AMAT 111

Lecture Notes Sheets

Course Outline

PART ONE: Linear Algebra

1.1 Vectors in R^n and Matrix Algebra

1.1.1 Vectors in R^n

1.1.2 Vectors Spaces

1.1.3 Matrices

1.1.3.1 Determinants

PART ONE

Basics of Linear Algebra

1.1 Vectors in $\mathbb{R}^n$ and Matrix Algebra

1.1.1 Vectors

- $\mathbb{R}^n$ is the set of all ordered n-tuples of real numbers, which can be assembled as columns or as rows.

- Let $x_1, \dots, x_n$ be n real numbers. Then the **column-vector** (or just **vector**) is an ordered n-tuple of the form

  $$\mathbf{v} = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix},$$

  and the **row-vector** (also called a **covector**) is an ordered n-tuple of the form

  $$\mathbf{v}^T = (v_1, v_2, \dots, v_n).$$

  The real numbers $x_1, \dots x_n$ are called the **components** of the vectors.

- The operation that converts column-vectors into row-vectors and vice versa preserving the order of the components is called the **transposition** and denoted by T. That is

  $$\begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}^T = (v_1, v_2, \dots, v_n) \quad \text{and} \quad (v_1, v_2, \dots, v_n)^T = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}.$$

  Of course, for any vector $\mathbf{v}$

  $$(\mathbf{v}^T)^T = \mathbf{v}.$$

- The addition of vectors is defined by

$$\mathbf{u} + \mathbf{v} = \begin{pmatrix} u_1 + v_1 \\ u_2 + v_2 \\ \vdots \\ u_n + v_n \end{pmatrix},$$

and

$$\mathbf{u} + \mathbf{v} = (u_1 + v_1, \ldots, u_n + v_n).$$

- Notice that one cannot add a column-vector and a row-vector!
- The multiplication of vectors by a real constant, called a scalar, is defined by

$$a\mathbf{v} = \begin{pmatrix} av_1 \\ av_2 \\ \vdots \\ av_n \end{pmatrix}, \qquad a\mathbf{v} = (av_1, \ldots, av_n).$$

- The vectors that have only zero elements are called zero vectors, that is

$$\mathbf{0} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \qquad \mathbf{0}^T = (0, \ldots, 0).$$

- The set of column-vectors

$$\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \qquad \mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \qquad \ldots, \quad \mathbf{e}_n = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 0 \\ 1 \end{pmatrix}$$

and the set of row-vectors

$$\mathbf{e}_1^T = (1, 0, \ldots, 0), \qquad \mathbf{e}_2^T = (0, 1, \ldots, 0), \qquad \mathbf{e}_n^T = (0, 0, \ldots, 1)$$

are called the standard (or canonical) bases in $\mathbb{R}^n$.

- A real vector space E is called an inner product space if there is a function $(\cdot,\cdot): E \times E \to \mathbb{R}$, called the inner product, that assigns to every two vectors u and v a real number (u, v) and satisfies the conditions: $\forall u, v, w \in E$, $\forall a \in \mathbb{R}$:

  1. $(v, v) \geq 0$
  2. $(v, v) = 0$ if and only if $v = 0$
  3. $(u, v) = (v, u)$
  4. $(u + v, w) = (u, w) + (v, w)$
  5. $(au, v) = (u, av) = a(u, v)$

  A finite-dimensional inner product space is called a Euclidean space.

- The inner product is often called the dot product, or the scalar product, and is denoted by

$$(u, v) = u \cdot v .$$

- Scalar product is defined by

$$v \cdot u = v_1 u_1 + v_2 u_2$$

- The Euclidean norm is a function $\|\cdot\|: E \to \mathbb{R}$ that assigns to every vector $v \in E$ a real number $\|v\|$ defined by

$$\|v\| = \sqrt{(v, v)}.$$

- The norm of a vector is also called the length.
- A vector with unit norm is called a unit vector.
- Theorem 1.1.1: For any $u, v \in E$ there holds

$$\|u + v\|^2 = \|u\|^2 + 2(u, v) + \|v\|^2 .$$

- Theorem 1.1.2: Cauchy-Schwarz's Inequality. For any $u, v \in E$ there holds

$$|(u, v)| \leq \|u\| \, \|v\| .$$

  The equality

$$|(u, v)| = \|u\| \, \|v\|$$

  holds if and only if u and v are parallel.

- Corollary: Triangle Inequality. For any $u, v \in E$ there holds

$$\|u + v\| \leq \|u\| + \|v\| .$$

Exercises

1. Parallelogram Law. Show that for any $u, v \in E$

$$\|u + v\|^2 + \|u - v\|^2 = 2\ \|u\|^2 + \|v\|^2$$

2. Pythagorean Theorem. Show that if $u \perp v$, then

$$\|u + v\|^2 = \|u\|^2 + \|v\|^2 .$$

- The angle between two non-zero vectors $\mathbf{u}$ and $\mathbf{v}$ is defined by

$$\cos\theta = \frac{(\mathbf{u}, \mathbf{v})}{\|\mathbf{u}\| \, \|\mathbf{v}\|}, \qquad 0 \le \theta \le \pi .$$

  Then the inner product can be written in the form

$$(\mathbf{u}, \mathbf{v}) = \|\mathbf{u}\| \, \|\mathbf{v}\| \cos\theta .$$

- Two non-zero vectors $\mathbf{u}, \mathbf{v} \in E$ are orthogonal, denoted by $\mathbf{u} \perp \mathbf{v}$, if

$$(\mathbf{u}, \mathbf{v}) = 0.$$

- A basis $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ is called orthonormal if each vector of the basis is a unit vector and any two distinct vectors are orthogonal to each other, that is,

$$(\mathbf{e}_i, \mathbf{e}_j) = \begin{cases} 1, & \text{if } i = j \\ 0, & \text{if } i \neq j \end{cases} .$$

- Theorem 1.1.3 : Every Euclidean space has an orthonormal basis.

- We denote the standard orthonormal basis in $\mathbb{R}^3$ by

$$\mathbf{e}_1 = \mathbf{i}, \qquad \mathbf{e}_2 = \mathbf{j}, \qquad \mathbf{e}_3 = \mathbf{k},$$

  so that

$$\mathbf{e}_i \cdot \mathbf{e}_j = \delta_{ij}.$$

- Each vector $\mathbf{v}$ is decomposed as

$$\mathbf{v} = v_1 \mathbf{i} + v_2 \mathbf{j} + v_3 \mathbf{k} .$$

  The components are computed by

$$v_1 = \mathbf{v} \cdot \mathbf{i}; \qquad v_2 = \mathbf{v} \cdot \mathbf{j}; \qquad v_3 = \mathbf{v} \cdot \mathbf{k}.$$

- The norm of the vector

$$\|\mathbf{v}\| = \sqrt{v_1^2 + v_2^2 + v_3^2}.$$

- Scalar product is defined by

$$\mathbf{v} \cdot \mathbf{u} = v_1 u_1 + v_2 u_2 + v_3 u_3 .$$

- The angle between vectors

$$\cos\theta = \frac{\mathbf{u} \cdot \mathbf{v}}{\|\mathbf{u}\| \, \|\mathbf{v}\|}.$$

- The orthogonal decomposition of a vector $\mathbf{v}$ with respect to a given unit vector $\mathbf{u}$ is

$$\mathbf{v} = \mathbf{v}_{\parallel} + \mathbf{v}_{\perp},$$

where

$$\mathbf{v}_{\parallel} = \mathbf{u}(\mathbf{u} \cdot \mathbf{v}), \qquad \mathbf{v}_{\perp} = \mathbf{v} - \mathbf{u}(\mathbf{u} \cdot \mathbf{v}).$$

The radius vector (the position vector) is

$$\mathbf{r} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}.$$

- The parametric equation of a line parallel to a vector $\mathbf{u} = a\mathbf{i} + b\mathbf{j} + c\mathbf{k}$ is

$$\mathbf{r} = \mathbf{r}_0 + t\mathbf{u},$$

where $\mathbf{r}_0 = x_0\mathbf{i} + y_0\mathbf{j} + z_0\mathbf{k}$ is a fixed vector and t is a real parameter. In components,

$$x = x_0 + at, \quad y = y_0 + bt, \quad z = z_0 + ct.$$

The non-parametric equation of a line (if a, b, c are non-zero) is

$$\frac{x - x_0}{a} = \frac{y - y_0}{b} = \frac{z - z_0}{c}.$$

- the positive orientation, the other side has the negative (left-handed) orientation.

$\mathbf{v}$ is

$$\mathbf{r} = \mathbf{r}_0 + t\mathbf{u} + s\mathbf{v},$$

where t and s are real parameters.

- A vector $\mathbf{n}$ that is perpendicular to both vectors $\mathbf{u}$ and $\mathbf{v}$ is normal to the plane.
- The non-parametric equation of a plane with the normal $\mathbf{n} = a\mathbf{i} + b\mathbf{j} + c\mathbf{k}$ is

$$(\mathbf{r} - \mathbf{r}_0) \cdot \mathbf{n} = 0$$

or

$$a(x - x_0) + b(y - y_0) + (z - z_0) = 0,$$

which can also be written as

$$ax + by + cz = d,$$

where

$$d = ax_0 + by_0 + cz_0.$$

- The positive (right-handed) orientation of a plane is defined by the right hand (or counterclockwise) rule. That is, if $\mathbf{u}_1$ and $\mathbf{u}_2$ span a plane then we orient the plane by saying which vector is the first and which is the second. The orientation is positive if the rotation from $\mathbf{u}_1$ to $\mathbf{u}_2$ is counterclockwise and negative if it is clockwise.

- The vector product of two vectors is defined by

$$\mathbf{w} = \mathbf{u} \times \mathbf{v} = \det \begin{pmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{pmatrix},$$

  or, in components,

$$w^i = \varepsilon^{ijk} u_j v_k = \frac{1}{2} \varepsilon^{ijk} (u_j v_k - u_k v_j) .$$

- The vector products of the basis vectors are

$$\mathbf{e}_i \times \mathbf{e}_j = \varepsilon_{ijk} \mathbf{e}_k .$$

  The Levi-Civita symbol in three dimensions

$$\varepsilon_{ijk} = \varepsilon^{ijk} = \begin{cases} +1 & \text{if } (i, j, k) = (1, 2, 3), (2, 3, 1), (3, 1, 2) \\ -1 & \text{if } (i, j, k) = (2, 1, 3), (3, 2, 1), (1, 3, 2) \\ 0 & \text{otherwise} \end{cases}$$

  has the following properties:

$$\varepsilon_{ijk} = -\varepsilon_{jik} = -\varepsilon_{ikj} = -\varepsilon_{kji}$$

$$\varepsilon_{ijk} = \varepsilon_{jki} = \varepsilon_{kij}$$

- If $\mathbf{u}$ and $\mathbf{v}$ are two nonzero nonparallel vectors, then the vector $\mathbf{w} = \mathbf{u} \times \mathbf{v}$ is orthogonal to both vectors, $\mathbf{u}$ and $\mathbf{v}$, and, hence, to the plane spanned by these vectors. It defines a normal to this plane.

- The area of the parallelogram spanned by two vectors $\mathbf{u}$ and $\mathbf{v}$ is

$$\text{area}(\mathbf{u}, \mathbf{v}) = |\mathbf{u} \times \mathbf{v}| = \|\mathbf{u}\| \, \|\mathbf{v}\| \sin \theta .$$

- The signed volume of the parallelepiped spanned by three vectors $\mathbf{u}$, $\mathbf{v}$ and $\mathbf{w}$ is

$$\text{vol}(\mathbf{u}, \mathbf{v}, \mathbf{w}) = \mathbf{u} \cdot (\mathbf{v}; \mathbf{w}) = \det \begin{pmatrix} u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{pmatrix} = \varepsilon^{ijk} u_i v_j w_k .$$

  The signed volume is also called the scalar triple product and denoted by

$$[\mathbf{u}, \mathbf{v}, \mathbf{w}] = \mathbf{u} \cdot (\mathbf{v} \times \mathbf{w}) .$$

- The signed volume is zero if and only if the vectors are linearly dependent, that is, coplanar.

- For linearly independent vectors its sign depends on the orientation of the triple of vectors $\{\mathbf{u}, \mathbf{v}\mathbf{w}\}$

$$\mathrm{vol}(\mathbf{u}, \mathbf{v}, \mathbf{w}) = \mathrm{sign}\,(\mathbf{u}, \mathbf{v}, \mathbf{w})|\mathrm{vol}(\mathbf{u}, \mathbf{v}, \mathbf{w})|\,,$$

where

$$\mathrm{sign}\,(\mathbf{u}, \mathbf{v}, \mathbf{w}) = \begin{cases} 1 & \text{if } \{\mathbf{u}, \mathbf{v}, \mathbf{w}\} \text{ is positively oriented} \\ -1 & \text{if } \{\mathbf{u}, \mathbf{v}, \mathbf{w}\} \text{ is negatively oriented} \end{cases}$$

- The scalar triple product is linear in each argument, anti-symmetric

$$[\mathbf{u}, \mathbf{v}, \mathbf{w}] = -[\mathbf{v}, \mathbf{u}, \mathbf{w}] = -[\mathbf{u}, \mathbf{w}, \mathbf{v}] = -[\mathbf{w}, \mathbf{v}, \mathbf{u}]$$

cyclic

$$[\mathbf{u}, \mathbf{v}, \mathbf{w}] = [\mathbf{v}, \mathbf{w}, \mathbf{u}] = [\mathbf{w}, \mathbf{u}, \mathbf{v}]\,.$$

It is normalized so that

$$[\mathbf{i}, \mathbf{j}, \mathbf{k}] = 1\,.$$

- The orthogonal decomposition of a vector $\mathbf{v}$ with respect to a unit vector $\mathbf{u}$ can be written in the form

$$\mathbf{v} = \mathbf{u}(\mathbf{u}\cdot\mathbf{v}) - \mathbf{u}\times(\mathbf{u}\times\mathbf{v})\,.$$

- This leads to many vector identities that express double vector product in terms of scalar product. For example,

$$\mathbf{u}\times(\mathbf{v}\times\mathbf{w}) = (\mathbf{u}\cdot\mathbf{w})\mathbf{v} - (\mathbf{u}\cdot\mathbf{v})\mathbf{w}$$
$$\mathbf{u}\times(\mathbf{v}\times\mathbf{w}) + \mathbf{v}\times(\mathbf{w}\times\mathbf{u}) + \mathbf{w}\times(\mathbf{u}\times\mathbf{v}) = \mathbf{0}$$
$$(\mathbf{u}\times\mathbf{v})\times(\mathbf{w}\times\mathbf{n}) = \mathbf{v}[\mathbf{u}, \mathbf{w}, \mathbf{n}] - \mathbf{u}[\mathbf{v}, \mathbf{w}, \mathbf{n}]$$
$$(\mathbf{u}\times\mathbf{v})\cdot(\mathbf{w}\times\mathbf{n}) = (\mathbf{u}\cdot\mathbf{w})(\mathbf{v}\cdot\mathbf{n}) - (\mathbf{u}\cdot\mathbf{n})(\mathbf{v}\cdot\mathbf{w})$$

1.1.2 Vector Spaces

- A **real vector space** consists of a set E, whose elements are called **vectors**, and the set of real numbers $\mathbb{R}$, whose elements are called **scalars**. There are two operations on a vector space:

  1. **Vector addition**, $+ : E \times E \to E$, that assigns to two vectors $u, v \in E$ another vector $u + v$, and
  2. **Multiplication by scalars**, $\cdot : \mathbb{R} \times E \to E$, that assigns to a vector $v \in E$ and a scalar $a \in \mathbb{R}$ a new vector $av \in E$.

  The vector addition is an **associative commutative** operation with an **additive identity**. It satisfies the following conditions:

  1. $u + v = v + u, \quad \forall u, v, \in E$
  2. $(u + v) + w = u + (v + w), \quad \forall u, v, w \in E$
  3. There is a vector $0 \in E$, called the **zero vector**, such that for any $v \in E$ there holds $v + 0 = v$.
  4. For any vector $v \in E$, there is a vector $(-v) \in E$, called the **opposite** of v, such that $v + (-v) = 0$.

  The multiplication by scalars satisfies the following conditions:

  1. $a(bv) = (ab)v, \quad \forall v \in E, \forall a, b \in \mathbb{R}$,
  2. $(a + b)v = av + bv, \quad \forall v \in E, \forall a, b \in \mathbb{R}$,
  3. $a(u + v) = au + av, \quad \forall u, v \in E, \forall a \in \mathbb{R}$,
  4. $1\,v = v \quad \forall v \in E$.

- The zero vector is unique.

- For any $u, v \in E$ there is a unique vector denoted by $w = v - u$, called the difference of v and u, such that $u + w = v$.

- For any $v \in E$,

  $$0v = 0, \qquad \text{and} \qquad (-1)v = -v.$$

- Let E be a real vector space and $A = \{e_1, \ldots, e_k\}$ be a finite collection of vectors from E. A **linear combination** of these vectors is a vector

  $$a_1 e_1 + \cdots + a_k e_k,$$

  where $\{a_1, \ldots, a_n\}$ are scalars.

- A finite collection of vectors $A = \{e_1, \ldots, e_k\}$ is **linearly independent** if

  $$a_1 e_1 + \cdots + a_k e_k = 0$$

  implies $a_1 = \cdots = a_k = 0$.

- A collection A of vectors is linearly dependent if it is not linearly independent.

- Two non-zero vectors u and v which are linearly dependent are also called parallel, denoted by $u \| v$.

- A collection A of vectors is linearly independent if no vector of A is a linear combination of a finite number of vectors from A.

- Let A be a subset of a vector space E. The span of A, denoted by span A, is the subset of E consisting of all finite linear combinations of vectors from A, i.e.

$$\operatorname{span} A = \{ v \in E \mid v = a^1 e_1 + \cdots + a^k e_k, \; e_i \in A, \; a^i \in \mathbb{R} \} .$$

  We say that the subset span A is spanned by A.

- Theorem 1.1.2.1 The span of any subset of a vector space is a vector space.

- A vector subspace of a vector space E is a subset $S \subseteq E$ of E which is itself a vector space.

- Theorem 1.1.2.2 A subset S of E is a vector subspace of E if and only if span $S = S$.

- Span of A is the smallest subspace of E containing A.

- A collection B of vectors of a vector space E is a basis of E if B is linearly independent and span $B = E$.

- A vector space E is finite-dimensional if it has a finite basis.

- Theorem 1.1.2.3 If the vector space E is finite-dimensional, then the number of vectors in any basis is the same.

- The dimension of a finite-dimensional real vector space E, denoted by dim E, is the number of vectors in a basis.

- Theorem 1.1.2.4 If $\{e_1, \dots, e_n\}$ is a basis in E, then for every vector $v \in E$ there is a unique set of real numbers $(v^i) = (v^1, \dots, v^n)$ such that

$$v = \sum_{i=1}^{n} v^i e_i = v^1 e_1 + \cdots + v^n e_n .$$

- The real numbers v^i, $i = 1, \dots, n$, are called the components of the vector v with respect to the basis $\{e_i\}$.

- It is customary to denote the components of vectors by superscripts, which should not be confused with powers of real numbers

$$v^2 \neq (v)^2 = vv, \quad \dots, \quad v^n \neq (v)^n .$$

Examples of Vector Subspaces

- Zero subspace $\{\mathbb{0}\}$.
- **Line** with a tangent vector $\mathbb{u}$:

$$S_1 = \text{span}\{\mathbb{u}\} = \{\mathbb{v} \in E \mid \mathbb{v} = t\mathbb{u},\ t \in \mathbb{R}\}.$$

- **Plane** spanned by two nonparallel vectors $\mathbb{u}_1$ and $\mathbb{u}_2$

$$S_2 = \text{span}\{\mathbb{u}_1, \mathbb{u}_2\} = \{\mathbb{v} \in E \mid \mathbb{v} = t\mathbb{u}_1 + s\mathbb{u}_2,\ t, s \in \mathbb{R}\}.$$

- More generally, a **k-plane** spanned by a linearly independent collection of k vectors $\{\mathbb{u}_1, \dots, \mathbb{u}_k\}$

$$S_k = \text{span}\{\mathbb{u}_1, \dots, \mathbb{u}_k\} = \{\mathbb{v} \in E \mid \mathbb{v} = t_1\mathbb{u}_1 + \cdots + t_k\mathbb{u}_k,\ t_1, \dots, t_k \in \mathbb{R}\}.$$

- An $(n-1)$-plane in an n-dimensional vector space is called a **hyperplane**.

1.2.1 Exercises

1. Show that if $\lambda\mathbb{v} = 0$, then either $\mathbb{v} = 0$ or $\lambda = 0$.
2. Prove that the span of a collection of vectors is a vector subspace.

1.1.3 Matrices

- A set of n^2 real numbers A_{ij}, $i, j = 1, \dots, n$, arranged in an array that has n columns and n rows

$$A = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nn} \end{pmatrix}$$

  is called a square $n \times n$ real matrix.

- The set of all real square $n \times n$ matrices is denoted by $\mathrm{Mat}(n, \mathbb{R})$.

- The number A_{ij} (also called an entry of the matrix) appears in the i-th row and the j-th column of the matrix A

$$A = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1j} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2j} & \cdots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ A_{i1} & A_{i2} & \cdots & \boxed{A_{ij}} & \cdots & A_{in} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nj} & \cdots & A_{nn} \end{pmatrix}$$

- Remark. Notice that the first index indicates the row and the second index indicates the column of the matrix.

- The matrix whose all entries are equal to zero is called the zero matrix.

- The addition of matrices is defined by

$$A + B = \begin{pmatrix} A_{11} + B_{11} & A_{12} + B_{12} & \cdots & A_{1n} + B_{1n} \\ A_{21} + B_{21} & A_{22} + B_{22} & \cdots & A_{2n} + B_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{n1} + B_{n1} & A_{n2} + B_{n2} & \cdots & A_{nn} + B_{nn} \end{pmatrix}$$

and the **multiplication by scalars** by

$$cA = \begin{pmatrix} cA_{11} & cA_{12} & \cdots & cA_{1n} \\ cA_{21} & cA_{22} & \cdots & cA_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ cA_{n1} & cA_{n2} & \cdots & cA_{nn} \end{pmatrix}.$$

- The numbers A_{ii} are called the **diagonal entries**. Of course, there are n diagonal entries. The set of diagonal entries is called the **diagonal** of the matrix A.
- The numbers A_{ij} with $i \neq j$ are called **off-diagonal entries**; there are $n(n-1)$ off-diagonal entries.
- The numbers A_{ij} with $i < j$ are called the **upper triangular entries**. The set of upper triangular entries is called the **upper triangular part** of the matrix A.
- The numbers A_{ij} with $i > j$ are called the **lower triangular entries**. The set of lower triangular entries is called the **lower triangular part** of the matrix A.
- The number of upper-triangular entries and the lower-triangular entries is the same and is equal to $n(n-1)/2$.
- A matrix whose only non-zero entries are on the diagonal is called a **diagonal matrix**. For a diagonal matrix

$$A_{ij} = 0 \qquad \text{if} \qquad i \neq j.$$

- The diagonal matrix

$$A = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}$$

is also denoted by

$$A = \operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$$

- A diagonal matrix whose all diagonal entries are equal to 1

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

is called the **identity matrix**. The elements of the identity matrix are

$$I_{ij} = \begin{cases} 1, & \text{if } i = j \\ 0, & \text{if } i \neq j. \end{cases}$$

- A matrix A of the form
$$A = \begin{pmatrix} * & * & \cdots & * \\ 0 & * & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & * \end{pmatrix}$$
where $*$ represents nonzero entries is called an upper triangular matrix. Its lower triangular part is zero, that is,
$$A_{ij} = 0 \qquad \text{if} \qquad i < j.$$

- A matrix A of the form
$$A = \begin{pmatrix} * & 0 & \cdots & 0 \\ * & * & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ * & * & \cdots & * \end{pmatrix}$$
whose upper triangular part is zero, that is,
$$A_{ij} = 0 \qquad \text{if} \qquad i > j,$$
is called a lower triangular matrix.

- The transpose of a matrix A whose ij-th entry is A_{ij} is the matrix A^T whose ij-th entry is A_{ji}. That is, A^T obtained from A by switching the roles of rows and columns of A:
$$A^T = \begin{pmatrix} A_{11} & A_{21} & \cdots & A_{j1} & \cdots & A_{n1} \\ A_{12} & A_{22} & \cdots & A_{j2} & \cdots & A_{n2} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ A_{1i} & A_{2i} & \cdots & \boxed{A_{ji}} & \cdots & A_{ni} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ A_{1n} & A_{2n} & \cdots & A_{jn} & \cdots & A_{nn} \end{pmatrix}$$
or
$$(A^T)_{ij} = A_{ji}.$$

- A matrix A is called symmetric if
$$A^T = A$$
and anti-symmetric if
$$A^T = -A.$$

- The number of independent entries of an anti-symmetric matrix is $n(n-1)/2$.

- The number of independent entries of a symmetric matrix is $n(n+1)/2$.

- A matrix A of the form

$$A = \begin{pmatrix} * & * & \cdots & * \\ 0 & * & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & * \end{pmatrix}$$

where $*$ represents nonzero entries is called an upper triangular matrix. Its lower triangular part is zero, that is,

$$A_{ij} = 0 \qquad \text{if} \qquad i < j.$$

- A matrix A of the form

$$A = \begin{pmatrix} * & 0 & \cdots & 0 \\ * & * & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ * & * & \cdots & * \end{pmatrix}$$

whose upper triangular part is zero, that is,

$$A_{ij} = 0 \qquad \text{if} \qquad i > j,$$

is called a lower triangular matrix.

- The transpose of a matrix A whose ij-th entry is A_{ij} is the matrix A^T whose ij-th entry is A_{ji}. That is, A^T obtained from A by switching the roles of rows and columns of A:

$$A^T = \begin{pmatrix} A_{11} & A_{21} & \cdots & A_{j1} & \cdots & A_{n1} \\ A_{12} & A_{22} & \cdots & A_{j2} & \cdots & A_{n2} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ A_{1i} & A_{2i} & \cdots & \boxed{A_{ji}} & \cdots & A_{ni} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ A_{1n} & A_{2n} & \cdots & A_{jn} & \cdots & A_{nn} \end{pmatrix}$$

or

$$(A^T)_{ij} = A_{ji}.$$

- A matrix A is called symmetric if

$$A^T = A$$

and anti-symmetric if

$$A^T = -A.$$

- The number of independent entries of an anti-symmetric matrix is $n(n-1)/2$.

- The number of independent entries of a symmetric matrix is $n(n+1)/2$.

- Every matrix A can be uniquely decomposed as the sum of its diagonal part A_D, the lower triangular part A_L and the upper triangular part A_U

$$A = A_D + A_L + A_U .$$

- For an anti-symmetric matrix

$$A_U^T = -A_L \qquad \text{and} \qquad A_D = 0 .$$

- For a symmetric matrix

$$A_U^T = A_L .$$

- Every matrix A can be uniquely decomposed as the sum of its symmetric part A_S and its anti-symmetric part A_A

$$A = A_S + A_A ,$$

where

$$A_S = \frac{1}{2}(A + A^T) , \qquad A_A = \frac{1}{2}(A - A^T) .$$

- The product of matrices is defined as follows. The ij-th entry of the product $C = AB$ of two matrices A and B is

$$C_{ij} = \sum_{k=1}^{n} A_{ik} B_{kj} = A_{i1} B_{1j} + A_{i2} B_{2j} + \cdots + A_{in} B_{nj} .$$

This is again a multiplication of the "i-th row of the matrix A by the j-th column of the matrix B".

- Theorem 1.1.3.1: The product of matrices is associative, that is, for any matrices A, B, C

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

- Theorem 1.1.3.2: For any two matrices A and B

$$(AB)^T = B^T A^T .$$

- A matrix A is called invertible if there is another matrix A^{-1} such that

$$AA^{-1} = A^{-1} A = I .$$

The matrix A^{-1} is called the inverse of A

- Theorem 1.1.3.3: For any two invertible matrices A and B

$$(AB)^{-1} = B^{-1} A^{-1} ,$$

and

$$(A^{-1})^T = (A^T)^{-1} .$$

- A matrix A is called orthogonal if

$$A^T A = AA^T = I,$$

which means $A^T = A^{-1}$.

- The trace is a map $\mathrm{tr} : \mathrm{Mat}(n, \mathbb{R})$ that assigns to each matrix $A = (A_{ij})$ a real number $\mathrm{tr}\, A$ equal to the sum of the diagonal elements of a matrix

$$\mathrm{tr}\, A = \sum_{k=1}^{n} A_{kk}.$$

- Theorem 1.1.3.4 The trace has the properties

$$\mathrm{tr}\,(AB) = \mathrm{tr}\,(BA),$$

and

$$\mathrm{tr}\, A^T = \mathrm{tr}\, A$$

- Obviously, the trace of an anti-symmetric matrix is equal to zero.

- Finally, we define the multiplication of column-vectors by matrices from the left and the multiplication of row-vectors by matrices from the right as follows.

- Each matrix defines a natural left action on a column-vector and a right action on a row-vector.

- For each column-vector v and a matrix $A = (A_{ij})$ the column-vector $u = Av$ is given by

$$\begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_i \\ \vdots \\ u_n \end{pmatrix} = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{i1} & A_{i2} & \cdots & A_{in} \\ \vdots & \vdots & \vdots & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nn} \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_i \\ \vdots \\ v_n \end{pmatrix} = \begin{pmatrix} A_{11} v_1 + A_{12} v_2 + \cdots + A_{1n} v_n \\ A_{21} v_1 + A_{22} v_2 + \cdots + A_{2n} v_n \\ \vdots \\ A_{i1} v_1 + A_{i2} v_2 + \cdots + A_{in} v_n \\ \vdots \\ A_{n1} v_1 + A_{n2} v_2 + \cdots + A_{nn} v_n \end{pmatrix}$$

- The components of the vector u are

$$u_i = \sum_{j=1}^{n} A_{ij} v_j = A_{i1} v_1 + A_{i2} v_2 + \cdots + A_{in} v_n.$$

- Similarly, for a row vector v^T the components of the row-vector $u^T = v^T A$ are defined by

$$u_i = \sum_{j=1}^{n} v_j A_{ji} = v_1 A_{1i} + v_2 A_{2i} + \cdots + v_n A_{ni}.$$

1.1.3.1 Determinant

- Consider the set $Z_n = \{1, 2, \ldots, n\}$ of the first n integers. A permutation φ of the set $\{1, 2, \ldots, n\}$ is an ordered n-tuple $(\varphi(1), \ldots, \varphi(n))$ of these numbers.
- That is, a permutation is a bijective (one-to-one and onto) function
$$\varphi : Z_n \to Z_n$$
that assigns to each number i from the set $Z_n = \{1, \ldots, n\}$ another number $\varphi(i)$ from this set.
- An elementary permutation is a permutation that exchanges the order of only two numbers.
- Every permutation can be realized as a product (or a composition) of elementary permutations. A permutation that can be realized by an even number of elementary permutations is called an even permutation. A permutation that can be realized by an odd number of elementary permutations is called an odd permutation.
- Proposition 1.1.3.1: The parity of a permutation does not depend on the representation of a permutation by a product of the elementary ones.
- That is, each representation of an even permutation has even number of elementary permutations, and similarly for odd permutations.
- The sign of a permutation φ, denoted by $\mathrm{sign}(\varphi)$ (or simply $(-1)^\varphi$), is defined by
$$\mathrm{sign}(\varphi) = (-1)^\varphi = \begin{cases} +1, & \text{if } \varphi \text{ is even,} \\ -1, & \text{if } \varphi \text{ is odd} \end{cases}$$
- The set of all permutations of n numbers is denoted by S_n.
- Theorem 1.1.3.5: The cardinality of this set, that is, the number of different permutations, is
$$|S_n| = n! \,.$$
- The determinant is a map $\det : \mathrm{Mat}(n, \mathbb{R}) \to \mathbb{R}$ that assigns to each matrix $A = (A_{ij})$ a real number $\det A$ defined by
$$\det A = \sum_{\varphi \in S_n} \mathrm{sign}(\varphi) A_{1\varphi(1)} \cdots A_{n\varphi(n)} \,,$$
where the summation goes over all $n!$ permutations.
- The most important properties of the determinant are listed below:

Theorem 1.1.3.6: 1. The determinant of the product of matrices is equal to the product of the determinants:
$$\det(AB) = \det A \det B \,.$$

2. The determinants of a matrix A and of its transpose A^T are equal:

$$\det A = \det A^T.$$

3. The determinant of the inverse A^{-1} of an invertible matrix A is equal to the inverse of the determinant of A:

$$\det A^{-1} = (\det A)^{-1}$$

4. A matrix is invertible if and only if its determinant is non-zero.

- The set of real invertible matrices (with non-zero determinant) is denoted by $GL(n, \mathbb{R})$. The set of matrices with positive determinant is denoted by $GL_+(n, \mathbb{R})$.
- A matrix with unit determinant is called **unimodular**.
- The set of real matrices with unit determinant is denoted by $SL(n, \mathbb{R})$.
- The set of real orthogonal matrices is denoted by $O(n)$.
- **Theorem 1.1.3.7** The determinant of an orthogonal matrix is equal to either 1 or -1.
- An orthogonal matrix with unit determinant (a unimodular orthogonal matrix) is called a **proper orthogonal matrix** or just a **rotation**.
- The set of real orthogonal matrices with unit determinant is denoted by $SO(n)$.
- A set G of invertible matrices forms a **group** if it is closed under taking inverse and matrix multiplication, that is, if the inverse A^{-1} of any matrix A in G belongs to the set G and the product AB of any two matrices A and B in G belongs to G.

Exercises

1. Show that the product of invertible matrices is an invertible matrix.

2. Show that the product of matrices with positive determinant is a matrix with positive determinant.

Show that the product of matrices with unit determinant is a matrix with unit determinant.

Show that the inverse of a matrix with unit determinant is a matrix with unit determinant.

Linear combination

Let $\mathbf{A}$ be a matrix. The i^{th} row is said *linear combination* of the other rows if each of its element $a_{i,j}$ can be expressed as weighted sum of the other elements of the j^{th} column by means of the same scalars $\lambda_1, \lambda_2, \ldots, \lambda_{i-1}, \lambda_{i+1}, \ldots \lambda_n$:

$$\mathbf{a_i} = \lambda_1 \mathbf{a_1} + \lambda_2 \mathbf{a_2} + \cdots + \lambda_{i-1} \mathbf{a_{i-1}} + \lambda_{i+1} \mathbf{a_{i+1}} + \ldots + \lambda_n \mathbf{a_n}.$$

Equivalently, we may express the same concept by considering each row element:

$$\forall j : \exists \lambda_1, \lambda_2, \ldots, \lambda_{i-1}, \lambda_{i+1}, \ldots \lambda_n |$$
$$a_{i,j} = \lambda_1 a_{1,j} + \lambda_2 a_{2,j} + \ldots \lambda_{i-1} a_{i-1,j} + \lambda_{i+1} a_{i+1,j} + \ldots \lambda_n a_{n,j}.$$

Example 2.27. Let us consider the following matrix:

$$\mathbf{A} = \begin{pmatrix} 0 & 1 & 1 \\ 3 & 2 & 1 \\ 6 & 5 & 3 \end{pmatrix}.$$

The third row is a linear combination of the first two by means of scalars $\lambda_1, \lambda_2 = 1, 2$, the third row is equal to the weighted sum obtained by multiplying the first row by 1 and summing to it the second row multiplied by 2:

$$(6,5,3) = (0,1,1) + 2(3,2,1)$$

that is

$$\mathbf{a_3} = \mathbf{a_1} + 2\mathbf{a_2}.$$

Let $\mathbf{A}$ be a matrix. The j^{th} column is said *linear combination* of the other column if each of its element $a_{i,j}$ can be expressed as weighted sum of the other elements of the i^{th} row by means of the same scalars $\lambda_1, \lambda_2, \ldots, \lambda_{j-1}$, $\lambda_{j+1}, \ldots \lambda_n$:

$$\mathbf{a^j} = \lambda_1 \mathbf{a^1} + \lambda_2 \mathbf{a^2} + \cdots + \lambda_{j-1} \mathbf{a^{j-1}} + \lambda_{j+1} \mathbf{a^{j+1}} + \ldots + \lambda_n \mathbf{a^n}.$$

Equivalently, we may express the same concept by considering each row element:

$$\forall i : \exists \lambda_1, \lambda_2, \ldots, \lambda_{j-1}, \lambda_{j+1}, \ldots \lambda_n |$$
$$a_{i,j} = \lambda_1 a_{i,1} + \lambda_2 a_{i,2} + \ldots \lambda_{i-1} a_{i,j-1} + \lambda_{i+1} a_{i,j+1} + \ldots \lambda_n a_{i,n}.$$

Example 2.28. Let us consider the following matrix:

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 2 & 4 \\ 1 & 3 & 0 \end{pmatrix}.$$

The third column is a linear combination of the first two by means of scalars $\lambda_1, \lambda_2 = 3, -1$, the third column is equal to the weighted sum obtained by multiplying the first column by 3 and summing to it the second row multiplied by -1:

$$\begin{pmatrix} 1 \\ 4 \\ 0 \end{pmatrix} = 3 \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} - \begin{pmatrix} 2 \\ 2 \\ 3 \end{pmatrix}.$$

that is

$$\mathbf{a}^3 = 3\mathbf{a}^1 - \mathbf{a}^2.$$

Let $\mathbf{A} \in \mathbb{R}_{m,n}$ be a matrix. The m rows (n columns) are *linearly dependent* if a row (column) composed of all zeros $\mathbf{o} = (0, 0, \ldots, 0)$ can be expressed as the linear combination of the m rows (n columns) by means of nun-null scalars.

In the case of linearly dependent rows, if the matrix $\mathbf{A}$ is represented as a vector of row vectors:

$$\mathbf{A} = \begin{pmatrix} \mathbf{a_1} \\ \mathbf{a_2} \\ \ldots \\ \mathbf{a_m} \end{pmatrix}$$

the rows are linearly dependent if

$$\exists \lambda_1, \lambda_2, \ldots, \lambda_m \neq 0, 0, \ldots, 0$$

such that

$$\mathbf{o} = \lambda_1 \mathbf{a_1} + \lambda_2 \mathbf{a_2} + \ldots + \lambda_m \mathbf{a_m}.$$

Submatrices, Cofactors, Adjugate Matrices

Submatrices. Let us consider a matrix $\mathbf{A} \in \mathbb{R}_{m,n}$. Let r, s be two positive integer numbers such that $1 \leq r \leq m$ and $1 \leq s \leq n$. A *submatrix* is a matrix obtained from $\mathbf{A}$ by cancelling $m - r$ rows and $n - s$ columns.

Example 2.41. Let us consider the following matrix:

$$\mathbf{A} = \begin{pmatrix} 3 & 3 & 1 & 0 \\ 2 & 4 & 1 & 2 \\ 5 & 1 & 1 & 1 \end{pmatrix}.$$

The submatrix obtained by cancelling the second row, the second and fourth columns is

$$\begin{pmatrix} 3 & 1 \\ 5 & 1 \end{pmatrix}.$$

•**Cofactor.** Let us consider a matrix $\mathbf{A} \in \mathbb{R}_{n,n}$, its generic element $a_{i,j}$ and corresponding complement minor $M_{i,j}$. The *cofactor* $A_{i,j}$ of the element $a_{i,j}$ is defined as $A_{i,j} = (-1)^{i+j} M_{i,j}$.

Example 2.44. From the matrix of the previous example, the cofactor $A_{1,2} = (-1)M_{1,2}$.

Adjugate Matrix. Let us consider a matrix $\mathbf{A} \in \mathbb{R}_{n,n}$:

$$\mathbf{A} = \begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} \\ \dots & \dots & \dots & \dots \\ a_{n,1} & a_{n,2} & \dots & a_{n,n} \end{pmatrix}.$$

Let us compute the transpose matrix $\mathbf{A}^{\mathbf{T}}$:

$$\mathbf{A}^{\mathbf{T}} = \begin{pmatrix} a_{1,1} & a_{2,1} & \dots & a_{n,1} \\ a_{1,2} & a_{2,2} & \dots & a_{n,2} \\ \dots & \dots & \dots & \dots \\ a_{1,n} & a_{2,n} & \dots & a_{n,n} \end{pmatrix}.$$

Let us substitute each element of the transpose matrix with its correspondi cofactor $A_{i,j}$. The resulting matrix is said *adjugate matrix* (or adjunct or adjoint) the matrix $\mathbf{A}$ and is indicated with $\text{adj}(\mathbf{A})$:

$$\text{adj}(\mathbf{A}) = \begin{pmatrix} A_{1,1} & A_{2,1} & \dots & A_{n,1} \\ A_{1,2} & A_{2,2} & \dots & A_{n,2} \\ \dots & \dots & \dots & \dots \\ A_{1,n} & A_{2,n} & \dots & A_{n,n} \end{pmatrix}.$$

Example 2.45. Let us consider the following matrix $\mathbf{A} \in \mathbb{R}_{3,3}$:

$$\mathbf{A} = \begin{pmatrix} 1 & 3 & 0 \\ 5 & 3 & 2 \\ 0 & 1 & 2 \end{pmatrix}$$

and compute the corresponding Adjugate Matrix. In order to achieve this purpo let us compute $\mathbf{A}^{\mathbf{T}}$:

$$\mathbf{A}^{\mathbf{T}} = \begin{pmatrix} 1 & 5 & 0 \\ 3 & 3 & 1 \\ 0 & 2 & 2 \end{pmatrix}.$$

Let us compute the nine complements minors: $M_{1,1} = 4$, $M_{1,2} = 6$, $M_{1,3} =$ $M_{2,1} = 10$, $M_{2,2} = 2$, $M_{2,3} = 2$, $M_{3,1} = 5$, $M_{3,2} = 1$, and $M_{3,3} = -12$. The Adjug Matrix $\text{adj}(\mathbf{A})$ is:

$$\text{adj}(\mathbf{A}) = \begin{pmatrix} 4 & -6 & 6 \\ -10 & 2 & -2 \\ 5 & -1 & -12 \end{pmatrix}.$$

Invertible matrices

Definition

Let $\mathbf{A} \in \mathbb{R}_{n,n}$. If $\det \mathbf{A} = 0$ the matrix is said singular. If $\det \mathbf{A} \neq 0$ the matrix is said non-singular.

Definition

Let $\mathbf{A} \in \mathbb{R}_{n,n}$. The matrix $\mathbf{A}$ is said *invertible* if $\exists$ a matrix $\mathbf{B} \in \mathbb{R}_{n,n} | \mathbf{AB} = \mathbf{I} = \mathbf{BA}$. The matrix $\mathbf{B}$ is said *inverse* matrix of the matrix $\mathbf{A}$.

Theorem

If $\mathbf{A} \in \mathbb{R}_{n,n}$ is an invertible matrix and $\mathbf{B}$ is its inverse. It follows that the inverse matrix is unique: $\exists! \mathbf{B} \in \mathbb{R}_{n,n} | \mathbf{AB} = \mathbf{I} = \mathbf{BA}$.

Proof. Let us assume by contradiction that the inverse matrix is not unique. Thus, besides $\mathbf{B}$, there exists another inverse of $\mathbf{A}$, indicated as $\mathbf{C} \in \mathbb{R}_{n,n}$.

This would mean that for the hypothesis $\mathbf{B}$ is inverse of $\mathbf{A}$ and thus

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

For the contradiction hypothesis also $\mathbf{C}$ is inverse of $\mathbf{A}$ and thus

$$\mathbf{AC} = \mathbf{CA} = \mathbf{I}.$$

Considering that $\mathbf{I}$ is the neutral element with respect to the product of matrices ($\forall \mathbf{A} : \mathbf{AI} = \mathbf{IA} = \mathbf{A}$) and that the product of matrices is associative, it follows that

$$\mathbf{C} = \mathbf{CI} = \mathbf{C}(\mathbf{AB}) = (\mathbf{CA})\mathbf{B} = \mathbf{IB} = \mathbf{B}.$$

In other words, if $\mathbf{B}$ is an inverse matrix of $\mathbf{A}$ and another inverse matrix $\mathbf{C}$ exists, then $\mathbf{C} = \mathbf{B}$. Thus, the inverse matrix is unique. $\square$

The only inverse matrix of the matrix $\mathbf{A}$ is indicated with $\mathbf{A}^{-1}$.

Theorem

Let $\mathbf{A} \in \mathbb{R}_{n,n}$ and $A_{i,j}$ its generic cofactor. The inverse matrix $\mathbf{A}^{-1}$ is

$$\mathbf{A}^{-1} = \frac{1}{\det \mathbf{A}} \operatorname{adj}(\mathbf{A}).$$

Example 2.49. Let us calculate the inverse of the matrix

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

The determinant of this matrix is $\det \mathbf{A} = 1$. The transpose of this matrix is

$$\mathbf{A}^{\mathbf{T}} = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix},$$

which, in this case, is equal to $\mathbf{A}$.

The adjugate matrix is

$$\operatorname{adj}(\mathbf{A}) = \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix}.$$

The inverse of the matrix $\mathbf{A}$ is then

$$\mathbf{A}^{-1} = \frac{1}{\det \mathbf{A}} \operatorname{adj}(\mathbf{A}) = \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix}.$$

Let $\mathbf{A}$ *and* $\mathbf{B}$ *be two square and invertible matrices. it follows that*

$$(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}.$$

Orthogonal Matrix

Definition

A matrix $\mathbf{A} \in \mathbb{R}_{n,n}$ is said *orthogonal* if the product between it and its transpose is the identity matrix:

$$\mathbf{A}\mathbf{A}^{\mathbf{T}} = \mathbf{I} = \mathbf{A}^{\mathbf{T}}\mathbf{A}.$$

An orthogonal matrix is always non-singular and its determinant is either 1 *or* -1.

Proof. Let $\mathbf{A} \in \mathbb{R}_{n,n}$ be an orthogonal matrix. Then,

$$\mathbf{A}\mathbf{A}^\mathbf{T} = \mathbf{I}.$$

Thus, the determinants are still equal:

$$\det\left(\mathbf{A}\mathbf{A}^\mathbf{T}\right) = \det \mathbf{I}.$$

For the properties of the determinant

$$\det\left(\mathbf{A}\mathbf{A}^\mathbf{T}\right) = \det \mathbf{A} \det \mathbf{A}^\mathbf{T}$$
$$\det \mathbf{A} = \det \mathbf{A}^\mathbf{T}$$
$$\det \mathbf{I} = 1$$

Thus,

$$(\det \mathbf{A})^2 = 1.$$

This can happen only when $\det \mathbf{A} = \pm 1$. $\square$

Properties of Orthogonal matrices

- $\sum_{j=1}^{n} a_{i,j}^2 = 1$ Sum of squared row elements
  $\forall i, j\ \mathbf{a_i}\mathbf{a_j} = 0.$ Dot product of any two rows or columns

Example 2.57. The following matrices are orthogonal:

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

and

$$\begin{pmatrix} sin(\alpha) & 0 & cos(\alpha) \\ 0 & 1 & 0 \\ -cos(\alpha) & 0 & sin(\alpha) \end{pmatrix}.$$

Example 2.59. The following matrix is orthogonal

$$\mathbf{A} = \frac{1}{3}\begin{pmatrix} 2 & -2 & 1 \\ 1 & 2 & 2 \\ 2 & 1 & -2 \end{pmatrix}.$$

Let us verify that the matrix is orthogonal by calculating $\mathbf{A}\mathbf{A}^\mathbf{T}$:

$$\mathbf{A}\mathbf{A}^\mathbf{T} = \frac{1}{3}\begin{pmatrix} 2 & -2 & 1 \\ 1 & 2 & 2 \\ 2 & 1 & -2 \end{pmatrix} \frac{1}{3}\begin{pmatrix} 2 & 1 & 2 \\ -2 & 2 & 1 \\ 1 & 2 & -2 \end{pmatrix} = \frac{1}{9}\begin{pmatrix} 9 & 0 & 0 \\ 0 & 9 & 0 \\ 0 & 0 & 9 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \mathbf{I}.$$

Let us verify the properties of orthogonal matrices. The sum of the squares of the rows (columns) is equal to one, e.g.

$$a_{1,1}^2 + a_{1,2}^2 + a_{1,3}^2 = \left(\frac{2}{3}\right)^2 + \left(-\frac{2}{3}\right)^2 + \left(\frac{1}{3}\right)^2 = \frac{4}{9} + \frac{4}{9} + \frac{1}{9} = 1.$$

The scalar product of each pair of rows (columns) is zero, e.g.

$$\mathbf{a_1 a_2} = a_{1,1}a_{2,1} + a_{1,2}a_{2,2} + a_{1,3}a_{2,3} = \left(\frac{2}{3}, -\frac{2}{3}, \frac{1}{3}\right)\left(\frac{1}{3}, \frac{2}{3}, \frac{2}{3}\right) = \frac{2}{9} - \frac{4}{9} + \frac{2}{9} = 0.$$

Rank of a matrix

Definition

Let $\mathbf{A} \in \mathbb{R}_{m,n}$ with $\mathbf{A}$ assumed to be different from the null matrix.
The *rank* of the matrix $\mathbf{A}$, indicated as $\rho_{\mathbf{A}}$, is the highest order of the non-singular submatrix $\mathbf{A}_\rho \subset \mathbf{A}$. If $\mathbf{A}$ is the null matrix then its rank is taken equal to 0.

Example 2.60. The rank of the matrix

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

is 2 as the submatrix

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

is non-singular.

Theorem

Let $\mathbf{A} \in \mathbb{R}_{n,n}$ and ρ its rank. The matrix $\mathbf{A}$ has ρ linearly independent rows (columns).