3. *Positive Semi-definiteness:* Recall that a n-by-n real symmetric matrix K is called positive semi-definite (*p.s.d.* or $K \succeq 0$) iff for every $x \in \mathbb{R}^n$, $x^T K x \geq 0$.

(a) Show that $K \succeq 0$ if and only if its eigenvalues are all nonnegative.

(b) Show that $d_{ij} = K_{ii} + K_{jj} - 2K_{ij}$ is a squared distance function, *i.e.* there exists vectors $u_i, v_j \in \mathbb{R}^n$ $(1 \leq i, j \leq n)$ such that $d_{ij} = \|u_i - u_j\|^2$.

(c) Let $\alpha \in \mathbb{R}^n$ be a signed measure s.t. $\sum_i \alpha_i = 1$ (or $e^T\alpha = 1$) and $H_\alpha = I - e\alpha^T$ be the Householder centering matrix. Show that $B_\alpha = -\frac{1}{2}H_\alpha D H_\alpha^T \succeq 0$ for matrix $D = [d_{ij}]$.

(d) If $A \succeq 0$ and $B \succeq 0$ $(A, B \in \mathbb{R}^{n\times n})$, show that $A + B = [A_{ij} + B_{ij}]_{ij} \succeq 0$ (elementwise sum), and $A \circ B = [A_{ij}B_{ij}]_{ij} \succeq 0$ (Hadamard product or elementwise product).

(a) The eigenvalue λ and corresponding eigenvector v satisfies

$$Kv = \lambda v$$

① $\Rightarrow$

If $K \succeq 0$ $\quad v^T K v = v^T \lambda v = \lambda \cdot v^T v \geq 0$

$$\Rightarrow \lambda \geq 0$$

② $\Leftarrow$ If $\lambda_1 \geq \lambda_2 \cdots \geq \lambda_n \geq 0$ $\quad T = \mathrm{diag}\,\lambda$

$$K = Q^T T Q$$

$\forall x \in \mathbb{R}^n$ $\quad x^T K x = (Qx)^T T (Qx)$ $\quad Qx = \begin{pmatrix} p_1 \\ \vdots \\ p_n \end{pmatrix}$

$x^T K x = \sum_{i=1}^{n} \lambda_i p_i^2 \geq 0$ $\quad$ Thus $K \succeq 0$ $\blacksquare$

(b) $\|u_i - v_j\|^2 = (u_i - v_j)^T (u_i - v_j) = u_i^T v_i + u_j^T v_j - u_i^T v_j - u_i v_j^T$

We Assign.

u_i to be the i-th row of K

v_j to be the n-dim a zero vectors with only j-th element to be 1

Then for $\forall i, j$ we always have

$$\|u_i - v_j\|^2 = d_{ij} = K_{ii} + K_{jj} - 2K_{ij} \qquad \blacksquare$$

(c) We consider the proof of C later.

(d) If $A \succeq 0$, $B \succeq 0$

① $\forall x \in \mathbb{R}^n$

$x^T(A+B)x = x^TAx + x^TBx > 0+0 = 0$

$(A+B) \succeq 0$

② $B \succeq 0 \Rightarrow \exists T$ s.t. $B = TT^T$

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

$x^T(A \circ B)x = x^T(A \circ (TT^T))x$

$= \sum_{i,j=1}^{n} x_i a_{ij} \left(\sum_{k=1}^{n} t_{ik} t_{jk}\right) x_j$

$= \sum_{k=1}^{n} (x * t_k)' A (x * t_k) \geq \sum_{k=1}^{n} 0 = 0$

Where t_k is the k-th column of T

So $A * B \succeq 0$ ∎

(c) $D = ke^T + ek^T - 2K$ $\quad k = \operatorname{diag} K \in \mathbb{R}^n$

suppose $K = X^TX$

$B_\alpha = -\frac{1}{2} H_\alpha D H_\alpha^T = -\frac{1}{2} H_\alpha (ke^T + ek^T - 2K) H_\alpha^T$

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

$= (I - e\alpha^T) k (e^T - (e^T\alpha) e^T) = 0$

Likewise $H_\alpha e k^T H_\alpha^T = (e - e(\alpha^T e)) k H_\alpha^T = 0$

Thus. $B_\alpha = H_\alpha K H_\alpha^T$

Choose any $x \in \mathbb{R}^n$ $\quad x^T B_\alpha x = x^T H_\alpha K H_\alpha^T x$

$= (H_\alpha^T x)^T K (H_\alpha^T x) \geq 0$

Since K is p.s.d. Thus B_α is p.s.d. ∎

4. *Distance:* Suppose that $d : \mathbb{R}^p \times \mathbb{R}^p \to \mathbb{R}$ is a distance function.

(a) Is d^2 a distance function? Prove or give a counter example.

(b) Is $\sqrt{d}$ a distance function? Prove or give a counter example.

(a) d^2 is not a distance function.

Counter example

Pick three points on the plane $x=0$ $y=2$ $z=4$

$d(a,b) \triangleq |a-b|$

$d(x,y) = d(y,z) = 2$

d is a distance function satisfies:

1) $d(a,b) \geq 0$ and $d(a,b)=0$ iff $a=b$

2) $d(a,b) = d(b,a)$

3) $d(c,a) + d(c,b) \geq d(a,b)$

However, for $d^2(a,b) \triangleq (d(a,b))^2 = |a-b|^2$

$d^2(x,z) = 4^2 = 16 > 8 = d^2(x,y) + d^2(y,z)$

Condition 3 triangle inequality is broken

Thus d^2 is not distance function.

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

we prove it by verifying that $\phi(d) = d^{1/2}$ can be written in the form of Schoenberg Transformation.

because according to Note p45 Thm 5.2 Schoenberg transform characterizes all the transforms between squared distance matrices.

Assign $\quad g(\lambda) \overset{\Delta}{=} \dfrac{\frac{1}{2}}{\Gamma(\frac{1}{2})} \lambda^{-\frac{1}{2}}$

then $\quad \displaystyle\int_0^\infty \frac{1-\exp(-\lambda d)}{\lambda} g(\lambda)\, d\lambda$

$$= \int_0^\infty \frac{1-\exp(-\lambda d)}{\lambda} \cdot \frac{\frac{1}{2}}{\Gamma(\frac{1}{2})} \cdot \lambda^{-\frac{1}{2}}$$

$$= \frac{1}{2}\int_0^\infty \frac{1}{\Gamma(\frac{1}{2})} \lambda^{-\frac{3}{2}} \left(1-\exp(-\lambda d)\right)$$

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

So $\phi(d) = d^{\frac{1}{2}}$ is a schoenberg transform.

So $d^{\frac{1}{2}}$ is a distance function.