99.0 Review / practice questions

Disclaimer: this set of summary remarks and questions is far from being comprehensive!

The goal is to highlight some of the main take-aways, while also reiterating some subtle points.

$\begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$

Ch01: Matrices

shapes / transpose / symmetry / multiplication / orthogonality / determinant / eigenvalues / trace

$\mathbb{R}$

1. If $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_N$ are **orthogonal** vectors in $\mathbb{F}^N$, then $\boldsymbol{U} = \begin{bmatrix} \boldsymbol{u}_1 & \ldots & \boldsymbol{u}_N \end{bmatrix}$ is an **orthogonal matrix**. (?)
A: True B: False $\boxed{??}$

$\mathbb{R}$

2. If $\boldsymbol{q}_1, \ldots, \boldsymbol{q}_N$ are **orthogonal** vectors in $\mathbb{F}^N$, then there exist nonzero scalars $\alpha_1, \ldots, \alpha_N \in \mathbb{F}$ such that $\boldsymbol{Q} = \begin{bmatrix} \alpha_1 \boldsymbol{q}_1 & \ldots & \alpha_N \boldsymbol{q}_N \end{bmatrix}$ is an **orthogonal matrix**. (?)
A: True B: False $\boxed{??}$

$q_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ $q_2 = \begin{bmatrix} 0 \\ 2 \end{bmatrix}$

$Q = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$

Z.' nonzero

True: $\alpha_n = \dfrac{1}{\|q_n\|}$

Ch02: Matrix decompositions

Venn diagram of matrices and decompositions

Letting $|\lambda_{(n)}|$ denote the nth largest magnitude eigenvalue of a square matrix, we have the following Venn diagram to summarize the decompositions discussed so far.

Rectangular SVD: $\boldsymbol{A} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}'$

Square $\boldsymbol{A}\boldsymbol{V} = \boldsymbol{V}\boldsymbol{\Lambda}$

Diagonalizable $\boldsymbol{A} = \boldsymbol{V}\boldsymbol{\Lambda}\boldsymbol{V}^{-1}$ iff $\boldsymbol{V}$ linearly ind.

Normal $\boldsymbol{A} = \boldsymbol{V}\boldsymbol{\Lambda}\boldsymbol{V}'$ $\boldsymbol{V}$ unitary $\sigma_n = |\lambda_{(n)}|$

Hermitian $\boldsymbol{\Lambda}$ real

PSD: $\boldsymbol{A} \succeq \boldsymbol{0}$ $\boldsymbol{\Sigma} = \boldsymbol{\Lambda} \succeq \boldsymbol{0}$ $\boldsymbol{U} = \boldsymbol{V}$

For another Venn diagram, see this online "matrix world" figure from Gilbert Strang's book "Linear algebra for everyone."

3. If $\boldsymbol{A}$ is positive semidefinite in $\mathbb{F}^{N\times N}$ with rank r, then there exists a set of orthonormal vectors $\{\boldsymbol{z}_1, \ldots, \boldsymbol{z}_r\}$ in $\mathbb{F}^N$ such that $\boldsymbol{A} = \sum_{k=1}^{r} \sigma_k \boldsymbol{z}_k \boldsymbol{z}_k'$ is a **compact SVD** of $\boldsymbol{A}$.

A: True B: False **??**

V_r

$A = V \Sigma V'$

4. If a companion matrix has repeated eigenvalues, then it is not diagonalizable. $\iff$

A: True B: False ??

5. If a square matrix has repeated eigenvalues, then it is not diagonalizable. $\Rightarrow$ I

A: True B: False ??

If A has distinct eigvals, then A is diagonalizable,

6. If matrix $\boldsymbol{B}$ has orthonormal columns, then its singular values are all unity.

A: True $B'B = I$ B: False $B'B = I$ $B = U\Sigma V'$ $= VV'$??

$B = BII$ U_r $[B \; \square] \begin{bmatrix} I \\ 0 \end{bmatrix} I'$ $M \times (M-N)$

2021-12-09

7. If $\boldsymbol{B} \in \mathbb{F}^{M \times K}$ and $\boldsymbol{C} \in \mathbb{F}^{K \times N}$ each have orthonormal columns then $\boldsymbol{A} = \boldsymbol{BC}$ has orthonormal columns.

$A'A = (BC)'(BC) = C'B'BC = I$ gen. of HW

A: True B: False ??

$A'A = (BC)'(BC) = C'\underbrace{B'B}_{I}C = C'C = I$ ✓ $U_1 U_2 \cdots U_N$

8. If tall matrix $\boldsymbol{X}$ has singular values that are all unity, then $\boldsymbol{X} \in \mathbb{F}^{M \times N}$ has orthonormal columns.

A: True $[\;\;]$ B: False $X = UIV'$ $X'X = VU'UV' = I$??

$X = U_N I V'$ $X'X = (U_N V')'(U_N V') = V\underbrace{U_N' U_N}_{I} V' = \cancel{V}V' = I$

$U'U = UU' = I$ $U_r' U_r = I_r$ $U_r U_r' \neq I$

rank(A) ≤ min(m,n) = N

Ch03: Subspaces

span / linear dependence / basis / dimension / direct sum / orthogonal complement / range / nullspace / rank / four fundamental spaces / orthogonal bases / projection

9. If $\boldsymbol{A}$ is a $M \times N$ matrix with $M > N$, then $\mathrm{dim}(\mathcal{R}(\boldsymbol{A})) = N$. (rank)

rank(A) ≤ min(M,N) = N

A: True B: False ??

A = 0

10. If $\mathcal{S}$ is a subspace with basis vectors $\{\boldsymbol{b}_1, \dots, \boldsymbol{b}_K\}$ then $\mathrm{dim}(\mathcal{S}) = K$ (?)

lin. ind.

dim(S) ≤ N ✓

A: True B: False ??

If S = span({v₁, …, v_N}) dim(S) = N ? False

S = R(B)

11. If $\mathcal{S}$ is a **subspace** with **basis** $\{\boldsymbol{b}_1, \dots, \boldsymbol{b}_K\}$ and we define $\boldsymbol{B} = \begin{bmatrix} \boldsymbol{b}_1 & \dots & \boldsymbol{b}_K \end{bmatrix}$, then $\boldsymbol{P}_{\mathcal{S}}$ is (?)

A: $\boldsymbol{B}\boldsymbol{B}'$ B: $\boldsymbol{B}'\boldsymbol{B}$ C: $\boldsymbol{B}^+\boldsymbol{B}$ D: $\boldsymbol{B}\boldsymbol{B}^+$ E: None of these. ??

If B has orthonormal columns

then $P_{R(B)} = BB'$

$P_S(y) = B \cdot \underbrace{\left(\arg\min_x \|Bx - y\|\right)}_{B^+ y} = BB^+ y$

$P_S^\perp = I - BB^+$

12. If $\mathcal{S}$ is a subspace with **orthogonal basis** $\{\boldsymbol{b}_1, \ldots, \boldsymbol{b}_K\}$ then $\boldsymbol{P}_{\mathcal{S}}^\perp = \boldsymbol{I} - \boldsymbol{B}\boldsymbol{B}'$. (?)

A: True B: False $\boxed{??}$

orthonormal

$B^+ = B'$

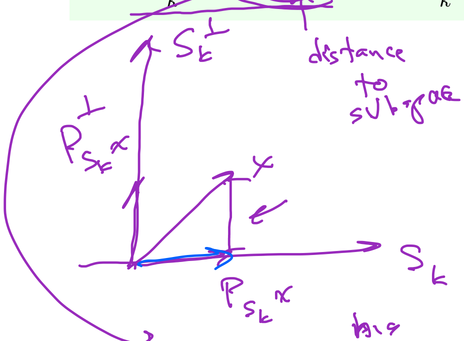
$b_1 = 1_N \quad S = \operatorname{span}(b_1) \quad P_S^\perp = I - \frac{1}{N} 1 1'$

$I - \frac{1}{\|b_1\|^2} b_1 b_1' \quad = I - \left(\frac{1}{\sqrt{N}} 1\right)\left(\frac{1}{\sqrt{N}} 1'\right)$

13. If $\{\mathcal{S}_1, \ldots, \mathcal{S}_K\}$ are K subspaces, then an appropriate nearest-subspace classifier of $\boldsymbol{x}$ is:

A: $\underset{k}{\arg\min} \left\| \boldsymbol{P}_{\mathcal{S}_k}^\perp \boldsymbol{x} \right\|$ B: $\underset{k}{\arg\min} \left\| \boldsymbol{P}_{\mathcal{S}_k} \boldsymbol{x} \right\|$ C: $\underset{k}{\arg\max} \left\| \boldsymbol{P}_{\mathcal{S}_k}^\perp \boldsymbol{x} \right\|$ D: $\boxed{\underset{k}{\arg\max} \left\| \boldsymbol{P}_{\mathcal{S}_k} \boldsymbol{x} \right\|}$ E: None of these $\boxed{??}$

distance to subspace



$x = P_S^\perp x + \underline{P_S x}$

$\|x\|^2 = \underbrace{\|P_{S_k}^\perp x\|^2}_{\text{min}} + \underbrace{\|P_{S_k} x\|^2}_{\text{max}}$

$(I - \overbrace{Q_k Q_k'}^{\text{big}}) x$

$x - Q_k \overbrace{(Q_k' x)}$

$\|x\|_p = \left(\sum_i |x_i|^p\right)^{1/p}$

$\left\| \begin{bmatrix} x \\ y \end{bmatrix} \right\|_p = \left(\sum_i |x_i|^p + \sum_j |y_j|^p \right)^{1/p} = \left(\|x\|_p^p + \|y\|_p^p \right)^{1/p}$

$\left\| \begin{bmatrix} x \\ y \end{bmatrix} \right\|_\infty = \max\left(\|x\|_\infty, \|y\|_\infty \right)$

Ch04: Linear least-squares

pseudo-inverse / Tikhonov regularization / truncated SVD / frames / projection / orthogonal projection

all !

14. If $\boldsymbol{A} \in \mathbb{F}^{M \times N}$ has compact SVD $\boldsymbol{U}_r \boldsymbol{\Sigma}_r \boldsymbol{V}_r'$, then a solution to the **LS problem** $\arg\min_{\boldsymbol{x}} \|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{y}\|_2$ is $\boldsymbol{x}_* = \underbrace{\boldsymbol{V}_r \boldsymbol{\Sigma}_r^{-1} \boldsymbol{U}_r'}_{A^+} \boldsymbol{y} + \underbrace{\underbrace{(\boldsymbol{I} - \boldsymbol{V}_r \boldsymbol{V}_r')}_{P_{N(A)}} \boldsymbol{z}}_{\in N(A)}$ for any vector $\boldsymbol{z} \in \mathbb{F}^N$. (?)

A: True B: False ??

15. A solution to $\arg\min_{\boldsymbol{x}} \left(\|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{y}\|_2^2 + \|\boldsymbol{B}\boldsymbol{x} - \boldsymbol{z}\|_2^2 \right)$ is: $= \arg\min_x \left\| \underbrace{\begin{bmatrix} A \\ B \end{bmatrix}}_{\text{vcat}} x - \begin{bmatrix} y \\ z \end{bmatrix} \right\|_2^2$

$\left\| \begin{bmatrix} x \\ y \end{bmatrix} \right\|_1 = \|x\|_1 + \|y\|_1$ $p=1$

A: $\boldsymbol{A}^+\boldsymbol{y} + \boldsymbol{B}^+\boldsymbol{z}$ B: $(\boldsymbol{A} + \boldsymbol{B})^+(\boldsymbol{y} + \boldsymbol{z})$ C: $[\boldsymbol{A}, \boldsymbol{B}]^+ [\boldsymbol{y}, \boldsymbol{z}]$ D: $\begin{bmatrix} \boldsymbol{A} \\ \boldsymbol{B} \end{bmatrix}^+ \begin{bmatrix} \boldsymbol{y} \\ \boldsymbol{z} \end{bmatrix}$ E: None of these. ??

$0 = \nabla \cdots = 2\left(A'(Ax - y) + B'(Bx - z) \right)$

16. If $\boldsymbol{\Phi} \in \mathbb{F}^{N \times M}$ is a **tight frame** and $\boldsymbol{U}$ is a $M \times M$ unitary matrix, then $\boldsymbol{B} = \boldsymbol{\Phi}\boldsymbol{U}$ is a frame. (?)

$\Phi \Phi' = \alpha I$

A: True B: False ??

$BB' = (\Phi U)(\Phi U)' = \Phi \Phi' = \alpha I \Rightarrow$ tight frame

$BB' = (\Phi U)(\Phi U)' = \alpha I$

17. $\mathbf{0}_{N\times N}$, $\boldsymbol{I}_N$ and $(\mathbf{1}_N\mathbf{1}_N')/N$ are **orthogonal projection matrices** on $\mathbb{F}^N$. (?)

A: T,T,T B: T,T,F C: T,F,F D: F,T,T E: F,F,F **??**

$P P = P$

$\frac{x x'}{\|x\|^2} = P_{\text{Span}(x)}$

$\underbrace{\frac{\mathbf{1}\overbrace{\mathbf{1}' \cdot \mathbf{1}}^{N}}{N}}_{P} \underbrace{\frac{\mathbf{1}'}{N}}_{P} = \underbrace{\frac{\mathbf{1}\mathbf{1}'}{N}}_{P}$

$$\begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{bmatrix}$$

$P = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix}'$

18. If P is an **orthogonal projection matrix**, then $P = P^+$.
A: True B: False ??

??

Q

tall ~ socially

- $\underbrace{P = QQ'}_{r \neq 0,\ Q'Q = I} = \underbrace{QIQ'}_{\text{compact SVD}} \Rightarrow P^+ = Q I^{-1} Q' = P$
- $P = 0 = 0^+$

If $P = 0$

$PP = P = P^+$

If $P \neq 0$

$P = QQ'$

19. A converse: If $\underbrace{A = A^+}_{\text{square}} \underbrace{\succeq 0}_{\text{Herm. sym.}}$, then A is an **orthogonal projection matrix**.
A: True B: False ??

??

$A = U_r \Sigma_r$

- $A = 0 = A^+$
- $A \neq 0$ $\quad A = U_r \Sigma_r U_r' = A^+ = U_r \Sigma_r^{-1} U_r'$

$\Rightarrow \underbrace{U_r' U_r}\Sigma_r \underbrace{U_r' U_r} = \underbrace{U_r' U_r}\Sigma_r^{-1} \underbrace{U_r' U_r}$

$\Rightarrow \Sigma_r = \Sigma_r^{-1} \Rightarrow \Sigma_r = I$

$\Rightarrow A = U_r U_r'$ ✓

20. If $\boldsymbol{A} = \boldsymbol{A}^+$ is a Hermitian matrix, then $\boldsymbol{A}$ is an **orthogonal projection matrix**.

A: True B: False **??**

$A = [-1]$ $= \begin{bmatrix} -1 & 0 \\ 0 & 0 \end{bmatrix} = A^+ \neq QQ'$

$u = \frac{\begin{bmatrix} 1 \\ 1 \end{bmatrix}}{\sqrt{2}}$ $P_S = uu'$

$P_S \begin{bmatrix} 3 \\ 1 \end{bmatrix} = uu' \begin{bmatrix} 3 \\ 1 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 \\ 1 \end{bmatrix} [1\ 1] \begin{bmatrix} 3 \\ 1 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 \\ 1 \end{bmatrix} 4 = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$

matrix

21. Let $\mathcal{S} = \{(x_1, x_2) \in \mathbb{R}^2 \ : \ x_1 = x_2\}$. Determine $\mathbf{1}' \boldsymbol{P}_{\mathcal{S}} \left[\begin{smallmatrix} 3 \\ 1 \end{smallmatrix}\right]$.

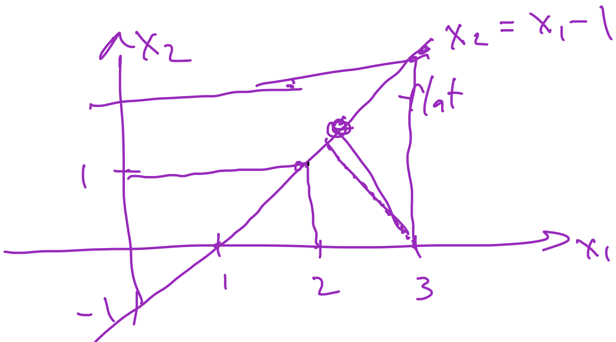
A: 0 B: 1 C: 2 D: 3 E: 4 **??**

span $(\mathbf{1}_2)$

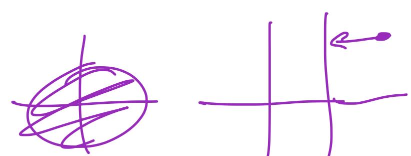
$P_S = \frac{\mathbf{1}_2 \mathbf{1}_2'}{2}$

$$\boxed{P_S \begin{bmatrix} 3 \\ 1 \end{bmatrix} = \frac{\mathbf{1}_2}{2} \underbrace{[1\ 1] \begin{bmatrix} 3 \\ 1 \end{bmatrix}}_{4} = 2 \cdot \mathbf{1}_2 = \begin{bmatrix} 2 \\ 2 \end{bmatrix}}$$

$2 \left(\frac{\mathbf{1}_2' \mathbf{1}_2}{2} \right) = 4$



$S = \{ (x_1, x_2) : x_1 = x_2 + 1 \}$



$P_S \left(\begin{bmatrix} 3 \\ 1 \end{bmatrix} \right)$

not bold

22. Let $\mathcal{S} = \{\boldsymbol{A} \in \mathbb{F}^{N\times N} : \boldsymbol{A} = \boldsymbol{A}'\}$ (Hermitian matrices). Then $\mathcal{S}$ is a (convex set, subspace)?

A: (F,F) B: (F,T) C: (T,F) D: (T,T) ??

??

$\mathbb{R}^{N\times N}$

$A = \begin{bmatrix} 2 & 3i \\ -3i & -4 \end{bmatrix}$

$A = A'$

$B = iA$ ← not Hermitian

$B' = -i A' = -i A \neq B$

x

S $P_S(x) \in S$

23. Let $\mathcal{S} = \{\boldsymbol{A} \in \mathbb{R}^{N\times N} : \boldsymbol{A} = \boldsymbol{A}'\}$. Determine $\mathcal{P}_{\mathcal{S}}(\boldsymbol{B})$.

A: $\boldsymbol{B}$ B: $\boldsymbol{B}'$ C: $2\boldsymbol{B}$ D: $2\boldsymbol{B}'$ E: $(\boldsymbol{B} + \boldsymbol{B}')/2$??

$\arg\min_{A \in S} \|A - B\|_F$ b_{ij} b_{ji}

24. If $\boldsymbol{U}$ and $\boldsymbol{Q}$ are $N \times N$ unitary matrices and $\boldsymbol{A} = [\gamma \boldsymbol{U} \;\; \delta \boldsymbol{Q}]$ with α,β not both zero, then $\boldsymbol{A}$ is a

A: Parseval tight frame B: tight frame with frame bound $\gamma + \delta$

C: tight frame with frame bound $\gamma^2 + \delta^2$ D: tight frame with frame bound $\sqrt{\gamma^2 + \delta^2}$

E: None of these TF with bound $|\gamma|^2 + |\delta|^2$??

$$A A' = [\gamma U \;\; \delta Q] \begin{bmatrix} \gamma^* U' \\ \delta^* Q' \end{bmatrix} = |\gamma|^2 I + |\delta|^2 I = (|\gamma|^2 + |\delta|^2) I$$

$(\alpha A)' = \alpha^* A'$ $(2+3i)^* = (2-3i)$ $(2+3im)'$

A : 30 , B : 60 C, Sat. D 2022-12/17

Ch05: Norms

vector norms / inner products / Cauchy-Schwarz inequality / matrix norms / induced norms / unitary invariance / norm equivalence / spectral radius / orthogonal Procrustes problem / Stiefel manifold

25. If $\boldsymbol{W}$ is a (**positive semidefinite**, **positive definite**) matrix, then $\|\boldsymbol{x}\|_{\boldsymbol{W}} \triangleq \sqrt{\boldsymbol{x}'\boldsymbol{W}\boldsymbol{x}}$ is a norm. (?)

A: (F,F) B: (F,T) C: (T,F) D: (T,T) $\boxed{??}$

$\boxed{??}$

$W=0$ $\|x\|_W = 0$ even for $x \neq 0$

$Q = UV'$ where $BA' = U\Sigma V'$

26. If $\boldsymbol{B}$ and $\boldsymbol{A}$ are $M \times N$ then $\operatorname{tr}((B - UV'A)(B-UV'A)')$ $= \operatorname{tr}(BB')$

$$\min_{\boldsymbol{Q}\,:\,\boldsymbol{Q}'\boldsymbol{Q}=\boldsymbol{I}_M} \|\boldsymbol{B} - \boldsymbol{Q}\boldsymbol{A}\|_{\mathrm{F}}^2 = \|\boldsymbol{A}\|_{\mathrm{F}}^2 + \|\boldsymbol{B}\|_{\mathrm{F}}^2 - 2\,\|\boldsymbol{B}\boldsymbol{A}'\|_* .$$

$- BA'VU - \operatorname{tr}(UV'AB')$

(?)

A: True B: False $\boxed{??}$

$\boxed{??}$

argmin $Q = UV'$ where $BA' = U\Sigma V'$

$\|BA'\|_*$

$\|B - UV'A\|_F^2 = \operatorname{tr}((B-UV'A)(B-UV'A)')$

$= \operatorname{tr}(BB') + \operatorname{tr}(AA') - \operatorname{tr}(\overset{U\Sigma V'}{BA'}VU') - \operatorname{tr}(UV'AB') \Rightarrow -2\operatorname{tr}(\Sigma) = \|BA'\|_*$

$\|X\|_{S,4}^4 = \sum_{k=1}^{r=1} \sigma_k^4$ $\|x-y\|_{S,4}^4 = (s_k - \sigma_k)^4 + \beta \mathbb{I}$

$\hat{X} = \sum w_k \cdot u_k v_k'$

Ch06: Low-rank approximation

Frobenius norm / generalizations to UI norms / proximal operators / hard thresholding / soft thresholding / choosing β or rank / OptShrink / Subspace learning

$\hat{w}_k = \arg\min (\sigma_k - s_k)^4 + \beta \mathbb{I}_{\{s \neq 0\}}$

V^4 unitless

27. Determine $\hat{X} = \underset{X \in \mathbb{F}^{M\times N}}{\arg\min} \underbrace{\|X - Y\|_{S,4}^4}_{\text{UI}} + \beta\, \mathrm{rank}(X)$ and consider the *units* of β and $\hat{w}_k$.

$\boxed{??}$

$\hat{X} = \sum_{k=1}^{r} \hat{w}_k u_k v_k'$

$Y = U_r \Sigma_r V_r'$ units of Volts

$\hat{w}_k = \underset{s \geq 0}{\arg\min} \underbrace{(s - \sigma_k)^4}_{\sigma} + \underbrace{\beta \mathbb{I}_{\{s \neq 0\}}}_{\beta}$ $\sigma_k^4 \gtrless \beta$

$= \begin{cases} \sigma_k, & \sigma_k^4 > \beta \\ 0, & \sigma_k^4 < \beta \end{cases}$

$\sigma_k \gtrless \beta^{1/4}$

V $\underbrace{(V^4)^{1/4}}_{V}$

Y units Volts

28. Determine $\hat{\boldsymbol{X}} = \underset{\boldsymbol{X}\in\mathbb{F}^{M\times N}}{\arg\min} \frac{1}{4} \|\boldsymbol{X}-\boldsymbol{Y}\|_{S,4}^4 + \beta \|\boldsymbol{X}\|_*$. Think about units again.

$V^4 \quad V^3 \quad V$

??

$\hat{w}_k = \underset{s\ge 0}{\arg\min} \frac{1}{4}(s-\sigma_k)^4 + \beta s$

$(s-\sigma_k)^3 + \beta \cdot 1 = 0$

$-(\sigma_k - s)^3 + \beta = 0$

$\beta = (\sigma_k - s)^3$

$\beta^{1/3} = \sigma_k - s$

$s = \sigma_k - \beta^{1/3}$

$\hat{w}_k = \begin{cases} \sigma_k - \beta^{1/3}, & \sigma_k \ge \beta^{1/3} \\ 0, & \text{else} \end{cases} = \max(\sigma_k - \beta^{1/3}, 0)$

$\frac{1}{(p-1)}$

29. The ϵ-rank of a matrix is defined as follows for $0 < \epsilon < 1$ [1]:

$$\text{rank}_\epsilon(\boldsymbol{X}) = \underset{k\ge 0}{\arg\min} \{k : \sigma_{k+1}(\boldsymbol{X}) \le \epsilon \underbrace{\|\boldsymbol{X}\|_2}_{\sigma_1(X)}\}.$$

$\hat{X} - X_{\text{true}}$ Hard

$Y = X_{\text{true}} + \varepsilon$

For a $M \times N$ matrix $\boldsymbol{A}$ with rank r and some given ϵ, let $K = \text{rank}_\epsilon(\boldsymbol{A})$ and define $\hat{\boldsymbol{A}}_K$ to be the usual rank-at-most-K approximation to $\boldsymbol{A}$. Give an upper bound for $\|\hat{\boldsymbol{A}}_K - \boldsymbol{A}\|_F$.

A: ϵ B: ϵr C: ϵK D: $\epsilon(r-k)$ E: None of these

??

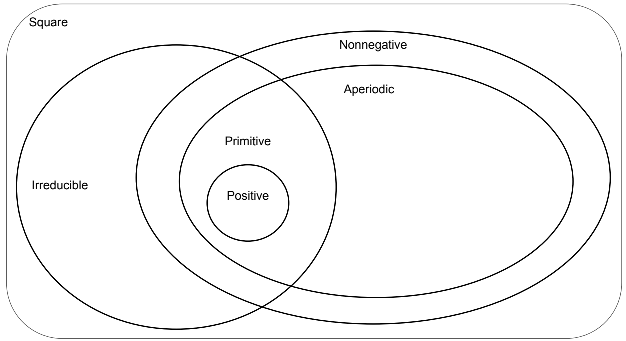
$\|\hat{X} - Y\|_?$

$\|\hat{A}_K - A\|_F = \sqrt{\sum_{k=K+1}^r \sigma_k^2} \le \sqrt{\sum_{k=K+1}^r (\varepsilon\sigma_1)^2}$

$\boxed{= \varepsilon\sigma_1\sqrt{r-K}}$

Ch07: Special matrices

companion / Vandermond / circulant / power iteration / primitive / irreducible / Markov chains / PageRank



30. If C is a **Hermitian circulant** matrix, then the vector $\boldsymbol{x}$ with elements $x_n = \cos(2\pi kn/N)$ is an **eigenvector** of C for any $k \in \mathbb{Z}$.

A: True B: False **??** ♦♦

??

31. We have two Markov chains (with M,N states) and we define a new Markov chain as follows. We flip a coin and start in one of the two chains; at each time we roll a die and if we get a 5 we jump to a random state in the other chain, otherwise we follow the transition probability for the current chain.
This new Markov chain (has a unique equilibrium distribution, to which the power iteration for its transition matrix always converges). (?)

A: (F,F) B: (F,T) C: (T,F) D: (T,T) **??**

$$Q = \begin{bmatrix} \frac{5}{6} P_1 & \frac{1}{6M} 1_M 1_N' \\ \frac{1}{6N} 1_N 1_M' & \frac{5}{6} P_2 \end{bmatrix}$$

Q^2 is positive

5

32. The **nuclear norm** of the $N \times N$ **circulant matrix** $\boldsymbol{C}$ whose first column is $(0, 0, \ldots, 1)$ is

A: 1 B: N C: N^2 D: $N^2/2$ E: None of these **??**

??

$$C_3 = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$$

permutation matrix

$C'C = I$

$\sigma_k^2(C) = \mathrm{eig}_{(k)}(C'C)$ $\sigma_k = 1$

$$\mathrm{DFT}\left(\begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}\right) = e^{-i \frac{2\pi}{N} (N-1) k}$$

eigs of C $\sigma_k = |e^{i \cdots}| = 1$

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

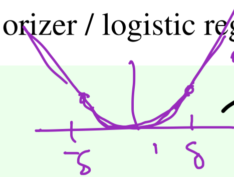
normal

$$\dot\psi(x) = \begin{cases} x & |x| < \delta \\ \delta & x > \delta \\ -\delta & x < \delta \end{cases}$$

2: $\dfrac{\|f(x) - f(z)\|}{\|x - z\|} \le L$, i.e. $\|f(x) - f(z)\| \le L \|x - z\|$

Ch08: Optimization

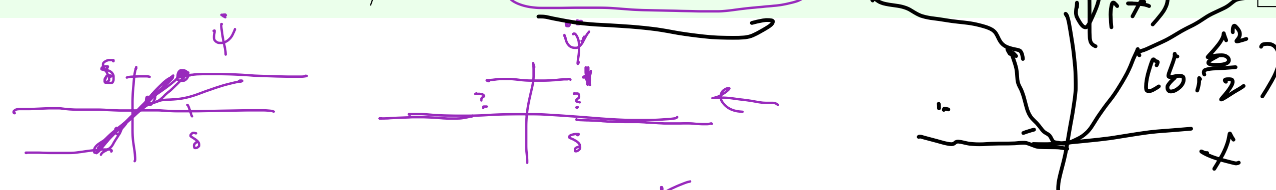
GD / preconditioning / step size / Lipschitz continuity / GP / convexity / majorizer / logistic regression



33. The Huber **potential function** is $\psi(x) = \begin{cases} \frac{1}{2}x^2, & |x| \le \delta \\ \delta|x| - \frac{1}{2}\delta^2, & \text{otherwise.} \end{cases}$

Is the derivative of ψ **Lipschitz continuous**?

A: No B: Yes with $L = 1/2$ C: Yes with $L = 1$ D: Yes with $L = \delta$??



34. A Lipschitz constant for the gradient of $f(\boldsymbol{x}) \triangleq \mathbf{1}'_M \psi.(\boldsymbol{A}\boldsymbol{x})$, where $\boldsymbol{A}$ is $M \times N$ and ψ is the Huber function above, is:

A: $\|\boldsymbol{A}'\boldsymbol{A}\|_2$ B: $\|\boldsymbol{A}\|_2$ C: $M\|\boldsymbol{A}\|_2$ D: $N\|\boldsymbol{A}\|_2$ E: None of these ??

$\nabla f(x) = A' \dot\psi.(Ax)$ Ch 8

$\nabla^2 f(x) = A' D(x) A$ $D(x) = \mathrm{Diag}(\ddot\psi([Ax]_i)) \preceq I$

$\preceq A' I A = A'A$ $L \le \|A'A\|_2 =$

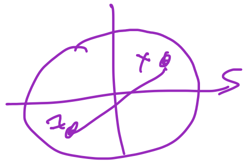
$D(x) = \begin{bmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 0 \end{bmatrix}$

$I - D \succeq 0$

$L = \|A'A\|_2 \cdot L_{\dot\psi}$

35. $\mathcal{C} = \{ \boldsymbol{x} \in \mathbb{R}^N : \|\boldsymbol{x}\|_2 \leq 5 \}$ is **convex** set (?)

A: True   B: False   ??



$\|x\| \leq 5 \qquad \|z\| \leq 5$

convex combo

$y = \alpha x + (1-\alpha) z \overset{?}{\in} C \qquad \alpha \in [0,1]$

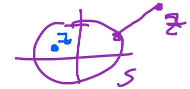
$\|y\| = \|\alpha x + (1-\alpha) z\|_2 \leq \|\alpha x\| + \|(1-\alpha) z\| = \alpha \|x\| + (1-\alpha)\|z\| \leq \alpha 5 + (1-\alpha) 5 = 5$ ✓

$\|x\|_2 \leq 5 \qquad \|z\| \leq 5$

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

$$\|\alpha x + (1-\alpha) z\|_2 \leq \|\alpha x\|_2 + \|(1-\alpha) z\|_2$$
$$\leq \alpha \|x\| + (1-\alpha)\|z\|_2$$
$$\leq 5\alpha + (1-\alpha)\cdot 5$$
$$\leq 5$$

$\therefore$



36. Which of these is the **projection** of a nonzero point $\boldsymbol{z} \in \mathbb{R}^N$ onto $\mathcal{C} = \{\boldsymbol{x} \in \mathbb{R}^N : \|\boldsymbol{x}\|_2 \leq 5\}$?

A: `min(1,5/norm(z))*z` B: `max(1,5/norm(z))*z` C: `(5/norm(z))*z`

D: `min.(abs.(z),5).*sign(z)` E: None of these ??

A $\sigma = \frac{2}{\sigma_1(A)+\sigma_N(A)} = \frac{2}{}$ proof?

37. When solving a LS problem via **gradient descent**, where $\boldsymbol{A}$ is unitary, what is the optimal **step size**?

A: $1/4$ B: $1/2$ C: 1 D: 2 E: None of these ??

$\alpha_* = \frac{2}{\sigma_1^2(A) + \sigma_N^2(A)} = \frac{2}{1+1} = 1$

$\hat{x} = A^+ y = A' y$

$x_1 = x_0 - \alpha A'(\overset{I}{A} x_0 - y) = x_0 - \alpha x_0 + \alpha A' y = A' y$ if $\alpha = 1$

38. Suppose $\boldsymbol{A}$ is a wide matrix and consider the regularized LS cost function $f(\boldsymbol{x}) \triangleq \frac{1}{2} \|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{y}\|_2^2 + \beta \|\boldsymbol{U}\boldsymbol{x}\|_2^2$ where $\beta > 0$ and $\boldsymbol{U}$ is a unitary matrix. Then f is a **strictly convex** function. (?)

A: True B: False ??

$\nabla^2 f_{(x)} = A'A + \beta 2 U'U = \underbrace{A'A}_{\succeq 0} + \underbrace{2\beta I}_{\succ 0} \succ 0$

$\underbrace{\frac{1}{2}\|Ax - y\|_2^2}_{\text{convex}} + \underbrace{\beta \|x\|_2^2}_{\text{strictly convex}} \Rightarrow$ strictly convex