Calendar

	MON	TUE	WED	THU	FRI	SAT	SUN
			21/9				
				29/9			
			5/10				

Dmitry Gromov
gromov@lu.lv

Math 4 Data Analytics Week 1

Systems of linear equations

$a_{11}x_1 + a_{12}x_2 + \dots + a_nx_n = b$

$\vdots$

$a_{k1}x_1 + a_{k2}x_2$ - - - - - - - etc

k amount of equations
n amount of equations

1st subscript is rows
2nd subscript is variables

Example of page rank

$$s_1 = s_3 + s_4$$
$$s_2 = s_1 + s_4$$
$$s_3 = s_2$$
$$s_4 = s_2 + s_3$$

$$s_3 + s_4 - s_1 = 0$$
$$s_1 + s_4 - s_2 = 0$$
$$s_2 - s_3 = 0$$
$$s_2 + s_3 - s_4 = 0$$
$$s_1 + s_2 + s_3 + s_4 = 1$$

{

If all $b_i = 0$ system is homogeneous
If (*) has at least one solution, we call it determined

Keywords

Gaussian Elimination

Elementary operations

1. $Eq_i \to \alpha Eq_i, \alpha \neq 0$
2. $Eq_i + Eq_j \to Eq_i$
3. $Eq_i \leftrightarrow Eq_j$

Reduced Row Form

Triangle Form

$$x_1 + 2x_2 + 4x_3 = 3$$
$$3x_1 + 8x_2 + 14x_3 = 13$$
$$2x_1 + 6x_2 + 13x_3 = 4$$

$$x_1 + 2x_2 + 4x_3 = 3$$
$$2x_2 + 2x_3 = 4$$
$$2x_2 + 5x_3 = -2$$

$$x_1 + 2x_2 + 4x_3 = 3$$
$$2x_2 + 2x_3 = 4$$
$$3x_3 = -6$$

$x_3 = -2$

$x_2 = 4$

$x_1 = 3$

$$x_1 + 3x_2 + x_3 = 3$$
$$x_2 - x_3 = 1$$
$$x_1 - 3x_3 = 5 \xrightarrow{-1}$$

$$3x_2 + 4x_3 = 2$$
$$x_3 = -5$$

$$x_1 - 3x_2 + 7x_3 = 1$$
$$-2x_1 + 6x_2 + x_3 = 1$$
$$3x_2 - 2x_3 = 1$$

$$x_1 - 3x_2 + 7x_3 = 1$$
$$3x_2 + x_3 = 1$$
$$-2x_1 + 6x_2 + x_3 = 1$$

$x_1 - 14/15 + 14/5 = 1$ $\quad 15x_3 = 1$ $\quad 3x_2 + 1/15 = 1$

$15x_1 - 14 + 42 = 1$ $\quad x_3 = 1/15$; $x_2 = 14/45$ $\quad 3x_2 = 14/15$

$$\frac{15x}{15} = \frac{-40}{}$$

$2x_1 - 5x_2 + x_3 = 2$ ← infinite many solutions
$-x_1 + 2x_2 - x_3 = 3$
$3x_1 - 8x_2 + x_3 = 7$

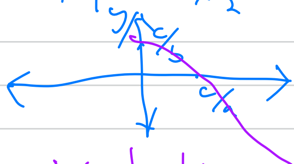
$0 = 0$

$2x_1 - 5x_2 + x_2 = 2$ ← no solution
$-x_1 + 2x_2 - x_3 = 3$
$3x_1 - 8x_2 + x_3 = 7$

$0 = 3$

$ax_1 + bx_2 = c$

$x_1 = 0$



1 solution

$a_{11}x_1 + a_{12}x_2 = b_1$
$a_{21}x_1 + a_{22}x_2 = b_2$

$\emptyset$ sol.

∞ sol.

Matrices

Rules

Orthoginal to Matrices is a array of #s

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & \vdots \\ \vdots & \vdots & & \vdots \\ a_{k1} & a_{k2} & \cdots & a_{kn} \end{bmatrix}$$

a_{ij}

row column

1. Dimension $[k \times n]$
   [rows × columns]
2. $[1 \times n]$ - row vector
   $[K \times 1]$ - column vector *
   scaler

Matrices Operations

(1). Multiplication w. a scalar

αA, $\alpha \in \mathbb{R}$

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \quad \alpha A = \begin{bmatrix} \alpha a_{11} & \alpha_{12} \\ \alpha a_{21} & \alpha_{22} \end{bmatrix}$$

(2) Addition

$A+B$

Dimensions must coincide

$$A = \begin{bmatrix} a_{11} & a_{21} \\ a_{21} & a_{22} \end{bmatrix} \quad B \begin{bmatrix} b_{11} & b_{21} \\ b_{21} & b_{22} \end{bmatrix}$$

$$A+B = \begin{bmatrix} a_{11}+b_{11} & \cdots \\ a_{21}+b_{21} & \cdots \end{bmatrix}$$

(3) Multiplication

$A \times B$ is defined only if their dimensions are consistent

$A \times B$

$[n \times m]\,[k \times l] = [n \times l]$

$m = k$

$n=1,\ \ell=1$

$A = \begin{bmatrix} 1 & 2 & 3 \end{bmatrix}$ $B = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ $AB = \begin{bmatrix} 1 & 2 & 3 \end{bmatrix}\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$

$\begin{bmatrix} 1\times1 + 2\times2 + 3\times3 \end{bmatrix}$

14

$A = [a_1 \ \cdots \ a_n]$

$B = \begin{bmatrix} b_{11} \\ \vdots \\ b_{n1} \end{bmatrix}$

$AB = a_{11} b_{11} + a_{12} b_{12} + \cdots + a_{1n} b_{n1} =$

$= \sum_{k=1}^{n} a_{1k} b_{k1}$

Problems

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

$-1 + 2 - 3 + 4$

2

$\begin{bmatrix} 1 & 2 & 0 & 4 \end{bmatrix}\begin{bmatrix} 2 \\ 0 \\ -4 \\ 5 \end{bmatrix}$

$2 \quad 20$

22

$2\times3 \quad 3\times2$

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

Vectors has atleast one dimension

$3\times3 \quad 3\times2 \quad 3\times2$

$\begin{bmatrix} 1 & 2 & -1 \\ 2 & 0 & 1 \\ 1 & 3 & 0 \end{bmatrix}\begin{bmatrix} 1 & 3 \\ 2 & 0 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} 6 & 2 \\ 1 & 7 \\ 7 & 3 \end{bmatrix}$

$1 \quad 4 \quad 1$

Connection to systems

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix}$$

$$Ax = \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 \end{bmatrix}$$

$$x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \quad b = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$$

$$a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = b_1$$
$$a_{21}x_1 + a_{22}x_2 + a_{23}x_3 = b_2$$

$$\boxed{Ax = b}$$

$$\overset{A}{\begin{bmatrix} 1 & 2 & 0 & 4 \end{bmatrix}} \underset{B}{\begin{bmatrix} 2 \\ 0 \\ -1 \\ 5 \end{bmatrix}}$$

$AB = 22$

$$BA = \begin{bmatrix} \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \end{bmatrix}$$

$A = [2 \times 3]$

$B = [2 \times 3]$

$C = [3 \times 3]$

1. $(A+B)C$ ✓
2. $(A+C)B$ ✗
3. $AC + B$ ✓
4. $CA + B$ ✗

1. Solving three equations in three unknowns

The easiest set of three simultaneous linear equations to solve is of the following type:

$$3x_1 = 6,$$

$$2x_2 = 5,$$

$$4x_3 = 7$$

which obviously has solution $[x_1, x_2, x_3]^T = \left[2, \frac{5}{2}, \frac{7}{4}\right]^T$ or $x_1 = 2, x_2 = \frac{5}{2}, x_3 = \frac{7}{4}$.
In matrix form $AX = B$ the equations are

$$\begin{bmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 6 \\ 5 \\ 7 \end{bmatrix}$$

where the matrix of coefficients, A, is clearly diagonal.

Task

Solve the equations

$$\begin{bmatrix} 2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 8 \\ 2 \\ -6 \end{bmatrix}.$$

Your solution

Answer

$[x_1, x_2, x_3]^T = [4, -2, -2]^T$.

The next easiest system of equations to solve is of the following kind:

$$\begin{aligned} 3x_1 + x_2 - x_3 &= 0 \\ 2x_2 + x_3 &= 12 \\ 3x_3 &= 6. \end{aligned}$$

The last equation can be solved immediately to give $x_3 = 2$.
Substituting this value of x_3 into the second equation gives

$$2x_2 + 2 = 12 \qquad \text{from which} \qquad 2x_2 = 10 \qquad \text{so that} \qquad x_2 = 5$$

Substituting these values of x_2 and x_3 into the first equation gives

$$3x_1 + 5 - 2 = 0 \qquad \text{from which} \qquad 3x_1 = -3 \qquad \text{so that} \qquad x_1 = -1$$

Hence the solution is $[x_1, x_2, x_3]^T = [-1, 5, 2]^T$.
This process of solution is called **back-substitution**.
In matrix form the system of equations is

$$\begin{bmatrix} 3 & 1 & -1 \\ 0 & 2 & 1 \\ 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 12 \\ 6 \end{bmatrix}.$$

The matrix of coefficients is said to be **upper triangular** because all elements below the leading diagonal are zero. Any system of equations in which the coefficient matrix is triangular (whether upper or lower) will be particularly easy to solve.

Task

Solve the following system of equations by back-substitution.

$$\begin{bmatrix} 2 & -1 & 3 \\ 0 & 3 & -1 \\ 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 7 \\ 5 \\ 2 \end{bmatrix}.$$

Write the equations in expanded form:

Your solution

Answer

$$\begin{aligned} 2x_1 - x_2 + 3x_3 &= 7 \\ 3x_2 - x_3 &= 5 \\ 2x_3 &= 2 \end{aligned}$$

Now find the solution for x_3:

Your solution

$x_3 =$

Answer

The last equation can be solved immediately to give $x_3 = 1$.

Using this value for x_3, obtain x_2 and x_1:

Your solution

$x_2 =$ $x_1 =$

Answer

$x_2 = 2, \quad x_1 = 3$. Therefore the solution is $x_1 = 3, x_2 = 2$ and $x_3 = 1$.

Although we have worked so far with integers this will not always be the case and fractions will enter the solution process. We must then take care and it is always wise to check that the equations balance using the calculated solution.

2. The general system of three simultaneous linear equations

In the previous subsection we met systems of equations which could be solved by back-substitution alone. In this Section we meet systems which are not so amenable and where preliminary work must be done before back-substitution can be used.

Consider the system

$$\begin{aligned} x_1 + 3x_2 + 5x_3 &= 14 \\ 2x_1 - x_2 - 3x_3 &= 3 \\ 4x_1 + 5x_2 - x_3 &= 7 \end{aligned}$$

We will use the solution method known as **Gauss elimination**, which has three stages. In the first stage the equations are written in matrix form. In the second stage the matrix equations are replaced by a system of equations having the same solution but which are in **triangular form**. In the final stage the new system is solved by **back-substitution**.

Stage 1: Matrix Formulation

The first step is to write the equations in matrix form:

$$\begin{bmatrix} 1 & 3 & 5 \\ 2 & -1 & -3 \\ 4 & 5 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 14 \\ 3 \\ 7 \end{bmatrix}.$$

Then, for conciseness, we combine the matrix of coefficients with the column vector of right-hand sides to produce the **augmented matrix**:

$$\left[\begin{array}{rrr|r} 1 & 3 & 5 & 14 \\ 2 & -1 & -3 & 3 \\ 4 & 5 & -1 & 7 \end{array}\right]$$

If the general system of equations is written $AX = B$ then the augmented matrix is written $[A|B]$.

Hence the first equation

$$x_1 + 3x_2 + 5x_3 = 14$$

is replaced by the first row of the augmented matrix,

$1 \quad 3 \quad 5 \quad | \quad 14$ and so on.

Stage 1 has now been completed. We will next triangularise the matrix of coefficients by means of **row operations**. There are three possible row operations:

- interchange two rows;
- multiply or divide a row by a non-zero constant factor;
- add to, or subtract from, one row a multiple of another row.

Note that interchanging two rows of the augmented matrix is equivalent to interchanging the two corresponding equations. The shorthand notation we use is introduced by example. To interchange row 1 and row 3 we write $R1 \leftrightarrow R3$. To divide row 2 by 5 we write $R2 \div 5$. To add three times row 1 to row 2, we write $R2 + 3R1$. In the Task which follows you will see where these annotations are placed.

Note that these operations neither create nor destroy solutions so that at every step the system of equations has the same solution as the original system.

Stage 2: Triangularisation

The second stage proceeds by first eliminating x_1 from the second and third equations using row operations.

$$\left[\begin{array}{rrr|r} 1 & 3 & 5 & 14 \\ 2 & -1 & -3 & 3 \\ 4 & 5 & -1 & 7 \end{array}\right] \begin{array}{l} \\ R2 - 2\times R1 \\ R3 - 4\times R1 \end{array} \quad \Rightarrow \quad \left[\begin{array}{rrr|r} 1 & 3 & 5 & 14 \\ 0 & -7 & -13 & -25 \\ 0 & -7 & -21 & -49 \end{array}\right]$$

In the above we have subtracted twice row (equation) 1 from row (equation) 2.
In full these operations would be written, respectively, as

$$(2x_1 - x_2 - 3x_3) - 2(x_1 + 3x_2 + 5x_3) = 3 - 2\times 14 \quad \text{or} \quad -7x_2 - 13x_3 = -25$$

and

$$(4x_1 + 5x_2 - x_3) - 4(x_1 + 3x_2 + 5x_3) = 7 - 4\times 14 \quad \text{or} \quad -7x_2 - 21x_3 = -49.$$

Now since all the elements in rows 2 and 3 are negative we multiply throughout by -1:

$$\left[\begin{array}{rrr|r} 1 & 3 & 5 & 14 \\ 0 & -7 & -13 & -25 \\ 0 & -7 & -21 & -49 \end{array}\right] \begin{array}{l} \\ R2\times(-1) \\ R3\times(-1) \end{array} \quad \Rightarrow \quad \left[\begin{array}{rrr|r} 1 & 3 & 5 & 14 \\ 0 & 7 & 13 & 25 \\ 0 & 7 & 21 & 49 \end{array}\right]$$

Finally, we eliminate x_2 from the third equation by subtracting equation 2 from equation 3
i.e. $R3 - R2$:

$$\left[\begin{array}{rrr|r} 1 & 3 & 5 & 14 \\ 0 & 7 & 13 & 25 \\ 0 & 7 & 21 & 49 \end{array}\right] \begin{array}{l} \\ \\ R3 - R2 \end{array} \quad \Rightarrow \quad \left[\begin{array}{rrr|r} 1 & 3 & 5 & 14 \\ 0 & 7 & 13 & 25 \\ 0 & 0 & 8 & 24 \end{array}\right]$$

The system is now in triangular form.

Stage 3: Back Substitution

Here we solve the equations from bottom to top. At each step of the back substitution process we encounter equations which only have a **single** unknown and so can be easily solved.

Task

Now complete the solution to the above system by back-substitution.

Your solution

Answer

In full the equations are

$$\begin{aligned} x_1 + 3x_2 + 5x_3 &= 14 \\ 7x_2 + 13x_3 &= 25 \\ 8x_3 &= 24 \end{aligned}$$

From the last equation we see that $x_3 = 3$.

Substituting this value into the second equation gives

$$7x_2 + 39 = 25 \quad \text{or} \quad 7x_2 = -14 \quad \text{so that} \quad x_2 = -2.$$

Finally, using these values for x_2 and x_3 in equation 1 gives $x_1 - 6 + 15 = 14$. Hence $x_1 = 5$. The solution is therefore $[x_1, x_2, x_3]^T = [5, -2, 3]^T$

Check that these values satisfy the original system of equations.

Task

Solve

$$\begin{aligned} 2x_1 - 3x_2 + 4x_3 &= 2 \\ 4x_1 + x_2 + 2x_3 &= 2 \\ x_1 - x_2 + 3x_3 &= 3 \end{aligned}$$

Write down the augmented matrix for this system and then interchange rows 1 and 3:

Your solution

Answer

Augmented matrix

$$\left[\begin{array}{rrr|r} 2 & -3 & 4 & 2 \\ 4 & 1 & 2 & 2 \\ 1 & -1 & 3 & 3 \end{array}\right] \quad R1 \leftrightarrow R3 \quad \Rightarrow \left[\begin{array}{rrr|r} 1 & -1 & 3 & 3 \\ 4 & 1 & 2 & 2 \\ 2 & -3 & 4 & 2 \end{array}\right]$$

Now subtract suitable multiples of row 1 from row 2 and from row 3 to eliminate the x_1 coefficient from rows 2 and 3:

Your solution

Answer

$$\left[\begin{array}{rrr|r} 1 & -1 & 3 & 3 \\ 4 & 1 & 2 & 2 \\ 2 & -3 & 4 & 2 \end{array}\right] \quad \begin{array}{l} \\ R2 - 4R1 \\ R3 - 2R1 \end{array} \quad \Rightarrow \left[\begin{array}{rrr|r} 1 & -1 & 3 & 3 \\ 0 & 5 & -10 & -10 \\ 0 & -1 & -2 & -4 \end{array}\right]$$

Now divide row 2 by 5 and add a suitable multiple of the result to row 3:

Your solution

Answer

$$\left[\begin{array}{ccc|c} 1 & -1 & 3 & 3 \\ 0 & 5 & -10 & -10 \\ 0 & -1 & -2 & -4 \end{array}\right] \quad R2 \div 5 \Rightarrow \left[\begin{array}{ccc|c} 1 & -1 & 3 & 3 \\ 0 & 1 & -2 & -2 \\ 0 & -1 & -2 & -4 \end{array}\right] \quad \begin{array}{r} \\ \Rightarrow \\ R3+R2 \end{array} \left[\begin{array}{ccc|c} 1 & -1 & 3 & 3 \\ 0 & 1 & -2 & -2 \\ 0 & 0 & -4 & -6 \end{array}\right]$$

Now complete the solution using back-substitution:

Your solution

Answer

The equations in full are

$$\begin{aligned} x_1 - x_2 + 3x_3 &= 3 \\ x_2 - 2x_3 &= -2 \\ -4x_3 &= -6. \end{aligned}$$

The last equation reduces to $x_3 = \frac{3}{2}$.

Using this value in the second equation gives $x_2 - 3 = -2$ so that $x_2 = 1$.

Finally, $x_1 - 1 + \frac{9}{2} = 3$ so that $x_1 = -\frac{1}{2}$.

The solution is therefore $[x_1, x_2, x_3]^T = \left[-\frac{1}{2}, 1, \frac{3}{2}\right]^T$.

You should check these values in the original equations to ensure that the equations balance.
Again we emphasise that we chose a particular set of procedures in Stage 1. This was chosen mainly to keep the arithmetic simple by delaying the introduction of fractions. Sometimes we are courageous and take fewer, harder steps.

An important point to note is that when in Stage 2 we wrote $R2 - 2R1$ against row 2; what we meant is that row 2 is replaced by the combination (row $2) - 2\times$(row 1).
In general, the operation

$$\text{row } i - \alpha \times \text{row } j$$

means replace **row** i by the combination

$$\text{row } i - \alpha \times \text{row } j.$$

3. Equations which have an infinite number of solutions

Consider the following system of equations

$$\begin{aligned} x_1 + x_2 - 3x_3 &= 3 \\ 2x_1 - 3x_2 + 4x_3 &= -4 \\ x_1 - x_2 + x_3 &= -1 \end{aligned}$$

In augmented form we have:

$$\left[\begin{array}{rrr|r} 1 & 1 & -3 & 3 \\ 2 & -3 & 4 & -4 \\ 1 & -1 & 1 & -1 \end{array}\right]$$

Now performing the usual Gauss elimination operations we have

$$\left[\begin{array}{rrr|r} 1 & 1 & -3 & 3 \\ 2 & -3 & 4 & -4 \\ 1 & -1 & 1 & -1 \end{array}\right] \begin{array}{c} \\ R2 - 2\times R1 \\ R3 - R1 \end{array} \quad \Rightarrow \quad \left[\begin{array}{rrr|r} 1 & 1 & -3 & 3 \\ 0 & -5 & 10 & -10 \\ 0 & -2 & 4 & -4 \end{array}\right]$$

Now applying $R2 \div -5$ and $R3 \div -2$ gives

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

Then $R2 - R3$ gives

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

We see that all the elements in the last row are zero. This means that the variable x_3 can take any value whatsoever, so let $x_3 = t$ then using back substitution the second row now implies

$$x_2 = 2 + 2x_3 = 2 + 2t$$

and then the first row implies

$$x_1 = 3 - x_2 + 3x_3 = 3 - (2 + 2t) + 3(t) = 1 + t$$

In this example the system of equations has an infinite number of solutions:

$$x_1 = 1 + t, \qquad x_2 = 2 + 2t, \qquad x_3 = t \quad \text{or} \quad [x_1, x_2, x_3]^T = [1 + t, 2 + 2t, t]^T$$

where t can be assigned any value. For every value of t these expressions for x_1, x_2 and x_3 will simultaneously satisfy each of the three given equations.

Systems of linear equations arise in the modelling of electrical circuits or networks. By breaking down a complicated system into simple loops, Kirchhoff's law can be applied. This leads to a set of linear equations in the unknown quantities (usually currents) which can easily be solved by one of the methods described in this Workbook.

Engineering Example 3

Currents in three loops

In the circuit shown find the currents (i_1, i_2, i_3) in the loops.

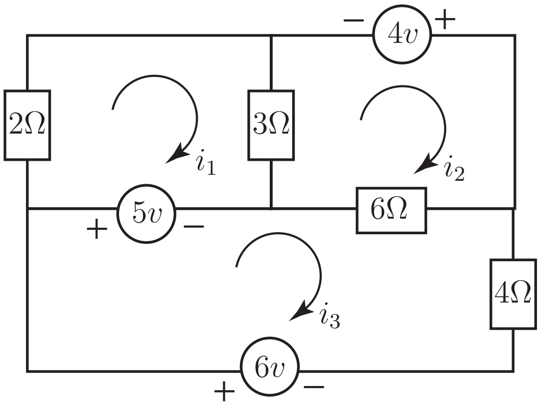


Figure 2

Solution

Loop 1 gives

$$2(i_1) + 3(i_1 - i_2) = 5 \quad \rightarrow \quad 5i_1 - 3i_2 = 5$$

Loop 2 gives

$$6(i_2 - i_3) + 3(i_2 - i_1) = 4 \quad \rightarrow \quad -3i_1 + 9i_2 - 6i_3 = 4$$

Loop 3 gives

$$6(i_3 - i_2) + 4(i_3) = 6 - 5 \quad \rightarrow \quad -6i_2 + 10i_3 = 1$$

Note that in loop 3, the current generated by the $6v$ cell is positive and for the $5v$ cell negative in the direction of the arrow.
In matrix form

$$\begin{bmatrix} 5 & -3 & 0 \\ -3 & 9 & -6 \\ 0 & -6 & 10 \end{bmatrix} \begin{bmatrix} i_1 \\ i_2 \\ i_3 \end{bmatrix} = \begin{bmatrix} 5 \\ 4 \\ 1 \end{bmatrix}$$

Solving gives

$$i_1 = \frac{34}{15}, \quad i_2 = \frac{19}{9}, \quad i_3 = \frac{41}{30}$$

Engineering Example 4

Velocity of a rocket

The upward velocity of a rocket, measured at 3 different times, is shown in the following table

Time, t (seconds)	Velocity, v (metres/second)
5	106.8
8	177.2
12	279.2

The velocity over the time interval $5 \leq t \leq 12$ is approximated by a quadratic expression as

$$v(t) = a_1t^2 + a_2t + a_3$$

Find the values of a_1, a_2 and a_3.

Solution

Substituting the values from the table into the quadratic equation for $v(t)$ gives:

$$\begin{aligned} 106.8 &= 25a_1 + 5a_2 + a_3 \\ 177.2 &= 64a_1 + 8a_2 + a_3 \\ 279.2 &= 144a_1 + 12a_2 + a_3 \end{aligned} \qquad \text{or} \qquad \begin{bmatrix} 25 & 5 & 1 \\ 64 & 8 & 1 \\ 144 & 12 & 1 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} 106.8 \\ 177.2 \\ 279.2 \end{bmatrix}$$

Applying one of the methods from this Workbook gives the solution as

$$a_1 = 0.2905 \qquad a_2 = 19.6905 \qquad a_3 = 1.0857 \quad \text{to 4 d.p.}$$

As the original values were all **experimental observations** then the values of the unknowns are all **approximations**. The relation $v(t) = 0.2905t^2 + 19.6905t + 1.0857$ can now be used to predict the approximate position of the rocket for any time within the interval $5 \leq t \leq 12$.

Exercises

Solve the following using Gauss elimination:

1.
$$\begin{array}{rcrcrcr} 2x_1 & + & x_2 & - & x_3 & = & 0 \\ x_1 & & & + & x_3 & = & 4 \\ x_1 & + & x_2 & + & x_3 & = & 0 \end{array}$$

2.
$$\begin{array}{rcrcrcr} x_1 & - & x_2 & + & x_3 & = & 1 \\ -x_1 & & & + & x_3 & = & 1 \\ x_1 & + & x_2 & - & x_3 & = & 0 \end{array}$$

3.
$$\begin{array}{rcrcrcr} x_1 & + & x_2 & + & x_3 & = & 2 \\ 2x_1 & + & 3x_2 & + & 4x_3 & = & 3 \\ x_1 & - & 2x_2 & - & x_3 & = & 1 \end{array}$$

4.
$$\begin{array}{rcrcrcr} x_1 & - & 2x_2 & - & 3x_3 & = & -1 \\ 3x_1 & + & x_2 & + & x_3 & = & 4 \\ 11x_1 & - & x_2 & - & 3x_3 & = & 10 \end{array}$$

You may need to think carefully about this system.

Answers

(1) $x_1 = \dfrac{8}{3},\ x_2 = -4,\ x_3 = \dfrac{4}{3}$

(2) $x_1 = \dfrac{1}{2},\ x_2 = 1,\ x_3 = \dfrac{3}{2}$

(3) $x_1 = 2,\ x_2 = 1,\ x_3 = -1$

(4) infinite number of solutions: $x_1 = t,\ x_2 = 11 - 10t,\ x_3 = 7t - 7$

1.

$$\begin{aligned} 2x_1 + x_2 - x_3 &= 0 \\ x_1 + x_3 &= 4 \\ x_1 + x_2 + x_3 &= 0 \end{aligned}$$

$x_1 + x_2 + x_3 = 0$
$x_1 + x_3 = 4$ $\xrightarrow{-R1}$
$2x_1 + x_2 - x_3 = 0$ $\xrightarrow{-2R1}$

$x_1 + x_2 + x_3 = 0$
$-x_2 = 4$
$-x_2 - 3x_3 = 0$

$\Downarrow -R2$

$x_1 + x_2 + x_3 = 0$
$-x_2 = 4$
$-3x_3 = -4$

$4/3 \cdot 12/3$

$\boxed{x_3 = 4/3 \quad x_2 = -4 \quad x_1 = 8/3}$

2.

$$\begin{aligned} x_1 - x_2 + x_3 &= 1 \\ -x_1 + x_3 &= 1 \\ x_1 + x_2 - x_3 &= 0 \end{aligned}$$

$+R1 \Downarrow$ $\quad -R1 \Downarrow$

$x_1 - x_2 + x_3 = 1$
$-x_2 + 2x_3 = 2$
$2x_2 - 2x_3 = -1$

$+2R2 \Downarrow$

$x_1 - x_2 + x_3 = 1$
$-x_2 + 2x_3 = 2$
$2x_3 = 3$

$x_3 = 3/2$

$-x_2 + 2 \cdot 3/2 = 2$
$-x_2 = -1$
$x_2 = 1$

$x_1 - 1 + 3/2 = 1$
$x_1 + 3/2 = 2$
$x_1 = 1/2$

$\boxed{x_1 = 1/2 \quad x_2 = 1 \quad x_3 = 3/2}$

3.

$$\begin{aligned} x_1 + x_2 + x_3 &= 2 \\ 2x_1 + 3x_2 + 4x_3 &= 3 \\ x_1 - 2x_2 - x_3 &= 1 \end{aligned}$$

$-2R2 \Downarrow$ $\quad -R1 \Downarrow$

$x_1 + x_2 + x_3 = 2$
$x_2 + 2x_3 = -1$
$-3x_2 - 2x_3 = -1$ $\xrightarrow{+3R2}$

$x_1 + x_2 + x_3 = 2$
$x_2 + 2x_3 = -1$
$4x_3 = -4$

$x_2 - 2 = -1$

$\boxed{x_3 = -1, x_2 = 1 \quad x_1 = 2}$

4.

$$\begin{aligned} x_1 - 2x_2 - 3x_3 &= -1 \\ 3x_1 + x_2 + x_3 &= 4 \\ 11x_1 - x_2 - 3x_3 &= 10 \end{aligned}$$

$-3R1 \Downarrow$ $\quad -11R1 \Downarrow$

$x_1 - 2x_2 - 3x_3 = -1$
$7x_2 + 10x_3 = 7$
$21x_2 + 30x_3 = 21$ $\xrightarrow{-3R2}$

$x_1 - 2x_2 - 3x_3 = -1$
$7x_2 + 10x_3 = 7$
$0 = 0$

$\boxed{\infty \text{ solutions}}$

$7x_2 + 10x_3 = 7$

$x_3 = \dfrac{-7x_2 + 7}{10}$

$x_2 = \dfrac{-10x_3 + 7}{7}$

$x_1 - 2\left(\dfrac{-10x_3 + 7}{7}\right) - 3\left(\dfrac{-7x_2 + 7}{10}\right) = -1$

Homework – 1

1. Read the document "Gaussian_elimination.pdf". Use the exercises in the text to check your understanding. **Solve the problems at the end of the document.**

2. Given matrix $A=\begin{bmatrix} 0 & -1 & 1 & 1 \\ 2 & 1 & 0 & 1 \\ -1 & 2 & -1 & 0 \\ 2 & -1 & -2 & 1 \end{bmatrix}$ and vectors $x=\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}, b=\begin{bmatrix} 3 \\ 0 \\ -1 \\ 2 \end{bmatrix}$. Write the system of linear algebraic equations **Ax=b** and solve it using the Gaussian elimination.

3. Consider the following matrices: $A=\begin{bmatrix} 0 & -1 & 1 \\ 1 & 1 & -1 \end{bmatrix}$, $B=\begin{bmatrix} 0 & 1 \\ -1 & 2 \\ 1 & 0 \\ -1 & 1 \end{bmatrix}$, $C=\begin{bmatrix} 0 & -1 \\ 1 & 1 \end{bmatrix}$ and $D=\begin{bmatrix} 0 & -1 \\ 2 & 1 \\ -1 & 2 \end{bmatrix}$. Which of the following expressions are well defined? Perform the computations whenever possible:

   a) AB+C;
   b) BC+3A;
   c) 2AD+C;
   d) DCA;
   e) ACD.

2. Given matrix $A=\begin{bmatrix} 0 & -1 & 1 & 1 \\ 2 & 1 & 0 & 1 \\ -1 & 2 & -1 & 0 \\ 2 & -1 & -2 & 1 \end{bmatrix}$ and vectors $x=\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}, b=\begin{bmatrix} 3 \\ 0 \\ -1 \\ 2 \end{bmatrix}$. Write the system of linear algebraic equations **Ax=b** and solve it using the Gaussian elimination.

$$\left[\begin{array}{cccc|c} 0 & -1 & 1 & 1 & 3 \\ 2 & 1 & 0 & 1 & 0 \\ -1 & 2 & -1 & 0 & -1 \\ 2 & -1 & -2 & 1 & 2 \end{array}\right] \xrightarrow{\cdot -1} \left[\begin{array}{cccc|c} 1 & -2 & 1 & 0 & 1 \\ 0 & 1 & -1 & -1 & -3 \\ 2 & 1 & 0 & 1 & 0 \\ 2 & -1 & -2 & 1 & 2 \end{array}\right] \underset{-2R1}{\overset{-2R1}{\Longrightarrow}} \left[\begin{array}{cccc|c} 1 & -2 & 1 & 0 & 1 \\ 0 & 1 & -1 & -1 & -3 \\ 0 & 5 & -2 & 0 & -2 \\ 0 & 3 & -4 & 1 & 0 \end{array}\right] \underset{-3R2}{\overset{-5R2}{\Longrightarrow}}$$

$$\left[\begin{array}{cccc|c} 1 & -2 & 1 & 0 & 1 \\ 0 & 1 & -1 & -1 & -3 \\ 0 & 0 & 3 & 5 & 13 \\ 0 & 0 & -1 & 4 & 9 \end{array}\right] \xrightarrow{\cdot -1} \left[\begin{array}{cccc|c} 1 & -2 & 1 & 0 & 1 \\ 0 & 1 & -1 & -1 & -3 \\ 0 & 0 & 1 & -4 & -9 \\ 0 & 0 & 3 & 5 & 13 \end{array}\right] \overset{-3R3}{\Longrightarrow} \left[\begin{array}{cccc|c} 1 & -2 & 1 & 0 & 1 \\ 0 & 1 & -1 & -1 & -3 \\ 0 & 0 & 1 & -4 & -9 \\ 0 & 0 & 0 & 17 & 30 \end{array}\right]$$

$x_4 = 30/17$

$x_3 - 4(30/17) = -9$

$x_3 - 120/17 = -9$

$x_3 = -33/17$

$x_2 + 33/17 - 30/17 = -3$

$x_2 + 3/17 = -3$

$x_2 = -54/17$

$x_1 + 2(-54/17) - 33/17 = 1$

$x_1 = -58/17$

$$\boxed{\begin{aligned} x_1 &= -58/17 \\ x_2 &= -54/17 \\ x_3 &= -33/17 \\ x_4 &= 30/17 \end{aligned}}$$

3. Consider the following matrices: $A=\begin{bmatrix} 0 & -1 & 1 \\ 1 & 1 & -1 \end{bmatrix}$ (2×3), $B=\begin{bmatrix} 0 & 1 \\ -1 & 2 \\ 1 & 0 \\ -1 & 1 \end{bmatrix}$ (4×2), $C=\begin{bmatrix} 0 & -1 \\ 1 & 1 \end{bmatrix}$ (2×2) and $D=\begin{bmatrix} 0 & -1 \\ 2 & 1 \\ -1 & 2 \end{bmatrix}$ (3×2). Which of the following expressions are well defined? Perform the computations whenever possible:

a) AB+C; ✗
b) BC+3A; ✗
c) 2AD+C; ✓
d) DCA; ✓
e) ACD. ✗

(c) $\begin{bmatrix} 0 & -2 & 2 \\ 2 & 2 & -2 \end{bmatrix}\begin{bmatrix} 0 & -1 \\ 2 & 1 \\ -1 & 2 \end{bmatrix}$

$\begin{bmatrix} -6 & 2 \\ 6 & -4 \end{bmatrix} + \begin{bmatrix} 0 & -1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} -6 & 1 \\ 7 & -3 \end{bmatrix}$

(d) $\begin{bmatrix} -1 & -1 \\ 1 & -1 \\ 2 & 3 \end{bmatrix}\begin{bmatrix} 0 & -1 & 1 \\ 1 & 1 & -1 \end{bmatrix} = \begin{bmatrix} -1 & 0 & 0 \\ -1 & -2 & 2 \\ 3 & 1 & -1 \end{bmatrix}$

Week 2 Notes

Classification of matrices

① Square Matrix

$[n \times n]$

Most cases we will used

② Diagonal Matrix

$$A = \begin{bmatrix} \ddots & & 0 \\ & \ddots & \\ 0 & & \ddots \end{bmatrix}$$

$a_{ii} \in R, \forall i$

$a_{ij} = 0, \; i \neq j$

Identity matrix

$$I \begin{bmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{bmatrix}$$

③ Zero Matrix

$$0 = \begin{bmatrix} 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{bmatrix}$$

④ Triangular Matrix

$$A_{\triangledown} = \begin{cases} a_{ij} \in \mathbb{R}, & i \leq j \\ a_{ij} = 0, & i > j \end{cases}$$ upper triangular

Example

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 0 & 0 & a_{33} \end{bmatrix}$$

Why special for I

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = A$$

$$AI = IA = A$$

Matrix operation

Transposition

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}^T = \begin{bmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{bmatrix}$$

$$[n \times m]^T = [m \times n]$$

$$[a_{ij}]^T = [a_{ji}]$$

Properties of Transposition

① $[\alpha A]^T = \alpha A^T$

② $[A+B]^T = A^T + B^T$

③ $[AB]^T = B^T A^T$

3. Consider the following matrices: $A=\begin{bmatrix} 0 & -1 & 1 \\ 1 & 1 & -1 \end{bmatrix}$, $B=\begin{bmatrix} 0 & 1 \\ -1 & 2 \\ 1 & 0 \\ -1 & 1 \end{bmatrix}$, $C=\begin{bmatrix} 0 & -1 \\ 1 & 1 \end{bmatrix}$ and $D=\begin{bmatrix} 0 & -1 \\ 2 & 1 \\ -1 & 2 \end{bmatrix}$. Which of the following expressions are well defined? Perform the computations whenever possible:

$$(AD)^T + C \qquad \begin{bmatrix} -3 & 1 \\ 3 & -2 \end{bmatrix}$$

$$D^T A^T + C$$

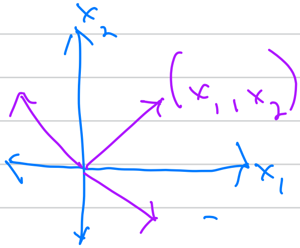
$$\begin{bmatrix} 0 & 2 & -1 \\ -1 & 1 & 2 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ -1 & 1 \\ 1 & -1 \end{bmatrix} + \begin{bmatrix} 0 & -1 \\ 1 & 1 \end{bmatrix}$$

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

⑤ Block Matrix

$$\left[\begin{array}{cc|cc} 1 & 2 & 3 & 4 \\ 1 & 3 & 4 & 5 \\ \hline 0 & 0 & 1 & 2 \\ 0 & 0 & -1 & 2 \end{array}\right]$$

Intro to vectors



Operations

1. Multiply by $\alpha \in \mathbb{R}$
2. Addition

$a = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix}, \; b = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$ $\quad a+b = \begin{bmatrix} a+b \\ \cdots \\ \cdots \\ \cdots \end{bmatrix}$

$x = \begin{bmatrix} x_1 \\ \vdots \\ \vdots \\ x_n \end{bmatrix}, \; x \in \mathbb{R}^n$

Matrix as an operator

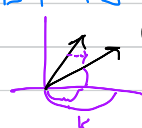
$A = [n \times n]$

$x \in \mathbb{R}^n$

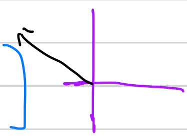
$Ax = y$

$[n \times n][n \times 1] = [n \times 1]$

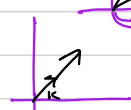
1. $A = \begin{bmatrix} k & 0 \\ 0 & 1 \end{bmatrix}$



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



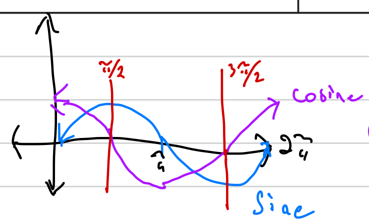
2. $A = \begin{bmatrix} k & 0 \\ 0 & k \end{bmatrix}$



Rotation Matrix

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

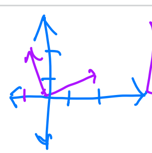
$\theta \in [0, 2\pi]$



$\theta = 0 \quad A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

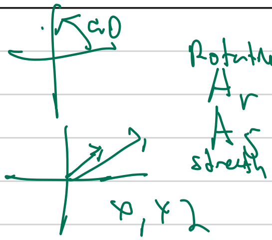
$\theta = \pi \quad A = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$

$\theta = \pi/2 \quad A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$



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

Rotation by θ by counterclockwise



$A_R = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$

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

$\boxed{AB \neq BA}$

$A_S(A_R x) = (A_S A_R) x$

$x = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ $\quad \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$

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

$\begin{bmatrix} -2 \\ 1 \end{bmatrix}$ $\quad \begin{bmatrix} -2 \\ 1 \end{bmatrix}$

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

1. Matrices can transform vectors
2. Order matters

Inverse Matrices

Given A, $[n \times n]$,

A^{-1} is said to be inverse of A if $A^{-1} \cdot A = I$

Inverse Matrices cont.

Properties

1. Not every sq matrix has an inverse

Assume A^{-1} exist

2. $(\alpha A)^{-1} = \alpha^{-1} A^{-1}, \ \alpha \neq 0$

Assume A^{-1}, B^{-1} exist

3. $(AB)^{-1} = B^{-1} A^{-1}$
4. $(A^T)^{-1} = (A^{-1})^T$

Contents 7

Matrices

Learning outcomes

In this Workbook you will learn about matrices. In the first instance you will learn about the algebra of matrices: how they can be added, subtracted and multiplied. You will learn about a characteristic quantity associated with square matrices - the determinant. Using knowledge of determinants you will learn how to find the inverse of a matrix. Also, a second method for finding a matrix inverse will be outlined - the Gaussian elimination method.

A working knowledge of matrices is a vital attribute of any mathematician, engineer or scientist. You will find that matrices arise in many varied areas of science.

Introduction to Matrices

7.1

Introduction

When we wish to solve large systems of simultaneous linear equations, which arise for example in the problem of finding the forces on members of a large framed structure, we can isolate the coefficients of the variables as a block of numbers called a matrix. There are many other applications matrices. In this Section we develop the terminology and basic properties of a matrix.

Prerequisites

Before starting this Section you should ...

- be familiar with the rules of number algebra

Learning Outcomes

On completion you should be able to ...

- express a system of linear equations in matrix form
- recognise and use the basic terminology associated with matrices
- carry out addition and subtraction with two given matrices or state that the operation is not possible

1. Applications of matrices

The solution of simultaneous linear equations is a task frequently occurring in engineering. In electrical engineering the analysis of circuits provides a ready example.
However the simultaneous equations arise, we need to study two things:

(a) how we can conveniently represent large systems of linear equations

(b) how we might find the solution of such equations.

We shall discover that knowledge of the theory of matrices is an essential mathematical tool in this area.

Representing simultaneous linear equations

Suppose that we wish to solve the following three equations in three unknowns x_1, x_2 and x_3:

$$\begin{aligned} 3x_1 + 2x_2 - x_3 &= 3 \\ x_1 - x_2 + x_3 &= 4 \\ 2x_1 + 3x_2 + 4x_3 &= 5 \end{aligned}$$

We can isolate three facets of this system: the **coefficients** of x_1, x_2, x_3; the **unknowns** x_1, x_2, x_3; and the **numbers** on the right-hand sides.
Notice that in the system

$$\begin{aligned} 3x + 2y - z &= 3 \\ x - y + z &= 4 \\ 2x + 3y + 4z &= 5 \end{aligned}$$

the only difference from the first system is the names given to the unknowns. It can be checked that the first system has the solution $x_1 = 2,\ x_2 = -1,\ x_3 = 1$. The second system therefore has the solution $x = 2, y = -1, z = 1$.
We can isolate the three facets of the first system by using **arrays** of numbers and of unknowns:

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

Even more conveniently we represent the arrays with letters (usually capital letters)

$$AX = B$$

Here, to be explicit, we write

$$A = \begin{bmatrix} 3 & 2 & -1 \\ 1 & -1 & 1 \\ 2 & 3 & 4 \end{bmatrix} \qquad X = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \qquad B = \begin{bmatrix} 3 \\ 4 \\ 5 \end{bmatrix}$$

Here A is called the **matrix of coefficients**, X is called the **matrix of unknowns** and B is called the **matrix of constants**.
If we now append to A the column of right-hand sides we obtain the **augmented matrix** for the system:

$$\left[\begin{array}{rrr|r} 3 & 2 & -1 & 3 \\ 1 & -1 & 1 & 4 \\ 2 & 3 & 4 & 5 \end{array}\right]$$

The order of the entries, or elements, is crucial. For example, all the entries in the second row relate to the second equation, the entries in column 1 are the coefficients of the unknown x_1, and those in the last column are the constants on the right-hand sides of the equations.
In particular, the entry in row 2 column 3 is the coefficient of x_3 in equation 2.

Representing networks

Shortest-distance problems are important in communications study. Figure 1 illustrates schematically a system of four towns connected by a set of roads.

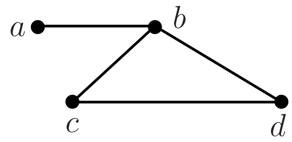


Figure 1

The system can be represented by the matrix

$$\begin{array}{c} \\ a \\ b \\ c \\ d \end{array}\begin{array}{c} \begin{array}{cccc} a & b & c & d \end{array} \\ \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{bmatrix} \end{array}$$

The row refers to the town from which the road starts and the column refers to the town where the road ends. An entry of 1 indicates that two towns are directly connected by a road (for example b and d) and an entry of zero indicates that there is no direct road (for example a and c). Of course, if there is a road from b to d (say) it is also a road from d to b.

In this Section we shall develop some basic ideas about matrices.

2. Definitions

An array of numbers, rectangular in shape, is called a **matrix**. The first matrix below has 3 rows and 2 columns and is said to be a '3 by 2' matrix (written 3×2). The second matrix is a '2 by 4' matrix (written 2×4).

$$\begin{bmatrix} 1 & 4 \\ -2 & 3 \\ 2 & 1 \end{bmatrix} \qquad \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 9 \end{bmatrix}$$

The general 3×3 matrix can be written

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

where a_{ij} denotes the element in row i, column j.
For example in the matrix:

$$A = \begin{bmatrix} 0 & -1 & -3 \\ 0 & 6 & -12 \\ 5 & 7 & 123 \end{bmatrix}$$

$$a_{11} = 0, \qquad a_{12} = -1, \qquad a_{13} = -3, \quad \ldots \quad a_{22} = 6, \quad \ldots \quad a_{32} = 7, \quad a_{33} = 123$$

Key Point 1

The General Matrix

A general $m \times n$ matrix A has m rows and n columns.

The entries in the matrix A are called the **elements** of A.

In matrix A the element in row i and column j is denoted by a_{ij}.

A matrix with only one column is called a **column vector** (or **column matrix**).

For example, $\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ and $\begin{bmatrix} 3 \\ 4 \\ 5 \end{bmatrix}$ are both 3×1 column vectors.

A matrix with only one row is called a **row vector** (or **row matrix**). For example $[2, -3, 8, 9]$ is a 1×4 row vector. Often the entries in a row vector are separated by commas for clarity.

Square matrices

When the number of rows is the same as the number of columns, i.e. $m = n$, the matrix is said to be **square** and of **order** n (or m).

- In an $n \times n$ square matrix A, the **leading diagonal** (or **principal diagonal**) is the 'north-west to south-east' collection of elements $a_{11}, a_{22}, \ldots, a_{nn}$. The sum of the elements in the leading diagonal of A is called the **trace** of the matrix, denoted by tr(A).

$$A = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{bmatrix} \qquad \text{tr}(A) = a_{11} + a_{22} + \cdots + a_{nn}$$

- A square matrix in which all the elements below the leading diagonal are zero is called an **upper triangular matrix**, often denoted by U.

$$U = \begin{bmatrix} u_{11} & u_{12} & \dots & \dots & u_{1n} \\ 0 & u_{22} & \dots & \dots & u_{2n} \\ 0 & 0 & \dots & \vdots & \vdots \\ 0 & 0 & \dots & 0 & u_{nn} \end{bmatrix} \qquad u_{ij} = 0 \quad \text{when} \quad i > j$$

- A square matrix in which all the elements above the leading diagonal are zero is called a **lower triangular matrix**, often denoted by L.

$$L = \begin{bmatrix} l_{11} & 0 & 0 & \dots & 0 \\ l_{21} & l_{22} & 0 & \dots & 0 \\ \vdots & \vdots & \dots & \dots & 0 \\ l_{n1} & l_{n2} & \vdots & \dots & l_{nn} \end{bmatrix} \qquad l_{ij} = 0 \quad \text{when} \quad i < j$$

- A square matrix where all the non-zero elements are along the leading diagonal is called a diagonal matrix, often denoted by D.

$$D = \begin{bmatrix} d_{11} & 0 & 0 & \dots & 0 \\ 0 & d_{22} & 0 & \dots & 0 \\ 0 & 0 & \dots & \dots & 0 \\ 0 & 0 & 0 & \dots & d_{nn} \end{bmatrix} \qquad d_{ij} = 0 \quad \text{when} \quad i \neq j$$

Some examples of matrices and their classification

$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$ is 2×3. It is not square.

$B = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ is 2×2. It is square.

Also, tr(A) does not exist, and tr$(B) = 1 + 4 = 5$.

$C = \begin{bmatrix} 1 & 2 & 3 \\ 0 & -2 & -5 \\ 0 & 0 & 1 \end{bmatrix}$ and $D = \begin{bmatrix} 4 & 0 & 3 \\ 0 & -2 & 5 \\ 0 & 0 & 1 \end{bmatrix}$ are both 3×3, square and upper triangular.

Also, tr$(C) = 0$ and tr$(D) = 3$.

$E = \begin{bmatrix} 1 & 0 & 0 \\ 2 & -2 & 0 \\ 3 & -5 & 1 \end{bmatrix}$ and $F = \begin{bmatrix} -1 & 0 & 0 \\ 1 & 4 & 0 \\ 0 & 1 & 1 \end{bmatrix}$ are both 3×3, square and lower triangular.

Also, tr$(E) = 0$ and tr$(F) = 4$.

$G = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -3 \end{bmatrix}$ and $H = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ are both 3×3, square and diagonal.

Also, tr$(G) = 0$ and tr$(H) = 6$.

Task

Classify the following matrices (and, where possible, find the trace):

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} \quad B = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ -1 & -3 & -2 & -4 \end{bmatrix} \quad C = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \\ 13 & 14 & 15 & 16 \end{bmatrix}$$

Your solution

Answer

A is 3×2, B is 3×4, C is 4×4 and square.

The trace is not defined for A or B. However, tr$(C) = 34$.

Task

Classify the following matrices:

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \quad B = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \quad C = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \quad D = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Your solution

Answer

A is 3×3 and square, B is 3×3 lower triangular, C is 3×3 upper triangular and D is 3×3 diagonal.

Equality of matrices

As we noted earlier, the terms in a matrix are called the **elements** of the matrix.

The elements of the matrix $A = \begin{bmatrix} 1 & 2 \\ -1 & -4 \end{bmatrix}$ are $1, 2, -1, -4$

We say two matrices A, B are **equal** to each other only if A and B have the same number of rows and the same number of columns and if each element of A is equal to the corresponding element of B. When this is the case we write $A = B$. For example if the following two matrices are equal:

$$A = \begin{bmatrix} 1 & \alpha \\ -1 & -\beta \end{bmatrix} \qquad B = \begin{bmatrix} 1 & 2 \\ -1 & -4 \end{bmatrix}$$

then we can conclude that $\alpha = 2$ and $\beta = 4$.

The unit matrix

The **unit matrix** or the **identity matrix**, denoted by I_n (or, often, simply I), is the diagonal matrix of order n in which all diagonal elements are 1.

Hence, for example, $I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ and $I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$.

The zero matrix

The **zero matrix** or **null matrix** is the matrix all of whose elements are zero. There is a zero matrix for every size. For example the 2×3 and 2×2 cases are:

$$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

Zero matrices, of whatever size, are denoted by $\underline{0}$.

The transpose of a matrix

The **transpose** of a matrix A is a matrix where the rows of A become the columns of the new matrix and the columns of A become its rows. For example

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

The resulting matrix is called the **transposed matrix** of A and denoted A^T. In the previous example it is clear that A^T is not equal to A since the matrices are of different sizes. If A is square $n \times n$ then A^T will also be $n \times n$.

Example 1

Find the transpose of the matrix $B = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$

Solution

Interchanging rows with columns we find

$$B^T = \begin{bmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{bmatrix}$$

Both matrices are 3×3 but B and B^T are clearly different.

When the transpose of a matrix is equal to the original matrix i.e. $A^T = A$, then we say that the matrix A is **symmetric**. (This is because it has symmetry about the leading diagonal.)
In Example 1 B is **not** symmetric.

Example 2

Show that the matrix $C = \begin{bmatrix} 1 & -2 & 3 \\ -2 & 4 & -5 \\ 3 & -5 & 6 \end{bmatrix}$ is symmetric.

Solution

Taking the transpose of C:

$$C^T = \begin{bmatrix} 1 & -2 & 3 \\ -2 & 4 & -5 \\ 3 & -5 & 6 \end{bmatrix}.$$

Clearly $C^T = C$ and so C is a symmetric matrix. Notice how the leading diagonal acts as a "mirror"; for example $c_{12} = -2$ and $c_{21} = -2$. In general $c_{ij} = c_{ji}$ for a symmetric matrix.

Task

Find the transpose of each of the following matrices. Which are symmetric?

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

$D = \begin{bmatrix} 1 & 2 \\ 4 & 5 \\ 7 & 8 \end{bmatrix} \quad E = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

Your solution

Answer

$A^T = \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix}, \quad B^T = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \quad C^T = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} = C,$ symmetric

$D^T = \begin{bmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \end{bmatrix} \quad E^T = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = E,$ symmetric

3. Addition and subtraction of matrices

Under what circumstances can we add two matrices i.e. define $A + B$ for given matrices A, B?

Consider

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \qquad \text{and} \qquad B = \begin{bmatrix} 5 & 6 & 9 \\ 7 & 8 & 10 \end{bmatrix}$$

There is no sensible way to define $A + B$ in this case since A and B are different sizes.
However, if we consider matrices of the same size then addition can be defined in a very natural way. Consider $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ and $B = \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix}$. The 'natural' way to add A and B is to add corresponding elements together:

$$A + B = \begin{bmatrix} 1+5 & 2+6 \\ 3+7 & 4+8 \end{bmatrix} = \begin{bmatrix} 6 & 8 \\ 10 & 12 \end{bmatrix}$$

In general if A and B are both $m \times n$ matrices, with elements a_{ij} and b_{ij} respectively, then their sum is a matrix C, also $m \times n$, such that the elements of C are

$$c_{ij} = a_{ij} + b_{ij} \qquad i = 1, 2, \ldots, m \quad j = 1, 2, \ldots, n$$

In the above example

$$c_{11} = a_{11} + b_{11} = 1 + 5 = 6 \qquad c_{21} = a_{21} + b_{21} = 3 + 7 = 10 \quad \text{and so on.}$$

Subtraction of matrices follows along similar lines:

$$D = A - B = \begin{bmatrix} 1-5 & 2-6 \\ 3-7 & 4-8 \end{bmatrix} = \begin{bmatrix} -4 & -4 \\ -4 & -4 \end{bmatrix}$$

4. Multiplication of a matrix by a number

There is also a natural way of defining the product of a matrix with a number. Using the matrix A above, we note that

$$A + A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} + \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} 2 & 4 \\ 6 & 8 \end{bmatrix}$$

What we see is that $2A$ (which is the shorthand notation for $A + A$) is obtained by multiplying *every* element of A by 2.
In general if A is an $m \times n$ matrix with typical element a_{ij} then the product of a number k with A is written kA and has the corresponding elements ka_{ij}.

Hence, again using the matrix A above,

$$7A = 7\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = \begin{bmatrix} 7 & 14 \\ 21 & 28 \end{bmatrix}$$

Similarly:

$$-3A = \begin{bmatrix} -3 & -6 \\ -9 & -12 \end{bmatrix}$$

Task

For the following matrices find, where possible, $A+B,\ A-B,\ B-A,\ 2A$.

1. $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \qquad B = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$

2. $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} \qquad B = \begin{bmatrix} 1 & 1 & 1 \\ -1 & -1 & -1 \\ 1 & 1 & 1 \end{bmatrix}$

3. $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} \qquad B = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix}$

Your solution

Answer

1. $A+B = \begin{bmatrix} 2 & 3 \\ 4 & 5 \end{bmatrix} \qquad A-B = \begin{bmatrix} 0 & 1 \\ 2 & 3 \end{bmatrix} \qquad B-A = \begin{bmatrix} 0 & -1 \\ -2 & -3 \end{bmatrix} \qquad 2A = \begin{bmatrix} 2 & 4 \\ 6 & 8 \end{bmatrix}$

2. $A+B = \begin{bmatrix} 2 & 3 & 4 \\ 3 & 4 & 5 \\ 8 & 9 & 10 \end{bmatrix} \qquad A-B = \begin{bmatrix} 0 & 1 & 2 \\ 5 & 6 & 7 \\ 6 & 7 & 8 \end{bmatrix} \qquad B-A = \begin{bmatrix} 0 & -1 & -2 \\ -5 & -6 & -7 \\ -6 & -7 & -8 \end{bmatrix}$

   $2A = \begin{bmatrix} 2 & 4 & 6 \\ 8 & 10 & 12 \\ 14 & 16 & 18 \end{bmatrix}$

3. None of $A+B,\ A-B,\ B-A,$ are defined. $\qquad 2A = \begin{bmatrix} 2 & 4 & 6 \\ 8 & 10 & 12 \\ 14 & 16 & 18 \end{bmatrix}$

5. Some simple matrix properties

Using the definition of matrix addition described above we can easily verify the following properties of matrix addition:

Key Point 2

Basic Properties of Matrices

Matrix addition is **commutative**: $A + B = B + A$

Matrix addition is **associative**: $A + (B + C) = (A + B) + C$

The **distributive law** holds: $k(A + B) = k\,A + k\,B$

These Key Point results follow from the fact that $a_{ij} + b_{ij} = b_{ij} + a_{ij}$ etc.

We can also show that the transpose of a matrix satisfies the following simple properties:

Key Point 3

Properties of Transposed Matrices

$$\begin{aligned}(A+B)^T &= A^T + B^T\\ (A-B)^T &= A^T - B^T\\ (A^T)^T &= A\end{aligned}$$

Example 3

Show that $(A^T)^T = A$ for the matrix $A = \begin{bmatrix} 1 & 2 & 3\\ 4 & 5 & 6\end{bmatrix}$

Solution

$A^T = \begin{bmatrix} 1 & 4\\ 2 & 5\\ 3 & 6\end{bmatrix}$ so that $(A^T)^T = \begin{bmatrix} 1 & 2 & 3\\ 4 & 5 & 6\end{bmatrix} = A$

Task

For the matrices $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$, $B = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$ verify that

(i) $3(A+B) = 3A + 3B$ (ii) $(A-B)^T = A^T - B^T$.

Your solution

Answer

(i) $A + B = \begin{bmatrix} 2 & 1 \\ 2 & 5 \end{bmatrix}$; $3(A+B) = \begin{bmatrix} 6 & 3 \\ 6 & 15 \end{bmatrix}$; $3A = \begin{bmatrix} 3 & 6 \\ 9 & 12 \end{bmatrix}$;

$3B = \begin{bmatrix} 3 & -3 \\ -3 & 3 \end{bmatrix}$; $3A + 3B = \begin{bmatrix} 6 & 3 \\ 6 & 15 \end{bmatrix}$.

(ii) $A - B = \begin{bmatrix} 0 & 3 \\ 4 & 3 \end{bmatrix}$; $(A-B)^T = \begin{bmatrix} 0 & 4 \\ 3 & 3 \end{bmatrix}$; $A^T = \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix}$;

$B^T = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$; $A^T - B^T = \begin{bmatrix} 0 & 4 \\ 3 & 3 \end{bmatrix}$.

Exercises

1. Find the coefficient matrix A of the system:

$$\begin{aligned} 2x_1 + 3x_2 - x_3 &= 1 \\ 4x_1 + 4x_2 &= 0 \\ 2x_1 - x_2 - x_3 &= 0 \end{aligned}$$

If $B = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 0 & 0 & 1 \end{bmatrix}$ determine $(3A^T - B)^T$.

2. If $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$ and $B = \begin{bmatrix} -1 & 4 \\ 0 & 1 \\ 2 & 7 \end{bmatrix}$ verify that $3(A^T - B) = (3A - 3B^T)^T$.

Answers

1. $A = \begin{bmatrix} 2 & 3 & -1 \\ 4 & 4 & 0 \\ 2 & -1 & -1 \end{bmatrix}, \quad A^T = \begin{bmatrix} 2 & 4 & 2 \\ 3 & 4 & -1 \\ -1 & 0 & -1 \end{bmatrix}, \quad 3A^T = \begin{bmatrix} 6 & 12 & 6 \\ 9 & 12 & -3 \\ -3 & 0 & -3 \end{bmatrix}$

$3A^T - B = \begin{bmatrix} 5 & 10 & 3 \\ 5 & 7 & -9 \\ -3 & 0 & -4 \end{bmatrix} \quad (3A^T - B)^T = \begin{bmatrix} 5 & 5 & -3 \\ 10 & 7 & 0 \\ 3 & -9 & -4 \end{bmatrix}$

2. $A^T = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}, \quad A^T - B = \begin{bmatrix} 2 & 0 \\ 2 & 4 \\ 1 & -1 \end{bmatrix}, \quad 3(A^T - B) = \begin{bmatrix} 6 & 0 \\ 6 & 12 \\ 3 & -3 \end{bmatrix}$

$B^T = \begin{bmatrix} -1 & 0 & 2 \\ 4 & 1 & 7 \end{bmatrix}, \quad 3A - 3B^T = \begin{bmatrix} 3 & 6 & 9 \\ 12 & 15 & 18 \end{bmatrix} - \begin{bmatrix} -3 & 0 & 6 \\ 12 & 3 & 21 \end{bmatrix} = \begin{bmatrix} 6 & 6 & 3 \\ 0 & 12 & -3 \end{bmatrix}$

Matrix Multiplication 7.2

Introduction

When we wish to multiply matrices together we have to ensure that the operation is possible - and this is not always so. Also, unlike number arithmetic and algebra, even when the product exists the order of multiplication may have an effect on the result. In this Section we pick our way through the minefield of matrix multiplication.

Prerequisites

Before starting this Section you should ...

- understand the concept of a matrix and associated terms.

Learning Outcomes

On completion you should be able to ...

- decide when the product AB exists
- recognise that $AB \neq BA$ in most cases
- carry out the multiplication AB
- explain what is meant by the identity matrix I

1. Multiplying row matrices and column matrices together

Let A be a 1×2 row matrix and B be a 2×1 column matrix:

$$A = \begin{bmatrix} a & b \end{bmatrix} \qquad B = \begin{bmatrix} c \\ d \end{bmatrix}$$

The product of these two matrices is written AB and is the 1×1 matrix defined by:

$$AB = \begin{bmatrix} a & b \end{bmatrix} \times \begin{bmatrix} c \\ d \end{bmatrix} = [ac + bd]$$

Note that corresponding elements are multiplied together and the results are then added together. For example

$$\begin{bmatrix} 2 & -3 \end{bmatrix} \times \begin{bmatrix} 6 \\ 5 \end{bmatrix} = [12 - 15] = [-3]$$

This matrix product is easily generalised to other row and column matrices. For example if C is a 1×4 row matrix and D is a 4×1 column matrix:

$$C = \begin{bmatrix} 2 & -4 & 3 & 2 \end{bmatrix} \qquad B = \begin{bmatrix} 3 \\ 3 \\ -2 \\ 5 \end{bmatrix}$$

then we define the product of C with D as

$$CD = \begin{bmatrix} 2 & -4 & 3 & 2 \end{bmatrix} \times \begin{bmatrix} 3 \\ 3 \\ -2 \\ 5 \end{bmatrix} = [6 - 12 - 6 + 10] = [-2]$$

The only requirement is that the number of elements of the row matrix is the same as the number of elements of the column matrix.

2. Multiplying two 2×2 matrices

If A and B are two matrices then the product AB is obtained by multiplying the rows of A with the columns of B in the manner described above. This will only be possible if the number of elements in the rows of A is the same as the number of elements in the columns of B. In particular, we define the product of two 2×2 matrices A and B to be another 2×2 matrix C whose elements are calculated according to the following pattern

$$\underset{A}{\begin{bmatrix} a & b \\ c & d \end{bmatrix}} \times \underset{B}{\begin{bmatrix} w & x \\ y & z \end{bmatrix}} \underset{=}{=} \underset{C}{\begin{bmatrix} aw + by & ax + bz \\ cw + dy & cx + dz \end{bmatrix}}$$

The rule for calculating the elements of C is described in the following Key Point:

Key Point 4

Matrix Product

$$AB = C$$

The element in the i^{th} row and j^{th} column of C is obtained by multiplying the i^{th} row of A with the j^{th} column of B.

We illustrate this construction for the abstract matrices A and B given above:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \times \begin{bmatrix} w & x \\ y & z \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} a & b \end{bmatrix}\begin{bmatrix} w \\ y \end{bmatrix} & \begin{bmatrix} a & b \end{bmatrix}\begin{bmatrix} x \\ z \end{bmatrix} \\ \begin{bmatrix} c & d \end{bmatrix}\begin{bmatrix} w \\ y \end{bmatrix} & \begin{bmatrix} c & d \end{bmatrix}\begin{bmatrix} x \\ z \end{bmatrix} \end{bmatrix} = \begin{bmatrix} aw + by & ax + bz \\ cw + dy & cx + dz \end{bmatrix}$$

For example

$$\begin{bmatrix} 2 & -1 \\ 3 & -2 \end{bmatrix} \times \begin{bmatrix} 2 & 4 \\ 6 & 1 \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} 2 & -1 \end{bmatrix}\begin{bmatrix} 2 \\ 6 \end{bmatrix} & \begin{bmatrix} 2 & -1 \end{bmatrix}\begin{bmatrix} 4 \\ 1 \end{bmatrix} \\ \begin{bmatrix} 3 & -2 \end{bmatrix}\begin{bmatrix} 2 \\ 6 \end{bmatrix} & \begin{bmatrix} 3 & -2 \end{bmatrix}\begin{bmatrix} 4 \\ 1 \end{bmatrix} \end{bmatrix} = \begin{bmatrix} -2 & 7 \\ -6 & 10 \end{bmatrix}$$

Task

Find the product AB where $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ $\quad B = \begin{bmatrix} 1 & -1 \\ -2 & 1 \end{bmatrix}$

First write down row 1 of A, column 2 of B and form the first element in product AB:

Your solution

Answer

$[1, 2]$ and $\begin{bmatrix} -1 \\ 1 \end{bmatrix}$; their product is $1 \times (-1) + 2 \times 1 = 1$.

Now repeat the process for row 2 of A, column 1 of B:

Your solution

Answer

$[3, 4]$ and $\begin{bmatrix} 1 \\ -2 \end{bmatrix}$. Their product is $3 \times 1 + 4 \times (-2) = -5$

Finally find the two other elements of $C = AB$ and hence write down the matrix C:

Your solution

Answer

Row 1 column 1 is $1 \times 1 + 2 \times (-2) = -3$. Row 2 column 2 is $3 \times (-1) + 4 \times 1 = 1$

$C = \begin{bmatrix} -3 & 1 \\ -5 & 1 \end{bmatrix}$

Clearly, matrix multiplication is tricky and not at all 'natural'. However, it is a very important mathematical procedure with many engineering applications so must be mastered.

3. Some surprising results

We have already calculated the product AB where

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \text{ and } B = \begin{bmatrix} 1 & -1 \\ -2 & 1 \end{bmatrix}$$

Now complete the following task in which you are asked to determine the product BA, i.e. with the matrices in reverse order.

Task

For matrices $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & -1 \\ -2 & 1 \end{bmatrix}$ form the products of

row 1 of B and column 1 of A row 1 of B and column 2 of A

row 2 of B and column 1 of A row 2 of B and column 2 of A

Now write down the matrix BA:

Your solution

Answer

row 1, column 1 is $1 \times 1 + (-1) \times 3 = -2$ row 1, column 2 is $1 \times 2 + (-1) \times 4 = -2$

row 2, column 1 is $-2 \times 1 + 1 \times 3 = 1$ row 2, column 2 is $-2 \times 2 + 1 \times 4 = 0$

$$BA \text{ is } \begin{bmatrix} -2 & -2 \\ 1 & 0 \end{bmatrix}$$

It is clear that AB and BA are **not** in general the same. In fact it is the **exception** that $AB = BA$. In the special case in which $AB = BA$ we say that the matrices A and B **commute**.

Task

Calculate AB and BA where

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \text{ and } B = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

Your solution

Answer

$AB = BA = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$

We call B the 2×2 **zero matrix** written $\underline{0}$ so that $A \times \underline{0} = \underline{0} \times A = 0$ for any matrix A.

Now in the multiplication of numbers, the equation

$$ab = 0$$

implies that either a is zero or b is zero or both are zero. The following task shows that this is not necessarily true for matrices.

Task

Carry out the multiplication AB where

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$$

Your solution

Answer

$AB = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$

Here we have a zero product yet neither A nor B is the zero matrix! Thus the statement $AB = 0$ does **not** allow us to conclude that either $A = \underline{0}$ or $B = \underline{0}$.

Task Find the product AB where $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

Your solution

Answer
$AB = \begin{bmatrix} a & b \\ c & d \end{bmatrix} = A$

The matrix $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ is called the **identity matrix** or **unit matrix** of order 2, and is usually denoted by the symbol I. (Strictly we should write I_2, to indicate the size.) I plays the same role in matrix multiplication as the number 1 does in number multiplication.

Hence

just as $a \times 1 = 1 \times a = a$ for any number a, so $AI = IA = A$ for any matrix A.

4. Multiplying two 3×3 matrices

The definition of the product $C = AB$ where A and B are two 3×3 matrices is as follows

$$C = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \begin{bmatrix} r & s & t \\ u & v & w \\ x & y & z \end{bmatrix} = \begin{bmatrix} ar + bu + cx & as + bv + cy & at + bw + cz \\ dr + eu + fx & ds + ev + fy & dt + ew + fz \\ gr + hu + ix & gs + hv + iy & gt + hw + iz \end{bmatrix}$$

This looks a rather daunting amount of algebra but in fact the construction of the matrix on the right-hand side is straightforward if we follow the simple rule from Key Point 4 that the element in the i^{th} row and j^{th} column of C is obtained by multiplying the i^{th} row of A with the j^{th} column of B.

For example, to obtain the element in row 2, column 3 of C we take row 2 of A: $[d, e, f]$ and multiply it with column 3 of B in the usual way to produce $[dt + ew + fz]$.

By repeating this process we obtain every element of C.

Task

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

First find the element in row 2 column 1 of the product:

Your solution

Answer

Row 2 of A is $(3, 4, 0)$ column 1 of B is $\begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}$

The combination required is $3 \times 2 + 4 \times 1 + (0) \times (0) = 10$.

Now complete the multiplication to find all the elements of the matrix AB:

Your solution

Answer

In full detail, the elements of AB are:

$$\begin{bmatrix} 1\times 2 + 2\times 1 + (-1)\times 0 & 1\times(-1) + 2\times(-2) + (-1)\times 3 & 1\times 3 + 2\times 1 + (-1)\times(-2) \\ 3\times 2 + 4\times 1 + 0\times 0 & 3\times(-1) + 4\times(-2) + 0\times 3 & 3\times 3 + 4\times 1 + 0\times(-2) \\ 1\times 2 + 5\times 1 + (-2)\times 0 & 1\times(-1) + 5\times(-2) + (-2)\times 3 & 1\times 3 + 5\times 1 + (-2)\times(-2) \end{bmatrix}$$

i.e. $AB = \begin{bmatrix} 4 & -8 & 7 \\ 10 & -11 & 13 \\ 7 & -17 & 12 \end{bmatrix}$

The 3×3 unit matrix is: $I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ and as in the 2×2 case this has the property that $AI = IA = A$

The 3×3 zero matrix is $\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$

5. Multiplying non-square matrices together

So far, we have just looked at multiplying 2×2 matrices and 3×3 matrices. However, products between non-square matrices may be possible.

Key Point 5

General Matrix Products

The general rule is that an $n \times p$ matrix A can be multiplied by a $p \times m$ matrix B to form an $n \times m$ matrix $AB = C$.

In words:

For the matrix product AB to be defined the number of columns of A must equal the number of rows of B.

The elements of C are found in the usual way:

The element in the i^{th} row and j^{th} column of C is obtained by multiplying the i^{th} row of A with the j^{th} column of B.

Example 4

Find the product AB if $A = \begin{bmatrix} 1 & 2 & 2 \\ 2 & 3 & 4 \end{bmatrix}$ and $B = \begin{bmatrix} 2 & 5 \\ 6 & 1 \\ 4 & 3 \end{bmatrix}$

Solution

Since A is a 2×3 and B is a 3×2 matrix the product AB can be found and results in a 2×2 matrix.

$$AB = \begin{bmatrix} 1 & 2 & 2 \\ 2 & 3 & 4 \end{bmatrix} \times \begin{bmatrix} 2 & 5 \\ 6 & 1 \\ 4 & 3 \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} 1 & 2 & 2 \end{bmatrix}\begin{bmatrix} 2 \\ 6 \\ 4 \end{bmatrix} & \begin{bmatrix} 1 & 2 & 2 \end{bmatrix}\begin{bmatrix} 5 \\ 1 \\ 3 \end{bmatrix} \\ \begin{bmatrix} 2 & 3 & 4 \end{bmatrix}\begin{bmatrix} 2 \\ 6 \\ 4 \end{bmatrix} & \begin{bmatrix} 2 & 3 & 4 \end{bmatrix}\begin{bmatrix} 5 \\ 1 \\ 3 \end{bmatrix} \end{bmatrix} = \begin{bmatrix} 22 & 13 \\ 38 & 25 \end{bmatrix}$$

Task

Obtain the product AB if $A = \begin{bmatrix} 1 & -2 \\ 2 & -3 \end{bmatrix}$ and $B = \begin{bmatrix} 2 & 4 & 1 \\ 6 & 1 & 0 \end{bmatrix}$

Your solution

Answer

AB is a 2×3 matrix.

$$AB = \begin{bmatrix} 1 & -2 \\ 2 & -3 \end{bmatrix} \times \begin{bmatrix} 2 & 4 & 1 \\ 6 & 1 & 0 \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} 1 & -2 \end{bmatrix}\begin{bmatrix} 2 \\ 6 \end{bmatrix} & \begin{bmatrix} 1 & -2 \end{bmatrix}\begin{bmatrix} 4 \\ 1 \end{bmatrix} & \begin{bmatrix} 1 & -2 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix} \\ \begin{bmatrix} 2 & -3 \end{bmatrix}\begin{bmatrix} 2 \\ 6 \end{bmatrix} & \begin{bmatrix} 2 & -3 \end{bmatrix}\begin{bmatrix} 4 \\ 1 \end{bmatrix} & \begin{bmatrix} 2 & -3 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix} \end{bmatrix}$$

$$= \begin{bmatrix} -10 & 2 & 1 \\ -14 & 5 & 2 \end{bmatrix}$$

6. The rules of matrix multiplication

It is worth noting that the process of multiplication can be continued to form products of more than two matrices.

Although two matrices may not commute (i.e. in general $AB \neq BA$) the **associative law** always holds i.e. for matrices which **can** be multiplied,

$$A(BC) = (AB)C.$$

The general principle is **keep the left to right order**, but within that limitation any two adjacent matrices can be multiplied.

It is important to note that it is not always possible to multiply together any two given matrices.

For example if $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ and $B = \begin{bmatrix} a & b & c \\ d & e & f \end{bmatrix}$ then $AB = \begin{bmatrix} a+2d & b+2e & c+2f \\ 3a+4d & 3b+4e & 3c+4f \end{bmatrix}$.

However $BA = \begin{bmatrix} a & b & c \\ d & e & f \end{bmatrix}\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ is **not defined** since each row of B has three elements whereas each column of A has two elements and we cannot multiply these elements in the manner described.

Task

Given $A = \begin{bmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, \quad C = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}$

State which of the products $AB,\ BA,\ AC,\ CA,\ BC,\ CB,\ (AB)C,\ A(CB)$ is defined and state the size $(n \times m)$ of the product when defined.

Your solution

AB
BA
AC
CA
BC
CB
$(AB)C$
$A(CB)$

Answer

$\begin{matrix} A & B \\ 2\times 3 & 2\times 2 \end{matrix}$	not possible	$\begin{matrix} B & A \\ 2\times 2 & 2\times 3 \end{matrix}$	possible; result 2×3
$\begin{matrix} A & C \\ 2\times 3 & 3\times 2 \end{matrix}$	possible; result 2×2	$\begin{matrix} C & A \\ 3\times 2 & 2\times 3 \end{matrix}$	possible; result 3×3
$\begin{matrix} B & C \\ 2\times 2 & 3\times 2 \end{matrix}$	not possible	$\begin{matrix} C & B \\ 3\times 2 & 2\times 2 \end{matrix}$	possible; result 3×2
$(AB)C$	not possible, AB not defined.	$\begin{matrix} A & (C\ B) \\ 2\times 3 & 3\times 2 \end{matrix}$	possible; result 2×2

We now list together some properties of matrix multiplication and compare them with corresponding properties for multiplication of numbers.

Key Point 6

Matrix algebra	Number algebra
$A(B+C) = AB + AC$	$a(b+c) = ab + ac$
$AB \neq BA$ in general	$ab = ba$
$A(BC) = (AB)C$	$a(bc) = (ab)c$
$AI = IA = A$	$1.a = a.1 = a$
$A\underline{0} = \underline{0}A = \underline{0}$	$0.a = a.0 = 0$
AB may not be possible	ab is always possible
$AB = \underline{0}$ does not imply $A = \underline{0}$ or $B = \underline{0}$	$ab = 0 \ \rightarrow\ a = 0$ or $b = 0$

Application of matrices to networks

A network is a collection of points (nodes) some of which are connected together by lines (paths). The information contained in a network can be conveniently stored in the form of a matrix.

Example 5

Petrol is delivered to terminals T_1 and T_2. They distribute the fuel to 3 storage depots ($S_1,\ S_2,\ S_3$). The network diagram below shows what fraction of the fuel goes from each terminal to the three storage depots. In turn the 3 depots supply fuel to 4 petrol stations (P_1, P_2, P_3, P_4) as shown in Figure 2:

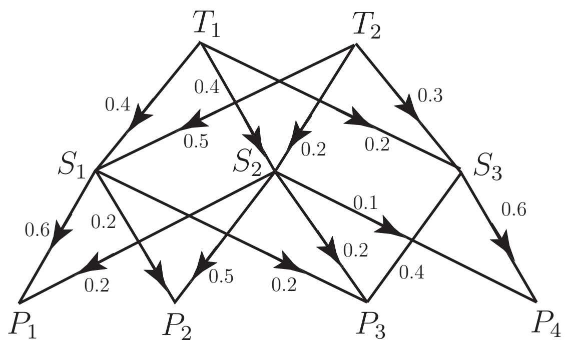
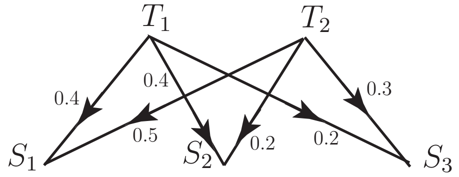


Figure 2

Show how this situation may be described using matrices.

Solution

Denote the amount of fuel, in litres, flowing from T_1 by t_1 and from T_2 by t_2 and the quantity being received at S_i by s_i for $i = 1, 2, 3$. This situation is described in the following diagram:

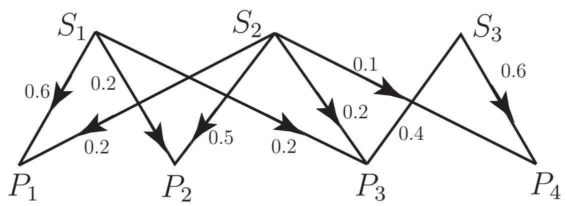


From this diagram we see that

$$\begin{aligned} s_1 &= 0.4t_1 + 0.5t_2 \\ s_2 &= 0.4t_1 + 0.2t_2 \\ s_3 &= 0.2t_1 + 0.3t_2 \end{aligned} \qquad \text{or, in matrix form:} \quad \begin{bmatrix} s_1 \\ s_2 \\ s_3 \end{bmatrix} = \begin{bmatrix} 0.4 & 0.5 \\ 0.4 & 0.2 \\ 0.2 & 0.3 \end{bmatrix} \begin{bmatrix} t_1 \\ t_2 \end{bmatrix}$$

Solution (contd.)

In turn the 3 depots supply fuel to 4 petrol stations as shown in the next diagram:



If the petrol stations receive p_1, p_2, p_3, p_4 litres respectively then from the diagram we have:

$$\begin{aligned} p_1 &= 0.6s_1 + 0.2s_2 \\ p_2 &= 0.2s_1 + 0.5s_2 \\ p_3 &= 0.2s_1 + 0.2s_2 + 0.4s_3 \\ p_4 &= 0.1s_2 + 0.6s_3 \end{aligned}$$ or, in matrix form: $$\begin{bmatrix} p_1 \\ p_2 \\ p_3 \\ p_4 \end{bmatrix} = \begin{bmatrix} 0.6 & 0.2 & 0 \\ 0.2 & 0.5 & 0 \\ 0.2 & 0.2 & 0.4 \\ 0 & 0.1 & 0.6 \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \\ s_3 \end{bmatrix}$$

Combining the equations, substituting expressions for s_1, s_2, s_3 in the equations for p_1, p_2, p_3, p_4

we get:

$$\begin{aligned} p_1 &= 0.6s_1 + 0.2s_2 \\ &= 0.6(0.4t_1 + 0.5t_2) + 0.2(0.4t_1 + 0.2t_1) \\ &= 0.32t_1 + 0.34t_2 \end{aligned}$$

with similar results for p_2, p_3 and p_4.

This is equivalent to combining the two networks. The results can be obtained more easily by multiplying the matrices:

$$\begin{aligned} \begin{bmatrix} p_1 \\ p_2 \\ p_3 \\ p_4 \end{bmatrix} &= \begin{bmatrix} 0.6 & 0.2 & 0 \\ 0.2 & 0.5 & 0 \\ 0.2 & 0.2 & 0.4 \\ 0 & 0.1 & 0.6 \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \\ s_3 \end{bmatrix} \\ &= \begin{bmatrix} 0.6 & 0.2 & 0 \\ 0.2 & 0.5 & 0 \\ 0.2 & 0.2 & 0.4 \\ 0 & 0.1 & 0.6 \end{bmatrix} \begin{bmatrix} 0.4 & 0.5 \\ 0.4 & 0.2 \\ 0.2 & 0.3 \end{bmatrix} \begin{bmatrix} t_1 \\ t_2 \end{bmatrix} \\ &= \begin{bmatrix} 0.32 & 0.34 \\ 0.28 & 0.20 \\ 0.24 & 0.26 \\ 0.16 & 0.20 \end{bmatrix} \begin{bmatrix} t_1 \\ t_2 \end{bmatrix} = \begin{bmatrix} 0.32t_1 + 0.34t_2 \\ 0.28t_1 + 0.20t_2 \\ 0.24t_1 + 0.26t_2 \\ 0.16t_1 + 0.20t_2 \end{bmatrix} \end{aligned}$$

Engineering Example 1

Communication network

Problem in words

Figure 3 represents a communication network. Vertices a, b, f and g represent offices. Vertices c, d and e represent switching centres. The numbers marked along the edges represent the number of connections between any two vertices. Calculate the number of routes from a and b to f and g

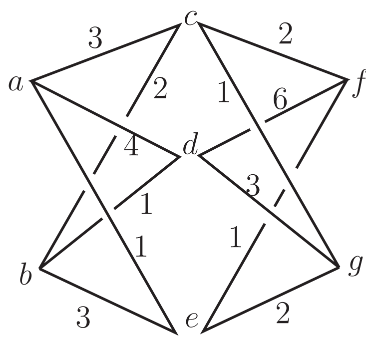


Figure 3: A communication network where a, b, f and g are offices and c, d and e are switching centres

Mathematical statement of the problem

The number of routes from a to f can be calculated by taking the number via c plus the number via d plus the number via e. In each case this is given by multiplying the number of connections along the edges connecting a to c, c to f etc. This gives the result:
Number of routes from a to $f = 3 \times 2 + 4 \times 6 + 1 \times 1 = 31$.

The nature of matrix multiplication means that the number of routes is obtained by multiplying the matrix representing the number of connections from ab to cde by the matrix representing the number of connections from cde to fg.

Mathematical analysis

The matrix representing the number of routes from ab to cde is:

$$\begin{array}{cc} & \begin{array}{ccc} c & d & e \end{array} \\ \begin{array}{c} a \\ b \end{array} & \begin{pmatrix} 3 & 4 & 1 \\ 2 & 1 & 3 \end{pmatrix} \end{array}$$

The matrix representing the number of routes from cde to fg is:

$$\begin{array}{cc} & \begin{array}{cc} f & g \end{array} \\ \begin{array}{c} c \\ d \\ e \end{array} & \begin{pmatrix} 2 & 1 \\ 6 & 3 \\ 1 & 2 \end{pmatrix} \end{array}$$

The product of these two matrices gives the total number of routes.

$$\begin{pmatrix} 3 & 4 & 1 \\ 2 & 1 & 3 \end{pmatrix}\begin{pmatrix} 2 & 1 \\ 6 & 3 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} 3\times 2+4\times 6+1\times 1 & 3\times 1+4\times 3+1\times 2 \\ 2\times 2+1\times 6+3\times 1 & 2\times 1+1\times 3+3\times 2 \end{pmatrix} = \begin{pmatrix} 31 & 17 \\ 13 & 11 \end{pmatrix}$$

Interpretation

We can interpret the resulting (product) matrix by labelling the columns and rows.

$$\begin{array}{cc} & \begin{array}{cc} f & g \end{array} \\ \begin{array}{c} a \\ b \end{array} & \begin{pmatrix} 31 & 17 \\ 13 & 11 \end{pmatrix} \end{array}$$

Hence there are 31 routes from a to f, 17 from a to g, 13 from b to f and 11 from b to g.

Exercises

1. If $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ $\quad B = \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix}$ $\quad C = \begin{bmatrix} 0 & -1 \\ 2 & -3 \end{bmatrix}$ find

   (a) AB, (b) AC, (c) $(A+B)C$, (d) $AC+BC$ (e) $2A-3C$

2. If a rotation through an angle θ is represented by the matrix $A = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}$ and a second rotation through an angle ϕ is represented by the matrix $B = \begin{bmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{bmatrix}$ show that both AB and BA represent a rotation through an angle $\theta+\phi$.

3. If $A = \begin{bmatrix} 1 & 2 & 3 \\ -1 & -1 & -1 \\ 2 & 2 & 2 \end{bmatrix}$, $B = \begin{bmatrix} 2 & 4 \\ -1 & 2 \\ 5 & 6 \end{bmatrix}$, $C = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$, find AB and BC.

4. If $A = \begin{bmatrix} 1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix}$, $\quad B = \begin{bmatrix} 1 & 2 & 3 \\ 5 & 0 & 0 \\ 1 & 2 & -1 \end{bmatrix}$, $\quad C = \begin{bmatrix} 0 \\ 1 \\ -2 \end{bmatrix}$,

   verify $A(BC) = (AB)C$.

5. If $A = \begin{bmatrix} 2 & 3 & -1 \\ 0 & 1 & 2 \\ 4 & 5 & 6 \end{bmatrix}$ then show that AA^T is symmetric.

6. If $A = \begin{bmatrix} 11 & 0 \\ 2 & 1 \end{bmatrix}$ $\quad B = \begin{bmatrix} 0 & 1 & 2 \\ 1 & 1 & 3 \end{bmatrix}$ verify that $(AB)^T = \begin{bmatrix} 0 & 1 \\ 11 & 3 \\ 22 & 7 \end{bmatrix} = B^TA^T$

Answers

1. (a) $AB = \begin{bmatrix} 19 & 22 \\ 43 & 50 \end{bmatrix}$ (b) $AC = \begin{bmatrix} 4 & -7 \\ 8 & -15 \end{bmatrix}$ (c) $(A+B)C = \begin{bmatrix} 16 & -30 \\ 24 & -46 \end{bmatrix}$

   (d) $AC+BC = \begin{bmatrix} 16 & -30 \\ 24 & -46 \end{bmatrix}$ (e) $\begin{bmatrix} 2 & 7 \\ 0 & 17 \end{bmatrix}$

2. $$AB = \begin{bmatrix} \cos\theta\cos\phi - \sin\theta\sin\phi & \cos\theta\sin\phi + \sin\theta\cos\phi \\ -\sin\theta\cos\phi - \cos\theta\sin\phi & -\sin\theta\sin\phi + \cos\theta\cos\phi \end{bmatrix}$$

   $$= \begin{bmatrix} \cos(\theta+\phi) & \sin(\theta+\phi) \\ -\sin(\theta+\phi) & \cos(\theta+\phi) \end{bmatrix}$$

   which clearly represents a rotation through angle $\theta+\phi$. BA gives the same result.

3. $AB = \begin{bmatrix} 15 & 26 \\ -6 & -12 \\ 12 & 24 \end{bmatrix}$, $\quad BC = \begin{bmatrix} 8 & 10 \\ 0 & 3 \\ 16 & 17 \end{bmatrix}$

4. $A(BC) = (AB)C = \begin{bmatrix} -8 \\ 8 \end{bmatrix}$

Exercises

1. If $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ $B = \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix}$ $C = \begin{bmatrix} 0 & -1 \\ 2 & -3 \end{bmatrix}$ find

(a) AB, (b) AC, (c) $(A+B)C$, (d) $AC+BC$ (e) $2A-3C$

(a) $\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}\begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix}$ $5+14$ $6+16$ $\begin{bmatrix} 19 & 22 \\ 43 & 50 \end{bmatrix}$ $15+28$ $18+32$

(b) $\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}\begin{bmatrix} 0 & -1 \\ 2 & -3 \end{bmatrix}$ $0+4$ $-1-6$ $\begin{bmatrix} 4 & -7 \\ 8 & -15 \end{bmatrix}$ $0+8$ $-3-12$

(c) $\begin{bmatrix} 6 & 8 \\ 10 & 12 \end{bmatrix}\begin{bmatrix} 0 & -1 \\ 2 & -3 \end{bmatrix}$ $0+16$ $-6-24$ $\begin{bmatrix} 16 & -30 \\ 24 & -46 \end{bmatrix}$ $0+24$ $-10-36$

(d) $\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}\begin{bmatrix} 0 & -1 \\ 2 & 3 \end{bmatrix} + \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix}\begin{bmatrix} 0 & -1 \\ 2 & -3 \end{bmatrix}$ $0+12$ $-5-18$

$\begin{bmatrix} 4 & -7 \\ 8 & -15 \end{bmatrix} + \begin{bmatrix} 12 & -23 \\ 16 & -31 \end{bmatrix}$ $0+16$ $-7-24$

$\begin{bmatrix} 16 & -30 \\ 24 & -46 \end{bmatrix}$

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

$\begin{bmatrix} 2 & 4 \\ 6 & 8 \end{bmatrix} + \begin{bmatrix} 0 & +3 \\ -6 & +9 \end{bmatrix}$

$\begin{bmatrix} 2 & 7 \\ 0 & 17 \end{bmatrix}$

2. If a rotation through an angle θ is represented by the matrix $A = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}$ and a second rotation through an angle ϕ is represented by the matrix $B = \begin{bmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{bmatrix}$ show that both AB and BA represent a rotation through an angle $\theta + \phi$.

Sine and Cosine Addition and Subtraction Formulas

$\sin(a+b) = \sin a\cos b + \cos a \sin b$

$\sin(a-b) = \sin a\cos b - \cos a \sin b$

$\cos(a+b) = \cos a\cos b - \sin a \sin b$

$\cos(a-b) = \cos a\cos b + \sin a \sin b$

AB

$\begin{bmatrix} \cos\theta\cos\phi - \sin\theta\sin\phi & \cos\theta\sin\phi + \sin\theta\cos\phi \\ -\sin\theta\cos\phi - \cos\theta\sin\phi & -\sin\theta\sin\phi + \cos\theta\cos\phi \end{bmatrix}$

$\begin{bmatrix} \cos(\theta+\phi) & \sin(\theta+\phi) \\ -\sin(\theta+\phi) & \cos(\theta+\phi) \end{bmatrix}$

BA

$\begin{bmatrix} \cos\phi\cos\theta - \sin\phi\sin\theta & \cos\phi\sin\theta + \sin\phi\cos\theta \\ -\sin\phi\cos\theta - \cos\phi\sin\theta & -\sin\phi\sin\theta + \cos\phi\cos\theta \end{bmatrix}$

$\begin{bmatrix} \cos(\phi+\theta) & \sin(\phi+\theta) \\ -\sin(\phi+\theta) & \cos(\phi+\theta) \end{bmatrix}$

3. If $A = \begin{bmatrix} 1 & 2 & 3 \\ -1 & -1 & -1 \\ 2 & 2 & 2 \end{bmatrix}$, $B = \begin{bmatrix} 2 & 4 \\ -1 & 2 \\ 5 & 6 \end{bmatrix}$, $C = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$, find AB and BC.

AB

$2-2+15$ $4+4+18$

-2 +1 -5 | -4 -2 -6

$\begin{bmatrix} 15 & 26 \\ -6 & -12 \\ 12 & 24 \end{bmatrix}$

+4 -2 +10 | +8 +4 +12

BC

$4+4$ $2+8$

-2 +2 | -1 +4

$\begin{bmatrix} 8 & 10 \\ 0 & 3 \\ 16 & 17 \end{bmatrix}$

+10 +6 | +5 +12

6. If $A = \begin{bmatrix} 11 & 0 \\ 2 & 1 \end{bmatrix}$ $B = \begin{bmatrix} 0 & 1 & 2 \\ 1 & 1 & 3 \end{bmatrix}$ verify that $(AB)^T = \begin{bmatrix} 0 & 1 \\ 11 & 3 \\ 22 & 7 \end{bmatrix} = B^T A^T$

AB

$0+0$ $11+0$ $22+0$

$\begin{bmatrix} 0 & 11 & 22 \\ 1 & 3 & 7 \end{bmatrix}$

$0+1$ $2+1$ $4+3$

$(AB)^T$

$\begin{bmatrix} 0 & 1 \\ 11 & 3 \\ 22 & 7 \end{bmatrix}$

B^T A^T

$\begin{bmatrix} 0 & 1 \\ 1 & 1 \\ 2 & 3 \end{bmatrix}$ $\begin{bmatrix} 11 & 2 \\ 0 & 1 \end{bmatrix}$

$0+6$ $0+1$

+11 +0 | +2 +1

$\begin{bmatrix} 0 & 1 \\ 11 & 3 \\ 22 & 7 \end{bmatrix}$

+22 +0 | +4 +3

$2 \to \begin{pmatrix} \cos(-\phi) & -\sin(-\phi) \\ \sin(\phi) & \cos(\phi) \end{pmatrix}$

odd func.

$f(-x) = -f(x)$

remember the properties of sin & cos

Homework-2

Based upon Chapter 7 of the HELM course “Matrices”:

1. **Skim over Sections 7.1** “*Introduction to matrices*” and 7.2 “*Matrix multiplication*”. Make sure that you understand the contents and can solve / reproduce all examples. This part must be sufficiently simple to you. However, let me know in the class if you encountered any difficulties. We will discuss your problems in the class. We did not discuss the following topics: *trace*, *symmetric matrices*. You may skip them when reading the course.
2. **On p. 29, solve the problems 1-3, 6**.
   In **Problem 2**, the rotation matrix is defined differently from what we had in the class. What is the difference between these two rotation matrices? Solving Problem 2 requires some knowledge of trigonometric transformations. If it is difficult to you, try to solve the problem geometrically.
3. **Skip Section 7.3** “Determinants”.
4. **Read pp. 38–43 of Section 7.4** “Inverse matrices”. Do not worry if you cannot follow this part completely. We will return to inverse matrices and discuss them in more detail later on.

The following problems may require some *mathematical* imagination. Do not be afraid if cannot solve some of them.

1. Considering all that we know about rotation matrix, answer the following:
   Given the rotation matrix $\begin{bmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{bmatrix}$, determine its inverse. Explain your solution.
2. Simplify the following expressions using the properties of transposed and inverse matrices:
   $a.\ \ A^T(B^T - A^{-1})^T B^{-1}, \quad b.\ \ B^{-1}(B^T B^{-1} - I)^{-1} B^{-1}.$
3. Find a general form of the matrix that commutes with $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$. A matrix B says to commute with A if AB=BA.
   Hint: write the matrix B in a general form as, e.g., $B = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. Calculate both sides of the equality AB=BA and equate the entries of the resulting matrices component-wise. Solve the resulting equations to determine the values of **a, b, c**, and **d**.

1. Considering all that we know about rotation matrix, answer the following:

Given the rotation matrix $A=\begin{bmatrix}\cos\phi & -\sin\phi \\ \sin\phi & \cos\phi\end{bmatrix}$, determine its inverse. Explain your solution.

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad-bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

Pythagorean Identities
$\sin^2\theta+\cos^2\theta=1$
$1+\tan^2\theta=\sec^2\theta$
$1+\cot^2\theta=\csc^2\theta$

$$\frac{1}{\cos^2\theta+\sin^2\theta}\begin{bmatrix}\cos\theta & \sin\theta \\ -\sin\theta & \cos\theta\end{bmatrix}$$

$$A^{-1}=\begin{bmatrix}\cos\phi & \sin\phi \\ -\sin\phi & \cos\phi\end{bmatrix}$$

2. Simplify the following expressions using the properties of transposed and inverse matrices:

$a.\ A^T(B^T-A^{-1})^T B^{-1}$, $\quad b.\ B^{-1}(B^T B^{-1}-I)^{-1}B^{-1}$.

$$A^T(B-A^{-1^T})B^{-1}$$

$$(A^T B - A^T A^{-1^T})B^{-1}$$

$$A^T B B^{-1} - (AA^{-1})^T B^{-1}$$

$$A^T I - (I)^T B^{-1}$$

$$A^T - B^{-1}$$

$$B^{-1}((B^T B^{-1})^{-1} - I)B^{-1}$$

$$(B^{-1} B B^{T^{-1}} - B^{-1} I)B^{-1}$$

$$I B^{T^{-1}} - B^{-1} I B^{-1}$$

$$(BB^T - BB)^{-1} \quad B^{T^{-1}} B^{-1} - B^{-1} B^{-1}$$

$$(BB^T)^{-1} - (BB)^{-1}$$

3. Find a general form of the matrix that commutes with $A=\begin{bmatrix}0 & -1 \\ 1 & 0\end{bmatrix}$. A matrix B says to commute with A if AB=BA.

Hint: write the matrix B in a general form as, e.g., $B=\begin{bmatrix}a & b \\ c & d\end{bmatrix}$. Calculate both sides of the equality AB=BA and equate the entries of the resulting matrices component-wise. Solve the resulting equations to determine the values of **a, b, c**, and **d**.

$$AB = BA$$

$$\begin{bmatrix}0 & -1 \\ 1 & 0\end{bmatrix}\begin{bmatrix}a & b \\ c & d\end{bmatrix} = \begin{bmatrix}a & b \\ c & d\end{bmatrix}\begin{bmatrix}0 & -1 \\ 1 & 0\end{bmatrix}$$

$$\begin{bmatrix}-c & -d \\ a & b\end{bmatrix} = \begin{bmatrix}b & -a \\ d & -c\end{bmatrix}$$

$-c = b$

$-d = -a$

$a = d$

$b = -c$

Week 3 Notes

① Linear combination of vectors

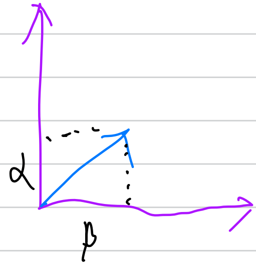
Given n vectors $v_1, v_2, \dots, v_n \in \mathbb{R}^m$ their l.c. is defined as $\alpha_1 v_1 + \alpha_2 v_2 + \dots \alpha_3 v_3$, where $\alpha_1, \dots, \alpha_n \in \mathbb{R}$

② Linear independance of vectors. Vectors $v_1, \dots, v_n$ are said to be linear independant. If their linear comb = 0 only if $\alpha_1, \dots, \alpha_n = 0$ otherwise vectors are linear dependent

Linear spaces

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix} \in \mathbb{R}^m$$

Vectors



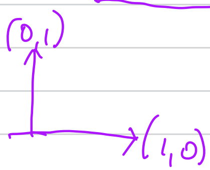
$$0 = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} \in \mathbb{R}^m$$

Ex

$$\alpha_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \alpha_2 \begin{bmatrix} 0 \\ 1 \end{bmatrix} + \alpha_3 \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

1 1 −1

linear dependent



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

linear independent

Linear dependant means you can express one vector by other vectors

Linear independant is minimal amount of vectors

$$v_1 = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, \quad v_2 = \begin{bmatrix} -1 \\ 0 \\ -1 \end{bmatrix}, \quad v_3 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

linear independent

$$\alpha_1 \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} + \alpha_2 \begin{bmatrix} -1 \\ 0 \\ -1 \end{bmatrix} + \alpha_3 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\begin{cases} \alpha_1 - 1\alpha_2 + \alpha_3 = 0 \\ 2\alpha_1 + 0\alpha_2 + \alpha_3 = 0 \\ 0\alpha_1 - 1\alpha_2 + \alpha_3 = 0 \end{cases}$$

$$\left[\begin{array}{ccc|c} 1 & -1 & 1 & 0 \\ 2 & 0 & 1 & 0 \\ 0 & -1 & 1 & 0 \end{array}\right] \xrightarrow{-2R} \left[\begin{array}{ccc|c} 1 & -1 & 1 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & -2 & -1 & 0 \end{array}\right] \xrightarrow{2R2} \left[\begin{array}{ccc|c} 1 & -1 & 1 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & -3 & 0 \end{array}\right]$$

$\alpha_1 = 0$
$\alpha_2 = 0$
$\alpha_3 = 0$

$v_1 = \begin{bmatrix} 0 \\ -1 \\ 2 \end{bmatrix}, v_2 = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}, v_3 = \begin{bmatrix} -2 \\ -5 \\ 7 \end{bmatrix}, v_4 = \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}$

linear independancy?

linear dependant

$$\left[\begin{array}{cccc|c} 0 & 1 & -2 & 2 & 0 \\ -1 & 2 & -5 & 0 & 0 \\ 2 & -1 & 7 & 1 & 0 \end{array}\right]$$

$$\left[\begin{array}{cccc|c} 1 & -2 & 5 & 0 & 0 \\ 0 & 1 & -2 & 2 & 0 \\ 2 & -1 & 7 & 1 & 0 \end{array}\right] \xrightarrow{-2R1} \left[\begin{array}{cccc|c} 1 & -2 & 5 & 0 & 0 \\ 0 & 1 & -2 & 2 & 0 \\ 0 & 3 & -3 & 1 & 0 \end{array}\right] \xrightarrow{-3R2} \left[\begin{array}{cccc|c} 1 & -2 & 5 & 0 & 0 \\ 0 & 1 & -2 & 2 & 0 \\ 0 & 0 & 3 & -5 & 0 \end{array}\right]$$

In an n-dim. space ($\mathbb{R}^n$) there can be at most n linear dependant vectors

Leading elements to show how many linear independant vectors

$v_1 = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix}$ $v_2 = \begin{bmatrix} -2 \\ -4 \\ 2 \end{bmatrix}$ $v_3 = \begin{bmatrix} 2 \\ 1 \\ -1 \end{bmatrix}$ $v_4 = \begin{bmatrix} 0 \\ -3 \\ 1 \end{bmatrix}$ ← linear dependant

$$\left[\begin{array}{cccc|c} 1 & -2 & 2 & 0 & 0 \\ 2 & -4 & 1 & -3 & 0 \\ -1 & 2 & -1 & 1 & 0 \end{array}\right] \xrightarrow{R3+R1} \left[\begin{array}{cccc|c} 1 & -2 & 2 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 2 & -4 & 1 & -3 & 0 \end{array}\right]$$

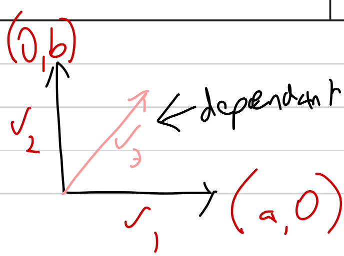
$-2R1$

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

leading depend.

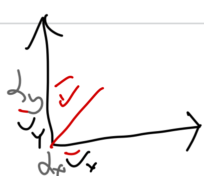
linear dependant

v_2 & v_4 can be expressed by v_1 & v_3



$v_3 = \alpha_1 v_1 + \alpha_2 v_2$

We express v_3 in terms of v_1 & v_2



$$\alpha_x \bar{V}_x + \alpha_y \bar{V}_y = \bar{V}$$

R^2 are linear independant

$v_1 = \begin{bmatrix}1\\2\end{bmatrix}$, $v_2 = \begin{bmatrix}2\\-1\end{bmatrix}$

$v_3 = \begin{bmatrix}1\\-1\end{bmatrix}$

$$\left[\begin{array}{cc|c}1 & 2 & 1\\ 2 & -1 & -1\end{array}\right]$$

$x + 2 = 1$
$x = 3$

$$\left[\begin{array}{cc|c}1 & 2 & 1\\ 0 & -5 & -3\end{array}\right]$$

$x = -1/5$
$y = 3/8$

$v_1 = \begin{pmatrix}1\\2\\0\end{pmatrix}$, $v_2 = \begin{pmatrix}-1\\0\\-1\end{pmatrix}$, $v_3 = \begin{pmatrix}1\\1\\1\end{pmatrix}$ $w = \begin{pmatrix}-1\\0\\-2\end{pmatrix}$

$$\left[\begin{array}{ccc|c}1 & -1 & 1 & -1\\ 2 & 0 & 1 & 0\\ 0 & -1 & 1 & -2\end{array}\right] \Rightarrow \left[\begin{array}{ccc|c}1 & -1 & 1 & -1\\ 0 & 1 & -1 & 2\\ 2 & 0 & 1 & 0\end{array}\right] \xRightarrow{-2R1} \left[\begin{array}{ccc|c}1 & -1 & 1 & -1\\ 0 & 1 & -1 & 2\\ 0 & 2 & -1 & 2\end{array}\right]$$

$\xrightarrow{-2R2}$

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

$x = 1$
$y = 0$
$z = -2$

$$\begin{pmatrix}1\\2\\0\end{pmatrix} - \begin{pmatrix}2\\2\\2\end{pmatrix} = \begin{pmatrix}-1\\0\\-2\end{pmatrix} \checkmark$$

Homework 3: Linear spaces

Some theoretical background

The space of n-dimensional real vectors is denoted by $\mathbb{R}^n$. An element of $\mathbb{R}^n$ writes as

$$\mathbf{v} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix},$$

where $x_1, x_2, \ldots, x_n$ are real numbers. In the following, we will use straight boldface letters (e.g., $\mathbf{v}$) to denote vectors, and regular italic ones (e.g., x) to denote scalars. Particularly, $\mathbf{0}$ is a vector of zeros. The dimension of this vector must be clear from the context.

Definition (Linear combination). Given k vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k \in \mathbb{R}^n$, their linear combination is defined as

$$\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \ldots \alpha_k\mathbf{v}_k,$$

where $\alpha_1, \alpha_2, \ldots, \alpha_k$ are real numbers (scalars).

Attention: linear combination of any number of vectors from $\mathbb{R}^n$ always belongs to $\mathbb{R}^n$.

Definition (Linear independence). Given k vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k \in \mathbb{R}^n$, they said to be linearly independent if the equality

$$\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \ldots \alpha_k\mathbf{v}_k = \mathbf{0} \qquad (*)$$

holds **only** for $\alpha_1 = \alpha_2 = \cdots = \alpha_k = 0$. Otherwise, i.e., if $(*)$ holds for **at least one** non-zero α, the vectors are called linearly dependent.

We have several properties of linear (in)dependence:

1. The set of vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k \in \mathbb{R}^n$ is linearly dependent if some vector $\mathbf{v}_i$ is equal to $\mathbf{0}$.
2. The set of vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k \in \mathbb{R}^n$ is linearly dependent if any two vectors $\mathbf{v}_i$, $\mathbf{v}_j$ are equal.
3. If the set of vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k \in \mathbb{R}^n$ is linearly dependent, there is always a vector that can be expressed as a linear combination of the remaining ones.
4. The space $\mathbb{R}^n$ contains **at most** n linearly independent vectors.

Try to explain the first two properties. For simplicity, consider only 3 vectors. Assume that one vector, say, the first one, is zero. Can we find α's, **not all equal to** 0 such that the linear combination is $\mathbf{0}$? Try to think about the second property in the same way.

$\{v_1, v_2, v_3, v_1 + v_2\}$

$(-1, -1, 0, 1)$

How to check linear independency?

1. Write the equation $(*)$ as a system of linear algebraic equations.

2. Perform Gaussian elimination to reduce the system to a triangular form.

3. The number of non-zero rows corresponds to the number of **linearly independent** vectors. If the number of linearly independent vectors equals the total number of vectors, then all vectors are linearly independent.

4. If the number of linearly independent vectors is less than the total number of vectors, we can identify the linearly independent ones as follows:

   Linearly independent vectors correspond to the leading elements in non-zero rows.

   For instance, after performing the Gaussian elimination, the system of equations has turned into

$$\begin{cases} \mathbf{2x_1} & +3x_2 & & -x_4 & = 0 \\ & & -\mathbf{x_3} & +x_4 & = 0 \\ & & & \mathbf{x_4} & = 0. \end{cases}$$

   Here, the leading elements are put in bold. Respectively, the independent vectors are the 1st, the 3rd, and the 4th.

Definition (Basis of $\mathbb{R}^n$) We say that any set of n linearly independent vectors in $\mathbb{R}^n$ forms a basis of $\mathbb{R}^n$.

Attention: there can be infinitely many bases in $\mathbb{R}^n$. Most often, we work with the Cartesian basis. For instance, the Cartesian basis of $\mathbb{R}^3$ is

$$C_3 = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}.$$

Expanding a vector with respect to a basis. The following theorem provides a theoretical foundation for the subsequent analysis.

Theorem. Given a basis of $\mathbb{R}^n$: $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$. Any vector $\mathbf{x} \in \mathbb{R}^n$ can be uniquely reperesented as a linear combination of the basis vectors:

$$\mathbf{x} = \alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \ldots \alpha_n\mathbf{v}_n.$$

The numbers $\alpha_1, \ldots, \alpha_n$ are called the coordinates of $\mathbf{x}$ in the basis $\mathcal{B}$.

Consider the situation, when we wish to express a vector in terms of others. This can happen in two cases:

1. Given a basis of $\mathbb{R}^n$, we wish to express some vector $\mathbf{x}$ as a linear combination of the basis vectors.

2. Given a set of linearly dependent vectors. We identified the subset of linearly independent vectors and want to express other vectors as linear combinations of independent vectors.

In either case, the procedure is the same: given a set of linearly independent vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k$, and a vector $\mathbf{x}$, we need to solve the vector equation

$$\alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2 + \ldots \alpha_n \mathbf{v}_n = \mathbf{x}.$$

This equation can be written as a system of linear algebraic equations (now with non-zero free terms) and solved using the standard Gaussian elimination.

Problems for the homework

1. Consider the following sets of vectors. Check if the vectors are linearly independent. **If no**, identify the linearly independent vectors and express the remaining vectors in terms of them.

$$a) \left\{ \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} \right\}, \; b) \left\{ \begin{bmatrix} 2 \\ -1 \\ 0 \\ 2 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ -2 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 3 \\ -1 \\ 2 \end{bmatrix} \right\}, \; c) \left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -2 \\ 2 \\ 1 \end{bmatrix} \right\}$$

$$d) \left\{ \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \end{bmatrix} \right\}, \; e) \left\{ \begin{bmatrix} 2 \\ -1 \\ 2 \\ -1 \end{bmatrix}, \begin{bmatrix} -2 \\ 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ -2 \\ 5 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \right\}$$

a) $\left\{ \begin{bmatrix}1\\-1\\1\end{bmatrix}, \begin{bmatrix}-1\\1\\1\end{bmatrix}, \begin{bmatrix}1\\1\\-1\end{bmatrix} \right\}$,

$$\left[\begin{array}{ccc|c}1&-1&1&0\\-1&1&1&0\\1&1&-1&0\end{array}\right] \xrightarrow[-R1]{+R1} \left[\begin{array}{ccc|c}1&-1&1&0\\0&0&2&0\\0&2&-2&0\end{array}\right]$$

vectors are linear independant

b) $\left\{ \begin{bmatrix}2\\-1\\0\\2\end{bmatrix}, \begin{bmatrix}-1\\1\\-2\\0\end{bmatrix}, \begin{bmatrix}0\\3\\-1\\2\end{bmatrix} \right\}$

$$\left[\begin{array}{ccc|c}2&-1&0&0\\-1&1&3&0\\0&-2&-1&0\\2&0&2&0\end{array}\right] \rightarrow \left[\begin{array}{ccc|c}1&-1&-3&0\\2&-1&0&0\\0&2&1&0\\2&0&2&0\end{array}\right] \xrightarrow{-2R1} \left[\begin{array}{ccc|c}1&-1&-3&0\\0&1&6&0\\0&2&1&0\\0&2&8&0\end{array}\right] \xrightarrow{-2R2} \left[\begin{array}{ccc|c}1&-1&-3&0\\0&1&6&0\\0&0&-11&0\\0&0&-4&0\end{array}\right]$$

vectors are linear independand

c) $\left\{ \begin{bmatrix}1\\0\\1\end{bmatrix}, \begin{bmatrix}1\\1\\-1\end{bmatrix}, \begin{bmatrix}2\\1\\0\end{bmatrix}, \begin{bmatrix}-2\\2\\1\end{bmatrix} \right\}$

$$\left[\begin{array}{cccc|c}1&1&2&-2&0\\0&1&1&2&0\\1&-1&0&1&0\end{array}\right] \xrightarrow{-1R1} \left[\begin{array}{cccc|c}1&1&2&-2&0\\0&1&1&2&0\\0&-2&-2&3&0\end{array}\right] \xrightarrow{+2R2} \left[\begin{array}{cccc|c}1&1&2&-2&0\\0&1&1&2&0\\0&0&0&7&0\end{array}\right]$$

1st, 2nd & 4th vector independant

3rd vector composed of 1st & 2nd

$$\left[\begin{array}{ccc|c}1&1&-2&2\\0&1&2&1\\1&-1&1&6\end{array}\right] \xrightarrow{-R1} \left[\begin{array}{ccc|c}1&1&-2&2\\0&1&2&1\\0&-2&3&-2\end{array}\right] \xrightarrow{+2R2} \left[\begin{array}{ccc|c}1&1&-2&2\\0&1&2&1\\0&0&7&0\end{array}\right]$$

$x=1$
$y=1$
$z=0$

d) $\left\{ \begin{bmatrix}1\\-1\end{bmatrix}, \begin{bmatrix}0\\1\end{bmatrix}, \begin{bmatrix}2\\1\end{bmatrix}, \begin{bmatrix}1\\2\end{bmatrix} \right\}$

$$\left[\begin{array}{cccc|c}1&0&2&1&0\\-1&1&1&2&0\end{array}\right] \xrightarrow{+1R1} \left[\begin{array}{cccc|c}1&0&2&1&0\\0&1&3&3&0\end{array}\right]$$

1st & 2nd vector are lin. indep.

$$\left[\begin{array}{cc|c}1&0&2\\-1&1&1\end{array}\right] \xrightarrow{+1R1} \left[\begin{array}{cc|c}1&0&2\\0&1&3\end{array}\right]$$

$2\vec{v}_1 + 3\vec{v}_2 = \vec{v}_3$

$$\left[\begin{array}{cc|c}1&0&1\\-1&1&2\end{array}\right] \xrightarrow{+1R1} \left[\begin{array}{cc|c}1&0&1\\0&1&3\end{array}\right]$$

$1\vec{v}_1 + 3\vec{v}_2 = \vec{v}_4$

e) $\left\{ \begin{bmatrix}2\\-1\\2\\-1\end{bmatrix}, \begin{bmatrix}-2\\0\\1\\1\end{bmatrix}, \begin{bmatrix}2\\-2\\5\\-1\end{bmatrix}, \begin{bmatrix}1\\1\\1\\1\end{bmatrix} \right\}$

$$\left[\begin{array}{cccc|c}1&-1&1&\tfrac{1}{2}&0\\-1&0&-2&1&0\\2&1&5&1&0\\-1&1&-1&1&0\end{array}\right]$$

$+1R1 \downarrow \quad -2R1 \downarrow \quad +1R1 \downarrow$

$$\left[\begin{array}{cccc|c}1&-1&1&\tfrac{1}{2}&0\\0&-1&-1&\tfrac{3}{2}&0\\0&3&3&0&0\\0&0&0&\tfrac{3}{2}&0\end{array}\right] \xrightarrow{3R2} \left[\begin{array}{cccc|c}1&-1&1&\tfrac{1}{2}&0\\0&-1&-1&\tfrac{3}{2}&0\\0&3&3&0&0\\0&0&0&\tfrac{3}{2}&0\end{array}\right] \xrightarrow{+R2} \left[\begin{array}{cccc|c}1&-1&1&\tfrac{1}{2}&0\\0&-1&-1&\tfrac{3}{2}&0\\0&0&0&\tfrac{9}{2}&0\\0&0&0&\tfrac{3}{2}&0\end{array}\right]$$

1st, 2nd, 4th are linear ind.

$\vec{v}_3 = 2\vec{v}_1 + \vec{v}_2$

$$\left[\begin{array}{ccc|c}2&-2&1&2\\-1&0&1&-2\\2&1&1&5\\-1&1&1&-1\end{array}\right] \quad \left[\begin{array}{ccc|c}1&0&-1&2\\2&-2&1&2\\2&1&1&5\\0&1&0&1\end{array}\right] \rightarrow \left[\begin{array}{ccc|c}1&0&-1&2\\0&-2&3&2\\0&1&3&1\\0&1&0&1\end{array}\right] \quad \left[\begin{array}{ccc|c}1&0&-1&2\\0&1&0&1\\0&0&3&0\\0&0&0&0\end{array}\right]$$

$c)$ $\left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -2 \\ 2 \\ 1 \end{bmatrix} \right\}$

$$\left[\begin{array}{ccc|c} 1 & 1 & 2 & -2 \\ 0 & 1 & 1 & 2 \\ 1 & -1 & 0 & 1 \end{array}\right] \xrightarrow{-R1} \left[\begin{array}{ccc|c} 1 & 1 & 2 & -2 \\ 0 & 1 & 1 & 2 \\ 0 & -2 & -2 & 3 \end{array}\right] \Rightarrow \left[\begin{array}{ccc|c} 1 & 1 & 2 & -2 \\ 0 & 1 & 1 & 2 \\ 0 & 0 & 0 & 2 \end{array}\right]$$

Week 4

Rank

Def. Given a set of vectors $V = \{v_1, \dots, v_n\}$ the num of un-ind vectors is called the rank of V

$$k = \text{rank}(V)$$

Rank shows richness of data

$Ax = b$ $(A^{-1})A = I$ ← computing matrices are error prone for finding inverse

$(A^{-1})Ax = (A)^{-1}b$

$x = \underline{A^{-1}b}$

Gaussian elemination is happening in background, Rank is important for data analysis

Example of recommender systems

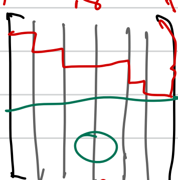
100

← vectors few linear independant vectors help predict patterns for users

10 000

What is the connection?

A =



k

Def Rank of Matrix A is equal to the number of linear independant columns

↑ pretty much what HW was

$\text{rank}(A) = k$

Solutions to $Ax = b$

single, infinite, none

The respective options are characterized as follows,

① Unique Solution ① Rank A = # variables (rank A = # of columns)

② Rank(A) = rank(A|b)

[A|b]

② ∞ solns: ① Rank A < variables

② Rank(A) = rank(A|b)

③ No solns: rank(A) < rank(A|b)

$$Ax=b \quad \begin{array}{ccc} x_1 & x_2 & x_3 \end{array} \begin{bmatrix} | & | & | \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} \; \end{bmatrix}$$

These conditions help in computation

Existence of inv. matrix

A matrix $A[n\times n]$ has the inverse if & only if $\operatorname{rank} A = n$

$Ax=b$

$[n\times n]$ A^{-1} exist if $\operatorname{rank}(A)=n$

Rank : transposition

$\operatorname{rank}(A) = \operatorname{rank}(A^T)$

Rank decomposition

Given a matrix $A, [n\times m]$ such that $\operatorname{rank} A = k < \min(n,)$

Then there exist matrices $B[n\times k]$ and $C[k\times m]$ s.t. $\operatorname{rank} B = \operatorname{rank} C = k$

$BC = A$

reduced

$Ax=b$

$$A = \left[\begin{array}{ccc|c} 1 & -4 & 1 & 1 \\ -2 & 0 & 2 & 1 \\ -1 & 4 & -3 & 1 \end{array}\right]$$

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

$+2R1 \downarrow\downarrow -3R1$

$$\left[\begin{array}{ccc|c} 1 & -4 & 3 & -1 \\ 0 & -8 & 8 & -1 \\ 0 & 8 & -8 & 4 \end{array}\right]$$

$$\left[\begin{array}{ccc|c} 1 & -4 & 3 & -1 \\ 0 & -8 & 8 & -1 \\ 0 & 0 & 0 & 3 \end{array}\right]$$

$\operatorname{Rank}(A) < \operatorname{Rank}(A|B)$

No soln

$\operatorname{Rank}(A) = 2$

$\operatorname{Rank}(A|b) = 3$

Linear spaces

Linear (in)dependence. Basis

The space of n-dimensional real vectors is denoted by $\mathbb{R}^n$. An element of $\mathbb{R}^n$ writes as

$$\mathbf{v} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix},$$

where $x_1, x_2, \ldots, x_n$ are real numbers. In the following, we will use straight boldface letters (e.g., $\mathbf{v}$) to denote vectors, and regular italic ones (e.g., x) to denote scalars. Particularly, $\mathbf{0}$ is a vector of zeros. The dimension of this vector must be clear from the context.

Definition (Linear combination). Given k vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k \in \mathbb{R}^n$, their linear combination is defined as

$$\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \ldots \alpha_k\mathbf{v}_k,$$

where $\alpha_1, \alpha_2, \ldots, \alpha_k$ are real numbers (scalars).

Attention: linear combination of any number of vectors from $\mathbb{R}^n$ always belongs to $\mathbb{R}^n$.

Definition (Linear independence). Given k vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k \in \mathbb{R}^n$, they said to be linearly independent if the equality

$$\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \ldots \alpha_k\mathbf{v}_k = \mathbf{0} \qquad (*)$$

holds **only** for $\alpha_1 = \alpha_2 = \cdots = \alpha_k = 0$. Otherwise, i.e., if $(*)$ holds for **at least one** non-zero α, the vectors are called linearly dependent.

We have several properties of linear (in)dependence:

1. The set of vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k \in \mathbb{R}^n$ is linearly dependent if some vector $\mathbf{v}_i$ is equal to $\mathbf{0}$.
2. The set of vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k \in \mathbb{R}^n$ is linearly dependent if any two vectors $\mathbf{v}_i$, $\mathbf{v}_j$ are equal.
3. If the set of vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k \in \mathbb{R}^n$ is linearly dependent, there is always a vector that can be expressed as a linear combination of the remaining ones.
4. The space $\mathbb{R}^n$ contains **at most** n linearly independent vectors.

Try to explain the first two properties. For simplicity, consider only 3 vectors. Assume that one vector, say, the first one, is zero. Can we find α's, **not all equal to** 0 such that the linear combination is $\mathbf{0}$? Try to think about the second property in the same way.

How to check linear independency?

1. Write the equation $(*)$ as a system of linear algebraic equations.
2. Perform Gaussian elimination to reduce the system to a triangular form.
3. The number of non-zero rows corresponds to the number of **linearly independent** vectors. If the number of linearly independent vectors equals the total number of vectors, then all vectors are linearly independent.
4. If the number of linearly independent vectors is less than the total number of vectors, we can identify the linearly independent ones as follows:

   Linearly independent vectors correspond to the leading elements in non-zero rows.

   For instance, after performing the Gaussian elimination, the system of equations has turned into

$$\begin{cases} \mathbf{2x_1} & +3x_2 & & -x_4 & = 0 \\ & & -\mathbf{x_3} & +x_4 & = 0 \\ & & & \mathbf{x_4} & = 0. \end{cases}$$

   Here, the leading elements are put in bold. Respectively, the independent vectors are the 1st, the 3rd, and the 4th.

Definition (Basis of $\mathbb{R}^n$) We say that any set of n linearly independent vectors in $\mathbb{R}^n$ forms a basis of $\mathbb{R}^n$.

Attention: there can be infinitely many bases in $\mathbb{R}^n$. Most often, we work with the Cartesian basis. For instance, the Cartesian basis of $\mathbb{R}^3$ is

$$C_3 = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}.$$

Expanding a vector with respect to a basis. The following theorem provides a theoretical foundation for the subsequent analysis.

Theorem. Given a basis of $\mathbb{R}^n$: $\mathcal{B} = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$. Any vector $\mathbf{x} \in \mathbb{R}^n$ can be uniquely reperesented as a linear combination of the basis vectors:

$$\mathbf{x} = \alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2 + \dots \alpha_n \mathbf{v}_n.$$

The numbers $\alpha_1, \dots, \alpha_n$ are called the coordinates of $\mathbf{x}$ in the basis $\mathcal{B}$.

Consider the situation, when we wish to express a vector in terms of others. This can happen in two cases:

1. Given a basis of $\mathbb{R}^n$, we wish to express some vector $\mathbf{x}$ as a linear combination of the basis vectors.

2. Given a set of linearly dependent vectors. We identified the subset of linearly independent vectors and want to express other vectors as linear combinations of independent vectors.

In either case, the procedure is the same: given a set of linearly independent vectors $\mathbf{v}_1, \ldots, \mathbf{v}_k$, and a vector $\mathbf{x}$, we need to solve the vector equation

$$\alpha_1 \mathbf{v}_1 + \alpha_2 \mathbf{v}_2 + \ldots \alpha_n \mathbf{v}_n = \mathbf{x}.$$

This equation can be written as a system of linear algebraic equations (now with non-zero free terms) and solved using the standard Gaussian elimination.

Rank

Definition. (Rank-1) Given a set of vectors $V = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$, $\mathbf{v}_i \in \mathbb{R}^k$, the number of linearly independent vectors in V is referred to as the rank of V and denoted by $\mathrm{rank}(V)$.

Note that while the linearly independent vectors (potentially) can be chosen in many ways, their number remains *unchanged*.

The notion of rank extends naturally to matrices.

Definition. (Rank-2) Given a matrix A, $[m \times n]$, the rank of A is defined as the number of linearly independent columns of A.

Below, we will see several properties of the rank. Let A, $[m \times n]$, and B, $[n \times k]$ be two real-valued matrices, then

1. $\mathrm{rank}(\mathrm{A}) \leq n$.
2. $\mathrm{rank}(A) = \mathrm{rank}(\mathrm{A}^\top)$.
3. $\mathrm{rank}(\mathrm{A}) \leq \min(m, n)$.
4. $\mathrm{rank}(\mathrm{AB}) \leq \min(\mathrm{rank}(\mathrm{A}), \mathrm{rank}(\mathrm{B}))$.
5. $\mathrm{rank}(\mathrm{AA}^\top) = \mathrm{rank}(\mathrm{A}^\top\mathrm{A}) = \mathrm{rank}(\mathrm{A})$.

Item 1 follows immediately from the definition. **Explain why.** Item 3 follows from 1. and 2. **Why?** The remaining properties will be taken for granted.

Application of the rank to solving a system of linear algebraic equations. The notion of rank can help to classify the systems of linear algebraic equations according to the number of solutions as shown in the folowing theorem.

Theorem. (Rouché-Capelli) Given a system

$$\mathrm{A}\mathbf{x} = \mathbf{b}, \qquad \text{(SLAU)}$$

where A is $[n \times m]$, while $\mathbf{x}$ and $\mathbf{b}$ have respective dimensions. There are 3 possible cases:

1. $\mathrm{rank}(A) = \#$ of variables $= n$ and $\mathrm{rank}(\mathrm{A}) = \mathrm{rank}(\mathrm{A}|\mathbf{b})$.

   $\Rightarrow$ (SLAU) has exactly 1 solution.

2. $\mathrm{rank}(A) = \#$ of variables $< n$ and $\mathrm{rank}(\mathrm{A}) = \mathrm{rank}(\mathrm{A}|\mathbf{b})$.

   $\Rightarrow$ (SLAU) has infinitely many solutions.

3. $\mathrm{rank}(\mathrm{A}) < \mathrm{rank}(\mathrm{A}|\mathbf{b})$.

   $\Rightarrow$ (SLAU) has no solution.

A particular case of the preceding theorem is formulated as follows.

Corollary. If A is a *square* matrix of dimensions $[n \times n]$ and $\mathrm{rank}(\mathrm{A}) = n$, then (SLAU) has exactly one solution.

The following theorems present further applications of the rank.

Theorem. (Existence of an inverse matrix) A square matrix A, $[n \times n]$, has an inverse if $\mathrm{rank}(\mathrm{A}) = n$.

Theorem. (Rank decomposition) Given a matrix A, $[n \times m]$ such that $\mathrm{rank}(\mathrm{A}) = k < \min(n, m)$. Then, there exist two matrices, B, $[n \times k]$, and C, $[k \times m]$ such that $\mathrm{rank}(\mathrm{B}) = \mathrm{rank}(\mathrm{C}) = k$ and $\mathrm{A} = \mathrm{BC}$.

Rank decomposition is computed according to the following algorithm (do not be afraid, we will consider it in detail in the class):

1. Identify the linearly independent columns of the matrix A. There must be exactly k such vectors, where $k = \mathrm{rank}(A)$.

2. Write the matrix B, whose columns coincide with the linearly independent columns of A.

3. The columns of the matrix C will correspond to the columns of A. So, we start from the first column and go till the last, while performing the following for the current, ith column:

   (a) If ith column of A is linearly dependent, then write the ith column of C as the coefficients of its expansion with respect to the columns of B.

   (b) If the ith column of A is linearly independent, then it must be included in B. In this case, write the ith column of C as $[0\ 0 \ldots 1 \ldots 0]^\top$, where 1 is located in the position, whose number coincide with the position of the respective column in B.

Homework 4: Rank

The second part of the homework is the sample test, which can also be found in the system.

1. Refresh the definition of the basis and the rank of a set of vectors. How are these related?

2. Determine the rank of the following matrices, identify the linearly independent columns:

$$\mathrm{A}_1 = \begin{bmatrix} -1 & 3 & 0 & 0 & -3 \\ -5 & 8 & -7 & 7 & -8 \\ -2 & 3 & -3 & 3 & -3 \\ -1 & 2 & -1 & 1 & -2 \end{bmatrix}, \quad \mathrm{A}_2 = \begin{bmatrix} -3 & -3 & -2 & -1 & 4 \\ -1 & 4 & -1 & 2 & 1 \\ -2 & -4 & 0 & -1 & 4 \\ 0 & -3 & 1 & -1 & 1 \end{bmatrix}.$$

3. Compute the rank of the matrix A and its transpose $\mathrm{A}^\top$. Was it equally difficult? Can we use the property $\mathrm{rank}(\mathrm{A}) = \mathrm{rank}(\mathrm{A}^\top)$ to simplify the computations? In which case this property can be most useful?

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

4. Assume that we are given an $[10^5 \times 10^4]$ matrix A. Let its rank be 100. How many elements do we need to store if we represent A using its rank decomposition? Now assume that the rank of A is 10. How much do we gain?

1. Refresh the definition of the basis and the rank of a set of vectors. How are these related?

Basis is the set of linear independant vectors of a vector space.

Rank can be found by the basis

2. Determine the rank of the following matrices, identify the linearly independent columns:

rank(A_1)=2 linear indep.

$$A_1 = \begin{bmatrix} -1 & 3 & 0 & 0 & -3 \\ -5 & 8 & -7 & 7 & -8 \\ -2 & 3 & -3 & 3 & -3 \\ -1 & 2 & -1 & 1 & -2 \end{bmatrix}, \qquad A_2 = \begin{bmatrix} -3 & -3 & -2 & -1 & 4 \\ -1 & 4 & -1 & 2 & 1 \\ -2 & -4 & 0 & -1 & 4 \\ 0 & -3 & 1 & -1 & 1 \end{bmatrix}.$$

5R1, 2R1, R1

$$\begin{bmatrix} 1 & -3 & 0 & 0 & 3 \\ 0 & -7 & -7 & 7 & 7 \\ 0 & -3 & -3 & 3 & 3 \\ 0 & -1 & -1 & 1 & 1 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & -3 & 0 & 0 & 3 \\ 0 & 1 & 1 & -1 & -1 \\ 0 & -1 & -1 & 1 & 1 \\ 0 & -1 & -1 & 1 & 1 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & -3 & 0 & 0 & 3 \\ 0 & 1 & 1 & -1 & -1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

rank(A_2)=3 linear indep.

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

+3R1, +2R1

$$\begin{bmatrix} 1 & -4 & 1 & -2 & -1 \\ 0 & 3 & -1 & 1 & -1 \\ 0 & -15 & 1 & -7 & 1 \\ 0 & -12 & 2 & -5 & 2 \end{bmatrix} \xrightarrow{+5R2,\ +4R2} \begin{bmatrix} 1 & -4 & 1 & -2 & 1 \\ 0 & 3 & -1 & 1 & -1 \\ 0 & 0 & -4 & -2 & -4 \\ 0 & 0 & -2 & -1 & -2 \end{bmatrix}$$

3. Compute the rank of the matrix A and its transpose $A^\top$. Was it equally difficult? Can we use the property $\mathrm{rank}(A) = \mathrm{rank}(A^\top)$ to simplify the computations? In which case this property can be most useful?

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

$$A = \begin{bmatrix} 1 & 3 \\ 0 & 1 \\ -3 & 1 \\ -1 & 2 \\ -2 & -4 \\ 4 & 1 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 3 \\ 0 & 1 \\ 0 & 10 \\ 0 & 5 \\ 0 & 2 \\ 0 & -11 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 3 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$$

$$A_2^T \begin{bmatrix} 1 & 0 & -3 & -1 & -2 & 4 \\ 3 & 1 & 1 & 2 & -4 & 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 & -3 & -1 & -2 & 4 \\ 0 & 1 & 10 & 5 & 2 & -11 \end{bmatrix}$$

4. Assume that we are given an $[10^5 \times 10^4]$ matrix A. Let its rank be 100. How many elements do we need to store if we represent A using its rank decomposition? Now assume that the rank of A is 10. How much do we gain?

Minimum matrix to using rank decomposition. $[10^2 \times 10^4]$, $[10^5 \times 10^2]$

Might have dif answer, called invarient

Test should be written without any auxiliary materials: only a pen and a sheet of paper. Duration of the test is 80 min.

Note: different problems have different weights. That is to say, you may get 1 point for the first problem and 3 for the third one. Exact weights will be indicated in the final version of the test.

Sample test

1. Given a system of linear algebraic equations. Write this system in the matrix form $\mathrm{A}\mathbf{x} = \mathbf{b}$.

$$\begin{aligned} 3x_1 \quad\quad + 2x_3 - 3x_4 &= 5 \\ -x_1 + 2x_2 + 3x_3 + x_4 &= 10 \\ 2x_1 - 2x_2 \quad\quad - 5x_4 &= 1, \end{aligned}$$

---

2. It is known that

$$\mathrm{A}^{-1} = \begin{bmatrix} 3 & -1 \\ 0 & 2 \end{bmatrix}, \text{ and } \mathrm{B}^{-1} = \begin{bmatrix} -1 & 1 \\ 1 & 0 \end{bmatrix},$$

compute $(\mathrm{AB})^{-1}$.

---

3. Given matrices A, $[2 \times 5]$ and B, $[4 \times 5]$. Determine, which of the following expressions are well-defined (= all operations can be performed).

$$a)\mathrm{AB}; \quad b)\mathrm{B} + \mathrm{A}^{\mathrm{T}}; \quad c)\mathrm{AB}^{\mathrm{T}}; \quad d)\mathrm{A} + \mathrm{AB}^{\mathrm{T}}; \quad e)\mathrm{A}(\mathrm{B}^{\mathrm{T}}\mathrm{B}).$$

---

4. Given matrices A, B, and C

$$\mathrm{A} = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}; \quad \mathrm{B} = \begin{bmatrix} 1 & 0 \\ 2 & -1 \end{bmatrix}; \quad \mathrm{C} = \begin{bmatrix} 1 & -1 \\ 0 & 2 \end{bmatrix}.$$

Compute the expression $(\mathrm{AB})^{\mathrm{T}} - 2\mathrm{C}$.

---

5. Given a set of vectors.

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

Determine whether these vectors are linearly independent. If not, choose from the given set a linearly dependent vector and write its representation as a linear combination of other vectors from the set.

6. Given the system

$$\begin{aligned} x_1 + 3x_2 - 2x_3 &= 1 \\ -2x_1 + x_2 + 3x_3 &= 4 \\ 3x_1 + 2x_2 - 5x_3 &= -3. \end{aligned}$$

Make a conclusion about the number of solutions based upon the Rouché-Capelli theorem.

1. Given a system of linear algebraic equations. Write this system in the matrix form $\mathbf{A}\mathbf{x} = \mathbf{b}$.

$$
\begin{aligned}
3x_1 \quad + 2x_3 - 3x_4 &= 5\\
-x_1 + 2x_2 + 3x_3 + x_4 &= 10\\
2x_1 - 2x_2 \quad - 5x_4 &= 1,
\end{aligned}
$$

$A \quad x = b$

$$\begin{bmatrix} 3 & 0 & 2 & -3 \\ -1 & 2 & 3 & 1 \\ 2 & -2 & 0 & -5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 5 \\ 10 \\ 1 \end{bmatrix}$$

2. It is known that

$$A^{-1} = \begin{bmatrix} 3 & -1 \\ 0 & 2 \end{bmatrix}, \text{ and } B^{-1} = \begin{bmatrix} -1 & 1 \\ 1 & 0 \end{bmatrix},$$

compute $(AB)^{-1}$.

Properties of inverse matrices

1. $(A^{-1})^{-1} = A$
2. $A^{-1}A = AA^{-1} = I$
3. $(AB)^{-1} = B^{-1}A^{-1}$ *
4. $(A^{-1})^T = (A^T)^{-1}$

Compute $(AB)^{-1}$

$B^{-1} \; A^{-1}$

$$\begin{bmatrix} 3 & -1 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} -1 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} -4 & 3 \\ 2 & 0 \end{bmatrix} = (AB)^{-1}$$

$-3-1 \quad 3+0$

$0+2 \quad 0+0$

3. Given matrices A, $[2 \times 5]$ and B, $[4 \times 5]$. Determine, which of the following expressions are well-defined (= all operations can be performed).

$A^T: [5 \times 2] \quad B^T: [5 \times 4]$

$[2 \times 4] \quad [5 \neq 5]$

a) AB; b) $B + A^T$; c) AB^T; d) $A + AB^T$; e) $A(B^TB)$.

✗ ✗ ✓ ✗ ✓

4. Given matrices A, B, and C

$$A = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}; \quad B = \begin{bmatrix} 1 & 0 \\ 2 & -1 \end{bmatrix}; \quad C = \begin{bmatrix} 1 & -1 \\ 0 & 2 \end{bmatrix}$$

Compute the expression $(AB)^T - 2C$.

Matrices - Operations

Properties of transposed matrices:

1. $(A+B)^T = A^T + B^T$
2. $(AB)^T = B^T A^T$ *
3. $(kA)^T = kA^T$
4. $(A^T)^T = A$

$B^T A^T - 2C$

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

$1+2 \quad -1+2$

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

$0-1 \quad 0-1$

$$\begin{bmatrix} 1 & 3 \\ -1 & -5 \end{bmatrix}$$

5. Given a set of vectors.

$$\vec{v}_1 = \begin{bmatrix} 1 \\ 0 \\ -1 \\ 0 \end{bmatrix}, \quad \vec{v}_2 = \begin{bmatrix} 2 \\ 1 \\ -2 \\ -1 \end{bmatrix}, \quad \vec{v}_3 = \begin{bmatrix} -2 \\ 1 \\ -1 \\ 1 \end{bmatrix}, \quad \vec{v}_4 = \begin{bmatrix} 1 \\ 1 \\ -1 \\ -1 \end{bmatrix}.$$

Determine whether these vectors are linearly independent. If not, choose from the given set a linearly dependent vector and write its representation as a linear combination of other vectors from the set.

How to check linear independency?

1. Write the equation $(*)$ as a system of linear algebraic equations.
2. Perform Gaussian elimination to reduce the system to a triangular form.
3. The number of non-zero rows corresponds to the number of **linearly independent** vectors. If the number of linearly independent vectors equals the total number of vectors, then all vectors are linearly independent.
4. If the number of linearly independent vectors is less than the total number of vectors, we can identify the linearly independent ones as follows:

   Linearly independent vectors correspond to the leading elements in non-zero rows.

   For instance, after performing the Gaussian elimination, the system of equations has turned into

$$\begin{cases} \mathbf{2x_1} + 3x_2 \quad - x_4 = 0 \\ \quad -\mathbf{x_3} + x_4 = 0 \\ \quad \mathbf{x_4} = 0. \end{cases}$$

   Here, the leading elements are put in bold. Respectively, the independent vectors are the 1st, the 3rd, and the 4th.

$$\begin{bmatrix} 1 & 2 & -2 & 1 \\ 0 & 1 & 1 & 1 \\ -1 & -2 & -1 & -1 \\ 0 & -1 & 1 & -1 \end{bmatrix} \begin{matrix} \\ \\ +R1 \\ +R2 \end{matrix} \Rightarrow \begin{bmatrix} 1 & 2 & -2 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & -3 & 0 \\ 0 & 0 & 2 & 0 \end{bmatrix}$$

$\div\frac{1}{3}R3 \quad -\frac{2}{3}R3$

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

$\vec{v}_1, \vec{v}_2, \vec{v}_3$ are linear independent

$$\left[\begin{array}{ccc|c} 1 & 2 & -2 & 1 \\ 0 & 1 & 1 & 1 \\ -1 & -2 & -1 & -1 \\ 0 & -1 & 1 & -1 \end{array}\right] \Rightarrow \left[\begin{array}{ccc|c} 1 & 2 & -2 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

$v_3 = 0$

$v_2 = 1$

$v_1 + 2 = 1$

$v_1 = -1$

$$-\vec{v}_1 + \vec{v}_2 = \vec{v}_4$$

6. Given the system

$$\begin{aligned} x_1 + 3x_2 - 2x_3 &= 1 \\ -2x_1 + x_2 + 3x_3 &= 4 \\ 3x_1 + 2x_2 - 5x_3 &= -3. \end{aligned}$$

Make a conclusion about the number of solutions based upon the Rouché-Capelli theorem.

Theorem. (Rouché-Capelli) Given a system

$$\mathrm{A}\mathbf{x} = \mathbf{b}, \qquad \text{(SLAU)}$$

where A is $[n \times m]$, while $\mathbf{x}$ and $\mathbf{b}$ have respective dimensions. There are 3 possible cases:

1. $\operatorname{rank}(A) =$ # of variables $= n$ and $\operatorname{rank}(\mathrm{A}) = \operatorname{rank}(\mathrm{A}|\mathbf{b})$.
   $\Rightarrow$ (SLAU) has exactly 1 solution.
2. $\operatorname{rank}(A) <$ # of variables $= n$ and $\operatorname{rank}(\mathrm{A}) = \operatorname{rank}(\mathrm{A}|\mathbf{b})$.
   $\Rightarrow$ (SLAU) has infinitely many solutions.
3. $\operatorname{rank}(\mathrm{A}) < \operatorname{rank}(\mathrm{A}|\mathbf{b})$.
   $\Rightarrow$ (SLAU) has no solution.

$$A \quad x = b$$

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

# of variables
3

Find Rank(A)

$$\begin{bmatrix} 1 & 3 & -2 \\ -2 & 1 & 3 \\ 3 & 2 & -5 \end{bmatrix} \xrightarrow[-3R1]{+2R1} \begin{bmatrix} 1 & 3 & -2 \\ 0 & 7 & -1 \\ 0 & -7 & 1 \end{bmatrix} \xrightarrow{+R2} \begin{bmatrix} 1 & 3 & -2 \\ 0 & 7 & -1 \\ 0 & 0 & 0 \end{bmatrix}$$

2 linear independent vectors
Rank(A) = 2

Find Rank(A|b)

$$\left[\begin{array}{ccc|c} 1 & 3 & -2 & 1 \\ -2 & 1 & 3 & 4 \\ 3 & 2 & 5 & -3 \end{array}\right] \xrightarrow[-3R1]{+2R1} \left[\begin{array}{ccc|c} 1 & 3 & -2 & 1 \\ 0 & 7 & -1 & 6 \\ 0 & -7 & 1 & -6 \end{array}\right] \xrightarrow{+R2} \left[\begin{array}{ccc|c} 1 & 3 & -2 & 1 \\ 0 & 7 & -1 & 6 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

2 linear independant vectors
rank(A|b) = 2

rank A < # of variables
& rank(A) = rank(A|b)
according to Rouché-Capelli theorem
the system has infinite solutions

Week 5

Idea of Rank decomposition

4×5 $\qquad A_2 = BC$

$$A_2 = \begin{bmatrix} -3 & -3 & -2 & -1 & 4 \\ -1 & 4 & -1 & 2 & 1 \\ -2 & -4 & 0 & -1 & 4 \\ 0 & -3 & 1 & -1 & 1 \end{bmatrix}.$$

this can be expressed as linear dep in decomp A_2

$\text{Rank}(A_2) = 3$

B $[4 \times 3]$ $\qquad$ C $[3 \times 5]$

$$\begin{bmatrix} -3 & -3 & -2 \\ -1 & 4 & -1 \\ -2 & -4 & 0 \\ 0 & -3 & 1 \end{bmatrix} \begin{array}{c} \begin{array}{ccccc} 1 & 2 & 3 & 4 & 5 \end{array} \\ \begin{bmatrix} 1 & 0 & 0 & -\frac{1}{2} & -2 \\ 0 & 1 & 0 & \frac{1}{2} & 0 \\ 0 & 0 & 1 & \frac{1}{2} & 1 \end{bmatrix} \end{array} = \begin{bmatrix} -3 & -3 & -2 \\ -1 & 4 & -1 \\ -2 & -4 & 0 \\ 0 & -3 & 1 \end{bmatrix}$$

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

$+3R1 \Downarrow \quad \Downarrow +2R1$ $\qquad\qquad$ $+2R1 \Downarrow \quad 3R1$

$$\left[\begin{array}{ccc|c} 1 & -4 & 1 & -2 \\ 0 & -15 & 1 & -7 \\ 0 & -12 & 2 & -5 \\ 0 & -3 & 1 & -1 \end{array}\right] \qquad \left[\begin{array}{ccc|c} 1 & -4 & 1 & -1 \\ 0 & 3 & -1 & -1 \\ 0 & -12 & 2 & 2 \\ 0 & -15 & 1 & -1 \end{array}\right]$$

$4R2 \Downarrow \quad 5R2$

$$\left[\begin{array}{ccc|c} 1 & -4 & 1 & -2 \\ 0 & 3 & -1 & 1 \\ 0 & -12 & 2 & -5 \\ 0 & -15 & 1 & -7 \end{array}\right] \qquad \left[\begin{array}{ccc|c} 1 & -4 & 1 & -1 \\ 0 & 3 & -1 & -1 \\ 0 & 0 & -2 & -2 \\ 0 & 0 & -5 & \end{array}\right]$$

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

Guassian elimination would be on the test

Rank decomposition

$$A_1 = \begin{bmatrix} -1 & 3 & 0 & 0 & -3 \\ -5 & 8 & -7 & 7 & -8 \\ -2 & 3 & -3 & 3 & -3 \\ -1 & 2 & -1 & 1 & -2 \end{bmatrix},$$

$\text{Rank}(A_1) = 2$

$A_1 = B \; C$

$[4 \times 2] \quad [2 \times 5]$

$$\begin{bmatrix} -1 & 3 \\ -5 & 8 \\ -2 & 3 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 0 & 3 & -3 & 0 \\ 0 & 1 & 1 & -1 & -1 \end{bmatrix} =$$

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

5R1↓ 2R1↓ 1R1↓

$$\left[\begin{array}{cc|c} 1 & -3 & 0 \\ 0 & -7 & -7 \\ 0 & -3 & -3 \\ 0 & -1 & -1 \end{array}\right]$$

$\boxed{\text{Singular value decomp}}$

Given A, $[n \times m]$, rank $A = p \leq \min(m,n)$

There exist orthogonal matrices, $U[n \times n]$, and $V[m \times m]$ and Σ, $[n \times m]$

where $\sigma_1 \geq \sigma_2 \geq \dots \sigma_p > 0$

$$\Sigma = \left[\begin{array}{cc|c} \sigma_1 & 0 & 0 \\ 0 & \sigma_p & \\ \hline 0 & & 0 \end{array}\right]$$

σ_i - singular numbers of A

orthogonal → $UU^T = U^T U = I$

$n \times n \qquad n \times m \qquad m \times n$

$$\underbrace{[\;\;]}_{p} \left[\begin{array}{c|c} \sigma_1 \dots \sigma_p & 0 \\ \hline 0 & 0 \end{array}\right] [\;\;] \}p$$

Use in recommender systems

$$U \Sigma V = [n \times p] \begin{bmatrix} \sigma_1 & & 0 \\ & \ddots & \\ 0 & & \sigma_p \end{bmatrix} [p \times m]$$

$$
\begin{array}{c|cccc}
 & 1 & 2 & 3 & M \cdots \\
\hline
1 & 1 & - & 4 & \cdots \\
2 & 2 & - & 4 & \cdot \\
3 & - & 2 & 3 & 3 \\
4 & & & &
\end{array}
$$

U

Netflix

$U \sim 500{,}000$

$M \sim 20{,}000$

Ratings $\sim 100{,}000{,}000$

Validations $\sim 2{,}000{,}000$

$\forall\, R = U \Sigma V \Rightarrow$

$$\Sigma = \begin{bmatrix} \sigma_1 & & 0 \\ & \ddots & \\ 0 & & \sigma_p \end{bmatrix}$$

$$r_{ij} = \sum_{k=1}^{p} \underbrace{U_{ik}}_{\text{user}}\, \overset{\text{population}}{\sigma_k}\, \overset{\text{movie}}{V_{kj}}$$

p latent features

① Eigenvector/Eigenvalue

$Av = w$

$A[n \times n]$ ← will only consider in rest of class

$Av = \lambda v$, $v \in \mathbb{R}^n$, $\lambda \in \mathbb{R}$

λ - eigenvalue

v - eigenvector

(λ, v) - eigenpair

Fact 1. A matrix $A[n \times n]$ has exactly n distinct eigenpairs (λ_i, v_i)

$Av = \lambda v$

fine → $\boxed{\lambda = 0}$ $Av = 0$

$v \neq 0$

Connection with Rank

Fact 2

Matrix A, $\operatorname{rank} A = k$ has exactly k non-zero eigvalues

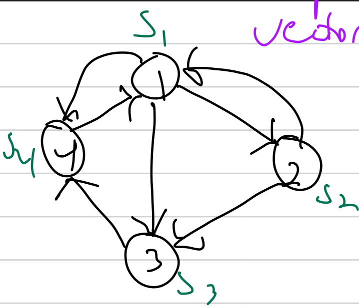
Fact 3 For a given λ, v is defined up to multiple by a non-zero scalar

$A_1 \cdots [10\times 10]$, rank $A = 5$

5 linear independent vectors

page rank by google

$\lambda_1 = 0 \quad v_1$
$\lambda_2 = 0 \quad v_2$
$\lambda_3 = 0 \quad v_3$
$\lambda_4 = 0 \quad v_4$
$\lambda_5 = 0 \quad v_5$



$s_2 + s_4 = s_1$
$s_1 \qquad\qquad 4 = s_2$
$s_1 + s_2 = s_3$
$s_1 + s_3 = s_4$

$$\begin{bmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \end{bmatrix} s = s$$

$As = s \qquad s = \begin{bmatrix} s_1 \\ \vdots \\ s_4 \end{bmatrix}$

Determinant

$\det : A \to \mathbb{R}$

Properties of det.

① If A' is obtained from A by swapping 2 cols (2 rows) then $\det A' = -\det A$

② If A' is obtained from A by multipl of any single row (column) of A by $k \neq 0$ then $\det A'' = k \det A$

③ If some column (row) of A can be written as a sum of 2 colums

$$A = \begin{bmatrix} | & | & | \\ c_1 & c_2 + c_3 & c_4 \\ | & | & | \end{bmatrix}, \text{ then } \det A = \det \begin{bmatrix} | & | & | \\ c_1 & c_2 & c_4 \\ | & | & | \end{bmatrix} + \begin{bmatrix} | & | & | \\ c_1 & c_3 & c_4 \\ | & | & | \end{bmatrix}$$

④ $\det I = 1$

$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \left[\begin{bmatrix} a \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ c \end{bmatrix} \quad \begin{matrix} b \\ d \end{matrix}\right]$

$\det A = \det \begin{bmatrix} a & b \\ 0 & d \end{bmatrix} + \det \begin{bmatrix} 0 & b \\ c & d \end{bmatrix} = \cancel{\det \begin{bmatrix} a & b \\ 0 & 0 \end{bmatrix}} + \det \begin{bmatrix} a & 0 \\ 0 & d \end{bmatrix} + \det \begin{bmatrix} 0 & b \\ c & 0 \end{bmatrix} + \cancel{\det \begin{bmatrix} 0 & 0 \\ c & d \end{bmatrix}}$

$ab \det \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} = 0 \qquad ad \det \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

$ad - bc$

$\det A = ad - bc$

Properties of determinates

1. If rank $A < n \Longleftrightarrow \det A = 0$
2. $\det A = \lambda_1 \cdot \lambda_2 \cdots \lambda_n$
3. $(I\lambda - A)$

$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \quad \begin{bmatrix} \lambda - a & -b \\ -c & \lambda - b \end{bmatrix}$

$\det(\lambda I - A) = (\lambda - a)(\lambda - b) - bc = \lambda^2 - \lambda(a+d) + \underbrace{(ad - bc)}_{\det A}$

Homework 5: Rank decomposition

1. Determine the rank of the following matrices, identify the linearly independent columns. Compute the rank decomposition using the previously discussed algorithm:

$$\mathrm{A}_1 = \begin{bmatrix} 3 & 4 & 7 & -8 & -5 \\ -2 & -3 & -5 & 5 & 3 \\ 1 & 3 & 4 & -1 & 0 \\ 1 & 1 & 2 & -3 & -2 \\ -2 & -3 & -5 & 5 & 3 \end{bmatrix}, \quad \mathrm{A}_2 = \begin{bmatrix} 2 & -5 & -5 & 3 & -5 \\ 1 & 0 & 0 & 4 & 0 \\ 0 & 2 & 2 & 2 & 2 \\ -2 & 4 & 4 & -4 & 4 \\ -2 & 5 & 5 & -3 & 5 \end{bmatrix}.$$

2. Multiply two matrices of dimensions $[5 \times 2]$ and $[2 \times 5]$, with *arbitrarily chosen entries*. Check that despite the resulting matrix has the dimensions $[5 \times 5]$, its rank does not exceed 2.

For the rest of the homework you are invited to refresh the material we covered so far and get ready for the test on this Thuesday.

1. Determine the rank of the following matrices, identify the linearly independent columns. Compute the rank decomposition using the previously discussed algorithm;

linear independant

$$A_1 = \begin{bmatrix} 3 & 4 & 7 & -8 & -5 \\ -2 & -3 & -5 & 5 & 3 \\ 1 & 3 & 4 & -1 & 0 \\ 1 & 1 & 2 & -3 & -2 \\ -2 & -3 & -5 & 5 & 3 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 2 & -5 & -5 & 3 & -5 \\ 1 & 0 & 0 & 4 & 0 \\ 0 & 2 & 2 & 2 & 2 \\ -2 & 4 & 4 & -4 & 4 \\ -2 & 5 & 5 & -3 & 5 \end{bmatrix}.$$

$\mathrm{rank}(A_2) = 3$

$$A_1' = \begin{bmatrix} 1 & 1 & 2 & -3 & -2 \\ 1 & 3 & 4 & -1 & 0 \\ 3 & 4 & 7 & -8 & -5 \\ -2 & -3 & -5 & 5 & 3 \\ -2 & -3 & -5 & 5 & 3 \end{bmatrix} \begin{matrix} \\ -R1 \\ -3R1 \\ +2R1 \\ -R4 \end{matrix} \Rightarrow \begin{bmatrix} 1 & 1 & 2 & -3 & -2 \\ 0 & 2 & 2 & 2 & 2 \\ 0 & 1 & 1 & 1 & 1 \\ 0 & -1 & -1 & -1 & -1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

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

$\mathrm{rank}(A_1) = 2$

$[5 \times 2] \quad [2 \times 5]$

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

$$\begin{bmatrix} 3 & 4 & -8 \\ -2 & -3 & 5 \\ 1 & 3 & -1 \\ 1 & 1 & -3 \\ -2 & -3 & 5 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 1 & -3 \\ 1 & 3 & -1 \\ 3 & 4 & -8 \\ -2 & -3 & 5 \\ -2 & -3 & -5 \end{bmatrix} \begin{matrix} \\ -R1 \\ -3R1 \\ +2R1 \\ +R4 \end{matrix} \Rightarrow \begin{bmatrix} 1 & 1 & -3 \\ 0 & 2 & 2 \\ 0 & 1 & 1 \\ 0 & -1 & -1 \\ 0 & 0 & 0 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 1 & -2 \\ 1 & 3 & 0 \\ 3 & 4 & -5 \\ -2 & -3 & 3 \\ -2 & -3 & 3 \end{bmatrix} \quad \begin{bmatrix} 1 & 1 & -2 \\ 0 & 2 & 2 \\ 0 & 1 & 1 \\ 0 & -1 & -1 \\ 0 & 0 & 0 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 & 0 & 4 & 0 \\ 0 & 1 & 1 & 1 & 1 \\ 2 & 4 & 4 & -4 & 4 \\ 2 & -5 & -5 & 3 & -5 \\ -2 & 5 & 5 & -3 & 5 \end{bmatrix} \begin{matrix} \\ \\ -2R1 \\ -2R1 \\ +R4 \end{matrix} \Rightarrow \begin{bmatrix} 1 & 0 & 0 & 4 & 0 \\ 0 & 1 & 1 & 1 & 1 \\ 0 & 4 & 4 & -12 & 4 \\ 0 & -5 & -5 & -5 & -5 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 & 0 & 4 & 0 \\ 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & -3 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \xrightarrow{-R2} \begin{bmatrix} 1 & 0 & 0 & 4 & 0 \\ 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & -2 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$5 \times 3 \quad 3 \times 5$

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

Week 6

Combinatorics

① permutation w. repetition
In how many ways we can fill K positions with the elements from a set of n elements?

② Permutations of n elements $\boxed{P_n = n!}$ Permutations

③ In how many ways we can choose k elements from n if the order is important? $\boxed{A_n^k = \frac{k!}{(n-k)!}}$ Arrangements

④ In how many ways we can choose K elements from n if the order is not important

Probability

$S = \{1, 2, 3, 4\}$

$\underline{\quad} \; \underline{\quad} \; \underline{\quad}$
$4 \cdot 4 \cdot 4$

$4^3 = $ 16 × 4 = 64

$\Pi_n^K = n^K$

Factorial
4!

$\underline{\quad} \; \underline{\quad} \; \underline{\quad} \; \underline{\quad}$
$4 \cdot 3 \cdot 2 \cdot 1$

$n! = 1 \cdot 2 \cdot \ldots \cdot n$
$0! = 1$

$\underline{\quad} \; \underline{\quad}$
$4 \quad 3$ $\quad \frac{4!}{2!}$

Write 2 digit int from set not repeating

$$\boxed{C_n^K = \frac{n!}{(n-k)!\,k!}}$$
Combinations

1. $\Pi_n^K = n^K$
2. $P_n = n!$
3. $A_n^K = n!/(n-k)!$
4. $C_n^K = n!/(n-k)!\,k!$

① How many ways are there to write a 4-digit #?
1111 ✓ 0111 ✗

$\underline{\quad} \; \underline{\quad} \; \underline{\quad} \; \underline{\quad}$
$9 \quad 10 \quad 10 \quad 10$

② There are 5 player in a team, in how many ways can we arrange them in a row?

$5! = 5 \times 4 \times 3 \times 2 \times 1 = 60$

Here we will need to translate word problems to math formula

③ ① ② ③ ④ ⑤

Pulling out 2 random balls out

$$\frac{5!}{(5-2)!} = \frac{5!}{3!} = 5 \times 4 = 20$$

Sometimes irrelevant information

④ There are 3 apples, 4 oranges and 2 lemons on the table. In how many ways can we take 2 oranges

$$C_4^2 = \frac{4!}{(2)!\,2!} = \frac{24}{4} = 6$$

Special cases can be brought out

⑤ Compute the number of all 5-digit with all different digits

A_9^5

$9 \quad \frac{9!}{5!} \qquad 9 \times 9 \times 8 \times 7 \times 6$

$9 \times 9 \times 8 \times 7 \times 3$

$8! \times 8 = 648$, $\times 42$, 6

Can split problems apart

⑥ 5 boys How many mixed pairs?
8 girls

$C_5^2 \qquad C_8^2$

If we want 2 dance pairs

$40 \times 7 = 280$

$8 \times 7 \times 5$

$$\frac{5!}{(3)!\,2!} \qquad \frac{8!}{(6)!\,2!} \qquad \frac{8 \times 7 \times 5 \times \cancel{4}}{\cancel{4}}$$

1 Combinatorics

First, we repeat the main combinatoric formulae. In there we will use the factorial, which is defined as folows:

$$n! = 1 \cdot 2 \cdot \dots \cdot (n-1) \cdot n, \quad 0! = 1.$$

Here we go!

1. In how many ways can one fill k[1] positions with the elements from the set of n elements $N = \{1, \dots, n\}$ if the elements of the set N can be used repeatedly? There is no standard name for this number. We will call it *permutations with repetitions*:

$$\Pi_n^k = n^k.$$

2. Given a set of n elements. The elements of this set can be written in $n!$ possible ways. This number is called the *number of permutations*:

$$P_n = n!.$$

3. In how many ways can we choose k elements fron the set of n elements if the order of elements **is important**? This number calls the *number of arrangements*[2]:

$$A_n^k = \frac{n!}{(n-k)!}.$$

4. Consider the previous problem, but assume that the order of elements **is not important**. So, we get the *number of combinations*:

$$C_n^k = \binom{n}{k} = \frac{A_n^k}{P_k} = \frac{n!}{k!(n-k)!}$$

Problems

Lets consider some combinatorial problems. Note that we do not deal with probabilities yet!

1. In how many ways can we write a 4-digit number if we additionally require that it does not start with 0?

   Hint: here we should use permutations with repetitions (as digits can be arbitrary), but we need to consider the first digit separately.

[1] Here and henceforth we assume $k < n$.

[2] The word *arrangements* is mostly used in French and Russian literature. In English literature sometines the name *partial permutations* is used, but I feel it may lead to a confusion.

2. There are 5 players in the team. In how many ways can we place them in a row? Consider a variation of this problem. Assume that there is one particular player X that has to be in the center (i.e., his/her place is fixed: $**X**$). How many options to place the remaining players do we have then?

   Hint: here we use permutations without repetitions (we cannot place the same player in two different positions).

3. There are numbered 5 balls in a backet, each ball has a unique number, i.e., either 1, 2, 3,4, or 5. We draw two balls. How many 2-digit numbers can we obtain in this way?

   Hint: Here the order plays a role. Indeed, 12 is different from 21.

4. There are 3 apples, 4 oranges, and 2 lemons on the table. In how many ways can we take 2 oranges? 2 citrus fruits?

   Hint: we do not care about the order in which we take the fruits.

5. (This problem is more complicated). Compute the number of all 5-digit even numbers with **all different** digits. For instance, 12346 is OK, but 01234 and 23454 are not accepted.

   Hint: the first and the last digits should be considered separately.

2 Probability

First, we define the main notions.

1. The *sample space* or the space of *elementary outcomes*

$$\Omega = \{\omega_1, \dots, \omega_n\}.$$

   Set

   For a die, we have $\Omega = \{$⚀, ⚁, ⚂, ⚃, ⚄, ⚅$\}$, or $\Omega = \{1, 2, 3, 4, 5, 6\}$ for simplicity.

   For a coin, $\Omega = \{H, T\}$. If we toss the coin twice, we have $\Omega = \{HH, HT, TH, TT\}$ and so on. . .

2. The set of all events contains all possible **subsets** of elements from Ω along with $\emptyset$ and Ω itself:

$$\mathcal{A} = \{\emptyset, \{1\}, \{2\}, \dots, \{6\}, \{1, 2\}, \{1, 3\}, \dots, \{1, 3, 5, 6\}, \dots \Omega\}.$$

3. The probability of an event $A \in \mathcal{A}$ is computed as the ratio

$$P(A) = \frac{m}{n},$$

where m is the number of *favorable outcomes* that correspond to the event A (i.e., the number of elements in the set A), and n is the total number of *elementary outcomes*.

Example: Compute the probability of getting an odd number of dots when tossing a die. We have $\Omega = \{1, 2, 3, 4, 5, 6\}$, i.e., $n = \#\Omega = 6$, and the event $A_{odd} = \{1, 3, 5\}$. So, $m = \#A_{odd} = 3$. Finally, we obtain $P(A_{odd}) = \frac{3}{6} = \frac{1}{2}$.

Here and henceforth we will use the symbol # to denote the number of elements in a set. Often, this is written using bars: $\#\Omega = |\Omega|$. The number of elements in a set is called the *cardinality* of a set.

Properties of probability

Now we formulate some basic properties of probability and introduce several additional definitions.

Rolling 1-11 on a dice

$\frac{1}{6} \cdot \frac{1}{6} = \frac{1}{36}$

$\frac{35}{36}$

$P(A_1) = 1/3$

$P(A_2) = 1/3$

$P(A_1 A_2) = 1/2$

1. Probability **always** takes values between 0 and 1:

$$P(\emptyset) = 0, \quad P(\Omega) = 1.$$

2. **[Complementary event].** Event $\bar{A} = \Omega \setminus A$ is said to be a *complement* of A. It holds that

   $A \subseteq \Omega$ $\quad$ $A = \{1, 2\}$

   $\bar{A} = \Omega \setminus A$ $\quad$ $\bar{A} = \{3, 4, 5, 6\}$

$$P(A) + P(\bar{A}) = 1.$$

$A_1 = \{1, 2\}$

$A_2 = \{2, 3\}$

$A_1 + A_2 = \{1, 2, 3\}$

3. **[Mutually exclusive events].** Events A_1 and A_2 are said to be *mutually exclusive* if the appearance of A_1 excludes the appearance of A_2 and vice versa.

   Example: Consider the event *"get an odd number of dots when tossing a die"*, $A_{odd} = \{1, 3, 5\}$, and the event *"get an even number of dots"*, $A_{even} = \{2, 4, 6\}$. These events are mutually exclusive. This follows from the fact that their intersection is empty, $A_{odd} \cap A_{even} = \emptyset$.

   $\cap$: intersection

$A_1 = \{1\}$

$A_2 = \{2\}$

$A_1 + A_2 = \{1, 2\}$

$A = \{1, 2\}$

$B = \{3, 4\}$

$A + B = \{1, 2, 3, 4\}$

4. **[Sum of events].** Event A is said to be a sum of events A_1 and A_2, denoted $A = A_1 \vee A_2 = A_1 + A_2$ if it occurs when **either** A_1 **or** A_2 occur. This corresponds to the **union** of the sets of elementary outcomes that correspond to A_1 and A_2.

   Example: Consider the events *"get an odd number of dots"*, $A_{odd} = \{1, 3, 5\}$, and the event *"get the number of dots divisible by 3"*, $A_{div3} = \{3, 6\}$. The sum of these events is $A = A_1 \vee A_2 = \{1, 3, 5, 6\}$. Note that A_1 and A_2 **are not** mutually exclusive as follows from $A_1 \cap A_2 = \{3\}$.

5. Probability of a sum of mutually exclusive events is equal to the sum of respective probabilities:

$$P(A_1 \vee A_2) = P(A_1) + P(A_2).$$

$\{1, 2\}, \{3, 4, 5\}, \{6\}$

A_1 or $A_2 = A_1 + A_2 = A_1 \cup A_2$

A_1 and $A_2 = A_1 \cdot A_2 = A_1 \cap A_2$

6. The set of mutually exclusive events $A_1, \ldots A_k$ is said to be *complete* if

$$P(A_1) + \ldots P(A_k) = 1.$$

7. **[Product of events].** Event A is said to be a product of events A_1 and A_2, denoted $A = A_1 \wedge A_2 = A_1 \cdot A_2$ if it occurs when **both** A_1 **or** A_2 occur. This corresponds to the **intersection** of the sets of elementary outcomes that correspond to A_1 and A_2.

   Example: Consider the same events $A_{odd} = \{1, 3, 5\}$ and $A_{div3} = \{3, 6\}$ as in item 4. The product of these events is $A = A_1 \wedge A_2 = \{3\}$.

8. If two events A_1 and A_2 are mutually exclusive,

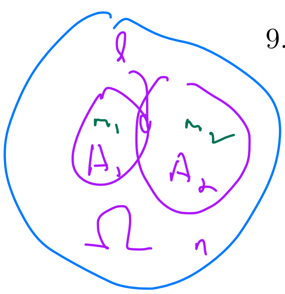
$$P(A_1 \wedge A_2) = 0.$$

   In particular, $P(A \wedge \bar{A}) = 0$ for any $A \in \mathcal{A}$.

$A_1 = \{2, 4, 6\}$

$A_2 = \{3, 6\}$

$\frac{1}{2} \cdot \frac{1}{3} = \frac{1}{6}$



9. **[Independent events].** Two events are said to be *independent* if

$$P(A_1 \cdot A_2) = P(A_1) \cdot P(A_2).$$

   Doesn't effect the other

   The notion of independence may sometimes appear counterintuitive, but this is rather a result of our incomplete understanding of the nature of the problem, rather than a sign of inconsistency in the theory.

10. **[Probability of the sum of events].** Given two, not necessarily mutually exclusive events, probability of their sum is equal to

$$P(A_1 \vee A_2) = P(A_1) + P(A_2) - P(A_1 \wedge A_2).$$

   For three events, this expression becomes larger, but its structure remains simple:

$$P(A_1 \vee A_2 \vee A_3) = P(A_1) + P(A_2) + P(A_3) \\ - P(A_1 \wedge A_2) - P(A_1 \wedge A_3) - P(A_2 \wedge A_3) + P(A_1 \wedge A_2 \wedge A_3).$$

   To understand how we obtain this formula, search in the Internet for the *inclusion–exclusion principle.*

$P(A_1) = \frac{m_1}{n}$

$P(A_2) = \frac{m_2}{n}$

$P(A_1 + A_2) = \frac{m_1 + m_2 - l}{n}$

11. **[The law of total probability].** For any two events $A, B \in \mathcal{A}$ it holds that

$$P(A) = P(A \wedge B) + P(A \wedge \bar{B}).$$

   This law can be generalized as follows. Let $B_1, B_2 \ldots, B_k$ be a complete set (of mutually exclusive events satisfying $\bigcup B_i = \Omega$). Then we have

$$P(A) = P(A \wedge B_1) + P(A \wedge B_2) + \cdots + P(A \wedge B_k).$$

$P(A_1) \cap P(A_2)$

$\frac{m_1}{n} + \frac{m_2}{n} - \frac{l}{n}$

$(A_1 \cdot A_2)$ $P(\square + \square = 5) = \frac{4}{36}$

1,4 4,1

2,3 3,2

$\square + \square =$

2,4,6,8,10

2,3,4,6,9,12

Toss Coin 10 times probability to get exact 4 heads

$$C_{10}^{4} = \frac{10!}{(6!)4!} = \frac{10\times \overset{3}{\cancel{9}}\times \overset{2}{\cancel{8}}\times 7}{\cancel{4}\times\cancel{3}\times\cancel{2}} = \frac{210}{2^{10}}$$

$$C_{10}^{6} = \frac{10!}{(4!)6!} = \frac{210}{1024}$$

Same as 4 tails

Problems

1. Let $P(A)$ and $P(B)$ be the probabilities of the events A and B. Which of the following situations are possible? For the possible situations, suggest a possible interpretation using the die model.

	$P(A)$	$P(B)$	$P(A \vee B)$
a)	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$
b)	$\frac{1}{2}$	$\frac{1}{3}$	$\frac{1}{6}$
c)	$\frac{1}{6}$	$\frac{1}{3}$	$\frac{1}{2}$
d)	$\frac{2}{3}$	$\frac{2}{3}$	1
e)	$\frac{1}{3}$	$\frac{1}{3}$	1

2. Assume that we toss 3 coins. What is the probability that the number of heads is larger than the number of tails?

3. What is the probability to win the game 5/35? In this game, one should guess 5 numbers out of 35.

4. Let there be 3 red and 5 green balls in a basket. We randomly take 3 balls from the basket. What is the probability that get exactly 1 red and 2 green balls?

3 Random variables

Consider the situation when we toss two dice and count the total number of dots on its faces. We can compute the probabilities of getting each particular number of dots, starting from 2 to 12.

The number of the dots can be seen as *random variable* X that takes its values within the set $\{2, 3, \ldots, 11, 12\}$ and each value appears with certain probability. The function f that assigns to each value of X its probability is called the *probability mass function* or the *discrete probability distribution*.

The values of the probability mass function can be conveniently written as a table:

X	2	3	4	5	6	7	8	9	10	11	12
$f(X)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

Any (discrete) probability distribution function **must** satisfy the following properties:

1. All probabilities are **non-negative**: $f(x) \geq 0$.

2. The probabilities of all values of X **must sum to** 1: $\sum_x f(x) = 1$.

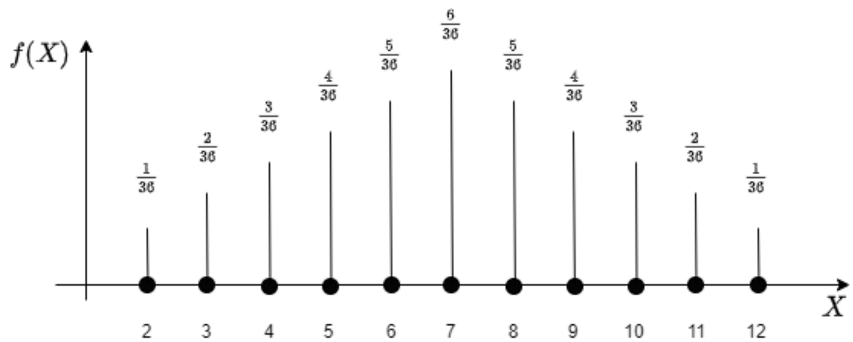


Рис. 1: The graph of a probability mass function.

The same rules remain valid for any, not necessarily discrete probabbility distribution, but in general the sum is replaced with an integral.

To illustrate a given discrete probability distribution we often use a “comb”-plot as shown in Fig. 1.

Consider the following three most common discrete probability distributions:

1. *Uniform* distribution. X takes values in the set $\mathcal{X} = \{1, 2, \ldots, n\}$ with equal probability. The probability function is thus

$$f(x) = \begin{cases} \frac{1}{n}, & x \in \mathcal{X}, \\ 0, & \text{otherwise.} \end{cases}$$

2. *Binomial* distribution. Assume that we repeat the same experiment n times. This experiment has only two outcomes[3], a “success” with the probability p, and a “failure” with the probability $q = 1 - p$. The binomial distribution describes the probability of getting exactly k positive outcomes in a series of n experiments:

$$f_{\text{bin}}(k) = \begin{cases} C_n^k \, p^k \, q^{n-k}, & k \in \{0,\, 1,\, \ldots,\, n\}, \\ 0, & \text{otherwise.} \end{cases}$$

Example. Let X be the number of “Heads” in 10 consecutive tosses of an unbiased coin (i.e., P(“Heads”) $= P$(“Tail”) $= 0.5$). Compute the probabilities of getting exactly k heads’. Do not forget to use combinatorics when computing the number of ways in which you can get a given number of heads. Note that they do not need to come up in a row.

[3] Obviously, the words *success* and *failure* should not be considered literally.

3. *Geometric* distribution. In the previously described setting, the geometric distribution gives the probabilities of the **first** success at the kth step. We have

$$f_{\text{geom}}(k) = q^{k-1}p.$$

Note that the variable k takes all positive integer values starting from 1. Thus, to check that the discrete probability distribution $f_{\text{geom}}(k)$ is consistent, we would need to show that

$$\sum_{k=1}^{\infty} q^{k-1}p = 1$$

for any two p and q such that $p + q = 1$. This amounts to computing a geometric series, whence the name. Recall that well known formula:

$$p + pq + pq^2 + \ldots = \frac{p}{1-q} = \frac{p}{p} = 1.$$

3.1 Numerical characteristics of a random variable

So far, we characterised a random variable by presenting the probabilities of all its realisations. In general, such information is not very useful, so we prefer to have some more "condensed" characteristics. We consider the most important ones.

Mean. Let the random variable X takes the values $x_1, \ldots, x_k$ with the probabilities $f(x_1), \ldots, f(x_k)$. The *mean value* or the *expectation* of X is denoted by $\bar{X}$ or $\mathrm{E}[X]$ and defined as follows:

$$\bar{X} = x_1 f(x_1) + \cdots + x_k f(x_k) = \sum_{i=1}^{k} x_i f(x_i).$$

One can easily check the linearity properties of the mean:

$$\begin{aligned} \mathrm{E}[X+Y] &= \mathrm{E}[X] + \mathrm{E}[Y], \\ \mathrm{E}[aX] &= a\mathrm{E}[X], \quad a \in \mathbb{R}. \end{aligned}$$

The mean value can be interpreted in many ways, the simples one is to think of it as an expected average value after a sufficiently large number of experiments. Returning to the example with tossing 2 dice, we can say that the **average number** of dots on two dice will approach 7 (= the mean of X) with an increasing number of experiments.

Often, the mean is used to define a centered random variable X^o, which is obtained by subtracting the mean $\bar{X}$ from X:

$$X^o = X - \bar{X}.$$

The table below shows both the original random variable X and its centered version X^o.

X	2	3	4	5	6	7	8	9	10	11	12
X^o	-5	-4	-3	-2	-1	0	1	2	3	4	5
$f(X)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

One can easily check that the expected value of X^o is **always** equal to 0.

Variance. The *variance* describes how far the realisations of X are spread out from its average value. The variance is denoted by $Var(X)$ and defined as

$$Var(X) = \mathrm{E}\left[(X^o)^2\right] = \mathrm{E}\left[(X - \bar{X})^2\right] = \sum_{i=1}^{k}(x_i - \bar{X})^2 f(x_i).$$

In practice, instead of variance one often uses the *standard deviation*

$$\sigma(X) = \sqrt{Var(X)}.$$

For the same reason, one sometimes defines the variation as $Var(X) = \sigma^2$.

For the previously described distributions, both the mean and the variance can be computed explicitely.

- Uniform distribution:
$$\bar{X} = \frac{n+1}{2}, \quad var(X) = \frac{n^2-1}{12}.$$
- Binomial distribution:
$$\bar{X} = np, \quad Var(X) = npq.$$
- Geometric distribution:
$$\bar{X} = \frac{1}{p}, \quad Var(X) = \frac{1-p}{p^2}.$$

4 Bayes' theorem

The Bayes theorem is a particularly important tool that allows us to compute the probabilities of some events from the probabilities of other events.

To start with, recall that two events $A, B \in \mathcal{A}$ are said to be independent if $P(A \wedge B) = P(A)P(B)$. Otherwise, the events are called dependent. For such events we introduce the notion of *conditional probability*: $P(A|B)$ is the conditional probability of A given B. It is defined as

$$P(A|B) = \frac{P(A \wedge B)}{P(B)}.$$

The preceding formula can be rewritten as

$$P(A \wedge B) = P(A|B)P(B).$$

That is to say, the joint probability of A and B is equal to the probuct of the conditional probability of A given B, $P(A|B)$, and the probability of B.

In the same way we can define the conditional probability of B given A:

$$P(B|A) = \frac{P(A \wedge B)}{P(A)},$$

whence we get

$$P(A \wedge B) = P(B|A)P(A).$$

Combining two different expressions for $P(A \wedge B)$, we recover a formula that relates $P(A|B)$ and $P(B|A)$:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}.$$

This formula is referred to as the *Bayes theorem*. In practice, any of the above formulas ca be used depending on the context.

For some applications of the Bayes theorem see the Wikipedia article. The main rule of solving problems consists in correct determination of the events and their probabilities (also, conditional probabilities). Do not forget to use the law of total probability (see Section 2).

4.1 Problems

1. Consider a test to detect a disease that 0.1% of the population have. The test is 99% effective in detecting an infected person. However, the test gives a false positive result in 0.5% of cases. If a person tests positive for the disease what is the probability that they actually have it?

2. Suppose $P(A)$, $P(\bar{A})$, $P(B|A)$, and $P(B|\bar{A})$ are known. Find an expression for $P(A|B)$ in terms of these four probabilities.

3. In a casino in Blackpool there are two slot machines: one that pays out 10% of the time, and one that pays out 20% of the time. Obviously, you would like to play on the machine that pays out 20% of the time but you do not know which of the two machines is the more generous. You thus adopt the following strategy: you assume initially that the two machines are equally likely to be the generous machine. You then select one of the two machines at random and put a coin into it. Given that you loose that first bet estimate the probability that the machine you selected is the more generous of the two machines.

4. Consider 3 coins where two are fair, yielding heads with probability 0.50, while the third yields heads with probability 0.75. If one randomly selects one of the coins and tosses it 3 times, yielding 3 heads - what is the probability this is the biased coin?

Homework #6

Do not be afraid if you cannot solve some of the problems. Get your act together and off we go!

1. A coin is tossed 8 times. In how many ways can we get a sequence of tosses that contains **exactly** 3 tails?
2. How many distinct sequences can we make using 3 letter "A"s and 5 letter "B"s? (AAABBBBB, AABABBBB, etc.)
3. Given a full set of cards (52 cards). We draw 5 random cards from the set. In how many ways can we draw a *straight flush* (10, 9, 8, 7, 6 of the **same** suit)
4. There are 10 people working at the department of mathematics. In how many ways can we choose a reexamination committee that consists of a committee chair and 3 members?
5. How many different words can be composed from the letters of the word MATHEMATICS, the word MISSISSIPPI?

1. A coin is tossed 8 times. In how many ways can we get a sequence of tosses that contains **exactly** 3 tails?

$C_n^k = \frac{n!}{(n-k)!\,k!}$ $C_8^3 = \frac{8!}{(5)!\,3!} = \frac{8 \times 7 \times \cancel{6}}{\cancel{3 \times 2 \times 1}} = 8 \times 7 = 56$

2. How many distinct sequences can we make using 3 letter "A"s and 5 letter "B"s? (AAABBBBB, AABABBBB, etc.)

$C_n^k = \frac{n!}{(n-k)!\,k!}$ $C_8^3 = \frac{8!}{(5)!\,3!} = \frac{8 \times 7 \times \cancel{6}}{\cancel{3 \times 2 \times 1}} = 8 \times 7 = 56$

3. Given a full set of cards (52 cards). We draw 5 random cards from the set. In how many ways can we draw a *straight flush* (10, 9, 8, 7, 6 of the **same** suit)

4 $5! \, 4$ ← correct

4. There are 10 people working at the department of mathematics. In how many ways can we choose a reexamination committee that consists of a committee chair and 3 members?

10 for chair

For members

C_n^k $C_{10}^3 = \frac{10!}{(7)!\,3!} = \frac{10 \times \cancel{9}^{3} \times \cancel{8}^{4}}{\cancel{3} \times \cancel{2} \times 1} = 120$

$120 \times 10 = 1200$

5. How many different words can be composed from the letters of the word MATHEMATICS, the word MISSISSIPPI?

2M
2A
2T

4S
4I
2P

$\frac{11!}{4!\,4!\,2!}$

$\{T\ T\ T\ T\ T\ T\ T\ T\ T\ T\}$

4. Let there be 3 red and 5 green balls in a basket. We randomly take 3 balls from the basket. What is the probability that get exactly 1 red and 2 green balls?

$$C_3^1 = \frac{3}{2!\,1!} = 3 \qquad C_8^2 = \frac{8!}{(3!)\,2!} = \frac{8 \times \cancel{4}^{2}}{\cancel{2}} \; 10$$

$$\frac{\cancel{30}}{\cancel{56}} = \frac{15}{28}$$

3. What is the probability to win the game 5/35? In this game, one should guess 5 numbers out of 35.

$$C_{35}^5 = \frac{35!}{(30!)\,5!} = \frac{\cancel{35}^{7} \times 34 \times \cancel{33}^{11} \times \cancel{32}^{16^4} \times 31}{\cancel{5} \times \cancel{4} \times \cancel{3} \times \cancel{2} \times 1}$$

$$34 \times 11 \times$$

$$\frac{120}{4 \times}$$

Homework #7

Dear students, congratulations on Latvia’s birthday!

Let’s start with rather elementary problems to recall the material:

1. A coin is thrown 5 times. What is the probability that at least one head is obtained?
2. Find the probability of getting a numbered card when a card is drawn from the pack of 52 cards.
3. There are 5 green 7 red balls. Two balls are selected one by one without replacement (i.e., the balls are not returned to the bin). Find the probability that first is green and second is red.
4. What is the probability of getting a sum of 8 when two dice are thrown?
5. Two cards are drawn from the pack of 52 cards. Find the probability that both are diamonds or both are kings.
6. Three dice are rolled together. What is the probability as getting at least one **4**?
7. Find the probability of getting at most two heads when five coins are tossed.
8. What is the probability of getting a sum of 22 or more when four dice are thrown?
9. From a pack of cards, three cards are drawn at random. Find the probability that each card is from different suit.
10. Fifteen people sit around a circular table. What is the probability that two particular people are sitting together?

You might also think about the following problem that appears very difficult first, but actually has a simple solution (although it requires some computation):

- Given a group of 20 people. What is the probability that there are two people, who share the same birthday?

1. A coin is thrown 5 times. What is the probability that at least one head is obtained?

All tails subtracting from 1

$\left(\frac{1}{2}\right)^5 = \frac{1}{32}$ $\quad 1 - \frac{1}{32} = 31/32$

2. Find the probability of getting a numbered card when a card is drawn from the pack of 52 cards.

9 # cards
4 suites

$\frac{\cancel{36}\ \cancel{18}\ 9}{\cancel{52}\ \cancel{26}\ 13} = 9/13$

3. There are 5 green 7 red balls. Two balls are selected one by one without replacement (i.e., the balls are not returned to the bin). Find the probability that first is green and second is red.

$\frac{5}{12} \cdot \frac{7}{11} = \frac{35}{132}$

4. What is the probability of getting a sum of 8 when two dice are thrown?

(2, 6)
(6, 2)
3, 5
5, 3
4, 4

5/36

5. Two cards are drawn from the pack of 52 cards. Find the probability that both are diamonds or both are kings.

$\Diamond = \frac{13}{52} \cdot \frac{12}{51}$
$+$
$K = \frac{4}{52} \cdot \frac{3}{51}$
$-$
$K\Diamond = \frac{1}{52} \cdot \frac{1}{51}$

$\frac{167}{2652}$

6. Three dice are rolled together. What is the probability as getting at least one **4**?

not roll a 4

$\frac{5}{6} \cdot \frac{5}{6} \cdot \frac{5}{6} = \frac{125}{216}$

$1 - \frac{125}{216}$

$\frac{216}{216} - \frac{125}{216} = \frac{91}{216}$

7. Find the probability of getting at most two heads when five coins are tossed.

$\left(\frac{1}{2}\right)^4$

8. What is the probability of getting a sum of 22 or more when four dice are thrown?

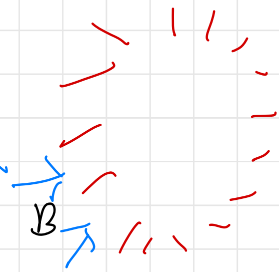
9. From a pack of cards, three cards are drawn at random. Find the probability that each card is from different suit.

$\frac{13^3 \, C^3_4}{C^3_{52}}$

10. Fifteen people sit around a circular table. What is the probability that two particular people are sitting together?

$$\frac{14!\cdot 2}{15!}$$

15! combos to sit



You might also think about the following problem that appears very difficult first, but actually has a simple solution (although it requires some computation):

- Given a group of 20 people. What is the probability that there are two people, who share the same birthday?

Approach from opposite

365 days

$$\frac{365\cdot 364\cdots 346}{365^{20}} = \frac{1}{2}$$

$$1-\frac{1}{2}=\boxed{\frac{1}{2}}$$

Probability distribution function

aka (mass function)

1 2 3 4 5 6

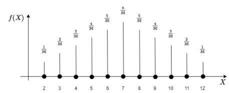


Рис. 1: The graph of a probability mass function.

$\frac{1}{36}$ $\frac{2}{36}$ $\frac{3}{36}$ $\frac{4}{36}$ $\frac{5}{36}$ $\frac{6}{36}$ $\frac{5}{36}$ $\frac{4}{36}$ $\frac{3}{36}$ $\frac{2}{36}$ $\frac{1}{36}$

2 3 4 5 6 7 8 9 10 11 12

2. *Binomial* distribution. Assume that we repeat the same experiment n times. This experiment has only two outcomes[3], a "success" with the probability p, and a "failure" with the probability $q = 1 - p$. The binomial distribution describes the probability of getting exactly k positive outcomes in a series of n experiments:

$$f_{\text{bin}}(k) = \begin{cases} C_n^k \, p^k \, q^{n-k}, & k \in \{0, 1, \ldots, n\}, \\ 0, & \text{otherwise.} \end{cases}$$

Success

Fail

q^4 $4pq^3$ $6p^2q^2$ $4q^3p$ p^4

0 1 2 3 4

$P(H = S) = p$

$P(T = F) = q$

$P(TTTT) = q^4$

$P(THTT + HTTT + TTHT + TTTH) = 4pq^3$

X - random variable

mean : f(x) - p.df

variance

E[X] = $\bar{X}$

(expectation)

Mean. Let the random variable X takes the values $x_1, \ldots, x_k$ with the probabilities $f(x_1), \ldots, f(x_k)$. The *mean value* or the *expectation* of X is denoted by $\bar{X}$ or $\mathrm{E}[X]$ and defined as follows:

$$\bar{X} = x_1 f(x_1) + \cdots + x_k f(x_k) = \sum_{i=1}^{k} x_i f(x_i).$$

One can easily check the linearity properties of the mean:

$$\mathrm{E}[X+Y] = \mathrm{E}[X] + \mathrm{E}[Y],$$

$$\mathrm{E}[aX] = a\mathrm{E}[X], \quad a \in \mathbb{R}.$$

$$E[X] = \frac{1}{6}(1 + \cdots 6) = \frac{1}{6} 21 = \frac{7}{2}$$

Variance. The *variance* describes how far the realisations of X are spread out from its average value. The variance is denoted by $Var(X)$ and defined as

$$Var(X) = \mathrm{E}\left[(X^o)^2\right] = \mathrm{E}\left[(X - \bar{X})^2\right] = \sum_{i=1}^{k} (x_i - \bar{X})^2 f(x_i).$$

Know how to compute mean: variance

Bayes Theorem

$P(A \wedge B) = P(A)P(B)$ if A & B are independent

$\wedge$ will be dropped

Conditional Probability

$$P(A|B) = \frac{P(AB)}{P(B)}$$

1 A and B are inp.

$$P(A|B) = \frac{P(AB)}{P(B)} = \frac{PA\,P(B)}{P(B)} = P(A)$$

$P(A|B)P(B) = P(AB)$

$P(B|A)PA = P(BA)$

$P(A|B)P(B) = P(B|A)) \cdot P(A)$

$$P(A|B) = \frac{P(B|A) \cdot P(A)}{P(B)}$$

1. **[The law of total probability].** For any two events $A, B \in \mathcal{A}$ it holds that

$$P(A) = P(A \wedge B) + P(A \wedge \bar{B}).$$

This law can be generalized as follows. Let $B_1, B_2 \ldots, B_k$ be a complete set (of mutually exclusive events satisfying $\bigcup B_i = \Omega$). Then we have

$$P(A) = P(A \wedge B_1) + P(A \wedge B_2) + \cdots + P(A \wedge B_k).$$

1. Consider a test to detect a disease that 0.1% of the population have. The test is 99% effective in detecting an infected person. However, the test gives a false positive result in 0.5% of cases. If a person tests positive for the disease what is the probability that they actually have it?

.1%
99%
.5%

condition

B - test positive
A - person is inf

$P(A) = .001$

$P(B|A) = .99$

$P(B|\bar{A}) = .005$

$P(\bar{A}) = .999$

$$\frac{P(B|A) \cdot P(A)}{P(B)} = \frac{.99 * .001}{P(BA) + P(B\bar{A})} = .1654$$

$$P(B|A) \cdot P(A) + P(B|\bar{A}) P(\bar{A})$$

$$.99 * .001 + .005 * .99$$

4. Consider 3 coins where two are fair, yielding heads with probability 0.50, while the third yields heads with probability 0.75. If one randomly selects one of the coins and tosses it 3 times, yielding 3 heads - what is the probability this is the biased coin?

A - probability coin is biased
B - HHH

$$P(A|B) = \frac{P(B|A)\, P(A)}{P(B)}$$

$P(A) = 1/3$

$P(B|A) = .75^3$

$P(B|\bar{A}) = .5^3$

Homework #8

The problems, listed below, are similar to those that will appear in the second test. However, not all of these problems will appear in the test.

1. We toss 3 dice. What is the probability that the total number of dots equals 11?
2. Again, assume that we toss 3 dice. Consider 2 events: $\mathbf{E_1}$ = "*total number of dots is divisible by 2*" and $\mathbf{E_2}$ = "*total number of dots is greater than 10*". Are these two events
   a. mutually exclusive?
   b. independent?
3. Assume that we play the following game: two dice are tossed. If the number of dots (**= n**) is odd, we **get n** euro. If the total number of dots is even, we **pay n/2** euro. What is the expected profit of this game? (**Hint**: *write the probability distribution function for the gain and compute the expected value*).
4. The random variable X takes the values 1...6 with probabilities {0.1 0.2 0.2 0.2 0.2 0.1}, while the random variable Y takes the same values, but with probabilities {0.05 0.1 0.25 0.2 0.25 0.1 0.05}. Compare the means and the variances of these two variables. What can we conclude?
5. Next, there are 2 problems related to Bayes' theorem.
   When solving the second problem, recall that we can consider more that two events. That is to say, instead of A and $\bar{A}$, we can have A_1, A_2, and A_3 that form a complete set of events (i.e., $P(A_1)+P(A_2)+P(A_3) = 1$). In this case, the law of total probability writes as $P(B\ A_1)+P(B\ A_2)+P(B\ A_3) = P(B)$. However, the remaining formulas do not change.
   a. Given two boxes. The first one contains 3 green and 7 blue balls, while the second one contains 6 green and 4 blue balls. Somebody has drawn a blue ball. What is the probability that this ball was drawn from the 2^{nd} box?
   b. Company A supplies 20% of the computers sold and is late 20% of the time. Company B supplies 30% of the computers sold and is late 3% of the time. Company C supplies another 50% and is late 2.5% of the time. A computer arrives late -- what is the probability that it came from Company A?

1. We toss 3 dice. What is the probability that the total number of dots equals 11?

Permutation

(1, 5, 5) 3 (2, 3, 6) 6

(1, 6, 4) 6 (3, 4, 4) 3 = 27

(2, 4, 5) 6 (3, 3, 5) 3

$\frac{27}{6^3}$

2. Again, assume that we toss 3 dice. Consider 2 events: $\mathbf{E_1}$ = "*total number of dots is divisible by 2*" and $\mathbf{E_2}$ = "*total number of dots is greater than 10*". Are these two events
   a. mutually exclusive?
   b. independent?

(a) not mutually exclusive since rolling 12 will be both events

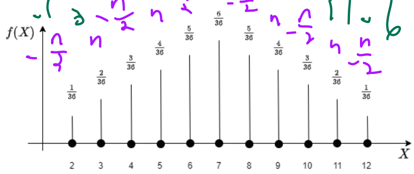
(b) independent

check

b. independent?

3. Assume that we play the following game: two dice are tossed. If the number of dots (= **n**) is odd, we **get n** euro. If the total number of dots is even, we **pay n/2** euro. What is the expected profit of this game? (**Hint**: *write the probability distribution function for the gain and compute the expected value*).

-1 3 -2 5 -3 7 -4 9 -5 11 -6



1.75

4. The random variable X takes the values 1...6 with probabilities {0.1 0.2 0.2 0.2 0.2 0.1}, while the random variable Y takes the same values, but with probabilities {0.05 0.1 0.25 0.2 0.25 0.1 0.05}. Compare the means and the variances of these two variables. What can we conclude?

$\bar{X} = \frac{1}{6} = .1\bar{6}$ $\bar{Y} = \frac{1}{7} = .1429$

$Var(X) = \frac{.013\bar{3}}{6}$ $Var(Y) = \frac{.047}{7}$

5. Next, there are 2 problems related to Bayes' theorem.
When solving the second problem, recall that we can consider more that two events. That is to say, instead of A and $\bar{A}$, we can have A_1, A_2, and A_3 that form a complete set of events (i.e., $P(A_1)+P(A_2)+P(A_3) = 1$). In this case, the law of total probability writes as $P(B A_1)+P(B A_2)+P(B A_3) = P(B)$. However, the remaining formulas do not change.
   a. Given two boxes. The first one contains 3 green and 7 blue balls, while the second one contains 6 green and 4 blue balls. Somebody has drawn a blue ball. What is the probability that this ball was drawn from the 2nd box?
   b. Company A supplies 20% of the computers sold and is late 20% of the time. Company B supplies 30% of the computers sold and is late 3% of the time. Company C supplies another 50% and is late 2.5% of the time. A computer arrives late -- what is the probability that it came from Company A?

A = Drawn from second box
B = blue ball ← condition given

$$P(A \mid B) = \frac{P(B \mid A)\,P(A)}{P(B)}$$

$P(B) = 11/20$

$P(A) = \frac{1}{2}$

$P(B|A) = 4/10$

$$\frac{\frac{\cancel{4}^{2}}{10} \cdot \frac{1}{\cancel{2}}}{11/20},\quad \frac{1}{\cancel{5}} \cdot \frac{\cancel{20}^{4}}{11}$$

$$\boxed{4/11}$$

✓

b. Company A supplies 20% of the computers sold and is late 20% of the time. Company B supplies 30% of the computers sold and is late 3% of the time. Company C supplies another 50% and is late 2.5% of the time. A computer arrives late -- what is the probability that it came from Company A?

A : computer from Company A

B : computer arrive late

$$P(A|B) = \frac{P(B|A)\,P(A)}{P(B)}$$

$P(A) = \frac{1}{5}$

$P(B|A) = \frac{1}{5}$

$$P(B) = P(B \cdot A) + P(B \cdot A_B) + P(B \cdot A_C)$$

$$\frac{1}{5} \cdot \frac{1}{5} + \frac{3}{10} \cdot \frac{3}{100} + \frac{50}{100} \cdot \frac{2.5}{100}$$

$$\frac{\cancel{50}}{\cancel{100}} \cdot \frac{\cancel{5}}{\cancel{200}} \quad \frac{1}{20} \cdot \frac{1}{4}$$

$$\frac{1}{25} + \frac{9}{1000} + \frac{1}{80}$$

$$\frac{\frac{1}{25}}{\frac{123}{2000}} = \frac{80}{123}$$

65.04%

① independance, mutually exclusive

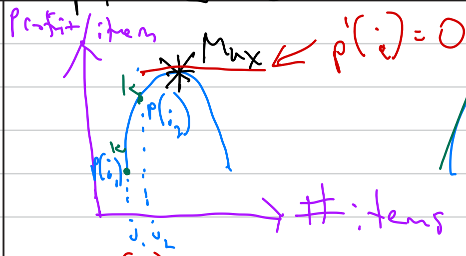
indep. $P(AB) = P(A)\,P(B)$

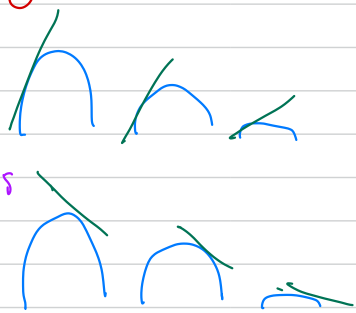
$P(A+B) = P(A) + P(B) - P(A \cdot B)$

mutually exclusive: $P(A \cup B) = \emptyset$

(2) computation for mean & var
bernoulli
(3) Bayes

Optemization





$p(i)$

$p(i) = k(i) + a$

slope aka derivative

$$k = \frac{\Delta p(i)}{\Delta i}$$

Deratives

Given a function $f(x)$, its derivative is denoted by $f'(x)$ or $\frac{d}{dx} f(x)$

Properties

1. $(af(x))' = a \cdot f'(x)$
2. $(f(x)+g(x))' = f'(x) + g'(x)$
3. $(f(x) \cdot g(x))' = f'(x)g(x) + g'(x)f(x)$
4. $(c)' = 0$
5. $(x^n)' = nx^{n-1}$
6. $(\sin\theta)' = \cos(\theta)$
7. $(\cos\theta)' = -\sin(\theta)$
8. $(e^x)' = e^x$

Polynomials

$p(x) = a_n x^n + a_{n-1} x^{n-1} \dots a_1 x + a_0$

$n = 2, 3$

parabola

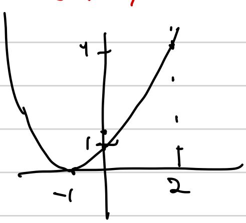
$p(x) = x^2 + a_1 x + a_0$

2nd degree poly

quadratic

a_n ∪ $-a_n$ ∩

$f(x) = x^2 + 2x + 1$



$f'(x) = (x^2)' + (2x)' + 1'$

$2x + 2 + 0$

$f'(x) = 0$

$0 = 2x + \cancel{2}$

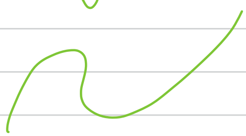
$-2 \qquad \cancel{-2}$

$\frac{-2}{2} = \frac{\cancel{2}x}{\cancel{2}}$

$-1 = x$

$(-1, 0)$

3rd degree polynomial



$f(x) = x^3$

$f'(x) = 3x^2$

$f'(0) = 0 \Rightarrow 3x^2 = 0 \Rightarrow x^* = 0$

$f'(-1) = 3, \; f'(1) = 3$

$f(x) = x^3 - 3x$

$f'(x) = 3x^2 - 3$

$f'(0) = -3$

$f'(2) = 9$

$f'(-2) = 9$

$0 = 3(x^2) - 3$

$0 = 3(x^2 - 1))$

$0 = 3(x+1)(x-1)$

$x+1 = 0 \qquad x-1 = 0$

$x = 1 \qquad x = -1$

min $\qquad$ max

$(1, -2) \qquad (-1, 2)$

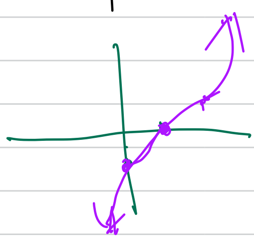
$f(x) = x^3 - 3x^2 + 3x - 1$

$f'(x) = 3x^2 - 6x + 3$

$0 = 3(x^2 - 2x + 1)$

$0 = (x-1)(x-1)$

$x = 1$



$1 - 3 + 3 - 1$

$f'(0) = 3$

$f'(2) = +$

$8 - 12 + 6 - 1$

$-4 + 6 - 1$

$-1 - 3 - 3 - 1$

$$f(x,y) = x^2 + 2y^2 + 3xy + 2x - 3y + 1$$

$f(x,y) = x^2 + y$

2 dimensions can have more min, max and saddle points

calculate partial derivative

$$\frac{\partial}{\partial x} f(x,y) = 2x + 3y + 2$$

$$\frac{\partial}{\partial y} f(x,y) = 4y + 3x - 3$$

set $\begin{cases} \frac{\partial}{\partial x} f = 0 \\ \frac{\partial}{\partial y} f = 0 \end{cases}$ $\qquad \begin{cases} 2x + 3y + 2 = 0 \\ 4y + 3x - 3 = 0 \end{cases}$

Reverse gaussian elim

$$\begin{cases} 2x + 3y = -2 \\ 3x + 4y = 3 \end{cases}$$

$$\left[\begin{array}{cc|c} 2 & 3 & -2 \\ 3 & 4 & 3 \end{array}\right] \rightarrow \left[\begin{array}{cc|c} 1 & \frac{3}{2} & -1 \\ 3 & 4 & 3 \end{array}\right] \xrightarrow{-3R1} \left[\begin{array}{cc|c} 1 & 3/2 & -1 \\ 0 & -1/2 & 6 \end{array}\right]$$

$\boxed{y = -12}$ $\boxed{x = 17}$

critical point

$$f(x,y) = x^2 - 2y^2 + 3xy - 4x + 2$$

$$\frac{\partial}{\partial x} f = 2x + 3y - 4$$

$$\frac{\partial}{\partial y} f = -4y + 3x + 0$$

$$\left[\begin{array}{cc|c} 2 & 3 & -4 \\ 3 & -4 & 0 \end{array}\right]$$

$$-3R_1 + R_2 \quad \left[\begin{array}{cc|c} 1 & 3/2 & -2 \\ 3 & -4 & 0 \end{array}\right]$$

$$\frac{9}{2} + \frac{8}{2}$$

$$\left[\begin{array}{cc|c} 1 & 3/2 & -2 \\ 0 & -17/2 & 6 \end{array}\right]$$

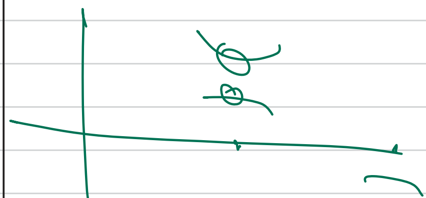
$$y = -\frac{12}{17}$$

$$\frac{18}{17} = -2 \qquad \frac{3}{2}\cdot\frac{12}{17}$$

$$\frac{18}{32}$$

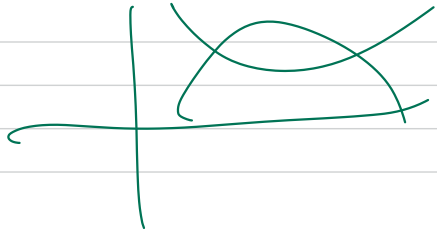
$$\boxed{x = \frac{16}{17}}$$

$$p(x) = a_n(x^n) + \dots + a_1 x + a_0$$



$$(x^*) \quad f(x^*)$$

f(x) has a max at x^* then $-f(x)$ has a minimum at x^*



Homework #9

Below, there is a sample test. In the "real" test, the formulations of the probability-related questions will slightly differ (**will be simpler**). Try to solve these problems **independently** to estimate your level. We need to identify the most difficult problems to give them more attention.

1. Consider the electrical circuit shown in Fig. 1. The resistors R_1, R_2, and R_3 fail **independently** with probabilities $p_1=0.05$, $p_2=0.2$, and $p_3=0.1$. Determine the probability that the lamp lights up when the power is turned on. For this, in a *parallel* connection, *at least one* of the parallel resistors must work, while in a *serial* connection, *all* components must work.
   (**Hint**: recall how we define the probability of the sum and the product of two events. OR corresponds to the sum, AND corresponds to the product).
2. Suppose that we toss 4 unbiased coins. Consider two events: E_1 = "Number of heads strictly exceeds the number of tails", and E_2 = "There are exactly 3 heads **in a row**". Are these two events
   1. Mutually exclusive?
   2. Independent?
3. Again, suppose that we toss 4 unbiased coins. The random variable X takes the value equal to the defference between the number of tails and the number of heads: X = #tails - #heads.
   1. Which values does this variable take?
   2. Compute the probability distribution for X.
4. Compute the expectation and variance for the probability distribution from Item 3.
5. An insurance company insured 2000 scooter drivers, 4000 car drivers, and 6000 truck drivers. The probability of an accident involving a scooter driver, car driver, and a truck is 0.01, 0.03, and 0.015 respectively. One of the insured persons meets with an accident. What is the probability that this person is a scooter driver?
6. Find all critical points of the following function: $f(x)=x^4+3x^3+x^2-18$. Determine their type (min/max/neigher).
7. Find the critical point of the folowing function of 3 variables:
   $f(x,y,z)=x^2+y^2+z^2+xy+yz-3x+2z$.
8. *Based upon the material that will be covered in the coming lecture.*

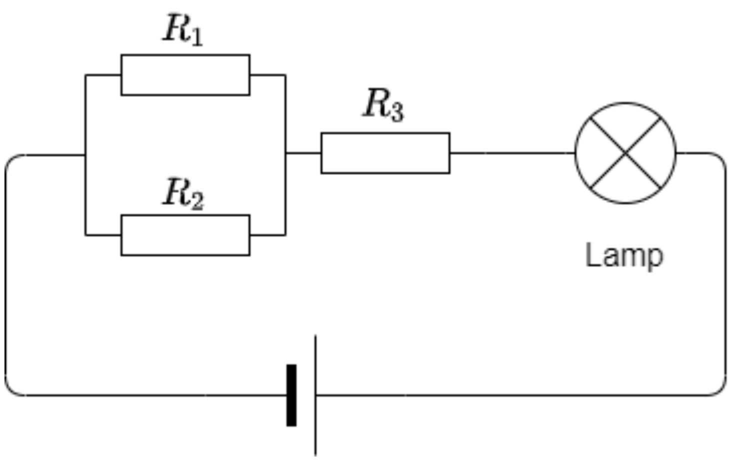


Figure 1. Aschematic picture of an electrical circuit.

1. Consider the electrical circuit shown in Fig. 1. The resistors R_1, R_2, and R_3 fail **independently** with probabilities $p_1=0.05$, $p_2=0.2$, and $p_3=0.1$. Determine the probability that the lamp lights up when the power is turned on. For this, in a *parallel* connection, *at least one* of the parallel resistors must work, while in a *serial* connection, *all* components must work.
(**Hint**: recall how we define the probability of the sum and the product of two events. OR corresponds to the sum, AND corresponds to the product).

$R1+R2$ $\quad$ $R3$

$.25$ $\quad$ $.1$

$\frac{1}{4} \times \frac{1}{10} = \frac{1}{40}$ chance of not turning on

$1 - \frac{1}{40} = \frac{39}{40}$

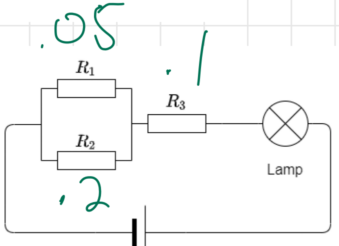


Figure 1. Aschematic picture of an electrical circuit.

2. Suppose that we toss 4 unbiased coins. Consider two events: E_1 = "Number of heads strictly exceeds the number of tails", and E_2 = "There are exactly 3 heads **in a row**". Are these two events
   1. Mutually exclusive?
   2. Independent?

① HHHT is part of both
Not mutually exclusive

② (H,H,H,H)
(H,H,H,T)×4

(HHH,T)
(T,H,H,H)

$P(E_1) = \frac{5}{16}$ $\quad$ $P(E_2) = \frac{2}{16}$

$P(E_1 \cap E_2) = P(E_1)P(E_2)$

$\frac{2}{16} \neq \frac{10}{256} = \frac{5}{128}$

Not independant

$\boxed{P(A \cap B) \overset{?}{=} P(A)P(B)}$

Check independance

3. Again, suppose that we toss 4 unbiased coins. The random variable X takes the value equal to the defference between the number of tails and the number of heads: X = #tails - #heads.
   1. Which values does this variable take?
   2. Compute the probability distribution for X.

$\frac{1}{16}$	$\frac{4}{16}$	$\frac{6}{16}$	$\frac{4}{16}$	$\frac{1}{16}$	← probability distribution
-4	-2	0	2	4	← values
HHHH	HHHT	HHTT	HTTT	TTTT	

4. Compute the expectation and variance for the probability distribution from Item 3.

$$E(x) = (-4)\frac{1}{16} + (-2)\frac{1}{4} + (0) + 2\left(\frac{1}{4}\right) + 4\frac{1}{16}$$

$$-\frac{1}{4} - \frac{2}{4} + \frac{2}{4} + \frac{1}{4} = \boxed{0}$$

$$Var(X) = \frac{16}{16} + \frac{4}{4} + 0 + \frac{4}{4} + \frac{16}{16} = \boxed{4}$$

5. An insurance company insured 2000 scooter drivers, 4000 car drivers, and 6000 truck drivers. The probability of an accident involving a scooter driver, car driver, and a truck is 0.01, 0.03, and 0.015 respectively. One of the insured persons meets with an accident. What is the probability that this person is a scooter driver?

A: is a insured scooter driver

B: an insured person had an accident

$$P(A|B) = \frac{P(A)P(B|A)}{P(B)}$$

$$\frac{\frac{1}{6\cancel{00}}}{\frac{11.5}{\cancel{6}00}} \quad \boxed{\frac{1}{11.5}}$$

$$P(A) = \frac{2\cancel{000}}{12\cancel{000}} = \frac{1}{6}$$

$$P(B|A) = \frac{1}{100}$$

$$P(B) = \frac{1}{6}\left(\frac{1}{100}\right) + \frac{1}{\cancel{3}}\left(\frac{\cancel{3}}{100}\right) + \frac{1}{2}\left(\frac{1.5}{100}\right)$$

$$\frac{1}{600} + \frac{1}{100} + \frac{1.5}{200}$$

$$\frac{1}{600} + \frac{6}{600} + \frac{4.5}{600} \quad \frac{11.5}{600}$$

5. An insurance company insured 2000 scooter drivers, 4000 car drivers, and 6000 truck drivers. The probability of an accident involving a scooter driver, car driver, and a truck is 0.01, 0.03, and 0.015 respectively. One of the insured persons meets with an accident. What is the probability that this person is a scooter driver?

6. Find all critical points of the following function: $f(x)=x^4+3x^3+x^2-18$. Determine their type (min/max/neigher).

$f'(x) = 4x^3+9x^2+2x$

$f'(0) = x(4x^2+9x+2)$

$x(4x+1)(x+2)$

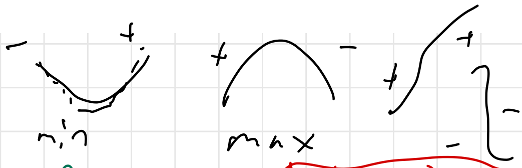
$x=0$ min, $x=-\frac{1}{4}$ max, $x=-2$ min

$f'(1) = +$

$f\left(-\frac{1}{10}\right) = -\frac{4}{1000} + \frac{90}{1000} - \frac{200}{1000}$ $-$

$f(-1) = -4+9-2$ $+$

$f(-3) = -27(4)+81-6$ $-$



$f(0) = -18$

$f\left(-\frac{1}{4}\right) = -17.98$

$-\frac{4603}{256}$

$f(-2) = 16-24+4-18$

$16-38$

-22

$$\boxed{\begin{array}{l}(0,-18) \text{ min} \\ \left(-\frac{1}{4}, -17.98\right) \text{ max} \\ (-2,-22) \text{ min}\end{array}}$$

7. Find the critical point of the folowing function of 3 variables: $f(x,y,z)=x^2+y^2+z^2+xy+yz-3x+2z$.

$\frac{\partial}{\partial x} f(\,) = 2x+y-3$

$\frac{\partial}{\partial y} f(\,) = 2y+x+z$

$\frac{\partial}{\partial z} f(\,) = 2z+y+2$

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

$$\left[\begin{array}{ccc|c} 1 & 2 & 1 & 0 \\ 2 & 1 & 0 & 3 \\ 0 & 1 & 2 & -2 \end{array}\right] \xRightarrow{-2R2} \left[\begin{array}{ccc|c} 1 & 2 & 1 & 0 \\ 0 & -3 & -2 & 3 \\ 0 & 1 & 2 & -2 \end{array}\right] \xRightarrow{3R3+R2} \left[\begin{array}{ccc|c} 1 & 2 & 1 & 0 \\ 0 & 1 & 2 & -2 \\ 0 & 0 & 4 & -3 \end{array}\right]$$

$z = -\frac{3}{4}$

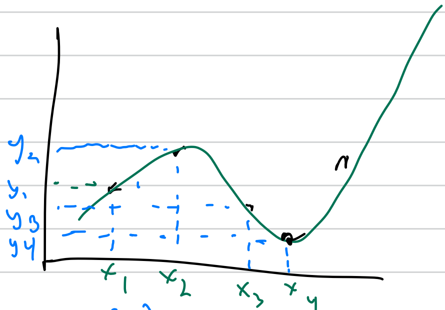
$-2 = y - \frac{3}{2}$

$-\frac{1}{2} = y$

$x - 1 - \frac{3}{4} = 0$ $\quad x = \frac{7}{4}$

critical point at $\left(\frac{7}{4}, -\frac{1}{2}, -\frac{3}{4}\right)$

Writing polynomial from given points



$$p(x) = a_3x^3 + a_2x^2 + a_1x + a_0$$

$$\begin{cases} a_3x_1^3 + a_2x_1^2 + a_1x_1 + a_0 = y_1 \\ a_3x_2^3 + a_2x_2^2 + a_1x_2 + a_0 = y_2 \\ a_3x_3^3 + a_2x_3^2 + a_1x_3 + a_0 = y_3 \\ a_3x_4^3 + a_2x_4^2 + a_1x_4 + a_0 = y_4 \end{cases}$$

4 points for 3 variables

x	y
-2	-1
-1	1
0	2
1	-1

$$\begin{cases} -8a_3 + 4a_2 - 2a_1 + a_0 = -1 \\ -a_3 + a_2 - a_1 + a_0 = 1 \\ a_0 = 2 \\ a_3 + a_2 + a_1 + a_0 = -1 \end{cases}$$

$a_3 - 2 - \frac{1}{2} = -3$

$a_3 - \frac{5}{2} = -3$

$$\left[\begin{array}{ccc|c} -8 & 4 & -2 & -3 \\ -1 & 1 & -1 & -1 \\ 1 & 1 & 1 & -3 \end{array}\right] \quad \left[\begin{array}{ccc|c} 1 & 1 & 1 & -3 \\ 0 & 2 & 0 & -4 \\ 0 & 12 & 6 & -27 \end{array}\right] \quad \left[\begin{array}{ccc|c} 1 & 1 & 1 & -3 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 6 & 3 \end{array}\right]$$

$a_1 = -\frac{1}{2}$

$a_2 = -2$

$a_3 = \frac{1}{2}$

$$\boxed{f(x) = \frac{1}{2}x^3 - 2x^2 - \frac{1}{2}x + 2}$$

Homework #10

As we discussed in the last class, first, there are several problems in discrete probability theory. These are aimed at letting you excersise in solving such problems, but they **will not be necessarily be given at the final test**. Some problems are simple. In this case, I do not show the right answer. For more complicated problems, there is an answer, so that you can check yourself.

1. A coint is tossed 5 times. What is the probability that at least one "Heads" will appear?
2. A random two-digit number is chosen. What is the probability that it will have different digits (e.g., 12, 36, 85, etc)?
3. There are 5 numbered[1] dice in the box. We pick the dice one by one, whithout returning them to the box and note the numbers shown on faces. What is the probability that the sequence of numbers will contain a **continuously growing sequence of 4 numbers**? (*Answer: p=1/72*)
4. Suppose that we toss 3 dice. What is the probability that we get a "6" and two **distinct** numbers, not equal to "6"? (*Answer: p=1/2*)
5. A box contains 10 pieces numbered from 1 to 10. We pick 6 pieces at random. What is the probability that this sample contains a) the piece #1; b) the pieces #1 and #2? (*Answer: p=3/5; 1/3*)
6. There are two smoke detection sensors. The first one recognizes 95% of cases, the second one – 90% of cases. What is the probability that there will be undetected smoke?
7. The probability that a randomly picked detail is broken is 0.1. What is the probability that of two randomly picked details **exactly one** is broken? (*Answer: p=0.18*)
8. How many times should one toss a die to ensure that the probability of **not having a single "6" is less than 0.25**? (*Answer: 8 times*)
9. In a warehouse, there are 100 eco-lamps and 200 glow-lamps. Eco-lamps fail in 3% of cases, while glow-lamps fail in 9% of cases. Assume that we pick up a lamp at random. What is the probability that it will fail?
10. Three students solved a difficult problem. Two students solved it correctly. What is the probability that the first student solved the problem if the chances of students to solve it were 0.4, 0.3, and 0.5. (*Answer: p=20/29*)

Finally, one problem devoted to curve fitting as we discussed in the last class. You can easily invent similar problems, solve them and check the results by plotting the resulted polynomal with, e.g., GeoGebra and inspectine the graphic of the function at the given points.

Find the coefficients of a 3rd order polynomial **p(x)** if we know that it passes through the following points: **(-1,0), (0,2), (1,-2), (2,1)**.

Note that the polynomial interpolation should be considered only within the interval between the outmost left and the outmost right points, i.e., in our case in the interval **[-1,2]**.

1 There are numbers from 1 to 6 instead of dots.

1. A coint is tossed 5 times. What is the probability that at least one “Heads” will appear?

T T T T T

$\frac{1}{2}$ $\frac{1}{2}$ $\frac{1}{2}$ $\frac{1}{2}$ $\frac{1}{2}$ $\frac{1}{32}$ $\frac{31}{32}$

2. A random two-digit number is chosen. What is the probability that it will have different digits (e.g., 12, 36, 85, etc)?

11, 22, 33, 44, 55, 66, 77, 88, 99

$1-\frac{9}{99}=\frac{90}{99}=\frac{1}{11}$

3. There are 5 numbered[1] dice in the box. We pick the dice one by one, whithout returning them to the box and note the numbers shown on faces. What is the probability that the sequence of numbers will contain a **continuously growing sequence of 4 numbers**?

(*Answer: p=1/72*)

4. Suppose that we toss 3 dice. What is the probability that we get a “6” and two **distinct** numbers, not equal to “6”?

(*Answer: p=1/2*)

5. A box contains 10 pieces numbered from 1 to 10. We pick 6 pieces at random. What is the probability that this sample contains a) the piece #1; b) the pieces #1 and #2? (*Answer*: *p*=3/5; 1/3)

$$\frac{\cancel{9}}{10}\cdot\frac{\cancel{8}}{\cancel{9}}\cdot\frac{\cancel{7}}{\cancel{8}}\cdot\frac{\cancel{6}}{\cancel{7}}\cdot\frac{\cancel{5}}{\cancel{6}}\cdot\frac{4}{\cancel{8}}=\frac{3}{5}$$

6. There are two smoke detection sensors. The first one recognizes 95% of cases, the second one – 90% of cases. What is the probability that there will be undetected smoke?

$$\frac{1}{20}\cdot\frac{1}{10}=\frac{1}{200}$$

7. The probability that a randomly picked detail is broken is 0.1. What is the probability that of two randomly picked details **exactly one** is broken? (*Answer*: *p*=0.18)

8. How many times should one toss a die to ensure that the probability of **not having a single "6" is less than 0.25**? (*Answer*: *8 times*)

9. In a warehouse, there are 100 eco-lamps and 200 glow-lamps. Eco-lamps fail in 3% of cases, while glow-lamps fail in 9% of cases. Assume that we pick up a lamp at random. What is the probability that it will fail?

$$\frac{1}{3}\cdot\frac{3}{100}+\frac{2}{3}\cdot\frac{9}{100}$$

$$\frac{3}{300}+\frac{18}{300}$$

$$\frac{21}{300}=\frac{7}{300}$$

10. Three students solved a difficult problem. Two students solved it correctly. What is the probability that the first student solved the problem if the chances of students to solve it were 0.4, 0.3, and 0.5. (*Answer: p=20/29*)

A: first student solved problem

B: 2 students solved correct

$$P(A|B) = \frac{P(B|A)\,P(A)}{P(B)}$$

Addition Rule (OR)

$P(A \text{ or } B) = P(A \cup B) = P(A) + P(B) - P(A \text{ and } B)$

Mutually exclusive: $P(A \text{ or } B) = P(A) + P(B)$

Multiplication Rule (AND)

$P(A \text{ and } B) = P(B|A) \cdot P(A)$

$P(B \text{ and } A) = P(A|B) \cdot P(B)$

Independence

If independent

$P(A \text{ and } B) = P(A) \times P(B)$

Mean. Let the random variable X takes the values $x_1, \ldots, x_k$ with the probabilities $f(x_1), \ldots, f(x_k)$. The *mean value* or the *expectation* of X is denoted by $\bar{X}$ or $\mathrm{E}[X]$ and defined as follows:

$$\bar{X} = x_1 f(x_1) + \cdots + x_k f(x_k) = \sum_{i=1}^{k} x_i f(x_i).$$

One can easily check the linearity properties of the mean:

$$\mathrm{E}[X+Y] = \mathrm{E}[X] + \mathrm{E}[Y],$$

$$\mathrm{E}[aX] = a\mathrm{E}[X], \quad a \in \mathbb{R}.$$

Variance. The *variance* describes how far the realisations of X are spread out from its average value. The variance is denoted by $Var(X)$ and defined as

$$Var(X) = \mathrm{E}\left[(X^o)^2\right] = \mathrm{E}\left[(X - \bar{X})^2\right] = \sum_{i=1}^{k}(x_i - \bar{X})^2 f(x_i).$$

Bayes Theorem

$$p(A|B) = \frac{p(B|A)\,p(A)}{p(B)}$$

A: Man has cancer

B: Tested positive

$$p(B) = p(A)p(B|A) + p(\bar{A})p(B|\bar{A})$$

$p(A) = .12$

$p(B|A) = .95$

$p(B|\bar{A}) = .06$

$$p(A|B) = \frac{.06 \cdot .12}{.12(.95) + (.06)(.88)}$$

Suppose a balanced six-sided die is about to be rolled, and we define the following events:

$A = \{1, 2\}$

$C = \{2, 4, 6\}$

$p(A) = \frac{2}{6} = \frac{1}{3}$

$p(C) = \frac{3}{6} = \frac{1}{2}$

$p(A \cap C) = \frac{1}{6}$

$p(A)p(C) = \frac{1}{6}$ ✓