# Linear Algebra
# Home assignment 5: Matrix factorisation

**Problem 1** (Rank of a matrix; 2 pt). (a) Assume that $A$ is an $m \times n$ matrix of rank $r$. Without using the singular value decomposition, prove that $A$ can be written as a sum of $r$ summands $\mathbf{u}_j \mathbf{v}_j^\top$ of rank 1. Is such a representation unique?

(b) For the matrices below, find their ranks $r$ and represent them as the sum of $r$ rank-one summands (do not use SVD yet!).

$$A = \begin{pmatrix} 1 & -2 & 3 & -4 \\ -2 & 4 & -6 & 8 \\ 3 & -6 & 9 & -12 \\ -4 & 8 & -12 & 16 \end{pmatrix}, \qquad B = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 5 \\ 3 & 4 & 5 & 6 \\ 4 & 5 & 6 & 7 \end{pmatrix}.$$

**Problem 2** (Singular values; 3 pt). (a) Prove that matrices $A$ and $A^\top$ have the same non-zero singular values.

(b) Find singular values of the following matrices:

$$(i) \quad \begin{pmatrix} 0 & 1 & 2 \end{pmatrix}; \qquad (ii) \quad \begin{pmatrix} 0 & 1 & 2 \end{pmatrix}^\top; \qquad (iii) \quad \begin{pmatrix} 1 & 2 & 2 \\ 1 & 3 & 3 \end{pmatrix}; \qquad (iv) \quad \begin{pmatrix} 2 & 0 \\ 1 & 2 \end{pmatrix}.$$

Hint: Use (a) in (i) and (iii)

**Problem 3** (Singular value decomposition; 4 pt). (a) If $A = U\Sigma V^\top$ is the SVD of $A$, what is the SVD of $A^\top$?

(b) Assume that $A = \mathbf{u}\mathbf{v}^\top$ is an $m \times n$ matrix of rank 1. Find the SVD of $A$.

(c) Find SVD of the following matrices:

$$(i) \quad \begin{pmatrix} 0 & 1 & 2 \end{pmatrix}; \qquad (ii) \quad \begin{pmatrix} 0 & 1 & 2 \end{pmatrix}^\top; \qquad (iii) \quad \begin{pmatrix} 2 & 1 & -2 \\ -2 & -1 & 2 \end{pmatrix}.$$

**Problem 4** (Singular value decomposition, 4 pt). Find the singular value decomposition of the matrix

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

To this end, complete the following steps:

(a) How many singular values $\sigma_j$ does $A$ have? How many of them are non-zero? Find the nonzero singular values $\sigma_1, \dots, \sigma_r$.

(b) find the right singular vectors $A\mathbf{v}_j = \sigma_j \mathbf{u}_j$, $j = 1, \dots, r$;

(c) find the left singular vectors $A^\top \mathbf{u}_j = \sigma_j \mathbf{v}_j$, $j = 1, \dots, r$;

(d) form the unitary matrices $U$ and $V$ and write the singular value decomposition of $A$.

(Hint: you may find it easier to work with $A^\top$)

**Problem 5** (Low-rank approximation; 2 pt). (a) For the matrix $A$ in Problem 4, find a unit vector $\mathbf{x} \in \mathbb{R}^3$ for which $A\mathbf{x}$ has the maximal length $\alpha$. What is $\alpha$? Justify your answer.

(b) Find the best rank one approximation for the matrix $A$ of Problem 4 in the Frobenius norm. Justify your answer

**Problem 6** (Low-rank approximation; 4 pt). Prove that the best rank $k$ approximation $A_k$ of an $m \times n$ matrix $A$ in the Frobenius norm is given by the first $k$ terms in the SVD of $A$, i.e.,

$$A_k = \sigma_1 \mathbf{u}_1 \mathbf{v}_1^\top + \cdots + \sigma_k \mathbf{u}_k \mathbf{v}_k^\top.$$

**Problem 7** (SVD and image compression; 5 pt). Take a jpg-picture $A$ of yourselves of reasonable size (say $1000 \times 1000$ pixels), perform the SVD (use `Python` or `R` or any other program of your choice) and find the best rank-$k$ approximation of $A$ with $k = 1, 2, 5, 10, 20, 50$. For what $k$ can one recognize the picture? Comment on how much the quality and the size increase along with $k$.

(Hint: one can use the Frobenius distance to the original picture as a quality measure)

**Problem 8** (Polar decomposition; 3 pt). Find the positive definite square root $S = V\Sigma V^\top$ of $A^\top A$ and its polar decomposition $A = QS$:

$$A = \frac{1}{\sqrt{10}} \begin{pmatrix} 10 & 6 \\ 0 & 8 \end{pmatrix}$$

**Problem 9** (Pseudo-inverse; 3 pt). (a) Assume that $A$ has linearly independent columns. Show that the matrix $A^+ := (A^\top A)^{-1} A^\top$ is well defined and gives a left inverse of $A$. What is $AA^+$?

(b) Show that the above matrix $A^+$ is the Moore–Penrose pseudo-inverse of $A$ (i.e., that $AA^+A = A$, $A^+AA^+ = A^+$, and the matrices $AA^+$ and $A^+A$ are symmetric)

(c) Assume that $A = QR$ is a $QR$-factorization of $A$. Express the above $A^+$ in terms of $Q$ and $R$.

**Problem 10** (Pseudo-inverse; 3 pt). Find the SVD and the Moore–Penrose pseudo-inverse $V\Sigma^+U^\top$ of the matrices

$$A = \begin{pmatrix} 1 & 1 & 1 & 1 \end{pmatrix}, \qquad B = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \qquad C = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}.$$

**Problem 11** (Shortest solution; 2 pt). When columns of a matrix $A$ are linearly dependent and a vector $\mathbf{b}$ is not in the column space of $A$, then the least square solution of $A\mathbf{x} = \mathbf{b}$ is not unique. Find the minimum-length least square solution $\mathbf{x}^+ = A^+\mathbf{b}$ of the equation $A\mathbf{x} = \mathbf{b}$ with

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \end{pmatrix}, \qquad \mathbf{b} = \begin{pmatrix} 0 \\ 2 \\ 2 \end{pmatrix}.$$

**Problem 12** (Cholesky decomposition; 4 pt). Let

$$A = \begin{pmatrix} 4 & 2 & 2 \\ 2 & 10 & 10 \\ 2 & 10 & 14 \end{pmatrix} \qquad \text{and} \qquad B = \begin{pmatrix} 1 & -4 & 2 \\ -4 & 1 & -2 \\ 2 & -2 & -2 \end{pmatrix}$$

(a) Without computing the eigenvalues of $A$, show that $A$ is positive definite.

(b) Factor $A$ into a product $LDL^T$ where $L$ is lower triangular with 1 on the diagonal and $D$ is diagonal.

(c) Compute the Cholesky factorization of $A$.

(d) Without computing the eigenvalues of $B$, show that $B$ is indefinite and thus does not possess a Cholesky decomposition. Then determine its *generalized Cholesky decomposition* $B = LDL^\top$ where $L$ is lower triangular with 1 on the diagonal and $D$ is diagonal.

**Problem 13** (Principal component analysis; 3 pt)**.** (a) Convert the matrix of observation $A$ to the row zero-mean form and then construct the sample covariance matrix:

$$A = \begin{pmatrix} 19 & 22 & 6 & 3 & 2 & 20 \\ 12 & 6 & 9 & 15 & 13 & 5 \end{pmatrix}$$

(b) Find the principal components of the data in matrix $A$.

(c) Let $x_1$, $x_2$ denote the variables for the two-dimensional data in (a). Find the new variable $y_1 = a_1x_1 + a_2x_2$ such that $y_1$ has maximum possible variance over the given data. How much of the variance in the data is explained by $y_1$?

**Problem 14** (PCA; 5 pt)**.** (a) Simulate $N = 100$ data $(x_k, y_k)$ from the two-dimensional Gaussian (normal) distribution $\mathcal{N}(\mu_1, \mu_2; \sigma_1^2, \sigma_2^2, \rho)$ with $\mu_1 = 1$, $\mu_2 = 2$, $\sigma_1 = 4$, $\sigma_2 = 9$, $\rho = \frac{1}{3}$.

Hint: Can you do this easily if $\rho = 0$? If $(Z_1, Z_2)^\top \sim \mathcal{N}(0, 0; 1, 1, 0)$, show that $(X, Y)^\top$ with $X = \mu_1 + \sigma_1 Z_1$ and $Y = \mu_2 + \sigma_2 \rho Z_1 + \sqrt{1 - \rho^2}\sigma_2 Z_2$ has the required distribution

(b) Form the empirical covariance matrix $C$ for the data simulated and compare it with the theoretical covariance matrix. Find the eigenvalues and eigenvectors of $C$.

(c) Perform the PCA on the data generated. Calculate the variance along the first component; what fraction of the total variance does it include?

(d) Predict how the above fraction depends on $\rho$ and confirm your reasoning numerically.