Lecture 4

Optimization problems with LMI constraints

Many optimization and control issues may be expressed as a series of LMIs with a feasible solution. The majority of issues, on the other hand, are best expressed as a simple objective function optimization over a set of LMIs. This lecture presents the formulation of certain optimization problems that arise when utilizing LMIs to solve some problems in control area.

4.1 Eigenvalue problem (EVP)

Eigenvalue problem (EVP) consists in the minimization of the maximum eigenvalue of a $n \times n$ symmetric matrix $A(P)$ that depends affinely on a variable, subject to LMI (symmetric) constraint $B(P) > 0$, i.e.,

$$\boxed{\begin{array}{c} \lambda_{\max}(A(P)) \to \min\limits_{P=P^{\intercal}} \\ B(P) > 0. \end{array}} \tag{4.1}$$

Since for large enough λ

$$A(P) \leq \lambda I_{n\times n}\,,$$

this problem can be equivalently represented as follows:

$$\boxed{\left.\begin{array}{c} \lambda \to \min\limits_{\lambda, P=P^{\intercal}} \\ \\ \begin{bmatrix} \lambda I_{n\times n} - A(P) > 0 & 0 \\ 0 & B(P) \end{bmatrix} > 0. \end{array}\right\}} \tag{4.2}$$

4.2 Tolerance level optimization

The *tolerance level optimization problem* can be represented in the following manner:

$$\boxed{\begin{array}{c} \gamma \to \min\limits_{0<\gamma, 0<P=P^{\intercal}} \\ \\ PA + A^{\intercal}P + C^{\intercal}C + \gamma^{-1}PBB^{\intercal}P < 0 \end{array}} \tag{4.3}$$

Equivalently, it can be rewritten by the Schur's complement (Theorem 1.1) as an optimization problem with LMI constraints:

$$\boxed{\begin{array}{c} \gamma \to \min\limits_{0<\gamma, 0<P=P^{\intercal}} \\ \\ \begin{bmatrix} -PA - A^{\intercal}P - C^{\intercal}C & PB & 0 \\ B^{\intercal}P & \gamma I & 0 \\ 0 & 0 & P \end{bmatrix} > 0. \end{array}} \tag{4.4}$$

4.3 Maximization of the quadratic stability degree

The *quadratic stability degree* of a stable $n \times n$ matrix A is defined as a positive value α satisfying the matrix inequality

$$A^{\intercal}P + PA < -\alpha P$$

for some positive definite matrix P, involved into the quadratic Lyapunov function $V = x^{\intercal}Px$, and for the matrix A, defining the linear dynamics $\dot{x} = Ax$. The problem of the maximization of the quadratic stability degree consists in the following optimization problem

$$\boxed{\begin{array}{c} \alpha \to \max\limits_{0<\alpha, 0<P=P^{\intercal}} \\ \\ A^{\intercal}P + PA + \alpha P < 0, \end{array}} \tag{4.5}$$

which can be expressed by the Schur's complement (Theorem 1.1) as an optimization with LMI constraint, namely,

$$\boxed{\begin{array}{c} \alpha \to \max\limits_{0<\alpha, 0<P=P^{\intercal}} \\ \\ \begin{bmatrix} -A^{\intercal}P - PA - \alpha P & 0 \\ 0 & P \end{bmatrix} > 0. \end{array}} \tag{4.6}$$

4.4 Minimization of linear function $\mathrm{Tr}\,(CPC^{\intercal})$ under the Lyapunov-type constraint

Lemma 4.1 ([8]) *Let*

1) *the matrix $A \in \mathbb{R}^{n\times n}$ is Hurwitz;*

2) *the pair (A, B) is controllable, i.e., there exists a matrix K such that $(A + KB)$ is Hurwitz.*

Then for any matrix $C \in \mathbb{R}^{k\times n}$ the solution of the problem

$$\boxed{\mathrm{Tr}\,(CPC^{\intercal}) \to \min_{P\geq 0}} \tag{4.7}$$

under the constraint

$$\boxed{AP + PA^{\intercal} + BB^{\intercal} \leq 0} \tag{4.8}$$

is attained on the Lyapunov matrix equation

$$\boxed{\begin{aligned} &AP^{*} + P^{*}A^{\intercal} + BB^{\intercal} = 0, \\ &P^{*} = \int_{t=0}^{\infty} e^{At}BB^{\intercal}e^{A^{\intercal}t}dt. \end{aligned}} \tag{4.9}$$

Proof. Suppose that the minimizing solution satisfies the equation

$$AP + PA^{\intercal} + BB^{\intercal} = -Q < 0.$$

Then

$$P = \int_{t=0}^{\infty} e^{At}\left(Q + BB^{\intercal}\right)e^{A^{\intercal}t}dt \geq \int_{t=0}^{\infty} e^{At}BB^{\intercal}e^{A^{\intercal}t}dt = P^{*},$$

and, hence,

$$\mathrm{Tr}\,(CPC^{\intercal}) = \mathrm{Tr}\,(CP^{*}C^{\intercal}) + \mathrm{Tr}\left(C\int_{t=0}^{\infty} e^{At}Qe^{A^{\intercal}t}dtC^{\intercal}\right) \geq \mathrm{Tr}\,(CP^{*}C^{\intercal}).$$

This means that P^{*} is minimizer. Lemma is proven. ■

4.5 The convex function $\log \det A^{-1}(X)$ minimization

First notice that $\log \det A^{-1}(X)$ is a convex function of A. We will encounter the following problem*:*

$$\boxed{\log\det A^{-1}(X) \to \min_{X=X^{\intercal}\in\mathbb{R}^{n\times n}}} \tag{4.10}$$

subjected to the constraints

$$\boxed{A(X) > 0,\ B(X) > 0,} \tag{4.11}$$

where $A(X)$, $B(X)$ are symmetric matrices that depend affinely (linearly) on X.

Example 4.1 *As an example of the problem (4.10)-(4.11) consider the following one:* ***find a minimal ellipsoid***

$$\boxed{\mathcal{E} := \{z \mid z^{\intercal}Pz \leq 1\},\ P > 0,} \tag{4.12}$$

containing the set of given points $v_i\,(i = 1, ..., L)$*, i.e.,* $v_i \in \mathcal{E}$*. Since the volume of* $\mathcal{E}$ *is proportional to* $(\det P)^{-1/2}$ *, minimizing* $\log\det P^{-1}$ *is the same as minimizing the volume of* $\mathcal{E}$*, this problem is converted into the following one:*

$$\boxed{\left.\begin{array}{c} \log\det P^{-1} \to \min\limits_{P\in\mathbb{R}^{n\times n}} \\ \\ P > 0,\ v_i^{\intercal}Pv_i \leq 1\ \ (i = 1, ..., L)\,. \end{array}\right\}} \tag{4.13}$$

4.6 Numerical methods for LMIs resolution

Numerical methods for LMI's resolution are discussed in details in this section.

4.6.1 What does it mean "to solve LMI"?

There exist several efficient methods for LMIs resolution. By "*solve an LMI*" we mean here:

- determine whether or not the LMI (or, the corresponding problem) is feasible;

- if it is, compute then a feasible point with "*an objective value*" that exceeds the global minimum by less than some prespecified accuracy.

What does it mean "*an objective value*", it depends on each concrete problem to be solved. Here we will assume that the problem we are solving has at least one "*optimal point*", i.e., the constraints are feasible.

To realize the numerical methods describe below, first, let us represent the matrix $X \in \mathbb{R}^{n \times n}$ as the corresponding extended vector $x \in \mathbb{R}^{n^2}$ obtained by the simple implementation of the operator col, that is,

$$\boxed{x := \operatorname{col} X.} \tag{4.14}$$

4.6.2 Ellipsoid algorithm

In a feasible problem, we may consider any feasible point as being optimal. The **basic idea** of the *ellipsoid algorithm* is as follows:

1. One may start with an ellipsoid $\mathcal{E}^{(0)}$ that is guaranteed to contain an optimal point.

2. Then the *cutting plane* for our problem is computed that passes through the center point $x^{(0)}$ of the initial ellipsoid $\mathcal{E}^{(0)}$. This means that we need to find a nonzero vector $g^{(0)}$ (namely, a vector orthogonal to the plane to be computed) such that an optimal point lies in the half-space

$$\left\{z \in \mathbb{R}^{n^2} \mid g^{(0)\intercal}\left(z - x^{(0)}\right) < 0\right\} \tag{4.15}$$

(below, we shall present some example how calculate $g^{(0)}$ in some concrete problems).

3. After that one may conclude that the sliced half-ellipsoid

$$\mathcal{E}^{(0)} \cap \left\{z \in \mathbb{R}^{n^2} \mid g^{(0)\intercal}\left(z - x^{(0)}\right) < 0\right\}$$

contains an optimal point.

4. Then we compute the ellipsoid $\mathcal{E}^{(1)}$ of a minimum volume that contains this sliced half-ellipsoid. This ellipsoid $\mathcal{E}^{(1)}$ is guaranteed to contain an optimal point, but its volume is expected to be less then the volume of the initial ellipsoid $\mathcal{E}^{(0)}$, i.e.

$$\operatorname{vol}\mathcal{E}^{(1)} < \operatorname{vol}\mathcal{E}^{(0)}.$$

5. The process is then iterated.

More explicitly, this algorithm may be describe as follows. Any *ellipsoid* $\mathcal{E}$ may be associated with some positive definite matrix A, that is,

$$\boxed{\mathcal{E} := \left\{ z \in \mathbb{R}^{n^2} \mid (z-a)^\intercal A^{-1} (z-a) \le 1 \right\}} \tag{4.16}$$

where $A = A^\intercal > 0$. The *minimum volume ellipsoid* $\tilde{\mathcal{E}}$ containing the slice half-ellipsoid

$$\left\{ z \in \mathbb{R}^{n^2} \mid (z-a)^\intercal A^{-1} (z-a) \le 1,\ g^\intercal (z-a) < 0 \right\}$$

is given by the symmetric matrix $\tilde{A}$ and the vector $\tilde{a}$ as its center, namely,

$$\boxed{\begin{gathered} \tilde{\mathcal{E}} := \left\{ z \in \mathbb{R}^{n^2} \mid (z-\tilde{a})^\intercal \tilde{A}^{-1} (z-\tilde{a}) \le 1 \right\} \\ \tilde{a} = a - \frac{A\tilde{g}}{m+1},\ m := n^2 > 1 \\ \tilde{A} = \frac{m^2}{m^2-1} \left(A - \frac{2}{m+1} A^\intercal \tilde{g} \tilde{g}^\intercal A \right) \\ \tilde{g} = \frac{g}{\sqrt{g^\intercal A g}} \end{gathered}} \tag{4.17}$$

(In the case of one variable ($m = 1$) the minimal length interval containing a half-interval is the half-interval itself). So, the ellipsoid algorithm starts with the initial points $x^{(0)}$ and the initial matrix $A^{(0)}$. Then for each intermediate pair $x^{(k)}$ and $A^{(k)}$ $(k = 0, 1, 2...)$ one may compute a vector $g^{(k)}$ and then calculate

$$x^{(k+1)} = x^{(k)} - \frac{A^{(k)}\tilde{g}}{m+1},\ m := n^2 > 1,$$

$$A^{(k+1)} = \frac{m^2}{m^2-1} \left(A^{(k)} - \frac{2}{m+1} \left(A^{(k)} \right)^\intercal \tilde{g} \tilde{g}^\intercal A^{(k)} \right),$$

$$\tilde{g} = \frac{g^{(k)}}{\sqrt{g^{(k)\intercal} A^{(k)} g^{(k)}}}.$$

It turns out that the volume

$$\mathrm{vol}\mathcal{E}^{(k)} = \det A^{(k)}$$

of these ellipsoids decreases geometrically, that is,

$$\mathrm{vol}\mathcal{E}^{(k+1)} \le e^{-k/2m}\mathrm{vol}\mathcal{E}^{(k)}$$

This means that the recursion above generates a sequence of ellipsoids that are guaranteed to contain an optimal point and converges to it geometrically. It may be proven that this algorithm converges more quickly, namely, in "polynomial time" (see [9] and references within).

Next examples illustrate the rule of selection of the nonzero vector g orthogonal to the cutting plane which is specified for each concrete problem.

Example 4.2 *Let LMI is represented in the form*

$$\boxed{F(x) := F_0 + \sum_{i=1}^{m} x_i F_i > 0} \tag{4.18}$$

where $F_i\,(i = 0, 1, 2, ..., m)$ are symmetric matrices. If x is infeasible, this means that at least there exists a nonzero vector u such that

$$u^\intercal F(x)\, u \le 0$$

Define $g = (g_1, ..., g_m)^\intercal$ by

$$\boxed{g_i = -u^\intercal F_i u} \tag{4.19}$$

Then for any z satisfying $g^\intercal(z - x) \ge 0$ it follows

$$u^\intercal F(z)\, u = u^\intercal \left[F_0 + \sum_{i=1}^{m} z_i F_i\right] u = u^\intercal F_0 u + \sum_{i=1}^{m} z_i u^\intercal F_i u =$$

$$u^\intercal F_0 u - \sum_{i=1}^{m} z_i g_i = u^\intercal F_0 u - g^\intercal z = u^\intercal F_0 u + g^\intercal x - g^\intercal(z - x) =$$

$$u^\intercal F(x)\, u - g^\intercal(z - x) \le 0$$

So, any feasible point belongs to the half-space

$$\{z \in \mathbb{R}^m \mid g^\intercal(z - x) < 0\}$$

or in other words, this g, given by (4.19), is a cutting plane for this LMI problem at the point x.

Example 4.3 *If we deal with the minimization problem of linear function $c^{\mathsf{T}}x$ subjected LMI (4.18), that is,*

$$c^{\mathsf{T}}x \to \min_{x \in \mathbb{R}^{n^2}}$$

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

one may encounter two possible situation:

1) x is infeasible, i.e., $F(x) \leq 0$; in this case g can be taken as in the previous example (4.19) since we are discarding the half-space

$$\left\{ z \in \mathbb{R}^{n^2} \mid g^{\mathsf{T}}(z - x) > 0 \right\}$$

because all such points are infeasible;

2) x is feasible, i.e., $F(x) > 0$; in this case g can be taken as

$$g = c$$

since we are discarding the half-space

$$\left\{ z \in \mathbb{R}^{n^2} \mid g^{\mathsf{T}}(z - x) > 0 \right\}$$

because all such points have an objective value larger than x and hence cannot be optimal.

4.6.3 Interior-point method

For LMI problem

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

let us define the, so-called, *barrier function* $\phi(x)$ for the feasible set:

$$\boxed{\phi(x) := \begin{cases} \log\det F^{-1}(x) & \text{if } F(x) > 0, \\ \infty & \text{if } F(x) \leq 0. \end{cases}} \tag{4.20}$$

Suppose then that the feasible set is nonempty and bounded. This implies that the matrices $F_1,\ldots,F_m$ are linearly independent (otherwise the feasible set will contain a line, i.e., be unbounded). It can be shown that $\phi(x)$ is strictly convex on the feasible set and, hence, it has a unique minimizer which we denote by x^*, that is,

$$x^* := \arg \min_{x:\, F(x)>0} \phi(x).$$

This point is referred as the *analytical center of the LMI* $F(x) > 0$. It is evident that

$$x^* := \arg \max_{F(x)>0} \det F(x).$$

Remark 4.1 *Two LMI´s $F(x) > 0$ and $T^{\intercal} F(x) T > 0$ have the same analytical center provided T is nonsingular.*

Let us apply the Newton's method for the search of the analytical center x^* of LMI, starting from a feasible initial point:

$$\boxed{x^{(k+1)} = x^{(k)} - \alpha^{(k)} H^{-1}\left(x^{(k)}\right) g\left(x^{(k)}\right),} \tag{4.21}$$

where $0 < \alpha^{(k)}$ is a damping factor at the k-th iteration, $H\left(x^{(k)}\right)$ is the Hessian and $g\left(x^{(k)}\right)$ is the gradient, respectively, of $\phi(x)$ at the point $x^{(k)}$. In [9] it is shown that if the *damping factor* is

$$\boxed{\begin{array}{c} \alpha^{(k)} := \left\{ \begin{array}{cc} 1 & \text{if } \delta\left(x^{(k)}\right) \leq 1/4 \\ 1/\left(1 + \delta\left(x^{(k)}\right)\right) & \text{otherwise} \end{array} \right. , \\ \\ \delta\left(x^{(k)}\right) := \sqrt{g^{\intercal}\left(x^{(k)}\right) H^{-1}\left(x^{(k)}\right) g\left(x^{(k)}\right)}, \end{array}} \tag{4.22}$$

then this step length always results in $x^{(k+1)}$ which is feasible, namely,

$$F\left(x^{(k+1)}\right) > 0$$

and leads to the convergence of $x^{(k)}$ to x^* when $k \to \infty$.

There exist another interior-point methods (for details, see [10]).

Important comment: the MATLAB packages realizing the numerical solutions of LMI's are

$\boxed{\textbf{SEDUMI} \text{ and } \textbf{YALMIP.}}$

References to Part 1

1. P. Finsler. Uber das vorkommen definiter und semidefiniter formen in scharen quadratischer formen. Comment. Math. Helv., 9:188–192, 1936/37.

2. V. A. Yakubovich. The *s*-procedure in nonlinear control theory. Vestnik Leningr. Univ., 4:73–93, 1977. in Russian — 1971, No.1, 62–77.

3. Derinkuyu, K., Pınar, M.Ç. On the *S*-procedure and Some Variants. Math Meth Oper Res 64, 55–77 (2006).

4. Imre Pólik and Tamás Terlaky, A Survey of the S-Lemma, SIAM Review, Volume 49, 2007, Pages 371–418.

5. Balandin, D., Kogan, M. (2007). Control laws synthesis (in Russian). Moscow: Fismatlit.

6. Poznyak, A. (2008). Advanced mathematical tools for automatic control engineers: Deterministic techniques. V.1, Amsterdam: Elsevier.

7. Rasvan, V. (1975), Absolute Stability of Time Lag Control Systems, Ed. Academiei, Bucharest (in Romanian, Russian revised edition by Nauka, Moscow, 1983).

8. Polyak, B. and Sherbakov, P. (2002), Robust Stability and Control, Nauka, Moscow, Russia.

9. Nesterov, Y. and Nemirovsky, A. (1994), Interior-point Polynomial Methods in Convex Programming, Vol. 13 of Studies in Applied Mathematics, SIAM, Philadelphia, Pennsylvania, USA.

10. Boyd, S., El Chaoui, L., Feron, E. and Balakrishnan, V. (1994), Linear Matrix Inequalities in System and Control Theory, Vol. 15 of SIAM Studies in Applied Mathematics, SIAM, Philadelphia, Pennsylvania, USA.

11. Scherer, Carsten, and Siep Weiland (2000), Linear matrix inequalities in control, Lecture Notes, Dutch Institute for Systems and Control, Delft, The Netherlands 3.2 .

12. Nie, Jiawang, and Markus Schweighofer (2007), On the complexity of Putinar's Positivstellensatz, Journal of Complexity 23.1 : 135-150.