Homework 1

WU Jiamin 20659352

3. (a) ($\Rightarrow$) if $K \succeq 0$, we want it's eigenvalue λ and eigenvector x.

where $Kx = \lambda x$.

Then for every $x \in \mathbb{R}^n$, $x^T K x \geq 0$

$$x^T K x = x^T \lambda x = \lambda x^T x \geq 0 \Rightarrow \lambda \geq 0$$

($\Leftarrow$) if all $\lambda_i \geq 0$ for $i = 1, \dots, n$, for the vector x, by the spectral decomposition, we have

$$x^T K x = (x^T U) D (U^T x) = \sum_{i=1}^{n} \lambda_i (x^T T_i)^2$$

as $\lambda_i \geq 0$, $x^T K x \geq 0$ so K is positive semi definite.

(b) By the spectral decomposition, we have

$K = U^T D U$ for some U and diagonal matrix D.

then $K_{ij} = \sum_{i=1}^{n} D_{ll} U_{il} U_{jl}$

$$\begin{aligned} d_{ij} &= K_{ii} + K_{jj} - 2K_{ij} \\ &= \sum_{i=1}^{n} D_{ll} U_{il}^2 + \sum_{i=1}^{n} D_{ll} U_{jl}^2 - 2 \sum_{i=1}^{n} D_{ll} U_{il} U_{jl} \\ &= \sum_{i=1}^{n} D_{ll} (U_{il} - U_{jl})^2 \\ &= \|U_i - U_j\|^2 \end{aligned}$$

(C) $D = [d_{ij}]$ $B_a = -\frac{1}{2} H_\alpha D H_\alpha^T$

$$= -\frac{1}{2} H_a (k_{ii} + k_{jj} - 2k_{ij}) H_\alpha^T$$

Let $K = M^T M$

$$\begin{aligned} B_a &= -\frac{1}{2} (H_\alpha K_{ii} H_\alpha^T + H_\alpha K_{jj} H_\alpha^T - 2 H_\alpha K_{ij} H_\alpha^T) \\ &= -\frac{1}{2} [H_\alpha K (H_\alpha I)^T + H_\alpha K^T (H_\alpha I)] + H_\alpha 2K H_\alpha^T \end{aligned}$$

$$H_\alpha \cdot I = (I - e\alpha^T) \cdot I = I - e\alpha^T \cdot I$$

$$e^T\alpha = 1 \Rightarrow e = (I \cdot \alpha^T)^T$$

$$H_\alpha \cdot I = I - (I \cdot \alpha^T)^T \cdot \alpha^T \cdot I = 0$$

$$\Rightarrow B_\alpha = H_\alpha 2K H_\alpha^T$$

$$= 2 H_\alpha M^T M H_\alpha^T$$

$$= 2 H_\alpha M^T (H_\alpha M^T)^T$$

$$v^T B_\alpha v = 2 v^T H_\alpha M^T (v^T H_\alpha M^T)^T \geq 0$$

So $B_\alpha \succeq 0$

(d). $A \succeq 0 \Rightarrow x^T A x \geq 0$

$B \succeq 0 \Rightarrow x^T B x \geq 0$

$x^T(A+B)x = x^T A x + x^T B x \geq 0$

So $A + B \succeq 0$

$A \succeq 0 \Rightarrow A = \sum_{i=1}^n \lambda_{A_i} e_{A_i} e_{A_i}^T$

$B \succeq 0 \Rightarrow B = \sum_{i=1}^n \lambda_{B_i} e_{B_i} e_{B_i}^T$

where λ_{A_i}, λ_{B_i} are eigenvalues for A and B,

e_{A_i}, e_{B_i} are eigenbasis for A and B.

Then $A \circ B = \sum_{i=1}^n \lambda_{A_i} e_{A_i} e_{A_i}^T \cdot \sum_{j=1}^n \lambda_{B_j} e_{B_j} e_{B_j}^T$

$$= \sum_{i,j} \lambda_{A_i} \lambda_{B_j} (e_{A_i} e_{A_i}^T)(e_{B_j} e_{B_j}^T)$$

$$= \sum_{i,j} \lambda_{A_i} \lambda_{B_j} (e_{A_i} e_{B_j})(e_{A_i} e_{B_j})^T$$

So $A \circ B \succeq 0$.

4. (a) d^2 may not be a distance function.

e.g. $X = (1, 0)$ $\quad Y = (2, 0)$ $\quad Z = (3, 0)$

$d(X, Y) = 1 \quad d(\cancel{X}\,y, Z) = 1 \quad d(X, Z) = 2$

$d^2(x,y)=1 \quad d^2(y,z)=1 \quad d^2(x,z)=4$

$d^2(x,y) + d^2(y,z) < d^2(x,z)$ contradicts triangle inequality.

(b). $\sqrt{d}$ is a distance function.

proof: M_1: $d \in \mathbb{R} \Rightarrow \sqrt{d} \in \mathbb{R} \quad d<\infty \Rightarrow \sqrt{d}<\infty$

$d \geq 0 \Rightarrow \sqrt{d} \geq 0$

M_2: $d(x,y)=0 \Leftrightarrow x=y$

$\sqrt{d(x,y)}=0 \Leftrightarrow x=y$

M_3: $d(x,y)=d(y,x) \Rightarrow \sqrt{d(x,y)} = \sqrt{d(y,x)}$

M_4: $d(x,y) \leq d(x,z) + d(z,y)$

$(\sqrt{d(x,y)})^2 \leq \sqrt{d(x,z)}^2 + \sqrt{d(z,y)}^2$

$(\sqrt{d(x,y)})^2 \leq \sqrt{d(x,z)}^2 + \sqrt{d(z,y)}^2 + 2\sqrt{d(x,z)}\sqrt{d(z,y)}$

$\leq (\sqrt{d(x,z)} + \sqrt{d(y,z)})^2$

$\Rightarrow \quad \sqrt{d(x,y)} \leq \sqrt{d(x,z)} + \sqrt{d(y,z)}$

$\sqrt{d}$ is a distance function because M_1, M_2, M_3, M_4 hold.

5. (a) A is an $m \times n$ real valued matrix, so $A^TA \succeq 0$

$\Rightarrow A^TA = V\Lambda V^T = \sum_{i=1}^{n} (\sigma_i)^2 v_i v_i^T$

where V_i is the eigenvector of A^TA, Λ is a diagonal matrix.

$A^TA v_i = (\sigma_i)^2 v_i$

Define $u_i = \frac{Av_i}{\sigma_i}$ $\quad U$ is $m \times m$ where the ith column is u_i.

Σ is the diagonal matrix with elements σ_i

$AA^T u_i = AA^T \left(\frac{Av_i}{\sigma_i}\right)$

$= A(\sigma_i)^2 v_i \frac{1}{\sigma_i^2}$

$= (\sigma_i)^2 u_i$

then $u_i^T u_i = \left(\frac{Av_i}{\sigma_i}\right)^T \left(\frac{Av_i}{\sigma_i}\right)$

$= v_i^T v_i$

$= 1$

Hence $U = AV\Sigma^{-1}$

$\Rightarrow U\Sigma = AV$

$\Rightarrow A = U\Sigma V^T$

(b). We have $A = U\Sigma V^T$, $A_k = U^T \Sigma_k V^T$

$\|A - A_k\| = \| \operatorname{diag}(0, \dots, 0, \sigma_{k+1}, \dots, \sigma_n) \|$

For $B \in \mathbb{R}^{m \times n}$ $\quad \operatorname{rank}_B \le k$

$\|A - A_k\|_2^2 \geq \|(A-B) w\|_2^2 = \|Aw\|^2 = \|A V_{k+1} V_{k+1}^T w\|^2$

$= \|U_{k+1} \Sigma_{k+1} V_{k+1}^T w\|^2$

$= \sum_{i=1}^{r+1} \sigma_i |v_i^T w|^2 \geq \sigma_{k+1} \sum_{i=1}^{k+1} |v_i^T w|^2$

$= \sigma_{k+1}$

(C) Let B minimize $\|A - B\|_F^2$ for all matrices whose ranks $\le k$.
V is the space spanned by the rows of B.
$\|A - B\|_F^2 \le \|A - M\|_F^2$, then B is the projection of rows for A on V. $\|A - B\|_F^2$ is the sum of A_i^2 on V.
A_k minimizes the sum of A_i^2 distance, then
$\|A - A_k\|_F \le \|A - B\|_F$.

(d) $\|A\|_p = \left(\sum_i \sigma_i^p\right)^{\frac{1}{p}}$

$QAZ = QU^T \Sigma V Z$

$= (UQ^T)^T \Sigma (VZ)$

$\Rightarrow \|QAZ\|_p = \|A\|_p$

f) $A = U\Sigma V^T \Rightarrow \Sigma = U^TAV$

$\|\Sigma - V^TRV\|_F^2 = \sum_{i=1}^{n}\sum_{j=1}^{n} R'^2_{ij} + \sum_{i=1}^{n}\sigma_i^2 - 2\sum_{i=1}^{n}\sigma_i^2 R_{ii}$

$\sum_{i=1}^{n} q_i^T R p_i = tr(Q^TRP) = tr(PQ^TR)$

$PQ^T = U\Sigma \cdot V^T$

$tr(\Sigma V^T R U) = \sum_{i=1}^{n}\sigma_i^2 V^T R_{ii} U$

then $R = VU^T$