Semidefinite Programming Duality and Linear Time-invariant Systems

Venkataramanan Balakrishnan
Lieven Vandenberghe[1]

School of Electrical and Computer Engineering
Purdue University
West Lafayette, IN 47907–1285
Email: ragu@ecn.purdue.edu

[1] Dr. Vandenberghe is an Associate Professor at the Department of Electrical Engineering, University of California, Los Angeles, CA 90095–1594.

Contents

Abstract

Several important problems in control theory can be reformulated as semidefinite programming problems, i.e., minimization of a linear objective subject to Linear Matrix Inequality (LMI) constraints. From convex optimization duality theory, conditions for infeasibility of the LMIs as well as dual optimization problems can be formulated. These can in turn be re-interpreted in control or system theoretic terms, often yielding new results or new proofs for existing results from control theory. We explore such connections for a few problems associated with linear time-invariant systems.

1 Introduction

Over the past few years, convex optimization, and semidefinite programming[2] (SDP) in particular, have come to be recognized as valuable numerical tools for control system analysis and design. A number of publications can be found in the control literature that survey applications of SDP to the solution of system and control problems (see for example [BEFB94, SI95, BE, DP00, EN00]). In parallel, there has been considerable recent research on algorithms and software for the numerical solution of SDPs (for surveys, see [NN94, Ali95, LO96, VB96b, VB96a, WSV00]). This interest was primarily motivated by applications of SDP in combinatorial optimization but, more recently, also by the applications in control.

Thus far, the application of SDP in systems and control has been mainly motivated by the possibilities it offers for the numerical solution of analysis and synthesis problems for which no analytical solutions are known [GNLC95, GDN95, WB96]. In this paper, we explore another application of SDP: We discuss the application of duality theory to obtain new theoretical insight or to provide new proofs to existing results from system and control theory. Specifically, we discuss the following applications of SDP duality.

- Theorems of alternatives provide systematic and unified proofs of *necessary and sufficient conditions for solvability* of LMIs. As examples, we investigate the conditions for the existence of feasible solutions to Lyapunov and Riccati inequalities. As a by-product, we obtain a simple new proof of the Kalman-Yakubovich-Popov lemma.

  Several of the results that we use from convex duality require technical conditions (so-called *constraint qualifications*). We show that for problems involving Riccati inequalities these constraint qualifications are related to controllability and observability. In particular, we will obtain a new criterion for the controllability of an LTI system realization.

- The optimal solution of an SDP is characterized by *necessary and sufficient optimality conditions* that involve the dual variables. As an example, we show that the properties of the solution of the LQR problem can be derived directly from the SDP optimality conditions.

- The dual problem associated with an SDP can be used to derive *lower bounds* on the optimal value. As an example, we give new easily computed bounds on the H_∞-norm of an LTI system, and a duality-based proof of the Enns-Glover lower bound.

Several researchers have recently applied notions from convex optimization duality toward the re-interpretation of existing results and the derivation of new results in system theory. Rantzer [Ran96] uses ideas from convexity theory to give a new proof of the Kalman-Yakubovich-Popov

[2] We shall use SDP to mean both "semidefinite programming", as well as a "semidefinite program", i.e., a semidefinite programming problem.

Lemma. Henrion and Meinsma [HM01] apply SDP to provide a new proof of a generalized form of Lyapunov's matrix inequality on the location of the eigenvalues of a matrix in some region of the complex plane. Yao, Zhao, and Zhang [YZZ99] apply SDP optimality conditions to derive properties of the optimal solution of a stochastic linear-quadratic control problem. Our work is similar in spirit to these; however, the scope of our paper is wider, as we present (and in many cases generalize) some of the results in these papers.

The notation and terminology are standard. $\mathbf{R}$ ($\mathbf{R}_+$) denotes the set of real (nonnegative real) numbers. $\mathbf{C}$ denotes the set of complex numbers. The matrix inequalities $A > B$ and $A \geq B$ mean A and B are square, Hermitian, and that $A - B$ is positive definite and positive semi-definite, respectively. $\Re(\cdot)$ and $\Im(\cdot)$ denote respectively the real and imaginary parts of a complex scalar, vector or matrix. $\mathbf{L}_2$ is the Hilbert space of square-integrable signals defined over $\mathbf{R}_+$ (see for example [DV75]). $\mathbf{L}_{2e}$ denotes the extended space associated with $\mathbf{L}_2$.

2 Duality

Let $\mathcal{S}^n$ denote the set of Hermitian $n \times n$ matrices with an associated inner product $\langle \cdot, \cdot \rangle_{\mathcal{S}^n}$. While the development in this section and the sequel are applicable to any inner product on $\mathcal{S}^n$, we will assume that the standard inner product, given by $\langle A, B \rangle_{\mathcal{S}^n} = \mathbf{Tr}\, A^* B = \mathbf{Tr}\, AB$ is in effect. Let $\mathcal{S}$ denote the set of block diagonal Hermitian matrices with given dimensions, $\mathcal{S} = \mathcal{S}^{n_1} \times \cdots \times \mathcal{S}^{n_L}$, with inner product $\langle \mathbf{diag}(A_1, \ldots, A_L), \mathbf{diag}(B_1, \ldots, B_L) \rangle_{\mathcal{S}} = \sum_{k=1}^{L} \mathbf{Tr}\, A_k B_k$.

Suppose that $\mathcal{V}$ is a finite-dimensional vector space with an inner product $\langle \cdot, \cdot \rangle_{\mathcal{V}}$, $\mathcal{A} : \mathcal{V} \to \mathcal{S}$ is a linear mapping and $A_0 \in \mathcal{S}$. Then, the inequality

$$\mathcal{A}(x) + A_0 \geq 0 \tag{1}$$

is called a *Linear Matrix Inequality* or LMI. We let $\mathcal{A}^{\mathrm{adj}}$ denote the adjoint mapping of $\mathcal{A}$. That is, $\mathcal{A}^{\mathrm{adj}} : \mathcal{S} \to \mathcal{V}$ such that for all $x \in \mathcal{V}$ and $Z \in \mathcal{S}$, $\langle \mathcal{A}(x), Z \rangle_{\mathcal{S}} = \langle x, \mathcal{A}^{\mathrm{adj}}(Z) \rangle_{\mathcal{V}}$.

2.1 Theorems of alternatives

We first examine criteria for solvability of different types of LMIs. We consider the following three feasibility problems.

- *Strict feasibility:* there exists an $x \in \mathcal{V}$ with $\mathcal{A}(x) + A_0 > 0$.
- *Nonzero feasibility:* there exists an $x \in \mathcal{V}$ with $\mathcal{A}(x) + A_0 \gneqq 0$ (i.e., positive semidefinite and nonzero).
- *Feasibility:* there exists an $x \in \mathcal{V}$ with $\mathcal{A}(x) + A_0 \geq 0$.

By properly choosing $\mathcal{A}$ we will be able to address a wide variety of LMI feasibility problems. For example, when $\mathcal{V} = \mathbf{R}^m$, we can express $\mathcal{A}$ as

$$\mathcal{A}(x) = x_1A_1 + x_2A_2 + \cdots + x_mA_m, \tag{2}$$

where $A_i \in \mathcal{S}$ are given. With this parametrization, the three problems described above reduce to the following three basic LMIs:

$$A_0 + x_1A_1 + x_2A_2 + \cdots + x_mA_m \quad > \quad 0, \tag{3}$$

$$A_0 + x_1A_1 + x_2A_2 + \cdots + x_mA_m \quad \gneq \quad 0, \tag{4}$$

$$A_0 + x_1A_1 + x_2A_2 + \cdots + x_mA_m \quad \geq \quad 0. \tag{5}$$

There exists a rich literature on theorems of alternatives for generalized inequalities (i.e., inequalities with respect to nonpolyhedral convex cones), and linear matrix inequalities in particular. For our purposes the following three theorems will be sufficient. We refer to [BI69, BBI71, CK77, BW81] for more background on theorems of alternative for nonpolyhedral cones, and to [Wol81, Las95, Las97] for results on linear matrix inequalities.

Theorem 1 (ALT 1) *Exactly one of the following statements is true.*

1. *There exists an $x \in \mathcal{V}$ with $\mathcal{A}(x) + A_0 > 0$.*
2. *There exists a $Z \in \mathcal{S}$ with $Z \gneq 0$, $\mathcal{A}^{\mathrm{adj}}(Z) = 0$, and $\langle A_0, Z\rangle_{\mathcal{S}} \leq 0$.*

We refer to Appendix A for a proof of this theorem and the other theorems in this section.

Theorem ALT 1 is the first example of a *theorem of alternatives.* The two statements in the theorem are called *strong alternatives*, because exactly one of them is true.

Example 1 The adjoint $\mathcal{A}_0^{\mathrm{adj}} : \mathcal{S} \to \mathbf{R}^m$ of the mapping defined by (2) is given by

$$\mathcal{A}_0^{\mathrm{adj}}(Z) = \begin{bmatrix} \mathbf{Tr}\, A_1Z & \mathbf{Tr}\, A_2Z & \cdots & \mathbf{Tr}\, A_mZ \end{bmatrix}^T.$$

Theorem **ALT 1** therefore implies that either there exists $x \in \mathbf{R}^m$ such that LMI (3) holds, or there exists $Z \in \mathcal{S}$ with $Z \gneq 0$ such that $\mathbf{Tr}\, A_iZ = 0$, $i = 1, 2, \ldots, m$, and $\mathbf{Tr}\, A_0Z \leq 0$.

Example 2 As an example of an application of Theorem **ALT 1** in matrix algebra, consider Finsler's Theorem [Fin37, Jac77], which states that given $F \in \mathbf{C}^{n\times m}$ with rank $r < n$, and $G \in \mathcal{S}^n$, the condition that there exists $\mu \in \mathbf{R}$ such that $\mu FF^* - G > 0$, is equivalent to $(F^\perp)^*GF^\perp < 0$, where $F^\perp$ is a full-rank matrix whose columns span the left nullspace of F, i.e., $(F^\perp)^*F = 0$.

Let $\mathcal{A} : \mathbf{R} \to \mathcal{S}^n$ be defined by $\mathcal{A}(\mu) = \mu FF^*$, and let $A_0 = -G$. Then $\mathcal{A}^{\mathrm{adj}} : \mathcal{S}^n \to \mathbf{R}$ is given by $\mathcal{A}^{\mathrm{adj}}(Z) = \mathbf{Tr}(F^*ZF)$. Then, there does not exist $\mu \in \mathbf{R}$ such that $\mathcal{A}(\mu) + A_0 > 0$, if and only if there exists a $Z \in \mathcal{S}^n$ with $Z \gneq 0$, $\mathbf{Tr}(F^*ZF) = 0$, $\mathbf{Tr}(ZG) \geq 0$. Factoring Z as $Z = \sum_{i=1}^k \lambda_i u_i u_i^*$, where $\lambda_i > 0$, we must have for some i, $u_i^* F = 0$ and $u_i^* G u_i \geq 0$, which immediately means $(F^\perp)^* G F^\perp < 0$ is violated. Conversely, if $(F^\perp)^* G F^\perp \not< 0$, then for some nonzero $u \in \mathbf{C}^n$, we must have $u^*(F^\perp)^* G F^\perp u \geq 0$. Then, with $Z = (F^\perp u)(F^\perp u)^*$, it is readily verified that $Z \gneqq 0$ and $\mathbf{Tr}(F^* Z F) = 0$.

Theorem 2 (ALT 2) *At most one of the following statements is true.*

1. *There exists an $x \in \mathcal{V}$ with $\mathcal{A}(x) + A_0 \gneqq 0$.*
2. *There exists a $Z \in \mathcal{S}$ with $Z > 0$, $\mathcal{A}^{\mathrm{adj}}(Z) = 0$, and $\langle A_0, Z \rangle_{\mathcal{S}} \leq 0$.*

Moreover, if $A_0 = \mathcal{A}(x_0)$ for some $x_0 \in \mathcal{V}$, or if there exists no $x \in \mathcal{V}$ with $\mathcal{A}(x) \gneqq 0$, then exactly one of the two statements is true.

The theorem gives a pair of *weak alternatives*, i.e., two statements at most one of which is true. It also gives additional assumptions under which the statements become strong alternatives. These additional assumptions are called *constraint qualifications*.

Remark 1 Note that if $A_0 = \mathcal{A}(x_0)$ for some x_0, the theorem can be paraphrased as follows: Exactly one of the following statements is true.

1. There exists an $x \in \mathcal{V}$ with $\mathcal{A}(x) \gneqq 0$.
2. There exists a $Z \in \mathcal{S}$ with $Z > 0$ and $\mathcal{A}^{\mathrm{adj}}(Z) = 0$.

If in addition the mapping $\mathcal{A}$ has full rank, i.e., $\mathcal{A}(x) = 0$ implies $x = 0$, then the first statement is equivalent to $\mathcal{A}(x) \geq 0$, $x \neq 0$. $\diamond$

Example 3 Theorem **ALT 2** implies that at most one of the following are possible: either there exists $x \in \mathbf{R}^m$ such that LMI (4) holds, or there exists $Z \in \mathcal{S}$ with $Z > 0$, $\mathbf{Tr}\, A_i Z = 0$ for $i = 1, \ldots, m$, and $\mathbf{Tr}\, A_0 Z \leq 0$. However, it is possible that neither condition holds. As an example, take $\mathcal{S} = \mathcal{S}^2$ and

$$A_0 = \begin{bmatrix} 0 & 0 \\ 0 & -1 \end{bmatrix}, \quad A_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}.$$

The LMI $A_0 + x A_1 \gneqq 0$ is infeasible. The alternative is that there exists $Z \in \mathcal{S}^2$ with

$$Z = \begin{bmatrix} z_{11} & z_{12} \\ z_{12}^* & z_{22} \end{bmatrix} > 0, \quad z_{11} = 0, \quad z_{22} \geq 0$$

which is also infeasible.

Theorem 3 (ALT 3) *At most one of the following statements is true.*

1. *There exists an $x \in \mathcal{V}$ with $\mathcal{A}(x) + A_0 \geq 0$.*

2. *There exists a* $Z \in \mathcal{S}$ *with* $Z \geq 0$, $\mathcal{A}^{\text{adj}}(Z) = 0$, *and* $\langle A_0, Z\rangle_{\mathcal{S}} < 0$.

Moreover, if $A_0 = \mathcal{A}(x_0)$ *for some* $x_0 \in \mathcal{V}$*, or if there exists no* $x \in \mathcal{V}$ *such that* $\mathcal{A}(x) \gneqq 0$*, then exactly one of the two statements is true.*

Again, the theorem states a pair of weak alternatives, and additional assumptions under which the statements are strong alternatives.

Note that the theorem is trivial if $A_0 = \mathcal{A}(x_0)$ for some x_0: the first statement is true because we can take $x = -x_0$; the second statement is obviously false because $\mathcal{A}^{\text{adj}}(Z) = 0$ implies that

$$\langle A_0, Z\rangle_{\mathcal{S}} = \langle \mathcal{A}(x_0), Z\rangle_{\mathcal{S}} = \langle x_0, \mathcal{A}^{\text{adj}}(Z)\rangle_{\mathcal{V}} = 0.$$

Example 4 Theorem **ALT 3**, applied to the linear mapping (2), implies that at most one of the following are possible: either there exists $x \in \mathbf{R}^m$ such that LMI (5) holds, or there exists $Z \in \mathcal{S}$ with $Z \geq 0$ such that $\mathbf{Tr}\, A_i Z = 0$, $i = 1, 2, \ldots, m$, and $\mathbf{Tr}\, A_0 Z < 0$. It is possible that neither condition holds. Consider the following example, taken from [BI69, p.378]), where $\mathcal{S} = \mathcal{S}^2$ and

$$A_0 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad A_1 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}.$$

Then, the LMI $A_0 + xA_1 \geq 0$ is infeasible. The alternative is that there exists

$$Z = \begin{bmatrix} z_{11} & z_{12} \\ z_{12}^* & z_{22} \end{bmatrix} \geq 0, \text{ with } z_{22} = 0 \text{ and } z_{12} + z_{12}^* < 0$$

which is also infeasible.

For each of the theorems of alternatives **ALT 1–ALT 3**, we can formulate a version with equality constraints. Let $\mathcal{W}$ be a finite-dimensional vector space with the inner product $\langle\cdot,\cdot\rangle_{\mathcal{W}}$. Let $\mathcal{B} : \mathcal{V} \to \mathcal{W}$ be a linear mapping, and let $\mathcal{B}^{\text{adj}}$ denote the adjoint mapping of $\mathcal{B}$. Then, we have the following theorems.

Theorem 4 (ALT 4) *Exactly one of the following statements is true.*

1. *There exists an* $x \in \mathcal{V}$ *with* $\mathcal{A}(x) + A_0 > 0$ *and* $\mathcal{B}(x) = 0$.

2. *There exists a* $Z \in \mathcal{S}^n$ *with* $Z \gneqq 0$, *and* $w \in \mathcal{W}$, *with* $\mathcal{A}^{\text{adj}}(Z) + \mathcal{B}^{\text{adj}}(w) = 0$, *and* $\langle A_0, Z\rangle_{\mathcal{S}^n} \leq 0$.

Theorem 5 (ALT 5a) *Exactly one of the following statements is true.*

1. *There exists an* $x \in \mathcal{V}$ *with* $\mathcal{A}(x) \gneqq 0$ *and* $\mathcal{B}(x) = 0$.

2. *There exists a* $Z \in \mathcal{S}^n$ *with* $Z > 0$, *and* $w \in \mathcal{W}$, *with* $\mathcal{A}^{\text{adj}}(Z) + \mathcal{B}^{\text{adj}}(w) = 0$.

Theorem 6 (ALT 5b) *At most one of the following statements is true.*

1. *There exists an* $x \in \mathcal{V}$ *with* $\mathcal{A}(x) + A_0 \gneqq 0$ *and* $\mathcal{B}(x) = 0$.

2. *There exists a* $Z \in \mathcal{S}^n$ *with* $Z > 0$, *and* $w \in \mathcal{W}$, *with* $\mathcal{A}^{\mathrm{adj}}(Z) + \mathcal{B}^{\mathrm{adj}}(w) = 0$, *and* $\langle A_0, Z \rangle_{\mathcal{S}^n} \leq 0$.

Moreover if there exists no $x \in \mathcal{V}$ *with* $\mathcal{A}(x) \gneqq 0$ *and* $\mathcal{B}(x) = 0$, *then exactly one of the two statements is true.*

Theorem 7 (ALT 6) *At most one of the following statements is true.*

1. *There exists an* $x \in \mathcal{V}$ *with* $\mathcal{A}(x) + A_0 \geq 0$ *and* $\mathcal{B}(x) = 0$.

2. *There exists a* $Z \in \mathcal{S}^n$ *with* $Z \geq 0$, *and* $w \in \mathcal{W}$, *with* $\mathcal{A}^{\mathrm{adj}}(Z) + \mathcal{B}^{\mathrm{adj}}(w) = 0$, *and* $\langle A_0, Z \rangle_{\mathcal{S}^n} < 0$.

Moreover if there exists no $x \in \mathcal{V}$ *with* $\mathcal{A}(x) \gneqq 0$ *and* $\mathcal{B}(x) = 0$, *then exactly one of the two statements is true.*

2.2 Semidefinite programming duality

A *semidefinite programming* problem (SDP) requires minimizing a linear function subject to an LMI constraint:

$$\begin{array}{ll} \text{minimize} & \langle c, x \rangle_{\mathcal{V}} \\ \text{subject to} & \mathcal{A}(x) + A_0 \geq 0 \end{array} \tag{6}$$

From convex duality, we can associate with the SDP the *dual* problem

$$\begin{array}{ll} \text{maximize} & -\langle A_0, Z \rangle_{\mathcal{S}} \\ \text{subject to} & \mathcal{A}^{\mathrm{adj}}(Z) = c, \quad Z \geq 0 \end{array} \tag{7}$$

where the variable is the matrix $Z \in \mathcal{S}$. In the context of duality we refer to the SDP (6) as the *primal problem* associated with (7).

The following theorem relates the optimal values of the primal and dual SDPs. Let p_{opt} be the optimal value of (6) and d_{opt} the optimal value of (7). We allow values $\pm\infty$: $p_{\mathrm{opt}} = +\infty$ if the primal problem is infeasible and $p_{\mathrm{opt}} = -\infty$ if it is unbounded below; $d_{\mathrm{opt}} = +\infty$ if the dual problem is unbounded above, $d_{\mathrm{opt}} = -\infty$ if it is infeasible.

Theorem 8 $p_{\mathrm{opt}} \geq d_{\mathrm{opt}}$. *If the primal problem is strictly feasible, (i.e., there exists x with* $\mathcal{A}(x) + A_0 > 0$*), or the dual problem is strictly feasible (i.e., there exists* $Z > 0$ *with* $\mathcal{A}^{\mathrm{adj}}(Z) = c$*), then* $p_{\mathrm{opt}} = d_{\mathrm{opt}}$.

The first property ($p_{\mathrm{opt}} \geq d_{\mathrm{opt}}$) is called *weak duality*. If $p_{\mathrm{opt}} = d_{\mathrm{opt}}$, we say the primal and dual SDPs satisfy *strong duality*. A proof of Theorem 8 is given in Appendix B.

Theorem 8 is the standard Lagrange duality result for semidefinite programming. An alternative duality theory, which does not require a constraint qualification, was developed by Ramana, Tunçel, and Wolkowicz [RTW97].

2.3 Optimality conditions

Suppose strong duality holds. The following facts are useful when studying the properties of the optimal solutions of the primal and dual SDP.

- A primal feasible x and a dual feasible Z are optimal if and only if $(\mathcal{A}(x)+A_0)Z=0$. This property is called *complementary slackness*.
- If the primal problem is strictly feasible, then the dual optimum is attained, i.e., there exists a dual optimal Z.
- If the dual problem is strictly feasible, then the primal optimum is attained, i.e., there exists a primal optimal x.

A proof of this result is given in Appendix C

We combine these properties to state necessary and sufficient conditions for optimality. For example, it follows that if the primal problem is strictly feasible (hence strong duality obtains), then a primal feasible x is optimal if and only if there exists a dual feasible Z with $(\mathcal{A}(x)+A_0)Z=0$.

Note that complementary slackness between optimal solutions is only satisfied when strong duality holds. Consider the following example from [VB96b], where $\mathcal{V}=\mathbf{R}^2$ and $\mathcal{S}=\mathcal{S}^3$, and the primal SDP is

$$\begin{array}{ll}\text{minimize} & x_1\\ \text{subject to} & \begin{bmatrix} 0 & x_1 & 0\\ x_1 & x_2 & 0\\ 0 & 0 & x_1+1\end{bmatrix}\succeq 0.\end{array}$$

This problem is not strictly feasible. Its optimal value is $p_{\rm opt}=0$, and any $x\in\mathbf{R}^2$ with $x_1=0$, $x_2\geq 0$ is optimal. The dual SDP is

$$\begin{array}{ll}\text{maximize} & -z_{33}\\ \text{subject to} & z_{22}=0,\quad z_{12}+z_{12}^*+z_{33}=1,\quad \begin{bmatrix} z_{11} & z_{12} & z_{13}\\ z_{12}^* & z_{22} & z_{23}\\ z_{13}^* & z_{23}^* & z_{33}\end{bmatrix}\succeq 0.\end{array}$$

This problem is not strictly feasible, because of the constraint $z_{22}=0$. This constraint also implies that $z_{23}=z_{12}=0$, and hence, $z_{33}=1$. All feasible Z therefore have the form

$$Z=\begin{bmatrix} z_{11} & 0 & z_{13}\\ 0 & 0 & 0\\ z_{13}^* & 0 & 1\end{bmatrix}$$

with $z_{11}\geq 0$ and $z_{11}\geq |z_{13}|^2$. The optimal value is $d_{\rm opt}=-1$. Comparing the two optimal solutions $x_1=x_2=0$ and $z_{11}=z_{13}=0$, we note that complementary slackness is not satisfied.

2.4 Some useful preliminaries

We will encounter four specific linear mappings several times in the sequel. For easy reference, we define these here, and derive the expression for their adjoints.

Example 5 Let $\mathcal{A}_1 : \mathcal{S}^n \to \mathcal{S}^n$ be defined by $\mathcal{A}_1(P) = -(A^*P + PA)$. Then, it is easily verified that $\mathcal{A}_1^{\text{adj}} : \mathcal{S}^n \to \mathcal{S}^n$ is given by $\mathcal{A}_1^{\text{adj}}(Z) = -(ZA^* + AZ)$.

Example 6 Let $\mathcal{A}_2 : \mathcal{S}^n \to \mathcal{S}^n \times \mathcal{S}^n$ be defined by $\mathcal{A}(P) = \mathbf{diag}(-(A^*P + PA), P)$, Then, it is easily verified that $\mathcal{A}_2^{\text{adj}} : \mathcal{S}^n \times \mathcal{S}^n \to \mathcal{S}^n$ is given by $\mathcal{A}_2^{\text{adj}}(Z) = -(Z_1A^* + AZ_1 - Z_2)$, where $Z = \mathbf{diag}(Z_1, Z_2)$.

Example 7 Let $\mathcal{A}_3 : \mathcal{S}^n \to \mathcal{S}^{n+m}$ be defined by

$$\mathcal{A}_3(P) = -\begin{bmatrix} A^*P + PA & PB \\ B^*P & 0 \end{bmatrix}.$$

Then, it is easily verified that $\mathcal{A}_3^{\text{adj}} : \mathcal{S}^{n+m} \to \mathcal{S}^n$ is given by

$$\mathcal{A}_3^{\text{adj}}(\begin{bmatrix} Z_{11} & Z_{12} \\ Z_{12}^* & Z_{22} \end{bmatrix}) = -Z_{11}A^* - AZ_{11} - BZ_{12}^* - Z_{12}B^*.$$

Example 8 Let $\mathcal{A}_4 : \mathcal{S}^n \to \mathcal{S}^{n+m} \times \mathcal{S}^n$ be defined by

$$\mathcal{A}_4(P) = \mathbf{diag}(-\begin{bmatrix} A^*P + PA & PB \\ B^*P & 0 \end{bmatrix}, P).$$

Then, it is easily verified that $\mathcal{A}_4^{\text{adj}} : \mathcal{S}^{n+m} \times \mathcal{S}^n \to \mathcal{V}$ is given by

$$\mathcal{A}_4^{\text{adj}}(\mathbf{diag}(\begin{bmatrix} Z_{11} & Z_{12} \\ Z_{12}^* & Z_{22} \end{bmatrix}, Z_2)) = -Z_{11}A - A^*Z_{11} - BZ_{12}^* - Z_{12}B^* + Z_2.$$

3 Lyapunov inequalities, stability, and controllability

As our first application of the theorem of alternatives to the analysis of linear time-invariant (LTI) systems, we consider the LTI system

$$\dot{x} = Ax, \tag{8}$$

where $A \in \mathbf{C}^{n\times n}$. Lyapunov equations, i.e., equations of the form $A^*P + PA + Q = 0$, and Lyapunov inequalities, i.e., LMIs of the form $A^*P + PA < 0$ or $A^*P + PA \leq 0$ play a fundamental role in establishing the stability of system (8); see any text on linear systems, for instance, [Rug96].

We consider some well known results on Lyapunov inequalities. Although these results are readily proved using standard techniques, we give a proof using SDP duality to illustrate the techniques that will be used later in the paper.

3.1 Strict Lyapunov inequalities

Proposition 1 *Exactly one of the following two statements is true.*

1. *There exists a $P \in \mathcal{S}^n$ such that $A^*P + PA < 0$.*

2. *A has an imaginary eigenvalue.*

Proof. With $\mathcal{A}_1$ as in Example 5 and with $A_0 = 0$, the first statement of the theorem is equivalent to the existence of $P \in \mathcal{S}^n$ such that $\mathcal{A}_1(P) + A_0 > 0$. Then, applying Theorem **ALT 1**, the alternative is that there exists a $Z \in \mathcal{S}^n$ with

$$Z \gneq 0, \quad AZ + ZA^* = 0. \tag{9}$$

We now show that this condition is equivalent to A having imaginary eigenvalues, establishing the proposition.

Suppose A has an imaginary eigenvalue, i.e., there exist nonzero $v \in \mathbf{C}^n$, and $\omega \in \mathbf{R}$ with $Av = j\omega v$. It is easily shown that $Z = vv^*$ satisfies (9).

Conversely, suppose that (9) holds. Let $Z = UU^*$ where $U \in \mathbf{C}^{n\times r}$ and $\mathbf{Rank}\, U = \mathbf{Rank}\, Z = r$. From (9), we note that AZ is skew-Hermitian, so that we must have $AUU^* = USU^*$ where S is skew-Hermitian. Therefore $AU = US$. The eigenvalues of S are all on the imaginary axis because S is skew-Hermitian. Therefore, the columns of U span an invariant subspace of A associated with a set of imaginary eigenvalues. Thus A has at least one imaginary eigenvalue. □

Remark 2 In Proposition 1, it is easy to show directly that both statements cannot hold; this is the "easy" part. For instance, if A has an imaginary eigenvalue, i.e., if $Av = j\omega v$ for some $\omega \in \mathbf{R}$ and nonzero $v \in \mathbf{C}^n$, it is easy to show that $A^*P + PA < 0$ cannot hold for any $P \in \mathcal{S}^n$. (In the proof, we prove this "easy" implication with the second alternative.) The hard part is the converse, and the theorems of alternatives give a "constructive" proof: We exhibit the eigenspace of A corresponding to one or more imaginary eigenvalues. It is also worthy of note that (numerical) convex optimization algorithms operate similarly: Given a convex feasibility problem, they either find a feasible point, or provide a constructive proof of infeasibility.

Proposition 1 is representative of most of the results in the sequel, with an easy part and a hard part, with the theorems of alternatives providing a constructive proof of the hard part. ◇

Proposition 2 *Exactly one of the following two statements is true.*

1. *There exists a $P \in \mathcal{S}^n$ such that $P > 0$ and $A^*P + PA < 0$.*

2. *A has an eigenvalue with non-negative real part.*

Remark 3 This is a restatement of the celebrated Lyapunov stability theorem for LTI systems. ◇

Proof. With $\mathcal{A}_2$ as in Example 6 and with $A_0 = 0$, the first statement of the theorem is equivalent to the existence of $P \in \mathcal{S}^n$ such that $\mathcal{A}_2(P) + A_0 > 0$. Then, applying Theorem **ALT 1**, the alternative is that there exist $Z_1 \in \mathcal{S}^n$ and $Z_2) \in \mathcal{S}^n$ with

$$\mathbf{diag}(Z_1, Z_2) \gneq 0, \quad Z_1 A^* + AZ_1 - Z_2 = 0. \tag{10}$$

We now show that this condition is equivalent to A having eigenvalues with non-negative real part, establishing the proposition.

Suppose that A has an eigenvalue with non-negative real part, i.e., there exist nonzero $v \in \mathbf{C}^n$, $\sigma \geq 0$ and $\omega \in \mathbf{R}$ with $Av = (\sigma + j\omega)v$. It is easily shown that $Z_1 = vv^*$, $Z_2 = 2\sigma vv^*$ satisfy (10).

Conversely, suppose that (10) holds. We can write $Z_1 = UU^*$ with $U \in \mathbf{C}^{n\times r}$ and $\mathbf{Rank}\, U = \mathbf{Rank}\, Z = r$. From (10), we note that the symmetric part of AZ_1 is positive semidefinite, so that we must have $AUU^* = USU^*$ where S is the sum of a skew-Hermitian and a positive semidefinite matrix. Then, $AU = US$. The eigenvalues of S are all in the closed right-half plane because S is the sum of a skew-Hermitian and a positive semidefinite matrix. Therefore U spans a (nonempty) invariant subspace of A associated with a set eigenvalues of A with non-negative real part. □

Remark 4 Theorem **ALT 1**, besides offering a simple proof to Lyapunov's theorem, also enables the extension of Proposition 2 to more general settings. Consider the problem of the existence of P satisfying

$$P > 0, \quad A_1^* P + PA_1 < 0, \quad A_2^* P + PA_2 < 0. \tag{11}$$

The matrix P can be interpreted as defining a common or simultaneous quadratic Lyapunov function [BY89, BEFB94, SN98, SN99, SN00] that proves the stability of the time-varying system

$$\dot{x} = A(t)x, \quad A(t) = \lambda(t)A_1 + (1 - \lambda(t))A_2, \quad \lambda(t) \in [0, 1] \text{ for all } t.$$

An application of Theorem **ALT 1** immediately yields a necessary and sufficient condition for (11) to be feasible: There do not exist $Z_1, Z_2 \in \mathcal{S}^n$ such that

$$\mathbf{diag}(Z_1, Z_2) \gneq 0, \quad Z_1 A_1^* + A_1 Z_1 + Z_2 A_2^* + A_2 Z_2 \geq 0. \tag{12}$$

It is easy to show that if $A_1 + \sigma A_2$ has a nonnegative eigenvalue for some $\sigma \in \mathbf{C}$, then (12) is feasible, or there does not exist P satisfying (11). References [SN98, SN99, SN00] explore sufficient conditions, using algebraic techniques, for the existence of P satisfying (11) for the special case when the matrices A_i are 2×2 and real.

3.2 Nonstrict Lyapunov inequalities

We saw in §3.1 that the alternatives to strict Lyapunov inequalities involving a matrix A are equivalent to a condition on *some* eigenvalue of A. We will see in this section that the alternatives to nonstrict Lyapunov inequalities result in conditions that are to be satisfied by *all* eigenvalues of A.

Proposition 3 *Exactly one of the following two statements is true.*

1. *There exists $P \in \mathcal{S}^n$ such that $A^*P + PA \lneqq 0$.*

2. *A is similar to a purely imaginary diagonal matrix.*

Proof. With $\mathcal{A}_1$ as in Example 5 and with $A_0 = 0$, the first statement of the theorem is equivalent to the existence of $P \in \mathcal{S}^n$ such that $\mathcal{A}_1(P) + A_0 \gneqq 0$. Then, applying Theorem **ALT 2**, the alternative is that there exists a $Z \in \mathcal{S}^n$ with $Z > 0, \quad AZ + ZA^* = 0$. We now show that this condition is equivalent to A being similar to a purely imaginary diagonal matrix.

Suppose A is similar to an imaginary diagonal matrix, i.e., there exists V such that $A = V\Lambda V^{-1}$ with Λ diagonal and imaginary. Then $Z = VV^* > 0$ and $AZ + ZA^* = 0$.

Conversely, suppose that there exists $Z > 0$ with $AZ + ZA^* = 0$, i.e., $AZ = S$ where S is skew-Hermitian. Therefore $A = SZ^{-1}$, which has the same eigenvalues as $Z^{-1/2}SZ^{-1/2}$, i.e., A has n non-defective imaginary eigenvalues. In fact, a similarity transformation that maps A to an imaginary diagonal matrix is easily constructed from Z. Let a Schur decomposition of the matrix $Z^{-1/2}AZ^{1/2}$ be given by $Z^{-1/2}AZ^{1/2} = WTW^*$, where $W^*W = WW^* = I$ and T is upper triangular. From $AZ + ZA^* = 0$, we have $W^*Z^{-1/2}(AZ + Z^*A^*)Z^{-1/2}W = T + T^* = 0$. Therefore T must be diagonal, with purely imaginary diagonal elements. In other words, if we define $V = Z^{1/2}W$, then the matrix $V^{-1}AV = T$ is a purely imaginary diagonal matrix. □

Proposition 4 *Exactly one of the following two statements is true.*

1. *There exists $P \in \mathcal{S}^n$ such that $A^*P + PA \leq 0, \quad P \gneqq 0$*

2. *The eigenvalues of A are in the open right half plane.*

Proof. With $\mathcal{A}_2$ as in Example 6 and with $A_0 = 0$, the first statement of the theorem is equivalent to the existence of $P \in \mathcal{S}^n$ such that $\mathcal{A}_2(P) + A_0 \gneqq 0$. Then, applying Theorem **ALT 2**, the alternative is that there exists a $Z \in \mathcal{S}^n$ with $Z > 0, \quad AZ + ZA^* > 0$. From Proposition 2 this is true if and only if A has no eigenvalue with non-positive real part, i.e., if all eigenvalues of A are in the open right half plane. □

3.3 Generalized Lyapunov inequalities

Propositions 1–4 deal with the issue of whether the eigenvalues of A lie in or on the boundary of the left-half complex plane. Standard techniques can be used to extend these results to handle more general regions in the complex plane; an indirect route is through conformal mapping techniques from complex analysis (see for instance, [Con78]). For example, the mapping $A \mapsto (I+A)(I-A)^{-1}$ can be used to derive theorems of alternatives that address whether the eigenvalues of A lie in or on the boundary of the unit disk in the complex plane; the underlying control-theoretic interpretation then concerns the stability of discrete-time linear systems.

We now demonstrate how Proposition 2 can be directly extended to handle generalized complex half-planes and generalized circles.

Proposition 5 *For some* $\theta \in [0, 2\pi)$ *and* $\beta \in \mathbf{R}$, *consider the complex half-plane*

$$\mathcal{H}_{\theta,\beta} = \left\{ s \in \mathbf{C} \mid \Re\left(e^{j\theta} s + \beta\right) < 0 \right\}.$$

Then, exactly one of the following two statements is true.

1. *There exists a* $P \in \mathcal{S}^n$ *such that*

$$P > 0, \quad e^{-j\theta} A^* P + P A e^{j\theta} + 2\beta P < 0. \tag{13}$$

2. *A has an eigenvalue that does not lie in* $\mathcal{H}_{\theta,\beta}$.

Proof. It is readily verified that A has an eigenvalue that does not lie in $\mathcal{H}_{\theta,\beta}$ if and only if $\tilde{A} = Ae^{j\theta} + \beta I$ has an eigenvalue with nonnegative real part. Using Proposition 2, we have that $\tilde{A}$ does not have an eigenvalue with nonnegative real part if and only if there exists P satisfying

$$P > 0, \quad \tilde{A}^* P + P\tilde{A} < 0,$$

which is equivalent to (13). □

Proposition 6 *For some* $\rho > 0$ *and* $s_0 \in \mathbf{C}$, *consider the disk*

$$\mathcal{C}_{\rho,s_0} = \left\{ s \in \mathbf{C} \mid |s - s_0|^2 < \rho^2 \right\}.$$

Then, exactly one of the following two statements is true.

1. *There exists a* $P \in \mathcal{S}^n$ *such that*

$$P > 0, \quad A^* P A - s_0 A^* P - s_0^* P A - (\rho^2 - |s_0|^2) P < 0. \tag{14}$$

2. *A has an eigenvalue that does not lie in* $\mathcal{C}_{\rho,s_0}$, *that is, A has an eigenvalue* λ *that satisfies* $|s - s_0|^2 \geq \rho^2$.

Proof. Let $\theta \in [0, 2\pi)$ be such that $(A - s_0 I)e^{j\theta} + \rho I$ is nonsingular. Then, it is readily verified that A has an eigenvalue λ that satisfies $|s - s_0|^2 \geq \rho^2$ if and only if

$$\tilde{A} = \left((A - s_0 I)e^{j\theta} + \rho I\right)^{-1} \left((A - s_0 I)e^{j\theta} - \rho I\right)$$

has an eigenvalue with nonnegative real part. Using Proposition 2, we have that $\tilde{A}$ does not have an eigenvalue with nonnegative real part if and only if there exists P satisfying

$$P > 0, \quad \tilde{A}^* P + P\tilde{A} < 0,$$

which, after routine manipulations, yields (14). □

Remark 5 It is straightforward to derive other results similar to the above two, using Propositions 1, 3, and 4. $\diamond$

Remark 6 Consider the subset of the complex plane

$$\mathcal{D} = \left\{ s \middle| \begin{bmatrix} 1 \\ s \end{bmatrix}^* \begin{bmatrix} a & b \\ b^* & c \end{bmatrix} \begin{bmatrix} 1 \\ s \end{bmatrix} < 0 \right\},$$

where $b \in \mathbf{C}$ and $a, c \in \mathbf{R}$ with $c \geq 0$. LMI conditions that are necessary and sufficient for the eigenvalues of a given $A \in \mathcal{S}^n$ to lie in $\mathcal{D}$ are given in [HM01].

Note that when $c = 0$, $\mathcal{D}$ equals $\mathcal{H}_{\theta,\beta}$ with $a = 2\beta$ and $b = e^{j\theta}$. When $c > 0$, $\mathcal{D}$ equals $\mathcal{C}_{\rho,s_0}$ with $s_0 = \overline{b}/c$ and $\rho = \sqrt{|b|^2 - ac}/|c|$. Thus, the development in this section serve to provide an alternate proof and to extend the results in [HM01]. $\diamond$

3.4 Lyapunov inequalities with equality constraints

We next consider an LTI system with an input:

$$\dot{x} = Ax + Bu, \tag{15}$$

where $A \in \mathbf{C}^{n\times n}$ and $B \in \mathbf{C}^{n\times m}$. The pair (A,B) is said to be *controllable* if for every initial condition $x(0)$, there exists an input u and T such that $x(T) = 0$. While, there are several equivalent characterizations and conditions for controllability of (A,B) (see for example [Rug96]), we will use the following: The pair (A,B) is not controllable if and only if there exists a left eigenvector v^* of A such that $v^*B = 0$.

If (A,B) is controllable, then given any monic polynomial $a : \mathbf{C} \to \mathbf{C}$ of degree n with complex coefficients, there exists $K \in \mathbf{C}^{m\times n}$ such that $\det(sI - A - BK) = a(s)$ for all $s \in \mathbf{C}$. In other words, with "state-feedback" $u = Kx$ in (15), the eigenvalues of $A + BK$ can be arbitrarily assigned. When (A,B) is not controllable, there exists a nonsingular matrix $T \in \mathbf{C}^{n\times n}$ such that

$$T^{-1}AT = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix}, \quad T^{-1}B = \begin{bmatrix} B_1 \\ 0 \end{bmatrix}, \tag{16}$$

where $A_{11} \in \mathbf{C}^{r\times r}$ and $B_1 \in \mathbf{C}^{r\times m}$, with $r < n$ and (A_{11}, B_1) being controllable. (This is called the "Kalman form".) The eigenvalues of A_{22} are called the uncontrollable modes. An uncontrollable mode is called nondefective if its algebraic multiplicity as an eigenvalue of A_{22} equals its geometric multiplicity. The matrix T in (16) has the interpretation of a state coordinate transformation $\overline{x} = T^{-1}x$ such that in the new coordinates, only the first r components of the state are controllable.

Proposition 7 *Exactly one of the following two statements is true.*

1. *There exists $P \in \mathcal{S}^n$ satisfying $A^*P + PA \lneqq 0, \quad PB = 0$.*

2. *All uncontrollable modes of (A,B) are nondefective and correspond to imaginary eigenvalues.*

Proof. With $\mathcal{A}_3$ as in Example 7 and with $A_0 = 0$, the first statement of the theorem is equivalent to the existence of $P \in \mathcal{S}^n$ such that $\mathcal{A}_3(P) + A_0 \gneqq 0$. Then, applying Theorem **ALT 2**, the alternative is that there exists $Z \in \mathcal{S}^{n+m}$ such that

$$Z = \begin{bmatrix} Z_{11} & Z_{12} \\ Z_{12}^* & Z_{22} \end{bmatrix} > 0, \quad AZ_{11} + Z_{11}A^* + BZ_{12}^* + Z_{12}B^* = 0.$$

Defining $K = Z_{12}^* Z_{11}^{-1}$, we can write this equivalently as

$$Z_{11} > 0, \quad (A+BK)Z_{11} + Z_{11}(A+BK)^* = 0. \tag{17}$$

In other words, the first statement of the Proposition is false if and only if there exist $K \in \mathbf{R}^{n\times m}$ and $Z_{11} \in \mathcal{S}^n$ that satisfy (17). We now establish that this condition is equivalent to the second statement. We will assume, without loss of generality, that (A,B) is in Kalman form, and that K and Z_{11} are appropriately partitioned as

$$K = \begin{bmatrix} K_1 & K_2 \end{bmatrix}, \quad Z_{11} = \begin{bmatrix} \tilde{Z}_{11} & \tilde{Z}_{12} \\ \tilde{Z}_{12}^* & \tilde{Z}_{22} \end{bmatrix}. \tag{18}$$

Suppose that the uncontrollable modes of (A,B) (if any) are nondefective and correspond to imaginary eigenvalues. We will establish that we can find $Z_{11} > 0$ and K satisfying (17). By assumption A_{22} is similar to a purely imaginary diagonal matrix. The pair (A_{11}, B_1) is controllable, so there exists K_1 such that the eigenvalues of $A_{11} + B_1K_1$ are distinct, purely imaginary, and different from the eigenvalues of A_{22}. Therefore there exist V_{11} and V_{22} such that

$$V_{11}(A_{11} + B_1K_1)V_{11}^{-1} = \Lambda_1, \quad V_{22}A_{22}V_{22}^{-1} = \Lambda_2$$

where Λ_1 and Λ_2 are diagonal and purely imaginary. The spectra of Λ_1 and A_{22} are disjoint, so the Sylvester equation $-\Lambda_1V_{12} + V_{12}A_{22} = -V_{11}A_{12}$ has a unique solution V_{12} (see [HJ91, Th. 4.4.5]). If we take $K_2 = 0$, it is easily verified that $V = \begin{bmatrix} V_{11} & V_{12} \\ 0 & V_{22} \end{bmatrix}$ satisfies

$$V(A+BK)V^{-1} = \begin{bmatrix} V_{11} & V_{12} \\ 0 & V_{22} \end{bmatrix} \begin{bmatrix} A_{11} + B_1K_1 & A_{12} + B_1K_2 \\ 0 & A_{22} \end{bmatrix} \begin{bmatrix} V_{11} & V_{12} \\ 0 & V_{22} \end{bmatrix}^{-1} = \begin{bmatrix} \Lambda_1 & 0 \\ 0 & \Lambda_2 \end{bmatrix},$$

i.e., $A+BK$ is similar to a purely imaginary diagonal matrix. We can now proceed as in the proof of Proposition 3 and show that the matrix $Z_{11} = VV^*$ satisfies (17).

Conversely, suppose that Z_{11} and K satisfy (17). In particular, $\tilde{Z}_{22} > 0$, and $A_{22}\tilde{Z}_{22} + \tilde{Z}_{22}A_{22}^* = 0$. As in the proof of Proposition 3 we can construct from $\tilde{Z}_{22}$ a similarity transformation that makes A_{22} diagonal with purely imaginary diagonal elements. Hence all the uncontrollable modes are nondefective and correspond to imaginary eigenvalues. □

Proposition 8 *Exactly one of the following two statements is true.*

1. *There exists $P \in \mathcal{S}^n$ satisfying*

$$P \gneqq 0, \quad A^*P + PA \leq 0, \quad PB = 0. \tag{19}$$

2. *All uncontrollable modes of (A,B) correspond to eigenvalues with positive real part.*

Proof. With $\mathcal{A}_4$ as in Example 8 and with $A_0 = 0$, the first statement of the theorem is equivalent to the existence of $P \in \mathcal{S}^n$ such that $\mathcal{A}_4(P) + A_0 \gneqq 0$. Then, applying Theorem **ALT 2**, the alternative is that there exists $Z_{11} \in \mathcal{S}^n$, $Z_{12} \in \mathbf{C}^{n\times m}$, $Z_{22} \in \mathcal{S}^m$, and $Z_2 \in \mathcal{S}^n$ with

$$\left[\begin{array}{cc} Z_{11} & Z_{12} \\ Z_{12}^* & Z_{22} \end{array}\right] > 0, \quad Z_2 > 0, \quad Z_{11}A^* + AZ_{11} + BZ_{12}^* + Z_{12}B^* = Z_2.$$

Defining $K = Z_{12}^* Z_{11}^{-1}$ this is equivalent to the existence of Z_{11} and K such that

$$Z_{11} > 0, \quad Z_{11}(A+BK)^* + (A+BK)Z_{11} > 0. \tag{20}$$

We now show that this is equivalent to the second statement in the Proposition. We will assume, without loss of generality, that (A,B) is in Kalman form, and that K and Z_{11} are appropriately partitioned as in (18).

First suppose that the uncontrollable modes of (A,B) (if any) correspond to eigenvalues of A with positive real part, i.e., the eigenvalues of A_{22} are in the open right half plane. An argument similar to the one in the proof of Proposition 7 can be given (in turn, using arguments from the proof of Proposition 4) to construct Z_{11} and K such that (20) holds.

Conversely, suppose that (20) holds. In particular, $\tilde{Z}_{22} > 0$, and $\tilde{Z}_{22}A_{22}^* + A_{22}\tilde{Z}_{22} > 0$. By Proposition 2 this implies that the eigenvalues A_{22} have a positive real part. □

Finally, we present a condition for controllability. We first note the following result, which can be interpreted as a theorem of alternatives for linear equations.

Proposition 9 *Exactly one of the following two statements is true.*

1. *There exists $P \in \mathcal{S}^n$ satisfying*

$$P \neq 0, \quad A^*P + PA = 0, \quad PB = 0 \tag{21}$$

2. *With $\lambda_1, \cdots, \lambda_p$ denoting the uncontrollable modes of (A,B), $\lambda_i + \lambda_j^* \neq 0$, $1 \leq i, j \leq p$.*

Proof. Without loss of generality we can assume that (A,B) is in the Kalman form (16), with $A_{22} \in \mathbf{C}^{p\times p}$. We partition P accordingly as

$$P = \begin{bmatrix} P_{11} & P_{12} \\ P_{12}^* & P_{22} \end{bmatrix}.$$

First suppose $\lambda_i + \lambda_j^* = 0$ for two eigenvalues λ_i and λ_j of A_{22}. Then the Lyapunov equation $A_{22}^* P_{22} + P_{22} A_{22} = 0$ has a nonzero solution P_{22} (see [HJ91, Th. 4.4.5]). Taking $P_{11} = 0$ and $P_{12} = 0$, we obtain a nonzero P that satisfies $A^*P + PA = 0$, $PB = 0$.

Conversely, if P satisfies (21), then $(A+BK)P + P(A+BK)^* = 0$ for all K. This is only possible if for all K,

$$A + BK = \begin{bmatrix} A_{11} + B_1 K_1 & A_{22} + B_1 K_2 \\ 0 & A_{22} \end{bmatrix}$$

has eigenvalues μ_i and μ_j that satisfy $\mu_i + \mu_j^* = 0$ (again, see [HJ91, Th. 4.4.5]). The spectrum of $A+BK$ is the union of the spectrum of $A_{11} + B_1 K_1$ and the spectrum of A_{22}. Therefore we must have $\lambda_i + \lambda_j^* = 0$ for two eigenvalues of A_{22}. □

Proposition 10 *Exactly one of the following two statements is true.*

1. *There exists $P \in \mathcal{S}^n$ satisfying $P \neq 0, \quad A^*P + PA \leq 0, \quad PB = 0$.*

2. *The pair (A,B) is controllable.*

Proof. Statement 1 is true if the statements 1a or 1b listed below are true.

1a. There exists $P \in \mathcal{S}^n$ satisfying $A^*P + PA \lneq 0$, $PB = 0$.

1b. There exists $P \in \mathcal{S}^n$ satisfying $P \neq 0$, $A^*P + PA = 0$, $PB = 0$.

By Propositions 9 and 7 the alternatives to these statements are the following:

2a. All uncontrollable modes are nondefective, and correspond to eigenvalues on the imaginary axis.

2b. With $\lambda_1, \cdots \lambda_p$ denoting the uncontrollable modes of (A,B), $\lambda_i + \lambda_j^* \neq 0$, $1 \leq i, j \leq p$.

The alternative to 1 is therefore that 2a *and* 2b are true, i.e., that there are no uncontrollable modes. □

Remark 7 Alternative proofs of this result appeared in [GND99] and [VB99, Lemma 1]. ◇

4 Riccati inequalities

We next consider convex Riccati inequalities, which take the form

$$\begin{bmatrix} A^*P+PA & PB \\ B^*P & 0 \end{bmatrix} - M \le 0, \tag{22}$$

with $A \in \mathbf{C}^{n\times n}$, $B \in \mathbf{C}^{n\times m}$. Let M be partitioned as

$$M = \begin{bmatrix} M_{11} & M_{12} \\ M_{12}^* & M_{22} \end{bmatrix},$$

where $M_{11} = M_{11}^* \in \mathcal{S}^n$. Then, when $M_{22} > 0$, inequality (22) is equivalent to

$$A^*P + PA - M_{11} + (PB - M_{12})M_{22}^{-1}(B^*P - M_{12}^*)^{-1} \le 0.$$

Such inequalities are widely encountered in quadratic optimal control, estimation theory, and H_∞ control; see for example [Wil71, LR91, BLW91].

4.1 Strict Riccati inequalities

Proposition 11 *Suppose $M_{22} > 0$. Then exactly one of the following two statements is true.*

1. *There exists $P \in \mathcal{S}^n$ such that*

$$\begin{bmatrix} A^*P+PA & PB \\ B^*P & 0 \end{bmatrix} - M < 0. \tag{23}$$

2. *For some full-rank $U \in \mathbf{C}^{n\times r}$, $V \in \mathbf{C}^{m\times r}$, and $S \in \mathbf{C}^{r\times r}$ with $S + S^* = 0$,*

$$US - AU = BV, \quad \mathbf{Tr}\left(\begin{bmatrix} U^* & V^* \end{bmatrix} M \begin{bmatrix} U \\ V \end{bmatrix} \right) \le 0.$$

Proof. With $\mathcal{A}_3$ as in Example 7 and with $A_0 = M$, the first statement of the theorem is equivalent to the existence of $P \in \mathcal{S}^n$ such that $\mathcal{A}_3(P) + A_0 > 0$. Then, applying Theorem **ALT 1**, the alternative is that there exists a $Z \in \mathcal{S}^{n+m}$ with

$$Z = \begin{bmatrix} Z_{11} & Z_{12} \\ Z_{12}^* & Z_{22} \end{bmatrix} \gneq 0, \quad Z_{11}A^* + AZ_{11} + Z_{12}B^* + BZ_{12}^* = 0, \quad \mathbf{Tr}\, ZM \le 0 \tag{24}$$

We now show that this condition is equivalent to the existence of $U \in \mathbf{C}^{n\times r}$, $V \in \mathbf{C}^{m\times r}$, and $S \in \mathbf{C}^{r\times r}$ with $S + S^* = 0$ such that

$$US - AU = BV, \quad \mathbf{Tr}\left(\begin{bmatrix} U^* & V^* \end{bmatrix} M \begin{bmatrix} U \\ V \end{bmatrix} \right) \le 0. \tag{25}$$

We must have $Z_{11} \gneqq 0$, as otherwise we would have $Z_{12} = 0$, and the last inequality in (24) would imply that $Z_{22} = 0$, and consequently $Z = 0$, a contradiction. Therefore, there exist $U \in \mathbf{C}^{n\times r}$ and $V \in \mathbf{C}^{m\times r}$, where $r = \mathbf{Rank}\, Z_{11} \geq 1$. such that

$$\begin{bmatrix} Z_{11} & Z_{12} \\ Z_{12}^* & Z_{22} \end{bmatrix} = \begin{bmatrix} U & 0 \\ V & \hat{V} \end{bmatrix} \begin{bmatrix} U^* & V^* \\ 0 & \hat{V}^* \end{bmatrix}$$

where U has full rank. The equation $Z_{11}A^* + AZ_{11} + Z_{12}B^* + BZ_{12}^* = 0$, represented in terms of U and V means that $AUU^* + BVU^*$ is skew-Hermitian, i.e., it can be written as $AUU^* + BVU^* = USU^*$, where S is skew-Hermitian. Since U has full rank, this last equation implies $AU + BV = US$. Expressing inequality $\mathbf{Tr}\, ZM \leq 0$ in terms of U and V, we obtain

$$\mathbf{Tr}\left(\begin{bmatrix} U^* & V^* \\ 0 & \hat{V}^* \end{bmatrix} M \begin{bmatrix} U & 0 \\ V & \hat{V} \end{bmatrix} \right) \leq 0,$$

which, since $M_{22} > 0$, implies that

$$\mathbf{Tr}\left(\begin{bmatrix} U^* & V^* \end{bmatrix} M \begin{bmatrix} U \\ V \end{bmatrix} \right) \leq 0,$$

completing the proof. □

The conclusion of Proposition 11 can be further developed to yield the Kalman-Yakubovich-Popov Lemma.

Lemma 1 (KYP Lemma) *Suppose $M_{22} > 0$. There exists $P \in \mathcal{S}^n$ such that*

$$\begin{bmatrix} A^*P + PA & PB \\ B^*P & 0 \end{bmatrix} - M < 0, \tag{26}$$

if and only for all $\omega \in \mathbf{R}$,

$$(j\omega I - A)u = Bv,\ (u,v) \neq 0 \Longrightarrow \begin{bmatrix} u^* & v^* \end{bmatrix} M \begin{bmatrix} u \\ v \end{bmatrix} > 0. \tag{27}$$

Proof. Suppose that there does not exist $P \in \mathcal{S}^n$ such that (26) holds. From Proposition 11, this is equivalent to the existence of a full-rank $U \in \mathbf{C}^{n\times r}$, $V \in \mathbf{C}^{m\times r}$, and $S \in \mathbf{C}^{r\times r}$ with $S + S^* = 0$, such that

$$US - AU = BV, \quad \mathbf{Tr}\left(\begin{bmatrix} U^* & V^* \end{bmatrix} M \begin{bmatrix} U \\ V \end{bmatrix} \right) \leq 0. \tag{28}$$

We show that (28) is equivalent to the existence of $u \in \mathbf{C}^n$ and $v \in \mathbf{C}^m$, not both zero, such that (27) does not hold at some ω.

Suppose there exist $u \in \mathbf{C}^n$ and $v \in \mathbf{C}^m$, not both zero, such that (27) does not hold at some ω. Then, it is easy to verify that (28) holds with

$$U = [\Re u \ \ \Im u], \quad V = [\Re v \ \ \Im v], \quad S = \begin{bmatrix} 0 & -\omega \\ \omega & 0 \end{bmatrix}.$$

Conversely suppose that there exist full-rank $U \in \mathbf{C}^{n\times r}$, $V \in \mathbf{C}^{m\times r}$, and $S \in \mathbf{C}^{r\times r}$ with $S+S^* = 0$, such that (28) holds. We then take the Schur decomposition of S: $S = \sum_{i=1}^{m} j\omega_i q_i q_i^*$, where $\sum_i q_i q_i^* = I$. We then have

$$0 \geq \mathbf{Tr}\left(\left[\begin{array}{cc} U^* & V^* \end{array}\right] M \left[\begin{array}{c} U \\ V \end{array}\right]\right) = \mathbf{Tr}\left(\left[\begin{array}{cc} U^* & V^* \end{array}\right] M \left[\begin{array}{c} U \\ V \end{array}\right] \sum_i q_i q_i^*\right)$$
$$= \sum_{i=1}^{m} q_i^* \left[\begin{array}{cc} U^* & V^* \end{array}\right] M \left[\begin{array}{c} U \\ V \end{array}\right] q_i.$$

At least one of the m terms in this last expression must be less than or equal to zero. Let k be the index of that term, and define $u = Uq_k$, $v = Vq_k$. (u is nonzero because U has full rank.) We have

$$\left[\begin{array}{cc} u^* & v^* \end{array}\right] M \left[\begin{array}{c} u \\ v \end{array}\right] \leq 0$$

and, by multiplying $US - AU = BV$ with q_k on the right, $Au + Bv = j\omega_k u$. In other words we have constructed a u and v showing that (27) does not hold at $\omega = \omega_k$. □

Remark 8 Our statement of the KYP Lemma is more general than standard versions (see for example, [Ran96]), as we allow A to have imaginary eigenvalues. If A has no imaginary eigenvalues, then (27) simply means that

$$\left[\begin{array}{cc} B^*(-j\omega I - A^*)^{-1} & I \end{array}\right] M \left[\begin{array}{c} (j\omega I - A)^{-1}B \\ I \end{array}\right] > 0. \tag{29}$$

The following form of the frequency-domain condition is more commonly found in the literature: the inequality (29) holds for all ω where $j\omega I - A$ is invertible. If A has imaginary eigenvalues, then this condition is weaker than requiring that (27) holds for all ω, and it is not equivalent to feasibility of the LMI (23). Consider for example

$$A = \left[\begin{array}{cc} 0 & 1 \\ -1 & 0 \end{array}\right], \quad B = 0, \quad M = \mathbf{diag}(-I, I).$$

It is readily verified that the LMI (26) does not hold for any P. The frequency condition (27) does not hold at $\omega = 1$, $u = (1, j)$, $v = 0$:

$$(j\omega I - A)u = \left[\begin{array}{cc} j & -1 \\ 1 & j \end{array}\right] \left[\begin{array}{c} 1 \\ j \end{array}\right] = 0 = Bv, \quad \left[\begin{array}{cc} u^* & v^* \end{array}\right] M \left[\begin{array}{c} u \\ v \end{array}\right] = -2.$$

However the inequality (29) is clearly valid for all $\omega \neq \pm 1$, since $B = 0$. ◇

Remark 9 We may give a geometric interpretation to the proof of Lemma 1. The set

$$\mathcal{C} = \left\{ Z = \left[\begin{array}{cc} Z_{11} & Z_{12} \\ Z_{12}^* & Z_{22} \end{array}\right] \,\middle|\, Z \geq 0,\ Z_{11}^* A^* + AZ_{11} + Z_{12}B^* + BZ_{12}^* = 0 \right\}$$

is a closed convex cone. Its extreme directions have either the form

$$\begin{bmatrix} 0 \\ w \end{bmatrix} \begin{bmatrix} 0 & w^* \end{bmatrix}$$

for $w \neq 0$, or

$$\Re\left(\begin{bmatrix} u \\ v \end{bmatrix} \begin{bmatrix} u^* & v^* \end{bmatrix}\right)$$

where u, v are not both zero and satisfy $(j\omega I - A)u = Bv$ for some ω. The construction in the proof provides a way to decompose any element in $\mathcal{C}$ as a positive combination of those extreme directions.

The inequality $\mathbf{Tr}\, ZM \leq 0$ defines a halfspace. If $\mathcal{C}$ intersects this halfspace (i.e., (24) is feasible), then there must be an extreme direction in the halfspace. Any extreme direction in the halfspace provides a frequency ω and vectors u, v where

$$\begin{bmatrix} u^* & v^* \end{bmatrix} M \begin{bmatrix} u \\ v \end{bmatrix} \leq 0.$$

We can find at least one extreme direction in the halfspace by taking any Z that satisfies (24), and decomposing it as a positive combination of extreme directions of $\mathcal{C}$. $\diamond$

We next use the theorem of alternatives to exhibit the well-known connection between the KYP lemma and a certain Hamiltonian matrix.

Proposition 12 *Suppose that A has no imaginary eigenvalues and that $M_{22} > 0$. Then, exactly one of the following statements is true.*

1. *There exists $P \in \mathcal{S}^n$ such that (23) holds.*

2. *The Hamiltonian matrix*

$$H = \begin{bmatrix} A - BM_{22}^{-1}M_{12}^* & BM_{22}^{-1}B^* \\ M_{11} - M_{12}M_{22}^{-1}M_{12}^* & -(A - BM_{22}^{-1}M_{12}^*)^* \end{bmatrix}$$

   has an imaginary eigenvalue.

Proof. We established in the proof of Proposition 11 that the condition that there does not exist $P \in \mathcal{S}^n$ such that (23) holds is equivalent to the existence of $Z \in \mathcal{S}^{n+m}$ such that (24) holds. We now show that this condition is equivalent to H having imaginary eigenvalues.

First suppose that H has an imaginary eigenvalue $\pm j\omega$. We show that we can construct Z_{11}, Z_{12}, Z_{22} that satisfy (24). Let $V_1 \in \mathbf{C}^{n\times 2}$ and $V_2 \in \mathbf{C}^{n\times 2}$ be such that

$$\begin{bmatrix} A - BM_{22}^{-1}M_{12}^* & BM_{22}^{-1}B^* \\ M_{11} - M_{12}M_{22}^{-1}M_{12}^* & -(A - BM_{22}^{-1}M_{12}^*)^* \end{bmatrix} \begin{bmatrix} V_1 \\ V_2 \end{bmatrix} = \begin{bmatrix} V_1 \\ V_2 \end{bmatrix} \begin{bmatrix} 0 & -\omega \\ \omega & 0 \end{bmatrix},$$

with V_1 and V_2 not both zero. Then, it is readily verified that with

$$Z_{11} = V_1V_1^*, \;\; Z_{12} = V_1(V_2^*B - V_1^*M_{12})M_{22}^{-1}, \;\; Z_{22} = M_{22}^{-1}(B^*V_2 - M_{12}^*V_1)(V_2^*B - V_1^*M_{12})M_{22}^{-1},$$

condition (24) holds. (Indeed the last inequality holds with equality.)

Conversely suppose that there exists $Z \in \mathcal{S}^{n+m}$ such that (24) holds. From the KYP Lemma, there exist $\omega_0 \in \mathbf{R}$, $u \in \mathbf{C}^n$, and $v \in \mathbf{C}^m$ such that

$$Au + Bv = j\omega_0 u, \quad \begin{bmatrix} u^* & v^* \end{bmatrix} M \begin{bmatrix} u \\ v \end{bmatrix} \leq 0.$$

Eliminating u from the first equality yields

$$v^* \begin{bmatrix} B^*(-j\omega_0 I - A^*)^{-1} & I \end{bmatrix} M \begin{bmatrix} (j\omega_0 I - A)^{-1}B \\ I \end{bmatrix} v \leq 0.$$

Define

$$G(\omega) = \begin{bmatrix} B^*(-j\omega I - A^*)^{-1} & I \end{bmatrix} M \begin{bmatrix} (j\omega I - A)^{-1}B \\ I \end{bmatrix}.$$

Now, as $\omega \to \infty$, $G(\omega) \to M_{22} > 0$, and it follows from elementary continuity arguments that for some frequency ω_1, $G(\omega_1)$ must be singular. Thus, for some ω_1 and $w \in \mathbf{C}^m$, we must have

$$\begin{bmatrix} B^*(-j\omega_1 I - A^*)^{-1} & I \end{bmatrix} M \begin{bmatrix} (j\omega_1 I - A)^{-1}B \\ I \end{bmatrix} w = 0.$$

Defining

$$\begin{bmatrix} \tilde{u} \\ \tilde{v} \end{bmatrix} = \begin{bmatrix} (j\omega I - A)^{-1}Bw \\ (j\omega I + A^*)^{-1}\left(M_{11}(j\omega I - A)^{-1}B + M_{12}\right)w \end{bmatrix},$$

it is readily verified that

$$\begin{bmatrix} A - BM_{22}^{-1}M_{12}^* & BM_{22}^{-1}B^* \\ M_{11} - M_{12}M_{22}^{-1}M_{12}^* & -(A - BM_{22}^{-1}M_{12}^*)^* \end{bmatrix} \begin{bmatrix} \tilde{u} \\ \tilde{v} \end{bmatrix} = j\omega_1 \begin{bmatrix} \tilde{u} \\ \tilde{v} \end{bmatrix},$$

i.e., H has an imaginary eigenvalue $j\omega_1$. □

4.2 Strict Riccati inequality with positive definite P

Proposition 13 *Suppose $M_{22} > 0$. Exactly one of the following two statements is true.*

1. *There exists $P \in \mathcal{S}^n$ such that*

$$P > 0, \quad \begin{bmatrix} A^*P + PA & PB \\ B^*P & 0 \end{bmatrix} - M < 0.$$

2. *For some full-rank $U \in \mathbf{C}^{n\times r}$, $V \in \mathbf{C}^{m\times r}$, and $S \in \mathbf{C}^{r\times r}$ with $S + S^* \geq 0$,*

$$US - AU = BV, \quad \mathbf{Tr}\left(\begin{bmatrix} U^* & V^* \end{bmatrix} M \begin{bmatrix} U \\ V \end{bmatrix}\right) \leq 0. \tag{30}$$

Proof. With $\mathcal{A}_4$ as in Example 8 and with $A_0 = M$, the first statement of the theorem is equivalent to the existence of $P \in \mathcal{S}^n$ such that $\mathcal{A}_4(P) + A_0 > 0$. Then, applying Theorem **ALT 1**, the alternative is that there exists a $Z \in \mathcal{S}^{n+m} \times \mathcal{S}^n$ with

$$Z = \mathbf{diag}\left(\begin{bmatrix} Z_{11} & Z_{12} \\ Z_{12}^* & Z_{22} \end{bmatrix}, Z_2\right) \gneq 0,\ Z_{11}A^* + AZ_{11} + Z_{12}B^* + BZ_{12}^* - Z_2 = 0,\ \mathbf{Tr}\, Z\, \mathbf{diag}(M,0) \leq 0.$$

or equivalently, there exist Z_{11}, Z_{12}, Z_{22}, not all zero, such that

$$\begin{bmatrix} Z_{11} & Z_{12} \\ Z_{12}^* & Z_{22} \end{bmatrix} \geq 0, \quad Z_{11}A^* + AZ_{11} + Z_{12}B^* + BZ_{12}^* \geq 0, \quad \mathbf{Tr} \begin{bmatrix} Z_{11} & Z_{12} \\ Z_{12}^* & Z_{22} \end{bmatrix} M \leq 0. \tag{31}$$

We now show that this condition is equivalent to the existence of $U \in \mathbf{C}^{n\times r}$, $V \in \mathbf{C}^{m\times r}$, and $S \in \mathbf{C}^{r\times r}$ with $S + S^* \geq 0$ such

$$US - AU = BV, \quad \mathbf{Tr}\left(\begin{bmatrix} U^* & V^* \end{bmatrix} M \begin{bmatrix} U \\ V \end{bmatrix}\right) \leq 0.$$

We must have $Z_{11} \gneq 0$, as otherwise the last inequality in (31) would imply that $Z_{22} \leq 0$, a contradiction. Therefore, there exist $U \in \mathbf{C}^{n\times r}$ and $V \in \mathbf{C}^{m\times r}$ such that

$$\begin{bmatrix} Z_{11} & Z_{12} \\ Z_{12}^* & Z_{22} \end{bmatrix} = \begin{bmatrix} U & 0 \\ V & \hat{V} \end{bmatrix} \begin{bmatrix} U^* & V^* \\ 0 & \hat{V}^* \end{bmatrix}$$

where U has full rank. The equation $Z_{11}A^* + AZ_{11} + Z_{12}B^* + BZ_{12}^* \geq 0$, represented in terms of U and V means that $AUU^* + BVU^*$ has a positive semidefinite symmetric part, i.e., it can be written as $AUU^* + BVU^* = USU^*$, where $S + S^* \geq 0$. Since U has full rank, this last equation implies $AU + BV = US$. And inequality $\mathbf{Tr}\, ZM \leq 0$, expressed in terms of U and V, implies that (see proof of Proposition 11)

$$\mathbf{Tr}\left(\begin{bmatrix} U^* & V^* \end{bmatrix} M \begin{bmatrix} U \\ V \end{bmatrix}\right) \leq 0,$$

which completes the proof. □

Frequency-domain interpretations

Recall that we were able to extend Proposition 11 to yield the KYP Lemma, which establishes the connection between an LMI and a certain frequency-domain condition. Unfortunately, no such extensions are possible in general with Proposition 13. For example, the existence of full-rank $U \in \mathbf{C}^{n\times r}$, $V \in \mathbf{C}^{m\times r}$, and $S \in \mathbf{C}^{r\times r}$ with $S + S^* \geq 0$ such that

$$US - AU = BV, \quad \mathbf{Tr}\left(\begin{bmatrix} U^* & V^* \end{bmatrix} M \begin{bmatrix} U \\ V \end{bmatrix}\right) \leq 0$$

does not imply that for some $s \in \mathbf{C}$ with $\Re s \geq 0$, there exist $u \in \mathbf{C}^n$, and $v \in \mathbf{C}^m$ such that

$$(sI - A)u = Bv, \quad \begin{bmatrix} u^* & v^* \end{bmatrix} M \begin{bmatrix} u \\ v \end{bmatrix} \leq 0. \tag{32}$$

The reverse implication is true, however. These facts are well-known; see for example, [Wil74, Ran96].

In other words, conjectures such as "There exists $P \in \mathcal{S}^n$ such that

$$P > 0, \quad \begin{bmatrix} A^*P + PA & PB \\ B^*P & 0 \end{bmatrix} - M < 0, \tag{33}$$

if and only if

$$\begin{bmatrix} B^*(sI - A)^{-*} & I \end{bmatrix} M \begin{bmatrix} (sI - A)^{-1}B \\ I \end{bmatrix} > 0 \tag{34}$$

for all $s \in \mathbf{C}$ with $\Re s \geq 0$", are false. Here is a simple counterexample:

$$A = \begin{bmatrix} -1 & 1 \\ -10 & -1 \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ \sqrt{10} \end{bmatrix}, \quad M = \begin{bmatrix} -10 & -2 & 0 \\ -2 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

It is readily verified that while there does not exist $P \in \mathcal{S}^n$ such that (33) holds, the inequality (34) holds for all $s \in \mathbf{C}$ with $\Re s \geq 0$.

However, when M satisfies additional constraints, it is possible to provide a frequency-domain interpretation for Proposition 13.

Proposition 14 *Suppose $M_{22} > 0$, $M_{11} \leq 0$, and all the eigenvalues of A have negative real part. There exists $P \in \mathcal{S}^n$ such that*

$$P > 0, \quad \begin{bmatrix} A^*P + PA & PB \\ B^*P & 0 \end{bmatrix} - M < 0 \tag{35}$$

if and only if for all $s \in \mathbf{C}$ with $\Re s \geq 0$,

$$\begin{bmatrix} B^*(sI - A)^{-*} & I \end{bmatrix} M \begin{bmatrix} (sI - A)^{-1}B \\ I \end{bmatrix} > 0. \tag{36}$$

Proof. Suppose (36) does not hold for some s, i.e., there exists a nonzero $v \in \mathbf{C}^m$ such that

$$v^* \begin{bmatrix} B^*(sI - A)^{-*} & I \end{bmatrix} M \begin{bmatrix} (sI - A)^{-1}B \\ I \end{bmatrix} v \leq 0.$$

Hence, (30) is satisfied by $U = (sI - A)^{-1}Bv, \quad V = v, \quad S = s$, and by Proposition 13 this implies that (35) is infeasible.

Conversely, suppose (35) is infeasible. As we have seen in the proof of Proposition 13, this implies that there exists a nonzero Z such that

$$\begin{bmatrix} Z_{11} & Z_{12} \\ Z_{12}^* & Z_{22} \end{bmatrix} \geq 0, \quad Z_{11}A^* + AZ_{11} + Z_{12}B^* + BZ_{12}^* = Q, \quad \mathbf{Tr} \begin{bmatrix} Z_{11} & Z_{12} \\ Z_{12}^* & Z_{22} \end{bmatrix} M \leq 0$$

for some $Q = Q^* \geq 0$. Since all the eigenvalues of A have negative real part, the Lyapunov equation $WA^* + AW + Q = 0$ has a positive semidefinite solution W. Hence the matrix $\tilde{Z}$, defined as

$$\tilde{Z} = \begin{bmatrix} Z_{11} + W & Z_{12} \\ Z_{12}^* & Z_{22} \end{bmatrix},$$

satisfies $\tilde{Z} \geq 0$, $\quad \tilde{Z}_{11}A^* + A\tilde{Z}_{11} + \tilde{Z}_{12}B^* + B\tilde{Z}_{12}^* = 0$, and, because $M_{11} \leq 0$, also

$$\mathbf{Tr}\, M_{11}\tilde{Z}_{11} + 2\,\mathbf{Tr}\, M_{12}^*\tilde{Z}_{12} + \mathbf{Tr}\, M_{22}\tilde{Z}_{22} \leq 0.$$

We can now proceed as in the proof of Proposition 4.1 and Lemma 1, and construct from $\tilde{Z}$ two vectors u and v, not both zero, such that for some ω,

$$(j\omega I - A)u = Bv, \quad \begin{bmatrix} u^* & v^* \end{bmatrix} M \begin{bmatrix} u \\ v \end{bmatrix} \leq 0.$$

This means that

$$\begin{bmatrix} (-j\omega I - A^*)^{-1} & I \end{bmatrix} M \begin{bmatrix} (j\omega I - A)^{-1} \\ I \end{bmatrix} \not> 0,$$

and hence, (36) does not hold for $s = j\omega$.

4.3 Nonstrict Riccati inequalities

Proposition 15 *Suppose $M_{22} \geq 0$ and that all uncontrollable modes of (A, B) are nondefective and correspond to imaginary eigenvalues. Then, exactly one of the following two statements is true.*

1. *There exists $P \in \mathcal{S}^n$ such that*

$$\begin{bmatrix} A^*P + PA & PB \\ B^*P & 0 \end{bmatrix} - M \lneqq 0. \qquad (37)$$

2. *For some full-rank $U \in \mathbf{C}^{n\times n}$, $V \in \mathbf{C}^{m\times n}$, and $S \in \mathbf{C}^{n\times n}$ with $S + S^* = 0$,*

$$US - AU = BV, \quad \mathbf{Tr}\left(\begin{bmatrix} U^* & V^* \end{bmatrix} M \begin{bmatrix} U \\ V \end{bmatrix}\right) \leq 0.$$

Proof. With $\mathcal{A}_3$ as in Example 7 and with $A_0 = M$, the first statement of the theorem is equivalent to the existence of $P \in \mathcal{S}^n$ such that $\mathcal{A}_2(P) + A_0 \gneqq 0$. From Proposition 7, the condition that all uncontrollable modes of (A, B) are on the imaginary axis, and have geometric multiplicity one is equivalent to the nonexistence of P satisfying $\mathcal{A}(P) \gneqq 0$. Therefore, from Theorem **ALT 2**, the alternative is that there exists a $Z \in \mathcal{S}^{n+m}$ with

$$Z = \begin{bmatrix} Z_{11} & Z_{12} \\ Z_{12}^* & Z_{22} \end{bmatrix} > 0, \quad Z_{11}A^* + AZ_{11} + Z_{12}B^* + BZ_{12}^* = 0, \quad \mathbf{Tr}\, ZM \leq 0$$

It can be shown, using arguments similar to the ones in the proof of Proposition 11, that the second condition is equivalent to existence of $U \in \mathbf{C}^{n\times r}$, $V \in \mathbf{C}^{m\times r}$, and $S \in \mathbf{C}^{r\times r}$ with $S+S^* = 0$ such that

$$US - AU = BV, \quad \mathbf{Tr}\left(\begin{bmatrix} U^* & V^* \end{bmatrix} M \begin{bmatrix} U \\ V \end{bmatrix}\right) \leq 0. \tag{38}$$

□

Proposition 16 *Suppose $M_{22} \geq 0$, and that all modes of (A,B) corresponding to eigenvalues with nonnegative real part are controllable. Then, exactly one of the following two statements is true.*

1. *There exists $P \in \mathcal{S}^n$ such that*

$$P = P^* \leq 0, \quad \begin{bmatrix} A^*P + PA & PB \\ B^*P & 0 \end{bmatrix} - M \leq 0 \tag{39}$$

   with the matrices on the left-hand sides of the inequalities (39) not both zero.

2. *For some full-rank $U \in \mathbf{C}^{n\times n}$, $V \in \mathbf{C}^{m\times n}$, and $S \in \mathbf{C}^{n\times n}$ with $S+S^* > 0$,*

$$US - AU = BV, \quad \mathbf{Tr}\left(\begin{bmatrix} U^* & V^* \end{bmatrix} M \begin{bmatrix} U \\ V \end{bmatrix}\right) \leq 0.$$

Proof. Similar to the proof of Proposition 15, but using Theorem **ALT 2**, Proposition 8 and steps from the proof of Proposition 13. □

Remark 10 The only difference between Propositions 11 and 15 (and respectively Propositions 13 and 16) is that the requirement on the sizes of the matrices U, V and S. ◇

We are unaware of any simple frequency domain interpretations of the conclusions of Propositions 15 and 16. However, it turns out that when constraint qualifications are invoked that enable the application of Theorem **ALT 3**, the resulting alternatives to the LMI (37) and (39) have interpretations from control theory. We explore these next.

Proposition 17 *Suppose $M_{22} \geq 0$ and that all uncontrollable modes of (A,B) are nondefective and correspond to imaginary eigenvalues. Then, exactly one of the following two statements is true.*

1. *There exists $P \in \mathcal{S}^n$ such that*

$$\begin{bmatrix} A^*P + PA & PB \\ B^*P & 0 \end{bmatrix} - M \leq 0. \tag{40}$$

2. *For some full-rank $U \in \mathbf{C}^{n\times n}$, $V \in \mathbf{C}^{m\times r}$, and $S \in \mathbf{C}^{r\times r}$ with $S+S^* = 0$,*

$$US - AU = BV, \quad \mathbf{Tr}\left(\begin{bmatrix} U^* & V^* \end{bmatrix} M \begin{bmatrix} U \\ V \end{bmatrix}\right) < 0.$$

Proof. With $\mathcal{A}_3$ as in Example 7 and with $A_0 = M$, the first statement of the theorem is equivalent to the existence of $P \in \mathcal{S}^n$ such that $\mathcal{A}_3(P) + A_0 \geq 0$. From Proposition 7, the condition that all uncontrollable modes of (A,B) are nondefective and correspond to imaginary eigenvalues, means that there exists no P such that $\mathcal{A}_3(P) \gneq 0$. Therefore, from Theorem **ALT 3**, we have a necessary and sufficient condition for the infeasibility of (40): There exist Z_{11}, Z_{12}, Z_{22} such that

$$Z = \begin{bmatrix} Z_{11} & Z_{12} \\ Z_{12}^* & Z_{22} \end{bmatrix} \geq 0, \quad Z_{11}A^* + AZ_{11} + Z_{12}B^* + BZ_{12}^* = 0, \quad \mathbf{Tr}\, ZM < 0.$$

It can be shown, using arguments similar to the ones in the proof of Proposition 11, that this condition is equivalent to existence of $U \in \mathbf{C}^{n\times r}$, $V \in \mathbf{C}^{m\times r}$, and $S \in \mathbf{C}^{r\times r}$ with $S + S^* = 0$ such that

$$US - AU = BV, \quad \mathbf{Tr}\left(\begin{bmatrix} U^* & V^* \end{bmatrix} M \begin{bmatrix} U \\ V \end{bmatrix} \right) < 0. \tag{41}$$

□

Remark 11 The conclusions of Proposition 17 are closely related to conditions for the solvability of Algebraic Riccati Equations (AREs) and Inequalities (ARIs), derived by Scherer [Sch95a, Sch95b], for systems with uncontrollable modes on the imaginary axis. Scherer's approach is to reduce the original problem to that of solvability of an ARE for a smaller controllable system, with auxiliary LMIs of the form $A^*P + PA + S \geq 0$ where A has purely imaginary eigenvalues, and P is required to be "arbitrarily large". ◇

As with Proposition 11, the conclusion of Proposition 17 can be further developed to yield the nonstrict version of the Kalman-Yakubovich-Popov Lemma.

Lemma 2 (KYP Lemma, nonstrict version) *Suppose $M_{22} \geq 0$ and that all uncontrollable modes of (A,B) are nondefective and correspond to imaginary eigenvalues. There exists $P \in \mathcal{S}^n$ such that*

$$\begin{bmatrix} A^*P + PA & PB \\ B^*P & 0 \end{bmatrix} - M \leq 0,$$

if and only if for all $\omega \in \mathbf{R}$,

$$(j\omega I - A)u = Bv, \ (u,v) \neq 0 \Longrightarrow \begin{bmatrix} u^* & v^* \end{bmatrix} M \begin{bmatrix} u \\ v \end{bmatrix} \geq 0.$$

Remark 12 The conclusions of Proposition 17 have implications for quadratic optimal control. The mathematical setting of the following discussion is taken from the paper by Willems [Wil71]. Suppose that the linear system $\dot{x} = Ax + Bu$, is controllable, and consider the following optimal control problem.

$$\inf_{u \in \mathbf{L}_{2e}} \int_0^\infty \left(\begin{bmatrix} x(t)^* & u(t)^* \end{bmatrix} M \begin{bmatrix} x(t) \\ u(t) \end{bmatrix} \right) dt, \quad \text{s.t.} \quad \lim_{t\to\infty} x(t) \to 0. \tag{42}$$

Willems has shown that the infimum in (42) is bounded if and only if there exists $P \in \mathcal{S}^n$ such that

$$\begin{bmatrix} A^*P + PA & PB \\ B^*P & 0 \end{bmatrix} - M \leq 0. \tag{43}$$

We will now show using the theorems of alternatives that if there does not exist $P \in \mathcal{S}^n$ such that (43) is feasible, then there exists a state-feedback input $u = Kx$ such that the objective in (42) is unbounded below. If (43) is infeasible, then from Theorem **ALT 3**, there exist Z_{11}, Z_{12}, Z_{22} such that

$$Z = \begin{bmatrix} Z_{11} & Z_{12} \\ Z_{12}^* & Z_{22} \end{bmatrix} \geq 0, \quad Z_{11}A^* + AZ_{11} + Z_{12}B^* + BZ_{12}^* = 0, \quad \mathbf{Tr}\, MZ < 0.$$

Proceeding along the lines of the proof of Proposition 11, it is easy to show that there exist full-rank $U \in \mathbf{C}^{n\times r}$, $V \in \mathbf{C}^{m\times r}$, and $S \in \mathbf{C}^{r\times r}$ with $S + S^* = 0$ such that

$$US - AU = BV, \quad \mathbf{Tr}\left(\begin{bmatrix} U^* & V^* \end{bmatrix} M \begin{bmatrix} U \\ V \end{bmatrix} \right) < 0. \tag{44}$$

Let $K = V(U^*U)^{-1}U^*$. Then (44) can be rewritten as

$$(A + BK)U = US, \quad \mathbf{Tr}\left(\begin{bmatrix} U^* & V^* \end{bmatrix} M \begin{bmatrix} U \\ V \end{bmatrix} \right) < 0.$$

Thus, there exists a state-feedback input $u(t) = Kx(t)$ such that $A + BK$ has imaginary eigenvalues (those of S). Moreover as it is easy to show that for some initial condition $x(0)$ that lies in the column space of U (this is the invariant subspace of $A + BK$ corresponding to pure imaginary eigenvalues), the corresponding objective value

$$\int_0^\infty \left(\begin{bmatrix} x(t)^* & u(t)^* \end{bmatrix} M \begin{bmatrix} x(t) \\ u(t) \end{bmatrix} \right) dt < 0. \tag{45}$$

Of course, this input $u(t) = Kx(t)$ is inadmissible, as $\lim_{t\to\infty} x(t) \neq 0$. However, it is easy to establish using continuity-based arguments that we can construct an input $u(t) = \tilde{K}x(t)$ such that $x(t) \to 0$ arbitrarily slowly, yet with the objective being negative, as in (45). This magnitude of the objective can be made arbitrarily large (owing to the slow decay of $x(t)$). Thus, the objective in (42) is unbounded below. $\diamond$

Next, we have another variation of Proposition 17, where we impose constraints on P.

Proposition 18 *Suppose $M_{22} \geq 0$ and that all uncontrollable modes of (A, B) correspond to eigenvalues with positive real part. Then, exactly one of the following two statements is true.*

1. *Then there exists $P \in \mathcal{S}^n$ such that*

$$P \geq 0, \quad \begin{bmatrix} A^*P + PA & PB \\ B^*P & 0 \end{bmatrix} - M \leq 0. \tag{46}$$

2. *For some full-rank* $U \in \mathbf{C}^{n\times n}$, $V \in \mathbf{C}^{m\times r}$, *and* $S \in \mathbf{C}^{r\times r}$ *with* $S + S^* \geq 0$,

$$US - AU = BV, \quad \mathbf{Tr}\left(\left[\begin{array}{cc} U^* & V^* \end{array}\right] M \left[\begin{array}{c} U \\ V \end{array}\right]\right) < 0. \tag{47}$$

Proof. With $\mathcal{A}_4$ as in Example 8 and with $A_0 = M$, the first statement of the theorem is equivalent to the existence of $P \in \mathcal{S}^n$ such that $\mathcal{A}_4(P) + A_0 \geq 0$. From Proposition 8, the condition that all uncontrollable modes of (A, B) correspond to eigenvalues with positive real part means that there exists no P such that $\mathcal{A}_4(P) \gneq 0$. Therefore, from Theorem **ALT 3**, we have a necessary and sufficient condition for the infeasibility of (46): There exist Z_{11}, Z_{12}, Z_{22}, not all zero, such that

$$Z = \left[\begin{array}{cc} Z_{11} & Z_{12} \\ Z_{12}^* & Z_{22} \end{array}\right] \geq 0, \quad Z_{11}A^* + AZ_{11} + Z_{12}B^* + BZ_{12}^* \geq 0, \quad \mathbf{Tr}\, MZ < 0.$$

The rest of the proof proceeds along the lines of the proof of Proposition 13. □

Remark 13 When there does not exist $P \in \mathcal{S}^n$ such that (46) is feasible, it turns out that the infimum in the quadratic optimal control problem

$$\inf_{u\in\mathbf{L}_{2e}} \int_0^\infty \left(\left[\begin{array}{cc} x(t)^* & u(t)^* \end{array}\right] M \left[\begin{array}{c} x(t) \\ u(t) \end{array}\right]\right) dt \tag{48}$$

is unbounded below (see [Wil71]). (Note the important difference from the problem in (42): In the problem in (48), there are no terminal constraints on the state $x(t)$.) Using the theorem of alternatives, it is possible to construct a state-feedback input $u(t) = Kx(t)$ that demonstrates that the objective in (48) is unbounded below. The construction of the state-feedback here is much more tedious than with the remark following Lemma 2. ◇

5 The linear quadratic regulator problem

In §4, we considered convex Riccati inequalities, and explored system-theoretic interpretations of conditions for their feasibility via the theorems of alternatives. In this section, we consider the Linear Quadratic Regulator (LQR) problem, which is a classical semidefinite program with convex Riccati inequalities.

Consider the semidefinite program

$$\begin{array}{ll} \text{maximize} & x_0^* P x_0 \\ \text{subject to} & \left[\begin{array}{cc} A^*P + PA + Q & PB \\ B^*P & I \end{array}\right] \geq 0, \quad P \geq 0, \end{array} \tag{49}$$

with $Q \geq 0$. The SDP (49) can be rewritten as

$$\begin{array}{ll} \text{minimize} & \langle c, x \rangle_{\mathcal{V}} \\ \text{subject to} & \mathcal{A}(x) + A_0 \geq 0 \end{array} \tag{50}$$

where $\mathcal{A} : \mathcal{S}^n \to \mathcal{S}^{n+m} \times \mathcal{S}^n$, A_0, and c are defined as

$$\mathcal{A}(P) = \mathbf{diag}\left(\begin{bmatrix} A^*P + PA & PB \\ B^*P & 0 \end{bmatrix}, P\right), \quad A_0 = \mathbf{diag}(\begin{bmatrix} Q & 0 \\ 0 & I \end{bmatrix}, 0), \quad c = -x_0 x_0^*.$$

(Of course, the optimal value of Problem (50) is the negative of the optimal value of Problem (49).)

The dual problem of (50) (see §2.2) is

$$\begin{array}{ll} \text{maximize} & -\langle A_0, Z \rangle_{\mathcal{S}^n}, \\ \text{subject to} & \mathcal{A}^{\text{adj}}(Z) = c, \quad Z = Z^* \geq 0 \end{array} \tag{51}$$

It is readily verified that $\mathcal{A}^{\text{adj}} : \mathcal{S}^{n+m} \times \mathcal{S}^n \to \mathcal{V}$ is given by

$$\mathcal{A}^{\text{adj}}(Z) = Z_{11}A^* + AZ_{11} + Z_{12}B^* + BZ_{12}^* + Z_2,$$

where $Z \in \mathcal{S}^{n+m} \times \mathcal{S}^n$ is partitioned as $Z = \mathbf{diag}\left(\begin{bmatrix} Z_{11} & Z_{12} \\ Z_{12}^* & Z_{22} \end{bmatrix}, Z_2\right)$, with $Z_{11} \in \mathcal{S}^n$. Thus, problem (51) can be rewritten as

$$\begin{array}{ll} \text{maximize} & -\mathbf{Tr}\, QZ_{11} - \mathbf{Tr}\, Z_{22} \\ \text{subject to} & AZ_{11} + BZ_{12}^* + Z_{11}A^* + Z_{12}B^* + x_0x_0^* \leq 0, \quad \begin{bmatrix} Z_{11} & Z_{12} \\ Z_{12}^* & Z_{22} \end{bmatrix} \geq 0, \end{array} \tag{52}$$

with variables $Z_{11} \in \mathcal{S}^n$, $Z_{12} \in \mathbf{C}^{n \times m}$, $Z_{22} \in \mathcal{S}^m$.

5.1 Interpretation of the primal problem

Consider the following optimal control problem (cf. Remarks 12 and 13): For the system

$$\dot{x} = Ax + Bu, \quad x(0) = x_0, \tag{53}$$

$$\text{find } u \in \mathbf{L}_{2e} \text{ that minimizes} \quad J = \int_0^\infty (x(t)^*Qx(t) + u(t)^*u(t))\, dt, \tag{54}$$

with $Q \geq 0$, subject to $\lim_{t\to\infty} x(t) = 0$. Let J_{opt} denote the minimum value.

We can write down a lower bound for J_{opt} using quadratic functions. Suppose for $P \geq 0$ we have

$$\frac{d}{dt} x(t)^*Px(t) \geq -(x(t)^*Qx(t) + u(t)^*u(t)), \tag{55}$$

for all $t \geq 0$, and for all x and u satisfying $\dot{x} = Ax + Bu$, $x(T) = 0$. Then, integrating both sides from 0 to T, we get

$$x_0^*Px_0 \leq \int_0^T (x(t)^*Qx(t) + u(t)^*u(t))\, dt,$$

or we have a lower bound for J_{opt}.

Condition (55) holds for *all* x and u (not necessarily those that steer state to zero) if the LMI

$$\begin{bmatrix} A^*P + PA + Q & PB \\ B^*P & I \end{bmatrix} \geq 0$$

is satisfied. Thus, the optimal value of the SDP

$$\begin{array}{ll} \text{maximize} & x_0^* P x_0 \\ \text{subject to} & \begin{bmatrix} A^*P + PA + Q & PB \\ B^*P & I \end{bmatrix} \geq 0, \quad P \geq 0, \end{array}$$

provides a lower bound to the optimal value of Problem (54). This SDP is the same as (49).

5.2 Interpretation of the dual problem

Suppose (A, B) is stabilizable. Consider system (53) with a constant, linear state-feedback $u = Kx$ that stabilizes the system: $\dot{x} = (A + BK)x$, $x(0) = x_0$, with all the eigenvalues of $A + BK$ having negative real part. Then the LQR objective J reduces to

$$J_K = \int_0^\infty x(t)^* (Q + K^*K) x(t)\, dt.$$

Clearly, for every K, J_K yields an upper bound on the optimum LQR objective J_{opt}. From standard results in control theory, J_K can be evaluated as $\mathbf{Tr}\, \tilde{Z}(Q + K^*K)$, where $\tilde{Z}$ satisfies

$$(A + BK)\tilde{Z} + \tilde{Z}(A + BK)^* + x_0 x_0^* = 0,$$

with all the eigenvalues of $A + BK$ having negative real part. Thus, the best upper bound on J_{opt}, achievable using state-feedback control, is given by the optimization problem with the optimization variables $\tilde{Z}$ and K:

$$\begin{array}{ll} \text{minimize} & \mathbf{Tr}\, \tilde{Z}(Q + K^*K) \\ \text{subject to} & \tilde{Z} \geq 0, \quad (A + BK)\tilde{Z} + \tilde{Z}(A + BK)^* + x_0 x_0^* = 0, \end{array}$$

which has the same objective value as (52) evaluated at $Z_{11} = \tilde{Z}$, $Z_{12} = \tilde{Z}K^*$, $Z_{22} = K\tilde{Z}K^*$.

5.3 Condition for strict primal feasibility

From Proposition 13, strict primal feasibility is equivalent to the condition that there does not exist a full-rank $U \in \mathbf{C}^{n \times r}$, $V \in \mathbf{C}^{m \times r}$, and $S \in \mathbf{C}^{r \times r}$ with $S + S^* \geq 0$, such that

$$US - AU = BV, \quad \mathbf{Tr}\, U^*QU + V^*V \leq 0. \tag{56}$$

As $Q \geq 0$, condition (56) is equivalent to $QU = 0$ and $V = 0$, or we have $AU = US$, $\quad QU = 0$, which is equivalent to (Q, A) having unobservable modes in the closed-right half complex plane [Rug96]. In other words, primal feasibility is equivalent to (Q, A) having no unobservable modes corresponding to eigenvalues with nonnegative real part.

5.4 Condition for strict dual feasibility

Suppose the dual problem is strictly feasible, that is, there exist $Z_{11} \in \mathcal{S}^n$ and $Z_{12} \in \mathbf{C}^{n\times m}$ such that $Z_{11} > 0$ and $AZ_{11} + BZ_{12}^* + Z_{11}A^* + Z_{12}B^* + x_0x_0^* < 0$. With $K = Z_{12}^*Z_{11}^{-1}$, we then have

$$Z_{11} > 0, \quad (A+BK)Z_{11} + Z_{11}(A+BK) < 0,$$

or (A,B) is stabilizable, that is, all the uncontrollable modes are in the open left-half complex plane [Rug96].

5.5 Optimality conditions

If primal and dual are strictly feasible, then strong duality holds, and primal and dual optima are attained. By complementary slackness, we have

$$\begin{bmatrix} Z_{11} & Z_{12} \\ Z_{12}^* & Z_{22} \end{bmatrix} \begin{bmatrix} A^*P+PA+Q & PB \\ B^*P & I \end{bmatrix} = 0,$$

i.e.,

$$\begin{bmatrix} I \\ K \end{bmatrix} \begin{bmatrix} I & K^* \end{bmatrix} \begin{bmatrix} A^*P+PA+Q & PB \\ B^*P & I \end{bmatrix} = 0,$$

or

$$\begin{bmatrix} I & K^* \end{bmatrix} \begin{bmatrix} A^*P+PA+Q & PB \\ B^*P & I \end{bmatrix} = 0,$$

or $K = -B^*P$, with all the eigenvalues of $A+BK$ having negative real part, and

$$A^*P + PA + Q - PBB^*P = 0. \tag{57}$$

This is the classical LQR result, that states that when (A,B) is stabilizable and (Q,A) is detectable, then the optimal control u that solves Problem (54) is a constant state-feedback, with the feedback gain given via the stabilizing solution to the Algebraic Riccati Equation (57).

6 SDP duality and bounds on the H_∞-norm

Consider the LTI system

$$\dot{x} = Ax + Bu, \;\; x(0) = 0, \quad y = Cx, \tag{58}$$

where $A \in \mathbf{C}^{n\times n}$, $B \in \mathbf{C}^{n\times m}$, and $C \in \mathbf{C}^{p\times n}$, with all the eigenvalues of A having a negative real part. Let (A,B,C) be a minimal realization, and let H denote the transfer function, i.e., $H(s) = C(sI-A)^{-1}B$.

The $\mathbf{H}_\infty$ norm of H is defined as $\|H\|_\infty = \sup_{\Re s>0} \sigma_{\max}(H(s))$, where $\sigma_{\max}(\cdot)$ denotes the maximum singular value. It turns out that we also have

$$\begin{aligned} \|H\|_\infty &= \sup_{\omega\in\mathbf{R}} \sigma_{\max}(H(j\omega)) && (59) \\ &= \sqrt{\sup_{u,T_1,T_2} \left\{ \int_{T_1}^{T_2} y(t)^* y(t)\, dt \;\middle|\; \int_{T_1}^{T_2} u(t)^* u(t)\, dt \le 1 \right\}}. && (60) \end{aligned}$$

Equality (60) means that $\|H\|_\infty$ is the $\mathbf{L}_2$ gain of system (58), and equality (59) means that $\|H\|_\infty$ is the $\mathbf{L}_2$ gain of system (58) over all possible sinusoidal inputs, i.e., it is the $\mathbf{L}_2$-gain of system (58) over all frequencies.

It is well-known (see for example [BEFB94]) that the the optimal value of the SDP

$$\begin{array}{ll} \text{minimize} & \beta \\ \text{subject to} & \begin{bmatrix} A^*P+PA+C^*C & PB \\ B^*P & -\beta I \end{bmatrix} \le 0 \end{array} \qquad (61)$$

in the variables $P \in \mathcal{S}^n$ and $\beta \in \mathbf{R}$ is equal to $\|H\|_\infty^2$. If we take $\mathcal{V} = \mathcal{S}^n \times \mathbf{R}$, and define $\mathcal{A} : \mathcal{V} \to \mathcal{S}^{n+m}$, $A_0 \in \mathcal{S}^{n+m}$, $c \in \mathcal{V}$ as

$$\mathcal{A}(P,\beta) = -\begin{bmatrix} A^*P+PA & PB \\ B^*P & -\beta I \end{bmatrix}, \quad A_0 = -\begin{bmatrix} C^*C & 0 \\ 0 & 0 \end{bmatrix} \in \mathcal{S}^{n+m}, \quad c = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix},$$

the SDP (61) can be rewritten as

$$\begin{array}{ll} \text{minimize} & \langle c, x\rangle_{\mathcal{V}} \\ \text{subject to} & \mathcal{A}(x) + A_0 \ge 0. \end{array} \qquad (62)$$

The dual problem of (62) is

$$\begin{array}{ll} \text{minimize} & \langle A_0, Z\rangle_{\mathcal{S}^n} \\ \text{subject to} & \mathcal{A}^{\text{adj}}(Z) = c, \quad Z \ge 0. \end{array} \qquad (63)$$

It is readily verified that $\mathcal{A}^{\text{adj}} : \mathcal{S}^{n+m} \to \mathcal{V}$ is given by

$$\mathcal{A}^{\text{adj}}\left(\begin{bmatrix} Z_{11} & Z_{12} \\ Z_{12}^* & Z_{22} \end{bmatrix}\right) = \mathbf{diag}\left(Z_{11}A^* + AZ_{11} + Z_{12}B^* + BZ_{12}^*, \mathbf{Tr}\, Z_{22}\right).$$

Thus, Problem (63) can be rewritten as

$$\begin{array}{ll} \text{maximize} & \mathbf{Tr}\, CZ_{11}C^* \\ \text{subject to} & Z_{11}A^* + AZ_{11} + Z_{12}B^* + BZ_{12}^* = 0, \quad \begin{bmatrix} Z_{11} & Z_{12} \\ Z_{12}^* & Z_{22} \end{bmatrix} \ge 0, \quad \mathbf{Tr}\, Z_{22} = 1, \end{array} \qquad (64)$$

with variables $Z_{11} \in \mathcal{S}^n$, $Z_{12} \in \mathbf{C}^{n\times m}$, $Z_{22} \in \mathcal{S}^m$.

6.1 Control-theoretic interpretations of the lower bound

Any feasible point to Problem (64) yields a lower bound on $\|H\|_\infty^2$. We now provide control-theoretic interpretations of such a lower bound.

Time-domain interpretation

We establish the connection between the time-domain control-theoretic interpretation of $\|H\|_\infty$ from (60), and the lower bound based on the dual problem (64).

Let $u(t)$ be any input that steers the state of system (58) from $x(T_1) = 0$ to $x(T_2) = 0$ for some $T_1, T_2 \in \mathbf{R}$, with $\int_{T_1}^{T_2} u(t)^* u(t)\, dt = 1$. Let $y(t)$ be the corresponding output. Then, from (60), the quantity $\int_{T_1}^{T_2} y(t)^* y(t)\, dt$ serves as a lower bound to $\|H\|_\infty^2$.

Define

$$Z_{11} = \int_{T_1}^{T_2} x(t)x(t)^*\, dt, \quad Z_{12} = \int_{T_1}^{T_2} x(t)u(t)^*\, dt, \quad Z_{22} = \int_{T_1}^{T_2} u(t)u(t)^*\, dt.$$

We have

$$\begin{bmatrix} Z_{11} & Z_{12} \\ Z_{12}^* & Z_{22} \end{bmatrix} = \int_{T_1}^{T_2} \begin{bmatrix} x(t) \\ u(t) \end{bmatrix} \begin{bmatrix} x(t)^* & u(t)^* \end{bmatrix} \geq 0, \quad \mathbf{Tr}\, Z_{22} = \int_{T_1}^{T_2} u(t)^* u(t)\, dt = 1,$$

and

$$AZ_{11} + BZ_{12}^* + Z_{11}A^* + Z_{12}B^* = \int_{T_1}^{T_2} \frac{d}{dt}(x(t)x(t)^*)\, dt = x(T)x(T)^* - x(0)x(0)^* = 0.$$

Thus, Z_{11}, Z_{12} and Z_{22} are dual feasible. The corresponding dual objective is

$$\mathbf{Tr}\, CZ_{11}C^* = \int_{T_1}^{T_2} y(t)^* y(t)\, dt,$$

completing the connection between the control-theoretic interpretation (60), and the dual problem (64).

Frequency-domain interpretation

We next establish the connection between the frequency-domain control-theoretic interpretation of $\|H\|_\infty$ from (59), and the lower bound based on the dual problem (64).

Let $\omega \in \mathbf{R}$, and let $U \in \mathbf{C}^m$ with $U^*U = 1$. Define $X = (j\omega I - A)^{-1}BU$, $Z_{11} = \Re XX^*$, $Z_{12} = \Re XU^*$, and $Z_{22} = \Re UU^*$. Then,

$$\begin{bmatrix} Z_{11} & Z_{12} \\ Z_{12}^* & Z_{22} \end{bmatrix} = \Re \left(\begin{bmatrix} X \\ U \end{bmatrix} \begin{bmatrix} X^* & U^* \end{bmatrix} \right) \geq 0,$$

and

$$\begin{aligned} AZ_{11}+BZ_{12}^*+Z_{11}A^*+Z_{12}B^* &= \Re\left(AXX^*+XX^*A^*+XU^*B^*+BUX^*\right) \\ &= \Re\left((-j\omega I+A)XX^*+XX^*(-j\omega I+A)^*+XU^*B^*+BUX^*\right) \\ &= 0. \end{aligned}$$

Thus, Z_{11}, Z_{12} and Z_{22} are dual feasible. The value of the dual objective function is

$$\mathbf{Tr}\, C^*CZ_{11} = X^*C^*CX = U^*B^*(-j\omega I - A^*)^{-1}C^*C(j\omega I - A)BU,$$

which, from (59), is a lower bound on $\|H\|_\infty^2$. The control-theoretic interpretation of the above development is as follows. Suppose the input to system (58) is a complex exponential $u(t) = e^{j\omega t}U$. (Note that u is not in $\mathbf{L}_2$, i.e., $\int_0^T u(t)^*u(t)\,dt$ is unbounded with T. This problem can be addressed, using the standard technique of restricting u to have finite support, and then normalizing it so that it has unit $\mathbf{L}_2$ norm. We will henceforth ignore such technical issues, and just give the basic idea.) Then, the output of system (58) is $y(t) = C(j\omega I - A)^{-1}Be^{j\omega t}U$, and $\sqrt{U^*B^*(-j\omega I - A^*)^{-1}C^*C(j\omega I - A)BU}$ is the corresponding $\mathbf{L}_2$ gain. Thus, the above development demonstrates that for every $\omega \in \mathbf{R}$, $\sigma_{\max}(H(j\omega))$ can be proven to be a lower-bound on the $\mathbf{H}_\infty$ via the construction of a feasible solution for Problem (64).

6.2 Relation to Enns-Glover lower bound

Let W_c and W_o be the controllability and observability Gramians of the system (58) respectively, that is, $AW_{\rm c} + W_{\rm c}A^* + BB^* = 0$, and $W_{\rm o}A + A^*W_{\rm o} + C^*C = 0$. Let z be a unit-norm eigenvector corresponding to the largest eigenvalue of $W_{\rm c}^{1/2}W_{\rm o}W_{\rm c}^{1/2}$, and let X and Y be the solutions of the two Lyapunov equations

$$AY + YA^* + W_{\rm c}^{1/2}zz^*W_{\rm c}^{1/2} = 0, \quad A^*X + XA + W_{\rm c}^{-1/2}zz^*W_{\rm c}^{-1/2} = 0.$$

Define Z as

$$Z = \begin{bmatrix} Z_{11} & Z_{12} \\ Z_{12}^* & Z_{22} \end{bmatrix} = \begin{bmatrix} Y + W_{\rm c}XW_{\rm c} & W_{\rm c}XB \\ B^*XW_{\rm c} & B^*XB \end{bmatrix}.$$

We verify that Z is dual feasible. Obviously,

$$Z = \begin{bmatrix} Y & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} W_{\rm c} \\ B^* \end{bmatrix} X \begin{bmatrix} W_{\rm c} & B \end{bmatrix} \geq 0.$$

Secondly,

$$\mathbf{Tr}\, Z_{22} = \mathbf{Tr}\, BB^*X = -\mathbf{Tr}(AW_{\rm c} + W_{\rm c}A^*)X = -\mathbf{Tr}\, W_{\rm c}(A^*X + XA) = z^*z = 1.$$

Finally, it is easily verified that

$$AY + AW_{\rm c}XW_{\rm c} + YA^* + W_{\rm c}XW_{\rm c}A^* + W_{\rm c}XBB^* + BB^*XW_{\rm c} = 0,$$

so that $Z_{11}A^* + AZ_{11} + Z_{12}B^* + BZ_{12}^* = 0$. Moreover the objective value is

$$\mathbf{Tr}\, CZ_{11}C^* = \mathbf{Tr}\, CYC^* + \mathbf{Tr}\, CW_{\mathrm{c}}XW_{\mathrm{c}}C^* \geq \mathbf{Tr}\, CYC^* = \bar{\sigma}$$

where $\bar{\sigma}$ is the largest eigenvalue of $W_{\mathrm{c}}W_{\mathrm{o}}$. This lower bound on $\|H\|_\infty^2$ is the well-known Enns-Glover lower bound [Enn84, Glo84]. Note that the actual duality-based bound

$$\mathbf{Tr}\, CYC^* + \mathbf{Tr}\, CW_{\mathrm{c}}XW_{\mathrm{c}}C^*$$

is *guaranteed* to be at least as good as the Enns Glover bound.

Time-domain interpretation

We may interpret the Enns-Glover lower bound in the context of the time-domain interpretation for the dual objective, given in §6.1. Here $T_1 = -\infty$, $T_2 = \infty$, and

$$u(t) = \begin{cases} B^* e^{-A^*t} W_{\mathrm{c}}^{-1/2} z & (t \leq 0) \\ 0 & (t > 0) \end{cases}$$

Then,

$$x(t) = \begin{cases} W_{\mathrm{c}} e^{-A^*t} W_{\mathrm{c}}^{-1/2} z & (t \leq 0) \\ e^{At} W_{\mathrm{c}}^{1/2} z & (t > 0). \end{cases}$$

It is then readily verified that $\int_{T_1}^{T_2} u(t)^* u(t)\, dt = 1$, and $\int_{T_1}^{T_2} y(t)^* y(t)\, dt \geq \bar{\sigma}$.

6.3 New duality-based upper and lower bounds

Noting that every primal feasible point gives an upper bound and every dual feasible point gives a lower bound, it is possible to generate new bounds for $\|H\|_\infty$. It is readily checked that these bounds are often better than existing bounds.

New upper bounds. It is easily checked that $(2W_{\mathrm{o}}, 4\lambda_{\max}(W_{\mathrm{o}}BB^*W_{\mathrm{o}}, C^*C)) \in \mathcal{S}^n \times \mathbf{R}$ is a primal feasible point, where $\lambda_{\max}(R,S)$ is the maximum generalized eigenvalue of (R,S). Therefore one upper bound on $\|H\|_\infty$ is given by $2\sqrt{\lambda_{\max}(W_{\mathrm{o}}BB^*W_{\mathrm{o}}, C^*C)}$.

Let $\tilde{H}$ be defined by $\tilde{H}(s) = H(s)^T$; then we have $\|H\|_\infty = \|\tilde{H}\|_\infty$, which yields another upper found for $\|H\|_\infty$: $2\sqrt{\lambda_{\max}(W_{\mathrm{c}}C^*CW_{\mathrm{c}}, BB^*)}$.

New lower bounds. It is easily verified that $Z_{11} = W_{\mathrm{c}}/\alpha$, $Z_{12} = B/(2\alpha)$, $Z_{22} = B^*W_{\mathrm{c}}^{-1}B/(4\alpha)$, where $\alpha = \mathbf{Tr}(B^*W_{\mathrm{c}}^{-1}B/4)$, are dual feasible. Therefore a lower bound on $\|H\|_\infty$ is given by $2\sqrt{\mathbf{Tr}\, CW_{\mathrm{c}}C^*/(\mathbf{Tr}\, B^*W_{\mathrm{c}}^{-1}B)}$.

Once again noting $\|H\|_\infty = \|\tilde{H}\|_\infty$, where $\tilde{H}(s) = H(s)^T$, we have another lower bound $\|H\|_\infty$: $2\sqrt{\mathbf{Tr}\, B^*W_{\mathrm{o}}B/(\mathbf{Tr}\, CW_{\mathrm{o}}^{-1}C^*)}$.

7 Conclusions

We have explored the application of semidefinite programming duality in order to obtain new insight, as well as to provide new and simple proofs for some classical results for linear time-invariant systems. We have also shown how SDP duality can be used to derive new results, such as new LMI criteria for controllability (and observability) properties, as well as new upper and lower bounds for the $\mathbf{H}_\infty$ norm.

References

[Ali95] F. Alizadeh. Interior point methods in semidefinite programming with applications to combinatorial optimization. *SIAM Journal on Optimization*, 5(1):13–51, February 1995.

[BBI71] A. Berman and A. Ben-Israel. More on linear inequalities with applications to matrix theory. *Journal of Mathematical Analysis and Applications*, 33:482–496, 1971.

[BE] V. Balakrishnan and E. Feron (Eds). *Linear Matrix Inequalities in Control Theory and Applications*. special issue of the *International Journal of Robust and Nonlinear Control*, vol. 6, no. 9/10, pp. 896–1099, November-December, 1996.

[BEFB94] S. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan. *Linear Matrix Inequalities in System and Control Theory*, volume 15 of *Studies in Applied Mathematics*. SIAM, Philadelphia, PA, June 1994.

[BI69] A. Ben-Israel. Linear equations and inequalities on finite dimensional, real or complex vector spaces: a unified theory. *Journal of Mathematical Analysis and Applications*, 27:367–389, 1969.

[BLW91] S. Bittanti, A. J. Laub, and J. C. Willems, editors. *The Riccati Equation*. Springer Verlag, Berlin, Germany, 1991.

[BW81] J. Borwein and H. Wolkowicz. Regularizing the abstract convex program. *Journal of Mathematical Analysis and Applications*, 83:495–530, 1981.

[BY89] S. Boyd and Q. Yang. Structured and simultaneous Lyapunov functions for system stability problems. *Int. J. Control*, 49(6):2215–2240, 1989.

[CK77] B. D. Craven and J. J. Koliha. Generalizations of Farkas' theorem. *SIAM Journal of Mathematical Analysis*, 8:983–997, 1977.

[Con78] J. B. Conway. *Functions of One Complex Variable*. Springer-Verlag, New York, 1978.

[DP00] G. E. Dullerud and F. Paganini. *A Course in Robust Control Theory. A Convex Approach*. Springer-Verlag, 2000.

[DV75] C. A. Desoer and M. Vidyasagar. *Feedback Systems: Input-Output Properties*. Academic Press, New York, 1975.

[EN00] L. El Ghaoui and S.-I. Niculescu, editors. *Advances in Linear Matrix Inequality Methods in Control*. Advances in Control and Design. SIAM, Philadelphia, PA, 2000.

[Enn84] D. F. Enns. Model reduction with balanced realizations: An error bound and a frequency weighted generalization. In *Proc. IEEE Conf. on Decision and Control*, pages 127–132, Las Vegas, NV, December 1984.

[Fin37] P. Finsler. Über das Vorkommen definiter und semi-definiter Formen in Scharen quadratischer Formen. *Comentarii Mathematici Helvetici*, 9:199–192, 1937.

[GDN95] L. El Ghaoui, F. Delebecque, and R. Nikoukhah. *LMITOOL: A User-friendly Interface for LMI Optimization*. ENSTA/INRIA, 1995. Software available via anonymous FTP from `ftp.inria.fr`, under directory `pub/elghaoui/lmitool`.

[Glo84] K. Glover. All optimal Hankel-norm approximations of linear multivariable systems and their $\mathbf{L}_\infty$-error bounds. *Int. J. Control*, 39(6):1115–1193, 1984.

[GND99] Y. Genin, Y. Nesterov, and P. Van Dooren. The central point of LMI's and Riccati equations. In *Proceedings of the European Control Conference*, 1999.

[GNLC95] P. Gahinet, A. Nemirovskii, A. Laub, and M. Chilali. *The LMI Control Toolbox*. The MathWorks, Inc., 1995.

[HJ91] R. Horn and C. Johnson. *Topics in Matrix Analysis*. Cambridge University Press, Cambridge, 1991.

[HM01] D. Henrion and G. Meinsma. Rank-one LMIs and Lyapunov's inequality. *IEEE Trans. Aut. Control*, 46(8):1285–1288, August 2001.

[Jac77] D. H. Jacobson. *Extensions of Linear-Quadratic Control, Optimization and Matrix Theory*, volume 133 of *Mathematics in Science and Engineering*. Academic Press, London, 1977.

[Las95] J. B. Lasserre. A new Farkas lemma for positive semidefinite matrices. *IEEE Trans. Aut. Control*, 40(6):1131–1133, June 1995.

[Las97] J. B. Lasserre. A Farkas lemma without a standard closure condition. *SIAM J. on Control*, 35:265–272, 1997.

[LO96] A. S. Lewis and M. L. Overton. Eigenvalue optimization. *Acta Numerica*, pages 149–190, 1996.

[LR91] P. Lancaster and L. Rodman. Solutions of the continuous and discrete time algebraic Riccati equations: A review. In S. Bittanti, A. J. Laub, and J. C. Willems, editors, *The Riccati equation*, pages 11–51. Springer Verlag, Berlin, Germany, 1991.

[NN94] Yu. Nesterov and A. Nemirovsky. *Interior-point polynomial methods in convex programming*, volume 13 of *Studies in Applied Mathematics*. SIAM, Philadelphia, PA, 1994.

[Ran96] A. Rantzer. On the Kalman-Yacubovich-Popov lemma. *Syst. Control Letters*, 28(1):7–10, 1996.

[RTW97] M. Ramana, L. Tunçel, and H. Wolkowicz. Strong duality for semidefinite programming. *SIAM J. on Optimization*, 7, August 1997.

[Rug96] W. J. Rugh. *Linear System Theory*. Prentice Hall, New Jersey, 1996.

[Sch95a] C. W. Scherer. The algebraic Riccati equation and inequality for systems with uncontrollable modes on the imaginary axis. *SIAM J. on Matrix Analysis and Applications*, 16(4):1308–1327, 1995.

[Sch95b] C. W. Scherer. The general nonstrict algebraic Riccati inequality. *Linear Algebra and Appl.*, 219:1–33, 1995.

[SI95] R. E. Skelton and T. Iwasaki. Increased roles of linear algebra in control education. *IEEE Control Syst. Mag.*, 15(4):76–89, 1995.

[SN98] R. Shorten and K. S. Narendra. On the stability and existence of common Lyapunov functions for linear switching systems. In *Proc. IEEE Conf. on Decision and Control*, Tampa, FL, 1998.

[SN99] R. Shorten and K. S. Narendra. Necessary and sufficient conditions for the existence of CQLF's for two stable second order linear systems. In *Proc. American Control Conf.*, San Diego, CA, 1999.

[SN00] R. Shorten and K. S. Narendra. Necessary and sufficient conditions for the existence of CQLF's for M stable second order linear systems. In *Proc. American Control Conf.*, Chicago, IL, 2000.

[VB96a] L. Vandenberghe and V. Balakrishnan. Algorithms and software tools for LMI problems in control. session overview. In *Proc. IEEE CACSD Symposium*, Detroit, MI, September 1996. Invited session *Algorithms and Software Tools for LMI Problems in Control*.

[VB96b] L. Vandenberghe and S. Boyd. Semidefinite programming. *SIAM Review*, 38(1):49–95, March 1996.

[VB99] L. Vandenberghe and V. Balakrishnan. Semidefinite programming duality and linear system theory: connections and implications for computation. In *Proc. IEEE Conf. on Decision and Control*, pages 989–994, 1999.

[WB96] S.-P. Wu and S. Boyd. SDPSOL: *A Parser/Solver for Semidefinite Programming and Determinant Maximization Problems with Matrix Structure. User's Guide, Version Beta*. Stanford University, June 1996.

[Wil71] J. C. Willems. Least squares stationary optimal control and the algebraic Riccati equation. *IEEE Trans. Aut. Control*, AC-16(6):621–634, December 1971.

[Wil74] A. N. Willson Jr. A stability criterion for nonautonomous difference equations with application to the design of a digital FSK oscillator. *IEEE Trans. Circuits Syst.*, 21:124–130, 1974.

[Wol81] H. Wolkowicz. Some applications of optimization in matrix theory. *Linear Algebra and Appl.*, 40:101–118, 1981.

[WSV00] H. Wolkowicz, R. Saigal, and L. Vandenberghe, editors. *Handbook of Semidefinite Programming*, volume 27 of *International Series in Operations Research and Management Science*. Kluwer Academic Publishers, Boston, MA, 2000.

[YZZ99] D. D. Yao, S. Zhang, and X. Y. Zhou. LQ control via semidefinite programming. In *Proc. IEEE Conf. on Decision and Control*, pages 1027–1032, 1999.

A Proofs of the theorems of alternatives

A.1 Theorem ALT 1

The two statements contradict each other:

$$0 < \langle \mathcal{A}(x) + A_0, Z\rangle_{\mathcal{S}} = \langle x, \mathcal{A}^{\mathrm{adj}}(Z)\rangle_{\mathcal{V}} + \langle A_0, Z\rangle_{\mathcal{S}} = \langle A_0, Z\rangle_{\mathcal{S}} \leq 0.$$

(The first inequality follows from $\mathcal{A}(x) + A_0 > 0$ and $Z \gneqq 0$.) Therefore at most one of the two statements is true.

To complete the proof, we show that if statement 1 is false, then statement 2 must be true. Consider the set

$$C = \{U \in \mathcal{S} \mid \mathcal{A}(y) + U > 0 \text{ for some } y \in \mathcal{V}\}.$$

Suppose statement 1 is false, i.e., $A_0 \notin C$. Since C is open, nonempty, and convex, there must be a hyperplane strictly separating A_0 from C, i.e., there exists a nonzero $Z \in \mathcal{S}$ that satisfies

$$\langle A_0, Z\rangle_{\mathcal{S}} < \langle U, Z\rangle_{\mathcal{S}}$$

for all $U \in C$. In other words, Z must satisfy $Z \neq 0$ and

$$\langle A_0, Z\rangle_{\mathcal{S}} < \langle -\mathcal{A}(y) + X, Z\rangle_{\mathcal{S}} = -\langle y, \mathcal{A}^{\text{adj}}(Z)\rangle_{\mathcal{V}} + \langle X, Z\rangle_{\mathcal{S}} \tag{65}$$

for all $X > 0$ and all $y \in \mathcal{V}$. The first term in the right-hand side is unbounded below as a function of y if $\mathcal{A}^{\text{adj}}(Z) \neq 0$, and equal to zero if $\mathcal{A}^{\text{adj}}(Z) = 0$. Therefore if Z defines a separating hyperplane, it must satisfy $\mathcal{A}^{\text{adj}}(Z) = 0$. The second term is unbounded below as a function of $X > 0$ if $Z \not\geq 0$. This yields a second condition: $Z \geq 0$. If Z satisfies both conditions, the right-hand side of (65) is positive for all X and y, and can take values arbitrarily close to 0. The inequality is therefore satisfied for all y and all $X > 0$ if $\langle A_0, Z\rangle_{\mathcal{S}^n} \leq 0$. In summary, Z satisfies

$$Z \gneqq 0, \quad \mathcal{A}^{\text{adj}}(Z) = 0, \quad \langle A_0, Z\rangle_{\mathcal{S}} \leq 0.$$

A.2 Theorem ALT 2

The two statements clearly contradict each other:

$$0 < \langle \mathcal{A}(x) + A_0, Z\rangle_{\mathcal{S}} = \langle x, \mathcal{A}^{\text{adj}}(Z)\rangle_{\mathcal{V}} + \langle A_0, Z\rangle_{\mathcal{S}} \leq 0.$$

Therefore at most one of the statements is true.

Let $\mathcal{B} : \mathcal{W} \to \mathcal{S}$ be a linear mapping spanning the nullspace of $\mathcal{A}^{\text{adj}}$, i.e.,

$$\begin{aligned} \mathcal{A}^{\text{adj}}(Z) = 0 \quad &\Longleftrightarrow \quad Z = \mathcal{B}(y) \text{ for some } y \in \mathcal{W} \\ X = \mathcal{A}(x) \text{ for some } x \in \mathcal{V} \quad &\Longleftrightarrow \quad \mathcal{B}^{\text{adj}}(X) = 0. \end{aligned}$$

We can express any $A_0 \in \mathcal{S}$ as $A_0 = \mathcal{A}(x_0) + A_0^{\perp}$, where $x_0 \in \mathcal{V}$ and $\mathcal{A}^{\text{adj}}(A_0^{\perp}) = 0$. It is clear that statement 1 holds if and only if there exists $\tilde{x}$ satisfying $\mathcal{A}(\tilde{x}) + A_0^{\perp} \gneqq 0$. Statement 2 holds if and only if there exist $Z > 0$ such that $\mathcal{A}^{\text{adj}}(Z) = 0$, and $\langle A_0, Z\rangle_{\mathcal{S}} = \langle A_0^{\perp}, Z\rangle_{\mathcal{S}} \leq 0$. The theorem therefore holds if and only if it holds with A_0 replaced by $A_0^{\perp}$.

Suppose $A_0 = \mathcal{A}(x_0)$ for some $x_0 \in \mathcal{V}$, i.e., $A_0^{\perp} = 0$. By the definition of $\mathcal{B}$, we can reformulate the theorem as follows. Exactly one of the following two statements is true.

1. There exists $X \gneqq 0$ with $\mathcal{B}^{\text{adj}}(X) = 0$.

2. There exists $y \in \mathcal{W}$ with $\mathcal{B}(y) > 0$.

This result follows immediately from Theorem **ALT 1**.

Next, suppose A_0 is not in the range of $\mathcal{A}$, i.e., $A_0^\perp \neq 0$, and that there exists no $x \in \mathcal{V}$ with $\mathcal{A}(x) \gneqq 0$. Suppose the second statement is false. In particular, this means there is no $Z \in \mathcal{S}$ with $Z > 0$, $\mathcal{A}^{\mathrm{adj}}(Z) = 0$, and $\langle A_0, Z\rangle_{\mathcal{S}} < 0$. Therefore there exists no $y \in \mathcal{W}$ such that

$$\mathbf{diag}(\mathcal{B}(y), -\langle A_0, \mathcal{B}(y)\rangle_{\mathcal{S}}) > 0.$$

By Theorem **ALT1** this implies there exists $\mathbf{diag}(X, \lambda) \in \mathcal{S} \times \mathbf{R}$, such that

$$\mathbf{diag}(X, \lambda) \gneqq 0, \quad \mathcal{B}^{\mathrm{adj}}(X) - \lambda \mathcal{B}^{\mathrm{adj}}(A_0) = 0.$$

The last equality holds if and only if $X - \lambda A_0 = \mathcal{A}(x)$ for some $x \in \mathcal{V}$. Therefore there exist $x \in \mathcal{V}$, $\lambda \in \mathbf{R}$ such that

$$\mathbf{diag}(\mathcal{A}(x) + \lambda A_0, \lambda) \gneqq 0.$$

By assumption, $\lambda = 0$ is impossible. Therefore $\lambda > 0$, and dividing by λ yields an $\tilde{x} = x/\lambda$ satisfying $\mathcal{A}(\tilde{x}) + A_0 \geq 0$. Finally, we note that we must have $\mathcal{A}(\tilde{x}) + A_0 \neq 0$ because $A_0 = \mathcal{A}(x_0) + A_0^\perp$ with $A_0^\perp$ nonzero and orthogonal to the range of $\mathcal{A}$. Hence $\tilde{x}$ satisfies the conditions in the first statement.

A.3 Theorem ALT 3

The two statements clearly contradict each other:

$$0 \leq \langle \mathcal{A}(x) + A_0, Z\rangle_{\mathcal{S}} = \langle A_0, Z\rangle_{\mathcal{S}} < 0,$$

so at most one of the two statements is true.

If remains to show that at least one of the two statements is true. This is clearly the case if $A_0 = \mathcal{A}(x_0)$ for some x_0: statement 1 holds with $x = -x_0$; statement 2 is false.

Next, assume that $A_0 = A(x_0) + A_0^\perp$, where $\mathcal{A}^{\mathrm{adj}}(A_0^\perp) = 0$ and $A_0^\perp \neq 0$, and that there exists no $x \in \mathcal{V}$ such that $\mathcal{A}(x) \gneqq 0$. Suppose statement 1 is false. Then statement 1 of Theorem **ALT 2** is also false, and by Theorem **ALT 2**, there exists $Z \in \mathcal{S}$, with $Z > 0$, $\mathcal{A}^{\mathrm{adj}}(Z) = 0$, $\langle A_0, Z\rangle_{\mathcal{S}} \leq 0$. Since $A_0^\perp \neq 0$, there exists a small positive t, such that $\tilde{Z} = Z - tA_0^\perp$ satisfies $\tilde{Z} > 0$ and $\mathcal{A}^{\mathrm{adj}}(\tilde{Z}) = 0$, and moreover

$$\langle A_0, \tilde{Z}\rangle_{\mathcal{S}} = \langle A_0, Z\rangle_{\mathcal{S}} - t\langle A_0, A_0^\perp\rangle_{\mathcal{S}} = \langle A_0, Z\rangle_{\mathcal{S}} - t\langle A_0^\perp, A_0^\perp\rangle_{\mathcal{S}} < 0.$$

Therefore statement 2 of Theorem **ALT 3** is true.

A.4 Theorems ALT 4, ALT 5a, ALT 5b, and ALT 6

Choose a linear mapping $\mathcal{C} : \mathcal{U} \to \mathcal{V}$, where $\mathcal{U}$ is some vector space, that satisfies

$$\mathcal{B}(x) = 0 \Longleftrightarrow x = \mathcal{C}(u) \text{ for some } u \in \mathcal{U},$$

i.e., the range of $\mathcal{C}$ is the nullspace of $\mathcal{B}$. The adjoint of $\mathcal{C}$ satisfies

$$\mathcal{C}^{\text{adj}}(x) = 0 \Longleftrightarrow x = \mathcal{B}^{\text{adj}}(w) \text{ for some } w \in \mathcal{W}.$$

Define $\tilde{A} : \mathcal{U} \to \mathcal{S}^n$ by $\tilde{\mathcal{A}}(u) = \mathcal{A}(\mathcal{C}(u))$. Its adjoint is given by $\tilde{\mathcal{A}}^{\text{adj}}(Z) = \mathcal{C}^{\text{adj}}(\mathcal{A}^{\text{adj}}(Z))$. Therefore,

$$\tilde{\mathcal{A}}^{\text{adj}}(Z) = 0 \Longleftrightarrow \mathcal{A}^{\text{adj}}(Z) = \mathcal{B}^{\text{adj}}(w) \text{ for some } w \in \mathcal{W}.$$

The four theorems now follow from Theorems **ALT 1–ALT 3** applied to $\tilde{A}$.

B Proof of the duality theorem (Theorem 8)

Weak duality

Weak duality is straightforward. If x is primal feasible and Z is dual feasible, then

$$\langle c, x\rangle_{\mathcal{V}} = \langle \mathcal{A}^{\text{adj}}(Z), x\rangle_{\mathcal{V}} = \langle Z, \mathcal{A}(x)\rangle_{\mathcal{S}} \geq -\langle Z, A_0\rangle_{\mathcal{S}}.$$

Therefore

$$p_{\text{opt}} = \inf_{\mathcal{A}(x)+A_0 \geq 0} \langle c, x\rangle_{\mathcal{V}} \geq \sup_{\mathcal{A}^{\text{adj}}(Z)=c,\, Z\geq 0} -\langle Z, A_0\rangle_{\mathcal{S}} = d_{\text{opt}}.$$

Strict primal feasibility implies strong duality

Suppose the primal problem is strictly feasible, i.e., there exists an x_0 with $\mathcal{A}(x_0) + A_0 > 0$. Define $X_0 = \mathcal{A}(x_0) + A_0$ and $t_0 = \langle c, x_0\rangle_{\mathcal{V}}$. Consider the set

$$C = \{(X,t) \in \mathcal{S} \times \mathbf{R} \mid \mathcal{A}(x) + A_0 \geq X, \langle c, x\rangle_{\mathcal{V}} \geq t, \text{ for some } x \in \mathcal{V}\}.$$

C is a nonempty convex set.

Suppose p_{opt} is finite. Then the point $(X,t) = (0, p_{\text{opt}})$ is in the boundary of C. Therefore there exists a supporting hyperplane to C at $(0, p_{\text{opt}})$, i.e., there exist $Z \in \mathcal{S}$ and $\mu \in \mathbf{R}$, not both zero, that satisfy

$$\langle Z, X\rangle + \mu t \leq \mu p_{\text{opt}} \tag{66}$$

for all $(X,t) \in C$. Note that $(X,t) \in C$ for all $X \leq X_0$ and all $t \geq t_0$. If we fix $t = t_0$, the left-hand side of (66) is bounded above as a function of $X \leq X_0$ only if $Z \geq 0$. If we fix $X = X_0$, it is bounded above as a function of $t \geq t_0$ only if $\mu \leq 0$. Next, note that $(\mathcal{A}(x)+A_0, \langle c,x\rangle_{\mathcal{V}}) \in C$ for all $x \in \mathcal{V}$. Therefore,

$$\langle Z, \mathcal{A}(x)+A_0\rangle_{\mathcal{S}} + \mu\langle c,x\rangle_{\mathcal{V}} = \langle \mathcal{A}^{\mathrm{adj}}(Z)+\mu c, x\rangle_{\mathcal{V}} + \langle A_0, Z\rangle_{\mathcal{S}} \leq \mu p_{\mathrm{opt}}$$

for all $x \in \mathcal{V}$. This is only possible if $\mathcal{A}^{\mathrm{adj}}(Z)+\mu c = 0$. In summary, Z and μ are not both zero and satisfy

$$Z \geq 0, \quad \mu \leq 0, \quad \mathcal{A}^{\mathrm{adj}}(Z)+\mu c = 0, \quad \langle A_0, Z\rangle_{\mathcal{S}} \leq \mu p_{\mathrm{opt}}.$$

If $\mu = 0$, this reduces to

$$Z \gneqq 0, \quad \mathcal{A}^{\mathrm{adj}}(Z) = 0, \quad \langle A_0, Z\rangle_{\mathcal{S}} \leq 0.$$

By Theorem **ALT 1** this contradicts our assumption that the primal problem is strictly feasible. Therefore we must have $\mu < 0$, and $\tilde{Z} = -Z/\mu$ satisfies

$$\tilde{Z} \geq 0, \quad \mathcal{A}^{\mathrm{adj}}(\tilde{Z}) = c, \quad -\langle A_0, \tilde{Z}\rangle_{\mathcal{S}} \geq p_{\mathrm{opt}},$$

i.e., $\tilde{Z}$ is dual feasible with an objective value greater than or equal to p_{opt}. By weak duality, this is only possible if $\tilde{Z}$ is dual optimal, i.e., $-\langle A_0, \tilde{Z}\rangle_{\mathcal{S}} = d_{\mathrm{op}} = p_{\mathrm{opt}}$.

Next, suppose $p_{\mathrm{opt}} = -\infty$. This means that the primal problem is unbounded below, i.e., for all t, there exist x such that

$$\mathcal{A}(x)+A_0 > 0, \quad \langle c,x\rangle_{\mathcal{V}} < t.$$

By Theorem **ALT 1** this implies that there exist no t, Z, μ with Z and μ not both zero, that satisfy

$$Z \geq 0, \quad \mu \geq 0, \quad \mathcal{A}^{\mathrm{adj}}(Z) = \mu c, \quad \langle A_0, Z\rangle_{\mathcal{S}} + \mu t \leq 0.$$

In particular, taking $\mu = 1$, we see that there is no Z that satisfies $Z \geq 0$, $\mathcal{A}^{\mathrm{adj}}(Z) = c$, i.e., the dual problem is infeasible and $d_{\mathrm{opt}} = -\infty$.

Strict dual feasibility implies strong duality

Let $\mathcal{B} : \mathcal{W} \to \mathcal{S}$ be a linear mapping satisfying

$$\begin{aligned} \mathcal{A}^{\mathrm{adj}}(Z) = 0 \quad &\Longleftrightarrow \quad Z = \mathcal{B}(y) \text{ for some } y \in \mathcal{W} \\ X = \mathcal{A}(x) \text{ for some } x \in \mathcal{V} \quad &\Longleftrightarrow \quad \mathcal{B}^{\mathrm{adj}}(X) = 0. \end{aligned}$$

Suppose the dual problem is strictly feasible, i.e., there exists a $Z_0 > 0$ with $\mathcal{A}^{\mathrm{adj}}(Z_0) = c$.

Z satisfies $\mathcal{A}^{\mathrm{adj}}(Z) = c$ if and only if $Z - Z_0 = \mathcal{B}(y)$ for some y. Therefore the dual problem can be reformulated as

$$\begin{array}{ll} \text{maximize} & -\langle A_0, Z_0\rangle_{\mathcal{S}} - \langle \mathcal{B}^{\mathrm{adj}}(A_0), y\rangle_{\mathcal{W}} \\ \text{subject to} & \mathcal{B}(y) + Z_0 \geq 0. \end{array}$$

In other words $d_{\text{opt}} = -\langle A_0, Z_0\rangle_{\mathcal{S}} - \tilde{p}_{\text{opt}}$, where $\tilde{p}_{\text{opt}}$ is the optimal value of the SDP

$$\begin{array}{ll} \text{minimize} & \langle \mathcal{B}^{\text{adj}}(A_0), y\rangle_{\mathcal{W}} \\ \text{subject to} & \mathcal{B}(y) + Z_0 \geq 0. \end{array}$$

This problem is strictly feasible ($y = 0$ is strictly feasible), so it satisfies strong duality, i.e., its optimal value $\tilde{p}_{\text{opt}}$ is equal to the optimal value $\tilde{d}_{\text{opt}}$ of the corresponding dual problem

$$\begin{array}{ll} \text{maximize} & -\langle Z_0, X\rangle_{\mathcal{S}} \\ \text{subject to} & \mathcal{B}^{\text{adj}}(X) = \mathcal{B}^{\text{adj}}(A_0) \\ & X \geq 0. \end{array} \tag{67}$$

X satisfies the equality constraint if and only if $X - A_0 = \mathcal{A}(x)$ for some x. The SDP (67) is therefore equivalent to (i.e., has the same optimal value as)

$$\begin{array}{ll} \text{maximize} & -\langle Z_0, \mathcal{A}(x) + A_0\rangle_{\mathcal{S}} = -\langle c, x\rangle_{\mathcal{V}} - \langle A_0, Z\rangle_{\mathcal{S}} \\ \text{subject to} & \mathcal{A}(x) + A_0 \geq 0. \end{array}$$

Comparing this with the original primal problem (6) we conclude that

$$p_{\text{opt}} = -\langle A_0, Z_0\rangle_{\mathcal{S}} - \tilde{d}_{\text{opt}} = -\langle A_0, Z_0\rangle_{\mathcal{S}} - \tilde{p}_{\text{opt}} = d_{\text{opt}}.$$

C Proof of the optimality conditions

Suppose $p_{\text{opt}} = d_{\text{opt}}$ and x and Z are primal and dual optimal. Then

$$\langle c, x\rangle_{\mathcal{V}} = \langle \mathcal{A}^{\text{adj}}(Z), x\rangle_{\mathcal{V}} = \langle Z, \mathcal{A}(x)\rangle_{\mathcal{S}} = -\langle Z, A_0\rangle_{\mathcal{S}}.$$

Therefore $\langle Z, \mathcal{A}(x) + A_0\rangle_{\mathcal{S}} = 0$. Since $Z \geq 0$ and $\mathcal{A}(x) + A_0 \geq 0$, this is only possible if

$$Z(\mathcal{A}(x) + A_0) = (\mathcal{A}(x) + A_0)Z = 0.$$

The remaining two facts were already proved in Appendix B. For example, we have established strong duality for a strictly feasible primal problem with finite optimal value p_{opt}, by showing that there exists a dual feasible Z with objective value p_{opt}.