MATH1050 Linear algebra beyond systems of linear equations and manipulation of matrices.

0. (a) The main objects of study in a beginning *linear algebra* course are systems of linear equations, and the manipulations with matrices.

   In a 'second' *linear algebra* course, the focus will be shifted to the 'linear structures' in various mathematical objects and 'linear transformations' between such objects which 'preserve' 'linear relations' within such objects.

   Systems of linear equations and matrices arise in a natural way when we study concrete examples of such 'linear structures' and 'linear transformations', and when we deduce theoretical results through the manipulations of 'linear relations'.

   Various (theoretical) results about systems of linear equations and matrices in the beginning *linear algebra* courses will serve as examples or illustrations for more general results in such a 'second' *linear algebra* course.

   (b) Everything in a beginning *linear algebra* course about systems of linear equations and matrices which can be described in terms of the arithmetical operations $+, -, \times, \div$ in $\mathbb{R}$ alone (such as Gaussian elimination, row-operations, reduced row-echelon forms, matrix addition and multiplication) can be carried over when the field $\mathbb{R}$ is replaced by a general field $\mathbb{F}$, and $+, -, \times, \div$ in $\mathbb{R}$ are replaced by the corresponding operations in $\mathbb{F}$.

   (c) This handout is meant to provide a glimpse into the rudiments in a 'second' *linear algebra* course.

   As the definitions and the theoretical results are about abstract mathematical objects, it is natural to formulate them in *set language*.

   Concrete examples which illustrates these definitions and results are drawn from material covered in a beginning *calculus* course, or a beginning *linear algebra* course, or whatever have been done in this course.

   The justification for the theoretical results and the claims in the concrete examples are left as exercises in the use of sets and functions in *set language*.

1. Recall the results Theorem (1), Theorem (2), Theorem (3) from a beginning *linear algebra* course.

Theorem (1).

Let A be an $(m \times n)$-matrix with real entries. Denote its null space by $\mathcal{N}(A)$. (By definition, $\mathcal{N}(A) = \{\mathbf{x} \in \mathbb{R}^n : A\mathbf{x} = \mathbf{0}_m\}$.) The statements below hold:

(a) *$(\mathcal{N}(A), +)$ is an abelian group, with additive identity $\mathbf{0}_n$.*

(b) *For any $\mathbf{x} \in \mathcal{N}(A)$, for any $\alpha, \beta \in \mathbb{R}$, $(\alpha\beta)\mathbf{x} = \alpha(\beta\mathbf{x})$.*

(c) *$1\mathbf{x} = \mathbf{x}$ for any $\mathbf{x} \in \mathcal{N}(A)$.*

(d) *For any $\mathbf{x} \in \mathcal{N}(A)$, for any $\alpha, \beta \in \mathbb{R}$, $(\alpha + \beta)\mathbf{x} = \alpha\mathbf{x} + \beta\mathbf{x}$.*

(e) *For any $\mathbf{x}, \mathbf{y} \in \mathcal{N}(A)$, for any $\alpha \in \mathbb{R}$, $\alpha(\mathbf{x} + \mathbf{y}) = \alpha\mathbf{x} + \alpha\mathbf{y}$.*

Theorem (2).

Let B be an $(m \times n)$-matrix with real entries. Denote its column space by $\mathcal{C}(B)$. (By definition, $\mathcal{C}(B) = \left\{\mathbf{x} \in \mathbb{R}^m : \begin{array}{l}\text{There exists some } \mathbf{u} \in \mathbb{R}^n \\ \text{such that } \mathbf{x} = B\mathbf{u}.\end{array}\right\}$.) The statements below hold:

(a) *$(\mathcal{C}(B), +)$ is an abelian group, with additive identity $\mathbf{0}_m$.*

(b) *For any $\mathbf{x} \in \mathcal{C}(B)$, for any $\alpha, \beta \in \mathbb{R}$, $(\alpha\beta)\mathbf{x} = \alpha(\beta\mathbf{x})$.*

(c) *$1\mathbf{x} = \mathbf{x}$ for any $\mathbf{x} \in \mathcal{C}(B)$.*

(d) *For any $\mathbf{x} \in \mathcal{C}(B)$, for any $\alpha, \beta \in \mathbb{R}$, $(\alpha + \beta)\mathbf{x} = \alpha\mathbf{x} + \beta\mathbf{x}$.*

(e) *For any $\mathbf{x}, \mathbf{y} \in \mathcal{C}(B)$, for any $\alpha \in \mathbb{R}$, $\alpha(\mathbf{x} + \mathbf{y}) = \alpha\mathbf{x} + \alpha\mathbf{y}$.*

Theorem (3).

Denote by $\mathsf{Mat}_{m\times n}(\mathbb{R})$ the set of all $(m \times n)$-matrices with real entries. The statements below hold:

(a) *$(\mathsf{Mat}_{m\times n}(\mathbb{R}), +)$ is an abelian group, with additive identity being the zero matrix $\mathbf{0}_{m\times n}$.*

(b) *For any $A \in \mathsf{Mat}_{m\times n}(\mathbb{R})$, for any $\alpha, \beta \in \mathbb{R}$, $(\alpha\beta)A = \alpha(\beta A)$.*

(c) $1A = A$ *for any* $A \in \mathsf{Mat}_{m\times n}(\mathbb{R})$.

(d) *For any* $A \in \mathsf{Mat}_{m\times n}(\mathbb{R})$, *for any* $\alpha, \beta \in \mathbb{R}$, $(\alpha+\beta)A = \alpha A + \beta A$.

(e) *For any* $A, B \in \mathsf{Mat}_{m\times n}(\mathbb{R})$, *for any* $\alpha \in \mathbb{R}$, $\alpha(A+B) = \alpha A + \alpha B$.

Theorem (1), Theorem (2), Theorem (3) suggest the presence of some common algebraic structure for various mathematical objects. This mathematical structure is usually referred to as **linear structure** or **vector space structure**.

2. **Definition.**

Let V *be a non-empty set,* $\oplus$ *be a closed binary operation on* V, $\mathbb{F}$ *be a field with addition* $+$ *and multiplication* $\times$, *and* $\odot$ *be a function from* $\mathbb{F}\times V$ *to* V.

For convenience, we agree to write $\alpha\times\beta$ *as* $\alpha\beta$ *and to write* $\odot(\alpha, \mathbf{x})$ *as* $a\odot\mathbf{x}$ *for any* $\mathbf{x}\in V$, *for any* $\alpha,\beta\in\mathbb{F}$.

We say V *is a* **vector space** *over the field* $\mathbb{F}$ *with vector addition* $\oplus$ *and scalar multiplication* $\odot$ *if it satisfies the all the conditions below:*

(VS0) $(V, \oplus)$ *is an abelian group, with additive identity* $\mathbf{0}$.

(VS1) *For any* $\mathbf{x}\in V$, *for any* $\alpha,\beta\in\mathbb{F}$, $(\alpha\beta)\odot\mathbf{x} = \alpha\odot(\beta\odot\mathbf{x})$.

(VS2) $1\odot\mathbf{x} = \mathbf{x}$ *for any* $\mathbf{x}\in V$.

(VS3) *For any* $\mathbf{x}\in V$, *for any* $\alpha,\beta\in\mathbb{F}$, $(\alpha+\beta)\odot\mathbf{x} = (\alpha\odot\mathbf{x}) + (\beta\odot\mathbf{x})$.

(VS4) *For any* $\mathbf{x},\mathbf{y}\in V$, *for any* $\alpha\in\mathbb{F}$, $\alpha\odot(\mathbf{x}+\mathbf{y}) = (\alpha\odot\mathbf{x}) + (\alpha\odot\mathbf{y})$.

Remarks on terminologies and notations.

- The elements of V are referred to as **vectors** in V. The special vector $\mathbf{0}$ is called the **zero vector** in V.
- The field $(\mathbb{F}, +, \times)$ is referred to as the **underlying field** of the vector space V.
- The statement (b) is referred to as the **Law of Associativity** for scalar multiplication of vectors.
- The statement (d) and the statement (e) are collectively known as the **Distributive Laws** for vector addition and scalar multiplication.
- From now on, where there is no ambiguity, we write $\mathbf{x}\oplus\mathbf{y}$ as $\mathbf{x}+\mathbf{y}$ and $\alpha\odot\mathbf{x}$ as $\alpha\mathbf{x}$ for any $\mathbf{x},\mathbf{y}\in V$, for any $\alpha\in\mathbb{F}$.

3. **Examples of vector spaces.**

(a) Every subspace of $\mathbb{R}^n$ is a vector space over $\mathbb{R}$.

(b) $\mathsf{Mat}_{m\times n}(\mathbb{R})$ is a vector space over $\mathbb{R}$.

(c) Suppose $\mathbb{F}$ is a field. Regard $\mathbb{F}^n$ as the set of all column vectors with n entries in the field $\mathbb{F}$. Define vector addition $+$ and scalar multiplication $\cdot$ for such column vectors, through addition and multiplication in the field $\mathbb{F}$, in an analogous way as that for column vectors with real entries:

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} x_1+y_1 \\ x_2+y_2 \\ \vdots \\ x_n+y_n \end{bmatrix}, \qquad c\cdot\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} cx_1 \\ cx_2 \\ \vdots \\ cx_n \end{bmatrix}$$

$\mathbb{F}^n$ is a vector space over $\mathbb{F}$ with vector addition $+$ and scalar multiplication $\cdot$.

(d) Suppose $\mathbb{F}$ is a field. Define $\mathsf{Mat}_{m\times n}(\mathbb{F})$ to be the set of all $(m\times n)$-matrix in the field $\mathbb{F}$. Define matrix addition $+$ and scalar multiplication $\cdot$ for such matrices, through addition and multiplication in the field $\mathbb{F}$, in an analogous way as that for matrices with real entries:

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ \vdots & \vdots & & \vdots \\ b_{m1} & b_{m2} & \cdots & b_{mn} \end{bmatrix} = \begin{bmatrix} a_{11}+b_{11} & a_{12}+b_{12} & \cdots & a_{1n}+b_{1n} \\ a_{21}+b_{21} & a_{22}+b_{22} & \cdots & a_{2n}+b_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1}+b_{m1} & a_{m2}+b_{m2} & \cdots & a_{mn}+b_{mn} \end{bmatrix},$$

$$c\cdot\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} = \begin{bmatrix} ca_{11} & ca_{12} & \cdots & ca_{1n} \\ ca_{21} & ca_{22} & \cdots & ca_{2n} \\ \vdots & \vdots & & \vdots \\ ca_{m1} & ca_{m2} & \cdots & ca_{mn} \end{bmatrix}$$

$\mathsf{Mat}_{m\times n}(\mathbb{F})$ is a vector space over $\mathbb{F}$ with vector addition $+$ and scalar multiplication $\cdot$.

(e) Denote by $\mathbb{R}[x]$ the set of all polynomials with real coefficients.

$\mathbb{R}[x]$ is a vector space over $\mathbb{R}$ with vector addition given by polynomial addition and with scalar multiplication given by multiplication of polynomials by constants.

(f) Let I be an interval in $\mathbb{R}$. Denote by $\mathsf{Map}(I,\mathbb{R})$ the set of all real-valued functions with one real variable, with domain I.

Define (pointwise) addition $+$ and scalar multiplication $\cdot$ for such functions, through addition and multiplication in $\mathbb{R}$, in a natural way:

- For any $f, g \in \mathsf{Map}(I,\mathbb{R})$, $f+g$ is the real-valued function of one real-variable with domain I given by $(f+g)(x) = f(x)+g(x)$ for any $x \in I$.
- For any $f \in \mathsf{Map}(I,\mathbb{R})$, for any $a \in \mathbb{R}$, $a\cdot f$ is the real-valued function of one real-variable with domain I given by $(a\cdot f)(x) = af(x)$ for any $x \in I$.

$\mathsf{Map}(I,\mathbb{R})$ is a vector space over $\mathbb{R}$ with vector addition $+$ and scalar multiplication $\cdot$.

(g) Let S be a set, and $\mathbb{F}$ be a field. Denote by $\mathsf{Map}(S,\mathbb{F})$ the set of all functions with domain S and range $\mathbb{F}$.

Define (pointwise) addition $+$ and scalar multiplication $\cdot$ for such functions, through addition and multiplication in $\mathbb{F}$, in an analogous way as that for real-valued functions of one real variable:

- For any $f, g \in \mathsf{Map}(S,\mathbb{F})$, $f+g$ is the function with domain S given by $(f+g)(x) = f(x)+g(x)$ for any $x \in S$.
- For any $f \in \mathsf{Map}(S,\mathbb{F})$, for any $a \in \mathbb{F}$, $a\cdot f$ is the function of one real-variable with domain S given by $(a\cdot f)(x) = af(x)$ for any $x \in S$.

$\mathsf{Map}(S,\mathbb{F})$ is a vector space over $\mathbb{R}$ with vector addition $+$ and scalar multiplication $\cdot$.

(h) Denote by $\mathsf{Map}(\mathbb{N},\mathbb{F})$ the set of all infinite sequences with entries in $\mathbb{F}$.

$\mathsf{Map}(\mathbb{N},\mathbb{F})$ is a vector space over $\mathbb{F}$ with vector addtion and scalar multiplication, given respectively by (term-by-term) addition $+$ and scalar multiplication $\cdot$ for such infinite sequences, through addition and multiplication in the field $\mathbb{F}$ in a natural way:

$$\{a_n\}_{n=0}^{\infty} + \{b_n\}_{n=0}^{\infty} = \{a_n + b_n\}_{n=0}^{\infty}, \qquad c\cdot\{a_n\}_{n=0}^{\infty} = \{ca_n\}_{n=0}^{\infty}.$$

4. **Definition.**

Let V be a vector space over a field $\mathbb{F}$. Let W be a non-empty subset of V.

W is said to form a **(vector) subspace of** V **over** $\mathbb{F}$ *(or simply,* **$\mathbb{F}$-linear subspace of** *V) if $\alpha\mathbf{x} + \beta\mathbf{y}$ for any $\mathbf{x}, \mathbf{y} \in W$ for any $\alpha, \beta \in \mathbb{F}$.*

5. **Theorem (4).**

Let V be a vector space over a field $\mathbb{F}$. The statements below hold:

(a) *$\{\mathbf{0}\}$ and V are subspaces of V over $\mathbb{F}$.*

(b) *Suppose W is a subset of V. Then W is a vector space over $\mathbb{F}$ iff W is a subspace of V over $\mathbb{F}$.*

(c) i. *Suppose W is a subspace of V over $\mathbb{F}$. Then W is a subspace of W over $\mathbb{F}$.*

ii. *Let W_1, W_2 be subspaces of V over $\mathbb{F}$. Suppose W_1 is a subspace of W_2 over $\mathbb{F}$, and W_2 is a subspace of W_1 over $\mathbb{F}$. Then $W_1 = W_2$.*

iii. *Let W_1, W_2, W_3 are subspaces of V over $\mathbb{F}$. Suppose W_1 is a subspace of W_2 over $\mathbb{F}$, and W_2 is a subspace of W_3 over $\mathbb{F}$. Then W_1 is a subspace of W_3 over $\mathbb{F}$.*

(d) *Suppose W_1, W_2 are subspaces of V over $\mathbb{F}$. Then $W_1 \cap W_2$ is a subspace of V over $\mathbb{F}$.*

(e) *Suppose W_1, W_2 are subspaces of V over $\mathbb{F}$. Then $W_1 \cup W_2$ is a subspace of V over $\mathbb{F}$ iff one of W_1, W_2 is a subspace of the other over $\mathbb{F}$.*

6. **More examples of vector spaces.**

(a) Suppose $\mathbb{F}$ is a field. Define matrix multiplication for matrices with entries in $\mathbb{F}$, through addition and multiplication in the field $\mathbb{F}$, in an analogous way as that for matrices with real entries:

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1p} \\ b_{21} & b_{22} & \cdots & b_{2p} \\ \vdots & \vdots & & \vdots \\ b_{n1} & b_{m2} & \cdots & b_{np} \end{bmatrix} = \begin{bmatrix} \sum_{k=1}^{n} a_{1k}b_{k1} & \sum_{k=1}^{n} a_{1k}b_{k2} & \cdots & \sum_{k=1}^{n} a_{1k}b_{kp} \\ \sum_{k=1}^{n} a_{2k}b_{k1} & \sum_{k=1}^{n} a_{2k}b_{k2} & \cdots & \sum_{k=1}^{n} a_{2k}b_{kp} \\ \vdots & \vdots & & \vdots \\ \sum_{k=1}^{n} a_{mk}b_{k1} & \sum_{k=1}^{n} a_{mk}b_{k2} & \cdots & \sum_{k=1}^{n} a_{mk}b_{kp} \end{bmatrix}.$$

Let A be an $(m \times n)$-matrix with entries in the field $\mathbb{F}$. Define the null space $\mathcal{N}(A)$ of the matrix A by $\mathcal{N}(A) = \{\mathbf{x} \in \mathbb{F}^n : A\mathbf{x} = \mathbf{0}_n\}$. (Here $\mathbf{0}_n$ is the zero vector in $\mathbb{F}^n$.)

i. $\mathcal{N}(A)$ is a subspace of $\mathbb{F}^n$ over $\mathbb{F}$.

ii. Further suppose B is an $(p \times n)$-matrix with entries in the field $\mathbb{F}$, and C is the $((m+p) \times n)$-matrix given by

$$C = \left[\frac{A}{B} \right].$$

Then $\mathcal{N}(A) \cap \mathcal{N}(B) = \mathcal{N}(C)$.

(b) For each non-negative integer n, denote by $\mathbb{R}_{\leq n}[x]$ the set of all polynomials with real coefficients and of degree at most n.

For each non-negative integer n, $\mathbb{R}_{\leq n}[x]$ is a subspace of $\mathbb{R}[x]$ over $\mathbb{R}$.

For any non-negative integers m, n, if $m \leq n$ then $\mathbb{R}_{\leq m}[x]$ is a subspace of $\mathbb{R}_{\leq n}[x]$.

(c) Let I be an interval in $\mathbb{R}$. Denote by $C(I)$ the set of all real-valued functions of one real variable with domain I which are continuous on I.

i. $C(I)$ is a subspace of $\mathsf{Map}(I, \mathbb{R})$ over $\mathbb{R}$.

ii. For each positive integer n, denote by $C^n(I)$ the set of all real-valued functions of one real variable with domain I which are n-times continuously differentiable on I.
For each positive integer n, $C^n(I)$ is a subspace of $C(I)$ over $\mathbb{R}$.
For any positive integers m, n, if $m \leq n$, then $C^n(I)$ is a subspace of $C^m(I)$ over $\mathbb{R}$.

(d) Denote by $\ell_2(\mathbb{R})$ the set of all square-summable infinite sequences of real numbers.

By Triangle Inequality for square-summable infinite sequences, the (term-by-term) sum of any two square-summable infinite sequences of real numbers is also a square-summable infinite sequence of real numbers.

We can also deduce that the scalar multiple of a square-summable infinite sequence of real numbers is also a square-summable infinite sequence of real numbers.

Then $\ell_2(\mathbb{R})$ is a subspace of $\mathsf{Map}(\mathbb{N}, \mathbb{R})$ over $\mathbb{R}$.

(e) Let $\mathbb{E}$ be a field with addition $+$ and multiplication $\times$. Suppose $\mathbb{F}$ is a subfield of $\mathbb{E}$. (So by definition, $\mathbb{E}$ is a field extension of $\mathbb{F}$.)

Then E is a vector space over $\mathbb{F}$ with vector addition and scalar multiplication given respectively by addition in E and multiplication in E.

For example, each of $\mathbb{R}, \mathbb{C}$ is a vector space over $\mathbb{Q}$, while $\mathbb{R}$ is a subspace of $\mathbb{C}$ over $\mathbb{Q}$.

7. **Theorem (5).**

Let V be a vector space over $\mathbb{F}$. Suppose W_1, W_2 are subspaces of V over $\mathbb{F}$.

Define $W_1 + W_2 = \{\mathbf{u} \in V : \mathbf{u} = \mathbf{x} + \mathbf{y}$ for some $\mathbf{x} \in W_1, \mathbf{y} \in W_2\}$.

Then $W_1 + W_2$ is a subspace of V over $\mathbb{F}$.

Remark on terminology.

As vector spaces, $W_1 + W_2$ is called the **sum** of W_1, W_2.

Theorem (6).

Let V be a vector space over a field $\mathbb{F}$. Suppose W_1, W_2, W_3 subspaces of V over $\mathbb{F}$. The statements below hold:

(a) *W_1, W_2 is a subspace of $W_1 + W_2$ over $\mathbb{F}$, and $W_1 \cup W_2$ is a subset of $W_1 + W_2$ over $\mathbb{F}$.*

(b) *$W_1 + W_2 = W_2$ iff W_1 is a subspace of W_2 over $\mathbb{F}$.*

(c) *$W_1 + W_2 = W_2 + W_1$.*

(d) *$(W_1 + W_2) + W_3 = W_1 + (W_2 + W_3)$.*

(e) *Suppose W_1 is a subspace of W_3 over $\mathbb{F}$. Then $(W_1 + W_2) \cap W_3 = W_1 + (W_2 \cap W_3)$.*

8. **Further examples of vector spaces.**

(a) Let A be an $(m \times n)$-matrix with entries in a field $\mathbb{F}$. Define the column space $\mathcal{C}(A)$ of the matrix A by $\mathcal{C}(A) = \left\{ \mathbf{x} \in \mathbb{F}^m : \begin{array}{l} \text{There exists some } \mathbf{u} \in \mathbb{F}^n \\ \text{such that } \mathbf{x} = A\mathbf{u}. \end{array} \right\}$.

i. $\mathcal{C}(A)$ is a subspace of $\mathbb{F}^m$ over $\mathbb{F}$.

ii. Further suppose B is an $(m \times p)$-matrix with entries in the field $\mathbb{F}$, and C is the $(m \times (n+p))$-matrix given by

$$C = \left[\begin{array}{c|c} A & B \end{array}\right].$$

Then $\mathcal{C}(A) + \mathcal{C}(B) = \mathcal{C}(C)$.

(b) Let $\mathbb{F}$ be a field. Let A be an $(n \times n)$-square matrix with entries in $\mathbb{F}$.

- A is said to be **symmetric** if $A^t = A$.
- A is said to be **skew-symmetric** if $A^t = -A$.

Denote by $\mathrm{Sym}_n(\mathbb{F})$, $\mathrm{Skew}_n(\mathbb{F})$ respectively the set of all symmetric $(n \times n)$-square matrices and the set of all skew-symmetric $(n \times n)$-square matrices with entries in the field $\mathbb{F}$.

Each of $\mathrm{Sym}_n(\mathbb{F})$, $\mathrm{Skew}_n(\mathbb{F})$ is a subspace of $\mathsf{Mat}_{n\times n}(\mathbb{F})$ over $\mathbb{F}$.

It happens that $\mathrm{Sym}_n(\mathbb{F}) + \mathrm{Skew}_n(\mathbb{F}) = \mathsf{Mat}_{n\times n}(\mathbb{F})$.

9. **Definition.**

Suppose that A is an $(m \times n)$-matrix with real entries.

Define the function $L_A : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ by $L_A(\mathbf{x}) = A\mathbf{x}$ for any $\mathbf{x} \in \mathbb{R}^n$.

Then L_A is called the **linear transformation defined by matrix multiplication from the left by** *A.*

The name 'linear transformation' is due to the validity of the result Theorem (7) and Corollary to Theorem (7) below.

Theorem (7).

Let A be an $(m\times n)$-matrix with real entries. For any $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, for any $\alpha, \beta \in \mathbb{R}$, $L_A(\alpha x+\beta y) = \alpha L_A(x)+\beta L_A(y)$.

Corollary to Theorem (7).

Let A be an $(m \times n)$-matrix with real entries. Let $\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_k \in \mathbb{R}^n$, and $\alpha_1, \alpha_2, \cdots, \alpha_k \in \mathbb{R}$.

Suppose $\alpha_1\mathbf{x}_1 + \alpha_2\mathbf{x}_2 + \cdots + \alpha_k\mathbf{x}_k = \mathbf{0}_n$. Then $\alpha_1 L_A(\mathbf{x}_1) + \alpha_2 L_A(\mathbf{x}_2) + \cdots + \alpha_k L_A(\mathbf{x}_k) = \mathbf{0}_m$.

Remark. So L_A preserves an arbitrary valid linear relation in $\mathbb{R}^n$, say, $\alpha_1\mathbf{x}_1 + \alpha_2\mathbf{x}_2 + \cdots + \alpha_k\mathbf{x}_k = \mathbf{0}_n$, in the sense that the linear relation in $\mathbb{R}^m$ resultant from replacing each $\mathbf{x}_j$ with $L_A(\mathbf{x}_j)$ and replacing $\mathbf{0}_n$ by $\mathbf{0}_m$ in the original one will remain valid.

10. Matrix multiplication and composition of linear transformation are linked up by Theorem (8).

Theorem (8).

Suppose A is an $(m \times n)$-matrix with real entries and B is an $(n \times p)$-matrix with real entries.

Then $L_A \circ L_B = L_{AB}$ as functions.

11. The result below tells us that every linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^m$ comes from matrix multiplication from the left.

Theorem (9).

For each positive integer p, for each $k = 1, 2, \cdots, p$, denote by $\mathbf{e}_k^{(p)}$ the column vector with p real entries, amongst which the k-th entry is 1 and all other entries are 0.

Let $T : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ be a function. Suppose for any $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, for any $\alpha, \beta \in \mathbb{R}$, $T(\alpha\mathbf{x} + \beta\mathbf{y}) = \alpha T(\mathbf{x}) + \beta T(\mathbf{y})$.

Define $A = \left[\; T(\mathbf{e}_1^{(n)}) \;\middle|\; T(\mathbf{e}_2^{(n)}) \;\middle|\; \cdots \;\middle|\; T(\mathbf{e}_n^{(n)}) \; \right]$.

Then $T = L_A$ as functions.

12. The preservation of 'linear relation' by 'transformations', in the sense of Theorem (7) and Corollary to Theorem (7), is common-place in mathematics that it motivates the definition for the notion of linear transformation from a general vector space to a general vector space.

Definition.

Let V, W be vector spaces over a field $\mathbb{F}$, and $T : V \longrightarrow W$ be a function. The function T is called a **linear transformation over $\mathbb{F}$** *if for any $\mathbf{x}, \mathbf{y} \in V$, for any $\alpha, \beta \in \mathbb{F}$, $T(\alpha\mathbf{x} + \beta\mathbf{y}) = \alpha T(\mathbf{x}) + \beta T(\mathbf{y})$.*

Theorem (10).

Let V, W be vector spaces over a field $\mathbb{F}$, whose respective zero vectors are $\mathbf{0}_V$, $\mathbf{0}_W$. Let $T : V \longrightarrow W$ be a linear tranformation over $\mathbb{F}$. Let $\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_k \in V$, and $\alpha_1, \alpha_2, \cdots, \alpha_k \in \mathbb{F}$.

Suppose $\alpha_1\mathbf{x}_1 + \alpha_2\mathbf{x}_2 + \cdots + \alpha_k\mathbf{x}_k = \mathbf{0}_V$. Then $\alpha_1 T(\mathbf{x}_1) + \alpha_2 T(\mathbf{x}_2) + \cdots + \alpha_k T(\mathbf{x}_k) = \mathbf{0}_W$.

Remark. Hence T 'transforms' the 'linear relation' $\alpha_1\mathbf{x}_1 + \alpha_2\mathbf{x}_2 + \cdots + \alpha_k\mathbf{x}_k = \mathbf{0}_V$ in the vector space V into the 'linear relation' $\alpha_1 T(\mathbf{x}_1) + \alpha_2 T(\mathbf{x}_2) + \cdots + \alpha_k T(\mathbf{x}_k) = \mathbf{0}_W$ in the vector space W.

13. **Theorem (11).**

Let V_1, V_2, V_3 be vector spaces over $\mathbb{F}$, and $T : V_1 \longrightarrow V_2$, $S : V_2 \longrightarrow V_3$ be functions.

Suppose T, S are linear transformations over $\mathbb{F}$. Then $S \circ T : V_1 \longrightarrow V_3$ is a linear transformation over $\mathbb{F}$.

14. **Examples of linear transformation.**

(a) Let $\mathbb{F}$ be a field.

i. Suppose that A is an $(m \times n)$-matrix with entries in $\mathbb{F}$.
Define the function $L_A : \mathbb{F}^n \longrightarrow \mathbb{F}^m$ by $L_A(\mathbf{x}) = A\mathbf{x}$ for any $\mathbf{x} \in \mathbb{F}^n$.
Then L_A is a linear transformation from $\mathbb{F}^n$ to $\mathbb{F}^m$.
We also call L_A the **linear transformation defined by matrix multiplication from the left by A.**

ii. For each positive integer p, for each $k = 1, 2, \cdots, p$, denote by $\mathbf{e}_k^{(p)}$ the column vector with p real entries, amongst which the k-th entry is 1 and all other entries are 0.
Let $T : \mathbb{F}^n \longrightarrow \mathbb{F}^m$ be a function. Suppose for any $\mathbf{x}, \mathbf{y} \in \mathbb{F}^n$, for any $\alpha, \beta \in \mathbb{R}$, $T(\alpha\mathbf{x} + \beta\mathbf{y}) = \alpha T(\mathbf{x}) + \beta T(\mathbf{y})$.
Define the $(m \times n)$-matrix A with entries in $\mathbb{F}$ by $A = \left[\; T(\mathbf{e}_1^{(n)}) \;\middle|\; T(\mathbf{e}_2^{(n)}) \;\middle|\; \cdots \;\middle|\; T(\mathbf{e}_n^{(n)}) \; \right]$.
Then $T = L_A$ as functions.

(b) Let $\mathbb{F}$ be a field.

i. For each $(n \times n)$-square matrix A with entries in $\mathbb{F}$, we define the trace of A to be the sum of the diagonal entries of A, and we denote it by $\mathrm{tr}(A)$.
Define the function $T : \mathsf{Mat}_{n\times n}(\mathbb{F}) \longrightarrow \mathbb{F}$ by $T(A) = \mathrm{tr}(A)$ for any $A \in \mathsf{Mat}_{n\times n}(\mathbb{F})$.
T is a linear transformation from $\mathsf{Mat}_{n\times n}(\mathbb{F})$ to $\mathbb{F}$.

ii. Let A be an $(m \times n)$-matrix with entries in $\mathbb{F}$.
Define the function $M_A : \mathsf{Mat}_{n\times p}(\mathbb{F}) \longrightarrow \mathsf{Mat}_{m\times p}(\mathbb{F})$ by $M_A(B) = AB$ for any $B \in \mathsf{Mat}_{n\times p}(\mathbb{F})$.
M_A is a linear transformation from $\mathsf{Mat}_{n\times p}(\mathbb{F})$ to $\mathsf{Mat}_{m\times p}(\mathbb{F})$.

iii. Let G be an $(n \times n)$-square matrix with entries in $\mathbb{F}$.
Define $\mathfrak{L}_G : \mathsf{Mat}_{n\times n}(\mathbb{F}) \longrightarrow \mathsf{Mat}_{n\times n}(\mathbb{F})$ by $\mathfrak{L}_G(H) = GH - HG$ for any $H \in \mathsf{Mat}_{n\times n}(\mathbb{F})$.
$\mathfrak{L}_G$ is a linear transformation from $\mathsf{Mat}_{n\times n}(\mathbb{F})$ to $\mathsf{Mat}_{n\times n}(\mathbb{F})$.

(c) i. Let $c \in \mathbb{R}$. Define the function $E_c : \mathbb{R}[x] \longrightarrow \mathbb{R}$ by $E_c(f) = f(c)$ for any $f(x) \in \mathbb{R}[x]$.
E_c is a linear transformation from $\mathbb{R}[x]$ to $\mathbb{R}$.

ii. Define the function $T : \mathbb{R}[x] \longrightarrow \mathbb{R}[x]$ by $(T(f))(x) = xf(x)$ for any $f(x) \in \mathbb{R}[x]$.
T is a linear transformation from $\mathbb{R}[x]$ to $\mathbb{R}[x]$.

iii. Define the function $S : \mathbb{R}[x] \longrightarrow \mathbb{R}[x]$ by $(S(f))(x) = f(x) - f(0)$ for any $f(x) \in \mathbb{R}[x]$.
S is a linear transformation from $\mathbb{R}[x]$ to $\mathbb{R}[x]$.

(d) Let J be an open interval in $\mathbb{R}$.

i. Let $c \in J$. Define the function $D_c : C^1(J) \longrightarrow \mathbb{R}$ by $D_c(\varphi) = \varphi'(c)$ for any $\varphi \in C^1(J)$.
D_c is a linear transformation from $C^1(J)$ to $\mathbb{R}$.

ii. Define the function $D : C^1(J) \longrightarrow C(J)$ by $(D(\varphi))(x) = \varphi'(x)$ for any $\varphi \in C^1(J)$ for any $x \in J$.
D is a linear transformation from $C^1(J)$ to $C(J)$.

iii. Let $f \in C(J)$. Define the function $T : C^1(J) \longrightarrow C(J)$ by $(T(\varphi))(x) = \varphi'(x) + f(x)\varphi(x)$ for any $\varphi \in C^1(J)$ for any $x \in J$.
T is a linear transformation from $C^1(J)$ to $C(J)$.

iv. Let $f, g \in C(J)$. Define the function $S : C^2(J) \longrightarrow C(J)$ by $(S(\varphi))(x) = \varphi''(x) + f(x)\varphi'(x) + g(x)\varphi(x)$ for any $\varphi \in C^2(J)$ for any $x \in J$.
S is a linear transformation from $C^2(J)$ to $C(J)$.

(e) Let J be an interval in $\mathbb{R}$.

i. Let $a, b \in \mathbb{R}$. Suppose $a < b$.
Define the function $I_a^b : C(J) \longrightarrow \mathbb{R}$ by $I_a^b(\varphi) = \displaystyle\int_a^b \varphi$ for any $\varphi \in C(J)$.
I_a^b is a linear transformation from $C(J)$ to $\mathbb{R}$.

ii. Let $c \in J$.
Define the function $I_c : C(J) \longrightarrow C^1(J)$ by $I_c(\varphi)(x) = \displaystyle\int_c^x \varphi$ for any $\varphi \in C(J)$ for any $x \in J$.
I_c is a linear transformation from $C(J)$ to $C^1(J)$.

iii. Let $c \in J$. Let $f \in C(J)$.
Define the function $T : C(J) \longrightarrow C^1(J)$ by $T(\varphi)(x) = \displaystyle\int_c^x \varphi \cdot f$ for any $\varphi \in C(J)$ for any $x \in J$.
T is a linear transformation from $C(J)$ to $C^1(J)$.

15. **Theorem (12).**

Let V, W be vector spaces over a field $\mathbb{F}$.

(a) *Suppose $\alpha, \beta \in \mathbb{F}$, and $S, T : V \longrightarrow W$ be linear transformation over $\mathbb{F}$.*
Define the function $\alpha S + \beta T : V \longrightarrow W$ by $(\alpha S + \beta T)(\mathbf{x}) = \alpha S(\mathbf{x}) + \beta T(\mathbf{x})$ for any $\mathbf{x} \in V$.
Then $\alpha S + \beta T$ is a linear transformation from V to W over $\mathbb{F}$.

(b) *Define the set $Hom_{\mathbb{F}}(V, W)$ by $Hom_{\mathbb{F}}(V, W) = \{T \mid T$ is a linear transformation from V to W over $\mathbb{F}\}$.*
$Hom_{\mathbb{F}}(V, W)$ is a vector space over $\mathbb{F}$.

Definition.

Let V be a vector space over $\mathbb{F}$.

(a) *A linear transformation from V to $\mathbb{F}$ is called a* **linear functional on V over $\mathbb{F}$**.

(b) *The* **dual space of V over $\mathbb{F}$** *is the vector space $Hom_{\mathbb{F}}(V, \mathbb{F})$, and is denoted by $V_{\mathbb{F}}^*$.*

Theorem (13).

Let V, W be vector spaces over a field $\mathbb{F}$, and $T : V \longrightarrow W$ be a function. Suppose T is a linear transformation from V to W over $\mathbb{F}$.

Define the function $T^* : W_{\mathbb{F}}^* \longrightarrow V_{\mathbb{F}}^*$ by $(T^*(f))(\mathbf{x}) = f(T(\mathbf{x}))$ for any $\mathbf{x} \in V$ for any $f \in W_{\mathbb{F}}^*$.

T^* is a linear transformation from $W_{\mathbb{F}}^*$ to $V_{\mathbb{F}}^*$.

Remark. T^* is called the **dual linear transformation** of T.

16. **What is 'solving a linear equation'?**

Suppose an equation (with some unknown), say, $(\star)$, can be presented in the form

$$T(\mathbf{u}) = \mathbf{b} \quad \text{——} \quad (\star)$$

in which T is some known linear transformation from some vector space V to some vector space W, $\mathbf{b}$ is some known vector in W, and $\mathbf{u}$ is an unknown vector in V. Then we will say that the equation $(\star)$ is a **linear equation**.

When $\mathbf{b}$ is the zero vector in W, the equation $(\star)$ is said to be **homogeneous**; otherwise it is said to be **non-homogeneous**.

Every vector $\mathbf{a}$ in V, which upon substitution into $(\star)$ results in a true statement, is called a **solution of the equation** $(\star)$.

To solve $(\star)$ is to give a full description of the **set of all solutions of the linear equation** $(\star)$.

17. **Examples of linear equations.**

(a) Every system of linear equations (in a beginning *linear algebra* course) is a linear equation.

(b) Let J be an open interval in $\mathbb{R}$, $p \in J$, and $\varphi : J \longrightarrow \mathbb{R}$ be a function. Suppose φ is continuous on J.
The equation

$$\frac{dy}{dx} = \varphi(x) \quad \text{on } J$$

with 'unknown function y on J' is a linear equation. In terms of the notations given in earlier examples, we can present this equation as

$$D(y) = \varphi \quad \text{——} \quad (*)$$

in which $D : C^1(J) \longrightarrow C(J)$ is the linear transformation assigning each $\psi \in C^1(J)$ to $\psi' \in C(J)$.
Solving this equation amounts to finding all real-valued function of one real variable with domain J and continuously differentiable on J which satisfies this equation.
According to the Fundamental Theorem of the Calculus, $\psi : J \longrightarrow \mathbb{R}$ is a solution of $(\star)$ iff there exists some $C \in \mathbb{R}$ such that $\psi(x) = \int_p^x \varphi(t)dt + C$ for any $x \in J$.

When presented in the language of school maths, what we mean is that

all possible y's which satisfy '$\frac{dy}{dx} = \varphi(x)$ on J' are given by $y = \int_p^x \varphi(t)dt + C$ where C is an arbitrary constant.

$(*)$ is the prototype of all 'linear ordinary differential equations'.

18. **Theorem (14).**

Let V, W be vector spaces over $\mathbb{F}$. Suppose T be a linear transformation from V to W over $\mathbb{F}$.
Define $\mathcal{N}(T) = \{\mathbf{x} \in V : T(\mathbf{x}) = \mathbf{0}\}$. ($\mathcal{N}(T)$ is called the **null space of** *T, or* **kernel of** *T.)*
The statements below hold:

(a) *$\mathcal{N}(T)$ is a subspace of V over $\mathbb{F}$.*

(b) *For any $\mathbf{x}, \mathbf{y} \in V$, $T(\mathbf{x}) = T(\mathbf{y})$ iff $\mathbf{x} - \mathbf{y} \in \mathcal{N}(T)$.*

Remark. Translated into the language of solving linear equations, this result says (for the given linear transformation $T : V \longrightarrow W$):

(a) The solution set of the homogeneous linear equation

$$T(\mathbf{u}) = \mathbf{0} \quad \text{——} \quad (\star_0)$$

with unknown $\mathbf{u}$ in V forms a subspace of V over $\mathbb{F}$, namely, the null space of T.

(b) Let $\mathbf{b} \in W$ and $\mathbf{x}, \mathbf{y} \in V$. Consider the linear equation

$$T(\mathbf{u}) = \mathbf{b} \quad \text{——} \quad (\star)$$

with unknown $\mathbf{u}$ in V.
Suppose '$\mathbf{u} = \mathbf{x}$' is a solution of $(\star)$.
Then '$\mathbf{u} = \mathbf{y}$' is a solution of $(\star)$ iff '$\mathbf{u} = \mathbf{x} - \mathbf{y}$' is a solution of $(\star_0)$.
Moreover the solution set of $(\star)$ is given by $\{\mathbf{z} \in V : \mathbf{z} = \mathbf{x} + \mathbf{h} \text{ for some } \mathbf{h} \in \mathcal{N}(T)\}$.

When $V = \mathbb{R}^n$, $W = \mathbb{R}^m$ and T is the linear transformation defined by matrix multiplication from the left by some $(m \times n)$-matrix with real entries, the above result reduces to the 'structural' result on the relation between the solution set of a system of linear equation and that of its associated homogeneous system, which you should be familiar with in a beginning *linear algebra* course.