ORF 523	LECTURE 10	Spring 2016, Princeton University
Instructor: A.A. Ahmadi		
Scribe: G. Hall		Thursday, March 24, 2016

When in doubt on the accuracy of these notes, please cross check with the instructor's notes, on `aaa.princeton.edu/orf523`. *Any typos should be emailed to gh4@princeton.edu.*

In this lecture, we consider some applications of SDP:

- Stability and stabilizability of linear systems.
  - The idea of a Lyapunov function.
- Eigenvalue and matrix norm minimization problems.

1 Stability of a linear system

Let's start with a concrete problem. Given a matrix $A \in \mathbb{R}^{n \times n}$, consider the linear dynamical system

$$x_{k+1} = Ax_k,$$

where x_k is the state of the system at time k. When is it true that $\forall x_0 \in \mathbb{R}^n$, $x_k \to 0$ as $k \to \infty$? This property called *global asymptotic stability (GAS)*[1].

The choice of $x = 0$ as the "attractor" is arbitrary here. If the system has a different equilibrium point (i.e., a point where $x_{k+1} = x_k$) then we could shift it to the origin by an affine change of coordinates. Stability is a fundamental concept in many areas of science and engineering. For example, in economics, we may want to know if deviations from some equilibrium price are forced back to the equilibrium under given price dynamics.

A standard result in linear algebra tells us that the origin of the system $x_{k+1} = Ax_k$ is GAS if and only if all eigenvalues of A have norm strictly less than one; i.e. the *spectral radius* $\rho(A)$ of A is less than one. In this, we call the matrix A *stable* (or *Schur stable*).

Here we give a different characterization of stable matrices that relates to semidefinite programming (SDP) and is much more useful than the eigenvalue characterization when we go beyond simple stability questions (e.g. "robust" stability or "stabilizability").

[1] The precise definition of global asymptotic stability requires a second condition (the so-called *stability in the sense of Lyapunov*), but the distinction is a non-issue for linear systems.

Theorem 1.

$$\text{The dynamical system } x_{k+1} = Ax_k \text{ is GAS}$$
$$\Leftrightarrow$$
$$\exists P \in S^{n\times n}, \text{ s.t. } P \succ 0 \text{ and } P \succ A^T P A. \tag{1}$$

(Note that given A, the search for the matrix P is an SDP.)

Proof: The proof is based on the fundamental concept of a *Lyapunov function*.

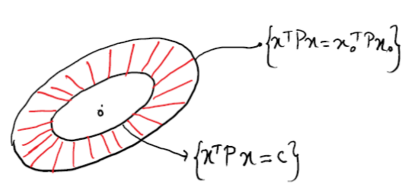
Consider the (Lyapunov) function $V(x) = x^T P x$. We have $V(0) = 0$ and $V(x) > 0\ \forall x \neq 0$ (because of (1)). Condition (1) also implies

$$V(Ax) < V(x),\ \forall x \neq 0.$$

In other words, the function V monotonically decreases along all trajectories of our dynamical system:

Take any x_0 and consider the sequence $\{V(x_k)\}$ of the function V evaluated on the trajectory starting at x_0. Since $\{V(x_k)\}$ is nonnegative and lower bounded, it converges to some $c \geq 0$. If $c = 0$, $V(x_k) \to 0$ implies that $x_k \to 0$. This is because V is only zero at zero and it is radially unbounded which implies that its sublevel sets are compact.
It remains to show that we cannot have $c > 0$. Indeed, if $c > 0$, then the trajectory starting at x_0 would forever be contained (because of (1)) in the compact set

$$S := \{x |\ c \leq x^T P x \leq x_0^T P x_0\}.$$



Let

$$\delta := \min_{x \in S} V(x) - V(Ax).$$

Since the objective is continuous and positive definite and since S is compact, δ exists and is positive. Therefore, in each iteration $V(x_k)$ decreases by at least δ. But this means that $\{V(x_k)\} \to -\infty$ which contradicts nonnegativity of V.

To prove the converse, suppose the dynamical system $x_{k+1} = Ax_k$ is GAS. Consider the quadratic function

$$\begin{aligned} V(x) &= \sum_{j=0}^{\infty} ||A^j x||^2 \\ &= \sum_{j=0}^{\infty} x^T A^{jT} A^j x \\ &= x^T \left(\sum_{j=0}^{\infty} A^{jT} A^j \right) x, \end{aligned}$$

which is well-defined since $\rho(A) < 1$. The function $V(x)$ is clearly positive definite since even its first term $||x||^2$ is positive definite. We also have

$$V(Ax) - V(x) = \sum_{j=1}^{\infty} ||A^j x||^2 - \sum_{j=0}^{\infty} ||A^j x||^2 = -||x||^2 < 0.$$

Letting $P = \sum_{j=0}^{\infty} A^{jT} A^j$, we have indeed established that $P \succ 0$ and $A^T P A \prec P$. $\square$

Remark: One can derive the same result in *continuous time*. The origin of the differential equation

$$\dot{x} = Ax$$

is GAS iff $\exists P \in S^{n \times n}$ s.t. $P \succ 0$ and $A^T P + PA \prec 0$. These LMIs imply that $V(x) = x^T P x$ satisfies $\dot{V}(x) = \langle \nabla V(x), \dot{x} \rangle < 0,\ \forall x \neq 0$.

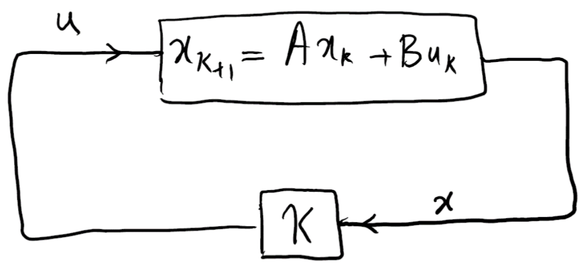
2 Stabilization with state feedback

We now consider a scenario where we can design the matrix A (under some restrictions) in such a way that the dynamical system $x_{k+1} = Ax_k$ becomes GAS. Let us once again pose a

concrete problem.
Given matrices $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times k}$, does there exist a matrix $K \in \mathbb{R}^{k\times n}$ such that

$$A + BK$$

is stable; i.e., such that $\rho(A + BK) < 1$?



This is a basic problem in control theory. In the controls jargon, we would like to design a linear controller $u = Kx$ which is in feedback with a "plant" $x_{k+1} = Ax_k + Bu_k$ and makes the closed-loop system stable:

$$\begin{aligned} x_{k+1} &= Ax_k + Bu_k \\ &= Ax_k + BKx_k \\ &= (A + BK)x_k. \end{aligned}$$

From our discussion before, $A + BK$ will be stable iff $\exists P \succ 0$ such that

$$(A + BK)^T P(A + BK) \prec P.$$

Unfortunately, this is not an SDP since the matrix inequality is not linear in the decision variables P and K. (It is in fact "bilinear", meaning that it becomes linear if you fix either P or K and search for the other.)

Nevertheless we are going to show an exact reformulation of this problem as an SDP by applying a few nice tricks!
Let's recall our Schur complement theorem first.

Lemma 1. *Consider a block matrix* $X = \begin{pmatrix} A & B \\ B^T & C \end{pmatrix}$ *and let* $S := C - B^T A^{-1} B$.

- *If* $A \succ 0$, *then* $X \succeq 0 \Leftrightarrow S \succeq 0$.

- $X \succ 0 \Leftrightarrow A \succ 0$ *and* $S \succ 0$.

In the previous lecture, we proved the first part of the theorem. The proof of the second part is very similar.

<u>Trick 1:</u> $A + BK$ is stable $\Leftrightarrow$ $A^T + K^T B^T$ is stable.

More generally, a matrix E is stable iff E^T is stable. This is clear since E and E^T have the same eigenvalues. It's also useful to see how the Lyapunov functions for the dynamics defined by E and E^T relate. Suppose we have $P \succ 0$ and $E^T P E \prec P$ (i.e., $V(x) = x^T P x$ is a Lyapunov function for $x_{k+1} = E x_k$) then by applying the Schur complement twice (starting from different blocks) we get

$$E^T P E \prec P \Leftrightarrow \left[\begin{array}{c|c} P^-1 & E \\ \hline E^T & P \end{array}\right] \succ 0 \Leftrightarrow P^{-1} - E P^{-1} E^T \succ 0.$$

Hence $V(x) = x^T P^{-1} x$ is our desired Lyapunov function for the dynamics $x_{k+1} = E^T x_k$. Note that P^{-1} exists and is postiive definite as eigenvalues of P^{-1} are the reciprocal eigenvalues of P. In summary, we will instead be looking for a Lyapunov function for the dynamics defined by $A^T + K^T B^T$.

<u>Trick 2:</u> Schur complements again.

We have

$$P - (A^T + K^T B^T)^T P (A^T + K^T B^T) \succ 0$$

$$\Updownarrow$$

$$\left[\begin{array}{c|c} P & P(A^T + K^T B^T) \\ \hline (A^T + K^T B^T)^T P & P \end{array}\right] \succ 0$$

$$\Updownarrow$$

$$\left[\begin{array}{c|c} P & PA^T + PK^T B^T \\ \hline AP + BKP & P \end{array}\right] \succ 0.$$

<u>Trick 3:</u> A change of variables.

Let $L = KP$. Then we have

$$\left[\begin{array}{c|c} P & PA^T + L^T B^T \\ \hline AP + BL & P \end{array}\right] \succ 0.$$

This is now a linear matrix inequality (LMI) in P and L! We can solve this semidefinite program for P and L and then we can simply recover the controller K as

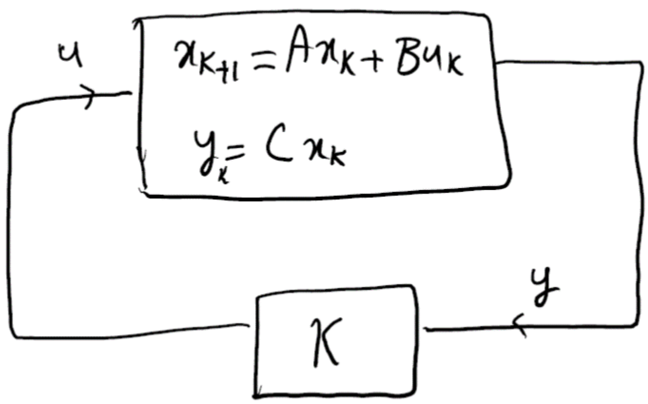
$$K = LP^{-1}.$$

3 Stabilization with output feedback

Here is another concrete problem of similar flavor.
Given matrices $A \in \mathbb{R}^{n\times n}, B \in \mathbb{R}^{n\times k}, C \in \mathbb{R}^{r\times n}$, does there exist a matrix $K \in \mathbb{R}^{k\times r}$ such that

$$A + BKC$$

is stable, i.e., such that $\rho(A + BKC) < 1$?



The problem is similar to the previous one, except that instead of feeding back the full state x to the controller K, we feeback an output y which is obtained from a (possibly non-invertible) linear mappping C from x. For this reason, the question of existence of a K that makes the closed-loop system (i.e., $A + BKC$) stable is known as the "stabilization with output feedback" problem.

Can this problem be formulated as an SDP via some "tricks"? We don't know! In fact, the exact complexity of this problem is regarded as a "major open problem in systems and control theory" [2].

If one in addition requires lower and upper bounds on the entries of the controller

$$l_{ij} \leq K_{ij} \leq u_{ij}$$

then Blondel and Tsitsiklis [3] have shown that the problem is NP-hard. In absence of these constraints, however, the complexity of the problem is unknown.
You will deservingly receive an A in ORF 523 if you present a polynomial-time algorithm for this problem or show that it is NP-hard.

4 Take-away message

It is not always obvious to determine whether a problem admits a formulation as a semidefinite program. A problem which looks non-convex in its original formulation can sometimes be formulated as an SDP via a sequence of transformations. In recent years, a lot of research has been done to understand these transformations more systematically. While some progress has been made, a complete answer is still out of reach. For example, we do not currently have a full answer to the following basic geometric question: Under what conditions can a convex set be written as the feasible set of an SDP or the projection of the feasible set of a higher dimensional SDP?

5 Eigenvalue and matrix norm optimization

Semidefinite programming is often the right tool for optimization problems involving eigenvalues of matrices or matrix norms. This is hardly surprising in view of the fact that positive semidefiniteness of a matrix has a direct characterization in terms of eigenvalues.

5.1 Maximizing the minimum eigenvalue

Let $A(x) = A_0 + \sum_{i=0}^{m} x_i A_i$, where $A_i \in S^{n\times n}$ are given. Consider the problem

$$\max_{x}.\ \lambda_{\min} A(x).$$

This problem can be written as the SDP

$$\begin{aligned} &\max_{x,t}.\ t \\ &\text{s.t. } tI \preceq A_0 + \sum_i x_i A_i \end{aligned}$$

This is simply because for a general matrix $B \in S^{n\times n}$ we have the relation

$$\lambda_i(B + \alpha I) = \lambda_i(B) + \alpha,$$

for the i^{th} eigenvalue λ_i. This is equally easy to see from the definition of eigenvalues as roots of the characteristic polynomial.

5.2 Minimizing the maximum eigenvalue

Similarly with $A(x)$ defined as before, we can formulate the problem

$$\min_{x}. \ \lambda_{\max} A(x)$$

as the SDP

$$\begin{aligned} &\min_{t,x}. \ t \\ &\text{s.t. } A(x) \preceq tI. \end{aligned}$$

Question for you:
Can we minimize the second largest eigenvalue of $A(x)$ using SDP?
(Hint: Convince yourself that if you could do this, you could find for example the largest independent set in a graph using SDP.
Hint on the hint: write the problem as an SDP with a rank-1 constraint.)

5.3 Minimizing the spectral norm

Given $A_0, A_1, \ldots, A_m \in \mathbb{R}^{n \times p}$, let $A(x) := A_0 + \sum_{i=1}^{n} x_i A_i$ and consider the optimization problem

$$\min_{x \in \mathbb{R}^m} ||A(x)||.$$

Here, the norm $||.||$ is the induced 2-norm (aka the spectral norm). We have already shown that $||B|| = \sqrt{\lambda_{\max}(B^T B)}$ for any matrix B.
Let us minimize the square of the norm instead, which does not change the optimal solution.

So our problem is

$$\begin{aligned}
&\min_{t,x}.\ t \\
&\text{s.t. } ||A(x)||^2 \leq t \\
&\Updownarrow \\
&\min_{t,x}.\ t \\
&\text{s.t. } A^T(x)A(x) \preceq tI_p \\
&\Updownarrow \\
&\min_{t,x}.\ t \\
&\text{s.t. } \left[\begin{array}{c|c} I_n & A(X) \\ \hline A^T(x) & tI_p \end{array}\right] \succeq 0.
\end{aligned}$$

This is an SDP.

Practice: With $A(x)$ defined as before, formulate the minimization of the Frobenius norm as an SDP:

$$\min_{x\in\mathbb{R}^m}.\ ||A(x)||_F.$$

Notes

Further reading for this lecture can include Chapter 4 of [1] and chapters 2 and 9 of [4].

References

[1] A. Ben-Tal and A. Nemirovski. *Lectures on Modern Convex Optimization: Analysis, Algorithms, and Engineering Applications*, volume 2. SIAM, 2001.

[2] V. Blondel, M. Gevers, and A. Lindquist. Survey on the state of systems and control. *European Journal of Control*, 1(1):5–23, 1995.

[3] V. Blondel and J.N. Tsitsiklis. NP-hardness of some linear control design problems. *SIAM Journal on Control and Optimization*, 35(6):2118–2127, 1997.

[4] M. Laurent and F. Vallentin. *Semidefinite Optimization*. 2012.