5.10.2023

Q: Kronecker Product?

→ of 2 Matrices

$\underline{\underline{A}} \in \mathbb{R}^{m\times n}$ $\quad$ $\underline{\underline{B}} \in \mathbb{R}^{l\times k}$

$m, n, l, k \in \mathbb{N}$

$\underline{\underline{A}} \otimes \underline{\underline{B}} \in \mathbb{R}^{ml \times nk}$

$$A_{11}\begin{bmatrix} B_{11} & B_{12} \cdots & B_{1k} \\ \vdots & & \\ B_{l1} & B_{l2} \cdots & B_{lk} \end{bmatrix}$$

$$\begin{bmatrix} \boxed{A_{11}\underline{\underline{B}}} & A_{12}\underline{\underline{B}} & \cdots & A_{1n}\underline{\underline{B}} \\ \vdots & & & \\ A_{m1}\underline{\underline{B}} & A_{m2}\underline{\underline{B}} & \cdots & A_{mm}\underline{\underline{B}} \end{bmatrix}$$

Example

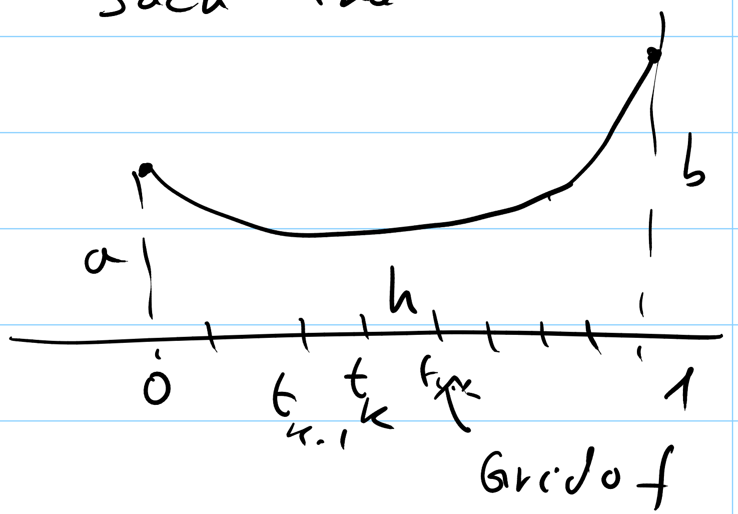
$$I_m \otimes \underline{\underline{A}} = \begin{bmatrix} \underline{\underline{A}} & 0 & \cdots & 0 \\ 0 & \underline{\underline{A}} & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & & \underline{\underline{A}} \end{bmatrix}$$

$$\underline{\underline{A}} \otimes \underline{I} = \begin{bmatrix} \boxed{\begin{matrix} a_{11} & & 0 \\ & a_{11} \ddots & \\ 0 & & a_{11} \end{matrix}} & \boxed{\begin{matrix} a_{12} & & \\ & a_{12} \ddots & \\ & & a_{12} \end{matrix}} & \cdots \\ \vdots & \ddots & \ddots \\ \vdots & & \boxed{a_{ij}\underline{I}} \end{bmatrix}$$

Use: Partial Differential equations

1D Find $u : [0,1] \to \mathbb{R}$ such that

$$\begin{cases} \frac{d^2 u}{dt} = F(t) \\ u(0) = a \\ u(1) = b \end{cases}$$



Grid of points

equidistant

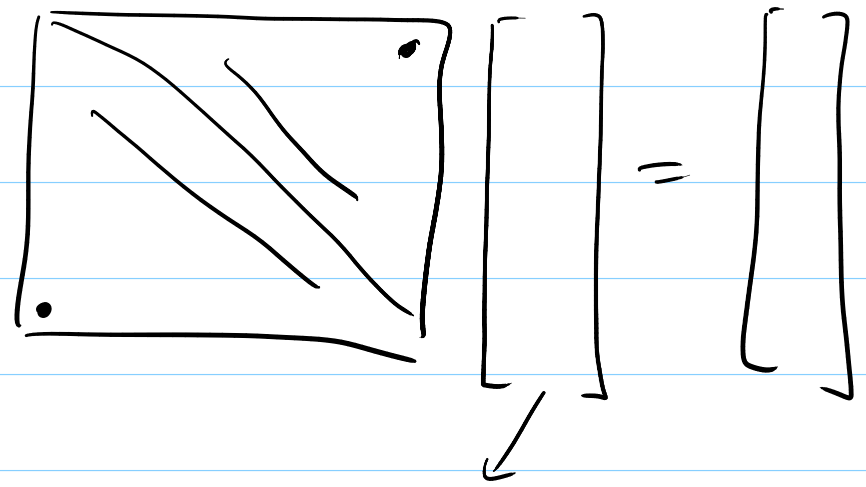
replace $\frac{d^2 u}{dt}$ with a quotient

$$\frac{du}{dt}(t_k) \approx \frac{u(t_{k+1}) - u(t_{k-1})}{2h}$$

$$\frac{d^2 u}{dt^2}(t_k) \approx \frac{u(t_{k+1}) - 2u(t_k) + u(t_{k+1})}{2h}$$

"Finite difference method" FD

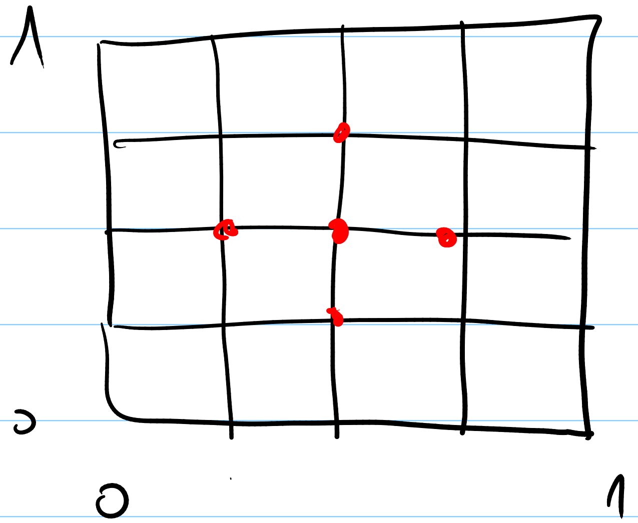
$\Rightarrow$ replace the diff. eq. by a lin. System



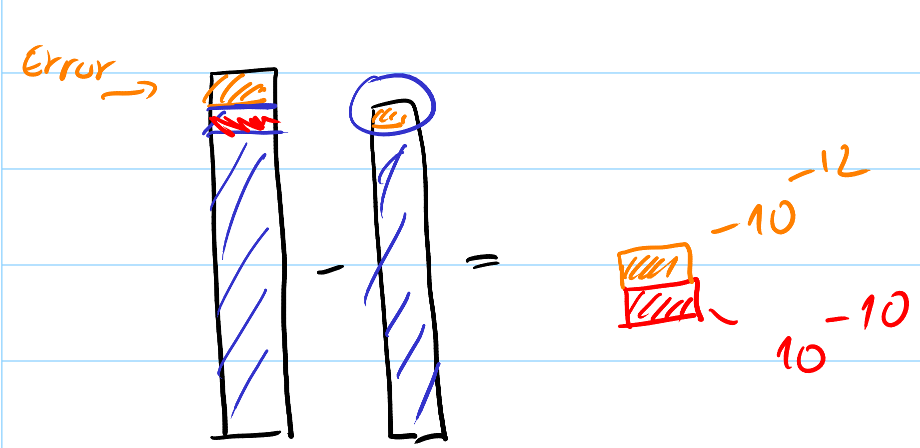
$u_k \approx u(t_k)$

$\underline{\underline{C}}\,\underline{u} = \underline{b}$

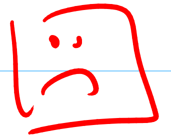
2D:

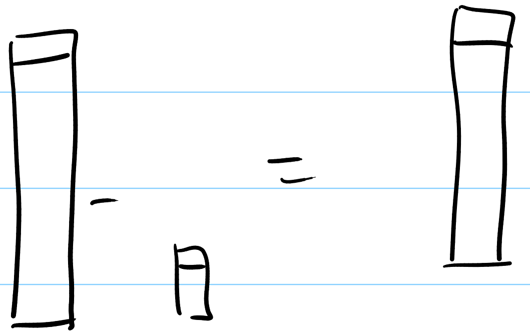


$\underline{\underline{C}} = \underline{\underline{A}} \otimes \underline{\underline{B}}$



$\frac{1}{10^2}$ or 10^2





$u \in$ Space of smooth functions
V ↘ ?

$P_n \subset V$ infinite dimensional C^∞

look for $u_n(t)$ = polynomial of degree $\le n-1$

$u_n \in P_n \subset V$

Replace:

find $u \in V$ s.t. $\begin{cases} \frac{d^2u}{dt^2} = f(t) \\ u(0) = a \\ u(1) = b \end{cases}$

with find $u_n \in P_n$ s.t. $\Big\{$ "same"

Basis in P_n: e.g. monomials

$$u_n(t) = a_0 + a_1 t + \dots + a_{n-1} t^{n-1}$$

$$\frac{du_n}{dt} = a_1 + 2a_2 t + \dots + \underset{(n-1)}{a_{n-1}} t^{n-1} \quad \Big\{ \Rightarrow$$

$$\frac{d^2u}{dt^2} = \dots$$

replace $\Rightarrow$ $a_0, a_1, \dots, a_{n-1}$

(LA).

12.10.2023

(Q2.3.2.21.H) [Schur complement] Consider the following block partitioning of a matrix $\mathbf{A} \in \mathbb{R}^{n+m,n+m}$

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix}, \quad \begin{matrix} \mathbf{A}_{11} \in \mathbb{R}^{n,n}, & \mathbf{A}_{12} \in \mathbb{R}^{n,m}, \\ \mathbf{A}_{21} \in \mathbb{R}^{m,n}, & \mathbf{A}_{22} \in \mathbb{R}^{m,m}. \end{matrix}$$

We assume that $\mathbf{A}_{11}$ is *regular*, which renders the **Schur complement**

$$\mathbf{S} := \mathbf{A}_{22} - \mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12} \in \mathbb{R}^{m,m}$$

well-defined.

Show that **$\mathbf{A}$ is singular, if and only if $\mathbf{S}$ is singular**:

$$\mathcal{N}(\mathbf{A}) \neq \{\mathbf{0}\} \iff \mathcal{N}(\mathbf{S}) \neq \{\mathbf{0}\}.$$

Proof

$\Rightarrow$: [Assume $\mathcal{N}(A) \neq \{0\}$ and prove $\mathcal{N}(S) \neq \{0\}$]

↓

there is $v \in \mathcal{N}(A)$, $v \neq 0$

$$0 \neq v = \begin{bmatrix} x \\ y \end{bmatrix} \text{ such that } Av = 0$$

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \iff$$

↑ A_{11} is regular

$$L \begin{bmatrix} A_{11} & A_{12} \\ 0 & S \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \iff$$

↓ L^{-1}

$$\left\{ \begin{aligned} A_{11} x + A_{12} y &= 0 \\ S y &= 0 \end{aligned} \right.$$

x solves this for given y

if $y \neq 0$ then we found $y \in \mathcal{N}(S)$, $y \neq 0$

[if $y = 0 \Rightarrow A_{11} x = 0$, A_{11} invertible $\Big\} \Rightarrow x = 0 \Rightarrow v = \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ ↯

This means $y \neq 0$]

$\Leftarrow$: Same $0 \neq y \in \mathcal{N}(S)$ and construct $\boxed{x = -A_{11}^{-1} A_{12} y}$

$\Rightarrow v = \begin{bmatrix} x \\ y \end{bmatrix}$ fullfills $Av = 0 \Rightarrow 0 \neq v \in \mathcal{N}(A)$.

e.g. $$\begin{array}{ll} r \rightarrow & \\ & \\ & \\ & \\ & \\ & \\ & \\ & \\ s \rightarrow & \end{array} \mathbf{A} = \begin{bmatrix} * & * & * & * & * & * & * & * & * \\ & & & & & & & * & \\ & & & & & & * & & \\ & & & & & * & & & \\ & & & & * & \nwarrow d. & & & \\ & & & * & & & & & \\ & & * & & & & & & \\ & * & & & & & & & \\ * & * & * & * & * & * & * & * & * \end{bmatrix}.$$

1. Outline an efficient algorithm for solving a linear system of equations $\mathbf{A}\mathbf{x} = \mathbf{b}$, $\mathbf{b} \in \mathbb{R}^{n,n}$.
2. Give a sufficient and necessary condition for $\mathbf{A}$ being regular/invertible.

$$\begin{bmatrix} r_1 & r_2 & r_3 \cdots & & r_{n-1} & r_n \\ & & & & d_{n-1} & \\ & & & d_{n-2} & & \\ & & d_3 & & & \\ & d_2 & & & & \\ s_1 & s_2 & s_3 & \cdots\cdots & s_{n-1} & s_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ \\ b_n \end{bmatrix}$$

$$\left\{ \begin{aligned} d_{n-1} x_{n-1} &= b_2 \\ d_{n-2} x_{n-2} &= b_3 \\ &\vdots \\ d_2 x_2 &= b_{n-1} \end{aligned} \right. \Rightarrow \quad \begin{aligned} x_{n-1} &= \frac{b_2}{d_{n-1}} \\ &\vdots \\ x_2 &= \frac{b_{n-1}}{d_2} \end{aligned}$$

for $k = 2, 3, \dots, n-1$: $x_k = \dfrac{b_{n+1-k}}{d_k}$ $\quad O(n-2)$

$$\left\{ \begin{aligned} r_1 x_1 + \boxed{r_2 x_2 + \dots + r_{n-1} x_{n-1}} + r_n x_n &= b_1 \\ s_1 x_1 + \boxed{s_2 x_2 + \dots + s_{n-1} x_{n-1}} + s_n x_n &= b_n \end{aligned} \right.$$

known now!

2 eq. with 2 unknowns. $\rightarrow O(2)$

$\rightarrow O(n)$ Operations !

(Q2.6.0.25.G) [A banded linear system] Sketch an efficient algorithm for the solution of the $n \times n$ linear system of equations

$$\begin{bmatrix} 1 & 0 & \cdots & \cdots & 0 & 1 \\ 1 & 1 & 0 & \cdots & & 0 \\ 0 & \ddots & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & 1 & 1 & 0 \\ 0 & \cdots & \cdots & 0 & 1 & 1 \end{bmatrix} \mathbf{x} = \mathbf{b} \in \mathbb{R}^n .$$

When will a solution exists for every right-hand side vector $\mathbf{b}$?

$$\begin{bmatrix} 1 & 0 & \cdots & \cdots & 0 & 1 \\ 1 & 1 & 0 & \cdots & & 0 \\ 0 & \ddots & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & 1 & 1 & 0 \\ 0 & \cdots & \cdots & 0 & 1 & 1 \end{bmatrix}$$

$$= \underbrace{\begin{bmatrix} 1 & 0 & \cdots & \cdots & 0 & 0 \\ 1 & 1 & 0 & \cdots & & 0 \\ 0 & \ddots & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & 1 & 1 & 0 \\ 0 & \cdots & \cdots & 0 & 1 & 1 \end{bmatrix}}_{O(n)} \underbrace{\begin{bmatrix} 1 & 0 & \cdots & & \cdots & 0 & 1 \\ 0 & 1 & \ddots & & & \vdots & -1 \\ \vdots & \ddots & \ddots & & & & 1 \\ & & & \ddots & \ddots & \vdots & -1 \\ & & & \ddots & \ddots & 0 & \vdots \\ \vdots & & & & \ddots & \ddots & (-1)^{n-2} \\ 0 & \cdots & & & \cdots & 0 & \boxed{1+(-1)^{n-1}} \end{bmatrix}}_{O(n)} .$$

(Block-Gauss as for Arrow!)

27.1.c

1) Build Part with $\underline{I}$ in $\underline{B}$:

```
for i = 0, 1, .., n-1
    store (i, i, 1)
```

2) Build Part with $\underline{A}$ in $\underline{B}$

$$\begin{array}{c|c} I & A \\ \hline A^T & 0 \end{array}$$

```
for i = 0, 1, .., n-1
  for j = 0, 1, .., n-1
      store (n+i, j, $a_{ij}$)
```

3) Buid Part with $\underline{A}^T$ in $\underline{B}$

```
for i = 0, 1, .., n-1
    for j = 0, 1, .., n-1
        store (i, n+j, $a_{ji}$)
```

Better: replace (2)&(3) by.

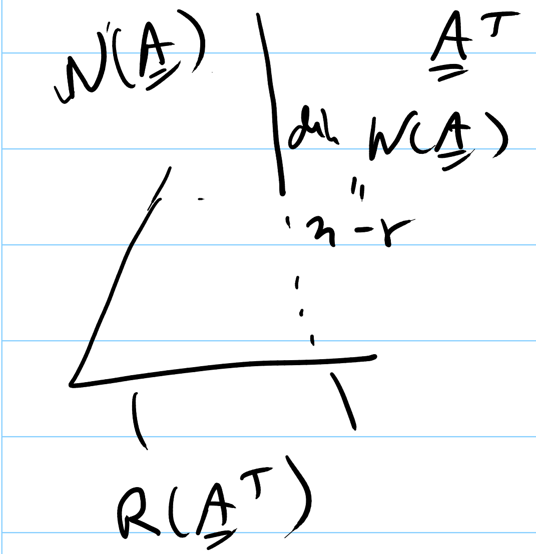
```
for $i = 0, 1, \dots, n-1$
  for $j = 0, 1, \dots, n-1$
        store($n+i$, $j$, $a_{ij}$)
        store($i$, $n+j$, $a_{ji}$)
```

"Q+A NCS" 19.10.2023

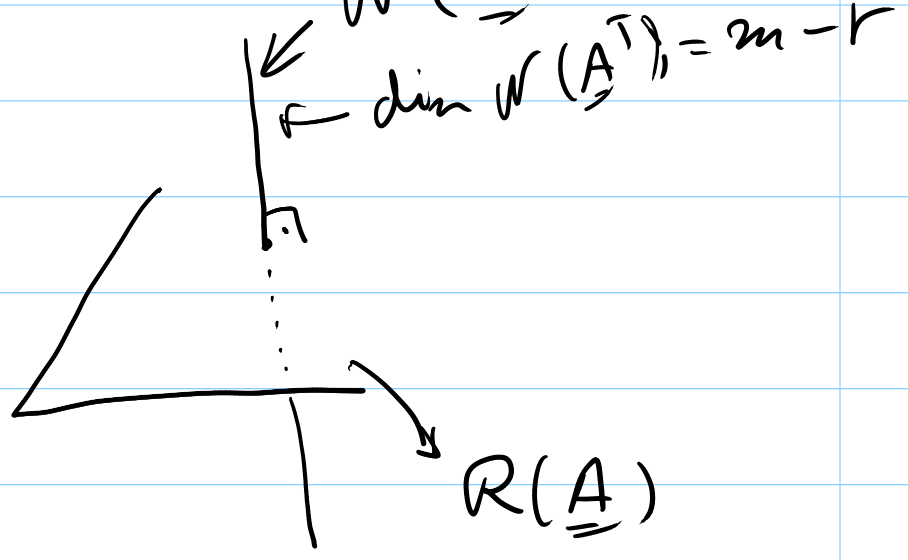
① Why are the solutions of the normal eq. $\underline{x}_0 + N(\underline{\underline{A}})$. Why $\underline{x}^+ \in N(\underline{\underline{A}})^\perp$

$\underline{\underline{A}}$ m rows, n columns

$$\begin{tikzcd} \mathbb{R}^n \arrow[r, bend left] & \mathbb{R}^m \arrow[l, bend left, "\underline{\underline{A}}^T"] \end{tikzcd}$$

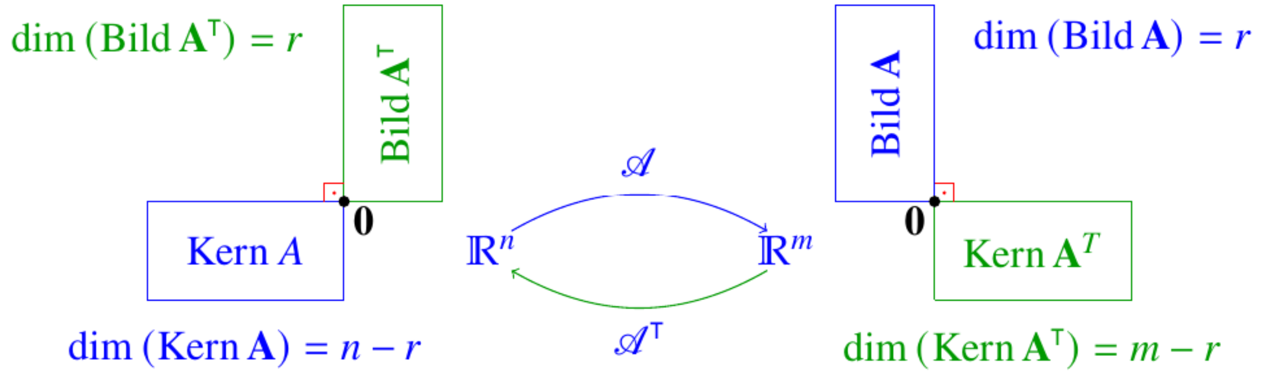


$\dim R(\underline{\underline{A}}^T) = r$



$\dim R(\underline{\underline{A}}) = \text{Rank}(\underline{\underline{A}}) = r$

$\underline{\underline{A}} \in \mathbb{R}^{3\times3}$, $\dim N(\underline{\underline{A}}) = 1$

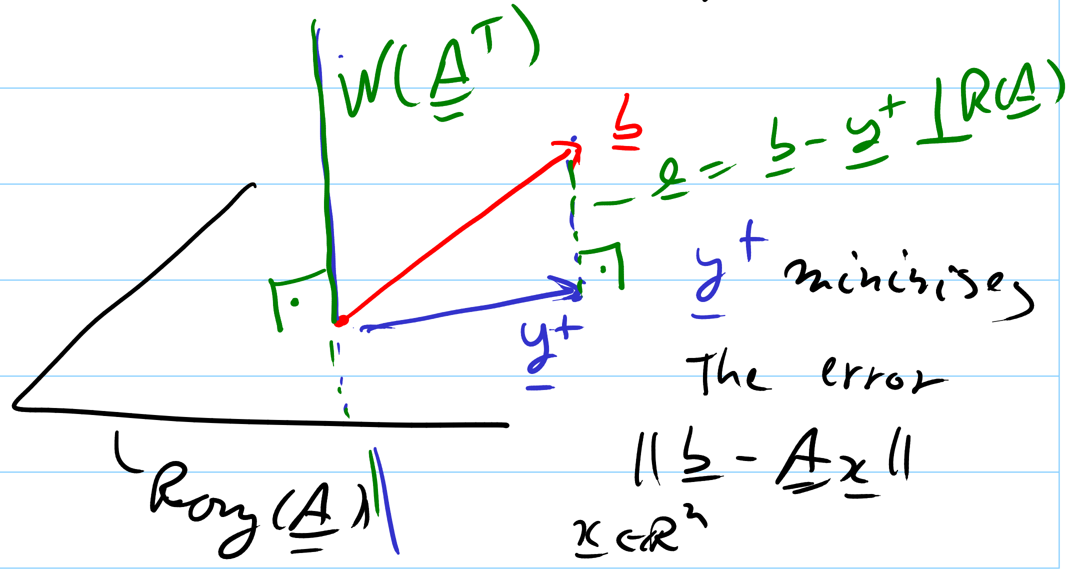


$\underline{\underline{A}}\,\underline{x} = \underline{b}$ $\quad \underline{\underline{A}} \in \mathbb{R}^{m\times n}$, $\underline{b} \in \mathbb{R}^m$

$\underline{x} \in \mathbb{R}^n$

$\underline{\underline{A}}^T \underline{\underline{A}}\,\underline{x} = \underline{\underline{A}}^T \underline{b}$

Rem 1) $R(\underline{\underline{A}})$ = lin. subspace of $\mathbb{R}^m$



$\underline{y}^+$ minimises

The error

$\|\underline{b} - \underline{\underline{A}}\,\underline{x}\|$

$x \in \mathbb{R}^n$

$$\underline{\underline{A}}^T(\underline{\underline{A}}\,\underline{x}-\underline{b}) = y^+ = \underline{\underline{A}}\,\underline{x}^+$$

The solution x^+ of $\underline{\underline{A}}^T\underline{\underline{A}}\,\underline{x} = \underline{\underline{A}}^T\underline{b}$

has the property:

$$\|\underline{\underline{A}}\,\underline{x}^+ - \underline{b}\| = \min_{\underline{x}\in\mathbb{R}^n} \|\underline{\underline{A}}\,\underline{x} - \underline{b}\|$$

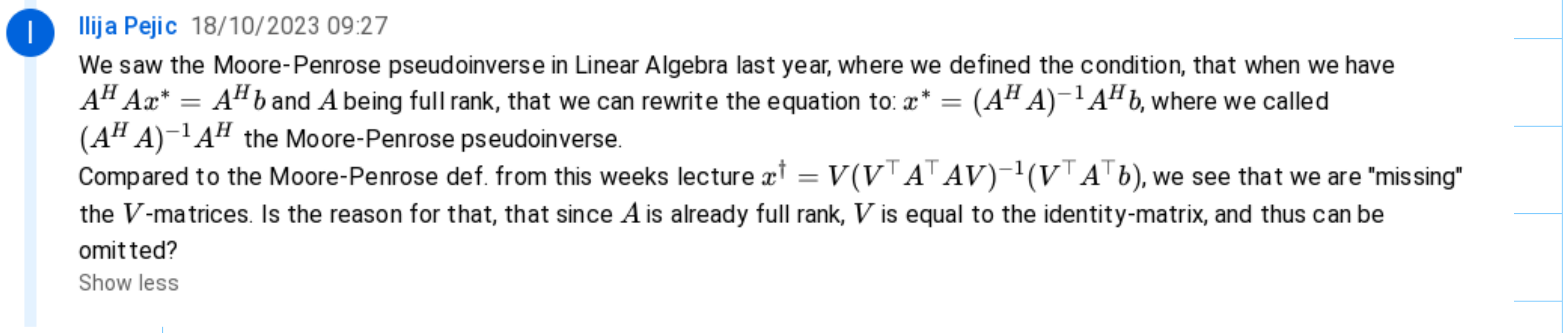
every other $x \in x^+ + N(\underline{\underline{A}})$ has the same property.

$x = x^+ + x_0$ with $x_0 \in N(\underline{\underline{A}}) \Rightarrow$

$$\underline{\underline{A}}\,\underline{x} = \underline{\underline{A}}\,\underline{x}^+ + \underline{\underline{A}}\,x_0 = \underline{\underline{A}}\,\underline{x}^+ + \underline{0} = \underline{\underline{A}}\,x^+$$

$e = \underline{b} - \underline{y}^+ \in N(\underline{\underline{A}}^+)$

because $N(\underline{\underline{A}}^T) \perp R(\underline{\underline{A}})$.



Ilija Pejic 18/10/2023 09:27

We saw the Moore-Penrose pseudoinverse in Linear Algebra last year, where we defined the condition, that when we have $A^HAx^* = A^Hb$ and A being full rank, that we can rewrite the equation to: $x^* = (A^HA)^{-1}A^Hb$, where we called $(A^HA)^{-1}A^H$ the Moore-Penrose pseudoinverse.

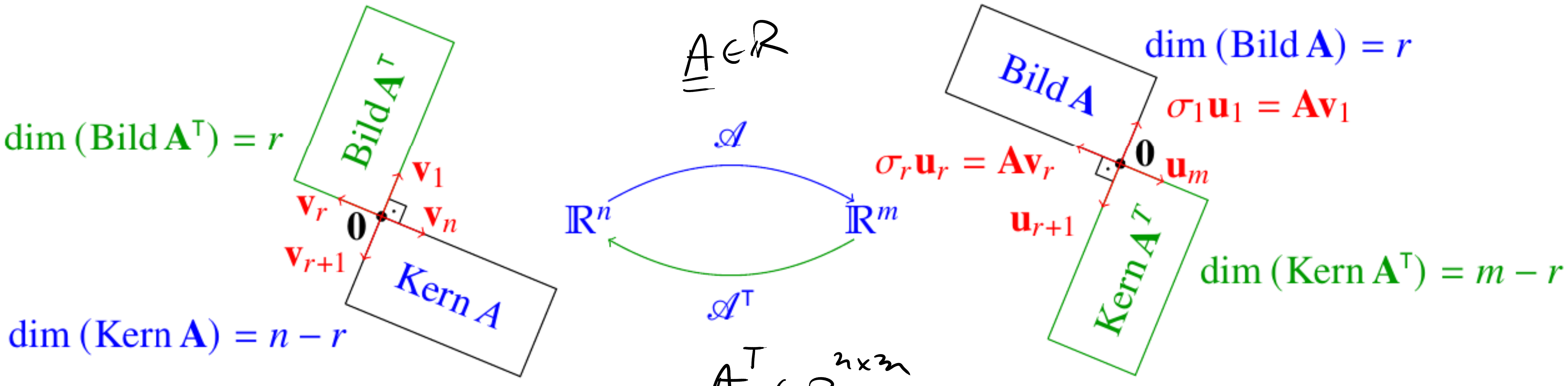
Compared to the Moore-Penrose def. from this weeks lecture $x^\dagger = V(V^\top A^\top AV)^{-1}(V^\top A^\top b)$, we see that we are "missing" the V-matrices. Is the reason for that, that since A is already full rank, V is equal to the identity-matrix, and thus can be omitted?

Show less

If $\underline{\underline{A}}$ is full rank $\Rightarrow$ $\underline{\underline{A}}^T\underline{\underline{A}}$ symmetric pos. def.

$\Rightarrow$ normal eq. has an unique sol!

What to do if $\underline{\underline{A}}$ is not full rank?

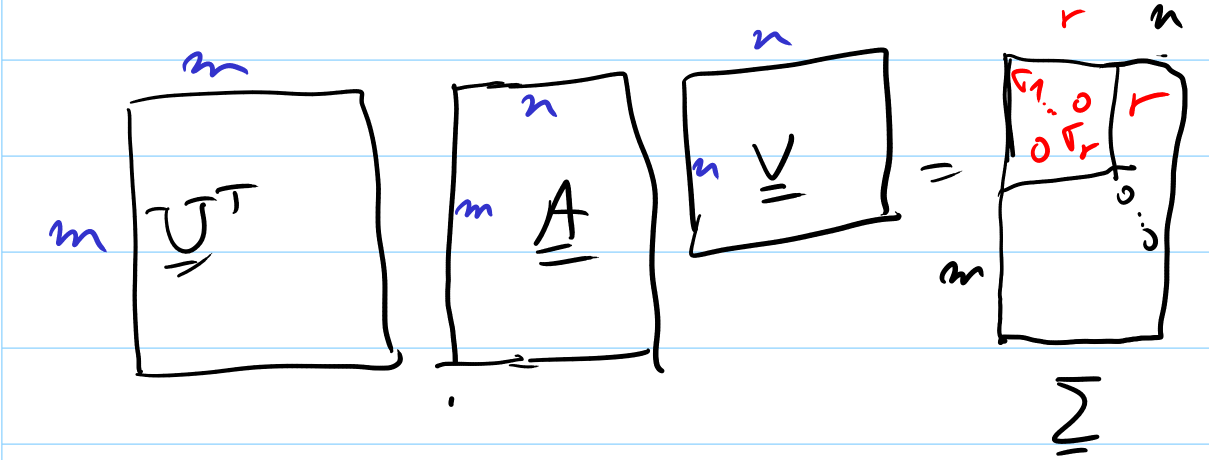


$\underline{A} \in \mathbb{R}^{m\times n}$

$\underline{A}^T \in \mathbb{R}^{n\times m}$

$$\underbrace{\underline{v}_1, \underline{v}_2, \dots, \underline{v}_r}_{\text{span } R(\underline{A}^T)}, \underbrace{\underline{v}_{r+1}, \dots, \underline{v}_n}_{\text{span } N(\underline{A})} \text{ ONB in } \mathbb{R}^n$$

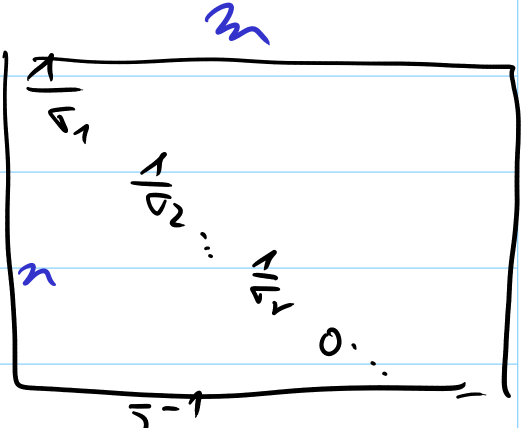
$$\underbrace{\underline{u}_1, \dots, \underline{u}_r}_{\text{span } R(\underline{A})}, \underbrace{\underline{u}_{r+1}, \dots, \underline{u}_m}_{\text{span } N(\underline{A}^T)} \text{ ONB in } \mathbb{R}^m$$



$\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r > 0$

$\underline{A} = \underline{U}\,\underline{\Sigma}\,\underline{V}^T$

Moore Penrose-Inverse: $\underline{A}^+ = \underline{V}$



$\underline{A}^+ \underline{A} = \underline{I}$

③ LU-Decomposition & pivoting

Suppose we do not have any pivoting

$$\underline{\underline{A}} = \underline{\underline{L}}\,\underline{\underline{U}}$$

⇒ LU-Decomposition is unique!

L_1, L_2 lower triangular with 1 in main diag.

⇒ $L_1 L_2$ is also —"— ______

U_1, U_2 upper triangular ⇒

$U_1 U_2$ —"— ______

$\underline{\underline{U}}^{-T}$

<u>Assume</u> $\underline{\underline{L}}_1 \underline{\underline{U}}_1 = \underline{\underline{L}}_2 \underline{\underline{U}}_2 \quad \Big|\, L_1^{-1} \quad\quad \Big|\, \underline{\underline{U}}_2^{-1}$ ⇒

$$\boxed{\begin{matrix} & * \\ 0 & \end{matrix}} = \underline{\underline{U}}_1 \underline{\underline{U}}_2^{-1} = \underline{\underline{L}}_1^{-1} \underline{\underline{L}}_2 = \boxed{\begin{matrix} 1 & 0 \\ * & 1 \;\ddots \end{matrix}} \Rightarrow$$

$$\Rightarrow \underline{\underline{L}}_1^{-1} \underline{\underline{L}}_2 = \underline{\underline{I}} \Rightarrow \underline{\underline{L}}_2 = \underline{\underline{L}}_1!$$

$$\underline{\underline{U}}_1 \underline{\underline{U}}_2^{-1} = \underline{\underline{I}} \Rightarrow \underline{\underline{U}}_1 = \underline{\underline{U}}_2$$

Hence LU-Decomposition is unique

(till a permutation!)

④

$$\operatorname{argmin}\left(\|\underline{\underline{A}}\,\underline{x} - \underline{b}\|^2 + \boxed{\lambda \|\underline{x}\|^2}\right)$$

reformulate as a standard lsp:

$$\begin{array}{cc} & n \\ 2n & \underbrace{\begin{bmatrix} \underline{\underline{A}} \\ \sqrt{\lambda}\, I \end{bmatrix}}_{\underline{\underline{M}}} \end{array} \underline{x} = \underbrace{\begin{bmatrix} b \\ 0 \end{bmatrix}}_{\tilde{\underline{b}}}$$

$$\operatorname{argmin} \|\underline{\underline{M}}\,\underline{z} - \underline{\underline{b}}\|^2 \qquad \underline{z} \in \mathbb{R}^{2n}$$

$$\operatorname{Rang}(\underline{\underline{M}}) = n$$

$$\underline{\underline{M}}^T \underline{\underline{M}}\, \underline{z} = \underline{\underline{M}}^T \underline{\underline{\tilde{b}}}$$

$$\begin{bmatrix} \underline{\underline{A}}^T & \sqrt{\lambda}\,\underline{I} \end{bmatrix} \begin{bmatrix} A \\ \sqrt{\lambda}\,\underline{I} \end{bmatrix} x = \begin{bmatrix} \underline{\underline{A}}^+ & \sqrt{\lambda}\,\underline{I} \end{bmatrix} \begin{bmatrix} b \\ 0 \end{bmatrix}$$

$$\left(\underline{\underline{A}}^T \underline{\underline{A}} + \lambda \underline{\underline{I}} \right) \underline{x} = \underline{\underline{A}}^T \underline{b} + \underset{0}{\boxed{\sqrt{\lambda}\,\underline{I} \cdot 0}}$$

26.10.2023

① Constrained least squares

$$L(y,m) = \boxed{\frac{1}{2}}\|Ay-b\|^2 + m^T(Cx-d)$$

$$\frac{\partial}{\partial y_1} L(y,m) = \frac{\partial}{\partial y_1}\left[\frac{1}{2}\sum_{j=1}^{n}\left((Ay)_j - b_j\right)^2\right] + 0$$

$$= \cancel{\frac{1}{2}}\sum_{j=1}^{n} \cancel{2}\left((Ay)_j - b_j\right)\frac{\partial}{\partial y_1}(Ay)_j$$

⇒ $\frac{1}{2}$ is only for math. beauty/simplicity.

Min. is attained for the same y^*.

② Givens Rotation: choice of the direction / order of operations.

→ 2 choices are possible for one rotation

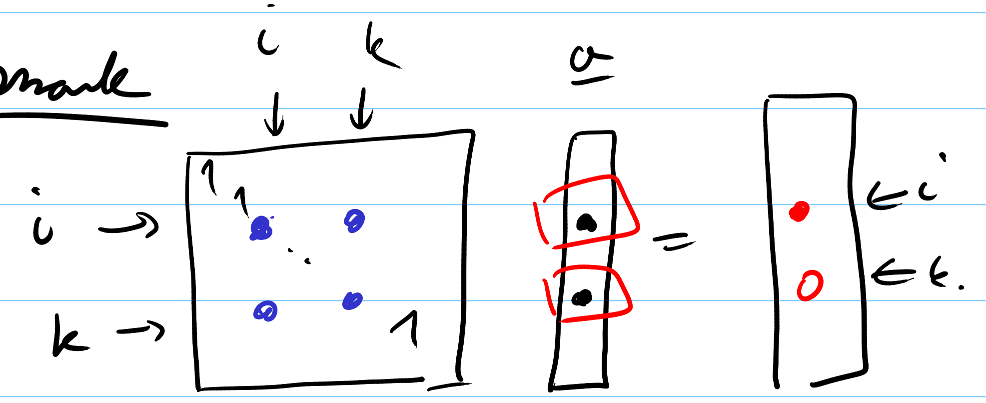
⇒ choose the one with less roundoff

$\left(\pm \frac{a_1}{\|a_0\|}\right)$ → see in the lecture document for the right choice

→ store only c or s or t at the position of A that gets 0.

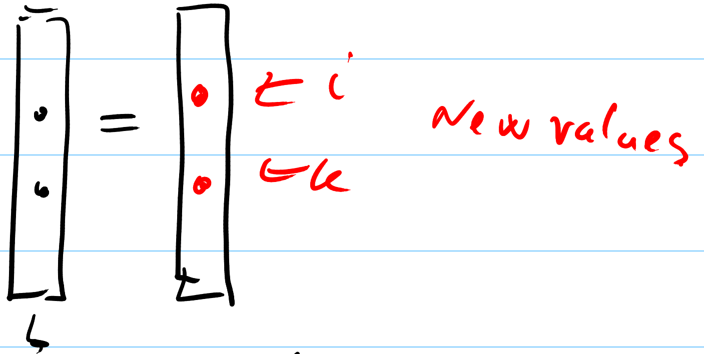
⇒ "economical" enough. ☺

Remark



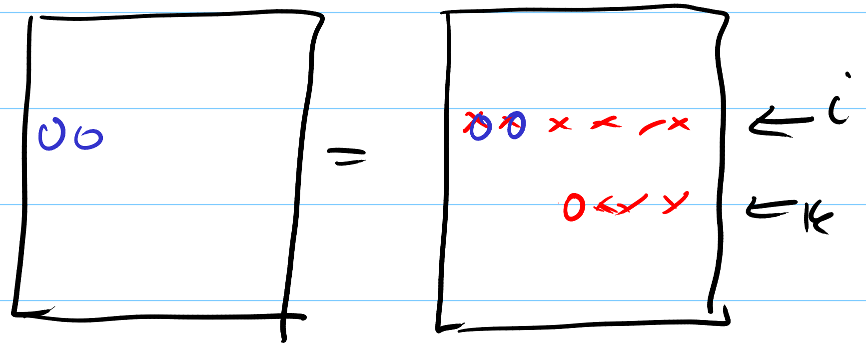
$$\begin{matrix} c & s \\ -s & c \end{matrix} \qquad G_{ik} = G_{ik}[a_i, a_k]$$

$\underline{G}_{ik}[a_i, a_k]$



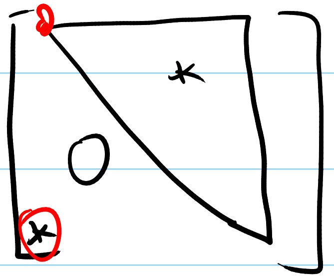
in top-down QR-decomposition:

$\underline{G}_{ik}[a_i, a_k]$

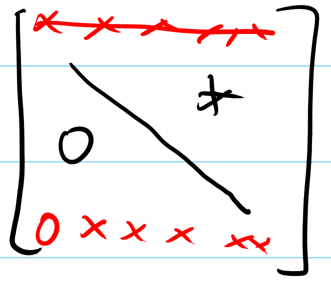


③ Discuss Q 3.3.3.28 B

$A =$

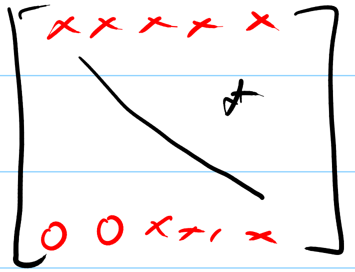


$\underline{Q}_{1n}[a_{11}, a_{n1}]\,\underline{A} =$



$\Rightarrow$ use $\underline{Q}_{1n}[a^{(1)}_{12}, a^{(1)}_{n2}]$

$\Rightarrow$

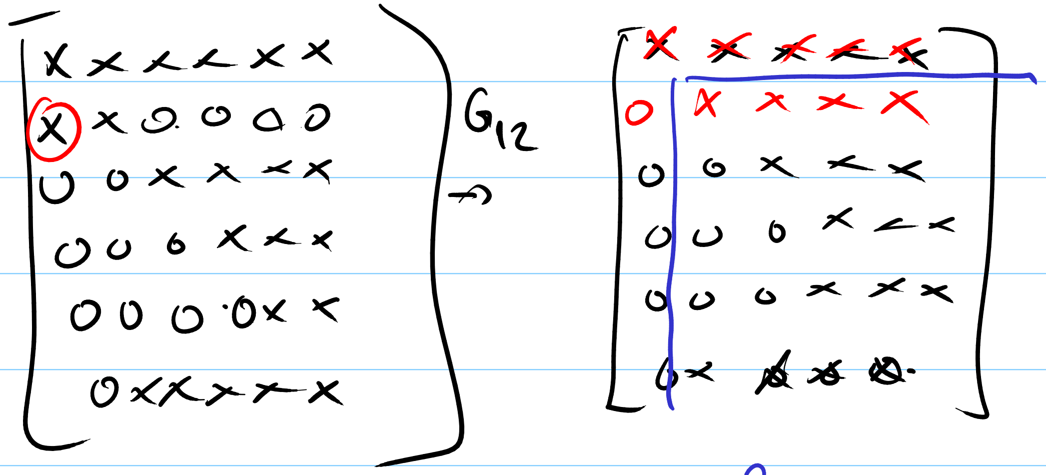


$\Rightarrow$

$$\underline{Q}_{1n}^{(n-1)} \dots \underline{Q}_{12}^{(1)} \underline{Q}_{1n} \underline{A} \Rightarrow \underline{\underline{R}}$$

Another way:

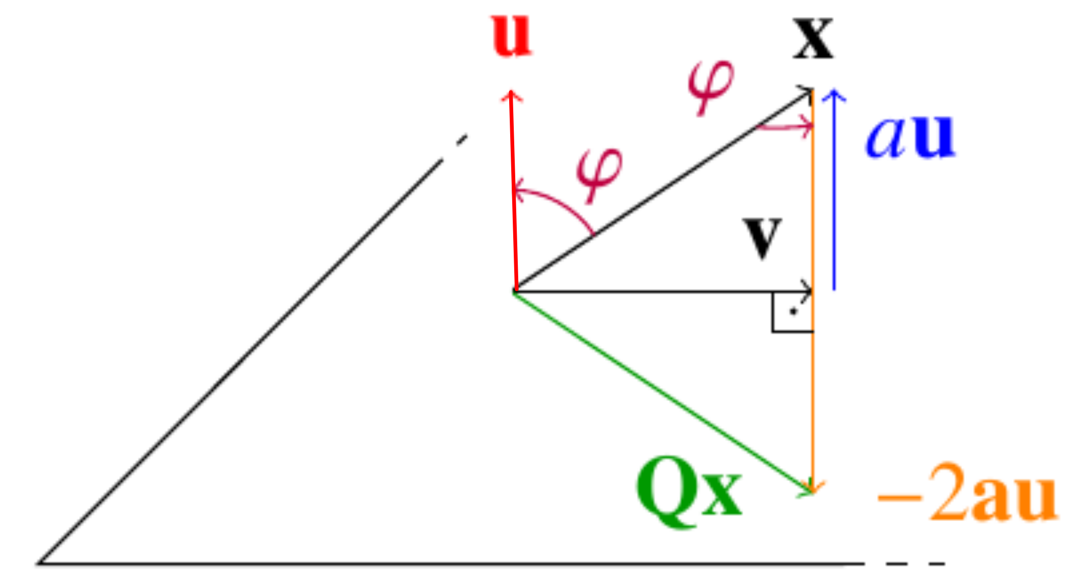
swap 2nd & last line.



exchg & G_{12} etc. ☹

not cheaper!

④ Householder Matrix



Given a hyper-plane (for the reflection)

|||

given the orthonormal vector $\underline{u}$

$\underline{u} \perp$ Plane & $\|\underline{u}\| = 1$

$\underline{\underline{Q}}\,\underline{x}$ = reflexion of $\underline{x}$ w.r.t. the plane.

Project $\underline{x}$ onto the plane => $\underline{v}$

"Direction of Projection": $-a\underline{u}$

The result: goes further in Direction $-\underline{u}$

$-2a\underline{u}$

$$\underline{\underline{Q}}\,\underline{x} = \underline{x} + (-2a\underline{u})$$

$$\underline{u}^T \underline{x} = \|u\| \cdot \|\underline{x}\| \cdot \cos\varphi = 1 \, \|x\| \, \frac{a}{\|x\|} = a$$

$$\Rightarrow \boxed{a = \underline{u}^T \underline{x}}$$

Hence $\underline{\underline{Q}}\,\underline{x} = \underline{x} - 2(\underline{u}^T \underline{x})\underline{u}$ ← because $\underline{u}^T\underline{x}$ is a scalar.

$$= \underline{x} - 2\,\underline{u}\,(\underline{u}^T \underline{x})$$

associativity

$$= \underline{x} - 2(\underline{u}\,\underline{u}^T)\underline{x}$$

$$= (\underline{\underline{I}} - 2\underline{u}\,\underline{u}^T)\,\underline{x} \Rightarrow \underline{\underline{Q}} = \underline{\underline{I}} - 2\underline{u}\,\underline{u}^T$$

02.11.2023

Consider 2 linear spaces, each with a norm

V with $\|\cdot\|_V$

W with $\|\cdot\|_W$

Consider a linear operator (function)

$f : V \to W$

$$\boxed{f(\lambda x + \mu y) = \lambda f(x) + \mu f(y) \quad \text{for } x, y \in V \text{ arbitrary}, \ \lambda, \mu \in \mathbb{R}}$$

$$\|f\| = \sup_{0 \neq x \in V} \frac{\|f(x)\|_W}{\|x\|_V} = \sup_{\|x\|_V = 1} \|f(x)\|_W$$

↗ Norm on the linear space $\{ f : V \to W, f \text{ linear} \}$ induced by $\|\cdot\|_W$ and $\|\cdot\|_V$

Application: Consider $V = \mathbb{R}^n$, $W = \mathbb{R}^m$

linear operator $\mathcal{A} : V \to W$

$\mathcal{A} : \mathbb{R}^n \to \mathbb{R}^m$

↕ $\underline{A}$

$x \in \mathbb{R}^n$, choose a basis in $\mathbb{R}^n$, a basis in $\mathbb{R}^m$ →

↓

coordinate vector for x: $\underline{x}$

coordinate vector for $\mathcal{A}(x) = y$: $\underline{y}$

There is a matrix $\underline{A} \in \mathbb{R}^{m \times n}$ s.t. $\underline{y} = \underline{A}\,\underline{x}$

(and vice versa)

$\underline{A} : \mathbb{R}^n \to \mathbb{R}^m$

choose Norms in $\mathbb{R}^n, \mathbb{R}^m$ →

induced norm is the norm of the Matrix.

Choose $\|\cdot\|_2$ in $\mathbb{R}^n$ and $\mathbb{R}^m$:

$$\|\underline{A}\|_2 = \sup_{\underline{0} \neq \underline{x} \in \mathbb{R}^n} \frac{\|\underline{A}\,\underline{x}\|_2}{\|\underline{x}\|_2} = \sup_{\|\underline{x}\|_2 = 1} \|\underline{A}\,\underline{x}\|_2$$

If $n=m$ => even more naturally

$$\|\underline{A}\|_2 = \sup_{\|\underline{x}\|_2=1} \|\underline{A}\,\underline{x}\|_2 \neq \sqrt{\sum_{i,j=1}^{n} |A_{ij}|^2} = \|\underline{A}\|_F$$

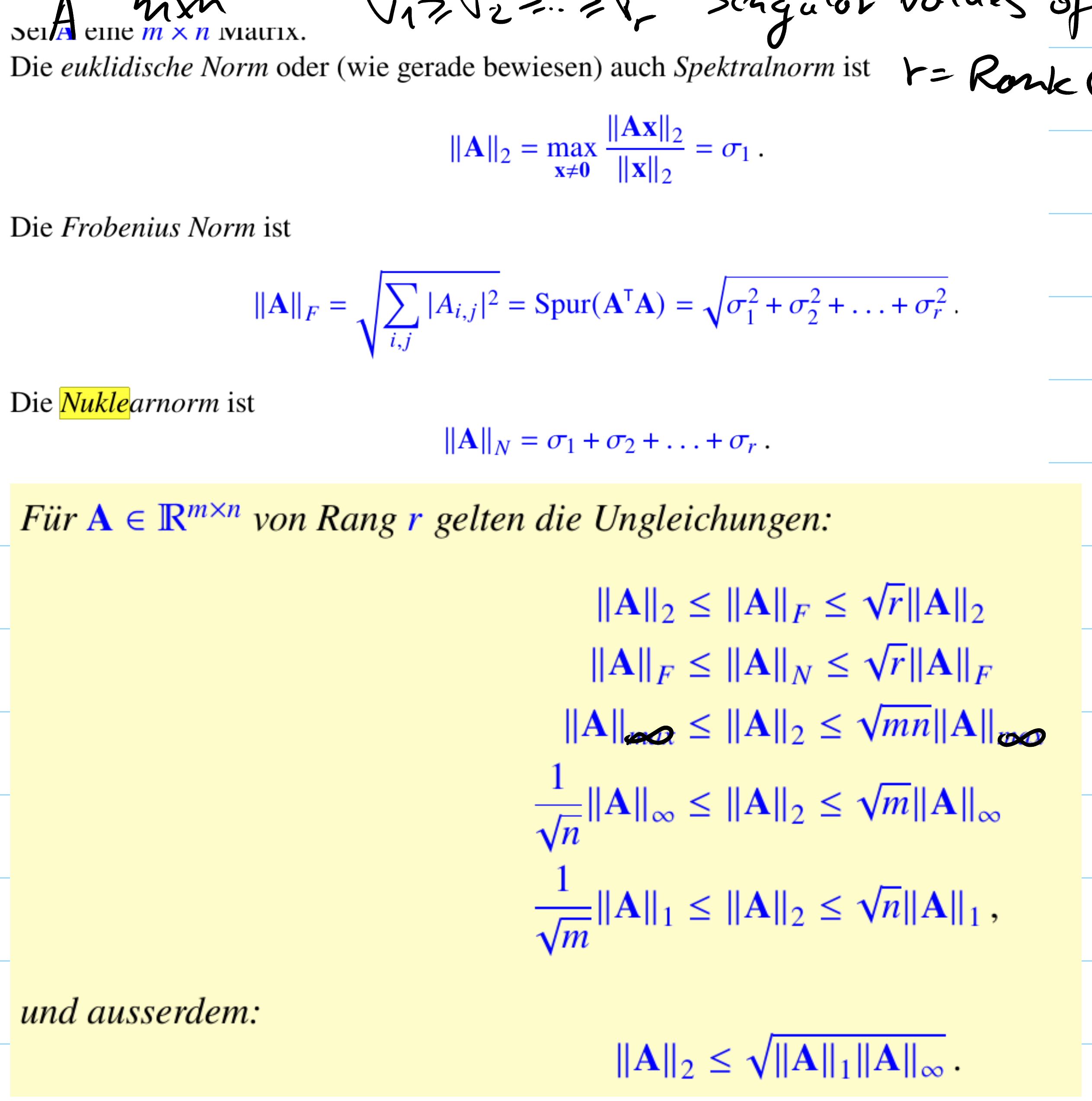
Euclidian norm.

Euclidian norm.

$\|\underline{A}\|_2 = \lambda_n$ = Spektral Norm.

$\underline{A}$ sym. pos.def => n real positive EW

$\lambda_1 < \dots < \lambda_n$



$\underline{A}$ $m \times n$ $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r$ singular values of $\underline{A}$

$r = \text{Rank}(\underline{A})$.

Sei $\mathbf{A}$ eine $m \times n$ Matrix.

Die *euklidische Norm* oder (wie gerade bewiesen) auch *Spektralnorm* ist

$$\|\mathbf{A}\|_2 = \max_{\mathbf{x} \neq \mathbf{0}} \frac{\|\mathbf{A}\mathbf{x}\|_2}{\|\mathbf{x}\|_2} = \sigma_1 .$$

Die *Frobenius Norm* ist

$$\|\mathbf{A}\|_F = \sqrt{\sum_{i,j} |A_{i,j}|^2} = \text{Spur}(\mathbf{A}^\top\mathbf{A}) = \sqrt{\sigma_1^2 + \sigma_2^2 + \ldots + \sigma_r^2} .$$

Die *Nuklearnorm* ist

$$\|\mathbf{A}\|_N = \sigma_1 + \sigma_2 + \ldots + \sigma_r .$$

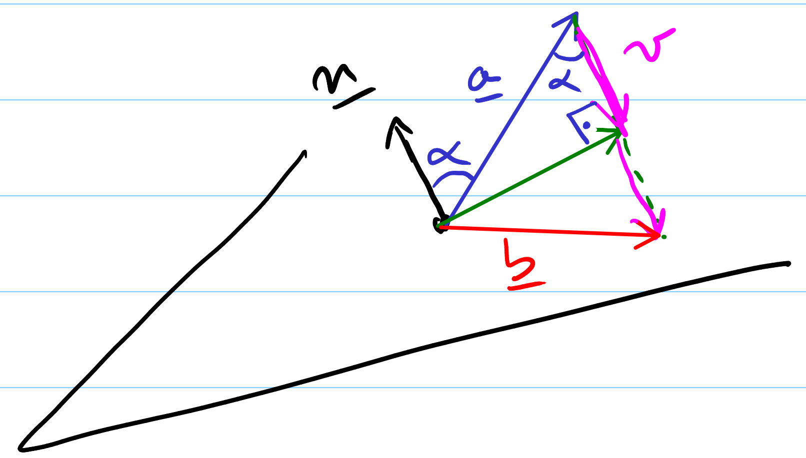
Für $\mathbf{A} \in \mathbb{R}^{m \times n}$ von Rang r gelten die Ungleichungen:

$$\|\mathbf{A}\|_2 \leq \|\mathbf{A}\|_F \leq \sqrt{r}\|\mathbf{A}\|_2$$
$$\|\mathbf{A}\|_F \leq \|\mathbf{A}\|_N \leq \sqrt{r}\|\mathbf{A}\|_F$$
$$\|\mathbf{A}\|_\infty \leq \|\mathbf{A}\|_2 \leq \sqrt{mn}\|\mathbf{A}\|_\infty$$
$$\frac{1}{\sqrt{n}}\|\mathbf{A}\|_\infty \leq \|\mathbf{A}\|_2 \leq \sqrt{m}\|\mathbf{A}\|_\infty$$
$$\frac{1}{\sqrt{m}}\|\mathbf{A}\|_1 \leq \|\mathbf{A}\|_2 \leq \sqrt{n}\|\mathbf{A}\|_1 ,$$

und ausserdem:

$$\|\mathbf{A}\|_2 \leq \sqrt{\|\mathbf{A}\|_1\|\mathbf{A}\|_\infty} .$$

(House holder) reflexions



Hyperplane to reflect on ad $\underline{n}$ Normal vektor of $\|\underline{n}\|_2 = 1$

$$\underline{b} = \underline{a} + 2\underline{v} = \underline{a} - 2\underline{n}\|\underline{v}\| = \underline{a} - 2\underline{n}\boxed{\underline{n}^T\underline{a}}$$

$$= \underbrace{(\underline{\underline{I}} - 2\underline{n}\underline{n}^T)}_{\underline{\underline{H}}}\underline{a}$$

$$\cos\alpha = \frac{\|\underline{v}\|}{\|\underline{a}\|}$$

$$\underline{a}^T\underline{n} \quad \langle \underline{a}, \underline{n} \rangle = \|\underline{a}\| \cdot \|\underline{n}\| \cdot \cos\alpha = \cancel{\|\underline{a}\|} \cdot \underbrace{\|n\|}_{1} \cdot \frac{\|v\|}{\cancel{\|a\|}} = \|v\|$$

⊛ Power method / EW / inverse iteration

Consider $\underline{\underline{A}}$ a matrix with EW $\lambda_1 > \lambda_2 \geq \dots \geq \lambda_n$ EW

$\underline{\underline{A}} \in \mathbb{R}^{n \times n}$ $\qquad \underline{v}_1 \; \underline{v}_2 \dots \underline{v}_n \in V$

$\hookrightarrow$ Basis in $\mathbb{R}^n$

Take:

$\underline{u} \in \mathbb{R}^n$ arbitrary. but $u_1 \neq 0$ $\Rightarrow$ $\underline{u} = u_1\underline{v}_1 + u_2\underline{v}_2 + \dots + u_n\underline{v}_n$

$$\underline{\underline{A}}\,\underline{u} = u_1\underline{\underline{A}}v_1 + u_2\underline{\underline{A}}v_2 + \dots + u_n\underline{\underline{A}}\underline{v}_n$$

$$= u_1\lambda_1\underline{v}_1 + u_2\lambda_2\underline{v}_2 + \dots + u_n\lambda_n\underline{v}_n$$

$$\underline{\underline{A}}^N\underline{u} = u_1\lambda_1^N\underline{v}_1 + u_2\lambda_2^N\underline{v}_2 + \dots + u_n\lambda_n^N\underline{v}_n$$

$$\underline{\underline{A}}^N\frac{\underline{u}}{\|u\|} = \lambda_1^N\left(\frac{u_1}{\|u\|}v_1 + \frac{u_2}{\|u\|}\underbrace{\left(\frac{\lambda_2}{\lambda_1}\right)^N}_{\to 0} + \dots + \frac{u_n}{\|u\|}\underbrace{\left(\frac{\lambda_n}{\lambda_1}\right)^N}_{0}v_n\right)$$

take u, $\|u\|=1$

$\underline{v} = \underline{\underline{A}}\,\underline{u}$, $v := v/\|v\|$

Here power iterations → $\underline{u}_n$, $\|u_n\|=1$, $\underline{u}_n \to \underline{v}_1$

belong to the biggest ev.

⊛ What if we want λ_n, $\underline{v}_n$ and not $\lambda_1, \underline{v}_1$?

Do power iteration for $\underline{\underline{A}}^{-1}$

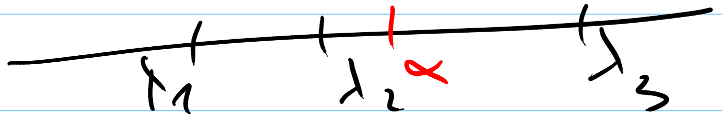
$$\frac{1}{\lambda_n} \boxed{>} \frac{1}{\lambda_{n-1}} ?\ldots > \frac{1}{\lambda_1}$$

⊛ Speed up with shifts.

$$\underline{\underline{A}},\ \underline{\underline{A}}^{-1},\ (\underline{\underline{A}} - \alpha\,\underline{\underline{I}})^{-1}$$

$\alpha - \lambda_2$ small.

$\boxed{\frac{1}{\lambda_2 - \alpha}}$ large.



09.11.2023

6. (M) Denote by $(\mathbf{x}_k) \in \mathbb{K}^k$ a sequence $[x_0, \cdots, x_{k-1}]$. Which of the following statements regarding discrete convolution (denoted by $*$) and discrete periodic convolution (denoted by $*_n$ with period n) is **wrong**?

a) $(\mathbf{x}_m) * (\mathbf{y}_n) \in \mathbb{K}^{m+n}$. False : by def. Length is $m+n-1$

b) $(\mathbf{x}_n) *_n (\mathbf{y}_n) \in \mathbb{K}^{2n}$. True : by def.

c) $(\mathbf{x}_m) * (\mathbf{y}_n) = (\mathbf{y}_n) * (\mathbf{x}_m)$ ← Commutativity: True by Remark 4.1.3.5

d) $(\mathbf{x}_n) *_n (\mathbf{y}_n) = (\mathbf{y}_n) *_n (\mathbf{x}_n)$ ←

e) $(\mathbf{x}_n) * (\mathbf{x}_n) = \mathbf{0}$ implies that $(\mathbf{x}_n) = \mathbf{0}$. → True ☺

f) $(\mathbf{x}_n) *_n (\mathbf{x}_n) = \mathbf{0}$ implies that $(\mathbf{x}_n) = \mathbf{0}$.

g) $(\mathbf{x}_m) * (\mathbf{y}_n) = \mathbf{0}$ implies that either $(\mathbf{x}_m) = \mathbf{0}$ or $(\mathbf{y}_n) = \mathbf{0}$. → False ! ☺

h) $(\mathbf{x}_n) *_n (\mathbf{y}_n) = \mathbf{0}$ implies that either $(\mathbf{x}_n) = \mathbf{0}$ or $(\mathbf{y}_n) = \mathbf{0}$.

← Fourier Transform of result

$$\widehat{(x_n) * (x_n)} = \underbrace{\widehat{(x_n)}}_{\underline{v}} \cdot \widehat{(x_n)} = \text{Skalar Product}$$

of $\underline{v}$ mit $\underline{v}$

$\underline{v}\cdot\underline{v} = \underline{v}^T\underline{v} = 0$

$\|\underline{v}\|_2^2 = 0 \Rightarrow \underline{v} = 0$

$\Rightarrow (x_n) = 0$

$\Rightarrow x_k = 0$ for all

$k = 0, 1, \ldots, n-1$

Remark No big difference in theory between periodic & non-periodic convolution because every conv. can be written as a periodic convolution by simply extending the signal with 0!

$$\widehat{(x_n) *_n (y_n)} = \widehat{(x_n)} \cdot \widehat{(y_n)} = \underline{v} \cdot \underline{w} = 0 \iff \underline{v} \perp \underline{w}$$

ok!

16.11.2023

Q: $$f'(t) \approx \frac{f(t+\delta) - f(t-\delta)}{2\delta} \quad \leftarrow !$$

$$f'(t) \approx \frac{f(t+\delta) - f(t)}{\delta}$$

$$f(t+\delta) = \boxed{f(t)} + \underline{\delta f'(t)} + \frac{\delta^2}{2} f''(t) + c \cdot \delta^3 \qquad \text{Taylor}$$

$$c(\delta) = \frac{1}{3!} f'''(\xi) \qquad \xi \in [a,b]$$

$$\Rightarrow f'(t) = \frac{f(t+\delta) - f(t)}{\delta} \;\Big|\; + \underbrace{\frac{\delta}{2} f''(t) + c \cdot \delta^2}$$

error : $O(\delta)$

$$f(t-\delta) = \boxed{f(t)} - \underline{\delta f'(t)} + \frac{\delta^2}{2} f''(t) + c \cdot \delta^3$$

Subtract the 2 expressions:

$$f(t+\delta) - f(t-\delta) = 2\delta f'(t) + 0 \cdot f''(t) + c\delta^3 \Rightarrow$$

$$f'(t) = \frac{f(t+\delta) - f(t-\delta)}{2\delta} \boxed{- \frac{c}{2} \cdot \delta^2}$$

error : $o(\delta^2)$

☺

⊛ Convolution Theorem:

$[u_0, u_1, \dots, u_{n-1}]$, $[v_0, v_1, \dots, v_{n-1}]$

$u \circledast v = \underline{\underline{C}}\,\underline{v}$ $\qquad \underline{\underline{C}} = \text{ciculant}(u_0, u_1, \dots, u_{n-1})$

$$\underline{\underline{C}} = F_n^{-1}\,\text{dig}(\underline{d})\,F_n$$

$$u \circledast v = \underbrace{\underline{\underline{F_n}}^{-1}\,\text{dig}(\underline{d})}\,\underline{\underline{F_n}}\,\underline{v} \qquad = (\underline{\underline{F_n}}\,\underline{u})^T\,\underline{\underline{F_n}}\,\underline{v} = \langle \underline{\underline{F_n}}\,\underline{u}, \underline{F_n}\,\underline{v}\rangle.$$

$$\text{dig}(\underline{d}) = \begin{bmatrix} d_0 & & & \\ & d_1 & & \\ & & \ddots & \\ & & & d_{n-1} \end{bmatrix} \qquad \underline{d} = F_n\,\underline{u}$$

A different view on Fourier Transform

Stone-Weierstrass: $f \in C[0,1] \Rightarrow \exists (f_n) \subset P$

such that $\max\limits_{t\in[0,1]} |f_n(t) - f(t)| \to 0$ ☹

How to find f

How fast $f_n \to f$?

Taylor: $f(t-t_0) = f(t_0) + (t-t_0) f'(t_0) + \ldots + \frac{(t-t_0)^{n-1}}{(n-1)!} f^{(n-1)}(t_0)$

$$+ \int_{t_0}^{t} \frac{(t-z)^{n-1}}{(n-1)!} f^{(n)}(z)\, dz$$

$f \in C^n[t_0-\delta, t_0+\delta]$

↪ very smooth around t_0

- locally good
- not stable for long n ☹

Fourier: replace monomial t^k by a wave

$$f_k(t) = e^{2\pi i k t} = \cos(2\pi k t) + i\sin(2\pi k t)$$

$$L^2[0,1] = \{ f : [0,1] \to \mathbb{C} ;\ \|f\|_{L^2[0,1]} < \infty \}$$

where

$$\|f\|^2_{L^2[0,1]} = \langle f, f \rangle_{L^2[0,1]}$$

where.

$$\langle f, g \rangle_{L^2[0,1]} = \int_0^1 \overline{f(t)}\, g(t)\, dt$$

$L^2[0,1]$ linear space with norm $\|\cdot\|_{L^2}$ coming from the scalar product $\langle \cdot, \cdot \rangle_{L^2[0,1]}$.

Theorem $f:\mathbb{R}\to\mathbb{C}$ periodical of period 1.

$f\in L^2[0,1) \Rightarrow$

$$f(t)=\sum_{k=-\infty}^{\infty}\hat{f}(k)\, e^{2\pi i k t}$$ convergence in $\|\cdot\|_{L^2}$

$\downarrow$ Fourier coefficient

$\left((e_k)_{k\in\mathbb{Z}} \text{ complete } \overbrace{\text{basis}}^{\text{orthonormal}} \text{ of } L^2[0,1) \right)$

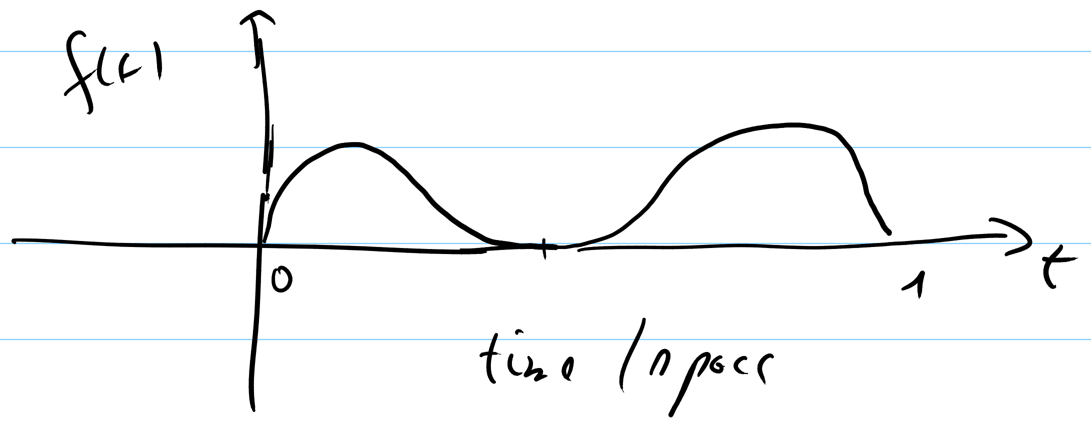
Parseval: $$\|f\|_{L^2}^2 = \sum_{k=-\infty}^{\infty} |\hat{f}(k)|^2$$

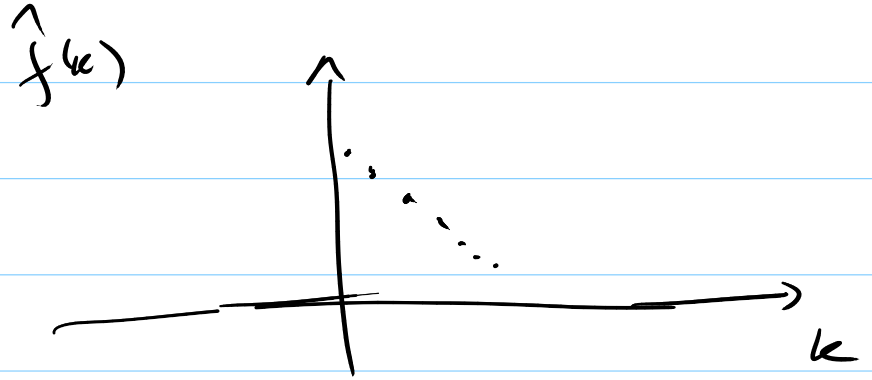
If operations are possible:

$$\widehat{f^{(n)}}(k) = (2\pi i k)^n \hat{f}(k)$$

$$\|f^{(n)}\|_2^2 = (2\pi)^n \sum_{k=-\infty}^{\infty} k^{2n} |\hat{f}(k)|^2$$

smooth $\Longleftrightarrow$ $\hat{f}(k)$ decay faster than k^{-n}
$\approx O(k^{-n})$





Analog signal $f(t)$

Discrete signal $f(t_0), f(t_1), \dots, f(t_{n-1})$

How to compute $\hat{f}(k)$ from samples only?

$$\underset{\underset{\text{ONB}}{\uparrow}}{\hat{f}(k)} = \langle \varphi_k , f \rangle = \int_0^1 f(t)\, e^{-2\pi i k t}\, dt$$

take samples at $t_j = \frac{j}{N}$, $j=0,1,\dots,N$

TR

sum

$$\to \frac{1}{N}\left[\frac{1}{2} f(0) + \sum_{j=1}^{N-1} f\left(\frac{j}{N}\right) + \frac{1}{2} f(1)\right] =$$

$$e^{-2\pi i k \frac{j}{N}}$$

$$\left(\varphi_k\left(-\frac{j}{N}\right)\right)$$

Periodicity: $f(1)=f(0)$

$$\downarrow$$

$$= \frac{1}{N} \sum_{j=0}^{N} f\left(\frac{j}{N}\right) e^{-2\pi i k \frac{j}{N}} = \hat{f}_N(k)$$

$\hat{f}(k) \approx \hat{f}_N(k)$; consider $N=2n$

$$f(t) \approx p_N(t) = \sum_{k=-\frac{N}{2}}^{\frac{N}{2}-1} \hat{f}_N(k)\, e^{2\pi i k t} ;$$

$$\omega_N = e^{-\frac{2\pi i}{N}} , \quad i^2=-1 \quad ; \quad w_N = \overline{\omega}_N = e^{\frac{2\pi i}{N}}$$

$\Rightarrow$ This is the Discrete Fourier Transform!

$$\hat{f}(k) = \frac{1}{N} \sum_{j=0}^{N-1} f\left(\frac{j}{N}\right) \omega_N^{kj}$$

One can show

$$\begin{cases} p_N(t_j) = f(t_j) \\ j=0,1,\dots,N-1 \end{cases} \Rightarrow p_N \text{ interpolates } f \text{ at } \left(\frac{j}{N}\right)_{j=0,1,\dots,N-1}$$

$\underline{\underline{F}}_N$ = as in the lecture

$$\underline{v}_k = \begin{bmatrix} w_N^{0k} \\ w_N^{1\cdot k} \\ \vdots \\ w_N^{(N-1)k} \end{bmatrix} \qquad \underline{\underline{F}}_N = [\underline{v}_0, \underline{v}_1, \dots, \underline{v}_{N-1}]$$

Fourier Matrix

$\underline{v}_0, \dots \underline{v}_{N-1}$ ONB in $\mathbb{C}^N$ etc.