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

Exercises on Analytic Geometry

Exercises

Solve Exercises 3.4, 3.5, 3.6, 3.8, 3.9 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. Compute the inner product of real functions $f(x)=2x$ and $g(x)=\sin(x)$ over the closed interval $[-\pi, \pi]$.
2b. Compute the length of the functions f and g.
2c. Compute the angle between the two functions.
2d. Show that the Cauchy-Schwarz inequality holds true.

Note: These exercises are intended for self-assessment purposes.

Solutions:

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

Exercises on Analytic Geometry

Exercises

Solve Exercises 3.4, 3.5, 3.6, 3.8, 3.9 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. Compute the inner product of real functions $f(x)=2x$ and $g(x)=\sin(x)$ over the closed interval $[-\pi, \pi]$.
2b. Compute the length of the functions f and g.
2c. Compute the angle between the two functions.
2d. Show that the Cauchy-Schwarz inequality holds true.

Note: These exercises are intended for self-assessment purposes.

Solutions:

3.4 Compute the angle between

$$\boldsymbol{x} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \quad \boldsymbol{y} = \begin{bmatrix} -1 \\ -1 \end{bmatrix}$$

using

a. $\langle \boldsymbol{x}, \boldsymbol{y} \rangle := \boldsymbol{x}^\top \boldsymbol{y}$

b. $\langle \boldsymbol{x}, \boldsymbol{y} \rangle := \boldsymbol{x}^\top \boldsymbol{B} \boldsymbol{y}, \quad \boldsymbol{B} := \begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix}$

It holds that

$$\cos \omega = \frac{\langle \boldsymbol{x}, \boldsymbol{y} \rangle}{\|\boldsymbol{x}\| \, \|\boldsymbol{y}\|},$$

where ω is the angle between $\boldsymbol{x}$ and $\boldsymbol{y}$.

a.

$$\cos \omega = \frac{-3}{\sqrt{5}\sqrt{2}} = -\frac{3}{\sqrt{10}} \approx 2.82 \, \mathrm{rad} = 161.5^\circ$$

b.

$$\cos \omega = \frac{\boldsymbol{x}^\top \boldsymbol{B} \boldsymbol{y}}{\sqrt{\boldsymbol{x}^\top \boldsymbol{B} \boldsymbol{y}} \sqrt{\boldsymbol{y}^\top \boldsymbol{B} \boldsymbol{y}}} = \frac{-11}{\sqrt{18}\sqrt{7}} = -\frac{11}{126} \approx 1.66 \, \mathrm{rad} = 95^\circ$$

3.5 Consider the Euclidean vector space $\mathbb{R}^5$ with the dot product. A subspace $U \subseteq \mathbb{R}^5$ and $\boldsymbol{x} \in \mathbb{R}^5$ are given by

$$U = \text{span}[\begin{bmatrix}0\\-1\\2\\0\\2\end{bmatrix}, \begin{bmatrix}1\\-3\\1\\-1\\2\end{bmatrix}, \begin{bmatrix}-3\\4\\1\\2\\1\end{bmatrix}, \begin{bmatrix}-1\\-3\\5\\0\\7\end{bmatrix}], \quad \boldsymbol{x} = \begin{bmatrix}-1\\-9\\-1\\4\\1\end{bmatrix}$$

a. Determine the orthogonal projection $\pi_U(\boldsymbol{x})$ of $\boldsymbol{x}$ onto U

First, we determine a basis of U. Writing the spanning vectors as the columns of a matrix $\boldsymbol{A}$, we use Gaussian elimination to bring $\boldsymbol{A}$ into (reduced) row echelon form:

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

From here, we see that the first three columns are pivot columns, i.e., the first three vectors in the generating set of U form a basis of U:

$$U = \text{span}[\begin{bmatrix}0\\-1\\2\\0\\2\end{bmatrix}, \begin{bmatrix}1\\-3\\1\\-1\\2\end{bmatrix}, \begin{bmatrix}-3\\4\\1\\2\\1\end{bmatrix}].$$

Now, we define

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

where we define three basis vectors $\boldsymbol{b}_i$ of U as the columns of $\boldsymbol{B}$ for $1 \leqslant i \leqslant 3$.

We know that the projection of $\boldsymbol{x}$ on U exists and we define $\boldsymbol{p} := \pi_U(\boldsymbol{x})$. Moreover, we know that $\boldsymbol{p} \in U$. We define $\boldsymbol{\lambda} := [\lambda_1, \lambda_2, \lambda_3]^\top \in \mathbb{R}^3$, such that $\boldsymbol{p}$ can be written $\boldsymbol{p} = \sum_{i=1}^3 \lambda_i \boldsymbol{b}_i = \boldsymbol{B\lambda}$.

As $\boldsymbol{p}$ is the orthogonal projection of $\boldsymbol{x}$ onto U, then $\boldsymbol{x} - \boldsymbol{p}$ is orthogonal to all the basis vectors of U, so that

$$\boldsymbol{B}^\top(\boldsymbol{x} - \boldsymbol{B\lambda}) = \boldsymbol{0}.$$

Therefore,

$$\boldsymbol{B}^\top\boldsymbol{B\lambda} = \boldsymbol{B}^\top\boldsymbol{x}.$$

Solving in $\boldsymbol{\lambda}$ the inhomogeneous system $\boldsymbol{B}^\top\boldsymbol{B\lambda} = \boldsymbol{B}^\top\boldsymbol{x}$ gives us a single solution

$$\boldsymbol{\lambda} = \begin{bmatrix} -3 \\ 4 \\ 1 \end{bmatrix}$$

and, therefore, the desired projection

$$\boldsymbol{p} = \boldsymbol{B\lambda} = \begin{bmatrix} 1 \\ -5 \\ -1 \\ -2 \\ 3 \end{bmatrix} \in U\,.$$

b. Determine the distance $d(\boldsymbol{x}, U)$
The distance is simply the length of $\boldsymbol{x} - \boldsymbol{p}$:

$$\|\boldsymbol{x} - \boldsymbol{p}\| = \left\| \begin{bmatrix} 2 \\ 4 \\ 0 \\ -6 \\ 2 \end{bmatrix} \right\| = \sqrt{60}$$

3.6 Consider $\mathbb{R}^3$ with the inner product

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle := \boldsymbol{x}^\top \begin{bmatrix} 2 & 1 & 0 \\ 1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix} \boldsymbol{y}\,.$$

Furthermore, we define $\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3$ as the standard/canonical basis in $\mathbb{R}^3$.

a. Determine the orthogonal projection $\pi_U(\boldsymbol{e}_2)$ of $\boldsymbol{e}_2$ onto

$$U = \operatorname{span}[\boldsymbol{e}_1, \boldsymbol{e}_3]\,.$$

Hint: Orthogonality is defined through the inner product.
Let $\boldsymbol{p} = \pi_U(\boldsymbol{e}_2)$. As $\boldsymbol{p} \in U$, we can define $\Lambda = (\lambda_1, \lambda_3) \in \mathbb{R}^2$ such that $\boldsymbol{p}$ can be written $\boldsymbol{p} = U\Lambda$. In fact, $\boldsymbol{p}$ becomes $\boldsymbol{p} = \lambda_1 \boldsymbol{e}_1 + \lambda_3 \boldsymbol{e}_3 = [\lambda_1, 0, \lambda_3]^\top$ expressed in the canonical basis.
Now, we know by orthogonal projection that

$$\begin{aligned} \boldsymbol{p} = \pi_U(\boldsymbol{e}_2) &\implies (\boldsymbol{p} - \boldsymbol{e}_2) \perp U \\ &\implies \begin{bmatrix} \langle \boldsymbol{p} - \boldsymbol{e}_2, \boldsymbol{e}_1 \rangle \\ \langle \boldsymbol{p} - \boldsymbol{e}_2, \boldsymbol{e}_3 \rangle \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \\ &\implies \begin{bmatrix} \langle \boldsymbol{p}, \boldsymbol{e}_1 \rangle - \langle \boldsymbol{e}_2, \boldsymbol{e}_1 \rangle \\ \langle \boldsymbol{p}, \boldsymbol{e}_3 \rangle - \langle \boldsymbol{e}_2, \boldsymbol{e}_3 \rangle \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \end{aligned}$$

We compute the individual components as

$$\langle \boldsymbol{p}, \boldsymbol{e}_1 \rangle = \begin{bmatrix} \lambda_1 & 0 & \lambda_3 \end{bmatrix} \begin{bmatrix} 2 & 1 & 0 \\ 1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = 2\lambda_1$$

$$\langle \boldsymbol{p}, \boldsymbol{e}_3 \rangle = \begin{bmatrix} \lambda_1 & 0 & \lambda_3 \end{bmatrix} \begin{bmatrix} 2 & 1 & 0 \\ 1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = 2\lambda_3$$

$$\langle \boldsymbol{e}_2, \boldsymbol{e}_1 \rangle = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \begin{bmatrix} 2 & 1 & 0 \\ 1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = 1$$

$$\langle \boldsymbol{e}_2, \boldsymbol{e}_3 \rangle = \begin{bmatrix} 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 2 & 1 & 0 \\ 1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = -1$$

This now leads to the inhomogeneous linear equation system

$$\begin{aligned} 2\lambda_1 &= 1 \\ 2\lambda_3 &= -1 \end{aligned}$$

This immediately gives the coordinates of the projection as

$$\pi_U(\boldsymbol{e}_2) = \begin{bmatrix} \frac{1}{2} \\ 0 \\ -\frac{1}{2} \end{bmatrix}$$

b. Compute the distance $d(\boldsymbol{e}_2, U)$.
The distance of $d(\boldsymbol{e}_2, U)$ is the distance between $\boldsymbol{e}_2$ and its orthogonal projection $\boldsymbol{p} = \pi_U(\boldsymbol{e}_2)$ onto U. Therefore,

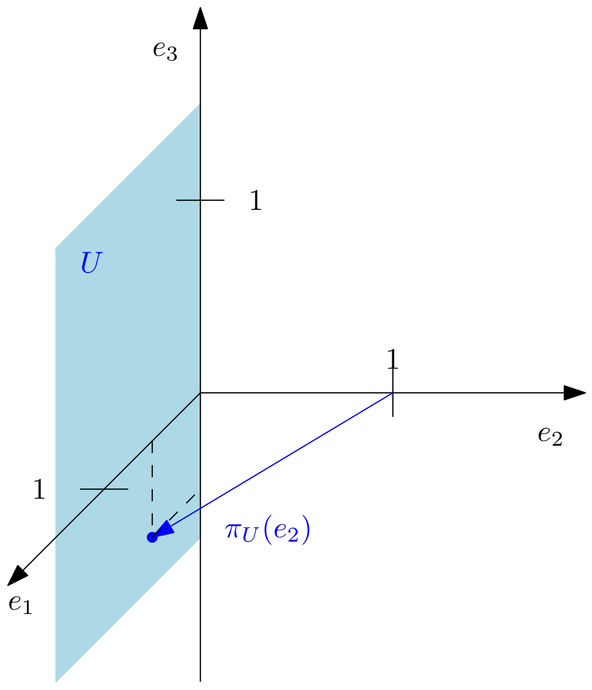
$$d(\boldsymbol{e}_2, U) = \sqrt{\langle \boldsymbol{p} - \boldsymbol{e}_2, \boldsymbol{p} - \boldsymbol{e}_2 \rangle^2}\,.$$

However,

$$\langle \boldsymbol{p} - \boldsymbol{e}_2, \boldsymbol{p} - \boldsymbol{e}_2 \rangle = \begin{bmatrix} \frac{1}{2} & -1 & -\frac{1}{2} \end{bmatrix} \begin{bmatrix} 2 & 1 & 0 \\ 1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix} \begin{bmatrix} \frac{1}{2} \\ -1 \\ -\frac{1}{2} \end{bmatrix} = 1\,,$$

which yields $d(\boldsymbol{e}_2, U) = \sqrt{\langle \boldsymbol{p} - \boldsymbol{e}_2, \boldsymbol{p} - \boldsymbol{e}_2 \rangle} = 1$

c. Draw the scenario: standard basis vectors and $\pi_U(\boldsymbol{e}_2)$
See Figure 3.1.



ure 3.1
jection $\pi_U(\boldsymbol{e}_2)$.

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

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

We start by normalizing $\boldsymbol{b}_1$

$$\boldsymbol{c}_1 := \frac{\boldsymbol{b}_1}{\|\boldsymbol{b}_1\|} = \frac{1}{\sqrt{3}}\begin{bmatrix}1\\1\\1\end{bmatrix}. \tag{3.1}$$

To get $\boldsymbol{c}_2$, we project $\boldsymbol{b}_2$ onto the subspace spanned by $\boldsymbol{c}_1$. This gives us (since $\|\boldsymbol{c}_1 = 1\|$)

$$\boldsymbol{c}_1^\top \boldsymbol{b}_2 \boldsymbol{c}_1 = \frac{1}{3}\begin{bmatrix}1\\1\\1\end{bmatrix} \in U\,.$$

By subtracting this projection (a multiple of $\boldsymbol{c}_1$) from $\boldsymbol{b}_2$, we get a vector that is orthogonal to $\boldsymbol{c}_1$:

$$\begin{bmatrix}-1\\2\\0\end{bmatrix} - \frac{1}{3}\begin{bmatrix}1\\1\\1\end{bmatrix} = \frac{1}{3}\begin{bmatrix}-4\\5\\-1\end{bmatrix} = \frac{1}{3}(-\boldsymbol{b}_1 + 3\boldsymbol{b}_2) \in U\,.$$

Normalizing $\tilde{\boldsymbol{c}}_2$ yields

$$\boldsymbol{c}_2 = \frac{\tilde{\boldsymbol{c}}_2}{\|\tilde{\boldsymbol{c}}_2\|} = \frac{3}{\sqrt{42}}\begin{bmatrix}-4\\5\\-1\end{bmatrix}.$$

We see that $\boldsymbol{c}_1 \perp \boldsymbol{c}_2$ and that $\|\boldsymbol{c}_1\| = 1 = \|\boldsymbol{c}_2\|$. Moreover, $\boldsymbol{c}_1, \boldsymbol{c}_2 \in U$ it follows that $(\boldsymbol{c}_1, \boldsymbol{c}_2)$ are a basis of U.

3.9 Let $n \in \mathbb{N}$ and let $x_1, \ldots, x_n > 0$ be n positive real numbers so that $x_1 + \cdots + x_n = 1$. Use the Cauchy-Schwarz inequality and show that

a. $\sum_{i=1}^n x_i^2 \geqslant \frac{1}{n}$
b. $\sum_{i=1}^n \frac{1}{x_i} \geqslant n^2$

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

Recall Cauchy-Schwarz inequality expressed with the dot product in $\mathbb{R}^n$. Let $\boldsymbol{x} = [x_1, \ldots, x_n]^\top$ and $\boldsymbol{y} = [y_1, \ldots, y_n]^\top$ be two vectors of $\mathbb{R}^n$. Cauchy-Schwarz tells us that

$$\langle \boldsymbol{x}, \boldsymbol{y}\rangle^2 \leqslant \langle \boldsymbol{x}, \boldsymbol{x}\rangle \cdot \langle \boldsymbol{y}, \boldsymbol{y}\rangle\,,$$

which, applied with the dot product in $\mathbb{R}^n$, can be rephrased as

$$\left(\sum_{i=1}^n x_i y_i\right)^2 \leqslant \left(\sum_{i=1}^n x_i^2\right) \cdot \left(\sum_{i=1}^n y_i^2\right).$$

a. Consider $\boldsymbol{x} = [x_1, \ldots, x_n]^\top$ as defined in the question. Let us choose $\boldsymbol{y} = [1, \ldots, 1]^\top$. Then, the Cauchy-Schwarz inequality becomes

$$\left(\sum_{i=1}^n x_i \cdot 1\right)^2 \leqslant \left(\sum_{i=1}^n x_i^2\right) \cdot \left(\sum_{i=1}^n 1^2\right)$$

and thus

$$1 \leqslant \left(\sum_{i=1}^n x_i^2\right) \cdot n\,,$$

which yields the expected result.

b. Let us now choose both vectors differently to obtain the expected result. Let $\boldsymbol{x} = [\frac{1}{\sqrt{x_1}}, \ldots, \frac{1}{\sqrt{x_n}}]^\top$ and $\boldsymbol{y} = [\sqrt{x_1}, \cdots, \sqrt{n}]^\top$. Note that our choice is legal since all x_i and y_i are strictly positive. The Cauchy-Schwarz inequality now becomes

$$\left(\sum_{i=1}^{n} \frac{1}{\sqrt{x_i}} \cdot \sqrt{x_i}\right)^2 \leqslant \left(\sum_{i=1}^{n} \left(\frac{1}{\sqrt{x_i}}\right)^2\right) \cdot \left(\sum_{i=1}^{n} \sqrt{x_i}^2\right)$$

so that

$$n^2 \leqslant \left(\sum_{i=1}^{n} \frac{1}{x_i}\right) \cdot \left(\sum_{i=1}^{n} x_i\right) .$$

This yields $n^2 \leqslant \sum_{i=1}^{n} \frac{1}{x_i} \cdot 1$, which gives the expected result.

2a. Compute the inner product of real functions f(x)=2x and g(x)=sin(x) over the closed interval [-π, π].

$$\langle f|g\rangle = \int_{-\pi}^{\pi} 2x \sin x \, dx = 4\pi$$

2b. Compute the length of the functions f and g.

$$\langle f|f\rangle = \int_{-\pi}^{\pi} 2x \cdot 2x \, dx = \frac{8}{3}\pi^3,$$

$$\|f\| = \sqrt{\frac{8}{3}\pi^3},$$

$$\langle g|g\rangle = \int_{-\pi}^{\pi} \sin x \cdot \sin x \, dx = \pi,$$

$$\|g\| = \sqrt{\pi}$$

2c. Compute the angle between the two functions.

$$\cos\vartheta = \frac{\langle f|g\rangle}{\|f\|\,\|g\|} = \frac{4\pi}{\sqrt{\frac{8}{3}\pi^3}\,\sqrt{\pi}} = \frac{\sqrt{6}}{\pi}$$

2d. Show that the Cauchy-Schwarz inequality holds true.

$$\langle f|g\rangle = 4\pi \leq \sqrt{\frac{8}{3}\pi^2} = \|f\|\,\|g\|,$$

and we see that Cauchy-Schwarz inequality holds.