Answer the following questions.

1. How is a *norm* $L : \mathbb{F}^n \to [0, \infty)$ defined?
2. Assume you are given the singular value decomposition $U\Sigma V^\top = A$ of some matrix $A \in \mathbb{R}^{m \times n}$, where the entries on the diagonal of Σ are given in a descending order. Denote the best rank-k approximation of A for $k \leq \min\{m, n\}$ and denote the criterion with respect to which this is the best approximation.
3. Give an example for a matrix $A \in \mathbb{R}^{2\times 2}$ where the LU-decomposition algorithm necessarily needs a permutation step.
4. Let $\lambda_1 \neq \lambda_2$ be two eigenvalues of a *symmetric* matrix $A \in \mathbb{R}^{n\times n}$, and let $v_1, v_2 \in \mathbb{R}^n$ be corresponding eigenvectors. Proof that $v_1^\top v_2 = 0$.
5. How is positive definiteness of a matrix $A \in \mathbb{R}^{n\times n}$ defined? What does this mean for the angle between a vector $x \in \mathbb{R}^n$ and the vector $z := Ax$?
6. How is *injectivity* of a function $f \colon X \to Y$ defined?
7. How is a *scalar product* $P \colon \mathbb{R}^n \times \mathbb{R}^n \mapsto \mathbb{R}$ defined?
8. Assume you are given the singular value decomposition (SVD) $U\Sigma V^\top = A$ of some matrix $A \in \mathbb{R}^{m\times n}$. What is the singular value decomposition of $A^\top$? Give a basis for $\Im(A)$ and $\Im(A^\top)$.
9. Assume you are given the singular value decomposition (SVD) $U\Sigma V^\top = A$ of some matrix $A \in \mathbb{R}^{m\times n}$. Determine the pseudoinverse A^+ of A with the help of the SVD. Explain why $A^+ = A^{-1}$, if A is invertible (*hint:* note that $m = n$ in this case).
10. Let $A \in \mathbb{R}^{n\times n}$ be a matrix. Are the notions of injectivity and surjectivity of A equivalent? Give a short justification.
11. What is the normal equation? Where is it applied?
12. Give an example of a vector space other than $\mathbb{R}^n$?
13. Let V be a vector space over the field $\mathbb{F}$. Give the definition of a basis.
14. Draw the sets $\{x \in \mathbb{R}^2 \colon \|x\|_p = 1\}$ for $p = 1, 2, \infty$.
15. What is the purpose of the power method? Write down its iteration instruction.
16. What is the purpose of the Richardson iteration? Write down its iteration instruction. When does it converge?
17. Let $R = (r_{ij})_{ij} \in \mathbb{R}^{n\times n}$ be a (lower or upper) triangular matrix with $r_{nn} = 0$. Is R invertible? Explain your answer.
18. Let $A \in \mathbb{R}^{n\times n}$ be a symmetric matrix. Are its singular values equal to its eigenvalues?
19. Let $A \in \mathbb{R}^{m\times n}$ have independent columns. How can we use the QR-decomposition of A to solve the least squares problem $\min\limits_{x\in\mathbb{R}^n} \|Ax - b\|_2^2.$, where $b \in \mathbb{R}^m$.
20. How is the rank of a real matrix $A \in \mathbb{R}^{m\times n}$ defined?
21. What does the dimension formula say?
22. Denote the optimization problem which is related to the principal component analysis.
23. When is a diagonal matrix invertible? Write down the inverse in this case.
24. What is the purpose of the QR Algorithm? Write down its iteration instruction.
25. Consider the iteration $x_{k+1} = Mx_k + b$ for some matrix $M \in \mathbb{R}^{n\times n}$ and vector $b \in \mathbb{R}^n$. Name a sufficient condition for the convergence of this sequence. What is the limit in this case?
26. What is the definition of an orthogonal matrix? What does it mean for the columns of the matrix?

Solution:

1. $(2P)$ $L : \mathbb{F}^n \to [0, +\infty)$ norm: $\Leftrightarrow$

$$\begin{aligned} &\text{i) } L(x) = 0 \quad \Rightarrow \quad x = 0 \\ &\text{ii) } L(\lambda x) = |\lambda| L(x) \\ &\text{iii) } L(x+y) \leq L(x) + L(y) \end{aligned}$$

2. $A = U\Sigma V^T = \sum_{j=1}^{\min(n,m)} \sigma_j u_j v_j^T$
   best rank-k approximation is given by truncated SVD

$$\begin{aligned} (1P) \quad A_k &:= \sum_{j=1}^{k} \sigma_j u_j v_j^T \text{ for which} \\ (1P) \quad A_k &:= \min_{B, \text{rank}(B)=k} \|B - A\|_F^2. \end{aligned}$$

3. $(2P)$ $A = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$

4. $(2P)$ $A \in \mathbb{R}^{n \times n}$ symmetric , $\lambda_1 \neq \lambda_2 \in \sigma(A)$ with v_1, v_2

$$\begin{aligned} &\Rightarrow \quad v_1^T A v_2 = \lambda_2 v_1^T v_2 \\ &\text{and} \quad v_1^T A v_2 = v_2^T A^T v_1 = \lambda_1 v_1^T v_2 \\ &\Rightarrow \quad (\lambda_1 - \lambda_2) v_1^T v_2 = 0 \\ &\overset{\lambda_1 \neq \lambda_2}{\Rightarrow} \quad v_1^T v_2 = 0 \end{aligned}$$

5. $A \in \mathbb{R}^{n \times n}$ positive definite $:\Leftrightarrow \quad \forall x \in \mathbb{R}^n \setminus \{0\} : \; x^T A x > 0$

$$\begin{aligned} &\Rightarrow \quad 0 < x^T \underbrace{(Ax)}_{:=z} = \cos(\alpha) \underbrace{\|x\| \|Ax\|}_{\geq 0} \quad (1P) \\ &\Rightarrow \quad \cos(\alpha) > 0 \quad \Rightarrow \quad \alpha \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right), \; |\alpha| < 90^\circ \quad (1P) \end{aligned}$$

6. $(1P)$ $f : X \to Y$ injective $:\Leftrightarrow \quad f(x) = f(y) \quad \Rightarrow \quad x = y \quad \forall x, y \in X$

7. $P : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ scalar product, if

$$\begin{aligned} &(1P) \quad \text{i) } \forall x, y \in \mathbb{R}^n : \; P(x,y) = P(y,x) \\ &(1P) \quad \text{ii) } \forall x \in \mathbb{R}^n \setminus \{0\} : \; P(x,x) > 0 \\ &(1P) \quad \text{iii) } \forall x, y, z \in \mathbb{R}^n : \; P(x, y+z) = P(x,y) + P(x,z) \\ &(1P) \quad \text{iv) } \forall x, y \in \mathbb{R}^n, \lambda \in \mathbb{R} : \; P(x, \lambda y) = \lambda P(x,y) \end{aligned}$$

8.
   - $(1P)$ pseudoinverse: $A^+ = V\Sigma^+ U^T$, where $\Sigma^+ = \text{diag}(\frac{1}{\sigma_i} : \; \sigma_i \neq 0)$
   - Let $A \in GL_n(\mathbb{R})$, then $(1P)$ $\sigma_{ii} \neq 0 \; \forall i$ and $A^{-1} = (U\Sigma V^T)^{-1} = V\Sigma^{-1}U^T$ $(1P)$ $(V, U$ orthogonal). Since $\Sigma^{-1} = \text{diag}(\frac{1}{\sigma_i}) = \Sigma^+$ we find $A^{-1} = A^+$ $(1P)$

9.

10.

11. $(1P)$ $\mathbb{R}^{m \times n}$, $\; P_n(\mathbb{R}) := \{x \mapsto \sum_{i=0}^n \alpha_i x^i : \; (\alpha_0, \ldots, \alpha_n) \in \mathbb{R}^{n+1}\}$

12. $\{v_1, \ldots, v_n\} \subset V$ basis $:\Leftrightarrow$

$$\begin{aligned} &(1P) \quad \text{i) } \{v_1, \ldots, v_n\} \text{ linearly independent } \quad (\Leftrightarrow \; \sum \lambda_j v_j = 0 \; \Rightarrow \; \lambda_j = 0 \; \forall \, j \\ &(1P) \quad \text{ii) } \text{span}(v_1, \ldots, v_n) = V \quad (\Leftrightarrow \; \{\sum \lambda_j v_j : \; \lambda_j \in \mathbb{R}\} = V) \end{aligned}$$

13. norms.pdf

14. $(2P)$ The power iteration is an algorithm to find an eigenvector of a matrix $A \in \mathbb{R}^{n\times n}$ corresponding to its largest eigenvalue. The drawback of this algorithm is the previous mentioned fact, that the resulting eigenvector is corresponded to the largest eigenvalue.
The iterations are computed by $x^{k+1} := \frac{Ax^k}{\|Ax^k\|}$.
As an alternative you can use the Inverse Power iteration.

15. $(2P)$ R is not invertible, because triangular matrices are invertible if and only if all diagonal entries are nonzero (see backward/forward substitution)

16. $(2P)$ No!

$$\lambda_i \in \sigma(A^TA),\ \sigma_i := \sqrt{\lambda} = |\tilde{\lambda}_i|$$
$$\tilde{\lambda}_i \in \sigma(A) \quad \Rightarrow \quad \tilde{\lambda}_i = \lambda_i \in \sigma(A^TA) = \sigma(A^2)$$

<u>Thus:</u> They are only equal up to the sign.

17. $(2P)$ Insert $A = QR$ into the normal equation:

$$A^TAx = A^Tb \quad \Leftrightarrow \quad (QR)^TQRx = (QR)^Tb \quad \Leftrightarrow \quad R^TRx = R^TQ^Tb$$
$$\overset{[R^T \text{ invertible, since A has independent columns}]}{\Leftrightarrow} Rx = Q^Tb$$

18. $(2P)$ The rank of a real matrix $A \in \mathbb{R}^{n\times m}$ is defined as the number of positive singular values ($= \dim(\mathrm{Im}A)$ $=$ number of linear independent columns).

19. $(2P)$ $\mathrm{rank}(A) + \dim(\ker(A)) = n$ (for $A \in \mathbb{R}^{m\times n}$)

20. $(2P)$ $A \in \mathbb{R}^{m\times n}$, then the first k principal components are given by

$$U_k := \underset{z\, \in\, \mathbb{R}^{m\times k}\ \text{orthogonal}}{\mathrm{argmin}} \|A - zz^TA\|_F^2.$$

21. $(1P)$ $D =$ diag(d_{ii}) invertible $\Leftrightarrow$ $d_{ii} \neq 0\ \forall i$
$(1P)$ Then $D^{-1} =$ diag$(\frac{1}{d_{ii}})$

22. $(1P)$ Purpose: Compute eigenvalues of a matrix $A \in \mathbb{R}^{n\times n}$

$$\begin{aligned} (1P)\quad & A_0 := A \\ & \text{for } i = 1, \dots, n \\ & \quad Q_iR_i := A_i \\ & \quad A_{i+1} := R_iQ_i \end{aligned}$$

23. $\rho(M) < 1$ $(1P)$ $\Rightarrow$ $(x_k)_k$ converges to fixed point $x^* = Mx^* + b$ $(1P)$

24. $(1P)$ $Q \in \mathbb{R}^{n\times n}$ orthogonal $:\Leftrightarrow$ $Q^TQ = I$
$(1P)$ Thus the columns of Q are mutually orthonormal.

1. $(2P)$ The rank of a real matrix $A \in \mathbb{R}^{n\times m}$ is defined as the number of positive singular values ($= \dim(\mathrm{Im} A) =$ number of linear independent columns).

2. $(2P)$ The power iteration is an algorithm to find an eigenvector of a matrix $A \in \mathbb{R}^{n\times n}$ corresponding to its largest eigenvalue. The drawback of this algorithm is the previous mentioned fact, that the resulting eigenvector is corresponded to the largest eigenvalue.
The iterations are computed by $x^{k+1} := \frac{Ax^k}{\|Ax^k\|}$.
As an alternative you can use the Inverse Power iteration.

3. $(2P)$ $A = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}$

4. $(2P)$ $\mathrm{rank}(A) + \dim(\ker(A)) = n \quad (\text{for } A \in \mathbb{R}^{m\times n})$

5. $(2P)$ No!

$$\lambda_i \in \sigma(A^TA),\ \sigma_i := \sqrt{\lambda} = |\tilde{\lambda}_i|$$
$$\tilde{\lambda}_i \in \sigma(A) \quad \Rightarrow \quad \tilde{\lambda}_i = \lambda_i \in \sigma(A^TA) = \sigma(A^2)$$

<u>Thus:</u> They are only equal up to the sign.

6. $(2P)$ $A \in \mathbb{R}^{m\times n}$, then the first k principal components are given by

$$U_k := \underset{z\,\in\,\mathbb{R}^{m\times k}\ \text{orthogonal}}{\operatorname{argmin}} \|A - zz^TA\|_F^2.$$

7. $(2P)$ $A \in \mathbb{R}^{n\times n}$ positive definite $\quad \Rightarrow \quad \forall x \neq 0:\ x^T \underbrace{Ax}_{:=z} > 0$

$$0 < x^Tz = \underbrace{\cos(\alpha)}_{>0}\,\underbrace{\|x\|\|z\|}_{>0} \quad \Rightarrow \quad \cos(\alpha) > 0$$
$$\Rightarrow \quad \alpha \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right),\ |\alpha| < 90°$$