Linear Algebra I

Overview notes

G30 Program, Nagoya University (Fall 2020)

Henrik Bachmann (Math. Building Room 457, henrik.bachmann@math.nagoya-u.ac.jp)
Lecture notes and exercises are available at: https://www.henrikbachmann.com/la1_2020.html

---

These notes serve as a compact overview of the definitions, propositions, lemmas, corollaries, and theorems given in the lectures. It will be updated regularly (This is Version 6 from November 6, 2020). The proofs, examples, and explanations are provided in the handwritten notes/lectures. The reference book for this course is [B], and we will probably cover Chapters 1,2,3 and 5 during this semester. (Chapter 4 will be part of Linear Algebra II)

If you find any typos in this note, please let me know!

Contents

References

[B] O. Bretscher: *Linear Algebra with Applications*, 4th edition, Pearson 2009.

1 Linear systems

By $\mathbb{R}$ we will denote the set of real numbers. $\mathbb{R}$ contains all numbers you usually considered in high school, such as $1, -1, 0, 2, 3, \frac{3}{8}, \pi, \sqrt{2}, e, \dots$. There are rigorous definitions of the real numbers, which would be part of a pure mathematics lecture (See for example `https://en.wikipedia.org/wiki/Construction_of_the_real_numbers`). But in this course, we just assume that they exist and that everyone is familiar with them.

Definition 1.1. *i) For real numbers $a_1, a_2, \dots, a_n, b \in \mathbb{R}$ an equation of the form*

$$a_1 x_1 + \cdots + a_n x_n = b$$

is called a **linear equation***.*

ii) A finite collection of linear equations is called a **linear system***.*

iii) A **solution of a linear system** *is a simultaneous solution to all of its equations.*

Definition 1.2. *The following operations on a linear system are called* **elementary row operations***.*

(R1) Add a multiple of an equation to another.

(R2) Multiply an equation with a <u>non-zero</u> number.

(R3) Change the order of the equations.

Proposition 1.3. *Applying an elementary row operation to a linear system does not change the set of all solutions of this linear system.*

2 Matrices & Vectors

Definition 2.1. *i) A $m \times n$-***matrix** *is given by an array (m rows, n columns) of numbers $a_{ij} \in \mathbb{R}$*

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} = (a_{ij})_{\substack{1 \le i \le m \\ 1 \le j \le n}} .$$

Notation: We often just write $A = (a_{ij})$ if the size of A, i.e. m and n, are known from context. By $\mathbb{R}^{m \times n}$ we denote the set all of all $m \times n$-matrices.

ii) A (column-) **vector** *of size n is a $n \times 1$-matrix*

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

and the set of all vectors of size n is denoted by $\mathbb{R}^n = \mathbb{R}^{n \times 1}$.

Definition 2.2. *For matrices $A = (a_{ij}), B = (b_{ij}) \in \mathbb{R}^{m\times n}$ and a real number $\lambda \in \mathbb{R}$ we define*

$$A + B = (a_{ij} + b_{ij}) \in \mathbb{R}^{m\times n} \qquad \text{(Sum of two matrices)},$$
$$\lambda A = (\lambda a_{ij}) \in \mathbb{R}^{m\times n} \qquad \text{(Scalar multiplication)}.$$

In the case $\lambda = -1$ we write $(-1)A = -A$ and $A - B$ means $A + (-1)B$.

The matrices A and B need to be of the same size, otherwise the sum $A + B$ is not defined. A special case of the addition of matrices is given by the addition of vectors. For $u, v \in \mathbb{R}^n$ and $\lambda \in \mathbb{R}$ we have

$$u = \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix}, v = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}, \quad u + v = \begin{pmatrix} u_1 + v_1 \\ \vdots \\ u_n + v_n \end{pmatrix}, \quad \lambda v = \begin{pmatrix} \lambda v_1 \\ \vdots \\ \lambda v_n \end{pmatrix}.$$

Definition 2.3. *The product of a matrix $A = (a_{ij}) \in \mathbb{R}^{m\times n}$ and a vector $v \in \mathbb{R}^n$ is defined by*

$$Av = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = \begin{pmatrix} a_{11}v_1 + a_{12}v_2 + \dots + a_{1n}v_n \\ a_{21}v_1 + a_{22}v_2 + \dots + a_{2n}v_n \\ \vdots \\ a_{m1}v_1 + a_{m2}v_2 + \dots + a_{mn}v_n \end{pmatrix} \in \mathbb{R}^m.$$

We have: ($m\times n$-matrix) $\cdot$ (vector of size n) = (vector of size m).

Proposition 2.4. *We have for $A \in \mathbb{R}^{m\times n}$, $x, y \in \mathbb{R}^n$ and $\lambda \in \mathbb{R}$*

i) $A(x + y) = Ax + Ay$,

ii) $A(\lambda x) = \lambda(Ax)$.

Definition 2.5. *For a matrix $A = (a_{ij}) \in \mathbb{R}^{m\times n}$ and a vector $b \in \mathbb{R}^m$ the matrix*

$$(A \mid b) = \left(\begin{array}{ccc|c} a_{11} & \dots & a_{1n} & b_1 \\ \vdots & \ddots & \vdots & \vdots \\ a_{m1} & \dots & a_{mn} & b_m \end{array}\right) \in \mathbb{R}^{m\times(n+1)}$$

is called the **augmented matrix** *of the linear system $Ax = b$.*

The augmented matrix $(A \mid b)$ is just the matrix A where we append the vector b as a column. The line $\mid$ is a useful notation to distinguish between the left- and right-hand side of the corresponding linear system but it has no mathematical meaning. We will view $(A \mid b)$ as a usual matrix with m rows and $n + 1$ columns.

Definition 2.6. *The following operations on a matrix are called* **elementary row operations**.

(R1) Add a multiple of one row to another row.

(R2) Multiply a row with a non-zero number.

(R3) Interchange two rows.

Applying a row operation to a linear system (Definition 1.2) corresponds to the same row operation (Definition 2.6) on the corresponding augmented matrix of this linear system.

Definition 2.7. *Two matrices A and B are called* **row equivalent***, if B can be obtained from A by elementary row operations. In this case we write*

$$A \sim B\,.$$

Notice that if $A \sim B$, then also $B \sim A$, i.e. A can be obtained from B by elementary row operations.

Proposition 2.8. *Let $A, B \in \mathbb{R}^{m\times n}$ and $b, c \in \mathbb{R}^m$. If $(A \mid b) \sim (B \mid c)$ then the linear systems $Ax = b$ and $Bx = c$ have the same solutions.*

Definition 2.9. *A matrix $A = (a_{ij}) \in \mathbb{R}^{m\times n}$ is on* **row-reduced echelon form** *if*

i) The first non-zero element on each row (if any) is equal to 1.

ii) If there is a leading 1 in a row, then all rows above contain a leading 1 further to the left.

iii) If a_{ij} is the first non-zero element in row i, then there are no other non-zero elements in the j-th column.

The first non-zero element in a row of a matrix in row-reduced echelon form is called **pivot element***.*

Theorem 2.10. *Every matrix A is row equivalent to a unique matrix B on row-reduced echelon form and we write*

$$B = \mathrm{rref}(A)\,.$$

Definition 2.11. *Let $A \in \mathbb{R}^{m\times n}$ be a matrix. The* **rank** $\mathrm{rk}(A)$ *of A is the number of pivot elements in* $\mathrm{rref}(A)$*.*

3 Sets & Functions

Definition 3.1. *Let X and Y be two sets.*

i) A **function** *$f : X \to Y$ is a rule, assigning to* <u>each</u> *element $x \in X$ an element $f(x) \in Y$. This is also denoted by*

$$\begin{aligned} f : X &\longrightarrow Y \\ x &\longmapsto f(x)\,. \end{aligned}$$

ii) For $f : X \to Y$ the set X is called the **domain of** *f and Y is called the* **codomain of** *f.*

A function is also sometimes called a **map**. These two names (for the exact same mathematical object) are used interchangeably in the literature.

Definition 3.2. *For a function $f : X \to Y$ the* **image of** *f is defined by*

$$\mathrm{im}(f) = \{y \in Y \mid \exists x \in X : y = f(x)\}\,.$$

Another notation is $\mathrm{im}(f) = f(X)$. The image is a subset of Y, i.e. $\mathrm{im}(f) \subset Y$.

Definition 3.3. *A function* $f : X \to Y$ *is called*

i) **injective** *if* $f(x_1) \neq f(x_2)$ *whenever* $x_1 \neq x_2$. *(*$x_1, x_2 \in X$*)*

ii) **surjective** *if* $\mathrm{im}(f) = Y$.

iii) **bijective** *if it is injective and surjective.*

4 Linear maps

Definition 4.1. *A function* $F : \mathbb{R}^n \to \mathbb{R}^m$ *is a* **linear map** *if for all* $u, v \in \mathbb{R}^n$, $\lambda \in \mathbb{R}$ *we have*

i) $F(u + v) = F(u) + F(v)$,

ii) $F(\lambda u) = \lambda F(u)$.

Theorem 4.2. *Let* $F : \mathbb{R}^n \to \mathbb{R}^m$ *be a linear map. Then there exists a unique matrix* $[F] \in \mathbb{R}^{m\times n}$, *such that for all* $x \in \mathbb{R}^n$ *we have*

$$F(x) = [F]x .$$

Here the left-hand side is the evaluation of the function F at x and the right-hand side is the multiplication of the matrix $[F]$ with the vector x.

Definition 4.3. *The matrix* $[F]$ *in Theorem 4.2 is called* **the matrix of F**.

5 Linear maps in geometry

Definition 5.1. *Let* $u = \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix}, v = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \in \mathbb{R}^n$.

i) The **dot product** *of* u *and* v *is defined by*

$$u \bullet v = u_1 v_1 + \cdots + u_n v_n .$$

ii) u *and* v *are called* **orthogonal** *if* $u \bullet v = 0$.

iii) The **norm** *of* u *is defined by*

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

Proposition 5.2. *The dot product satisfies the following properties for all* $u, v, w \in \mathbb{R}^n$ *and* $\lambda \in \mathbb{R}$:

i) $u \bullet v = v \bullet u$,

ii) $u \bullet (v + w) = u \bullet v + u \bullet w$,

iii) $u \bullet (\lambda v) = \lambda(u \bullet v)$.

Version 6 (November 6, 2020)

Definition 5.3. *Let $u \in \mathbb{R}^n$ with $u \neq 0$. We define the* **orthogonal projection** P_u *onto the line spanned by u as*

$$\begin{aligned} P_u : \mathbb{R}^n &\longrightarrow \mathbb{R}^n \\ x &\longmapsto \frac{u \bullet x}{u \bullet u} u \,. \end{aligned}$$

Proposition 5.4. *$P_u : \mathbb{R}^n \to \mathbb{R}^n$ is a linear map.*

Definition 5.5. *Let $u \in \mathbb{R}^n$ with $u \neq 0$. We define the* **reflection** ρ_u *along the line spanned by u as*

$$\begin{aligned} \rho_u : \mathbb{R}^n &\longrightarrow \mathbb{R}^n \\ x &\longmapsto 2\frac{u \bullet x}{u \bullet u} u - x \,. \end{aligned}$$

Proposition 5.6. *$\rho_u : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ is a linear map.*

Definition 5.7. *For $\varphi \in \mathbb{R}$ the counterclockwise* **rotation by an angle** φ *(in $\mathbb{R}^2$) is given by*

$$\begin{aligned} \mathrm{rot}_\varphi : \mathbb{R}^2 &\longrightarrow \mathbb{R}^2 \\ x &\longmapsto \begin{pmatrix} \cos(\varphi) & -\sin(\varphi) \\ \sin(\varphi) & \cos(\varphi) \end{pmatrix} x \,. \end{aligned}$$

6 Composition of linear maps & Matrix multiplication

Theorem 6.1. *If $F : \mathbb{R}^n \to \mathbb{R}^m$ and $G : \mathbb{R}^m \to \mathbb{R}^l$ are linear, then $GF : \mathbb{R}^n \to \mathbb{R}^l$ is linear.*

Definition 6.2. *Let $A \in \mathbb{R}^{l \times m}$ and $B \in \mathbb{R}^{m \times n}$, where B has the columns $v_1, \dots, v_n \in \mathbb{R}^m$, i.e.*

$$B = \begin{pmatrix} | & & | \\ v_1 & \dots & v_n \\ | & & | \end{pmatrix} .$$

Then the **product of** A **and** B *is the $l \times n$-matrix with columns $Av_1, \dots, Av_n \in \mathbb{R}^l$, i.e.*

$$A \cdot B = \begin{pmatrix} | & & | \\ Av_1 & \dots & Av_n \\ | & & | \end{pmatrix} \in \mathbb{R}^{l \times n} .$$

Theorem 6.3. *Let $F : \mathbb{R}^n \to \mathbb{R}^m$ and $G : \mathbb{R}^m \to \mathbb{R}^l$ be linear maps. Then the matrix of GF is given by the product of the matrices of G and F, i.e.*

$$[GF] = [G] \cdot [F] \,.$$

Proposition 6.4. *For all $A \in \mathbb{R}^{l \times m}, B, D \in \mathbb{R}^{m \times n}, C \in \mathbb{R}^{n \times p}$ and $\lambda \in \mathbb{R}$ we have*

i) $A \cdot I_m = I_l \cdot A = A$, *where I_m denotes the $m \times m$-identity matrix.*

ii) $(A \cdot B) \cdot C = A \cdot (B \cdot C)$.

iii) $A \cdot (B + D) = A \cdot B + A \cdot D$.

iv) $(B + D) \cdot C = B \cdot C + D \cdot C$.

v) $\lambda(A \cdot B) = (\lambda A) \cdot B = A \cdot (\lambda B)$.