MATRIX RANK MINIMIZATION WITH APPLICATIONS

A DISSERTATION

SUBMITTED TO THE DEPARTMENT OF ELECTRICAL ENGINEERING

AND THE COMMITTEE ON GRADUATE STUDIES

OF STANFORD UNIVERSITY

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

Maryam Fazel

March 2002

I certify that I have read this dissertation and that in my opinion it is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Professor Stephen P. Boyd
(Principal Adviser)

I certify that I have read this dissertation and that in my opinion it is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Professor Michael A. Saunders

I certify that I have read this dissertation and that in my opinion it is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Professor John T. Gill III

Approved for the University Committee on Graduate Studies:

Abstract

We consider the problem of minimizing the rank of a matrix over a convex set. The Rank Minimization Problem (RMP) arises in diverse areas such as control, system identification, statistics, signal processing, and combinatorial optimization, and is known to be computationally NP-hard. As a special case, it includes the problem of finding the sparsest vector in a convex set.

In this dissertation, we propose two heuristics based on convex optimization that approximately solve the RMP. We refer to them as the *trace/nuclear norm* and the *log-det* heuristics. Unlike the existing methods, these heuristics can handle any general matrix, are numerically very efficient, do not require a user-specified initial point, and yield a global lower bound on the RMP if the feasible set is bounded. We show that the nuclear norm heuristic is optimal in the sense that it minimizes the *convex envelope* of the rank function, thus providing theoretical support for its use. In the special case of finding sparse vectors, these heuristics reduce to the ℓ_1-norm and iterative ℓ_1-norm minimization methods.

We catalog many practical applications of the RMP. By giving numerical examples of problems from different fields, we demonstrate that the proposed heuristics work very well in practice.

Acknowledgements

First and foremost, I would like to thank my adviser Professor Stephen Boyd. His wealth of ideas, clarity of thought, enthusiasm and energy have made working with him an exceptional experience for me. He is also a masterful teacher. Attending a guest lecture he gave at one of my very first classes at Stanford was an inspiring experience that influenced the path of my studies and research. I feel very fortunate to have had him as an adviser and a teacher.

I am thankful to Professor Michael Saunders for being on my defense and reading committees, for many helpful discussions, and for his remarkably kind and caring attitude. I thank Professor John Gill for kindly accepting to be on my reading committee on short notice, and for providing much valuable feedback. I would also like to thank Professor Claire Tomlin for acting as the chair of my defense committee, and Professor David Donoho for being a member of the committee and for his thought-provoking questions and comments on my work.

I am very grateful to Professor Tom Kailath for his excellent academic guidance during my first two years, and for his wisdom and advice on how to embark on academic research. I also thank him for his generous financial support that enabled me to come to Stanford.

It has been a pleasure to be a part of ISL, where I have made many great friends. My deepest thanks go to my former and current officemates in Durand 110 and

Packard 243. I would like to especially acknowledge Haitham Hindi, for his collaborations on this research and co-authorship of several papers, and also for his friendship and constant support, Miguel Lobo for joint work on the portfolio optimization problem and for teaching me about finance and other topics, Cesar Crusius for being the computer guru of our group, Bob Lorenz for his always helpful advice, and his thorough reading of the draft of this thesis, Lin Xiao for many fruitful discussions, and Shao-Po Wu for writing the parser/solver SDPSOL, which greatly helped my research. Special thanks go to Arash Hassibi for being a caring mentor and a great friend over the years. I would also like to thank Costis Maglaras, Mike Grant, Babak Hassibi, Bijit Halder, and Yao-Ting Wang for helping me get settled at ISL in my first year. I am grateful to Denise Murphy for her efficient administrative assistance and helpful attitude.

My heartfelt acknowledgments go to my teachers in all stages of my education, the individuals who have inspired me to learn and grow. I specially like to thank Professor Homayoun Hashemi of Sharif University, for his valuable support and encouragements in my undergraduate years, and Mr. Jafari, my high-school teacher, for his stimulating teaching that sparked the love of mathematics and physics in me.

I am grateful to all friends who have supported me through the years and have made my time at Stanford so enjoyable. In particular, I would like to thank Ardavan, Arjang, Atul, Azita, Dara, Fati, Kristin, Kaveh, Nima, Nogol, Vida, and Yasi. My special thanks go to my close friend Tina, with whom I share many unforgettable memories.

My deepest gratitude and love, of course, belong to my parents Reza and Shahnaz, my sister Fati, and my brother Hossein, for their unconditional love and support all though my life. To them I owe all that I am and all that I have ever accomplished, and it is to them that I dedicate this thesis.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

In this dissertation, we study optimization problems that involve minimizing the rank of a matrix over a convex set. We refer to this as the Rank Minimization Problem (RMP). The problem arises in diverse areas such as control, system identification, statistics, signal processing, computational geometry, and combinatorial optimization.

In many applications, notions such as order, complexity, or dimension of a model or design can be expressed as the rank of a matrix. If the set of feasible models or designs is described by convex constraints, then choosing the simplest model can often be expressed as an RMP. For example, a low-rank matrix could correspond to a low-order controller for a system, a low-order statistical model fit for a random process, a shape that can be embedded in a low-dimensional space, or a design with a small number of components. It is not surprising that rank minimization has such a wide range of applications across all disciplines of engineering and computational sciences: we are often interested in *simple* models. This idea is well captured by the principle known as *Occam's razor*, which states that "Among competing explanations for a phenomenon, the simplest one is the best."

There are several special cases of the RMP that have well known solutions. For example, approximating a given matrix with a low-rank matrix in spectral or Frobenius norm is an RMP that can be solved via singular value decomposition. However, in general, the RMP is known to be computationally intractable (NP-hard). Therefore, we do not expect to find a computationally efficient (polynomial-time) method that can solve all instances of the problem exactly. What we look for, instead, are *heuristics* that solve the problem approximately but efficiently.

There exist several ad hoc and heuristic methods for the RMP. The major drawbacks of the existing methods are that they are highly sensitive to the choice of initial point, generally converge very slowly, and do not provide any information on the global minimum (*e.g.*, a global lower bound).

In this dissertation, we propose new heuristics for the RMP, based on convex approximations, that do not require an initial point, are numerically very efficient, and provide a global lower bound on the optimal value. We also provide theoretical results in support of their use. In their original form, these heuristics are applicable only to the case where the matrix is positive semidefinite. We present a new result that shows any general RMP can be embedded in a larger, positive semidefinite one. With this embedding, our heuristics are readily extended to the general RMP.

We also show that the problem of maximizing the *sparsity* of a vector over a convex set, which has many practical applications, is a special case of the RMP. We refer to this as the Cardinality Minimization Problem (CMP). In this case, our heuristics reduce to the well-known ℓ_1-norm minimization heuristic, and to a new iterative ℓ_1-norm minimization approach. We show that the iterative heuristic yields very sparse solutions in practice.

Throughout the chapters, we list many practical applications of the RMP, covering some well known applications and also many new ones. By giving numerical examples of selected problems from various fields, from control and computational geometry to finance, we demonstrate that our proposed heuristics work very well in practice.

1.1 Semidefinite Programming

Our approach to the RMP is based on convex optimization, specially semidefinite programming. A semidefinite program in the variable x is the optimization problem

$$\begin{array}{ll}\text{minimize} & c^T x \\ \text{subject to} & A_0 + x_1 A_1 + \cdots + x_n A_n \leq B,\end{array}$$

where $A_i,\ B \in \mathbf{R}^{m\times m}$ are symmetric matrices and $\leq$ is matrix inequality. In other words, a semidefinite program (SDP) minimizes a linear function subject to linear matrix inequalities (LMIs). See for example [74, 73, 27, 2, 92, 100, 76, 93, 94]. SDPs can be solved globally using interior-point methods with great efficiency. Software for solving SDPs is widely available [91, 27, 101, 30, 3, 10, 88]. Computation time grows gracefully with problem size and required accuracy.

However, if the problem is large-scale, *i.e.*, if the values of m and/or n are very large, special methods may be needed to handle the computation. Developing methods for large-scale but sparse SDPs has been an active area of research in recent years. Various methods have been developed to exploit the sparsity in problems with special structure; for example, dual-scaling interior-point algorithms [8, 18], the inexact Gauss-Newton method with preconditioned conjugate gradients [98], primal-dual interior-point methods using conjugate residuals [89], non-linear programming methods [14], and a matrix completion method [28, 72].

1.2 Outline of the dissertation

In Chapter 2, we study the RMP and its practical significance, list many examples from a wide range of applications, and also discuss the CMP. In Chapter 3 we present and prove a lemma that shows how a general RMP can be embedded in another, larger, positive semidefinite RMP. Chapter 4 gives an overview of existing approaches to the problem, and points out their drawbacks. Some special cases of the RMP with analytical solutions are also listed. Chapter 5 presents the new methods, the *trace* and *log-det* heuristics, first for the positive semidefinite case, and then for the general case using the embedding lemma. We also provide theoretical support for the use of the trace heuristic and its extension, and illustrative examples that compare the performance of these heuristics with existing methods. Finally, in Chapter 6 we demonstrate the effectiveness of these heuristics on numerical examples from various fields: system identification, control, statistics and psychometrics, and finance.

Parts of the material in Chapters 3, 5 and 6 appear in [26]. The portfolio optimization example in Chapter 6 appears in [63].

Chapter 2

The Rank Minimization Problem

We study the general matrix rank minimization problem (RMP) expressed as

$$\text{RMP:} \quad \begin{array}{ll} \text{minimize} & \mathbf{Rank}\, X \\ \text{subject to} & X \in \mathcal{C}, \end{array} \tag{2.1}$$

where $X \in \mathbf{R}^{m \times n}$ is the optimization variable and $\mathcal{C}$ is a convex set denoting the constraints. As a generic example of the RMP, suppose we are trying to estimate or reconstruct the covariance matrix

$$X = \mathbf{E}(z - \mathbf{E}\, z)(z - \mathbf{E}\, z)^T$$

of a random vector $z \in \mathbf{R}^n$, from measurements and prior assumptions. Here $\mathbf{E}(z)$ denotes the expectation of the random vector z. The constraint $X \in \mathcal{C}$ expresses the condition that the estimated covariance matrix is consistent with (or not improbable for) our measurements or observed data and prior assumptions. For example, it could mean that entries in X should lie in certain intervals. The rank of X is a measure of the complexity of the stochastic model of z, in the sense that it gives the number of underlying independent random variables needed to explain the covariance of z.

The RMP (2.1) is therefore the problem of finding the least complex stochastic model (*i.e.*, covariance) that is consistent with the observations and prior assumptions. As we will show in the examples, this problem has many practical applications.

In other applications, rank can have other meanings such as embedding dimension, controller order, or number of signals present. These will be explored further in the Section 2.2.

In problem (2.1), we allow any constraints on the matrix as long as they describe a convex set. Thus, we cover a large number of constraints and specifications that come up in practice. For example, constraints on the accuracy of a model or the performance of a design are common; *e.g.*, $f(X) \leq t$, where $f(\cdot)$ is a (convex) measure of performance, and $t \in \mathbf{R}$ is the tolerance. We give many examples of these constraints in the next section.

In practice, there is a fundamental trade-off between the model complexity, captured by $\mathbf{Rank}\, X$, and its accuracy, captured by $f(X)$. We can obtain a trade-off curve by solving the RMP

$$\begin{array}{ll} \text{minimize} & \mathbf{Rank}\, X \\ \text{subject to} & X \in \mathcal{C} \\ & f(X) \leq t, \end{array} \tag{2.2}$$

for a range of values of t. Alternatively, if we solve the rank-constrained problem

$$\begin{array}{ll} \text{minimize} & f(X) \\ \text{subject to} & X \in \mathcal{C} \\ & \mathbf{Rank}\, X \leq m, \end{array} \tag{2.3}$$

for various values of the integer m, the same trade-off curve is obtained. A typical curve is shown in Figure (2.1), where f^* shows the minimum value of $f(X)$ in

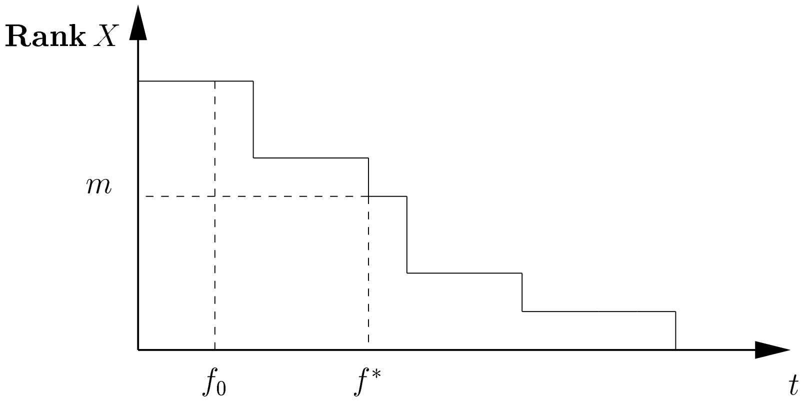

Figure 2.1: A typical curve showing the trade-off between complexity and accuracy of a model.

problem (2.3) for the given m.

The two problems above are closely related; we can solve problem (2.3) using the RMP (2.2). Let $f_0 = \min_{X \in \mathcal{C}} f(X)$; this provides a lower bound on (2.3). To solve problem (2.3), we can pick $t = \alpha f_0$ for some $\alpha > 1$, and solve (2.2) to find the optimum value, n_1. If $n_1 \leq m$, then f^*, the minimum of (2.3), is between f_0 and αf_0 and can be found by bisection (*i.e.*, iteratively dividing the interval in half). If $n_1 > m$, we increase α and repeat. Thus, if we can solve problem (2.2), we can solve problem (2.3) as well. Some examples of rank-constrained problems are given in Section 2.2.

2.1 Computational complexity

It is well known that the RMP is an NP-hard problem [20, 92]. For example, Boolean linear programming, which is known to be NP-hard, can be formulated as an RMP [92, §7.3]. To see this, consider the problem of determining whether there is an $x \in \mathbf{R}^m$ such that $Cx + d \geq 0$ and $x_i \in \{0,\ 1\}$, where $C \in \mathbf{R}^{n \times m}$ and $d \in \mathbf{R}^m$ are given. This problem is NP-hard. It can be formulated as the RMP

$$\begin{array}{ll} \text{minimize} & \mathbf{Rank}\,\mathbf{diag}(x_1,\ldots,x_m,1-x_1,\ldots,1-x_m) \\ \text{subject to} & Cx+d\geq 0, \end{array} \tag{2.4}$$

where $\mathbf{diag}(x_1,\ldots,x_n)$ refers to a diagonal matrix with the x_i as its diagonal entries. Here the rank of $\mathbf{diag}(x_1,\ldots,x_m,1-x_1,\ldots,1-x_m)$ is always at least m and equals m only when $x_i \in \{0,1\}$.

There are special cases of the RMP, though, where an exact solution can be found, *e.g.*, through singular value decomposition (SVD). These special cases are discussed in Chapter 4.

2.2 Examples

In this section we catalog many applications of the RMP. The purpose is to convey the general nature of the problem and the practical meaning of the rank, to emphasize the generality of the problem, and to point out connections between applications in different areas. Some of these problems are treated in greater detail in Chapter 6, along with numerical examples.

2.2.1 Rank of a covariance matrix

We begin by listing some examples that deal with the rank of a covariance matrix. These problems arise in statistics, econometrics, signal processing, and other fields where second-order statistics for random processes are used.

Multivariate statistical data analysis

Second-order statistical data analysis methods, such as principal component analysis and factor analysis (see, *e.g.*, [59]), deal with covariance matrices estimated from noisy data. Because of noise, the estimated covariance matrices have full rank (with probability one). Finding a covariance matrix of low rank comes up naturally in these methods. As mentioned in the beginning of this chapter, a low-rank covariance matrix corresponds to a simple explanation or model for the data. For example, consider the following constrained factor analysis problem:

$$\begin{array}{ll} \text{minimize} & \mathbf{Rank}(\Sigma) \\ \text{subject to} & \|\Sigma - \hat{\Sigma}\|_F \leq \epsilon, \\ & \Sigma \geq 0 \\ & \Sigma \in \mathcal{C}, \end{array}$$

where $\Sigma \in \mathbf{R}^{n \times n}$ is the optimization variable, $\hat{\Sigma}$ is the measured covariance matrix, $\mathcal{C}$ is a convex set denoting the prior information or assumptions on Σ, and $\|\cdot\|_F$ denotes the Frobenius norm of a matrix (other matrix norms can be handled as well). The constraint $\|\Sigma - \hat{\Sigma}\|_F \leq \epsilon$ means that the error, *i.e.*, the difference between Σ and the measured covariance in Frobenius norm, must be less than a given tolerance ϵ. The constraint $\Sigma \geq 0$ ensures that we obtain a valid covariance matrix. In the statistics terminology, the objective function, $\mathbf{Rank}\,\Sigma$ corresponds to the number of *factors* that explain Σ.

If $\mathcal{C} = \mathbf{R}^{n \times n}$ (*i.e.*, no prior information), this problem has an SVD-based analytical solution; see Section 4.1. However, extra constraints such as upper and lower bounds on the entries of Σ result in a computationally hard problem.

Sensor array processing

In sensor array processing, data containing the superposition of a number of signals, corrupted by additive noise, is measured at p spatially separated sensors. The vector of observations or measurements $y(t) \in \mathbf{R}^p$ can be modeled as

$$y(t) = \sum_{i=1}^{k} x_i(t) a_i + v(t),$$

where $x_i(t) \in \mathbf{R}$ is the ith signal, $v(t) \in \mathbf{R}^p$ is the noise, and $a_i \in \mathbf{R}^p$ is the response of the sensor array to the ith signal and is typically a function of some signal-dependent parameter. Equivalently,

$$y(t) = Ax(t) + v(t),$$

where $x(t) = [x_1(t), \ldots, x_k(t)]^T$ and $A = [a_i \;\ldots\; a_k]$. Each vector $y(t_i)$ is a snapshot across the array of sensors at time t_i. Given observations $y(t_1), \ldots, y(t_N)$, it is desired to estimate the unknown number of signals k, where $k < p$.

We assume that $x(t)$ has covariance matrix Σ_x, and $v(t)$ is white Gaussian noise, independent of $x(t)$, with covariance matrix $\sigma^2 I$. The covariance of $y(t)$ is given by $\Sigma_y = \Psi + \sigma^2 I$, where $\Psi = A\Sigma_x A^T$. If we assume A to have full rank, we have $k = \mathbf{Rank}\,\Sigma_x = \mathbf{Rank}(A\Sigma_x A^T) = \mathbf{Rank}\,\Psi$. Thus, the *number of signals* is expressed as the *rank* of a covariance matrix.

This problem comes up in antenna arrays [97], harmonic retrieval [62], and many other applications. For example, Figure 2.2 shows the set up in a general antenna array processing system. Here $x_i(t)$ is the complex amplitude of a plane-wave signal impinging upon the array at an angle θ_i. Each column of A, marked as $a(\theta_i)$ in the figure, gives the response of the antenna array in the direction θ_i.

We now formulate the problem of estimating the number of signals k, as well as the covariances Σ_y, Ψ, and σ^2. One constraint for Σ_y is to be consistent with

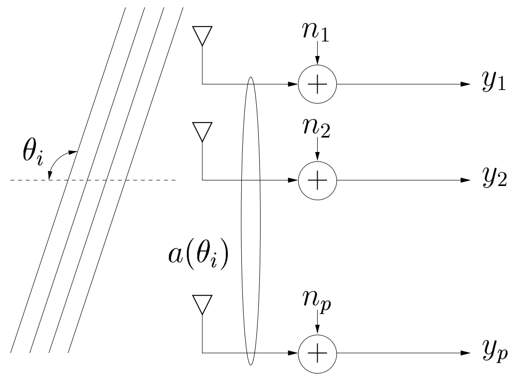

Figure 2.2: A general antenna array processing system.

the observations, *e.g.*, to maximize the likelihood of observing $y(t_1), \ldots, y(t_N)$, or to render this likelihood higher than a given threshold. By taking the log of the Gaussian joint distribution $f(y(t_1), \ldots, y(t_N))$, we obtain the log-likelihood function as

$$L(\Sigma_y) = -\frac{N}{2} \log\det \Sigma_y - \frac{N}{2} \mathbf{Tr}(\Sigma_y)^{-1}\hat{\Sigma}_y - \frac{Np}{2}\log(2\pi),$$

where $\hat{\Sigma}_y = (1/N)\sum_{i=1}^N y(t_i)y(t_i)^T$ is the covariance estimated from the observations.

Various RMPs and rank-constrained problems arise in this context. If the number of signals is known to be less than or equal to k, the maximum likelihood (ML) estimates of Ψ and σ^2 are found by solving the optimization problem

$$\begin{array}{ll} \text{maximize} & L(\Psi + \sigma^2 I) \\ \text{subject to} & \mathbf{Rank}\,\Psi \leq k \\ & \Psi \geq 0, \end{array} \tag{2.5}$$

with variables Ψ and σ^2. The solution to this problem is known to be

$$\Psi_{\text{opt}} = \textstyle\sum_{i=1}^k [\lambda_i(\hat{\Sigma}) - \sigma^2] v_i(\hat{\Sigma}) v_i(\hat{\Sigma})^T, \qquad \sigma^2_{\text{opt}} = \frac{1}{p-k}\sum_{i=k+1}^p \lambda_i(\hat{\Sigma}), \tag{2.6}$$

where λ_i and v_i denote the eigenvalues and eigenvectors, respectively [5]. Note that the above is also the solution to the rank-constrained problem

$$\begin{array}{ll} \text{minimize} & \|\hat{\Sigma}_y - \Psi - \sigma^2 I\|_F \\ \text{subject to} & \mathbf{Rank}\,\Psi \leq k \\ & \Psi \geq 0, \end{array}$$

which minimizes the least-squares error between Σ and $\hat{\Sigma}$. As we show in Chapter 4, this rank-constrained problem is readily solved for Ψ using SVD and yields the same Ψ_{opt} as above. The objective value will then be

$$\| \sum_{i=k+1}^{p} [\lambda_i(\hat{\Sigma}) - \sigma^2] v_i(\hat{\Sigma}) v_i(\hat{\Sigma})^T \|_F^2 = \sum_{i=k+1}^{p} (\lambda_i(\hat{\Sigma}) - \sigma^2)^2,$$

which attains its minimum value if σ^2 is chosen as the σ^2_{opt} given in (2.6).

In [97], the problem of estimating the number of signals is formulated as

$$\begin{array}{ll} \text{minimize} & -\log\det\Sigma_y - \mathbf{Tr}(\Sigma_y)^{-1}\hat{\Sigma}_y + g(k) \\ \text{subject to} & \Sigma_y - \sigma^2 I \geq 0, \end{array} \tag{2.7}$$

where $g(k)$ is a measure of 'complexity' of the model as a function of the number of free parameters in the model (*i.e.*, in Σ_y), which is in turn a function of $k = \mathbf{Rank}(\Sigma_y - \sigma^2 I) = \mathbf{Rank}\,\Psi$.

Two common choices for $g(k)$ come from information theoretic measures: the Akaike Information Criterion (AIC) [1] and Rissanen's Minimum Description Length (MDL) [80] criterion. For this problem, both criteria result in quadratic functions of the rank k: we have $\mathrm{AIC}(k) = k(2p-k)$ and $\mathrm{MDL}(k) = \frac{1}{2}(\log N)k(2p-k)$. (Note that these problems, although related to the RMP, are not RMPs themselves.)

Note that if k is fixed, the problem reduces to (2.5), whose optimal objective value in terms of k can be found using (2.6). To find the optimal k in (2.7), we simply need to check the value of this objective for $k = 1, \ldots, p$ (see [97] and references therein).

These are examples of RMPs and related problems that can be solved analytically. However, this may not be possible if we have additional constraints. We may have upper and lower bounds on the variances of some y_i, or know that, for example, y_i and y_j have a higher correlation than y_k and y_l. With such additional constraints, the resulting RMP is computationally hard.

The Frisch problem

Let $x \in \mathbf{R}^n$ be a random vector, with covariance matrix Σ_x. Suppose we have measurements of

$$y(t) = x(t) + v(t),$$

where the measurement noise v has zero mean, is uncorrelated with x, and has an unknown but diagonal covariance matrix $D = \mathbf{diag}\, d$. It follows that

$$\Sigma_y = \Sigma_x + D,$$

where Σ_y denotes the covariance of y. The problem is to identify, from noisy observations, the largest number of linear relations among the underlying data. This corresponds to the minimum rank Σ_x, since the rank drops by one for each linear relation that exists among the x_is. We assume we can estimate Σ_y with high confidence; *i.e.*, we consider it known. This problem can be expressed as the RMP

$$\begin{array}{ll} \text{minimize} & \mathbf{Rank}(\Sigma_y - D) \\ \text{subject to} & \Sigma_y - D \geq 0 \\ & D \geq 0 \\ & D \text{ diagonal}, \end{array} \tag{2.8}$$

where Σ_y and $d = \mathbf{diag}\, D$ are the optimization variables. Problem (2.8) also arises in a deterministic setting, where the matrices Σ_y and Σ_x are interpreted as Gramians rather than covariances [20].

For the special case where Σ^{-1} is Frobenius equivalent (*i.e.*, there exists $J = \mathbf{diag}\{\pm 1\}$ such that $J\Sigma^{-1}J$ has positive entries), Kalman [57] has shown that the minimum possible rank is $n-1$, and has given a complete characterization of the solutions. In [99], a lower bound on the minimum rank in problem (2.8) is given. However, in general no analytical solution is known for (2.8).

2.2.2 Rank of a Hankel matrix

We saw in the previous section that the rank of a covariance matrix plays a central role in many statistical methods as a notion of complexity of the stochastic model.

The rank of a *Hankel* matrix has similar significance in model identification problems in system theory and signal processing. It comes up commonly in problems that deal with recursive sequences, where the order of the recursion is expressed by the rank of an appropriate Hankel matrix. Two examples of RMPs involving Hankel matrices are given in this section.

Reconstructing polygons from moments

Consider a polygonal region P in the complex plane with vertices $z_1, \ldots, z_m$ ordered counterclockwise. *Complex moments* of P are defined as

$$\tau_k \triangleq k(k-1) \iint_P z^{k-2}\, dx\, dy, \qquad \tau_0 = \tau_1 = 0,$$

where $z = x + jy$. Using a theorem by Davis [21], we can write the moments as

$$\tau_k = \sum_{i=1}^{m} a_i z_i^k = a^T \begin{bmatrix} z_1^k \\ z_2^k \\ \vdots \\ z_m^k \end{bmatrix},$$

where the a_is are some complex constants. Let H_n be the $n \times n$ Hankel matrix consisting of the τ_i as follows:

$$H_n = \begin{bmatrix} \tau_0 & \tau_1 & \tau_2 & \dots & \tau_{n-1} \\ \tau_1 & \tau_2 & \tau_3 & \dots & \tau_n \\ \tau_2 & \tau_3 & \tau_4 & \dots & \tau_{n+1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \tau_{n-1} & \tau_n & \tau_{n+1} & \dots & \tau_{2n-2} \end{bmatrix}. \tag{2.9}$$

We now show that $m = \textbf{Rank}\, H_n$; that is, the number of vertices of P is the rank of the Hankel matrix consisting of the complex moments. Let H_m be the leading $m \times m$ block in the above matrix. Note that H_m can be written as $H_m = V_m(\textbf{diag}\, a)V_m^T$, where V_m is the Vandermonde matrix

$$V_m = \begin{bmatrix} 1 & 1 & \dots & 1 \\ z_1 & z_2 & \dots & z_m \\ \vdots & \vdots & \ddots & \vdots \\ z_1^m & z_2^m & \dots & z_m^m \end{bmatrix}.$$

Since the z_i are distinct, V_m has full rank. Thus, $\textbf{Rank}\, H_m = \textbf{Rank}\, \textbf{diag}\, a = m$. For any $n > m$, the rank of H_n stays equal to m, because the new rows and columns are linearly dependent on the previous ones. To see this, we write H_n as $H_n =$

$V_n(\mathbf{diag}\, a)V_n^T$, from which it is clear that $\mathbf{Rank}\, H_n \leq m$. Noting that the rank of the leading block in H_n is m, we conclude $\mathbf{Rank}\, H_n = m$, for any $n \geq m$. This shows that, in fact, the sequence of the τ_is is recursive of order m.

We now show that estimating the number of vertices from noisy measurements of the complex moments can be expressed as an RMP. Suppose we have

$$y_k = \tau_k + v_k, \qquad k = 0, \ldots, 2n-2,$$

with some model for the noise v_k, for example, white Gaussian with zero mean and a known variance. Suppose also that the number of vertices m is unknown (we assume $m \leq n$, otherwise the n measurements are not enough to determine the vertices). The goal is to estimate P, the region with the smallest number of vertices that is consistent with the measured moments. Let $\hat{H}_n$ be the Hankel matrix formed using the noisy measurements $y_0, \ldots, y_{2n-2}$. The problem can be expressed as the RMP

$$\begin{array}{ll} \text{minimize} & \mathbf{Rank}\, H \\ \text{subject to} & \|H - \hat{H}_n\|_F \leq \epsilon \\ & H \ \text{Hankel}, \end{array}$$

where the optimization variables are $\tau_0, \ldots, \tau_{2n-2}$ that form the Hankel matrix H, and ϵ is the desired tolerance or the noise variance. The first constraint can also be written as $\sum_{i=0}^{2n-2}(\tau_i - y_i)^2 \leq \epsilon^2$. Without the constraint that H must be Hankel, this problem can be solved via SVD. With the Hankel constraint, however, the problem is hard.

In [37, §5] and [69], more details on this problem and some interesting connections to the sensor array processing problem mentioned in Section (2.2.1) are given. This problem comes up in various applications such as signal recovery [83], computed tomography [69], and inverse potential theory [87].

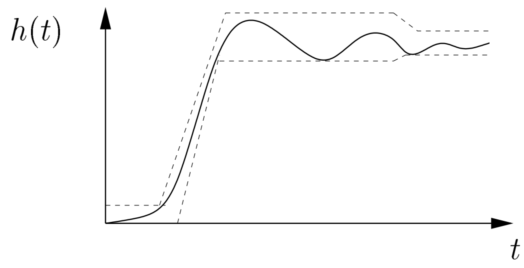

Figure 2.3: Typical step response specifications.

System realization

Problems involving minimizing the rank of a Hankel matrix also come up in system realization, *e.g.*, in designing a low-order linear, time-invariant system directly from convex specifications on its impulse response.

As an example, suppose desired specifications are given as upper and lower bounds on the first n samples of the step response, as in Figure 2.3. We would like to find a linear system with lowest order that fits the constraints. The dashed lines in the figure are meant to capture a typical set of time domain step response specifications: certain rise-time, slew-rate, overshoot, and settling characteristics. We will discuss this problem in detail in Chapter 6, along with numerical examples. There we show that it can be expressed as the RMP

$$\begin{array}{ll} \text{minimize} & \mathbf{Rank}\, H_n \\ \text{subject to} & l_i \leq s_i \leq u_i, \quad k = 1, \ldots, n \\ & h_{n+1}, \ldots, h_{2n-1} \in \mathbf{R}, \end{array} \tag{2.10}$$

where $s_k = \sum_{i=1}^{k} h_i$ denote the terms in the step response, l_i and u_i are, respectively, samples of the lower and upper time domain specifications, and

$$H_n = \begin{bmatrix} h_0 & h_1 & \dots & h_{n-1} \\ h_1 & h_2 & \dots & h_n \\ \vdots & \vdots & \ddots & \vdots \\ h_{n-1} & h_n & \dots & h_{2n-2} \end{bmatrix}.$$

We can readily extend this problem to MIMO (Multi-Input, Multi-Output) systems by using block-Hankel matrices.

2.2.3 Quadratic and bilinear matrix inequalities as rank constraints

Quadratic and Bilinear Matrix Inequalities (BMI) arise in many areas, especially in control and combinatorial optimization, where they play a central role. In control, the BMI has been extensively studied as a general framework for many NP-hard problems. In combinatorial optimization, non-convex quadratic inequalities are essential because they can capture Boolean or integer constraints on variables. For example, the constraint $x_i \in \{-1, 1\}$ is equivalent to the two constraints $x_i^2 \leq 1,\ x_i^2 \geq 1$, where the second inequality is non-convex.

In this section, we show how optimization problems involving quadratic matrix inequalities and BMIs can be cast as rank constrained problems.

Bilinear Matrix Inequality (BMI) problems

Consider the following problem with a *quadratic* matrix inequality:

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & C + \sum_{i=1}^{m} x_i A_i + \sum_{i,j=1}^{m} x_i x_j B_{ij} \leq 0, \end{array} \tag{2.11}$$

where $x \in \mathbf{R}^n$ is the optimization variable, and $c \in \mathbf{R}^n$ and the symmetric matrices A_i, B_{ij}, $C \in \mathbf{R}^{n\times n}$ are given. This problem is very general, but also non-convex. For example, if the matrices C, A_i, and B_{ij} are diagonal, the constraint in (2.11) reduces to a set of n (possibly indefinite) quadratic constraints in x. Problem (2.11) therefore includes all quadratic optimization problems. It also includes all polynomial problems (since by introducing new variables, one can reduce any polynomial inequality to a set of quadratic inequalities), and all $\{0,1\}$ and integer programs.

In control theory, a more restricted *bilinear* form is often suffieiently general. Here we split the variables in two vectors x and y, and replace the constraint by a *bilinear* matrix inequality (BMI):

$$\begin{array}{ll} \text{minimize} & c^T x + b^T y \\ \text{subject to} & D + \sum_{i=1}^{m} x_i A_i + \sum_{k=1}^{l} y_k B_k + \sum_{i=1}^{m}\sum_{k=1}^{l} x_i y_k C_{ik} \leq 0. \end{array} \tag{2.12}$$

BMIs include a wide variety of control problems, including synthesis with structured uncertainty [32, 34, 31], and fixed-order and fixed-structure controller design [23]. For more on the BMI and its computational methods, see [33, 81, 35].

We show that problem (2.12) can be cast as a rank-constrained problem. We first express the problem as

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & C + \sum_{i=1}^{m} x_i A_i + \sum_{i,j=1}^{m} w_{ij} B_{ij} \leq 0 \\ & w_{ij} = x_i x_j. \quad i,j = 1,\ldots,m \end{array}$$

The second constraint can be written as $W = xx^T$. This equality is equivalent to the following:

$$\textbf{Rank} \begin{bmatrix} W & x \\ x^T & 1 \end{bmatrix} = 1, \tag{2.13}$$

since the block matrix above is rank one if and only if the Schur complement of the (2,2) block is equal to zero, *i.e.*, $W - xx^T = 0$. To see this, we use the result (see, *e.g.*, [50, §2.2]) that the rank of a Hermitian block matrix is equal to the rank of a diagonal block plus the rank of its Schur complement, *i.e.*, if C is invertible, then

$$\textbf{Rank} \begin{bmatrix} A & B \\ B^T & C \end{bmatrix} = \textbf{Rank}\, C + \textbf{Rank}(A - BC^{-1}B^T). \tag{2.14}$$

The constraint (2.13) implies $\textbf{Rank}\, 1 + \textbf{Rank}(W - xx^T) = 1$, or $W = xx^T$. Therefore, problem (2.11) is equivalent to

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & C + \sum_{i=1}^m x_i A_i + \sum_{i,j=1}^m w_{ij} B_{ij} \leq 0 \\ & \textbf{Rank} \begin{bmatrix} W & x \\ x^T & 1 \end{bmatrix} \leq 1, \end{array} \tag{2.15}$$

where we have replaced the equality in (2.13) with an inequality so that the problem has the form of (2.3).

Combinatorial optimization problems

Many combinatorial optimization problems can be expressed as rank-constrained problems in the form of (2.3). We present some examples in this section. Consider the quadratic optimization problem

$$\begin{array}{ll} \text{minimize} & x^T A_0 x + 2b_0^T x + c_0 \\ \text{subject to} & x^T A_i x + 2b_i^T x + c_i \leq 0, \qquad i = 1, \dots, L, \end{array} \tag{2.16}$$

where $x \in \mathbf{R}^k$ is the optimization variable. The matrices A_i can be indefinite, and therefore problem (2.16) is a non-convex optimization problem. For example, it includes all problems with polynomial objective function and polynomial constraints [73, §6.4.4], [84].

Define the new variable $X \in \mathbf{R}^{k \times k}$ as $X = xx^T$. As shown in the previous section, this can be written as

$$\mathbf{Rank} \begin{bmatrix} X & x \\ x^T & 1 \end{bmatrix} = 1.$$

Noting that

$$x^T A_i x = \mathbf{Tr}\, A_i x x^T = \mathbf{Tr}\, AX,$$

we can write the quadratic terms in the objective function and the constraints in terms of X. Thus, problem (2.16) is equivalent to

$$\begin{array}{ll} \text{minimize} & \mathbf{Tr}\, A_0 X + 2b_0^T x \\ \text{subject to} & \mathbf{Tr}\, A_i X + 2b_i^T x + c_i \leq 0 \qquad i = 1, \dots, L \\ & \mathbf{Rank} \begin{bmatrix} X & x \\ x^T & 1 \end{bmatrix} \leq 1. \end{array}$$

Except for the rank constraint, all constraints and the objective function are convex in the optimization variables X and x, therefore this is a rank constrained optimization problem as in (2.3). As a simple example, consider the $\{-1, 1\}$ quadratic program

$$\begin{array}{ll} \text{minimize} & x^T A x + 2b^T x \\ \text{subject to} & x_i^2 = 1, \qquad i = 1, \ldots, k, \end{array}$$

which is equivalent to the rank constrained problem

$$\begin{array}{ll} \text{minimize} & \mathbf{Tr}\, AX + 2b^T x \\ \text{subject to} & X_{ii} = 1 \qquad i = 1, \ldots, k \\ & \mathbf{Rank} \begin{bmatrix} X & x \\ x^T & 1 \end{bmatrix} \leq 1. \end{array} \tag{2.17}$$

There exist convex relaxations to problem (2.16), which have been popular in recent years and are referred to as semidefinite relaxations [42, 58, 79]. For a survey of semidefinite programming in combinatorial optimization, see [2]. The basic relaxation can be obtained by simply relaxing the non-convex constraint $X = xx^T$ to the convex constraint $X \geq xx^T$, which can be written as the Linear Matrix Inequality (LMI) constraint

$$\begin{bmatrix} X & x \\ x^T & 1 \end{bmatrix} \geq 0.$$

We point out that the same relaxation is obtained if the trace heuristic discussed in Chapter 5 is applied to problem (2.17).

2.2.4 Other examples

Problems in systems and control

RMPs have been studied extensively in the control literature, since many important problems in controller design and system identification can be expressed as an RMP. We pointed out various control applications of the BMI in the previous section. Here we list several other examples.

Minimum-order controller design is perhaps the mostly widely studied problem among these. Its formulation as an RMP is well known. We discuss this problem in Chapter 6, along with numerical examples. Another problem is model order reduction in system identification. The problem of minimum order system identification from frequency domain data is formulated as an RMP and discussed in Chapter 6.

Other applications include reduced-order $\mathcal{H}_\infty$ synthesis and reduced-order μ synthesis with constant scalings [24], problems with intertia constraints [46], exact reducibility of uncertain systems [7], and simultaneous stabilization of linear systems [48].

Fast matrix computations

Low-rank matrix approximations are sometimes used to save computational effort. As a simple example, suppose we want to compute $y = Ax$, where $A \in \mathbf{R}^{m\times n}$, for various values of x, and suppose m and n are large. This requires mn multiplications. If $\mathbf{Rank}\, A = r$, then A can be factored as $A = RL^T$, where $R \in \mathbf{R}^{m\times r}$ and $L \in \mathbf{R}^{m\times r}$. Thus, $y = RL^Tx$ can be computed with only $(m+n)r$ multiplications. If r is much smaller than m and n, this could lead to significant savings in computation.

The simplest matrix approximation problem is

$$\begin{array}{ll} \text{minimize} & \mathbf{Rank}\, \hat{A} \\ \text{subject to} & \|A - \hat{A}\| \leq \epsilon, \end{array} \tag{2.18}$$

where $\hat{A}$ is the optimization variable and ϵ is the tolerance. This problem can readily be solved via SVD, as we see in Chapter 4. However, often when A has a particular structure (*e.g.*, Hankel or Toeplitz), $\hat{A}$ is desired to retain that structure. Such additional contraints typically make the problem hard.

One application arises in fast image simulation in microlithography. Optical image calculation using the Hopkins model involves a double-convolution operation with the kernel of the imaging system (see [38]). After discretization, this reduces to computing the quadratic $y = x^T W x$ for various inputs x and their shifted values. Using a low-rank approximation to W, referred to as Optimal Coherent Decomposition (OCD), a fast image simulator is made possible [96].

2.3 The cardinality minimization problem

An important special case of the RMP is minimizing the number of nonzero components of a vector $x \in \mathbf{R}^n$ (its *cardinality*, denoted $\mathbf{card}\, x$). Other common terms for cardinality include weight, sparsity, and 0-norm. The term 0-norm can be confusing since $\mathbf{card}\, x$ is certainly not a norm in the usual sense; it is neither homogeneous of degree one nor convex (the terminology will be explained below). We will also use the term *sparse* in the usual qualitative sense, to describe a vector $x \in \mathbf{R}^n$ with relatively few nonzero entries, *i.e.*, one with $\mathbf{card}\, x \ll n$.

To see that this problem is a special case of the RMP, let the matrix X in (2.1) be diagonal, *i.e.*, $X = \mathbf{diag}\, x$. Then $\mathbf{Rank}\, X$ is the same as the number of nonzero entries of the vector x. The constraint $X \in \mathcal{C}$ reduces to $x \in \bar{\mathcal{C}}$, where $\bar{\mathcal{C}}$ is the pre-image of the previous set under the mapping $x \to \mathbf{diag}\, x$. In this case, problem (2.1) is equivalent to searching for the *sparsest* vector in a convex set. We refer to this as the Cardinality Minimization Problem (CMP):

$$\begin{array}{ll} \text{minimize} & \mathbf{card}\, x \\ \text{subject to} & x \in \bar{\mathcal{C}}. \end{array} \tag{2.19}$$

This problem comes up in many application areas. In engineering design problems, x might represent some design variables and $\mathcal{C}$ the constraints and specifications. If $x_i = 0$ corresponds to an element or degree of freedom not used, then a sparse x

corresponds to an efficient design, *i.e.*, one that uses a small number of elements. The problem (2.19) is then to find the most efficient (or least complex) design that meets all the specifications.

A related interpretation occurs in modeling problems. Here x gives the coefficients of some model, and $x \in \mathcal{C}$ is the constraint that the model is consistent with (or not improbable for) the measured or observed data. In this case the problem (2.19) is to find the simplest model, *i.e.*, the one involving the least number of terms.

Another example comes up in wavelet decomposition of signals. In the decomposition of a signal (*e.g.*, an image) as a linear combination of known basis signals (*e.g.*, wavelet, edgelet, Fourier), sparse coefficients lead to signal compression [16].

The CMP (2.19) is also sometimes referred to as the ℓ_0-norm minimization problem, for the following reason. For $p \geq 1$, the standard p-norm is given by

$$\|x\|_p = \left(\sum_{i=1}^{n} |x_i|^p \right)^{1/p}.$$

The p-norm minimization problem is to minimize $\|x\|_p$ subject to $x \in \bar{\mathcal{C}}$. This is equivalent to minimizing the pth power of the p-norm, *i.e.*, the problem

$$\begin{array}{ll} \text{minimize} & \sum_{i=1}^{n} |x_i|^p \\ \text{subject to} & x \in \bar{\mathcal{C}}. \end{array} \tag{2.20}$$

For $p \geq 1$ this is a convex optimization problem. For $0 < p < 1$ the objective makes sense, but is neither convex nor the pth power of a norm. Now as $p \to 0$, the objective converges to the cardinality of x:

$$\lim_{p \to 0} \sum_{i=1}^{n} |x_i|^p = \mathbf{card}(x),$$

so the CMP can be considered as a limit of the p-norm minimization problem (2.20) as $p \to 0$. The similarity between the CMP (2.19) and the p-norm minimization problem (2.20) does not extend too far. For $p \geq 1$ the objective in the p-norm minimization problem is convex, so the problem can usually be solved globally and efficiently. For $p < 1$ the objective in the p-norm minimization problem is not convex, and the problem can have a very large number of local, but not global, minimizers.

The CMP is in general an NP-hard combinatorial problem. To find the global optimum, we need to check the feasibility of all 2^n sparsity patterns for x. There are special cases though where the solution can be found efficiently, as discussed in Chapter 4.

The CMP arises in sparse design problems, *e.g.*, truss design [95, 60] and power-ground mesh design [13]; in signal processing problems, *e.g.*, recovering sparse signals in noise [53, 39] and best basis selection [19, 61]; in wavelet decomposition problems [22, 16]; and many other applications. An application in portfolio optimization with fixed costs is discussed in detail in Chapter 6, along with numerical examples.

Chapter 3

Semidefinite Embedding

In Chapter 2, we stated the rank minimization problem and gave examples of its wide range of applications. Beginning in Chapter 4, we describe approaches to (approximately) solve the RMP. Before discussing solution approaches, however, we present and prove a useful property of the RMP, *semidefinite embedding*, that will be used in the next chapters to extend methods applicable to positive semidefinite RMPs to any general RMP.

3.1 Positive semidefinite RMP

If the matrix variable X in the RMP (2.1) is constrained to be positive semidefinite (PSD), *i.e.*, if the feasible set is a subset of the positive semidefinite cone, we call the problem a positive semidefinite RMP. The PSD cone has properties that aid us in finding a low-rank matrix; for example, we see later that such a matrix will always lie on the boundary of the cone. In fact, this is the basis of the analytical anti-centering and potential reduction methods that we discuss in Chapter 4. In Chapter 5, we present the trace and log-det heuristics for the RMP and give theoretical results in support of their use. These methods are also applicable only to PSD matrices.

There are many applications where X is not necessarily PSD, or even square. Thus, it becomes important to find a way to deal with the general RMP, problem (2.1). One of the contributions of this dissertation is to resolve this issue by showing that any general RMP can be embedded in a larger, positive semidefinite RMP. We refer to this as the *semidefinite embedding* lemma. In this chapter, we state and prove this lemma and point out its implications.

3.2 The semidefinite embedding lemma

We show that it is possible to associate with any nonsquare matrix X, a positive semidefinite matrix whose rank is exactly twice the rank of X.

Lemma 1 *Let $X \in \mathbf{R}^{m\times n}$ be a given matrix. Then $\mathbf{Rank}\, X \leq r$ if and only if there exist matrices $Y = Y^T \in \mathbf{R}^{m\times m}$ and $Z = Z^T \in \mathbf{R}^{n\times n}$ such that*

$$\mathbf{Rank}\, Y + \mathbf{Rank}\, Z \leq 2r, \qquad \begin{bmatrix} Y & X \\ X^T & Z \end{bmatrix} \geq 0. \tag{3.1}$$

This result means that minimizing the rank of a general nonsquare matrix X, problem (2.1), is equivalent to minimizing the rank of the semidefinite, block diagonal matrix $\mathbf{diag}(Y, Z)$:

$$\begin{array}{ll} \text{minimize} & \frac{1}{2}\, \mathbf{Rank}\, \mathbf{diag}(Y, Z) \\ \text{subject to} & \begin{bmatrix} Y & X \\ X^T & Z \end{bmatrix} \geq 0 \\ & X \in \mathcal{C}, \end{array} \tag{3.2}$$

with variables X, Y and Z. The equivalence is in the following sense: the tuple $(X^\star, Y^\star, Z^\star)$ is optimal for (3.2) if and only if $X^\star$ is optimal for problem (2.1), and the objective values in both problems are the same (which is why we keep the factor

$\frac{1}{2}$ in the objective).

It is possible to refine the result of Lemma 1 when X is known to have some structure:

Corollary 1 *If X has a block diagonal structure $X = \mathbf{diag}(X_1, \ldots, X_N)$, where $X_i \in \mathbf{R}^{m_i \times n_i}$, then without loss of generality, we may assume that the slack variables have the structure $Y = \mathbf{diag}(Y_1, \ldots, Y_N)$, where $Y_i = Y_i^T \geq 0 \in \mathbf{R}^{m_i \times m_i}$, and $Z = \mathbf{diag}(Z_1, \ldots, Z_N)$, where $Z_i = Z_i^T \geq 0 \in \mathbf{R}^{n_i \times n_i}$.*

To see this, note that $\mathbf{Rank}\, X = \sum_i \mathbf{Rank}\, X_i$ and apply Lemma 1 to each block to get

$$\begin{bmatrix} Y_i & X_i \\ X_i^T & Z_i \end{bmatrix} \geq 0 \quad i = 1, \ldots, N. \tag{3.3}$$

Corollary 2 *If X is symmetric, then without loss of generality, we can take $Y = Z$.*

This is because for any feasible Y and Z for (3.2), say with $\mathbf{Rank}\, Y \leq \mathbf{Rank}\, Z$, it is possible to choose a real number $\alpha > 0$, such that replacing Y and Z in (3.2) with αY is feasible, with lower or equal objective value; see Section 3.3.

3.2.1 Vector case

We showed in Chapter 2 that the cardinality minimization problem is a special case of the RMP. To obtain this directly from problem (3.2), let $X = \mathbf{diag}\, x$. Since X is diagonal and symmetric, it follows from the above corollaries that we can take $Y = Z = \mathbf{diag}\, y$, where $y \in \mathbf{R}^n$. The problem then reduces to

$$\begin{array}{ll} \text{minimize} & \mathbf{Rank}\,\mathbf{diag}\, y \\ \text{subject to} & \begin{bmatrix} \mathbf{diag}\, y & \mathbf{diag}\, x \\ \mathbf{diag}\, x & \mathbf{diag}\, y \end{bmatrix} \geq 0 \\ & \mathbf{diag}\, x \in \mathcal{C}, \end{array} \tag{3.4}$$

in the variables x, y and z. The first constraint above is equivalent to

$$\begin{bmatrix} y_i & x_i \\ x_i & y_i \end{bmatrix} \geq 0 \quad i = 1, \ldots, n, \tag{3.5}$$

from which it is easy to see that either $x_i = y_i = 0$ or $y_i \geq |x_i|$. Therefore, problem (3.4) reduces to

$$\begin{array}{ll} \text{minimize} & \mathbf{card}\, y \\ \text{subject to} & |x_i| \leq y_i \\ & \mathbf{diag}\, x \in \mathcal{C}, \end{array} \tag{3.6}$$

or equivalently,

$$\begin{array}{ll} \text{minimize} & \mathbf{card}\, x \\ \text{subject to} & x \in \bar{\mathcal{C}}, \end{array} \tag{3.7}$$

where $\bar{\mathcal{C}}$ is the pre-image of $\mathcal{C}$ under the mapping $x \to \mathbf{diag}\, x$.

In Chapter 5, we use the same approach to specialize results about the rank problem to the cardinality problem. This is useful because it automatically provides us with heuristic solution methods for the CMP, which has many applications.

3.3 Proof of Lemma

We begin with the following lemma [12, p.28], which is a generalization of the well known Schur complement condition for positive semidefiniteness [36]:

Let X, Y, and Z be real matrices of appropriate dimensions. Then we have the following equivalence:

$$\begin{bmatrix} Y & X \\ X^T & Z \end{bmatrix} \geq 0 \quad \Leftrightarrow \quad \begin{cases} \text{(i)} & Y \geq 0 \\ \text{(ii)} & X^T(I - YY^\dagger) = 0 \\ \text{(iii)} & Z - X^T Y^\dagger X \geq 0 \end{cases}, \tag{3.8}$$

where $Y^\dagger$ denotes the Moore-Penrose psuedoinverse of Y.

Also note that for any $X \in \mathbf{R}^{m\times n}$,

$$\mathbf{Rank}\, X = n - \dim \mathcal{N}(X) = m - \dim \mathcal{N}(X^T). \tag{3.9}$$

We now proceed with the proof of Lemma 1:

Lemma 1 *Let $X \in \mathbf{R}^{m\times n}$ be a given matrix. Then $\mathbf{Rank}\, X \leq r$ if and only if there exist matrices $Y \in \mathbf{R}^{m\times m}$ and $Z \in \mathbf{R}^{n\times n}$ such that*

$$\mathbf{Rank}\, Y + \mathbf{Rank}\, Z \leq 2r, \qquad \begin{bmatrix} Y & X \\ X^T & Z \end{bmatrix} \succeq 0. \tag{3.10}$$

Proof: We show each direction separately:

($\Rightarrow$) Suppose that $\mathbf{Rank}\, X = r_0 \leq r$. Then X can be factored as $X = L\,R$, where $L \in \mathbf{R}^{m\times r_0}$ and $R \in \mathbf{R}^{r_0\times n}$, and $\mathbf{Rank}\, L = \mathbf{Rank}\, R = r_0$. Set Y and Z to be the rank r_0 matrices $L\,L^T$ and $R^T\,R$, respectively. Then we have

$$\begin{bmatrix} Y & X \\ X^T & Z \end{bmatrix} = \begin{bmatrix} L \\ R^T \end{bmatrix} \begin{bmatrix} L^T & R \end{bmatrix} \succeq 0.$$

($\Leftarrow$) In this direction, conditions (i), (ii) and (iii) in (3.8) must hold. Our goal is to show that these conditions imply that $\mathbf{Rank}\, Y \geq \mathbf{Rank}\, X$ and $\mathbf{Rank}\, Z \geq \mathbf{Rank}\, X$.

Assume, without loss of generality, that $\mathbf{Rank}\, Y \leq \mathbf{Rank}\, Z$ (if this were not the case, we could write the conditions in (3.8) with Y and Z interchanged). From condition (ii) of (3.8), since $(I - YY^\dagger)$ is a projection operator for $\mathcal{N}(Y)$, it follows that

$$\mathcal{N}(X^T) \supseteq \mathcal{N}(Y) \quad \Rightarrow \quad \dim \mathcal{N}(X^T) \geq \dim \mathcal{N}(Y).$$

Using (3.9), we conclude that $\mathbf{Rank}\, Y \geq \mathbf{Rank}\, X^T = \mathbf{Rank}\, X$.

□

Finally, we prove the claim that when X is symmetric, we can take $Y = Z$ in (3.2). Specifically, we show that given any feasible Y and Z, we can construct a matrix W that is feasible when inserted in place of Y and Z in (3.2) and yields an equal or smaller objective value.

Proof: Again assume, without loss of generality, that $\mathbf{Rank}\, Y \leq \mathbf{Rank}\, Z$. Now let α be a positive real number and consider the matrix αY. Then for any $\alpha > 0$, $\mathbf{Rank}\, \alpha Y = \mathbf{Rank}\, Y$ and αY satisfies conditions (i) and (ii) of (3.8). If we can show that for some $\alpha_0 > 0$ condition (iii) is also satisfied, then we can take $W = \alpha_0 Y$ and we are done.

Toward that end, consider the expression for condition (iii), with αY in place of Y and Z. Noting that $(\alpha Y)^\dagger = \frac{1}{\alpha} Y^\dagger$, we can be write this as

$$\alpha^2 Y - X Y^\dagger X \geq 0. \tag{3.11}$$

Recall that $Y^\dagger$ can be decomposed as

$$Y^\dagger = \begin{bmatrix} U_1 & U_2 \end{bmatrix} \begin{bmatrix} \Sigma^{-1} & \\ & 0 \end{bmatrix} \begin{bmatrix} U_1^T \\ U_2^T \end{bmatrix} = U_1 \Sigma^{-1} U_1^T, \tag{3.12}$$

where Σ contains the nonzero eigenvalues of Y, U_1 and U_2 are orthonormal matrices that span the range space of Y, $\mathcal{R}(Y)$, and the nullspace of Y, $\mathcal{N}(Y)$, respectively, and satisfy the identity:

$$U_2 U_2^T + U_1 U_1^T = I. \tag{3.13}$$

Note that when X is symmetric, condition (ii) in (3.8) is equivalent to $XU_2 = 0$. Using this relation, and pre- and post-multiplying (3.11) by $[U_1\, U_2]^T$ and $[U_1\, U_2]$, respectively, we see that (3.11) holds if and only if the following equivalent condition holds:

$$\alpha^2\Sigma - U_1^T X Y^\dagger X U_1 \geq 0.$$

This condition can be satisfied by any $\alpha^2 \geq \lambda_{\max}(\Sigma^{-1/2} U_1^T X Y^\dagger X U_1 \Sigma^{-1/2})$. $\square$

Chapter 4

Existing Approaches

This chapter and the next deal with solution approaches to the RMP. In Section 4.1, we discuss special cases that can be solved efficiently, *e.g.*, analytically or in polynomial time. In general, however, the RMP is NP-hard, and there is little hope of finding the global minimum efficiently in all instances. What we look for, instead, are efficient *heuristics*. Section 4.2 lists the existing heuristic approaches organized into three groups.

4.1 Cases that can be solved efficiently

There are special cases of the RMP where the global minimum can be found efficiently. By *efficiently* we mean that a solution can be found in polynomial time, using, for example, Singular Value Decomposition (SVD) or convex optimization. In this section, we briefly describe some of these special cases and their solutions. For each problem, we also give its vector analog, *i.e.*, the corresponding cardinality minimization problem. Note that in these examples, we are concerned with finding the global minimum and *a* minimizer (which may not be unique). Although we will not attempt to do it here, in many cases it is possible to characterize all minimizers.

4.1.1 Cases solved via SVD

RMP with a norm-ball constraint

Perhaps the most commonly occurring RMP in the literature and in practice is the following:

$$\begin{array}{ll} \text{minimize} & \mathbf{Rank}\, X \\ \text{subject to} & \|X - A\| \leq \epsilon, \end{array} \tag{4.1}$$

where $X \in \mathbf{R}^{m\times n}$ is the optimization variable. The idea is to find a matrix with lowest rank that approximates A in matrix 2-norm within a tolerance of ϵ.

The solution to this problem is well known, and is based on the singular value decomposition (SVD). Here we show this in a simple way. Let $A = \sum_{i=1}^{\min\{m,n\}} \sigma_i u_i v_i^T$ be the SVD of A. Suppose $\mathbf{Rank}\, X \leq r$, which means $\dim \mathcal{N}(X) \geq n - r$. Since the v_i are orthonormal, we have $\dim \operatorname{span}\{v_1, \ldots, v_{r+1}\} = r + 1$. So the two subspaces intersect, *i.e.*, there is a vector $z \in \mathbf{R}^n$ in $\operatorname{span}\{v_1, \ldots, v_{r+1}\}$ such that $Xz = 0$. Thus,

$$\begin{aligned} (X - A)z &= -Az = -\textstyle\sum_{i=1}^{r+1} \sigma_i u_i v_i^T z, \\ \|(X - A)z\|^2 &= \textstyle\sum_{i=1}^{r+1} \sigma_i^2 (v_i^T z)^2 \geq \sigma_{r+1}^2 \|z\|^2, \end{aligned}$$

which yields $\|X - A\| \geq \sigma_{r+1}$. Equality holds for $X = \sum_{i=1}^{r} \sigma_i u_i v_i^T$. Therefore, the smallest possible rank in problem (4.1) is the smallest number r such that $\sigma_{r+1} \leq \epsilon$, and a matrix with this rank is given by the sum of the first r dyads in the SVD of A, *i.e.*,

$$X^\star = \sum_{i=1}^{r} \sigma_i u_i v_i^T.$$

The vector analog to the above problem is

$$\begin{array}{ll} \text{minimize} & \mathbf{card}\, x \\ \text{subject to} & \|x - a\|_\infty \leq \epsilon, \end{array} \tag{4.2}$$

where $x \in \mathbf{R}^n$. This problem has a trivial solution: for any $|a_i| \leq \epsilon$, set the corresponding x_i to zero.

The solution to problem (4.1) remains the same if the spectral norm is replaced by the Frobenius norm, or, in fact, by any unitarily invariant matrix norm [51, p. 449]. The vector analog of problem 4.1 with Frobenius norm is

$$\begin{array}{ll} \text{minimize} & \mathbf{card}(x) \\ \text{subject to} & \|x - a\|_2 \leq \epsilon, \end{array} \tag{4.3}$$

which means finding the sparsest vector in a given Euclidean ball. We introduce the (standard) notation $z_{[i]}$ for the ith largest component of the vector $z \in \mathbf{R}^n$, so that $z_{[1]} \geq z_{[2]} \geq \cdots \geq z_{[n]}$ is a permutation of $z_1, \ldots, z_n$. Using this notation we can describe the optimal solution of the problem (4.3) above as follows. The optimal value is the smallest r such that $\sum_{i=r+1}^{n} a_{[i]}^2 \leq \epsilon^2$, and an optimal x is obtained by taking $x_i = 0$ if a_i appears in the sum above, and $x_i = a_i$ otherwise.

In [66], a more general class of functions is described that can replace the norm in problem (4.1) and still yield the same optimal solution.

RMP with a single linear constraint

As another example of an RMP with an SVD-based solution, consider the problem

$$\begin{array}{ll} \text{minimize} & \mathbf{Rank}\, X \\ \text{subject to} & \mathbf{Tr}\, A^T X \geq b \\ & \|X\| \leq 1, \end{array} \tag{4.4}$$

where $X \in \mathbf{R}^{m \times n}$ is the optimization variable and $A \in \mathbf{R}^{m \times n}$ and $b \in \mathbf{R}$ are given.

Note that $\mathbf{Tr}\, A^T X$ defines an inner product in the space of $m \times n$ matrices:

$$\langle A, X \rangle = \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij} x_{ij} = \mathbf{Tr}(A^T X).$$

(The notation $A \bullet X$ is also sometimes used to denote this inner product.) The idea is to find a matrix of norm less than one that is close to A in the inner product sense, *i.e.*, has a large inner product with A.

The optimal solution to this problem can be described as follows. Let $X = U \Sigma V^T$ and $A = \bar{U} \bar{\Sigma} \bar{V}^T$ be the SVDs of X and A. Matrices Σ and $\bar{\Sigma}$ are $p \times p$ where $p = \min\{m, n\}$. By von Neumann's trace theorem [50, §3.3.21], we have $\mathbf{Tr}\, A^T X \leq \sum_{i=1}^{p} \bar{\sigma}_i \sigma_i$, where equality holds if $U = \bar{U}$ and $V = \bar{V}$. Since the objective function and the last constraint in (4.4) do not depend on U and V, we can choose $U = \bar{U}$ and $V = \bar{V}$ to make the trace term large. The problem is thus reduced to

$$\begin{array}{ll} \text{minimize} & \mathbf{card}\, \sigma \\ \text{subject to} & \sum_{i=1}^{p} \bar{\sigma}_i \sigma_i \geq b \\ & \sigma_i \leq 1, \quad i = 1, \ldots, n, \end{array} \tag{4.5}$$

where $\sigma \in \mathbf{R}^p$ denotes the vector of singular values of X. Let r be the smallest number of terms that satisfy $\sum_{i=1}^{r} \bar{\sigma}_i \geq b$. The optimal solution is then

$$X^\star = \bar{U} \Sigma \bar{V}^T, \quad \Sigma = \mathbf{diag}(\underbrace{1, \ldots, 1}_{r}, \underbrace{0, \ldots, 0}_{p-r}),$$

that is, $\sigma_i = 1$ for $i = 1, \ldots, r$ and $\sigma_i = 0$ for $i = r+1, \ldots, p$. Thus, the problem of minimizing the rank of a matrix subject to one linear constraint can be solved analytically. The solution is nontrivial if there is more than one linear constraint.

If $X = \mathbf{diag}\, x$, the above problem reduces to

$$\begin{array}{ll} \text{minimize} & \mathbf{card}\, x \\ \text{subject to} & a^T x \geq b, \quad |x_i| \leq 1. \end{array} \tag{4.6}$$

If $b \leq 0$, then $x = 0$ is optimal. Otherwise the problem is to get the weighted sum $\sum_i a_i x_i$ up to or above b using as few nonzero x_is as possible. Once we have $x_i \neq 0$, we might as well have $x_i = \text{sign}(a_i)$ in order to get the maximum contribution to the sum $\sum_i a_i x_i$ for a fixed cost in the objective. By making the ith component of x nonzero (which adds a cost of one to the objective) we can add the amount $|a_i|$ to the sum. This suggests the following method for obtaining the optimal solution. Order the $|a_i|$ from largest to smallest, and then take as many as needed (starting with the largest) until the sum is equal to or more than b. Take $x_i = \text{sign}(a_i)$ for each of these, and $x_i = 0$ for the others. This solution is in fact optimal, and the optimal value is the smallest r such that $\sum_{i=1}^{r} |a|_{[i]} \geq b$.

4.1.2 Cases that reduce to convex problems

Sometimes a rank minimization problem does not have an analytical solution, but can be reduced to a convex problem where the global minimum can be found efficiently.

This happens, for example, in the problem studied by Mesbahi and Papavassilopoulos in [68]. They consider minimizing the rank of a positive semidefinite matrix, subject to the constraint that a certain affine transformation of it is also positive semidefinite. If the feasible set has a special lattice structure, then there exists a feasible X that has the smallest rank as well as the smallest trace in the feasible set. Thus, the minimum-rank solution is obtained by solving a trace minimization problem, which is convex. As a specific example that has been of interest recently because of its possible relation to the problem of designing minimum order controllers, we consider the problem

$$\begin{array}{ll} \text{minimize} & \mathbf{Rank}\, X \\ \text{subject to} & X - \sum_{i=1}^{k} M_i X M_i^T \geq Q \\ & X \geq 0, \end{array} \tag{4.7}$$

where $X \in \mathbf{R}^{n\times n}$ is the variable, $M_i \in \mathbf{R}^{n\times n}$ and $Q \geq 0$ are given. It is shown in [68] that the feasible set of this problem has the desired lattice structure, and that the minimum-rank solution can be obtained by solving

$$\begin{array}{ll} \text{minimize} & \mathbf{Tr}\, X \\ \text{subject to} & X - \sum_{i=1}^{k} M_i X M_i^T \geq Q \\ & X \geq 0, \end{array} \tag{4.8}$$

which is a semidefinite program. Parrilo in [78] shows, using properties of cone-invariant LMIs, that the solution to this SDP can be obtained directly by solving a set of linear equations.

4.2 Heuristic methods

For small problem sizes, global optimization methods (*e.g.*, branch and bound [6]) may be applied to find the global minimum-rank solution. However, if the problem has more than a handful of variables, there is little hope of finding methods that can solve all instances of the RMP exactly and efficiently. What we look for, instead, are efficient *heuristics*. A good heuristic is a tractable method that in practice yields very low-rank solutions, although there is no guarantee on their optimality.

Several heuristics for the RMP have appeared in the literature, often in the context of a particular application (most often in the area of systems and control). In this section, we give an overview of the existing heuristics, organized into three groups. We list these methods and present the basic idea behind each, independent of the particular application areas they arise in.

4.2.1 Alternating projections method

The method of alternating projections is based on the fact that the sequence of orthogonal projections onto two closed, convex sets converges to a point in the intersection of the sets, if the intersection is non-empty [43]. It was first applied to statistical estimation and image restoration problems in [102, 103]. If the sets do not intersect, the sequence converges to a limit cycle, *i.e.*, a periodic iteration between two points in the disjoint sets. The distance between these two points gives the shortest distance between the sets. The hyperplane passing through the midpoint of these points is a separating hyperplane and yields a proof of infeasibility for the problem.

If one or more of the sets are non-convex, convergence to the intersection is no longer guaranteed. In this case, we can have a situation where the sets intersect but the sequence of projections converges to a limit cycle, as depicted in figure 4.1. However, *local* convergence is still guaranteed and the method may be used as a *heuristic*.

Here we mention how the alternating projections method can be applied to the RMP (this approach is used in [9] for the low-order controller design problem). We first fix the desired rank r. The goal is to find a matrix in the intersection of the following two sets, or determine that the intersection is empty: (i) the set of matrices of rank r, and (ii) the constraint set $\mathcal{C}$. Note that the first set is nonconvex, and therefore convergence to the intersection is not guaranteed. Orthogonal projections onto these sets are described as follows.

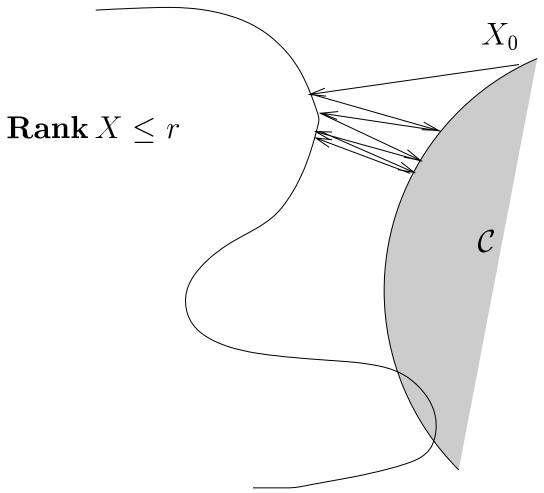

Figure 4.1: Illustration of the alternating projections method for the RMP.

Projection onto the set of matrices of rank r, *i.e.*, finding the closest rank r matrix to the current iterate X_{k-1}, can be done by solving

$$\begin{array}{ll} \text{minimize} & \|X - X_{k-1}\| \\ \text{subject to} & \mathbf{Rank}\, X \leq r. \end{array}$$

Commonly used norms are matrix 2-norm and Frobenius norm. This problem can be solved via SVD and keeping the first r dyads. We denote the solution by $\tilde{X}_k$. Projection onto the constraint set $\mathcal{C}$ can be done by minimizing the distance from $\tilde{X}_k$ to the set $\mathcal{C}$,

$$\begin{array}{ll} \text{minimize} & \|X - \tilde{X}_k\| \\ \text{subject to} & X \in \mathcal{C}, \end{array} \tag{4.9}$$

which is a convex optimization problem. Note that the norm used in the previous step should be used in this step. In summary, given a desired value of rank r, we use the following algorithm to check whether there is any $X \in \mathcal{C}$ such that $\mathbf{Rank}\, X \leq r$:

- Choose $X_0 \in \mathcal{C}$. Set $k = 1$.
- **repeat**

$$\begin{array}{rcl} \tilde{X}_k & = & \sum_{i=1}^{r} \lambda_i u_i v_i^T, \quad \text{where } X_{k-1} = U\Sigma V^T, \\ X_k & = & \operatorname{argmin}_{X\in\mathcal{C}} \|X - \tilde{X}_k\|, \\ e_k & = & \|X_k - \tilde{X}_k\|, \end{array}$$

until $|e_k - e_{k-1}| \le \epsilon$.

If the iterations converge to a feasible point, we stop. If they converge to a limit cycle, either the problem is infeasible or the method has failed to detect the feasibility. This ambiguity is due to the fact that the method searches for a feasible point only locally. Thus, using a different initial point a different result may be obtained, which means the choice of a suitable initial point is crucial.

This process can be repeated to check the feasibility of other values of rank. We can vary r from 1 to $n-1$, or use bisection on r (*i.e.*, iteratively halve the interval).

See [41, chapter 10] and references therein for a detailed discussion of the alternating projection method and its variations and their application to low-order control design.

We now briefly comment about two drawbacks of the alternating projections method. A few quantitative examples that illustrate these points are given in Section 5.3 of Chapter 5. In general, this method is known to have slow convergence [85] [1]. Note that in each iteration, in addition to an SVD, we need to solve problem (4.9). In some special cases, projection onto $\mathcal{C}$ has a simple analytical expression (see [85]). In these cases, we can afford a large number of iterations since the computation required per iteration is very low; but in general, each iteration involves solving a convex

[1]A variation of the method, called *directional alternating projections*, has improved convergence properties. However, the number of iterations required can still be quite high as suggested by the numerical examples given in [85, chapter 10].

problem, *e.g.*, a semidefinite program.

Another drawback of the method is that the result is highly dependent on the choice of the initial point, and in many applications, finding a good initial point is nontrivial.

4.2.2 Interior-point-based methods

Consider a positive semidefinite RMP, *i.e.*, a special case of the RMP with the extra constraint that $X \geq 0$. Reference [20] proposes two heuristics for this problem that use ideas from interior point methods for convex optimization [73].

The first heuristic, called analytic anti-centering, is based on the properties of convex logarithmic barrier functions used in interior point (IP) methods [73]. These barrier functions have the property that they grow to infinity as the boundary of the feasible set is approached. Minimization of a log-barrier function using the Newton method produces a point in the interior of the feasible set, known as the analytic center. Now note that any rank-deficient solution to the positive semidefinite RMP must lie on the boundary of the semidefinite cone. Hence, the analytic anti-centering approach takes steps in the reverse Newton direction, in order to maximize the log-barrier function. This tends to produce points that are on the boundary of the feasible set, and hence rank deficient. Since this approach involves the maximization of a convex function, the solutions are not necessarily global optima.

The second heuristic, called potential reduction, is based on interior point potential reduction methods [73]. The idea is based on the intuitive notion that minimizing the determinant of a semidefinite matrix tends to produce low-rank solutions. The rank objective in the positive semidefinite RMP is replaced by the determinant objective, $\det X$. The resulting optimization problem is then solved using standard potential reduction [73], exactly as if the objective were convex: the Newton method is used to minimize a potential function, constructed from a weighted logarithm of the objective ($\det X$) and a log-barrier function for the constraints. Of course, now that the $\det X$ objective is nonconvex, there is no guarantee that minimizing the potential function will produce a global optimum. This approach is easier to implement than analytic anti-centering and is also reported to have better performance [20].

These heuristics, as given, are applicable only to positive semidefinite RMPs. However, they can be extended to handle general non-PSD matrices via the semidefinite embedding lemma that we introduced in Chapter 3.

The main drawback of these methods is that the result is highly sensitive to the choice of the initial point. The initial point is typically chosen in the vicinity of the analytic center of the feasible region. The iterations may follow a completely different path to a different point on the boundary if the initial point is slightly changed. See reference [20] for more details and examples, and for the application of these methods to low-order control design.

4.2.3 Factorization, coordinate descent and linearization methods

The idea behind factorization methods is that $\mathbf{Rank}(X) \leq r$ if and only if X can be factored as $X = FG^T$, where $F \in \mathbf{R}^{m\times r}$ and $G \in \mathbf{R}^{n\times r}$. That is,

$$\mathbf{Rank}(X) \leq r \iff \text{ there exists } F \in \mathbf{R}^{m\times r},\ G \in \mathbf{R}^{n\times r} \text{ such that } X = FG^T.$$

For each given r, we check if there exists a feasible X of rank less than or equal to r by checking if any $X \in \mathcal{C}$ can be factored as above.

The expression $X = FG^T$ is not convex in X, F, and G simultaneously, but it is convex in (X, F) when G is fixed, and convex in (X, G) when F is fixed.

Various heuristics can be applied to handle this non-convex equality constraint. We consider the following simple heuristic: Fix F and G one at a time and iteratively solve a convex problem at each step. This can be expressed as

- Choose $F_0 \in \mathbf{R}^{m\times r}$. Set $k = 1$.
- **repeat**

$$
\begin{aligned}
(\tilde{X}_k, G_k) &= \operatorname*{argmin}_{X\in\mathcal{C},\ G\in\mathbf{R}^{n\times r}} \|X - F_{k-1}G^T\|_F \\
(X_k, F_k) &= \operatorname*{argmin}_{X\in\mathcal{C},\ F\in\mathbf{R}^{m\times r}} \|X - FG_k^T\|_F \\
e_k &= \|X_k - F_kG_k^T\|_F,
\end{aligned}
$$

until $e_k \leq \epsilon$, or e_k and e_{k-1} are approximately equal.

This is a coordinate descent method, since some variables (*i.e.*, coordinates) are fixed during each minimization step. The errors e_k form a monotonically non-increasing sequence (since at each minimization step, the previous minimizers are still feasible). This sequence is not guaranteed to converge to the global minimum of the error $\|X - FG^T\|_F$ as a function of X, F, and G; thus, similar to the methods described in the previous sections, it is not guaranteed to find an X with rank r even if one exists.

Another heuristic to handle the non-convex constraint $X = FG^T$ is to linearize this equation in F and G. Assuming the perturbations δF, δG are small enough so that the second order term is negligible, we get $X = FG^T + F\delta G^T + \delta F G^T$. This constraint can be handled easily since it is linear in both δF and δG. The method is useful if the initial choice for FG^T is close enough to a rank r matrix for the small perturbations assumption to be valid. This method has been used in BMI problems that come up in low-authority controller design [45].

Some other heuristics, similar to the ones described here, have been applied to the problem of reduced order controller design in the control literature (see Chapter 6 for more details on this problem). This problem has a particular structure, allowing for different choices for the variables in a coordinate descent or linearization method. For example, the dual iteration method in [54] and the successive minimization approach in [85] are coordinate descent methods applied to this problem, and [25] gives a linearization method based on a cone-complementarity formulation.

4.2.4 Summary and Remarks

In this section, we gave an overview of existing heuristics for the RMP. These heuristics were grouped as: alternating projections, interior-point-based methods, and factorization methods. We presented the basic idea behind each group of methods, and pointed out some of their properties. Note that, as is typical for any local optimization method, all the methods mentioned above require a suitable initial point and are sensitive to it, that is, the result may be significantly affected by choosing a slightly different initial point. Although in some cases special properties of the problem can be exploited to pick a suitable starting point, in general this is a non-trivial task.

In Chapter 5, we present other heuristics for the RMP, that can be applied to any general RMP, do not require a user-specified initial point, and have several other benefits. They also perform very well in practice. In Section 5.3, we give several examples to illustrate how these heuristics work compared to the methods mentioned in this chapter.

Chapter 5

Trace and Log-det Heuristics

In the first half of this chapter, we focus on the *trace heuristic*. We start from the well-known fact that minimizing the trace of a PSD matrix over a convex set tends to yield a low-rank solution, and then we develop a new, general heuristic that can handle *any* matrix. We refer to this general method as the *nuclear norm* heuristic. In the second half of the chapter, we introduce and discuss a related heuristic that we refer to as the *log-det heuristic*. We then give illustrative examples to demonstrate how these heuristics work compared to the ones described in Chapter 4.

5.1 Trace heuristic

In this section, we first state the trace heuristic for the positive semidefinite case. We use the semidefinite embedding lemma of Chapter 3 to extend this heuristic to the general case. We then show that the resulting general heuristic is equivalent to minimizing the sum of the singular values of the matrix. This quantity is a matrix norm called the *nuclear norm*. For the special case of minimizing the cardinality of a vector (*i.e.*, the CMP), we show that this heuristic reduces to minimizing the ℓ_1 norm of the vector. Furthermore, we provide insight into the nuclear norm heuristic by showing that, in fact, it minimizes the convex envelope of the rank function over a bounded set of matrices.

5.1.1 Positive semidefinite case

A well-known heuristic for the RMP when the variable $X \in \mathbf{R}^{n\times n}$ is positive semidefinite is to replace the rank objective in (2.1) with the trace of X and solve

$$\begin{array}{ll} \text{minimize} & \mathbf{Tr}\, X \\ \text{subject to} & X \in \mathcal{C} \\ & X \geq 0. \end{array} \tag{5.1}$$

One way to see why this heuristic works is to note that $\mathbf{Tr}\, X = \sum_{i=1}^{n} \lambda_i(X)$, which is the same as $\|\lambda(X)\|_1 = \sum_{i=1}^{n} |\lambda_i(X)|$ for a PSD matrix where the eigenvalues are non-negative. It is known that to obtain a sparse vector, minimizing the ℓ_1-norm of the vector is an effective heuristic (see Section 5.1.7). Thus, minimizing the ℓ_1-norm of $\lambda(X)$ renders many of the eigenvalues as zero, resulting in a low-rank matrix.

The trace heuristic has been used in many applications; see for example [68, 77, 78]. Its popularity stems from the fact that problem (5.1) is a convex optimization problem, which can be solved very efficiently and reliably in practice.

5.1.2 Symmetric non-PSD case

We can extend the trace heuristic to handle problems where X is symmetric but not necessarily positive semidefinite. Intuitively, the extension is to minimize the sum of absolute values of the eigenvalues, *i.e.*,

$$\begin{array}{ll}\text{minimize} & \sum_{i=1}^{n} |\lambda_i(X)| \\ \text{subject to} & X \in \mathcal{C}.\end{array} \tag{5.2}$$

It turns out that this problem can be written as an SDP and thus readily solved. It is equivalent to the SDP

$$\begin{array}{ll}\text{minimize} & \mathbf{Tr}\, X_+ + \mathbf{Tr}\, X_- \\ \text{subject to} & X = X_+ - X_- \\ & X_+ \geq 0,\ X_- \geq 0 \\ & X \in \mathcal{C}.\end{array} \tag{5.3}$$

To see this, note that the function $\sum_{i=1}^{n} |\lambda_i(X)|$ is convex in X (in fact, it is a matrix norm; see Section 5.1.3 for details). If $X = X_+ - X_-$, convexity implies that

$$\begin{array}{rcl}\sum_i |\lambda_i(X)| & \leq & \frac{1}{2}(\sum_i |\lambda_i(X_+)| + \sum_i |\lambda_i(X_-)|) \\ & = & \frac{1}{2}(\sum_i \lambda_i(X_+) + \sum_i \lambda_i(X_-)) \\ & = & \frac{1}{2}(\mathbf{Tr}\, X_+ + \mathbf{Tr}\, X_-).\end{array} \tag{5.4}$$

We show that there exist feasible X_+ and X_- for which the above inequality is tight. Let $X = Q\Lambda Q^T$ be the eigenvalue decomposition of X. We group the non-negative and negative eigenvalues as the diagonal entries of Λ_+ and Λ_-, respectively. We group the corresponding eigenvectors as Q_+ and Q_-, to obtain

$$X = Q\Lambda Q^T = [Q_+\ Q_-] \begin{bmatrix} \Lambda_+ & 0 \\ 0 & \Lambda_- \end{bmatrix} \begin{bmatrix} Q_+^T \\ Q_-^T \end{bmatrix}.$$

The inequality in (5.4) is tight if we choose $X_+ = Q_+\Lambda_+Q_+^T$ and $X_- = Q_-\Lambda_-Q_-^T$, which are feasible for problem (5.3). Thus, we have shown that (5.3) is equivalent to

(5.2).

5.1.3 General case: nuclear norm heuristic

As it stands, the trace heuristic (5.1) can be applied only to problems where the matrix whose rank is to be minimized is positive semidefinite. The extension to problems where X is non-PSD, or more generally, non-square, is not obvious, as the trace is not even defined for non-square matrices. Nevertheless, there are various important applications of the RMP where the variable is not square. Two examples that are described in Chapter 6 are system realization with time-domain specifications and low-order system approximation. A natural question is whether this simple and effective heuristic can be extended to handle the general RMP.

The answer is indeed yes. The lemma we introduced in Chapter 3 enables us to embed any general RMP

$$\begin{array}{ll} \text{minimize} & \mathbf{Rank}\, X \\ \text{subject to} & X \in \mathcal{C}, \end{array} \tag{5.5}$$

where $X \in \mathbf{R}^{m\times n}$ is the optimization variable and $\mathcal{C}$ is a convex set, in a larger, positive semidefinite one:

$$\begin{array}{ll} \text{minimize} & \mathbf{Rank}\,\mathbf{diag}(Y, Z) \\ \text{subject to} & \begin{bmatrix} Y & X \\ X^T & Z \end{bmatrix} \geq 0 \\ & X \in \mathcal{C}, \end{array} \tag{5.6}$$

where $Y \in \mathbf{R}^{m\times m}$ and $Z \in \mathbf{R}^{n\times n}$ are additional (slack) variables. Since the arguments of the rank function in (5.6), Y and Z, are known to be positive semidefinite, direct application of the trace heuristic in (5.1) yields

$$\begin{array}{ll} \text{minimize} & \mathbf{Tr}\,\mathbf{diag}(Y, Z) \\ \text{subject to} & \begin{bmatrix} Y & X \\ X^T & Z \end{bmatrix} \geq 0 \\ & X \in \mathcal{C}, \end{array} \tag{5.7}$$

which is again a convex optimization problem in variables X, Y and Z, and hence can be solved efficiently.

Next, we derive an equivalent form of (5.7) that provides insight into this heuristic and its relation to the original RMP. We show that (5.7) is equivalent to

$$\begin{array}{ll} \text{minimize} & \|X\|_* \\ \text{subject to} & X \in \mathcal{C}, \end{array} \tag{5.8}$$

where $\|X\|_* = \sum_{i=1}^{\min\{m,n\}} \sigma_i(X)$ is called the *nuclear norm* or the *Ky-Fan n-norm* of X; see, *e.g.*, [50]. This norm is the dual of the spectral (or the maximum singular value) norm. Since the spectral norm is denoted by $\|\cdot\|$, we use $\|\cdot\|_*$ to denote its dual. The equivalence of problems (5.7) and (5.8) results from the following lemma.

Lemma 2 *For $X \in \mathbf{R}^{m\times n}$ and $t \in \mathbf{R}$, we have $\|X\|_* \leq t$ if and only if there exist matrices $Y \in \mathbf{R}^{m\times m}$ and $Z \in \mathbf{R}^{n\times n}$ such that*

$$\begin{bmatrix} Y & X \\ X^T & Z \end{bmatrix} \geq 0, \qquad \mathbf{Tr}\, Y + \mathbf{Tr}\, Z \leq 2t. \tag{5.9}$$

Proof: ($\Longleftarrow$) Let Y and Z satisfy the relations (5.9) above, and let $X = U\Sigma V^T$

be the SVD of X. Here, Σ is of size $r \times r$, where r is the rank of X. We have

$$\mathbf{Tr} \begin{bmatrix} UU^T & -UV^T \\ -VU^T & VV^T \end{bmatrix} \begin{bmatrix} Y & X \\ X^T & Z \end{bmatrix} \geq 0,$$

since the trace of the product of two PSD matrices is always non-negative. This yields

$$\mathbf{Tr}\, UU^T Y - \mathbf{Tr}\, UV^T X^T - \mathbf{Tr}\, VU^T X + \mathbf{Tr}\, VV^T Z \geq 0. \tag{5.10}$$

Since the columns of U are orthonormal, we can add more columns to complete them to a full basis; *i.e.*, there exists $\tilde{U}$ such that $[U\ \tilde{U}][U\ \tilde{U}]^T = I$, or $UU^T + \tilde{U}\tilde{U}^T = I$. Since $\mathbf{Tr}\, \tilde{U}\tilde{U}^T Y \geq 0$, we have

$$\mathbf{Tr}\, UU^T Y \leq \mathbf{Tr}(UU^T + \tilde{U}\tilde{U}^T)Y = \mathbf{Tr}\, Y.$$

Similarly, for V we have $\mathbf{Tr}\, VV^T Z \leq \mathbf{Tr}\, Z$. Also, $\mathbf{Tr}\, VU^T X = \mathbf{Tr}\, V\Sigma V^T = \mathbf{Tr}\, \Sigma$, and $\mathbf{Tr}\, UV^T X^T = \mathbf{Tr}\, U\Sigma U^T = \mathbf{Tr}\, \Sigma$. If we substitute for all the terms in (5.10), we get

$$\begin{aligned} &\mathbf{Tr}\, Y + \mathbf{Tr}\, Z - 2\,\mathbf{Tr}\, \Sigma \geq 0, \\ &\mathbf{Tr}\, \Sigma \leq \tfrac{1}{2}(\mathbf{Tr}\, Y + \mathbf{Tr}\, Z), \\ &\mathbf{Tr}\, \Sigma = \|X\|_* \leq t. \end{aligned}$$

($\Longrightarrow$) Suppose $\|X\|_* \leq t$. We show Y and Z can be chosen to satisfy the relations (5.9). If $Y = U\Sigma U^T + \gamma I$ and $Z = V\Sigma V^T + \gamma I$, then

$$\mathbf{Tr}\, Y + \mathbf{Tr}\, Z = 2\,\mathbf{Tr}\, \Sigma + \gamma(p+q) = 2\|X\|_* + \gamma(p+q),$$

so if we choose $\gamma = \frac{2(t - \|X\|_*)}{p+q}$, we have $\mathbf{Tr}\, Y + \mathbf{Tr}\, Z = 2t$.

Also note that

$$\begin{aligned}
\begin{bmatrix} Y & X \\ X^T & Z \end{bmatrix} &= \begin{bmatrix} U\Sigma U^T & U\Sigma V^T \\ V\Sigma U^T & V\Sigma V^T \end{bmatrix} + \gamma \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix} \\
&= \begin{bmatrix} U \\ V \end{bmatrix} \Sigma \begin{bmatrix} U^T & V^T \end{bmatrix} + \gamma I,
\end{aligned}$$

which is PSD. □

In other words, the condition $\|X\|_* \leq t$ can be represented as an LMI. This observation was made also in [92, §3.1]. It is now straightforward to show that the generalized trace heuristic (5.7) and the nuclear norm minimization problem (5.8) are equivalent. We simply write (5.7) as

$$\begin{array}{ll}
\text{minimize} & t \\
\text{subject to} & \mathbf{Tr}\, Y + \mathbf{Tr}\, Z \leq 2t \\
& \begin{bmatrix} Y & X \\ X^T & Z \end{bmatrix} \geq 0 \\
& X \in \mathcal{C},
\end{array}$$

with variables X, Y, Z and t. Then we apply the above lemma to get

$$\begin{array}{ll}
\text{minimize} & t \\
\text{subject to} & \|X\|_* \leq t \\
& X \in \mathcal{C},
\end{array}$$

with variables X and t. This is equivalent to (5.8). In the next section we show how this problem is related to the original RMP.

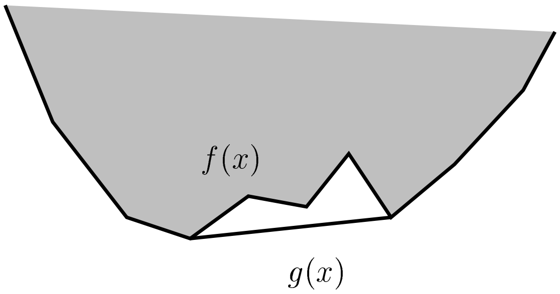

Figure 5.1: Illustration of convex envelope of a function. $g(x)$ is the convex envelope of $f(x)$.

5.1.4 Convex envelope of rank

In this section we explore the relation between the RMP (5.5) and the nuclear norm (or the generalized trace) heuristic (5.8) in more detail. This leads to an interesting interpretation of the nuclear norm heuristic: in effect, this heuristic minimizes the *convex envelope* of the rank function over a bounded set.

The *convex envelope* of $f : \mathcal{C} \to \mathbf{R}$ is defined as the largest convex function g such that $g(x) \leq f(x)$ for all $x \in \mathcal{C}$; see, *e.g.*, [49]. This means that among all convex functions, g is the one that is closest (pointwise) to f.

In situations such as problem (5.5) where the objective function is non-convex, its convex envelope can serve as a tractable convex approximation that can be minimized efficiently. The minimum of the convex envelope can then serve as a lower bound on the true minimum, and the minimizing argument can serve as an initial point for a more complicated non-convex local search, if needed. This is discussed in Section 5.2. The following theorem gives the convex envelope of the rank function over the set of matrices with bounded norm.

Theorem 1 *On the set* $\mathcal{S} = \{X \in \mathbf{R}^{m\times n} \mid \|X\| \leq 1\}$, *the convex envelope of the function* $\phi(X) = \mathbf{Rank}\, X$ *is* $\phi_{\text{env}}(X) = \|X\|_* = \sum_{i=1}^{\min\{m,n\}} \sigma_i(X)$.

Thus, the nuclear norm heuristic in effect minimizes the convex envelope of the rank function. The proof of the theorem is given in the next section.

Theorem 1 has the following implications for the RMP and the heuristic (5.8). Suppose the feasible set is bounded by M; *i.e.*, for all $X \in \mathcal{C}$, we have $\|X\| \le M$. The convex envelope of $\mathbf{Rank}\, X$ on $\{X \in \mathbf{R}^{m\times n} \mid \|X\| \le M\}$ is given by $\frac{1}{M}\|X\|_*$, so $\mathbf{Rank}\, X \ge \frac{1}{M}\|X\|_*$ for all $X \in \mathcal{C}$. Let p_{rmp} denote the optimal value of the rank minimization problem (5.5) and p_{tr} the optimal value of the trace minimization problem (5.8). Then

$$p_{\mathrm{rmp}} \ge \frac{1}{M}\, p_{\mathrm{tr}}.$$

In other words, by solving the heuristic problem, we obtain a lower bound on the optimal value of the original problem (provided we can identify a bound M on the feasible set). Note that the nuclear norm is the *tightest* convex lower approximation to the rank function over the set $\mathcal{S}$; thus among all convex approximations, it yields the tightest global lower bound on rank.

For comparison purposes, here we examine another approach to extending the trace heuristic to the general case. This approach yields a different convex approximation to rank, which does not have the convex envelope property. Perhaps the easiest way to relate the rank of a general, non-square matrix to that of a PSD matrix is to observe that $\mathbf{Rank}\, X = \mathbf{Rank}(XX^T)$. Since XX^T is positive semidefinite, one can directly apply the trace heuristic to get

$$\begin{array}{ll} \text{minimize} & \mathbf{Tr}(XX^T) \\ \text{subject to} & X \in \mathcal{C}. \end{array}$$

Here $\mathbf{Tr}(XX^T) = \sum_i \sigma_i(X)^2$, the squared Frobenius norm of X, serves as a convex approximation to rank. Note that this is the ℓ_2-norm of the vector of singular values. It is known that minimizing the ℓ_2-norm of a vector, unlike the ℓ_1-norm, often does not yield a sparse vector. Therefore we cannot expect to obtain a sparse set of singular values or a low-rank matrix using the Frobenius norm. Minimizing $\mathbf{Tr}(XX^T)$ over $\mathcal{C}$ can also be written as the SDP

$$\begin{array}{ll} \text{minimize} & \mathbf{Tr}\, Y \\ \text{subject to} & \begin{bmatrix} Y & X \\ X^T & I \end{bmatrix} \geq 0 \\ & X \in \mathcal{C}. \end{array}$$

Comparing this problem to (5.7) shows that these two problems are the same, except that the variable Z in (5.7) is replaced by the identity, thus reducing the degrees of freedom in the optimization. We conclude that the convex envelope property of the nuclear norm plays an important role in the effectiveness of heuristic (5.8).

5.1.5 Proof of the convex envelope theorem

We now prove Theorem 1, using the notion of *conjugate functions*. The conjugate f^* of a function $f : \mathcal{C} \to \mathbf{R}$, where $\mathcal{C} \subseteq \mathbf{R}^n$, is defined as

$$f^*(y) = \sup\{\langle y, x\rangle - f(x) \mid x \in C\},$$

where $\langle y, x\rangle$ denotes the inner product in $\mathbf{R}^n$. A basic result of convex analysis is that the conjugate of the conjugate, f^{**}, is the convex envelope of the function f, provided some technical conditions (which are valid here) hold. See theorem 1.3.5 in [49] for more details.

Part 1. *Computing* ϕ^*: The conjugate of the rank function ϕ, on the set of matrices with spectral norm less than or equal to one, is

$$\phi^*(Y) = \sup_{\|X\| \leq 1} (\mathbf{Tr}\, Y^T X - \phi(X)), \tag{5.11}$$

where $\langle Y, X\rangle = \mathbf{Tr}\, Y^T X$ is the inner product in $\mathbf{R}^{m\times n}$. Let $q = \min\{m, n\}$. By von Neumann's trace theorem [50],

$$\mathbf{Tr}\, Y^T X \le \sum_{i=1}^{q} \sigma_i(Y)\sigma_i(X), \tag{5.12}$$

where $\sigma_i(X)$ denotes the ith largest singular value of X. Given Y, equality in (5.12) is achieved if U_X and V_X are chosen equal to U_Y and V_Y, respectively, where $X = U_X\Sigma_X V_X^T$ and $Y = U_Y\Sigma_Y V_Y^T$ are the SVDs of X and Y. The term $\phi(X)$ in (5.11) is independent of U_X and V_X, therefore to find the supremum, we pick $U_X = U_Y$ and $V_X = V_Y$ to maximize the first term. It follows that

$$\phi^*(Y) = \sup_{\|X\|\le 1} \left(\sum_{i=1}^{q} \sigma_i(Y)\sigma_i(X) - \mathbf{Rank}\, X \right).$$

If $X = 0$, we have $\phi^*(Y) = 0$ for all Y. If $\mathbf{Rank}\, X = r$, $1 \le r \le q$, then $\phi^*(Y) = \sum_{i=1}^{r} \sigma_i(Y) - r$. Hence, $\phi^*(Y)$ can be expressed as

$$\phi^*(Y) = \max\left\{0,\ \sigma_1(Y) - 1, \ldots,\ \sum_{i=1}^{r}\sigma_i(Y) - r, \ldots,\ \sum_{i=1}^{q}\sigma_i(Y) - q\right\}.$$

The largest term in this set is the one that sums all *positive* $(\sigma_i(Y) - 1)$ terms. We conclude that

$$\begin{aligned} \phi^*(Y) &= \begin{cases} 0 & \|Y\| \le 1 \\ \sum_{i=1}^{r}\sigma_i(Y) - r & \sigma_r(Y) > 1 \text{ and } \sigma_{r+1}(Y) \le 1 \end{cases} \\ &= \sum_{i=1}^{q} \left(\sigma_i(Y) - 1\right)_+, \end{aligned} \tag{5.13}$$

where a_+ denotes the positive part of a, *i.e.*, $a_+ = \max\{0, a\}$.

Part 2. *Computing* ϕ^{**}: We now find the conjugate of ϕ^*, defined as

$$\phi^{**}(Z) = \sup_Y (\mathbf{Tr}\, Z^T Y - \phi^*(Y)),$$

for all $Z \in \mathcal{C}^{m\times n}$. As before, we choose $U_Y = U_Z$ and $V_Y = V_Z$ to get

$$\phi^{**}(Z) = \sup_Y \left(\sum_{i=1}^{q} \sigma_i(Z)\sigma_i(Y) - \phi^*(Y) \right).$$

We will consider two cases, $\|Z\| > 1$ and $\|Z\| \leq 1$:

If $\|Z\| > 1$, we can choose $\sigma_1(Y)$ large enough so that $\phi^{**}(Z) \to \infty$. To see this, note that in

$$\phi^{**}(Z) = \sup_Y \left(\sum_{i=1}^{q} \sigma_i(Z)\sigma_i(Y) - (\sum_{i=1}^{r} \sigma_i(Y) - r) \right),$$

the coefficient of $\sigma_1(Y)$ is $(\sigma_1(Z) - 1)$ which is positive.

Now let $\|Z\| \leq 1$. If $\|Y\| \leq 1$, then $\phi^*(Y) = 0$ and the supremum is achieved for $\sigma_i(Y) = 1$, $i = 1, \ldots, q$, yielding

$$\phi^{**}(Z) = \sum_{i=1}^{q} \sigma_i(Z) = \|Z\|_*.$$

We now show that if $\|Y\| > 1$, the argument of the sup is always smaller than the value given above. By adding and subtracting the term $\sum_{i=1}^{q} \sigma_i(Z)$ and rearranging the terms, we get

$$\begin{aligned}
&\sum_{i=1}^{q} \sigma_i(Y)\sigma_i(Z) - \sum_{i=1}^{r} (\sigma_i(Y) - 1) \\
&= \sum_{i=1}^{q} \sigma_i(Y)\sigma_i(Z) - \sum_{i=1}^{r} (\sigma_i(Y) - 1) - \sum_{i=1}^{q} \sigma_i(Z) + \sum_{i=1}^{q} \sigma_i(Z) \\
&= \sum_{i=1}^{r} (\sigma_i(Y) - 1)(\sigma_i(Z) - 1) + \sum_{i=r+1}^{q} (\sigma_i(Y) - 1)\sigma_i(Z) + \sum_{i=1}^{q} \sigma_i(Z) \\
&< \sum_{i=1}^{q} \sigma_i(Z).
\end{aligned}$$

The last inequality holds because the first two sums on the third line always have a negative value.

In summary, we have shown

$$\phi^{**}(Z) = \|Z\|_*$$

over the set $\{Z \mid \|Z\| \leq 1\}$. Thus, over this set, $\|Z\|_*$ is the convex envelope of the function $\mathbf{Rank}\, Z$. $\square$

5.1.6 Intuitive interpretation

The following provides a simple way to visualize and understand the nuclear norm heuristic. The rank of a matrix X in $\mathbf{R}^{m\times n}$ equals the number of singular values greater than zero, which can be written as

$$\mathbf{Rank}\, X \quad = \quad \sum_{i=1}^{\min\{m,n\}} I_+(\sigma_i(X)), \tag{5.14}$$

where the function I_+ is the indicator function for the positive reals $\mathbf{R}_+$:

$$I_+(x) \triangleq \begin{cases} 1 & x > 0, \\ 0 & x \leq 0. \end{cases}$$

Figure 5.2 shows the rank function and its convex envelope for the case of a 1×1 (scalar) matrix. Note that in this case X has only one singular value $\sigma(X) = |X|$, and $\mathbf{Rank}\, X = I_+(|X|)$. The figure suggests that the convex envelope of rank over the set of matrices bounded by M may be obtained by replacing each $I_+(\sigma_i(X))$ term

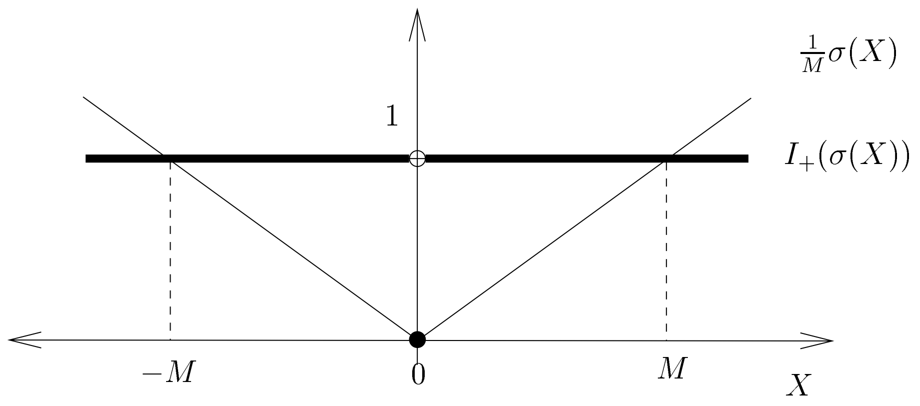

Figure 5.2: Basic idea behind the convex envelope approximation of $\mathbf{Rank}\, X$: when X is a 1×1 (scalar) matrix, it has only one singular value $\sigma(X) = |X|$; then $\mathbf{Rank}\, X = I_+(|X|)$, which has the convex envelope $\frac{1}{M}|X| = \frac{1}{M}\sigma(X)$.

in (5.14) by $\frac{1}{M}\sigma_i(X)$ to get

$$\phi_{\text{env}}(X) = \frac{1}{M} \sum_{i=1}^{\min\{m,n\}} \sigma_i(X) = \frac{1}{M}\|X\|_*, \tag{5.15}$$

which agrees with the results in Section 5.1.4.

5.1.7 Vector case: ℓ_1-norm minimization

We now consider the special case where the matrix X is diagonal, *i.e.*, $X = \mathbf{diag}\, x$, $x \in \mathbf{R}^n$. Recall from Chapter 3 that in this case, the RMP reduces to the cardinality minimization problem (CMP).

Consider the CMP in the form of (3.4). Applying the trace heuristic (5.1) to this problem yields

$$\begin{array}{ll} \text{minimize} & \mathbf{Tr}\,\mathbf{diag}\,y \\ \text{subject to} & \begin{bmatrix} \mathbf{diag}\,y & \mathbf{diag}\,x \\ \mathbf{diag}\,x & \mathbf{diag}\,y \end{bmatrix} \geq 0 \\ & \mathbf{diag}\,x \in \mathcal{C}. \end{array}$$

Using the same discussion as in Section 3.2.1 of Chapter 3, we can simplify this problem to

$$\begin{array}{ll} \text{minimize} & \sum_{i=1}^n y_i \\ \text{subject to} & |x_i| \leq y_i \\ & \mathbf{diag}\,x \in \mathcal{C}, \end{array}$$

which is equivalent to

$$\begin{array}{ll} \text{minimize} & \sum_{i=1}^n |x_i| \\ \text{subject to} & x \in \bar{\mathcal{C}}, \end{array} \tag{5.16}$$

where $\bar{\mathcal{C}}$ is the pre-image of $\mathcal{C}$ under the mapping $x \to \mathbf{diag}\,x$, and $\|x\|_1 = \sum_i |x_i|$ denotes the ℓ_1-norm of x. Problem (5.16) is the well-known ℓ_1-norm heuristic for cardinality minimization and has been used in various applications; *e.g.*, actuator/sensor placement in low-authority control [44], wavelet decomposition of signals using basis pursuit [17, 16], and robust estimators in statistics [52].

Not surprisingly, $\|x\|_1$ is also the convex envelope of $\mathbf{card}\,x$ over $\{x \mid \|x\|_\infty \leq 1\}$. The nuclear norm heuristic can thus be considered as an extension of the ℓ_1 heuristic to the matrix case. In Chapter 6, Section 6.5, we apply the ℓ_1 heuristic to the problem of portfolio optimization with fixed costs.

5.2 Log-det heuristic

We now discuss another heuristic for the RMP that we refer to as the *log-det* heuristic. We first state the heuristic for the case of positive semidefinite matrices, where we use the log-det function as a smooth surrogate for rank, and propose an iterative linearization and minimization scheme for finding a local minimum. We show that the resulting heuristic can be viewed as a refinement of the trace heuristic. We then apply the log-det heuristic to the general (non-square) case using the semidefinite embedding lemma of Chapter 3, and give an intuitive justification for the heuristic. We then show how applying this heuristic to diagonal matrices yields a new iterative heuristic for the CMP.

5.2.1 Positive semidefinite case

Consider the RMP with $X \in \mathbf{R}^{n\times n}$, $X \geq 0$. The log-det heuristic can be described as follows: rather than solving the RMP, use the function $\log\det(X+\delta I)$ as a *smooth surrogate* for $\mathbf{Rank}\, X$ and instead solve the problem

$$\begin{array}{ll} \text{minimize} & \log\det(X+\delta I) \\ \text{subject to} & X \in \mathcal{C}, \end{array} \tag{5.17}$$

where $\delta > 0$ can be interpreted as a small regularization constant. The idea of using a log-det type function to obtain low-rank solutions to LMI problems is not entirely new—a similar idea also appears in the potential reduction method of [20] for positive semidefinite matrices (see Chapter 4). However, we take a different approach to finding a local minimum of this function over the constraint set $\mathcal{C}$.

Note that the surrogate function $\log\det(X+\delta I)$ is not convex (in fact, it is concave). However, since it is smooth on the positive definite cone, it can be minimized (locally) using any local minimization method. We use iterative linearization to find a local minimum. Let X_k denote the kth iterate of the optimization variable X. The first-order Taylor series expansion of $\log\det(X+\delta I)$ about X_k is given by

$$\log\det(X+\delta I) \approx \log\det(X_k+\delta I) \ + \ \mathbf{Tr}(X_k+\delta I)^{-1}(X-X_k). \tag{5.18}$$

Here we have used the fact that $\nabla \log\det X = X^{-1}$, when $X > 0$. Hence, one could attempt to minimize $\log\det(X+\delta I)$ over the constraint set $\mathcal{C}$ by iteratively minimizing the local linearization (5.18). This leads to

$$X_{k+1} = \operatorname*{argmin}_{X\in\mathcal{C}} \mathbf{Tr}(X_k+\delta I)^{-1}X. \tag{5.19}$$

The new optimal point is X_{k+1}, and we have ignored the constants in (5.18) because they do not affect the minimization.

Since the function $\log\det(X+\delta I)$ is concave in X, at each iteration its value decreases by an amount more than the decrease in the value of the linearized objective. Based on this observation, it can be shown (*e.g.*, using the global convergence theorem in [65, p.187]) that the sequence of the function values generated converges to a local minimum of $\log\det(X+\delta I)$.

Note that the trace heuristic can be viewed as the first iteration in (5.19), starting from the initial point $X_0 = I$. Therefore, we always pick $X_0 = I$, so that X_1 is the result of the trace heuristic, and the iterations that follow try to reduce the rank of X_1 further.

5.2.2 General case

In order to extend the log-det heuristic to the general case, we appeal to Lemma 1 again, and recall the equivalence between the RMP (2.1) and its PSD form (3.2). Since the matrix $\mathbf{diag}(Y, Z)$ is semidefinite, the log-det heuristic (5.17) can be applied. This yields

$$\begin{array}{ll} \text{minimize} & \log\det(\mathbf{diag}(Y,Z)+\delta I) \\ \text{subject to} & \begin{bmatrix} Y & X \\ X^T & Z \end{bmatrix} \geq 0 \\ & X \in \mathcal{C}. \end{array} \tag{5.20}$$

Linearizing as before, we obtain the following iterations for solving (5.20) locally:

$$\begin{array}{rll} \mathbf{diag}(Y_{k+1},Z_{k+1}) = & \text{argmin} & \mathbf{Tr}(\mathbf{diag}(Y_k,Z_k)+\delta I)^{-1}\,\mathbf{diag}(Y,Z) \\ & \text{subject to} & \begin{bmatrix} Y & X \\ X^T & Z \end{bmatrix} \geq 0 \\ & & X \in \mathcal{C}, \end{array} \tag{5.21}$$

where each iteration is an SDP in the variables X, Y and Z.

Figure 5.3 provides an intuitive interpretation for the heuristic. It shows the basic idea behind the $\mathbf{Tr}\,X$ and $\log\det(X+\delta I)$ approximations of $\mathbf{Rank}\,X$. The objective functions for the trace and log-det heuristics are shown for the scalar case, *i.e.*, when $x \in \mathbf{R}$ and $\sigma(x) = |x|$.

5.2.3 Vector case: iterative ℓ_1-norm minimization

We now study the log-det heuristic in the vector case, as we did with the trace heuristic in Section 5.1.7. Consider the special case of the diagonal rank minimization problem (3.4). Applying the log-det heuristic to this problem yields

$$\begin{array}{ll} \text{minimize} & \sum_i \log(y_i+\delta) \\ \text{subject to} & |x_i| \leq y_i \quad,\ i=1,\ldots,n \\ & x \in \bar{\mathcal{C}}, \end{array}$$

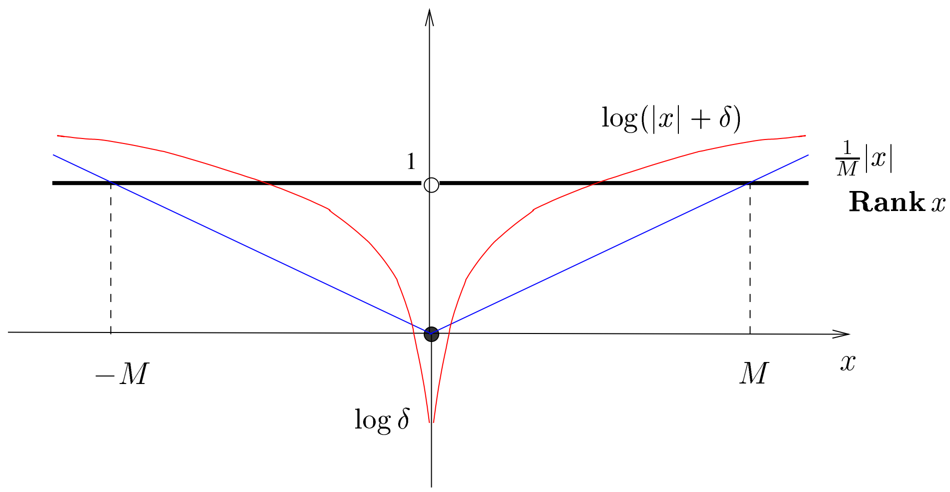

Figure 5.3: The rank, trace, and log-det objectives in the scalar case

or, equivalently,

$$\begin{array}{ll}\text{minimize} & \sum_i \log(|x_i|+\delta) \\ \text{subject to} & x \in \bar{\mathcal{C}},\end{array}$$

where $x \in \mathbf{R}^n$ is the optimization variable. Iterative linearization of the concave objective function gives the following heuristic for vector cardinality minimization:

$$x^{(k+1)} = \operatorname*{argmin}_{x\in\bar{\mathcal{C}}} \sum_i \frac{|x_i|}{|x_i^{(k)}|+\delta}. \tag{5.22}$$

Note that if the initial point is chosen as $x^{(0)} = [1,\ 1, \ldots,\ 1]$, the first iteration will minimize $\|x\|_1$. Thus the first iteration is the same as the ℓ_1 heuristic that we derived in Section 5.1.7 as the vector version of the trace heuristic.

A closer look at this iterative procedure shows that in each step, a weighted ℓ_1 norm of the vector x is minimized. This yields an intuitive interpretation of the method: if $x_i^{(k)}$ is small, its weighting factor in the next minimization step, $(x_i^{(k)}+\delta)^{-1}$, is large. So the small entries in x are generally pushed towards zero as far as the constraints on x allow, and thus yield a sparse solution.

See [63] for more about the iterative ℓ_1 minimization procedure, and its application to the problem of portfolio optimization with fixed transaction costs.

5.3 Illustrative examples and comparisons

In the previous chapter, we listed several groups of heuristics for the RMP: alternating projections, factorization, and analytic anti-centering/potential reduction. In this chapter we presented the trace and log-det heuristics. The goal of the following examples is to give an idea of how well various heuristics work, in terms of finding a low-rank solution. In these examples, we consider the RMP

$$\begin{array}{ll} \text{minimize} & \mathbf{Rank}\, X \\ \text{subject to} & X = A_0 + \sum_{i=1}^{n} x_i A_i \\ & X \geq 0, \end{array} \tag{5.23}$$

where $x \in \mathbf{R}^n$ is the optimization variable.

Example 1. We first present a simple example with only two variables so that a graphical representation is possible. The purpose is to illustrate how various heuristics work. Consider problem (5.23) with $x \in \mathbf{R}^2$, and with A_0, A_1 and A_2 given such that

$$A(x) = A_0 + x_1 A_1 + x_2 A_2 = \begin{bmatrix} 9 - x_1 & 3 & 0 & 0 & 0 \\ 3 & 9 - x_2 & 0 & 0 & 0 \\ 0 & 0 & x_1 - 1 & 0 & 0 \\ 0 & 0 & 0 & -x_1 + x_2 & 0 \\ 0 & 0 & 0 & 0 & x_1 + x_2 - 2 \end{bmatrix}.$$

The feasible region is shown in Figure 5.4, where the solid lines depict the boundary of the set. The curved boundary corresponds to the constraint that the top 2×2 block in A be positive semidefinite. Similarly, the linear boundaries correspond to

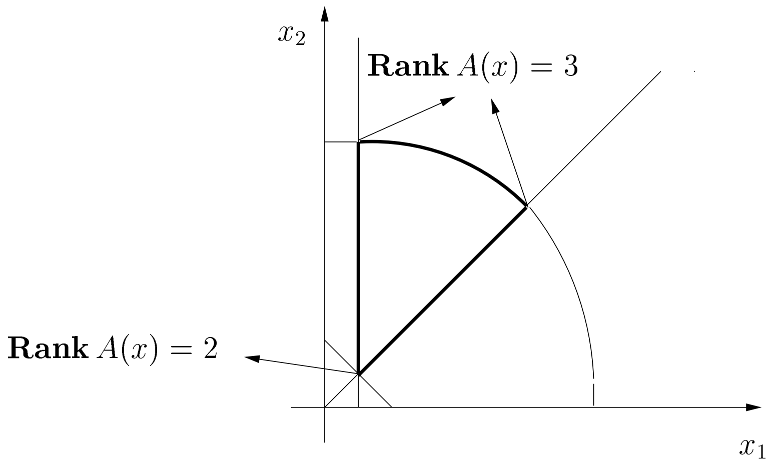

Figure 5.4: Feasible region for the problem in Example 1.

the other diagonal terms in A being non-negative. The goal is to pick an x in this set such that $A(x)$ has the lowest possible rank. Such a point has to lie on the intersection of the largest number of boundaries. Thus, it can also be solved by solving a set of linear and quadratic problems that check if the boundaries intersect. But since every combination of boundaries has to be taken into account, it is a combinatorial problem. We can see from the figure that the point $[1,\ 1]^T$, which is on the crossing of three boundary lines, yields the global minimum of 2 for **Rank** $A(x)$. The rank at $[6,\ 6]^T$ and $[1,\ 7.875]^T$ is 3, and at all other (non-corner) boundary points it is 4. In the interior of the set, $A(x)$ is clearly of full rank.

We apply the alternating projections method, the factorization method described in section 4.2.3 of Chapter 4, and the trace heuristic described in this chapter, to this RMP.

In the alternating projections method, in the first step ($r = 4$) the initial point x_0 is picked randomly, and in the next steps it is chosen to be the result of the previous step. Similarly, in the factorization method, the initial F_0 is random, and in the next steps it is chosen as a square-root of the $A(x)$ obtained in the previous step.

Method	r	x_{opt}	# of iterations
Factorization	4	[1.000, 5.688]	2 (4 SDPs)
	3	[1.000, 7.875]	4 (8 SDPs)
	2	feasible point not found	
Alternating projections	4	[2.765, 7.557]	2
	3	[1.000, 7.875]	37
	2	feasible point not found	
Trace	2	[1.000, 1.000]	1

Table 5.1: Results for Example 1.

The results are shown in Table 5.1. We see that the global minimum, with a rank of 2, is found by the trace heuristic (*i.e.*, one iteration of the log-det heuristic). This involves solving an SDP with 2 variables and 5 constraints. The two other methods manage to find only a rank 3 matrix, with higher computational effort. Note that same behavior is observed with various randomly chosen initial points, although for both methods, there exist initial points that do yield a rank 2 solution. This fact further emphasizes that these two methods are sensitive to the choice of the starting point.

Example 2. As a larger example, we consider problem (5.23) with $x \in \mathbf{R}^8$. The matrices $A_i \in \mathbf{R}^{15\times 15}$ are diagonal, and are chosen randomly. Thus the constraints are simply a set of 15 linear inequalities. As before, the minimum rank is achieved at a point where the largest number of the constraints are tight.

We again apply the three heuristics and compare the results. The initial points for the factorization and alternating projections methods are chosen as in the previous example. The results are given in Table 5.2.

Example 3. In the following examples, we compare the trace and log-det heuristics with the interior-point-based methods on two simple examples. These methods are introduced and studied in detail in the dissertation of Johan David [20], where many small numerical examples are given. We quote two examples of [20] here for

Method	r	# of iterations
Factorization	7	3
	6	feasible point not found
	5	feasible point not found
Alternating projections	7	9
	6	feasible point not found
	5	feasible point not found
Trace	5	1

Table 5.2: Results for Example 2.

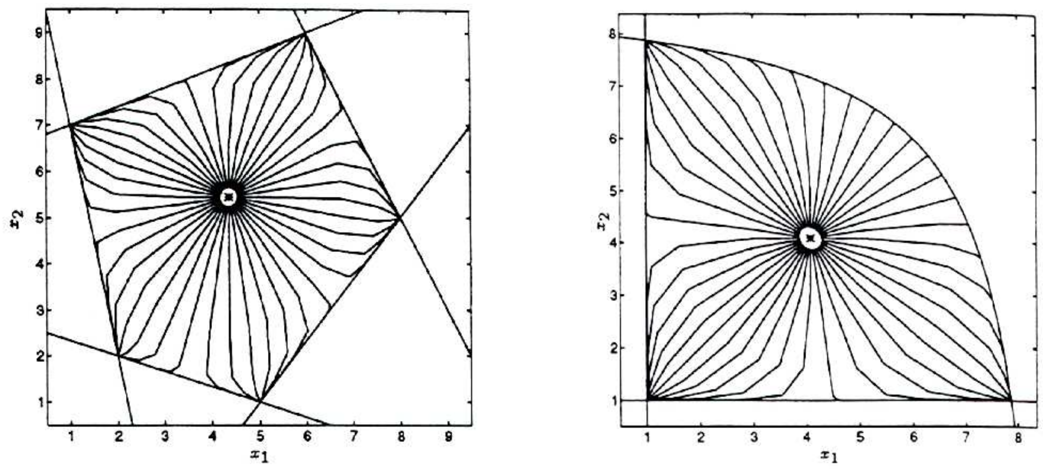

Figure 5.5: Figure shows feasible region, the analytic center, and the paths followed by the potential reduction iterations towards the boundary for two examples taken from [20].

comparison purposes. Both problems are RMPs in the form of (5.23) with two variables, with feasible regions shown in Figure 5.5. The initial point is chosen in the vicinity of the analytic center of the constraints. The lines show the paths followed by the iterations in the potential reduction method towards the boundary. We see that in both examples, the point that the paths converge to is highly sensitive to the initial point, *i.e.*, a small change in the initial point may significantly change the results.

For comparison, we apply the trace and log-det heuristics to these two examples. In example shown on the left, the trace heuristic (*i.e.*, one iteration of log-det) yields $x = [2,\ 2]$ where $\mathbf{Rank}\, A(x) = 3$, the global minimum in this case. In example shown on the right, the log-det heuristic converges in 3 iterations and yields $x = [1,\ 1]$ where $\mathbf{Rank}\, A(x) = 2$, again the global minimum.

5.3.1 Conclusions and remarks

We note that the methods presented in this and the previous chapter are all heuristics; we can always find special examples in which one method outperforms the others. Therefore, we do not claim that the trace/log-det heuristics yield lower rank solutions than other methods in all problem cases. However, the random examples given in this section suggest that the trace/log-det heuristics, besides having several important benefits over the other methods, often do perform better as well.

The main benefits of the new heuristics are as follows:

- They can be applied to any general RMP.
- There is no need for user-specified initial points, because the nuclear norm heuristic provides a low-rank solution that can be used as an initial point for further log-det iterations.

- For the nuclear norm heuristic, it is possible to show (through the convex envelope result) in what sense the heuristic is optimal, and what kind of behavior could be expected from it. It also provides global lower bound on the minimum rank, if the feasible set is bounded (which is often the case in practice). This is in contrast to other heuristics that do not provide such information. Such bounds can be used in branch-and-bound methods in global optimization.

- The log-det iterations require solving a convex problem at each step, which can be done very efficiently. Typically only a few steps are needed as the log-det iterations converge very fast in practice.

- Unlike the alternating projections and factorization methods, these heuristics do not require checking the feasibility for all values of rank, thus saving the extra computational effort.

Chapter 6

Applications

In this chapter, several rank minimization problems are considered that arise in a variety of areas, ranging from control and system identification, to statistics and psychometrics, and finance. The problems are solved (approximately) using the trace and log-det heuristics presented in the previous chapter. The goal is to demonstrate the effectiveness of these heuristics in producing low-rank solutions in practice.

Recall that in general, the trace heuristic, as well as each iteration of the log-det heuristic, require solution of an SDP. We use the software SDPSOL [101]. In some cases, *e.g.*, minimum-order system approximation in the single-input, single-output (SISO) case, each iteration can be written as a second-order cone program (SOCP), which can be solved more efficiently than an SDP. Some existing SOCP solvers include SOCP [64] and MOSEK [4].

6.1 System realization with time-domain constraints

In this section, we discuss the problem of designing a low-order, discrete-time, linear time-invariant (LTI) dynamical system, directly from convex specifications on the first n time samples of its impulse response. Some typical specifications are bounds on the rise-time, settling-time, slew-rate, overshoot, etc. This problem can be posed as one of minimizing the rank of a Hankel matrix over a convex set.

We begin with a fact about linear systems that can be derived from standard results in [15, 86]. We denote by H_n the Hankel matrix with parameters $h_1, h_2, \ldots, h_{2n-1} \in \mathbf{R}$,

$$H_n = \begin{bmatrix} h_1 & h_2 & h_3 & \ldots & h_n \\ h_2 & h_3 & h_4 & \ldots & h_{n+1} \\ h_3 & h_4 & h_5 & \ldots & h_{n+2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ h_n & h_{n+1} & h_{n+2} & \ldots & h_{2n-1} \end{bmatrix}. \tag{6.1}$$

Fact 1 *Let $h_1, h_2, \ldots, h_n$ be given real numbers. Then there exists a minimal LTI system of order r, with state space matrices $A \in \mathbf{R}^{r\times r}$, $B \in \mathbf{R}^{r\times 1}$ and $C \in \mathbf{R}^{1\times r}$, such that*

$$CA^{i-1}B = h_i \quad i = 1, \ldots, n,$$

if and only if

$$r = \min_{h_{n+1},\ldots,h_{2n-1}\in\mathbf{R}} \mathbf{Rank}\, H_n,$$

where H_n is a Hankel matrix whose first n parameters are the given $h_1, h_2, \ldots, h_n$, and whose last $n-1$ parameters, $h_{n+1}, \ldots, h_{2n-1} \in \mathbf{R}$, are free variables.

In other words, there exists a linear time invariant system of order r whose first n impulse response samples are $h_1, \ldots, h_n$, if and only if the minimal-rank Hankel matrix has rank r. Once $h_1, \ldots, h_{2n-1}$ are known, a state space description $\{A, B, C\}$ can be easily obtained [75].

Note that the constraints in the Fact 1 are only on the first n samples, even though $h_{n+1}, \ldots, h_{2n-1}$ also appear in the Hankel matrix. These extra variables are left free in the optimization. Thus, they are chosen in a way so as to minimize the overall

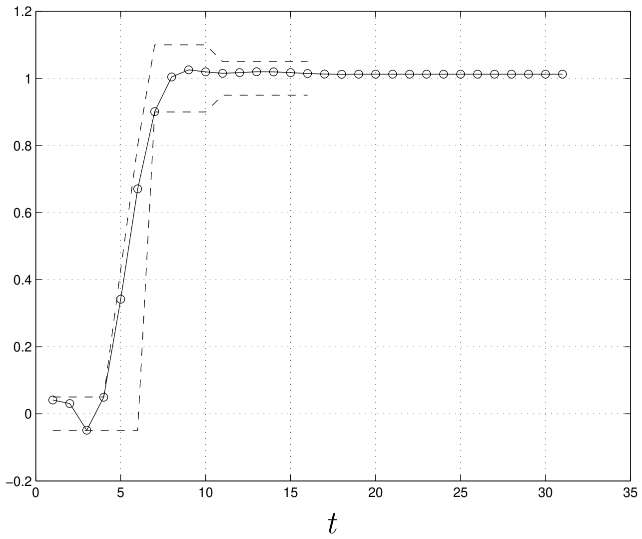

Figure 6.1: Step response specifications (dashed) and actual step response obtained after 5 iterations of the log-det heuristic

rank of the Hankel matrix.

To see how the facts above can be used to design low-order systems directly from specifications, consider the specifications on the step response shown in Figure 6.1. The goal is to find the minimum-order system whose step response fits in the region defined by the dashed lines, up to the 16th sample. The dashed lines are meant to capture a typical set of time-domain step response specifications: certain rise-time, slew-rate, overshoot, and settling characteristics and an approximate delay of four samples. The problem can be expressed as

$$\begin{array}{ll} \text{minimize} & \textbf{Rank}\, H_n \\ \text{subject to} & l_i \le s_i \le u_i, \quad k = 1, \ldots, n \\ & h_{n+1}, \ldots, h_{2n-1} \in \mathbf{R}, \end{array} \tag{6.2}$$

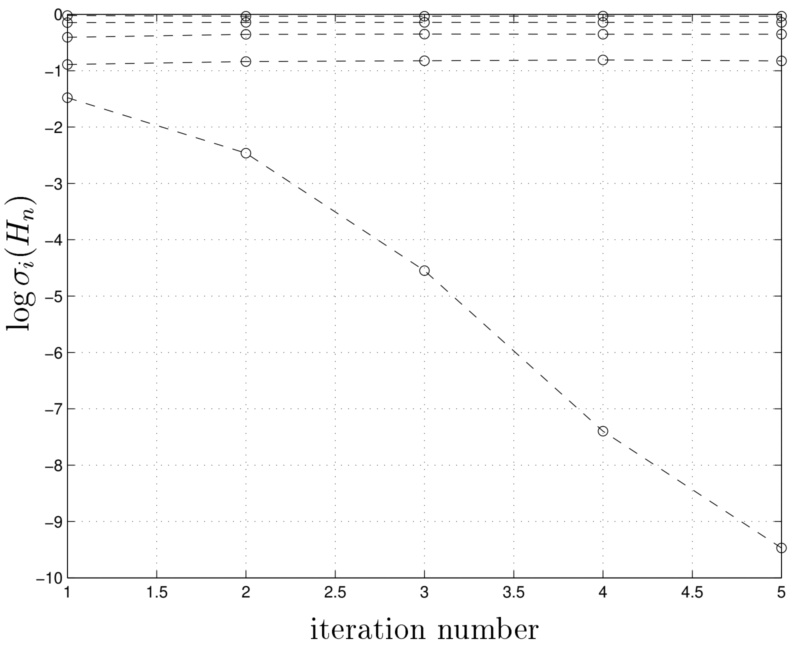

Figure 6.2: Log of the singular values $\sigma_1, \ldots, \sigma_5$ of H_n at each iteration

where $s_k = \sum_{i=1}^{k} h_i$ denote the terms in the step response, and l_i and u_i are, respectively, samples of the lower and upper time domain specifications (shown by the dashed lines).

This problem is an RMP with no analytical solution. Note also that the optimization variable H_n is not positive semidefinite. We apply the generalized trace and log-det heuristics described in Chapter 5 to this problem. Because of the approximate four-sample delay specification, we do not expect that the specifications can be met by a system of order less than four.

After five iterations of the log-det heuristic, a fourth-order system is obtained with the step response shown in Figure 6.1. Thus, all the specifications can be met by a linear time-invariant system of order exactly four. In this example, we set $\delta = 10^{-6}$. Figure 6.2 shows the logarithm of the nonzero Hankel singular values. We see that the rank of the 16×16 matrix H_n drops to 5 after the first iteration, and the next four iterations bring the rank to 4, which in this case happens to be the global minimum.

6.2 Minimum-order system approximation

In this section, we apply the nuclear norm heuristic to the problem of minimum-order system approximation. Such problems arise, for example, in model reduction problems that come from overparametrization in subspace system identification [56, 67, 75]. They also arise in $\mathcal{H}_\infty$ model reduction [47].

Let $p_1, \ldots, p_N \in \mathbf{C}$ be a set of complex numbers with conjugate symmetry, *i.e.*, whenever p_i is complex, there is some j such that $p_j = \bar{p_i}$. We consider the family of proper rational matrices given by

$$H(s) = R_0 + \sum_{i=1}^{N} \frac{R_i}{s - p_i}, \tag{6.3}$$

where $R_i \in \mathbf{C}^{m \times n}$ satisfy conjugate symmetry: whenever $p_i = \bar{p_j}$, we have $R_i = \bar{R_j}$. We consider p_i, the poles of the rational matrix H, as fixed; the residues R_i are the variables that we will use for approximation (subject to the conjugate symmetry constraint). The McMillan degree, *i.e.*, the order of a minimal state-space realization, of the rational matrix H is given by

$$\deg(H) = \sum_{i=1}^{N} \mathbf{Rank}\, R_i = \mathbf{Rank}\, \mathbf{diag}(R_1, \ldots, R_N).$$

Our goal is to determine values of the residue matrices R_i that minimize the McMillan degree over some set of acceptable approximations.

Let $\omega_1, \ldots, \omega_K \in \mathbf{R}$, and suppose $G_k \in \mathbf{C}^{m \times n}$ are given. We can interpret the ω_k

and G_k as sampled frequencies and the measured frequency response matrix, respectively. As a criterion for acceptable fit, we use the simple conditions

$$\|H(j\omega_k) - G_k\| \leq \epsilon, \qquad k = 1, \ldots, K,$$

i.e., that the matrix H, evaluated at the given frequencies, should approximate the given data (in spectral norm) within a tolerance ϵ.

The problem of finding the minimum-order approximation is then given by

$$\begin{array}{ll} \text{minimize} & \mathbf{Rank}\,\mathbf{diag}(R_1, \ldots, R_N) \\ \text{subject to} & \|H(j\omega_k) - G_k\| \leq \epsilon, \quad k = 1, \ldots, K \\ & R_j = \bar{R}_i \text{ for } p_j = \bar{p}_i, \end{array} \tag{6.4}$$

where the optimization variables are the $R_i \in \mathbf{C}^{m \times n}$. Note that $H(j\omega_k)$ is a linear function of the variables R_i. The objective can also be expressed as the rank of the block-diagonal matrix with blocks $R_1, \ldots, R_N$, so this problem has the form of the RMP (2.1) (with complex matrices instead of real matrices). For a discussion of optimization over an affine family of linear systems, see [12, §10.1].

Using Schur complements, we can replace the first constraint in (6.4) by its LMI equivalent. Then applying the generalized trace heuristic, we obtain the following SDP:

$$\begin{array}{ll} \text{minimize} & \sum_{i=1}^N \mathbf{Tr}\,Y_i + \mathbf{Tr}\,Z_i \\ \text{subject to} & \begin{bmatrix} Y_i & R_i \\ R_i^* & Z_i \end{bmatrix} \geq 0 \quad i = 1, \ldots, N \\ & \begin{bmatrix} \epsilon I & H(j\omega_k) - G_k \\ (H(j\omega_k) - G_k)^* & \epsilon I \end{bmatrix} \geq 0 \quad k = 1, \ldots, K \\ & R_j = \bar{R}_i \quad \text{for } p_j = \bar{p}_i, \end{array} \tag{6.5}$$

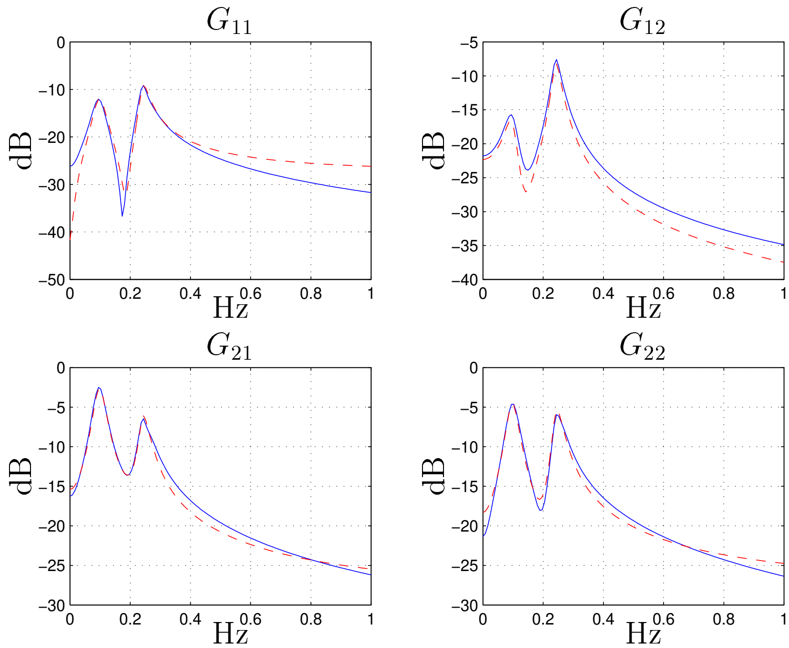

Figure 6.3: Original 8th order data (solid), and 6th order approximation (dashed).

where $R_i \in \mathbf{C}^{m\times n}$, $Y = Y^* \in \mathbf{C}^{m\times m}$, and $Z = Z^* \in \mathbf{C}^{n\times n}$ are the variables.

Note that (6.5) is a complex SDP. Since Y_i and Z_i are Hermitian, their traces are real, so the objective is real. The complex constraints in (6.5) can in turn be expressed as real LMIs, using the fact that for any Hermitian matrix $X \in \mathbf{C}^{n\times n}$, the matrix inequality $X \geq 0$ is equivalent to

$$\begin{bmatrix} \Re X & -\Im X \\ \Im X & \Re X \end{bmatrix} \geq 0,$$

which is an ordinary (real) LMI in the (real) matrix variables $\Re X$ and $\Im X$.

We demonstrate the techniques above on numerical data, generated from a generic system model. We use an 8th-order, 2-input 2-output transfer matrix F, which is normalized so that $\|F\|_\infty = \sup_\omega \|F(j\omega)\| = 1$. The frequencies ω_k, $k = 1, \ldots, K =$

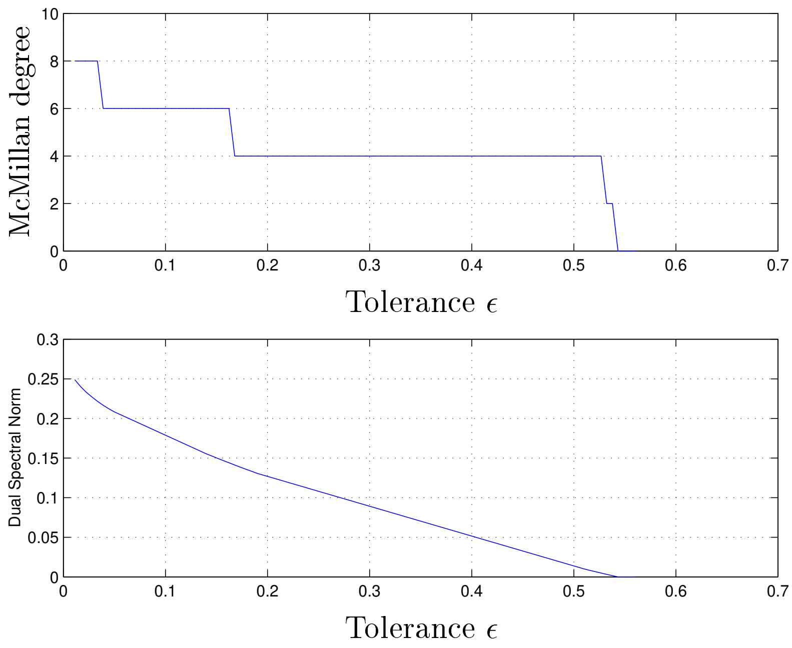

Figure 6.4: Tradeoff curves. The horizontal axis gives the approximation tolerance ϵ. The top plot shows the MacMillan degree obtained by the nuclear norm heuristic. The bottom plot shows the minimum nuclear norm.

128 are chosen as linearly spaced from 0 Hz to 1 Hz, and G_k was taken as the value of the 8th-order model at ω_k: $G_k = F(j\omega_k)$. For the poles $p_1, \ldots, p_8$, we take the poles of F, which appear in four complex conjugate pairs. Two pairs are clustered at ± 0.10 Hz, and the other two are around ± 0.24 Hz. The system approximation problem then becomes a model reduction problem: we keep the poles of the original system and modify the residue matrices. The goal is to reduce the order while respecting a given error level.

As an example, (6.5) is solved with $\epsilon = 0.05$ (-26dB). The result is a 6th-order approximation. Figure 6.3 shows the magnitude plot of the original system (F) and the approximation result (*i.e.*, H). Using the generalized trace (or nuclear norm) heuristic (6.5) for a range of values of the tolerance ϵ from very small to 0.55, we obtain the tradeoff curve in Figure 6.4. The staircase curve is the actual rank objective from (6.4), evaluated for the optimizer of (6.5). This provides an upper bound on the optimal rank objective in (6.4). The lower curve is the objective value of (6.5). For more details on this problem, see [26].

6.3 Reduced-order controller design

In this section we study the application of trace and log-det heuristics to the well-known problem of reduced-order controller design. The most important considerations in the practical implementation of controllers using fixed-point DSP processors today are computation time, memory usage, and the effects of finite precision arithmetic. All of these are directly related to the order of the controller being implemented. Thus, in practice, it is highly desirable to achieve the specified performance with the lowest possible controller order.

As a specific example, consider the discrete-time linear time-invariant system P,

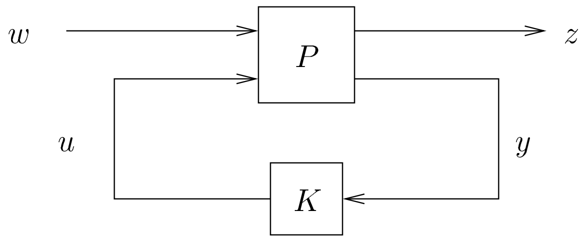

Figure 6.5: Closed loop feedback system with plant P and controller K.

described by the state-space equations

$$\begin{aligned} x(k+1) &= Ax(k) + B_1 w(k) + B_2 u(k) \\ z(k) &= C_1 x(k) + D_{11} w(k) + D_{12} u(k) \\ y(k) &= C_2 x(k) + D_{21} w(k). \end{aligned}$$

Here $x \in \mathbf{R}^n$ is the state; $w \in \mathbf{R}^{n_w}$ and $u \in \mathbf{R}^{n_u}$ are the disturbance and control inputs, respectively; $z \in \mathbf{R}^{n_z}$ and $y \in \mathbf{R}^{n_y}$ are the performance and measured outputs, respectively. Assume that the system is stabilizable and detectable. The goal is to find a stabilizing linear time-invariant controller K of minimum order, which when hooked up to the system as in Figure 6.5, makes the closed loop l_2-gain from w to z less than some prescribed level γ.

It is shown in [29, 55] that there exists a linear time-invariant stabilizing controller K of order $n_K \leq n$ that achieves a performance level γ, if and only if there exist symmetric positive definite matrices $R, S \in \mathbf{R}^{n \times n}$ such that

$$\mathbf{Rank} \begin{bmatrix} R & I \\ I & S \end{bmatrix} \leq n + n_K, \tag{6.6}$$

and

$$
\begin{bmatrix} R & I \\ I & S \end{bmatrix} \geq 0
$$

$$
\left[\begin{array}{c|c} \mathcal{N}_{12} & 0 \\ \hline 0 & I \end{array}\right]^T \left[\begin{array}{cc|c} ARA^T - R & ARC_1^T & B_1 \\ C_1RA^T & -\gamma I + C_1RC_1^T & D_{11} \\ \hline B_1^T & D_{11}^T & -\gamma I \end{array}\right] \left[\begin{array}{c|c} \mathcal{N}_{12} & 0 \\ \hline 0 & I \end{array}\right] < 0 \tag{6.7}
$$

$$
\left[\begin{array}{c|c} \mathcal{N}_{21} & 0 \\ \hline 0 & I \end{array}\right]^T \left[\begin{array}{cc|c} A^TSA - S & A^TSB_1 & C_1^T \\ B_1^TSA & -\gamma I + B_1^TSB_1 & D_{11}^T \\ \hline C_1 & D_{11} & -\gamma I \end{array}\right] \left[\begin{array}{c|c} \mathcal{N}_{21} & 0 \\ \hline 0 & I \end{array}\right] < 0,
$$

where $\mathcal{N}_{12}$ and $\mathcal{N}_{21}$ are orthonormal bases for the null spaces of $\left[B_2^T \; D_{12}^T\right]$ and $[C_2 \; D_{21}]$, respectively. Once appropriate R and S have been found, the state-space matrices of the controller K can be obtained directly from R and S and the state-space matrices of the plant P; see [29, 55].

The computation of the minimum-order controller K that achieves a performance level γ can thus be cast as the following RMP:

$$
\begin{array}{ll} \text{minimize} & \textbf{Rank} \begin{bmatrix} R & I \\ I & S \end{bmatrix} \\ \text{subject to} & (6.7). \end{array} \tag{6.8}
$$

This is a well-studied problem, and various heuristics have appeared in the literature, *e.g.*, the potential reduction method [20], the alternating projections method [41], the cone-complementarity approach [25], and the dual iteration method [54]. Still, the search for new and better heurisics for this important problem continues. In this section, we apply the trace and log-det heuristics to problem (6.8) and present some numerical results.

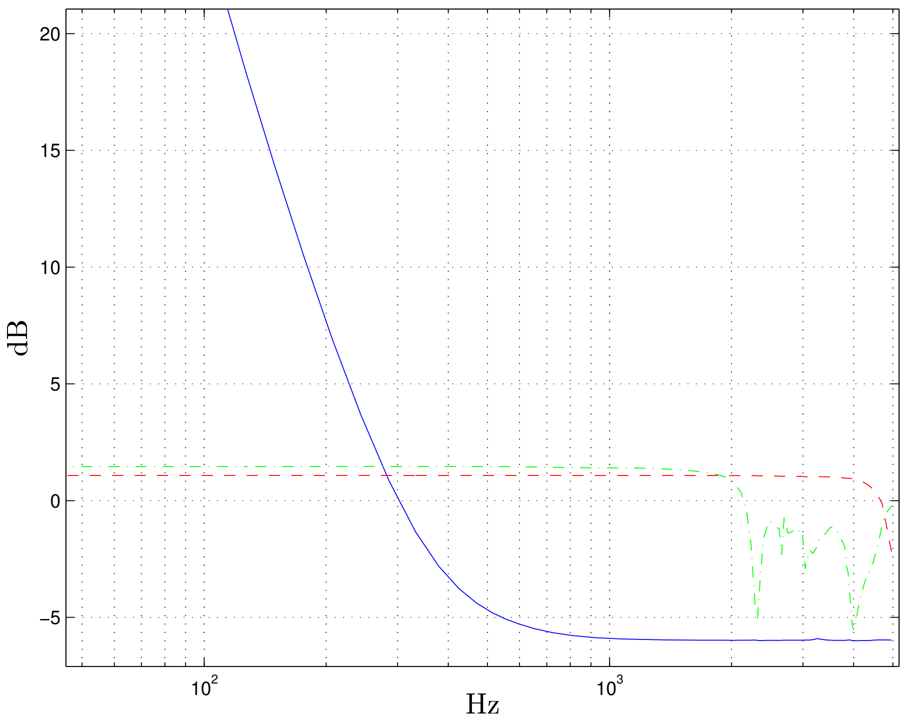

Figure 6.6: Maximum singular value plots: open loop system (solid), closed loop system with 29th-order controller corresponding to γ_{opt} (dashed), and closed loop system with 20th-order controller corresponding to a 5% relaxation of γ_{opt} (dash-dot). (Note that the open loop response does not appear to be very high order. This is because the higher order dynamics show up significantly in the smaller singular values, which are not plotted here.)

γ_{opt} Relaxation (%)	0	2	5	10
Controller Order	29	25	20	19

Table 6.1: Tradeoff between performance level γ and controller order.

As a specific example, we take P to be a model of a MIMO high-speed flexible positioning mechanism. The model P has five inputs ($n_w = 4, n_u = 1$), three outputs ($n_z = 2, n_y = 1$), and overall order $n = 29$.

The following procedure is used to design a low-order controller for P. First, the minimum achievable performance level γ_{opt} is computed by minimizing γ subject to (6.7)—this is a standard semidefinite program in γ, R and S. The controller computed from the R and S associated with γ_{opt} has order 29, and gives the closed loop response shown by the dashed line in Figure 6.6. We then relax the performance level γ_{opt} by 2%, 5%, and 10%, and solve (6.8) approximately using the trace and log-det heuristics for each of the relaxed γ values. This yields the approximate tradeoff between minimum controller order and achievable performance shown in Table 6.1.

Beyond a 5% relaxation of the performance level, it is not possible to get significant reductions in the controller order without large degradation in performance. Thus, a reasonable choice might be the controller associated with the 5% relaxation of γ_{opt}. This controller has order 20, and results in the closed loop response shown by the dash-dotted line in Figure 6.6. To within less than 0.5 dB, the controller achieves essentially the same performance, but with a 30% reduction in the number of states.

6.4 Euclidean distance matrix problems

Euclidean distance matrix (EDM) problems deal with constructing configurations of points from information about interpoint (Euclidean) distances. A simple example is reconstruction of the geographical map of a set of cities given pairwise inter-city distances [11, p. 16].

A matrix $D \in \mathbf{R}^{n\times n}$ is called a *Euclidean distance matrix* if there exist points $x_1, \ldots, x_n$ in $\mathbf{R}^r$ such that $D_{ij} = \|x_i - x_j\|^2$. The dimension of the space in which the points lie, r, is called the *embedding dimension*. Let $X \in \mathbf{R}^{r\times n}$ denote the matrix containing the x_i as columns, *i.e.*, $X = [x_1 \ \ldots \ x_n]$. The relation between the matrix of inner products $B = X^T X$ and the distance matrix D is then

$$D = \mathbf{diag}\, B\ \mathbf{1}^T + \mathbf{1}\ (\mathbf{diag}\, B)^T - 2B,$$

where

$$D_{ij} = \|x_i\|^2 + \|x_j\|^2 - 2x_i^T x_j = B_{ii} + B_{jj} - 2B_{ij}.$$

Let $V = I - \frac{1}{n}\mathbf{1}\mathbf{1}^T$ be the projection matrix onto the hyperplane $\mathbf{1}^T x = 0$. Multiplying a vector by V "centers" the vector by subtracting the mean of all coordinates from each coordinate, *i.e.*, by shifting the origin to the centroid of the points. Multiplying D by V on both sides yields

$$\begin{aligned} VDV &= V(\mathbf{diag}\, B\ \mathbf{1}^T + \mathbf{1}\ (\mathbf{diag}\, B)^T - 2B)V \\ &= -2VBV \\ &= -2\tilde{X}^T \tilde{X}, \end{aligned}$$

where $\tilde{X} = XV$, and columns of $\tilde{X}$ are the centered x_is. The matrix $-\frac{1}{2}VDV$ is sometimes called the double-centered distance matrix.

Schoenberg in 1935 [82] gave the necessary and sufficient conditions for a matrix to be an EDM with given embedding dimension. In our notation, this result shows that $D = D^T \in \mathbf{R}^{n\times n}$ is an EDM with embedding dimension r if and only if the following hold:

$$D_{ii} = 0, \tag{6.9}$$

$$VDV \preceq 0, \tag{6.10}$$

$$\mathbf{Rank}(VDV) \leq r. \tag{6.11}$$

The conditions above make intuitive sense; for example, we can show that the distance properties

$$D_{ii} = 0, \quad D_{ij} = D_{ji} \geq 0, \quad \text{and } D_{ij}^{1/2} \leq D_{ik}^{1/2} + D_{kj}^{1/2}, \quad \text{for any } i, j, k,$$

can be derived from (6.9) and (6.10). To show the positivity of D_{ij}, note that (6.10) is equivalent to $x^T Dx \leq 0$ for any x on the hyperplane $\mathbf{1}^T x = 0$. Let x be a vector with 1 in the ith position, -1 in the jth position, and zeros everywhere else. Then $x^T Dx \leq 0$ gives $D_{ij} \geq 0$.

To derive the triangle inequality, let x be a vector with entries x_i and x_j at the i and jth positions, and $-(x_i + x_j)$ at the kth position. From $x^T Dx \leq 0$ it follows

$$\begin{bmatrix} x_i & x_j & -(x_i + x_j) \end{bmatrix} \begin{bmatrix} 0 & D_{ij} & D_{ik} \\ D_{ij} & 0 & D_{jk} \\ D_{ik} & D_{jk} & 0 \end{bmatrix} \begin{bmatrix} x_i \\ x_j \\ -(x_i + x_j) \end{bmatrix} \leq 0,$$

which can be re-written as

$$\begin{bmatrix} x_i & x_j \end{bmatrix} \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} 0 & D_{ij} & D_{ik} \\ D_{ij} & 0 & D_{jk} \\ D_{ik} & D_{jk} & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ -1 & -1 \end{bmatrix} \begin{bmatrix} x_i \\ x_j \end{bmatrix} \leq 0,$$

for all $x_i, x_j \in \mathbf{R}$, which means the product of the three matrices above is negative semidefinite. Multiplying out the matrices, we have

$$\begin{bmatrix} -2D_{ik} & D_{ij} - D_{jk} - D_{ik} \\ D_{ij} - D_{jk} - D_{ik} & -2D_{jk} \end{bmatrix} \preceq 0.$$

To be negative semidefinite, the determinant of this 2×2 matrix should be positive, which yields $4D_{ik}D_{jk} \geq (D_{ij} - D_{jk} - D_{ik})^2$. Taking square-roots, we get

$$2D_{ik}^{1/2}D_{jk}^{1/2} \geq |D_{ij} - D_{jk} - D_{ik}| \geq D_{ij} - D_{jk} - D_{ik}.$$

Rearranging the terms as $(D_{ik}^{1/2} + D_{jk}^{1/2})^2 \geq D_{ij}$ and taking square-roots again yields the triangle inequality $D_{ik}^{1/2} + D_{kj}^{1/2} \geq D_{ij}^{1/2}$ for any i, j and k.

Condition (6.11) deals with the embedding dimension. It implies that VDV can be factorized as $VDV = -2\tilde{X}^T\tilde{X}$ with $\tilde{X} \in \mathbf{R}^{r\times n}$, where the columns of $\tilde{X}$ give an embedding in $\mathbf{R}^r$.

Problems involving EDMs arise in a variety of fields, such as Multi-Dimensional Scaling (MDS) in psychometrics and statistics, and in computational chemistry. In psychometrics, the information about interpoint distances is usually gathered through a set of experiments where subjects are asked to make quantitative (*e.g.*, in *metric* MDS) or qualitative (*e.g.*, in *non-metric* MDS) comparisons of objects. In statistics, such problems occur in extracting the underlying geometric structure of distance data. They have also been used in marketing research, in order to detect the number and nature of dimensions underlying the perceptions of different brands or products [40]. In chemistry, they come up in inferring the 3-dimensional structure of a molecule (molecular conformation) from information about its interatomic distances [71, 70, 90].

If the EDM D is known exactly, the corresponding configuration of points (up to a unitary transform) can be obtained by finding a square-root of $-\frac{1}{2}VDV$. However, in practice, typically only partial data, noisy measurements or incomplete information on D is available. It is often desired to find an EDM that not only is consistent with the measurements, but also requires the smallest number of coordinates to represent the data, *i.e.*, has the smallest embedding dimension. This problem can be expressed as the RMP

$$\begin{array}{ll} \text{minimize} & \mathbf{Rank}(-VDV) \\ \text{subject to} & D_{ii} = 0 \\ & -VDV \geq 0 \\ & D \in \mathcal{C}, \end{array} \tag{6.12}$$

where $\mathcal{C}$ is a convex set denoting the prior information on D. For example, we may have interval constraints on the distances, *i.e.*,

$$L_{ij} \leq D_{ij} \leq U_{ij},$$

where matrices L and U denote the lower and upper bounds. Another common constraint is for D to be close to the measured distance matrix $\hat{D}$ (*e.g.*, in matrix 2-norm or Frobenius norm),

$$\|D - \hat{D}\|_{2,F} \leq \epsilon,$$

where ϵ is a given tolerance. The measure of closeness can also be the Lipschitz distance, which is defined as

$$\delta(D, \hat{D}) = \log\left(\max_{i,j} \frac{D_{ij}}{\hat{D}_{ij}} \max_{i,j} \frac{\hat{D}_{ij}}{D_{ij}}\right).$$

Bounding the Lipschitz distance by some ϵ results in the following set of linear constraints on D:

$$\frac{D_{kl}}{\hat{D}_{kl}} \leq (\exp \epsilon) \frac{D_{ij}}{\hat{D}_{ij}}, \quad \text{for all } i, j, k, l.$$

Another type of constraint comes up in non-metric MDS, where only the "order" of the measured distances is considered, rather than the absolute distances themselves. This happens, for example, in non-metric MDS in psychometrics, where the data are human judgments on a pair of stimuli. The human mind may distort distances in a monotonic fashion; therefore only the information on the order of distances is reliable. The order information translates simply to linear inequality constraints on the entries of D, which is convex and can be easily handled.

The trace and log-det heuristics can be applied to this RMP. Our numerical experiments show that they work well, yielding EDMs with very low embedding dimensions.

As an example, consider 30 randomly generated points in $\mathbf{R}^5$, with all coordinates distributed uniformly over the interval $[0, 1]$. Let $\hat{D}$ be the matrix of squared distances corrupted by additive Gaussian noise, with zero mean and covariance 0.01. This matrix has full rank (with probability one) because of the noise, which obscures the underlying geometric structure. We would like to find the D close to $\hat{D}$ in Frobenuis norm, with the smallest embedding dimension. This can be expressed as the RMP

$$\begin{array}{ll}\text{minimize} & \mathbf{Rank}(-VDV) \\ \text{subject to} & D_{ii} = 0, \\ & -VDV \geq 0 \\ & \|D - \hat{D}\|_F \leq \epsilon,\end{array}$$

where we assume the tolerance ϵ to be $0.05\|\hat{D}\|_F$. Applying the log-det heuristic to this problem results in a D with a (correct) embedding dimension of 5 after 2 iterations (with $\delta = 10^{-6}$).

6.5 Portfolio optimization with fixed transaction costs

The Cardinality Minimization Problem (CMP) was discussed in Section 2.3 of Chapter 2 as a special case of the RMP. In this section we consider an application that arises in finance, the problem of portfolio optimization with fixed transaction costs. We show that the fixed-costs term can be expressed as a cardinality constraint, then we apply a variation of the ℓ_1-norm and iterative ℓ_1-norm heuristics (*i.e.*, the trace and log-det heuristics for the vector case) to obtain a sparse solution and global bounds on the optimal value. For more details on this problem and other portfolio selection problems, as well as the iterative ℓ_1-norm heuristic, see [63].

Consider an investment portfolio that consists of holdings in some or all of n assets. This portfolio is to be adjusted by performing a number of transactions, after which the portfolio will be held over a fixed time period. The investor's goal is to maximize the expected wealth at the end of the period, while satisfying a set of constraints on the portfolio. These constraints typically include limits on exposure to risk, and bounds on the amount held in each asset. Let $w \in \mathbf{R}^n$ denote the vector of current holdings in the assets. The total current wealth is then $\mathbf{1}^T w$. The dollar amount transacted in the ith asset is specified by x_i, with $x_i > 0$ for buying, and $x_i < 0$ for selling. Then $x \in \mathbf{R}^n$ is the vector of transactions. After transactions, the adjusted portfolio is $w + x$. Representing the sum of all transaction costs associated with x by $\phi(x)$, the budget constraint is $\mathbf{1}^T x + \phi(x) = 0$.

The adjusted portfolio $w + x$ is then held for a fixed period of time. At the end of that period, the return on asset i is the random variable a_i. All random variables are on a given probability space, for which $\mathbf{E}$ denotes expectation. We assume that we know the first and second moments of the joint distribution of $a = [a_1, \ldots, a_n]^T$, *i.e.*, $\mathbf{E}\, a = \bar{a}$ and $\mathbf{E}(a - \bar{a})(a - \bar{a})^T = \Sigma$. A riskless asset can be included, in which case the corresponding $\bar{a}_i$ is equal to its (certain) return, and the ith row and column of Σ are zero. The end of period wealth is a random variable, $W = a^T(w+x)$, with expected value and variance given by

$$\mathbf{E}\,W = \bar{a}^T(w+x), \qquad \mathbf{E}(W - \mathbf{E}\,W)^2 \;=\; (w+x)^T\Sigma\,(w+x).$$

The budget constraint can also be written as the inequality $\mathbf{1}^T x + \phi(x) \le 0$. With some obvious assumptions ($\bar{a}_i > 0$, $\phi \ge 0$), solving an expected wealth maximization problem with either form of the budget constraint yields the same result.

We summarize the portfolio selection problem as

$$\begin{array}{ll} \text{maximize} & \bar{a}^T(w+x) \\ \text{subject to} & \mathbf{1}^T x + \phi(x) \le 0 \\ & w + x \in \mathcal{S}, \end{array} \tag{6.13}$$

where $\mathcal{S} \subseteq \mathbf{R}^n$ is the set of feasible portfolios. Typical constraints on the portfolio include upper bounds on the variance, bounds on the amount of shorting allowed on each asset (*i.e.*, $w_i + x_i \ge -s_i$), bounds on total shorting, etc.

We assume the transaction costs to be separable, *i.e.*, the sum of the transaction costs associated with each trade is

$$\phi(x) = \sum_{i=1}^{n} \phi_i(x_i),$$

where ϕ_i is the transaction cost function for asset i. The simplest model for transaction costs is that there are none, *i.e.*, $\phi(x) = 0$. A better model of real transactions costs is a linear one, with the costs for each transaction proportional to the amount traded. However, in practice, transaction costs are not linear, and a fixed charge for any nonzero trade is common.

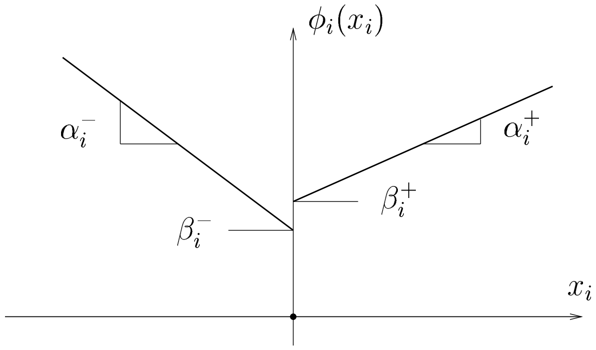

Figure 6.7: Fixed plus linear transaction costs $\phi_i(x_i)$ as a function of transaction amount x_i. There is no cost for no transaction, *i.e.*, $\phi_i(0) = 0$.

We consider a simple model that includes fixed plus linear costs. Let β_i^+ and β_i^- be the fixed costs, and α_i^+ and α_i^- be the cost rates associated with buying and selling asset i. The fixed-plus-linear transaction costs function is illustrated in Figure 6.7.

To simplify notation, we assume equal costs for buying and selling. The transaction costs function is then $\phi(x) = \sum_{i=1}^n \phi_i(x_i)$, with

$$\phi_i(x_i) = \begin{cases} 0, & x_i = 0 \\ \beta_i + \alpha_i|x_i|, & x_i \neq 0. \end{cases} \tag{6.14}$$

In general, costs of this form lead to a hard combinatorial problem. The simplest way to obtain an approximate solution is to ignore the fixed costs, and solve for $\phi_i(x_i) = \alpha_i|x_i|$. If the β_i are very small, this may lead to an acceptable approximation. In general, however, it will generate inefficient solutions with too many transactions. On the other hand, by considering the fixed costs, we discourage trading small amounts of a large number of assets. Thus, we obtain a *sparse* vector of trades; *i.e.*, one with many zero entries, or a small cardinality. This means most of the trading will be concentrated in a few assets, which is a desirable property. Thus, problem (6.13) can be considered as a problem with a constraint on the cardinality of the vector of trades. To see this directly, we can rewrite the budget constraint as

$$\mathbf{1}^T x + \sum_{i=1}^{n} \alpha_i |x_i| + \sum_{i=1}^{n} \beta_i I_+(x_i) \leq 0,$$

where $I_+(.)$ in the last term is the indicator function defined in Section 5.1.6. The term $\sum_i \beta_i I_+(x_i)$ is the same as the cardinality function $\mathbf{card}\, x = \sum_i I_+(x_i)$ except that each term has a weight β_i. We show that the ℓ_1-norm and iterative ℓ_1-norm heuristics can readily be extended to this problem.

We use the basic idea of replacing ϕ_i with its convex envelope. The convex envelope of ϕ_i in the interval $[-l_i, u_i]$ is

$$\phi_i^{\text{c.e.}}(x_i) = \begin{cases} \left(\dfrac{\beta_i}{u_i} + \alpha_i\right) x_i\,, & x_i \geq 0 \\ -\left(\dfrac{\beta_i}{l_i} + \alpha_i\right) x_i\,, & x_i \leq 0, \end{cases} \tag{6.15}$$

as shown in Figure 6.8. Using $\phi_i^{\text{c.e.}}$ for ϕ_i relaxes the budget constraint, in the sense that it enlarges the search set. Consider the portfolio selection problem (6.13), with $\phi_i^{\text{c.e.}}$ in place of ϕ_i,

$$\begin{array}{ll} \text{maximize} & \bar{a}^T(w + x) \\ \text{subject to} & \mathbf{1}^T x + \sum_{i=1}^{n} \phi_i^{\text{c.e.}}(x_i) \leq 0 \\ & w + x \in \mathcal{S}. \end{array} \tag{6.16}$$

This corresponds to optimizing the same objective subject to the same portfolio constraints, but with a looser budget constraint. Therefore the optimal value of (6.16) is an upper bound on the optimal value of the unmodified problem (6.13).

Note that in most cases the optimal transactions vector for the relaxed problem (6.16) will not be feasible for the original problem (6.13). The unmodified budget

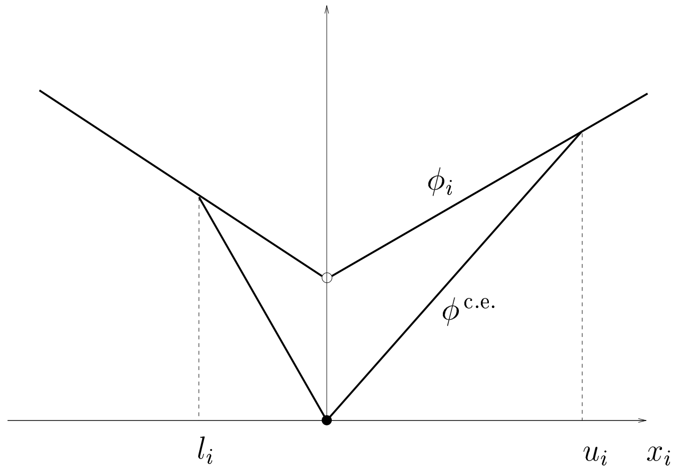

Figure 6.8: The convex envelope of ϕ_i over the interval $[l_i, u_i]$, is the largest convex function smaller than ϕ_i over the interval. For fixed plus linear costs, as shown here, the convex envelope is a linear transaction costs function.

constraint will not be satisfied by the solution of the modified program, except in the very special case when each transaction amount x_i is either l_i, u_i, or 0. These are the three values for which the convex envelope and the true transaction costs function agree.

We now show how to apply the iterative ℓ_1-norm heuristic for finding a feasible, suboptimal portfolio. The feasibility of the portfolio is obtained by ensuring that the modified transaction costs function $\widehat{\phi}_i^k$ agrees with the true ϕ_i at the solution transactions x_i^*.

Consider the following procedure: Let x^0 be the solution to problem (6.16), and apply the iterations

$$\begin{aligned} \widehat{\phi}_i^k(x_i) &= \left(\frac{\beta_i}{|x_i^{k-1}| + \delta} + \alpha_i \right) |x_i|, \\ x^k &= \underset{\substack{\mathbf{1}^T x + \sum_{i=1}^n \widehat{\phi}_i^k(x_i) \leq 0 \\ w + x \in \mathcal{S}}}{\operatorname{argmax}} \bar{a}^T(w + x) \end{aligned} \tag{6.17}$$

If the iterations converge, *i.e.*, if two successive iterates are close to each other, the solution x^* will be nearly feasible for the original problem (see Figure 6.9). This is seen by noting that for $x_i^* \gg \delta$,

$$\widehat{\phi}_i(x_i^*) \;=\; \left(\frac{\beta_i}{|x_i^*| + \delta} + \alpha_i\right)|x_i^*| \;\approx\; \beta_i + \alpha_i |x_i^*| \;=\; \phi_i(x_i^*),$$

and for $x_i^* = 0$,

$$\widehat{\phi}_i(x_i^*) \;=\; 0 \;=\; \phi_i(x_i^*).$$

Note that while $\widehat{\phi}_i(x_i^*) \leq \phi_i(x_i^*)$ for all x_i^*, this inequality is tight except when x_i^* is on the order of δ. In a sense, δ in this problem defines a *soft threshold* for deciding whether a given x_i^* is considered zero, *i.e.*, whether the corresponding transaction should be performed or not. In a practical implementation of the portfolio trades, a hard threshold is needed, and the x_i^* of order δ or smaller should be taken as zero.

Note that upper and lower bounds on the global optimum for the expected end of period wealth are given by $\bar{a}^T x^0$ and $\bar{a}^T x^*$.

For numerical examples, we consider problem (6.13) with fixed plus linear transaction costs, a limit on shorting of s_i per asset, and a bound on the variance of σ. We first describe an example with 10 stocks, plus a riskless asset. The mean and covariance of the risky assets was estimated from one year of daily closing prices of 10 stocks from the S&P 500. The parameters used are

$$\begin{array}{ll}
w_1, \ldots, w_{11} = 1/11 & \\
\alpha_1, \ldots, \alpha_{10} = 1\%, & \alpha_{11} = 0 \\
\beta_1, \ldots, \beta_{10} = 0.01, & \beta_{11} = 0 \\
s_1, \ldots, s_{10} = 0.05, & s_{11} = 0.5.
\end{array}$$

The small size of this problem allows us to compute the exact solution, that is the global optimum, by combinatorial search. Figure 6.10 displays the resulting tradeoff

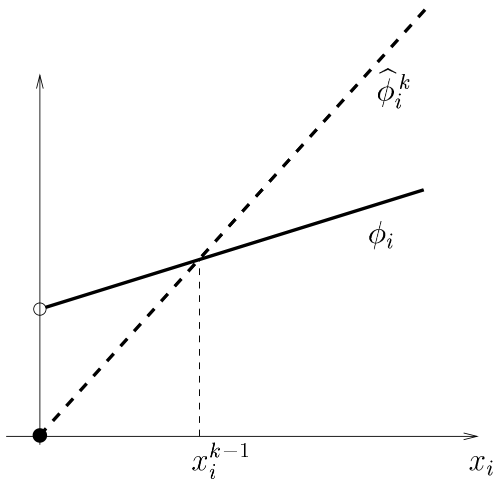

Figure 6.9: One iteration of the iterative ℓ_1 algorithm. Each of the nonconvex transaction costs (plotted as a solid line) is replaced by a convex one (plotted as a dashed line) that agrees with the nonconvex one at the current iterate. If two successive iterates are the same, then the iterates are feasible for the original nonconvex problem.

curve, with expected return plotted against standard deviation of return. Four curves are shown: the upper bound; the exact solution; the heuristic solution; and the solution computed without regard for fixed cost. Note that the upper bound is close to the heuristic solution. Note also that the heuristic nearly coincides with the exact solution. For the heuristic, we used $\delta = 10^{-3}$.

In Figure 6.11, still for the same 11 assets example, $\sigma_{\max}$ was kept constant at 0.15, and the problem was solved for a range of fixed costs β. The optimal expected return is plotted as a function of fixed costs, with the four curves obtained by the same procedure as in the previous figure. Again we can see that the difference between our heuristic and the optimal is very small. In this figure we can also see the cost of ignoring the transaction costs, which, naturally, increases with increasing fixed transaction costs.

As a second and larger example, we considered 100 stocks, plus a riskless asset,

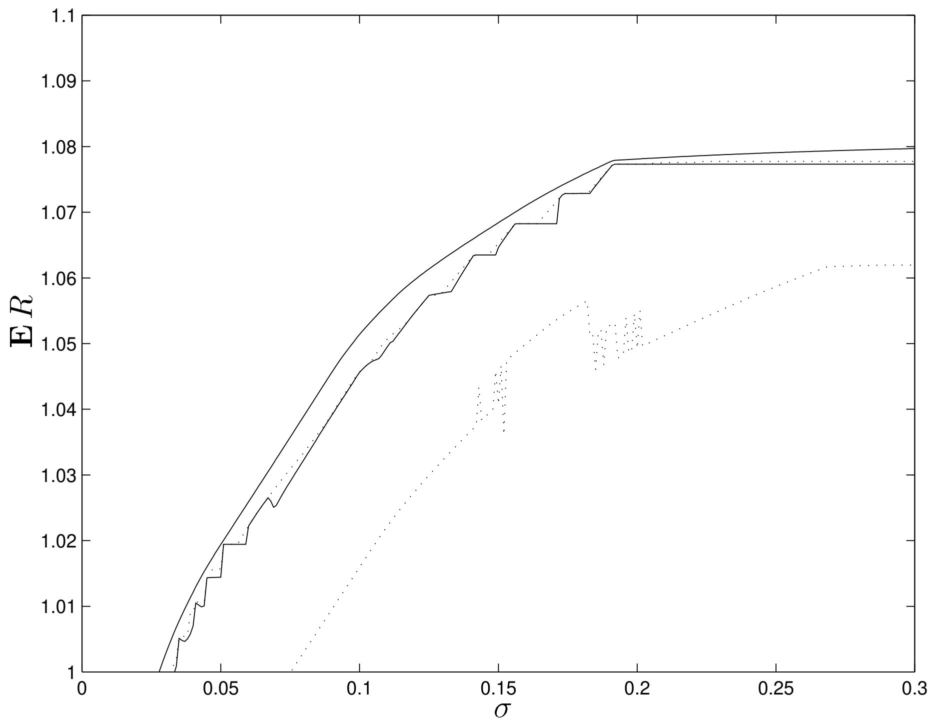

Figure 6.10: Example with 10 stocks plus riskless asset, plot of expected return as a function of standard deviation. Curves from top to bottom are: 1. global upper bound (*solid*), 2. true optimum by exhaustive search (*dotted*), 3. heuristic solution (*solid*), and 4. solution computed without regard for fixed costs (*dotted*). Note that curves 2 and 3 nearly coincide.

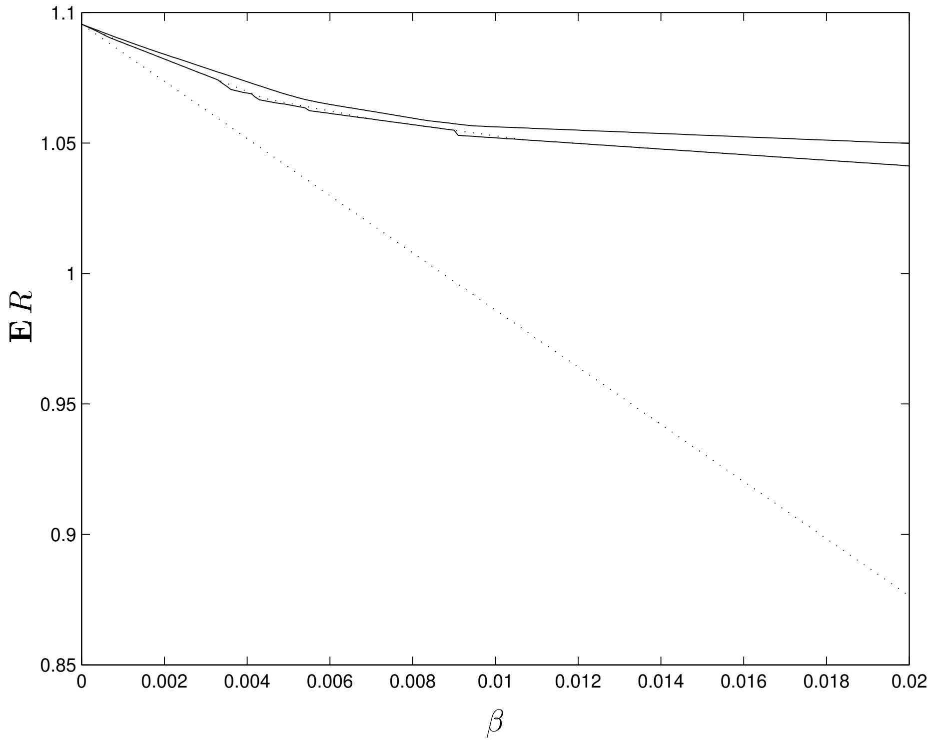

Figure 6.11: Example with 10 stocks plus riskless asset, plot of expected return as a function of fixed transaction costs. Curves from top to bottom are: 1. global upper bound (*solid*), 2. true optimum by exhaustive search (*dotted*), 3. heuristic solution (*solid*), and 4. solution computed without regard for fixed costs (*dotted*). Note that curves 2 and 3 nearly coincide.

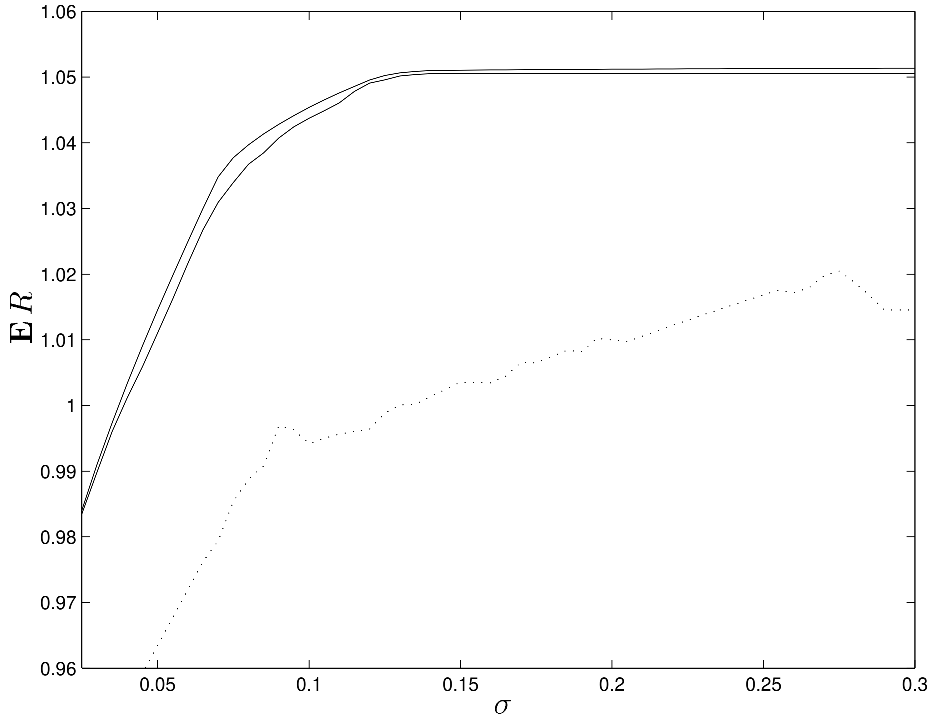

Figure 6.12: Example with 100 stocks plus riskless asset, plot of expected return as a function of standard deviation. Curves from top to bottom are: 1. global upper bound (*solid*), 2. heuristic solution (*solid*), and 3. solution computed without regard for fixed costs (*dotted*).

using the parameters

$$\begin{aligned} &w_1, \ldots, w_{101} = 1/101 \\ &\alpha_1, \ldots, \alpha_{100} = 1\%, &&\alpha_{101} = 0 \\ &\beta_1, \ldots, \beta_{100} = 0.001, &&\beta_{101} = 0 \\ &s_1, \ldots, s_{100} = 0.005, &&s_{101} = 0.5. \end{aligned}$$

Figure 6.12 displays the resulting tradeoff curve. The curves shown are the upper bound, the heuristic solution, and the solution computed without regard for fixed costs. Our numerical experiments indicate that convergence occurs in about 4 iterations or less for a wide range of problems.

Chapter 7

Conclusions

We studied the Rank Minimization Problem and listed many applications from a wide range of fields. Since the RMP is known to be NP-hard in general, we focused on efficient, but approximate, heuristic methods. We summarized the existing heuristics in three groups: alternating projections, interior-point-based methods, and factorization and linearization methods. Briefly, the main shortcomings of these methods are as follows: they are highly sensitive to the choice of initial point; are slow in general; and do not provide any information on the global minimum.

We presented new heuristics based on convex optimization. We started from the well-known fact that minimizing the trace of a positive semidefinite matrix tends to yield a low-rank matrix. We extended this heuristic to the case of general matrices using our semidefinite embedding lemma, which enabled us to embed a general RMP in a (larger) positive semidefinite one. The generalized heuristic is equivalent to minimizing the nuclear norm of the matrix. We showed that the nuclear norm is the convex envelope of the rank function over the set of matrices with spectral norm less than one, a result that provides theoretical backing for the heuristic. In the vector case (*i.e.*, the CMP), this heuristic reduces to the known ℓ_1-norm heuristic for obtaining sparse vectors.

As another heuristic, we used the (non-convex) log-det function as an approximation to rank. The trace heuristic provides a low-rank solution that can serve as a suitable starting point for a local optimization method to minimize the log-det function. Any local method can be used here; we used an iterative linearization and minimization method. This results in iterative weighted-trace minimization, which tends to reduce the rank further and refine the result of the trace heuristic. The log-det heuristic readily extends to general matrices using the semidefinite embedding lemma. In the vector case, this heuristic reduces to a new iterative ℓ_1-norm heuristic for the CMP.

The main benefits of the new heuristics are as follows:

- They can be applied to any general RMP.
- There is no need for user-specified initial points, because the nuclear norm heuristic provides a low-rank solution that can be used as an initial point for further log-det iterations.
- The nuclear norm heuristic is optimal in the sense that it minimizes the convex envelope of the rank function over the set of matrices with bounded spectral norm. It also provides a global lower bound on the minimum rank, if the feasible set is bounded (which is often the case in practice).
- The log-det iterations require solving a convex problem at each step, which can be done very efficiently. Typically only a few steps are needed as the log-det iterations converge very rapidly in practice.
- Unlike the alternating projections and factorization methods, these heuristics do not require checking feasibility for all values of the rank, thus saving the extra computational effort.

Finally, we demonstrated the effectiveness of the proposed heuristics on applied problems from different fields, *e.g.*, system identification, control, statistics and finance.

7.1 Future Research

- *Large-scale problems.* Many rank minimization problems that arise in practice are large-scale. Typically, they are also very sparse and highly structured. Examples include problems in combinatorial optimization, network flow optimization, image processing, and system approximation. If solvers for large-scale semidefinite programming become available, the heuristics we presented can be applied to large-scale RMPs that are important in practice. Large-scale semidefinite programming is currently an active area of research; see list of references given in Chapter 1.

- *Extensive numerical experiments.* Since the methods we discussed are all heuristics, none can be claimed to yield a better solution in all problem cases. A careful study of the computational behavior of the methods requires benchmarking problems of various types and sizes, and a large number of simulations for each. This is a topic to be addressed.

- *New applications.* We showed that the RMP arises in many different fields. Searching for new applications in these fields, and also fields that we did not cover, *e.g.*, biological systems, can be a direction for further work.

Appendix A

Notation and Glossary

$\mathbf{R}$	The set of real numbers.
$\mathbf{R}^m$	The set of real m-vectors.
$\mathbf{R}^{m \times n}$	The set of real $m \times n$ matrices.
$\mathbf{Rank}\, X$	The rank of X.
$\mathbf{Tr}\, X$	The trace of X.
$\det X$	The determinant of X.
$\|X\|$	The spectral (maximum singular value) norm of X.
$\|X\|_*$	The nuclear norm of X.
$\|X\|_F$	The Frobenius norm of X.
$\lambda_i(X)$	The ith largest eigenvalue of X.
$\sigma_i(X)$	The ith largest singular value of X.
I	The identity matrix (of appropriate dimensions).
$\mathbf{diag}(X_1, \ldots, X_n)$	The block diagonal matrix with diagonal blocks $X_1, \ldots, X_n$.
$X > 0$ ($X \geq 0$)	X is positive (semi-)definite, *i.e.*, $X = X^T$ and $z^T X z > 0$ ($z^T X z \geq 0$) for all nonzero z.
$X > Y$ ($X \geq Y$)	$X - Y$ is positive (semi-)definite.
$\mathbf{card}\, x$	The cardinality of the vector x.

PSD	Positive Semidefinite
SDP	Semidefinite Programming
LMI	Linear Matrix Inequality
RMP	Rank Minimization Problem
CMP	Cardinality Minimization Problem

Bibliography

[1] H. Akaike. A new look at the statistical model identification. *IEEE Trans. Aut. Control*, AC-19:716–723, 1974.

[2] F. Alizadeh. Interior point methods in semidefinite programming with applications to combinatorial optimization. *SIAM Journal on Optimization*, 5(1):13–51, February 1995.

[3] F. Alizadeh, J. P. Haeberly, M. V. Nayakkankuppam, and M. L. Overton. SDP-PACK *User's Guide, Version 0.8 Beta.* NYU, June 1997.

[4] E. Andersen. *MOSEK v1.0b User's manual*, 1999. Available from the URL `http://www.mosek.com`.

[5] T. W. Anderson. Asymptotic theory for principal component analysis. *Annals of Mathematical Statistics*, 34:122–148, 1963.

[6] V. Balakrishnan and S. Boyd. Global optimization in control system analysis and design. In C. T. Leondes, editor, *Control and Dynamic Systems: Advances in Theory and Applications*, volume 53, pages 1–56. Academic Press, New York, New York, 1992.

[7] C. Beck and J. Doyle. A necessary and sufficient minimality condition for uncertain systems. *IEEE Trans. Aut. Control*, 44(10):1802–1813, October 1999.

[8] S. J. Benson, Y. Ye, and X. Zhang. Solving large-scale sparse semidefinite programs for combinatorial optimization. *SIAM Journal on Optimization*, 10(2):443–461, 2000.

[9] E. Beran and K. Grigoriadis. A combined alternating projection and semidefinite programming algorithm for low-order control design. In *Proceedings of IFAC 96*, volume C, pages 85–90, July 1996.

[10] B. Borchers. CSDP*, a C library for semidefinite programming.* New Mexico Tech, March 1997.

[11] I. Borg and P. Groenen. *Modern Multidimensional Scaling: Theory and Applications.* Springer, 1997.

[12] S. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan. *Linear Matrix Inequalities in System and Control Theory*, volume 15 of *Studies in Applied Mathematics.* SIAM, Philadelphia, PA, June 1994.

[13] S. Boyd, L. Vandenberghe, A. El Gamal, and S. Yun. Design of robust global power and ground networks. *Proc. ACM/SIGDA Int. Symp. on Physical Design*, pages 60–65, April 2001.

[14] S. Burer, R. D. C. Monteiro, and Y. Zhang. Solving semidefinite programs via nonlinear programming. Part II: Interior-point methods for a subclass of SDPs. Technical report, Department of Computational and Applied Mathematics, Rice University, December 1999.

[15] C. T. Chen. *Linear System Theory and Design.* Holt, Rinehart and Winston, 1984.

[16] S. S. Chen and D. Donoho. Basis pursuit. In *28th Asilomar Conference Proceedings*, volume 1, November 1994.

[17] S. S. Chen, D. L. Donoho, and M. A. Saunders. Atomic decomposition by basis pursuit. *SIAM Review*, 43, 2001.

[18] C. Choi and Y. Ye. Solving sparse semidefinite programs using the dual scaling algorithm with an iterative solver. Technical report, The University of Iowa, 2000.

[19] R. R. Coifman and M. V. Wickerhauser. Entropy-based algorithms for best basis selection. *IEEE Transactions on information theory*, 38(2):713–718, March 1992.

[20] J. David. *Algorithms for analysis and design of robust controllers*. PhD thesis, Kat. Univ. Leuven, ESAT, 3001 Leuven, Belgium, 1994.

[21] P. Davis. Plane regions determined by complex moments. *Journal of Approximation Theory*, 19:148–153, 1977.

[22] D. L. Donoho. Sparse components of images and optimal atomic decompositions. Technical report, Statistics Department, Stanford University, December 1998.

[23] L. El Ghaoui and V. Balakrishnan. Synthesis of fixed-structure controllers via numerical optimization. In *Proc. IEEE Conf. on Decision and Control*, pages 2678–2683, December 1994.

[24] L. El Ghaoui and P. Gahinet. Rank minimization under LMI constraints: A framework for output feedback problems. In *Proc. European Control Conf.*, The Netherlands, 1993.

[25] L. El Ghaoui, F. Oustry, and M. AitRami. A cone complementarity linearization algorithm for static output-feedback and related problems. In *Proceedings of*

the 1996 IEEE International Symposium on Computer-Aided Control System Design, pages 246–251, Dearborn, MI, 1996.

[26] M. Fazel, H. Hindi, and S. P. Boyd. A rank minimization heuristic with application to minimum order system approximation. In *Proc. American Control Conf.*, Arlington, Virginia, 2001.

[27] K. Fujisawa and M. Kojima. SDPA (semidefinite programming algorithm) user's manual. Technical Report B-308, Department of Mathematical and Computing Sciences. Tokyo Institute of Technology, 1995.

[28] M. Fukuda, M. Kojima, K. Murota, and K. Nakata. Exploiting sparsity in semidefinite programming via matrix completion I: General framework. Technical Report B-358, Tokyo Institute of Technology, 2000.

[29] P. Gahinet and P. Apkarian. A linear matrix inequality approach to $\mathcal{H}_\infty$ control. *Int. J. Robust and Nonlinear Control*, 4:421–488, 1994.

[30] P. Gahinet and A. Nemirovskii. *LMI Lab: A Package for Manipulating and Solving LMIs.* INRIA, 1993.

[31] K.-C. Goh. *Robust Control Synthesis via Bilinear Matrix Inequalities.* PhD thesis, University of Southern California, May 1995.

[32] K. C. Goh, J. H. Ly, L. Turand, and M. G. Safonov. μ/k_m-synthesis via bilinear matrix inequalities. In *Proceedings of the 33rd IEEE Conference on Decision and Control*, December 1994.

[33] K. C. Goh, M. G. Safonov, and J. H. Ly. A global optimization approach for the BMI problem. In *Proc. IEEE Conf. on Decision and Control*, Lake Buena vista, FL, December 1994.

[34] K. C. Goh, M. G. Safonov, and J. H. Ly. Robust synthesis via bilinear matrix inequalities. 1995. Submitted to the *International Journal of Robust and Nonlinear Control*.

[35] K. C. Goh, L. Turan, M. G. Safonov, G. P. Papavassilopoulos, and J. H. Ly. Biaffine matrix inequality properties and computational methods. In *Proc. American Control Conf.*, pages 850–855, 1994.

[36] G. Golub and C. Van Loan. *Matrix Computations.* Johns Hopkins Univ. Press, Baltimore, second edition, 1989.

[37] G. H. Golub, P. Milanfar, and J. Varah. A stable numerical method for inverting shape from moments. *SIAM J. Sci. Comput.*, 21(4):1222–1243, 1999.

[38] J. W. Goodman. *Statistical Optics.* John Wiley & Sons, 1985.

[39] I. F. Gorodnitsky and B. D. Rao. Sparse signal reconstruction from limited data using FOCUSS: a re-weighted minimum norm algorithm. *IEEE Trans. Signal Processing*, 45(3):600–616, March 1997.

[40] P. E. Green, F. J. Carmone, and S. M. Smith. *Multidimensional scaling: Concepts and Applications.* Allyn & Bacon, Boston, MA, 1989.

[41] K. M. Grigoriadis and E. B. Beran. Alternating projection algorithms for linear matrix inequalities problems with rank constraints. In L. El Ghaoui and S. Niculescu, editors, *Advances in Linear Matrix Inequality Methods in Control*, chapter 13, pages 251–267. SIAM, 2000.

[42] M. Grötschel, L. Lovász, and A. Schrijver. *Geometric Algorithms and Combinatorial Optimization.* Algorithms and Combinatorics. Springer-Verlag, New York, 1988.

[43] L. G. Gubin, B. T. Polyak, and E. V. Raik. The method of projections for finding the common point of convex sets. *USSR Comp. Math. Phys.*, (7):1–24, 1967.

[44] A. Hassibi, J. How, and S. Boyd. Low-authority controller design via convex optimization. *AIAA J. Guidance, Control, and Dynamics*, 22(6):862–72, November 1999.

[45] A. Hassibi, J. P. How, and S. P. Boyd. A path-following method for solving BMI problems in control. In *Proceedings of American Control Conference*, volume 2, pages 1385–9, June 1999.

[46] A. Helmersson. IQC synthesis based on inertia constraints. In *Proceedings of the 14th IFAC conference*, pages 163–168, 1999.

[47] A. J. Helmicki, C. A. Jacobsen, and C. N. Nett. Worst-case/deterministic identification in $\mathcal{H}_\infty$: The continuous time case. *IEEE Trans. Aut. Control*, AC-37(5):604–610, May 1992.

[48] D. Henrion, S. Tarbouriech, and M. Sebek. Rank-one lmi approach to simultaneous stabilization of linear systems. *Syst. Control Letters*, 38(2):79–89, 1999.

[49] J.-B. Hiriart-Urruty and C. Lemaréchal. *Convex Analysis and Minimization Algorithms II: Advanced Theory and Bundle Methods*, volume 306 of *Grundlehren der mathematischen Wissenschaften.* Springer-Verlag, New York, 1993.

[50] R. Horn and C. Johnson. *Topics in Matrix Analysis.* Cambridge University Press, Cambridge, 1991.

[51] R. A. Horn and C. A. Johnson. *Matrix Analysis.* Cambridge University Press, 1985.

[52] P. J. Huber. *Robust Statistics*. Wiley, New York, 1981.

[53] P. Ishwar, K. Ratakonda, P. Moulin, and N. Ahuja. Image denoising using multiple compaction domains, 1998.

[54] T. Iwasaki. The dual iteration for fixed-order control. *IEEE Trans. Aut. Control*, 44(4):783–788, 1999.

[55] T. Iwasaki and R. E. Skelton. All controllers for the general $\mathbf{H}_\infty$ control problem: LMI existence conditions and state space formulas. *Automatica*, 30(8):1307–1317, 1994.

[56] R. N. Jacques. *On-line system identification and control design for flexible structures*. PhD thesis, Department of Aeronautics and Astronautics, MIT, 1994.

[57] R. E. Kalman. System identification from noisy data. In A. R. Bednarek and L. Cesari, editors, *Dynamical systems II*, pages 135–164. Academic press, New York, 1982.

[58] D. Karger, R. Motwani, and M. Sudan. Approximate graph coloring by semidefinite programming. Technical report, Department of Computer Science, Stanford University, 1994.

[59] M.G. Kendall. *Multivariate Analysis*. New York, MacMillan, 1980.

[60] U. Kirsch. *Structural Optimization, Fundamentals and Applications*. Springer-Verlag, 1993.

[61] K. Kreutz-Delgado and B. D. Rao. Measures and algorithms for best basis selection. In *ICASSP Proceedings*, 1998.

[62] R. Kumaresan and D. W. Tufts. Estimation of arrival of multiple plane waves. *IEEE Transactions on on Aerospace and Electronic Systems*, AES-19:123–133, 1983.

[63] M. S. Lobo, M. Fazel, and S. Boyd. Portfolio optimization with linear and fixed transaction costs and bounds on risk. Submitted to *Operations Research*, June 2000.

[64] M. S. Lobo, L. Vandenberghe, S. Boyd, and H. Lebret. Applications of second-order cone programming. *Linear Algebra and Applications*, 284(1-3):193–228, November 1998.

[65] D. G. Luenberger. *Linear and Nonlinear Programming*. Addison-Wesley, Reading, Mass., 2nd edition, 1984.

[66] R. Mathar and R. Meyer. Preorderings, monotone functions, and best rank r approximations with applications to classical mds. *Journal of statistical planning and inference*, 37:291–305, 1993.

[67] T. McKelvey. *Identification of state-space models from time and frequency data.* PhD thesis, Linköping University, Sweden, Department of Electrical Engineering, 1995.

[68] M. Mesbahi and G. P. Papavassilopoulos. On the rank minimization problem over a positive semidefinite linear matrix inequality. *IEEE Transactions on Automatic Control*, AC-42(2):239–43, February 1997.

[69] P. Milanfar, G. C. Verghese, W. C. Karl, and A. S. Willsky. Reconstructing polygons from moments with connections to array processing. *IEEE Transactions on Signal Processing*, 43(2):432–442, February 1995.

[70] J. More and Z. Wu. Distance geometry optimization for protein structures. Technical Report MCS-P628-1296, Mathematics and Computer Science division, Argonne national labratory, 9700 South Cass Ave, Argonne, IL 60439, December 1996.

[71] J. More and Z. Wu. Issues in large-scale global molecular optimization. Technical Report MCS-P539-1059, Mathematics and Computer Science division, Argonne National Labratory, 9700 South Cass Ave, Argonne, IL 60439, March 1996.

[72] K. Nakata, K. Fujisawa, M. Fukuda, M. Kojima, and K. Murota. Exploiting sparsity in semidefinite programming via matrix completion II: Implementation and numerical results, 2001.

[73] Y. Nesterov and A. Nemirovskii. *Interior-Point Polynomial Algorithms in Convex Programming*. SIAM Studies in Applied Mathematics. SIAM, 1994.

[74] Yu. Nesterov and A. Nemirovsky. *Optimization over positive semidefinite matrices: Mathematical background and user's manual*. USSR Acad. Sci. Centr. Econ. & Math. Inst., 32 Krasikova St., Moscow 117418 USSR, 1990.

[75] P. Van Overschee and B. De Moor. *Subspace Identification for Linear Systems: Theory, Implementation, Applications*. Kluwer, 1996.

[76] M. Overton and H. Wolkowicz. Semidefinite programming (foreword). *Mathematical Programming*, 77(2):105–110, May 1997.

[77] T. E. Pare. *Analysis and control of nonlinear systems*. PhD thesis, Dept. of Aeronautics and Astronautics, Stanford University, August 2000.

[78] P. A. Parrilo. *Structured Semidefinite Programs and Semialgebraic Geometry Methods in Robustness and Optimization.* PhD thesis, California Institute of Technology, Pasadena, California, 2000.

[79] S. Poljak, F. Rendl, and H. Wolkowicz. A recipe for best semidefinite relaxation for $(0, 1)$-quadratic programming. *Journal of Global Optimization*, 1994. To appear.

[80] J. Rissanen. Modeling by shortest data description. *Automatica*, 14:465–471, 1978.

[81] M. G. Safonov, K. C. Goh, and J. H. Ly. Control system synthesis via bilinear matrix inequalities. In *Proceedings of American Control Conference*, volume 1, pages 45–9, Baltimore, MD, 1994.

[82] I. J. Schoenberg. Remarks to Maurice Frechet's article: Sur la d'efinition axiomatique d'une classe d'espace distances vectoriellement applicable sur l'espace de hilbert. *The Annals of Mathematics*, 36(3):724–732, July 1935.

[83] M. Sezan and H. Stark. Incorporation of a priori moment information into signal recovery and synthesis problems. *J. Math. Anal. Appl*, 122:172–186, 1987.

[84] N. Z. Shor. Quadratic optimization problems. *Soviet Journal of Circuits and Systems Sciences*, 25(6):1–11, 1987.

[85] R. E. Skelton, T. Iwasaki, and K. Grigoriadis. *A Unified Algebraic Approach to Linear Control Design.* Taylor and Francis, 1998.

[86] E. D. Sontag. *Mathematical Control Theory*, volume 6 of *Texts in Applied Mathematics.* Springer-Verlag, 1990.

[87] V. N. Strakhov and M. A. Brodsky. On the uniqueness of the inverse logarithmic potential problem. *SIAM Journal on Applied Mathematics*, 46(2):324–344, 1986.

[88] J. F. Sturm. Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones. *Optimization Methods and Software*, 11–12:625–653, 1999. Special issue on Interior Point Methods (CD supplement with software).

[89] K. C. Toh and M. Kojima. Solving some large scale semidefinite programs via the conjugate residual method. Technical report, National University of Singapore, 2000.

[90] M. W. Trosset. Distance matrix completion by numerical optimization. Technical report, Dept. of Computational and Applied mathematics, Rice University, Houston, TX 77005-1892, June 1997.

[91] L. Vandenberghe and S. Boyd. SP*: Software for Semidefinite Programming. User's Guide, Beta Version.* Dept. of Electrical Engineering, Stanford University, October 1994. Available at `www.stanford.edu/~boyd`.

[92] L. Vandenberghe and S. Boyd. Semidefinite programming. *SIAM Review*, 38(1):49–95, March 1996.

[93] L. Vandenberghe, S. Boyd, and S.-P. Wu. Determinant maximization with linear matrix inequality constraints. Technical report, Information Systems Laboratory. Stanford University, February 1996.

[94] L. Vandenberghe, S. Boyd, and S.-P. Wu. Determinant maximization with linear matrix inequality constraints. *SIAM J. on Matrix Analysis and Applications*, 19(2):499–533, April 1998.

[95] R. J. Vanderbei. *Linear Programming, Foundations and Extensions.* Kluwer, 1996.

[96] Y. T. Wang. *Automated design of phase-shifting masks for microlithography.* PhD thesis, Dept. of Electrical Engineering, Stanford University, June 1997.

[97] M. Wax and T. Kailath. Detection of signals by information theoretic criteria. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 33(2):387–392, April 1985.

[98] H. wolkowicz. Simple efficient solutions for semidefinite programming. Technical Report CORR 2001-49, University of Waterloo, 2001.

[99] K. Woodgate. An upper bound on the number of linear relations identified from noisy data using the frisch scheme. *Syst. Control Letters*, 24:153–158, 1995.

[100] S.-P. Wu and S. Boyd. Design and implementation of a parser/solver for SDPs with matrix structure. In *Proceedings of the 1996 IEEE International Symposium on Computer-Aided Control System Design*, pages 240–245, Dearborn, Michigan, 1996.

[101] S.-P. Wu and S. Boyd. SDPSOL: *A Parser/Solver for Semidefinite Programming and Determinant Maximization Problems with Matrix Structure. User's Guide, Version Beta.* Stanford University, June 1996.

[102] D. C. Youla. Generalized image restoration by the method of alternating orthogonal projections. *IEEE Trans. Circuits Syst.*, CAS-25(9):694–702, September 1978.

[103] D. C. Youla and H. Webb. Image restoration by the method of convex projections. *IEEE Transactions on Medical Imaging*, (1):89–94, 1982.