Exercises of Mathematics for Artificial Intelligence and Data Science

Tomás A. Revilla* and Jan Valdman†

September 6, 2024

Faculty of Science,
University of South Bohemia,
Czech Republic

*trevillarimbach@jcu.cz
†jvaldman@jcu.cz

ISBN 978-80-7694-083-3

Contents

Preface

The Mathematics for Artificial Intelligence and Data Science course was created in the academic year 2021/22 as part of the MSc course on Artificial Intelligence and Data Science (MAID) organized jointly by the University of South Bohemia in České Budějovice and Deggendorf Institute of Technology. This collection of tutorial exercises accompanies the main textbook [1] and provides students with enough practical exercises. If you have any questions or improvement tips, feel free to contact us.

T. A. Revilla, J. Valdman - the authors.

1 Systems of linear equations

We aim at solving the system of linear equations

$$A\vec{x} = \vec{b},$$

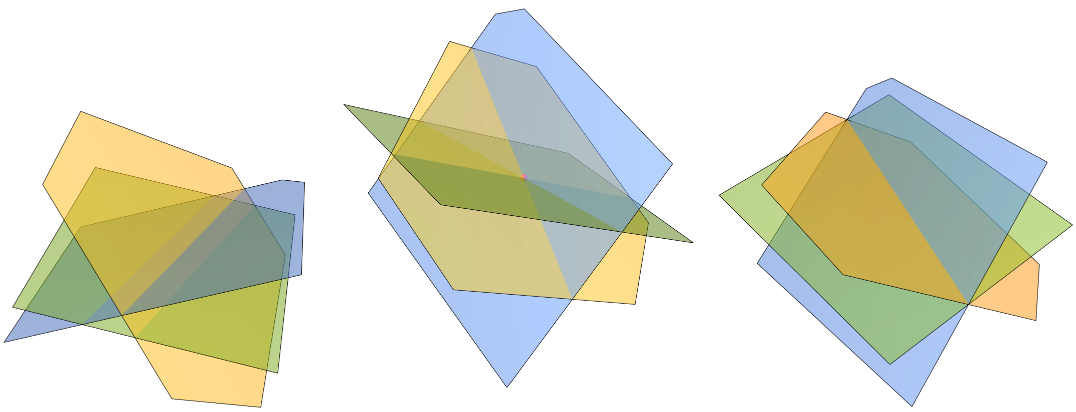
where $A \in \mathbb{R}^{m\times n}$ is a rectangular matrix with m rows and n columns, $\vec{b} \in \mathbb{R}^m$ is the right-hand side vector and $\vec{x} \in \mathbb{R}^n$ is a solution vector. There are three options: there is either no solution, one solution, or infinitely many solutions. We demonstrate is graphically for the problem in three space dimensions. Then the solution of the linear system of equations corresponds to the intersection of three planes. We consider three different problems (taken from [1], example 2.2, page 20) in forms

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & 2 \\ 2 & 0 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix}, \tag{1.1}$$

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & 2 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \\ 2 \end{bmatrix}, \tag{1.2}$$

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & 2 \\ 2 & 0 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \\ 5 \end{bmatrix}. \tag{1.3}$$

The first system (1.1) corresponds to no intersection of all planes (although any pair of planes intersect in various lines), the second system (1.2) to one intersection point and the third system (1.3) to intersection of all plane in one line. See Figure 1.1.



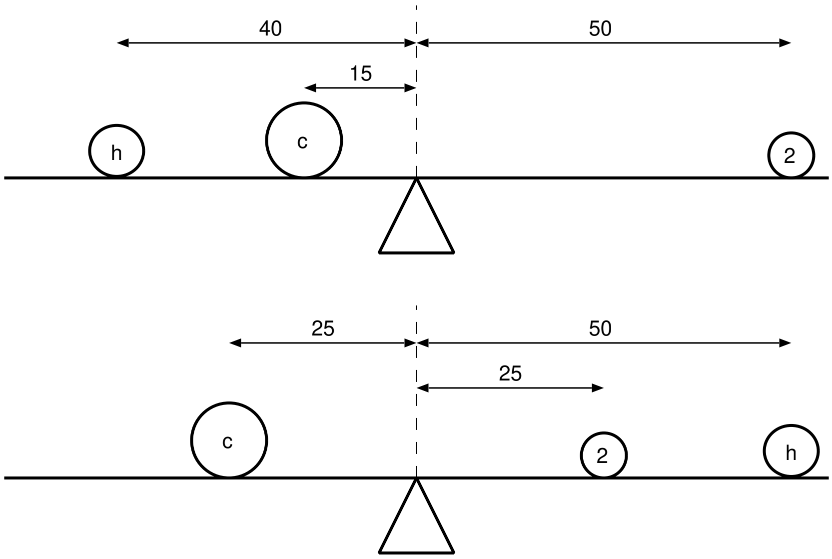
(a) (b) (c)

Figure 1.1: The planes in (a), (b) and (c) correspond to the three equations defined by (1.1), (1.2) and (1.3), respectively.

The distinction of three cases is easy to see from the corresponding reduced echelon forms of all three system providing

$$\left[\begin{array}{ccc|c} 1 & 0 & \frac{3}{2} & 0 \\ 0 & 1 & -\frac{1}{2} & 0 \\ 0 & 0 & 0 & 1 \end{array}\right], \qquad \left[\begin{array}{ccc|c} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{array}\right], \qquad \left[\begin{array}{ccc|c} 1 & 0 & \frac{3}{2} & \frac{5}{2} \\ 0 & 1 & -\frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 0 & 1 \end{array}\right].$$

Exercise 1.1. Three masses c, h and 2 are balanced as shown below. What are the values of c and h?



Solution. Balance requires that *torques* on the left and right side of the lever are equal:

$$\begin{aligned} 40h + 15c &= 50 \cdot 2 \\ 25c &= 25 \cdot 2 + 50h \end{aligned}$$

Divide both sides of the 1st equation by 2 and both sides of the 2nd by 25 to get

$$\begin{aligned} 8h + 3c &= 20 \\ c &= 2 + 2h \end{aligned}$$

Substitute c from the 2nd in the 1st and solve h

$$8h + 3(2 + 2h) = 20 \to 8h + 6 + 6h = 20 \to 14h = 14 \to \boxed{h = 1}$$

Solve $c = 2 + 2 \cdot 1 \to \boxed{c = 4}$

Exercise 1.2. The following reaction between toluene and nitric acid produces TNT and dihydrogen monoxide

$$x\,\mathrm{C_7\,H_8} + y\,\mathrm{HNO_3} \longrightarrow z\,\mathrm{C_7\,H_5\,O_6\,N_3} + w\,\mathrm{H_2\,O}$$

Find x, y, z, w.

Solution. From the *law of conservation of matter*

$$\begin{array}{ll} \text{conservation of C:} & 7x = 7z \\ \text{conservation of H:} & 8x + 1y = 5z + 2w \\ \text{conservation of N:} & 1y = 3z \\ \text{conservation of O:} & 3y = 6z + 1w \end{array}$$

Or the same thing as

$$\begin{array}{ccccccccc} 7x & + & 0y & - & 7z & - & 0w & = & 0 \\ 8x & + & 1y & - & 5z & - & 2w & = & 0 \\ 0x & + & 1y & - & 3z & - & 0w & = & 0 \\ 0x & + & 3y & - & 6z & - & 1w & = & 0 \end{array}$$

This system of equations is homogeneous because the independent terms are all 0. We solve this system using row operations

$$\left[\begin{array}{rrrr|r} 7 & 0 & -7 & 0 & 0 \\ 8 & 1 & -5 & -2 & 0 \\ 0 & 1 & -3 & 0 & 0 \\ 0 & 3 & -6 & -1 & 0 \end{array}\right]$$

This is the *augmented matrix* of coefficients. The vertical line represents the equal sign (=). and the columns on the left side correspond to the variables (the order maters!)

- Subtract row 1 multiplied by 8/7 from row 2, *i.e.*, $R_2 = R_2 - \frac{8}{7}R_1$:

$$\begin{array}{r|ccccc|c} R_2 & 8 & 1 & -5 & -2 & 0 \\ -\frac{8}{7}R_1 & -\frac{8}{7}\cdot 7 & -\frac{8}{7}\cdot 0 & -\frac{8}{7}\cdot(-7) & -\frac{8}{7}\cdot 0 & -\frac{8}{7}\cdot 0 \\ \hline R_2 & \mathbf{0} & \mathbf{1} & \mathbf{3} & \mathbf{-2} & \mathbf{0} \end{array} \quad \to \quad \left[\begin{array}{rrrr|r} 7 & 0 & -7 & 0 & 0 \\ \mathbf{0} & \mathbf{1} & \mathbf{3} & \mathbf{-2} & \mathbf{0} \\ 0 & 1 & -3 & 0 & 0 \\ 0 & 3 & -6 & -1 & 0 \end{array}\right]$$

- $R_3 = R_3 - R_2$:

$$\left[\begin{array}{rrrr|r} 7 & 0 & -7 & 0 & 0 \\ 0 & 1 & 3 & -2 & 0 \\ 0 & 0 & -6 & 2 & 0 \\ 0 & 3 & -6 & -1 & 0 \end{array}\right]$$

- $R_4 = R_4 - 3R_2$:

$$\left[\begin{array}{rrrr|r} 7 & 0 & -7 & 0 & 0 \\ 0 & 1 & 3 & -2 & 0 \\ 0 & 0 & -6 & 2 & 0 \\ 0 & 3 & -15 & 5 & 0 \end{array}\right]$$

- $R_4 = R_4 - \frac{5}{2}R_3$:

$$\left[\begin{array}{rrrr|r} 7 & 0 & -7 & 0 & 0 \\ 0 & 1 & 3 & -2 & 0 \\ 0 & 0 & -6 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

This is a *row echelon form* (REF). The 4th row is *redundant*, we have three *independent* rows/equations

$$\begin{aligned} 7x - 7z &= 0 \\ y + 3z - 2w &= 0 \\ -6z + 2w &= 0 \end{aligned}$$

From the 3rd $z = \frac{1}{3}w$, from the 1st $x = \frac{1}{3}w$, and from the 2nd $y = 2w - 3 \cdot \frac{1}{3}w = w$.

We are left with w, it can be anything! Positive, negative, zero! Of course, since we are dealing with *matter* here, let us consider a "sensible" value, like $w = 3$. Then $z = 1, x = 1$ and $y = 3$. Now

$$C_7\,H_8 + 3\,HNO_3 \longrightarrow C_7\,H_5\,O_6\,N_3 + 3\,H_2\,O$$

is *mass balanced*.

Remark. In the mechanical example there is only one solution. There are three masses and two *constraints*, *i.e.*, the two equations, but one of the masses ("2") was already set. In the chemical example, there are infinite solutions because the total mass of the system is not constrained.

Exercise 1.3. Solve the system from the mechanical exercise 1.1 using *Cramer's rule*.

Solution. Given

$$\underbrace{\begin{bmatrix} 8 & 3 \\ -2 & 1 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} h \\ c \end{bmatrix}}_{\vec{X}} = \underbrace{\begin{bmatrix} 20 \\ 2 \end{bmatrix}}_{\vec{b}}$$

$$h = \frac{\Delta_h}{\Delta} = \frac{\begin{vmatrix} 20 & 3 \\ 2 & 1 \end{vmatrix}}{\begin{vmatrix} 8 & 3 \\ -2 & 1 \end{vmatrix}} = \frac{20 \times 1 - 2 \times 3}{8 \times 1 - (-2 \times 3)} = \frac{14}{14} = \boxed{1}$$

$$c = \frac{\Delta_c}{\Delta} = \frac{\begin{vmatrix} 8 & 20 \\ -2 & 2 \end{vmatrix}}{\begin{vmatrix} 8 & 3 \\ -2 & 1 \end{vmatrix}} = \frac{8 \times 2 - (-2 \times 20)}{8 \times 1 - (-2 \times 3)} = \frac{56}{14} = \boxed{4}$$

- Δ: the determinant of A
- Δ_h: like Δ but 1st column replaced by $\vec{b}$
- Δ_c: like Δ but 2nd column replaced by $\vec{b}$

Exercise 1.4. Solve the system from the mechanical exercise 1.1 using multiplication by the inverse.

Solution. Given

$$\underbrace{\begin{bmatrix} 8 & 3 \\ -2 & 1 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} h \\ c \end{bmatrix}}_{\vec{X}} = \underbrace{\begin{bmatrix} 20 \\ 2 \end{bmatrix}}_{\vec{b}}$$

We want to do this

$$A\vec{X} = \vec{b} \to A^{-1}A\vec{X} = A^{-1}\vec{b} \to I\vec{X} = A^{-1}\vec{b} \to \vec{X} = A^{-1}\vec{b}$$

where I is the identity matrix $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

Find the matrix inverse. There is a very simple formula for 2×2 matrices

$$A = \begin{bmatrix} 8 & 3 \\ -2 & 1 \end{bmatrix} \to A^{-1} = \underbrace{\frac{1}{8 \cdot 1 - (-2 \cdot 3)}}_{\text{determinant}} \underbrace{\begin{bmatrix} 1 & -3 \\ 2 & 8 \end{bmatrix}}_{\text{adjugate matrix}} = \frac{1}{14}\begin{bmatrix} 1 & -3 \\ 2 & 8 \end{bmatrix} = \begin{bmatrix} \frac{1}{4} & \frac{-3}{14} \\ \frac{2}{14} & \frac{8}{14} \end{bmatrix}$$

Then

$$\vec{X} = A^{-1}\vec{b} = \frac{1}{14}\begin{bmatrix} 1 & -3 \\ 2 & 8 \end{bmatrix}\begin{bmatrix} 20 \\ 2 \end{bmatrix} = \frac{1}{14}\begin{bmatrix} 20 - 6 \\ 40 + 16 \end{bmatrix} = \frac{1}{14}\begin{bmatrix} 14 \\ 56 \end{bmatrix} = \boxed{\begin{bmatrix} 1 \\ 4 \end{bmatrix}}$$

Exercise 1.5. Solve the system from the chemical exercise 1.2 by *gaussian elimination.*

Solution. Start with the REF (disregard the last row full of zeros)

$$\left[\begin{array}{cccc|c} \boxed{7} & 0 & -7 & 0 & 0 \\ 0 & \boxed{1} & 3 & -2 & 0 \\ 0 & 0 & \boxed{-6} & 2 & 0 \end{array}\right]$$

The numbers with squares are *pivots*, they are the first *non-zero* elements of their rows and the last *non-zero* elements of their columns. We want the *reduced row echelon form* (RREF) where all pivots are equal to 1 and the only entries of their columns

- $R_1 = \frac{1}{7}R_1$ and $R_3 = -\frac{1}{6}R_3$

$$\left[\begin{array}{cccc|c} 1 & 0 & -1 & 0 & 0 \\ 0 & 1 & 3 & -2 & 0 \\ 0 & 0 & 1 & \frac{-1}{3} & 0 \end{array}\right]$$

- $R_1 = R_1 + R_3$ and $R_2 = R_2 - 3R_3$

$$\left[\begin{array}{cccc|c} 1 & 0 & 0 & \frac{-1}{3} & 0 \\ 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 1 & \frac{-1}{3} & 0 \end{array}\right]$$

Now see

$$\underbrace{\left[\begin{array}{cccc|c} 1 & 0 & 0 & \frac{-1}{3} & 0 \\ 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 1 & \frac{-1}{3} & 0 \end{array}\right]}_{RREF} \to \begin{array}{ccccccccc} 1x & + & 0y & + & 0z & -\frac{1}{3}w & = & 0 \\ 0x & + & 1y & + & 0z & -1w & = & 0 \\ 0x & + & 0y & + & 1z & -\frac{1}{3}w & = & 0 \end{array} \to \begin{array}{ccc} x & = & \frac{1}{3}w \\ y & = & w \\ z & = & \frac{1}{3}w \end{array}$$

The general solution of the equation $M\vec{V} = \vec{0}$ is

$$\boxed{\vec{V} = \begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = \left\{ \alpha \begin{bmatrix} \frac{1}{3} \\ 1 \\ \frac{1}{3} \\ 1 \end{bmatrix} \middle| \, \alpha \in \mathbb{R} \right\}}$$

Exercise 1.6. (ex. 2.5a from [1]). Find the general solution of

$$\begin{bmatrix} 1 & 1 & -1 & -1 \\ 2 & 5 & -7 & -5 \\ 2 & -1 & 1 & 3 \\ 5 & 2 & -4 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 1 \\ -2 \\ 4 \\ 6 \end{bmatrix}$$

Solution. Write the augmented matrix

$$\left[\begin{array}{cccc|c} 1 & 1 & -1 & -1 & 1 \\ 2 & 5 & -7 & -5 & -2 \\ 2 & -1 & 1 & 3 & 4 \\ 5 & 2 & -4 & 2 & 6 \end{array}\right]$$

Perform gaussian elimination

$$\left[\begin{array}{rrrr|r} 1 & 1 & -1 & -1 & 1 \\ 2 & 5 & -7 & -5 & -2 \\ 2 & -1 & 1 & 3 & 4 \\ 5 & 2 & -4 & 2 & 6 \end{array}\right] \to \cdots \to \left[\begin{array}{rrrr|r} 1 & 0 & \frac{2}{3} & 0 & \frac{7}{3} \\ 0 & 1 & -\frac{5}{3} & -1 & -\frac{4}{3} \\ 0 & 0 & -2 & 2 & -2 \\ 0 & 0 & 0 & 0 & 1 \end{array}\right]$$

The last row of the REF tells us that $0 \cdot x_4 = 1$ which makes no sense. This system is inconsistent, it has **no solutions**.

Exercise 1.7. (ex. 2.5b from [1]). Find the general solution of

$$\begin{bmatrix} 1 & -1 & 0 & 0 & 1 \\ 1 & 1 & 0 & -3 & 0 \\ 2 & -1 & 0 & 1 & -1 \\ -1 & 2 & 0 & -2 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 3 \\ 6 \\ 5 \\ -1 \end{bmatrix}$$

Solution. Notice that the matrix of coefficients is not squared, there are 5 variables but just 4 equations. Yet, we can find a *general solution* by gaussian elimination.

$$\left[\begin{array}{rrrrr|r} 1 & -1 & 0 & 0 & 1 & 3 \\ 1 & 1 & 0 & -3 & 0 & 6 \\ 2 & -1 & 0 & 1 & -1 & 5 \\ -1 & 2 & 0 & -2 & -1 & -1 \end{array}\right] \to \cdots \to \left[\begin{array}{rrrrr|r} 1 & 0 & 0 & 0 & -1 & 3 \\ 0 & 1 & 0 & 0 & -2 & 0 \\ 0 & 0 & 0 & 1 & -1 & -1 \end{array}\right]$$

The last one is in *reduced row echelon form* (RREF), and we can ignore the last row. Put it back into matrix–vector form and as equations

$$\begin{bmatrix} 1 & 0 & 0 & 0 & -1 \\ 0 & 1 & 0 & 0 & -2 \\ 0 & 0 & 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 3 \\ 0 \\ -1 \end{bmatrix} \to \begin{array}{rrcr} x_1 & -x_5 & = & 3 \\ x_2 & -2x_5 & = & 0 \\ x_4 & -x_5 & = & -1 \end{array}$$

Notice that x_3 does not appear on the right version. This means that any $x_3 = \alpha \in \mathbb{R}$ is valid. Next, x_5 can be chosen arbitrarily, and x_1, x_2, x_4 be given in terms of $x_5 = \beta \in \mathbb{R}$, *i.e.*, $x_1 = 3 + \beta$, $x_2 = 2\beta$ and $x_4 = \beta - 1$. Now watch the "miracle"

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} 3+\beta \\ 2\beta \\ \alpha \\ \beta - 1 \\ \beta \end{bmatrix} = \underbrace{\begin{bmatrix} 3 \\ 0 \\ 0 \\ -1 \\ 0 \end{bmatrix}}_{\vec{U}} + \alpha \underbrace{\begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}}_{\vec{V}} + \beta \underbrace{\begin{bmatrix} 1 \\ 2 \\ 0 \\ 1 \\ 1 \end{bmatrix}}_{\vec{W}}$$

The general solution is

$$\boxed{\vec{X} = \{\vec{U} + \alpha\vec{V} + \beta\vec{W} | \alpha, \beta \in \mathbb{R}\} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \left\{ \begin{bmatrix} 3 \\ 0 \\ 0 \\ -1 \\ 0 \end{bmatrix} + \alpha \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} + \beta \begin{bmatrix} 1 \\ 2 \\ 0 \\ 1 \\ 1 \end{bmatrix} | \alpha, \beta \in \mathbb{R} \right\}}$$

Exercise 1.8. Use gaussian elimination to find the inverse of

$$A = \begin{bmatrix} 8 & 3 \\ -2 & 1 \end{bmatrix}$$

Solution. The inverse has four unknown entries x, y, z, w

$$\underbrace{\begin{bmatrix} 8 & 3 \\ -2 & 1 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} x & y \\ z & w \end{bmatrix}}_{A^{-1}} = \underbrace{\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}}_{I} \to \left[\begin{array}{cc|cc} 8 & 3 & 1 & 0 \\ -2 & 1 & 0 & 1 \end{array}\right]$$

Perform row operations until the left side of the augmented matrix is the identity matrix I

$$\left[\begin{array}{cc|cc} 8 & 3 & 1 & 0 \\ -2 & 1 & 0 & 1 \end{array}\right] \begin{matrix} \\ \times 4 \end{matrix} \to \left[\begin{array}{cc|cc} 8 & 3 & 1 & 0 \\ -8 & 4 & 0 & 4 \end{array}\right] \begin{matrix} \\ +R_1 \end{matrix} \to \left[\begin{array}{cc|cc} 8 & 3 & 1 & 0 \\ 0 & 7 & 1 & 4 \end{array}\right] \begin{matrix} \times 7 \\ \times(-3) \end{matrix}$$

$$\to \left[\begin{array}{cc|cc} 56 & 21 & 7 & 0 \\ 0 & -21 & -3 & -12 \end{array}\right] \begin{matrix} +R_2 \\ \times(-1) \end{matrix} \to \left[\begin{array}{cc|cc} 56 & 0 & 4 & -12 \\ 0 & 21 & 3 & 12 \end{array}\right] \begin{matrix} \times 1/56 \\ \times 1/21 \end{matrix}$$

$$\to \left[\begin{array}{cc|cc} 1 & 0 & \frac{4}{56} & \frac{-12}{56} \\ 0 & 1 & \frac{3}{21} & \frac{12}{21} \end{array}\right] \to \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x & y \\ z & w \end{bmatrix} = \begin{bmatrix} \frac{1}{14} & \frac{-3}{14} \\ \frac{1}{7} & \frac{3}{7} \end{bmatrix} \to \boxed{A^{-1} = \begin{bmatrix} \frac{1}{14} & \frac{-3}{14} \\ \frac{1}{7} & \frac{3}{7} \end{bmatrix}}$$

Remark. We actually solved a system of four equations in four unknowns

$$\begin{bmatrix} 8 & 3 \\ -2 & 1 \end{bmatrix} \begin{bmatrix} x & y \\ z & w \end{bmatrix} = \begin{bmatrix} 8x+3z & 8y+3w \\ -2x+z & -2y+w \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \to \begin{aligned} 8x+3z &= 1 \\ 8y+3w &= 0 \\ -2x+z &= 0 \\ -2y+w &= 1 \end{aligned}$$

Exercise 1.9. (ex. 2.10a from [1]). Check if the following vectors are linear independent

$$\vec{X}_1 = \begin{bmatrix} 2 \\ -1 \\ 3 \end{bmatrix}, \vec{X}_2 = \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix}, \vec{X}_1 = \begin{bmatrix} 3 \\ -3 \\ 8 \end{bmatrix}$$

Solution. Let $\vec{0}$ be the *zero vector*. Vectors $\vec{X}_1, \vec{X}_2, \vec{X}_3$ are linear independent if the solution of

$$a\vec{X}_1 + b\vec{X}_2 + c\vec{X}_3 = \vec{0},$$

can **only** be $a = b = c = 0$.

$$a\begin{bmatrix} 2 \\ -1 \\ 3 \end{bmatrix} + b\begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix} + c\begin{bmatrix} 3 \\ -3 \\ 8 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \to \underbrace{\begin{bmatrix} 2a+b+3c \\ -1a+b-3c \\ 3a-2b+8c \end{bmatrix}}_{\text{this is a single column!}} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \to \begin{bmatrix} 2 & 1 & 3 \\ -1 & 1 & -3 \\ 3 & -2 & 8 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Gaussian elimination

$$\left[\begin{array}{ccc|c} 2 & 1 & 3 & 0 \\ -1 & 1 & -3 & 0 \\ 3 & -2 & 8 & 0 \end{array}\right] \to \cdots \to \left[\begin{array}{ccc|c} 1 & 0 & 2 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

Now we have

$$a\begin{bmatrix}1\\0\\0\end{bmatrix}+b\begin{bmatrix}0\\1\\0\end{bmatrix}+c\begin{bmatrix}1\\-1\\0\end{bmatrix}=\begin{bmatrix}0\\0\\0\end{bmatrix}\rightarrow\begin{matrix}a+2c=0\\b-c=0\end{matrix}\rightarrow\begin{matrix}a=-2c\\b=c\end{matrix}$$

There are infinite solutions, pick any c and you get valid a and b. For instance, try $c = 1$, $a = -2$, $b = 1$. Thus, the vectors are not independent. You can conclude this from the fact that one of the rows from the REF is full of zeros.

2 Vector spaces

Exercise 2.1. Check if $\mathbb{R}^3 = \{(x, y, z)|x, y, z \in \mathbb{R}\}$ is a vector space.

Solution. Let $a, b, c, d, e, f, \lambda \in \mathbb{R}$. Then $\mathbf{u} = (a, b, c) \in \mathbb{R}^3$ and $\mathbf{v} = (d, e, f) \in \mathbb{R}^3$

1. Closure under addition:

$$\mathbf{u} + \mathbf{v} = (a + d, b + e, c + f) \Rightarrow a + d, b + e, c + f \in \mathbb{R} \Rightarrow \mathbf{u} + \mathbf{v} \in \mathbb{R}^3$$

   Check.

2. Closure under scalar multiplication:

$$\lambda\mathbf{u} = \lambda(a, b, c) = (\lambda a, \lambda b, \lambda c) \Rightarrow \lambda a, \lambda b, \lambda c \in \mathbb{R} \Rightarrow \lambda\mathbf{u} \in \mathbb{R}^3$$

   Check.

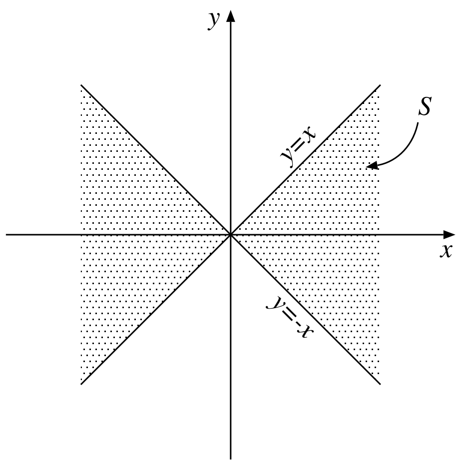
3. Contains zero element:

$$\mathbf{0} = (0, 0, 0) \Rightarrow 0 \in \mathbb{R} \Rightarrow \mathbf{0} \in \mathbb{R}^3$$

   Check.

$\mathbb{R}^3$ is a vector space.

Exercise 2.2. Check if $S = \{(x, ax) \,|\, x, a \in \mathbb{R}; |a| \leq 1\}$ is a vector space.



Solution. Let $x, y, \lambda \in \mathbb{R}$ and $|\alpha| \leq 1$. Then $\mathbf{u} = (x, \alpha x) \in S$ and $\mathbf{v} = (y, -\alpha y) \in S$

1. Closure under addition:

$$\mathbf{u} + \mathbf{v} = (x + y, \alpha x - \alpha y) \Rightarrow \text{Let } x = 1, y = -1 \Rightarrow \mathbf{u} + \mathbf{v} = (0, 2) \notin S$$

   Nope.

2. Closure under scalar multiplication:

$$\lambda\mathbf{u} = \lambda(x, ax) = (\lambda x, a \cdot \lambda x) \Rightarrow \lambda x \in \mathbb{R} \Rightarrow \lambda\mathbf{u} \in S$$

   Check.

3. Contains zero element:
$$(0, a \cdot 0) \in S \Rightarrow (0,0) = \mathbf{0} \in S$$
Check.

S is not a vector space.

Exercise 2.3. Check if the collection of 2×2 matrices with 0's in the upper right and lower left entries
$$\left\{ \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix} \middle| a, b \in \mathbb{R} \right\}$$
is a vector space.

Solution. Let $a_1, a_2, b_1, b_2, \lambda \in \mathbb{R}$

1. Closure under addition:
$$\begin{bmatrix} a_1 & 0 \\ 0 & b_1 \end{bmatrix} + \begin{bmatrix} a_2 & 0 \\ 0 & b_2 \end{bmatrix} = \begin{bmatrix} a_1 + a_2 & 0 \\ 0 & b_1 + b_2 \end{bmatrix}$$
has 0's in the upper rigth and lower left entries. Check.
2. Closure under scalar multiplication:
$$\lambda \begin{bmatrix} a_1 & 0 \\ 0 & b_1 \end{bmatrix} = \begin{bmatrix} \lambda a_1 & 0 \\ 0 & \lambda b_1 \end{bmatrix}$$
has 0's in the upper rigth and lower left entries. Check.
3. Contains zero element: let $a = 0, b = 0$
$$\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$
has 0's in the upper rigth and lower left entries. Check.

It is a vector space.

Exercise 2.4. Check if the collection of cubic polynomials with no quadratic term,
$$a_0 + a_1 x + a_3 x^3$$
is a vector space.

Solution. Let $\lambda \in \mathbb{R}$, $p = \begin{bmatrix} 1 & x & x^3 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_3 \end{bmatrix}$ and $q = \begin{bmatrix} 1 & x & x^3 \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \\ b_3 \end{bmatrix}$

1. Closure under addition:
$$\begin{aligned} p + q &= \begin{bmatrix} 1 & x & x^3 \end{bmatrix} \left(\begin{bmatrix} a_0 \\ a_1 \\ a_3 \end{bmatrix} + \begin{bmatrix} b_0 \\ b_1 \\ b_3 \end{bmatrix} \right) = \begin{bmatrix} 1 & x & x^3 \end{bmatrix} \begin{bmatrix} a_0 + b_0 \\ a_1 + b_1 \\ a_3 + b_3 \end{bmatrix} \\ &= a_0 + b_0 + (a_1 + b_1)x + (a_3 + b_3)x^3 \end{aligned}$$
is a cubic polynomial with no quadratic term. Check.

2. Closure under scalar multiplication:

$$\lambda p = \lambda \begin{bmatrix} 1 & x & x^3 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_3 \end{bmatrix} = a_0\lambda + a_1\lambda x + a_3\lambda x^3$$

is a cubic polynomial with no quadratic term. Check.

3. Contains zero element: let $a_0 = a_1 = a_3 = 0$

$$\begin{bmatrix} 1 & x & x^3 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} = 0 + 0 \cdot x + 0 \cdot x^3$$

is a cubic polynomial with no quadratic term. Check.

It is a vector space.

Exercise 2.5. (ex. 2.9 from [1]). Which of the following sets are subspaces of $\mathbb{R}^3$

$A = \{(\lambda, \lambda + \mu^3, \lambda - \mu^3) \mid \lambda, \mu \in \mathbb{R}\}$
$B = \{(\lambda^2, -\lambda^2, 0) \mid \lambda \in \mathbb{R}\}$
$C = \{(\varepsilon_1, \varepsilon_2, \varepsilon_3) \in \mathbb{R}^3 \mid \varepsilon_1 - 2\varepsilon_2 + 3\varepsilon_3 = \gamma\}$
$D = \{(\varepsilon_1, \varepsilon_2, \varepsilon_3) \in \mathbb{R}^3 \mid \varepsilon_2 \in \mathbb{Z}\}$

Solution. For A:

$$A = \left\{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \lambda \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + \mu^3 \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} \middle| \lambda, \mu \in \mathbb{R} \right\}$$

Setting $\lambda = \mu = 0$ we have the zero vector $[0, 0, 0]^{\mathsf{T}}$. Let

$$\vec{u}_1 = \lambda_1 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + \mu_1^3 \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}, \quad \vec{u}_2 = \lambda_2 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + \mu_2^3 \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}$$

$$\vec{u}_3 = \vec{u}_1 + \vec{u}_2 = (\lambda_1 + \lambda_2) \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + (\mu_1^3 + \mu_2^3) \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} = \lambda_3 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + \mu_3^3 \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}$$

where $\lambda_3 = \lambda_1 + \lambda_2 \in \mathbb{R}$ and $\mu_3 = \sqrt[3]{\mu_1^3 + \mu_2^3} \in \mathbb{R}$. The set is closed under addition. Now let $r \in \mathbb{R}$

$$\vec{u}_4 = r \left(\lambda \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + \mu^3 \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} \right) = r\lambda \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + r\mu^3 \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} = \lambda_4 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + \mu_4^3 \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}$$

where $\lambda_4 = r\lambda \in \mathbb{R}$ and $\mu_4 = \mu\sqrt[3]{r} \in \mathbb{R}$. The set is closed under multiplication by scalar. A is a subspace of $\mathbb{R}^3$. A is a plane inside a volume.

Solution. For B:

$$B = \left\{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \lambda^2 \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} \middle| \lambda \in \mathbb{R} \right\}$$

Setting $\lambda = 0$ we have the zero vector $[0, 0, 0]^\mathsf{T}$. Let

$$\vec{u}_1 = \lambda_1^2 \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \vec{u}_2 = \lambda_2^2 \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$$

$$\vec{u}_3 = \vec{u}_1 + \vec{u}_2 = (\lambda_1^2 + \lambda_2^2) \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} = \lambda_3^2 \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$$

where $\lambda_3^2 = \lambda_1^2 + \lambda_2^2 \in \mathbb{R}$. The set is closed under addition. Now let $r \in \mathbb{R}$

$$\vec{u}_4 = r\lambda^2 \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} = |r|\lambda^2 \begin{bmatrix} \text{sign}(r) \\ -\text{sign}(r) \\ 0 \end{bmatrix} = \lambda_4 \begin{bmatrix} \text{sign}(r) \\ -\text{sign}(r) \\ 0 \end{bmatrix}$$

where $\lambda_4 = |r|\lambda^2 \geq 0$ and sign(r) is the *sign function.* If $r < 0$ the 2nd entry of $\vec{u}_4$ is negative and $\vec{u}_4 \notin B$. The set is not closed under multiplication by scalar. B is a not subspace of $\mathbb{R}^3$. B is a ray in the $z = 0$ plane where $y = -x$ and $x \geq 0$.

Solution. For C:

$$C = \left\{ \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \end{bmatrix} \in \mathbb{R}^3 \middle| \begin{bmatrix} 1 & -2 & 3 \end{bmatrix} \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \end{bmatrix} = \gamma \right\}$$

The zero vector $[\varepsilon_1, \varepsilon_2, \varepsilon_3]^\mathsf{T} = [0, 0, 0]^\mathsf{T}$ exists if $\gamma = 0$. Let $\gamma = 0$ then

$$\vec{x} = \begin{bmatrix} 1 & -2 & 3 \end{bmatrix} \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \end{bmatrix} = \vec{0}, \vec{y} = \begin{bmatrix} 1 & -2 & 3 \end{bmatrix} \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \epsilon_3 \end{bmatrix} = \vec{0}$$

$$\vec{x} + \vec{y} = \vec{0} + \vec{0} = \vec{0}$$

The set is closed under addition. Now let $r \in \mathbb{R}$

$$r\vec{x} = r\vec{0} = \vec{0}$$

The set is closed under multiplication by scalar. C is a subspace of $\mathbb{R}^3$ if and only if $\gamma = 0$.

Solution. For D: If an element of D is multiplied a non-integer number such as $\frac{1}{2}$ or π, the result is not in D. Thus, D is not a subspace of $\mathbb{R}^3$.

Exercise 2.6. Consider the following subspace of $\mathbb{R}^4$

$$U = \text{span}\left(\begin{bmatrix} 1 \\ 1 \\ -3 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ -1 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ -1 \\ 1 \end{bmatrix} \right)$$

Determine a basis for U.

Solution. First check linear independence of

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \\ -3 \\ 1 \end{bmatrix}, \mathbf{v}_2 = \begin{bmatrix} 2 \\ -1 \\ 0 \\ -1 \end{bmatrix}, \mathbf{v}_3 = \begin{bmatrix} -1 \\ 1 \\ -1 \\ 1 \end{bmatrix}$$

by proving there is a nontrivial solution for $\lambda_1\mathbf{v}_1 + \lambda_2\mathbf{v}_2 + \lambda_3\mathbf{v}_3 = \mathbf{0}$

$$\begin{bmatrix} 1 & 2 & -1 \\ 1 & -1 & 1 \\ -3 & 0 & -1 \\ 1 & -1 & 1 \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \to \left[\begin{array}{ccc|c} 1 & 2 & -1 & 0 \\ 1 & -1 & 1 & 0 \\ -3 & 0 & -1 & 0 \\ 1 & -1 & 1 & 0 \end{array}\right] \to \cdots \to \left[\begin{array}{ccc|c} 1 & 2 & -1 & 0 \\ 0 & -3 & 2 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

from this we get $\lambda_1 + 2\lambda_2 = \lambda_3$ and $3\lambda_2 = 2\lambda_3$. You can solve the 2nd with $\lambda_2 = 2, \lambda_3 = 3$, then it must be $\lambda_1 = -1$ in the first. Since this solution is non trivial the vectors are not independent, i.e., you can get any by combining the other two. Now consider the first two vectors

$$\alpha_1 \begin{bmatrix} 1 \\ 1 \\ -3 \\ 1 \end{bmatrix} + \alpha_2 \begin{bmatrix} 2 \\ -1 \\ 0 \\ -1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

they are independent because the only way that the sum $\mathbf{0}$ is when $\alpha_1 = \alpha_2 = 0$ which is trivial. Thus, $\mathbf{v}_1$ and $\mathbf{v}_2$ are independent and can be a basis for U.

$$\boxed{\mathcal{B}_U = \left\langle \begin{bmatrix} 1 \\ 1 \\ -3 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ -1 \\ 0 \\ -1 \end{bmatrix} \right\rangle} \Longleftrightarrow U = \text{span}\left(\begin{bmatrix} 1 \\ 1 \\ -3 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ -1 \\ 0 \\ -1 \end{bmatrix} \right)$$

Exercise 2.7. Find a basis and the dimension of this vector space

$$V = \left\{ \begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} \in \mathbb{R}^4 \middle| x - w + z = 0 \right\}$$

Solution. Use the condition/constraint/equation $x - w + z = 0$ to parameterize V as the sum of three vectors

$$V = \left\{ \begin{bmatrix} w - z \\ y \\ z \\ w \end{bmatrix} = y \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} + z \begin{bmatrix} -1 \\ 0 \\ 1 \\ 0 \end{bmatrix} + w \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} \middle| y, z, w \in \mathbb{R} \right\}$$

Each of the vectors above has a "1" that can't be generated by combining the other two, i.e., they are all independent. Thus,

$$\boxed{\mathcal{B}_V = \left\langle \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix} \right\rangle}$$

The dimension of V is the number of vectors of the basis, $\boxed{\dim(V) = 3}$. Notice that $V \subset \mathbb{R}^4$ is a subspace, a 3-dimensional space embedded inside a 4-dimensional space.

Exercise 2.8. Find a basis and dimension of

$$\left\{a_0 + a_1 x + a_2 x^2 + a_3 x^3 \mid a_0 + a_1 = 0\,,\ a_2 - 2a_3 = 0\right\} \subseteq \mathcal{P}_3$$

Solution. Parameterize $a_0 = -a_1, a_2 = 2a_3$. We get

$$\left\{-a_1 + a_1 x + 2a_3 x^2 + a_3 x^3\right\} = \left\{a_1(-1+x) + a_3\left(2x^2 + x^3\right)\right\}$$

The polynomial subspace is spanned by $-1+x$ and $2x^2+x^3$ which are independent, i.e., neither can be obtained from the other by multiplication with a scalar. Thus, we have a basis

$$\left\langle -1+x, 2x^2+x^3 \right\rangle \Longleftrightarrow \boxed{\left\langle \begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 2 \\ 1 \end{bmatrix} \right\rangle}$$

and the dimension of this subspace is 2.

3 Linear mappings

Exercise 3.1. Give a basis for the column space of this matrix. Give the matrix's rank and nullity

$$\begin{bmatrix} 1 & 3 & -1 & 2 \\ 2 & 1 & 1 & 0 \\ 0 & 1 & 1 & 4 \end{bmatrix}$$

Solution. We want the basis of the following span

$$\text{span}\left(\begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \\ 4 \end{bmatrix}\right) \subseteq \mathbb{R}^3$$

Check column independence by solving the homogeneous system

$$x\begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} + y\begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix} + z\begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix} + w\begin{bmatrix} 2 \\ 0 \\ 4 \end{bmatrix} = \begin{bmatrix} 1 & 3 & -1 & 2 \\ 2 & 1 & 1 & 0 \\ 0 & 1 & 1 & 4 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

That is

$$\left[\begin{array}{cccc|c} 1 & 3 & -1 & 2 & 0 \\ 2 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 4 & 0 \end{array}\right] \to \cdots \to \left[\begin{array}{cccc|c} 1 & 0 & 0 & -2 & 0 \\ 0 & 1 & 0 & 2 & 0 \\ 0 & 0 & 1 & 2 & 0 \end{array}\right]$$

There are three independent columns. Thus, a basis can be

$$\boxed{\mathcal{B} = \left\langle \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 3 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix} \right\rangle}$$

The matrix's rank is the dimension of the basis, i.e., the number of elements of the basis, so <u>the rank is 3</u>. By the *rank–nullity theorem* the matrix's nullity is the number of columns (of the matrix) minus the rank, so <u>the nullity is 1</u>. The nullity is the dimension of the *kernel* or *null space*, the (general) solution of the homogeneous system (use the RREF)

$$\begin{aligned} x - 2w &= 0 \\ y + 2w &= 0 \\ z + 2w &= 0 \end{aligned}$$

$$\text{Ker} = \left\{ \begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} = \lambda \begin{bmatrix} 1 \\ -2 \\ -2 \\ 1 \end{bmatrix} \middle| \lambda \in \mathbb{R}^2 \right\}$$

See, the matrix's *domain* is 4-dimensional (columns) and the kernel/nullspace is a line (1-dimensional subset of $\mathbb{R}^4$) that *maps* to the $\vec{0}$ vector in the matrix's *image* which is of 3-dimensional.

Exercise 3.2. Give a basis for the span of the following set

$$\{x + x^2, 2 - 2x, 7, 4 + 3x + 2x^2\}$$

Give also the rank and the nullity.

Solution. Parameterize in matrix form

$$\begin{bmatrix} 0 \cdot x^0 & + & 1 \cdot x^1 & + & 1 \cdot x^2 \\ 2 \cdot x^0 & - & 2 \cdot x^1 & + & 0 \cdot x^2 \\ 7 \cdot x^0 & + & 0 \cdot x^1 & + & 0 \cdot x^2 \\ 4 \cdot x^0 & + & 3 \cdot x^1 & + & 2 \cdot x^2 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 1 \\ 2 & -2 & 0 \\ 7 & 0 & 0 \\ 4 & 3 & 2 \end{bmatrix} \cdot \begin{bmatrix} x^0 \\ x^1 \\ x^2 \end{bmatrix}$$

We want to know the number of independent polynomials, which translates to independent rows

$$\begin{bmatrix} 0 & 1 & 1 \\ 2 & -2 & 0 \\ 7 & 0 & 0 \\ 4 & 3 & 2 \end{bmatrix} \to \begin{bmatrix} 0 & 1 & 1 \\ 1 & -1 & 0 \\ 1 & 0 & 0 \\ 4 & 3 & 2 \end{bmatrix} \to \begin{bmatrix} 0 & 1 & 1 \\ 0 & -1 & 0 \\ 1 & 0 & 0 \\ 4 & 3 & 2 \end{bmatrix} \to \begin{bmatrix} 0 & 0 & 1 \\ 0 & -1 & 0 \\ 1 & 0 & 0 \\ 4 & 3 & 2 \end{bmatrix} \to \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

There are three independent rows, i.e., a basis can be

$$\boxed{\mathcal{B} = \langle (1,0,0), (0,1,0), (0,0,1) \rangle \Longleftrightarrow \mathcal{B} = \left\langle 1, x, x^2 \right\rangle}$$

The rank of the span is 3 and by the *rank–nullity theorem* the nullity is the number of columns minus the rank, so the nullity is 0. All the polynomials spanned by original set can be produced with the three elements of the basis.

Exercise 3.3. Verify that $h : \mathcal{P}_3 \to \mathbb{R}^2$ given by

$$ax^2 + bx + c \longmapsto \begin{bmatrix} a + b \\ a + c \end{bmatrix}$$

is a *homomorphism.* Hint: check that $h(\alpha A + \beta B) = \alpha h(A) + \beta h(B)$.

Solution. We need to map two polynomials

$$\text{mapping of polynomial 1: } a_1 x^2 + b_1 x + c_1 \longmapsto \begin{bmatrix} a_1 + b_1 \\ a_1 + c_1 \end{bmatrix}$$

$$\text{mapping of polynomial 2: } a_2 x^2 + b_2 x + c_2 \longmapsto \begin{bmatrix} a_2 + b_2 \\ a_2 + c_2 \end{bmatrix}$$

Linear combination of the polynomials

$$\lambda_1 \left(a_1 x^2 + b_1 x + c_1\right) + \lambda_2 \left(a_2 x^2 + b_2 x + c_2\right) = \left(\lambda_1 a_1 + \lambda_2 a_2\right) x^2 + \left(\lambda_1 b_1 + \lambda_2 b_2\right) x + \left(\lambda_1 c_1 + \lambda_2 c_2\right)$$

Mapping of the combination

$$\left(\lambda_1 a_1 + \lambda_2 a_2\right) x^2 + \left(\lambda_1 b_1 + \lambda_2 b_2\right) x + \left(\lambda_1 c_1 + \lambda_2 c_2\right) \longmapsto \begin{bmatrix} \lambda_1 a_1 + \lambda_2 a_2 + \lambda_1 b_1 + \lambda_2 b_2 \\ \lambda_1 a_1 + \lambda_2 a_2 + \lambda_1 c_1 + \lambda_2 c_2 \end{bmatrix}$$

Rewrite the mapping

$$\begin{bmatrix} \lambda_1 a_1 + \lambda_2 a_2 + \lambda_1 b_1 + \lambda_2 b_2 \\ \lambda_1 a_1 + \lambda_2 a_2 + \lambda_1 c_1 + \lambda_2 c_2 \end{bmatrix} = \begin{bmatrix} \lambda_1 \left(a_1 + b_1\right) + \lambda_2 \left(a_2 + b_2\right) \\ \lambda_1 \left(a_1 + c_1\right) + \lambda_2 \left(a_2 + c_2\right) \end{bmatrix} = \lambda_1 \begin{bmatrix} a_1 + b_1 \\ a_1 + c_1 \end{bmatrix} + \lambda_2 \begin{bmatrix} a_2 + b_2 \\ a_2 + c_2 \end{bmatrix}$$

The mapping of the linear combination of polynomials is the linear combination of the mappings of the polynomials. Thus, h is a homomorphism.

Exercise 3.4. Verify that $f : \mathbb{R}^2 \to \mathbb{R}^3$ given by

$$\begin{bmatrix} x \\ y \end{bmatrix} \longmapsto \begin{bmatrix} 0 \\ x-y \\ 3y \end{bmatrix}$$

is a *homomorphism*. Hint: check that $h(\alpha A + \beta B) = \alpha h(A) + \beta h(B)$.

Solution. Quick and dirty

$$\begin{aligned} f\left(a_1 \begin{bmatrix} x_1 \\ y_1 \end{bmatrix} + a_2 \begin{bmatrix} x_2 \\ y_2 \end{bmatrix}\right) &= f\left(\begin{bmatrix} a_1x_1 + a_2x_2 \\ a_1y_1 + a_2y_2 \end{bmatrix}\right) \\ &= \begin{bmatrix} 0 \\ a_1x_1 + a_2x_2 - a_1y_1 + a_2y_2 \\ 3\left(a_1y_1 + a_2y_2\right) \end{bmatrix} \\ &= \begin{bmatrix} 0+0 \\ a_1(x_1 - y_1) + a_2(x_2 + y_2) \\ 3a_1y_1 + 3a_2y_2 \end{bmatrix} \\ &= a_1 \begin{bmatrix} 0 \\ x_1 - y_1 \\ 3y_1 \end{bmatrix} + a_2 \begin{bmatrix} 0 \\ x_2 + y_2 \\ 3y_2 \end{bmatrix} = a_1 f\left(\begin{bmatrix} x_1 \\ y_1 \end{bmatrix}\right) + a_2 f\left(\begin{bmatrix} x_2 \\ y_2 \end{bmatrix}\right) \end{aligned}$$

f is a homomorphism.

Exercise 3.5. Assume each matrix below represents a map $h : \mathbb{R}^m \to \mathbb{R}^n$

a) $$\begin{bmatrix} 2 & 1 \\ -1 & 3 \end{bmatrix}$$

b) $$\begin{bmatrix} 0 & 1 & 3 \\ 2 & 3 & 4 \\ -2 & -1 & 2 \end{bmatrix}$$

c) $$\begin{bmatrix} 1 & 1 \\ 2 & 1 \\ 3 & 1 \end{bmatrix}$$

For each state

- m and n
- range space and rank
- null space and nullity

Hint: study/solve/analyze the system $M\vec{u} = \vec{v}$.

Solution. For (a):

$$\begin{bmatrix} 2 & 1 \\ -1 & 3 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} a \\ b \end{bmatrix}$$

The dimension of the *domain* space $\mathbb{R}^2$ is the number of columns $\boxed{m = 2}$. The dimension of the *codomain* space $\mathbb{R}^2$ is the number of rows $\boxed{n = 2}$. Next, solve the system

$$\left[\begin{array}{cc|c} 2 & 1 & a \\ -1 & 3 & b \end{array}\right] \to \cdots \to \left[\begin{array}{cc|c} 1 & 0 & \frac{3}{7}a - \frac{1}{7}b \\ 0 & 1 & \frac{1}{7}a + \frac{2}{7}b \end{array}\right]$$

Note that for any vector in the codomain (right side of the |) there is a solution in the domain (left side of the |). Thus, the range is all of the codomain $\mathcal{R}(h) = \mathbb{R}^2$. The map's rank is the dimension of the range, $\boxed{\text{rank}(h) = 2}$. By setting $a = b = 0$ the only solution is $x = y = 0$, i.e., the null space (kernel) is the trivial subspace of the domain

$$\mathcal{N}(h) = \ker(h) = \left\{ \begin{bmatrix} 0 \\ 0 \end{bmatrix} \right\}$$

and the nullity is the dimension of the null space, $\boxed{\text{nullity}(h) = 0}$.

For (b):

$$\begin{bmatrix} 0 & 1 & 3 \\ 2 & 3 & 4 \\ -2 & -1 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$$

The dimension of the *domain* space $\mathbb{R}^3$ is the number of columns $\boxed{m = 3}$. The dimension of the *codomain* space $\mathbb{R}^3$ is the number of rows $\boxed{n = 3}$. Next, solve the system

$$\left[\begin{array}{ccc|c} 0 & 1 & 3 & a \\ 2 & 3 & 4 & b \\ -2 & -1 & 2 & c \end{array}\right] \to \cdots \to \left[\begin{array}{ccc|c} 1 & 0 & -\frac{5}{2} & -\frac{3}{2}a + \frac{1}{2}b \\ 0 & 1 & 3 & a \\ 0 & 0 & 0 & -2a + b + c \end{array}\right]$$

From the last row $0 = -2a + b + c$ we get $a = (b + c)/2$, so the range is

$$\mathcal{R}(h) = \left\{ \begin{bmatrix} a \\ b \\ c \end{bmatrix} \in \mathbb{R}^3 \middle| a = \frac{b+c}{2} \right\} = \left\{ b \begin{bmatrix} \frac{1}{2} \\ 1 \\ 0 \end{bmatrix} + c \begin{bmatrix} \frac{1}{2} \\ 0 \\ 1 \end{bmatrix} \middle| b, c \in \mathbb{R} \right\}$$

The map's rank is the dimension of the range, $\boxed{\text{rank}(h) = 2}$. By setting $a = b = c = 0$

$$\left[\begin{array}{ccc|c} 1 & 0 & -\frac{5}{2} & 0 \\ 0 & 1 & 3 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right] \Longleftrightarrow \begin{bmatrix} 1 & 0 & -\frac{5}{2} \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

the solution is $x = \frac{5}{2}z, y = -3z$, i.e., the null space (kernel) is

$$\mathcal{N}(h) = \ker(h) = \left\{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} \in \mathbb{R}^3 \middle| x = \frac{5}{2}z, y = -3z \right\} = \left\{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \lambda \begin{bmatrix} \frac{5}{2} \\ -3 \\ 0 \end{bmatrix} \middle| \lambda \in \mathbb{R} \right\}$$

and the nullity is the dimension of the null space, $\boxed{\text{nullity}(h) = 1}$ (the kernel is a line in $\mathbb{R}^3$).

For (c):

$$\begin{bmatrix} 1 & 1 \\ 2 & 1 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} a \\ b \\ c \end{bmatrix}$$

The dimension of the *domain* space $\mathbb{R}^2$ is the number of columns $\boxed{m = 2}$. The dimension of the *codomain* space $\mathbb{R}^3$ is the number of rows $\boxed{n = 3}$. Next, solve the system

$$\left[\begin{array}{cc|c} 1 & 1 & a \\ 2 & 1 & b \\ 3 & 1 & c \end{array}\right] \rightarrow \cdots \rightarrow \left[\begin{array}{cc|c} 1 & 0 & -a+b \\ 0 & 1 & 2a-b \\ 0 & 0 & a-2b+c \end{array}\right]$$

From the last row $0 = a - 2b + c$ we get $a = 2b - c$, so the range is

$$\mathcal{R}(h) = \left\{ \begin{bmatrix} a \\ b \\ c \end{bmatrix} \in \mathbb{R}^3 \middle| a = 2b - c \right\} = \left\{ b \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} + c \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \middle| b, c \in \mathbb{R} \right\}$$

The map's rank is the dimension of the range, $\boxed{\operatorname{rank}(h) = 2}$. By setting $a = b = c = 0$

$$\left[\begin{array}{cc|c} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{array}\right] \iff \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

the solution is $x = y = 0$, i.e., the null space (kernel) is

$$\mathcal{N}(h) = \ker(h) = \left\{ \begin{bmatrix} 0 \\ 0 \end{bmatrix} \right\}$$

and the nullity is the dimension of the null space, $\boxed{\operatorname{nullity}(h) = 0}$.

Exercise 3.6. For each of the mappings $f : \mathbb{R}^2 \to \mathbb{R}^3$ below

a) $$\begin{bmatrix} x \\ y \end{bmatrix} \longmapsto \begin{bmatrix} x \\ y \\ x + 2y \end{bmatrix}$$

b) $$\begin{bmatrix} x \\ y \end{bmatrix} \longmapsto \begin{bmatrix} x \\ y \\ x + 2y - 1 \end{bmatrix}$$

c) $$\begin{bmatrix} x \\ y \end{bmatrix} \longmapsto \begin{bmatrix} x \\ y \\ x^2 + y^2 \end{bmatrix}$$

Determine if the *range* (i.e., right side of the arrow, the "destination") is a vector space.

Solution. For (a): The range set is a plane spanned by two vectors

$$\mathcal{R}(f) = \operatorname{span}\left(\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix} \right) = \left\{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} = a \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + b \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix} \middle| a, b \in \mathbb{R} \right\}$$

Take two elements from the set

$$u = \begin{bmatrix} x \\ y \\ x + 2y \end{bmatrix}, \ v = \begin{bmatrix} a \\ b \\ a + 2b \end{bmatrix}$$

It is easy to show that the zero element exists: make $x = y = 0$ (or $a = b = 0$) and you will get $[0, 0, 0]^\mathsf{T}$. Check closure under addition and multiplication by scalar (both at the same time)

$$\alpha u + \beta v = \begin{bmatrix} \alpha x + \beta a \\ \alpha y + \beta b \\ \alpha x + \beta a + 2(\alpha y + \beta b) \end{bmatrix} = \begin{bmatrix} ☹ \\ ☺ \\ ☹ + 2☺ \end{bmatrix} \in \mathcal{R}(f)$$

The <u>range set is a vector space</u>.

For (b): The range set is a plane

$$\mathcal{R}(f) = \left\{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} = a \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + b \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix} - \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \middle| a, b \in \mathbb{R} \right\}$$

Let's check if $\mathcal{R}(f)$ has the zero element. It doesn't because there are no solutions a, b for the system

$$\begin{aligned} x &= 0 = a \\ y &= 0 = b \\ z &= 0 = a + 2b - 1 \end{aligned}$$

The <u>range set is not a vector space</u>.

For (c): The range set is a paraboloid

$$\mathcal{R}(f) = \left\{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} a \\ b \\ a^2 + b^2 \end{bmatrix} \middle| a, b \in \mathbb{R} \right\}$$

Let's check if $\mathcal{R}(f)$ has the zero element. Make $a = b = 0$ and you will get $[0, 0, 0]^\mathsf{T}$, the zero element exists. To check closure under addition consider two vectors from $\mathcal{R}(f)$

$$u_1 = \begin{bmatrix} a_1 \\ b_1 \\ a_1^2 + b_1^2 \end{bmatrix}, \quad u_2 = \begin{bmatrix} a_2 \\ b_2 \\ a_2^2 + b_2^2 \end{bmatrix}$$

Sum the vectors

$$u_3 = u_1 + u_2 = \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} a_1 + a_2 \\ b_1 + b_2 \\ a_1^2 + b_1^2 + a_2^2 + b_2^2 \end{bmatrix}$$

If $u_3 \in \mathcal{R}(f)$ then $z = x^2 + y^2$, but this implies that

$$\begin{aligned} (a_1 + a_2)^2 + (b_1 + b_2)^2 &= a_1^2 + b_1^2 + a_2^2 + b_2^2 \\ a_1^2 + 2a_1a_2 + a_2^2 + b_1^2 + 2b_1b_2 + b_2^2 &= a_1^2 + b_1^2 + a_2^2 + b_2^2 \\ a_1a_2 + b_1b_2 &= 0 \end{aligned}$$

which is not always true. There is not closure under addition. The <u>range set is not a vector space</u>.

4 Inner products

Exercise 4.1. Show that $\langle \cdot, \cdot \rangle$ defined for all $\vec{x} = [x_1, x_2]^\mathsf{T} \in \mathbb{R}$ and $\vec{y} = [y_1, y_2]^\mathsf{T} \in \mathbb{R}$ by

$$\langle \vec{x}, \vec{y} \rangle := x_1 y_1 - (x_1 y_2 + x_2 y_1) + 2 x_2 y_2$$

is an inner product.

Solution. We need to show that $\langle \vec{x}, \vec{y} \rangle$ is

- symmetric: $\langle \vec{x}, \vec{y} \rangle = \langle \vec{y}, \vec{x} \rangle$
- positive definite: $\langle \vec{x}, \vec{x} \rangle > 0$ and $\left\langle \vec{0}, \vec{0} \right\rangle = 0$
- bi-linear: $\langle \vec{x} + \vec{y}, \vec{z} \rangle = \langle \vec{x}, \vec{z} \rangle + \langle \vec{y}, \vec{z} \rangle$ and $\langle \lambda \vec{x}, \vec{y} \rangle = \lambda \langle \vec{x}, \vec{z} \rangle$

Symmetry:

$$\begin{aligned}\langle \vec{x}, \vec{y} \rangle &= x_1 y_1 - (x_1 y_2 + x_2 y_1) + 2 x_2 y_2 \\ &= y_1 x_1 - (y_2 x_1 + y_1 x_2) + 2 y_2 x_2 = \langle \vec{y}, \vec{x} \rangle\end{aligned}$$

It is symmetric.

Positive definite:

$$\begin{aligned}\langle \vec{x}, \vec{x} \rangle &= x_1 x_1 - (x_1 x_2 + x_2 x_1) + 2 x_2 x_2 \\ &= x_1^2 - 2 x_1 x_2 + 2 x_2^2 \\ &= x_1^2 - 2 x_1 x_2 + x_2^2 + x_2^2 \\ &= (x_1 + x_2)^2 + x_2^2\end{aligned}$$

$\langle \vec{x}, \vec{x} \rangle$ is positive for $\vec{x} \neq \vec{0}$ and zero for $\vec{x} = \vec{0}$ only.

Bi-linearity: we can do both tests at the same time

$$\begin{aligned}\langle \lambda (\vec{x} + \vec{y}), \vec{z} \rangle &= \lambda (x_1 + y_1) z_1 - [\lambda (x_1 + y_1) z_2 + \lambda (x_2 + y_2) z_1] + \lambda 2 (x_2 + y_2) z_2 \\ &= \lambda \{x_1 z_1 + y_1 z_1 - [x_1 z_2 + y_1 z_2 + x_2 z_1 + y_2 z_1] + 2 x_2 z_2 + 2 y_2 z_2\} \\ &= \lambda \{x_1 z_1 - x_1 z_2 - x_2 z_1 + 2 x_2 z_2 + y_1 z_1 - y_1 z_2 - y_2 z_1 + 2 y_2 z_2\} \\ &= \lambda \{x_1 z_1 - (x_1 z_2 + x_2 z_1) + 2 x_2 z_2 + y_1 z_1 - (y_1 z_2 + y_2 z_1) + 2 y_2 z_2\} \\ &= \lambda \{x_1 z_1 - (x_1 z_2 + x_2 z_1) + 2 x_2 z_2\} + \lambda \{y_1 z_1 - (y_1 z_2 + y_2 z_1) + 2 y_2 z_2\} \\ &= \lambda \langle \vec{x}, \vec{z} \rangle + \lambda \langle \vec{y}, \vec{z} \rangle\end{aligned}$$

It is bi-linear.

Thus, we conclude that

$$\langle \vec{x}, \vec{y} \rangle := x_1 y_1 - (x_1 y_2 + x_2 y_1) + 2 x_2 y_2$$

is an inner product.

Exercise 4.2. Consider $\langle\cdot,\cdot\rangle$ defined for all $\vec{x}$ and $\vec{y}$ in $\mathbb{R}^2$ as

$$\langle\vec{x},\vec{y}\rangle := \vec{x}^\mathsf{T}\begin{bmatrix} 2 & 0 \\ 1 & 2 \end{bmatrix}\vec{y}$$

Is $\langle\cdot,\cdot\rangle$ an inner product? Is the matrix M above positive definite?

Solution. That the matrix is not symmetric gives a hint that the symmetry property is violated. Test symmetry using vectors

$$\vec{x} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \text{ and } \vec{y} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

$$\langle\vec{x},\vec{y}\rangle = \begin{bmatrix} 1 & 0 \end{bmatrix}\begin{bmatrix} 2 & 0 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \end{bmatrix}\begin{bmatrix} 0 \\ 2 \end{bmatrix} = 0$$

$$\langle\vec{y},\vec{x}\rangle = \begin{bmatrix} 0 & 1 \end{bmatrix}\begin{bmatrix} 2 & 0 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 \end{bmatrix}\begin{bmatrix} 2 \\ 1 \end{bmatrix} = 1$$

There is no symmetry, $\langle\cdot,\cdot\rangle$ <u>is not an inner product</u>.

Let's check if the matrix is positive definite, i.e., prove that $\vec{x}^\mathsf{T}M\vec{x} > 0$ and $\vec{0}^\mathsf{T}M\vec{0} = 0$

$$\vec{x}^\mathsf{T}M\vec{x} = \begin{bmatrix} x_1 & x_2 \end{bmatrix}\begin{bmatrix} 2 & 0 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 2x_1^2 + x_1x_2 + 2x_2^2$$

now let $x_2 = \lambda x_1$ so that $\vec{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_1\begin{bmatrix} 1 \\ \lambda \end{bmatrix}$. Accordingly

$$2x_1^2 + x_1\lambda x_1 + 2\lambda^2x_1^2 = x_1^2\left(2 + \lambda + 2\lambda^2\right)$$

x_1^2 cannot be negative and $2 + \lambda + 2\lambda^2$ is always positive. If $\vec{x} \neq \vec{0}$ then $\vec{x}^\mathsf{T}M\vec{x} > 0$ and if $\vec{x} = \vec{0}$ then $\vec{x}^\mathsf{T}M\vec{x} = 0$. Matrix M <u>is positive definite</u>.

Exercise 4.3. Compute the distance between vectors

$$\vec{x} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \quad \vec{y} = \begin{bmatrix} -1 \\ -1 \\ 0 \end{bmatrix}$$

using

(a) $\langle\vec{x},\vec{y}\rangle := \vec{x}^\mathsf{T}\vec{y}$

(b) $\langle\vec{x},\vec{y}\rangle := \vec{x}^\mathsf{T}A\vec{y}$ where $A = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 3 & -1 \\ 0 & -1 & 2 \end{bmatrix}$

Solution. We need to calculate the length of the vector difference

$$\vec{z} = \vec{x} - \vec{y} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} - \begin{bmatrix} -1 \\ -1 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \\ 3 \end{bmatrix}$$

Part (a):

$$\langle \vec{z}, \vec{z} \rangle = \begin{bmatrix} 2 & 3 & 4 \end{bmatrix} \begin{bmatrix} 2 \\ 3 \\ 3 \end{bmatrix} = 4 + 9 + 9 = 22$$

and the distance is $\sqrt{22} \approx 4.49$.

Part (b):

$$\langle \vec{z}, \vec{z} \rangle = \begin{bmatrix} 2 & 3 & 4 \end{bmatrix} \begin{bmatrix} 2 & 1 & 0 \\ 1 & 3 & -1 \\ 0 & -1 & 2 \end{bmatrix} \begin{bmatrix} 2 \\ 3 \\ 3 \end{bmatrix} = \begin{bmatrix} 2 & 3 & 4 \end{bmatrix} \begin{bmatrix} 7 \\ 8 \\ 3 \end{bmatrix} = 14 + 24 + 12 = 50$$

and the distance is $\sqrt{50} \approx 7.07$.

Exercise 4.4. Compute the angle between

$$\vec{x} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \text{ and } \vec{y} = \begin{bmatrix} -1 \\ -1 \end{bmatrix}$$

using

(a) $\langle \vec{x}, \vec{y} \rangle := \vec{x}^{\mathsf{T}} \vec{y}$

(b) $\langle \vec{x}, \vec{y} \rangle := \vec{x}^{\mathsf{T}} A \vec{y}$ where $A = \begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix}$

You will need

$$\cos\theta = \frac{\langle \vec{x}, \vec{y} \rangle}{\|\vec{x}\| \, \|\vec{y}\|} = \frac{\langle \vec{x}, \vec{y} \rangle}{\sqrt{\langle \vec{x}, \vec{x} \rangle} \sqrt{\langle \vec{y}, \vec{y} \rangle}}.$$

Solution. Part (a):

$$\langle \vec{x}, \vec{y} \rangle = \begin{bmatrix} 1 & 2 \end{bmatrix} \begin{bmatrix} -1 \\ -1 \end{bmatrix} = -1 - 2 = -3$$

$$\langle \vec{x}, \vec{x} \rangle = \begin{bmatrix} 1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = 1 + 4 = 5$$

$$\langle \vec{y}, \vec{y} \rangle = \begin{bmatrix} -1 & -1 \end{bmatrix} \begin{bmatrix} -1 \\ -1 \end{bmatrix} = 1 + 1 = 2$$

Then

$$\cos\theta = \frac{-3}{\sqrt{5}\sqrt{2}} = \frac{-3}{\sqrt{10}} = -\frac{3}{10}\sqrt{10} \approx -0.94868$$

$$\theta = \arccos(-0.94868) \approx 2.82\,\text{rad} \approx 161.5\,\text{deg}$$

Part (b):

$$\langle \vec{x}, \vec{y} \rangle = \begin{bmatrix} 1 & 2 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} -1 \\ -1 \end{bmatrix} = -11$$
$$\langle \vec{x}, \vec{x} \rangle = \begin{bmatrix} 1 & 2 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = 18$$
$$\langle \vec{y}, \vec{y} \rangle = \begin{bmatrix} -1 & -1 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} -1 \\ -1 \end{bmatrix} = 7$$

Then

$$\cos\theta = \frac{-11}{\sqrt{18}\sqrt{7}} = -\frac{11}{126} \approx -0.087302$$
$$\theta = \arccos(-0.087302) \approx 1.66\,\mathrm{rad} \approx 95\,\mathrm{deg}$$

5 Projections

Exercise 5.1. Project the vector orthogonaly into the line

(a) $$\begin{bmatrix} 2 \\ -1 \\ 4 \end{bmatrix}, \left\{ c \begin{bmatrix} -3 \\ 1 \\ -3 \end{bmatrix} \middle| c \in \mathbb{R} \right\}$$

(b) $$\begin{bmatrix} -1 \\ -1 \end{bmatrix}, y = 3x$$

Solution. The projection of vector $\mathbf{u}$ into vector $\mathbf{v}$ is

$$\pi_{\mathbf{v}}(\mathbf{u}) = \lambda \mathbf{v} = \left(\frac{\mathbf{u} \cdot \mathbf{v}}{\mathbf{v} \cdot \mathbf{v}} \right) \mathbf{v}$$

Part (a): Let

$$\mathbf{u} = \begin{bmatrix} 2 \\ -1 \\ 4 \end{bmatrix}, \mathbf{v} = c \begin{bmatrix} -3 \\ 1 \\ -3 \end{bmatrix}$$

Get the scale factor

$$\lambda = \frac{(2, -1, 4) \cdot (-3c, c, -3c)}{(-3c, c, -3c) \cdot (-3c, c, -3c)} = \frac{\cancel{c}(-6 - 1 - 12)}{c^{\cancel{2}}(9 + 1 + 9)} = \frac{-19}{19c} = -\frac{1}{c}$$

The projection is

$$\pi_{\mathbf{v}}(\mathbf{u}) = -\frac{1}{\cancel{c}} \cancel{c} \begin{bmatrix} -3 \\ 1 \\ -3 \end{bmatrix} = \begin{bmatrix} 3 \\ -1 \\ 3 \end{bmatrix}$$

Notice: the choice of c determines the scalar λ, but the projection $\pi_{\mathbf{v}}$ is independent of c and λ.

Part (b): Let

$$\mathbf{u} = \begin{bmatrix} -1 \\ -1 \end{bmatrix}, \mathbf{v} = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$$

The projection is

$$\pi_{\mathbf{v}}(\mathbf{u}) = \frac{(-1, -1) \cdot (1, 3)}{(1, 3) \cdot (1, 3)} \begin{bmatrix} 1 \\ 3 \end{bmatrix} = \frac{-1 - 3}{1 + 9} \begin{bmatrix} 1 \\ 3 \end{bmatrix} = \frac{-4}{10} \begin{bmatrix} 1 \\ 3 \end{bmatrix} = - \begin{bmatrix} \frac{2}{5} \\ \frac{6}{5} \end{bmatrix}$$

Exercise 5.2. In $\mathbb{R}^4$ project point $p = (1, 2, 1, 3)$ into the line $l = \{c(-1, 1, -1, 1) \,|\, c \in \mathbb{R}\}$.

Solution.

$$\begin{aligned} \pi_l(p) &= \left[\frac{(1, 2, 1, 3) \cdot (-1, 1, -1, 1)}{(-1, 1, -1, 1) \cdot (-1, 1, -1, 1)} \right] (-1, 1, -1, 1) \\ &= \left[\frac{-1 + 2 - 1 + 3}{1 + 1 + 1 + 1} \right] (-1, 1, -1, 1) = \frac{5}{4}(-1, 1, -1, 1) \\ &= \left(-\frac{5}{4}, \frac{5}{4}, -\frac{5}{4}, \frac{5}{4} \right) \end{aligned}$$

Exercise 5.3. Consider the transformation of $\mathbb{R}^2$ resulting from fixing $\mathbf{s} = [3, 1]^\mathsf{T}$ and projecting $\mathbf{v}$ into the line spanned by $\mathbf{s}$. Show that in general the projection transformation is

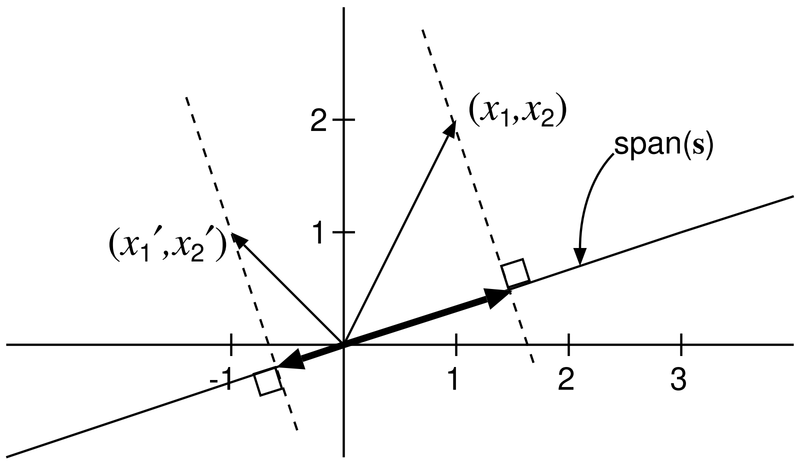
$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \mapsto \begin{bmatrix} \frac{9x_1+3x_2}{10} \\ \frac{3x_1+x_2}{10} \end{bmatrix}$$

and find the projection matrix.

Solution.

$$\pi_{\text{line}}(\mathbf{v}) = \frac{\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}^\mathsf{T} \begin{bmatrix} 3 \\ 1 \end{bmatrix}}{\begin{bmatrix} 3 \\ 1 \end{bmatrix}^\mathsf{T} \begin{bmatrix} 3 \\ 1 \end{bmatrix}} \begin{bmatrix} 3 \\ 1 \end{bmatrix} = \frac{3x_1 + x_2}{10} \begin{bmatrix} 3 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{9x_1+3x_2}{10} \\ \frac{3x_1+x_2}{10} \end{bmatrix}$$
$$= \begin{bmatrix} \frac{9x_1}{10} \\ \frac{3x_1}{10} \end{bmatrix} + \begin{bmatrix} \frac{3x_2}{10} \\ \frac{x_2}{10} \end{bmatrix} = \begin{bmatrix} \frac{9}{10} & \frac{3}{10} \\ \frac{3}{10} & \frac{1}{10} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

The picture bellow shows the projection (thick arrows) of two vectors (thin arrows) into the line spanned by $\mathbf{s}$.



Exercise 5.4. Consider the euclidean vector space $\mathbb{R}^5$ with the dot product. A subspace $U \subseteq \mathbb{R}^5$ and $\vec{x} \in \mathbb{R}^5$ given by

$$U = \text{span}\left(\begin{bmatrix} 0 \\ -1 \\ 2 \\ 0 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ -3 \\ 1 \\ -1 \\ 2 \end{bmatrix}, \begin{bmatrix} -3 \\ 4 \\ 1 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ -3 \\ 5 \\ 0 \\ 7 \end{bmatrix} \right), \vec{x} = \begin{bmatrix} -1 \\ -9 \\ -1 \\ 4 \\ 1 \end{bmatrix}$$

- determine the orthogonal projection $\pi_U(\vec{x})$
- determine the distance $d(\vec{x}, U)$

Solution. First determine a basis for U

$$\begin{bmatrix} 0 & 1 & -3 & -1 \\ -1 & -3 & 4 & -3 \\ 2 & 1 & 1 & 5 \\ 0 & -1 & 2 & 0 \\ 2 & 2 & 1 & 7 \end{bmatrix} \to \cdots \to \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

There are only three independent vectors. Pick the first three for the basis so that

$$U = \text{span}\left(\begin{bmatrix} 0 \\ -1 \\ 2 \\ 0 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ -3 \\ 1 \\ -1 \\ 2 \end{bmatrix}, \begin{bmatrix} -3 \\ 4 \\ 1 \\ 2 \\ 1 \end{bmatrix} \right)$$

In other words U could be spanned by the columns of this matrix

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

The projection $\pi_U(\vec{x}) = \vec{p}$ is some vector $\vec{p} = B\vec{\lambda} \in \mathbb{R}^5$. We need a $\vec{\lambda} = [\lambda_1, \lambda_2, \lambda_3]^\mathsf{T} \neq \vec{0}$ in $\in \mathbb{R}^3$. How? Let

$$\vec{x} = \vec{p} + \vec{q} \,\text{and}\, \vec{p} \perp \vec{q}$$

this means

$$\begin{aligned} \vec{p} \cdot (\vec{x} - \vec{p}) &= 0 && (\vec{p} \perp \vec{q}) \\ (B\vec{\lambda})^\mathsf{T}(\vec{x} - B\vec{\lambda}) &= 0 && (\vec{p} = B\vec{\lambda}) \\ \vec{\lambda}^\mathsf{T} B^\mathsf{T}(\vec{x} - B\vec{\lambda}) &= 0 \\ B^\mathsf{T}(\vec{x} - B\vec{\lambda}) &= \vec{0} && (\vec{\lambda} \neq \vec{0}) \\ B^\mathsf{T}\vec{x} - B^\mathsf{T}B\vec{\lambda} &= \vec{0} \\ B^\mathsf{T}B\vec{\lambda} &= B^\mathsf{T}\vec{x} \end{aligned}$$

In practice

$$\underbrace{\begin{bmatrix} 0 & -1 & 2 & 0 & 2 \\ 1 & -3 & 1 & -1 & 2 \\ -3 & 4 & 1 & 2 & 1 \end{bmatrix}}_{B^\mathsf{T}} \underbrace{\begin{bmatrix} 0 & 1 & -3 \\ -1 & -3 & 4 \\ 2 & 1 & 1 \\ 0 & -1 & 2 \\ 2 & 2 & 1 \end{bmatrix}}_{B} \underbrace{\begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \end{bmatrix}}_{\vec{\lambda}} = \begin{bmatrix} 0 & -1 & 2 & 0 & 2 \\ 1 & -3 & 1 & -1 & 2 \\ -3 & 4 & 1 & 2 & 1 \end{bmatrix} \underbrace{\begin{bmatrix} 1 \\ -3 \\ 1 \\ -1 \\ 2 \end{bmatrix}}_{\vec{x}}$$

$$\vec{\lambda} = \underbrace{\left(B^\mathsf{T}B\right)^{-1} B^\mathsf{T}}_{\text{pseudoinverse}} \vec{x}$$

The scalars are $\lambda_1 = -3, \lambda_1 = 4, \lambda_1 = 1$. The orthogonal projection of $\vec{x}$ into U is

$$\vec{p} = \pi_U(\vec{x}) = B\vec{\lambda} = \begin{bmatrix} 0 & 1 & -3 \\ -1 & -3 & 4 \\ 2 & 1 & 1 \\ 0 & -1 & 2 \\ 2 & 2 & 1 \end{bmatrix} \begin{bmatrix} -3 \\ 4 \\ 1 \end{bmatrix} = \boxed{\begin{bmatrix} 1 \\ -5 \\ -1 \\ -2 \\ 3 \end{bmatrix}}$$

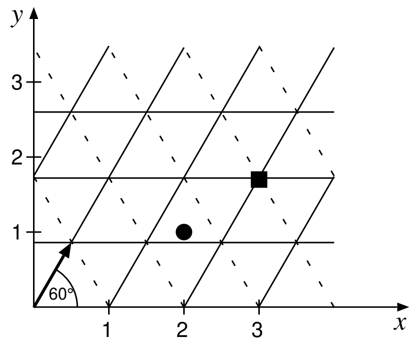
The distance between $\vec{x}$ and U is the distance between $\vec{x}$ and $\vec{p}$

$$d(\vec{x}, U) = ||\vec{x} - \vec{p}||$$

$$\vec{x} - \vec{p} = \begin{bmatrix} 1 \\ -3 \\ 1 \\ -1 \\ 2 \end{bmatrix} - \begin{bmatrix} 1 \\ -5 \\ -1 \\ -2 \\ 3 \end{bmatrix} = \begin{bmatrix} -2 \\ -4 \\ 0 \\ 6 \\ -2 \end{bmatrix}$$

$$d(\vec{x}, U) = \sqrt{2^2 + 4^4 + 6^2 + 2^2} = \boxed{\sqrt{60}}$$

Exercise 5.5. Project the "round point" $(x, y) = (2, 1)$ into the "rhomboid grid" shown below. The length of the vector forming the $60°$ angle is 1.



Find the (x, y) of "squarish point". Find the transformation projecting "from rectangular to rhomboid".

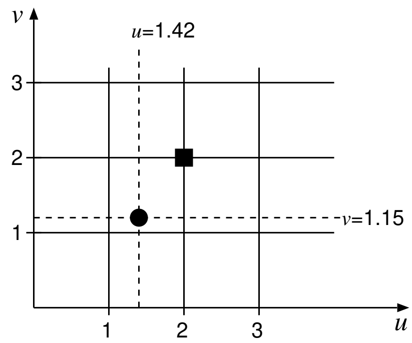
Solution. The coordinates of the unit vector are $x = \cos 60° = \frac{1}{2}$ and $y = \sin 60° = \frac{\sqrt{3}}{2}$. You can use the vectors $\left(\frac{1}{2}, \frac{\sqrt{3}}{2}\right)$ and $(1, 0)$ as basis for the rhomboid grid

$$\begin{bmatrix} 2 \\ 1 \end{bmatrix} = u \begin{bmatrix} 1 \\ 0 \end{bmatrix} + v \begin{bmatrix} \frac{1}{2} \\ \frac{\sqrt{3}}{2} \end{bmatrix}$$

We can solve

$$u + \frac{1}{2}v = 2$$
$$\frac{\sqrt{3}}{2}v = 1$$

$u = \frac{6-\sqrt{3}}{3} \approx 1.42$ and $v = \frac{2\sqrt{3}}{3} \approx 1.15$, these are the "rhomboid coordinates"



The squarish point lies at $(u, v) = (2, 2)$ and its canonical coordinates are

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 & \frac{1}{2} \\ 0 & \frac{\sqrt{3}}{2} \end{bmatrix} \begin{bmatrix} 2 \\ 2 \end{bmatrix} = \begin{bmatrix} 3 \\ \sqrt{3} \end{bmatrix}$$

If the transformation from rhomboid to rectangular is

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1 & \frac{1}{2} \\ 0 & \frac{\sqrt{3}}{2} \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix}$$

the inverse transformation from rectangular to rhomboid requires the matrix inverse

$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} 1 & -\frac{1}{3}\sqrt{3} \\ 0 & \frac{2}{3}\sqrt{3} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

Exercise 5.6. Use Gram–Schmidt orthogonalization to create a basis for $\mathbb{R}^2$ using

$$\left\langle \underbrace{\begin{bmatrix} 1 \\ 1 \end{bmatrix}}_{\vec{b}_1}, \underbrace{\begin{bmatrix} -1 \\ 2 \end{bmatrix}}_{\vec{b}_2} \right\rangle$$

Solution. Let

$$\vec{b}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \vec{b}_2 = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$$

It is easier to orthogonalize with respect to $\vec{b}_1$. The vectors for the new basis will be

$$\vec{u}_1 = \vec{b}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \vec{u}_2 = \pi_2(\vec{b}_2)$$

so that

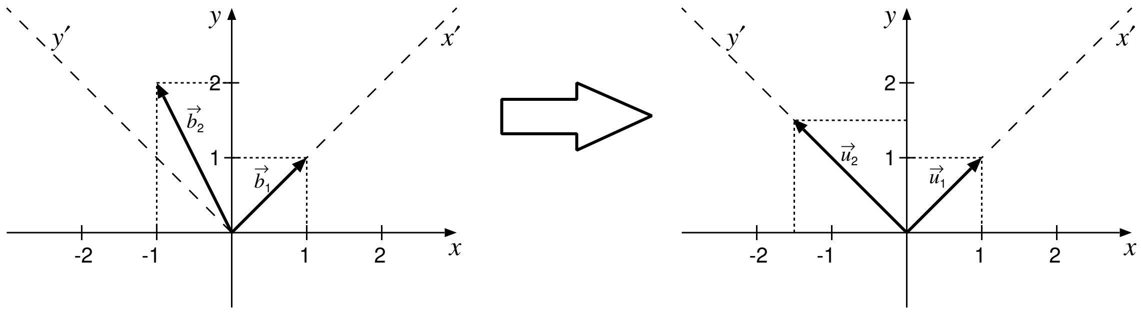
$$\vec{b}_2 = \pi_1(\vec{b}_2) + \pi_2(\vec{b}_2) \Longleftrightarrow \pi_2(\vec{b}_2) = \vec{b}_2 - \pi_1(\vec{b}_2)$$

Thus,

$$\vec{u}_2 = \vec{b}_2 - \left(\frac{\vec{b}_2 \cdot \vec{b}_1}{\vec{b}_1 \cdot \vec{b}_1} \right) \vec{b}_1 = \begin{bmatrix} -1 \\ 2 \end{bmatrix} - \frac{(-1, 2) \cdot (1, 1)}{(1, 1) \cdot (1, 1)} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -\frac{3}{2} \\ \frac{3}{2} \end{bmatrix}$$

The orthogonal basis is

$$\boxed{\left\langle \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -\frac{3}{2} \\ \frac{3}{2} \end{bmatrix} \right\rangle}$$



You can check that $\vec{u}_1$ and $\vec{u}_2$ are indeed orthogonal

$$\vec{u}_1 \cdot \vec{u}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} -\frac{3}{2} \\ \frac{3}{2} \end{bmatrix} = -\frac{3}{2} + \frac{3}{2} = 0$$

Exercise 5.7. Find an orthonormal basis for this subspace of $\mathbb{R}^3$: the plane $x - y + z = 0$.

Solution. First thing to do is to parameterize the plane

$$\left\{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} \in \mathbb{R}^3 \middle| x = y - z \right\} = \left\{ y \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + z \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \middle| y, z \in \mathbb{R} \right\}$$

Let's make

$$\vec{b}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$$

one of our basis vectors. Use the Gram–Schmidt method to find our second orthogonal basis vector

$$\vec{b}_2 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} - \left(\frac{[1,1,0] \cdot [-1,0,1]}{[-1,0,1] \cdot [-1,0,1]} \right) \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} + \frac{1}{2} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} -\frac{1}{2} \\ \frac{1}{2} \\ 1 \end{bmatrix}$$

Next, we must normalize $\vec{b}_1$ and $\vec{b}_2$

$$\vec{u}_1 = \frac{\vec{b}_1}{\left\| \vec{b}_1 \right\|} = \frac{1}{\sqrt{1^2 + 1^2}} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix}$$

$$\vec{u}_2 = \frac{\vec{b}_2}{\left\| \vec{b}_2 \right\|} = \frac{1}{\sqrt{(-\frac{1}{2})^2 + (\frac{1}{2})^2 + 1^2}} \begin{bmatrix} -\frac{1}{2} \\ \frac{1}{2} \\ 1 \end{bmatrix} = \begin{bmatrix} -\frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \\ \frac{2}{\sqrt{6}} \end{bmatrix}$$

The orthonormal basis is

$$\boxed{\left\langle \begin{bmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix}, \begin{bmatrix} -\frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \\ \frac{2}{\sqrt{6}} \end{bmatrix} \right\rangle}$$

Exercise 5.8. Rotate the vectors

$$\mathbf{x}_1 = \begin{bmatrix} 2 \\ 3 \end{bmatrix}, \mathbf{x}_2 = \begin{bmatrix} 0 \\ -1 \end{bmatrix}$$

by 30°.

Solution. We use the rotation matrix

$$\mathbf{A} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

with $\theta = 30° = \frac{\pi}{6}$: $\cos(\frac{\pi}{6}) = \frac{\sqrt{3}}{2}$ and $\sin(\frac{\pi}{6}) = \frac{1}{2}$

$$\mathbf{r}_1 = \begin{bmatrix} \frac{\sqrt{3}}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{\sqrt{3}}{2} \end{bmatrix} \begin{bmatrix} 2 \\ 3 \end{bmatrix} = \begin{bmatrix} \sqrt{3} - \frac{3}{2} \\ 1 + \frac{\sqrt{3}}{6} \end{bmatrix} = \begin{bmatrix} \frac{2\sqrt{3}-3}{2} \\ \frac{6+\sqrt{3}}{6} \end{bmatrix} = \boxed{\frac{1}{6} \begin{bmatrix} 6\sqrt{3} - 9 \\ 6 + \sqrt{3} \end{bmatrix}}$$

$$\mathbf{r}_2 = \begin{bmatrix} \frac{\sqrt{3}}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{\sqrt{3}}{2} \end{bmatrix} \begin{bmatrix} 0 \\ -1 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} \\ -\frac{\sqrt{3}}{2} \end{bmatrix} = \boxed{\frac{1}{2} \begin{bmatrix} 1 \\ -\sqrt{3} \end{bmatrix}}$$

6 Eigenvectors and eigenvalues

Exercise 6.1. Compute the determinant of

$$A = \begin{bmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \\ 0 & 2 & 4 \end{bmatrix}$$

using the Laplace expansion, the rule of Sarrus and row operations.

Solution. Determinant by Laplace expansion of the 1st, 2nd or 3rd row

$$\begin{vmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \\ 0 & 2 & 4 \end{vmatrix} = \begin{cases} \mathbf{1} \cdot \begin{vmatrix} 4 & 6 \\ 2 & 4 \end{vmatrix} - \mathbf{3} \begin{vmatrix} 2 & 6 \\ 0 & 4 \end{vmatrix} + \mathbf{5} \begin{vmatrix} 2 & 4 \\ 0 & 2 \end{vmatrix} & = 1 \cdot (16 - 12) - 3 \cdot (8 - 0) + 5 \cdot (4 - 0) = 0 \\ \mathbf{-2} \cdot \begin{vmatrix} 3 & 5 \\ 4 & 6 \end{vmatrix} + \mathbf{4} \begin{vmatrix} 1 & 5 \\ 0 & 4 \end{vmatrix} - \mathbf{6} \begin{vmatrix} 1 & 3 \\ 0 & 2 \end{vmatrix} & = -2 \cdot (18 - 20) + 4 \cdot (4 - 0) - 6 \cdot (3 - 0) = 0 \\ \mathbf{0} \cdot \begin{vmatrix} 3 & 5 \\ 4 & 6 \end{vmatrix} - \mathbf{2} \begin{vmatrix} 1 & 5 \\ 2 & 6 \end{vmatrix} + \mathbf{4} \begin{vmatrix} 1 & 3 \\ 2 & 4 \end{vmatrix} & = -2 \cdot (6 - 10) + 4 \cdot (4 - 6) = 0 \end{cases}$$

Determinant by the rule of Sarrus: write out the first two columns of the matrix to the right of the third column, giving five columns in a row. Then add the products of the diagonals going from top to bottom (solid) and subtract the products of the diagonals going from bottom to top (dashed)

$$\begin{array}{ccc:cc} + & + & + & & \\ 1 & 3 & 5 & 1 & 3 \\ 2 & 4 & 6 & 2 & 4 \\ 0 & 2 & 4 & 0 & 2 \\ - & - & - & & \end{array}$$

$$\det(A) = 1 \cdot 4 \cdot 4 + 3 \cdot 6 \cdot 0 + 5 \cdot 2 \cdot 2 - (0 \cdot 4 \cdot 5 + 2 \cdot 6 \cdot 1 + 4 \cdot 2 \cdot 3) = 0$$

Determinant by row operations

$$\begin{vmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \\ 0 & 2 & 4 \end{vmatrix} \begin{matrix} \\ -2R_1 \\ \\ \end{matrix} = \begin{vmatrix} 1 & 3 & 5 \\ 0 & -2 & -4 \\ 0 & 2 & 4 \end{vmatrix} \begin{matrix} \\ \\ +R_2 \end{matrix} = \begin{vmatrix} 1 & 3 & 5 \\ 0 & -2 & -4 \\ 0 & 0 & 0 \end{vmatrix} = 1 \times (-2) \times 0 = 0$$

$$\boxed{\det(A) = \begin{vmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \\ 0 & 2 & 4 \end{vmatrix} = 0}$$

Exercise 6.2. Compute the following determinant efficiently

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

Solution. Perform row operations to get the matrix into echelon form

$$\begin{vmatrix} 2 & 0 & 1 & 2 & 0 \\ 2 & -1 & 0 & 1 & 1 \\ 0 & 1 & 2 & 1 & 2 \\ -2 & 0 & 2 & -1 & 2 \\ 2 & 0 & 0 & 1 & 1 \end{vmatrix} \begin{matrix} \\ -R_1 \\ \\ +R_1 \\ -R_1 \end{matrix} = \begin{vmatrix} 2 & 0 & 1 & 2 & 0 \\ 0 & -1 & -1 & -1 & 1 \\ 0 & 1 & 2 & 1 & 2 \\ 0 & 0 & 3 & 1 & 2 \\ 0 & 0 & -1 & -1 & 1 \end{vmatrix} \begin{matrix} \\ \\ +R_2 \\ \\ \\ \end{matrix}$$

$$= \begin{vmatrix} 2 & 0 & 1 & 2 & 0 \\ 0 & -1 & -1 & -1 & 1 \\ 0 & 0 & 1 & 0 & 3 \\ 0 & 0 & 3 & 1 & 2 \\ 0 & 0 & -1 & -1 & 1 \end{vmatrix} \begin{matrix} \\ \\ \\ -3R_3 \\ +R_3 \end{matrix}$$

$$= \begin{vmatrix} 2 & 0 & 1 & 2 & 0 \\ 0 & -1 & -1 & -1 & 1 \\ 0 & 0 & 1 & 0 & 3 \\ 0 & 0 & 0 & 1 & -7 \\ 0 & 0 & 0 & -1 & 4 \end{vmatrix} \begin{matrix} \\ \\ \\ \\ +R_4 \end{matrix}$$

$$= \begin{vmatrix} 2 & 0 & 1 & 2 & 0 \\ 0 & -1 & -1 & -1 & 1 \\ 0 & 0 & 1 & 0 & 3 \\ 0 & 0 & 0 & 1 & -7 \\ 0 & 0 & 0 & 0 & -3 \end{vmatrix}$$

$$= 2 \times (-1) \times 1 \times 1 \times (-3) = \boxed{6}$$

Exercise 6.3. Compute the eigenspaces of

(a) $$A = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$$

(b) $$B = \begin{bmatrix} -2 & 2 \\ 2 & 1 \end{bmatrix}$$

Solution. For this we need to solve the characteristic equation, then the eigenvalue equation.

Part (a) Characteristic equation

$$\det\left(\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right) = \begin{vmatrix} 1-\lambda & 0 \\ 1 & 1-\lambda \end{vmatrix} = (1-\lambda)^2 = 0$$

There is only one solution $\lambda = 1$ with algebraic multiplicity 2. Eigenvalue equation

$$\left(\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} - \mathbf{1}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right)\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$x\begin{bmatrix} 0 \\ 1 \end{bmatrix} + y\begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

See that x must be zero but y can be any number. Let's set $x = 0$ and $y = 1$ for the eigenvector. The eigenspace will be

$$\boxed{E_1 = \text{span}\left(\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right)}$$

Part (b) Characteristic equation

$$\det\left(\begin{bmatrix} -2 & 2 \\ 2 & 1 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right) = \begin{vmatrix} 2-\lambda & 2 \\ 2 & 1-\lambda \end{vmatrix} = (\lambda - 2)(\lambda + 3) = 0$$

The eigenvalues are $\lambda_1 = 2$ and $\lambda_2 = -3$. The eigenvector associated with $\lambda_1 = 2$ is the solution of

$$\begin{bmatrix} -2-\mathbf{2} & 2 \\ 2 & 1-\mathbf{2} \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Rightarrow \begin{bmatrix} -4 & 2 \\ 2 & -1 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Rightarrow y = 2x \Rightarrow \mathbf{v}_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

and the corresponding eigenspace of $\lambda_1 = 2$ is $E_1 = \text{span}\left(\begin{bmatrix} 1 \\ 2 \end{bmatrix}\right)$.
The eigenvector associated with $\lambda_2 = -3$ is the solution of

$$\begin{bmatrix} -2-(-\mathbf{3}) & 2 \\ 2 & 1-(-\mathbf{3}) \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Rightarrow x = -2y \Rightarrow \mathbf{v}_2 = \begin{bmatrix} 2 \\ -1 \end{bmatrix}$$

and the corresponding eigenspace of $\lambda_2 = -3$ is $\boxed{E_2 = \text{span}\left(\begin{bmatrix} 2 \\ -1 \end{bmatrix}\right)}$

Exercise 6.4. Consider the map

$$\begin{bmatrix} x \\ y \end{bmatrix}_{k+1} = \begin{bmatrix} 1 & 0 \\ \frac{1}{4} & -2 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}_k$$

Given $[x, y]_0^\top = [1, 1]^\top$ calculate $[x, y]_3^\top$. What happens if $k \to \infty$?

Solution. We can do this

$$\begin{bmatrix} x \\ y \end{bmatrix}_1 = \begin{bmatrix} 1 & 0 \\ \frac{1}{4} & -2 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}_0 = \begin{bmatrix} 1 & 0 \\ \frac{1}{4} & -2 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ -\frac{7}{4} \end{bmatrix}$$
$$\begin{bmatrix} x \\ y \end{bmatrix}_2 = \begin{bmatrix} 1 & 0 \\ \frac{1}{4} & -2 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}_1 = \begin{bmatrix} 1 & 0 \\ \frac{1}{4} & -2 \end{bmatrix}\begin{bmatrix} 1 \\ -\frac{7}{4} \end{bmatrix} = \begin{bmatrix} 1 \\ \frac{15}{4} \end{bmatrix}$$
$$\begin{bmatrix} x \\ y \end{bmatrix}_3 = \begin{bmatrix} 1 & 0 \\ \frac{1}{4} & -2 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}_2 = \begin{bmatrix} 1 & 0 \\ \frac{1}{4} & -2 \end{bmatrix}\begin{bmatrix} 1 \\ \frac{15}{4} \end{bmatrix} = \boxed{\begin{bmatrix} 1 \\ \frac{-29}{4} \end{bmatrix}}$$

But not a good indea for k in general because that requires $\mathbf{v}_k = \mathbf{M}^k\mathbf{v}_0$ and multiplying many matrices is not efficient. Instead, decompose the matrix $\mathbf{M} = \mathbf{PDP}^{-1}$, where $\mathbf{D}$ is the diagonal matrix of eigenvalues and $\mathbf{P}$ the matrix with corresponding eigenvectors as columns. Then

$$\mathbf{M}^k = \left(\mathbf{PDP}^{-1}\right)^k = \underbrace{\mathbf{PD}\cancel{\mathbf{P}^{-1}}\cancel{\mathbf{P}}\mathbf{DP}^{-1}\cdots\mathbf{PD}\cancel{\mathbf{P}^{-1}}\cancel{\mathbf{P}}\mathbf{DP}^{-1}}_{k\text{ times}} = \mathbf{PD}^k\mathbf{P}^{-1},$$

diagonal matrices are easier to raise to powers.

For this matrix one can read the eigenvalues from the diagonal (it is triangular!) $\lambda_1 = 1, \lambda_2 = -2$.

For the eigenvectors of $\lambda_1 = 1$ solve

$$\left(\begin{bmatrix} 1 & 0 \\ \frac{1}{4} & -2 \end{bmatrix} - 1\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right)\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Rightarrow \begin{bmatrix} 0 & 0 \\ \frac{1}{4} & -3 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Rightarrow \mathbf{v}_1 = \begin{bmatrix} 12 \\ 1 \end{bmatrix}.$$

For the eigenvectors of $\lambda_2 = -2$ solve

$$\left(\begin{bmatrix} 1 & 0 \\ \frac{1}{4} & -2 \end{bmatrix} - (-2)\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right)\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Rightarrow \begin{bmatrix} 3 & 0 \\ \frac{1}{4} & 0 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Rightarrow \mathbf{v}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

The mapping matrix can be factored as

$$\mathbf{M} = \underbrace{\begin{bmatrix} 12 & 0 \\ 1 & 1 \end{bmatrix}}_{\mathbf{P}}\underbrace{\begin{bmatrix} 1 & 0 \\ 0 & -2 \end{bmatrix}}_{\mathbf{D}}\underbrace{\frac{1}{12}\begin{bmatrix} 1 & 0 \\ -1 & 12 \end{bmatrix}}_{\mathbf{P}^{-1}}.$$

Check for $k = 3$

$$\begin{aligned}
\begin{bmatrix} x \\ y \end{bmatrix}_3 &= \begin{bmatrix} 1 & 0 \\ \frac{1}{4} & -2 \end{bmatrix}^3 \begin{bmatrix} x \\ y \end{bmatrix}_0 = \underbrace{\begin{bmatrix} 12 & 0 \\ 1 & 1 \end{bmatrix}}_{\mathbf{P}}\underbrace{\begin{bmatrix} 1 & 0 \\ 0 & (-2)^3 \end{bmatrix}}_{\mathbf{D}^3}\underbrace{\frac{1}{12}\begin{bmatrix} 1 & 0 \\ -1 & 12 \end{bmatrix}}_{\mathbf{P}^{-1}}\begin{bmatrix} 1 \\ 1 \end{bmatrix} \\
&= \frac{1}{12}\begin{bmatrix} 12 & 0 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & -8 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ -1 & 12 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \frac{1}{12}\begin{bmatrix} 12 & 0 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 8 & -8 \cdot 12 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix} \\
&= \frac{1}{12}\begin{bmatrix} 12 & 0 \\ 9 & -8 \cdot 12 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ \frac{3}{4} & -8 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ \frac{3}{4} - 8 \end{bmatrix} = \begin{bmatrix} 1 \\ \frac{-29}{4} \end{bmatrix}
\end{aligned}$$

But if $k \to \infty$

$$\begin{aligned}
\begin{bmatrix} x \\ y \end{bmatrix}_k &= \begin{bmatrix} 1 & 0 \\ \frac{1}{4} & -2 \end{bmatrix}^k \begin{bmatrix} x \\ y \end{bmatrix}_0 = \underbrace{\begin{bmatrix} 12 & 0 \\ 1 & 1 \end{bmatrix}}_{\mathbf{P}}\underbrace{\begin{bmatrix} 1 & 0 \\ 0 & (-2)^k \end{bmatrix}}_{\mathbf{D}^k}\underbrace{\frac{1}{12}\begin{bmatrix} 1 & 0 \\ -1 & 12 \end{bmatrix}}_{\mathbf{P}^{-1}}\begin{bmatrix} 1 \\ 1 \end{bmatrix} \\
&= \frac{1}{12}\begin{bmatrix} 12 & 0 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & (-2)^k \end{bmatrix}\begin{bmatrix} 1 & 0 \\ -1 & 12 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \frac{1}{12}\begin{bmatrix} 12 & 0 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0 & (-2)^k \end{bmatrix}\begin{bmatrix} 1 \\ 11 \end{bmatrix} \\
&= \frac{1}{12}\begin{bmatrix} 12 & 0 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 11(-2)^k \end{bmatrix} = \frac{1}{12}\begin{bmatrix} 12 \\ 1 + 11(-2)^k \end{bmatrix} = \frac{1}{12}\begin{bmatrix} 12 \\ 1 + 11(-1)^k 2^k \end{bmatrix}
\end{aligned}$$

Thus,

$$\lim_{k\to\infty} \frac{1}{12}\begin{bmatrix} 12 \\ 1 + 11(-1)^k 2^k \end{bmatrix} = \frac{1}{12}\begin{bmatrix} \lim_{k\to\infty} 12 \\ 1 + 11\lim_{k\to\infty}(-1)^k 2^k \end{bmatrix} = \frac{1}{12}\begin{bmatrix} 12 \\ \text{diverges} \end{bmatrix} = \boxed{\begin{bmatrix} 1 \\ \text{diverges} \end{bmatrix}}$$

and the limit does not exist.

Exercise 6.5. Picture an animal that can live 4 years. Survival from ages $1 \to 2$, $2 \to 3$ and $3 \to 4$ occur with probability 0.6, 0.4 and 0.3 respectively. An individual of age 2 or 3 produces 80 and 50 eggs respectively, but only a fraction $s = 0.1$ survive predation. Let n_1, n_2, n_3, n_4 be the number of individuals of each age class. The same numbers one year later are n_1', n_2', n_3', n_4'. If a population starts with a $n_1 = n_2 = n_4 = 0, n_3 = 10$, what are the numbers after $t = 20$ years? What is the stable age-structure?

Solution. We start by writing the dependences of n'_1, n'_2, n'_3, n'_4 on the n_1, n_2, n_3, n_4

$$\begin{aligned} n'_1 &= s80n_2 + s50n_3 \\ n'_2 &= 0.6n_1 \\ n'_3 &= 0.4n_2 \\ n'_4 &= 0.3n_3 \end{aligned}$$

This mapping can be represented as

$$\underbrace{\begin{bmatrix} n'_1 \\ n'_2 \\ n'_3 \\ n'_4 \end{bmatrix}}_{\mathbf{N}'} = \underbrace{\begin{bmatrix} 0 & 80s & 50s & 0 \\ 0.6 & 0 & 0 & 0 \\ 0 & 0.4 & 0 & 0 \\ 0 & 0 & 0.3 & 0 \end{bmatrix}}_{\mathbf{L}} \underbrace{\begin{bmatrix} n_1 \\ n_2 \\ n_3 \\ n_4 \end{bmatrix}}_{\mathbf{N}}$$

$\mathbf{N}, \mathbf{N}'$ are population structure vectors, $\mathbf{L}$ is a "Leslie matrix". If the population vector in year 0 is $\mathbf{N}_0$

$$\begin{aligned} \mathbf{N}_1 &= \mathbf{L}\mathbf{N}_0 \\ \mathbf{N}_2 &= \mathbf{L}\mathbf{N}_1 = \mathbf{L}\mathbf{L}\mathbf{N}_0 \\ \mathbf{N}_3 &= \mathbf{L}\mathbf{N}_3 = \mathbf{L}\mathbf{L}\mathbf{L}\mathbf{N}_0 \\ &\vdots \\ \mathbf{N}_t &= \mathbf{L}^t\mathbf{N}_0 \end{aligned}$$

$$\begin{bmatrix} n_{1,t} \\ n_{2,t} \\ n_{3,t} \\ n_{4,t} \end{bmatrix} = \begin{bmatrix} 0 & 80s & 50s & 0 \\ 0.6 & 0 & 0 & 0 \\ 0 & 0.4 & 0 & 0 \\ 0 & 0 & 0.3 & 0 \end{bmatrix}^t \begin{bmatrix} n_{1,0} \\ n_{2,0} \\ n_{3,0} \\ n_{4,0} \end{bmatrix}$$

We have to diagonalize the Leslie matrix as $\mathbf{L} = \mathbf{VRW}$ where $\mathbf{R}$ is the diagonal matrix of eigenvalues of $\mathbf{L}$, $\mathbf{V}$ is the a matrix of eigenvectors whose columns correspond to eigenvalues in $\mathbf{R}$, and $\mathbf{W} = \mathbf{V}^{-1}$. Using $s = 0.1$ Octave or Matlab gives the following eigenvalues

$$\mathbf{R} = \begin{bmatrix} 0.0000 & 0.0000 & 0.0000 & 0.0000 \\ 0.0000 & \mathbf{2.3066} & 0.0000 & 0.0000 \\ 0.0000 & 0.0000 & -2.0532 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 & -0.2534 \end{bmatrix},$$

eigenvectors (columns)

$$\mathbf{V} = \begin{bmatrix} 0.0000 & \mathbf{-0.9669} & 0.9584 & -0.1578 \\ 0.0000 & \mathbf{-0.2515} & -0.2801 & 0.3736 \\ 0.0000 & \mathbf{-0.0436} & 0.0546 & -0.5898 \\ 1.0000 & \mathbf{-0.0057} & -0.0080 & 0.6983 \end{bmatrix},$$

and

$$\mathbf{W} = \begin{bmatrix} -0.0600 & 0.0000 & 1.2000 & 1.0000 \\ \mathbf{-0.4930} & \mathbf{-1.8954} & \mathbf{-1.0688} & \mathbf{0.0000} \\ 0.5606 & -1.9182 & -1.3651 & 0.0000 \\ 0.0883 & -0.0373 & -1.7427 & 0.0000 \end{bmatrix}.$$

For $\mathbf{N}_0 = [0, 0, 10, 0]$ the population after $t = 20$ years is

$$\mathbf{N}_{20} = \mathbf{L}^{20}\mathbf{N}_0 = \mathbf{P}\mathbf{D}^{20}\mathbf{W}\mathbf{N}_0 = \begin{bmatrix} 1.6456 \times 10^8 \\ 5.5613 \times 10^7 \\ 7.1492 \times 10^6 \\ 1.2944 \times 10^6 \end{bmatrix} \approx 10^6 \times \begin{bmatrix} 164.7 \\ 55.6 \\ 7.1 \\ 1.3 \end{bmatrix}$$

For very large t

$$\mathbf{N}_t \approx \mathbf{v}_d \boldsymbol{\lambda}_d^t \mathbf{w}_d \mathbf{N}_0 = \mathbf{2.3066}^t \begin{bmatrix} -\mathbf{0.9669} \\ -\mathbf{0.2515} \\ -\mathbf{0.0436} \\ -\mathbf{0.0057} \end{bmatrix} \begin{bmatrix} -\mathbf{0.4930} & -\mathbf{1.8954} & -\mathbf{1.0688} & \mathbf{0.0000} \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 10 \\ 0 \end{bmatrix}$$

$$\approx 2.3066^t \begin{bmatrix} 10.333 \\ 2.688 \\ 0.466 \\ 0.061 \end{bmatrix}$$

where $\boldsymbol{\lambda}_d$ is the dominant eigenvalue and $\mathbf{v}_d, \mathbf{w}_d$ are associated dominant left and right eigenvectors. The stable age-structure is given by the dominant eigenvector

$$\begin{bmatrix} -\mathbf{0.9669} \\ -\mathbf{0.2515} \\ -\mathbf{0.0436} \\ -\mathbf{0.0057} \end{bmatrix} \Rightarrow \text{divide by 1st element, multiply by 10000, round to integer} \Rightarrow \begin{bmatrix} 10000 \\ 2601 \\ 451 \\ 59 \end{bmatrix}$$

Scaling by 10000 is convenient for communication because it says e.g., “there are 59 individuals of age 4 per 10000 individuals of age 1”.

7 Vector calculus

Exercise 7.1. Find the gradient and jacobian of $f(\mathbf{x}) = \sin(x_1)\cos(x_2)$, $\mathbf{x} \in \mathbb{R}^2$.

Solution. We need to get the partial derivatives with respect to the components of $\mathbf{x}$

$$\frac{\partial f}{\partial x_1} = \cos(x_1)\cos(x_2)$$
$$\frac{\partial f}{\partial x_2} = -\sin(x_1)\sin(x_2)$$

The gradient is the vector of partials

$$\nabla f = \begin{bmatrix} \frac{\partial f}{\partial x_1} & \frac{\partial f}{\partial x_2} \end{bmatrix} = \boxed{\begin{bmatrix} \cos(x_1)\cos(x_2) & -\sin(x_1)\sin(x_2) \end{bmatrix}}$$

the gradient is a row vector. The jacobian is a column vector of gradients. Since there is only one gradient, the jacobian is just the gradient

$$\mathbf{J} = \nabla f = \boxed{\begin{bmatrix} \cos(x_1)\cos(x_2) & -\sin(x_1)\sin(x_2) \end{bmatrix}}$$

Exercise 7.2. Find the gradient and jacobian of $f(\mathbf{x}, \mathbf{y}) = \mathbf{x}^\mathsf{T}\mathbf{y}$, $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$.

Solution. This function is a scalar multiplication of vectors, the dot product

$$f(\mathbf{x}, \mathbf{y}) = \mathbf{x}^\mathsf{T}\mathbf{y} = \sum_{i=1}^{n} x_i y_i$$

Note that $f : \mathbb{R}^{2n} \mapsto \mathbb{R}$. The partial derivatives with respect to $\mathbf{x}$ and to$\mathbf{y}$

$$\frac{\partial f}{\partial \mathbf{x}} = \begin{bmatrix} \frac{\partial f}{\partial x_1} & \dots & \frac{\partial f}{\partial x_n} \end{bmatrix} = \begin{bmatrix} y_1 & \dots & y_n \end{bmatrix} = \mathbf{y}^\mathsf{T} \in \mathbb{R}^n$$
$$\frac{\partial f}{\partial \mathbf{y}} = \begin{bmatrix} \frac{\partial f}{\partial y_1} & \dots & \frac{\partial f}{\partial y_n} \end{bmatrix} = \begin{bmatrix} x_1 & \dots & x_n \end{bmatrix} = \mathbf{x}^\mathsf{T} \in \mathbb{R}^n$$

The gradient is the vector of partials

$$\nabla f = \begin{bmatrix} \frac{\partial f}{\partial \mathbf{x}} & \frac{\partial f}{\partial \mathbf{y}} \end{bmatrix} = \begin{bmatrix} \mathbf{y}^\mathsf{T} & \mathbf{x}^\mathsf{T} \end{bmatrix} = \boxed{\begin{bmatrix} y_1 & \dots & y_n & x_1 & \dots & x_n \end{bmatrix}}$$

the gradient is a row vector. The jacobian is a column vector of gradients. Since there is only one gradient, the jacobian is just the gradient

$$\mathbf{J} = \nabla f = \boxed{\begin{bmatrix} y_1 & \dots & y_n & x_1 & \dots & x_n \end{bmatrix}}$$

Exercise 7.3. Find the jacobian of $f(\mathbf{x}) = \mathbf{x}\mathbf{x}^\mathsf{T}$, $\mathbf{x} \in \mathbb{R}^n$.

Solution. This function maps $\mathbb{R}^n$ to $\mathbb{R}^{n\times n}$

$$f(\mathbf{x}) = \mathbf{x}\mathbf{x}^\mathsf{T} = \underbrace{\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}}_{\mathbf{x}} \underbrace{\begin{bmatrix} x_1 & \dots & x_n \end{bmatrix}}_{\mathbf{x}^\mathsf{T}} = \begin{bmatrix} x_1\mathbf{x}^\mathsf{T} \\ \vdots \\ x_n\mathbf{x}^\mathsf{T} \end{bmatrix} = \begin{bmatrix} \mathbf{x}x_1 & \dots & \mathbf{x}x_n \end{bmatrix} = \begin{bmatrix} x_1^2 & \cdots & x_1x_n \\ \vdots & \ddots & \vdots \\ x_nx_1 & \cdots & x_n^2 \end{bmatrix}$$

The partial derivative with respect to x_1 is

$$\begin{aligned}
\frac{\partial}{\partial x_1}\left(\mathbf{x}\mathbf{x}^\mathsf{T}\right) &= \frac{\partial \mathbf{x}}{\partial x_1}\mathbf{x}^\mathsf{T} + \mathbf{x}\frac{\partial \mathbf{x}^\mathsf{T}}{\partial x_1} = \frac{\partial}{\partial x_1}\left(\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}\right)\mathbf{x}^\mathsf{T} + \mathbf{x}\frac{\partial}{\partial x_1}\left(\begin{bmatrix} x_1 & \dots & x_n \end{bmatrix}\right) \\
&= \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}\mathbf{x}^\mathsf{T} + \mathbf{x}\begin{bmatrix} 1 & 0 & \dots & 0 \end{bmatrix} = \begin{bmatrix} \mathbf{x}^\mathsf{T} \\ 0 \\ \vdots \\ 0 \end{bmatrix} + \begin{bmatrix} \mathbf{x} & 0 & \dots & 0 \end{bmatrix} \\
&= \begin{bmatrix} x_1 & x_2 & \cdots & x_n \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix} + \begin{bmatrix} x_1 & 0 & \dots & 0 \\ x_2 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ x_n & 0 & \dots & 0 \end{bmatrix} = \begin{bmatrix} 2x_1 & x_2 & \dots & x_n \\ x_2 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ x_n & 0 & \dots & 0 \end{bmatrix}
\end{aligned}$$

The other partials follow the same pattern

$$\frac{\partial}{\partial x_i}\left(\mathbf{x}\mathbf{x}^\mathsf{T}\right) = \begin{bmatrix} \mathbf{0}_{(i-1)\times n} \\ \mathbf{x}^\mathsf{T} \\ \mathbf{0}_{(n-i+1)\times n} \end{bmatrix} + \begin{bmatrix} \mathbf{0}_{n\times(i-1)} & \mathbf{x} & \mathbf{0}_{n\times(n-i+1)} \end{bmatrix}$$

and the jacobian matrix is $\mathbf{J} = \begin{bmatrix} \frac{\partial f}{\partial x_1} & \frac{\partial f}{\partial x_2} & \cdots & \frac{\partial f}{\partial x_n} \end{bmatrix}$

$$\boxed{\mathbf{J} = \left[\begin{array}{cccc|cccc|c|cccc} 2x_1 & x_2 & \dots & x_n & 0 & x_1 & \dots & 0 & \dots & 0 & 0 & \dots & x_1 \\ x_2 & 0 & \dots & 0 & x_1 & 2x_2 & \dots & x_n & \dots & 0 & 0 & \dots & x_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \dots & \vdots & \vdots & \ddots & \vdots \\ x_n & 0 & \dots & 0 & x_n & 0 & \dots & 0 & \dots & x_1 & x_2 & \dots & 2x_1 \end{array}\right]}$$

Exercise 7.4. Find the jacobian matrix for the Lorenz system of differential equations

$$\begin{aligned}
\dot{x} &= f(x,y,z) = \sigma(y-x) \\
\dot{y} &= g(x,y,z) = x(\rho-z) \\
\dot{z} &= h(x,y,z) = xy - \beta z
\end{aligned}$$

and evaluate the jacobian at stationary points (x,y,z) where $\dot{x} = \dot{y} = \dot{z} = 0$.

Solution. We need to find the three gradients first

$$\begin{aligned}
\nabla f &= \begin{bmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} & \frac{\partial f}{\partial z} \end{bmatrix} = \begin{bmatrix} \frac{\partial}{\partial x}(\sigma y - \sigma x) & \frac{\partial}{\partial y}(\sigma y - \sigma x) & \frac{\partial f}{\partial z}(\sigma y - \sigma x) \end{bmatrix} = \begin{bmatrix} -\sigma & \sigma & 0 \end{bmatrix} \\
\nabla g &= \begin{bmatrix} \frac{\partial g}{\partial x} & \frac{\partial g}{\partial y} & \frac{\partial g}{\partial z} \end{bmatrix} = \begin{bmatrix} \frac{\partial}{\partial x}(\rho x - xz) & \frac{\partial}{\partial y}(\rho x - xz) & \frac{\partial f}{\partial z}(\rho x - xz) \end{bmatrix} = \begin{bmatrix} \rho - z & 0 & -x \end{bmatrix} \\
\nabla h &= \begin{bmatrix} \frac{\partial h}{\partial x} & \frac{\partial h}{\partial y} & \frac{\partial h}{\partial z} \end{bmatrix} = \begin{bmatrix} \frac{\partial}{\partial x}(xy - \beta z) & \frac{\partial}{\partial y}(xy - \beta z) & \frac{\partial f}{\partial z}(xy - \beta z) \end{bmatrix} = \begin{bmatrix} y & x & -\beta \end{bmatrix}
\end{aligned}$$

The jacobian matrix is

$$\boxed{\mathbf{J} = \begin{bmatrix} -\sigma & \sigma & 0 \\ \rho - z & 0 & -x \\ y & x & -\beta \end{bmatrix}}.$$

At a stationary point

$$\begin{aligned} 0 &= \sigma(y - x) \\ 0 &= x(\rho - z) \\ 0 &= xy - \beta z \end{aligned}$$

There is a trivial solution $(x, y, z) = (0, 0, 0)$ where the jacobian is

$$\mathbf{J} = \begin{bmatrix} -\sigma & \sigma & 0 \\ \rho & 0 & 0 \\ y & 0 & -\beta \end{bmatrix},$$

a non/trivial solution $(x, y, z) = (\sqrt{\beta\rho}, \sqrt{\beta\rho}, \rho)$ where the jacobian is

$$\mathbf{J} = \begin{bmatrix} -\sigma & \sigma & 0 \\ 0 & 0 & -\sqrt{\beta\rho} \\ \sqrt{\beta\rho} & \sqrt{\beta\rho} & -\beta \end{bmatrix},$$

and a non/trivial solution $(x, y, z) = (-\sqrt{\beta\rho}, -\sqrt{\beta\rho}, \rho)$ where the jacobian is

$$\mathbf{J} = \begin{bmatrix} -\sigma & \sigma & 0 \\ 0 & 0 & \sqrt{\beta\rho} \\ -\sqrt{\beta\rho} & -\sqrt{\beta\rho} & -\beta \end{bmatrix}.$$

Exercise 7.5. Find the jacobian matrix for

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{b}$$

where $\mathbf{x}, \mathbf{y}, \mathbf{b} \in \mathbb{R}^n$ and $\mathbf{A} \in \mathbb{R}^{n \times n}$.

Solution. Let's first expand this too see what's happening here

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^n a_{1i}x_i + b_1 \\ \sum_{i=1}^n a_{2i}x_i + b_2 \\ \vdots \\ \sum_{i=1}^n a_{ni}x_i + b_n \end{bmatrix}$$

The derivative of y_i $(i = 1, \ldots, n)$ with respect to x_j $(j = 1, \ldots, n)$ is

$$\frac{\partial y_i}{\partial x_j} = \frac{\partial}{\partial x_j} \left(\sum_{k=1}^n a_{ik}x_k + b_i \right) = a_{ij},$$

i.e., it's all zero except when $k = j$. This means that

$$\boxed{\mathbf{J} = \mathbf{A}}$$

Exercise 7.6. Compute the derivatives of f respect to $\mathbf{x}$ for

$$f(z) = e^{-z}$$

where $z = \mathbf{x}^\mathsf{T}\mathbf{x}$, $\mathbf{x} \in \mathbb{R}^n$.

Solution. We need to use the chain rule

$$\frac{\partial f}{\partial \mathbf{x}} = \frac{\partial f}{\partial z} \cdot \frac{\partial z}{\partial \mathbf{x}}$$

$\frac{\partial f}{\partial z} = \frac{\partial}{\partial z}\left(e^{-z}\right) = -e^{-z}$ this was easy. Now for $\frac{\partial z}{\partial \mathbf{x}}$, this is a gradient

$$\begin{aligned}\nabla z = \frac{\partial z}{\partial \mathbf{x}} &= \begin{bmatrix} \frac{\partial z}{\partial x_1} & \frac{\partial z}{\partial x_2} & \cdots & \frac{\partial z}{\partial x_n} \end{bmatrix} = \begin{bmatrix} \frac{\partial \mathbf{x}^\mathsf{T}\mathbf{x}}{\partial x_1} & \frac{\partial \mathbf{x}^\mathsf{T}\mathbf{x}}{\partial x_2} & \cdots & \frac{\partial \mathbf{x}^\mathsf{T}\mathbf{x}}{\partial x_n} \end{bmatrix} \\ &= \begin{bmatrix} \frac{\partial \sum x_i^2}{\partial x_1} & \frac{\partial \sum x_i^2}{\partial x_2} & \cdots & \frac{\partial \sum x_i^2}{\partial x_n} \end{bmatrix} = \begin{bmatrix} 2x_1 & 2x_2 & \cdots & 2x_n \end{bmatrix} = 2\mathbf{x}^\mathsf{T}\end{aligned}$$

So,

$$\begin{aligned}\frac{\partial f}{\partial \mathbf{x}} &= -\,e^{-z}\nabla z \\ &= -\,e^{-\mathbf{x}^\mathsf{T}\mathbf{x}}\,2\mathbf{x}^\mathsf{T} \\ &= -\,2\begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix} e^{-\sum x_i^2} \\ &= \boxed{\begin{bmatrix} -2x_1e^{-x_1^2} & -2x_2e^{-x_2^2} & \cdots & -2x_ne^{-x_n^2} \end{bmatrix}}\end{aligned}$$

Exercise 7.7. Compute the derivatives of f respect to $\mathbf{x}$ for

$$f(\mathbf{y}) = \ln \mathbf{y}$$

where $\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{b}$, $\mathbf{x} \in \mathbb{R}^3$, $\mathbf{A} = \begin{bmatrix} 1 & -1 & 1 \\ -1 & 1 & -1 \end{bmatrix}$ and $\mathbf{b} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

Solution. We need to use the chain rule

$$\frac{\partial f}{\partial \mathbf{x}} = \frac{\partial f}{\partial \mathbf{y}} \cdot \frac{\partial \mathbf{y}}{\partial \mathbf{x}}$$

For $\frac{\partial f}{\partial \mathbf{y}}$:

$$\ln \mathbf{y} = \begin{bmatrix} \ln y_1 \\ \ln y_2 \end{bmatrix} \Rightarrow \frac{\partial f}{\partial \mathbf{y}} = \begin{bmatrix} \frac{\partial f}{\partial y_1} & \frac{\partial f}{\partial y_2} \end{bmatrix} \Rightarrow \frac{\partial f}{\partial y_1} = \begin{bmatrix} \frac{1}{y_1} \\ 0 \end{bmatrix}, \frac{\partial f}{\partial y_2} = \begin{bmatrix} 0 \\ \frac{1}{y_2} \end{bmatrix} \Rightarrow \frac{\partial f}{\partial \mathbf{y}} = \begin{bmatrix} \frac{1}{y_1} & 0 \\ 0 & \frac{1}{y_2} \end{bmatrix}$$

For $\frac{\partial \mathbf{y}}{\partial \mathbf{x}}$:

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} x_1 - x_2 + x_3 + 1 \\ -x_1 + x_2 - x_3 + 1 \end{bmatrix} \Rightarrow \frac{\partial \mathbf{y}}{\partial \mathbf{x}} = \begin{bmatrix} \frac{\partial y_1}{\partial \mathbf{x}} \\ \frac{\partial y_1}{\partial \mathbf{x}} \end{bmatrix} \Rightarrow \frac{\partial \mathbf{y}}{\partial \mathbf{x}} = \begin{bmatrix} \frac{\partial y_1}{\partial x_1} & \frac{\partial y_1}{\partial x_2} & \frac{\partial y_1}{\partial x_3} \\ \frac{\partial y_2}{\partial x_1} & \frac{\partial y_2}{\partial x_2} & \frac{\partial y_2}{\partial x_3} \end{bmatrix} = \begin{bmatrix} 1 & -1 & 1 \\ -1 & 1 & -1 \end{bmatrix}$$

So,

$$\begin{aligned}\frac{\partial f}{\partial \mathbf{x}} &= \begin{bmatrix} \frac{1}{y_1} & 0 \\ 0 & \frac{1}{y_2} \end{bmatrix}\begin{bmatrix} 1 & -1 & 1 \\ -1 & 1 & -1 \end{bmatrix} = \begin{bmatrix} \frac{1}{y_1} & -\frac{1}{y_1} & \frac{1}{y_1} \\ -\frac{1}{y_2} & \frac{1}{y_2} & -\frac{1}{y_2} \end{bmatrix} \\ &= \boxed{\begin{bmatrix} \frac{1}{1+x_1-x_2+x_3} & \frac{-1}{1+x_1-x_2+x_3} & \frac{1}{1+x_1-x_2+x_3} \\ \frac{-1}{1-x_1+x_2-x_3} & \frac{1}{1-x_1+x_2-x_3} & \frac{-1}{1-x_1+x_2-x_3} \end{bmatrix}}\end{aligned}$$

Exercise 7.8. Compute the derivatives of f respect to $\mathbf{x}$

(a) $f(z) = \log(1+z)$, $z = \mathbf{x}^\mathsf{T}\mathbf{x}$

(b) $f(\mathbf{z}) = \sin(\mathbf{z})$, $\mathbf{z} = \mathbf{A}\mathbf{x} + \mathbf{b}$

$$\mathbf{A} \in \mathbb{R}^{E\times D},\ \mathbf{x} \in \mathbb{R}^D,\ \mathbf{b} \in \mathbb{R}^E$$

Solution. (a) See that $f(z(\mathbf{x}))$ is a scalar-valued function of a vector, so the derivative takes the gradient vector form

$$\frac{\partial f}{\partial \mathbf{x}} = \frac{2\mathbf{x}^\mathsf{T}}{1+\mathbf{x}^\mathsf{T}\mathbf{x}} = \begin{bmatrix} \frac{2x_1}{1+\sum_{i=1}^D x_i^2} & \frac{2x_2}{1+\sum_{i=1}^D x_i^2} & \cdots & \frac{2x_D}{1+\sum_{i=1}^D x_i^2} \end{bmatrix}$$

(b) See that $f(\mathbf{z}(\mathbf{x}))$ is a vector-valued function of a vector, so the derivative takes the jacobian matrix form. Using the Chen-Lu

$$\frac{\partial f}{\partial \mathbf{x}} = \frac{\partial f}{\partial \mathbf{z}}\frac{\partial \mathbf{z}}{\partial \mathbf{x}} = \cos(\mathbf{z})\mathbf{A},$$

but $\cos(\mathbf{z})$ has E rows and $\mathbf{A}$ has D columns, we cannot multiply $\cos(\mathbf{z})\mathbf{A}$ as shown unless $E = D$. But if we lay out $\cos(\mathbf{z})$ as a diagonal matrix $E \times E$ we can multiply it by $\mathbf{A}$ which has E rows

$$\begin{aligned}\frac{\partial f}{\partial \mathbf{x}} &= \operatorname{diag}\left(\cos\left(\mathbf{A}\mathbf{x}+\mathbf{b}\right)\right)\mathbf{A}\\ &= \begin{bmatrix} \cos\left(b_1 + \sum_{i=1}^D a_{1i}x_i\right) & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \cos\left(b_E + \sum_{i=1}^D a_{Ei}x_i\right) \end{bmatrix} \begin{bmatrix} a_{11} & \cdots & a_{1D} \\ \vdots & \ddots & \vdots \\ a_{E1} & \cdots & a_{ED} \end{bmatrix}\end{aligned}$$

If you are not convinced, let's do it the long way

$$f(\mathbf{z}) = \underbrace{f(\begin{bmatrix} \sin(z_1) \\ \vdots \\ \sin(z_E) \end{bmatrix})}_{E\times 1} = \underbrace{f(\underbrace{\begin{bmatrix} x_1 \\ \vdots \\ x_D \end{bmatrix}}_{D\times 1})}_{E\times 1} = \underbrace{\begin{bmatrix} \sin\left(b_1 + \sum_{j=1}^D a_{1j}x_j\right) \\ \vdots \\ \sin\left(b_D + \sum_{j=1}^D a_{Ej}x_j\right) \end{bmatrix}}_{\text{Hey! that's still } E\times 1}$$

the derivative is

$$\frac{\partial f}{\partial \mathbf{x}} = \begin{bmatrix} \frac{\partial}{\partial \mathbf{x}}\left(\sin\left(b_1 + \sum_{j=1}^D a_{1j}x_j\right)\right) \\ \vdots \\ \frac{\partial}{\partial \mathbf{x}}\left(\sin\left(b_D + \sum_{j=1}^D a_{Ej}x_j\right)\right) \end{bmatrix} = \begin{bmatrix} \cos\left(b_1 + \sum_{j=1}^D a_{1j}x_j\right)\frac{\partial}{\partial \mathbf{x}}\left(b_1 + \sum_{j=1}^D a_{1j}x_j\right) \\ \vdots \\ \cos\left(b_D + \sum_{j=1}^D a_{Ej}x_j\right)\frac{\partial}{\partial \mathbf{x}}\left(b_D + \sum_{j=1}^D a_{Ej}x_j\right) \end{bmatrix}$$

now see that $\frac{\partial}{\partial \mathbf{x}}\left(b_i + \sum_{j=1}^D a_{ij}x_j\right)$ is a gradient vector. Continue but using $z_i = b_i + \sum_{j=1}^D a_{ij}x_j$ down below

$$\frac{\partial f}{\partial \mathbf{x}} = \begin{bmatrix} \cos(z_1)\nabla_\mathbf{x}(z_1) \\ \vdots \\ \cos(z_E)\nabla_\mathbf{x}(z_E) \end{bmatrix} = \begin{bmatrix} \cos(z_1)\begin{bmatrix} a_{11} & \cdots & a_{1D} \end{bmatrix} \\ \vdots \\ \cos(z_E)\begin{bmatrix} a_{E1} & \cdots & a_{ED} \end{bmatrix} \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} a_{11}\cos(z_1) & \cdots & a_{1D}\cos(z_1) \end{bmatrix} \\ \vdots \\ \begin{bmatrix} a_{E1}\cos(z_E) & \cdots & a_{ED}\cos(z_E) \end{bmatrix} \end{bmatrix}$$

This is a column vector or row vectors, so $\frac{\partial f}{\partial \mathbf{x}}$ is a jacobian matrix. Row i of this jacobian is row i of $\mathbf{A}$ scaled by $\cos(z_i)$. This is why we can factor the jacobian as the product of a diagonal $\cos(\mathbf{z})$ matrix times $\mathbf{A}$.

8 Continuous optimization

Exercise 8.1. Find and classify all the critical points of the following function

$$f(x, y) = 3y^3 - x^2y^2 + 8y^2 + 4x^2 - 20y.$$

Solution. Get all first and second partial derivatives

$$f_x = -2xy^2 + 8x \qquad f_y = 9y^2 - 2x^2y + 16y - 20$$
$$f_{xx} = -2y^2 + 8 \qquad f_{xy} = -4xy \qquad f_{yy} = 18y - 2x^2 + 16$$

Find all critical points by setting the gradient to zero $(f_x, f_y) = (0, 0)$. Equation $f_x = 0$ has three solutions $x = 0$ and $y = \pm 2$. Use them to find the solutions of the $f_y = 0$ equation:

- if $x = 0$: the solution of $9y^2 - 2 \cdot 0^2 y + 16y - 20 = 0$ is $y = \frac{-16 \pm \sqrt{976}}{18}$
- if $y = 2$: the solution of $9 \cdot 2^2 - 2x^2 \cdot 2 + 16 \cdot 2 - 20 = 0$ is $x = \pm 2\sqrt{3}$
- if $y = -2$: the solution of $9 \cdot (-2)^2 - 2x^2 \cdot (-2) + 16 \cdot (-2) - 20 = 0$ is $x = \pm 2$

There are six critical points (x, y)

1. $\left(0, \frac{-16+\sqrt{976}}{18}\right)$
2. $\left(0, \frac{-16-\sqrt{976}}{18}\right)$
3. $\left(2\sqrt{3}, 2\right)$
4. $\left(-2\sqrt{3}, 2\right)$
5. $(2, -2)$
6. $(-2, -2)$

For classification we need

$$\begin{aligned} D(x, y) =& f_{xx} f_{yy} - f_{xy}^2 = \left(-2y^2 + 8\right)\left(18y - 2x^2 + 16\right) - (-4xy)^2 \\ =& 4\left[\left(4 - y^2\right)\left(8 + 9y - x^2\right) - 4x^2y^2\right] \end{aligned}$$

Use a calculator

1. $D\left(0, \frac{-16+\sqrt{976}}{18}\right) = 205.1 > 0$ and $f_{xx}\left(0, \frac{-16+\sqrt{976}}{18}\right) = 6.6 > 0$. This is a (relative) **minimum**
2. $D\left(0, \frac{-16-\sqrt{976}}{18}\right) = 180.4 > 0$ and $f_{xx}\left(0, \frac{-16-\sqrt{976}}{18}\right) = -5.8 < 0$. This is a (relative) **maximum**
3. $D\left(2\sqrt{3}, 2\right) = -768 < 0$. This is a **saddle point**
4. $D\left(-2\sqrt{3}, 2\right) = -768 < 0$. This is a **saddle point**
5. $D(2, -2) = -256 < 0$. This is a **saddle point**
6. $D(-2, -2) = -256 < 0$. This is a **saddle point**

Exercise 8.2. Find and classify all the critical points of the following function

$$f\left(x,y\right) = 8x - x\sqrt{y-1} + x^3 + \frac{1}{2}y - 12x^2.$$

Solution. Get all first and second partial derivatives

$$f_x = 8 - \sqrt{y-1} + 3x^2 - 24x \qquad\qquad f_y = -\frac{x}{2\sqrt{y-1}} + \frac{1}{2}$$

$$f_{xx} = 6x - 24 \qquad\qquad f_{xy} = 0 \qquad\qquad f_{yy} = \frac{x}{4}(y-1)^{-3/2}$$

Find all critical points by setting the gradient to zero $(f_x, f_y) = (0,0)$. From $f_x = 0$ we get that $\sqrt{y-1} = 3x^2 - 24x + 8$. Substituting this in $f_y = 0$ gives $3x^2 - 24x + 8 = 0$ and this quadratic equation has two solutions $x = \frac{1}{3}$ and $x = 8$. Replacing $x = \frac{1}{3}$ in $f_x = 0$ gives

$$\sqrt{y-1} = 3\cdot\frac{1}{3^2} - 24\cdot\frac{1}{3} + 8 = 8 - \frac{23}{3} + 8 = \frac{1}{3} \Rightarrow y - 1 = \frac{1}{9} \Rightarrow y = \frac{10}{9}$$

and replacing $x = 8$ in $f_x = 0$ gives

$$\sqrt{y-1} = \cancel{3\cdot 8^2} - \cancel{24\cdot 8} + 8 = 8 \Rightarrow y - 1 = 64 \Rightarrow y = \sqrt{65}$$

There are two critical points (x,y)

1. $\left(\frac{1}{3}, \frac{10}{9}\right)$
2. $\left(8, \sqrt{65}\right)$

For classification we need

$$D(x,y) = f_{xx}f_{yy} - f_{xy}^2 = (6x-24)\left(\frac{x}{4}(y-1)^{-3/2}\right) - (0)^2 = \frac{3x(x-4)}{2(y-1)^{3/2}}$$

A calculator is not required because it is very easy to determine the sign of D at both critical points

1. $D\left(\frac{1}{3}, \frac{10}{9}\right) = \frac{3}{2}\frac{1}{3}\left(\frac{1}{3} - 4\right)\left(\frac{10}{9} - 1\right)^{-3/2} < 0$. This is a **saddle point**
2. $D\left(8, \sqrt{65}\right) = \frac{3}{2}8\,(8-4)\left(\sqrt{65} - 1\right)^{-3/2} > 0$ and $f_{xx}\left(8, \sqrt{65}\right) = 6\cdot 8 - 24 = 24 > 0$. This is a **minimum**

Exercise 8.3. Find the equation of the tangent plane to

$$z = x^2\cos\left(\pi y\right) - \frac{6}{xy^2}$$

at $(2, -1)$.

Solution. Get the partial derivatives

$$f_x = 2x\cos(\pi y) + \frac{6}{x^2y^2} \qquad\qquad f_y = -\pi x^2\sin(xy) + \frac{12}{xy^3}$$

Evaluate at $(2, -1)$

$$f(2,-1) = -7 \qquad\qquad f_x(2,-1) = -\frac{5}{2} \qquad\qquad f_y(2,-1) = -6$$

The tangent plane is

$$\boxed{z = -7 - \frac{5}{2}(x-2) - 6(y+1) = -\frac{5}{2}x - 6y - 8}$$

Exercise 8.4. Find the equation of the tangent plane to

$$z = x^2y^4 - \frac{12x}{y}$$

at $(-1, 6)$.

Solution. Get the partial derivatives

$$f_x = 2xy^4 - \frac{12}{y} \qquad\qquad f_y = 4x^2y^3 + \frac{12x}{y^2}$$

Evaluate at $(-1, 6)$

$$\begin{aligned} f(-1,6) &= (-1)^2 6^4 - \frac{12\cdot(-1)}{6} \\ &= 6^4 + 2 = 1298 \end{aligned} \qquad \begin{aligned} f_x(-1,6) &= 2(-1)6^4 - \frac{12}{6} \\ &= 2\cdot 6^4 - 2 = 2590 \end{aligned} \qquad \begin{aligned} f_y(-1,6) &= 4(-1)^2 6^3 + \frac{12\cdot(-2)}{6^3} \\ &= 4\cdot 6^3 - \frac{24}{6^3} = \frac{2590}{3} \end{aligned}$$

The tangent plane is

$$\begin{aligned} z &= f(-1,6) + f_x(-1,6)(x+1) + f(-1,6)(y-6) \\ &= 1298 + 2590(x+1) + \frac{2590(y-6)}{3} \\ &= 1298 + 2590x + 2590 + \frac{2590}{3}y - 2\cdot 2590 \\ &= \boxed{-1292 + 2590x + \frac{2590}{3}y} \end{aligned}$$

Exercise 8.5. Find the maximum and minimum values of $f(x, y) = 81x^2+y^2$ subject to the constraint $4x^2 + y^2 = 9$.

Solution. Before we start notice from the constraint that $-\frac{3}{2} \leq x \leq \frac{3}{2}$ and $-3 \leq y \leq 3$. We need to solve the following system of equations

$$\begin{aligned} \nabla f(x,y) &= \lambda \nabla g(x,y) \\ g(x,y) &= c \end{aligned}$$

where $g(x, y) = 4x^2 + y^2$ and $c = 9$. The first means that the gradients of f and g are parallel at critical points, and the scalar λ (Lagrange multiplier) makes them equal. The gradients are

$$\begin{aligned} \nabla f(x,y) &= (162x, 2y) \\ \nabla g(x,y) &= (8x, 2y) \end{aligned}$$

We need to satisfy

$$\begin{aligned} 162x &= \lambda 8x \\ 2y &= \lambda 2y \\ 4x^2 + y^2 &= 9 \end{aligned}$$

In the 2nd equation the solutions are $y = 0$ and $\lambda = 1$:

- substituting $y = 0$ on the 3rd we get $x = \pm\frac{3}{2}$
- substituting $\lambda = 1$ on the 1st we get $x = 0$, which replaced on the 3rd gives $y = \pm 3$

We have three solutions

1. $(x, y) = \left(\frac{3}{2}, 0\right)$ and $f\left(\frac{3}{2}, 0\right) = \frac{729}{4}$ that's between 175 and 200
2. $(x, y) = \left(-\frac{3}{2}, 0\right)$ and $f\left(-\frac{3}{2}, 0\right) = \frac{729}{4}$ that's between 175 and 200
3. $(x, y) = (0, 3)$ and $f(0, 3) = 9$
4. $(x, y) = (0, -3)$ and $f(0, -3) = 9$

We can conclude that points 1 and 2 are absolute maxima while 3 and 4 are absolute minima.

Exercise 8.6. Find the maximum and minimum values of $f(x, y, z) = 3x^2+y$ subject to the constraints $4x - 3y = 9$ and $x^2 + z^2 = 9$.

Solution. First note that because of the 2nd constraint $-3 \leq x \leq 3$ and $-3 \leq z \leq 3$. Then from the 1st constraint it follows that $-7 \leq z \leq 7$. In this problem we need satisfaction of the following equations

$$\begin{aligned}
\nabla f(x, y, z) &= \lambda \nabla g(x, y, z) + \mu \nabla h(x, y, z) \\
g(x, y, z) &= 9 \\
h(x, y, z) &= 9
\end{aligned}$$

with $g(x, y, z) = 4x - 3y$ and $h(x, y, z) = x^2 + z^2$. Notice that

$$\nabla f = \underbrace{\begin{bmatrix} \lambda & \mu \end{bmatrix}}_{\text{vector of multipliers}} \underbrace{\begin{bmatrix} \nabla g \\ \nabla h \end{bmatrix}}_{\text{jacobian matrix}}$$

The gradients

$$\begin{aligned}
\nabla f &= \left[\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}\right] = [6x, 1, 0] \\
\nabla g &= \left[\frac{\partial g}{\partial x}, \frac{\partial g}{\partial y}, \frac{\partial g}{\partial z}\right] = [4, -3, 0] \\
\nabla h &= \left[\frac{\partial h}{\partial x}, \frac{\partial h}{\partial y}, \frac{\partial h}{\partial z}\right] = [2x, 0, 2z]
\end{aligned}$$

That's a total of five equations

$$\begin{aligned}
6x &= 4\lambda + 2\mu x \\
1 &= -3\lambda \\
0 &= 2\mu z \\
4x - 3y &= 9 \\
x^2 + z^2 &= 9
\end{aligned}$$

From the 2nd equation $\boxed{\lambda = -\frac{1}{3}}$. From the 3rd $\mu = 0$ or $z = 0$. We need to see what happens with each of these two

- For $z = 0$ the 5th equation (the 2nd constraint) gives $x = \pm 3$. Now use these two on the 4th equation (1st constraint)

$$\begin{array}{llll}
x = -3: & -12 - 3y = 9 & \rightarrow & y = -7 \\
x = 3: & 12 - 3y = 9 & \rightarrow & y = 1
\end{array}$$

  So, we have points $P_1 = (-3, -7, 0)$ and $P_2 = (3, 1, 0)$

- For $\mu = 0$ the 1st equation (remember that $\lambda = -\frac{1}{3}$!!!) produces $x = -\frac{2}{9}$. Using this value in the 4th and 5 equations (1st and 2nd constraints) makes

$$4\left(-\frac{2}{9}\right) - 3y = 9 \qquad \rightarrow \qquad y = -\frac{89}{27}$$
$$\left(-\frac{2}{9}\right)^2 + z^2 = 9 \qquad \rightarrow \qquad z = \pm\frac{5\sqrt{29}}{9}$$

So, we have points $P_3 = \left(-\frac{2}{9}, -\frac{89}{27}, -\frac{5\sqrt{29}}{9}\right)$ and $P_4 = \left(-\frac{2}{9}, -\frac{89}{27}, \frac{5\sqrt{29}}{9}\right)$

Evaluation at the points

1. $f(-3, -7, 0) = 20$
2. $f(3, 1, 0) = 28$
3. $f\left(-\frac{2}{9}, -\frac{89}{27}, -\frac{5\sqrt{29}}{9}\right) = -\frac{85}{27}$
4. $f\left(-\frac{2}{9}, -\frac{89}{27}, \frac{5\sqrt{29}}{9}\right) = -\frac{85}{27}$

The absolute maximum is 28 at $(x, y, z) = (3, 1, 0)$ and the absolute minimum is $-\frac{85}{27}$ which occurs at $\left(-\frac{2}{9}, -\frac{89}{27}, \pm\frac{5\sqrt{29}}{9}\right)$.

Exercise 8.7. Find the maximum and minimum values of $f(x, y) = 3x - 6y$ subject to the constraint $4x^2 + 2y^2 = 25$.

Solution. The constraint indicates that $|x| \leq 5/2$ and $|y| \leq 5\sqrt{2}/2$. The following equations must be satisfied

$$\nabla f(x, y) = \lambda \nabla g(x, y)$$
$$g(x, y) = 0$$

with $g(x, y) = 4x^2 + 2y^2 - 25$. This gives

$$3 = 8\lambda x$$
$$-3 = 2\lambda y$$
$$4x^2 + 2y^2 = 25$$

The 1st and 2nd equations produce $x = \frac{3}{8\lambda}$ and $y = \frac{3}{2\lambda}$, respectively. Substite these into the 3rd to get λ

$$4\left(\frac{3}{8\lambda}\right)^2 + 2\left(\frac{3}{2\lambda}\right)^2 = 25$$
$$\frac{36}{64\lambda^2} + \frac{9}{2\lambda^2} = 25$$
$$25\lambda^2 = \frac{81}{16}$$
$$\lambda = \pm\frac{9}{20}.$$

Using $\lambda = \pm\frac{9}{20}$ gives $x = \pm\frac{5}{12}$ and $y = \pm\frac{10}{3}$, all valid (within the constraints). Thus,

1. $(x,y)=\left(\frac{5}{12},\frac{10}{3}\right)$ and $f\left(\frac{5}{12},\frac{10}{3}\right)=3\frac{5}{12}-6\frac{10}{3}=-\frac{75}{4}\approx-18.75$
2. $(x,y)=\left(-\frac{5}{12},-\frac{10}{3}\right)$ and $f\left(-\frac{5}{12},-\frac{10}{3}\right)=-3\frac{5}{12}+6\frac{10}{3}=\frac{75}{4}\approx 18.75$

We can conclude that point 1 is an absolute minimum and point 2 is an absolute maximum.

Exercise 8.8. Determine the quadratic polynomial approximation of

$$f(x,y)=\sin 2x+\cos y$$

near the point $(0,0)$.

Solution. Get all first and second partial derivatives

$$f_x=2\cos 2x \qquad\qquad f_y=-\sin y$$
$$f_{xx}=-4\sin 2x \qquad f_{xy}=0 \qquad f_{yy}=-\cos y$$

At $(x_o,y_o)=(0,0)$ we get $f(0,0)=\sin 2\cdot 0+\cos 0=1$ and

$$f_x(0,0)=2 \qquad\qquad f_y(0,0)=0$$
$$f_{xx}(0,0)=0 \qquad f_{xy}(0,0)=0 \qquad f_{yy}(0,0)=-1$$

The linear approximation is

$$L(x,y)=f(x,y)+f_x(x,y)(x-x_o)+f_y(x,y)(y-y_o)$$

and at $(x_o,y_o)=(0,0)$ that's

$$L(0,0)=f(0,0)+f_x(0,0)x+f_y(0,0)y=1+2x$$

The quadratic approximation is

$$Q(x,y)=L(x,y)+\frac{f_x(x,y)}{2}(x-x_o)^2+f_{xy}(x,y)(x-x_o)(y-y_o)+\frac{f_y(x,y)}{2}(y-y_o)^2$$

and at $(x_o,y_o)=(0,0)$ that's

$$Q(x,y)=L(0,0)+\frac{f_x(0,0)}{2}x^2+f_{xy}(0,0)xy+\frac{f_y(0,0)}{2}y^2=\boxed{1+2x-\frac{1}{2}y^2}$$

Exercise 8.9. Determine the quadratic polynomial approximation of

$$f(x,y)=1+xe^y$$

near the point $(1,0)$.

Solution. Get all first and second partial derivatives

$$f_x=e^y \qquad\qquad f_y=x$$
$$f_{xx}=0 \qquad f_{xy}=0 \qquad f_{yy}=0$$

At $(x_o,y_o)=(1,0)$ we get $f(1,0)=1+1\cdot e^0=2$ and

$$f_x(1,0)=1 \qquad\qquad f_y(1,0)=1$$
$$f_{xx}(1,0)=0 \qquad f_{xy}(1,0)=0 \qquad f_{yy}(1,0)=0$$

The linear approximation is

$$L(x,y)=f(x,y)+f_x(x,y)(x-x_o)+f_y(x,y)(y-y_o)$$

and at $(x_o,y_o)=(1,0)$ that's

$$L(1,0)=f(1,0)+f_x(1,0)(x-1)+f_y(1,0)(y-0)=2+1\cdot(x-1)+1\cdot y=1+x+y$$

The quadratic approximation is

$$Q(x,y)=L(x,y)+\frac{f_{xx}(x,y)}{2}(x-x_o)^2+f_{xy}(x,y)(x-x_o)(y-y_o)+\frac{f_{yy}(x,y)}{2}(y-y_o)^2$$

but at $(x_o,y_o)=(1,0)$ all 2nd order derivates are equal to 0. Thus, $Q(1,0)=L(1,0)=\boxed{1+x+y}$

Bibliography

[1] Deisenroth, M. P., Faisal, A. A. and Ong, C. S. (2000) *Mathematics for Machine Learning*, Cambridge University Press.