Problem set 3: Solutions

Mandatory part

Workout 0.1 With SVD factors

$$U = \begin{pmatrix} 0 & -\frac{1}{\sqrt{5}} & 0 & 0 & -\frac{2}{\sqrt{5}} \\ 0 & 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & -\frac{2}{\sqrt{5}} & 0 & 0 & \frac{1}{\sqrt{5}} \end{pmatrix} \quad \Sigma = \begin{pmatrix} 3 & 0 & 0 & 0 \\ 0 & \sqrt{5} & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad V = \begin{pmatrix} 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 0 \end{pmatrix}$$

a. The eigenvalues of A^TA are squares of singular values of A. The singular values of A are $\sigma_1 = 3$, $\sigma_2 = \sqrt{5}$, $\sigma_3 = 2$, and $\sigma_4 = 0$. Therefore, eigenvalues of A^TA are $\{9, 5, 4, 0\}$. Eigenvectors of A^TA are columns of V,

$$\boldsymbol{v}_1 = \begin{bmatrix} 0 \\ -1 \\ 0 \\ 0 \end{bmatrix}, \quad \boldsymbol{v}_2 = \begin{bmatrix} -1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad \boldsymbol{v}_3 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ -1 \end{bmatrix}, \quad \boldsymbol{v}_4 = \begin{bmatrix} 0 \\ 0 \\ -1 \\ 0 \end{bmatrix}.$$

b. The eigenvalues of AA^T are squares of singular values of A plus additional $m - n = 5 - 4 = 1$ zero. Therefore, the eigenvalue set of AA^T is $\{9, 5, 4, 0, 0\}$. Columns of U form the eigenvectors:

$$\boldsymbol{u}_1 = \begin{bmatrix} 0 \\ 0 \\ -1 \\ 0 \\ 0 \end{bmatrix}, \quad \boldsymbol{u}_2 = \begin{bmatrix} -\frac{1}{\sqrt{5}} \\ 0 \\ 0 \\ 0 \\ -\frac{2}{\sqrt{5}} \end{bmatrix}, \quad \boldsymbol{u}_3 = \begin{bmatrix} 0 \\ -1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad \boldsymbol{u}_4 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \quad \boldsymbol{u}_5 = \begin{bmatrix} -\frac{2}{\sqrt{5}} \\ 0 \\ 0 \\ 0 \\ \frac{1}{\sqrt{5}} \end{bmatrix}.$$

c. From Lecture 8 we know $rank(A^TA) = rank(AA^T) = rank(A) = number\ of\ non-zero\ sigular\ values$. In this case $rank(A^TA) = rank(AA^T) = 3$.

d. Similar to item c: $rank(A) = 3$

e. The condition number of a matrix is defined as the quotient of largest and smallest positive singular values:

$$cond_2(A) = \frac{largest\ singular\ value}{smallest\ positive\ singular\ value} = \frac{3}{2}$$

f. The reduced form (or "economy" form) SVD: Since A has dimensions 5×4, it results in a Σ with one zero row, specifically the last row. In our case here, two rows in Σ are zeros, but one of them is a consequence of the last singular value being zero. Consequently, in the reduced form, only the last column of U and the last row of Σ are eliminated:

$$A = \begin{pmatrix} 0 & -\frac{1}{\sqrt{5}} & 0 & 0 \\ 0 & 0 & -1 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & -\frac{2}{\sqrt{5}} & 0 & 0 \end{pmatrix} \begin{pmatrix} 3 & 0 & 0 & 0 \\ 0 & \sqrt{5} & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 0 \end{pmatrix}$$

g. Write the matrix A as a series of rank-1 matrices:

$$\begin{aligned} A &= \sigma_1 \boldsymbol{u}_1 \boldsymbol{v}_1^T + \sigma_2 \boldsymbol{u}_2 \boldsymbol{v}_2^T + \sigma_3 \boldsymbol{u}_3 \boldsymbol{v}_3^T \\ &= 3 \begin{bmatrix} 0 \\ 0 \\ -1 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} 0 & -1 & 0 & 0 \end{bmatrix} + \sqrt{5} \begin{bmatrix} -\frac{1}{\sqrt{5}} \\ 0 \\ 0 \\ 0 \\ -\frac{2}{\sqrt{5}} \end{bmatrix} \begin{bmatrix} -1 & 0 & 0 & 0 \end{bmatrix} + 2 \begin{bmatrix} 0 \\ -1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 & -1 \end{bmatrix} \\ &= \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \end{aligned}$$

h. The closest rank 2 matrix to A:

$$A_2 = \sigma_1 \boldsymbol{u}_1 \boldsymbol{v}_1^T + \sigma_2 \boldsymbol{u}_2 \boldsymbol{v}_2^T = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & 0 \end{bmatrix}$$

The error is $\|A - A_2\|_2 = \sigma_3 = 2$

i. Python command for item (h):

```
U,S,Vt = numpy.linalg.svd(A)
A2 = U[:,:2] @ np.diag(S[:2]) @ Vt[:2,:]
```

Workout 0.2

a. The pseudoinverse is $A^+ = V_1\Sigma_1^{-1}U_1^T$ where U_1 and V_1 are the first r columns of U and V, respectively, and Σ_1 is the leading $r \times r$ block of Σ. Here $r = 3$ is the rank of A. Thus we have:

$$A^+ = \begin{pmatrix} 0 & -1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} \frac{1}{3} & 0 & 0 \\ 0 & \frac{1}{\sqrt{5}} & 0 \\ 0 & 0 & \frac{1}{2} \end{pmatrix} \begin{pmatrix} 0 & 0 & -1 & 0 & 0 \\ -\frac{1}{\sqrt{5}} & 0 & 0 & 0 & -\frac{2}{\sqrt{5}} \\ 0 & -1 & 0 & 0 & 0 \end{pmatrix}$$
$$= \begin{pmatrix} \frac{1}{5} & 0 & 0 & 0 & \frac{2}{5} \\ 0 & 0 & \frac{1}{3} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 0 & 0 \end{pmatrix}$$

b. Find all least squares solutions of $A\boldsymbol{x} = \boldsymbol{b}$: The matrix is rank-deficient, $rank(A) = 3 < 4$. Referring to page 2 (case 2) in Lecture 9, we set $\boldsymbol{y}_1 = \Sigma_1^{-1}U_1^T\boldsymbol{b}$ and $\boldsymbol{y}_2$ an arbitrary vector (here an arbitrary scalar because deficiency is 1). If $\boldsymbol{y} = \begin{bmatrix} \boldsymbol{y}_1 \\ \boldsymbol{y}_2 \end{bmatrix}$ then $\boldsymbol{x} = V\boldsymbol{y}$ represents all least squares solutions.

$$\boldsymbol{y}_1 = \begin{pmatrix} \frac{1}{3} & 0 & 0 \\ 0 & \frac{1}{\sqrt{5}} & 0 \\ 0 & 0 & \frac{1}{2} \end{pmatrix} \begin{pmatrix} 0 & 0 & -1 & 0 & 0 \\ -\frac{1}{\sqrt{5}} & 0 & 0 & 0 & -\frac{2}{\sqrt{5}} \\ 0 & -1 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 4 \\ 3 \\ 0 \\ 2 \end{pmatrix} = \begin{pmatrix} -1 \\ -1 \\ -2 \end{pmatrix}$$

We assume that $\boldsymbol{y}_2 = c$ (an arbitrary scalar) to get

$$\boldsymbol{y} = \begin{bmatrix} \boldsymbol{y}_1 \\ \boldsymbol{y}_2 \end{bmatrix} = \begin{pmatrix} -1 \\ -1 \\ -2 \\ c \end{pmatrix}$$

and finally

$$\boldsymbol{x} = V\boldsymbol{y} = \begin{pmatrix} 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 0 \end{pmatrix} \begin{pmatrix} -1 \\ -1 \\ -2 \\ c \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ -c \\ 2 \end{pmatrix}$$

This means that all vectors of the form $\boldsymbol{x} = \begin{bmatrix} 1 & 1 & -c & 2 \end{bmatrix}^T$, with c arbitrary, are least squares solutions of $A\boldsymbol{x} = \boldsymbol{b}$ (infinite number of solutions).

c. The norm minimal solution is obtained by setting $\boldsymbol{y}_2 = c = 0$, because in this case the norm 2 of $\boldsymbol{x}$ is minimum. So $\boldsymbol{x} = \begin{bmatrix}1 & 1 & 0 & 2\end{bmatrix}^T$ is the norm-minimal solution. Equivalent method to obtain the norm-minimal solution is

$$\boldsymbol{x} = A^+\boldsymbol{b} = \begin{pmatrix} \frac{1}{5} & 0 & 0 & 0 & \frac{2}{5} \\ 0 & 0 & \frac{1}{3} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 4 \\ 3 \\ 0 \\ 2 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 2 \end{pmatrix}$$

Workout 0.3 Assume that the SVD of A is given by $A = U\Sigma V^T$.

a. $A^T = (U\Sigma V^T)^T = (V^T)^T\Sigma^T U^T = V\Sigma^T U^T$. Schematic for a matrix A of size 4×2:

$$A = \begin{bmatrix} \times & \times \\ \times & \times \\ \times & \times \\ \times & \times \end{bmatrix} = \begin{bmatrix} \times & \times & \times & \times \\ \times & \times & \times & \times \\ \times & \times & \times & \times \\ \times & \times & \times & \times \end{bmatrix} \begin{bmatrix} \times & 0 \\ 0 & \times \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \times & \times \\ \times & \times \end{bmatrix}$$

For A^T we have

$$A^T = \begin{bmatrix} \times & \times & \times & \times \\ \times & \times & \times & \times \end{bmatrix} = \begin{bmatrix} \times & \times \\ \times & \times \end{bmatrix} \begin{bmatrix} \times & 0 & 0 & 0 \\ 0 & \times & 0 & 0 \end{bmatrix} \begin{bmatrix} \times & \times & \times & \times \\ \times & \times & \times & \times \\ \times & \times & \times & \times \\ \times & \times & \times & \times \end{bmatrix}$$

b. $\alpha A = U(\alpha\Sigma)V^T$. Singular values of αA are $\alpha\sigma_1, \alpha\sigma_2, \ldots, \alpha\sigma_n$.

c. $A^{-1} = (U\Sigma V^T)^{-1} = (V^T)^{-1}\Sigma^{-1}U^{-1} = V\Sigma^{-1}U^T$. Note that since A is square, Σ is also square, and since A is non-singular, all singular values are positive, so Σ^{-1} exists.

d. Assume that the rank of A is r and U_1 and V_1 are the first r columns of U and V, respectively. Besides assume that Σ_1 is the leading $r \times r$ block of Σ. We know $A^+ = V_1\Sigma_1^{-1}U_1^T$ which gives $(A^+)^T = (V_1\Sigma_1^{-1}U_1^T)^T = U_1\Sigma_1^{-1}V_1^T$. On the other side, $A^T = V\Sigma^T U^T = V_1\Sigma_1 U_1^T$ (reduced form). With the same argument $(A^T)^+ = U_1\Sigma_1^{-1}V_1^T$. Therefore, $(A^+)^T = (A^T)^+$.

Workout 0.4 First, we compute $A^TA = VDV^T$ where $D = \Sigma^T\Sigma$:

$$A^TA = \begin{pmatrix} 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 0 \end{pmatrix} \begin{pmatrix} 9 & 0 & 0 & 0 \\ 0 & 5 & 0 & 0 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & -1 & 0 \end{pmatrix} = \begin{pmatrix} 5 & 0 & 0 & 0 \\ 0 & 9 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 4 \end{pmatrix}$$

The matrix is diagonal and exact eigenvalues are $\{9, 5, 4, 0\}$. Eigenvectors are columns of V. The dominant eigenvalue is $\lambda_1 = 9$ with corresponding eigenvector $\boldsymbol{v}_1 = \begin{bmatrix} 0 & -1 & 0 & 0 \end{bmatrix}^T$. Now we apply the power method with initial guess $\boldsymbol{v}^{(0)} = \begin{bmatrix} 1 & 1 & 1 & 1 \end{bmatrix}^T$. Assume $B = A^T A$.

$$\boldsymbol{v} = B\boldsymbol{v}^{(0)} = \begin{pmatrix} 5 & 0 & 0 & 0 \\ 0 & 9 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 4 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 5 \\ 9 \\ 0 \\ 4 \end{pmatrix}, \quad \boldsymbol{v}^{(1)} = \frac{\boldsymbol{v}}{\|\boldsymbol{v}\|_2} \doteq \begin{pmatrix} 0.4527 \\ 0.8148 \\ 0 \\ 0.3621 \end{pmatrix}, \quad \lambda^{(1)} = \boldsymbol{v}^{(1)T} B \boldsymbol{v}^{(1)} \doteq 7.5246$$

$$\boldsymbol{v} = B\boldsymbol{v}^{(1)} = \begin{pmatrix} 2.2634 \\ 7.3334 \\ 0 \\ 1.4486 \end{pmatrix}, \quad \boldsymbol{v}^{(2)} = \frac{\boldsymbol{v}}{\|\boldsymbol{v}\|_2} \doteq \begin{pmatrix} 0.2898 \\ 0.9389 \\ 0 \\ 0.1855 \end{pmatrix}, \quad \lambda^{(2)} = \boldsymbol{v}^{(2)T} B \boldsymbol{v}^{(2)} \doteq 8.4921$$

$$\boldsymbol{v} = B\boldsymbol{v}^{(2)} = \begin{pmatrix} 1.4490 \\ 8.4505 \\ 0 \\ 0.7419 \end{pmatrix}, \quad \boldsymbol{v}^{(3)} = \frac{\boldsymbol{v}}{\|\boldsymbol{v}\|_2} \doteq \begin{pmatrix} 0.1684 \\ 0.9819 \\ 0 \\ 0.0862 \end{pmatrix}, \quad \lambda^{(3)} = \boldsymbol{v}^{(3)T} B \boldsymbol{v}^{(3)} \doteq 8.8494$$

We observe the sequence $\boldsymbol{v}^{(k)}$ approaches a negative multiple of $\boldsymbol{v}_1$ and the sequence $\lambda^{(k)}$ approaches $\lambda_1 = 9$.

Workout 0.5 Same as the previous workout, we apply the power method on matrix

$$A = \begin{pmatrix} 5 & 0 & 0 \\ 0 & -5 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

The 10^{th} iteration is:

$$\boldsymbol{v}^{(10)} \doteq \begin{pmatrix} 0.7071 \\ 0.7071 \\ 0.0000 \end{pmatrix}, \quad \lambda^{(10)} \doteq 0.0000$$

which is not close to the exact dominant eigenpair. The divergence occurs due to the fact that, for this particular matrix, the magnitudes of λ_1 and λ_2 are equal ($|\lambda_1| = |\lambda_2|$). The convergence rate of the power method is characterized by $\left(\frac{|\lambda_2|}{|\lambda_1|}\right)^k$ as k increases. For this specific matrix, the quotient is equal to 1 resulting to a non-convergent iteration.

Non-mandatory part

Workout 0.6 Refer to Lecture 9 (Application of SVD to image compression)

Workout 0.7

a. The vector $\boldsymbol{x}$ must satisfy $\boldsymbol{v}_1^T \boldsymbol{x} = 1$ where $\boldsymbol{v}_1$ is the first eigenvector which has already been computed. A simple candidate for $\boldsymbol{x}$ is $\boldsymbol{x} = \boldsymbol{v}_1$.

b. We apply the power method on deflation matrix $B = A - \lambda_1 \boldsymbol{v}_1 \boldsymbol{x}^T = A - \lambda_1 \boldsymbol{v}_1 \boldsymbol{v}_1^T$. From the theorem, λ_2 is the dominant eigenvalue of B. Iteration must converge because $|\lambda_2| > |\lambda_3|$ by assumption.

c. The application of power method to B results in eigenvalue λ_2 and a vector $\boldsymbol{u}_2$ (the dominant eigenvector of B). To compute $\boldsymbol{v}_2$ we use the given formula as below

$$\begin{aligned}
\boldsymbol{u}_2 =& \boldsymbol{v}_2 - \left(\frac{\lambda_1}{\lambda_2}\boldsymbol{x}^T \boldsymbol{v}_2\right)\boldsymbol{v}_1 \\
&= \boldsymbol{v}_2 - \boldsymbol{v}_1 \left(\frac{\lambda_1}{\lambda_2}\boldsymbol{x}^T \boldsymbol{v}_2\right) \\
&= \boldsymbol{v}_2 - \frac{\lambda_1}{\lambda_2}(\boldsymbol{v}_1 \boldsymbol{x}^T)\boldsymbol{v}_2 \\
&= \left(I - \frac{\lambda_1}{\lambda_2}\boldsymbol{v}_1 \boldsymbol{x}^T\right)\boldsymbol{v}_2
\end{aligned}$$

We can solve the linear system $E\boldsymbol{v}_2 = \boldsymbol{u}_2$ where $E = I - \frac{\lambda_1}{\lambda_2}\boldsymbol{v}_1 \boldsymbol{x}^T$ to obtain $\boldsymbol{v}_2$.

d. Yes! this process can be extended to calculate other eigenvalues and their corresponding eigenvectors. However, eigenvalues must be distinct in magnitude; they must be real and differ in their magnitudes. Another issue with this approach is the necessity to solve a linear system of equations at each iteration to determine the corresponding eigenvector.

Workout 0.8 Yes, it is an iterative method: it iteratively produces approximations to a quasi-triangular matrix that is similar to A. For the second part of the question refer to Lecture 10.