Lecture 2. Linear Systems

Ivan Papusha

CDS270–2: Mathematical Methods in Control and System Engineering

April 6, 2015

Logistics

- hw1 due this **Wed, Apr 8**
- hw due every Wed
  - in class, or
  - my mailbox on 3rd floor of Annenberg
- reading: BV Appendix A, pay attention to
  - linear algebra, notation
  - Schur complements (also in hw1)
  - $\geq$ vs $\succeq$
- hw2+ will be "choose your own adventure":

do an assigned problem

or

pick and do a problem from the catalog

Autonomous systems

Consider the *autonomous* linear dynamical system

$$\dot{x}(t) = Ax(t), \quad x(0) = x_0$$

- solution is matrix exponential

$$x(t) = e^{At}x(0),$$

where

$$\begin{aligned} e^{At} &\overset{\Delta}{=} I + At + \frac{1}{2!}A^2t^2 + \frac{1}{3!}A^3t^3 + \cdots \\ &= \sinh(At) + \cosh(At) \end{aligned}$$

Formal derivation from discrete time

Original continuous equation approximated by forward Euler for small timestep $\delta \ll 1$

$$\frac{x_{k+1} - x_k}{\delta} \approx A x_k, \quad x_k = x(k\delta), \quad k = 0, 1, 2, \ldots$$

Classic pattern for discrete time systems:

$$\begin{aligned}
x_0 &= x(0) = x_0 \\
x_1 &= x_0 + A\delta\, x_0 \\
x_2 &= (I + A\delta)^2 x_0 \\
&\vdots \\
x_k &= (I + A\delta)^k x_0 \\
&\vdots \\
x(t) &= \lim_{\delta \to 0^+} (I + A\delta)^{\lfloor t/\delta \rfloor} x_0 = e^{At} x_0
\end{aligned}$$

State propagation

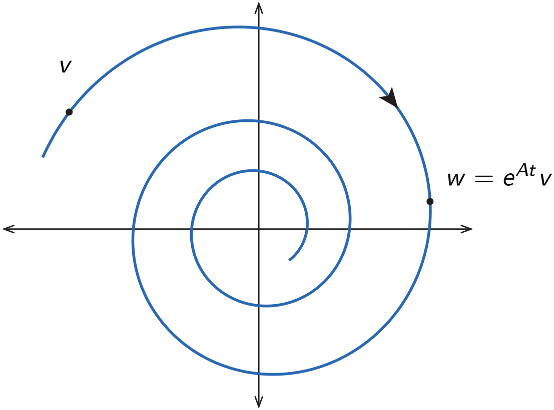
propagator. Multiplying by e^{At} propagates the autonomous state forward by time t. For $v, w \in \mathbf{R}^n$,

$$w = e^{At} v \quad \text{implies} \quad v = e^{-At} w.$$

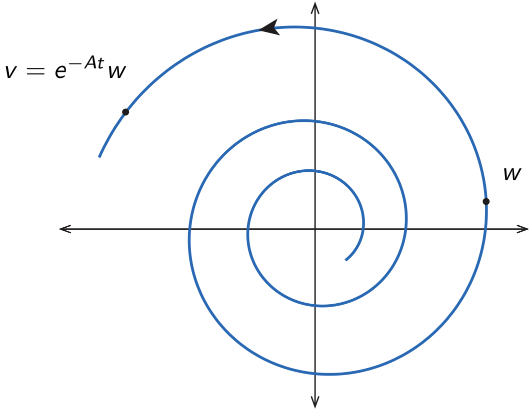
- the point w is v propagated by time t
- equivalently: the point v is w propagated by time $-t$
- current state contains all information
- matrix exponential is a time propagator (*huge* deal in physics, *e.g.*, Hamiltonians in quantum mechanics)

Markov property. future is independent of past given present

State propagation: forward



State propagation: backward



Example: output prediction from threshold alarms

An autonomous dynamical system evolves according to

$$\dot{x}(t) = Ax(t), \quad y(t) = Cx(t), \quad x(0) = x_0,$$

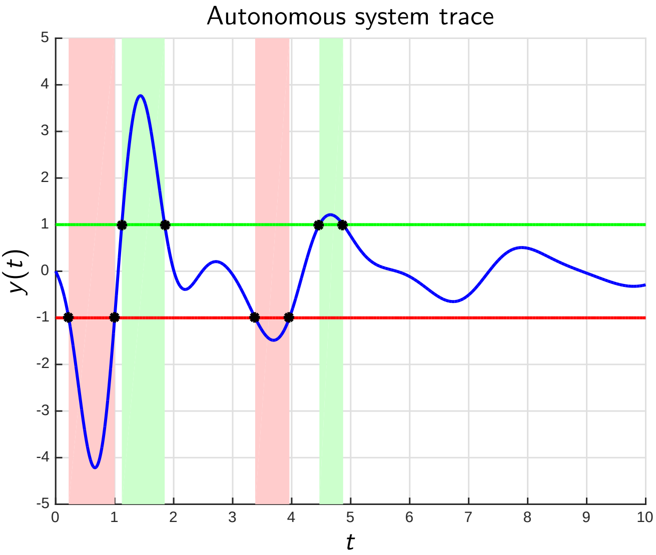
where

- $A \in \mathbf{R}^{n \times n}$, and $C \in \mathbf{R}^{1 \times n}$ are known, $n = 6$
- $x_0 \in \mathbf{R}^n$ is not known
- $x(t)$ and $y(t)$ are not directly measurable
- we have threshold information,

$$\begin{aligned} y(t) \geq 1, &\quad t \in [1.12, 1.85] \cup [4.47, 4.87], \\ y(t) \leq -1, &\quad t \in [0.22, 1.00] \cup [3.38, 3.96]. \end{aligned}$$

question: $x(10) =??$ (and is it unique?)

Example: output prediction from threshold alarms



Example: output prediction from threshold alarms

method 1. determine x_0, then propagate forward by 10s:

$$1 = Ce^{At_i}x_0, \quad t_i = 1.12, 1.85, 4.47, 4.87$$
$$-1 = Ce^{At_i}x_0, \quad t_i = 0.22, 1.00, 3.38, 3.96$$

$$\text{8 eqns, 6 unknowns} \quad \Longrightarrow \quad x(10) = e^{A\cdot 10}x_0$$

method 2. determine $x(10)$ directly:

$$1 = Ce^{-A(10-t_i)}x(10), \quad t_i = 1.12, 1.85, 4.47, 4.87$$
$$-1 = Ce^{-A(10-t_i)}x(10), \quad t_i = 0.22, 1.00, 3.38, 3.96$$

8 eqns, 6 unknowns

State propagation with inputs

For continuous time input-output system

$$\dot{x}(t) = Ax(t) + Bu(t)$$
$$y(t) = Cx(t) + Du(t)$$
$$x(0) = x_0,$$

if $u(\cdot) \not\equiv 0$, then the state propagator is a convolution operation,

$$x(t_0 + t) = e^{At}x(t_0) + \int_{t_0}^{t_0+t} e^{A(t_0+t-\tau)}Bu(\tau)\,d\tau$$

interpreted elementwise.

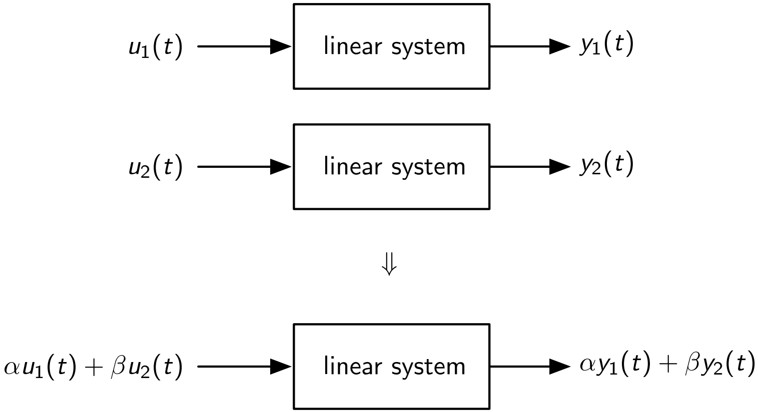
Impulse response

suppose the input is an impulse $u(t) = \delta(t)$

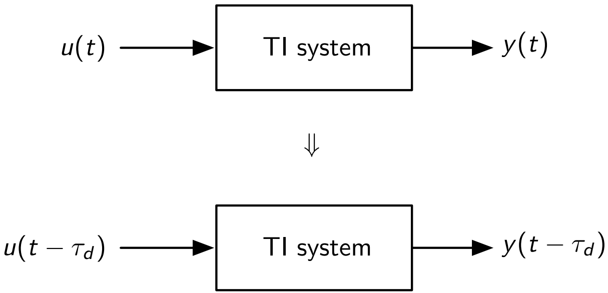
$$
\begin{aligned}
y(t) &= Ce^{At}x_0 + \int_{0^-}^{t} Ce^{A(t-\tau)}Bu(\tau)\,d\tau + Du(t) \\
&= Ce^{At}x_0 + \int_{0^-}^{t} Ce^{A(t-\tau)}B\delta(\tau)\,d\tau + D\delta(t) \\
&= Ce^{At}x_0 + Ce^{At}B + D\delta(t)
\end{aligned}
$$

- $Ce^{At}x_0$ is due to initial condition
- $h(t) = Ce^{At}B + D\delta(t)$ is the *impulse response*

Linearity


$u_1(t)$
linear system
$y_1(t)$
$u_2(t)$
linear system
$y_2(t)$
$\Downarrow$
$\alpha u_1(t) + \beta u_2(t)$
linear system
$\alpha y_1(t) + \beta y_2(t)$

Time invariance



Linear + Time Invariant (LTI) systems

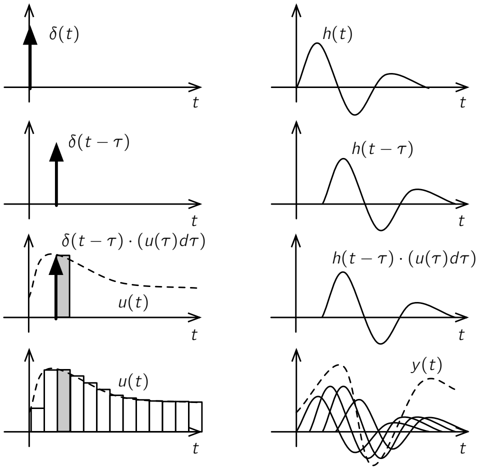
fact. (ÅM §5.3) if a system is LTI, its output is a convolution

$$y(t) = \int_{-\infty}^{\infty} h(t-\tau)u(\tau)\, d\tau$$

- $u(t)$: input
- $y(t)$: output
- $h(t)$: impulse response fully characterizes system for any input

$$\begin{aligned} y(t) = (h * u)(t) &= \int_{-\infty}^{\infty} h(t-\tau)u(\tau)\ d\tau \\ &= \int_{-\infty}^{\infty} h(\tau)u(t-\tau)\ d\tau \end{aligned}$$

Graphical interpretation



$$u(t) = \int_{-\infty}^{\infty} \delta(t-\tau)u(\tau)\ d\tau \qquad y(t) = \int_{-\infty}^{\infty} h(t-\tau)u(\tau)\ d\tau$$

Singular value decomposition

fact. every $m \times n$ matrix A can be factored as

$$A = U\Sigma V^T = \sum_{i=1}^{r} \sigma_i u_i v_i^T$$

where $r = \mathbf{rank}(A)$, $U \in \mathbf{R}^{m\times r}$, $U^T U = I$, $V \in \mathbf{R}^{n\times r}$, $V^T V = I$, $\Sigma = \mathbf{diag}(\sigma_1, \ldots, \sigma_r)$, and

$$\sigma_1 \geq \cdots \geq \sigma_r \geq 0.$$

- $u_i \in \mathbf{R}^m$ are the *left singular vectors*
- $v_i \in \mathbf{R}^n$ are the *right singular vectors*
- $\sigma_i \geq 0$ are the *singular values*

Singular value decomposition

The "thin" decomposition

$$A = \begin{bmatrix} | & & | \\ u_1 & \cdots & u_r \\ | & & | \end{bmatrix} \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_r \end{bmatrix} \begin{bmatrix} - v_1^T - \\ \vdots \\ - v_r^T - \end{bmatrix}$$

can be extended to a "thick" decomposition by completing the basis for $\mathbf{R}^m$ and $\mathbf{R}^n$ and making U, V square.

- $\{u_1, \ldots, u_r\}$ is an orthonormal basis for $\mathbf{range}(A)$
- $\{v_{r+1}, \ldots, v_n\}$ is an orthonormal basis for $\mathbf{null}(A)$

Controllability: testing for membership in span

Given a desired $y \in \mathbf{R}^m$, we have

$$y \in \textbf{range}(A) \iff \textbf{rank}\begin{bmatrix} A & y \end{bmatrix} = \textbf{rank}(A)$$
$$\iff y \in \textbf{span}\{u_1, \ldots, u_r\}.$$

The component of y in **range**(A) is $\sum_{i=1}^{r} u_i u_i^T y$.

$$y \in \textbf{range}(A) \iff y - \sum_{i=1}^{r} u_i u_i^T y = 0$$
$$\iff (I - UU^T)y = 0$$

Linear mappings and ellipsoids

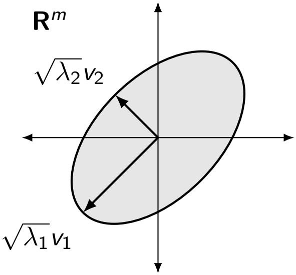
An $m \times n$ matrix A maps the unit ball in $\mathbf{R}^n$ to an ellipsoid in $\mathbf{R}^m$,

$$\mathcal{B} = \{x \in \mathbf{R}^n : \|x\| \leq 1\} \mapsto A\mathcal{B} = \{Ax : x \in \mathcal{B}\}.$$

- an ellipsoid induced by a matrix $P \in \mathbf{S}^m_{++}$ is the set

$$\mathcal{E}_P = \{x \in \mathbf{R}^m : x^T P^{-1} x \leq 1\}$$

- λ_i, v_i: eigenvalues and eigenvectors of P

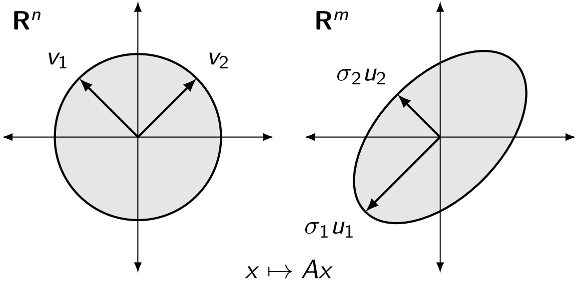


SVD Mapping

The SVD is a decomposition of the linear mapping $x \mapsto Ax$, such that right singular vectors v_i map to left singular vectors $\sigma_i u_i$

$$A = U \Sigma V^T = \sum_{i=1}^{r} = \sigma_i u_i v_i^T, \quad r = \textbf{rank}(A)$$

- equivalent "eigenvalue problem" is $Av_i = \sigma_i u_i$



Pseudoinverse

For $A = U\Sigma V^T \in \mathbf{R}^{m\times n}$ full rank, the pseudoinverse is

$$\begin{aligned} A^\dagger &= V\Sigma^{-1}U^T \\ &= \lim_{\mu\to 0}(A^TA+\mu I)^{-1}A^T = (A^TA)^{-1}A^T, \quad m \geq n \\ &= \lim_{\mu\to 0} A^T(AA^T+\mu I)^{-1} = A^T(AA^T)^{-1}, \quad m \leq n \end{aligned}$$

- **least norm (control):** if A is fat and full rank,

$$x^\star = \arg \min_{x\in\mathbf{R}^n, Ax=y} \|x\|_2^2 = A^\dagger y$$

- **least squares (estimation):** if A is skinny and full rank,

$$x^\star = \arg\min_{x\in\mathbf{R}^n} \|Ax-y\|_2^2 = A^\dagger y$$

Discrete convolution

Discrete linear system with impulse coefficients $h_0, \dots, h_{n-1}$,

$$y_k = \sum_{i=0}^{k} h_{k-i} u_i, \quad k = 0, \dots, n-1$$

or written as a matrix equation,

$$\begin{bmatrix} y_0 \\ y_1 \\ \vdots \\ y_{n-1} \end{bmatrix} = \begin{bmatrix} h_0 & 0 & \cdots & 0 \\ h_1 & h_0 & \cdots & 0 \\ & & \ddots & \\ h_{n-1} & h_{n-2} & \cdots & h_0 \end{bmatrix} \begin{bmatrix} u_0 \\ u_1 \\ \vdots \\ u_{n-1} \end{bmatrix}$$

Discrete convolution

Matrix structure gives rise to familiar system properties:

$$\underbrace{\begin{bmatrix} y_0 \\ y_1 \\ \vdots \\ y_{n-1} \end{bmatrix}}_{y} = \underbrace{\begin{bmatrix} h_0 & 0 & \cdots & 0 \\ h_1 & h_0 & \cdots & 0 \\ & & \ddots & \\ h_{n-1} & h_{n-2} & \cdots & h_0 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} u_0 \\ u_1 \\ \vdots \\ u_{n-1} \end{bmatrix}}_{u}$$

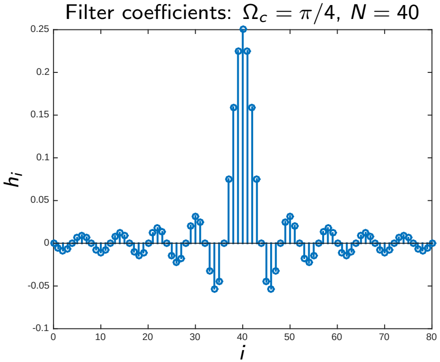
- A is a matrix: system is *linear*
- A lower triangular: system is *causal*
- A Toeplitz: system is *time-invariant*

open problem. how do you spell Otto Toeplitz's name?

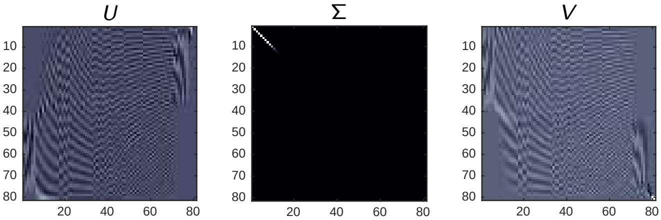
Typical filter

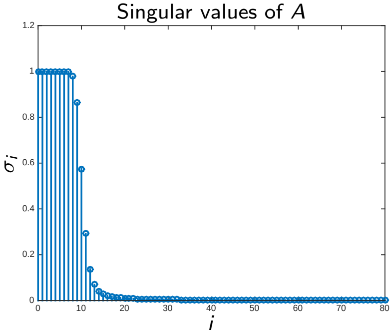
Causal finite $2N + 1$ point truncation of ideal low pass filter

$$h_i = \frac{\Omega_c}{\pi} \operatorname{sinc}\left(\frac{\Omega_c}{\pi}(i - N)\right) \quad i = 0, \ldots, 2N$$



SVD of filter matrix





Closely related: circulant matrices

A circulant matrix is a "folded over" Toeplitz matrix

$$C = \begin{bmatrix} c_0 & c_1 & c_2 & \cdots & c_{n-1} \\ c_{n-1} & c_0 & c_1 & \cdots & c_{n-2} \\ c_{n-2} & c_{n-1} & c_0 & \ddots & c_{n-3} \\ \vdots & & \ddots & \ddots & \vdots \\ c_1 & c_2 & c_3 & \cdots & c_0 \end{bmatrix}$$

- eigenvalues are the DFT of the sequence $\{c_0, \ldots, c_{n-1}\}$

$$\lambda^{(k)} = \sum_{i=0}^{n-1} c_i e^{-2\pi jki/n}, \quad v^{(k)} = \frac{1}{\sqrt{n}} \begin{bmatrix} 1 \\ e^{-2\pi jk/n} \\ \vdots \\ e^{-2\pi jk(n-1)/n} \end{bmatrix}$$

- for more, see R. M. Gray "Toeplitz and Circulant Matrices"

Equalizer design

causal deconvolution problem. pick filter coefficients g such that $g * h$ is as close as possible to the identity

$$\begin{bmatrix} u_0 \\ u_1 \\ \vdots \\ u_{n-1} \end{bmatrix} \approx \begin{bmatrix} g_0 & 0 & \cdots & 0 \\ g_1 & g_0 & \cdots & 0 \\ & & \ddots & \\ g_{n-1} & g_{n-2} & \cdots & g_0 \end{bmatrix} \begin{bmatrix} h_0 & 0 & \cdots & 0 \\ h_1 & h_0 & \cdots & 0 \\ & & \ddots & \\ h_{n-1} & h_{n-2} & \cdots & h_0 \end{bmatrix} \begin{bmatrix} u_0 \\ u_1 \\ \vdots \\ u_{n-1} \end{bmatrix}$$

- $G^\star = A^\dagger$ minimizes $\|GA - I\|_F^2$ over all $G \in \mathbf{R}^{n \times n}$
- often want G causal, *i.e.*, lower triangular Toeplitz (ltt.)

$$\begin{array}{ll} \text{minimize} & \|GA - I\|_F^2 \\ \text{subject to} & G \text{ is ltt.} \end{array}$$

- for more, see Wiener–Hopf filter

Markov parameters

Consider the discrete-time LTI system

$$
\begin{aligned}
x_{k+1} &= Ax_k + Bu_k \\
y_k &= Cx_k \\
x(0) &= x_0.
\end{aligned}
$$

For each $k = 0, 1, 2, \ldots$, we have

$$
\begin{bmatrix} y_0 \\ y_1 \\ y_2 \\ \vdots \\ y_k \end{bmatrix} = \begin{bmatrix} 0 & & & & \\ CB & 0 & & & \\ CAB & CB & 0 & & \\ \vdots & & & \ddots & \\ CA^{k-1}B & CA^{k-2}B & CA^{k-3}B & \cdots & 0 \end{bmatrix} \begin{bmatrix} u_0 \\ u_1 \\ u_2 \\ \vdots \\ u_k \end{bmatrix} + \begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \\ CA^k \end{bmatrix} x_0.
$$

Minimum energy control

Given matrices A, B, C, and initial condition x_0, find a minimum energy input sequence $u_0, \ldots, u_k$ that achieves desired outputs $y_{k_1}^{\text{des}}, \ldots, y_{k_\ell}^{\text{des}}$ at times $k_1, \ldots, k_\ell$.

$$\underbrace{\begin{bmatrix} y_{k_1}^{\text{des}} \\ y_{k_2}^{\text{des}} \\ \vdots \\ y_{k_\ell}^{\text{des}} \end{bmatrix}}_{y} = \underbrace{\begin{bmatrix} CA^{k_1-1}B & CA^{k_1-2}B & \ldots & 0 \\ CA^{k_2-1}B & CA^{k_2-2}B & \ldots & 0 \\ \vdots & & \ddots & \\ CA^{k_\ell-1}B & CA^{k_\ell-2}B & \ldots & 0 \end{bmatrix}}_{H} \underbrace{\begin{bmatrix} u_0 \\ u_1 \\ \vdots \\ u_k \end{bmatrix}}_{u} + \underbrace{\begin{bmatrix} CA^{k_1} \\ CA^{k_2} \\ \vdots \\ CA^{k_\ell} \end{bmatrix}}_{G} x_0.$$

- if $y - Gx_0 \in \textbf{range}(H)$, a minimizing sequence is

$$u = H^\dagger(y - Gx_0).$$

- *left* singular vectors of H with large σ_i are modes with lots of actuator authority

Least squares estimation

Given matrices A, C, $(B = 0)$ and measured (noisy) outputs $y^{\text{meas}}_{k_1}, \dots, y^{\text{meas}}_{k_\ell}$ at times $k_1, \dots, k_\ell$, find the best least squares estimate of the initial condition x_0.

$$\underbrace{\begin{bmatrix} y^{\text{meas}}_{k_1} \\ y^{\text{meas}}_{k_2} \\ \vdots \\ y^{\text{meas}}_{k_\ell} \end{bmatrix}}_{y} = \underbrace{\begin{bmatrix} CA^{k_1} \\ CA^{k_2} \\ \vdots \\ CA^{k_\ell} \end{bmatrix}}_{G} x_0 + \text{noise}$$

- the least squares estimate is

$$x_0 = G^\dagger y$$

- *right* singular vectors of G with large σ_i are modes with lots of sensitivity
- **null**(G) is unobservable space