# ASSIGNMENT 1

Mathematics for Machine Learning 811 — 17 January 2022

---

Unless specified otherwise, the problems in this assignment are intended for pen-and-paper and your submission should include steps. I do encourage you to verify your answers in Python.

What you submit must be your own work, and sources other than the lecture material must be cited. Remember to append a signed plagiarism declaration to your submission.

---

**1.** Using Gaussian elimination, find all solutions of $\boldsymbol{Ax} = \boldsymbol{b}$ with

$$\boldsymbol{A} = \begin{bmatrix} 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 \end{bmatrix}, \quad \boldsymbol{b} = \begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix}$$

**2.** Determine the inverses of the following matrices if possible:

$$\textbf{(a)} \quad \begin{bmatrix} 2 & 3 & 4 \\ 3 & 4 & 5 \\ 4 & 5 & 6 \end{bmatrix} \qquad \textbf{(b)} \quad \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix}$$

**3.** **(a)** Write $\boldsymbol{y} = \begin{bmatrix} 1 \\ -2 \\ 5 \end{bmatrix}$ as a linear combination of $\boldsymbol{x}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$, $\boldsymbol{x}_2 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$, $\boldsymbol{x}_3 = \begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix}$.

**(b)** Do the vectors $\boldsymbol{x}_1$, $\boldsymbol{x}_2$, $\boldsymbol{x}_3$ in part (a) form a generating set of $\mathbb{R}^3$?

**4.** Consider two subspaces $U_1$ and $U_2$, where $U_1$ is spanned by the columns of $\boldsymbol{A}_1$ and $U_2$ is spanned by the columns of $\boldsymbol{A}_2$ with

$$\boldsymbol{A}_1 = \begin{bmatrix} 1 & 0 & 1 \\ 1 & -2 & -1 \\ 2 & 1 & 3 \\ 1 & 0 & 1 \end{bmatrix}, \quad \boldsymbol{A}_2 = \begin{bmatrix} 3 & -3 & 0 \\ 1 & 2 & 3 \\ 7 & -5 & 2 \\ 3 & -1 & 2 \end{bmatrix}$$

**(a)** Determine the dimensions of $U_1$ and $U_2$.

**(b)** Determine bases of $U_1$ and $U_2$.

**(c)** Determine a basis of $U_1 \cap U_2$, that is, the intersection of $U_1$ and $U_2$.

**5.** Consider the linear mapping $\Phi : \mathbb{R}^3 \to \mathbb{R}^4$ with

$$\Phi\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = \begin{bmatrix} 3x_1 + 2x_2 + x_3 \\ x_1 + x_2 + x_3 \\ x_1 - 3x_2 \\ 2x_1 + 3x_2 + x_3 \end{bmatrix}$$

**(a)** Find the transformation matrix $\boldsymbol{A}_\Phi$.

**(b)** Determine $\text{rk}(\boldsymbol{A}_\Phi)$.

**(c)** Compute the kernel and image of $\Phi$. What are $\dim(\ker(\Phi))$ and $\dim(\text{Im}(\Phi))$?

**6.** Implement the modified Richardson iteration method for solving linear systems, in Python. Read up on how the method's one free parameter can be chosen optimally in terms of the eigenvalues of the coefficient matrix, and demonstrate how your code solves one or two example systems of your own choosing.

**7.** Using the Gram-Schmidt method, turn the basis $B = (\boldsymbol{b}_1, \boldsymbol{b}_2)$ of a 2-dimensional subspace $U \subseteq \mathbb{R}^3$ into an orthonormal basis $C = (\boldsymbol{c}_1, \boldsymbol{c}_2)$ of $U$, where

$$\boldsymbol{b}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \quad \boldsymbol{b}_2 = \begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix}$$

**8.** Let $x_1, \ldots, x_n > 0$ be $n$ positive real numbers so that $x_1 + \ldots + x_n = 1$. Use the Cauchy-Schwarz inequality and show that

$$\textbf{(a)} \quad \sum_{i=1}^{n} x_i^2 \geq \frac{1}{n} \qquad \sum_{i=1}^{n} \frac{1}{x_i} \geq n^2$$

**Hint:** think about the dot product in $\mathbb{R}^n$, and choose specific vectors $\boldsymbol{x}, \boldsymbol{y}$ for the Cauchy-Schwarz inequality.

**9.** Compute the eigenspaces of $\boldsymbol{A} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$.

**10.** Diagonalisability and invertibility are unrelated. Determine for each of the following matrices whether it is diagonalisable and/or invertible:

$$\textbf{(a)} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \qquad \textbf{(b)} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \qquad \textbf{(c)} \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \qquad \textbf{(d)} \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$

**11.** **(a)** Find the SVD of $\boldsymbol{A} = \begin{bmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{bmatrix}$.

**(b)** Find a good rank-1 approximation of the matrix $\boldsymbol{A}$ in part (a).

**12.** **(a)** Let $\boldsymbol{A}$ be an $m \times n$ greyscale image of your own choosing. Compute the SVD of $\boldsymbol{A}$ in Python, plot its singular values and observe that they decrease quickly. Approximate $\boldsymbol{A}$ with a rank-$k$ matrix $\widehat{\boldsymbol{A}}_k$, and display resulting images for a few values of $k$.

**(b)** This approximation seems like effective image compression. What information is needed to construct $\widehat{\boldsymbol{A}}_k$? In terms of $m$ and $n$, what is the largest value of $k$ such that the storage required to build $\widehat{\boldsymbol{A}}_k$ is less than the $mn$ 8-bit integers in $\boldsymbol{A}$? Considering results from choosing $k$ less than this upper limit, would you say low-rank approximation offers effective image compression?

**13.** Determine the derivative $f'(x)$ of the logistic sigmoid $f(x) = \dfrac{1}{1 + \exp(-x)}$.

**14.** Compute the Taylor polynomials $T_n$, $n = 0, 1, 2, 3$ of $f(x) = \sin(x) + \cos(x)$ at $x_0 = 0$. Plot them in Python on the same set of axes, along with $f(x)$, for $x \in [-\pi, \pi]$.

**15.** Find the derivatives $\mathrm{d}f/\mathrm{d}\boldsymbol{x}$ of the following function by using the chain rule. Provide the dimensions of every single partial derivative.

$$f(z) = \log(1 + z), \quad z = \boldsymbol{x}^\top \boldsymbol{x}, \quad \boldsymbol{x} \in \mathbb{R}^D$$

---