

LMI's and SDP. Notes by Eduardo Sontag, Rutgers, Nov 2006

Let \mathcal{S}^n denote the subspace of $\mathbb{R}^{n \times n}$ consisting of symmetric matrices, let \mathcal{P}_{\geq}^n denote the subset of \mathcal{S}^n consisting of positive semidefinite matrices (that is, $x'Px \geq 0$ for all $x \in \mathbb{R}^n$), and let $\mathcal{P}_{>}^n$ the subset of positive definite matrices ($x'Px > 0$ for all nonzero $x \in \mathbb{R}^n$).

We also write $P \geq 0$ or $P > 0$ to indicate that a matrix is in \mathcal{P}_{\geq}^n or $\mathcal{P}_{>}^n$ respectively.

Recall that an *affine* map $f : V \rightarrow W$ between two real vector spaces is one that satisfies

$$f(\alpha v_1 + (1 - \alpha)v_2) = \alpha f(v_1) + (1 - \alpha)f(v_2) \quad \forall v_1, v_2 \in V, \quad \forall \alpha \in \mathbb{R}.$$

Equivalently, $f(v) = w_0 + g(v)$, where g is linear and w_0 is some element of W . (Linear maps are a particular case: $w_0 = 0$.)

An inequality

$$f(v) \in \mathcal{P}_{\geq}^n,$$

where $f : V \rightarrow \mathcal{S}^n$ is affine, is called a *linear matrix inequality (LMI)*. One considers also *strict* LMI's given by

$$f(v) \in \mathcal{P}_{>}^n.$$

Often LMI problems involve multiple constraints, but from a theoretical point of view there is no loss of generality in considering single constraints, because several constraints can be “stacked” into one. For example, the requirements that both $f(v) \in \mathcal{P}_{\geq}^n$ and $g(v) \in \mathcal{P}_{\geq}^m$ can be equivalently represented by the one constraint that

$$\begin{pmatrix} f(v) & 0 \\ 0 & g(v) \end{pmatrix}$$

be negative definite, and similarly for strict LMI's.

A semidefinite optimization problem or *semidefinite program (SDP)* is a problem of minimizing a linear function subject to an LMI constraint:

$$\min_{v \in V} Lv \quad \text{subject to a constraint} \quad f(v) \in \mathcal{P}_{\geq}^n$$

where L is a scalar linear mapping and $f : V \rightarrow \mathbb{R}^{n \times n}$ is affine.

Observation: if $f : V \rightarrow \mathbb{R}^{n \times n}$ is affine, then the sublevel set

$$\{v \mid f(v) \in \mathcal{P}_{\geq}^n\}$$

is convex. This is because \mathcal{P}_{\geq}^n is convex: for any $\alpha \in [0, 1]$:

$$x'(\alpha P_1 + (1 - \alpha)P_2)x = \alpha x'P_1x + (1 - \alpha)x'P_2x$$

and each term on the right-hand side is nonnegative.

Similarly, the sets of v so that $f(v) \in \mathcal{P}_{>}^n$, $-f(v) \in \mathcal{P}_{\geq}^n$, or $-f(v) \in \mathcal{P}_{>}^n$, are each convex.

Thus, SDP's (and feasibility problems represented by LMI's) are convex optimization problems. In general, the minimization of a convex function is relatively easy to do numerically: since a convex function has no spurious (non-global) local minima, gradient descent techniques converge to minima.

More generally, convex constraints can be transformed into convex penalty terms (barrier functions). For example, one may define

$$\phi(v) := -\log \det f(v) \quad \text{if } f(v) \in \mathcal{P}_{>}^n$$

and $\phi(v) := +\infty$ otherwise. Adding the term $\phi(v)$ to an optimization problem has the following effect: if we start with a v such that $f(v) \in \mathcal{P}_{>}^n$, then approaching the boundary of the feasible set $\{v \mid f(v) \in \mathcal{P}_{>}^n\}$ means that $\phi(v)$ becomes infinite. Thus an interior method will stay inside the constraint set. Interior methods as well as ellipsoid and other methods are used to efficiently (polynomial time) solve LMI's. SeDuMi is a popular (interior-point) free package for MATLAB.

One usually presents LMI's and SDP's in matrix form. Let us explain this through an example. Consider the property that the quadratic function $V(x) = x'Px$ is a Lyapunov function for $\dot{x} = Ax$. This translates into the requirement that P is positive definite and $A'P + PA$ is negative definite, i.e.

$$-(A'P + PA) \in \mathcal{P}_{>}^n \quad \text{and} \quad P \in \mathcal{P}_{>}^n.$$

Let $V = \mathcal{S}^n$. Clearly, for any fixed matrix A , and any $P \in \mathcal{S}^n$, $-(A'P + PA) \in \mathcal{S}^n$, and the map $f : P \mapsto -(A'P + PA)$ is linear, hence affine. Also the map $g(P) = P$ is obviously linear $V \rightarrow \mathcal{S}^n$, and thus the Lyapunov requirement is the strict LMI

$$f(v) \in \mathcal{P}_{>}^n, \quad g(v) \in \mathcal{P}_{>}^n$$

(as we remarked earlier, the two constraints can be stacked into just as one).

Similarly, for discrete-time systems $x^+ = Ax$, a symmetric matrix P gives a quadratic Lyapunov function provided that

$$-(A'PA - I) \in \mathcal{P}_{>}^n \quad \text{and} \quad P \in \mathcal{P}_{>}^n,$$

which is also a strict LMI.

For Lyapunov functions, this is not terribly interesting, since the theory guarantees that, if A is a Hurwitz matrix, one can just find P by solving the linear equation $A'P + PA = -I$. However, there are a large number of problems in control theory that admit no simple closed-form solution yet can be formulated as LMI problems. Let us discuss a couple of these now, as illustration of the power of these techniques.

As a first example, take two (the same arguments work for more than two) $n \times n$ matrices A_1 and A_2 , and consider a *switched system*

$$\dot{x} = A_{\sigma(t)}x$$

where $\sigma(t) \in \{1, 2\}$ is an arbitrary switching signal. (To avoid technical issues, let us suppose that the piecewise constant function σ can only switch values at a discrete set of times $t_1 < t_2 < \dots < t_n \rightarrow \infty$. Solutions are then unique and well-defined.)

How can one guarantee that $x(t) \rightarrow 0$ as $t \rightarrow \infty$, for all initial conditions and all switching functions? One sufficient condition is that there exist a common Lyapunov function, for example a common quadratic Lyapunov function. That is, there should exist a matrix $P \in \mathcal{P}_{>}^n$ such that

$$-(A_1'P + PA_1) \in \mathcal{P}_{>}^n, \quad -(A_2'P + PA_2) \in \mathcal{P}_{>}^n \quad \text{and} \quad P \in \mathcal{P}_{>}^n.$$

This is a strict LMI, and hence can be checked easily. If such a P exists, then we can find (repeating arguments done for a single system) a constant $\alpha > 0$ such that, along all trajectories,

$$\frac{dV(x(t))}{dt} \leq -\alpha V(x(t))$$

for all non-switching times. Since $V(x(t))$ is continuous, this means that

$$V(x(t)) \leq e^{-\alpha t} V(x(0)) \rightarrow 0$$

(proving the inequality in each interval between switchings), and hence $x(t) \rightarrow 0$ as claimed.

A second example is the computation of L^2 gains of systems. We do this for the system

$$(\Sigma) \quad \dot{x} = Ax + Bu$$

but similar arguments work for discrete time systems and for systems with outputs. Assume that A is Hurwitz, and we start from the initial state $x(0) = 0$. It is then possible to prove that, for each input $u(\cdot) \in L^2$, the state trajectory $x(\cdot)$ is also in L^2 . The induced L_2 norm of the operator $u(\cdot) \mapsto x(\cdot)$ is called the “ H_∞ gain” of the system (as discussed later in the course). We’d like to compute this norm,

$$\|\Sigma\| = \sup_{u \neq 0, u \in L^2} \frac{\|x\|}{\|u\|}.$$

A sufficient (also, it turns out, necessary for controllable systems, but we do not prove that fact here) condition for $\|\Sigma\| \leq \gamma_0$, for a nonnegative number γ_0 , is that there exist a positive definite matrix P such that, defining $V(x) = x'Px$ (typically an “energy storage function” for electrical or mechanical systems) such that, for every input $u(\cdot)$ and every solution $x(\cdot)$ with $x(0) = 0$,

$$\frac{dV(x(t))}{dt} \leq \gamma_0^2 \|u(t)\|^2 - \|x(t)\|^2. \quad (1)$$

Indeed, if this inequality holds, then

$$0 \leq V(x(t)) = V(x(t)) - V(x(0)) = \int_0^t \gamma_0^2 \|u(s)\|^2 ds - \int_0^t \|x(s)\|^2 ds$$

for all $t \geq 0$, and hence taking limits as $t \rightarrow \infty$ there results

$$\|x\| \leq \gamma_0 \|u\|.$$

Since

$$\frac{dx(t)'Px(t)}{dt} = \dot{x}(t)'Px(t) + x(t)'P\dot{x}(t) = (Ax(t) + Bu(t))'PAx(t) + x(t)'P(Ax(t) + Bu(t)),$$

we see that P will work provided that

$$x'(A'P + PA)x + u'B'Px + x'PBu \leq \gamma_0^2 u'u - x'x$$

for all constant vectors $x \in \mathbb{R}^n$ and $u \in \mathbb{R}^m$. This can be summarized by asking that

$$\begin{pmatrix} A'P + PA + I & PB \\ B'P & -\gamma_0^2 I \end{pmatrix}$$

be negative semidefinite (note that the matrix is automatically symmetric, if P is). Together with the requirement $P > 0$, we have an LMI problem.

In fact, we can go a step further and ask what is the *best* possible value of γ that can be found in this way. (Since the existence of V is equivalent to the gain condition –we have not proved that fact– it follows that this actually will find the exact value of the norm.) All we need to do is see the positive number γ as a positive definite matrix, and consider the composite of P and γ as a positive definite matrix \tilde{P} to be solved for. Now we look at the LMI

$$\begin{pmatrix} A'P + PA + I & PB \\ B'P & -\gamma^2 I \end{pmatrix} \leq 0$$

and consider the problem of maximizing the linear function $L(\tilde{P}) = \gamma$. This is a semidefinite program.

Some remarks

One can allow equality constraints in SDP's and LMI's, simply by eliminating variables.

Linear programming is a particular case of SDP: if we want to minimize Lv subject to $Av \geq b$, we simply view the linear constraint as requiring the stacked constraints $A_i v - b_i \geq 0$ for all i , where A_i is the i th row of A and b_i is the i th entry of b . Note that each $A_i v - b_i$ is a scalar, and therefore a "symmetric matrix".

Far more interestingly, any convex quadratic program can be seen as an SDP. Suppose that we wish to minimize a quadratic form on \mathbb{R}^n :

$$v'Qv + c'v$$

where Q is a symmetric positive semidefinite matrix, and $c \in \mathbb{R}^n$, subject to a linear constraint $Av \geq b$. In order to reduce this to an SDP, we first review the *Schur complement* construction.

For any $Z \in \mathcal{S}^m$, and any $Y \in \mathbb{R}^{m \times n}$, consider the following nonsingular matrix:

$$T = \begin{pmatrix} I & 0 \\ -Z^{-1}Y & I \end{pmatrix}.$$

Then, for any $X \in \mathcal{S}^n$, we have that

$$T' \begin{pmatrix} X & Y' \\ Y & Z \end{pmatrix} T = \begin{pmatrix} X - Y'Z^{-1}Y & 0 \\ 0 & Z \end{pmatrix}.$$

Under equivalence ($M \mapsto T'MT$, T nonsingular) of matrices, positive definiteness (or semidefiniteness) is preserved (why? prove it!). Therefore, asking

$$Z > 0 \quad \text{and} \quad X - Y'Z^{-1}Y > 0$$

is equivalent to the requirement

$$\begin{pmatrix} X & Y' \\ Y & Z \end{pmatrix} > 0$$

and similarly for negative definiteness, or for semidefiniteness.*

Back to the quadratic programming problem, we first factor $Q = R'R$ (this can always be done for a symmetric semidefinite matrix). Now we note that $\gamma \geq v'R'Rv + c'v$ if and only if

$$(\gamma - c'v) - v'R'Rv = X - Y'Z^{-1}Y \geq 0$$

where

$$X = \gamma - c'v, \quad Y = Rv, \quad Z = I.$$

So the problem is equivalent to the SDP

$$\min_{v \in \mathbb{R}^n, \gamma \in \mathbb{R}} \gamma$$

subject to the constraints that

$$\begin{pmatrix} \gamma - c'v & v'R \\ Rv & I \end{pmatrix} \geq 0$$

and that $A_i v - b_i \geq 0$ for all i .

The Schur construction appears often in optimization problems in control theory.

*For 2×2 matrices, this is just asking that all principal minors be positive (or ≥ 0).