1 LMIs

A linear matrix inequality (LMI) is an affine matrix-valued function,

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

where $x \in \mathbb{R}^m$ are called the decision variables and $F_i = F_i^T \in \mathbb{R}^{n\times n}$ are symmetric matrices.

A very important aspect of the LMI is that it defines a *convex set*. To see this, let x_1 and x_2 be two solutions to an LMI problem, i.e. $F(x_1) \succ 0$ and $F(x_2) \succ 0$. Then also any convex combination $x = (1-\lambda)x_1 + \lambda x_2$, with $\lambda \in [0,1]$ solves the LMI:

$$F(x) = F((1-\lambda)x_1 + \lambda x_2) = (1-\lambda)F(x_1) + \lambda F(x_2) \succ 0. \tag{2}$$

Efficient numerical methods have been developed to solve these kind of problems. Some of thesemost efficient algorithms for solving LMI problems are based on interior point methods. These methods are iterative and each iteration includes a least squares minimization problem.

Example 1.1 Most LMIs are not formulated in the standard form (1) but they can be rewritten as is shown in this example. Let us consider the following Lyapunov problem:

$$\begin{cases} P = P^T \succ 0 \\ A^T P + PA \prec 0. \end{cases}$$

Note that P enters linearly in both inequalities. To show how to rewrite this into the standard LMI form, we assume that $P \in \mathbb{R}^{2\times 2}$. Parametrize P as a linear function in x:

$$P(x) = x_1 \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + x_2 \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} + x_3 \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} = x_1 P_1 + x_2 P_2 + x_3 P_3.$$

Then $F(x) = \operatorname{diag}\left[P(x), -A^T P(x) - P(x)A\right]$, or

$$\begin{aligned} F(x) = x_1 &\begin{bmatrix} P_1 & 0 \\ 0 & -A^T P_1 - P_1 A \end{bmatrix} + x_2 \begin{bmatrix} P_2 & 0 \\ 0 & -A^T P_2 - P_2 A \end{bmatrix} \\ &+ x_3 \begin{bmatrix} P_3 & 0 \\ 0 & -A^T P_3 - P_3 A \end{bmatrix}, \end{aligned}$$

which is in the form (1) with $F_0 = 0$. In this case three decision variables are needed. In general we need $n(n+1)/2$ decision variables for symmetric matrices and n^2 for full square matrices of size $n \times n$. □

1.1 Some Standard LMI Problems

Some standard LMI problems are listed in [2]. The most important ones are

LMIP: The LMI problem is to find a feasible x such that $F(x) \succ 0$ or to determine if the LMI is infeasible.

As an example of an LMIP we take the problem in example 1.1 of finding a Lyapunov matrix $P = P^T \succ 0$ such that

$$A^T P + PA \prec 0. \tag{3}$$

EVP/PDP: The eigenvalue problem (EVP) is to minimize the maximum eigenvalue of a matrix $A(x)$ that depends affinely on a variable subject to an LMI constraint. That is

$$\min_{\lambda,x}\{\lambda : \lambda I - A(x) \succ 0, B(x) \succ 0\}. \tag{4}$$

This is equivalent to minimizing a linear function of x subject to an LMI constraint:

$$\min_{x}\{c^T x : F(x) \succ 0\}. \tag{5}$$

The latter formulation is called positive definite programming (PDP) or, if the inequality is nonstrict, semidefinite programming (SDP) [12].

As an example of an PDP we consider the bounded real lemma (lemma 2), which determines the $\mathcal{H}_\infty$ norm of a system $G(s) = D + C(sI - A)^{-1}B$ by minimizing γ with respect to $P \succ 0$ subject to

$$\begin{bmatrix} PA + A^T P & PB & C^T \\ B^T P & -\gamma I & D^T \\ C & D & -\gamma I \end{bmatrix} \prec 0.$$

Note that the LMI problem, $F(x) \succ 0$ can be formulated as an EVP by

$$\min_{\lambda,x}\{\lambda : F(x) + \lambda I \succ 0\} \leq 0, \tag{6}$$

for which a feasible point, (λ, x), can be easily found for any initial x by choosing λ sufficiently large.

GEVP: The generalized eigenvalue problem is to minimize the maximum generalized eigenvalue of a pair of matrices that depend affinely on a variable, subject to an LMI constraint. The general form of a GEVP is

$$\min_{\lambda,x}\{\lambda : \lambda B(x) - A(x) \succ 0, B(x) \succ 0, C(x) \succ 0\} \tag{7}$$

where $A(x)$, $B(x)$ and $C(x)$ are symmetric matrices that depend affinely (linearly) on $x \in \mathbb{R}^m$.

As an example of a GEVP we take the problem of finding the upper bound ν of the complex μ value of a matrix M. This problem is solved by minimizing $\nu > 0$ (or equivalently ν^2) with respect to $P \succ 0$ subject to

$$M^* PM \prec \nu^2 P.$$

The GEVP problem is not convex but quasi-convex.

1.2 Interior Point Methods

1.2.1 Analytic Center of an Affine Matrix Inequality

We will here consider a general affine matrix inequality $F(x) \succ 0$, where

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

and $F_i = F_i^T \in \mathbb{R}^{n \times n}$. Without loss of generality, assume that the matrices $F_1, \dots, F_m$ are linearly independent. Denote the feasible set by $\mathcal{X}$:

$$\mathcal{X} = \{x \in \mathbb{R}^m : F(x) \succ 0\} \tag{8}$$

The function

$$\phi(x) = \begin{cases} \log\det F^{-1}(x) & x \in \mathcal{X} \\ \infty & x \notin \mathcal{X} \end{cases}$$

is finite if and only if $x \in \mathcal{X}$, and becomes infinite as x approaches the boundary of $\mathcal{X}$; ϕ is called a *barrier function* for $\mathcal{X}$. There are other barrier functions, but this one enjoys many special properties including that it is analytic and strictly convex when $x \in \mathcal{X}$.

Suppose now that $\mathcal{X}$ is nonempty and bounded. Denote the unique minimizer with

$$x^\star = \arg\min_x \phi(x). \tag{9}$$

We refer to $x^\star$ as the analytic center of the affine matrix inequality $F(x) \succ 0$. Equivalently, $F(x^\star)$ has maximum determinant among all positive definite matrices of the form $F(x)$. Note that the analytic center $x^\star$ is a property of $F(x)$ and not of $\mathcal{X}$. Two Fs may define the same $\mathcal{X}$ but may have different $x^\star$.

Newton's method, with appropriate step length selection, can be used efficiently to compute $x^\star$, given an initial point in $\mathcal{X}$. Note that a feasible point can be found by first solving an auxiliary problem defined by $\lambda I + F(x) \succ 0$, where λ is chosen sufficiently large. Then λ is minimized until it becomes less than zero.

We consider the algorithm:

$$x^{(k+1)} = x^{(k)} - \alpha^{(k)} H(x^{(k)})^{-1} g(x^{(k)})$$

where $\alpha^{(k)}$ is the damping factor of the kth iteration, $g(x)$ and $H(x)$ are the gradient and the Hessian respectively:

$$\begin{aligned} g_i(x) &= -\operatorname{tr} F(x)^{-1} F_i \\ &= -\operatorname{tr} F(x)^{-1/2} F_i F(x)^{-1/2} \end{aligned}$$

$$\begin{aligned} H_{ij}(x) &= \operatorname{tr} F(x)^{-1} F_i F(x)^{-1} F_j \\ &= \operatorname{tr} \left(F(x)^{-1/2} F_i F(x)^{-1/2}\right) \left(F(x)^{-1/2} F_j F(x)^{-1/2}\right). \end{aligned}$$

Nesterov and Nemirovsky [10] give a simple step length rule appropriate for a general class of so called self-concordant barrier functions. Their damping factor depends on a quantity that they call the *Newton decrement* of ϕ at x:

$$\delta(x) = \left\| H(x)^{-1/2} g(x) \right\|.$$

The Nesterov-Nemirovsky damping factor is

$$\alpha^{(k)} = \begin{cases} 1 & \delta(x^{(k)}) \leq 1/4 \\ 1/(1+\delta(x^{(k)})) & \delta(x^{(k)}) > 1/4. \end{cases}$$

Nesterov and Nemirovsky show that this step length always results in $x^{(k+1)} \in \mathcal{X}$. Moreover, for $\delta(x^{(k)}) < 1/4$, we have $\delta(x^{(k+1)}) \leq 2\delta(x^{(k)})^2$. Thus, the algorithm converges quadratically once we start taking undamped Newton steps. A better step length can be obtained by exact line-search, i.e.

$$\alpha^{(k)} = \arg\min_{\alpha} \phi(x^{(k)} + \alpha v^{(k)}),$$

where $v^{(k)} = -H(x^{(k)})^{-1} g(x^{(k)})$.

The undamped Newton step $-H(x)^{-1}g(x)$ can be interpreted as the solution of the weighted least squares problem:

$$\begin{aligned} H(x)^{-1}g(x) &= \arg\min_{v} \left\| F(x)^{-1/2} F(x-v) F(x)^{-1/2} \right\|_F \\ &= \arg\min_{v} \left\| I - \sum_{i=1}^{m} v_i F(x)^{-1/2} F_i F(x)^{-1/2} \right\|_F . \end{aligned}$$

where $\|.\|_F$ denotes the Frobenius norm of a matrix. The Frobenius norm is the square root of the sum of squares of all elements of the matrix, or equivalently $\|F\|_F^2 = \operatorname{tr} F^T F$.

1.2.2 The Path of Centers

Consider the standard EVP/PDP:

$$\lambda^{\text{opt}} = \min_{x} \{c^T x : F(x) \succ 0\}.$$

For each $\lambda > \lambda^{\text{opt}}$ the LMI

$$F(x) \succ 0, \qquad c^T x < \lambda \tag{10}$$

is feasible. Assuming that (10) has a bounded feasible set the analytic center exists, which is denoted by $x^\star(\lambda)$. The curve defined by $x^\star(\lambda)$ for $\lambda > \lambda^{\text{opt}}$ is called the *path of centers* for the EVP/PDP. It is analytic and has a limit as $\lambda \to \lambda^{\text{opt}}$, which is the optimal solution.

1.2.3 Methods of Centers

Perhaps the simplest optimization algorithm based on the notion of analytic center is the method of centers [8]. Consider the GEVP:

$$\min_{\lambda, x} \{\lambda : \lambda B(x) - A(x) \succ 0, B(x) \succ 0, C(x) \succ 0\}.$$

The algorithm is initialized with $\lambda^{(0)}$ and $x^{(0)}$ with $\lambda^{(0)} B(x^{(0)}) - A(x^{(0)}) \succ 0$ and $C(x^{(0)}) \succ 0$, and proceeds as follows:

$$\begin{aligned} \lambda^{(k+1)} &= (1-\theta)\, \lambda_{\max}(A(x^{(k)}), B(x^{(k)})) + \theta\, \lambda^{(k)} \\ x^{(k+1)} &= x^\star(\lambda^{(k+1)}) \end{aligned}$$

where $0 < \theta < 1$, $\lambda_{\max}$ is the maximum generalized eigenvalue and $x^\star(\lambda)$ denotes the analytic center of $\mathrm{diag}[\lambda B(x) - A(x),\ C(x)]$. Simple proofs of the convergence of this algorithm are given in [1, 2]. Among interior-point methods, the method of centers is not the most efficient. The most efficient methods today appear to be primal-dual methods and projective methods [10].

1.2.4 Primal and Dual Methods

Primal-dual methods have been developed for positive definite programming (PDP) and semidefinite programming (SDP) [12].

1.2.5 Duality

Consider the EVP/PDP

$$p^* = \inf_x \left\{ c^T x : F(x) = F_0 + \sum_i^m x_i F_i \succ 0 \right\}. \tag{11}$$

Associated with this, the so called primal problem, there is a dual problem defined by

$$d^* = \sup_{Z=Z^T \geq 0} \left\{ -\operatorname{tr} F_0 Z : \operatorname{tr} F_i Z = c_i \right\}. \tag{12}$$

The dual problem itself can be reformulated as an LMI. It is easy to show $p^* \geq d^*$. If there exists a strict solution to either of the problems then it can also be shown that $p^* = d^*$. The primal-dual formulation offers a number of nice properties. One of them is that the interval of the optimal solution is given, since both an upper bound, $c^T x$, and a lower bound, $\operatorname{tr} F_0 Z$, are provided. Specifically, nonfeasibility can be shown using the primal-dual formulation. Also, computational advantages can be made in the solution of the LMI problem [12].

Example 1.2 We want to determine if

$$F(x) = \begin{bmatrix} x-4 & x-3 \\ x-3 & 2 \end{bmatrix} \succ 0, \tag{13}$$

is feasible or not. Introduce the auxiliary EVP/PDP as defined in (6) with $x_2 = \lambda$:

$$p^* = \inf_x \left\{ \begin{bmatrix} 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} : \tilde{F}(x) = \begin{bmatrix} -4 & -3 \\ -3 & 2 \end{bmatrix} + x_1 \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} + x_2 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \succ 0 \right\}.$$

Then (13) is feasible if $p^* < 0$. The dual problem is given by (12). We first solve Z with respect to $\operatorname{tr} \tilde{F}_1 Z = 0$ and $\operatorname{tr} \tilde{F}_2 Z = 1$. Then

$$d^* = \sup_{Z(\zeta) \geq 0} \left\{ -\operatorname{tr} \tilde{F}_0 Z(\zeta) = -2 - 6\zeta \right\}, \qquad \text{with} \qquad Z(\zeta) = \begin{bmatrix} -2\zeta & \zeta \\ \zeta & 1+2\zeta \end{bmatrix} \geq 0.$$

Since any $\zeta \in [-0.4, 0]$ satisfies $Z \geq 0$, we obtain $d^* = 0.4$ and infer that (13) is not feasible ($p^* \geq 0.4$). Note that only one Z is needed to show the infeasibility of (13). $\square$

1.3 Complexity

The best methods available for solving LMIs are efficient, even if they are more complex than most matrix manipulations, such as matrix inversions and solving Riccati equations. The LMI solvers based on interior point methods are iterative and solve a least squares problem in each iteration. Rather surprisingly, the number of iterations is practically almost independent of the size of the LMI problem. Typically, 5 to 50 iterations are needed. An upper bound on the number of iterations needed can be found but the typical performance is better than that bound. To improve efficiency further, the structure of the LMI can be exploited for reducing the computational effort to solve the least squares problem.

The primal-dual interior point method presented in [12] has a proved the worst-case complexity in terms of arithmetic operations of $\mathcal{O}(m^{2.75}L^{1.5})$, where m is the number of decision variables and L is the number of constraints. This result applies to a set of L Lyapunov inequalities. Typical performance is much better, for which a complexity of $\mathcal{O}(m^{2.1}L^{1.2})$ is reported.

1.4 Software Packages

A number of toolboxes are available for solving LMIs. One such toolbox is the LMI Control Toolbox [6] to be used within Matlab. Other toolboxes are the LMItool [7, 4] and sdpsol [3].

A few years ago, a Matlab interface called Yalmip [9] became available. Yalmip works as an high-level interface to some of the most popular LMI solvers. See `http://control.ee.ethz.ch/~joloef/yalmip.php` for more information.

1.5 Some Matrix Problems

1.5.1 Minimizing Matrix Norms

A simple problem that can be formulated as an LMI is to minimize the maximum singular value of a matrix that is an affine function of some parameters. Assume that $F \in \mathbb{R}^m \to \mathbb{R}^{n\times n}$ is an affine matrix function. Then the problem of minimizing $\bar{\sigma}(F(x))$ is equivalent to minimizing

$$\sqrt{\min_{\substack{x\in\mathbb{R}^m\\ \beta>0}} \{\beta : F^T(x)F(x) \leq \beta I\}}, \tag{14}$$

or

$$\min_{\substack{x\in\mathbb{R}^m\\ \gamma>0}} \left\{\gamma : \begin{bmatrix} -\gamma I & F^T(x) \\ F(x) & -\gamma I \end{bmatrix} \leq 0\right\}, \tag{15}$$

which are LMIs in (x, β) and (x, γ) respectively. The equivalences follow by observing that the eigenvalues of $F(x)^T F(x)$ are equal to the square of the singular values of $F(x)$. Specifically, $\bar{\sigma}(F(x)) \leq \gamma$ is equivalent to $F(x)^T F(x) \leq \gamma^2 I$. We then use the Schur complement formula followed by a diagonal pre and post scalings by $\mathrm{diag}[\gamma^{-1/2}I, \gamma^{1/2}I]$, to obtain (15).

If $F(x)$ is a vector then the maximum singular value is equal to the Euclidian norm, that is the square root of the sum of squares. Thus, least squares problems can be combined LMI constraints.

1.5.2 Minimizing Condition Number

When solving LMIs it is important to consider the numerical aspects of the problem in order to get a reliable solution. The problem we are focusing on here is generally of the type $F(P) \succ 0$ where F is a function of a matrix P. In order to find a reliable solution we can try to keep the condition number of P as low as possible [2]. The condition number is defined as the ratio of the largest and the smallest singular value. If we assume that P is a symmetric and positive definite matrix the singular values and the eigenvalues coincide. Thus the condition number of P is less than γ if and only if there exists a μ, such that

$$\mu I \prec P \prec \gamma\mu I.$$

Suppose that

$$F(x) = F_0 + \sum_{i=1}^{m} x_i F_i, \qquad P(x) = P_0 + \sum_{i=1}^{m} x_i P_i.$$

Defining new variables $\bar{x} = x/\mu$ and $\nu = 1/\mu$, we obtain an EVP. Minimize γ subject to

$$\nu F_0 + \sum_{i=1}^{m} \bar{x}_i F_i \succ 0, \qquad I \prec \nu P_0 + \sum_{i=1}^{m} \bar{x}_i P_i \prec \gamma I.$$

Another similar problem is to find a nonsingular scaling matrix T of a given structure, such that the condition number of $T^T PT$ is minimized. This problem is equivalent to minimizing γ subject to

$$\mu I \prec T^T PT \prec \gamma\mu I \quad \text{or} \quad \mu T^{-T}T^{-1} \prec P \prec \gamma\mu T^{-T}T^{-1}$$

with respect to T, $\mu > 0$ and γ. Introducing $W = \gamma\mu T^{-T}T^{-1}$ yields

$$W \prec P, \qquad \text{and} \qquad \gamma^{-1}P \prec W,$$

which is an EVP with respect to γ^{-1}.

1.5.3 Treating Complex-Valued LMIs

When using an LMI solver it usually only accepts real-valued problems. In order to handle complex-valued LMIs, we need to turn these into real-valued ones. This can be done by using the following observation. The field of complex numbers, $a + jb \in \mathbb{C}$, is isomorphic to the field of 2×2 matrices with the structure $\left[\begin{smallmatrix} a & b \\ -b & a \end{smallmatrix}\right] \in \mathbb{R}^{2\times 2}$. Thus, the product $P = X + jY$ of two complex matrices $M = A + jB$ and $N = C + jD$ can be computed using a real-valued matrix multiplication

$$\begin{bmatrix} X & Y \\ -Y & X \end{bmatrix} = \begin{bmatrix} A & B \\ -B & A \end{bmatrix} \begin{bmatrix} C & D \\ -D & C \end{bmatrix}. \tag{16}$$

Also, a Hermitian matrix $P = P^* = X + jY$ is positive definite if and only if

$$\begin{bmatrix} X & Y \\ -Y & X \end{bmatrix} \succ 0. \tag{17}$$

Since the dimensions of the matrices double, the complexity of a complex-valued LMI is significantly higher than a real-valued LMI of the same dimension. Thus we should always try to use as much of the structure of the problem as possible and, for instance, avoid to treat a real-valued LMI as a complex one.

2 Performance Bounds

Consider a dynamic system described by a differential equation

$$\dot{x} = f(x, w) \tag{18}$$

and a performance criterion, $J_w \in \mathbb{R}^n \to \mathbb{R}$:

$$J_w(x(t)) = \int_t^T g(x(\tau), w(\tau)) d\tau \tag{19}$$

where $x \in \mathbb{R} \to \mathbb{R}^n$ is the state vector as a function of time, $w \in \mathbb{R} \to \mathbb{R}^m$ is the input (or disturbance) vector and $g \in \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ is the cost function.

The time variable t can be included in the state-space vector x by having a state x_i such that $\dot{x}_i = 1$ and $x_i(t) = t$.

Theorem 1. *A strict upper bound V for the performance criterion J_w, such that $J_w(x) < V(x)$ for all w, can be established if there exists a continuously differentiable, positive definite Lyapunov or storage function (see [13]), V, that makes the Hamiltonian, H, negative for all x and w:*

$$H = g(x, w) + V_x(x) f(x, w) < 0, \ \forall x, w, \tag{20}$$

where $V_x = \frac{\partial}{\partial x} V$ denotes the partial derivative of V with respect to x.

Proof.

$$\begin{aligned}
J_w(x(t)) &= \int_t^T g(x, w) d\tau \\
&= \int_t^T g(x, w) d\tau + V(x(T)) - V(x(t)) + V(x(t)) - V(x(T)) \\
&\leq \int_t^T \Big(g(x, w) + \dot{V}(x) \Big) d\tau + V(x(t)) \\
&= \int_t^T \left(g(x, w) + V_x(x) f(x, w) \right) d\tau + V(x(t)) < V(x(t)).
\end{aligned}$$

□

This theorem can be modified to a nonstrict version by replacing $<$ with $\leq$.

Here we will use this inequality to provide conditions for stability and performance bounds on linear systems subject to nonlinear disturbances. We will use the $\mathcal{L}_2[t, T]$ norm as a performance criterion. We assume that $x(t) = 0$ and that $t = 0$ and $T = \infty$ if nothing otherwise is stated. However, the analysis is general and the $\mathcal{L}_2$ norm can easily be extended to any interval $[t, T]$.

3 Matrix Inequalities

3.1 Continuous Time

We will here study linear, stable systems subject to nonlinear uncertainties:

$$\begin{aligned}
\dot{x} &= Ax + Bw \\
z &= Cx + Dw,
\end{aligned} \tag{21}$$

where w is the disturbance input and z is the performance output.

The aim of this section is to give criteria for assuring upper bounds of the H_∞ or $\mathcal{L}_2$-induced norm from w to z for LTI system, i.e. to show that

$$\|z\|_2 < \gamma \|w\|_2,$$

or equivalently

$$\|z\|_2 - \gamma^2 \|w\|_2 = \int z^T(t)z(t) - \gamma^2 w^T(t)w(t)dt < 0.$$

For this problem the following cost function can be used

$$g(x, w) = \|z\|^2 - \gamma^2 \|w\|^2 = z^T z - \gamma^2 w^T w, \tag{22}$$

and a quadratic Lyapunov function is chosen

$$V(x) = x^T P x. \tag{23}$$

To assure internal stability, it is assumed that the Lyapunov matrix P is symmetric and positive definite ($P \succ 0$), that is $x^T P x > 0, \forall x \neq 0$. If $x(0) = 0$ the $\mathcal{L}_2$-induced norm from w to z is less than one if the Hamiltonian for (21) and (22) is negative for all x:

$$\begin{aligned} H &= \dot{V} + g(x, w) \\ &= \dot{x}^T P x + x^T P \dot{x} + z^T z - w^T w \\ &= x^T P(Ax + Bw) + (Ax + Bw)^T P x + (Cx + Dw)^T (Cx + Dw) - \gamma^2 w^T w. \end{aligned} \tag{24}$$

In order to assure that $\|z\|_2 < \gamma \|w\|_2$ then $H < 0$ must hold for all x and w.

3.2 The Riccati Inequality

One way of arriving at the related Riccati inequality is by completing the squares in (24). First observe that by letting $x = 0$ it can be inferred that $D^T D < I$ and thus $R = \gamma^2 I - D^T D$ is invertible. Then

$$\begin{aligned} H = x^T &\left(A^T P + PA + (B^T P + D^T C)^T R^{-1}(B^T P + D^T C) + C^T C\right) x \\ &- \left(w - R^{-1}(B^T P + D^T C)x\right)^T R \left(w - R^{-1}(B^T P + D^T C)x\right) \\ \leq x^T &\left(A^T P + PA + (B^T P + D^T C)^T R^{-1}(B^T P + D^T C) + C^T C\right) x. \end{aligned}$$

Equality is obtained for

$$w = R^{-1}(B^T P + D^T C)x, \tag{25}$$

which can be interpreted as the worst-case disturbance.

3.3 Linear Matrix Inequalities (LMIs)

Instead of completing the squares, the Hamiltonian (24) can be rewritten into

$$H = \begin{bmatrix} x \\ w \end{bmatrix}^T \begin{bmatrix} PA + A^T P + C^T C & PB + C^T D \\ B^T P + D^T C & D^T D - \gamma^2 I \end{bmatrix} \begin{bmatrix} x \\ w \end{bmatrix} < 0, \tag{26}$$

which shall hold for all nonzero x, w. This implies that

$$\begin{bmatrix} PA + A^T P + C^T C & PB + C^T D \\ B^T P + D^T C & D^T D - \gamma^2 I \end{bmatrix} \prec 0 \tag{27}$$

which is a linear matrix inequality (LMI) in P, for given (A, B, C, D). This implies that the set of P satisfying the LMI is convex, which substantially simplifies the search for P.

3.3.1 Schur Complements

The equivalence between the Riccati inequality and the LMI can be seen by the following well-known fact:

Lemma 1 (Schur Complement)**.** *Suppose R and S are Hermitian, i.e. $R = R^*$ and $S = S^*$. Then, the following conditions are equivalent:*

$$R \prec 0, \qquad S - G^T R^{-1} G \prec 0; \tag{28}$$

and

$$\begin{bmatrix} S & G^T \\ G & R \end{bmatrix} \prec 0. \tag{29}$$

Proof. Post-multiply (29) by the nonsingular $\left[\begin{smallmatrix} I & 0 \\ -R^{-1}G & I \end{smallmatrix}\right]$ and pre-multiply by its transpose:

$$\begin{bmatrix} I & -G^T R^{-1} \\ 0 & I \end{bmatrix} \begin{bmatrix} S & G^T \\ G & R \end{bmatrix} \begin{bmatrix} I & 0 \\ -R^{-1}G & I \end{bmatrix} = \begin{bmatrix} S - G^T R^{-1} G & 0 \\ 0 & R \end{bmatrix} \prec 0,$$

which is equivalent to the conditions in (28). □

The Schur complement result can be generalized to nonstrict inequalities. Using Schur complements we can infer that if a matrix is positive definite then an arbitrary diagonal square sub-block is also positive definite. For instance, if any diagonal element p_{ii} of a matrix P is negative or zero the matrix P is not positive definite.

3.3.2 Congruence Transformations

When proving the Schur complements a so called *congruence transformation* is employed. Let U be a nonsingular matrix, then

$$F \succ 0 \qquad \text{and} \qquad U^T F U \succ 0, \tag{30}$$

are equivalent statements. The inequality can be replaced by equality ($=$) or nonstrict inequality ($\geq$).

Applying the Schur complement on (27), which can be written as

$$\begin{bmatrix} PA + A^T P & PB \\ B^T P & -\gamma^2 I \end{bmatrix} + \begin{bmatrix} C^T \\ D^T \end{bmatrix} I \begin{bmatrix} C & D \end{bmatrix} \prec 0, \tag{31}$$

yields

$$\begin{bmatrix} PA + A^T P & PB & C^T \\ B^T P & -\gamma^2 I & D^T \\ C & D & -I \end{bmatrix} \prec 0. \tag{32}$$

By scaling P we can rewrite this into

$$\begin{bmatrix} PA + A^T P & PB & C^T \\ B^T P & -\gamma I & D^T \\ C & D & -\gamma I \end{bmatrix} \prec 0. \tag{33}$$

3.3.3 Equivalent Matrix Inequalities

We have shown the equivalence between the Riccati inequality and the corresponding LMI. They have different virtues, which will be employed when convenient. One of the objectives for choosing the LMI is that it in many cases provides a simple tool for showing that the set of solutions is convex.

By repeating the Schur complement we arrive at the following equivalent conditions:

(i) $$\begin{cases} \bar{\sigma}(D) < \gamma, \\ A^T P + PA + (B^T P + D^T C)^T (\gamma^2 I - D^T D)^{-1} (B^T P + D^T C) + C^T C \prec 0; \end{cases}$$

(ii) $$\begin{bmatrix} PA + A^T P + C^T C & PB + C^T D \\ B^T P + D^T C & D^T D - \gamma^2 I \end{bmatrix} \prec 0;$$

(iii) $$\begin{bmatrix} PA + A^T P & PB & C^T \\ B^T P & -\gamma I & D^T \\ C & D & -\gamma I \end{bmatrix} \prec 0.$$

All but the first one of these inequalities are linear in P if (A, B, C, D) are kept fixed. The last one of these inequalities (iii) is linear in (A, B, C, D) for a given P, from which we conclude that the set of system matrices satisfying the Riccati inequality or equivalently the LMI is convex. The bounded real lemma states an extension of these results.

Lemma 2 (Bounded Real Lemma [11, 5]). *The following statements are equivalent*

(i) $\|G\|_\infty < \gamma$ *and* A *stable with* $G(s) = D + C(sI - A)^{-1}B$*;*

(ii) there exists a solution $P \succ 0$ *to the LMI*

$$\begin{bmatrix} PA + A^T P & PB & C^T \\ B^T P & -\gamma I & D^T \\ C & D & -\gamma I \end{bmatrix} \prec 0. \tag{34}$$

Proof. We have already shown that (ii) $\Rightarrow$ (i). By using the Kalman-Yakubovich-Popov lemma, the opposite direction can proved. $\square$

4 The Elimination Lemma

Let $U^\perp$ denote any matrix of maximum rank that satisfies $U^\perp U = 0$. To be more precise: $\operatorname{range} U^{\perp\mathrm{T}} = \operatorname{null} U^{\mathrm{T}}$.

Lemma 3. *Let Q, U and V be given matrices. Then*

$$Q + UKV^{\mathrm{T}} + VK^{\mathrm{T}}U^{\mathrm{T}} \prec 0, \tag{35}$$

has a solution K if and only if

(i) $V^{\perp}QV^{\perp\mathrm{T}} \prec 0$,

(ii) $U^{\perp}QU^{\perp\mathrm{T}} \prec 0$,

If $V^{\perp}$ or $U^{\perp}$ does not exist the corresponding condition is assumed to be satisfied.

Proof. We show the lemma by construction. Make a congruence transformation of (35) using a nonsingular $T = \begin{bmatrix} T_1 & T_2 & T_3 & T_4 \end{bmatrix}$ where the transformed rows (and columns) correspond to the following disjunct spaces

(1) $\operatorname{range} T_1 = \operatorname{null} U^{\mathrm{T}} \cap \operatorname{null} V^{\mathrm{T}}$

(2) $\operatorname{range} T_2 = \operatorname{range} U \cap \operatorname{null} V^{\mathrm{T}}$

(3) $\operatorname{range} T_3 = \operatorname{null} U^{\mathrm{T}} \cap \operatorname{range} V$

(4) $\operatorname{range} T_4 = \operatorname{range} U \cap \operatorname{range} V$

Then, (35), is equivalent to

$$T^{\mathrm{T}}(Q+\ldots)T = \left[\begin{array}{ccc|c} \tilde{Q}_{11} & \tilde{Q}_{12} & \tilde{Q}_{13} & \tilde{Q}_{14} \\ \tilde{Q}_{12}^{\mathrm{T}} & \tilde{Q}_{22} & \tilde{Q}_{23}+\tilde{K}_{23} & \tilde{Q}_{24}+\tilde{K}_{24} \\ \tilde{Q}_{13}^{\mathrm{T}} & \tilde{Q}_{23}^{\mathrm{T}}+\tilde{K}_{23}^{\mathrm{T}} & \tilde{Q}_{33} & \tilde{Q}_{34}+\tilde{K}_{43}^{\mathrm{T}} \\ \hline \tilde{Q}_{14}^{\mathrm{T}} & \tilde{Q}_{24}^{\mathrm{T}}+\tilde{K}_{24}^{\mathrm{T}} & \tilde{Q}_{34}^{\mathrm{T}}+\tilde{K}_{43} & \tilde{Q}_{44}+\tilde{K}_{44}+\tilde{K}_{44}^{\mathrm{T}} \end{array}\right] \prec 0 \tag{36}$$

where

$$\begin{bmatrix} \tilde{K}_{23} & \tilde{K}_{24} \\ \tilde{K}_{43} & \tilde{K}_{44} \end{bmatrix} = \begin{bmatrix} T_2 & T_3 \end{bmatrix}^{\mathrm{T}} UKV^{\mathrm{T}} \begin{bmatrix} T_3 & T_4 \end{bmatrix} \tag{37}$$

can be chosen freely.

We first consider the the sub-matrix of $\tilde{Q}$ that comprises the first three rows and columns. Eliminating the first row and column by Schur complement yields $\tilde{Q}_{11} \prec 0$ and

$$\begin{bmatrix} \tilde{Q}_{22} - \tilde{Q}_{12}^{\mathrm{T}}\tilde{Q}_{11}^{-1}\tilde{Q}_{12} & \tilde{Q}_{23} - \tilde{Q}_{12}^{\mathrm{T}}\tilde{Q}_{11}^{-1}\tilde{Q}_{13} + \tilde{K}_{23} \\ \tilde{Q}_{23}^{\mathrm{T}} - \tilde{Q}_{13}^{\mathrm{T}}\tilde{Q}_{11}^{-1}\tilde{Q}_{12} + \tilde{K}_{23}^{\mathrm{T}} & \tilde{Q}_{33} - \tilde{Q}_{13}^{\mathrm{T}}\tilde{Q}_{11}^{-1}\tilde{Q}_{13} \end{bmatrix} \prec 0 \tag{38}$$

We can choose $\tilde{K}_{23} = -\tilde{Q}_{23} + \tilde{Q}_{12}^{\mathrm{T}}\tilde{Q}_{11}^{-1}\tilde{Q}_{13}$, which yields $\tilde{Q}_{22} - \tilde{Q}_{12}^{\mathrm{T}}\tilde{Q}_{11}^{-1}\tilde{Q}_{12} \prec 0$ and $\tilde{Q}_{33} - \tilde{Q}_{13}^{\mathrm{T}}\tilde{Q}_{11}^{-1}\tilde{Q}_{13} \prec 0$ as remaining conditions together with $\tilde{Q}_{11} \prec 0$. These are equivalent to

$$\begin{bmatrix} \tilde{Q}_{11} & \tilde{Q}_{12} \\ \tilde{Q}_{12}^{\mathrm{T}} & \tilde{Q}_{22} \end{bmatrix} \prec 0 \qquad \text{and} \qquad \begin{bmatrix} \tilde{Q}_{11} & \tilde{Q}_{13} \\ \tilde{Q}_{13}^{\mathrm{T}} & \tilde{Q}_{33} \end{bmatrix} \prec 0 \tag{39}$$

which in turn is equivalent to condition (i) and (ii), respectively. Finally, including the fourth row and column of (36), we can always find a constant $\tilde{K}_{44} = -\sigma I$ provided conditions (i) and (ii) hold, by choosing σ large enough. □

5 H_∞ Synthesis using LMIs

The problem addressed here is the following. Suppose we are given a linear time-invariant (LTI) plant, G, with state-space realization

$$\begin{aligned} \dot{x} &= Ax + B_1 w + B_2 u \\ z &= C_1 x + D_{11} w + D_{12} u \\ y &= C_2 x + D_{21} w + D_{22} u \end{aligned} \tag{40}$$

where $A \in \mathbb{R}^{n \times n}$, $D_{11} \in \mathbb{R}^{p_1 \times m_1}$ and $D_{22} \in \mathbb{R}^{p_2 \times m_2}$ define the problem dimension.

From now on we assume that D_{22} is zero. If this is not the case we can find a controller, $\bar{K}$, for a modified G in which D_{22} is set to zero. Then the controller for $D_{22} \neq 0$ is $K = \bar{K}(I + D_{22}\bar{K})^{-1}$. Hence there is now loss of generality in assuming $D_{22} = 0$. See also [14, Section 17.2, page 454].

The output-feedback control problem consists of finding a dynamic controller with state-space equations

$$\begin{aligned} \dot{x}_K &= K_A x + K_B y \\ u &= K_C x + K_D y \end{aligned} \tag{41}$$

where $K_A \in \mathbb{R}^{r \times r}$ that ensures internal stability and a guaranteed performance bound, γ. The performance bound is defined as the H_∞ norm of the closed loop system from disturbance input signal, w, to the performance output, z.

We use the index K to denote the state-space matrices of the closed loop system. If we assume that $D_{22} = 0$, we can write the closed loop system as

$$\begin{bmatrix} A_K & B_K \\ C_K & D_K \end{bmatrix} = \left[\begin{array}{cc|c} A & 0 & B_1 \\ 0 & 0 & 0 \\ \hline C_1 & 0 & D_{11} \end{array}\right] + \left[\begin{array}{cc} B_2 & 0 \\ 0 & I \\ \hline D_{12} & 0 \end{array}\right] \begin{bmatrix} K_D & K_C \\ K_B & K_A \end{bmatrix} \left[\begin{array}{cc|c} C_2 & 0 & D_{21} \\ 0 & I & 0 \end{array}\right] \tag{42}$$

where $A_K \in \mathbb{R}^{(n+r) \times (n+r)}$. The closed loop system is internally stable and has an H_∞ norm of γ if there exists a symmetric $P = P^{\mathrm{T}} \succ 0$ such that Lemma 2 holds or, equivalently,

$$\begin{aligned} &\begin{bmatrix} PA_K + A_K^{\mathrm{T}} P & PB_K & C_K^{\mathrm{T}} \\ B_K^{\mathrm{T}} P & -\gamma I & D_K^{\mathrm{T}} \\ C_K & D_K & -\gamma I \end{bmatrix} \\ &= \begin{bmatrix} 0 & 0 & 0 \\ 0 & -\gamma I & 0 \\ 0 & 0 & -\gamma I \end{bmatrix} + \begin{bmatrix} P & 0 \\ 0 & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} A_K & B_K \\ C_K & D_K \end{bmatrix} \begin{bmatrix} I & 0 & 0 \\ 0 & I & 0 \end{bmatrix} + \begin{bmatrix} \end{bmatrix}^{\mathrm{T}} \prec 0 \end{aligned} \tag{43}$$

Inserting this in (43) together with

$$P = \begin{bmatrix} X & N \\ N^{\mathrm{T}} & L \end{bmatrix} = \begin{bmatrix} Y & M \\ M^{\mathrm{T}} & * \end{bmatrix}^{-1} \succ 0. \tag{44}$$

where $N, M \in \mathbb{R}^{n\times r}$ yields

$$\begin{bmatrix} XA + A^{\mathrm{T}}X & A^{\mathrm{T}}N & XB_1 & C_1^{\mathrm{T}} \\ N^{\mathrm{T}}A & 0 & N^{\mathrm{T}}B_1 & 0 \\ B_1^{\mathrm{T}}X & B_1^{\mathrm{T}}N & -\gamma I & D_{11}^{\mathrm{T}} \\ C_1 & 0 & D_{11} & -\gamma I \end{bmatrix}$$

$$+ \begin{bmatrix} XB_2 & N \\ N^{\mathrm{T}}B_2 & L \\ 0 & 0 \\ D_{12} & 0 \end{bmatrix} \begin{bmatrix} K_D & K_C \\ K_B & K_A \end{bmatrix} \begin{bmatrix} C_2 & 0 & D_{21} & 0 \\ 0 & I & 0 & 0 \end{bmatrix} + \begin{bmatrix} \quad \end{bmatrix}^{\mathrm{T}} \prec 0 \tag{45}$$

Using Lemma 3, the existence of $K = \begin{bmatrix} K_D & K_C \\ K_B & K_A \end{bmatrix}$ is equivalent to

$$\begin{bmatrix} \mathcal{N}_X & 0 \\ 0 & I \end{bmatrix}^{\mathrm{T}} \left[\begin{array}{cc|c} XA + A^{\mathrm{T}}X & XB_1 & C_1^{\mathrm{T}} \\ B_1^{\mathrm{T}}X & -\gamma I & D_{11}^{\mathrm{T}} \\ \hline C_1 & D_{11} & -\gamma I \end{array}\right] \begin{bmatrix} \mathcal{N}_X & 0 \\ 0 & I \end{bmatrix} \prec 0 \tag{46}$$

and

$$\begin{bmatrix} \mathcal{N}_Y & 0 \\ 0 & I \end{bmatrix}^{\mathrm{T}} \left[\begin{array}{cc|c} AY + YA^{\mathrm{T}} & YC_1^{\mathrm{T}} & B_1 \\ YC_1 & -\gamma I & D_{11} \\ \hline B_1^{\mathrm{T}} & D_{11}^{\mathrm{T}} & -\gamma I \end{array}\right] \begin{bmatrix} \mathcal{N}_Y & 0 \\ 0 & I \end{bmatrix} \prec 0 \tag{47}$$

where $\mathcal{N}_X$ and $\mathcal{N}_Y$ designate any bases of the null spaces of $\begin{bmatrix} C_2 & D_{21} \end{bmatrix}$ and $\begin{bmatrix} B_2^{\mathrm{T}} & D_{12}^{\mathrm{T}} \end{bmatrix}$, respectively.

For showing (47) we have used

$$\left[\begin{array}{cc} XB_2 & N \\ N^{\mathrm{T}}B_2 & L \\ \hline 0 & 0 \\ D_{12} & 0 \end{array}\right]^{\perp} = \left[\begin{array}{cc} B_2 & 0 \\ 0 & I \\ \hline 0 & 0 \\ D_{12} & 0 \end{array}\right]^{\perp} \begin{bmatrix} P^{-1} & 0 \\ 0 & I \end{bmatrix} = \begin{bmatrix} \mathcal{N}_Y & 0 \\ 0 & I \end{bmatrix}^{\mathrm{T}} \left[\begin{array}{cc|cc} Y & M & 0 & 0 \\ 0 & 0 & 0 & I \\ \hline 0 & 0 & I & 0 \end{array}\right]$$

A third LMI is required to link X and Y according to Lemma 4 below

$$\begin{bmatrix} X & I \\ I & Y \end{bmatrix} \geq 0. \tag{48}$$

Note that (48) is equivalent to $X - Y^{-1} \geq 0$.

Lemma 4. *Suppose $X = X^{\mathrm{T}} \in \mathbb{R}^{n\times n}$ and $Y = Y^{\mathrm{T}} \in \mathbb{R}^{n\times n}$. Let r be a positive integer. The following statements are equivalent:*

(i)

$$\begin{bmatrix} X & I \\ I & Y \end{bmatrix} \geq 0 \qquad \textit{and} \qquad \operatorname{rank}(X - Y^{-1}) \leq r \tag{49}$$

(ii) There exists $P = P \in \mathbb{R}^{(n+r)\times(n+r)}$ such that

$$P = \begin{bmatrix} X & N \\ N^{\mathrm{T}} & L \end{bmatrix} = \begin{bmatrix} Y & M \\ M^{\mathrm{T}} & * \end{bmatrix}^{-1} \succ 0 \tag{50}$$

Proof. (i) $\Rightarrow$ (ii): Factor $X - Y^{-1} = NN^{\mathrm{T}}$, where $N \in \mathbb{R}^{n\times r}$ and let $M = -YN$ and $L = I$. Then

$$P = \begin{bmatrix} X & N \\ N^{\mathrm{T}} & I \end{bmatrix} = \begin{bmatrix} Y & -YN \\ -N^{\mathrm{T}}Y & I + N^{\mathrm{T}}YN \end{bmatrix}^{-1} \succ 0 \tag{51}$$

(ii) $\Rightarrow$ (i): Using the Schur formulas for matrix inversion gives that $Y^{-1} = X - NL^{-1}N^{\mathrm{T}}$. Hence, $X - Y^{-1} = NL^{-1}N^{\mathrm{T}} \geq 0$, and indeed, $\operatorname{rank}(X - Y^{-1}) = \operatorname{rank}(NL^{-1}N^{\mathrm{T}}) \leq r$. $\square$

We conclude by stating the following theorem:

Theorem 2. *The following statements are equivalent.*

(i) There exists an controller of the system (40) *or order r that achieves closed-loop stability with H_∞ norm of γ.*

(ii) There exists $X = X^{\mathrm{T}}, Y = Y^{\mathrm{T}} \in \mathbb{R}^{n\times n}$, such that (46)*,* (47) *and* (49) *hold.*

In order to reduce $\operatorname{rank}(X - Y^{-1})$ we can instead minimize

$$\operatorname{rank} \begin{bmatrix} X & I \\ I & Y \end{bmatrix} = n + \operatorname{rank}(X - Y^{-1}). \tag{52}$$

One way to try to reduce the rank is to minimize $\operatorname{tr} X + \operatorname{tr} Y$, even if this does not guarantee to find the minimum order controller for a given γ.

References

[1] S. Boyd and L. El Ghaoui. Method of centers for minimizing generalized eigenvalues. *Linear Algebra and Applications, special issue on Numerical Linear Algebra Methods in Control, Signals and Systems*, 188-189:63–111, July 1993.

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

[3] S. Boyd and S. Wu. *sdpsol – A Parser/Solver Semidefinite Programs with Matrix Structure – User's Guide, Version Alpha.* Stanford University, 1995.

[4] F. Delebecque, R. Nikoukhah, and L. El Ghaoui. *LMITOOL: A Package for LMI Optimization in Scilab – User's Guide*, 1995.

[5] P. Gahinet and P. Apkarian. An LMI-based parametrization of all H_∞ controllers with applications. In *IEEE Proceedings of the 32nd Conference on Decision and Control*, volume 1, pages 656–661, San Antonio, Texas, December 1993.

[6] P. Gahinet, A. Nemirovski, A. J. Laub, and M. Chilali. The LMI control toolbox. In *Proceedings of the 3rd European Control Conference*, volume 4, part 1, pages 3206–3211, Rome, Italy, September 1995.

[7] L. El Ghaoui, F. Delebecque, and R. Nikoukhah. *LMITOOL: A User-Friendly Interface for LMI Optimization – User's Guide, Beta version*, 1995.

[8] B. Lieu and P. Huard. La méthode des centres dans un espace topologique. *Numerische Mathematik*, 8:56–67, 1965.

[9] J. Löfberg. YALMIP : A toolbox for modeling and optimization in MATLAB. In *Proceedings of the CACSD Conference*, Taipei, Taiwan, 2004. Available from `http://control.ee.ethz.ch/~joloef/yalmip.php`.

[10] Yu. Nesterov and A. Nemirovsky. Interior point polynomial methods in convex programming: Theory and application. *SIAM*, 1993.

[11] C. Scherer. *The Riccati Inequality and State-Space H_∞-Optimal Control*. PhD thesis, Universitat Wurtzburg, Wurtzburg, Germany, 1990.

[12] L. Vandenberghe and S. Boyd. Positive-definite programming. 1994. submitted to SIAM review.

[13] J. C. Willems. The generation of Lyapunov functions for input-output stable systems. *SIAM Journal of Control*, 9:105–133, 1971.

[14] K. Zhou, J. C. Doyle, and K. Glover. *Robust and Optimal Control*. Prentice Hall, 1995.