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>

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

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

$$\begin{aligned} \mathbf{x} \bullet \mathbf{y} &= x_1 y_1 + x_2 y_2 + \cdots + x_n y_n \\ ||\mathbf{x}|| &= \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2} \end{aligned}$$

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

- $\mathbf{x_1}, \mathbf{x_2}, \dots, \mathbf{x_k} \in \mathbb{R}^n$
  - If $a_1, a_2, \dots, a_k \in \mathbb{R}$, then $\mathbf{z} = a_1\mathbf{x_1} + a_2\mathbf{x_2} + \cdots + a_k\mathbf{x_k}$ is called a <u>linear combination</u> of $\mathbf{x_1}, \mathbf{x_2}, \dots, \mathbf{x_k}$.
  - $\mathbf{x_1}, \mathbf{x_2}, \dots, \mathbf{x_k}$ are called <u>linearly dependent</u>, if there exist $b_1, b_2, \dots, b_k \in \mathbb{R}$ such that $b_1\mathbf{x_1} + b_2\mathbf{x_2} + \cdots + b_k\mathbf{x_k} = \mathbf{0}$ <u>and</u> not all $b_j = 0$.
  - $\mathbf{x_1}, \mathbf{x_2}, \dots, \mathbf{x_k}$ are called <u>linearly independent</u>, 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 $b_1 = b_2 = \cdots = b_k = 0$.

1.2 Real Matrices

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

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

is called an <u>$n \times m$ matrix</u>.

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

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

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

- 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 <u>determinant</u> $\det \mathbf{A}$ is given (recursively) by

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

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

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 <u>linear transformation</u> 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:*

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

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

$$T_{\mathbf{A}}(\mathbf{x}) \;=\; \mathbf{A} \cdot \mathbf{x} \;=\; x_1 \mathbf{a}_1 + \cdots + x_m \mathbf{a}_m.$$

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

1.4 Complex matrices and vectors

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:

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

1.5 Matrix calculus

1a. $\mathbf{A}+\mathbf{B}=\mathbf{B}+\mathbf{A}$ 1b. $\mathbf{AB} \neq \mathbf{BA}$
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. $\mathbf{AI}=\mathbf{IA}=\mathbf{A}$, ($\mathbf{A}$ square)

4. $\mathbf{AB}=\mathbf{0} \quad \not\Rightarrow \quad \mathbf{A}=\mathbf{0}$ or $\mathbf{B}=\mathbf{0}$
5. $\mathbf{AB}=\mathbf{AC} \quad \not\Rightarrow \quad \mathbf{B}=\mathbf{C}$

6. $\lambda(\mathbf{A}+\mathbf{B}) = \lambda\mathbf{A}+\lambda\mathbf{B} \quad \lambda \in \mathbb{R}$
7. $\mathbf{A}(\mathbf{B}+\mathbf{C}) = \mathbf{AB}+\mathbf{AC}$
8. $(\mathbf{A}+\mathbf{B})\mathbf{C} = \mathbf{AC}+\mathbf{BC}$

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

For $\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ with $ad-bc \neq 0$ is $\mathbf{A}^{-1} = \frac{1}{ad-bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$.

All these definitions and results can be generalized to vectors and matrices with complex entries.

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* <u>*eigenvalue*</u> *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* <u>*eigenvector*</u> *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 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 linear system of 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 it has determinant equal 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 <u>characteristic equation</u> of $\mathbf{A}$. The function $p_A(\lambda)$ is a polynomial of degree n in λ, called the <u>characteristic polynomial</u> of $\mathbf{A}$.

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 so called <u>eigenspace</u> for λ_i
$$V(\lambda_i) \;=\; \{\, \mathbf{x} \in \mathbb{C}^n \mid (\mathbf{A} - \lambda_i\mathbf{I})\,\mathbf{x} = \mathbf{0} \,\}.$$
The dimension of the vector space $V(\lambda_i)$ is called the <u>geometric multiplicity</u> of the eigenvalue λ_i.

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* <u>*generalized eigenvector*</u> *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 2.1 *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, $\mathbf{e}_2$ a generalized eigenvector of degree 2 *and $\mathbf{e}_3$ a generalized eigenvector of degree* 3.

Theorem 2.1 *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 $i = 1, 2, \ldots, r$.

Example 2.2 *Let* $n = 2$ *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 $\lambda_{1,2} = \frac{a+d}{2} \pm \sqrt{\frac{(a+d)^2}{4} - \det(A)}$.

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

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

If $n = 2$, *we have four different cases:*

1. $\lambda_1, \lambda_2 \in \mathbb{R}$, $\lambda_1 \neq \lambda_2$
   Example: $\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$
2. $\lambda = \lambda_1 = \lambda_2 \in \mathbb{R}$ *with* $\dim V(\lambda) = 2$
   Example: $\begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$
3. $\lambda = \lambda_1 = \lambda_2 \in \mathbb{R}$ *with* $\dim V(\lambda) = 1$
   Example: $\begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}$
4. $\lambda_2 = \overline{\lambda_1} \in \mathbb{C} - \mathbb{R}$
   Example: $\begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix}$ *with* $\phi \neq k\pi$

3 Diagonalization

Let $\mathbf{A}$ and $\mathbf{P}$ be $n \times n$ matrices with $\mathbf{P}$ invertible. Then $\mathbf{A}$ and $\mathbf{P^{-1}AP}$ 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*

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

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,*

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

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.

Many of the 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 orthogonal and real matrix $\mathbf{P}$ ($\mathbf{P^{-1}} = \mathbf{P^T}$) such that*

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

The columns $\mathbf{p_1}, \ldots, \mathbf{p_n}$ of the matrix $\mathbf{P}$ are eigenvectors of unit length corresponding to the eigenvalues $\lambda_1, \ldots, \lambda_n$.

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:*

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

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 $n \times n$ matrix.

Quadratic forms are important examples of multi-variate functions.

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*

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

Unfortunatly, 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

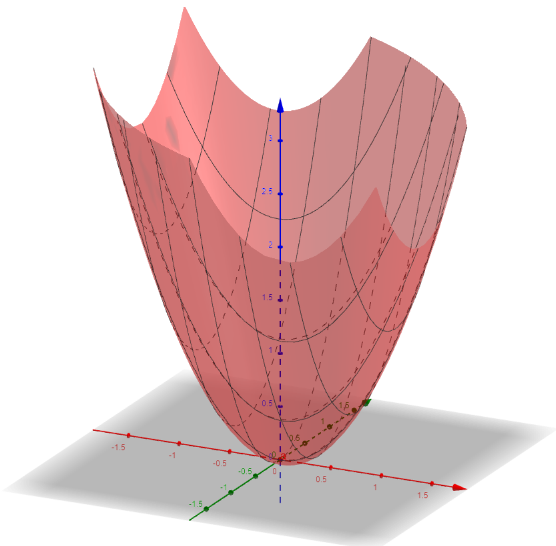
- *The quadratic form* $Q(x_1, x_2) = x_1^2 + x_1x_2 + x_2^2$ *can be written as*

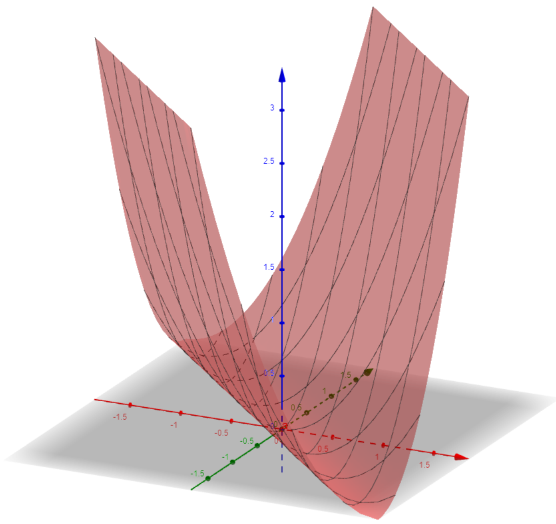
$$\left(x_1 + \frac{x_2}{2}\right)^2 + \frac{3}{4}x_2^2.$$

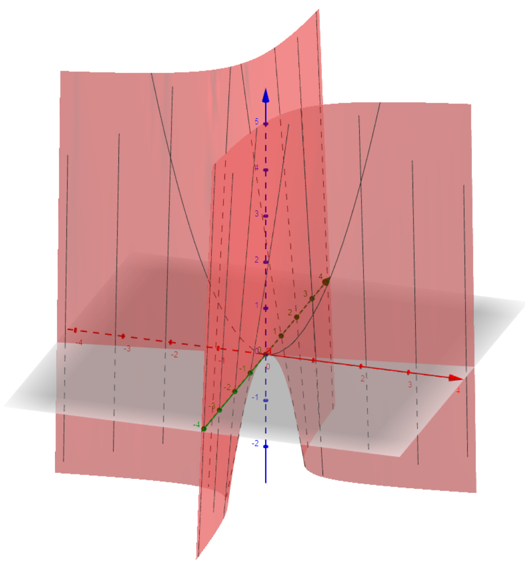
  As a sum of squares, it can not be negative and can only be zero when $x_1 + \frac{x_2}{2} = 0$ *and* $x_2 = 0$*, or* $x_1 = x_2 = 0$*. We call this a positive definite quadratic form.*

- *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 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 quadrtic 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 quadrtic 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 problem of signing 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

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

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

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

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 definitness 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 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 $\mathbf{A}$.

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* $\det(\mathbf{A})$.

Theorem 4.2
Let $\mathbf{A}$ be a symmetric $n \times n$ matrix. We denote by D_k the leading principal minor of order k and let Δ_k denote an arbitrary principal minor of order k. Then the quadratic form $Q_{\mathbf{A}}(\mathbf{x}) = \mathbf{x}^T\mathbf{A}\mathbf{x}$ is

$$\begin{aligned} \underline{\textit{positive definite}} &\iff D_k > 0 \textit{ for } k = 1, \ldots, n \\ \underline{\textit{positive semi-definite}} &\iff \Delta_k \geq 0 \textit{ for all principal minors of order } k = 1, \ldots, n \\ \underline{\textit{negative definite}} &\iff (-1)^k D_k > 0 \textit{ for } k = 1, \ldots, n \\ \underline{\textit{negative semi-definite}} &\iff (-1)^k \Delta_k \geq 0 \textit{ for all principal minors of order } k = 1, \ldots, n. \end{aligned}$$

Special case $n = 2$ The quadratic form

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

- is positiv definite if $a_{11} > 0$ and $\det \mathbf{A} = a_{11}a_{22} - a_{12}^2 > 0$;
- is positiv 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 $\det \mathbf{A} = a_{11}a_{22} - a_{12}^2 \geq 0$;