3 (a) Given $K \succeq 0 \Leftrightarrow x \in R^n, x^T K x \geq 0$, we will prove:

$x \in R^n, x^T K x \geq 0 \Leftrightarrow$ all eigenvalues are nonnegative.

$\Leftarrow$: $\exists Q$ s.t. $K = Q \Lambda Q^T$

$x^T K x = x^T Q \Lambda Q^T x = y^T \Lambda y = \lambda_1 y_1^2 + \cdots + \lambda_n y_n^2 \geq 0$

where $y = Q^T x = (y_1, \ldots, y_n) \neq 0$.

$\Rightarrow$: K is real symmetric matrix. $\Rightarrow$ $K = Q \Lambda Q^T$

$x^T K x \geq 0 \Rightarrow x^T Q \Lambda Q^T x \geq 0$

If $y = Q^T x$, $y \Lambda y^T \geq 0 \Rightarrow \lambda_i \geq 0$

QED

(b) $d_{ij} = K_{ii} + K_{jj} - 2K_{ij}$

$= u_i^T K u_i + v_j^T K v_j - 2 u_i^T K v_j$

$= (u_i - v_j)^T K (u_i - v_j)$

K is a symmetric matrix $\Rightarrow$ d_{ij} is a squared distance function.

(c) Squared distance function is conditionally negative define (c.n.d)

We will prove $B_\alpha = -\frac{1}{2} H_\alpha D H_\alpha^T \succeq 0$ for $D = [d_{ij}] \Leftrightarrow [d_{ij}]$ is c.n.d.

$\Leftarrow$: $\exists x \in R^n$, $x^T B_\alpha x = \frac{1}{2} x^T H_\alpha D H_\alpha^T x = -\frac{1}{2} (H_\alpha^T x)^T D (H_\alpha^T x)$

Now, we are going to show that $y = H_\alpha^T x$ satisfies $e^T y = 0$

In fact, $\because e^T H_\alpha^T x = e^T (I - e^T) x = (e - e^T \cdot \alpha) e^T \cdot x = 0$

as $e^T \cdot \alpha = 1$ for α.

Therefore: $x^T B_\alpha x = -\frac{1}{2} (H_\alpha^T x)^T D (H_\alpha^T x) \geq 0$

as D is c.n.d.

$\Rightarrow$: For $\forall x \in R^n$ s.t. $e^T x = 0$, we have

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

Therefore: $x^T D x = (H_\alpha^T \alpha)^T D (H_\alpha^T x) = x^T H_\alpha D H_\alpha^T x = -2 x^T B_\alpha x \leq 0$

QED.

(d) $A \succeq 0$ and $B \succeq 0 \Rightarrow A + B \succeq 0$ (property of positive semi-define matrix)

4 distance function:

① $d(A,B) \geq 0$
② $d(A,A)=0$
③ $d(A,B)=d(B,A)$
④ Triangle inequality: $d(A,B) \leq d(A,C) + d(C,B)$

(a) No. counter example:

$d(A,B)=5$. $d(A,C)=2$. $d(B,C)=3$.

$d^2(A,B)=25 > d^2(A,C)=4. + d^2(B,C)=9$.

which not satisfy triangle inequality

(b) Yes.

① $d(A,B) \geq 0 \Rightarrow \sqrt{d(A,B)} \geq 0$
② $d(A,A)=0 \Rightarrow \sqrt{d(A,A)} = 0$
③ $d(A,B) = d(B,A) \Rightarrow \sqrt{d(A,B)} = \sqrt{d(B,A)}$
④ $d(A,B) \leq d(A,C) + d(C,B) \Rightarrow$

$d(A,B) \leq d(A,C) + d(C,B) + 2\sqrt{d(A,C)\,d(C,B)} \Rightarrow$

$(\sqrt{d(A,B)})^2 \leq (\sqrt{d(A,C)} + \sqrt{d(C,B)})^2$

QED

5.(a) Let $x \in R^n$ and $y \in R^m$ be unit 2-norm vectors that satisfy $Ax = \sigma y$ with $\sigma = \|A\|_2$,

From Theorem 2.5.1 in Golub and Van Loan; there exists $V_2 \in R^{n\times(n-1)}$ and $U_2 \in R^{m\times(m-1)}$. So $V = [X_2\ V_2] \in R^{n\times n}$ and $U = [y\ U_2] \in R^{m\times m}$ are orthogonal. It is not hard to show that U^TAV has the following structure:

$$U^TAV = \begin{bmatrix} \sigma & w^T \\ 0 & B \end{bmatrix} \equiv A_1$$

Since:

$$\left\| A_1\left(\begin{bmatrix} \sigma \\ w \end{bmatrix}\right)\right\|_2^2 \geq (\sigma^2 + w^Tw)^2$$

we have $\|A_1\|_2^2 \geq (\sigma^2 + w^Tw)$. But $\sigma^2 = \|A\|_2^2 = \|A_1\|_2^2$ and so we must have $w=0$. An obvious induction argument complete the proof of the theorem

(b) Since $U^T A_k V = \mathrm{diag}(\sigma_1, \dots \sigma_k, 0, \dots, 0)$, it follows that $\mathrm{rank}(A_k) = k$ and that $U^T(A - A_k)V = \mathrm{diag}(0, \dots, 0, \sigma_{k+1}, \dots, \sigma_p)$ and so $\|A - A_k\|_2 = \sigma_{k+1}$

Now suppose $\mathrm{rank}(B) = k$ for some $B \in \mathbb{R}^{m \times n}$. It follows that we can find orthonormal vectors $x_1, \dots, x_{n-k}$, so $\mathrm{null}(B) = \mathrm{span}\{x_1, \dots, x_{n-k}\}$.

A dimension argument shows that

$$\mathrm{span}\{x_1, \dots, x_{n-k}\} \cap \mathrm{span}\{v_1, \dots, v_{k+1}\} \neq \{0\}$$

Let z be a unit 2-norm vetor in this intersection. Since $Bz = 0$ and

$$Az = \sum_{i=1}^{k+1} \sigma_i (v_i^T z) u_i$$

we have $\|A - B\|_2^2 \geq \|(A-B)z\|_2^2 = \|Az\|_2^2 = \sum_{i=1}^{k+1} \sigma_i^2 (v_i^T z)^2 \geq \sigma_{k+1}^2$

Q.E.D.

(c) Let B minimize $\|A - B\|_F^2$ among all rank or less matrices. Let V be the space spanned by the rows of B. The dimension of V is at most k.

Since B minimizes $\|A - B\|_F^2$, it must be that each row of B is the projection of the corresponding row of A onto V. otherwise replacing the row of B with the projection of the corresponding row of A onto V does not change V and hence the rank of B but would reduce $\|A - B\|_F^2$.

Since each row of B is the projection of the corresponding row of A, it would follow that $\|A - B\|_F^2$ is the sum of squared distances of rows of A to V.

Since A_k minimizes the sum of squared distance of rows of A to any k-dimensional subspace, it follows that $\|A - A_k\|_F \leq \|A - B\|_F$.

Q.E.D.