

Homework 5 Solutions

Math 198: Math for Machine Learning

Due Date: March 11

Name:

Student ID:

1 Working with Adjoint

Let $\mathbf{A} \in \mathbb{R}^{m \times n}$.

- (a) Show that $\ker \mathbf{A}^\top \mathbf{A} = \ker \mathbf{A}$.

$$\begin{aligned} \mathbf{v} \in \ker(\mathbf{A}^\top \mathbf{A}) &\iff \mathbf{A}^\top \mathbf{A} \mathbf{v} = \mathbf{0} \\ &\iff \mathbf{A} \mathbf{v} \in \ker(\mathbf{A}^\top) \\ &\iff \mathbf{A} \mathbf{v} \in \text{range}(\mathbf{A})^\perp \text{ by FTLA} \\ &\iff \mathbf{A} \mathbf{v} = \mathbf{0} \text{ since } \mathbf{A} \mathbf{v} \in \text{range}(\mathbf{A}) \\ &\iff \mathbf{v} \in \ker(\mathbf{A}) \end{aligned}$$

Therefore, $\ker(\mathbf{A}^\top \mathbf{A}) = \ker(\mathbf{A})$. □

- (b) Deduce that $\text{rank}(\mathbf{A}^\top \mathbf{A}) = \text{rank}(\mathbf{A})$.

We have that $\text{rank}(\mathbf{A}) + \dim \ker(\mathbf{A}) = n = \text{rank}(\mathbf{A}^\top \mathbf{A}) + \dim \ker(\mathbf{A}^\top \mathbf{A})$. Since $\dim \ker(\mathbf{A}) = \dim \ker(\mathbf{A}^\top \mathbf{A})$, $\text{rank}(\mathbf{A}) = \text{rank}(\mathbf{A}^\top \mathbf{A})$.

- (c) Suppose \mathbf{A} is square. Show that \mathbf{A} and \mathbf{A}^\top have the same eigenvalues.

$$\begin{aligned} p_{\mathbf{A}}(\lambda) &= \det(\mathbf{A} - \lambda \mathbf{I}) \\ &= \det(\mathbf{A} - \lambda \mathbf{I})^\top \\ &= \det(\mathbf{A}^\top - \lambda \mathbf{I}) \\ &= p_{\mathbf{A}^\top}(\lambda) \end{aligned}$$

as determinant is invariant under transpose. Since the two matrices have the same characteristic polynomial, they have the same eigenvalues.

- (d) Deduce that $\text{rank}(\mathbf{A}) = \text{rank}(\mathbf{A}^\top)$.

Because the matrices have the same eigenvalues, they have the same number of non-zero eigenvalues. Therefore, they have the same rank.

2 SVD

- (a) Let $\mathbf{A} \in \mathbb{R}^{m \times n}$. Let $\mathbf{A} = \mathbf{U} \mathbf{\Sigma} \mathbf{V}^\top$ be its SVD. Find the spectral decompositions of $\mathbf{A}^\top \mathbf{A}$ and $\mathbf{A} \mathbf{A}^\top$ in terms of \mathbf{U} , $\mathbf{\Sigma}$, \mathbf{V} .

$$\begin{aligned} \mathbf{A}^\top \mathbf{A} &= (\mathbf{U} \mathbf{\Sigma} \mathbf{V}^\top)^\top \mathbf{U} \mathbf{\Sigma} \mathbf{V}^\top \\ &= \mathbf{V} \mathbf{\Sigma} \mathbf{U} \mathbf{U}^\top \mathbf{\Sigma} \mathbf{V}^\top \\ &= \mathbf{V} \mathbf{\Sigma}^2 \mathbf{V}^\top \end{aligned}$$

$$\begin{aligned}
\mathbf{A}\mathbf{A}^\top &= \mathbf{U}\mathbf{\Sigma}\mathbf{V}^\top(\mathbf{U}\mathbf{\Sigma}\mathbf{V}^\top)^\top \\
&= \mathbf{U}\mathbf{\Sigma}\mathbf{V}^\top\mathbf{V}\mathbf{\Sigma}\mathbf{U}^\top \\
&= \mathbf{U}\mathbf{\Sigma}\mathbf{U}^\top
\end{aligned}$$

- (b) Prove: If $\mathbf{A} \in \mathbb{R}^{n \times n}$ is PSD, then the spectral decomposition of \mathbf{A} coincides with the SVD of \mathbf{A} . (The technical details involving rows or columns of zeros are omitted.) Let $\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^\top$ by SVD, and $\mathbf{A} = \mathbf{U}'\mathbf{\Lambda}\mathbf{U}'^\top$ by the Spectral Theorem. We seek to show that $\mathbf{U} = \mathbf{U}' = \mathbf{V}$ and $\mathbf{\Sigma} = \mathbf{\Lambda}$. Recall that \mathbf{V} is the unitary matrix in the spectral decomposition $\mathbf{A}^\top\mathbf{A} = \mathbf{V}\mathbf{\Sigma}^2\mathbf{V}^\top$. But $\mathbf{A}^\top\mathbf{A} = (\mathbf{U}'\mathbf{\Lambda}\mathbf{U}'^\top)^\top\mathbf{U}'\mathbf{\Lambda}\mathbf{U}'^\top = \mathbf{U}'\mathbf{\Lambda}^2\mathbf{U}'^\top$. So $\mathbf{V} = \mathbf{U}'$, and since $\mathbf{\Lambda}$ and $\mathbf{\Sigma}$ are PSD, $\mathbf{\Lambda} = \mathbf{\Sigma}$. So $\mathbf{U}\mathbf{\Sigma}\mathbf{V}^\top = \mathbf{U}'\mathbf{\Sigma}\mathbf{V}^\top$, and so $\mathbf{U} = \mathbf{U}'$.
- (c) Let $\mathbf{A} \in \mathbb{R}^{m \times n}$, and let $\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^\top$ be its SVD. If $\mathbf{V} = (\mathbf{v}_1, \dots, \mathbf{v}_n)$, $\mathbf{U} = (\mathbf{u}_1, \dots, \mathbf{u}_m)$, and $r = \text{rank}(\mathbf{A})$, then let

$$\begin{aligned}
\mathbf{V}_r &= (\mathbf{v}_1, \dots, \mathbf{v}_r) \\
\mathbf{U}_r &= (\mathbf{u}_1, \dots, \mathbf{u}_r).
\end{aligned}$$

Show that $\mathbf{v}_1, \dots, \mathbf{v}_r$ “diagonalize” \mathbf{A} in the following way: For $i = 1, \dots, r$, show that $\mathbf{A}\mathbf{v}_i = \sigma_i\mathbf{u}_i$. Because \mathbf{V} is unitary, $\mathbf{V}^\top\mathbf{v}_i$ will be 0 in every index except $(\mathbf{V}\mathbf{v}_i)_i = 1$. Since $\mathbf{A}\mathbf{v}_i = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^\top\mathbf{v}_i$,

$$\mathbf{A}\mathbf{v}_i = \mathbf{U} \begin{bmatrix} 0 & \dots & \sigma_i & \dots & 0 \end{bmatrix}^\top = \sigma_i\mathbf{u}_i$$

- (d) Let

$$\mathbf{A} = \begin{pmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{pmatrix}.$$

Compute the SVD of \mathbf{A} .

We first compute $\mathbf{A}^\top\mathbf{A}$:

$$\mathbf{A}^\top\mathbf{A} = \begin{bmatrix} 3 & 2 \\ 2 & 3 \\ 2 & -2 \end{bmatrix} \begin{bmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{bmatrix} = \begin{bmatrix} 13 & 12 & 2 \\ 12 & 13 & -2 \\ 2 & -2 & 8 \end{bmatrix}$$

Next, we find the eigenvalues of $\mathbf{A}^\top\mathbf{A}$:

$$\begin{aligned}
p_{\mathbf{A}^\top\mathbf{A}}(\lambda) &= \det \begin{bmatrix} 13 - \lambda & 12 & 2 \\ 12 & 13 - \lambda & -2 \\ 2 & -2 & 8 - \lambda \end{bmatrix} \\
&= (13 - \lambda) \det \begin{bmatrix} 13 - \lambda & -2 \\ -2 & 8 - \lambda \end{bmatrix} - 12 \det \begin{bmatrix} 12 & -2 \\ 2 & 8 - \lambda \end{bmatrix} + 2 \det \begin{bmatrix} 12 & 13 - \lambda \\ 2 & -2 \end{bmatrix} \\
&= (13 - \lambda)((13 - \lambda)(8 - \lambda) - 4) - 12(12(8 - \lambda) + 4) + 2(-24 - 2(13 - \lambda)) \\
&= (13 - \lambda)(100 - 21\lambda + \lambda^2) - 12(100 - 12\lambda) + 2(-50 + 2\lambda) \\
&= -\lambda^3 + 34\lambda^2 - 225\lambda
\end{aligned}$$

It is immediately clear that one eigenvalue is 0, as expected, since $\text{rank}(\mathbf{A}) = 2$. To find the other two, we factor:

$$-\lambda^3 + 34\lambda^2 - 225\lambda = -(\lambda - 9)(\lambda - 25)$$

So the eigenvalues of $\mathbf{A}^\top\mathbf{A}$ are 0, 9, and 25. We now find the eigenvectors of $\mathbf{A}^\top\mathbf{A}$. We first find the eigenvector corresponding to $\lambda_1 = 25$:

$$\begin{aligned}
\mathbf{v}_1 &\in \ker(\mathbf{A}^\top\mathbf{A} - 25\mathbf{I}) \\
&\in \ker \begin{bmatrix} -12 & 12 & 2 \\ 12 & -12 & -2 \\ 2 & -2 & -17 \end{bmatrix}
\end{aligned}$$

Observe that $(\mathbf{A}^\top \mathbf{A} - 25\mathbf{I})_1 = -(\mathbf{A}^\top \mathbf{A} - 25\mathbf{I})_2$. Therefore, $[1 \ 1 \ 0]^\top$ is an eigenvector of \mathbf{A} corresponding to $\lambda_1 = 25$. We turn next to $\lambda_2 = 9$:

$$\begin{aligned} \mathbf{v}_2 &\in \ker(\mathbf{A}^\top \mathbf{A} - 9\mathbf{I}) \\ &\in \ker \begin{bmatrix} 4 & 12 & 2 \\ 12 & 4 & -2 \\ 2 & -2 & -1 \end{bmatrix} \end{aligned}$$

Observe that $(\mathbf{A}^\top \mathbf{A} - 9\mathbf{I})_2 = (\mathbf{A}^\top \mathbf{A} - 9\mathbf{I})_1 + 4(\mathbf{A}^\top \mathbf{A} - 9\mathbf{I})_3$. Therefore, $[1 \ -1 \ 4]^\top$ is an eigenvector of \mathbf{A} corresponding to $\lambda_2 = 9$. We turn next to $\lambda_3 = 0$:

$$\begin{aligned} \mathbf{v}_3 &\in \ker(\mathbf{A}^\top \mathbf{A}) \\ &\in \ker \begin{bmatrix} 13 & 12 & 2 \\ 12 & 13 & -2 \\ 2 & -2 & 8 \end{bmatrix} \end{aligned}$$

Observe that $2(\mathbf{A}^\top \mathbf{A})_1 - 2(\mathbf{A}^\top \mathbf{A})_2 = (\mathbf{A}^\top \mathbf{A})_3$. Therefore, $[2 \ -2 \ -1]^\top$ is an eigenvector of \mathbf{A} corresponding to $\lambda_3 = 0$. Note that the eigenvectors we have collected are already an orthogonal basis for \mathbb{R}^3 ; we just need to normalize them to find the column vectors of \mathbf{V} . Therefore,

$$\mathbf{V} = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{3\sqrt{2}} & \frac{2}{3} \\ \frac{1}{\sqrt{2}} & -\frac{1}{3\sqrt{2}} & -\frac{2}{3} \\ 0 & \frac{4}{3\sqrt{2}} & -\frac{1}{3} \end{bmatrix}, \mathbf{\Lambda} = \begin{bmatrix} 25 & 0 & 0 \\ 0 & 9 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \mathbf{A}^\top \mathbf{A} = \mathbf{V} \mathbf{\Lambda} \mathbf{V}^\top$$

So, for our SVD, we have

$$\mathbf{\Sigma} = \begin{bmatrix} 5 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \mathbf{V} = \mathbf{V} \text{ from spectral decomposition}$$

We can then solve for \mathbf{U} :

$$\begin{aligned} \mathbf{U}' &= \mathbf{A} \mathbf{V} \mathbf{\Sigma}^{-1} \\ &= \begin{bmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{3\sqrt{2}} & \frac{2}{3} \\ \frac{1}{\sqrt{2}} & -\frac{1}{3\sqrt{2}} & -\frac{2}{3} \\ 0 & \frac{4}{3\sqrt{2}} & -\frac{1}{3} \end{bmatrix} \begin{bmatrix} \frac{1}{5} & 0 & 0 \\ 0 & \frac{1}{3} & 0 \\ 0 & 0 & 0 \end{bmatrix} \\ &= \begin{bmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{bmatrix} \begin{bmatrix} \frac{1}{5\sqrt{2}} & \frac{1}{9\sqrt{2}} & 0 \\ \frac{1}{5\sqrt{2}} & -\frac{1}{9\sqrt{2}} & 0 \\ 0 & \frac{4}{9\sqrt{2}} & 0 \end{bmatrix} \\ &= \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix} \end{aligned}$$

So the SVD of \mathbf{A} is

$$\mathbf{A} = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \end{bmatrix} \begin{bmatrix} 5 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{3\sqrt{2}} & \frac{2}{3} \\ \frac{1}{\sqrt{2}} & -\frac{1}{3\sqrt{2}} & -\frac{2}{3} \\ 0 & \frac{4}{3\sqrt{2}} & -\frac{1}{3} \end{bmatrix}^\top$$

(e) Let

$$\mathbf{A} = \begin{pmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{pmatrix}.$$

Find orthonormal bases for the four fundamental subspaces of \mathbf{A} .

From FTLA part II and the SVD from (d), we have

$$\begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix}, \begin{bmatrix} \frac{1}{3\sqrt{2}} \\ -\frac{1}{3\sqrt{2}} \\ \frac{4}{3\sqrt{2}} \end{bmatrix} \text{ are a basis for } \text{range}(\mathbf{A}^\top)$$

$$\begin{bmatrix} \frac{2}{3} \\ -\frac{2}{3} \\ -\frac{1}{3} \end{bmatrix} \text{ is a basis for } \ker(\mathbf{A})$$

$$\begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}, \begin{bmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{bmatrix} \text{ are a basis for } \text{range}(\mathbf{A})$$

\mathbf{A}^\top has a trivial kernel

3 PCA

Let

$$\mathbf{X} = \begin{bmatrix} -1 & 1 \\ 1 & -1 \\ -2 & -2 \\ 0 & 0 \\ 2 & 2 \end{bmatrix}$$

Note that the SVD of \mathbf{X} is

$$\text{SVD}(\mathbf{X}) = \begin{bmatrix} 0 & 0.7 & -0.5 & 0 & 0.5 \\ 0 & -0.7 & -0.5 & 0 & 0.5 \\ -0.7 & 0 & 0.5 & 0 & 0.5 \\ 0 & 0 & 0 & 1 & 0 \\ 0.7 & 0 & 0.5 & 0 & 0.5 \end{bmatrix} \begin{bmatrix} 4 & 0 \\ 0 & 2 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0.7 & 0.7 \\ -0.7 & 0.7 \end{bmatrix}$$

Compute the principal components of \mathbf{X} .

The principal components of \mathbf{X} are unit eigenvectors of $\mathbf{X}^\top \mathbf{X}$. In this case,

$$\mathbf{v}_1 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}$$