & \mu_{l+1} \\ \cdots & \cdots & \cdots & \cdots \\ \mu_l & \mu_{l+1} & \cdots & \mu_{2l} \end{pmatrix},$$

where $\mu_i = \int_{-1}^{1} x^i \, d\xi(x)$. Then we have

$$B_l(\xi(-x)) = \begin{pmatrix} 1 & -\mu_1 & \cdots & (-1)^l \mu_l \\ -\mu_1 & \mu_2 & \cdots & (-1)^{l+1}\mu_{l+1} \\ \cdots & \cdots & \cdots & \cdots \\ (-1)^{l+1}\mu_{l+1} & (-1)^{l+2}\mu_{l+2} & \cdots & (-1)^{2l}\mu_{2l} \end{pmatrix}.$$

By Lemma 4.2.1, we have $|B_l(\xi(x))| = |B_l(\xi(-x))|$ for any natural number l, especially for $l = m, \ldots, m+k$.

Let $B_m^{-1}(\xi(x)) = (b_{ij}^{+})_{(m+1)\times(m+1)}$ and $B_m^{-1}(\xi(-x)) = (b_{ij}^{-})_{(m+1)\times(m+1)}$. By Lemma 4.2.2, we have $b_{ij}^{-} = (-1)^{i+j} b_{ij}^{+}$. Hence we get $trB^{-1}(\xi(x)) = trB^{-1}(\xi(-x))$.

Note that

$$d_m(x, \xi(x)) = (1, x, \ldots, x^m) B_m^{-1}(\xi(x)) \begin{pmatrix} 1 \\ x \\ \vdots \\ x^m \end{pmatrix} = \sum_{i=0}^{m}\sum_{j=0}^{m} b_{ij}^{+} x^{i+j}$$

and

$$d_m(x, \xi(-x)) = (1, x, \ldots, x^m) B_m^{-1}(\xi(-x)) \begin{pmatrix} 1 \\ x \\ \vdots \\ x^m \end{pmatrix} = \sum_{i=0}^{m}\sum_{j=0}^{m} b_{ij}^{-} x^{i+j}.$$

Although $d_m(x, \xi(x))$ and $d_m(x, \xi(-x))$ are different for the terms when $i+j$ is odd, but they are the same for the terms when $i+j$ is even. Hence we have

$$\begin{aligned} \int_{-1}^{1} d_m(x, \xi(-x))\, dx &= \textstyle\sum_{i=0}^{m}\sum_{j=0}^{m} \int_{-1}^{1} (-1)^{i+j} b_{ij}^{+} x^{i+j}\, dx = \sum\sum_{i+j \text{ even}} \int_{-1}^{1} b_{ij}^{+} x^{i+j}\, dx \\ &= \textstyle\int_{-1}^{1} d_m(x, \xi(x))\, dx. \end{aligned}$$

Part (ii) of Lemma 4.2.4 follows by the above results. □

The consequence of Lemma 4.2.4 is that there is a symmetrical optimal design. In fact, let ξ_1 be a $\underset{\sim}{C}^k$-restricted optimal design and define $\xi_2(x) = \xi_1(-x)$; then it follows that $(\xi_1 + \xi_2)/2$ is a symmetrical $\underset{\sim}{C}^k$-restricted optimal design. Hence, for

$\underset{\sim}{C}^k$-restricted D-, A-, and Q- optimal designs, we only need search optimal solutions within the class of symmetrical designs.

The problem of search for $\underset{\sim}{C}^k$-restricted optimal designs can be further simplified by some well-known results (see Karlin and Shapley(1953), Shohat and Tamarkin(1943), and also Stigler(1971)) which will reduce the problem to a non-linear programming problem with l variables ($l+1$ if l is even). These variables are the $(l-1)/2$ points ($l/2$ of l is even). x_i, $i = 1, \ldots, (l-1)/2$ in $[0, 1]$ such that the optimal ξ is supported by $\{\pm x_i\}$ together with ± 1 and 0; and the $(l+1)/2$ weights ($(l+2)/2$ if l is even) $\xi(x_i)$, $i = 1, \ldots, (l-1)/2$ and $\xi(0)$. However, we are not going to use this result, since it may not sufficiently reduce the problem. Instead, we would like to provide a direct approach to simplify the problem of finding $\underset{\sim}{C}^k$-restricted optimal designs. But first, we introduce some necessary notations. Let

$$\mathcal{F}_0 = \left\{\xi : \xi \in \mathcal{F},\, \xi = \frac{\alpha}{2}\Delta_{\pm\sqrt{x}} + (1-\alpha)\Delta_0,\ 0 \leq \alpha \leq 1,\ 0 \leq x \leq 1\right\}$$

and

$$\mathcal{F}_1 = \{\xi : \xi \in \mathcal{F},\, \xi = \frac{\alpha}{2}\Delta_{\pm 1} + \frac{\beta}{2}\Delta_{\pm\sqrt{x}} + (1-\alpha-\beta)\Delta_0,\ 0 \leq \alpha \leq 1,\ 0 \leq \beta \leq 1,$$
$$0 \leq x \leq 1,\ \alpha+\beta \leq 1\},$$

where $\mathcal{F}$ is the class of all design measures supported on $[-1, 1]$, and let

$$\mathcal{E} = \left\{\xi : \xi \in \mathcal{F},\, \xi = \frac{1}{2}\Delta_{\sqrt{x}} + \frac{1}{2}\Delta_{-\sqrt{x}},\ 0 \leq x \leq 1\right\}.$$

We define a functional T_0 as

$$T_0(\xi) : \mathcal{F} \to \mathbf{R}^2 \text{ by } T_0(\xi) = \left(\int_{-1}^{1} x^2 d\xi(x), \int_{-1}^{1} x^4 d\xi(x)\right)$$

and a functional T_1 as

$$T_1(\xi) : \mathcal{F} \to \mathbf{R}^3 \text{ by } T_1(\xi) = \left(\int_{-1}^{1} x^2 d\xi(x), \int_{-1}^{1} x^4 d\xi(x),\ \int_{-1}^{1} x^6 d\xi(x)\right).$$

It is clear that

$$T_0(\mathcal{E}) = \{(x, y) : x = x,\ y = x^2,\ 0 \leq x \leq 1\} := S_0$$

and

$$T_1(\mathcal{E}) = \{(x, y, z) : x = x,\ y = x^2,\ z = x^3,\ 0 \le x \le 1\} := S_1.$$

Furthermore, we denote $\hat{S}_0$ to be the convex hull of S_0 and $\hat{S}_1$ to be the convex hull of S_1 . Then we can prove that the next two theorems are true.

Theorem 4.2.5 $T_0(\mathcal{F}_0) = T_0(\mathcal{F}) = \hat{S}_0$.

Proof: Theorem 4.2.5 is a special case of Theorem 2.2.4. □

Theorem 4.2.6 $T_1(\mathcal{F}_1) = T_1(\mathcal{F}) = \hat{S}_1$.

We would like to present a lemma before we prove Theorem 4.2.6. Let $p_0 = (0,0,0)$, $p_1 = (1,1,1)$, $p_k = (k, k^2, k^3)$, where $0 \le k \le 1$, and

$$P = \{A_k : A_k \text{ is the interior and boundary of the triangle with vertices } p_0, p_1, p_k,\ 0 \le k \le 1\}.$$

The following lemma plays an important role to the proof of Theorem 4.2.6.

Lemma 4.2.7 $P = \hat{S}_1$.

Proof: (i) It is obvious that $P \subseteq \hat{S}_1$, since $p_0, p_1, p_k \in S_1$, and any point $p \in A_k$ is a convex combination of p_0, p_1, and p_k.

(ii) We are going to show that "if $p \notin A_k$ for any $k \in [0,1]$, then $p \notin \hat{S}_1$."

Let $L = \{(x, y, z) : x = t,\ y = t,\ z = t,\ 0 \le t \le 1\}$, and $\Pi = \{\pi_k : z = -kx + (1+k)y,\ k \in \mathbf{R}^1\}$. It is obvious that π_k goes through p_0 and p_1, and $\bigcup_{k \in \mathbf{R}^1} \pi_k = \mathbf{R}^3$. Moreover, we have $\pi_k \cap S_1 = \{p_0,\ p_1\}$ for $k \le 0$ or $k \ge 1$, since the following system of equations

$$\begin{cases} z = -kx + (1+k)y \\ x = x,\ y = x^2,\ z = x^3 \end{cases}$$

has solutions $x_1 = 0$, $x_2 = 1$, and $x_3 = k$.

For any point $p \in \pi_k \backslash L$, where $k \le 0$ or $k \ge 1$, it is easy to see that there is a π'_k such that $p \in \pi'_k$ and $\pi'_k \cap S_1 = \emptyset$. Hence $\pi'_k \cap \hat{S}_1 = \emptyset$ and $p \notin \hat{S}_1$.

We are now going to show that for any point $p \in \pi_k \backslash A_k$ yields $p \notin \hat{S}_1$ when $0 < k < 1$.

Let S_1^* and A_k^* be the projections of $\hat{S}_1$ and A_k onto the plane xOy, i.e., $z = 0$. We have

$$S_1^* = \{(x,y) : 0 \leq x \leq 1,\ x^2 \leq y \leq x\}$$

and

$$A_k^* = \{(x,y) : 0 \leq x \leq k,\ kx \leq y \leq x;\ \text{and}\ k \leq x \leq 1,\ (1+k)x - k \leq y \leq x\}.$$

We define $B_1 = \{(x,y) : 0 < x < k,\ x^2 \leq y < kx\}$, and $B_2 = \{(x,y) : k < x < 1,\ x^2 \leq y < (1+k)x - k\}$. It is clear that $S_1^* = A_k^* \cup B_1 \cup B_2$.

Let p^* be the projection of p onto xOy. For any point $p \in \pi_k \backslash L$, $p^* \notin S_1^*$ implies $p \notin \hat{S}_1$. Hence, we only need to consider the point $p \in \pi_k \backslash L$, such that $p^* \in B_1 \cup B_2$.

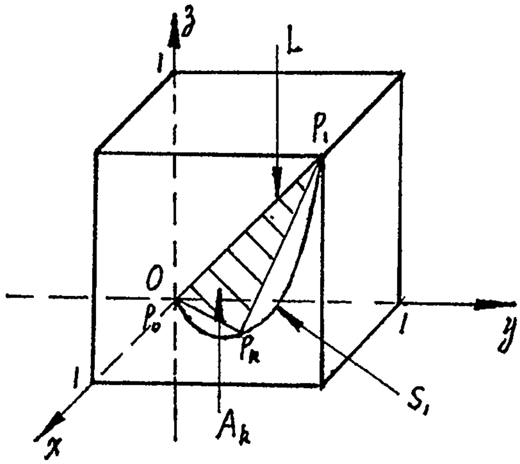

Figure 4.2.1

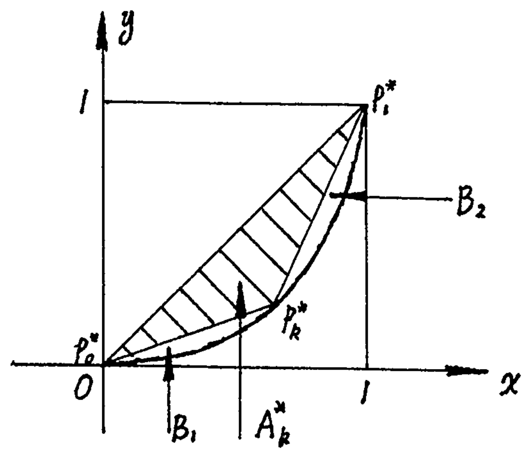

Figure 4.2.2

Case 1 $\quad p^* \in B_1$

For $p = (x_0,\ y_0,\ z_0) \in \pi_k$ such that $p^* \in B_1$, we have

$$y_0 < kx_0. \tag{4.2.1}$$

Let $\pi_1(a)$ be a plane going through p_0 and p. Then $\pi_1(a)$ satisfies the following:

$$\begin{cases} \dfrac{x}{a} + \dfrac{y}{b} + z = 0 & (4.2.2) \\[2mm] \dfrac{x_0}{a} + \dfrac{y_0}{b} + z_0 = 0\ . & (4.2.3) \end{cases}$$

Since $p \in \pi_k$, we also have

$$z_0 = -kx_0 + (1+k)y_0. \tag{4.2.4}$$

Put (4.2.4) into (4.2.3) to get

$$\frac{x_0}{a} + \frac{y_0}{b} - kx_0 + (1+k)y_0 = 0 \qquad \text{and} \qquad \frac{1}{b} = \frac{[kx_0 - (1+k)y_0]a - x_0}{y_0 a}. \tag{4.2.5}$$

Put (4.2.5) into (4.2.2) to get

$$\frac{x}{a} + \frac{\{[kx_0 - (1+k)y_0]a - x_0\}y}{y_0 a} + z = 0,$$

and hence we have

$$y_0 x + \{[kx_0 - (1+k)y_0]a - x_0\}y + y_0 a z = 0. \tag{4.2.6}$$

Thus (4.2.6) is the equation of $\pi_1(a)$. Varying a within $\mathbf{R}^1$ we get the family of all the planes going through p_0 and p. We denote this family as Π_1, i.e., $\Pi_1 = \{\pi_1(a) : a \in \mathbf{R}^1\}$.

We are now seeking a plane $\pi_1(a^*)$ such that $\pi_1(a^*) \cap S_1 = \{p_0\}$. Solving the following system of equations

$$\begin{cases} y_0 x + \{[kx_0 - (1+k)y_0]a - x_0\}y + y_0 a z = 0 \\ x = x, \quad y = x^2, \quad z = x^3 \end{cases},$$

we get

$$\begin{aligned} & y_0 x + \{[kx_0 - (1+k)y_0]a - x_0\}x^2 + y_0 a x^3 \\ = {} & x\{y_0 a x^2 + [(kx_0 - (1+k)y_0)a - x_0]x - y_0\} = 0 . \end{aligned} \tag{4.2.7}$$

We know that $x = 0$ is a solution of (4.2.7), and we hope that there exists a real number a such that

$$y_0 a x^2 + \{[kx_0 - (1+k)y_0]a - x_0\}x + y_0 = 0 \tag{4.2.8}$$

has no solution. (What we really need is that (4.2.8) has no solution in [0,1].)

The necessary and sufficient condition for (4.2.8) to have no solution is

$$\delta_1 := \{[kx_0 - (1+k)y_0]a - x_0\}^2 - 4y_0^2 a < 0. \tag{4.2.9}$$

(1) For $kx_0 - (1+k)y_0 \neq 0$, we have

$$\begin{aligned}
\delta_1 &= [kx_0 - (1+k)y_0]^2 a^2 + x_0^2 - 2x_0[kx_0 - (1+k)y_0]a - 4y_0^2 a \\
&= [kx_0 - (1+k)y_0]^2 a^2 - 2\{x_0[kx_0 - (1+k)y_0] + 2y_0^2\}a + x_0^2 \\
&= [kx_0 - (1+k)y_0]^2 \left\{ a^2 - 2\frac{x_0[kx_0 - (1+k)y_0] + 2y_0^2}{[kx_0 - (1+k)y_0]^2} a \right. \\
&\quad \left. + \frac{x_0^2}{[kx_0 - (1+k)y_0]^2} \right\} \\
&= [kx_0 - (1+k)y_0]^2 \left\{ \left(a - \frac{x_0[kx_0 - (1+k)y_0] + 2y_0^2}{[kx_0 - (1+k)y_0]^2} \right)^2 \right. \\
&\quad \left. - \frac{\{x_0[kx_0 - (1+k)y_0] + 2y_0^2\}^2}{[kx_0 - (1+k)y_0]^4} + \frac{x_0^2}{[kx_0 - (1+k)y_0]^2} \right\}.
\end{aligned}$$

Note that

$$\begin{aligned}
& -\frac{\{x_0[kx_0 - (1+k)y_0] + 2y_0^2\}^2}{[kx_0 - (1+k)y_0]^4} + \frac{x_0^2}{[kx_0 - (1+k)y_0]^2} \\
= & \frac{x_0^2[kx_0 - (1+k)y_0]^2 - \{x_0[kx_0 - (1+k)y_0] + 2y_0^2\}^2}{[kx_0 - (1+k)y_0]^4} \\
= & \frac{1}{[kx_0 - (1+k)y_0]^4} \cdot \{x_0^2[kx_0 - (1+k)y_0]^2 - x_0^2[kx_0 - (1+k)y_0]^2 \\
& -4y_0^4 - 4x_0y_0^2[kx_0 - (1+k)y_0]\} \\
= & \frac{-4y_0^2}{[kx_0 - (1+k)y_0]^4}\{y_0^2 + x_0[kx_0 - (1+k)y_0]\} \\
= & \frac{-4y_0^2}{[kx_0 - (1+k)y_0]^4} \cdot \{y_0^2 + kx_0^2 - (1+k)x_0y_0\} \\
= & \frac{-4y_0^2}{[kx_0 - (1+k)y_0]^4}(x_0 - y_0)(kx_0 - y_0) < 0.
\end{aligned}$$

For $(x_0, y_0) \in B_1$, we have $x_0 > y_0$ and $y_0 < kx_0$ (4.2.1). Hence, we get $\delta_1 < 0$ if we choose

$$a^* = \frac{x_0[kx_0 - (1+k)y_0] + 2y_0^2}{[kx_0 - (1+k)y_0]^2}.$$

(2) For $kx_0 - (1+k)y_0 = 0$, we have $\delta_1 = x_0^2 - 4y_0^2 a$. Hence $\delta_1 < 0$ if we choose

$$a^* > x_0^2/4y_0^2.$$

The conclusion of the above argument is that there exists a real number a^* such that (4.2.8) has no solution. Hence, there is a $\pi_1(a^*) \in \Pi_1$, such that $\pi_1(a^*) \cap S_1 = \{p_0\}$. Note that p_0 is an end point of S_1, and S_1 is located at one side of $\pi_1(a^*)$. It is easy to see that there is a plane π_1^* such that $p \in \pi_1^*$ and $\pi_1^* \cap S_1 = \emptyset$. This implies that $\pi_1^* \cap \hat{S}_1 = \emptyset$, and hence we conclude that $p \notin \hat{S}_1$.

Case 2 $p^* \in B_2$

For $p = (x_0, y_0, z_0) \in \pi_k$ such that $p^* \in B_2$, we have

$$y_0 < (1+k)x_0 - k \ . \tag{4.2.10}$$

Let $\pi_2(a)$ be a plane going through p_1 and p. Then $\pi_2(a)$ satisfies the following:

$$\begin{cases} \dfrac{x-1}{a} + \dfrac{y-1}{b} + (z-1) = 0 & (4.2.11) \\ \dfrac{x_0-1}{a} + \dfrac{y_0-1}{b} + (z_0-1) = 0 \ . & (4.2.12) \end{cases}$$

Put (4.2.4) into (4.2.12) to get

$$\frac{x_0-1}{a} + \frac{y_0-1}{b} - kx_0 + (1+k)y_0 - 1 = 0,$$

and

$$\frac{1}{b} = \frac{[kx_0 - (1+k)y_0 + 1]a - (x_0-1)}{(y_0-1)a} \tag{4.2.13}$$

Put (4.2.13) into (4.2.11) to get

$$\frac{x-1}{a} + \frac{(y-1)\{[kx_0 - (1+k)y_0 + 1]a - (x_0-1)\}}{(y_0-1)a} + (z-1) = 0$$

and hence, we have

$$(y_0-1)(x-1) + \{[kx_0-(1+k)y_0+1]a-(x_0-1)\}(y-1) + (y_0-1)a(z-1) = 0. \tag{4.2.14}$$

Thus (4.2.14) is the equation of $\pi_2(a)$. Varying a within $\mathbf{R}^1$ we get the family of all the planes going through p_1 and p. We denote this family as Π_2, i.e., $\Pi_2 = \{\pi_2(a) : a \in \mathbf{R}^1\}$.

We are now seeking a plane $\pi_2(a^*) \in \Pi_2$ such that $\pi_2(a^*) \cap S_1 = \{p_1\}$. Solving the following system of equations

$$\begin{cases} (y_0-1)(x-1) + \{[kx_0 - (1+k)y_0 + 1]a - (x_0-1)\}(y-1) \\ \qquad +(y_0-1)a(z-1) = 0 \\ x = x, \quad y = x^2, \quad z = x^3 \end{cases},$$

we get

$$
\begin{aligned}
& (y_0-1)(x-1)+\{[kx_0-(1+k)y_0+1]a-(x_0-1)\}(x^2-1) \\
& \qquad +(y_0-1)a(x^3-1) \\
= \; & (x-1)\{(y_0-1)+[(kx_0-(1+k)y_0+1)a-(x_0-1)](x+1) \\
& \qquad +(y_0-1)a(x^2+x+1)\}=0\,. \qquad (4.2.15)
\end{aligned}
$$

We know that $x=1$ is a solution of (4.2.15), and we hope that there exists a real number a such that

$$
\begin{aligned}
& (y_0-1)+\{[kx_0-(1+k)y_0+1]a-(x_0-1)\}(x+1) \\
& \qquad +(y_0-1)a(x^2+x+1) \\
= \; & (y_0-1)ax^2+\{(y_0-1)a+[kx_0-(1+k)y_0+1]a-(x_0-1)\}x \\
& +(y_0-1)+(y_0-1)a+\{[kx_0-(1+k)y_0+1]a-(x_0-1)\} \\
= \; & (y_0-1)ax^2+(kx_0a-ky_0a-x_0+1)x \\
& +(y_0+kx_0a-ky_0a-x_0)=0 \qquad (4.2.16)
\end{aligned}
$$

has no solution. (Again, what we really want is that (4.2.16) has no solution in [0,1].)

The necessary and sufficient condition for (4.2.16) to have no solution is $\delta_2<0$, where

$$
\begin{aligned}
\delta_2 = \; & (kx_0a-ky_0a-x_0+1)^2-4(y_0-1)a(y_0+kx_0a-ky_0a-x_0) \\
= \; & [ka(x_0-y_0)+(1-x_0)]^2-4(y_0-1)a[ka(x_0-y_0)-(x_0-y_0)] \\
= \; & k^2(x_0-y_0)a^2+(1-x_0)^2+2k(x_0-y_0)(1-x_0)a \\
& -4k(y_0-1)(x_0-y_0)a^2+4(y_0-1)(x_0-y_0)a \\
= \; & [k^2(x_0-y_0)^2-4k(y_0-1)(x_0-y_0)]a^2+2[k(x_0-y_0)(1-x_0) \\
& +2(y_0-1)(x_0-y_0)]a+(1-x_0)^2 \\
= \; & k(x_0-y_0)[k(x_0-y_0)+4(1-y_0)] \\
& \cdot\left\{a^2+2\cdot\frac{[k(1-x_0)-2(1-y_0)]}{k[k(x_0-y_0)+4(1-y_0)]}\right. \\
& \left.+\frac{(1-x_0)^2}{k(x_0-y_0)[k(x_0-y_0)+4(1-y_0)]}\right\} \\
= \; & k(x_0-y_0)[k(x_0-y_0)+4(1-y_0)] \\
& \cdot\left\{\left(a+\frac{k(1-x_0)-2(1-y_0)}{k[k(x_0-y_0)+4(1-y_0)]}\right)^2-\frac{[k(1-x_0)-2(1-y_0)]^2}{k^2[k(x_0-y_0)+4(1-y_0)]^2}\right. \\
& \left.+\frac{(1-x_0)^2}{k(x_0-y_0)[k(x_0-y_0)+4(1-y_0)]}\right\}.
\end{aligned}
$$

Note that $k(x_0 - y_0) + 4(1 - y_0) \neq 0$ when $p^* \in B_2$, and

$$
\begin{aligned}
& -\frac{[k(1-x_0)-2(1-y_0)]^2}{k^2[k(x_0-y_0)+4(1-y_0)]^2} + \frac{(1-x_0)^2}{k(x_0-y_0)[k(x_0-y_0)+4(1-y_0)]} \\
= \; & \frac{(1-x_0)^2 \cdot k[k(x_0-y_0)+4(1-y_0)] - (x_0-y_0)[k(1-x_0)-2(1-y_0)]^2}{k^2(x_0-y_0)[k(x_0-y_0)+4(1-y_0)]^2} \\
= \; & \frac{-4(1-y_0)}{k^2(x_0-y_0)[k(x_0-y_0)+4(1-y_0)]^2} \cdot [(1-y_0)(x_0-y_0) - k(1-x_0)^2 \\
& -k(1-x_0)(x_0-y_0)] \\
= \; & \frac{-4(1-y_0)[(1+k)x_0-k-y_0]}{k^2(x_0-y_0)[k(x_0-y_0)+4(1-y_0)]^2} < 0,
\end{aligned}
$$

since for $(x_0, y_0) \in B_2$ we have $y_0 < (1+k)x_0 - k$ (4.2.10) and $y_0 < x_0$. Hence we have $\delta_2 < 0$ if we choose

$$a^* = -\frac{k(1-x_0)-2(1-y_0)}{k[k(x_0-y_0)+4(1-y_0)]}.$$

The conclusion of the above argument is that there exists a real number a^* such that (4.2.16) has no solution. Hence, there is a $\pi_2(a^*) \in \prod_2$ such that $\pi_2(a^*) \cap S_1 = \{p_1\}$. Note that p_1 is an end point of S_1, and S_1 is located at one side of $\pi_2(a^*)$. It is easy to see that there is a plane π_2^* such that $p \in \pi_2^*$ and $\pi_2^* \cap S_1 = \emptyset$. This implies that $\pi_2^* \cap \hat{S}_1 = \emptyset$, and hence we conclude that $p \notin \hat{S}_1$. □

With the aid of Lemma 4.2.7, we are now able to prove Theorem 4.2.6.

The proof of Theorem 4.2.6:

We are going to prove Theorem 4.2.6 by showing (i) $T_1(\mathcal{F}_1) \subseteq T_1(\mathcal{F})$, (ii) $T_1(\mathcal{F}) \subseteq \hat{S}_1$, and (iii)$\hat{S}_1 \subseteq T_1(\mathcal{F}_1)$.

(i) It is obvious that $T_1(\mathcal{F}_1) \subseteq T_1(\mathcal{F})$, since $\mathcal{F}_1 \subseteq \mathcal{F}$.

(ii) By Lemma 2.2.2, we know that for any $\xi \in \mathcal{F}$ there exists $\xi_n = \sum_{i=1}^{n} \frac{p_{ni}}{2} \Delta_{\pm x_{ni}}$, such that $\xi_n \Rightarrow \xi$.

Let $p_n = (\int_{-1}^{1} x^2 \, d\xi_n, \int_{-1}^{1} x^4 \, d\xi_n, \int_{-1}^{1} x^6 \, d\xi_n)$, and
$p = (\int_{-1}^{1} x^2 \, d\xi, \int_{-1}^{1} x^4 \, d\xi, \int_{-1}^{1} x^6 \, d\xi)$. According to Lemma 2.2.3, we have $p_n \to p$.

$$
\begin{aligned}
\text{Note that } p_n &= (\textstyle\int_{-1}^{1} x^2 \, d\xi_n, \int_{-1}^{1} x^4 \, d\xi_n, \int_{-1}^{1} x^6 \, d\xi_n) \\
&= (\textstyle\sum_{i=1}^{n} p_{ni} x_{ni}^2, \sum_{i=1}^{n} p_{ni} x_{ni}^4, \sum_{i=1}^{n} p_{ni} x_{ni}^6) \\
&= \textstyle\sum_{i=1}^{n} p_{ni} (x_{ni}^2, x_{ni}^4, x_{ni}^6) \in \hat{S}_1,
\end{aligned}
$$

since $(x_{ni}^2, x_{ni}^4, x_{ni}^6) \in S_1$, $i = 1, \ldots, n$. Note also p is a limit point of p_n and $\hat{S}_1$ is a closed set. This implies that $p \in \hat{S}_1$. Hence we have $T_1(\mathcal{F}) \subseteq \hat{S}_1$.

(iii) By Lemma 4.2.7, we have $P = \hat{S}_1$. Hence, for any $p \in \hat{S}_1$, there exists a number $k \in [0,1]$ such that $p \in A_k$. This implies that $p = \alpha p_1 + \beta p_k + (1-\alpha-\beta)p_0$ for some $0 \le \alpha \le 1$, $0 \le \beta \le 1$, and $\alpha + \beta \le 1$, where $p_k \in S_1$.

$$\begin{aligned}
\text{Let } \xi_0 &= \frac{\alpha}{2}\Delta_{\pm 1} + \frac{\beta}{2}\Delta_{\pm\sqrt{k}} + (1-\alpha-\beta)\Delta_0. \text{ Then } \xi_0 \in \mathcal{F}_1, \text{ and} \\
T_1(\xi_0) &= \left(\textstyle\int_{-1}^1 x^2 d\xi_0(x),\ \int_{-1}^1 x^4 d\xi_0(x),\ \int_{-1}^1 x^6 d\xi_0(x)\right) \\
&= (\alpha + \beta k, \alpha + \beta k^2, \alpha + \beta k^3) \\
&= \alpha(1,1,1) + \beta(k, k^2, k^3) + (1-\alpha-\beta)(0,0,0) \\
&= \alpha p_1 + \beta p_k + (1-\alpha-\beta)p_0.
\end{aligned}$$

Hence we have $p = T_1(\xi_0) \in T_1(\mathcal{F}_1)$, and $\hat{S}_1 \subseteq T_1(\mathcal{F}_1)$. □

4.3 Approximately Linear Regression Model: Case I

We are now going to find the $\underset{\sim}{C}^k$-restricted optimal designs for different optimality criteria. In this section, we only consider the case when $m = 1$ and $k = 1$. We simply write "C-restricted" rather than "$\underset{\sim}{C}^1$-restricted", since in this case there is only one component in the "vector" $\underset{\sim}{C}^1$.

According to Lemma 4.2.4, we only need search C-restricted optimal designs within the class of symmetric design measures supported on $[-1, 1]$. We define $\mathcal{F}_s = \{\xi : \xi \in \mathcal{F} \quad \xi(-x) = \xi(x)\}$. For any $\xi \in \mathcal{F}_s$, we have

$$B_1(\xi) = \begin{pmatrix} 1 & 0 \\ 0 & \mu_2 \end{pmatrix} \qquad \text{and} \qquad B_2(\xi) = \begin{pmatrix} 1 & 0 & \mu_2 \\ 0 & \mu_2 & 0 \\ \mu_2 & 0 & \mu_4 \end{pmatrix},$$

where $\mu_2 = \int_{-1}^1 x^2 d\xi$ and $\mu_4 = \int_{-1}^1 x^4 d\xi$. Consequently, we find $|B_1(\xi)| = \mu_2$, $|B_2(\xi)| = \mu_2(\mu_4 - \mu_2^2)$, $trB_1^{-1}(\xi) = 1 + \frac{1}{\mu_2}$, $d_1(x, \xi) = (1, x)B_1^{-1}(\xi)\begin{pmatrix} 1 \\ x \end{pmatrix} = 1 + \frac{x^2}{\mu^2}$. Hence we have $\int_{-1}^1 d_1(x, \xi)dx = 2 + \frac{2}{3\mu_2}$, and $\max_{-1 \le x \le 1} d_1(x, \xi) = 1 + \frac{1}{\mu_2}$. Consider the loss functions $\mathcal{L}_D(\xi) = |B_m^{-1}(\xi)|$, $\mathcal{L}_A(\xi) = trB_m^{-1}(\xi)$, $\mathcal{L}_Q(\xi) = \int_{-1}^1 d_m(x, \xi)dx$, and $\mathcal{L}_G(\xi) = \max_{-1 \le x \le 1} d_m(x, \xi)$. For $m = 1$ and $\mathcal{L}(\xi) \in \{\mathcal{L}_D(\xi),\ \mathcal{L}_A(\xi),\ \mathcal{L}_Q(\xi),\ \mathcal{L}_G(\xi)\}$, it is obvious that

$$\min_{\xi \in \mathcal{F}_s} \mathcal{L}(\xi) \quad \text{subject to} \quad |B_1(\xi)| \le c\,|B_2(\xi)|$$

is equivalent to

$$\max_{\xi \in \mathcal{F}_s} \mu_2 \quad \text{subject to} \quad \frac{1}{c} \le \mu_4 - \mu_2^2.$$

The solution of this problem has been presented in Stigler (1971). Stigler claimed that the restricted optimal designs can be obtained by searching within the class of symmetric designs supported at three points $-1, 0, 1$. This conclusion is based on the well known results we mentioned in Section 4.2. But here we are going to solve the problem based on the theorems we provided in Section 4.2 which will lead to the following theorem, the same result as Stigler's (1971).

Theorem 4.3.1 *For $c \geq 4$, the C-restricted D-, A-, Q-optimal designs over $\mathcal{F}$, and G-optimal design over $\mathcal{F}_s$ for the model P_1 is given by*

$$\xi_0(-1) = \xi_0(1) = \frac{1}{4} + \frac{1}{2}\sqrt{\frac{1}{4} - \frac{1}{c}}, \quad \xi_0(0) = \frac{1}{2} - \sqrt{\frac{1}{4} - \frac{1}{c}}.$$

Proof: (Method 1)

According to Theorem 4.2.5, we have $T_0(\mathcal{F}_0) = \hat{S}_0$, where $\mathcal{F}_0 = \{\xi : \xi = \frac{\alpha}{2}\Delta_{\pm\sqrt{x}} + (1-\alpha)\Delta_0, 0 \leq \alpha \leq 1, 0 \leq x \leq 1\}$, $T_0(\xi) = (\int_{-1}^{1} x^2 d\xi, \int_{-1}^{1} x^4 d\xi) = (\mu_2, \mu_4)$, and $S_0 = \{(\mu_2, \mu_4) : \mu_4 = \mu_2^2, 0 \leq \mu_2 \leq 1\}$. $\hat{S}_0$ is the convex hull of S_0. On the "plane" of $\mu_2 \bigcirc \mu_4$ (imagine μ_2 and μ_4 could be any real numbers), the regions of $\hat{S}_0$ and $|B_1(\xi)| \leq c|B_2(\xi)|$ can be shown by the following Figures:

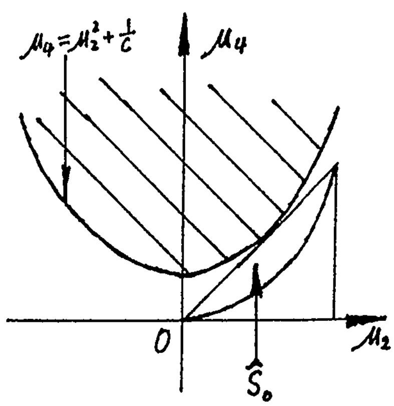

Figure 4.3.1

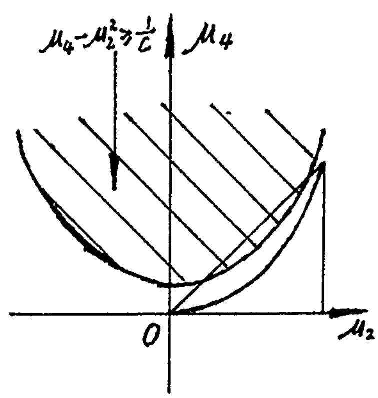

Figure 4.3.2

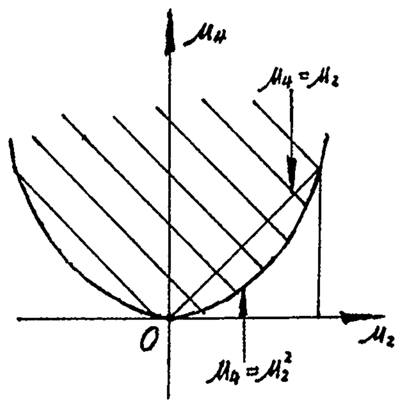

Figure 4.3.3

It is easy to see that the minimum value of c, such that the problem has feasible solutions, corresponds to the unique solution of the system equations of $\mu_4 = \mu_2^2 + \frac{1}{c}$, $\mu_4 = \mu_2$, and $2\mu_2 = 1$. In this case, we get $\mu_4 = \mu_2 = \frac{1}{2}$ and $c = 4$.

For any $4 \leq c < \infty$, the C-restricted optimal design corresponds to the solution of

$$\begin{cases} \mu_4 &= \mu_2^2 + \dfrac{1}{c} \\ \mu_4 &= \mu_2 \end{cases},$$

which yields $\mu_2 = \mu_4 = \frac{1}{2} \pm \sqrt{\frac{1}{4} - \frac{1}{c}}$. We choose $\mu_2 = \mu_4 = \frac{1}{2} + \sqrt{\frac{1}{4} - \frac{1}{c}}$, since we want to maximize μ_2.

On the other hand, any point on the "line" $\mu_4 = \mu_2$ can be realized by a design measure ξ_0 of the form $\xi_0 = \frac{\alpha}{2}\Delta_{\pm 1} + (1-\alpha)\Delta_0$ for some $0 \le \alpha \le 1$. This implies that $\alpha = \int_{-1}^{1} x^2 d\xi_0 = \mu_2 = \frac{1}{2} + \sqrt{\frac{1}{4} - \frac{1}{c}}$. Hence we get $\xi_0 = \left(\frac{1}{4} + \frac{1}{2}\sqrt{\frac{1}{4} - \frac{1}{c}}\right)\Delta_{\pm 1} + \left(\frac{1}{2} - \sqrt{\frac{1}{4} - \frac{1}{c}}\right)\Delta_0$, or we can write $\xi_0(-1) = \xi_0(1) = \frac{1}{4} + \frac{1}{2}\sqrt{\frac{1}{4} - \frac{1}{c}}$ and $\xi_0(0) = \frac{1}{2} - \sqrt{\frac{1}{4} - \frac{1}{c}}$. □

Proof: (Method 2)

Again, by Theorem 4.2.5, we know that

$$\max_{\xi \in \mathcal{F}} \mu_2 \qquad \text{subject to } \frac{1}{c} \le \mu_4 - \mu_2^2$$

is equivalent to

$$\max_{\xi \in \mathcal{F}_0} \mu_2 \qquad \text{subject to } \frac{1}{c} \le \mu_4 - \mu_2^2.$$

For any $\xi \in \mathcal{F}_0$, we have $\xi = \frac{\alpha}{2}\Delta_{\pm\sqrt{x}} + (1-\alpha)\Delta_0$ for some $0 \le \alpha \le 1$ and $0 \le x \le 1$. Hence, we get $\mu_2 = \int_{-1}^{1} x^2 d\xi = \alpha x$ and $\mu_4 = \int_{-1}^{1} x^4 d\xi = \alpha x^2$. The problem now becomes

$$\max\{\alpha x\} \text{ subject to } \frac{1}{c} \le \alpha x^2 - \alpha^2 x^2 \quad \text{for } 0 \le \alpha \le 1 \text{ and } 0 \le x \le 1. \tag{4.3.1}$$

It is easy to see that (4.3.1) is equivalent to

$$\max\{\alpha x\} \text{ subject to } \frac{1}{c} = \alpha x^2 - \alpha^2 x^2 \quad \text{for } 0 \le \alpha \le 1 \text{ and } 0 \le x \le 1. \tag{4.3.2}$$

Let $L(\alpha, x, \lambda) = \alpha x + \lambda[(\alpha - \alpha^2)x^2 - \frac{1}{c}]$, and set

$$\begin{cases} \dfrac{\partial L(\alpha,x,\lambda)}{\partial \alpha} = x + \lambda[(1-2\alpha)x^2] \quad = 0 & (4.3.3) \\ \dfrac{\partial L(\alpha,x,\lambda)}{\partial \beta} = \alpha + \lambda[2x\alpha(1-\alpha)] \quad = 0 & (4.3.4) \\ \dfrac{\partial L(\alpha,x,\lambda)}{\partial \lambda} = \alpha(1-\alpha)x^2 - \dfrac{1}{c} \quad = 0 & (4.3.5) \end{cases}$$

We find that the system of equations (4.3.3), (4.3.4), and (4.3.5) has no solution, since (4.3.3) and (4.3.4) contradict each other.

Consider the boundary cases (i) $\alpha = 0$, (ii) $\alpha = 1$, (iii) $x = 0$, and (iv) $x = 1$. We find that the maximum value will occur in the case when $x = 1$. In this case, we have $L(\alpha, 1, \lambda) = \alpha + \lambda[\alpha - \alpha^2 - \frac{1}{c}]$. Solving

$$\begin{cases} \dfrac{\partial L}{\partial \alpha} &= 1 + \lambda(1 - 2\alpha) &= 0 \\ \dfrac{\partial L}{\partial \lambda} &= \alpha(1-\alpha) - \dfrac{1}{c} &= 0 \end{cases},$$

we get

$$\alpha^2 - \alpha + \frac{1}{c} = 0 \quad \text{and} \quad \alpha = \frac{1}{2} \pm \sqrt{\frac{1}{4} - \frac{1}{c}}.$$

Again, we choose $\alpha = \frac{1}{2} + \sqrt{\frac{1}{4} - \frac{1}{c}}$, since we wnat to maximize αx. Therefore, we get $\xi_0 = \left(\frac{1}{4} + \frac{1}{2}\sqrt{\frac{1}{4} - \frac{1}{c}}\right)\Delta_{\pm 1} + \left(\frac{1}{2} - \sqrt{\frac{1}{4} - \frac{1}{c}}\right)\Delta_0$, i.e. $\xi_0(\pm 1) = \frac{1}{4} + \frac{1}{2}\sqrt{\frac{1}{4} - \frac{1}{c}}$ and $\xi_0(0) = \frac{1}{2} - \sqrt{\frac{1}{4} - \frac{1}{c}}$. □

Remark 1. The design measure ξ_0 which we found in Theorem 4.3.1 is the C-restricted D-, A-, and Q- optimal design over $\mathcal{F}$. This fact follows by Lemma 4.2.4. But the design measure ξ_0 may not be the C-restricted G- optimal design over $\mathcal{F}$, although Stigler (1971) showed the convexity of $G(\underset{\sim}{C}^k)$, where $G(\underset{\sim}{C}^k) = \{\xi : \max_{-1 \le x \le 1} d(x, \xi_0) = \min_{\xi \in V(\underset{\sim}{C}^k)} \max_{-1 \le x \le 1} d(x, \xi_0)\}$. But we do not know whether $\xi(x) \in G(\underset{\sim}{C}^k)$ will imply $\xi(-x) \in G(\underset{\sim}{C}^k)$ or not. To be safe, we only consider ξ_0 to be the C-restricted G-optimal design over $\mathcal{F}_S$.

Remark 2. For $c = \infty$, the design becomes $\xi_0(-1) = \xi_0(1) = \frac{1}{2}$ which is the usual optimal design for the model P_1. When $c = 4$, we get $\xi_0(-1) = \xi_0(1) = \frac{1}{4}$ and $\xi_0(0) = \frac{1}{2}$, which is the best design for estimating β_2 in the model P_2. For $4 < c < \infty$, we get a compromise between these two extreme cases.

The question remains: how should c be chosen? As we mentioned in Chapter 3, the choice of c should reflect both the desire for efficiency of the model P_1 and the wish to check the fit of this model. Stigler (1971) considered the following measures of the efficiency of a design for a model.

Definition 4.3.2 *The model P_m D-efficiency of a design ξ is given by*

$$E_m^D(\xi) = \left(\frac{|B_m(\xi)|}{\max_{\eta \in \mathcal{F}} |B_m(\eta)|} \right)^{1/(m+1)} .$$

The model P_m G-efficiency of a design ξ is given by

$$E_m^G(\xi) = \frac{m+1}{\max_{-1 \le x \le 1} d_m(x,\xi)}.$$

Stigler (1971) provided some tables and pictures to show the relationship between the choice of c and the efficiencies $E_m^D(\xi)$ and $E_m^G(\xi)$. It is easy to extend this definition to the A- and Q- optimal situations.

Definition 4.3.3 *The model P_m A-efficiency of a design ξ is given by*

$$E_m^A(\xi) = \frac{\min_{\eta \in \mathcal{F}} trB_m^{-1}(\eta)}{trB_m^{-1}(\xi)}.$$

The model P_m Q-efficiency of a design ξ is given by

$$E_m^Q(\xi) = \frac{\min_{\eta \in \mathcal{F}} \int_{-1}^{1} d_m(x,\eta)\, dx}{\int_{-1}^{1} d_m(x,\xi)\, dx} .$$

4.4 Approximately Linear Regression Model: Case II

The C-restricted optimal design we discussed in Section 4.3 is useful for determining the presence of a quadratic term in the regression function, but it is no good at all for estimating cubic or higher-order coefficients. The $\underset{\sim}{C}^k$-restricted optimal design is proposed to deal with this situation. In this section, we are going to find the $\underset{\sim}{C}^k$-restricted optimal designs for the model P_m with $m = 1$ and $k = 2$.

As we noted in Section 4.3, for $m = 1$ and $\mathcal{L}(\xi) \in \{\mathcal{L}_D(\xi), \mathcal{L}_A(\xi), \mathcal{L}_Q(\xi), \mathcal{L}_G(\xi)\}$, we have that

$$\min_{\xi\in\mathcal{F}} \mathcal{L}(\xi) \quad \text{subject to} \quad |B_1(\xi)| \le c_1\,|B_2(\xi)| \text{ and } |B_2(\xi)| \le c_2\,|B_3(\xi)|$$

is equivalent to

$$\max_{\xi\in\mathcal{F}} \mu_2 \quad \text{subject to} \quad \frac{1}{c_1} \le \mu_4 - \mu_2^2 \text{ and } \mu_2 \le c_2(\mu_2\mu_6 - \mu_4^2).$$

According to Theorem 4.2.6, the above problem is equivalent to

$$\max_{\xi\in\mathcal{F}_1} \mu_2 \quad \text{subject to} \quad \frac{1}{c_1} \le \mu_4 - \mu_2^2 \text{ and } \mu_2 \le c_2(\mu_2\mu_6 - \mu_4^2),$$

where $\mathcal{F}_1 = \{\xi : \xi = \frac{\alpha}{2}\Delta_{\pm 1} + \frac{\beta}{2}\Delta_{\pm\sqrt{x}} + (1-\alpha-\beta)\Delta_0,\ 0 \le \alpha \le 1, 0 \le \beta \le 1, \alpha+\beta \le 1, 0 \le x \le 1\}$.

For any $\xi \in \mathcal{F}_1$, we have

$$\mu_2 = \int_{-1}^{1} x^2\, d\xi = \alpha + \beta x,\ \mu_4 = \int_{-1}^{1} x^4\, d\xi = \alpha + \beta x^2,\ \mu_6 = \int_{-1}^{1} x^6\, d\xi = \alpha + \beta x^3.$$

Hence, the problem can be further transformed to the following form:

$$\max\,\{\alpha + \beta x\} \quad \text{subject to} \quad \frac{1}{c_1} \le \alpha + \beta x^2 - (\alpha+\beta x)^2 \text{ and } \alpha + \beta x \le c_2\alpha\beta x(1-x)^2$$
$$\text{for } 0 \le \alpha \le 1,\ 0 \le \beta \le 1,\ \alpha + \beta \le 1,\ 0 \le x \le 1.$$

The problem may not have feasible solutions for some values of c_1 and c_2. If we consider the restriction c_1 alone, we know that the lower bound for c_1 is $c_1^* = 4$. Now the question is: What is the lower bound for c_2? The following lemma answers the question.

Lemma 4.4.1 *The lower bound for c_2 is $c_2^* = 16$, and the corresponding design measure is $\xi = \frac{1}{6}\Delta_{\pm 1} + \frac{1}{3}\Delta_{\pm\frac{1}{2}}$.*

Proof: It is easy to see that the lower bound of c_2 ,denoted by c_2^*, should satisfy the following:

$$\frac{1}{c_2^*} = \max_{(\alpha,\beta,x)\in A} \frac{\alpha\beta x(1-x)^2}{\alpha+\beta x},$$

where $A = \{(\alpha,\beta,x) : 0 \le \alpha \le 1,\ 0 \le \beta \le 1,\ \alpha+\beta \le 1,\ 0 \le x \le 1\}$. We denote

$$L(\alpha,\beta,x) = \frac{\alpha\beta x(1-x)^2}{\alpha+\beta x},$$

and we should exclude the case when $\alpha = 0$ and $\beta = 0$, or $\alpha = 0$ and $x = 0$ in which our objective function $L(\alpha,\beta,x)$ will approach zero and the corresponding $c_2 = \infty$. Solving

$$\begin{cases} \dfrac{\partial L}{\partial \alpha} &= \dfrac{\beta x(1-x)^2}{(\alpha+\beta x)^2}[(\alpha+\beta x)-\alpha] = \dfrac{\beta^2x^2(1-x)^2}{(\alpha+\beta x)^2} = 0 \\ \dfrac{\partial L}{\partial \beta} &= \dfrac{\alpha x(1-x)^2}{(\alpha+\beta x)^2}[(\alpha+\beta x)-\beta x] = \dfrac{\alpha^2x(1-x)^2}{(\alpha+\beta x)^2} = 0 \\ \dfrac{\partial L}{\partial x} &= \dfrac{\alpha\beta}{(\alpha+\beta x)^2}[(\alpha+\beta x)(1-4x+3x^2) - x(1-x)^2\beta] \\ &= \dfrac{\alpha\beta(1-x)}{(\alpha+\beta x)^2}(\alpha-3\alpha x-2\beta x^2) \ = \ 0 \end{cases},$$

we find that there is no solution inside A. Hence, the extremum points must occur on the boundary of A. It is obvious that the boundary cases (i) $\alpha = 0$, (ii) $\beta = 0$, (iii) $x = 0$, and (iv) $x = 1$ are corresponding to the minimum values of $L(\alpha,\beta,x)$. We only need to consider the situation when $\alpha+\beta = 1$. In this case, we have

$$L_0(1-\beta,\beta,x) = \frac{\beta(1-\beta)x(1-x)^2}{(1-\beta+\beta x)},$$

and

$$\begin{cases} \dfrac{\partial L_0}{\partial \beta} &= \dfrac{x(1-x)^2}{(1-\beta+\beta x)^2}[(1-\beta+\beta x)(1-2\beta)-\beta(1-\beta)(-1+\mu)] \\ &= \dfrac{x(1-x)^2}{(1-\beta+\beta x)^2}(1-2\beta+\beta^2-\beta^2 x) \\ \dfrac{\partial L_0}{\partial x} &= \dfrac{\beta(1-\beta)}{(1-\beta+\beta x)^2}[(1-\beta+\beta x)(1-4x+3x^2)-x(1-x)^2\beta] \\ &= \dfrac{\beta(1-\beta)(1-x)}{(1-\beta+\beta x)^2}(1-3x-\beta+3\beta x-2\beta x^2) \end{cases}.$$

We exclude the cases of $\beta = 0$, $\beta = 1$, $x = 0$, or $x = 1$ which will cause $\alpha = 0$. By solving

$$\begin{cases} 1-2\beta+\beta^2-\beta^2 x &= 0 \qquad (4.4.1) \\ 1-3x-\beta-3\beta x-2\beta x^2 &= 0 \quad , \qquad (4.4.2) \end{cases}$$

we find $\beta = 1$, $x = 0$, or $\beta = \frac{2}{3}$, $x = \frac{1}{4}$. We choose the solution $\beta = \frac{2}{3}$, $x = \frac{1}{4}$ which yields $\alpha = \frac{1}{3}$ and

$$L(\frac{1}{3}, \frac{2}{3}, \frac{1}{4}) = \frac{1}{16} = \max_{(\alpha,\beta,x)\in A} L(\alpha, \beta, x) = \frac{1}{c_2^*}.$$

Hence, we get $c_2^* = 16$ and the corresponding design measure is $\xi = \frac{\alpha}{2}\Delta_{\pm 1} + \frac{\beta}{2}\Delta_{\pm\sqrt{x}} + (1-\alpha-\beta)\Delta_0 = \frac{1}{6}\Delta_{\pm 1} + \frac{1}{3}\Delta_{\pm\frac{1}{2}}$. □

Remark 1. It is interesting to note that the design measure ξ corresponding to $c_1 = 4$ (the lower bound of c_1) corresponds to $c_2 = \infty$ (the upper bound of c_2). But the design measure ξ corresponding to $c_2 = 16$ (the lower bound of c_2) corresponds to $c_1 = 8$ (not the upper bound of c_1). This fact implies that the design measure ξ which provides the largest opportunity for determining the presence of a quadratic term in the regression function is no good for estimating cubic coefficients. But on the other hand, the design measure ξ which provides the largest opportunity for determining the presence of a cubic term still allows us to estimate the quadratic term with some degree of precision.

There are two special cases linked with the original problem. One considers the restriction c_1 only and the other c_2 only, i.e.

$$\text{(i)} \quad \max_{\xi\in\mathcal{F}} \mu_2 \quad \text{subject to} \quad |B_1(\xi)| \le c_1 |B_2(\xi)|,$$

and

$$\text{(ii)} \quad \max_{\xi \in \mathcal{F}} \mu_2 \quad \text{subject to} \quad |B_2(\xi)| \le c_2 |B_3(\xi)| .$$

Case (i) has already been studied in Section 4.3. For case (ii), we first point out the following fact:

Lemma 4.4.2 *The problem of* $\max_{\xi \in \mathcal{F}} \mu_2$ *subject to* $|B_2(\xi)| \le c_2 |B_3(\xi)|$ *is equivalent to the problem of* $\max_{\xi \in \mathcal{F}} \mu_2$ *subject to* $|B_2(\xi)| = c_2 |B_3(\xi)|$.

Proof: The constraint $|B_2(\xi)| \le c_2 |B_3(\xi)|$ is the same as $\mu_2 \le c_2(\mu_2\mu_6 - \mu_4^2)$ or $\mu_6 \ge \frac{\mu_4^2}{\mu_2} + \frac{1}{c_2}$, if we exclude the case $\mu_2 = 0$. For convenience, we use (x, y, z) instead of (μ_2, μ_4, μ_6). For any fixed $x \in [0, 1]$, we denote $D(c_2) = \{(x, y, z) : z \ge \frac{y^2}{x} + \frac{1}{c_2}\}$ and $E(c_2) = \{(x, y, z) : z = \frac{y^2}{x} + \frac{1}{c_2}\}$. Let $\pi_k : z = -kx + (1 + k)y$ and $\hat{S}_1 = P = \{A_k : A_k$ is the interior and boundary of the triangle with vertices $p_0, p_1, p_k,\ 0 \le k \le 1\}$, where $p_0 = (0, 0, 0)$, $p_1 = (1, 1, 1)$, and $p_k = (k, k^2, k^3)$ (See Section 4.2). It is obvious that $\hat{S}_1 \subseteq \{\pi_k : 0 \le k \le 1\}$. In order to get $D(c_2) \cap \hat{S}_1 \ne \emptyset$, it is necessary that

$$\frac{y^2}{x} + \frac{1}{c_2} \le -kx + (1 + k)y \tag{4.4.3}$$

for some $0 \le k \le 1$ and some $(x, y) \in S^* = \{(x, y) : 0 \le x \le 1,\ x^2 \le y \le x\}$. The inequality (4.4.3) can also be written as

$$y^2 - (1 + k)xy + (kx^2 + \frac{x}{c_2}) \le 0. \tag{4.4.4}$$

Solving

$$y^2 - (1 + k)xy + (kx^2 + \frac{x}{c_2}) = 0, \tag{4.4.5}$$

we get

$$y = \frac{(1 + k)x \pm \sqrt{(1 + k)^2x^2 - 4(kx^2 + \frac{x}{c_2})}}{2} = \frac{(1 + k)x \pm \sqrt{(1 - k)^2x^2 - \frac{4x}{c_2}}}{2}.$$

Hence, (4.4.3) will hold if $(1 - k)^2x^2 - \frac{4x}{c_2} \ge 0$. In this case we have

$$\frac{(1 + k)x - \sqrt{(1 - k)^2x^2 - \frac{4x}{c_2}}}{2} \le y \le \frac{(1 + k)x + \sqrt{(1 - k)^2x^2 - \frac{4x}{c_2}}}{2}$$

$$\leq \frac{(1+k)x + \sqrt{(1-k)^2x^2}}{2} = \frac{(1+k)x + (1-k)x}{2} = x \,.$$

It is not hard to see that $D(c_2) \cap \hat{S}_1 \neq \emptyset$ implies $E(c_2) \cap \hat{S}_1 \neq \emptyset$, since we have $y \leq x$.

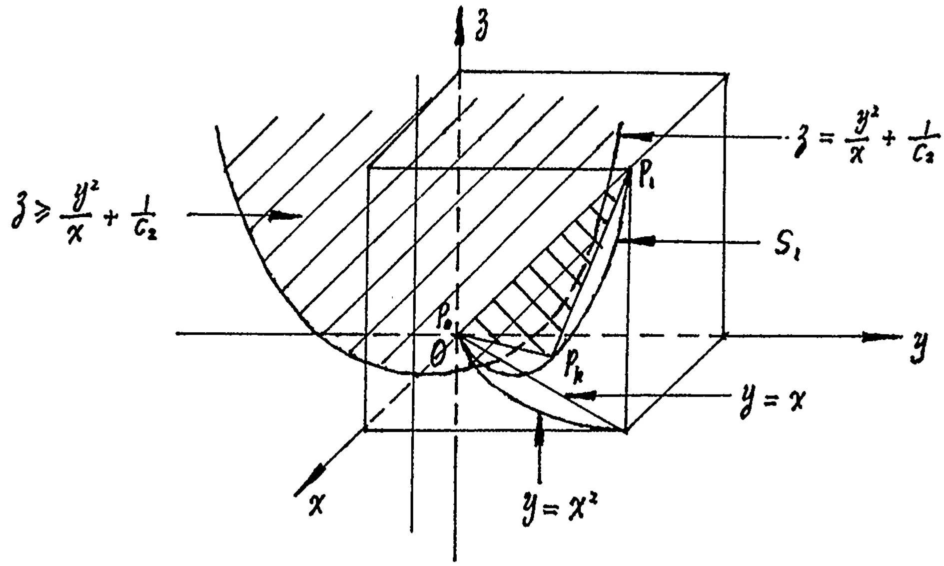

Figure 4.4.1

Note also our objective function is simply μ_2. Hence, it is sufficient for us to search for the maximum μ_2 on the boundary of the constraint which is $|B_2(\xi)| = c_2\,|B_3(\xi)|$.□

According to Lemma 4.4.2, we can use Lagrange's method of multipliers to solve the following problem:

$$\max_{\xi \in \mathcal{F}} \mu_2 \quad \text{subject to} \quad \mu_2 = c_2(\mu_2\mu_6 - \mu_4^2),$$

or

$$\max_{(\alpha,\beta,x)\in A} \{\alpha + \beta x\} \quad \text{subject to} \quad \alpha + \beta x = c_2\alpha\beta x(1-x)^2.$$

We define $L(\alpha, \beta, x, \lambda) = \alpha + \beta x + \lambda[\alpha + \beta x - c_2\alpha\beta x(1-x)^2]$. By solving

$$\begin{cases} \dfrac{\partial L}{\partial \alpha} = 1+\lambda[1-c_2\beta x(1-x)^2] = 0 \\ \dfrac{\partial L}{\partial \beta} = x+\lambda[x-c_2\alpha x(1-x)^2] = 0 \\ \dfrac{\partial L}{\partial x} = \beta+\lambda[\beta-c_2\alpha\beta(1-x)(1-3x)] = 0 \\ \dfrac{\partial L}{\partial \lambda} = \alpha+\beta x-c_2\alpha\beta x(1-x)^2 = 0 \end{cases},$$

we find that the above system of equations has no solution inside A. There are five boundary cases to be considered. They are (i) $\alpha = 0$, (ii) $\beta = 0$, (iii) $x = 0$, (iv) $x = 1$ and (v) $\alpha+\beta = 1$. We find that the extremum points should occur on the boundary $\alpha+\beta=1$. For $\alpha+\beta=1$, we define $L_0(1-\beta,\beta,x,\lambda) = 1-\beta+\beta x+\lambda[1-\beta+\beta x-c_2\beta(1-\beta)x(1-x)^2]$, and let

$$\begin{cases} \dfrac{\partial L_0}{\partial \beta} = -1+x+\lambda[-1+x-c_2(1-2\beta)x(1-x)^2] = 0 & (4.4.6) \\ \dfrac{\partial L_0}{\partial x} = \beta+\lambda[\beta-c_2\beta(1-\beta)(1-x)(1-3x)] = 0 & (4.4.7) \\ \dfrac{\partial L_0}{\partial \lambda} = 1-\beta+\beta x-c_2\beta(1-\beta)x(1-x)^2 = 0 \ . & (4.4.8) \end{cases}$$

Combining (4.4.6) and (4.4.7) gives us

$$\frac{-1}{1-c_2(1-\beta)(1-x)(1-3x)} = \frac{1-x}{-1+x-c_2(1-2\beta)x(1-x)^2},$$

and hence we have

$$x = \frac{1-\beta}{2-\beta}. \tag{4.4.9}$$

Put (4.4.9) into (4.4.8), we get

$$(1-\beta)+\beta\cdot\frac{(1-\beta)}{(2-\beta)} - c_2\beta(1-\beta)\cdot\frac{(1-\beta)}{(2-\beta)} = 0$$

which can be simplified as the following

$$(2+c_2)\beta^2-(8+c_2)\beta+8 = 0 \ . \tag{4.4.10}$$

Solving (4.4.10), we find

$$\beta = \frac{8+c_2\pm\sqrt{(8+c_2)^2-32(2+c_2)}}{2(2+c_2)} = \frac{8+c_2\pm\sqrt{c_2^2-16c_2}}{2(2+c_2)}.$$

For $\beta_1 = \dfrac{8 + c_2 + \sqrt{c_2^2 - 16c_2}}{(4 + 2c_2)}$, we get

$$\alpha_1 = 1 - \beta_1 = 1 - \frac{8 + c_2 + \sqrt{c_2^2 - 16c_2}}{4 + 2c_2} = \frac{4 + 2c_2 - (8 + c_2 + \sqrt{c_2^2 - 16c_2})}{4 + 2c_2}$$
$$= \frac{c_2 - 4 - \sqrt{c_2^2 - 16c_2}}{4 + 2c_2},$$

and

$$x_1 = \frac{1 - \beta_1}{2 - \beta_1} = \frac{1 - \frac{8+c_2+\sqrt{c_2^2-16c_2}}{(4+2c_2)}}{2 - \frac{8+c_2+\sqrt{c_2^2-16c_2}}{(4+2c_2)}} = \frac{c_2 - 4 - \sqrt{c_2^2 - 16c_2}}{3c_2 - \sqrt{c_2^2 - 16c_2}}$$
$$= \frac{(c_2 - 4 - \sqrt{c_2^2 - 16c_2})(3c_2 + \sqrt{c_2^2 - 16c_2})}{9c_2^2 - (c_2^2 - 16c_2)}$$
$$= \frac{2c_2^2 + 4c_2 - (2c_2 + 4)\sqrt{c_2^2 - 16c_2}}{8c_2^2 + 16c_2} = \frac{c_2 - \sqrt{c_2^2 - 16c_2}}{4c_2}.$$

Hence, we have

$$\mu_2^{(1)} = \alpha_1 + \beta_1 x_1 = \frac{c_2 - 4 - \sqrt{c_2^2 - 16c_2}}{4 + 2c_2} + \frac{8 + c_2 + \sqrt{c_2^2 - 16c_2}}{4 + 2c_2} \cdot \frac{c_2 - \sqrt{c_2^2 - 16c_2}}{4c_2}$$
$$= \frac{4c_2(c_2 - 4 - \sqrt{c_2^2 - 16c_2}) + (8 + c_2 + \sqrt{c_2^2 - 16c_2})(c_2 - \sqrt{c_2^2 - 16c_2})}{4c_2(4 + 2c_2)}$$
$$= \frac{4c_2^2 + 8c_2 - (4c_2 + 8)\sqrt{c_2^2 - 16c_2}}{8c_2(2 + c_2)} = \frac{c_2 - \sqrt{c_2^2 - 16c_2}}{2c_2}.$$

Similarly, for $\beta_2 = \dfrac{8 + c_2 - \sqrt{c_2^2 - 16c_2}}{4 + 2c_2}$, we get

$$\alpha_2 = 1 - \beta_2 = 1 - \frac{8 + c_2 - \sqrt{c_2^2 - 16c_2}}{4 + 2c_2} = \frac{4 + 2c_2 - (8 + c_2 - \sqrt{c_2^2 - 16c_2})}{4 + 2c_2}$$
$$= \frac{c_2 - 4 + \sqrt{c_2^2 - 16c_2}}{4 + 2c_2},$$

and

$$x_2 = \frac{1 - \beta_2}{2 - \beta_2} = \frac{1 - \frac{8+c_2-\sqrt{c_2^2-16c_2}}{4+2c_2}}{2 - \frac{8+c_2-\sqrt{c_2^2-16c_2}}{4+2c_2}} = \frac{c_2 - 4 + \sqrt{c_2^2 - 16c_2}}{3c_2 + \sqrt{c_2^2 - 16c_2}}$$

$$= \frac{(c_2 - 4 + \sqrt{c_2^2 - 16c_2})(3c_2 - \sqrt{c_2^2 - 16c_2})}{9c_2^2 - (c_2^2 - 16c_2)}$$

$$= \frac{2c_2^2 + 4c_2 + (2c_2 + 4)\sqrt{c_2^2 - 16c_2}}{8c_2^2 + 16c_2} = \frac{c_2 + \sqrt{c_2^2 - 16c_2}}{4c_2}.$$

Hence, we have

$$\mu_2^{(2)} = \alpha_2 + \beta_2 x_2 = \frac{c_2 - 4 + \sqrt{c_2^2 - 16c_2}}{4 + 2c_2} + \frac{8 + c_2 - \sqrt{c_2^2 - 16c_2}}{4 + 2c_2} \cdot \frac{c_2 + \sqrt{c_2^2 - 16c_2}}{4c_2}$$

$$= \frac{4c_2(c_2 - 4 + \sqrt{c_2^2 - 16c_2}) + (8 + c_2 - \sqrt{c_2^2 - 16c_2})(c_2 + \sqrt{c_2^2 - 16c_2})}{4c_2(4 + 2c_2)}$$

$$= \frac{4c_2^2 + 8c_2 + (4c_2 + 8)\sqrt{c_2^2 - 16c_2}}{8c_2(c_2 + 2)} = \frac{c_2 + \sqrt{c_2^2 - 16c_2}}{2c_2}.$$

For $c_2 \geq 16$, we always have $\mu_2^{(2)} \geq \mu_2^{(1)}$. Hence we choose

$$\alpha = \frac{c_2 - 4 + \sqrt{c_2^2 - 16c_2}}{4 + 2c_2}, \quad \beta = \frac{8 + c_2 - \sqrt{c_2^2 - 16c_2}}{4 + 2c_2}, \text{ and } x = \frac{c_2 + \sqrt{c_2^2 - 16c_2}}{4c_2}.$$

Therefore, the restricted optimal design is given by

$$\xi_0(\pm 1) = \frac{c_2 - 4 + \sqrt{c_2^2 - 16c_2}}{4c_2 + 8} \text{ and } \xi_0\left(\pm \frac{\sqrt{c_2 + \sqrt{c_2^2 - 16c_2}}}{2\sqrt{c_2}}\right) = \frac{8 + c_2 - \sqrt{c_2^2 - 16c_2}}{4c_2 + 8}.$$

When $c_2 = 16$, we have $\alpha = \frac{1}{3}$, $\beta = \frac{2}{3}$, $x = \frac{1}{4}$, and $\xi_0 = \frac{1}{6}\Delta_{\pm 1} + \frac{1}{3}\Delta_{\pm\frac{1}{2}}$. When $c_2 = \infty$, we have $\alpha = 1$, $\beta = 0$, $x = \frac{1}{2}$, and $\xi_0 = \frac{1}{2}\Delta_{+1} + \frac{1}{2}\Delta_{-1}$.

We have proved the following:

Theorem 4.4.3 *For any* $16 \leq c_2 \leq \infty$*, the optimal design for*

$$\max_{\xi \in \mathcal{F}} \mu_2 \quad \textit{subject to} \quad |B_2(\xi)| \leq c_2 \, |B_3(\xi)|$$

is given by

$$\xi_0(\pm 1) = \frac{c_2 - 4 + \sqrt{c_2^2 - 16c_2}}{4c_2 + 8} \textit{ and } \xi_0\left(\pm \frac{\sqrt{c_2 + \sqrt{c_2^2 - 16c_2}}}{2\sqrt{c_2}}\right) = \frac{8 + c_2 - \sqrt{c_2^2 - 16c_2}}{4c_2 + 8}.$$

Remark 2. According to Theorem 4.4.3, the lower bound of c_2 can be spotted right away from the form of ξ_0. It seems as if Lemma 4.4.1 is not necessary. However, the restricted optimal designs sometime cannot be solved explicitly. In this case, the proof of Lemma 4.4.1 provides an alternative way to find the lower bounds for c_i's without solving the restricted optimal designs.

In general, we can only find the $\underset{\sim}{C}^2$-restricted optimal designs numerically by

$$\max_{(\alpha,\beta,x)\in A}\{\alpha+\beta x\} \quad \text{subject to} \quad \frac{1}{c_1}\le \alpha+\beta x^2-(\alpha+\beta x)^2 \text{ and } \alpha+\beta x\le c_2\alpha\beta x(1-x)^2$$

for any $c_1\in[4,\infty)$ and $c_2\in[16,\infty)$. However, when $c_1$and c_2 have some special relationship, we are still able to solve the problem explicitly. For any $4\le c_1\le\infty$ and $16\le c_2\le\infty$, we define $\mathcal{F}(c_1,c_2)=\{\xi:\xi\in\mathcal{F}_s,\ |B_1(\xi)|\le c_1|B_2(\xi)|,\ |B_2(\xi)|\le c_2|B_3(\xi)|\}$, and let $\xi_0(c_1,c_2)$ be the design measure such that $\mu_2(\xi_0(c_1,c_2)) = \max_{\xi\in\mathcal{F}(c_1,c_2)}\mu_2(\xi)$. Then we have the following:

Lemma 4.4.4 *Let c_1 and c_2 be any real numbers such that $4\le c_1\le\infty$ and $16\le c_2\le\infty$. If $c_1\ge\frac{c_2}{2}$, then $\xi_0(\infty,c_2)\in\mathcal{F}(c_1,c_2)$.*

Proof: We know that $\xi_0(\infty,c_2)$ is the design measure such that $\mu_2(\xi(\infty,c_2)) = \max_{\xi\in\mathcal{F}(\infty,c_2)}\mu_2(\xi)$. According to Theorem 4.4.3, we have $\xi_0(\infty,c_2)=\frac{\alpha}{2}\Delta_{\pm1}+\frac{\beta}{2}\Delta_{\pm\sqrt{x}}$, where

$$\alpha=\frac{c_2-4+\sqrt{c_2^2-16c_2}}{4+2c_2},\ \beta=\frac{8+c_2-\sqrt{c_2^2-16c_2}}{4+2c_2}, \text{and } x=\frac{c_2+\sqrt{c_2^2-16c_2}}{4c_2}.$$

In order to prove $\xi_0(\infty,c_2)\in\mathcal{F}(c_1,c_2)$, it is sufficient to show $|B_1(\xi_0(\infty,c_2))|\le c_1|B_2(\xi_0(\infty,c_2))|$. We find

$$\mu_2(\xi_0(\infty,c_2))=\alpha+\beta x=\frac{c_2-4+\sqrt{c_2^2-16c_2}}{4+2c_2}$$
$$+\frac{8+c_2-\sqrt{c_2^2-16c_2}}{4+2c_2}\cdot\frac{c_2+\sqrt{c_2^2-16c_2}}{4c_2}=\frac{c_2+\sqrt{c_2^2-16c_2}}{2c_2},$$

and

$$\mu_4(\xi_0(\infty,c_2))=\alpha+\beta x^2=\frac{c_2-4+\sqrt{c_2^2-16c_2}}{4+2c_2}$$

$$+\frac{8+c_2-\sqrt{c_2^2-16c_2}}{4+2c_2}\cdot\left(\frac{c_2+\sqrt{c_2^2-16c_2}}{4c_2}\right)^2=\frac{c_2-4+\sqrt{c_2^2-16c_2}}{2c_2}.$$

Hence, we get

$$\mu_4(\xi_0(\infty,c_2))-\mu_2^2(\xi_0(\infty,c_2))=\frac{c_2-4+\sqrt{c_2^2-16c_2}}{2c_2}-\left(\frac{c_2+\sqrt{c_2^2-16c_2}}{2c_2}\right)^2=\frac{2}{c_2},$$

and

$$\frac{|B_1(\xi_0)|}{|B_2(\xi_0)|}=\frac{1}{\mu_4(\xi_0(\infty,c_2))-\mu_2^2(\xi_0(\infty,c_2))}=\frac{c_2}{2}\leq c_1.$$

Therefore, we have proved $\xi_0(\infty,c_2)\in\mathcal{F}(c_1,c_2)$. □

Lemma 4.4.4 plays an important role in the proof of the next theorem.

Theorem 4.4.5 *Let c_1 and c_2 be any real numbers such that $4\leq c_1\leq\infty$ and $16\leq c_2\leq\infty$.*

(i) If $c_1\geq\frac{c_2}{2}$, then we have $\mu_2(\xi_0(\infty,c_2))=\max_{\xi\in\mathcal{F}(c_1,c_2)}\mu_2(\xi)$.

(ii) If $c_1<\frac{c_2}{2}$, then we have

$$\frac{1}{2}+\sqrt{\frac{1}{4}-\frac{2}{c_1}}\leq\max_{\xi\in\mathcal{F}(c_1,c_2)}\mu_2(\xi)\leq\begin{cases}\frac{1}{2}+\sqrt{\frac{1}{4}-\frac{4}{c_2}} & \text{if } 2c_1<c_2\leq 4c_1\\ \frac{1}{2}+\sqrt{\frac{1}{4}-\frac{1}{c_1}} & \text{if } c_2>4c_1\end{cases}.$$

Proof: (i) For any $c_1\in[\frac{c_2}{2},\infty)$, it is clear that $\mathcal{F}(\frac{c_2}{2},c_2)\subseteq\mathcal{F}(c_1,c_2)\subseteq\mathcal{F}(\infty.c_2)$. By Lemma 4.4.4, we have $\xi_0(\infty,c_2)\in\mathcal{F}(\frac{c_2}{2},c_2)$. Combining these facts along with the definition of $\xi_0(\infty,c_2)$, we have

$$\max_{\xi\in\mathcal{F}(\frac{c_2}{2},c_2)}\mu_2(\xi)\leq\max_{\xi\in\mathcal{F}(c_1,c_2)}\mu_2(\xi)\leq\max_{\xi\in\mathcal{F}(\infty,c_2)}\mu_2(\xi)=\mu_2(\xi_0(\infty,c_2))\leq\max_{\xi\in\mathcal{F}(\frac{c_2}{2},c_2)}\mu_2(\xi).$$

Hence we have shown $\mu_2(\xi_0(\infty,c_2))=\max_{\xi\in\mathcal{F}(c_1,c_2)}\mu_2(\xi)$.

(ii) For any $c_1\in[4,\frac{c_2}{2})$, we have $2c_1<c_2$ and $\mathcal{F}(c_1,2c_1)\subseteq\mathcal{F}(c_1,c_2)$ which implies that $\max_{\xi\in\mathcal{F}(c_1,2c_1)}\mu_2(\xi)\leq\max_{\xi\in\mathcal{F}(c_1,c_2)}\mu_2(\xi)$. According to part (i) of Theorem 4.4.5 and Theorem 4.4.3, we have

$$\max_{\xi\in\mathcal{F}(c_1,2c_1)}\mu_2(\xi)=\mu_2(\xi_0(\infty,2c_1))=\frac{2c_1+\sqrt{(2c_1)^2-16(2c_1)}}{2(2c_1)}$$

$$= \frac{1}{2} + \sqrt{\frac{1}{4} - \frac{2}{c_1}}.$$

Hence, we get $\frac{1}{2} + \sqrt{\frac{1}{4} - \frac{2}{c_1}} \le \max_{\xi \in \mathcal{F}(c_1,c_2)} \mu_2(\xi)$.

On the other hand, we have (1) $\mathcal{F}(c_1,c_2) \subseteq \mathcal{F}(\infty,c_2)$, and (2) $\mathcal{F}(c_1,c_2) \subseteq \mathcal{F}(c_1,\infty)$. By (1), we have

$$\max_{\xi \in \mathcal{F}(c_1,c_2)} \mu_2(\xi) \le \max_{\xi \in \mathcal{F}(\infty,c_2)} \mu_2(\xi) = \mu_2(\xi_0(\infty,c_2)) = \frac{c_2 + \sqrt{c_2^2 - 16c_2}}{2c_2}$$
$$= \frac{1}{2} + \sqrt{\frac{1}{4} - \frac{4}{c_2}}.$$

By (2), we have

$$\max_{\xi \in \mathcal{F}(c_1,c_2)} \mu_2(\xi) \le \max_{\xi \in \mathcal{F}(c_1,\infty)} \mu_2(\xi) = \mu_2(\xi_0(c_1,\infty)) = \frac{1}{2} + \sqrt{\frac{1}{4} - \frac{1}{c_1}}.$$

It is clear that $\frac{1}{2} + \sqrt{\frac{1}{4} - \frac{4}{c_2}} \le \frac{1}{2} + \sqrt{\frac{1}{4} - \frac{1}{c_1}}$ when $2c_1 < c_2 \le 4c_1$ and $\frac{1}{2} + \sqrt{\frac{1}{4} - \frac{4}{c_2}} > \frac{1}{2} + \sqrt{\frac{1}{4} - \frac{1}{c_1}}$ when $c_2 > 4c_1$. Hence, we have shown

$$\max_{\xi \in \mathcal{F}(c_1,c_2)} \mu_2(\xi) \le \begin{cases} \frac{1}{2} + \sqrt{\frac{1}{4} - \frac{4}{c_2}} & \text{if } 2c_1 < c_2 \le 4c_1 \\ \frac{1}{2} + \sqrt{\frac{1}{4} - \frac{1}{c_1}} & \text{if } c_2 > 4c_1 \end{cases} . \qquad \square$$

Remark 3. In fact, part (i) of Theorem 4.4.5 tells us $\mu_2(\xi_0(\infty,c_2)) = \mu_2(\xi_0(c_1,c_2))$ for any $c_1 \ge \frac{c_2}{2}$. When $c_1 < \frac{c_2}{2}$, part (ii) of Theorem 4.4.5 gives us a range for $\mu_2(\xi_0(c_1,c_2))$.

Again, the choice of c_i in the $\underset{\sim}{C}^k$-restricted optimal designs should reflect both the desire for efficiency for the model P_m and the wish to check the fit of this model.

Let $\xi_P(c_i)$ be a design such that

$$\mathcal{L}_P(\xi_P(c_i)) = \min_{\xi \in \mathcal{F}} \mathcal{L}_P(\xi) \quad \text{subject to} \quad |B_{m+i-1}(\xi)| \le c_i\,|B_{m+i}(\xi)|$$
$$\text{for some } i \in \{1,\ldots,k\},$$

where $P \in \{D, A, Q, G\}$.

We propose the following:

Definition 4.4.6 *The ith-relative D-efficiency of the model P_m is given by*

$$RE_m^D(\xi_D(c_i)) = \left(\frac{|B_m(\xi_D(c_i))|}{\max\limits_{\eta \in \mathcal{F}} |B_m(\eta)|} \right)^{1/(m+1)} ;$$

The ith-relative A-efficiency of the model P_m is given by

$$RE_m^A(\xi_A(c_i)) = \frac{\min\limits_{\eta \in \mathcal{F}} trB_m^{-1}(\eta)}{trB_m^{-1}(\xi_A(c_i))};$$

The ith-relative Q-efficiency of the model P_m is given by

$$RE_m^Q(\xi_Q(c_i)) = \frac{\min\limits_{\eta \in \mathcal{F}} \int_{-1}^{1} d_m(x,\eta)\, dx}{\int_{-1}^{1} d_m(x,\xi_Q(c_i))\, dx};$$

The ith-relative G-efficiency of the model P_m is given by

$$RE_m^G(\xi_G(c_i)) = \frac{m+1}{\max\limits_{-1 \le x \le 1} d_m(x,\xi_G(c_i))},$$

where $i = 1, \ldots, k$.

For the $\underset{\sim}{C}^k$-restricted optimal designs, the efficiency defined by Definition 4.4.6 should serve the same purpose as Definition 4.3.2 and Definition 4.3.3 did for the C-restricted optimal designs.

4.5 Approximately Quadratic Regression Model

In this section, we briefly discuss the optimal designs for the approximately quadratic polynomial regression model. We consider the following problem:

$$\min_{\xi\in\mathcal{F}} \mathcal{L}(\xi) \quad \text{subject to} \quad |B_2(\xi)| \le c\,|B_3(\xi)|\,, \tag{4.5.1}$$

where $\mathcal{L}(\xi) \in \{\mathcal{L}_D(\xi), \mathcal{L}_A(\xi), \mathcal{L}_Q(\xi)\}$.

For any $\xi \in \mathcal{F}_s$, we have

$$B_2(\xi) = \begin{pmatrix} 1 & 0 & \mu_2 \\ 0 & \mu_2 & 0 \\ \mu_2 & 0 & \mu_4 \end{pmatrix}, \quad B_2^{-1}(\xi) = \begin{pmatrix} \frac{\mu_4}{\mu_4-\mu_2^2} & 0 & \frac{-\mu_2}{\mu_4-\mu_2^2} \\ 0 & \frac{1}{\mu_2} & 0 \\ \frac{-\mu_2}{\mu_4-\mu_2^2} & 0 & \frac{1}{\mu_4-\mu_2^2} \end{pmatrix},$$

$$B_3(\xi) = \begin{pmatrix} 1 & 0 & \mu_2 & 0 \\ 0 & \mu_2 & 0 & \mu_4 \\ \mu_2 & 0 & \mu_4 & 0 \\ 0 & \mu_4 & 0 & \mu_6 \end{pmatrix},$$

and

$$d(x,\xi) = (1,x,x^2)B_2^{-1}(\xi)\begin{pmatrix} 1 \\ x \\ x^2 \end{pmatrix} = \frac{\mu_4}{\mu_4-\mu_2^2} + \left(\frac{1}{\mu_2} - \frac{2\mu_2}{\mu_4-\mu_2^2}\right)x^2 + \frac{1}{\mu_4-\mu_2^2}x^4.$$

According to Lemma 4.4.1, we know that the lower bound for c is $c^* = 16$. Hence (4.5.1) has feasible solutions when $c \in [16,\infty)$.

For D-optimality, (4.5.1) becomes

$$\max_{\xi\in\mathcal{F}_s} |B_2(\xi)| \ \left(\text{or } \min_{\xi\in\mathcal{F}_s} \left|B_2^{-1}(\xi)\right|\right) \quad \text{subject to} \quad |B_2(\xi)| \le c\,|B_3(\xi)|\,,$$

which is the same as

$$\max_{\xi\in\mathcal{F}_s} \{\mu_2(\mu_4-\mu_2^2)\} \quad \text{subject to} \quad \mu_2 \le c(\mu_2\mu_6-\mu_4^2). \tag{4.5.2}$$

For A-optimality, (4.5.1) becomes

$$\min_{\xi\in\mathcal{F}_s} trB_2^{-1}(\xi) \quad \text{subject to} \quad |B_2(\xi)| \le c|B_3(\xi)| \ .$$

i.e.

$$\min_{\xi\in\mathcal{F}_s} \left\{\frac{1+\mu_4}{\mu_4-\mu_2^2}+\frac{1}{\mu_2}\right\} \quad \text{subject to} \quad \mu_2 \le c(\mu_2\mu_6-\mu_4^2) \tag{4.5.3}$$

For Q-optimality, we have

$$\begin{aligned}\int_{-1}^{1} d(x,\xi)\,dx &= \int_{-1}^{1}\left[\frac{\mu_4}{\mu_4-\mu_2^2}+\left(\frac{1}{\mu_2}-\frac{2\mu_2}{\mu_4-\mu_2^2}\right)x^2+\frac{1}{\mu_4-\mu_2^2}x^4\right]dx \\ &= 2\left[\frac{\mu_4}{\mu_4-\mu_2^2}+\frac{1}{3}\left(\frac{1}{\mu_2}-\frac{2\mu_2}{\mu_4-\mu_2^2}\right)+\frac{1}{5}\cdot\frac{1}{\mu_4-\mu_2^2}\right] \\ &= \frac{30\mu_4-20\mu_2+6}{15(\mu_4-\mu_2^2)}+\frac{2}{3\mu_2}.\end{aligned}$$

Hence, (4.5.1) becomes

$$\min_{\xi\in\mathcal{F}_s}\left\{\frac{30\mu_4-20\mu_2+6}{15(\mu_4-\mu_2^2)}+\frac{2}{3\mu_2}\right\} \quad \text{subject to} \quad \mu_2 \le c(\mu_2\mu_6-\mu_4^2). \tag{4.5.4}$$

In light of Lemma 4.2.4, we know that the optimal designs over $\mathcal{F}$ are the same as those over $\mathcal{F}_s$. Moreover, followed by Theorem 4.2.6, we can change (4.5.2), (4.5.3), and (4.5.4) into the following forms respectively:

$$\max_{(\alpha,\beta,x)\in A}\{(\alpha+\beta x)[\alpha+\beta x^2-(\alpha+\beta x)^2]\} \quad \text{subject to} \quad \alpha+\beta x \le c\alpha\beta x(1-x)^2, \tag{4.5.5}$$

$$\min_{(\alpha,\beta,x)\in A}\left\{\frac{1+\alpha+\beta x^2}{\alpha+\beta x^2-(\alpha+\beta x)^2}+\frac{1}{\alpha+\beta x}\right\} \quad \text{subject to} \quad \alpha+\beta x \le c\alpha\beta x(1-x)^2, \tag{4.5.6}$$

and

$$\min_{(\alpha,\beta,x)\in A}\left\{\frac{30(\alpha+\beta x^2)-20(\alpha+\beta x)+6}{15[\alpha+\beta x^2-(\alpha+\beta x)^2]}+\frac{2}{3(\alpha+\beta x)}\right\} \tag{4.5.7}$$

$$\text{subject to} \quad \alpha+\beta x \le c\alpha\beta x(1-x)^2 \ .$$

The C-restricted D-, A-, and Q- optimal designs can be solved numerically by searching for optimal solutions according to (4.5.5), (4.5.6), and (4.5.7) respectively for any $c \in [16,\infty)$.

4.6 Some Numerical Results

In Section 4.3, we found the C-restricted optimal designs for model P_1. We noted that the D-, A-, and Q-optimal designs are the same. But the efficiencies of these designs are different. For different values of c_1, we present the optimal designs and their efficiencies in Table 4.6.1. The calculations are based on Definition 4.3.2, Definition 4.3.3 and Theorem 4.3.1. When $c_1 = \infty$, the "restricted" optimal design is $\xi_0 = \frac{1}{2}\,\Delta_{\pm 1}$ which is the usual optimal design (without the restriction).

In Section 4.4, we discussed the $\underset{\sim}{C}^2$-restricted optimal designs for model P_1, where $\underset{\sim}{C}^2 = (c_1, c_2)^T$. We found that the optimal design ξ_0 has the form of $\xi_0 = \frac{\alpha}{2}\,\Delta_{\pm 1} + \frac{\beta}{2}\,\Delta_{\pm\sqrt{x}} + (1-\alpha-\beta)\Delta_0$ for some $0 \le \alpha \le 1$, $0 \le \beta \le 1$, $\alpha+\beta \le 1$, and $0 \le x \le 1$. In Table 4.6.2, we provide the $\underset{\sim}{C}^2$-restricted optimal designs for model P_1. Note that, in Table 4.6.2, there are three numbers in each cell. The first one is the value of α; the second β; the third x. We provide the efficiencies of $\underset{\sim}{C}^2$-restricted optimal designs for model P_1 in Table 4.6.3. The calculations are based on Definition 4.3.2 and Definition 4.3.3 rather than Definition 4.4.6. Again there are three numbers in each cell of Table 4.6.3. The first one is the model P_1 D-efficiency of design ξ_0; the second A-efficiency; the third Q-efficiency. There are some missing values in Table 4.6.2 and Table 4.6.3 which are the cases when $c_1 = 4$ and $c_2 < \infty$. This is simply because when $c_1 = 4$, there is only one design measure $\xi_0 = \frac{1}{4}\,\Delta_{\pm 1} + \frac{1}{2}\,\Delta_0$ satisfying $|B_1(\xi)| \le c_1\,|B_2(\xi)|$ and also satisfying $|B_2(\xi)| \le c_2\,|B_3(\xi)|$ only if $c_2 = \infty$ (see Remark 1 in Section 4.4). Some results in Table 4.6.2 and Table 4.6.3 are the same due to the fact that $\mu_2(\xi_0(\infty, c_2)) = \max_{\xi \in \mathcal{F}(c_1,c_2)} \mu_2(\xi)$ when $2c_1 \ge c_2$ (see Theorem 4.4.5 Part (i)). In Table 4.6.2, there are two "$*$" values for x, which means x can be any number between 0 and 1. The reason is, in these two cases, we have $\beta = 0$. When $c_1 = 4$ and $c_2 = \infty$, we have $\xi_0 = \frac{1}{4}\,\Delta_{\pm 1} + \frac{1}{2}\,\Delta_0$ which coincides with the c_1-restricted optimal design for model P_1 with $c_1 = 4$. When $c_1 = \infty$ and $c_2 = \infty$, we have $\xi_0 = \frac{1}{2}\,\Delta_{\pm 1}$.

Based on the formulas in Section 4.5, C-restricted D-, A-, and Q-optimal designs and their efficiencies for model P_2 are calculated in Table 4.6.4 and Table 4.6.5. In this case, we again find the fact that the optimal design ξ_0 has the form of $\xi_0 = \frac{\alpha}{2}\,\Delta_{\pm 1} + \frac{\beta}{2}\,\Delta_{\pm\sqrt{x}} + (1-\alpha-\beta)\Delta_0$ for some $0 \le \alpha \le 1,\ 0 \le \beta \le 1,\ \alpha+\beta \le 1$, and $0 \le x \le 1$. For some different c_2, the corresponding α, β, and x values are presented. When $c_2 = \infty$, we get $\xi_0 = \frac{1}{3}\,\Delta_{\pm 1} + \frac{1}{3}\,\Delta_0$ for D-optimality and $\xi_0 = \frac{1}{4}\,\Delta_{\pm 1} + \frac{1}{2}\,\Delta_0$ for A- and Q-optimality. (In this case, A- and Q-optimal designs are the same). These are the usual D-, A- and Q-optimal designs for the quadratic polynomial regression model.

Table 4.6.1

C-restricted Optimal Designs and Their Efficiencies for Model P_1

c_1	$\xi_0(-1), \xi_0(+1)$	$\xi_0(0)$	$E_1^D(\xi_0)$	$E_1^A(\xi_0), E_1^G(\xi_0)$	$E_1^Q(\xi_0)$
4.0	0.2500	0.5000	0.7071	0.6667	0.8000
4.5	0.3333	0.3333	0.8165	0.8000	0.8889
5.0	0.3618	0.2764	0.8507	0.8396	0.9128
5.5	0.3816	0.2389	0.8724	0.8644	0.9272
6.0	0.3943	0.2113	0.8881	0.8819	0.9372
7.0	0.4137	0.1727	0.9096	0.9055	0.9504
8.0	0.4268	0.1464	0.9239	0.9210	0.9589
9.0	0.4363	0.1273	0.9342	0.9320	0.9648
10.0	0.4436	0.1127	0.9420	0.9403	0.9692
15.0	0.4641	0.0718	0.9634	0.9627	0.9810
20.0	0.4736	0.0528	0.9732	0.9729	0.9863
50.0	0.4898	0.0204	0.9897	0.9897	0.9948
100.0	0.4949	0.0101	0.9949	0.9949	0.9975
∞	0.5000	0.0000	1.0000	1.0000	1.0000

Table 4.6.2

$\underset{\sim}{C}^2$-restricted Optimal Designs for Model P_1

c_1 / c_2	4	10	20	50	100	200	∞
16		0.3333 0.6667 0.2500	0.3333 0.6667 0.2500	0.3333 0.6667 0.2500	0.3333 0.6667 0.2500	0.3333 0.6667 0.2500	0.3333 0.6667 0.2500
20		0.5669 0.4331 0.3618	0.5669 0.4331 0.3618	0.5669 0.4331 0.3618	0.5669 0.4331 0.3618	0.5669 0.4331 0.3618	0.5669 0.4331 0.3618
30		0.7400 0.2600 0.2700	0.7265 0.2735 0.4208	0.7265 0.2735 0.4208	0.7265 0.2735 0.4208	0.7265 0.2735 0.4208	0.7265 0.2735 0.4208
50		0.8200 0.1800 0.1700	0.8500 0.1500 0.3700	0.8388 0.1612 0.4562	0.8388 0.1612 0.4562	0.8388 0.1612 0.4562	0.8388 0.1612 0.4562
100		0.8500 0.0800 0.2300	0.9100 0.0700 0.2900	0.9199 0.0801 0.4791	0.9199 0.0801 0.4791	0.9199 0.0801 0.4791	0.9199 0.0801 0.4791
200		0.8700 0.0400 0.2100	0.9300 0.0400 0.2300	0.9600 0.0400 0.2700	0.9600 0.0400 0.4898	0.9600 0.0400 0.4898	0.9600 0.0400 0.4898
500		0.8800 0.0200 0.1700	0.9400 0.0300 0.1200	0.9700 0.0200 0.2700	0.9800 0.0200 0.2800	0.9800 0.0200 0.4900	0.9840 0.0160 0.4960
∞	0.5000 0.0000 *	0.8850 0.0050 0.2300	0.9450 0.0050 0.2350	0.9750 0.0050 0.6950	0.9850 0.0050 0.8450	0.9900 0.0050 0.8450	1.0000 0.0000 *

Table 4.6.3

The Efficiencies of $\underset{\sim}{C}^2$–restricted Optimal Designs for Model P_1

c_1 / c_2	4	10	20	50	100	200	∞
16		0.7071 0.6667 0.8000	0.7071 0.6667 0.8000	0.7071 0.6667 0.8000	0.7071 0.6667 0.8000	0.7071 0.6667 0.8000	0.7071 0.6667 0.8000
20		0.8507 0.8396 0.9128	0.8507 0.8396 0.9128	0.8507 0.8396 0.9128	0.8507 0.8396 0.9128	0.8507 0.8396 0.9128	0.8507 0.8396 0.9128
30		0.9001 0.8951 0.9447	0.9174 0.9140 0.9550	0.9174 0.9140 0.9550	0.9174 0.9140 0.9550	0.9174 0.9140 0.9550	0.9174 0.9140 0.9550
50		0.9223 0.9193 0.9579	0.9516 0.9504 0.9746	0.9551 0.9541 0.9765	0.9551 0.9541 0.9765	0.9551 0.9541 0.9765	0.9551 0.9541 0.9765
100		0.9319 0.9296 0.9635	0.9645 0.9639 0.9816	0.9789 0.9787 0.9892	0.9789 0.9787 0.9892	0.9789 0.9787 0.9892	0.9789 0.9787 0.9892
200		0.9372 0.9353 0.9665	0.9691 0.9686 0.9841	0.9853 0.9852 0.9925	0.9897 0.9897 0.9948	0.9897 0.9897 0.9948	0.9897 0.9897 0.9948
500		0.9399 0.9381 0.9681	0.9714 0.9710 0.9853	0.9876 0.9875 0.9937	0.9928 0.9927 0.9964	0.9949 0.9949 0.9974	0.9960 0.9960 0.9980
∞	0.7071 0.6667 0.8000	0.9414 0.9396 0.9689	0.9727 0.9723 0.9860	0.9892 0.9891 0.9945	0.9946 0.9946 0.9973	0.9971 0.9971 0.9985	1.0000 1.0000 1.0000

Table 4.6.4

C-restricted D-Optimal Design and its Efficiency for Model P_2

c_2	α	β	x	$E_2^D(\xi_0)$
16.1	0.37	0.63	0.25	0.7758
20	0.53	0.47	0.19	0.8807
25	0.57	0.43	0.14	0.9170
30	0.60	0.39	0.12	0.9351
50	0.63	0.35	0.07	0.9649
75	0.64	0.26	0.06	0.9776
100	0.65	0.28	0.04	0.9837
200	0.66	0.18	0.03	0.9921
500	0.66	0.21	0.01	0.9968
∞	0.67	0.33	0.00	1.0000

Table 4.6.5

C-restricted A- and Q-optimal Design and Their Efficiencies for Model P_2

c_2	α	β	x	$E_2^A(\xi_0)$	$E_2^Q(\xi_0)$
16.1	0.34	0.66	0.23	0.6469	0.8417
20	0.41	0.58	0.14	0.8058	0.9233
25	0.43	0.56	0.10	0.8642	0.9488
30	0.45	0.54	0.08	0.8947	0.9608
50	0.47	0.47	0.05	0.9424	0.9792
75	0.48	0.49	0.03	0.9636	0.9870
100	0.49	0.28	0.04	0.9726	0.9902
200	0.49	0.18	0.03	0.9868	0.9953
500	0.50	0.21	0.01	0.9948	0.9981
∞	0.50	0.00	0.00	1.0000	1.0000

Chapter 5

Bounded Bias Optimal Designs for Approximately Linear Regression Models

5.1 Introduction and Preliminaries

In 1992, Douglas Wiens suggested a problem that considers the optimal design minimizing the variance of the estimator of the parameters of the regression function when the fitted model is correct, subject to a bound on the bias term which occurs when the true model is different from the assumed one. The corresponding optimal designs can be called bounded bias optimal designs. We are now going to formulate the problem in detail.

Consider the following regression model:

$$y_i = y(x_i) = \underset{\sim}{\theta}^T \underset{\sim}{f}\,(x_i) + \epsilon_i, \quad i = 1, ..., n, \tag{5.1.1}$$

where ϵ_i's are independent and identically distributed with mean 0 and some common variance $\sigma^2 > 0$. $\underset{\sim}{\theta}^T = (\theta_0, \theta_1, ..., \theta_k)$ and $\underset{\sim}{f}^T\,(x) = (f_0(x), f_1(x), ..., f_k(x))$, $x \in S \subseteq \mathbf{R}^1$.

Let $\hat{\underset{\sim}{\theta}}$ be the least squares estimator of $\underset{\sim}{\theta}$. For a given design measure ξ, we have

$$M(\xi) = E[(\hat{\underset{\sim}{\theta}} - \underset{\sim}{\theta})(\hat{\underset{\sim}{\theta}} - \underset{\sim}{\theta})^T] = \frac{\sigma^2}{n} B^{-1}(\xi), \tag{5.1.2}$$

where $B(\xi) = \int_S \underset{\sim}{f}(x) \cdot \underset{\sim}{f}^T(x) d\xi(x)$.

Suppose that (5.1.1) is only an approximation of the real situation. Instead of (5.1.1), the real model is

$$y_i = y(x_i) = \underset{\sim}{\theta}^T \underset{\sim}{f}(x_i) + \psi(x_i) + \epsilon_i, \quad i = 1, ..., n, \tag{5.1.3}$$

where $\psi(x)$ is the bias factor, departure from (5.1.1).

Under (5.1.3), we know that

$$E[\hat{\underset{\sim}{\theta}} - \underset{\sim}{\theta}] = B^{-1}(\xi) \cdot \underset{\sim}{b}(\psi, \xi) \tag{5.1.4}$$

and

$$\begin{aligned} M(\psi, \xi) &= E[(\hat{\underset{\sim}{\theta}} - \underset{\sim}{\theta})(\hat{\underset{\sim}{\theta}} - \underset{\sim}{\theta})^T] \\ &= \frac{\sigma^2}{n} B^{-1}(\xi) + B^{-1}(\xi) \underset{\sim}{b}(\psi, \xi) \underset{\sim}{b}^T(\psi, \xi) B^{-1}(\xi), \end{aligned} \tag{5.1.5}$$

where $\underset{\sim}{b}(\psi, \xi) = \int_S \underset{\sim}{f}(x) \psi(x) d\xi(x)$.

Let $v(\xi) = \frac{\sigma^2}{n} B^{-1}(\xi)$, and $bias(\psi, \xi) = B^{-1}(\xi) \underset{\sim}{b}(\psi, \xi) \underset{\sim}{b}^T(\psi, \xi) B^{-1}(\xi)$. For some loss function $\mathcal{L}$, our objective is the following:

$$\min_{\xi \in \mathcal{F}} \mathcal{L}[B^{-1}(\xi)] \text{ subject to } \max_{\psi \in \Psi} ||E[\hat{\underset{\sim}{\theta}} - \underset{\sim}{\theta}]|| \leq c, \tag{5.1.6}$$

where $|| \cdot ||$ is the norm of vector $E[\hat{\underset{\sim}{\theta}} - \underset{\sim}{\theta}]$. As we noted, $B^{-1}(\xi)$ is proportional to $v(\xi)$ and $E[\hat{\underset{\sim}{\theta}} - \underset{\sim}{\theta}]$ is closely related to $bias(\psi, \xi)$. (In fact, $bias(\psi, \xi) = \left(E\left[\hat{\underset{\sim}{\theta}} - \underset{\sim}{\theta}\right]\right) \cdot \left(E\left[\hat{\underset{\sim}{\theta}} - \underset{\sim}{\theta}\right]\right)^T$). There are many different ways to choose $\mathcal{L}$, $\mathcal{F}$, and Ψ.

First, we consider the following situation. Let

$$\mathcal{F} = \{\xi(x) : \xi(x) \text{ is a design measure such that } \sigma(\xi) \subseteq S \subseteq \mathbf{R}^1\}$$

and

$$\begin{aligned} \Psi(\phi) &= \{\psi(x) : |\psi(x)| \leq \phi(x), \text{ where } \phi(x) \text{ is a known function such that} \\ &\qquad \phi(-x) = \phi(x) \text{ and } \phi(x) \geq 0 \text{ on } S\}. \end{aligned}$$

For the loss function, we choose $\mathcal{L} \in \{\mathcal{L}_D, \mathcal{L}_A, \mathcal{L}_Q, \mathcal{L}_G\}$. We understand that $\mathcal{L}_D(\xi) = |B^{-1}(\xi)|$, $\mathcal{L}_A(\xi) = trB^{-1}(\xi)$, $\mathcal{L}_Q(\xi) = \int_S d(x, \xi) dx$ and $\mathcal{L}_G(\xi) = \max_{x \in S} d(x, \xi)$, where

$d(x,\xi) = \underset{\sim}{f}^T(x)B^{-1}(\xi)\underset{\sim}{f}(x)$. It is clear that $d(x,\xi)$ is proportional to the MSE of $\hat{\underset{\sim}{\theta}}^T \underset{\sim}{f}(x)$ as an estimator of $\underset{\sim}{\theta}^T \underset{\sim}{f}(x)$ under the model (5.1.1).

In general, it is hard to find a solution of (5.1.6). In Section 5.2 and Section 5.3, we further restrict our attention to the one dimensional linear regression case, i.e., $\underset{\sim}{f}^T(x) = (1,x)$. For the set of Ψ, we consider two cases:
(i) $\Psi_s(\phi) = \{\psi : \psi(0) = \psi'(0) = 0,\ \psi(-x) = \psi(x),\ |\psi(x)| \leq \phi(x)$ on $[-1,1]$ and $\phi(x) = \sum_{i=1}^{p} a_{2i}x^{2i} \geq 0$ on $[-1,1]\}$ and
(ii) $\Psi_a(\phi) = \{\psi : \psi(0) = \psi'(0) = 0,\ \psi(-x) = -\psi(x),\ |\psi(x)| \leq |\phi(x)|$ on $[-1,1]$ and $\phi(x) = \sum_{i=1}^{p} a_{2i+1}x^{2i+1} \geq 0$ on $[0,1]\}$.
The condition of $\psi(0) = \psi'(0) = 0$ is to insure the identifiability of the parameters to be estimated. We refer to $\Psi_s(\phi)$ as the set of "symmetric contamination" and $\Psi_a(\phi)$ "antisymmetric contamination". These two cases are treated in Section 5.2 and Section 5.3, respectively. For some choices of the "upper bound" function $\phi(x)$, we find the solution to the problem (5.1.6). It is possible to extend the problem to the higher dimensional case and high order polynomial case. But the details will be very tedious. Similar to the problem of $\underset{\sim}{C}^k$-restricted optimal designs, we have that the bounded bias optimal designs over $\mathcal{F}$ is equivalent to that over $\mathcal{F}_S$, where $\mathcal{F}_S = \{\xi(x) : \xi(x) \in \mathcal{F}$ and $\xi(-x) = \xi(x)$ on $[-1,1]\}$. The problem can be further reduced to a non-linear programming problem with a finite number of variables or a search for optimal solutions within the subclass of $\mathcal{F}_S$ that have design measures supported on a finite number of points. Some relevant results will be provided later in this section.

Second, we consider the choice of $\mathcal{F}$ and Ψ as the following. Let

$$\mathcal{F} = \{\xi(x) : \frac{d\xi(x)}{dx} = m(x), \int_S m(x)dx = 1, m(x) \geq 0 \text{ and } m(-x) = m(x) \text{ on } S\},$$

and

$$\Psi = \{\psi(x) : \int_S \psi^2(x)dx \leq \eta^2, \int_S \underset{\sim}{f}(x)\psi(x)ds = \underset{\sim}{0}\},$$

where η is a preassigned constant and the side condition $\int_S \underset{\sim}{f}(x)\psi(x)dx = \underset{\sim}{0}$ is to insure the identifiability of the parameters to be estimated.

Let $\hat{y} = \hat{\underset{\sim}{\theta}}^T \underset{\sim}{f}(x)$ be the estimator of $\underset{\sim}{\theta}^T \underset{\sim}{f}(x)$. Under (5.1.3), we have

$$MSE(\hat{y}) = \frac{\sigma^2}{n} \underset{\sim}{f}^T(x) B^{-1}(\xi) \underset{\sim}{f}(x) + \underset{\sim}{f}^T(x) B^{-1}(\xi) \underset{\sim}{b}(\psi,\xi) \underset{\sim}{b}^T(\psi,\xi) B^{-1}(\xi) \underset{\sim}{f}(x),$$

and

$$\int_S MSE(\hat{y})dx = \frac{\sigma^2}{n} trAB^{-1}(\xi) + \underset{\sim}{b}^T(\psi,\xi) B^{-1}(\xi) A B^{-1}(\xi) \underset{\sim}{b}(\psi,\xi),$$

where $A = \int_S \underset{\sim}{f}(x) \underset{\sim}{f}^T(x)dx$.

Let $V(\xi) = \frac{\sigma^2}{n} trAB^{-1}(\xi)$ and $Bias(\psi,\xi) = \underset{\sim}{b}^T(\psi,\xi) B^{-1}(\xi) A B^{-1}(\xi) \underset{\sim}{b}(\psi,\xi)$. We consider the following problem:

$$\min_{\xi \in \mathcal{F}} V(\xi) \quad \text{subject to} \quad \max_{\psi \in \Psi} Bias(\psi,\xi) \leq c, \tag{5.1.7}$$

where c is a preassigned positive number.

The maximization of $Bias(\psi,\xi)$ over the class Ψ was done by Huber (1975). The solution to the problem (5.1.7) is the main result in Section 5.4. The problem (5.1.7) can be easily extended to multiple linear regression case, since the maximization of $Bias(\psi,\xi)$ in multiple linear regression case was solved by Wiens (1990).

We are now going to provide some results which can be used to reduce the problem (5.1.6). For the cases we will consider in Section 5.2 and Section 5.3, we always have

$$\max_{\psi \in \Psi} ||E(\hat{\underset{\sim}{\theta}} - \underset{\sim}{\theta})|| = \sum_{i=0}^{p} c_i \mu_{2i}, \tag{5.1.8}$$

where c_i's are some coefficients and $\mu_{2i} = \int_{-1}^{1} x^{2i} d\xi(x)$.

Let $V(\phi, c)$ be the set of all design measures satisfying the constraint $\sum_{i=0}^{p} c_i \mu_{2i} \leq c$, where c is a preassigned positive number and the range of c will be specified case by case. We claim the following:

Lemma 5.1.1 *(i) $V(\phi,c)$ is convex.*

(ii) $\xi(x) \in V(\phi,c)$ if and only if $\xi(-x) \in V(\phi,c)$.

Proof: (i) Let $\xi_1, \xi_2 \in V(\phi,c)$ and $\xi^* = \lambda\xi_1 + (1-\lambda)\xi_2$ where $0 < \lambda < 1$. We have

$$\begin{aligned}
\textstyle\sum_{i=0}^{p} c_i \mu_{2i} &= \textstyle\sum_{i=0}^{p} c_i \int_{-1}^{1} x^{2i} d\xi^*(x) \\
&= \textstyle\sum_{i=0}^{p} c_i \int_{-1}^{1} x^{2i} d(\lambda\xi_1 + (1-\lambda)\xi_2) \\
&= \lambda \textstyle\sum_{i=0}^{p} c_i \int_{-1}^{1} x^{2i} d\xi_1 + (1-\lambda) \textstyle\sum_{i=0}^{p} c_i \int_{-1}^{1} x^{2i} d\xi_2 \\
&\leq \lambda c + (1-\lambda)c = c.
\end{aligned}$$

Hence, $\xi^* \in V(\phi, c)$.

(ii) Note that

$$\sum_{i=0}^{p} c_i \int_{-1}^{1} x^{2i} d\xi(-x) = \sum_{i=0}^{p} c_i \int_{-1}^{1} x^{2i} d\xi(x).$$

Hence, $\xi(x) \in V(\phi, c)$ if and only if $\xi(-x) \in V(\phi, c)$. $\square$

Let $D = \{\xi_0 : |B^{-1}(\xi_0)| = \min_{\xi \in V(\phi,c)} |B^{-1}(\xi)|\}$, $A = \{\xi_0 : trB^{-1}(\xi_0) = \min_{\xi \in V(\phi,c)} trB^{-1}(\xi)\}$, and $Q = \{\xi_0 : \int_{-1}^{1} d(x, \xi_0)dx = \min_{\xi \in V(\phi,c)} \int_{-1}^{1} d(x, \xi)dx\}$, that is, D, A, and Q are the sets of bounded *bias* D-, A-, and Q-optimal designs respectively. We state the following lemma without proof, since the proof is very similar to the proof of Lemma 4.2.4.

Lemma 5.1.2 *(i) D, A, Q are convex sets.*

(ii) $\xi(x) \in D, A, Q$ if and only if $\xi(-x) \in D, A, Q$ respectively.

Assume that ξ_1 is a bounded *bias* D-, A-, or Q-optimal design measure and we define ξ_2 by $\xi_2(x) = \xi_1(-x)$. The consequence of Lemma 5.1.2 is that there is a symmetrical optimal design, namely $(\xi_1 + \xi_2)/2$. The problem (5.1.6) can be further simplified by using Theorem 4.2.5 and Theorem 4.2.6 as we will see in Section 5.2 and Section 5.3.

5.2 Approximately Linear Regression Model with Norm 1 Bounded and Symmetric Contamination Functions

In this section, we consider the one dimensional linear regression model and the class of symmetric design measures on $[-1,1]$. Let $\underset{\sim}{f}^T(x) = (1, x)$, $\underset{\sim}{\theta}^T = (\theta_0, \theta_1)$, and $\mathcal{F}_S = \{\xi : \xi \in \mathcal{F}, \xi(-x) = \xi(x)\}$. For any $\xi \in \mathcal{F}_S$ we have

$$B(\xi) = \int_{-1}^{1} \underset{\sim}{f}(x)\, \underset{\sim}{f}^T(x) d\xi(x) = \begin{pmatrix} 1 & 0 \\ 0 & \mu_2 \end{pmatrix}, \quad B^{-1}(\xi) = \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{\mu_2} \end{pmatrix}$$

and

$$d(x,\xi) = \underset{\sim}{f}^T(x) B^{-1}(\xi)\, \underset{\sim}{f}(x) = (1, x) \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{\mu_2} \end{pmatrix} \begin{pmatrix} 1 \\ x \end{pmatrix} = 1 + \frac{x^2}{\mu_2},$$

where $\mu_2 = \int_{-1}^{1} x^2 d\xi(x)$. We know that $\mathcal{L}_D(\xi) = \dfrac{1}{\mu_2}$, $\mathcal{L}_A(\xi) = 1 + \dfrac{1}{\mu_2}$, $\mathcal{L}_Q(\xi) = 2 + \dfrac{2}{3\mu_2}$, and $\mathcal{L}_G(\xi) = 1 + \dfrac{1}{\mu_2}$. Hence, for $\mathcal{L} \in \{\mathcal{L}_D, \mathcal{L}_A, \mathcal{L}_Q, \mathcal{L}_G\}$,

$$\min_{\xi \in \mathcal{F}_S} \mathcal{L}(\xi) \text{ subject to } \max_{\psi \in \Psi} ||E(\hat{\underset{\sim}{\theta}} - \underset{\sim}{\theta})|| \leq c$$

is equivalent to

$$\max_{\xi \in \mathcal{F}_S} \mu_2 \text{ subject to } \max_{\psi \in \Psi} ||E(\hat{\underset{\sim}{\theta}} - \underset{\sim}{\theta})|| \leq c.$$

For the class of contamination functions Ψ, we choose

$$\begin{aligned} \Psi_s(\phi) = \ & \{\psi : \psi(0) = \psi'(0) = 0, \psi(-x) = \psi(x), \\ & |\psi(x)| \leq \phi(x) \text{ on } [-1,1], \text{ and } \phi(x) = \textstyle\sum_{i=1}^{p} a_{2i} x^{2i} \geq 0 \text{ on } [0,1]\}. \end{aligned}$$

In this case, we have

$$
\begin{aligned}
E[\hat{\underset{\sim}{\theta}} - \underset{\sim}{\theta}] &= B^{-1}(\xi)\, \underset{\sim}{b}\,(\psi, \xi) \\
&= \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{\mu_2} \end{pmatrix} \begin{pmatrix} \int \psi d\xi \\ \int x\psi d\xi \end{pmatrix} \\
&= \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{\mu_2} \end{pmatrix} \begin{pmatrix} \int \psi d\xi \\ 0 \end{pmatrix} \\
&= \begin{pmatrix} \int \psi d\xi \\ 0 \end{pmatrix}
\end{aligned}
$$

and $||E[\hat{\underset{\sim}{\theta}} - \underset{\sim}{\theta}]|| = |\int \psi(x) d\xi(x)|$. It is clear that

$$
\begin{aligned}
\max_{\psi \in \Psi_s} ||E[\hat{\underset{\sim}{\theta}} - \underset{\sim}{\theta}]|| &= \max_{\psi \in \Psi_s} |\int \psi(x) d\xi(x)| \\
&\leq \max_{\psi \in \Psi_S} \int |\psi(x)| d\xi(x) \\
&\leq \int \phi(x) d\xi(x) \\
&= \int (\textstyle\sum_{i=1}^{p} a_{2i} x^{2i}) d\xi(x) \\
&= \textstyle\sum_{i=1}^{p} a_{2i} \mu_{2i}.
\end{aligned}
$$

On the other hand, $\phi(x) \in \Psi_s(\phi)$. Hence we have

$$
\max_{\psi \in \Psi_S} ||E[\hat{\underset{\sim}{\theta}} - \underset{\sim}{\theta}]|| = \int \phi(x) d\xi(x) = \sum_{i=1}^{p} a_{2i} \mu_{2i}.
$$

In the case considered here, the problem (5.1.6) becomes

$$
\max_{\xi \in \mathcal{F}_S} \mu_2 \quad \text{subject to } \sum_{i=1}^{p} a_{2i} \mu_{2i} \leq c \tag{5.2.1}
$$

for some a_{2i}'s such that $\phi(x) = \sum_{i=1}^{p} a_{2i} x^{2i} \geq 0$ on $[-1, 1]$.

It is necessary to specify the range of c so that the problem (5.2.1) has feasible solutions. Let

$$
c_* = \min_{\xi \in \mathcal{F}_S} \{\sum_{i=1}^{p} a_{2i} \mu_{2i}\}, \text{ and } c^* = \max_{\xi \in \mathcal{F}_S} \{\sum_{i=1}^{p} a_{2i} \mu_{2i}\}.
$$

Then (5.2.1) has solutions for any $c_* \leq c \leq c^*$. (5.2.1) has no solution for $c < c_*$. For $c > c^*$, (5.2.1) has the same solution as the unconditional optimal design problem. Hence c^* is not the "real upper bound".

We are now seeking the solution of the problem (5.2.1) for some specific choices of $\phi(x)$.

Case I: $\phi(x) = a_2x^2$, $a_2 > 0$

When $\phi(x) = a_2x^2$, the problem (5.2.1) becomes

$$\max_{\xi\in\mathcal{F}_S} \mu_2 \text{ subject to } a_2\mu_2 \le c. \tag{5.2.2}$$

Denote μ_2^* to be the maximum value of μ_2 in (5.2.1) and ξ^* to be the corresponding design measure. The following result is trivial.

Theorem 5.2.1 *In (5.2.1), let $\phi(x) = a_2x^2$ and $a_2 > 0$. Then we have*

(i) $\xi^* = \frac{1}{2}\Delta_{\pm\sqrt{\tau}}$ and $\mu_2^* = \tau$, where $\tau = \dfrac{c}{a_2}$, if $0 < c < a_2$

(ii) $\xi^* = \frac{1}{2}\Delta_{\pm 1}$ and $\mu_2^* = 1$, if $c \ge a_2$.

Remark 1. The usual optimal design is $\xi_0 = \frac{1}{2}\Delta_{\pm 1}$ which is supported at ± 1. For Case I, the violation occurs at ± 1. Hence the bounded bias optimal design ξ^* supported at $\pm\sqrt{\tau}$, somehow stays away from ± 1.

Case II: $\phi(x) = a_2x^2 + a_4x^4$, $a_4 \ne 0$

In order to make $\phi(x) \ge 0$ on $[-1, 1]$, we must have (i) $a_4 > 0$, $a_2 \ge 0$ or (ii) $a_4 < 0$, $a_2 + a_4 \ge 0$. In Case II, the side condition of (5.2.1) is $a_2\mu_2 + a_4\mu_4 \le c$. It is clear that $c_* = 0$. We have $c^* = a_2 + a_4$ if $a_4 > 0$, $a_2 \ge 0$. For the case that $a_4 < 0$, $a_2 + a_4 \ge 0$, we have

$$c^* := \max_{\xi\in\mathcal{F}_S}\{a_2\mu_2 + a_4\mu_4\} = \max_{\xi\in\mathcal{F}_0}\{a_2\mu_2 + a_4\mu_4\},$$

where $\mathcal{F}_0 = \{\xi : \xi = \frac{\alpha}{2}\Delta_{\pm\sqrt{x}} + (1-\alpha)\Delta_0,\ 0 \le \alpha \le 1,\ 0 \le x \le 1\}$. The second equality is followed by Theorem 4.2.5. For any $\xi \in \mathcal{F}_0$, we have $\mu_2 = \alpha x$ and $\mu_4 = \alpha x^2$. Hence

$$\max_{\xi\in\mathcal{F}_0}\{a_2\mu_2 + a_4\mu_4\} = \max_{(\alpha,x)\in A_0}\{a_2\alpha x + a_4\alpha x^2\},$$

where $A_0 = \{(\alpha, x) : 0 \le \alpha \le 1,\ 0 \le x \le 1\}$. We define $L(\alpha, x) = a_2\alpha x + a_4\alpha x^2$, and let

$$\begin{cases} \dfrac{\partial L}{\partial \alpha} = a_2 x + a_4 x^2 := 0 \\ \dfrac{\partial L}{\partial x} = a_2\alpha + 2a_4\alpha x := 0. \end{cases}$$

We find that there is no solution in the interior of A_0. There are four boundary cases. They are (i) $\alpha = 0$, (ii) $x = 0$, (iii) $\alpha = 1$, and (iv) $x = 1$. For case (i) and (ii), we have $L(0, x) = L(\alpha, 0) = 0$. Case (iii) gives us $L(1, x) = a_2 x + a_4 x^2$. Let $\dfrac{dL}{dx} = a_2 + 2a_4 x = 0$, we get $x = -\dfrac{a_2}{2a_4}$. Under the condition that $a_4 < 0$, $a_2 + a_4 \ge 0$, we always have $x = -\dfrac{a_2}{2a_4} > 0$. If $-\dfrac{a_2}{2a_4} \le 1$, i.e. $a_2 + 2a_4 \le 0$, then $L(1, -\dfrac{a_2}{2a_4}) = a_2(-\dfrac{a_2}{2a_4}) + a_4(-\dfrac{a_2}{2a_4})^2 = -\dfrac{a_2^2}{4a_4}$. For case (iv), we have $L(\alpha, 1) = a_2\alpha + a_4\alpha = (a_2 + a_4)\alpha$. The fact $a_2 + a_4 \ge 0$ implies $\max L(\alpha, 1) = L(1, 1) = a_2 + a_4$. Note that $L(1, -\dfrac{a_2}{2a_4}) - L(1, 1) = -\dfrac{a_2^2}{4a_4} - a_2 - a_4 = \dfrac{(a_2 + 2a_4)^2}{-4a_4} \ge 0$. Hence, we have

$$c^* = -\frac{a_2^2}{4a_4} \text{ if } a_4 < 0,\ a_2 + a_4 \ge 0, \text{ and } a_2 + 2a_4 \le 0$$

and

$$c^* = a_2 + a_4 \text{ if } a_4 < 0,\ a_2 + 2a_4 > 0.$$

We summarize these results as the following lemma:

Lemma 5.2.2 *(i) $\phi(x) = a_2x^2 + a_4x^4 \ge 0$ on $[-1, 1]$ if and only if one of the following is true:.*

(a) $a_4 > 0$, and $a_2 \ge 0$; (b) $a_4 < 0$, and $a_2 + a_4 \ge 0$.

(ii) When $\phi(x) \ge 0$ on $[-1, 1]$, we have $c_* = 0$ and

$$c^* = \begin{cases} a_2 + a_4 & \text{if } a_4 > 0,\ a_2 \ge 0, \text{ or } a_4 < 0, a_2 + 2a_4 > 0 \\ -\dfrac{a_2^2}{4a_4} & \text{if } a_4 < 0, a_2 + a_4 \ge 0, \text{ and } a_2 + 2a_4 \le 0. \end{cases}$$

In the case of $\phi(x) = a_2x^2 + a_4x^4$, the problem (5.2.1) becomes

$$\max_{\xi \in \mathcal{F}_S} \mu_2 \text{ subject to } a_2\mu_2 + a_4\mu_4 \le c, \tag{5.2.3}$$

which is equivalent to

$$\max_{\xi \in \mathcal{F}_0} \mu_2 \quad \text{subject to } a_2\mu_2 + a_4\mu_4 \le c. \tag{5.2.4}$$

Let $\hat{S}_0 = \{(\mu_2, \mu_4) : \mu_2^2 \le \mu_4 \le \mu_2, 0 \le \mu_2 \le 1\}$ and $\tilde{S}_0 = \{(\mu_2, \mu_4) : (\mu_2, \mu_4) \in \hat{S}_0$, and $a_2\mu_2 + a_4\mu_4 \le c\}$. Let (a_2, a_4) satisfy one of the conditions (a), (b) in Lemma 5.2.2, and $0 < c \le c^*$. It is clear that "$a_2\mu_2 + a_4\mu_4 = c$" divides $\hat{S}_0$ into $\tilde{S}_0$ and $\hat{S}_0 \backslash \tilde{S}_0$ where $\tilde{S}_0 \ne \phi$. Moreover, we have

$$\mu_2^* = \max_{(\mu_2,\mu_4)\in\tilde{S}_0} \mu_2, \tag{5.2.5}$$

and the design measure ξ^* will be the solution to the problem (5.2.3), where $\mu_2^* = \int x^2 d\xi^*(x)$.

The values of μ_2^* and the corresponding design measures ξ^* are summarized in the next theorem.

Theorem 5.2.3 *In (5.2.1), let $\phi(x) = a_2x^2 + a_4x^4$, where (a_2, a_4) satisfies one of the conditions (a) and (b) in Lemma 5.2.2. Given $0 < c \le c^*$, we have*

(i) $\xi^* = \frac{1}{2}\Delta_{\pm 1}$ and $\mu_2^* = 1$, if $a_2 + a_4 \le c$;

(ii) $\xi^* = \frac{1}{2}\Delta_{\pm\sqrt{z}}$ and $\mu_2^* = z$, where $z = \dfrac{-a_2 + \sqrt{a_2^2 + 4a_4c}}{2a_4}$, if $a_2 + a_4 > c$, $a_4 > 0$, and $a_2 \ge 0$;

(iii) $\xi^* = \frac{\tau}{2}\Delta_{\pm 1} + (1-\tau)\Delta_0$ and $\mu_2^* = \tau$, where $\tau = \dfrac{c}{a_2 + a_4}$, if $a_2 + a_4 > c$, $a_4 < 0$, and $a_2 + a_4 \ge 0$.

Proof: It is obvious that $(\mu_2, \mu_4) = (1, 1) \in \tilde{S}_0$ if and only if $a_2 + a_4 \le c$. In this case, we have $\mu_2^* = \max_{(\mu_2,\mu_4)\in\tilde{S}_0} \mu_2 = 1$ which can be achieved by $\xi^* = \frac{1}{2}\Delta_{\pm 1}$. This gives us case (i). Next, we consider the situation that $(1, 1) \notin \tilde{S}_0$, i.e. $a_2 + a_4 > c$. Assume $a_4 > 0$, $a_2 \ge 0$. Then the side condition becomes $\mu_4 \le \frac{c}{a_4} - \frac{a_2}{a_4}\mu_2$, and μ_2^* is the solution of the system equations $\mu_4 = \frac{c}{a_4} - \frac{a_2}{a_4}\mu_2$, and $\mu_4 = \mu_2^2$ (see Figure 5.2.1). We find that $\mu_2^* = \dfrac{-a_2 + \sqrt{a_2^2 + 4a_4c}}{2a_4} := z$, which can be achieved by $\xi^* = \frac{1}{2}\Delta_{\pm\sqrt{z}}$. This is case (ii). For case (iii), i.e. $a_2 + a_4 > c$, $a_4 < 0$, and $a_2 + a_4 \ge 0$, the side condition becomes $\mu_4 \geq \dfrac{c}{a_4} - \dfrac{a_2}{a_4}\mu_2$, and μ_2^* is the solution of the system equations $\mu_4 = \dfrac{c}{a_4} - \dfrac{a_2}{a_4}\mu_2$ and $\mu_4 = \mu_2$, (see Figure 5.2.2). We find $\mu_2^* = \dfrac{c}{a_2 + a_4} := \tau$ and μ_2^* can be realized by $\xi^* = \dfrac{\tau}{2}\Delta_{\pm 1} + (1-\tau)\Delta_0$. □

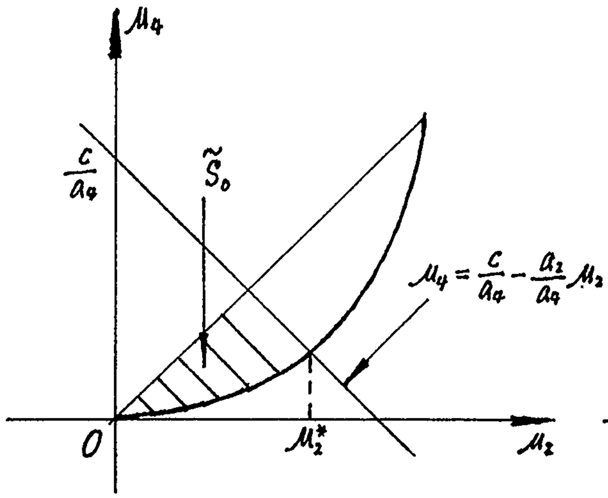

Figure 5.2.1

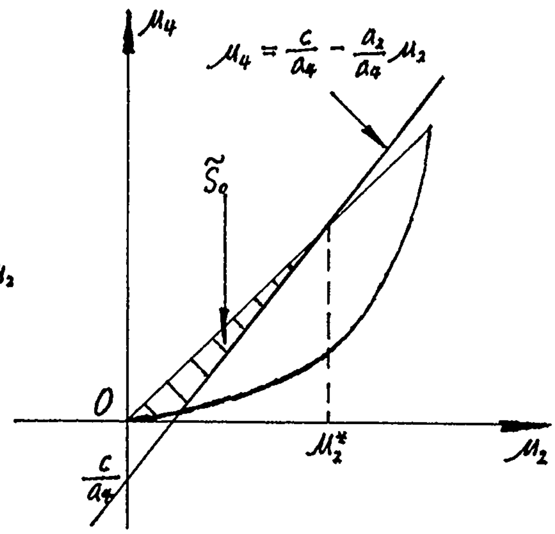

Figure 5.2.2

Case III: $\phi(x) = a_2x^2 + a_4x^4 + a_6x^6$, $a_6 \neq 0$

In the case that $\phi(x) = a_2x^2 + a_4x^4 + a_6x^6$, the problem (5.2.1) becomes

$$\max_{\xi \in \mathcal{F}_S} \mu_2 \quad \text{subject to } a_2\mu_2 + a_4\mu_4 + a_6\mu_6 \leq c, \tag{5.2.6}$$

which is equivalent to

$$\max_{\xi \in \mathcal{F}_1} \mu_2 \quad \text{subject to } a_2\mu_2 + a_4\mu_4 + a_6\mu_6 \leq c, \tag{5.2.7}$$

where $\mathcal{F}_1 = \{\xi : \xi \in \mathcal{F}_S, \xi = \frac{\alpha}{2}\Delta_{\pm 1} + \frac{\beta}{2}\Delta_{\pm\sqrt{x}} + (1-\alpha-\beta)\Delta_0, 0 \leq \alpha \leq 1, 0 \leq \beta \leq 1, \alpha + \beta \leq 1, 0 \leq x \leq 1\}$. The equivalence of (5.2.6) and (5.2.7) follows by Theorem 4.2.6.

Before we solve the problem (5.2.7), we would like to provide a necessary and sufficient condition for $\phi(x) \geq 0$ on $[-1, 1]$ in terms of the coefficients of $\phi(x)$, namely a_2, a_4, and a_6. We state the result as below.

Lemma 5.2.4 *Let $\phi(x) = a_2x^2 + a_4x^4 + a_6x^6$. Then $\phi(x) \geq 0$ on $[-1, 1]$ if and only if one of the following is true:*

(i) $a_6 < 0,\ a_2 \geq 0,$ *and* $a_2 + a_4 + a_6 \geq 0;$

(ii) $a_6 > 0,\ a_4 \geq 0,$ *and* $a_2 \geq 0;$

(iii) $a_6 > 0,\ a_4 < 0,\ a_2 > 0,\ a_4 + 2a_6 \geq 0,$ *and* $4a_2a_6 - a_4^2 \geq 0;$

(iv) $a_6 > 0,\ a_4 < 0,\ a_2 > 0,\ a_4 + 2a_6 < 0,$ *and* $a_2 + a_4 + a_6 \geq 0.$

Proof: Let $\phi_0(x) = a_2 + a_4x^2 + a_6x^4$. Then $\phi(x) \geq 0$ on $[-1, 1]$ if and only if $\phi_0(x) \geq 0$ on $[-1, 1]$. We discuss different cases according to the signs of the coefficients of $\phi_0(x)$.

(1) $a_6 < 0$

In this case, it is necessary that $a_2 \geq 0$. Note that $\phi_0(x) = 0$ has only one pair of solutions

$$x_0 = \pm\sqrt{\frac{-a_4 - \sqrt{a_4^2 - 4a_2a_6}}{2a_6}}.$$

It is easy to see that the necessary and sufficient condition for $\phi_0(x) \geq 0$ on $[-1, 1]$ is $|x_0| \geq 1$ which is equivalent to $a_2 + a_4 + a_6 \geq 0$. In fact, we can simply require $\phi_0(1) \geq 0$. Again, we get $a_2 + a_4 + a_6 \geq 0$. This gives us case (i)

(2) $a_6 > 0,\ a_4 \geq 0$

It is obvious that the necessary and sufficient condition for $\phi_0(x) \geq 0$ on $[-1, 1]$ is $a_2 \geq 0$ which is case (ii).

(3) $a_6 > 0,\ a_4 < 0$

In this case, it is necessary that $a_2 > 0$. Let $\phi_0'(x) = 2a_4x + 4a_6x^3 := 0$. We get $x = 0$, and $x = \pm\sqrt{-\frac{a_4}{2a_6}}$. If $\sqrt{-\frac{a_4}{2a_6}} \leq 1$, i.e., $a_4 + 2a_6 \geq 0$, then we need

$$\phi_0\left(\sqrt{-\frac{a_4}{2a_6}}\right) = a_2 + a_4\left(-\frac{a_4}{2a_6}\right) + a_6\left(-\frac{a_4}{2a_6}\right)^2 = a_2 - \frac{a_4^2}{4a_6} \geq 0,$$

i.e. $4a_2a_6 - a_4^2 \geq 0$, which gives us case (iii).

If $\sqrt{-\frac{a_4}{2a_6}} > 1$, i.e. $a_4 + 2a_6 < 0$, then we need $\phi_0(1) = a_2 + a_4 + a_6 \geq 0$. This is case (iv). □

We are now going to find the range of c such that (5.2.7) has feasible solution. The range of c is indicated by the following:

Lemma 5.2.5 *Let (a_2, a_4, a_6) satisfy one of the four conditions (i)-(iv) in Lemma 5.2.4. Let* $x_0 = \dfrac{-a_4 - \sqrt{a_4^2 - 3a_2a_6}}{3a_6}$. *Then we have*

(i) $c_* = 0$

(ii) $c^* = \begin{cases} \dfrac{1}{27a_6^2}(2a_4^3 + 2a_4^2\sqrt{a_4^2 - 3a_2a_6} \\ \qquad -9a_2a_4a_6 - 6a_2a_6\sqrt{a_4^2 - 3a_2a_6}) & \text{if } 0 < x_0 < 1 \\ a_2 + a_4 + a_6 & \text{otherwise} \end{cases}$

Proof: For any $\xi \in \mathcal{F}_1$, we have $\mu_2 = \alpha + \beta x$, $\mu_4 = \alpha + \beta x^2$, and $\mu_6 = \alpha + \beta x^3$. Let

$$\begin{aligned} L(\alpha, \beta, x) &= a_2(\alpha + \beta x) + a_4(\alpha + \beta x^2) + a_6(\alpha + \beta x^3) \\ &= (a_2 + a_4 + a_6)\alpha + (a_2 x + a_4 x^2 + a_6 x^3)\beta. \end{aligned}$$

It is clear that $c_* = \min_{(\alpha,\beta,x)\in A} L(\alpha, \beta, x)$ and $c^* = \max_{(\alpha,\beta,x)\in A} L(\alpha, \beta, x)$, where $A = \{(\alpha, \beta, x) : 0 \le \alpha \le 1, 0 \le \beta \le 1, \alpha + \beta \le 1, 0 \le x \le 1\}$. Solving

$$\begin{cases} \dfrac{\partial L}{\partial \alpha} = a_2 + a_4 + a_6 := 0 \\ \dfrac{\partial L}{\partial \beta} = x(a_2 + a_4 x + a_6 x^2) := 0 \\ \dfrac{\partial L}{\partial x} = \beta(a_2 + 2a_4 x + 3a_6 x^2) := 0, \end{cases}$$

we find that the system of equations has no solution in the interior of A. We now consider the boundary cases (i) $\alpha = 0$, (ii) $\beta = 0$, (iii) $x = 0$, (iv) $x = 1$, and (v) $\alpha + \beta = 1$.

(i) For $\alpha = 0$, we have $L(0, \beta, x) = (a_2 x + a_4 x^2 + a_6 x^3)\beta$. Let

$$\begin{cases} \dfrac{\partial L}{\partial \beta} = x(a_2 + a_4 x + a_6 x^2) := 0 \\ \dfrac{\partial L}{\partial x} = \beta(a_2 + 2a_4 x + 3a_6 x^2) := 0. \end{cases}$$

The above equations have no solution for $0 < x < 1$ and $0 < \beta < 1$. For $\beta = 1$, we have $L(0, 1, x) = a_2 x + a_4 x^2 + a_6 x^3$. Solving $\dfrac{dL}{dx} = a_2 + 2a_4 x + 3a_6 x^2 = 0$, we get $x = \dfrac{-a_4 \pm \sqrt{a_4^2 - 3a_2a_6}}{3a_6}$. We denote $x_0 = \dfrac{-a_4 - \sqrt{a_4^2 - 3a_2a_6}}{3a_6}$ and $x_1 = \dfrac{-a_4 + \sqrt{a_4^2 - 3a_2a_6}}{3a_6}$. We have to discuss the maximum and minimum values of $L(0, 1, x)$ in each of the four cases of Lemma 5.2.4. We assume that $a_4^2 - 3a_2a_6 \ge 0$.

(1) $a_6 < 0$, $a_2 \geq 0$, and $a_2 + a_4 + a_6 \geq 0$

In this case, we have $x_0 > 0$ and $x_1 \leq 0$. It is clear that

$$\min_{0 \leq x \leq 1} L(0,1,x) = L(0,1,0) = 0$$

and

$$\max_{0 \leq x \leq 1} L(0,1,x) = \begin{cases} L(0,1,x_0) & \text{if } x_0 < 1 \\ L(0,1,1) = a_2 + a_4 + a_6 & \text{if } x_0 \geq 1. \end{cases}$$

(2) $a_6 > 0$, $a_4 \geq 0$, and $a_2 \geq 0$

In this case, we have $x_0, x_1 \leq 0$. Hence we have $\frac{dL}{dx} \geq 0$ for $x \geq 0$, and

$$\min_{0 \leq x \leq 1} L(0,1,x) = L(0,1,0) = 0$$

and

$$\max_{0 \leq x \leq 1} L(0,1,x) = L(0,1,1) = a_2 + a_4 + a_6.$$

In Case (3) and (4) of Lemma 5.2.4, we always have $a_6 > 0$, $a_4 < 0$, and $a_2 > 0$. This implies that $x_1 > x_0 > 0$. It is clear that

$$\min_{0 \leq x \leq 1} L(0,1,x) = L(0,1,0) = 0$$

and

$$\max_{0 \leq x \leq 1} L(0,1,x) = \begin{cases} L(0,1,x_0) & \text{if } 0 < x_0 < 1 \\ L(0,1,1) = a_2 + a_4 + a_6 & \text{if } x_0 \geq 1. \end{cases}$$

We find that

$$L(0,1,x_0) = \frac{1}{27a_6^2}(2a_4^3 + 2a_4^2\sqrt{a_4^2 - 3a_2a_6} - 9a_2a_4a_6 - 6a_2a_6\sqrt{a_4^2 - 3a_2a_6}).$$

If $a_4^2 - 3a_2a_6 < 0$, then $L(0,1,x) = a_2x + a_4x^2 + a_6x^3$ is monotone increasing when $a_6 > 0$ and monotone decreasing when $a_6 < 0$. When $L(0,1,x)$ is monotone increasing, we have $\min_{0 \leq x \leq 1} L(0,1,x) = 0$ and $\max_{0 \leq x \leq 1} L(0,1,x) = a_2 + a_4 + a_6$. When $L(0,1,x)$ is monotone decreasing, we have $L(0,1,1) < L(0,1,0) = 0$. This is not possible since $\phi(x) \geq 0$ on $[-1,1]$ implies that $\phi(1) = a_2 + a_4 + a_6 \geq 0$.

(ii) For $\beta = 0$, we have $L(\alpha, 0, x) = (a_2 + a_4 + a_6)\alpha$. It is obvious that $\min_{0\le x\le 1} L(\alpha, 0, x) = L(0, 0, x) = 0$ and $\max_{0\le x\le 1} L(\alpha, 0, x) = a_2+a_4+a_6$, since $a_2+a_4+a_6 \ge 0$ implies that $L(\alpha, 0, x)$ is non-decreasing in α.

(iii) For $x = 0$, we have $L(\alpha, \beta, 0) = (a_2 + a_4 + a_6)\alpha$ which is the same as case (ii).

(iv) For $x = 1$, we have $L(\alpha, \beta, 1) = (a_2+a_4+a_6)(\alpha+\beta) := (a_2+a_4+a_6)\gamma$, where $\gamma = \alpha + \beta$, $0 \le \gamma \le 1$. This is similar to the case (ii).

(v) When $\alpha + \beta = 1$, we have

$$\begin{aligned} L(1-\beta, \beta, x) &= (a_2 + a_4 + a_6)(1-\beta) + (a_2 x + a_4 x^2 + a_6 x^3)\beta \\ &= a_2 + a_4 + a_6 + [a_2(x-1) + a_4(x^2-1) + a_6(x^3-1)]\beta. \end{aligned}$$

Let

$$\begin{cases} \dfrac{\partial L}{\partial \beta} = (x-1)[a_2 + a_4(x+1) + a_6(x^2+x+1)] := 0 \\ \dfrac{\partial L}{\partial x} = \beta[a_2 + 2a_4 x + 3a_6 x^2] := 0, \end{cases}$$

and let $(\beta^*, x^*) \in (0,1) \times (0,1)$ be a solution of the above equations. Then we always have $L(1-\beta^*, \beta^*, x^*) = a_2 + a_4 + a_6$. Summarizing the above results, we have proved Lemma 5.2.5. □

In order to solve the problem (5.2.7), the next lemma is needed. It indicates that the optimal solution of (5.2.7) is achieved on the boundary of the side condition.

Lemma 5.2.6 *Let (a_2, a_4, a_6) satisfy one of the four conditions in Lemma 5.2.4 and $0 < c \le c^*$, where c^* is indicated in Lemma 5.2.5. If $a_2 + a_4 + a_6 > c$, then (5.2.7) is equivalent to*

$$\max_{\xi \in \mathcal{F}_1} \mu_2 \quad \text{subject to } a_2\mu_2 + a_4\mu_4 + a_6\mu_6 = c. \tag{5.2.8}$$

Proof: By Lemma 4.2.7, we have

$$\begin{aligned} \hat{S}_1 = P = \{A_k : &\ A_k \text{ is the interior and boundary of the triangle with vertices} \\ &\ p_0, p_1, p_k, \quad 0 \le k \le 1\}, \end{aligned}$$

where $p_0 = (0,0,0)$, $p_1 = (1,1,1)$, and $p_k = (k, k^2, k^3)$.

Let $\tilde{S}_1 = \{(x,y,z) : (x,y,z) \in \hat{S}_1$ and $a_2 x + a_4 y + a_6 z \le c\} = \{(x,y,z) : (x,y,z) \in A_k$ and $a_2 x + a_4 y + a_6 z \le c$, $0 \le k \le 1\}$. Let $\pi_k : z = -kx + (1+k)y, k \in (0,1)$. Then

$A_k \subseteq \pi_k$. Using (x, y, z) instead of (μ_2, μ_4, μ_6), the side condition of (5.2.8) becomes $a_2x + a_4y + a_6z = c$. Projecting

$$l_k : \begin{cases} z = -kx + (1+k)y \\ a_2x + a_4y + a_6z = c \end{cases}$$

onto the plane xOy, i.e., $z = 0$, we have $-ka_6x + (1+k)a_6y = c - a_2x - a_4y$, and hence $(a_2 - ka_6)x + [a_4 + (1+k)a_6]y = c$. Let $a := a_2 - ka_6$ and $b := a_4 + (1+k)a_6$. Then we have $ax + by = c$. We make the following statement:

$S1$: " $\max\limits_{0 \le x \le 1} \{x\}$ subject to $(x, y) \in P[\tilde{S}_1 \cap A_k]$ (or $(x, y, z) \in \tilde{S}_1 \cap A_k$)
is equivalent to
$\max\limits_{0 \le x \le 1} \{x\}$ subject to $(x, y) \in P[l_k \cap A_k]$ (or $(x, y, z) \in l_k \cap A_k$)".

Note $P[S]$ refers to the projection of set S onto xOy. Note also that, for any possible values of a_2, a_4, a_6, and $0 \le k \le 1$, the possible signs of a and b are (i) $a > 0$, $b > 0$, (ii) $a < 0$, $b > 0$, (iii) $a < 0$, $b < 0$, (iv) $a > 0$, $b < 0$, (v) $a = 0$ or $b = 0$, and (vi) $a = 0$ and $b = 0$.

It is obvious that $S1$ is true in case (i) and case (ii) (see Figure 5.2.3 and Figure 5.2.4). Under the assumption of Lemma 5.2.6, we have $\hat{S}_1 \cap B \ne \phi$, where $B = \{(x, y, z) : a_2x + a_4y + a_6z = c\}$. Hence $a < 0$ and $b < 0$ is not possible. In case (iv), we solve

$$\begin{cases} (a_2 - ka_6)x + [a_4 + (1+k)a_6]y = c \\ y = x, \end{cases}$$

and get

$$x = \frac{c}{a_2 - ka_6 + a_4 + (1+k)a_6} = \frac{c}{a_2 + a_4 + a_6} < 1.$$

Hence $S1$ is also true in this case. (See Figure 5.2.5). For case (v), we have $y = \dfrac{c}{a_4 + (1+k)a_6}$ when $a = 0$, and $x = \dfrac{c}{a_2 - ka_6}$ when $b = 0$. It is clear that in both cases $S1$ is true. (See Figure 5.2.6 and Figure 5.2.7). In case (vi) we have $a = 0$ and $b = 0$. Note that $a = 0$ implies $a_2 = ka_6$, and $b = 0$ implies $a_4 = -(1+k)a_6$. Consequently, we have $a_2 + a_4 + a_6 = ka_6 - (1+k)a_6 + a_6 = 0$. This contradicts the assumption that $a_2 + a_4 + a_6 > c \ge 0$.

We have shown that $S1$ is true for any $k \in [0,1]$ and all the (a_2, a_4, a_6) values that satisfy the conditions of Lemma 5.2.4 and $0 < c \leq c^*$. Hence Lemma 5.2.6 is true. □

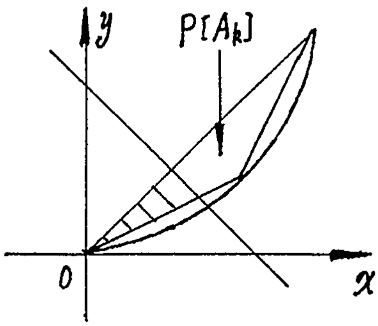

Figure 5.2.3

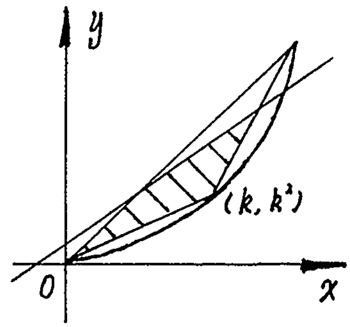

Figure 5.2.4

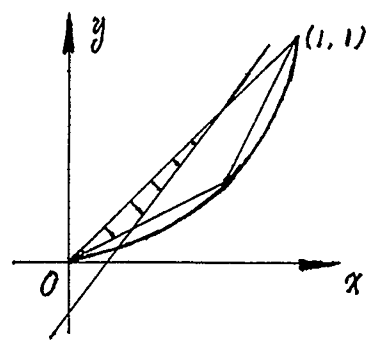

Figure 5.2.5

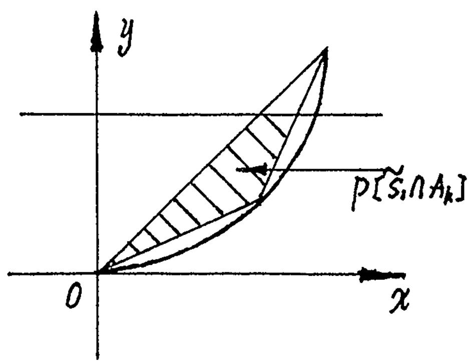

Figure 5.2.6

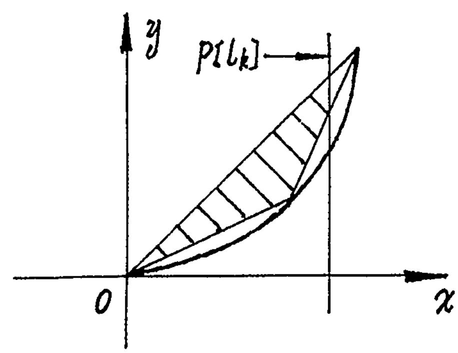

Figure 5.2.7

The consequence of Lemma 5.2.6 is that the inequality constraint in(5.2.7) becomes an equality constraint. Hence, the method of Lagrange multiplier can be used to solve the problem (5.2.7).

For any $\xi \in \mathcal{F}_1$, we have $\mu_2 = \alpha + \beta x$, $\mu_4 = \alpha + \beta x^2$, and $\mu_6 = \alpha + \beta x^3$, where $0 \leq \alpha \leq 1$, $0 \leq \beta \leq 1$, $\alpha + \beta \leq 1$, and $0 \leq x \leq 1$. Let $A = \{(\alpha, \beta, x) : 0 \leq \alpha \leq 1, 0 \leq \beta \leq 1, \alpha + \beta \leq 1$, and $0 \leq x \leq 1\}$, and $\tilde{A} = \{(\alpha, \beta, x) : (\alpha, \beta, x) \in A$, and $a_2(\alpha + \beta x) + a_4(\alpha + \beta x^2) + a_6(\alpha + \beta x^3) \leq c\}$, where (a_2, a_4, a_6) satisfies one of the four conditions in Lemma 5.2.4 and $0 < c \leq c^*$, where c^* is indicated in Lemma 5.2.5. It is obvious that (5.2.7) is equivalent to

$$\max_{(\alpha,\beta,x)\in A} \{\alpha + \beta x\} \text{ subject to } a_2(\alpha + \beta x) + a_4(\alpha + \beta x^2) + a_6(\alpha + \beta x^3) \leq c. \quad (5.2.9)$$

Let $\mu_2^* = \alpha^* + \beta^* x^*$ to be the maximum value of (5.2.9). We have $\mu_2^* = \max_{(\alpha,\beta,x)\in \tilde{A}}$

$\{\alpha + \beta x\}$. Note that $(\mu_2, \mu_4, \mu_6) = (1, 1, 1)$ can be attained by design measure $\xi = \frac{1}{2}\Delta_{\pm 1}$ which corresponds to $(\alpha_0, \beta_0, 1)$ for any α_0, β_0 such that $\alpha_0 + \beta_0 = 1$. Moreover, $(\alpha_0, \beta_0, 1) \in \tilde{A}$ if and only if $a_2 + a_4 + a_6 \leq c$, and in this case we have $\mu_2^* = 1$ and the corresponding design measure is $\xi^* = \frac{1}{2}\Delta_{\pm 1}$.

From now on, we assume that $a_2 + a_4 + a_6 > c$. In this situation, the application of Lemma 5.2.6 will simplify (5.2.9) to be the following:

$$\max_{(\alpha,\beta,x)\in A}\{\alpha + \beta x\} \text{ subject to } a_2(\alpha + \beta x) + a_4(\alpha + \beta x^2) + a_6(\alpha + \beta x^3) = c. \quad (5.2.10)$$

The method of Lagrange multipliers now can be used to solve (5.2.10). Let

$$L(\alpha, \beta, x, \lambda) = \alpha + \beta x + \lambda[(a_2 + a_4 + a_6)\alpha + (a_2 x + a_4 x^2 + a_6 x^3)\beta - c],$$

and

$$\begin{cases} \dfrac{\partial L}{\partial \alpha} = 1 + \lambda(a_2 + a_4 + a_6) := 0 \\ \dfrac{\partial L}{\partial \beta} = x + \lambda x(a_2 + a_4 x + a_6 x^2) := 0 \\ \dfrac{\partial L}{\partial x} = \beta + \lambda\beta(a_2 + 2a_4 x + 3a_6 x^2) := 0 \\ \dfrac{\partial L}{\partial \lambda} = (a_2 + a_4 + a_6)\alpha + (a_2 x + a_4 x^2 + a_6 x^3)\beta - c := 0. \end{cases}$$

We find that the above equations have no solution in the interior of A. We now consider the five boundary cases (i) $\alpha = 0$, (ii) $\beta = 0$, (iii) $x = 0$, (iv) $x = 1$, and (v) $\alpha + \beta = 1$.

(i) $\alpha = 0$

In this case, we have $L(0, \beta, x, \lambda) = \beta x + \lambda[(a_2 x + a_4 x^2 + a_6 x^3)\beta - c]$. Let

$$\begin{cases} \dfrac{\partial L}{\partial \beta} = x + \lambda x(a_2 + a_4 x + a_6 x^2) := 0 & \text{(i-1)} \\ \dfrac{\partial L}{\partial x} = \beta + \lambda\beta(a_2 + 2a_4 x + 3a_6 x^2) := 0 & \text{(i-2)} \\ \dfrac{\partial L}{\partial \lambda} = (a_2 x + a_4 x^2 + a_6 x^3)\beta - c := 0. & \text{(i-3)} \end{cases}$$

For $x \neq 0$, and $\beta \neq 0$, we can rewrite (i-1) and (i-2) as

$$\begin{cases} 1 + \lambda(a_2 + a_4 x + a_6 x^2) = 0 & \text{(i-4)} \\ 1 + \lambda(a_2 + 2a_4 x + 3a_6 x^2) = 0 & \text{(i-5)} \end{cases}$$

Solving (i-4) and (i-5), we find that $x^* = -\dfrac{a_4}{2a_6}$. Put x^* into (i-3) and (i-4), we have

$$\beta^* = \frac{c}{x^*(a_2 + a_4x^* + a_6x^{*2})} = \frac{c}{-\frac{a_4}{2a_6}(a_2 - \frac{a_4^2}{2a_6} + \frac{a_4^2}{4a_6})} = \frac{8a_6^2c}{a_4(a_4^2 - 4a_2a_6)}$$

and

$$\lambda^* = \frac{-1}{a_2 + a_4x^* + a_6x^{*2}} = \frac{-1}{a_2 - \frac{a_4^2}{2a_6} + \frac{a_4^2}{4a_6}} = \frac{4a_6}{a_4^2 - 4a_2a_6}.$$

If $x^* = -\dfrac{a_4}{2a_6} \in (0.1)$, i.e. (1) $a_6 < 0$, $a_4 > 0$, and $a_4 + 2a_6 < 0$, or (2) $a_6 > 0$, $a_4 < 0$, and $a_4 + 2a_6 > 0$, then we always have $\beta^* > 0$ (see Lemma 5.2.4 case (i) and case (iii)). Moreover, if we assume that $\beta^* = \dfrac{8ca_6^2}{a_4(a_4^2 - 4a_2a_6)} \le 1$, then $L(0, \beta^*, x^*, \lambda^*) = \dfrac{4ca_6}{4a_2a_6 - a_4^2}$ is a possible maximum value of $L(\alpha, \beta, x, \lambda)$. The corresponding design measure is $\xi^* = \dfrac{\beta^*}{2}\Delta_{\pm\sqrt{x^*}} + (1 - \beta^*)\Delta_0$.

(ii) $\beta = 0$

When $\beta = 0$, we have $L(\alpha, 0, x, \lambda) = \alpha + \lambda[(a_2 + a_4 + a_6)\alpha - c]$. Let

$$\begin{cases} \dfrac{\partial L}{\partial \alpha} = 1 + \lambda(a_2 + a_4 + a_6) := 0 & \text{(ii-1)} \\ \dfrac{\partial L}{\partial \lambda} = (a_2 + a_4 + a_6)\alpha - c := 0 & \text{(ii-2).} \end{cases}$$

From equations (ii-1) and (ii-2), we find that $\alpha^* = \dfrac{c}{a_2 + a_4 + a_6}$ and $\lambda^* = \dfrac{-1}{a_2 + a_4 + a_6}$. Hence, $L(\alpha^*, 0, x^*, \lambda^*) = \alpha^* = \dfrac{c}{a_2 + a_4 + a_6}$ is a possible maximum value of $L(\alpha, \beta, x, \lambda)$, and the corresponding design measure is $\xi^* = \dfrac{\alpha^*}{2}\Delta_{\pm 1} + (1 - \alpha^*)\Delta_0$.

For case (iii), we have $x = 0$, and $L(\alpha, \beta, 0, \lambda) = \alpha + \lambda[(a_2 + a_4 + a_6)\alpha - c]$ which is the same as case (ii). Similarly, for case (iv), we have $x = 1$, and $L(\alpha, \beta, 1, \lambda) = \alpha + \beta + \lambda[(a_2 + a_4 + a_6)(\alpha + \beta) - c] = \gamma + \lambda[(a_2 + a_4 + a_6)\gamma - c]$, where $\gamma = \alpha + \beta$. This yields the same maximum value as case (ii).

(v) $\alpha + \beta = 1$

In this case, we have $L(1 - \beta, \beta, x, \lambda) = 1 - \beta + \beta x + \lambda[(a_2 + a_4 + a_6)(1 - \beta) + (a_2x + a_4x^2 + a_6x^3)\beta - c] = 1 + (x - 1)\beta + \lambda\{(a_2 + a_4 + a_6) + [a_2(x - 1) + a_4(x^2 -$

$1) + a_6(x^3 - 1)]\beta - c\}$. Let

$$\begin{cases} \dfrac{\partial L}{\partial \beta} = (x-1) + \lambda(x-1)[a_2 + a_4(x+1) + a_6(x^2+x+1)] := 0 & \text{(iii-1)} \\ \dfrac{\partial L}{\partial x} = \beta + \lambda\beta(a_2 + 2a_4 x + 3a_6 x^2) := 0 & \text{(iii-2)} \\ \dfrac{\partial L}{\partial \lambda} = (a_2 + a_4 + a_6) + (x-1)[a_2 + a_4(x+1) + a_6(x^2+x+1)]\beta - c := 0. & \text{(iii-3)} \end{cases}$$

For $x \neq 1$ and $\beta \neq 0$, (iii-1) and (iii-2) can be rewritten as

$$\begin{cases} 1 + \lambda[(a_2 + a_4 + a_6) + (a_4 + a_6)x + a_6 x^2] = 0 & \text{(iii-4)} \\ 1 + \lambda(a_2 + 2a_4 x + 3a_6 x^2) = 0. & \text{(iii-5)} \end{cases}$$

Solving (iii-4) and (iii-5), we find that $x^* = -\dfrac{a_4 + a_6}{2a_6}$,

$$\begin{aligned} \lambda^* &= \frac{-1}{a_2 + 2a_4 x^* + 3a_6 x^{*2}} \\ &= \frac{-1}{a_2 - \dfrac{2a_4(a_4+a_6)}{2a_6} + \dfrac{3(a_4+a_6)^2}{4a_6}} \\ &= \frac{-4a_6}{3a_6^2 + 2a_4 a_6 + 4a_2 a_6 - a_4^2}, \end{aligned}$$

$$\begin{aligned} \beta^* &= \frac{c - (a_2 + a_4 + a_6)}{(x^* - 1)[a_2 + a_4(x^*+1) + a_6(x^{*2} + x^* + 1)]} \\ &= \frac{c - (a_2 + a_4 + a_6)}{(x^* - 1)(a_2 + 2a_4 x^* + 3a_6 x^{*2})} \\ &= \frac{c - (a_2 + a_4 + a_6)}{(-\dfrac{a_4 + a_6}{2a_6} - 1)[a_2 - \dfrac{2a_4(a_4+a_6)}{2a_6} + \dfrac{3a_6(a_4+a_6)^2}{4a_6^2}]} \\ &= \frac{8a_6^2(a_2 + a_4 + a_6 - c)}{(a_4 + 3a_6)(3a_6^2 + 2a_4 a_6 + 4a_2 a_6 - a_4^2)}, \end{aligned}$$

and

$$\alpha^* = 1 - \beta^* = \frac{(a_4 + a_6)(a_6^2 + 4a_2 a_6 - a_4^2) + 8a_6^2 c}{(a_4 + 3a_6)(3a_6^2 + 2a_4 a_6 + 4a_2 a_6 - a_4^2)}.$$

If we have $x^* \in (0,1)$ and $\beta^* \in (0,1)$, then $L(1-\beta^*, \beta^*, x^*, \lambda^*) = 1 + (z^* - 1)\beta^* = \dfrac{4a_6 c - (a_4 + a_6)^2}{3a_6^2 + 2a_4 a_6 + 4a_2 a_6 - a_4^2}$ is a possible maximum value of $L(\alpha, \beta, x, \lambda)$ and the corresponding design measure is $\xi^* = \dfrac{\alpha^*}{2}\Delta_{\pm 1} + \dfrac{\beta^*}{2}\Delta_{\pm\sqrt{x^*}}$.

We now list some conditions:

C1: $a_2 + a_4 + a_6 \le c$.

C2: $x^* = -\dfrac{a_4}{2a_6} \in (0,1)$ and $\beta^* = \dfrac{8a_6^2 c}{a_4(a_4^2 - 4a_2a_6)} \le 1$.

C3: $x^* = -\dfrac{a_4 + a_6}{2a_6} \in (0,1)$ and $\beta^* = \dfrac{8a_6^2(a_2 + a_4 + a_6 - c)}{(a_4 + 3a_6)(3a_6^2 + 2a_4a_6 + 4a_2a_6 - a_4^2)} \in (0,1)$.

Furthermore, we define some notations:

N1: $\alpha_1^* = \dfrac{c}{a_2 + a_4 + a_6}$, $\beta_1^* = x_1^* = 0$, $v_1^* = \dfrac{c}{a_2 + a_4 + a_6}$, and $\xi_1^* = \dfrac{\alpha_1^*}{2}\Delta_{\pm 1} + (1 - \alpha_1^*)\Delta_0$.

N2: $\alpha_2^* = 0$, $\beta_2^* = \dfrac{8a_6^2 c}{a_4(a_4^2 - 4a_2a_6)}$, $x_2^* = -\dfrac{a_4}{2a_6}$, $v_2^* = \dfrac{4a_6 c}{4a_2a_6 - a_4^2}$, and $\xi_2^* = \dfrac{\beta_2^*}{2}\Delta_{\pm\sqrt{x_2^*}} + (1 - \beta_2^*)\Delta_0$.

N3: $\alpha_3^* = \dfrac{(a_4 + a_6)(a_6^2 + 4a_2a_6 - a_4^2) + 8a_6^2 c}{(a_4 + 3a_6)(3a_6^2 + 2a_4a_6 + 4a_2a_6 - a_4^2)}$,

$\beta_3^* = \dfrac{8a_6^2(a_2 + a_4 + a_6 - c)}{(a_4 + 3a_6)(3a_6^2 + 2a_4a_6 + 4a_2a_6 - a_4^2)}$, $x_3^* = -\dfrac{a_4 + a_6}{2a_6}$,

$v_3^* = \dfrac{4a_6 c - (a_4 + a_6)^2}{3a_6^2 + 2a_4a_6 + 4a_2a_6 - a_4^2}$, and $\xi_3^* = \dfrac{\alpha_3^*}{2}\Delta_{\pm 1} + \dfrac{\beta_3^*}{2}\Delta_{\pm\sqrt{x_3^*}}$.

We have proved the next theorem which provides the solution to the problem (5.2.7).

Theorem 5.2.7 *Let (a_2, a_4, a_6) satisfy one of the four conditions in Lemma 5.2.4, and $0 < c \le c^*$, where c^* is indicated in Lemma 5.2.5. The solution to the problem (5.2.7) is provided as follows:*

(i) If C1 is true, then $\xi^* = \dfrac{1}{2}\Delta_{\pm 1}$ and $\mu_2^* = 1$.

(ii) If C1 is not true, but C2 and C3 are true, then $\mu_2^* = \max\{v_1^*, v_2^*, v_3^*\} := v_i^*$, and the corresponding design measure is ξ_i^*, $i \in \{1,2,3\}$.

(iii) If C1 and C2 are not true, but C3 is true, then $\mu_2^* = \max\{v_1^*, v_3^*\} := v_j^*$, and the corresponding design measure is ξ_j^*, $j \in \{1,3\}$.

(iv) If C1 and C3 are not true, but C2 is true, then $\mu_2^* = \max\{v_1^*, v_2^*\} := v_k^*$, and the corresponding design measure is ξ_k^*, $k \in \{1,2\}$.

(v) If C1, C2, and C3 are not true, then $\mu_2^* = v_1^*$, and the corresponding design measure is $\xi_1^* = \dfrac{\alpha_1^*}{2}\Delta_{\pm 1} + (1 - \alpha_1^*)\Delta_0$.

The following corollary is obvious.

Corollary 5.2.8 *Let $\phi(x) = a_2x^2 + a_4x^4 + a_6x^6 \geq 0$ on $[-1, 1]$, and $a_2 + a_4 + a_6 = 0$. Then the solution of the problem (5.2.7) is $\xi^* = \frac{1}{2}\Delta_{\pm 1}$, and $\mu_2^* = 1$.*

Remark 2. Note that $a_2 + a_4 + a_6 = 0$ if and only if $\phi(\pm 1) = 0$, i.e. there is no violation at ± 1. Note also that the usual optimal design is supported at ± 1. The implication of Corollary 5.2.8 is that the bounded bias optimal design will be the same as usual optimal design if the regression model is not violated at ± 1.

Remark 3. For D-, A-, and Q-optimality, the optimal design we get in Theorem 5.2.7 is not only the optimal design over $\mathcal{F}_s$ but also the optimal design over $\mathcal{F}$ as well. This fact follows by Lemma 5.1.2. For G-optimality we can only say that the result in Theorem 5.2.7 is only the optimal design over $\mathcal{F}_s$.

Let us define the efficiency of bounded bias optimal design ξ^* as the following

$$e(\xi^*) = \frac{\mathcal{L}(\xi^0)}{\mathcal{L}(\xi^*)},$$

where ξ^0 is the usual optimal design and $\mathcal{L} \in \{\mathcal{L}_D, \mathcal{L}_A, \mathcal{L}_Q, \mathcal{L}_G\}$. It is clear that $e(\xi^*)$ is "increasing in μ_2^*". Consider the situation that $\mu_2^* = \dfrac{c}{a_2 + a_4 + a_6}$. For fixed c, $e(\xi^*)$ is decreasing when $\phi(1) = a_2 + a_4 + a_6$ is increasing; for fixed $\phi(1) = a_2 + a_4 + a_6$, $e(\xi^*)$ is decreasing when c is decreasing. This implies that we lose efficiency (smaller $e(\xi^*)$) to gain more protection on the possible bias (smaller c) when the amount of model violations at ± 1 are fixed.

5.3 Approximately Linear Regression Model with Norm 1 Bounded and Antisymmetric Contamination Functions

In this section, we consider again the one dimensional linear regression model and the class of symmetric design measures on $[-1,1]$. But for the class of contamination functions Ψ we choose

$$\Psi_a(\phi) = \{\psi : \psi(0) = \psi'(0) = 0, \psi(-x) = -\psi(x), |\psi(x)| \leq |\phi(x)|$$
$$\text{on } [-1,1], \text{ and } \phi(x) = \textstyle\sum_{i=1}^{p} a_{2i+1}x^{2i+1} \geq 0 \text{ on } [0,1]\}.$$

In the case of antisymmetric contamination, we have

$$\begin{aligned}
E[\underset{\sim}{\hat{\theta}} - \underset{\sim}{\theta}] &= B^{-1}(\xi)\ \underset{\sim}{b}\,(\psi,\xi) \\
&= \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{\mu_2} \end{pmatrix} \begin{pmatrix} \int \psi d\xi \\ \int x\psi d\xi \end{pmatrix} \\
&= \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{\mu_2} \end{pmatrix} \begin{pmatrix} 0 \\ \int x\psi d\xi \end{pmatrix} \\
&= \begin{pmatrix} 0 \\ \frac{1}{\mu_2} \int x\psi d\xi \end{pmatrix}
\end{aligned}$$

and

$$\begin{aligned}
||E[\underset{\sim}{\hat{\theta}} - \underset{\sim}{\theta}]|| &= |\frac{1}{\mu_2} \int x\psi(x)\xi(x)| \\
&\leq \frac{1}{\mu_2} \int |x\psi(x)| d\xi(x) \\
&\leq \frac{1}{\mu_2} \int |x\phi(x)| d\xi(x) \\
&= \frac{1}{\mu_2} \int x\phi(x) d\xi(x).
\end{aligned}$$

On the other hand, we have $\phi(x) \in \Psi_a(\phi)$. Hence

$$\begin{aligned}\max_{\psi\in\Psi_a} \|E[\hat{\underset{\sim}{\theta}} - \underset{\sim}{\theta}]\| &= \frac{1}{\mu_2} \int x\phi(x)d\xi(x) \\ &= \frac{1}{\mu_2} \int \textstyle\sum_{i=1}^p a_{2i+1}x^{2i+2}d\xi(x) \\ &= \frac{1}{\mu_2} \textstyle\sum_{i=1}^p a_{2i+1}\mu_{2i+2}.\end{aligned}$$

Under the case of antisymmetric contamination, the original problem (5.1.6) now becomes

$$\max_{\xi\in\mathcal{F}_s} \mu_2 \quad \text{subject to } \frac{1}{\mu_2}\sum_{i=1}^{p} a_{2i+1}\mu_{2i+2} \le c, \tag{5.3.1}$$

for some a_{2i+1}'s such that $\phi(x) = \sum_{i=1}^p a_{2i+1}x^{2i+1} \ge 0$ on $[0,1]$.

We are seeking the solution to the problem (5.3.1) for some specific choice of $\phi(x)$. The situation is very similar to Section 5.2. Hence, we only consider one case, namely when $\phi(x) = a_3x^3 + a_5x^5$. In this case, the problem (5.3.1) becomes

$$\max_{\xi\in\mathcal{F}_s} \mu_2 \quad \text{subject to } a_3\mu_4 + a_5\mu_6 \le c\mu_2. \tag{5.3.2}$$

It is clear that $\phi(x) \ge 0$ on $[0,1]$ if and only if $\phi_1(x) := a_3 + a_5x^2 \ge 0$ on $[0,1]$ if and only if (i) $a_5 > 0$, $a_3 \ge 0$; or (ii) $a_5 < 0$, $a_3 + a_5 \ge 0$. Again we define $c_* = \inf_{\xi\in\mathcal{F}_S}\{\frac{1}{\mu_2}(a_3\mu_4 + a_5\mu_6)\}$ and $c^* = \max_{\xi\in\mathcal{F}_S}\{\frac{1}{\mu_2}(a_3\mu_4 + a_5\mu_6)\}$. It follows by Theorem 4.2.6 that we have

$$c_* = \inf_{(\alpha,\beta,x)\in A}\{\frac{a_3(\alpha+\beta x^2) + a_5(\alpha+\beta x^3)}{\alpha+\beta x}\}$$

and

$$c^* = \max_{(\alpha,\beta,x)\in A}\{\frac{a_3(\alpha+\beta x^2) + a_5(\alpha+\beta x^3)}{\alpha+\beta x}\},$$

where $A = \{(\alpha,\beta,x) : 0 \le \alpha \le 1, 0 \le \beta \le 1, \alpha+\beta \le 1, \text{ and } 0 \le x \le 1\}$. For c_*, we use "inf" instead of "min", because c_* is not achievable. We define

$$L(\alpha,\beta,x) = \frac{a_3(\alpha+\beta x^2) + a_5(\alpha+\beta x^3)}{\alpha+\beta x} = \frac{(a_3+a_5)\alpha + (a_3x^2 + a_5x^3)\beta}{\alpha+\beta x}.$$

We exclude the cases when $\alpha = 0$, $\beta = 0$, or $\alpha = 0, x = 0$, where $L(\alpha,\beta,x)$ is undefined. Let

$$
\begin{cases}
\begin{aligned}
\frac{\partial L(\alpha,\beta,x)}{\partial \alpha} &= \frac{1}{(\alpha+\beta x)^2}\{(\alpha+\beta x)(a_3+a_5) - [(a_3+a_5)\alpha + (a_3x^2+a_5x^3)\beta]\} \\
&= \frac{1}{(\alpha+\beta x)^2}[\beta x(a_3+a_5) - \beta x^2(a_3+a_5x)] \\
&= \frac{\beta x}{(\alpha+\beta x)^2}(a_3+a_5-a_3x-a_5x^2) \\
&= \frac{\beta x(1-x)}{(\alpha+\beta x)^2}(a_3+a_5+a_5x) \\
&:= 0 && \text{(i-1)} \\
\frac{\partial L(\alpha,\beta,x)}{\partial \beta} &= \frac{1}{(\alpha+\beta x)^2}\{(\alpha+\beta x)(a_3x^2+a_5x^3) \\
&\quad -[(a_3+a_5)\alpha + (a_3x^2+a_5x^3)\beta]x\} \\
&= \frac{x}{(\alpha+\beta x)^2}[\alpha(a_3x+a_5x^2) - (a_3+a_5)\alpha] \\
&= \frac{\alpha x(x-1)}{(\alpha+\beta x)^2}(a_3+a_5+a_5x) \\
&:= 0 && \text{(i-2)} \\
\frac{\partial L(\alpha,\beta,x)}{\partial x} &= \frac{1}{(\alpha+\beta x)^2}\{(\alpha+\beta x)\beta(2a_3x+3a_5x^2) \\
&\quad -[(a_3+a_5)\alpha + (a_3x^2+a_5x^3)\beta]\beta\} \\
&= \frac{\beta}{(\alpha+\beta x)^2}(2a_3\alpha x + 3a_5\alpha x^2 + a_3\beta x^2 + 2a_5\beta x^3 - a_3\alpha - a_5\alpha) \\
&:= 0. && \text{(i-3)}
\end{aligned}
\end{cases}
$$

From (i-1) and (i-2), we find $x^* = -\frac{a_3+a_5}{a_5}$. Let $x^* \in (0,1)$, we get $a_5 < 0$, $a_3+a_5 > 0$, and $a_3 + 2a_5 < 0$. Put x^* into (i-3), we have

$$
\begin{aligned}
&2a_3\alpha(-\frac{a_3+a_5}{a_5}) + (3a_5\alpha + a_3\beta)(-\frac{a_3+a_5}{a_5})^2 + 2a_5\beta(-\frac{a_3+a_5}{a_5})^3 - (a_3+a_5)\alpha \\
&= \frac{(a_3+a_5)}{a_5^2}(-2a_3a_5\alpha + 3a_3a_5\alpha + 3a_5^2\alpha + a_3^2\beta + a_3a_5\beta - 2a_3^2\beta - 4a_3a_5\beta - 2a_5^2\beta - a_5^2\alpha) \\
&= \frac{(a_3+a_5)}{a_5^2}[a_5(a_3+2a_5)\alpha - (a_3^2+3a_3a_5+2a_5^2)\beta] \\
&= \frac{1}{a_5^2}(a_3+a_5)(a_3+2a_5)[a_5\alpha - (a_3+a_5)\beta] \neq 0,
\end{aligned}
$$

since $a_3+a_5 > 0$, $a_3+2a_5 < 0$, and $a_5\alpha - (a_3+a_5)\beta < 0$ (Note that $a_5 < 0$).

Therefore the system of equations (i-1), (i-2), and (i-3) has no solution in the interior of A. There are five boundary cases to be considered.

(i) $\alpha = 0$

In this case, we have $L(0,\beta,x) = (a_3 + a_5 x)x$. Let $\frac{dL}{dx} = a_3 + 2a_5 x = 0$; then we have $x^* = -\frac{a_3}{2a_5}$. If $x^* = -\frac{a_3}{2a_5} \in (0,1)$, then $L(0,\beta,x^*) = (a_3 + a_5 \cdot \frac{-a_3}{2a_5})(-\frac{a_3}{2a_5}) = -\frac{a_3^2}{4a_5}$ is a possible maximum value of $L(\alpha,\beta,x)$.

For the cases (ii) $\beta = 0$, (iii) $x = 0$, and (iv) $x = 1$, we have $L(\alpha,0,x) = L(\alpha,\beta,0) = L(\alpha,\beta,1) = a_3 + a_5$.

(v) $\alpha + \beta = 1$

In this case, we have

$$\begin{aligned} L(1-\beta,\beta,x) &= \frac{(a_3+a_5)(1-\beta) + (a_3x^2 + a_5x^3)\beta}{1-\beta+\beta x} \\ &= \frac{(a_3+a_5) + [a_3(x^2-1) + a_5(x^3-1)]\beta}{1+\beta(x-1)}. \end{aligned}$$

. Let

$$\begin{aligned} \frac{\partial L}{\partial \beta} &= \frac{1}{[1+\beta(x-1)]^2}\{[1+\beta(x-1)][a_3(x^2-1)+a_5(x^3-1)] \\ &\quad -[(a_3+a_5)+(a_3(x^2-1)+a_5(x^3-1))\beta](x-1)\} \\ &\quad \frac{x(x-1)(a_3+a_5+a_5x)}{[1+\beta(z-1)]^2} := 0 \end{aligned} \qquad \text{(ii-1)}$$

and

$$\begin{aligned} \frac{\partial L}{\partial x} &= \frac{1}{[1+\beta(x-1)]^2}\{[1+\beta(x-1)]\beta[2a_3x+3a_5x^2] \\ &\quad -[(a_3+a_5)+(a_3(x^2-1)+a_5(x^3-1))\beta]\beta\} \\ &= \frac{\beta}{[1+\beta(x-1)]^2}[(1-\beta)(2a_3x+3a_5x^2-a_3-a_5)+\beta(a_3+2a_5x)x^2] := 0. \end{aligned} \qquad \text{(ii-2)}$$

From (ii-1), we find $x^* = -\frac{a_3+a_5}{a_5}$. Let $x^* \in (0,1)$; then we get $a_5 < 0$, $a_3 + a_5 > 0$, and $a_3 + 2a_5 < 0$. Putting x^* into (ii-2), we have

$$\begin{aligned}
&(1-\beta)[2a_3(-\frac{a_3+a_5}{a_5}) + 3a_5(-\frac{a_3+a_5}{a_5})^2 - (a_3+a_5)] \\
&\qquad +\beta[a_3 + 2a_5(-\frac{a_3+a_5}{a_5})](-\frac{a_3+a_5}{a_5})^2 \\
&= \frac{(1-\beta)(a_3+a_5)}{a_5}[-2a_3 + 3(a_3+a_5) - a_5] + \frac{\beta(a_3+a_5)^2}{a_5^2}(a_3 - 2a_3 - 2a_5) \\
&= \frac{(1-\beta)(a_3+a_5)}{a_5}(a_3+2a_5) - \frac{\beta(a_3+a_5)^2}{a_5^2}(a_3+2a_5) \\
&= \frac{1}{a_5^2}(a_3+a_5)(a_3+2a_5)[a_5(1-\beta) - (a_3+a_5)\beta] \\
&= \frac{1}{a_5^2}(a_3+a_5)(a_3+2a_5)[a_5 - (a_3+2a_5)\beta] := 0. \qquad \text{(ii-3)}
\end{aligned}$$

From (ii-3), we find $\beta^* = \frac{a_5}{a_3+2a_5}$. Let $\beta^* < 1$; then we get $a_5 > a_3 + 2a_5$, i.e., $a_3 + a_5 < 0$, which is a contradiction. Note that the case of $\beta^* = 1$ corresponds to $x^* = 0$ and $\alpha^* = 0$ which has been excluded. Hence, there is no extreme value in case (v).

Note that

$$(a_3+a_5) - (-\frac{a_3^2}{4a_5}) = \frac{4a_3a_5 + 4a_5^2 + a_3^2}{4a_5} = \frac{(a_3+2a_5)^2}{4a_5} < 0$$

if $a_5 < 0$. Hence, we have the conclusion,

$$c^* = \begin{cases} -\frac{a_3^2}{4a_5} & \text{if } x^* = -\frac{a_3}{2a_5} \in (0,1] \text{ i.e. } a_5 < 0 \text{ and } a_3 + 2a_5 \le 0 \\ a_3 + a_5 & \text{otherwise.} \end{cases}$$

Note also, we have $L(\alpha,\beta,x) \ge 0$, and $\lim_{x\to 0} L(0,1,x) = \lim_{x\to 0}(a_3 + a_5x)x = 0$. Hence $c_* = 0$. We have proved the next lemma.

Lemma 5.3.1 *(i) $\varphi(x) = a_3x^3 + a_5x^5 \ge 0$ on $[0,1]$ if and only if one of the following is true: (a) $a_5 > 0$, and $a_3 \ge 0$; (b) $a_5 < 0$, and $a_3 + a_5 \ge 0$.*

(ii) When $\varphi(x) \ge 0$ on $[0,1]$, we have $c_* = 0$, and

$$c^* = \begin{cases} a_3 + a_5 & \text{if } a_5 > 0, a_3 \ge 0, \text{ or } a_5 < 0, a_3 + 2a_5 > 0 \\ \\ -\frac{a_3^2}{4a_5} & \text{if } a_5 < 0, a_3 + a_5 \ge 0, \text{ and } a_3 + 2a_5 \le 0. \end{cases}$$

Similar to Lemma 5.2.6, we indicate that the optimal solution to the problem (5.3.2) is achieved on the boundary of the side condition of (5.3.2). We formally state this fact as the next Lemma.

Lemma 5.3.2 *Let (a_3, a_5) satisfy one of the two conditions (a) and (b) in Lemma 5.3.1, and $0 < c \le c^*$ where c^* is indicated in Lemma 5.3.1 (ii). If $a_3 + a_5 > c$, then (5.3.2) is equivalent to*

$$\max_{\xi \in \mathcal{F}_1} \mu_2 \quad \textit{subject to } a_3\mu_4 + a_5\mu_6 = c\mu_2. \tag{5.3.3}$$

Proof: By Lemma 4.2.7, we have

$$\hat{S}_1 = P = \{A_k : A_k \text{ is the interior and boundary of the triangle with vertices } p_0, p_1, p_k,\ 0 \le k \le 1\},$$

where $p_0 = (0,0,0)$, $p_1 = (1,1,1)$, and $p_k = (k, k^2, k^3)$. Let $\tilde{S}_1 = \{(x,y,z) : (x,y,z) \in \hat{S}_1$ and $a_3y + a_5z \le cx\} = \{(x,y,z) : (x,y,z) \in A_k$ and $a_3y + a_5z \le cx,\ 0 \le k \le 1\}$. Let $\pi_k : z = -kx + (1+k)y$, $k \in (0,1)$. Then we have $A_k \subseteq \pi_k$. Using (x,y,z) instead of (μ_2, μ_4, μ_6), the side condition of (5.3.3) becomes $a_3y + a_5z = cx$. Projecting

$$l_k : \begin{cases} z = -kx + (1+k)y \\ a_3y + a_5z = cx \end{cases}$$

onto the plane xOy, i.e., $z = 0$, we have $a_3y + a_5[-kx + (1+k)y] = cx$ and hence $[a_3 + a_5(1+k)]y = (c + a_5k)x$. Let $a = c + a_5k$, $b = a_3 + a_5(1+k)$. Then we have $P[l_k] : ax = by$, where $P[S]$ refers to the projection of set S onto the plane xOy. We make a statement as follows:

$$S2 : \text{“} \max_{0 \le x \le 1} \{x\} \text{ subject to } (x,y) \in P[\tilde{S}_1 \cap A_k] \text{ (or } (x,y,z) \in \tilde{S}_1 \cap A_k)$$

is equivalent to

$$\max_{0 \le x \le 1} \{x\} \text{ subject to } (x,y) \in P[l_k \cap A_k] \text{ (or } (x,y,z) \in l_k \cap A_k)\text{”}.$$

If for some $k \in [0,1]$, we have one of the following: (i) $a = 0$, (ii) $b = 0$, (iii) $ab < 0$, or (iv) $ab > 0$ but the slope of $P[l_k]$ is greater than 1 or less than k. Then

$P[l_k] \cap P[A_k] = \{(0,0)\}$. Hence $\tilde{S}_1 \cap A_k = \{(0,0,0)\}$. In this case, it is trivial that $S2$ is true. If for some $k \in [0,1]$, we have (v) $ab > 0$ and the slope of $P[l_k]$ is greater or equal to k and less than 1, again $S2$ is obviously true (see Figure 5.3.1). Note that $c + a_5 k = a_3 + a_5(1+k)$ implies that $c = a_3 + a_5$ which contradicts to $a_3 + a_5 > c$. Hence $a = b$ will not occur, i.e. the slop of $P[l_k]$ will never be 1. Also, $a = b = 0$ will never occur.

We have shown that $S2$ is true for any $k \in [0,1]$. Therefore, we have proved Lemma 5.3.2. □

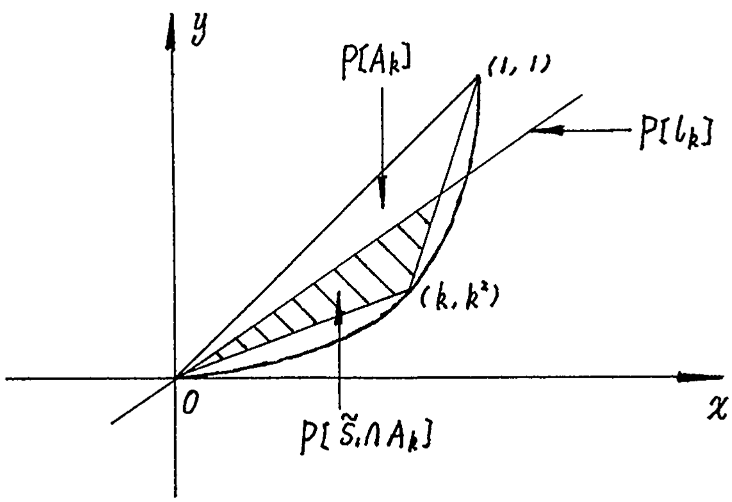

Figure 5.3.1

We are now going to find the solution to the problem (5.3.2). Similar to the proof of Theorem 5.2.7, we have $\mu_2^* = 1$ if and only if $a_3 + a_5 \le c$. The corresponding design measure is $\xi^* = \dfrac{1}{2}\Delta_{\pm 1}$.

From now on, we assume that $a_3 + a_5 > 0$. In this case, the application of Theorem 4.2.6 and Lemma 5.3.2 will simplify the problem (5.3.2) to be the following:

$$\max_{(\alpha,\beta,x)\in A} \{\alpha + \beta x\} \quad \text{subject to } c(\alpha + \beta x) - a_3(\alpha + \beta x^2) - a_5(\alpha + \beta x^3) = 0. \quad (5.3.4)$$

The method of Lagrange multipliers now can be used to solve (5.3.4). Let

$$\begin{aligned} L(\alpha,\beta,x,\lambda) &= \alpha + \beta x + \lambda[c(\alpha + \beta x) - a_3(\alpha + \beta x^2) - a_5(\alpha + \beta x^3)] \\ &= \alpha + \beta x + \lambda[(c - a_3 - a_5)\alpha + (cx - a_3 x^2 - a_5 x^3)\beta], \end{aligned}$$

and

$$\begin{cases} \dfrac{\partial L}{\partial \alpha} &= 1 + \lambda(c - a_3 - a_5) := 0 \\ \dfrac{\partial L}{\partial \beta} &= x + \lambda x(c - a_3 x - a_5 x^2) := 0 \\ \dfrac{\partial L}{\partial x} &= \beta + \lambda\beta(c - 2a_3 x - 3a_5 x^2) := 0 \\ \dfrac{\partial L}{\partial \lambda} &= (c - a_3 - a_5)\alpha + (cx - a_3 x^2 - a_5 x^3) := 0. \end{cases}$$

We find that the above system of equations has no solution in the interior of A. There are five boundary cases to be considered. They are (i) $\alpha = 0$, (ii) $\beta = 0$, (iii) $x = 0$, (iv) $x = 1$, and (v) $\alpha + \beta = 1$.

(i) $\alpha = 0$

In this case, we have $L(0, \beta, x, \lambda) = \beta x + \lambda\beta(cx - a_3 x^2 - a_5 x^3)$. Let

$$\begin{cases} \dfrac{\partial L}{\partial \beta} = x + \lambda x(c - a_3 x - a_5 x^2) := 0 & \text{(i-1)} \\ \dfrac{\partial L}{\partial x} = \beta + \lambda\beta(c - 2a_3 x - 3a_5 x^2) := 0 & \text{(i-2)} \\ \dfrac{\partial L}{\partial \lambda} = \beta x(c - a_3 x - a_5 x^2) := 0. & \text{(i-3)} \end{cases}$$

Note that (i-1) and (i-3) contradict each other when $\beta \neq 0$ or $z \neq 0$. For the special case when $\alpha = 0$ and $\beta = 1$, we have $L(0, 1, x, \lambda) = x + \lambda(cx - a_3 x^2 - a_5 x^3)$, and

$$\begin{cases} \dfrac{\partial L}{\partial x} = 1 + \lambda(c - 2a_3 x - 3a_5 x^2) := 0 & \text{(i-4)} \\ \dfrac{\partial L}{\partial \lambda} = x(c - a_3 x - a_5 x^2) := 0. & \text{(i-5)} \end{cases}$$

Solving (i-4) and (i-5), we find $x_1 = \dfrac{-a_3 - \sqrt{a_3^2 + 4a_5 c}}{2a_5}$ and $x_2 = \dfrac{-a_3 + \sqrt{a_3^2 + 4a_5 c}}{2a_5}$. When (a_3, a_5) and c satisfy the restriction in Lemma 5.3.1, we always have $x_2 \in [0, 1]$, and $x_1 < 0$, or $x_1 > 1$. Let $x^* = x_2$. Then $L(0, 1, x^*, \lambda) = x^*$ is a possible maximum value of $L(\alpha, \beta, x, \lambda)$.

For the cases: (ii) $\beta = 0$, (iii) $x = 0$, and (iv) $x = 1$, we have $L(\alpha, 0, x, \lambda) = L(\alpha, \beta, 0, \lambda) = \alpha + \lambda(c - a_3 - a_5)\alpha$ and $L(\alpha, \beta, 1, \lambda) = \alpha + \beta + \lambda(c - a_3 - a_5)(\alpha + \beta)$. We find no maximum values in these cases.

(v) $\alpha + \beta = 1$

In this case, we have $L(1 - \beta, \beta, x, \lambda) = 1 - \beta + \beta x + \lambda[(c - a_3 - a_5)(1 - \beta) + (cx - a_3 x^2 - a_5 x^3)\beta] = 1 + \beta(x - 1) + \lambda\{(c - a_3 - a_5) + \beta(x - 1)[c - a_3(x + 1) - a_5(x^2 + x + 1)]\}$.

Let

$$\begin{cases} \dfrac{\partial L}{\partial \beta} = (x-1) + \lambda(x-1)[c - a_3(x+1) - a_5(x^2+x+1)] := 0 & \text{(ii-1)} \\ \dfrac{\partial L}{\partial x} = \beta + \lambda\beta(c - 2a_3x - 3a_5x^2) := 0 & \text{(ii-2)} \\ \dfrac{\partial L}{\partial \lambda} = (c - a_3 - a_5) + \beta(x-1)[c - a_3(x+1) - a_5(x^2+x+1)] := 0. & \text{(ii-3)} \end{cases}$$

From (ii-1) and (ii-2), we get $c - 2a_3x - 3a_5x^2 = c - a_3(x+1) - a_5(x^2+x+1)$, and hence $(x-1)(x + \frac{a_3+a_5}{2a_5}) = 0$. Let $x^* = -\frac{a_3+a_5}{2a_5} \in (0,1)$. Then we have $a_5 < 0$ and $a_3 + 3a_5 < 0$. Putting x^* into (ii-3), we find

$$\begin{aligned} \beta^* &= \frac{a_3+a_5-c}{(x^*-1)[c - a_3(x^*+1) - a_5(x^{*2}+x^*+1)]} \\ &= \frac{a_3+a_5-c}{(x^*-1)[c - 2a_3x^* - 3a_5x^{*2}]} \\ &= \frac{a_3+a_5-c}{(-\frac{a_3+a_5}{2a_5} - 1)[c - 2a_3(-\frac{a_3+a_5}{2a_5}) - 3a_5(-\frac{a_3+a_5}{2a_5})^2]} \\ &= \frac{8a_5^2(c-a_3-a_5)}{(a_3+3a_5)(a_3^2 - 2a_3a_5 - 3a_5^2 + 4a_5c)}. \end{aligned}$$

It is easy to see that $\beta^* > 0$. If we also have $\beta^* < 1$, then

$$\begin{aligned} L(1-\beta^*, \beta^*, x^*, \lambda^*) &= 1 - \beta^* + \beta^* x^* \\ &= \frac{(a_3+a_5)^2}{a_3^2 - 2a_3a_5 - 3a_5^2 + 4a_5c}, \end{aligned}$$

is also a possible maximum value of $L(\alpha, \beta, x, \lambda)$.

We introduce the following conditions:

C1: $a_3 + a_5 \le c$.

C2: $x^* = -\frac{a_3+a_5}{2a_5} \in (0,1)$ and $\beta^* = \frac{8a_5^2(c-a_3-a_5)}{(a_3+3a_5)(a_3^2-2a_3a_5-3a_5^2+4a_5c)} < 1$.

Furthermore, we define the following notations:

N1: $\alpha_1^* = 0$, $\beta_1^* = 1$, $x_1^* = \frac{-a_3 + \sqrt{a_3^2+4a_5c}}{2a_5}$, $v_1^* = \frac{-a_3+\sqrt{a_3^2+4a_5c}}{2a_5}$, and $\xi_1^* = \frac{1}{2}\Delta_{\pm\sqrt{x_1^*}}$.

N2: $\alpha_2^* = \frac{(a_3+a_5)(a_3^2 - a_5^2 + 4a_5c)}{(a_3+3a_5)(a_3^2-2a_3a_5-3a_5^2+4a_5c)}$, $\beta_2^* = \frac{8a_5^2(c-a_3-a_5)}{(a_3+3a_5)(a_3^2-2a_3a_5-3a_5^2+4a_5c)}$, $x_2^* = -\frac{a_3+a_5}{2a_5}$, $v_2^* = \frac{(a_3+a_5)^2}{a_3^2-2a_3a_5-3a_5^2+4a_5c}$, and $\xi_2^* = \frac{\alpha_2^*}{2}\Delta_{\pm 1} + \frac{\beta_2^*}{2}\Delta_{\pm\sqrt{x_2^*}}$.

We have proved the next theorem which provides the solution to the problem (5.3.2).

Theorem 5.3.3 *Let (a_3, a_5) satisfy one of the two conditions (a) and (b) in Lemma 5.3.1 (i), and $0 < c \leq c^*$, where c^* is indicated in Lemma 5.3.1 (ii). The solution to the problem (5.3.2) is the following:*

(i) If C1 is true, then $\xi^* = \frac{1}{2}\Delta_{\pm 1}$ and $\mu_2^* = 1$.

(ii) If C1 is not true, but C2 is true, then $\mu_2^* = \max\{v_1^*, v_2^*\} := v_i^*$, and the corresponding design measure is ξ_i, $i \in \{1, 2\}$.

(iii) If both C1 and C2 are not true, then $\mu_2^* = v_1^*$, and the corresponding design measure is $\xi_1^* = \dfrac{1}{2}\Delta_{\pm\sqrt{x_1^*}}$.

The following corollary is obviously true.

Corollary 5.3.4 *Let $\varphi(x) = a_3x^3 + a_5x^5 \geq 0$ on $[0,1]$ and $a_3 + a_5 = 0$. Then the solution of the problem (5.3.2) is $\xi^* = \dfrac{1}{2}\Delta_{\pm 1}$, and $\mu_2^* = 1$.*

5.4 Approximately Linear Regression Models with Norm 2 bounded Contamination Functions

In this section, we consider the problem of bounded bias optimal designs of approximately linear and multiple linear regression models with different class of $\mathcal{F}$ and Ψ as we discussed in Section 5.2 and Section 5.3. We define

$$\Psi = \{\psi(x) : \int_S \psi^2(x)dx \leq \eta^2, \int_S \underset{\sim}{f}(x)\psi(x)dx = \underset{\sim}{0}\},$$

where η is a preassigned constant and the side condition $\int_S \underset{\sim}{f}(x)\psi(x)dx = \underset{\sim}{0}$ is to insure the identifiability of the parameters to be estimated. Also, we define

$$\mathcal{F} = \{\xi(x) : \frac{d\xi(x)}{dx} = m(x), \int_S m(x)dx = 1, m(x) \geq 0, \text{ and } m(-x) = m(x) \text{ on } S\}.$$

In this section, we choose $S = [-\frac{1}{2}, \frac{1}{2}]$.

Let $\hat{y} = \hat{\underset{\sim}{\theta}}^T \underset{\sim}{f}(x)$ to be the estimator of $\underset{\sim}{\theta}^T \underset{\sim}{f}(x)$, where $\underset{\sim}{\theta}^T = (\theta_0, \theta_1)$, $\hat{\underset{\sim}{\theta}}^T = (\hat{\theta}_0, \hat{\theta}_1)$, and $\underset{\sim}{f}^T(x) = (1, x)$. Under the model (5.1.3), we know that

$$MSE(\hat{y}) = \frac{\sigma^2}{n} \underset{\sim}{f}^T(x)B^{-1}(\xi)\underset{\sim}{f}(x) + \underset{\sim}{f}^T(x)B^{-1}(\xi)\underset{\sim}{b}(\psi,\xi)\underset{\sim}{b}^T(\psi,\xi)B^{-1}(\xi)\underset{\sim}{f}(x),$$

where $B(\xi) = \int_S \underset{\sim}{f}(x)\underset{\sim}{f}^T(x)d\xi(x)$, $\underset{\sim}{b}(\psi,\xi) = \int_S \underset{\sim}{f}(x)\psi(x)d\xi(x)$, and

$$\int_S MSE(\hat{y})dx = \frac{\sigma^2}{n} trAB^{-1}(\xi) + \underset{\sim}{b}^T(\psi,\xi)B^{-1}(\xi)AB^{-1}(\xi)\underset{\sim}{b}(\psi,\xi),$$

with $A = \int_S \underset{\sim}{f}(x)\underset{\sim}{f}^T(x)dx$.

For the approximately linear regression model, we have

$$B(\xi) = \begin{pmatrix} 1 & 0 \\ 0 & \mu_2 \end{pmatrix}, B^{-1}(\xi) = \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{\mu_2} \end{pmatrix}, A = \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{12} \end{pmatrix}, \text{ and}$$

$\underset{\sim}{b}^T(\psi,\xi) = (\int_S \psi(x)d\xi(x), \int_S x\psi(x)d\xi(x))$. Denote

$$\begin{aligned} V(\xi) = \frac{\sigma^2}{n} trAB^{-1}(\xi) &= \frac{\sigma^2}{n} tr \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{12} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{\mu_2} \end{pmatrix} \\ &= \frac{\sigma^2}{n} tr \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{12\mu_2} \end{pmatrix} \\ &= \frac{\sigma^2}{n}(1 + \frac{1}{12\mu_2}) \end{aligned}$$

and

$$\begin{aligned} Bias(\psi,\xi) &= \underset{\sim}{b}^T(\psi,\xi)B^{-1}(\xi)AB^{-1}(\xi)\, \underset{\sim}{b}(\psi,\xi) \\ &= (\int \psi(x)d\xi(x), \int x\psi(x)d\xi(x)) \\ &\quad \cdot \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{\mu_2} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{12} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{\mu_2} \end{pmatrix} \begin{pmatrix} \int \psi(x)d\xi(x) \\ \int x\psi(x)d\xi(x) \end{pmatrix} \\ &= (\int \psi(x)d\xi(x))^2 + \frac{1}{12\mu_2^2}(\int x\psi(x)d\xi(x))^2. \end{aligned}$$

We consider the following problem:

$$\min_{\xi \in \mathcal{F}} V(\xi) \quad \text{subject to} \quad \max_{\psi \in \Psi} Bias(\psi,\xi) \leq c, \tag{5.4.1}$$

where c is a preassigned positive constant.

The maximization of $Bias(\psi,\xi)$ over the class Ψ has been done by Huber in 1975. The result can also be found in Huber (1981). We state the result in the next lemma.

Lemma 5.4.1 *(Huber 1975) Let $\mathcal{F}_l = \{\xi : \xi \in \mathcal{F}$ and $\mu_2 = \int_S x^2 d\xi(x) \geq \frac{1}{12}\}$. For any $\xi \in \mathcal{F}_l$, we have*

$$\max_{\psi \in \Psi} Bias(\psi,\xi) = \eta^2 \int_S (m(x)-1)^2 dx,$$

where $m(x) = \dfrac{d\xi(x)}{dx}$.

The fact $V(\xi) = \frac{\sigma^2}{n}(1 + \frac{1}{12\mu_2})$ implies that minimizing $V(\xi)$ is equivalent to maximizing μ_2. By Lemma 5.4.1, we conclude that, within the class $\mathcal{F}_l$, (5.4.1) is equivalent to

$$\max_{\xi \in \mathcal{F}_l} \mu_2 \text{ subject to } \int_S m(x)dx = 1, \text{ and } \eta^2 \int_S (m(x)-1)^2 dx \le c. \tag{5.4.2}$$

Furthermore, we can write $\mu_2 = \int_S x^2 d\xi(x) = \int_S x^2 m(x)dx$, and $\int_S (m(x)-1)^2 dx = \int_S (m^2(x) - 2m(x) + 1)dx = \int_S m^2(x)dx - 1$. Hence (5.4.2) becomes

$$\max_{\xi \in \mathcal{F}_l} \int_S x^2 m(x)dx \text{ subject to } \int_S m(x)dx = 1 \text{ and } \int_S m^2(x)dx \le \frac{c}{\eta^2} + 1. \tag{5.4.3}$$

We first consider the subproblem

$$\max_{\xi \in \mathcal{F}_l} \int_S x^2 m(x)dx \text{ subject to } \int_S m(x)dx = 1 \text{ and } \int_S m^2(x)dx = \frac{c}{\eta^2} + 1. \tag{5.4.4}$$

For some multipliers a, b, we maximize

$$\int [x^2 m(x) + \frac{b}{a} m(x) - \frac{1}{2a} m^2(x)]dx$$

by maximizing the integrand pointwise. We find that

$$m(x) = [ax^2 + b]^+,$$

with a, b determined by $\int_S m(x)dx = 1$ and $\int_S m^2(x)dx = \frac{c}{\eta_2} + 1$.

Before we solve the problem (5.4.3), we are going to show that the problem (5.4.3) and the problem (5.4.4) are equivalent.

Lemma 5.4.2

$$\max_{\xi \in \mathcal{F}_l} \int_S x^2 m(x)dx \textit{ subject to } \int_S m(x)dx = 1 \textit{ and } \int_S m^2(x)dx \le \frac{c}{\eta^2} + 1$$

is equivalent to

$$\max_{\xi \in \mathcal{F}_l} \int_S x^2 m(x)dx \textit{ subject to } \int_S m(x)dx = 1 \textit{ and } \int_S m^2(x)dx = \frac{c}{\eta^2} + 1.$$

Proof: Let μ_2^* be the maximum value of (5.4.3). We are going to prove Lemma 5.4.2 by showing that μ_2^* is a function of λ and increasing in λ where $0 \le \lambda \le \frac{c}{\eta_2}$.

For any $\xi \in \mathcal{F}_l$, we have $\mu_2 = \int_S x^2 d\xi(x) \ge \frac{1}{12}$ which implies $a > 0$. Hence, we consider two cases.

(i) $a > 0$, $b \geq 0$

In this case, we have $m(x) = [ax^2 + b]^+ = ax^2 + b$. Hence, we get

$$\textstyle\int_S m(x)dx = \int_{-\frac{1}{2}}^{\frac{1}{2}} (ax^2 + b)dx = \frac{a}{12} + b := 1, \qquad \text{(i-1)}$$

$$\textstyle\int_S m^2(x)dx = \int_{-\frac{1}{2}}^{\frac{1}{2}} (a^2x^4 + 2abx^2 + b^2)dx = \frac{1}{80}a^2 + \frac{1}{6}ab + b^2 := 1 + \lambda, \quad \text{(i-2)}$$

and

$$\textstyle\mu_2 = \int_S x^2 m(x)dx = \int_{-\frac{1}{2}}^{\frac{1}{2}} (ax^4 + bx^2)dx = \frac{a}{80} + \frac{b}{12}. \quad \text{(i-3)}$$

Put (i-1) into (i-2) to obtain

$$\frac{1}{80}a^2 + \frac{1}{6}a(1 - \frac{a}{12}) + (1 - \frac{a}{12})^2 = 1 + \lambda; \text{ and hence } a = \sqrt{180\lambda}. \quad \text{(i-4)}$$

Put (i-1) and (i-4) into (i-3) to obtain

$$\mu_2^* = \frac{1}{80}a + \frac{1}{12}(1 - \frac{1}{12}a) = \frac{1}{12} + \frac{1}{180}a = \frac{1}{12} + \frac{1}{180}\sqrt{180\lambda} = \frac{1}{12} + \frac{\sqrt{\lambda}}{6\sqrt{5}}. \quad \text{(i-5)}$$

Note $b = 1 - \frac{a}{12} = 1 - \frac{\sqrt{180\lambda}}{12} \geq 0$ implies that $\lambda \leq \frac{4}{5}$. Hence μ_2^* is a function of λ and increasing in λ when $\lambda \in (0, \frac{4}{5}]$.

(ii) $a > 0$, $b < 0$

In this case, there exists a $x_0 \in (0, \frac{1}{2})$ such that

$$m(x) = \begin{cases} ax^2 + b & x \in [-\frac{1}{2}, -x_0) \text{ or } x \in (x_0, \frac{1}{2}] \\ 0 & x \in [-x_0, x_0] \end{cases}.$$

Hence, we get

$$ax_0^2 + b = 0, \qquad \text{(ii-1)}$$

$$\textstyle\int_S m(x)ds = 2\int_{x_0}^{\frac{1}{2}} (ax^2 + b)dx = 2[\frac{a}{3}x^3 + bx]_{x_0}^{\frac{1}{2}} = 2[\frac{a}{3}(\frac{1}{8} - x_0^3) + b(\frac{1}{2} - x_0)] := 1, \quad \text{(ii-2)}$$

$$\begin{aligned} \textstyle\int_S m^2(x)dx &= 2\textstyle\int_{x_0}^{\frac{1}{2}} (ax^2 + b)^2 dx = 2\int_{x_0}^{\frac{1}{2}} (a^2x^4 + 2abx^2 + b^2)dx \\ &= 2[\frac{a^2}{5}x^5 + \frac{2ab}{3}x^3 + b^2x]_{x_0}^{\frac{1}{2}} \\ &= 2[\frac{a^2}{5}(\frac{1}{32} - x_0^5) + \frac{2ab}{3}(\frac{1}{8} - x_0^3) + b^2(\frac{1}{2} - x_0)] \\ &:= 1 + \lambda, \end{aligned} \qquad \text{(ii-3)}$$

and

$$\begin{aligned}\mu_2 = \int_S x^2 m(x)dx \quad &= 2\int_{x_0}^{\frac{1}{2}} x^2(ax^2+b)dx \\ &= 2\int_{x_0}^{\frac{1}{2}}(ax^4+bx^2)dx \\ &= 2[\frac{a}{5}x^5+\frac{b}{3}x^3]_{x_0}^{\frac{1}{2}} \\ &= 2[\frac{a}{5}(\frac{1}{32}-x_0^5)+\frac{b}{3}(\frac{1}{8}-x_0^3)]. \end{aligned} \qquad \text{(ii-4)}$$

Put (ii-1) into (ii-2) to get $\frac{a}{3}(\frac{1}{8}-x_0^3)+(-ax_0^2)(\frac{1}{2}-x_0)=\frac{1}{2}$; and hence we find

$$\begin{aligned} a &= \frac{1}{2[\frac{1}{3}(\frac{1}{8}-x_0^3)-x_0^2(\frac{1}{2}-x_0)]} \\ &= \frac{1}{2(\frac{1}{2}-x_0)(\frac{1}{12}+\frac{1}{6}x_0-\frac{2}{3}x_0^2)} \\ &= \frac{1}{2(x_0-\frac{1}{2})^2(\frac{2}{3}x_0+\frac{1}{6})}. \end{aligned} \qquad \text{(ii-5)}$$

Putting (ii-2) and (ii-5) into (ii-3), we get

$$\frac{a^2}{5}(\frac{1}{32}-x_0^5)+\frac{2a}{3}(-ax_0^2)(\frac{1}{8}-x_0^3)+(-ax_0^2)^2(\frac{1}{2}-x_0)=\frac{1+\lambda}{2}$$

which can be simplified as the following:

$$\frac{3}{5}\frac{(32x_0^2+18x_0+3)}{(-32x_0^3+6x_0+1)}=1+\lambda. \qquad \text{(ii-6)}$$

Finally, putting (ii-1) and (ii-5) into (ii-4), we find

$$\begin{aligned} \mu_2^* &= 2[\frac{a}{5}(\frac{1}{32}-x_0^5)+\frac{1}{3}(-ax_0^2)(\frac{1}{8}-x_0^3)] \\ &= 2[\frac{1}{5}(\frac{1}{32}-x_0^5)-\frac{1}{3}x_0^2(\frac{1}{8}-x_0^3)]a \\ &= \frac{2\left(x_0-\frac{1}{2}\right)^2\left(\frac{2}{15}x_0^3+\frac{2}{15}x_0^2+\frac{1}{10}x_0+\frac{1}{40}\right)}{2\left(x_0-\frac{1}{2}\right)^2\left(\frac{2}{3}x_0+\frac{1}{6}\right)} \\ &= \frac{16x_0^3+16x_0^2+12x_0+3}{20(4x_0+1)}. \end{aligned} \qquad \text{(ii-7)}$$

From (ii-6) and (ii-7), we know that μ_2^* is a function of x_0, and x_0 is a function of λ. Hence μ_2^* is a function of λ through x_0. For the sake of argument, we write x instead of x_0 hereafter. Taking the derivative on both sides of (ii-6) with respect to λ, we get

$$\frac{3}{5(-32x^3+6x+1)^2}[(-32x^3+6x+1)(64x+18)\frac{dx}{d\lambda}-(32x^2+18x+3)(-96x^2+6)\frac{dx}{d\lambda}]$$

$$=\frac{96x(32x^3+36x^2+15x+2)}{5(-32x^3+6x+1)^2}\frac{dx}{d\lambda}=1.$$

It is easy to see that

$$\frac{dx}{d\lambda}=\frac{5(-32x^3+6x+1)^2}{96x(32x^3+36x^2+15x+2)}>0,$$

when $x\in(0,\frac{1}{2})$. On the other hand, $x\in(0,\frac{1}{2})$ if and only if $\lambda\in(\frac{4}{5},\infty)$. Hence, we conclude that $\frac{dx}{d\lambda}>0$, for $\lambda\in(\frac{4}{5},\infty)$. Taking the derivative on both sides of (ii-7) with respect to x, we get

$$\begin{aligned}\frac{d\mu_2^*}{dx}&=\frac{1}{20(4x+1)^2}[(4x+1)(48x^2+32x+12)-4(16x^3+16x^2+12x+3)]\\&=\frac{1}{5(4x+1)^2}(32x^3+28x^2+8x)\\&=\frac{4x(8x^2+7x+2)}{5(4x+1)^2}>0\text{ for }x\in(0,\frac{1}{2}).\end{aligned}$$

By the chain rule for derivatives, we have

$$\frac{d\mu_2^*}{d\lambda}=\frac{d\mu_2^*}{dx}\cdot\frac{dx}{d\lambda}>0\quad\text{for }\lambda\in(\frac{4}{5},\infty).$$

Hence μ_2^* is a function of λ and increasing in λ. □

Remark 1. There are two limiting cases:

(i) $\lambda\to 0$ which corresponds to $a=0$, and $b=1$. In this case, we get $m(x)=1$ on $[-\frac{1}{2},\frac{1}{2}]$ which is the uniform density function and $\mu_2^*=\frac{1}{12}$.

(ii) $\lambda\to\infty$ which corresponds to $x_0\to\frac{1}{2}$. In this case, we have $\mu_2^*=\frac{1}{4}$ and $\xi^*=\frac{1}{2}\Delta_{\pm\frac{1}{2}}$ which is the same as the usual optimal design.

The next theorem follows by Lemma 5.4.2, which provides the solution to the problem (5.4.3).

Theorem 5.4.3 *Let us maintain the same notations as we used in this section. The solution to the problem (5.4.3) depends on the value of* $\lambda := \frac{c}{\eta^2}$.

(i) If $\lambda \in (0, \frac{4}{5}]$, *then* $m(x) = ax^2 + b$ *with* a, b *determined by*

$$\int_S m(x)dx = 1, \int_S m^2(x)dx = 1 + \lambda,$$

and $\mu_2^* = \frac{1}{12} + \frac{\sqrt{\lambda}}{6\sqrt{5}}$.

(ii) If $\lambda \in (\frac{4}{5}, \infty)$, *then*

$$m(x) = \begin{cases} ax^2 + b & x \in [-\frac{1}{2}, x_0) \text{ or } x \in (x_0, \frac{1}{2}] \\ 0 & x \in [-x_0, x_0] \end{cases}$$

with a, b, determined by

$$\int_S m(x)dx = 1, \int_S m^2(x)dx = 1 + \lambda,$$

and

$$\mu_2^* = \frac{16x_0^3 + 16x_0^2 + 12x_0 + 3}{20(4x_0 + 1)},$$

where x_0 *is determined by*

$$\frac{3(32x_0^2 + 18x_0 + 3)}{5(-32x_0^3 + 6x_0 + 1)} = 1 + \lambda.$$

It is possible to extend the problem (5.4.3) to the higher dimensional case. Here we only discuss the multiple linear regression case.

Let $\underset{\sim}{f}^T(\underset{\sim}{x}) = (1, x_1, ..., x_p)$, $\underset{\sim}{\theta}^T = (\theta_0, \theta_1, ..., \theta_p)$, and $\hat{\underset{\sim}{\theta}}^T = (\hat{\theta}_0, \hat{\theta}_1, ..., \hat{\theta}_p)$ where $\hat{\underset{\sim}{\theta}}$ is the least squares estimator of $\underset{\sim}{\theta}$. Define

$$\Psi = \{\psi(\underset{\sim}{x}) : \int_R \psi^2(\underset{\sim}{x})d\underset{\sim}{x} \leq \eta^2, \text{ and } \int_R \underset{\sim}{f}(\underset{\sim}{x})\psi(\underset{\sim}{x})d\underset{\sim}{x} = \underset{\sim}{0}\},$$

where $R = \{\underset{\sim}{x} : ||\underset{\sim}{x}|| \leq \gamma_p := \frac{[\Gamma(\frac{p}{2}+1)]^{\frac{1}{p}}}{\sqrt{\pi}}\}$, and

$$\mathcal{F} = \{\xi(\underset{\sim}{x}) : \frac{d\xi(\underset{\sim}{x})}{d\underset{\sim}{x}} = m(\underset{\sim}{x}), \textstyle\int_R m(\underset{\sim}{x})d\underset{\sim}{x} = 1, m(\underset{\sim}{x}) \geq 0, \text{ and } m(x_1, ..., -x_i, ..., x_p)$$

$$= m(x_1, ..., x_i, ..., x_p) \text{ on } R, i = 1, ..., p\}.$$

Then we have

$$
\begin{aligned}
B(\xi(\underset{\sim}{x})) &= \int_R \underset{\sim}{f}(\underset{\sim}{x})\, \underset{\sim}{f}^T(\underset{\sim}{x}) d\xi(\underset{\sim}{x}) \\
&= \int_R \begin{pmatrix} 1 & x_1 & \cdots & x_p \\ x_1 & x_1^2 & \cdots & x_1 x_p \\ \cdots & \cdots & \cdots & \cdots \\ x_p & x_p x_1 & \cdots & x_p^2 \end{pmatrix} d\xi(\underset{\sim}{x}) \\
&= \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & \int_R x_1^2 m(\underset{\sim}{x}) d\underset{\sim}{x} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \int_R x_p^2 m(\underset{\sim}{x}) d\underset{\sim}{x} \end{pmatrix} \\
&= \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & \gamma & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \gamma \end{pmatrix},
\end{aligned}
$$

where $\gamma = \int_R x_i^2 m(\underset{\sim}{x}) d\underset{\sim}{x}$, $i = 1, ..., p$, and

$$
\begin{aligned}
A &= \int_R \underset{\sim}{f}(\underset{\sim}{x})\, \underset{\sim}{f}^T(\underset{\sim}{x}) d\underset{\sim}{x} \\
&= \int_R \begin{pmatrix} 1 & x_1 & \cdots & x_p \\ x_1 & x_1^2 & \cdots & x_1 x_p \\ \vdots & \vdots & \ddots & \vdots \\ x_p & x_p x_1 & \cdots & x_p^2 \end{pmatrix} d\underset{\sim}{x} \\
&= \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & \int_R x_1^2 d\underset{\sim}{x} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \int_R x_p^2 d\underset{\sim}{x} \end{pmatrix} \\
&= \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & \gamma_0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \gamma_0 \end{pmatrix},
\end{aligned}
$$

where $\gamma_0 = \int_R x_i^2 d\underset{\sim}{x} = \frac{\gamma_p^2}{p+2}$.

Let $\hat{y} = \hat{\underset{\sim}{\theta}}^T \underset{\sim}{f}(\underset{\sim}{x})$ be the estimator of $\underset{\sim}{\theta}^T \underset{\sim}{f}(\underset{\sim}{x})$. Under the model

$$y_i = y(\underset{\sim}{x}_i) = \underset{\sim}{\theta}^T \underset{\sim}{f}(\underset{\sim}{x}_i) + \psi(\underset{\sim}{x}_i) + \epsilon_i, \qquad i = 1, 2, ..., n, \tag{5.4.5}$$

we have

$$\begin{aligned} MSE(\hat{y}) &= \frac{\sigma^2}{n} \underset{\sim}{f}^T(\underset{\sim}{x}) B^{-1}(\xi(\underset{\sim}{x})) \underset{\sim}{f}(\underset{\sim}{x}) \\ &+ \underset{\sim}{f}^T(\underset{\sim}{x}) B^{-1}(\xi(\underset{\sim}{x})) \underset{\sim}{b}(\psi(\underset{\sim}{x}), \xi(\underset{\sim}{x})) \underset{\sim}{b}^T(\psi(\underset{\sim}{x}), \xi(\underset{\sim}{x})) B^{-1}(\xi(\underset{\sim}{x})) \underset{\sim}{f}(\underset{\sim}{x}), \end{aligned}$$

where

$$\begin{aligned} \underset{\sim}{b}^T(\psi(\underset{\sim}{x}), \xi(\underset{\sim}{x})) &= \int_R \underset{\sim}{f}^T(\underset{\sim}{x}) \psi(\underset{\sim}{x}) d\xi(\underset{\sim}{x}) \\ &= \left(\int_R \psi(\underset{\sim}{x}) m(\underset{\sim}{x}) d\underset{\sim}{x}, \int_R x_1 \psi(\underset{\sim}{x}) m(\underset{\sim}{x}) d\underset{\sim}{x}, ..., \int_R x_p \psi(\underset{\sim}{x}) m(\underset{\sim}{x}) d\underset{\sim}{x}\right), \end{aligned}$$

and

$$\int_R MSE(\hat{y}) d\underset{\sim}{x} = V(\xi(\underset{\sim}{x})) + Bias(\psi(\underset{\sim}{x}), \xi(\underset{\sim}{x})),$$

where

$$\begin{aligned} V(\xi(\underset{\sim}{x})) &= \frac{\sigma^2}{n} tr A B^{-1}(\xi(\underset{\sim}{x})) \\ &= \frac{\sigma^2}{n} tr \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & \gamma_0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \gamma_0 \end{pmatrix} \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & \frac{1}{\gamma} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{1}{\gamma} \end{pmatrix} \\ &= \frac{\sigma^2}{n}[1 + \frac{p\gamma_0}{\gamma}], \end{aligned}$$

and

$$Bias(\psi(\underset{\sim}{x}), \xi(\underset{\sim}{x})) = \underset{\sim}{b}^T(\psi(\underset{\sim}{x}), \xi(\underset{\sim}{x})) B^{-1}(\xi(\underset{\sim}{x})) A B^{-1}(\xi(\underset{\sim}{x})) \underset{\sim}{b}(\psi(\underset{\sim}{x}), \xi(\underset{\sim}{x})).$$

We consider the following problem:

$$\min_{\xi \in \mathcal{F}} V(\xi(\underset{\sim}{x})) \text{ subject to } \max_{\psi \in \Psi} Bias(\psi(\underset{\sim}{x}), \xi(\underset{\sim}{x})) \leq c, \tag{5.4.6}$$

where c is a preassigned positive constant.

The maximization of $Bias(\psi(\underset{\sim}{x}), \xi(\underset{\sim}{x}))$ over the class Ψ has been done by Wiens (1990). Let $H = B(\xi(\underset{\sim}{x})) A^{-1} B(\xi(\underset{\sim}{x}))$, and $K = \int_R \underset{\sim}{f}(\underset{\sim}{x}) \underset{\sim}{f}^T(\underset{\sim}{x}) m^2(\underset{\sim}{x}) d\underset{\sim}{x}$, and v_ξ be the largest solution to the equation $|K - vH| = 0$. Then we have:

Lemma 5.4.4 *(Wiens 1990) Let* $\mathcal{F}_m = \{\xi : \xi \in \mathcal{F},$ *and* $\gamma = \int_R x_i^2 d\xi(\underset{\sim}{x}) \geq \frac{\gamma_p^2}{p+2}\}$. *For any* $\xi \in \mathcal{F}_m$, *we have*

$$\max_{\psi \in \Psi} Bias(\psi(\underset{\sim}{x}), \xi(\underset{\sim}{x})) = \eta^2 (\int_R m^2(\underset{\sim}{x}) d \underset{\sim}{x} - 1),$$

where $m(\underset{\sim}{x}) = \frac{d\xi(\underset{\sim}{x})}{d \underset{\sim}{x}}$.

Similar to the one dimensional case, within the class of $\mathcal{F}_m$, we only need to consider the problem:

$$\min_{x \in \mathcal{F}_m} \frac{\sigma^2}{n}[1 + \frac{p\gamma_0}{\gamma}] \text{ subject to } \eta^2 (\int_R m^2(\underset{\sim}{x}) d \underset{\sim}{x} - 1) = c, \tag{5.4.7}$$

which is equivalent to

$$\begin{array}{c} \max \int_R x_1^2 m(\underset{\sim}{x}) d \underset{\sim}{x} \quad \text{subject to } \int_R m(\underset{\sim}{x}) d \underset{\sim}{x} = 1 \\ \text{and } \int_R m^2(\underset{\sim}{x}) d \underset{\sim}{x} = \frac{c}{\eta^2} + 1 := \lambda + 1. \end{array} \tag{5.4.8}$$

For some multipliers a, b we maximize

$$\int_R [x_1^2 m(\underset{\sim}{x}) + \frac{b}{a} m(\underset{\sim}{x}) - \frac{1}{2a} m^2(\underset{\sim}{x})] d \underset{\sim}{x}$$

by maximizing the integrand pointwise. We find $m(\underset{\sim}{x}) = [a \| \underset{\sim}{x} \|^2 + b]^+$, with a, b determined by

$$\int_R m(\underset{\sim}{x}) d \underset{\sim}{x} = 1 \text{ and } \int_R m^2(\underset{\sim}{x}) d \underset{\sim}{x} = \lambda + 1.$$

Chapter 6

Robust Designs for some Regression Models with Random Bias

6.1 Introduction and Preliminaries

In Chapter 4 and Chapter 5, we have assumed that a class of possible bias functions exists but that all functions in the class are equally likely to be the actual bias present in the regression model. In practice, it is possible that certain bias functions would be more likely than the others, and perhaps the experimenter can specify a prior probability distribution on the form of the possible bias functions in the model. In this chapter, we consider the following regression model:

$$y(\underset{\sim i}{x},\omega) = \underset{\sim}{\theta}^T \underset{\sim}{f}(\underset{\sim i}{x}) + \psi(\underset{\sim i}{x},\omega) + \epsilon_i(\omega), \qquad i = 1, ..., n. \tag{6.1.1}$$

The $\epsilon_i(\omega)$'s, for a given ω, are uncorrelated random variables with mean 0 and variance $\sigma^2 > 0$. $\underset{\sim i}{x} \in S := \{(x_1, ..., x_q) : -1 \leq x_j \leq 1, \ j = 1, ..., q\} \subseteq \mathbb{R}^q$, $i = 1, ..., n$. $\underset{\sim}{\theta}^T = (\theta_0, \theta_1, ..., \theta_p)$, and $\underset{\sim}{f}^T(\underset{\sim}{x}) = \left(f_0(\underset{\sim}{x}), f_1(\underset{\sim}{x}), ..., f_p(\underset{\sim}{x})\right)$. ω is a random variable defined on Ω with distribution $\prod(\omega)$. For $A \in \mathcal{F}_\Omega$, $\mathcal{F}_\Omega$ is a σ−finite field defined on Ω, $\prod(A)$ represents the probability that $\psi(\cdot,\omega)$ falls in the set $\{\psi(\cdot,\omega), \ \omega \in A\}$. $\prod(\omega)$ represents our prior knowledge or opinion about the distribution of possible functions

$\psi(\cdot,\omega)$. We additionally assume that $E[\psi(\underset{\sim}{x},\omega)] = \int \psi(\underset{\sim}{x},\omega)d\prod(\omega) = 0$ and $E[\psi^2(\underset{\sim}{x},\omega)] < \infty$ for all $\underset{\sim}{x}\in S$. The assumption $E[\psi(\underset{\sim}{x},\omega)] = 0$ reflects the notion that the model

$$y(\underset{\sim i}{x},\omega) = \underset{\sim}{\theta}^T \underset{\sim}{f}(\underset{\sim i}{x}) + \epsilon_i(\omega), \qquad i = 1, ..., n,$$

is correct on the average, but any particular realization (choice of ω) may induce the bias $\psi(\underset{\sim i}{x},\omega)$. Note that if $\infty > \left|E[\psi(\underset{\sim}{x},\omega)]\right| \neq 0$, and one can write $E[\psi(\underset{\sim}{x},\omega)] = \underset{\sim}{\mu}^T \cdot \underset{\sim}{g}(\underset{\sim}{x})$ where $\underset{\sim}{\mu}$ is a vector of unknown constants, and $\underset{\sim}{g}(\underset{\sim}{x})$ a known function. One can then define $\phi(\underset{\sim}{x},\omega) = \psi(\underset{\sim}{x},\omega) - E[\psi(\underset{\sim}{x},\omega)]$ and write

$$y(\underset{\sim}{x},\omega) = (\underset{\sim}{\theta}^T, \underset{\sim}{\mu}^T)\begin{pmatrix} \underset{\sim}{f}(\underset{\sim}{x}) \\ \underset{\sim}{g}(\underset{\sim}{x}) \end{pmatrix} + \phi(\underset{\sim}{x},\omega) + \epsilon(\underset{\sim}{x},\omega),$$

which is in the form of (6.1.1) with $E[\phi(\underset{\sim}{x},\omega)] = 0$. Note also that the condition $E[\psi(\underset{\sim}{x},\omega)] = 0$ insures the identifiability of the parameters $\underset{\sim}{\theta}$.

Let $\xi(\underset{\sim}{x})$ be a design measure defined on S. We define $B(\xi(\underset{\sim}{x})) = \int_S \underset{\sim}{f}(\underset{\sim}{x}) \cdot \underset{\sim}{f}^T(\underset{\sim}{x})d\xi(\underset{\sim}{x})$ and $\underset{\sim}{b}^T(\psi(\underset{\sim}{x},\omega),\ \xi(\underset{\sim}{x})) = \left(\int_S f_0(\underset{\sim}{x})\psi(\underset{\sim}{x},\omega)d\xi(\underset{\sim}{x}), ..., \int_S f_p(\underset{\sim}{x})\psi(\underset{\sim}{x},\omega)d\xi(\underset{\sim}{x})\right)$. The least squares estimator of $\underset{\sim}{\theta}$ is then $\underset{\sim}{\hat{\theta}} = B^{-1}(\xi(\underset{\sim}{x}))\int_S \underset{\sim}{f}(\underset{\sim}{x})y(\underset{\sim}{x})d\,\underset{\sim}{x}$ with bias vector and mean squared error matrix as follows:

$$E[\underset{\sim}{\hat{\theta}} - \underset{\sim}{\theta}] = B^{-1}(\xi(\underset{\sim}{x}))\cdot \underset{\sim}{b}(\psi(\underset{\sim}{x},\omega),\ \xi(\underset{\sim}{x})) \tag{6.1.2}$$

and

$$\begin{aligned} MSE(\underset{\sim}{\hat{\theta}}) &= E[(\underset{\sim}{\hat{\theta}} - \underset{\sim}{\theta})(\underset{\sim}{\hat{\theta}} - \underset{\sim}{\theta})^T] \\ &= \frac{\sigma^2}{n} B^{-1}(\xi(\underset{\sim}{x})) + B^{-1}(\xi(\underset{\sim}{x}))\, \underset{\sim}{b}(\psi(\underset{\sim}{x},\omega),\ \xi(\underset{\sim}{x})) \\ &\quad \cdot \underset{\sim}{b}^T(\psi(\underset{\sim}{x},\omega),\ \xi(\underset{\sim}{x}))B^{-1}(\xi(\underset{\sim}{x})). \end{aligned} \tag{6.1.3}$$

Let $\hat{y} = \underset{\sim}{\hat{\theta}}^T \underset{\sim}{f}(\underset{\sim}{x})$ be the estimator of $\underset{\sim}{\theta}^T \underset{\sim}{f}(\underset{\sim}{x})$. Then the mean squared error of $\hat{y}$ is

$$\begin{aligned} MSE(\hat{y}) &= E[(\underset{\sim}{\hat{\theta}}^T \underset{\sim}{f}(\underset{\sim}{x}) - \underset{\sim}{\theta}^T \underset{\sim}{f}(\underset{\sim}{x}))^2] \\ &= \frac{\sigma^2}{n} \underset{\sim}{f}^T(\underset{\sim}{x})B^{-1}(\xi(\underset{\sim}{x}))\, \underset{\sim}{f}(\underset{\sim}{x}) + \underset{\sim}{f}^T(\underset{\sim}{x})B^{-1}(\xi(\underset{\sim}{x})) \\ &\quad \cdot \underset{\sim}{b}(\psi(\underset{\sim}{x},\omega),\ \xi(\underset{\sim}{x}))\, \underset{\sim}{b}^T(\psi(\underset{\sim}{x},\omega),\ \xi(\underset{\sim}{x}))B^{-1}(\xi(\underset{\sim}{x}))\, \underset{\sim}{f}(\underset{\sim}{x}). \end{aligned} \tag{6.1.4}$$

We consider the loss functions (i) $\mathcal{L}_D(\xi) = \int_\Omega \left| MSE(\hat{\underset{\sim}{\theta}}) \right| d\prod(\omega)$, (ii) $\mathcal{L}_A(\xi) = \int_\Omega tr\ MSE(\hat{\underset{\sim}{\theta}}) d\prod(\omega)$, and (iii) $\mathcal{L}_Q(\xi) = \int_\Omega \left(\int_S MSE(\hat{y}) d\underset{\sim}{x} \right) d\prod(\omega)$. It is clear that these loss functions correspond to the classical notions of $D-$, $A-$, and Q-optimality if $\psi(\underset{\sim}{x}, \omega) \equiv 0$. For convenience, we call a design measure ξ $D-$, $A-$, or Q-optimal if ξ minimizes $\mathcal{L}_D(\xi)$, $\mathcal{L}_A(\xi)$, or $\mathcal{L}_Q(\xi)$ respectively. In this chapter, we are going to find:

(i) $Q-$optimal design for one dimensional polynomial regression model;
(ii) $D-$, $A-$, and Q-optimal designs for multiple linear regression model;
(iii) $D-$, $A-$, and Q-optimal designs for two dimensional linear regression with interaction term.

The solutions are provided in Section 6.2, Section 6.3 and Section 6.4 respectively. The problem discussed here was posed by Notz in 1989. In his paper, he found optimal designs for one dimensional polynomial regression model with respect to $D-$ and $A-$optimal criteria. Hence this chapter extends Notz's results to some other loss function and to the high dimensional situations.

Before we start to solve the optimal design problems, we first provide some results which are useful in the later sections. For convenience, we set $v = \dfrac{\sigma^2}{n}$. We maintain the same notations as we made earlier in this section. We present some useful results in the next four lemmas.

Lemma 6.1.1 $\mathcal{L}_D(\xi) = \left| vB^{-1}(\xi(\underset{\sim}{x})) \right| \cdot \{1 + v^{-1} \int_\Omega \underset{\sim}{b}^T (\psi(\underset{\sim}{x}, \omega), \xi(\underset{\sim}{x})) \cdot B^{-1}(\xi(\underset{\sim}{x}))\ \underset{\sim}{b}\ (\psi(\underset{\sim}{x}, \omega), \xi(\underset{\sim}{x})) d\prod(\omega)\}$.

Proof: Let G be a nondegenerate $m \times m$ matrix and let F be an $m \times k$ matrix. Then we have

$$\left| G + FF^T \right| = |G| \left| I_k + F^T G^{-1} F \right|$$

(See Fedorov (1972)). Let $G = vB^{-1}(\xi(\underset{\sim}{x}))$ and $F = B^{-1}(\xi(\underset{\sim}{x}))\ \underset{\sim}{b}\ (\psi(\underset{\sim}{x}, \omega), \xi(\underset{\sim}{x}))$ (In this case, we have $k = 1$). Then we have

$$
\begin{aligned}
|\,MSE(\underset{\sim}{\hat{\theta}})\,| &= |\,vB^{-1}(\xi(\underset{\sim}{x})) + B^{-1}(\xi(\underset{\sim}{x}))\ \underset{\sim}{b}\ (\psi(\underset{\sim}{x},\omega),\xi(\underset{\sim}{x})) \\
&\quad \cdot\underset{\sim}{b}^T\ (\psi(\underset{\sim}{x},\omega),\xi(\underset{\sim}{x}))B^{-1}(\xi(\underset{\sim}{x}))\,| \\
&= |\,vB^{-1}(\xi(\underset{\sim}{x}))\,||\,I_k + (B^{-1}(\xi(\underset{\sim}{x}))\ \underset{\sim}{b}\ (\psi(\underset{\sim}{x},\omega),\xi(\underset{\sim}{x})))^T \\
&\quad \cdot(vB^{-1}(\xi(\underset{\sim}{x})))^{-1}(B^{-1}(\xi(\underset{\sim}{x}))\ \underset{\sim}{b}\ (\psi(\underset{\sim}{x},\omega),\xi(\underset{\sim}{x}))\,| \\
&= |\,vB^{-1}(\xi(\underset{\sim}{x}))\,|\cdot\{1+\ \underset{\sim}{b}^T\ (\psi(\underset{\sim}{x},\omega),\ \xi(\underset{\sim}{x}))B^{-1}(\xi(\underset{\sim}{x}))\cdot\frac{1}{v}B(\xi(\underset{\sim}{x})) \\
&\quad \cdot B^{-1}(\xi(\underset{\sim}{x}))\ \underset{\sim}{b}\ (\psi(\underset{\sim}{x},\omega),\xi(\underset{\sim}{x}))\} \\
&= |\,vB^{-1}(\xi(\underset{\sim}{x}))\,|\cdot\{1+v^{-1}\cdot\ \underset{\sim}{b}^T\ (\psi(\underset{\sim}{x},\omega),\xi(\underset{\sim}{x}))B^{-1}(\xi(\underset{\sim}{x}))\ \underset{\sim}{b}\ (\psi(\underset{\sim}{x},\omega),\xi(\underset{\sim}{x}))\}.
\end{aligned}
$$

Hence, we get

$$
\begin{aligned}
\mathcal{L}_D(\xi) &= \textstyle\int_\Omega \left|MSE(\underset{\sim}{\hat{\theta}})\right| d\prod(\omega) \\
&= |\,vB^{-1}(\xi(\underset{\sim}{x}))\,|\cdot\{1+v^{-1}\textstyle\int_\Omega\ \underset{\sim}{b}^T\ (\psi(\underset{\sim}{x},\omega),\xi(\underset{\sim}{x})) \\
&\quad \cdot B^{-1}(\xi(\underset{\sim}{x}))\ \underset{\sim}{b}\ (\psi(\underset{\sim}{x},\omega),\xi(\underset{\sim}{x}))d\prod(\omega)\}.
\end{aligned}
$$

□

Lemma 6.1.2 $\mathcal{L}_A(\xi) = v\cdot trB^{-1}(\xi(\underset{\sim}{x})) + \int_\Omega \underset{\sim}{b}^T\ (\psi(\underset{\sim}{x},\omega),\xi(\underset{\sim}{x}))B^{-2}(\xi(\underset{\sim}{x}))$
$\cdot\underset{\sim}{b}\ (\psi(\underset{\sim}{x},\omega),\xi(\underset{\sim}{x}))d\prod(\omega)$.

Proof: Let G be an $m\times m$ matrix and $\underset{\sim}{u}$ be a m-dimensional column vector. We are going to show that the following equation is true:

$$
trG\ \underset{\sim}{u}\underset{\sim}{u}^T\ G = \underset{\sim}{u}^T\ G^2\ \underset{\sim}{u}\ . \tag{6.1.5}
$$

Let g_{ij} be the i^{th} row and j^{th} column element of G, and $\underset{\sim}{u}^T = (u_1,...,u_m)$. The j^{th} row and k^{th} column element of $\underset{\sim}{u}\underset{\sim}{u}^T$ is u_ju_k. Hence, the i^{th} row and k^{th} column element of $G\ \underset{\sim}{u}\underset{\sim}{u}^T$ is $\sum_{j=1}^m g_{ij}u_ju_k$, and the i^{th} row and l^{th} column element of $G\ \underset{\sim}{u}\underset{\sim}{u}^T\ G$ is $\sum_{k=1}^m\left(\sum_{j=1}^m g_{ij}u_ju_k\right)g_{kl}$. Therefore, we have

$$
trG\ \underset{\sim}{u}\underset{\sim}{u}^T\ G = \sum_{i=1}^m\left[\sum_{k=1}^m\left(\sum_{j=1}^m g_{ij}u_ju_k\right)g_{ki}\right] = \sum_{k=1}^m\left[\sum_{j=1}^m u_j\left(\sum_{i=1}^m g_{ki}g_{ij}\right)u_k\right].
$$

On the other hand, the k^{th} row and j^{th} column element of G^2 is $\sum_{i=1}^m g_{ki}g_{ij}$. Hence

$$
\underset{\sim}{u}^T\ G^2\ \underset{\sim}{u} = \sum_{k=1}^m\left[\sum_{j=1}^m u_j\left(\sum_{i=1}^m g_{ki}g_{ij}\right)u_k\right] = trG\ \underset{\sim}{u}\underset{\sim}{u}^T\ G.
$$

We have proved that (6.1.5) is true.

Let $G = B^{-1}(\xi(\underset{\sim}{x}))$ and $\underset{\sim}{u}=\underset{\sim}{b}\,(\psi(\underset{\sim}{x},\omega),\xi(\underset{\sim}{x}))$. By (6.1.5), we have

$$\begin{aligned} trMSE(\hat{\theta}) &= tr[vB^{-1}(\xi(\underset{\sim}{x})) + B^{-1}(\xi(\underset{\sim}{x}))\,\underset{\sim}{b}\,(\psi(\underset{\sim}{x},\omega),\xi(\underset{\sim}{x})) \\ &\qquad \cdot \underset{\sim}{b}^T\,(\psi(\underset{\sim}{x},\omega),\xi(\underset{\sim}{x}))B^{-1}(\xi(\underset{\sim}{x}))] \\ &= v\cdot trB^{-1}(\xi(\underset{\sim}{x})) + \underset{\sim}{b}^T\,(\psi(\underset{\sim}{x},\omega),\xi(\underset{\sim}{x}))B^{-2}(\xi(\underset{\sim}{x}))\,\underset{\sim}{b}\,(\psi(\underset{\sim}{x},\omega),\xi(\underset{\sim}{x})). \end{aligned}$$

Hence, we get

$$\begin{aligned} \mathcal{L}_A(\xi) &= \textstyle\int_\Omega trMSE(\hat{\theta})d\prod(\omega) \\ &= v\cdot trB^{-1}(\xi(\underset{\sim}{x})) + \textstyle\int_\Omega \underset{\sim}{b}^T\,(\psi(\underset{\sim}{x},\omega),\xi(\underset{\sim}{x})) \\ &\quad \cdot B^{-2}(\xi(\underset{\sim}{x}))\,\underset{\sim}{b}\,(\psi(\underset{\sim}{x},\omega),\xi(\underset{\sim}{x}))d\prod(\omega). \end{aligned}$$ □

Lemma 6.1.3 *Let* $\underset{\sim}{f}^T\,(\underset{\sim}{x}) = (f_0(\underset{\sim}{x}),...,f_p(\underset{\sim}{x}))$ *and* B *be a* $(p+1)\times(p+1)$ *matrix. Then*

$$\int_S \underset{\sim}{f}^T\,(\underset{\sim}{x})B\,\underset{\sim}{f}\,(\underset{\sim}{x})d\,\underset{\sim}{x} = trAB,$$

where $A = \int_S \underset{\sim}{f}\,(\underset{\sim}{x})\,\underset{\sim}{f}^T\,(\underset{\sim}{x})d\,\underset{\sim}{x}$.

Proof: Let $a_{ij} = \int_S f_i(\underset{\sim}{x})f_j(\underset{\sim}{x})d\,\underset{\sim}{x}$ be the i^{th} row and j^{th} column element of A, b_{jk} be the j^{th} row and k^{th} column element of B and c_{ik} be the i^{th} row and k^{th} column element of AB. We have

$$c_{ik} = \sum_{j=0}^{p} a_{ij}b_{jk} = \sum_{j=0}^{p}\int_S f_i(\underset{\sim}{x})f_j(\underset{\sim}{x})b_{jk}d\,\underset{\sim}{x}\,.$$

Consequently, we get

$$\begin{aligned} trAB &= \sum_{i=0}^{p} c_{ii} \\ &= \sum_{i=0}^{p}\left(\sum_{j=0}^{p}\int_S f_i(\underset{\sim}{x})b_{ji}f_j(\underset{\sim}{x})d\,\underset{\sim}{x}\right) \\ &= \int_S \left(\sum_{i=0}^{p}\sum_{j=0}^{p} f_i(\underset{\sim}{x})b_{ji}f_j(\underset{\sim}{x})\right)d\,\underset{\sim}{x} \\ &= \int_S \underset{\sim}{f}^T\,(\underset{\sim}{x})B\,\underset{\sim}{f}\,(\underset{\sim}{x})d\,\underset{\sim}{x}\,. \end{aligned}$$ □

Suppose ξ is a probability measure supported on $[-1,1]$ and $\mu_i = \int_{-1}^{1} x^i d\xi(x)$ for integers $i \geq 0$. If $i \geq j \geq 0$ and i, j are even integers, then Notz (1989) showed

$$\mu_{i+j} \geq \mu_i \cdot \mu_j. \tag{6.1.6}$$

The application of (6.1.6) will give us the next lemma.

Lemma 6.1.4 $\mu_2^k \leq \mu_{2k} \leq \mu_2$ *for* $k \geq 1$.

Proof: (i) For $k = 1$, we have $\mu_2 \leq \mu_2$.

Assume that $\mu_2^{k-1} \leq \mu_{2k-2}$. Then we have $\mu_2^k = \mu_2^{k-1} \cdot \mu_2 \leq \mu_{2k-2} \cdot \mu_2 \leq \mu_{2k}$.

Note that the first inequality follows by the assumption and the second inequality follows by (6.1.6). By mathematical induction, we have

$$\mu_2^k \leq \mu_{2k} \text{ for all } k \geq 1.$$

(ii) $\mu_{2k} = \int_{-1}^{1} x^{2k} d\xi(x) \leq \int_{-1}^{1} x^2 d\xi(x) = \mu_2$.

We have proved $\mu_2^k \leq \mu_{2k} \leq \mu_2$ for $k \geq 1$. □

6.2 Q-optimal Design for One Dimensional Polynomial Regression

We first consider the one dimensional linear regression model. Let $\underset{\sim}{f}^T(x) = (1, x)$, $\psi(x, \omega) = \sum_{i=0}^{r} c_i(\omega)x^i$, where $x \in S := [-1, 1]$. $c_i(\omega)$'s are integrable real valued function on Ω satisfying $\gamma_i = \int_\Omega c_i^2(\omega)d\prod(\omega) < \infty$, and $\int_\Omega c_i(\omega)c_j(\omega)d\prod(\omega) = 0$ for $i \neq j$ and $i + j$ even. Let $\mathcal{F}_S$ be the set of all the symmetric design measures defined on S. For any $\xi \in \mathcal{F}_S$, we have

$$B(\xi(x)) = \int_S \underset{\sim}{f}(x)\,\underset{\sim}{f}^T(x)d\xi(x) = \int_S \begin{pmatrix} 1 & x \\ x & x^2 \end{pmatrix} d\xi(x) = \begin{pmatrix} 1 & 0 \\ 0 & \mu_2 \end{pmatrix}$$

and

$$\begin{aligned} \underset{\sim}{b}^T(\psi(x,\omega), \xi(x)) &= \left(\int \psi(x,\omega)d\xi(x),\ \int x\psi(x,\omega)d\xi(x)\right) \\ &= \left(\sum_{\substack{i=0 \\ i \text{ even}}}^{r} c_i(\omega)\mu_i,\ \sum_{\substack{i=0 \\ i \text{ odd}}}^{r} c_i(\omega)\mu_{i+1}\right), \end{aligned}$$

where $\mu_i = \int_S x^i d\xi(x)$. We also have

$$A = \int_S \underset{\sim}{f}^T(x)\,\underset{\sim}{f}(x)dx = \int_{-1}^{1} \begin{pmatrix} 1 & x \\ x & x^2 \end{pmatrix} dx = \begin{pmatrix} 2 & 0 \\ 0 & \frac{2}{3} \end{pmatrix}$$

and

$$\int_S MSE(\hat{y})dx = \frac{\sigma^2}{n} tr \begin{pmatrix} 2 & 0 \\ 0 & \frac{2}{3} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{\mu_2} \end{pmatrix} + tr \begin{pmatrix} 2 & 0 \\ 0 & \frac{2}{3} \end{pmatrix}$$

$$\cdot \begin{pmatrix} \left(\sum_{\substack{i=0 \\ i \text{ even}}}^{r} c_i(\omega)\mu_i \right)^2 & \frac{1}{\mu_2} \left(\sum_{\substack{i=0 \\ i \text{ even}}}^{r} c_i(\omega)\mu_i \right) \left(\sum_{\substack{i=0 \\ i \text{ odd}}}^{r} c_i(\omega)\mu_{i+1} \right) \\ \frac{1}{\mu_2} \left(\sum_{\substack{i=0 \\ i \text{ even}}}^{r} c_i(\omega)\mu_i \right) \left(\sum_{\substack{i=0 \\ i \text{ odd}}}^{r} c_i(\omega)\mu_{i+1} \right) & \frac{1}{\mu_2^2} \left(\sum_{\substack{i=0 \\ i \text{ odd}}}^{r} c_i(\omega)\mu_{i+1} \right)^2 \end{pmatrix}$$

$$= \frac{2\sigma^2}{n} \left(1 + \frac{1}{3\mu_2} \right) + 2 \left(\sum_{\substack{i=0 \\ i \text{ even}}}^{r} c_i(\omega)\mu_i \right)^2 + \frac{2}{3\mu_2^2} \left(\sum_{\substack{i=0 \\ i \text{ odd}}}^{r} c_i(\omega)\mu_{i+1} \right)^2 .$$

Therefore, we find

$$\mathcal{L}_Q(\xi) = \int_\Omega \left(\int_S MSE(\hat{y})dx \right) d\prod(\omega)$$

$$= 2v \left(1 + \frac{1}{3\mu_2} \right) + 2 \sum_{\substack{i=0 \\ i \text{ even}}}^{r} \gamma_i \mu_i^2 + \frac{2}{3\mu_2^2} \sum_{\substack{i=0 \\ i \text{ odd}}}^{r} \gamma_i \mu_{i+1}^2,$$

where $v = \dfrac{\sigma^2}{n}$. Our objective is to find a symmetric design measure ξ_0 minimizing $\mathcal{L}_Q(\xi)$ over $\mathcal{F}_S$. This is equivalent to minimizing

$$\frac{v}{3\mu_2} + \sum_{\substack{i=0 \\ i \text{ even}}}^{r} \gamma_i \mu_i^2 + \frac{1}{3\mu_2^2} \sum_{\substack{i=0 \\ i \text{ odd}}}^{r} \gamma_i \mu_{i+1}^2 \tag{6.2.1}$$

over $\mathcal{F}_S$.

In light of Lemma 6.1.4, (6.2.1) can be minimized when ξ is such that $\mu_i = \mu_2^{\frac{i}{2}}$ for even $i \geq 2$. This occurs when ξ is of the form $\xi(z) = \xi(-z) = \frac{1}{2}$ for some $0 < z \leq 1$ or $\xi(0) = 1$. We exclude the case $\xi(0) = 1$ since then $\mu_2 = 0$ and $\mathcal{L}_Q(\xi)$ is not defined. For any ξ of the form $\xi(z) = \xi(-z) = \frac{1}{2}$, we have $\mu_i = z^i$ when i is even. For such ξ, (6.2.1) becomes

$$\frac{v}{3z^2} + \sum_{\substack{i=0 \\ i \text{ even}}}^{r} \gamma_i z^{2i} + \frac{1}{3z^4} \sum_{\substack{i=0 \\ i \text{ odd}}}^{r} \gamma_i z^{2i+2}$$
$$= \frac{v}{3z^2} + \sum_{\substack{i=0 \\ i \text{ even}}}^{r} \gamma_i z^{2i} + \frac{1}{3} \sum_{\substack{i=0 \\ i \text{ odd}}}^{r} \gamma_i z^{2i-2}$$
$$= \frac{v}{3x} + \sum_{\substack{i=0 \\ i \text{ even}}}^{r} \gamma_i x^{i} + \frac{1}{3} \sum_{\substack{i=0 \\ i \text{ odd}}}^{r} \gamma_i x^{i-1} := L(x),$$

where $x = z^2$ and $0 < x \leq 1$. Let

$$\frac{dL(x)}{dx} = -\frac{v}{3x^2} + \sum_{\substack{i=0 \\ i \text{ even}}}^{r} i\gamma_i x^{i-1} + \frac{1}{3} \sum_{\substack{i=0 \\ i \text{ odd}}}^{r} (i-1)\gamma_i x^{i-2} := 0.$$

We get

$$v = 3 \sum_{\substack{i=0 \\ i \text{ even}}}^{r} i\gamma_i x^{i+1} + \sum_{\substack{i=0 \\ i \text{ odd}}}^{r} (i-1)\gamma_i x^{i}. \tag{6.2.2}$$

Thus we have proved the following theorem.

Theorem 6.2.1 *Assume* $\underset{\sim}{f}^T(x) = (1, x)$ *and* $\psi(x, \omega) = \sum_{i=0}^{r} c_i(\omega) x^i$ *for some integer* $r \geq 0$, *where the* $c_i(\omega)'s$ *are integrable real valued functions defined on* Ω *satisfying* $\gamma_i = \int_\Omega c_i^2(\omega) d\prod(\omega) < \infty$ *for* $i = 0, ..., r$ *and* $\int_\Omega c_i(\omega) c_j(\omega) d\prod(\omega) = 0$ *for all* $i \neq j$ *and* $i + j$ *even. Then there exists a design measure* $\xi_0 \in \mathcal{F}_S$ *that minimizes* $\mathcal{L}_Q(\xi)$ *and is of the form* $\xi(z) = \xi(-z) = \frac{1}{2}$, *where* $z = \min\{1, \sqrt{x}\}$ *and* $0 < x \leq 1$ *satisfies (6.2.2).*

Probably, the most useful special case is $\psi(x, \omega) = \sum_{i=0}^{2} c_i(\omega) x^i$. In this case, the more explicit result can be found. We state it as a corollary.

Corollary 6.2.2 *Suppose the assumptions of Theorem 6.2.1 are hold with* $r = 2$. *Then* $\mathcal{L}_Q(\xi)$ *is minimized by* ξ_0 *of the form* $\xi(z) = \xi(-z) = \frac{1}{2}$ *where* $z = \min\left\{1, \sqrt[6]{\frac{v}{6\gamma_2}}\right\}$.

Proof: In (6.2.2), we put $r = 2$. Then we have $v = 6\gamma_2 x^3$ and hence $x = \sqrt[3]{\frac{v}{6\gamma_2}}$. The corollary follows by Theorem 6.2.1. □

Now we consider the one dimensional quadratic regression model. Let $\underset{\sim}{f}^T(x) = (1, x, x^2)$. For $\psi(x,\omega)$, we consider a special case when $\psi(x,\omega) = c_0(\omega) + c_1(\omega)x + c_2(\omega)x^2 + c_3(\omega)x^3$. For $\xi \in \mathcal{F}_S$, we find

$$B(\xi(x)) = \int_S \underset{\sim}{f}(x)\,\underset{\sim}{f}^T(x)d\xi(x) = \begin{pmatrix} 1 & 0 & \mu_2 \\ 0 & \mu_2 & 0 \\ \mu_2 & 0 & \mu_4 \end{pmatrix}$$

and

$$B^{-1}(\xi(x)) = \begin{pmatrix} \dfrac{\mu_4}{\mu_4-\mu_2^2} & 0 & \dfrac{-\mu_2}{\mu_4-\mu_2^2} \\ 0 & \dfrac{1}{\mu_2} & 0 \\ \dfrac{-\mu_2}{\mu_4-\mu_2^2} & 0 & \dfrac{1}{\mu_4-\mu_2^2} \end{pmatrix}.$$

We also find

$$A = \int_S \underset{\sim}{f}(x)\,\underset{\sim}{f}^T(x)dx = \begin{pmatrix} 2 & 0 & \dfrac{2}{3} \\ 0 & \dfrac{2}{3} & 0 \\ \dfrac{2}{3} & 0 & \dfrac{2}{5} \end{pmatrix},$$

$$\begin{aligned} \underset{\sim}{b}^T(\psi(x,\omega),\xi(x)) &= \left(\int_S \psi(x,\omega)d\xi(x),\ \int_S x\psi(x,\omega)d\xi(x),\ \int_S x^2\psi(x,\omega)d\xi(x)\right) \\ &= \left(c_0(\omega)+c_2(\omega)\mu_2,\ c_1(\omega)\mu_2+c_3(\omega)\mu_4,\ c_0(\omega)\mu_2+c_2(\omega)\mu_4\right), \end{aligned}$$

and

$$\begin{aligned} tr\ AB^{-1}(\xi(x)) &= \frac{2\mu_4}{\mu_4-\mu_2^2} - \frac{2\mu_2}{3(\mu_4-\mu_2^2)} + \frac{2}{3\mu_2} - \frac{2\mu_2}{3(\mu_4-\mu_2^2)} + \frac{2}{5(\mu_4-\mu_2^2)} \\ &= \frac{30\mu_4 - 20\mu_2 + 6}{15(\mu_4-\mu_2^2)} + \frac{2}{3\mu_2}. \end{aligned}$$

Note that

$$
\begin{aligned}
&\underset{\sim}{f}^T(x)B^{-1}(\xi(x))\,\underset{\sim}{b}\,(\psi(x,\omega),\xi(x)) \\
&= (1,x,x^2)\begin{pmatrix} \frac{\mu_4}{\mu_4-\mu_2^2} & 0 & \frac{-\mu_2}{\mu_4-\mu_2^2} \\ 0 & \frac{1}{\mu_2} & 0 \\ \frac{-\mu_2}{\mu_4-\mu_2^2} & 0 & \frac{1}{\mu_4-\mu_2^2} \end{pmatrix}\begin{pmatrix} c_0(\omega)+c_2(\omega)\mu_2 \\ c_1(\omega)\mu_2+c_3(\omega)\mu_4 \\ c_0(\omega)\mu_2+c_2(\omega)\mu_4 \end{pmatrix} \\
&= (1,x,x^2)\begin{pmatrix} c_0(\omega) \\ c_1(\omega)+c_3(\omega)\frac{\mu_4}{\mu_2} \\ c_2(\omega) \end{pmatrix} \\
&= c_0(\omega)+c_1(\omega)x+c_2(\omega)x^2+c_3(\omega)\frac{\mu_4}{\mu_2}x.
\end{aligned}
$$

Hence, we have

$$
\begin{aligned}
&tr\ AB^{-1}(\xi(x))\,\underset{\sim}{b}\,(\psi(x,\omega),\xi(x))\,\underset{\sim}{b}^T(\psi(x,\omega),\xi(x))B^{-1}(\xi(x)) \\
&= \int_S(\underset{\sim}{f}^T(x)B^{-1}(\xi(x))\,\underset{\sim}{b}\,(\psi(x,\omega),\xi(x))][\underset{\sim}{f}^T(x)B^{-1}(\xi(x))\,\underset{\sim}{b}\,(\psi(x,\omega),\xi(x))]^T dx \\
&= \int_S[c_0(\omega)+c_1(\omega)x+c_2(\omega)x^2+c_3(\omega)\frac{\mu_4}{\mu_2}x]^2dx \\
&= 2c_0^2(\omega)+\frac{2}{3}c_1^2(\omega)+\frac{2}{5}c_2^2(\omega)+\frac{2}{3}c_3^2(\omega)\frac{\mu_4^2}{\mu_2^2}+\frac{2}{3}c_0(\omega)c_2(\omega)+\frac{2}{3}c_1(\omega)c_3(\omega)\frac{\mu_4}{\mu_2}.
\end{aligned}
$$

Finally, we get

$$
\begin{aligned}
\mathcal{L}_Q(\xi) &= \int_\Omega\left(\int_S MSE(\hat{y})dx\right)d\prod(\omega) \\
&= v\cdot trAB^{-1}(\xi(x))+\int_\Omega trAB^{-1}(\xi(x))\,\underset{\sim}{b}\,(\psi(x,\omega),\xi(x)) \\
&\qquad \cdot\underset{\sim}{b}^T(\psi(x,\omega),\xi(x))B^{-1}(\xi(x))d\prod(\omega) \\
&= v\left[\frac{30\mu_4-20\mu_2+6}{15(\mu_4-\mu_2^2)}+\frac{2}{3\mu_2}\right]+\int_\Omega[2c_0^2(\omega)+\frac{2}{3}c_1^2(\omega)+\frac{2}{5}c_2^2(\omega) \\
&\qquad +\frac{2}{3}c_3^2(\omega)+\frac{2}{3}c_0(\omega)c_2(\omega)+\frac{2}{3}c_1(\omega)c_3(\omega)\frac{\mu_4}{\mu_2}]d\prod(\omega) \\
&= v\left[\frac{30\mu_4-20\mu_2+6}{15(\mu_4-\mu_2^2)}+\frac{2}{3\mu_2}\right]+2\gamma_0+\frac{2}{3}\gamma_1+\frac{2}{5}\gamma_2+\frac{2}{3}\gamma_3\frac{\mu_4^2}{\mu_2^2},
\end{aligned}
$$

if $\int_\Omega c_0(\omega)c_2(\omega)d\prod(\omega) = \int_\Omega c_1(\omega)c_3(\omega)d\prod(\omega) = 0$ and $\gamma_i = \int_\Omega c_i^2(\omega)d\prod(\omega) < \infty,\quad i=0,1,2,3.$

It is clear that a symmetric design measure ξ_0 which minimizes $\mathcal{L}_Q(\xi)$ over $\mathcal{F}_S$ is equivalent to minimizing

$$L_0(\mu_2,\mu_4) := v\left[\frac{15\mu_4 - 10\mu_2 + 3}{15(\mu_4 - \mu_2^2)} + \frac{1}{3\mu_2}\right] + \frac{1}{3}\gamma_3\frac{\mu_4^2}{\mu_2^2} \tag{6.2.3}$$

over $\mathcal{F}_S$. $L_0(\mu_2,\mu_4)$ depends on ξ only through (μ_2,μ_4). The application of Theorem 4.2.5 yields

$$\min_{\xi\in\mathcal{F}_S} L_0(\mu_2,\mu_4) = \min_{\xi\in\mathcal{F}_0} L_0(\mu_2,\mu_4),$$

where $\mathcal{F}_0 = \{\xi:\ \xi\in\mathcal{F}_S,\ \ \xi = \frac{\alpha}{2}\Delta_{\pm\sqrt{x}} + (1-\alpha)\Delta_0,\ 0\le\alpha\le 1,\ 0\le x\le 1\}$. For any $\xi\in\mathcal{F}_0$, we have $\mu_2 = \int_S x^2 d\xi(x) = \alpha x$, and $\mu_4 = \int_S x^4 d\xi(x) = \alpha x^2$. Put $\mu_2 = \alpha x$, and $\mu_4 = \alpha x^2$ into (6.2.3), we find

$$L_0(\mu_2,\mu_4) := L_1(\alpha,x) = v\left[\frac{15\alpha x^2 - 10\alpha x + 3}{15\alpha(1-\alpha)x^2} + \frac{1}{3\alpha x}\right] + \frac{\gamma_3 x^2}{3}\,. \tag{6.2.4}$$

Hence, we have

$$\min_{\xi\in\mathcal{F}_S} L_0(\mu_2,\mu_4) = \min_{(\alpha,x)\in A} L_1(\alpha,x),$$

where $A = \{(\alpha,x):\ \ 0\le\alpha\le 1,\ 0\le x\le 1\}$. This yields the next theorem.

Theorem 6.2.3 *Assume* $\underset{\sim}{f}^T(x) = (1,x,x^2)$ *and* $\psi(x,\omega) = \sum_{i=0}^{3} c_i(\omega)x^i$, *where the* $c_i(\omega)$*'s are integrable real valued functions defined on* Ω *satisfying* $\gamma_i = \int_\Omega c_i^2(\omega)d\prod(\omega) < \infty$ *for* $i = 0,1,2,3$, *and* $\int_\Omega c_0(\omega)c_2(\omega)d\prod(\omega) = \int_\Omega c_1(\omega)c_3(\omega)d\prod(\omega) = 0$. *Then there exists a design measure* $\xi_0\in\mathcal{F}_S$ *that minimizes* $\mathcal{L}_Q(\xi)$ *over* $\mathcal{F}_S$ *and is of the form* $\xi_0 = \frac{\alpha_0}{2}\Delta_{\pm\sqrt{x_0}} + (1-\alpha_0)\Delta_0$, *where* $0\le\alpha_0\le 1$ *and* $0\le x_0\le 1$ *minimize* $L_1(\alpha,x)$ *in (6.2.4).*

Remark 1. The assumptions $\int_\Omega c_0(\omega)c_2(\omega)d\prod(\omega) = \int_\Omega c_1(\omega)c_3(\omega)d\prod(\omega) = 0$ in Theorem 6.2.3 are not necessary. The similar result can be found without these assumptions. However, when we consider the case $\psi(x,\omega) = \sum_{i=0}^{r} c_i(\omega)x^i$ for $r > 3$, the assumptions $\int_\Omega c_i(\omega)c_j(\omega)d\prod(\omega) = 0$ for $i\ne j$ and $i+j$ even will greatly simplify the problem.

Remark 2. The solutions in Theorem 6.2.1 and Theorem 6.2.3 are not unique. For example, (6.2.4) can also be minimized by a design measure ξ of the form $\xi = \frac{\alpha}{2}\triangle_{\pm 1} + \frac{(1-\alpha)}{2}\triangle_{\pm\sqrt{x}}$ for some $0 \le \alpha \le 1$ and $0 \le x \le 1$.

6.3 Optimal Designs for Approximately Multiple Linear Regression

The results in Section 6.2 and in Notz (1989) can be extended to the high dimensional case. In this section, we only consider the approximately linear regression situation. In this case, the problem of finding optimal designs can be easily solved. Let

$$y_i = y_i(\underset{\sim}{x}_i, \omega) = \underset{\sim}{\theta}^T \underset{\sim}{f}(\underset{\sim}{x}_i) + \psi(\underset{\sim}{x}_i, \omega) + \epsilon_i(\omega), \qquad i = 1, ..., n,$$

where $\epsilon_i(\omega)$ are independent and identically distributed with mean 0 and finite variance $\sigma^2 > 0$. We also assume $\underset{\sim}{x} = (x_1, ..., x_p) \in S := \{(x_1, ..., x_p) : -1 \leq x_j \leq 1, j = 1, ..., p\} \subseteq \mathbb{R}^p$, $\underset{\sim}{f}^T(\underset{\sim}{x}) = (f_0(\underset{\sim}{x}), ..., f_p(\underset{\sim}{x})) = (1, x_1, ..., x_p)$, $\underset{\sim}{\theta}^T = (\theta_0, \theta_1, ..., \theta_p)$ with $p \geq 2$, and

$$\psi(\underset{\sim}{x}, \omega) = a_0(\omega) + \sum_{i=1}^{p} a_i(\omega) x_i + \sum_{i=1}^{p} b_i(\omega) x_i^2 + \sum_{i \neq j} c_{ij} x_i x_j.$$

We again restrict ourself to consider the symmetric design measures defined on S. We denote $\mathcal{F}_p$ to be the set of all the symmetric design measures on S. For any $\xi \in \mathcal{F}_p$, we have

$$B(\xi(\underset{\sim}{x})) = \int_S \underset{\sim}{f}(\underset{\sim}{x}) \underset{\sim}{f}^T(\underset{\sim}{x}) d\xi(\underset{\sim}{x}) = \int_S \begin{pmatrix} 1 & x_1 & \dots & x_p \\ x_1 & x_1^2 & \dots & x_1 x_p \\ \dots & \dots & \dots & \dots \\ x_p & x_p x_1 & \dots & x_p^2 \end{pmatrix} d\xi(\underset{\sim}{x})$$

$$= \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & \int x_1^2 d\xi(\underset{\sim}{x}) & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \int x_p^2 d\xi(\underset{\sim}{x}) \end{pmatrix} = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & \lambda & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \lambda \end{pmatrix},$$

where $\lambda = \int_S x_i^2 d\xi(\underset{\sim}{x}) \quad i = 1, ..., p$, and

$$\begin{aligned}
\underset{\sim}{b}^T\ (\psi(\underset{\sim}{x}, \omega), \xi(\underset{\sim}{x})) &= \left(\int_S \psi(\underset{\sim}{x}, \omega) d\xi(\underset{\sim}{x}), \int_S x_1 \psi(\underset{\sim}{x}, \omega) d\xi(\underset{\sim}{x}), \ \dots\ , \int_S x_p \psi(\underset{\sim}{x}, \omega) d\xi(\underset{\sim}{x})\right) \\
&= \left(a_0(\omega) + \lambda \sum_{i=1}^p b_i(\omega),\ a_1(\omega)\lambda, ..., a_p(\omega)\lambda\right).
\end{aligned}$$

According to Lemma 6.1.1, we have

$$\begin{aligned}
\mathcal{L}_D(\xi) &= \left|vB^{-1}(\xi(\underset{\sim}{x}))\right| \cdot \Big\{1 + v^{-1} \int_\Omega \underset{\sim}{b}^T\ (\psi(\underset{\sim}{x}, \omega), \xi(\underset{\sim}{x})) \\
&\qquad \cdot B^{-1}(\xi(\underset{\sim}{x}))\ \underset{\sim}{b}\ (\psi(\underset{\sim}{x}, \omega), \xi(\underset{\sim}{x})) d\textstyle\prod(\omega)\Big\} \\
&= \frac{v}{\lambda^p}\{1 + v^{-1} \int_\Omega \left(a_0(\omega) + \lambda \sum_{i=1}^p b_i(\omega),\ a_1(\omega)\lambda,\ \dots\ , a_p(\omega)\lambda\right) \\
&\qquad \cdot \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & \frac{1}{\lambda} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{1}{\lambda} \end{pmatrix} \begin{pmatrix} a_0(\omega) + \lambda \sum_{i=1}^p b_i(\omega) \\ a_1(\omega)\lambda \\ \vdots \\ a_p(\omega)\lambda \end{pmatrix} d\textstyle\prod(\omega)\} \\
&= \frac{v}{\lambda^p}\{1 + v^{-1} \int_\Omega [(a_0(\omega) + \lambda \sum_{i=1}^p b_i(\omega))^2 \\
&\qquad + a_1^2(\omega)\lambda + \dots + a_p^2(\omega)\lambda] d\textstyle\prod(\omega)\} \\
&= \frac{1}{\lambda^p}\left(v + \alpha_0 + \lambda^2 \sum_{i=1}^p \beta_i + \lambda \sum_{i=1}^p \alpha_i\right),
\end{aligned} \tag{6.3.1}$$

if $\int_\Omega a_0(\omega) b_i(\omega) d\prod(\omega) = 0, \quad i = 1, ..., p$, and $\int_\Omega b_i(\omega) b_j(\omega) d\prod(\omega) = 0$ for all $i \neq j$; $\alpha_i = \int_\Omega a_i^2(\omega) d\prod(\omega) < \infty, \quad i = 0, ..., p$, and $\beta_i = \int_\Omega b_i^2(\omega) d\prod(\omega) < \infty, \quad i = 1, ..., p$.

It is clear that the range of λ is between 0 and 1, and any value of λ within its range can be achieved by choosing a design measure ξ of the form $\xi(x_1, ..., x_p) = \frac{1}{2^p}$ where $x_i = \pm\sqrt{z}$ for some $0 \le z \le 1,\ i = 1, ..., p$. Putting $\lambda = \int_S x_i^2 d(\xi(\underset{\sim}{x})) = z$ into (6.3.1), we get

$$\begin{aligned}
\mathcal{L}_D(\xi) &= \frac{1}{z^p}\left(v + \alpha_0 + z^2 \sum_{i=1}^p \beta_i + z \sum_{i=1}^p \alpha_i\right) \\
&= (v + \alpha_0)\frac{1}{z^p} + \frac{1}{z^{p-2}} \cdot \sum_{i=1}^p \beta_i + \frac{1}{z^{p-1}} \cdot \sum_{i=1}^p \alpha_i := L_D(z).
\end{aligned}$$

We find

$$\frac{dL_D(z)}{dz} = (v+\alpha_0)\frac{(-p)}{z^{p+1}} + \frac{-(p-2)}{z^{p-1}}\sum_{i=1}^{p}\beta_i + \frac{-(p-1)}{z^p}\sum_{i=1}^{p}\alpha_i < 0,$$

since we assume $p \geq 2$. Hence $L_D(\xi)$ is decreasing in z, and

$$\min_{0\leq z\leq 1} L_D(z) = L_D(1) = v + \sum_{i=0}^{p}\alpha_i + \sum_{i=1}^{p}\beta_i\ ,$$

and the corresponding optimal design measure is ξ_0, where $\xi_0(x_1, ..., x_p) = \dfrac{1}{2^p}$ with $x_i = \pm 1,\ i = 1, ..., p$.

According to Lemma 6.1.2, we have

$$\begin{aligned}
\mathcal{L}_A(\xi) &= \int_\Omega tr\ MSE(\hat{\theta})d\prod(\omega) \\
&= v\cdot trB^{-1}(\xi(\underset{\sim}{x})) + \int_\Omega \underset{\sim}{b}^T\ (\psi(\underset{\sim}{x},\omega),\xi(\underset{\sim}{x})) \\
&\qquad \cdot B^{-2}(\xi(\underset{\sim}{x}))\ \underset{\sim}{b}\ (\psi(\underset{\sim}{x},\omega),\xi(\underset{\sim}{x}))d\prod(\omega) \\
&= v(1+\frac{p}{\lambda}) + \int_\Omega (a_0(\omega)+\lambda\sum_{i=1}^{p}b_i(\omega),\ a_1(\omega)\lambda,\ \dots\ , a_p(\omega)\lambda) \\
&\qquad \cdot \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & \frac{1}{\lambda^2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \frac{1}{\lambda^2} \end{pmatrix} \begin{pmatrix} a_0(\omega)+\lambda\sum_{i=1}^{p}b_i(\omega) \\ a_1(\omega)\lambda \\ \vdots \\ a_p(\omega)\lambda \end{pmatrix} d\prod(\omega) \\
&= v(1+\frac{p}{\lambda}) + \textstyle\int_\Omega[(a_0(\omega)+\lambda\sum_{i=1}^{p}b_i(\omega))^2 + \sum_{i=1}^{p}a_i^2(\omega)]d\prod(\omega) \\
&= v(1+\frac{p}{\lambda}) + \alpha_0 + \lambda^2\sum_{i=1}^{p}\beta_i + \sum_{i=1}^{p}\alpha_i \\
&= \left(v+\sum_{i=0}^{p}\alpha_i\right) + \frac{vp}{\lambda} + \lambda^2\sum_{i=1}^{p}\beta_i,
\end{aligned} \tag{6.3.2}$$

if $\int_\Omega a_0(\omega)b_i(\omega)d\prod(\omega) = 0,\ i = 1, ..., p$, and $\int_\Omega b_i(\omega)b_j(\omega)d\prod(\omega) = 0$ for all $i \neq j$; $\alpha_i = \int_\Omega a_i^2(\omega)d\prod(\omega) < \infty,\ i = 0, ..., p$, and $\beta_i = \int_\Omega b_i^2(\omega)d\prod(\omega) < \infty,\ i = 1, ..., p$.

Similar to D-optimal case, (6.3.2) can be minimized by a design measure ξ of the form $\xi(x_1, ..., x_p) = \frac{1}{2^p}$ with $x_i = \pm\sqrt{z}$ for some $0 \leq z \leq 1,\ i = 1, ..., p$. Put

$\lambda = \int_S x_i^2 d\xi(\underset{\sim}{x}) = z$ into (6.3.2), we have

$$\mathcal{L}_A(\xi) = \left(v + \sum_{i=0}^{p} \alpha_i\right) + \frac{vp}{z} + z^2 \sum_{i=1}^{p} \beta_i := L_A(z).$$

Solving $\dfrac{dL_A(z)}{dz} = -\dfrac{vp}{z^2} + 2z\sum_{i=1}^{p}\beta_i := 0$, we find $2z\sum_{i=1}^{p}\beta_i = \dfrac{vp}{z^2}$ and hence $z = \sqrt[3]{vp/2\sum_{i=1}^{p}\beta_i}$.

Note that

$$\begin{aligned}
&\underset{\sim}{f}^T(\underset{\sim}{x})B^{-1}(\xi(\underset{\sim}{x}))\,\underset{\sim}{b}\,(\psi(\underset{\sim}{x},\omega),\xi(\underset{\sim}{x}))\\
&= (1, x_1, ..., x_p)\begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & \frac{1}{\lambda} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{1}{\lambda} \end{pmatrix}\begin{pmatrix} a_0(\omega) + \lambda\sum_{i=1}^{p} b_i(\omega) \\ a_1(\omega)\lambda \\ \vdots \\ a_p(\omega)\lambda \end{pmatrix}\\
&= a_0(\omega) + \lambda\sum_{i=1}^{p} b_i(\omega) + \sum_{i=1}^{p} a_i(\omega)x_i
\end{aligned}$$

and

$$\begin{aligned}
A = \int_S \underset{\sim}{f}\,(\underset{\sim}{x})\,\underset{\sim}{f}^T\,(\underset{\sim}{x})d\,\underset{\sim}{x} \;&=\; \int_S \begin{pmatrix} 1 & x_1 & \dots & x_p \\ x_1 & x_1^2 & \dots & x_1x_p \\ \dots & \dots & \dots & \dots \\ x_p & x_px_1 & \dots & x_p^2 \end{pmatrix} d\,\underset{\sim}{x}\\
&= \begin{pmatrix} 2^p & 0 & \dots & 0 \\ 0 & \frac{1}{3}2^p & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \frac{1}{3}2^p \end{pmatrix}.
\end{aligned}$$

Hence, for Q−optimality, we have

$$\begin{aligned}
\int_S MSE(\hat{y})d\underset{\sim}{x} &= v\int_S \underset{\sim}{f}^T(\underset{\sim}{x})B^{-1}(\xi(\underset{\sim}{x}))\underset{\sim}{f}(\underset{\sim}{x})d\underset{\sim}{x} \\
&\quad + \int_S \underset{\sim}{f}^T(\underset{\sim}{x})B^{-1}(\xi(\underset{\sim}{x}))\underset{\sim}{b}(\psi(\underset{\sim}{x},\omega),\xi(\underset{\sim}{x})) \\
&\quad \cdot \underset{\sim}{b}^T(\psi(\underset{\sim}{x},\omega),\xi(\underset{\sim}{x}))B^{-1}(\xi(\underset{\sim}{x}))\underset{\sim}{f}(\underset{\sim}{x})d\underset{\sim}{x} \\
&= v\cdot trAB^{-1}(\xi(\underset{\sim}{x})) + \int_S \left(a_0(\omega)+\lambda\sum_{i=1}^p b_i(\omega)+\sum_{i=1}^p a_i(\omega)x_i\right)^2 d\underset{\sim}{x} \\
&= v\left[2^p+\frac{2^p}{3}(\frac{1}{\lambda}+\cdots+\frac{1}{\lambda})\right] + 2^p\left(a_0(\omega)+\lambda\sum_{i=1}^p b_i(\omega)\right)^2 \\
&\quad + \int_S \sum_{i=1}^p a_i^2(\omega)x_i^2 d\underset{\sim}{x} + 2\int_S \sum\sum_{i<j} a_i(\omega)a_j(\omega)x_ix_j d\underset{\sim}{x} \\
&= 2^p v\left(1+\frac{p}{3\lambda}\right) + 2^p\left(a_0(\omega)+\lambda\sum_{i=1}^p b_i(\omega)\right)^2 + \frac{2^p}{3}\sum_{i=1}^p a_i^2(\omega).
\end{aligned}$$

Therefore, we get

$$\begin{aligned}
\mathcal{L}_Q(\xi) &= \int_\Omega\left(\int_S MSE(\hat{y})d\underset{\sim}{x}\right)d\prod(\omega) \\
&= 2^p v\left(1+\frac{p}{3\lambda}\right) + 2^p\left(\alpha_0+\lambda^2\sum_{i=1}^p \beta_i\right) + \frac{2^p}{3}\sum_{i=1}^p \alpha_i,
\end{aligned}$$

if $\int_\Omega a_0(\omega)b_i(\omega)d\prod(\omega) = 0$, $i = 1,...,p$, and $\int_\Omega b_i(\omega)b_j(\omega)d\prod(\omega) = 0$ for all $i \neq j$; $\alpha_i = \int_\Omega a_i^2(\omega)d\prod(\omega) < \infty$, $i = 0,...,p$, and $\beta_i = \int_\Omega b_i^2(\omega)d\prod(\omega) < \infty$, $i = 1,...,p$.

It is clear that the minimization of $\mathcal{L}_Q(\xi)$ is equivalent to the minimization of the following:

$$\frac{vp}{3\lambda} + \lambda^2\sum_{i=1}^p \beta_i, \tag{6.3.3}$$

which can be minimized by a design measure ξ of the form $\xi(x_1,...,x_p) = \frac{1}{2^p}$ with $x_i = \pm\sqrt{z}$ for some $0 \le z \le 1$, $i = 1,...,p$. Putting $\lambda = \int_S x_i^2 d(\xi(\underset{\sim}{x})) = z$ into (6.3.3), we have

$$\frac{vp}{3z} + z^2\sum_{i=1}^p \beta_i := L_Q(z).$$

Solving $\frac{dL_Q(z)}{dz} = -\frac{vp}{3z^2} + 2z\sum_{i=1}^p \beta_i = 0$, we find $z = \sqrt[3]{vp/6\sum_{i=1}^p \beta_i}$.

We summarize the above results as the next theorem.

Theorem 6.3.1 *Assume that* $\underset{\sim}{f}^T(\underset{\sim}{x}) = (1, x_1, ..., x_p)$, *and* $\psi(\underset{\sim}{x}, \omega) = a_0(\omega) + \sum_{i=1}^{p} a_i(\omega)x_i + \sum_{i=1}^{p} b_i(\omega)x_i^2 + \sum\sum_{i \neq j} c_{ij}x_ix_j$ *where* $\alpha_i = \int_\Omega a_i^2(\omega)d\prod(\omega) < \infty$, $i = 0, 1, ..., p$, $\beta_i = \int_\Omega b_i^2(\omega)d\prod(\omega) < \infty$, $i = 1, ..., p$, *and* $\int_\Omega a_0(\omega)b_i(\omega)d\prod(\omega) = 0$, $i = 1, ..., p$, $\int_\Omega b_i(\omega)b_j(\omega)d\prod(\omega) = 0$ *for all* $i \neq j$. *Within the class of* $\mathcal{F}_p$, *we have that*

(i) $\mathcal{L}_D(\xi)$ *is minimized by* $\xi_0(x_1, ..., x_p) = \frac{1}{2^p}$, *where* $x_i = \pm 1 \quad i = 1, ..., p$, *and*

$$\min_{\xi \in \mathcal{F}_p} \mathcal{L}_D(\xi) = \mathcal{L}_D(\xi_0) = v + \sum_{i=0}^{p} \alpha_i + \sum_{i=1}^{p} \beta_i.$$

(ii) $\mathcal{L}_A(\xi)$ *is minimized by* $\xi_0(x_1, ..., x_p) = \frac{1}{2^p}$ *with* $x_i = \pm\sqrt{z} \quad i = 1, ..., p$, *where* $z = \min\left\{1, \sqrt[3]{vp/2\sum_{i=1}^{p}\beta_i}\right\}$.

(iii) $\mathcal{L}_Q(\xi)$ *is minimized by* $\xi_0(x_1, ..., x_p) = \frac{1}{2^p}$ *with* $x_i = \pm\sqrt{z} \quad i = 1, ..., p$, *where* $z = \min\left\{1, \sqrt[3]{vp/6\sum_{i=1}^{p}\beta_i}\right\}$.

6.4 Optimal Designs for Two Dimensional Linear Regression with Interaction Term

Linear regression with an interaction term is a very common and useful model in regression analysis. There is an advantage to considering this model. For any symmetric design measure, we have that $B^{-1}(\xi(\underset{\sim}{x}))$ is a diagonal matrix. In this case, the problem of finding optimal designs is very easy. The approach to the problem in this section is very similar to Section 6.3.

Let $\underset{\sim}{f}^T(\underset{\sim}{x}) = (1, x_1, x_2, x_1x_2)$, $\psi(\underset{\sim}{x}, \omega) = a_0(\omega) + a_1(\omega)x_1 + a_2(\omega)x_2 + b_1(\omega)x_1^2 + b_2(\omega)x_2^2 + c_{12}x_1x_2$, $S = \{(x_1, x_2) : -1 \le x_j \le 1,\ j = 1, 2\} \subseteq \mathbf{R}^2$, and $\mathcal{F}_2 = \{\xi : \xi$ is symmetric design measure defined on $S\}$. We find

$$
\begin{aligned}
B(\xi(\underset{\sim}{x})) &= \int_S \underset{\sim}{f}(\underset{\sim}{x})\, \underset{\sim}{f}^T(\underset{\sim}{x}) d\xi(\underset{\sim}{x}) \\
&= \int_S \begin{pmatrix} 1 & x_1 & x_2 & x_1x_2 \\ x_1 & x_1^2 & x_1x_2 & x_1^2x_2 \\ x_2 & x_1x_2 & x_2^2 & x_1x_2^2 \\ x_1x_2 & x_1^2x_2 & x_1x_2^2 & x_1^2x_2^2 \end{pmatrix} d\xi(\underset{\sim}{x}) \\
&= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \int_S x_1^2 d\xi(\underset{\sim}{x}) & 0 & 0 \\ 0 & 0 & \int_S x_2^2 d\xi(\underset{\sim}{x}) & 0 \\ 0 & 0 & 0 & \int_S x_1^2x_2^2 d\xi(\underset{\sim}{x}) \end{pmatrix} \\
&= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \lambda & 0 & 0 \\ 0 & 0 & \lambda & 0 \\ 0 & 0 & 0 & \lambda_{12} \end{pmatrix},
\end{aligned}
$$

where $\lambda = \int_S x_i^2 d\xi(\underset{\sim}{x})$ $i = 1,2$ and $\lambda_{12} = \int_S x_1^2 x_2^2 d\xi(\underset{\sim}{x})$, and

$$\begin{aligned}
\underset{\sim}{b}\left(\psi(\underset{\sim}{x},\omega),\xi(\underset{\sim}{x})\right) &= (\int_S \psi(\underset{\sim}{x},\omega)d\xi(\underset{\sim}{x}), \int_S x_1\psi(\underset{\sim}{x},\omega)d\xi(\underset{\sim}{x}),\\
&\qquad \int_S x_2\psi(\underset{\sim}{x},\omega)d\xi(\underset{\sim}{x}), \int_S x_1x_2\psi(\underset{\sim}{x},\omega)d\xi(\underset{\sim}{x}))\\
&= (a_0(\omega)+(b_1(\omega)+b_0(\omega))\lambda,\ a_1(\omega)\lambda, a_2(\omega)\lambda, c_{12}(\omega)\lambda_{12}).
\end{aligned}$$

For $D-$optimality, we have

$$\begin{aligned}
\mathcal{L}_D(\xi) &= \int_\Omega \left|MSE(\hat{\theta})\right| d\prod(\omega)\\
&= \det B^{-1}(\xi(\underset{\sim}{x}))\cdot[v + \int_\Omega \underset{\sim}{b}^T\left(\psi(\underset{\sim}{x},\omega),\xi(\underset{\sim}{x})\right)\\
&\qquad \cdot B^{-1}(\xi(\underset{\sim}{x}))\ \underset{\sim}{b}\left(\psi(\underset{\sim}{x},\omega),\xi(\underset{\sim}{x})\right)d\prod(\omega)]\\
&= \frac{1}{\lambda^2\lambda_{12}}\{v + \int_\Omega (a_0(\omega)+(b_1(\omega)+b_2(\omega))\lambda,\ a_1(\omega)\lambda, a_2(\omega)\lambda, c_{12}(\omega)\lambda_{12})\\
&\qquad \cdot \begin{pmatrix} 1 & 0 & 0 & 0\\ 0 & \frac{1}{\lambda} & 0 & 0\\ 0 & 0 & \frac{1}{\lambda} & 0\\ 0 & 0 & 0 & \frac{1}{\lambda}\end{pmatrix}\begin{pmatrix} a_0(\omega)+(b_1(\omega)+b_2(\omega))\lambda\\ a_1(\omega)\lambda\\ a_2(\omega)\lambda\\ c_{12}(\omega)\lambda_{12}\end{pmatrix} d\prod(\omega)\}\\
&= \frac{1}{\lambda^2\lambda_{12}}\{v + \int_\Omega [(a_0(\omega)+(b_1(\omega)+b_2(\omega))\lambda)^2\\
&\qquad +(a_1^2(\omega)+a_2^2(\omega))\lambda + c_{12}^2(\omega)\lambda_{12}]d\prod(\omega)\}\\
&= \frac{1}{\lambda^2\lambda_{12}}\{v+\alpha_0+(\alpha_1+\alpha_2)\lambda+(\beta_1+\beta_2)\lambda^2+\gamma_{12}\lambda_{12}\},
\end{aligned}$$

(6.4.1)

if $\alpha_i = \int_S a_i^2(\omega)d\prod(\omega) < \infty$, $i = 0,1,2$, $\beta_i = \int_S b_i^2(\omega)d\prod(\omega) < \infty$, $i = 1,2$, $\gamma_{12} = \int_S c_{12}^2(\omega)d\prod(\omega) < \infty$ and $\int_S a_0(\omega)b_1(\omega)d\prod(\omega) = \int_S a_0(\omega)b_2(\omega)d\prod(\omega) = \int_S b_1(\omega)b_2(\omega)d\prod(\omega) = 0$.

It is clear that Theorem 4.2.5 can be extended to the high dimensional case. For any $\xi(\underset{\sim}{x}) \in \mathcal{F}_2$, let $T(\xi) = \left(\int_S x_1^2 d\xi(\underset{\sim}{x}),\ \int_S x_1^2x_2^2 d\xi(\underset{\sim}{x})\right)$ and $\mathcal{E}_2 = \{\xi : \xi(\pm\sqrt{z},\ \pm\sqrt{z}) = \frac{1}{4},\ 0 \le z \le 1\}$ which is the set of extreme points of $\mathcal{F}_2$. Then we have $T(\mathcal{E}_2) = \{(u,v) :\ 0 \le u \le 1,\ v = u^2\} := S_2$ and $T(\mathcal{F}_2) = \hat{S}_2$ where $\hat{S}_2$ is the convex hull of S_2. It is easy to see that any point $p \in \hat{S}_2$ is the image of ξ under T where $\xi(\pm\sqrt{z},\ \pm\sqrt{z}) = \alpha,\ \xi(0,0) = 1-4\alpha$ for some $0 \le \alpha \le \frac{1}{4}$ and $0 \le z \le 1$. If we denote $\mathcal{F}^*$ to be the set containing all these design measures, i.e. $\mathcal{F}^* = \{\xi : \xi \in \mathcal{F}_2$ $\xi(\pm\sqrt{z}, \pm\sqrt{z}) = \alpha,\ \xi(0,0) = 1 - 4\alpha,\ 0 \le \alpha \le \frac{1}{4}$ and $0 \le z \le 1\}$, then we know that the minimization of (6.4.1) over $\mathcal{F}_2$ is equivalent to the minimization of (6.4.1) over $\mathcal{F}^*$. For any $\xi \in \mathcal{F}^*$, we have

$$\lambda = \int_S x_i^2 d\xi(\underset{\sim}{x}) = 4\alpha z \text{ and } \lambda_{12} = \int_S x_1^2 x_2^2 d\xi(\underset{\sim}{x}) = 4\alpha z^2. \tag{6.4.2}$$

Putting (6.4.2) into (6.4.1), we find

$$\begin{aligned} \mathcal{L}_D(\xi) \ &:= \ L_D(\alpha, z) = \frac{1}{64\alpha^3 z^4}[v + \alpha_0 + 4(\alpha_1 + \alpha_2)\alpha z \\ &\qquad + 16(\beta_1 + \beta_2)\alpha^2 z^2 + 4\gamma_{12}\alpha z^2]. \end{aligned} \tag{6.4.3}$$

For $A-$optimality, we have

$$\begin{aligned} \mathcal{L}_A(\xi) &= \int_\Omega tr\ MSE(\hat{\theta}) d\prod(\omega) \\ &= v \cdot trB^{-1}(\xi(\underset{\sim}{x})) + \int_\Omega \underset{\sim}{b}^T\ (\psi(\underset{\sim}{x}, \omega), \xi(\underset{\sim}{x})) \\ &\qquad \cdot B^{-2}(\xi(\underset{\sim}{x}))\ \underset{\sim}{b}\ (\psi(\underset{\sim}{x}, \omega), \xi(\underset{\sim}{x})) d\prod(\omega) \\ &= v(1 + \frac{2}{\lambda} + \frac{1}{\lambda_{12}}) + \int_\Omega (a_0(\omega) + (b_1(\omega) + b_2(\omega))\lambda,\ a_1(\omega)\lambda, a_2(\omega)\lambda, \\ &\qquad c_{12}(\omega)\lambda_{12}) \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \frac{1}{\lambda^2} & 0 & 0 \\ 0 & 0 & \frac{1}{\lambda^2} & 0 \\ 0 & 0 & 0 & \frac{1}{\lambda_{12}^2} \end{pmatrix} \begin{pmatrix} a_0(\omega) + (b_1(\omega) + b_2(\omega))\lambda \\ a_1(\omega)\lambda \\ a_2(\omega)\lambda \\ c_{12}(\omega)\lambda_{12} \end{pmatrix} d\prod(\omega) \\ &= v(1 + \frac{2}{\lambda} + \frac{1}{\lambda_{12}}) + \int_\Omega [(a_0(\omega) + (b_1(\omega) + b_2(\omega))\lambda)^2 \\ &\qquad + a_1^2(\omega) + a_2^2(\omega) + c_{12}^2(\omega)] d\prod(\omega) \\ &= v(1 + \frac{2}{\lambda} + \frac{1}{\lambda_{12}}) + \alpha_0 + (\beta_1 + \beta_2)\lambda^2 + \alpha_1 + \alpha_2 + \gamma_{12} \\ &= v + \sum_{i=0}^{2} \alpha_i + \gamma_{12} + v(\frac{2}{\lambda} + \frac{1}{\lambda_{12}}) + (\beta_1 + \beta_2)\lambda^2. \end{aligned} \tag{6.4.4}$$

Similar to the case of $D-$optimality, we only need to search for the optimal result within $\mathcal{F}^*$. Putting (6.4.2) into (6.4.4) we find

$$\begin{aligned} \mathcal{L}_A(\xi) \ &:= \ L_A(\alpha, z) \\ &= \ v + \sum_{i=0}^{2} \alpha_i + \gamma_{12} + \frac{v}{4\alpha z}(2 + \frac{1}{z}) + 16(\beta_1 + \beta_2)\alpha^2 z^2. \end{aligned} \tag{6.4.5}$$

It is obvious that the minimization of (6.4.5) is equivalent to the minimization of the following:

$$\frac{v}{4\alpha z}(2+\frac{1}{z})+16(\beta_1+\beta_2)\alpha^2 z^2. \tag{6.4.6}$$

For $Q-$optimality, we have

$$\begin{aligned}
\int_S MSE(\hat{y})d\,\underset{\sim}{x} &= v\int_S \underset{\sim}{f}^T(\underset{\sim}{x})B^{-1}(\xi(\underset{\sim}{x}))\,\underset{\sim}{f}(\underset{\sim}{x})d\,\underset{\sim}{x} \\
&+\int_S \underset{\sim}{f}^T(\underset{\sim}{x})B^{-1}(\xi(\underset{\sim}{x}))\,\underset{\sim}{b}(\psi(\underset{\sim}{x},\omega),\xi(\underset{\sim}{x})) \\
&\qquad \cdot\underset{\sim}{b}^T(\psi(\underset{\sim}{x},\omega),\xi(\underset{\sim}{x}))B^{-1}(\xi(\underset{\sim}{x}))\,\underset{\sim}{f}(\underset{\sim}{x})d\,\underset{\sim}{x} \\
&= v\int_S\left(1+\frac{x_1^2}{\lambda}+\frac{x_2^2}{\lambda}+\frac{x_1^2x_2^2}{\lambda_{12}}\right)d\,\underset{\sim}{x} \\
&+\int_S[a_0(\omega)+(b_1(\omega)+b_2(\omega))\lambda \\
&+a_1(\omega)x_1+a_2(\omega)x_2+c_{12}(\omega)x_1x_2]^2 d\,\underset{\sim}{x} \\
&= 4v\left(1+\frac{2}{3\lambda}+\frac{1}{9\lambda_{12}}\right)+4\{[a_0(\omega)+(b_1(\omega)+b_2(\omega))\lambda]^2 \\
&+\frac{1}{3}a_1^2(\omega)+\frac{1}{3}a_2^2(\omega)+\frac{1}{9}c_{12}^2(\omega)\}.
\end{aligned}$$

Hence, we get

$$\begin{aligned}
\mathcal{L}_Q(\xi) &= \int_\Omega(\int_S MSE(\hat{y})d\,\underset{\sim}{x})d\prod(\omega) \\
&= 4v\left(1+\frac{2}{3\lambda}+\frac{1}{9\lambda_{12}}\right) \\
&+4\left[\alpha_0+(\beta_1+\beta_2)\lambda^2+\frac{1}{3}\alpha_1+\frac{1}{3}\alpha_2+\frac{1}{9}\gamma_{12}\right] \\
&= 4\left(v+\alpha_0+\frac{1}{3}\alpha_1+\frac{1}{3}\alpha_2+\frac{1}{9}\gamma_{12}\right) \\
&+4\left[v\left(\frac{2}{3\lambda}+\frac{1}{9\lambda_{12}}\right)+(\beta_1+\beta_2)\lambda^2\right].
\end{aligned} \tag{6.4.7}$$

Again, put (6.4.2) into (6.4.7), we find

$$\begin{aligned}
\mathcal{L}_Q(\xi) &:= L_Q(\alpha,z)=4\left(v+\alpha_0+\frac{1}{3}\alpha_1+\frac{1}{3}\alpha_2+\frac{1}{9}\gamma_{12}\right) \\
&+4\left[v\left(\frac{1}{6\alpha z}+\frac{1}{36\alpha z^2}\right)+16(\beta_1+\beta_2)\alpha^2z^2\right],
\end{aligned} \tag{6.4.8}$$

and the minimization of $\mathcal{L}_Q(\xi)$ is equivalent to the minimization of the following:

$$\frac{v}{6\alpha z}\left(1+\frac{1}{6z}\right)+16(\beta_1+\beta_2)\alpha^2z^2. \tag{6.4.9}$$

We summarize these results as the next theorem.

Theorem 6.4.1 *Assume that* $\underset{\sim}{f}^T(\underset{\sim}{x}) = (1, x_1, x_2, x_1x_2)$, $\psi(\underset{\sim}{x}, \omega) = a_0(\omega) + a_1(\omega)x_1 + a_2(\omega)x_2 + b_1(\omega)x_1^2 + b_2(\omega)x_2^2 + c_{12}(\omega)x_1x_2$, *where* $\alpha_i = \int_\Omega a_i^2(\omega)d\prod(\omega) < \infty$, $i = 0, 1, 2$, $\beta_i = \int_\Omega b_i^2(\omega)d\prod(\omega) < \infty$, $i = 1, 2$, $\gamma_{12} = \int_\Omega c_{12}^2(\omega)d\prod(\omega) < \infty$ *and* $\int_\Omega a_0(\omega)b_1(\omega)d\prod(\omega) = \int_\Omega a_0(\omega)b_2(\omega)d\prod(\omega) = \int_\Omega b_1(\omega)b_2(\omega)d\prod(\omega) = 0$. *Let* ξ_0 *be a design measure of the form* $\xi_0(\pm\sqrt{z}, \pm\sqrt{z}) = \alpha$, $\xi_0(0,0) = 1 - 4\alpha$. *Within the class of* $\mathcal{F}_2$, *we have* $\mathcal{L}_D(\xi)$, $\mathcal{L}_A(\xi)$, *or* $\mathcal{L}_Q(\xi)$ *is minimized by* ξ_0 *for some* $0 \le \alpha \le \frac{1}{4}$ *and* $0 \le z \le 1$ *which minimize (6.4.3), (6.4.6), or (6.4.9) respectively.*

Remark. The above result can be easily extended to the case when $p > 2$, where p is the dimension of the design space.

Bibliography

[1] ANDREWS, D.F., BICKEL, P.J., HAMPEL, F.R., HUBER, P.J., ROGERS, W.H. and TUKEY, J.W. (1972). Robust Estimates of Location: Survey and Advances. Princeton Univ. Press.

[2] ATKINSON, A.C. (1972). Planning experiments to detect inadequate regression models. Biometrika 59, 2, 275-293.

[3] ATKINSON, A.C. (1982). Developments in the design of experiments. International Statistical Review 50, 161-177.

[4] ATWOOD, C.L. (1969). Optimal and efficient designs of experiments. Ann. Math. Statist. 40, 5, 1570-1602.

[5] ATWOOD, C.L. (1971). Robust procedures for estimating polynomial regression. J. Amer. Statist. Assoc. 66, 855-860.

[6] BICKEL, P.J. (1965). On some robust estimates of location. Ann. Math. Statist., 43, 847-858.

[7] BICKEL, P.J. and HERZBERG, A.M. (1979). Robustness of design against autocorrelation in time I: Asymptotic theory, optimality for location and linear regression. Ann. Statist. 7, 1, 77-95.

[8] BICKEL, P.J., HERZBERG, A.M. and SCHILLING, M.F. (1981). Robustness of design against autocorrelation in time II: Optimality, theoretical and numerical results for the first-order autoregressive process. J. Amer. Statist. Assoc. 76, 870-877.

[9] BILLINGSLEY, P. (1986). Probability and Measure. Second Edition, Wiley, New York.

[10] BOOS, D.D. and SERFLING, R.J. (1980). A note on differentials and the CLT and LIL for statistical functions, with application to M-estimates. Ann. Statist. 8, 618-624.

[11] BOX, G.E.P. and DRAPER, N.R. (1959). A basis for the selection of a response surface design. J. Amer. Statist. Assoc. 54, 622-654.

[12] BOX, G.E.P. and DRAPER, N.R. (1963). The choice of a second order rotatable design. Biometrika 50, 3 and 4, 335-352.

[13] BOX, G.E.P. and DRAPER, N.R. (1975). Robust designs. Biometrika 62, 2, 347-352.

[14] BOX, G.E.P., HUNTER, W.G. and HUNTER, J.S. Statistics for Experimenters: An Introduction to Design, Data Analysis, and Model Building, Wiley, New York.

[15] CARROLL, R.J. and RUPPERT, D. (1979). On the asymptotic normality of robust regression estimates. Mimeo Series No. 1242. Institute of Statistics, Chapel Hill, North Carolina.

[16] COCHRAN, W.G. and COX, G.M. (1957). Experimental Designs Second Edition. Wiley, New York.

[17] COLLINS, J.R. (1976). Robust estimation of a location parameter in the presence of asymmetry. Ann. Statist. 4, 1, 68-85.

[18] COLLINS, J.R. (1977). Upper bounds on asymptotic variances of M-estimators of location. Ann. Statist. 5, 1, 646-657.

[19] COLLINS, J.R. (1986). Maximum asymptotic variances of trimmed means under asymmetric contamination. Ann. Statist. 14, 1, 348-354.

[20] COLLINS, J.R. (1991). Maximal asymptotic biases and variances of symmetrized interquantile ranges under asymmetric contamination. Statistics 22, 3, 379-402.

[21] COLLINS, J.R. and PORTNOY, S.L. (1981). Maximizing the variance of M-estimators using the generalized method of moment spaces. Ann. Statist. 9, 3, 567-577.

[22] COLLINS, J.R. and WIENS, D.P. (1989). Minimax properties of M-, R- and L-estimators of location in Lévy neighbourhoods. Ann. Statist. 17, 1, 327-336.

[23] DASGUPTA, A. (1990). Bounds on asymptotic relative efficiencies of robust estimates of locations for random contaminations. Technical Report #90-64, Department of Statistics, Purdue University.

[24] DEFEO, P. and MYERS, R.H. (1992). A new look at experimental design robustness. Biometrika 79, 2, 375-380.

[25] FEDOROV, V.V. (1972). Theory of Optimal Experiments. Academic Press, New York.

[26] FERNHOLZ, L.T. (1983). Von Mises Calculus for Statistical Functionals. Lecture Notes in Statistics 19, Springer, New York.

[27] GASTWIRTH, J.L. and RUBIN, H. (1975). The behavior of robust estimators on dependent data. Ann. Statist. 3, 5, 1070-1100.

[28] GRAYBILL, F.A. (1976). Theory and Application of the Linear Model. Duxbury Press, North Scituate.

[29] HAMPEL, F.R. (1968). Contributions to the theory of robust estimation. Ph.D. thesis. University of California, Berkeley.

[30] HAMPEL, F.R. (1971). A general qualitative difinition of robustness. Ann. Math. Statist. 42, 1887-1896.

[31] HAMPEL, F.R. (1974). The influence curve and its role in robust estimation. J. Amer. Statist. Assoc. 69, 383-393.

[32] HAMPEL, F.R., RONCHETTI, E.M., ROUSSEEUW, P.J. and STAHEL, W.A. (1986). Robust Statistics : The Approach based on Influence Functions. Wiley, New York.

[33] HERZBERG, A.M. and ANDREWS, D.F. (1976). Some considerations in the optimal design of experiments in non-optimal situations. J.R. Statist. Soc. B38, 284-289.

[34] HORN, R.A. and JOHNSON, C.R. (1985). Matrix Analysis. Cambridge University Press, Cambridge.

[35] HUBER, P.J. (1964). Robust estimation of a location parameter. Ann. Math. Statist. 35, 73-101.

[36] HUBER, P.J. (1965). A robust version of the probability ratio test. Ann. math. Statist. 36, 1753-1758.

[37] HUBER, P.J. (1968). Robust confidence limits. Z. Wahrsch. verw Geb. 10, 269-278.

[38] HUBER, P.J. (1975). Robustness and designs. In Srivastava, J.N. (Ed.) A Survey of Statistical Design and Linear Models, 287-303. North Holland, Amsterdam.

[39] HUBER, P.J. (1981). Robust Statistics. Wiley, New York.

[40] HUBER, P.J. (1983). Minimax aspects of bounded-influence regression. J. Amer. Statist. Assoc. 78, 66-80.

[41] JAECKEL, L.A. (1971). Robust estimates of location: Symmetry and asymmetric contamination. Ann. Math. Statist. 42, 3, 1020-1034.

[42] JONES, E.R. and MITCHELL, T.J. (1978). Design criteria for detecting model inadequacy. Biometrika 65, 541-551.

[43] KARLIN, S. and SHAPLEY, L.S. (1953). Geometry of Moment Spaces, No. 12 of Memoirs of the American Mathematical Society.

[44] KARLIN, S. and STUDDEN, W.J. (1966). Optimal experimental designs. Ann. Math. Statist. 37, 783-815.

[45] KIEFER, J. (1959). Optimal experimental designs. J.R. Statist. Soc. B21, 272-319.

[46] KIEFER, J. (1961). Optimum designs in regression problems, II. Ann. Math. Statist. 32, 298-325.

[47] KIEFER, J. and WOLFOWITZ (1960). The equivalence of two extremum problems. Canad. J. Math. 12, 3, 363-366.

[48] KIEFER, J. and WYNN, H.P. (1984). Optimum and minimax exact treatment designs for one-dimensional autoregressive error processes. Ann. Statist. 12, 2, 414-450.

[49] LEHMANN, E.L. (1991). Theory of Point Estimation. Wadsworth & Brooks/Cole Advanced Books & Software Pacific Grove, California.

[50] LI, K.C. (1984). Robust regression designs with the design space consists of finitely many points. Ann. Statist. 12, 1, 269-282.

[51] LOH, W.Y. (1984). Bounds on AREs for restricted classes of distributions defined via tail-orderings. Ann. Statist. 12, 2, 685-701.

[52] MARCUS, M.B. and SACKS, J. (1977). Robust designs for regression problems. In S.S. Gupta and D.S. Moore (eds.). Statistical Theory and Related Topics II, 245-268. Academic Press, New York.

[53] MARONNA, R.A. and YOHAI, V.J. (1981). Asymptotic behavior of general M-estimates for regression and scale with random carriers. Z. Wahrsch. verw. Gebiete 58, 7-20.

[54] MARTIN, R.D., YOHAI, V.J. and ZAMAR, R.H. (1989). Min-max bias robust regression. Ann. Statist. 17, 4, 1608-1630.

[55] MYERS, R.H., KHURI, A.I. and CARTER, W.H. Jr. (1989). Response surface methodology: 1966-1988. Techometrics 31, 137-157.

[56] NOTZ, W.I. (1989). Optimal designs for regression models with possible bias. J. Statist. Planning and Inference 22, 43-54.

[57] PESOTCHINSKY, L. (1982). Optimal robust designs : Linear regression in R^k. Ann. Statist. 10, 511-525.

[58] PORTNOY, S.L. (1977). Robust estimation in dependent situations. Ann. Statist. 5, 1, 22-43.

[59] PORTNOY, S.L. (1979). Further remarks on robust estimation in dependent situations Ann. Statist. 7, 1, 224-231.

[60] PROHOROV, Y.V. (1956). Convergence of random processes and limit theorems in probability theory. Theor. Prob. Appl. 1, 157-214.

[61] RAO, C.R. and MITRA, S.K. (1971). Generalized Inverse of Matrices and its Applications. Wiley, New York.

[62] ROCKAFELLAR, R.T. (1970). Convex Analysis. Princeton University Press, Princeton.

[63] ROUSSEEUW, P.J. (1981). A new infinitesimal approach to robust estimation. Z. Wahrsch. verw. Gebiete 56, 127-132.

[64] ROUSSEEUW, P.J. and LEROY, A.M. (1987). Robust Regression and Outlier Detection. Wiley New York.

[65] SACKS, J. and YLVISAKER, D. (1966). Designs for regression problems with correlated errors. Ann. Math. Statist. 37, 66-89.

[66] SACKS, J. and YLVISAKER, D. (1968). Designs for regression problems with correlated errors : Many parameters. Ann. Math. Statist. 39, 1, 49-69.

[67] SACKS, J. and YLVISAKER, D. (1978). Linear estimation for approximately linear models. Ann. Statist. 6, 5, 1122-1137.

[68] SACKS, J. and YLVISAKER, D. (1984). Some model robust designs in regression. Ann. Statist. 12, 4, 1324-1348.

[69] SHOHAT, J.A. and TAMARKIN, J.D. (1943). The Problem of Moments. American Mathematical Society Surveys, No. 1, New York.

[70] STIGLER, S.M. (1971). Optimal experimental design for polynomial regression. J. Amer. Statist. Assoc. 66, 311-318.

[71] TUKEY, J.W. (1960). A survey of sampling from contaminated distributions. In Contributions to Probability and Statistics, Ed. I. Olkin, Stanford University Press, Stanford, California, 448-485.

[72] WHITTLE, P. (1973). Some general points in the theory of optimal experimental design. J.R. Statist. Soc. B 35, 123-130.

[73] WIENS, D.P. (1986). Minimax variance M-estimators of location in Kolmogorov neighbourhoods. Ann. Statist. 14, 2, 724-732.

[74] WIENS, D.P. (1990). Robust minimax designs for multiple linear regression. Linear Algebra and Its Applications, Second Special Issue on Algebra and Statistics, 127, 327-340.

[75] WIENS, D.P. (1991). Robust designs for approximately linear regression: M-estimated parameters. Technical Report 91-08 Dept. of Statistics and Applied Probability, Uni. of Alberta, Edmonton, Alberta.

[76] WIENS, D.P. (1992). Minimax designs for approximately linear regression. J. Statist. Planning and Inference 31, 353-371.

[77] WIENS, D.P. WU, K.H.E. (1991). Distributions minimizing Fisher information for scale in Kolmogorov neighbourhoods. Canad. J. Statist. 19, 2, 179-190.

[78] YLVISAKER, N.D. (1964). Lower bounds for minimum covariance matrices in time series regression problems. Ann. math. Statist. 35, 362-368.

[79] YOHAI, V.J. and MARONNA, R.A. (1979). Asymptotic behavior of M-estimators for the linear model. Ann. Statist. 7, 2, 258-268.

[80] ZAMAR, R.H. (1985). Orthogonal regression M-estimators. Technical Report No. 62, Dept. of Statistics, University of Washington, Seattle.