ECE285: Semidefinite and sum-of-squares optimization **Winter 2023**

Lecture 9: Applications of SDPs in Control

Lecturer: Yang Zheng *Scribe: Yang Zheng*

Disclaimer: *These notes have not been subjected to the usual scrutiny reserved for formal publications. Any typos should be sent to* `zhengy@eng.ucsd.edu`.

Learning goals:

1. Stability of linear systems and Lyapunov functions
2. Stabilization of linear systems
3. Eigenvalue and matrix norm minimization

1 Stability of linear systems

1.1 Lyapunov stability of linear time-invariant systems

Consider a discrete-time linear dynamical system

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

where $x_k \in \mathbb{R}^n$ is the state of system, evolving over time k from an initial condition x_0. One natural and important question is to ask when $x_k \to 0$ as $k \to \infty, \forall x_0 \in \mathbb{R}^n$. Then, we call system (1) is asymptotically stable. From linear algebra, it is well-known that x_k converges to zero for all initial conditions if and only if the spectral radius of A is smaller than one, i.e., all the eigenvalues of A, $|\lambda_i(A)| < 1, i = 1, \ldots, n$. In this case, we say system (1) or matrix A Schur stable.

This spectral characterization is very useful. Here, we give another different characterization based on SDPs, which is convenient when we go beyond simple stability. The basic idea is to consider a generalization of the notion of energy, often known as Lyapunov functions. A Lyapunov function is a particular function of the state x_k, which satisfies the property that it decreases monotonically along any trajectories of the system (1). It turns out that for linear systems, the stability is equivalent to the existence of quadratic Lyapunov functions $V(x_k) = x_k^{\mathsf{T}} P x_k$. We have the following result.

Theorem 9.1. *Given a matrix $A \in \mathbb{R}^{n\times n}$, the following conditions are equivalent:*

1. *A is stable, i.e., all its eigenvalues satisfy $|\lambda_i(A)| < 1, i = 1, \ldots, n$.*
2. *There exists a matrix $P \in \mathbb{S}^n$ such that*

$$P \succ 0,\ A^{\mathsf{T}} P A - P \prec 0.$$

Proof (2) $\Rightarrow$ (1) Consider the eigenvalue $Av = \lambda v$. Then

$$v^*(A^{\mathsf{T}} P A - P)v = (|\lambda|^2 - 1)v^* P v.$$

Since $P \succ 0$, we have $v^* P v > 0$. Considering the fact $A^{\mathsf{T}} P A - P \prec 0$, we have $|\lambda|^2 < 1$.

$(1) \Rightarrow (2)$: Suppose the system $x_{k+1} = Ax_k$ is asymptotically stable. We consider a quadratic function

$$V(x_k) = \sum_{j=k}^{\infty} \|x_j\|_2^2 = x_k^\mathsf{T} \left(\sum_{j=0}^{\infty} (A^j)^\mathsf{T} A^j \right) x_k,$$

which is well-defined since $\rho(A) < 1$. It is easy to verify that $V(x_k) > 0, \forall x_k \neq 0$, and

$$V(x_{k+1}) - V(x_k) = \sum_{j=k+1}^{\infty} \|x_j\|_2^2 - \sum_{j=k}^{\infty} \|x_j\|_2^2 = -\|x_k\|_2^2 < 0, \forall x_k \neq 0.$$

This indeed verifies that $P = \sum_{j=0}^{\infty} (A^j)^\mathsf{T} A^j$ satisfies

$$P \succ 0, \quad A^\mathsf{T} P A - P = -I \prec 0.$$

Another useful perspective for proving stability is to consider a Lypuanov function

$$V(x_k) = x_k^\mathsf{T} P x_k.$$

which satisfies $V(0) = 0, V(x) > 0, \forall x \neq 0$. Furthermore, it is easy to check that

$$V(x_{k+1}) - V(x_k) = V(Ax_k) - V(x_k) = x_k^\mathsf{T}(A^\mathsf{T} P A - P)x_k < 0, \quad \forall x_k \neq 0.$$

In other words, consider any trajectory $x_0, x_1, \ldots, x_k, \ldots$. The function $V(x_k)$ decreases monotonically. Furthermore, since $V(x_k)$ is nonnegative and lower bounded by 0, it must converges to some constant $c \geq 0$. If $c = 0$, $V(x_k) \to 0$ as $k \to \infty$ implies that $x_k \to 0$.

It remains to show that c cannot be strictly positive. Indeed, if $c > 0$, then the trajectory starting at x_0 will stay in the compact region

$$S = \{x \in \mathbb{R}^n \mid c \leq x^\mathsf{T} P x \leq x_0^\mathsf{T} P x_0\}.$$

Now, we let

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

Since the cost function $V(x) - V(Ax)$ is continuous and strictly positive over S. The set S is also compact. Thus δ is strictly positive. Consequently, at each iteration, $V(x_k)$ decreases by at least $\delta > 0$. This implies that $V(x_k) \to -\infty$, which contradicts with the fact that $V(x_k) \geq 0, \forall x_k \in \mathbb{R}^n$.

One can further derive an analogue result for linear systems in continuous time $\dot{x} = Ax$. The system is asymptotically stable (i.e., $\lim_{t\to\infty} x(t) = 0$) if and only if

$$P \succ 0, \quad A^\mathsf{T} P + PA \prec 0.$$

These conditions imply a Lyapunov function $V(x) = x^\mathsf{T} P x$ satisfying $V(x) > 0, \forall x \neq 0$ and $\dot{V}(x) < 0, \forall x \neq 0$.

1.2 Quadratic stability of time-varying linear systems

We consider a time-varying linear system

$$x_{k+1} = A_k x_k, \quad A_k \in \{A_1, \ldots, A_m\}. \tag{2}$$

We ask whether all trajectories of system (2) converge to zero as $k \to \infty$. This problem is much harder than the problem in the previous subsection. A simple sufficient condition is the existence of a quadratic function $V(x_k) = x_k^\mathsf{T} P x_k, P \succ 0$ that decreases along every trajectory. In this case, we say (2) is quadratically stable and we call V as a quadratic Lyapunov function.

It is easy to verify that

$$V(x_{k+1}) - V(x_k) = x_k^\mathsf{T}(A_k^\mathsf{T} P A_k - P)x_k.$$

A sufficient and necessary condition for quadratic stability of (2) is

$$P \succ 0, \quad A_k^\mathsf{T} P A_k - P \prec 0, \forall A_k \in \{A_1, \ldots, A_m\}.$$

This is equivalent to

$$P \succ 0, \quad A_k^\mathsf{T} P A_k - P \prec 0, \quad \forall k = 1, \ldots m,$$

which is a semidefinite program.

2 Stabilization with state feedback

We now consider the case of linear systems with control input u_k,

$$x_{k+1} = Ax_k + Bu_k, \tag{3}$$

starting from an initial state $x_0 \in \mathbb{R}^n$. The problem here is to choose the control input u_k at each time step k such that the system behavior x_k converges to zero globally. Here, we consider static state feedback $u_k = Kx_k$. Then the closed-loop system becomes

$$x_{k+1} = Ax_k + Bu_k = (A + BK)x_k.$$

Given matrices $A \in \mathbb{R}^{n\times n}, B \in \mathbb{R}^{n\times m}$, find a matrix $K \in \mathbb{R}^{m\times n}$ such that $A+BK$ is stable, i.e., $\rho(A+BK) < 1$. This problem is equivalent to solve

$$P \succ 0, \qquad (A + BK)^\mathsf{T} P(A + BK) - P \prec 0. \tag{4}$$

Unfortunately, this formulation is not convex since it is bilinear in (P, K), meaning that it becomes linear if fixing either P or K and searching for the other. This problem is actually fairly complicated in its present form.

We will show an exact reformulation of the problem above into an SDP by introducing a change of variables. Let us first recall the Schur complement theorem.

Lemma 9.1. *Consider a block symmetric matrix*

$$X = \begin{bmatrix} A & B \\ B^\mathsf{T} & C \end{bmatrix}.$$

We have

- *If $A \succ 0$, then $X \succeq 0$ if and only if $C - B^\mathsf{T} A^{-1} B \succeq 0$.*
- *$X \succ 0$ if and only if $A \succ 0$ and $C - B^\mathsf{T} A^{-1} B \succ 0$.*

Step 1: By the Schur complement theorem, (4) is equivalent to

$$\begin{bmatrix} P & (A + BK)^\mathsf{T} P \\ P(A + BK) & P \end{bmatrix} \succ 0.$$

Step 2: define $Q = P^{-1}$. The condition above is equivalent to

$$\begin{bmatrix} Q & Q(A + BK)^\mathsf{T} \\ (A + BK)Q & Q \end{bmatrix} \succ 0.$$

Step 3: A change of variables. Define $Y = KQ$, and we have

$$\begin{bmatrix} Q & QA^\top + Y^\top B^\top \\ AQ + BY & Q \end{bmatrix} \succ 0,$$

which is linear in the new variables (Q, Y). This is indeed an SDP. After solving it, the controller is recovered as $K = YQ^{-1}$.

Semidefinite optimization techniques have become central in many analysis and design of control systems. A great collection of problems in control that can be handled by semidefinite optimization can be found in [5].

It is not always obvious to see whether a problem admit a reformulation as an SDP. Many non-convex problems in their natural forms can be reformulated into SDPs via a sequence of "tricks" (which often happens in control). Still, a systematic understanding of these tricks is far from complete. As mentioned in the previous lecture, the following basic geometric question is not fully understood: when can a convex set be written as the feasible region of an SDP or the projection of the feasible region of a higher dimensional SDP?

2.1 Some "simple" control problems that are hard

Similar to stability and stabilization problems above, many other control problems admit efficient SDP reformulations; see [5] for an excellent survey. However, there are some seemingly "benign" control problems that are fundamentally harder to solve. Here are a few of them [3]

- *Static output stabilization*: Given matrices $A \in \mathbb{R}^{n\times n}, B \in \mathbb{R}^{n\times m}, C \in \mathbb{R}^{p\times n}$, find a matrix $K \in \mathbb{R}^{m\times p}$ such that $A + BKC$ is stable.
- *State feedback with bounds*: Given matrices $A \in \mathbb{R}^{n\times n}, B \in \mathbb{R}^{n\times m}$, and bounds $\bar{k}_{ij}, \underline{k}_{ij}$, find a matrix $K \in \mathbb{R}^{m\times p}$ with $\underline{k}_{ij} \leq k_{ij} \leq \bar{k}_{ij}$such that $A + BK$ is stable.
- *Decentralized control*: Given matrices $A \in \mathbb{R}^{n\times n}, B \in \mathbb{R}^{n\times m}$, and a binary matrix $S \in \{0, 1\}^{m\times n}$, find a matrix $K \in \mathbb{R}^{m\times p}$ with $k_{ij} = 0$ if $S_{ij} = 0$ such that $A + BK$ is stable.

3 Eigenvalue and matrix norm optimization

Semidefinite optimization is often the right method for problems involving eigenvalues of matrices or matrix norms. This is not surprising considering the fact that positive semidefiniteness of a matrix is a direct characterization of eigenvalues.

3.1 Maximizing the minimum eigenvalue

Define $A(x) = A_0 + x_1A_1 + \ldots + x_mA_m$, where $A_i \in \mathbb{S}^n$ are problem data. We consider the problem of maximizing the minimum eigenvalue

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

This problem is equivalent to an SDP

$$\begin{aligned} \max_{x,t} \quad & t \\ \text{subject to} \quad & A_0 + x_1A_1 + \ldots + x_mA_m - tI \succeq 0. \end{aligned}$$

This is simply due to the fact that

$$\lambda_i(B + \alpha I) = \lambda_i(B) + \alpha, \quad \forall B \in \mathbb{S}^n, \alpha \in \mathbb{R}.$$

3.2 Minimizing the maximum eigenvalue

Similarly, we consider the problem of minimizing the maximum eigenvalue

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

where $A(x) = A_0 + x_1 A_1 + \ldots + x_m A_m$, with $A_i \in \mathbb{S}^n$. This problem is equivalent to an SDP

$$\begin{aligned} \min_{x,t} \quad & t \\ \text{subject to} \quad & A_0 + x_1 A_1 + \ldots + x_m A_m - tI \preceq 0. \end{aligned}$$

3.3 Minimizing the spectral norm

Here we define $A(x) = A_0 + x_1 A_1 + \ldots + x_m A_m$, with matrices $A_i \in \mathbb{R}^{n\times p}$. We consider the optimization problem

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

Here, $\|\cdot\|$ denotes the induced-2 norm, i.e., the spectral norm, $\|B\| = \sqrt{\lambda_{\max}(B^\mathsf{T} B)}$ for any matrix B. Let us minimize the square of the norm, which does not change the optimal solution. Then, we have

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

By definition, this is equivalent to

$$\begin{aligned} \min_{x,t} \quad & t \\ \text{subject to} \quad & \lambda_{\max}(A(x)^\mathsf{T} A(x)) \leq t. \end{aligned}$$

Similarly to the previous section, we now have

$$\begin{aligned} \min_{x,t} \quad & t \\ \text{subject to} \quad & A(x)^\mathsf{T} A(x) - tI_p \preceq 0. \end{aligned}$$

We now apply the Schur complement theorem, leading to

$$\begin{aligned} \min_{x,t} \quad & t \\ \text{subject to} \quad & \begin{bmatrix} I_n & A(x) \\ A(x)^\mathsf{T} & tI_p \end{bmatrix} \succeq 0, \end{aligned}$$

which is an SDP.

3.4 Minimizing the nuclear norm

The nuclear norm of a matrix $A \in \mathbb{R}^{n\times p}$ is defined as

$$\|A\|_* = \sum_{i=1}^{r} \sigma_i(A),$$

where σ_i is the ith singular value of A and r is the rank of A.

The nuclear norm is alternatively known by several other names including the Schatten 1-norm, the Ky Fan r-norm, and the trace class norm. The nuclear norm is particularly useful in optimization problems involving ranks of matrices (see [6] for an excellent discussion).

In Homework 4, you will prove that the nuclear norm is the dual norm of the spectral norm, i.e.,

$$\begin{aligned} \|X\|_* = \max_{Y \in \mathbb{R}^{n \times p}} \quad & \langle X, Y \rangle \\ \text{subject to} \quad & \|Y\| \leq 1. \end{aligned} \tag{5}$$

Furthermore, in Homework 4, you will prove that the nuclear norm $\|X\|_*$ corresponds to the optimal value of the primal-dual pair of SDPs

$$\begin{aligned} \max_{Y \in \mathbb{R}^{n \times p}} \quad & \langle X, Y \rangle \\ \text{subject to} \quad & \begin{bmatrix} I_n & Y \\ Y^\mathsf{T} & I_p \end{bmatrix} \succeq 0, \end{aligned} \tag{6}$$

and

$$\begin{aligned} \min_{W_1, W_2} \quad & \frac{1}{2}\text{trace}(W_1) + \frac{1}{2}\text{trace}(W_2) \\ \text{subject to} \quad & \begin{bmatrix} W_1 & X \\ X^\mathsf{T} & W_2 \end{bmatrix} \succeq 0. \end{aligned} \tag{7}$$

Remark 9.1. *This dual norm characterization can be viewed as a generalization of vector norms in $\mathbb{R}^n$ in the following sense*

- *When $n = 1$, i.e., $X \in \mathbb{R}^{m \times 1}$ is a column vector, it is easy to verify that*

  $$\sigma_{\max}(X) = \sqrt{X_1^2 + \ldots + X_m^2}.$$

  Thus both the spectral norm $\|X\|$ and the nuclear norm $\|X\|_*$ correspond to the standard l_2 norm of vectors in $\mathbb{R}^m$. Thus (5) is reduced to the case where the dual norm of l_2 norm is itself.

- *When $m = n$ and X is a diagonal matrix of the form*

  $$X = \begin{bmatrix} x_1 & & \\ & \ddots & \\ & & x_n \end{bmatrix},$$

  it is easy to verify that

  $$\sigma_{\max}(X) = \max_{i=1,\ldots,n} \; |x_i|, \qquad \|X\|_* = \sum_{i=1}^{n} |x_i|.$$

  Now, (5) is reduced to the case where the dual norm of l_1 norm in $\mathbb{R}^n$ is the l_∞ norm.

Notes

The preparation of this lecture was based on [1, Lecture 10]. Further reading for this lecture can refer to [2, Chapter 2.2] and [4, Chapter 4].

References

[1] Amir Ali Ahmadi. ORF523: Convex and Conic Optimization, Spring 2021.

[2] Grigoriy Blekherman, Pablo A Parrilo, and Rekha R Thomas. *Semidefinite optimization and convex algebraic geometry.* SIAM, 2012.

[3] Vincent Blondel and John N Tsitsiklis. Np-hardness of some linear control design problems. *SIAM journal on control and optimization*, 35(6):2118–2127, 1997.

[4] Stephen Boyd, Stephen P Boyd, and Lieven Vandenberghe. *Convex optimization*. Cambridge university press, 2004.

[5] Stephen Boyd, Laurent El Ghaoui, Eric Feron, and Venkataramanan Balakrishnan. *Linear matrix inequalities in system and control theory*. SIAM, 1994.

[6] Benjamin Recht, Maryam Fazel, and Pablo A Parrilo. Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization. *SIAM review*, 52(3):471–501, 2010.