Title
Structured Low-Rank Matrix Approximation in Signal Processing: Semidefinite Formulations and Entropic First-Order Methods

Author
Chao, Hsiao-Han

Publication Date
2018

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA
Los Angeles

Structured Low-Rank Matrix Approximation in Signal Processing: Semidefinite Formulations and Entropic First-Order Methods

A dissertation submitted in partial satisfaction
of the requirements for the degree
Doctor of Philosophy in Electrical and Computer Engineering

by

Hsiao-Han Chao

2018

ABSTRACT OF THE DISSERTATION

Structured Low-Rank Matrix Approximation in Signal Processing: Semidefinite Formulations and Entropic First-Order Methods

by

Hsiao-Han Chao

Doctor of Philosophy in Electrical and Computer Engineering

University of California, Los Angeles, 2018

Professor Lieven Vandenberghe, Chair

Applications of semidefinite optimization in signal processing are often derived from the Kalman–Yakubovich–Popov lemma and its extensions, which give sum-of-squares theorems of nonnegative trigonometric polynomials and generalized polynomials. The dual semidefinite programs involve optimization over positive semidefinite matrices with Toeplitz structure or extensions of the Toeplitz structure. In recent applications, these techniques have been used in continuous-domain sparse signal approximations. These applications are commonly referred to as super-resolution, gridless compressed sensing, continuous 1-norm, or total-variation norm minimization. The semidefinite formulations of these problems introduce a large number of auxiliary variables and are expensive to solve using general-purpose or even customized interior-point solvers.

The thesis can be divided into two parts. As a first contribution, we extend the semidefinite penalty formulations in super-resolution applications to more general types of structured low-rank matrix approximations. The penalty functions for structured symmetric and nonsymmetric matrices are discussed. The connection via duality between these penalty functions and the (generalized) Kalman–Yakubovich–Popov lemma from linear system theory is further clarified, which leads to a more systematic proof for the equivalent semidefinite formulations. In the second part of the thesis, we propose a new class of efficient first-order splitting methods based on an appropriate choice of a generalized distance function, the

Itakura–Saito distance, for optimizations over the cone of nonnegative trigonometric polynomials. The Itakura–Saito distance is the Bregman distance defined by the negative entropy function. The choice for this distance function is motivated by the fact that the associated generalized projection on the set of normalized nonnegative trigonometric polynomials can be computed at a cost that is roughly quadratic in the degree of the polynomial. This should be compared to the cubic per-iteration-complexity of standard first-order methods (the cost of a Euclidean projection on the positive semidefinite cone) and customized interior-point solvers. The quadratic complexity is confirmed by numerical experiments with Auslender and Teboulle's accelerated proximal gradient method for Bregman distances.

The dissertation of Hsiao-Han Chao is approved.

Tetsuya Iwasaki

Paulo Tabuada

Kung Yao

Lieven Vandenberghe, Committee Chair

University of California, Los Angeles

2018

To My Family

TABLE OF CONTENTS

1 Introduction **1**

2 Semidefinite Duality and the KYP Lemma **4**

2.1 Conic duality 4

2.2 SDP theorem of alternatives 5

2.3 Nonnegative trigonometric polynomials 7

2.3.1 Spectral factorization 7

2.3.2 LMI characterizations 8

2.3.3 Duality: positive semidefinite Toeplitz matrices 10

2.4 Structured positive semidefinite matrix factorization 12

2.4.1 Main result: conic decomposition 14

2.5 Examples 15

2.5.1 Trigonometric polynomials 15

2.5.2 Polynomials 18

2.5.3 Rational functions 20

2.6 Duality 21

2.7 Kalman–Yakubovich–Popov Lemma 22

2.7.1 Strict version 23

2.7.2 Nonstrict version 24

2.7.3 Linear time-invariant systems 25

3 Gauge Penalties for Structured Symmetric Matrices **27**

3.1 Sparse signal reconstruction 29

3.2 Related works 32

3.3 Semidefinite representation of gauges for structured symmetric matrices . . . 34
3.4 Duality . . . 38
3.4.1 Conjugate of symmetric matrix gauge . . . 38
3.4.2 Dual problem interpretation . . . 40
3.5 Line spectrum estimation examples . . . 41
3.5.1 Gaussian white noise model . . . 42
3.5.2 Moving average noise model . . . 43

4 Gauge Penalties for Structured Nonsymmetric matrices . . . 47
4.1 Semidefinite representation of gauges for structured nonsymmetric matrices . 48
4.2 Duality . . . 53
4.2.1 Conjugate of nonsymmetric matrix gauge . . . 54
4.2.2 Dual problem interpretation . . . 55
4.3 Numerical examples . . . 56
4.3.1 Line spectrum estimation by penalty approximation . . . 56
4.3.2 Direction of arrival estimation . . . 59
4.3.3 Direction of arrival from multiple measurement vectors . . . 61
4.3.4 Structured matrix decomposition . . . 63

5 Itakura–Saito Generalized Distance . . . 66
5.1 Forward and backward Levinson–Durbin algorithm . . . 68
5.1.1 Levinson–Durbin algorithm . . . 69
5.1.2 Jury stability test . . . 72
5.1.3 Factorization of Jury matrix . . . 73
5.2 Entropy . . . 74
5.2.1 Semidefinite representations . . . 76

5.2.2 Gradients . . . 79
5.2.3 Legendre property . . . 80
5.3 Itakura–Saito distance . . . 81
5.3.1 Itakura–Saito and Kullback–Leibler distance . . . 81
5.3.2 Strong convexity . . . 82

6 Entropic Proximal Operators for Nonnegative Trigonometric Polynomials 84
6.1 Entropic proximal operators . . . 86
6.1.1 Projection . . . 88
6.1.2 Proximal operator . . . 90
6.2 Numerical experiments . . . 91
6.2.1 Covariance estimation . . . 91
6.2.2 Euclidean projection on nonnegative polynomials . . . 93

7 Conclusion . . . 98

A Classical methods for line spectrum estimation . . . 99

B Subsets of the complex plane . . . 107

C Matrix factorization results . . . 112

D Strict feasibility . . . 117

E Generalized proximal gradient method . . . 122

References . . . 126

LIST OF FIGURES

3.1 Line spectrum estimation by Toeplitz covariance fitting (Section 3.5.1). The red dots represent the frequencies and magnitudes of the true model. The blue lines show the estimated parameters obtained by solving (3.36). 42

3.2 Line spectrum estimation by Toeplitz covariance fitting (Section 3.5.2) with white noise model (MA(0)) (*Top*) and with MA(3) model (*Bottom*). The red (blue) stems represent the true (estimated) line spectrum, and the red (blue) curve represents the true (estimated) noise spectrum. The dotted vertical line indicates the cutoff frequency $\omega_c = \pi/6$. 46

4.1 The data for the example in Section 4.3.1. The red dashed lines show the exact, noise-free signal. The circles show the signal corrupted by Gaussian white noise (in blue), plus a few larger errors in 20 positions (in black). The green dots show the recovered signal y from (4.19). 57

4.2 Line spectrum models estimated from the signal in Figure 4.1 by solving the optimization problem (4.19) (*Left*) and using the matrix pencil method (*Right*). 58

4.3 Direction-of-arrival estimation with (*Left*) and without (*Right*) interval constraints (Section 4.3.2). 61

4.4 Comparison of recovery rate for different number of available measurements with interval constraints (red) and without (blue), in the example of Section 4.3.2. . . 62

4.5 The results with 1 (*Top*), 15 (*Middle*) and 30 (*Bottom*) measurement vector(s) in the DOA estimation problem of Section 4.3.3. The figures on the right show the magnitude of the trigonometric polynomials obtained from the dual solution. The red dots show the true directions of arrival (and magnitudes). 64

4.6 Structured matrix decomposition of a matrix with rank 3 plus a Gaussian noise matrix, with (*Left*) and without (*Right*) interval constraint (Section 4.3.4). . . . 65

5.1 Contour lines of the function $\phi(1, x_1, x_2)$ defined in (5.26) (*Left*), and contour lines of the function $\tilde{\phi}(1, x_1, x_2)$ defined in (5.49) (*Right*), on the set $\{(x_1, x_2) \mid (1, x_1, x_2) \in K\} = \{(x_1, x_2) \mid 1 + 2x_1 \cos\omega + 2x_2 \cos 2\omega \geq 0 \; \forall\omega\}$ 82

6.1 True signal (solid line) and noisy samples (circles). 93

6.2 True and estimated line spectrum (red circles and blue stem lines, respectively) and dual optimal polynomial $F_x(e^{\mathrm{j}\omega})$ (solid curve). 94

6.3 Relative suboptimality versus iteration number of the generalized proximal gradient method applied to the dual problem (6.18). The optimality gap is computed as $(f(x^k) - f^{\mathrm{opt}})/|f^{\mathrm{opt}}|$, where $f(x)$ denotes the negative of the dual objective value in (6.18) and f^{opt} is the optimal solution computed by CVX. 94

6.4 Euclidean projection on the normalized nonnegative trigonometric polynomials of order $p = 9$. The red curve is $F_a(e^{\mathrm{j}\omega})$; the blue curve is $F_x(e^{\mathrm{j}\omega})$, where x is the Euclidean projection of a on $\{x \in K \mid x_0 = 1\}$. 95

6.5 Time for proximal gradient method and general-purpose interior-point methods (IPM) versus problem size (*Left*), and time per iteration for the proximal gradient method (*Right*). 96

6.6 Time for the proximal gradient method (*Left*) and time per iteration (*Right*) versus problem size. 97

LIST OF TABLES

B.1 Common choices of Ψ with $\Phi = \Phi_u$ (λ on the unit circle). 108

B.2 Common choices of Ψ with $\Phi = \Phi_i$ (λ imaginary). 109

B.3 Common choices of Ψ with $\Phi = \Phi_r$ (λ real). 109

ACKNOWLEDGMENTS

First and foremost, I thank my advisor, Professor Lieven Vandenberghe, to whom this dissertation owes its existence. In his course on convex optimization, he introduced and started my interest in the subject through his elegant and clear exposition. During my graduate study, his efforts to always delve deeper and think out of the box have helped tremendously. It would only do justice to give all credits to him, and I am deeply grateful and honored that he has let me participate in the work. His determination and dedication as a researcher and educator has also been a great inspiration. He never lacks the courage to break and rebuild when it comes to polishing works in his hands, be it a paper, a presentation, or a course assignment. He treats how his students are learning with extra care, by providing appropriate context and motivation, as well as by taking and addressing their questions seriously. Moreover, I appreciate the examples he has set as a leader and person. He holds the highest standard for himself yet remains down-to-earth, understanding, and gracious toward the undeserving. I am forever thankful for the opportunity to work with and to learn first-hand from Professor Vandenberghe.

In addition, I would like to thank my committee members, Professors Tetsuya Iwasaki, Paulo Tabuada, and Kung Yao, for their help and scholarly examples. I also thank the staff of ECE department for their helpful assistance and student-centered service.

I am very grateful for my labmates, Daniel O'Connor, Yifan Sun, Jinchao Li, Rong Rong, Cameron Gunn, and Xin Jiang, for the discussions and fun times we have shared. Moreover, I have been immensely blessed by roommates and friends during my time at UCLA. They have held me up through prayers, encouragements, or even accommodations. I sincerely thank them for sharing their lives with me.

For my parents' unconditional love and support throughout my life, I am forever indebted to them. Many thanks also for Hsin-Hao's care, encouragements, and admonishments for me in times of need. In addition, I thank Mochi for his company during the last stages of my study, for the joy and motivation he has brought me. Finally, I give thanks to God for orchestrating events in my life, and for humbling and sustaining me throughout.

VITA

2011	B.S. (Electrical Engineering and Physics) National Taiwan University
2013	M.S. (Electrical Engineering) UCLA
2013	Teaching Assistant Programming in Computing UCLA
2013–2017	Teaching Assistant Electrical Engineering Department UCLA
2014–present	Graduate Student Researcher Electrical Engineering Department UCLA

PUBLICATIONS

H.-H. Chao and L. Vandenberghe. *Extensions of semidefinite programming methods for atomic decomposition*. In *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2016.

H.-H. Chao and L. Vandenberghe. *Semidefinite representations of gauge functions for structured low-rank matrix decomposition*. *SIAM Journal on Optimization*, 27:1362–1389, 2017.

H.-H. Chao and L. Vandenberghe. *Entropic proximal operators for nonnegative trigonometric polynomials*. 2018. Submitted for publication.

CHAPTER 1

Introduction

Non-polyhedral conic extensions of linear programming (LP), such as second-order cone programming (SOCP) and semidefinite programming (SDP), have been widely studied in the last few decades, for applications in many areas including operations research, statistics, machine learning, and engineering. The popularity is mainly due to two reasons. The first reason is the development of reliable and efficient solvers, which are mostly based on some variation of interior-point methods (IPMs) [Kar84, Meh92, NN94]. Theoretically IPMs have a worst-case polynomial complexity, and in practice they are quite efficient and robust. Secondly, despite the very simple standard forms of LP, SOCP and SDP, there exists a surprising variety of problems they can cover, either through an equivalent reformulation, approximation, or as a heuristic. See [BN01, BV04] for example, for modeling problems arising in a wide spectrum of applications. In fact, powerful modeling softwares for convex optimization, such as CVX [GB14] and CVXPY [DB16], transform problems into these three forms.

LP, SOCP and SDP all fit in to the conic optimization framework, with specific choices of the cone. The conic linear program has the following standard form

$$\begin{array}{ll} \text{minimize} & \langle c, x\rangle \\ \text{subject to} & \mathcal{A}(x) = b \\ & x \in \mathcal{K}, \end{array} \tag{1.1}$$

with variable $x \in \mathbf{E}$, where $\mathbf{E}$ is a finite-dimensional vector space over $\mathbf{R}$, and $\langle \cdot, \cdot \rangle$ denotes an inner product on $\mathbf{E}$. The problem data are $c \in \mathbf{E}$, $b \in \mathbf{R}^m$, linear mapping $\mathcal{A} : \mathbf{E} \to \mathbf{R}^m$, and a proper cone $\mathcal{K} \subset \mathbf{E}$ (the definition is given in Section 2.1). With the choices $\mathbf{E} = \mathbf{R}^n$ and $\mathcal{K} = \mathbf{R}^n_+$ (the nonnegative orthant), the mapping can be represented as $\mathcal{A}(x) = Ax$ with

$A \in \mathbf{R}^{m\times n}$, and problem (1.1) reduces to an LP. With the same choices of $\mathbf{E}$ and the mapping $\mathcal{A}$, and taking $\mathcal{K} = K^{n_1} \times \cdots \times K^{n_l}$, where $K^{n_k} = \{y \in \mathbf{R}^{n_k} \mid \sqrt{y_1^2 + \cdots + y_{n_k-1}^2} \leq y_{n_k}\}$ and $\sum_{k=1}^{l} n_k = n$, problem (1.1) reduces to an SOCP (although its dual form is more commonly seen). If we choose $\mathbf{E} = \mathbf{S}^n$, the space of $n \times n$ symmetric matrices, and $\mathcal{K} = \mathbf{S}^n_+$, the positive semidefinite (PSD) cone, the mapping can be defined as $\mathcal{A}(x) = (\langle A_1, x\rangle, \ldots, \langle A_m, x\rangle)$ with $A_1, \ldots, A_m \in \mathbf{S}^n$, and problem (1.1) reduces to an SDP.

The Lagrange dual of (1.1) is also a conic LP

$$\begin{array}{ll} \text{maximize} & -\langle b, z\rangle \\ \text{subject to} & c + \mathcal{A}^{\text{adj}}(z) \in \mathcal{K}^*, \end{array} \tag{1.2}$$

with variable $z \in \mathbf{R}^m$, and here $\langle \cdot, \cdot \rangle$ denotes an inner product on $\mathbf{R}^m$. The adjoint $\mathcal{A}^{\text{adj}} : \mathbf{R}^m \to \mathbf{E}$ is the linear mapping satisfying $\langle \mathcal{A}(x), z\rangle = \langle x, \mathcal{A}^{\text{adj}}(z)\rangle$ for all $x \in \mathbf{E}$ and $z \in \mathbf{R}^m$, and $\mathcal{K}^* \subset \mathbf{E}$ denotes the cone dual to $\mathcal{K}$ with respect to the inner product on $\mathbf{E}$. With the standard choices of inner products, the corresponding cones of LP, SOCP and SDP are self-dual. Popular general-purpose solvers like SDPT3 [TTT02], SeDuMi [Stu99], and MOSEK [MOS02] are based on implementations of primal-dual IPMs for LP, SOCP and SDP.

Many SDP applications in signal processing and control are based on SDP formulations of problems involving the cone of nonnegative trigonometric polynomials, its dual cone, the PSD Toeplitz matrices, and their extensions [WBV96, SMM00, DTS01, DLS02, AV02, Hac03, Dum07, GL08]. These applications were made possible by the development of IPMs for SDP in the 1990s. Dumitrescu provides in his book [Dum07] a good overview of signal processing applications of nonnegative trigonometric polynomials for the first 15 years since then. More recently, there have emerged new SDP applications to problems of continuous-domain sparse signal approximation, commonly referred to as super-resolution, gridless compressed sensing, continuous 1-norm, or total-variation norm minimization [CG12, CF14, BTR13, TBS13, YX16, LC16, YX15, CC15, CGH17, MCK15]. Interestingly, these problems are related to the earlier formulations of nonnegative polynomials by duality. Owing to the fact that LMI representations of these cones introduce a large number of auxiliary variables, in either the classical and the more recent applications,

there has been a strong research activity to develop faster solvers than the general-purpose ones [AV00, AV02, Hac03, RV06, LV07, HV14, BTR13].

The thesis can be divided into two subjects. First, inspired by the connection of the SDPs used in super-resolution applications with the bounded real lemma, we extend the SDP penalty formulations to more general types of structured low-rank matrix approximations, as well as clarify their connections with the (generalized) Kalman–Yakubovich–Popov (KYP) lemma [Kal63, Yak62, Pop62, Ran96, IMF00, BV02, BV03, IH05, Sch06, PV11] from linear system theory. More specifically, in Chapter 2, we give the LMI characterization and conic decomposition of the extensions to the PSD Toeplitz matrices, and explain the connections with the KYP lemma. In Chapter 3 and 4, for symmetric and nonsymmetric matrices, respectively, the SDP formulations of gauge penalty functions of continuous 1-norm for structured low-rank matrix approximation are presented, and their primal-dual interpretations are explained. This part of work has appeared in the papers [CV16, CV17].

The second subject of the thesis addresses the algorithmic aspect. Earlier works in this area include customized IPMs that exploit the structures of the SDPs involving nonnegative trigonometric polynomials and extensions, which achieve $O(n^3)$ per-iteration complexity [AV00, AV02, Hac03, RV06, LV07, HV14], as well as first-order splitting methods that depend on a Euclidean projection on the PSD cone and therefore cannot offer improvement over the $O(n^3)$ per-iteration complexity (the cost of an eigenvalue decomposition) [BTR13]. We develop a new class of efficient first-order splitting methods with an appropriate choice of the Bregman distance function. We show that its complexity is at least as efficient as $O(n^2)$ per iteration, as compared to the earlier works of $O(n^3)$, whereas general-purpose IPMs exhibit $O(n^6)$ and $O(n^4)$ per-iteration complexities on the primal and dual SDP formulations, respectively. In Chapter 5, we present the Itakura–Saito distance, and a pair of entropy functions associated with it, as well as classical results useful for their calculations. In Chapter 6, we discuss first-order methods based on generalized proximal operators defined in terms of the Itakura–Saito distance, and include numerical results with Auslender and Teboulle's accelerated proximal gradient method for Bregman distances.

Chapter 7 concludes the thesis with some final remarks.

CHAPTER 2

Semidefinite Duality and the KYP Lemma

The purpose of this chapter is to provide enough background for the development of the following chapters. While the main result directly needed can be found in Section 2.4, the other sections give relevant explanations on the duality in the Kalman–Yakubovich–Popov (KYP) lemmas. Specifically, the sections are divided as the following. Sections 2.1 and 2.2 provide a short review on conic duality and the theorems of alternatives. In Section 2.3, we first discuss the duality between the nonnegative trigonometric polynomials and the positive semidefinite Toeplitz matrices. Then as a nontrivial extension, the duality between the cone of nonnegative Popov functions and a convex cone of structured positive semidefinite matrices is explained. The conic decomposition of these structured positive semidefinite matrices, which is the main result necessary for the development of Chapter 3 and 4, is presented in Section 2.4. Specific examples of the decomposition are given in Section 2.5, and the duality is pointed out in Section 2.6. It turns out that this decomposition result also constitutes a key step in the proofs (that are based on SDP duality) of the KYP lemma and its generalizations [Ran96, IMF00, BV02, BV03, IH05, Sch06, PV11]. The KYP lemmas provide the foundation for the characterization of nonnegative Popov functions as linear matrix inequalities (LMIs). For completeness, the general forms of the KYP lemmas and the proofs are given in Section 2.7.

2.1 Conic duality

This section provides some background on convex cones and duality. The definitions and properties are standard and can be found, for example, in [BN01, §2] [BV04, §2].

Let $\mathbf{E}$ denote a finite-dimensional vector space over real numbers $\mathbf{R}$. A convex cone $\mathcal{K} \subset \mathbf{E}$ is a convex set that is invariant under nonnegative scalar multiplication, *i.e.*, given any $x, y \in \mathcal{K}$ and scalars $\alpha, \beta \geq 0$, it holds that $\alpha x + \beta y \in \mathcal{K}$. A set is *closed* if and only if it contains the limit point of every convergent sequence in it. A convex cone $\mathcal{K}$ is *pointed* if and only if it contains no line, *i.e.*, $x \in \mathcal{K}$ and $-x \in \mathcal{K}$ only if $x = 0$. A set is called *solid* if and only if it has a nonempty interior.

Definition 2.1 (**Proper cone**)**.** A cone is *proper* if it is convex, closed, pointed, and solid.

It is useful to consider the dual cones for several reasons. As pointed out in Chapter 1, dual cones appear in the dual conic optimization problems.

Definition 2.2 (**Dual cone**)**.** The dual cone of a cone $\mathcal{K}$ is defined as

$$\mathcal{K}^* = \{z \in \mathbf{E} \mid \langle x, z \rangle \geq 0 \ \ \forall x \in \mathcal{K}\},$$

where $\langle \cdot, \cdot \rangle$ denotes an inner product on $\mathbf{E}$. The dual cone $\mathcal{K}^*$ is by definition closed and convex even if $\mathcal{K}$ is not.

The following properties regarding the dual cones are relevant to the rest of the chapter.

Lemma 2.1. *The dual cone $\mathcal{K}^*$ is proper if and only if $\mathcal{K}$ is proper.*

Lemma 2.2. *If $\mathcal{K}$ is a closed convex cone, then $\mathcal{K}^{**} = \mathcal{K}$. The cones $\mathcal{K}$ and $\mathcal{K}^*$ are therefore a pair of dual cones.*

Moreover, if the dual cones are identical, *i.e.*, $\mathcal{K}^* = \mathcal{K}$, it is said to be self-dual. Examples of self-dual proper cones are the nonnegative orthant, the second-order cone, the positive semidefinite cone, and their direct products.

2.2 SDP theorem of alternatives

The theorems of alternatives deal with feasibility problems involving systems of (generalized) inequalities and equalities. There exists a rich literature on different variants and

applications of theorems of alternatives. The best known one is the Farkas' lemma for linear feasibility problems [Far02]. More background and non-polyhedral conic extensions can be found in [Ben69,BB71,CK77,BW81]. Results on LMIs appear in [Wol81,Las95,Las97,BV03].

For our purpose we present three theorems of alternatives for LMIs in [BV03], which will be useful for the derivations in the following sections. Denote the space of block diagonal Hermitian matrices as $\mathcal{H} = \mathbf{H}^{n_1} \times \cdots \times \mathbf{H}^{n_l}$, with the standard inner product

$$\langle \mathbf{diag}(X_1, \ldots, X_l), \mathbf{diag}(Z_1, \ldots, Z_l) \rangle = \sum_{k=1}^{l} \mathbf{tr}\, X_k Z_k.$$

Let $\mathcal{A} : \mathbf{E} \to \mathcal{H}$ be a linear mapping and its adjoint $\mathcal{A}^{\mathrm{adj}} : \mathcal{H} \to \mathbf{E}$ be the linear mapping satisfying $\langle \mathcal{A}(x), Z \rangle = \langle x, \mathcal{A}^{\mathrm{adj}}(Z) \rangle$ for all $x \in \mathbf{E}$ and $Z \in \mathcal{H}$, and let $A_0 \in \mathcal{H}$.

Theorem 2.1. *Exactly one of the following statements is true.*

(a) There exists an $x \in \mathbf{E}$ such that $\mathcal{A}(x) + A_0 \succ 0$.

(b) There exists a $Z \in \mathcal{H}$ with $0 \neq Z \succeq 0$ such that $\mathcal{A}^{\mathrm{adj}}(Z) = 0$ and $\langle A_0, Z \rangle \leq 0$.

Theorem 2.2. *Exactly one of the following statements is true.*

(a) There exists an $x \in \mathbf{E}$ such that $0 \neq \mathcal{A}(x) \succeq 0$.

(b) There exists a $Z \in \mathcal{H}$ with $Z \succ 0$ such that $\mathcal{A}^{\mathrm{adj}}(Z) = 0$.

Theorem 2.3. *At most one of the following statements is true.*

(a) There exists an $x \in \mathbf{E}$ such that $\mathcal{A}(x) + A_0 \succeq 0$.

(b) There exists a $Z \in \mathcal{H}$ with $Z \succeq 0$ such that $\mathcal{A}^{\mathrm{adj}}(Z) = 0$ and $\langle A_0, Z \rangle < 0$.

Moreover, if $\mathcal{A}(x) \succeq 0$ implies $\mathcal{A}(x) = 0$, then exactly one of the two statements is true.

Theorems 2.1 and 2.2 present strong alternatives, whereas Theorem 2.3 presents weak alternatives that become strong alternatives if a certain condition holds.

2.3 Nonnegative trigonometric polynomials

Trigonometric polynomials are the simplest examples of Popov functions (whose formal definition comes later in Section 2.7) with many important applications. The material presented here has been well known in signal processing and control [SMM00, DTS01, AV02, Dum07]. We include this section as a tangible and classical example, and in particular, as a preview for the rest of the chapter.

Consider a vector of complex exponentials parametrized by $\omega \in [0, 2\pi)$:

$$a(e^{\mathrm{j}\omega}) = (1, e^{\mathrm{j}\omega}, e^{\mathrm{j}2\omega}, \ldots, e^{\mathrm{j}p\omega}),$$

and define an inner product on $\mathbf{R} \times \mathbf{C}^p$:

$$\begin{aligned} \langle x, z \rangle &= \mathrm{Re}(x_0 z_0 + 2\bar{x}_1 z_1 + \cdots + 2\bar{x}_p z_p) \\ &= x_0 z_0 + (\bar{x}_1 z_1 + x_1 \bar{z}_1) + \cdots + (\bar{x}_p z_p + x_p \bar{z}_p). \end{aligned}$$

A trigonometric polynomial of degree p or less can be expressed as

$$\begin{aligned} F_x(e^{\mathrm{j}\omega}) &= \langle a(e^{\mathrm{j}\omega}), x \rangle \\ &= x_0 + 2\big(\mathrm{Re}(x_1) \cos\omega + \mathrm{Im}(x_1) \sin\omega + \cdots + \mathrm{Re}(x_p) \cos p\omega + \mathrm{Im}(x_p) \sin p\omega \big). \end{aligned}$$

The cone of nonnegative trigonometric polynomials

$$K_{\mathrm{trig}} = \{ x \in \mathbf{R} \times \mathbf{C}^p \mid F_x(e^{\mathrm{j}\omega}) \geq 0 \ \ \forall \omega \in [0, 2\pi) \}$$

is a proper cone, and it appears naturally as constraints in signal processing applications [WBV96, SMM00, DTS01, DLS02, AV02, Hac03, Dum07, GL08, SDL10].

2.3.1 Spectral factorization

A useful classical result concerning nonnegative trigonometric polynomials is the spectral factorization theorem (or Riesz-Fejér theorem, see [AM79, §9] [Dum07, §1], for example). Denote with superscript a^H the conjugate transpose of a vector or matrix a. The theorem

states that $F_x(e^{\mathrm{j}\omega})$ is nonnegative if and only if it is the square of a causal polynomial $\mathcal{B}(e^{\mathrm{j}\omega}) = a(e^{\mathrm{j}\omega})^H b = \sum_{k=0}^{p} b_k e^{-\mathrm{j}k\omega}$ for some $b \in \mathbf{R} \times \mathbf{C}^p$, *i.e.*, a sum-of-squares with one term,

$$x \in K_{\mathrm{trig}} \iff F_x(e^{\mathrm{j}\omega}) = |\mathcal{B}(e^{\mathrm{j}\omega})|^2 = a(e^{\mathrm{j}\omega})^H b b^H a(e^{\mathrm{j}\omega}). \tag{2.1}$$

Several efficient algorithms exist for computing a (minimum phase) spectral factor b given an $x \in K_{\mathrm{trig}}$ (see *e.g.* [Dum07, appendix B] and references therein).

2.3.2 LMI characterizations

The following equivalence enables solving problems involving the constraint $x \in K_{\mathrm{trig}}$ as SDPs when the objective function and the other constraints are SDP representable. It is the foundation of many signal processing applications of SDP [Dum07], where $x = \mathcal{D}(X)$ is called the Gram matrix parametrization of any trigonometric polynomial, with the linear mapping $\mathcal{D} : \mathbf{H}^{p+1} \to \mathbf{R} \times \mathbf{C}^p$ defined as the diagonal sums

$$\mathcal{D}(\begin{bmatrix} X_{00} & X_{01} & \cdots & X_{0p} \\ X_{10} & X_{11} & \cdots & X_{1p} \\ \vdots & \vdots & \ddots & \vdots \\ X_{p0} & X_{p1} & \cdots & X_{pp} \end{bmatrix}) = (\sum_{i=0}^{p} X_{ii}, \sum_{i=0}^{p-1} X_{i+1,i}, \ldots, X_{p-1,0} + X_{p1}, X_{p0}). \tag{2.2}$$

Theorem 2.4. *The following statements are equivalent.*

(a) $F_x(e^{\mathrm{j}\omega}) \geq 0$ *for all* $\omega \in [0, 2\pi)$.

(b) There exists an $X \in \mathbf{H}^{p+1}$ *such that* $x = \mathcal{D}(X)$ *and* $X \succeq 0$.

(c) There exists a $P \in \mathbf{H}^p$ *such that* $M + F^H P F - G^H P G \succeq 0$, *where* $M \in \mathbf{H}^{p+1}$ *with*

$$M = \begin{bmatrix} 0 & \cdots & 0 & \bar{x}_p \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 0 & \bar{x}_1 \\ x_p & \cdots & x_1 & x_0 \end{bmatrix}$$

and $F, G \in \mathbf{R}^{p \times (p+1)}$ *with*

$$F = \begin{bmatrix} 0 & I_p \end{bmatrix}, \qquad G = \begin{bmatrix} I_p & 0 \end{bmatrix}. \tag{2.3}$$

Proof. The proof proceeds as (c) $\Rightarrow$ (b) $\Rightarrow$ (a) and $\neg$(c) $\Rightarrow$ $\neg$(a). Showing (c) $\Rightarrow$ (b) is straightforward by taking

$$X = M + F^H P F - G^H P G \succeq 0$$

and noting that $x = \mathcal{D}(X)$. Showing (b) $\Rightarrow$ (a) is also straightforward. Note that if $x = \mathcal{D}(X)$ and $X \succeq 0$, then $F_x(e^{\mathrm{j}\omega}) = a(e^{\mathrm{j}\omega})^H X a(e^{\mathrm{j}\omega}) \geq 0$ for all ω.

To show $\neg$(c) $\Rightarrow$ $\neg$(a), we invoke Theorem 2.3 with $\mathcal{A}(P) = F^H P F - G^H P G$, $A_0 = M$, and note that constraint qualification is satisfied. This is easily seen by looking at the diagonal of $\mathcal{A}(P) \succeq 0$, which enforces the constraints $0 \geq P_{11} \geq P_{22} \geq \cdots \geq P_{pp} \geq 0$ on the diagonal elements of P and implies $\mathcal{A}(P) = 0$. The negation of (c) is therefore equivalent to the existence of a $Z \succeq 0$ with rank $r \geq 1$ such that $FZF^H = GZG^H$ and $\mathbf{tr}\, MZ < 0$. The equality on Z states that Z is Toeplitz. Hence, we use the classical result of Vandermonde decomposition, which states that every positive semidefinite $n \times n$ Toeplitz matrix of rank r can be decomposed as

$$Z = \sum_{k=1}^{r} d_k a(e^{\mathrm{j}\omega_k}) a(e^{\mathrm{j}\omega_k})^H$$

with distinct ω_k and positive weights d_k for $k = 1, \ldots, r$ [SM97, page 170]. Thus, the strict inequality means $\mathbf{tr}\, MZ = \sum_{k=1}^{r} d_k a(e^{\mathrm{j}\omega_k})^H M a(e^{\mathrm{j}\omega_k}) < 0$, which implies that

$$F_x(e^{\mathrm{j}\omega_k}) = a(e^{\mathrm{j}\omega_k})^H M a(e^{\mathrm{j}\omega_k}) < 0$$

for some k. □

The equivalence between (a) and (c) in Theorem 2.4 is a result of KYP lemma applied to the discrete-time finite-impulse response case. As in the proof of Theorem 2.4, it is shown in Section 2.7 that the nontrivial direction in the proof of the general non-strict KYP lemma requires the use of Theorem 2.3, and hence, a constraint qualification. Secondly, as the Vandermonde decomposition is needed to factorize a positive semidefinite Toeplitz matrix, we need the decomposition result for more generally structured positive semidefinite matrices, which is described precisely in Section 2.4.

Note that from the last expression in (2.1), we see the relation $x = \mathcal{D}(bb^H)$ and conclude

that $x \in K_{\text{trig}} \setminus \{0\}$ if and only if there exists a $X \in \mathbf{H}^{p+1}$ with rank 1 such that $x = \mathcal{D}(X)$ and $X \succeq 0$. This also allows a stronger statement than (b) in Theorem 2.4 to be made.

The LMI characterization provides more flexibility in the problems we can solve by incorporating other SDP constraints, for example, problems with noisy or missing data.

Example: MA estimation. Suppose we are modeling a time series as an order p moving average process, MA(p), parametrized by $b \in \mathbf{R} \times \mathbf{C}^p$, given an estimated autocorrelation sequence of length $p+1$, $\hat{r} = (\hat{r}_0, \hat{r}_1, \ldots, \hat{r}_p) \in \mathbf{R} \times \mathbf{C}^p$. It is possible that the estimated sequence $\hat{r}$ is not a valid autocorrelation sequence, *i.e.*, the estimated power spectral density $F_{\hat{r}}(e^{\mathrm{j}\omega})$ may be negative at some frequency ω. Therefore, it is desirable to find a valid autocorrelation sequence $r \in K_{\text{trig}}$ that is 'close to' the estimated sequence $\hat{r}$, for example, by solving the optimization problem

$$\begin{array}{ll} \text{minimize} & \|r - \hat{r}\| \\ \text{subject to} & r \in K_{\text{trig}}. \end{array}$$

After obtaining the optimal r, spectral factorization algorithms can be applied to obtain b from $\mathcal{D}(bb^H) = r$.

2.3.3 Duality: positive semidefinite Toeplitz matrices

Next we will see that the decomposition result is also important in the characterization of the dual cone. The dual cone also appears in interesting applications, albeit not as prominent, but they are certainly useful from an optimization point of view. From Theorem 2.4, we obtain equivalent expressions for K_{trig}. In particular, the expression

$$K_{\text{trig}} = \{\mathcal{D}(X) \mid X \succeq 0\}$$

shows that K_{trig} is the image of the positive semidefinite cone under linear mapping. We can then easily obtain an expression for the dual cone as

$$\begin{aligned} K_{\text{trig}}^* = K_{\text{Toep}} &= \{z \in \mathbf{R} \times \mathbf{C}^p \mid \langle x, z\rangle \geq 0 \ \ \forall x \in K_{\text{trig}}\} \\ &= \{z \in \mathbf{R} \times \mathbf{C}^p \mid \langle \mathcal{D}(X), z\rangle = \mathbf{tr}(X\mathcal{T}(z)) \geq 0 \ \ \forall X \succeq 0\} \\ &= \{z \in \mathbf{R} \times \mathbf{C}^p \mid \mathcal{T}(z) \succeq 0\}, \end{aligned}$$

where the mapping $\mathcal{T} : \mathbf{R} \times \mathbf{C}^p \to \mathbf{H}^{p+1}$ denotes the adjoint of $\mathcal{D}$ and maps a vector to a Hermitian Toeplitz matrix

$$\mathcal{T}(z) = \begin{bmatrix} z_0 & \bar{z}_1 & \cdots & \bar{z}_p \\ z_1 & z_0 & \cdots & \bar{z}_{p-1} \\ \vdots & \vdots & \ddots & \vdots \\ z_p & z_{p-1} & \cdots & z_0 \end{bmatrix}. \tag{2.4}$$

Since K_{trig} is a proper cone, Lemmas 2.1 and 2.2 tell us that K_{Toep} is also proper and that the cones K_{trig} and K_{Toep} are dual to each other.

As a side note, using the Vandermonde decomposition in the proof of Theorem 2.4, we see that the dual cone admits the expression

$$K_{\text{Toep}} = \{z \in \mathbf{R} \times \mathbf{C}^p \mid z = \sum_{k=1}^{r} d_k a(e^{\mathrm{j}\omega_k}), \quad d_k \geq 0, \quad k = 1, \ldots, r\},$$

which is the (truncated) trigonometric moment cone.

Example: line spectrum estimation. Parametric line spectrum estimation is concerned with fitting signal models of the form

$$y(t) = \sum_{k=1}^{r} c_k e^{\mathrm{j}\omega_k t} + v(t), \tag{2.5}$$

where $v(t)$ is noise. If the phase angles of c_k are independent random variables, uniformly distributed on $[-\pi, \pi]$, and $v(t)$ is circular white noise with $\mathbf{E}\,|v(t)|^2 = \sigma^2$, then the covariance matrix of $y(t)$ of order $p+1$ is given by

$$\begin{bmatrix} r_0 & r_{-1} & \cdots & r_{-p} \\ r_1 & r_0 & \cdots & r_{-p+1} \\ \vdots & \vdots & \ddots & \vdots \\ r_p & r_{p-1} & \cdots & r_0 \end{bmatrix} = \sigma^2 I + \sum_{k=1}^{r} |c_k|^2 \begin{bmatrix} 1 \\ e^{\mathrm{j}\omega_k} \\ \vdots \\ e^{\mathrm{j}p\omega_k} \end{bmatrix} \begin{bmatrix} 1 \\ e^{\mathrm{j}\omega_k} \\ \vdots \\ e^{\mathrm{j}p\omega_k} \end{bmatrix}^H, \tag{2.6}$$

where $r_k = \mathbf{E}\,(y(t)\overline{y(t-k)})$ [SM97, §4.1] [PM96, §12.5]. Many classical subspace methods are designed for this model (see Appendix A for a short review), and active researches are

still being conducted to adapt these methods to more general situations. An optimization approach is considered here instead. A given estimated covariance matrix $\hat{R}$ may not be PSD, and it is often desirable to find $z \in K_1^*$ such that $\sigma^2 I + \mathcal{T}(z)$ is 'close to' $\hat{R}$, for example, by solving the optimization problem

$$\begin{array}{ll} \text{minimize} & \gamma \|\sigma^2 I + \mathcal{T}(z) - \hat{R}\| + z_0 \\ \text{subject to} & z \in K_{\text{Toep}}, \end{array}$$

where γ is a positive regularization parameter. An estimate of the model parameters $|c_k|^2$, ω_k for $k = 1, \ldots, r$ can therefore be obtained via Vandermonde decomposition of $\mathcal{T}(z)$.

2.4 Structured positive semidefinite matrix factorization

The section is adapted from [CV17]. In this thesis, the extension is made by observing that the set of complex exponentials

$$\{a(e^{\mathrm{j}\omega}) = (1, e^{\mathrm{j}\omega}, \ldots, e^{\mathrm{j}p\omega}) \mid \omega \in [0, 2\pi)\}$$

can be parameterized as

$$\{a(\lambda) \mid (\lambda G - F)a = 0,\ \lambda \in \mathcal{C}',\ a_0 = 1\}$$

where $\mathcal{C}'$ is the unit circle in the complex plane, and F and G are the $p \times (p+1)$ matrices defined in (2.3). Note that the parameter λ for the vector a is dropped at times to simplify notation. In the rest of this chapter, we remove the restriction $a_0 = 1$ and generalize the set in two ways. The first generalization is to allow F and G to be arbitrary matrices of equal size, *i.e.*, to replace $\lambda G - F$ with an arbitrary matrix pencil (a matrix polynomial of degree one). Second, we allow $\mathcal{C}'$ to be an arbitrary circle or line in the complex plane, or a segment of a line or a circle. Specific examples of these extensions, with different choices of F, G, and $\mathcal{C}'$, are discussed in Section 2.5.

Throughout the rest of the thesis we assume that F and G are complex matrices of size $p \times n$, and Φ and Ψ are Hermitian 2×2 matrices with $\det \Phi < 0$. We define

$$\mathcal{A} = \{a \in \mathbf{C}^n \mid (\mu G - \nu F)a = 0,\ (\mu, \nu) \in \mathcal{C}\}, \tag{2.7}$$

where

$$\mathcal{C} = \left\{(\mu,\nu) \in \mathbf{C}^2 \mid (\mu,\nu) \neq 0,\ q_\Phi(\mu,\nu) = 0,\ q_\Psi(\mu,\nu) \leq 0\right\}. \tag{2.8}$$

Here q_Φ, q_Ψ are the quadratic forms defined by Φ and Ψ:

$$q_\Phi(\mu,\nu) = \begin{bmatrix} \mu \\ \nu \end{bmatrix}^H \Phi \begin{bmatrix} \mu \\ \nu \end{bmatrix}, \qquad q_\Psi(\mu,\nu) = \begin{bmatrix} \mu \\ \nu \end{bmatrix}^H \Psi \begin{bmatrix} \mu \\ \nu \end{bmatrix}. \tag{2.9}$$

The set $\mathcal{C}$ is a subset of a line or circle in the complex plane, expressed in homogeneous coordinates, as explained in appendix B. Three important special cases of Φ are

$$\Phi_{\mathrm{u}} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \qquad \Phi_{\mathrm{i}} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \qquad \Phi_{\mathrm{r}} = \begin{bmatrix} 0 & \mathrm{j} \\ -\mathrm{j} & 0 \end{bmatrix},$$

for the unit circle, imaginary axis, and real axis, respectively. If $\Phi_{11} \neq 0$ or $\Psi_{11} > 0$, then $\nu \neq 0$ for all elements $(\mu,\nu) \in \mathcal{C}$, and we can simplify the definition of $\mathcal{A}$ as

$$\mathcal{A} = \{a \in \mathbf{C}^n \mid (\lambda G - F)a = 0,\ (\lambda,1) \in \mathcal{C}\}. \tag{2.10}$$

If $\Phi_{11} = 0$ and $\Psi_{11} \leq 0$, then the pair $(1,0)$ is also in $\mathcal{C}$ and the set $\mathcal{A}$ in (2.7) is the union of the right-hand side of (2.10) and the nullspace of G, *i.e.*,

$$\mathcal{A} = \{a \in \mathbf{C}^n \mid (\lambda G - F)a = 0,\ (\lambda,1) \in \mathcal{C}\} \cup \{a \in \mathbf{C}^n \mid Ga = 0\}.$$

Examples of sets $\mathcal{A}$ are given in Section 2.5. When referring to a specific element in $\mathcal{A}$ corresponding to a parameter $\lambda \in \mathbf{C} \cup \{\infty\}$, we use the notation $a(\lambda)$ to indicate that

$$(\lambda G - F)a(\lambda) = 0$$

with a finite λ satisfying $(\lambda,1) \in \mathcal{C}$, or

$$Ga(\lambda) = 0$$

with λ being a point at infinity and $(1,0) \in \mathcal{C}$.

The purpose of this section is to discuss a class of structured positive semidefinite matrices, *i.e.*, the convex hull of the set of matrices aa^H with $a \in \mathcal{A}$,

$$\operatorname{conv}\{aa^H \mid a \in \mathcal{A}\} = \{\sum_{k=1}^{r} a_k a_k^H \mid a_k \in \mathcal{A},\ k = 1,\ldots,r\}. \tag{2.11}$$

2.4.1 Main result: conic decomposition

The key result (Theorem 2.5) is known under various forms in system theory, signal processing, and moment theory [KS66, KN77, GS84]. Our purpose is to give a simple semidefinite formulation that encompasses a wide variety of interesting special cases, and to present a constructive proof that can be implemented using the basic decompositions of numerical linear algebra (specifically, symmetric eigenvalue, singular value, and Schur decompositions).

Lemma 2.3 provides the matrix factorization result needed in the proof of Theorem 2.5.

Lemma 2.3. *Let* Φ, $\Psi \in \mathbf{H}^2$ *with* $\det \Phi < 0$. *If* $U, V \in \mathbf{C}^{p\times r}$ *satisfy*

$$\Phi_{11}UU^H + \Phi_{21}UV^H + \Phi_{12}VU^H + \Phi_{22}VV^H = 0, \tag{2.12}$$

$$\Psi_{11}UU^H + \Psi_{21}UV^H + \Psi_{12}VU^H + \Psi_{22}VV^H \preceq 0, \tag{2.13}$$

then there exist a $W \in \mathbf{C}^{p\times r}$, *a unitary* $Q \in \mathbf{C}^{r\times r}$, *and vectors* $\mu, \nu \in \mathbf{C}^r$ *such that*

$$U = W\,\mathbf{diag}(\mu)Q^H, \qquad V = W\,\mathbf{diag}(\nu)Q^H,$$

and $q_\Phi(\mu_i, \nu_i) = 0$, $q_\Psi(\mu_i, \nu_i) \leq 0$, $(\mu_i, \nu_i) \neq 0$ *for* $i = 1, \ldots, r$.

Proof. See Appendix C. □

Theorem 2.5. *Let* $\mathcal{A}$ *be defined by (2.7) and (2.8), where* F, $G \in \mathbf{C}^{p\times n}$ *and* Φ, $\Psi \in \mathbf{H}^2$ *with* $\det \Phi < 0$. *A matrix* $X \in \mathbf{H}^n$ *is positive semidefinite of rank* $r \geq 1$ *and satisfies*

$$\Phi_{11}FXF^H + \Phi_{21}FXG^H + \Phi_{12}GXF^H + \Phi_{22}GXG^H = 0 \tag{2.14}$$

$$\Psi_{11}FXF^H + \Psi_{21}FXG^H + \Psi_{12}GXF^H + \Psi_{22}GXG^H \preceq 0, \tag{2.15}$$

if and only if X *can be decomposed as* $X = \sum_{k=1}^r a_k a_k^H$, *with linearly independent vectors* a_1, $\ldots$, $a_r \in \mathcal{A}$.

Proof. Sufficiency is readily proved by substituting $X = \sum_{k=1}^r a_k a_k^H$ in (2.14) and (2.15),

and verifying that if $(\mu_k G - \nu_k F)a_k = 0$ with $(\mu_k, \nu_k) \neq 0$, then

$$\begin{aligned}\Phi_{11}FXF^H + \Phi_{21}FXG^H + \Phi_{12}GXF^H + \Phi_{22}GXG^H &= \sum_{k=1}^{r} \alpha_k q_\Phi(\mu_k, \nu_k) y_k y_k^H \\ \Psi_{11}FXF^H + \Psi_{21}FXG^H + \Psi_{12}GXF^H + \Psi_{22}GXG^H &= \sum_{k=1}^{r} \alpha_k q_\Psi(\mu_k, \nu_k) y_k y_k^H\end{aligned}$$

where $\alpha_k = 1/|\nu_k|^2$, $y_k = Ga_k$ if $\nu_k \neq 0$, and $\alpha_k = 1/|\mu_k|^2$, $y_k = Fa_k$ if $\nu_k = 0$.

To show necessity, we start from any factorization $X = YY^H$ where $Y \in \mathbf{C}^{n\times r}$ has rank r. It follows from Lemma 2.3 , applied to $U = FY$ and $V = GY$, that there exist a matrix $W \in \mathbf{C}^{p\times r}$, a unitary matrix $Q \in \mathbf{C}^{r\times r}$, and two vectors $\mu, \nu \in \mathbf{C}^r$ such that

$$FYQ = W\,\mathbf{diag}(\mu), \qquad GYQ = W\,\mathbf{diag}(\nu), \qquad (\mu_i, \nu_i) \in \mathcal{C}, \quad i = 1, \ldots, r.$$

Choosing a_k equal to the kth column of YQ gives the decomposition of X. □

Viewed geometrically, the theorem says that (2.11) is the set of positive semidefinite matrices X that satisfy (2.14) and (2.15).

It is useful to note that the proof of Lemma 2.3 in the appendix is constructive and gives a simple algorithm, based on singular value and Schur decompositions, for computing the matrices W, Q and the vectors μ, ν.

2.5 Examples

In this section we illustrate the decomposition in Theorem 2.5 with different choices of the matrices F, G, Φ, and Ψ. The section is adapted from [CV17].

2.5.1 Trigonometric polynomials

Complex exponentials As a first example, we take $p = n - 1$,

$$F = \begin{bmatrix} 0 & I_{n-1} \end{bmatrix}, \quad G = \begin{bmatrix} I_{n-1} & 0 \end{bmatrix}, \quad \Phi = \Phi_{\mathrm{u}} \triangleq \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \quad \Psi = 0. \tag{2.16}$$

A nonzero pair (μ, ν) satisfies $q_\Phi(\mu, \nu) = |\mu|^2 - |\nu|^2 = 0$ only if μ and ν are nonzero and $\lambda = \mu/\nu$ is on the unit circle. The condition $(\lambda G - F)a = 0$ in the definition of $\mathcal{A}$ gives

a recursion $\lambda a_1 = a_2$, $\lambda a_2 = a_3$, $\ldots$, $\lambda a_{n-1} = a_n$. Defining $\exp(\mathrm{j}\omega) = \lambda$, we find that $\mathcal{A}$ contains the vectors

$$a = c\,(1, e^{\mathrm{j}\omega}, e^{\mathrm{j}2\omega}, \ldots, e^{\mathrm{j}(n-1)\omega}), \tag{2.17}$$

for all $\omega \in [0, 2\pi)$ and $c \in \mathbf{C}$. The matrix constraints (2.14)–(2.15) reduce to $FXF^H = GXG^H$, *i.e.*, X is a Toeplitz matrix. Theorem 2.5 therefore reduces to the well known Vandermonde decomposition of every positive semidefinite Toeplitz matrix

$$X = \sum_{k=1}^{r} |c_k|^2 \begin{bmatrix} 1 \\ e^{\mathrm{j}\omega_k} \\ \vdots \\ e^{\mathrm{j}(n-1)\omega_k} \end{bmatrix} \begin{bmatrix} 1 \\ e^{\mathrm{j}\omega_k} \\ \vdots \\ e^{\mathrm{j}(n-1)\omega_k} \end{bmatrix}^H, \tag{2.18}$$

with $c_k \neq 0$ and distinct $\omega_1, \ldots, \omega_r$ [SM97, page 170].

Restricted complex exponentials Define F, G, Φ as in (2.16), and

$$\Psi = \begin{bmatrix} 0 & -e^{\mathrm{j}\alpha} \\ -e^{-\mathrm{j}\alpha} & 2\cos\beta \end{bmatrix}$$

with $\alpha \in [0, 2\pi)$ and $\beta \in [0, \pi)$. The elements $a \in \mathcal{A}$ have the same general form (2.17), with the added constraint that $\cos\beta \leq \cos(\omega - \alpha)$. Since we can restrict ω to the interval $[\alpha - \pi, \alpha + \pi]$, this is equivalent to $|\omega - \alpha| \leq \beta$. The constraints (2.14)–(2.15) specify that X is Toeplitz and satisfies the matrix inequality

$$-e^{-\mathrm{j}\alpha} FXG^H - e^{\mathrm{j}\alpha} GXF^H + 2(\cos\beta) GXG^H \preceq 0. \tag{2.19}$$

The theorem states that a positive semidefinite Toeplitz matrix of rank r satisfies (2.19) if and only if it can be decomposed as (2.18) with nonzero c_k and $|\omega_k - \alpha| \leq \beta$.

Real trigonometric functions Next consider $p = n - 1$,

$$G = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 2 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 2 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 2 & 0 \end{bmatrix}, \qquad F = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 & 0 & 0 \\ 1 & 0 & 1 & \cdots & 0 & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 0 & 1 \end{bmatrix},$$

and

$$\Phi = \Phi_{\mathrm{r}} \triangleq \begin{bmatrix} 0 & \mathrm{j} \\ -\mathrm{j} & 0 \end{bmatrix}, \qquad \Psi = \Phi_{\mathrm{u}} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.$$

A nonzero pair (μ, ν) satisfies $q_\Phi(\mu, \nu) = \mathrm{j}(\bar{\mu}\nu - \mu\bar{\nu}) = 0$ and $q_\Psi(\mu, \nu) = |\mu|^2 - |\nu|^2 \leq 0$ only if $\nu \neq 0$ and $\lambda = \mu/\nu$ is real with $|\lambda| \leq 1$. The condition $(\lambda G - F)a = 0$ gives a recursion $\lambda a_1 = a_2$, $2\lambda a_k = a_{k-1} + a_{k+1}$ for $k = 2, \ldots, n-1$. If we write $\lambda = \cos\omega$, we recognize the recursion $2\cos\omega\cos k\omega = \cos(k-1)\omega + \cos(k+1)\omega$ and find that $\mathcal{A}$ contains the vectors

$$a = c\,(1,\, \cos\omega,\, \cos 2\omega,\, \ldots,\, \cos(n-1)\omega),$$

for all $\omega \in [0, 2\pi)$ and all c. With the same F and $G = [\ 2I_{n-1} \quad 0\ ]$, the condition $(\lambda G - F)a = 0$ reduces to $2\lambda a_1 = a_2$, $2\lambda a_k = a_{k-1} + a_{k+1}$ for $k = 2, \ldots, n-1$. If we write $\lambda = \cos\omega$, the solutions are the vectors

$$a = c\,(1, \frac{\sin 2\omega}{\sin\omega}, \frac{\sin 3\omega}{\sin\omega}, \ldots, \frac{\sin n\omega}{\sin\omega}),$$

for all $\omega \in [0, 2\pi)$ and all c.

Trigonometric vector polynomials We take $p = (k-1)l$, $n = kl$, and replace F and G in (2.16) with

$$F = \begin{bmatrix} 0 & I & 0 & \cdots & 0 \\ 0 & 0 & I & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & I \end{bmatrix}, \qquad G = \begin{bmatrix} I & 0 & \cdots & 0 & 0 \\ 0 & I & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & I & 0 \end{bmatrix},$$

and blocks of size $l \times l$. Then $\mathcal{A}$ contains the vectors of the form

$$a = (1,\, e^{\mathrm{j}\omega},\, e^{\mathrm{j}2\omega},\, \ldots,\, e^{\mathrm{j}(k-1)\omega}) \otimes c,$$

for all $c \in \mathbf{C}^l$ and $\omega \in [0, 2\pi)$, where $\otimes$ denotes Kronecker product.

2.5.2 Polynomials

See [KS66, KN77], for example, for a classical treatment of sequence of polynomials and moments.

Real powers Define F, G as in (2.16), and $\Phi = \Phi_{\mathrm{r}}$, $\Psi = 0$. A pair (μ, ν) satisfies $q_\Phi(\mu, \nu) = 0$ if and only if $\bar{\mu}\nu$ is real. If $(\mu, \nu) \neq 0$, we either have $\nu = 0$ and μ arbitrary, or $\nu \neq 0$ and $\lambda = \mu/\nu$ real. The set $\mathcal{A}$ therefore contains the vectors

$$a = c\,(1, \lambda, \lambda^2, \ldots, \lambda^{n-1}), \qquad a = c\,(0, 0, \ldots, 0, 1)$$

for all $\lambda \in \mathbf{R}$ and c. The matrix constraints (2.14)–(2.15) reduce to $FXG^H = GXF^H$, *i.e.*, X is a symmetric (real) Hankel matrix. Hence, a real symmetric positive semidefinite Hankel matrix of rank r can be decomposed in one of two forms

$$X = \sum_{k=1}^{r} |c_k|^2 \begin{bmatrix} 1 \\ \lambda_k \\ \vdots \\ \lambda_k^{n-2} \\ \lambda_k^{n-1} \end{bmatrix} \begin{bmatrix} 1 \\ \lambda_k \\ \vdots \\ \lambda_k^{n-2} \\ \lambda_k^{n-1} \end{bmatrix}^T, \quad X = \sum_{k=1}^{r-1} |c_k|^2 \begin{bmatrix} 1 \\ \lambda_k \\ \vdots \\ \lambda_k^{n-2} \\ \lambda_k^{n-1} \end{bmatrix} \begin{bmatrix} 1 \\ \lambda_k \\ \vdots \\ \lambda_k^{n-2} \\ \lambda_k^{n-1} \end{bmatrix}^T + |c_r|^2 \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}^T,$$

with distinct real λ_k and nonzero c_k.

Restricted polynomials If F, G are defined as in (2.16) and $\Phi = \Phi_{\mathrm{r}}$,

$$\Psi = \begin{bmatrix} 2 & -(\alpha+\beta) \\ -(\alpha+\beta) & 2\alpha\beta \end{bmatrix}$$

where $-\infty < \alpha < \beta < \infty$, then $\mathcal{A}$ contains all vectors $a = c(1, \lambda, \ldots, \lambda^{n-1})$ with $\lambda \in [\alpha, \beta]$. The matrix constraints require X to be a real symmetric Hankel matrix that satisfies

$$2FXF^H - (\alpha+\beta)(FXG^H + GXF^H) + 2\alpha\beta GXG^H \preceq 0.$$

Orthogonal polynomials Let $p_0(\lambda)$, $p_1(\lambda)$, $p_2(\lambda)$, ... be a sequence of real polynomials on $\mathbf{R}$, with p_i of degree i. It is well known that the polynomials are orthonormal with respect to an inner product that satisfies the property

$$\langle f(\lambda), \lambda g(\lambda)\rangle = \langle \lambda f(\lambda), g(\lambda)\rangle \tag{2.20}$$

(for example, an inner product of the form $\langle f, g\rangle = \int f(\lambda)g(\lambda)w(\lambda)d\lambda$ with $w(\lambda) \geq 0$) if and only if the polynomials satisfy a three-term recurrence

$$\beta_{i+1}p_{i+1}(\lambda) = (\lambda - \alpha_i)p_i(\lambda) - \beta_i p_{i-1}(\lambda), \tag{2.21}$$

with $p_{-1}(\lambda) = 0$ and $p_0(\lambda) = 1/d_0$, where $d_0^2 = \langle 1, 1\rangle$. This can be seen as follows [GK83].

Suppose $p_0, \ldots, p_{n-1}$ is any set of polynomials, with p_i of degree i. Then $\lambda p_i(\lambda)$ can be expressed as a linear combination of the polynomials $p_0(\lambda), \ldots, p_{i+1}(\lambda)$, and therefore

$$\lambda \begin{bmatrix} p_0(\lambda) \\ p_1(\lambda) \\ \vdots \\ p_{n-2}(\lambda) \end{bmatrix} = \begin{bmatrix} J & \beta_{n-1}e_{n-1} \end{bmatrix} \begin{bmatrix} p_0(\lambda) \\ p_1(\lambda) \\ \vdots \\ p_{n-1}(\lambda) \end{bmatrix} \tag{2.22}$$

for some lower-Hessenberg matrix J (*i.e.*, satisfying $J_{ij} = 0$ for $j > i+1$). Let $\langle \cdot, \cdot \rangle$ be an inner product on the space of polynomials of degree $n-1$ or less. Taking inner products on both sides of (2.22), we find that

$$H = JG + \beta_{n-1}e_{n-1}g^T$$

where

$$H_{ij} = \langle \lambda p_{i-1}(\lambda), p_{j-1}(\lambda)\rangle, \qquad G_{ij} = \langle p_{i-1}(\lambda), p_{j-1}(\lambda)\rangle, \qquad g_j = \langle p_{n-1}(\lambda), p_{j-1}(\lambda)\rangle,$$

for $i, j = 1, \ldots, n-1$. The polynomials are orthonormal for the inner product if and only if $G = I$ and $g = 0$. The inner product satisfies the property (2.20) if and only if H is symmetric. Hence if the polynomials are orthonormal for an inner product that satisfies (2.20), then J is a symmetric tridiagonal matrix. If we use the notation

$$J = \begin{bmatrix} \alpha_0 & \beta_1 & 0 & \cdots & 0 & 0 \\ \beta_1 & \alpha_1 & \beta_2 & \cdots & 0 & 0 \\ 0 & \beta_2 & \alpha_2 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \alpha_{n-3} & \beta_{n-2} \\ 0 & 0 & 0 & \cdots & \beta_{n-2} & \alpha_{n-2} \end{bmatrix}, \tag{2.23}$$

the recurrence (2.21) follows. Conversely, if the three-term recurrence holds, and we define the inner product by setting $G = I$, $g = 0$, then H is symmetric and the inner product satisfies (2.20).

Now consider (2.7) and (2.8), with $p = n - 1$ and

$$\Phi = \Phi_{\mathrm{r}}, \qquad \Psi = 0, \qquad G = \begin{bmatrix} I_{n-1} & 0 \end{bmatrix}, \qquad F = \begin{bmatrix} J & \beta_{n-1} e_{n-1} \end{bmatrix},$$

where J is the Jacobi matrix (2.23) of a system of orthogonal polynomials. Then $(\mu, \nu) \in \mathcal{C}$ if and only if either $\nu \neq 0$ and $\lambda = \mu/\nu \in \mathbf{R}$, or $\nu = 0$. The set contains the vectors a of the following form for all $\lambda \in \mathbf{R}$:

$$a = c\,(p_0(\lambda),\, p_1(\lambda),\, p_2(\lambda), \ldots,\, p_{n-1}(\lambda)), \qquad a = c\,(0, 0, \ldots, 0, 1).$$

2.5.3 Rational functions

As a final example, we consider the controllability pencil of a linear system:

$$G = \begin{bmatrix} I & 0 \end{bmatrix}, \qquad F = \begin{bmatrix} A & B \end{bmatrix}, \tag{2.24}$$

where $A \in \mathbf{C}^{n_{\mathrm{s}} \times n_{\mathrm{s}}}$ and $B \in \mathbf{C}^{n_{\mathrm{s}} \times m}$. With this choice, $\mathcal{A}$ contains the vectors $a = (x, u)$ that satisfy the equality $(\mu I - \nu A)x = \nu B u$ for some $(\mu, \nu) \in \mathcal{C}$. Since $(\mu, \nu) \neq 0$, we either have

$\nu = 0$ and $x = 0$, or $\nu \neq 0$ and $((\mu/\nu)I - A)x = Bu$. If A has no eigenvalues λ that satisfy $(\lambda, 1) \in \mathcal{C}$, then $\mathcal{A}$ contains the vectors

$$a = \begin{bmatrix} (\lambda I - A)^{-1}Bu \\ u \end{bmatrix}$$

for all $(\lambda, 1) \in \mathcal{C}$ and all $u \in \mathbf{C}^m$. If $\mathcal{C}$ includes the point $(1, 0)$ at infinity, then $\mathcal{A}$ also contains the vectors $(0, u)$ for all $u \in \mathbf{C}^m$.

This can be extended to the controllability pencil of a descriptor system

$$G = \begin{bmatrix} E & 0 \end{bmatrix}, \qquad F = \begin{bmatrix} A & B \end{bmatrix},$$

where $E \in \mathbf{C}^{n_s \times n_s}$ is possibly singular. With this choice, $\mathcal{A}$ contains the vectors $a = (x, u)$ that satisfy the equality $(\mu E - \nu A)x = \nu Bu$ for some $(\mu, \nu) \in \mathcal{C}$. If $\det(\mu E - \nu A) \neq 0$ for all $(\mu, \nu) \in \mathcal{C}$, then $\mathcal{A}$ contains all vectors

$$a = \begin{bmatrix} (\lambda E - A)^{-1}Bu \\ u \end{bmatrix}$$

for $(\lambda, 1) \in \mathcal{C}$ and $u \in \mathbf{C}^m$. If $(1, 0) \in \mathcal{C}$, then $\mathcal{A}$ also contains $(0, u)$ for all $u \in \mathbf{C}^m$.

2.6 Duality

Using Theorem 2.5, we obtain an LMI expression for the cone of structured positive semi-definite matrices defined in (2.11) as

$$\begin{aligned} K_{\text{stru}} &= \{\sum_{k=1}^{r} a_k a_k^H \mid a_k \in \mathcal{A},\ k = 1, \ldots, r\} \\ &= \{X \in \mathbf{H}^n \mid X \succeq 0,\ (2.14),\ (2.15)\}. \end{aligned}$$

The LMI expression shows that the cone K_{stru} is closed and convex. Therefore, according to Lemma 2.2, it forms a dual pair with its dual cone,

$$\begin{aligned} K_{\text{stru}}^* = K_{\text{Popov}} &= \{M \in \mathbf{H}^n \mid \sum_{k=1}^{r} a_k^H M a_k \geq 0 \ \ \forall a_k \in \mathcal{A},\ k = 1, \ldots, r\} \\ &= \{M \in \mathbf{H}^n \mid F_M(\lambda) = a(\lambda)^H M a(\lambda) \geq 0 \ \ \forall a(\lambda) \in \mathcal{A}\}. \end{aligned}$$

This is the cone of nonnegative Popov functions. Section 2.7 gives more discussion on Popov functions and the equivalent expressions for K_{Popov}.

2.7 Kalman–Yakubovich–Popov Lemma

Although it is not directly related to the main contributions of the thesis, this section presents general forms of the KYP lemma [Kal63, Yak62, Pop62, Ran96, IMF00, BV02, BV03, IH05, Sch06, PV11] consistent with the notation of this thesis and gives specific examples in their well-known forms. The theorems in this section state the same results as [IH05, theorem 2], except that a less restrictive constraint qualification is given for the nonstrict KYP lemma. We note that there are further variations of the KYP lemma that are not covered here, for example, the sampling formulation that results in a low-rank structure in the LMIs [LP04, RV06, LV07, RDV07], as well as the generalization to certain curves in the complex plane described by polynomial equality and inequality of order higher than quadratic [PIH14].

The trigonometric polynomials in Section 2.3 are extended to Popov functions (see [IOW99, HSK99], for example) defined as

$$F_M(\lambda) = a(\lambda)^H M a(\lambda) \tag{2.25}$$

where $M \in \mathbf{H}^n$ is called the central matrix, and $a(\lambda) \in \mathcal{A}$. Many properties in control and signal processing applications can be characterized as the nonnegativity of a Popov function over a set of parameters, *i.e.*,

$$F_M(\lambda) \geq 0 \qquad \forall \lambda \in \mathcal{C}' = \{\mu/\nu \in \mathbf{C} \cup \{\infty\} \mid (\mu, \nu) \in \mathcal{C}\}$$

where $\mathcal{C}$ represents a subset of a line of circle in the complex plane, as described in Section 2.4 and Appendix B. Here we assume the set $\mathcal{C}$ defined with Φ and Ψ in (2.8) is not empty and not a singleton, since otherwise the nonnegative Popov constraint would be trivial. We also assume the inequality $q_\Psi(\mu, \nu) \leq 0$ is not redundant, which means there exist points $\lambda = \mu/\nu$ with $q_\Phi(\mu, \nu) = 0$ and $q_\Psi(\mu, \nu) < 0$. The KYP lemma [IH05, theorem 2] establishes equivalence of the nonnegative Popov constraint to a matrix inequality that is linear in the central matrix M, thus allowing $M(x)$ to be any linear mapping of the decision variable x in an optimization problem.

2.7.1 Strict version

We first state the KYP lemma with strict inequality.

Theorem 2.6. *The following statements are equivalent.*

(a) $F_M(\lambda) > 0$ *for all* $\lambda \in \mathcal{C}'$.

(b) There exist $P, Q \in \mathbf{H}^p$ *with* $Q \succ 0$ *and*

$$M + \begin{bmatrix} F \\ G \end{bmatrix}^H (\Phi \otimes P + \Psi \otimes Q) \begin{bmatrix} F \\ G \end{bmatrix} \succ 0. \tag{2.26}$$

Proof. Suppose (b) holds. Consider any $\lambda = \mu/\nu \in \mathbf{C} \cup \{\infty\}$ with $(\mu, \nu) \in \mathcal{C}$ and $a \neq 0$ such that $(\mu G - \nu F)a = 0$. Define

$$w = \begin{cases} (1/\nu)Ga & \nu \neq 0 \\ (1/\mu)Fa & \nu = 0. \end{cases}$$

Therefore $Ga = \nu w$ and $Fa = \mu w$, and.

$$\begin{aligned} F_M(\lambda) \;=\; a^H M a \;&>\; -\begin{bmatrix} Fa \\ Ga \end{bmatrix}^H (\Phi \otimes P + \Psi \otimes Q) \begin{bmatrix} Fa \\ Ga \end{bmatrix} \\ &=\; -\begin{bmatrix} \mu w \\ \nu w \end{bmatrix}^H (\Phi \otimes P + \Psi \otimes Q) \begin{bmatrix} \mu w \\ \nu w \end{bmatrix} \\ &=\; -q_\Phi(\mu, \nu) w^H P w - q_\Psi(\mu, \nu) w^H Q w \\ &\geq\; 0. \end{aligned}$$

The first line follows from (2.26), and the last line from $q_\Phi(\mu, \nu) = 0$, $q_\Psi(\mu, \nu) \leq 0$, and $Q \succ 0$.

Conversely, suppose (b) is false, then by applying Theorem 2.1 with $\mathcal{A} : \mathbf{H}^p \times \mathbf{H}^p \to \mathbf{H}^n \times \mathbf{H}^p$ and

$$\mathcal{A}(\mathbf{diag}(P, Q)) = \mathbf{diag}(\begin{bmatrix} F \\ G \end{bmatrix}^H (\Phi \otimes P + \Psi \otimes Q) \begin{bmatrix} F \\ G \end{bmatrix}, Q), \quad A_0 = \mathbf{diag}(M, 0), \tag{2.27}$$

there exist $X \in \mathbf{H}^n$ and $Z \in \mathbf{H}^p$ such that

$$\mathbf{diag}(X, Z) \gneqq 0, \qquad \mathcal{A}^{\mathrm{adj}}(\mathbf{diag}(X, Z)) = 0, \qquad \mathbf{tr}(MX) \leq 0.$$

Equivalently, there exists $X \succeq 0$ with rank $r \geq 1$ satisfying (2.14) and (2.15). Theorem 2.5 and $\mathbf{tr}(MX) \leq 0$ then imply $a^H M a \leq 0$ for some $a \in \mathcal{A}$. □

2.7.2 Nonstrict version

A constraint qualification (CQ) is needed in the nonstrict KYP lemma. The proof of Theorem 2.7 uses strong alternatives of Theorem 2.3, which is ensured by an assumption we refer to as CQ. It states that the image of the linear mapping $\mathcal{A}$ does not intersect with the PSD cone except for the origin. Consider the linear mapping $\mathcal{A}$ defined in (2.27), CQ says that there exists no P and Q such that $\mathcal{A}(\mathbf{diag}(P, Q)) \gneqq 0$, or equivalently, via the application of Theorem 2.2, there exist X and Z such that $\mathbf{diag}(X, Z) \succ 0$, and $\mathcal{A}^{\mathrm{adj}}(\mathbf{diag}(X, Z)) = 0$. More explicitly, there exists $X \succ 0$ such that

$$\Phi_{11} F X F^H + \Phi_{21} F X G^H + \Phi_{12} G X F^H + \Phi_{22} G X G^H = 0$$

$$\Psi_{11} F X F^H + \Psi_{21} F X G^H + \Psi_{12} G X F^H + \Psi_{22} G X G^H \prec 0.$$

This CQ can also be characterized exactly in terms of the problem data F and G, which states that the following two conditions must hold:

1. The normal rank of $\lambda G - F$ is p.
2. The generalized eigenvalues of $\lambda G - F$ are nondefective and lie in the interior of the one-dimensional set $\mathcal{C}'$. More accurately, if λ is a finite generalized eigenvalue, then it satisfies $q_\Phi(\lambda, 1) = 0$ and $q_\Psi(\lambda, 1) < 0$. If it is an infinite generalized eigenvalue, then $q_\Phi(1, 0) = 0$ and $q_\Psi(1, 0) < 0$.

We refer interested readers to Appendix D for details and a proof. A sufficient and more easily verified condition is that $\mathbf{rank}(\mu G - \nu F) = p$, *i.e.*, full row rank, for all $(\mu, \nu) \neq 0$.

Theorem 2.7. *If constraint qualification holds, the following statements are equivalent.*

(a) $F_M(\lambda) \succeq 0$ *for all* $\lambda \in \mathcal{C}'$.

(b) There exist $P, Q \in \mathbf{H}^p$ *with* $Q \succeq 0$ *and*

$$M + \begin{bmatrix} F \\ G \end{bmatrix}^H (\Phi \otimes P + \Psi \otimes Q) \begin{bmatrix} F \\ G \end{bmatrix} \succeq 0.$$

Proof. The proof is similar to that of Theorem 2.6, except that we invoke Theorem 2.3 with constraint qualification. □

2.7.3 Linear time-invariant systems

This section presents the KYP lemma as it is more commonly seen in system and control theory. Consider again the controllability pencil (2.24)

$$G = \begin{bmatrix} I & 0 \end{bmatrix}, \qquad F = \begin{bmatrix} A & B \end{bmatrix},$$

where $A \in \mathbf{C}^{n_s \times n_s}$ and $B \in \mathbf{C}^{n_s \times m}$ describe a linear system. For the discrete-time systems, with $\Phi = \Phi_{\mathrm{u}}$, the set $\mathcal{C}'$ is (a subset of) the unit circle in the complex plane. For the continuous-time systems, with

$$\Phi = \Phi_{\mathrm{i}} \triangleq \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix},$$

the set $\mathcal{C}'$ is (a subset of) the imaginary axis in the complex plane. The CQ translates into system theoretic properties.

Lemma 2.4. *If all uncontrollable modes of the system described by the pair* (A, B) *are nondefective and corresponding to eigenvalues in the interior of the one-dimensional set* $\mathcal{C}'$*, the following statements are equivalent.*

(a) For all $\lambda \in \mathcal{C}'$ *and* $(x, u) \in \mathbf{C}^{n_s+m}$ *such that* $(\lambda I - A)x = Bu$ *(which means* $x = 0$ *for* λ *at infinity), it holds that*

$$\begin{bmatrix} x \\ u \end{bmatrix}^H M \begin{bmatrix} x \\ u \end{bmatrix} \geq 0.$$

(b) There exist $P, Q \in \mathbf{H}^p$ *with* $Q \succeq 0$ *and*

$$M + \begin{bmatrix} A & B \\ I & 0 \end{bmatrix}^H (\Phi \otimes P + \Psi \otimes Q) \begin{bmatrix} A & B \\ I & 0 \end{bmatrix} \succeq 0.$$

Note that if the eigenvalues of A do not lie in $\mathcal{C}'$, *i.e.*, $(\lambda I - A)^{-1}$ exists for all $\lambda \in \mathcal{C}'$, then statement (a) can be replaced with the following: for all $\lambda \in \mathcal{C}'$, it holds that

$$\begin{bmatrix} (\lambda I - A)^{-1} B \\ I \end{bmatrix}^H M \begin{bmatrix} (\lambda I - A)^{-1} B \\ I \end{bmatrix} \succeq 0.$$

CHAPTER 3

Gauge Penalties for Structured Symmetric Matrices

Few optimization problems have attracted as much interest in recent years as the problem of minimizing the sum of a convex function and an ℓ_1-norm regularization term. A general form of problems of this type is

$$\begin{array}{ll} \text{minimize} & f(\sum_{k=1}^{r} \theta_k a_k) + \sum_{k=1}^{r} |\theta_k| \\ \text{subject to} & a_k \in C, \quad k = 1, \ldots, r, \end{array} \tag{3.1}$$

where f is a convex function and C is a set (or *dictionary*) of vectors in $\mathbf{C}^n$ or $\mathbf{R}^n$. The unknowns in problem (3.1) are the real or complex coefficients θ_k, the vectors (or *atoms*) a_1, $\ldots$, a_r selected from C, and the number r of selected dictionary elements. If C is a finite set of vectors, it can be represented by a matrix A with the elements of C as its columns, and the problem can be written as

$$\text{minimize} \quad f(A\theta) + \|\theta\|_1. \tag{3.2}$$

This includes as special cases the LASSO problem [Tib96], basis pursuit [CDS98], noisy basis pursuit [Tro06, DE06], and numerous other applications [CRT06, Don06, Ela10, HTW15].

When reviewing the literature on ℓ_1-norm methods in signal processing [CRT06, CM73, SS86, SBT12], it is striking that many of the underlying applications involve signals in continuous domains (time, space, or frequency domain), and the ℓ_1-norm problems arise after discretizing and truncating an infinite dictionary. The discretization is used when no exact method for the continuous problem is known, or when the discretized problem is believed to be easier to solve numerically by convex optimization techniques.

It was recently noted that certain problems of the form (3.1) with infinite dictionaries can be exactly solved by semidefinite optimization, if the function f is also semidefinite

representable. In particular, the authors of [CF14, TBS13, BR11, CF13, YX15, Fer15] consider 1-norm minimization with dictionaries of vectors of undamped complex exponentials,

$$C = C_{\mathrm{e}} = \{\gamma\,(1, e^{\mathrm{j}\omega}, \ldots, e^{\mathrm{j}(n-1)\omega}) \in \mathbf{C}^n \mid \omega \in [0, 2\pi),\ |\gamma| = 1/\sqrt{n}\}$$

and use the fact that problem (3.1) is equivalent to the finite-dimensional convex optimization problem

$$\begin{array}{ll} \text{minimize} & f(x) + (\mathbf{tr}\, V + w)/2 \\ \text{subject to} & \left[\begin{array}{cc} V & x \\ x^H & w \end{array}\right] \succeq 0 \\ & V \text{ is Toeplitz.} \end{array} \tag{3.3}$$

The variables in this problem are $V \in \mathbf{H}^n$, $x \in \mathbf{C}^n$, $w \in \mathbf{R}$. (Here $\mathbf{H}^n$ denotes the set of Hermitian $n \times n$ matrices.) The formulations were proposed for super-resolution, gridless compressed sensing, and other applications in signal processing, and allow the continuous sparse optimization problems to be posed directly, as a finite-dimensional convex semidefinite optimization problem, without discretization.

This chapter and Chapter 4 present extensions of semidefinite programming formulations (3.3) of 1-norm optimization problems over infinite dictionaries of vectors of complex exponentials. In particular, the ℓ_1-norm penalty is extended to convex penalties promoting certain types of structure. We distinguish between two cases. Penalties for structured symmetric matrices are treated in this chapter. Specifically, Chapter 3 is concerned with the problem (3.1) where the set C contains structured positive semidefinite matrices. Based on the development of this chapter, Chapter 4 discusses penalties for structured nonsymmetric matrices, *i.e.*, when the set C contains structured nonsymmetric matrices, which include vectors as a special case. The results in the previous chapter provide simple, constructive proofs of the semidefinite representations of the penalty functions used in the aforementioned applications. The connection also leads the extensions to penalty functions for sets of vectors parameterized via the nullspace of matrix pencils. The techniques are illustrated with examples of low-rank matrix approximation problems arising in spectral estimation and array processing.

Outline

Section 3.1 gives a brief overview of sparse signal reconstruction via ℓ_1-norm and atomic norm minimizations. Section 3.2 outlines recent works on extensions to certain continuous domains. In Section 3.3, we define and present semidefinite representations of the atomic norms and gauge functions for structured symmetric matrices, based on the direct and constructive proofs of Theorem 2.5 in Section 2.4 and Appendix C. In Section 3.4, we derive the convex conjugates of the atomic norms and gauge functions, and discuss the relation between the dual SDP representations and the Kalman–Yakubovich–Popov lemma. In Section 3.5, the SDP techniques are illustrated with some applications in signal processing. Appendix D discusses the technical results on the properties of the matrix pencil $\lambda F - G$ that are needed to ensure strong duality in the dual problems. Most content of this chapter is adapted from [CV17].

3.1 Sparse signal reconstruction

Techniques for sparse signal reconstruction via ℓ_1-norm minimization have been a very active research topic over the past decades. The optimization of ℓ_1-norm has been widely studied in statistics, signal processing, and machine learning, and forms the basis of the celebrated LASSO method for regressor selection, compressed sensing, basis pursuit, and many other techniques. From a computational viewpoint, ℓ_1-norm minimization is attractive because it leads to tractable optimization problems that can be solved as linear programs (LPs) or second-order cone programs (SOCPs). In addition, an extensive theory has been developed that explains when and why ℓ_1-norm methods are successful.

One of the best known examples is the basis pursuit problem, in which we seek the minimum ℓ_1-norm solution to an underdetermined linear equation [CDS98]:

$$\begin{array}{ll} \text{minimize} & \|\theta\|_1 \\ \text{subject to} & A\theta = y. \end{array} \tag{3.4}$$

This is problem (3.2) with $f(x)$ being the indicator function of the singleton $\{y\}$, and it

can be written as an LP, for which many efficient algorithms exist. The matrix $A \in \mathbf{C}^{n\times m}$ in (3.4) may consist of an over-complete (*i.e.*, $n < m$) dictionary of basis signals from which an element is 'selected' if the corresponding element in θ is nonzero. The goal of sparse signal recovery is to select a small number of basis signals that are sufficient to represent the data vector y. One possible formulation is to compute the *sparsest* vector of coefficients θ by minimizing the cardinality (number of nonzero elements) of θ:

$$\begin{array}{ll} \text{minimize} & \|\theta\|_0 \\ \text{subject to} & A\theta = y. \end{array} \tag{3.5}$$

The problem (3.5) is nonconvex and combinatorial in nature, and it is in general NP-hard. The ℓ_1 minimization (3.4), on the other hand, is a tractable convex problem and often has very sparse solutions θ. It can therefore be interpreted as a convex heuristic for (3.5). Moreover, problems (3.5) and (3.4) can be shown to be *equivalent* when the problem data A, y satisfy certain properties; see [FN03, CRT06].

Many variations of the basis pursuit problem (3.4) exist that account for noisy or missing data, or add convex constraints on the coefficients θ. For example, in the robust recovery problem we allow noise in the data vector y and replace the equality constraint with a norm inequality, *i.e.*, choosing $f(x)$ in (3.2) as the indicator function of $\{x \mid \|x - y\| \leq \delta\}$,

$$\begin{array}{ll} \text{minimize} & \|\theta\|_1 \\ \text{subject to} & \|A\theta - y\| \leq \delta, \end{array} \tag{3.6}$$

or penalty in the objective,

$$\text{minimize} \quad \|\theta\|_1 + \frac{\gamma}{2}\|A\theta - y\|^2, \tag{3.7}$$

which is $f(x) = (\gamma/2)\|x - y\|^2$ in (3.2). The scalars δ and γ are nonnegative parameters. The latter formulation (3.7) is also well known as the LASSO [Tib96], and can be viewed as either robust basis pursuit or ℓ_1-norm regularized least-squares regression.

Building on the foundations and success of basis pursuit, methods for exact signal recovery with incomplete data solve a problem of the same form (3.4), where A is a randomly chosen measurement matrix, and each equality constraint represents a linear measurement on θ.

Exact recovery with $n < m$ is guaranteed under certain conditions [CRT06], and these conditions can be shown to hold with high probability for certain distributions of A. This gives rise to the famous compressed sensing (or compressive sampling) theory and finds many interesting applications [Don06, CW08, Bar07, Ela10, EK12].

More recent work has focused on extensions of these results from sparse optimization to other types of 'low-dimensional structure'. An example is the minimization of the trace norm (or nuclear norm, *i.e.*, the sum of singular values) of a matrix,

$$\begin{array}{ll} \text{minimize} & \|X\|_* \\ \text{subject to} & \mathcal{A}(X) = y, \end{array} \tag{3.8}$$

as a convex heuristic to the rank minimization problem

$$\begin{array}{ll} \text{minimize} & \mathbf{rank}\, X \\ \text{subject to} & \mathcal{A}(X) = y. \end{array} \tag{3.9}$$

Here $X \in \mathbf{C}^{n\times m}$ is a matrix variable, $y \in \mathbf{C}^l$ is given, and $\mathcal{A}$ denotes a linear map. The problem (3.8) can be formulated as a semidefinite program (SDP). The rank minimization problems is in general NP-hard, but, as for ℓ_1-norm minimization, there are proven conditions under which the two problems are equivalent [Faz02, RFP10].

The notion of atomic norm introduced in [CRP12] gives a unified description of convex penalty functions that extend the ℓ_1-norm penalty, used to promote sparsity in the solution of an optimization problem, to various other types of structure. The atomic norm associated with a non-empty (finite or infinite) set C is defined as the gauge of its convex hull, *i.e.*, the convex function

$$\begin{aligned} g(x) &= \inf\{t \geq 0 \mid x \in t\,\mathrm{conv}\, C\} \\ &= \inf\{\sum_{k=1}^{r} \theta_k \mid x = \sum_{k=1}^{r} \theta_k a_k,\ \theta_k \geq 0,\ a_k \in C\}. \end{aligned} \tag{3.10}$$

This function is convex, nonnegative, and positively homogeneous. It is not necessarily a norm, but it is common to use the term 'atomic norm' even when g is not a norm. When used as a regularization term in an optimization problem, the function $g(x)$ defined in (3.10)

promotes the property that x can be expressed as a nonnegative linear combination of a small number of elements (or 'atoms') of C.

The vector ℓ_1-norm and the matrix trace norm are the best known examples of atomic norms. The ℓ_1-norm of a real or complex n-vector is the atomic norm associated with $C = \{se_k \mid |s| = 1,\ k = 1, \ldots, n\}$, where e_k is the kth unit vector of length n. The matrix trace norm (or nuclear norm) is the atomic norm for the set of rank-1 matrices with unit norm. Specifically, the trace norm on $\mathbf{C}^{n\times m}$ is the atomic norm for $C = \{vw^H \mid v \in \mathbf{C}^n,\ w \in \mathbf{C}^m,\ \|v\| = \|w\| = 1\}$, where w^H is the conjugate transpose and $\|\cdot\|$ denotes the Euclidean norm. Many other examples are discussed in [CRP12, BTR13, TBS13].

3.2 Related works

It is interesting to note that applications of ℓ_1-norm optimization for sparse signal processing often involve a discretization of a sparse optimization problem over the underlying continuous domains (time, space, or frequency domain) [CRT06, CM73, SS86, SBT12]. The reason for adopting a discretization procedure was either that no exact method for the continuous problem was known, or that solving the discretized version was believed to be numerically cheaper or easier to implement. However, several issues arise because of the discretization. The continuous-domain signal may not be sparsely representable after discretization (often referred to as 'basis mismatch,' see [TBS13, CSP11] for example), resulting in inaccurate or non-sparse estimation. As the grid gets finer, the discretization problem typically becomes more ill-conditioned, which leads to numerical difficulties in optimization algorithms. Finally, the discretized problem may be very large and expensive to solve (see [TBS13], and references therein).

For these reasons, several researchers have recently studied exact formulations for certain sparse signal reconstruction problems in continuous domains. In particular, the atomic norm associated with the set

$$C_{\mathrm{e}} = \{\gamma\,(1, e^{\mathrm{j}\omega}, \ldots, e^{\mathrm{j}(n-1)\omega}) \in \mathbf{C}^n \mid \omega \in [0, 2\pi),\ |\gamma| = 1/\sqrt{n}\} \tag{3.11}$$

has been studied by several groups [CG12, CGH17, CF13, CF14, Fer15, BTR13, TBS13, YX16, LC16, MCK14, MCK15]. Problems of this type are widely encountered in signal processing and system theory. They include very classical problems, for example the estimation of line spectra, direction of arrival estimation in sensor array processing, and the estimation of spike signals. It is known that the atomic norm for the set C_{e} is the optimal value of the semidefinite program (SDP)

$$\begin{array}{ll} \text{minimize} & (\mathbf{tr}\, V + w)/2 \\ \text{subject to} & \left[\begin{array}{cc} V & x \\ x^H & w \end{array}\right] \succeq 0 \\ & V \text{ is Toeplitz,} \end{array} \tag{3.12}$$

with variables $w \in \mathbf{R}$ and $V \in \mathbf{H}^n$ (the $n \times n$ Hermitian matrices). This result can be shown via convex duality and semidefinite characterizations of bounded trigonometric polynomials [BTR13, CF14, CG12], or directly by referring to the classical Vandermonde decomposition of positive semidefinite Toeplitz matrices as a positive sum of outer product of vectors in C_{e} [TBS13, STY14]. More generally, one can consider the atomic norm of the set of matrices $C = \{vw^H \in \mathbf{C}^{n\times m} \mid v \in C_{\mathrm{e}},\ \|w\| = 1\}$. The atomic norm for this set, evaluated at a matrix $X \in \mathbf{C}^{n\times m}$, is the optimal value of the SDP

$$\begin{array}{ll} \text{minimize} & (\mathbf{tr}\, V + \mathbf{tr}\, W)/2 \\ \text{subject to} & \left[\begin{array}{cc} V & X \\ X^H & W \end{array}\right] \succeq 0 \\ & V \text{ is Toeplitz,} \end{array} \tag{3.13}$$

with variables $V \in \mathbf{H}^n$ and $W \in \mathbf{H}^m$; see [YX16, LC16, Fer15, CV16, CV17]. Further extensions, that place restrictions on the parameter ω in the definition (3.11), can be found in [MCK14, MCK15, CV16, CV17].

3.3 Semidefinite representation of gauges for structured symmetric matrices

In this thesis we discuss extensions of the SDP representations (3.12) and (3.13) to a larger class of atomic norms and gauge functions. Recall for easy reference definitions (2.7) and (2.8) of the set in Section 2.4,

$$\mathcal{A} = \{a \in \mathbf{C}^n \mid (\mu G - \nu F)a = 0,\ (\mu,\nu) \in \mathcal{C}\} \tag{3.14}$$

where $F, G \in \mathbf{C}^{p\times n}$, and

$$\mathcal{C} = \left\{(\mu,\nu) \in \mathbf{C}^2 \mid (\mu,\nu) \neq 0,\ q_\Phi(\mu,\nu) = 0,\ q_\Psi(\mu,\nu) \leq 0\right\} \tag{3.15}$$

where $\Phi, \Psi \in \mathbf{H}^2$ such that $\det \Phi < 0$, with the quadratic forms

$$q_\Phi(\mu,\nu) = \begin{bmatrix} \mu \\ \nu \end{bmatrix}^H \Phi \begin{bmatrix} \mu \\ \nu \end{bmatrix}, \qquad q_\Psi(\mu,\nu) = \begin{bmatrix} \mu \\ \nu \end{bmatrix}^H \Psi \begin{bmatrix} \mu \\ \nu \end{bmatrix}.$$

We extend the gauge function representations by making a similar observation as in Chapter 2 that C_{e} (3.11) can be parameterized as

$$C_{\mathrm{e}} = \{a \mid a \in \mathcal{A},\ \|a\| = 1\} \tag{3.16}$$

where the matrices F, G, Φ, Ψ are defined in (2.16). In addition to the extensions already mentioned in Chapter 2, where we allow more general choices of F, G, Φ, and Ψ, we also replace the normalization $\|a\| = 1$ with a condition of the type $\|Ea\| \leq 1$ where E is not necessarily full column rank.

A function g is called a *gauge* if it is convex, positively homogeneous ($g(tx) = tg(x)$ for $t > 0$), nonnegative, and vanishes at the origin [Roc70, §15], [KN77, §1]. Examples are the *(Minkowski) gauges* of nonempty convex sets C, which are defined as

$$g(x) = \inf\{t \geq 0 \mid x \in tC\}.$$

Conversely, if g is a gauge, then it is the Minkowski gauge of the set $C = \{x \mid g(x) \leq 1\}$. A gauge is a norm if it is defined everywhere, positive except at the origin, and symmetric ($g(x) = g(-x)$).

The gauge of the convex hull $\operatorname{conv} C$ of a set C can be expressed as

$$g(x) = \inf\{\sum_{k=1}^{r} \theta_k \mid x = \sum_{k=1}^{r} \theta_k a_k,\ \theta_k \geq 0,\ a_k \in C,\ k = 1, \ldots, r\}.$$

The minimum is over all possible decompositions of x as a nonnegative combination of a finite number of elements of C. The gauge of the convex hull of a compact set is also called the *atomic norm* associated with the set [CRP12].

Let F, G, Φ, Ψ be defined as in Theorem 2.5, where F, $G \in \mathbf{C}^{p\times n}$ and Φ, $\Psi \in \mathbf{H}^2$ with $\det \Phi < 0$. We assume that the set $\mathcal{C}$ defined in (3.15) is not empty. In this section we first discuss the gauge of the convex hull of the set

$$C = \{aa^H \in \mathbf{H}^n \mid a \in \mathcal{A},\ \|a\| = 1\},$$

where $\mathcal{A}$ is defined in (3.14). The gauge of the convex hull of C is the function

$$\begin{aligned} g(X) &= \inf\{\sum_{k=1}^{r} \theta_k \mid X = \sum_{k=1}^{r} \theta_k a_k a_k^H,\ \theta_k \geq 0,\ a_k \in \mathcal{A},\ \|a_k\| = 1\} && (3.17)\\ &= \inf\{\sum_{k=1}^{r} \|a_k\|^2 \mid X = \sum_{k=1}^{r} a_k a_k^H,\ a_k \in \mathcal{A}\}. && (3.18) \end{aligned}$$

The second expression follows from the fact that if $a \in \mathcal{A}$ then $\beta a \in \mathcal{A}$ for all β.

The expressions $\sum_k \theta_k$ and $\sum_k \|a_k\|^2$ in these minimizations take only two possible values: $\mathbf{tr}\, X$ if X can be decomposed as in (3.17) and (3.18), and $+\infty$ otherwise. Theorem 2.5 tells us that a decomposition exists if and only if X is positive semidefinite and satisfies the two constraints

$$\Phi_{11} F X F^H + \Phi_{21} F X G^H + \Phi_{12} G X F^H + \Phi_{22} G X G^H = 0 \tag{3.19}$$

$$\Psi_{11} F X F^H + \Psi_{21} F X G^H + \Psi_{12} G X F^H + \Psi_{22} G X G^H \preceq 0. \tag{3.20}$$

Therefore

$$g(X) = \begin{cases} \mathbf{tr}\, X & X \succeq 0,\ (3.19),\ (3.20) \\ +\infty & \text{otherwise.} \end{cases} \tag{3.21}$$

Now consider an optimization problem in which we minimize the sum of a function $f : \mathbf{H}^n \to \mathbf{R}$ and the gauge defined in (3.18) and (3.21),

$$\text{minimize} \quad f(X) + g(X). \tag{3.22}$$

If we substitute the definition (3.18), this can be written as

$$\begin{array}{ll} \text{minimize} & f(X) + \sum_{k=1}^{r} \|a_k\|^2 \\ \text{subject to} & X = \sum_{k=1}^{r} a_k a_k^H \\ & a_k \in \mathcal{A}, \quad k = 1, \ldots, r. \end{array} \tag{3.23}$$

The variables are X and the parameters a_1, $\ldots$, a_r, and r of the decomposition of X. This formulation shows that the function $g(X)$ in (3.22) acts as a regularization term that promotes a structured low rank property in X. If we substitute the expression (3.21) we obtain the equivalent formulation

$$\begin{array}{ll} \text{minimize} & f(X) + \mathbf{tr}\, X \\ \text{subject to} & \Phi_{11} F X F^H + \Phi_{21} F X G^H + \Phi_{12} G X F^H + \Phi_{22} G X G^H = 0 \\ & \Psi_{11} F X F^H + \Psi_{21} F X G^H + \Psi_{12} G X F^H + \Psi_{22} G X G^H \preceq 0 \\ & X \succeq 0. \end{array} \tag{3.24}$$

This problem is convex if f is convex.

A useful generalization of (3.18) is the gauge of the convex hull of

$$C = \{aa^H \mid a \in \mathcal{A},\ \|Ea\| \leq 1\}$$

where E may have rank less than n. The gauge of $\operatorname{conv} C$ is

$$g(X) = \inf \{ \sum_{k=1}^{r} \theta_k \mid X = \sum_{k=1}^{r} \theta_k a_k a_k^H,\ \theta_k \geq 0,\ a_k \in \mathcal{A},\ \|Ea_k\| \leq 1 \}. \tag{3.25}$$

The variables θ_k in this definition can be eliminated by making the following observation. Suppose that the directions of the vectors a_k in the decomposition of X in (3.25) are given, but not their norms or the coefficients θ_k. If $0 < \|Ea_k\| < 1$, we can decrease θ_k by scaling a_k until $\|Ea_k\| = 1$. If $Ea_k = 0$, θ_k can be made arbitrarily small by scaling a_k. Hence, we obtain the same result if we use $\sqrt{\theta}_k a_k$ as variables and write the infimum as:

$$g(X) = \inf \{ \sum_{k=1}^{r} \|Ea_k\|^2 \mid X = \sum_{k=1}^{r} a_k a_k^H,\ a_k \in \mathcal{A},\ k = 1, \ldots, r \}. \tag{3.26}$$

Therefore $g(X) = \sum_k \|Ea_k\|^2 = \mathbf{tr}(EXE^H)$ if X can be decomposed as in (3.26) and $+\infty$ otherwise. Using Theorem 2.5 we can express this result as

$$g(X) = \begin{cases} \mathbf{tr}(EXE^H) & X \succeq 0,\ (3.19),\ (3.20) \\ +\infty & \text{otherwise.} \end{cases} \tag{3.27}$$

Minimizing $f(X) + g(X)$ is equivalent to the optimization problem

$$\begin{array}{ll} \text{minimize} & f(X) + \sum_{k=1}^{r} \|Ea_k\|^2 \\ \text{subject to} & X = \sum_{k=1}^{r} a_k a_k^H \\ & a_k \in \mathcal{A}, \quad k = 1, \ldots, r, \end{array} \tag{3.28}$$

with variables X and the parameters a_1, $\ldots$, a_r, r of the decomposition of X. When $E^H E = I$ this is the same as (3.23). By choosing different E we assign different weights to the vectors a_k. Using the expression (3.27), the problem (3.28) can be written as

$$\begin{array}{ll} \text{minimize} & f(X) + \mathbf{tr}\,(EXE^H) \\ \text{subject to} & \Phi_{11} FXF^H + \Phi_{21} FXG^H + \Phi_{12} GXF^H + \Phi_{22} GXG^H = 0 \\ & \Psi_{11} FXF^H + \Psi_{21} FXG^H + \Psi_{12} GXF^H + \Psi_{22} GXG^H \preceq 0 \\ & X \succeq 0. \end{array} \tag{3.29}$$

Example Consider the line spectrum estimation example in Section 2.3.3. Classical methods, such as MUSIC and ESPRIT, are based on the eigenvalue decomposition of an estimated covariance matrix (see Appendix A). With the formulation outlined in this section one can solve related but more general covariance fitting problems, expressed as

$$\begin{array}{ll} \text{minimize} & f(R) + n \sum_{k=1}^{r} |c_k|^2 \\ \text{subject to} & R = \sigma^2 I + \sum_{k=1}^{r} |c_k|^2 \begin{bmatrix} 1 \\ e^{\mathrm{j}\omega_k} \\ \vdots \\ e^{\mathrm{j}(n-1)\omega_k} \end{bmatrix} \begin{bmatrix} 1 \\ e^{\mathrm{j}\omega_k} \\ \vdots \\ e^{\mathrm{j}(n-1)\omega_k} \end{bmatrix}^H, \end{array}$$

with variables $R \in \mathbf{H}^n$, σ^2, $|c_k|$, ω_k, and r, where f is a convex penalty or indicator function that measures the quality of the fit between R and the estimated covariance matrix. This is equivalent to the convex optimization problem

$$\begin{array}{ll} \text{minimize} & f(X + tI) + \mathbf{tr}\, X \\ \text{subject to} & X \succeq 0, \; t \geq 0 \\ & X \text{ is Toeplitz.} \end{array}$$

A numerical example is given in Section 3.5.

3.4 Duality

In this section we derive the conjugates of the gauge functions defined in Section 3.3 and show that they can be interpreted as indicator functions of sets of nonnegative or bounded generalized polynomials. This gives a useful interpretation of the dual problems for (3.22). Using only SDP duality, we derive the complementary slackness relations expressed in terms of points and polynomials on the one-dimensional continuous domain $\mathcal{C}$. Special cases of the relations with vectors of complex exponentials C_{e} were previously shown via either infinite-dimensional convex duality [CG12, CF14] or atomic norm duality [CRP12, TBS13].

We assume that the subset of the complex plane represented by $\mathcal{C}$ in (3.15) is one-dimensional, *i.e.*, $\mathcal{C}$ is not a singleton and not the empty set. Equivalently, the inequality $q_\Psi(\mu,\nu) \leq 0$ in the definition is either redundant (and $\mathcal{C}$ represents a line or circle), or it is not redundant and then there exist elements of $\mathcal{C}$ with $q_\Psi(\mu,\nu) < 0$. When stating and analyzing the dual problems, we will need to distinguish these two cases ($q_\Psi(\mu,\nu) \leq 0$ is redundant or not). For the sake of brevity we only give the formulas for the case where the inequality is not redundant. The dual problems for the other case follow by setting $\Psi = 0$ and making obvious simplifications.

We also assume that $\mu G - \nu F$ has full row rank ($\mathbf{rank}(\mu G - \nu F) = p$) for all nonzero (μ,ν)). This condition will serve as a 'constraint qualification' that guarantees strong duality.

3.4.1 Conjugate of symmetric matrix gauge

We first consider the conjugate of the function g defined in (3.27). The conjugate is defined as $g^*(Z) = \sup_X (\mathbf{tr}(XZ) - g(X))$, *i.e.*, the optimal value of the SDP

$$\begin{array}{ll} \text{maximize} & \mathbf{tr}\left((Z - E^H E)X\right) \\ \text{subject to} & X \succeq 0 \\ & \Phi_{11} F X F^H + \Phi_{21} F X G^H + \Phi_{12} G X F^H + \Phi_{22} G X G^H = 0 \\ & \Psi_{11} F X F^H + \Psi_{21} F X G^H + \Psi_{12} G X F^H + \Psi_{22} G X G^H \preceq 0. \end{array} \tag{3.30}$$

The dual of this problem is

$$\begin{array}{ll} \text{minimize} & 0 \\ \text{subject to} & Z - \left[\begin{array}{c} F \\ G \end{array}\right]^H (\Phi \otimes P + \Psi \otimes Q) \left[\begin{array}{c} F \\ G \end{array}\right] \preceq E^H E \\ & Q \succeq 0, \end{array} \tag{3.31}$$

with variables $P, Q \in \mathbf{H}^p$. It is shown in Appendix D that strong duality holds under the assumptions listed at the top of Section 3.4.

If strong duality holds, then $g^*(Z)$ is the optimal value of (3.31), *i.e.*, equal to zero if there exist P, Q that satisfy the constraints in (3.31), and $+\infty$ otherwise. In other words, $g^*(Z)$ is the indicator function of the set described by the constraints in (3.31). To complete the picture, we now show that $g^*(Z)$ can be expressed as

$$g^*(Z) = \begin{cases} 0 & a^H Z a \leq \|Ea\|^2 \text{ for all } a \in \mathcal{A} \\ +\infty & \text{otherwise.} \end{cases} \tag{3.32}$$

This expression of g^* follows directly from the definition of the conjugate and (3.26), since

$$g^*(Z) = \sup_X \left(\mathbf{tr}\,(XZ) - g(X)\right) = \sup_{a_1,\ldots,a_r \in \mathcal{A}} \sum_{k=1}^{r} (a_k^H Z a_k - \|Ea_k\|^2),$$

which is the same as (3.32). This is consistent with a property from gauge duality: the conjugate of the gauge of a set is the indicator of the unit level set of the polar gauge [FMP14, proposition 2.1]. It is also instructive to derive (3.32) from the dual SDP (3.31). The expression (3.32) is obtained immediately by applying Theorem 2.7; however, we include the derivation here to make the presentation in this chapter self-contained. Suppose P and Q are feasible in (3.31). Consider any $a \in \mathcal{A}$ and $(\mu, \nu) \in \mathcal{C}$ with $\mu G a = \nu F a$. Define $y = (1/\nu)Ga$ if $\nu \neq 0$ and $y = (1/\mu)Fa$ otherwise. Then

$$\begin{array}{rcl} a^H Z a - \|Ea\|^2 & \leq & \left[\begin{array}{c} Fa \\ Ga \end{array}\right]^H (\Phi \otimes P + \Psi \otimes Q) \left[\begin{array}{c} Fa \\ Ga \end{array}\right] \\ & = & \left[\begin{array}{c} \mu y \\ \nu y \end{array}\right]^H (\Phi \otimes P + \Psi \otimes Q) \left[\begin{array}{c} \mu y \\ \nu y \end{array}\right] \\ & = & (y^H P y) q_\Phi(\mu, \nu) + (y^H Q y) q_\Psi(\mu, \nu) \\ & \leq & 0. \end{array}$$

The last line follows from $Q \succeq 0$ and $q_\Phi(\mu,\nu) = 0$, $q_\Psi(\mu,\nu) \leq 0$. Conversely, if problem (3.31) is infeasible, then the optimal value is $+\infty$ and, since strong duality holds, there exist matrices X that are feasible for (3.30) with $\mathbf{tr}((Z - E^H E)X) > 0$. Using Theorem 2.5, we see that there exist $a_1, \ldots, a_r \in \mathcal{A}$ with

$$\sum_{k=1}^{r} (a_k^H Z a_k - \|Ea_k\|^2) > 0.$$

Therefore $a_k^H Z a_k > \|Ea_k\|^2$ for at least one a_k.

3.4.2 Dual problem interpretation

The interpretation of the conjugate gives useful insight in problem (3.22), where g is defined in (3.27). The dual problem is

$$\text{maximize} \quad -f^*(Z) - g^*(-Z).$$

Expanding $g^*(-Z)$ using (3.31) gives the equivalent problem

$$\begin{array}{ll} \text{maximize} & -f^*(Z) \\ \text{subject to} & -Z - \left[\begin{array}{c} F \\ G \end{array}\right]^H (\Phi \otimes P + \Psi \otimes Q) \left[\begin{array}{c} F \\ G \end{array}\right] \preceq E^H E \\ & Q \succeq 0, \end{array} \tag{3.33}$$

with variables Z, P, Q, and using the expression (3.32) we can put the constraints in this problem more succinctly as

$$\begin{array}{ll} \text{maximize} & -f^*(Z) \\ \text{subject to} & \|Ea\|^2 + a^H Z a \geq 0 \quad \text{for all } a \in \mathcal{A}. \end{array} \tag{3.34}$$

This last form leads to an interesting set of optimality conditions. Suppose X and Z are feasible for (3.28) and (3.34), respectively. Then

$$\begin{aligned} f(X) + \sum_{k=1}^{r} \|Ea_k\|^2 &\geq -f^*(Z) + \mathbf{tr}(XZ) + \sum_{k=1}^{r} \|Ea_k\|^2 \\ &= -f^*(Z) + \sum_{k=1}^{r} (\|Ea_k\|^2 + a_k^H Z a_k) \\ &\geq -f^*(Z). \end{aligned}$$

The first inequality follows by definition of $f^*(Z)$, and the second and third line from primal and dual feasibility. If X and Z are optimal and strong duality holds, then

$$f(X) + \sum_{k=1}^{r} \|Ea_k\|^2 = -f^*(Z).$$

This is only possible if $f(X) + f^*(Z) = \mathbf{tr}(XZ)$ and $\|Ea_k\|^2 + a_k^H Z a_k = 0$ for $k = 1, \ldots, r$. Hence only the vectors $a \in \mathcal{A}$ at which the inequality in (3.34) is active, can be used to form an optimal $X = \sum_k a_k a_k^H$.

Example: Generalized Kalman–Yakubovich–Popov lemma When specialized to the controllability pencil (2.24), the equivalence between the constraints in (3.34) and (3.33) is known as the (generalized) Kalman–Yakubovich–Popov lemma [Kal63, Yak62, Pop62, Sch06, IH05]; see Section 2.7 for more details.

We assume that A has no eigenvalues λ with $(\lambda, 1) \in \mathcal{C}$, and that the pair (A, B) is controllable, so the pencil satisfies the rank condition that $\mathbf{rank}(\lambda F - G) = n_{\mathrm{s}}$ for all λ. The dual problem (3.34) becomes

$$\begin{array}{ll} \text{maximize} & -f^*(Z) \\ \text{subject to} & \mathcal{F}(\lambda, Z) \succeq 0 \quad \text{for all } (\lambda, 1) \in \mathcal{C} \\ & M_{22} + Z_{22} \succeq 0 \quad \text{if } (1, 0) \in \mathcal{C} \end{array}$$

where $M = E^H E$ and

$$\mathcal{F}(\lambda, Z) = \begin{bmatrix} (\lambda I - A)^{-1} B \\ I \end{bmatrix}^H \begin{bmatrix} M_{11} + Z_{11} & M_{12} + Z_{12} \\ M_{21} + Z_{21} & M_{22} + Z_{22} \end{bmatrix} \begin{bmatrix} (\lambda I - A)^{-1} B \\ I \end{bmatrix}.$$

The function $\mathcal{F}$ is called the *Popov function* with central matrix $M + Z$ [IOW99, HSK99].

3.5 Line spectrum estimation examples

The formulations in Section 3.3 will now be illustrated with examples of Toeplitz covariance fitting from signal processing. The optimization problems were solved with CVX [GB14].

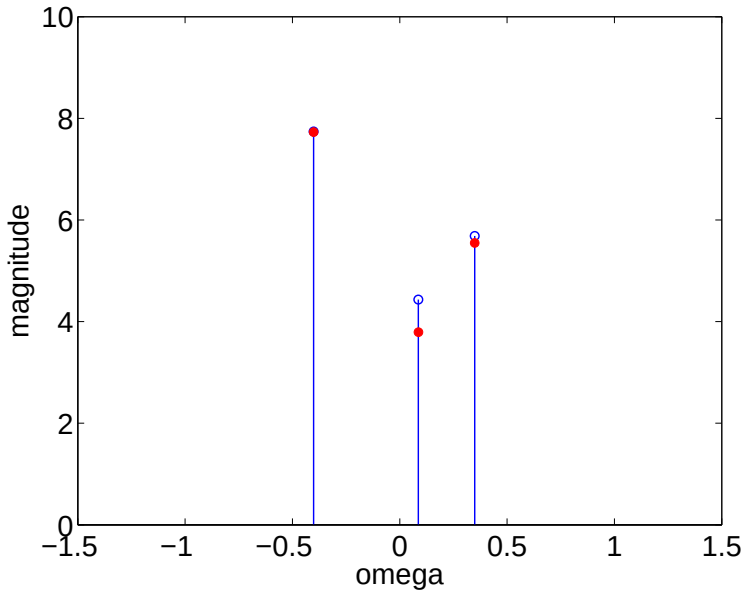

Figure 3.1: Line spectrum estimation by Toeplitz covariance fitting (Section 3.5.1). The red dots represent the frequencies and magnitudes of the true model. The blue lines show the estimated parameters obtained by solving (3.36).

3.5.1 Gaussian white noise model

We fit a covariance matrix of the form (2.6) to an estimated covariance matrix R_{m}. The estimate R_{m} is constructed from $N = 150$ samples of the time series $y(t)$ defined in (2.5), with $r = 3$, and frequencies ω_k and magnitudes $|c_k|$ shown in Figure 3.1. The noise is Gaussian white noise with variance $\sigma^2 = 64$. The sample covariance matrix is constructed with $n = 30$ as

$$R_{\mathrm{m}} = \frac{1}{N - n + 1} Y Y^H \tag{3.35}$$

where Y is the $n \times (N - n + 1)$ Hankel matrix with $y(1), \ldots, y(N - n + 1)$ in its first row. To estimate the model we solve

$$\begin{array}{ll} \text{minimize} & \gamma\|R - R_{\text{m}}\|_2 + \sum_{k=1}^{r} |c_k|^2 \\ \text{subject to} & R = \sigma^2 I + \sum_{k=1}^{r} |c_k|^2 \begin{bmatrix} 1 \\ e^{\text{j}\omega_k} \\ \vdots \\ e^{\text{j}(n-1)\omega_k} \end{bmatrix} \begin{bmatrix} 1 \\ e^{\text{j}\omega_k} \\ \vdots \\ e^{\text{j}(n-1)\omega_k} \end{bmatrix}^H, \end{array} \tag{3.36}$$

with variables σ^2, $|c_k|^2$, ω_k, r, and R. The norm $\|\cdot\|_2$ in the objective is the spectral norm. The regularization parameter γ is set to 0.25. As can be seen from Figure 3.1, the recovered parameters ω_k and $|c_k|$ are quite accurate, despite the very low signal-to-noise ratio. The estimated noise variance σ^2 is 79.6.

The semidefinite optimization approach allows us to fit a covariance matrix with the structure prescribed in (2.6) to a sample covariance matrix that may not be Toeplitz or positive semidefinite.

3.5.2 Moving average noise model

The covariance fitting formulation can also be extended to applications where the noise $v(t)$ is modeled as a moving-average process [Geo06, SDL10].

Moving average model Consider a zero-mean moving average model that describes a time series as a weighted sum of a white noise series. Specifically, a moving average model of order m (denoted MA(m)) takes the form

$$v(t) = e(t) + \sum_{k=1}^{m} \sigma_k e(t-k)$$

where σ_k are scalar (real or complex) coefficients and $e(t)$ is a (circular) white noise series. It is a stationary process, and the covariance sequence vanishes for $|k| > m$, *i.e.*,

$$\mathbf{E}[v(t)\overline{v(t-k)}] = 0 \text{ for } |k| > m.$$

In other words, the covariance matrix of $v(t)$ is a banded Toeplitz matrix.

Consider fitting signal models of the form (2.5)

$$y(t) = \sum_{k=1}^{r} c_k e^{\mathrm{j}\omega_k t} + v(t),$$

where the frequencies are restricted to $|\omega_k| \leq \omega_{\mathrm{c}}$ with a cutoff frequency ω_{c}, and the noise $v(t)$ is a moving average process of order m. One can formulate a covariance fitting problem

$$\begin{aligned}
\text{minimize} \quad & \gamma \|R - R_{\mathrm{m}}\|_2 + \sum_{k=1}^{r} |c_k|^2 \\
\text{subject to} \quad & R = P + \sum_{k=1}^{r} |c_k|^2 \begin{bmatrix} 1 \\ e^{\mathrm{j}\omega_k} \\ \vdots \\ e^{\mathrm{j}(n-1)\omega_k} \end{bmatrix} \begin{bmatrix} 1 \\ e^{\mathrm{j}\omega_k} \\ \vdots \\ e^{\mathrm{j}(n-1)\omega_k} \end{bmatrix}^H \\
& |\omega_k| \leq \omega_{\mathrm{c}}, \quad k = 1, \ldots, r \\
& P \text{ is a covariance of MA}(m).
\end{aligned}$$

The term $\|R - R_{\mathrm{m}}\|_2$ in the objective promotes closeness of the identified covariance matrix R to the sample covariance matrix R_{m}, which may be non-Toeplitz or not positive semidefinite, depending on how it is constructed. The closeness can also be measured with other matrix norms or distance metrics such as discussed in [Geo07]. The second term is in effect the same as maximizing $\mathbf{tr}\, P$ (or minimizing $-\mathbf{tr}\, P$) with an appropriate choice of γ. It encourages a larger noise variance in the identified model, and a smaller sum of the signal powers. In fact, at the optimum, the second term in the identified covariance is always singular, because if it is nonsingular, we can increase the diagonal elements in P such that it is still a covariance matrix of MA(m) [Geo06].

It can be shown that the problem is equivalent to the SDP

$$
\begin{array}{ll}
\text{minimize} & \gamma\|R - R_{\mathrm{m}}\|_2 + \mathbf{tr}\, X/n \\
\text{subject to} & R = P + X \\
& X \text{ is Toeplitz} \\
& -FXG^T - GXF^T + 2(\cos\omega_{\mathrm{c}})GXG^T \preceq 0 \\
& X \succeq 0 \\
& P = \begin{bmatrix}
p_0 & \cdots & \bar{p}_m & & & \\
\vdots & \ddots & & \ddots & & \\
p_m & & \ddots & & \ddots & \\
& \ddots & & \ddots & & \bar{p}_m \\
& & \ddots & & \ddots & \vdots \\
& & & p_m & \cdots & p_0
\end{bmatrix} \\
& (p_0, \ldots, p_m) = \mathcal{D}(Q) \\
& Q \succeq 0,
\end{array}
$$

where F, G are defined as in (2.16) and $\mathcal{D}$ is defined in (2.2).

As a numerical example, we construct an estimated covariance matrix R_{m} as in (3.35) with $n = 30$ from $N = 250$ samples of $y(t)$, which is a real-valued signal generated from the superposition of line spectrum and a MA(3) noise with variance 51.75, shown in red stems and curves, respectively, in Figure 3.2. We compare estimation results between fitting with white noise model (MA(0)) and with MA(3) noise model. In Figure 3.2, the regularization parameter $\gamma = 8.0$ and the estimated noise variance is 51.41 for the top figure, and for the bottom figure, $\gamma = 18$ and the estimated noise variance is 50.77. As is visually notable, fitting with MA(3) noise model can provide a more accurate estimate of the signal model.

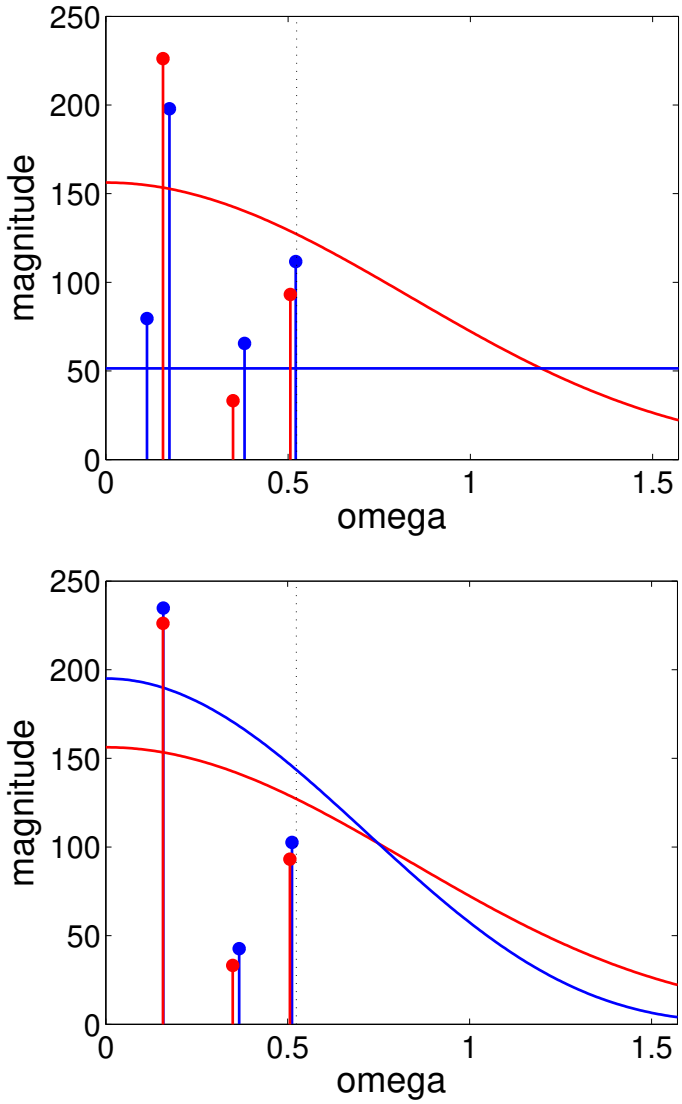

Figure 3.2: Line spectrum estimation by Toeplitz covariance fitting (Section 3.5.2) with white noise model (MA(0)) (*Top*) and with MA(3) model (*Bottom*). The red (blue) stems represent the true (estimated) line spectrum, and the red (blue) curve represents the true (estimated) noise spectrum. The dotted vertical line indicates the cutoff frequency $\omega_c = \pi/6$.

CHAPTER 4

Gauge Penalties for Structured Nonsymmetric matrices

We now extend the results of the previous chapter to nonsymmetric matrices. The penalty functions discussed in this chapter can also be interpreted as extensions to the trace norm to take into account certain matrix structures. In particular, it is well known that the trace norm $\|Y\|_*$ of a matrix Y is the optimal value of the SDP

$$\begin{array}{ll} \text{minimize} & (\mathbf{tr}\, V + \mathbf{tr}\, W)/2 \\ \text{subject to} & \begin{bmatrix} V & Y \\ Y^H & W \end{bmatrix} \succeq 0, \end{array} \tag{4.1}$$

where V and W are variables of compatible sizes. We will show that this SDP (4.1) is a special case of the SDP representations of our penalty functions. In Section 4.1, we derive SDP representations of atomic norms and gauge functions for structured nonsymmetric matrices.

The chapter is organized as follows. In Section 4.2, we derive the convex conjugates of these gauge functions and discuss optimality conditions. In Section 4.3, the techniques are illustrated with examples of low-rank matrix approximation problems arising in spectral estimation and array processing. Most content of this chapter is adapted from [CV16,CV17]. Some background and related works have been given in Sections 3.1 and 3.2.

4.1 Semidefinite representation of gauges for structured nonsymmetric matrices

We define F, G, E, Φ, Ψ, and $\mathcal{A}$ as in the previous chapter, but add the assumption that the matrices F, G, and E are block-diagonal:

$$G = \begin{bmatrix} G_1 & 0 \\ 0 & G_2 \end{bmatrix}, \qquad F = \begin{bmatrix} F_1 & 0 \\ 0 & F_2 \end{bmatrix}, \qquad E = \begin{bmatrix} E_1 & 0 \\ 0 & E_2 \end{bmatrix}. \tag{4.2}$$

Here F_1, $G_1 \in \mathbf{C}^{p_1 \times n_1}$ and F_2, $G_2 \in \mathbf{C}^{p_2 \times n_2}$ (possibly with $p_1 = 0$ or $p_2 = 0$). The matrices E_1 and E_2 have n_1 and n_2 columns, respectively. We discuss the function

$$h(Y) = \frac{1}{2} \inf_{V,W} g(\begin{bmatrix} V & Y \\ Y^H & W \end{bmatrix})$$

of $Y \in \mathbf{C}^{n_1 \times n_2}$, where g is defined in (3.26) and (3.27). Using (3.26) we can write $h(Y)$ as

$$h(Y) = \inf \{ \frac{1}{2} \sum_{k=1}^{r} (\|E_1 v_k\|^2 + \|E_2 w_k\|^2) \mid Y = \sum_{k=1}^{r} v_k w_k^H, \ (v_k, w_k) \in \mathcal{A} \}, \tag{4.3}$$

while the characterization (3.27) shows that $h(Y)$ is the optimal value of the SDP

$$\begin{array}{ll} \text{minimize} & \left(\mathbf{tr}(E_1 V E_1^H) + \mathbf{tr}(E_2 W E_2^H)\right)/2 \\ \text{subject to} & \Phi_{11} F X F^H + \Phi_{21} F X G^H + \Phi_{12} G X F^H + \Phi_{22} G X G^H = 0 \\ & \Psi_{11} F X F^H + \Psi_{21} F X G^H + \Psi_{12} G X F^H + \Psi_{22} G X G^H \preceq 0 \\ & X = \begin{bmatrix} V & Y \\ Y^H & W \end{bmatrix} \succeq 0, \end{array} \tag{4.4}$$

with V and W as variables. This can be seen as an extension of the well-known SDP formulation (4.1) of the trace norm of a matrix. If we take F and G to have zero row dimensions (equivalently, define $\mathcal{A} = \mathbf{C}^{n_1} \times \mathbf{C}^{n_2}$ and omit the first two constraints in (4.4)) and choose $E_1 = I$, $E_2 = I$, then $h(Y) = \|Y\|_*$, the trace norm of Y.

The block-diagonal form of F and G implies that if $(v, w) \in \mathcal{A}$, then $(\alpha v, \beta w) \in \mathcal{A}$ for all α, β. This observation leads to a number of useful equivalent expressions for (4.3). First, we note that $h(Y)$ can be written as

$$h(Y) = \inf \{ \sum_{k=1}^{r} \|E_1 v_k\| \|E_2 w_k\| \mid Y = \sum_{k=1}^{r} v_k w_k^H, \ (v_k, w_k) \in \mathcal{A} \}. \tag{4.5}$$

This follows from the fact $\|E_1 v_k\|^2 + \|E_2 w_k\|^2 \geq 2\|E_1 v_k\|\|E_2 w_k\|$, with equality if $\|E_1 v_k\| = \|E_2 w_k\|$. If the decomposition of Y in (4.3) involves a term $v_k w_k^H$ with $E_1 v_k$ and $E_2 w_k$ nonzero, then replacing v_k and w_k with

$$\tilde{v}_k = (\frac{\|E_2 w_k\|}{\|E_1 v_k\|})^{1/2} v_k, \qquad \tilde{w}_k = (\frac{\|E_1 v_k\|}{\|E_2 w_k\|})^{1/2} w_k$$

gives another valid decomposition with

$$\frac{1}{2}(\|E_1 \tilde{v}_k\|^2 + \|E_2 \tilde{w}_k\|^2) = \|E_1 v_k\|\|E_2 w_k\| \leq \frac{1}{2}(\|E_1 v_k\|^2 + \|E_2 w_k\|^2).$$

If $E_1 v_k = 0$ and $E_2 w_k \neq 0$, then replacing v_k and w_k with $\tilde{v}_k = \alpha v_k$, $\tilde{w}_k = (1/\alpha) w_k$ gives an equivalent decomposition with

$$\frac{1}{2}(\|E_1 \tilde{v}_k\|^2 + \|E_2 \tilde{w}_k\|^2) = \frac{1}{2\alpha^2}\|E_2 w_k\|^2 \to 0$$

as α goes to infinity. The same argument applies when $E_1 v_k \neq 0$ and $E_2 w_k = 0$. In all cases, therefore, the two expressions (4.3) and (4.5) give the same result.

From (4.5) we obtain two other useful expressions:

$$h(Y) = \inf\{\sum_{k=1}^{r} \|E_1 v_k\| \mid Y = \sum_{k=1}^{r} v_k w_k^H,\ (v_k, w_k) \in \mathcal{A},\ \|E_2 w_k\| \leq 1\} \tag{4.6}$$

$$= \inf\{\sum_{k=1}^{r} \|E_2 w_k\| \mid Y = \sum_{k=1}^{r} v_k w_k^H,\ (v_k, w_k) \in \mathcal{A},\ \|E_1 v_k\| \leq 1\}. \tag{4.7}$$

This again follows from the property that the components v_k, w_k of elements (v_k, w_k) in $\mathcal{A}$ can be scaled independently. At the optimal decomposition in (4.6), all terms satisfy $E_2 w_k = 0$ or $\|E_2 w_k\| = 1$. In (4.7), all terms satisfy $E_1 v_k = 0$ or $\|E_1 v_k\| = 1$.

A final interpretation of h is

$$\begin{aligned} h(Y) = \inf\{\sum_{k=1}^{r} \theta_k \mid Y = \sum_{k=1}^{r} \theta_k v_k w_k^H, \\ \theta_k \geq 0,\ (v_k, w_k) \in \mathcal{A},\ \|E_1 v_k\| \leq 1,\ \|E_2 w_k\| \leq 1\}. \end{aligned} \tag{4.8}$$

The equivalence with (4.5) follows from the fact that if the optimal decomposition of $Y = \sum_{k=1}^{r} \theta_k v_k w_k^H$ involves the term $v_k w_k^H$, then the norms $\|E_1 v_k\|$ and $\|E_2 w_k\|$ will be either zero

or one. (If $0 < \|E_1 v_k\| < 1$ we can decrease θ_k by scaling v_k until $\|E_1 v_k\| = 1$, and similarly for w_k.) The expression (4.8) shows that $h(Y)$ is the gauge of the convex hull of the set

$$\{vw^H \in \mathbf{C}^{n_1 \times n_2} \mid (v, w) \in \mathcal{A},\ \|E_1 v\| \leq 1,\ \|E_2 w\| \leq 1\}.$$

The SDP representation of h in (4.4) allows us to reformulate problems

$$\text{minimize} \quad f(Y) + h(Y), \tag{4.9}$$

where f is convex and h is the gauge (4.3)–(4.8), as a convex problem. Minimizing $f(Y) + h(Y)$ is equivalent to

$$\begin{array}{ll} \text{minimize} & f(Y) + \sum_{k=1}^{r} \|E_1 v_k\| \|E_2 w_k\| \\ \text{subject to} & Y = \sum_{k=1}^{r} v_k w_k^H \\ & (v_k, w_k) \in \mathcal{A}, \quad k = 1, \ldots, r. \end{array} \tag{4.10}$$

Alternatively, one can replace the second term in the objective with $\sum_k \|E_2 w_k\|$ and add constraints $\|E_1 v_k\| \leq 1$, as in

$$\begin{array}{ll} \text{minimize} & f(Y) + \sum_{k=1}^{r} \|E_2 w_k\| \\ \text{subject to} & Y = \sum_{k=1}^{r} v_k w_k^H \\ & (v_k, w_k) \in \mathcal{A}, \quad k = 1, \ldots, r \\ & \|E_1 v_k\| \leq 1, \quad k = 1, \ldots, r, \end{array} \tag{4.11}$$

or vice versa. When E_1 and E_2 are identity matrices, we can interpret $h(Y)$ as a convex penalty that promotes a structured low-rank property of Y. The outer products $v_k w_k^H$ are constrained by the set $\mathcal{A}$; the penalty term in the objective is the sum of the norms $\|v_k w_k^H\|_2 = \|v_k\| \|w_k\|$. The matrices E_1 and E_2 can be chosen to assign a different weight to different terms $v_k w_k^H$.

Problems (4.10) and (4.11) can be reformulated as

$$\begin{array}{ll} \text{minimize} & f(Y) + (\mathbf{tr}(E_1 V E_1^H) + \mathbf{tr}(E_2 W E_2^H))/2 \\ \text{subject to} & \Phi_{11} F X F^H + \Phi_{21} F X G^H + \Phi_{12} G X F^H + \Phi_{22} G X G^H = 0 \\ & \Psi_{11} F X F^H + \Psi_{21} F X G^H + \Psi_{12} G X F^H + \Psi_{22} G X G^H \preceq 0 \\ & X = \left[\begin{array}{cc} V & Y \\ Y^H & W \end{array} \right] \succeq 0. \end{array} \tag{4.12}$$

Example: column structure When $p_2 = 0$, the matrices F and G in (4.2) have the form $F = \left[\begin{array}{cc} F_1 & 0 \end{array} \right]$ and $G = \left[\begin{array}{cc} G_1 & 0 \end{array} \right]$. This means that $\mathcal{A} = \mathcal{A}_1 \times \mathbf{C}^{n_2}$ where

$$\mathcal{A}_1 = \{v \in \mathbf{C}^{n_1} \mid (\mu G_1 - \nu F_1)v = 0,\ (\mu, \nu) \in \mathcal{C}\}.$$

There are no restrictions on the w-component in $(v, w) \in \mathcal{A}$. Problem (4.10) simplifies:

$$\begin{array}{ll} \text{minimize} & f(Y) + \sum_{k=1}^{r} \|E_1 v_k\| \|E_2 w_k\| \\ \text{subject to} & Y = \sum_{k=1}^{r} v_k w_k^H \\ & v_k \in \mathcal{A}_1, \quad k = 1, \ldots, r. \end{array} \tag{4.13}$$

The equivalent semidefinite formulation (4.12) simplifies to

$$\begin{array}{ll} \text{minimize} & f(Y) + (\mathbf{tr}(E_1 V E_1^H) + \mathbf{tr}(E_2 W E_2^H))/2 \\ \text{subject to} & \Phi_{11} F_1 V F_1^H + \Phi_{21} F_1 V G_1^H + \Phi_{12} G_1 V F_1^H + \Phi_{22} G_1 V G_1^H = 0 \\ & \Psi_{11} F_1 V F_1^H + \Psi_{21} F_1 V G_1^H + \Psi_{12} G_1 V F_1^H + \Psi_{22} G_1 V G_1^H \preceq 0 \\ & \left[\begin{array}{cc} V & Y \\ Y^H & W \end{array} \right] \succeq 0. \end{array}$$

This SDP formulation of (4.13) (with $E_1 = I$, $E_2 = I$) was studied in [CV16].

As an example, we again consider the signal model (2.5). A natural idea for estimating the parameters ω_k and c_k is to solve a nonlinear least squares problem

$$\text{minimize} \quad \sum_{t=0}^{n-1} |y_{\mathrm{m}}(t) - \sum_{k=1}^{r} c_k e^{\mathrm{j}\omega_k t}|^2,$$

where $y_{\rm m}(t)$ is the observed signal. This problem is not convex and difficult to solve iteratively without a good starting point [SM97, page 148]. Instead of fixing r, we can also impose a penalty on $\sum_k |c_k|$, and consider the optimization problem

$$\begin{array}{ll} \text{minimize} & \gamma\|y - y_{\rm m}\|^2 + \sum_{k=1}^{r} |c_k| \\ \text{subject to} & y = \sum_{k=1}^{r} c_k \begin{bmatrix} 1 \\ e^{{\rm j}\omega_k} \\ \vdots \\ e^{{\rm j}(n-1)\omega_k} \end{bmatrix}. \end{array} \tag{4.14}$$

The optimization variables are y and the parameters c_k, ω_k, r in the decomposition of y. The vector $y_{\rm m}$ has elements $y_{\rm m}(0)$, $\ldots$, $y_{\rm m}(n-1)$. This is a special case of (4.11) with $f(y) = \gamma\|y - y_{\rm m}\|^2$, $n_1 = n$, $n_2 = 1$, $\Phi = \Phi_{\rm u}$, $\Psi = 0$, and

$$E_1 = (1/\sqrt{n})I, \qquad E_2 = 1, \qquad F_1 = \begin{bmatrix} 0 & I_{n_1-1} \end{bmatrix}, \qquad G_1 = \begin{bmatrix} I_{n_1-1} & 0 \end{bmatrix},$$

so that $\mathcal{A}_1$ is the set of all multiples of vectors $(1, e^{{\rm j}\omega}, \ldots, e^{{\rm j}(n-1)\omega})$. The problem is therefore equivalent to the convex problem

$$\begin{array}{ll} \text{minimize} & \gamma\|y - y_{\rm m}\|^2 + (\mathbf{tr}\, V)/(2n) + w/2 \\ \text{subject to} & \begin{bmatrix} V & y \\ y^H & w \end{bmatrix} \succeq 0 \\ & V \text{ is Toeplitz.} \end{array}$$

A related numerical example will be given in Section 4.3.1.

Example: joint column and row structure To illustrate the general problem (4.10), we consider a variation on the previous example. Suppose we arrange the observations in an $n \times m$ Hankel matrix

$$Y_{\rm m} = \begin{bmatrix} y_{\rm m}(0) & y_{\rm m}(1) & \cdots & y_{\rm m}(m-1) \\ y_{\rm m}(1) & y_{\rm m}(2) & \cdots & y_{\rm m}(m) \\ \vdots & \vdots & & \vdots \\ y_{\rm m}(n-1) & y_{\rm m}(n) & \cdots & y_{\rm m}(m+n-2) \end{bmatrix},$$

and we fit to this matrix a matrix Y with the same Hankel structure and with elements $y(t) = \sum_{k=1}^{r} c_k \exp(\mathrm{j}\omega_k t)$. We formulate the problem as

$$\begin{array}{ll} \text{minimize} & \gamma\|Y - Y_{\mathrm{m}}\|_F^2 + \displaystyle\sum_{k=1}^{r} |c_k| \\ \text{subject to} & Y = \displaystyle\sum_{k=1}^{r} c_k \begin{bmatrix} 1 \\ e^{\mathrm{j}\omega_k} \\ \vdots \\ e^{\mathrm{j}(n-1)\omega_k} \end{bmatrix} \begin{bmatrix} 1 \\ e^{-\mathrm{j}\omega_k} \\ \vdots \\ e^{-\mathrm{j}(m-1)\omega_k} \end{bmatrix}^H . \end{array} \tag{4.15}$$

This is an instance of (4.10) with $n_1 = n$, $n_2 = m$, $\Phi = \Phi_{\mathrm{u}}$, $\Psi = 0$, $E_1 = (1/\sqrt{n})I$, $E_2 = (1/\sqrt{m})I$, and

$$G_1 = \begin{bmatrix} I_{n-1} & 0 \end{bmatrix}, \quad F_1 = \begin{bmatrix} 0 & I_{n-1} \end{bmatrix}, \quad G_2 = \begin{bmatrix} 0 & I_{m-1} \end{bmatrix}, \quad F_2 = \begin{bmatrix} I_{m-1} & 0 \end{bmatrix}.$$

With these parameters, the set $\mathcal{A}$ contains the pairs (v, w) of the form

$$v = \alpha(1, e^{\mathrm{j}\omega}, \ldots, e^{\mathrm{j}(n-1)\omega}), \qquad w = \beta(1, e^{-\mathrm{j}\omega}, \ldots, e^{-\mathrm{j}(m-1)\omega}).$$

The convex formulation is

$$\begin{array}{ll} \text{minimize} & \gamma\|Y - Y_{\mathrm{m}}\|_F^2 + (\mathbf{tr}\, V)/(2n) + (\mathbf{tr}\, W)/(2m) \\ \text{subject to} & \begin{bmatrix} V & Y \\ Y^H & W \end{bmatrix} \succeq 0 \\ & \begin{bmatrix} F_1 & 0 \\ 0 & F_2 \end{bmatrix} \begin{bmatrix} V & Y \\ Y^H & W \end{bmatrix} \begin{bmatrix} F_1 & 0 \\ 0 & F_2 \end{bmatrix}^T = \begin{bmatrix} G_1 & 0 \\ 0 & G_2 \end{bmatrix} \begin{bmatrix} V & Y \\ Y^H & W \end{bmatrix} \begin{bmatrix} G_1 & 0 \\ 0 & G_2 \end{bmatrix}^T . \end{array}$$

A related example is discussed in Section 4.3.1.

4.2 Duality

In this section we derive the conjugates of the gauge functions defined in Section 4.1 and show that they can be interpreted as indicator functions of sets of bounded generalized polynomials. This gives a useful interpretation of the dual problems for (4.9).

We make the same assumptions as in the beginning of Section 3.4 to ensure that strong duality holds.

4.2.1 Conjugate of nonsymmetric matrix gauge

Consider the conjugate of the gauge defined in (4.3)–(4.8). We have

$$h^*(Z) = \sup_Y \left(\mathrm{Re}\,(\mathbf{tr}\, Z^H Y) - h(Y)\right)$$

where $h(Y)$ is the optimal value of (4.4). Therefore $h^*(Z)$ is the optimal value of the SDP

$$\begin{array}{ll} \text{maximize} & \frac{1}{2}\,\mathbf{tr}(\begin{bmatrix} -E_1^H E_1 & Z \\ Z^H & -E_2^H E_2 \end{bmatrix} X) \\ \text{subject to} & \Phi_{11} F X F^H + \Phi_{21} F X G^H + \Phi_{12} G X F^H + \Phi_{22} G X G^H = 0 \\ & \Psi_{11} F X F^H + \Psi_{21} F X G^H + \Psi_{12} G X F^H + \Psi_{22} G X G^H \preceq 0 \\ & X \succeq 0. \end{array} \tag{4.16}$$

The dual of this problem is

$$\begin{array}{ll} \text{minimize} & 0 \\ \text{subject to} & \begin{bmatrix} 0 & Z \\ Z^H & 0 \end{bmatrix} - \begin{bmatrix} F \\ G \end{bmatrix}^H (\Phi \otimes P + \Psi \otimes Q) \begin{bmatrix} F \\ G \end{bmatrix} \preceq \begin{bmatrix} E_1^H E_1 & 0 \\ 0 & E_2^H E_2 \end{bmatrix} \\ & Q \succeq 0. \end{array} \tag{4.17}$$

As in Section 3.4.1, with the assumptions we make, it follows from Appendix D that strong duality holds. Therefore $h^*(Z)$ is equal to the optimal value of (4.17), *i.e.*, zero if there exist P and Q that satisfy the constraints of this problem, and $+\infty$ otherwise. This will now be shown to be equivalent to

$$\begin{aligned} h^*(Z) &= \begin{cases} 0 & \mathrm{Re}\,(v^H Z w) \le (\|E_1 v\|^2 + \|E_2 w\|^2)/2 \quad \text{for all } (v,w) \in \mathcal{A} \\ +\infty & \text{otherwise} \end{cases} \\ &= \begin{cases} 0 & \mathrm{Re}\,(v^H Z w) \le \|E_1 v\| \|E_2 w\| \quad \text{for all } (v,w) \in \mathcal{A} \\ +\infty & \text{otherwise.} \end{cases} \end{aligned} \tag{4.18}$$

To see this, first assume P and Q are feasible in (4.17), and $a = (v, w) \in \mathcal{A}$ satisfies $(\mu G - \nu F)a = 0$ with $(\mu, \nu) \in \mathcal{C}$. Then

$$\begin{aligned} v^H Z w + w^H Z^H v - \|E_1 v\|^2 - \|E_2 w\|^2 &\leq \begin{bmatrix} Fa \\ Ga \end{bmatrix}^H (\Phi \otimes P + \Psi \otimes Q) \begin{bmatrix} Fa \\ Ga \end{bmatrix} \\ &= (y^H P y) q_\Phi(\mu, \nu) + (y^H Q y) q_\Psi(\mu, \nu) \\ &\leq 0, \end{aligned}$$

where $y = (1/\nu)Ga$ if $\nu \neq 0$ and $y = (1/\mu)Fa$ otherwise. Conversely, if problem (4.17) is infeasible, then (4.16) is unbounded above, so there exists a feasible X with positive objective value. If we decompose X as in Theorem 2.5, with $a_k = (v_k, w_k)$, we find that

$$\begin{aligned} 0 &< \mathbf{tr}\Big(\begin{bmatrix} -E_1^H E_1 & Z \\ Z^H & -E_2^H E_2 \end{bmatrix} \sum_{k=1}^r \begin{bmatrix} v_k \\ w_k \end{bmatrix} \begin{bmatrix} v_k \\ w_k \end{bmatrix}^H \Big) \\ &= \sum_{k=1}^r (v_k^H Z w_k + w_k^H Z^H v_k - \|E_1 v_k\|^2 - \|E_2 w_k\|^2) \end{aligned}$$

so at least one term in the sum is positive. The second expression for $h^*(Z)$ in (4.18) follows from the block diagonal structure of F and G. Following similar arguments as in Section 3.4.1, the expression (4.18) can also be derived directly from definition of the conjugate, (4.3), and (4.5), or via gauge duality.

4.2.2 Dual problem interpretation

The interpretation of the conjugate h^* can be applied to interpret the dual of (4.9),

$$\text{maximize} \quad -f^*(Z) - h^*(-Z).$$

Substituting the expression (4.17) for $h^*(-Z)$, one can write this as

$$\begin{array}{ll} \text{maximize} & -f^*(Z) \\ \text{subject to} & \begin{bmatrix} 0 & -Z \\ -Z^H & 0 \end{bmatrix} - \begin{bmatrix} F \\ G \end{bmatrix}^H (\Phi \otimes P + \Psi \otimes Q) \begin{bmatrix} F \\ G \end{bmatrix} \preceq \begin{bmatrix} E_1^H E_1 & 0 \\ 0 & E_2^H E_2 \end{bmatrix} \\ & Q \succeq 0, \end{array}$$

with variables Z, P, Q. Substituting the expression (4.18) for $h^*(-Z)$, we obtain

$$\begin{array}{ll} \text{maximize} & -f^*(Z) \\ \text{subject to} & \text{Re}\,(v^H Z w) \leq \|E_1 v\|\|E_2 w\| \quad \text{for all } (v,w) \in \mathcal{A}. \end{array}$$

As in Section 3.4.2, the primal-dual optimality conditions provide a useful set of complementary slackness relations between primal optimal Y and dual optimal Z. Specifically, the optimal Y can be decomposed as $Y = \sum_k v_k w_k^H$ with elements $(v_k, w_k) \in \mathcal{A}$ at which $\text{Re}\,(v_k^H Z w_k) = \|E_1 v_k\|\|E_2 w_k\|$.

Example Suppose $A \in \mathbf{C}^{n_s \times n_s}$, $B \in \mathbf{C}^{n_s \times m}$, $C \in \mathbf{C}^{l \times n_s}$, $D \in \mathbf{C}^{l \times m}$ are matrices in a state-space model with (A, B) controllable, and A has no eigenvalues that satisfy $(\lambda, 1) \in \mathcal{C}$. We take $p_1 = 0$, $n_1 = l$, $p_2 = n_s$, $n_2 = n_s + m$,

$$G_2 = \begin{bmatrix} I & 0 \end{bmatrix}, \qquad F_2 = \begin{bmatrix} A & B \end{bmatrix}, \qquad E_1 = I, \qquad E_2 = \begin{bmatrix} 0 & I \end{bmatrix}.$$

With this choice of parameters, $\mathcal{A} = \mathbf{C}^l \times \mathcal{A}_2$, where $\mathcal{A}_2$ contains the vectors

$$w = \begin{bmatrix} (\lambda I - A)^{-1} B u \\ u \end{bmatrix}$$

for all $u \in \mathbf{C}^m$ and all $(\lambda, 1) \in \mathcal{C}$, plus the vectors $(0, u)$ if $(1, 0) \in \mathcal{C}$. Since v is arbitrary and $E_1 = I$, the inequality in (4.18) reduces to $\|Zw\| \leq \|E_2 w\|$ for all $w \in \mathcal{A}_2$. With $Z = \begin{bmatrix} C & D \end{bmatrix}$, this is equivalent to a bound on the transfer function

$$\|D + C(\lambda I - A)^{-1} B\|_2 \leq 1 \quad \text{for all } (\lambda, 1) \in \mathcal{C}, \qquad \|D\|_2 \leq 1 \quad \text{if } (1, 0) \in \mathcal{C}.$$

4.3 Numerical examples

The formulations in Section 4.1 will now be illustrated with examples from signal processing. The optimization problems were solved with CVX [GB14].

4.3.1 Line spectrum estimation by penalty approximation

This example is a variation on problem (4.14). We take $n = 50$ consecutive measurements of the signal defined in (2.5). There are three sinusoids with frequencies and magnitudes shown

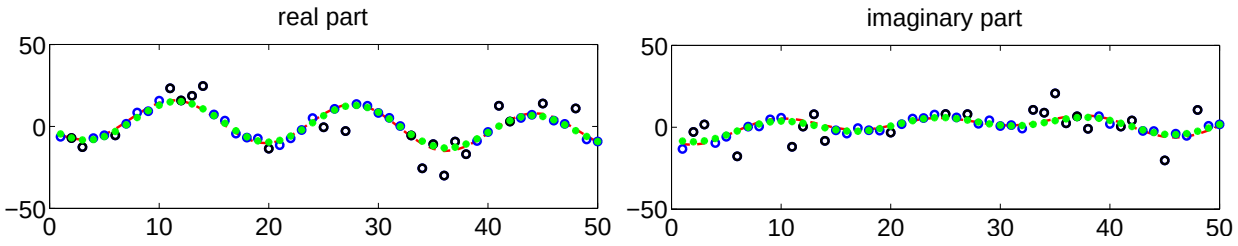

Figure 4.1: The data for the example in Section 4.3.1. The red dashed lines show the exact, noise-free signal. The circles show the signal corrupted by Gaussian white noise (in blue), plus a few larger errors in 20 positions (in black). The green dots show the recovered signal y from (4.19).

in Figure 4.2. The noise $v(t)$ is a superposition of white noise and a sparse corruption of 20 elements (see Figure 4.1). The model parameters are estimated by solving

$$\begin{array}{ll} \text{minimize} & \gamma \sum_{i=1}^{n} \phi(y_i - y_{\mathrm{m},i}) + \sum_{k=1}^{r} |c_k| \\ \text{subject to} & y = \sum_{k=1}^{r} c_k \begin{bmatrix} 1 \\ e^{\mathrm{j}\omega_k} \\ \vdots \\ e^{\mathrm{j}(n-1)\omega_k} \end{bmatrix} \\ & |\omega_k| \leq \omega_{\mathrm{c}}, \quad k = 1, \ldots, r, \end{array} \tag{4.19}$$

where ϕ is the Huber penalty, $\gamma = 0.071$, and $\omega_{\mathrm{c}} = \pi/6$. The variables are y and the parameters r, c_k, ω_k in the decomposition of y. The problem is equivalent to the convex problem

$$\begin{array}{ll} \text{minimize} & \gamma \sum_{i=1}^{n} \phi(y_i - y_{\mathrm{m},i}) + (\mathbf{tr}\, V)/(2n) + w/2 \\ \text{subject to} & \begin{bmatrix} V & y \\ y^H & w \end{bmatrix} \succeq 0 \\ & FVF^H - GVG^H = 0 \\ & -FVG^H - GVF^H + 2(\cos\omega_{\mathrm{c}})GVG^H \preceq 0 \end{array} \tag{4.20}$$

with F and G defined in (2.16). The variables are the n-vector y, the Hermitian $n \times n$ matrix V, and the scalar w. Figure 4.2 shows the result, and the estimates obtained from a

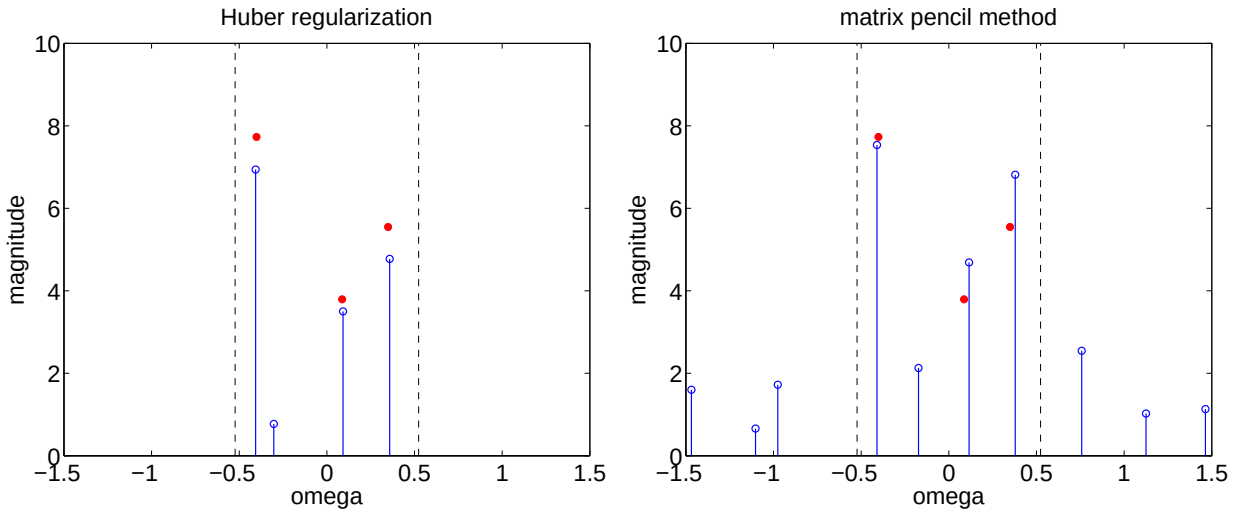

Figure 4.2: Line spectrum models estimated from the signal in Figure 4.1 by solving the optimization problem (4.19) (*Left*) and using the matrix pencil method (*Right*).

simple implementation (without filtering) of the matrix pencil method described in [HS90, SP95], where we form a 30×21 Hankel matrix Y_{m} from the measurements and compute the generalized eigenvalues of $\lambda Y_{\mathrm{m1}} - Y_{\mathrm{m2}}$ as estimates of $e^{\mathrm{j}\omega_k}$ (Y_{m1} and Y_{m2} represent the matrix Y_{m} with the last and the first column removed, respectively). The comparison illustrates the usefulness of incorporating the prior frequency constraint and the Huber penalty in (4.19).

It is interesting to note that problem (4.19) can be equivalently formulated as

$$\begin{aligned} \text{minimize} \quad & \gamma \sum_{i=1}^{n} \phi(y_i - y_{\mathrm{m},i}) + \sum_{k=1}^{r} |c_k| \\ \text{subject to} \quad & \begin{bmatrix} y_1 & y_2 & \cdots & y_{n_2} \\ y_2 & y_3 & \cdots & y_{n_2+1} \\ \vdots & \vdots & & \vdots \\ y_{n_1} & y_{n_1+1} & \cdots & y_{n_1+n_2-1} \end{bmatrix} = \sum_{k=1}^{r} c_k \begin{bmatrix} 1 \\ e^{\mathrm{j}\omega_k} \\ \vdots \\ e^{\mathrm{j}(n_1-1)\omega_k} \end{bmatrix} \begin{bmatrix} 1 \\ e^{-\mathrm{j}\omega_k} \\ \vdots \\ e^{-\mathrm{j}(n_2-1)\omega_k} \end{bmatrix}^H \\ & |\omega_k| \le \omega_{\mathrm{c}}, \quad k = 1, \ldots, r, \end{aligned} \tag{4.21}$$

where $n_1 + n_2 - 1 = n$. This problem is equivalent to

$$\begin{array}{ll} \text{minimize} & \gamma \sum_{i=1}^{m} \phi(y_i - y_{\mathrm{m},i}) + (\mathbf{tr}\, V)/(2n_1) + (\mathbf{tr}\, W)/(2n_2) \\ \text{subject to} & X = \left[\begin{array}{cc} V & Y \\ Y^H & W \end{array} \right] \succeq 0 \\ & FXF^T = GXG^T \\ & -FXG^T - GXF^T + 2(\cos\omega_{\mathrm{c}})GXG^T \preceq 0 \end{array} \tag{4.22}$$

where G and F are block diagonal with blocks

$$G_1 = \left[\begin{array}{cc} I_{n_1-1} & 0 \end{array} \right], \qquad F_1 = \left[\begin{array}{cc} 0 & I_{n_1-1} \end{array} \right], \qquad G_2 = \left[\begin{array}{cc} 0 & I_{n_2-1} \end{array} \right], \qquad F_2 = \left[\begin{array}{cc} I_{n_2-1} & 0 \end{array} \right].$$

The variables in (4.22) are the matrices V, Y, W. The elements y_i in the objective are the elements in the first row and last column of the matrix variable Y. The two SDPs (4.20) and (4.22) give the same result y, but may have different numerical properties (in terms of accuracy or complexity).

4.3.2 Direction of arrival estimation

This example illustrates the use of frequency interval constraints in direction-of-arrival (DOA) estimation. We consider a uniform linear array of n sensors. The signal arriving at the array is a superposition of a small number of planar waves arriving from different directions in $[-\pi/2, \pi/2]$. We take $2d/\lambda_{\mathrm{c}} = 1$, where d is the distance between the sensors and λ_{c} the signal wavelength [SM97, §6.2]. When all sensor measurements are available, the directions of arrival can be estimated by classical methods, such as MUSIC and ESPRIT [SM97,Sch86,PRK86]; see Appendix A for a brief overview. In this example, however, we assume that only a randomly selected subset of sensors is used. Moreover, the sensors are not omnidirectional. They are randomly partitioned in two groups of equal size, measuring two different ranges of directions. To simplify notation, we will assume the measurements

are noise-free. Consider the DOA estimation problem:

$$\begin{array}{ll} \text{minimize} & \sum_{j=1}^{3}\sum_{k=1}^{r_j}|x_{jk}| \\ \text{subject to} & y_j = \sum_{k=1}^{r_j} x_{jk}\begin{bmatrix} 1 \\ e^{\mathrm{j}\pi\sin\theta_{jk}} \\ \vdots \\ e^{\mathrm{j}(n-1)\pi\sin\theta_{jk}} \end{bmatrix} \\ & \theta_{jk}\in\Theta_j, \quad k=1,\ldots,r_j, \quad j=1,2,3 \\ & (y_1+y_2)_{I_1} = b_1, \quad (y_2+y_3)_{I_2} = b_2, \end{array} \tag{4.23}$$

with variables y_j and their decomposition parameters r_j, x_{jk} and θ_{jk}. The vectors b_1 and b_2 contain the outputs of two subsets of the elements in a linear array of n non-isotropic antennas. Elements in the first group, indexed by the index set I_1, measure input signals arriving from angles in $\Theta_1\cup\Theta_2 = [-\pi/2,-\pi/6]\cup[-\pi/6,\pi/6]$. Elements in the second group, indexed by the index set I_1, measure input signals arriving from $\Theta_2\cup\Theta_3 = [-\pi/6,\pi/6]\cup[\pi/6,\pi/2]$. The problem can be equivalently cast as the SDP

$$\begin{array}{ll} \text{minimize} & \sum_{j=1}^{3}(\mathbf{tr}(V_j)+w_j)/2 \\ \text{subject to} & \begin{bmatrix} V_j & y_j \\ y_j^H & w_j \end{bmatrix} \succeq 0, \quad j=1,2,3 \\ & V_1,\, V_2,\, V_3 \text{ are Toeplitz} \\ & -e^{-\mathrm{j}\alpha_j}FV_jG^T - e^{\mathrm{j}\alpha_j}GV_jF^T + 2(\cos\beta_j)GV_jG^T \preceq 0, \quad j=1,2,3 \\ & (y_1+y_2)_{I_1} = b_1, \quad (y_2+y_3)_{I_2} = b_2 \end{array} \tag{4.24}$$

with variables $V_j\in\mathbf{H}^n$, $y_j\in\mathbf{C}^n$, $w_j\in\mathbf{R}$, for $j=1,2,3$. The matrices F, G are defined in (2.16). The three intervals $[\alpha_j-\beta_j,\alpha_j+\beta_j]$ are the images of the intervals Θ_j after the transformation $\omega=\pi\sin\theta$, which represents the spatial frequency associated with a direction of arrival θ.

Figure 4.3 shows the results of an instance with $n=500$ elements in the array, but using only a total of 40 randomly selected measurements ($|I_1|=|I_2|=20$). The red dots show the angles and magnitudes of 7 signals used to compute the measurement vectors b_1, b_2. The

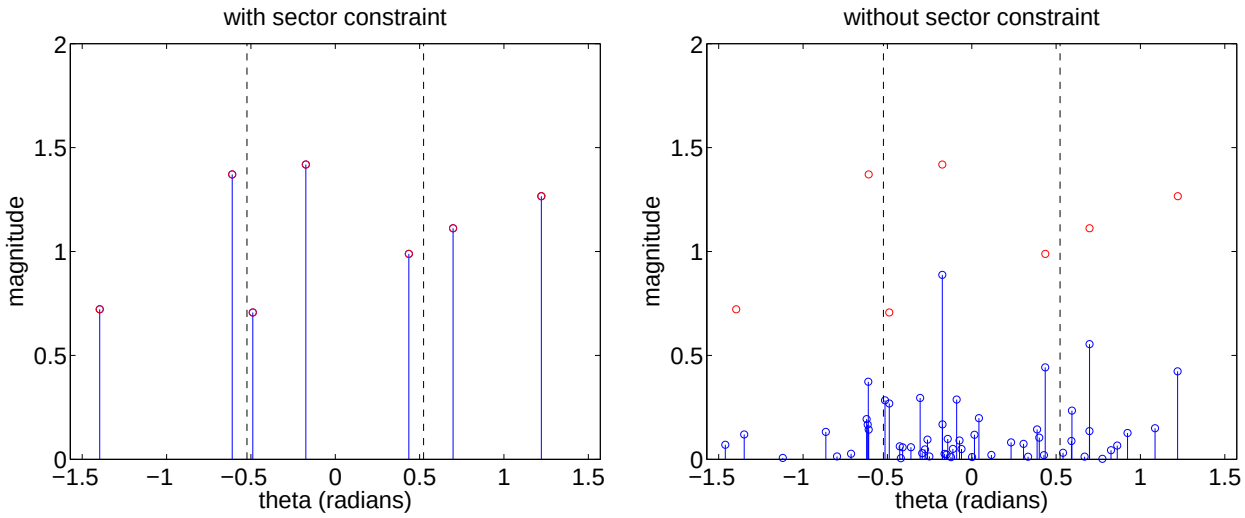

Figure 4.3: Direction-of-arrival estimation with (*Left*) and without (*Right*) interval constraints (Section 4.3.2).

estimated angles and coefficients $|c_{jk}|$ are shown with blue lines. The right-hand plot shows the solution if we omit the interval constraints in (4.23).

Figure 4.4 shows the success rate as a function of the number $|I_1| + |I_2|$ of available measurements, for an example with $n = 50$ elements, and the same angles as in Figure 4.3. Each data point is the average of 100 trials, with different, randomly generated coefficients, and different random selections of the two sensor groups. We observe that solving the optimization problem with the interval constraints has a higher rate of exact recovery. For example, with 30 available measurements, including the interval constraints gave the exact answer in all instances, whereas the method without the interval constraints was successful in only about 25% of the instances.

4.3.3 Direction of arrival from multiple measurement vectors

This example demonstrates the advantage of using multiple measurement vectors (or snapshots) in DOA estimation, as pointed out in [LC16, YX16]. Suppose we have K omnidirectional sensors placed at randomly chosen positions of a linear grid of length n. The measurements of the K sensors at one time instance form one measurement vector. We

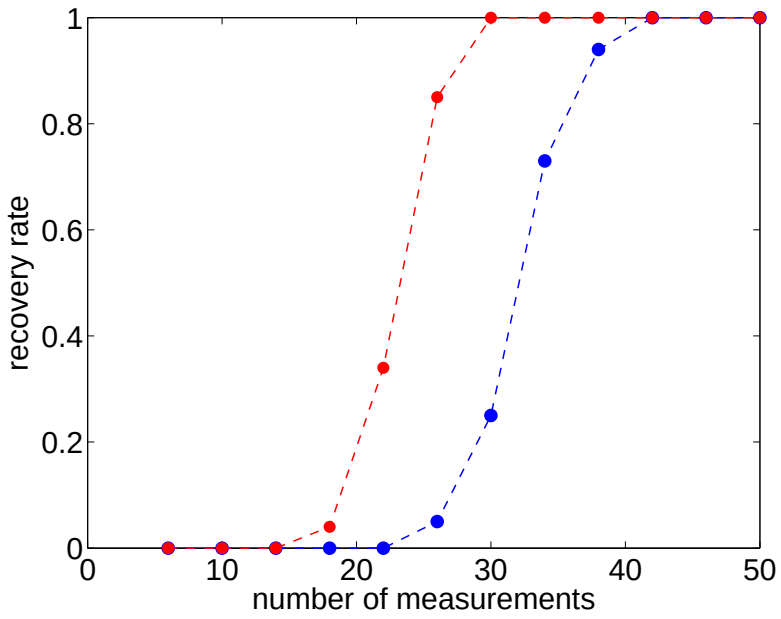

Figure 4.4: Comparison of recovery rate for different number of available measurements with interval constraints (red) and without (blue), in the example of Section 4.3.2.

collect m of these measurement vectors, at m different times, and assume that the directions of arrival and the source magnitudes remain constant while the measurements are taken. The problem is formulated as

$$\begin{array}{ll} \text{minimize} & \displaystyle\sum_{k=1}^{r} \|c_k\| \\ \text{subject to} & \displaystyle Y = \sum_{k=1}^{r} \begin{bmatrix} 1 \\ e^{\mathrm{j}\alpha \sin\theta_k} \\ \vdots \\ e^{\mathrm{j}(n-1)\alpha \sin\theta_k} \end{bmatrix} c_k^H \\ & |\theta_k| \leq \theta_{\mathrm{c}}, \quad k = 1, \ldots, r \\ & Y_I = B, \end{array} \tag{4.25}$$

with variables $Y \in \mathbf{C}^{n\times m}$, $c_k \in \mathbf{C}^m$, θ_k, and r. Here $\alpha = 2\pi d/\lambda_{\mathrm{c}}$, where d is the distance between the grid points and λ_{c} is the signal wavelength, and θ_{c} is a given cutoff angle. The columns of the $K \times m$ matrix B are the measurement vectors. The matrix Y_I is the submatrix

of Y containing the K rows indexed by $I \subset \{1, \ldots, n\}$. The convex formulation is

$$\begin{array}{ll} \text{minimize} & (\mathbf{tr}\, V)/(2n) + (\mathbf{tr}\, W)/2 \\ \text{subject to} & \begin{bmatrix} V & Y \\ Y^H & W \end{bmatrix} \succeq 0 \\ & FVF^H - GVG^H = 0 \\ & -FVG^H - GVF^H + 2(\cos\omega_{\mathrm{c}})GVG^H \preceq 0 \\ & Y_I = B \end{array}$$

with F and G defined in (2.16) and $\omega_{\mathrm{c}} = \alpha \sin\theta_{\mathrm{c}}$.

Figure 4.5 shows an instance with $n = 30$, $K = 7$, $\alpha = 2$, and $\theta_{\mathrm{c}} = \pi/4$. We show the solution for $m = 1$, $m = 15$ and $m = 30$. The blue lines show the values of θ_k and $\|c_k\|/\sqrt{m}$ computed by solving problem (4.25). In an experiment of 150 trials with randomly chosen index sets I, the signal was recovered accurately in 67.3% of the trials for $m = 15$ and 85.3% for $m = 30$.

4.3.4 Structured matrix decomposition

We generate a 30×30 matrix $C = AB + N$ as a product of a 30×3 matrix A with entries $A_{ij} = \exp(\mathrm{j}(i-1)\omega_j)$, for given values of ω_1, ω_2, ω_3, and a randomly generated complex 3×30 matrix B with entries from a normal distribution, plus a Gaussian noise matrix N. The goal is to estimate the parameters ω_j and the matrix B from the noisy measurements C.

We compare two methods. In the first method we assume we are given a narrow interval that includes the parameters ω_j. We consider the optimization problem

$$\begin{array}{ll} \text{minimize} & \gamma\|Y - C\|_F + \sum_{k=1}^{r} \|x_k\| \\ \text{subject to} & Y = \sum_{k=1}^{r} \begin{bmatrix} 1 \\ e^{\mathrm{j}\omega_k} \\ \vdots \\ e^{\mathrm{j}(n-1)\omega_k} \end{bmatrix} x_k^H \\ & |\omega_k - \alpha| \leq \beta, \quad k = 1, \ldots, r \end{array} \tag{4.26}$$

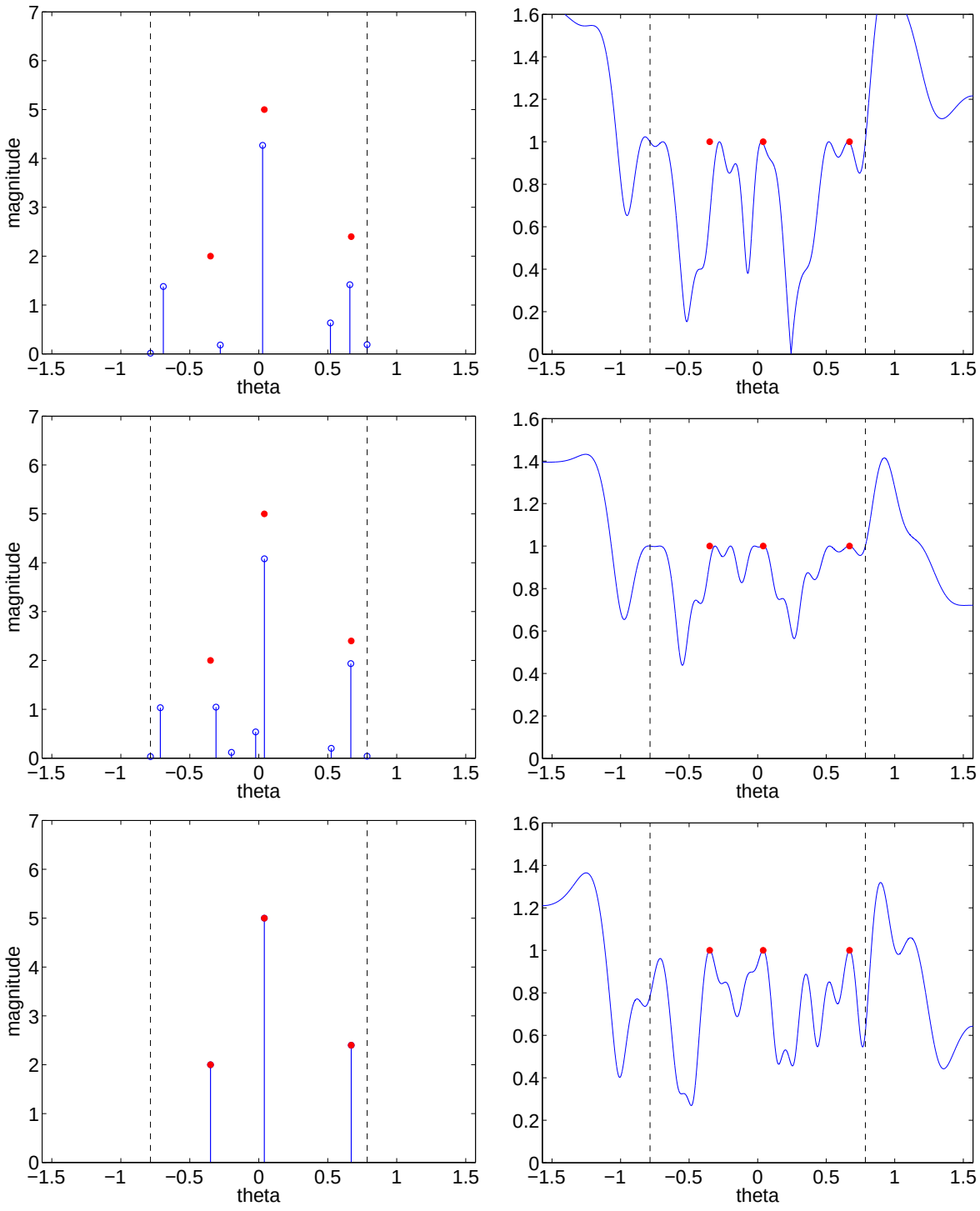

Figure 4.5: The results with 1 (*Top*), 15 (*Middle*) and 30 (*Bottom*) measurement vector(s) in the DOA estimation problem of Section 4.3.3. The figures on the right show the magnitude of the trigonometric polynomials obtained from the dual solution. The red dots show the true directions of arrival (and magnitudes).

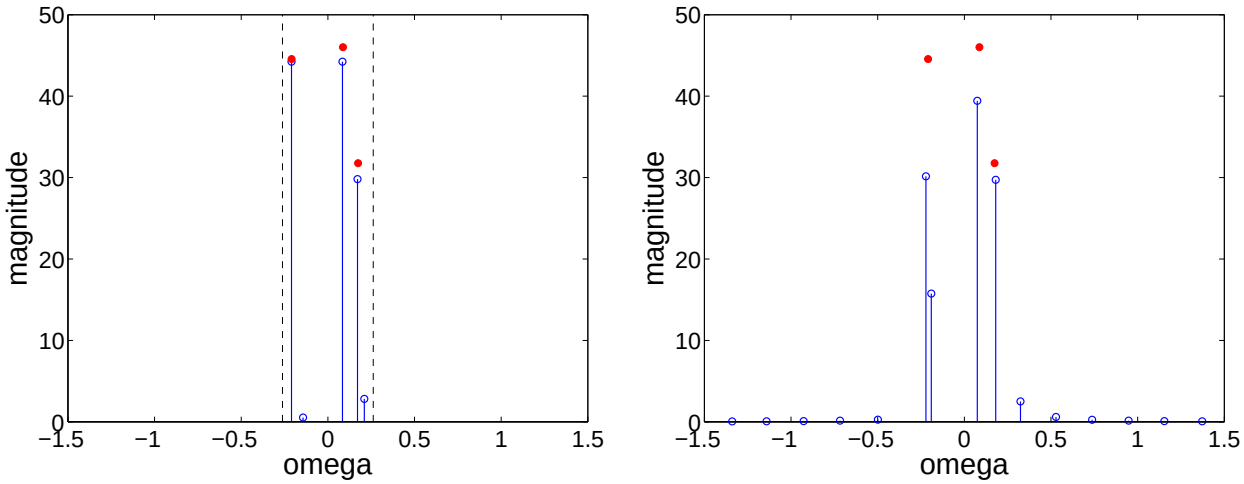

Figure 4.6: Structured matrix decomposition of a matrix with rank 3 plus a Gaussian noise matrix, with (*Left*) and without (*Right*) interval constraint (Section 4.3.4).

with γ a positive parameter. In the example, we use $\alpha = 0$, $\beta = \pi/12$. The problem can be converted to the SDP

$$\begin{array}{ll} \text{minimize} & \gamma\|Y - C\|_F + (\mathbf{tr}(V) + \mathbf{tr}(W))/(2\sqrt{n}) \\ \text{subject to} & \begin{bmatrix} V & Y \\ Y^H & W \end{bmatrix} \succeq 0 \\ & V \text{ is Toeplitz} \\ & -FVG^H - GVF^H + 2\cos\beta\, GVG^H \preceq 0 \end{array}$$

with F and G defined in (2.16). In the second method, we omit the third constraint in the SDP, *i.e.*, solve (4.26) without the interval constraint. Figure 4.6 shows the two solutions, for independently tuned values of γ, with the estimates of ω_k on the horizontal axis, and the norms of the vectors x_k on the vertical axis. As can be seen, adding the interval constraints allowed the method to identify the parameters ω_k, and thus $\|x_k\|$, more accurately.

CHAPTER 5

Itakura–Saito Generalized Distance

The second subject of the thesis is of the algorithmic aspect. In this chapter, we introduce a generalized distance defined for the cone of nonnegative trigonometric polynomials and discuss the properties and calculation techniques regarding this distance. In Chapter 6, we will explain how the distance is used in first-order proximal methods and see the techniques discussed in this chapter in action. Chapters 5 and 6 are adapted from [CV18].

As mentioned in the previous chapters, optimization problems over the cone of nonnegative trigonometric polynomials or its dual cone, the cone of positive semidefinite Toeplitz matrices, are common in signal processing and system identification [MAK95, WBV98, SMM00, DTS01, DLS02, AV02, Dum07, GL08, ALH17]. Recent examples include superresolution techniques for spectrum estimation and gridless compressed sensing [CF13, TBS13, CF14]. If the cost function admits an efficient semidefinite representation, such problems can be solved by general-purpose interior-point solvers for semidefinite optimization. Special-purpose interior-point solvers and first-order splitting methods such as alternating direction method of multipliers (ADMM) have also been explored [AV00, AV02, Hac03, RV06, LV07, HV14, BTR13]. However, these methods have at best a per-iteration-complexity that is cubic in the degree of the polynomial. In particular, at each iteration, the first-order splitting methods involve a Euclidean projection on the positive semidefinite cone, which takes an eigenvalue decomposition. Therefore, in order to solve these problems more efficiently, we explore generalized first-order methods that are based on a generalized distance, the *Itakura–Saito distance*.

We consider real coefficients in the following for the simplicity of notation. The results can be extended to complex coefficients by making natural changes. To be specific, let K be

the cone of nonnegative trigonometric polynomials of degree p or less:

$$K = \{x = (x_0, \ldots, x_p) \in \mathbf{R}^{p+1} \mid F_x(e^{\mathrm{j}\omega}) \geq 0 \ \ \forall \omega\} \tag{5.1}$$

where $F_x(z)$ is the Laurent polynomial

$$F_x(z) = x_p z^p + \cdots + x_1 z + x_0 + x_1 z^{-1} + \cdots + x_p z^{-p}. \tag{5.2}$$

The convex cone K can be expressed as the image of the positive semidefinite cone under a linear transformation,

$$K = \{\mathcal{D}(X) \mid X \in \mathbf{S}_+^{p+1}\} \tag{5.3}$$

where $\mathbf{S}_+^{p+1}$ is the set of symmetric positive semidefinite matrices of order $p+1$, and the linear mapping $\mathcal{D}$ maps X to the vector of its diagonal sums, *i.e.*,

$$\mathcal{D}\left(\begin{bmatrix} X_{00} & X_{01} & \cdots & X_{0p} \\ X_{10} & X_{11} & \cdots & X_{1p} \\ \vdots & \vdots & \ddots & \vdots \\ X_{p0} & X_{p1} & \cdots & X_{pp} \end{bmatrix}\right) = \left(\sum_{i=0}^{p} X_{ii}, \sum_{i=0}^{p-1} X_{i+1,i}, \ldots, X_{p-1,0} + X_{p1}, X_{p0}\right) \tag{5.4}$$

(see [DTS01, AV02, Dum07] and Theorem 2.4).

The Itakura–Saito distance is defined as

$$d(x, v) = \frac{1}{2\pi} \int_0^{2\pi} \left(\frac{F_x(e^{\mathrm{j}\omega})}{F_v(e^{\mathrm{j}\omega})} - \log \frac{F_x(e^{\mathrm{j}\omega})}{F_v(e^{\mathrm{j}\omega})} - 1\right) d\omega, \tag{5.5}$$

with domain $\mathbf{dom}\, d = (K \setminus \{0\}) \times (\operatorname{int} K)$. It is the Bregman distance

$$d_\phi(x, v) = \phi(x) - \phi(v) - \langle \nabla\phi(v), x - v\rangle,$$

associated with the negative entropy function

$$\phi(x) = -\frac{1}{2\pi} \int_0^{2\pi} \log F_x(e^{\mathrm{j}\omega})\, d\omega. \tag{5.6}$$

The chapter is organized as follows. In Section 5.1 we review some background material from statistical signal processing and numerical linear algebra related to positive definite Toeplitz systems. The negative entropy function (5.6) and its conjugate are discussed in Section 5.2, and the associated Bregman distance (5.5) in Section 5.3.

5.1 Forward and backward Levinson–Durbin algorithm

In this section we review some classical results and algorithms from statistical signal processing. We denote by $\mathcal{T}(y)$, where $y = (y_0, y_1, \ldots, y_p)$, the symmetric Toeplitz matrix

$$\mathcal{T}(y) = \begin{bmatrix} y_0 & y_1 & \cdots & y_p \\ y_1 & y_0 & \cdots & y_{p-1} \\ \vdots & \vdots & \ddots & \vdots \\ y_p & y_{p-1} & \cdots & y_0 \end{bmatrix}, \tag{5.7}$$

and by $\mathcal{J}(b)$, where $b = (b_0, b_1, \ldots, b_p)$, the matrix

$$\mathcal{J}(b) = \begin{bmatrix} b_0/2 & 0 & \cdots & 0 & 0 \\ b_1/2 & b_0 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ b_{p-1}/2 & b_{p-2} & \cdots & b_0 & 0 \\ b_p/2 & b_{p-1} & \cdots & b_1 & b_0 \end{bmatrix} + \begin{bmatrix} b_0/2 & b_1 & \cdots & b_{p-1} & b_p \\ b_1/2 & b_2 & \cdots & b_p & 0 \\ \vdots & \vdots & \iddots & \vdots & \vdots \\ b_{p-1}/2 & b_p & \cdots & 0 & 0 \\ b_p/2 & 0 & \cdots & 0 & 0 \end{bmatrix}. \tag{5.8}$$

This matrix is known as the *Jury matrix* [DM90]. Note that $\mathcal{T}(y)b = \mathcal{J}(b)y$ for all y and b.

The results in this section may be summarized as follows. Suppose $\mathcal{T}(y)$ is positive definite, and let $b = (b_0, \ldots, b_p)$ be the solution of the linear equation

$$\mathcal{T}(y)b = e, \tag{5.9}$$

where $e = (1, 0, \ldots, 0)$. Then the polynomial $b_0 z^p + b_1 z^{p-1} + \cdots + b_p$ is stable (has all its zeros inside the unit circle). The classical algorithm for solving this equation is the Levinson–Durbin algorithm, which computes a Cholesky factorization of $\mathcal{T}(y)^{-1}$ in order p^2 operations. Several more recent algorithms for positive definite Toeplitz systems are even faster, with an order $p(\log p)^2$ complexity [BA80, BGY80, Hoo87, AG88, Ste03].

Second, suppose the polynomial $\mathcal{B}(z) = b_0 z^p + b_1 z^{p-1} + \cdots + b_p$ is stable. Then the Jury matrix $\mathcal{J}(b)$ is nonsingular, and the unique solution y of the equation

$$\mathcal{J}(b)y = e \tag{5.10}$$

defines a positive definite Toeplitz matrix $\mathcal{T}(y)$. This equation can be solved by a recursive algorithm that is essentially the Jury stability test applied to $\mathcal{B}(z)$. The algorithm computes a factorization of $\mathcal{J}(b)$ and can be interpreted as a backward Levinson–Durbin algorithm.

More details on these algorithms may be found in textbooks on statistical signal processing, for example, [Sch91, §10] or [PM96, §11].

5.1.1 Levinson–Durbin algorithm

The Levinson–Durbin algorithm [GV96, §4.7] is a fast algorithm for the Cholesky factorization of the inverse of a positive definite Toeplitz matrix $\mathcal{T}(y)$, given its first column $y = (y_0, \ldots, y_p)$. The computed factorization is

$$U\mathcal{T}(y)U^T = \mathbf{diag}(\sigma_p^2, \ldots, \sigma_0^2), \tag{5.11}$$

where $0 < \sigma_p \leq \cdots \leq \sigma_0$ and U is a unit upper triangular matrix

$$U = \begin{bmatrix} 1 & a_{p1} & a_{p2} & \cdots & a_{p,p-1} & a_{pp} \\ 0 & 1 & a_{p-1,1} & \cdots & a_{p-1,p-2} & a_{p-1,p-1} \\ 0 & 0 & 1 & \cdots & a_{p-2,p-3} & a_{p-2,p-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & a_{11} \\ 0 & 0 & 0 & \cdots & 0 & 1 \end{bmatrix}. \tag{5.12}$$

For theoretical purposes later in the thesis, it is useful to note that the Levinson–Durbin algorithm can be extended to Toeplitz matrices that are positive semidefinite, but not positive definite. In that case we can still compute a factorization of the form (5.11), where $0 \leq \sigma_p \leq \cdots \leq \sigma_0$.

Algorithm 5.1. *Levinson–Durbin algorithm.*

Input. The coefficients $y_0, \ldots, y_p$ of a Toeplitz matrix $\mathcal{T}(y)$.

Output. If $\mathcal{T}(y)$ is positive semidefinite, a matrix U and coefficients $\sigma_0, \ldots, \sigma_p$ that satisfy (5.11), (5.12), and $0 \leq \sigma_p \leq \cdots \leq \sigma_0$.

Algorithm. Define $\sigma_0 = \sqrt{y_0}$. For $k = 0, \ldots, p-1$, execute the following steps.

- If $\sigma_k \neq 0$, define

$$\kappa_k = -\frac{y_{k+1} + y_k a_{k1} + \cdots + y_1 a_{kk}}{\sigma_k^2}. \tag{5.13}$$

  Otherwise, set $\kappa_k = 0$.

- If $|\kappa_k| > 1$, terminate. The matrix $\mathcal{T}(y)$ is not positive semidefinite.

- Compute $\sigma_{k+1} = \sigma_k(1 - \kappa_k^2)^{1/2}$ and

$$\begin{bmatrix} a_{k+1,1} \\ \vdots \\ a_{k+1,k} \\ a_{k+1,k+1} \end{bmatrix} = \begin{bmatrix} a_{k1} \\ \vdots \\ a_{kk} \\ 0 \end{bmatrix} + \kappa_k \begin{bmatrix} a_{kk} \\ \vdots \\ a_{k1} \\ 1 \end{bmatrix}. \tag{5.14}$$

The coefficients κ_k are known as the *reflection coefficients*. The algorithm has complexity $O(p^2)$.

The update (5.14) can be written concisely using polynomial notation, if we define the polynomials

$$\mathcal{A}_0(z) = 1, \qquad \mathcal{A}_k(z) = z^k + a_{k1} z^{k-1} + \cdots + a_{k,k-1} z + a_{kk}, \quad k = 1, \ldots, p, \tag{5.15}$$

and the reversed polynomials $\hat{\mathcal{A}}_k(z) = z^k \mathcal{A}_k(1/z) = a_{kk} z^k + \cdots + a_{k1} z + 1$. With this notation, the update (5.14) can be written

$$\begin{bmatrix} \mathcal{A}_{k+1}(z) \\ \hat{\mathcal{A}}_{k+1}(z) \end{bmatrix} = \begin{bmatrix} 1 & \kappa_k \\ \kappa_k & 1 \end{bmatrix} \begin{bmatrix} z\mathcal{A}_k(z) \\ \hat{\mathcal{A}}_k(z) \end{bmatrix}, \quad k = 0, \ldots, p-1, \tag{5.16}$$

starting at $\mathcal{A}_0(z) = 1$.

Another useful form is in terms of the $(p+1)$-vectors $a^{(k)} = (1, a_{k1}, \ldots, a_{kk}, 0, \ldots, 0)$. The recursion (5.14) can be written as a sequence of linear transformations

$$a^{(k+1)} = H_k a^{(k)}, \quad k = 0, \ldots, p-1, \tag{5.17}$$

starting at $a^{(0)} = (1, 0, \ldots, 0)$, where

$$H_k = \begin{bmatrix} I_{k+2} + \kappa_k J_{k+2} & 0 \\ 0 & I_{p-k-1} \end{bmatrix} \tag{5.18}$$

and J_r is the $r \times r$ reverser matrix

$$J_r = \begin{bmatrix} 0 & \cdots & 0 & 1 \\ 0 & \cdots & 1 & 0 \\ \vdots & \iddots & \vdots & \vdots \\ 1 & \cdots & 0 & 0 \end{bmatrix}.$$

We mention two properties of the factorization (5.11) that will be useful later in the thesis. The first column of the equation $\mathcal{T}(y)U^T = U^{-1}\,\mathbf{diag}(\sigma_p^2, \ldots, \sigma_0^2)$ is

$$\begin{bmatrix} y_0 & y_1 & \cdots & y_p \\ y_1 & y_0 & \cdots & y_{p-1} \\ \vdots & \vdots & \ddots & \vdots \\ y_p & y_{p-1} & \cdots & y_0 \end{bmatrix} \begin{bmatrix} 1 \\ a_{p1} \\ \vdots \\ a_{pp} \end{bmatrix} = \begin{bmatrix} \sigma_p^2 \\ 0 \\ \vdots \\ 0 \end{bmatrix}. \tag{5.19}$$

This is known as the *Yule–Walker equation*. The Levinson–Durbin algorithm solves the Yule–Walker equation (*i.e.*, computes $a_{p1}, \ldots, a_{pp}, \sigma_p^2$, given $y_0, \ldots, y_p$) in $O(p^2)$ operations. The solution is unique if the $p \times p$ Toeplitz matrix with first column $y_0, \ldots, y_{p-1}$ is positive definite. If $\mathcal{T}(y)$ is positive definite, then $\sigma_p^2 \neq 0$, and $b = \sigma_p^{-2}(1, a_{p1}, \ldots, a_{pp})$ is the solution of (5.9).

Second, if $\mathcal{T}(y)$ is positive definite, then the polynomials $\mathcal{A}_k(z)$ defined in (5.15) are stable. In particular, the polynomial

$$b_0 z^p + b_1 z^{p-1} + \cdots + b_p = \frac{1}{\sigma_p^2}\mathcal{A}_p(z)$$

is stable. To show this, one can note that if $\mathcal{A}_k(z)$ is a stable polynomial, then $|\mathcal{A}_k(z)| \geq |\hat{\mathcal{A}}_k(z)|$ holds for $|z| \geq 1$. (This is easily seen by factoring $\mathcal{A}_k(z)$ and using the fact that $|z - a|/|1 - \bar{a}z| \geq 1$ if $|a| < 1$ and $|z| \geq 1$.) Therefore, if $\mathcal{A}_k(z)$ is stable ($\mathcal{A}_k(z) \neq 0$ for $|z| \geq 1$) and $|\kappa_k| < 1$, then the polynomial $\mathcal{A}_{k+1}(z)$ defined in (5.16) is nonzero for $|z| \geq 1$. Since $\mathcal{A}_0(z) = 1$, stability of the polynomials $\mathcal{A}_k(z)$ follows by induction.

5.1.2 Jury stability test

In this section we discuss an algorithm that can be seen as the Levinson–Durbin algorithm run backwards. The algorithm is equivalent to the Jury test for determining the stability of a real polynomial. The connection between the Jury test and the Levinson–Durbin algorithm was noted by Vieira and Kailath [VK77]. In Section 5.1.3 we will see that the algorithm also computes a factorization of the Jury matrix defined in (5.8).

We use the same notation (5.12) as in the previous section.

Algorithm 5.2. *Jury stability test.*

Input. The coefficients $b_0, \ldots, b_p$ of a polynomial $\mathcal{B}(z) = b_0 z^p + \cdots + b_p$, with $b_0 > 0$.

Output. If $\mathcal{B}(z)$ is stable, a unit upper triangular matrix U with first row $(b_0, \ldots, b_p)/b_0$ and coefficients $0 < \sigma_p \leq \cdots \leq \sigma_0$ such that $U^{-1}\,\mathbf{diag}(\sigma_p^2, \ldots, \sigma_0^2)U^{-T}$ is Toeplitz.

Algorithm. Define

$$(a_{p1}, \ldots, a_{pp}) = (b_1/b_0, \ldots, b_p/b_0), \qquad \sigma_p = 1/\sqrt{b_0}.$$

For $k = p-1, p-2, \ldots, 0$, execute the following steps.

- Define $\kappa_k = a_{k+1,k+1}$. If $|\kappa_k| \geq 1$, terminate. The polynomial $\mathcal{B}(z)$ is not stable.
- Otherwise, compute $\sigma_k = \sigma_{k+1}/(1-\kappa_k^2)^{1/2}$ and

$$\begin{bmatrix} a_{k1} \\ \vdots \\ a_{kk} \end{bmatrix} = \frac{1}{1-\kappa_k^2} \begin{bmatrix} a_{k+1,1} \\ \vdots \\ a_{k+1,k} \end{bmatrix} - \frac{\kappa_k}{1-\kappa_k^2} \begin{bmatrix} a_{k+1,k} \\ \vdots \\ a_{k+1,1} \end{bmatrix}. \tag{5.20}$$

The Jury stability test is successful if $|\kappa_k| < 1$ for $k = p-1, \ldots, 0$, or, equivalently, $\sigma_0 \geq \sigma_1 \geq \cdots \geq \sigma_p > 0$. The complexity of this algorithm is $O(p^2)$.

The relation with the Levinson–Durbin algorithm is clear if we write the update (5.20) as a recursion for the polynomials $\mathcal{A}_k(z)$ (with coefficients a_{ki}, as defined in (5.15)),

$$\begin{bmatrix} z\mathcal{A}_k(z) \\ \hat{\mathcal{A}}_k(z) \end{bmatrix} = \frac{1}{1-\kappa_k^2} \begin{bmatrix} 1 & -\kappa_k \\ -\kappa_k & 1 \end{bmatrix} \begin{bmatrix} \mathcal{A}_{k+1}(z) \\ \hat{\mathcal{A}}_{k+1}(z) \end{bmatrix} = \begin{bmatrix} 1 & \kappa_k \\ \kappa_k & 1 \end{bmatrix}^{-1} \begin{bmatrix} \mathcal{A}_{k+1}(z) \\ \hat{\mathcal{A}}_{k+1}(z) \end{bmatrix},$$

and compare this with (5.16). The choice $\kappa_k = a_{k+1,k+1}$ ensures that $\mathcal{A}_k(z)$ is a polynomial of degree k. This form of the recursion also explains why the Jury test works. Suppose $\mathcal{A}_{k+1}(z)$ is stable. Then $|\kappa_k| < 1$, since $|a_{k+1,k+1}|$ is the product of the absolute values of the zeros of $\mathcal{A}_{k+1}(z)$. Moreover, since $|\mathcal{A}_{k+1}(z)| \geq |\hat{\mathcal{A}}_{k+1}(z)|$ for $|z| \geq 1$, the polynomial $\mathcal{A}_k(z)$ is nonzero for $|z| \geq 1$. Therefore if the recursion starts with a stable polynomial $\mathcal{A}_p(z) = (1/b_0)\mathcal{B}(z)$, then the polynomials $\mathcal{A}_k(z)$ are stable and $|\kappa_k| < 1$ for $k = p-1, \ldots, 0$. The converse can be shown as in the proof of stability of the polynomials generated by the Levinson–Durbin algorithm. If $|\kappa_k| < 1$ for $k = 0, \ldots, p-1$, then the recursion $\mathcal{A}_{k+1}(z) = z\mathcal{A}_k(z) + \kappa_k \hat{\mathcal{A}}_k(z)$ started at $\mathcal{A}_0(z) = 1$ generates a sequence of stable polynomials.

In terms of the vectors $a^{(k)} = (1, a_{k1}, \ldots, a_{kk}, 0, \ldots, 0)$, the recursion (5.20) can be written as

$$\begin{aligned} a^{(k)} &= \begin{bmatrix} (1-\kappa_k^2)^{-1}(I_{k+2} - \kappa_k J_{k+2}) & 0 \\ 0 & I_{p-k-1} \end{bmatrix} a^{(k+1)} \\ &= H_k^{-1} a^{(k+1)}, \quad k = p-1, \ldots, 0, \end{aligned} \tag{5.21}$$

with H_k defined in (5.18), where κ_k is the reflection coefficient computed by Algorithm 5.2.

5.1.3 Factorization of Jury matrix

Vostrý in [Vos75] points out that Algorithm 5.2 computes a factorization of the Jury matrix (5.8). Using the formula (5.21) for the recursion of Algorithm 5.2, we find that

$$\frac{1}{b_0} H_0^{-1} \cdots H_{p-1}^{-1} \mathcal{J}(b) = \begin{bmatrix} 1 & 0 \\ 0 & L \end{bmatrix}, \tag{5.22}$$

where

$$L = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 & 0 & 0 \\ a_{11} & 1 & 0 & \cdots & 0 & 0 & 0 \\ a_{22} & a_{21} & 1 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ a_{p-3,p-3} & a_{p-3,p-4} & a_{p-3,p-5} & \cdots & 1 & 0 & 0 \\ a_{p-2,p-2} & a_{p-2,p-3} & a_{p-2,p-4} & \cdots & a_{p-2,1} & 1 & 0 \\ a_{p-1,p-1} & a_{p-1,p-2} & a_{p-1,p-3} & \cdots & a_{p-1,2} & a_{p-1,1} & 1 \end{bmatrix}.$$

The factorization shows that the Jury matrix is nonsingular if the vector b defines a stable polynomial. It also provides an $O(p^2)$ algorithm for solving equations with coefficient matrix $\mathcal{J}(b)$. In particular, it can be used to solve (5.10). To compute $y = \mathcal{J}(b)^{-1}e$, we first compute

$$\frac{1}{b_0}H_0^{-1}\cdots H_{p-1}^{-1}e = \begin{bmatrix} \sigma_0^2 \\ -\kappa_0\sigma_0^2 \\ \vdots \\ -\kappa_{p-1}\sigma_{p-1}^2 \end{bmatrix}$$

and then calculate y using forward substitution with the triangular matrix on the right-hand side of (5.22). In other words, from the output of Algorithm 5.2, the solution of (5.10) can be computed as

$$y_0 = \sigma_0^2, \qquad y_{k+1} = -\sigma_k^2\kappa_k - y_1 a_{kk} - \cdots - y_k a_{k1}, \quad k = 0, \ldots, p-1. \tag{5.23}$$

5.2 Entropy

Recall the definition of the cone of nonnegative trigonometric polynomials in (5.1) and its semidefinite characterization (5.3). To simplify notation, we use the inner product

$$\begin{aligned} \langle x, y\rangle &= x_0y_0 + 2x_1y_1 + \cdots + 2x_py_p \\ &= \frac{1}{2\pi}\int_0^{2\pi} F_x(e^{\mathrm{j}\omega})F_y(e^{\mathrm{j}\omega})\,d\omega, \end{aligned} \tag{5.24}$$

on $\mathbf{R}^{p+1}$. The adjoint of the linear mapping $\mathcal{D}$, for this inner product on $\mathbf{R}^{p+1}$ and the trace inner product on $\mathbf{S}^{p+1}$, is the function $\mathcal{T} : \mathbf{R}^{p+1} \to \mathbf{S}^{p+1}$ that maps $y = (y_0, \ldots, y_p)$ to the

symmetric Toeplitz matrix (5.7). The dual cone of K is the cone of positive semidefinite Toeplitz matrices

$$K^* = \{y \mid \langle y, x\rangle \geq 0 \ \forall x \in K\} = \{y \mid \mathcal{T}(y) \succeq 0\}. \tag{5.25}$$

We will discuss two related convex functions, associated with the cones K and K^*. The first function is

$$\phi(x) = -\frac{1}{2\pi}\int_0^{2\pi} \log F_x(e^{\mathrm{j}\omega})d\omega, \tag{5.26}$$

with domain $\mathbf{dom}\,\phi = K \setminus \{0\}$. This is the negative of the *(differential) entropy rate* of the moving-average process with power spectrum $F_x(e^{\mathrm{j}\omega})$. The second function is

$$\psi(y) = \log(e^T\mathcal{T}(y)^{-1}e) \tag{5.27}$$

with domain $\mathbf{dom}\,\psi = \operatorname{int} K^* = \{y \mid \mathcal{T}(y) \succ 0\}$. (Recall that $e = (1, 0, \ldots, 0)$, so $e^T\mathcal{T}(y)^{-1}e = (\mathcal{T}(y)^{-1})_{00}$.) The function ψ can be evaluated by solving the Yule–Walker equation (5.19) with coefficient matrix $\mathcal{T}(y)$, since

$$e^T\mathcal{T}(y)^{-1}e = 1/\sigma_p^2. \tag{5.28}$$

Another useful expression is

$$\psi(y) = \log(e^T\mathcal{T}(y)^{-1}e) = -\log(y_0 - \tilde{y}^T\tilde{\mathcal{T}}(y)^{-1}\tilde{y}) \tag{5.29}$$

where on the right-hand side we refer to a partition of $\mathcal{T}(y)$ as

$$\mathcal{T}(y) = \left[\begin{array}{c|ccc} y_0 & y_1 & \cdots & y_p \\ \hline y_1 & y_0 & \cdots & y_{p-1} \\ \vdots & \vdots & \ddots & \vdots \\ y_p & y_{p-1} & \cdots & y_0 \end{array}\right] = \left[\begin{array}{cc} y_0 & \tilde{y}^T \\ \tilde{y} & \tilde{\mathcal{T}}(y) \end{array}\right].$$

The second expression in (5.29) shows that ψ is a convex function, since the argument of the logarithm is concave in y. The function $\psi(y)$ is equal to the negative entropy of the autoregressive process defined by the solution a_{p1}, …, a_{pp}, σ_p of (5.19), *i.e.*,

$$\psi(y) = -\frac{1}{2\pi}\int_0^{2\pi} \log \frac{\sigma_p^2}{|1 + a_{p1}e^{-\mathrm{j}\omega} + \cdots + a_{pp}e^{-\mathrm{j}p\omega}|^2}\, d\omega = -\log\sigma_p^2.$$

Up to a change of sign and a constant, the two functions form a pair of conjugates; we will see that

$$\phi^*(y) = \psi(-y) - 1, \qquad \psi^*(x) = \phi(-x) - 1. \tag{5.30}$$

Discussions of the duality relations between the two functions and their importance in signal processing can be found in [BGL98, BGL01]. In Section 5.3 the function ϕ is used as the kernel to define a Bregman distance.

5.2.1 Semidefinite representations

It will be useful to express the functions $\phi(x)$ and $\psi(y)$ as optimal values of convex optimization problems.

We first consider the negative entropy function ϕ. If $x \in K \setminus \{0\}$, then $F_x(z)$ has a spectral factorization

$$F_x(z) = \mathcal{B}_*(z)\mathcal{B}(z) \tag{5.31}$$

where

$$\mathcal{B}(z) = b_0 + b_1 z^{-1} + \cdots + b_p z^{-p}, \qquad \mathcal{B}_*(z) = \mathcal{B}(1/z) = b_0 + b_1 z + \cdots + b_p z^p,$$

with real coefficients $b_0, \ldots, b_p$ and $b_0 > 0$. The factor $\mathcal{B}(z)$ can be chosen to have all its zeros on or inside the unit circle ($\mathcal{B}(z) \neq 0$ for $|z| > 1$). If $x \in \operatorname{int} K$, then $\mathcal{B}(z)$ can be chosen to have its zeros inside the unit circle. This choice of $\mathcal{B}(z)$ is known as the *minimum-phase* spectral factor and is unique. Substituting $z = e^{\mathrm{j}\omega}$, one can write the spectral factorization (5.31) as $F_x(e^{\mathrm{j}\omega}) = |\mathcal{B}(e^{\mathrm{j}\omega})|^2$. From the minimum-phase spectral factors we immediately obtain the value of the negative entropy function:

$$\phi(x) = -2 \log b_0. \tag{5.32}$$

The minimum-phase spectral factorization of positive trigonometric polynomials is efficiently computed by the *cepstral method*, described in [Vai93, appendix D] [SN97, §5.4], or by the Newton method proposed by Tunnicliffe Wilson [Wil69]. Tunnicliffe Wilson's method finds the coefficients b in the spectral factorization (5.31) by solving the equivalent set of

quadratic equations

$$\mathcal{D}(bb^T) = x \tag{5.33}$$

via Newton's method. The two methods are compared in [FSR03].

It is also known that spectral factorization problems can be formulated as semidefinite programming problems with low-rank solutions. Replacing bb^T in (5.33) with a positive semidefinite matrix $X \in \mathbf{S}^{p+1}$ gives a convex relaxation

$$\mathcal{D}(X) = x, \qquad X \succeq 0.$$

The feasible solution X with maximum element $X_{00} = e^T X e$ can be shown to be equal to $X = bb^T$, where b is the vector of coefficients of the minimum-phase spectral factor; see [MW01] [Hac03, theorem 6.6] [Dum07, theorem 2.15]. If we combine this fact with the expression (5.32), we see that the negative entropy $\phi(x)$ is the optimal value of the convex optimization problem

$$\begin{array}{ll} \text{minimize} & -\log(e^T X e) \\ \text{subject to} & \mathcal{D}(X) = x \\ & X \succeq 0 \end{array} \tag{5.34}$$

in the variable X, as a function of the right-hand side x of the equality constraint.

Convex duality then gives another expression for $\phi(x)$. A convenient dual for (5.34) can be derived starting from the reformulation

$$\begin{array}{ll} \text{minimize} & -\log v \\ \text{subject to} & e^T X e = v \\ & \mathcal{D}(X) = x \\ & X \succeq 0, \end{array}$$

with an extra scalar variable v. The Lagrangian of this problem is

$$\begin{aligned} L(X, v, w, y, Z) &= -\log v - w(e^T X e - v) + \langle y, \mathcal{D}(X) - x\rangle - \mathbf{tr}(ZX) \\ &= -\log v + wv + \mathbf{tr}(X(\mathcal{T}(y) - wee^T - Z)) - \langle x, y\rangle, \end{aligned}$$

and the dual function is

$$\inf_{v>0,X} L(X, v, w, y, Z) = \begin{cases} \log w - \langle x, y\rangle + 1 & w > 0,\ \mathcal{T}(y) - wee^T = Z \\ -\infty & \text{otherwise.} \end{cases}$$

The Lagrange dual of (5.34) is therefore

$$\begin{array}{ll} \text{maximize} & \log w - \langle x, y\rangle + 1 \\ \text{subject to} & wee^T \preceq \mathcal{T}(y), \end{array} \tag{5.35}$$

with a scalar variable w and a vector variable $y = (y_0, \ldots, y_p)$. The variables w and y are the Lagrange multipliers for the equality constraints in (5.34). Since strong duality holds (the dual problem is strictly feasible), $\phi(x)$ is also equal to the optimal value of (5.35).

The dual problem (5.35) can be further simplified by eliminating w. Dual feasibility requires the Toeplitz matrix $\mathcal{T}(y)$ to be positive semidefinite. It therefore has a factorization (5.11), and the inequality in the dual problem can be written as

$$\mathbf{diag}(\sigma_p^2, \ldots, \sigma_0^2) = U\mathcal{T}(y)U^T \succeq w(Ue)(Ue)^T = wee^T.$$

If $\mathcal{T}(y)$ is singular, we have $\sigma_p^2 = 0$ and there exists no solution with positive w, so the problem is infeasible. If $\mathcal{T}(y)$ is nonsingular, we have $0 < \sigma_p^2 \leq \cdots \leq \sigma_0^2$, so $w = \sigma_p^2 = 1/(e^T\mathcal{T}(y)^{-1}e)$ at the optimum and

$$\log w = -\log(e^T\mathcal{T}(y)^{-1}e) = -\psi(y),$$

where ψ is defined in (5.27). The result of this elimination step is an unconstrained optimization problem in the variable y:

$$\text{maximize} \quad -\psi(y) - \langle x, y\rangle + 1. \tag{5.36}$$

The optimal value of this problem is again equal to $\phi(x)$.

Using similar arguments, we derive semidefinite programming representations of $\psi(y)$. If $\mathcal{T}(y)$ is positive definite, then $\psi(y)$ is the optimal value of the convex problem

$$\begin{array}{ll} \text{minimize} & -\log w \\ \text{subject to} & wee^T \preceq \mathcal{T}(y), \end{array} \tag{5.37}$$

with variable w. The dual of this problem is

$$\begin{array}{ll} \text{maximize} & \log(e^TXe) - \langle \mathcal{D}(X), y\rangle + 1 \\ \text{subject to} & X \succeq 0, \end{array} \tag{5.38}$$

with a symmetric variable X. Since $\phi(x)$ is the optimal value of (5.34), the dual can be written as

$$\text{maximize} \quad -\phi(x) - \langle x, y\rangle + 1, \tag{5.39}$$

with variable x. By strong duality, the optimal values of (5.38) and (5.39) are also equal to $\psi(y)$.

5.2.2 Gradients

We have seen how $\phi(x)$ can be evaluated via spectral factorization, and $\psi(y)$ by solving a Yule–Walker equation. We now discuss algorithms for computing the gradients of the two functions.

Suppose $x \in \operatorname{int} K = \operatorname{int} \mathbf{dom}\, \phi$. The optimal value of (5.34) and of the dual problem (5.35) is $\phi(x)$. From convex duality theory, if the dual has a unique optimal solution y, then the optimal value ϕ is differentiable at x and

$$\nabla\phi(x) = -y.$$

The techniques described in Sections 5.1.2 and 5.1.3 allow us to construct the unique dual optimal solution y from the primal optimal solution, as follows. A primal feasible X and dual feasible y, w are optimal for (5.34) and (5.35) if they satisfy

$$w = X_{00}^{-1}, \qquad (\mathcal{T}(y) - wee^T)X = 0 \tag{5.40}$$

(see [BV04, §5]). The second equality is known as *complementary slackness*. Now let b be the vector of coefficients of the minimum-phase spectral factor, so $X = bb^T$ is optimal for (5.34). Then, from (5.40) the dual optimal solution w, y satisfies $w = 1/b_0^2$ and

$$\mathcal{J}(b)y = \mathcal{T}(y)b = \frac{1}{b_0}e.$$

The solution $y = (1/b_0)\mathcal{J}(b)^{-1}e$ can be computed using Algorithm 5.2 and the factorization of $\mathcal{J}(b)$ given in Section 5.1.3. Algorithm 5.2 thus provides an $O(p^2)$ algorithm for computing the gradient of ϕ at a point $x \in \operatorname{int} K$, from its spectral factor.

The function ψ is clearly differentiable, with gradient

$$\nabla\psi(y) = -\frac{1}{e^T\mathcal{T}(y)^{-1}e}\mathcal{D}(\mathcal{T}(y)^{-1}ee^T\mathcal{T}(y)^{-1}). \tag{5.41}$$

The gradient is easily obtained from the solution of the Yule–Walker equation (5.19), since $\mathcal{T}(y)^{-1}e = \sigma_p^{-2}(1, a_{p1}, \ldots, a_{pp})$. If we define $a = (1, a_{p1}, \ldots, a_{pp})$, then

$$\nabla\psi(y) = -\frac{1}{\sigma_p^2}\mathcal{D}(aa^T). \tag{5.42}$$

5.2.3 Legendre property

We have shown that $\phi(x)$ is the optimal value of (5.36), and therefore

$$\phi(x) = \sup_y\left(-\langle x, y\rangle - \psi(y)\right) + 1 = \psi^*(-x) + 1. \tag{5.43}$$

This is the second of the conjugacy relations (5.30). Similarly, from the fact that $\psi(y)$ is the optimal value of (5.39) we conclude that

$$\psi(y) = \sup_x\left(-\langle x, y\rangle - \phi(x)\right) + 1 = \phi^*(-y) + 1. \tag{5.44}$$

This gives the first identity in (5.30). The relation (5.44) can also be obtained directly from (5.43) by noting that ψ is a closed convex function (closed because its domain is open, and its value $\psi(y)$ tends to infinity as y approaches the boundary of its domain and $\sigma_p \to 0$). Therefore $\psi^{**} = \psi$ [Roc70, theorem 12.2], and the identity $\phi^*(y) = \psi(-y) - 1$ follows by taking the conjugates of the two sides of (5.43).

In addition to being closed, convex, and differentiable on an open domain $\operatorname{int} K^*$, the function ψ is strictly convex. It is therefore a convex function of *Legendre type* [Roc70, page 258]. By [Roc70, theorem 26.5] the pair $(\operatorname{int} K, \phi)$ is also of Legendre type, and the gradient $\nabla\psi$ is a one-to-one mapping from $\operatorname{int} K^*$ to $-\operatorname{int} K$, with inverse

$$(\nabla\psi)^{-1}(x) = -\nabla\phi(-x). \tag{5.45}$$

Algorithms 5.1 and 5.2 give efficient algorithms for evaluating the two gradient mappings. Even though the function ϕ is finite on the boundary of K (except at the origin), it is *essentially smooth, i.e.*, the norm of $\nabla\phi(x)$ grows unboundedly as x approaches the boundary [Roc70, page 251].

5.3 Itakura–Saito distance

Let $h : \mathbf{R}^n \to \mathbf{R}$ be a closed strictly convex function with $\operatorname{int}(\mathbf{dom}\, h) \neq \emptyset$, and assume h is differentiable on $\operatorname{int}(\mathbf{dom}\, h)$. The *Bregman distance* with *kernel* h is the function

$$d_h(x, v) = h(x) - h(v) - \langle \nabla h(v), x - v \rangle, \tag{5.46}$$

with domain $\mathbf{dom}\, d_h = \mathbf{dom}\, h \times \operatorname{int}(\mathbf{dom}\, h)$. For example, the squared Euclidean distance $d_h(x, v) = (1/2)\|x - v\|_2^2$ is the Bregman distance for $h(x) = (1/2)\|x\|_2^2$ and the standard inner product $\langle u, v \rangle = u^T v$. The best known non-quadratic example is the relative entropy

$$d_h(x, v) = \sum_{i=1}^{n} (x_i \log(x_i/v_i) - x_i + v_i), \tag{5.47}$$

which is the Bregman distance for the negative entropy function $h(x) = \sum_i x_i \log x_i$ and the standard inner product.

From the definition (5.46) it is clear that d_h is convex in x for fixed v. By convexity of h, we also have $d_h(x, v) \geq 0$ for all $(x, v) \in \mathbf{dom}\, d_h$. Strict convexity of h further implies that $d_h(x, v) = 0$ only if $x = v$. However, $d_h(x, v) \neq d_h(v, x)$ in general, so $d_h(x, v)$ is not a true distance.

5.3.1 Itakura–Saito and Kullback–Leibler distance

The Bregman distance d_ϕ for the negative entropy kernel (5.26) and the inner product (5.24) is called the *Itakura–Saito distance*. To simplify notation we omit the subscript in d_ϕ, and define

$$\begin{aligned} d(x, v) &= \phi(x) - \phi(v) - \langle \nabla \phi(v), x - v \rangle \\ &= \frac{1}{2\pi} \int_0^{2\pi} \left(\frac{F_x(e^{\mathrm{j}\omega})}{F_v(e^{\mathrm{j}\omega})} - \log \frac{F_x(e^{\mathrm{j}\omega})}{F_v(e^{\mathrm{j}\omega})} - 1 \right) d\omega. \end{aligned} \tag{5.48}$$

The domain of d is $(K \setminus \{0\}) \times (\operatorname{int} K)$. The Itakura–Saito distance was first proposed and has been studied extensively in speech processing [GM76, GBG80]. For surveys of the Itakura–Saito and other spectral distance measures, see [Bas89, GKT09, Bas13].

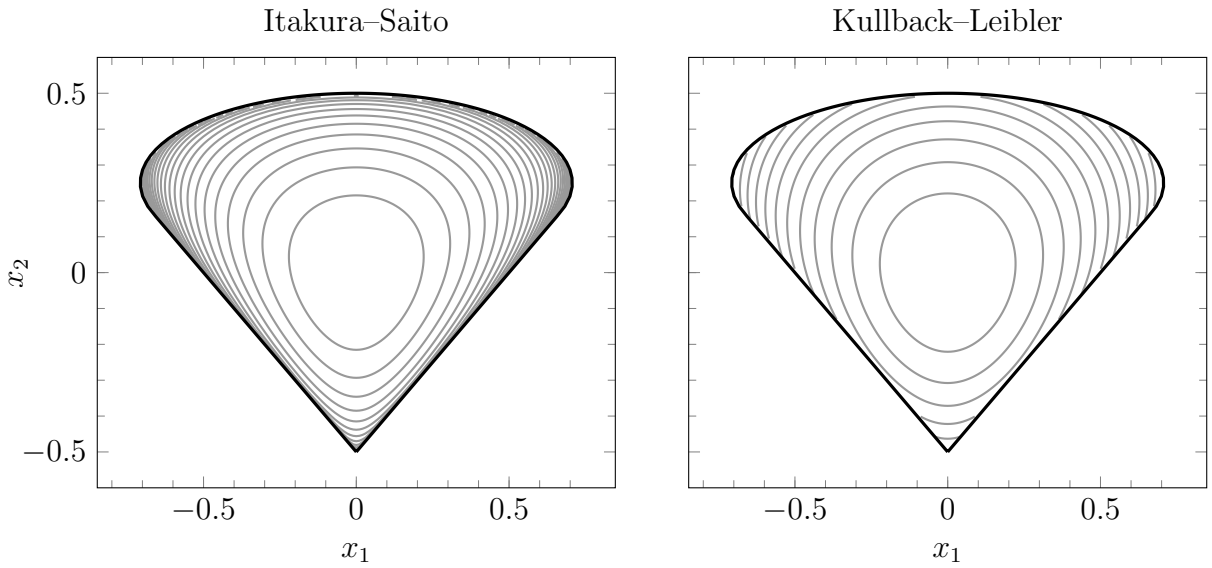

Figure 5.1: Contour lines of the function $\phi(1, x_1, x_2)$ defined in (5.26) (*Left*), and contour lines of the function $\tilde{\phi}(1, x_1, x_2)$ defined in (5.49) (*Right*), on the set $\{(x_1, x_2) \mid (1, x_1, x_2) \in K\} = \{(x_1, x_2) \mid 1 + 2x_1 \cos\omega + 2x_2 \cos 2\omega \geq 0 \ \forall \omega\}$

The Legendre property of the underlying kernel function ϕ makes the Itakura–Saito distance well suited for the generalized proximal methods discussed in Sections 6.1 and 6.2. This is a key difference with the better known *Kullback–Leibler divergence,*

$$d_{\mathrm{kl}}(x, v) = \frac{1}{2\pi} \int_0^{2\pi} (F_x(e^{\mathrm{j}\omega}) \log \frac{F_x(e^{\mathrm{j}\omega})}{F_v(e^{\mathrm{j}\omega})} - F_x(e^{\mathrm{j}\omega}) + F_v(e^{\mathrm{j}\omega}))d\omega,$$

which is also a Bregman distance, namely for the kernel function

$$\tilde{\phi}(x) = \frac{1}{2\pi} \int_0^{2\pi} F_x(e^{\mathrm{j}\omega}) \log F_x(e^{\mathrm{j}\omega}) d\omega. \tag{5.49}$$

However, the function $\tilde{\phi}$ is not essentially smooth, *i.e.*, the norm of $\tilde{\phi}(x)$ does not necessarily go to infinity as x approaches the boundary of K. Figure 5.1 illustrates the different behavior of ϕ and $\tilde{\phi}$ near the boundary of K.

5.3.2 Strong convexity

Another important property of the Itakura–Saito distance, required for its use in generalized proximal gradient methods, follows from the strong convexity of the negative entropy

function $\phi(x)$ when restricted to a bounded set. We define the norm

$$\|x\|_1 = \frac{1}{2\pi}\int_0^{2\pi} |F_x(e^{\mathrm{j}\omega})|d\omega.$$

With respect to this norm the function ϕ is 1-strongly convex on the set $\{x \mid \|x\| \leq 1\}$ where $\|x\| = \langle x, x\rangle^{1/2}$. In other words,

$$d(x,v) \geq \frac{1}{2}\|x-v\|_1^2 \quad \forall (x,v) \in \mathbf{dom}\, d,\ \|x\| \leq 1,\ \|v\| \leq 1. \tag{5.50}$$

To see this, we consider $v \in \operatorname{int} K$ and $x \in K \setminus \{0\}$, and define

$$g(t) = \phi(v + t(x-v)) = -\frac{1}{2\pi}\int_0^{2\pi} \log(F_{v+t(x-v)}(e^{\mathrm{j}\omega}))d\omega$$

for $v + t(x-v) \in K$. The second derivative is

$$\begin{aligned} g''(t) &= \frac{1}{2\pi}\int_0^{2\pi} \frac{F_{x-v}(e^{\mathrm{j}\omega})^2}{F_{v+t(x-v)}(e^{\mathrm{j}\omega})^2}d\omega \\ &\geq \left(\frac{1}{2\pi}\int_0^{2\pi} \frac{F_{x-v}(e^{\mathrm{j}\omega})^2}{F_{v+t(x-v)}(e^{\mathrm{j}\omega})^2}d\omega\right)\left(\frac{1}{2\pi}\int_0^{2\pi} F_{v+t(x-v)}(e^{\mathrm{j}\omega})^2 d\omega\right) \\ &\geq \left(\frac{1}{2\pi}\int_0^{2\pi} |F_{x-v}(e^{\mathrm{j}\omega})|d\omega\right)^2 \\ &= \|v-x\|_1^2. \end{aligned}$$

The first inequality follows from $\|v + t(x-v)\| \leq 1$, and the second inequality from the Cauchy–Schwarz inequality. Integrating the inequality $g''(t) \geq \|v-x\|_1^2$ twice gives (5.50). More generally,

$$d(x,v) \geq \frac{\sigma}{2}\|x-v\|_1^2 \quad \forall (x,v) \in \mathbf{dom}\, d,\ \|x\| \leq 1/\sqrt{\sigma},\ \|v\| \leq 1/\sqrt{\sigma}.$$

CHAPTER 6

Entropic Proximal Operators for Nonnegative Trigonometric Polynomials

We now discuss algorithms for solving optimization problems over the cone K (5.1) of nonnegative trigonometric polynomials. If f is a cost function with an epigraph $\{(x,t) \mid f(x) \leq t\}$ that can be represented by linear matrix inequalities, then the semidefinite representation of K (5.3) allows us to formulate the problem of minimizing $f(x)$ over K as a semidefinite program (SDP), and solve it using general-purpose SDP solvers. The interior-point algorithms implemented in these solvers have a complexity of $O(p^4)$ per iteration, if we assume that the complexity is dominated by the cost of handling the constraint $x \in K$ (*i.e.*, ignoring the cost of handling the constraints that represent the epigraph of f). The special-purpose interior-point algorithms developed in [AV02, GHN03, RV06] reduce the complexity to $O(p^3)$ per iteration. First-order proximal algorithms such as the proximal gradient algorithm [Nes04, BT09b] or the alternating direction method of multipliers (ADMM) [BPC11] offer no immediate improvement over the $O(p^3)$ per-iteration-complexity of the customized interior-point methods, since they require at each iteration a Euclidean projection on the positive semidefinite cone (*i.e.*, an eigenvalue decomposition of order $p+1$) and, moreover, converge more slowly than interior-point methods.

The purpose of this chapter is to describe faster first-order methods, based on the generalized distance introduced in Chapter 5, with a complexity of roughly $O(p^2)$ or $O(p(\log p)^2)$ operations per iteration. Specifically, the algorithms are based on generalized proximal operators defined in terms of the Itakura–Saito distance

$$d(x,v) = \frac{1}{2\pi}\int_0^{2\pi} \left(\frac{F_x(e^{\mathrm{j}\omega})}{F_v(e^{\mathrm{j}\omega})} - \log\frac{F_x(e^{\mathrm{j}\omega})}{F_v(e^{\mathrm{j}\omega})} - 1\right) d\omega, \tag{6.1}$$

with domain $\mathbf{dom}\, d = (K \setminus \{0\}) \times (\operatorname{int} K)$. We present an efficient method for computing a generalized projection $x = \Pi(a, v)$, defined as the solution of the problem

$$\begin{array}{ll} \text{minimize} & \langle a, x\rangle + d(x, v) \\ \text{subject to} & x_0 = 1 \end{array} \tag{6.2}$$

for an arbitrary $(p+1)$-vector a and a vector $v \in \operatorname{int} K$. If we interpret $F_x(e^{\mathrm{j}\omega})$ as a power spectrum, then the constraint $x_0 = 1$ normalizes the total power

$$x_0 = \frac{1}{2\pi}\int_0^{2\pi} F_x(e^{\mathrm{j}\omega})d\omega. \tag{6.3}$$

Our method for (6.2) reduces the problem to a nonlinear equation in one variable (equivalently, an unconstrained differentiable convex optimization problem in one variable) that can be solved using Newton's method. Each Newton iteration requires the solution of a positive definite Toeplitz equation, which takes $O(p^2)$ operations using Levinson's algorithm, as summarized in Section 5.1.1, or $O(p(\log p)^2)$ operations using superfast Toeplitz solvers. Since the number of Newton steps is small and weakly dependent on problem size, we conclude that the complexity of solving problem (6.2) is roughly $O(p^2)$ or $O(p(\log p)^2)$.

The Itakura–Saito projection operation (6.2) should be compared with the Euclidean projection of the vector $v - a$ on the set $\{x \in K \mid x_0 = 1\}$, *i.e.*, the solution of

$$\begin{array}{ll} \text{minimize} & \langle a, x\rangle + \frac{1}{2}\|x - v\|^2 \\ \text{subject to} & x \in K \\ & x_0 = 1 \end{array} \tag{6.4}$$

where $\|u\|^2 = \langle u, u\rangle$. This is a non-trivial convex optimization problem [AV00, DTS01, Dum07].

To test the effectiveness of the entropic projection operator, we use it in an accelerated proximal gradient method [AT06, Tse08] for optimization problems of the form

$$\begin{array}{ll} \text{minimize} & f(x) \\ \text{subject to} & x \in K \\ & x_0 = 1, \end{array} \tag{6.5}$$

where f is a differentiable convex function. The entropic projection operator can also be used in other types of first-order methods, for example, the mirror descent algorithm [BT03].

The generalized projection can be further extended to define generalized proximal operators, which map vectors a and v to the solution of

$$\text{minimize} \quad \langle a, x\rangle + \tilde{g}(x_0) + \frac{1}{\tau} d(x, v) \tag{6.6}$$

where $\tilde{g}$ is a possibly nondifferentiable convex function of one variable and $\tau > 0$. This is useful for optimization problems

$$\begin{array}{ll} \text{minimize} & f(x) + \tilde{g}(x_0) \\ \text{subject to} & x \in K, \end{array}$$

with differentiable f. The second term in the cost function assigns a cost to the total power (6.3).

The main results of the chapter are in Section 6.1, where we describe the algorithm for the Itakura–Saito projection (6.2). Section 6.2 contains numerical examples with a generalized proximal gradient method based on the Itakura–Saito distance. To make the thesis self-contained, Appendix E gives more details and a proof of convergence of this proximal gradient method. This chapter is adapted from [CV18].

6.1 Entropic proximal operators

Proximal algorithms, such as the projected and proximal gradient methods and their accelerated variants [Nes04, BT09a], the Douglas-Rachford method and alternating direction method of multipliers [CP07, BPC11], or Dykstra's sequential projection method [BD86, Han88], depend on efficient methods for evaluating the proximal operators of cost functions. The proximal operator of a convex function g is the mapping

$$\text{prox}_g(u) = \operatorname*{argmin}_x \left(g(x) + \frac{1}{2}\|x - u\|_2^2\right), \tag{6.7}$$

where $\|\cdot\|_2$ is the Euclidean norm. If g is the indicator function of a set, this is the Euclidean projection of u on the set.

Useful extensions of the proximal methods are obtained by replacing the squared Euclidean distance in the definition by a generalized Bregman distance function, in the hope of making the generalized proximal operators or projections easier to compute. There is an extensive literature on methods of this type (see, for example, the book [CZ97]), and the properties that the generalized distance function must satisfy depend on the algorithm in which they are applied [BL00]. In the numerical experiments of the next section we will apply one specific algorithm, Auslender and Teboulle's generalization of an accelerated proximal gradient method due to Nesterov [AT06, Tse08]. The generalized proximal operator used in this method is defined as

$$\text{prox}_g^h(a, v) = \underset{x}{\text{argmin}}\,(\langle a, x\rangle + g(x) + d_h(x, v)), \tag{6.8}$$

where d_h is a Bregman distance (5.46). On the right-hand side of (6.8), the vectors a and v are given, with $v \in \text{int}\,(\mathbf{dom}\, h)$. The variable in the minimization problem is x and the feasible set is $\mathbf{dom}\, g \cap \mathbf{dom}\, h$. This is a generalization of (6.7): if $d_h(x, v) = (1/2)\|x - v\|_2^2$ and $\langle a, x\rangle = a^T x$, then the solution of (6.8) is $\text{prox}_g(v - a)$.

Proximal algorithms that use the generalized definition (6.8) require that for every a and every $v \in \text{int}\,(\mathbf{dom}\, h)$, the minimizer in (6.8) is a unique and easily computed point $\hat{x} \in \text{int}\,(\mathbf{dom}\, h)$. The classical example is the indicator $g(x) = \delta_C(x)$ of the probability simplex $C = \{x \in \mathbf{R}^n \mid x \geq 0,\ \mathbf{1}^T x = 1\}$, and the relative entropy function (5.47). With this choice of g and d_h, the solution of the optimization problem in (6.8) is

$$\hat{x}_i = \frac{v_i e^{-a_i}}{\sum_{j=1}^n v_j e^{-a_j}}, \quad i = 1, \ldots, n.$$

Sufficient conditions that guarantee existence in $\text{int}\,(\mathbf{dom}\, h)$ and uniqueness of the solution of (6.8) are discussed in papers on generalized distances (for example, [BL00, BT03]). We will return to this question in the context of the specific applications studied in the chapter.

In the following sections we consider the generalized proximal operator (6.8) defined by the Itakura–Saito distance (6.1) and the inner product (5.24)

$$\langle x, y\rangle = x_0 y_0 + 2x_1 y_1 + \cdots + 2x_p y_p$$

on $\mathbf{R}^{p+1}$.

6.1.1 Projection

We first take for g the indicator function of

$$\{x \in \mathbf{R}^{p+1} \mid x_0 = 1\} = \{x \in \mathbf{R}^{p+1} \mid \frac{1}{2\pi}\int_0^{2\pi} F_x(e^{\mathrm{j}\omega})d\omega = 1\},$$

and denote the generalized proximal operator by

$$\Pi(a, v) = \operatorname*{argmin}_{x_0=1} \left(\langle a, x\rangle + d(x, v)\right) \tag{6.9}$$

To simplify notation we define $c = a - \nabla\phi(v)$ and write the minimization problem in the definition as

$$\begin{array}{ll} \text{minimize} & \langle c, x\rangle + \phi(x) \\ \text{subject to} & x_0 = 1. \end{array} \tag{6.10}$$

The feasible set is a compact set $\{x \in K \mid x_0 = 1\}$. Since ϕ is strictly convex and essentially smooth, the problem has a unique solution in $\operatorname{int} K$, for every c. The optimality conditions for the projection problem are

$$\nabla\phi(x) = -c - \lambda e, \qquad \langle e, x\rangle = 1.$$

The variable λ is a Lagrange multiplier for the equality constraint in (6.10). The unknown x can be eliminated from the first equation, using the inverse gradient mapping in (5.45). Substituting $x = -\nabla\psi(c + \lambda e)$ in the second equation gives a nonlinear equation in λ:

$$\langle e, \nabla\psi(c + \lambda e)\rangle + 1 = 0.$$

More explicitly, in view of (5.41), λ is the root of the equation

$$-\frac{e^T(\mathcal{T}(c) + \lambda I)^{-2}e}{e^T(\mathcal{T}(c) + \lambda I)^{-1}e} + 1 = 0 \tag{6.11}$$

in the interval $(-\lambda_{\min}(\mathcal{T}(c)), \infty)$. After solving the nonlinear equation for λ, we compute the solution of the Yule–Walker equation with coefficient matrix $\mathcal{T}(c+\lambda e) = \mathcal{T}(c)+\lambda I$, and obtain x from the expression (5.42).

Solving (6.11) is equivalent to solving the dual of problem (6.10), which is given by

$$\text{maximize} \quad -\phi^*(-c - \lambda e) - \lambda = -\psi(c + \lambda e) - \lambda + 1.$$

As we have seen, the negative of the cost function

$$h(\lambda) = \psi(c + \lambda e) + \lambda - 1 = \log(e^T(\mathcal{T}(c) + \lambda I)^{-1}e) + \lambda - 1$$

is strictly convex and differentiable on the interval $(-\lambda_{\min}(\mathcal{T}(c)), \infty)$. It increases to ∞ as $\lambda \to -\lambda_{\min}(\mathcal{T}(c))$ and as $\lambda \to \infty$. The optimal λ can therefore be found by setting the derivative of h to zero. The derivative $h'(\lambda)$ is the left-hand side of (6.11).

To solve the nonlinear equation (6.11), one can minimize $h(\lambda)$ by Newton's method with a backtracking line search, or use a safeguarded Newton method, similar to algorithms used for secular equations [CGT00, §7.3] [NW06, §4.3]. To check whether $\lambda > -\lambda_{\min}(\mathcal{T}(c))$, one can use the Levinson–Durbin algorithm and terminate the recursion early, as soon as a reflection coefficient with $|\kappa_k| \geq 1$ is found.

A feasible starting value $\lambda > -\lambda_{\min}(\mathcal{T}(c))$ is easily found by embedding the Toeplitz matrix $\mathcal{T}(c)$ in a symmetric circulant matrix. The smallest eigenvalue of the circulant matrix is a lower bound on the smallest eigenvalue of $\mathcal{T}(c)$ and can be computed by the discrete Fourier transform.

The second derivative is

$$h''(\lambda) = -\frac{(e^T(\mathcal{T}(c) + \lambda I)^{-2}e)^2}{(e^T(\mathcal{T}(c) + \lambda I)^{-1}e)^2} + 2\frac{e^T(\mathcal{T}(c) + \lambda I)^{-3}e}{e^T(\mathcal{T}(c) + \lambda I)^{-1}e}.$$

The value of $h(\lambda)$ and its derivatives follow from the solution of the Yule–Walker equation with coefficient matrix $\mathcal{T}(c) + \lambda I$. They can be computed in order p^2 operations by the Levinson–Durbin algorithm, or in order $p(\log p)^2$ operations by superfast algorithms for positive definite Toeplitz systems.

Let $\lambda^\star$ be the solution of (6.11). The derivative $h'(\lambda)$, which is given by the left-hand side of (6.11), increases monotonically from $-\infty$ to zero on the interval $(-\lambda_{\min}(\mathcal{T}(c)), \lambda^\star]$ and from zero to one on the interval $[\lambda^\star, \infty)$. When started at a point $\lambda^{(0)} \in (-\lambda_{\min}(\mathcal{T}(c)), \lambda^\star)$, Newton's method with unit steps produces an increasing sequence of values that converges to $\lambda^\star$ from the left. When started at a point $\lambda^{(0)} \in (\lambda^\star, \infty)$, the Newton update may be infeasible, and backtracking or bisection steps can be taken to find a point in $(-\lambda_{\min}(\mathcal{T}(c)), \lambda^{(0)})$.

In practice, a small number of Newton iterations (on the order of 10) is sufficient, almost independent of problem size. The computational cost of the projection algorithm is therefore a small multiple of the cost of a positive definite Toeplitz factorization, *i.e.*, $O(p^2)$ for the Levinson–Durbin algorithm and $O(p(\log p)^2)$ for the superfast algorithms.

6.1.2 Proximal operator

The method of the previous section can be extended to generalized proximal operators (6.8) where $g(x)$ has the form $g(x) = \tilde{g}(x_0)$, with $\tilde{g}$ a convex function of one variable. This generalized proximal operator maps vectors $a \in \mathbf{R}^{p+1}$ and $v \in \operatorname{int} K$ to the vector

$$\operatorname*{argmin}_x \left(\langle a, x\rangle + \tilde{g}(x_0) + d(x, v)\right). \tag{6.12}$$

The projection operator discussed in Section 6.1.1 is a special case with $\tilde{g}(t) = \delta_{\{1\}}(t)$, the indicator function of $\{1\}$. Other interesting choices are

$$\tilde{g}(t) = \delta_{[0,1]}(t) = \begin{cases} 0 & 0 \le t \le 1 \\ +\infty & \text{otherwise,} \end{cases} \qquad \tilde{g}(t) = \frac{t^2}{2} + \delta_{[0,\infty)}(t) = \begin{cases} t^2/2 & t \ge 0 \\ +\infty & \text{otherwise.} \end{cases}$$

We will assume that $\tilde{g}$ increases faster than linearly as $t \to \infty$ (*i.e.*, that $\lim_{t\to\infty} g(t)/t = \infty$) and, without loss of generality, that $\mathbf{dom}\, \tilde{g} \subseteq \mathbf{R}^+$. This implies that $\tilde{g}$ is *co-finite* and therefore its conjugate $\tilde{g}^*(\lambda)$ is defined for all λ [Roc70, corollary 13.3.1].

If we define $c = a - \nabla\phi(v)$ and introduce an auxiliary variable u, the minimization problem in the definition (6.12) is

$$\begin{array}{ll} \text{minimize} & \langle c, x\rangle + \tilde{g}(u) + \phi(x) \\ \text{subject to} & \langle e, x\rangle = u. \end{array}$$

The Lagrange dual of this problem is

$$\text{maximize} \quad -\psi(c + \lambda e) - \tilde{g}^*(\lambda) + 1.$$

If $\tilde{g}^*$ is a simple function, as in the examples mentioned above, this concave maximization problem with one variable can be solved by modifying the methods described in Section 6.1.1.

6.2 Numerical experiments

In this section we use the generalized projection $\Pi(a, v)$ (6.9) in an accelerated proximal gradient method for solving convex problems of the form

$$\begin{array}{ll} \text{minimize} & f(x) \\ \text{subject to} & x \in K \\ & x_0 = 1, \end{array} \tag{6.13}$$

where f is convex and differentiable on $\{x \in K \mid x_0 = 1\}$. The algorithm is algorithm IGA (Improved Interior Gradient Algorithm) from [AT06], and is also discussed in [Tse08, algorithm 1]. It is an extension to non-Euclidean projections of an accelerated proximal gradient algorithm by Nesterov. The algorithm generates three strictly feasible sequences v^k, x^k, y^k, using the following recursion started at a strictly feasible $v^0 = x^0$:

$$y^k = (1-\theta_k)x^{k-1} + \theta_k v^{k-1} \tag{6.14a}$$

$$v^k = \Pi(\tau_k \nabla f(y^k), v^{k-1}) \tag{6.14b}$$

$$x^k = (1-\theta_k)x^{k-1} + \theta_k v^k. \tag{6.14c}$$

Appendix E describes the algorithm in detail, including different strategies for choosing the parameters $\theta_k \in (0, 1)$ and $\tau_k > 0$. (In the experiments we used the monotonic search strategy.)

6.2.1 Covariance estimation

As a first example, we consider a variation of the line spectrum estimation example in Section 3.5.1. We estimate the parameters in a signal model

$$s(t) = \sum_{k=1}^{\rho} c_k e^{\mathrm{j}\omega_k t} + w(t), \tag{6.15}$$

where $w(t)$ is white noise with variance σ^2. Under standard assumptions [SM97, §4.1] [PM96, §12.5] the covariance matrix of $s(t)$ of order $p+1$ is given by

$$\begin{bmatrix} r_0 & r_{-1} & \cdots & r_{-p} \\ r_1 & r_0 & \cdots & r_{-p+1} \\ \vdots & \vdots & \ddots & \vdots \\ r_p & r_{p-1} & \cdots & r_0 \end{bmatrix} = \sigma^2 I + \sum_{k=1}^{\rho} |c_k|^2 \begin{bmatrix} 1 \\ e^{\mathrm{j}\omega_k} \\ \vdots \\ e^{\mathrm{j}p\omega_k} \end{bmatrix} \begin{bmatrix} 1 \\ e^{\mathrm{j}\omega_k} \\ \vdots \\ e^{\mathrm{j}p\omega_k} \end{bmatrix}^H, \tag{6.16}$$

i.e., a positive multiple of the identity plus a rank-ρ positive semidefinite Toeplitz matrix. If the line spectrum has Hermitian symmetry, the signal is real and the covariance matrix is symmetric ($r_k = r_{-k}$).

To fit a covariance matrix of this structure to observed data, we introduce variables $t = \sigma^2$ and $y = r - te$, and solve a convex problem

$$\begin{array}{ll} \text{minimize} & y_0 + \gamma \tilde{f}(y + te) \\ \text{subject to} & y \in K^*. \end{array} \tag{6.17}$$

The second term in the objective measures the quality of the fit of the matrix $\mathcal{T}(y) + tI = \mathcal{T}(y+te)$ to the observed data. The first term in the objective is a multiple of the trace of $\mathcal{T}(y)$ and is added to encourage low-rank solutions. The coefficient γ is a positive regularization parameter. The dual of this problem can be written as

$$\begin{array}{ll} \text{maximize} & -\gamma \tilde{f}^*((x-e)/\gamma) \\ \text{subject to} & x \in K \\ & x_0 = 1, \end{array} \tag{6.18}$$

where $\tilde{f}^*$ is the conjugate of $\tilde{f}$. If $\tilde{f}^*$ is differentiable, this is of the form (6.13) with $f(x) = \gamma \tilde{f}^*((x-e)/\gamma)$.

In the example we take a simple quadratic penalty function $\tilde{f}(r) = \|\mathcal{T}(r) - R\|_F^2$, where R is a sample covariance matrix. With this choice, $\tilde{f}^*$ is quadratic. The sample covariance matrix R is constructed from $N = 150$ samples of a time series $s(t)$ of the form (6.15), shown in Figure 6.1. We take $\rho = 4$, and the frequencies ω_k and magnitudes $|c_k|$ indicated with red circles in Figure 6.2. The noise is Gaussian white noise with variance $\sigma^2 = 64$. The sample

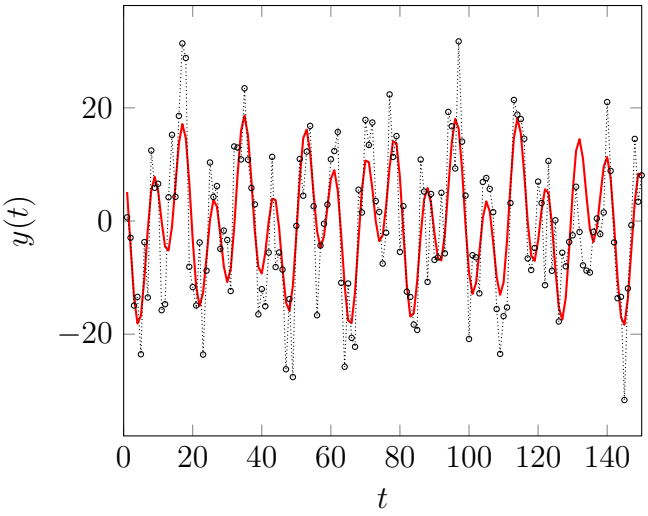

Figure 6.1: True signal (solid line) and noisy samples (circles).

covariance matrix of order $p+1 = 30$ is constructed as $R = HH^T/(N-p)$ where H is the $(p+1) \times (N-p)$ Hankel matrix with $s(1), \ldots, s(N-p)$ in its first row.

Figure 6.2 shows the solution of the primal and dual problems (6.18) and (6.17), for $\gamma = 2 \cdot 10^{-4}$. As can be seen, the recovered spectrum is quite accurate. The estimated noise variance σ^2 is 77.2. Figure 6.3 shows the relative optimality gap in the dual problem (6.18) versus iteration number. To estimate the optimality gap we use the optimal value computed by CVX [GB14]. The error decreases roughly as $1/k^2$.

6.2.2 Euclidean projection on nonnegative polynomials

To evaluate the complexity for large p, we test the generalized proximal gradient method on a family of test problems

$$\begin{array}{ll} \text{minimize} & \sum_{k=1}^{p} (x_k - a_k)^2 \\ \text{subject to} & x \in K \\ & x_0 = 1. \end{array}$$

This problem arises in signal processing, as the problem of finding the normalized autocorrelation sequence closest to a given sequence [AV00, DTS01, Dum07]. Figure 6.4 shows a small example.

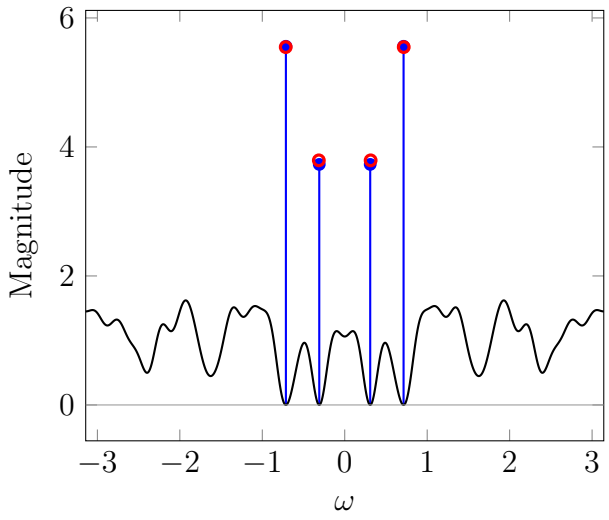

Figure 6.2: True and estimated line spectrum (red circles and blue stem lines, respectively) and dual optimal polynomial $F_x(e^{\mathrm{j}\omega})$ (solid curve).

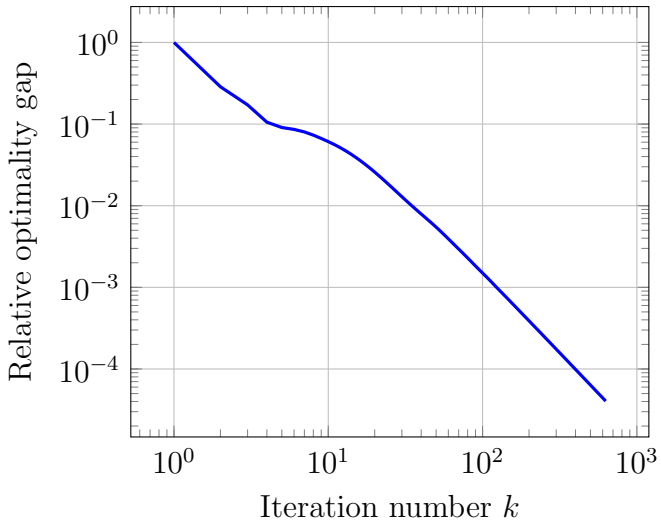

Figure 6.3: Relative suboptimality versus iteration number of the generalized proximal gradient method applied to the dual problem (6.18). The optimality gap is computed as $(f(x^k) - f^{\mathrm{opt}})/|f^{\mathrm{opt}}|$, where $f(x)$ denotes the negative of the dual objective value in (6.18) and f^{opt} is the optimal solution computed by CVX.

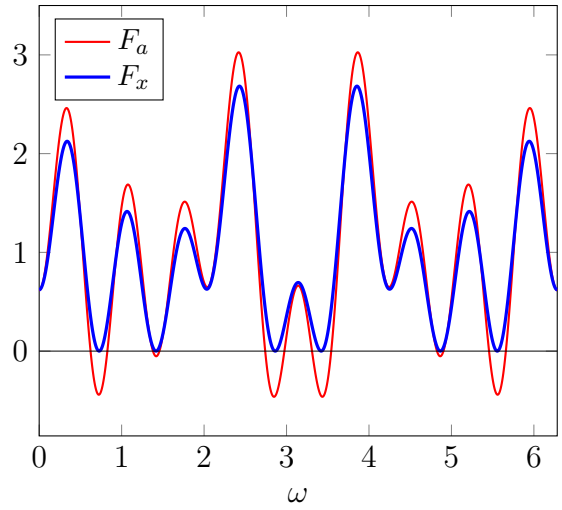

Figure 6.4: Euclidean projection on the normalized nonnegative trigonometric polynomials of order $p = 9$. The red curve is $F_a(e^{j\omega})$; the blue curve is $F_x(e^{j\omega})$, where x is the Euclidean projection of a on $\{x \in K \mid x_0 = 1\}$.

The experiment was performed on an Intel Core i5-2410M 2.30GHz CPU with 6GB RAM and 64-bit operating system, using MATLAB version 7.12 (R2011a). The initial proximal stepsize in the algorithm is $\tau_0 = 10/p$. The monotone search strategy in Appendix E (with $\beta = 2$) is used. In most problems less than 5 line search steps during the first few iterations of the algorithm were needed.

In Figure 6.5 we compare the complexity of the generalized proximal gradient method (6.14) with general-purpose interior-point solvers called via CVX. The problem instances are randomly generated, with a from the normal distribution $N(0, I)$. For the first three data points ($p+1 = 200, 400, 800$), SDPT3 [TTT02] was used as the interior-point method. Each of these data points is an average over 10 instances. For $p + 1 = 1000, 1200, 1600, 2000$, SeDuMi [Stu99] with the low-precision option was used. The first three of these data points are averages over five instances. For $p + 1 = 2000$, only one instance was used. The blue curve is the total time for the proximal gradient method, averaged over the same instances as the interior-point solvers. The iteration was terminated when the relative suboptimality was less than 10^{-4}. The CVX solution was used to evaluate the suboptimality.

The number of iterations for the interior-point solvers was generally between 10 and 30,

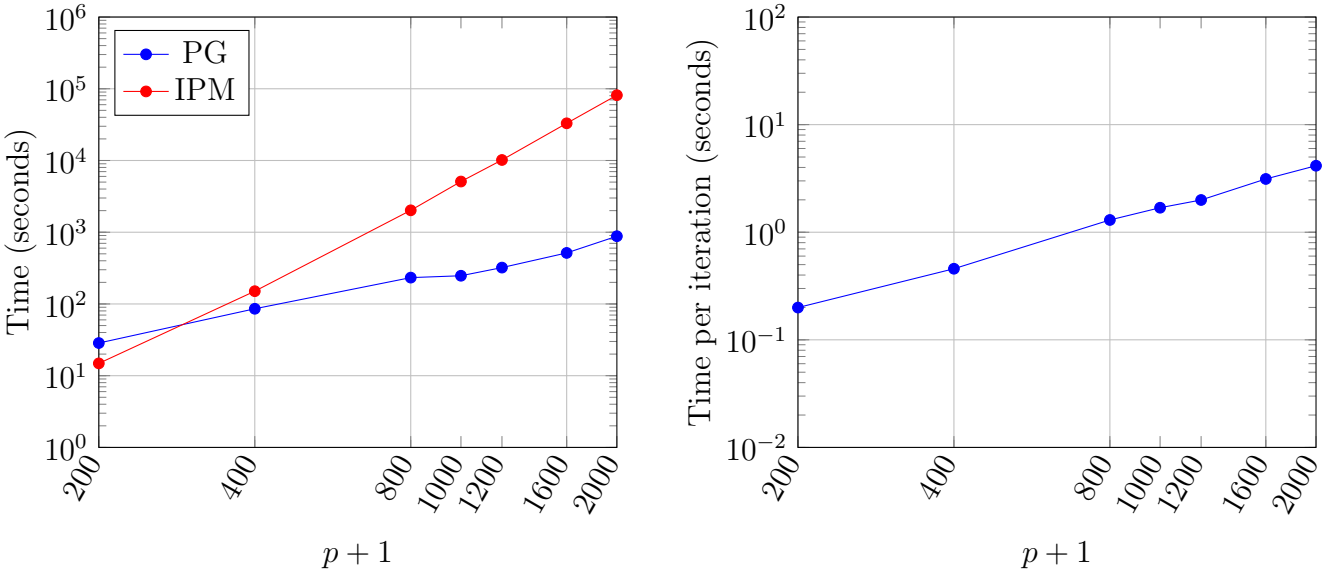

Figure 6.5: Time for proximal gradient method and general-purpose interior-point methods (IPM) versus problem size (*Left*), and time per iteration for the proximal gradient method (*Right*).

and for the proximal gradient method between 100 and 200. On average about 10 Newton iterations were sufficient to evaluate the generalized projections. From Figure 6.5, it can be observed that the proximal gradient method exhibits a complexity under $O(p^2)$, whereas the SDP solvers have a complexity close to $O(p^4)$.

In Figure 6.6 we show results for larger problems of size up to 8000. Each data point is an average over five instances, and the iteration was terminated when the relative improvement in the cost function, defined as $|\min_{i<k} f(x^i) - f(x^k)|/\min_{i\leq k} f(x^i)$, was below 10^{-6}.

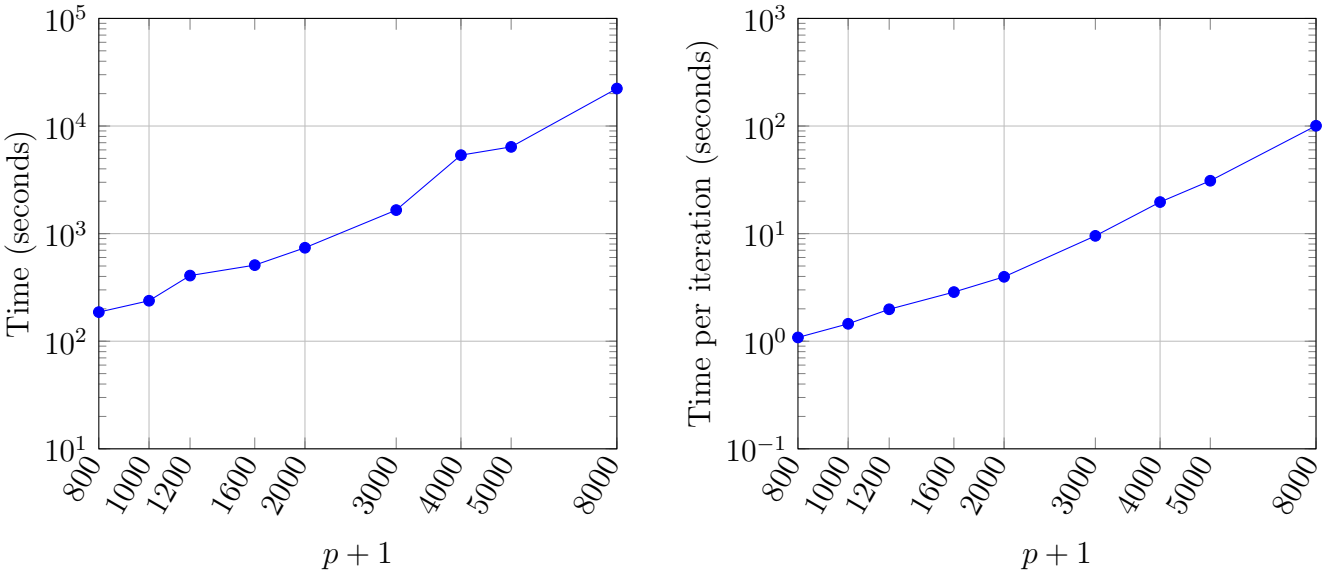

Figure 6.6: Time for the proximal gradient method (*Left*) and time per iteration (*Right*) versus problem size.

CHAPTER 7

Conclusion

In the first part of the thesis we developed semidefinite representations of a class of gauge functions and atomic norms for sets parameterized by linear matrix pencils. The formulations extend the semidefinite representation of the atomic norm associated with the trigonometric moment curve, which underlies recent results in continuous or 'off-the-grid' compressed sensing. The main contribution is a self-contained constructive proof of the semidefinite representations, using techniques developed in the literature on the Kalman–Yakubovich–Popov (KYP) lemma. In addition to opening new possible areas of applications in system theory and control, the connection with the KYP lemma is important for numerical algorithms. Specialized techniques for solving semidefinite programs (SDPs) derived from the KYP lemma, for example, by exploiting real symmetries and rank-one structure [GHN03, LP04, RV06, LV07, HV14], are useful in the development of customized interior-point solvers for the SDPs discussed in the thesis.

The second part of the thesis discussed a generalized proximal operator for the cone of nonnegative trigonometric polynomials, based on the Itakura–Saito distance. Analysis and numerical experiments show that projections in this distance have a complexity that is roughly quadratic in the degree of the polynomial. Proximal algorithms based on the generalized distance therefore scale better than standard (Euclidean) proximal algorithms, which require eigenvalue decompositions, and interior-point methods, which have a complexity that is cubic or higher at each iteration. In consequence, the approach discussed in the thesis is very promising for a wider range of large-scale SDP applications and algorithms.

APPENDIX A

Classical methods for line spectrum estimation

Methods for fitting data to the model

$$y = \sum_{k=1}^{r} x_k \begin{bmatrix} 1 \\ e^{-\mathrm{j}\omega_k} \\ e^{-\mathrm{j}2\omega_k} \\ \vdots \\ e^{-\mathrm{j}(n-1)\omega_k} \end{bmatrix}, \tag{A.1}$$

have been very extensively studied in various areas of applications. Some of the representative and popular ones are presented here. A list of more complete references can be found in [SM97, §4].

Prony's approach

Fitting given exact data y to the model in (A.1) has been studied at least as early as by Prony [Pro95]. The key idea is the following Prony's polynomial $p : \mathbf{C} \to \mathbf{C}$:

$$p(s) = \prod_{k=1}^{r}(1 - \frac{s}{s_k}) = 1 + \sum_{k=1}^{r} p_k s^k = \begin{bmatrix} 1 & s & s^2 & \cdots & s^r \end{bmatrix} \begin{bmatrix} 1 \\ p_1 \\ p_2 \\ \vdots \\ p_r \end{bmatrix},$$

where s is a complex argument and this r-degree polynomial has roots at $s_k = e^{\mathrm{j}\omega_k}$, or equivalently, $p(s_k) = 0$ for $k = 1, \ldots, r$. Then for *any* $r + 1$ consecutive elements of y, we

have

$$\begin{bmatrix} y_l \\ y_{l+1} \\ y_{l+2} \\ \vdots \\ y_{l+r} \end{bmatrix}^H \begin{bmatrix} 1 \\ p_1 \\ p_2 \\ \vdots \\ p_r \end{bmatrix} = \sum_{k=1}^{r} \bar{x}_k s_k^{l-1} \left(\begin{bmatrix} 1 & s_k & s_k^2 & \cdots & s_k^r \end{bmatrix} \begin{bmatrix} 1 \\ p_1 \\ p_2 \\ \vdots \\ p_r \end{bmatrix} \right)$$

$$= \sum_{k=1}^{r} \bar{x}_k s_k^{-l-r+1} p(s_k) = 0.$$

This states that the vector $p = \begin{bmatrix} 1 & p_1 & p_2 & \cdots & p_r \end{bmatrix}^T$ is orthogonal to the 'signal subspace', which is a subspace of $\mathbf{C}^{r+1}$ spanned by all possible vectors consisting of $r+1$ consecutive elements of y.

Therefore, assuming that r, the number of components, is known, the following procedure recovers parameters in (A.1) with exact data $y \in \mathbf{C}^n$ and $n \geq 2r$.

Exact Prony's method

1. Solve a (possibly over-complete but consistent) system of linear equations to obtain Prony's polynomial p.

$$\begin{bmatrix} y_1 & y_2 & \cdots & y_{n-r} \\ y_2 & y_3 & \cdots & y_{n-r+1} \\ y_3 & y_4 & \cdots & y_{n-r+2} \\ \vdots & \vdots & \vdots & \vdots \\ y_{r+1} & y_{r+2} & \cdots & y_n \end{bmatrix}^H \begin{bmatrix} 1 \\ p_1 \\ p_2 \\ \vdots \\ p_r \end{bmatrix}$$

$$= \begin{bmatrix} \bar{y}_2 & \bar{y}_3 & \cdots & \bar{y}_{r+1} \\ \bar{y}_3 & \bar{y}_4 & \cdots & \bar{y}_{r+2} \\ \vdots & \vdots & \cdots & \vdots \\ \bar{y}_{n-r+1} & \bar{y}_{n-r+2} & \cdots & \bar{y}_n \end{bmatrix} \begin{bmatrix} p_1 \\ p_2 \\ \vdots \\ p_r \end{bmatrix} + \begin{bmatrix} \bar{y}_1 \\ \bar{y}_2 \\ \vdots \\ \bar{y}_{n-r} \end{bmatrix} = 0$$

2. Compute the r roots of $p(s)$, for example, by computing the eigenvalues of its companion matrix. For each k, $e^{\mathrm{j}\omega_k} = s_k$.

3. Solve an (over-complete but consistent) system of linear equations to obtain $x = \begin{bmatrix} x_1 & x_2 & \cdots & x_r \end{bmatrix}^T$.

$$y = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ s_1^{-1} & s_2^{-1} & \cdots & s_r^{-1} \\ \vdots & \vdots & \cdots & \vdots \\ s_1^{-(n-1)} & s_2^{-(n-1)} & \cdots & s_r^{-(n-1)} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_r \end{bmatrix}$$

◇

In the presence of noise, however, the linear equations at step 1 is often no longer consistent. Although it could be replaced by computing the (total) least-squares solution, its stability cannot be easily characterized. Moreover, at step 2, finding roots of a polynomial from its coefficients can be an extremely ill-conditioned problem, *i.e.*, a small perturbation of the coefficients may well leave some of the roots falling far from the unit circle. As a result, much works have been done to robustify Prony's method. The MUSIC method is perhaps the most popular one among them.

MUSIC

The MUltiple SIgnal Classification method follows the same spirit as the exact Prony's method, but it computes the signal subspace and its orthogonal complement (sometimes referred to as the 'noise subspace') in a robust way, achieving asymptotic stability under certain statistical assumptions. Consider the model (A.1) with additive noise,

$$y = \sum_{k=1}^{r} x_k \begin{bmatrix} 1 \\ e^{-\mathrm{j}\omega_k} \\ e^{-\mathrm{j}2\omega_k} \\ \vdots \\ e^{-\mathrm{j}(m-1)\omega_k} \end{bmatrix} + w, \tag{A.2}$$

where the data y are measurements corrupted by *i.i.d.* white Gaussian noise w with zero mean and covariance matrix $\sigma^2 I$. Here we abuse the notation and use y to denote a length-

m vector of random variables. Assume also that the phase of each x_k is random, and their distributions are such that

$$\mathbf{E}\left[xx^H\right] = D = \begin{bmatrix} |x_1|^2 & 0 & \cdots & 0 \\ 0 & |x_2|^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \cdots & |x_r|^2 \end{bmatrix}$$

This is satisfied, for example, when the phase of each x_k is drawn independently and uniformly from $\{\pm 1\}$ or $[0, 2\pi)$.

We use the following notation to simplify matters,

$$a(s) = \begin{bmatrix} 1 \\ s^{-1} \\ s^{-2} \\ \vdots \\ s^{-m+1} \end{bmatrix} \quad \text{and} \quad A = \begin{bmatrix} a(s_1) & a(s_2) & \cdots & a(s_r) \end{bmatrix},$$

where $m > r$ and $s_k = e^{\mathrm{j}\omega_k}$. The covariance matrix of $y \in \mathbf{C}^m$ is then

$$\mathbf{E}\left[yy^H\right] = \mathbf{E}\left[(Ax + w)(Ax + w)^H\right] = ADA^H + \sigma^2 I.$$

The first term ADA^H has rank r, and suppose it has eigenvalue decomposition

$$ADA^H = \begin{bmatrix} P_S & P_N \end{bmatrix} \mathbf{diag}\left(\lambda_1, \ldots, \lambda_r, 0, \ldots, 0\right) \begin{bmatrix} P_S & P_N \end{bmatrix}^H,$$

where the columns of $P_S \in \mathbf{C}^{m \times r}$ span the signal subspace and the column(s) of $P_N \in \mathbf{C}^{m \times (m-r)}$ span the noise subspace. Then $A^H P_N = 0$ implies

$$a(s_k)^H P_N = a(s_k^{-1})^T P_N = 0, \qquad k = 1, \ldots, r,$$

as well as $a(s^{-1})^T P_N \neq 0$ and $P_N^H a(s) \neq 0$ for $s \notin \{s_1, \ldots, s_r\}$. Therefore, the only roots of $a(s^{-1})^T P_N P_N^H a(s)$ on the unit circle are $s_1, \ldots, s_r$. Through the eigenvalue decomposition

of $\mathbf{E}\,[yy^H]$,

$$\begin{aligned}\mathbf{E}\,[yy^H] &= ADA^H + \sigma^2 I \\ &= \begin{bmatrix} P_S & P_N \end{bmatrix} \begin{bmatrix} \lambda_1 + \sigma^2 & & & & & \\ & \ddots & & & 0 & \\ & & \lambda_r + \sigma^2 & & & \\ & & & \sigma^2 & & \\ & 0 & & & \ddots & \\ & & & & & \sigma^2 \end{bmatrix} \begin{bmatrix} P_S & P_N \end{bmatrix}^H,\end{aligned}$$

we can obtain P_N corresponding to the smallest eigenvalue of multiplicity $m - r$.

Following the analysis, we conclude that the supports ω_k's can be recovered exactly by computing the roots of $a(s^{-1})^T P_N P_N^H a(s)$ on the unit circle, provided that a true covariance matrix $\mathbf{E}\,[yy^H]$ is available. In practice, however, it is common to instead use a sample covariance matrix

$$\frac{1}{N}\sum_{l=1}^{N} \begin{bmatrix} y_l \\ y_{l+1} \\ y_{l+2} \\ \vdots \\ y_{l+m-1} \end{bmatrix} \begin{bmatrix} y_l \\ y_{l+1} \\ y_{l+2} \\ \vdots \\ y_{l+m-1} \end{bmatrix}^H = \frac{1}{N} Y Y^H,$$

where we denote the $m \times N$ Hankel matrix

$$Y = \begin{bmatrix} y_1 & y_2 & \cdots & y_N \\ y_2 & y_3 & \cdots & y_{N+1} \\ y_3 & y_4 & \cdots & y_{N+2} \\ \vdots & \vdots & \vdots & \vdots \\ y_m & y_{m+1} & \cdots & y_{N+m-1} \end{bmatrix}.$$

Note that when n measurements are available, we typically take $N = n - m + 1$.

When the sample covariance matrix is used, two major issues need to be considered. The first is the determination of r, as it may not be clear how many eigenvalues correspond to the noise subspace P_N. The second issues come up when we are trying to find r roots of

$a(s^{-1})^T P_N P_N^H a(s)$ on the unit circle, as they may deviate from it due to inaccurate estimate of P_N. The following procedure describes heuristics to go around them and achieves stability as $N \to \infty$.

MUSIC method

1. Compute the eigenvalue decomposition of the sample covariance matrix $\frac{1}{N} Y Y^H$, or equivalently, the singular value decomposition (SVD) of Y.

2. If r is not given, then determine it by finding a cutoff point of the eigenvalues or singular values, perhaps with some knowledge about the noise level σ^2. Then from the partition

$$\frac{1}{N} Y Y^H = \begin{bmatrix} P_S & P_N \end{bmatrix} \begin{bmatrix} \Lambda_S & 0 \\ 0 & \Lambda_N \end{bmatrix} \begin{bmatrix} P_S & P_N \end{bmatrix}^H$$

or

$$Y = \begin{bmatrix} P_S & P_N \end{bmatrix} \begin{bmatrix} \Sigma_S & 0 \\ 0 & \Sigma_N \end{bmatrix} V^H,$$

we can determine P_N.

3. Determine the supports by either of the following.

   (a) (*Spectral MUSIC* [Sch86] [Bie79]) Find the r locations of the highest peaks of the function

$$\frac{1}{a(e^{-\mathrm{j}\omega})^T P_N P_N^H a(e^{\mathrm{j}\omega})}, \qquad \omega \in [0, 2\pi].$$

   (b) (*Root MUSIC* [Bar83]) Find the r (pairs of reciprocal) roots of the equation $a(s^{-1})^T P_N P_N^H a(s) = 0$ that are closest to the unit circle. Then take ω_k's as the angular positions of them, *i.e.*, $s_k = |s_k| e^{\mathrm{j}\omega_k}$. ◇

As in Prony's analysis, this method requires $n \geq 2r$. In fact, we obtain Prony's polynomial in step 2, $a(s^{-1})^T P_N = cp(s)$ (where c is a scalar constant) if we take $n = 2r$ and $m = r + 1$. To have advantage over the exact Prony's method, we need $n > 2r$, which is often the case in practice.

ESPRIT and matrix pencil method

We describe another subspace method, Estimation of Signal Parameters via Rotational Invariance Techniques [PRK86], that also achieves asymptotic stability based on the same statistical model. It has similar computational cost as MUSIC but is reported to obtain slightly more accurate estimates in most cases [SM97, page 164].

The key idea is the following observation of rotational invariance. Let

$$A_1 = \begin{bmatrix} I_{m-1} & 0 \end{bmatrix} A \quad \text{and} \quad A_2 = \begin{bmatrix} 0 & I_{m-1} \end{bmatrix} A,$$

where A is as defined before, then

$$A_1 = A_2 \, \mathbf{diag}\,(s_1, \ldots, s_r).$$

From $ADA^H = P_S \Lambda_S P_S^H$ we know that $P_S = AC$ with some nonsingular C $(= DA^H P_S \Lambda_S^{-1})$, so if we define

$$P_{S1} = \begin{bmatrix} I_{m-1} & 0 \end{bmatrix} P_S \quad \text{and} \quad P_{S2} = \begin{bmatrix} 0 & I_{m-1} \end{bmatrix} P_S,$$

then

$$\begin{aligned} P_{S1} \;=\; A_1 C \;&=\; A_2 \, \mathbf{diag}\,(s_1, \ldots, s_r) C \\ &=\; P_{S2} (C^{-1} \, \mathbf{diag}\,(s_1, \ldots, s_r) C). \end{aligned}$$

Therefore, the s_k's are obtained as the eigenvalues of $P_{S2}^\dagger P_{S1}$.

The first two steps of the ESPRIT method are identical to those of the MUSIC method.

ESPRIT method

1. Compute the eigenvalue decomposition of the sample covariance matrix $\frac{1}{N} YY^H$, or equivalently, the singular value decomposition (SVD) of Y. (When working directly with Y, this procedure is equivalent to a matrix pencil method [HS88]. This procedure actually also works in a more general case of estimating damped complex sinusoids.)

2. Do the decomposition as in step 2 of the MUSIC method, but will use P_S.

3. Let $P_{S1} = \begin{bmatrix} I_{m-1} & 0 \end{bmatrix} P_S$ and $P_{S2} = \begin{bmatrix} 0 & I_{m-1} \end{bmatrix} P_S$. Compute the eigenvalues of $P_{S2}^\dagger P_{S1}$ to get s_k's, and take ω_k's as the angular positions of them, *i.e.*, $s_k = |s_k| e^{\mathrm{j}\omega_k}$. $\diamond$

In all the above methods, we solve for x as the least-squares solution of

$$y \approx \left[\begin{array}{cccc} a(e^{\mathrm{j}\omega_1}) & a(e^{\mathrm{j}\omega_2}) & \cdots & a(e^{\mathrm{j}\omega_r}) \end{array} \right] x.$$

Note that there is no restriction on how close the ω_k's can be with respect to each other. However, the problems of determining the subspaces and eigenvalues become extremely ill-conditioned when any two ω_k's are too close together. Hence with the presence of noise and numerical error, the estimates could be quite inaccurate.

APPENDIX B

Subsets of the complex plane

This appendix is from [CV17]. In this appendix we explain the notation used in equation (2.8) to describe subsets of the closed complex plane. Recall that we use the notation

$$q_\Theta(\mu,\nu) = \begin{bmatrix} \mu \\ \nu \end{bmatrix}^H \begin{bmatrix} \Theta_{11} & \Theta_{12} \\ \Theta_{21} & \Theta_{22} \end{bmatrix} \begin{bmatrix} \mu \\ \nu \end{bmatrix}$$

for the quadratic form defined by a Hermitian 2×2 matrix Θ.

Lines and circles If Φ is a 2×2 Hermitian matrix with $\det \Phi < 0$, then the quadratic equation

$$q_\Phi(\lambda, 1) = 0 \tag{B.1}$$

defines a straight line (if $\Phi_{11} = 0$) or a circle (if $\Phi_{11} \neq 0$) in the complex plane. Three important special cases are

$$\Phi_{\mathrm{u}} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \qquad \Phi_{\mathrm{i}} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \qquad \Phi_{\mathrm{r}} = \begin{bmatrix} 0 & \mathrm{j} \\ -\mathrm{j} & 0 \end{bmatrix},$$

for the unit circle, imaginary axis, and real axis, respectively. Curves defined by two different matrices Φ, $\tilde{\Phi}$ can be mapped to one another by applying a nonsingular congruence transformation $\tilde{\Phi} = R\Phi R^H$.

When $\Phi_{11} = 0$, we include the point $\lambda = \infty$ in the solution set of (B.1). Alternatively, one can define points in the closed complex plane as directions $(\mu, \nu) \neq 0$. If $\nu \neq 0$, the pair (μ, ν) represents the complex number $\lambda = \mu/\nu$. If $\nu = 0$, it represents the point at infinity. Using this notation, a circle or line in the closed complex plane is defined as the nonzero

$\angle\lambda$	Ψ	Assumptions
$[a-b, a+b]$	$\begin{bmatrix} 0 & -e^{\mathrm{j}a} \\ -e^{-\mathrm{j}a} & 2\cos b \end{bmatrix}$	$0 \leq b \leq \pi$
$[a, 2\pi - a]$	$\begin{bmatrix} 0 & 1 \\ 1 & -2\cos a \end{bmatrix}$	$0 \leq a \leq \pi$

Table B.1: Common choices of Ψ with $\Phi = \Phi_{\mathrm{u}}$ (λ on the unit circle).

solution set of a quadratic equation

$$q_\Phi(\mu, \nu) = \begin{bmatrix} \mu \\ \nu \end{bmatrix}^H \Phi \begin{bmatrix} \mu \\ \nu \end{bmatrix} = 0,$$

with $\det \Phi < 0$. A congruence transformation $\tilde{\Phi} = R\Phi R^H$ corresponds to a linear transformation between the sets associated with the matrices Φ and $\tilde{\Phi}$.

Segments of lines and circles The second type of set we encounter is defined by a quadratic equality and inequality

$$q_\Phi(\lambda, 1) = 0, \qquad q_\Psi(\lambda, 1) \leq 0. \tag{B.2}$$

We assume that $\det \Phi < 0$. If the inequality is redundant (*e.g.*, $\Psi = 0$) the solution set of (B.2) is the line or circle defined by the equality. Otherwise it is an arc of a circle, a closed interval of a line, or the complement of an open interval of a line. It includes the point at infinity if $\Phi_{11} = 0$ and $\Psi_{11} \leq 0$. Alternatively, one can use homogeneous coordinates and consider sets of points (μ, ν) that satisfy

$$q_\Phi(\mu, \nu) = 0, \qquad q_\Psi(\mu, \nu) \leq 0, \qquad (\mu, \nu) \neq 0. \tag{B.3}$$

For easy reference, we list the most common combinations of Φ and Ψ in tables B.1–B.3 [IH03, IH05].

As for circles and lines, we can apply a congruence transformation to reduce (B.2) to a simple canonical case. We mention two examples. Iwasaki and Hara [IH05, lemma 2] show

$\operatorname{Im}\lambda$	Ψ	Assumptions
$[a,b]$	$\begin{bmatrix} 2 & -\mathrm{j}(a+b) \\ \mathrm{j}(a+b) & 2ab \end{bmatrix}$	$a \leq b$
$[-\infty,-a]\cup[a,\infty]$	$\begin{bmatrix} -1 & 0 \\ 0 & a^2 \end{bmatrix}$	$a \geq 0$

Table B.2: Common choices of Ψ with $\Phi = \Phi_{\mathrm{i}}$ (λ imaginary).

λ	Ψ	Assumptions
$[a,b]$	$\begin{bmatrix} 2 & -(a+b) \\ -(a+b) & 2ab \end{bmatrix}$	$a \leq b$
$[-\infty,a]\cup[b,\infty]$	$\begin{bmatrix} -2 & a+b \\ a+b & -2ab \end{bmatrix}$	$a \leq b$
$[a,\infty]$	$\begin{bmatrix} 0 & -1 \\ -1 & 2a \end{bmatrix}$	
$[-\infty,a]$	$\begin{bmatrix} 0 & 1 \\ 1 & -2a \end{bmatrix}$	

Table B.3: Common choices of Ψ with $\Phi = \Phi_{\mathrm{r}}$ (λ real).

that for every Φ, Ψ with $\det \Phi < 0$, there exists a nonsingular R such that

$$\Phi = R^H \Phi_{\mathrm{i}} R, \qquad \Psi = R^H \begin{bmatrix} \alpha & \beta \\ \beta & \gamma \end{bmatrix} R \tag{B.4}$$

with α, β, γ real, and $\alpha \geq \gamma$. To see this, we first apply a congruence transformation $\Phi = R_1^H \Phi_{\mathrm{i}} R_1$ to transform Φ to Φ_{i}. Define

$$R_1^{-H} \Psi R_1^{-1} = \begin{bmatrix} x & \beta + \mathrm{j}z \\ \beta - \mathrm{j}z & y \end{bmatrix}$$

with real x, y, z, β, and consider the eigenvalue decomposition

$$\begin{bmatrix} x & \mathrm{j}z \\ -\mathrm{j}z & y \end{bmatrix} = Q \begin{bmatrix} \alpha & 0 \\ 0 & \gamma \end{bmatrix} Q^H, \tag{B.5}$$

with eigenvalues sorted as $\alpha \geq \gamma$. Since the $2,1$ element of the matrix on the left-hand side of (B.5) is purely imaginary, the columns of Q can be normalized to be of the form

$$Q = \begin{bmatrix} u & \mathrm{j}v \\ \mathrm{j}v & u \end{bmatrix}$$

with u and v real, and $u^2 + v^2 = 1$. This implies that $Q \Phi_{\mathrm{i}} Q^H = Q^H \Phi_{\mathrm{i}} Q = \Phi_{\mathrm{i}}$ and

$$Q^H \begin{bmatrix} x & \beta + \mathrm{j}z \\ \beta - \mathrm{j}z & y \end{bmatrix} Q = Q^H \begin{bmatrix} x & \mathrm{j}z \\ -\mathrm{j}z & y \end{bmatrix} Q + \begin{bmatrix} 0 & \beta \\ \beta & 0 \end{bmatrix} = \begin{bmatrix} \alpha & \beta \\ \beta & \gamma \end{bmatrix}.$$

The transformation (B.4) now follows by taking $R = Q^H R_1$.

Applying the congruence defined by R, we can reduce the conditions (B.3) to an equivalent system

$$\begin{bmatrix} \mu' \\ \nu' \end{bmatrix}^H \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \mu' \\ \nu' \end{bmatrix} = 0, \qquad \begin{bmatrix} \mu' \\ \nu' \end{bmatrix}^H \begin{bmatrix} \alpha & 0 \\ 0 & \gamma \end{bmatrix} \begin{bmatrix} \mu' \\ \nu' \end{bmatrix} \leq 0, \qquad (\mu', \nu') \neq 0, \tag{B.6}$$

where $(\mu', \nu') = R(\mu, \nu)$. In non-homogeneous coordinates,

$$\operatorname{Re} \lambda' = 0, \qquad \alpha |\lambda'|^2 + \gamma \leq 0. \tag{B.7}$$

Keeping in mind that $\alpha \geq \gamma$, we can distinguish four cases. If $0 < \gamma \leq \alpha$ the solution set of (B.7) is empty. If $\gamma = 0 < \alpha$ the solution set is a singleton $\{0\}$. If $\gamma < 0 < \alpha$, the solution

set of (B.7) is the interval of the imaginary axis defined by $|\lambda'| \leq (-\gamma/\alpha)^{1/2}$. If $\gamma \leq \alpha \leq 0$, the inequality is redundant and the solution set is the entire imaginary axis.

Another useful canonical form of (B.2) is obtained by transforming the solution set to a subset of the unit circle. If we define

$$T = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} R, \qquad \epsilon = \frac{1}{2}(\alpha + \gamma), \qquad \delta = \frac{1}{2}(\alpha - \gamma), \qquad \eta = \beta.$$

then it follows from from (B.4) that

$$\Phi = T^H \Phi_{\mathrm{u}} T, \qquad \Psi = T^H \begin{bmatrix} \epsilon + \eta & -\delta \\ -\delta & \epsilon - \eta \end{bmatrix} T.$$

The coefficients ϵ, δ, η are real, with $\delta \geq 0$. The congruence defined by T therefore transforms the conditions (B.3) to an equivalent system

$$\begin{bmatrix} \mu' \\ \nu' \end{bmatrix}^H \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} \mu' \\ \nu' \end{bmatrix} = 0, \qquad \begin{bmatrix} \mu' \\ \nu' \end{bmatrix}^H \begin{bmatrix} 0 & -\delta \\ -\delta & 2\epsilon \end{bmatrix} \begin{bmatrix} \mu' \\ \nu' \end{bmatrix} \leq 0,$$

where $(\mu', \nu') = T(\mu, \nu)$. In non-homogeneous coordinates, this is

$$|\lambda'|^2 = 1, \qquad \delta \operatorname{Re} \lambda' \geq \epsilon.$$

The solution set is empty if $\epsilon > \delta$. It is the unit circle if $\epsilon \leq -\delta$. It is the singleton $\{1\}$ if $\epsilon = \delta > 0$. It is a segment of the unit circle if $-\delta < \epsilon < \delta$.

APPENDIX C

Matrix factorization results

This appendix is from [CV17] and contains a self-contained proof of Lemma 2.3, needed in the proof of Theorem 2.5, and some other matrix factorization results that have appeared in papers on the Kalman–Yakubovich–Popov (KYP) lemma [Ran96,IMF00,BV02,BV03,PV11]. We include the proofs because their constructive character is important for the result in Theorem 2.5. Lemma C.1 is based on [Ran96, lemma 3] and [IH05, lemma 5]. Lemma 2.3 can be found in [PV11, corollary 1].

Lemma C.1. *Let U and V be two matrices in $\mathbf{C}^{p\times r}$.*

(a) If $UU^H = VV^H$, then $U = V\Lambda$ for some unitary matrix $\Lambda \in \mathbf{C}^{r\times r}$.

(b) If $UU^H = VV^H$ and $UV^H + VU^H = 0$, then $U = V\Lambda$ for some unitary and skew-Hermitian matrix $\Lambda \in \mathbf{C}^{r\times r}$.

(c) If $UU^H \preceq VV^H$ and $UV^H + VU^H = 0$, then $U = V\Lambda$ for some skew-Hermitian matrix $\Lambda \in \mathbf{C}^{r\times r}$ with $\|\Lambda\|_2 \leq 1$.

Proof. *Part (a).* If $UU^H = VV^H$, then U and V have singular value decompositions of the form

$$U = P\Sigma Q_u^H, \qquad V = P\Sigma Q_v^H, \tag{C.1}$$

with unitary P, Q_u, Q_v. The unitary matrix $\Lambda = Q_v Q_u^H$ satisfies $U = V\Lambda$.

Part (b). If we substitute the singular value decompositions (C.1) in the equation $UV^H + VU^H = 0$, we obtain

$$\Sigma(Q_u^H Q_v + Q_v^H Q_u)\Sigma^T = 0. \tag{C.2}$$

If U and V, and therefore Σ, have full column rank, this implies that the matrix $\tilde{\Lambda} = Q_u^H Q_v$ is skew-Hermitian. The matrix $\Lambda = Q_v \tilde{\Lambda} Q_v^H = Q_v Q_u^H$ is skew-Hermitian and unitary, and satisfies $U = V\Lambda$. If U and V do not have full column rank, we modify $\tilde{\Lambda}$ as follows. We write (C.2) as

$$\begin{bmatrix} \Sigma_1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \tilde{\Lambda}_{11} + \tilde{\Lambda}_{11}^H & \tilde{\Lambda}_{12} + \tilde{\Lambda}_{21}^H \\ \tilde{\Lambda}_{21} + \tilde{\Lambda}_{12}^H & \tilde{\Lambda}_{22} + \tilde{\Lambda}_{22}^H \end{bmatrix} \begin{bmatrix} \Sigma_1 & 0 \\ 0 & 0 \end{bmatrix} = 0$$

with Σ_1 positive diagonal of size $q \times q$, where $q = \mathbf{rank}(U) = \mathbf{rank}(V)$, and $\tilde{\Lambda}_{11}$ the $q \times q$ leading diagonal block of $\tilde{\Lambda}$. This shows that $\tilde{\Lambda}_{11} + \tilde{\Lambda}_{11}^H = 0$, so $\tilde{\Lambda}$ is unitary with a skew-Hermitian $1,1$ block. Since $\tilde{\Lambda}_{11}$ is skew-Hermitian it has a Schur decomposition $\tilde{\Lambda}_{11} = Q\Delta Q^H$ with unitary $Q \in \mathbf{C}^{q\times q}$, and Δ diagonal and purely imaginary. Moreover $\Delta\Delta^H \preceq I$ because $\tilde{\Lambda}_{11}$ is a submatrix of the unitary matrix $\tilde{\Lambda}$. Partition Q and Δ as

$$\tilde{\Lambda}_{11} = \begin{bmatrix} Q_1 & Q_2 \end{bmatrix} \begin{bmatrix} \Delta_1 & 0 \\ 0 & \Delta_2 \end{bmatrix} \begin{bmatrix} Q_1 & Q_2 \end{bmatrix}^H \tag{C.3}$$

with $\Delta_1\Delta_1^H \prec I$ and $\Delta_2\Delta_2^H = I$. Since $\tilde{\Lambda}$ is unitary, we have

$$\begin{aligned} \tilde{\Lambda}_{12}\tilde{\Lambda}_{12}^H &= I - \tilde{\Lambda}_{11}\tilde{\Lambda}_{11}^H \\ &= Q_1Q_1^H + Q_2Q_2^H - Q_1\Delta_1\Delta_1^HQ_1^H - Q_2\Delta_2\Delta_2^HQ_2^H \\ &= Q_1(I - \Delta_1\Delta_1^H)Q_1^H, \end{aligned}$$

and by part (a),

$$\tilde{\Lambda}_{12} = Q_1(I - \Delta_1\Delta_1^H)^{1/2}\Omega \tag{C.4}$$

for some unitary Ω. Therefore the matrix

$$\begin{bmatrix} \tilde{\Lambda}_{11} & \tilde{\Lambda}_{12} \\ -\tilde{\Lambda}_{12}^H & \Omega^H\Delta_1^H\Omega \end{bmatrix} = \begin{bmatrix} Q_1 & Q_2 & 0 \\ 0 & 0 & \Omega^H \end{bmatrix} \begin{bmatrix} \Delta_1 & 0 & \Gamma \\ 0 & \Delta_2 & 0 \\ -\Gamma & 0 & \Delta_1^H \end{bmatrix} \begin{bmatrix} Q_1^H & 0 \\ Q_2^H & 0 \\ 0 & \Omega \end{bmatrix},$$

where $\Gamma = (I - \Delta_1\Delta_1^H)^{1/2}$, is skew-Hermitian (from the expression on the left-hand side and the fact that $\tilde{\Lambda}_{11}$ is skew-Hermitian and Δ_1 is purely imaginary) and unitary (the right-hand side is a product of three unitary matrices). If we now define

$$\Lambda = Q_v \begin{bmatrix} \tilde{\Lambda}_{11} & \tilde{\Lambda}_{12} \\ -\tilde{\Lambda}_{12}^H & \Omega^H\Delta_1^H\Omega \end{bmatrix} Q_v^H$$

then Λ is unitary and skew-Hermitian, and

$$\begin{aligned} U &= P \begin{bmatrix} \Sigma_1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \tilde{\Lambda}_{11} & \tilde{\Lambda}_{12} \\ \tilde{\Lambda}_{21} & \tilde{\Lambda}_{22} \end{bmatrix} Q_v^H \\ &= P \begin{bmatrix} \Sigma_1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \tilde{\Lambda}_{11} & \tilde{\Lambda}_{12} \\ -\tilde{\Lambda}_{12}^H & \Omega^H \Delta_1^H \Omega \end{bmatrix} Q_v^H \\ &= P \begin{bmatrix} \Sigma_1 & 0 \\ 0 & 0 \end{bmatrix} Q_v^H \Lambda \\ &= V\Lambda. \end{aligned}$$

Part (c). Assume $UU^H \preceq VV^H$ and $VV^H - UU^H$ has rank s. We factorize $VV^H - UU^H = \tilde{U}\tilde{U}^H$ with $\tilde{U} \in \mathbf{C}^{p\times s}$ and write $UU^H \preceq VV^H$ and $UV^H + VU^H = 0$ as

$$\begin{bmatrix} U & \tilde{U} \end{bmatrix} \begin{bmatrix} U & \tilde{U} \end{bmatrix}^H = \begin{bmatrix} V & 0 \end{bmatrix} \begin{bmatrix} V & 0 \end{bmatrix}^H \tag{C.5}$$

and

$$\begin{bmatrix} U & \tilde{U} \end{bmatrix} \begin{bmatrix} V & 0 \end{bmatrix}^H + \begin{bmatrix} V & 0 \end{bmatrix} \begin{bmatrix} U & \tilde{U} \end{bmatrix}^H = 0.$$

It follows from part (b) that there exists a unitary skew-Hermitian matrix $\tilde{\Lambda}$ for which

$$\begin{bmatrix} U & \tilde{U} \end{bmatrix} = \begin{bmatrix} V & 0 \end{bmatrix} \begin{bmatrix} \tilde{\Lambda}_{11} & \tilde{\Lambda}_{12} \\ \tilde{\Lambda}_{21} & \tilde{\Lambda}_{22} \end{bmatrix}.$$

The subblock $\Lambda = \tilde{\Lambda}_{11}$ satisfies $U = V\Lambda$, $\Lambda + \Lambda^H = 0$ and $\Lambda^H\Lambda \preceq I$. □

The following small example illustrates the use of the proof of part (b) in Lemma C.1. Consider $r = 3$, $q = \mathbf{rank}(U) = \mathbf{rank}(V) = 2$, $Q_v = I$ and,

$$\tilde{\Lambda} = Q_u^H = \begin{bmatrix} 0.8\mathrm{j} & 0 & 0.6\mathrm{j} \\ 0 & \mathrm{j} & 0 \\ 0.6 & 0 & -0.8 \end{bmatrix},$$

which is unitary but not skew-Hermitian. Then $\Delta_1 = 0.8\mathrm{j}$, $\Delta_2 = \mathrm{j}$, and

$$\Omega = (I - \Delta_1\Delta_1^H)^{-1/2} Q_1^H \tilde{\Lambda}_{12} = (1 - 0.64)^{-1/2} \begin{bmatrix} 1 \\ 0 \end{bmatrix}^H \begin{bmatrix} 0.6\mathrm{j} \\ 0 \end{bmatrix} = \mathrm{j}.$$

The resulting Λ is therefore

$$\Lambda = Q_v \begin{bmatrix} \tilde{\Lambda}_{11} & \tilde{\Lambda}_{12} \\ -\tilde{\Lambda}_{12}^H & \Omega^H \Delta_1^H \Omega \end{bmatrix} Q_v^H = \begin{bmatrix} 0.8\mathrm{j} & 0 & 0.6\mathrm{j} \\ 0 & \mathrm{j} & 0 \\ 0.6\mathrm{j} & 0 & -0.8\mathrm{j} \end{bmatrix}.$$

Proof of Lemma 2.3. Suppose U and V are $p \times r$ matrices that satisfy (2.12) and (2.13). As explained in appendix B, there exists a nonsingular R such that

$$\Phi = R^H \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} R, \qquad \Psi = R^H \begin{bmatrix} \alpha & \beta \\ \beta & \gamma \end{bmatrix} R$$

with β real and $\gamma \leq \alpha$. Define $S = R_{11}U + R_{12}V$ and $T = R_{21}U + R_{22}V$. Then

$$ST^H + TS^H = \begin{bmatrix} U & V \end{bmatrix} \begin{bmatrix} \Phi_{11}I & \Phi_{21}I \\ \Phi_{12}I & \Phi_{22}I \end{bmatrix} \begin{bmatrix} U^H \\ V^H \end{bmatrix} = 0 \tag{C.6}$$

and

$$\alpha SS^H + \gamma TT^H = \begin{bmatrix} U & V \end{bmatrix} \begin{bmatrix} \Psi_{11}I & \Psi_{21}I \\ \Psi_{12}I & \Psi_{22}I \end{bmatrix} \begin{bmatrix} U^H \\ V^H \end{bmatrix} \preceq 0. \tag{C.7}$$

We show that this implies that $S = W\,\mathbf{diag}(s)Q^H$, $T = W\,\mathbf{diag}(t)Q^H$, for some $W \in \mathbf{C}^{p\times r}$, unitary $Q \in \mathbf{C}^{r\times r}$, and vectors $s, t \in \mathbf{C}^r$ that satisfy

$$s_i\bar{t}_i + \bar{s}_i t_i = 0, \qquad \alpha|s_i|^2 + \gamma|t_i|^2 \leq 0, \qquad (s_i, t_i) \neq 0, \qquad i = 1, \ldots, r.$$

The result is trivial if S and T are zero, since in that case we can choose $W = 0$, and arbitrary Q, s, t. If at least one of the two matrices is nonzero, then (C.7), combined with $\alpha \geq \gamma$, implies that $\gamma \leq 0$. Therefore there are three cases to consider.

- If $\alpha \leq 0$, we write (C.6) as $(S+T)(S+T)^H = (S-T)(S-T)^H$. From Lemma C.1, $S + T = (S - T)\Lambda$ with Λ unitary. Let $\Lambda = Q\,\mathbf{diag}(\rho)Q^H$ be the Schur decomposition of Λ, with $|\rho_i| = 1$ for $i = 1, \ldots, r$. Define $W = (S-T)Q$, $s = (1/2)(\rho + \mathbf{1})$, $t = (1/2)(\rho - \mathbf{1})$.

- If $\gamma = 0 < \alpha$, then $S = 0$, and we can take $Q = I$, $W = T$, $s = 0$, $t = \mathbf{1}$.

- If $\gamma < 0 < \alpha$, then from Lemma C.1, $S = (-\gamma/\alpha)^{1/2} T\Lambda$ for some skew-Hermitian Λ with $\Lambda^H \Lambda \preceq I$. This matrix has a Schur decomposition $\Lambda = Q\,\mathbf{diag}(\rho) Q^H$ with $|\rho_i| \leq 1$ for $i = 1, \ldots, r$. Define $W = TQ$, $s = (-\gamma/\alpha)^{1/2}\rho$, and $t = \mathbf{1}$.

The factorizations of U and V now follow from

$$\begin{bmatrix} U \\ V \end{bmatrix} = (R^{-1} \otimes I) \begin{bmatrix} S \\ T \end{bmatrix} = (R^{-1} \otimes I) \begin{bmatrix} W\,\mathbf{diag}(s) \\ W\,\mathbf{diag}(t) \end{bmatrix} Q^H = \begin{bmatrix} W\,\mathbf{diag}(\mu) \\ W\,\mathbf{diag}(\nu) \end{bmatrix} Q^H$$

where μ and ν are defined as

$$\begin{bmatrix} \mu_i \\ \nu_i \end{bmatrix} = R^{-1} \begin{bmatrix} s_i \\ t_i \end{bmatrix}, \quad i = 1, \ldots, r.$$

These pairs (μ_i, ν_i) are nonzero and satisfy

$$\begin{bmatrix} \mu_i \\ \nu_i \end{bmatrix}^H \Phi \begin{bmatrix} \mu_i \\ \nu_i \end{bmatrix} = \begin{bmatrix} s_i \\ t_i \end{bmatrix}^H \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} s_i \\ t_i \end{bmatrix} = \bar{s}_i t_i + s_i \bar{t}_i = 0$$

and

$$\begin{bmatrix} \mu_i \\ \nu_i \end{bmatrix}^H \Psi \begin{bmatrix} \mu_i \\ \nu_i \end{bmatrix} = \begin{bmatrix} s_i \\ t_i \end{bmatrix}^H \begin{bmatrix} \alpha & \beta \\ \beta & \gamma \end{bmatrix} \begin{bmatrix} s_i \\ t_i \end{bmatrix} = \alpha |s_i|^2 + \beta(\bar{s}_i t_i + s_i \bar{t}_i) + \gamma |t_i|^2 \leq 0.$$

□

APPENDIX D

Strict feasibility

In this appendix from [CV17] we discuss strict feasibility of the constraints $X \succeq 0$, (2.14), (2.15) in Theorem 2.5. We assume that the set $\mathcal{C}$ defined in (2.8) is not empty and not a singleton. This means that if the inequality $q_\Psi(\mu,\nu) \leq 0$ in the definition is not redundant, then there exist points in $\mathcal{C}$ with $q_\Psi(\mu,\nu) < 0$. We will distinguish these two cases.

- *Line or circle.* If the inequality $q_\Psi(\mu,\nu) \leq 0$ is redundant, we have $\mathcal{C} = \{(\mu,\nu) \in \mathbf{C}^2 \mid (\mu,\nu) \neq 0,\ q_\Phi(\mu,\nu) = 0\}$, a line or circle in homogeneous coordinates. In this case we understand by strict feasibility of X that

$$X \succ 0, \quad \Phi_{11} F X F^H + \Phi_{21} F X G^H + \Phi_{12} G X F^H + \Phi_{22} G X G^H = 0. \tag{D.1}$$

  We also define $\mathcal{C}^\circ = \mathcal{C}$.

- *Segment of line or circle.* In the second case, $\mathcal{C}$ is a proper one-dimensional subset of the line or circle defined by $q_\Phi(\mu,\nu) = 0$. In this case we define strict feasibility of X as

$$\text{(D.1)}, \quad \Psi_{11} F X F^H + \Psi_{21} F X G^H + \Psi_{12} G X F^H + \Psi_{22} G X G^H \prec 0. \tag{D.2}$$

  We also define $\mathcal{C}^\circ = \{(\mu,\nu) \neq 0 \mid q_\Phi(\mu,\nu) = 0,\ q_\Psi(\mu,\nu) < 0\}$.

The conditions on F and G that guarantee strict feasibility will be expressed in terms of the Kronecker structure of the matrix pencil $\lambda G - F$ [Gan05, Van79]. For every matrix pencil

there exist nonsingular matrices P and Q such that

$$
\begin{aligned}
&P(\lambda G - F)Q \\
&= \begin{bmatrix}
L_{\eta_1}(\lambda)^T & 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 \\
0 & L_{\eta_2}(\lambda)^T & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \vdots & & \vdots \\
0 & 0 & \cdots & L_{\eta_l}(\lambda)^T & 0 & 0 & 0 & \cdots & 0 \\
0 & 0 & \cdots & 0 & \lambda B - A & 0 & 0 & \cdots & 0 \\
0 & 0 & \cdots & 0 & 0 & L_{\epsilon_1}(\lambda) & 0 & \cdots & 0 \\
0 & 0 & \cdots & 0 & 0 & 0 & L_{\epsilon_2}(\lambda) & \cdots & 0 \\
\vdots & \vdots & & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & L_{\epsilon_r}(\lambda)
\end{bmatrix}
\end{aligned}
\tag{D.3}
$$

where $L_\epsilon(\lambda)$ is the $\epsilon \times (\epsilon + 1)$ pencil

$$
L_\epsilon(\lambda) = \begin{bmatrix}
\lambda & -1 & 0 & \cdots & 0 & 0 \\
0 & \lambda & -1 & \cdots & 0 & 0 \\
\vdots & \vdots & \vdots & & \vdots & \vdots \\
0 & 0 & 0 & \cdots & -1 & 0 \\
0 & 0 & 0 & \cdots & \lambda & -1
\end{bmatrix},
$$

and $\lambda B - A$ is a regular pencil, *i.e.*, it is square and $\det(\lambda B - A)$ is not identically zero. The parameters $\epsilon_1, \ldots, \epsilon_r$, and $\eta_1, \ldots, \eta_l$ are the right and the left Kronecker indices of the pencil. The *normal rank* of the pencil is $p - l$, where p is the row dimension.

We show that there exists a strictly feasible X if and only if the following two conditions hold.

1. The normal rank of $\lambda G - F$ is p. This means that $l = 0$ in (D.3).

2. The generalized eigenvalues of $\lambda B - A$ are nondefective and lie in $\mathcal{C}^\circ$. (More accurately, if λ is a finite generalized eigenvalue, then $(\lambda, 1) \in \mathcal{C}^\circ$. If it is an infinite generalized eigenvalue, then $(1, 0) \in \mathcal{C}^\circ$.).

A sufficient but more easily verified condition is that $\mathbf{rank}\,(\mu G - \nu F) = p$ for all $(\mu, \nu) \neq 0$, *i.e.*, $l = 0$ and the block $\lambda B - A$ in (D.3) is not present.

Proof. Without loss of generality we assume that the pencil is in the Kronecker canonical form ($P = I$, $Q = I$ in (D.3)) and $\Phi = \Phi_{\mathrm{u}}$, so the equality in (D.1) is

$$FXF^H = GXG^H. \tag{D.4}$$

We first show that the conditions are necessary. Assume X is strictly feasible. Partition X as an $(l+1+r) \times (l+1+r)$ block matrix, with block dimensions equal to the column dimensions of the $l+1+r$ block columns in (D.3). Suppose $l \geq 1$ and consider the kth diagonal block X_{kk} with $1 \leq k \leq l$. The kth diagonal block of the pencil is

$$\lambda G_k - F_k = L_{\eta_k}(\lambda)^T = \lambda \begin{bmatrix} I_{\eta_k} \\ 0_{1\times\eta_k} \end{bmatrix} - \begin{bmatrix} 0_{1\times\eta_k} \\ I_{\eta_k} \end{bmatrix}.$$

The kth diagonal block of (D.4) is $F_k X_{kk} F_k^H = G_k X_{kk} G_k^H$ or

$$\begin{bmatrix} 0_{1\times\eta_k} \\ I_{\eta_k} \end{bmatrix} X_{kk} \begin{bmatrix} 0_{\eta_k\times 1} & I_{\eta_k} \end{bmatrix} = \begin{bmatrix} I_{\eta_k} \\ 0_{1\times\eta_k} \end{bmatrix} X_{kk} \begin{bmatrix} I_{\eta_k} & 0_{\eta_k\times 1} \end{bmatrix}.$$

This is impossible since $X_{kk} \succ 0$. Hence, if (D.4) holds with $X \succ 0$, then $l = 0$.

Next suppose $\det(\mu B - \nu A) = 0$ for some $(\mu, \nu) \neq 0$. If $\nu \neq 0$, then μ/ν is a finite generalized eigenvalue of the pencil $\lambda B - A$; if $\nu = 0$ then the pencil has a generalized eigenvalue at infinity. Let y be a corresponding left generalized eigenvector, *i.e.*, $y^H(\mu B - \nu A) = 0$, while $y^H B$ and $y^H A$ are not both zero (since $y^H B = y^H A = 0$ would imply that the pencil $\lambda B - A$ is singular). Define $u^H = y^H B$ if $\nu \neq 0$ and $u^H = y^H A$ otherwise. This is a nonzero vector. The first diagonal block of (D.4) is

$$AX_{11}A^H = BX_{11}B^H. \tag{D.5}$$

From this it follows that $|\mu|^2 u^H X_{11} u = |\nu|^2 u^H X_{11} u$, and, since $X_{11} \succ 0$, we have $q_\Phi(\mu, \nu) = |\mu|^2 - |\nu|^2 = 0$, *i.e.*, the generalized eigenvalues are on the unit circle. In addition, if the inequality in (D.2) holds, then

$$\Psi_{11} A X_{11} A^H + \Psi_{21} A X_{11} B^H + \Psi_{12} B X_{11} A^H + \Psi_{22} B X_{11} B^H \prec 0$$

and from this, $q_\Psi(\mu,\nu)(u^H X_{11} u) < 0$. This is only possible if $q_\Psi(\mu,\nu) < 0$. We conclude that if $\det(\mu B - \nu A) = 0$ for nonzero (μ,ν), then $(\mu,\nu) \in \mathcal{C}^\circ$.

Next we show that the generalized eigenvalues of the pencil $\lambda B - A$ are nondefective. Since C° is the unit circle or a subset of the unit circle, there are no infinite generalized eigenvalues. Assume the pencil is in Weierstrass canonical form, *i.e.*,

$$\lambda B - A = \begin{bmatrix} (\lambda-\rho_1)I - J_{s_1} & 0 & \cdots & 0 \\ 0 & (\lambda-\rho_2)I - J_{s_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & (\lambda-\rho_t)I - J_{s_t} \end{bmatrix},$$

where $\rho_1, \ldots, \rho_t$ are the generalized eigenvalues (which satisfy $|\rho_i| = 1$), and J_s is the $s \times s$ matrix

$$J_s = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & 0 & \cdots & 0 & 0 \end{bmatrix}.$$

Then (D.5) implies that

$$(\rho_i I + J_{s_i}) X_{11,i} (\rho_i I + J_{s_i})^H = X_{11,i}$$

where $X_{11,i}$ is the ith diagonal block of X_{11}, if we partition X_{11} as a $t \times t$ block matrix with i,j block of size of $s_i \times s_j$. Expanding this gives

$$|\rho_i|^2 X_{11,i} + \rho_i X_{11,i} J_{s_i}^T + \bar{\rho}_i J_{s_i} X_{11,i} + J_{s_i} X_{11,i} J_{s_i}^T = X_{11,i}.$$

Since $|\rho_i| = 1$ this simplifies to

$$\rho_i X_{11,i} J_{s_i}^T + \bar{\rho}_i J_{s_i} X_{11,i} + J_{s_i} X_{11,i} J_{s_i}^T = 0.$$

The last rows of the second and third matrices are zero. Therefore the last row of the first matrix is zero. However the element in column $s_i - 1$ is the last diagonal element of the

positive definite matrix $X_{11,i}$. This is a contradiction unless $s_i = 1$, *i.e.*, the eigenvalue ρ_i is nondefective. We conclude that the two conditions are necessary.

It remains to show sufficiency. Suppose $\lambda G - F$ has the Kronecker canonical form

$$\lambda G - F = \begin{bmatrix} \lambda - \rho_1 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & & \vdots \\ 0 & \cdots & \lambda - \rho_t & 0 & \cdots & 0 \\ 0 & \cdots & 0 & L_{\epsilon_1}(\lambda) & \cdots & 0 \\ \vdots & & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & L_{\epsilon_r}(\lambda) \end{bmatrix}$$

with $\rho_i \in \mathcal{C}^\circ$. Define a block diagonal matrix

$$X = \begin{bmatrix} 1 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & & \vdots \\ 0 & \cdots & 1 & 0 & \cdots & 0 \\ 0 & \cdots & 0 & X_{11} & \cdots & 0 \\ \vdots & & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & X_{rr} \end{bmatrix}$$

with diagonal blocks

$$X_{kk} = \sum_{i=1}^{\epsilon_k+1} \begin{bmatrix} 1 \\ \lambda_{ki} \\ \vdots \\ \lambda_{ki}^{\epsilon_k} \end{bmatrix} \begin{bmatrix} 1 \\ \lambda_{ki} \\ \vdots \\ \lambda_{ki}^{\epsilon_k} \end{bmatrix}^H, \qquad k = 1, \ldots, r,$$

where $\lambda_{k1}, \ldots, \lambda_{k,\epsilon_k+1}$ are distinct and in $\mathcal{C}^\circ$. This matrix X is strictly feasible. □

APPENDIX E

Generalized proximal gradient method

The appendix is from [CV18] and describes the accelerated proximal gradient method used in the experiments, including a convergence proof. The proof follows [Tse08] and is included to clarify where our assumptions on the problem and the Bregman distance are needed. These conditions are slightly weaker than the ones stated in [Tse08, page 17]. The proof also justifies the third parameter selection strategy discussed below.

We consider an optimization problem

$$\text{minimize} \quad F(x) = f(x) + g(x), \tag{E.1}$$

in which the objective is split as a sum of two convex functions. We assume that $\emptyset \neq \mathbf{dom}\, g \subseteq \mathbf{dom}\, f$ and that f is differentiable with a Lipschitz continuous gradient on $\mathbf{dom}\, g$, *i.e.*, there exists a constant L such that

$$f(x) \leq f(y) + \langle \nabla f(y), x - y \rangle + \frac{L}{2}\|x - y\|^2 \tag{E.2}$$

for all $x, y \in \mathbf{dom}\, g$. In addition, we assume that d_h is a Bregman distance with kernel h, and that for every a and every $v \in \operatorname{int}(\mathbf{dom}\, h)$, the generalized proximal operator $\operatorname{prox}^h_{\tau g}(\tau a, v)$ defined in (6.8), is well defined, *i.e.*, the optimization problem

$$\text{minimize} \quad \langle a, x \rangle + g(x) + \frac{1}{\tau} d_h(x, v) \tag{E.3}$$

has a unique solution in $\mathbf{dom}\, g \cap \operatorname{int}(\mathbf{dom}\, h)$. Here τ is a positive proximal stepsize. We also assume that

$$d_h(x, y) \geq \frac{1}{2}\|x - y\|^2 \tag{E.4}$$

for all $x \in \mathbf{dom}\, g \cap \mathbf{dom}\, h$ and $y \in \mathbf{dom}\, g \cap \operatorname{int}(\mathbf{dom}\, h)$. The norm on the right-hand side of (E.4) is the same norm as in (E.2). Finally, we assume that the problem (E.1) is solvable and has a solution $x^\star \in \mathbf{dom}\, g \cap \mathbf{dom}\, h$.

The following algorithm is IGA in [AT06] and Algorithm 1 in [Tse08]. We start at $x^0 = v^0 \in \mathbf{dom}\, g \cap \mathrm{int}\,(\mathbf{dom}\, h)$ and run the iteration

$$y^k = (1-\theta_k)x^{k-1} + \theta_k v^{k-1} \tag{E.5a}$$

$$v^k = \mathrm{prox}^h_{\tau_k g}(\tau_k \nabla f(y^k), v^{k-1}) \tag{E.5b}$$

$$x^k = (1-\theta_k)x^{k-1} + \theta_k v^k. \tag{E.5c}$$

Suitable choices for the parameters $\theta_k \in [0,1]$ and $\tau_k > 0$ are discussed below. Since the minimizer v^k in step (E.5b) is in the convex set $\mathbf{dom}\, g \cap \mathrm{int}\,(\mathbf{dom}\, h)$, all iterates y^k, v^k, x^k are in $\mathbf{dom}\, g \cap \mathrm{int}\,(\mathbf{dom}\, h)$. The update in the second step (E.5b) is therefore well defined at all iterations.

We discuss three strategies for choosing θ_k and τ_k. The first option requires knowledge of L, the Lipschitz constant in (E.2) with respect to a norm that also satisfies (E.4). Several strategies have been proposed to avoid this and replace L with an adaptively adjusted estimate λ_k [BT09a, Gul92, Tse08, Nes13, SGB14]. The second and third methods below are examples of this.

In each method we will choose $\theta_1 = 1$, $\theta_k \in (0,1)$ for $k > 1$, and $\tau_k > 0$ subject to the two conditions:

$$\begin{aligned} F(x^k) \leq (1-\theta_k)F(x^{k-1}) + \theta_k(g(v^k) + f(y^k) \\ + \langle \nabla f(y^k), v^k - y^k \rangle + \frac{1}{\tau_k} d_h(v^k, v^{k-1})), \end{aligned} \tag{E.6}$$

and

$$\tau_k(1-\theta_k)\theta_{k-1} \leq \tau_{k-1}\theta_k. \tag{E.7}$$

We will see that these conditions imply that

$$F(x^k) - F(x^\star) \leq \frac{\theta_k}{\tau_k} d_h(x^\star, x^0). \tag{E.8}$$

Each of the following three parameter selection methods satisfies (E.6) and (E.7), with $\theta_k/\tau_k = O(1/k^2)$.

Known Lipschitz constant We choose $\tau_k = 1/(L\theta_k)$ and a sequence θ_k that satisfies $\theta_1 = 1$ and

$$(1-\theta_k)\theta_{k-1}^2 \leq \theta_k^2, \quad k > 1. \tag{E.9}$$

A simple choice is $\theta_k = 2/(k+1)$. The sequence that decreases most quickly, subject to the constraint (E.9), is obtained by imposing equality in (E.7). This gives the recursion

$$\theta_k = \frac{-\theta_{k-1}^2 + \sqrt{\theta_{k-1}^4 + 4\theta_{k-1}^2}}{2}.$$

To show that (E.6) holds, we apply (E.2) with $x = x^k$ and $y = y^k$, substitute (E.5c) for x^k on the right-hand side, simplify the argument of the norm using (E.5a), and apply (E.4) to obtain

$$\begin{aligned} f(x^k) \leq{}& (1-\theta_k)(f(y^k) + \langle \nabla f(y^k), x^{k-1} - y^k \rangle) \\ &+ \theta_k(f(y^k) + \langle \nabla f(y^k), v^k - y^k \rangle + \frac{1}{\tau_k} d(v^k, v^{k-1})). \end{aligned}$$

The inequality (E.6) now follows from convexity of f and Jensen's inequality for g applied to (E.5c).

Monotonic search This is the strategy of [Tse08, BT09a]. We choose a fixed sequence θ_k that satisfies (E.9), as in the previous strategy. We choose $\lambda_0 > 0$, and at iteration k choose for λ_k the smallest element of $\{\beta^i \lambda_{k-1} \mid i = 0, 1, 2, \ldots\}$, for which $\tau_k = 1/(\lambda_k \theta_k)$ satisfies (E.6). Here $\beta > 1$.

The inequality (E.7) holds because

$$\tau_k \theta_k = 1/\lambda_k \leq 1/\lambda_{k-1} = \tau_{k-1}\theta_{k-1}$$

and (E.9) holds. The procedure guarantees that $\lambda_k \leq \lambda_{\max} = \max\{\lambda_0, \beta L\}$ because, as shown above, (E.6) holds for $\theta_k \tau_k \leq 1/L$. Therefore

$$\frac{\theta_k}{\tau_k} \leq \theta_k^2 \lambda_{\max} \leq \frac{4\lambda_{\max}}{(k+1)^2} = O(\frac{1}{k^2}).$$

In this method, testing a candidate λ_k requires the evaluation of the generalized proximal operator in step (E.5b), and evaluations of $f(x^k)$, $g(x^k)$, $g(v^k)$, and $d_h(v^k, v^{k-1})$. These function values are needed to verify whether the inequality (E.6) holds.

Non-monotonic search The third method does not force λ_k to be monotonically increasing as in the second method. At each iteration, choose some $\hat{\lambda}_k > 0$, and take the smallest λ_k in $\{\beta^i \hat{\lambda}_k \mid i = 0, 1, 2, \ldots\}$ that satisfies (E.6) with θ_k defined as the positive root of

$$\lambda_k \theta_k^2 = \lambda_{k-1} \theta_{k-1}^2 (1 - \theta_k),$$

and $\tau_k = 1/(\theta_k \lambda_k)$.

Lipschitz continuity of ∇f guarantees that (E.6) holds if $\lambda_k = 1/(\theta_k \tau_k) \geq L$. Hence the selected parameter satisfies $\lambda_k \leq \max\{\hat{\lambda}_k, \beta L\}$. The second condition (E.7) is satisfied by construction of θ_k. Finally, it can be shown that $\theta_k/\tau_k = O(1/k^2)$ [Gul92, lemma 2.2]. The steps in this method are more expensive than in the second method. When testing a candidate λ_k, we also change θ_k and therefore y^k, so we need to recompute $f(y^k)$ and $\nabla f(y^k)$.

We now prove the inequality (E.8). We will need the following lemma [Tse08, proposition 1]. If $\hat{x} \in \operatorname{int}(\mathbf{dom}\, h)$ is a solution of (E.3), then for all $x \in \mathbf{dom}\, g \cap \mathbf{dom}\, h$,

$$\begin{aligned} &\langle a, \hat{x} \rangle + g(\hat{x}) - \langle a, x \rangle - g(x) \\ &\quad \leq \frac{1}{\tau}(d_h(x, v) - d_h(\hat{x}, v) - d_h(x, \hat{x})). \end{aligned} \tag{E.10}$$

Suppose (E.6) holds. By definition, v^k satisfies an inequality of the form (E.10), *i.e.*, for $x \in \mathbf{dom}\, g \cap \mathbf{dom}\, h$,

$$\begin{aligned} &\langle \nabla f(y^k), v^k \rangle + g(v^k) - \langle \nabla f(y^k), x \rangle - g(x) \\ &\leq \frac{1}{\tau_k}(d_h(x, v^{k-1}) - d_h(v^k, v^{k-1}) - d_h(x, v^k)). \end{aligned}$$

Evaluating this at $x = x^\star$ and combining the result with (E.6) gives

$$\begin{aligned} F(x^k) - (1 - \theta_k) F(x^{k-1}) + \frac{\theta_k}{\tau_k}(d_h(x^\star, v^k) - d_h(x^\star, v^{k-1})) \\ \leq \; \theta_k (f(y^k) + \langle \nabla f(y^k), x^\star - y^k \rangle + g(x^\star)) \\ \leq \; \theta_k F(x^\star). \end{aligned}$$

Re-arranging gives

$$\begin{aligned} &\frac{\tau_k}{\theta_k}(F(x^k) - F(x^\star)) + d_h(x^\star, v^k) \\ &\quad \leq \; \frac{(1 - \theta_k)\tau_k}{\theta_k}(F(x^{k-1}) - F(x^\star)) + d_h(x^\star, v^{k-1}). \end{aligned}$$

Combining these inequalities recursively using (E.7) gives (E.8).

[AG88] G. S. Ammar and W. B. Gragg. "Superfast solution of real positive definite Toeplitz systems." *SIAM Journal on Matrix Analysis and Applications*, **9**(1):61–76, 1988.

[ALH17] M. Annergren, C. A. Larsson, H. Hjalmarsson, X. Bombois, and B. Wahlberg. "Application-oriented input design in system identification optimal input design for control." *IEEE Control Systems Magazine*, **37**:31–56, 2017.

[AM79] B. Anderson and J. B. Moore. *Optimal Filtering*. Prentice-Hall, 1979.

[AT06] A. Auslender and M. Teboulle. "Interior gradient and proximal methods for convex and conic optimization." *SIAM Journal on Optimization*, **16**(3):697–725, 2006.

[AV00] B. Alkire and L. Vandenberghe. "Handling nonnegative constraints in spectral estimation." In *Proceedings of the 34th Asilomar Conference on Signals, Systems, and Computers*, pp. 202–206, 2000.

[AV02] B. Alkire and L. Vandenberghe. "Convex optimization problems involving finite autocorrelation sequences." *Mathematical Programming Series A*, **93**:331–359, 2002.

[BA80] R. R. Bitmead and B. D. O. Anderson. "Asymptotically fast solution of Toeplitz and related systems of linear equations." *Linear Algebra and Its Applications*, **34**:103–116, 1980.

[Bar83] A. Barabell. "Improving the resolution performance of eigenstructure-based direction-finding algorithms." In *Acoustics, Speech, and Signal Processing, IEEE International Conference on ICASSP '83.*, volume 8, pp. 336–339, 1983.

[Bar07] R. Baraniuk. "Compressive sensing." *IEEE Signal Processing Magazine*, **24**(4):118–121, 2007.

[Bas89] M. Basseville. "Distance measures for signal processing and pattern recognition." *Signal Processing*, **18**:349–369, 1989.

[Bas13] M. Basseville. "Divergence measures for statistical data processing—An annotated bibliography." *Signal Processing*, **93**:621–633, 2013.

[BB71] A. Berman and A. Ben-Israel. "More on linear inequalities with applications to matrix theory." *Journal of Mathematical Analysis and Applications*, **33**:482–496, 1971.

[BD86] J. P. Boyle and R. L. Dykstra. "A method for finding projections onto the intersection of convex sets in Hilbert spaces." In R. Dykstra, T. Robertson, and F. T. Wright, editors, *Advances in Order Restricted Statistical Inference*, volume 37 of *Lecture Notes in Statistics*, pp. 28–47. Springer-Verlag, 1986.

[Ben69] A. Ben-Israel. "Linear equations and inequalities on finite dimensional, real or complex vector spaces: a unified theory." *Journal of Mathematical Analysis and Applications*, **27**:367–389, 1969.

[BGL98] C. I. Byrnes, S. V. Gusev, and A. Lindquist. "A convex optimization approach to the rational covariance extension problem." *SIAM Journal on Control and Optimization*, **37**(1):211–229, 1998.

[BGL01] C. I. Byrnes, S. V. Gusev, and A. Lindquist. "From finite covariance windows to modeling filters: a convex optimization approach." *SIAM Review*, **4**(4):645–675, 2001.

[BGY80] R. P. Brent, F. G. Gustavson, and D. Y. Y. Yun. "Fast solution of Toeplitz systems of equations and computation of Padé approximants." *Journal of Algorithms*, **1**:259–195, 1980.

[Bie79] G. Bienvenu. "Influence of the spatial coherence of the background noise on high resolution passive methods." In *IEEE International Conference on Acoustics, Speech, and Signal Processing*, volume 4, pp. 306–309, 1979.

[BL00] H. H. Bauschke and A. S. Lewis. "Dykstra's algorithm with Bregman projections: A convergence proof." *Optimization: A Journal of Mathematical Programming and Operations Research*, **48**(4):409–427, 2000.

[BN01] A. Ben-Tal and A. Nemirovski. *Lectures on Modern Convex Optimization. Analysis, Algorithms, and Engineering Applications*. SIAM, 2001.

[BPC11] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein. "Distributed optimization and statistical learning via the alternating direction method of multipliers." *Foundations and Trends in Machine Learning*, **3**(1):1–122, 2011.

[BR11] B. N. Bhaskar and B. Recht. "Atomic norm denoising with applications to line spectral estimation." In *Communication, Control, and Computing (Allerton), 2011 49th Annual Allerton Conference on*, pp. 261–268, 2011.

[BT03] A. Beck and M. Teboulle. "Mirror descent and nonlinear projected subgradient methods for convex optimization." *Operations Research Letters*, **31**:167–175, 2003.

[BT09a] A. Beck and M. Teboulle. "A fast iterative shrinkage-thresholding algorithm for linear inverse problems." *SIAM Journal on Imaging Sciences*, **2**(1):183–202, 2009.

[BT09b] A. Beck and M. Teboulle. "Gradient-based algorithms with applications to signal recovery." In Y. Eldar and D. Palomar, editors, *Convex Optimization in Signal Processing and Communications*. Cambridge University Press, 2009.

[BTR13] B. N. Bhaskar, G. Tang, and B. Recht. "Atomic norm denoising with applications to line spectral estimation." *IEEE Transactions on Signal Processing*, **61**(23):5987–5999, 2013.

[BV02] V. Balakrishnan and L. Vandenberghe. "Semidefinite Programming Duality and Linear Time-Invariant Systems." Technical Report TR-ECE-02-02, School of Electrical and Computer Engineering, Purdue University, 2002.

[BV03] V. Balakrishnan and L. Vandenberghe. "Semidefinite programming duality and linear time-invariant systems." *IEEE Transactions on Automatic Control*, **48**:30–41, 2003.

[BV04] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, Cambridge, 2004.

[BW81] J. Borwein and H. Wolkowicz. "Regularizing the Abstract Convex Program." *Journal of Mathematical Analysis and Applications*, **83**:495–530, 1981.

[CC15] Y. Chi and Y. Chen. "Compressive two-dimensional harmonic retrieval via atomic norm minimization." *IEEE Transactions on Signal Processing*, **63**(4):1030–1042, 2015.

[CDS98] S. S. Chen, D. L. Donoho, and M. A. Saunders. "Atomic decomposition by basis pursuit." *SIAM Journal on Scientific Computing*, **20**:33–61, 1998.

[CF13] E. J. Candès and C. Fernandez-Granda. "Super-resolution from noisy data." *Journal of Fourier Analysis and Applications*, **19**:1229–1254, 2013.

[CF14] E. J. Candès and C. Fernandez-Granda. "Towards a mathematical theory of super-resolution." *Communications of Pure and Applied Mathematics*, **67**(6):906–956, 2014.

[CG12] Y. de Castro and F. Gamboa. "Exact reconstruction using Beurling minimal extrapolation." *Journal of Mathematical Analysis and Applications*, **395**(1):336 – 354, 2012.

[CGH17] Y. de Castro, F. Gamboa, D. Henrion, and J. B. Lasserre. "Exact solutions to super resolution on semi-algebraic domains in higher dimensions." *IEEE Transactions on Information Theory*, **63**(1):621–630, 2017.

[CGT00] A. R. Conn, N. I. M. Gould, and Ph. L. Toint. *Trust-Region Methods*. SIAM, 2000.

[CK77] B. D. Craven and J. J. Koliha. "Generalizations of Farkas' theorem." *SIAM Journal on Mathematical Analysis*, **8**(6):983–997, 1977.

[CM73] J. F. Claerbout and F. Muir. "Robust modeling with erratic data." *Geophysics*, **38**(5):826–844, 1973.

[CP07] P. L. Combettes and J.-C. Pesquet. "A Douglas-Rachford splitting approach to nonsmooth convex variational signal recovery." *IEEE Journal of Selected Topics in Signal Processing*, **1**(4):564–574, 2007.

[CRP12] V. Chandrasekaran, B. Recht, P. A. Parrilo, and A. S. Willsky. "The convex geometry of linear inverse problems." *Foundations of Computational Mathematics*, **12**:805–849, 2012.

[CRT06] E. J. Candès, J. Romberg, and T. Tao. "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information." *IEEE Transactions on Information Theory*, **52**(2):489–509, 2006.

[CSP11] Y. Chi, L. L. Scharf, A. Pezeshki, and A. R. Calderbank. "Sensitivity to basis mismatch in compressed sensing." *IEEE Transactions on Signal Processing*, **59**(5):2182–2195, 2011.

[CV16] H.-H. Chao and L. Vandenberghe. "Extensions of semidefinite programming methods for atomic decomposition." In *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2016.

[CV17] H.-H. Chao and L. Vandenberghe. "Semidefinite representations of gauge functions for structured low-rank matrix decomposition." *SIAM Journal on Optimization*, **27**:1362–1389, 2017.

[CV18] H.-H. Chao and L. Vandenberghe. "Entropic proximal operators for nonnegative trigonometric polynomials." 2018. Submitted for publication.

[CW08] E. J. Candès and M. B. Wakin. "An introduction to compressive sampling." *IEEE Signal Processing Magazine*, **25**(2):21–30, 2008.

[CZ97] Y. Censor and S. A. Zenios. *Parallel Optimization: Theory, Algorithms, and Applications*. Numerical Mathematics and Scientific Computation. Oxford University Press, New York, 1997.

[DB16] S. Diamond and S. Boyd. "CVXPY: A Python-Embedded Modeling Language for Convex Optimization." *Journal of Machine Learning Research*, **17**(83):1–5, 2016.

[DE06] D. L. Donoho and M. Elad. "On the stability of the basis pursuit in the presence of noise." *Signal Processing*, **86**(3):511 – 532, 2006.

[DLS02] T. N. Davidson, Z.-Q. Luo, and J. F. Sturm. "Linear matrix inequality formulation of spectral mask constraints." *IEEE Transactions on Signal Processing*, **50**(11):2702–2715, 2002.

[DM90] C. J. Demeure and C. T. Mullis. "A Newton-Raphson method for moving-average spectral factorization using the Euclid algorithm." *IEEE Transactions on Acoustics, Speech, and Signal Processing*, **18**:1697–1709, 1990.

[Don06] D. L. Donoho. "Compressed sensing." *IEEE Transactions on Information Theory*, **52**(4):1289–1306, 2006.

[DTS01] B. Dumitrescu, I. Tabus, and P. Stoica. "On the parametrization of positive real sequences and MA parameter estimation." *IEEE Transactions on Signal Processing*, **49**(11):2630–2639, 2001.

[Dum07] B. Dumitrescu. *Positive Trigonometric Polynomials and Signal Processing Applications*. Springer, 2007.

[EK12] Y. C. Eldar and G. Kutyniok, editors. *Compressed Sensing: Theory and Applications*. Cambridge University Press, 2012.

[Ela10] M. Elad. *Sparse and Redundant Representations: From Theory to Applications in Signal and Image Processing*. Springer, 2010.

[Far02] J. Farkas. "Theorie der einfachen Ungleichungen." *Journal fr die reine und angewandte Mathematik*, **124**:1–27, 1902.

[Faz02] M. Fazel. *Matrix Rank Minimization with Applications*. PhD thesis, Stanford University, 2002.

[Fer15] C. Fernandez-Granda. "Super-resolution of point sources via convex programming." 2015. `arXiv:1507.07034`.

[FMP14] M. P. Friedlander, I. Macêdo, and T. K. Pong. "Gauge optimization and duality." *SIAM Journal on Optimization*, **24**(4):1999–2022, 2014.

[FN03] A. Feuer and A. Nemirovski. "On sparse representation in pairs of bases." *IEEE Transactions on Information Theory*, **49**(6):1579–1581, 2003.

[FSR03] S. Fomel, P. Sava, J. Rickett, and J. F. Claerbout. "The Wilson-Burg method of spectral factorization with application to helical filtering." *Geophysical Prospecting*, **51**:409–420, 2003.

[Gan05] F. R. Gantmacher. *Applications of the Theory of Matrices*. Dover Publications, 2005. Originally published in 1959 by Interscience Publishers, Inc., New York.

[GB14] M. Grant and S. Boyd. "CVX: Matlab Software for Disciplined Convex Programming, version 2.1." http://cvxr.com/cvx, March 2014.

[GBG80] R. Gray, A. Buzo, A. Gray, and Y. Matsuyama. "Distortion measures for speech processing." *IEEE Transactions on Acoustics, Speech, and Signal Processing*, **28**(4):367–376, 1980.

[Geo06] T. T. Georgiou. "Decomposition of Toeplitz matrices via convex optimization." *IEEE Signal Processing Letters*, **13**(9):537–540, 2006.

[Geo07] T. T. Georgiou. *Distances Between Time-Series and Their Autocorrelation Statistics*, pp. 113–122. Springer Berlin Heidelberg, Berlin, Heidelberg, 2007.

[GHN03] Y. Genin, Y. Hachez, Yu. Nesterov, and P. Van Dooren. "Optimization problems over positive pseudopolynomial matrices." *SIAM Journal on Matrix Analysis and Applications*, **25**(1):57–79, 2003.

[GK83] G. H. Golub and J. Kautsky. "Calculation of Gauss quadratures with multiple free and fixed knots." *Numerische Mathematik*, **41**:147–163, 1983.

[GKT09] T. T. Georgiou, J. Karlsson, and M. S. Takyar. "Metrics for power spectra: an axiomatic approach." *IEEE Transactions on Signal Processing*, **57**(3):859–867, 2009.

[GL08] T. T. Georgiou and A. Lindquist. "A convex optimization approach to ARMA modeling." *IEEE Transactions on Automatic Control*, **53**:1108–1119, 2008.

[GM76] A. H. Gray and J. D. Markel. "Distance measures for speech processing." *IEEE Transactions on Acoustics, Speech, and Signal Processing*, **ASSP-24**:380–391, 1976.

[GS84] U. Grenander and G. Szegö. *Toeplitz Forms and Their Applications*. Chelsea, New York, 1984. First published in 1958.

[Gul92] O. Güler. "New proximal point algorithm for convex minimization." *SIAM Journal on Optimization*, **2**(4):649–664, 1992.

[GV96] G. H. Golub and C. F. Van Loan. *Matrix Computations*. Johns Hopkins University Press, 3rd edition, 1996.

[Hac03] Y. Hachez. *Convex Optimization over Non-Negative Polynomials: Structured Algorithms and Applications*. PhD thesis, Université catholique de Louvain, 2003.

[Han88] S.-P. Han. "A successive projection method." *Mathematical Programming*, **40**:1–14, 1988.

[Hoo87] F. de Hoog. "A new algorithm for solving Toeplitz systems of equations." *Linear Algebra and Its Applications*, **88/89**:123–138, 1987.

[HS88] Y. Hua and T. K. Sakar. "Matrix pencil method and its performance." *IEEE International Conference on Acoustics, Speech, and Signal Processing*, **4**:2476–2479, 1988.

[HS90] Y. Hua and T. K. Sarkar. "Matrix pencil method for estimating parameters of exponentially damped/undamped sinusoids in noise." *IEEE Transactions on Acoustics, Speech, and Signal Processing*, **38**(5):814–824, 1990.

[HSK99] B. Hassibi, A. H. Sayed, and T. Kailath. *Indefinite-Quadratic Estimation and Control. A Unified Approach to H^2 and H^∞ Theories*. Society for Industrial and Applied Mathematics, 1999.

[HTW15] T. Hastie, R. Tibshirani, and M. Wainwright. *Statistical Learning with Sparsity. The Lasso and Generalizations*. CRC Press, 2015.

[HV14] A. Hansson and L. Vandenberghe. "Sampling method for semidefinite programmes with non-negative Popov function constraints." *International Journal of Control*, **87**(2):330–345, 2014.

[IH03] T. Iwasaki and S. Hara. "Generalization of Kalman-Yakubovič-Popov lemma for restricted frequency inequalities." In *Proceedings of the American Control Conference*, pp. 3828–3833, 2003.

[IH05] T. Iwasaki and S. Hara. "Generalized KYP lemma: unified frequency domain inequalities with design applications." *IEEE Transactions on Automatic Control*, **50**(1):41–59, 2005.

[IMF00] T. Iwasaki, G. Meinsma, and M. Fu. "Generalized S-procedure and finite-frequency KYP lemma." *Mathematical Problems in Engineering*, **6**:305–320, 2000.

[IOW99] V. Ionescu, C. Oară, and M. Weiss. *Generalized Riccati Theory and Robust Control*. John Wiley and Sons, 1999.

[Kal63] R. E. Kalman. "Lyapunov Functions for the Problem of Lur'e in Automatic Control." *Proc. Nat. Acad. Sci., USA*, **49**:201–205, 1963.

[Kar84] N. Karmarkar. "A new polynomial-time algorithm for linear programming." *Combinatorica*, **4**(4):373–395, 1984.

[KN77] M. G. Krein and A. A. Nudelman. *The Markov Moment Problem and Extremal Problems*, volume 50 of *Translations of Mathematical Monographs*. American Mathematical Society, Providence, Rhode Island, 1977.

[KS66] S. Karlin and W. J. Studden. *Tchebycheff Systems: With Applications in Analysis and Statistics*. Wiley-Interscience, 1966.

[Las95] J. B. Lasserre. "A new Farkas lemma for positive semidefinite matrices." *IEEE Transactions on Automatic Control*, **40**(6):1131–1133, June 1995.

[Las97] J. B. Lasserre. "A Farkas lemma without a standard closure condition." *SIAM Journal on Control and Optimization*, **35**(1):265–272, 1997.

[LC16] Y. Li and Y. Chi. "Off-the-grid line spectrum denoising and estimation with multiple measurement vectors." *IEEE Transactions on Signal Processing*, **64**(5):1257–1269, 2016.

[LP04] J. Löfberg and P. A. Parrilo. "From coefficients to samples: a new approach to SOS optimization." In *Proceedings of the 43rd IEEE Conference on Decision and Control*, pp. 3154–3159, 2004.

[LV07] Z. Liu and L. Vandenberghe. "Low-rank structure in semidefinite programs derived from the KYP lemma." In *Proceedings of the 46th IEEE Conference on Decision and Control*, pp. 5652–5659, 2007.

[MAK95] P. Moulin, M. Anitescu, K. O. Kortanek, and F. A. Potra. "The role of linear semi-infinite programming in signal-adapted QMF bank design." *IEEE Transactions on Signal Processing*, **45**(9):2160–2174, 1995.

[MCK14] K. V. Mishra, M. Cho, A. Kruger, and W. Xu. "Off-the-grid spectral compressed sensing with prior information." In *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2014.

[MCK15] K. V. Mishra, M. Cho, A. Kruger, and W. Xu. "Spectral super-resolution with prior knowledge." *IEEE Transactions on Signal Processing*, **63**(20):5342–5357, 2015.

[Meh92] S. Mehrotra. "On the implementation of a primal-dual interior point method." *SIAM Journal on Optimization*, **2**(4):575–601, 1992.

[MOS02] MOSEK ApS. *The MOSEK Optimization Tools Version 2.5. User's Manual and Reference*, 2002. Available from `www.mosek.com`.

[MW01] J. W. McLean and H. J. Woerdeman. "Spectral factorizations and sums of squares representations via semidefinite programming." *SIAM Journal on Matrix Analysis and Applications*, **23**(3):646–655, 2001.

[Nes04] Yu. Nesterov. *Introductory Lectures on Convex Optimization*. Kluwer Academic Publishers, Dordrecht, The Netherlands, 2004.

[Nes13] Yu. Nesterov. "Gradient methods for minimizing composite functions." *Mathematical Programming, Series B*, **140**:125–161, 2013.

[NN94] Y. Nesterov and A. Nemirovskii. *Interior-Point Polynomial Algorithms in Convex Programming*. Society for Industrial and Applied Mathematics, 1994.

[NW06] J. Nocedal and S. J. Wright. *Numerical Optimization*. Springer, 2nd edition, 2006.

[PIH14] G. Pipeleers, T. Iwasaki, and S. Hara. "Generalizing the KYP lemma to multiple frequency intervals." *SIAM Journal on Control and Optimization*, **52**(6):3618–3638, 2014.

[PM96] J. G. Proakis and D. G. Manolakis. *Digital Signal Processing. Principles, Algorithms, and Applications*. Prentice-Hall, third edition, 1996.

[Pop62] V. M. Popov. "Absolute Stability of Nonlinear Systems of Automatic Control." *Automation and Remote Control*, **22**:857–875, 1962.

[PRK86] A. Paulraj, R. Roy, and T. Kailath. "A subspace rotation approach to signal parameter estimation." *Proceedings of the IEEE*, **74**(7):1044–1046, July 1986.

[Pro95] R. Prony. "Essai expérimental et analytique: sur les lois de la dilatabilité de fluides élastiques et sur celles de la force expansive de la vapeur de lalkool, à différentes températures." *Journal de lEcole Polytechnique*, **1**(2):24–76, 1795.

[PV11] G. Pipeleers and L. Vandenberghe. "Generalized KYP lemma with real data." *IEEE Transactions on Automatic Control*, **56**(12):2942–2946, 2011.

[Ran96] A. Rantzer. "On the Kalman-Yakubovich-Popov Lemma." *Systems and Control Letters*, **28**(1):7–10, 1996.

[RDV07] T. Roh, B. Dumitrescu, and L. Vandenberghe. "Interior-point algorithms for sum-of-squares optimization of multidimensional trigonometric polynomials." In *Proceedings of the International Conference on Acoustics, Speech, and Signal Processing*, volume 3, pp. 905–908, 2007.

[RFP10] B. Recht, M. Fazel, and P. A. Parrilo. "Guaranteed Minimum-Rank Solutions of Linear Matrix Equations via Nuclear Norm Minimization." *SIAM Review*, **52**(3):471–501, 2010.

[Roc70] R. T. Rockafellar. *Convex Analysis*. Princeton Univ. Press, Princeton, 1970.

[RV06] T. Roh and L. Vandenberghe. "Discrete transforms, semidefinite programming, and sum-of-squares representations of nonnegative polynomials." *SIAM Journal on Optimization*, **16**(4):939–964, 2006.

[SBT12] P. Shah, B. N. Bhaskar, G. Tang, and B. Recht. "Linear system identification via atomic norm regularization." In *Proceedings of the 51st IEEE Conference on Decision and Control*, pp. 6265–6270, 2012.

[Sch86] R. O. Schmidt. "Multiple emitter location and signal parameter estimation." *IEEE Transactions on Antennas and Propagation*, **34**(3):276–280, Mar 1986.

[Sch91] L. L. Scharf. *Statistical Signal Processing*. Addison-Wesley, 1991.

[Sch06] C. Scherer. "LMI relaxations in robust control." *European Journal of Control*, **12**(1):3–29, 2006.

[SDL10] P. Stoica, L. Du, J. Li, and T. Georgiou. "A new method for moving-average parameter estimation." In *2010 Conference Record of the Forty Fourth Asilomar Conference on Signals, Systems and Computers*, pp. 1817–1820, Nov 2010.

[SGB14] K. Scheinberg, D. Goldfarb, and X. Bai. "Fast first-order methods for composite convex optimization with backtracking." *Foundations of Computational Mathematics*, **14**:389–417, 2014.

[SM97] P. Stoica and R. L. Moses. *Introduction to Spectral Analysis*. Prentice Hall, London, 1997.

[SMM00] P. Stoica, T. McKelvey, and J. Mari. "MA estimation in polynomial time." *IEEE Transactions on Signal Processing*, **48**(7):1999–2012, July 2000.

[SN97] G. Strang and T. Nguyen. *Wavelets and Filter Banks*. Wellesley-Cambridge Press, revised edition, 1997.

[SP95] T. K. Sarkar and O. Pereira. "Using the matrix pencil method to estimate the parameters of a sum of complex exponentials." *IEEE Antennas and Propagation Magazine*, **37**(1):48–55, 1995.

[SS86] F. Santosa and W. H. Symes. "Linear inversion of band-limited reflection seismograms." *SIAM Journal on Scientific and Statistical Computing*, **7**(4):1307–1330, 1986.

[Ste03] M. Stewart. "A superfast Toeplitz solver with improved numerical stability." *SIAM Journal on Matrix Analysis and Applications*, **25**(3):669–693, 2003.

[Stu99] J. F. Sturm. "Using SEDUMI 1.02, a Matlab Toolbox for Optimization Over Symmetric Cones." *Optimization Methods and Software*, **11-12**:625–653, 1999.

[STY14] P. Stoica, G. Tang, Z. Yang, and D. Zachariah. "Gridless compressive-sensing methods for frequency estimation: Points of tangency and links to basics." In *2014 22nd European Signal Processing Conference (EUSIPCO)*, pp. 1831–1835, 2014.

[TBS13] G. Tang, B. N. Bhaskar, P. Shah, and B. Recht. "Compressed sensing off the grid." *IEEE Transactions on Information Theory*, **59**(11):7465–7490, 2013.

[Tib96] R. Tibshirani. "Regression shrinkage and selection via the Lasso." *Journal of the Royal Statistical Society. Series B (Methodological)*, **58**(1):267–288, 1996.

[Tro06] J. A. Tropp. "Just relax: Convex programming methods for identifying sparse signals in noise." *IEEE Transactions on Information Theory*, **52**(3):1030–1051, 2006.

[Tse08] P. Tseng. "On accelerated proximal gradient methods for convex-concave optimization." 2008.

[TTT02] K. C. Toh, R. H. Tütüncü, and M. J. Todd. *SDPT3 version 3.02. A Matlab software for semidefinite-quadratic-linear programming*, 2002. Available from `www.math.nus.edu.sg/~mattohkc/sdpt3.html`.

[Vai93] P. P. Vaidyanathan. *Multirate Systems and Filter Banks*. Prentice Hall, 1993.

[Van79] P. Van Dooren. "The computation of Kronecker's canonical form of a singular pencil." *Linear Algebra and Its Applications*, **27**:103–140, 1979.

[VK77] A. Vieira and T. Kailath. "On another approach to the Schur-Cohn criterion." *IEEE Transactions on Circuits and Systems*, **24**:218–220, 1977.

[Vos75] Z. Vostrý. "New algorithm for polynomial spectral factorization with quadratic convergence. Part I." *Kybernetika*, **11**:415–422, 1975.

[WBV96] S.-P. Wu, S. Boyd, and L. Vandenberghe. "FIR filter design via semidefinite programming and spectral factorization." In *Proc. IEEE Conf. on Decision and Control*, pp. 271–276, 1996.

[WBV98] S.-P. Wu, S. Boyd, and L. Vandenberghe. "FIR Filter Design via Spectral Factorization and Convex Optimization." In B. Datta, editor, *Applied and Computational Control, Signals, and Circuits*, volume 1, pp. 215–245. Birkhauser, 1998.

[Wil69] G. Wilson. "Factorization of the covariance generating function of a pure moving average process." *SIAM J. on Numerical Analysis*, **6**:1–7, 1969.

[Wol81] H. Wolkowicz. "Some Applications of Optimization in Matrix Theory." *Linear Algebra and Appl.*, **40**:101–118, 1981.

[Yak62] V. A. Yakubovich. "The solution of certain matrix inequalities in automatic control theory." *Soviet Math. Dokl.*, **3**:620–623, 1962.

[YX15] Z. Yang and L. Xie. "On gridless sparse methods for line spectral estimation from complete and incomplete data." *IEEE Transactions on Signal Processing*, **63**(12):3139–3153, 2015.

[YX16] Z. Yang and L. Xie. "Exact joint sparse frequency recovery via optimization methods." *IEEE Transactions on Signal Processing*, **64**(19):5145–5157, 2016.