

Sampled-Data Control Systems (036709)

lecture no. 5

Leonid Mirkin

Faculty of Mechanical Engineering
Technion — IIT



Dynamics and complexity

Causal systems may be classified as

- static¹, those where the past does not affect the present and the future
- dynamic, those where it does.

The complexity of dynamic phenomena is naturally expressed in terms of

— the “size of their memory,”

i.e. the amount of information required to store the past.

The notion of **state space**, that we start studying today, plays an important role in characterizing this size.

¹Aka memoryless.

Dynamics and complexity

Causal systems may be classified as

- static¹, those where the past does not affect the present and the future
- dynamic, those where it does.

The complexity of dynamic phenomena is naturally expressed in terms of

- the “size of their memory,”

i.e. the amount of information required to store the past.

The notion of **state space**, that we start studying today, plays an important role in characterizing this size.

¹Aka memoryless.

Dynamics and complexity (contd)

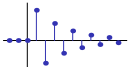
An LSI system $G : u \mapsto y$ is completely determined by its impulse response,

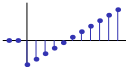
$$y[t] = \sum_{s \in \mathbb{Z}} g[t - s]u[s].$$

Static LSI systems have the impulse response $g[t] = k\delta[t]$, dynamic systems have $g[t] \neq 0$ for at least some $t > 0$. A relevant question is

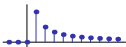
- can memory be quantified in terms of $g[t]$?

For example, let

$$g_1[t] = a^{t-1}b\mathbb{1}[t-1] =$$


$$g_2[t] = (b_0 + b_1 t)\mathbb{1}[t] =$$


$$g_3[t] = \delta[t-3] =$$


$$g_4[t] = \frac{1}{t}\mathbb{1}[t-1] =$$


Which of them has the largest memory?

Outline

Linear algebra preliminaries (self study)

State space

Structural properties of state-space realizations

Lyapunov stability

Outline

Linear algebra preliminaries (self study)

State space

Structural properties of state-space realizations

Lyapunov stability

Notation and terminology

A matrix $A \in \mathbb{R}^{n \times m}$ is an $n \times m$ table of real numbers,

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{bmatrix} := [a_{ij}],$$

accompanied by their manipulation rules (addition, multiplication, etc). A matrix $A \in \mathbb{R}^{n \times m}$ is said to be

- square if $n = m$, tall if $n > m$, and fat if $n < m$;
- diagonal ($A = \text{diag}\{a_i\}$) if it is square and $a_{ij} = 0$ whenever $i \neq j$;
- lower (upper) triangular if $a_{ij} = 0$ whenever $i > j$ ($i < j$);
- symmetric if $A = A'$, where the transpose $A' := [a_{ji}]$.

The identity matrix $I_n := \text{diag}\{1\} \in \mathbb{R}^{n \times n}$ (if the dimension is clear, just I). It is a convention that $A^0 = I$ for every nonzero square A .

- trace of a square $A \in \mathbb{R}^{n \times n}$ is $\text{tr}(A) = \sum_{i=1}^n a_{ii}$.

Rank

The rank of a matrix $A \in \mathbb{R}^{n \times m}$, denoted $\text{rank } A$, is the number of linearly independent columns of A . Clearly, $\text{rank } A \leq \min\{n, m\}$. We say that

- A is a full-rank matrix if $\text{rank } A = \min\{n, m\}$
(square full-rank matrices are also termed nonsingular or invertible)
- A is a reduced-rank matrix if $\text{rank } A < \min\{n, m\}$
- A has full row (column) rank if $\text{rank } A = n$ ($\text{rank } A = m$)

The rank of a matrix may be seen as its **complexity measure**.

Properties

- $\text{rank } A$ equals the number of linearly independent rows of A
- $\text{rank}(AB) \leq \min\{\text{rank } A, \text{rank } B\}$
- $\text{rank } A = \text{rank}(SAT)$ if $A \in \mathbb{R}^{n \times m}$, $\text{rank } S = n$, and $\text{rank } T = m$
- A has full row rank iff there is no $\eta \neq 0$ such that $\eta^T A = 0$
- A has full column rank iff there is no $\eta \neq 0$ such that $A\eta = 0$

Rank

The rank of a matrix $A \in \mathbb{R}^{n \times m}$, denoted $\text{rank } A$, is the number of linearly independent columns of A . Clearly, $\text{rank } A \leq \min\{n, m\}$. We say that

- A is a full-rank matrix if $\text{rank } A = \min\{n, m\}$
(square full-rank matrices are also termed nonsingular or invertible)
- A is a reduced-rank matrix if $\text{rank } A < \min\{n, m\}$
- A has full row (column) rank if $\text{rank } A = n$ ($\text{rank } A = m$)

The rank of a matrix may be seen as its complexity measure.

Properties:

- $\text{rank } A$ equals the number of linearly independent rows of A
- $\text{rank}(AB) \leq \min\{\text{rank } A, \text{rank } B\}$
- $\text{rank } A = \text{rank}(SAT)$ if $A \in \mathbb{R}^{n \times m}$, $\text{rank } S = n$, and $\text{rank } T = m$
- A has full row rank iff there is no $\eta \neq 0$ such that $\eta' A = 0$
- A has full column rank iff there is no $\eta \neq 0$ such that $A\eta = 0$

Rank decomposition

Let $A \in \mathbb{R}^{n \times m}$. If $\text{rank } A = r$, then there are rank- r matrices $B \in \mathbb{R}^{n \times r}$ and $C \in \mathbb{R}^{r \times m}$ such that

$$A = BC.$$

Example

Rank-1

$$A = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 2 & 2 & 2 \\ 3 & 3 & 3 & 3 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 \end{bmatrix}.$$

Lemma

Let $A = BC$ be a rank decomposition of a rank- r matrix A . $B_1 \in \mathbb{R}^{n \times r}$ and $C_1 \in \mathbb{R}^{r \times m}$ are rank- r matrices satisfying $A = B_1 C_1$ iff

$$C_1 = CT^{-1} \quad \text{and} \quad B_1 = TB$$

for a nonsingular $T \in \mathbb{R}^{r \times r}$.

Rank decomposition

Let $A \in \mathbb{R}^{n \times m}$. If $\text{rank } A = r$, then there are rank- r matrices $B \in \mathbb{R}^{n \times r}$ and $C \in \mathbb{R}^{r \times m}$ such that

$$A = BC.$$

Example

Rank-1

$$A = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 2 & 2 & 2 \\ 3 & 3 & 3 & 3 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 \end{bmatrix}.$$



Lemma

Let $A = BC$ be a rank decomposition of a rank- r matrix A . $B_1 \in \mathbb{R}^{n \times r}$ and $C_1 \in \mathbb{R}^{r \times m}$ are rank- r matrices satisfying $A = B_1 C_1$ iff

$$C_1 = CT^{-1} \quad \text{and} \quad B_1 = TB$$

for a nonsingular $T \in \mathbb{R}^{r \times r}$.

Rank decomposition

Let $A \in \mathbb{R}^{n \times m}$. If $\text{rank } A = r$, then there are rank- r matrices $B \in \mathbb{R}^{n \times r}$ and $C \in \mathbb{R}^{r \times m}$ such that

$$A = BC.$$

Example

Rank-1

$$A = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 2 & 2 & 2 \\ 3 & 3 & 3 & 3 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 \end{bmatrix}.$$



Lemma

Let $A = BC$ be a rank decomposition of a rank- r matrix A . $B_1 \in \mathbb{R}^{n \times r}$ and $C_1 \in \mathbb{R}^{r \times m}$ are rank- r matrices satisfying $A = B_1 C_1$ iff

$$C_1 = CT^{-1} \quad \text{and} \quad B_1 = TB$$

for a nonsingular $T \in \mathbb{R}^{r \times r}$.

Linear equations

Solve in $x \in \mathbb{R}^m$ the equation $Ax = b$ for given $A \in \mathbb{R}^{n \times m}$ and $B \in \mathbb{R}^n$.

- $Ax = b$ is solvable iff $\text{rank } A = \text{rank} \begin{bmatrix} A & b \end{bmatrix}$
- $Ax = b$ has a unique solution iff $\text{rank } A = \text{rank} \begin{bmatrix} A & b \end{bmatrix} = m$
- $Ax = b$ has infinitely many solutions iff $\text{rank } A = \text{rank} \begin{bmatrix} A & b \end{bmatrix} < m$

Example

Let (here $n = 3$ and $m = 2$)

$$\begin{bmatrix} A & b \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & \alpha & 0 \\ 1 & 0 & 1 \end{bmatrix} \implies \begin{bmatrix} 1 & 0 \\ 0 & \alpha \\ 1 & 0 \end{bmatrix} x = \begin{bmatrix} x_1 \\ \alpha x_2 \\ x_1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$$

and $\text{rank } A = \text{rank} \begin{bmatrix} A & b \end{bmatrix}$ for all $\alpha \in \mathbb{R}$. Two cases are possible:

$\alpha \neq 0 \implies \text{rank } A = \text{rank} \begin{bmatrix} A & b \end{bmatrix} = 2$ and $x = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ is unique;

$\alpha = 0 \implies \text{rank } A = \text{rank} \begin{bmatrix} A & b \end{bmatrix} = 1$ and all $x = \begin{bmatrix} 1 \\ x_2 \end{bmatrix}$ solve $Ax = b$.



Linear equations

Solve in $x \in \mathbb{R}^m$ the equation $Ax = b$ for given $A \in \mathbb{R}^{n \times m}$ and $B \in \mathbb{R}^n$.

- $Ax = b$ is solvable iff $\text{rank } A = \text{rank} \begin{bmatrix} A & b \end{bmatrix}$
- $Ax = b$ has a unique solution iff $\text{rank } A = \text{rank} \begin{bmatrix} A & b \end{bmatrix} = m$
- $Ax = b$ has infinitely many solutions iff $\text{rank } A = \text{rank} \begin{bmatrix} A & b \end{bmatrix} < m$

Example

Let (here $n = 3$ and $m = 2$)

$$\left[A \mid b \right] = \left[\begin{array}{cc|c} 1 & 0 & 1 \\ 0 & \alpha & 0 \\ 1 & 0 & 1 \end{array} \right] \implies \begin{bmatrix} 1 & 0 \\ 0 & \alpha \\ 1 & 0 \end{bmatrix} x = \begin{bmatrix} x_1 \\ \alpha x_2 \\ x_1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$$

and $\text{rank } A = \text{rank} \begin{bmatrix} A & b \end{bmatrix}$ for all $\alpha \in \mathbb{R}$. Two cases are possible:

$\alpha \neq 0 \implies \text{rank } A = \text{rank} \begin{bmatrix} A & b \end{bmatrix} = 2$ and $x = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ is unique;

$\alpha = 0 \implies \text{rank } A = \text{rank} \begin{bmatrix} A & b \end{bmatrix} = 1$ and all $x = \begin{bmatrix} 1 \\ x_2 \end{bmatrix}$ solve $Ax = b$.



Eigenvalues and vectors

Given $A \in \mathbb{R}^{n \times n}$, its **eigenvalues** are the solutions $\lambda \in \mathbb{C}$ to

$$\chi_A(\lambda) := \det(\lambda I - A) = \lambda^n + \chi_{n-1}\lambda^{n-1} + \cdots + \chi_1\lambda + \chi_0 = 0$$

(**characteristic equation**). There are n (not necessarily distinct) eigenvalues of A and $\text{spec}(A)$ denotes their set. If $\lambda_i \in \text{spec}(A)$ is such that $\text{Im } \lambda_i \neq 0$, then $\bar{\lambda}_i \in \text{spec}(A)$ too. But $A = A' \implies \text{spec}(A) \subset \mathbb{R}$.

- the spectral radius of A is $\rho(A) := \max_{\lambda_i \in \text{spec}(A)} |\lambda_i|$.

Right and left **eigenvectors** of $\lambda_i \in \text{spec}(A)$ are nonzero vectors such that

$$(\lambda_i I - A)\eta_i = 0 \quad \text{and} \quad \eta_i'(\lambda_i I - A) = 0 \quad \eta_i \neq 0 = \eta_i'$$

respectively. Note that

$$A^k \eta_i = \eta_i \lambda_i^k \quad \text{and} \quad \eta_i' A^k = \lambda_i^k \eta_i' \quad \forall k \in \mathbb{N}$$

so $\lambda_i \in \text{spec}(A) \implies \lambda_i^k \in \text{spec}(A^k)$, keeping left and right eigenvectors.

Eigenvalues and vectors

Given $A \in \mathbb{R}^{n \times n}$, its **eigenvalues** are the solutions $\lambda \in \mathbb{C}$ to

$$\chi_A(\lambda) := \det(\lambda I - A) = \lambda^n + \chi_{n-1}\lambda^{n-1} + \cdots + \chi_1\lambda + \chi_0 = 0$$

(**characteristic equation**). There are n (not necessarily distinct) eigenvalues of A and $\text{spec}(A)$ denotes their set. If $\lambda_i \in \text{spec}(A)$ is such that $\text{Im } \lambda_i \neq 0$, then $\bar{\lambda}_i \in \text{spec}(A)$ too. But $A = A' \implies \text{spec}(A) \subset \mathbb{R}$.

— the spectral radius of A is $\rho(A) := \max_{\lambda_i \in \text{spec}(A)} |\lambda_i|$.

Right and left **eigenvectors** of $\lambda_i \in \text{spec}(A)$ are nonzero vectors such that

$$A\eta_i = \lambda_i\eta_i \iff (\lambda_i I - A)\eta_i = 0 \quad \text{and} \quad \eta'_i(\lambda_i I - A) = 0 \iff \eta'_i A = \lambda_i \eta'_i$$

respectively. Note that

$$A^k \eta_i = \lambda_i^k \eta_i \quad \text{and} \quad \eta'_i A^k = \lambda_i^k \eta'_i, \quad \forall k \in \mathbb{N}$$

so $\lambda_i \in \text{spec}(A) \implies \lambda_i^k \in \text{spec}(A^k)$, keeping left and right eigenvectors.

Similar matrices

$A_1, A_2 \in \mathbb{R}^{n \times n}$ are said to be **similar** if there is a nonsingular $T \in \mathbb{R}^{n \times n}$ for which

$$A_2 = T A_1 T^{-1}.$$

The mapping $A \mapsto T A T^{-1}$ is known as the similarity transformation.

- similarity does not affect the spectrum,

$$\begin{aligned}\chi_{T A T^{-1}}(\lambda) &= \det(\lambda I - T A T^{-1}) = \det T(\lambda I - A)T^{-1} = \det(\lambda I - A) \\ &= \chi_A(\lambda);\end{aligned}$$

- similarity is preserved under matrix power, for every $k \in \mathbb{N}$,

$$(T A T^{-1})^k = T A T^{-1} T A T^{-1} \dots T A T^{-1} = T A A \dots A T^{-1} = T A^k T^{-1};$$

- if $\det A \neq 0$, then $T A^{-1} T^{-1} \cdot T A T^{-1} = T A^{-1} A T^{-1} = T T^{-1} = I$, i.e.

$$(T A T^{-1})^{-1} = T A^{-1} T^{-1}.$$

Cayley–Hamilton

In essence, each square matrix satisfies its own characteristic equation:

$$\chi_A(A) := A^n + \chi_{n-1}A^{n-1} + \cdots + \chi_1A + \chi_0I_n = 0.$$

Important consequences:

- A^k for all $k \geq n$ is a linear combination of A^i , $i = 0, \dots, n-1$, like

$$\begin{aligned} A^n &= -\chi_{n-1}A^{n-1} - \cdots - \chi_1A - \chi_0I_n \\ A^{n+1} &= -\chi_{n-1}A^n - \cdots - \chi_1A^2 - \chi_0A \\ &= \chi_{n-1}(\chi_{n-1}A^{n-1} + \cdots + \chi_1A + \chi_0I_n) - \cdots - \chi_1A^2 - \chi_0A \\ &= (\chi_{n-1}^2 - \chi_{n-2})A^{n-1} + \cdots + (\chi_{n-1}\chi_1 - \chi_0)A + \chi_{n-1}\chi_0I_n \\ &\vdots \end{aligned}$$

- A^{-1} , if exists, is also a linear combination of A^i , $i = 0, \dots, n-1$:

$$A^{-1} = -\frac{1}{\chi_0}(\chi_1I + \cdots + \chi_{n-1}A^{n-2} + A^{n-1}).$$

Matrix functions

Let $f(x) = \sum_{i \in \mathbb{Z}_+} f_i x^i$ be analytic. Its matrix version $f(A)$ is defined as

$$f(A) := \sum_{i \in \mathbb{Z}_+} f_i A^i.$$

Properties:

- similarity is preserved under function mapping,

$$f(TAT^{-1}) = \sum_{i \in \mathbb{Z}_+} f_i (TAT^{-1})^i = T \left(\sum_{i \in \mathbb{Z}_+} f_i A^i \right) T^{-1} = Tf(A)T^{-1};$$

- if λ_j is an eigenvalue of A with the eigenvector η_j , then

$$f(A)\eta_j = \sum_{i \in \mathbb{Z}_+} f_i A^i \eta_j = \sum_{i \in \mathbb{Z}_+} f_i \eta_j \lambda_j^i = \eta_j \sum_{i \in \mathbb{Z}_+} f_i \lambda_j^i = \eta_j f(\lambda_j);$$

- $f(A)$ is singular iff $f(\lambda_j) = 0$ for some $\lambda_j \in \text{spec}(A)$;
- A commutes with any its function, $Af(A) = f(A)A$.

Matrix exponential

The matrix exponential is defined (here $t \in \mathbb{R}$, we shall need it later on) as

$$\exp(At) = e^{At} := I + At + \frac{1}{2!}(At)^2 + \frac{1}{3!}(At)^3 + \dots$$

Properties:

- $(e^{At})' = e^{A't}$
- e^{At} is invertible for all $A \in \mathbb{R}^{n \times n}$ and $t \in \mathbb{R}$, with $(e^{At})^{-1} = e^{-At}$
- $AB = BA \implies e^{(A+B)t} = e^{At}e^{Bt}$ (e.g. true for all A if $B = f(A)$)
- $e^{\text{diag}\{\lambda_i\}t} = \text{diag}\{e^{\lambda_i t}\}$

$$\text{— } \exp \left(\begin{bmatrix} \lambda & 1 & 0 & \dots & 0 \\ 0 & \lambda & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ 0 & 0 & 0 & \dots & \lambda \end{bmatrix} t \right) = \begin{bmatrix} 1 & t & t^2/2! & \dots & t^{n-1}/(n-1)! \\ 0 & 1 & t & \dots & t^{n-2}/(n-2)! \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & t \\ 0 & 0 & 0 & \dots & 1 \end{bmatrix} e^{\lambda t}$$

Sign-definite matrices

A symmetric matrix $A = A' \in \mathbb{R}^{n \times n}$ is said to be

- positive definite ($A > 0$) if $x'Ax > 0$ for all $x \neq 0$,
- positive semidefinite ($A \geq 0$) if $x'Ax \geq 0$ for all x ,
- negative definite ($A < 0$) if $x'Ax < 0$ for all $x \neq 0$ (or if $-A > 0$),
- negative semidefinite ($A \leq 0$) if $x'Ax \leq 0$ for all x (or $-A \geq 0$).

Clearly, $A > 0 \implies \det A \neq 0$. In fact,

- $A > 0$ (≥ 0) iff all its eigenvalues are positive (nonnegative)

Given two matrices $A = A' \in \mathbb{R}^{n \times n}$ and $B = B' \in \mathbb{R}^{n \times n}$, we say

- $A > B$ if $A - B > 0$,
- $A \geq B$ if $A - B \geq 0$,
- $A < B$ if $B - A > 0$ (equivalently, $A - B < 0$),
- $A \leq B$ if $B - A \geq 0$ (equivalently, $A - B \leq 0$).

Sign-definite matrices (contd)

Example

Let $A = \begin{bmatrix} \alpha & -1 \\ -1 & 1 \end{bmatrix}$ for some $\alpha \in \mathbb{R}$, then

$$\begin{aligned} x'Ax &= \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} \alpha & -1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \alpha x_1^2 - 2x_1x_2 + x_2^2 \\ &= (\alpha - 1)x_1^2 + (x_1 - x_2)^2 \end{aligned}$$

and if

$\alpha > 1$ then $x'Ax > 0$ unless $x_1 = x_2 = 0$, for which $x'Ax = 0 \implies A > 0$

$\alpha = 1$ then $x'Ax > 0$ unless $x_1 = x_2$, for which $x'Ax = 0 \implies A \geq 0$

$\alpha < 1$ then $x'Ax$ might be both positive (e.g. if $x_1 = 0 \neq x_2$)
and negative (e.g. if $x_1 = x_2 \neq 0$) $\implies A \not\geq 0$



Block matrices

Let $A \in \mathbb{R}^{n \times m}$ and let $\sum_{i=1}^v n_i = n$ and $\sum_{i=1}^{\mu} m_i = m$ for some $n_i, m_i \in \mathbb{N}$. Then A can be formally split as

$$A = \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1\mu} \\ A_{21} & A_{22} & \cdots & A_{2\mu} \\ \vdots & \vdots & \ddots & \vdots \\ A_{v1} & A_{v2} & \cdots & A_{v\mu} \end{bmatrix} = [A_{ij}]$$

for $A_{ij} = [a_{ij,kl}] \in \mathbb{R}^{n_i \times m_j}$ with entries $a_{ij,kl} = a_{pq}$, where $p = \sum_{r=1}^{i-1} n_r + k$ and $q = \sum_{r=1}^{j-1} m_r + l$. For example, we may present

$$A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \end{bmatrix}$$

for $A_{11} = \begin{bmatrix} 1 & 2 \\ 5 & 6 \end{bmatrix}$, $A_{12} = \begin{bmatrix} 3 \\ 7 \end{bmatrix}$, $A_{13} = \begin{bmatrix} 4 \\ 8 \end{bmatrix}$, $A_{21} = \begin{bmatrix} 9 & 10 \end{bmatrix}$, $A_{22} = 11$, $A_{23} = 12$. This split is not unique and is normally done merely for convenience.

Vector and matrix norms

Norms serve to measure sizes of vectors and matrices. We need

- Euclidean vector norm, $\|v\| := (\sum_{i=1}^n |v_i|^2)^{1/2}$ for a given $v \in \mathbb{C}^n$
- spectral matrix norm, $\|A\| := \sqrt{\rho(A'A)}$ for a given $A \in \mathbb{C}^{n \times m}$
- Frobenius matrix norm, $\|A\|_F := \sqrt{\text{tr}(A'A)}$ for a given $A \in \mathbb{C}^{n \times m}$

Properties:

- $\|A\| = \max_{v \in \mathbb{C}^m, \|v\|=1} \|Av\|$
(i.e. the spectral matrix norm is induced by the Euclidean vector norm)
- $\|AB\| \leq \|A\| \|B\|$ for all $A \in \mathbb{C}^{n \times m}$ and $B \in \mathbb{C}^{m \times p}$ (not true for $\|\cdot\|_F$)
- $\|A\| \leq \|A\|_F \leq \sqrt{\text{rank } A} \|A\|$

A rough, but nevertheless useful, bound:

$$\|e^{At}\| \leq e^{\|A\|t} \text{ for all } t \geq 0.$$

Vector and matrix norms

Norms serve to measure sizes of vectors and matrices. We need

- Euclidean vector norm, $\|v\| := (\sum_{i=1}^n |v_i|^2)^{1/2}$ for a given $v \in \mathbb{C}^n$
- spectral matrix norm, $\|A\| := \sqrt{\rho(A'A)}$ for a given $A \in \mathbb{C}^{n \times m}$
- Frobenius matrix norm, $\|A\|_F := \sqrt{\text{tr}(A'A)}$ for a given $A \in \mathbb{C}^{n \times m}$

Properties:

- $\|A\| = \max_{v \in \mathbb{C}^m, \|v\|=1} \|Av\|$
(i.e. the spectral matrix norm is induced by the Euclidean vector norm)
- $\|AB\| \leq \|A\| \|B\|$ for all $A \in \mathbb{C}^{n \times m}$ and $B \in \mathbb{C}^{m \times p}$ (not true for $\|\cdot\|_F$)
- $\|A\| \leq \|A\|_F \leq \sqrt{\text{rank } A} \|A\|$

A rough, but nevertheless useful, bound:

- $\|e^{At}\| \leq e^{\|A\|t}$ for all $t \geq 0$.

Block matrices (contd)

We say that a block matrix $A = [A_{ij}]$ is

- block-diagonal ($A = \text{diag}\{A_i\}$) if $v = \mu$ and $A_{ij} = 0$ whenever $i \neq j$
- block lower (upper) triangular if $A_{ij} = 0$ whenever $i > j$ ($i < j$)

Operations on block matrices are performed in terms of their blocks, like

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} + \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} = \begin{bmatrix} A_{11} + B_{11} & A_{12} + B_{12} \\ A_{21} + B_{21} & A_{22} + B_{22} \end{bmatrix},$$

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix} = \begin{bmatrix} A_{11}B_{11} + A_{12}B_{21} & A_{11}B_{12} + A_{12}B_{22} \\ A_{21}B_{11} + A_{22}B_{21} & A_{21}B_{12} + A_{22}B_{22} \end{bmatrix}.$$

Sum, product, inverse of block-diagonal (block-triangular) matrices remain block-diagonal (block-triangular). Hence,

- functions of block-diagonal (block-triangular) matrices remain such.

Also, $\text{rank} \begin{bmatrix} A_{11} & A_{12} \end{bmatrix} \geq \text{rank } A_{1i}$ and $\text{rank} \begin{bmatrix} A_{11} \\ A_{21} \end{bmatrix} \geq \text{rank } A_{i1}$, $i \in \mathbb{Z}_{1..2}$.

Outline

Linear algebra preliminaries (self study)

State space

Structural properties of state-space realizations

Lyapunov stability

Matrix representation

We know that a causal $m \times p$ LSI $G : u \mapsto y$ can be described as

$$\begin{bmatrix} \vdots \\ y[t-2] \\ y[t-1] \\ y[t] \\ y[t+1] \\ \vdots \end{bmatrix} = \underbrace{\begin{bmatrix} \vdots & \vdots & \vdots & \vdots & \vdots \\ \cdots & g[0] & 0 & 0 & 0 & \cdots \\ \cdots & g[1] & g[0] & 0 & 0 & \cdots \\ \cdots & g[2] & g[1] & g[0] & 0 & \cdots \\ \cdots & g[3] & g[2] & g[1] & g[0] & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{bmatrix}}_{\llbracket G \rrbracket \text{ (matrix representation of } G\rrbracket} \begin{bmatrix} \vdots \\ u[t-2] \\ u[t-1] \\ u[t] \\ u[t+1] \\ \vdots \end{bmatrix}.$$

The response y starting from a given time instance t can be presented as

$$\underbrace{\begin{bmatrix} y[t] \\ y[t+1] \\ y[t+2] \end{bmatrix}}_{\text{past}} = \underbrace{\begin{bmatrix} g[1] & g[2] & g[3] \\ g[2] & g[3] & g[4] \\ g[3] & g[4] & g[5] \end{bmatrix}}_{\text{past}} \underbrace{\begin{bmatrix} u[t-1] \\ u[t-2] \\ u[t-3] \end{bmatrix}}_{\text{past}} + \underbrace{f(u[t], u[t+1], \dots)}_{\text{present+future}}.$$

Matrix representation

We know that a causal $m \times p$ LSI $G : u \mapsto y$ can be described as

$$\begin{bmatrix} \vdots \\ y[t-2] \\ y[t-1] \\ y[t] \\ y[t+1] \\ \vdots \end{bmatrix} = \underbrace{\begin{bmatrix} \vdots & \vdots & \vdots & \vdots & \vdots \\ \cdots & g[0] & 0 & 0 & 0 & \cdots \\ \cdots & g[1] & g[0] & 0 & 0 & \cdots \\ \cdots & g[2] & g[1] & g[0] & 0 & \cdots \\ \cdots & g[3] & g[2] & g[1] & g[0] & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}}_{\llbracket G \rrbracket \text{ (matrix representation of } G\rrbracket} \begin{bmatrix} \vdots \\ u[t-2] \\ u[t-1] \\ u[t] \\ u[t+1] \\ \vdots \end{bmatrix}.$$

The response y starting from a given time instance t can be presented as

$$\begin{bmatrix} y[t] \\ y[t+1] \\ y[t+2] \\ \vdots \end{bmatrix} = \underbrace{\begin{bmatrix} g[1] & g[2] & g[3] & \cdots \\ g[2] & g[3] & g[4] & \cdots \\ g[3] & g[4] & g[5] & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}}_{\text{past}} \underbrace{\begin{bmatrix} u[t-1] \\ u[t-2] \\ u[t-3] \\ \vdots \end{bmatrix}}_{\text{present+future}} + f(u[t], u[t+1], \dots)$$

Hankel operator

The semi-infinite matrix

$$\llbracket H_G \rrbracket = \begin{bmatrix} g[1] & g[2] & g[3] & \cdots \\ g[2] & g[3] & g[4] & \cdots \\ g[3] & g[4] & g[5] & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

is the matrix representation of the

— **Hankel operator** $H_G : u_{\text{past}} \mapsto y_{\text{present}+\text{future}}$

associated with G . The Hankel operator represents the **memory** of G .

Systems having

— finite-rank Hankel operators are said to be finite dimensional
and $n = \text{rank} \llbracket H_G \rrbracket$ is known as the dimension of G .

Hankel operator

The semi-infinite matrix

$$\llbracket H_G \rrbracket = \begin{bmatrix} g[1] & g[2] & g[3] & \cdots \\ g[2] & g[3] & g[4] & \cdots \\ g[3] & g[4] & g[5] & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

is the matrix representation of the

- **Hankel operator** $H_G : u_{\text{past}} \mapsto y_{\text{present}+\text{future}}$

associated with G . The Hankel operator represents the **memory** of G .

Systems having

- **finite-rank Hankel** operators are said to be **finite dimensional**

and $n = \text{rank} \llbracket H_G \rrbracket$ is known as the dimension of G .

Hankel operator of G_1 , with $g_1[t] = a^{t-1}b\mathbb{1}[t-1]$

In this case

$$\llbracket H_{G_1} \rrbracket = \begin{bmatrix} b & ab & a^2b & \cdots \\ ab & a^2b & a^3b & \cdots \\ a^2b & a^3b & a^4b & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

and its j th column

$$\begin{bmatrix} a^{j-1}b \\ a^j b \\ a^{j+1}b \\ \vdots \end{bmatrix} = \begin{bmatrix} b \\ ab \\ a^2b \\ \vdots \end{bmatrix} a^{j-1}.$$

Hence, $\llbracket H_{G_1} \rrbracket$ has rank 1 (provided $b \neq 0$, of course) and the system G_1 is 1-dimensional.

Hankel operator of G_2 , with $g_2[t] = (b_0 + b_1 t)\mathbb{1}[t]$

If $b_1 \neq 0$ (otherwise it's the previous case), then

$$\llbracket H_{G_2} \rrbracket = \begin{bmatrix} b_0 + b_1 & b_0 + 2b_1 & b_0 + 3b_1 & \cdots \\ b_0 + 2b_1 & b_0 + 3b_1 & b_0 + 4b_1 & \cdots \\ b_0 + 3b_1 & b_0 + 4b_1 & b_0 + 5b_1 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix},$$

its first two columns are linearly independent, and its j th column, $j \geq 3$,

$$\begin{bmatrix} b_0 + jb_1 \\ b_0 + (j+1)b_1 \\ b_0 + (j+2)b_1 \\ \vdots \end{bmatrix} = \begin{bmatrix} b_0 + b_1 \\ b_0 + 2b_1 \\ b_0 + 3b_1 \\ \vdots \end{bmatrix} (2-j) + \begin{bmatrix} b_0 + 2b_1 \\ b_0 + 3b_1 \\ b_0 + 4b_1 \\ \vdots \end{bmatrix} (j-1).$$

Hence, $\text{rank} \llbracket H_{G_2} \rrbracket = 2$ and the system G_2 is 2-dimensional.

Hankel operator of G_3 , with $g_3[t] = \delta[t - 3]$

In this case

$$\llbracket H_{G_3} \rrbracket = \begin{bmatrix} 0 & 0 & 1 & 0 & \cdots \\ 0 & 1 & 0 & 0 & \cdots \\ 1 & 0 & 0 & 0 & \cdots \\ 0 & 0 & 0 & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

has rank 3. Hence, the system G_3 is 3-dimensional.

Hankel operator of G_4 , with $g_4[t] = \frac{1}{t} \mathbb{1}[t-1]$

In this case

$$\llbracket H_{G_4} \rrbracket = \begin{bmatrix} 1 & 1/2 & 1/3 & 1/4 & \cdots \\ 1/2 & 1/3 & 1/4 & 1/5 & \cdots \\ 1/3 & 1/4 & 1/5 & 1/6 & \cdots \\ 1/4 & 1/5 & 1/6 & 1/7 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

is an infinite-rank matrix (i.e. all its rows are linearly independent). Hence, the system G_4 is infinite dimensional.

State vector

If the rank of (infinite-dimensional) H_G is finite, say n , factorize

$$\llbracket H_G \rrbracket = \begin{bmatrix} O_1 \\ O_2 \\ O_3 \\ \vdots \end{bmatrix} \begin{bmatrix} R_1 & R_2 & R_3 & \cdots \end{bmatrix} =: \mathbf{O}_G \mathbf{R}_G$$

($O_i \in \mathbb{R}^{p \times n}$ and $R_i \in \mathbb{R}^{n \times m}$ for each $i \in \mathbb{N}$), where $\mathbf{O}_G$ and $\mathbf{R}_G$ are rank- n matrices having n columns and n rows, respectively. The n -dimensional

$$x[t] := \mathbf{R}_G \begin{bmatrix} u[t-1] \\ u[t-2] \\ u[t-3] \\ \vdots \end{bmatrix} = \sum_{s=-\infty}^{t-1} R_{t-s} u[s]$$

is called the **state vector** of G at t .

State vector (contd)

The state vector is the **history accumulator**. Indeed,

$$\begin{bmatrix} y[t] \\ y[t+1] \\ y[t+2] \\ \vdots \end{bmatrix} = \begin{bmatrix} O_1 \\ O_2 \\ O_3 \\ \vdots \end{bmatrix} x[t] + \begin{bmatrix} g[0] & 0 & 0 & \cdots \\ g[1] & g[0] & 0 & \cdots \\ g[2] & g[1] & g[0] & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} \begin{bmatrix} u[t] \\ u[t+1] \\ u[t+2] \\ \vdots \end{bmatrix},$$

meaning that no knowledge of $u[s]$ for $s < t$ is required to calculate $y[s]$ for $s \geq t$ if $x[t]$ is known. This also suggests that

- causal systems may be viewed as mappings $(x[0], u) \mapsto y$ for u and y on $\mathbb{Z}_+$, rather than $u \mapsto y$ for u and y on $\mathbb{Z}$, with the initial condition $x[0]$ explaining the effect of prehistoric inputs (having a starting point may be intuitive in control applications, this is the time moment where the control law is applied). Causal systems with zero initial conditions may be viewed as operators $u \mapsto y$ for u and y on $\mathbb{Z}_+$.

State vector (contd)

The state vector is the **history accumulator**. Indeed,

$$\begin{bmatrix} y[t] \\ y[t+1] \\ y[t+2] \\ \vdots \end{bmatrix} = \begin{bmatrix} O_1 \\ O_2 \\ O_3 \\ \vdots \end{bmatrix} x[t] + \begin{bmatrix} g[0] & 0 & 0 & \cdots \\ g[1] & g[0] & 0 & \cdots \\ g[2] & g[1] & g[0] & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} \begin{bmatrix} u[t] \\ u[t+1] \\ u[t+2] \\ \vdots \end{bmatrix},$$

meaning that no knowledge of $u[s]$ for $s < t$ is required to calculate $y[s]$ for $s \geq t$ if $x[t]$ is known. This also suggests that

- causal systems may be viewed as mappings $(x[0], u) \mapsto y$

for u and y on $\mathbb{Z}_+$, rather than $u \mapsto y$ for u and y on $\mathbb{Z}$, with the

- initial condition $x[0]$ explaining the effect of prehistoric inputs

(having a starting point may be intuitive in control applications, this is the time moment where the control law is applied). Causal systems with zero initial conditions may be viewed as operators $u \mapsto y$ for u and y on $\mathbb{Z}_+$.

Structure of O_G and R_G

Rank- n and the Hankel structure of H_G guarantee the existence of matrices $A \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{n \times m}$, and $C \in \mathbb{R}^{p \times n}$ such that

$$O_G = \begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \end{bmatrix} \quad \text{and} \quad R_G = [B \quad AB \quad A^2B \quad \cdots],$$

so that $g[t] = CA^{t-1}B$ for all $t \geq 1$,

$$\llbracket H_G \rrbracket = \begin{bmatrix} CB & CAB & CA^2B & \cdots \\ CAB & CA^2B & CA^3B & \cdots \\ CA^2B & CA^3B & CA^4B & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} = \begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \end{bmatrix} [B \quad AB \quad A^2B \quad \cdots],$$

and

$$x[t] = \sum_{s=-\infty}^{t-1} A^{t-1-s} Bu[s] = A^t x[0] + \sum_{s=0}^{t-1} A^{t-1-s} Bu[s].$$

Structure of O_G and R_G for $g_1[t] = a^{t-1}b\mathbb{1}[t-1]$

In this case

$$\llbracket H_{G_1} \rrbracket = \begin{bmatrix} b & ab & a^2b & \cdots \\ ab & a^2b & a^3b & \cdots \\ a^2b & a^3b & a^4b & \cdots \\ \vdots & \vdots & \vdots & \end{bmatrix} = \begin{bmatrix} 1 \\ a \\ a^2 \\ \vdots \end{bmatrix} \begin{bmatrix} b & ab & a^2b & \cdots \end{bmatrix}$$

Obviously, $A = a$, $B = b$, and $C = 1$ do the trick.

Structure of O_G and R_G for $g_2[t] = (b_0 + b_1 t)\mathbb{1}[t]$

In this case (again, $b_1 \neq 0$ is assumed)

$$\begin{aligned}\mathbb{H}_{G_2} &= \begin{bmatrix} b_0 + b_1 & b_0 + 2b_1 & b_0 + 3b_1 & \cdots \\ b_0 + 2b_1 & b_0 + 3b_1 & b_0 + 4b_1 & \cdots \\ b_0 + 3b_1 & b_0 + 4b_1 & b_0 + 5b_1 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} \\ &= \begin{bmatrix} b_1 & b_0 + b_1 \\ b_1 & b_0 + 2b_1 \\ b_1 & b_0 + 3b_1 \\ \vdots & \vdots \end{bmatrix} \begin{bmatrix} 0 & 1 & 2 & 3 & \cdots \\ 1 & 1 & 1 & 1 & \cdots \end{bmatrix}\end{aligned}$$

and we can choose

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad \text{and} \quad C = [b_1 \quad b_0 + b_1].$$

Structure of O_G and R_G for $g_3[t] = \delta[t - 3]$

In this case

$$\llbracket H_{G_3} \rrbracket = \begin{bmatrix} 0 & 0 & 1 & 0 & \cdots \\ 0 & 1 & 0 & 0 & \cdots \\ 1 & 0 & 0 & 0 & \cdots \\ 0 & 0 & 0 & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \\ \vdots & \vdots & \vdots \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 & \cdots \\ 0 & 1 & 0 & 0 & \cdots \\ 0 & 0 & 1 & 0 & \cdots \end{bmatrix}$$

and we can choose

$$A = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad \text{and} \quad C = [0 \ 0 \ 1].$$

Consequences of structure of O_G and R_G

1. The state vector can be propagated **recursively** (convenient):

$$\begin{aligned}
 x[t+1] &= \sum_{s=-\infty}^t A^{t-s} Bu[s] = \sum_{s=-\infty}^{t-1} \overbrace{A^{t-s}}^{\geq 1} Bu[s] + Bu[t] \\
 &= A \sum_{s=-\infty}^{t-1} A^{t-1-s} Bu[s] + Bu[t] \\
 &= Ax[t] + Bu[t].
 \end{aligned}$$

2. The output at each t is a linear function of current state and input:

$$\begin{aligned}
 y[t] &= \sum_{s=-\infty}^t g[t-s]u[s] = C \sum_{s=-\infty}^{t-1} A^{t-1-s} Bu[s] + g[0]u[t] \\
 &= Cx[t] + Du[t],
 \end{aligned}$$

where $D := g[0]$ is the feedthrough term.

Consequences of structure of O_G and R_G

1. The state vector can be propagated **recursively** (convenient):

$$\begin{aligned}x[t+1] &= \sum_{s=-\infty}^t A^{t-s} Bu[s] = \sum_{s=-\infty}^{t-1} \overbrace{A^{t-s}}^{\geq 1} Bu[s] + Bu[t] \\&= A \sum_{s=-\infty}^{t-1} A^{t-1-s} Bu[s] + Bu[t] \\&= Ax[t] + Bu[t].\end{aligned}$$

2. The output at each t is a linear function of current state and input:

$$\begin{aligned}y[t] &= \sum_{s=-\infty}^t g[t-s]u[s] = C \sum_{s=-\infty}^{t-1} A^{t-1-s} Bu[s] + g[0]u[t] \\&= Cx[t] + Du[t],\end{aligned}$$

where $D := g[0]$ is the **feedthrough** term.

Realizations

Thus, we end up with the following description (with $D := g[0]$):

$$G : \begin{cases} x[t+1] = Ax[t] + Bu[t], & x[0] = x_0 \\ y[t] = Cx[t] + Du[t] \end{cases}$$

called the **state representation** of G and the quadruple (A, B, C, D) is called the **state-space realization** of G .

This is a compact form of

$$\begin{bmatrix} y[0] \\ y[1] \\ y[2] \\ \vdots \end{bmatrix} = \begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \end{bmatrix} x_0 + \begin{bmatrix} D & 0 & 0 & \cdots \\ CB & D & 0 & \cdots \\ CAB & CB & D & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} \begin{bmatrix} u[0] \\ u[1] \\ u[2] \\ \vdots \end{bmatrix}$$

or

$$y[t] = CA^t x_0 + C \sum_{s=0}^{t-1} A^{t-1-s} Bu[s] + Du[t].$$

Similar realizations

The rank decomposition $\llbracket H_G \rrbracket = \mathbf{O}_G \mathbf{R}_G$ is not unique. All rank- n factors of $\llbracket H_G \rrbracket$ are $\mathbf{O}_G T^{-1}$ and $T \mathbf{R}_G$ for a nonsingular $T \in \mathbb{R}^{n \times n}$, viz.

$$\mathbf{O}_G T^{-1} = \begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \end{bmatrix} T^{-1} = \begin{bmatrix} CT^{-1} \\ CT^{-1}(TAT^{-1}) \\ CT^{-1}(TAT^{-1})^2 \\ \vdots \end{bmatrix}$$

and

$$T \mathbf{R}_G = T \begin{bmatrix} B & AB & A^2B & \cdots \end{bmatrix} = \begin{bmatrix} TB & (TAT^{-1})TB & (TAT^{-1})^2TB & \cdots \end{bmatrix}$$

In other words, given an n -dimensional system G ,

→ all its n -dimensional realizations are $(TAT^{-1}, TB, CT^{-1}, D)$

for $\det T \neq 0$, where (A, B, C, D) is any particular n -dimensional realization.

The realization $(TAT^{-1}, TB, CT^{-1}, D)$ is said to be similar to (A, B, C, D) and its state vector is $Tx[t]$ for all t . The impulse response is not affected by T , as $CT^{-1}B = CT^{-1}(TAT^{-1})^tTB$ for all $t \in \mathbb{N}$.

Similar realizations

The rank decomposition $\llbracket H_G \rrbracket = \mathbf{O}_G \mathbf{R}_G$ is not unique. All rank- n factors of $\llbracket H_G \rrbracket$ are $\mathbf{O}_G T^{-1}$ and $T \mathbf{R}_G$ for a nonsingular $T \in \mathbb{R}^{n \times n}$, viz.

$$\mathbf{O}_G T^{-1} = \begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \end{bmatrix} T^{-1} = \begin{bmatrix} CT^{-1} \\ CT^{-1}(TAT^{-1}) \\ CT^{-1}(TAT^{-1})^2 \\ \vdots \end{bmatrix}$$

and

$$T \mathbf{R}_G = T \begin{bmatrix} B & AB & A^2B & \cdots \end{bmatrix} = \begin{bmatrix} TB & (TAT^{-1})TB & (TAT^{-1})^2TB & \cdots \end{bmatrix}$$

In other words, given an n -dimensional system G ,

- all its n -dimensional realizations are $(TAT^{-1}, TB, CT^{-1}, D)$

for $\det T \neq 0$, where (A, B, C, D) is any particular n -dimensional realization. The realization $(TAT^{-1}, TB, CT^{-1}, D)$ is said to be **similar** to (A, B, C, D) and its state vector is $Tx[t]$ for all t . The impulse response is not affected by T , as $CA^tB = CT^{-1}(TAT^{-1})^tTB$ for all $t \in \mathbb{N}$.

From state space to transfer function

Assume zero initial condition and write

$$\begin{cases} q\mathbf{x}[t] = A\mathbf{x}[t] + B\mathbf{u}[t] \\ y[t] = C\mathbf{x}[t] + D\mathbf{u}[t] \end{cases} \implies \begin{cases} zX(z) = AX(z) + BU(z) \\ Y(z) = CX(z) + DU(z) \end{cases}$$

Hence, the transfer function

$$G(z) = D + C(zI - A)^{-1}B =: \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array} \right].$$

Properties:

- $\left[\begin{array}{c|c} A & B \\ \hline C & D \end{array} \right] = \left[\begin{array}{c|c} TAT^{-1} & TB \\ \hline CT^{-1} & D \end{array} \right]$ for every nonsingular T
- $G(z) = D + C(zI - A)^{-1}B$ is always proper (strictly proper iff $D = 0$)
- if $D = 0$, the pole excess is the smallest $k \in \mathbb{N}$ such that $CA^{k-1}B \neq 0$ (think of the impulse response, with $g[1] = CB$, $g[2] = CAB$, $g[3] = CA^2B$, etc)

System interconnections in state space

If

$$G_i : \begin{cases} x_i[t+1] = A_i x_i[t] + B_i u_i[t], & x_i[0] = x_{0i} \\ y_i[t] = C_i x_i[t] + D_i u_i[t] \end{cases}$$

for $i = 1, 2$, then interconnections of G_1 and G_2 follow the steps below:

1. determine what inputs / outputs to equate and then eliminate
2. unite state vectors x_1 and x_2

Example

The cascade $G_2 G_1 : u \mapsto y$ implies that $u_1 = u$, $u_2 = y_1$, and $y = y_2$, so

$$G_2 G_1 : \begin{cases} \begin{bmatrix} x_2[t+1] \\ x_1[t+1] \end{bmatrix} = \begin{bmatrix} A_2 & 0 \\ 0 & A_1 \end{bmatrix} \begin{bmatrix} x_2[t] \\ x_1[t] \end{bmatrix} + \begin{bmatrix} B_2 \\ 0 \end{bmatrix} u[t] \\ y[t] = \begin{bmatrix} 0 & C_1 \end{bmatrix} \begin{bmatrix} x_2[t] \\ x_1[t] \end{bmatrix} + (D_1) u[t] \end{cases}$$

System interconnections in state space

If

$$G_i : \begin{cases} x_i[t+1] = A_i x_i[t] + B_i u_i[t], & x_i[0] = x_{0i} \\ y_i[t] = C_i x_i[t] + D_i u_i[t] \end{cases}$$

for $i = 1, 2$, then interconnections of G_1 and G_2 follow the steps below:

1. determine what inputs / outputs to equate and then eliminate
2. unite state vectors x_1 and x_2

Example

The cascade $G_2 G_1 : u \mapsto y$ implies that $u_1 = u$, $u_2 = y_1$, and $y = y_2$, so

$$G_2 G_1 : \begin{cases} \begin{bmatrix} x_2[t+1] \\ x_1[t+1] \end{bmatrix} = \begin{bmatrix} & \\ & \end{bmatrix} \begin{bmatrix} x_2[t] \\ x_1[t] \end{bmatrix} + \begin{bmatrix} \\ \end{bmatrix} u[t] \\ y[t] = \begin{bmatrix} & \end{bmatrix} \begin{bmatrix} x_2[t] \\ x_1[t] \end{bmatrix} + (\quad) u[t] \end{cases}$$

System interconnections in state space

If

$$G_i : \begin{cases} x_i[t+1] = A_i x_i[t] + B_i u_i[t], & x_i[0] = x_{0i} \\ y_i[t] = C_i x_i[t] + D_i u_i[t] \end{cases}$$

for $i = 1, 2$, then interconnections of G_1 and G_2 follow the steps below:

1. determine what inputs / outputs to equate and then eliminate
2. unite state vectors x_1 and x_2

Example

The cascade $G_2 G_1 : u \mapsto y$ implies that $u_1 = u$, $u_2 = y_1$, and $y = y_2$, so

$$G_2 G_1 : \begin{cases} \begin{bmatrix} x_2[t+1] \\ x_1[t+1] \end{bmatrix} = \begin{bmatrix} 0 & A_1 \end{bmatrix} \begin{bmatrix} x_2[t] \\ x_1[t] \end{bmatrix} + \begin{bmatrix} B_1 \end{bmatrix} u[t] \\ y[t] = \begin{bmatrix} \end{bmatrix} \begin{bmatrix} x_2[t] \\ x_1[t] \end{bmatrix} + () u[t] \end{cases}$$

System interconnections in state space

If

$$G_i : \begin{cases} x_i[t+1] = A_i x_i[t] + B_i u_i[t], & x_i[0] = x_{0i} \\ y_i[t] = C_i x_i[t] + D_i u_i[t] \end{cases}$$

for $i = 1, 2$, then interconnections of G_1 and G_2 follow the steps below:

1. determine what inputs / outputs to equate and then eliminate
2. unite state vectors x_1 and x_2

Example

The cascade $G_2 G_1 : u \mapsto y$ implies that $u_1 = u$, $u_2 = y_1$, and $y = y_2$, so

$$G_2 G_1 : \begin{cases} \begin{bmatrix} x_2[t+1] \\ x_1[t+1] \end{bmatrix} = \begin{bmatrix} 0 & A_1 \end{bmatrix} \begin{bmatrix} x_2[t] \\ x_1[t] \end{bmatrix} + \begin{bmatrix} B_1 \end{bmatrix} u[t] \\ y[t] = \begin{bmatrix} \end{bmatrix} \begin{bmatrix} x_2[t] \\ x_1[t] \end{bmatrix} + () u[t] \end{cases}$$

$$(u_2[t] = C_1 x_1[t] + D_1 u[t])$$

System interconnections in state space

If

$$G_i : \begin{cases} x_i[t+1] = A_i x_i[t] + B_i u_i[t], & x_i[0] = x_{0i} \\ y_i[t] = C_i x_i[t] + D_i u_i[t] \end{cases}$$

for $i = 1, 2$, then interconnections of G_1 and G_2 follow the steps below:

1. determine what inputs / outputs to equate and then eliminate
2. unite state vectors x_1 and x_2

Example

The cascade $G_2 G_1 : u \mapsto y$ implies that $u_1 = u$, $u_2 = y_1$, and $y = y_2$, so

$$G_2 G_1 : \begin{cases} \begin{bmatrix} x_2[t+1] \\ x_1[t+1] \end{bmatrix} = \begin{bmatrix} A_2 & B_2 C_1 \\ 0 & A_1 \end{bmatrix} \begin{bmatrix} x_2[t] \\ x_1[t] \end{bmatrix} + \begin{bmatrix} B_2 D_1 \\ B_1 \end{bmatrix} u[t] \\ y[t] = \begin{bmatrix} C_2 & C_1 D_1 + D_2 \end{bmatrix} \begin{bmatrix} x_2[t] \\ x_1[t] \end{bmatrix} + (D_2 D_1) u[t] \end{cases}$$

$$(u_2[t] = C_1 x_1[t] + D_1 u[t])$$

System interconnections in state space

If

$$G_i : \begin{cases} x_i[t+1] = A_i x_i[t] + B_i u_i[t], & x_i[0] = x_{0i} \\ y_i[t] = C_i x_i[t] + D_i u_i[t] \end{cases}$$

for $i = 1, 2$, then interconnections of G_1 and G_2 follow the steps below:

1. determine what inputs / outputs to equate and then eliminate
2. unite state vectors x_1 and x_2

Example

The cascade $G_2 G_1 : u \mapsto y$ implies that $u_1 = u$, $u_2 = y_1$, and $y = y_2$, so

$$G_2 G_1 : \begin{cases} \begin{bmatrix} x_2[t+1] \\ x_1[t+1] \end{bmatrix} = \begin{bmatrix} A_2 & B_2 C_1 \\ 0 & A_1 \end{bmatrix} \begin{bmatrix} x_2[t] \\ x_1[t] \end{bmatrix} + \begin{bmatrix} B_2 D_1 \\ B_1 \end{bmatrix} u[t] \\ y[t] = \begin{bmatrix} C_2 & D_2 C_1 \end{bmatrix} \begin{bmatrix} x_2[t] \\ x_1[t] \end{bmatrix} + (D_2 D_1) u[t] \end{cases}$$

($u_2[t] = C_1 x_1[t] + D_1 u[t]$), with corresponding initial conditions.

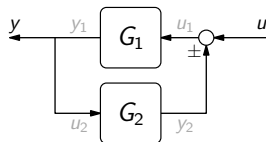


System interconnections in state space (contd)

Example

Consider now the feedback interconnection,

- $u_2 = y = y_1$
- $u_1 = u \pm y_2$



Assuming $D_1 = 0$ for simplicity, the system $G : u \mapsto y$ has the realization

$$G : \begin{cases} \begin{bmatrix} x_1[t+1] \\ x_2[t+1] \end{bmatrix} = \begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix} \begin{bmatrix} x_1[t] \\ x_2[t] \end{bmatrix} + \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} u[t] \\ y[t] = \begin{bmatrix} C_1 & C_2 \end{bmatrix} \begin{bmatrix} x_1[t] \\ x_2[t] \end{bmatrix} + (D_1) u[t] \end{cases}$$

because

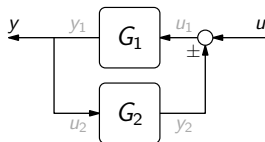
- $u_2[t] = y_1[t] = C_1 x_1[t]$
- $u_1[t] = u[t] \pm C_2 x_2[t] \pm D_2 C_1 x_1[t]$
- $y[t] = C_1 x_1[t]$

System interconnections in state space (contd)

Example

Consider now the feedback interconnection,

- $u_2 = y = y_1$
- $u_1 = u \pm y_2$



Assuming $D_1 = 0$ for simplicity, the system $G : u \mapsto y$ has the realization

$$G : \begin{cases} \begin{bmatrix} x_1[t+1] \\ x_2[t+1] \end{bmatrix} = \begin{bmatrix} A_1 & B_1 \\ A_2 & B_2 \end{bmatrix} \begin{bmatrix} x_1[t] \\ x_2[t] \end{bmatrix} + \begin{bmatrix} D_1 \\ D_2 \end{bmatrix} u[t] \\ y[t] = \begin{bmatrix} C_1 & C_2 \end{bmatrix} \begin{bmatrix} x_1[t] \\ x_2[t] \end{bmatrix} + () u[t] \end{cases}$$

because

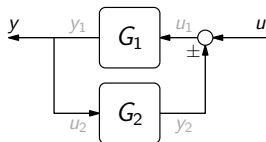
- $u_2[t] = y_1[t] = G_1 x_1[t]$
- $u_1[t] = u[t] \pm y_2[t] = u[t] \pm G_2 x_2[t] \pm D_2 G_1 x_1[t]$
- $y[t] = G_1 x_1[t]$

System interconnections in state space (contd)

Example

Consider now the feedback interconnection,

- $u_2 = y = y_1$
- $u_1 = u \pm y_2$



Assuming $D_1 = 0$ for simplicity, the system $G : u \mapsto y$ has the realization

$$G : \begin{cases} \begin{bmatrix} x_1[t+1] \\ x_2[t+1] \end{bmatrix} = \begin{bmatrix} A_1 & B_1 C_1 \\ B_2 C_1 & A_2 \end{bmatrix} \begin{bmatrix} x_1[t] \\ x_2[t] \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \end{bmatrix} u[t] \\ y[t] = \begin{bmatrix} C_1 & 0 \end{bmatrix} \begin{bmatrix} x_1[t] \\ x_2[t] \end{bmatrix} + (D_1) u[t] \end{cases}$$

because

- $u_2[t] = y_1[t] = C_1 x_1[t]$

$$u_1[t] = u[t] \pm y_2[t] = u[t] \pm G_2 x_2[t] \pm D_2 G_1 x_1[t]$$

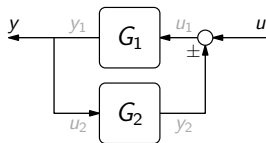
$$y[t] = G_1 x_1[t]$$

System interconnections in state space (contd)

Example

Consider now the feedback interconnection,

- $u_2 = y = y_1$
- $u_1 = u \pm y_2$



Assuming $D_1 = 0$ for simplicity, the system $G : u \mapsto y$ has the realization

$$G : \begin{cases} \begin{bmatrix} x_1[t+1] \\ x_2[t+1] \end{bmatrix} = \begin{bmatrix} A_1 \pm B_1 D_2 C_1 & \pm B_1 C_2 \\ B_2 C_1 & A_2 \end{bmatrix} \begin{bmatrix} x_1[t] \\ x_2[t] \end{bmatrix} + \begin{bmatrix} B_1 \\ 0 \end{bmatrix} u[t] \\ y[t] = \begin{bmatrix} \quad \quad \end{bmatrix} \begin{bmatrix} x_1[t] \\ x_2[t] \end{bmatrix} + (\quad) u[t] \end{cases}$$

because

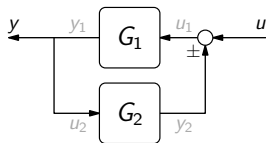
- $u_2[t] = y_1[t] = C_1 x_1[t]$
- $u_1[t] = u[t] \pm C_2 x_2[t] \pm D_2 u_2[t] = u[t] \pm C_2 x_2[t] \pm D_2 C_1 x_1[t]$
- $y[t] = C_1 x_1[t]$

System interconnections in state space (contd)

Example

Consider now the feedback interconnection,

- $u_2 = y = y_1$
- $u_1 = u \pm y_2$



Assuming $D_1 = 0$ for simplicity, the system $G : u \mapsto y$ has the realization

$$G : \begin{cases} \begin{bmatrix} x_1[t+1] \\ x_2[t+1] \end{bmatrix} = \begin{bmatrix} A_1 \pm B_1 D_2 C_1 & \pm B_1 C_2 \\ B_2 C_1 & A_2 \end{bmatrix} \begin{bmatrix} x_1[t] \\ x_2[t] \end{bmatrix} + \begin{bmatrix} B_1 \\ 0 \end{bmatrix} u[t] \\ y[t] = \begin{bmatrix} C_1 & 0 \end{bmatrix} \begin{bmatrix} x_1[t] \\ x_2[t] \end{bmatrix} + (0) u[t] \end{cases}$$

because

- $u_2[t] = y_1[t] = C_1 x_1[t]$
- $u_1[t] = u[t] \pm C_2 x_2[t] \pm D_2 u_2[t] = u[t] \pm C_2 x_2[t] \pm D_2 C_1 x_1[t]$
- $y[t] = C_1 x_1[t]$



Outline

Linear algebra preliminaries (self study)

State space

Structural properties of state-space realizations

Lyapunov stability

Reachability and controllability

Consider the state equation

$$x[t+1] = Ax[t] + Bu[t], \quad x[0] = x_0$$

for $A \in \mathbb{R}^{n \times n}$ and $B \in \mathbb{R}^{n \times m}$.

Definition (reachability from the origin)

A state x_f is said to be **reachable** if, given $x_0 = 0$, there exist a time $t_f \in \mathbb{N}$ and a control signal $u[t]$, $t \in \mathbb{Z}_{0..t_f-1}$, such that $x[t_f] = x_f$. If every state is reachable, we say that the system or (A, B) is (completely) reachable.

Definition (controllability to the origin)

An initial state x_0 is said to be **controllable** if there exist a time $t_f \in \mathbb{N}$ and a control signal $u[t]$, $t \in \mathbb{Z}_{0..t_f-1}$, such that $x[t_f] = 0$. If every initial state is controllable, we say that the system or (A, B) is (completely) controllable.

Reachability: the upper bound on t_f

Lemma

If $x_f \in \mathbb{R}^n$ is reachable in $t_f > n$ steps, then it is reachable in n steps.

Proof: By Cayley–Hamilton, $A^{n+j} = \sum_{i=0}^{n-1} \alpha_{j,i} A^{n-1-i}$ for some $\alpha_{j,i}$. Thus,

$$\begin{aligned}
 x[n+k] &= \sum_{i=0}^{n+k-1} A^{n+k-1-i} Bu[i] = \sum_{i=k}^{n+k-1} \overbrace{A^{n+k-1-i}}^{\in \mathbb{Z}_{0,n-1}} Bu[i] + \sum_{j=0}^{k-1} \overbrace{A^{n+k-1-j}}^{\in \mathbb{Z}_{n..n+k-1}} Bu[j] \\
 &= \sum_{i=0}^{n-1} A^{n-1-i} Bu[i+k] + \sum_{j=0}^{k-1} \sum_{i=0}^{n-1} \alpha_{k-1-j,i} A^{n-1-i} Bu[j] \\
 &= \sum_{i=0}^{n-1} A^{n-1-i} B \left(\underbrace{u[i+k] + \sum_{j=0}^{k-1} \alpha_{k-1-j,i} u[j]}_{\tilde{u}[i]} \right)
 \end{aligned}$$

So if $x[n+k] = x_f$, then the very same x_f is reachable in n steps by $\tilde{u}$. $\square$

Reachability: the upper bound on t_f

Lemma

If $x_f \in \mathbb{R}^n$ is reachable in $t_f > n$ steps, then it is reachable in n steps.

Proof: By Cayley–Hamilton, $A^{n+j} = \sum_{i=0}^{n-1} \alpha_{j,i} A^{n-1-i}$ for some $\alpha_{j,i}$. Thus,

$$\begin{aligned}
 x[n+k] &= \sum_{i=0}^{n+k-1} A^{n+k-1-i} B u[i] = \sum_{i=k}^{n+k-1} \overbrace{A^{n+k-1-i}}^{\in \mathbb{Z}_{0,n-1}} B u[i] + \sum_{j=0}^{k-1} \overbrace{A^{n+k-1-j}}^{\in \mathbb{Z}_{n..n+k-1}} B u[j] \\
 &= \sum_{i=0}^{n-1} A^{n-1-i} B u[i+k] + \sum_{j=0}^{k-1} \sum_{i=0}^{n-1} \alpha_{k-1-j,i} A^{n-1-i} B u[j] \\
 &= \sum_{i=0}^{n-1} A^{n-1-i} B \left(\underbrace{u[i+k] + \sum_{j=0}^{k-1} \alpha_{k-1-j,i} u[j]}_{\tilde{u}[i]} \right) = \sum_{i=0}^{n-1} A^{n-1-i} B \tilde{u}[i]
 \end{aligned}$$

So if $x[n+k] = x_f$, then the very same x_f is reachable in n steps by $\tilde{u}$. □

Reachability and the controllability matrix

The matrix

$$M_c := \begin{bmatrix} B & AB & \cdots & A^{n-1}B \end{bmatrix} \in \mathbb{R}^{n \times mn}$$

is known as the **controllability matrix** associated with (A, B) .

Theorem

(A, B) is reachable iff $\text{rank } M_c = n$ (i.e. M_c has full row rank).

Proof: We know that

$$x[n] = \sum_{i=0}^{n-1} A^{n-1-i} B u[i] = \overbrace{\begin{bmatrix} B & AB & \cdots & A^{n-1}B \end{bmatrix}}^{M_c} \overbrace{\begin{bmatrix} u[n-1] \\ \vdots \\ u[1] \\ u[0] \end{bmatrix}}^{U_n}.$$

if: $\text{rank } M_c = n \implies x_f = M_c U_n$ is solvable in U_n for all $x_f \in \mathbb{R}^n$.

only if: $\text{rank } M_c < n \implies \exists \eta \neq 0$ such that $\eta' M_c = 0 \implies \eta' x_f = 0$ for all $x_f = M_c U_n$ (i.e. reachable) $\implies x_f = \eta$ is not reachable.

The proof follows as reachability in n steps is equivalent to reachability. $\square$

Reachability from an arbitrary initial condition

Lemma

If (A, B) is reachable, then every x_f is reachable from all initial conditions x_0 .

Proof: Let x_f and x_0 be arbitrary and define $x_1 := x_f - A^n x_0 \in \mathbb{R}^n$. Because (A, B) is reachable, there is $u[t] = u_1[t]$, $t \in \mathbb{Z}_{0..n-1}$, such that

$$x_1 = x_f - A^n x_0 = \sum_{i=0}^{n-1} A^{n-1-i} B u_1[i].$$

But this implies that x_f is reachable from x_0 :

$$x_f = A^n x_0 + \sum_{i=0}^{n-1} A^{n-1-i} B u_1[i].$$

The result follows by the arbitrariness of x_f and x_0 . □

Controllability and the controllability matrix

A corollary of the last lemma is that

- reachability $\implies$ controllability.

The other direction, reachability $\Leftarrow$ controllability, is *not necessarily true*.

Example

Let

$$[A|B] = \left[\begin{array}{cc|c} 1 & 1 & 1 \\ 0 & 0 & 0 \end{array} \right] \implies \text{rank } M_c = \text{rank} \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} = 1,$$

meaning this (A, B) is not reachable. But the dynamics

$$\begin{cases} x_1[t+1] = x_1[t] + x_2[t] + u[t], & x_1[0] = x_{0,1} \\ x_2[t+1] = 0, & x_2[0] = x_{0,2} \end{cases}$$

can be controlled to the origin in one step, by $u[0] = -x_{01} - x_{02}$. Hence, in this case (A, B) is controllable. ∇

Controllability and the controllability matrix

A corollary of the last lemma is that

- reachability $\implies$ controllability.

The other direction, reachability $\Longleftarrow$ controllability, is *not necessarily* true:

Example

Let

$$\left[A \mid B \right] = \left[\begin{array}{cc|c} 1 & 1 & 1 \\ 0 & 0 & 0 \end{array} \right] \implies \text{rank } M_c = \text{rank} \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} = 1,$$

meaning this (A, B) is not reachable. But the dynamics

$$\begin{cases} x_1[t+1] = x_1[t] + x_2[t] + u[t], & x_1[0] = x_{0,1} \\ x_2[t+1] = 0, & x_2[0] = x_{0,2} \end{cases}$$

can be controlled to the origin in one step, by $u[0] = -x_{0,1} - x_{0,2}$. Hence, in this case (A, B) is controllable. 

Controllability and the controllability matrix (contd)

Theorem

(A, B) is controllable iff $\text{rank } M_c = \text{rank} \begin{bmatrix} A^n & M_c \end{bmatrix}$.

Proof: An initial condition $x_0 \neq 0$ is controllable iff

$$0 = A^n x_0 + \sum_{i=0}^{n-1} A^{n-1-i} B u[i] \iff M_c U_n = -A^n x_0$$

is solvable in $U_n \iff \text{rank } M_c = \text{rank} \begin{bmatrix} -A^n x_0 & M_c \end{bmatrix}$. Because

$$\begin{bmatrix} -A^n x_0 & M_c \end{bmatrix} = \begin{bmatrix} A^n & M_c \end{bmatrix} \begin{bmatrix} -x_0 & 0 \\ 0 & I_{mn} \end{bmatrix} \quad \text{and} \quad \text{rank} \begin{bmatrix} -x_0 & 0 \\ 0 & I_{mn} \end{bmatrix} = mn + 1,$$

$\text{rank } M_c \leq \text{rank} \begin{bmatrix} -A^n x_0 & M_c \end{bmatrix} \leq \text{rank} \begin{bmatrix} A^n & M_c \end{bmatrix}$ holds for all x_0 , whence *sufficiency* follows. If $\text{rank } M_c < \text{rank} \begin{bmatrix} A^n & M_c \end{bmatrix}$, then there is $\eta \neq 0$ such that $\text{rank } M_c < \text{rank} \begin{bmatrix} A^n \eta & M_c \end{bmatrix}$. Hence, $x_0 = -\eta$ is not controllable and we have *necessity*. □

Controllability and the controllability matrix (contd)

Example

Return to $[A \mid B] = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}$. In this case

$$M_c = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} \quad \text{and} \quad [A^n \mid M_c] = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad \left(\text{as } A^2 = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} \right)$$

and

$$\text{rank} \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} = \text{rank} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} = 1,$$

as expected. 

If $\det A \neq 0$, then $\text{rank} [A^n \mid M_c] = n$ irrespective of M_c , so that

$$\text{rank } M_c = \text{rank} [A^n \mid M_c] \iff \text{rank } M_c = n$$

and reachability $\iff$ controllability.

Controllability and the controllability matrix (contd)

Example

Return to $[A \mid B] = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}$. In this case

$$M_c = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} \quad \text{and} \quad [A^n \mid M_c] = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad \left(\text{as } A^2 = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} \right)$$

and

$$\text{rank} \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} = \text{rank} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} = 1,$$

as expected. 

If $\det A \neq 0$, then $\text{rank} [A^n \mid M_c] = n$ irrespective of M_c , so that

$$\text{rank } M_c = \text{rank} [A^n \mid M_c] \iff \text{rank } M_c = n$$

and **reachability** $\iff$ **controllability**.

Other reachability and controllability criteria

Theorem

(A, B) is reachable iff either of the conditions below holds:

1. $\text{rank } M_c = n$
2. the controllability Gramian $W_c[k] := \sum_{i=0}^{k-1} A^i B B' (A^i)' > 0, \forall k \geq n$
3. $\text{rank} \begin{bmatrix} A - \lambda I & B \end{bmatrix} = n$ for all $\lambda \in \mathbb{C}$ (PBH: Popov–Belevitch–Hautus)
4. $\text{spec}(A + BF)$ can be assigned arbitrarily by a choice of F

Remark: It is readily seen that $W_c[n] = M_c M_c'$. If A is Schur, then $W_c[\infty]$ exists and satisfies the Stein (discrete Lyapunov) equation $W_c[\infty] - A W_c[\infty] A' = B B'$. Moreover, in this case $W_c[\infty] = R_G R_G'$.

Theorem

(A, B) is controllable iff either of the conditions below holds:

1. $\text{rank } M_c = \text{rank} \begin{bmatrix} A^n & M_c \end{bmatrix}$
2. $\text{rank} \begin{bmatrix} A - \lambda I & B \end{bmatrix} = n$ for all $\lambda \in \mathbb{C} \setminus \{0\}$

Other reachability and controllability criteria

Theorem

(A, B) is reachable iff either of the conditions below holds:

1. $\text{rank } M_c = n$
2. the controllability Gramian $W_c[k] := \sum_{i=0}^{k-1} A^i B B' (A^i)' > 0, \forall k \geq n$
3. $\text{rank} \begin{bmatrix} A - \lambda I & B \end{bmatrix} = n$ for all $\lambda \in \mathbb{C}$ (PBH: Popov–Belevitch–Hautus)
4. $\text{spec}(A + BF)$ can be assigned arbitrarily by a choice of F

Remark: It is readily seen that $W_c[n] = M_c M_c'$. If A is Schur, then $W_c[\infty]$ exists and satisfies the Stein (discrete Lyapunov) equation $W_c[\infty] - A W_c[\infty] A' = B B'$. Moreover, in this case $W_c[\infty] = R_G R_G'$.

Theorem

(A, B) is controllable iff either of the conditions below holds:

1. $\text{rank } M_c = \text{rank} \begin{bmatrix} A^n & M_c \end{bmatrix}$
2. $\text{rank} \begin{bmatrix} A - \lambda I & B \end{bmatrix} = n$ for all $\lambda \in \mathbb{C} \setminus \{0\}$

More about PBH

If $\text{rank} \begin{bmatrix} A - \lambda I & B \end{bmatrix} < n$ for some λ , then there is $\eta \neq 0$ such that

$$\eta' \begin{bmatrix} A - \lambda I & B \end{bmatrix} = 0 \quad \Longleftrightarrow \quad (\eta' A = \lambda \eta') \wedge (\eta' B = 0)$$

Hence, PBH might fail only if $\lambda \in \text{spec}(A)$. Moreover, in this case

$$\eta'(A + BF) = \lambda \eta',$$

meaning that $\lambda \in \text{spec}(A + BF)$ for all F . As such,

eigenvalues of A at which PBH fails are called **unreachable modes** of (A, B) .

More about PBH

If $\text{rank} \begin{bmatrix} A - \lambda I & B \end{bmatrix} < n$ for some λ , then there is $\eta \neq 0$ such that

$$\eta' \begin{bmatrix} A - \lambda I & B \end{bmatrix} = 0 \quad \Longleftrightarrow \quad (\eta' A = \lambda \eta') \wedge (\eta' B = 0)$$

Hence, PBH might fail only if $\lambda \in \text{spec}(A)$. Moreover, in this case

$$\eta'(A + BF) = \lambda \eta',$$

meaning that $\lambda \in \text{spec}(A + BF)$ for all F . As such,

— eigenvalues of A at which PBH fails are called **unreachable modes** of (A, B) .

Observability

Consider the state equation with zero input

$$\begin{cases} x[t+1] = Ax[t], & x[0] = x_0 \\ y[t] = Cx[t] \end{cases}$$

for $A \in \mathbb{R}^{n \times n}$ and $C \in \mathbb{R}^{p \times n}$.

Definition (observability)

An initial state x_0 is said to be **observable** if there exists a time $t_f \in \mathbb{N}$ such that x_0 can be reconstructed from measurements $y[t]$, $t \in \mathbb{Z}_{0..t_f-1}$. If every initial state is observable, we say that the system or (C, A) is observable.

Observability: the upper bound on t_f

Lemma

If $x_0 \in \mathbb{R}^n$ is observable in $t_f > n$ steps, then it is observable in n steps.

Proof: By Cayley–Hamilton, $A^{n+k} = \sum_{i=0}^{n-1} \beta_{k,i} A^i$ for some $\beta_{k,i}$. Thus,

$$y[n+k] = CA^{n+k}x_0 = C \sum_{i=0}^{n-1} \beta_{k,i} A^i x_0 = \sum_{i=0}^{n-1} \beta_{k,i} CA^i x_0 = \sum_{i=0}^{n-1} \beta_{k,i} y[i]$$

for every x_0 . Hence, measurements $y[t]$ for $t \geq n$ are linear combinations of measurements $y[t]$ for $t \in \mathbb{Z}_{0..n-1}$ and contain no new information. $\square$

Observability and the observability matrix

The matrix

$$M_o := \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{bmatrix} \in \mathbb{R}^{pn \times n}$$

is known as the **observability matrix** associated with (C, A) .

Theorem

(C, A) is observable iff $\text{rank } M_o = n$ (i.e. M_o has full column rank).

Proof: Let $Y_h := [y'[0] \ y'[1] \ \cdots \ y'[n-1]]'$. The first n measurements satisfy $Y_h = M_o x_0$. We have then:

if: $\text{rank } M_o = n \implies Y_h = M_o x_0 \in \mathbb{R}^n$ has a unique x_0 satisfying it.

only if: $\text{rank } M_o < n \implies \exists \eta \neq 0$ such that $M_o \eta = 0 \implies y \equiv 0$ under both $x_0 = 0$ and $x_0 = \eta$ and we cannot tell whether measurements are caused by η or by 0. Hence, $x_0 = \eta$ is not observable. □

Other observability criteria

Theorem

(C, A) is observable iff either of the conditions below holds:

1. $\text{rank } M_o = n$
2. the observability Gramian $W_o[k] := \sum_{i=0}^{k-1} (A^i)' C' C A^i > 0, \forall k \geq n$
3. $\text{rank} \begin{bmatrix} A - \lambda I \\ C \end{bmatrix} = n$ for all $\lambda \in \mathbb{C}$ (PBH: Popov–Belevitch–Hautus)
4. $\text{spec}(A + LC)$ can be assigned arbitrarily by a choice of L
5. (A', C') is reachable

Remark: It is readily seen that $W_o[n] = M_o' M_o$. If A is Schur, then $W_o[\infty]$ exists and satisfies the Stein (discrete Lyapunov) equation $W_o[\infty] - A' W_o[\infty] A = C' C$. Moreover, in this case $W_o[\infty] = O_G' O_G$.

More about PBH

If $\text{rank} \begin{bmatrix} A - \lambda I \\ C \end{bmatrix} < n$ for some λ , then there is $\eta \neq 0$ such that

$$\begin{bmatrix} A - \lambda I \\ C \end{bmatrix} \eta = 0 \quad \Longleftrightarrow \quad (A\eta = \lambda\eta) \wedge (C\eta = 0)$$

Hence, PBH might fail only if $\lambda \in \text{spec}(A)$. Moreover, in this case

$$(A + LC)\eta = \lambda\eta,$$

meaning that $\lambda \in \text{spec}(A + LC)$ for all L . As such,

λ – eigenvalues of A at which PBH fails are called **unobservable modes** of (C, A) .

More about PBH

If $\text{rank} \begin{bmatrix} A - \lambda I \\ C \end{bmatrix} < n$ for some λ , then there is $\eta \neq 0$ such that

$$\begin{bmatrix} A - \lambda I \\ C \end{bmatrix} \eta = 0 \quad \Longleftrightarrow \quad (A\eta = \lambda\eta) \wedge (C\eta = 0)$$

Hence, PBH might fail only if $\lambda \in \text{spec}(A)$. Moreover, in this case

$$(A + LC)\eta = \lambda\eta,$$

meaning that $\lambda \in \text{spec}(A + LC)$ for all L . As such,

— eigenvalues of A at which PBH fails are called **unobservable modes** of (C, A) .

Stabilizability and detectability

Definition (stabilizability)

A pair (A, B) is said to be stabilizable if there exists F such that $A + BF$ is Schur, i.e. $\text{spec}(A + BF) \in \mathbb{D}_1$ (equivalently, $\rho(A + BF) < 1$).

It can be shown that the

- pair (A, B) is stabilizable iff $\text{rank} \begin{bmatrix} A - \lambda I & B \end{bmatrix} = n$ for all $\lambda \in \mathbb{C} \setminus \mathbb{D}_1$ (no unstable unreachable modes).

Definition (detectability)

A pair (C, A) is said to be detectable if there exists L such that $A + LC$ is Schur, i.e. $\text{spec}(A + LC) \in \mathbb{D}_1$ (equivalently, $\rho(A + LC) < 1$).

It can be shown that the

- pair (C, A) is detectable iff $\text{rank} \begin{bmatrix} A - \lambda I \\ C \end{bmatrix} = n$ for all $\lambda \in \mathbb{C} \setminus \mathbb{D}_1$ (no unstable unobservable modes).

Stabilizability and detectability

Definition (stabilizability)

A pair (A, B) is said to be stabilizable if there exists F such that $A + BF$ is Schur, i.e. $\text{spec}(A + BF) \in \mathbb{D}_1$ (equivalently, $\rho(A + BF) < 1$).

It can be shown that the

- pair (A, B) is stabilizable iff $\text{rank} \begin{bmatrix} A - \lambda I & B \end{bmatrix} = n$ for all $\lambda \in \mathbb{C} \setminus \mathbb{D}_1$ (no unstable unreachable modes).

Definition (detectability)

A pair (C, A) is said to be detectable if there exists L such that $A + LC$ is Schur, i.e. $\text{spec}(A + LC) \in \mathbb{D}_1$ (equivalently, $\rho(A + LC) < 1$).

It can be shown that the

- pair (C, A) is detectable iff $\text{rank} \begin{bmatrix} A - \lambda I \\ C \end{bmatrix} = n$ for all $\lambda \in \mathbb{C} \setminus \mathbb{D}_1$ (no unstable unobservable modes).

*ability and similarity transformations

Similarity transformations do not affect reachability, observability, et cetera.

If $(\tilde{A}, \tilde{B}, \tilde{C}, D) := (TAT^{-1}, TB, CT^{-1}, D)$, then

- $\tilde{M}_c = TM_c$ not similar
- $\tilde{W}_c[t] = TW_c[t]T'$ for all $t \in \mathbb{Z}_+$ not similar
- $\begin{bmatrix} \tilde{A} - \lambda I & \tilde{B} \end{bmatrix} = T \begin{bmatrix} A - \lambda I & B \end{bmatrix} \begin{bmatrix} T^{-1} & 0 \\ 0 & I \end{bmatrix}$ for all $\lambda \in \mathbb{C}$
- $\tilde{A} + \tilde{B}\tilde{F} = T(A + BF)T^{-1}$ if $\tilde{F} = FT^{-1}$ similar
- $\tilde{M}_o = M_o T^{-1}$ not similar
- $\tilde{W}_o[t] = (T^{-1})' W_o[t] T^{-1}$ for all $t \in \mathbb{Z}_+$ not similar
- $\begin{bmatrix} \tilde{A} - \lambda I \\ \tilde{C} \end{bmatrix} = \begin{bmatrix} T & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} A - \lambda I \\ C \end{bmatrix} T^{-1}$ for all $\lambda \in \mathbb{C}$
- $\tilde{A} + \tilde{L}\tilde{C} = T(A + LC)T^{-1}$ if $\tilde{L} = TL$ similar

Note, $\tilde{M}_c \tilde{M}_o = TM_c M_o T^{-1}$ and $\tilde{W}_c[t] \tilde{W}_o[t] = TW_c[t] W_o[t] T^{-1}$ are similar.

Minimality

Definition

Let G be an n -dimensional LSI system. Its realization is (A, B, C, D) is said to be minimal if its state vector $x[t] \in \mathbb{R}^n$.

Theorem

A realization is minimal iff it is both reachable and observable.

Proof (outline): Follows from the rank decomposition

$$[H_G] = O_G R_G = \begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \end{bmatrix} \begin{bmatrix} B & AB & A^2B & \cdots \end{bmatrix},$$

and observability and reachability criteria.

If a realization (A, B, C, D) of G is minimal, then

$$z_p \text{ is a pole of } G(z) \iff z_p \in \text{spec}(A)$$

Minimality

Definition

Let G be an n -dimensional LSI system. Its realization is (A, B, C, D) is said to be minimal if its state vector $x[t] \in \mathbb{R}^n$.

Theorem

A realization is minimal iff it is both reachable and observable.

Proof (outline): Follows from the rank decomposition

$$\llbracket H_G \rrbracket = \mathbf{O}_G \mathbf{R}_G = \begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \end{bmatrix} [B \quad AB \quad A^2B \quad \cdots],$$

and observability and reachability criteria. □

If a realization (A, B, C, D) of G is minimal, then

λ_p is a pole of $G(z) \iff \lambda_p \in \text{spec}(A)$

Minimality

Definition

Let G be an n -dimensional LSI system. Its realization is (A, B, C, D) is said to be minimal if its state vector $x[t] \in \mathbb{R}^n$.

Theorem

A realization is minimal iff it is both reachable and observable.

Proof (outline): Follows from the rank decomposition

$$\llbracket H_G \rrbracket = \mathbf{O}_G \mathbf{R}_G = \begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \end{bmatrix} [B \quad AB \quad A^2B \quad \cdots],$$

and observability and reachability criteria. □

If a realization (A, B, C, D) of G is minimal, then

- z_p is a pole of $G(z) \iff z_p \in \text{spec}(A)$

Outline

Linear algebra preliminaries (self study)

State space

Structural properties of state-space realizations

Lyapunov stability

Lyapunov stability

Definition (Lyapunov stability)

An equilibrium x_{eq} of autonomous dynamics $x[t+1] = f(x[t])$ is said to be **stable** if for every $\epsilon > 0$ there is $\delta = \delta(\epsilon) > 0$ such that $\|x[0] - x_{\text{eq}}\| < \delta$ implies that $\|x[t] - x_{\text{eq}}\| < \epsilon$ for all $t \in \mathbb{N}$.

Definition (asymptotic Lyapunov stability)

An equilibrium x_{eq} of autonomous dynamics $x[t+1] = f(x[t])$ is said to be **asymptotically stable** if it is stable and $\lim_{t \rightarrow \infty} \|x[t] - x_{\text{eq}}\| = 0$.

Example (pendulum)

Assuming no friction:

stable equilibrium

unstable equilibrium

Lyapunov stability

Definition (Lyapunov stability)

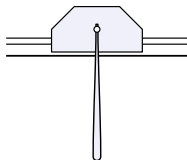
An equilibrium x_{eq} of autonomous dynamics $x[t+1] = f(x[t])$ is said to be **stable** if for every $\epsilon > 0$ there is $\delta = \delta(\epsilon) > 0$ such that $\|x[0] - x_{\text{eq}}\| < \delta$ implies that $\|x[t] - x_{\text{eq}}\| < \epsilon$ for all $t \in \mathbb{N}$.

Definition (asymptotic Lyapunov stability)

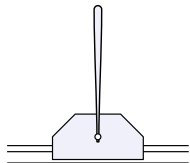
An equilibrium x_{eq} of autonomous dynamics $x[t+1] = f(x[t])$ is said to be **asymptotically stable** if it is stable and $\lim_{t \rightarrow \infty} \|x[t] - x_{\text{eq}}\| = 0$.

Example (pendulum)

Assuming no friction:



stable equilibrium



unstable equilibrium

Lyapunov stability: LSI case

If $f(x[t]) = Ax[t]$, then the (asymptotic) stability property is independent of the choice of the equilibrium and the system $x[t+1] = Ax[t]$ is

- stable iff $\text{spec}(A) \in \bar{\mathbb{D}}_1$ and all eigenvalues on $\mathbb{T}$, if exist, are simple
- asymptotically stable iff $\text{spec}(A) \in \mathbb{D}_1$

Although Lyapunov stability notions are concerned with unforced dynamics only, they are connected with i/o stability of systems. Namely, we have:

Theorem

If $x[t+1] = Ax[t]$ is asymptotically stable, then the system G having the transfer function $G(z) = D + C(zI - A)^{-1}B$ is ℓ_p -stable for all B , C , and D . The converse is true if (A, B) is stabilizable and (C, A) is detectable.

Lyapunov stability: LSI case

If $f(x[t]) = Ax[t]$, then the (asymptotic) stability property is independent of the choice of the equilibrium and the system $x[t+1] = Ax[t]$ is

- stable iff $\text{spec}(A) \in \bar{\mathbb{D}}_1$ and all eigenvalues on $\mathbb{T}$, if exist, are simple
- asymptotically stable iff $\text{spec}(A) \in \mathbb{D}_1$

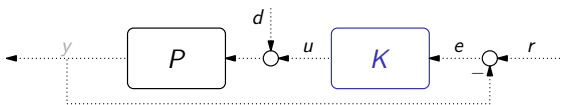
Although Lyapunov stability notions are concerned with unforced dynamics only², they are connected with i/o stability of systems. Namely, we have:

Theorem

If $x[t+1] = Ax[t]$ is asymptotically stable, then the system G having the transfer function $G(z) = D + C(zI - A)^{-1}B$ is ℓ_p -stable for all B , C , and D . The converse is true if (A, B) is stabilizable and (C, A) is detectable.

²Equivalently, only with the effect of the past on the evolution of the state.

Internal stability via Lyapunov stability



Internal stability requires GoF, $(r, d) \mapsto (e, u)$, to be stable. Let

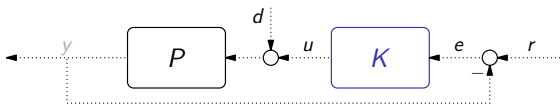
$$P : \begin{cases} x[t+1] = Ax[t] + Bu_p[t] \\ y[t] = Cx[t] \end{cases} \quad \text{and} \quad K : \begin{cases} x_k[t+1] = A_k x_k[t] + B_k e[t] \\ u[t] = C_k x_k[t] + D_k e[t] \end{cases}$$

Taking into account that $u_p = u + d$ and $e = r - y$, we end up with

$$\begin{cases} \begin{bmatrix} x[t+1] \\ x_k[t+1] \end{bmatrix} = \begin{bmatrix} A - BD_k C & BC_k \\ -B_k C & A_k \end{bmatrix} \begin{bmatrix} x[t] \\ x_k[t] \end{bmatrix} + \begin{bmatrix} BD_k & B \\ B_k & 0 \end{bmatrix} \begin{bmatrix} r[t] \\ d[t] \end{bmatrix} \\ \begin{bmatrix} e[t] \\ u[t] \end{bmatrix} = \begin{bmatrix} -C & 0 \\ -D_k C & C_k \end{bmatrix} \begin{bmatrix} x[t] \\ x_k[t] \end{bmatrix} + \begin{bmatrix} I & 0 \\ D_k & 0 \end{bmatrix} \begin{bmatrix} r[t] \\ d[t] \end{bmatrix} \end{cases}$$

(convince yourselves that this is true).

Internal stability via Lyapunov stability (contd)



$$A_{cl} := \begin{bmatrix} A - BD_k C & BC_k \\ -B_k C & A_k \end{bmatrix} \text{ is Schur} \implies \text{the system is internally stable.}$$

Lemma

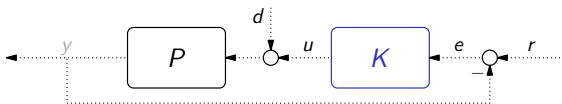
A_d can be made Schur only if (C, A, B) is stabilizable and detectable.

Proof: If (A, B) is not stabilizable, then there are λ_i , $|\lambda_i| \geq 1$, and $\eta_i' \neq 0$ such that $\eta_i' A = \lambda_i \eta_i'$ and $\eta_i' B = 0$. But then

$$\begin{bmatrix} \eta_i' & 0 \end{bmatrix} \begin{bmatrix} A - BD_k C & BC_k \\ -B_k C & A_k \end{bmatrix} = \begin{bmatrix} \eta_i' A - \eta_i' BD_k C & \eta_i' BC_k \end{bmatrix} = \lambda_i \begin{bmatrix} \eta_i' & 0 \end{bmatrix}$$

implies that $\lambda_i \in \text{spec}(A_{cl})$ and A_{cl} is not Schur regardless the controller K . Similar arguments apply if (C, A) is not detectable. $\square$

Internal stability via Lyapunov stability (contd)



$A_{cl} := \begin{bmatrix} A - BD_k C & BC_k \\ -B_k C & A_k \end{bmatrix}$ is Schur $\implies$ the system is internally stable.

Lemma

A_{cl} can be made Schur only if (C, A, B) is stabilizable and detectable.

Proof: If (A, B) is not stabilizable, then there are λ_i , $|\lambda_i| \geq 1$, and $\eta_i \neq 0$ such that $\eta_i' A = \lambda_i \eta_i'$ and $\eta_i' B = 0$. But then

$$\begin{bmatrix} \eta_i' & 0 \end{bmatrix} \begin{bmatrix} A - BD_k C & BC_k \\ -B_k C & A_k \end{bmatrix} = \begin{bmatrix} \eta_i' A - \eta_i' BD_k C & \eta_i' BC_k \end{bmatrix} = \lambda_i \begin{bmatrix} \eta_i' & 0 \end{bmatrix}$$

implies that $\lambda_i \in \text{spec}(A_{cl})$ and A_{cl} is not Schur regardless the controller K . Similar arguments apply if (C, A) is not detectable. $\square$