ENGG 5781 Matrix Analysis and Computations
Term 2, 2017-18

Instructor: Prof. Wing-Kin (Ken) Ma
Scribe: Zihao FU

2018.1.12

A subset S of $\mathbb{R}^m$ is called a subspace if

$$x, y \in S,\ \alpha, \beta \in \mathbb{R} \Rightarrow \alpha x + \beta y \in S$$

subspace是一种特殊的subset.

$$a_1, \cdots, a_n \in S \Rightarrow \sum_{i=1}^{n} \alpha_i a_i \in S$$

S is a subsp. $\quad \forall \alpha_1, \cdots, \alpha_n$

Span of $\{a_1, \cdots, a_n\} \subset \mathbb{R}^m$:

$$\alpha = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix}$$

$$\operatorname{span}\{a_1, \cdots, a_n\} = \{y = \sum_i \alpha_i a_i \mid \alpha \in \mathbb{R}^n\}$$

Thm: Any subsp. $S \subseteq \mathbb{R}^m$ can be represented by

$S = \operatorname{span}\{a_1, \cdots a_n\}$

for some $a_1, \cdots, a_n \in \mathbb{R}^n$ and positive integer n.

学习数学的过程:

Hypothesis or assumptions, or definition

物理意义是什么?

$\downarrow$

properties, facts, theorems, results

$\downarrow$

more results $\longrightarrow$ app

Thm A

$\downarrow$

Thm B

不能用B去证A.

Span is a subsp. (easy to verify)

proof: Suppose, $x, y \in \operatorname{span}\{a_1, \cdots, a_n\} \Rightarrow$

$x = \sum_i \alpha_i a_i$, $y = \sum_i \beta_i a_i$

$\Rightarrow \theta x + \gamma y$

$= \theta\left(\sum_i \alpha_i a_i\right) + \gamma\left(\sum_i \beta_i a_i\right)$

$= \sum_i (\theta \alpha_i + \gamma \beta_i) a_i \in \mathrm{span}\{a_1, \cdots, a_n\}$

But is subspace a span?

$\{a_1, \cdots, a_n\} \subseteq \mathbb{R}^m$ is said linearly independent (LI) if $\sum_{i=1}^n \alpha_i a_i \neq 0$. $\forall \alpha \in \mathbb{R}^n, \alpha \neq 0$ otherwise we say it linearly dependent (LD)

· Fact $\{a_1, \cdots a_n\} \subseteq \mathbb{R}^m \Rightarrow n \leq m$
is LI

A subset $\{a_{i_1}, \cdots, a_{i_k}\}$ of $\{a_1, \cdots, a_n\}$ is said maximal LI if

1. $\{a_{i_1}, \cdots, a_{i_k}\}$ is LI
2. $\{a_{i_1}, \cdots, a_{i_k}\} \not\subseteq \{a_{j_1}, \cdots, a_{j_l}\}$ for any LI subset $\{a_{j_1}, \cdots, a_{j_l}\}$ of $\{a_1, \cdots, a_n\}$

Example:

$a_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad a_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad a_3 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad a_4 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$

The LI subsets of $\{a_1, a_2, a_3, a_4\}$ are:

$\{a_1\} \quad \{a_2\} \quad \{a_3\} \quad \cancel{\{a_4\}}$

$\{a_1, a_2\}, \{a_2, a_3\}, \{a_1, a_3\}$

这三个都是 maximal LI ∵ 个数是 LI set 中最多的

$\{a_1\} \subseteq \{a_1, a_2\}$ ∴ $\{a_1\}$ 不是 maximal

A vector set $\{b_1, \dots, b_k\} \subseteq \mathbb{R}^m$ is called a basis for S if $\{b_1, \dots, b_k\}$ is LI and $S = \text{span}\{b_1, \dots, b_k\}$

← subspace of $\mathbb{R}^m$ and $S \neq \{0\}$

$\text{span}\{a_1, \dots, a_n\} = \text{span}\{a_{i_1}, \dots, a_{i_k}\}$ where $\{a_{i_1}, \dots, a_{i_k}\}$ is a maximal LI subset of $\{a_1, \dots, a_n\}$

The dim of a subsp. S, denoted $\dim S$, is the no. of elements (k) of a basis for S.

Orthogonal Complement.

Let $S \subseteq \mathbb{R}^m$ be a subset (may not be a subsp.)

$S^\perp = \{y \mid z^T y = 0, \forall z \in S\}$

- $S^\perp$ is a subsp.
- $S + S^\perp = \mathbb{R}^m$ if S is a subsp.
- $\dim S + \dim S^\perp = m$ if S is a subsp.

2018.6.18

- $\operatorname{span}\{a_1, \cdots, a_n\} = \{y = \sum_{i=1}^{n} \alpha_i a_i \mid \alpha \in \mathbb{R}^n\}$
- Let $S \subseteq \mathbb{R}^n$ be a subspace.
- $\operatorname{span}\{a_1, \cdots, a_n\}$ is a subsp.
- any S can be written as $S = \operatorname{span}\{a_1, \cdots, a_n\}$

Example: $S = \left\{ y = \begin{bmatrix} y_1 \\ y_2 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \middle| \; y_1, y_2 \in \mathbb{R} \right\}$

$= \operatorname{span}\{e_1, e_2\}$, where

$e_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad e_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$

a: 存在一个
the: 仅有一个

- $\{a_1, \cdots, a_n\}$ is called a basis for S if $\{a_1, \cdots, a_n\}$ is LI & $S = \operatorname{span}\{a_1, \cdots, a_n\}$

- $\dim \operatorname{span}\{a_1, \cdots, a_n\} = n$ if $\{a_1, \cdots, a_n\}$ is LI.

$\dim \operatorname{span}\{a_1, \cdots, a_n\} = k$ if k is the no. of elements of any max LI subset $\{a_{i_1}, \cdots, a_{i_k}\}$ of $\{a_1, \cdots, a_n\}$ $k \le n$

$\dim \mathbb{R}^m = m$, $\dim\{0\} = 0$

正课

Defin: Range space: Let $A \in \mathbb{R}^{m \times n}$

$$R(A) = \{ y = Ax \mid x \in \mathbb{R}^n \}$$

注意到:

$$Ax = [a_1, \cdots, a_n] \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \sum_{i=1}^{n} x_i a_i$$

$\therefore R(A) = \text{span}\{a_1, \cdots, a_n\}$ 实际就是A的列向量的span

Defin: Null space:

$$N(A) = \{ x \in \mathbb{R}^n \mid Ax = 0 \}$$

注: null space is a subspace.

证: Let $x_1, x_2 \in N(A)$

For $\alpha_1, \alpha_2 \in \mathbb{R}_2$

$$A(\alpha_1 x_1 + \alpha_2 x_2) = \alpha_1 Ax_1 + \alpha_2 Ax_2 = 0$$

$\Rightarrow N(A)$ is a subsp!

注: $\because N(A)$ 是 sub sp, $\therefore$ 它是一个 span,

$\therefore \exists B$, s.t. $N(A) = R(B)$

Define,

rank:

rank(A) = max no. of LI columns of A

$$= \dim \operatorname{span}\{a_1, \cdots, a_n\}$$

$$= \dim R(A)$$

性质见 ppt.

$$\operatorname{rank}(A) \le \min\{m, n\}$$

若 $\operatorname{rank}(A) < \min\{m, n\}$ 则称 rank deficient

rank deficient 不是好事

Define Inverse 见 ppt

这里，我们说 the Inverse, ∵ Inverse 唯一.

Define Determinant

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

$$\det(A) = a_{11} a_{22} - a_{12} a_{21}$$

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

A_{11}

A_{13}

$$\det(A) = a_{11} \cdot \det(A_{11}) - a_{12} \cdot \det(A_{12}) + a_{13} \cdot \det(A_{13})$$

$$A_{12} = \begin{bmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{bmatrix}$$

$$A=\begin{bmatrix} a_{11} & 0 & 0 \\ a_{21} & a_{22} & 0 \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \qquad \det(A)=a_{11}\det(A_{11})$$
$$=a_{11}\cdot a_{22}\cdot a_{33}$$

这种三角形式很容易算特征值，之后来看应用

性质：

$$Ax=0 \iff \det(A)=0$$

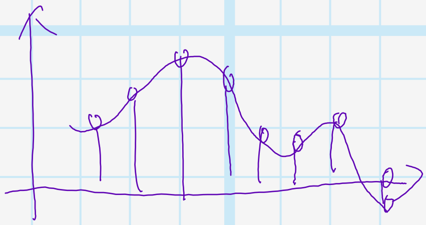
$|\det(A)|$ 是列向量张成的平行四面体的体积。

$$A=\begin{bmatrix} B & C \\ 0 & D \end{bmatrix} \qquad \det(A)=\det(B)\det(D)$$

2018.01.09

Vector Norms

例：



如何评估这两个sequence的相似?

Norm满足的四条property见ppt.

比较难证的是第三条，三角不等式.

1-norm $\|x\|_1=\sum_{i=1}^{n}|x_i|$ 就是曼哈顿距离.

例：证明1-norm的三角不等式.

$$f(x)=\sum_{i=1}^{n}|x_i|$$

$$f(x+y)=\sum_{i=1}^{n}|x_i+y_i|$$

$$\le\sum_{i=1}^{n}|x_i|+|y_i|$$

$$=f(x)+f(y)$$

三角不等式就是指和和f的操作大小问题?

Cauchy-Schwartz inequality

$$|x^Ty|\le\|x\|_2\|y\|_2$$

当且仅当 $x=\alpha y$，即它们共线时，取等号

$\Rightarrow$: $x^T y = x^T(\alpha x)$

$= \alpha \cdot x^T x$

$= \alpha \cdot \|x\|_2^2$

$= \alpha \|x\|_2 \|x\|_2$

$= \|y\|_2 \cdot \|x\|_2$

$\Leftarrow$: 略.

$\square$

下面看两个向量的距离.

$$\|x-y\|_2^2$$

$$= (x-y)^T(x-y)$$

$$= x^T x - y^T x - x^T y + y^T y$$

$$= \|x\|_2^2 - 2 y^T x + \|y\|_2^2$$

可见 $x^T y$ 的内积对它们的距离有影响.

Hölder inequality

$$|x^T y| \le \|x\|_p \|y\|_q$$

for any p, q s.t. $\frac{1}{p} + \frac{1}{q} = 1$ $\quad p \ge 1$

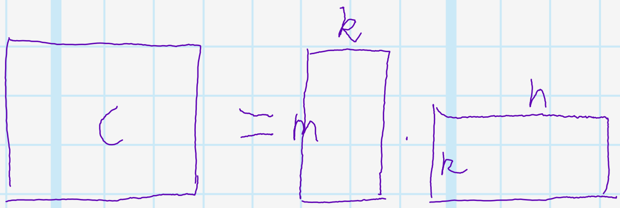
当 $(p,q) = (2,2)$ 时, 得出 Cauchy-Schwartz 不等式

证明秩不等式 $\quad A \in \mathbb{R}^{m\times n}$

(1) $\text{rank}(A) = \text{rank}(A^T) \iff \dim \underbrace{R(A)}_{\subseteq \mathbb{R}^m} = \dim \underbrace{R(A^T)}_{\subseteq \mathbb{R}^n}$

(2) $\text{rank}(AB) \le \min\{\text{rank}(A), \text{rank}(B)\}$

(2) 的作用:



$\Rightarrow \operatorname{rank}(C) \le k$

若C能分解为两个矩阵的乘积

当A has full col rank

B has full row rank

时, $C = AB \Leftrightarrow \operatorname{rank}(C) = k$

(a) $\dim R(AB) \le \dim R(A)$

Suppose (a) is true

$$\dim R(AB) \overset{(1)}{=} \dim R((AB)^T)$$
$$= \dim R(B^T A^T)$$
$$\underset{(2)}{\le} \dim R(B^T)$$
$$= \dim R(B)$$

proof of (a).

$$R(AB) = \{ y = ABx \mid x \in \mathbb{R}^n \}$$
$$= \{ y = Az \mid z \in R(B) \}$$
$$\subseteq \{ y = Az \mid z \in \mathbb{R}^k \}$$
$$\Rightarrow \dim(AB) \le \dim R(A)$$

补充 Fact: Let S_1, S_2 be subspace

1、 $S_1 \subseteq S_2 \Rightarrow \dim S_1 \le \dim S_2$

2、 $S_1 \subseteq S_2$, $\dim S_1 = \dim S_2 \Rightarrow S_1 = S_2$

3、 $\dim S_1 = m \Leftrightarrow S_1 = \mathbb{R}^m$

证明 (1) $\mathrm{rank}(A)=\mathrm{rank}(A^T)$

(1) $\mathrm{rank}(A)=\mathrm{rank}(A^T) \Leftrightarrow \dim R(A)=\dim R(A^T)$

$\Leftrightarrow \dim R(A^T) \geq \dim R(A)$

∴ 等价于证明如下三个:

1° show $A\alpha=0 \Leftrightarrow A^TA\alpha=0$

2° show $\dim R(A^T) \geq \dim R(\underbrace{A^TA}_{n\times n})$

3° show $\dim R(A^TA) \geq \dim R(\underbrace{A}_{m\times n})$

1°. $A\alpha=0 \Rightarrow A^T(A\alpha)=A^T0=0$

$A^TA\alpha=0 \Rightarrow \alpha^TA^TA\alpha=\alpha^T0=0$

$\therefore \|A\alpha\|_2^2=0 \Rightarrow A\alpha=0$

2° $R(A^T)=\{y=A^Tx \mid x\in\mathbb{R}^m\}$

$\supseteq \{y=A^Tx \mid x\in R(A)\}$

$=\{y=A^Tx \mid x=Az,\ z\in\mathbb{R}^n\}$

$=\{y=A^TAz \mid z\in\mathbb{R}^n\}=R(A^TA)$

3° Let $k=\dim R(A)$ Let

$\{a_{i_1},\cdots,a_{i_k}\}$ be a max LI subset of

$\{a_1,\cdots,a_n\}$ I want to prove $\dim R(A^TA)\geq k$

$\sum_{j=1}^{k}\alpha_j a_{i_j}\neq 0\ \forall\alpha\in\mathbb{R}^k \overset{1^\circ}{\Longleftrightarrow}$

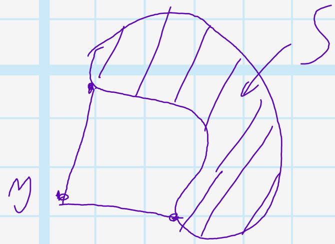
$\sum_{j=1}^{k}\alpha_j A^Ta_{i_j}\neq 0\ \forall\alpha\in\mathbb{R}^k \Leftrightarrow \{A^Ta_{i_1},\cdots,A^Ta_{i_k}\}$ is LI

$$\Rightarrow \dim \operatorname{span}\{A^T a_1, \cdots, A^T a_n\} \geq k$$
$$= R(A^T A)$$

一般书上直接给出 $\operatorname{rank}(A) = \operatorname{rank}(A^T)$，老师自己记出来的.

Let $S \subseteq \mathbb{R}^m$ be nonempty & closed. Proj of $y \in \mathbb{R}^m$ onto S:

$$\Pi_S(y) = \underset{z \in S}{\operatorname{argmin}} \|z - y\|_2^2$$



非凸的话，投影有多个

本课程关注在subspace上的投影

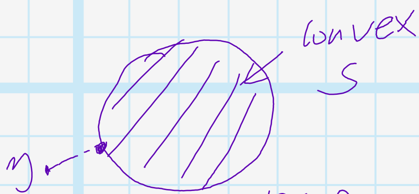
Proj. Thm, If S is a subsp.

1° proj. of y onto S is unique;

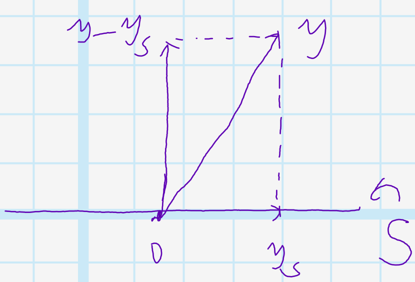
2° $y_s = \Pi_s(y) \Leftrightarrow \begin{cases} y_s \in S \\ z^T(y_s - y) = 0 \quad \forall z \in S \end{cases}$

$$\Updownarrow$$

$$y_s - y \in S^{\perp}$$



凸的话，投影唯一



Orthogonal Complement

$$S^{\perp} = \{y \in \mathbb{R}^m \mid z^T y = 0, \ \forall z \in S\}$$

2018.1.25

内积

$$\langle x,y\rangle=\sum_{i=1}^{n}x_i^{*}y_i=x^{H}y$$

为什么这么定义?

当 $x\in\mathbb{R}^n$ 时 $\|x\|_2^2=\langle x,x\rangle=\sum_{i=1}^{n}|x_i|^2$

当 $x\in\mathbb{C}^n$ 时, $\|x\|_2^2=\sum_{i=1}^{n}x_i^{*}x_i=\sum_{i=1}^{n}|x_i|^2$

Orthogonal Bases and Matrices.

如何求线性组合的系数?

$$y=\sum_{i=1}^{n}\alpha_i a_i$$

$(a_1,a_2,\cdots a_n)$ orthonormal

(相互正交,且模长为1)

$$a_j^Ty=a_j^T\left(\sum_{i=1}^{n}\alpha_i a_i\right)$$

$$=\sum_{i=1}^{n}\alpha_i a_j^Ta_i=\alpha_j$$

用矩阵形式来理解上式:

$$y=A\alpha$$

$$A^Ty=A^TA\alpha$$

$$=\begin{bmatrix}a_1^T\\a_2^T\\\vdots\\a_n^T\end{bmatrix}[a_1,a_2,\cdots,a_n]\alpha$$

$$=\begin{bmatrix}1&&\\&\ddots&\\&&1\end{bmatrix}=I\alpha=\alpha$$

Orthogonal 和 semi-orthogonal

所有 orthogonal matrix 都能写成

$$Q = \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{bmatrix}$$ 的形式，三维再加一个变量.

$QQ^T = I$ 当 Q 是 orthogonal 的时候是成立.
当 Q 是 semi-orthogonal 时不成立

$Q^TQ = I$ 对 orthognal 和 semi-orthogonal 都成立.

Let $S = R(Q_1)$ $Q_1 \in \mathbb{R}^{n \times k}$, Note: $\dim S = k$

- $\dim S + \dim S^\perp = n$
  $\dim S^\perp = n - k$

- I can always write $S^\perp = R(Q_2)$ for some semi-ortho Q_2

- Is it true $Q_1^TQ_2 = 0$? 或者说 $\underbrace{\begin{bmatrix} Q_1^T \\ Q_2^T \end{bmatrix}}_{Q^T} \underbrace{[Q_1 \; Q_2]}_{Q} \stackrel{?}{=} I$

$$Q^TQ = \begin{bmatrix} Q_1^TQ_1 & Q_1^TQ_2 \\ Q_2^TQ_1 & Q_2^TQ_2 \end{bmatrix}$$

since $y^Tx = 0 \quad \forall x \in S, y \in S^\perp$

$\Rightarrow (Q_2\beta)^T(Q_1\alpha)=0 \quad \forall \alpha\in\mathbb{R}^k, \beta\in\mathbb{R}^{n-k}$ 写成 range space 的形式，

$\Leftrightarrow \beta^T Q_2^T Q_1 \alpha = 0 \quad \forall \alpha, \beta$

$\Rightarrow Q_2^T Q_1 = 0$ $\because \alpha, \beta$ 可以任取

Complexities of Matrix Computations

2018.1.26

Complexities of Matrix Computations

$$y^T x = \sum_{i=1}^{n} x_i y_i = x_1 y_1 + \cdots x_n y_n$$

计算flops

$$\left.\begin{array}{l} \text{"}\times\text{"}: n \\ \text{"}+\text{"}: n-1 \end{array}\right\} \Rightarrow \text{flops} = 2n-1$$

$$Ax = \begin{bmatrix} a_1^T \\ a_2^T \\ \vdots \\ a_n^T \end{bmatrix} x = \begin{bmatrix} a_1^T x \\ a_2^T x \\ \vdots \\ a_n^T x \end{bmatrix} : \quad \text{flops} = (2n-1) \times m$$

$$\underset{m\times n}{A}\underset{n\times k}{X} = A[x_1, \cdots, x_k] \qquad \text{flops} = (2n-1)\times m \times k$$

$$= [Ax_1, \cdots, Ax_k] \qquad = O(nmk)$$

Big O Notation

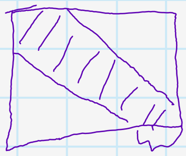
How to save Computation

调整矩阵连乘的顺序能节省时间。

$$ABC \longrightarrow A(BC)$$

$$\underset{n\times n}{A}\underset{n\times m}{X} \quad \text{复杂度为 } O(n^2 m)$$

若A是对角阵 $A=\begin{bmatrix} a_1 & & \\ & \ddots & \\ & & a_n\end{bmatrix}$ 或 band diagonal



p+1

复杂度为 $O(pn)$

稀疏矩阵

Aim: Given $a_1, \cdots, a_n \in \mathbb{R}^m$ (LI), find an orthonormal set $\{q_1, \cdots, q_n\}$ such that.

$$\mathrm{span}\{a_1, \cdots, a_n\} = \mathrm{span}\{q_1, \cdots, q_n\}$$

Suppose $\mathrm{span}\{a_1, \cdots, a_{i-1}\} = \mathrm{span}\{q_1, \cdots, q_{i-1}\}$

$$\mathrm{span}\{a_1, \cdots, a_i\} = \left\{ y = \sum_{j=1}^{i-1} \alpha_j a_j + \cdots + \alpha_i a_i \,\middle|\, \alpha \in \mathbb{R}^i \right\}$$

$$= \mathrm{span}\{a_1, \cdots, a_{i-1}\} + \mathrm{span}\{a_i\}$$

$$= \mathrm{span}\{q_1, \cdots, q_{i-1}\} + \mathrm{span}\{a_i\}$$

$$= \left\{ y = \sum_{j=1}^{i-1} \alpha_j q_j + \alpha_i a_i \,\middle|\, \alpha \in \mathbb{R}^i \right\}$$

$$= \left\{ y = \sum_{j=1}^{i-1} \alpha_j q_j + \alpha_i \left(\tilde{q}_i + \sum_{j=1}^{i-1} (q_j^T a_i) q_j\right) \,\middle|\, \alpha \in \mathbb{R}^i \right\}$$

$$= \left\{ y = \sum_{j=1}^{i-1} (\alpha_j + \alpha_i (q_j^T a_j)) q_j + \alpha_i \tilde{q}_i \,\middle|\, \alpha \in \mathbb{R}^i \right\}$$

$$= \text{span}\{a_1, a_2, \cdots a_{i-1}, \tilde{a}_i\}$$

$$\dim(\ ) = i-1 \text{ if } \tilde{a}_i = 0$$

矩阵乘法的外积表示。

一个矩阵乘法也可以理解为外积的和

$$C = \underset{m\times n}{A}\ \underset{k\times n}{B} = \sum_{i=1}^{k} a_i b_i^T$$

可以用 rank 的性质来给出 rank(C) 的上界

Extension to $\mathbb{R}^{m\times n}$

基也可以选为矩阵

$$\text{span}\{A_1, \cdots, A_k\} = \{Y \in \mathbb{R}^{m\times n} \mid Y = \sum_{i=1}^{k} \alpha_i A_i, \alpha \in \mathbb{R}^k\}$$

人脸识别中的 eigen face

定义矩阵间的内积

$$\langle X, Y \rangle = \text{tr}(Y^T X)$$

第一章结束。

2018.2.2

证明范德蒙矩阵满秩，
课件上把上面切出来为 B

$$B=\begin{bmatrix} 1 & y_1 & y_1^2 & \cdots & y_1^{k-1} \\ 1 & y_2 & y_2^2 & \cdots & y_2^{k-1} \\ & & \cdots & & \\ 1 & y_k & y_k^2 & \cdots & y_k^{k-1} \end{bmatrix}$$

$\text{rank}(B)=k$ if $y_i \neq y_j \ \forall i \neq j$

proof: show $B\alpha \neq 0 \quad \forall \alpha \neq 0$

复习下矩阵分解：

$$f(y)=1-3y+2y^2$$

$$=(1-y)(1-2y)$$

$$y(y)=0 \iff y=1 \text{ or } y=\frac{1}{2}$$

$$f(y)=\alpha_1+\alpha_2 y+\alpha_3 y^2+\cdots \alpha_k y^{k-1}$$

$$f(y)=c\prod_{i=1}^{k-1}(\underset{\substack{\uparrow \\ \text{root}}}{y_i}-y)$$

Suppose for some $\alpha \neq 0$, $B\alpha=0$

$\Rightarrow \alpha_1+\alpha_2 y_i+\alpha_3 y_i^2+\cdots+\alpha_k y_i^{k-1}=0, \quad i=1,\ldots k$

y_i is a rook of the polynomial

$$f(y) = \alpha_1 + \alpha_2 y + \cdots \alpha_k y^{k-1}$$

这个多项式不能有k个相异的根，

∴矛盾.

Basis Representation.

$$[h_0 \;\; h_{n-1} \;\; \cdots \;\; h_2 \;\; h_1] \begin{bmatrix} 1 \\ e^{j\frac{2\pi k}{n}} \\ \vdots \\ \vdots \\ e^{j\frac{2\pi k}{n} n-2} \\ e^{j\frac{2\pi k}{n} n-1} \end{bmatrix} = h_0 + h_{n-1} e^{\frac{j2\pi k}{n}} + \cdots + h_2 e^{\frac{j2\pi k(n-2)}{n}} + h_1 e^{\frac{j2\pi k(n-1)}{n}}$$

$$= h_0 + h_{n-1} e^{-\frac{j2\pi k}{n}(n-1)} + \cdots + h_2 e^{-\frac{j2\pi k}{n}2} + h_1 e^{-\frac{j2\pi k}{n}1}$$

$$= h_0 + h_1 e^{-\frac{j2\pi k}{n}} + h_2 e^{-\frac{j2\pi k}{n}\times 2} + \cdots + h_{n-1} e^{-\frac{j2\pi k}{n}(n-1)}$$

倒过来写.

$$= \underbrace{\sum_{i=0}^{n-1} h_i e^{-\frac{j2\pi k}{n} i}}_{\text{DFT with freq. at } \frac{2\pi k}{n}}$$

2018.2.8

我们找 basis 都找 tall matrix，不找 fat.

矩阵分解.

$$\min_{A}\left\{\sum_{i=1}^{n}\left[\min_{b_i}\|y_i - Ab_i\|_2^2\right]\right\}$$

2018.3.1

Review

Schur dec: Let $A \in \mathbb{R}^{n\times n}$ or $A \in \mathbb{C}^{n\times n}$ Then,

$$A = \underset{\substack{n\times n \\ \text{unitary}}}{U}\ \underset{\substack{\text{upper}\,\triangle \\ n\times n \\ \text{diagonals}= \\ \lambda_1,\cdots,\lambda_n}}{T}\ U^H$$

Eigen dec: Let $A \in \mathbb{H}^n$, ← set of all $n\times n$ complex Hamitian matrices

是Schur dec的特殊形式，可能不存在

Then

$$A = \underset{\substack{n\times n \\ \text{unitary}}}{V}\ \underset{\substack{\Lambda = \mathrm{Diag}(\lambda_1,\cdots,\lambda_n) \\ \lambda_i\text{'s are real}}}{\Lambda}\ V^H$$

这里是对角 而 Schur 是三角.

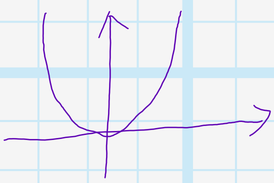
Let $A \in S^n$ ← set of all $n\times n$ real symmetric matrices

Then

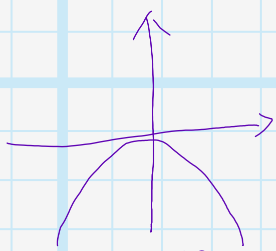
$$A = \underset{\substack{n\times n \\ \text{orthognal}}}{V}\ \Lambda\ V^T$$

今日，讲

$$f(x) = ax^2$$



$a > 0$



$a < 0$.

$f(x) = x^T A x$ $\quad$ A is real, $x \in \mathbb{R}^n$

$$[Ax]_i = \sum_j a_{ij} x_j$$

$$x^T(Ax) = \sum_i x_i \sum_j a_{ij} x_j$$

$$= \sum_i \sum_j x_i x_j a_{ij}$$

当 A 是 1×1 时 就退化为了 ax^2

$$x^T(A^T)x = \sum_i \sum_j x_i x_j a_{ji} = x^T A x$$

$$x^T(A)x = x^T\left(\frac{1}{2}A + \frac{1}{2}A^T\right)x$$

不论 A 是否是对称的，$\frac{1}{2}A + \frac{1}{2}A^T$ 都是对称的.

∴ 之后可以只研究当 A 为对称时的情形.

We will assume $A \in S^n$

$A \in S^n$ is said to be positive semidefinite (PSD) if $x^T A x \geq 0 \quad \forall x \in \mathbb{R}^n$

PD: $x^T A x > 0 \quad \forall x \in \mathbb{R}^n \setminus \{0\}$

Notation: $A \succeq 0 \Rightarrow$ A PSD

$A \succ 0 \Rightarrow$ A PD

$A \not\succeq 0 \Rightarrow$ A indef

Thm, $A \succeq 0$ iff $\lambda_i \geq 0$, $i=1,\dots,n$

$A \succ 0$ iff $\lambda_i > 0$, $i=1,\dots,n$

如何理解：

$$x^T A x = x^T V \Lambda V^T x$$

$$= z^T \Lambda z \qquad z \in R(V^T) = \mathbb{R}^n$$

$$= \sum_i \lambda_i (z_i)^2$$

2018.3.8

- $A \in S^n$ is PSD if

$$x^T A x \geq 0 \quad \forall x \in \mathbb{R}^n$$

- Any $A \in S^n$ can be factored as

$$A = V \Lambda V^T \qquad V \text{ ortho}, \ \Lambda = \mathrm{Diag}(\lambda_1, \cdots, \lambda_n)$$

- $A \in S^n$ is PSD iff $\lambda_i \geq 0 \ \forall i$

- $A \in S^n$ is PSD 可以把 $A = V\Lambda V^T$ 拆成 $V\Lambda^{1/2}\Lambda^{1/2}V^T$

$$A = B^T B$$

for some B (does not need to be square)

$A^{1/2} = V\Lambda^{1/2}V^T$ is the unique PSD square root of A.

- Let $B \in \mathbb{R}^{n \times k}$, $C \in \mathbb{R}^{n \times k}$ be full col. rank matrices

Let $R = BB^T$ Then

1) $\mathcal{R}(R) = \mathcal{R}(B)$

2) $BB^T = CC^T$ implies $C = BQ$ for some orthogonal Q

下面看分解 $A=B^TB$, $A\in S^n$

$\Rightarrow$:

A is PSD $\Rightarrow A=V\Lambda V^T = V\Lambda^{1/2}\Lambda^{1/2}V^T$

$$= \underbrace{(V\Lambda^{1/2})}_{B^T}\underbrace{(V\Lambda^{1/2})^T}_{B}$$

$\Lambda^{1/2} = \mathrm{Diag}(\sqrt{\lambda_1}, \cdots \sqrt{\lambda_n})$

$\Leftarrow$:

if $A=B^TB$ then

$$x^TAx = x^TB^TBx = \|Bx\|_2^2 \geq 0.$$

$\therefore$ A is PSD.

$\square$

注意到 A 的那个分解.

$$A=V\Lambda V^T = V\Lambda^{1/2}\Lambda^{1/2}V^T$$

$$= \underbrace{(V\Lambda^{1/2}V^T)}_{B}\underbrace{(V\Lambda^{1/2}V^T)}_{B} = B^2$$

B is symmetric and PSD

更一般的

$$A = V\Lambda^{1/2}\Lambda^{1/2}V^T$$

$$= V\Lambda^{1/2}Q^TQ\Lambda^{1/2}V^T$$

$$= (V\Lambda^{1/2}Q^T)(V\Lambda^{1/2}Q^T)^T$$

可以不是正方形矩阵.

$$A = B^T B \qquad B = V\Lambda^{1/2}$$

$$= \underbrace{(QB)}_{=C}{}^T (QB) = C^T C$$

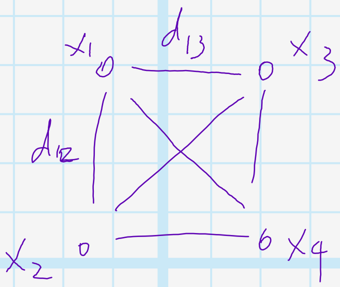
Suppose we have B, C, s.t.

$$A = B^T B = C^T C.$$

Is $C = QB$ for some orthogonal Q?

不同的分解之间能否通过一个Q来转换?

Application: Euclidean Distance Matrices



$$d_{ij} = \|x_i - x_j\|_2 \quad \forall i,j \in \{1,\dots,n\}$$

Let $\tilde{x}_i = \underbrace{Q}_{\text{Orthonormal}} x_i + b$

问题: 已知 d_{ij},

求 $X = [x_1, \dots, x_n]$

有 $d_{ij} = \|x_i - x_j\|_2$

$$= \|Q(x_i - x_j)\|_2$$

$$= \|Qx_i - Qx_j\|_2$$

$$= \|Qx_i + b - (Qx_j + b)\|_2$$

思路:

$$r_{ij} = d_{ij}^2 = \|x_i - x_j\|_2^2 = \|x_i\|_2^2 - 2x_i^T x_j + \|x_j\|_2^2$$

$$G = X^T X \to G_{ij} = x_i^T x_j$$

如果我们知道了G, 可以用分解的方法来求X

Suppose $G = \hat{X}^T \hat{X}$ Then $\hat{X} = QX$

Assume $x_1 = 0$ Then $r_{11} = 0 = \|x_1\|_2^2$

$$r_{12} = \|x_2\|_2^2 \cdots\cdots \quad r_{1n} = \|x_n\|_2^2$$

$$R = \mathbf{1} r^T - 2G + r \mathbf{1}^T$$

$$G = -\frac{1}{2}\left(R - \mathbf{1} r^T - r \mathbf{1}^T\right)$$

第一个例子: spectral Analysis. 2018.3.9

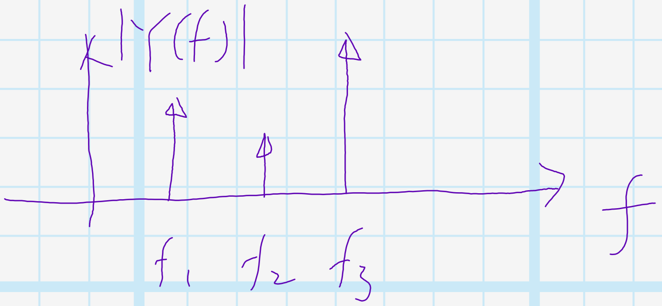
$$y_t = \sum_{i=1}^{k} \alpha_i e^{j2\pi f_i t} + V_t, \quad t = 0, 1, \cdots, T-1$$

where $\alpha_i \in \mathbb{C}$, $f_i \in [-\frac{1}{2}, \frac{1}{2})$

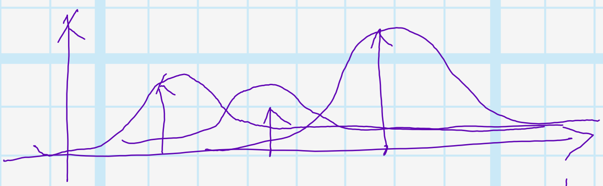
Fourier tranform $\quad f_i \neq f_j \ \forall i, j$

$$Y(f) = \sum_{t=-\infty}^{+\infty} y_t e^{-j2\pi f t} \quad f \in [-\frac{1}{2}, \frac{1}{2})$$

其实就是投影到指定频率上?



在实际情况下, 不可能取到无穷的时间,
那么把无穷变为了有限时间 变成了:



durintion 会大于 $\frac{1}{T}$

For convenience, $\alpha_i = 1 \ \forall i$

Let $z_i = e^{j2\pi f t}$, Then

$$y_t = \sum_{i=1}^{k} z_i^t + V_t, \quad t = 0, 1, \cdots, T-1$$

$$y_t = \begin{bmatrix} y_t \\ y_{t+1} \\ \vdots \\ y_{t+d-1} \end{bmatrix} \Big\updownarrow d = \sum_{i=1}^{k} \begin{bmatrix} y_i^t \\ y_i^{t+1} \\ \vdots \\ y_i^{t+d-1} \end{bmatrix} + V_t = \sum_{i=1}^{k} \overbrace{\begin{bmatrix} 1 \\ y_i \\ \vdots \\ y_i^{d-1} \end{bmatrix}}^{a_i} + V_t$$

$$= A s_t + V_t, \quad \text{where } s_t = \begin{bmatrix} y_1^t \\ y_2^t \\ \vdots \\ y_k^t \end{bmatrix}$$

$$A = [a_1, \cdots, a_k]_{k \times d}$$

Let $Y = [y_0, y_1, \cdots, y_{T_d}] = AS + V$ ← concaniate noise

$$S = [s_0, s_1, \cdots, s_{T_d}]$$

$$= \begin{bmatrix} 1 & z_1 & z_1^2 & \cdots & z_1^{T_d} \\ 1 & z_2 & z_2^2 & \cdots & z_2^{T_d} \\ \vdots & & & & \\ 1 & z_k & z_k^2 & \cdots & z_k^{T_d} \end{bmatrix}$$

由范德蒙矩阵性质 ↓ 有 full column rank.

If $d \geq k$, A has full col rank

If $T_d \geq k$, S has full row rank.

Step 1: find a basis V_2 from Y

s.t. $R(V_2) = R(A)^{\perp}$

step 2: find f_i's from V_2

Let $a(z) = \begin{bmatrix} 1 \\ z \\ \vdots \\ z^{d-1} \end{bmatrix}$

If $z = z_i$ for any $i \in \{1, \dots, k\}$

$$a(z)^H x = 0 \quad \forall x \in R(V_2)$$

$$\iff a(z)^H V_2 \alpha = 0 \quad \forall \alpha$$

$$\iff a(z)^H V_2 = 0$$

$$\iff \|V_2^H a(z)\|_2^2 = 0 \quad \iff \frac{1}{\|V_2^H a(z)\|_2^2} \to 0$$

Does it exist $z \notin \{z_1, \dots, z_k\}$

s.t.

$$a(z)^H x = 0 \quad \forall x \in R(V_2) = R(A)^\perp \, ?$$

The above is the same as

$$a(z) \in R(A)$$

That implies

$$a(z) = \sum_{i=1}^{k} \alpha_i a(z_i) \text{ for some } \alpha$$

That also imply

$$\underbrace{[\, a(z) \; a_1, a_2, \dots a_k \,]}_{d \times (k+1)} \text{ is LD}$$

$[a(z), a_1, \cdots, a_k]$ is Vandermonde with distinct roots. It is LD only if $d < k+1$

Not happening if we choose $d \geq k+1$

$$\min_{B,C} \|Y - BC\|_F^2$$

$$\hat{R}_y = \frac{1}{T_d} \sum_{t=0}^{T_d - 1} y_t y_t^H = \frac{1}{T_d} Y Y^H$$

Assume no noise, $v = 0$

$$R_y = \frac{1}{T_d} A S S^H A^H = A \underbrace{\left(\frac{1}{T_d} S S^H\right)}_{\Phi} A^H$$

Φ is nonsingular

$$x^H \Phi x = \frac{1}{T_d} \| S_x^H \|_2^2 > 0, \forall x \neq 0$$

$\Rightarrow \Phi$ is PD

$$\mathcal{R}(R_y) = \mathcal{R}(A)$$

$$\mathcal{R}(R_y) = \{ z \mid z = R_y x, x \in \mathbb{C}^d \}$$

$$= \{ z \mid z = A \Phi A^H x,\ x \in \mathbb{C}^d \}$$

$$= \{ z \mid z = A \Phi c,\ c \in R(A^H) \}$$

full row rank

$$= \{ z \mid z = A d,\ d \in \mathbb{C}^k \}$$

$$= R(A)$$

$\text{rank}(R_y) =$ no. of nozero eigenvalues.

EVD of R_y:

$$R_y = V \Lambda V^T$$

$$= [v_1, \cdots, v_k, v_{k+1}, \cdots, v_d] \begin{bmatrix} \lambda_1 & & & & & \\ & \ddots & & & & \\ & & \lambda_k & & & \\ & & & \lambda_{k+1} & & \\ & & & & \ddots & \\ & & & & & \lambda_d \end{bmatrix} \begin{bmatrix} v_1^H \\ \vdots \\ v_k^H \\ v_{k+1}^H \\ \vdots \\ v_d^H \end{bmatrix}$$

Assume $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_k \cdots \geq \lambda_d$

$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_k > 0$

$\lambda_{k+1} = \cdots = \lambda_d = 0$

$$= [V_1 \quad V_2] \begin{bmatrix} \Lambda_1 & \\ & 0 \end{bmatrix} \begin{bmatrix} V_1^H \\ V_2^H \end{bmatrix}$$

$$= V_1 \Lambda_1 V_1^H$$

We have $R(R_y) = R(V_1)$

So. $R(R_y)^{\perp} = R(V_2)$

Assume $V \neq 0$, but is i.i.d.

$$R_y = \frac{1}{T_d} Y Y^H = \frac{1}{T_d} (AS+V)(AS+V)^H$$

$$= \frac{1}{T_d} (ASS^H A^H + VS^H A^H + ASV^H + VV^H)$$

If $T_d \to \infty$,

$$R_y = A \Phi A^H + \sigma^2 I$$

$$R_y = V \Lambda V^H + \sigma^2 I$$

$$= V(\Lambda + \sigma^2 I) V^H$$

V is the ematrix of $A \Phi A^H$

V is also the ematrix of R_y.

2018.3.15

Singular Value Decomp (SVD)

Any $A \in \mathbb{R}^{m\times n}$ can be decomposed as

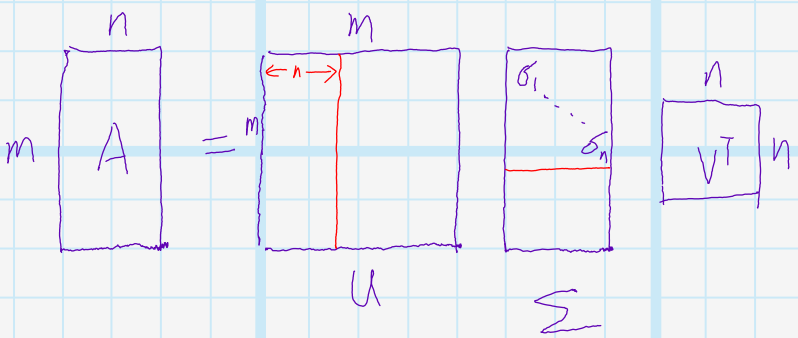
$$A = \underset{m\times m}{U}\,\underset{m\times n}{\Sigma}\,\underset{n\times n}{V^T}$$

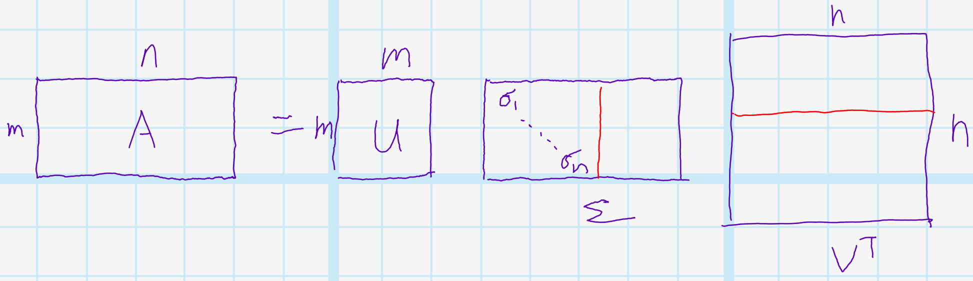
where U, V are ortho, Σ is sort of diagonal w/

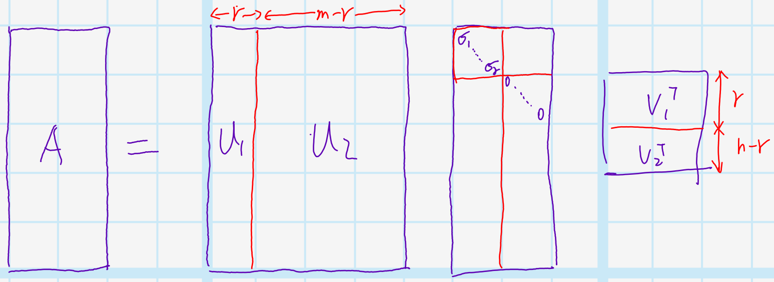
$$[\Sigma]_{ii} = \sigma_i, \quad \sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_{\min(m,n)} \geq 0$$

$$[\Sigma]_{ij} = 0 \quad \forall i \neq j$$

对高矩阵：







可以写为更简单的形式：

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

Properties

a) $R(A) = R(U_1)$

Proof: $R(A) = \{ y = U_1 \tilde{\Sigma} V_1^T x \mid x \in \mathbb{R}^n \}$

$= \{ y = U_1 \tilde{\Sigma} z \mid z \in R(V_1^T) \}$

$= R(U_1)$

b) $R(A)^\perp = R(U_2)$

Proof: $R(A) = R(U_1)$

$R(U_1)^\perp = \{ y \in \mathbb{R}^m \mid y^T U_1 x = 0, \ \forall x \in \mathbb{R}^r \}$

Any $y \in \mathbb{R}^n$ can be written as

$y = U\alpha, \ \alpha \in \mathbb{R}^m$

$= [U_1 \vdots U_2] \begin{bmatrix} \alpha_1 \\ \hdashline \alpha_2 \end{bmatrix}$

$= U_1\alpha_1 + U_2\alpha_2$

$$0 = y^T U_1 x = (U_1\alpha_1 + U_2\alpha_2)^T U_1 x$$

$$= \alpha_1^T \underbrace{U_1^T U_1}_{=I} x + \alpha_2^T \underbrace{U_2^T U_1}_{=0} x$$

$$= \alpha_1^T x$$

$$0 = \alpha_1^T x \quad \forall x \in \mathbb{R}^r,$$

$$\Leftrightarrow \alpha_1 = 0$$

$$\therefore R(U_1)^{\perp} = \{ y = U_2\alpha_2 \mid \alpha_2 \in \mathbb{R}^{m-r} \} = R(U_2)$$

$\boxtimes$

c) $R(A^T) = R(V_1)$

d) $R(A^T)^{\perp} = N(A) = R(V_2)$ 这个等式给出了如何求一个矩阵A的核空间，通过SVD.

e) $\text{rank}(A) = \dim R(A) = r$ 这个等式直接可以证明行秩等于列秩

$\text{rank}(A^T) = \text{ran}(A)$

$\dim R(A^T) = \dim R(V_1) = r.$

还可以用来证明 $\dim N(A) = n - \text{rank}(A)$

即 $\dim R(V_2) = n - r$

$= n - \text{rank}(A)$

用SVD来理解求解线性方程组.

$$Ax=y \qquad A\in\mathbb{R}^{m\times m}$$

$$U\Sigma V^T x = y$$

$$x = V\Sigma^{-1}U^T y$$

SVD是 schur decomp. 的 by product.

Suppos $A = U\Sigma V^T$

Let $R = AA^T$ $\because R$ is PSD,

$\therefore R = U\Sigma V^T V \Sigma^T U^T$ 注意 $\Sigma \neq \Sigma^T$

$$= U\Sigma\Sigma^T U$$

当 $m\geq n$ 时 当 $m\leq n$ 时

$$\Sigma\Sigma^T = \begin{bmatrix} \sigma_1^2 & & & & \\ & \ddots & & & \\ & & \sigma_n^2 & & \\ & & & 0 & \\ & & & & \ddots \\ & & & & & 0 \end{bmatrix} \qquad \Sigma\Sigma^T = \begin{bmatrix} \sigma_1^2 & & \\ & \ddots & \\ & & \sigma_m^2 \end{bmatrix}$$

Proof sketch of SVD:

Do Eigen Decomposition

$$R = U\Lambda U^T$$

Assume R is PD.

$$\Sigma = \mathrm{Diag}(\sqrt{\lambda}, \cdots, \sqrt{\lambda_m})$$

Let $V = A^T U \Sigma^{-1}$

We have $U\Sigma V^T = U\Sigma(\Sigma^{-1}U^TA)$

$$= A$$

2018.3.16

Notation:

- $A, B \in S^n$
- $\lambda_1(A) \geq \lambda_2(A) \geq \cdots \geq \lambda_n(A)$ denote evalues of A
- if not specified,

  $\lambda_1 \geq \cdots \geq \lambda_n$ are evalues of A

  $\mu_1 \geq \cdots \geq \mu_n$ are evalues of B

  V, U are eigen matrices of A, B resp.

Questions for Mat People:

- How are $\lambda_k(A+B)$, $\lambda_k(A)$, $\lambda_k(B)$ related?
- Suppose B is noise. Can we bound

$$|\lambda_k(A+B) - \lambda_k(A)| \leq \max\{|\lambda_1(B)|, |\lambda_n(B)|\}$$

$$\leq \|B\|_F$$

$$\|B\|_F^2 = tr(B^TB) = tr(BB) = tr(UDU^TUDU^T) = tr(U|D|^2U^T) = tr(|D|^2)$$

- It is true $\sum_{i=1}^n a_{ii} = \sum_{i=1}^n \lambda_i$. Do we have

$$\sum_{i=1}^{r} a_{ii} \le \sum_{i=1}^{r} \lambda_i$$

where $1 \le r \le n$

$\rightarrow tr(A) = tr(V \Lambda V^T)$
$= tr(V^T V \Lambda)$
$= tr(\Lambda)$

Eigenvalue Inequalities:

- $\lambda_k(A) + \lambda_n(B) \le \lambda_k(A+B) \le \lambda_k(A) + \lambda_1(B)$
  $1 \le k \le n$ B的最小特征值 B的最大特征值 (Weyl)

- if $rank(B) = 1$, $\lambda_{k+1}(A) \le \lambda_k(A+B) \le \lambda_{k-1}(A)$

- $\lambda_{j+k-1}(A+B) \le \lambda_j(A) + \lambda_k(B)$ (Weyl)

- Let $A = \begin{array}{cc} & \overset{\leftarrow n-1 \rightarrow}{} \\ n-1 & \begin{bmatrix} B & \times \\ \times & \times \end{bmatrix} \end{array}$ $\lambda_{k+1}(A) \le \lambda_k(B) \le \lambda_k(A)$ (Cauchy)
  可继续推论：$\lambda_1(A) \ge \lambda_1(B) \ge \lambda_2(A) \ge \lambda_2(B) \ge \dots$

- Let $U \in \mathbb{R}^{n \times r}$ be semi-ortho $\lambda_{k+n-r}(A) \le \lambda_k(U^T A U) \le \lambda_k(A)$

- $tr(AB) \le \sum_{i=1}^{n} \lambda_i(A) \lambda_i(B)$

- for any $1 \le i_1 \le \dots \le i_k \le n$, (Lidski)
$$\sum_{k=1}^{k} \lambda_{i_j}(A+B) \le \sum_{j=1}^{k} \lambda_{i_j}(A) + \sum_{j=1}^{k} \lambda_j(B)$$

应用例子：

假设A是一个信号，如果在上面加一个噪声信号B

那么如果假设 $\mathrm{rank}(B)=1$，那么A的第k个

特征值的变化不会超过 $\lambda_{k-1}(A)$ 和 $\lambda_{k+1}(A)$

$$U=[e_1, e_2 \cdots e_r]$$

$$\begin{aligned}
tr(U^TAU) &= \sum_{i=1}^{r} [U^TAU]_{ii} \\
&= \sum_{i=1}^{r} e_i^T U^TAU e_i \\
&= \sum_{i=1}^{r} u_i^T A u_i \\
&= \sum_{i=1}^{r} a_{ii}
\end{aligned}$$

$$\sum_{i=1}^{r} a_{ii} = tr(U^TAU) = \sum_{i=1}^{r} \lambda_i(U^TAU) \le \sum_{i=1}^{r} \lambda_i(A)$$

Rayleigh-Ritz：

$$\lambda_n \|x\|_2^2 \le x^TAx \le \lambda_1 \|x\|_2^2$$

$$\lambda_1 = \max_{\|x\|_2=1} x^TAx \qquad \lambda_n = \min_{\|x\|_2=1} x^TAx$$

Courant Fischer: let S_r be a subsp. with dim r,

$$\lambda_k = \min_{S_{n-k+1} \subseteq \mathbb{R}^n} \max_{\substack{x \in S_{n-k+1},\\ \|x\|_2 = 1}} x^T A x = \max_{S_k \subseteq \mathbb{R}^n} \min_{\substack{x \in S_k,\\ \|x\|_2 = 1}} x^T A x$$

$$\max_{\substack{U \in \mathbb{R}^{n \times r}\\ U^T U = I}} tr(U^T A U) = \sum_{i=1}^{r} \lambda_i(A)$$

证: (Rayleigh-Ritz)

$$x^T A x = x^T V \Lambda \underbrace{V^T x}_{=y} \qquad \text{y只是x的旋转.}$$

$$= y^T \Lambda y$$

$$= \underbrace{\sum_{i=1}^{n} \lambda_i |y_i|^2}$$

$$\lambda_n \sum_{i=1}^{n} |y_i|^2 \le \downarrow \le \lambda_1 \sum_{i=1}^{n} |y_i|^2$$

$$\Downarrow$$

$$\lambda_1 \|y\|_2^2 = \lambda_1 \|V^T x\|_2 = \lambda_1 \|x\|_2^2$$

For all x have $\|x\|_2 = 1$

$$x^T A x \le \lambda_1$$

But is there a unital-norm x that attains $x^T A x = \lambda_1$?

存在x可以取到极值.

Yes, $y = e_1 \iff x = v_1$

$y = V^T x \quad e_1 = V^T x$

$v_1 = Ve_1 = VV^Tx = x$

上面研究了λ_1和λ_n，那么$\lambda_2 \cdots \lambda_{n-1}$怎么办？

我们来看如何用优化问题来找λ_k，

我们希望$\lambda_1, \cdots \lambda_{k-1}$都消失，那么$\lambda_k$此时最大

$$\lambda_k = \max_{\substack{x \in \text{span}\{v_k, \cdots, v_n\} \\ \|x\|_2 = 1}} x^T A x$$

限制了取不到$v_1, \cdots v_{k-1}$

$\dim(\text{span}\{v_k, \cdots, v_n\}) = n-k+1$

下面看C-F

$$\min_{S_{n-k+1} \subseteq \mathbb{R}^n} \underbrace{\max_{x \in S_{n-k+1}, \|x\|_2 = 1} x^T A x}_{\text{是一个关于 } S_{n-k+1} \text{ 的函数, } f(S_{n-k+1})}$$

找一个合适的subspa. s.t. $f(S_{n-k+1})$最小。

但这个问题是属于流形上的优化？

证明放到后面，现在用这个 C-F 来证明一个 eigen Inequality.

$$\lambda_k(A+B) \leq \lambda_k(A) + \lambda_1(B)$$

$$\lambda_k(A+B) = \min_{S_{n-k+1} \subseteq \mathbb{R}^n} \max_{\substack{x \in S_{n-k+1} \\ \|x\|_2 = 1}} x^T(A+B)x$$

注意到 $x^T(A+B)x$

$$= x^TAx + x^TBx$$

$$= x^TAx + \lambda_1(B)\|x\|_2^2 \quad (\text{Releigh})$$

$\therefore$ 原式 $\leq \min\limits_{S_{n-k+1} \subseteq \mathbb{R}^n} \max\limits_{\substack{x \in S_{n-k+1} \\ \|x\|_2 = 1}} x^TAx + \lambda_1(B)$

和 x 无关，甩出去.

$$\leq \lambda_1(A) + \lambda_1(B)$$

直接套 C-F 那个等式.

min max 和 max min 分别用来证右边和左边.

$$\lambda_k = \max_{\substack{x \in \text{span}\{v_k, \cdots, v_n\} \\ \|x\|_2 = 1}} x^T A x$$

$$\min_{S_{n-k+1}} \max_{\substack{x \in S_{n-k+1} \\ \|x\|_2 = 1}} x^T A x \le \lambda_k .$$

Let $V = \{ y = U^T x \mid x \in S_{n-k+1} \}$, $\dim V = n-k+1$

$W = \text{span}\{e_1, \cdots, e_k\}$ $\dim W = k$

$y \in W \Rightarrow y_{k+1} = y_{k+2} \cdots = y_n = 0$

$$\min_{S_{n-k+1}} \max_{\substack{x \in S_{n-k+1} \\ \|x\|_2 = 1}} x^T A x$$

$$= \min_{\substack{V \subseteq \mathbb{R}^n \\ \dim V = n-k+1}} \max_{\substack{y \in V \\ \|y\|_2 = 1}} \sum_{i=1}^n \lambda_i |y_i|^2$$

$$\geq \min_{\substack{V \in \mathbb{R}^n \\ \dim V = n-k+1}} \max_{\substack{y \in V \cap W \\ \|y\|_2 = 1}} \sum_{i=1}^n \lambda_i (y_i)^2$$

$$= \sum_{i=1}^n \lambda_i \|y\|^2$$

$$\geq \lambda_k \|y\|_2^2$$

$$\dim (U \cap W) = \dim U + \dim W - \dim (U + W)$$

$\geq n - h + 1 + k - n = 1$

2018.3.22

Recall

$$\lambda_1(A) = \max_{\substack{\|x\|_2=1 \\ x\in\mathbb{R}^n}} x^TAx, \quad \text{attained when } x=v_1$$

$$\lambda_2(A) = \max_{\substack{x\in \text{Span}\{u_2\cdots u_n\} \\ =\text{Span}(v_1)^\perp \\ \|x\|_2=1}} x^TAx, \quad \text{attained when } x=v_2.$$

Problem:

Solve

$$\max_{\substack{U\in\mathbb{R}^{n\times r} \\ U^TU=I}} tr(U^TAU) = \sum_{i=1}^{r} u_i^TAu_i$$

这是上面那个 Colant-Fisher 的更general的形式.

Example: $r=2$ 可转化为如下的 nested optimization problem

$$\max_{\|u_1\|_2=1} u_1^TAu_1 + \max_{\substack{u_2^Tu_1=0 \\ \|u_2\|_2=1}} u_2^TAu_2$$

下面来证明.

$$\lambda_k(U^TAU) \le \lambda_k(A)$$

U - semi-ortho

$$tr(U^TAU) = \sum_{i=1}^{r}\lambda_i(U^TAU) \le \sum_{i=1}^{r}\lambda_i(A)$$

choose $U=[v_1,\cdots,v_r]$

$$U^T A U = \begin{bmatrix} \lambda_1(A) & & \\ & \ddots & \\ & & \lambda_r(A) \end{bmatrix}$$

$$tr(\cdots) = \sum_{i=1}^{r} \lambda_i(A)$$

Matrix Inequalities.

例：

$$x = \begin{bmatrix} 0.9 \\ 0.7 \\ 0.6 \end{bmatrix} \begin{array}{l} \leftarrow \text{score of bio} \\ \leftarrow \text{score of physics} \\ \leftarrow \text{score of music} \end{array}$$

x是一个学生的成绩.

如何比较两个学生谁优秀？

$x \geq y$ means $x_i - y_i \geq 0, \ \forall i$

$x \ngeq y$ means $x_i - y_i \geq 0$ is not true for some i.

在矩阵中，

$A \succeq B$ means that $A - B$ is PSD

Schur Complement.

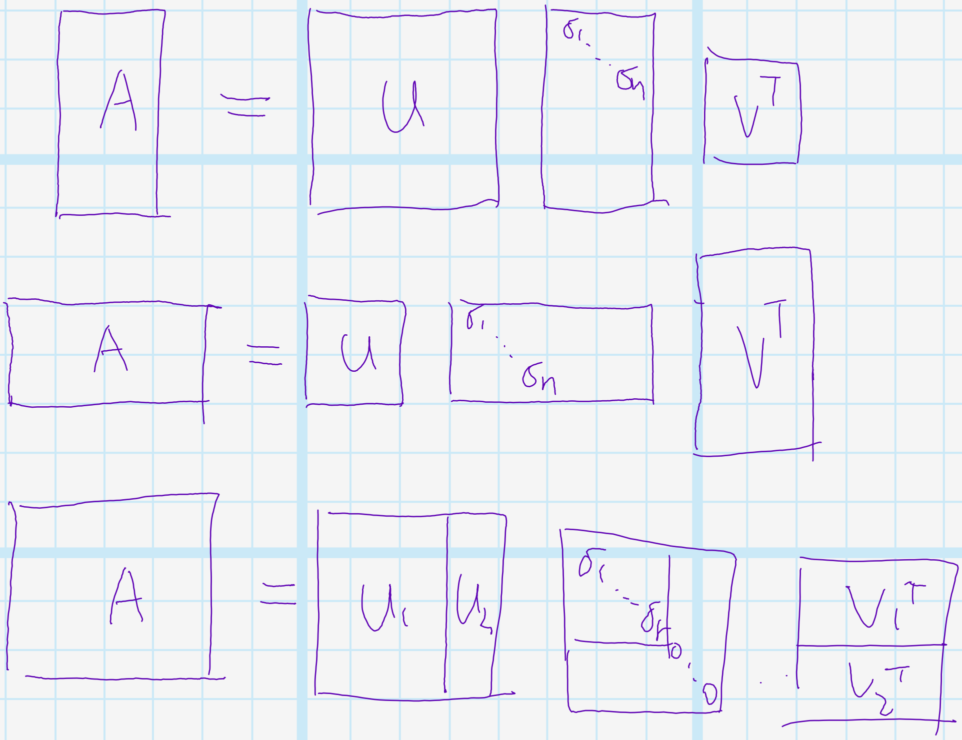
$$X = \begin{bmatrix} 1 & b^T \\ b & C \end{bmatrix} \succeq 0 \iff 1 - b^T C^{-1} b \geq 0$$

$$\Updownarrow$$

$$\begin{bmatrix} C & b \\ b^T & 1 \end{bmatrix} \succeq 0 \iff C - b 1^{-1} b^T \succeq 0$$

2018.3.23

SVD: $A = U \Sigma V^T$



$= U_1 \tilde{\Sigma} V_1^T$

Note: $\sigma_1 \geq \sigma_2 \geq \cdots \sigma_{\min\{m,n\}}$

rank(A) = no. of +ve σ_i's

$\lambda_i(A^T A) = \sigma_i(A)^2$ (for some "principal" i's)

Linear Sys.

$$y = Ax = U \Sigma \underbrace{V^T x}_{\tilde{x}} = U \underbrace{\Sigma \tilde{x}}_{\tilde{y}} = U \tilde{y}$$

这个相当于:

$$x \longrightarrow \boxed{V^T} \xrightarrow{\tilde{x}} \boxed{\Sigma} \xrightarrow{\tilde{y}} \boxed{U} \longrightarrow y$$

$m \geq n \quad \tilde{y} = \Sigma \tilde{x} \iff \begin{bmatrix} \tilde{y}_1 \\ \tilde{y}_2 \\ \vdots \\ \tilde{y}_n \\ \tilde{y}_{n+1} \\ \vdots \\ \tilde{y}_m \end{bmatrix} = \begin{bmatrix} \sigma_1 \tilde{x}_1 \\ \sigma_2 \tilde{x}_2 \\ \vdots \\ \sigma_n \tilde{x}_n \\ 0 \\ \vdots \\ 0 \end{bmatrix}$

$m \leq n: \quad \tilde{y} = \Sigma \tilde{x} \iff \begin{bmatrix} \tilde{y}_1 \\ \vdots \\ \tilde{y}_m \end{bmatrix} = \begin{bmatrix} \sigma_1 \tilde{x}_1 \\ \vdots \\ \sigma_m \tilde{x}_m \end{bmatrix}$

Example: A is square no singular.

$$y = U \Sigma V^T x$$

$$U^T y = \Sigma V^T x$$

$$\Sigma^{-1} U^T y = V^T x$$

$$V \Sigma^{-1} U^T y = x$$

$\longrightarrow \text{Diag}\left\{\frac{1}{\sigma_1}, \dots, \frac{1}{\sigma_n}\right\}$

Matrix Norms:

$$\|A\|_F = \sqrt{\sum_i \sum_j |a_{ij}|^2} = tr(A^T A)^{1/2}$$

$$= tr\left((U\Sigma V^T)^T (U\Sigma V^T) \right)^{1/2}$$

$$= tr\left(V\Sigma^T \underbrace{U^T U}_{= I} \Sigma V^T \right)^{1/2}$$

$$= tr\left(\Sigma^T \Sigma \right)^{1/2}$$

$$= \left(\sum_{i=1}^{\min(m,n)} \sigma_i^2 \right)^{1/2}$$

Induced norm

$$f(A) = \max_{\|x\|_\beta \le 1} \|Ax\|_\alpha$$

控制输入信号的能量

matrix 2-norm (spectral norm)

$$\|A\|_2 = \max_{\|x\|_2 \le 1} \|Ax\|_2 = \sigma_1$$

$$\max_{\|x\|_2 \le 1} \|Ax\|_2^2 = \max_{\|x\|_2 \le 1} x^T A^T A x = \lambda_1 (A^T A) = \sigma_1(A)^2$$

Attained $x = v_1$

扰动对矩阵的影响，

定义条件数 $\kappa(A) = \dfrac{\sigma_{max}(A)}{\sigma_{min}(A)}$

$$\frac{\|\hat{x}-x\|_2}{\|x\|_2} \le 4\varepsilon\kappa(A)$$ 给出了 bound.

Pseudo-Inverse

$$A^{+} = V_1 \Sigma^{-1} U_1^{T}$$

对于普通的逆，有 $AB=I$, $BA=I$

但对于伪逆 可能只有一个式子成立

$$x = A^{+}y + n$$

null space component

2018.4.6

Let $A \in \mathbb{R}^{m\times n}$ If we write $\text{Krank}(A)=k$, it means that for any $\{i_1, \cdots i_l\} \subseteq \{1, \cdots, n\}$ with $l \le k$,

$$\{a_{i_1}, \cdots, a_{i_l}\} \text{ is linear independent}$$

Problem: Suppose $y = Ax$ and

$$\text{krank}(A) \geq 2\|x\|_0$$

If x' is such that $y = Ax'$, then $\|x'\|_0 > \|x\|_0$

注：就是任意已们中的那个，矩阵 A，任选 k 列，它们都线性独立。也可以记作：

$$\text{spark}(A) = \text{krank}(A) - 1$$

通常 A 是一个 fat matrix.

例：$A = [0 \quad \underbrace{a_2, \cdots a_n}_{LI}]$ linear independent

rank $= n-1$ krank $= 0$

可见 krank 更严格.

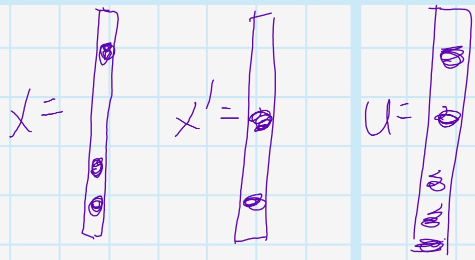
$$\min_{\bar{x}} \|\bar{x}\|_0$$

$$\text{s.t. } y = A\bar{x}$$

$y = Ax$

$y = Ax'$

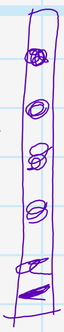
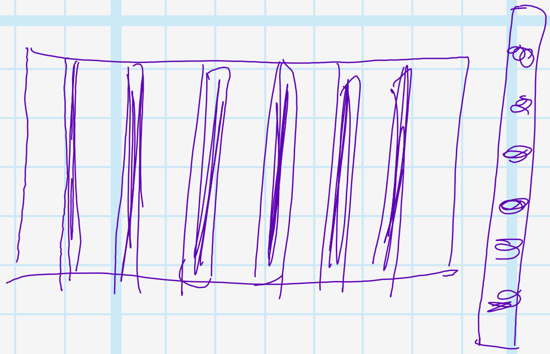
$\Rightarrow A\underbrace{(x - x')}_{=u} = 0$



We have $\|u\|_0 \le \|x\|_0 + \|x'\|_0$

Suppose $\|x'\|_0 < \|x\|_0$. Then

$\|u\| \le 2\|x\|_0$

$u =$  $\quad 0 = Au =$ 

$\|u\|_0 \le 2\|x\|_0 = 6$

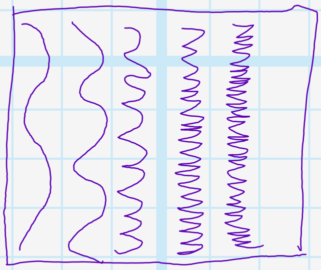
Then, let $\text{supp}(u) = \{ i \in \{1, \cdots, n\} \mid u_i \neq 0 \}$

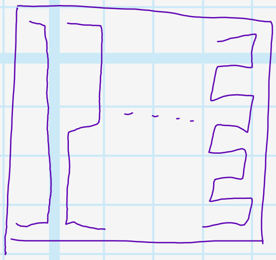
By letting $\text{supp}(u) = \{ i_1, \cdots, i_k \}$, $k = \text{card}(\text{supp}(u)) \le 2\|x\|_0$

$Au = 0 \Leftrightarrow a_{i_1}, \cdots, a_{i_k}$ is linearly dependent

傅立叶变换

傅立叶基：　　　　Harr 基





下面来看如何求解开始提的问题.

Brute-Force Search

Example $A=[a_1, a_2, a_3]$

1
- try if $\exists x_1$, s.t. $a_1x_1=y$
- ...... x_2 ...... $a_2x_2=y$
- x_3 ...... $a_3x_3=y$

2
- x_1, x_2 ...... $a_1x_1+a_2x_2=y$
- x_2, x_3 ...... $a_2x_2+a_3x_3=y$
- $\vdots$

3
- x_1, x_2, x_3, s.t. $a_1x_1+a_2x_2+a_3x_3=y$

让x的非0元素个数不断增多，看是否有解.

2018.4.12

$y = Ax_0 \longrightarrow x = A^{+}y$ if A has full col rank

加入噪声

$$y = Ax_0 + v \longrightarrow \min_x \|y - Ax\|_2^2$$

$$\Rightarrow x_{LS} = (A^T A)^{-1} A^T y$$

$$x_{RLS} = (A^T A + \lambda I)^{-1} A^T y$$

$$\min_x \|y - Ax\|_2^2 + \lambda \|x\|_2^2$$ 加一下稀疏限制

现在考虑新的问题：

$$y = (A + E)x_0 + v$$

uncertainty 在A上存在误差.

$\Downarrow$

$$\min_{x,B} \|y - Bx\|_2^2 + \|B - A\|_F^2 \quad (TLS)$$

另个问题：

$$\min_x \max_{\Delta \in \mathcal{U}} \|y - (A + \Delta)x\|_2^2 \quad (\text{Robust LS})$$

uncertainty set

例：类似于GAN网络，每次给一个x，我把干扰最大，使得在最差环境下效果最好

$$\mathcal{U}=\{\Delta \mid \|\Delta\|_2 \le \lambda\}$$

$$\max_{\Delta\in\mathcal{U}} \|y-(A+\Delta)x\|_2 = \|y-Ax\|_2 + \lambda\|x\|_2$$

GAN 的做法：先放大 Δ，把目标变大，再调 x 让目标变小，反复迭代。

$$\max_{\Delta\in\mathcal{U}} \|y-(A+\Delta)x\|_2 = \|y-Ax\|_2 + \lambda\|x\|_2$$

证：
$$\|y-(A+\Delta)x\|_2 \le \|y-Ax\|_2 + \|\Delta x\|_2$$
$$\le \|y-Ax\|_2 + \|\Delta\|_2\,\|x\|_2$$
$$\le \|y-Ax\|_2 + \lambda\|x\|_2$$

$$\Delta = -\lambda\frac{u}{\|u\|_2}\cdot\frac{x^T}{\|x\|_2}$$

$$\Delta = ux^T$$
$$\Delta x = ux^Tx = \|x\|_2^2\, u$$

$$\Delta x = -\lambda\frac{u}{\|u\|_2}\|x\|_2$$

$$\|\Delta x\|_2 = \lambda\|x\|_2$$

TLS 可以用 SVD 来求解。

$$\min_{B,x} \|[A\ \ y] - [B\ \ Bx]\|_F^2$$

$m\times(n+1)$

$m \ge n+1$

令 $C = [A\ y]$, $D = [B, Bx]$

$$D = [B\ \ Bx] \Leftrightarrow D\begin{bmatrix} x \\ -1 \end{bmatrix} = 0 \Rightarrow \operatorname{rank}(D) \le n$$

relaxition

$$\min_D \|C - D\|_F^2$$

$$\text{s.t. } \operatorname{rank}(D) \le n$$

就是在约束之下，把误差降到最小。

2018.4.13

Example: Do Gauss Elimination with

$$3x_1 + 5x_2 = 9 \qquad (a)$$

$$6x_1 + 7x_2 = 4 \qquad (b)$$

Solve $Ax = b$, $A \in \mathbb{R}^{n \times n}$ being nonsingular

$$x = A^{-1} b$$

computing A^{-1} by $\underbrace{\text{Crammer's}}_{O(\frac{2}{3}n^3)}$ rule is expensive

下面看如何求解上面那个Example.

(a) × 2 − (b):

$$3x_2 = 14 \rightarrow x_2 = 14/3$$

Subs. $x_2 = 14/3$ into (a):

$$3x_1 + 5 \times 14/3 = 9 \rightarrow x_1 = \cdots$$

$$\begin{bmatrix} 6 & 7 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 4 \\ 14 \end{bmatrix}$$

一般求解方式是

- $A = LU$

$$Ax = b \iff LUx = b$$

$$\iff \left.\begin{array}{l} Lz = b \\ Ux = z \end{array}\right\} O(2n) = O(n)$$

变成求解两个上(下)三角矩阵.

Gauss Transform.

$$\underbrace{\begin{array}{c} \\ k \to \\ k+1 \to \\ \\ \end{array}\begin{bmatrix} 1 & & & & & \\ & \ddots & & \vdots & & \\ & & 1 & & & \\ & & -\tau_{k+1} & 1 & & \\ & & \vdots & & \ddots & \\ & & -\tau_n & & & 1 \end{bmatrix}}_{M} \begin{bmatrix} x_1 \\ \vdots \\ x_k \\ x_{k+1} \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_k \\ x_{k+1} - \tau_{k+1} x_k \\ \vdots \\ x_n - \tau_n x_k \end{bmatrix} = \begin{bmatrix} x_1 \\ \vdots \\ x_k \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

choose $\tau_{k+1} = \dfrac{x_{k+1}}{x_k}, \cdots, \tau_n = \dfrac{x_n}{x_k}$

$$My = \begin{bmatrix} y_1 \\ \vdots \\ y_k \\ x \\ \vdots \\ x \end{bmatrix}$$

下面是LU分解

Stage 1: find M_1 such that $Ma_1 = \begin{bmatrix} a_{11} \\ 0 \\ \vdots \\ 0 \end{bmatrix}$

$$A^{(1)} = M_1 A = M_1 [a_1, \cdots, a_n]$$

$$= [M_1 a_1, M_1 a_2, \cdots, M_1 a_n]$$

$$= \begin{bmatrix} a_{11} & x & \cdots & x \\ 0 & \vdots & & \vdots \\ \vdots & \vdots & & \vdots \\ 0 & x & \cdots & x \end{bmatrix}$$ x表示不care

$$A^{(1)} = \begin{bmatrix} a_{11}^{(1)} & a_{12}^{(1)} & \cdots & a_{1n}^{(1)} \\ 0 & a_{22}^{(1)} & & \vdots \\ \vdots & \vdots & & \vdots \\ 0 & a_{2n}^{(1)} & \cdots & a_{nn}^{(1)} \end{bmatrix}$$

Stage 2: find M_2 such that

$$M_2 a_2^{(1)} = \begin{bmatrix} a_{12}^{(1)} \\ a_{22}^{(1)} \\ 0 \\ \vdots \\ 0 \end{bmatrix} \qquad a_{22}^{(1)} \neq 0$$

$$A^{(2)} = M_2 A^{(1)} = \begin{bmatrix} a_{11}^{(1)} & a_{12}^{(1)} & x & \cdots & x \\ 0 & a_{22}^{(1)} & \vdots & & \vdots \\ \vdots & 0 & \vdots & & \vdots \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & x & \cdots & x \end{bmatrix}$$

$$\vdots$$

$$U = A^{(n-1)} = M_{n-1} A^{(n-2)} = \begin{bmatrix} a_{11}^{(1)} & a_{12}^{(1)} & a_{13}^{(1)} \\ 0 & a_{22}^{(1)} & a_{23} \\ \vdots & & \\ 0 & & \end{bmatrix}$$

$$U = \underbrace{M_{n-1} M_{n-2} \cdots M_1}_{M} A$$

if M is invertable

$$M^{-1} U = A$$

补充下几个性质

if A, B is lower Δ,

$C = AB$ is lower Δ with

$c_{ii} = a_{ii} b_{ii} \quad \forall i$.

if A is lower Δ, $\det(A) = \prod_{i=1}^{n} a_{ii}$

if A is lower Δ, and nonsingular $\Leftrightarrow a_{ii} \neq 0$

Then, $B = A^{-1}$ is lower Δ with $b_{ii} = 1/a_{ii}$

证: $AB = I \Leftrightarrow A[b_1, \cdots b_n] = [e_1, \cdots, e_n]$

$\Leftrightarrow A b_i = e_i, \; i = 1, \cdots, n$

$$\begin{bmatrix} a_{11} & & & \\ a_{21} & a_{22} & & \\ \vdots & & \ddots & \\ a_{n1} & & & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

$$x_1 = \frac{1}{a_{11}}$$

$$x_2 = \frac{1}{a_{22}}(-a_{21}x_1)$$

$$\vdots$$

$\leftarrow$ 解出 b_1

接着算 b_2：

$$Ax = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

$$a_{11}x_1 = 0 \Rightarrow x_1 = 0$$

$$a_{22}x_2 = 1 \Rightarrow x_2 = \frac{1}{a_{22}}$$

$$\vdots$$

根据这些初元的性质，

$$U = \underbrace{M_{n-1}M_{n-2}\cdots M_1}_{\text{Lower } \Delta \text{ with diagonals} = 1, \ldots, 1} A$$

下面看 LU 分解的存在条件，

由上面可知 $a_{22}^{(1)} \neq 0$，下面来刻画这个，

例如：

$$\begin{bmatrix} a_{11}^{(1)} & a_{12}^{(1)} \\ 0 & a_{22}^{(1)} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ x & 1 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

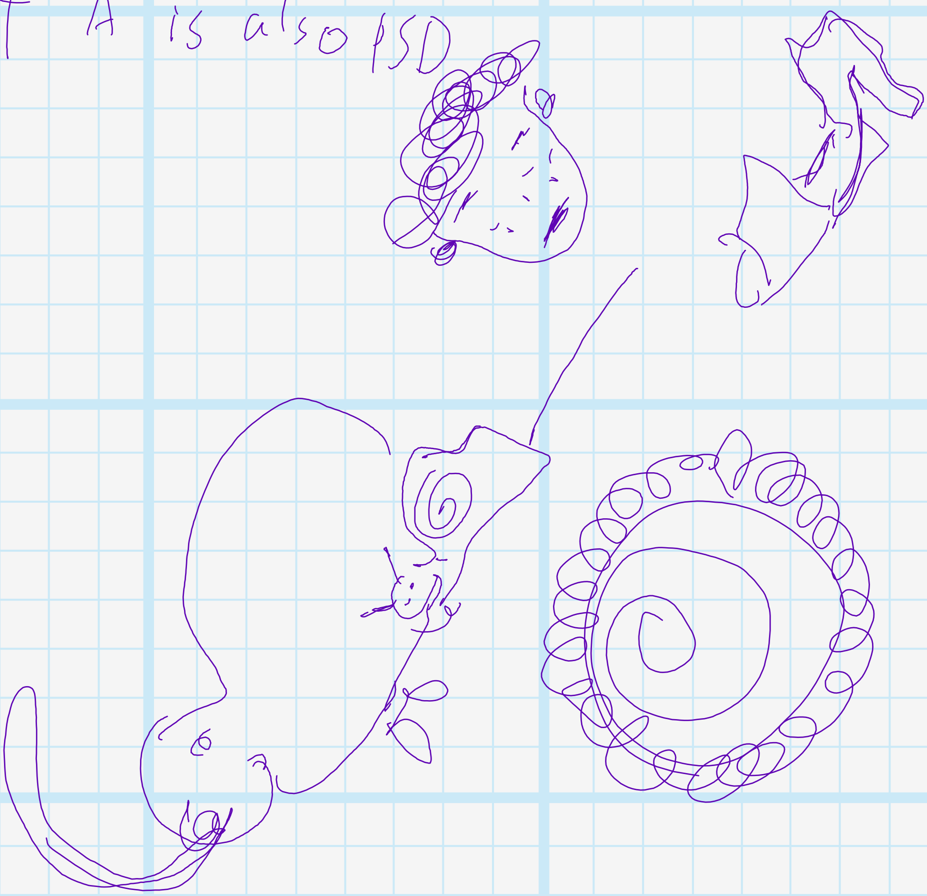
$$\downarrow$$

$$\det(\cdot) = \det(\cdot)\det(\cdot)$$

$$= a_{11}a_{22}^{(1)} \qquad \overset{\shortparallel}{1} \qquad \overset{\shortparallel}{\det(A_{\{1,2\}})}$$

$a_{11} a_{22}^{(1)} = \det(A_{\{1,2\}})$

当A是PSD时, all principal submatrix of A is also PSD



Compute $B = A^{-1}$

$AB = I \iff Ab_i = e_i, \ i = 1, \cdots, n$

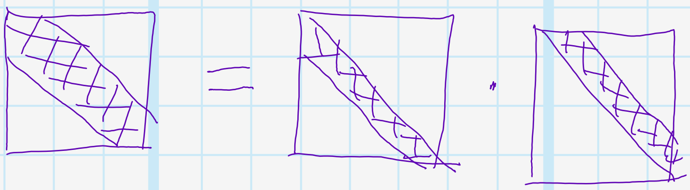
$\iff LUb_i = e_i, \ i = 1, \cdots, n$

· Compute $\det(A)$,

$\det(A) = \det(L) \det(U) = 1 \times \prod_{i=1}^{n} u_{ii}$

· Compute LU for huge A:

- sparse A ⇏ sparse (L, U)



对 $A \in S^n$ $A = LDL^T$ $O(\frac{1}{3}n^3)$

$$A^T = (LDL^T)^T$$
$$= LD^TL^T$$
$$= LDL^T$$

if $A \in S^n$ is PD

$$d_1 \cdots d_n > 0$$

$$A = \underbrace{LD^{\frac{1}{2}}}_{G} D^{\frac{1}{2}} L^T$$
$$= GG^T$$

这就是乔里斯基分解 cholesky

在 matlab 中, 用 x=linsolve(A, b, opt)

$\uparrow$ 矩阵特征

判断一个矩阵是否是正定, 可以直接在matlab中调用 cholesky, 如果报错就不正定.

Jacobi and Gauss-Seidel (GS)

$$Ax = b$$

$\Leftrightarrow \sum_j a_{ij}x_j = b_i \Rightarrow x_1 = \frac{1}{a_{11}}(b_1 - \sum_{j\neq 1} a_{1j}x_j)$

$$\sum_j a_{2j} x_j = b_2 \Rightarrow x_2 = \frac{1}{a_{22}}\Big(b_2 - \sum_{j \neq 2} a_{2j} x_j\Big)$$

$$\vdots$$

$$\sum_j a_{nj} x_j = b_n \Rightarrow x_n = \frac{1}{a_{nn}}\Big(b_n - \sum_{j \neq n} a_{nj} x_j\Big)$$

initialize $x^{(0)}$

for $k = 0, 1, 2, \ldots$

for $i = 1, 2, \ldots, n$

$$x_i^{(k+1)} = \frac{1}{a_{ii}}\Big(b_i - \sum_{j \neq i} a_{ij} x_j^{(k)}\Big)$$

end

end

$\leftarrow$ Jacobi

GS:

$$x_i^{(k+1)} = \frac{1}{a_{ii}}\Big(b_i - \sum_{j=1}^{i-1} a_{ij} x_j^{(k+1)} - \sum_{j=i+1}^{n} a_{ij} x_j^{(k)}\Big)$$

$$A = L + D + U$$

$$L = \begin{bmatrix} 0 & & & \\ a_{21} & \ddots & & \\ \vdots & & \ddots & \\ a_{n1} & \cdots & a_{n,n-1} & 0 \end{bmatrix} \qquad D = \begin{bmatrix} a_{11} & & \\ & \ddots & \\ & & a_{nn} \end{bmatrix}$$

$$U = \begin{bmatrix} 0 & a_{12} & \cdots & a_{1n} \\ & \ddots & & \vdots \\ & & \ddots & a_{n-1,n} \\ & & & 0 \end{bmatrix}$$

Jacobi:

Let $H = D^{-1}(L+U)$

if $|\lambda_i(H)| < 1 \quad \forall i$

Then J iterative convergency for x.

GS:

If A is PD,

GS iterative always convergence

2018.4.26

考试复习

Lec 1:

Subspaces; Linear independence

subspace dim, rank; norm, inner product

ortho. complement + projection; Ortho basis

inner/outer rep.; ~~flops, 算法复杂度~~

21页 property 证明.

37页 证明. 39页证明.

45页 53页.

Lec 2:

如何证 gradient descent 会收敛?

Application 都可以 skip.

LS Solution; circulant matrices.

* Vandermonde matrices (证线性独立)

可能会根据 Vandermonde 考证 Topizt

ortho. projector 45页可以用 SVD 来证.

pseudo-inverse

Lec 3:

characteristic polynomial; (Assignment 上那个奇怪的问题) eigen decomposition for sym. matrices (or not)

12页证明, 16页.

power method; Schur's decomposition

用 Schur 来证 23页.

~~Non-neg Matrix Theo~~

Lec 4:

properties of PSD matrices; 性质证明 9页, 13页, 14, 15, 16, 17, 18

Variational characterizations

40页, 41页.

~~PSD Theq~~

Schur's complement

46页证明, 和那些东西类似.

☆Lec 5: almost everything, except application

7, 8, 11, 12, 13(知道)

~~Sensitivity analysis~~

Singular value ineq

(这个可把 eigen value 和 singular value 组合到一起)

Lec 6:

~~sparsity~~

一些 regulization 5页

例: $\min_x \|y - Ax\|_2^2 + \|Wx\|_2^2$

LS with robustness

Lec 7~8

"common sense" w/ forward & backward subs.

at least know what is LU, QR, Cholesky...

Assignment 5 Prob 2, 对角占优存在LU解

~~Lec 9~10~~

~~10 是 application~~ skip.