In this lecture, we study two pillars of controls: $\mathcal{H}_2$ and $\mathcal{H}_\infty$ control. We frame LQR and LQG as types of $\mathcal{H}_2$ control. We also cover the bounded real lemma, an important ingredient in $\mathcal{H}_\infty$ control.

1 System norms

We previously discussed vector norms and matrix norms. We will review these concepts and add one more; the concept of a system norm.

1.1 Vector Norms

- $L_1$ norm: $\|x\|_1 = \sum_i |x_i|$
- $L_2$ norm: $\|x\|_2 = \sqrt{\sum_i x_i^2}$
- $L_p$ norm: $\|x\|_p = \left(\sum_i |x_i|^p\right)^{1/p}$
- $L_\infty$ norm: $\|x\|_\infty = \max_i |x_i|$

1.2 Matrix Norms

- Spectral norm / induced 2-norm / matrix norm: $\|A\|_2 = \|Ax\|_2 = \max_{x \neq 0} \frac{\|Ax\|}{\|x\|} = \sigma_1(A)$
- Frobenius norm: $\|A\|_F = \sum_{ij} |A_{ij}|^2 = \text{tr}(A^T A) = \sum_i \sigma_i(A)^2$

1.3 System Norms

$$u \rightarrow \begin{align*}
x_{t+1} &= A_tx_t + Bu_t \\
y_t &= Cx_t \end{align*} \rightarrow y$$

- The $\mathcal{H}_2$ norm squared is $\|G\|_2^2 = \sum_{t=0}^\infty \|Y_t\|_F^2$, where either $U_0 = I$ and $U_t = 0$ for $t \geq 1$ (Frobenius norm of the matrix impulse response), or it is the steady-state average covariance when the input is white noise. In other words, $\|G\|_2^2 = \lim_{N \to \infty} \frac{1}{N} \sum_{t=0}^{N-1} \mathbb{E}[y_t^2]$, where $u_t \sim \mathcal{N}(0, I)$. Both definitions are equivalent.
The $H_\infty$ norm is the maximum gain of the system $G$ when the input is square-summable. In other words, $\|G\|_\infty = \sup_{u \in \ell_2} \frac{\|Gu\|}{\|u\|}$.

We can view $\|G\|_2$ and $\|G\|_\infty$ as analogous to $\|A\|_F$ and $\|A\|$ for matrices, respectively. It is unfortunate that the naming conventions are not more consistent. The spectral norm for matrices is often written as $\|A\|_2$ whereas the corresponding norm for systems is $\|G\|_\infty$. Meanwhile, $\|G\|_2$ for systems means something completely different, and is not an induced norm at all.

## 2 The $H_2$ norm

### 2.1 Impulse response

The (matrix) impulse response of the system assumes matrix inputs of $U_0 = I$ and $U_t = 0$ for $t \geq 1$, with an initial state of $X_0 = 0$. This yields matrix outputs of:

\[
Y_0 = 0, \quad Y_1 = CB, \quad Y_2 = CAB, \quad Y_3 = CA^2B, \ldots \quad Y_k = CA^{k-1}B
\]

Therefore, the $H_2$ norm of a system is given by:

\[
\|G\|^2_2 = \text{tr} \left( B^T C^T CB + B^T A^T C^T CAB + B^T (A^T)^2 C^T CA^2 B + \ldots \right)
\]

Call the term inside the inner bracket $Q$, and note that $Q$ satisfies a Lyapunov equation. In fact, $Q$ is the observability Gramian of $G$. We conclude that

\[
\|G\|^2_2 = \text{tr} (B^TQB), \quad \text{where } Q \text{ satisfies } A^T QA - Q + C^T C = 0
\]

Using the fact that $\text{tr}(XY) = \text{tr}(YX)$, write $\text{tr}(B^T(A^T)^k C^T CA^k B) = \text{tr}(CA^k BB^T(A^T)^k C^T)$. This leads us to an alternative but equivalent definition in terms of the controllability Gramian:

\[
\|G\|^2_2 = \text{tr}(CPC^T), \quad \text{where } P \text{ satisfies } APA^T - P + BB^T = 0
\]

### 2.2 LQR interpretation

Recall the LQR problem is to find inputs $\{u_0, u_1, \ldots\}$ that minimize the cost

\[
J = \sum_{t=0}^{\infty} x_t^T Q x_t + u_t^T R u_t
\]

where $x_{t+1} = Ax_t + Bu_t$. We can view this as a case of minimizing $\|y\|^2$, where we defined the output signal as

\[
y_t = \begin{bmatrix} Q^{1/2} & 0 \\ 0 & R^{1/2} \end{bmatrix} \begin{bmatrix} x_t \\ u_t \end{bmatrix} = Fx_t + Hu_t
\]
When using the control law \( u_t = Kx_t \), we can rewrite the closed-loop system \( G_{CL} \) as:

\[
\begin{align*}
    x_{t+1} &= (A + BK)x_t + w_t \\
    y_t &= \begin{bmatrix} Q^{1/2} \\ R^{1/2}K \end{bmatrix} x_t 
\end{align*}
\]

Now consider the \( \mathcal{H}_2 \) norm of this system. Specifically, imagine that \( w_{-1} = x_0 \) and we let \( w_t = 0 \) for all future timesteps. Then using the observability Gramian formulation (renaming \( Q \) to \( X \)), we have

\[
\|G_{CL}\|_2^2 = x_0^T X x_0,
\]

where \( X \) satisfies

\[
(A - BK)^T X (A - BK) - X + (Q + K^T R K) = 0
\]

Note that when \( K \) is the optimal LQR gain, the above Lyapunov equation becomes the DARE for the LQR problem. Therefore, we can interpret LQR control as finding the controller that minimizes the \( \mathcal{H}_2 \) norm of the closed-loop map from \( w \) to \( y \).

### 2.3 LQG interpretation

We omit the details, but the LQG interpretation is similar to the LQR interpretation. The optimal LQG controller also minimizes the \( \mathcal{H}_2 \) norm of the closed-loop map from the external disturbance \( w \) to the performance output channel.

### 3 The \( \mathcal{H}_\infty \) norm

The \( \mathcal{H}_\infty \) is a worst-case norm. It is defined as

\[
\|G\|_\infty = \sup_{u \in \ell^2} \frac{\|G u\|}{\|u\|}
\]

We can view this as the matrix norm of the infinite-dimensional matrix that maps \( u \) to \( y \):

\[
\begin{bmatrix}
    y_0 \\
    y_1 \\
    y_2 \\
    y_3 \\
    \vdots
\end{bmatrix} = \underbrace{egin{bmatrix}
    0 & 0 & 0 & 0 \\
    CB & 0 & 0 & 0 \\
    CAB & CB & 0 & 0 \\
    CA^2B & CAB & CB & 0 \\
    \vdots & \ddots & \ddots & \ddots & \ddots
\end{bmatrix}}_{\Phi} \begin{bmatrix}
    u_0 \\
    u_1 \\
    u_2 \\
    u_3 \\
    \vdots
\end{bmatrix}
\]

In other words, \( \|G\|_\infty = \|\Phi\| \). This matrix \( \Phi \) is block lower-triangular (because the map \( u \to y \) is causal), and it has a block-Toeplitz structure (because the map \( u \to y \) is time-invariant). This is different from the Hankel matrix we saw when discussing the Hankel singular values of a system (that matrix mapped past inputs to future outputs, whereas the matrix \( \Phi \) maps future inputs to future outputs). Unlike the Hankel case, the matrix \( \Phi \) does not have a nice factorization. Instead, we will use a Lyapunov-like approach.
Suppose we can find a function $V : \mathbb{R}^n \to \mathbb{R}$ such that

(i) $V(x) > 0 \quad \forall x \neq 0$

(ii) $V(x_{t+1}) - V(x_t) \leq \gamma^2 \|u_t\|^2 - \|y_t\|^2$

Then we can sum the second inequality and use the first inequality to deduce:

$$-V(x_0) \leq V(x_N) - V(x_0) \leq \gamma^2 \sum_{t=0}^{N-1} \|u_t\|^2 - \sum_{t=0}^{N-1} \|y_t\|^2$$

Rearranging, taking the limit $N \to \infty$, and assuming $u \in \ell_2$, we obtain

$$\sum_{t=0}^{N-1} \|y_t\|^2 - V(x_0) \leq \gamma^2 \sum_{t=0}^{N-1} \|u_t\|^2 \leq \gamma^2 \|u\|^2$$

This holds for all $N$, so we deduce that $y \in \ell_2$ as well, and we have

$$\|y\|^2 - V(x_0) \leq \gamma^2 \|u\|^2$$

This leads to:

$$\frac{\|y\|^2}{\|u\|^2} - \frac{V(x_0)}{\|u\|^2} \leq \gamma^2$$

This holds for all $u$, so in particular, since the system is linear, we can scale $u$ and $y$ will scale as well. We deduce that

$$\|y\| \leq \gamma \|u\|$$

Let’s look for a quadratic $V(x)$. Specifically, pick $V(x) = x^T P x$, with $P > 0$. Substituting this into the second inequality, we obtain:

$$x_{t+1}^T P x_{t+1} - x_t^T P x_t \leq \gamma^2 \|u_t\|^2 - \|y_t\|^2$$

Substituting $x_{t+1} = Ax_t + Bu_t$ and $y_t = Cx_t$, we obtain a quadratic inequality in $(x_t, u_t)$, which we can write as

$$\begin{bmatrix} x_t^T P A - P + C^T C & A^T P B \\ B^T P A & B^T P B - \gamma^2 I \end{bmatrix} \begin{bmatrix} x_t \\ u_t \end{bmatrix} \succeq 0$$

this will be satisfied if the matrix in the middle is negative semidefinite. All we have proven so far is that if we can find $P > 0$ such that this matrix condition holds, we will have proven that the gain of the system $G$ is less than $\gamma$. It turns out the converse is true as well! A quadratic function $P$ with these properties exists if and only if the system has gain bound less than $\gamma$. This result is known as the bounded real lemma. Here is a complete statement.

**Theorem 3.1.** Suppose $(A, B, C)$ is a minimal realization. In other words, $(A, B)$ is controllable and $(A, C)$ is observable. Then the following statements are equivalent.

(i) $\|y\| \leq \gamma \|u\|$ for all $u \in \ell_2$

(ii) There exists a matrix $P > 0$ that satisfies

$$\begin{bmatrix} A^T P A - P + C^T C & A^T P B \\ B^T P A & B^T P B - \gamma^2 I \end{bmatrix} \succeq 0$$

(2)
Eq. (2) in Theorem 3.1 is a linear matrix inequality. The set of $P$ that satisfy it is a convex set, since $P$ appears linearly. Moreover, it is also linear in $\gamma^2$. So the optimization problem of finding the smallest $\gamma^2$ such that there exists a $P > 0$ satisfying (2) is a convex optimization problem.

For a detailed proof of Theorem 3.1, please see the supplementary notes.

3.1 Schur Complements

A useful property of Schur complements is that:

$$\begin{bmatrix} A & B \\ B^T & D \end{bmatrix} > 0 \iff \begin{cases} A > 0 \\ D - B^T A^{-1} B > 0 \end{cases} \iff \begin{cases} A > 0 \\ D > 0 \end{cases}$$

This result follows from the block-LDU and block-UDL factorizations seen earlier in the class. We can therefore apply this to the LMI (2) and conclude that:

$$\begin{bmatrix} A^T P A - P + C^T C & A^T P B \\ B^T P A & B^T P B - \gamma^2 I \end{bmatrix} > 0$$

$$\begin{bmatrix} P - A^T P A & -A^T P B \\ -B^T P A & \gamma^2 I - B^T P B \end{bmatrix} - \begin{bmatrix} C^T \\ 0 \end{bmatrix} \begin{bmatrix} C \\ 0 \end{bmatrix} > 0$$

$$\begin{bmatrix} P & 0 & C^T \\ 0 & \gamma^2 I & 0 \\ C & 0 & I \end{bmatrix} - \begin{bmatrix} A^T \\ B^T \end{bmatrix} P \begin{bmatrix} A & B & 0 \end{bmatrix} > 0$$

$$\begin{bmatrix} P & 0 & C^T & A^T \\ 0 & \gamma^2 I & 0 & B^T \\ C & 0 & I & 0 \\ A & B & 0 & P^{-1} \end{bmatrix} > 0$$

$$\begin{bmatrix} P & 0 & C^T & A^T P \\ 0 & \gamma^2 I & 0 & B^T P \\ C & 0 & I & 0 \\ P A & P B & 0 & P \end{bmatrix} > 0$$

We can also make the observation that since $A^T P A - P + C^T C$ is negative definite, then $A^T P A - P + C^T C + H = 0$ for some $H > 0$, and we can conclude from our standard results on Lyapunov equations that $A$ must be Schur-stable. In other words, having a finite system gain implies stability. For linear finite-dimensional systems, the converse is also true (stable systems must have a finite gain bound).