

State-Space Control Fundamentals (00860721)

lecture no. 6

Leonid Mirkin

The Stephen B. Klein Faculty of Aerospace Engineering
Technion—IIT



Outline

Background: Lyapunov Stability and Lyapunov's direct method

Analysis via quadratic Lyapunov functions

State-feedback regulation via quadratic Lyapunov functions and LMIs

LQR design via LMIs

State observers

Outline

Background: Lyapunov Stability and Lyapunov's direct method

Analysis via quadratic Lyapunov functions

State-feedback regulation via quadratic Lyapunov functions and LMIs

LQR design via LMIs

State observers

Stability of unforced motion (Lyapunov stability)

Let

$$\dot{x} = f(x).$$

Its equilibrium $x_{\text{eq}} \in \mathbb{R}^n$, satisfying $f(x_{\text{eq}}) = 0$, is said to be

- **stable** if $\forall \epsilon > 0$ there is $\delta = \delta(\epsilon) > 0$ such that

$$\|x(0) - x_{\text{eq}}\| < \delta \implies \|x(t) - x_{\text{eq}}\| < \epsilon, \quad \forall t \in \mathbb{R}_+$$

- **asymptotically stable** if stable and $\exists \delta > 0$ such that

$$\|x(0) - x_{\text{eq}}\| < \delta \implies \lim_{t \rightarrow \infty} \|x(t) - x_{\text{eq}}\| = 0$$

The **region of attraction** of an asymptotically stable equilibrium is the set of initial conditions $x(0)$ that generate states x converging to x_{eq} . If the region of attraction is the whole set $\mathbb{R}^n$, then the equilibrium is said to be **globally asymptotically stable**.

Remark. We may assume w.l.o.g. that $x_{\text{eq}} = 0$. Otherwise, just rewrite the dynamics in terms of deviations $x_\delta = x - x_{\text{eq}}$.

Lyapunov stability for LTI systems

The system $\dot{x} = Ax$ is

- stable iff $\text{spec}(A) \subset \{z \in \mathbb{C} \mid \text{Re } z \leq 0\}$ and every imaginary eigenvalue is not defective (i.e. its algebraic and geometric multiplicities are equal)
- asymptotically stable iff A is Hurwitz, i.e. $\text{spec}(A) \subset \{z \in \mathbb{C} \mid \text{Re } z < 0\}$

Moreover,

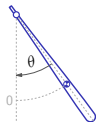
- stability properties are always global
- asymptotic stability is equivalent to BIBO stability of $\dot{x} = Ax + Bu$

in the finite-dimensional LTI case.

Lyapunov's direct method: motivating example

Nonlinear dynamics of planar damped pendulum:

$$J\ddot{\theta}(t) + c\dot{\theta}(t) + mgl \sin \theta(t) = 0$$



where

- m is the pendulum mass
- J is its moment of inertia about its pivot point
- c is the viscous friction coefficient
- l is the distance from the pendulum center of mass to its pivot point
- g is the standard gravity

Total mechanical energy (kinetic + potential, zero at the lowest CoM point)

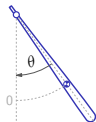
$$E(t) = \frac{1}{2} J \dot{\theta}^2(t) + mgl(1 - \cos \theta(t)).$$

The minimum $E = 0$ is at the equilibrium $\theta = \dot{\theta} = 0$ (is stable, by physics).
The unstable equilibrium, $\theta = \pi$ and $\dot{\theta} = 0$, has a potential energy of $2mgl$.

Lyapunov's direct method: motivating example

Nonlinear dynamics of planar damped pendulum:

$$J\ddot{\theta}(t) + c\dot{\theta}(t) + mgl \sin \theta(t) = 0$$



where

- m is the pendulum mass
- J is its moment of inertia about its pivot point
- c is the viscous friction coefficient
- l is the distance from the pendulum center of mass to its pivot point
- g is the standard gravity

Total mechanical energy (kinetic + potential, zero at the lowest CoM point)

$$E(t) = \frac{1}{2}J\dot{\theta}^2(t) + mgl(1 - \cos \theta(t)).$$

The minimum $E = 0$ is at the equilibrium $\theta = \dot{\theta} = 0$ (is stable, by physics).
The unstable equilibrium, $\theta = \pi$ and $\dot{\theta} = 0$, has a potential energy of $2mgl$.

Lyapunov's direct method: motivating example (contd)

The time derivative of the energy E along the pendulum motion

$$\dot{E} = \frac{\partial E}{\partial \theta} \dot{\theta} + \frac{\partial E}{\partial \dot{\theta}} \ddot{\theta} = mgl \dot{\theta} \sin \theta + J \underbrace{\dot{\theta} \frac{-c\dot{\theta} - mgl \sin \theta}{J}}_{\ddot{\theta}} = -c\dot{\theta}^2 \leq 0$$

Analysis (a bit sketchy)

- E strictly monotonically decays, i.e. dissipates, unless $\dot{\theta} = 0$
- $\dot{\theta} \equiv 0$ only at the equilibria, otherwise $\dot{\theta} = 0$ is instantaneous
- if $E(0) < 2mgl$, the “up” equilibrium cannot be attained
- hence, asymptotic convergence to the “down” equilibrium under initial conditions sufficiently close to $\theta = \dot{\theta} = 0$ (i.e. asymptotic stability)

Moral

- stability can be evaluated w/o solving dynamics explicitly
- helpful: analyse an energy-like function along trajectories of the system

Lyapunov's direct method: motivating example (contd)

The time derivative of the energy E along the pendulum motion

$$\dot{E} = \frac{\partial E}{\partial \theta} \dot{\theta} + \frac{\partial E}{\partial \dot{\theta}} \ddot{\theta} = mgl \dot{\theta} \sin \theta + J \underbrace{\dot{\theta} \frac{-c\dot{\theta} - mgl \sin \theta}{J}}_{\ddot{\theta}} = -c\dot{\theta}^2 \leq 0$$

Analysis (a bit sketchy)

- E strictly monotonically decays, i.e. dissipates, unless $\dot{\theta} = 0$
- $\dot{\theta} \equiv 0$ only at the equilibria, otherwise $\dot{\theta} = 0$ is instantaneous
- if $E(0) < 2mgl$, the “up” equilibrium cannot be attained
- hence, asymptotic convergence to the “down” equilibrium under initial conditions sufficiently close to $\theta = \dot{\theta} = 0$ (i.e. asymptotic stability)

Moral:

- stability can be evaluated w/o solving dynamics explicitly
- helpful: analyse an energy-like function along trajectories of the system

Lyapunov's direct method in a nutshell

Let $x_{\text{eq}} = 0$ and $\mathbb{X} \subset \mathbb{R}^n$ be such that $0 \in \mathbb{X}$. A *continuously differentiable* function $V : \mathbb{X} \rightarrow \mathbb{R}$ is a Lyapunov function candidate (“ersatz energy”) if

- $V(x) > 0$ for all $x(t) \in \mathbb{X} \setminus \{0\}$ and $V(0) = 0$.

Its derivative along trajectories of the system

$$\dot{V}(x) := \frac{dV(x)}{dt} = \frac{\partial V(x)}{\partial x} \frac{dx}{dt} = \frac{\partial V(x)}{\partial x} f(x).$$

The system $\dot{x} = f(x)$ is

- stable if $\dot{V}(x) \leq 0$ for all $x(t) \in \mathbb{X}$
 x converges to the largest invariant set¹ s.t. $\dot{V}(x) \equiv 0$ LaSalle's invariance principle
- asymptotically stable if $\dot{V}(x) < 0$ for all $x(t) \in \mathbb{X} \setminus \{0\}$

Global versions require, in addition,

– $\mathbb{X} = \mathbb{R}^n$

– radial unboundedness of V , i.e. that $V(x) \rightarrow \infty$ whenever $\|x\| \rightarrow \infty$.

¹A set $\mathbb{S}$ is said to be invariant for $\dot{x} = f(x)$ if $x(0) \in \mathbb{S} \implies x(t) \in \mathbb{S}$ for all $t > 0$.

Lyapunov's direct method in a nutshell

Let $x_{\text{eq}} = 0$ and $\mathbb{X} \subset \mathbb{R}^n$ be such that $0 \in \mathbb{X}$. A *continuously differentiable* function $V : \mathbb{X} \rightarrow \mathbb{R}$ is a Lyapunov function candidate (“ersatz energy”) if

- $V(x) > 0$ for all $x(t) \in \mathbb{X} \setminus \{0\}$ and $V(0) = 0$.

Its derivative along trajectories of the system

$$\dot{V}(x) := \frac{dV(x)}{dt} = \frac{\partial V(x)}{\partial x} \frac{dx}{dt} = \frac{\partial V(x)}{\partial x} f(x).$$

The system $\dot{x} = f(x)$ is

- stable if $\dot{V}(x) \leq 0$ for all $x(t) \in \mathbb{X}$
 x converges to the largest invariant set¹ s.t. $\dot{V}(x) \equiv 0$ LaSalle's invariance principle
- asymptotically stable if $\dot{V}(x) < 0$ for all $x(t) \in \mathbb{X} \setminus \{0\}$

Global versions require, in addition,

- g $\mathbb{X} = \mathbb{R}^n$,
- g radial unboundedness of V , i.e. that $V(x) \rightarrow \infty$ whenever $\|x\| \rightarrow \infty$.

¹A set $\mathbb{S}$ is said to be invariant for $\dot{x} = f(x)$ if $x(0) \in \mathbb{S} \implies x(t) \in \mathbb{S}$ for all $t > 0$.

Pendulum example (contd)

Choosing $x = (\theta, \dot{\theta})$, the state equation reads $\dot{x} = \begin{bmatrix} -(cx_2 + \frac{x_2^2}{J} \sin x_1)/J \end{bmatrix}$.

- If $V(x) = E = 0.5Jx_2^2 + mgl(1 - \cos x_1)$, then $\dot{V}(x) = -cx_2^2 \leq 0$ and we have local stability, but not asymptotic stability (need LaSalle).

If

$$V(x) = \frac{J}{2}x_2^2 + \beta x_1 x_2 + \frac{\alpha}{2}x_1^2 + mgl(1 - \cos x_1), \quad \alpha > \frac{\beta^2}{J}$$

(the last inequality ensures that $V(x) > 0$ for all $x \neq 0$), then

$$\dot{V}(x) = -\beta \frac{mgl}{J} x_1 \sin x_1 - (c - \beta)x_2^2 + \left(\alpha - \beta \frac{c}{J}\right)x_1 x_2.$$

We may want to zero the cross-term $x_1 x_2$, as its sign is uncertain, and keep the first two terms negative. Hence, $\alpha = \beta c/J$ and $\beta \in (0, c)$, in which case also the condition $\alpha > \beta^2/J$ is satisfied and we end up with

$$\dot{V} = -\beta \frac{mgl}{J} x_1 \sin x_1 - (c - \beta)x_2^2 < 0$$

for all $x \in ((-\pi, \pi) \times \mathbb{R}) \setminus \{0\}$. This confirms local asymptotic stability.

Pendulum example (contd)

Choosing $x = (\theta, \dot{\theta})$, the state equation reads $\dot{x} = \begin{bmatrix} -(cx_2 + \frac{x_2^2}{2} \sin x_1)/J \end{bmatrix}$.

- If $V(x) = E = 0.5Jx_2^2 + mgl(1 - \cos x_1)$, then $\dot{V}(x) = -cx_2^2 \leq 0$ and we have local stability, but not asymptotic stability (need LaSalle).

- If

$$V(x) = \frac{J}{2}x_2^2 + \beta x_1 x_2 + \frac{\alpha}{2}x_1^2 + mgl(1 - \cos x_1), \quad \alpha > \frac{\beta^2}{J}$$

(the last inequality ensures that $V(x) > 0$ for all $x \neq 0$), then

$$\dot{V}(x) = -\beta \frac{mgl}{J} x_1 \sin x_1 - (c - \beta)x_2^2 + \left(\alpha - \beta \frac{c}{J}\right) x_1 x_2.$$

We may want to zero the cross-term $x_1 x_2$, as its sign is uncertain, and keep the first two terms negative. Hence, $\alpha = \beta c/J$ and $\beta \in (0, c)$, in which case also the condition $\alpha > \beta^2/J$ is satisfied and we end up with

$$\dot{V} = -\beta \frac{mgl}{J} x_1 \sin x_1 - (c - \beta)x_2^2 < 0$$

for all $x \in ((-\pi, \pi) \times \mathbb{R}) \setminus \{0\}$. This confirms local asymptotic stability.

Outline

Background: Lyapunov Stability and Lyapunov's direct method

Analysis via quadratic Lyapunov functions

State-feedback regulation via quadratic Lyapunov functions and LMIs

LQR design via LMIs

State observers

Quadratic Lyapunov functions for linear systems

A popular choice is

$$V(x) = x'Px$$

for some $P = P' > 0$, which is a design parameter. Because $f(x) = Ax$,

$$\begin{aligned}\dot{V}(x) &= \frac{\partial V(x)}{\partial x} Ax = 2x'PAx = x'PAx + (x'PAx)' = x'PAx + x'A'Px \\ &= x'(PA + A'P)x\end{aligned}$$

Hence, by Lyapunov's direct method the system is

- asymptotically stable if there is $P > 0$ such that $PA + A'P < 0$.

This condition is known as the **Lyapunov inequality** on P . If $PA + A'P < 0$, then $\exists Q = Q' > 0$ such that

$$PA + A'P + Q = 0,$$

which is a matrix Lyapunov equation.

Quadratic Lyapunov functions for linear systems

A popular choice is

$$V(x) = x'Px$$

for some $P = P' > 0$, which is a design parameter. Because $f(x) = Ax$,

$$\begin{aligned}\dot{V}(x) &= \frac{\partial V(x)}{\partial x} Ax = 2x'PAx = x'PAx + (x'PAx)' = x'PAx + x'A'Px \\ &= x'(PA + A'P)x\end{aligned}$$

Hence, by Lyapunov's direct method the system is

- asymptotically stable if there is $P > 0$ such that $PA + A'P < 0$.

This condition is known as the **Lyapunov inequality** on P . If $PA + A'P < 0$, then $\exists Q = Q' > 0$ such that

$$PA + A'P + Q = 0,$$

which is a matrix Lyapunov equation.

Quadratic Lyapunov functions: necessity

Finding a Lyapunov function might be a hard problem in itself, even though its existence is necessary for the stability of a large class of systems. But in the LTI case it may be sufficient to seek for a quadratic V .

Lemma 6.1

$\dot{x} = Ax$ is asymptotically stable *iff* there is a quadratic Lyapunov function of the form $V = x'Px$ for some $P = P' > 0$ for it.

Proof.

if: By Lyapunov's direct method.

only if: If A is Hurwitz, then the Lyapunov equation $PA + A'P + Q = 0$ is solvable for every Q and $P > 0$ whenever $Q > 0$. Hence, $V = x'Px$ is a quadratic Lyapunov function, $\dot{V} = -x'Qx < 0$ for all $x \neq 0$. □

Lemma 6.2 (alternative formulation)

A is Hurwitz *iff* there is $P = P' > 0$ such that $PA + A'P < 0$.

Quadratic Lyapunov functions: necessity

Finding a Lyapunov function might be a hard problem in itself, even though its existence is necessary for the stability of a large class of systems. But in the LTI case it may be sufficient to seek for a quadratic V .

Lemma 6.1

$\dot{x} = Ax$ is asymptotically stable *iff* there is a quadratic Lyapunov function of the form $V = x'Px$ for some $P = P' > 0$ for it.

Proof.

if: By Lyapunov's direct method.

only if: If A is Hurwitz, then the Lyapunov equation $PA + A'P + Q = 0$ is solvable for every Q and $P > 0$ whenever $Q > 0$. Hence, $V = x'Px$ is a quadratic Lyapunov function, $\dot{V} = -x'Qx < 0$ for all $x \neq 0$. □

Lemma 6.2 (alternative formulation)

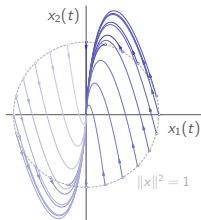
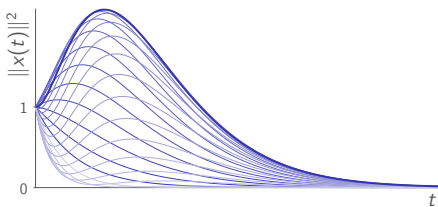
A is Hurwitz *iff* there is $P = P' > 0$ such that $PA + A'P < 0$.

The norm of states of stable systems

Example: If

$$A = \begin{bmatrix} -1 & 0 \\ 3.75 & -1 \end{bmatrix} \Rightarrow e^{At} = \begin{bmatrix} e^{-t} & 0 \\ 3.75te^{-t} & e^{-t} \end{bmatrix},$$

then



from initial conditions $x_0 = \begin{bmatrix} \sin \theta \\ \cos \theta \end{bmatrix}$ for various $\theta \in [0, 2\pi]$.

Moral: if a system is asymptotically stable, so that $\lim_{t \rightarrow \infty} x(t) = 0$,

- $\|x(t)\|$ is not necessarily a monotonically decreasing function of t
this is why switching between stable systems might produce an unstable response...

Geometry of quadratic Lyapunov functions

The Lyapunov inequality essentially says that

- level curves $x'(t)Px(t) = \text{const}$ (ellipsoids) strictly decrease with time for any $P > 0$ such that $PA + A'P < 0$.

Another viewpoint: Let $T \in \mathbb{R}^{n \times n}$ be any matrix such that $P = T'T$. This T is nonsingular and may be taken as $P^{1/2}$ or the Cholesky factor of P , for instance. Changing coordinates to $\tilde{x} = Tx$ results in

$$\dot{\tilde{x}} = \tilde{A}\tilde{x} \quad \text{for } \tilde{A} := TAT^{-1}.$$

The Lyapunov candidate $\tilde{V}(\tilde{x}) = \|\tilde{x}\|^2$ for it satisfies

$$\dot{\tilde{V}}(\tilde{x}) = \tilde{x}'(\tilde{A} + \tilde{A}')\tilde{x} = \tilde{x}'(TAT^{-1} + T^{-1}A'T')\tilde{x} = \tilde{x}'T^{-1}(PA + A'P)T^{-1}\tilde{x} \stackrel{PA + A'P < 0}{< 0}.$$

Thus, each quadratic Lyapunov function offers

- a coordinate basis in which the state norm decreases monotonically
- a Lyapunov ellipsoid $\tilde{x}'\tilde{x} = 1$ (or $x'Px = 1$) around the origin.

Geometry of quadratic Lyapunov functions

The Lyapunov inequality essentially says that

- level curves $x'(t)Px(t) = \text{const}$ (ellipsoids) strictly decrease with time for any $P > 0$ such that $PA + A'P < 0$.

Another viewpoint: Let $T \in \mathbb{R}^{n \times n}$ be any matrix such that $P = T'T$. This T is nonsingular and may be taken as $P^{1/2}$ or the Cholesky factor of P , for instance. Changing coordinates to $\tilde{x} = Tx$ results in

$$\dot{\tilde{x}} = \tilde{A}\tilde{x} \quad \text{for } \tilde{A} := TAT^{-1}.$$

The Lyapunov candidate $\tilde{V}(\tilde{x}) = \|\tilde{x}\|^2$ for it satisfies

$$\dot{\tilde{V}}(\tilde{x}) = \tilde{x}'(\tilde{A} + \tilde{A}')\tilde{x} = \tilde{x}'(TAT^{-1} + T^{-1}A'T')\tilde{x} = \tilde{x}'T^{-1}(\overbrace{PA + A'P}^{<0})T^{-1}\tilde{x}.$$

Thus, each quadratic Lyapunov function offers

- a coordinate basis in which the state norm decreases monotonically
this motivates analysing switched systems via common Lyapunov functions...

Geometry of quadratic Lyapunov functions (contd)

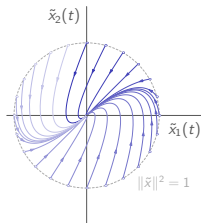
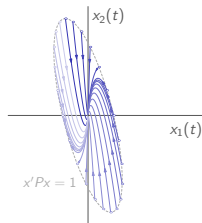
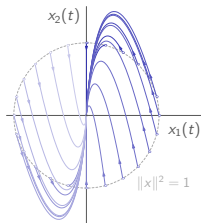
Example (contd): Choosing P via $PA + A'P = -2I$ results in

$$P = \frac{1}{32} \begin{bmatrix} 257 & 60 \\ 60 & 32 \end{bmatrix} \quad \text{and} \quad T = \frac{1}{8\sqrt{257}} \begin{bmatrix} 257 & 60 \\ 0 & 68 \end{bmatrix} \approx \begin{bmatrix} 2 & 0.468 \\ 0 & 0.53 \end{bmatrix}$$

The level curves $x'Px = \text{const}$ are tilted ellipses with the axes ratio $1/4$ and $\|\tilde{x}(t)\|^2$ is a strictly monotonically decreasing function of t , with

$$\frac{d}{dt} \|\tilde{x}(t)\|^2 = -2\|\tilde{x}(t)\|^2 = -2e^{-2t}(x_{0,1}^2 + (x_{0,2} + 3.75tx_{0,1})^2) < 0$$

Compare



Quadratic Lyapunov functions and decay rate

Suppose $V(x) = x'Px$ for some $P = P' > 0$ is such that $\dot{V}(x) \leq -x'Qx$ for some $Q > 0$. Because $x'Qx \geq \lambda_{\min}(Q)\|x\|^2$ and $x'Px \leq \lambda_{\max}(P)\|x\|^2$,

$$\dot{V}(x) \leq -\lambda_{\min}(Q)\|x\|^2 \leq -\frac{\lambda_{\min}(Q)}{\lambda_{\max}(P)}V(x).$$

Hence, by Grönwall's inequality,

$$V(x(t)) \leq e^{-t\lambda_{\min}(Q)/\lambda_{\max}(P)}V(x_0).$$

Using $\lambda_{\min}(P)\|x\|^2 \leq V(x) \leq \lambda_{\max}(P)\|x\|^2$, we end up with

$$\|x(t)\|^2 \leq \frac{\lambda_{\max}(P)}{\lambda_{\min}(P)}e^{-t\lambda_{\min}(Q)/\lambda_{\max}(P)}\|x_0\|^2 =: \kappa_P e^{-\alpha_{qp}t}\|x_0\|^2.$$

This is the definition of **exponential convergence**.

Remark. Nothing in this analysis required the linearity of the system. Hence, we have exponential convergence whenever we have a quadratic Lyapunov function with a quadratic upper bound of its derivative along trajectories of the system.

Quadratic Lyapunov functions and decay rate (contd)

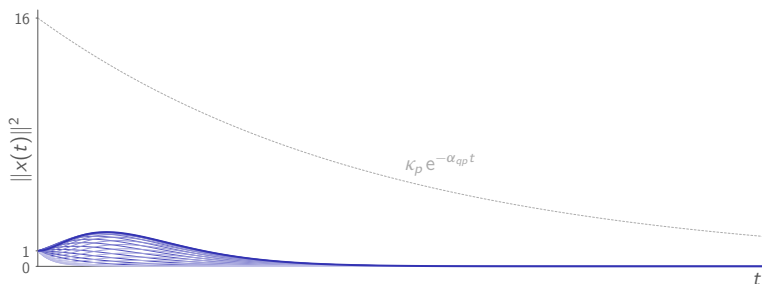
The upper bound $\|x(t)\|^2 \leq \kappa_p e^{-\alpha_{qp} t} \|x_0\|^2$ might be very loose. It is thus

- more a qualitative result than a practical measure.

Example (contd): If P satisfies $PA + A'P = -2I$, then

$$\kappa_p = \frac{\lambda_{\max}(P)}{\lambda_{\min}(P)} = 16 \quad \text{and} \quad \alpha_{qp} = \frac{\lambda_{\min}(Q)}{\lambda_{\max}(P)} = \frac{4}{17} \approx 0.236$$

so



Lyapunov inequality for guaranteed decay rate

If instead of $\dot{V}(x) < 0$ we require $\dot{V}(x) < -2\alpha V(x)$ for some $\alpha > 0$, then

$$V(x(t)) < e^{-2\alpha t} V(x_0)$$

and, therefore,

$$\|x(t)\|^2 < \frac{\lambda_{\max}(P)}{\lambda_{\min}(P)} e^{-2\alpha t} \|x_0\|^2 = \kappa e^{-2\alpha t} \|x_0\|^2.$$

In other words, in this case we can

- guarantee the exponential convergence faster than that of $e^{-2\alpha t}$.

In the linear case, $\dot{V}(x) < -2\alpha V(x)$ reads

$$PA + A^T P < -2\alpha P \iff P(A + \alpha I) + (A + \alpha I)^T P < 0,$$

which equals to Hurwitz $A + \alpha I \iff \text{spec}(A) \in \{z \in \mathbb{C} \mid \text{Re } z < -\alpha\}$.

Lyapunov inequality for guaranteed decay rate

If instead of $\dot{V}(x) < 0$ we require $\dot{V}(x) < -2\alpha V(x)$ for some $\alpha > 0$, then

$$V(x(t)) < e^{-2\alpha t} V(x_0)$$

and, therefore,

$$\|x(t)\|^2 < \frac{\lambda_{\max}(P)}{\lambda_{\min}(P)} e^{-2\alpha t} \|x_0\|^2 = \kappa e^{-2\alpha t} \|x_0\|^2.$$

In other words, in this case we can

- guarantee the exponential convergence faster than that of $e^{-2\alpha t}$.

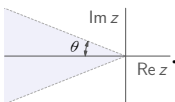
In the linear case, $\dot{V}(x) < -2\alpha V(x)$ reads

$$PA + A'P < -2\alpha P \iff P(A + \alpha I) + (A + \alpha I)'P < 0,$$

which equals to Hurwitz $A + \alpha I \iff \text{spec}(A) \in \{z \in \mathbb{C} \mid \text{Re } z < -\alpha\}$.

Lyapunov inequality for guaranteed damping ratio

Given $\theta \in (0, \pi/2]$, define

$$\text{sect}_\theta := \{z \in \mathbb{C} \mid \text{Re } z < -\cot \theta |\text{Im } z|\} =$$


It can be shown that $\text{spec}(A) \in \text{sect}_\theta$ iff

$$\exists P > 0 \text{ such that } \begin{bmatrix} \sin \theta (PA + A'P) & \cos \theta (PA - A'P) \\ -\cos \theta (PA - A'P) & \sin \theta (PA + A'P) \end{bmatrix} < 0$$

($\theta = \arccos \zeta$ if $\zeta \in [0, 1]$ is the damping ratio of an underdamped system).

Remark. This result is connected to the relation (holds whenever $0 < \theta \leq \pi/2$, check it)

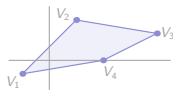
$$\text{sect}_\theta = \left\{ z \in \mathbb{C} \mid z \begin{bmatrix} \sin \theta & \cos \theta \\ -\cos \theta & \sin \theta \end{bmatrix} + \bar{z} \begin{bmatrix} \sin \theta & -\cos \theta \\ \cos \theta & \sin \theta \end{bmatrix} < 0 \right\}$$

which reads $\text{Re } z < 0$ and $(\text{Re } z)^2 < \cot^2 \theta (\text{Im } z)^2$ via Schur complement arguments.

Polytopic uncertainty

Given k elements (vertices) V_i , their convex hull is the set

$$\text{conv}\{V_i\} := \sum_{i=1}^k \mu_i V_i \text{ for all } \mu_i \geq 0 \text{ s.t. } \sum_{i=1}^k \mu_i = 1, \quad \text{e.g.}$$



This is a way to define uncertainty in matrices. For example, let

$$A = \begin{bmatrix} -\alpha/\underline{a}_1 & \alpha/\underline{a}_1 \\ \alpha/\underline{a}_2 & -\alpha(1+\rho^2)/\underline{a}_2 \end{bmatrix} \quad \text{under } a_i \in [\underline{a}_i, \bar{a}_i] \text{ s.t. } \bar{a}_i > \underline{a}_i > 0,$$

where $\alpha := r_1 \sqrt{1+\rho^2}/(2\rho\sqrt{h_{\text{eq1}}})$. There are 4 vertices

$$A_1 = \begin{bmatrix} -\alpha/\underline{a}_1 & \alpha/\underline{a}_1 \\ \alpha/\underline{a}_2 & -\alpha(1+\rho^2)/\underline{a}_2 \end{bmatrix}, \quad A_2 = \begin{bmatrix} -\alpha/\bar{a}_1 & \alpha/\bar{a}_1 \\ \alpha/\bar{a}_2 & -\alpha(1+\rho^2)/\bar{a}_2 \end{bmatrix},$$

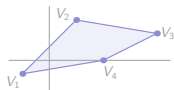
$$A_3 = \begin{bmatrix} -\alpha/\underline{a}_1 & \alpha/\bar{a}_1 \\ \alpha/\bar{a}_2 & -\alpha(1+\rho^2)/\bar{a}_2 \end{bmatrix}, \quad A_4 = \begin{bmatrix} -\alpha/\bar{a}_1 & \alpha/\underline{a}_1 \\ \alpha/\underline{a}_2 & -\alpha(1+\rho^2)/\underline{a}_2 \end{bmatrix},$$

and then $A \in \text{conv}\{A_1, A_2, A_3, A_4\}$ for every admissible choice of a_1 and a_2 .

Polytopic uncertainty

Given k elements (vertices) V_i , their convex hull is the set

$$\text{conv}\{V_i\} := \sum_{i=1}^k \mu_i V_i \text{ for all } \mu_i \geq 0 \text{ s.t. } \sum_{i=1}^k \mu_i = 1, \quad \text{e.g.}$$



This is a way to define uncertainty in matrices. For example, let

$$A = \begin{bmatrix} -\alpha/a_1 & \alpha/a_1 \\ \alpha/a_2 & -\alpha(1+\rho^2)/a_2 \end{bmatrix} \quad \text{under } a_i \in [\underline{a}_i, \bar{a}_i] \text{ s.t. } \bar{a}_i > \underline{a}_i > 0,$$

where $\alpha := r_1 \sqrt{1 + \rho^2} / (2\rho \sqrt{h_{\text{eq1}}})$. There are 4 vertices

$$A_1 = \begin{bmatrix} -\alpha/\underline{a}_1 & \alpha/\underline{a}_1 \\ \alpha/\underline{a}_2 & -\alpha(1+\rho^2)/\underline{a}_2 \end{bmatrix}, \quad A_2 = \begin{bmatrix} -\alpha/\bar{a}_1 & \alpha/\bar{a}_1 \\ \alpha/\bar{a}_2 & -\alpha(1+\rho^2)/\bar{a}_2 \end{bmatrix},$$

$$A_3 = \begin{bmatrix} -\alpha/\underline{a}_1 & \alpha/\bar{a}_1 \\ \alpha/\bar{a}_2 & -\alpha(1+\rho^2)/\bar{a}_2 \end{bmatrix}, \quad A_4 = \begin{bmatrix} -\alpha/\bar{a}_1 & \alpha/\underline{a}_1 \\ \alpha/\underline{a}_2 & -\alpha(1+\rho^2)/\underline{a}_2 \end{bmatrix},$$

and then $A \in \text{conv}\{A_1, A_2, A_3, A_4\}$ for every admissible choice of a_1 and a_2 .

Lyapunov inequalities and polytopic uncertainty

Assume that $A \in \text{conv}\{A_i\}$ for some known $A_i \in \mathbb{R}^{n \times n}$, $i \in \mathbb{Z}_{1..k}$. We may be interested to know if

- $\dot{x} = Ax$ is stable for every A from that class.

By Lyapunov's direct method, it is *sufficient* to have $P > 0$ such that

$$PA + A'P < 0 \quad \text{for every admissible } \mu_i.$$

Because

$$PA + A'P = \sum_{i=1}^v \mu_i (PA_i + A'_i P) \quad \text{and} \quad \mu_i \geq 0,$$

we conclude that

- stability is ensured if $\exists P > 0$ such that $PA_i + A'_i P < 0$ for all $i \in \mathbb{Z}_{1..k}$, i.e. it's enough to check a finite number of Lyapunov inequalities. The same applies to inequalities for guaranteed decay rate, damping ratio, ...

Remark. These arguments do not exploit the information that $\sum_i \mu_i = 1$. This implies that the result will most probably be quite conservative. Still, this is just a flavor of the approach, *robust control* has many tools.

Outline

Background: Lyapunov Stability and Lyapunov's direct method

Analysis via quadratic Lyapunov functions

State-feedback regulation via quadratic Lyapunov functions and LMIs

LQR design via LMIs

State observers

Linear matrix inequalities

LMIs are inequalities of the form

$$M_0 + x_1 M_1 + \cdots + x_k M_k < 0$$

for given $M_i = M_i' \in \mathbb{R}^{r \times r}$, $i \in \mathbb{Z}_{0..k}$, and unknowns x_i , $i \in \mathbb{Z}_{1..k}$ (**decision variables**). Two key problems are

feasibility: determine whether there are x_i for which the LMI is satisfied,

optimization: minimize $\mu_1 x_1 + \cdots + \mu_k x_k$ over all feasible x_i for given μ_i .

Advantages:

- there are efficient numerical tools for solving LMI problems
- several LMI's combine as an LMI seamlessly, like

$$\left. \begin{array}{l} x_1 M_1 + x_2 M_2 < 0 \\ x_1 N_1 + x_2 N_2 < 0 \end{array} \right\} \iff x_1 \begin{bmatrix} M_1 & 0 \\ 0 & N_1 \end{bmatrix} + x_2 \begin{bmatrix} M_2 & 0 \\ 0 & N_2 \end{bmatrix} < 0$$

how all this is relevant to seeking for $P > 0$ rendering $PA + A'P < 0$?

Linear matrix inequalities

LMIs are inequalities of the form

$$M_0 + x_1 M_1 + \cdots + x_k M_k < 0$$

for given $M_i = M_i' \in \mathbb{R}^{r \times r}$, $i \in \mathbb{Z}_{0..k}$, and unknowns x_i , $i \in \mathbb{Z}_{1..k}$ (**decision variables**). Two key problems are

feasibility: determine whether there are x_i for which the LMI is satisfied,

optimization: minimize $\mu_1 x_1 + \cdots + \mu_k x_k$ over all feasible x_i for given μ_i .

Advantages:

- there are efficient numerical tools for solving LMI problems
- several LMI's combine as an LMI seamlessly, like

$$\left. \begin{array}{l} x_1 M_1 + x_2 M_2 < 0 \\ x_1 N_1 + x_2 N_2 < 0 \end{array} \right\} \iff x_1 \begin{bmatrix} M_1 & 0 \\ 0 & N_1 \end{bmatrix} + x_2 \begin{bmatrix} M_2 & 0 \\ 0 & N_2 \end{bmatrix} < 0$$

Q: how all this is relevant to seeking for $P > 0$ rendering $PA + A'P < 0$?

Stability via LMIs: example

Let $n = 2$. In this case any symmetric

$$P = \begin{bmatrix} x_1 & x_3 \\ x_3 & x_2 \end{bmatrix} = x_1 \underbrace{\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}}_{E_1} + x_2 \underbrace{\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}}_{E_2} + x_3 \underbrace{\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}}_{E_3}.$$

Hence,

$$PA + A'P < 0 \iff x_1(E_1A + A'E_1) + x_2(E_2A + A'E_2) + x_3(E_3A + A'E_3) < 0$$

and

$$P > 0 \iff -x_1E_1 - x_2E_2 - x_3E_3 < 0.$$

Thus, there is $P > 0$ such that $PA + A'P < 0$ iff the LMI

$$\sum_{i=1}^3 x_i \underbrace{\begin{bmatrix} -E_i & 0 \\ 0 & E_iA + A'E_i \end{bmatrix}}_{M_i} < 0$$

is feasible.

Software to solve LMIs

Lucky us,

- we may never need to decompose LMIs to the defining form manually.

Software can be divided into two categories:

1. interface, e.g. with Matlab we use CVX, <https://cvxr.com/cvx/>
2. solver, under the hood e.g. MOSEK, <https://www.mosek.com>

Matlab with CVX code example:

```
A = [0 1 0; 0 0 1; -1 -3 -3];  
n = size(A,1);  
cvx_begin SDP  
    variable X(n,n) symmetric  
    subject to  
        X >= 1E-5*eye(n)      % workaround to handle X>0  
        A'*X+X*A <= -eye(n)  
cvx_end
```

Design problems

Given:

- stabilizable $(A, B) \in \mathbb{R}^{n \times n} \times \mathbb{R}^{n \times m}$

Find:

- gain $K \in \mathbb{R}^{m \times n}$ so that $A_K = A + BK$ is Hurwitz
- gain $K \in \mathbb{R}^{m \times n}$ so that $\text{spec}(A_K) \in \{z \in \mathbb{C} \mid \text{Re } z < -\alpha\}$
- gain $K \in \mathbb{R}^{m \times n}$ so that $\text{spec}(A_K) \in \{z \in \mathbb{C} \mid \text{Re } z < -\frac{\xi}{\sqrt{1-\xi^2}} |\text{Im } z|\}$

Equivalently, find K and $P > 0$ such that

$$P(A + BK) + (A + BK)'P < 0$$

$$P(A + BK) + (A + BK)'P + 2\alpha P < 0$$

$$\begin{bmatrix} \sqrt{1-\xi^2}(PA_K + A_K'P) & \xi(PA_K - A_K'P) \\ -\xi(PA_K - A_K'P) & \sqrt{1-\xi^2}(PA_K + A_K'P) \end{bmatrix} < 0$$

These are not LMIs in design variables P and K , because of the presence of

- bilinear terms PBK and $K'B'P$.

Design problems

Given:

- stabilizable $(A, B) \in \mathbb{R}^{n \times n} \times \mathbb{R}^{n \times m}$

Find:

- gain $K \in \mathbb{R}^{m \times n}$ so that $A_K = A + BK$ is Hurwitz
- gain $K \in \mathbb{R}^{m \times n}$ so that $\text{spec}(A_K) \in \{z \in \mathbb{C} \mid \text{Re } z < -\alpha\}$
- gain $K \in \mathbb{R}^{m \times n}$ so that $\text{spec}(A_K) \in \{z \in \mathbb{C} \mid \text{Re } z < -\frac{\xi}{\sqrt{1-\xi^2}} |\text{Im } z|\}$

Equivalently, find K and $P > 0$ such that

$$P(A + BK) + (A + BK)'P < 0$$

$$P(A + BK) + (A + BK)'P + 2\alpha P < 0$$

$$\begin{bmatrix} \sqrt{1-\xi^2}(PA_K + A_K'P) & \xi(PA_K - A_K'P) \\ -\xi(PA_K - A_K'P) & \sqrt{1-\xi^2}(PA_K + A_K'P) \end{bmatrix} < 0$$

These are not LMIs in design variables P and K , because of the presence of

- bilinear terms PBK and $K'B'P$.

The trick

We know (Property 2 in Lecture 1, Slide 28) that

$$P(A + BK) + (A + BK)'P = P((A + BK)P^{-1} + P^{-1}(A + BK)')P < 0$$

$$\Leftrightarrow$$

$$AP^{-1} + BKP^{-1} + P^{-1}A' + P^{-1}K'B' < 0$$

Introducing $M \equiv P^{-1}$ and $N \equiv KP^{-1}$, the problem is equivalent to finding $M \equiv M' > 0$ and N such that

$$AM + BN + MA' + N'B' < 0.$$

This is a feasibility LMI in the decision variables M and N . Once solved,

$$K \equiv NM^{-1}$$

is the required stabilizing state-feedback gain.

The trick

We know (Property 2 in Lecture 1, Slide 28) that

$$P(A + BK) + (A + BK)'P = P((A + BK)P^{-1} + P^{-1}(A + BK)')P < 0$$

$$\Leftrightarrow$$

$$AP^{-1} + BKP^{-1} + P^{-1}A' + P^{-1}K'B' < 0$$

Introducing $M := P^{-1}$ and $N := KP^{-1}$, the problem is equivalent to finding $M = M' > 0$ and N such that

$$AM + BN + MA' + N'B' < 0.$$

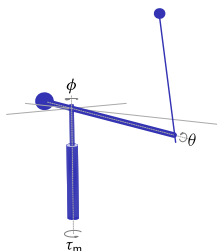
This is a **feasibility LMI** in the **decision variables M and N** . Once solved²,

$$K = NM^{-1}$$

is the required stabilizing state-feedback gain.

²The LMI must be feasible by stabilizability, although this is not evident from its form.

Example: inverted Furuta pendulum stabilization

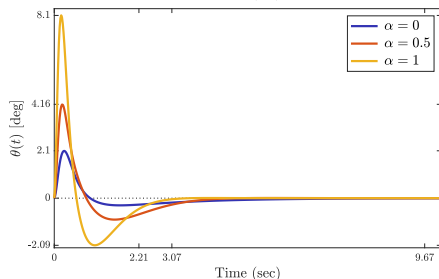
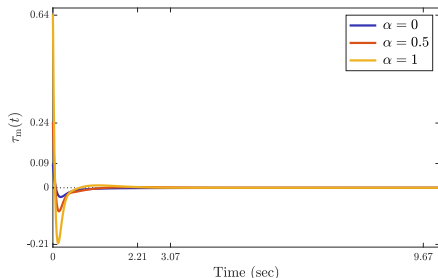
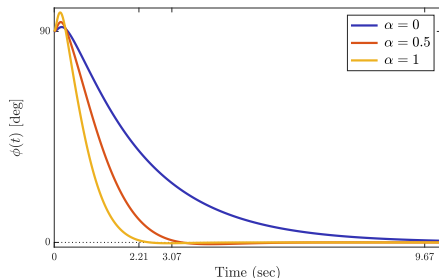


whose model is 4-order, with the state vector $x = (\phi, \dot{\phi}, \theta, \dot{\theta})$ and

$$[A \ B] = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & J_b + mr^2 & 0 & -mrl \\ 0 & 0 & 1 & 0 \\ 0 & -mrl & 0 & J_p + ml^2 \end{bmatrix}^{-1} \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & -f & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & gml & -\zeta & 0 \end{bmatrix}.$$

We take $J_b = 0.01$, $f = 0.01$, $J_p = 0.0003$, $\zeta = 0.0005$, $l = 0.3226$, $r = 0.3$, $m = 0.04$, and $g = 9.81$. In this case there is an ORHP pole at $s = 5.9582$.

Example: design with no damping constraints



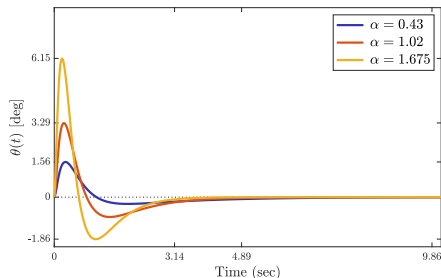
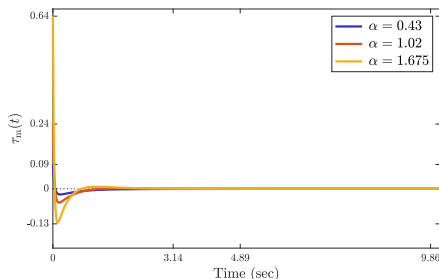
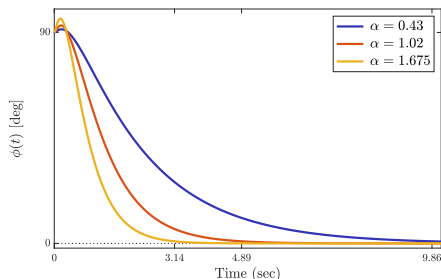
with

$$K = \begin{bmatrix} 0.056 & 0.15 & -2.565 & -0.403 \\ 0.156 & 0.209 & -3.749 & -0.497 \\ 0.41 & 0.383 & -4.921 & -0.754 \end{bmatrix}$$

and

$$\text{spec}(A+BK) : \begin{cases} \zeta_1 = 0.813, \zeta_{2,3} = 1 \\ \zeta_1 = 0.701, \zeta_2 = 0.84 \\ \zeta_1 = 0.768, \zeta_2 = 0.878 \end{cases}$$

Example: design with damping constraint $\zeta > 0.99$



with

$$K = \begin{bmatrix} 0.056 & 0.158 & -3.052 & -0.555 \\ 0.154 & 0.244 & -3.883 & -0.705 \\ 0.408 & 0.44 & -5.506 & -1 \end{bmatrix}$$

and

$$\text{spec}(A + BK) : \begin{cases} \zeta_{1,2,3,4} = 1 \\ \zeta_{1,2,3,4} = 1 \\ \zeta_{1,2,3,4} = 1 \end{cases}$$

Outline

Background: Lyapunov Stability and Lyapunov's direct method

Analysis via quadratic Lyapunov functions

State-feedback regulation via quadratic Lyapunov functions and LMIs

LQR design via LMIs

State observers

Weighted state energy

Consider the system $\dot{x}(t) = Ax(t)$, $x(0) = x_0$, with a Hurwitz A and the cost

$$\mathcal{J} := \int_0^\infty x'(t) Q_x x(t) dt$$

for some $Q_x = Q_x' \geq 0$. Because $x(t) = e^{At} x_0$,

$$\mathcal{J} = \int_0^\infty x_0' e^{A't} Q_x e^{At} x_0 dt = x_0' \int_0^\infty e^{A't} Q_x e^{At} dt x_0 = x_0' W x_0,$$

where $W \geq 0$ solves the Lyapunov equation

$$A'W + WA + Q_x = 0$$

and is thus the controllability Gramian of $(A', Q_x^{1/2})$. Note that

$$- \quad Q_x > 0 \implies Q_x^{1/2} > 0 \implies (A', Q_x^{1/2}) \text{ is controllable} \implies W > 0.$$

LQR revisited

Assume that

- $\begin{bmatrix} Q & S \\ S' & R \end{bmatrix} > 0$ to simplify technicalities, can be ≥ 0
- $u = Kx$ for a stabilizing K

The standard LQR is then the problem of minimizing

$$\mathcal{J} = \int_0^\infty x'(t) \begin{bmatrix} I & K' \end{bmatrix} \begin{bmatrix} Q & S \\ S' & R \end{bmatrix} \begin{bmatrix} I \\ K \end{bmatrix} x(t) dt$$

for $\dot{x} = (A + BK)x$ and a Hurwitz $A + BK$. Hence,

$$\mathcal{J} = x_0' W_K x_0$$

for $W_K = W_K' \geq 0$ solving

$$(A + BK)' W_K + W_K (A + BK) + \overbrace{\begin{bmatrix} I & K' \end{bmatrix} \begin{bmatrix} Q & S \\ S' & R \end{bmatrix} \begin{bmatrix} I \\ K \end{bmatrix}}^{>0} = 0.$$

By Lemma 6.1, we know that $W_K > 0$ then.

Replacing equality with inequality

Let $P = P' > 0$ solve

$$(A + BK)'P + P(A + BK) + \overbrace{\begin{bmatrix} I & K' \end{bmatrix} \begin{bmatrix} Q & S \\ S' & R \end{bmatrix} \begin{bmatrix} I \\ K \end{bmatrix}}^{>0} + Q_1 = 0$$

for some $Q_1 = Q_1' \geq 0$. The matrix $\Delta := P - W_K$ satisfies

$$(A + BK)'\Delta + \Delta(A + BK) + Q_1 = 0.$$

Because $A + BK$ is Hurwitz,

$$- Q_1 \geq 0 \implies \Delta \geq 0 \implies x_0' P x_0 \geq x_0' W_K x_0 \text{ for all } x_0.$$

Hence,

$$\mathcal{J} = \min x_0' P x_0$$

over all $P = P' > 0$ such that

$$(A + BK)'P + P(A + BK) + \begin{bmatrix} I & K' \end{bmatrix} \begin{bmatrix} Q & S \\ S' & R \end{bmatrix} \begin{bmatrix} I \\ K \end{bmatrix} \leq 0$$

for every K .

LQR revisited (contd)

Thus, LQR reduces to solving

$$\begin{aligned} & \underset{P=P', K}{\text{minimize}} && x_0' P x_0 \\ & \text{subject to} && P > 0 \\ & && (A + BK)'P + P(A + BK) + \begin{bmatrix} I & K' \end{bmatrix} \begin{bmatrix} Q & S \\ S' & R \end{bmatrix} \begin{bmatrix} I \\ K \end{bmatrix} \leq 0 \end{aligned}$$

Introducing $M = P^{-1}$ and $N = KP^{-1}$, it is equivalent to

$$\begin{aligned} & \underset{M=M', N}{\text{minimize}} && x_0' M^{-1} x_0 \\ & \text{subject to} && M > 0 \\ & && MA' + N'B' + AM + BN + \begin{bmatrix} M & N' \end{bmatrix} \begin{bmatrix} Q & S \\ S' & R \end{bmatrix} \begin{bmatrix} M \\ N \end{bmatrix} \leq 0 \end{aligned}$$

Still two problems:

1. cost is *not linear* in M
2. the second condition is still *not linear* in M and N

LQR revisited (contd)

Thus, LQR reduces to solving

$$\begin{aligned} & \underset{P=P', K}{\text{minimize}} && x_0' P x_0 \\ & \text{subject to} && P > 0 \\ & && (A + BK)'P + P(A + BK) + \begin{bmatrix} I & K' \end{bmatrix} \begin{bmatrix} Q & S \\ S' & R \end{bmatrix} \begin{bmatrix} I \\ K \end{bmatrix} \leq 0 \end{aligned}$$

Introducing $M = P^{-1}$ and $N = KP^{-1}$, it is equivalent to

$$\begin{aligned} & \underset{M=M', N}{\text{minimize}} && x_0' M^{-1} x_0 \\ & \text{subject to} && M > 0 \\ & && MA' + N'B' + AM + BN + \begin{bmatrix} M & N' \end{bmatrix} \begin{bmatrix} Q & S \\ S' & R \end{bmatrix} \begin{bmatrix} M \\ N \end{bmatrix} \leq 0 \end{aligned}$$

Still two problems:

1. cost is *not linear* in M
2. the second condition is still *not linear* in M and N

Schur complement arguments (Theorem 1.4, Lecture 1)

$$1. \quad x_0' M^{-1} x_0 \leq \gamma \iff \begin{bmatrix} \gamma & x_0' \\ x_0 & M \end{bmatrix} \geq 0 \text{ whenever } M > 0$$

2. Decompose

$$\begin{bmatrix} Q & S \\ S' & R \end{bmatrix} = \begin{bmatrix} C_z' \\ D_{zu}' \end{bmatrix} \begin{bmatrix} C_z & D_{zu} \end{bmatrix}$$

for some $C_z \in \mathbb{R}^{(n+m) \times n}$ and $D_{zu} \in \mathbb{R}^{(n+m) \times m}$. Hence

$$\begin{bmatrix} M & N' \end{bmatrix} \begin{bmatrix} Q & S \\ S' & R \end{bmatrix} \begin{bmatrix} M \\ N \end{bmatrix} = (C_z M + D_{zu} N)' (C_z M + D_{zu} N)$$

and

$$MA' + N'B' + AM + BN + (C_z M + D_{zu} N)' (C_z M + D_{zu} N) \leq 0$$

$\iff$

$$\begin{bmatrix} MA' + N'B' + AM + BN & MC_z' + N'D_{zu}' \\ C_z M + D_{zu} N & -I \end{bmatrix} \leq 0$$

Schur complement arguments (Theorem 1.4, Lecture 1)

1. $x_0' M^{-1} x_0 \leq \gamma \iff \begin{bmatrix} \gamma & x_0' \\ x_0 & M \end{bmatrix} \geq 0$ whenever $M > 0$

2. Decompose

$$\begin{bmatrix} Q & S \\ S' & R \end{bmatrix} = \begin{bmatrix} C_z' \\ D_{zu}' \end{bmatrix} \begin{bmatrix} C_z & D_{zu} \end{bmatrix}$$

for some $C_z \in \mathbb{R}^{(n+m) \times n}$ and $D_{zu} \in \mathbb{R}^{(n+m) \times m}$. Hence

$$\begin{bmatrix} M & N' \end{bmatrix} \begin{bmatrix} Q & S \\ S' & R \end{bmatrix} \begin{bmatrix} M \\ N \end{bmatrix} = (C_z M + D_{zu} N)' (C_z M + D_{zu} N)$$

and

$$MA' + N'B' + AM + BN + (C_z M + D_{zu} N)' (C_z M + D_{zu} N) \leq 0$$

$\iff$

$$\begin{bmatrix} MA' + N'B' + AM + BN & MC_z' + N'D_{zu}' \\ C_z M + D_{zu} N & -I \end{bmatrix} \leq 0$$

LQR: solution via LMI

Thus, the LQR problem can be solved via the LMI

$$\begin{aligned}
 & \text{minimize} \quad \gamma \\
 & \quad \quad \quad M=M', N \\
 & \text{subject to} \quad M > 0 \\
 & \quad \quad \quad \begin{bmatrix} \gamma & x_0' \\ x_0 & M \end{bmatrix} \geq 0 \iff \gamma \geq x_0' M^{-1} x_0 \\
 & \quad \quad \quad \begin{bmatrix} MA' + N'B' + AM + BN & MC_z' + N'D_{zu}' \\ C_z M + D_{zu} N & -I \end{bmatrix} \leq 0
 \end{aligned}$$

and the resulted state-feedback gain $K = NM^{-1}$.

Remark. We can add other LMIs to the “subject to” part above, incorporating additional requirements, like guaranteed damping or decay rate. Uncertainty can be incorporated as well. However, in such cases the assumption $u = Kx$ might become restrictive.

Remark. The dependence of one of the resulted LMIs on x_0 is inconvenient (compare with the Riccati based solution). A way to circumvent that is to minimize a size of M^{-1} , which would yield small $x_0' M^{-1} x_0$ for all x_0 . A common approach in LQR (H_2 -approach) is to minimize $\text{tr}(M^{-1})$ or, equivalently, to minimize $\text{tr}(\Gamma)$ such that $\begin{bmatrix} \Gamma & I \\ I & M \end{bmatrix} \geq 0$.

Riccati connections

We know that the CARE $A'\bar{X} + \bar{X}A + Q - (S + \bar{X}B)R^{-1}(S' + B'\bar{X}) = 0$ is equivalent to (Slide 14, Lecture 4)

$$(A + BK)'\bar{X} + \bar{X}(A + BK) + (C_z + D_{zu}K)'(C_z + D_{zu}K) = 0$$

for $K = -R^{-1}(S' + B'\bar{X})$. It is readily seen that

$$\begin{bmatrix} (A + BK)'\bar{X} + \bar{X}(A + BK) & C_z' + K'D_{zu}' \\ C_z + D_{zu}K & -I \end{bmatrix} \leq 0.$$

This suggests³ that $P = \bar{X}$ and $K = -R^{-1}(S' + B'\bar{X})$ solve

$$\begin{array}{ll} \underset{P=P', K}{\text{minimize}} & \text{tr } P \\ \text{subject to} & P > 0 \\ & (A + BK)'P + P(A + BK) + \begin{bmatrix} I & K' \end{bmatrix} \begin{bmatrix} Q & S \\ S' & R \end{bmatrix} \begin{bmatrix} I \\ K \end{bmatrix} \leq 0 \end{array}$$

³The rigorous proof is quite technical.

Outline

Background: Lyapunov Stability and Lyapunov's direct method

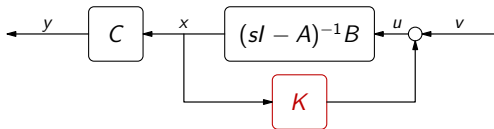
Analysis via quadratic Lyapunov functions

State-feedback regulation via quadratic Lyapunov functions and LMIs

LQR design via LMIs

State observers

The elephant in the room



State feedback is

- (relatively) easy to design
- static, so easy to implement
- fit for powerful Lyapunov, LQR, H_∞ (to be learned) design methods

Q: what if the state vector cannot be measured directly?

State reconstruction

Consider

$$\begin{cases} \dot{x}(t) = Ax(t) + Bu(t), & x(0) = x_0, \\ y(t) = Cx(t). \end{cases}$$

for $A \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{n \times m}$, $C \in \mathbb{R}^{p \times n}$, and $\text{rank } C = p < n$.

If x comprises derivatives of y , we may try to recover x by calculating those derivatives numerically. This may be feasible, but is

- prone to measurement inaccuracies.

Moreover, in some situations (e.g. in the two-tank example from Lecture 2) the state does not contain derivatives of y .

A control way is to

- build an implementable (i.e. having a proper transfer function) system reconstructing x as the response to measurements of u and y only.

Such a system is known as *state observer* or simply *observer*.

State reconstruction

Consider

$$\begin{cases} \dot{x}(t) = Ax(t) + Bu(t), & x(0) = x_0, \\ y(t) = Cx(t). \end{cases}$$

for $A \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{n \times m}$, $C \in \mathbb{R}^{p \times n}$, and $\text{rank } C = p < n$.

If x comprises derivatives of y , we may try to recover x by calculating those derivatives numerically. This may be feasible, but is

- prone to measurement inaccuracies.

Moreover, in some situations (e.g. in the two-tank example from Lecture 2) the state does not contain derivatives of y .

A control way is to

- build an implementable (i.e. having a proper transfer function) system reconstructing x as the response to measurements of u and y only.

Such a system is known as **state observer** or simply **observer**.

Naïve state observer

A possible approach is to construct a **virtual plant**, like

$$\dot{\hat{x}}(t) = A\hat{x}(t) + Bu(t), \quad \hat{x}(0) = \hat{x}_0,$$

for some initial guess $\hat{x}_0$. The observation error $\epsilon := x - \hat{x}$ satisfies

$$\dot{\epsilon}(t) = A\epsilon(t), \quad \epsilon(0) = x_0 - \hat{x}_0$$

which are autonomous dynamics, driven only by the mismatch between $\hat{x}(0)$ and $x(0)$, with $\epsilon(t) = e^{At}(x_0 - \hat{x}_0)$.

Good news:

- if A is Hurwitz, then $\lim_{t \rightarrow \infty} \epsilon(t) = 0$, i.e. $\hat{x}(t) \rightarrow x(t)$ asymptotically no matter what u is, provided we know it, of course

Bad news:

- we cannot affect error dynamics
- if A is unstable, then $\hat{x}$ does not converge to x

Naïve state observer

A possible approach is to construct a **virtual plant**, like

$$\dot{\hat{x}}(t) = A\hat{x}(t) + Bu(t), \quad \hat{x}(0) = \hat{x}_0,$$

for some initial guess $\hat{x}_0$. The observation error $\epsilon := x - \hat{x}$ satisfies

$$\dot{\epsilon}(t) = A\epsilon(t), \quad \epsilon(0) = x_0 - \hat{x}_0$$

which are autonomous dynamics, driven only by the mismatch between $\hat{x}(0)$ and $x(0)$, with $\epsilon(t) = e^{At}(x_0 - \hat{x}_0)$.

Good news:

- if A is Hurwitz, then $\lim_{t \rightarrow \infty} \epsilon(t) = 0$, i.e. $\hat{x}(t) \rightarrow x(t)$ asymptotically no matter what u is, provided we know it, of course

Bad news:

- we cannot affect error dynamics
- if A is unstable, then $\hat{x}$ does not converge to x

Naïve state observer

A possible approach is to construct a **virtual plant**, like

$$\dot{\hat{x}}(t) = A\hat{x}(t) + Bu(t), \quad \hat{x}(0) = \hat{x}_0,$$

for some initial guess $\hat{x}_0$. The observation error $\epsilon := x - \hat{x}$ satisfies

$$\dot{\epsilon}(t) = A\epsilon(t), \quad \epsilon(0) = x_0 - \hat{x}_0$$

which are autonomous dynamics, driven only by the mismatch between $\hat{x}(0)$ and $x(0)$, with $\epsilon(t) = e^{At}(x_0 - \hat{x}_0)$.

Good news:

- if A is Hurwitz, then $\lim_{t \rightarrow \infty} \epsilon(t) = 0$, i.e. $\hat{x}(t) \rightarrow x(t)$ asymptotically no matter what u is, provided we know it, of course

Bad news:

- we cannot affect error dynamics
- if A is unstable, then $\hat{x}$ does not converge to x

Naïve state observer

A possible approach is to construct a **virtual plant**, like

$$\dot{\hat{x}}(t) = A\hat{x}(t) + Bu(t), \quad \hat{x}(0) = \hat{x}_0,$$

for some initial guess $\hat{x}_0$. The observation error $\epsilon := x - \hat{x}$ satisfies

$$\dot{\epsilon}(t) = A\epsilon(t), \quad \epsilon(0) = x_0 - \hat{x}_0$$

which are autonomous dynamics, driven only by the mismatch between $\hat{x}(0)$ and $x(0)$, with $\epsilon(t) = e^{At}(x_0 - \hat{x}_0)$.

Good news:

- if A is Hurwitz, then $\lim_{t \rightarrow \infty} \epsilon(t) = 0$, i.e. $\hat{x}(t) \rightarrow x(t)$ asymptotically no matter what u is, provided we know it, of course

Bad news:

- we cannot affect error dynamics
- if A is unstable, then $\hat{x}$ does not converge to x

Full-order state observer

Naïve observer

- ignores the information about x , available in the measurement y .

With the unity-feedback architecture in mind, consider adding the measured mismatch $y - C\hat{x}$ (aka innovations signal) to the algorithm,

$$\begin{aligned}\dot{\hat{x}}(t) &= A\hat{x}(t) + Bu(t) - L(y(t) - C\hat{x}(t)), & \hat{x}(0) &= \hat{x}_0 \\ &= (A + LC)\hat{x}(t) + Bu(t) - Ly(t), & \hat{x}(0) &= \hat{x}_0\end{aligned}$$

for some gain $L \in \mathbb{R}^{n \times p}$. In this case,

$$\dot{\epsilon}(t) = \underbrace{(A + LC)}_{A_e} \epsilon(t), \quad \epsilon(0) = x_0 - \hat{x}_0.$$

Now we potentially

- can affect the error dynamics via a design parameter L .

Q: what freedom we have in assigning $\text{spec}(A + LC)$ by the choice of L ?

Full-order state observer

Naïve observer

- ignores the information about x , available in the measurement y .

With the unity-feedback architecture in mind, consider adding the measured mismatch $y - C\hat{x}$ (aka **innovations signal**) to the algorithm,

$$\begin{aligned}\dot{\hat{x}}(t) &= A\hat{x}(t) + Bu(t) - L(y(t) - C\hat{x}(t)), & \hat{x}(0) &= \hat{x}_0 \\ &= (A + LC)\hat{x}(t) + Bu(t) - Ly(t), & \hat{x}(0) &= \hat{x}_0\end{aligned}$$

for some gain⁴ $L \in \mathbb{R}^{n \times p}$. In this case,

$$\begin{aligned}\dot{\epsilon}(t) &= (A + LC)\epsilon(t), & \epsilon(0) &= x_0 - \hat{x}_0, \\ \epsilon(t) &= e^{(A + LC)t} \epsilon(0).\end{aligned}$$

Now we potentially

- can affect the error dynamics via a design parameter L .

Q: what freedom we have in assigning $\text{spec}(A + LC)$ by the choice of L ?

⁴This L is for David G. Luenberger.

Full-order state observer

Naïve observer

- ignores the information about x , available in the measurement y .

With the unity-feedback architecture in mind, consider adding the measured mismatch $y - C\hat{x}$ (aka **innovations signal**) to the algorithm,

$$\begin{aligned}\dot{\hat{x}}(t) &= A\hat{x}(t) + Bu(t) - L(y(t) - C\hat{x}(t)), & \hat{x}(0) &= \hat{x}_0 \\ &= (A + LC)\hat{x}(t) + Bu(t) - Ly(t), & \hat{x}(0) &= \hat{x}_0\end{aligned}$$

for some gain $L \in \mathbb{R}^{n \times p}$. In this case,

$$\dot{\epsilon}(t) = \underbrace{(A + LC)}_{:= A_L} \epsilon(t), \quad \epsilon(0) = x_0 - \hat{x}_0.$$

Now we potentially

- can affect the error dynamics via a design parameter L .

Q: what freedom we have in assigning $\text{spec}(A + LC)$ by the choice of L ?

Full-order state observer

Naïve observer

- ignores the information about x , available in the measurement y .

With the unity-feedback architecture in mind, consider adding the measured mismatch $y - C\hat{x}$ (aka **innovations signal**) to the algorithm,

$$\begin{aligned}\dot{\hat{x}}(t) &= A\hat{x}(t) + Bu(t) - L(y(t) - C\hat{x}(t)), & \hat{x}(0) &= \hat{x}_0 \\ &= (A + LC)\hat{x}(t) + Bu(t) - Ly(t), & \hat{x}(0) &= \hat{x}_0\end{aligned}$$

for some gain $L \in \mathbb{R}^{n \times p}$. In this case,

$$\dot{\epsilon}(t) = \underbrace{(A + LC)}_{:= A_L} \epsilon(t), \quad \epsilon(0) = x_0 - \hat{x}_0.$$

Now we potentially

- can affect the error dynamics via a design parameter L .

Q: what freedom we have in assigning $\text{spec}(A + LC)$ by the choice of L ?

Special case: observer form

Assume that

$$A = \begin{bmatrix} -a_{n-1} & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ -a_1 & 0 & \cdots & 1 \\ -a_0 & 0 & \cdots & 0 \end{bmatrix} \quad \text{and} \quad C = \begin{bmatrix} 1 & 0 & \cdots & 0 \end{bmatrix}.$$

Choosing

$$L = \begin{bmatrix} l_{n-1} \\ \vdots \\ l_1 \\ l_0 \end{bmatrix} \quad \Rightarrow \quad A_L = A + LC = \begin{bmatrix} -(a_{n-1} - l_{n-1}) & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ -(a_1 - l_1) & 0 & \cdots & 1 \\ -(a_0 - l_0) & 0 & \cdots & 0 \end{bmatrix}$$

is still an observer form (companion matrix) and its characteristic polynomial

$$\chi_{A_L}(\lambda) = \lambda^n + (a_{n-1} - l_{n-1})\lambda^{n-1} + \cdots + (a_1 - l_1)\lambda + (a_0 - l_0).$$

Special case: observer form (contd)

Therefore, any desired observer characteristic polynomial, say

$$\hat{\chi}(s) = s^n + \hat{\chi}_{n-1}s^{n-1} + \cdots + \hat{\chi}_1s + \hat{\chi}_0$$

for some coefficients $\hat{\chi}_i > 0$, can be assigned by

$$L = \begin{bmatrix} a_{n-1} - \hat{\chi}_{n-1} \\ \vdots \\ a_1 - \hat{\chi}_1 \\ a_0 - \hat{\chi}_0 \end{bmatrix} \implies A_L = \begin{bmatrix} -\hat{\chi}_{n-1} & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ -\hat{\chi}_1 & 0 & \cdots & 1 \\ -\hat{\chi}_0 & 0 & \cdots & 0 \end{bmatrix}$$

Q: under what condition this can be said about an arbitrary realization?

A: iff the pair (A', C') is controllable

Q: what is the observer gain L for the observer form?

Special case: observer form (contd)

Therefore, any desired observer characteristic polynomial, say

$$\hat{\chi}(s) = s^n + \hat{\chi}_{n-1}s^{n-1} + \cdots + \hat{\chi}_1s + \hat{\chi}_0$$

for some coefficients $\hat{\chi}_i > 0$, can be assigned by

$$L = \begin{bmatrix} a_{n-1} - \hat{\chi}_{n-1} \\ \vdots \\ a_1 - \hat{\chi}_1 \\ a_0 - \hat{\chi}_0 \end{bmatrix} \implies A_L = \begin{bmatrix} -\hat{\chi}_{n-1} & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ -\hat{\chi}_1 & 0 & \cdots & 1 \\ -\hat{\chi}_0 & 0 & \cdots & 0 \end{bmatrix}$$

Q: under what condition this can be said about an arbitrary realization?

A: iff the pair (A', C') is controllable

Special case: observer form (contd)

Therefore, any desired observer characteristic polynomial, say

$$\hat{\chi}(s) = s^n + \hat{\chi}_{n-1}s^{n-1} + \cdots + \hat{\chi}_1s + \hat{\chi}_0$$

for some coefficients $\hat{\chi}_i > 0$, can be assigned by

$$L = \begin{bmatrix} a_{n-1} - \hat{\chi}_{n-1} \\ \vdots \\ a_1 - \hat{\chi}_1 \\ a_0 - \hat{\chi}_0 \end{bmatrix} \implies A_L = \begin{bmatrix} -\hat{\chi}_{n-1} & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ -\hat{\chi}_1 & 0 & \cdots & 1 \\ -\hat{\chi}_0 & 0 & \cdots & 0 \end{bmatrix}$$

Q: under what condition this can be said about an arbitrary realization?

A: iff the pair (A', C') is controllable

$\text{spec}(A + LC) = \text{spec}(A' + C'L')$ and the latter is in the “state feedback” form