MAST10018
Linear Algebra
Extension Studies

Edward Wang

Semester 1, 2024

Contents

1 Matrices

A **matrix** is a rectangular array of numbers. If a matrix has m rows and n columns, then we call it a $(m \times n)$-matrix, and we call $(m \times n)$ the **dimension** of the matrix. The set of all square $(n \times n)$-matrices with real entries is denoted $M_n(\mathbb{R})$. The entry in the i^{th} row and j^{th} column of a matrix A is denoted A_{ij}.

1.1 Matrix multiplication

Matrix multiplication can be thought of as taking the dot product of the rows of the first matrix with the columns of the second matrix.

Definition (Matrix multiplication)

Given a $(m \times n)$-matrix A and a $(n \times p)$-matrix B, their product is a $(m \times p)$-matrix AB, where the entries are

$$(AB)_{ij} = \sum_{k=1}^{n} A_{ik} B_{kj}.$$

Example 1.1

We have

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} e & f \\ g & h \end{bmatrix} = \begin{bmatrix} ae + bg & af + bh \\ ce + dg & cf + dh \end{bmatrix}.$$

Remark — The matrix product AB is only defined if the number of columns of A is equal to the number of rows of B.

1.2 Matrix transpose

The matrix transpose is a useful operation on a single matrix. There are many properties of the matrix transpose.

Definition

Given a matrix A, the **transpose** of A is denoted A^T, such that

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

Definition

A matrix A is **symmetric** if $A = A^T$.

Theorem 1.1

For matrices A and B where AB is defined, we have

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

Proof. We know by definition that $((AB)^T)_{ij} = (AB)_{ji}$. Using the definition of matrix multiplication, we have

$$\begin{aligned}(AB)_{ji} &= \sum_{k=1}^{n} A_{jk}B_{ki} \\ &= \sum_{k=1}^{n} B_{ki}A_{jk} \\ &= \sum_{k=1}^{n} (B^T)_{ik}(A^T)_{kj} \\ &= (B^T A^T)_{ij}.\end{aligned}$$

As this holds for all i, j, we must have that $(AB)^T = B^T A^T$. $\square$

1.3 Trace

The **trace** of a square matrix A is the sum of the diagonal entries of A, and is denoted $\mathrm{tr}(A)$.

Theorem 1.2
Given $(n \times n)$-matrices A and B, we have

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

1.4 Matrix inverses

A square matrix A has an **inverse** if there exists a matrix B such that $AB = BA = I$, where I is the identity matrix.

Remark — Matrix multiplication is **not** commutative, that is $AB \neq BA$ in general.

If there exists such a matrix B, then we call it the **inverse** of A, and denote it $B = A^{-1}$. A matrix is **invertible** if it has an inverse, and **singular** if it does not have an inverse.

Theorem 1.3
Matrix inverses are unique. That is, if A has an inverse A^{-1}, then A^{-1} is unique.

Proof. Suppose A has inverse B and C, such that

$$\begin{aligned}AB = BA = I \\ AC = CA = I.\end{aligned}$$

We have

$$\begin{aligned}C = CI &= C(AB) \\ &= (CA)B \\ &= IB \\ &= B.\end{aligned}$$

Hence $B = C$, so A^{-1} is unique. $\square$

THEOREM 1.4

If A and B are matrices such that AB is defined and AB is invertible, then

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

Proof. Using associativity, we have

$$\begin{aligned}(AB)(B^{-1}A^{-1}) &= A(BB^{-1})A^{-1} \\ &= AIA^{-1} \\ &= AA^{-1} \\ &= I.\end{aligned}$$

Moreover, we have

$$\begin{aligned}(B^{-1}A^{-1})(AB) &= B^{-1}(A^{-1}A)B \\ &= B^{-1}IB \\ &= B^{-1}B \\ &= I.\end{aligned}$$

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

THEOREM 1.5

If A and B are invertible matrices of the same dimension, then the following all hold:

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

THEOREM 1.6

The inverse of a general (2×2)-matrix is given by

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

1.5 Determinants

The **determinant** of a square matrix A is a function of its entries that outputs a real number, denoted $\det(A)$. In other words, we have

$$\det\colon \{\text{all square matrices}\} \to \mathbb{R}.$$

THEOREM 1.7

If A is a square matrix, then

$$\det(A) \neq 0 \iff A \text{ is invertible.}$$

For (2×2)-matrices, the determinant is defined as

$$\det\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc.$$

The determinant satisfies many properties.

THEOREM 1.8

Given two square matrices A and B, then the following properties always hold:

1. $\det(AB) = \det(A)\det(B)$.
2. $\det(A^T) = \det(A)$.

THEOREM 1.9

If I is an identity matrix, then

$$\det(I) = 1.$$

Corollary 1.9.1. If A is invertible, then

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

1.6 Matrices as transformations

Let A be a $(m \times n)$ matrix. Consider a column vector B of length n, such that AB is defined. Then AB has dimension $(m \times 1)$, so it is a column vector of length m.

$$\begin{bmatrix} a_{11} & \dots & \dots & a_{1n} \\ \vdots & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ a_{m_1} & \dots & \dots & a_{mn} \end{bmatrix} \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix} = \begin{bmatrix} c_1 \\ \vdots \\ \vdots \\ c_m \end{bmatrix}$$

Thus we can think of a $(m \times n)$matrix A as a mapping from $\mathbb{R}^n$ to $\mathbb{R}^m$, that is,

$$A \colon \mathbb{R}^n \to \mathbb{R}^m.$$

2 Vectors

Euclidean space is commonly encountered. In general, we may have n dimensions, in which algebra behaves much the same.

Definition

We define n-dimensional Euclidean space to be the set of all tuples

$$R^n = \{(x_1, x_2, \dots, x_n) : x_i \in \mathbb{R} \text{ for all } 1 \le i \le n\}.$$

Definition (Linear dependency)

If $S = \{\mathbf{v}_1, \dots, \mathbf{v}_k\}$ is a collection of vectors, then S is **linearly dependent** if there exist scalars $\lambda_1, \dots, \lambda_k$ such that

$$\lambda_1 \mathbf{v}_1 + \lambda_2 \mathbf{v}_2 + \cdots + \lambda_k \mathbf{v}_k = \mathbf{0},$$

and λ_i are not all equal to zero. If such λ_i do not exist, then S is **linearly independent**.

Definition

The **magnitude** of a vector $\mathbf{u} \in \mathbb{R}^n$ is defined to be

$$\|u\| = \sqrt{u_1^2 + \cdots + u_n^2}.$$

Definition (Dot product)

Given two vectors $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$, we define the **dot product** to be $\mathbf{u} \cdot \mathbf{v}$ such that

$$\mathbf{u} \cdot \mathbf{v} = u_1 v_1 + u_2 v_2 + \cdots + u_n v_n.$$

Remark — The dot product gives us a quantitiative measure of linear dependence.

Theorem 2.1 (Dot product properties)

Suppose we have two vectors $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$. Then the following properties all hold:

1. $\mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u}$,
2. $k(\mathbf{u} \cdot \mathbf{v}) = k\mathbf{u} \cdot \mathbf{v} = \mathbf{u}k\mathbf{v}$,
3. $\mathbf{u} \cdot (\mathbf{v} + \mathbf{w}) = \mathbf{u} \cdot \mathbf{v} + \mathbf{u} \cdot \mathbf{w}$,
4. $\mathbf{u} \cdot \mathbf{u} = \|\mathbf{u}\|^2 \ge 0$, and $\mathbf{u} \cdot \mathbf{u} = 0$ iff $\mathbf{u} = \mathbf{0}$.

Definition (Distance)

The **distance** between two vectors $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$ is given by

$$d(\mathbf{u}, \mathbf{v}) = \|\mathbf{u} - \mathbf{v}\| = \sqrt{(u_1 - v_1)^2 + \cdots + (u_n - v_n)^2}.$$

Remark — The distance function satisfies commutativity, and is non-negative.

Theorem 2.2 (Cauchy-Schwarz)

If $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$, then

$$|\mathbf{u} \cdot \mathbf{v}| \leq \|\mathbf{u}\|\|\mathbf{v}\|.$$

Proof. If $\mathbf{v} = \mathbf{0}$, then the result is obviously true as the left hand side is equal to 0. Hence we assume that $\mathbf{v} \neq \mathbf{0}$.

Define $f \colon \mathbb{R} \to \mathbb{R}, f(t) = (\mathbf{u} + t\mathbf{v}) \cdot (\mathbf{u} + t\mathbf{v})$. Using the distributive property of the dot product, we have

$$\begin{aligned} f(t) &= \mathbf{u} \cdot \mathbf{u} + 2t(\mathbf{u} \cdot \mathbf{v}) + t^2(\mathbf{v} \cdot \mathbf{v}) \\ &= \|\mathbf{u}\|^2 + 2t(\mathbf{u} \cdot \mathbf{v}) + t^2\|\mathbf{v}\|^2, \end{aligned}$$

so f is a quadratic in t. Since $f(t) \geq 0$ by positive definiteness, we have that the descriminant $\Delta \leq 0$, giving us

$$\begin{aligned} \Delta = 4(\mathbf{u} \cdot \mathbf{v})^2 - 4\|\mathbf{u}\|\|\mathbf{v}\|^2 &\leq 0 \\ |\mathbf{u} \cdot \mathbf{v}| &\leq \|\mathbf{u}\|\|\mathbf{v}\|, \end{aligned}$$

with inequality iff either $\mathbf{u}$ or $\mathbf{v}$ is equal to $\mathbf{0}$, or $\mathbf{u} = t\mathbf{v}$ for some scalar t. □

Theorem 2.3

If $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$, then

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

where θ is the angle between $\mathbf{u}$ and $\mathbf{v}$.

Remark — Note that the 'angle' between the two vectors are only well defined for $n = 2, 3$. In higher dimensions, we may *define* the angle using the theorem.

Theorem 2.4 (Generalised Pythagoras)

If $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$, and $\mathbf{u} \cdot \mathbf{v} = 0$, then

$$\|\mathbf{u} + \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 = .$$

Proof. We have

$$\begin{aligned} \|\mathbf{u} + \mathbf{v}\|^2 &= (\mathbf{u} + \mathbf{v}) \cdot (\mathbf{u} + \mathbf{v}) \\ &= \mathbf{u} \cdot \mathbf{u} + 2(\mathbf{u} \cdot \mathbf{v}) + \mathbf{v} \cdot \mathbf{v} \\ &= \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2, \end{aligned}$$

as required. □

Theorem 2.5 (Triangle Inequality)
If $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$, then
$$\|\mathbf{u}+\mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|.$$

Proof. We have
$$\begin{aligned}
\|\mathbf{u}+\mathbf{v}\|^2 &= \|\mathbf{u}\|^2 + 2(\mathbf{u}\cdot\mathbf{v}) + \|\mathbf{v}\|^2 \\
&\leq \|\mathbf{u}\|^2 + 2(\|\mathbf{u}\|\|\mathbf{v}\|) + \|\mathbf{v}\|^2 && \text{(Cauchy-Schwarz)} \\
&= (\|\mathbf{u}\| + \|\mathbf{v}\|)^2,
\end{aligned}$$
from where the result follows. □

Definition (Vector resolute)
The **vector resolute** of a vector $\mathbf{u}$ in the direction of $\mathbf{v}$ is $(\mathbf{u}\cdot\hat{\mathbf{v}})\hat{\mathbf{v}}$.

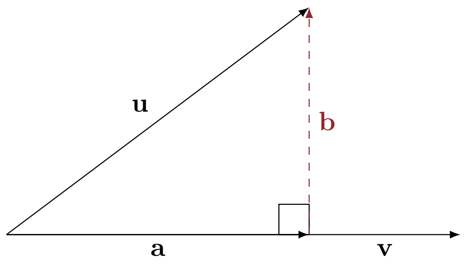

Figure 2.1: Here, $\mathbf{b}$ is the vector resolute of $\mathbf{u}$ in the direction of $\mathbf{v}$.

We can then decompose the vector $\mathbf{u}$ into two components, one being parallel ot $\mathbf{v}$, and one perpendicular:
$$\mathbf{u} = (\mathbf{u}\cdot\hat{\mathbf{v}})\hat{\mathbf{v}} + (\mathbf{u} - (\mathbf{u}\cdot\hat{\mathbf{v}})\hat{\mathbf{v}}).$$
We can now verify that these vectors are indeed perpendicular by evaluating the dot product:
$$\begin{aligned}
((\mathbf{u}\cdot\hat{\mathbf{v}})\hat{\mathbf{v}})\cdot(\mathbf{u} - (\mathbf{u}\cdot\hat{\mathbf{v}})\hat{\mathbf{v}}) &= ((\mathbf{u}\cdot\hat{\mathbf{v}})\hat{\mathbf{v}})\cdot\mathbf{u} - ((\mathbf{u}\cdot\hat{\mathbf{v}})\hat{\mathbf{v}})\cdot((\mathbf{u}\cdot\hat{\mathbf{v}})\hat{\mathbf{v}}) \\
&= (\mathbf{u}\cdot\hat{\mathbf{v}})(\hat{\mathbf{v}}\cdot\mathbf{u}) - \|(\mathbf{u}\cdot\hat{\mathbf{v}})\hat{\mathbf{v}}\|^2 \\
&= (\mathbf{u}\cdot\hat{\mathbf{v}})^2 - (\mathbf{u}\cdot\hat{\mathbf{v}})^2 \\
&= 0,
\end{aligned}$$
showing that they are perpendicular.