University of Cyprus
MAI613 - Research Methodologies and Professional Practices in AI
Lecturer: Stelios Timotheou

Exercises on Matrix Decompositions

Exercises

1. Solve Exercises 4.2, 4.4, 4.7, 4.8, 4.10 and 4.12 from the book:

- M.P. Deisenroth, A. A. Faisal, and C. S. Ong, "Mathematics for Machine Learning," Cambridge University Press, 2020. https://mml-book.github.io/book/mml-book.pdf

2a. Show that matrix **A** is positive definite.

$$\mathbf{A} = \begin{bmatrix} 25 & 15 & -5 \\ 15 & 18 & 0 \\ -5 & 0 & 11 \end{bmatrix}$$

2b. Find the Cholesky factor of **A**, **L**, such that $\mathbf{A} = \mathbf{L}\mathbf{L}^{\mathrm{T}}$.

Note: These exercises are intended for self-assessment purposes. Their solutions will be uploaded in one week from now.

4.2 Compute the following determinant efficiently:

$$\begin{bmatrix} 2 & 0 & 1 & 2 & 0 \\ 2 & -1 & 0 & 1 & 1 \\ 0 & 1 & 2 & 1 & 2 \\ -2 & 0 & 2 & -1 & 2 \\ 2 & 0 & 0 & 1 & 1 \end{bmatrix}.$$

This strategy shows the power of the methods we learned in this and the previous chapter. We can first apply Gaussian elimination to transform $\boldsymbol{A}$ into a triangular form, and then use the fact that the determinant of a triangular matrix equals the product of its diagonal elements.

$$\begin{vmatrix} 2 & 0 & 1 & 2 & 0 \\ 2 & -1 & 0 & 1 & 1 \\ 0 & 1 & 2 & 1 & 2 \\ -2 & 0 & 2 & -1 & 2 \\ 2 & 0 & 0 & 1 & 1 \end{vmatrix} = \begin{vmatrix} 2 & 0 & 1 & 2 & 0 \\ 0 & -1 & -1 & -1 & 1 \\ 0 & 1 & 2 & 1 & 2 \\ 0 & 0 & 3 & 1 & 2 \\ 0 & 0 & -1 & -1 & 1 \end{vmatrix} = \begin{vmatrix} 2 & 0 & 1 & 2 & 0 \\ 0 & -1 & -1 & -1 & 1 \\ 0 & 0 & 1 & 0 & 3 \\ 0 & 0 & 3 & 1 & 2 \\ 0 & 0 & -1 & -1 & 1 \end{vmatrix}$$

$$= \begin{vmatrix} 2 & 0 & 1 & 2 & 0 \\ 0 & -1 & -1 & -1 & 1 \\ 0 & 0 & 1 & 0 & 3 \\ 0 & 0 & 0 & 1 & -7 \\ 0 & 0 & 0 & -1 & 4 \end{vmatrix} = \begin{vmatrix} 2 & 0 & 1 & 2 & 0 \\ 0 & -1 & -1 & -1 & 1 \\ 0 & 0 & 1 & 0 & 3 \\ 0 & 0 & 0 & 1 & -7 \\ 0 & 0 & 0 & 0 & -3 \end{vmatrix} = 6\,.$$

Alternatively, we can apply the Laplace expansion and arrive at the same solution:

$$\begin{vmatrix} 2 & 0 & 1 & 2 & 0 \\ 2 & -1 & 0 & 1 & 1 \\ 0 & 1 & 2 & 1 & 2 \\ -2 & 0 & 2 & -1 & 2 \\ 2 & 0 & 0 & 1 & 1 \end{vmatrix} = \begin{vmatrix} 2 & 0 & 1 & 2 & 0 \\ 0 & -1 & -1 & -1 & 1 \\ 0 & 1 & 2 & 1 & 2 \\ 0 & 0 & 3 & 1 & 2 \\ 0 & 0 & -1 & -1 & 1 \end{vmatrix}$$

$$\overset{\text{1st col.}}{=} (-1)^{1+1} 2 \cdot \begin{vmatrix} -1 & -1 & -1 & 1 \\ 1 & 2 & 1 & 2 \\ 0 & 3 & 1 & 2 \\ 0 & -1 & -1 & 1 \end{vmatrix}\,.$$

If we now subtract the fourth row from the first row and multiply (-2) times the third column to the fourth column we obtain

$$2 \begin{vmatrix} -1 & 0 & 0 & 0 \\ 1 & 2 & 1 & 0 \\ 0 & 3 & 1 & 0 \\ 0 & -1 & -1 & 3 \end{vmatrix} \overset{\text{1st row}}{=} -2 \begin{vmatrix} 2 & 1 & 0 \\ 3 & 1 & 0 \\ -1 & -1 & 3 \end{vmatrix} \overset{\text{3rd col.}}{=} (-2) \cdot 3(-1)^{3+3} \cdot \begin{vmatrix} 2 & 1 \\ 3 & 1 \end{vmatrix} = 6\,.$$

4.4 Compute all eigenspaces of

$$A = \begin{bmatrix} 0 & -1 & 1 & 1 \\ -1 & 1 & -2 & 3 \\ 2 & -1 & 0 & 0 \\ 1 & -1 & 1 & 0 \end{bmatrix}.$$

(i) Characteristic polynomial:

$$p(\lambda) = \begin{vmatrix} -\lambda & -1 & 1 & 1 \\ -1 & 1-\lambda & -2 & 3 \\ 2 & -1 & -\lambda & 0 \\ 1 & -1 & 1 & -\lambda \end{vmatrix} = \begin{vmatrix} -\lambda & -1 & 1 & 1 \\ 0 & -\lambda & -1 & 3-\lambda \\ 0 & 1 & -2-\lambda & 2\lambda \\ 1 & -1 & 1 & -\lambda \end{vmatrix}$$

$$= \begin{vmatrix} -\lambda & -1-\lambda & 0 & 1 \\ 0 & -\lambda & -1-\lambda & 3-\lambda \\ 0 & 1 & -1-\lambda & 2\lambda \\ 1 & 0 & 0 & -\lambda \end{vmatrix}$$

$$= (-\lambda)\begin{vmatrix} -\lambda & -1-\lambda & 3-\lambda \\ 1 & -1-\lambda & 2\lambda \\ 0 & 0 & -\lambda \end{vmatrix} - \begin{vmatrix} -1-\lambda & 0 & 1 \\ -\lambda & -1-\lambda & 3-\lambda \\ 1 & -1-\lambda & 2\lambda \end{vmatrix}$$

$$= (-\lambda)^2\begin{vmatrix} -\lambda & -1-\lambda \\ 1 & -1-\lambda \end{vmatrix} - \begin{vmatrix} -1-\lambda & 0 & 1 \\ -\lambda & -1-\lambda & 3-\lambda \\ 1 & -1-\lambda & 2\lambda \end{vmatrix}$$

$$= (1+\lambda)^2(\lambda^2 - 3\lambda + 2) = (1+\lambda)^2(1-\lambda)(2-\lambda)$$

Therefore, the eigenvalues of A are $\lambda_1 = -1, \lambda_2 = 1, \lambda_3 = 2$.

(ii) The corresponding eigenspaces are the solutions of $(A - \lambda_i I)x = \mathbf{0}$, $i = 1, 2, 3$, and given by

$$E_{-1} = \text{span}[\begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix}], \quad E_1 = \text{span}[\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}], \quad E_2 = \text{span}[\begin{bmatrix} 1 \\ 0 \\ 1 \\ 1 \end{bmatrix}].$$

4.7 Are the following matrices diagonalizable? If yes, determine their diagonal form and a basis with respect to which the transformation matrices are diagonal. If no, give reasons why they are not diagonalizable.

a.

$$\boldsymbol{A} = \begin{bmatrix} 0 & 1 \\ -8 & 4 \end{bmatrix}$$

We determine the characteristic polynomial as

$$p(\lambda) = \det(\boldsymbol{A} - \lambda \boldsymbol{I}) = -\lambda(4-\lambda) + 8 = \lambda^2 - 4\lambda + 8\,.$$

The characteristic polynomial does not decompose into linear factors over $\mathbb{R}$ because the roots of $p(\lambda)$ are complex and given by $\lambda_{1,2} = 2 \pm \sqrt{-4}$. Since the characteristic polynomial does not decompose into linear factors, $\boldsymbol{A}$ cannot be diagonalized (over $\mathbb{R}$).

b.

$$\boldsymbol{A} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

(i) The characteristic polynomial is $p(\lambda) = \det(\boldsymbol{A} - \lambda I)$.

$$p(\lambda) = \begin{vmatrix} 1-\lambda & 1 & 1 \\ 1 & 1-\lambda & 1 \\ 1 & 1 & 1-\lambda \end{vmatrix} \overset{\text{subtr. } R_1 \text{ from } R_2, R_3}{=} \begin{vmatrix} 1-\lambda & 1 & 1 \\ \lambda & -\lambda & 0 \\ \lambda & 0 & -\lambda \end{vmatrix}$$

$$\overset{\text{develop last row}}{=} \lambda \begin{vmatrix} 1 & 1 \\ -\lambda & 0 \end{vmatrix} - \lambda \begin{vmatrix} 1-\lambda & 1 \\ \lambda & -\lambda \end{vmatrix}$$

$$= \lambda^2 + \lambda(\lambda(1-\lambda) + \lambda) = \lambda(-\lambda^2 + 3\lambda) = \lambda^2(\lambda - 3)\,.$$

Therefore, the roots of $p(\lambda)$ are 0 and 3 with algebraic multiplicities 2 and 1, respectively.

(ii) To determine whether $\boldsymbol{A}$ is diagonalizable, we need to show that the dimension of E_0 is 2 (because the dimension of E_3 is necessarily 1: an eigenspace has at least dimension 1 by definition, and its dimension cannot exceed the algebraic multiplicity of its associated eigenvalue).

Let us study $E_0 = \ker(\boldsymbol{A} - 0\boldsymbol{I})$:

$$\boldsymbol{A} - 0\boldsymbol{I} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \rightsquigarrow \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Here, $\dim E_0 = 2$, which is identical to the algebraic multiplicity of the eigenvalue 0 in the characteristic polynomial. Thus A is diagonalizable. Moreover, we can read from the reduced row echelon form that

$$E_0 = \mathrm{span}[\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}].$$

(iii) For E_3, we obtain

$$\boldsymbol{A} - 3\boldsymbol{I} = \begin{bmatrix} -2 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & -2 \end{bmatrix} \rightsquigarrow \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix},$$

which has rank 2, and, therefore (using the rank-nullity theorem),

E_3 has dimension 1 (it could not be anything else anyway, as justified above) and

$$E_1 = \text{span}[\begin{bmatrix}-1\\-1\\-1\end{bmatrix}].$$

(iv) Therefore, we can find a new basis P as the concatenation of the spanning vectors of the eigenspaces. If we identify the matrix

$$\boldsymbol{P} = \begin{bmatrix}-1 & 1 & 1\\-1 & -1 & 0\\-1 & 0 & -1\end{bmatrix},$$

whose columns are composed of the basis vectors of basis P, then our endomorphism will have the diagonal form: $\boldsymbol{D} = \text{diag}[3, 0, 0]$ with respect to this new basis. As a reminder, $\text{diag}[3, 0, 0]$ refers to the 3×3 diagonal matrix with $3, 0, 0$ as values on the diagonal. Note that the diagonal form is not unique and depends on the order of the eigenvectors in the new basis. For example, we can define another matrix $\boldsymbol{Q}$ composed of the same vectors as $\boldsymbol{P}$ but in a different order:

$$\boldsymbol{Q} = \begin{bmatrix}1 & -1 & 1\\-1 & -1 & 0\\0 & -1 & -1\end{bmatrix}.$$

If we use this matrix, our endomorphism would have another diagonal form: $\boldsymbol{D}' = \text{diag}[0, 3, 0]$. Sceptical students can check that $\boldsymbol{Q}^{-1}\boldsymbol{A}\boldsymbol{Q} = \boldsymbol{D}'$ and $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P} = \boldsymbol{D}$.

c.

$$\boldsymbol{A} = \begin{bmatrix}5 & 4 & 2 & 1\\0 & 1 & -1 & -1\\-1 & -1 & 3 & 0\\1 & 1 & -1 & 2\end{bmatrix}$$

$p(\lambda) = (\lambda - 1)(\lambda - 2)(\lambda - 4)^2$

$$E_1 = \text{span}[\begin{bmatrix}-1\\1\\0\\0\end{bmatrix}], \quad E_2 = \text{span}[\begin{bmatrix}1\\-1\\0\\1\end{bmatrix}], \quad E_4 = \text{span}[\begin{bmatrix}1\\0\\-1\\1\end{bmatrix}].$$

Here, we see that $\dim(E_4) = 1 \neq 2$ (which is the algebraic multiplicity of the eigenvalue 4). Therefore, $\boldsymbol{A}$ cannot be diagonalized.

d.

$$\boldsymbol{A} = \begin{bmatrix}5 & -6 & -6\\-1 & 4 & 2\\3 & -6 & -4\end{bmatrix}$$

(i) We compute the characteristic polynomial $p(\lambda) = \det(\boldsymbol{A} - \lambda\boldsymbol{I})$ as

$$
\begin{aligned}
p(\lambda) &= \begin{vmatrix} 5-\lambda & -6 & -6 \\ -1 & 4-\lambda & 2 \\ 3 & -6 & -4-\lambda \end{vmatrix} \\
&= (5-\lambda)(4-\lambda)(-4-\lambda) - 36 - 36 + 18(4-\lambda) \\
&\quad + 12(5-\lambda) - 6(-4-\lambda) \\
&= -\lambda^3 + 5\lambda^2 - 8\lambda + 4 = (1-\lambda)(2-\lambda)^2\,,
\end{aligned}
$$

where we used Sarrus rule. The characteristic polynomial decomposes into linear factors, and the eigenvalues are $\lambda_1 = 1, \lambda_2 = 2$ with (algebraic) multiplicity 1 and 2, respectively.

(ii) If the dimension of the eigenspaces are identical to multiplicity of the corresponding eigenvalues, the matrix is diagonalizable. The eigenspace dimension is the dimension of $\ker(\boldsymbol{A} - \lambda_i\boldsymbol{I})$, where λ_i are the eigenvalues (here: 1,2). For a simple check whether the matrices are diagonalizable, it is sufficient to compute the rank r_i of $\boldsymbol{A} - \lambda_i\boldsymbol{I}$ since the eigenspace dimension is $n - r_i$ (rank-nullity theorem).

Let us study E_2 and apply Gaussian elimination on $\boldsymbol{A} - 2\boldsymbol{I}$:

$$
\begin{bmatrix} 3 & -6 & -6 \\ -1 & 2 & 2 \\ 3 & -6 & -6 \end{bmatrix}\,. \tag{4.1}
$$

(iii) We can immediately see that the rank of this matrix is 1 since the first and third row are three times the second. Therefore, the eigenspace dimension is $\dim(E_2) = 3 - 1 = 2$, which corresponds to the algebraic multiplicity of the eigenvalue $\lambda = 2$ in $p(\lambda)$. Moreover, we know that the dimension of E_1 is 1 since it cannot exceed its algebraic multiplicity, and the dimension of an eigenspace is at least 1. Hence, $\boldsymbol{A}$ is diagonalizable.

(iv) The diagonal matrix is easy to determine since it just contains the eigenvalues (with corresponding multiplicities) on its diagonal:

$$
\boldsymbol{D} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}\,.
$$

(v) We need to determine a basis with respect to which the transformation matrix is diagonal. We know that the basis that consists of the eigenvectors has exactly this property. Therefore, we need to determine the eigenvectors for all eigenvalues. Remember that $\boldsymbol{x}$ is an eigenvector for an eigenvalue λ if they satisfy $\boldsymbol{A}\boldsymbol{x} = \lambda\boldsymbol{x} \iff (\boldsymbol{A} - \lambda\boldsymbol{I})\boldsymbol{x} = \boldsymbol{0}$. Therefore, we need to find the basis vectors of the eigenspaces E_1, E_2.

For $E_1 = \ker(\boldsymbol{A} - \boldsymbol{I})$ we obtain:

$$
\begin{bmatrix} 4 & -6 & -6 \\ -1 & 3 & 2 \\ 3 & -6 & -5 \end{bmatrix}
\begin{array}{l} +4R_2 \\ \\ +3R_2 \end{array}
\rightsquigarrow
\begin{bmatrix} 0 & 6 & 2 \\ -1 & 3 & 2 \\ 0 & 3 & 1 \end{bmatrix}
\begin{array}{l} \cdot(\frac{1}{6}) \\ \cdot(-1)|\text{swap with } R_1 \\ -\frac{1}{2}R_1 \end{array}
$$

$$\rightsquigarrow \begin{bmatrix} 1 & -3 & -2 \\ 0 & 1 & \frac{1}{3} \\ 0 & 0 & 0 \end{bmatrix} \begin{matrix} +3R_2 \\ \\ \\ \end{matrix} \rightsquigarrow \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & \frac{1}{3} \\ 0 & 0 & 0 \end{bmatrix}$$

The rank of this matrix is 2. Since $3 - 2 = 1$ it follows that $\dim(E_1) = 1$, which corresponds to the algebraic multiplicity of the eigenvalue $\lambda = 1$ in the characteristic polynomial.

(vi) From the reduced row echelon form we see that

$$E_1 = \mathrm{span}[\begin{bmatrix} 3 \\ -1 \\ 3 \end{bmatrix}],$$

and our first eigenvector is $[3, -1, 3]^\top$.

(vii) We proceed with determining a basis of E_2, which will give us the other two basis vectors that we need (remember that $\dim(E_2) = 2$). From (4.1), we (almost) immediately obtain the reduced row echelon form

$$\begin{bmatrix} 1 & -2 & -2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

and the corresponding eigenspace

$$E_2 = \mathrm{span}[\begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}].$$

(viii) Overall, an ordered basis with respect to which $\boldsymbol{A}$ has diagonal form $\boldsymbol{D}$ consists of all eigenvectors is

$$B = (\begin{bmatrix} 3 \\ -1 \\ 3 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}).$$

4.8 Find the SVD of the matrix

$$\boldsymbol{A} = \begin{bmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{bmatrix}.$$

$$\boldsymbol{A} = \underbrace{\begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}}_{=\boldsymbol{U}} \underbrace{\begin{bmatrix} 5 & 0 & 0 \\ 0 & 3 & 0 \end{bmatrix}}_{=\boldsymbol{\Sigma}} \underbrace{\begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{3\sqrt{2}} & -\frac{2}{3} \\ \frac{1}{\sqrt{2}} & \frac{1}{3\sqrt{2}} & \frac{2}{3} \\ 0 & -\frac{2\sqrt{2}}{3} & \frac{1}{3} \end{bmatrix}}_{=\boldsymbol{V}}$$

4.10 Find the rank-1 approximation of

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

To find the rank-1 approximation we apply the SVD to $\boldsymbol{A}$ (as in Exercise 4.7) to obtain

$$\boldsymbol{U} = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$$

$$\boldsymbol{V} = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{3\sqrt{2}} & -\frac{2}{3} \\ \frac{1}{\sqrt{2}} & \frac{1}{3\sqrt{2}} & \frac{2}{3} \\ 0 & -\frac{2\sqrt{2}}{3} & \frac{1}{3} \end{bmatrix}.$$

We apply the construction rule for rank-1 matrices

$$\boldsymbol{A}_i = \boldsymbol{u}_i \boldsymbol{v}_i^\top .$$

We use the largest singular value ($\sigma_1 = 5$, i.e., $i = 1$ and the first column vectors of the $\boldsymbol{U}$ and $\boldsymbol{V}$ matrices, respectively:

$$\boldsymbol{A}_1 = \boldsymbol{u}_1 \boldsymbol{v}_1^\top = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \end{bmatrix}.$$

To find the rank-1 approximation, we apply the SVD to $\boldsymbol{A}$ to obtain

$$\boldsymbol{U} = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$$

$$\boldsymbol{V} = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{3\sqrt{2}} & -\frac{2}{3} \\ \frac{1}{\sqrt{2}} & \frac{1}{3\sqrt{2}} & \frac{2}{3} \\ 0 & -\frac{2\sqrt{2}}{3} & \frac{1}{3} \end{bmatrix}.$$

We apply the construction rule for rank-1 matrices

$$\boldsymbol{A}_i = \boldsymbol{u}_i \boldsymbol{v}_i^\top .$$

We use the largest singular value ($\sigma_1 = 5$, i.e., $i = 1$) and therefore, the first column vectors of $\boldsymbol{U}$ and $\boldsymbol{V}$, respectively, which then yields

$$\boldsymbol{A}_1 = \boldsymbol{u}_1 \boldsymbol{v}_1^\top = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \end{bmatrix}.$$

4.12 Show that for $\boldsymbol{x} \neq \boldsymbol{0}$ Theorem 4.24 holds, i.e., show that

$$\max_{\boldsymbol{x}} \frac{\|\boldsymbol{A}\boldsymbol{x}\|_2}{\|\boldsymbol{x}\|_2} = \sigma_1 \,,$$

where σ_1 is the largest singular value of $\boldsymbol{A} \in \mathbb{R}^{m \times n}$.

(i) We compute the eigendecomposition of the symmetric matrix

$$\boldsymbol{A}^\top \boldsymbol{A} = \boldsymbol{P}\boldsymbol{D}\boldsymbol{P}^\top$$

for diagonal $\boldsymbol{D}$ and orthogonal $\boldsymbol{P}$. Since the columns of $\boldsymbol{P}$ are an ONB of $\mathbb{R}^n$, we can write every $\boldsymbol{y} = \boldsymbol{P}\boldsymbol{x}$ as a linear combination of the eigenvectors $\boldsymbol{p}_i$ so that

$$\boldsymbol{y} = \boldsymbol{P}\boldsymbol{x} = \sum_{i=1}^{n} x_i \boldsymbol{p}_i \,, \quad \boldsymbol{x} \in \mathbb{R} \,. \tag{4.8}$$

Moreover, since the orthogonal matrix $\boldsymbol{P}$ preserves lengths (see Section 3.4), we obtain

$$\|\boldsymbol{y}\|_2^2 = \|\boldsymbol{P}\boldsymbol{x}\|_2^2 = \|\boldsymbol{x}\|_2^2 = \sum_{i=1}^{n} x_i^2 \,. \tag{4.9}$$

(ii) Then,

$$\|\boldsymbol{A}\boldsymbol{x}\|_2^2 = \boldsymbol{x}^\top(\boldsymbol{P}\boldsymbol{D}\boldsymbol{P}^\top)\boldsymbol{x} = \boldsymbol{y}^\top \boldsymbol{D}\boldsymbol{y} = \left\langle \sum_{i=1}^{n} \sqrt{\lambda_i} x_i \boldsymbol{p}_i, \sum_{i=1}^{n} \sqrt{\lambda_i} x_i \boldsymbol{p}_i \right\rangle \,,$$

where we used $\langle \cdot, \cdot \rangle$ to denote the dot product.

(iii) The bilinearity of the dot product gives us

$$\|\boldsymbol{A}\boldsymbol{x}\|_2^2 = \sum_{i=1}^{n} \lambda_i \langle x_i \boldsymbol{p}_i, x_i \boldsymbol{p}_i \rangle = \sum_{i=1}^{n} \lambda_i x_i^2$$

where we exploited that the $\boldsymbol{p}_i$ are an ONB and $\boldsymbol{p}_i^\top \boldsymbol{p}_i = 1$.

(iv) With (4.8) we obtain

$$\|\boldsymbol{A}\boldsymbol{x}\|_2^2 \leqslant \big(\max_{1 \leqslant j \leqslant n} \lambda_j\big) \sum_{i=1}^{n} x_i^2 \overset{(4.9)}{=} \max_{1 \leqslant j \leqslant n} \lambda_j \, \|\boldsymbol{x}\|_2^2$$

so that

$$\frac{\|\boldsymbol{A}\boldsymbol{x}\|_2^2}{\|\boldsymbol{x}\|_2^2} \leqslant \max_{1 \leqslant j \leqslant n} \lambda_j \,,$$

where λ_j are the eigenvalues of $\boldsymbol{A}^\top \boldsymbol{A}$.

(v) Assuming the eigenvalues of $\boldsymbol{A}^\top \boldsymbol{A}$ are sorted in descending order, we get

$$\frac{\|\boldsymbol{A}\boldsymbol{x}\|_2}{\|\boldsymbol{x}\|_2} \leqslant \sqrt{\lambda_1} = \sigma_1 \,,$$

where σ_1 is the maximum singular value of $\boldsymbol{A}$.

2a. We can show that matrix A is positive definite if $x^T A x \geq 0$, for all x.

$$A = \begin{bmatrix} 25 & 15 & -5 \\ 15 & 18 & 0 \\ -5 & 0 & 11 \end{bmatrix}$$

$$x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

$$x^T A x = [x_1 \quad x_2 \quad x_3] \begin{bmatrix} 25 & 15 & -5 \\ 15 & 18 & 0 \\ -5 & 0 & 11 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

$$= \begin{bmatrix} 25x_1 + 15x_2 - 5x_3 \\ 15x_1 + 18x_2 + 0x_3 \\ -5x_1 + 0x_2 + 11x_3 \end{bmatrix}^T \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

$$\begin{aligned} = 25x_1^2 + 15x_1x_2 - 5x_1x_3 \\ +15x_1x_2 + 18x_2^2 \\ -5x_1x_3 + 11x_3^2 \end{aligned}$$

$$= 25x_1^2 + 18x_2^2 + 11x_3^2 + 30x_1x_2 - 10x_1x_3$$

We can prove that the above is always ≥ 0:

$$x^T A x = 25x_1^2 + 18x_2^2 + 11x_3^2 + 30x_1x_2 - 10x_1x_3$$

$$\begin{aligned} &= (25 - \alpha - \gamma)x_1^2 + (18 - \beta)x_2^2 + (11 - \delta)x_3^2 \\ &\quad + \left(\sqrt{\alpha}x_1 + \sqrt{\beta}x_2\right)^2 + \left(\sqrt{\gamma}x_1 - \sqrt{\delta}x_3\right)^2 \end{aligned}$$

$$\begin{aligned} &= (25 - \alpha - \gamma)x_1^2 + (18 - \beta)x_2^2 + (11 - \delta)x_3^2 \\ &\quad + \left(\alpha x_1^2 + \beta x_2^2 + 2\sqrt{\alpha\beta}x_1x_2\right) \\ &\quad + \left(\gamma x_1^2 + \delta x_3^2 - 2\sqrt{\gamma\delta}x_1x_3\right) \end{aligned}$$

with:

$$\alpha + \gamma \leq 25, \qquad \beta \leq 18, \qquad \delta \leq 11$$

$$2\sqrt{\alpha\beta} = 30, \qquad 2\sqrt{\gamma\delta} = 10$$

Thus,

$$\alpha + \gamma \leq 25, \qquad \beta \leq 18, \qquad \delta \leq 11$$

$$\alpha\beta = 225, \qquad \gamma\delta = 25$$

Which are valid for $\alpha = \beta = 15$, and $\gamma = \delta = 5$, resulting in:

$$\begin{aligned} x^T A x &= (25 - 15 - 5)x_1^2 + (18 - 15)x_2^2 + (11 - 5)x_3^2 \\ &\quad + \left(\sqrt{15}x_1 + \sqrt{15}x_2\right)^2 + \left(\sqrt{5}x_1 - \sqrt{5}x_3\right)^2 \\ &= 5x_1^2 + 3x_2^2 + 6x_3^2 + \left(\sqrt{15}x_1 + \sqrt{15}x_2\right)^2 + \left(\sqrt{5}x_1 - \sqrt{5}x_3\right)^2 \end{aligned}$$

Which is always ≥ 0 for any values of x_1, x_2, x_3. Also, it can only be $x^T A x = 0$ if and only if $x_1 = x_2 = x_3 = 0$. <u>Therefore A is a Positive Definite matrix.</u>

2b. To find the Cholesky factor L of A, such that $A = LL^T$:

$$A = \begin{bmatrix} 25 & 15 & -5 \\ 15 & 18 & 0 \\ -5 & 0 & 11 \end{bmatrix} = \begin{bmatrix} L_{11} & 0 & 0 \\ L_{12} & L_{22} & 0 \\ L_{13} & L_{23} & L_{33} \end{bmatrix} \begin{bmatrix} L_{11} & L_{12} & L_{13} \\ 0 & L_{22} & L_{23} \\ 0 & 0 & L_{33} \end{bmatrix}$$

$$= \begin{bmatrix} L_{11}^2 & L_{11}L_{12} & L_{11}L_{13} \\ L_{11}L_{12} & L_{12}^2 + L_{22}^2 & L_{12}L_{13} + L_{22}L_{23} \\ L_{11}L_{13} & L_{12}L_{13} + L_{22}L_{23} & L_{13}^2 + L_{23}^2 + L_{33}^2 \end{bmatrix}$$

We first calculate the first row/column of A:

$$\begin{bmatrix} \mathbf{25} & \mathbf{15} & \mathbf{-5} \\ \mathbf{15} & 18 & 0 \\ \mathbf{-5} & 0 & 11 \end{bmatrix} = \begin{bmatrix} 5 & 0 & 0 \\ 3 & L_{22} & 0 \\ -1 & L_{23} & L_{33} \end{bmatrix} \begin{bmatrix} 5 & 3 & -1 \\ 0 & L_{22} & L_{23} \\ 0 & 0 & L_{33} \end{bmatrix}$$

Then, we calculate the second row/column of A:

$$\begin{bmatrix} 25 & 15 & -5 \\ 15 & \mathbf{18} & \mathbf{0} \\ -5 & \mathbf{0} & 11 \end{bmatrix} = \begin{bmatrix} 5 & 0 & 0 \\ 3 & 3 & 0 \\ -1 & 1 & L_{33} \end{bmatrix} \begin{bmatrix} 5 & 3 & -1 \\ 0 & 3 & 1 \\ 0 & 0 & L_{33} \end{bmatrix}$$

Finally, we calculate the last row/column of A:

$$\begin{bmatrix} 25 & 15 & -5 \\ 15 & 18 & 0 \\ -5 & 0 & \mathbf{11} \end{bmatrix} = \begin{bmatrix} 5 & 0 & 0 \\ 3 & 3 & 0 \\ -1 & 1 & 3 \end{bmatrix} \begin{bmatrix} 5 & 3 & -1 \\ 0 & 3 & 1 \\ 0 & 0 & 3 \end{bmatrix}$$