Linear Algebra

Lecture Notes

Rostyslav Hryniv

Ukrainian Catholic University
Computer Science Programme

3[rd] term
Autumn 2017

Lecture 4. Vector spaces

Outline

$\mathbb{R}^n$ as a vector space

$\mathbb{R}^n$ is the set of all $n \times 1$ column vectors $\mathbf{x} = (x_1, x_2, \dots, x_n)^\top$

Two operations on $\mathbb{R}^n$: addition and multiplication by scalars

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{pmatrix}, \qquad k \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} kx_1 \\ \vdots \\ kx_n \end{pmatrix}$$

Properties of addition and multiplication ($\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{R}^n$, $c, d \in \mathbb{R}$):

1. $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$ commutative
2. $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$ associative
3. $\mathbf{u} + \mathbf{0} = \mathbf{0} + \mathbf{u} = \mathbf{u}$ additive zero
4. $-\mathbf{v} + \mathbf{v} = \mathbf{v} + (-\mathbf{v}) = \mathbf{0}$ additive inverse
5. $c(\mathbf{u} + \mathbf{v}) = c\mathbf{u} + c\mathbf{v}$ distributive
6. $(c + d)\mathbf{u} = c\mathbf{u} + d\mathbf{u}$ distributive
7. $c(d\mathbf{u}) = (cd)\mathbf{u}$ associative
8. $1 \cdot \mathbf{u} = \mathbf{u}$ multiplicative identity

Example ($m \times n$ matrices)

On the set $M_{m\times n}(\mathbb{R})$ of all $m \times n$ matrices with real entries addition and multiplication by scalars are well defined

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} + \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} = \begin{pmatrix} a_{11}+b_{11} & a_{12}+b_{12} \\ a_{21}+b_{21} & a_{22}+b_{22} \end{pmatrix}$$

$$k \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \begin{pmatrix} ka_{11} & ka_{12} \\ ka_{21} & ka_{22} \end{pmatrix}$$

and enjoy properties 1–8:

$$A + B = B + A, \qquad (A + B) + C = A + (B + C)$$

$$A + \mathbf{0} = \mathbf{0} + A = A \qquad (-A) + A = \mathbf{0} \ldots$$

Example (Functions over $[a, b]$)

For functions f, g over $[a, b]$ and $c \in \mathbb{R}$, define $f + g$ and $c \cdot f$ as

$$(f + g)(x) := f(x) + g(x), \qquad (c \cdot f)(x) = c \cdot (f(x))$$

These addition and multiplication enjoy properties 1–8:

$$f + g = g + f, \qquad (f + g) + h = f + (g + h)$$

$$f + \mathbf{0} = \mathbf{0} + f = f \qquad (-f) + f = \mathbf{0} \ldots$$

Definition of vector space

Definition

A linear vector space V is a set of elements $\mathbf{u}, \mathbf{v}, \mathbf{w} \ldots$ (vectors) with two operations:

addition $\quad + : V \times V \to V$

multiplication by scalar $\quad \cdot : \mathbb{R} \times V \to V$

satisfying the following properties:

1. $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$ (commutative)
2. $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$ (associative)
3. $\exists$ zero vector $\mathbf{0}$ in V s.t. $\mathbf{u} + \mathbf{0} = \mathbf{u}$ (additive zero)
4. $\forall \mathbf{v} \in V, \exists (-\mathbf{v}) \in V$ s.t. $(-\mathbf{v}) + \mathbf{v} = \mathbf{0}$ (additive inverse)
5. $c(\mathbf{u} + \mathbf{v}) = c\mathbf{u} + c\mathbf{v}$ (distributive)
6. $(c + d)\mathbf{u} = c\mathbf{u} + d\mathbf{u}$ (distributive)
7. $c(d\mathbf{u}) = (cd)\mathbf{u}$ (associative)
8. $1 \cdot \mathbf{u} = \mathbf{u}$ (multiplicative identity)

Examples of linear vector spaces

- the set $\mathbb{R}^n$ of $n \times 1$ column vectors
- the set $M_{m\times n}(\mathbb{R})$ of $m \times n$ matrices
- the set $\mathcal{F}[a,b]$ of functions over $[a,b]$
- the set $\mathcal{P}_\infty$ of polynomials $a_0 + a_1 x + \cdots + a_n x^n$ with real coefficients ($n \geq 0$ arbitrary)
- the set $\mathbb{R}^\mathbb{N}$ of infinite sequences $\mathbf{x} := (x_1, x_2, x_3, \dots)$, $\mathbf{y} := (y_1, y_2, y_3, \dots)$, **. . .** with term-wise addition and multiplication by scalars:

$$\mathbf{x} + \mathbf{y} = (x_1 + y_1, x_2 + y_2, x_3 + y_3, \dots)$$
$$k \cdot \mathbf{x} = (kx_1, kx_2, kx_3 \dots)$$

- $V = \{\mathbf{0}\}$ is the smallest vector space
- $V = \mathbb{R}_+$ with operation $x + y = xy$, $k \cdot x := x^k$
- $V = \{(x,y,z)^\top \in \mathbb{R}^3 \mid x + y + z = 1\}$ is not a l. v. s.

Theorem

Let V be a vector space, $\mathbf{u}$ a vector in V, and k a scalar; then:

(i) $0 \cdot \mathbf{u} = \mathbf{0}$;

(ii) $k \cdot \mathbf{0} = \mathbf{0}$;

(iii) $(-1) \cdot \mathbf{u} = (-\mathbf{u})$

(iv) *if* $k \cdot \mathbf{u} = \mathbf{0}$, *then either* $k = 0$ *or* $\mathbf{u} = \mathbf{0}$

Proof.

(i) Set $\mathbf{v} := 0 \cdot \mathbf{u}$; then

$$\mathbf{v} + \mathbf{v} = 0 \cdot \mathbf{u} + 0 \cdot \mathbf{u} = (0 + 0) \cdot \mathbf{u} = 0 \cdot \mathbf{u} = \mathbf{v}$$

Adding the inverse $(-\mathbf{v})$ of $\mathbf{v}$ to both sides gives

$$\mathbf{v} = \mathbf{0} + \mathbf{v} = (-\mathbf{v} + \mathbf{v}) + \mathbf{v} = (-\mathbf{v}) + (\mathbf{v} + \mathbf{v}) = (-\mathbf{v}) + \mathbf{v} = \mathbf{0}$$

(ii) $k \cdot \mathbf{0} + \mathbf{v} = k \cdot \mathbf{0} + (k\frac{1}{k}) \cdot \mathbf{v} = k \cdot (\mathbf{0} + \frac{1}{k}\mathbf{v}) = k \cdot (\frac{1}{k} \cdot \mathbf{v}) = \mathbf{v}$

(iii) $(-1) \cdot \mathbf{u} + \mathbf{v} = (-1) \cdot \mathbf{u} + 1 \cdot \mathbf{v} = (-1 + 1) \cdot \mathbf{v} = 0 \cdot \mathbf{v} = \mathbf{0}$ $\square$

Subspaces; definition and examples

Definition

A non-empty subset W of a v.s. V is called its subspace if it is closed under vector addition and multiplication by scalars (ie, if $\mathbf{u}, \mathbf{v} \in W$ and $k \in \mathbb{R}$, then $\mathbf{u} + \mathbf{v} \in W$ and $k \cdot \mathbf{v} \in W$).

Remark

- $\mathbf{0}$ *always belongs to a subspace*
- $\{\mathbf{0}\}$ *is the smallest subspace*
- *A subspace of a linear space is a linear space by itself*

Example

- Subspaces of $\mathbb{R}^2$ are
  - $\mathbb{R}^2$ itself
  - every line through the origin $\{(x, y)^\top \mid ax + by = 0\}$
  - $\{(0, 0)^\top\}$
- Subspaces of $\mathbb{R}^3$ are ...

Which of these sets are subspaces?

- In the vector space $M_{n\times n}(\mathbb{R})$ of square $n \times n$ matrices:
  - diagonal matrices
  - upper/lower-triangular matrices
  - non-singular matrices
  - singular matrices
- In the vector set $\mathcal{F}[a, b]$ of functions over $[a, b]$:
  - continuous functions $C[a, b]$
  - continuously differentiable functions $C^1[a, b]$
  - functions vanishing at $x = a$
  - nonnegative functions
- In the vector space $\mathcal{P}_\infty$ of all polynomials:
  - the set $\mathcal{P}_n$ of all polynomials of degree *at most* n
  - the set of all polynomials of degree *exactly* n
- For a given $A \in M_{m\times n}(\mathbb{R})$,
  - the nullspace $\mathcal{N}(A)$ (the solution set of the equation $A\mathbf{x} = \mathbf{0}$)
  - column space $C(A)$ (linear combinations of its columns)

  are subspaces

Linear combinations and linear spans

Definition (Linear combination)

A linear combination of vectors $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ in a v.s. V is a vector $c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_k\mathbf{v}_k$ with any scalars $c_1, c_2, \dots, c_k$

Definition (Linear span)

For a set $S \subset V$, the collection of all finite linear combinations of vectors from S is called the linear span of S and denoted $\mathrm{ls}\{S\}$

Example

- $\mathbb{R}^n$ is the linear span of the standard unit vectors $\mathbf{e}_1 = (1, 0 \dots, 0)^\top, \dots, \mathbf{e}_n = (0, \dots, 0, 1)^\top$
- $\mathcal{P}_n = \mathrm{ls}\{1, x, \dots, x^n\}$
- $\mathrm{ls}\{1, x, \dots, x^n, \dots\} = \mathcal{P}_\infty \neq C(\mathbb{R})$
- The column space $C(A)$ of an $m \times n$ matrix A is the linear span of columns of A.

  $C(A)$ coicides with the range $\mathcal{R}(A) = \{A\mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n\}$ of A

Properties of linear spans

Theorem

(i) *For any subset S of a v.s. V, $\operatorname{ls}(S)$ is a subspace of V*

(ii) *if W is a subspace of V containing S, then $\operatorname{ls}(S) \subset W$*

(iii) $\operatorname{ls}(S)$ *is the smallest subspace of V containing S*

(iv) $\operatorname{ls}(\operatorname{ls}(S)) = \operatorname{ls}(S)$

Proof.

$\mathbf{u}, \mathbf{v} \in \operatorname{ls}(S) \iff \exists \mathbf{w}_1, \ldots, \mathbf{w}_k \in S,\ a_1, \ldots, a_k, b_1, \ldots, b_k \in \mathbb{R}$ s.t. $\mathbf{u} = a_1\mathbf{w}_1 + \cdots + a_k\mathbf{w}_k, \qquad \mathbf{v} = b_1\mathbf{w}_1 + \cdots + b_k\mathbf{w}_k$. Then

$$\mathbf{u} + \mathbf{v} = (a_1 + b_1)\mathbf{w}_1 + \cdots + (a_k + b_k)\mathbf{w}_k \in \operatorname{ls}(S)$$
$$c\mathbf{u} = ca_1\mathbf{w}_1 + \cdots + ca_k\mathbf{w}_k \in \operatorname{ls}(S)$$

For (ii): any linear combination of elements of S belongs to W

For (iii): the intersection of all subspaces of V containing S is a subspace, contains S, cannot be larger than $\operatorname{ls}(S)$ and cannot be smaller than $\operatorname{ls}(S)$ $\square$

Coincidence of linear spans

Proposition:

$\mathrm{ls}(S_1) = \mathrm{ls}(S_2)$ if and only if

- every $\mathbf{v}_1 \in S_1$ belongs to $\mathrm{ls}(S_2)$ and
- every $\mathbf{v}_2 \in S_2$ belongs to $\mathrm{ls}(S_1)$

Proof.

If every $\mathbf{v}_1 \in S_1$ belongs to $\mathrm{ls}(S_2)$, then $\mathrm{ls}(S_1) \subset \mathrm{ls}(S_2)$ as $\mathrm{ls}(S_1)$ is the smallest subspace containing S_1. By the same reason, $\mathrm{ls}(S_2) \subset \mathrm{ls}(S_1)$. □

Example

Polynomials $(1-x), (1-x)^2, \dots, (1-x)^n, \dots$ do not span $\mathcal{P}_\infty$

Linear independence

Definition (Linear independence)

The set S of vectors of a v.s. V is called linearly independent if no *nontrivial* linear combination of vectors from S equals $\mathbf{0}$

A set S is linearly independent $\iff$

no vector in S is a linear combination of the other vectors in S

Example (Linearly independent sets)

- The basis unit vectors $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$ in $\mathbb{R}^n$
- $S = \{1, x, x^2, \dots, x^n, \dots\}$ is linearly independent in $\mathcal{P}_\infty$
- Columns of $A \in M_{m\times n}(\mathbb{R})$ are linearly independent $\iff$ $A\mathbf{x} = \mathbf{0}$ has no nontrivial solutions $\iff$ $\mathcal{N}(A) = \{\mathbf{0}\}$
- In an echelon form U of A, the rows with pivots and the columns with pivots are linearly independent
- Matrices $E_{ij} = \delta_{ij}$ are linearly independent

Example

Are $1 - x$, $5 + 3x - 2x^2$ and $1 + 3x - x^2$ linearly dependent?

Look for a linear combination $c_1 p_1(x) + c_2 p_2(x) + c_3 p_3(x)$ that is identically equal to zero:

$$\begin{aligned} c_1 p_1(x) + c_2 p_2(x) + c_3 p_3(x) \\ &= (c_1 + 5c_2 + c_3) \cdot 1 \\ &+ (-c_1 + 3c_2 + 3c_3) \cdot x \\ &+ (-2c_2 - c_3) \cdot x^2 \end{aligned}$$

Thus

$$\begin{aligned} c_1 + 5c_2 + c_3 &= 0 \\ -c_1 + 3c_2 + 3c_3 &= 0 \\ -2c_2 - c_3 &= 0 \end{aligned} \iff \begin{pmatrix} 1 & 5 & 1 \\ -1 & 3 & 3 \\ 0 & -2 & -1 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

The matrix is singular thus there is a nontrivial solution $(c_1, c_2, c_3) \implies$ the functions p_1, p_2, p_3 are linearly dependent

Wronskian and independence

Definition

Wronskian The Wronskian of functions $f_1, \ldots, f_n \in C^{n-1}[a, b]$ is the determinant

$$W(x) := \begin{vmatrix} f_1(x) & f_1'(x) & \ldots & f_1^{(n-1)}(x) \\ f_2(x) & f_2'(x) & \ldots & f_2^{(n-1)}(x) \\ \cdots & \cdots & \cdots & \cdots \\ f_n(x) & f_n'(x) & \ldots & f_n^{(n-1)}(x) \end{vmatrix}$$

Theorem

If functions $f_1, f_2, \ldots, f_n \in C^{n-1}[a, b]$ are linearly dependent, then their Wronskian is identically equal to zero.

Proof.

Assume e.g. that $f_1 = c_2 f_2 + \cdots + c_n f_n$. Then for every x, the first row of the matrix in the Wronskian is a linear combination of the other rows and thus $W(x) \equiv 0$. □

Independence

Corollary (Test for independence)

Assume that for functions $f_1, f_2, \ldots, f_n \in C^{n-1}[a, b]$ there is a point $x_0 \in [a, b]$ such that $W(x_0) \neq 0$. Then $f_1, f_2, \ldots, f_n$ are linearly independent.

Example

For the functions

$$p_1(x) = x(x-1), \quad p_2(x) = x(x-2), \quad p_3(x) = (x-1)(x-2)$$

the Wronskian at the point $x = 0$ is

$$W(0) = \begin{vmatrix} 0 & -1 & 2 \\ 0 & -2 & 2 \\ 2 & -3 & 2 \end{vmatrix} \neq 0$$

so they are linearly independent.