ON SYSTEMS-THEORETIC PROBLEMS WITH SYMMETRIC MULTILINEAR PARAMETER DEPENDENCE

by

Sesan Iwarere

A dissertation submitted in partial fulfillment of
the requirements for the degree of

Doctor of Philosophy

(Electrical Engineering)

at the

UNIVERSITY OF WISCONSIN–MADISON

2015

Date of final oral examination: 10/09/15

The dissertation is approved by the following members of the Final Oral Committee:

B. Ross Barmish, Professor, Electrical and Computer Engineering

John A. Gubner, Professor, Electrical and Computer Engineering

Nigel Boston, Professor, Electrical and Computer Engineering & Mathematics

Eric Bach, Professor, Computer Sciences

Douglass L. Henderson, Professor, Engineering Physics

ACKNOWLEDGMENTS

Many thanks to my advisor, Professor B. Ross Barmish for his mentorship and guidance over many years. I am indebted to his enduring patience with me during this long process.

I would also like to thank Professors Eric Bach, Nigel Boston, John Gubner and Douglass Henderson for serving on my dissertation committee and providing great feedback and corrections.

I would like to thank the Graduate Engineering Research Scholars and its former director Professor Douglass Henderson for financial and academic support. I am also grateful for the National Physical Science Consortium for their financial support over many years.

I am thankful for my fellow graduate school colleague Shirzad Malekpour for his helpful ideas and suggestions and for edits from my friend Mark Zimmerman.

I would also like to thank my parents, Dr. and Mrs. Iwarere, as well as my sister Seyi and brother Seye for their unending encouragement and support.

I am grateful for my fíance Melissa Brandt for her companionship and steadfast support over the last three years. She endured many evenings and weekends of work.

Finally, and most importantly, I would like to thank God for sustaining me through this long journey. There were many ups and downs that He carried me through.

TABLE OF CONTENTS

Abstract

This dissertation addresses problems involving multilinear functions which arise in Systems Theory and satisfy certain symmetry conditions. To this end, in the first part of the dissertation, we consider several problems in optimization and probability involving a so-called *symmetric pair* $(f, \mathfrak{X})$ where $f(x)$ is to be optimized over constraint set $\mathfrak{X}$. In later chapters, we generalize the formulation to address problems involving a so-called *multi-group symmetric pair*. This generalization is motivated by the fact that symmetry may fail to be satisfied, but there may be subsets of variables entering into the problem in a symmetric way. We also consider variations on the themes above. For example, $f(x)$ may be a sum of symmetric functions or enjoy certain affine linearity properties.

For systems described via our symmetric $(f, \mathfrak{X})$ pair framework, we show that the complexity of a problem solution can often be significantly reduced via exploitation of symmetry. By this we mean the following: While the solution for the non-symmetric case requires us to perform function evaluations on all 2^n extreme points of $\mathfrak{X}$, for the symmetric case, we may only need to perform function evaluations on a much "smaller" subset of them. For example, for a symmetric pair $(f, \mathfrak{X})$ with $f(x)$ being multilinear and $\mathfrak{X} \subseteq \mathbb{R}^n$ being a symmetric hypercube, finding the maximum or minimum of $f(x)$ on $\mathfrak{X}$ requires considering only a linear number of extreme points in n versus 2^n for the non-symmetric case. We provide a "Mapping Theorem" along these lines that gives a characterization of the convex hull of the image of a hypercube under a multilinear mapping. The dissertation also includes a number of probability-related applications associated with a symmetric function of a random vector with independent and identically distributed components. When finding the probability mass function of a symmetric function, we see that symmetry can lead to far fewer point masses than the non-symmetric case.

For the more general case of multi-group symmetry with m groups each of size N and $\mathfrak{X}$ being a product of identical hyperrectangles, we show that with $M = 2^N$, the maximization and minimization of $f(x)$ on $\mathfrak{X}$ can be attained by considering

$$N_{ext} = \binom{m+M-1}{m}$$

extremes points rather than 2^n. For many applications where the group size N is small in comparison to the number of groups m, the formula N_{ext} leads to significant savings. Subsequently, we extend these results to address a so-called groupwise affine symmetric case with $\mathfrak{X}$ being a product of m polytopes.

One highlight of this dissertation is the result in Chapter 6. There, we consider a class of resource allocation problems which involves n suppliers, m identical warehouses and a single resource. In this context, the function $f(x)$ to be minimized is a separable, concave sum of symmetric functions on a polytope constraint set $\mathfrak{X}$. Although each instance of this problem can be solved via extreme point function evaluation, by making use of symmetry in conjunction with Schur concavity, again, great savings in computation can result. For cases with a relatively small number of warehouses m in comparison with the number of suppliers n, the number of extreme points one needs to check to find a solution is drastically less than the number of extreme points of $\mathfrak{X}$.

In addition to illustrative applications already mentioned, we also consider finding the minimum and maximum of a symmetric trading gain over a hypercube, problems involving probability mass functions for binomial and quadrinomial lattice models, analysis involving a confinement polygon model for a multi-group symmetric trading gain and distribution of a commodity such as corn among identical warehouses.

Chapter 1

Overview of Dissertation Research

This dissertation concentrates on systems whose mathematical description involves parameters entering into various equations and set descriptions both multilinearly and symmetrically. The systems under consideration are described by a so-called *symmetric pair* $(f, \mathfrak{X})$. Much of the research herein is devoted to demonstrating how results from the literature can be improved if symmetry is brought into play. In particular, when a multilinear function $f(x)$ is maximized or minimized on $\mathfrak{X}$ via extreme point function evaluations, we see that symmetry often leads to a "drastic" reduction in computation versus that required for a generic multilinear function. For applications, we develop results exploiting symmetry in the areas of probability, optimization and finance.

Generalizing in later chapters, we define a new concept called *multi-group symmetry*. In this case, subgroups of variables may enter into $f(x)$ a symmetric way which can be exploited. In this framework, we provide theorems involving the maximization and minimization of multilinear functions on hyperrectangles and polytopes. Finally, one highlight of this dissertation is a result for a class of resource allocation problems involving n suppliers, a single resource and the inventory carrying costs for m identical warehouses. We seek to minimize a separable, concave sum of symmetric functions over a polytope representing the constraints. For any instance of this problem, when finding a solution via extreme point function evaluation, we show that because of symmetry and Schur concavity, for cases where there are a small number of warehouses in comparison to suppliers, the number of extreme points one needs to check can be orders of magnitude less than the number of extreme points of the constraint polytope.

In the next section, we review some standard definitions in mathematics. Then, in the sections to follow, we look at existing literature on multilinear functions addressing problems with similar flavor to those in this dissertation in the areas of optimization, probability and robust stability. We also provide motivating examples where symmetric multilinear problems naturally occur.

1.1 Some Basic Definitions

In this section we provide definitions which serve as the takeoff point for the chapters to follow.

1.1.1 Definition (Affine Linear Function): A function $f\colon \mathbb{R}^n \to \mathbb{R}$, is said to be *affine* if there exist real scalars $\alpha_1, \alpha_2, \ldots, \alpha_n$ and $\beta \in \mathbb{R}$ such that for $x \doteq (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$,

$$f(x) = \sum_{i=1}^{n} \alpha_i x_i + \beta.$$

1.1.2 Definition (Multilinear Function): A function $f\colon \mathbb{R}^n \to \mathbb{R}$ is said to be *multilinear* if it is affine in each of its individual components x_i. In other words, when all variables but one are held fixed, then the resulting single-variable function is affine.

1.1.3 Example (Multilinear Function): An example of a multilinear function is

$$f(x_1, x_2, x_3) = 2x_1 + 4x_1x_2 + 6x_1x_2x_3 + 7.$$

That is, if we hold x_2 and x_3 fixed, then $f(x_1, x_2, x_3)$ is affine linear in x_1. Similarly, if we hold x_1 and x_3 fixed, then $f(x_1, x_2, x_3)$ is affine linear in x_2. Finally, if we hold x_1 and x_2 fixed, then $f(x_1, x_2, x_3)$ is affine linear in x_3.

1.1.4 Definition (Symmetric Function): A function $f\colon \mathbb{R}^n \to \mathbb{R}$ is said to be *symmetric* if it is unchanged by any permutation of its variables; e.g., see Macdonald [1] and Terr [2].

1.1.5 Example (Symmetric Multilinear Function): The function

$$f(x_1, x_2, x_3) = x_1 + x_2 + x_3 + 2x_1x_2x_3$$

is multilinear by inspection and symmetric, i.e.,

$$f(x_1, x_2, x_3) = f(x_1, x_3, x_2) = f(x_2, x_1, x_3) = f(x_2, x_3, x_1) = f(x_3, x_1, x_2) = f(x_3, x_2, x_1).$$

1.1.6 Remarks on Symmetric Functions: In the field of abstract algebra, when symmetric functions are addressed, it is usually in reference to symmetric polynomials such as the function

$$f(x) = x_1^2 x_2^2 \ + \ x_1^2 x_3^2 \ + \ x_2^2 x_3^2 \ + \ 10 x_1 x_2 x_3 \ + \ 5.$$

In this dissertation, although our primary focus is on symmetric multilinear functions, a special kind of symmetric polynomial, some of our results will also hold for rather general symmetric functions such as

$$f(x) = \cos x_1 x_2 x_3 \ + \ \sin^2(x_1 x_2 + x_1 x_3 + x_2 x_3).$$

1.1.7 Definition (Symmetric Set): A set $\mathfrak{X} \subseteq \mathbb{R}^n$ is said to be *symmetric* if, for any given vector $x = (x_1, x_2, \ldots, x_n) \in \mathfrak{X}$, any permutation of its components also belongs to $\mathfrak{X}$.

1.1.8 Example (Symmetric Set): In $\mathbb{R}^2$, the quarter-circle described by

$$\mathfrak{X} = \left\{ x \in \mathbb{R}^2 : x_1^2 + x_2^2 \leq 5^2, x_1 \geq 0, x_2 \geq 0 \right\}$$

is symmetric; see Figure 1.1. Notice that given any $(x_1, x_2) \in \mathfrak{X}$ we also have $(x_2, x_1) \in \mathfrak{X}$.

1.1.9 Definition (Symmetric Pair): Given a symmetric function $f\colon \mathbb{R}^n \to \mathbb{R}$ and a symmetric set $\mathfrak{X} \subseteq \mathbb{R}^n$, we call $(f, \mathfrak{X})$ a *symmetric pair*.

1.1.10 The Big Picture: In this dissertation, for systems fitting into our $(f, \mathfrak{X})$ pair framework, we show that the complexity of the problem solution can often be significantly reduced via exploitation of symmetry. That is, while the non-symmetric case requires us to perform function evaluations on all 2^n extreme points of $\mathfrak{X}$ to find a solution, for the symmetric case we may only need to perform function evaluations on a much "smaller" subset of them. For example, when $(f, \mathfrak{X})$ is a symmetric pair with $f(x)$ multilinear and $\mathfrak{X} \subseteq \mathbb{R}^n$ a hypercube, to find the minimum or maximum of f on $\mathfrak{X}$, we only have to consider a linear number of extreme points in n

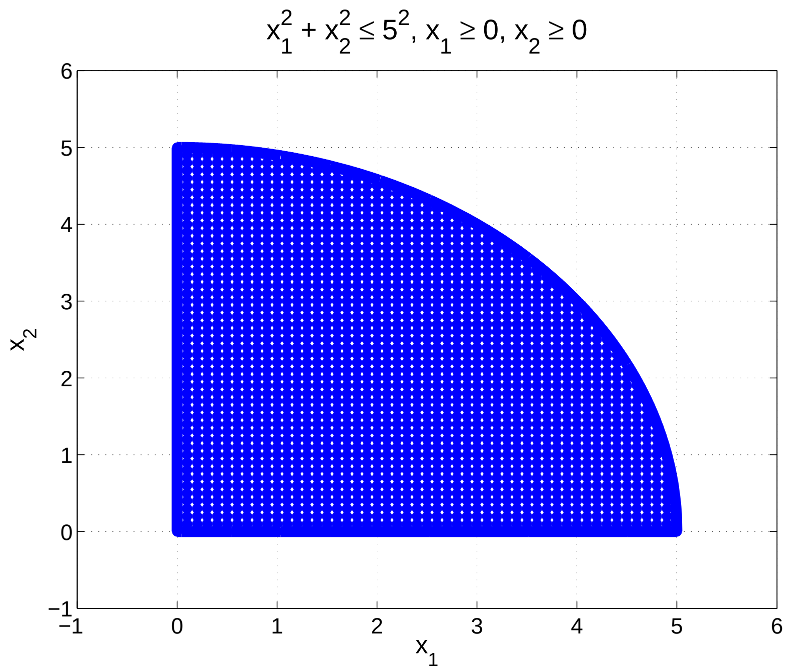

Figure 1.1 Quarter-Circle Symmetric Constraint Set $\mathfrak{X}$ in $\mathbb{R}^2$

rather than 2^n in the classical solution when symmetry is not in play. We also demonstrate how to exploit symmetry to simplify the computation of probabilities. When finding the probability mass function of a symmetric function, we see that symmetry can lead to far fewer point masses than the non-symmetric case.

1.2 Motivating Examples

In this section, we present four motivating examples from Systems Theory. One of them involves financial markets, two cover discrete-time bilinear systems and one involves the transfer function of a circuit.

1.2.1 Financial Markets: For the first motivating example, we see where symmetric multilinear uncertainty structures arise in financial markets. To illustrate, we consider the trading gain equation that results from a linear feedback strategy with feedback gain K; see [54]. When a trader begins with initial investment I_0, the trading gain or loss is given by

$$f(x) = \frac{I_0}{K}\left[\prod_{i=0}^{n-1}(1 + Kx(i)) - 1\right]$$

where $x \doteq (x(0), x(1), \ldots, x(n-1))$ is the sequence of daily returns and n represents the number of trading days. Notice that $f(x)$ is a symmetric multilinear function. Furthermore, when the $x(i)$ above are independent and identically constrained, the resulting pair $(f, \mathfrak{X})$ is symmetric. In Section 4.9, we find the probability mass function of $f(x)$ when the $x(i)$ above are independent and identically distributed Bernoulli-like random variables. In Section 4.11, we find the probability mass function for an extended version of the trading gain above.

1.2.2 Discrete Bilinear System: We first consider the discrete-time bilinear system shown in Figure 1.2. In the figure, Z^{-1} denotes the delay operator. In this system, we have two interacting subsystems. The first is bilinear in $x(k)$ and $y_1(k)$, and the second is bilinear in $x(k)$ and $y_2(k)$. The two interacting subsystems share the same bounded uncertain time-varying perturbation $x(k)$. The output $z(k)$ of the system is the sum of the outputs of the interacting subsystems. Mathematically, we have

$$\begin{aligned} y_1(k+1) &= (a_1 + \lambda_1 x(k))\, y_1(k); \\ y_2(k+1) &= (a_2 + \lambda_2 x(k))\, y_2(k); \\ z(k) &= y_1(k) + y_2(k) \end{aligned}$$

where a_1, a_2, λ_1, and λ_2 are constants. Now, beginning with $y_1(0) = y_2(0) = 1$, for any time instant $k \geq 0$, it is readily verified that the output $z(k)$ is a symmetric multilinear function of the perturbations $x(k)$. For example, if $k = 3$, then a straightforward calculation leads to

$$\begin{aligned} z(3) = (a_1^3 + a_2^3) &+ (\lambda_1 a_1^2 + \lambda_2 a_2^2)(x(0) + x(1) + x(2)) \\ &+ (\lambda_1^2 a_1 + \lambda_2^2 a_2)(x(0)x(1) + x(0)x(2) + x(1)x(2)) \\ &+ (\lambda_1^3 + \lambda_2^3)(x(0)x(1)x(2)) \end{aligned}$$

which is a symmetric multilinear function of $x(0)$, $x(1)$ and $x(2)$.

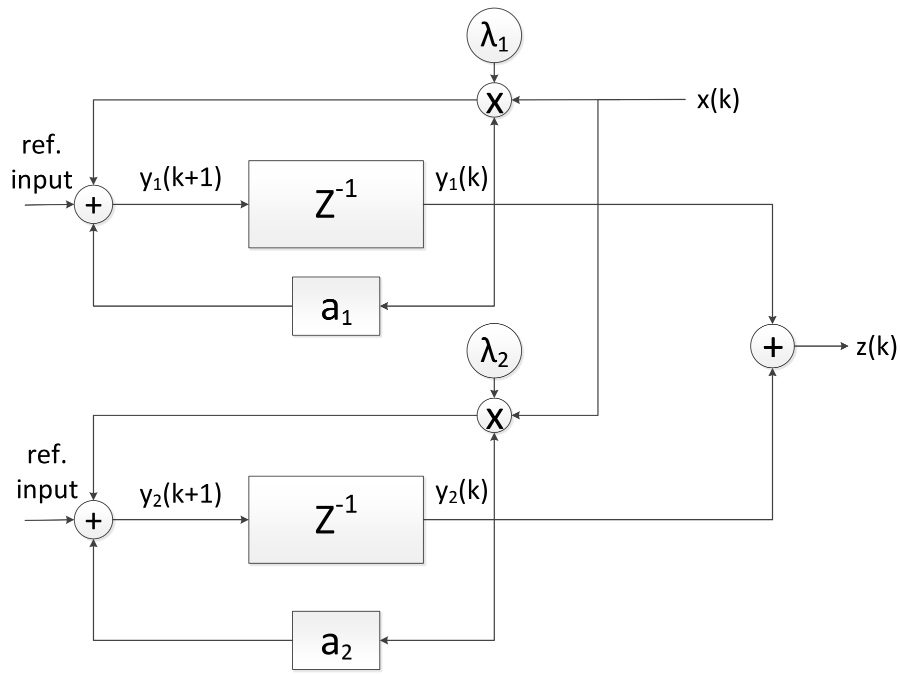

Figure 1.2 Two-dimensional Bilinear Control System with One Time-varying Perturbation

Now suppose the two bilinear subsystems in Figure 1.2 are not interacting and each have their own bounded uncertain time-varying perturbation. Let $x_1(k)$ be the perturbation associated with the first system and let $x_2(k)$ be the perturbation associated with the second system. Then the system is described mathematically by

$$\begin{aligned} y_1(k+1) &= (a_1 + \lambda_1 x_1(k))\, y_1(k); \\ y_2(k+1) &= (a_2 + \lambda_2 x_2(k))\, y_2(k); \\ z(k) &= y_1(k) + y_2(k), \end{aligned}$$

where a_1, a_2, λ_1 and λ_2 are constants. Now, beginning with $y_1(0) = y_2(0) = 1$, for any time instant $k \geq 0$, the output $z(k)$ can be readily shown to be a so-called multi-group symmetric multilinear function involving $x_1(k)$ and $x_2(k)$. As formally defined in Chapter 5, a multi-group symmetric function is symmetric with respect to the permutations of subgroups of its variables, rather than all its variables as in the symmetric case. Now suppose $k = 3$, then

$$\begin{aligned} z(3) \;=\; & a_1^3 \;+\; \lambda_1 a_1^2(x_1(0)+x_1(1)+x_1(2)) \;+\; \lambda_1^2 a_1(x_1(0)x_1(1)+x_1(0)x_1(2)+x_1(1)x_1(2)) \\ & +\; \lambda_1^3(x_1(0)x_1(1)x_1(2)) \;+\; a_2^3 \;+\; \lambda_2 a_2^2(x_2(0)+x_2(1)+x_2(2)) \\ & +\; \lambda_2^2 a_2(x_2(0)x_2(1)+x_2(0)x_2(2)+x_2(1)x_2(2)) \;+\; \lambda_2^3(x_2(0)x_2(1)x_2(2)) \end{aligned}$$

is a multi-group symmetric multilinear function that is symmetric with respect to the three groups given by $\{x_1(0), x_2(0)\}$, $\{x_1(1), x_2(1)\}$ and $\{x_1(2), x_2(2)\}$. In Chapter 5, we provide a formal definition of a multi-group symmetric function. We also define other multi-group symmetric concepts such as a multi-group symmetric set. For optimization theorems involving a multi-group symmetric pair $(f, \mathfrak{X})$, we describe how an optimum is found using a subset of the extreme points of $\mathfrak{X}$. We also provide a formula for the number of such points which are needed.

1.2.3 Systems with Uncertain Poles: Another class of problems giving rise to symmetric pairs $(f, \mathfrak{X})$ is obtained for the feedback control system, shown in Figure 1.3, that is derived from a resistor-capacitor network of lag compensators. Also, in this system, the α_i are given fixed zeros, and $x = (x_1, x_2, x_3)$ are the uncertain poles that are restricted to the bounding set $\mathfrak{X}$ described by $.01 \leq x_i \leq 1$ for $i = 1, 2, 3$. Via a straightforward calculation, the denominator of the closed loop transfer function is readily verified to be

$$p(s,x) = s^3 + \left(\frac{1}{2}(x_1+x_2+x_3) + \tilde{\alpha}_2\right)s^2 + \left(\frac{1}{2}(x_1x_2+x_1x_3+x_2x_3) + \tilde{\alpha}_1\right)s + \frac{1}{2}x_1x_2x_3 + \tilde{\alpha}_0$$

where

$$\tilde{\alpha}_0 \doteq \frac{1}{2}\alpha_1\alpha_2\alpha_3; \quad \tilde{\alpha}_1 \doteq \frac{1}{2}(\alpha_1\alpha_2+\alpha_1\alpha_3+\alpha_2\alpha_3); \quad \tilde{\alpha}_2 \doteq \frac{1}{2}(\alpha_1+\alpha_2+\alpha_3).$$

The key points to note are that all the coefficients of this polynomial are symmetric multilinear functions of x and $(f, \mathfrak{X})$ defines a symmetric pair.

1.3 Motivating Background Literature

Over the last few decades, there has been a great deal of work done in the study of dynamic systems with a so-called "multilinear uncertainty structure." In reviewing the literature on this topic, we

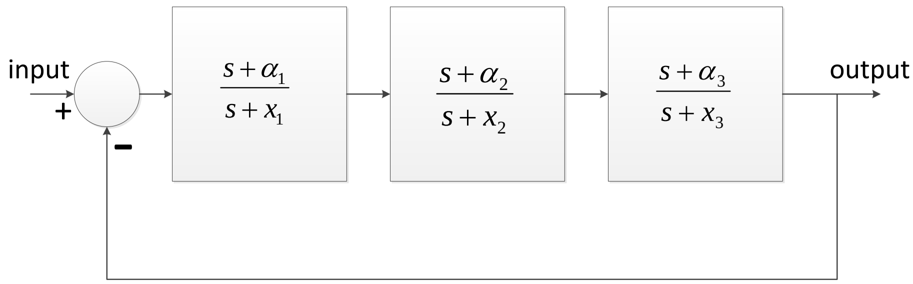

Figure 1.3 System Leading to Closed Loop Symmetric Multilinear Uncertainty Structure

focus on the areas most closely related to the work in this dissertation: mapping of multilinear uncertainty, probability and mathematical programming. We address situations where symmetric multilinearity can arise in each of these areas.

1.4 Literature on Multilinear Mapping of Uncertainty

The Mapping Theorem, given in [33], is a powerful result that can be used in the analysis of multilinear functions. It is stated below.

1.4.1 Mapping Theorem: *Suppose $\mathfrak{X} \subset \mathbb{R}^n$ is a hyperrectangle with extreme points $\{x^i\}$ and $f : \mathbb{R}^n \to \mathbb{R}^k$ is multilinear. Then the range $f(\mathfrak{X}) = \{f(x) : x \in \mathfrak{X}\}$ has convex hull*

$$\text{conv } f(\mathfrak{X}) = \text{conv } \left\{ f\left(x^i\right) \right\}.$$

1.4.2 Remarks on Mapping Theorem: According to the Mapping Theorem, for a multilinear function defined on a hyperrectangle, one can compute the convex hull of the range of the function by finding the convex hull of the images of the extreme points of $\mathfrak{X}$. In [34], with the aid of a two-dimensional version of the Mapping Theorem, the robust stability of a family of characteristic polynomials whose coefficients $a_i(x)$ are multilinear functions, the convex hull above can be determined graphically. The uncertain polynomials in [34] come from several real-world systems that include a circuit from a motor inertia system, a circuit from an electrical circuit with a feedback network, a model of a DC motor with resonant load and the model of a supersonic transport plane. In [35], several theorems for robust stability are given for a family of polynomials with coefficients $a_i(x)$, which are multilinear. The proofs of several of these theorems require the use of the Mapping Theorem.

1.4.3 Example (Use of Mapping Theorem): To provide an example to illustrate the use of the Mapping Theorem, we consider the family of uncertain polynomials described by

$$\begin{aligned} p(s,x) &= s^3 - (x_1+x_2+x_3)\, s^2 + (x_1x_2+x_1x_3+x_2x_3-36)\, s \\ &\quad + 12\,(x_1+x_2+x_3) - x_1x_2x_3 - 114 \end{aligned}$$

with $\mathfrak{X}$ described by $0 \le x_i \le 1$ for $i = 1,2,3$. In control theory, with frequency variable $\omega \ge 0$, we typically substitute $s = j\omega$ above and construct $p\,(j\omega, \mathfrak{X})$, the so-called value set. Using the Mapping Theorem we obtain an approximation for $p\,(j\omega, \mathfrak{X})$, namely its convex hull. For $p(s,x)$ above, $\mathfrak{X}$ has 8 extreme points

$$\begin{aligned} &x^1 = (0,0,0)\,; \quad x^2 = (0,0,1)\,; \quad x^3 = (0,1,0)\,; \quad x^4 = (0,1,1)\,; \\ &x^5 = (1,0,0)\,; \quad x^6 = (1,0,1)\,; \quad x^7 = (1,1,0)\,; \quad x^8 = (1,1,1)\,. \end{aligned}$$

Hence, to obtain $\operatorname{conv}\{p\,(j\omega, \mathfrak{X})\}$, we require 8 points in the complex plane. Namely we need $p\,(j\omega, x^i)$ for $i = 1, 2, \ldots, 8$. For example when $i = 4$, straightforward calculation yields

$$p(j\omega, x^4) = 2\omega^2 - 90 - j\left(\omega^3 + 35\right).$$

To further illustrate, we take $\omega = 1$. Then $p(j1, x^4) = -88 - j36$ and the remaining 7 points in the complex plane are readily found and the resulting convex hull is obtained; see Figure 1.4, where $\operatorname{conv}\{p\,(j1, \mathfrak{X})\}$ is compared with the true set $p\,(j1, \mathfrak{X})$.

1.4.4 Mapping Theorem Exploiting Symmetry: It is no accident that $\operatorname{conv}\{p\,(j1, \mathfrak{X})\}$ above is obtained by using only four extreme points instead of the eight predicted by the Mapping Theorem. This is a consequence of symmetry. With this motivation in mind, in Section 3.6, we give a

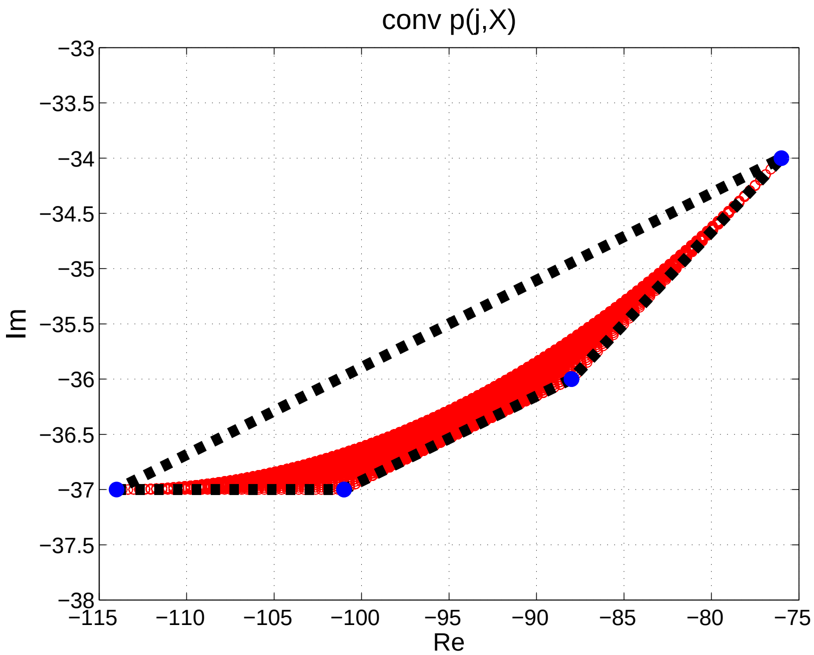

Figure 1.4 Convex Hull of $p(j\omega, \mathfrak{X})$ and True Set at $\omega = 1$

"symmetric multilinear version" of the Mapping Theorem. In this case, to compute the convex hull, as a result of the symmetry of the pair $(f, \mathfrak{X})$, one only needs to evaluate the function $f(x)$ at $n + 1$ extreme points instead of the original 2^n. Since many systems problems involve large n, this reduction of complexity is critical for computational tractability. When $\mathfrak{X}$ satisfies the non-triviality requirement that $p(s, x^0)$ is stable for at least one $x^0 \in \mathfrak{X}$, the well known necessary and sufficient condition for robust stability is that $0 \notin p\,(j\omega, \mathfrak{X})$ for all $0 \leq \omega < \infty$; e.g., see [3].

1.5 Probability Problems Involving Multilinear Dependence

In the systems area, multilinear uncertainty structures not only arise in a robustness context but also in a probabilistic context. Using the notation from Section 1.4.3, we now take X to be a random vector with components X_i. In stability analysis, we consider a random polynomial whose coefficients $a_i\,(X)$ are real multilinear functions of

$$p\,(s, X)\; \doteq\; \sum_{i=0}^{m} a_i\,(X)\, s^i$$

In this setting, $\mathfrak{X}$ corresponds to the support of X and is a hypercube with center

$$x^0 \doteq \left(x_1^0, x_2^0, \ldots, x_n^0\right)$$

described by $|x_i - x_i^0| \leq r_i$ where $r_i \geq 0$ for $i = 1, 2, \ldots, m$. With this formulation, in [36] the notion of "distributional robustness" is considered and applications include a mass-spring damper system with fifteen uncertain parameters. In a similar context, the work in [37] considers the H_∞ norm of a proper stable transfer function matrix $H\left(s, X\right)$. In [38], for the case of a random resistive network, the worst-case probability distribution for its equivalent conductance is derived. This conductance is written as the quotient of multilinear functions, and uncertain conductances are viewed as independent Gaussian distributed random variables.

1.6 Mathematical Programming with Multilinear Functions

In Systems Theory, multilinear functions arise in many mathematical programming applications. For example, in [41], multilinear functions are used in bilinear and quadratic programs. Multilinear functions are also present in Reformulation-Linearization techniques in discrete and continuous optimization; e.g., see [42]. In [43], several schemes for bounding multilinear functions are studied. They include the arithmetic method, the logarithmic method, and the exponent transformation method; see [44]. Multilinear functions are also used as objective functions, as in the feedback analog neural networks studied in [45]. Finally, we mention literature involving the transfer function of a large class of lumped linear time-invariant electrical networks, can be written in the form

$$H(s,x) = \frac{N(s,x)}{D(s,x)}$$

where

$$N(s,x) = \sum_{i=0}^{m_1} n_i(x) s^i; \quad D(s,x) = \sum_{i=0}^{m_1} d_i(x) s^i$$

are polynomials with real coefficients $n_i(x)$ and $d_i(x)$ being multilinear functions. The formulation also includes a hyperrectangle set $\mathfrak{X}$ for the uncertain parameter vector x. In this setting, references [46] and [47] involve nonlinear programming to compute the worst-case for various classes of transfer functions.

1.7 Overview of Chapters to Follow

Now that we have provided the "big picture," we conclude this chapter with a brief overview of the technical results to follow. Indeed, in Chapter 2, we provide a first principles derivation of the general form of a symmetric multilinear function that we believe is useful for the systems community. That is, in contrast to more abstract approaches to the derivation of a more general result for polynomials that can be found from the field of algebra, for example, see [1], [64], and [65], the proofs become much easier for multilinear functions.

In Chapter 3, we study the maximization and minimization of a symmetric multilinear $f(x)$ over a symmetric hypercube $\mathfrak{X} \subseteq \mathbb{R}^n$. For this symmetric pair $(f, \mathfrak{X})$, we prove that we only need $n+1$ extreme point function evaluations to find the minimum or maximum. This compares with the classical solution which requires 2^n such evaluations for a generic multilinear function. Included in Chapter 3 is an application involving a trading gain $f(x)$ arising in a financial market scenario. We find the minimum and maximum of this function with a symmetric hypercube constraint. Finally, we provide a "new" version of the Mapping Theorem that applies for a symmetric pair $(f, \mathfrak{X})$.

In Chapter 4, we derive the probability mass function (PMF) associated with a symmetric function $S(X)$ with X being a random vector with independent and identically distributed random components. As an application, in Section 4.3, we find the CDF and PMF associated with the "Simultaneous Long-Short" analysis in [53]. In this case, the components of X are independent and identically distributed random variables that take on only two values. We also find the probability mass function of a trading gain $S(X)$ arising from a linear feedback strategy where the underlying price trajectory is determined by a binomial lattice model. In addition, we find the PMF of a function $S(X)$ that is a sum of two symmetric functions and is associated with a so-called quadrinomial lattice model for the price trajectories of two correlated stocks. We also discuss how to extend our PMF result for three or more correlated stocks.

In Chapter 5, we generalize the notion of a *symmetric pair* $(f, \mathfrak{X})$ to a so-called *multi-group symmetric pair* with $\mathfrak{X} \subseteq \mathbb{R}^n$ being the product of m identical hyperrectangles. We address the maximization and minimization of $f(x)$ over $\mathfrak{X}$ and show that in many cases, particularly when m is large in comparison to the size of each group N, exploitation of multi-group symmetry requires much fewer extreme point function evaluations than that required for a generic multilinear function. That is, with $M = 2^N$, one need only consider

$$N_{ext} = \binom{m+M-1}{m}$$

extreme points. Also in Chapter 5, we consider a so-called *groupwise affine symmetric pair* where $\mathfrak{X}$ is the product of m identical polytopes each with M extreme points. We first show that as a result of $f(x)$ being groupwise affine, a maximum and minimum occur at the extreme points of the polytope $\mathfrak{X}$. Then, by taking advantage of multi-group symmetry, we arrive at the same formula of N_{ext} for the number of extreme points to be considered. This compares to M^m extreme points which are required without multi-group symmetry. As an application, we consider the case where $f(x)$ is a groupwise affine trading gain that results from trading two correlated stocks using a linear feedback strategy. We estimate the best and worst case values of $f(x)$ after m days of trading.

In Chapter 6, we present a highlight of this dissertation. Namely, we study a class of inventory carrying cost problems which fall under the umbrella of resource allocation. The minimization problems we consider involve n suppliers m warehouses and a single resource. The i-th supplier can send resource amount $x_{i,j} \geq 0$ to a prescribed subset of warehouses. We seek to minimize $f(x)$, the total "inventory carrying costs," which is seen to be a separable, concave sum of symmetric functions. The constraint set $\mathfrak{X}$ is taken to be a simplex associated with admissible resource allocations. Then, by making use of Schur concavity, consistent with other results in this thesis, it is shown that the number of extreme point function evaluations one needs to perform to find a solution can be dramatically less than the number of extreme points of $\mathfrak{X}$. To achieve these results, we introduce a new technical concept called the *Ordered Loading Property* (OLP). We prove that the number of such extreme points having this property is $m!$ regardless of the value of n, while for fixed m, the number of extreme points of $\mathfrak{X}$ can be many orders of magnitudes larger than the number of OLP points.

In Chapter 7, we provide concluding remarks and directions for further research. In particular, we consider an alternate version of the resource allocation problem in Chapter 6 including the geographical location of the suppliers. We also consider how we can generalize our multi-group symmetric theory to cases where only a subset of variables enter into a multi-group function symmetrically. Finally, we consider some robust stability problems where symmetry can be exploited.

Chapter 2

On Representation of a Symmetric Multilinear Function

In this chapter, we provide a characterization of symmetric multilinear functions. We show that a generic symmetric multilinear function $f(x)$, for $x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$, is a linear combination of elementary symmetric functions. These elementary symmetric functions come from the field of abstract algebra.

2.1 Elementary Symmetric Functions

2.1.1 Definition: For $x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$, the *elementary symmetric functions* $e_0(x)$, $e_1(x)$, $e_2(x)$, …, $e_{n-1}(x)$, $e_n(x)$, see [2], are given by

$$e_0(x) \equiv 1; \qquad e_4(x) \doteq \sum_{1\le i<j<k<l\le n} x_i x_j x_k x_l;$$

$$e_1(x) \doteq \sum_{i=1}^{n} x_i; \qquad e_5(x) \doteq \sum_{1\le i<j<k<l<m\le n} x_i x_j x_k x_l x_m;$$

$$e_2(x) \doteq \sum_{1\le i<j\le n} x_i x_j; \qquad \vdots$$

$$e_3(x) \doteq \sum_{1\le i<j<k\le n} x_i x_j x_k; \qquad e_n(x) \doteq \prod_{i=1}^{n} x_i.$$

Note that the i-th such function has $\binom{n}{i}$ terms and $e_i(x) = 0$ if $i > n$.

2.1.2 Properties: The elementary symmetric functions have several nice properties. Affine linear combinations of elementary symmetric functions are also symmetric functions. Finally, elementary symmetric functions satisfy a recursive relationship. That is, if $x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$ and

$\bar{x} = (x_1, x_2, \ldots, x_{n-1}) \in \mathbb{R}^{n-1}$, then

$$e_i(x) = e_i(\bar{x}) \; + \; e_{i-1}(\bar{x})\, x_n$$

for $i = 1, 2, \ldots, n$. Note that $e_n(\bar{x}) = 0$ and the number of terms in the above relation is in agreement with the identity

$$\binom{n}{i} = \binom{n-1}{i} \; + \; \binom{n-1}{i-1}.$$

2.2 Main Result: Representation of a Multilinear Symmetric Function

Per the discussion in Chapter 1, in this section, we see that the desired characterization of symmetric multilinear functions is easy to obtain under the strengthened hypothesis that $f(x)$ is multilinear. No abstract algebra concepts are needed, making the ideas more accessible to the engineering community.

2.2.1 Theorem: *A multilinear function $f\colon \mathbb{R}^n \to \mathbb{R}$ is symmetric if and only if it can be expressed in the form*

$$f(x) = \sum_{i=0}^{n} a_i e_i(x),$$

where the $e_i(x)$ are the elementary symmetric functions and $a_0, a_1, \ldots, a_{n-1}, a_n$ are real constants.

2.2.2 Preliminaries: To facilitate the proof, we define the mappings

$$\sigma_i\colon \; \{1, 2, \ldots, n-1, n\} \to \{1, 2, \ldots, n-1, n\}$$

as follows: For $i = 1$, we take σ_1 to be the identity mapping. Then, for $i = 2, 3, \ldots, n-1$

$$\sigma_i(j) \doteq \begin{cases} i & \text{if} \quad j = 1; \\ j-1 & \text{if} \quad 1 < j \leq i; \\ j & \text{if} \quad i < j \leq n. \end{cases}$$

Finally for $i = n$, let

$$\sigma_n(j) \doteq \begin{cases} n & \text{if} \quad j = 1; \\ j-1 & \text{if} \quad 1 < j \leq n. \end{cases}$$

Now using $\sigma_1, \sigma_2, \ldots, \sigma_n$, we define the *distinguished permutations* of $x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$ as mappings with components

$$x_j^{\sigma_i} \doteq x_{\sigma_i(j)},$$

for $i = 1, 2, \ldots, n$ and $j = 1, 2, \ldots, n$. For example, if $n = 5$, the distinguished permutations are

$$\begin{aligned} x^{\sigma_1} &= (x_1, x_2, x_3, x_4, x_5)\,; \\ x^{\sigma_2} &= (x_2, x_1, x_3, x_4, x_5)\,; \\ x^{\sigma_3} &= (x_3, x_1, x_2, x_4, x_5)\,; \\ x^{\sigma_4} &= (x_4, x_1, x_2, x_3, x_5)\,; \\ x^{\sigma_5} &= (x_5, x_1, x_2, x_3, x_4)\,. \end{aligned}$$

Now, for a symmetric function $f(x)$, we observe that

$$f(x^{\sigma_j}) = f(x)$$

for $j = 1, 2, \ldots, n$.

2.2.3 Proof of Theorem 2.2.1: Sufficiency is trivial since every multilinear function of the form

$$f(x) = \sum_{i=0}^{n} a_i e_i(x).$$

is symmetric by inspection. In order to establish necessity, suppose $f(x)$ is an arbitrary symmetric multilinear function. Then, it can be expressed as

$$f(x) = a_0 \; + \; \sum_{i=1}^{n} b_i x_i \; + \sum_{\substack{1 \le j < k \le n \\ 1 \le i \le \binom{n}{2}}} c_i x_j x_k \; + \sum_{\substack{1 \le j < k < l \le n \\ 1 \le i \le \binom{n}{3}}} d_i x_j x_k x_l \; + \; \cdots \; + \; a_n e_n(x)$$

where $a_0, b_i, c_i, d_i, \ldots, a_n$ are constants. Since $f(x)$ is symmetric, we claim that the coefficients must satisfy the equations

$$\begin{aligned} b_1 = b_2 = \;\; \cdots \;\; &= b_{\binom{n}{1}}; \\ c_1 = c_2 = \;\; \cdots \;\; &= c_{\binom{n}{2}}; \\ d_1 = d_2 = \;\; \cdots \;\; &= d_{\binom{n}{3}}. \\ &\;\vdots \end{aligned}$$

To prove the equalities above, we invoke symmetry. That is, using the distinguished permutations defined above, the necessary conditions

$$\begin{aligned} f(x^{\sigma_1}) - f(x) &= 0; \\ f(x^{\sigma_2}) - f(x) &= 0; \\ &\vdots \\ f(x^{\sigma_n}) - f(x) &= 0 \end{aligned}$$

must be satisfied. That is

$$f(x^{\sigma_1}) - f(x^{\sigma_j}) = 0$$

for $j = 1, 2, \ldots, n$. Now, to show all the b_i coefficients are the same, all the c_i coefficients are the same, all the d_i coefficients are the same and so on, in each case we use n distinguished vectors with components that consists of ones and zeros and substitute them into equation above. There are n distinguished vectors that show all the b_i coefficients are the same, n distinguished vectors that show that all the c_i coefficients are the same, n distinguished vectors that show all the d_i coefficients are the same, etc. For notational simplicity, we demonstrate how this procedure works for $n = 3$ noting that the same method of proof works for arbitrary n. Now, for $n = 3$, the arbitrary symmetric multilinear function is given by

$$f(x) = a_0 \ + \ b_1x_1 \ + \ b_2x_2 \ + \ b_3x_3 + \ c_1x_1x_2 \ + \ c_2x_1x_3 \ + \ c_3x_2x_3 \ + \ a_3x_1x_2x_3.$$

Now, the three necessary conditions $f(x^{\sigma_1}) - f(x^{\sigma_j}) = 0$ reduce to the two equations

$$\begin{aligned} (b_1 - b_2)x_1 \ + \ (b_2 - b_1)x_2 \ + \ (c_2 - c_3)x_1x_3 \ + \ (c_3 - c_2)x_2x_3 &= 0; \\ (b_1 - b_2)x_1 \ + \ (b_2 - b_3)x_2 \ + \ (b_3 - b_1)x_3 \ + \ (c_1 - c_3)x_1x_2 \ + \ (c_2 - c_1)x_1x_3 & \\ + \ (c_3 - c_2)x_2x_3 &= 0. \end{aligned}$$

since the first equation is redundant. To show all the b_i coefficients are the same, we use the 3 distinguished vectors

$$(1, 0, 0), \ (0, 1, 0), \ (0, 0, 1).$$

Substituting $(1, 0, 0)$ into our system of equations forces

$$b_1 = b_2.$$

Now our original set of equations reduce to

$$\begin{aligned} (c_2 - c_3)x_1x_3 \ + \ (c_3 - c_2)x_2x_3 &= 0; \\ (b_2 - b_3)x_2 \ + \ (b_3 - b_1)x_3 \ + \ (c_1 - c_3)x_1x_2 \ + \ (c_2 - c_1)x_1x_3 \ + \ (c_3 - c_2)x_2x_3 &= 0. \end{aligned}$$

Substituting $(0, 1, 0)$ into previous equations we get

$$b_2 = b_3.$$

Since $b_1 = b_2$ and $b_2 = b_3$, we can conclude that $b_1 = b_2 = b_3$. Now, to consider the c_i, note that our set of equations now reduces to

$$\begin{aligned} (c_2 - c_3)x_1x_3 \ + \ (c_3 - c_2)x_2x_3 &= 0; \\ (c_1 - c_3)x_1x_2 \ + \ (c_2 - c_1)x_1x_3 \ + \ (c_3 - c_2)x_2x_3 &= 0. \end{aligned}$$

Now to show all the c_i coefficients are the same, we use the 3 distinguished vectors

$$(1, 1, 0), \ (1, 0, 1), \ (0, 1, 1).$$

Substituting $(1, 1, 0)$ leads to

$$c_1 = c_3.$$

With $c_1 = c_3$, we can further reduce our set of equations to

$$\begin{aligned} (c_2 - c_3)x_1x_3 \ + \ (c_3 - c_2)x_2x_3 &= 0; \\ (c_2 - c_1)x_1x_3 \ + \ (c_3 - c_2)x_2x_3 &= 0. \end{aligned}$$

Now substituting the second distinguished vector $(1, 0, 1)$ leads to

$$\begin{aligned} c_2 &= c_3; \\ c_2 &= c_1, \end{aligned}$$

above. Therefore $c_1 = c_2 = c_3$. Since we now have all b_i equal and all c_i equal, $f(x)$ reduces to

$$f(x) \;=\; a_0 \,+\, b_1(x_1 + x_2 + x_3) \,+\, c_1(x_1x_2 + x_1x_3 + x_2x_3) \,+\, a_3x_1x_2x_3.$$

Now via the change of notation $a_1 = b_1$ and $a_2 = c_1$, we obtain the form

$$f(x) \;=\; a_0 \,+\, a_1e_1(x) \,+\, a_2e_2(x) \,+\, a_3e_3(x).$$

Now, for general n, using a similar procedure with distinguished vectors each having $n - i$ entries $x_i = 1$ and the remaining $x_i = 0$, the symmetric multilinear function $f(x)$ can be simplified to

$$f(x) \;=\; \sum_{i=0}^{n} a_ie_i(x).$$

Chapter 3

Symmetric Multilinear Optimization Problems

This chapter considers several optimization problems involving the maximization and minimization of a symmetric multilinear function $f(x)$ on a symmetric hypercube $\mathfrak{X}$ with center $x^0 \in \mathbb{R}^n$ and radius r. For notational convenience, using the l^∞ norm, we write

$$\mathfrak{X} \doteq \left\{ x \in \mathbb{R}^n : \|x - x^0\|_\infty \leq r \right\}.$$

With $\mathfrak{X}$ symmetric, it is easy to see that all components of the center point x^0 must have the same value; i.e., the center is of the form $x^0 = (c, c, \ldots, c)$ for some $c \in \mathbb{R}$. It is well known that the maximum and minimum of a multilinear function on a hypercube is attained among its 2^n extreme points; e.g., see [3]. For the symmetric pair $(f, \mathfrak{X})$ above, it will be shown that we only need to consider at most $n + 1$ extreme points as opposed to 2^n extreme points for a general multilinear problem. In other words, in the symmetric case, one need only consider a linear number of extreme points in contrast to an exponential number in the general case. Since real-world problems often involve large n, this reduction in computational complexity is dramatic.

We apply this result to the example involving financial markets described in Section 1.2.1. To this end, we find the minimum and maximum of a symmetric multilinear trading gain function that arises from using the "Simultaneous Long-Short" strategy in [53]. Finally, in this chapter, we establish a specialized Mapping Theorem for the symmetric pair $(f, \mathfrak{X})$; recall Section 1.4.1 and see [33]. In this case, we see that the convex hull of the range of $f(x)$ is obtained by using only $n + 1$ extreme points as opposed to 2^n extreme points for a general multilinear function.

3.1 Transformation to Unit Hypercube

In the sequel, it is often convenient to work on the unit hypercube. This simplifies both notation and proofs. For the unit hypercube, in terms of the notation above, we have center $x^0 = 0$ and

$$\mathfrak{X} \doteq \{x \in \mathbb{R}^n \colon \|x\|_\infty \leq 1\}.$$

3.1.1 Relationship Between General Hypercube and Unit Hypercube: Let $(f, \mathfrak{X})$ be a symmetric pair with $f(x)$ multilinear with $\mathfrak{X}$ being a hypercube having center x^0 and radius $r \geq 0$. Given $f(x)$ expressed in terms of elementary symmetric functions as

$$f(x) = \sum_{i=0}^{n} a_i e_i(x),$$

define the multilinear function $\tilde{f} \colon \mathbb{R}^n \to \mathbb{R}$ by

$$\tilde{f}(x) \doteq \sum_{i=0}^{n} \tilde{a}_i e_i(x)$$

where

$$\tilde{a}_i \doteq r^i \sum_{k=0}^{n-i} \binom{n-i}{k} c^k a_{k+i}$$

for $i = 0, 1, \ldots, n$. Then, it is easy to see that

$$\begin{aligned} \min_{x \in \mathfrak{X}} f(x) &= \min_{\|x\|_\infty \leq 1} \tilde{f}(x); \\ \max_{x \in \mathfrak{X}} f(x) &= \max_{\|x\|_\infty \leq 1} \tilde{f}(x). \end{aligned}$$

Further if $\tilde{x}^*$ is an optimal element for $\tilde{f}(x)$, it can be readily shown that

$$x^* = r\tilde{x}^* \ + \ x^0$$

is an optimal element for $f(x)$.

It is interesting to note that the coefficient vectors a and $\tilde{a}$ representing $f(x)$ and $\tilde{f}(x)$ above can be represented in a compact matrix form. Indeed, a lengthy but straightforward calculation leads to

$$\tilde{a} = (AP \circ B)\, a$$

where $\circ$ is the Hadamard product for matrices, P is an upper triangular matrix whose non-zero entries are taken from Pascal's triangle,

$$A \doteq \operatorname{diag}\left(1, r^1, r^2, \ldots, r^n\right),$$

and B is an upper triangular matrix

$$B \doteq \begin{bmatrix} 1 & c & c^2 & \cdots & c^{n-1} & c^n \\ 0 & 1 & c & \cdots & c^{n-2} & c^{n-1} \\ 0 & 0 & 1 & \cdots & c^{n-3} & c^{n-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & c \\ 0 & 0 & 0 & \cdots & 0 & 1 \end{bmatrix}.$$

3.1.2 Example (Generic Hypercube to Unit Hypercube): With $n = 2$, $r = 2$ and $c = 2$, the symmetric hypercube $\mathfrak{X}$ is described by

$$\mathfrak{X} = \left\{x \in \mathbb{R}^2 \ : \ \|x - x^0\|_\infty \leq 2\right\}$$

with center $x^0 = (2, 2)$. Now given the symmetric multilinear function

$$f(x) = 5 \ + \ 5(x_1 + x_2) \ - \ 2x_1x_2,$$

a straightforward calculation leads to

$$\tilde{f}(x) = -8x_1x_2 \ + \ 2(x_1 + x_2) \ + \ 17$$

and the maximum of $\tilde{f}$ occurs at $\tilde{x}^* = (-1, 1)$. In addition, the transformation above leads to $x^* = (0, 4)$ and optimal values $f(x^*) = \tilde{f}(\tilde{x}^*) = 25$.

3.2 Extreme Points of $\mathfrak{X}$

The set of extreme points of a symmetric hypercube $\mathfrak{X}$, denoted by $\mathfrak{X}_{ext}$, are straightforward to describe. That is, $\mathfrak{X}_{ext}$ consists of the 2^n points $x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$ with

$$x_i \in \{c - r, c + r\}$$

for $i = 1, 2, \ldots, n$. As mentioned earlier in this chapter, it is well known that the maximum and minimum of a multilinear function $f(x)$ defined on a hypercube is found among its extreme points. In other words,

$$\begin{aligned} \min_{x \in \mathfrak{X}} f(x) &= \min_{x \in \mathfrak{X}_{ext}} f(x); \\ \max_{x \in \mathfrak{X}} f(x) &= \max_{x \in \mathfrak{X}_{ext}} f(x). \end{aligned}$$

In the theorem to follow, we see that symmetry of the pair $(f, \mathfrak{X})$ makes it possible to solve the maximization and minimization problems above using only $n+1$ distinguished extreme points.

3.3 Maximization and Minimization of a Symmetric Multilinear Function

In this section, we see how symmetry is exploited to simplify the solution of the problems above.

3.3.1 Preliminaries: To facilitate the proof of the theorem to follow, we define the sets

$$\mathfrak{X}_k \doteq \{x : k \text{ components } x_i = c + r \text{ and } n - k \text{ components } x_i = c - r\},$$

for $k = 0, 1, \ldots, n$ and notice that

$$\mathfrak{X}_{ext} = \bigcup_{k=0}^{n} \mathfrak{X}_k$$

partitions $\mathfrak{X}_{ext}$ into $n+1$ disjoint sets each of size $\|\mathfrak{X}_k\| = \binom{n}{k}$.

We also define the *distinguished extreme point set* $\mathfrak{X}_\mathcal{K}$ by selecting one element from each of the sets $\mathfrak{X}_k$ defined above. For $\mathfrak{X}_k$, with $1 \le k \le n-1$, the selection we make has components

$$x_i = \begin{cases} c+r & \text{if } 1 \le i \le k; \\ c-r & \text{if } k+1 \le i \le n. \end{cases}$$

The vectors $(c-r, c-r, \ldots, c-r)$ and $(c+r, c+r, \ldots, c+r)$ are used for $k = 0$ and $k = n$ respectively. Note that the cardinality of $\|\mathfrak{X}_\mathcal{K}\| = n+1$.

To illustrate construction of the sets $\mathfrak{X}_k$ and $\mathfrak{X}_\mathcal{K}$ defined above, suppose that $n = 3$, $c = 55$, and $r = 45$. Then the associated hypercube $\mathfrak{X}$, has extreme point set

$$\mathfrak{X}_{ext} = \bigcup_{k=0}^{3} \mathfrak{X}_k$$

and moreover,

$$\begin{aligned} \mathfrak{X}_0 &= \{(10,10,10)\}; \\ \mathfrak{X}_1 &= \{(100,10,10),\ (10,100,10),\ (10,10,100)\}; \\ \mathfrak{X}_2 &= \{(100,100,10),\ (100,10,100),\ (10,100,100)\}; \\ \mathfrak{X}_3 &= \{(100,100,100)\} \end{aligned}$$

and the distinguished extreme point set is

$$\mathfrak{X}_\mathcal{K} = \{(10,10,10),\ (100,10,10),\ (100,100,10),\ (100,100,100)\},$$

with cardinality $\|\mathfrak{X}_\mathcal{K}\| = 4$ while $\|\mathfrak{X}_{ext}\| = 8$.

3.3.2 Theorem: *Let $(f, \mathfrak{X})$ be a symmetric pair with $f(x)$ multilinear and $\mathfrak{X}$ being a symmetric hypercube with distinguished extreme point set $\mathfrak{X}_\mathcal{K}$. Then*

$$\begin{aligned} \min_{x \in \mathfrak{X}} f(x) &= \min_{x \in \mathfrak{X}_\mathcal{K}} f(x); \\ \max_{x \in \mathfrak{X}} f(x) &= \max_{x \in \mathfrak{X}_\mathcal{K}} f(x). \end{aligned}$$

3.3.3 Proof: We only provide a proof of the maximization case since the proof for the minimization case is nearly identical. Per the discussion in Section 3.1, without loss of generality, we take $\mathfrak{X}$ to be the unit hypercube. From the "ordinary" Mapping Theorem, we know that an optimal element x^* can be found in $\mathfrak{X}_{ext}$; see Section 3.2. Now in view of symmetry, $f(x)$ is constant on $\mathfrak{X}_k$ for each $k = 0, 1, \ldots, n$. In other words all the $\binom{n}{k}$ extreme points associated with the set $\mathfrak{X}_k$ all evaluate to the same function value. Hence, to obtain the maximum of $f(x)$, we need only consider one extreme point from each set $\mathfrak{X}_k$. Since the set $\mathfrak{X}_{\mathcal{K}}$ is comprised of $n+1$ such points, we conclude that

$$\max_{x\in\mathfrak{X}} f(x) \;=\; \max_{x\in\mathfrak{X}_{\mathcal{K}}} f(x).$$

3.4 Illustrative Example

To illustrate application of the theorem above, we take $n = 4$ and consider finding the maximum of the symmetric multilinear function

$$\begin{aligned} f(x) \;=\; & 7 \;-\; 3(x_1+x_2+x_3+x_4) \;+\; 5(x_1x_2+x_1x_3+x_1x_4+x_2x_3+x_2x_4+x_3x_4) \\ & +\; 8(x_1x_2x_3+x_1x_2x_4+x_1x_3x_4+x_2x_3x_4) \;-\; 7x_1x_2x_3x_4 \end{aligned}$$

on the symmetric hypercube $\mathfrak{X}$ with center $x^0 = (5,5,5,5)$ and radius $r = 2$. According to the theorem, we first generate the distinguished extreme point set

$$\mathfrak{X}_{\mathcal{K}} = \{(3,3,3,3)\,,\ (7,3,3,3)\,,\ (7,7,3,3)\,,\ (7,7,7,3)\,,\ (7,7,7,7)\}\,.$$

Now evaluating $f(x)$ at each of these extreme points, we find a maximizer occurs at $x^* = (7,7,3,3)$ and associated optimal value $f(x^*) = 930$. By taking advantage of symmetry, instead of having to consider $2^4 = 16$ extreme points we only used 5 extreme points.

3.5 Application: Stock Trading Example

In this section, we analyze the symmetric multilinear stock trading gain resulting from a so-called "Simultaneous Long Short" linear feedback controller introduced in [53]. The symmetric function of interest is given by

$$f(x) = \left[\prod_{i=0}^{n-1}(1 + x(i)) \;+\; \prod_{i=0}^{n-1}(1 - x(i)) \;-\; 2\right]$$

where $x = (x(0), x(1), \ldots, x(n-1))$ represents the daily returns and n represents the number of trading days. We consider the symmetric hypercube $\mathfrak{X}$ with radius $0 < r < 1$ and our goal is to find the minimum and maximum value of $f(x)$. Now, because $(f, \mathfrak{X})$ is a symmetric pair, according to Theorem 3.3.2, both the maximum and minimum of $f(x)$ will occur at one of the $n + 1$ distinguished extreme points in the set

$$\mathfrak{X}_{\mathcal{K}} = \{(-r, -r, -r, \ldots, -r)\,,\; (r, -r, -r, \ldots, -r)\,, \ldots, (r, r, \ldots, r, -r)\,,\; (r, r, r, \ldots, r)\}\,.$$

If $f(x)$ were not symmetric, we would have to evaluate $f(x)$ at each of the 2^n extreme points of $\mathfrak{X}_{ext}$ to find its minimum or maximum. Now, as a result of the structure of $f(x)$ for this problem, we can actually find a formula for the optimum without recourse to multiple extreme point evaluations. By substituting r for k components of x and $-r$ for $n - k$ components, it suffices to study the function

$$\tilde{f}(k) \doteq \left[(1 + r)^k (1 - r)^{n-k} \;+\; (1 - r)^k (1 + r)^{n-k} \;-\; 2\right]$$

where $0 \leq k \leq n$. In the theorem below, we see that the maximum and minimum value of $\tilde{f}(x)$ can be found in closed form.

3.5.1 Theorem: *The minimum and maximum values of the trading gain $f(x)$ are given by*

$$\min_{x \in \mathfrak{X}} f(x) \;=\; \begin{cases} 2\left[(1 - r^2)^{\frac{n}{2}} \;-\; 1\right] & \text{for } n \text{ even;} \\[2em] 2\left[(1 - r^2)^{\lfloor \frac{n}{2} \rfloor} \;-\; 1\right] & \text{for } n \text{ odd;} \end{cases}$$

$$\max_{x \in \mathfrak{X}} f(x) \;=\; \left[(1 + r)^n \;+\; (1 - r)^n \;-\; 2\right] \;\; \text{for all } n.$$

For n even, every point $x \in \mathfrak{X}_k$ with $k = \frac{n}{2}$ is a minimizer and one such point is given by

$$x_i^* = \begin{cases} r & \text{for } 1 \le i \le \frac{n}{2}; \\ -r & \text{for } \frac{n}{2}+1 \le i \le n. \end{cases}$$

For n odd, the minimum occurs at every point $x \in \mathfrak{X}_{k_1} \cup \mathfrak{X}_{k_2}$ with $k_1 = \lfloor \frac{n}{2} \rfloor$ and $k_2 = \lceil \frac{n}{2} \rceil$. Two such points are given by

$$x_i^* = \begin{cases} r & \text{for } 1 \le i \le \lfloor \frac{n}{2} \rfloor; \\ -r & \text{for } \lfloor \frac{n}{2} \rfloor + 1 \le i \le n \end{cases}$$

and

$$x_i^* = \begin{cases} r & \text{for } 1 \le i \le \lceil \frac{n}{2} \rceil; \\ -r & \text{for } \lceil \frac{n}{2} \rceil + 1 \le i \le n. \end{cases}$$

The maximum occurs at $x^* = (-r, -r, -r, \ldots, -r)$ and $(r, r, r, \ldots, r)$.

3.5.2 Proof: Since $(f, \mathfrak{X})$ is a symmetric pair with $f(x)$ multilinear and $\mathfrak{X}$ being a hypercube, we know from Theorem 3.3.2, that the minimum and maximum of $f(x)$ occurs at one of the distinguished extreme points of $\mathfrak{X}_{\mathcal{K}}$ given by

$$\mathfrak{X}_{\mathcal{K}} = \{(-r, -r, -r, \ldots, -r)\,,\ (r, -r, -r, \ldots, -r)\,, \ldots, (r, r, \ldots, r, -r)\,,\ (r, r, r, \ldots, r)\}\,.$$

Now, recalling the function $\tilde{f}(k)$ above, notice that the values $\tilde{f}(0), \tilde{f}(1), \ldots, \tilde{f}(n)$ correspond to $f(x)$ evaluated at each of the elements of the distinguished extreme point set $\mathfrak{X}_{\mathcal{K}}$. Noting that $\tilde{f}(k)$ is convex when k is treated as a continuous variable, if the minimum occurs at an integer, the minimum for continuous k will be the same as the minimum for discrete k. If the minimum for continuous $\tilde{f}(k)$ does not occur at an integer, as a result of the convexity of the function, for discrete $\tilde{f}(k)$, the minimum will occur at one or both of the two closest integers bracketing this minimum. By calculus, it is straightforward to verify that for continuous k,

$$\left.\frac{d\tilde{f}}{dk}\right|_{k=\frac{n}{2}} = 0;$$

$$\left.\frac{d^2\tilde{f}}{dk^2}\right|_{k=\frac{n}{2}} \ge 0.$$

Therefore the minimum of $\tilde{f}(k)$ for continuous k occurs at $k^* = \frac{n}{2}$. When n is even this corresponds to discrete minimize $k^* = \frac{n}{2}$ as well, and

$$\tilde{f}\left(\frac{n}{2}\right) = 2\left[\left(1 - r^2\right)^{\frac{n}{2}} \; - \; 1\right].$$

Note that $k^* = \frac{n}{2}$ corresponds to distinguished extreme point

$$x_i^* = \begin{cases} r & \text{for } 1 \leq i \leq \frac{n}{2}; \\ -r & \text{for } \frac{n}{2} + 1 \leq i \leq n. \end{cases}$$

For $\frac{n}{2}$, if k^* is not an integer, the minimum for $\tilde{f}(k)$ for discrete k will occur at either $k^* = \lceil \frac{n}{2} \rceil$ or $k^* = \lfloor \frac{n}{2} \rfloor$. However, since

$$\tilde{f}\left(\lceil \frac{n}{2} \rceil\right) = \tilde{f}\left(\lfloor \frac{n}{2} \rfloor\right) = 2\left[\left(1 - r^2\right)^{\lfloor \frac{n}{2} \rfloor} \; - \; 1\right],$$

both of these candidates are optimal. In summary, the optimal elements $k^* = \lceil \frac{n}{2} \rceil$ and $k^* = \lfloor \frac{n}{2} \rfloor$ correspond to distinguished extreme point

$$x_i^* = \begin{cases} r & \text{for } 1 \leq i \leq \lceil \frac{n}{2} \rceil; \\ -r & \text{for } \lceil \frac{n}{2} \rceil + 1 \leq i \leq n \end{cases}$$

and

$$x_i^* = \begin{cases} r & \text{for } 1 \leq i \leq \lfloor \frac{n}{2} \rfloor; \\ -r & \text{for } \lfloor \frac{n}{2} \rfloor + 1 \leq i \leq n. \end{cases}$$

As a result of the convexity of $\tilde{f}(k)$, for continuous k, its maximum occurs at either $k^* = 0$ or $k^* = n$, or both. However since $\tilde{f}(0) = \tilde{f}(n)$, it occurs at both. Hence we obtain maximum value

$$\tilde{f}(0) = \tilde{f}(n) = \left[(1 + r)^n \; + \; (1 - r)^n \; - \; 2\right]$$

and corresponding optimal elements $x^* = (-r, -r, -r, \ldots, -r)$ and $x^* = (r, r, r, \ldots, r)$.

3.5.3 Examples: When $n = 4$,

$$\begin{aligned} \min_{x \in \mathfrak{X}} f(x) &= 2\left[(1 - r)^2 \; - \; 1\right]; \;\; x^* = (r, r, -r, -r); \\ \max_{x \in \mathfrak{X}} f(x) &= \left[(1 + r)^4 \; + \; (1 - r)^4 \; - \; 2\right]; \;\; x^* = \{(-r, -r, -r, -r), \; (r, r, r, r)\}. \end{aligned}$$

When $n = 5$,

$$\min_{x\in\mathfrak{X}} f(x) \;=\; 2\left[(1-r)^2 \;-\; 1\right]; \;\; x^* = \{(r,r,-r,-r,-r)\,,\; (r,r,r,-r,-r)\}\,;$$
$$\max_{x\in\mathfrak{X}} f(x) \;=\; \left[(1+r)^5 \;+\; (1-r)^5 \;-\; 2\right]; \;\; x^* = \{(-r,-r,-r,-r,-r)\,,\; (r,r,r,r,r)\}\,.$$

3.6 Mapping Theorem for Symmetric Multilinear Functions

In this section, we provide a specialized version of the Mapping Theorem which applies to a symmetric pair $(f, \mathfrak{X})$; see Section 1.4.1 where this topic is first discussed.

3.6.1 Theorem (Symmetric Mapping): *Suppose $f\colon \mathbb{R}^n \to \mathbb{R}^k$ has components $f_i(x)$ that are symmetric multilinear functions and $\mathfrak{X} \subseteq \mathbb{R}^n$ is an symmetric hypercube with distinguished extreme point set $\mathfrak{X}_{\mathcal{K}}$. Then*

$$\text{conv}\;\, f\,(\mathfrak{X}) = \text{conv}\;\, f\,(\mathfrak{X}_{\mathcal{K}})\,.$$

3.6.2 Preliminaries: In preparation for the proof of the theorem, we first review the well-known definition of a support function; e.g., see [67]. For $\mathfrak{X} \subseteq \mathbb{R}^n$ the *support function* $h\colon \mathbb{R}^n \to \mathbb{R}$ on $\mathfrak{X}$ is defined by

$$h(y) \doteq \sup_{x\in\mathfrak{X}}\; y^T x$$

where x and y are column vectors and y^T is the transpose of y. Now for a polytope,

$$X = \text{conv}\,(\mathfrak{X}_{ext})\,,$$

given any $y \in \mathbb{R}^n$, it is also well known and easy to show that we can obtain the support function of X by using $\mathfrak{X}_{ext}$, the extreme points of X. That is,

$$h(y) = \max_{x\in\mathfrak{X}_{ext}} y^T x.$$

Finally, another important well-known property which we need is that if two closed sets $\mathfrak{X}_1 \subseteq \mathbb{R}^n$ and $\mathfrak{X}_2 \subseteq \mathbb{R}^n$ have the same support function $h(y)$, then

$$\text{conv}\;\, \mathfrak{X}_1 = \text{conv}\;\, \mathfrak{X}_2.$$

3.6.3 Proof of Theorem 3.6.1: We prove the theorem by making use of the preliminaries above. Indeed, we first define the finite point set

$$\mathcal{F} \doteq f\left(\mathfrak{X}_{\mathcal{K}}\right).$$

It suffices to show that $f\left(\mathfrak{X}\right)$ and $\mathcal{F}$ have the same support function. Indeed, let $y \in \mathbb{R}^k$ be arbitrarily fixed and let $h_f(y)$ and $h_{\mathcal{F}}(y)$ denote the support functions on $f\left(\mathfrak{X}\right)$ and $\mathcal{F}$ respectively. Beginning with

$$h_f(y) = \sup_{x \in \mathfrak{X}} y^T f\left(x\right),$$

and observing that $\tilde{f}(x) = y^T f(x)$ is a symmetric function for each fixed $y \in \mathbb{R}^n$, in view of Theorem 3.3.2 and the compactness of $\mathfrak{X}$, we obtain

$$h_f(y) = \max_{x \in \mathfrak{X}} \tilde{f}(x) = \max_{x \in \mathfrak{X}_{\mathcal{K}}} y^T f(x) = h_{\mathcal{F}}(y),$$

where we have also used the fact that $f\left(\mathfrak{X}_{\mathcal{K}}\right)$ and $\operatorname{conv} f\left(\mathfrak{X}_{\mathcal{K}}\right)$ have the same support function.

3.7 Further Research

In this chapter, we have shown how to efficiently find the maximum and minimum of a symmetric multilinear function $f(x)$ over a symmetric hypercube $\mathfrak{X}$. Exploiting symmetry, one only need to consider $n+1$ extreme points as opposed to the 2^n extreme points required for the generic non-symmetric case. We now discuss directions for future research along this line.

3.7.1 Min and Max of Multi-Group Symmetric Multilinear Functions: In Chapter 5, we will introduce the notion of an multi-group symmetric function $f(x)$. This will allow us to extend the results in this chapter to cover optimization problems under a less stringent requirement. In our multi-group symmetric theory, the variables $x_1, x_2, \ldots, x_n$ of x, get partitioned into m groups, each of equal size $N \doteq n/m$ where n is divisible by m. For the case where $f(x)$ is multilinear and $\mathfrak{X}$ is the Cartesian product of m identical hyperrectangles, each with $M = 2^N$ extreme points, we prove that to find an optimum for both the minimization and maximization cases, $f(x)$ only needs to be evaluated at

$$N_{ext} \doteq \binom{m+M-1}{m}$$

extreme points rather than 2^n for a generic multilinear function. For the limiting case when $m = n$, each group has size $N = 1$, $\mathfrak{X}$ becomes a hypercube, multi-group symmetry reduces to symmetry and $N_{ext} = n + 1$.

3.7.2 Minimization of Symmetric Convex Function: In this section, we give a conjecture for a minimizer of a a symmetric convex function $f(x)$ on a symmetric convex set, not necessarily a hypercube, $\mathfrak{X}\subseteq\mathbb{R}^n$. If the conjecture is proved to be true, then an n-dimensional optimization problem becomes a one-dimensional optimization problem. Suppose $f(x)$ is such that a minimizer exists. Then, we conjecture that $f(x)$ has a minimizer $x^* \in \mathfrak{X}$ of the form

$$x^* = (\gamma, \gamma, \ldots, \gamma)$$

for some $\gamma \in \mathbb{R}$. To illustrate the ramifications if this conjecture proves to be true, suppose $\mathfrak{X} = \mathbb{R}^3$ and

$$f(x) = 5\left(x_1^2 + x_2^2 + x_3^2\right) \; + \; 5\left(x_1x_2 + x_1x_3 + x_2x_3\right) \; + \; 2\left(x_1 + x_2 + x_3\right).$$

Now, $f(x)$ is symmetric by inspection and it is also a strictly convex function since its Hessian

$$H(x) = \begin{bmatrix} 10 & 5 & 5 \\ 5 & 10 & 5 \\ 5 & 5 & 10 \end{bmatrix}$$

is positive definite. To find the minimum of $f(x)$, we set $x_1 = x_2 = x_3 = \gamma$. Then a straightforward calculation yields

$$f(\gamma, \gamma, \gamma) = 30\gamma^2 \; + \; 6\gamma,$$

whose global minimum is trivially obtained with $\gamma = -0.1$. In this case the corresponding optimal value is readily found to be

$$f^* = f\left(-0.1, -0.1, -0.1\right) = -0.3.$$

Chapter 4

On Symmetric Functions of Discrete Random Variables

In this chapter, we consider the probability distribution functions, both the cumulative distribution function (CDF) and probability mass function (PMF), for a symmetric function $\mathcal{S}(X)$ with underlying random vector X whose components X_i are independent and identically distributed. As a result of the symmetry of $\mathcal{S}(X)$, the number of terms required to compute the PMF and CDF is seen to be "drastically" reduced in comparison to the non-symmetric case. Note that in this chapter we represent symmetric functions with the notation $\mathcal{S}(X)$ rather than $f(x)$ which we reserve for probability density functions. To this end, we denote the CDF of $\mathcal{S}(X)$ by $F_{\mathcal{S}}(x)$ and the PMF by $f_{\mathcal{S}}(x)$. As an application, we also revisit the example involving financial markets that was considered in Chapter 3. That is, we find the PMF and CDF of the symmetric multilinear function that arises from stock market trading using the so-called "Simultaneous Long-Short" method. In this context, the random vector X has i.i.d. components X_i that take on only two values.

Subsequently, later in the chapter, we consider another financial market application. We first look at the case where $\mathcal{S}(X)$ is the trading gain that occurs from using linear feedback and a binomial lattice, see [59], for propagating stock prices. Then, we consider a second case where $\mathcal{S}(X)$ is the trading gain that occurs when a so-called quadrinomial lattice, see [58], is used to propagate a pair of correlated stock prices. For both cases, we find the PMF of $\mathcal{S}(X)$. By way of future research, we consider the form of the PMF of $\mathcal{S}(X)$ with $m = 3$ stocks. In this case, $\mathcal{S}(X)$ is the sum of three symmetric multilinear functions. These special sums show promise for more general PMF results involving generic sums of symmetric functions.

4.1 Description of Random Variables

In this chapter, the random vector $X = (X_1, X_2, \ldots, X_n)$ has components X_i which are assumed to be i.i.d. discrete random variables having common mass concentration points $x_1 < x_2 < \cdots < x_l$, for $i = 1, 2, \ldots, n$. That is, each X_i takes on l distinct values and its probability mass function is described by

$$p_j \doteq P(X_i = x_j); \quad j = 1, 2, \ldots, l$$

with all $p_j \geq 0$ and $p_1 + p_2 + \cdots + p_l = 1$.

4.2 Derivation of CDF and PMF of $\mathcal{S}(X)$

4.2.1 Preliminaries: To state Theorem 4.2.3 to follow, we require some notation. With X defined as above and $\mathcal{S}(X)$ a symmetric multilinear function, we first define the index set

$$\mathcal{K} \doteq \left\{ K = (k_1, k_2, \ldots, k_l) \in \mathbb{Z}_+^l : \sum_{j=1}^{l} k_j = n \right\}.$$

Now, for each $K = (k_1, k_2, \ldots, k_l) \in \mathcal{K}$, we take $\mathfrak{X}_K$ to be the event

$$\mathfrak{X}_K \doteq \{x \in \mathbb{R}^n : \text{exactly } k_j \text{ components of } X \text{ are } x_j \text{ for } j = 1, 2, \ldots, l\}.$$

The *cardinality of* $\mathfrak{X}_K$ is

$$\|\mathfrak{X}_K\| = \frac{n!}{k_1! k_2! \cdots k_l!}$$

and we note that $\mathfrak{X}$ is partitioned as

$$\mathfrak{X} \doteq \bigcup_{K \in \mathcal{K}} \mathfrak{X}_K.$$

Now, for each $K \in \mathcal{K}$ we define the *distinguished vector* $x^K \in \mathfrak{X}_K$ by

$$x_i^K \doteq \begin{cases} x_1 & \text{for } 1 \le i \le k_1; \\ x_2 & \text{for } k_1 + 1 \le i \le k_1 + k_2; \\ x_3 & \text{for } k_1 + k_2 + 1 \le i \le k_1 + k_2 + k_3; \\ \vdots & \\ x_l & \text{for } \sum_{j=1}^{l-1} k_j + 1 \le i \le n \end{cases}$$

for $i = 1, 2, \ldots, n$, with appropriate deletions above for $k_i = 0$.

4.2.2 Example Illustrating Notation: To provide an example of the construction above, suppose $l = 3$ and $n = 4$. Then

$$\mathcal{K} \;=\; \left\{K = (k_1, k_2, k_3) \,\in\, \mathbb{Z}_+^3 : k_1 + k_2 + k_3 = 4\right\}$$

and

$$\mathcal{K} \;=\; \begin{Bmatrix} (0,0,4)\,, & (0,1,3)\,, & (0,2,2) \\ (0,3,1)\,, & (0,4,0)\,, & (1,0,3) \\ (1,1,2)\,, & (1,2,1)\,, & (1,3,0) \\ (2,0,2)\,, & (2,1,1)\,, & (2,2,0) \\ (3,0,1)\,, & (3,1,0)\,, & (4,0,0) \end{Bmatrix}.$$

Now to illustrate construction of x^K, suppose $K = (0, 2, 2)$, then

$$\mathfrak{X}_K = \begin{Bmatrix} (x_2, x_2, x_3, x_3)\,, & (x_2, x_3, x_2, x_3)\,, & (x_2, x_3, x_3, x_2) \\ (x_3, x_2, x_2, x_3)\,, & (x_3, x_2, x_3, x_2)\,, & (x_3, x_3, x_2, x_2) \end{Bmatrix},$$

$\|\mathfrak{X}_K\| = 6$ and

$$x^K = (x_2, x_2, x_3, x_3)\,.$$

4.2.3 Theorem: *Given the symmetric function $\mathcal{S} : \mathbb{R}^n \to \mathbb{R}$ and the discrete i.i.d. random variables $X_1, X_2, \ldots, X_n$, the CDF $F_{\mathcal{S}}(x)$ and PMF $f_{\mathcal{S}}(x)$ are given respectively by*

$$F_{\mathcal{S}}(x) = \sum_{K \in \mathcal{K}} p_1^{k_1} p_2^{k_2} \cdots p_l^{k_l} \left(\frac{n!}{k_1! k_2! \cdots k_l!} \right) u\left(x - \mathcal{S}\left(x^K\right)\right);$$

$$f_{\mathcal{S}}(x) = \sum_{K \in \mathcal{K}} p_1^{k_1} p_2^{k_2} \cdots p_l^{k_l} \left(\frac{n!}{k_1! k_2! \cdots k_l!} \right) \delta\left(x - \mathcal{S}\left(x^K\right)\right)$$

where $u(\cdot)$ denotes the unit step function and $\delta(\cdot)$ denotes the Dirac delta function.

4.2.4 Proof of Theorem 4.2.3: Recalling the partition of X above in terms of $\mathfrak{X}_K$, we invoke the *law of total probability* to obtain

$$F_{\mathcal{S}}(x) = P\left(\mathcal{S}(X) \le x\right) = \sum_{K \in \mathcal{K}} P\left(\mathcal{S}(X) \le x \,|X \in \mathfrak{X}_K\right) P\left(X \in \mathfrak{X}_K\right).$$

Now, since $\mathcal{S}(X)$ is symmetric and constant on $\mathfrak{X}_K$, taking representative $x^K \in \mathfrak{X}_K$, we have

$$\begin{aligned} P\left(\mathcal{S}(X) \le x \,|X \in \mathfrak{X}_K\right) &= \begin{cases} 0 & \text{if } \mathcal{S}\left(x^K\right) > x; \\ 1 & \text{if } \mathcal{S}\left(x^K\right) \le x; \end{cases} \\ &= u\left(x - \mathcal{S}\left(x^K\right)\right). \end{aligned}$$

Now substituting in $F_{\mathcal{S}}(x)$ above yields

$$F_{\mathcal{S}}(x) = \sum_{K \in \mathcal{K}} u\left(x - \mathcal{S}\left(x^K\right)\right) P\left(X \in \mathfrak{X}_K\right).$$

Next using the fact that the X_i are i.i.d., we have

$$\begin{aligned} P\left(X \in \mathfrak{X}_K\right) &= p_1^{k_1} p_2^{k_2} \cdots p_l^{k_l} \|\mathfrak{X}_K\| \\ &= p_1^{k_1} p_2^{k_2} \cdots p_l^{k_l} \left(\frac{n!}{k_1! k_2! \cdots k_l!} \right). \end{aligned}$$

This leads to the final expression for the CDF which is

$$F_{\mathcal{S}}(x) = \sum_{K \in \mathcal{K}} p_1^{k_1} p_2^{k_2} \cdots p_l^{k_l} \left(\frac{n!}{k_1! k_2! \cdots k_l!} \right) u\left(x - \mathcal{S}\left(x^K\right)\right)$$

and the PMF formula follows trivially via differentiation.

4.2.5 Remarks on Complexity of PMF Calculation: As a consequence of symmetry, the number of terms in $F_{\mathcal{S}}(x)$ can be dramatically less than the number of terms obtained without symmetry. Without symmetry, one would need to work with the formula for $\|\mathfrak{X}_K\|$, given in Section 4.2.1, to determine $P\left(\mathcal{S}\left(X\right) \leq x \,|X \in \mathfrak{X}_K\right)$. Also as a consequence of symmetry, the number of terms required to compute both the CDF and PMF of $\mathcal{S}(X)$ in Theorem 4.2.3 is given by

$$N(l,n) = \binom{n+l-1}{n}$$

rather than l^n terms without symmetry; see [66]. Note that $\|\mathcal{K}\| = N(l,n)$.

4.3 Example: Trading Gain with a Two-Mass Distribution

In this section, we find the PMF associated with the symmetric multilinear trading gain equation which arises from a "Simultaneous Long-Short" stock trading described in Section 3.5. Indeed, we begin with the symmetric function for the trading gain given by

$$\mathcal{S}(X) = \left[\prod_{i=0}^{n-1}(1+X(i)) \;+\; \prod_{i=0}^{n-1}(1-X(i)) \;-\; 2\right].$$

We now consider the case when the components of

$$X = \left(X(0), X(1), \ldots, X(n-1)\right),$$

which represent the daily returns take on only two values like the Bernoulli distribution. Our analysis corresponds to $l=2$ in Theorem 4.2.3.

To illustrate further, suppose $0 < r < 1$, $x_1 = -r$ and $x_2 = r$. Then we begin with

$$p_1 = P\left(X(i) = -r\right);\; p_2 = P\left(X(i) = r\right)$$

for $i = 0, 1, 2, \ldots, n-1$ with $p_1 \geq 0$ and $p_2 \geq 0$ as given with $p_1 + p_2 = 1$. Now, to apply Theorem 4.2.3, use

$$\mathcal{K} = \left\{K = (k_1, k_2) \in \mathbb{Z}_+^2 : k_1 + k_2 = n\right\}.$$

and distinguished vectors x^K with components

$$x_i^K = \begin{cases} -r & \text{for } 1 \le i \le k_1; \\ r & \text{for } k_1 + 1 \le i \le k_1 + k_2 \end{cases}$$

This leads to PMF

$$f_{\mathcal{S}}(x) = \sum_{K \in \mathcal{K}} p_1^{k_1} p_2^{k_2} \left(\frac{n!}{k_1! k_2!} \right) \delta\left(x - \mathcal{S}\left(x^K\right)\right)$$

where, for $K = (k_1, k_2) \in \mathcal{K}$,

$$\mathcal{S}\left(x^K\right) = (1+r)^{k_1}(1-r)^{k_2} \; + \; (1-r)^{k_1}(1+r)^{k_2} \; - \; 2.$$

Note that for this special case where $l = 2$ and the discrete random variable $X(i)$ takes on two values $-r$ and r, the number of total Dirac functions is much less than the number of elements in $\mathcal{K}$. More specifically $\|\mathcal{K}\| = n + 1$ while the number of Dirac functions is given by $\lceil \frac{n+1}{2} \rceil$.

Now for further illustration, suppose $n = 5$, then there are 6 terms in $f_S(x)$ obtained from

$$\mathcal{K} = \Big\{(0,5)\,,\ (1,4)\,,\ (2,3)\,,\ (3,2)\,,\ (4,1)\,,\ (5,0)\Big\}$$

corresponding distinguished vectors x^K given by

$$x^{(0,5)} = (r, r, r, r, r)\,; \qquad x^{(1,4)} = (-r, r, r, r, r)\,; \qquad x^{(2,3)} = (-r, -r, r, r, r)\,;$$
$$x^{(3,2)} = (-r, -r, -r, r, r)\,; \quad x^{(4,1)} = (-r, -r, -r, -r, r)\,; \quad x^{(5,0)} = (-r, -r, -r, -r, -r)$$

and the resulting PMF is given by

$$\begin{aligned} f_{\mathcal{S}}(x) \;\; = \;\; & \big(p_1^5 + p_2^5\big)\delta\left(x - \big[(1+r)^5 + (1-r)^5 - 2\big]\right) \\ & + 5p_1p_2^4\delta\left(x - \big[(1-r)(1+r)^4 + (1+r)(1-r)^4 - 2\big]\right) \\ & + 10p_1^2p_2^3\delta\left(x - \big[(1-r)^2(1+r)^3 + (1+r)^2(1-r)^3 - 2\big]\right) \\ & + 10p_1^3p_2^2\delta\left(x - \big[(1-r)^3(1+r)^2 + (1+r)^3(1-r)^2 - 2\big]\right) \\ & + 5p_1^4p_2\delta\left(x - \big[(1-r)^4(1+r) + (1+r)^4(1-r) - 2\big]\right). \end{aligned}$$

After expanding the products in r associated with each point mass and combining like terms, the PMF reduces to

$$f_S(x) \;=\; \left(p_1^5+p_2^5\right)\delta\left(x-10r^4-20r^2\right) \;+\; 5\left(p_1p_2^4+p_1^4p_2\right)\delta\left(x+6r^4-4r^2\right) \\ +\;10\left(p_1^2p_2^3+p_1^3p_2^2\right)\delta\left(x-2r^4+4r^2\right).$$

and there are 3 Dirac functions as expected. To provide a more concrete final result, we take $p_1 = 0.4$, $p_2 = 0.6$ and $r = 0.1$ in the equation above and we find that

$$f_S(x) \;\approx\; 0.088\delta\left(x-0.201\right) \;+\; 0.336\delta\left(x-0.0394\right) \;+\; 0.576\delta\left(x+0.0398\right).$$

Since our example above is only for 5 trading days, the PMF of the trading gain is very sparse. Figure 4.1 shows the central portion of the PMF based on a new calculation using a one-year period with $n = 252$ and the same parameters p_1, p_2 and r. Each stem in the figure represents a Dirac function. Although not all visible in Figure 4.1, the total number of Dirac functions is 127 rather than 253 in the general case.

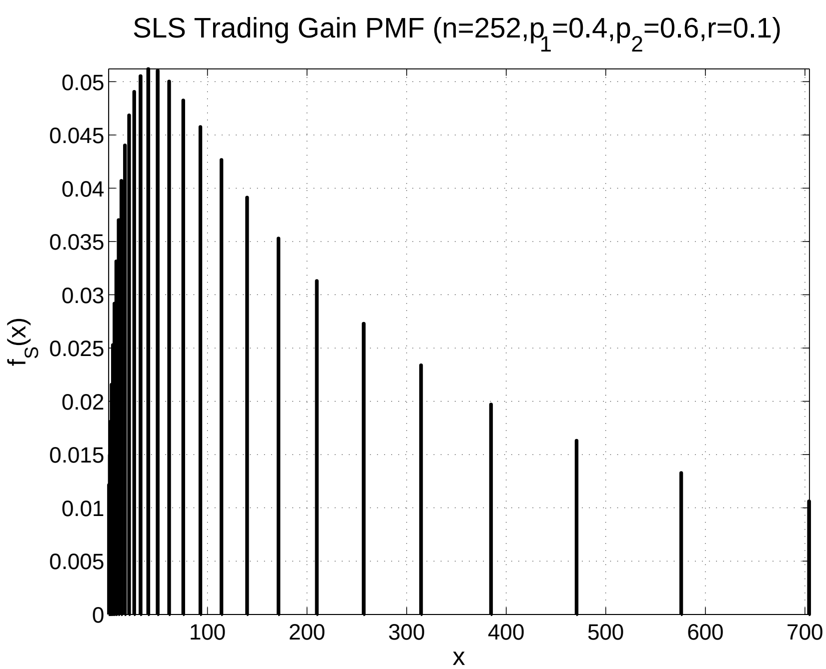

Figure 4.1 PMF of $f_S(x)$ for $n = 252$ Trading Days

4.4 Symmetry Over a Stock-Price Lattice

In the remaining sections of this chapter, we consider another example involving financial markets. In particular, we focus on the stochastic characterization of the trading gain garnered from the use of linear feedback with stock prices varying over a lattice. We begin with the simple case of a single stock and assume that the underlying stock price is governed by the classical binomial model of Cox, Ross and Rubinstein [59]. From this starting point, we derive the resulting PMF for the discrete-time trading gain. We then consider the case of two correlated stocks governed by a so-called quadrinomial lattice and derive the PMF its discrete-time trading gain.

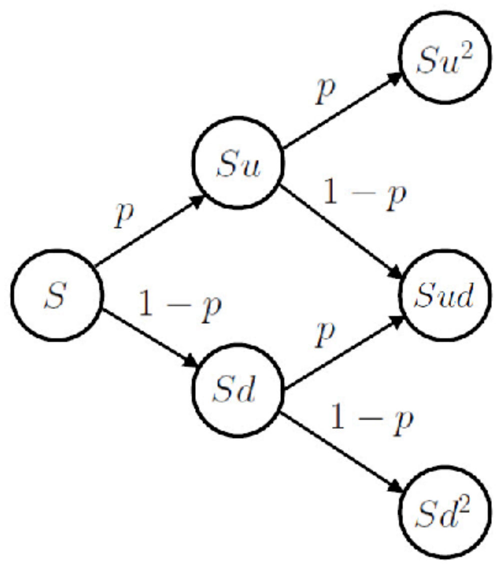

Figure 4.2 Two Stages of a Binomial Lattice

4.5 Binomial Lattice Model

This model is described in terms of two parameters $u > 1$, $0 < d < 1$ and a given probability p of "up"; see [59]. Then, going forward in time from some stock price S at stage k, it is assumed that the next price at stage $k+1$ is Su with probability p and Sd with probability $1-p$. Figure 4.4

illustrates the first two stages of a binomial lattice. At any stage k, $S(k+1)$ can either be $S(k)u$ or $S(k)d$. Later in this chapter, when working with a symmetric function over this lattice, thanks to symmetry, we see that the PMF involves only $n+1$ point masses rather than 2^n.

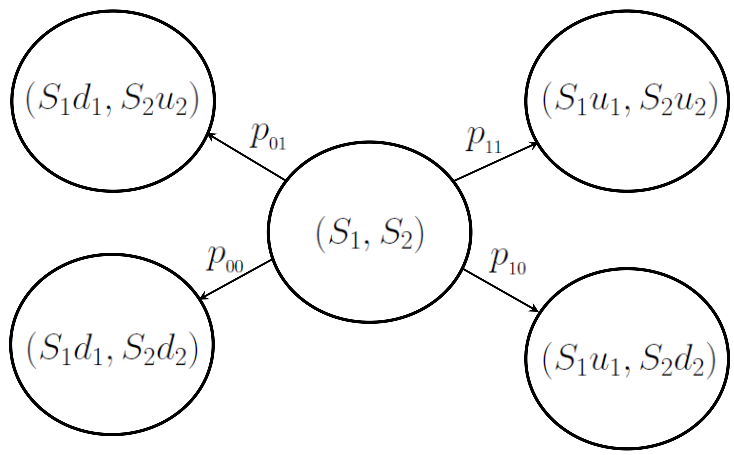

Figure 4.3 Single Transition Possibilities in Quadrinomial Lattice

4.6 Quadrinomial Lattice Model

The quadrinomial lattice model is a generalization of the binomial lattice model in [59] which is used for a single stock. We now consider two correlated stocks. To this end, we work with a model very similar to the one considered in [60] in option valuation; see also [61]-[62] for other examples of option valuation via lattices and [63] for an introduction to lattice models in finance. For the two-stock case under consideration, we have $u_1 > 1$ and $0 < d_1 < 1$ as the "up-down parameters" for the first stock price $S_1(k)$ and $u_2 > 1$ and $0 < d_2 < 1$ for the second stock price $S_2(k)$. In this more general case, beginning with price pair (S_1, S_2) at stage k in the lattice, there are four possible branches and associated probabilities for the next price pair at stage $k+1$. Namely, we transition to (S_1d_1, S_2d_2) with probability p_{00}, (S_1d_1, S_2u_2) with probability p_{01}, (S_1u_1, S_2d_2) with probability p_{10} and (S_1u_1, S_2u_2) with probability p_{11}. These transitions are depicted in Figure 4.3 and the main result, a description of the PMF for trading gains, is given in Theorem 4.11.1.

4.7 Trading Gain PMF Overview for Lattice Models

In this section, we first analyze the trading gains using a binomial lattice and then provide the generalization to the quadrinomial lattice. At stage k, the amount invested is given by

$$I(k) = I_0 + Kg(k)$$

with I_0 being the initial amount, K being the feedback and $g(k)$ the trading gain with $g(k) < 0$ denoting a loss. When I_0 and K are positive, the interpretation of the feedback law above is that the trader is "long" and benefits from increases in the stock price $S(k)$. Conversely, when I_0 and K are negative, the trader is "short" and profits are accrued if $S(k)$ decreases. Given this setting, the problem formulations in papers [53]-[57] are closest to the one given here.

This PMF of the trading gain for the binomial lattice is seen to have the same form as the PMF we found in the scenario in Section 4.3 for the "Simultaneous Long-Short" case. Our analysis is then generalized to the case of two correlated stocks whose prices are governed by a quadrinomial lattice model. We note that the PMF for the binomial lattice result can be obtained as a special case of the quadrinomial analysis using two "perfectly" correlated stocks. In this chapter, we exploit the symmetry property of the trading gains at stage $k = n$. They are a function of the random vector X whose components are single-step returns from each stock. For the single stock case, we used

$$X \doteq (X(0), X(1), \ldots, X(n-1))$$

where

$$X(k) \doteq \frac{S(k+1) - S(k)}{S(k)}$$

for $k = 0, 1, \ldots, n-1$. For the two stock case, we now take

$$X \doteq (X_1, X_2)$$

with random vector components

$$X_i \doteq (X_i(0), X_i(1), \ldots, X_i(n-1))$$

for $i = 1, 2$ and

$$X_i(k) \doteq \frac{S_i(k+1) - S_i(k)}{S_i(k)}$$

for $k = 0, 1, \ldots, n-1$. To denote the stochastic dependence of $g(n)$ on X, in the sequel, when convenient, we use the notation $\mathcal{S}(X) \doteq g(n)$. Since $\mathcal{S}(X)$ is invariant under permutations of the random single-step returns $X_i(k)$ for each stock, it is therefore a sum of symmetric functions. Hence, instead of having 2^{2n} point masses describing its probability distribution, only $(n+1)^2$ are required. For the single-stock case where $m = 1$, instead of having 2^n point masses for the PMF of the trading gain, as a result of the symmetry of $\mathcal{S}(X)$, only $(n+1)$ are required. After obtaining the PMF of $\mathcal{S}(X)$ we provide numerical examples showing how it is found from both the binomial and quadrinomial lattice models.

4.8 Trading Gain Dynamics for the Binomial Lattice

In this section, we discuss the trading gain dynamics of a single stock resulting from the use of the linear feedback control $I(k) = I_0 + Kg(k)$ over a binomial lattice described by the triple (u, d, p). We first derive sample path solutions for $g(k)$ which result from a realization of the (u, d) sequences. Indeed, suppose that $X(0), X(1), \ldots, X(n-1), X(n)$ is a sample path of stock returns generated from some underlying stochastic process. Then, when we arrive at stage k with the cumulative trading gains $g(k)$, to update to stage $k+1$, the increment to the trading gains, call it $\Delta g(k)$, is obtained by multiplying the percentage change in the stock price $X(k)$ by the amount being invested $I(k)$. That is,

$$\Delta g(k) = g(k+1) - g(k) = X(k)I(k).$$

Now substituting the linear feedback $I(k) = I_0 + Kg(k)$ above and simplifying, we obtain update dynamics

$$g(k+1) = (1 + KX(k))g(k) + X(k)I_0.$$

By viewing the recursion above as a linear time-varying system with input $X(k)I_0$, a formula for $\mathcal{S}(X)$ is readily obtained via classical state-space methods. Indeed, via a lengthy but straightforward calculation, the solution is

$$\mathcal{S}(X) = \frac{I_0}{K}\left[\prod_{i=0}^{n-1}(1+KX(i))-1\right],$$

where, considering the binomial lattice parameters u, d and p for $i = 0, 1, \ldots, n-1$ above, the random variables $X(i)$ are described by

$$\begin{aligned} P\left(X(i) = d-1\right) &= 1-p; \\ P\left(X(i) = u-1\right) &= p. \end{aligned}$$

4.9 The PMF of $\mathcal{S}(X)$ for the Binomial Lattice

The first result of this chapter, the PMF of the trading gain $\mathcal{S}(X)$ for the binomial lattice case, is given in the lemma below. Later in the chapter, this simple result is generalized to the more complex quadrinomial lattice.

4.9.1 Lemma: *Given the linear feedback control stock trading strategy $I(k) = I_0 + Kg(k)$ and binomial lattice triple (u, d, p), let*

$$x_i \doteq \frac{I_0}{K}\left[(1+K\,(u-1))^i(1+K\,(d-1))^{n-i}-1\right]$$

for $i = 0, 1, 2, ..., n$. Then, the probability mass function for the trading gain or loss $\mathcal{S}(X)$ is the sum of Dirac Delta functions given by

$$f_{\mathcal{S}}(x) = \sum_{i=0}^{n}\binom{n}{i}p^i(1-p)^{n-i}\delta\left(x-x_i\right).$$

4.9.2 Proof: To find $f_{\mathcal{S}}(x)$, for the random variable $\mathcal{S}(X)$, we first note that there are 2^n possible paths through the lattice associated with the price $S(k)$. However, since $\mathcal{S}(X)$ is a symmetric function of the $X(k)$, $\mathcal{S}(X)$ is invariant to any permutation of these returns. Hence, $\mathcal{S}(X)$ can only take on $n+1$ possible values given by

$$x_i \doteq \frac{I_0}{K}\left[(1+K(u-1))^i(1+K(d-1))^{n-i}-1\right]$$

where $i=0,1,\ldots,n$. These values span from the case where the stock goes down for n consecutive periods to the case where the stock goes up for n consecutive periods. Now for these $n+1$ values, we find the probability mass for each of them. Indeed, for x_i, there are $N_i=\binom{n}{i}$ possible paths with each such path having probability $p_i=p^i(1-p)^{n-i}$. Now summing over these $n+1$ possibilities leads to

$$\begin{aligned} f_{\mathcal{S}}(x) &= \sum_{i=0}^{n} N_i p_i \delta(x-x_i); \\ &= \sum_{i=0}^{n}\binom{n}{i}p^i(1-p)^{n-i}\delta(x-x_i). \end{aligned}$$

4.10 Analysis of the Quadrinomial Lattice

For the quadrinomial lattice with two stocks, as described in Section 4.6, the function $\mathcal{S}(X)$ is the sum of two symmetric functions, one for the first stock and one for the second. At stage k, the amount invested is given by

$$I(k)=I_1(k)+I_2(k)$$

for $k=0,1,\ldots,n-1$ with components

$$I_1(k)=I_{0,1}+K_1g_1(k);\quad I_2(k)=I_{0,2}+K_2g_2(k)$$

with $g_1(k)$ and $g_2(k)$ representing the trading gain of the stock, $I_{0,1}$ and $I_{0,2}$ being the initial amounts invested, and K_1 and K_2 being the feedback gains. When the $I_{0,i}$ and K_i are positive, the interpretation of the feedback law above is that the trader is “long” and benefits from increases in the stock price $S_i(k)$. Conversely, when $I_{0,i}$ and K_i are negative, the trader is “short” and profits are accrued if $S_i(k)$ decreases; see [53]-[57]. The overall trading gain at stage k is given by

$$g(k)=g_1(k)+g_2(k).$$

For the evolution of the stock price pair $(S_1(k), S_2(k))$, to denote stochastic dependence on X at stage $k = n$, when convenient, we write

$$\mathcal{S}(X) \doteq g_{1,n}(X_1) + g_{2,n}(X_2)$$

where

$$g_{i,n}(X_i) = \frac{I_{0,i}}{K_i}\left[\prod_{j=0}^{n-1}(1 + K_i X_i(j)) - 1\right]$$

for $i = 1, 2$ and $X_i(k)$ is the return on Stock i at stage k. Noting that the trading gains $g_{i,n}(X_i)$ above are symmetric functions, in the theorem to follow, we see that the number of terms comprising its PMF $f_S(x)$ is much smaller than what would be needed in the non-symmetric case. Now, consistent with Figure 4.3, the PMF for each $X_i(k)$ above is described by

$$\begin{aligned}
P\left(X_1(k) = d_1 - 1, X_2(k) = d_2 - 1\right) &= p_{00};\\
P\left(X_1(k) = d_1 - 1, X_2(k) = u_2 - 1\right) &= p_{01};\\
P\left(X_1(k) = u_1 - 1, X_2(k) = d_2 - 1\right) &= p_{10};\\
P\left(X_1(k) = u_1 - 1, X_2(k) = u_2 - 1\right) &= p_{11}
\end{aligned}$$

with the $p_{ij} \geq 0$ and $p_{00} + p_{01} + p_{10} + p_{11} = 1$ being satisfied.

4.11 The PMF of $\mathcal{S}(X)$ for the Quadrinomial Lattice

In this section we provide the main result of this chapter, the PMF of the trading gain $\mathcal{S}(X)$ for the Quadrinomial Lattice case.

4.11.1 Theorem: *Given the pair of linear feedback stock trading strategies which are represented by $I_1(k) = I_{0,1} + K_1 g_1(k)$ and $I_2(k) = I_{0,2} + K_2 g_2(k)$ and the quadrinomial lattice with parameters (u_1, d_1, u_2, d_2) and transition probabilities $(p_{00}, p_{01}, p_{10}, p_{11})$, let*

$$\begin{aligned}
x_{1,i} &\doteq \frac{I_{0,1}}{K_1}\left[(1 + K_1(u_1 - 1))^i (1 + K_1(d_1 - 1))^{n-i} - 1\right];\\
x_{2,i} &\doteq \frac{I_{0,2}}{K_2}\left[(1 + K_2(u_2 - 1))^i (1 + K_2(d_2 - 1))^{n-i} - 1\right]
\end{aligned}$$

for $i = 0, 1, 2, ..., n$. Then, the probability mass function $f_{\mathcal{S}}(x)$ for the overall trading gain $\mathcal{S}(x)$ is the sum of Dirac Delta functions given by

$$f_{\mathcal{S}}(x) = \sum_{i=0}^{n}\sum_{k=0}^{i}\sum_{j=k}^{n-i+k}\binom{n}{i}\binom{i}{k}\binom{n-i}{j-k}p_{00}^{n-i-j+k}p_{01}^{j-k}p_{10}^{i-k}p_{11}^{k}\delta\left(x - x_{1,i} - x_{2,j}\right).$$

4.11.2 Proof: To find $f_{\mathcal{S}}(x)$ for the random variable $\mathcal{S}(X)$, we first note that there are 2^{2n} possible paths for the lattice associated with the price pair $(S_1(k), S_2(k))$. However, since each $g_i(X_i)$ is a symmetric function, it is invariant to any permutation of these returns. Hence, each of these gains can only take on $n+1$ possible values at most. Now, for $i = 1, 2$, there are $n+1$ possible values for the $g_{i,n}(X_i)$ obtained by spanning all the "up-down" possibilities for $x_{i,j}$, given by

$$x_{i,j} \doteq \frac{I_{0,i}}{K_i}\left[(1 + K_i\left(u_i - 1\right))^j(1 + K_i\left(d_i - 1\right))^{n-j} - 1\right]$$

for $j = 0, 1, \ldots, n-1$. Consequently, there are at most $(n+1)^2$ possible values for $\mathcal{S}(X)$ obtained from sums of the form $x_{1,i} + x_{2,j}$. Now, for these $(n+1)^2$ points comprising the sample space for $\mathcal{S}(X)$, we need to find the probability mass associated with each of them.

Indeed, count up all the outcomes where after n periods of trading, i values of the single-step returns $(X_1(0), X_1(1)\ldots, X_1(n-1))$ in $g_{1,n}(X)$ are $u_1 - 1$ and j values of the single-step returns $(X_2(0), X_2(1)\ldots, X_2(n-1))$ in $g_{2,n}(X_2)$ are $u_2 - 1$ where i and j range from 0 to $n-1$. Now for fixed i and j, let k represent the number of periods where both the first stock and the second stock go up on the same period. For a particular outcome of $g_{1,n}(X_1)$ where there are i values of $X_i(k)$ set to $u_1 - 1$ and $n-i$ values set to $d_1 - 1$ and there are

$$N_{ijk} \doteq \binom{n}{i}\binom{i}{k}\binom{n-i}{j-k}$$

possible sample paths for $\mathcal{S}(X)$ for $k = 0, 1, \ldots, i$. Notice that $\binom{i}{k}$ is the number of ways for the first and second stock to have k out of i periods where both their prices go up on the same period and $\binom{n-i}{j-k}$ is the number of ways where the first stock price goes down and the second stock price either goes up or goes down over the remaining $n-i$ periods of trading. Since each trading period is independent, the probabilities associated with these sample paths are all the same and given by

$$p_{ijk} \doteq p_{00}^{n-i-j+k} p_{01}^{j-k} p_{10}^{i-k} p_{11}^{k}$$

for $k = 0, 1, \ldots, i$. Now summing over all the $(n+1)^2$ possibilities leads to

$$\begin{aligned} f_{\mathcal{S}}(x) &= \sum_{i=0}^{n} \sum_{j=0}^{n} \sum_{k=0}^{i} N_{ijk} p_{ijk} \delta\left(x - x_{1,i} - x_{2,j}\right) \\ &= \sum_{i=0}^{n} \sum_{j=0}^{n} \sum_{k=0}^{i} \binom{n}{i} \binom{i}{k} \binom{n-i}{j-k} \\ &\quad \times \ p_{00}^{n-i-j+k} p_{01}^{j-k} p_{10}^{i-k} p_{11}^{k} \delta\left(x - x_{1,i} - x_{2,j}\right) \\ &= \sum_{i=0}^{n} \sum_{k=0}^{i} \sum_{j=k}^{n-i+k} \binom{n}{i} \binom{i}{k} \binom{n-i}{j-k} \\ &\quad \times \ p_{00}^{n-i-j+k} p_{01}^{j-k} p_{10}^{i-k} p_{11}^{k} \delta\left(x - x_{1,i} - x_{2,j}\right) \end{aligned}$$

where, in the summations above, we use the convention that $\binom{i}{k} = 0$ if $k > i$ and $\binom{n-i}{j-k} = 0$ if $j - k > n - i$ or $k > j$. The proof of the theorem is now complete.

4.11.3 Remarks on Term Count and PMF Complexity: As stated earlier, as a result of symmetry, the number of point masses for the PMF of $\mathcal{S}(X)$ in Theorem 4.11.1 is reduced to $(n+1)^2$ as opposed to a potential 2^{2n} point masses without symmetry. In addition, via a straightforward computation, the number of terms in the PMF for $\mathcal{S}(X)$ is found to be

$$N = \sum_{i=0}^{n} \sum_{k=0}^{i} \sum_{j=k}^{n-i+k} 1 = \frac{1}{6}n^3 + \frac{1}{2}n^2 + \frac{4}{3}n + 1.$$

4.11.4 Binomial Lattice as a Special Case: As mentioned earlier in the chapter, the PMF formula for the quadrinomial lattice is a generalization of the one obtained for the binomial lattice. To see this, we consider the case of two "perfectly" correlated stocks where

$$p_{11} = p,\ p_{00} = 1 - p,\ p_{10} = p_{01} = 0,\ K_1 = K_2 = K,\ I_{0,1} = I_{0,2} = I_0/2,$$

$u_1 = u_2$ and $d_1 = d_2$. Then the PMF formula for the quadrinomial reduces to that given for the binomial lattice.

4.11.5 SLS as a Special Case: A second special case of interest involves the "Simultaneous Long-Short" (SLS) linear feedback control strategy considered in Section 3.5. In the context of the quadrinomial lattice framework from Section 4.10, the long investment $I_1(k)$ has initial investment $I_{0,1} = I_0$ and feedback parameter $K_1 = K$, while the short investment $I_2(k)$ has initial investment $I_{0,2} = -I_0$ and feedback parameter $K_2 = -K$. With these trading rules, we recover the PMF of $\mathcal{S}(X)$ for SLS by making the substitutions

$$p_{11} = p,\ p_{00} = 1 - p,\ p_{01} = p_{10} = 0,\ u_1 = u_2,\ d_1 = d_2$$

in the quadrinomial lattice PMF formula for $f_{\mathcal{S}}(x)$. Now the PMF becomes

$$f_{\mathcal{S}}(x) = \sum_{i=0}^{n} \binom{n}{i} p^i (1-p)^{n-i} \delta\left(x - x_i\right)$$

where

$$\begin{aligned} x_i \ &\doteq\ \frac{I_0}{K}\Big[(1 + K\left(u-1\right))^i (1 + K\left(d-1\right))^{n-i} \\ &\qquad +\ (1 - K\left(u-1\right))^i (1 - K\left(d-1\right))^{n-i} - 2\Big]. \end{aligned}$$

The SLS PMF result we found in Section 4.3 is a special case of the PMF above with $I_0 = 1$, $K = 1$, $(u - 1) = r$ and $(d - 1) = -r$.

4.12 Numerical Examples

In this section, we provide numerical examples illustrating the generation of the PMF for trading scenarios using the binomial and quadrinomial lattice models. In order to simulate realistic stock prices, we size the parameters u_i and d_i to represent trading on the order of every minute. We select u_i and d_i so that the PMF $f_{\mathcal{S}}(x)$, representing $n = 100$ periods of trading, roughly corresponds to 100 minutes of trading.

4.12.1 Binomial Lattice Example: For the single stock trading scenario, we assign up parameter $u = 1.001$, down parameter $d = 0.999$ and "bullish" transition probability $p = 0.55$. This corresponds to a daily move in the stock price which can be up to 0.1 percent up or down. For the long case, we set the initial investment to $I_0 = 1$ and the feedback gain to $K = 2$, while for the short case, the initial investment is $I_0 = -1$ and the feedback gain is $K = -2$. Now for long trading, in accordance with Theorem 4.9.1, the formula for the PMF is

$$f_{\mathcal{S}}(x) = \sum_{i=0}^{100} \binom{100}{i} (0.55)^i (0.45)^{100-i} \delta\left(x - 0.5\left[(1.002)^i (0.998)^{100-i} - 1\right]\right).$$

Since $u > d$, the probability of winning, that is, having $\mathcal{S}(X) > 0$, is greater than 0.5. In fact, based on the PMF above, we obtain $P(\mathcal{S}(X) > 0) \approx 0.8198$. We also find that $E[\mathcal{S}(X)] \approx 0.01017$. Finally, the minimum trading gain is obtained as $g_{min} \approx -0.09072$ while the maximum is $g_{max} \approx 0.1106$. Similarly, for short trading, the formula for the PMF is

$$f_{\mathcal{S}}(x) = \sum_{i=0}^{100} \binom{100}{i} (0.55)^i (0.45)^{100-i} \delta\left(x - 0.5\left[(0.998)^i (1.002)^{100-i} - 1\right]\right).$$

Since $u > d$, the probability of winning, is smaller than 0.5. Using the PMF above, we ob-

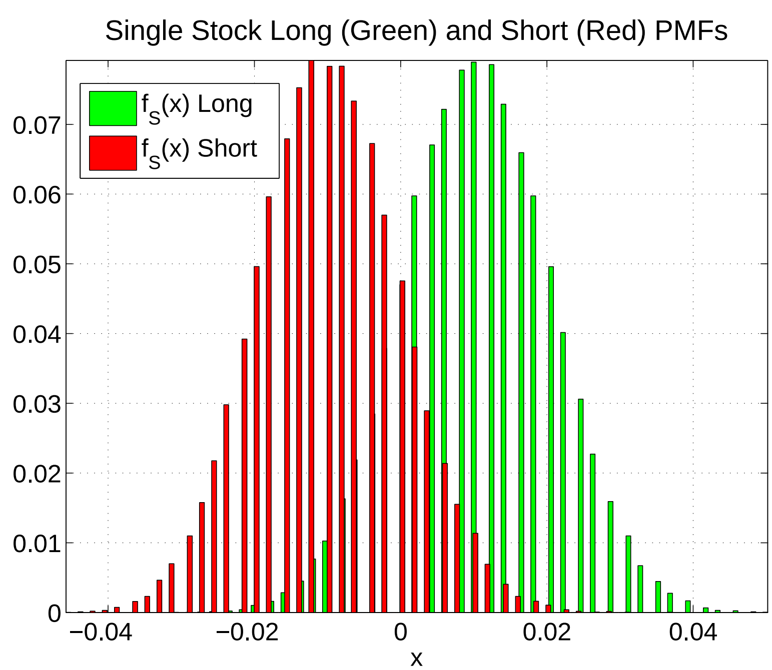

Figure 4.4 Binomial Lattice Examples

tain $P(\mathcal{S}(X) > 0) \approx 0.1321$. We also find that $E[\mathcal{S}(X)] \approx -9.979 \times 10^{-3}$. Finally, the minimum value of the gain is $g_{min} \approx -0.09072$ while $g_{max} \approx 0.1106$. For comparison purposes, we show plots of the PMF for both the long and short cases in Figure 4.4. The long case is represented by the green bars while the short case is represented by the red bars. Note that for both PMF plots, we only show the central portions of the PMF.

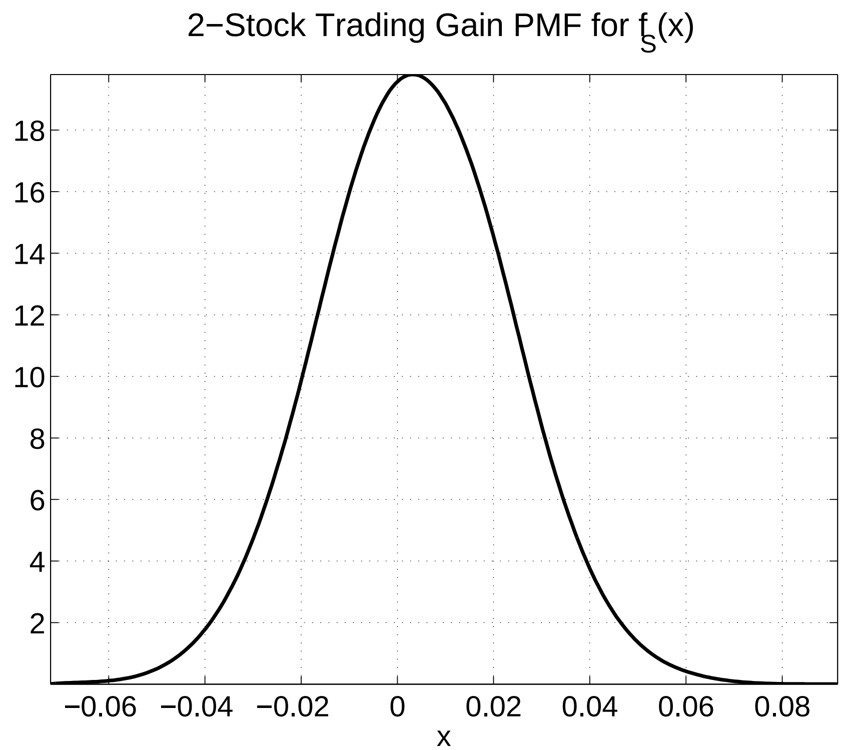

Figure 4.5 Quadrinomial Lattice Example

4.12.2 Quadrinomial Lattice Examples: For our first example, we consider a scenario where the first stock is a long trade and the second stock is a short trade. The quadrinomial lattice parameters are $u_1 = 1.001$, $u_2 = 1.002$, $d_1 = 0.999$, $d_2 = 0.998$, $p_{00} = 0.25$, $p_{01} = p_{10} = 0.125$ and $p_{11} = 0.5$. For the long trade, we take initial investment $I_{0,1} = 2$ and the feedback gain $K_1 = 2$, while for the short trade, we take $I_{0,2} = -1$ and $K_2 = -2$. Using Theorem 4.11.1, we obtain PMF

$$\begin{aligned} f_S(x) &= \sum_{i=0}^{100}\sum_{k=0}^{i}\sum_{j=k}^{100-i+k}\binom{100}{i}\binom{i}{k}\binom{100-i}{j-k}(0.25)^{100-i-j+k}(0.125)^{i+j-2k}(0.5)^{k} \\ &\quad \times\ \delta\left(x - x_{1i} - x_{2j}\right); \\ x_{1,i} &= \left[(1.002)^{i}(0.998)^{100-i} - 1\right]; \\ x_{2,j} &= 0.5\left[(0.996)^{j}(1.004)^{100-j} - 1\right]. \end{aligned}$$

The number of terms in the PMF is $N = 171,801$ while the number of point masses is $10,201$. The difference between the number of terms and the number of point masses is due to the repetition of Dirac function values. Using the PMF above, we obtain $P(\mathcal{S}(X) > 0) \approx 0.4273$ for the probability of winning. We also find that $E[\mathcal{S}(X)] \approx 0.0037$. Finally, $g_{min} \approx -0.3465$ while $g_{max} \approx 0.4665$. A plot of the PMF is shown in Figure 4.5. Note that we only show a smoothed envelope of the central portion of the distribution since the actual distribution has over $10,000$ point masses.

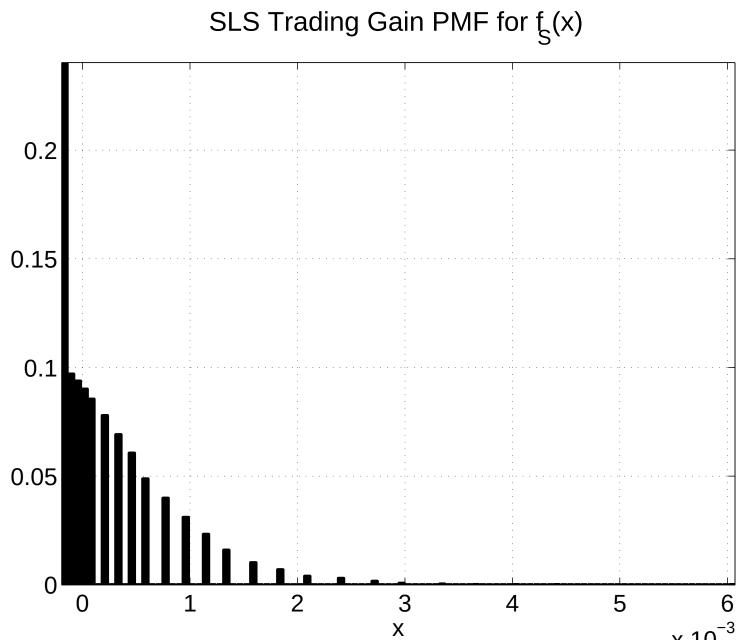

Figure 4.6 SLS Quadrinomial Lattice Example

Finally, we give an example of SLS trading as considered in Section 4.3. Now with parameters $u_1 = u_2 = 1.001$, $d_1 = d_2 = 0.999$, $p_{00} = 0.45$, $p_{01} = p_{10} = 0$, $p_{11} = 0.55$, $I_{0,1} = 1$, $I_{0,2} = -1$, $K_1 = 2$ and $K_2 = -2$, using the formulae in Section 4.10, we enhance the information in Section 4.3 with the PMF for the "combined" trading gain $\mathcal{S}(X) = g_{1,n}(X_1) + g_{2,n}(X_2)$. Using the PMF, we obtain the probability of winning, $P(\mathcal{S}(X) > 0) \approx 0.479$. We also find that $E[\mathcal{S}(X)] \approx 1.997 \times 10^{-4}$. Finally, $g_{min} \approx -0.1814$ while $g_{max} \approx 0.2212$. Figure 4.6 is a plot of the central portion of the PMF of $\mathcal{S}(X)$ for this SLS case. Notice that the probability of winning for SLS is more than that of the binomial lattice short trade, but smaller than that of the long stock trading case.

4.13 Further Research

In this chapter, we first derived the PMF for a symmetric function $S(X)$ with components X_i of X are discrete i.i.d. random variables. We applied this result to the specific case when the X_i are i.i.d. Bernoulli random variables. Later in the chapter, we provided a formula for the PMF of the trading gains resulting from a linear feedback control strategy with prices generated via a quadrinomial lattice model for a pair of correlated stock prices $(S_1(k), S_2(k))$. We believe that it should be possible to extend the ideas used for $m = 2$ stocks to $m = 3$ and beyond. To sketch the key ideas for an octonomial lattice obtained with $m = 3$ stocks, beginning with the transition probabilities $(p_{000}, p_{001}, p_{010}, p_{011}, p_{100}, p_{101}, p_{110}, p_{111})$, corresponding to combinations of up-down stock moves (u_i, d_i) of the correlated stock prices $(S_1(k), S_2(k), S_3(k))$ for $i = 1, 2, 3$, the PMF for the trading gain or loss

$$\mathcal{S}(X) \doteq g_{1,n}(X_1) + g_{2,n}(X_2) + g_{3,n}(X_3)$$

has the form

$$\begin{aligned} f_{\mathcal{S}}(x) &= \sum_{i_1=0}^{n}\sum_{i_2=0}^{n}\sum_{i_3=0}^{n}\sum_{k=0}^{i_1} C\left(n, i_1, i_2, i_3, k\right) p_{000}^{n-i_1-i_2-i_3+k} \cdots p_{111}^{k} \\ &\quad \times \ \delta\left(x - x_{1,i_1} - x_{2,i_2} - x_{3,i_3}\right), \end{aligned}$$

where

$$x_{i,j} \doteq \frac{I_{0,i}}{K_i}\left[(1 + K_i\left(u_i - 1\right))^j (1 + K_i\left(d_i - 1\right))^{n-j} - 1\right]$$

for $i = 1, 2, 3$ and $j = 0, 1, \ldots, n$. Note that the coefficient $C\left(n, i_1, i_2, i_3, k\right)$ is a product of appropriately constructed binomial coefficients and there are $(n + 1)^3$ point masses comprising the PMF.

As a result of the symmetry of $\mathcal{S}(X)$ with respect to the returns $X_1(k)$ and $X_2(k)$, out of a possible 2^{2n} return sequences comprising the sample space, we only end up dealing with $(n + 1)^2$ point masses. More generally, for the case of m stocks with $\mathcal{S}(X)$ being a sum of m symmetric functions, using our method and fully exploiting symmetry, the 2^{mn} possible return sequences collapse down to $(n + 1)^m$ probability mass points. This number, while manageable for small (m, n)

combinations, can easily become prohibitive for larger n and m. For example, with $n = 100$, in the example given in Section 4.12, the number of point masses was over $10,000$ and this number increases to over $1,000,000$ for three stocks.

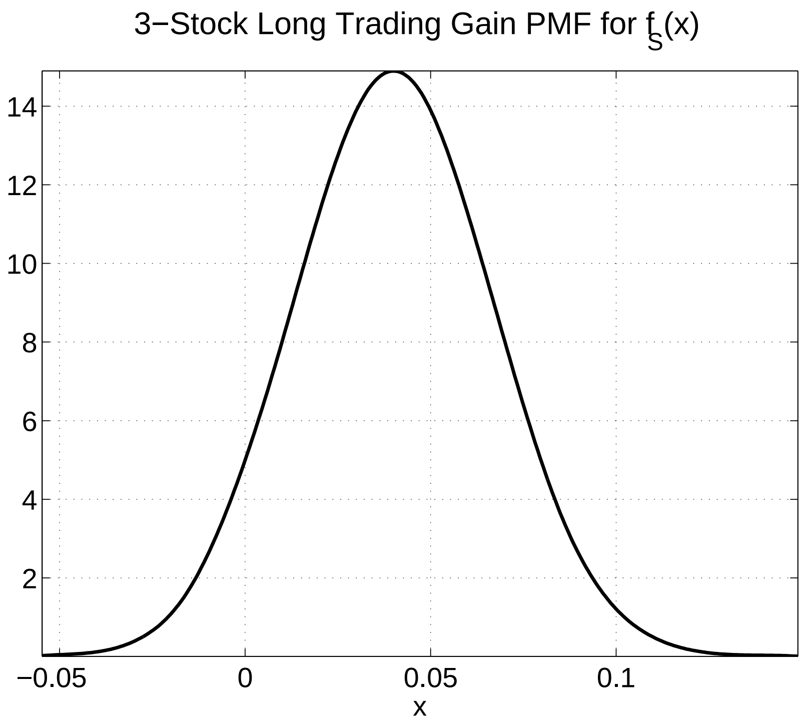

Figure 4.7 Octonomial Lattice Example

Given the growth rate of point masses described above, for $m > 2$ correlated stocks, in many cases, development of a general formula for the PMF of $\mathcal{S}(X)$ may be more of an academic exercise than of practical importance because the central portion of the PMF can readily be estimated via Monte Carlo simulation. To illustrate, for the case when $m = 3$ with $n = 100$, using $100,000$ sample paths we generated an approximation for the PMF of $\mathcal{S}(X)$ using lattice parameters

$$u_1 = 1.001,\ d_1 = 0.999,\ u_2 = 1.002,\ d_2 = 0.998,\ u_3 = 1.001,\ d_3 = 0.999;$$
$$p_{000} = p_{001} = p_{010} = p_{011} = 1/8,\ p_{100} = p_{101} = 1/16,\ p_{110} = 1/4; p_{111} = 1/8$$

and feedback and initial investment parameters are given by

$$K_1 = K_2 = K_3 = 2,\ I_{0,1} = I_{0,2} = I_{0,3} = 1.$$

Using the resulting approximation of the probability mass function, we estimated the probability of winning, $P(\mathcal{S}(X) > 0) \approx 0.96315$. We also find that $E\left[\mathcal{S}(X)\right] \approx 0.040073$. Finally, the minimum and maximum value of the trading gain are $g_{min} \approx -0.3465$ and $g_{max} \approx 0.4665$ respectively. Figure 4.7 shows a plot of $\mathcal{S}(X)$. Notice that we only show a smoothed envelope of the central portion of the PMF since it consists of over a million point masses.

Chapter 5

Multi-Group Symmetric Multilinear Optimization Problems

This chapter addresses some minimization and maximization problems for a so-called multi-group symmetric pair $(f, \mathfrak{X})$ which will be defined below. We begin the chapter by providing the definitions of multi-group symmetric functions and multi-group symmetric sets. Although a function $f(x)$ may not be symmetric, subgroups of its variables may enter into it in a symmetric way. In our multi-group symmetric framework, the variables $x_1, x_2, \ldots, x_n$ of x, get partitioned into m groups, each of equal size $N = n/m$ where n is divisible by m. We consider this setting in order to extend the notions of a symmetric pair $(f, \mathfrak{X})$ from Chapter 3.

For the initial results in this chapter, we work with the multi-group symmetric pair $(f, \mathfrak{X})$ where $f(x)$ is multilinear and $\mathfrak{X} \subset \mathbb{R}^n$ is the Cartesian product of m identical hyperrectangles $\tilde{\mathfrak{X}} \subset \mathbb{R}^N$ with non-empty interiors. That is, $\tilde{\mathfrak{X}}$ is not necessarily a hypercube as in Chapter 3; the "width" in each coordinate direction can be different. The results in this chapter specialize to those in Chapter 3 when we consider the special case when $m = n$. Then each group has size $N = 1$, the set $\mathfrak{X}$ itself becomes a hypercube and multi-group symmetry reduces to symmetry.

To provide perspective for the analysis to follow, we begin with the well-known fact that the maximum and minimum of a multilinear function on $\mathfrak{X}$ is attained among its 2^n extreme points; for example see [3]. However, as a result of multi-group symmetry, for our particular product of hyperrectangles, noting that each hyperrectangle has $M = 2^N$ extreme points, it will be shown that to obtain the optimum, we only need to consider at most

$$N_{ext} = \binom{m+M-1}{m}$$

extreme points. In many cases, especially when the group size N is not large, this can result in huge savings in the number of extreme points one needs to consider; see below. For the limiting case when $m = n$, described previously, $N_{ext} = n + 1$.

5.0.1 Extreme Point Savings: We now elaborate on the extreme point savings one can obtain in the multi-group symmetric case versus a generic multilinear function. With $M = 2^N$ extreme points for each group, we can compare N_{ext} given above with $N_{max} \doteq 2^n = 2^{mN}$. Figure 5.1 shows a logarithmic plot of N_{max}/N_{ext} for fixed N and variable m. The four plots shown in the figure correspond to the cases where $n = m$, $n = 2m$, $n = 5m$ and $n = 10m$. There are many instances that illustrate that N_{max} is orders of magnitude larger than N_{ext}. For example, when $N = 2$ and $m = 100$, $N_{ext} = 176,851$ while $N_{max}{\approx}176,851{\times}10^{55}$.

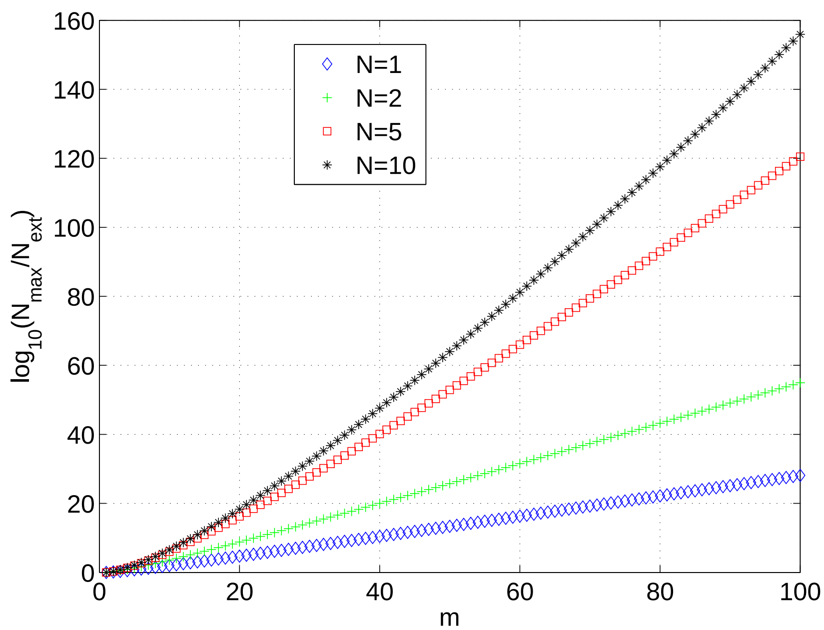

Figure 5.1 Extreme Point Comparison between N_{max} and N_{ext}

5.0.2 Extension to Polytopes: In the latter part of this chapter, we provide a theorem involving the maximization and minimization of a so-called groupwise affine function on a polytope. A groupwise affine function $f(x)$ is a special type of multilinear function that is defined below. The motivation for consideration of a groupwise affine pair is derived from applications with $f(x)$ having special structure -- over and above multilinearity. For this case, $\mathfrak{X}$ can be a product of polytopes and we obtain a theorem that shows that both the maximum and minimum of $f(x)$ are found among the extreme points of $\mathfrak{X}$. This leads to the same formula for N_{ext} as given above, but in this case, M corresponds to the number of extreme points of each group's polytope. In many cases, N_{ext} is much smaller than M^m, the number of extreme points required without exploiting multi-group symmetry. As an application of this result, we estimate the best case and worst case values of the trading gain resulting from using a polytopic containment model for the returns resulting from trading a pair of stocks. This generalizes results for a single stock given in [54].

5.1 Multi-Group Definitions

In this section, we provide the definitions for the multi-group framework. Of particular interest are the definitions of multi-group symmetric functions and multi-group symmetric sets.

5.1.1 Preliminaries: In the sequel, we often work with a partition $\mathcal{I}_m \doteq \{I_1, I_2, \ldots, I_m\}$ of the set $\{1, 2, \ldots, n\}$; i.e., the I_i are nonempty, pairwise disjoint and

$$\bigcup_{i=1}^{m} I_i = \{1, 2, \ldots, n\}.$$

We use notation $N_i \doteq \|I_i\|$ for $i = 1, 2, \ldots, m$ to denote sizes. Note that $N_1 + N_2 + \cdots + N_m = n$. To illustrate, for $m = 3$ and $n = 9$, the sets $I_1 = \{1, 3\}$, $I_2 = \{2, 4, 7\}$ and $I_3 = \{5, 6, 8, 9\}$ form a partition with $N_1 = 2$, $N_2 = 3$ and $N_3 = 4$. When all N_i are equal, we call this an *equi-partition*, and for convenience of notation we often do not refer to the underlying set $\{1, 2, \ldots, n\}$. In this case $N_i = N$ for $i = 1, 2, \ldots, m$. Associated with partition $\mathcal{I}_m$ are the *group variables* associated with $x \in \mathbb{R}^n$. That is, we work with subvectors $x^{[1]}$,$x^{[2]}$,. . .,$x^{[m]}$ of x where $x^{[j]}$ consists of the components x_i with $i \in I_j$.

5.1.2 Definition (Multi-Group Symmetric Function): Let function $f\colon \mathbb{R}^n \rightarrow \mathbb{R}$ and an equi-partition $\mathcal{I}_m = \{I_1, I_2, \ldots, I_m\}$ be given. Then f is said to be *multi-group symmetric* with respect to $\mathcal{I}_m$ if it is unchanged by any permutation of its group variables $x^{[1]}$,$x^{[2]}$,. . .,$x^{[m]}$. For simplicity, we often suppress all the qualifiers about the partition $\mathcal{I}_m$ defining the groups and simply refer to f as "multi-group symmetric."

5.1.3 Examples (Multi-Group Symmetric Multilinear Function): The multilinear function

$$f(x_1, x_2, x_3, x_4) = 7(x_1x_2 + x_3x_4) \; + \; (x_2 + x_4)$$

with partition $I_1 = \{1, 2\}$, $I_2 = \{3, 4\}$ is multi-group symmetric. That is, by inspection it satisfies

$$f\left(x^{[1]}, x^{[2]}\right) = f\left(x^{[2]}, x^{[1]}\right).$$

Notice that the function $f\colon \mathbb{R}^9 \rightarrow \mathbb{R}$ given by

$$f(x) \;=\; x_1x_3 \;+\; x_2x_4x_7 \;+\; x_5x_6x_8x_9$$

with partition $I_1 = \{1, 3\}$, $I_2 = \{2, 4, 7\}$ and $I_3 = \{5, 6, 8, 9\}$ does not qualify as a multi-group symmetric function since the group sizes are not equal.

5.1.4 Definition (Multi-Group Symmetric Set): Let set $\mathfrak{X} \subseteq \mathbb{R}^n$ be given. Also, let equi-partition $\mathcal{I}_m = \{I_1, I_2, \ldots, I_m\}$ of $\{1, 2, \ldots, n\}$ also be given. Then, the set $\mathfrak{X}$ is said to be *multi-group symmetric* with respect to $\mathcal{I}_m$ if any permutation of its groups $x^{[1]}$,$x^{[2]}$,. . .,$x^{[m]}$ belong to $\mathfrak{X}$ whenever $x \in \mathfrak{X}$.

5.1.5 Definition (Multi-Group Symmetric Pair): A *multi-group symmetric pair* $(f, \mathfrak{X})$ consists of a function $f\colon \mathbb{R}^n \rightarrow \mathbb{R}$ and a set $\mathfrak{X} \subseteq \mathbb{R}^n$, both of which are symmetric with respect to the same equi-partition $\mathcal{I}_m$.

5.1.6 Assumption: In this chapter, when dealing with symmetric functions and sets, without loss of generality we assume that a partition $\mathcal{I}_m$ is arranged in ascending order from 1 to n. That is, without loss of generality, we assume that the groups $x^{[j]}$ satisfy $x = \left(x^{[1]}, x^{[2]}, \ldots, x^{[m]}\right)$.

5.1.7 Remark on Assumption: This assumption is made without loss of generality because we can always relabel the x_i before proceeding. For example, for the function $f\colon \mathbb{R}^9 \to \mathbb{R}$ given by

$$f(x) \;=\; x_1 x_3 \;+\; x_2 x_4 x_7 \;+\; x_5 x_6 x_8 x_9$$

with partition $I_1 = \{1, 8\}$, $I_2 = \{3, 2, 6\}$ and $I_3 = \{4, 5, 7, 9\}$, to obtain ascending order, we relabel x_8 as x_2, x_2 as x_4, x_6 as x_5, x_4 as x_6, x_5 as x_7 and x_7 as x_8 so that

$$f(x) \;=\; x_1 x_3 \;+\; x_4 x_6 x_8 \;+\; x_2 x_5 x_7 x_9,$$

would be respect to the partition $I_1 = \{1, 2\}$, $I_2 = \{3, 4, 5\}$, $I_3 = \{6, 7, 8, 9\}$.

A major implication of this assumption above is that given a partition $\mathcal{I}_m$, a function $f\colon \mathbb{R}^n \to \mathbb{R}$ can be expressed as

$$f(x) = f\left(x^{[1]}, x^{[2]}, \ldots, x^{[m]}\right).$$

Therefore, a multi-group symmetric function with partition $\mathcal{I}_m = \{1, 2, \ldots, n\}$ is symmetric. In this case we have $I_j = j$ and $x^{[j]} = x_j$ for $j = 1, 2, \ldots, n$ since $m = n$.

5.2 Constraint Set for Multi-Group Symmetric Problem

As previously stated, the initial constraint set $\mathfrak{X}$ for our first multi-group symmetric problem is the product of m identical hyperrectangles. Specifically,

$$\begin{aligned}
\mathfrak{X} &\doteq \bigtimes_{i=1}^{m} \tilde{\mathfrak{X}} = \tilde{\mathfrak{X}}^m; \\
\tilde{\mathfrak{X}} &\doteq \bigtimes_{i=1}^{N} \left[c_i - r_i, c_i + r_i\right],
\end{aligned}$$

with all $r_i > 0$ and

$$x^0 \doteq (c_1, c_2, \ldots, c_N) \in \mathbb{R}^N$$

being the *center* of $\tilde{\mathfrak{X}}$. Note that for the special case where $N = 1$, $\tilde{\mathfrak{X}}$ is an interval and we recover the formulation in Chapter 3. The set of extreme points of $\mathfrak{X}$, denoted by $\mathfrak{X}_{ext}$, are the 2^n points of the form

$$x = \left(x^{[1]}, x^{[2]}, \ldots, x^{[m]}\right)$$

with each $x^{[i]}$ above being an extreme point of $\tilde{\mathfrak{X}}$; i.e., each $x^{[i]}$ is an N-tuple of the form

$$(c_1 \pm r_1, c_2 \pm r_2, \ldots, c_N \pm r_N)\,.$$

In the sequel, we use x^j, where $j \in J = \{1, 2, \ldots, M\}$, to index these extreme points. We note that the results to follow do not depend on the particular labelling scheme which is used.

5.3 Obtaining Distinguished Extreme Points

In this section, using multi-group symmetry, we construct a "smaller" distinguished subset of $\mathfrak{X}_{ext}$, called $\mathfrak{X}_{\mathcal{K}}$. In Theorem 5.4.1 to follow, we see that the maximum or minimum can be found when working with a multi-group symmetric pair.

5.3.1 Construction of Distinguished Extremes: Recall that the number of points of $\mathfrak{X}_{ext}$ is given by $2^n = 2^{mN}$ and the number of extreme points of $\tilde{\mathfrak{X}}$ is $M = 2^N$, where N is the size of each group. We now construct a distinguished subset of $\mathfrak{X}_{ext}$. Indeed, letting

$$\mathcal{K} \doteq \left\{ k = (k_1, k_2, \ldots, k_M) \in \mathbb{Z}_+^M : \sum_{i=1}^{M} k_i = m \right\},$$

recalling x^i denotes the extreme points of $\tilde{\mathfrak{X}}$, for each $k = (k_1, k_2, \ldots, k_M) \in \mathcal{K}$, we choose the extreme point x with the first k_1 groups assigned the vector x^1, the next k_2 groups assigned the vector x^2 and so on until the last k_M groups are assigned x^M. For example, if $(k_1, k_2) = (3, 2)$, then the first three subvectors of x, $x^{[1]}$,$x^{[2]}$ and $x^{[3]}$, are assigned the extreme point x^1 and the last two subvectors $x^{[4]}$ and $x^{[5]}$ are assigned the extreme point x^2. These points comprise the *distinguished extreme point set* $\mathfrak{X}_{\mathcal{K}}$ and the number of such points is given by

$$N_{ext} = \binom{m+M-1}{m}.$$

5.3.2 Example of $\mathfrak{X}_{ext}$ and $\mathfrak{X}_{\mathcal{K}}$: To illustrate construction of the set $\mathfrak{X}_{\mathcal{K}}$ defined above, suppose that $N = 2$ and $m = 2$ so that $n = 4$. Also, suppose that $c_1 = 2$, $c_2 = 4$, $r_1 = 1$ and $r_2 = 2$. Then the hyperrectangle $\mathfrak{X}$ has extreme point set

$$\mathfrak{X}_{ext} = \begin{Bmatrix} (1,2,1,2)\,, & (1,6,1,2)\,, & (1,2,1,6)\,, & (1,6,1,6) \\ (3,2,1,2)\,, & (1,2,3,2)\,, & (3,2,1,6)\,, & (1,6,3,2) \\ (3,2,3,2)\,, & (3,6,1,2)\,, & (1,2,3,6)\,, & (3,6,1,6) \\ (1,6,3,6)\,, & (3,6,3,2)\,, & (3,2,3,6)\,, & (3,6,3,6) \end{Bmatrix}$$

and the distinguished extreme point set is readily obtained as

$$\mathfrak{X}_{\mathcal{K}} = \begin{Bmatrix} (1,2,1,2)\,, & (1,6,1,2)\,, & (1,6,1,6)\,, & (3,2,1,2)\,, & (3,2,1,6) \\ (3,2,3,2)\,, & (3,6,1,2)\,, & (3,6,1,6)\,, & (3,6,3,2)\,, & (3,6,3,6) \end{Bmatrix}$$

with cardinality $\|\mathfrak{X}_{\mathcal{K}}\| = 10$ while $\|\mathfrak{X}_{ext}\| = 16$.

5.4 Hyperrectangle Extreme Point Theorem

In this section, we give our first result on multi-group symmetry and provide an example illustrating its application.

5.4.1 Theorem: *Let $(f, \mathfrak{X})$ be a multi-group symmetric pair with respect to equi-partition $\mathcal{I}_m$ where $f(x)$ is multilinear and $\mathfrak{X}$ is the product of m identical hyperrectangles with corresponding distinguished extreme point set $\mathfrak{X}_{\mathcal{K}}$. Then*

$$\begin{aligned} \min_{x\in\mathfrak{X}} f(x) &= \min_{x\in\mathfrak{X}_{\mathcal{K}}} f(x); \\ \max_{x\in\mathfrak{X}} f(x) &= \max_{x\in\mathfrak{X}_{\mathcal{K}}} f(x). \end{aligned}$$

5.4.2 Proof: We only provide a proof of the maximization case since the proof for the minimization case is nearly identical. Beginning with the fact that an optimal element

$$x^* \doteq \left(x_*^{[1]}, x_*^{[2]}, \ldots, x_*^{[m]}\right)$$

can be found in $\mathfrak{X}_{ext}$, to reduce the set of extremes to $\mathfrak{X}_{\mathcal{K}}$ we first partition $\mathfrak{X}_{ext}$ as

$$\mathfrak{X}_{ext} = \bigcup_{k \in \mathcal{K}} \mathfrak{X}_k$$

with $\mathfrak{X}_k$ consisting of the set of vectors $x \in \mathfrak{X}$ satisfying the following condition: k_1 groups are assigned the vector x^1, k_2 groups are assigned the vector x^2 and so on until k_M groups are assigned x^M. We see that the sets $\mathfrak{X}_k$ do indeed form a partition of $\mathfrak{X}_{ext}$ since the $\mathfrak{X}_k$ are all disjoint,

$$\|\mathfrak{X}_k\| = \binom{m}{k_1, k_2, \ldots, k_M},$$

and

$$\sum_{k_1 + k_2 + \cdots + k_M = m} \binom{m}{k_1, k_2, \ldots, k_M} = M^m = 2^{mN} = 2^n.$$

Since $(f, \mathfrak{X})$ is a multi-group symmetric pair with respect to $\mathcal{I}_m$, $f(x)$ is constant on $\mathfrak{X}_k$. In other words all the extreme points of $\mathfrak{X}_k$ all evaluate to the same function value. Hence to obtain the maximum of $f(x)$, we need only consider one extreme point from each set $\mathfrak{X}_k$. Since the set $\mathfrak{X}_{\mathcal{K}}$ is comprised of such points, it follows that

$$\max_{x \in \mathfrak{X}} f(x) = \max_{x \in \mathfrak{X}_{\mathcal{K}}} f(x).$$

5.4.3 Illustrative Example: To illustrate application of the theorem above, we consider finding the maximum of the multi-group symmetric multilinear function

$$f(x) \;=\; 10 \;+\; 5(x_1x_2 + x_2x_3 + x_4x_5 + x_5x_6) \;-\; 10(x_1x_2x_3 + x_4x_5x_6),$$

with partition $I_1 = \{1, 2, 3\}$ and $I_2 = \{4, 5, 6\}$, and hyperrectangle $\mathfrak{X}$ defined by

$$c_1 = 3, c_2 = 4, c_3 = 5;\; r_1 = r_2 = r_3 = 5.$$

According to the theorem, we generate the distinguished extreme point set $\mathfrak{X}_{\mathcal{K}}$ shown below. Note that this set has $N_{ext} = 36$ with $M = 2^3 = 8$ and $m = 2$, whereas $2^n = 2^6 = 64$ extreme points are needed without exploiting symmetry. Evaluating $f(x)$ at each of these extreme points, we find a maximizer at $x^* = (-2, 9, 10, -2, 9, 10) \in \mathfrak{X}_{\mathcal{K}}$ with optimal value $f(x^*) = 4330$. The set $\mathfrak{X}_{\mathcal{K}}$ is given by

$$\mathfrak{X}_{\mathcal{K}} = \left\{ \begin{array}{c} (8,9,0,8,9,0)\,,\ (8,9,10,8,9,0)\,,\ (8,9,10,8,9,10) \\ (8,-1,10,8,-1,10)\,,\ (8,9,0,8,-1,10)\,,\ (8,9,10,8,-1,10) \\ (8,-1,10,-2,9,10)\,,\ (8,9,0,-2,9,10)\,,\ (8,9,10,-2,9,10) \\ (8,9,-2,8,-1,0)\,,\ (8,9,10,8,-1,0)\,,\ (-2,9,10,-2,9,10) \\ (8,-1,0,8,-1,0)\,,\ (-2,9,10,8,-1,0)\,,\ (8,-1,10,8,-1,0) \\ (8,-1,10,-2,9,0)\,,\ (8,9,0,-2,9,0)\,,\ (8,9,10,-2,9,0) \\ (-2,9,0,-2,9,0)\,,\ (8,-1,0,-2,9,0)\,,\ (-2,9,10,-2,9,0) \\ (8,-1,10,-2,-1,10)\,,\ (8,9,0,-2,-1,10)\,,\ (8,9,10,-2,-1,10) \\ (-2,9,0,-2,-1,10)\,,\ (8,-1,0,-2,-1,10)\,,\ (-2,9,10,-2,-1,10) \\ (8,9,0,-2,-1,0)\,,\ (8,9,10,-2,-1,0)\,,\ (-2,-1,10,-2,-1,10) \\ (9,-1,0,-2,-1,0)\,,\ (-2,9,10,-2,-1,0)\,,\ (8,-1,10,-2,-1,0) \\ (-2,-1,0,-2,-1,0)\,,\ (-2,-1,10,-2,-1,0)\,,\ (-2,9,0,-2,-1,0) \end{array} \right\}.$$

5.5 Groupwise Affine Pairs and Polytopes

Previously in this chapter, we studied the multi-group symmetric pair $(f, \mathfrak{X})$ where $f(x)$ is a multilinear function on $\mathbb{R}^n$ and $\mathfrak{X}$ is the product of m hyperrectangles. Further to the preview in Section 5.0.2, in this section, we define the notion of a groupwise affine pair $(f, \mathfrak{X})$ with respect to partition $\mathcal{I}_m = \{I_1, I_2, \ldots, I_m\}$ where $f(x)$ is a so-called groupwise affine function and $\mathfrak{X}$ is a product of m polytopes $\mathfrak{X}_i$, i.e.,

$$\mathfrak{X} = \bigtimes_{i=1}^{m} \mathfrak{X}_i$$

with each $\mathfrak{X}_i \subseteq \mathbb{R}^N$. Now letting $\mathfrak{X}_{ext}$ denote the extreme points of $\mathfrak{X}_i$, we have

$$\mathfrak{X}_{ext} \doteq \bigtimes_{i=1}^{m} \mathfrak{X}_{ext,i}$$

where $\mathfrak{X}_{ext,i}$ are the extreme points of the polytope $\mathfrak{X}_i$.

5.5.1 Definition (Groupwise Affine Function): Let the function $f\colon \mathbb{R}^n \to \mathbb{R}$ be given and suppose the partition $\mathcal{I}_m = \{I_1, I_2, \ldots, I_m\}$ be given. Then f is said to be *groupwise affine* if it is affine in each of its group variables $x^{[1]}$,$x^{[2]}$,. . .,$x^{[m]}$. In other words, when all the variables associated with the groups $x^{[j]}$, where $j \neq i$, are held fixed, the resulting function is affine in $x^{[i]}$. Similar to the multilinear case, we often suppress all the qualifiers about the partition $\mathcal{I}_m$ defining the groups and simply refer to f as "groupwise affine."

5.5.2 Definition (Groupwise Affine Pair): A *groupwise affine pair* $(f, \mathfrak{X})$ consists of a groupwise affine function $f\colon \mathbb{R}^n \to \mathbb{R}$ and set $\mathfrak{X} \subseteq \mathbb{R}^n$, both of which are respect to the same partition $\mathcal{I}_m$.

5.5.3 Definition (Groupwise Affine Symmetric Pair): A *groupwise affine symmetric pair* $(f, \mathfrak{X})$ consists of a groupwise affine function $f\colon \mathbb{R}^n \to \mathbb{R}$ and set $\mathfrak{X} \subseteq \mathbb{R}^n$, both of which are symmetric with respect to the same equi-partition $\mathcal{I}_m$.

5.5.4 Remarks on Groupwise Functions: The multilinear function

$$f(x) = 6(x_1 x_4 + x_2 x_3) \; + \; 2(x_1 + x_4) \; + \; 3(x_2 + x_3) \; + \; 1$$

with partition $I_1 = \{1, 2\}$, $I_2 = \{3, 4\}$ is groupwise affine by inspection. Now, if instead we consider the multilinear function

$$f(x) \;\; = \;\; 10 \; + \; 5(x_1 x_2 + x_2 x_3 + x_4 x_5 + x_5 x_6) \; - \; 10(x_1 x_2 x_3 + x_4 x_5 x_6),$$

from Section 5.4.3 with partition $I_1 = \{1, 2, 3\}$ and $I_2 = \{4, 5, 6\}$, it is easy to see that it is not groupwise affine. That is, if $x^{[2]}$ is held fixed, the resulting function is multilinear rather than affine in (x_1, x_2, x_3).

5.5.5 Remarks on Multilinear Functions on Polytopes: To motivate the theorem to follow, we note the following: To find the maximum or minimum of a multilinear function on a polytope, one cannot simply consider the extreme points. For example, if $f(x) = x_1 x_2$, $x_1 \geq 0$, $x_2 \geq 0$ and $x_1 + x_2 \leq 1$, the maximum occurs at $(1/2, 1/2)$ which is not one of the extreme points $(0,0)$, $(0,1)$ and $(1,0)$. However, in the theorem below, for our particular multilinear function $f(x)$ and polytope $\mathfrak{X}$, it turns out that both its maximum and minimum do actually occur on an extreme.

5.5.6 Theorem: *Let $(f, \mathfrak{X})$ be a groupwise affine pair with respect to partition $\mathcal{I}_m$ where $f(x)$ is a groupwise affine function and $\mathfrak{X}$ is a polytope with extreme point set $\mathfrak{X}_{ext}$. Then*

$$\begin{aligned} \min_{x \in \mathfrak{X}} f(x) &= \min_{x \in \mathfrak{X}_{ext}} f(x); \\ \max_{x \in \mathfrak{X}} f(x) &= \max_{x \in \mathfrak{X}_{ext}} f(x). \end{aligned}$$

5.5.7 Proof: We only consider the maximization case since the proof for the minimization case is nearly identical. Indeed, let

$$f^* \doteq \max_{x \in \mathfrak{X}} f(x)$$

and suppose

$$x^* \doteq \left(x_*^{[1]}, x_*^{[2]}, \ldots, x_*^{[m]}\right)$$

is an optimal element above; i.e., $f(x^*) = f^*$. Now we consider two cases. First, if all the $x_*^{[i]}$ are extreme points of $\mathfrak{X}_i$, the proof is complete. On the other hand, if not all of the $x_*^{[i]}$ are extreme, let $x_*^{[j]}$ denote one of the components that is *not* extreme and define the function F

$$F\left(x^{[j]}\right) \doteq f\left(x_*^{[1]}, x_*^{[2]}, \ldots, x_*^{[j-1]}, x^{[j]}, x_*^{[j+1]}, \ldots, x_*^{[m]}\right)$$

which is an affine function in $x^{[j]}$. Combining this with the fact that $\mathfrak{X}_j$ is a polytope, F is maximized on $\mathfrak{X}_{ext,j}$. Now let $\overline{x}_*^{[j]} \in \mathfrak{X}_{ext,j}$ be an extreme point where F is maximized. Then letting $\overline{x}^*$ denote x^* with $\overline{x}_*^{[j]}$ substituted for $x_*^{[j]}$ it follows that $\overline{x}^*$ is an optimal point too; i.e.,

$$f(\overline{x}^*) = F\left(\overline{x}_*^{[j]}\right) = f\left(x_*^{[1]}, x_*^{[2]}, \ldots, x_*^{[j-1]}, \overline{x}_*^{[j]}, x_*^{[j+1]}, \ldots, x_*^{[m]}\right) \geq f^*.$$

Now, the argument above is repeated; i.e., if all the $x_*^{[i]}$ are extreme points the proof is complete. Otherwise, we repeat the procedure above for another component $x_*^{[j]}$ that is not extreme. Continuing this process, eventually all the $x^{[i]}$ will be perturbed to extreme points of their respective $\mathfrak{X}_i$, while preserving optimality at each step. Hence, we conclude

$$\max_{x\in\mathfrak{X}} f(x) = \max_{x\in\mathfrak{X}_{ext}} f(x).$$

5.6 Extreme Point Theorem for Groupwise Affine Case

In this section, we give a theorem that builds upon the results given in Section 5.5. We consider the groupwise affine symmetric pair $(f, \mathfrak{X})$ where $f(x)$ is a groupwise affine function on $\mathbb{R}^n$ and $\mathfrak{X} \subseteq \mathbb{R}^n$ is a product of m identical polytopes. That is, $\mathfrak{X}$ is described by

$$\mathfrak{X} \doteq \bigtimes_{i=1}^{m} \tilde{\mathfrak{X}} = \tilde{\mathfrak{X}}^m$$

and we take $\tilde{\mathfrak{X}}\subset\mathbb{R}^N$ to be described by the convex hull of its extreme points; i.e.,

$$\tilde{\mathfrak{X}} \doteq conv\left\{x^1, x^2, \ldots, x^M\right\}.$$

Hence, the set of extreme points of $\mathfrak{X}$ is given by

$$\mathfrak{X}_{ext} \doteq \bigtimes_{i=1}^{m} \left\{x^1, x^2, \ldots, x^M\right\}.$$

5.6.1 Distinguished Extreme Points: We already know from Theorem 5.5.6 that the minimum and maximum of $f(x)$ is attained at the extreme points of $\mathfrak{X}$; i.e., for some $x\in\mathfrak{X}_{ext}$. However, when $(f, \mathfrak{X})$ is the groupwise affine symmetric pair as described above, we obtain a stronger result. In this case, we define the *distinguished extreme point set* $\mathfrak{X}_{\mathcal{K}}$ in the same manner as Section 5.3.1. That is, in this case x^1,x^2,. . .,x^M are the extreme points of the polytope $\tilde{\mathfrak{X}}$ and we prove in the sequel that multi-group symmetry makes it possible to find the maximum or minimum of $f(x)$ on $\mathfrak{X}$ using

$$N_{ext} = \binom{m+M-1}{m}$$

distinguished extreme points rather than M^m extreme points.

5.6.2 Example of $\mathfrak{X}_{ext}$ and $\mathfrak{X}_{\mathcal{K}}$: Suppose $M = 3$, $m = 2$, $N = 2$, $n = 4$ and $\tilde{\mathfrak{X}}$ is the triangle in Figure 5.2 with extreme points

$$x^1 = (1, 2);\ x^2 = (3, 4);\ x^3 = (5, 2).$$

Whereas the polytope $\mathfrak{X}$ has extreme point set

$$\mathfrak{X}_{ext} = \begin{Bmatrix} (5,2,5,2)\,,\ (5,2,3,4)\,,\ (3,4,5,2) \\ (3,4,3,4)\,,\ (5,2,1,2)\,,\ (1,2,5,2) \\ (3,4,1,2)\,,\ (1,2,3,4)\,,\ (1,2,1,2) \end{Bmatrix},$$

with cardinality $\|\mathfrak{X}_{ext}\| = 9$, the distinguished extreme point set is

$$\mathfrak{X}_{\mathcal{K}} = \{(5,2,5,2)\,,\ (5,2,3,4)\,,\ (3,4,3,4)\,,\ (5,2,1,2)\,,\ (3,4,1,2)\,,\ (1,2,1,2)\}\,,$$

with cardinality $\|\mathfrak{X}_{\mathcal{K}}\| = 6$.

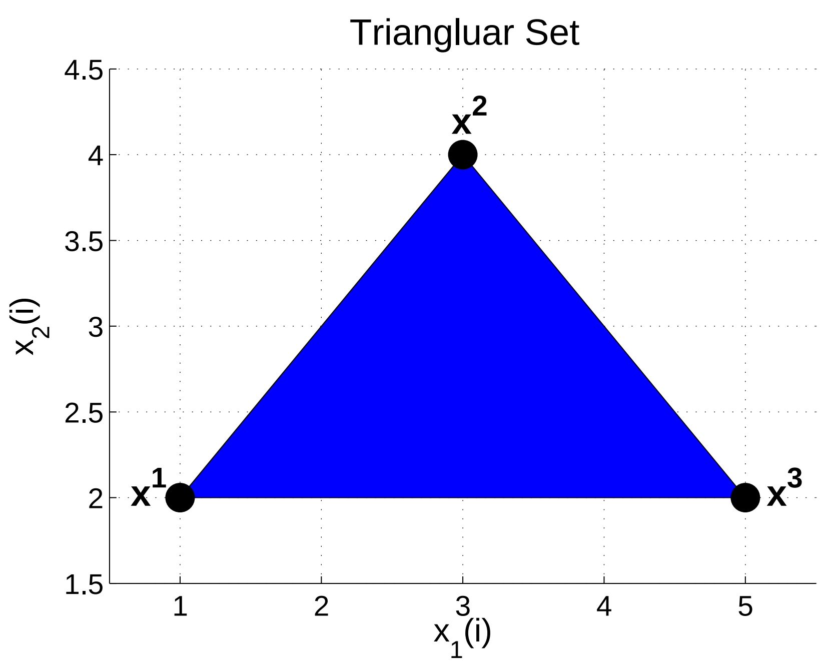

Figure 5.2 Triangle Set $\tilde{\mathfrak{X}}$ in $\mathbb{R}^2$

5.6.3 Theorem: *Let $(f, \mathfrak{X})$ be the groupwise affine symmetric pair with respect to equi-partition $\mathcal{I}_m$ where $f(x)$ is groupwise affine and $\mathfrak{X}$ is a product of identical polytopes with distinguished extreme point set $\mathfrak{X}_{\mathcal{K}}$. Then*

$$\begin{aligned} \min_{x \in \mathfrak{X}} f(x) &= \min_{x \in \mathfrak{X}_{\mathcal{K}}} f(x); \\ \max_{x \in \mathfrak{X}} f(x) &= \max_{x \in \mathfrak{X}_{\mathcal{K}}} f(x). \end{aligned}$$

5.6.4 Proof: We consider maximization noting that a nearly identical proof holds for the minimum. By Theorem 5.5.6, we know that $f(x)$ is maximized or minimized on an extreme point of $\mathfrak{X}$. The remainder of the proof is the same as that given in Section 5.4.2 except that $\mathfrak{X}$ is a polytope rather than a hyperrectangle with x^1, x^2, ..., x^M being its set of extreme points of $\tilde{\mathfrak{X}}$. In this case, M is not necessarily equal to 2^N.

5.7 Multi-Stock Trading Example

As an illustration of Theorem 5.6.3, in this section, we look at the groupwise affine symmetric pair $(f, \mathfrak{X})$ where $f(x)$ is a trading gain function and $\mathfrak{X} \subseteq \mathbb{R}^n$ is a product of m identical polytopes $\tilde{\mathfrak{X}}$ each of which represents a constraint on the daily return vector. This "polytope containment set" is constructed by taking the convex hull of the daily returns of stocks from historical data. Then moving forward, to demonstrate our theory, we assume the returns for each trading day fall within the estimated polytope $\tilde{\mathfrak{X}}$. Our goal is to investigate how large or small the trading gain could become after trading for m days.

The function $f(x)$ which we consider is derived from a linear feedback trading strategy, for example, see [54], and is given by

$$f(x) \doteq \sum_{i=1}^{N} \frac{I_{0,i}}{K_i} \left[\prod_{j=0}^{m-1} (1 + K_i x_i(j)) - 1 \right]$$

where

$$x \doteq (x(0), x(1), \ldots, x(m-1))$$

and x has subvectors

$$x^{[k+1]} = x(k) \doteq (x_1(k), x_2(k), \ldots, x_N(k))$$

for $k = 0, 1, \ldots, m-1$ with components $x_i(k)$ representing the daily returns of N correlated stocks, m being the number of trading days, $I_{0,i}$ being the initial investment and K_i being the feedback gain for Stock i respectively. Since, we also require self-financing, at stage k, the trader is not allowed to invest an amount $I(k)$ leading to the possibility that the account value $V(k+1)$ is negative. This condition means that if the returns are bounded by $|x_i(j)| < x_{max,i}$ then the feedback K_i which the trader uses for each stock satisfies $|K_i| \leq 1/x_{max,i}$. As in Section 5.6, the polytope $\mathfrak{X}$ for x is the product of m identical polytopes $\tilde{\mathfrak{X}} \subset \mathbb{R}^N$ given by the convex hull of its extreme points $x^1, x^2, \ldots, x^M$.

5.7.1 Two-Stock Example: As an illustrative application of Theorem 5.6.3, we estimate the best and worst case values that the trading gain $f(x)$ can attain when trading two stocks. We choose two stocks from the same sector with one of the stocks traded long while the other stock is traded short. In this context, the trading gain is given by

$$f(x) = \frac{I_{0,1}}{K_1}\left[\prod_{j=0}^{m-1}(1 + K_1 x_1(j)) - 1\right] + \frac{I_{0,2}}{K_2}\left[\prod_{j=0}^{m-1}(1 + K_2 x_2(j)) - 1\right]$$

where $I_{0,1} > 0$, $K_1 > 0$, $I_{0,2} < 0$ and $K_2 < 0$. To estimate the best and worst case values that $f(x)$ can possibly attain over m days of trading, we work with a polytope bound $\tilde{\mathfrak{X}}$, a polygon in $\mathbb{R}^2$ in this case, for $x(k)$ which partially accounts for correlation. It is assumed that at each period, the return $x(k) = (x_1(k), x_2(k))$ is confined to the polytope $\tilde{\mathfrak{X}}$ that is the convex hull of in-sample returns from the previous year, defined by 252 trading days. Note that $f(x)$ represents what could happen if the returns are confined to $\tilde{\mathfrak{X}}$ for the next m days with no further restriction regarding probabilities and $\mathfrak{X} = \tilde{\mathfrak{X}}^m$.

We now claim that $(f, \mathfrak{X})$ is a groupwise affine symmetric pair so that Theorem 5.6.3 is applicable. Indeed, we have m groups with partition $I_k = \{2(k-1)+1, 2(k-1)+2\}$ for $k = 1, 2, \ldots, m$

and associated group variables given by $x^{[k+1]} = (x_1(k), x_2(k))$ for $k = 0, 1, \ldots, m-1$. Now, by inspection, $f(x)$ is groupwise affine symmetric.

In the two examples below, we estimate the best case and worse case values that $f(x)$ could attain for $m = 10$ days of trading beginning August 28, 2014. This corresponds to two weeks of trading. In constructing the polytope $\tilde{\mathfrak{X}}$, we use return data from August 27, 2013 to August 27, 2014. In addition, in all cases, the initial investment in the long trading stock is $I_{0,1} = 1$ with feedback gain $K_1 = 2$ while for the short case $I_{0,2} = -1$ and $K_2 = -2$. Therefore, the trading gain becomes

$$f(x) \;=\; \frac{1}{2}\left[\prod_{j=0}^{9}(1+2x_1(j)) \;+\; \prod_{j=0}^{9}(1-2x_2(j)) \;-\; 2\right]$$

where $x = (x_1(0), x_2(0), x_1(1), x_2(1), \ldots, x_1(9), x_2(9))$.

5.7.2 Confinement Polygon for Honda-Toyota: For our first example, we consider Honda Motor Company (HMC) and Toyota Motor Corporation (TM), with HMC being the long trade and TM the short trade. These two stocks are part of the Consumer Goods sector of the stock market. The polygon $\tilde{\mathfrak{X}}$ for this example is shown in Figure 5.3. To estimate $\mathfrak{X}$, we eliminated any outliers using the Mahalanobis distance metric; see [49]-[52]. Displaying the daily return data, the x axis shows returns from the HMC stock while the y axis shows returns from the TM stock. The extreme points of the polygon are given by

$$\begin{array}{ccc} x^1 = (.0291, .0357)\,; & x^2 = (.0014, .0282)\,; & x^3 = (-.0227, .0012)\,; \\ x^4 = (-.0310, -.0112)\,; & x^5 = (-.0285, -.0175)\,; & x^6 = (-.0199, -.0254)\,; \\ x^7 = (-.0044, -.0203)\,; & x^8 = (.0097, -.0147)\,; & x^9 = (.0127, -.0126)\,; \\ x^{10} = (.0239, 0)\,; & x^{11} = (.0385, .0203)\,. & \end{array}$$

Therefore $\|\mathfrak{X}_{\mathcal{K}}\| = 184,756$ while $\|\mathfrak{X}_{ext}\| = 25,937,424,601$. After generating the distinguished extreme point set $\mathfrak{X}_{\mathcal{K}}$ via a lengthy computation, a minimum value of the trading gain is found to be given by

$$x^*(k) = \begin{cases} x^2 & \text{for } k = 0,1,2,3,4; \\ x^3 & \text{for } k = 5,6,7,8,9 \end{cases}$$

with $f(x^*) = -0.2284$ while the maximum of the trading gain occurs at $x^*(k) = x^{11}$ for all k with $f(x^*) = 0.3807$.

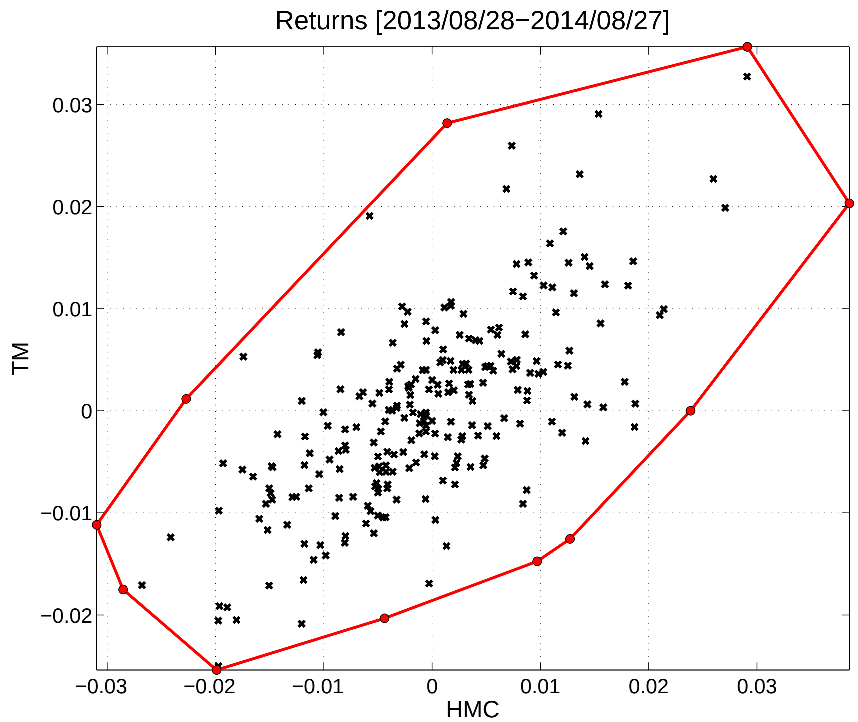

Figure 5.3 Polytope Set $\tilde{\mathfrak{X}}$ for HMC and TM

5.7.3 Confinement Polygon for FedEx and UPS: For our second example, the long trade is done on FedEx (FDX) while the short trade is done on United Parcel Service (UPS). The polygon $\tilde{\mathfrak{X}}$ for this example is shown in Figure 5.4. Again, outliers are eliminated. The x axis shows returns for FedEx stock while the y axis shows returns from the UPS stock. The polygon's extreme points are given by

$$x^1 = (.0220, .0178)\,; \quad x^2 = (.0035, .0227)\,;$$
$$x^3 = (-.0404, -.0165)\,; \quad x^4 = (-.0166, -.0227)\,;$$
$$x^5 = (-.0150, -.0223)\,; \quad x^6 = (.0116, -.0134)\,;$$
$$x^7 = (.0279, -.0002)\,; \quad x^8 = (.0282, .0010)\,;$$
$$x^9 = (.0278, .0126)\,; \quad x^{10} = (.0252, .0151)\,.$$

Figure 5.4 Polytope Set $\tilde{\mathfrak{X}}$ for FDX and UPS

Therefore $\|\mathfrak{X}_{\mathcal{K}}\| = 92,378$ while $\|\mathfrak{X}_{ext}\| = 10^{10}$. After generating the distinguished extreme point set $\mathfrak{X}_{\mathcal{K}}$, via a lengthy computation, a minimum value of the trading gain is found to be given by

$$x^*(k) = \begin{cases} x^2 & \text{for } k = 0, 1, \ldots, 5; \\ x^3 & \text{for } k = 6, 7, 8, 9 \end{cases}$$

with $f(x^*) = -0.1969$ while for the maximum case $f(x^*) = 0.3624$ with $x^*(k) = x^7$ for all k.

5.8 Conclusion and Further Research

In this chapter, we studied the maximization and minimization of multi-group symmetric functions. Our first theorem involved the multi-group symmetric pair $(f, \mathfrak{X})$ where f was multilinear and $\mathfrak{X}$ was a product of m identical hyperrectangles each with $M = 2^N$ extreme points. We showed that as a result of multi-group symmetry, instead of 2^n extreme points for a general multilinear function, we only need to consider $N_{ext} = \binom{m+M-1}{m}$ extreme points to find the minimum or maximum of f on $\mathfrak{X}$. The second theorem involved the groupwise affine pair $(f, \mathfrak{X})$ where f was a groupwise affine function and $\mathfrak{X}$ was the product of m polytopes. We showed that as a result of being groupwise affine, f was maximized and minimized on the extreme points of the polytope $\mathfrak{X}$. The final theorem of the chapter involves the groupwise affine symmetric pair $(f, \mathfrak{X})$ where $\mathfrak{X}$ is the product of m identical polytopes each with M extreme points. We showed that as a result of multi-group symmetry, instead of M^m extreme points, we only need to consider $N_{ext} = \binom{m+M-1}{m}$ extreme points to maximize or minimize f on $\mathfrak{X}$. As an illustration of this result, we considered a polytope confinement set for daily stock returns and estimated the best and worst case values of the resulting groupwise affine trading gain.

Based on the results in this chapter, one interesting direction for further research involves an investigation of the computational complexity associated with finding the distinguished extreme point set $\mathfrak{X}_{\mathcal{K}}$ for the polytope $\mathfrak{X}$ for the two-stock application of the polytope containment model. We can characterize the computational time as a function of the number of extreme points M of the polytope $\tilde{\mathfrak{X}}$ and the number of trading days m used to estimate the best and worst case values of the trading gain. For the examples given in this chapter with $M = 10$ extreme points and $m = 10$ trading days, it took about 10 minutes to compute the best and worse case values of the trading gain using MATLAB.

Chapter 6

On a Class of Resource Allocation Problems

The results in this chapter constitute an extension of work in previous chapters on optimization problems involving the pair $(f, \mathfrak{X})$. In Chapter 3, we looked at the case where $f(x)$ was a symmetric multilinear function and $\mathfrak{X}$ was a symmetric hypercube. In Chapter 5, we extended the results to address the so-called multi-group symmetric case. In this chapter, we consider a class of resource allocation problems involving minimization of total "inventory carrying costs" obtained with $f(x)$ being a separable, concave sum of symmetric functions. The constraint set $\mathfrak{X}$, is taken to be a simplex associated with admissible resource allocations.

6.1 The Inventory Carrying Cost Problem

The Inventory Carrying Cost Problem (ICCP) which we consider involves n suppliers, m identical warehouses and a single resource. Each supplier begins with an amount $r_i > 0$ of this resource which can be allocated to a prescribed subset of the warehouses. Each warehouse has a carrying cost associated with its new and pre-existing inventory and the goal is to minimize the sum of the carrying costs.

6.1.1 Suppliers and Warehouses: We formulate the ICCP using the multiple continuous resource allocation model in [68]. As previously stated, each supplier $i \in \{1, 2, \dots, n\}$ is assumed to have an amount of the resource $r_i > 0$. We let I_i be the set of indices of warehouses to which supplier i can supply some portion of r_i. Also, we let $n_i \doteq \|I_i\|$ be the number of warehouses associated with the i-th supplier. To avoid trivialities, we assume each I_i is non-empty. Without loss of generality, it is assumed that no two I_i are identical otherwise we can sum the r_i levels of two identical suppliers and treat them as one "super supplier." As a result of this assumption, the number of suppliers is $n \leq 2^m - 1$. Now for each warehouse $j \in \{1, 2, \ldots, m\}$, we let J_j be the set of indices indicating the suppliers from which warehouse j can receive amounts of the resource. This notation is purely for convenience since the J_j can be derived from the I_i. That is, if $j \in I_i$ then $i \in J_j$.

6.1.2 Variables and Constraints: For each $j \in I_i$, we let $x_{i,j} \geq 0$ be the amount of the resource which supplier i decides to send to warehouse j. Then, for supplier i, letting $x^{[i]}$ be the allocation vector with components $x_{i,j}$ arranged in ascending order on j, the decision vector for this problem can be represented by the $\overline{n}$-tuple

$$x \doteq \left(x^{[1]}, x^{[2]}, \ldots, x^{[n]}\right).$$

where

$$\overline{n} \doteq \sum_{i=1}^{n} n_i.$$

The constraint set for supplier i is the simplex $\mathfrak{X}_i$ described by

$$\sum_{j \in I_i} x_{i,j} = r_i$$

and the overall constraint set is the polytope described by the product of simplices

$$\mathfrak{X} \doteq \bigtimes_{i=1}^{n} \mathfrak{X}_i.$$

The extreme points of $\mathfrak{X}$, denoted by $\mathfrak{X}_{ext}$ consists of tuples

$$\left(x^{l_1}, x^{l_2}, \ldots, x^{l_n}\right)$$

where x^{l_i} is an extreme point of $\mathfrak{X}_i$ for $i = 1, 2, \ldots, n$. Each component of x^{l_i} belongs to $\{0, r_i\}$. Finally, taking $x_{0,j}$ to be the pre-existing inventory for warehouse j, its total inventory is given by

$$X_j(x) \doteq x_{0,j} + \sum_{i \in J_j} x_{i,j}.$$

6.1.3 Carrying Costs and Optimization Problem: It is assumed that each of the warehouses have identical concave carrying costs determined by the concave function $g\colon \mathbb{R}_+ \to \mathbb{R}$. Hence, the carrying cost for warehouse j is

$$f_j(x) \doteq g(X_j(x))$$

and the total overall carrying cost, summed over all warehouses, is the concave function

$$f(x) \doteq \sum_{j=1}^{m} f_j(x).$$

Our goal is to find an optimal allocation vector x^* minimizing $f(x)$ with respect to $x \in \mathfrak{X}$. That is, we seek $x^* \in \mathfrak{X}$ such that

$$f(x^*) \doteq \min_{x \in \mathfrak{X}} f(x).$$

6.2 Extreme Points and Computational Considerations

Since $\mathfrak{X}$ is a polytope and $f(x)$ is concave, it is well known that a minimizer is obtained at one or more of the extreme points of $\mathfrak{X}$; see [75]. Accordingly, for any proposed algorithm aimed at finding a solution via function evaluation at the extreme points, in order to evaluate the required computational effort, it is important to count the number of such points.

6.2.1 Preview of Main Result: Our main result in this chapter is an algorithm which finds the optimum using

$$N_*(m) \doteq m!$$

extreme points of $\mathfrak{X}$ at most. Notice that this result is independent of n, the number of suppliers. Therefore, for problems with a relatively small number of warehouses m in comparison with the number of suppliers n, N_* can be orders of magnitude less than the total number of extreme points of $\mathfrak{X}$. To elaborate, for fixed m, the number of extreme points of $\mathfrak{X}$ can be as large as

$$N_{max}(m) \doteq \prod_{i=1}^{m} i^{\binom{m}{i}}.$$

For this worst-case scenario, even for small m, the reduction in the number of extreme points for our method can be dramatic. For example, if $m = 4$, $N_{max} = 20,736$ while $N_* = 24$.

6.2.2 Counting Extreme Points (More Details): Recall from earlier in the chapter that the number of warehouses associated with the i-th supplier is $n_i = \|I_i\|$. Now noting that $\mathfrak{X}$ is the product of simplices $\mathfrak{X}_i$ of dimension n_i, it follows that its number of extreme points $\|\mathfrak{X}_{ext}\| \doteq \prod_{i=1}^{n} n_i$. For a given number of warehouses m and suppliers n, we ask the question: How large can the number of extreme points, call it $\mathcal{N}_{max}(m,n)$, be? To answer this, we want to determine what choice of admissible I_i leads to the largest product of the $\mathcal{N}_{max}(m,n)$ above. Indeed, given that no two I_i are identical, $\mathcal{N}_{max}(m,n)$ is maximized with each supplier "connected" to as many warehouses as possible. This leads to one supplier with $n_i = m$, then up to $\binom{m}{m-1}$ suppliers with $n_i = m-1$, up to $\binom{m}{m-2}$ suppliers with $n_i = m-2$, etc. This process continues until $n_1 + n_2 + \cdots + n_n = n$. For example, if $m = 4$ and $n = 4$, then one scenario where the largest number of extreme points is reached is with $I_1 = \{1,2,3,4\}$, $I_2 = \{1,2,3\}$, $I_3 = \{1,2,4\}$ and $I_4 = \{1,3,4\}$. This leads to 108 extreme points whereas our results to follow lead to $N_*(4) = 24$ extremes.

6.2.3 Formula for $\mathcal{N}_{max}(m,n)$: To summarize the argument above, for a given $n \leq 2^m - 1$, via a maximization over the choice of I_i, the number of extreme points can be as high as

$$\mathcal{N}_{max}(m,n) \quad = \quad (m - k^* - 1)^{n - \sum_{i=0}^{k^*} \binom{m}{i}} \prod_{i=0}^{k^*} (m-i)^{\binom{m}{i}}$$

where

$$k^* \doteq k^*(m,n) = \max\left\{k \colon \sum_{i=0}^{k} \binom{m}{i} \leq n\right\}.$$

When the number of suppliers is $n = 2^m - 1$, the maximum possible, we refer to this as a "full problem." We would like to determine how large the number of extremes can be by choice of I_i. In this case $\mathcal{N}_{max}(m,n)$ depends only on m and

$$N_{max}(m) \doteq \mathcal{N}_{max}(m, 2^m - 1) = \prod_{i=0}^{m-1} (m-i)^{\binom{m}{i}} = \prod_{i=1}^{m} i^{\binom{m}{i}}.$$

For this worst-case scenario, even for small m, the number of extreme points is quite large. For example, if $m = 5$, there are $309, 586, 821, 100$ extreme points which contrasts to $N_*(5) = 120$ using the results to follow.

6.2.4 Contribution of this Chapter: To obtain the result $N_*(m) = m!$ in this chapter, we show that there is a "small" distinguished set of extreme points which determine the minimum of $f(x)$. In order to compare the difference in extreme points between $N_{max}(m)$ and our solution with $N_*(m)$, we use the metric

$$\log\left(\frac{N_{max}(m)}{N_*(m)}\right) = \sum_{i=1}^{m}\binom{m}{i}\log(i) \;-\; \log(m!)\,.$$

Figure 6.1 shows a plot of this function using a base 10 logarithmic scale. The fact that $N_{max}(m)$ is orders of magnitude larger than $N_*(m)$ as is illustrated by a problem with seven warehouses. Whereas $N_*(7) = 5040$, which is in the realm of modest desktop computation, a full problem with the same number of warehouses involves $N_{max}(7) = 5040 \times 10^{65}$.

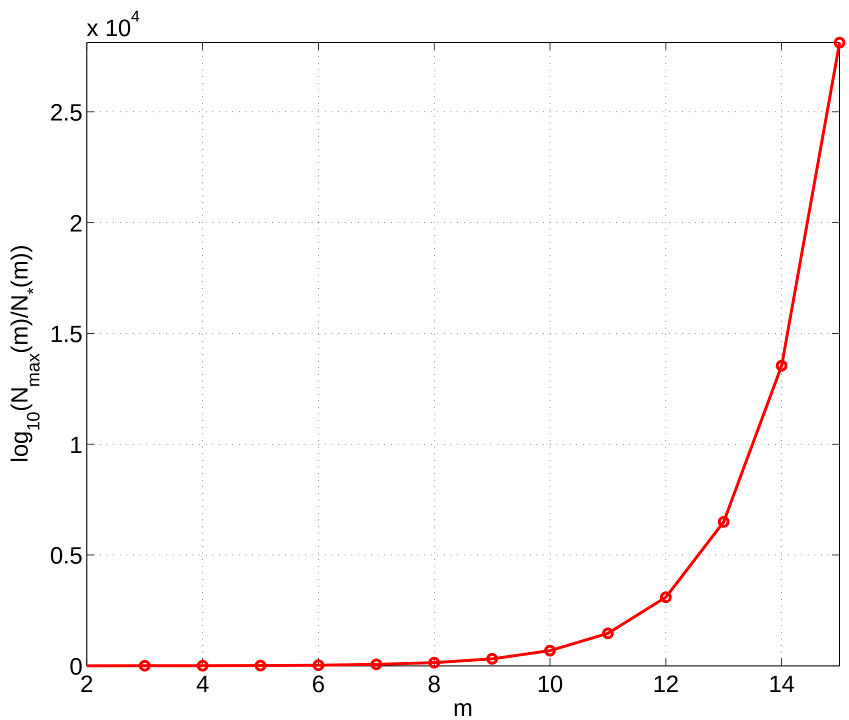

Figure 6.1 Extreme Point Comparison between $N_{max}(m)$ and $N_*(m)$

6.3 Related Literature

There are many resource allocation results in the literature for the separable objective function case; e.g., see [68]-[74]. Most of these papers involve minimization of a convex separable objective function over a linear constraint set; e.g., see [68] and [69].

As our resource allocation problems involve a concave objective function and a compact set $\mathfrak{X}$ obtained via linear constraints, extreme points of $\mathfrak{X}$ are in play. The concave minimization problems closest to the formulation in this chapter are the papers in the literature involving a concave separable objective function and linear constraints; e.g., see [75]-[85]. Most of the problems in these papers have exponential complexity and only special cases of them can be solved in polynomial time; e.g., see [75]. In most cases, branch and bound algorithms involving linear approximations of a concave objective function are used as part of the solution; e.g., see [75]-[81].

Another class of concave minimization problems that give rise to extreme point solutions involve Leontief substitution models for multi-inventory systems; see [82]. Some examples include deterministic multi-facility economic lot size and warehousing problems with concave costs. For the general Leontief formulation, the computational effort needed to search extreme points increases exponentially with the size of the problem. However, for certain special cases, one can use dynamic programming with computational effort increasing only polynomially with the size of the problem. One example of this involves certain arborescence multi-echelon structures that have separable concave production costs; see [82].

Finally we mention the large class of Minimum Concave-Cost Network Flow Problems (MCCNFP) which also admit extreme point solutions; see [85]. MCCNFP arises from a directed graph problem consisting of nodes and arcs. In MCCNFP, the objective function is a separable concave cost function and most of these problems have exponential complexity; see [85]. There are a few exceptions involving specially structured versions of the MCCNFP for which polynomial time algorithms have been developed. For example, in [84], for a fixed number of sources h and nonlinear arc costs k, a polynomial time algorithm is given for small values of h and k. However, the problem becomes exponentially hard for larger values of h and k.

6.4 Motivating Example

To provide an example of Inventory Carrying Cost Problem to be solved, we consider $m = 3$ suppliers and $n = 4$ warehouses and imagine the resource being a "divisible" commodity. In this case, let the commodity be rice. The rice is measured in pounds. Figure 6.2 above shows a directed graph depicting the situation at hand. The warehouses are represented by the rectangles which contain their pre-existing inventories $x_{0,i}$, while the suppliers are represented by the circles which contain their r_i values. The black arrows protruding from each supplier circle indicate the warehouses that can receive their rice quantities.

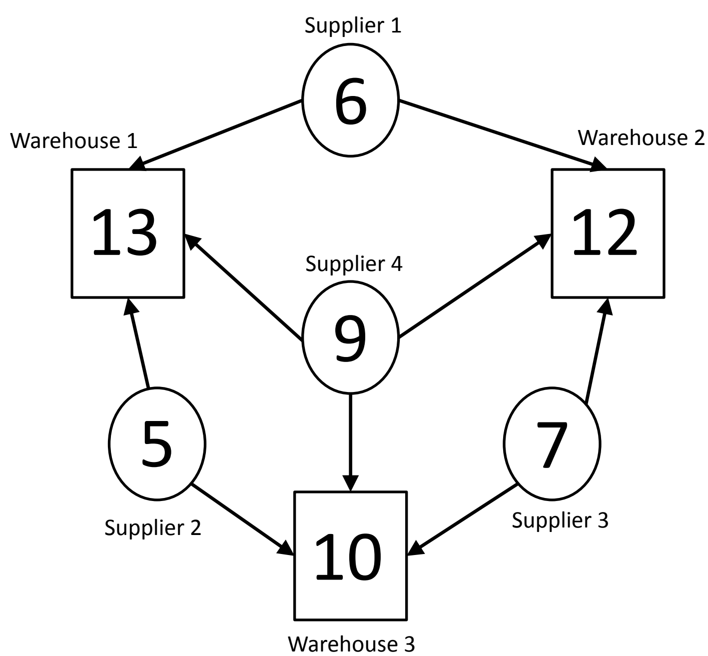

Figure 6.2 Directed Graph of Rice Supplier-Warehouse Problem

Now our Inventory Carrying Cost Problem is to find an x^* minimizing

$$f(x) = \sum_{j=1}^{3} \log(1 + X_j(x));$$

$$\begin{aligned} X_1(x) &= 13 + x_{1,1} + x_{2,1} + x_{4,1}; \\ X_2(x) &= 12 + x_{1,2} + x_{3,2} + x_{4,2}; \\ X_3(x) &= 10 + x_{2,3} + x_{3,3} + x_{4,3} \end{aligned}$$

and $x = (x_{1,1}, x_{1,2}, x_{2,1}, x_{2,3}, x_{3,2}, x_{3,3}, x_{4,1}, x_{4,2}, x_{4,3})$. The warehouses that each supplier can provide rice quantities to are $I_1 = \{1, 2\}$, $I_2 = \{1, 3\}$, $I_3 = \{2, 3\}$ and $I_4 = \{1, 2, 3\}$. The overall constraint set is

$$\mathfrak{X} = \bigtimes_{i=1}^{4} \mathfrak{X}_i$$

where the constraint set $\mathfrak{X}_i$ for supplier $i \in \{1, 2, 3, 4\}$ is described by

$$\begin{aligned} x_{1,1} + x_{1,2} &= 6; \\ x_{2,1} + x_{2,3} &= 5; \\ x_{3,2} + x_{3,3} &= 7; \\ x_{4,1} + x_{4,2} + x_{4,3} &= 9. \end{aligned}$$

The minimizer to this instance of our ICCP occurs at an extreme point of $\mathfrak{X}$. This corresponds to each of the suppliers allocating all of their rice quantities to only one of the warehouses to which they can supply. For example, one extreme point allocation would be if Warehouse 1 ended up with 24 pounds of rice (6 pounds from Supplier 1 and 5 pounds from Supplier 2), Warehouse 2 with 19 pounds of rice (7 pounds from Supplier 3) and Warehouse 3 with 19 pounds of rice (9 pounds from Supplier 4). Now this problem has a total of 24 extreme points while our theory to follow only requires an enumeration of 6 extreme points to find a minimizer. Each of these special extreme points has a property which we will call Ordered Loading Property (OLP). The OLP extreme point which minimizes this ICCP assigns 18 pounds of rice to Warehouse 1, 34 pounds of rice to Warehouse 2 and 10 pounds of rice to Warehouse 3. It is constructed by supplying Warehouse 2 with as many pounds of rice as possible, then Warehouse 1 with as many pounds of rice as possible and finally Warehouse 3 with as many pounds of rice as possible.

6.5 Ordered Loading Property Definition and Theorem

In this section, we give the main result of this chapter, the Ordered Loading Property (OLP) Theorem. We begin by giving a formal definition of the OLP extreme points. In the sections to follow, we prove the theorem using results in the literature involving majorization and Schur concavity in [86].

6.5.1 OLP Distinguished Points: The set of OLP extreme points are a distinguished set of $m!$ points of $\mathfrak{X}_{ext}$ which will be denoted by $\mathfrak{X}_{OLP}$. Each OLP distinguished point $x \in \mathfrak{X}_{OLP}$ is constructed using one of the $m!$ bijections ω from $\{1, 2, \ldots, m\}$ to $\{1, 2, \ldots, m\}$. This corresponds to specifying the ordering which will be used to service the warehouses. More specifically, under an OLP assignment, Warehouse $\omega(1)$ will be allocated as much of the resource as allowed by the constraints. Recalling that J_j is the set of indices indicating which suppliers can deliver to Warehouse j, this forces

$$x_{i,\omega(1)} = r_i \quad \text{for all} \quad i \in J_{\omega(1)}.$$

To obtain the remaining components of this OLP extreme point, supplier inventories are updated and the process is repeated using $\omega(2)$, $\omega(3)$, etc. That is,

$$x_{i,\omega(j)} = r_i \quad \text{for all} \quad i \in \ J_{\omega(j)} \ - \bigcup_{k=1}^{j-1} J_{\omega(k)}$$

for $j = 2, 3, \ldots, m$. This results in $m!$ distinguished extreme points comprising $\mathfrak{X}_{OLP}$.

6.5.2 Theorem (OLP): *Let $\mathfrak{X}_{OLP}$ be the set of OLP distinguished points of $\mathfrak{X}_{ext}$. Then*

$$\min_{x \in \mathfrak{X}} f(x) \quad = \quad \min_{x \in \mathfrak{X}_{OLP}} f(x).$$

6.6 Preliminaries for the Proof of the OLP Theorem

The proof of the theorem is facilitated via some preliminaries. First, let $T_j : \mathbb{R}_+^{\overline{n}} \rightarrow \mathbb{R}_+^m$ be the affine linear transformation taking x to $X_j(x)$ defined in Section 6.1. Then, its vectorized version $T(x)$ with components $T_j(x)$ takes the original minimization to an equivalent problem over $T(\mathfrak{X})$. Namely,

$$\min_{x\in\mathfrak{X}} f(x) \;=\; \min_{X\in T(\mathfrak{X})} F(X)$$

where $T(\mathfrak{X})$ is the polytope with extreme points which describe all possible m-tuples of warehouse inventories resulting from each supplier directing all resources to only one warehouse and

$$F(X) \;\doteq\; \sum_{j=1}^{m} g(X_j),$$

is easily seen to be symmetric and concave. The definitions to follow will be used in the proof of the OLP Theorem. They can be found in [86].

6.6.1 Definition (Majorization): For any $X = (X_1, X_2, \ldots, X_m) \in \mathbb{R}^m$, let

$$X_\downarrow \doteq \left(X_{[1]}, X_{[2]}, \ldots, X_{[m]}\right)$$

denote the re-ordering of the components of X in non-decreasing order. Now for

$$X, Y = (Y_1, Y_2, \ldots, Y_m) \in \mathbb{R}^m,$$

we say X *weakly majorizes* Y if for each $j \in \{1, 2, \ldots, m\}$

$$X_{[1]} + X_{[2]} + \cdots + X_{[j]} \geq Y_{[1]} + Y_{[2]} + \cdots + Y_{[j]}.$$

If in addition

$$\sum_{j=1}^{m} X_j = \sum_{j=1}^{m} Y_j,$$

then we say X *majorizes* Y or alternatively, Y is majorized by X. We denote this by writing $X \succ Y$.

6.6.2 Definition (Schur Concave): A real-valued function F defined on a set $\tilde{\mathfrak{X}} \subset \mathbb{R}^m$ is said to be *Schur concave* on $\tilde{\mathfrak{X}}$ if whenever $X = (X_1, X_2, \ldots, X_m)$ and $Y = (Y_1, Y_2, \ldots, Y_m)$ are vectors in $\tilde{\mathfrak{X}}$ such that X majorizes Y, then $F(X) \leq F(Y)$.

6.6.3 Remark (Symmetric Concave Functions): Suppose $g\colon \mathbb{R}_+ \to \mathbb{R}$ is concave, then the function $F\colon \mathbb{R}_+^m \to \mathbb{R}$ given by

$$F(X) \doteq \sum_{j=1}^{m} g(X_j),$$

is Schur concave; see [86].

6.7 Proof of the OLP Theorem

Recalling the transformation T from Section 6.6, we know that

$$\min_{x \in \mathfrak{X}} f(x) = \min_{X \in T(\mathfrak{X})} F(X).$$

Now let $Y = (Y_1, Y_2, \ldots, Y_m)$ be an arbitrary point in $T(\mathfrak{X})$. Then, to prove the theorem, it suffices to show that there exists some $X \in T(\mathfrak{X}_{OLP})$ such that $F(X) \leq F(Y)$. Indeed, let

$$\omega\colon \{1, 2, \ldots, m\} \to \{1, 2, \ldots, m\}$$

be the bijection that reorders the components of Y in decreasing order; i.e.,

$$Y_{\downarrow} = \left(Y_{\omega(1)}, Y_{\omega(2)}, \ldots, Y_{\omega(m)}\right).$$

Subsequently let $X = (X_1, X_2, \ldots, X_m)$ be the image of the corresponding OLP distinguished point under T with bijection ω. Then X has components

$$X_{\omega(1)} \doteq x_{0,\omega(1)} + \sum_{i \in J_{\omega(1)}} r_i,$$

and for $j = 2, 3, \ldots, m$,

$$X_{\omega(j)} \doteq x_{0,\omega(j)} + \sum_{i \in J_{\omega(j)} - \bigcup_{k=1}^{j-1} J_{\omega(k)}} r_i.$$

We now claim that $F(X) \le F(Y)$. Indeed, in view of [86], it suffices to show that X majorizes Y. By construction, we know that for the image of the OLP distinguished point,

$$X_{\omega(1)} \ge Y_{\omega(1)},$$

and for each $j = 2, 3, \ldots, m-1$,

$$X_{\omega(1)} + X_{\omega(2)} + \cdots + X_{\omega(j)} \ge Y_{\omega(1)} + Y_{\omega(2)} + \cdots + Y_{\omega(j)}$$

and finally,

$$X_{\omega(1)} + X_{\omega(2)} + \cdots + X_{\omega(m)} = Y_{\omega(1)} + Y_{\omega(2)} + \cdots + Y_{\omega(m)}.$$

Recalling the notation $X_\downarrow = \left(X_{[1]}, X_{[2]}, \ldots, X_{[m]}\right)$ for the components of X in decreasing order, it is immediate that

$$X_{[1]} + X_{[2]} + \cdots + X_{[j]} \ge X_{\omega(1)} + X_{\omega(2)} + \cdots + X_{\omega(j)}$$

for $j = 2, 3, \ldots, m$ and

$$X_{[1]} + X_{[2]} + \cdots + X_{[m]} = X_{\omega(1)} + X_{\omega(2)} + \cdots + X_{\omega(m)}.$$

Combining this with the previous inequalities involving $X_{\omega(i)}$ and $Y_{\omega(i)}$, we can conclude that X majorizes Y. Hence $F(X) \le F(Y)$ and the proof is complete.

6.8 Illustrative Example

To illustrate the application of the Ordered Loading Property Theorem in the previous section, we revisit the motivating example in Section 6.4 with

$$\begin{aligned} f(x) &= \log(14 + x_{1,1} + x_{2,1} + x_{4,1}) + \log(13 + x_{1,2} + x_{3,2} + x_{4,2}) \\ &\quad + \log(11 + x_{2,3} + x_{3,3} + x_{4,3}), \\ x &= (x_{1,1}, x_{1,2}, x_{2,1}, x_{2,3}, x_{3,2}, x_{3,3}, x_{4,1}, x_{4,2}, x_{4,3}) \end{aligned}$$

and constraint set $\mathfrak{X}$ described by

$$\begin{aligned} x_{1,1}+x_{1,2} &= 6; \\ x_{2,1}+x_{2,3} &= 5; \\ x_{3,2}+x_{3,3} &= 7; \\ x_{4,1}+x_{4,2}+x_{4,3} &= 9. \end{aligned}$$

The warehouses to which each supplier provides rice quantities are

$$I_1=\{1,2\}\,;\ I_2=\{1,3\}\,;\ I_3=\{2,3\}\,;\ I_4=\{1,2,3\}$$

while the suppliers from which warehouse can receive rice quantities are

$$J_1=\{1,2,4\}\,;\ J_2=\{1,3,4\}\,;\ J_3=\{2,3,4\}\,.$$

By the theorem, since $m=3$, $f(x)$ is minimized at one of $3!=6$ OLP points of $\mathfrak{X}_{OLP}$ which are given by

$$\begin{aligned} x^1 &= (6,0,5,0,7,0,9,0,0)\,; \\ x^2 &= (6,0,5,0,0,7,9,0,0)\,; \\ x^3 &= (0,6,5,0,7,0,0,9,0)\,; \\ x^4 &= (0,6,0,5,7,0,0,9,0)\,; \\ x^5 &= (6,0,0,5,0,7,0,0,9)\,; \\ x^6 &= (0,6,0,5,0,7,0,0,9)\,. \end{aligned}$$

The minimum occurs at $x=x^3$ with $\omega(1)=2, \omega(2)=1, \omega(3)=3$ and associated optimal value

$$f^*=f\left(x^3\right)=8.8977.$$

Recall from Section 6.4 that the optimal solution corresponds to Warehouse 1 holding 18 pounds of rice, Warehouse 2 holding 34 pounds of rice and Warehouse 3 holding 10 pounds of rice. The optimal allocation to each of the warehouses is determined by the bijection ω which requires Warehouse 2 to be allocated as much rice quantities from suppliers as possible, then Warehouse 1 to be allocated all the rice quantities from suppliers not taken by Warehouse 2 which leaves Warehouse 3 with only its pre-existing inventory. Warehouse 1 is allocated 5 pounds of rice from Supplier 2 and has a pre-existing inventory of 13 pounds of rice. Warehouse 2 is allocated 6 pounds of rice from Supplier 1, 7 pounds of rice from Supplier 3 and 9 pounds of rice from Supplier 4 and has a pre-existing inventory of 12 pounds of rice. Finally, Warehouse 3 has a pre-existing inventory of 10 pounds of rice and is not allocated any rice quantities from any of the suppliers.

Figure 6.3 shows a plot of the 24 possible allocations associated with the extreme points of $\mathfrak{X}_{ext}$. The OLP allocations are shown in black. The OLP allocation that corresponds to x^* is the black diamond. The allocations that do not correspond to the OLP points are shown in white.

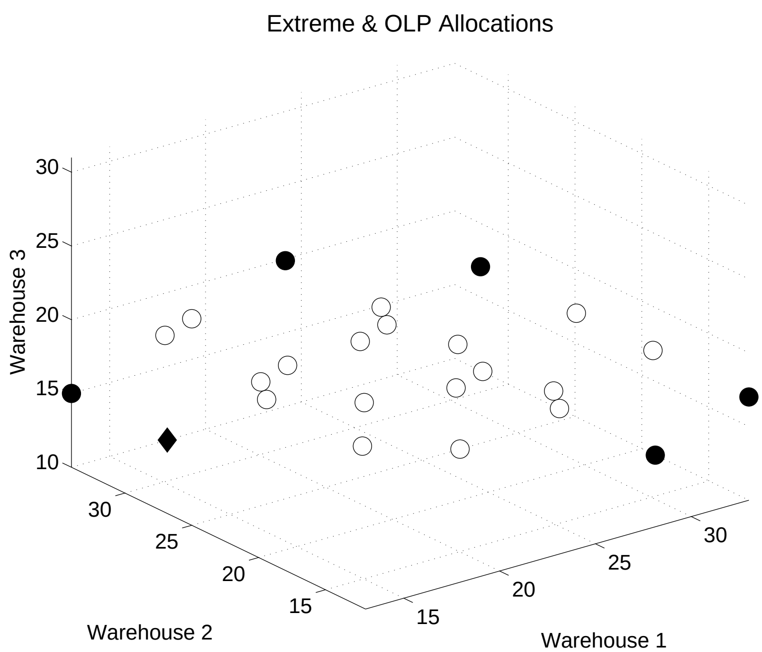

Figure 6.3 Graphical Depiction of Extreme Point and OLP Allocations

6.9 Larger-Scale Example

In this section, we consider an instance of the ICCP where $m = 8$ warehouses and $n = 2^8 - 1 = 255$ suppliers. We imagine our "divisible" commodity to be corn. The corn is measured in pounds. The collection of sets $I_1, I_2, \ldots, I_{255}$ indicate the warehouses from which each supplier can provide some quantity of corn. They are represented by the powerset of $\{1, 2, \ldots, 8\}$ excluding the empty set. The first $\binom{8}{1}$ sets $I_1, I_2, \ldots, I_8$ are given by

$$\{1\}, \{2\}, \ldots, \{8\}.$$

The next $\binom{8}{2}$ sets $I_9, I_{10}, \ldots, I_{36}$ are given by

$$\{1, 2\}, \{1, 3\}, \ldots, \{7, 8\}.$$

The next $\binom{8}{3}$ sets $I_{37}, I_{38}, \ldots, I_{92}$ are given by

$$\{1, 2, 3\}, \{1, 2, 4\}, \ldots, \{6, 7, 8\}.$$

The next $\binom{8}{4}$ sets $I_{93}, I_{94}, \ldots, I_{162}$ are given by

$$\{1, 2, 3, 4\}, \{1, 2, 3, 5\}, \ldots, \{5, 6, 7, 8\}.$$

The next $\binom{8}{5}$ sets $I_{163}, I_{164}, \ldots, I_{218}$ are given by

$$\{1, 2, 3, 4, 5\}, \{1, 2, 3, 4, 6\}, \ldots, \{4, 5, 6, 7, 8\}.$$

The next $\binom{8}{6}$ sets $I_{219}, I_{220}, \ldots, I_{246}$ are given by

$$\{1, 2, 3, 4, 5, 6\}, \{1, 2, 3, 4, 5, 7\}, \ldots, \{3, 4, 5, 6, 7, 8\}.$$

The next $\binom{8}{7}$ sets $I_{247}, I_{248}, \ldots, I_{254}$ are given by

$$\{1, 2, 3, 4, 5, 6, 7\}, \{1, 2, 3, 4, 5, 6, 8\}, \ldots, \{2, 3, 4, 5, 6, 7, 8\}.$$

Finally, the last set $I_{255} = \{1, 2, 3, 4, 5, 6, 7, 8\}$. Now the suppliers from which each warehouse can receive quantities of corn from are the J_i derived from the I_i above. We take the allocation vector x to be

$$x = (x_{1,1}, x_{2,2}, \ldots, x_{8,8}, \ldots, x_{255,1}, x_{255,2}, \ldots, x_{255,8}).$$

Each component of x has the form $x_{i,j}$ where i corresponds to I_i and j is an element of I_i. The components of x are ordered according to the elements in the sets $I_1, I_2, \ldots, I_{255}$. The quantity of corn r_i (in pounds) that each supplier has is taken from a uniform distribution of integers ranging from 10 to 40. Letting $r(i:j)$ denote the components of the vector r from r_i to r_j we have

$$r(1:5)=10; r(6:12)=11; r(13:20)=12; r(21:29)=13; r(30:39)=14; r(40:51)=15;$$

$$r(52:56)=16;\ r(57:71)=17;\ r(72:81)=18;\ r(82:89)=19;\ r(90:97)=20;$$

$$r(98:102)=21;\ r(103:109)=22;\ r(110:122)=23;\ r(123:126)=24;\ r(127:139)=25;$$

$$r(140:148)=26;\ r(149:156)=27;\ r(157:162)=28;\ r(163:168)=29;\ r(169:175)=30;$$

$$r(176:183)=31;\ r(184:186)=32;\ r(187:195)=33;\ r(196:207)=34;\ r(208:216)=35;$$

$$r(217:222)=36;\ r(223:226)=37;\ r(227:239)=38;\ r(240:252)=39;\ r(253:255)=40.$$

The pre-existing inventory of the warehouses was taken from a uniform distribution of numbers ranging from 25 to 50 pounds. We obtained

$$(x_{0,1}, x_{0,2}, \ldots, x_{0,8}) \quad = \quad (41, 42, 35, 34, 50, 25, 48, 48)\,.$$

In this example, due to economies of scale (see [75]), we use the carrying cost function

$$g(x) = \log(1+x)$$

and therefore

$$f(x) \quad = \quad \sum_{j=1}^{8} \log\left(1 + X_j(x)\right).$$

From the OLP Theorem, we know that the optimal allocation is found among the $8! = 40,320$ OLP points whereas the set $\mathfrak{X}$ has 7.6959×10^{145} extreme points! For this data set, we find that the optimal allocation assigns 58 pounds of corn to Warehouse 1, 69 pounds of corn to Warehouse 2, 3592 pounds of corn to Warehouse 3, 170 pounds of corn to Warehouse 4, 72 pounds of corn to Warehouse 5, 730 pounds of corn to Warehouse 6, 331 pounds of corn to Warehouse 7 and finally, 1618 pounds of corn to Warehouse 8. Now $f^* = 19.8620$ and the bijection associated with the optimum OLP allocation is given by

$$\omega(1)=3;\ \omega(2)=8;\ \omega(3)=6;\ \omega(4)=7;\ \omega(5)=4;\ \omega(6)=5;\ \omega(7)=2;\ \omega(8)=1.$$

6.10 Further Research

In this chapter, we studied an Inventory Carrying Cost Problem, denoted as ICCP, which involves the distribution of a resource between n suppliers and m identical warehouses. We showed that the optimum allocation for this problem is found among the set of $m!$ OLP points. As an extension of our current research, we consider the case where each warehouse's carrying cost function is different. More specifically, we plan to consider a modified version of the ICCP where the objective function is given by

$$f(x) = \sum_{j=1}^{m} f_j(x)$$

where

$$f_j(x) = g_j(X_j(x))$$

and each g_j is a concave function. In this case, the Schur concavity argument in our proof of optimality is no longer valid since each warehouse's carrying cost function is different. Nevertheless, we conjecture that it may still be the case that the set of OLP extreme points contain a solution to these problems. To illustrate this, consider the motivational example from Section 6.4, but with the objective function replaced by

$$\begin{aligned} f(x) &= \log(14 + x_{1,1} + x_{2,1} + x_{4,1}) + (12 + x_{1,2} + x_{3,2} + x_{4,2})^{1/2} \\ &\quad + \log(1 + 2(10 + x_{2,3} + x_{3,3} + x_{4,3})); \\ x &= (x_{1,1}, x_{1,2}, x_{2,1}, x_{2,3}, x_{3,2}, x_{3,3}, x_{4,1}, x_{4,2}, x_{4,3}), \end{aligned}$$

where $g_1(x) = \log(1+x)$, $g_2(x) = \sqrt{x}$ and $g_3(x) = \log(1+2x)$. Despite each warehouse having a different carrying cost function, a minimum still occurs at the OLP point

$$x^* = (6, 0, 5, 0, 0, 7, 9, 0, 0)$$

with $f(x^*) = 10.5458$. The optimal solution results in 33 pounds of rice allocated to Warehouse 1, 12 pounds of rice allocated to Warehouse 2 and 17 pounds of rice to Warehouse 3.

Chapter 7

Conclusion and Future Research

The main theme of this dissertation was the exploitation of symmetry for systems problems involving multilinear uncertainty structures. To this end, we considered various optimization and probability scenarios and demonstrated how symmetry is used to reduce computational complexity. A few of our theorems are more general and do not require multilinearity; only symmetry is exploited in these cases. This was the case in Chapter 4 where we considered the probability mass function of a general symmetric function of i.i.d random variables. We also provided a number of theorems involving optimization and probability for sums of symmetric functions and multi-group symmetric functions. One highlight of this dissertation is the extreme point result obtained via Schur concavity for the Inventory Carrying Cost Problem described in Chapter 6. We now present three avenues for future research.

7.1 Future Work on Resource Allocation

A first direction for future research involves a more general allocation problem in Chapter 6. One important extension would be to consider the geographical location of suppliers as optimization variables and take into account the distances to warehouses. This extended formulation will be similar in flavor to position location problems which are studied in optimal facility location problems in [80]. Facility location problems are often formulated as Minimum Concave-Cost Network Flow Problems (MCCNFP) which involve directed graphs; see [84] and [85].

7.1.1 Supplier Positioning Problem: Building on the setup in Chapter 6 we introduce a pair of variables (u_i, v_i) which denotes the position of supplier i in the u-v plane. Given all the positions of the warehouses, we would like to make an optimal selection of these variables using an augmented objective function which not only includes the distances mentioned above but also accounts for the amount of resource delivered by the suppliers; i.e., the augmented objective function should also reflect larger good shipments being more costly. This leads us to consider a function of the form

$$d(u, v) = \sum_{i=1}^{n} \sum_{j \in I_i} d_{i,j}\left(u_i, v_i, x_{i,j}\right)$$

where, as in Chapter 6, I_i are the set of indices of warehouses to which supplier i can supply some portion of r_i and $x_{i,j}$ is the amount of resource delivered from supplier i to warehouse j . We also plan to address the extent to which concavity and extremality can be brought into play and how difficult the problem becomes as we increase the number of warehouses. In other words, we need to determine how complex the problem becomes as m increases.

7.2 Future Work on Multi-Group Symmetry

For future research regarding multi-group functions and sets, it would be of interest to investigate computational complexity when only a subset of variables enter into $f(x)$ either symmetrically or in a groupwise manner. To motivate this topic, consider the pair $(f, \mathfrak{X})$ where $f\colon \mathbb{R}^5 \to \mathbb{R}$ is the multilinear function

$$f(x) = x_1x_2 + x_1x_3 + x_2x_3 \;+\; 2\left(x_4 + x_1x_2x_3x_4x_5\right)$$

and $\mathfrak{X} \subset \mathbb{R}^5$ is given by the product of hyperrectangles $\mathfrak{X}_1 \times \mathfrak{X}_2$ where $\mathfrak{X}_1 = [1,2] \times [1,2] \times [1,2]$ and $\mathfrak{X}_2 = [2,3] \times [2,3]$ with partition $I_1 = \{1,2,3\}$, $I_2 = \{4,5\}$. If we hold $x^{[2]}$ fixed, then the resulting function is symmetric in (x_1, x_2, x_3) and we only need to consider 4 extreme points to find its minimum or maximum on $\mathfrak{X}_1$. On the other hand, if we hold $x^{[1]}$ fixed, the resulting function is not symmetric in (x_4, x_5) and we need to consider all 4 extreme points to find its minimum or maximum on $\mathfrak{X}_2$. Therefore, to find a minimum or maximum of f on $\mathfrak{X}$, it suffices to consider $(4)(4) = 16$ extreme points as opposed to the $2^5 = 32$ for a general multilinear function.

One direction for future research involves generalizing results in this thesis to allow for subsets of group variables being symmetric or groupwise affine. To this end we will be working with "subset-group asymmetric" and "subset-groupwise affine symmetric" pairs. It should be possible to develop theorems characterizing extreme point savings when working with such pairs. One special case of interest for a "subset-group asymmetric pair" is obtained when $f(x)$ is multilinear and $\mathfrak{X}$ is the product of hyperrectangles that are not all identical. A "subset-groupwise affine symmetric pair" $(f, \mathfrak{X})$ would be the case which arises when $f(x)$ is groupwise and $\mathfrak{X}$ is the product of polytopes that are not all identical.

7.3 Future Research on Symmetric Multilinear Robust Stability Problems

In systems theory, multilinear functions are heavily used. Hence this section on future research is much more detailed than its predecessors. In particular, we focus on robust stability and related mathematical programming issues. Indeed, there is a large body of literature devoted to the study of the stability of the *family of polynomials* $\mathcal{P} \doteq \{p(\cdot, x) : x \in \mathfrak{X}\}$ with each member $p(\cdot, x)$ having the form

$$p(s, x) \doteq \sum_{i=0}^{m} a_i(x) s^i.$$

Typically, $p(s, x)$ is called an *uncertain polynomial*, called the vector of *uncertain parameters* and $\mathfrak{X}$ is called the *uncertainty bounding set*; e.g., see [3]. The polynomial above is typically assumed to have continuous coefficient functions $a_0(x), a_1(x), \ldots, a_m(x)$ and $\mathfrak{X} \subset \mathbb{R}^n$ is a hyperrectangle

$$\mathfrak{X} \doteq \left\{ x \in \mathbb{R}^n : x_i^- \leq x_i \leq x_i^+ \text{ for } i = 1, 2, \ldots, n \right\},$$

where ${x_i}^-$ and ${x_i}^+$ are given bounds for the i-th component x_i of x. Now, this family $\mathcal{P}$ is said to be *robustly stable* if, for all $x \in \mathfrak{X}$, $p(s, x)$ is stable; that is, all roots of $p(s, x)$ lie strictly in the left half plane; see [3] and its extensive bibliography for full coverage of this topic.

7.3.1 Robust Stability with Symmetry in Play: Many papers such as [4] and [5] consider the robust stability problem of the family of polynomials above with the coefficients $a_i(x)$ being multilinear functions. For example, this structure arises in a cascade connection of Single-Input-Single-Output transfer functions; see also [7]-[9], [10]-[14], [15]-[24] and [25]-[30] where problems of stability with multilinear dependence arise. Of particular note is reference [31], where the robust stability problem is considered with coefficients $a_i(x)$, which are symmetric multilinear functions. When the uncertainty bounding set $\mathfrak{X}$ satisfies the *nonoverlapping* condition

$$x_1^+ > x_1^- > x_2^+ > x_2^- > \cdots > x_n^+ > x_n^-,$$

it turns out one can test for the robust stability by using the so-called Edge Theorem in [32]. This theorem reduces the robust stability problem to testing the subset of polynomials associated with the one-dimensional edges of $\mathfrak{X}$.

7.3.2 Further Research on Robust Stability: To motivate how symmetry can be exploited in new settings not considered to date in the literature, we now consider the linear time-invariant state-space system

$$\dot{y} \;=\; \begin{bmatrix} x_1 & 1 & 2 \\ 12 & x_2 & 3 \\ 6 & 4 & x_3 \end{bmatrix} y \;\doteq\; A(x)y.$$

The uncertain characteristic polynomial of this system is easily calculated to be

$$\begin{aligned} p(s,x) &= \det\left(sI - A(x)\right) \\ &= s^3 \;-\; (x_1 + x_2 + x_3)\, s^2 \;+\; (x_1x_2 + x_1x_3 + x_2x_3 - 36)\, s \\ &\qquad + 12\,(x_1 + x_2 + x_3) - x_1x_2x_3 - 114 \end{aligned}$$

and it is noted that the coefficients of this polynomial are all symmetric multilinear functions of x.

To see the benefit of the symmetry above in a computational complexity context, we consider the simple case where each x_i can take on one of two possible values, a or b. That is, $x_i \in \{a, b\}$ for

$i = 1, 2, 3$. To determine if this system is robustly stable, it is necessary and sufficient to find the roots of the polynomials associated with the eight possibilities for x and check if they all lie in the strict left half plane. These eight possibilities are given by

$$x^1 = (a, a, a)\,; \quad x^2 = (a, a, b)\,; \quad x^3 = (a, b, a)\,; \quad x^4 = (a, b, b)\,;$$
$$x^5 = (b, a, a)\,; \quad x^6 = (b, a, b)\,; \quad x^7 = (b, b, a)\,; \quad x^8 = (b, b, b)$$

and the polynomials are readily generated. For example,

$$p(s, x^4) = s^3 \; - \; (a + 2b)\, s^2 \; + \; \left(b^2 + 2ab - 36\right) s \; + \; 12a + 24b - ab^2 - 114.$$

Notice that as a result of the symmetry of $p(s, x)$, we have

$$p(s, x^2) = p(s, x^3) = p(s, x^5);$$
$$p(s, x^4) = p(s, x^6) = p(s, x^7).$$

Therefore in view of symmetry, to determine if $p(s, x)$ is robustly stable we only need to check the polynomials associated with the extreme points, $p(s, x^1)$, $p(s, x^2)$, $p(s, x^4)$, and $p(s, x^8)$, instead of the original eight. This example, while very simple, illustrates the reduction in computational complexity that can be attained with symmetric multilinear functions. More generally, if the matrix $A(x)$ above has n uncertain parameters $x_1, x_2, \ldots, x_n$, each taking on values a or b; then we have 2^n extreme points. To determine if $p(s, x)$ is robustly stable, using symmetry, we only need to check the polynomials associated with $n + 1$ of the extreme points, instead of the original 2^n.

7.3.3 Symmetry Related to Hurwitz Matrices: One interesting new class of optimization problems which one could consider involves the so-called Hurwitz matrix which arises in the robust stability problems above. To provide motivation, we revisit the uncertain polynomial

$$\begin{aligned} p(s, x) \;\; &= \;\; s^3 \; - \; (x_1 + x_2 + x_3 - 4)\, s^2 \; + \; (x_1 x_2 + x_1 x_3 + x_2 x_3 - 36)\, s \\ &\quad + \; 12\,(x_1 + x_2 + x_3) - x_1 x_2 x_3 - 114, \end{aligned}$$

considered above with symmetric coefficient functions and uncertainty bounding set $\mathfrak{X}$ described by $0 \leq x_i \leq 1$ for $i = 1, 2, 3$. We now reformulate the robust stability problem in terms of optimization theory, by forming the Hurwitz matrix

$$\mathcal{H}(x) = \begin{bmatrix} a_1(x) & a_3(x) & 0 \\ a_0(x) & a_2(x) & 0 \\ 0 & a_1(x) & a_3(x) \end{bmatrix}$$

for $p(s,x)$ where the entries above are the coefficients $a_i(x)$ of $p(s,x)$; i.e.,

$$\begin{aligned} a_0(x) &= 12\,(x_1 + x_2 + x_3) - x_1 x_2 x_3 - 114; \\ a_1(x) &= x_1 x_2 + x_1 x_3 + x_2 x_3 - 36; \\ a_2(x) &= 4 - x_1 - x_2 - x_3; \\ a_3(x) &= 1. \end{aligned}$$

Now, noting that the Hurwitz stability criterion requires the leading principle minors of $\mathcal{H}(x)$ to all be positive, the necessary and sufficient conditions for robust stability are

$$a_1(x) > 0;$$
$$a_1(x)a_2(x) - a_0(x) > 0$$

for all $x \in \mathfrak{X}$. Equivalently, robust stability is assured if the following two optimality conditions are satisfied:

$$\min_{x \in \mathfrak{X}} \; a_1(x) > 0;$$
$$\min_{x \in \mathfrak{X}} \; [a_1(x)a_2(x) - a_0(x)] > 0.$$

We notice that the second Hurwitz minor, $\det \mathcal{H}(x)$, above is symmetric. That is,

$$\begin{aligned} \det \mathcal{H}(x) &= a_1(x)a_2(x) - a_0(x) \\ &= \left(x_1^2 x_2 + x_1 x_2^2 + x_1^2 x_3 + x_1 x_3^2 + x_2^2 x_3 + x_2 x_3^2\right) \; + \; 4\,(x_1 x_2 + x_1 x_3 + x_2 x_3) \\ &\qquad - \; 2\,(x_1 x_2 x_3) \; + \; 24\,(x_1 + x_2 + x_3) \; - \; 30. \end{aligned}$$

The symmetry in the coefficients $a_i(x)$ is preserved in the minors of the Hurwitz matrix. One could study this in greater detail for future work. This "symmetry preservation property" may be useful to solve new classes of robust stability problems.

LIST OF REFERENCES

[1] I. G. Macdonald, *Symmetric Functions and Hall Polynomials*, Clarendon Press, Oxford, 1979.

[2] D. Terr and E. Weisstein, “Symmetric Polynomial.”, From MathWorld–A Wolfram Web Resource, http://mathworld.wolfram.com/SymmetricPolynomial.html, 2011.

[3] B. R. Barmish, *New Tools for Robustness of Linear Systems*, Macmillan, New York, 1994.

[4] B. R. Barmish and Z. Shi, “Robust stability of a class of polynomials with coefficients depending multilinearly on perturbations,” *IEEE Transactions on Automatic Control*, vol. AC-35, pp. 1040-1043, 1990.

[5] H. Chapellat, L. H. Keels, and S. P. Bhattacharyya, “Extremal robustness properties of multilinear interval systems,” *Automatica*, vol. 30, pp. 1037-1042, 1994.

[6] V. L. Kharitonov, “Asymptotic stability of an equilibrium position of a family of systems of linear differential equations,” *Differential Equations*, vol. 14, pp. 1483-1485, 1979.

[7] E. R. Panier, M. K. H. Fan, and A. L. Tits, “On the stability of polynomials with no cross-coupling between the perturbations in the coefficients of even and odd powers,” *Systems & Control Letters*, vol. 12, pp. 291–299, 1989.

[8] T. E. Djaferis and C. V. Hollott, “Partitioning via shaping conditions for the stability of families of polynomials,” *Proceedings of the American Control Conference*, 1988.

[9] H. Chapellat, M. Dahleh, and S. P. Bhattacharyya, “Robust stability manifolds for multilinear interval systems,” *IEEE Transactions on Automatic Control*, vol. AC-38, pp. 314-318, 1993.

[10] H. Chapellat and S. P. Bhattacharyya, “A generalization of Kharitonovs theorem: Robust stability of interval plants,” *IEEE Transactions on Automatic Control*, vol. AC-34, pp. 306-311, 1989.

[11] J. D. Cobb and C. L. DeMarco, “The minimal dimension of stable faces required to guarantee stability of a matrix polytope,” *IEEE Transactions on Automatic Control*, vol. AC-34, pp. 990–992, 1989.

[12] M. Mansour, "Robust stability of interval matrices," *Proceedings of the 28th IEEE Conference on Decision and Control*, pp. 46–51, 1989.

[13] A. Sideris and R. R. E. de Gaston, "Multivariable stability margin calculation with uncertain correlated parameters," *Proceedings of the IEEE Conference on Decision and Control*, 1986.

[14] A. Sideris and R. S. Sänchez Peña, "Fast computation of the multivariable stability margin for real interrelated uncertain parameters," *Proceedings of the American Control Conference*, 1988.

[15] N. Tsing and A. Tits, "On the multiaffine image of a cube," in *Robustness of Dynamic Systems with Parameter Uncertainties*, pp. 105–110, Birkhauser, Boston, 1992.

[16] I. R. Petersen, "A collection of results on the stability of families of polynomials with multilinear parameter dependence," Department of Electrical Engineering, University of New South Wales, Australian Defence Force Academy, Technical Report EE8801, 1988.

[17] C. V. Hollot and Z. L. Xu. "When is the image of a multilinear function a polytope?—A conjecture," *Proceedings of the 28th IEEE Conference on Decision and Control*, pp. 1890–1891, 1989.

[18] F. J. Kraus, B. D. 0. Anderson, and M. Mansour, "Robust stability of polynomials with multilinear parameter dependence," *International Journal of Control*, vol. 50, pp. 1745–1762, 1989.

[19] H. Chapellat, L. H. Keel, and S. P. Bhattacharyya, "Robustness properties of multilinear interval systems," in *Robustness of Dynamic Systems with Parameter Uncertainties*, pp. 73–80, Birkhauser, Boston, 1992.

[20] B. T. Polyak, "Robustness analysis for multilinear perturbations," *Robustness of Dynamic Systems with Parameter Uncertainties*, pp. 93–104, Birkhauser, Boston, 1992.

[21] Y. Tian, C. Feng, and X. Xin, "Robust stability of polynomials with multilinearly dependent coefficient perturbations," *IEEE Transactions on Automatic Control*, vol. AC-39, pp. 554–558, 1994.

[22] Y. C. Soh, R. J. Evans, I. R. Petersen and R. J. Betz, "Robust pole assignment," *Automatica*, vol. 23, pp. 601–610, 1987.

[23] P. Padmanabhan, C. V. Hollot, and W. Siegl, "Robust pole placement using two pole dominant models," *Proceedings of the 12th IFAC World Congress*, Sydney, Australia, 1993.

[24] W. Chen and I. Petersen, "An easily testable sufficient condition for the robust stability of multilinear uncertain polynomials," *Proceedings of the European Control Conference*, 1997.

[25] T. E. Djaferis, "Shaping conditions for the robust stability of polynomials with multilinear parameter uncertainty," *Proceedings of the Conference on Decision and Control*, 1988.

[26] S. Dasgupta, "Perspectives on Kharitonov's Theorem: A view from the imaginary axis," Technical Report, Department of Electrical and Computer Engineering, University of Iowa, Iowa City, 1987.

[27] M. B. Argoun, "Stability of Hurwitz polynomials under coefficient perturbations: necessary and sufficient Conditions," *International Journal of Control*, vol. 45, pp. 739-744, 1987.

[28] M. Saeki, "A method of robust stability analysis with highly structured uncertainties," *IEEE Transactions on Automatic Control*, vol. AC-31, pp. 935-940, 1986.

[29] T. E. Djaferis and C. V. Hollot, "Characterization of the Hurwitz Region for Systems with Parametric Uncertainty," *Proceeding of the American Control Conference*, 1988.

[30] R.R.E. de Gaston and M. G. Safanov, "Exact Calculation of the Multiloop Stability Margin," *IEEE Transactions on Automatic Control*, vol. AC-33, 1988.

[31] W. Chen and I. R. Petersen, "A class of multilinear uncertain polynomials to which the edge theorem is applicable," *IEEE Transactions on Automatic Control*, vol. AC-40, pp. 2103-2107, 1995.

[32] A. C. Bartlett, C. V. Hollot, and H. Lin, "Root locations of an entire polytope of polynomials: It suffices to check the edges," *Mathematics of Control, Signals and Systems*, vol. 1, pp. 61–71, 1988.

[33] L. A. Zadeh and C. A. Desoer, *Linear System Theory*, McGraw-Hill, New York, 1963.

[34] R. Tempo, M. Barberis, M. Casales, and D. Cavallera, "Robust stability with multilinear perturbations: How good is the convex hull approximation?," *Proceedings of the IEEE Conference on Decision and Control*, 1990.

[35] B. D. O. Anderson, F. Kraus, M. Mansour, and S. Dasgupta, "Easily testable sufficient conditions for the robust stability of systems with multilinear parameter dependence," *Automatica*, pp. 25-40, 1995.

[36] S. R. Ross and B. R. Barmish, "A worst-case estimate of stability probability for polynomials with multilinear uncertainty structure," *Proceedings of the IEEE Conference on Decision and Control*, 2000.

[37] B. R. Barmish, "A probabilistic robustness result for a multilinearly parameterized H_∞ norm," *Proceedings of the American Control Conference*, 2000.

[38] S. R. Ross, "Contributions to the Theory of Robustness of Systems with Multilinear Uncertainty Structure," PhD dissertation, Department of Electrical and Computer Engineering, University of Wisconsin-Madison, 2006.

[39] B. R. Barmish, P. S. Shcherbakov, S. R. Ross, and F. Dabbene, "On positivity of polynomials: The dilation integral method," *IEEE Transactions on Automatic Control*, vol. AC-54, pp. 965–978, 2009.

[40] D. Henrion and A. Garulli, *Positive Polynomials in Control*, Springer, Berlin, 2005.

[41] R. Horst and H. Tuy, *Global Optimization: Deterministic Approaches*, Springer-Verlag, Berlin, 1996.

[42] H. D. Sherali and W. P. Adams, *Reformulation-Linearization Techniques in Discrete and Continuous Optimization*, Nonconvex Optimization and Its Applications, Kluwer Academic Publishers, London, 1999.

[43] H. S. Ryoo and N. V. Sahinidis, "Analysis of bounds for multilinear functions," *Journal of Global Optimization*, vol. 19, pp. 403-424, 2001.

[44] A. S. E. Hamed, *Calculation of Bounds on Variables and Underestimating Convex Functions for Nonconvex Functions*, PhD dissertation, The George Washington University, 1991.

[45] M. Vidyasagar, "Minimum-seeking properties of analog neural networks with multilinear objective functions," *IEEE Transactions on Automatic Control*, vol. AC-40, pp. 1359-1375, 1995.

[46] S. Skelboe,"True Worst-Case Analysis of Linear Electrical Circuits by Interval Arithmetic," *IEEE Transactions on Circuits and Systems*, vol. 26, pp. 874–878, 1979.

[47] K. Singhal and J. Vlach, "Symbolic analysis of analog and digital circuits," *Proceedings of the IEEE International Symposium on Circuits and Systems*, 1976.

[48] D. Elliot, *Bilinear Control Systems Matrices in Action*, Springer, New York, 2009.

[49] V. Barnett and T. Lewis, *Outliers on Statistical Data*, John Wiley & Sons, New York, 1978.

[50] A. C. Rencher, *Methods of Multivariate Analysis. 2nd. ed.*, John Wiley & Sons, New Jersey, 2002.

[51] S. S. Wilks, "Multivariate Statistical Outliers," *Sankhya, Series A*, vol 25, pp. 407–426, 1963.

[52] S. S. Yang and Y. Lee, "Identification of a Multivariate Outlier," *Presented at the Annual Meeting of the American Statistical Association*, San Francisco, 1987.

[53] B. R. Barmish, "On performance limits of feedback control-based stock strategies," *Proceedings of the American Control Conference*, pp. 3874–3879, San Francisco, July 2011.

[54] B. R. Barmish, "On Trading of Equities: A Robust Control Paradigm," *Proceedings IFAC World Congress*, pp. 1621–1626, Seoul, Korea, July 2008.

[55] B. R. Barmish and J. A. Primbs, "On Arbitrage Possibilities Via Linear Feedback in an Idealized Brownian Motion Stock Market," *Proc. IEEE Conference on Decision and Control*, pp. 2889–2894, Orlando, December 2011.

[56] S. Malekpour and B. R. Barmish, "How Useful are Mean-Variance Considerations in Stock Trading via Feedback Control?" *Proceedings of the IEEE Conference on Decision and Control*, pp. 2110–2115, Maui, December 2012.

[57] S. Iwarere and B. R. Barmish, "A Confidence Interval Triggering Method for Stock Trading via Feedback Control," *Proceedings of the American Control Conference*, pp. 6910–6916, Baltimore, July 2010.

[58] S. Iwarere and B. R. Barmish, "On Stock Trading Over a Lattice via Linear Feedback," *Proceedings IFAC World Congress*, Cape Town, South Africa, August 2014.

[59] J. Cox, S. Ross and M. Rubinstein. "Option Pricing: A Simplified Approach," *Journal of Financial Economics*, vol. 7 pp. 229–264, 1979.

[60] P. Boyle, "A Lattice Framework for Option Pricing with Two State Variables," *Journal of Financial and Quantitative Analysis*, vol. 23, pp. 1-12, 1988.

[61] A. Gamba and L. Trigeorgis, "An Improved Binomial Lattice Method for Multi-Dimensional Options," *Applied Mathematical Finance*, vol. 14, pp. 453-475, 2007.

[62] J. A. Prmbs, M. Rathinam and Y. Yamada, "Option Pricing with a Pentanomial Lattice Model that Incorporates Skewness and Kurtosis," *Applied Mathematical Finance*, vol. 14, pp. 1–17, 2007.

[63] D. G. Luenberger, *Investment Science*, Oxford University Press, New York, 1998.

[64] B. E. Sagan, *The Symmetric Group Representations, Combinatorial Algorithms, & Symmetric Functions*, Brooks/Cole Publishing Company, California, 1991.

[65] L. Manivel, *Symmetric Functions, Schubert Polynomials and Degeneracy Loci*, American Mathematical Society, Rhode Island, 2001.

[66] W. Feller, *An Introduction to Probability and Its Applications*, Wiley, New York, 1968.

[67] R. T. Rockafellar, *Convex Analysis*, Princeton University Press, Princeton, 1970.

[68] T. Ibaraki and N. Katoh, *Resource Allocation Problems, Algorithmic Approaches*, The MIT Press, Cambridge, MA, 1988.

[69] M. Patriksson, "A Survey on the Continuous Nonlinear Resource Allocation Problem," *European Journal of Operational Research*, vol. 185, pp. 1-46, 2008.

[70] A. Zubow, J. Marotzke, D. Camps-Mur, X. P. Costa, "sGSA: A SDMA-OFDMA Greedy Scheduling Algorithm for WiMAX Networks," *Computer Networks*, vol. 56, pp. 3511–3530, 2012.

[71] B. B. M. Shao, and H. R. Rao, "A Comparative Analysis of Information Acquisition Mechanisms for Discrete Resource Allocation," *IEEE Transactions on Systems, Man, and Cybernetics, Part A: Systems and Humans*, vol. 31, pp. 199–209, May 2001.

[72] J. F. Kurose and R. Simha,"A Microeconomic Approach to Optimal Resource Allocation in Distributed Computer Systems," *IEEE Transactions on Computers*, vol. 38, pp. 705-717, May 1989.

[73] G. R. Bitran and A. C. Hax, "Dissagregation and Resource Allocation Using Convex Knapsack Problems with Bounded Variables," *Management Science*, vol. 27, pp. 431-441, 1981.

[74] A. Balakrishnan, R. L. Francis, and S. J. Grotzinger, "Bottleneck Resource Allocation in Manufacturing" *Management Science*, vol. 42, pp. 1611-1625, 1996.

[75] R. Horst and P. Pardalos, *Handbook of Global Optimization*, Kluwer Academic Publishers, Dordrecht, Netherlands, 1995.

[76] C. D. Heising-Goodman, "A Survey of Methodology for the Global Minimization of Concave Functions Subject to Convex Constraints," *International Journal of Management Science*, vol. 9, pp. 313-319, 1981.

[77] J. Bracken, G. P. McCormick, *Selected Applications of Nonlinear Programming*, John Wiley and Sons, 1968.

[78] H. A. Taha, "Concave Minimization Over a Convex Polyhedron," *Naval Research Logistics Quarterly*, vol. 20, pp. 533–548, 1973.

[79] V. A. Cabot, "Variations on a Cutting Plane Method for Solving Concave Minimization Problems with Linear Constraints," *Naval Research Logistics Quarterly*, vol. 21, pp. 265–274, 1974.

[80] R. M. Soland, "Optimal Facility Location with Concave Costs," *Operations Research*, vol. 22, pp. 373–382, 1974.

[81] V. I. Manousiouthakis, N. Thomas and A. M. Justanieah, "On a Finite Branch and Bound Algorithm for the Global Minimization of a Concave Power Law Over a Polytope," *Journal of Optimization Theory and Applications*, vol. 151, pp. 121–134, October 2011.

[82] A. F. Veinott, "Minimum Concave Cost Solution of Leontief Substitution Models of Multi-Facility Inventory Systems," *Operations Research*, vol. 17, pp. 262–291, 1969.

[83] H. Tuy, S. Ghannadan, A. Migdalas and P. Varbrand, "Strongly Polynomial Algorithm for a Production-Transportation Problem with Conave Production Costs," Research Report, Department of Mathematics, Linkoping Institute of Technology, Linkoping, Sweden, 1992.

[84] H. Tuy, S. Ghannadan, A. Migdalas and P. Varbrand, "Minimum Concave Cost Network Flow Problems with a Fixed Number of Nonlinear Arc Costs and Sources," Research Report, Department of Mathematics, Linkoping Institute of Technology, Linkoping, Sweden, 1992.

[85] G. M. Guisewite and P. M. Pardalos, "Minimum Concave-cost Network Flow Problems: Applications, Complexity, and Algorithms," *Operations Research*, vol. 25, pp. 75–100, November 1990.

[86] A. W. Marshall and I. Olkin, *Inequalities: Theory of Majorization and its Applications*, Academic Press, New York, 1979.