Part I

Introduction

Chapter 1

Robust Decision Problems in Engineering: An LMI Approach[1]

L. El Ghaoui and S.-I. Niculescu

1.1 Introduction

1.1.1 Basic Idea

The basic idea of the LMI method is to formulate a given problem as an optimization problem with linear objective and *linear matrix inequality* (LMI) constraints. An LMI constraint on a vector $x \in \mathbf{R}^m$ is one of the form

$$F(x) = F_0 + \sum_{i=1}^{m} x_i F_i \geq 0 \tag{1.1}$$

where the symmetric matrices $F_i = F_i^T \in \mathbf{R}^{N \times N}$, $i = 0, \ldots, m$ are given. The minimization problem

$$\text{minimize } c^T x \text{ subject to } F(x) \geq 0 \tag{1.2}$$

where $c \in \mathbf{R}^m$, and $F \geq 0$ means the matrix F is symmetric and positive semidefinite, is called a semidefinite program (SDP). The above framework is particularly attractive for the following reasons.

Efficient numerical solution. SDP optimization problems can be solved very efficiently using recent interior-point methods for convex optimization (the global optimum is found in polynomial time). This brings a numerical solution to problems when no analytical or closed-form solution is known.

Robustness against uncertainty. The approach is very well suited for problems with uncertainty in the data. Based on a deterministic description of

[1]This work was partially funded by a support from EDF (Electricité de France), under contract P33L74/2M7570/EP777.

uncertainty (with detailed structure, and hard bounds), a systematic procedure enables to formulate an SDP optimization problem that yields a robust solution. This statement has implications for a wide scope of engineering problems, where measurements, modelling errors, etc, are often present.

Multicriteria problems. The approach enables to impose many different (possibly conflicting) specifications in the design process, allowing to explore trade-offs and analyze limits of performance and feasibility. This offers a drastic advantage over design methods that rely on a single criterium, deemed to reflect all design constraints—the choice of a relevant criterium is sometimes a non trivial task.

Wide applicability. The techniques used in the approach are relevant far beyond control and estimation. This opens exciting avenues of research where seemingly very different problems are analyzed and solved in a unified framework. For example, the method known in LMI-based control as the $\mathcal{S}$-procedure can be sucessfully applied in combinatorial optimization, leading to efficient relaxations of hard problems.

1.1.2 Approximation and complexity

At this point we should comment on a very important aspect of the LMI approach, which puts in perspective the above.

The LMI approach is deeply related to the branch of theoretical computer science that seeks to classify problems in terms of their *computational complexity* [77]. In a sense, the approach is an approximation technique for a class of (NP-) hard problems, via polynomial-time algorithms, and thus hinges on a dividing line between "hard" and "easy", in the sense of computational complexity.

This point of view shows that the LMI approach is not only a numerical approach to practical problems. It also requires one to consider the complexity analysis of a given problem, and to seek an approximation in the case of a "hard" problem. This concern is quite new in control; see [14] for a survey and references on complexity analysis of control problems.

1.1.3 Purpose of the Chapter

In this Chapter, we describe a general LMI-based method to cope with engineering decision problems with uncertainty, and illustrate its relevance, both in control and in other areas. This Chapter is intended to serve two kinds of objectives.

The first is to *introduce* the reader to the LMI method for control, and help understand some of the issues treated in more detail in the subsequent Chapters. References that would usefully complement this reading include a book on the LMI approach in system and control theory, by Boyd *et al.* [5], a book on algorithms for LMI problems by Nesterov and Nemirovskii [22], and [26]. Also, the course notes [19] contain many results on convex optimization and applications.

Another purpose of this Chapter is to *open* the method to other areas. In our presentation, we have tried to give an "optimization point of view" of the method, in an effort to put it in perspective with classical relaxation methods. This effort is mainly motivated by the belief that the approach could be useful in many other fields, especially regarding robustness issues. This topic of robustness is receiving a renewed attention in the field of optimization, as demonstrated by a series of recent papers [11, 75, 8].

1.1.4 Outline

Section 1.2 defines robust solutions to problems with uncertain data, and outlines the general ideas involved in Lagrange relaxations for these problems. Section 1.3 is devoted to models of uncertain systems. In section 1.4, we concentrate on SDP problems with uncertain data matrices, and detail a general methodology for computing appropriate bounds on these problems. Section 1.5 briefly describes algorithms, software and related issues for solving general SDPs. (This section, written with the help of Michael Overton and Francois Oustry, can be read independently of the rest of this Chapter.) Some illustrative examples taken from the control area, which complement those put forth in the other Chapters, are given in 1.6. Section 1.7 presents a number of examples taken out of control. Finally, Section 1.8 explores some perspectives and challenges that lie ahead.

1.2 Decision Problems with Uncertain Data

1.2.1 Decision problems

Many engineering analysis and design problems can be seen as decision problems. In control engineering, one must decide which controller gains to choose, in order to satisfy the desired specifications. This decision involves several trade-offs.

In this context, decision problems are based on two "objects": a set of *decision* variables, that reflect the engineer's choices (controller gains, actuator location, etc); and a set of *constraints* on these decision variables, that reflect the specifications imposed by the problem (desired closed-loop behavior, etc). Such problems can be sometimes translated in terms of mathematical programming problems of the form

$$\text{(1.3)} \qquad \text{minimize } f_0(x) \text{ subject to } x \in \mathcal{X}, \quad f_i(x) \leq 0, \quad i = 1, \ldots, p$$

where $f_0, f_1, \ldots, f_p$ are given scalar-valued functions of the decision vector $x \in \mathbf{R}^m$, and $\mathcal{X}$ is a subset of $\mathbf{R}^m$. In some problems, $\mathcal{X}$ is infinite-dimensional, and the decision vector x is a function.

In the language of control theory, we can interpret the above as a *multi-specification* control problem, where several (possibly conflicting) constraints have to be satisfied.

Let us comment on the limitations of the above mathematical framework. The desired specifications are sometimes easy to describe mathematically as in the above model (for example, when we seek to ensure a desired maximum overshoot). Others are more difficult to handle, and not less important, such as economic constraints (cost of design), reliability and ease of implementation (time needed to obtain a successful design), etc. These "soft" constraints can be neglected, or taken into account implicitely (for example, choosing a PID controller structure, as opposed to a more complicated one). A classical way to handle these constraints is by hierarchical decision, where soft constraints are taken into account at a higher level of managerial decision (*e.g.*, allowing a constrained budget for the whole control problem).

1.2.2 Using uncertainty models

In some applications, the "data" in problem (1.3) is not well-known; the previous "optimization model" falls short of handling this case, so it would be useful to consider decision problems with uncertainty. For this, it is necessary to introduce *models* for uncertain systems; this is the purpose of section 1.3.

There are other motivations for us to use such models. Real-world phenomena are very complex (nonlinear, time-varying, infinite-dimensional, etc). To cope with analysis and design problems for such systems, we almost always have to use (finite-dimensional, linear, stationary) approximations. We believe that a finer way to handle complexity is to use approximations with uncertainty.

We thus make the distinction between "physical" and "engineering" uncertainties. The first are due to ignorance of the precise physical behavior; the second come from a voluntary simplifying assumption. (This technique can be referred to as *embedding.*)

Assume for instance that in a complex system, two variables p, q are linked by a nonlinear relation $p = \delta(q)$. The function δ may be very difficult to identify (this is a physical uncertainty). Moreover, we may choose to "embed" this nonlinearity in a *family* of linear relations of the form $p = \delta q$, where the scalar δ is inside a known sector. (The latter may be inferred from physical knowledge.) We thus obtain a linear model with (engineering) uncertainty, that is more tractable than the nonlinear one. The technique of embedding may be conservative, but usually gives rise to tractable problems.

1.2.3 Robust decision problems

In some applications, the "data" in problem (1.3) is not well-known, that is, every f_i depends not only on x but also on a real "uncertainty matrix[2]" Δ. The latter is only known to belong to a set of admissible perturbations $\mathbf{\Delta}$. We may then define a variety of *robust decision problems,* and associated solutions. These problems are based on a worst-case analysis of the effect of perturbations.

[2]We take the uncertainty to be a matrix instead of a vector, for reasons to be made clear later.

Robustness analysis.

We define the *robust feasibility* set

$$(1.4) \quad \mathcal{X}(\mathbf{\Delta}) = \{x \in \mathcal{X} \mid f_i(x, \Delta) \leq 0, \quad i = 1, \ldots, p \text{ for every } \Delta \in \mathbf{\Delta}\} .$$

The analysis problem is to check if a given candidate solution x is robustly feasible.

Robust synthesis

A robust synthesis problem is to find a vector x that satisfies robustness constraints. We may distinguish:

- A *robust feasibility problem*, where we seek a vector x that belongs to $\mathcal{X}(\mathbf{\Delta})$.
- A *robust optimality problem*, defined as

  $$(1.5) \qquad \text{minimize } \max_{\Delta \in \mathbf{\Delta}} f_0(x, \Delta) \text{ subject to } x \in \mathcal{X}(\mathbf{\Delta}),$$

  Here, we seek to minimize the worst-case value of the objective $f_0(x, \Delta)$, over the set of robustly feasible solutions. (We note that the objective function f_0 may be assumed independent of the perturbation, without loss of generality.)
- *Parameter-scheduled synthesis.* In such a problem, we allow the decision vector to be a function of the perturbation. In this case, we assume that the decision vector is not finite-dimensional, but evolves in a set $\mathcal{X}$ of functions (of the perturbation Δ). The problem reads formally as the robust synthesis ones, but with $\mathcal{X}$ a set of functions.
- *Mixed parameter-scheduled/robust problems.* In some problems, the decision vector is a mixture of a set of real variables, and of functions of some parameters. In control, this kind of problem occurs when only some parameters are available for measurement.

In control theory, we interpret the above problems as *robust multi-specification* control problems, where several (possibly conflicting) constraints have to be satisfied in a robust manner. We note that these problems are in general very hard to solve (non computationally tractable). References on complexity of such problems in control include [21, 67, 20, 13, 92, 39].

Robust decision problems first arose (to our knowledge) in the field of automatic control. In fact, robustness is presented in the classical control textbooks as the main reason for using feedback. Good references on robust control include the book by Zhou, Doyle and Glover [29], that of Dahleh and Diaz-Bobillo [22], and Green and Limebeer [44].

In other areas of engineering, uncertainty is generally dealt with using stochastic models for uncertainty. In *stochastic optimization,* the uncertainty is assumed to have a known distribution, and a typical problem is to evaluate the distribution of the optimal value, of the solution, etc; a good reference on this subject is the book by Dempster [24].

Models with deterministic uncertainty are less classical in engineering optimization. Ben Tal and Nemirovskii consider a truss topology design problems, with uncertainty on the loading forces [10, 9]. In a more recent work [11], they introduce, and study the complexity of, robust optimality problems, in the sense defined above, in the context of convex optimization. Their approach is based on ellipsoidal bounds for the perturbation. The paper [75] addresses the computation of bounds for a class of robust problems, and uses a control-based approach to solve for these bounds, via LFRs and SDP.

1.2.4 Lagrange relaxations for min-max problems

Robust decision problems such as problem (1.5) belong to the class of min-max problems. To attack them, we can thus use a versatile technique, called *Lagrange relaxation*, that enables to approximate a set of complicated constraints by a "more tractable" set.

To understand the technique, let us assume that the uncertainty set $\mathbf{\Delta}$ is a subset of $\mathbf{R}^l$ that can be described by a finite number of (nonlinear) constraints

$$\mathbf{\Delta} = \left\{ \delta \in \mathbf{R}^l \;\middle|\; g_i(\delta) \le 0, \quad i = 1, \ldots, q \right\},$$

where the g_i's are given scalar-valued functions of the perturbation vector δ. Now, observe that the property

$$g_0(\delta) \le 0 \text{ for every } \delta, \quad g_i(\delta) \le 0, \quad i = 1, \ldots, q$$

holds if there exist nonnegative scalars $\tau_1, \ldots, \tau_q$ such that

$$\text{for every } \delta, \quad g_0(\delta) \le \sum_{i=1}^{q} \tau_i g_i(\delta).$$

The above condition is rewritten

$$\max_{\delta} \left\{ g_0(\delta) - \sum_{i=1}^{q} \tau_i g_i(\delta) \right\} \le 0. \tag{1.6}$$

The above idea can be used for general min-max problems of the form (1.5). The basic idea is to approximate a min-max problem by a (hopefully easier) minimization problem. To motivate this idea, we take an example of the robust optimality problem (1.5) with no constraints on x ($p = 0$). That is, we consider

$$\min_{x \in \mathcal{X}} \max_{\delta} \left\{ f_0(x, \delta) \;\mid\; g_i(\delta) \le 0, \quad i = 1, \ldots, q \right\}. \tag{1.7}$$

The Lagrange relaxation technique yields an *upper bound* on the above problem:

$$\begin{array}{ll} \text{minimize} & \displaystyle\max_{\delta} \left(f_0(x, \delta) - \sum_{i=1}^{q} \tau_i g_i(\delta) \right) \\ \text{subject to} & x \in \mathcal{X}, \quad \tau_i \ge 0, \quad i = 1, \ldots, q. \end{array} \tag{1.8}$$

In some cases, computing the (unconstrained) maximum in the above problem is very easy; then, the min-max problem is approximated (via an upper bound) by a minimization problem (the minimization variables are x and τ). (The scalars τ_i appearing in the above are often referred to as *Lagrange multipliers.*) Note that, for x fixed, the above problem is convex in τ; if f_0 is convex in x, then we may use (subgradient-based) convex optimization to solve the relaxed problem. The technique thus requires that computing values and sub-gradient of the "max" part (a function of x, τ) is computationnally tractable.

The upper bounds found via Lagrange relaxations can be improved as long as we are able to solve, and compute corresponding sub-gradients, for a constrained "max" problem in (1.8). For example, we may choose to relax only a subset of the constraints $g_i(\delta) \leq 0$, and keep the remaining ones as hard constraints in the maximization problem. Different choices of "relaxed" vs. "non relaxed" constraints lead to different Lagrange relaxations (and upper bounds).

A special case: the $\mathcal{S}$-procedure. The well-known $\mathcal{S}$-procedure is an application of Lagrange relaxation in the case when the g_i's are quadratic functions of $\delta \in \mathbf{R}^l$. In this case, checking (1.6) yields a simple LMI in the multipliers τ_i. The $\mathcal{S}$-procedure lemma, which we recall below, is widely used, not only in control theory [59, 37, 5], but also in connection with trust region methods in optimization [88, 90].

Lemma 1.1 ($\mathcal{S}$-procedure)
Let $g_0, \ldots, g_p$ be quadratic functions of the variable $\delta \in \mathbf{R}^m$:

$$g_i(\delta) \triangleq \begin{bmatrix} \delta \\ 1 \end{bmatrix}^T F_i \begin{bmatrix} \delta \\ 1 \end{bmatrix}, \quad i = 0, \ldots, p, \tag{1.9}$$

where $F_i = F_i^T$. The following condition on $g_0, \ldots, g_p$:

$$g_0(\delta) \geq 0 \text{ for all } \delta \text{ such that } g_i(\delta) \leq 0, \; i = 1, \ldots, p$$

holds if there exist scalars $\tau_i \geq 0$, $i = 1, \ldots, p$, such that

$$F(\tau) = F_0 + \sum_{i=1}^{p} \tau_i F_i \geq 0. \tag{1.10}$$

When $p = 1$, the converse holds, provided that there is some δ_0 such that $g_1(\delta_0) > 0$.

1.3 Uncertainty Models

A typical decision problem is defined via some data, which we collect for convenience in a matrix M. For example, if the problem is defined in terms of the transfer function of an LTI system $M(s) = D + C(sI - A)^{-1}B$, the matrix M contains the four matrices A, B, C, D.

When the data is uncertain, we need to describe as compactly as possible, the way the perturbation Δ affects the data matrix, and also the structure of the perturbation set $\mathbf{\Delta}$. This will result in an *uncertainty model* for the data. Formally, an uncertainty model for the data is defined as a matrix set

$$\{\mathbf{M}(\Delta) \mid \Delta \in \mathbf{\Delta}\}, \tag{1.11}$$

where $\mathbf{M}$ is a (possibly nonlinear) matrix-valued function of the "perturbation matrix" Δ, and $\mathbf{\Delta}$ is a matrix set.

1.3.1 Linear-fractional models

Definition

A linear-fractional model is a set of the form (1.11), where the following three assumptions are made.

Linear-fractional assumption. We assume that the matrix-valued function can be written, for almost every Δ, as

$$\mathbf{M}(\Delta) = M + L\Delta(I - D\Delta)^{-1}R$$

for appropriate constant matrices M, L, R and $D \in \mathbf{R}^{n_q \times n_p}$.

Uncertainty set assumption. In addition, the uncertainty set $\mathbf{\Delta}$ is required to satisfy the following.

For every vector pair p, q, the condition

$$p = \Delta q \text{ for some } \Delta \in \mathbf{\Delta}$$

can be characterized as a *linear matrix inequality* on a rank-one matrix:

$$\Phi_{\mathbf{\Delta}}\left(\begin{bmatrix} p \\ q \end{bmatrix}\begin{bmatrix} p \\ q \end{bmatrix}^T\right) \geq 0 \tag{1.12}$$

where $\Phi_{\mathbf{\Delta}}$ is a given linear operator "caracterizing" $\mathbf{\Delta}$.

Well-posedness assumption. To simplify our exposition, we make a last assumption from now on, that the above model is *well-posed* over $\mathbf{\Delta}$, meaning that

$$\det(I - D\Delta) \neq 0 \text{ for every } \Delta \in \mathbf{\Delta}.$$

Checking well-posedness is very difficult in general; using the theory developed in section 1.4, we can derive sufficient LMI conditions for well-posedness. Chapter 12 discusses the issue of well-posedness in more detail.

Motivations

The above uncertainty models for the data seem very specialized. However, they can cover a wide variety of uncertain matrices, as we now show.

First, the *linear-fractional assumption* made on the matrix function $\mathbf{M}(\Delta)$ is motivated by the following lemma, a constructive proof of which appears in *e.g.* [16, 29].

Lemma 1.2
Every (matrix-valued) rational function $\mathbf{M}(\delta)$ *of* $\delta \in \mathbf{R}^p$ *that is well-defined for* $\delta = 0$ *admits* a linear-fractional representation *(LFR) of the form*

$$\begin{array}{c} \mathbf{M}(\delta) = M + L\Delta(I - D\Delta)^{-1}R, \\ \textit{where } \Delta = \mathbf{diag}(\delta_1 I_{r_1}, \ldots, \delta_l I_{r_l}), \end{array} \tag{1.13}$$

valid for whenever $\det(I - D\Delta) \neq 0$, *for appropriate matrices* M, L, R, D *and integers* $r_1, \ldots, r_l$.

We note that, when $l = 1$ (that is, for a *monovariable* rational matrix function), the matrices M, L, R, D are simply a state-space realization of the transfer matrix $M + L(sI - D)^{-1}R$, where $s = 1/\delta$. An LFR of $\mathbf{M}$ can then always be constructed in a "minimal" way, such that the size of matrix Δ is exactly the order of $\mathbf{M}$ in s. In the multivariable case, the integers r_i are greater than the degrees of the corresponding variable δ_i in $\mathbf{M}$, and the issue of minimality is much more subtle. Finally, we note that when $\mathbf{M}$ is polynomial in its argument, an LFR can always be constructed such that D is strictly upper triangular, so that this LFR is everywhere well-posed—as is $\mathbf{M}$.

The LFR generalizes the state-space representation known for (monovariable) transfer matrices, to the multivariable case[3]. For other comments on the LFR formalism, see for example [6, 56].

The above lemma shows that we can handle almost arbitrary algebraic functions of perturbation parameters, provided we define the perturbation matrix Δ as a diagonal matrix, with repeated elements. (In the sequel, we pay a special attention to such diagonal perturbation structures.)

The *uncertainty set assumption* made on the set $\mathbf{\Delta}$ can appear also very specialized. In fact, it can handle a wide array of uncertainty bounds. Here is a short list of examples:

Unstructured case. Assume

$$\mathbf{\Delta} = \{\Delta \in \mathbf{R}^{n_p \times n_q} \mid \|\Delta\| \leq 1\} \tag{1.14}$$

This case is referred to as the "unstructured perturbations" case, and is a classic since the development of $\mathbf{H}_\infty$ control. The corresponding characterization is

$$\Phi_{\mathbf{\Delta}}\left(\begin{bmatrix} p \\ q \end{bmatrix}\begin{bmatrix} p \\ q \end{bmatrix}^T\right) = q^T q - p^T p.$$

[3] The above representation is usually referred to as the "linear-fractional transformation" (LFT) [29]. We believe the term representation is more appropriate here.

Euclidean-norm bounds. Assume

$$\mathbf{\Delta} = \left\{ \mathbf{diag}(\delta_1 I_{r_1}, \ldots, \delta_l I_{r_l}) \mid \delta \in \mathbf{R}^l,\ \|\delta\|_2 \le 1 \right\}. \tag{1.15}$$

The corresponding characterization is

$$\Phi_{\mathbf{\Delta}}\left(\begin{bmatrix} p \\ q \end{bmatrix} \begin{bmatrix} p \\ q \end{bmatrix}^T \right) = \mathbf{diag}(q_1 q_1^T, \ldots, q_l q_l^T) - pp^T.$$

(In the above, q_i denotes i-th $r_i \times 1$ block in vector q.)

Maximum-norm bounds. Assume

$$\mathbf{\Delta} = \left\{ \mathbf{diag}(\delta_1 I_{r_1}, \ldots, \delta_l I_{r_l}) \mid \delta \in \mathbf{R}^l,\ \|\delta\|_\infty \le 1 \right\}. \tag{1.16}$$

The corresponding characterization is

$$\Phi_{\mathbf{\Delta}}\left(\begin{bmatrix} p \\ q \end{bmatrix} \begin{bmatrix} p \\ q \end{bmatrix}^T \right) = \mathbf{diag}(q_1 q_1^T - p_1 p_1^T, \ldots, q_l q_l^T - p_l p_l^T).$$

(Again, p_i (resp. q_i) denotes i-th $r_i \times 1$ block in vector q (resp. p).)

Sector bounds. Assume $\mathbf{\Delta} = \{s \cdot I \mid s \in \mathbf{C},\ \ s + s^* \ge 0\}$. The corresponding characterization is

$$\Phi_{\mathbf{\Delta}}\left(\begin{bmatrix} p \\ q \end{bmatrix} \begin{bmatrix} p \\ q \end{bmatrix}^T \right) = pq^H + qp^H.$$

Of course, it is possible to mix different bounds to cover more complicated cases.

1.3.2 Operator point of view

The above uncertainty models for data matrices can be interpreted as input-output maps, as follows. Given $\Delta \in \mathbf{\Delta}$, the input-output relationship between two vectors u and y

$$y = \mathbf{M}(\Delta) u$$

can be rewritten, if $\det(I - D\Delta) \neq 0$, as a *quadratic* one:

$$\begin{array}{rcl} y &=& Mu + Lp, \\ q &=& Ru + Dp, \\ p &=& \Delta q. \end{array}$$

The above representation is not surprising: it shows that a nonlinear (algebraic) constraint can always be seen as a quadratic one, provided we introduce enough "dummy" variables (here, the vectors p, q).

In view of our assumption on the uncertainty set $\mathbf{\Delta}$, the "uncertain constraint"

$$y = \mathbf{M}(\Delta)u \text{ for some } \Delta \in \mathbf{\Delta}$$

can be rewritten in a quadratic matrix inequality form

$$(1.17) \qquad \begin{bmatrix} y \\ q \end{bmatrix} = \begin{bmatrix} M & L \\ R & D \end{bmatrix} \begin{bmatrix} u \\ p \end{bmatrix}, \quad \Phi_{\mathbf{\Delta}} \left(\begin{bmatrix} p \\ q \end{bmatrix} \begin{bmatrix} p \\ q \end{bmatrix}^T \right) \geq 0.$$

The operator point of view allows to consider (linear) dynamical relationships. For example, the LTI system

$$\dot{p} = Ap$$

where A is a constant matrix, can be represented in the frequency domain as

$$q = Ap, \quad q = s \cdot p.$$

When we need to analyze stability, we restrict ourselves to all complex s such that $\Re s \geq 0$. In this case, the uncertain system to be considered is

$$q = Ap, \quad pq^H + qp^H \geq 0.$$

In section 1.4.6, we follow up this idea in connection with Lyapunov theory.

1.3.3 Examples

Let us now show how the above tools can be used to describe compactly a wide array of uncertain matrices and dynamical systems. To simplify our description, we take the operator point of view mentioned previously.

Matrices.

The above framework covers the case when some (matrix) data occuring in the robust decision problem is affected by a perturbation vector δ, in an algebraic manner, and the vector δ is unknown-but-bounded.

Consider for example an "uncertain matrix" $\mathbf{M}(\delta)$, where $\mathbf{M}$ is an algebraic function of vector δ, and δ is unknown, but bounded in the maximum-norm sense. This matrix can be represented as an input-output operator $u \rightarrow y := \mathbf{M}(\delta)u$, in the form

$$(1.18) \begin{bmatrix} y \\ q \end{bmatrix} = \begin{bmatrix} M & L \\ R & D \end{bmatrix} \begin{bmatrix} u \\ p \end{bmatrix}, \quad \Delta = \mathbf{diag}(\delta_1 I_{r_1}, \ldots \delta_l I_{r_l}), \quad \|\Delta\| \leq 1.$$

The above form consitutes a *linear-fractional model* for the uncertain linear constraint $y = \mathbf{M}(\delta)u$, $\|\delta\|_\infty \leq 1$.

Model (1.18) can also be represented in the quadratic matrix inequality form (1.17), where $\Phi_{\mathbf{\Delta}}$ is defined in (1.16).

Dynamical systems.

The input-output relationship $y = \mathbf{M}(\Delta)u$ may involve a dynamical system; in this case also, LFR models can be used.

For example, the uncertain dynamical system

$$\dot{y} = \mathbf{M}(\Delta)y, \quad \Delta \in \mathbf{\Delta}$$

can be represented as, with the assumptions made previously in force, as

$$\begin{bmatrix} \dot{y} \\ q \end{bmatrix} = \begin{bmatrix} M & L \\ R & D \end{bmatrix} \begin{bmatrix} y \\ p \end{bmatrix}, \quad \Phi_{\mathbf{\Delta}} \left(\begin{bmatrix} p \\ q \end{bmatrix} \begin{bmatrix} p \\ q \end{bmatrix}^T \right) \geq 0.$$

It is possible to consider nonlinear systems with the above representations, using an "embedding" technique. Roughly speaking, if we know bounds on the trajectories of the above uncertain system, then these bounds will also hold for the nonlinear system

$$\dot{y} = \mathbf{M}(\Delta(y))y,$$

if $\Delta(y) \in \mathbf{\Delta}$ for every y. (This technique, which is more powerful than the classical linearization technique, is related to the *comparison principle* in differential equations; see [53] for more details.)

We take another example arising in the control of linear systems with delayed input, where we assume only time-delay *interval* uncertainty. Consider the following (frequency-domain) constraint:

$$y(s) = H(s)u(s), \text{ where } H(s) = H_0(s)e^{-\delta s}, \tag{1.19}$$

where H_0 is a transfer function that represents the nominal system of the form:

$$H_0(s) = H_{nd}(s)e^{-\delta_n s},$$

with H_{nd} free of delay and δ_n as the nominal delay. The uncertainty δ is of the form:

$$\delta = \rho\Delta,$$

where Δ is an unknown-but-bounded real scalar ($-1 \leq \Delta \leq 1$). A way to handle this class of systems is to aproximate the delay uncertainty term using a first-order real Padé approximant. An LFR of the approximation is:

$$e^{-\rho\Delta s} \approx \frac{1 - \frac{\rho}{2}s\Delta}{1 + \frac{\rho}{2}s\Delta} = 1 - \left(1 + \frac{\rho}{2}s\Delta\right)^{-1} \rho s\Delta, \ \Delta \in [-1, 1].$$

Using this LFR, we may obtain an LFR for the delayed system. (An alternate representation of this kind of uncertainty is given below.) Analysis and comparisons between several approximations of the delay uncertainty can be found in [99].

1.3.4 Towards other models

We now briefly mention other possible models. (For other remarks and comments on uncertainty modeling, we refer to [89].)

Integral quadratic constraint models.

An alternative to the class of the above parameter-dependent models is one where we assume that some signals in a known, LTI system are subject to integral quadratic constraints (IQCs) that are usually written in the frequency domain. Thus, the unknown part of the system is only subject to energy-type constraints. For other comments on this formalism, see for example [21, 81].

This class of systems is perhaps more versatile than the class of parameter-dependent systems seen above. We can use this models to "embed" a large class of complicated systems, such as systems with delays, parameter-dependent systems with bounds on the parameter's rate of variation, infinite-dimensional systems, etc.

Consider for example the following system:

$$\begin{array}{l} \dot{x}(t) = Ax(t) + Bp(t) \\ q(t) = Cx(t) + Dp(t) \\ p(t) = \Delta(t)q(t) \end{array}, \tag{1.20}$$

where the uncertainty block $\Delta(t)$ satisfies an IQC. This IQC is defined as follows: let $\Pi : j\mathbf{R} \mapsto \mathbb{C}^{r\times r}$ be a bounded measurable function, with appropriate r. The corresponding IQC is

$$\int_{-\infty}^{+\infty} \left[\begin{array}{c} \hat{p}(j\omega) \\ \hat{q}(j\omega) \end{array} \right]^* \Pi(j\omega) \left[\begin{array}{c} \hat{p}(j\omega) \\ \hat{q}(j\omega) \end{array} \right] d\omega \geq 0, \tag{1.21}$$

where $\hat{p}$, $\hat{q}$ are the Fourier transforms of $p, q \in \mathbf{L}_2[0,\infty)$.

As an example, we consider again the delayed system (1.19) with $\delta_n = 0$, and the uncertainty delay interval of the form $[0,\rho]$. A different way to handle the uncertainty on the delay, is to suppose that the uncertain "quantity" is $\Delta u(t) = u(t-h) - u(t)$, $h \leq \rho$ (here, $u(\cdot)$ is the input signal). This signal satisfies the IQC above, with (see [81]):

$$\Pi(j\omega) = \left[\begin{array}{cc} 0 & \alpha(j\omega) + j\beta(j\omega) \\ \alpha(j\omega) - j\beta(j\omega) & -\alpha(j\omega) - j\beta(j\omega) cot\left(\frac{\rho\omega}{2}\right) \end{array} \right],$$

where α and β are bounded real-valued functions on the imaginary axis such that

$$\left\{ \begin{array}{l} \omega\beta(j\omega) \geq 0 \;\text{ for }\; |\,\omega\,| \leq \frac{\pi}{\rho} \\ \beta(j\omega) = 0 \;\text{ otherwise.} \end{array} \right.$$

IQCs are examined in more detail in Chapter 6 of this book.

Polytopic models.

In a polytopic model [5], we assume that the state-space matrices of the system are only known to lie in a given polytope. For example:

$$\dot{x} = \mathbf{A}(t)x, \text{ where } \mathbf{A}(t) \in \mathbf{Co}\{A_1, \ldots, A_L\} \text{ for every } t \geq 0. \tag{1.22}$$

Here, the set of admissible perturbations is $\mathcal{D}$ is the polytope $\{\lambda \mid \lambda_i \geq 0, \ \sum \lambda_i = 1\}$, and the mapping ϕ assigns the linear combination $\sum \lambda_i A_i$ to the vector λ. Polytopic systems are considered in Chapter 9 of this book.

Refined perturbation sets.

The above models are very rough in that they impose no restriction on the rate of change of the perturbed elements. For example, the polytopic model (1.22) could be used for a system subject to possible failures; each matrix A_i in (1.22) represents a specific operating mode for the system. In practice, this model will be very conservative, as it does not prevent the system from jumping arbitrarily fast from a mode to any other. Some refined models simply impose a bound on the rate of variation of the perturbation parameters, and make use of LMI analysis which takes into account this bound.

Stochastic models.

These models are often used in engineering to describe uncertainty, and could be interpreted as more refined than deterministic ones. For example, for a system with several failure modes, a model that is more refined than the polytopic model could be based on systems with Markovian jumps. To the matrices A_i of the polytopic model, we associate a "transition probability matrix" that describes the probability of change from one mode to the other. Similarly, it is possible to consider a linear-fractional model, where the uncertain matrix Δ obeys a known white noise-type distribution; see [29] for an example.

The reader should not make a conflict out of the distinction deterministic/stochastic: it is possible to consider stochastic models with deterministic uncertainty. For example, we may consider a linear system with Markovian jumps, where the transition probability matrix is unknown-but-bounded (see [31]).

1.4 Robust SDP

Robust decision problems are defined in terms of (a) a "nominal" problem, of the form (1.3); and (b) an uncertainty model for the data. Equipped with (matrix-based) uncertainty models for the data described previously, we need to define more precisely the form of the nominal problem. It turns out that a class of problems very rich in applications is precisely SDP. This motivates the study of *robust SDP* problems, which correspond to an SDP with uncertain data.

1.4.1 Definition

Consider the case when the constraints $f_i(x,\delta) \le 0$, $i = 1,\ldots,p$, can be written as

$$\mathbf{F}(x,\Delta) \ge 0,$$

where $\mathbf{F}$ is affine in the decision variable x, and takes its values in the set of symmetric matrices.

We make the assumption that $\mathbf{F}$ is a linear-fractional function of Δ, with LFR

$$\mathbf{F}(x,\Delta) = F(x) + L(x)\Delta(I - D\Delta)^{-1}R + \left(L(x)\Delta(I - D\Delta)^{-1}R\right)^T,$$

where $F(x) = F(x)^T$ and $L(x)$ are affine matrix-valued functions of x, and R, D are given matrices, while the uncertainty matrix Δ is restricted to a set $\boldsymbol{\Delta}$. We assume that the above LFR is well-posed over $\boldsymbol{\Delta}$, and in addition, we assume that the set $\boldsymbol{\Delta}$ satisfies the *uncertainty set assumption* defined in section 1.3.1.

The problem

$$\text{minimize } c^T x \text{ subject to } \quad \mathbf{F}(x,\Delta) \ge 0 \text{ for every } \Delta \in \boldsymbol{\Delta}$$

is called a *robust SDP* problem. Note that we assumed that the "objective vector" c is independent of perturbation; this is done without loss of generality.

The above problem is convex, however it is in general very difficult to solve (NP-hard, see [19, 77]); checking if a given candidate solution x is robustly feasible is already very difficult in general. Our objective is to find lower bounds on this problem, in the form of SDP.

1.4.2 Lagrange relaxation of robust SDP

Let us fix the decision variable x, and seek sufficient condition for

$$\xi^T \mathbf{F}(x,\Delta)\xi \ge 0 \text{ for every } \xi, \quad \Delta \in \boldsymbol{\Delta}$$

Using the LFR, and using the well-posedness assumption, we rewrite the above as

$$\begin{array}{c} \xi^T\left(F(x)\xi + 2L(x)p\right) \ge 0, \text{ for every } \xi, p, q, \Delta \text{ such that} \\ q = R\xi + Dp, \quad p = \Delta q, \quad \Delta \in \boldsymbol{\Delta}. \end{array}$$

With the previous characterization $p = \Delta q$, the above is seen to be equivalent to

$$\begin{array}{c} \mathbf{Tr} \begin{bmatrix} F(x) & L(x) \\ L(x)^T & 0 \end{bmatrix} \begin{bmatrix} \xi \\ p \end{bmatrix} \begin{bmatrix} \xi \\ p \end{bmatrix}^T \ge 0, \text{ for every } \xi, p \\ \text{such that } \Phi_{\boldsymbol{\Delta}}\left(\begin{bmatrix} R & D \\ 0 & I \end{bmatrix} \begin{bmatrix} \xi \\ p \end{bmatrix} \begin{bmatrix} \xi \\ p \end{bmatrix}^T \begin{bmatrix} R & D \\ 0 & I \end{bmatrix} \right) \ge 0. \end{array} \tag{1.23}$$

We are ready now to use Lagrange relaxation, following the general ideas evoked in section 1.2.4. Previous condition holds if there exist a positive semidefinite matrix S such that : $\forall\ \xi, p$,

$$\begin{bmatrix} \xi \\ p \end{bmatrix}^T \begin{bmatrix} F(x) & L(x) \\ L(x)^T & 0 \end{bmatrix} \begin{bmatrix} \xi \\ p \end{bmatrix} \geq \mathbf{Tr} S \Phi_{\mathbf{\Delta}} \left(\begin{bmatrix} R & D \\ 0 & I \end{bmatrix} \begin{bmatrix} \xi \\ p \end{bmatrix} \begin{bmatrix} \xi \\ p \end{bmatrix}^T \begin{bmatrix} R & D \\ 0 & I \end{bmatrix} \right),$$

or, equivalently,

$$\begin{bmatrix} F(x) & L(x) \\ L(x)^T & 0 \end{bmatrix} \geq \begin{bmatrix} R & D \\ 0 & I \end{bmatrix}^T \Phi^*_{\mathbf{\Delta}}(S) \begin{bmatrix} R & D \\ 0 & I \end{bmatrix}, \tag{1.24}$$

where $\Phi^*_{\mathbf{\Delta}}$ is the dual map of $\Phi_{\mathbf{\Delta}}$.

Hence, the Lagrange relaxation of robust SDP is an SDP, of the form

$$\min\ c^T x\ \text{ subject to } S \geq 0 \text{ and } (1.24)$$

The variables in this SDP are x (the original decision variable) and S (the "multiplier" matrix). The SDP provides an approximation of the robust SDP, in the following senses: (a) it yields an upper bound on the robust SDP; (b) any optimal x is robust (*i.e.*, it is feasible for all $\Delta \in \mathbf{\Delta}$); (c) in the unstructured case ($\mathbf{\Delta}$ defined by (1.14)), the approximation is exact.

A special case: robust linear programming

Consider the LP

$$\text{minimize } c^T x \text{ subject to } a_i^T x \geq b_i,\ i = 1, \ldots, L. \tag{1.25}$$

Assume that the a_i's and b_i's are subject to unstructured perturbations. That is, the perturbed value of $[a_i^T\ b_i]^T$ is $[a_i^T\ b_i]^T + \delta_i$, where $\|\delta_i\|_2 \leq \rho$, $i = 1, \ldots, L$, and $\rho > 0$ is a given measure of the perturbation level. We seek a robust solution to our problem, that is, we seek x that minimizes $c^T x$ subject to the robustness constraints

$$\left(\begin{bmatrix} a_i \\ b_i \end{bmatrix} + \delta_i \right)^T \begin{bmatrix} x \\ -1 \end{bmatrix} \leq 0 \text{ for every } \delta_i, \quad \|\delta_i\|_2 \leq \rho, \quad i = 1, \ldots, L.$$

This is a special case of a robust optimality problem, as defined in §1.2.3.

Using the previous theory, it is straightforward to show that the robust LP is *equivalent* to the problem

$$\text{minimize } c^T x \text{ subject to } a_i^T x - \rho\sqrt{\|x\|_2^2 + 1} \geq b_i,\ i = 1, \ldots, L. \tag{1.26}$$

The above problem is second-order cone program (SOCP), for which efficient special-purpose interior-point methods are available (see section 1.5.1).

We note that for every $\rho > 0$, the above problem yields a unique solution, that is a regular (in particular, continuous) function of the problem's data (the

a_i's and the b_i's). Thus, the approach provides a regularization procedure for the LP (1.25).

A simple variation of the robust LP is *LP with implementation constraints*, where we seek a solution x such that for every δx, $\|\delta x\| \le \rho$, the modified vector $x + \delta x$ still satisfies the constraint of the original LP. This problem arises for example in the area of optimal filter design, where the filter's coefficients can only be implemented with finite precision.

1.4.3 A dual view: rank relaxation

Lagrange relaxation can be interpreted in a dual way, in terms of relaxation of a rank constraint of the rank-one, positive semidefinite matrix

$$X = \begin{bmatrix} \xi \\ p \end{bmatrix} \begin{bmatrix} \xi \\ p \end{bmatrix}^T$$

which appears in (1.23). Define

$$N := \begin{bmatrix} R & D \\ 0 & I \end{bmatrix}, \quad M(x) := \begin{bmatrix} F(x) & L(x) \\ L(x)^T & 0 \end{bmatrix} = M_0 + \sum_{i=1}^m x_i M_i.$$

Define also the convex set

$$\mathcal{X} := \left\{ X \;\middle|\; \Phi_{\mathbf{\Delta}}\left(NXN^T\right) \ge 0, \;\; X \ge 0 \right\}.$$

Let us fix the decision variable x again. Condition (1.23) holds if and only if

$$0 \ge \max \left\{ -\mathbf{Tr} M(x) X \;\mid X \in \mathcal{X}, \;\; \mathbf{rank} X = 1 \right\}.$$

The robust SDP problem writes as the max-min problem

$$\min_x \; \max_X \; \left\{ c^T x - \mathbf{Tr} M(x) X \;\mid X \in \mathcal{X}, \;\; \mathbf{rank} X = 1 \right\}.$$

Relaxing the rank constraint yields *lower* bound

$$\min_x \; \max_{X \in \mathcal{X}} \; c^T x - \mathbf{Tr} M(x) X,$$

which, by virtue of the saddle-point theorem (recall $\mathcal{X}$ is convex), is equivalent to the SDP in variable X

$$\max - \mathbf{Tr} M(0) X \quad : \quad \begin{array}{l} \mathbf{Tr} M_i X = c_i, \; 1 \le i \le m, \\ \Phi_{\mathbf{\Delta}}\left(NXN^T\right) \ge 0, \;\; X \ge 0 \end{array}$$

The above is the dual of the SDP obtained via Lagrange relaxation. We will see in section 1.7.1 how this connection between Lagrange and rank relaxation has been recognized earlier in the special case of combinatorial optimization problems. For a related point of view on the geometrical properties of rank-constrained symmetric matrices, see [48] and references therein.

1.4.4 Extension to non convex nominal problems

In many cases, notably in control synthesis, we encounter problems where the nominal problem is non convex. Most of these problems can be turned into *cone complementarity* problems (see section 1.5.2 for more details).

We have seen how to handle uncertainty in (convex) SDPs, which are *symmetric* eigenvalue problems. To illustrate the ideas, let us consider a (non convex) *unsymmetric* eigenvalue problem. For example, consider the following robust synthesis problem:

$$\min_x \max_{\Delta \in \mathbf{\Delta}} \rho\left(\mathbf{M}(x, \Delta)\right),$$

where $\rho(\cdot)$ denotes the spectral radius, and $\mathbf{M}(x, \Delta)$ is affine in x and rational in $\Delta \in \mathbf{\Delta}$, given in a LFR

$$\mathbf{M}(x, \Delta) = M(x) + L(x)\Delta(I - D\Delta)^{-1}R,$$

(We note that the fact that $\mathbf{M}$ is affine in x is done without much loss of generality.)

Robustness analysis

Let us fix λ, x, seek sufficient condition for

$$\rho\left(\mathbf{M}(x, \Delta)\right) \leq \lambda \text{ for every } \Delta \in \mathbf{\Delta}$$

Let us first "forget" uncertainty, and apply Lagrange relaxation to the problem of checking if

$$(zI - M)^H (zI - M) \leq \lambda^2 I \text{ for every } z,\ |z| \leq 1.$$

We obtain that λ is an upper bound on spectral radius iff there exist $X > 0$ such that

$$\lambda^2 X \geq M^T X M.$$

The above condition is an LMI in X, which implies that we can apply the robust semidefinite programming methodology to handle the case when M is uncertain. We obtain that λ is an upper bound on maximal spectral radius if there exist $P > 0$ and S such that

$$\begin{bmatrix} \lambda^2 P & 0 \\ 0 & 0 \end{bmatrix} \geq \begin{bmatrix} M^T(x) \\ L^T(x) \end{bmatrix} P \begin{bmatrix} M^T(x) \\ L^T(x) \end{bmatrix}^T + \begin{bmatrix} R & D \\ 0 & I \end{bmatrix}^T \Phi^*_{\mathbf{\Delta}}(S) \begin{bmatrix} R & D \\ 0 & I \end{bmatrix}$$

For fixed x, minimizing λ a (quasi-) convex problem (over S and P), resulting in an *upper* bound on the worst-case spectral radius.

Robust synthesis

We now seek a sufficient condition for, λ being fixed,

$$\exists\; x \text{ such that } \rho\left(\mathbf{M}(x,\Delta)\right) \leq \lambda \text{ for every } \Delta \in \boldsymbol{\Delta} \tag{1.27}$$

holds. Now, x becomes a variable. Due to possible cross-products between x and P in the previous sufficient condition, the problem is not convex, contrarily to what happens in the robust SDP case.

Let $X = \lambda^2 P$, $Y = P^{-1}$, and write problem as a *conic complementarity problem*

$$\min\ \mathbf{Tr} XY \text{ subject to } \begin{bmatrix} X & \lambda I \\ \lambda I & Y \end{bmatrix} \geq 0,$$

$$\begin{bmatrix} X & 0 & M(x)^T \\ 0 & 0 & L(x)^T \\ M(x) & L(x) & Y \end{bmatrix} \geq \begin{bmatrix} R & D & 0 \\ 0 & I & 0 \end{bmatrix}^T \Phi^*_{\boldsymbol{\Delta}}(S) \begin{bmatrix} R & D & 0 \\ 0 & I & 0 \end{bmatrix}$$

Condition (1.27) is true iff optimum is $\lambda^2 n$.

The problem we obtain is a generalization of SDP, in the sense that SDP is a special complementarity problem (see section 1.5). In general, the above problem is not convex; however, several efficient techniques can be used (see Part V). One of these techniques, described in Chapters 16 and 17, rely on a simple linearization of the quadratic objective, and lead to a sequence of SDPs. In some cases, this approximation technique is guaranteed to yield the global optimum of the original problem (see [16]).

1.4.5 On quality of relaxations

Associated with the relaxation methods, comes the need to evaluate the quality of the approximations involved. In the context of robust SDP, the first results are due to Nemirovskii [8].

To illustrate the ideas involved in this analysis, consider the following robustness analysis problem: given l symmetric $n \times n$ $F_1, \ldots, F_l$, and a scalar ρ, check if

$$\mathcal{P}_\rho\ :\ I \geq \sum_{i=1}^{l} \delta_i F_i \text{ for every } \delta,\ \|\delta\|_2 \leq \rho. \tag{1.28}$$

This seemingly simple problem is NP-hard. If we apply Lagrange relaxation, we obtain that a sufficient condition for property $\mathcal{P}_1$ (*i.e.*, $\rho = 1$ in (1.28)) to hold is:

$$\exists S_i > 0,\ \ i = 1, \ldots, l,\ \ 2 \cdot I \geq \textstyle\sum_{i=1}^{l} \left(F_i S_i^{-1} F_i + S_i \right) \tag{1.29}$$

To evaluate the quality of the above approximation, we seek the smallest ρ such that condition (1.29) *implies* $\mathcal{P}_\rho$ holds. Clearly, such a ρ satisfies $\rho \geq 1$; the closer to 1, the better the approximation is.

Nemirovskii has proved that $\rho \leq \min(\sqrt{n}, \sqrt{l})$. In terms of control, we interpret this result as a bound on the quality of the approximation to a kind of

"μ-analysis" problem [6], with affine dependences and Euclidean-norm bounds. More refined results, including for other norm bounds, are under way [65].

1.4.6 Link with Lyapunov Functions

There is more than one connection between the approach proposed here and Lyapunov theory; we will briefly hint at these connections here.

Lyapunov functions and Lagrange multipliers.

Lyapunov functions can be interpreted in the context of Lagrange relaxations, which establishes an interesting link between two "central" methods in control and optimization.

For the sake of simplicity, we first consider the problem of establishing stability of the LTI system

$$\dot{x} = Ax,$$

where $x \in \mathbf{R}^n$ is the state, and $A \in \mathbf{R}^{n\times n}$ is a (constant) square matrix. It is well-known that the system is stable if and only if

$$(sI - A)^H(sI - A) > 0 \text{ for every } s, \;\; s + s^* \geq 0.$$

This shows that the stability problem is a robustness analysis problem (with respect to "uncertain" parameter s).

To attack it using Lagrange relaxation, we first use a linear-fractional representation of the problem. Introduce $p = sx$, our robustness analysis problem is equivalent to checking if

$$\|p - Ax\|^2 > 0 \text{ for every } (p, x) \neq (0,0), \;\; p = sx \text{ for some } s, \;\; s + s^* \geq 0.$$

Let us eliminate the "uncertain parameter" s by noting that

$$p = sx \text{ for some } s, \;\; s + s^* \geq 0 \Longleftrightarrow px^H + xp^H \geq 0$$

Let us apply a Lagrange relaxation idea now. A sufficient condition for stability is that there exists a "multiplier" matrix $S > 0$ such that

$$\forall (p, x) \neq (0,0), \;\; \|p - Ax\|^2 > \mathbf{Tr} S \left(px^H + xp^H\right)$$

A simple computation leads to the classical Lyapunov condition

$$\exists \; S \text{ such that } A^T S + SA < 0, \;\; S > 0,$$

which corresponds to the Lyapunov function candidate $V(x) = x^T Sx$.

Let us give another example of the "Lagrange relaxation" method for dynamical systems, consider a system with delayed state:

$$\dot{x}(t) = Ax(t) + A_d x(t-h). \tag{1.30}$$

The system is stable *independent of the delay* $\tau \geq 0$ if

$$\det(sI - A - zA_d) \neq 0 \text{ for every } s, z, \quad s + s^* \geq 0 \text{ and } |z| \leq 1.$$

An equivalent stability condition is that

$$\begin{array}{c} \|p - Aq - A_d p_d\|^2 > 0 \text{ for every } (p, p_d, q) \neq (0,0,0) \text{ such that} \\ pq^H + qp^H \geq 0, \quad p_d p_d^H \leq qq^H \end{array}$$

This time, we may relax the *two* matrix inequality constraints, which after some matrix calculations, yields the *sufficient* (possibly conservative) condition for stability: There exists symmetric matrices S, Z such that $S > 0$, $Z > 0$, and

$$\left[\begin{array}{cc} A^T S + SA + Z & SA_d \\ A_d^T S & -Z \end{array} \right] < 0.$$

To this condition, we can associate a "Lyapunov function candidate", termed as a Lyapunov-Krasovskii functional, defined as

$$V = x(t)^T S x(t) + \int_{-\tau}^{0} x(t+\theta)^T Z x(t+\theta) \, d\theta.$$

Quadratic stability.

The robust decision methodology can also be applied in the quadratic stability analysis for an uncertain system subject to unknown-but-bounded uncertainty:

$$\dot{x} = \mathbf{A}(\Delta) x, \quad \Delta \in \mathbf{\Delta},$$

where $\mathbf{\Delta}$ satisfies the assumptions of section 1.3.1, and $\mathbf{A}$ is defined via the LFR

$$\mathbf{A}(\Delta) = A + B_p \Delta (I - D_{qp} \Delta)^{-1} C_q. \tag{1.31}$$

(Recall that we assume that this LFR is well-posed over $\mathbf{\Delta}$.)

To analyze stability, the idea is to apply Lagrange relaxation twice. In a first step, we "forget" uncertainty and apply Lagrange relaxation to the "frozen" system. This results in a sufficient condition for stability: There exists a matrix S such that

$$\mathbf{A}(\Delta)^T S + S\mathbf{A}(\Delta) < 0, \quad S > 0 \text{ for every } \Delta \in \mathbf{\Delta}.$$

This condition is known as a quadratic stability condition, since it requires the existence of a quadratic Lyapunov function proving stability of the uncertain system.

In general, the above condition is still not tractable, despite the fact it is a convex condition on S. However, the relaxed problem above now belongs to the realm of robust SDP. We can apply Lagrange relaxation again. We obtain that if there exist matrices S and W such that

$$\begin{array}{l} \left[\begin{array}{cc} A^T S + SA & B_p \\ B_p^T & 0 \end{array}\right] + \left[\begin{array}{cc} C_q & D_{qp} \\ 0 & I \end{array}\right]^T \Phi_{\mathbf{\Delta}}^*(W) \left[\begin{array}{cc} C_q & D_{qp} \\ 0 & I \end{array}\right] < 0, \\ W > 0, \end{array} \tag{1.32}$$

holds, where $\Phi_{\mathbf{\Delta}}$ is the linear operator that characterizes the uncertainty set $\mathbf{\Delta}$, then the system is quadratically stable.

The above problem is an SDP, with decision variable S (linked to the quadratic Lyapunov function proving quadratic stability), and W. When $\mathbf{\Delta}$ is the set defined by (1.14), the above condition is the well-known small gain theorem (see [5] and references therein).

The above idea can be extended in various ways, as illustrated in some other chapters (Parts III and IV) of this book. For example, we may consider parameter-dependent, or frequency-dependent Lyapunov functions, based on so-called Integral Quadratic Constraints (IQCs).

Stochastic Lyapunov functions

Stochastic Lyapunov functions have been introduced by Kushner [52], and their use in the context of LMI optimization has been introduced in [5, 30, 31]. Consider a stochastic system of the form

$$dx = (Adt + Gdw)\, x$$

where w is a scalar standard Wiener process (the "derivative" of this process can be interpreted as a random perturbation on the state-space matrix, hence the name multiplicative noise given to such perturbations). The above systems are useful models in a number of areas, most notably finance [57].

The second moment $X = \mathbf{E}(xx^T)$ obeys the deterministic equation

$$\dot{X} = AX + XA^T + GXG^T. \tag{1.33}$$

The stability of this system hinges on the existence of a *linear* Lyapunov function

$$V(X) = \mathbf{Tr}(XS) = \mathbf{E}(x^T Sx) \tag{1.34}$$

where $S > 0$, which proves stability of the deterministic system (1.33) (evolving over the positive semidefinite cone). It can be shown that the (second-moment) stability is guaranteed if and only if

$$S > 0, \quad A^T S + SA + G^T SG < 0$$

To this Lyapunov function, corresponds a stochastic Lyapunov function (in the sense of Kushner), namely $v(x) = \mathbf{E}(x^T Sx)$.

One of the advantages of LMI formulations to stochastic control problems is that, once an LMI formulation is known, it is possible to apply Lagrange relaxation in the case when the "data" (for example the "nominal" matrix A in the above) is uncertain.

1.5 Algorithms and Software for SDP[4]

Algorithms for SDPs are now the subject of intense research in the field of optimization, especially since 1995. This progress should have a great impact on LMI-based control, as algorithms become more efficient. Recent methods to solve a general SDP can be roughly divided in two classes: interior-point methods, and nondifferentiable optimization methods. These algorithms have been implemented, and several interior-point software packages are now generally available.

1.5.1 Facts about SDP and related problems

We briefly recall some important results on semidefinite programming and related problems.

Duality.

The problem dual to problem (1.37) is

$$\begin{array}{ll} \text{maximize} & -\mathbf{Tr} F_0 Z \\ \text{subject to} & Z \geq 0, \ \ \mathbf{Tr} F_i Z = c_i, \ \ i = 1, \ldots, m, \end{array} \tag{1.35}$$

where Z is a symmetric $N \times N$ matrix and c_i is the i-th coordinate of vector c. When both problems are strictly feasible (that is, when there exist x, Z which satisfy the constraints strictly), the existence of optimal points is guaranteed [22, thm.4.2.1], and both problems have equal optimal objectives.

Related convex problems.

We list several problems that are related to semidefinite programs.

Second-order cone programming problems are special cases of SDPs. They take the form

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & \|C_i x + d_i\| \leq e_i^T x + f_i, \ \ i = 1, \ldots, L, \end{array} \tag{1.36}$$

where $C_i \in \mathbf{R}^{n_i \times m}$, $d_i \in \mathbf{R}^{n_i}$, $e_i \in \mathbf{R}^m$, $f_i \in \mathbf{R}$, $i = 1, \ldots, L$. The special structure of SOCPs can be exploited in algorithms, to yield much faster methods than using general-purpose SDP methods, see *e.g.* [22, 54, 6], as well as Chapter 3 of this book.

Max-det problems are of the form

$$\text{minimize } c^T x - \log\det G(x) \text{ subject to } F(x) \geq 0 \tag{1.37}$$

where G is an affine function taking values in the set of symmetric matrices of order p, is called a max-det problem. Max-det problems are convex optimization problems and can be solved in polynomial-time with interior-point

[4]This section was written with precious help from Michael Overton and Francois Oustry.

methods [22, 95]. Chapter 2 of this book provides more details on max-det problems. Other interesting applications of max-det problems are reviewed in [97].

Generalized eigenvalue problems are generalizations of SDPs, of the form

$$\text{minimize } \lambda \text{ subject to } \lambda B(x) - A(x) \geq 0 \quad B(x) \geq 0, \quad C(x) \geq 0$$

Here, A, B, and C are symmetric matrices that depend affinely on $x \in \mathbf{R}^m$. GEVPs are not convex, but are quasi-convex, and are amenable to interior-point methods similar to those used for SDPs, see [4, 68].

1.5.2 Interior-point methods

One way to view interior-point methods is as a reduction of the original problem to a sequence of differentiable, unconstrained minimization subproblems. Each subproblem can then be efficiently solved, perhaps approximately. These methods can be described in the context of self-concordant barrier and potential functions; this is the point of view taken in [22, 26].

Primal-dual interior-point methods.

Without doubt, the most efficient interior-point methods for *linear programming* (LP) are the primal-dual interior-point methods [101]. When a primal-dual interior-point method is used, each subproblem can be described by a linear least squares problem, or equivalently a system of linear equations. These methods are very efficient for four reasons: (1) they enjoy robust convergence from infeasible starting points, with guaranteed complexity when implemented accordingly, though typically this is not done in practice; (2) they converge rapidly, even if the solutions are not unique; (3) the linear systems to be solved are sparse (or can be made sparse by, for example, using standard techniques to handle dense columns in the constraint matrix); consequently, very large problems can be solved; (4) the inherent ill-conditioning in the linear systems to be solved does not, in practice, greatly limit achievable accuracy. However, this ill-conditioning does mean that iterative methods (*e.g.*conjugate gradient) are not generally practical to solve the linear systems: the standard workhorse is sparse Cholesky.

Since 1995, there has been a flurry of papers on primal-dual interior-point methods for SDPs; a complete bibliography on this subject is outside the scope of this Chapter. Relevant references include [70, 5, 51, 62, 80, 91, 47, 26, 2, 50]. In these methods, the subproblem to be solved is again simply a system of linear equations, though this is not always equivalent to a least-squares problem. As in LP, these methods converge very rapidly; various complexity results are now available; software packages have also recently become available. However, primal-dual methods are not necessarily the best at exploting sparsity in the data. The linear system which must be solved at every iteration invariably has dimension m, the number of variables in the primal form of the SDP. Therefore, with present technology, primal-dual methods are highly efficient if m is not too large (say less than 1000), but if it is much larger, they are not practical unless sparsity can be exploited. See [12] for a primal alternative to primal-dual methods which quite effectively exploits sparsity in SDPs arising in applications to graph theory. Even this approach is, however, limited by the size of linear systems which can be solved.

In what follows, we briefly describe primal-dual path-following methods, following Kojima *et al.* [51].

Optimality conditions and central path.

When the primal problem (1.2) and its dual (1.35) are both strictly feasible, then we may write the following optimality conditions:

$$ZF = 0, \quad F \geq 0, \quad Z \geq 0, \quad (F, Z) \in \mathcal{L}, \tag{1.38}$$

where $\mathcal{L}$ is an affine subspace:

$$\mathcal{L} = \left\{ (F, Z) \;\middle|\; \begin{array}{c} F = F(x) \text{ for some } x \in \mathbf{R}^m, \\ c_i = \mathbf{Tr} Z F_i, \quad i = 1, \ldots, m \end{array} \right\}. \tag{1.39}$$

We can interpret the above conditions as complementarity conditions over the positive semidefinite cone, similar to those arising in linear programming [51]. The convexity of the original SDP is transported here into a property of monotonicity of $\mathcal{L}$ which is crucial for the algorithm outlined below to converge globally.

We introduce a parameter μ that modifies the above conditions, as follows:

$$ZF = \mu I, \quad F \geq 0, \quad Z \geq 0, \quad (F, Z) \in \mathcal{L}, \tag{1.40}$$

Under strict feasibility conditions, the above system defines a unique primal-dual pair (F_μ, Z_μ). Solving (1.40) for decreasing values of the parameter μ leads to an optimal solution. The path followed by (F_μ, Z_μ) as $\mu \to 0$ is referred to as the central path, which explains the qualification "path-following".

For each value of μ, the conditions (1.40) can be interpreted as the optimality conditions for a differentiable problem involving a barrier function. Thus, the above description can of interior-point methods can be interpreted in the context of barrier functions [22, 26].

Predictor and corrector steps.

Assume that we have found a pair (F_k, Z_k) that is close to the central path (*i.e.*, approximately solve (1.40)). To find the new iterate, two steps are taken. In the *predictor* step, we seek to approach the optimum, that is, to satisfy $ZF = 0$. In the *corrector* step, we seek to return close to the central path by ensuring $ZF = \mu I$.

The search directions $\delta x, \delta Z$ for each step (predictor and corrector) are computed by linearization of the constraint $ZF = 0$ (predictor) or $ZF = \mu I$

(corrector) around the current value of (x, Z). Each step thus gives rise to a linear system in the elements of δx, δZ. Note that ZF can be linearized in a number of ways, depending on the specific method used.

The step lengths may be chosen small enough so that the new iterates will be guaranteed to be close to the central path again (and thus, strictly feasible). Here, closeness is measured using a proximity function such as

$$\|Z^{1/2}FZ^{1/2} - \mu I\|.$$

(Many alternatives are possible, see [51].) Depending on the proximity function used, we may define long-step and short-step strategies (see for example [62]).

There are two different approaches to initialization which have become standard. One approach uses an "infeasible start", i.e. points which satisfy the cone inequality constraints but not the linear equality constraints. If the original problem turns out to be either primal or dual infeasible, iterates with large norms will typically be generated, alerting the user that the problem is most likely either primal or dual infeasible. The other approach is to construct a modified problem for which feasible starting points are known, which then guarantees that either the original problem will be solved or a certificate of infeasibility will be found [71]. Though very attractive theoretically, it remains unclear whether this approach is better than the infeasible-start approach in practice.

In fact, the whole issue of feasibility is much more subtle for SDP than it is for LP. An interesting discussion on this subject can be found in [58].

1.5.3 Nondifferentiable optimization methods

When the number of variables is sufficiently large, and sparsity cannot be effectively exploited, it may be advantageous to consider alternative methods to solve SDPs. For simplicity, we only discuss an LMI feasibility problem, where a solution to a set of LMIs is sought; such problems appear routinely in control, *e.g.* in stability analysis.

An LMI feasibility problem is easily formulated as eigenvalue minimization problem of the form

$$\text{minimize } \lambda_{\max}(F(x)), \tag{1.41}$$

where $F(x)$ is an affine function taking values in the set of real symmetric matrices, and $\lambda_{\max}$ denotes largest eigenvalue. (To see this, note that there exists an x such that $F(x) \leq 0$ if and only if the minimum of $\lambda_{\max}(F(x))$ is below zero.) The function $f(x) = \lambda_{\max}(F(x))$ is a nondifferentiable function; yet it is convex and all theoretical tools from convex analysis apply to this function: in particular first-order and second-order generalized derivatives can be explicitly quantified and numerically implemented for this class of functions [74, 94, 33, 29, 6].

In Chapter 3 of this book, an extensive review of existing NDOMs is proposed ranging from early *subgradient methods* to the modern *bundle methods.*

Chapter 3 also includes some very recent developments on *second-order bundle methods* as well as numerical results of a new type: a mixed polyhedral-semidefinite bundle method enables us to check the stability of polytopic linear differential inclusions with $1,000$ state variables (which implies more than $m = 500,000$ variables). A second-order bundle method is shown to bring in practice an asymptotic superlinear rate of convergence for smaller problems. Some perspectives are sketched to see how the latter methods could be made efficient on large scale problems.

1.5.4 Software

Software for solving SDPs is generally available, both freely on the Internet and commercially. (To our knowledge, most of these SDP algorithms use interior-point methods.) Much of this has become available quite recently. Here is a (probably incomplete) list:

- The Matlab control toolbox, developed by Gahinet and others [40], uses Nesterov and Nemirovskii's projective algorithm [22, 66].

- The code **sdpsol**, described in Chapter 2 of this book is a parser/solver for SDPs and maxdet-problems. **sdpsol** uses the code SP, developed by Vandenberghe and Boyd [95].
  `http://www.stanford.edu/~boyd/group_index.html`
  The toolbox `lmitool`, developed by El Ghaoui, Nikoukhah and Delebecque [15], and later by Commeau, is an interface to the codes SP, SDPpack and SDPHA.
  `http://www.ensta.fr/~gropco/`

- The code **SDPpack**, outlined in Chapter 3 of this book, implements a primal-dual algorithm due to Alizadeh, Haeberly and Overton [5, 3].
  `http://www.cs.nyu.edu/cs/faculty/overton/sdppack/sdppack.html`

- The code **SDPA**: developed by Fujisawa, Kojima and Nakata; the underlying algorithm is described in [51].
  `ftp://ftp.is.titech.ac.jp/pub/OpRes/software/SDPA`

- The code **SDPHA**: developed by Brixius, Sheng and Potra; for a description of the underlying algorithm, see the web page:
  `http://www.cs.uiowa.edu/ brixius/SDPHA`

- The code **SDPT3**: uses the algoritm of Toh, Todd and Tütüncü [91].
  `http://www.math.nus.sg/~mattohkc/SDPT3.tar.Z`

- The code **SeDuMi**: developed by Sturm.
  `http://www.crl.mcmaster.ca/ASPC/people/sturm/software/`

1.6 Illustrations in Control

This book abounds with examples of applications of the LMI framework to control problems. In this section, we seek to provide complementary examples.

1.6.1 Optimal control and SDP duality

Notions central to SDP theory, such as duality, optimality conditions, sensitivity analysis, etc, have usually interesting interpretations for the analysis of (optimal) control problems. For example, the perturbation theory for general SDP recently developed by Shapiro, Bonnans and Cominetti [85, 86, 15] can be used for the analysis of Lyapunov and Riccati inequalities and equations encountered in optimal control.

The following example shows how SDP duality and optimality theory can be used in an optimal control problem. In a variety of stochastic control problems, a Riccati equation of the following form [100] arises:

$$A^TP + PA - PB^TR^{-1}BP + G^TPG + Q = 0, \tag{1.42}$$

where $Q = Q^T \geq 0$, $R > 0$, and A, B, G are matrices of appropriate sizes. When $G = 0$, we recover a standard Riccati equation (which can be solved with via the eigenvectors of a Hamiltonian matrix). In the general case $G \neq 0$, the problem of computing the solution is non trivial.

Under hypotheses that are physically reasonable, detailed below, we can solve the above equation using SDP, as follows. Consider the SDP

$$\begin{array}{c} \text{maximize } \mathbf{Tr}P \text{ subject to} \\ \begin{bmatrix} A^TP + PA + Q + G^TPG & PB \\ B^TP & R \end{bmatrix} \geq 0, \quad P = P^T. \end{array} \tag{1.43}$$

We form the problem dual to (1.43) (see §1.5):

$$\begin{array}{c} \text{minimize } \mathbf{Tr}(XU + SR) \text{ subject to} \\ \begin{bmatrix} X & U^T \\ U & S \end{bmatrix} \geq 0, \quad AX + BU + (AX + BU)^T + GXG^T + I = 0. \end{array} \tag{1.44}$$

The above problem is strictly feasible if and only if the underlying stochastic system is stabilizable in a stochastic sense (with static state-feedback control law $u = Kx$, $K = UX^{-1}$). Likewise, it is easy to show that the primal problem (1.43) is strictly feasible when this system is detectable. Let us make the hypothesis of primal and dual strict feasibility.

Since both primal and dual problems are strictly feasible, it follows that a set of feasible primal-dual variables (P, X, U, S) are optimal if and only if

$$\begin{bmatrix} X & U^T \\ U & S \end{bmatrix} \begin{bmatrix} A^TP + PA + Q + G^TPG & PB \\ B^TP & R \end{bmatrix} = 0,$$

from which we get

$$\begin{array}{rcl} X(A^TP + PA + Q + G^TPG) + U^TB^TP & = & 0, \\ XPB + U^TR & = & 0. \end{array} \tag{1.45}$$

With $R > 0$, we obtain $U = KX$, where $K = -R^{-1}B^TP$. The first equation of (1.45) then writes

$$X(A^TP + PA - PB^TR^{-1}BP + Q + G^TPG) = 0.$$

Finally, we observe that $X > 0$: if $Xz = 0$ for some z, then $Uz = 0$ (from $U = KX$). Since U, X are feasible for (1.44), we obtain

$$0 = z^T(AX + BU + (AX + BU)^T + GXG^T + I)z \geq z^Tz,$$

which implies $z = 0$. The conclusion is that P satisfies the Riccati equation; furthermore, the control law $u = Kx$ is stabilizing, since $X > 0$, and

$$\begin{array}{c} (A + BK)X + X(A + BK)^T + GXG^T = \\ AX + BU + (AX + BU)^T + GXG^T = -I < 0. \end{array}$$

For other references where duality theory and control links are exploited, see for example [98].

The above approach can be extended to some other classes of nonstandard Riccati equations, involving for example jump linear systems [1], problems with indefinite control weighting matrix R [78], etc.

1.6.2 Impulse differential systems

Impulse differential systems are systems subject to state discontinuities (or jumps). These systems are *hybrid*, in the sense that they obey to a mixture of discrete- and continuous-time behaviors. See [7, 84] for references, and also [103] for an application in the context of telecommunications. In classical control, hybrid systems are very common: they arise when one controls a (continuous-time) system with a discrete-time controller.

A typical impulse linear differential system is of the form

$$\begin{array}{c} \dot{x} = Ax, \quad t \neq \tau_i, \\ x(\tau_i^+) = Mx(\tau_i^-), \end{array}$$

where A, M are given square $n \times n$ matrices, and τ_i, $i = 0, 1, \ldots$ is a strictly increasing infinite sequence of time instants, which determine when the (otherwise smooth) trajectory suddenly jumps from $x(\tau_i^-)$ to $x(\tau_i^+)$. By proper choice of the impulse matrix M, one seeks to guarantee stability, using small impulses at properly chose time instants. (In the sequel, we assume the jumps are separated by a fixed interval T.)

The stability analysis of impulse systems is non trivial in general. In [103], Yang and Chua have devised (based on an earlier result by [84]) a stability criterion, using quadratic Lyapunov functions. It turns out that it is easy to rewrite this criterion in the form of an LMI condition, as follows. If there exists $S > 0$ such that

$$\begin{array}{c} A^T S + SA + 2\alpha S \leq 0, \quad \alpha \geq 0, \quad 2\alpha T > \log \lambda \\ \lambda S \geq M^T S M \end{array}$$

then the system is asymptoticaly stable.

The above conditions can be interpreted in terms of the ellipsoid $\mathcal{E} = \{\xi \mid \xi^T S \xi \leq 1\}$: the first LMI says that, during the "continuous-time behavior", the trajectory never escapes the (possibly bigger) ellipsoid $\{\xi \mid \xi^T S \xi \leq e^{-2\alpha T}\}$; the last two conditions ensure that the jump brings the state back to $\mathcal{E}$, which implies some kind of invariance for $\mathcal{E}$. Note that we do not assume the continuous system to be stable, since the decay rate α may be negative; in some sense, we are requiring that the "degree of stability" of the discrete-time behavior dominates that of the continuous-time one.

The above condition is an LMI in S for fixed λ, α, which shows that the problem of checking stability (based on quadratic stability) is tractable, provided a plane search (on λ, α) is used. The main advantage of the LMI formulation is that it allows for *designing* an appropriate impulse M: with the change of variables $U = SM$, the above conditions become convex in *both* S, U.

With an LMI (sufficient) condition stability in hand, we can address a variety of control problems.

For example, we can extend the results to nonlinear and/or uncertain systems; for example, we may allow the matrix A to be unknown-but-bounded, or consider the case when the system without impulses is LTI with a sector-bounded nonlinearity, and apply Lur'e stability conditions [5]. Likewise, it is possible to deal with optimal (state-feedback) control problems, using the framework used in section 1.6.1.

1.6.3 Delay differential systems

Delay systems arise in many applications, especially when a physical or biological transport phenomenon occurs (see *e.g.* Murray [64] for interesting examples in biology).

A typical problem in control of delay system is to analyze stability as a function of delay, where the delay is viewed as a *parameter* of the system. The problem is still largely open [26]; even in the linear case, and with multiple (non commensurable) delays, the problem is NP-hard (see Toker and Ozbay [93] for a precise statement and results).

We may distinguish *delay-independent* stability analysis (*i.e.* no information on the delay size is used) and *delay-dependent* stability analysis (based on this information). Sufficient stability conditions, both delay-independent or not, have been devised using LMIs; see *e.g.* [102, 55, 38]. Important tools in stability analysis are based on extensions of classical Lyapunov theory, leading to Krasovskii functionals or Razumikhin functions [46]. The Razumikhin theorem has been linked to the $\mathcal{S}$-procedure of section 1.2.4 in [73]. A survey of results and further comments are given in [73].

In section 1.4.6, we have given a delay-independent stability result, based on Lagrange relaxations, which is related to a particular choice of Lyapunov-Krasovskii functional candidate. This approach can be extended to a delay-dependent result, using an embedding technique (comparison principle), as follows.

Consider the stability analysis problem for the delay-differential system 1.30. We introduce the following distributed delay system

$$\dot{x} = (A + A_d)x(t) - \int_{-h}^{0} \Big(A_d A x(t+\theta) + A_d^2 x(t+\theta-h) \Big) \, d\theta,$$

and argue that the above system is an embedding of the original system 1.30, in the sense that stability of the above guarantees that of 1.30 (in fact, the above is an integration of the original equations over one delay interval). Using Laplace transforms, and straightforward computations, we obtain a sufficient condition for stability

$$\begin{array}{c} \det\left(sI - (A + A_d) - z_1 A_d A - z_2 A_d^2\right) \neq 0 \\ \text{for every } s, z_1, z_2 \;\; s + s^* \geq 0 \text{ and } |z_i| \leq h, \;\; i = 1, 2. \end{array}$$

It is then straightforward to apply the Lagrange relaxation technique (see section 1.4.6), and obtain the following delay-dependent relaxed LMI condition

$$\left[\begin{array}{ccc} \left(\begin{array}{c} h^{-1}\left[(A+A_d)^T P + P(A+A_d)\right] \\ +S_1 + S_2 \end{array} \right) & P A_d A & P(A_d)^2 \\ A^T A_d^T P & -S_1 & 0 \\ (A_d^T)^2 P & 0 & -S_2 \end{array} \right] < 0.$$

Computing an lower bound on the largest allowable delay is a GEVP. Refinements of these results can be found in [72].

To the above LMI condition corresponds a Lyapunov-Krasovskii functional candidate, of the form

$$V(x_t) = x(t)^T P x(t) + \int_{-h}^{0} \int_{t+\theta}^{t} x(v)^T S_1 x(v) \, dv d\theta + \int_{-2h}^{-h} \int_{t+\theta}^{t} x(v)^T S_2 x(v) \, dv d\theta.$$

The above ideas can be extended to multiple delays (seen as independent uncertain parameters). A different idea, leading to similar results, is based on the *IQC* Theorem, combined with the Kalman-Yakubovich-Popov Lemma [38] (see Chapters 6 for more on IQCs). The basic idea here is to view the delay as an uncertainty that satisfies a given IQC (see §1.3.1), and then apply IQC analysis.

It is possible to refine the results using a more computationally intensive LMI condition, using an idea similar in some sense to the one exploited in the context of robustness analysis with multipliers (see Chapters 6 and 11 for example). Indeed, for the system (1.30), we may exhibit a Lyapunov-Krasovskii functional which yields a necessary and sufficient condition for stability [49]. Unfortunately, this functional cannot be parametrized using a finite-dimensional function (as it is the case when working with LTI systems and quadratic Lyapunov functions); rather, it depends on several matrix functions (defined over an interval $[-\tau\ 0]$). It is possible to discretize this parametrization [45], and obtain LMIs which degree of conservativeness is arbitrarily low, with an inversely proportional computational effort.

1.6.4 Open-loop and predictive robust control

Robust optimization opens up the possibility of designing robust open-loop controllers for uncertain systems. This idea is suggested in [19], and we briefly elaborate on it now. Consider for example a dynamical system

$$\begin{aligned} x_{k+1} &= \mathbf{A}_k x_k + \mathbf{B}_k u_k, \\ y_k &= \mathbf{C}_k x_k + \mathbf{D}_k u_k \end{aligned}$$

where $\mathbf{A}_k$, $\mathbf{B}_k \in \mathbf{R}^{n \times n_u}$, $\mathbf{C}_k \in \mathbf{R}^{n_y \times n}$ and $\mathbf{D}_k$ are unknown-but-bounded time-varying matrices. We are given a time horizon $T \geq 0$, and denote by u and y the vectors $(u_0, \ldots, u_{T-1}) \in \mathbf{R}^{Tn_u}$ and $(y_0, \ldots, y_{T-1}) \in \mathbf{R}^{Tn_y}$. We also define three polytopes (or ellipsoids) $\mathcal{U} \subseteq \mathbf{R}^{Tn_u}$, $\mathcal{X}_0$ and $\mathcal{Y} \in \mathbf{R}^{Tn_y}$, sets of admissible controls, initial conditions, and outputs constraints, respectively.

The problem is to find if any $u \in \mathcal{U}$, if any, such that for every $x_0 \in \mathcal{X}_0$, the resulting output satisfies $y \in \mathcal{Y}$.

If the state-space matrices are exactly known, the problem is equivalent to a simple linear program. In the case when the state-space matrices are uncertain, the robust counterpart can be approximated using the robust decision methodology. In terms of optimization, this versatile framework gives rise to challenging large-scale problems.

One way to avoid the curse of dimensionality is to construct ellipsoids of confidence for the future state, and optimize (over a few time steps) the control vector. This idea is similar to the one used in model predictive control. The construction of the corresponding ellipsoids of confidence can be done using the techniques proposed shortly in section 1.7.2, in a recursive manner.

1.7 Robustness out of Control

The title of this section is borrowed from that of a talk by J. C. Doyle [28].

The LMI formalism provides a unitary framework, not only for various control problems, but also in other fields, where LMIs have recently emerged as a useful tool. We give some examples of this interplay, concentrating on areas where control-related notions (*e.g.*, robustness) seem to play a role which goes sometimes unrecognized yet, such as interval linear algebra.

1.7.1 LMIs in combinatorial optimization

SDPs are now recognized as yielding very efficient relaxations for (hard) combinatorial problems [42]. It turns out that these SDP relaxations for combinatorial problems can be obtained as special cases of the methodology outlined in §1.4. (The connection between $\mathcal{S}$-procedure and such relaxations is presented in [19].) In turn, the SDP relaxations open up interesting perspectives for handling *combinatorial problems with uncertain data,* in view of the tools developed for robust SDP.

We illustrate this via a simple example. Consider the NP-hard problem

$$\max_{\delta \in \mathbf{R}^l} \delta^T W \delta \text{ subject to } \delta_i^2 = 1, \quad i = 1, \ldots, l, \tag{1.46}$$

where W is a given symmetric matrix. When W assumes some special structure, the above problem is known in the combinatorial optimization literature as "the maximum cut" (MAX-CUT) problem [42, 19].

This problem is a robustness analysis problem, as defined in 1.2.3. First we note that, without loss of generality, we may assume $W > 0$, and rewrite the problem a robust synthesis problem:

$$\text{minimize } x \text{ subject to } \delta^T W \delta \leq x \text{ for every } \delta, \ \|\delta\|_\infty \leq 1.$$

Let us now apply our results, with

$$\mathbf{F}(x, \Delta) = \begin{bmatrix} x & \delta^T \\ \delta & W^{-1} \end{bmatrix} = F + L\Delta R + (L\Delta R)^T,$$

for appropriate matrices F, L, R, and with $\Delta = \mathbf{diag}(\delta_1, \ldots, \delta_l)$. The uncertainty set is

$$\boldsymbol{\Delta} = \{\Delta \text{ diagonal}, \ \|\Delta\| \leq 1\}.$$

The above set satisfies the assumptions made in section 1.3.1.

We obtain easily the following SDP approximation, yielding an upper bound on the original problem:

$$\min \mathbf{Tr} S \text{ subject to } W \leq S, \ S \text{ diagonal}. \tag{1.47}$$

In [42], Goemans and Williamson have also proposed an SDP relaxation (*i.e.*, an upper bound) of the problem. The technique is actually a special case of the general technique described in §1.4.3. Introduce the variable $X = \delta\delta^T$, and rewriting problem (1.46) in the equivalent form

$$\max_X \mathbf{Tr} WX \text{ subject to } \mathbf{rank} X = 1, \ \ X_{ii} = 1, \ \ X \geq 0.$$

Relaxing the non convex constraint $\mathbf{rank} X = 1$, we obtain an SDP that is exactly the "dual" (in the sense of SDP, see §1.5.1) of (1.47). As shown in [42], the above relaxation is the most efficient currently available (in terms of closeness to the actual optimum, in a certain stochastic sense).

Once an SDP formulation of the original problem is known, we can consider problems in which the "data matrix" W is uncertain, starting from the "nominal problem" 1.47. (In practice, such problems often arise in telecommunications, and the matrix W is estimated from statistics.)

1.7.2 Interval calculus

A basic problem in interval computations is the following. We are given a function $\mathbf{f}$ from $\mathbf{R}^l$ to $\mathbf{R}^m$, and a set confidence $\mathcal{D}$ for $\delta \in \mathbf{R}^l$, in the form of a product of intervals. We seek to estimate intervals of confidence for the components of $x = \mathbf{f}(\delta)$ when δ ranges $\mathcal{D}$. (Sometimes, $\mathbf{f}$ is given in implicit form, as in the interval linear algebra problem; we return to this case in the next section.) Obtaining exact estimates for intervals of confidence for the elements of solutions x, is in general NP-hard (see [82, 83] for the interval linear algebra case).

One classical approach to this problem resorts to interval calculus, introduced by Moore [63], where each one of the basic operations $(+, -, \times, /)$ is replaced by an "interval counterpart", and standard (eg LU) linear algebra algorithms are adapted to this new "algebra". Many refinements of this basic idea have been proposed, but the algorithms based on this idea have in general exponential complexity. (For control applications of interval calculus, see *e.g.* [25].)

Robust semidefinite programming can be used for this problem, as follows. Assume we can describe $\mathbf{f}$ explicitely as a rational function of its arguments; from lemma 1.2, we can construct (in polynomial time) an LFR of $\mathbf{f}$, in the form

$$\mathbf{f}(\delta) = f + L\Delta\,(I - D\Delta)^{-1}\,r, \text{ where } \Delta = \mathbf{diag}\,(\delta_1 I_{r_1}, \ldots, \delta_l I_{r_l})\,.$$

Assume first that we seek an ellipsoid of confidence for the solution, in the form

$$\mathcal{E} = \left\{ x \;\middle|\; (x - x_0)(x - x_0)^T \leq P \right\}, \tag{1.48}$$

where $x_0 \in \mathbf{R}^n$ is the center and $P \geq 0$ defines the "shape" (our parametrization allows for degenerate, "flat", ellipsoids, to handle cases when some components of the solution are certain). We seek to minimize the "size" of $\mathcal{E}$ subject to $\mathbf{f}(\delta) \in \mathcal{E}$ for every $\delta \in \mathcal{D}$. Measuring the size of $\mathcal{E}$ by $\mathbf{Tr}P$ (other measures are possible), we obtain the following equivalent formulation of the problem.

$$\text{minimize } \mathbf{Tr}P \text{ subject to}$$

$$\begin{bmatrix} P & (\mathbf{f}(\delta) - x_0) \\ (\mathbf{f}(\delta) - x_0)^T & 1 \end{bmatrix} \geq 0 \text{ for every } \delta \in \mathcal{D}. \tag{1.49}$$

The above is obviously a robust SDP problem (in variables x_0, P), for which an explicit SDP counterpart (approximation) can be devised, provided $\mathcal{D}$ takes the form of a (general) ellipsoidal set. (A typical set is a product of intervals $\Pi[\underline{\delta}_i\ \overline{\delta}_i]$, where $\underline{\delta}_i$, $\overline{\delta}_i$ are given.)

The above method finds ellipsoids of confidence, but it is also possible to find intervals of confidence for the components of $\mathbf{f}(\delta)$, by modifying the objective of the above robust SDP suitably (for example, if we minimize the $(1,1)$ component of the matrix variable P instead of its trace, we will obtain an interval of confidence for the first component of $\mathbf{f}(\delta)$, when δ ranges $\mathcal{D}$).

1.7.3 Structured linear algebra

Many linear algebra problems involve matrices with a given structure (in identification for instance, we obtain least squares problems in which the coefficient matrix has a Toeplitz structure). One of the most active research areas in numerical linear algebra pertains to exploiting perturbation structure, both for error analysis and for computing the solution.

Consider for example a *polynomial interpolation problem,* leading to a linear system of the form $Ax = b$, where x contains the polynomial coefficients, and A is a matrix with a Vandermonde structure. In some applications, the interpolation points are not precisely known. This results in an "uncertain Vandermonde system", of the form $\mathbf{A}(\delta)x = \mathbf{b}(\delta)$, where $\mathbf{A}(\delta)$, $\mathbf{b}(\delta)$ are matrix- (resp. vector-) valued polynomial functions of a perturbation vector δ.

A structured linear equation is a model of the form

$$\mathbf{A}(\delta)x = \mathbf{b}(\delta), \tag{1.50}$$

where $\mathbf{A}(\delta)$, $\mathbf{b}(\delta)$ are (affine) functions of a parameter vector $\delta \in \mathbf{R}^l$, and x is the variable. We assume that δ is unknown-but-bounded, for example $\|\delta\|_\infty \leq 1$. (Note that (1.50) is not an equation, but a model of the way an equation is subject to perturbations.) The above model can be represented in the linear-fractional form:

$$\begin{array}{l} Ax + Lp = b, \\ q = R_A x + R_b + Dp, \\ p = \Delta(\delta) q, \quad \Delta(\delta) = \mathbf{diag}(\delta_1 I_{r_1}, \ldots, \delta_l I_{r_l}). \end{array}$$

A number of problems pertain to the structured equation (1.50).

Robust least squares

One such problem is robust least squares, where we seek a vector x which minimizes the worst-case residual

$$r_S(\mathbf{A}, \mathbf{b}, x) \stackrel{\Delta}{=} \max_{\|\delta\|_\infty \leq 1} \|\mathbf{A}(\delta)x - \mathbf{b}(\delta)\|. \tag{1.51}$$

Such a vector can be termed robust in the sense that it minimizes the effect of perturbations on the "performance" (measured by the residual). The above problem has many implications in control, notably in robust identification.

The robust least-squares problem is a robust SDP problem, since the nominal problem, *i.e.* the classical least-squares problem, is a special case of SDP. Details on this approach are given in [33].

Condition estimates and accurate solutions

Another problem is to *measure* the sensitivity of a candidate solution x. A way to quantify this sensitivity (or backward error) is to compute

$$\max_{\delta, y} \|y - x\| \text{ subject to (1.50)}, \quad \|\delta\|_\infty \leq 1.$$

It is also interesting to compute a vector x which minimizes (an upper bound on) the above quantity. Such a vector is a solution for (1.50) with minimum sensitivity with respect to the perturbation vector δ; in this sense, it is accurate (with respect to the given perturbation structure and bounds).

As seen in the next paragraph, the above problems are robust SDP problems, which can be attacked using the methodology described previously.

Interval linear algebra

The basic problem of interval linear algebra is a slight variation on the problem considered in section 1.7.2, and is as follows. We are given matrices $\mathbf{A} \in \mathbf{R}^{n \times n}$, $\mathbf{b} \in \mathbf{R}^n$ the elements of which are only known within intervals; in other words, $[\mathbf{A}\ \mathbf{b}]$ is only known to belong to an "interval matrix set" $\mathcal{U}$. we seek to compute intervals of confidence for the set of solutions, if any, to the equation $\mathbf{A}x = \mathbf{b}$.

To attack this problem we seek an ellipsoid of confidence in the form (1.48), where $P > 0$. For simplicity, we assume first that A, b are given; using the $\mathcal{S}$-procedure, we find that a necessary and sufficient condition for $\mathcal{E}$ to be an ellipsoid of confidence is that there exist a multiplier $\tau \geq 0$ such that

$$\text{for every } y, \ (y - x_0)^T P^{-1} (y - x_0) + \tau \|Ay - b\|^2 \leq 1.$$

This condition is easily converted in a LMI in x_0, P. From then on, we can use the robust SDP methodology to handle the case when A, b are uncertain.

For further reference, we detail the result in the special case when $[\mathbf{A}\ \mathbf{b}]$ is an uncertain matrix, subject to "unstructured additive perturbations". Assume

$$[\mathbf{A}\ \mathbf{b}] \in \mathcal{U} = \{[A + \Delta A\ \ b + \Delta b] \ | \ \|[\Delta A\ \Delta b]\| \leq \rho\},$$

where $[A\ b] \in \mathbf{R}^{n \times (m+1)}$ and $\rho \geq 0$ are given. In this case, our results are exact, and yield an ellipsoid center of the form

$$x_0 = (A^T A - \rho^2 I)^{-1} A^T b. \tag{1.52}$$

(We assume that $\sigma_{\min}([A\ b]) \geq \rho$, otherwise the ellipsoid of confidence is unbounded. Except in degenerate cases, this guarantees the existence of the inverse in the above.)

Structured total least-squares

The so-called *structured total least squares* (STLS) problem, which arises in many areas, *e.g.* in identification and other inverse problems in engineering [23], is as follows. We are given $\mathbf{A}(\delta) \in \mathbf{R}^{m\times n}$, $\mathbf{b}(\delta) \in \mathbf{R}^m$, two (algebraic) functions of a parameter $\delta \in \mathbf{R}^l$. The STLS problem is

$$\text{minimize } \|\delta\| \text{ subject to } \mathbf{A}(\delta)x = \mathbf{b}(\delta). \tag{1.53}$$

Here, we are interested in finding one value of the "perturbation" vector δ which makes the vector x feasible. In this sense, the problem belongs to the class of "backward error analysis" problems: a perturbation is sought, such that the linear equation above becomes consistent.

The STLS problem can solved (at least approximately) by resorting to robust optimization, as follows. A scalar ρ is an upper bound on the objective in (1.53) if there exists $x \in \mathbf{R}^m$ such that

$$\mathbf{A}(\delta)x = \mathbf{b}(\delta) \text{ for some } \delta, \quad \|\delta\| \le \rho.$$

Assume that we have computed an ellipsoid of confidence of center x_0 and shape matrix P, denoted by $\mathcal{E}_\rho$, for the set of solutions of the above constraint. The quantity

$$\min\ \rho \text{ subject to } \mathcal{E}_\rho \text{ is not empty}$$

is an upper bound on the STLS problem. The above problem turns out to be a GEVP, with variables ρ, x_0, P.

In the unstructured (additive) perturbations case, we recover the classical notion of total least-squares solution developed by Golub and Van Loan [43]. The upper bound computed via GEVP is exact; the ellipsoid of confidence can be shown to be reduced to the singleton, and the optimal center is given by (1.52), where $\rho = \sigma_{\min}([A\ b])$.

1.7.4 Robust portfolio optimization

A portfolio is a collection of financial assets, say $p_1, \ldots, p_n$. Each asset p_i has a (random) return over a given horizon, denoted r_i. If x_i is the proportion of the total value of an asset p invested in the i-th asset, then the return of the portfolio over the given horizon is $\rho = x_1 r_1 + \ldots + x_n r_n = r^T x$.

The return of each asset is of course, unknown a priori. However, it is possible to deduce from past measurements the values of the mean $\mu = \mathbf{E}(r)$ and covariance matrix $V = \mathbf{E}(rr^T)$. The standard mean-variance portfolio choice problem, introduced in the fifties by Markowitz [60], is to minimize the variance of the asset's return, $x^T V x$, subject to a given level α of mean return, $\alpha = \mu^T x$.

This problem can be written:

$$\text{minimize } x^T V x \text{ subject to } x_i \ge 0, \quad e^T x = 1, \quad \mu^T x = \alpha.$$

(In the above, e stands for the n-vector with every component equal to one.) The problem above belongs to the class of a convex quadratic problem (a subset of SDP problems), and as such, can be very efficiently solved using interior-point methods.

One may then consider variations of this problem, and solve them using the LMI framework. One interesting example obtains when we assume that the mean and covariance matrices are unknown-but-bounded. This leads to portfolio choices that are robust with respect to uncertainty in mean and covariance. Another interesting perspective is to formulate the problem in a dynamic (multi-stage) setting, and apply robust control techniques to devise a robust portfolio update.

1.8 Perspectives and Challenges

1.8.1 Algorithms and software

Efficient algorithms and software should exploit structure of the given problem, in the most systematic way. For example, many problems in control lead to SDPs of the form (1.2), where the coefficient matrices are sparse. A sparse SDP solver would thus be extremely useful. Interior-point methods seem difficult to adapt directly to sparse LMI problems. To understand why, recall from §1.5.2 that at each step, we have to find (approximate) solutions to a complementarity equation of the form

$$FZ = \mu I,$$

where μ is a parameter, and F (resp. Z) is the primal (resp. dual) variable. Sparsity of F does not imply that of Z, and this makes the use of sparse computations uneasy to propagate throughout the iterations.

In contrast, the non-differentiable methods, described in §1.5.3, and in more detail in Chapter 3, seem promising for sparse SDPs. In particular, first-order (bundle) methods only require the computation of a (few) subgradient(s), and this can be done with sparse calculus. One problem with such methods is slow convergence, but we believe that using second-order information (such as U-Lagrangians), it will be possible to increase convergence speed, at a controlled increased cost per iteration (recall that with U-Lagrangian method, we work on a subspace of dimension equal to the multiplicity at the optimum, and this number is often much smaller than the size of the problem). More theoretical studies are needed to make this statement anything else than a belief. In particular, more work is needed on generic multiplicity at the optimum for SDPs, along the lines of [29, 79, 4].

Sparsity is not the only structure that we may take advantage of. SDPs which pertain to matrices with Toeplitz, or Lyapunov, etc, structure can often be solved much faster than general SDPs. (Struture is also a crucial information to use for solving linear equations.) This issue is raised in detail in the Chapter 4 of this book.

1.8.2 Non convex LMI-constrained problems

A number of problems arising in control and other fields are not convex. We encounter the following classes. (We only discuss feasibility problems, to simplify;

see Part V of this book for more on these problems.)

Bilinear matrix inequality problems are of the form

$$\text{Find } x \text{ such that } F(x) = F_0 + \sum_{i=1}^{m} x_i F_i + \sum_{i,j} x_i x_j F_{ij} \geq 0, \tag{1.54}$$

where F_i, G_{ij} are given symmetric matrices. This class contains much of combinatorial optimization; a typical example of such a problem is the MAX-CUT problem mentioned in §1.7.1.

Rank-constrained problems are of the form

$$\text{Find } x \text{ such that } F(x) \geq 0, \;\; \mathbf{Rank}\ G(x) \leq k,$$

where $F(\cdot), G(\cdot)$ are two affine matrix functions (with $F(\cdot)$ symmetric), and k is a given integer. Rank-constrained problems arise in *e.g.*, reduced-order feedback control (see [6, 14, 16]).

Conic complementarity problems consist of solving

$$\text{minimize } \mathbf{Tr} ZF, \quad (F, Z) \in \mathcal{K}, \tag{1.55}$$

where $\mathcal{K}$ is a convex subset of the cone of semidefinite matrices, and F, Z are variables. When the problem is "monotone" (that is when the set $\mathcal{K}$ is a "self-dual" cone), then the above problem can be solved using primal-dual interior-point methods for SDPs. For general cones $\mathcal{K}$ however, the above problem is hard. Conic complementarity problems are generalizations of the famous linear complementarity problems, and arise in many situations, including robust control (see [16]); see section 1.4.4 for an example.

The above problems are in general NP-hard. As we will see, they are polynomially equivalent, in the sense that we may transform one into the other in polynomial time. This equivalence can be used to devise efficient relaxation algorithms.

First, we may rewrite any BMI in the form of a RCP, as follows. Define $X = xx^T$, rewrite the constraint in (1.54) as

$$\begin{array}{c} F_0 + \sum_{i=1}^m x_i F_i + \sum_{i,j} X_{ij} F_{ij} \geq 0, \;\; \left[\begin{array}{cc} X & x \\ x^T & 1 \end{array}\right] \geq 0, \\ \mathbf{Rank} \left[\begin{array}{cc} X & x \\ x^T & 1 \end{array}\right] = 1. \end{array} \tag{1.56}$$

We obtain a rank-constrained problem. Dropping the rank constraint yields an LMI condition, which provides us with a relaxation of the original BMI problems, very similar to the one used in combinatorial optimization (see §1.7.1).

Conversely, any rank-constrained problem is a BMI, and a CCP. First we may assume without loss of generality that the matrix $G(x)$ is square ($n_G \times n_G$) and symmetric. Next, the constraint $\mathbf{Rank} G(x) \leq k$ is equivalent to saying that $\mathbf{Tr} ZF = 0$, $F \geq 0$, $Z \geq 0$, where

$$F = \left[\begin{array}{cc} G(x) & M \\ M^T & I \end{array}\right], \;\; Z = \left[\begin{array}{cc} S & U \\ U^T & V \end{array}\right] \tag{1.57}$$

for some $M \in \mathbf{R}^{n_G \times k}$, $S \in \mathbf{R}^{n_G \times n_G}$, $S \geq I$, U, V of appropriate size. Note that the constraint $\mathbf{Tr} ZF =$ can be written as a BMI, which shows the polynomial equivalence between BMIs and RCPs.

By solving the CCP

$$\text{minimize } \mathbf{Tr} ZF \text{ subject to } (Z, F) \in \mathcal{K}$$

where $\mathcal{K}$ denotes the cone of positive semidefinite pairs (F, Z) of the form (1.57), with $S \geq I$, we solve the original RCP.

The conic complementarity formulation seems very similar to the problem we encounter in primal-dual interior-point methods for solving SDPs. This suggests an algorithmic approach to such problems, where we directly apply a primal-dual interior-point method to the problem. (Of course, this method will converge globally when the complementarity problem is monotone). For non monotone problems, the method will at least provide a local minimum.

1.8.3 Dynamical systems over cones

The developments made in section 1.4.6 seem to indicate that, to analyze a (possibly stochastic) dynamical system with a vector state $x \in \mathbf{R}^n$ in the context of LMIs, it is fruitful to consider the (equivalent) system describing the evolution of the matrix $X = xx^T$ (or its expected value, if a stochastic system is at hand).

By definition, the resulting (deterministic) dynamical system leaves the positive semidefinite cone invariant. This motivates a closer study of dynamical systems of the form

$$\dot{X} = \phi(X), \tag{1.58}$$

where ϕ is a deterministic map, with the property that the positive semidefinite cone is invariant for (1.58). There seem to be a number of (apparently unexplored) interesting connections between several optimization and control notions, all related to systems of the form above.

For example, for the LTI system $\dot{x} = Ax$, the resulting cone-invariant system takes the form (1.58), with $\phi(X) = A^T X + XA^T$. A (linear) Lyapunov candidate of the form $V(X) = \mathbf{Tr} XS$ (with $S > 0$ given) proving stability of system (1.58) can be interpreted as a *primal-dual* gap (X is the "primal" variable, S the "dual" one), and is related to a quadratic Lyapunov function proving stability of the original (state-vector) system. The complementarity condition of the form (1.40) encountered in SDP, where a product of primal and dual variables are inverse of each other, can be related to the notion of adjoint system. Indeed, we note that if $X(t)$ is a trajectory of (1.58), with $X(t)$ invertible for every $t \geq 0$, then the inverse matrix $S(t) = X(t)^{-1}$ satifies the adjoint system

$$-\dot{S} = \phi^*(S) = A^T S + SA.$$

It would be interesting to see how the above connections can be exploited in control and optimization problems, and if the study of cone-invariant systems (1.58) can be made for other cones than the positive semidefinite cone.

1.8.4 Quality of relaxations for uncertain systems

As seen in 1.4.5, the problem of estimating the quality of relaxations is crucial in robust SDP, and this opens several challenging new problems. Good quality estimates would be of great importance for robust control, as they would enable to bound the conservativeness of well-known tools such as μ-analysis.

Consider for example the uncertain dynamical system

$$\dot{x} = \mathbf{A}(\delta(t))x, \quad \|\delta(t)\|_2 \leq 1 \text{ for every } t \geq 0, \tag{1.59}$$

where the $n \times n$ matrix $\mathbf{A}$ is an algebraic function of its argument $\delta \in \mathbf{R}^l$.

A sufficient condition for stability is quadratic stability (see section 1.4.6):

$$\exists\ P > 0, \quad \mathbf{A}(\delta)^T P + P\mathbf{A}(\delta) < 0 \text{ for every } \delta, \quad \|\delta\|_2 \leq 1.$$

In turn, the above condition can be relaxed using robust SDP techniques, leading to a (tractable, sufficient) LMI condition, of the form (1.32).

In some cases, it is possible to assess the possible conservativeness of this LMI condition, over the quadratic stability one. For example, if $\mathbf{A}$ is affine in δ, then Nemirovskii's results apply (see section 1.4.5), leading to a quality bound (in the sense defined in section 1.4.5) of $\min(\sqrt{n}, \sqrt{l})$.

The more general case when the matrix $\mathbf{A}$ is nonlinear (algebraic) in the parameter vector δ can be perhaps attacked using a singular perturbation technique, as follows. Represent $\mathbf{A}$ in the LFR format

$$\mathbf{A}(\delta) = A + B_p\Delta(I - D_{qp}\Delta)^{-1}C_q, \quad \Delta = \mathbf{diag}(\delta_1 I_{r_1}, \ldots, \delta_l I_{r_l}).$$

As usual, we assume this LFR to be well-posed over $\mathbf{\Delta}$. Now let $\epsilon > 0$, and associate to system 1.59 the following "augmented" system

$$\begin{array}{rcl} \dot{x} & = & Ax + B_p p, \\ \epsilon\dot{p} & = & \Delta(t)C_q x + (\Delta(t)D_{qp} - I)p, \quad \Delta(t) \in \mathbf{\Delta} \text{ for every } t \geq 0, \end{array} \tag{1.60}$$

where $\mathbf{\Delta}$ is the set defined in (1.15).

Under certain conditions classical in singular perturbations theory, involving (robust) stability of the matrix $D_{qp}\Delta - I$ for $\Delta \in \mathbf{\Delta}$[8], the above system can serve as an approximation to system 1.59. For every $\epsilon > 0$, the above system is equivalent to an uncertain linear system, with *affine* dependence on the uncertain parameters δ. For this system, the conservativeness estimate for quadratic stability is $\min(\sqrt{n+N}, \sqrt{l})$, where $N = r_1 + \ldots + r_l$. Since by definition $r_i \geq 1$, we obtain the conservativeness estimate $\sqrt{l}$.

The challenge here is to see how this analysis can be extended to the case of arbitrarily small ϵ, in order to give a rigorous conservativeness estimate.

[8] If $\mathbf{A}$ is polynomial in δ, then D_{qp} can be constructed as a strictly upper triangular matrix, and $D_{qp}\Delta - I$ is stable for every $\Delta \in \mathbf{\Delta}$.

1.9 Concluding Remarks

In this Chapter, we sought to give an overview of the LMI method. We proposed a unitary framework based on optimization ideas, for addressing a large class of decision problems with uncertainty. We have indicated some possible research directions where this method could be applied.

In control theory, to borrow words from Doyle *et. al* [6],

> *LMIs play the same central role in the postmodern theory as Lyapunov function and Riccati equations played in the modern, and in turn various graphical techniques such as Bode, Nyquist and Nichols plots played in the classical.*

Thus, the LMI approach constitutes a basis for a "postmodern control theory" [6], which allows robust, multicriteria synthesis. Benefits of both classical (PID) or modern (LQG, $\mathbf{H}_\infty$) approaches are combined: reduced number, and clear physical interpretation, of design parameters, and simplicity of numerical solution, even for multivariable problems. In addition, this approach provides guarantees for the design, and allows (possibly competing) specifications.

In our opinion, LMI-based control methods have reached a certain degree of "academic" maturity, even though many control areas are yet unexplored. Both theoretical grounds and efficient algorithms are now available, and this should lead to more and more industrial applications, some of which are in progress.

We strongly believe that the LMI approach could be applied to many other engineering problems, especially design problems with robustness and multicriteria issues. We presented some of these problems, with an optimization point of view. The reason is that the language of optimization seems the best suited as a common framework for these problems.

In some ways, we could widen the scope of the above quote from [6] by saying that LMIs (and convex optimization ideas) should play the same central role in "postmodern" engineering as numerical linear algebra (least squares, eigenvalue problems, etc) played in the past 20 years.

Acknowledgments

We are grateful to Michael Overton and Francois Oustry, who completely rewrote the section on Algorithms (Section 1.5). The first author is greatly indebted to Arkadi Nemirovskii for many stimulating discussions on the subject of robust optimization.

Bibliography

[1] M. Aitrami and L. El Ghaoui. LMI optimization for non standard Riccati equations arising in stochastic control. *IEEE Trans. Autom. & Control*, pages 1666–1670, 1996.

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

[3] F. Alizadeh, J.–P. A. Haeberly, M. V. Nayakkankuppam, M. L. Overton, and S. Schmieta. SDPpack User Guide (Version 0.9 Beta). Technical Report 737, Courant Institute of Mathematical Sciences, New York NY, June 1997. Also available from the SDPpack web page: http://www.cs.nyu.edu/faculty/overton/sdppack.

[4] F. Alizadeh, J.-P. A. Haeberly, and M. L. Overton. Complementarity and nondegeneracy in semidefinite programming. Technical Report 681, NYU Computer Science Dept, 1995. To appear in Mathematical Programming (Series B).

[5] F. Alizadeh, J.-P. A. Haeberly, and M. L. Overton. Primal-dual interior-point methods for semidefinite programming: Convergence rates, stability and numerical results. *SIAM J. Optimization*, 8(3):746–768, 1998.

[6] K. D. Andersen. An efficient newton barrier method for minimizing a sum of euclidean norms. *SIAM J. on Optimization*, 6(1):74–95, February 1996.

[7] B. D. Bainov and P. S. Simeonov. *Systems with impulse effect: stability, theory and applications*. Wiley, New York, 1989.

[8] A. Ben-Tal, L. El Ghaoui, and A. Nemirovski. Robust semidefinite programming. In H. Wolkowicz R. Saigal, L. Vandenberghe, editor, *Semidefinite Programming and Applications*, 1998. To appear.

[9] A. Ben-Tal and A. Nemirovski. Robust convex programming. Technical Report 1/95, Optimization Laboratory, Faculty of Industrial Engineering and Management, Technion, Israel Institute of Technology, Technion City, Haifa 32000, Israel, 1995.

[10] A. Ben-Tal and A. Nemirovski. Robust truss topology design. *SIAM J. Optimization*, 1997. To appear.

[11] A. Ben-Tal and A. Nemirovski. Robust convex programming. To appear in IMA J. Numer. Anal., March 1998.

[12] S. Benson, Y. Ye, and X. Zhang. Solving large-scale sparse semidefinite programs for combinatorial optimization. Technical report, Department of Management Sciences, The University of Iowa, Iowa City, IA 52242, Revised October 13 1997.

[13] V. Blondel and J. N. Tsitsiklis. NP-hardness of some linear control design problems. *European Jour. Control*, 1, 1995.

[14] V. Blondel and J. N. Tsitsiklis. A survey of computational complexity results in systems and control. Technical report, LIDS, MIT, Cambridge, MA 02139, October 1998.

[15] J. F. Bonnans, R. Cominetti, and A. Shapiro. Pertubed optimization under the second order regularity hypothesis, application to semidefinite optimization. In preparation, 1996.

[16] N. K. Bose. *Applied Multidimensional System Theory.* Van Nostrand Rheinhold, New York, 1982.

[17] S. Boyd and L. El Ghaoui. Method of centers for minimizing generalized eigenvalues. *Linear Algebra and Applications, special issue on Numerical Linear Algebra Methods in Control, Signals and Systems*, 188, July 1993.

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

[19] S. P. Boyd and L. Vandenberghe. Introduction to convex optimization with engineering applications. Lecture notes for ee392x, Stanford University, 1996. Available via anonymous ftp into `isl.stanford.edu`, under `/pub/boyd`.

[20] R. D. Braatz, P. M. Young, J. C. Doyle, and M. Morari. Computational complexity of μ calculation. In *Proc. American Control Conf.*, pages 1682–1683, San Francisco, CA, June 1993.

[21] G. E. Coxson and C. L. DeMarco. Testing robust stability of general matrix polytopes is an NP-hard computation. In *Proc. Annual Allerton Conf. on Communication, Control and Computing*, pages 105–106, Allerton House, Monticello, Illinois, 1991.

[22] M. A. Dahleh and I. Diaz-Bobillo. *Control of uncertain Sytems: A Linear Programming Approach.* Prentice-Hall, 1995. achete.

[23] B. de Moor. Total least squares for affinely structured matrices and the noisy realization problem. *IEEE Trans. Acoust., Speech, Signal Processing*, 42(11):3104–3113, November 1994.

[24] M.A.H. Dempster. *Stochastic programming*. Academic Press, 1980.

[25] O. Didrit, L. Jaulin, and E. Walter. Guaranteed analysis and optimization of parametric systems, with application to their stability degree. *European J. Control*, 1997.

[26] O. Diekmann, S. A. von Gils, S. M. Verduyn Lunel, and H. O. Walther. *Delay equations, functional-, complex- and nonlinear analysis*. Applied Math. Sciences. Springer-Verlag, New York, 1995.

[27] J. Doyle, A. Packard, and K. Zhou. Review of LFT's, LMI's and μ. In *Proc. IEEE Conf. on Decision and Control*, volume 2, pages 1227–1232, Brighton, December 1991.

[28] J. C. Doyle. Uncertainty and robustness out of control. Talk delivered at the *Workshop on Robust Control*, Siena, Italy, July 1998.

[29] A. El Bouhtouri and A. J. Pritchard. Stability radii of linear systems with respect to stochastic perturbations. *Syst. Control Letters*, 19(1):29–33, July 1992.

[30] L. El Ghaoui. State-feedback control of systems with multiplicative noise via Linear Matrix Inequalities. *Syst. Control Letters*, 24(3):223–228, February 1995.

[31] L. El Ghaoui and M. Aitrami. Robust state-feedback stabilization of jump linear systems. *Int. J. Robust and Nonlienar Control*, November 1996.

[32] L. El Ghaoui and P. Gahinet. Rank minimization under LMI constraints: a framework for output feedback problems. In *Proc. of the European Contr. Conf.*, July 1993.

[33] L. El Ghaoui and H. Lebret. Robust solutions to least-squares problems with uncertain data matrices. *SIAM J. Matrix Anal. Appl.*, October 1997.

[34] L. El Ghaoui, R. Nikoukhah, and F. Delebecque. `LMITOOL`: *A front-end for LMI optimization, user's guide*, February 1995. Available via anonymous ftp to `ftp.ensta.fr`, under `/pub/elghaoui/lmitool`.

[35] L. El Ghaoui, F. Oustry, and M. Aït Rami. A cone complementary linearization algorithm for static output-feedback and related problems. *IEEE Trans. Autom. & Control*, 42(8), August 1997.

[36] R. Fletcher. Semidefinite matrix constraints in optimization. *SIAM J. Control and Opt.*, 23:493–513, 1985.

[37] A. Fradkov and V. A. Yakubovich. The $\mathcal{S}$-procedure and duality theorems for non convex problems of quadratic programming. Technical Report 1, Vestnik Leningrad University, 1973. In Russian.

[38] M. Fu, H. Li, and S. I. Niculescu. *Stability and Control of delay systems*, chapter Robust stability analysis and stabilization for time-delay systems using the integral quadratic constraint approach. LNCIS. Springer-Verlag, London, 1997. Chapter 4.

[39] M. Fu and Z. Q. Luo. Computational complexity of a problem arising in fixed order output feedback design. *Sys. Contr. Lett.*, 1997. To appear.

[40] P. Gahinet, A. Nemirovski, A. Laub, and M. Chilali. *LMI control toolbox*. The Mathworks, Inc., Natick, Mass., May 1995.

[41] M. R. Garey and D. S. Johnson. *Computers and Intractability: A guide to the theory of NP-Completeness*. W. H. Freeman, 1979.

[42] M. X. Goemans and D. P. Williamson. .878-approximation for MAX CUT and MAX 2SAT. In *Proc.* 26$^{\text{th}}$ *ACM Symp. Theor. Computing*, pages 422–431, 1994.

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

[44] M. J. Green and D. J. N. Limebeer. *Linear robust control*. Prentice-Hall, Englewood Cliffs, 1995.

[45] K. Gu. Discretized lmi set in the stability problem of linear uncertain time-delay systems. *Int. J. Control*, 68:923–934, 1997.

[46] J. K. Hale and S. Verduyn Lunel. *Intoduction to functional differential equations*. Applied Math. Sciences. Springer-Verlag, New York, 1991.

[47] C. Helmberg, F. Rendl, R. Vanderbei, and H. Wolkowicz. An interior-point method for semidefinite programming. *SIAM J. Optimization*, 6:342–361, 1996.

[48] U. Helmke and J. B. Moore. *Optimization and dynamical systems*. Springer-Verlag, London, 1994.

[49] E. F. Infante and W. B. Castelan. A Lyapunov functional for a matrix difference-differential equation. *J. Differential Equations*, 29:439–451, 1978.

[50] F. Jarre. An interior-point method for minimizing the maximum eigenvalue of a linear combination of matrices. *SIAM J. Control Optim.*, 31(5):1360–1377, September 1993.

[51] M. Kojima, M. Shida, and S. Shindoh. Local convergence of predictor-corrector infeasible-interior-point algorithms for sdps and sdlcps. *Mathematical Programming*, 80:129–161, 1998.

[52] H. Kushner. *Introduction to Stochastic Control*. Holt, Rinehart and Winston, Inc., 1971.

[53] V. Lakshmikantam and S. Leela. *Differential and integral inequalities*, volume I and II. Academic press, New-York, 1969.

[54] H. Lebret. Antenna pattern synthesis through convex optimization. In *Advanced Signal Processing Algorithms*, pages 182–192. SPIE - The International Society for Optical Engineering, 1995.

[55] H. Li, S. I. Niculescu, L. Dugard, and J. M. Dion. Robust $\mathbf{h}_\infty$ control of uncertain linear time-delay systems: A linear matrix inequality approach with guaranteed α-stability. part II. In *Proc. 35th IEEE Conf. Dec. Contr.*, December 1996.

[56] W. M. Lu, K. Zhou, and J. C. Doyle. Stabilization of uncertain linear systems: an LFT approach. *IEEE Trans. Automat. Contr.*, 1:50–65, 1996.

[57] D. G. Luenberger. *Investment Science*. Prentice-Hall, 1997.

[58] Z.-Q Luo, J. F. Sturm, and S. Zhang. Duality and self-duality for conic convex programming. Technical Report EI 9620/A, Econometric Institute, Erasmus University Rotterdam, The Netherlands, 1996.

[59] A. I. Lur'e and V. N. Postnikov. On the theory of stability of control systems. *Applied mathematics and mechanics*, 8(3), 1944. In Russian.

[60] H. M. Markowitz. *Mean-variance analysis in portfolio choice and capital markets*. Basil Blackwell, 1987.

[61] A. Megretsky and S. Treil. Power distribution inequalities in optimization and robustness of uncertain systems. *Journal of Mathematical Systems, Estimation, and Control*, 3(3):301–319, 1993.

[62] R. D. C. Monteiro. Polynomial convergence of primal-dual algorithms for semidefinite programming based on monteiro and zhang family of directions. *SIAM J. Optimization*, 8(3):797–812, 1998.

[63] R. E. Moore. *Methods and applications of interval analysis*. SIAM, Philadelphia, 1979.

[64] J. D. Murray. *Mathematical Biology*. Springer-Verlag, 1993.

[65] A. Nemirovski. Subject: Quality of approximations. Private communication.

[66] A. Nemirovski and P. Gahinet. The projective method for solving linear matrix inequalities. *Mathematical Programming*, 77:163–190, 1997.

[67] A. Nemirovskii. Several NP-hard problems arising in robust stability analysis. *MCSS*, 6:99–105, 1993.

[68] Yu. Nesterov and A. Nemirovskii. An interior-point method for generalized linear-fractional problems. *Math. Programming, series B*, 1993.

[69] Yu. Nesterov and A. Nemirovsky. *Interior point polynomial methods in convex programming: Theory and applications.* SIAM, Philadelphia, PA, 1994.

[70] Yu Nesterov and M.J. Todd. Primal-dual interior-point methods for self-scaled cones. *SIAM J. Optimization*, 8(2):324–364, 1998.

[71] Yu.E. Nesterov, M.J. Todd, and Y. Ye. Infeasible-start primal-dual dethods and infeasibility detectors for nonlinear programming problems. Technical report, Department of Management Sciences, The University of Iowa, Iowa City, IA 52242, February 1998. To appear in *Mathematical Programming*.

[72] S.-I. Niculescu. On some frequency sweeping tests for delay-dependent stability: a model transformation case study. In preparation.

[73] S. I. Niculescu. *Systèmes à retard. Aspects qualitatifs sur la stabilité et la stabilisation.* Nouveaux Essais. Diderot, Paris, 1997.

[74] F. Oustry. The $\mathcal{U}$-Lagrangian of the maximum eigenvalue function. *SIAM J. Optimization*, 1997. To appear.

[75] F. Oustry, L. El Ghaoui, and H. Lebret. Robust solutions to uncertain semidefinite programs. *SIAM J. Optimization*, 1998. To appear.

[76] M. L. Overton and R.S. Womersley. Second derivatives for optimizing eigenvalues of symmetric matrices. *SIAM J. Matrix Anal. Appl.*, 16(3):667–718, July 1995.

[77] C. H. Papadimitriou. *Computational complexity.* Addison-Wesley, Reading, Mass., 1994.

[78] P. Park and T. Kailath. Dual |-spectral factorization. *Int. J. Control*, 1997.

[79] G. Pataki. On the rank of extreme matrices in semidefinite programs and the multiplicity of optimal eigenvalues. Technical Report MSRR-604(R), Management Science Research Group, GSIA, Carnegie Mellon University, August 1995.

[80] F. A. Potra and R. Sheng. Superlinear convergence of interior-point algorithms for semidefinite programming. Technical Report 86, Department of Mathematics, The University of Iowa, Iowa City, Iowa 52242, USA, November 1996.

[81] A. Rantzer and A. Megretski. System analysis via integral quadratic constraints. In *Proc. 34th IEEE Conf. Dec. Contr.*, pages 3062–3067, December 1994.

[82] J. Rohn. Systems of linear interval equations. *Linear Algebra and its Applications*, 126:39–78, 1989.

[83] J. Rohn. Overestimations in bounding solutions of perturbed linear equations. *Linear Algebra Appl.*, 262:55–66, September 1997.

[84] A. M. Samoilenko and N. A. Perestyuk. *Impulsive differential equations*. World Scientific, Singapore, 1995.

[85] A. Shapiro. Perturbation theory of nonlinear programs when the set of optimal solutions is not a singleton. *Applied Mathematics and Optimization*, 18:215–229, 1988.

[86] A. Shapiro. First and second order analysis of nonlinear semidefinite programs. *Mathematical Programming (Series B)*, 1996. to appear.

[87] A. Shapiro. First and second order analysis of nonlinear semidefinite programs. *Mathematical Programming (Series B)*, 77:301–320, 1997.

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

[89] R. S. Smith and M. Dahleh. *The modeling of uncertainty in control systems*. LNCIS. Springer-Verlag, Berlin, 1993.

[90] R. J. Stern and H. Wolkowicz. Indefinite trust region subproblems and nonsymmetric eigenvalue perturbations. *SIAM J. Optimization*, 5(2):286–313, May 1995.

[91] M. J. Todd, K. C. Toh, and R. H. Tütüncü. On the Nesterov-Todd direction in semidefinite programming. *SIAM J. Optimization*, 8(3):769–796, 1998.

[92] O. Toker. On the complexity of the robust stability problem for linear parameter varying systems. In 13*th Triennial IFAC World Congress*, volume H, pages 53–57, 1996.

[93] O. Toker and H. Ozbay. Complexity issues in robust stability of linear delay-differential systems. *Mathematics of Control, Signals, and Systems*, 1996.

[94] M. Torki. Epi-différentiabilité du second ordre de certaines fonctions spectrales et conditions d'optimalité en optimisation de valeurs propres. Technical report, LAO, Université Paul Sabatier, Toulouse, France, Décembre 1997.

[95] L. Vandenberghe and S. Boyd. **SP,** *Software for semidefinite programming, user's guide*, December 1994. Available via anonymous ftp to `isl.stanford.edu` under `/pub/boyd/semidef_prog`.

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

[97] L. Vandenberghe, S. Boyd, and A. El Gamal. Optimizing dominant time constant in RC circuits. IEEE Trans. on Computer-Aided Design, 1996. Submitted.

[98] P. G. Voulgaris. Optimal $\mathbf{h}_2/\mathbf{l}_1$ control via duality theory. *IEEE Trans. Autom. & Control*, 11:1881–1888, 1995.

[99] Z. Q. Wang and S. Skogerstad. *Analysis and Optimization of Systems: State and Frequency domain approaches for infinite–dimensional systems*, chapter Robust controller design for uncertain time delay systems, pages 610–623. LNCIS. Springer-Verlag, Berlin, 1993.

[100] W. M. Wonham. Random differential equations in control theory. In A. T. Bharucha-Reid, editor, *Probabilistic methods in applied mathematics*, volume 2, pages 131–212. New York: academic, 1970.

[101] S. Wright. *Primal-Dual Interior-Point Methods*. SIAM, 1997.

[102] L. Xi and C. E. de Souza. Lmi approach to delay-dependent robust stability and stabilization of uncertain linear delay systems. In *Proc. 35th IEEE Conf. Dec. Contr.*, December 1995.

[103] T. Yang and L. O. Chua. Impulsive stabilization for control and synchronization of chaotic systems: theory and application to secure communication. *IEEE Trans. Circuits Syst.*, 44(10):976–988, October 1997.

[104] K. Zhou, J. Doyle, and K. Glover. *Robust and Optimal Control*. Prentice Hall, New Jersey, 1995.