Chapter 1: ~~Toep~~ Toeplitz: $\begin{bmatrix}1&2&4\\5&1&2\\6&5&1\end{bmatrix}$ circulant: $\begin{bmatrix}1&2&3&4\\4&1&2&3\\3&4&1&2\end{bmatrix}$

Transpose: A^T Hermitian transpose: A', A^* (有复数)

$(AB)' = B'A'$ $(a+b)' = a'+b'$ $(a-b)' = a'-b'$

dot/inner product: $\langle x,y\rangle = y'x$

$\langle \alpha x, y\rangle = \alpha\langle x,y\rangle$

outer product: $xy' = \begin{bmatrix}x_1\\ \vdots\\ x_M\end{bmatrix}[y_1^* \cdots y_N^*] = \begin{bmatrix}x_1y_1^* & \cdots\\ x_My_1^* & \cdots\end{bmatrix}$

$A_{:,1} = Ae_1$: first col $A_{1,:} = \tilde{e}_1' A$: first row

$Ax = \begin{bmatrix}A_{1,:}x\\ \vdots \\ A_{M,:}x\end{bmatrix} = \sum_{n=1}^{N} A_{:,n}x_n = A_{:,1}x_1 + \cdots A_{:,N}x_N$

$A = \begin{bmatrix}5&7\\6&8\end{bmatrix}$ 电脑 $\begin{matrix}5\\6\\7\\8\end{matrix}$ 这样存的 $A'y$ 会比较快

$y = \begin{bmatrix}5\\6\end{bmatrix}$ $y'x$ 会比较快 $(x\cdot y')\cdot z$: $O(N^2)$ $x\cdot(y'\cdot z)$: $O(N)$

P: $A(B+C) = AB + AC$ $A(BC) = (AB)C$

$AB \neq BA$ (even if size match) $A'A \neq AA'$ general

$IA = AI = A$, $A = B \Rightarrow AC = CB$ but symmetric

$vec(A) = \begin{bmatrix}A_{:,1}\\ \vdots \\ A_{:,N}\end{bmatrix}$

$vec(AXB) = (B^T \otimes A)vec(X)$

$\boxed{\text{Invertible}}$ (non-singular): invertible A: square

P: · A has linearly ind columns

☆ · A has full rank (i.e: all eigvals are non-zero)

· If A is unitary, $A' = A^{-1} = (A^+)$

· $\underline{det(A) \neq 0}$, $det(A^{-1}) = 1/det(A)$

· $Ax = 0 \Rightarrow x = 0$ · $(AB)^{-1} = B^{-1}A^{-1}$

$det(I - xy') = 1 - y'x$, $(A')^{-1} = (A^{-1})'$

If A is inv, $det\begin{pmatrix}A & B\\ C & D\end{pmatrix} = det(A)\cdot det(D - CA^{-1}B)$

If G is inv, $G + H = G(I + G^{-1}H)$

$det(B - \lambda I) = det(A + \theta xx' - \lambda I)$

$= det(A - \lambda I)(I + (A - \lambda I)^{-1}\theta xx')$

$= det(A - \lambda I)\, det(I + \theta(A - \lambda I)xx')$

$= det(A - \lambda I)(1 + \theta x'(A - \lambda I)^{-1}x)$

$\Rightarrow \theta x'(A - \lambda I)^{-1}x = 1$

$A = I_N + yy'$ $det(A - \lambda I) = det((1-\lambda)I_N + yy')$

$= (1-\lambda)^{N-1}(1 - \lambda + \|y\|_2^2)$

$N-1$ eigenvals are 1, one eigval is $1 + \|y\|_2^2$

$\boxed{\text{Orthogonal}}$:

vector x, y are orthogonal if $x'y = 0$

addition: If $x'x = y'y = 1$ (unit norm), they are orthonormal vectors.

$u \perp v \Longleftrightarrow \|u+v\|_2^2 = \|u\|_2^2 + \|v\|_2^2$

$= \|u\|_2^2 + 2\mathcal{R}\{u'v\} + \|v\|_2^2$

$\|x\|_2 = \sqrt{\langle x,x\rangle} = \sqrt{x'x} = \sqrt{\sum_{i=1}^{N}|x_i|^2}$

$\boxed{\text{Square}}$ $Q \in \mathbb{R}^{N\times N}$ is orthogonal matrix iff $Q^TQ = QQ^T = I$

$\boxed{\text{Square}}$ $Q \in \mathbb{C}^{N\times N}$ is unitary matrix iff $Q'Q = QQ' = I$

The set of col/row of orthogonal/unitary matrix is orthonormal set.

$Q^{-1} = Q' = Q^+$ $\|Qx\|_2 = \|x\|_2$

$\boxed{\text{determinant det}}$ $P(1,4)\cdot P(1,4) = P(1,4)\cdot P(1,4)' = I$ unitary

- If A is $\mathbb{F}^{N\times N}$ and upper/lower triangular, then $det\{A\} = a_{11}\cdot a_{22}\cdots a_{nn}$

- If $A, B \in \mathbb{F}^{N\times N}$, $det\{AB\} = det\{A\} det\{B\} = \underline{det\{BA\}}$.

- $det\{A'\} = det\{A\}^*$ complex conjugate of $det\{A\}$.

- $det\{P_{i,j}\} = -1$ $A = VP'PAP'(VP')'$

P: · $det(P_{ij}A) = -det(A) = det(AP_{ij})$ $A \in \mathbb{F}^{N\times N}$

· $det\{A^T\} = det\{A\}$. · $det\{cA\} = c^N det\{A\}$

· $det(I_N) = 1$ $A^+ = \begin{bmatrix}1 & 0\\ 0 & 1/4\end{bmatrix}$

$\boxed{\text{eigenvalues}}$

$det\{A - \lambda I\} = 0$ $A = \begin{bmatrix}1&0\\0&4\end{bmatrix}$ eigval: 1, 4

P: · If A is upper/lower triangular, or diagonal matrix, eigenvals of A are its diagonal elements.

· eigval(A') are the complex conjugate of eigvals(A)

· eigval(αA) are α · eigvals(A)

· If A is N×N and T is $\boxed{\text{inver}}$ N×N, eigvals(A) = eigvals(TAT^{-1}), and $\boxed{\text{unitary}}$ Q

eigvals(A) = eigvals(QAQ')

$A - \lambda I$ is singular

· $det(A) = \lambda_1\cdot\lambda_2\cdots\lambda_n$ $Tr(A) = \lambda_1 + \lambda_2 + \cdots + \lambda_n$

· If A has eigenvals $\{\lambda_1, \cdots \lambda_N\}$, then A^2 eigenvals $\{\lambda_1^2 \cdots \lambda_N^2\}$

A^k eigvals $\{\lambda_1^k \cdots \lambda_N^k\}$.

diff size · $eig(AB) - \{0\} = eig(BA) - \{0\}$.

不是所有 diag has N distinct eigenvalues.

$\boxed{\text{Trace}}$ of square

$Tr(A) = \sum_{i=1}^{n} A_{ii}$ (对角线的和)

$Tr(\alpha A + \beta B) = \alpha Tr(A) + \beta Tr\{B\}$

$Tr(XYZ) = Tr(ZXY) = Tr(YZX)$ $\boxed{\text{Cyclic}}$

$Tr(AB) = Tr(BA)$, B, A 不是 square.

$\boxed{\text{Chapter 2:}}$

for any $\boxed{\text{symmetric or normal A}}$

$A = Q\Lambda Q'$ orthogonal eigendecomposition

Q is unitary

$A = V\Lambda V^{-1}$ diagonalization (some square A)

V is linearly indep vector

$A = U\Sigma V'$. SVD

$AV = V\Lambda$ $V = [v_1 \cdots v_N]$ $\Lambda = diag\{\lambda_1 \ldots \lambda_n\}$.

$|V|$ is 不可数 uncountably infinite.

$\boxed{\text{Spectral Theorem}}$

Symmetric matrix: rank = # of non-zero eigenvals

$A'A = AA'$ (normal)

If $A \in \mathbb{F}^{N\times N}$ is (Hermitian) symmetric, (sufficient but no necessary)

⇓ normal

then · eigvals(A) are all real

· $A = V\Lambda V' = \sum_{n=1}^{N}\lambda_n v_n v_n'$

V is orthogonal (unitary) matrix

orthonormal basis for $\mathbb{F}^N$ consisting of eig vectors of A

($x'x$ or xx' are always symmetric and square)

$P^{-1} = P'$ is orthogonal matrix

$\boxed{\text{spectral theorem}}$

A square matrix A is diagonalizable by a unitary matrix, i.e: has unitary eigendecomposition, there exist V unitary and Λ diagonal such that $A = V\Lambda V'$ iff A is normal matrix

A is unitary eigendecomposition 条件:

unitary: necessary and suffient

square: necessary

diagonlizable: ~~ness~~ necessary

define. diagonalizaliable:

A sqare matrix is diagonalizable iff it is similar to a diagonal matrix

i.e: iff there exists an invertible matrix V such that $V^{-1}AV$ is diagonal.

(o.w. defective)

$A = V\Lambda V^{-1}$ (V is linearly indep eigenvectors)

some, but not all, $\boxed{\text{square asymmetric matrix}}$ that are not normal matrix are diagonalizable.

· If A is asymmetric, A cannot decompose eigen unitary

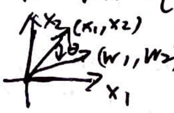
· If A has N $\boxed{\text{distinct}}$ eigenvals, then A is diagonalizable (not necessary condition)

· If A is both invertible and diagonalizable then $A^{-1} = V\Lambda^{-1}V^{-1}$

· Being diagonalizable 和 invertiable because eigvals can be 0 没关系

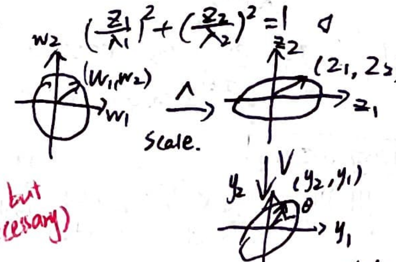
$A = \begin{bmatrix}3&1\\0&3\end{bmatrix}$ has ~~reapted~~ repeated values so it is not diagonalizable.

$W = V'x = \begin{bmatrix}\cos\theta & -q\cdot\sin\theta\\ \sin\theta & q\cos\theta\end{bmatrix}x$



$W \to Z = \Lambda W = \begin{bmatrix}\lambda_1 & 0\\ 0 & \lambda_2\end{bmatrix}W = \begin{bmatrix}\lambda_1 w_1\\ \lambda_2 w_2\end{bmatrix}$

$\left(\frac{z_1}{\lambda_1}\right)^2 + \left(\frac{z_2}{\lambda_2}\right)^2 = 1$



If's Hermitian Symmetric ⇒ normal ⇒ diagonalizable.

all normal → diagonalizable but not all diagonalizable are normal

SVD singular-value-decomposition

$Av_i = \sigma_i u_i$
$A'u_i = \sigma_i v_i$ singular vec 不可数

$A \in F^{M\times N}$, $\boxed{A = U\Sigma V' = \sum_{k=1}^{\min(M,N)} \sigma_k u_k v_k'}$

$A = \sum_{k=1}^{r} \sigma_k u_k v_k'$

U is $M\times M$ and unitary: $U'U = UU' = I$
columns consists of M left singular vector of A

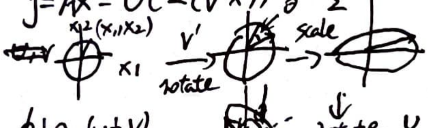
V is $N\times N$ and unitary: N right singular vec of A

Σ: 长方形. singular values of A

$\Sigma = \begin{bmatrix} \sigma_1 \ddots \sigma_N \\ 0_{M-N,N}\end{bmatrix}$ $M>N$ $\quad = [\sigma_1 \ddots \sigma_M \mid 0_{M,N-M}]$ $N>M$ $\quad \Sigma = [\sigma_1 \ddots \sigma_M]$ $M=N$

Singular are real and non-negative.

$y = Ax = U(\Sigma(V'x))$ θ Σ



$\phi \neq \theta$, $(U \neq V)$

U, V are not unique while Σ is unique

$\sigma_1 u_1 \perp \sigma_2 u_2$

Full SVD: svd(A, full=true)

• $A' = V\Sigma' U'$ • $Av_i = \sigma_i u_i$, $A'u_i = \sigma_i v_i$

• If $B = \alpha A$, then $B = U\tilde{\Sigma}V'$ where $\tilde{\Sigma} = \alpha\Sigma$ is a SVD of B

• If $B = AQ'$, Q is unitary
$B = U\Sigma\tilde{V}'$ where $\tilde{V} = QV$

The matrix 2-norm or spectral norm

① $x_* = \arg\max_{\|x\|_2=1} \|Ax\|$
$x_* = v_1$ or $e^{i\phi}v_1$

② $y_* = \arg\max_{\|y\|=1} \|A'y\| = u_1$

*x_*, y_* is unique when $N=1$ or ($N>1$, $\sigma_1(A) > \sigma_2$ -)

③ $\sigma_1(A) = \max_{\|x\|_2=1} \|Ax\|_2$

$j = \imath = \sqrt{-1}$ $\quad e^{j\pi} = \cos\pi + j\sin\pi$
$e^{\imath\phi} = \cos\phi + \sqrt{-1}\sin\phi$

matrix-2 norm

$|||A|||_2 \triangleq \max_{x:\|x\|=1} \|Ax\|_2 = \max_{x\neq 0} \frac{\|Ax\|_2}{\|x\|_2} = \sigma_1$

$\|Ax\|_2 \le \sigma_1\|x\|_2 = |||A|||_2 \|x\|_2$
upper bound is achieved when $x = v_1$

$\lambda_1 = \max_{\|x\|_2=1} x'Ax \quad \lambda_N = \min_{\|x\|_2=1} x'Ax$

$\sigma_{\min(M,N)} = \min_{x\neq 0}\frac{\|Ax\|_2}{\|x\|_2} = \min_{x:\|x\|_2=1}\|Ax\|_2 = \begin{cases} 0, & N>M \\ \sigma_N, & N\le M\end{cases}$

SVD ⟷ eigendecompositions

• V: right singular vectors of A is an eigenvectors of $A'A$
$A'Av_k = A'(\sigma_k u_k) = \sigma_k^2 v_k$

• U: left singular vectors of A is an eigenvectors of AA'
$AA'u_k = A(\sigma_k v_k) = \sigma_k^2 u_k$

• $\min(M,N)$ singular values of A are the square root of the eigenvalues of $A'A$ or AA'

$\sigma_1(A) = \sqrt{\lambda_1(A'A)}$ $\quad A'A = V\underbrace{\Sigma'\Sigma}_{\Lambda}V' = V\begin{bmatrix}\sigma_1^2 & & \\ & \ddots & \\ & & \sigma_N^2\end{bmatrix}V'$

$\sigma_1 = \sqrt{\lambda_1}$ $\quad A = V\Lambda V^{-1}$

① If A is diagonalizable square matrix, then eigendecomposition is unrelated to A

② If A is normal, related. A is $N\times N$ normal
$A = V\Lambda V' = \sum_{n=1}^{N}\lambda_n v_n v_n'$
$= \sum_{n=1}^{N} \underbrace{|\lambda_n|}_{\sigma_n} \underbrace{\mathrm{sign}(\lambda_n) v_n}_{u_n} v_n'$
$= U\Sigma V'$

$A = \begin{bmatrix}3&0\\0&-3\end{bmatrix} = \underbrace{\begin{bmatrix}1&0\\0&1\end{bmatrix}}_{V}\underbrace{\begin{bmatrix}3&0\\0&-3\end{bmatrix}}_{\Lambda}\underbrace{\begin{bmatrix}1&0\\0&1\end{bmatrix}}_{V'}$
$= \underbrace{\begin{bmatrix}1&0\\0&-1\end{bmatrix}}_{U}\underbrace{\begin{bmatrix}3&0\\0&3\end{bmatrix}}_{\Sigma}\underbrace{\begin{bmatrix}1&0\\0&1\end{bmatrix}}_{V'}$

$U = V$ iff A is square, $A = A'$, and $\lambda_i \ge 0$ $\forall i$

Positive semidefinite matrices PSD>0 is always invertible (eigenvals ≥ 0)

① square Hermitian A is positive semidefinite iff $x'Ax \ge 0$ $\forall x \in F^N$

② If eigvals > 0, A is positive definite iff $x'Ax > 0$

suff If $A = BB'$ for any B, then A is positive semidefinite
$A = BB'$, an eigendecomposition is also SVD

Rectangular	square	diagonalizable	Normal	Hermitian	PSD: $A \succeq 0$
SVD: $A = U\Sigma V'$	$AV = V\Lambda$	$A = V\Lambda V^{-1}$ iff V linearly ind	$A = V\Lambda V'$, V unitary, $\sigma_n = \lvert\lambda(n)\rvert$	Λ real, $A = A'$	$\Sigma = \Lambda \succeq 0$, $U = V$
	$\begin{bmatrix}0&1\\0&0\end{bmatrix}$	$\begin{bmatrix}1&1\\0&0\end{bmatrix}$, $\begin{bmatrix}1&0\\1&0\end{bmatrix}$	$[i]$	$[-1]$	

Chap3

Subspace → S always includes the zero vector

For a vector space V, a nonempty $S \subseteq V$ is a called subspace or linear subspace of V iff
• S is closed under vector addition: $u, v \in S \Rightarrow u+v \in S$
• S is closed under scalar multiplication. $v \in S \Rightarrow \alpha v \in S$ $\forall \alpha \in F$

span is a subspace of vector space V
$\mathrm{span}(\phi) = 0$

Any set of orthonormal vectors is linearly independent
Any set of non-zero orthogonal vectors is linearly independent.

Basis vector for V
iff • $\{b_1, \cdots b_z\}$ is a linear ind set
• $\mathrm{span}(\{b_1, b_2 \cdots\}) = V$

Every subspace in a vector space has a basis

$S = \mathrm{span}(\{u_1, \ldots u_r\})$ then $\dim(S) \le N$

If $S, T \subseteq V$, then
Sum two subspace. $S+T = \{s+t: s\in S, t \in T\}$
Intersection: $S \cap T = \{v \in V: v \in S \text{ and } v \in T\}$

$\oplus$ direct sum
$S \oplus T$ iff $S \cap T = \{0\} \Rightarrow S \perp T$
If $U = S\oplus T$. $\dim(U) = \dim(S) + \dim(T)$

Orthogonal complement

For a subspace S of a vector space V, the orthogonal complement of S is the subset of vectors in V that orthogonal to every vector in S

$S^\perp = \{v \in V: \langle s, v\rangle = v's = 0, \forall s \in S\}$

(P) • $S^\perp$ is subspace of V
• $(S^\perp)^\perp = S$ • $S \oplus S^\perp = V$
$S \cap S^\perp = \{0\}$
• $\dim(S) + \dim(S^\perp) = \dim(V)$

Range. range of A = column space
$R(A) = \mathrm{span}(\{a_1, \ldots a_N\}) = \{Ax: x \in F^N\}$

Row space: $R(A')$
$A \in F^{M\times N}, B \in F^{N\times K} \Rightarrow R(AB) \subseteq R(A)$
If B is invertible, then $R(AB) = R(A)$

☆ **Rank**. A: $M\times N$
col rank $\triangleq \dim(R(A))$ = # of linear ind cols $\le N$
row rank $\triangleq \dim(R(A'))$ = # of linear rows $\le M$
$\dim(R(A)) = \dim(R(A'))$ col rank = row rank

$0 \le \mathrm{rank}(AB) \le \min(\mathrm{rank}(A), \mathrm{rank}(B))$
$= \min(M, N, K)$

$\mathrm{rank}(A) = \mathrm{rank}(AQ) = \mathrm{rank}(QA)$

If A is Hermitian (Normal) with eigendecomposition $A = V\Lambda V'$
$\mathrm{rank}(A) = \mathrm{rank}(\Lambda)$ = number of non-zero eigenvalues of A

$\mathrm{rank}(A) = \mathrm{rank}(\Sigma)$

Nullspace or kernel
$N(A) = \mathrm{Ker}(A) \triangleq \{x \in F^N: Ax = 0_M\}$

(P): • $N(A) = \{0\} \Longleftrightarrow A$ has full rank
• $N(A) = F^N \Longleftrightarrow A = 0_{M\times N}$
• $N(B) \subseteq N(AB)$
If $N(A) = \{0\}$, then $N(B) = N(AB)$

$A \in F^{M\times N}$:
$N(A) \oplus N^\perp(A) = F^N$
$R(A) \oplus R^\perp(A) = F^M$

$\underbrace{\dim(N(A))}_{\text{Nullity}} + \underbrace{\dim(R(A))}_{\text{rank}} = N$

$N^\perp(A) = R(A')$ $\quad N(A) \perp R(A')$

$R^\perp(A) = N(A')$ $\quad R(A) \perp N(A')$

$A \in F^{M\times N}$

dim	input F^N		output F^M	dim
r	$N^\perp(A)=R(V_r)$	$\xrightarrow{A}$	$R(A)=R(U_r)$	r
	$\oplus\{0\}$		$\{0\}\oplus$	
$N-r$	$N(A)=R(V_0)$		$R^\perp(A)=R(U_0)$	$M-r$

$A\in F^{M\times N}$

$$A = \underset{M\times r}{[U_r} \,\vdots\, \underset{M\times(M-r)}{U_0]} \overset{r\times r}{\begin{bmatrix}\Sigma_r & 0\\ 0 & 0\end{bmatrix}}_{(M-r)\times(N-r)} \underset{N\times r}{[V_r} \,\vdots\, \underset{N\times(N-r)}{V_0]'}$$

$$A\cdot V_0 z = U\Sigma\begin{bmatrix}V_r'\\V_0'\end{bmatrix}V_0 x = U\begin{bmatrix}\Sigma_r & 0\\0&0\end{bmatrix}\begin{bmatrix}0\\I_{N-r}\end{bmatrix} = 0$$

Tall: $M>N$: $r\le N<M$

$$\underset{M\times N}{\boxed{A}} = \underset{M\times r}{[U_r} \,|\, \underset{M\times(M-r)}{U_0]} \underset{M\times r}{\overset{r\times r}{\begin{bmatrix}\Sigma_r\\0\end{bmatrix}}} \underset{r\times N}{V_r'}$$

underdetermined

Wide A: $N>M$ $\quad r\le M<N$ $\quad$ 永远 $r<N$

$$\boxed{A} = \underset{M\times r}{U_r}\underbrace{[\Sigma_r\,|\,0]}_{r\times N}\underset{N\times N}{\begin{bmatrix}V_r'\\V_0'\end{bmatrix}}\begin{matrix}r\times N\\(N-r)\times N\end{matrix}$$

If $r=N$, there is 0 V_0 $\quad N(A)=0$

orthogonal bases

- $\{b_1,b_2\cdots\}$ in V is orthogonal basis for V iff
  ① $\{b_1,b_2\cdots\}$ is a basis for V
  ② The basis vector are orthogonal $\langle b_n,b_m\rangle=0$.

N cols of orthogonal matrix $V\in R^{N\times N}$ are orthonormal basis for R^N

N cols of unitary matrix $V\in C^{N\times N}$ are orthonormal basis for C^N

$$A = bc' = \frac{b}{\|b\|_2}\|b\|_2\|c\|_2\left(\frac{c}{\|c\|}\right)'$$

$$= \underbrace{\left[\frac{b}{\|b\|}\,\vdots\, U_0\right]}_{U}\underbrace{\begin{bmatrix}\|b\|\|c\| & 0\\ 0 & 0\end{bmatrix}}_{\Sigma}\underbrace{\left[\frac{c}{\|c\|}\,\Big|\, V_0\right]'}_{V'}$$

Projection

S_1 $\quad P_{S_1}(v) \ldots v$

$$\hat v = P_S(v) \triangleq \arg\min_{s\in S}\|v-s\|_n$$

P_S is an idempotent operation: $P_S\circ P_S = P_S$

If $U\in F^{M\times K}$ has orthonormal columns: $U'U = I_K$ $\quad UU'\ne I_M$

$S=R(U)$ for which U is an orthonormal basis

$\hat v = P_S(v) = P_{R(U)}(v) = U(U'v)$

$P_S \triangleq UU'$ (projection matrix)

$P_S^\perp = I - P_S$ $\quad P_{S^\perp}(v) = v - P_S(v)$

$$P_{S^\perp}(v) = v - P_S(v) = v - P_S v = v - UU'v = (I-UU')v$$

U 是 orthonormal basis.

例: For non-zero x, and orthonormal basis for $S=\mathrm{span}(\{x\})$ is $u = x/\|x\|_2$, so $P_S = uu'$. Thus, the projection of y onto the orthogonal complement of $S=\mathrm{span}\{x\}$ is

$$P_{S^\perp}y = P_S^\perp y = (I-P_S)y = y-P_Sy = y-uu'y = y-\frac{x}{\|x\|_2}\frac{x'}{\|x\|_2}y = y-\frac{x'y}{x'x}x$$

P: $(P_S^\perp)^2 = P_S^\perp$
- $P_S' = P_S$ and $(P_S^\perp)' = P_S^\perp$, $\;\cdot\; P_S P_S^\perp = 0$

$$\hat c = \arg\min_{c\in\{1,\ldots,C\}} d_c(v),\quad d_c(v) = \|v - P_{S_c}(v)\|_2 = \|v - U_c(U_c'v)\|_2$$

Convex function

$$f(\alpha x + (1-\alpha)z) \le \alpha f(x) + (1-\alpha)f(z)$$

$f: V\to R$, and $x,z\in V$, $\alpha=[0,1]$

Chap 4

C is convex set iff $x,z\in C \Rightarrow \alpha x+(1-\alpha)z\in C$ $\quad \alpha\in[0,1]$

$$\hat x = \arg\min_{x\in F^N}\|Ax-y\|_2^2$$

$r = A\hat x - y$ is residual

$y=Ax$, $y=\begin{bmatrix}y_1\\\vdots\\y_n\end{bmatrix}$, $A=\begin{bmatrix}1 & t_1 & t_1^2\\1&t_2&t_2^2\\1&t_3&t_3^2\\\vdots&&\\1&t_n&t_n^2\end{bmatrix}$, $x=\begin{bmatrix}\alpha_0\\\alpha_1\\\alpha_2\end{bmatrix}$ $\quad S(t)=\alpha_0+\alpha_1 t+\alpha_2t^2$ $\quad f(t)=x_1t^0$

LLS cost function:

$$f(x) = \frac12\|Ax-y\|_2^2 = \frac12(Ax-y)'(Ax-y) = \frac12 x'A'Ax - y'Ax + \frac12 y'y$$

$$\nabla f(x) = A'Ax - A'y = A'(Ax-y)$$

Let $\nabla f(x)=0$

$$A'A\hat x = A'y$$

① If A is wide, $A'A$ rank at most M; $A'A$ is ~~invertible~~ singular (non-inv) 很多 $\hat x$

② If A is tall or square, $A\in F^{M\times N}$ and rank$(A)=N$

$$\Rightarrow \boxed{\hat x = (A'A)^{-1}(A'y)}$$

Using SVD:

$$\|Ax-y\|_2^2 = \|U_r\Sigma_rV_r'x\|_2^2 - 2\,\mathrm{real}((U_r\Sigma_rV_r'x)y') + \|y\|_2$$

$$= \|\Sigma_rV_r'x - U_r'y\|_2 + \|y\|_2^2 - \|U_r'y\|_2^2$$

$$\hat x = \arg\min_x\|\Sigma_rV_r'x - U_r'y\|_2^2$$

① when $r=N$, $\boxed{\hat x = V_r\Sigma_r^{-1}U_r'y = \sum_{k=1}^{r}\frac{1}{\sigma_k}v_k(u_k'y)}$ unique

② when $r<N$, not unique

$$\boxed{\hat x = V_r\Sigma_r^{-1}U_r'y + V_0z_0},\quad \forall z\in F^{N-r}$$

两种情况:
- A is square or tall but has linearly dependent columns.

不是 unique:
- If A is wide, $r\le\min(M,N)=M<N$

$x_h = V\cdot\mathrm{Diagonal}(1./s)\cdot(U' * y)$

$$\hat x = V\Sigma_N^{-1}U_N'y = V[\Sigma_N^{-1}\,|\,0_{N\times(M-N)}]\begin{bmatrix}U_N'\\U_0'\end{bmatrix}y = V[\Sigma_N^{-1}\,|\,0_{N\times(M-N)}]U'y$$

If $M\ge N$ and rank$(A)=N$ $\Rightarrow$

$$\boxed{\hat x = V\Sigma^+U'y}\qquad \Sigma^+\triangleq[\Sigma_N^{-1}\,|\,0_{N\times(M-N)}]\qquad \Sigma = \begin{bmatrix}\Sigma_N\\0_{M-N\times N}\end{bmatrix}$$

$1_M = [\,\vdots\,]$ $\quad 1_M^+ = \frac1M 1'$

Moore-Penrose Pseudoinverse

- $AA^+A = A$ $\quad\cdot\; A^+AA^+ = A^+$ $\quad\cdot\; (A^+A)' = A^+A$ $\quad (AA^+)' = AA^+$

$A\in F^{M\times N}$

P:
- A^+ is unique $\quad (A^+)^+ = A$
- If A is inver, then $A^+ = A^{-1}$
- $A^+ = (A'A)^+A' = A'(AA')^+$
- If A is unitary, then $A^+ = A^{-1} = A'$
- If A has full col rank, $A^+ = (A'A)^{-1}A'$
  $A^+A = I_N$, A^+ is left inverse
- If A has full row rank, then $A^+ = A'(AA')^{-1}$
  $AA^+ = I_M$, A^+ is right inverse
- $(A^+)' = (A')^+$ $\qquad 0^+_{M\times N} = 0_{N\times M}$

- If Q is $M\times K$ with orthonormal col: $Q'Q = I_K$, then $(QB)^+ = B^+Q'$ $\quad B\in F^{K\times L}$

If Q is $L\times N$ with ~ rows, $QQ' = I_{L\times L}$ then $(BQ)^+ = Q'B^+$

A $M\times N$ B $N\times K$ A has full col rank B has full row rank then $(AB)^+ = B^+A^+$

If Q is orthonormal col/row, $Q^+ = Q'$

$\Rightarrow U_r^+ = U_r'$ and $V_r^+ = V_r'$

$$\boxed{A^+} = V_r\Sigma_r^{-1}U_r' = (V')^+\Sigma^+U^+$$

$$\underset{M\times N}{\Sigma = \begin{bmatrix}\Sigma_r & 0_{r\times N-r}\\ 0_{M-r\times r} & 0_{M-r\times N-r}\end{bmatrix}}\qquad \underset{N\times M}{\Sigma^+ = \begin{bmatrix}\Sigma_r^{-1} & 0_{r\times M-r}\\ 0_{N-r\times r} & 0_{N-r\times M-r}\end{bmatrix}}$$

$r=N$:

$$\Sigma = \begin{bmatrix}\Sigma_N\\0_{M-N\times N}\end{bmatrix}\Rightarrow\Sigma^+ = [\Sigma_N^{-1}\,|\,0_{N\times M-N}]$$

A^+: $\mathrm{pinv}\left(\begin{bmatrix}0&0\\0&2\end{bmatrix}\right) = \begin{bmatrix}0&0\\0&\frac12\end{bmatrix}$ $\quad \mathrm{pinv}\begin{pmatrix}3&0\\0&5\end{pmatrix} = \begin{bmatrix}1/3&0\\0&1/5\end{bmatrix}$

A^+: Diagonal matrix $\Rightarrow$ matrix with pinv each diagonal element

$A\in F^{M\times N}$ and rank$(A)=N$

$$\boxed{\hat x = \arg\min_{x\in F^N}\|Ax-y\|_2^2 = (A'A)^{-1}A'y = A^+y}$$

$x_h = \mathrm{pinv}(A)\cdot y = \mathrm{inv}(A'\cdot A)\cdot(A'\cdot y)$

不是很有效 $\quad = A\backslash y$ (QR decom) preferable

$\hat x = V_r\Sigma_r^{-1}U_r'y + V_0\hat z_0$

$\hat x\in R(V_r) = N^\perp(A)$ $\quad \hat x_N\in N(A)$

$= A^+y + N(A) = A^+y + \hat x_N$

Pick smallest Euclidean norm

$r=N$, $r<N \Rightarrow \hat x = A^+y$

$N(A)=\{0\}$.

Truncated SVD Solution:

$$\hat{x} = \sum_{k=1}^{r} \frac{1}{\sigma_k} v_k u_k' y$$

σ_1, σ_k 差距很大, poorly conditioned

Since A has large condition number (κ)

$$\kappa(A) = \begin{cases} \infty, & M<N \\ \frac{\sigma_{max}(A)}{\sigma_{min}(A)} & M \ge N \end{cases} = \begin{cases} \infty, & m<N \\ \frac{\sigma_1}{\sigma_N}, & M \ge N \end{cases}$$

$$= \begin{cases} \infty & M<N \\ \sqrt{\|A'A\|_2 \, \|(A'A)^+\|_2} & M \ge N \end{cases} = \sqrt{\frac{\sigma_{max}(A'A)}{\sigma_{min}(A'A)}}$$

① truncated SVD solution.

discard any "too small" singular values.

$$\hat{x}_K = \sum_{k=1}^{K} v_k \frac{1}{\sigma_k} u_k' y$$

for $k<r$, $\sigma_k > \delta > 0$ for some tol δ

`pinv(A; tol = δ) * y`

$\hat{x} = A \backslash y$ (no tol)

可以先 $A_K = U_K \Sigma_K V_K'$ then $\hat{x} = A_K^+ y$

(principal component regression)

② Tikhonov regularization

$$\hat{x}_\beta = \underset{x \in \mathbb{F}^N}{\arg\min} \|Ax-y\|_2^2 + \beta \|x\|_2^2$$

energy of x

$$= \left\| \begin{bmatrix} A \\ \sqrt{\beta} I_N \end{bmatrix} x - \begin{bmatrix} y \\ 0_N \end{bmatrix} \right\|_2^2 = \|\hat{A}x - \hat{y}\|_2^2$$

$\hat{A} = \begin{bmatrix} A \\ \sqrt{\beta} I_N \end{bmatrix}$ $\quad \hat{y} = \begin{bmatrix} y \\ 0_N \end{bmatrix}$

$$\Rightarrow \boxed{\hat{x}_\beta = (A'A + \beta I)^{-1} A' y = \sum_{k=1}^{r} \left(\frac{\sigma_k}{\sigma_k^2 + \beta}\right) v_k (u_k' y)}$$

$\hat{x}_\beta \to A^+ y$ as $\beta \to 0$

Frame

① Frame: $\alpha \|x\|_2^2 \le \|\Phi' x\|_2^2 \le \beta \|x\|_2^2$

Φ is a frame

(P): • 若 Φ has $r = N$, $0 < \alpha \le \sigma_N^2$

• Φ must be square or wide

• $\Phi^+ = V_N \Sigma_N^{-1} U'$ $\quad \Phi \Phi^+ = I_N$

② Tight Frame, $\alpha = \beta$

$\alpha \|x\|_2^2 = \|\Phi' x\|_2^2$.

(P): $\Phi \Phi' = \alpha I_N$, $\alpha = \sigma_1^2 = \dots = \sigma_N^2$

$\Phi^+ = \frac{1}{\alpha} \Phi' = \frac{1}{\sigma_1^2} \Phi'$

③ Parseval tight Frame. $\alpha = \beta = 1$

(P). $\Phi \Phi' = I_N$, $\sigma_1 = \dots = \sigma_N = 1$

$\Phi^+ = \Phi'$

Gradient descent iteration $0 < \mu < \frac{2}{\sigma_1(A)^2}$

$$x_{k+1} = x_k - \mu A'(Ax_k - b)$$

Rectangular	Frame $\begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$	Tight frame	Parseval tight	unitary
$N \times M$ wide $N \le M$	$0 < \alpha \le \sigma_N^2 \le \sigma_1^2 \le \beta$	$0 < \alpha = \sigma_N^2 = \sigma_1^2$	$\sigma_N = \sigma_1 = 1$	$M = N$
$\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix}$	$\Phi\Phi'$ invertible	$\Phi\Phi' = \alpha I$	$\Phi\Phi' = I$	$U^+ = U^{-1} = U'$
	$\Phi^+ = \Phi'(\Phi\Phi')^{-1}$	$\Phi^+ = \frac{1}{\alpha}\Phi'$ $\begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$	$\Phi^+ = \Phi'$ $[1 \; 1]/\sqrt{2}$	

求A的SVD. ① 看 A是不是 normal Hermitian ② outer product.

~~Projection~~

Projection and orthogonal projection

A $\boxed{\text{Square}}$ P is called projection matrix or idempotent iff $P^2 = PP = P$

(P): • eigvals are 0 or 1

• every P is diagonalizable

orthogonal projector / projection matrix

If A is P and A is Hermitian unitary eigendecomposition

then A is orthogonal projector

I: 是 orthogonal matrix 和这个 matrix

$R^\perp(A)$, y, $A\hat{x}$

$r = y - A\hat{x} \in R^\perp(A)$ $\quad Az \perp y - A\hat{x}$

square matrix $[2]$	Idempotent Projection matrix $\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$	orthogonal Projection matrix $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$	Non-zero orthogonal projection matrix $\begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix} [\cos\theta \; \sin\theta]$	Diagonal with {0,1} elements $\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$	$[I]$
	$P^2 = P$, diagonalizable, e/g {0,1}.	$P = P'$, normal	$P = QQ'$ where $Q'Q = I$		

$$\boxed{\begin{aligned} \hat{x} &= \arg\min \|Ax - y\| = A^+ y + V_0 z \\ \hat{y} &= A\hat{x} = A(A^+ y + V_0 z) \\ &= AA^+ y \quad \text{(unique)} \end{aligned}}$$

$P_{R(A)} y = AA^+ y$

$P_{R(A)} \triangleq AA^+ = U_r U_r'$ $\quad$ Julia: `Ur = U[:, 1:r]`

※ If Q is a matrix with orthonormal columns,

$P_{R(Q)}(y) = Q(Q'y)$ (orthonormal bases)

If a basis B is not orthonormal,

$P_{R(B)} y = BB^+ y$

v, $r = v - P_S \in S^\perp$, $P_S(v)$, S

space	subspace	equivalent subspace	orthonormal basis matrix	orthogonal projector	alternate projector
F^N	$N(A)$	$R^\perp(A')$	V_0	V_0V_0'	$I - V_rV_r'$
F^N	$N^\perp(A)$	$R(A')$	V_r	V_rV_r'	$I - V_0V_0'$
F^M	$R(A)$	$N^\perp(A')$	U_r	U_rU_r'	$I - U_0U_0'$
F^M	$R^\perp(A)$	$N(A')$	U_0	U_0U_0'	$I - U_rU_r'$

rank $P_{R(A)} = r$ $\quad r(P_{R^\perp(A)}) = M - r$ $\quad P_{R(A)}^\perp = P_{R^\perp(A)}$

$$P_{R^\perp(A^+)} x = P_{R(V_0)} x = (I - V_r V_r') x$$

$= x - V[:, 1:r] \cdot (V[:, 1:r]' \cdot x)$

$A^+ = V_r \Sigma_r U_r'$

$R(A^+) = R(V_r)$

$R^\perp(A^+) = R(V_0)$

$AA^+ y = U_r U_r' y$

$A^+ x = A \backslash x$

`= pinv(A) * x`

```
InRange(B, z)   z = Rang(B) ?
(U, S, _) = svd(B)
r = rank(Diagonal(S))
Ur = U[:, 1:r]
~~Pz = U[:, 1:r]~~
Pz = Ur * (Ur' * z)
return isapprox(z, Pz)
```

chops.

$\|x+y\| \le \|x\| + \|y\|$ $\quad \|x\| \ge 0$ $\quad \|x\| = 0$ iff $x = 0$

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

$\|x\|_p \triangleq \left(\sum_{i=0} |x_i|^p\right)^{1/p}$

$\|x\|_2 \triangleq \sqrt{\sum_i |x_i|^2}$

$\|x\|_\infty \triangleq \max\{|x_1|, \dots |x_N|\}$.

$\|x\|_0 = \sum_i \mathbb{I}_{\{x_i \ne 0\}}$ # of non-zero

`norm(v, 0) = count(!=(0), v)`

$\langle A, B \rangle \triangleq tr(AB')$

$|\langle x, y \rangle| \le \|x\| \|y\|$ $\forall x, y \in V$

$\|A\|_F = \sqrt{tr(A'A)} = \sqrt{tr(AA')} = \sqrt{tr(\Sigma\Sigma')}$

$= \sqrt{\sum_{k=1}^{r} \sigma_k^2}$

$\|Ax\|_2 \le \|A\|_F \|x\|_2$.

$ax + by + cz = 0$

$[a \; b \; c] \begin{bmatrix} x \\ y \\ z \end{bmatrix}$

$\begin{bmatrix} x \\ y \\ z \end{bmatrix} \in N([a \; b \; c]) = R(V_0)$

nearest point: $P_{R(V_0)} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$

$= P_{R[v_2, v_3]} y$

$= [I - v_1 v_1'] y$

eigen real: Gram matrix, Hermitian matrix, positive semi-definite matrix

not real: permutation matrix, Unitary, rotation matrix

are 0

☆ $[\;][\;] = xy'$ ① one eigen is $y'x$, rest are 0

④ SVD $= \frac{x}{\|x\|_2} \|x\|_2 \|y\|_2 \frac{y'}{\|y\|_2}$ ② xy' symmetric

③ rank$(xy') \le 1$

⑤ $(xy')^+ = \frac{yx'}{\|x\|_2^2 \|y\|_2^2}$

$A'A + \delta I$ is always invertible whenever $\delta > 0$ invertible

diagonal of PD are all positive. PD is always

A, B is PD, $A+B$ is PD $\quad AB \ne PD$

$A^{\frac{1}{2}} A^{\frac{1}{2}} = A$

All eigenval $> 0 \Rightarrow PD \Rightarrow$ invertible = full rank = det ≠

Singular: 有 $\lambda = 0$, 情况 $n - r$

ellipse

$x'Ax = 1$ $\quad A$ = Hermitian

$A = V \Lambda V'$

$y = V'x \Rightarrow y' \Lambda y = 1$

$\lambda_1 x^2 + \lambda_2 x^2 = 1$.

SVD return economy of SVD

SVD(A) $U_{M \times min} \Sigma_{min \times min} V_{N \times min}$

PSD 一定是Hermitian $\begin{bmatrix} 1 & 0 \\ 0 & -5 \end{bmatrix}$ 若A is Hermitian, 不一定PSD 呢

$A \in 5 \times 5$

$-P(\lambda) = \det(\lambda I_5 - A)$. Highest degree is 5 of A.

so $A^5 = aA^4 + bA^3 + cA^2 + dA + eI$

A^+, `U, S, V = svd(A)` S是 vec

`r = rank(Diagonal(S))`

```
if r > 0
  Ur = U[:, 1:r]
  Vr = V[:, 1:r]
  S = S[1:r]    Sr = Diagonal(S)
```

if r = 0 return A' end

`Sr = Diagonal(...)`

`V[:, r:-1:1]`

`U[:, r:-1:1]`

`S[r:-1:1]`

`Sr = inv(Sr)`