TD — Linear Algebra for Machine Learning

Lebanese University - Faculty of Sciences
Dr. Youssef SALMAN

Exercise 1 — Determinant, Eigenvalues, Eigenvectors, and Diagonalization

Consider the matrix

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

1. Compute the determinant of A and verify if it is invertible.
2. Find all eigenvalues of A by solving the characteristic polynomial.
3. For each eigenvalue, determine a corresponding eigenvector.
4. Check whether A is diagonalizable, and if yes, find matrices P and D such that $A = PDP^{-1}$.

Solution Exercice 1

Consider

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

1) Determinant. To avoid fractions, set $B = 2A$:

$$B = \begin{pmatrix} 3 & 1 & -1 \\ -1 & 5 & 1 \\ -2 & 2 & 4 \end{pmatrix}, \qquad \det(B) = 8\ \det(A).$$

Compute $\det(B)$ by the 3×3 formula:

$$\det(B) = 3\begin{vmatrix} 5 & 1 \\ 2 & 4 \end{vmatrix} - 1\begin{vmatrix} -1 & 1 \\ -2 & 4 \end{vmatrix} + (-1)\begin{vmatrix} -1 & 5 \\ -2 & 2 \end{vmatrix}.$$

Therefore

$$\boxed{\det(A) = \frac{\det(B)}{8} = \frac{48}{8} = 6 \neq 0.}$$

2) Eigenvalues. Compute the characteristic polynomial $\chi_A(\lambda) = \det(\lambda I - A)$. Again, clear fractions by working with $M = 2(\lambda I - A)$:

$$\lambda I - A = \begin{pmatrix} \lambda - \frac{3}{2} & -\frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \lambda - \frac{5}{2} & -\frac{1}{2} \\ 1 & -1 & \lambda - 2 \end{pmatrix}, \quad M = 2(\lambda I - A) = \begin{pmatrix} 2\lambda - 3 & -1 & 1 \\ 1 & 2\lambda - 5 & -1 \\ 2 & -2 & 2\lambda - 4 \end{pmatrix}.$$

Since $\det(\lambda I - A) = \det(M)/8$,(note that $det(kA) = k^n det(A)$), compute $\det(M)$ by expansion:

$$\det(M) = (2\lambda-3)\Big((2\lambda-5)(2\lambda-4)-(-1)(-2)\Big)-(-1)\Big(1\cdot(2\lambda-4)-(-1)\cdot 2\Big)+1\Big(1\cdot(-2)-(2\lambda-5)\cdot 2\Big).$$

Inside:

$$(2\lambda - 5)(2\lambda - 4) = 4\lambda^2 - 18\lambda + 20, \quad \Rightarrow \quad (2\lambda - 5)(2\lambda - 4) - 2 = 4\lambda^2 - 18\lambda + 18,$$

$$1 \cdot (2\lambda - 4) - (-1) \cdot 2 = 2\lambda - 2, \qquad 1 \cdot (-2) - (2\lambda - 5) \cdot 2 = -2 - (4\lambda - 10) = 8 - 4\lambda.$$

Thus

$$\det(M) = (2\lambda - 3)(4\lambda^2 - 18\lambda + 18) + (2\lambda - 2) + (8 - 4\lambda) = 8\lambda^3 - 48\lambda^2 + 88\lambda - 48.$$

Therefore

$$\chi_A(\lambda) = \det(\lambda I - A) = \frac{\det(M)}{8} = \lambda^3 - 6\lambda^2 + 11\lambda - 6 = (\lambda - 1)(\lambda - 2)(\lambda - 3).$$

So the eigenvalues are

$$\boxed{\lambda_1 = 1, \quad \lambda_2 = 2, \quad \lambda_3 = 3.}$$

3) Eigenvectors. Solve $(A - \lambda I)x = 0$ for each λ.

For $\lambda = 1$: Solve $(A - I)x = 0$. One obtains the eigenvector

$$\boxed{v_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}} \quad \text{(any nonzero scalar multiple is valid).}$$

For $\lambda = 2$: Solve $(A - 2I)x = 0$. One obtains

$$\boxed{v_2 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}}.$$

For $\lambda = 3$: Solve $(A - 3I)x = 0$. One obtains

$$\boxed{v_3 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}}.$$

(These are easy to verify by direct multiplication $Av_i = \lambda_i v_i$.)

4) Diagonalization. Since A has three distinct eigenvalues, the eigenvectors v_1, v_2, v_3 are linearly independent, and A is diagonalizable. Let

$$P = \begin{bmatrix} v_1 & v_2 & v_3 \end{bmatrix} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix}, \qquad D = \operatorname{diag}(1, 2, 3).$$

Then

$$\boxed{A = PDP^{-1}.}$$

(Indeed, the columns of P are eigenvectors of A and D contains the corresponding eigenvalues.)

Exercice 2

Consider the shear matrix

$$S = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

1. Show that S has a single eigenvalue $\lambda = 1$ with only one independent eigenvector. Conclude that S is *not* diagonalizable.
2. Compute $S^\top S$ and its eigenvalues $\{\lambda_1, \lambda_2\}$; set the singular values $\sigma_i = \sqrt{\lambda_i}$.
3. Find the right singular vectors v_i as eigenvectors of $S^\top S$ and the left singular vectors $u_i = \frac{Sv_i}{\sigma_i}$.
4. Conclude the SVD $S = U\Sigma V^\top$ with $U = [u_1\ u_2]$, $V = [v_1\ v_2]$, $\Sigma = \operatorname{diag}(\sigma_1, \sigma_2)$.

Solution of Exercice 2

We study

$$S = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

1. **Eigenvalue(s), eigenvector(s), and non-diagonalizability.**

   Characteristic polynomial.

   $$\det(S - \lambda I) = \det \begin{pmatrix} 1-\lambda & 1 \\ 0 & 1-\lambda \end{pmatrix} = (1-\lambda)^2.$$

   Thus S has the single eigenvalue $\boxed{\lambda = 1}$ with algebraic multiplicity 2.

   Eigenvectors. Solve $(S - I)x = 0$:

   $$S - I = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \qquad \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 0 \implies y = 0, \text{ free } x.$$

   Hence the eigenspace is $\ker(S - I) = \operatorname{Span}\{(1, 0)^\top\}$, which is *one*-dimensional.

   Conclusion. The geometric multiplicity (1) is strictly less than the size (2), so $\boxed{S \text{ is not diagonalizable}}$ (over $\mathbb{R}$ or $\mathbb{C}$).

2. **Compute $S^\top S$ and its eigenvalues λ_1, λ_2 (singular values $\sigma_i = \sqrt{\lambda_i}$).**

$$S^\top S = \begin{pmatrix}1 & 0\\ 1 & 1\end{pmatrix}\begin{pmatrix}1 & 1\\ 0 & 1\end{pmatrix} = \begin{pmatrix}1 & 1\\ 1 & 2\end{pmatrix}.$$

<u>Eigenvalues of $S^\top S$.</u>

$$\det\begin{pmatrix}1-\lambda & 1\\ 1 & 2-\lambda\end{pmatrix} = (1-\lambda)(2-\lambda) - 1 = \lambda^2 - 3\lambda + 1 = 0.$$

Thus

$$\boxed{\lambda_{1,2} = \frac{3 \pm \sqrt{5}}{2}}.$$

By definition, the singular values are

$$\boxed{\sigma_1 = \sqrt{\lambda_1} = \sqrt{\frac{3+\sqrt{5}}{2}}, \qquad \sigma_2 = \sqrt{\lambda_2} = \sqrt{\frac{3-\sqrt{5}}{2}}}.$$

(Notice $\sigma_1 > \sigma_2 > 0$.)

3. **Right/left singular vectors.**

<u>Right singular vectors v_i.</u> They are unit eigenvectors of $S^\top S$.

For $\lambda_1 = \frac{3+\sqrt{5}}{2}$, solve $(S^\top S - \lambda_1 I)v_1 = 0$:

$$\begin{pmatrix}1-\lambda_1 & 1\\ 1 & 2-\lambda_1\end{pmatrix}\begin{pmatrix}x\\ y\end{pmatrix} = 0 \quad\Rightarrow\quad y = (\lambda_1 - 1)\,x.$$

A (non-unit) eigenvector is $v_1^{\text{raw}} = (1,\ \lambda_1 - 1)^\top$. Its squared norm is

$$\|v_1^{\text{raw}}\|^2 = 1 + (\lambda_1 - 1)^2 = 1 + \left(\tfrac{1+\sqrt{5}}{2}\right)^2 = \frac{5+\sqrt{5}}{2}.$$

Hence the *unit* vector is

$$\boxed{v_1 = \frac{1}{\sqrt{\frac{5+\sqrt{5}}{2}}}\begin{pmatrix}1\\ \frac{1+\sqrt{5}}{2}\end{pmatrix}}.$$

For $\lambda_2 = \frac{3-\sqrt{5}}{2}$, similarly $y = (\lambda_2 - 1)x = \frac{1-\sqrt{5}}{2}x$, so a (non-unit) eigenvector is $v_2^{\text{raw}} = (1,\ \frac{1-\sqrt{5}}{2})^\top$ with

$$\|v_2^{\text{raw}}\|^2 = 1 + \left(\tfrac{1-\sqrt{5}}{2}\right)^2 = \frac{5-\sqrt{5}}{2}.$$

Thus the *unit* vector is

$$\boxed{v_2 = \frac{1}{\sqrt{\frac{5-\sqrt{5}}{2}}}\begin{pmatrix}1\\ \frac{1-\sqrt{5}}{2}\end{pmatrix}}.$$

(As $S^\top S$ is symmetric, $v_1 \perp v_2$.)

Left singular vectors u_i. By definition,

$$u_i = \frac{S\,v_i}{\sigma_i}.$$

Using the raw (non-unit) v_i^{raw} is convenient to see directions:

$$S\,v_1^{\mathrm{raw}} = \begin{pmatrix}1 & 1\\ 0 & 1\end{pmatrix}\begin{pmatrix}1\\ \frac{1+\sqrt{5}}{2}\end{pmatrix} = \begin{pmatrix}1+\frac{1+\sqrt{5}}{2}\\ \frac{1+\sqrt{5}}{2}\end{pmatrix} = \begin{pmatrix}\frac{3+\sqrt{5}}{2}\\ \frac{1+\sqrt{5}}{2}\end{pmatrix},$$

which is proportional to $\left(\frac{1+\sqrt{5}}{2},\,1\right)^\top$. A unit choice is therefore

$$\boxed{u_1 = \frac{1}{\sqrt{\frac{5+\sqrt{5}}{2}}}\begin{pmatrix}\frac{1+\sqrt{5}}{2}\\ 1\end{pmatrix}.}$$

Likewise,

$$S\,v_2^{\mathrm{raw}} = \begin{pmatrix}1 & 1\\ 0 & 1\end{pmatrix}\begin{pmatrix}1\\ \frac{1-\sqrt{5}}{2}\end{pmatrix} = \begin{pmatrix}1+\frac{1-\sqrt{5}}{2}\\ \frac{1-\sqrt{5}}{2}\end{pmatrix} = \begin{pmatrix}\frac{3-\sqrt{5}}{2}\\ \frac{1-\sqrt{5}}{2}\end{pmatrix},$$

which is proportional to $\left(\frac{1-\sqrt{5}}{2},\,-1\right)^\top$. Normalize to get

$$\boxed{u_2 = \frac{1}{\sqrt{\frac{5+\sqrt{5}}{2}}}\begin{pmatrix}\frac{1-\sqrt{5}}{2}\\ -1\end{pmatrix}.}$$

(Indeed $u_1 \perp u_2$, and $\|u_i\| = 1$.)

4. **Assemble the SVD.**

   Collect $U = [u_1\ u_2]$, $V = [v_1\ v_2]$, and

$$\Sigma = \begin{pmatrix}\sigma_1 & 0\\ 0 & \sigma_2\end{pmatrix} = \begin{pmatrix}\sqrt{\frac{3+\sqrt{5}}{2}} & 0\\ 0 & \sqrt{\frac{3-\sqrt{5}}{2}}\end{pmatrix}.$$

   Then

$$\boxed{S = U\,\Sigma\,V^\top}$$

   with U, V orthogonal and $\sigma_1 \ge \sigma_2 > 0$. (You can verify numerically that $U^\top U = V^\top V = I$ and $U\Sigma V^\top = S$.)

Remark (Golden-ratio identities for intuition). Let $\varphi = \frac{1+\sqrt{5}}{2}$. Then

$$\lambda_1 = \varphi^2,\quad \lambda_2 = \varphi^{-2},\qquad \sigma_1 = \varphi,\quad \sigma_2 = \varphi^{-1},$$

and the (non-unit) singular directions can be written as

$$v_1 \propto \begin{pmatrix}1\\ \varphi\end{pmatrix},\quad v_2 \propto \begin{pmatrix}1\\ -\varphi^{-1}\end{pmatrix},\qquad u_1 \propto \begin{pmatrix}\varphi\\ 1\end{pmatrix},\quad u_2 \propto \begin{pmatrix}-\varphi\\ 1\end{pmatrix}.$$

This makes the shear's geometry particularly transparent.

Exercise 3 — PCA on a tiny 2D dataset

Given the four points (rows are samples)

$$X = \begin{pmatrix} 2 & 0 \\ 0 & 2 \\ 3 & 1 \\ 1 & 3 \end{pmatrix} \in \mathbb{R}^{4\times 2}.$$

1. Center the data: compute the column mean μ and $Z = X - \mathbf{1}\mu^\top$.
2. Compute the sample covariance matrix $C = \frac{1}{n-1} Z^\top Z$.
3. Compute the eigenvalues/eigenvectors of C (principal components).
4. Order components by decreasing eigenvalue, and give the explained variance ratio for $k = 1$ and $k = 2$.
5. Project the centered data on the first principal component: $T_1 = Zv_1$.

Exercise 3

Centering. $\mu = \dfrac{1}{4}(2+0+3+1,\ 0+2+1+3)^\top = (1.5,\ 1.5)^\top$.

$$Z = X - \mathbf{1}\mu^\top = \begin{pmatrix} 0.5 & -1.5 \\ -1.5 & 0.5 \\ 1.5 & -0.5 \\ -0.5 & 1.5 \end{pmatrix}.$$

Covariance.

$$C = \frac{1}{3} Z^\top Z = \begin{pmatrix} \frac{5}{3} & -\frac{3}{3} \\ -\frac{3}{3} & \frac{5}{3} \end{pmatrix} = \begin{pmatrix} 1.666\ldots & -1 \\ -1 & 1.666\ldots \end{pmatrix}.$$

Eigenpairs (PCs). The eigenvalues are

$$\boxed{\lambda_1 = \tfrac{8}{3} \approx 2.6667, \quad \lambda_2 = \tfrac{2}{3} \approx 0.6667}.$$

Unit eigenvectors can be chosen as

$$v_1 = \frac{1}{\sqrt{2}}\begin{pmatrix} -1 \\ 1 \end{pmatrix}, \qquad v_2 = \frac{1}{\sqrt{2}}\begin{pmatrix} -1 \\ -1 \end{pmatrix}.$$

Explained variance.

$$\mathrm{EVR}(k) = \frac{\sum_{i=1}^{k} \lambda_i}{\lambda_1 + \lambda_2}, \quad \mathrm{EVR}(1) = \frac{8/3}{10/3} = 0.8, \quad \mathrm{EVR}(2) = 1.$$

Projection on PC1. Scores on the first PC: $T_1 = Z\,v_1$ (a 4×1 vector of coordinates along v_1).