MATH247: Honours Applied Linear Algebra

Matthew He

January 21, 2025

Abstract

1 *Preliminaries*

1.1 *Fields*

Definition 1 (Field). *A (nonempty) set with two (inner) operations, addition and multiplication:*

$$\cdot : K \times K \mapsto K, \cdot(x, y) = x \cdot y$$
$$+ : K \times K \mapsto K, +(x, y) = x + y$$

is called a field if the following axioms hold for all $x, y, z \in K$:

$$\begin{aligned}
&(F1) \quad x + (y + z) = (x + y) + z, \quad x(yz) = (xy)z, \quad \textit{(Associativity)} \\
&(F2) \quad x + y = y + x, \quad xy = yx, \quad \textit{(Commutativity)} \\
&(F3) \quad x + (y + z) = (x + y) + z, \quad x(yz) = (xy)z, \quad \textit{(Distributivity)} \\
&(F4) \quad \exists o \in K \text{ such that } x + o = x, \exists e \in K \text{ such that } x \cdot e = x, \quad \textit{(Neutral elements)} \\
&(F5a) \quad \exists a \in K \text{ such that } x + a = o, \quad \textit{(Additive inverses)} \\
&(F5b) \quad \exists b \in K \text{ such that } x \cdot b = e, \quad \textit{(Multiplicative inverses)}
\end{aligned}$$

Subfields Write $(K, +, \cdot)$ to point out notation for a field.

Example 1 ($\mathbb{F}_2$ - *finite field*)

Let $K = \{0, 1\}$ be a set with $1 \neq 0$, and define the operations $+, \cdot$ as follows:

$$0 + x = x + 0, \quad 0 + 1 = 1, \quad 1 + 0 = 1, \quad 1 + 1 = 0$$
$$0 \cdot 0 = 0, \quad 0 \cdot 1 = 0, \quad 1 \cdot 0 = 0, \quad 1 \cdot 1 = 1$$

Theorem 1. *Let $(K, +, \cdot)$ be a field. Then for all x, y, z:*

$$\begin{aligned}
&(a) \quad x + y = x + z \Rightarrow y = z \quad (\text{Cancellation}) \\
&(b) \quad xy = xz \Rightarrow y = z, \forall x \in K \setminus \{0\} \\
&(c) \quad x \cdot o = o \\
&(d) \quad x \cdot y = o \Rightarrow x = o \lor y = o \quad \textit{(Free of zero divisors)}
\end{aligned}$$

Proof. (a) Let a be the odd

1.2 *The field of complex number*

Complex numbers $\mathbb{C}$ is born out of necessity to solve equations like

$$x^2 + t = 0.$$

Definition 2 (Complex number)**.** *We set*

$$\mathbb{C} = \{a + bi | a, b \in \mathbb{R}\}$$

where "i" is the imaginary unit

$$i^2 = -1 \quad (\star).$$

Using ordinary unit for addition and multiplication in $\mathbb{R}$ *and* $(\star)$ *,* $(\mathbb{C}, +, \cdot)$ *becomes a field containing* $\mathbb{R}$ *as a subfield.*

Given $z = a + bi \in \mathbb{C}$ we define:

$$\begin{aligned} &= a - bi \quad \text{conjugate} \\ R(z) &= a \\ Jm() &= b \end{aligned}$$

Theorem 2. *Let* $u, v \in \mathbb{C}$

1.3 *Matrices over a field*

Let $(K, +, \cdot)$ be a field. Generalizing the concept of a matrix over $K \in \{\mathbb{R}, \mathbb{C}\}$, for $m, n \in \mathbb{N}$, we define:

$$K^{m\times n} := \left\{ \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \middle| a_{ij} \in K \ (i = 1, \dots, m, \ j = 1, \dots, n) \right\}$$

of all $m \times n$ matrices with entries in K.
Given a matrix $A \in K^{m\times n}$, we write a_{ij} for its entry in the i-th row and j-th column or $A[i, j]$.

We make the convention

$$K^n := K^{n\times 1}.$$

Inspired by the case $K \in \{\mathbb{R}, \mathbb{C}\}$, we define the following calculus for matrices in $K^{m\times n}$:

- **Addition:** For $A, B \in K^{m\times n}$ we have $(A + B)[i, j] = A[i, j] + B[i, j]$.
- **Multiplication:** For $A \in K^{m\times n}$ and $B \in K^{n\times p}$ we have

$$(AB)[i, k] = \sum_{j=1}^{n} A[i, j]B[j, k].$$

- **Scalar Multiplication:** For $A \in K^{m\times n}$ and $\lambda \in K$ we have $(\lambda A)[i,j] = \lambda A[i,j]$.

The zero matrix in $K^{m\times n}$ is the matrix of all zeros, denoted by $\mathbf{0}_{m\times n}$ or simply $\mathbf{0}$. For $m = n$, we write $\mathbf{0}_n := \mathbf{0}_{n\times n}$.

The matrix units $E_{ij} \in K^{m\times n}$ with

$$E_{ij} = \begin{pmatrix} \cdots & \cdots & \cdots & \cdots \\ \cdots & 1 & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \end{pmatrix}$$

(having 1 at the (i,j)-th entry and 0 elsewhere).

The notion of the transpose of a matrix was also established in MATH 133: For $A \in K^{m\times n}$, its transpose $A^T \in K^{n\times m}$ is defined via

$$A^T[i,j] = A[j,i] \quad \text{for all } i = 1,\ldots,n \text{ and } j = 1,\ldots,m.$$

We call $A \in K^{n\times n}$ a **square matrix** or simply **square**. Important classes of square matrices are defined below:

- A matrix $A \in K^{n\times n}$ is called an **upper triangular matrix** if

$$a_{ij} = 0 \quad \text{for } i > j \quad (i,j = 1,\ldots,n).$$

- A matrix A is called a **lower triangular matrix** if

$$a_{ij} = 0 \quad \text{for } i < j \quad (i,j = 1,\ldots,n).$$

A matrix is called a **diagonal matrix** if and only if it is both upper and lower triangular matrix.

A matrix $A = (a_{ij})$ is clearly diagonal if and only if $a_{ij} = 0$ for all $i \neq j$. A special diagonal matrix in $K^{n\times n}$ is the **identity matrix**

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

Using the **Kronecker delta**

$$\delta_{ij} = \begin{cases} 1, & \text{if } i = j, \\ 0, & \text{otherwise,} \end{cases}$$

we can write $I = (\delta_{ij})$.

For a square matrix $A \in K^{n\times n}$, if there exists $B \in K^{n\times n}$ such that

$$AB = I_n = BA,$$

we say that A is **invertible** and write $A^{-1} = B$ and call A^{-1} the **inverse** of A, which is uniquely determined already by one of the conditions $AB = I$ or $BA = I$, as we will confirm in Section 4.1.

The following properties are established in MATH 133 for the case $K \in \{\mathbb{R}, \mathbb{C}\}$ and will be used throughout.

Theorem 3 (Matrix calculus)**.** *Let $A, A' \in K^{m\times n}$, $B, B' \in K^{n\times r}$, and $C \in K^{r\times s}$. Then the following hold:*

(a) $A(B + B') = AB + AB'$ *and* $(B + B')C = BC + B'C$ *(distributivity).*

(b) $\lambda(AB) = (\lambda A)B = A(\lambda B)$ *for all* $\lambda \in K$ *(homogeneity).*

(c) $(AB)C = A(BC)$ *(associativity).*

(d) $I_m A = AI_n = A$.

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

(f) A, B *are invertible, then* $(AB)^{-1} = B^{-1}A^{-1}$. *In particular,* AB *is invertible.*

(g) *If* A *is invertible, then* $(A^T)^{-1} = (A^{-1})^T$. *In particular,* A^T *is invertible.*

This works because the properties of the field K. However, matricies over a field is not a field.

2 *Vector spaces*

We start with the abstract definition of a vector space over a field K. Although we are mainly interested in the case $K \in \{\mathbb{R}, \mathbb{C}\}$, it is of some benefit to keep things more general where no extra effort is needed.

As advertised, we will now write 0 and 1 for the neutral elements of addition and multiplication in the field K.

Definition 3 (Vector space). *A nonempty set V is called a **vector space** over the field $(K, +, \cdot)$ (with additive/multiplicative neutral 0 and 1, respectively) if there exist operations*

$$\odot : K \times V \to V, \quad (\lambda, x) \mapsto \lambda \odot x \quad \textit{(scalar multiplication)}$$

and

$$\oplus : V \times V \to V, \quad (x, y) \mapsto x \oplus y \quad \textit{(vector addition)}$$

such that the following hold:

(V1) $(V, \oplus)$ *is an **abelian group**:*

- *(i)* $(x \oplus y) \oplus z = x \oplus (y \oplus z)$ *for all* $x, y, z \in V$ *(associativity);*
- *(ii) There exists* $o \in V$ *such that* $o \oplus x = x$ *for all* $x \in V$ *(neutral element);*
- *(iii) For every* $x \in V$ *there exists* $-x \in V$ *such that* $x \oplus (-x) = o$ *(inverse element);*
- *(iv)* $x \oplus y = y \oplus x$ *for all* $x, y \in V$ *(commutativity).*

(V2) For all $\lambda, \mu \in K$, $x \in V$ *we have*

$$(\lambda \cdot \mu) \odot x = \lambda \odot (\mu \odot x) \quad \textit{(mixed associativity).}$$

(V3) For all $\lambda, \mu \in K$, $x, y \in V$ *we have*

$$\lambda \odot (x \oplus y) = (\lambda \odot x) \oplus (\lambda \odot y) \quad \text{and} \quad (\lambda + \mu) \odot x = (\lambda \odot x) \oplus (\mu \odot x) \quad \textit{(mixed distributivity).}$$

(V4) For all $x \in V$ *we have*

$$1 \odot x = x \quad \textit{(identity law).}$$

Lemma 1. *Let M be a subset of the vector space V over K. Then the span M is a subspace of V.*

Theorem 4 (Linear hull = linear span). *Let V be a vector space over K and $M \subset V$.*

$$linM = spanM$$

Definition 4 (Sum of subspaces). *Let V be a vector space over K and let $U_1, U_2 \subset V$ be subspaces. Then*

$$U_1 + U_2 := \{u_1 + u_2 | u_i \in U_i \ (i = 1, 2)\}$$

is called the (Minkowski) sum of U_1 and U_2. If $U_1 \cap U_2 = \{0\}$ then this sum is called direct and we we write $U_1 \oplus U_2$.

Theorem 5 (Sum of subspaces). *Let V be a vector space over K and let $U_1, U_2 \subset V$ be subspaces. Then the following hold:*

1. *$U_1 + U_2$ is a subspace of V.*

2. *$U_1 + U_2 = lin(U_1 \cup U_2)$*

Definition 5 (Linear independence). *Let V be a vector space over a field K.*
a) Finitely many vectors** $x_1, \ldots, x_n \in V$ **are called** linearly independent **if the equation

$$\sum_{i=1}^{n} \lambda_i x_i = 0, \quad \lambda_i \in K \quad (i = 1, \ldots, n)$$

***only has the trivial solution** $\lambda_1 = \cdots = \lambda_n = 0$.*
b) An arbitrary set** $M \subset V$ **is called linearly independent if every collection of finitely many** (pairwise) different **vectors from** M **are linearly independent. In addition we let the empty set** $\emptyset$ **be linearly independent.

The only correct definition of a vector is it's an element of a vector space.

These are all abstract. It says nothings about how the scalar multiplication works and how it looks like. We are not defining anything here. We are saying if we have these operations satisfying these properties, then we have a vector space.

References