LionSealGrey

Linear Systems
Lecture 0
Some Math Background

Department Automatic Control
Lund University

Lecture 0

- Course Contents
- Vector spaces and mappings
- Matrix theory
- Norms

Material:

- Lecture slides
- R.A. Horn and C.R. Johnson. *Matrix Analysis*. Cambridge University Press, 2013.

Linear Systems I, 2016

- 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, 2016

Rugh, Linear System Theory, 2nd edition

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

J. P. Hespanha, *Linear Systems Theory*. Princeton University Press. 2009.
Some more handouts

Course Contents

Credits: 9hp

- 9 Lectures (including this intro)
- 8 Exercise sessions (1st one on Wednesday, this week)
- 8 Handins (7 best counts). Strict deadlines!
- 24 hour take-home exam (date tbd: 8-th Dec or mid-January 2017)

Vector spaces

A set of elements $\{v_k\}_{k=1}^n$ in a vector space $\mathcal{V}$ over field $\mathbb{F}$ is:

- *linearly independent*, if $\sum_{k=1}^n \alpha_k v_k = 0 \implies \alpha_k = 0, \forall k$.
- $\{v_k\}_{k=1}^n$ forms a basis for $\mathcal{V}$.
- If $\{v_k\}_{k=1}^n$ exists for finite n, $\mathcal{V}$ is *finite-dimensional*. Otherwise, $\mathcal{V}$ is *infinite dimensional*.
- A subset $\mathcal{U}$ of a vector space $\mathcal{V}$ is called a *subspace* if

$$au_1 + bu_2, \forall u_1, u_2 \in \mathcal{U}, \text{ and } a, b \in \mathbb{F}.$$

Mappings

A *functional* mapping A from subspace $\mathcal{U}$ into a vector space $\mathcal{W}$ is done by associating each $u \in \mathcal{U}$ with a *single* $w \in \mathcal{W}$. Usually denoted by $u \mapsto w = Au$.

w is the *range (image)* of u under A. The subspace is the *domain*, denoted by $\mathrm{dom}(A)$. The *range* of A is the set of all images

$$\mathrm{range}(A) := \{w \in \mathcal{W} : w = Au, u \in \mathrm{dom}(A)\}.$$

The *inverse image* $w_0 \in \mathcal{W}$ is the set of all $u \in \mathrm{dom}(A)$ such that $w_0 = Au$. We obtain the *inverse map* of A by associating each $w \in \mathrm{range}(A)$ with its inverse image.

A functional mapping $A : \mathcal{U} \to \mathcal{W}$ is *injective* (one-to-one) if, for every $u_1, u_2 \in \mathrm{dom}(A)$, $u_1 \neq u_2 \Rightarrow Au_1 \neq Au_2$. It is *surjective* if $\mathrm{range}(A) = \mathcal{W}$, and *bijective* if both.

Matrix representation of mappings

Given two vector spaces $\mathcal{V}$ and $\mathcal{W}$ over $\mathbb{F}$, a mapping $A : \mathcal{V} \to \mathcal{W}$ is linear if

$$A(av + bu) = aAv + bAu, \qquad \forall u, v \in \mathcal{V}, \text{ and } a, b \in \mathbb{F}.$$

Let $\{v_k\}_{k=1}^n$ and $\{w_k\}_{k=1}^m$ be bases for $\mathcal{V}$ and $\mathcal{W}$, respectively. For each basis vector v_k, let $\{a_{1k}, a_{2k}, \ldots, a_{mk}\}$ be the unique scalars satisfying

$$Av_k = a_{1k}w_1 + \cdots + a_{mk}w_m.$$

The mn scalars $a_{lk} \in \mathbb{F}$ completely characterises the map A. (why?)

Matrix representation of mappings

Let $\{v_k\}_{k=1}^n$ and $\{w_k\}_{k=1}^m$ be bases for $\mathcal{V}$ and $\mathcal{W}$, respectively. For each basis vector v_k, let $\{a_{1k}, a_{2k}, \ldots, a_{mk}\}$ be the unique scalars satisfying

$$Av_k = a_{1k}w_1 + \cdots + a_{mk}w_m.$$

The mn scalars $a_{lk} \in \mathbb{F}$ completely characterises the map A. Given any $v = \alpha_1 v_1 + \cdots + \alpha_n v_n$ and let $w = Av = \beta_1 w_1 + \cdots + \beta_n w_n$, by *linearity* we obtain

$$\begin{bmatrix} \beta_1 \\ \vdots \\ \beta_m \end{bmatrix} = \begin{bmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \ldots & a_{mn} \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix}.$$

The matrix $[a_{jk}] \in \mathbb{F}^{m \times n}$ is the *matrix representation* of the linear map A w.r.t. the *input basis* $\{v_k\}_{k=1}^n$ and *output basis* $\{w_k\}_{k=1}^m$.

Matrix Theory

Definition and standard rules

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

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

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

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

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

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

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

Eigenvalues

$Av = \lambda v$

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

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

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

Number of Jordan blocks J_i = total number of independent eigenvectors of A.

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 (i.e. $UU^T = I$ and $VV^T = I$) and

$\Sigma = \mathrm{diag}(\sigma_1, \ldots, \sigma_r) > 0$, where σ_i is the square-root of an eigenvalue of AA^T.

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 (kernel) $null(A) := \{x \mid Ax = 0\}$

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

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

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

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

  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 $t \Leftrightarrow AB = BA$. Note: In general, $e^{At}e^{Bt} \neq e^{(A+B)t}$.

Quadratic Forms $x^T Ax$

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

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

Positive semi-definite: $A > 0 \quad \Leftrightarrow \quad x^T Ax > 0, \forall x \neq 0$

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

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

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

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

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

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

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

$$\|x\|_1 = \sum |x_i|$$

$$\|x\|_2 = \left(\sum |x_i|^2\right)^{1/2}$$

$$\|x\|_\infty = \max |x_i|$$

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

Norms: why are they useful?

A sequence $\{v_k\}_{k=1}^{n}$ in a normed vector space $\mathcal{V}$ is said to converge, if $\exists v \in \mathcal{V}$ such that

$$\|v - v_k\|_{\mathcal{V}} \to 0, \text{ as } k \to \infty.$$

If such a v exists, it is unique (why?).

Note that norms quantify the 'closeness' of two elements in a vector space, as we have seen above, i.e. converts convergence of $\{v_k\}_{k=0}^{\infty}$ to a vector v to convergence of $\{\|v - v_k\|\}_{k=0}^{\infty}$ to 0!

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 (Inner Products)

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

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

Scalar Products (Inner Products)

A vector space $\mathcal{V}$ equipped with a scalar product is called a *scalar product (inner product) space*.

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)

Why are these concepts useful?

In this course, we use vector spaces equipped with an inner product and corresponding norm. All these vector spaces have an additional property which is useful in the study of sequence in the vector space (recall why a norm is useful).

A sequence $\{v_k\}_{k=0}^{\infty}$ in a normed vector space $\mathcal{V}$ is *Cauchy*, if for any $\epsilon > 0$, there exists $N(\epsilon)$ such that

$$\|v_k - v_m\|_{\mathcal{V}} < \epsilon, \qquad \forall k, m \geq N(\epsilon).$$

Note: Every convergent sequence is Cauchy, but not necessarily the converse.

Why are these concepts useful?

A normed vector space + every Cauchy sequence is convergent is called *complete* and known as a *Banach space*.

A Banach space + scalar product is called a *Hilbert space*.

In a complete vector space, it is possible to check whether a sequence is convergent by checking if it is Cauchy.

We can consider the modelling of a system in terms of *mappings* between signal vector spaces. In this course, we deal with mappings between *Banach spaces*.

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.