Selected topics in linear algebra.

- Essential basics

  - vector spaces
  - vector and matrix norms
  - definite and semidefinite matrices
  - invariant subspaces ◁

- Some of you might be less familiar with

  - singular value decomposition ◁
  - Schur inversion formulae ◁
  - Sylvester equation ◁

- Very important for this course

  - algebraic Riccati equation ◁

---

§2.4 * <u>Invariant subspaces</u>

- A subspace $V \subset \mathbb{R}^n$ is an invariant subspace of $A \in \mathbb{C}^{n\times n}$ if

$$Ax \in V \,, \quad \forall x \in V$$

- Trivial examples are

$$\{0\} \,, \ \mathrm{Ker}(A) \,, \ \mathrm{Im}(A)$$

- A less trivial example: Let $(\lambda_1, \dots, \lambda_k)$ be a part of a spectrum of $\mathcal{A}$ and $(v_1, \dots v_k)$ be all the associated generalized eigenvectors. Then, $\text{span}(v_1, \dots, v_n)$ is a subspace of $\mathcal{A}$.

- Any invariant subspace can be associated with a part of spectrum.

- Matrix criterion for $\mathcal{A}$-invariance:
  Consider $\mathcal{A} \in \mathbb{R}^{n \times n}$ and $\mathcal{M} \in \mathbb{R}^{n \times k}$ for $k \leq n$.
  Let V be a span of the columns of $\mathcal{M}$.
  Then:
  1) V is $\mathcal{A}$-invariant iff $\exists\, \mathcal{A}_V \in \mathbb{R}^{k \times k}$ such that $\mathcal{A}\mathcal{M} = \mathcal{M}\mathcal{A}_V$.
  2) $\text{spec}(\mathcal{A}_V) \subset \text{spec}(\mathcal{A})$
  3) V is a span of the generalized eigenvectors of $\mathcal{A}$ associated with $\text{spec}(\mathcal{A}_V)$.

- Invariant subspace is called <u>stable invariant subsp.</u> if it is associated with stable eigenvalues only. (i.e., if $\mathcal{A}_V$ is Hurwitz)

- The notion of invariant subspace will play a key role later on in the solution of Riccati equation

* <u>Unitary matrices and orthogonal complement</u>

- A square matrix $U \in \mathbb{R}^{n \times n}$ is called unitary if

$$U'U = I = UU'$$

- Columns of a unitary matrix constitute an orthonormal basis for $\mathbb{R}^n$. (The same for rows)

- A tall matrix D of a full column rank can always be „normalized" by post-multiplication

$$D_n := DT \quad , \text{ so that } \quad D_n' D_n = I$$

<u>What is the expression for T?</u>

$$T = (D'D)^{-1/2}$$

$$D_n' D_n = T' D' D T =$$

$$= (D'D)^{-1/2} D'D (D'D)^{-1/2} = I$$

- For any tall matrix D of a full column rank there exists $D_\perp$ such that

$$\left[D(D'D)^{-1/2} \quad D_\perp \right] \text{ is unitary.}$$

- $D_\perp$ is called orthogonal complement of D

<u>How to construct $D_\perp$?</u>

1. $\tilde{D}_\perp := \text{basis}(\text{Ker}(D'))$

2. $D_\perp := \tilde{D}_\perp (\tilde{D}_\perp' \tilde{D}_\perp)^{-1/2}$

§2.6 * Singular value decomposition

Given $A \in \mathbb{R}^{m \times n}$, there exist unitary matrices

$$U = [u_1, \dots, u_m] \in \mathbb{R}^{m \times m}$$
$$V = [v_1, \dots, v_n] \in \mathbb{R}^{n \times n}$$

such that $A = U \Sigma V^*$, where

$$\Sigma = \begin{bmatrix} \bar{\Sigma} & 0 \\ 0 & 0 \end{bmatrix}$$

$$\bar{\Sigma} = \mathrm{diag}(\sigma_1, \dots, \sigma_p)$$

with $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_p$, $p = \mathrm{rank}(A)$.

- Largest singular value : $\bar{\sigma}(A) = \sigma_1$
- Smallest singular value : $\underline{\sigma}(A) = \sigma_p$
- σ_i^2 are eigenvalues of AA' and $A'A$.
- Singular value decomp reveals principal input and output matrix directions
  - v_1 / v_n - highest/lowest gain input directions
  - u_1 / u_n - highest/lowest gain output directions

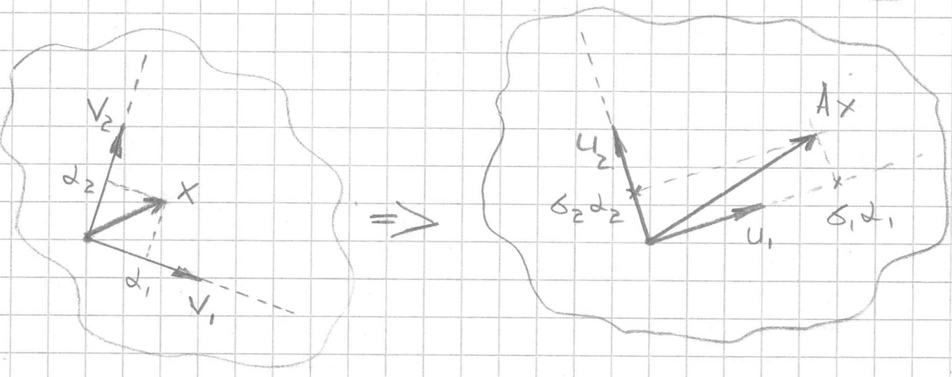
Calculate $A v_1$

$$\underline{A v_1} = U \Sigma V^* v_1 = U \Sigma \begin{bmatrix} v_1^* \\ \vdots \\ v_n^* \end{bmatrix} \cdot v_1$$

$$= U \Sigma \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = U \begin{bmatrix} \sigma_1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \underline{u_1 \sigma_1}$$

- The mapping $y = A x$ transforms unit ball in $\mathbb{R}^n$ into an elipsoid in $\mathbb{R}^m$ with semi-axes laying along u_i and having a length σ_i.

<u>How does this look for $A \in \mathbb{R}^{2\times 2}$?</u>



- Alternative definition for largest sing. value:

$$\bar{\sigma}(A) = \max_{\|x\|_2 = 1} \left(\| A x \|_2 \right) = \| A \|_2$$

- Some useful properties:

  - $A = \sum_{i=1}^{p} \sigma_i u_i v_i^*$
  - $\bar{\sigma}(A) = 1 / \underline{\sigma}(A^{-1})$
    (if A is invertible)
  - $\underline{\sigma}(A + B) - \underline{\sigma}(B) \leq \bar{\sigma}(A)$
  - $\underline{\sigma}(AB) \geq \underline{\sigma}(A)\,\underline{\sigma}(B)$

- The notion of singular values will play a key role later on for the definition of H_∞ system norm.

§2.3 * Schur complement and inversion formulae

- If $\mathcal{A}_{11}$ is invertible, we can define

$$\Delta_{11} := \mathcal{A}_{22} - \mathcal{A}_{21}\mathcal{A}_{11}^{-1}\mathcal{A}_{12}$$

and show

$$\begin{bmatrix} \mathcal{A}_{11} & \mathcal{A}_{12} \\ \mathcal{A}_{21} & \mathcal{A}_{22} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ \mathcal{A}_{21}\mathcal{A}_{11}^{-1} & 1 \end{bmatrix} \begin{bmatrix} \mathcal{A}_{11} & 0 \\ 0 & \Delta_{11} \end{bmatrix} \begin{bmatrix} 1 & \mathcal{A}_{11}^{-1}\mathcal{A}_{12} \\ 0 & 1 \end{bmatrix}$$

$$\begin{bmatrix} \mathcal{A}_{11} & \mathcal{A}_{12} \\ \mathcal{A}_{21} & \mathcal{A}_{22} \end{bmatrix}^{-1} = \begin{bmatrix} \mathcal{A}_{11}^{-1} + \mathcal{A}_{11}^{-1}\mathcal{A}_{12}\Delta_{11}^{-1}\mathcal{A}_{21}\mathcal{A}_{11}^{-1} & -\mathcal{A}_{11}^{-1}\mathcal{A}_{12}\Delta_{11}^{-1} \\ \Delta_{11}^{-1}\mathcal{A}_{21}\mathcal{A}_{11}^{-1} & \Delta_{11}^{-1} \end{bmatrix}$$

- If $\mathcal{A}_{22}$ is invertible, we can define

$$\Delta_{22} := \mathcal{A}_{11} - \mathcal{A}_{12}\mathcal{A}_{22}^{-1}\mathcal{A}_{21}$$

and show

$$\begin{bmatrix} \mathcal{A}_{11} & \mathcal{A}_{12} \\ \mathcal{A}_{21} & \mathcal{A}_{22} \end{bmatrix} = \begin{bmatrix} 1 & \mathcal{A}_{12}\mathcal{A}_{22}^{-1} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \Delta_{22} & 0 \\ 0 & \mathcal{A}_{22} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ \mathcal{A}_{22}^{-1}\mathcal{A}_{21} & 1 \end{bmatrix}$$

$$\begin{bmatrix} \mathcal{A}_{11} & \mathcal{A}_{12} \\ \mathcal{A}_{21} & \mathcal{A}_{22} \end{bmatrix}^{-1} = \begin{bmatrix} \Delta_{22}^{-1} & \Delta_{22}^{-1}\mathcal{A}_{12}\mathcal{A}_{22}^{-1} \\ -\mathcal{A}_{22}^{-1}\mathcal{A}_{21}\Delta_{22}^{-1} & \mathcal{A}_{22}^{-1} + \mathcal{A}_{22}^{-1}\mathcal{A}_{21}\Delta_{22}^{-1}\mathcal{A}_{12}\mathcal{A}_{22}^{-1} \end{bmatrix}$$

- If both $\mathcal{A}_{11}$ and $\mathcal{A}_{22}$ are invertible, then

$$\left(\mathcal{A}_{11} - \mathcal{A}_{12}\mathcal{A}_{22}^{-1}\mathcal{A}_{21}\right)^{-1}$$
$$= \mathcal{A}_{11}^{-1} + \mathcal{A}_{11}^{-1}\mathcal{A}_{12}\left(\mathcal{A}_{22} - \mathcal{A}_{21}\mathcal{A}_{11}^{-1}\mathcal{A}_{12}\right)^{-1}\mathcal{A}_{21}\mathcal{A}_{11}^{-1}$$

* Sylvester and Lyapunov equations

- Given $\mathcal{A} \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{m \times m}$ and $C \in \mathbb{R}^{n \times m}$ equation of a form

$$\mathcal{A} X + X \cdot B = C$$

is called Sylvester equation.

- Sylvester equation has a unique solution $X \in \mathbb{R}^{n \times m}$ iff

$$\forall i, j : \quad \lambda_i(\mathcal{A}) + \lambda_j(B) \neq 0$$

- Special case with $B = \mathcal{A}'$ is called Lyapunov eq.
- In Matlab use 'sylv' and 'lyap' functions.

* Algebraic Riccati equation

pp. 233-238

- Given $\mathcal{A} \in \mathbb{R}^{n \times n}$ and symmetric $R, Q \in \mathbb{R}^{n \times n}$ equation of a form:

$$\mathcal{A}' X + X \mathcal{A} + X R X + Q = 0$$

is called algebraic Riccati equation ($\mathcal{A}$RE).

- Solution $X \in \mathbb{R}^{n \times n}$, which renders $\mathcal{A} + RX$ Hurwitz is called stabilizing solution.
- Once stabilizing solution exists, it is unique.
- In Matlab use 'care' and 'are' functions.

* ARE and Hamiltonian matrix

- ARE is associated with the following Hamiltonian matrix:

$$H = \begin{bmatrix} A & R \\ -Q & -A' \end{bmatrix}$$

- The Hamiltonian matrix satisfies $H = -JH^*J^{-1}$, where

$$J = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}.$$

Namely, H is similar to $-H^*$ and λ is its eigenvalue iff $-\bar{\lambda}$ is.

- Any solution of ARE, X, satisfies:

$$\begin{bmatrix} A & R \\ -Q & -A' \end{bmatrix} \begin{bmatrix} 1 \\ X \end{bmatrix} = \begin{bmatrix} 1 \\ X \end{bmatrix} (A + RX)$$

Let's check:

1st row: $A + RX = A + RX$ (trivial)

2nd row: $-Q - A'X = XA + XRX$ (ARE)

- If X is a stabilizing solution, then $\operatorname{span}\begin{bmatrix} 1 \\ X \end{bmatrix}$ is a stable invariant subspace of H.

- Finding a stabilizing solution of ARE

$\Leftrightarrow$ Splitting the spectrum of H

More precisely, to find stabilizing solution:

1. Verify that there exists n-dimensional stable invariant subspace of H.
   (No eigenvalues of H on imaginary axis)

2. Construct a basis for n-dimensional stable invariant subspace of H and equally partition it as $\begin{bmatrix} X_1 \\ X_2 \end{bmatrix}$

3. Verify that X_1 is invertible

4. $X = X_2 X_1^{-1}$

- ARE has stabilizing solution iff

1. H has no eigenvalues on imaginary axis.

2. Its stable invariant subspace is complementary to $\operatorname{Im}\begin{bmatrix} 0 \\ I \end{bmatrix}$

- Stabilizing solution of ARE will be denoted by

$$X = \operatorname{Ric}(H) = \operatorname{Ric}\left(\begin{bmatrix} A & R \\ -Q & -A' \end{bmatrix}\right)$$

Vector spaces

- A linear vector space is a set $X = \{x\}$ with $\mathbb{F} = \mathbb{R}/\mathbb{C}$ and two operators:

$$+ : X \times X \to X$$
$$\cdot : \mathbb{F} \times X \to X$$

satisfying:

1. $x_1 + x_2 = x_2 + x_1$
2. $(x_1 + x_2) + x_3 = x_1 + (x_2 + x_3)$
3. $\exists 0 \in X : x + 0 = x, \ \forall x \in X$
4. $\forall x \in X : \exists -x \in X \mid x + (-x) = 0$
5. $(\lambda_1 + \lambda_2) x = \lambda_1 x + \lambda_2 x$
6. $\lambda (x_1 + x_2) = \lambda x_1 + \lambda x_2$
7. $\lambda_1 (\lambda_2 x) = (\lambda_1 \lambda_2) x$
8. $1 \cdot x = x$

- A normed vector space is a linear vector space equipped with norm $\|\cdot\|$ satisfying:

1. $\|x\| \geq 0$ and $\|x\| = 0 \iff x = 0$
2. $\|\lambda x\| = |\lambda| \, \|x\|$
3. $\|x_1 + x_2\| \leq \|x_1\| + \|x_2\|$

- Norm defines a „topology". The distance between x_1 and x_2 is $\|x_1 - x_2\|$

- This allows to define convergence of series and completeness.

- Complete normed space is called Banach space

- Sometimes the notion of „distance" is not enough and we need the notion of „angle". (provided by inner product)

- Inner product is a map $\langle \cdot, \cdot \rangle : X \times X \to \mathbb{F}$ satisfying:

  1. $\langle x, x \rangle \geq 0$ and $\langle x, x \rangle = 0 \iff x = 0$
  2. $\langle x_1, x_2 \rangle = \overline{\langle x_2, x_1 \rangle}$
  3. $\langle x_1 + x_2, x_3 \rangle = \langle x_1, x_3 \rangle + \langle x_2, x_3 \rangle$
  4. $\langle \lambda x_1, x_2 \rangle = \lambda \langle x_1, x_2 \rangle$

- It can be verified that $\sqrt{\langle x, x \rangle}$ satisfies all norm properties. $\|x\| = \sqrt{\langle x, x \rangle}$ is called <u>induced norm</u>.

- Complet linear space equipped with inner product and induced norm is called <u>Hilbert space</u>.

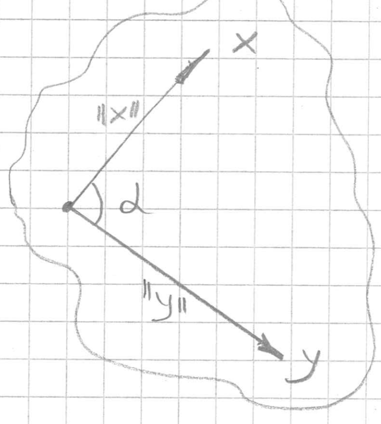
- Inner product provides abstract notion of „angle". We say that $x \perp y$ if $\langle x, y \rangle = 0$.

<u>$\mathbb{R}^2$ example:</u>

$$x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \quad y = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

$$\langle x, y \rangle = x_1 y_1 + x_2 y_2$$

$$\|x\| = \sqrt{x_1^2 + x_2^2}$$



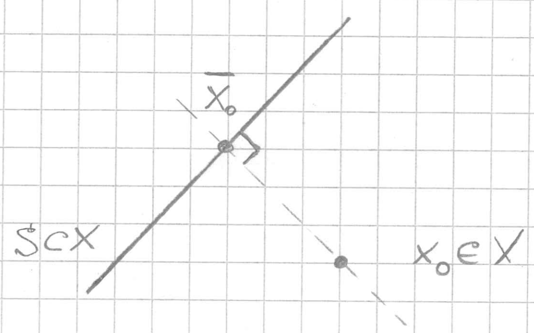
$$\cos(\alpha) = \frac{\langle x, y \rangle}{\|x\| \|y\|}$$

$$\langle x, y \rangle = 0 \iff \alpha = 90^\circ$$

- The key feature of Hilbert space is the „projection theorem".

Let $S \subset X$ be a subspace of a Hilbert space X.
Consider $x_0 \in X$, $x_0 \notin S$.

$$\bar{x}_0 = \underset{x \in S}{\operatorname{argmin}} \|x - x_0\| \iff (\bar{x}_0 - x_0) \perp S$$



This is a generalization of projection in $\mathbb{R}^2$ for abstract spaces.

§4.1

Signal spaces

- Signals in this course are vector functions $u : \mathbb{R} \to \mathbb{R}^n$, i.e., $u(t) = \begin{bmatrix} u_1(t) \\ \vdots \\ u_n(t) \end{bmatrix}$.

- The set of n-dimensional signals can be considered as a linear space. (with natural summation and scalar multiplication operators)

* L_2 signal space ($L_2(\mathbb{R})$)

$$L_2(\mathbb{R}) = \Big\{ u(t) \,\Big|\, \underbrace{\sqrt{\int_{-\infty}^{\infty} u'(t)u(t)\,dt}}_{\text{„signal energy"}} < \infty \Big\}$$

This is a Hilbert space with

$$<u(t), v(t)> := \int_{-\infty}^{\infty} v'(t)u(t)\,dt$$

$$\|u(t)\| := \sqrt{\int_{-\infty}^{\infty} u'(t)u(t)\,dt}$$

- In this course we will work with $L_2(\mathbb{R}^+)$ space

$$L_2(\mathbb{R}^+) = \{ u(t) \in L_2(\mathbb{R}) \,|\, u(t) = 0,\ \forall t < 0 \}$$

- How we can define „distance" and even a sort of „angle" between signals with finite energy.

§4.2

System spaces

- Dynamical system can be characterized by its impulse responce matrix

$$G(t) = \begin{bmatrix} g_{11}(t) & \dots & g_{1m}(t) \\ \vdots & & \\ g_{n1}(t) & \dots & g_{nm}(t) \end{bmatrix}$$

via the convolution integral.

- Applying Laplace transform, we move to the s-domain and get a transfer matrix

$$G(s) = \begin{bmatrix} g_{11}(s) & \dots & g_{1m}(s) \\ \vdots & & \\ g_{n1}(s) & \dots & g_{nm}(s) \end{bmatrix}$$

- The set of $n \times m$ systems can be considered as a linear space. (either in time or in s-domain under natural summation and scalar multiplication.)

* L_2 system space ($L_2(i\mathbb{R})$)

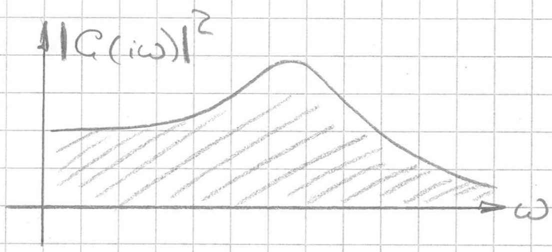
$$L_2(i\mathbb{R}) = \left\{ G(s) \,\middle|\, \int_{-\infty}^{\infty} tr\left(G^\sim(i\omega)G(i\omega)\right) d\omega < \infty \right\}$$

This is a Hilbert space with

$$< G(s), H(s) > = \frac{1}{2\pi} \int_{-\infty}^{\infty} tr\left(H^\sim(i\omega)G(i\omega)\right) d\omega$$

$$\| G(s) \|_2 = \underbrace{\sqrt{\frac{1}{2\pi} \int_{-\infty}^{\infty} tr\left(G^\sim(i\omega)G(i\omega)\right) d\omega}}_{\text{refered to as } L_2 \text{ norm of } G(s)}$$

- In SISO case the L_2 norm is proportional to the area under square of $|G(i\omega)|$.



* $\underline{H_2 \text{ and } H_2^{\perp} \text{ spaces}}$

$H_2(i\mathbb{R}) = \{ G(s) \in L_2 \text{ that are analytic in } \mathbb{C}^+ \}$

($\approx$ no unstable poles)

$H_2^{\perp}(i\mathbb{R}) = \{ G(s) \in L_2 \text{ that are analytic in } \mathbb{C}^- \}$

($\approx$ no stable poles)

It can be shown that these spaces are orthogonal complement.

- The L_2 norm of $G \in H_2$ is sometimes called the H_2 norm of this system.

- According to Parseval theorem, for $G \in H_2$

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} tr\left(G^{\sim}(i\omega)G(i\omega)\right) d\omega = \int_0^{\infty} tr\left(G'(t)G(t)\right) dt$$

This implies that

1. $H_2(i\mathbb{R}) \longleftrightarrow L_2(\mathbb{R}^+)$

2. H_2 system norm = „Energy" of impuls response.

* L_∞, H_∞ system spaces

$$L_\infty = \left\{ G(s) \;\middle|\; \sup_{\omega \in \mathbb{R}} \bar{\sigma}\big(G(i\omega)\big) < \infty \right\}$$

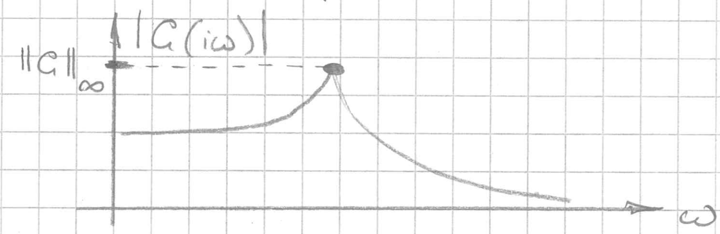
$$H_\infty = \left\{ G \in L_\infty \text{ that are analytic in } \mathbb{C}^+ \right\}$$

($\approx$ no unstable poles)

These are Banach (but not Hilbert) spaces with

$$\underbrace{\| G(s) \|_\infty = \sup_{\omega \in \mathbb{R}} \bar{\sigma}\big(G(i\omega)\big)}_{\text{refered to as } L_\infty \text{ (or } H_\infty\text{) system norm.}}$$

- In SISO case L_∞ norm is a peak of $|G(i\omega)|$



- It can be shown that $G \in H_\infty$ iff it maps $L_2(\mathbb{R}^+)$ into $L_2(\mathbb{R}^+)$. Namely, <u>H_∞ is a space of causal and stable systems.</u>

- H_∞ is the induced operator norm, i.e.,

$$\| G(s) \|_\infty = \sup_{\|u\|_2 = 1} \| Gu \|_2 \qquad (\text{for } G \in H_\infty)$$

So for $G \in H_\infty$, $\|G(s)\|_\infty$ is the „worst case" energy gain of the system.

What did we study today?

- Some topics in algebra
  - singular value decomposition
  - Riccati equation
- Banach and Hilbert spaces (abstract)
- Spaces of signals and systems.
  - norms of signals and systems provide the notion of „distance".