Math 218D Problem Session: Week 14

Answer Key

1. Rules of vector SVD

Which of the following $A = \sigma_1 u_1 v_1^T + \cdots + \sigma_r u_r v_r^T$ are valid singular value decompositions? Why/why not?

a) $A = 1(1,0)(1,0)^T + 3(0,1)(0,1)^T$

b) $A = 4(1,0)(0,1)^T + 3(0,1)(1,0)^T$

c) $A = 3(1,-1)(1,0)^T + 2(1,1)(0,1)^T$

d) $A = -3(1/\sqrt{2}, -1/\sqrt{2}, 0)(1,0)^T + 2(0,0,1)(0,1)^T$

e) $A = 3(-1/\sqrt{2}, 1/\sqrt{2}, 0)(1,0)^T + 2(0,0,1)(0,1)^T$

f) $A = 5(1,0,0)(0,1)^T + 3(0,1,0)(1,0)^T + 2(0,0,1)(0,1)^T$

Solution.

a) $A = 1(1,0)(1,0)^T + 3(0,1)(0,1)^T$ is not an SVD since $1 < 3$, but singular values must be in decreasing order.

b) $A = 4(1,0)(0,1)^T + 3(0,1)(1,0)^T$ is an SVD.

c) $A = 3(1,-1)(1,0)^T + 2(1,1)(0,1)^T$ is not an SVD, since $(1,-1)$ and $(1,1)$ are not unit vectors.

d) $A = -3(1/\sqrt{2}, -1/\sqrt{2}, 0)(1,0)^T + 2(0,0,1)(0,1)^T$ is not an SVD since $-3 < 0$, but singular values must be positive.

e) $A = 3(-1/\sqrt{2}, 1/\sqrt{2}, 0)(1,0)^T + 2(0,0,1)(0,1)^T$ is an SVD.

f) $A = 5(1,0,0)(0,1)^T + 3(0,1,0)(1,0)^T + 2(0,0,1)(0,1)^T$ is not an SVD, since the vectors $(0,1), (1,0), (0,1)$ are not orthogonal.

2. **The matrix SVD** Suppose that A is an $m \times n$ matrix of rank r, with SVD $A = U\Sigma V^T$.

a) U is a ___ $\times$ ___ matrix, Σ is a ___ $\times$ ___ matrix, and V is a ___ $\times$ ___ matrix. The matrices U and V are _______ matrices. The first ___ diagonal entries of Σ are > 0.

b) Expand A^TA using $A = U\Sigma V^T$ to see that the matrix A^TA has symmetric diagonalization $Q_1 D_1 Q_1^T$, with $Q_1 =$ ___ and $D_1 =$ ___. What are the eigenvectors of A^TA? What are the eigenvalues?

c) Expand AA^T using $A = U\Sigma V^T$ to see that the matrix AA^T has symmetric diagonalization $Q_2 D_2 Q_2^T$, with $Q_2 =$ ___ and $D_2 =$ ___. What are the eigenvectors of AA^T? What are the eigenvalues?

d) Suppose that $i \le r$. The left singular vector u_i is the ith column of U, the singular value σ_i is the ith diagonal entry of Σ, and the right singular vector v_i is the ith column of V. Explain why $Av_i = \sigma_i u_i$ by computing $V^T v_i, \Sigma V^T v_i$, and $U\Sigma V^T v_i$.

Solution.

a) U is an $m \times m$ matrix, Σ is a $m \times n$ matrix, and V is a $n \times n$ matrix. The matrices U and V are orthogonal matrices. The first r diagonal entries of Σ are > 0.

b) $A^TA = (U\Sigma V^T)^T U\Sigma V^T = V\Sigma^T U^T U\Sigma V^T = V(\Sigma^T\Sigma)V^T$. Therefore $Q_1 = V$ and $D_1 = \Sigma^T\Sigma$. The columns of V are eigenvectors of A^TA, and the eigenvalues are the diagonal entries of the $n \times n$ matrix $\Sigma^T\Sigma$, which are $\sigma_1^2, \ldots, \sigma_r^2, 0, \ldots, 0$.

c) $AA^T = U\Sigma V^T (U\Sigma V^T)^T = U\Sigma V^T V\Sigma^T U^T = U(\Sigma\Sigma^T)U^T$. Therefore $Q_2 = U$ and $D_2 = \Sigma\Sigma^T$. The columns of U are eigenvectors of AA^T, and the eigenvalues are the diagonal entries of the $n \times n$ matrix $\Sigma\Sigma^T$, which are $\sigma_1^2, \ldots, \sigma_r^2, 0, \ldots, 0$.

d) $V^T v_i = (v_1 \cdot v_i, \ldots, v_i \cdot v_i, \ldots, v_n \cdot v_i) = (0, \ldots, 1, \ldots, 0)$, $\Sigma V^T v_i = \Sigma(0, \ldots, 1, \ldots, 0) = (0, \ldots, \sigma_i, \ldots, 0)$, $Av_i = U\Sigma V^T v_i = U(0, \ldots, \sigma_i, \ldots, 0) = \sigma_i U e_i = \sigma_i u_i$.

3. Computing the vector SVD

To

(1) Find the non-zero eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \lambda_r > 0$ of A^TA.
(2) Find a orthonormal basis of each of the λ_i eigenspace of A^TA. Listed in order of decreasing eigenvalue, these are the right singular vectors $v_1, \cdots, v_r$.
(3) For $i = 1, \ldots, r$, set $\sigma_i = \sqrt{\lambda_i}$ and $u_i = \frac{Av_i}{\sigma_i}$. These are the singular values and left singular vectors.
(4) Write $A = \sigma_1 u_1 v_1^T + \cdots + \sigma_r u_r v_r^T$.

Compute the vector SVD of each of the following matrices:

a) $A = \begin{pmatrix} 0 & -1 \\ 3 & 0 \end{pmatrix}$

b) $A = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & -3 & 0 & 0 \end{pmatrix}$

c) $A = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}$

Solution.

a) $A = \begin{pmatrix} 0 & -1 \\ 3 & 0 \end{pmatrix}$. The matrix $A^TA = \begin{pmatrix} 9 & 0 \\ 0 & 1 \end{pmatrix}$ has eigenvalues $\lambda_1 = 9, \lambda_2 = 1$, with eigenvectors $v_1 = (1,0)$ and $v_2 = (0,1)$. The singular values are $\sigma_1 = \sqrt{\lambda_1} = 3$ and $\sigma_2 = \sqrt{\lambda_2} = 1$. The left singular vectors are $u_1 = \frac{1}{3}Av_1 = \frac{1}{3}(0,3) = (0,1)$ and $u_2 = \frac{Av_2}{1} = (-1,0)$. The vector SVD is

$$A = 3(0,1)(1,0)^T + 1(-1,0)(0,1)^T.$$

b) $A = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & -3 & 0 & 0 \end{pmatrix}$. The matrix $A^TA = \begin{pmatrix} 4 & 0 & 0 & 0 \\ 0 & 9 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$ has eigenvalues $\lambda_1 = 9, \lambda_2 = 4, \lambda_3 = 1, \lambda_4 = 0$. The orthonormal eigenvectors for the non-zero eigenvalues are $v_1 = (0,1,0,0)$, $v_2 = (1,0,0,0)$, and $v_3 = (0,0,0,1)$. The singular values are $\sigma_1 = 3, \sigma_2 = 2, \sigma_3 = 1$. The left singular vectors are $u_1 = \frac{1}{3}Av_1 = (0,0,-1), u_2 = \frac{1}{2}Av_2 = (1,0,0)$, $u_3 = \frac{1}{1}Av_3 = (0,1,0)$. The vector SVD is

$$A = 3(0,0,-1)(0,1,0,0)^T + 2(1,0,0)(1,0,0,0)^T + 1(0,1,0)(0,0,0,1)^T.$$

c) $A = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}$. The matrix $A^TA = \begin{pmatrix} 4 & 2 \\ 2 & 5 \end{pmatrix}$ has characteristic polynomial $\lambda^2 - 9\lambda + 16$, with eigenvalues $\lambda_1 = \frac{9+\sqrt{17}}{2}, \lambda_2 = \frac{9-\sqrt{17}}{2}$. We then have eigenvectors $v_1 = \frac{(-2,4-\lambda_1)}{||(-2,4-\lambda_1)||} \approx (-0.615, -0.788)$, and $v_2 = \frac{(-2,4-\lambda_2)}{||(-2,4-\lambda_2)||} \approx (-0.788, 0.615)$.

The singular values are $\sigma_1 = \sqrt{\lambda_1} \approx 2.562$ and $\sigma_2 = \sqrt{\lambda_2} \approx 1.562$.

The left singular vectors are $u_1 = \frac{Av_1}{\sigma_1} \approx (-0.788, -0.615)$ and $u_2 = \frac{Av_2}{\sigma_2} \approx (-0.615, 0.788)$.

The vector SVD is, approximately,

$$A = 2.562(-0.788, -0.615)(-0.615, -0.788)^T + 1.562(-0.615, 0.788)(-0.788, 0.615)^T.$$

(The fact that the u and v vectors look so similar seems to be a coincidence.)

4. Computing the matrix SVD

To find the matrix SVD $A = U\Sigma V^T$ of a matrix A:

(1) Find the symmetric diagonalization VDV^T of A^TA, where the eigenvalues are listed in decreasing order: $\lambda_1 \geq \ldots \geq \lambda_n$. The rank r of A is the same as the number of non-zero eigenvalues of A^TA (counted with multiplicity).

(2) The columns of V are the right singular vectors $v_1, \ldots, v_r$, followed by an orthonormal basis $v_{r+1}, \ldots, v_n$ of Nul(A).

(3) For $i = 1, \ldots, r$, set $\sigma_i = \sqrt{\lambda_i}$ and $u_i = \frac{Av_i}{\sigma_i}$. These are the singular values and left singular vectors.

(4) We still need the vectors $u_{r+1}, \ldots, u_m$: find these by computing an orthonormal basis of Nul(A^T) (using RREF to find a basis, Gram–Schmidt to replace it with an orthonormal basis).

(5) Finally, the matrix U is the matrix with columns $u_1, \ldots, u_m$, the matrix V was found in (1), and Σ has its first r diagonal entries as $\sigma_1, \ldots, \sigma_r$ and the remaining entries of Σ being zero.

Compute the matrix SVD of each of the following matrices:

a) $A = \begin{pmatrix} 0 & -1 \\ 3 & 0 \end{pmatrix}$

b) $A = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & -3 & 0 & 0 \end{pmatrix}$

c) $A = \begin{pmatrix} 1 & 2 \\ 1 & 2 \\ 2 & 4 \end{pmatrix}$

Solution.

a) $A = \begin{pmatrix} 0 & -1 \\ 3 & 0 \end{pmatrix}$. Since $m = n = r = 2$, we can just use the singular vectors and values found in problem 3a) (no need to find ONB for Nul(A) or Nul(A^T).) We have

$$A = U\Sigma V^T = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}^T.$$

b) $A = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & -3 & 0 & 0 \end{pmatrix}$. We found the vector SVD of this in 3b). Since $m = 3, n = 4, r = 3$, we need to find the additional vector v_4, an ONB of Nul(A). Since the matrix A^TA was rather simple, and Nul(A^TA) = Nul(A), we can use that A^TA had unit eigenvector $v_4 = (0,0,1,0)$ for the eigenvalue $\lambda_4 = 0$. Then

$$A = U\Sigma V^T = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 3 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}^T.$$

c) $A = \begin{pmatrix} 1 & 2 \\ 1 & 2 \\ 2 & 4 \end{pmatrix}$. The matrix $A^T A = \begin{pmatrix} 6 & 12 \\ 12 & 24 \end{pmatrix}$ has characteristic polynomial $\lambda^2 - 30\lambda$, with eigenvalues $\lambda_1 = 30$ and $\lambda_2 = 0$. The λ_1-eigenvector equals $v_1 = \frac{(12,24)}{||(12,24)||} = \frac{(1,2)}{||(1,2)||} = \frac{1}{\sqrt{5}}(1,2)$, while the λ_2-eigenvector equals $v_2 = \frac{(12,-6)}{||(12,-6)||} = \frac{1}{\sqrt{5}}(2,-1)$.

The only singular value is $\sigma_1 = \sqrt{\lambda_1} = \sqrt{30}$. We find the left singular vector $u_1 = \frac{1}{\sqrt{30}} A v_1 = \frac{1}{5\sqrt{6}}(5,5,10) = \frac{1}{\sqrt{6}}(1,1,2)$.

We already found the vector v_2 spanning $\mathrm{Nul}(A^T A) = \mathrm{Nul}(A)$. It remains to find an ONB u_2, u_3 of $\mathrm{Nul}(A^T)$. It is not hard to see that $(1,-1,0)$ and $(2,0,-1)$ are a basis of $\mathrm{Nul}(A^T)$, but they are not orthonormal.

Doing Gram-Schmidt, we first replace $(1,-1,0)$ with $(1,-1,0)$ and replace $(2,0,-1)$ with $(2,0,-1) - \frac{(2,0,-1)\cdot(1,-1,0)}{(1,-1,0)\cdot(1,-1,0)}(1,-1,0) = (2,0,-1) - (1,-1,0) = (1,1,-1)$.

(I avoided using the usual names for vectors in Gram-Schmidt, since it would be easy to confuse with the u and v vectors of SVD, which have a totally different meaning).

Making these unit vectors, we find that $u_2 = \frac{1}{\sqrt{2}}(1,-1,0)$ and $u_3 = \frac{1}{\sqrt{3}}(1,1,-1)$ form an ONB of $\mathrm{Nul}(A^T)$.

We conclude that

$$U = \begin{pmatrix} 1/\sqrt{6} & 1/\sqrt{2} & 1/\sqrt{3} \\ 1/\sqrt{6} & -1/\sqrt{2} & 1/\sqrt{3} \\ 2/\sqrt{6} & 0 & -1/\sqrt{3} \end{pmatrix}, V = \begin{pmatrix} 1/\sqrt{5} & 2/\sqrt{5} \\ 2/\sqrt{5} & -1/\sqrt{5} \end{pmatrix}, \Sigma = \begin{pmatrix} \sqrt{30} & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

Warning: Many other answers are possible for U and V. Your columns of V might be off by a sign, your first column of U might be off by a sign, and the final two columns of U can look quite different.

5. Sums of rank 1 matrices

This final problem is not about SVDs, but just about sums of rank one matrices.

a) Without computing A, explain why

$$A = (1,2,1)(1,1)^T + (1,-1,1)(-1,1)^T$$

is a rank 2 matrix.

Hint: compute $A(1,1)$ and $A(-1,1)$, and use this to show that $(1,2,1)$ and $(1,-1,1)$ are in the column space of A.

b) If $A = u_1 v_1^T + \cdots + u_r v_r^T$ for some vectors $u_i \in \mathbf{R}^m$ and $v_j \in \mathbf{R}^n$, explain why the rank of A is at most r.

Hint: Show that the subspace Col(A) is contained in the span Span$\{u_1, \ldots, u_r\}$, which is at most r-dimensional.

c) If the vectors $u_1, \ldots, u_r \in \mathbf{R}^m$ are a linearly independent set of vectors, and the vectors $v_1, \ldots, v_r \in \mathbf{R}^n$ are also linearly independent, prove that

$$A = u_1 v_1^T + \cdots + u_r v_r^T$$

has rank equal to r.

Hint: Show that there is a vector $v \in \mathbf{R}^n$ which is orthogonal to $v_2, \ldots, v_r$, but $v_1^T v \neq 0$. Compute Av, and use this to show that $u_1 \in$ Col(A). The same idea shows that $u_2, \ldots, u_r$ are all in Col(A).

Solution.

a) Without computing A, we will explain why

$$A = (1,2,1)(1,1)^T + (1,-1,1)(-1,1)^T$$

is a rank 2 matrix.

Since $A(1,1) = (1,2,1)((1,1)\cdot(1,1)) + (1,-1,1)((-1,1)\cdot(1,1)) = 2(1,2,1) + 0$, the vector $2(1,2,1)$ is in the column space of A. Similarly, $A(-1,1) = (1,2,1)((-1,1)\cdot(1,1)) + (1,-1,1)((-1,1)\cdot(-1,1)) = 0 + 2(1,-1,1)$, so $2(1,-1,1)$ is also in the column space of A since these two column space vectors are linearly independent, the rank of A is at least 2. Since A is a 3×2 matrix, its rank is at most 2. Therefore the rank of A equals 2.

b) If $A = u_1 v_1^T + \cdots + u_r v_r^T$ for some vectors $u_i \in \mathbf{R}^m$ and $v_j \in \mathbf{R}^n$, we will explain why the rank of A is at most r.

For any vector x, $Ax = (v_1 \cdot x)u_1 + \cdots + (v_r \cdot x)u_r$. Therefore any vector b for which $Ax = b$ is consistent must be a linear combination of $u_1, \ldots, u_r$. In other words, $\mathrm{Col}(A) \subset \mathrm{Span}\{u_1, \ldots, u_r\}$. Since $\dim \mathrm{Span}\{u_1, \ldots, u_r\} \leq r$, and $\dim \mathrm{Col}(A) \leq \dim \mathrm{Span}\{u_1, \ldots, u_r\}$, we conclude that $\mathrm{rank}(A) = \dim \mathrm{Col}(A) \leq r$.

c) Suppose that the vectors $u_1, \ldots, u_r \in \mathbf{R}^m$ are a linearly independent set of vectors, and the vectors $v_1, \ldots, v_r \in \mathbf{R}^n$ are also linearly independent. We consider the matrix $A = u_1 v_1^T + \cdots + u_r v_r^T$.

Since the v_i vectors are linearly independent, $\mathrm{Span}\{v_1, \ldots, v_r\}$ is r-dimensional, while $\mathrm{Span}\{v_2, \ldots, v_r\}$ is $(r-1)$-dimensional. Using Gram-Schmidt, we can find a vector $v \in \text{Span}\{v_1, \ldots, v_r\}$ which is orthogonal to $\text{Span}\{v_2, \ldots, v_r\}$ but not orthogonal to v_1 (i.e. we can project v_1 onto the orthogonal complement of $\text{Span}\{v_2, \ldots, v_r\}$).

Using this vector v, $Av = (v_1 \cdot v)u_1 + \cdots + (v_r \cdot v)u_r = (v_1 \cdot v)u_1$, since $v \cdot v_2 = 0, v \cdot v_3 = 0, \ldots$. In other words, $A\left(\frac{v}{v_1 \cdot v}\right) = u_1$, which verifies that u_1 is in Col(A).

A similar argument shows that each of the vectors u_i is in Col(A). Therefore $\text{Span}\{u_1, \ldots, u_r\} \subset \text{Col}(A)$. Since the u_i vectors are linearly independent, $r = \text{Span}\{u_1, \ldots, u_r\} \leq \dim \text{Col}(A) = \text{rank}(A)$. On the other hand, by **b)**, $\text{rank}(A) \leq r$. Therefore $\text{rank}(A) = r$.

6. Principal component analysis Consider the following data points

$$\begin{pmatrix}1\\2\\3\end{pmatrix}, \begin{pmatrix}3\\5\\7\end{pmatrix}, \begin{pmatrix}4\\2\\3\end{pmatrix}, \begin{pmatrix}-1\\0\\-1\end{pmatrix}, \begin{pmatrix}3\\1\\2\end{pmatrix}$$

a) Calculate the mean value vector $\overline{\mu}$ then recenter the data points and put them in the matrix form A.

b) Calculate the covariance matrix by the formula $S = \frac{1}{n-1} A \cdot A^T$. What does the diagonal of this matrix tells you?

c) Using a calculator, find the eigenvalus of S and its corresponding eigenvector.

d) Find the principal component and the variance along that direction. What is the geometric meaning of the principal component?

Solution.

a) $\overline{\mu} = \begin{pmatrix}2\\2\\2\end{pmatrix}, \quad A = \begin{pmatrix}-1 & 1 & 2 & -3 & 1\\ 0 & 3 & 0 & -2 & -1\\ 1 & 1 & 1 & -3 & 0\end{pmatrix}$

b) $S = \begin{pmatrix}4 & 2 & 11/4\\ 2 & 7/2 & 9/4\\ 11/4 & 9/4 & 3\end{pmatrix}$. The diagonals are variances for each row.

c) $\begin{pmatrix}\lambda_1\\ \lambda_2\\ \lambda_3\end{pmatrix} = \begin{pmatrix}0.5620\\ 1.7439\\ 8.1941\end{pmatrix}, v_1 = \begin{pmatrix}0.4825\\ 0.3014\\ -0.8224\end{pmatrix}, v_2 = \begin{pmatrix}0.6122\\ -0.7876\\ 0.0706\end{pmatrix}, v_3 = \begin{pmatrix}0.6264\\ 0.5375\\ 0.5645\end{pmatrix}$