Linear Systems
Lecture 0
Some Math Background

Department Automatic Control
Lund University

Lecture 0

- Course Contents
- Matrix theory
- Norms
- Distributions
- Fourier and Laplace Transforms

Material: Lecture slides

Linear Systems I, 2014

- Introduction
- Multivariable Time-varying Systems
- Transition Matrices
- Controllability and Observability
- Realization Theory
- Stability Theory
- Linear Feedback
- Multivariable input/output descriptions
- Some Bonus Material

Linear Systems I, 2014

Rugh, Linear System Theory, 2nd edition

- Most of 1-7,9-12,13-14,16-17
- Scan 15,20-23,25-29
- Skip 8,18-19, 24

Optimization by Vector Space Methods, Handout Lec 6

Some more handouts

Course Contents

Credits: 9hp

- 9 Lectures (incl this intro)
- 8 Exercise sessions (nothing this week)
- 8 Handins (7 best counts). Strict deadlines!
- Take home exam 24 hours

Matrix Theory

Definition and standard rules for $\det(A) = \prod \lambda_i$ and $\mathrm{tr}(A) = \sum \lambda_i$

$\det(AB) = \det(A)\det(B)$, $\mathrm{tr}(AB) = \mathrm{tr}(BA)$

$(AB)^{-1} = B^{-1}A^{-1}$ and $(AB)^T = B^T A^T$

$\det(A) = \sum_i a_{ij}c_{ij} = \sum_j a_{ij}c_{ij}$

cofactors $c_{ij} = (-1)^{i+j}\det(A')$ (delete row i and col j)

$\mathrm{adj}(A) = C^T$

$A\,\mathrm{adj}(A) = \det(A)I$, so $A^{-1} = \frac{\mathrm{adj}(A)}{\det(A)}$

$\frac{d}{dt}(AB) = \frac{dA}{dt}B + A\frac{dB}{dt}$

If in need: Google "Matrix Cookbook"

Eigenvalues

$Av = \lambda v$

Characteristic equation $p(\lambda) = \det(\lambda I - A) = 0$

Geometric multiplicity $\leq$ Algebraic multiplicity

If $A^T = A$ then eigenvalues are real and there are n orthogonal eigenvectors: $A = V\Lambda V^T$ with $V^T V = I$

General A: Jordan normal form

$$A = V \text{ blockdiag } (J_i)V^{-1} \text{ where } J_i = \begin{pmatrix} \lambda_i & 1 & \\ & \ddots & 1 \\ & & \lambda_i \end{pmatrix}$$

Cayley-Hamilton: $p(A) = 0$

Computation of e^{At}

Definition: $e^{At} = \sum_{k=0}^{\infty} \frac{1}{k!}(At)^k$. Satisfies $\frac{dX}{dt} = AX$.

$\frac{d}{dt}e^{At} = Ae^{At} = e^{At}A$

If $A = V\Lambda V^T$ then $e^{At} = V\mathrm{diag}(e^{\lambda_i t})V^T$

If $A = V$ blockdiag $(J_i)V^{-1}$ then $e^{At} = V$ blockdiag $(e^{J_i t})V^{-1}$

where $e^{J_i t} = \begin{pmatrix} e^{\lambda_i t} & te^{\lambda_i t} & \dots & \frac{t^{n_i-1}}{(n_i-1)!}e^{\lambda_i t} \\ & \ddots & \ddots & \\ & & e^{\lambda_i t} & te^{\lambda_i t} \\ & & & e^{\lambda_i t} \end{pmatrix}$

Laplace-transform $\mathcal{L}(e^{At}) = (sI - A)^{-1}$

$e^{(A+B)t} = e^{At}e^{Bt}$ for all $t \Leftrightarrow AB = BA$.

Singular Value Decomposition etc

If $A \in R^{m \times n}$ then

$$A = U \begin{pmatrix} \Sigma & 0 \\ 0 & 0 \end{pmatrix} V^T$$

where $U \in R^{m \times m}, V \in R^{n \times n}$ orthogonal and

$\Sigma = \mathrm{diag}(\sigma_1, \ldots, \sigma_r) > 0$ diagonal of size r=rank(A)

A symmetric $\Longrightarrow A = U \Sigma U^T$.

Geometric View

$$A = \begin{pmatrix} U_1 & \dots & U_r & \dots U_m \end{pmatrix} \begin{pmatrix} \Sigma & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} V_1^T \\ \vdots \\ V_r^T \\ \vdots \\ V_n^T \end{pmatrix}$$

Null space $N(A) := \{x \mid Ax = 0\}$

Range space $R(A) := \{y \mid y = Ax \text{ for some } x\}$

$$R^n = \underbrace{R(A^T)}_{\text{spanned by } V_1 \ldots V_r} \oplus \underbrace{N(A)}_{\text{spanned by } V_{r+1} \ldots V_n}$$

$$R^m = \underbrace{R(A)}_{\text{spanned by } U_1 \ldots U_r} \oplus \underbrace{N(A^T)}_{\text{spanned by } U_{r+1} \ldots U_m}$$

Quadratic Forms $x^T Ax$

Let's assume $A^T = A$ (note that $x^T Ax = x^T(A + A^T)x/2$)

$A \geq 0 \quad \Leftrightarrow \quad x^T Ax \geq 0, \forall x$

$A > 0 \quad \Leftrightarrow \quad x^T Ax > 0, \forall x \neq 0$

We say that $A \geq B$ iff $A - B \geq 0$.

Courant-Fisher formulas when $A^T = A$:

$\lambda_{max}(A) = \max_{x \neq 0} \frac{x^T Ax}{x^T x} = \max_{x^T x = 1} x^T Ax$

$\lambda_{min}(A) = \min_{x \neq 0} \frac{x^T Ax}{x^T x} = \min_{x^T x = 1} x^T Ax$

$\lambda_{min}(A)I \leq A \leq \lambda_{max} I$

$A > 0 \Leftrightarrow \lambda_i(A) > 0, \forall i$

Norms

A norm is a real-valued function satisfying

$$\|x\| \geq 0, \text{ with equality iff } x = 0 \tag{1}$$
$$\|\alpha x\| = |\alpha|\|x\| \tag{2}$$
$$\|x + y\| \leq \|x\| + \|y\| \tag{3}$$

Some vector norms on R^n

$$\|x\|_1 = \sum |x_i|$$
$$\|x\|_2 = \left(\sum |x_i|^2\right)^{1/2}$$
$$\|x\|_\infty = \max |x_i|$$
$$\|x\|_p = \left(\sum |x_i|^p\right)^{1/p}, \quad 1 \leq p \leq \infty$$

Signal Norms

$$\|f\|_p = \left(\int_{-\infty}^{\infty} |f(t)|^p dt \right)^{1/p}$$

For $p = 2$, called "signal-energy"

$L_p(I)$ denotes functions with $\int_I |f(t)|^p dt < \infty$

Matrix Norms

A matrix norm is a function satisfying (1)-(3) above

Examples: (induced matrix norms)

$$\|A\|_{\alpha,\beta} = \sup_{x \neq 0} \frac{\|Ax\|_\beta}{\|x\|_\alpha}$$

Induced 2-norm

$$\|A\|_2 = \sup_{x \neq 0} \frac{\|Ax\|_2}{\|x\|_2} = \sigma_{max}(A)$$

This is often the "default-norm".

Submultiplicative Matrix Norms

If the norm also satisfies $\|AB\| \leq \|A\|\|B\|$ it is called *submultiplicative*

All induced matrix norms are submultiplicative.

Frobenius-norm or Hilbert-Schmidt norm (submultiplicative, but not an induced norm)

$$\|A\|_F = \left(\sum_{i,j} |a_{ij}|^2 \right)^{1/2} = \Big(\mathrm{Trace}(A^T A) \Big)^{1/2}$$

Scalar Products

A scalar product $\langle \cdot , \cdot \rangle$ $V \times V \mapsto \mathcal{C}$ satisfies

$$\begin{aligned} \text{Positive definite} \quad & \langle x, x \rangle \geq 0 \text{ with equality iff } x = 0 \\ \text{Conjugate symmetric} \quad & \langle x, y \rangle = \overline{\langle y, x \rangle} \\ \text{Linearity} \quad & \langle x, \lambda_1 y_1 + \lambda_2 y_2 \rangle = \lambda_1 \langle x, y_1 \rangle + \lambda_2 \langle x, y_2 \rangle \end{aligned}$$

Examples

$$\begin{aligned} \langle x, y \rangle &= x^* y \\ \langle X, Y \rangle &= \text{Trace}(X^* Y) \\ \langle x(t), y(t) \rangle &= \int x(t)^* y(t) dt \end{aligned}$$

Scalar Products

We say that x and y are orthogonal, denoted $x \perp y$ if $\langle x, y \rangle = 0$

For subspace: $X \perp Y$ means that $x \perp y, \forall x \in X, y \in Y$

Example: $\cos t$ is orthogonal to $\sin t$ in $V = L_2([-\pi, \pi])$

Cauchy-Schwarz' inequality:

$$\sum_{i=1}^{n} |x_i y_i| = \langle x, y \rangle \leq \|x\|_2 \|y\|_2$$

(with equality if and only if x and y are proportional)

Summary of Distributions

A distribution S is a "generalized function".

One defines the distribution via its action on (suitably smooth) test-functions $\varphi \in \mathcal{D}$

Different notations: $S(\varphi) = \langle S, \varphi \rangle = \int S(t)\phi(t)dt$

Standard choice:

$$\mathcal{D} = C_0^\infty = \ \{\varphi \mid \ \varphi \in C^\infty, \ \varphi \text{ has compact support}\}$$

$S(\varphi)$ should be linear and "continuous" function of φ

Example

Dirac function defined via $\int \delta(t)\varphi(t)dt = \varphi(0)$

Derivative of S is defined via $S'(\varphi) = -S(\varphi')$

$\int \delta^{(k)}(t)\varphi(t)dt = (-1)^k \varphi^{(k)}(0)$

Step function $\theta(x) = 1_{x>0}$

$\theta'(x) = \delta(x)$

5 min Exercise

Verify that we for a smooth f have

$$\delta(x) f(x) = \delta(x) f(0)$$

where equality is interpreted in the sense of distributions

Also show that

$$\delta'(x) f(x) = \delta'(x) f(0) - \delta(x) f'(0)$$

Similarly simplify $\delta^{(k)}(x) f(x)$

Distributions with point support

It can be proved that if S is a distribution (with "order" k) and if $S(x) = 0$ for $x \neq 0$ then

$$S(x) = c_0\delta(x) + c_1\delta^{(1)}(x) + \ldots + c_k\delta^{(k)}(x)$$

Example:

$$xS(x) = 0 \quad \Longrightarrow \quad S(x) = c_0\delta(x)$$

Fourier Transforms

$$\mathcal{F}(\omega) = \widehat{f}(\omega) = \int_{-\infty}^{\infty} e^{-i\omega t} f(t) dt$$

Well defined in classical sense if $\int |f(t)| dt < \infty$

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i\omega t} \mathcal{F}(\omega) d\omega$$

$$\int_{-\infty}^{\infty} f(t)\overline{g(t)} dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} \mathcal{F}(\omega)\overline{\mathcal{G}(\omega)} d\omega$$

$$\int_{-\infty}^{\infty} f(x)\widehat{g}(x) dx = \int_{-\infty}^{\infty} \widehat{f}(x) g(x) dx$$

$$\widehat{f * g} = \widehat{f}\widehat{g}$$

$$\widehat{fg} = (2\pi)^{-1} \widehat{f} * \widehat{g}$$

Fourier Transforms with distributions

$\widehat{f}$ can be defined, as a distribution, via

$$\langle \widehat{f}, \varphi \rangle := \langle f, \widehat{\varphi} \rangle, \quad \varphi \in \mathcal{S}$$

where $\phi \in \mathcal{S}$ if $\phi \in C^\infty$ and $\sup |x^n \phi^{(k)}(x)| < \infty$ for $n, k \geq 0$

This gives meaning to Fourier transforms of e.g. functions satisfying $f \in C^\infty$ and $\sup_t |(1+|t|)^{-n} f(t)| < \infty, \text{ for some } n$ ("at most polynomial growth")

Examples

$$f(t) \equiv 1, \quad \mathcal{F}(\omega) = 2\pi\delta(\omega)$$

$$f(t) = \cos(\omega_0 t), \quad \mathcal{F}(\omega) = \pi\left(\delta(\omega - \omega_0) + \delta(\omega - \omega_0)\right)$$

$$f(t) = \delta(t), \quad \mathcal{F}(\omega) \equiv 1$$

$$f'(t) \mapsto i\omega\mathcal{F}(\omega)$$

$$f(at) \mapsto \frac{1}{|a|}\mathcal{F}\left(\frac{\omega}{a}\right)$$

$$f(t) = \sum_{n=-\infty}^{\infty} \delta(x - nT), \quad \mathcal{F}(\omega) = \frac{2\pi}{T}\sum_{k=-\infty}^{\infty} \delta\left(\omega - \frac{2\pi k}{T}\right)$$

Laplace Transforms

Double-sided Laplace-transform $F(s) := \int_{-\infty}^{\infty} e^{-st} f(t) dt$

Well defined in classical sense if $\int |e^{-st} f(t)| dt < \infty$

Typically satisfied in a strip $a < \mathrm{Re}(s) < b$

Examples

$f(t) = \theta(t), \quad F(s) = \frac{1}{s}$ for $\mathrm{Re} s > 0$

$f(t) = \frac{t^k}{k!}\theta(t), \quad F(s) = \frac{1}{s^{k+1}}$ for $\mathrm{Re} s > 0$

$f(t) = e^{At}\theta(t), \quad F(s) = (sI - A)^{-1}$ when $\mathrm{Re}(s) > \max \mathrm{Re} \lambda_i(A)$

One-sided Laplace Transform

$$F(s) := \int_{0-}^{\infty} e^{-st} f(t) dt$$

Convergence strip is now $a < \mathrm{Re}(s)$

Laplace transform of f' is $sF(s) - f(0)$

$\dot{x} = Ax, x(0) = x_0, \quad (sI - A)X(s) = x_0$

when $\mathrm{Re}(s) > \max \mathrm{Re}\lambda_i(A)$

10 min problem

What is the Fourier-transform of step function $f(t) = \theta(t)$?

Hint: First guess $\mathcal{F}(\omega) = \frac{1}{i\omega}$ is wrong

Exercise: Markov Parameters and Moments

Consider $h(t) = ce^{At}b\theta(t)$ and $H(s) = c(sI - A)^{-1}b$

- Expansion of $H(s)$ around $s = \infty$ gives

$$H(s) = c(I/s + A/s^2 + A^2/s^3 + \ldots)b = \sum_{k=1}^{\infty} h_k/s^k$$

where $h_k = h^{(k)}(0) = cA^{k-1}b$ are the *Markov parameters*.

- If A as. stable then Taylor-expansion of $H(s)$ around $s = 0$ is

$$H(s) = \sum_{k=0}^{\infty} m_k s^k,$$

where $m_k = \int_0^\infty \frac{(-t)^k}{k!} h(t) dt$ are the *moments* of $h(t)$.

Heat transfer in semi-infinite rod

$T(t, x)$ temperature at time t at x

$$\frac{dT}{dt} = \frac{d^2T}{dx^2}$$

Assume input: $T(t, 0) = u(t) = e^{st}$

Assume output at $x > 0$: $T(t, x) = \Psi(x)e^{st}$ (neglect transients)

$$s\Psi(x) = \frac{d^2\Psi}{dx^2}, \qquad \Psi(0) = 1$$

Gives $\Psi(x) = Ae^{x\sqrt{s}} + Be^{-x\sqrt{s}} \Longrightarrow A = 0$ and $B = 1$

Hence transfer function from $u(t)$ to $T(t, x_0)$ equals

$$H(s) = e^{-x_0\sqrt{s}}$$

Tools

Make sure you know how to simulate an ordinary differential system in e.g. Matlab/Simulink or Maple

You should also be familiar with using some symbolic manipulation program such as Maple

You should be able to use the Control System Toolbox (or similar)

Handin 1

1. Use Matlab and/or Maple to calculate characteristic polynomial, eigenvalues, eigenvectors and e^{At} both numerically and symbolically for $A = \begin{pmatrix} 0 & -1 \\ 1 & -1 \end{pmatrix}$.

2. The following frequency domain based code can be used (why?) to simulate the step response of the system $1/(s+1)$.

```
N=2^12; dt=0.01; T=N*dt; dw=2*pi/T;
t = dt*(0:N-1);
omega = -pi/dt:dw:(pi/dt-dw);
u = [ones(1,N/2) zeros(1,N/2)];
U = fft(u);
P = 1./(i*omega+1);
y = ifft(fftshift(P).*U);
plot(t+dt/2,real(y),'-bx');
hold on;grid on
plot(t,1-exp(-t),'-ro')
```

Handin 1 - continued

Simulate the step response of the open loop system $P(s) = \exp\left(-\sqrt{s}\right)$ and of the closed loop system $PC/(1+PC)$ under PI-control with $C(s) = 1 + 1/s$ (you might want to tune N and dt).

Compare the rise time to 50% and the settling times to 99% of the final value for open loop vs closed loop control.

3. See exercise session.