ECS231

Mathematics Review I: Linear Algebra

Reference: Chap.1 of Solomon

Vector spaces over $\mathbb{R}$

Denote a (abstract) vector by v. A vector space

$$\mathcal{V} = \{\text{a collection of vectors } v\}$$

which satisfies

- All $v, w \in \mathcal{V}$ can be *added* and *multiplied by* $\alpha \in \mathbb{R}$:

$$v + w \in \mathcal{V}, \quad \alpha \cdot v \in \mathcal{V}$$

- The operations '$+, \cdot$' must satisfy the *axioms*:

For arbitrary $u, v, w \in \mathcal{V}$,

1. '$+$' commutativity and associativity: $v + w = w + v$, $(u + v) + w = u + (v + w)$.
2. Distributivity: $\alpha(v + w) = \alpha v + \alpha w$, $(\alpha + \beta)v = \alpha v + \beta v$, for all $\alpha, \beta \in \mathbb{R}$.
3. '$+$' identity: there exists $0 \in \mathcal{V}$ with $0 + v = v$.
4. '$+$' inverse: for any $v \in \mathcal{V}$, there exists $w \in \mathcal{V}$ with $v + w = 0$.
5. '$\cdot$' identity: $1 \cdot v = v$.
6. '$\cdot$' compatibility: for all $\alpha, \beta \in \mathbb{R}$, $(\alpha\beta) \cdot v = \alpha \cdot (\beta \cdot v)$.

Example

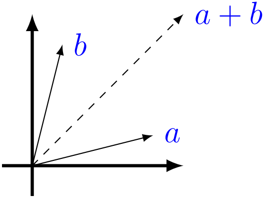
- Euclidean space:

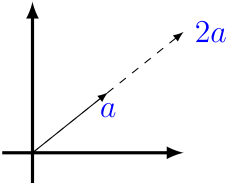
$$\mathbb{R}^n = \left\{ a \equiv (a_1, a_2, \ldots, a_n) \colon a_i \in \mathbb{R} \right\}.$$

  - Addition:
    $(a_1, \ldots, a_n) + (b_1, \ldots, b_n) = (a_1 + b_1, \ldots, a_n + b_n)$
  - Multiplication:

$$c \cdot (a_1, \ldots, a_n) = (ca_1, \ldots, ca_n)$$

- Illustration in $\mathbb{R}^2$:





Example

- Polynomials:

$$\mathbb{R}[x] = \left\{ p(x) = \sum_i a_i x^i \colon a_i \in \mathbb{R} \right\}.$$

- Addition and multiplication in the usual way, e.g. $p(x) = a_0 + a_1 x + a_2 x^2$, $q(x) = b_1 x$:
  - Addition:

$$p(x) + q(x) = a_0 + (a_1 + b_1)x + a_2 x^2.$$

  - Multiplication:

$$2p(x) = 2a_0 + 2a_1 x + 2a_2 x^2.$$

Span of vectors

- Start with $v_1, \ldots, v_n \in \mathcal{V}$, and $a_i \in \mathbb{R}$, we can define

$$v \equiv \sum_{i=1}^{n} a_i v_i = a_1 v_1 + a_2 v_2 + \cdots + a_n v_n,$$

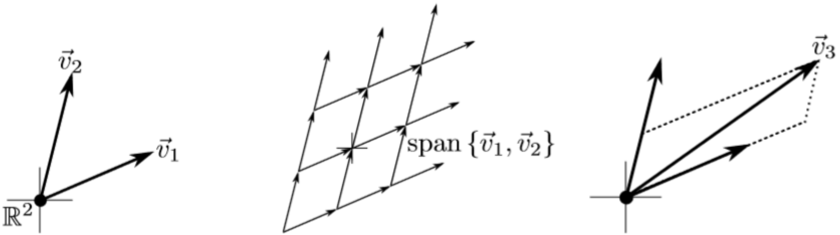
  Such a v is called a *linear combination* of $v_1, \ldots, v_n$.
- For a set of vectors

$$S = \{v_i \colon i \in \mathcal{I}\},$$

  all its linear combinations define

$$\operatorname{span} S \equiv \left\{ \sum_i a_i v_i \colon v_i \in S \text{ and } a_i \in \mathbb{R} \right\}$$

Example in $\mathbb{R}^2$



(a) $\vec{v}_1, \vec{v}_2 \in \mathbb{R}^2$ (b) span $\{\vec{v}_1, \vec{v}_2\}$ (c) span $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$

- Observation from (c): adding a new vector does not always increase the span.

Linear dependence

- A set S of vectors is *linearly dependent* if it contains a vector

$$v = \sum_{i=1}^{k} c_i v_i, \quad \text{for some } v_i \in S \backslash \{v\} \text{ and nonzero } c_i \in \mathbb{R}.$$

- Otherwise, S is called *linearly independent*.
- Two other equivalent defs. of linear dependence:
  - There exists $\{v_1, \ldots, v_k\} \subset S \backslash \{0\}$ such that

$$\sum_{i=1}^{k} c_i v_i = 0 \quad \text{where } c_i \neq 0 \text{ for all } i.$$

  - There exists $v \in S$ such that

$$\operatorname{span} S = \operatorname{span}(S \backslash \{v\}).$$

Dimension and basis

- Given a vector space $\mathcal{V}$, it is natural to build a finite set of linearly independent vectors:

$$\{v_1, \dots, v_n\} \subset \mathcal{V}.$$

- The max number n of such vectors defines the *dimension* of $\mathcal{V}$.
- Any set S of such vectors is a basis of $\mathcal{V}$, and satisfies

$$\operatorname{span} S = \mathcal{V}.$$

Examples

- The standard basis for $\mathbb{R}^n$ is given by the n vectors

$$e_i = (\underbrace{0, \ldots, 0}_{i-1}, 1, \underbrace{0, \ldots, 0}_{n-i}) \quad \text{for } i = 1, \ldots, n$$

  Since
  - e_i is not linear combination of the rest of vectors.
  - For all $c \in \mathbb{R}^n$, we have $c = \sum_{i=1}^n c_i e_i$.

  Hence, the dimension of $\mathbb{R}^n$ is n.

- A basis of polynomials $\mathbb{R}[x]$ is given by monomials

$$\{1, x, x^2, \ldots\}.$$

  The dimension of $\mathbb{R}[x]$ is ∞.

More about $\mathbb{R}^n$

- Dot product: for $a = (a_1, \ldots, a_n), b = (b_1, \ldots, b_n) \in \mathbb{R}^n$

$$a \cdot b = \sum_{i=1}^{n} a_i b_i.$$

- Length of a vector

$$\|a\|_2 = \sqrt{a_1^2 + \cdots + a_n^2} = \sqrt{a \cdot a}.$$

- Angle between two vectors

$$\theta = \arccos \frac{a \cdot b}{\|a\|_2 \|b\|_2}.$$

(*Motivating trigonometric in $\mathbb{R}^3$: $a \cdot b = \|a\|_2 \|b\|_2 \cos\theta$.)

- Vectors a, b are *orthogonal* if $a \cdot b = 0 = \cos 90°$.

Linear function

- Given two vector spaces $\mathcal{V}, \mathcal{V}'$, a function

$$\mathcal{L}\colon \mathcal{V} \to \mathcal{V}'$$

  is *linear*, if it preserves *linearity*.

- Namely, for all $v_1, v_2 \in \mathcal{V}$ and $c \in \mathbb{R}$,
  - $\mathcal{L}[v_1 + v_2] = \mathcal{L}[v_1] + \mathcal{L}[v_2]$.
  - $\mathcal{L}[cv_1] = c\mathcal{L}[v_1]$.
- $\mathcal{L}$ is completely defined by its action on a basis of $\mathcal{V}$:

$$\mathcal{L}[v] = \sum_i c_i \mathcal{L}[v_i],$$

  where $v = \sum_i c_i v_i$ and $\{v_1, v_2, \dots\}$ is a basis of $\mathcal{V}$.

Examples

- Linear map in $\mathbb{R}^n$:

$$\mathcal{L}\colon \mathbb{R}^2 \to \mathbb{R}^3$$

defined by

$$\mathcal{L}[(x, y)] = (3x, 2x + y, -y).$$

- Integration operator: linear map

$$\mathcal{L}\colon \mathbb{R}[x] \to \mathbb{R}[x]$$

defined by

$$\mathcal{L}[p(x)] = \int_0^1 p(x)dx.$$

Matrix

- Write vectors in $\mathbb{R}^m$ in '*column forms*', e.g.,

$$v_1 = \begin{bmatrix} v_{11} \\ \vdots \\ v_{m1} \end{bmatrix}, v_2 = \begin{bmatrix} v_{12} \\ \vdots \\ v_{m2} \end{bmatrix}, \dots, v_n = \begin{bmatrix} v_{1n} \\ \vdots \\ v_{mn} \end{bmatrix}.$$

- Put n columns together we obtain an $m \times n$ *matrix*

$$V \equiv \begin{bmatrix} | & | & & | \\ v_1 & v_2 & \dots & v_n \\ | & | & & | \end{bmatrix} = \begin{bmatrix} v_{11} & v_{12} & \dots & v_{1n} \\ v_{21} & v_{22} & \dots & v_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ v_{m1} & v_{m2} & \dots & v_{mn} \end{bmatrix}$$

- The space of all such matrices is denoted by $\mathbb{R}^{m \times n}$.

Unified notation: Scalars, Vectors, and Matrices

- A scalar $c \in \mathbb{R}$ is viewed as a 1×1 matrix

$$c \in \mathbb{R}^{1\times 1}.$$

- A column vector $v \in \mathbb{R}^n$ is viewed as an $n \times 1$ matrix

$$v \in \mathbb{R}^{n\times 1}.$$

Matrix vector multiplication

- A matrix $V \in \mathbb{R}^{m \times n}$ can be multiplied by a vector $c \in \mathbb{R}^n$:

$$\begin{bmatrix} | & | & & | \\ v_1 & v_2 & \dots & v_n \\ | & | & & | \end{bmatrix} \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} = c_1 v_1 + c_2 v_2 + \cdots + c_n v_n.$$

  Vc is a linear combination of the columns of V. This is fundamental.

Using matrix notation

- Matrix vector multiplication can be denoted by

$$\underbrace{A}_{\mathbb{R}^{m\times n}}\underbrace{x}_{\mathbb{R}^{n}} = \underbrace{b}_{\mathbb{R}^{m}}.$$

- $M \in \mathbb{R}^{m\times n}$ multiplied by another matrix in $\mathbb{R}^{n\times k}$ can be defined as

$$M[c_1, \ldots, c_k] \equiv [Mc_1, \ldots, Mc_k].$$

Example

- Identity matrix

$$I_n \equiv \begin{bmatrix} | & | & & | \\ e_1 & e_2 & \dots & e_n \\ | & | & & | \end{bmatrix} = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & 1 \end{bmatrix}.$$

It holds

$$I_n c = c \text{ for all } c \in \mathbb{R}^n.$$

Example

- Linear map $\mathcal{L}[(x,y)] = (3x, 2x+y, -y)$ satisfies

$$\mathcal{L}[(x,y)] = \underbrace{\begin{bmatrix} 3 & 0 \\ 2 & 1 \\ 0 & -1 \end{bmatrix}}_{\mathbb{R}^{3\times 2}} \cdot \underbrace{\begin{bmatrix} x \\ y \end{bmatrix}}_{\mathbb{R}^2} = \underbrace{\begin{bmatrix} 3x \\ 2x+y \\ -y \end{bmatrix}}_{\mathbb{R}^3}.$$

- All linear maps $\mathcal{L}\colon \mathbb{R}^n \to \mathbb{R}^m$ can be expressed as

$$\mathcal{L}[x] = Ax,$$

for some matrix $A \in \mathbb{R}^{m\times n}$.

Matrix transpose

- Use A_{ij} to denote the element of A at row i column j.
- The transpose of $A \in \mathbb{R}^{m \times n}$ is defined as $A^T \in \mathbb{R}^{n \times m}$

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

Example:

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} \qquad \Rightarrow \qquad A^T = \begin{bmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{bmatrix}.$$

- Basic identities:

$$(A^T)^T = A, \quad (A+B)^T = A^T + B^T, \quad (AB)^T = B^T A^T.$$

Examples: Matrix operations with transpose

- Dot product of $a, b \in \mathbb{R}^n$:

$$a \cdot b = \sum_{i=1}^{n} a_i b_i = \begin{bmatrix} a_1 & \dots & a_n \end{bmatrix} \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix} = a^T b.$$

- Residual norms of $r = Ax - b$:

$$\begin{aligned} \|Ax - b\|_2^2 &= (Ax - b)^T (Ax - b) \\ &= (x^T A^T - b^T)(Ax - b) \\ &= b^T b - b^T Ax - x^T A^T b + x^T A^T Ax \\ \text{(by } b^T Ax = x^T A^T b) \quad &= \|b\|_2^2 - 2b^T Ax + \|Ax\|_2^2. \end{aligned}$$

Computation aspects

- Storage of matrices in memory:

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} \Rightarrow \begin{cases} \text{Row-major:} & \boxed{1}\boxed{2}\boxed{3}\boxed{4}\boxed{5}\boxed{6} \\ \text{Column-major:} & \boxed{1}\boxed{3}\boxed{5}\boxed{2}\boxed{4}\boxed{6} \end{cases}$$

- Multiplication $b = Ax$ for $A \in \mathbb{R}^{m \times n}$ and $x \in \mathbb{R}^n$:

Access A row-by-row:

```
b = 0
for i = 1, ..., m do
    for j = 1, ..., n do
        b_i = b_i + A_ij x_j
    end for
end for
```

Access column-by-column:

```
b = 0
for j = 1, ..., n do
    for i = 1, ..., m do
        b_i = b_i + A_ij x_j
    end for
end for
```

Linear systems of equations in matrix form

- **Example**: find (x, y, z) satisfying

$$\begin{aligned} 3x + 2y + 5z &= 0 \\ -4x + 9y - 3z &= -7 \\ 2x - 3y - 3z &= 1. \end{aligned} \quad \Rightarrow \quad \begin{bmatrix} 3 & 2 & 5 \\ -4 & 9 & -3 \\ 2 & -3 & -3 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ -7 \\ 1 \end{bmatrix}$$

- Given $A = [a_1, \ldots, a_n] \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, find $x \in \mathbb{R}^n$:

$$Ax = b.$$

- Solution exists if b is in *column space* of A:

$$b \in \operatorname{col} A \equiv \{Ax \colon x \in \mathbb{R}^n\} = \left\{ \sum_{i=1}^{n} x_i a_i \colon x_i \in \mathbb{R} \right\}.$$

The dimension of $\operatorname{col} A$ is defined as the *rank* of A.

The square case

- Let $A \in \mathbb{R}^{n \times n}$ be a square matrix, and suppose $Ax = b$ has solution for all $b \in \mathbb{R}^n$. We can solve

$$Ax_i = e_i, \quad \text{for } i = 1, \ldots, n.$$
$$\Updownarrow$$
$$A \underbrace{\begin{bmatrix} x_1 & x_2 & \ldots & x_n \end{bmatrix}}_{A^{-1}} = I_n$$

- The *inverse* satisfies (*why*?)

$$AA^{-1} = A^{-1}A = I_n \quad \text{and} \quad (A^{-1})^{-1} = A.$$

- Hence, for any b, we can *express* the solution as

$$x = A^{-1}Ax = A^{-1}b.$$