Linear algebra

Keywords: vector, matrix, eigenvalue, eigenvector, diagonalization, linear transformation, quadratic forms and symmetric matrices

Compare: Vorlesung Mathematik 2

Contents

1 Matrices and vectors

1.1 Real Vectors

- n-dimensional space $\mathbb{R}^n$
- elements $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ are called <u>n-vectors</u>

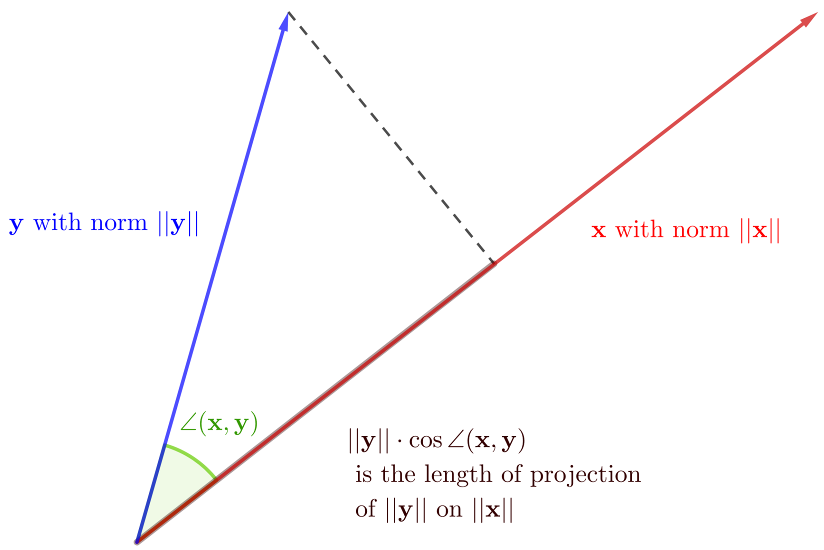
$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} x_1 & x_2 & \dots & x_n \end{pmatrix}^T \text{ and } \mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

- <u>scalar product</u> and <u>norm or length</u>:

$$\begin{aligned} \mathbf{x} \bullet \mathbf{y} &= \mathbf{x}^T\mathbf{y} = <\mathbf{x}, \mathbf{y}> &= x_1y_1 + x_2y_2 + \cdots + x_ny_n \\ ||\mathbf{x}|| &= \sqrt{\mathbf{x} \bullet \mathbf{x}} &= \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2} \end{aligned}$$

$$\boxed{\mathbf{x} \bullet \mathbf{y} = ||\mathbf{x}|| \cdot ||\mathbf{y}|| \cdot \cos\angle(\mathbf{x}, \mathbf{y})}$$

You may see, that $||\mathbf{y}|| \cdot \cos\angle(\mathbf{x}, \mathbf{y})$ is the length of the orthogonal projection of the vector $\mathbf{y}$ on $\mathbf{x}$, with the negative sign if the projection has an opposite direction with respect to $\mathbf{x}$.



- For each vector $\mathbf{x}$, the vector $\frac{1}{||\mathbf{x}||}\mathbf{x}$ is an <u>unit vector</u> (vector with norm 1), in the direction of $\mathbf{x}$.

- Let $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k \in \mathbb{R}^n$ be a family of vectors.
  - If $a_1, a_2, \ldots, a_k \in \mathbb{R}$, then $\mathbf{z} = a_1\mathbf{x}_1 + a_2\mathbf{x}_2 + \cdots + a_k\mathbf{x}_k$ is called a linear combination of $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k$.
  - The set of all linear combinations of the vectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k$ is called the vector space spanned by the vectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k$ and denoted by
$$V(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k) \;=\; \{a_1\mathbf{x}_1 + a_2\mathbf{x}_2 + \cdots + a_k\mathbf{x}_k \mid a_1, a_2, \ldots, a_k \in \mathbb{R}\}$$
  - $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k$ are called linearly dependent, if there exist $b_1, b_2, \ldots, b_k \in \mathbb{R}$ such that $b_1\mathbf{x}_1 + b_2\mathbf{x}_2 + \cdots + b_k\mathbf{x}_k = \mathbf{0}$ and not all $b_j = 0$.
  - $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k$ are called linearly independent, if a linear combination of the zero vector
$$b_1\mathbf{x}_1 + b_2\mathbf{x}_2 + \cdots + b_k\mathbf{x}_k = \mathbf{0}$$
is possible only with $b_1 = b_2 = \cdots = b_k = 0$.
  - Each family of exactly n linearly independent vectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n \in \mathbb{R}^n$ is a so called basis of $\mathbb{R}^n$. This means, that each vector $\mathbf{x} \in \mathbb{R}^n$ can uniquely expressed as a linear combination of the basis:
$$\mathbf{x} \;=\; b_1\mathbf{x}_1 + b_2\mathbf{x}_2 + \cdots + b_n\mathbf{x}_n$$
  - A family of n (linearly independent) vectors $\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_n \in \mathbb{R}^n$ is called orthonormal basis of $\mathbb{R}^n$ if
$$\mathbf{p}_i \bullet \mathbf{p}_j \;=\; \mathbf{p}_i^T\mathbf{p}_j \;=\; \begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{if } i = j \end{cases}$$
for all $i, j = 1, 2, \ldots, n$. This means, that each vector has length 1 and each pair of (different) vectors has a right angle. As before, each vector $\mathbf{x} \in \mathbb{R}^n$ can uniquely expressed as a linear combination of the ortogonal basis
$$\mathbf{x} \;=\; b_1\mathbf{p}_1 + b_2\mathbf{p}_2 + \cdots + b_n\mathbf{p}_n \;=\; \sum_{i=1}^{n} b_i\mathbf{p}_i$$
but the coefficients b_i have a nice interpretation (for orthogonal bases). We see
$$\mathbf{p}_j^T\mathbf{x} \;=\; \sum_{i=1}^{n} b_i\mathbf{p}_j^T\mathbf{p}_i \;=\; b_j\mathbf{p}_j^T\mathbf{p}_j \;=\; b_j$$
Hence the coefficient
$$b_j = \mathbf{p}_j^T\mathbf{x} = ||\mathbf{p}_j|| \cdot ||\mathbf{x}|| \cdot \cos\angle(\mathbf{p}_j, \mathbf{x}) = ||\mathbf{x}|| \cdot \cos\angle(\mathbf{p}_j, \mathbf{x})$$
is the length of the orthogonal projection of $\mathbf{x}$ on the basis vector $\mathbf{p}_j$.

1.2 Real Matrices

$\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_m \in \mathbb{R}^n$

$$\mathbf{a}_1 = \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{pmatrix}, \mathbf{a}_2 = \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{n2} \end{pmatrix}, \ldots, \mathbf{a}_m = \begin{pmatrix} a_{1m} \\ a_{2m} \\ \vdots \\ a_{nm} \end{pmatrix}$$

$$\rightarrow \quad \mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1m} \\ a_{21} & a_{22} & \ldots & a_{2m} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nm} \end{pmatrix} = (\ \mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_m\ )$$

is called an $n \times m$ matrix.

Notation: $\mathbf{A} \in \mathbb{R}^{n \times m}$

- The inverse matrix $\mathbf{A}^{-1}$ of the $n \times n$ matrix $\mathbf{A} = (a_{ij})$ is defined by

$$\mathbf{A}^{-1} \cdot \mathbf{A} = \mathbf{A} \cdot \mathbf{A}^{-1} = \mathbf{I_n} = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}.$$

- The rank of a matrix $\mathbf{A}$, written $rk(\mathbf{A})$, is the maximum number of linearly independent column or row vectors in $\mathbf{A}$.

- For the $n \times n$ matrix $\mathbf{A}$ let $\mathbf{A}_{ij}$ denote the $(n-1) \times (n-1)$ submatrix of $\mathbf{A}$ generated by cancelling the i-th row and the j-th column of $\mathbf{A}$. Then the determinant $\det(\mathbf{A})$ is given (recursively) by the so called cofactor expansion

$$\boxed{\det(\mathbf{A}) = |\mathbf{A}| = a_{11} \det \mathbf{A}_{11} - a_{12} \det \mathbf{A}_{12} + \cdots + (-1)^{n+1} a_{1n} \det \mathbf{A}_{1n}}$$

- $\det(\mathbf{A} \cdot \mathbf{B}) = \det(\mathbf{A}) \cdot \det(\mathbf{B})$

Example 1.1

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

$$= 1 \cdot \begin{vmatrix} 2 & 1 & 2 \\ -2 & 1 & -2 \\ 1 & -2 & -1 \end{vmatrix} - 1 \cdot \begin{vmatrix} 1 & 1 & 2 \\ 1 & 1 & -2 \\ 0 & -2 & -1 \end{vmatrix} + 3 \cdot \begin{vmatrix} 1 & 2 & 2 \\ 1 & -2 & -2 \\ 0 & 1 & -1 \end{vmatrix} - 3 \cdot \begin{vmatrix} 1 & 2 & 1 \\ 1 & -2 & 1 \\ 0 & 1 & -2 \end{vmatrix}.$$

1.3 Linear transformations and matrices

Definition 1.1 *A linear transformation is a map $T : \mathbb{R}^m \to \mathbb{R}^n$ such that for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^m$ and all $\lambda, \mu \in \mathbb{R}$ we have:*

$$\boxed{T(\lambda \cdot \mathbf{x} + \mu \cdot \mathbf{y}) \;=\; \lambda \cdot T(\mathbf{x}) + \mu \cdot T(\mathbf{y})}$$

Example 1.2

- *The map $T(\mathbf{x}) = T(x_1, x_2, x_3) = x_1 + 2x_2 + 4x_3$ is a linear transformation from $\mathbb{R}^3$ to $\mathbb{R}^1$:*

$$\begin{aligned}
T(\lambda \cdot \mathbf{x} + \mu \cdot \mathbf{y}) &= T(\lambda \cdot x_1 + \mu \cdot y_1, \lambda \cdot x_2 + \mu \cdot y_2, \lambda \cdot x_3 + \mu \cdot y_3) \\
&= \lambda \cdot x_1 + \mu \cdot y_1 + 2 \cdot (\lambda \cdot x_2 + \mu \cdot y_2) + 4 \cdot (\lambda \cdot x_3 + \mu \cdot y_3) \\
&= \lambda \cdot x_1 + 2 \cdot \lambda \cdot x_2 + 4 \cdot \lambda \cdot x_3 + \mu \cdot y_1 + 2 \cdot \mu \cdot y_2 + 4 \cdot \mu \cdot y_3 \\
&= \lambda \cdot (x_1 + 2 \cdot x_2 + 4 \cdot x_3) + \mu \cdot (y_1 + 2 \cdot y_2 + 4 \cdot y_3) \\
&= \lambda \cdot T(\mathbf{x}) + \mu \cdot L(\mathbf{y})
\end{aligned}$$

- *The map $L(\mathbf{x}) = L(x_1, x_2) = \begin{pmatrix} x_1 + x_2 \\ x_1 - x_2 \end{pmatrix}$ is a linear transformation from $\mathbb{R}^2$ to $\mathbb{R}^2$. Proof it!*

- *The map $L(\mathbf{x}) = L(x_1, x_2) = \begin{pmatrix} x_1^2 + x_2 \\ x_1 - x_2 \end{pmatrix}$ is* **not** *linear.*

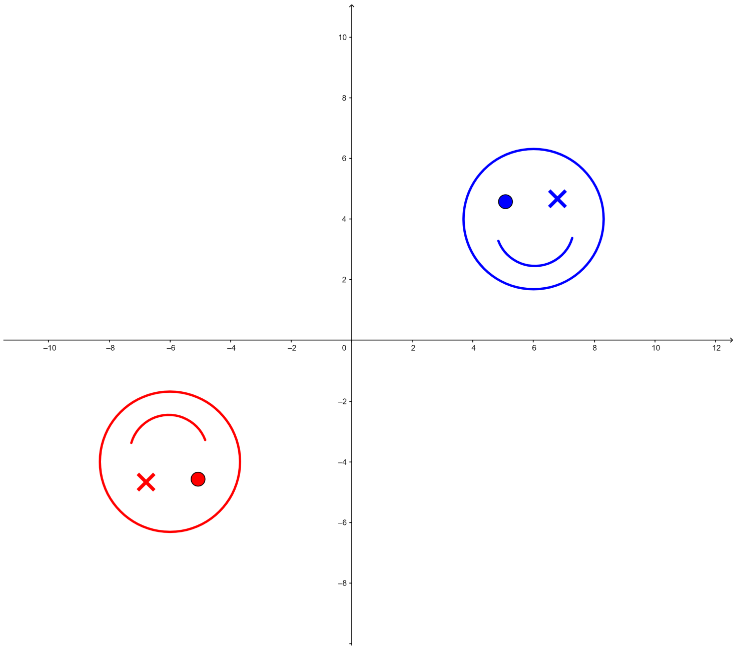
Each $n \times m$ matrix $\mathbf{A}$ defines a linear transformation by matrix multiplication

$$\begin{aligned}
T_{\mathbf{A}}(\mathbf{x}) \;=\; \mathbf{A} \cdot \mathbf{x} &= \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{pmatrix} \\
&= x_1 \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix} + x_2 \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix} + \cdots + x_n \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix}
\end{aligned}$$

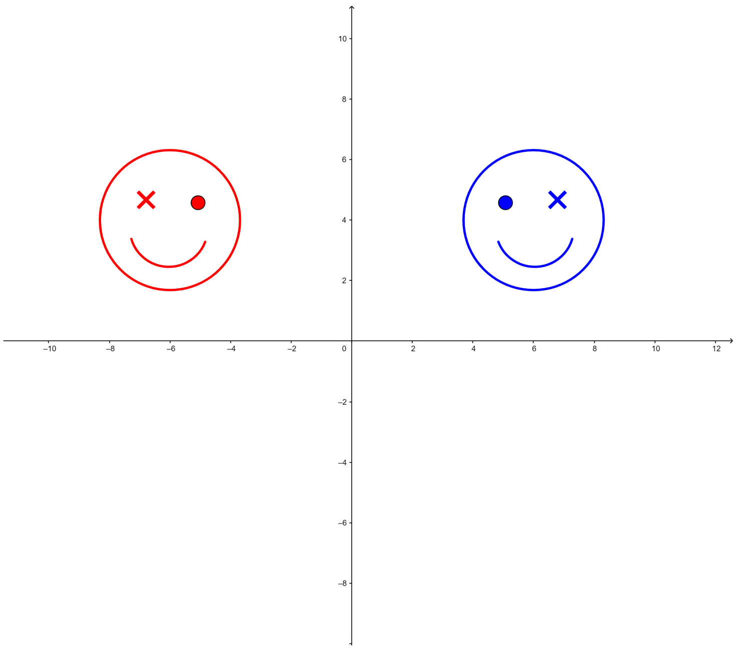
The image of the vector $\mathbf{x} \in \mathbb{R}^m$ is a linear combination of the column vectors of the matrix $\mathbf{A}$.

Example 1.3 *In the following picture you can see the original figur (blue) and the image of this figur under the linear map L_A. Each blue point (endpoint of the vector $\mathbf{x}$) is maped on the point $A\mathbf{x}$ (red).*

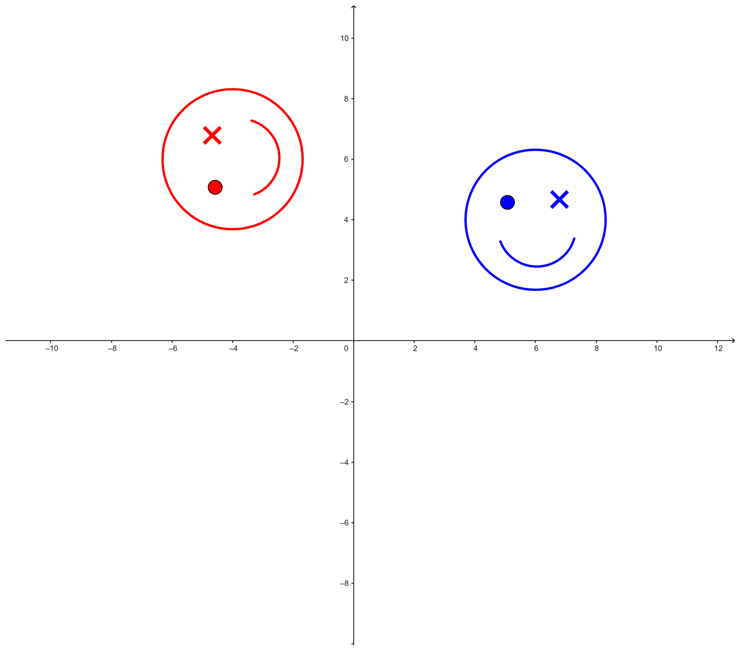
- $A = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$ *Rotation with center* $(0,0)$ *by* 180 *degree*



- $A = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$ *Reflection along the y-axis*



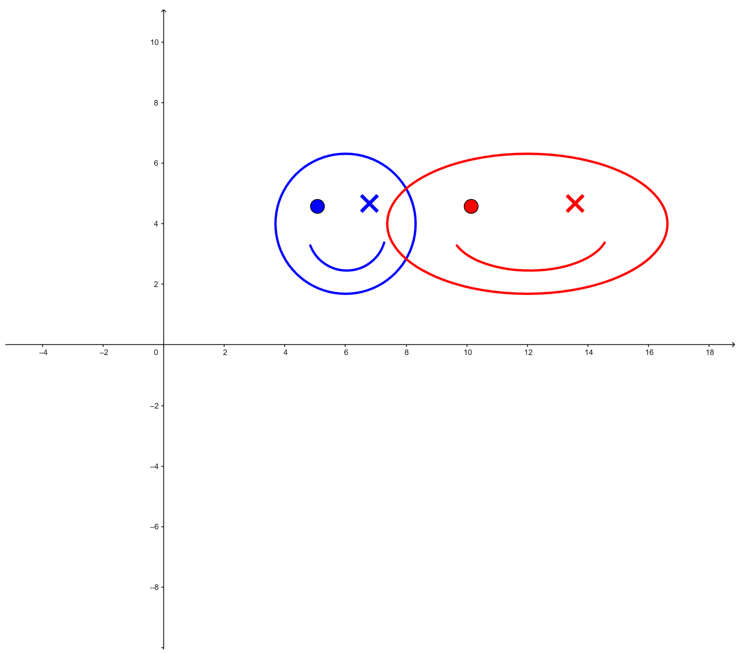
- $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ *Rotation with center* $(0,0)$ *by* 90 *degree*



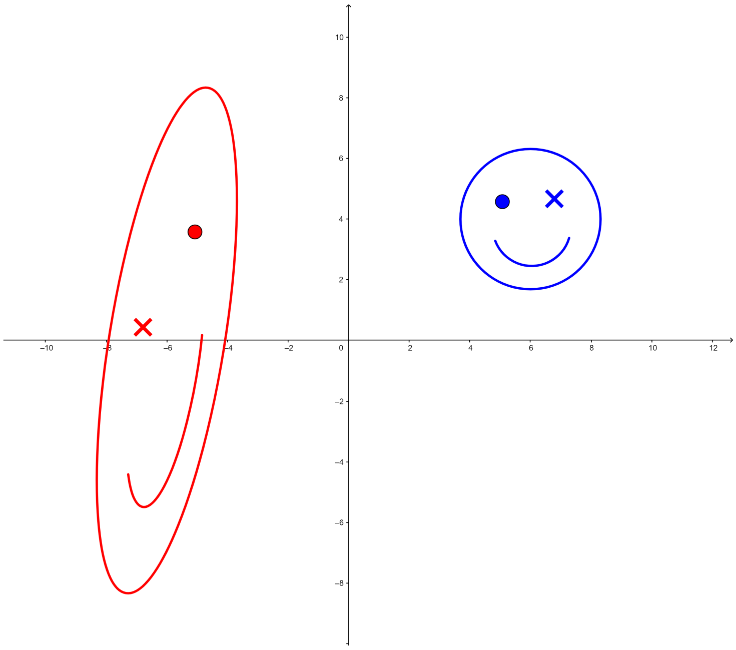
Remark: *The general rotation with center* $(0,0)$ *by* α *degree is given by the following matrix:*

$$\begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix}$$

- $A = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}$ *Scaling (by the factor* 2*) in* x*-direction*



- $\begin{pmatrix} -1 & 0 \\ -2 & 3 \end{pmatrix}$?



Projections on lines The following type of matrices is of the special interest. Let

$$\mathbf{p} = \begin{pmatrix} p_1 \\ p_2 \\ \vdots \\ p_n \end{pmatrix}$$

be an arbitrary non-zero vector. The direct calculation

$$\begin{aligned} \mathbf{p} \cdot \mathbf{p}^T &= \begin{pmatrix} p_1 \\ p_2 \\ \vdots \\ p_n \end{pmatrix} \cdot \begin{pmatrix} p_1 & p_2 & \dots & p_n \end{pmatrix} = \begin{pmatrix} p_1 \cdot p_1 & p_1 \cdot p_2 & \dots & p_1 \cdot p_n \\ p_2 \cdot p_1 & p_2 \cdot p_2 & \dots & p_2 \cdot p_n \\ \vdots & \vdots & \vdots & \vdots \\ p_n \cdot p_1 & p_n \cdot p_2 & \dots & p_n \cdot p_n \end{pmatrix} \\ &= \begin{pmatrix} - & p_1 \cdot \mathbf{p}^T & - \\ - & p_2 \cdot \mathbf{p}^T & - \\ \vdots & \vdots & \vdots \\ - & p_2 \cdot \mathbf{p}^T & - \end{pmatrix} = \begin{pmatrix} | & | & & | \\ p_1 \cdot \mathbf{p} & p_2 \cdot \mathbf{p} & \dots & p_n \cdot \mathbf{p} \\ | & | & & | \end{pmatrix} \end{aligned}$$

shows that $\mathbf{p} \cdot \mathbf{p}^T$ is a symmetric $n \times n$ matrix of rank 1 (all columns and all rows are multiples of the vector $\mathbf{p}$ resp. $\mathbf{p}^T$).

Theorem 1.1 *Let $\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_n \in \mathbb{R}^n$ an orthonormal basis of $\mathbb{R}^n$, this means*

$$\mathbf{p}_i \bullet \mathbf{p}_j = \mathbf{p}_i^T \mathbf{p}_j = \begin{cases} 0 & \textit{if } i \neq j \\ 1 & \textit{if } i = j \end{cases}$$

and $\mathbf{P}_i = \mathbf{p}_i \cdot \mathbf{p}_i^T$ for all $i, j = 1, 2, \dots, n$.
Then the linear map

$$T_{\mathbf{P}_i}(\mathbf{x}) = \mathbf{P}_i \cdot \mathbf{x}$$

given by matrix multiplication is a projection on the line spanned by the vector $\mathbf{p}_i$ for all $i = 1, 2, \dots, n$.

Proof: Let

$$\mathbf{x} = b_1 \mathbf{p}_1 + b_2 \mathbf{p}_2 + \dots + b_n \mathbf{p}_n = \sum_{j=1}^{n} b_j \mathbf{p}_j$$

be a vector expressed in the given orthonormal basis. Then by direct calculation

$$\begin{aligned} \mathbf{P}_i \cdot \mathbf{x} &= \left(\mathbf{p}_i \cdot \mathbf{p}_i^T\right) \cdot \left(\sum_{j=1}^{n} b_j \mathbf{p}_j\right) \\ &= \mathbf{p}_i \cdot \left(\mathbf{p}_i^T \cdot \sum_{j=1}^{n} b_j \mathbf{p}_j\right) \\ &= \mathbf{p}_i \cdot \left(\sum_{j=1}^{n} b_j \mathbf{p}_i^T \cdot \mathbf{p}_j\right) = \mathbf{p}_i b_i = b_i \mathbf{p}_i. \end{aligned}$$

In the sum between the brackets all terms $\mathbf{p}_i^T \cdot \mathbf{p}_j$ are equal to 0 if $i \neq j$.

□

1.4 Solutions of systems of linear equations

The general linear system of m equations in n unknowns is given by

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ \vdots \qquad\qquad & \quad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m. \end{aligned}$$

$$\boxed{\underbrace{\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}}_{\mathbf{A}} \underbrace{\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}}_{\mathbf{x}} = \underbrace{\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}}_{\mathbf{b}}}$$

Let $\mathbf{a}_j$ denote the jth column vector of $\mathbf{A}$. The matrix

$$\mathbf{A_b} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{pmatrix} = (\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n, \mathbf{b})$$

is called the augmented matrix of the system. The general linear system can also be written as

$$\boxed{x_1 \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix} + x_2 \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix} + \cdots + x_n \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}}$$

We understand: The problem is to express **the right side b as a linear combination of the column vectors $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$ of the matrix A!!**

Theorem 1.2

$$\mathbf{Ax} = \mathbf{b} \textit{ has a solution} \quad \longleftrightarrow \quad rk(\mathbf{A}) = rk(\mathbf{A_b})$$

1.5 Complex matrices and vectors

The set

$$\mathbb{C} \;=\; \left\{\, a+ib \mid a,b\in\mathbb{R}, i^2=-1 \,\right\}$$

is called the set of complex numbers. i is called imaginary unit and per definition is $i^2 = -1$. The real numbers $\mathbb{R}$ are a subset of $\mathbb{C}$ (all complex numbers with $b = 0$). Together with addition and multiplication, given by

$$\begin{aligned}(a_1+ib_1)+(a_2+ib_2) &= (a_1+a_2+i(b_1+b_2))\\ (a_1+ib_1)\cdot(a_2+ib_2) &= (a_1a_2-b_1b_2+i(a_1b_2+a_2b_1)),\end{aligned}$$

$\mathbb{C}$ is a very nice set for calculations.

Sometimes it is helpful to allow complex matrices and vectors (matrices whose elements are complex numbers). A complex matrix can be viewed as a combination of two real matrices:

$$\begin{aligned}\mathbf{A} &= \begin{pmatrix} a_{11}+ib_{11} & a_{12}+ib_{12} & \dots & a_{1m}+ib_{1m}\\ a_{21}+ib_{21} & a_{22}+ib_{22} & \dots & a_{2m}+ib_{2m}\\ \vdots & \vdots & \vdots & \vdots\\ a_{n1}+ib_{n1} & a_{n2}+ib_{n2} & \dots & a_{nm}+ib_{nm}\end{pmatrix}\\ &= \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1m}\\ a_{21} & a_{22} & \dots & a_{2m}\\ \vdots & \vdots & \vdots & \vdots\\ a_{n1} & a_{n2} & \dots & a_{nm}\end{pmatrix} + i\cdot \begin{pmatrix} b_{11} & b_{12} & \dots & b_{1m}\\ b_{21} & b_{22} & \dots & b_{2m}\\ \vdots & \vdots & \vdots & \vdots\\ b_{n1} & b_{n2} & \dots & b_{nm}\end{pmatrix}\end{aligned}$$

1.6 Matrix calculus

1a.	$\mathbf{A}+\mathbf{B}=\mathbf{B}+\mathbf{A}$	1b.	In general: $\mathbf{AB}\neq\mathbf{BA}$
2a.	$(\mathbf{A}+\mathbf{B})+\mathbf{C}=\mathbf{A}+(\mathbf{B}+\mathbf{C})$	2b.	$(\mathbf{AB})\mathbf{C}=\mathbf{A}(\mathbf{BC})$
3a.	$\mathbf{A}+\mathbf{0}=\mathbf{A}$	3b.	$\mathbf{A}I=I\mathbf{A}=\mathbf{A}$ ($\mathbf{A}$ square)

$$\begin{array}{lllll} 4. & \mathbf{AB}=\mathbf{0} & \not\Rightarrow & \mathbf{A}=\mathbf{0} \text{ or } \mathbf{B}=\mathbf{0} & \\ 5. & \mathbf{AB}=\mathbf{AC} & \not\Rightarrow & \mathbf{B}=\mathbf{C} & \end{array}$$

$$\begin{array}{lllll} 6. & \lambda(\mathbf{A}+\mathbf{B}) & = & \lambda\mathbf{A}+\lambda\mathbf{B} & \lambda\in\mathbb{R}\\ 7. & \mathbf{A}(\mathbf{B}+\mathbf{C}) & = & \mathbf{AB}+\mathbf{AC} & \\ 8. & (\mathbf{A}+\mathbf{B})\mathbf{C} & = & \mathbf{AC}+\mathbf{BC} & \end{array}$$

$$\begin{array}{llll} 9. & (\mathbf{A}^{-1})^{-1} & = & \mathbf{A}\\ 10. & (\mathbf{AB})^{\mathbf{-1}} & = & \mathbf{B}^{\mathbf{-1}}\mathbf{A}^{\mathbf{-1}}\\ 11. & (\mathbf{A}^T)^T & = & \mathbf{A}\\ 12. & (\mathbf{A}+\mathbf{B})^{\mathbf{T}} & = & \mathbf{A}^T+\mathbf{B}^{\mathbf{T}}\\ 13. & (\mathbf{AB})^{\mathbf{T}} & = & \mathbf{B}^{\mathbf{T}}\mathbf{A}^{\mathbf{T}}\\ 14. & (\mathbf{A}^{-1})^T & = & (\mathbf{A}^T)^{-1} \end{array}$$

For $\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ with $ad - bc \neq 0$ is $\mathbf{A}^{-1} = \dfrac{1}{ad-bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$.
All these definitions and results can be generalized to vectors and matrices with complex entries.

Exercise 1.1 *Why is* $(\mathbf{AB})^{-1}$ *equal to* $\mathbf{B}^{-1}\mathbf{A}^{-1}$ *and not equal to* $\mathbf{A}^{-1}\mathbf{B}^{-1}$*?*

Exercise 1.2 *Why is* $\mathbf{A}^{-1} = \dfrac{1}{ad-bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$ *for* $\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$*?*

2 Eigenvalues and eigenvectors

2.1 Definition and determination

Definition 2.1 *If* $\mathbf{A}$ *is a real (or complex)* $n \times n$ *matrix, then a (complex) number* λ *is an eigenvalue of* $\mathbf{A}$ *if there is a nonzero (complex) vector* $\mathbf{x} \in \mathbb{C}^{\mathbf{n}}$ *such that*

$$\boxed{\mathbf{A}\mathbf{x} \;=\; \lambda\mathbf{x}}$$

Then $\mathbf{x}$ *is an eigenvector of* $\mathbf{A}$ *(associated with* λ*).*

Remark: If $\mathbf{x}$ is an eigenvector associated with the eigenvalue λ, then so is $\alpha\mathbf{x}$ for every real (and complex) number $\alpha \neq 0$.

$$\mathbf{A}\,(\alpha\,\mathbf{x}) \;=\; \alpha\,\mathbf{A}\,\mathbf{x} \;=\; \alpha\,(\lambda\,\mathbf{x}) \;=\; \lambda\,(\alpha\,\mathbf{x})$$

How to find eigenvalues? The equation can be written as

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

This is a homogeneous system of linear equations. It has a solution $\mathbf{x} \neq \mathbf{0}$ if and only if the matrix $(\mathbf{A} - \lambda\,\mathbf{I})$ is singular which means that its determinant equals to 0.

$$\boxed{(\mathbf{A} - \lambda\,\mathbf{I}) \text{ singular} \quad \Leftrightarrow \quad \underbrace{\det(\mathbf{A} - \lambda\,\mathbf{I})}_{p_A(\lambda)} = 0}$$

$p_A(\lambda) = 0$ is called the characteristic equation of $\mathbf{A}$. The function $p_A(\lambda)$ is a polynomial of degree n in λ, called the characteristic polynomial of $\mathbf{A}$.

Theorem 2.1 *Are both* $\mathbf{x}$ *and* $\mathbf{y}$ *eigenvectors of* $\mathbf{A}$ *associated with the same eigenvalue* λ*, then all nontrivial linear combinations of* $\mathbf{x}$ *and* $\mathbf{y}$ *are eigenvectors associated with* λ *to. This means, that the set of all eigenvectors (and the* $\mathbf{0}$*-vector) associated with an eigenvalue* λ *is a vector space, called the eigenspace of* λ*:*

$$V(\lambda) \;=\; \{\, \mathbf{x} \in \mathbb{C}^n \mid (\mathbf{A} - \lambda\mathbf{I})\,\mathbf{x} = \mathbf{0} \,\}.$$

The dimension of the vector space $V(\lambda)$ *is called the geometric multiplicity of the eigenvalue* λ*.*

Proof: Let $\mathbf{A}\mathbf{x} \;=\; \lambda\,\mathbf{x}$ and $\mathbf{A}\mathbf{y} \;=\; \lambda\,\mathbf{y}$ and $a, b \in \mathbb{R}$, not both equal to 0. Then we have for $\mathbf{z} \;=\; a\mathbf{x} + b\mathbf{y}$:

$$\mathbf{A}\mathbf{z} \;=\; \mathbf{A}(a\mathbf{x} + b\mathbf{y}) \;=\; a\mathbf{A}\mathbf{x} + b\mathbf{A}\mathbf{y} \;=\; a\lambda\mathbf{x} + b\lambda\mathbf{y} \;=\; \lambda\mathbf{z}.$$

□

Determination of the eigenvalues and eigenvectors

1. The polynomial equation $p_A(\lambda) = 0$ has always n complex solutions (counted with multiplicity) and may have no real solutions. If $\lambda_1, \ldots, \lambda_r \in \mathbb{C}$ are the pairwise distinct solutions (the eigenvalues of $\mathbf{A}$) with the multiplicities $k_1, \ldots, k_r$ then the characteristic polynomial can be written as
$$p_A(\lambda) \;=\; (\lambda_1 - \lambda)^{k_1}\,(\lambda_2 - \lambda)^{k_2} \ldots (\lambda_r - \lambda)^{k_r}.$$
The multiplicity k_i of the zero λ_i is called <u>algebraic multiplicity</u> of the eigenvalue λ_i. Generally, the determination of the (exact) zeros is impossible for $n \geq 5$ and we have to use numerical methods.

2. For each eigenvalue λ_i $(1 \leq i \leq r)$ we compute the eigenspace of λ_i
$$V(\lambda_i) \;=\; \{\, \mathbf{x} \in \mathbb{C}^n \mid (\mathbf{A} - \lambda_i \mathbf{I})\, \mathbf{x} = \mathbf{0} \,\}.$$

Example 2.1

$$A = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

- $p_A(\lambda) \;=\; (2-\lambda)^2(1-\lambda)$
- *Zeros of the characteristic polynomial:* $\lambda_1 = 1$ *(algebraic multiplicity* 1*),* $\lambda_2 = 2$ *(algebraic multiplicity* 2*)*
- $$\left[\begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix} - 1 \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\right] \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \;\rightarrow\; \mathbf{x^{(1)}} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$
  and $V(-1) = \{\, t \cdot \mathbf{x^{(1)}} \mid t \in \mathbb{R} \,\}$ *with geometric multiplicity* 1*.*
- $$\left[\begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix} - 2 \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\right] \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$
  The 2*-dimensional vectorspace of all solutions is given by the single equation* $x_3 = 0$ *and there are infinitely many pairs of orthogonal vectors which span this space. We take the two standard vectors:*
  $$V(2) \;=\; \left\{\, t_1 \cdot \underbrace{\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}}_{\mathbf{x^{(2)}}} + t_2 \cdot \underbrace{\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}}_{\mathbf{x^{(3)}}} \mid t_1, t_2 \in \mathbb{R} \,\right\}$$

Example 2.2

$$A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -6 & -1 & 4 \end{pmatrix}$$

- $p_A(\lambda) \;=\; -\lambda^3 + 4\lambda^2 - \lambda - 6 \;=\; (\lambda+1)\cdot(-\lambda^2+5\lambda-6) \;=\; -(\lambda+1)\cdot(\lambda-2)\cdot(\lambda-3)$
- *Zeros of the characteristic polynomial:* $\lambda_1 = -1$, $\lambda_2 = 2$ *and* $\lambda_3 = 3$ *(all of algebraic multiplicity* 1*)*
- $$\left[\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -6 & -1 & 4 \end{pmatrix} - (-1)\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\right] \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \;\to\; \mathbf{x^{(1)}} = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$$

  and $V(-1) = \{\, t \cdot \mathbf{x^{(1)}} \mid t \in \mathbb{R} \,\}$ *with geometric multiplicity* 1.
- $$\left[\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -6 & -1 & 4 \end{pmatrix} - 2\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\right] \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \;\to\; \mathbf{x^{(2)}} = \begin{pmatrix} 1 \\ 2 \\ 4 \end{pmatrix}$$

  and $V(2) = \{\, t \cdot \mathbf{x^{(2)}} \mid t \in \mathbb{R} \,\}$ *with geometric multiplicity* 1.
- $$\left[\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -6 & -1 & 4 \end{pmatrix} - 3\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\right] \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \;\to\; \mathbf{x^{(3)}} = \begin{pmatrix} 1 \\ 3 \\ 9 \end{pmatrix}$$

  and $V(3) = \{\, t \cdot \mathbf{x^{(3)}} \mid t \in \mathbb{R} \,\}$ *with geometric multiplicity* 1.

Definition 2.2 *The* <u>*spectral radius*</u> *of a quadratic matrix* A *is the real number*

$$\rho(A) \;:=\; \max\{|\lambda_1|, \ldots, |\lambda_r|\,\}.$$

2.2 *Generalized Eigenvectors*

To solve some interesting problems we have to generalize the notion of eigenvectors.

Definition 2.3 *A vector* $\mathbf{x} \in \mathbb{C}^n$ *is called* *generalized eigenvector* *of degree* $l \in \mathbb{N}$ *associated to the eigenvalue* λ *of* $\mathbf{A}$*, if*

$$(\mathbf{A} - \lambda\mathbf{I})^l \, \mathbf{x} \;=\; \mathbf{0} \quad \textit{and} \quad (\mathbf{A} - \lambda\mathbf{I})^{l-1} \, \mathbf{x} \;\neq\; \mathbf{0}.$$

Of course, an eigenvector is a generalized eigenvector of degree 1.

Example 2.3 *The matrix*

$$\mathbf{A} \;=\; \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

has the eigenvalue 1 *of (algebraic) multiplicity* 3 *with* $\dim V(1) = 1$ *(geometric multiplicity). We have:*

$$\begin{array}{lll} (\mathbf{A} - \mathbf{I})\, \mathbf{e}_1 \;=\; \mathbf{0} & (\mathbf{A} - \mathbf{I})\, \mathbf{e}_2 \;=\; \mathbf{e}_1 & (\mathbf{A} - \mathbf{I})^2\, \mathbf{e}_2 \;=\; \mathbf{0} \\ (\mathbf{A} - \mathbf{I})\, \mathbf{e}_3 \;=\; \mathbf{e}_1 + \mathbf{e}_2 & (\mathbf{A} - \mathbf{I})^2\, \mathbf{e}_3 \;=\; \mathbf{e}_1 & (\mathbf{A} - \mathbf{I})^3\, \mathbf{e}_3 \;=\; \mathbf{0} \end{array}$$

This means, that $\mathbf{e}_1$ *is an eigenvector,* $\mathbf{e}_2$ *is a generalized eigenvector of degree* 2 *and* $\mathbf{e}_3$ *is a generalized eigenvector of degree* 3*.*

Theorem 2.2 *Let* $\mathbf{A} \in \mathbb{C}^{n\times n}$ *be a complex (or real) matrix with*

$$p_A(\lambda) \;=\; (\lambda_1 - \lambda)^{k_1} \, (\lambda_2 - \lambda)^{k_2} \ldots (\lambda_r - \lambda)^{k_r}.$$

- *Let* λ *be an eigenvalue of* $\mathbf{A}$ *of (algebraic) multiplicity* l*. Then there exist* l *linearly independent generalized eigenvectors (of degree* $\leq l$*). This means:*

  $$\dim\{\, \mathbf{x} \in \mathbb{C}^n \mid (\mathbf{A} - \lambda\mathbf{I})^l \, \mathbf{x} \;=\; \mathbf{0} \,\} \;=\; l.$$

- *Generalized eigenvectors associated to pairwise different eigenvalues of* $\mathbf{A}$ *are linearly independent.*

- *There exists a basis* $\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_n$ *of* $\mathbb{C}^n$ *consisting of generalized eigenvectors of* $\mathbf{A}$*. If* $\mathbf{P}$ *is the matrix with this basis as the columns, then*

  $$\mathbf{P}^{-1}\, \mathbf{A}\, \mathbf{P} \;=\; \begin{pmatrix} \boxed{\mathbf{A}_1} & & & \mathbf{0} \\ & \boxed{\mathbf{A}_2} & & \\ & & \ddots & \\ \mathbf{0} & & & \boxed{\mathbf{A}_r} \end{pmatrix}$$

  with $\mathbf{A}_i \in \mathbb{C}^{k_i\times k_i}$ *for all* $i = 1, 2, \ldots, r$*.*

Let us have a look at the case $n = 2$ and $\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$.

1. Characteristic polynomial:

$$\begin{aligned} p_A(\lambda) &= \det \begin{pmatrix} a-\lambda & b \\ c & d-\lambda \end{pmatrix} \\ &= \lambda^2 - \underbrace{(a+d)}_{=:tr(A)} \lambda + \underbrace{ad-bc}_{=\det(A)} = (\lambda_1 - \lambda)(\lambda_2 - \lambda) \end{aligned}$$

with $\lambda_{1,2} = \dfrac{a+d}{2} \pm \sqrt{\dfrac{(a+d)^2}{4} - \det(A)}$.

2. For each λ_i $(i = 1, 2)$ we solve the linear system

$$\begin{pmatrix} a-\lambda_i & b \\ c & d-\lambda_i \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

We have four different cases:

1. $\lambda_1, \lambda_2 \in \mathbb{R}$, $\lambda_1 \neq \lambda_2$

   Example: $\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$

   We have $p_A(\lambda) = (1-\lambda)^2 - 4 = (\lambda+1)(\lambda-3)$ (two different eigenvalues of algebraic multiplicity 1). A direct calculation shows, that $\dim V(-1) = 1$ and $\dim V(3) = 1$ and the geometric multiplicitiy are (of all eigenvalues) equal to the algebraic multiplicity.

2. $\lambda = \lambda_1 = \lambda_2 \in \mathbb{R}$ with $\dim V(\lambda) = 2$

   Example: $\mathbf{A} = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$

   We have $p_A(\lambda) = (2-\lambda)^2$ (one eigenvalue of algebraic multiplicity 2). A direct calculation shows, that $\dim V(2) = 2$ and the geometric multiplicitiy (of the eigenvalue 2) is equal to the algebraic multiplicity.

3. $\lambda = \lambda_1 = \lambda_2 \in \mathbb{R}$ with $\dim V(\lambda) = 1$

   Example: $\mathbf{A} = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}$

   We have $p_A(\lambda) = (2-\lambda)^2$ (one eigenvalue of algebraic multiplicity 2). A direct calculation shows, that $\dim V(2) = 1$ and the geometric multiplicitiy of the eigenvalue 2 is different of the algebraic multiplicity.

4. $\lambda_2 = \overline{\lambda_1} \in \mathbb{C} - \mathbb{R}$

   Example: $\mathbf{A} = \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix}$ with $\phi \neq k\pi$

   We have $p_A(\lambda) = (\lambda - \cos\phi)^2 + \sin^2\phi = \lambda^2 - 2\lambda\cos\phi + 1$ with the two different complex zeroes $\lambda_{1,2} = \cos\phi \pm i \sin\phi$.

3 Diagonalizable Matrices

3.1 Diagonalization

Let $\mathbf{A}$ and $\mathbf{P}$ be $n \times n$ matrices with $\mathbf{P}$ invertible. Then $\mathbf{A}$ and $\mathbf{P}^{-1}\mathbf{A}\mathbf{P}$ have the same eigenvalues (because they have the same characteristic polynomial).

Definition 3.1 *An $n \times n$ matrix $\mathbf{A}$ is diagonalizable if there is an invertible matrix $\mathbf{P}$ and a diagonal matrix $\mathbf{D}$ such that*

$$\mathbf{P}^{-1}\mathbf{A}\mathbf{P} \;=\; \mathbf{D}.$$

Two natural questions:

1. Which square matrices are diagonalizable?
2. If $\mathbf{A}$ is diagonalizable, how do we find the matrix $\mathbf{P}$?

Theorem 3.1 *An $n \times n$ matrix $\mathbf{A}$ is diagonalizable if and only if it has a set of n linearly independent eigenvectors $\mathbf{p}_1, \ldots, \mathbf{p}_n$. In this case,*

$$\mathbf{P}^{-1}\mathbf{A}\mathbf{P} \;=\; \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix},$$

where $\mathbf{P}$ is the matrix with $\mathbf{p}_1, \ldots, \mathbf{p}_n$ as its columns, and $\lambda_1, \ldots, \lambda_n$ are the corresponding eigenvalues.

Proof: We prove only one direction of the statement:

$\mathbf{A}$ has n linearly independent eigenvectors $\Longrightarrow$ $\mathbf{A}$ is diagonalizable.

Let $\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_n$ be the n linearly independent eigenvectors of $\mathbf{A}$ with corresponding eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$. We form the matrix

$$\mathbf{P} = \begin{pmatrix} | & | & & | \\ \mathbf{p}_1 & \mathbf{p}_2 & \ldots & \mathbf{p}_n \\ | & | & & | \end{pmatrix}$$

with the eigenvectors of $\mathbf{A}$ as the columns. Then

$$\mathbf{A}\mathbf{P} \;=\; \begin{pmatrix} | & | & & | \\ \mathbf{A}\mathbf{p}_1 & \mathbf{A}\mathbf{p}_2 & \ldots & \mathbf{A}\mathbf{p}_n \\ | & | & & | \end{pmatrix}$$

the column vectors of $\mathbf{AP}$ are the vectors $\mathbf{Ap}_1, \mathbf{Ap}_2, \ldots, \mathbf{Ap}_n$. Using the property of eigenvectors, we get

$$\begin{aligned}
\mathbf{AP} &= \begin{pmatrix} | & | & & | \\ \mathbf{Ap}_1 & \mathbf{Ap}_2 & \ldots & \mathbf{Ap}_n \\ | & | & & | \end{pmatrix} \\
&= \begin{pmatrix} | & | & & | \\ \lambda_1\mathbf{p}_1 & \lambda_2\mathbf{p}_2 & \ldots & \lambda_n\mathbf{p}_n \\ | & | & & | \end{pmatrix} \\
&= \begin{pmatrix} | & | & & | \\ \mathbf{p}_1 & \mathbf{p}_2 & \ldots & \mathbf{p}_n \\ | & | & & | \end{pmatrix} \begin{pmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix} \\
&= \mathbf{PD}.
\end{aligned}$$

where $\mathbf{D}$ is the diagonal matrix with diagonal entries equal to the eigenvalues of $\mathbf{A}$. The matrix $\mathbf{P}$ has maximal rank (and is invertible), because the column vectors are linearly independent. Hence the equation $\mathbf{AP} = \mathbf{PD}$ is equivalent to $\mathbf{P}^{-1}\mathbf{AP} = \mathbf{D}$.

□

Example 3.1 *The matrix* $\mathbf{A} = \begin{pmatrix} 1 & 1 \\ -2 & 4 \end{pmatrix}$ *has the eigenvalues and eigenvectors*

$$\lambda_1 = 2 \qquad \mathbf{p}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

$$\lambda_2 = 3 \qquad \mathbf{p}_2 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

Hence $\mathbf{P} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$, $\mathbf{P}^{-1} = \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix}$ *and:*

$$\mathbf{P}^{-1}\mathbf{AP} = \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ -2 & 4 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}$$

3.2 Spectral Theorem for symmetric matrices

Many matrices encountered in economics are (real) symmetric and for these matrices we have the following important result.

Theorem 3.2 (Spectral Theorem for symmetric matrices) *If the real $n \times n$ matrix $\mathbf{A}$ is symmetric ($\mathbf{A} = \mathbf{A}^T$), then:*

1. *All n eigenvalues $\lambda_1, \ldots, \lambda_n$ are real.*
2. *Eigenvectors that correspond to different eigenvalues are orthogonal.*
3. *There exists an <u>orthogonal</u> and real matrix $\mathbf{P}$ ($\mathbf{P}^{-1} = \mathbf{P}^T$) such that*

$$\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}.$$

The columns $\mathbf{p}_1, \ldots, \mathbf{p}_n$ of the matrix $\mathbf{P}$ are eigenvectors of unit length corresponding to the eigenvalues $\lambda_1, \ldots, \lambda_n$.

Proof: Let $\mathbf{A}$ be a real and symmetric $n \times n$ matrix.

1. Let $\mathbf{A}\mathbf{p}_i = \lambda_i \mathbf{p}_i$. By complex conjugation of this equation (complex conjugate all entries of the vector and matrix, but keep in mind that $\mathbf{A}$ has only real entries) we get

$$\overline{\mathbf{A}\mathbf{p}_i} = \overline{\mathbf{A}}\overline{\mathbf{p}_i} = \mathbf{A}\overline{\mathbf{p}_i} = \overline{\lambda_i}\overline{\mathbf{p}_i}$$

and

$$\lambda_i \mathbf{p}_i^T \overline{\mathbf{p}_i} = (\mathbf{A}\mathbf{p}_i)^T \overline{\mathbf{p}_i} = \mathbf{p}_i^T \mathbf{A}^T \overline{\mathbf{p}_i} = \mathbf{p}_i^T \mathbf{A} \overline{\mathbf{p}_i} = \mathbf{p}_i^T \overline{\lambda_i} \overline{\mathbf{p}_i} = \overline{\lambda_i} \mathbf{p}_i^T \overline{\mathbf{p}_i}$$

Because $\mathbf{p}_i^T \overline{\mathbf{p}_i} = ||\mathbf{p}_i||^2 \neq 0$, we have $\lambda_i = \overline{\lambda_i}$ and λ_i must be a real number.

2. Let $\mathbf{A}\mathbf{p}_i = \lambda_i \mathbf{p}_i$ and $\mathbf{A}\mathbf{p}_j = \lambda_j \mathbf{p}_j$ with $\lambda_i \neq \lambda_j$. Then

$$\begin{aligned} \lambda_i \, \mathbf{p}_i^T \mathbf{p}_j &= (\mathbf{A}\mathbf{p}_i)^T \mathbf{p}_j \\ &= \mathbf{p}_i^T \mathbf{A}^T \mathbf{p}_j \\ &= \mathbf{p}_i^T (\mathbf{A}^T \mathbf{p}_j) \\ &= \mathbf{p}_i^T (\mathbf{A}\mathbf{p}_j) && \text{because } \mathbf{A} = \mathbf{A}^T \\ &= \mathbf{p}_i^T \lambda_j \mathbf{p}_j \\ &= \lambda_j \, \mathbf{p}_i^T \mathbf{p}_j \end{aligned}$$

or

$$\lambda_i \, (\mathbf{p}_i^T \mathbf{p}_j) \;=\; \lambda_j \, (\mathbf{p}_i^T \mathbf{p}_j)$$

and because $\lambda_i \neq \lambda_j$, the scalar product of $\mathbf{p}_i$ and $\mathbf{p}_j$ must be zero: $\mathbf{p}_i^T \mathbf{p}_j = \mathbf{p}_i \bullet \mathbf{p}_j = 0$. Hence the two eigenvectors are orthogonal.

3. We give the proof of part 3 only for the case that all eigenvalues $\lambda_1, \ldots, \lambda_n$ are (pairwise) different (and real by part 1). In this case, the corresponding eigenvectors $\mathbf{p}'_1, \ldots, \mathbf{p}'_n$ are orthogonal (by part 2) and hence linearly independent. Now choose for $i = 1, \ldots, n$ an eigenvector of length 1 by

$$\mathbf{p}_i \;:=\; \frac{1}{||\mathbf{p}'_i||} \, \mathbf{p}'_i$$

It is easy to show, that

$$\mathbf{p}_i^T \mathbf{p}_j \;=\; \mathbf{p}_i \bullet \mathbf{p}_j \;=\; \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

The matrix

$$\mathbf{P} = \begin{pmatrix} | & | & & | \\ \mathbf{p}_1 & \mathbf{p}_2 & \cdots & \mathbf{p}_n \\ | & | & & | \end{pmatrix}$$

is an orthogonal matrix, because

$$\mathbf{P}^T \mathbf{P} \;=\; \begin{pmatrix} - & \mathbf{p}_1^T & - \\ - & \mathbf{p}_2^T & - \\ & \cdots & \\ - & \mathbf{p}_n^T & - \end{pmatrix} \begin{pmatrix} | & | & & | \\ \mathbf{p}_1 & \mathbf{p}_2 & \cdots & \mathbf{p}_n \\ | & | & & | \end{pmatrix}$$

$$= \begin{pmatrix} \mathbf{p}_1^T \mathbf{p}_1 & \mathbf{p}_1^T \mathbf{p}_2 & \cdots & \mathbf{p}_1^T \mathbf{p}_n \\ \mathbf{p}_2^T \mathbf{p}_1 & \mathbf{p}_2^T \mathbf{p}_2 & \cdots & \mathbf{p}_2^T \mathbf{p}_n \\ \cdots & \cdots & \ddots & \cdots \\ \mathbf{p}_n^T \mathbf{p}_1 & \mathbf{p}_n^T \mathbf{p}_2 & \cdots & \mathbf{p}_n^T \mathbf{p}_n \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \cdots & \cdots & \ddots & \cdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}.$$

Hence we have $\mathbf{P}^T = \mathbf{P}^{-1}$. From Theorem 3.1 we already know, that

$$\mathbf{P}^{-1} \mathbf{A} \mathbf{P} \;=\; \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}.$$

if $\mathbf{P}$ has a set of linearly independent eigenvectors of $\mathbf{A}$ as columns.

$\square$

3.3 Spectral decomposition of symmetric matrices

Theorem 3.3 (Spectral decomposition of symmetric matrices) *Let* $\mathbf{A}$ *be a symmetric matrix with the set* $\mathbf{p}_1, \ldots, \mathbf{p}_n$ *of orthonormal eigenvectors associated to the real eigenvalues* $\lambda_1, \ldots, \lambda_n$*. Then* $\mathbf{A}$ *can be written as*

$$\boxed{\mathbf{A} \;=\; \sum_{i=1}^{n} \lambda_i \mathbf{p}_i \mathbf{p}_i^T \;=\; \lambda_1 \mathbf{p}_1 \mathbf{p}_1^T \;+ \cdots +\; \lambda_n \mathbf{p}_n \mathbf{p}_n^T}$$

Proof: For each vector $\mathbf{p}_j$ (of the given ONB) we have

$$\mathbf{A}\mathbf{p}_j \;=\; \lambda_j \mathbf{p}_j$$

and

$$\left(\sum_{i=1}^{n} \lambda_i \mathbf{p}_i \mathbf{p}_i^T \right) \mathbf{p}_j \;=\; \sum_{i=1}^{n} \lambda_i \mathbf{p}_i \left(\mathbf{p}_i^T \mathbf{p}_j \right) \;=\; \lambda_j \mathbf{p}_j .$$

Both matrices $\mathbf{A}$ and $\left(\sum_{i=1}^{n} \lambda_i \mathbf{p}_i \mathbf{p}_i^T \right)$ have the same values, if we multiply with any basis vector. By linearrity of matrix multiplication, they must have the same values if we multiply with an arbitrary vector (expressed in the given basis). Hence, both matrices must be equal. $\square$

Example 3.2 *The matrix* $\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$ *is symmetric and has the eigenvalues* $\lambda_1 = -1$ *and* $\lambda_2 = 3$. *The corresponding eigenspaces are*

$$\begin{aligned} V(-1) &= \left\{ t \begin{pmatrix} 1 \\ -1 \end{pmatrix} \mid t \in \mathbb{R} \right\} \\ V(3) &= \left\{ t \begin{pmatrix} 1 \\ 1 \end{pmatrix} \mid t \in \mathbb{R} \right\} \end{aligned}$$

The two eigenspaces are orthogonal, because the scalar product of the two spanning vectors is 0. *In order to construct the matrix* $\mathbf{P}$, *we have to use eigenvectors of length* 1 *(unit vectors). A spanning vector of length* 1 *for* $V(-1)$ *is*

$$\mathbf{p}_1 = \frac{1}{\sqrt{1^2 + (-1)^2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

and for $V(3)$ *is*

$$\mathbf{p}_2 = \frac{1}{\sqrt{1^2 + 1^2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

Hence $\mathbf{P} = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$ *is an orthogonal matrix, because* $\mathbf{P}^{-1} = \mathbf{P}^T = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$ *and*

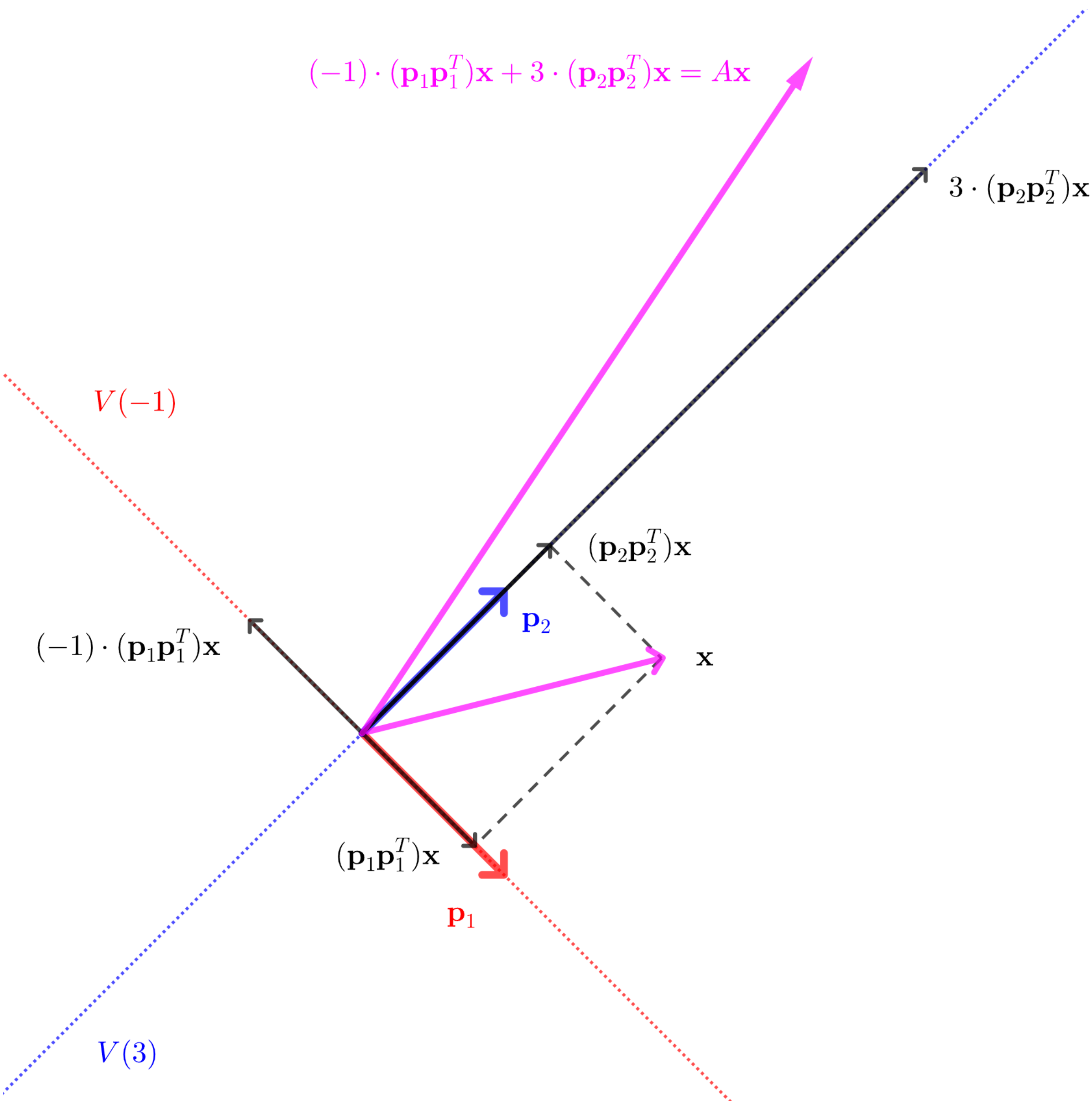
$$\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & 3 \end{pmatrix}$$

Furthermore

$$\begin{aligned} \mathbf{p}_1\mathbf{p}_1^T &= \begin{pmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{pmatrix} \begin{pmatrix} 1/\sqrt{2} & -1/\sqrt{2} \end{pmatrix} = \begin{pmatrix} 1/2 & -1/2 \\ -1/2 & 1/2 \end{pmatrix} \\ \mathbf{p}_2\mathbf{p}_2^T &= \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix} \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix} = \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix} \end{aligned}$$

and the spectral decomposition of $\mathbf{A}$ *ist given by*

$$\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} = -1 \cdot \begin{pmatrix} 1/2 & -1/2 \\ -1/2 & 1/2 \end{pmatrix} + 3 \cdot \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix}$$


$(-1)\cdot(\mathbf{p}_1\mathbf{p}_1^T)\mathbf{x} + 3\cdot(\mathbf{p}_2\mathbf{p}_2^T)\mathbf{x} = A\mathbf{x}$
$3\cdot(\mathbf{p}_2\mathbf{p}_2^T)\mathbf{x}$
$V(-1)$
$(\mathbf{p}_2\mathbf{p}_2^T)\mathbf{x}$
$\mathbf{p}_2$
$(-1)\cdot(\mathbf{p}_1\mathbf{p}_1^T)\mathbf{x}$
$\mathbf{x}$
$(\mathbf{p}_1\mathbf{p}_1^T)\mathbf{x}$
$\mathbf{p}_1$
$V(3)$

4 Quadratic forms and matrices

Definition 4.1 *A quadratic form in n variables $\mathbf{x} = (x_1, \ldots, x_n)^T$ is a function of the form*

$$Q_{\mathbf{A}}(\mathbf{x}) \;=\; \sum_{i,j=1}^{n} a_{ij} x_i x_j \;=\; \mathbf{x}^T \mathbf{A} \mathbf{x}$$

where $\mathbf{A} = (a_{ij})$ is an $n \times n$ matrix.

Quadratic forms are important examples of multi-variate functions and $Q_{\mathbf{A}}$ is a homogeneous function of degree 2 in n variables. Of course, $Q_{\mathbf{A}}(\mathbf{0}) = 0$ for all quadratic forms.

Example 4.1 $Q(x_1, x_2) = x_1^2 + x_1 x_2 + x_2^2$ *is a quadratic form and can be written as*

$$\begin{aligned}(x_1\ x_2)\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \;&=\; (x_1\ x_2)\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \\ &=\; (x_1\ x_2)\begin{pmatrix} 1 & 1/2 \\ 1/2 & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \;=\; \ldots..\end{aligned}$$

Unfortunaetly, there is no unique way to write a given quadratic form in matrix term. But we may resolve this situation by **always choosing A to be symmetric!**

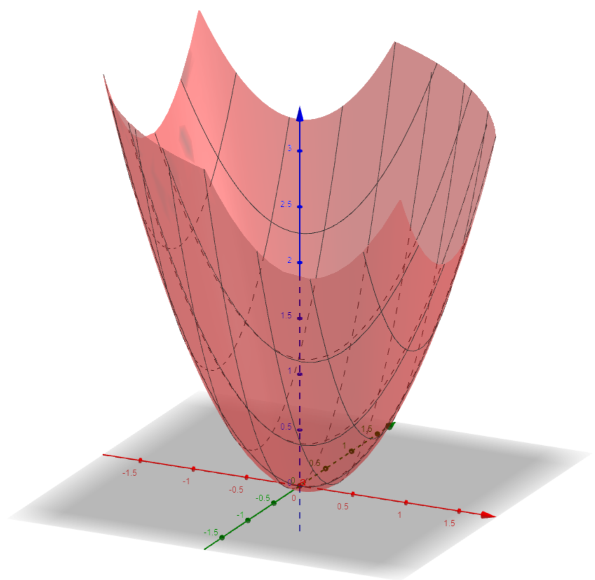
Exercise 4.1 *Let $Q(\mathbf{x}) = \mathbf{x}^T \mathbf{B} \mathbf{x}$ where $\mathbf{B}$ is not symmetric. Let $\mathbf{A} = (\mathbf{B} + \mathbf{B}^T)/2$ and $\mathbf{C} = (\mathbf{B} - \mathbf{B}^T)/2$. Show that $\mathbf{A}$ is symmetric and evaluate both $\mathbf{x}^T \mathbf{A} \mathbf{x}$ and $\mathbf{x}^T \mathbf{C} \mathbf{x}$.*

Example 4.2

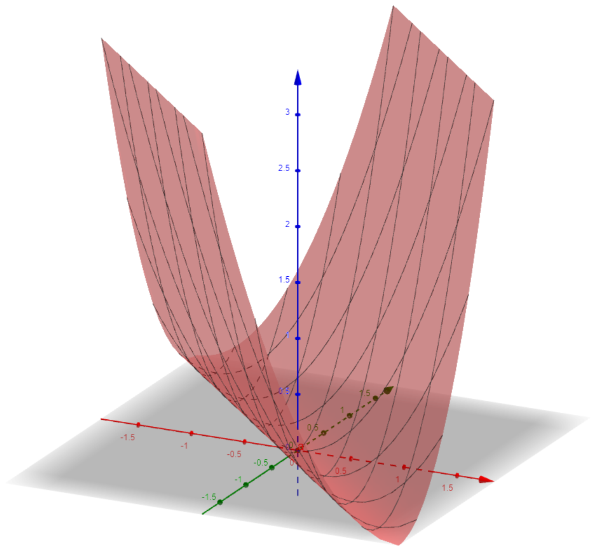
- *The quadratic form $Q(x_1, x_2) = x_1^2 + x_1 x_2 + x_2^2$ can be written as*

  $$\left(x_1 + \frac{x_2}{2}\right)^2 + \frac{3}{4} x_2^2.$$

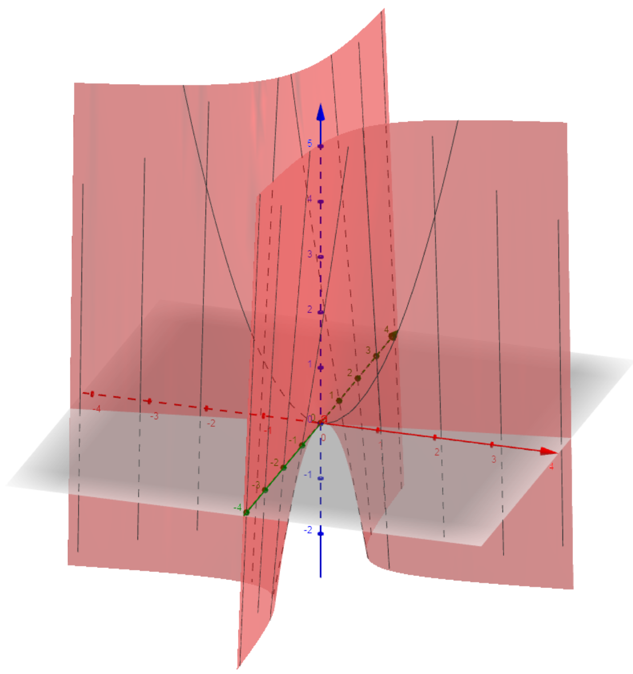
  As a sum of squares, it can not be negative and can only be zero when $x_1 + \frac{x_2}{2} = 0$ and $x_2 = 0$, or $x_1 = x_2 = 0$. We call this a positive definite quadratic form.



- *The quadratic form* $Q(x_1, x_2) = x_1^2 + 2x_1x_2 + x_2^2 \;=\; (x_1 + x_2)^2$ *is always non-negative, but it is zero whenever* $x_1 + x_2 = 0$ *or* $x_1 = -x_2$ *(it is zero for non-zero values of the variables). We call this a positive semi-definite quadratic form.*



- *The quadratic form* $Q(x_1, x_2) = x_1^2 - 6x_1x_2 \;=\; (x_1 - 3x_2)^2 - 9x_2^2$ *can be positive or negative. We call this an indefinite quadratic form.*



Definition 4.2 *A quadratic form* $Q_{\mathbf{A}}(\mathbf{x}) = \mathbf{x}^T\mathbf{A}\mathbf{x}$*, as well as its associated symmetric matrix* $\mathbf{A}$*, is said to be*

$$\begin{array}{rcl} \underline{\textit{positive definite}} & :\Longleftrightarrow & Q_{\mathbf{A}}(\mathbf{x}) > 0 \\ \underline{\textit{positive semi-definite}} & :\Longleftrightarrow & Q_{\mathbf{A}}(\mathbf{x}) \geq 0 \\ \underline{\textit{negative definite}} & :\Longleftrightarrow & Q_{\mathbf{A}}(\mathbf{x}) < 0 \\ \underline{\textit{negative semi-definite}} & :\Longleftrightarrow & Q_{\mathbf{A}}(\mathbf{x}) \leq 0 \end{array}$$

for **all** $\mathbf{x} \neq \mathbf{0}$.

The quadratic form is called <u>*indefinite*</u>*, if there are vectors* $\mathbf{a}$ *and* $\mathbf{b}$ *with* $Q_{\mathbf{A}}(\mathbf{a}) < 0$ *and* $Q_{\mathbf{A}}(\mathbf{b}) > 0$.

It is easy to see, that for $i = 1, \ldots, n$:

$$\boxed{Q_{\mathbf{A}}(\mathbf{e}_i) = a_{ii}.}$$

The technique used in the examples to examine the sign of the quadratic form is known as **completing the squares**. Let us examine the possible signs of a quadratic form $Q_{\mathbf{A}}(\mathbf{x}) = \mathbf{x}^T\mathbf{A}\mathbf{x}$ using the eigenvalues/eigenvectors of the **symmetric** matrix $\mathbf{A}$.

By the **Spectral Theorem for symmetric matrices** we can choose a matrix $\mathbf{P}$ of eigenvectors $\mathbf{p}_1, \ldots, \mathbf{p}_n$ of $\mathbf{A}$, such that $\mathbf{P}^{-1} = \mathbf{P}^T$ and

$$\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{P}^T\mathbf{A}\mathbf{P} = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix},$$

where $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of $\mathbf{A}$.

Now let $\mathbf{y} := \mathbf{P}^T\mathbf{x}$. This defines new variables $y_1, \ldots, y_n$ as linear combinations of the old ones

$$y_i = \sum_{j=1}^{n} p_{ji}x_j.$$

Further, since $\mathbf{P}\mathbf{P}^T = \mathbf{I}$ we have $\mathbf{x} = \mathbf{P}\mathbf{y}$ and

$$\begin{aligned} Q_{\mathbf{A}}(\mathbf{x}) &= \mathbf{x}^T\mathbf{A}\mathbf{x} \\ &= (\mathbf{P}\mathbf{y})^T\mathbf{A}(\mathbf{P}\mathbf{y}) \\ &= \mathbf{y}^T\mathbf{P}^T\mathbf{A}\mathbf{P}\mathbf{y} \\ &= \mathbf{y}^T(\mathbf{P}^T\mathbf{A}\mathbf{P})\mathbf{y} \\ &= \mathbf{y}^T \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} \mathbf{y} \\ &= \sum_{i=1}^{n} \lambda_i y_i^2. \end{aligned}$$

Thus we completed the squares. The quadratic form is expressed in terms of the new variables as a sum/difference of pure square terms. To determine the sign of the quadratic form, we simply inspect the signs of the eigenvalues of $\mathbf{A}$.

Theorem 4.1 (Sylvester)

If $\mathbf{A}$ is symmetric, then the quadratic form $Q_{\mathbf{A}}(\mathbf{x}) = \mathbf{x}^T\mathbf{A}\mathbf{x}$ is

$$
\begin{array}{rcl}
\underline{\textit{positive definite}} & \Longleftrightarrow & \forall\ \lambda_i > 0 \\
\underline{\textit{positive semi-definite}} & \Longleftrightarrow & \forall\ \lambda_i \geq 0 \\
\underline{\textit{negative definite}} & \Longleftrightarrow & \forall\ \lambda_i < 0 \\
\underline{\textit{negative semi-definite}} & \Longleftrightarrow & \forall\ \lambda_i \leq 0 \\
\underline{\textit{indefinite}} & \Longleftrightarrow & \exists\ \lambda_i > 0 \textit{ and } \lambda_j < 0.
\end{array}
$$

Checking eigenvalues can be tedious. There is a convenient condition on the matrix $\mathbf{A}$ in terms of certain sub-determinants, which can be used to identify the definiteness of $\mathbf{A}$.

An arbitrary principal minor of order r of an $n \times n$ matrix $\mathbf{A}$ is the determinant of a matrix obtained by deleting $n - r$ rows and $n - r$ columns of $\mathbf{A}$ such that if the ith row (column) is selected then so is the ith column (row). A principal minor is called a leading principal minor of order r if it consists of the first (leading) r rows and columns of $\mathbf{A}$.

Theorem 4.2

Let $\mathbf{A}$ be a symmetric $n \times n$ matrix. We denote by D_r the leading principal minor of order r and let Δ_r denote an arbitrary principal minor of order r. Then the quadratic form $Q_{\mathbf{A}}(\mathbf{x}) = \mathbf{x}^T\mathbf{A}\mathbf{x}$ is

$$
\begin{array}{rcl}
\underline{\textit{positive definite}} & \Longleftrightarrow & D_r > 0 \textit{ for } r = 1, \ldots, n \\
\underline{\textit{positive semi-definite}} & \Longleftrightarrow & \Delta_r \geq 0 \textit{ for all principal minors of order } r = 1, \ldots, n \\
\underline{\textit{negative definite}} & \Longleftrightarrow & (-1)^r D_r > 0 \textit{ for } r = 1, \ldots, n \\
\underline{\textit{negative semi-definite}} & \Longleftrightarrow & (-1)^r \Delta_r \geq 0 \textit{ for all principal minors of order } r = 1, \ldots, n.
\end{array}
$$

Example 4.3 *Let*

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

The principal minors of $\mathbf{A}$ *are* $\det(\mathbf{A})$, $\det\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$, $\det\begin{pmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{pmatrix}$, $\det\begin{pmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{pmatrix}$, a_{11}, a_{22} *and* a_{33}.

The leading principal minors are a_{11}, $\det\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ *and* $\det(\mathbf{A})$.

Example 4.4

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & 3 \\ 0 & 1 & 2 \\ 3 & 2 & 13 \end{pmatrix}$$

$\mathbf{A}$ is not negativ semi-definit (and not negativ definit), because $Q_{\mathbf{A}}(1,0,0) = 1$, $Q_{\mathbf{A}}(0,1,0) = 1$ or $Q_{\mathbf{A}}(0,0,1) = 13$. It may be positive semi-definit (positiv definit) or indefinit.

- *all principal minors of order* 1*:*

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & 3 \\ 0 & 1 & 2 \\ 3 & 2 & 13 \end{pmatrix} \qquad \det(1) \geq 0$$

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & 3 \\ 0 & 1 & 2 \\ 3 & 2 & 13 \end{pmatrix} \qquad \det(1) \geq 0$$

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & 3 \\ 0 & 1 & 2 \\ 3 & 2 & 13 \end{pmatrix} \qquad \det(13) \geq 0$$

- *all pricipal minors of order* 2

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & 3 \\ 0 & 1 & 2 \\ 3 & 2 & 13 \end{pmatrix} \qquad \det\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = 1 \geq 0$$

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & 3 \\ 0 & 1 & 2 \\ 3 & 2 & 13 \end{pmatrix} \qquad \det\begin{pmatrix} 1 & 3 \\ 3 & 13 \end{pmatrix} = 4 \geq 0$$

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & 3 \\ 0 & 1 & 2 \\ 3 & 2 & 13 \end{pmatrix} \qquad \det\begin{pmatrix} 1 & 2 \\ 2 & 13 \end{pmatrix} = 9 \geq 0$$

- *all principal minors of order* 3

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & 3 \\ 0 & 1 & 2 \\ 3 & 2 & 13 \end{pmatrix} \qquad \det\begin{pmatrix} 1 & 0 & 3 \\ 0 & 1 & 2 \\ 3 & 2 & 13 \end{pmatrix} = 0 \geq 0$$

Hence $\mathbf{A}$ is positive semi-definite and not positive definite.

Special case: $n = 2$ The quadratic form

$$Q_{\mathbf{A}}(\mathbf{x}) = (x_1\ x_2)\begin{pmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2$$

- is positive definite if $a_{11} > 0$ and $\det \mathbf{A} = a_{11}a_{22} - a_{12}^2 > 0$;
- is positive semi-definite if $a_{11} \geq 0$, $a_{22} \geq 0$ and $\det \mathbf{A} = a_{11}a_{22} - a_{12}^2 \geq 0$;
- is negative definite if $a_{11} < 0$ and $\det \mathbf{A} = a_{11}a_{22} - a_{12}^2 > 0$;
- is negative semi-definite if $a_{11} \leq 0$, $a_{22} \leq 0$ and $\det \mathbf{A} = a_{11}a_{22} - a_{12}^2 \geq 0$.