PRINCETON UNIVERSITY
ORFE

Instructor:
Amir Ali Ahmadi
Spring 2016

TA: Georgina Hall

Today: Some applications of SDP:

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

Stability of a linear system

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

$$x_{k+1} = A x_k,$$

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 is called "global asymptotic stability" (GAS).

The choice of $x=0$ as the "attractor" is arbitrary here. If the system had 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 a 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 eigen values of A have magnitude 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.

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

Thm. The dynamical system $x_{k+1} = Ax_k$ is GAS

$$\Longleftrightarrow$$

$$\exists P \in S^{n\times n}, \text{ s.t. } \quad P \succ 0 \quad \text{and} \quad A^TPA \prec P. \qquad (1)$$

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

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

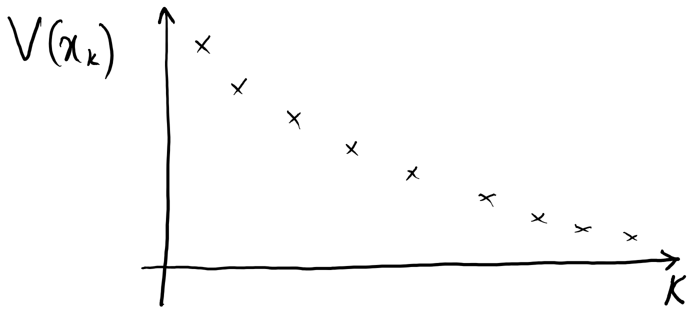
Consider the (Lyapunov) function $\quad V(x) = x^TPx.$

We have $V(0)=0$ and $V(x)>0 \quad \forall x \neq 0$ (because of (1)).

Condition (1) also implies:

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

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



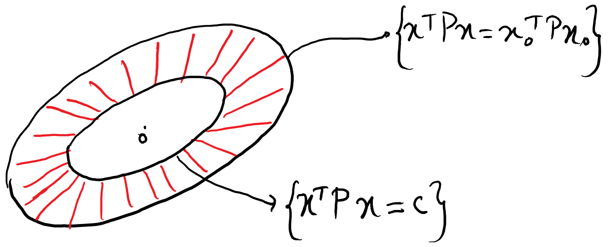
Take any x_0 and consider the sequence $V(x_k)$ of the function V evaluated on the trajectory starting at x_0.

Since $\{V(x_k)\}$ is positive and lower bounded, it converges to some $c \geq 0$.

If $c=0$, $V(x_k) \to 0$ implies that $x_k \to 0$ (because V has compact sublevel sets and is only zero at zero) and we would be done.

We claim that we cannot have $c>0$. Indeed if $c>0$, then the trajectory starting at x_0 would forever be traversing (b/c of ①) in the compact set

$$S := \{x \mid x^T P x \geq c\} \cap \{x \mid x^T P x \leq x_0^T P x_0\}.$$



Let $\delta := \min_{x \in S} V(x) - V(Ax)$. Since the objective is continuous and negative definite and since S is compact, δ exists and is negative.

Therefore, in each iteration $V(x_k)$ decreases by at least $|\delta|$.

This, however, implies 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

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

which is well-defined since $\rho(A) < 1$. The function $V(x)$ is clearly positive definite since it dominates $\|x\|^2$. 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 $P = \sum_{j=0}^{\infty} A^{j^T} A^j$, we have indeed established that $P \succ 0$ and $A^T P A \prec P$.

□

<u>Remark</u>. One can derive the same result in <u>continuous time</u>. 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 $\dot{V}(x) = \langle \nabla V(x), \dot{x} \rangle < 0$, $\forall x \neq 0$.

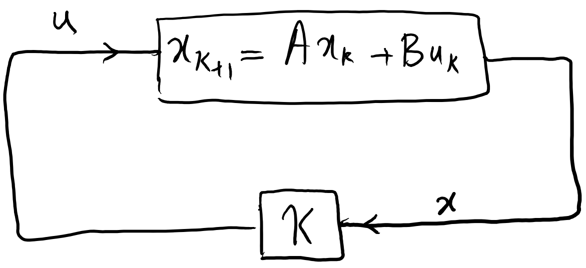
Stabilization with state feedback

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

Given $A \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{n \times k}$, does there exist $K \in \mathbb{R}^{k \times n}$ such that

$$A + BK$$

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



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

$$x_{k+1} = A x_k + B u_k = A x_k + B K x_k$$

$$= (A + BK) x_k.$$

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!

<u>Trick 1</u>. $A+BK$ is stable $\Leftrightarrow$ $A^T+K^TB^T$ 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 observe what's a Lyapunov function for the dynamical system $x_{k+1}=E^Tx_k$. Suppose we have $P \succ 0$ and $E^TPE \prec P$ (i.e., $V(x)=x^TPx$ is a Lyapunov function for $x_{k+1}=Ex_k$), then by applying the Schur complement twice (starting from different blocks) we get

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

Hence, $V(x)=x^TP^{-1}x$ is our desired Lyapunov function for the E^T dynamics. Note that P^{-1} exists and is positive definite as eigenvalues of P^{-1} are one over eigenvalues of P.

So we will instead look for a Lyapunov function for $A^T + K^T B^T$.

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

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

$$\Updownarrow$$

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

$$\Updownarrow$$

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

<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

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

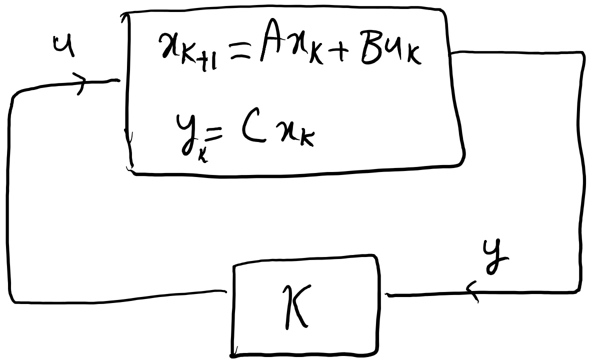
Stabilization with output feedback

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?



This problem is similar to our previous one, except that instead of feeding back the full state x to the controller K, we feed back an output y which is obtained from a (possibly non-invertible) linear mapping 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 open. It is regarded as a "major open problem in systems and control theory" [BGL95].

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 [BT97] have shown that the problem is NP-hard. In absence of these constraints, however, the complexity of the problem is unknown.

You would deservingly receive an A in ORF523 if you present a polynomial-time algorithm or show that it is NP-hard.

Take-away message

It is not always obvious to see when a problem admits a formulation as a semidefinite program. More concretely, we do not currently have a full answer to the following 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 an SDP?

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.

Maximizing the minimum eigenvalue

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

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

This problem can be written as the SDP

$$\max_{x,t} \quad t$$

$$tI \preceq A_0 + \sum x_i A_i$$

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

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

This is easy to see from the definition of eigenvalues as roots of the characteristic polynomial.

Minimizing the maximum eigenvalue

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

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

as the SDP

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

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) largest independent sets in graphs using SDP.)

Minimizing the spectral norm.

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

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

Here, the norm $\|\cdot\|$ is the induced 2-norm (aka the spectral norm). We have already showed 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

$$\begin{aligned} \min_{t,x} \quad & t \\ & \|A(x)\|^2 \leq t \end{aligned}$$

$\Updownarrow$ λ_{max} characterization

$$\begin{aligned} \min_{t,x} \quad & t \\ & A^T(x)A(x) \preceq t I_p \end{aligned}$$

$\Updownarrow$ Schur complements

$$\begin{aligned} \min_{t,x} \quad & t \\ & \left[\begin{array}{c|c} I_n & A(x) \\ \hline A^T(x) & t I_p \end{array}\right] \succeq 0 \end{aligned}$$

This is an SDP.

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

$$\min_{x} \quad \|A(x)\|_F .$$

Notes

Further reading for this lecture can include Chapter 4 of [BN01] and Chapters 2 and 9 of [LV12].

References

[BT97] V. Blondel and J.N. Tsitsiklis. NP-hardness of some linear control design problems. SIAM J. on Control and Optimization, Vol. 35, No. 6, 1997.

[BGL95] V. Blondel, M. Gevers, and A. Lindquist. Survey on the state of systems and control. European J. on Control, 1995.

[BN01] A. Ben-Tal and A. Nemirovski. Lecture Notes on Modern Convex Optimization. MPS/SIAM Series on Optimization, 2001.

[LV12] M. Laurent and F. Vallentin. Lecture Notes on Semidefinite Optimization, 2012.