Outline

• Motivation

• Some Linear Algebra

• A Tiny Bit of Optimization and Duality

• Linear Matrix Inequalities

• Semidefinite Programming Problems

• Some Examples

• Conclusion
Motivation

• Converse Lyapunov function theorems are not constructive

• Basic idea for Lyapunov function synthesis
  - Select Lyapunov function structure (e.g. quadratic, polynomial, piecewise quadratic, ...)
  - Parameterize Lyapunov function candidates
  - Find values of parameters to satisfy Lyapunov conditions

• Many Lyapunov synthesis problems can be formulated as Semidefinite Programming (SDP) problems.

• SDP has broad applications in many other fields as well
Real Symmetric Matrices

- $S^n$: set of real symmetric matrices

- All eigenvalues are real

- There exists a full set of orthogonal eigenvectors

- Spectral decomposition: If $A \in S^n$, then $A = Q \Lambda Q^T$, where $\Lambda$ diagonal and $Q$ is unitary.
Positive Semidefinite Matrices I

- $A \in S^n$ is called positive semidefinite (p.s.d.), denoted by $A \succeq 0$, if $x^T A x \geq 0, \forall x \in \mathbb{R}^n$

- $A \in S^n$ is called positive definite (p.d.), denoted by $A \succ 0$, if $x^T A x > 0$ for all nonzero $x \in \mathbb{R}^n$

- $S^n_+$: set of all p.s.d. (symmetric) matrices

- $S^n_{++}$: set of all p.d. (symmetric) matrices

- p.s.d. or p.d. matrices can also be defined for non-symmetric matrices.

  e.g.: \[
  \begin{bmatrix}
  1 & 1 \\
  -1 & 1 
  \end{bmatrix}
  \]

- We assume p.s.d. and p.d. are symmetric (unless otherwise noted)

- Notation: $A \succeq B$ (resp. $A \succ B$) means $A - B \in S^n_+$ (resp. $A - B \in S^n_{++}$)
Positive Semidefinite Matrices II

• Other equivalent definitions for symmetric p.s.d. matrices:
  - All $2^n - 1$ principal minors of $A$ are nonnegative
  - All eigs of $A$ are nonnegative
  - There exists a factorization $A = B^T B$

• Other equivalent definitions for p.d. matrices:
  - All $n$ leading principal minors of $A$ are positive
  - All eigs of $A$ are strictly positive
  - There exists a factorization $A = B^T B$ with $B$ square and nonsingular.
Positive Semidefinite Matrices III

• Useful facts:
  - If \( T \) nonsingular, \( A > 0 \) ⇔ \( T^T A T > 0 \); and \( A \geq 0 \) ⇔ \( T^T A T \geq 0 \)

- Inner product on \( \mathbb{R}^{m \times n} \): \( \langle A, B \rangle \triangleq \text{tr}(A^T B) \triangleq A \bullet B \).
Positive Semidefinite Matrices IV

- For $A, B \in S^n_+$, $\text{tr}(AB) \geq 0$

- For any symmetric $A \in S^n$,
  
  $\lambda_{\min}(A) \geq \mu \iff A \succeq \mu I$ \hspace{1cm} and \hspace{1cm} $\lambda_{\max}(A) \leq \beta \iff A \preceq \beta I$


Convex Set

Recall from Lecture 3.

- **Convex Set**: A set $S$ is convex if

  $$x_1, x_2 \in S \Rightarrow \alpha x_1 + (1 - \alpha) x_2 \in S, \forall \alpha \in [0, 1]$$

- **Convex combination of** $x_1, \ldots, x_k$:

  $$\left\{ \alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_k x_k : \alpha_i \geq 0, \text{ and } \sum_i \alpha_i = 1 \right\}$$

- **Convex hull**: $\text{co} \{ S \}$ set of all convex combinations of points in $S$
Convex Cone

• A set $S$ is called a cone if $\lambda > 0$, $x \in S \Rightarrow \lambda x \in S$.

• Conic combination of $x_1$ and $x_2$:
  
  $x = \alpha_1 x_1 + \alpha_2 x_2$ with $\alpha_1, \alpha_2 \geq 0$

• Convex cone:
  1. a cone that is convex

  2. equivalently, a set that contains all the conic combinations of points in the set
Positive Semidefinite Cone

- The set of positive semidefinite matrices (i.e. $S_n^+$) is a convex cone and is referred to as the positive semidefinite (PSD) cone.

- Recall that if $A, B \in S_n^+$, then $tr(AB) \geq 0$. This indicates that the cone $S_n^+$ is acute.
Operations that preserve convexity I

- Intersection of possibly infinite number of convex sets:
  - e.g.: polyhedron:
  - e.g.: PSD cone:
Operations that preserve convexity II

- **Affine mapping** $f : \mathbb{R}^n \to \mathbb{R}^m$ (i.e. $f(x) = Ax + b$)
  
  - $f(X) = \{f(x) : x \in X\}$ is convex whenever $X \subseteq \mathbb{R}^n$ is convex
  - E.g.: Ellipsoid: $E_1 = \{x \in \mathbb{R}^n : (x - x_c)^T P(x - x_c) \leq 1\}$ or equivalently $E_2 = \{x_c + Au : \|u\|_2 \leq 1\}$

- $f^{-1}(Y) = \{x \in \mathbb{R}^n : f(x) \in Y\}$ is convex whenever $Y \subseteq \mathbb{R}^m$ is convex
  - E.g.: $\{Ax \leq b\} = f^{-1}(\mathbb{R}^n_+)$, where $\mathbb{R}^n_+$ is nonnegative orthant
Convex Function I

Consider a finite dimensional vector space $\mathcal{X}$. Let $\mathcal{D} \subset \mathcal{X}$ be convex.

**Definition 1 (Convex Function).**

A function $f : \mathcal{D} \rightarrow \mathbb{R}$ is called convex if

$$f(\alpha x_1 + (1 - \alpha) x_2) \leq \alpha f(x_1) + (1 - \alpha) f(x_2), \forall x_1, x_2 \in \mathcal{D}, \forall \alpha \in [0, 1]$$

- $f : \mathcal{D} \rightarrow \mathbb{R}$ is called strictly convex if
  $$f(\alpha x_1 + (1 - \alpha) x_2) < \alpha f(x_1) + (1 - \alpha) f(x_2), \forall x_1 \neq x_2 \in \mathcal{D}, \forall \alpha \in [0, 1]$$

- $f : \mathcal{D} \rightarrow \mathbb{R}$ is called concave if $-f$ is convex
Convex Function II

How to check a function is convex?

• Directly use definition

• First-order condition: If $f$ is differentiable over an open set that contains $\mathcal{D}$, then $f$ is convex over $\mathcal{D}$ iff

\[ f(z) \geq f(x) + \nabla f(x)^T (z - x), \forall x, z \in \mathcal{D} \]

• Second-order condition: Suppose $f$ is twicely differentiable over an open set that contains $\mathcal{D}$, then $f$ is convex over $\mathcal{D}$ iff

\[ \nabla^2 f(x) \succeq 0, \quad \forall x \in \mathcal{D} \]

• Many other conditions, tricks,… see [BV04].
Convex Function III

Examples of convex functions:

- In general, affine functions are both convex and concave
  - e.g.: \( f(x) = a^T x + b, \) for \( x \in \mathbb{R}^n \)
  
  - e.g.: \( f(X) = tr(A^T X) + c = \sum_{i=1}^{m} \sum_{j=1}^{n} A_{ij}X_{ij} + c, \) for \( X \in \mathbb{R}^{m \times n} \)

- All norms are convex
  - e.g. in \( \mathbb{R}^n: \) \( f(x) = \|x\|_p = (\sum_{i=1}^{n} |x_i|^p)^{1/p}; \) \( f(x) = \|x\|_\infty = \max_k |x_k| \)
  
  - e.g. in \( \mathbb{R}^{m \times n}: \) \( f(X) = \|X\|_2 = \sigma_{\text{max}}(X) \)
Lemma 1 (Jensen’s Inequality).

If $f : \mathcal{D} \rightarrow \mathbb{R}$ is a convex function, then for any possible choices of $x_1, \ldots, x_m \in \mathcal{D}$ and $\alpha_1, \ldots, \alpha_m \geq 0$ with $\sum_{l=1}^{m} \alpha_l = 1$ we have

$$f(\alpha_1 x_1 + \cdots + \alpha_m x_m) \leq \alpha_1 f(x_1) + \cdots + \alpha_m f(x_m)$$

- $\alpha_k$’s can be viewed as the probability that $x_k$ occurs.

- Probabilistic version of Jensen’s inequality reads:

$$f(E(X)) \leq E(f(X))$$

for any random variable $X$ taking value in $\mathcal{D}$.
Nonlinear Optimization and Duality I

Nonlinear Optimization:

\[
\begin{aligned}
\text{minimize: } & \quad f_0(x) \\
\text{subject to: } & \quad f_i(x) \leq 0, \ i = 1, \ldots, m \\
& \quad h_i(x) = 0, \ i = 1, \ldots, q
\end{aligned}
\]

- decision variable \( x \in \mathbb{R}^n \), domain \( \mathcal{D} \), referred to as primal problem
- optimal value \( p^* \)
- is called a convex optimization problem if \( f_0, \ldots, f_m \) are convex and \( h_1, \ldots, h_q \) are affine
- typically convex optimization can be solved efficiently
Nonlinear Optimization and Duality II

Associated **Lagrangian**: \( L : \mathcal{D} \times \mathbb{R}^m \times \mathbb{R}^q \to \mathbb{R} \)

\[
L(x, \lambda, \nu) = f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x) + \sum_{i=1}^{q} \nu_i h_i(x)
\]

- weighted sum of objective and constraints functions

- \( \lambda_i \): Lagrangian multiplier associated with \( f_i(x) \leq 0 \)

- \( \nu_i \): Lagrangian multiplier associated with \( h_i(x) = 0 \)
Lagrange dual function: $g : \mathbb{R}^m \times \mathbb{R}^q \to \mathbb{R}$

$$g(\lambda, \nu) = \inf_{x \in D} L(x, \lambda, \nu)$$

$$= \inf_{x \in D} \left\{ f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x) + \sum_{i=1}^{q} \nu_i h_i(x) \right\}$$

- $g$ is concave, can be $-\infty$ for some $\lambda, \nu$

- **Lower bound property:** If $\lambda \succeq 0$ (elementwise), then $g(\lambda, \nu) \leq p^*$
Nonlinear Optimization and Duality IV

Lagrange Dual Problem:

\[
\begin{align*}
\text{maximize : } & \quad g(\lambda, \nu) \\
\text{subject to : } & \quad \lambda \succeq 0
\end{align*}
\]

- Find the best lower bound on \( p^* \) using the Lagrange dual function
- a convex optimization problem even when the primal is nonconvex
- optimal value denoted \( d^* \)
- \((\lambda, \nu)\) is called **dual feasible** if \( \lambda \succeq 0 \) and \((\lambda, \nu) \in \text{dom}(g)\)
- Often simplified by making the implicit constraint \((\lambda, \nu) \in \text{dom}(g)\) explicit
- Example: Linear Program and its dual

\[
\begin{align*}
\text{minimize: } & \quad c^T x \\
\text{subject to: } & \quad Ax = b \\
& \quad x \succeq 0
\end{align*}
\]

\[
\begin{align*}
\text{maximize: } & \quad -b^T \nu \\
\text{subject to: } & \quad A^T \nu + c \succeq 0
\end{align*}
\]
Nonlinear Optimization and Duality V

Duality Theorems

- **Weak Duality**: \( d^* \leq p^* \)
  - always hold (for convex and nonconvex problems)
  - can be used to find nontrivial lower bounds for difficult problems

- **Strong Duality**: \( d^* = p^* \)
  - not true in general, but typically holds for convex problems
  - conditions that guarantee strong duality in convex problems are called *constraint qualifications*
  - **Slater’s constraint qualification**: Primal is strictly feasible
Linear Matrix Inequalities I

• **Standard form**: Given symmetric matrices $F_0, \ldots, F_m \in S^n$,

$$F(x) = F_0 + x_1 F_1 + \cdots + x_m F_m \succeq 0$$

is called a *Linear Matrix Inequality* in $x = (x_1, \ldots, x_m)^T \in \mathbb{R}^m$

• The function $F(x)$ is affine in $x$

• The constraint set $\{x \in \mathbb{R}^n : F(x) \succeq 0\}$ is nonlinear but convex
Example 1 (LMI in Standard Form).

Characterize the constraint set: \( F(x) = \begin{bmatrix} x_1 + x_2 & x_2 + 1 \\ x_2 + 1 & x_3 \end{bmatrix} \succeq 0 \)
Linear Matrix Inequalities III

- General Linear Matrix Inequalities (LMI)
  - Let $\mathcal{X}$ be a finite-dimensional real vector space.
  
  - $F : \mathcal{X} \rightarrow S^n$ is an affine mapping from $\mathcal{X}$ to $n \times n$ symmetric matrices
  
  - Then $F(X) \succeq 0$ is called also an LMI in variable $X \in \mathcal{X}$

- Translation to standard form: Choose a basis $X_1, \ldots, X_m$ of $\mathcal{X}$ and represent $X = x_1X_1 + \cdots + x_mX_m$ for any $X \in \mathcal{X}$. For a given affine mapping $F : \mathcal{X} \rightarrow S^n$, we can define $\hat{F} : \mathbb{R}^m \rightarrow S^n$ as

  \[ \hat{F}(x) \triangleq F(X) = F(0) + \sum_{i=1}^{m} x_i[F(X_i) - F(0)] \]

  where $x$ is the coordinate of $X$ w.r.t. the basis $X_1, \ldots, X_m$. 
Example 2.

Find conditions on matrix $P$ to ensure that $V(x) = x^T Px$ is a Lyapunov function for a linear system $\dot{x} = Ax$. 

Lemma 2 (Schur Complement Lemma).

Define $M = \begin{bmatrix} A & B \\ B^T & C \end{bmatrix}$. The following three sets of inequalities are equivalent.

\[ M \succ 0 \iff \begin{cases} A \succ 0 \\ C - B^T A^{-1} B \succ 0 \end{cases} \iff \begin{cases} C \succ 0 \\ A - B C^{-1} B^T \succ 0 \end{cases} \]

• Proof: The lemma follows immediately from the following identities:

\[
\begin{bmatrix} I & 0 \\ -B^T A^{-1} & I \end{bmatrix} \begin{bmatrix} A & B \\ B^T & C \end{bmatrix} \begin{bmatrix} I & -A^{-1} B \\ 0 & I \end{bmatrix} = \begin{bmatrix} A & 0 \\ 0 & C - B^T A^{-1} B \end{bmatrix}
\]

\[
\begin{bmatrix} I & -B C^{-1} \\ 0 & I \end{bmatrix} \begin{bmatrix} A & B \\ B^T & C \end{bmatrix} \begin{bmatrix} I & 0 \\ -C^{-1} B^T & I \end{bmatrix} = \begin{bmatrix} A - B C^{-1} B^T & 0 \\ 0 & C \end{bmatrix}
\]
Schur Complement Lemma II

- The proof of Schur complement lemma also reveals more general relations between the numbers of negative, zero, positive eigenvalues of

  - \( M \) vs. \( A \) and \( C - B^T A^{-1} B \)

  - \( M \) vs. \( C \) and \( A - B C^{-1} B^T \)

- Schur complement lemma is a very useful result to transform nonlinear (quadratic or bilinear) matrix inequalities to linear ones.
Semidefinite Programming I

- **Semidefinite Programming (SDP) Problem**: Optimization problem with linear objective, and Linear Matrix Inequality and linear equality constraints:

\[
\begin{aligned}
\text{minimize:} & \quad c^T x \\
\text{subject to:} & \quad F_0 + x_1 F_1 + \cdots + x_m F_m \succeq 0 \\
& \quad Ax = b
\end{aligned}
\]  

(1)

- Linear *equality* constraint in (1) can be eliminated. So essentially SDP can be viewed as optimizing linear function subject to only LMI constraints.

- SDP is a particular class of convex optimization problem. Global optimal solution can be found efficiently.

- Optimizing nonlinear but convex cost function subject to LMI constraints is also a convex optimization that can often be solved efficiently.
Semidefinite Programming II

Standard forms of SDP in *matrix variable*:

- **SDP Standard Prime Form:**
  
  \[
  \begin{align*}
  \min_{X \in \mathbb{S}^n} : & \quad f_p(X) = C \cdot X \\
  \text{subject to:} : & \quad A_i \cdot X = b_i, i = 1, \ldots, m \\
  & \quad X \succeq 0
  \end{align*}
  \]  
  (2)

- **SDP Dual form:**
  
  \[
  \begin{align*}
  \max_{y \in \mathbb{R}^m} : & \quad f_d(y) = b^T y \\
  \text{subject to:} : & \quad \sum_{i=1}^m y_i A_i \preceq C
  \end{align*}
  \]  
  (3)

  - One can derive the dual from the prime using either standard Lagrange duality method or more specialized Fenchel duality results

  - The dual form (3) is equivalent to (1) (after eliminating the equality constraint $Ax = b$ in (1))
Semidefinite Programming III

• **SDP Weak Duality**: \( f_p(X) \geq f_d(y) \) for any primal and dual feasible \( X \) and \( y \)

• **SDP Strong Duality**: \( f_p(X^*) = f_d(y^*) \) holds under Slater’s condition:

• Many control and optimization problem can be formulated or translated into SDP problems

• Various computationally difficult optimization problems can be effectively approximated by SDP problems (SDP relaxation...)

• We will see some examples after introducing an important technique: \textit{S-procedure}
S-Procedure I

- Many stability/engineering problems require to certify that a given function is sign-definite over certain subset of the space.

- Mathematically, this condition can be stated as follows:

\[
g_0(x) \geq 0 \quad \text{on} \quad \{x \in \mathbb{R}^n | g_1(x) \geq 0, \ldots, g_m(x) \geq 0\} \tag{4}
\]

- Given functions \(g_0, \ldots, g_m\), we want to know whether the condition holds. Sometimes we may also want to find a \(g_0\) satisfying this condition for given \(g_1, \ldots, g_m\).

- Conservative but useful condition: \(\exists\) PSD functions \(s_i(x)\) s.t.

\[
g_0(x) - \sum_i s_i(x)g_i(x) \geq 0, \forall x \in \mathbb{R}^n
\]

This is the so-called **Generalized S-Procedure**
S-Procedure II

Now consider an important special case: \( g_i(x) = x^T G_i x, i = 0, 1, \ldots \) are quadratic functions

- Requirement (4) becomes:

\[
\forall x \in \mathbb{R}^n, \quad x^T G_1 x \geq 0, \ldots, x^T G_k x \geq 0 \quad \Rightarrow \quad x^T G_0 x \geq 0
\]

- Sufficient condition (S-procedure): \( \exists \alpha_1, \ldots, \alpha_m \geq 0 \) with

\[
G_0 \succeq \alpha_1 G_1 + \cdots + \alpha_m G_m
\]

- S-Procedure is lossless if \( m = 1 \) and \( \exists \hat{x} \) s.t. \( \hat{x}^T G_1 \hat{x} > 0 \) (constraint qualification)
Example 3 (Eigenvalue Optimization).

Given symmetric matrices $A_0, A_1, \ldots, A_m$. Let $S(w) = A_0 + \sum_i w_i A_i$. Find weights $\{w_i\}_{i=1}^m$ to minimize $\lambda_{\text{max}}(S(w))$. 
Example 4 (Ellipsoid inequality).

Given $R \in S^n_{++}$, the set $E = \{ x \in \mathbb{R}^n : (x - x_c)^T R (x - x_c) < 1 \}$ is an ellipsoid with center $x_c$. Find the point in $E$ that is the closet to the origin.
### Example 5 (Linear Feedback Control Gain Design).

Given a linear control system $\dot{x} = Ax + Bu$ with linear state feedback $u = Kx$. Find $K$ to stabilize the system.
Example 6 (Robust Stability).

Given system $\dot{x} = Ax + u$ with uncertain feedback $u = g(x)$. Suppose all we know is that the feedback law satisfies: $\|g(x)\|^2 \leq \beta \|x\|^2$. Find Lyapunov function $V(x) = x^T Px$ to ensure exponential stability.
Concluding Remarks

• Basic knowledge on optimization and Lagrange duality

• Linear matrix inequalities impose convex constraints

• Semidefinite programming problem: optimize linear cost subject to LMI constraints

• SDP has broad applications in various engineering fields: signal processing, networking, communication, control, machine learning, big data...

• Further reading [BV04; Boy+94; VB00]

• **Next Lecture:** Stability analysis for hybrid systems
References

