12

LINEAR MATRIX INEQUALITIES

This chapter gives an introduction to the use of linear matrix inequalities (LMIs) in control. LMI problems are defined and tools described for transforming such problems into suitable formats for solution. The chapter ends with a case study on anti-windup compensator synthesis.

12.1 Introduction to LMI problems

Linear matrix inequalities are matrix inequalities which are linear (or affine) in a set of *matrix* variables. Many problems in control theory can be stated in terms of linear matrix inequalities and their existence can be traced back over 100 years to the work of Lyapunov. However, until relatively recently, there were few (if any) routines available to solve LMIs numerically. During the past 10-15 years, the development of sophisticated numerical routines has made it possible to solve LMIs in a reasonably efficient manner. These routines exploit the convexity of LMI problems in order to obtain reliable numerical calculations.

From a control engineering perspective, one of the main attractions of LMIs is that they can be used to solve problems which involve several matrix variables, and, moreover, different structures can be imposed on these matrix variables. Another attractive feature of LMI methods is that they are flexible, so it is often relatively straightforward to pose a variety of problems as LMI problems, amenable to LMI methods. Furthermore, in many cases the use of LMIs can remove restrictions associated with conventional methods and aid their extension to more general scenarios. Often LMI methods can be applied in instances where conventional methods either fail, or struggle to find a solution.

Another advantage of LMIs, at least in a pedagogical sense, is that they are able to *unite* many previous results in a common framework. This can enable one to obtain additional insight into established areas.

The contents of this chapter are available as a University of Leicester technical report (Turner et al., 2004) which draws heavily on the material contained within the book by Boyd et al. (1994) and the Matlab LMI control toolbox by Gahinet et al. (1995). In order to convey the main points, the presentation is somewhat condensed and the interested reader should consult Boyd et al. (1994) for a more complete exposure to LMIs.

12.1.1 Fundamental LMI properties

A notion central to the understanding of matrix inequalities is *definiteness*. In particular, a matrix Q is defined to be *positive definite* if

$$x^T Q x > 0 \quad \forall x \neq 0 \tag{12.1}$$

Likewise, Q is said to be *positive semi-definite* if

$$x^T Q x \geq 0 \quad \forall x \tag{12.2}$$

It is common practice to write $Q > 0$ $(Q \geq 0)$ to indicate that it is positive (semi-) definite. A positive definite real matrix has three key features: it is square, symmetric and all of its eigenvalues are real and positive[1]. A positive semi-definite real matrix shares the first two attributes, but the last is relaxed to the requirement that all of its eigenvalues are positive or zero. A matrix $P = -Q$ is said to be *negative (semi) definite* if Q is postive (semi) definite. To indicate negative (semi) definiteness we write $P < 0$ $(P \leq 0)$.

The basic structure of an LMI is

$$F(x) = F_0 + \sum_{i=1}^{m} x_i F_i > 0 \tag{12.3}$$

where $x \in \mathbb{R}^m$ is a variable and F_0, F_i are given constant symmetric real matrices.

The basic LMI problem - the *feasibility* problem - is to find x such that inequality (12.3) holds. Note that $F(x) > 0$ in (12.3) describes an *affine* relationship in terms of the variable x. Normally the variable x, which we are interested in, is composed of one or many matrices whose columns have been ‘stacked’ as a vector. That is,

$$F(x) = F(X_1, X_2, \ldots, X_n) \tag{12.4}$$

where $X_i \in \mathbb{R}^{q_i \times p_i}$ is a matrix, $\sum_{i=1}^{n} q_i \times p_i = m$, and the columns of all the matrix variables are stacked up to form a single vector variable.

Hence, from now on, we will consider functions of the form

$$\begin{aligned} F(X_1, X_2, \ldots, X_n) &= F_0 + G_1 X_1 H_1 + G_2 X_2 H_2 + \ldots && (12.5) \\ &= F_0 + \sum_{i=1}^{n} G_i X_i H_i > 0 && (12.6) \end{aligned}$$

[1] Technically, a matrix is called positive-definite, if and only if all the eigenvalues of its’ symmetric part $(Q + Q^T)$ are positive, but this has little practical significance.

where F_0, G_i, H_i are given matrices and the X_i are the matrix variables which we seek.

12.1.2 Systems of LMIs

In general, we are frequently faced with LMI constraints of the form

$$
\begin{aligned}
F_1(X_1, \ldots, X_n) &> 0 \qquad (12.7)\\
&\vdots \\
F_p(X_1, \ldots, X_n) &> 0 \qquad (12.8)
\end{aligned}
$$

where

$$F_j(X_1, \ldots, X_n) = F_{0j} + \sum_{i=1}^{n} G_{ij} X_i H_{ij} \qquad (12.9)$$

However, it is easily seen that, by defining $\widetilde{F}_0, \widetilde{G}_i, \widetilde{H}_i, \widetilde{X}_i$ as

$$
\begin{aligned}
\widetilde{F}_0 &= \operatorname{diag}(F_{01}, \ldots, F_{0p}) \qquad (12.10)\\
\widetilde{G}_i &= \operatorname{diag}(G_{i1}, \ldots, G_{ip}) \qquad (12.11)\\
\widetilde{H}_i &= \operatorname{diag}(H_{i1}, \ldots, H_{ip}) \qquad (12.12)\\
\widetilde{X}_i &= \operatorname{diag}(X_i, \ldots X_i) \qquad (12.13)
\end{aligned}
$$

we actually have the inequality

$$F_{\text{big}}(X_1, \ldots, X_n) \triangleq \widetilde{F}_0 + \sum_{i=1}^{n} \widetilde{G}_i \widetilde{X}_i \widetilde{H}_i > 0 \qquad (12.14)$$

That is, we can represent a (big) system of LMIs as a single LMI. Therefore, we do not distinguish a single LMI from a system of LMIs; they are the same mathematical entity. We may also encounter systems of LMIs of the form:

$$
\begin{aligned}
F_1(X_1, \ldots, X_n) &> 0 \qquad (12.15)\\
F_2(X_1, \ldots, X_n) &> F_3(X_1, \ldots, X_n) \qquad (12.16)
\end{aligned}
$$

Again, it is easy to see that this can be written in the same form as inequality (12.14) above. For the remainder of the report we do not distinguish between LMIs which can be written as above, or those which are in the more generic form of inequality (12.14).

12.1.3 Types of LMI Problems

The term ‘LMI problem’ is rather vague and in fact there are several sub-groups of LMI problems including LMI feasibility problems, linear objective minimization problems and generalized eigenvalue problems. These will be described below in the same way that they are separated in the Matlab LMI toolbox. Note that by ‘LMI problem’ we normally mean solving an optimization problem or an eigenvalue problem with LMI constraints.

LMI feasibility problems

These are simply problems for which we seek a *feasible* solution $\{X_1, \ldots, X_n\}$ such that

$$F(X_1, \ldots, X_n) > 0 \tag{12.17}$$

We are not interested in the optimality of the solution, only in finding a solution, which may not be unique.

Example 12.1 **Determining stability of a linear system.** *Consider an autonomous linear system*

$$\dot{x} = Ax \tag{12.18}$$

then the Lyapunov LMI problem for proving stability of this system $(\mathrm{Re}\{\lambda_i(A)\} < 0, \forall i)$ *is to find a* $P > 0$ *such that (see, e.g., Boyd et al. (1994))*

$$A^H P + PA < 0 \tag{12.19}$$

This is an LMI feasibility problem in $P > 0$*. However, given any* $P > 0$ *which satisfies this, it is obvious that any matrix from the set*

$$\mathcal{P} = \{\beta P : \text{ scalar } \beta > 0\} \tag{12.20}$$

also solves the problem. In fact, as will be seen later, the matrix P *forms part of a Lyapunov function for the linear system.*

Linear objective minimization problems

These problems are also called eigenvalue problems. They involve the minimization (or maximization) of some *linear scalar* function, $\alpha(.)$, of the matrix variables, subject to LMI constraints:

$$\min \alpha(X_1, \ldots, X_n) \tag{12.21}$$

$$\text{s.t.} \qquad F(X_1, \ldots, X_n) > 0 \tag{12.22}$$

where we have used the abbreviation ‘s.t.’ to mean ‘such that’. In this case, we are therefore trying to optimize some quantity whilst ensuring some LMI constraints are satisfied.

Example 12.2 **Calculating the $\mathcal{H}_\infty$norm of a linear system.** *Consider a linear system*

$$\dot{x} = Ax + Bw \tag{12.23}$$
$$z = Cx + Dw \tag{12.24}$$

then the problem of finding the $\mathcal{H}_\infty$norm of the transfer function matrix T_{zw} from w to z is equivalent to the following optimization procedure (see, e.g., Gahinet and Apkarian (1994)):

$$\min \gamma \tag{12.25}$$
$$\text{s.t.} \quad \begin{bmatrix} A^TP + PA & PB & C^T \\ B^TP & -\gamma I & D^T \\ C & D & -\gamma I \end{bmatrix} < 0 \tag{12.26}$$

Note that although $\gamma > 0$ is unique, the uniqueness of $P > 0$ is, in general, not guaranteed.

Generalized eigenvalue problems

The generalized eigenvalue problem, or GEVP, is slightly different to the preceding problem in the sense that the objective of the optimization problem is not actually convex, but *quasi-convex*. However, the methods used to solve such problems are similar. Specifically a GEVP is formulated as

$$\min \lambda \tag{12.27}$$
$$\text{s.t.} \quad F_1(X_1, \ldots, X_n) - \lambda F_2(X_1, \ldots, X_n) < 0 \tag{12.28}$$
$$F_2(X_1, \ldots, X_n) < 0 \tag{12.29}$$
$$\vdots$$
$$F_n(X_1, \ldots, X_n) < 0 \tag{12.30}$$

The first two lines are equivalent to minimizing the largest 'generalized' eigenvalue of the matrix pencil $F_1(X_1, \ldots, X_n) - \lambda F_2(X_1, \ldots, X_n)$. In some cases, a GEVP problem can be reduced to a linear objective minimization problem, through an appropriate change of variables.

Example 12.3 **Bounding the decay rate of a linear system.** *Given a stable linear system $\dot{x} = Ax$, the decay rate is the largest α such that*

$$\|x(t)\| \leq \exp(-\alpha t)\beta\|x(0)\| \quad \forall x(t) \tag{12.31}$$

where β is a constant. If we choose $V(x) = x^TPx > 0$ as a Lyapunov function for the system and ensure that $\dot{V}(x) \leq -2\alpha V(x)$ it is easily shown that the system will have a decay rate of at least α. Hence, the problem of finding the decay rate could be posed as the optimization problem

$$\min -\alpha \tag{12.32}$$
$$\text{s.t.} \quad A^TP + PA + 2\alpha P \leq 0 \tag{12.33}$$

This problem is a GEVP with the functions

$$F_1(P) = A^T P + PA \tag{12.34}$$
$$F_2(P) = -2P \tag{12.35}$$

Example 12.4 Calculating upper bound on μ. *Consider the problem of calculating upper bound on the structured singular value, μ given as*

$$\mu(M) \leq \min_{D \in \mathcal{D}} \bar{\sigma}(DMD^{-1}) \tag{12.36}$$

Due to the presence of the inverse term, the optimization problem is difficult to solve in its present form; however, it can be transformed into an equivalent LMI problem.

Let $\mathcal{D}$ to be the set of matrices D which commute with Δ (i.e. satisfy $D\Delta = \Delta D$). There exists $D \in \mathcal{D}$ such that $\bar{\sigma}(DMD^{-1}) < \gamma$ if and only if

$$(DMD^{-1})^H(DMD^{-1}) < \gamma^2 I \quad \text{for some} \quad D \in \mathcal{D} \tag{12.37}$$
$$M^H D^H DM < \gamma^2 D^H D \quad \text{for some} \quad D \in \mathcal{D} \tag{12.38}$$
$$M^H PM < \gamma^2 P \quad \text{for some} \quad P = D^H D \tag{12.39}$$

The problem in (12.39) is a GEVP with $P > 0$.

12.2 Tricks in LMI problems

Although many control problems can be cast as LMI problems, a substantial number of these need to be manipulated before they are in a suitable LMI problem format. Fortunately, there are a number of common tools or 'tricks' which can be used to transform problems into suitable LMI forms. Some of the more useful ones are described below.

12.2.1 Change of variables

Many control problems can be posed in the form of a set of nonlinear matrix inequalities; that is the inequalities are nonlinear in the matrix variables we seek. However, by defining new variables it is sometimes possible to 'linearise' the nonlinear inequalities, hence making them solvable by LMI methods.

Example 12.5 State feedback control synthesis problem. *Consider the problem of finding a matrix $F \in \mathbb{R}^{m \times n}$ such that the matrix $A + BF \in \mathbb{R}^{n \times n}$ has all of its eigenvalues in the open left-half complex plane. By the theory of Lyapunov equations (see Zhou et al. (1996)), this is equivalent to finding a matrix F and a postive definite matrix $P \in \mathbb{R}^{n \times n}$ such that the following inequality holds*

$$(A + BF)^T P + P(A + BF) < 0 \tag{12.40}$$

or

$$A^T P + PA + F^T B^T P + PBF \quad < \quad 0 \tag{12.41}$$

This problem is not in LMI form due to the terms which contain products of F and P - these terms are 'nonlinear' and as there are products of two variables, they are said to be 'bilinear'. If we multiply on either side of equation (12.41) by $Q := P^{-1}$ (which does not change the definiteness of the expression since $\mathrm{rank}(P) = \mathrm{rank}(Q) = n$) we obtain

$$QA^T + AQ + QF^T B^T + BFQ < 0 \tag{12.42}$$

This is a new matrix inequality in the variables $Q > 0$ and F, but it is still nonlinear. To rectify this, we simply define a second new variable $L = FQ$ giving

$$QA^T + AQ + L^T B^T + BL < 0 \tag{12.43}$$

We now have an LMI feasibility problem in the new variables $Q > 0$ and $L \in \mathbb{R}^{m\times n}$. Once this LMI has been solved we can recover a suitable state-feedback matrix as $F = LQ^{-1}$ and our Lyapunov variable as $P = Q^{-1}$. Hence, by making a change of variables we have obtained an LMI from a nonlinear matrix inequality.

The key fact to consider when making a change of variables is the assurance that the original variables can be recovered and that they are not over-determined.

12.2.2 Congruence transformation

For a given positive definite matrix $Q \in \mathbb{R}^{n\times n}$, we know that, for another real matrix $W \in \mathbb{R}^{n\times n}$ such that $\mathrm{rank}(W) = n$, the following inequality holds

$$WQW^T > 0 \tag{12.44}$$

In other words, *definiteness* of a matrix is invariant under pre and post-multiplication by a full rank real matrix, and its transpose, respectively. The process of transforming $Q > 0$ into equation (12.44) using a real full rank matrix is called a 'congruence transformation'. It is very useful for 'removing' bilinear terms in matrix inequalities and is often used, in conjunction with a change of variables, to make a bilinear matrix inequality *linear*. Often W is chosen to have a diagonal structure.

Example 12.6 **Making a bilinear matrix inequality linear.** *Consider*

$$Q = \begin{bmatrix} A^T P + PA & PBF + C^T V \\ \star & -2V \end{bmatrix} < 0 \tag{12.45}$$

where the matrices $P > 0, V > 0$ and F (definiteness not specified) are the matrix variables and the remaining matrices are constant. The $\star$ in the bottom left entry of the matrix denotes the term required to make the expression symmetric (in this case, $\star = F^T B^T P^T + V^T C$) and will be used frequently hereafter. Notice that this inequality is bilinear in the variables P

and F which occur in the (1,2) and (2,1) blocks of the matrix $Q \in \mathbb{R}^{(n+l)\times(n+l)}$. However, if we choose the matrix

$$W = \begin{bmatrix} P^{-1} & 0 \\ 0 & V^{-1} \end{bmatrix} \in \mathbb{R}^{(n+l)\times(n+l)} \tag{12.46}$$

which is full rank ($\mathrm{rank}(W) = n + l$) by virtue of the invertibility of P and V (which exist as the matrices are positive definite), then calculating WQW^T gives

$$WQW^T = \begin{bmatrix} P^{-1}A^T + AP^{-1} & BFV^{-1} + P^{-1}C^T \\ \star & -2V^{-1} \end{bmatrix} < 0 \tag{12.47}$$

Hence, in the new variables $X = P^{-1}$, $U = V^{-1}$ and $L = FV^{-1}$ we have a linear *matrix inequality*

$$WQW^T = \begin{bmatrix} XA^T + AX & BL + XC^T \\ \star & -2U \end{bmatrix} \tag{12.48}$$

Notice that the original variables can be recovered by inverting X and U.

12.2.3 Schur complement

The main use of the Schur complement is to transform quadratic matrix inequalities into linear matrix inequalities, or at least as a step in this direction. Schur's formula says that the following statements are equivalent:

$$(i) \qquad \Phi = \begin{bmatrix} \Phi_{11} & \Phi_{12} \\ \Phi_{12}^T & \Phi_{22} \end{bmatrix} < 0$$

$$(ii) \qquad \begin{aligned} \Phi_{22} &< 0 \\ \Phi_{11} - \Phi_{12}\Phi_{22}^{-1}\Phi_{12}^T &< 0 \end{aligned}$$

A non-strict form involving a Moore-Penrose pseudo inverse also exists if Φ is only negative semi-definite; see Boyd et al. (1994).

Example 12.7 Making a quadratic inequality linear. *Consider the LQR-type matrix inequality (Riccati inequality)*

$$A^T P + PA + PBR^{-1}B^T P + Q < 0 \tag{12.49}$$

where $P > 0$ is the matrix variable and the other matrices are constant with $Q, R > 0$. This inequality can be used to minimize the cost function (seen in Chapter 9)

$$J = \int_0^\infty (x^T Qx + u^T Ru)dt \tag{12.50}$$

If we now define

$$\Phi_{11} := A^T P + PA + Q \tag{12.51}$$

$$\Phi_{12} := PB \tag{12.52}$$

$$\Phi_{22} := -R \tag{12.53}$$

and use the Schur complement identities we can transform our Riccati inequality into

$$\begin{bmatrix} A^T P + PA + Q & PB \\ \star & -R \end{bmatrix} < 0 \tag{12.54}$$

In other words, we have transformed a quadratic matrix inequality into a linear matrix inequality.

12.2.4 The S-procedure

The S-procedure is essentially a method which enables one to combine several quadratic inequalities into one single inequality (generally with some conservatism). There are many instances in control engineering when we would like to ensure that a single quadratic function of $x \in \mathbb{R}^m$ is such that

$$F_0(x) \le 0; \quad F_0(x) \triangleq x^T A_0 x + 2b_0 x + c_0 \tag{12.55}$$

whenever certain other quadratic functions are positive semi-definite, i.e. when

$$F_i(x) \ge 0 \quad F_i(x) \triangleq x^T A_i x + 2b_0 x + c_0, \quad i \in \{1, 2, \ldots, q\} \tag{12.56}$$

To illustrate the S-procedure, consider $i = 1$, for simplicity. That is, we would like to ensure $F_0(x) \le 0$ for all x such that $F_1(x) \ge 0$. Now, if there exists a positive (or zero) scalar, τ, such that

$$F_{\text{aug}}(x) \triangleq F_0(x) + \tau F_1(x) \le 0 \quad \forall x \quad \text{s.t.} \quad F_1(x) \ge 0 \tag{12.57}$$

it follows that our goal is achieved. To see this, note that $F_{aug}(x) \le 0$ implies that $F_0(x) \le 0$ if $\tau F_1(x) \ge 0$ because $F_0(x) \le F_{\text{aug}}(x)$ if $F_1(x) \ge 0$. Thus, extending this idea to q inequality constraints we have that

$$F_0(x) \le 0 \quad \text{whenever} \quad F_i(x) \ge 0 \tag{12.58}$$

holds if

$$F_0(x) + \sum_{i=1}^{q} \tau_i F_i(x) \le 0, \quad \tau_i \ge 0 \tag{12.59}$$

In general the S-procedure is conservative; inequality (12.59) implies inequality (12.58), but not vice versa. The usefulness of the S-procedure is in the possibility of including the τ_i's as variables in an LMI problem.

Example 12.8 Combining quadratic constraints to yield an LMI. *An instructive example, taken from Boyd et al. (1994), involves finding a matrix variable $P > 0$ such that*

$$\begin{bmatrix} x \\ z \end{bmatrix}^T \begin{bmatrix} A^T P + PA & PB \\ B^T P & 0 \end{bmatrix} \begin{bmatrix} x \\ z \end{bmatrix} < 0 \tag{12.60}$$

whenever $x \neq 0$ and z satisfy the constraint

$$z^T z \leq x^T C^T C x \tag{12.61}$$

Note that inequality (12.61) is equivalent to

$$(x^T C^T C x - z^T z) \quad \geq \quad 0 \tag{12.62}$$

or

$$\begin{bmatrix} x \\ z \end{bmatrix}^T \begin{bmatrix} C^T C & 0 \\ 0 & -I \end{bmatrix} \begin{bmatrix} x \\ z \end{bmatrix} \quad \geq \quad 0 \tag{12.63}$$

The two quadratic constraints (12.60) and (12.63) can thus be combined with the S-procedure to yield the LMI

$$\begin{bmatrix} A^T P + PA + \tau C^T C & PB \\ B^T P & -\tau I \end{bmatrix} < 0 \tag{12.64}$$

in the variables $P > 0$ and $\tau \geq 0$.

12.2.5 The projection lemma and Finsler's Lemma

In some types of control problems, particularly those seeking dynamic controllers, we encounter inequalities of the form

$$\Psi(X) + G(X)\Lambda H^T(X) + H(X)\Lambda^T G^T(X) < 0 \tag{12.65}$$

where X and Λ are the matrix variables and $\Psi(.), G(.), H(.)$ are (normally affine) functions of X but not of Λ.

In Gahinet and Apkarian (1994), it is proved that inequality (12.65) is satisfied, for some X, if and only if

$$\begin{cases} W^T_{G(X)} \Psi(X) W_{G(X)} & < \quad 0 \\ W^T_{H(X)} \Psi(X) W_{H(X)} & < \quad 0 \end{cases} \tag{12.66}$$

where $W_{G(X)}$ and $W_{H(X)}$ are matrices with columns which form bases for the null spaces of $G(X)$ and $H(X)$ respectively. Alternatively, $W_{G(X)}$ and $W_{H(X)}$ are sometimes called *orthogonal complements* of $G(X)$ and $H(X)$ respectively. Note that

$$W_{G(X)} G(X) = 0 \quad W_{H(X)} H(X) = 0 \tag{12.67}$$

The main point of this result (referred to as Gahinet and Apkarian's projection lemma) is that it enables one to transform a matrix inequality, which is not necessarily a linear function of *two* variables, into two inequalities which are functions of just *one* variable. This has two useful consequences:

(i) It can facilitate the derivation of an LMI and
(ii) There are less variables for computation

Finsler (1937) also proved that inequality (12.65) is equivalent to two inequalities

$$\left\{ \begin{array}{ll} \Psi(X) - \sigma G(X)G(X)^T & < 0 \\ \Psi(X) - \sigma H(X)H(X)^T & < 0 \end{array} \right. \tag{12.68}$$

for some real σ. In other words, inequalities (12.66) and (12.68) are equivalent. This result is often referred to as *Finsler's Lemma*.

Example 12.5 (state feedback) continued. Consider again the state feedback synthesis problem of finding $P > 0$ and F such that

$$(A + BF)^T P + P(A + BF) < 0 \tag{12.69}$$

Using the change of variables described earlier in Example 12.5, we can change this problem into that of finding $Q > 0$ and L such that

$$QA^T + AQ + L^T B^T + BL < 0 \tag{12.70}$$

If we choose to eliminate the variable L using the projection lemma we get

$$\left\{ \begin{array}{lll} W_B^T(AQ + QA^T)W_B & < & 0, \quad Q > 0 \\ W_I^T(AQ + QA^T)W_I & < & 0, \quad Q > 0 \end{array} \right. \tag{12.71}$$

However, as W_I is a matrix whose columns span the null space of the identity matrix which is $\mathcal{N}(I) = \{0\}$ the above equation simply reduces to

$$W_B^T(AQ + QA^T)W_B < 0, \quad Q > 0 \tag{12.72}$$

which is an LMI problem.

Alternatively, using Finsler's Lemma we get

$$\left\{ \begin{array}{lll} AQ + QA^T - \sigma BB^T & < & 0, \quad Q > 0 \\ AQ + QA^T - \sigma I & < & 0, \quad Q > 0 \end{array} \right. \tag{12.73}$$

However, we can neglect the second inequality because if we can find a σ satisfying the first inequality, we can always find one which satisfies the second.

Notice that the use of both the projection lemma and Finsler's Lemma effectively reduces our original LMI problem into two separate ones: the first LMI problem involves the calculation of $Q > 0$; the second ivolves the back-substitution of Q into the original problem in order for us to find L (and then F). The reader is, however, cautioned against the possibility of ill-conditioning in this two step approach. For some problems, normally those with large numbers of variables, X can be poorly conditioned, which can hinder the numerical determination of Λ from equation (12.65).

12.3 Case study: anti-windup compensator synthesis

Linear controllers can be very effective at controlling real plants until they encounter actuator saturation, which can cause the behaviour of the system to dramatically deteriorate, or even become unstable. To limit this loss of performance special compensators called *anti-windup compensators* are added which take action when the control signal saturates. As the anti-windup compensator is inactive for large periods of time, conventional linear methods are not always useful for designing such a compensator. However, as we will discover, LMIs can play an important part in this design.

Anti-windup was also discussed in Section 9.4.5, where the Hanus scheme was briefly introduced. The approach below is more general and rigourous.

12.3.1 Representing anti-windup compensators

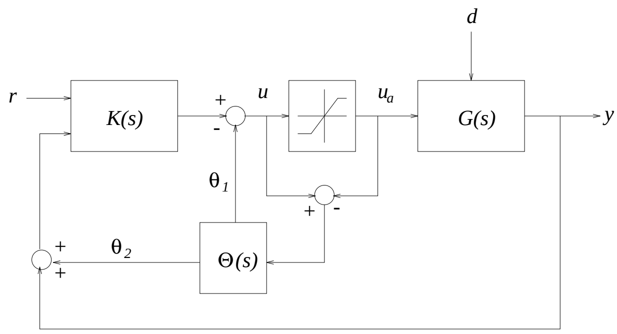

Figure 12.1: Generic anti-windup scheme

A generic anti-windup compensator is depicted in Figure 12.1. The plant $G(s) = [G_1(s) \quad G_2(s)]$ is assumed to be stable (to enable global results to be obtained - see Turner and Postlethwaite (2004) for more detail about this). $G_1(s)$ represents the disturbance feedforward part of the plant and therefore is the transfer function from the disturbance $d(s)$ to the output $y(s)$. Similarly $G_2(s)$ represents the feedback part of the plant and therefore is the transfer function from the actual control input $u_a(s)$ to the output $y(s)$. Only $G_2(s)$ plays a part in anti-windup synthesis and its state-space realisation is given by

$$G_2(s) \stackrel{s}{=} \left[\begin{array}{c|c} A_p & B_p \\ \hline C_p & D_p \end{array}\right] \tag{12.74}$$

The $K(s)$ is the linear controller, which we assume has been designed such that its closed loop interconnection with $G(s)$ is stable, in the absence of saturation, and such that some linear performance specifications have been satisfied.

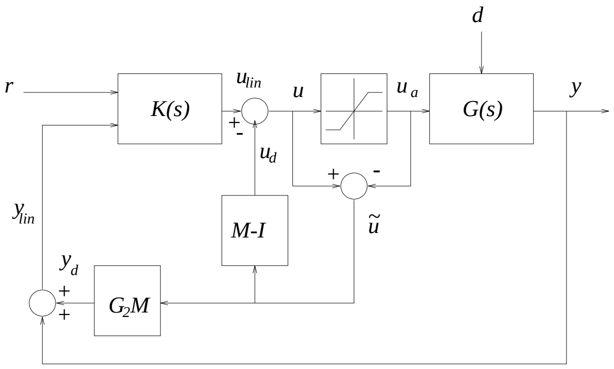

Figure 12.2: Conditioning with $M(s)$

The anti-windup compensator, $\Theta(s)$, adds extra signals to the controller input and output when control signal saturation occurs. By choosing $\Theta(s)$ in different ways, the closed loop properties during and following saturation are influenced. If we choose $\Theta(s)$ as shown in Figure 12.2, which is parameterised in terms of the transfer function $M(s)$, it can be proved using the identity

$$\mathrm{Dz}(u) = u - \mathrm{sat}(u) \tag{12.75}$$

where $\mathrm{Dz}(.)$ and $\mathrm{sat}(.)$ represent the saturation and deadzone functions, that Figures 12.2 and 12.3 are equivalent (Weston and Postlethwaite, 2000). Figure 12.3 is convenient to analyse the stability and performance of the system and, in particular, it can be seen that, provided that the nominal linear closed loop is stable, that overall stability is governed by the stability of the *nonlinear loop*. Moreover, the performance of the system can be measured by the 'size' of the map from u_{lin} to y_d. This map, call it $\mathcal{T}_p$ governs how much the linear output is perturbed by the saturation of the control signal. Hence, it would be useful to minimize the size of the norm of this - nonlinear - operator. For more information on the motivation behind this see, for example, Turner and Postlethwaite (2004) and Turner et al. (2003).

12.3.2 Lyapunov stability

The stability of nonlinear systems is more difficult to ascertain than that of linear systems. A sufficient (but not necessary) condition was given by Lyapunov; see, for example, Khalil (1996).

Theorem 12.1 Lyapunov's Theorem *Given a positive definite function* $V(x) > 0 \quad \forall x \neq 0$ *and an autonomous system* $\dot{x} = f(x)$*, then the system* $\dot{x} = f(x)$ *is stable if*

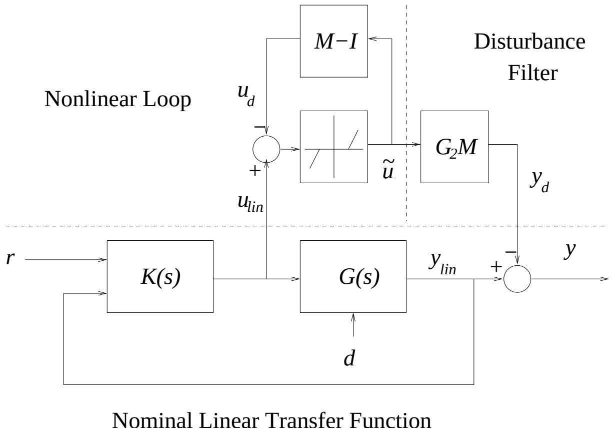

Figure 12.3: Equivalent representation of conditioning with $M(s)$

$$\dot{V}(x) = \frac{\partial V}{\partial x} f(x) < 0 \quad \forall x \neq 0 \tag{12.76}$$

As our anti-windup system is nonlinear due to the presence of the saturation function, we will use Lyapunov's theorem to establish stability.

12.3.3 $\mathcal{L}_2$ gain

In linear systems, the $\mathcal{H}_\infty$norm is equivalent to the maximum root mean square or RMS energy gain of the system. The equivalent measure for nonlinear systems is the so-called $\mathcal{L}_2$ gain, which is a bound on the RMS energy gain. Specifically a *nonlinear* system with input $u(t)$ and output $y(t)$ is said to have an $\mathcal{L}_2$ gain of γ if

$$\|y\|_2 < \gamma\|u\|_2 + \beta \tag{12.77}$$

where β is a positive constant and $\|(.)\|_2$ denotes the standard 2-norm-in-time ($\mathcal{L}_2$ norm) of a vector. Thus the $\mathcal{L}_2$ gain of a system can be taken as a measure of the size of the output a system exhibits relative to the size of its input.

12.3.4 Sector boundedness

The saturation function is defined as

$$sat(u) = [sat_1(u_1), \ldots, sat_m(u_m)]^T \tag{12.78}$$

and $sat_i(u_i) = \mathrm{sign}(u_i) \times \min\{|u_i|, \bar{u}_i\}$, $\quad \bar{u}_i > 0 \quad \forall i \in \{1, \ldots, m\}$, where $\bar{u}_i$ is the i'th saturation limit. From this, the deadzone function can be defined as

$$Dz(u) = u - sat(u) \tag{12.79}$$

It is easy to verify that the saturation function, $\mathrm{sat}_i(u_i)$ satisfies the following inequality

$$u_i \mathrm{sat}_i(u_i) \geq \mathrm{sat}_i^2(u_i) \tag{12.80}$$

or

$$\mathrm{sat}_i(u_i)[u_i - \mathrm{sat}_i(u_i)]w_i \geq 0 \tag{12.81}$$

for some $w_i > 0$. Collecting this inequality for all i we can write

$$\mathrm{sat}(u)^T W[u - \mathrm{sat}(u)] \geq 0 \tag{12.82}$$

for some diagonal $W > 0$. Similarly it follows that

$$\mathrm{Dz}(u)^T W[u - \mathrm{Dz}(u)] \geq 0 \tag{12.83}$$

for some diagonal $W > 0$. We will make use of this inequality in the derivation of our anti-windup compensator synthesis equations.

12.3.5 Full-order anti-windup compensators

The term 'full-order' anti-windup compensators has a similar meaning to the term 'full-order' $\mathcal{H}_\infty$controller; that is, the compensator is of order equal to the plant. We will confine our attention to full order anti-windup compensator synthesis. For a treatment of low-order and static anti-windup synthesis, see Turner and Postlethwaite (2004).

Assume that we factorise $G_2(s) = N(s)M(s)^{-1}$ i.e. the anti-windup parameter $M(s)$ is chosen as part of a coprime factorisation of $G_2(s)$; for example, see Section 4.1.5 or (Zhou et al., 1996). In this case, the operator $\mathcal{T}_p : u_{lin} \mapsto y_d$ is given by

$$\mathcal{T}_p \sim \begin{cases} \dot{x}_p &= (A_p + B_p F)x_p + B_p \widetilde{u} \\ u_d &= F x_p \\ y_d &= (C_p + D_p F)x_p + D_p \widetilde{u} \\ \widetilde{u} &= \mathrm{Dz}(u_{lin} - u_d) \end{cases} \tag{12.84}$$

The matrix F determines the coprime factorisation of $G_2(s)$, which in turn influences the performance of the anti-windup compensator. Hence, our goal in full-order anti-windup synthesis is to find an appropriate matrix F such that the closed-loop performance in the presence of saturation is good.

12.3.6 Anti-windup synthesis

We would like to choose F (and therefore $M(s)$) such that $\mathcal{T}_p$ is internally stable with sufficiently small $\mathcal{L}_2$. It can be verified (see Turner and Postlethwaite (2004)) that if we choose a Lyapunov function $V(x) = x_p^T P x_p > 0$ and ensure that

$$\dot{V}(x) + y_d^T y_d - \gamma^2 u_{lin}^T u_{lin} < 0 \tag{12.85}$$

then the operator $\mathcal{T}_p$ is indeed internally stable with an $\mathcal{L}_2$ gain of γ. Therefore, using the expression for $\mathcal{T}_p$ we can write inequality (12.85) as

$$z^T \begin{bmatrix} \bar{A}^T P + P\bar{A} + \bar{C}^T\bar{C} & PB_P + \bar{C}^T D_P & 0 \\ \star & D_p^T D_P & 0 \\ \star & \star & -\gamma^2 I \end{bmatrix} z < 0 \tag{12.86}$$

where

$$\bar{A} = A_p + B_p F \tag{12.87}$$

$$\bar{C} = C_p + D_p F \tag{12.88}$$

$$z = \begin{bmatrix} x_p & \widetilde{u} & u_{\text{lin}} \end{bmatrix}^T \tag{12.89}$$

However, from the sector boundedness of the deadzone we also have that

$$2\widetilde{u}^T W[u_{lin} - Fx_p - \widetilde{u}] \geq 0 \tag{12.90}$$

We will use the S-procedure to combine inequalities (12.86) and (12.90). First note that inequality (12.90) may be written as

$$z^T \begin{bmatrix} 0 & -F^T W & 0 \\ \star & -2W & W \\ \star & \star & 0 \end{bmatrix} z \geq 0 \tag{12.91}$$

We have added the factor of 2 into inequality (12.90) so that inequality (12.91) can be written in a tidier fashion; without this factor of 2, there would be several factors of $1/2$ present. Using the S-procedure described earlier, we can combine inequality (12.86) with (12.91) to obtain

$$\begin{bmatrix} \bar{A}^T P + P\bar{A} + \bar{C}^T\bar{C} & PB_P + \bar{C}^T D_P - F^T W\tau & 0 \\ \star & -2W\tau + D_p^T D_P & W\tau \\ \star & \star & -\gamma^2 I \end{bmatrix} < 0 \tag{12.92}$$

Notice that τ *only* appears adjacent to W, so we can define a new variable $V = W\tau$ and use this from now on. Applying the Schur complement we obtain

$$\begin{bmatrix} \bar{A}^T P + P\bar{A} & PB_P - F^T V & 0 & \bar{C}^T \\ \star & -2V & V & D_p^T \\ \star & \star & -\gamma I & 0 \\ \star & \star & \star & -\gamma I \end{bmatrix} < 0 \tag{12.93}$$

Next, using the congruence transformation $\mathrm{diag}(P^{-1}, V^{-1}, I, I)$ we obtain

$$\begin{bmatrix} P^{-1}A_p^T + A_pP^{-1} + P^{-1}F^TB_p^T + B_pFP^{-1} & B_PV^{-1} - P^{-1}F^T & 0 & P^{-1}C_p^T + P^{-1}F^TD_p^T \\ \star & -2V^{-1} & I & V^{-1}D_p^T \\ \star & \star & -\gamma I & 0 \\ \star & \star & \star & -\gamma I \end{bmatrix} \tag{12.94}$$

Finally, defining new variables $Q = P^{-1}, U = V^{-1}, L = QF$ we get

$$\begin{bmatrix} QA_p^T + A_pQ + L^TB_p^T + B_pL & B_PU - QF^T & 0 & QC_p^T + L^TD_p^T \\ \star & -2U & I & UD_p^T \\ \star & \star & -\gamma I & 0 \\ \star & \star & \star & -\gamma I \end{bmatrix} < 0$$

which is now an LMI in $Q > 0$, $U > 0$ and diagonal, $\gamma > 0$ and L. To obtain F we can thus compute $F = Q^{-1}L$, which allows us to construct our anti-windup compensator.

For applications of these and similar formulae see Turner and Postlethwaite (2004) and Herrmann et al.(2003a, 2003b).

12.4 Conclusion

In recent years, efficient interior-point algorithms have been developed to solve comvex LMI optimization problems of the type presented in this chapter. We have described the main (generic) LMI problems in control and the tools and tricks required to transform them into formats that can readily take advantage of the algorithms now available, especially in Matlab. By including this chapter, we have attempted to give the essential ingredients for developing an understanding of the power and usufulness of LMIs. More details can be found in Boyd et al. (1994). A cautionary note is that the complexity of LMI computations is high, and certainly higher, for example, than solving a Riccati equation in a conventional approach. Nevertheless, the LMI approach opens the way to solving problems that conventional methods cannot.