3

1. Since K is symmetric, there exists U and Λ such that $K = U^T \Lambda U$ where Λ is diagonal and U is orthogonal. Since $\Lambda_{ii} = (U^T e_i)^T K U^T e_i$ and $x^T K x = \sum_i \Lambda_{ii} (e_i^T U x)^2$, K is p.s.d. iff. $\Lambda_{ii} \geq 0$ for all i.

2. Since $K_{ij} = e_i^T K e_j$, $d_{ij} = (e_i - e_j)^T K (e_i - e_j)$. Let $\sqrt{\Lambda} = diag\left(\left(\sqrt{\Lambda_{ii}}\right)_i\right)$. Then, $d_{ij} = \|u_i - u_j\|^2$ where $u_i = \sqrt{\Lambda} U e_i$.

3. Let $y = H_\alpha^T x = x - \alpha e^T x$, then $e^T x = 0$.

$$\begin{aligned} y^T D y = & \sum_{i,j} y_i y_j (u_i - u_j)^T (u_i - u_j) \\ =& 2 \sum_i e^T y y_i u_i^T u_i - 2 \sum_{i,j} y_i y_j u_i^T u_j \\ =& -2 \left(\sum_i y_i u_i \right)^T \left(\sum_i y_i u_i \right) \leq 0 \end{aligned}$$

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

   $x^T (A \circ B) x = tr(A \operatorname{diag}(x) B \operatorname{diag}(x)) = tr(C C^T)$ where $C = \sqrt{A} \operatorname{diag}(x) \sqrt{B}$ which is p.s.d.. Thus $x^T (A \circ B) x \geq 0$.

4

1. No. Counter example: when d is Euclidean distance, let $d(x, y) = 1$, $d(x, z) = 2$, and $d(y, z) = 3$, d^2 violates the triangle inequality.

2. Yes. Identity and symmetry properties are trivial. Also note that $\sqrt{d(x, y)} \leq \sqrt{d(x, z) + d(z, y)} \leq \sqrt{d(x, z) + d(z, y) + \sqrt{4 d(x, z) d(z, y)}} = \sqrt{d(x, z)} + \sqrt{d(z, y)}$.

5

1. $A^T A$ admits orthogonal diagonalization $A^T A = V \sqrt{\Lambda}^2 V^T$. $\operatorname{rank}(\sqrt{\Lambda}) \leq \min\{m, n\}$. Thus, we may take $\Sigma = \operatorname{diag}\{\sqrt{\lambda_1}, \ldots, \sqrt{\lambda_1}\} \in \mathbb{R}^{m \times n}$ and $\Sigma^T \Sigma = \Lambda$. The column vectors of A and ΣV^T have the same Gram matrix. Thus, there is an orthogonal transformation that maps column vectors (in $\mathbb{R}^n$) of ΣV^T to column vectors (in $\mathbb{R}^m$) of A. Denote by U the matrix representation of the transformation. Then $A = U \Sigma V^T$.

   If $A = U \Sigma V^T = \hat{U} \hat{\Sigma} \hat{V}^T$, then $A^T A = V \Sigma^T \Sigma V^T = \hat{V} \hat{\Sigma}^T \hat{\Sigma} \hat{V}^T$. $\Sigma^T \Sigma$ and its tilde version are diagonal and they are similar, thus must be the same and hence $\Sigma = \tilde{\Sigma}$.

2. Since $\|x\| = 1, x \in \mathbb{R}^n$ iff. $\|V^T x\| = 1$, $\min_{rank(B)\leq k} \|A-B\| = \min_{rank(\tilde{B})\leq k} \|\Sigma - \tilde{B}\|$, where $\tilde{B} = U^T BV$ is $m \times n$. Since $\dim(\ker(\tilde{B})) + k \geq n$, there exists a unit vector $v \in \ker(B) \cap span\{e_1, \ldots, e_{k+1}\}$. Then $\|(\Sigma - B)v\| \geq \sigma_{k+1}$. On the other hand $\|\Sigma - \Sigma_k\| = \sigma_{k+1}$.

3. Note that $\|A\|_F^2 = tr(A^T A) = \sum_i \sigma_i^2$. Since $\text{rank}(B - (\Sigma - B)_{i-1}) \leq k+i-1$, $\sigma_i(\Sigma - B) = \sigma_1(\Sigma - B - (\Sigma - B)_{i-1}) = \|\Sigma - (B - (\Sigma - B)_{i-1})\| \geq \sigma_{i+k-1}$. Therefore, $\|A - B\|_F^2 = tr((\Sigma - \tilde{B})^T(\Sigma - \tilde{B})) = \sum_i \sigma_i(\Sigma - B)^2 \geq \sum_{i=k+1}^n \sigma_i^2 = \|A - A_k\|_F^2$.

4. Since $QAZ = QU\Sigma V^T Z$ is an SVD, QAZ and A have the same set of singular values.

5. $\|A - R\| = \|\Sigma - \tilde{R}\|_F$, $\tilde{R} = U^T RV$ also satisfies $\tilde{R}^T\tilde{R} = I$ and $\tilde{R}\tilde{R}^T = I$. $\|\Sigma - \tilde{R}\|_F^2 = tr(\Sigma^T\Sigma + I) - tr(\Sigma^T R + R^T\Sigma) = tr(\Sigma^T\Sigma + I) - 2\sum_i \sigma_i R_{ii} \geq tr(\Sigma^T\Sigma + I) - 2\sum_i \sigma_i$ as $R_{ii} \leq 1$. On the other hand $\|\Sigma - I\|_F = tr(\Sigma^T\Sigma + I) - 2\sum_i \sigma_i$.