
The Notebook
of
MATHEMATIC

HUY BUI



THE PUBLISHER

Contents

1	Matrices	2
1.1	Vector space	2
1.2	Definition	3
1.3	Basic operations	4
1.3.1	Matrix addition	4
1.3.2	Scalar multiplication	4
1.3.3	Matrix multiplication	4
1.3.4	Transpose	5
1.4	Square matrices	6
1.4.1	Definition	6
1.4.2	Unit matrix	6
1.4.3	Invertible matrix	7
1.4.4	Determinants	7
1.4.5	Eigenvectors and eigenvalues	9
1.5	Vector operations	10
1.5.1	Inner product	10
1.5.2	Cross product	11
1.5.3	Inner product and cross product	12
1.6	Matrix applications	12
1.6.1	Orthogonal projection	12
1.6.2	Rotation	12
1.7	Exercises	14
2	Linear equations	16
2.1	Linear independent set	16
2.2	System of linear equations	17
2.2.1	Inverse matrix	18
2.2.2	Crammer's rule	18
2.2.3	Gaussian elimination	19
2.3	Linear least squares	20
2.4	Exercise	21

Chapter 1

Matrices

1.1 Vector space

A vector space is a collection of objects called vectors, which may be added together and multiplied (“scaled”) by numbers, called scalars.

Let F be the field¹ and X be abelian group². The pair (X, \odot) is a vector space if $\odot : F \times X \rightarrow X$ satisfies $\forall s, t \in F$ and $\forall u, v \in X$:

$$\begin{cases} s \odot (t \odot u) &= (s \cdot t) \odot u \\ s \odot (u + v) &= (s \odot u) + (s \odot v) \odot u \\ (s + t) \odot u &= (s + u) \odot (s + v) \\ 1 \odot u &= u \end{cases}$$

The elements of F (meaning s, t) are called “scalars”. The elements of X (meaning u, v) are called “vectors”. The habit to denote vectors in this notebook, say $u \in X$, by \mathbf{u} .

In this notebook, $F = \mathbb{R}$, thus we only discuss vector spaces. Additionally, we omit the symbol \odot , thus $s \odot \mathbf{u}$ is writing $s\mathbf{u}$.

Example 1.1. \mathbb{R} is a vector space.

Example 1.2. \mathbb{R}^2 is also a vector space, where for $\mathbf{u} = (u_1, u_2)$ and $\mathbf{v} = (v_1, v_2)$ and $s \in \mathbb{R}$:

$$\begin{cases} s\mathbf{u} &= (su_1, su_2) \\ \mathbf{u} + \mathbf{v} &= (u_1 + u_2, u_2 + v_2) \\ \mathbf{0} &= (0, 0) \end{cases}$$

¹a field is a set on which addition, subtraction, multiplication, and division are defined, and behave as when they are applied to rational and real numbers.

²an abelian group, also called a commutative group, is a group in which the result of applying the group operation to two group elements does not depend on the order in which they are written.

Example 1.3. Any \mathbb{R}^n is a vector space, where for $\mathbf{u} = (u_1, u_2, \dots, u_n)$ and $\mathbf{v} = (v_1, v_2, \dots, v_n)$ and $s \in \mathbb{R}$ satisfies:

$$\begin{cases} s\mathbf{u} &= (su_1, su_2, \dots, su_n) \\ \mathbf{u} + \mathbf{v} &= (u_1 + v_1, u_2 + v_2, \dots, u_n + v_n) \\ \mathbf{0} &= (\underbrace{0, 0, \dots, 0}_{n \text{ times}}) \end{cases}$$

Note! If $\mathbf{u} \in \mathbb{R}^2$ and $\mathbf{v} \in \mathbb{R}^3$ then $\mathbf{u} + \mathbf{v}$ is not defined because \mathbf{u} and \mathbf{v} do not belong to the same vector space.

1.2 Definition

Let $\mathbb{R}^{m \times n}$ be the set of all tables with m rows and n columns such that they consist of real numbers. For example, $\begin{bmatrix} 2 & 3 & -1 \\ 2 & 1 & 4 \end{bmatrix}$ is a table of 2×3 . Let define a table \mathbf{A} such that:

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}$$

Note that the element a_{ij} , called the ij -entry, appears in the i th row and the j th column. We simply denote $\mathbf{A} = [a_{ij}]$, where $i \in 1, 2, \dots, m$ and $j \in 1, 2, \dots, n$.

Now, $\mathbb{R}^{m \times n}$ is a vector space if $\mathbf{A} = [a_{ij}]$ and $\mathbf{B} = [b_{ij}]$ and $s \in \mathbb{R}$ satisfies:

$$\begin{cases} \mathbf{A} + \mathbf{B} &= [a_{ij} + b_{ij}] \\ s\mathbf{A} &= [sa_{ij}] \\ [\mathbf{0}] &\text{is the zero-element} \end{cases}$$

In this case, $[\mathbf{0}]$ is an $m \times n$ table whose entries are all zero. $\mathbf{A}, \mathbf{B}, [\mathbf{0}]$ are now called matrices. Notice that a vector can be regarded as a special type of matrix. A row vector is a $1 \times n$ matrix, and a column vector is a $m \times 1$ matrix.

Example 1.4. Vector $\mathbf{s} = (1, 2)$ can be written as a matrix $\begin{bmatrix} 1 & 2 \end{bmatrix}$ (row vector)

or $\begin{bmatrix} 1 \\ 2 \end{bmatrix}$ (column vector).

1.3 Basic operations

1.3.1 Matrix addition

Let \mathbf{A} and \mathbf{B} be two matrices of the same vector space. The sum of \mathbf{A} and \mathbf{B} , written $\mathbf{A} + \mathbf{B}$, is the matrix obtained by adding corresponding elements from \mathbf{A} and \mathbf{B} :

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} & \dots & b_{1n} \\ b_{21} & b_{22} & \dots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m1} & b_{m2} & \dots & b_{mn} \end{bmatrix} = \begin{bmatrix} a_{11} + b_{11} & a_{12} + b_{12} & \dots & a_{1n} + b_{1n} \\ a_{21} + b_{21} & a_{22} + b_{22} & \dots & a_{2n} + b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} + b_{m1} & a_{m2} + b_{m2} & \dots & a_{mn} + b_{mn} \end{bmatrix}$$

1.3.2 Scalar multiplication

The product of a scalar k and a matrix \mathbf{A} , written $k\mathbf{A}$ or $\mathbf{A}k$, is the matrix obtained by multiplying each element of \mathbf{A} by k :

$$k \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} = \begin{bmatrix} ka_{11} & ka_{12} & \dots & ka_{1n} \\ ka_{21} & ka_{22} & \dots & ka_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ ka_{m1} & ka_{m2} & \dots & ka_{mn} \end{bmatrix}$$

We also define $-\mathbf{A} = (-1)\mathbf{A}$ and $\mathbf{A} - \mathbf{B} = \mathbf{A} + (-\mathbf{B})$. The matrix $-\mathbf{A}$ is called the *negative* of the matrix \mathbf{A} .

1.3.3 Matrix multiplication

Now suppose \mathbf{A} and \mathbf{B} are two matrices such that the number of columns of \mathbf{A} is equal to the number of rows of \mathbf{B} , say \mathbf{A} is an $m \times p$ matrix and \mathbf{B} is an $p \times n$. Then the product of \mathbf{A} and \mathbf{B} , written \mathbf{AB} , is the $m \times n$ matrix whose ij -entry is obtained by multiplying the elements of i th row of \mathbf{A} by the corresponding elements of the j th column of \mathbf{B} and then adding:

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1p} \\ \cdot & \cdot & \dots & \cdot \\ a_{i1} & a_{i2} & \dots & a_{ip} \\ \cdot & \cdot & \dots & \cdot \\ a_{m1} & a_{m2} & \dots & a_{mp} \end{bmatrix} \begin{bmatrix} b_{11} & \dots & b_{1j} & b_{1n} \\ \vdots & \vdots & \vdots & \vdots \\ b_{m1} & b_{m2} & \dots & b_{mn} \end{bmatrix} = \begin{bmatrix} c_{11} & \dots & c_{1n} \\ \vdots & \vdots & \vdots \\ \vdots & c_{ij} & \vdots \\ \vdots & \vdots & \vdots \\ c_{m1} & \dots & c_{mn} \end{bmatrix}$$

where

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \dots + a_{ip}b_{pj} = \sum_{k=1}^p a_{ik}b_{kj}$$

If the number of columns of \mathbf{A} is not equal to the number of rows of \mathbf{B} , then the product \mathbf{AB} is not defined.

Example 1.5. Find \mathbf{AB} , where

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

Since \mathbf{A} is a 2×2 matrix, \mathbf{B} is a 2×3 , the product matrix \mathbf{AB} is defined and is a 2×3 matrix. To obtain the elements in the first row of the product matrix \mathbf{AB} , multiply the first row $\begin{bmatrix} 1 & 3 \end{bmatrix}$ of \mathbf{A} by the columns:

$$\begin{bmatrix} 2 \\ 3 \end{bmatrix} \quad \begin{bmatrix} 0 \\ -2 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} -4 \\ 6 \end{bmatrix}$$

of \mathbf{B} respectively:

$$\begin{aligned} \begin{bmatrix} 1 & 3 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} 2 & 0 & -4 \\ 3 & -2 & 6 \end{bmatrix} &= \begin{bmatrix} 1 \cdot 2 + 3 \cdot 3 & 1 \cdot 0 + 3(-2) & 1(-4) + 3 \cdot 6 \\ 2 \cdot 2 + (-1) \cdot 3 & 2 \cdot 0 + (-1) \cdot (-2) & 2 \cdot (-4) + (-1) \cdot 6 \end{bmatrix} \\ &= \begin{bmatrix} 11 & -6 & 14 \\ 1 & 2 & -14 \end{bmatrix} \end{aligned}$$

To obtain the elements in the second row of the product matrix \mathbf{AB} , multiply the second row $\begin{bmatrix} 2 & -1 \end{bmatrix}$ of \mathbf{A} by the columns of \mathbf{B} respectively:

$$\begin{aligned} \begin{bmatrix} 1 & 3 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} 2 & 0 & -4 \\ 3 & -2 & 6 \end{bmatrix} &= \begin{bmatrix} 11 & -6 & 14 \\ 2 \cdot 2 + (-1) \cdot 3 & 2 \cdot 0 + (-1) \cdot (-2) & 2 \cdot (-4) + (-1) \cdot 6 \end{bmatrix} \\ &= \begin{bmatrix} 11 & -6 & 14 \\ 1 & 2 & -14 \end{bmatrix} \end{aligned}$$

1.3.4 Transpose

The transpose of a matrix is an operator which flips a matrix over its diagonal, that is, it switches the row and column indices of the matrix by producing another matrix.

Let $\mathbf{A} \in \mathbb{R}^{m \times n}$. The transpose of \mathbf{A} , denote \mathbf{A}^T , is a matrix $\mathbf{A}^T \in \mathbb{R}^{n \times m}$ such that for $\mathbf{A} = [a_{ij}]$, we have $\mathbf{A}^T = [a_{ji}]$.

Example 1.6. The transpose of $\mathbf{A} = \begin{bmatrix} 2 & 3 & 1 \\ 0 & 7 & 8 \end{bmatrix}$ is $\mathbf{A}^T = \begin{bmatrix} 2 & 0 \\ 3 & 7 \\ 1 & 8 \end{bmatrix}$.

For matrices \mathbf{A} and \mathbf{B} and scalar k , we have the following properties of transpose:

- The operation of taking the transpose is an involution (self-inverse), meaning that $(\mathbf{A}^T)^T = \mathbf{A}$.
- The transpose respects addition, $(\mathbf{A} + \mathbf{B})^T = \mathbf{A}^T + \mathbf{B}^T$.
- The order of the factors reverses when taking transpose of a product matrix, $\mathbf{AB}^T = \mathbf{B}^T \mathbf{A}^T$. From this, one can deduce that a square matrix \mathbf{A} is invertible if and only if its transpose \mathbf{A}^T is invertible, and in this case, we have $(\mathbf{A}^{-1})^T = (\mathbf{A}^T)^{-1}$.
- $(c\mathbf{A})^T = c\mathbf{A}^T$.

1.4 Square matrices

1.4.1 Definition

A matrix with the same number of rows as columns is called a *square matrix*. A square matrix with n rows and n columns is said to be of *order n* , and is called *n -square matrix*. The diagonal of an n -square matrix $\mathbf{A} = [a_{ij}]$ consists of elements $a_{11}, a_{22}, \dots, a_{nn}$.

Example 1.7. The matrix

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

is a square matrix of order 3. The numbers along the diagonal are 1, -4, and 2.

1.4.2 Unit matrix

The n -square matrix with 1s along the diagonal and 0s elsewhere, e.g.,

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

is called *unit matrix* and will be denoted by \mathbf{I} . The unit matrix \mathbf{I} plays the same role in matrix multiplication as the number 1 does in the usual multiplication of number. Specifically,

$$\mathbf{AI} = \mathbf{IA} = \mathbf{A}$$

for any square matrix \mathbf{A} .

We can form powers of a square matrix \mathbf{X} by defining:

$$\mathbf{X}^2 = \mathbf{XX}, \quad \mathbf{X}^3 = \mathbf{X}^2\mathbf{X} \quad \dots \quad \mathbf{X}^0 = \mathbf{I}$$

1.4.3 Invertible matrix

An n-square matrix \mathbf{A} is said to be invertible if there exists an n-square matrix \mathbf{B} with the property

$$\mathbf{AB} = \mathbf{BA} = \mathbf{I}_n$$

Such a matrix \mathbf{B} is unique, and is called the *inverse* of \mathbf{A} , and is denoted by \mathbf{A}^{-1} . Observe that \mathbf{B} is the inverse of \mathbf{A} if and only if \mathbf{A} is the inverse of \mathbf{B} .

Example 1.8. Suppose

$$\mathbf{A} = \begin{bmatrix} 2 & 5 \\ 1 & 3 \end{bmatrix} \quad \text{and} \quad \mathbf{B} = \begin{bmatrix} 3 & -5 \\ -1 & 2 \end{bmatrix}$$

Then

$$\begin{aligned} \mathbf{AB} &= \begin{bmatrix} 6-5 & -10+10 \\ 3-3 & -5+6 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \mathbf{I}_2 \\ \mathbf{BA} &= \begin{bmatrix} 6-5 & 15-15 \\ -2+2 & -5+6 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \mathbf{I}_2 \end{aligned}$$

Thus \mathbf{A} and \mathbf{B} are inverses.

1.4.4 Determinants

To each n-square matrix $\mathbf{A} = [a_{ij}]$ we assign a specific number called the *determinant* of \mathbf{A} , denoted $\det(\mathbf{A})$ or

$$\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}$$

The determinants of matrices of order one, two, three are defined as follows:

$$\begin{aligned} \begin{vmatrix} a_{11} \end{vmatrix} &= a_{11} \\ \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} &= a_{11}a_{22} - a_{12}a_{21} \\ \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} &= a_{11}a_{22}a_{33} + a_{21}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{12}a_{21}a_{33} - a_{11}a_{23}a_{32} \end{aligned}$$

Below is another way to calculate determinant of a 3×3 matrix:

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} a_{11} - \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} a_{12} + \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} a_{13} \quad (1.1)$$

In general, one can calculate determinant of an n -square matrix \mathbf{A} using the following formula:

$$\det(\mathbf{A}) = \sum_{j=1}^n (-1)^{i+j} a_{ij} \det(\mathbf{A}_{ij})$$

Using determinant, we can calculate the inverse of an invertible matrix. For example, the inverse of a 2×2 matrix with nonzero determinant can be calculated as follow:

$$\mathbf{A}^{-1} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{\det(\mathbf{A})} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

For n -square matrix $\mathbf{A} = [a_{ij}]$ with nonzero determinant, the inverse would be:

$$\mathbf{A}^{-1} = \frac{1}{\det(\mathbf{A})} \begin{vmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ c_{n1} & c_{n2} & \cdots & c_{nn} \end{vmatrix}$$

where $c_{ij} = (-1)^{i+j} \det(\mathbf{A}_{ij})$.

For matrices \mathbf{A}, \mathbf{B} and scalar $k \neq 0$, we have the following properties of determinants:

- $\det(\mathbf{AB}) = \det(\mathbf{A}) \cdot \det(\mathbf{B})$
- A *square* matrix is invertible if and only if it has a nonzero determinant.

- Swapping two rows of a matrix produces the negative matrix. If \mathbf{B} is a matrix \mathbf{A} with any two rows interchanged, then $\det(\mathbf{B}) = -\det(\mathbf{A})$.
- If \mathbf{B} is a matrix \mathbf{A} but one row multiplied by k then $\det(\mathbf{B}) = k \det(\mathbf{A})$.
- If \mathbf{B} is a matrix \mathbf{A} but row i th, say R_i , is replaced by $R_i + kR_t, \forall t \neq i$, then $\det(\mathbf{B}) = \det(\mathbf{A})$.

Example 1.9. Given matrix \mathbf{A} :

$$\mathbf{A} = \begin{bmatrix} -2 & -1 & 4 \\ 6 & -3 & -2 \\ 4 & 1 & 2 \end{bmatrix} \quad \det(\mathbf{A}) = 100$$

Interchange any two rows of \mathbf{A} , we have matrix \mathbf{B} :

$$\mathbf{B} = \begin{bmatrix} 6 & -3 & -2 \\ -2 & -1 & 4 \\ 4 & 1 & 2 \end{bmatrix} \quad \det(\mathbf{B}) = -\det(\mathbf{A}) = -100$$

Multiply first row of \mathbf{A} with 4, we have matrix \mathbf{B} :

$$\mathbf{B} = \begin{bmatrix} -8 & -4 & 16 \\ 6 & -3 & -2 \\ 4 & 1 & 2 \end{bmatrix} \quad \det(\mathbf{B}) = 4 \det(\mathbf{A}) = 400$$

Replace the first row R_1 by $R_1 + \frac{1}{2}R_3$

$$\mathbf{B} = \begin{bmatrix} 1 & -5/2 & 3 \\ 6 & -3 & -2 \\ 4 & 1 & 2 \end{bmatrix} \quad \det(\mathbf{B}) = \det(\mathbf{A}) = 100$$

1.4.5 Eigenvectors and eigenvalues

Let \mathbf{A} be an n -square matrix. The eigenvector of \mathbf{A} is a column vector \mathbf{x} that satisfies:

$$\mathbf{Ax} = \lambda \mathbf{x}$$

Let us start to find \mathbf{x} such that

$$\begin{aligned} \mathbf{Ax} &= \lambda \mathbf{x} \\ \Leftrightarrow \mathbf{Ax} - \lambda \mathbf{x} &= \mathbf{0} \\ \Leftrightarrow (\mathbf{A} - \lambda \mathbf{I}) \mathbf{x} &= \mathbf{0} \end{aligned}$$

where \mathbf{I} is the identity matrix living in the same vector space with \mathbf{A} . There are infinite number of solutions for \mathbf{x} if there exists λ that makes $\det(\mathbf{A} - \lambda\mathbf{I}) = 0$. In that case, λ is called eigenvalue associated with the eigenvector \mathbf{x} .

Note! There are no more than n distinct eigenvalues for an n -square matrix.

Example 1.10. To calculate the eigenvalues and eigenvectors of $\mathbf{A} = \begin{bmatrix} 2 & 1 \\ 1 & 4 \end{bmatrix}$,

we need to solve the following equation:

$$\begin{aligned} \det(\mathbf{A} - \lambda\mathbf{I}) = 0 &\Leftrightarrow \begin{vmatrix} 2-\lambda & 1 \\ 1 & 4-\lambda \end{vmatrix} = 0 \\ \Leftrightarrow (2-\lambda)(4-\lambda) = 0 &\Leftrightarrow \lambda^2 - 6\lambda + 7 = 0 \\ \Leftrightarrow \lambda_1 = 3 + \sqrt{2} \quad \text{and} \quad \lambda_2 = 3 - \sqrt{2} \end{aligned}$$

If the eigenvalue is $\lambda_1 = 3 + \sqrt{2}$, then

$$\begin{aligned} (\mathbf{A} - (3 + \sqrt{2})\mathbf{I})\mathbf{x} &= \mathbf{0} \\ \begin{bmatrix} -1 - \sqrt{2} & 1 \\ 1 & 1 - \sqrt{2} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} &= \mathbf{0} \\ \Leftrightarrow \begin{cases} (-1 - \sqrt{2})x_1 + x_2 = 0 \\ x_1 + (1 - \sqrt{2})x_2 = 0 \end{cases} & \\ \Leftrightarrow x_1 = (1 + \sqrt{2})x_2 & \end{aligned}$$

Eigenvectors corresponding to $\lambda_1 = 3 + \sqrt{2}$ are

$$\mathbf{x} = \begin{bmatrix} 1 \\ 1 + \sqrt{2} \end{bmatrix} k, \quad k \in \mathbb{R} \setminus \{0\}$$

Likewise, If the eigenvalue is $\lambda_1 = 3 - \sqrt{2}$, then the corresponding eigenvectors are

$$\mathbf{x} = \begin{bmatrix} 1 \\ 1 - \sqrt{2} \end{bmatrix} k, \quad k \in \mathbb{R} \setminus \{0\}$$

Note! The sum of diagonal members of the matrix equals to the sum of eigenvalues.

1.5 Vector operations

1.5.1 Inner product

Let X be a vector space. If one can define a binary operation $\cdot : X \times X \rightarrow \mathbb{R}$ (take 2 vectors and produce a scalar) such that $\forall \mathbf{w}, \mathbf{u}, \mathbf{v} \in X$ and $\forall s \in \mathbb{R}$, we

have:

$$\begin{cases} \mathbf{u} \cdot \mathbf{v} &= \mathbf{v} \cdot \mathbf{u} \\ \mathbf{u} \cdot (\mathbf{v} + \mathbf{w}) &= \mathbf{u} \cdot \mathbf{v} + \mathbf{u} \cdot \mathbf{w} \\ s(\mathbf{u} \cdot \mathbf{v}) &= (s\mathbf{u}) \cdot \mathbf{v} = \mathbf{u} \cdot (s\mathbf{v}) \\ \mathbf{u} \cdot \mathbf{u} &\geq 0 \\ \mathbf{u} \cdot \mathbf{u} = 0 &\Leftrightarrow \mathbf{u} = \mathbf{0} \end{cases}$$

then X is called inner product space. The binary operation \cdot is called inner product.

Inner product is used to calculate norm of vectors. For any vector $\mathbf{u} \in X$, one can define norm $\|\mathbf{u}\|$ by

$$\|\mathbf{u}\| = \sqrt{\mathbf{u} \cdot \mathbf{u}}$$

For any two vectors $\mathbf{u}, \mathbf{v} \in X$, we have

$$\mathbf{u} \cdot \mathbf{v} = \|\mathbf{u}\| \|\mathbf{v}\| \cos \alpha$$

A vector whose norm is 1, is called a unit vector. One can define a unit vector $\hat{\mathbf{x}}$ by the following formula:

$$\hat{\mathbf{x}} = \frac{1}{\|\mathbf{x}\|} \mathbf{x}$$

Example 1.11. Vector $\mathbf{x} = (-2, 5, 3, 6)$ has inner product norm of $\|\mathbf{x}\| = \sqrt{(-2)^2 + 5^2 + 3^2 + 6^2} = \sqrt{74}$. Thus,

$$\hat{\mathbf{x}} = \frac{1}{\|\mathbf{x}\|} \mathbf{x} = \frac{1}{\sqrt{74}} (-2, 5, 3, 6)$$

1.5.2 Cross product

Let \mathbb{R}^3 be a vector space. If one can define a binary operation $\times : X \times X \rightarrow X$ (take 2 vectors and produce a vector) such that $\forall \mathbf{a}, \mathbf{b} \in \mathbb{R}^3$, we have:

$$\mathbf{c} = \mathbf{a} \times \mathbf{b} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix} \quad (1.2)$$

where $\mathbf{i}, \mathbf{j}, \mathbf{k}$ are the base vectors, and $\mathbf{a} = (a_1, a_2, a_3)$, $\mathbf{b} = (b_1, b_2, b_3)$. Note that the determinant in (1.2) produces a vector instead of a scalar. It is also called *symbolic determinant*.

Example 1.12. Let $\mathbf{a} = (1, 0, 5)$, $\mathbf{b} = (-1, 2, 4)$. Applying the formula (1.1), the cross product $\mathbf{a} \times \mathbf{b}$ is

$$\begin{aligned} \mathbf{c} &= \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 1 & 0 & 5 \\ -1 & 2 & 4 \end{vmatrix} = \begin{vmatrix} 0 & 5 \\ 2 & 1 \end{vmatrix} \mathbf{i} - \begin{vmatrix} 1 & 5 \\ -1 & 1 \end{vmatrix} \mathbf{j} + \begin{vmatrix} 1 & 0 \\ -1 & 2 \end{vmatrix} \mathbf{k} \\ &= (-10)\mathbf{i} - 6\mathbf{j} + 2\mathbf{k} \end{aligned}$$

which means that $\mathbf{c} = (-10, -6, 2)$

If we draw \mathbf{a} , \mathbf{b} and $\mathbf{c} = \mathbf{a} \times \mathbf{b}$ in a right-handed three-dimensional coordinate system, then \mathbf{c} is perpendicular to both \mathbf{a} and \mathbf{b} . Moreover, if vectors \mathbf{a} and \mathbf{b} are considered as edges of a parallelogram then $A = \|\mathbf{a} \times \mathbf{b}\|$ is the area of that parallelogram.

For vectors \mathbf{a} , \mathbf{b} , \mathbf{c} and scalar $k \neq 0$, we have the following properties of cross products:

- $\|\mathbf{a} \times \mathbf{b}\| = \|\mathbf{a}\|\|\mathbf{b}\| \sin \alpha$, where $\alpha = \angle(\mathbf{a}, \mathbf{b})$.
- $\mathbf{a} \times \mathbf{b} = -(\mathbf{b} \times \mathbf{a})$
- $\mathbf{a} \times \mathbf{a} = \mathbf{0}$, where $\mathbf{0}$ is zero-vector (not a scalar).
- $(k\mathbf{a}) \times \mathbf{b} = \mathbf{a} \times (k\mathbf{b}) = k(\mathbf{a} \times \mathbf{b})$
- $\mathbf{c} \times (\mathbf{a} + \mathbf{b}) = \mathbf{c} \times \mathbf{a} + \mathbf{c} \times \mathbf{b}$
- Cross product is not associative, $\mathbf{c} \times (\mathbf{a} \times \mathbf{b})$ may not equal to $(\mathbf{c} \times \mathbf{a}) \times \mathbf{b}$.

1.5.3 Inner product and cross product

Given three vectors $\mathbf{a} = (a_1, a_2, a_3)$, $\mathbf{b} = (b_1, b_2, b_3)$ and $\mathbf{c} = (c_1, c_2, c_3)$. The combination of cross product and inner product as follow produces a scalar:

$$(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c} = \begin{bmatrix} c_1 & c_2 & c_3 \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{bmatrix} \quad (1.3)$$

1.6 Matrix applications

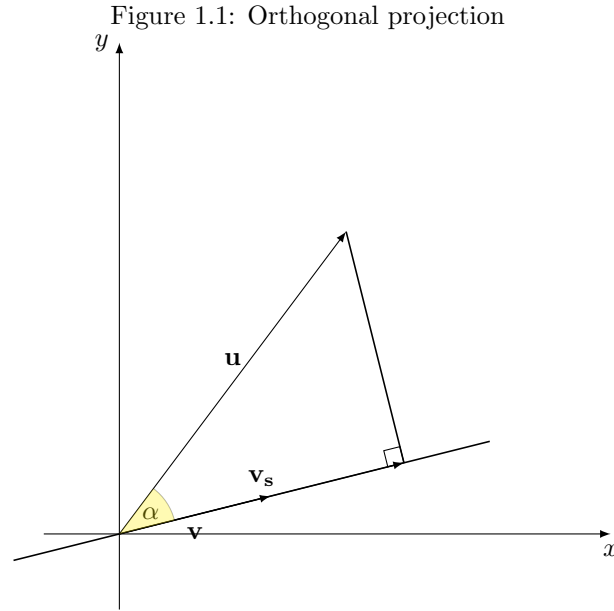
1.6.1 Orthogonal projection

The orthogonal projection of \mathbf{v} onto the line spanned by a nonzero \mathbf{s} (see figure 1.1) is:

$$\mathbf{v}_s = \frac{\mathbf{v} \cdot \mathbf{s}}{\|\mathbf{s}\|^2} \mathbf{s}$$

1.6.2 Rotation

In linear algebra, a rotation matrix is a matrix that is used to perform a rotation in Euclidean space.



In two dimensions

In two dimensions, every rotation matrix has the following form,

$$\mathbf{R} = \begin{bmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{bmatrix}$$

rotates points in the xy -plane counterclockwise through an angle α about the origin of the Cartesian coordinate system. The column vectors are calculated by means of the following matrix multiplication,

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

So the new coordinates (x', y') of the point (x, y) after rotation are

$$x' = x \cos \alpha - y \sin \alpha \quad y' = x \sin \alpha + y \cos \alpha .$$

In three dimensions

A basic rotation (also called elemental rotation) is a rotation about one of the axes of a Coordinate system. The following three basic rotation matrices rotate vectors by an angle θ about the x -, y -, or z -axis, in three dimensions, using the

right-hand rule:

$$R_x(\theta) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & -\sin \theta \\ 0 & \sin \theta & \cos \theta \end{bmatrix}$$

$$R_y(\theta) = \begin{bmatrix} \cos \theta & 0 & \sin \theta \\ 0 & 1 & 0 \\ -\sin \theta & 0 & \cos \theta \end{bmatrix}$$

$$R_z(\theta) = \begin{bmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

1.7 Exercises

Exercise 1.1. Consider an object living in 3D real place. Its current position is $P(2, 5, 3)$. The object is heading towards the point $Q(4, 10, 5)$, and moves to that direction exactly 3 length units. Calculate the position of this object after movement.

Proof. Let \mathbf{y} be the vector of the movement and \mathbf{x} be the distance between P and Q. We have:

$$\begin{aligned} \mathbf{x} &= \mathbf{Q} - \mathbf{P} = (4, 10, 5) - (2, 5, 3) = (2, 5, 2) \\ \|\mathbf{x}\| &= \sqrt{2^2 + 5^2 + 2^2} = \sqrt{33}, \quad \hat{\mathbf{x}} = \frac{1}{\sqrt{33}}(2, 5, 2) \\ \mathbf{y} &= \mathbf{P} + 3\hat{\mathbf{x}} = (2, 5, 3) + \frac{3}{\sqrt{33}}(2, 5, 2) \quad \square \end{aligned}$$

Exercise 1.2. Let $\mathbf{a} = (2, 4)$ and $\mathbf{b} = (1, -3)$ be vectors in \mathbb{R}^2 . If these vectors are considered as edges of a parallelogram then calculate its area A.

Proof. Let \mathbf{A} be the matrix formed by \mathbf{a} and \mathbf{b}

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

The area of the parallelogram is the absolute value of the determinant of \mathbf{A}

$$S = |\det(\mathbf{A})| = |2(-3) - 4 \cdot 1| = 10 \quad \square$$

Exercise 1.3. Let $\mathbf{p}_0 \in \mathbb{R}^2$ be such that $\mathbf{p}_0 = (5, 3)$ and the corresponding point in \mathbb{R}^2 is $P(5, 3)$. We define the function $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ such that $f(\mathbf{x}) = \mathbf{x} - \mathbf{p}_0$. Calculate $f((0, 0))$, $f((-5, 9))$, $f((\mathbf{p}_0))$. Define the meaning of this function.

Proof. This function moves the origin of the coordinate system from $O(0, 0)$ to $P(5, 3)$, and then calculates the new coordinate of vector \mathbf{x} . For example, the coordinate of $O(0, 0)$ in the new coordinate system is $f((0, 0)) = (-5, 3)$. Point $(-5, 9)$ now becomes $f((-5, 9)) = (-10, 6)$. And, of course, because $P(5, 3)$ becomes the origin of the coordinate system, its coordinate is $(0, 0)$. \square

Exercise 1.4. Suppose we have a cube in \mathbb{R}^3 such that the vertices are at the point $(0, 0, 0)$, $(0, 0, 1)$, $(1, 0, 0)$, $(1, 0, 1)$, $(1, 1, 0)$ and $(1, 1, 1)$. Determine a function $f : \mathbb{R}^{3 \times 8} \rightarrow \mathbb{R}^{3 \times 8}$ such that cube is centered at the origin of the coordinate system $O(0, 0, 0)$, thus the center of the cube is mapped to 0.

Proof. We write each point of the cube as a column in a single matrix \mathbf{C} :

$$\mathbf{C} = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \end{bmatrix}$$

The center of the cube is $C_0 = (1/2, 1/2, 1/2)$. The function f moves the origin of the coordinate from point $O(0, 0, 0)$ to C_0 . According to exercise 1.3, that function is

$$f(\mathbf{C}) = \mathbf{C} - \begin{bmatrix} 1/2 \\ 1/2 \\ 1/2 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix} \quad \square$$

Chapter 2

Linear equations

2.1 Linear independent set

Let X be any vector space. If vectors $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n \in X$, then for any scalars $s_1, s_2, \dots, s_n \in \mathbb{R}$, the vector

$$\mathbf{x} = \sum_{i=1}^n s_i \mathbf{a}_i$$

is called a *linear combination* of $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$. Let S be a set of vectors $S = \{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n\}$, where all vectors in S belong to the same vector space X . Denote the zero-element of vector space X as $\bar{0}$. If the set S is *linearly independent* then a linear combination of vectors in S equals zero-element if and only if all scalars in the linear combination are zero.

$$\sum_{i=1}^n s_i \mathbf{a}_i = \bar{0} \Leftrightarrow s_i = 0, \forall i \in \{1, 2, \dots, n\}$$

Example 2.1. Let $X = \mathbb{R}^2$ and $S = \{\mathbf{a}_1, \mathbf{a}_2\}$, where $\mathbf{a}_1 = (2, 4)$ and $\mathbf{a}_2 = (1, 2)$. We have

$$\begin{aligned} s_1 \mathbf{a}_1 + s_2 \mathbf{a}_2 = \bar{0} &\Leftrightarrow s_1 \begin{bmatrix} 2 \\ 4 \end{bmatrix} + s_2 \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \\ \Leftrightarrow \begin{cases} 2s_1 + s_2 &= 0 \\ 4s_1 + 2s_2 &= 0 \end{cases} &\Leftrightarrow s_2 = -2s_1 \end{aligned}$$

Thus, the set S is not linearly independent. Notice that the equation $s_2 = -2s_1$ indicates $\mathbf{a}_1, \mathbf{a}_2$ are on the same line in the two-dimensional coordinate system. We can use \mathbf{a}_1 to define \mathbf{a}_2 based on the equation $s_2 = -2s_1$, and vice versa \mathbf{a}_1 can be defined by \mathbf{a}_2 .

Example 2.2. Let $X = \mathbb{R}^2$ be a vector space and $S = \{\mathbf{a}_1, \mathbf{a}_2\}$, where $\mathbf{a}_1 = (2, 4)$ and $\mathbf{a}_2 = (-1, 2)$. We have

$$\begin{aligned} s_1 \mathbf{a}_1 + s_2 \mathbf{a}_2 = \bar{\mathbf{0}} &\Leftrightarrow s_1 \begin{bmatrix} 2 \\ 4 \end{bmatrix} + s_2 \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \\ &\Leftrightarrow \begin{cases} 2s_1 - s_2 = 0 \\ 4s_1 + 2s_2 = 0 \end{cases} \Leftrightarrow \begin{cases} s_2 = 2s_1 \\ s_2 = -2s_1 \end{cases} \Leftrightarrow s_2 = s_1 = 0 \end{aligned}$$

Thus, the set S is linearly independent. Notice that the equation $s_2 = s_1 = 0$ indicates there is no way to define \mathbf{a}_1 using \mathbf{a}_2 , or vice versa.

Example 2.3. Show that vectors $\mathbf{u}, \mathbf{v} \in \mathbb{R}^5$ does not form a linearly independent set if $\mathbf{u} = (-1, 0, 8, 5, 2)$ and $\mathbf{v} = (-6, 0, 48, 30, 12)$.

The set $\{\mathbf{u}, \mathbf{v}\}$ is linearly independent if and only if

$$s_1 \mathbf{u} + s_2 \mathbf{v} = \bar{\mathbf{0}} \Leftrightarrow s_1 = s_2 = 0 \quad (2.1)$$

We have $s_1 \mathbf{u} + s_2 \mathbf{v} = \bar{\mathbf{0}}$ is equivalent to

$$\begin{cases} -s_1 - 6s_2 = 0 \\ 8s_1 + 48s_2 = 0 \\ 5s_1 + 30s_2 = 0 \\ 2s_1 + 12s_2 = 0 \end{cases}$$

All of these equations are equivalent to $s_1 + 6s_2 = 0$. There are an infinite number of solutions to this equations, which does not satisfies (2.1).

Let X be a vector space and the set $A \subseteq X$ is linearly independent. If any vectors $\mathbf{u} \in X$ is a linear combination of vectors in A , then A is a *base* of X .

2.2 System of linear equations

Given \mathbf{A} an n -square matrix and two column vectors $\mathbf{x} = \begin{bmatrix} x_1 & x_2 & \dots & x_n \end{bmatrix}^T$ and $\mathbf{b} = \begin{bmatrix} b_1 & b_2 & \dots & b_n \end{bmatrix}^T$. The system of n linear equations in n unknowns is equivalent to the matrix equation:

$$\mathbf{Ax} = \mathbf{b}$$

The matrix \mathbf{A} is called the *coefficient matrix*. The properties of this matrix, especially determinant, decide the solution of the equation.

The collection of all columns in \mathbf{A} form a linearly independent set (rank of \mathbf{A} is n) if and only if $\det(\mathbf{A}) \neq 0$. This also means that there exists a unique solution for the equation. On the other hand, if $\det(\mathbf{A}) = 0$, then either of two

cases occurs: the system has no solutions, or there are an infinite number of solutions.

There are many ways to solve a system of equations, however, in this notebook, only three solutions related to matrix are mentioned: Inverse matrix, Cramer's rule, Gaussian elimination.

2.2.1 Inverse matrix

Given the matrix equation $\mathbf{Ax} = \mathbf{b}$ where \mathbf{A} is an n -square matrix, $\mathbf{x} = \begin{bmatrix} x_1 & x_2 & \dots & x_n \end{bmatrix}^T$ and $\mathbf{b} = \begin{bmatrix} b_1 & b_2 & \dots & b_n \end{bmatrix}^T$. If $\det(\mathbf{A}) \neq 0$, then \mathbf{A} is invertible and there is a unique solution to the equation:

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

Observe that this method only works if the determinant of the coefficient matrix is nonzero.

Example 2.4. Solve the following system of equation:

$$\begin{cases} 2x_1 + 4x_2 = 8 \\ 7x_1 - 2x_2 = 4 \end{cases} \quad (2.2)$$

Let define matrix \mathbf{A} , \mathbf{x} , and \mathbf{b} as follow:

$$\mathbf{A} = \begin{bmatrix} 2 & 4 \\ 7 & -2 \end{bmatrix} \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \quad \mathbf{b} = \begin{bmatrix} 8 \\ 4 \end{bmatrix}$$

The system of equations is equivalent to

$$\mathbf{Ax} = \mathbf{b}$$

The determinant of coefficient matrix is

$$\det(\mathbf{A}) = 2 \cdot (-2) - 7 \cdot 4 = -32 \neq 0$$

Thus, the equation has a unique solution, which is

$$\mathbf{x} = \mathbf{A}^{-1}\mathbf{b} = \frac{1}{-32} \begin{bmatrix} -2 & -4 \\ -7 & 2 \end{bmatrix} \begin{bmatrix} 8 \\ 4 \end{bmatrix} = \begin{bmatrix} 1 \\ 3/2 \end{bmatrix}$$

2.2.2 Cramer's rule

Suppose we need to solve the matrix equation $\mathbf{Ax} = \mathbf{b}$ where \mathbf{A} is an n -square matrix, $\mathbf{x} = \begin{bmatrix} x_1 & x_2 & \dots & x_n \end{bmatrix}^T$ and $\mathbf{b} = \begin{bmatrix} b_1 & b_2 & \dots & b_n \end{bmatrix}^T$. Let $\mathbf{A}^{(j)}$ be a matrix \mathbf{A} but replace j th column with \mathbf{b} . Then we have

$$x_j = \frac{\det(\mathbf{A}^{(j)})}{\det(\mathbf{A})}$$

Like Inverse matrix, this method only works if the determinant of the coefficient matrix is nonzero.

Example 2.5. The system of equation (2.2) can solved using Crammer's rule as follow:

$$\mathbf{A} = \begin{bmatrix} 2 & 4 \\ 7 & -2 \end{bmatrix}, \quad \det(\mathbf{A}) = -32$$

$$\det(\mathbf{A}^{(1)}) = \begin{vmatrix} 8 & 4 \\ 4 & -2 \end{vmatrix} = -32 \Rightarrow x_1 = \frac{-32}{-32} = 1$$

$$\det(\mathbf{A}^{(2)}) = \begin{vmatrix} 2 & 7 \\ 8 & 4 \end{vmatrix} = -48 \Rightarrow x_2 = \frac{-48}{-32} = \frac{3}{2}$$

2.2.3 Gaussian elimination

Gaussian elimination (also known as row reduction) is an algorithm for solving systems of linear equations. It is usually understood as a sequence of operations performed on the corresponding matrix of coefficients. This method can also be used to find the rank of a matrix, to calculate the determinant of a matrix, and to calculate the inverse of an invertible square matrix.

To perform Gaussian elimination, one uses a sequence of elementary row operations to modify the matrix until the lower left-hand corner of the matrix is filled with zeros, as much as possible. There are three types of elementary row operations:

- Swap the positions of two rows.
- Multiply a row by a nonzero scalar.
- Add to one row a scalar multiple of another.

These operations are already shown in section 1.4.4 as properties of determinants.

Example 2.6. The system of equation (2.2) can solved using Gaussian elimination. Let define matrix \mathbf{A} , \mathbf{x} , and \mathbf{b} as follow:

$$\mathbf{A} = \begin{bmatrix} 2 & 4 \\ 7 & -2 \end{bmatrix} \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \quad \mathbf{b} = \begin{bmatrix} 8 \\ 4 \end{bmatrix}$$

First, we create the augmented matrix which is formed by merging matrix \mathbf{A} and \mathbf{b} .

$$\left[\begin{array}{cc|c} 2 & 4 & 8 \\ 7 & -2 & 4 \end{array} \right]$$

Swap the positions of two rows, we have:

$$\left[\begin{array}{cc|c} 7 & -2 & 4 \\ 2 & 4 & 8 \end{array} \right]$$

Multiply the first row by $1/7$, we have:

$$\left[\begin{array}{cc|c} 1 & -2/7 & 4/7 \\ 2 & 4 & 8 \end{array} \right]$$

Replace the second row by $R_2 \leftarrow R_2 - 2R_1$

$$\left[\begin{array}{cc|c} 1 & -2/7 & 4/7 \\ 0 & 32/7 & 48/7 \end{array} \right]$$

Multiply the first row by $32/7$, we have:

$$\left[\begin{array}{cc|c} 1 & -2/7 & 4/7 \\ 0 & 1 & 48/32 \end{array} \right]$$

Replace the first row by $R_1 \leftarrow R_2 + \frac{2}{7}R_1$

$$\left[\begin{array}{cc|c} 1 & 0 & 1 \\ 0 & 1 & 48/32 \end{array} \right]$$

Thus, $x_1 = 1$ and $x_2 = 48/32$.

2.3 Linear least squares

Let \mathbf{A} be an $m \times n$ matrix, \mathbf{x} be an $n \times 1$ matrix and \mathbf{b} be an $m \times 1$ matrix. Solve the matrix equation $\mathbf{Ax} = \mathbf{b}$.

If $m = n$, we can use three methods mentioned in section 2.2. However, if $m > n$, the the solution usually does not exist. But if the columns of \mathbf{A} form a linearly independent set, then

$$\det(\mathbf{A}^T \mathbf{A}) \neq 0 \quad \text{and} \quad \mathbf{A}^T \mathbf{Ax} = \mathbf{A}^T \mathbf{b} \quad (2.3)$$

$$\mathbf{x} = \mathbf{A}^T \mathbf{A}^{-1} \mathbf{A}^T \mathbf{b} \quad (2.4)$$

The matrix equation now can be solved with the minimum error, meaning $\|\mathbf{Ax} - \mathbf{b}\|$ is minimum. Using formula (2.4) to solve a system of equations is called The method of linear least squares. See exercise 2.4 as an example.

Exercise 2.1. Suppose we want to draw a best fitting line through three points (4,-2), (5,0), (5,-1), (6,4), (7,5) and (8,5) so a linear function $y = t + kx$ should be given. Determine this function by means of the Method of Least Squares (LSM). Perform calculations by SAGE then plot points and the line in the same coordinate system.

2.4 Exercise

Proof. The linear function $y = kx + t$ should satisfies the following system of equations with the minimum error:

$$\begin{cases} 4k + t = -2 \\ 5k + t = 0 \\ 5k + t = -1 \\ 6k + t = 4 \\ 7k + t = 5 \\ 8k + t = 5 \end{cases} \quad (2.5)$$

The equivalent matrix equation of (2.5) is $\mathbf{Ax} = \mathbf{b}$, where

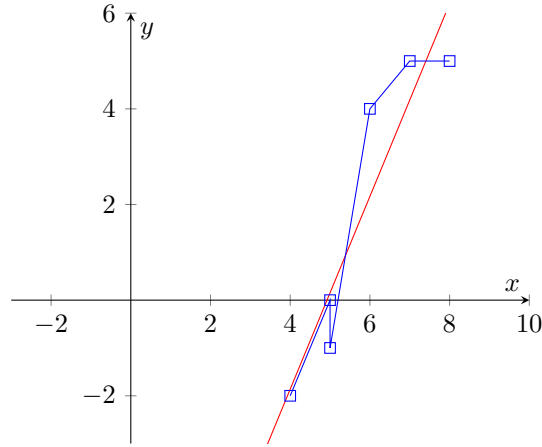
$$\mathbf{A} = \begin{bmatrix} 4 & 1 \\ 5 & 1 \\ 5 & 1 \\ 6 & 1 \\ 7 & 1 \\ 8 & 1 \end{bmatrix} \quad \mathbf{x} = \begin{bmatrix} k \\ t \end{bmatrix} \quad \mathbf{b} = \begin{bmatrix} -2 \\ 0 \\ -1 \\ 4 \\ 5 \\ 5 \end{bmatrix}$$

Using formula (2.4), we now can perform calculations by SAGE as follow:

```
A=matrix([[4,1], [5,1], [5,1], [6,1], [7,1], [8,1]])
b=matrix([ [-2,0,-1,4,5,5] ]).T
x=(A.T*A).inverse() * (A.T*b)

k=x[0][0]
t=x[1][0]
z=var("z")

S=plot(k*z+t, (z,-1,9))
P = scatter_plot([[4,-2], [5,0], [5,-1], [6,4], [7,5], [8,5]])
show(S+P)
```



□

Exercise 2.2. Suppose we want to draw a best 4th order polynomial line through 6 points $(7, 9)$, $(6, 12)$, $(4, 4)$, $(2, -1)$, $(-1, 0)$, $(6, 16)$. Determine this function by means of the Method of Least Squares (LSM). Perform calculations by SAGE and plot the points and the line in the same coordinate system.

Proof. We need to find the coefficients of the following polynomial

$$y = mx^4 + nx^3 + px^2 + qx + r$$

Substitute x and y to this polynomial, we obtain a system of equations:

$$\begin{cases} 2401m + 343n + 49p + 7q + r = 9 \\ 1296m + 216n + 36p + 6q + r = 12 \\ 256m + 64n + 16p + 4q + r = 4 \\ 16m + 8n + 4p + 2q + r = -1 \\ m - n + p - q + r = 0 \\ 1296m + 216n + 36p + 6q + r = 16 \end{cases}$$

,which is equivalent to the following matrix equation

$$\begin{bmatrix} 2401 & 343 & 49 & 7 & 1 \\ 1296 & 216 & 36 & 6 & 1 \\ 256 & 64 & 16 & 4 & 1 \\ 16 & 8 & 4 & 2 & 1 \\ 1 & -1 & 1 & -1 & 1 \\ 1296 & 216 & 36 & 6 & 1 \end{bmatrix} \begin{bmatrix} m \\ n \\ p \\ q \\ r \end{bmatrix} = \begin{bmatrix} 9 \\ 12 \\ 4 \\ -1 \\ 0 \\ 16 \end{bmatrix}$$

Using formula (2.4), we now can perform calculations by SAGE as follow:

```

A=matrix([[2401,343,49,7,1],[1296,216,36,6,1],[256,64,16,4,1],
          [16,8,4,2,1],[1,-1,1,-1,1],[1296,216,36,6,1]])
b=matrix([[9,12,4,-1,0,16]]).T
x=(A.T*A).inverse()* (A.T*b)

m=x[0][0]
n=x[1][0]
p=x[2][0]
q=x[3][0]
r=x[4][0]

z=var("z")
S=plot(m*z**4+n*z**3+p*z*z+q*z+r, (z,-2,8))
P = scatter_plot([[7,9],[6,12],[4,4],[2,-1],[-1,0],[6,16]])
show(S+P)

```

□