Homework 8

Graded

Student

Jaiden Rain Atterbury

Total Points

35.5 / 40 pts

Question 1

Question 1

9.25 / 10 pts

✔ **– 0.75 pts** (c) only Q or Q^T (but not both) should be modified along with Λ (or else we no longer have equality $C = Q\Lambda Q^T$)

Question 2

Question 2

9.5 / 10 pts

✔ **– 0.5 pts** (a) minor error

1. dim
2. span
3. col

Question 3

Question 3

7.75 / 10 pts

✔ **– 2 pts** (a)

✔ **– 0.25 pts** (b) the zero matrix has rank 0

Question 4

Question 4

9 / 10 pts

✔ **– 1 pt** Insufficient explanation/incorrect solution on (a)(i)

Question assigned to the following page: 1

Jaiden Atterbury 05-21-2022 Homework 8

(1) SVD of Symmetric and PSD matrices

(a) $B = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \\ 3 & 5 & 6 \end{bmatrix}$ $B = U \Sigma V^T$

$$B = \underbrace{\begin{bmatrix} -0.327985 & -0.736976 & -0.591009 \\ -0.591009 & -0.327985 & 0.736976 \\ -0.736976 & 0.591009 & -0.327985 \end{bmatrix}}_{U} \underbrace{\begin{bmatrix} 11.344814 & 0 & 0 \\ 0 & 0.515729 & 0 \\ 0 & 0 & 0.170915 \end{bmatrix}}_{\Sigma} \underbrace{\begin{bmatrix} -0.327985 & -0.591009 & -0.736976 \\ 0.736976 & 0.327985 & -0.591009 \\ -0.591009 & 0.736976 & -0.327985 \end{bmatrix}}_{V^T}$$

(b) A is $n \times n$ symmetric:

By lemma 8.1.4/8.1.5 to get the singular values of A we must take the positive square roots of the r positive eigenvalues of A^TA or AA^T. Since a symmetric matrix satisfies the property $A = A^T$ (definition 5.1.1) it follows that:

$$AA^T \text{ and } A^TA = A^2$$

given that $A\vec{x} = \lambda\vec{x}$ then $A^2\vec{x} = \lambda^2\vec{x}$ (Homework 1 and Lecture notes 3)

thus the eigenvalues of A^2 are λ^2, furthermore the singular values of A are the positive square roots of λ^2

$\therefore \sigma_i = \sqrt{\lambda_i^2} = |\lambda_i|$

(c) C is general $n \times n$ symmetric matrix;

Since C is symmetric, by the spectral theorem it can be diagonalized by $Q\Lambda Q^T$ such that Q is orthogonal and made of the eigenvectors of C, Since the SVD of C has the eigenvectors of A^2 in the matrices U and V^T, the only difference between the SVD and the diagonalization of C is that the singular values are the magnitudes of the eigenvalues of C, while the eigenvalues in the

Question assigned to the following page: 1

diagonal matrix Λ can be negative. With this said, in order to get from the diagonalization of C to the SVD of C, any column of Q corresponding to a negative eigenvalue in Λ needs to be multiplied by -1. Once you take Q^T, this negative sign will "carry over," and when multiplied together will return the same C and hence $Q = U$ so the SVD will then equal the orthogonal diagonalization. It is important to note that if $Q\Lambda Q^T$ is already expanded, then you will need to multiply the column in Q AND the row in Q^T corresponding to the negative eigenvalue, in order to get the SVD to equal the orthogonal diagonalization.

(d) A is PSD $n \times n$ matrix:

If A is a PSD matrix, then by definition 6.1.9, all eigenvalues of A are non-negative. Since A is also symmetric, by part (b) $\sigma_i = |\lambda_i|$, since the eigenvalues are non-negative it turns out $\sigma_i = \lambda_i$ if A is PSD.

Since, by the spectral theorem, we can choose orthonormal eigenvectors for the diagonalization of A it turns out that $Q = U = V$ (Q is from the orthogonal diagonalization, while U and V are from the SVD. Lastly, since $\sigma_i = \lambda_i$, if A is PSD/PD the SVD of A will coincide/equal the orthogonal diagonalization of A.

SVD of PSD matrix A: $A = U\Sigma U^T$ (where $U = V$)

Question assigned to the following page: 2

(2) Rank one matrices

(a) Prove that for any A and B, $\text{rank}(A+B) \leq \text{rank}(A) + \text{rank}(B)$

If $A = [a_1 \cdots a_n] \rightarrow \text{rank}(A) = \text{Col}(A) = \dim(\text{span}\{a_1 \cdots a_n\})$

$B = [b_1 \cdots b_n] \rightarrow \text{rank}(B) = \text{Col}(B) = \dim(\text{span}\{b_1 \cdots b_n\})$

$A+B = [a_1+b_1 \cdots a_n+b_n] \rightarrow \text{rank}(A+B) = \text{col}(A+B) = \dim(\text{span}\{a_1+b_1 \cdots a_n+b_n\}$

but since $\{a_1+b_1, \cdots, a_n+b_n\} \subseteq \{a_1, \cdots a_n, b_1, \cdots b_n\}$

Proof: if $x \in A+B$

then $x = c_1(\vec{a}_1+\vec{b}_1) + \cdots + c_n(\vec{a}_n+\vec{b}_n)$

$= c_1 a_1 + \cdots + c_n a_n + c_1 b_1 + \cdots + c_n b_n$

$\therefore x \in A \cup B$

So $\dim\{a_1+b_1, \cdots, a_n+b_n\} \leq \dim\{a_1 \cdots a_n, b_1 \cdots b_n\} = \dim\{A \cup B\}$

Proof: By example 4.2.9 of the 20B notes, a subspace can only have a dimension less than or equal to its original space.

Using the fact that $\dim(\text{span}\{S \cup T\}) \leq \dim(\text{span}\{S\}) + \dim(\text{span}\{T\})$ taking $S = \text{col}(A)$ and $T = \text{col}(B)$, and using the fact that rank of a matrix is the dimension of the column space of that matrix, we have shown that:

$\text{rank}(A+B) \leq \text{rank}(A) + \text{rank}(B)$

(b) $A \in \mathbb{R}^{m \times n}$ is rank 1 with SVD of $A = U \Sigma V^T$, SVD of A:

$A = [u_r \ u_{r+1} \cdots u_m] \begin{bmatrix} \sigma_r & & \\ & 0 & \\ & & \ddots & \\ & & & 0 \end{bmatrix} \begin{bmatrix} v_r^T \\ v_{r+1}^T \\ \vdots \\ v_n^T \end{bmatrix}$ which multiplies out to:

$A = \sigma_r u_r v_r^T$ where $u_r \in \mathbb{R}^m$ and $v_r \in \mathbb{R}^n$

Question assigned to the following page: 2

$\therefore$ Every rank one matrix is of the form uv^T (since $\sigma_r u_r v_r^T$ is just a scaling of the rank 1 matrix $u_r v_r^T$).

(c) (i) rank 1 + rank 1 = rank 1

$A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ $B = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}$

$A+B = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 3 & 0 \\ 0 & 0 \end{bmatrix} =$ rank 1

(ii) rank 1 + rank 1 = rank 2

$A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ $B = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$

$A+B = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} =$ rank 2

(d) $A \in \mathbb{R}^{m \times k}$ with columns $a_1, \dots, a_k$

$B \in \mathbb{R}^{k \times n}$ with rows $b_1^T, \dots, b_k^T$

Show that $AB = a_1 b_1^T + a_2 b_2^T + \dots + a_k b_k^T$

$$A = \begin{array}{cc} & k \\ m & \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & a_{22} & \cdots & a_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mk} \end{bmatrix} \\ & \begin{matrix} a_1 & a_2 & \cdots & a_k \end{matrix} \end{array} \qquad B = \begin{array}{ccc} & n & \\ k & \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{k1} & b_{k2} & \cdots & b_{kn} \end{bmatrix} & \begin{matrix} b_1^T \\ b_2^T \\ \vdots \\ b_k^T \end{matrix} \end{array}$$

$$AB = \begin{bmatrix} a_{11}b_{11} + a_{12}b_{21} + \dots + a_{1k}b_{k1} & a_{11}b_{12} + a_{12}b_{22} + \dots + a_{1k}b_{k2} & & a_{11}b_{1n} + a_{12}b_{2n} + \dots + a_{1k}b_{kn} \\ a_{21}b_{11} + a_{22}b_{21} + \dots + a_{2k}b_{k1} & a_{21}b_{12} + a_{22}b_{22} + \dots + a_{2k}b_{k2} & \cdots & a_{21}b_{1n} + a_{22}b_{2n} + \dots + a_{2k}b_{kn} \\ \vdots & \vdots & & \vdots \\ a_{m1}b_{11} + a_{m2}b_{21} + \dots + a_{mk}b_{k1} & a_{m1}b_{12} + a_{m2}b_{22} + \dots + a_{mk}b_{k2} & & a_{m1}b_{1n} + a_{m2}b_{2n} + \dots + a_{mk}b_{kn} \end{bmatrix}$$

right side = left side

$$AB = \underbrace{\begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix}}_{a_1} \underbrace{\begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \end{bmatrix}}_{b_1^T} + \underbrace{\begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix}}_{a_2} \underbrace{\begin{bmatrix} b_{21} & b_{22} & \cdots & b_{2n} \end{bmatrix}}_{b_2^T} + \dots + \underbrace{\begin{bmatrix} a_{1k} \\ a_{2k} \\ \vdots \\ a_{mk} \end{bmatrix}}_{a_k} \underbrace{\begin{bmatrix} b_{k1} & b_{k2} & \cdots & b_{kn} \end{bmatrix}}_{b_k^T}$$

$\therefore AB = a_1 b_1^T + a_2 b_2^T + \dots + a_k b_k^T$

Question assigned to the following page: 3

(3) Rank one decomposition of symmetric and PSD matrices

(a) By definition 6.1.4 if $A \in \mathbb{R}^{n\times n}$ is PSD, then A can be factored by a matrix $B \in \mathbb{R}^{k\times n}$ for some k such that $A = B^T B$. If we let $b \in \mathbb{R}^n$ be B^T and $b^T \in \mathbb{R}^m$ be B then A will be factorized by $A = bb^T$ where b is $n \times k$ and b^T is $k \times n$. since bb^T is $n\times n$, and rank one since all columns are scalar multiples of b, $A = bb^T$ is rank one and PSD.

Also, $\vec{x}^T bb^T \vec{x} = (b^T\vec{x})^T \cdot (b^T x) = \|b^T \vec{x}\|^2 \geq 0$

$\therefore$ bb^T is PSD and rank one

(b) Describe 3x3 matrices of rank 1 that have 0s and 1s on the diagonal:

rank 1 PSD matrices $\rightarrow bb^T$:

$$\begin{bmatrix} a \\ b \\ c \end{bmatrix} \begin{bmatrix} a & b & c \end{bmatrix} = \begin{bmatrix} a^2 & ba & ca \\ ba & b^2 & cb \\ ca & cb & c^2 \end{bmatrix}$$

in order to get only 1's and 0's on the diagonal, all entries a,b,c have to be either -1, 0, or 1, since the diagonal entries are a^2, b^2, c^2.

By the fundamental counting principal / multiplication rule since there are 3 choices (1, 0, -1) for 3 components (a, b, c) there are $3 \cdot 3 \cdot 3 = 3^3 = 27$ different matrices that satisfy this property.

(c) A is PSD and rank r:

The SVD of A will be the same as its orthogonal diagonalization.

Writing out the SVD of A:

$A = V\Sigma V^T$, which can be decomposed of rank 1 matrices:

$= \sigma_1 v_1 v_1^T + \dots + \sigma_r v_r v_r^T$

Question assigned to the following page: 3

since the matrices $v_r v_r^T$ are of the form bb^T, by part (a) we can conclude they are PSD, and since the number of singular values equals the rank of A, we can write A as the sum of r rank one PSD matrices.

(d) If a symmetric matrix $A \in \mathbb{R}^{n \times n}$ of rank r is not PSD (just a general symmetric matrix), then it can't be written as a sum of rank 1 PSD matrices as the sum of PSD matrices is also PSD:

$$x^T(M_1 + M_2 + \cdots + M_r)x$$

$$x^T M_1 x + x^T M x + \cdots + x^T M_r x \geq 0$$

↳ PSD and rank r

BUT, A is not PSD. Therefore, A cannot be decomposed into the sum or r rank 1 PSD matrices.

Question assigned to the following page: 4

(4) Projection with an orthonormal basis

(a) $q_1, \cdots, q_K$ is an orthonormal basis of V, $Q \in \mathbb{R}^{n \times k}$

(i) $P = Q(Q^TQ)^{-1}Q^T \rightarrow Q^TQ = I_n$

$P = QQ^T$

$Q = [q_1 \; q_2 \cdots q_K]$

$P = [q_1 \; q_2 \cdots q_K][q_1 \; q_2 \cdots q_K]^T = [q_1 \; q_2 \cdots q_K]\begin{bmatrix} q_1^T \\ q_2^T \\ \vdots \\ q_K^T \end{bmatrix} = q_1q_1^T + q_2q_2^T + \cdots + q_Kq_K^T$

(ii) $\text{Proj}_V b = Pb = QQ^Tb$

$= (q_1q_1^T + q_2q_2^T + \cdots + q_Kq_K^T)b$

$= q_1(q_1^Tb) + q_2(q_2^Tb) + \cdots + q_K(q_K^Tb)$

(iii) $[\text{Proj}_V b]_Q = \begin{bmatrix} q_1^Tb \\ q_2^Tb \\ \vdots \\ q_K^Tb \end{bmatrix}$

(iv) $I-P$ projects onto orthogonal subspace of P (Strang Page 209 P2)

Since P projects onto V

$I-P$ projects onto $V^\perp$

$\therefore$ matrix that projects onto $V^\perp$ is: $I - QQ^T$

(v) $\text{Proj}_{V^\perp} b = (I - QQ^T)b$

$= b - QQ^Tb$

$= b - q_1(q_1^Tb) - q_2(q_2^Tb) - \cdots - q_K(q_K^Tb)$

(b) $q_1, \cdots, q_K, q_{K+1}, \cdots, q_n$ is an orthonormal basis of $\mathbb{R}^n$

(i) $P = QQ^T$

$P = q_{K+1}q_{K+1}^T + q_{K+2}q_{K+2}^T + \cdots + q_nq_n^T$

$\therefore \text{Proj}_{V^\perp} b = Pb = QQ^Tb = q_{K+1}(q_{K+1}^Tb) + q_{K+2}(q_{K+2}^Tb) + \cdots + q_n(q_n^Tb)$

Question assigned to the following page: 4

(ii) $b - q_1(q_1^T b) - q_2(q_2^T b) - \cdots - q_k(q_k^T b) = q_{k+1}(q_{k+1}^T b) + q_{k+2}(q_{k+2}^T b) + \cdots + q_n(q_n^T b)$

$\therefore b = q_1(q_1^T b) + q_2(q_2^T b) + \cdots + q_k(q_k^T b) + q_{k+1}(q_{k+1}^T b) + q_{k+2}(q_{k+2}^T b) + \cdots + q_n(q_n^T b)$

(iii) $[b]_Q = \begin{bmatrix} q_1^T b \\ q_2^T b \\ \vdots \\ q_n^T b \end{bmatrix}$

(c) $L = \text{span}([1, 1, 1]^T)$

(i) $\text{proj}_L x \to x \longmapsto uu^T x$

by proposition 4.3.6 if $V = \text{span}\{u\}$ then $\text{proj}_V x$ is $P = \frac{uu^T}{u^T u}$ that sends $x \mapsto \frac{uu^T}{u^T u} x$

$\therefore$ we need u to be a unit vector ($\|u\| = 1$).

$\left\|\begin{bmatrix}1\\1\\1\end{bmatrix}\right\| = \sqrt{3}$ $\quad \frac{\begin{bmatrix}1\\1\\1\end{bmatrix}}{\left\|\begin{bmatrix}1\\1\\1\end{bmatrix}\right\|} = \begin{bmatrix}\sqrt{3}/3\\ \sqrt{3}/3\\ \sqrt{3}/3\end{bmatrix} = \vec{u}$

$\therefore \text{Proj}_L x = \begin{bmatrix}\sqrt{3}/3\\ \sqrt{3}/3\\ \sqrt{3}/3\end{bmatrix} \begin{bmatrix}\sqrt{3}/3 & \sqrt{3}/3 & \sqrt{3}/3\end{bmatrix} \vec{x}$

(ii) $\text{Proj}_L b = \underbrace{\begin{bmatrix} 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \end{bmatrix}}_{uu^T} \underbrace{\begin{bmatrix}2\\3\\4\end{bmatrix}}_{b} = \underbrace{\begin{bmatrix}3\\3\\3\end{bmatrix}}_{\text{proj}_L b}$

$\text{Proj}_{L^\perp} b = I - P = I - uu^T = \left(\underbrace{\begin{bmatrix}1&0&0\\0&1&0\\0&0&1\end{bmatrix}}_{I} - \underbrace{\begin{bmatrix}1/3&1/3&1/3\\1/3&1/3&1/3\\1/3&1/3&1/3\end{bmatrix}}_{uu^T}\right) \underbrace{\begin{bmatrix}2\\3\\4\end{bmatrix}}_{b}$

$= \underbrace{\begin{bmatrix}2/3 & -1/3 & -1/3\\ -1/3 & 2/3 & -1/3\\ -1/3 & -1/3 & 2/3\end{bmatrix}}_{I - uu^T} \underbrace{\begin{bmatrix}2\\3\\4\end{bmatrix}}_{b} = \underbrace{\begin{bmatrix}-1\\0\\1\end{bmatrix}}_{\text{proj}_{L^\perp} b}$