Lecture 2 Vectors

Note by Samion Suwito on 1/23/25

Vectors and Functions

Vector subspace $\mathcal{X}$ over real field $\mathbb{R}$ is a set closed under vector addition/multiplication

$$x, y \in \mathcal{X}$$
$$\alpha, \beta \in \mathbb{R}$$
$$\alpha x + \beta y \in \mathcal{X}$$

Elements of a vector subspace are vectors (think of as lists of $\mathbb{R}$)

Examples

Example 1

$$\mathcal{X} = \mathbb{R}$$
$$x \in \mathcal{X}$$
$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

Example 2

$$\mathcal{X} \in C_b(\mathbb{R}) = \{f : \mathbb{R} \to \mathbb{R} : f \text{ is bounded and continuous}\}$$
$$x \in \mathcal{X}$$
$$x = t \in R$$

Example 3

$$\mathcal{X} \in \mathbb{P}_n(\mathbb{R}) = \text{Polynomials on } \mathbb{R} \text{ of degree at most} n$$
$$p \in \mathcal{X}$$
$$p(t) = p_0 + p_1 t + p_2 t^2 + \cdots + p_n t^n$$

Vectors have three interpretations

1. Points in space
2. Directions
3. Functions (Esp. Linear)

1. Points in Space

Consider a 2D plane and each point on the plane being represented as a vector (x_1, x_2).

For a collection $x^{(1)}, x^{(2)}, \ldots$ of vectors, a linear combination is a sum of the following form, specifically finite.

$$\sum_{i=1}^{k} \alpha_i x^{(i)}$$

for some $\alpha_i \in \mathbb{R}$

A subspace $\mathcal{V} \subset \mathcal{X}$ is a "vector space within $\mathcal{X}$"
If $u, v \in \mathcal{V}\ \alpha, \beta \in \mathbb{R} \Rightarrow \alpha u + \beta v \in \mathcal{V}$
Example: Subspace of continuous bounded functions would be continuous bounded functions with continuous bounded first derivatives

For a collection of vectors $S = \{x^{(1)}, \ldots, x^{(k)}\}$. Span of S is all LC of vectors S. Written as

$$Span(S) = \left\{ \sum_{i=1}^{k} \alpha_i x^{(i)} : \alpha_i \ldots \alpha_k \in \mathbb{R} \right\}$$

Sum of Subspaces

If $\mathcal{U}, \mathcal{V} \subset \mathcal{X}$ are subspaces then so is

$$\mathcal{U} + \mathcal{V} = \{u + v : u \in \mathcal{U}, v \in \mathcal{V}\}$$

To show it's closed

$$\alpha(u_1 + v_1) + \beta(u_2 + v_2) = (\alpha u_1 + \beta u_2) + (\alpha v_1 + \beta v_2)$$

The first part is $\in \mathcal{U} + \mathcal{V}$ then the second and third are respectively in $\in \mathcal{U}$ and $\in \mathcal{V}$.

If they are disjointed such that $\mathcal{U} \cap \mathcal{V} = \{0\}$ then their **direct sum** (often used in cases of orthogonal vectors) is denoted as $\mathcal{U} \oplus \mathcal{V} = \mathcal{U} + \mathcal{V}$

If $x \in \mathcal{U} \oplus \mathcal{V}$ then x can be **uniquely** written as

$$x = u + v, u \in \mathcal{U}, v \in \mathcal{V}$$

Proof by contradiction
suppose $x = u + v = u' + v'$
$= u - u' = v - v' = 0$ as $u - u'$ can be a vector in $\mathcal{U}$

Definition: Vectors $x^{(1)}, \dots, x^{(k)}$ are linearly independent if

$$\sum_{i=1}^{k} \alpha_i x^{(i)} = 0 \Rightarrow \alpha_i = 0 \forall i$$

If $\mathcal{V}$ Is a subspace with $Span(\{x^{(1)}, \dots, x^{(k)}\})$ and x^i is LI then the set is a basis for $\mathcal{V}$.

If $\{x^{(1)}, \dots, x^{(k)}\}$ and $\{y^{(1)}, \dots, y^{(d)}\}$ are basis for $\mathcal{V}$ then $k = d$ which is the dimension of $\mathcal{V}$. This is because assume $k < d$ you can write the vectors of y as the vectors of x and you'll find a contradiction.

If $\{x^{(1)}, \dots, x^{(k)}\}$ is a basis of $\mathcal{V}$ and $x \in \mathcal{V}$ has a unique representation

$$x = \sum \alpha_i x^{(i)}$$

which can be proved as the sum of the disjoint subspaces earlier

$$\mathcal{V} = Span(x^{(1)}) \oplus \cdots \oplus Span(x^{(k)}))$$

Important consequence: Any n-dimensional vector space $\mathcal{X}$ can be equivalently represented as $\mathbb{R}^n$

$$x = \sum \alpha_i x^{(i)} \mapsto \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} \in \mathbb{R}^n$$

Example: $\{1, t, \dots, t^n\}$ is a basis for $P_n(\mathbb{R})$

$$p(t) = p_0 + p_1 t + \cdots + p_n t^n \mapsto \begin{pmatrix} p_0 \\ \vdots \\ P_n \end{pmatrix} \in \mathbb{R}^{n+1}$$

2. Directions

For vectors, ”length” is generally measured by “norms”
Definition: A norm $||\cdot|| : \mathcal{X} \to \mathbb{R}$ is a function satisfying

1. Positive Definiteness $||x|| \geq 0 \; \forall x \in \mathcal{X}$ w/ equality iff $x = 0$
2. Triangle Inequality $||x + y|| \leq ||x|| + ||y||$
3. Positive Homogenity $||\alpha x|| = |a|||x||$
   Norm is a good objective function
   Ex: l^p norms:

$$||x||_p := \left(\sum_{i=1}^{n} |x_i|^p\right)^{\frac{1}{p}} \quad x \in \mathbb{R}^n, p \geq 1$$

   When $p = +\infty$

$$||x||_\infty = \max_{i=1...n} |x_i|$$

   Special cases $p = 2$ is Euclidean norm
   when $p = 1$

$$||x||_1 = \sum |x_i|$$

   It’s the manhattan/taxi cab/grid distance norm

Angle of Vectors

Definition: A (real) inner product space $\mathcal{X}$ is a vector space equipped w/ inner product $\langle \cdot, \cdot \rangle : \mathcal{X} \to \mathbb{R}$
Criteria for inner product:

1. $\langle x, x \rangle \geq 0$ w/ equality iff x=0
2. $< \alpha x + \beta y, z >= \alpha < x, z > + \beta < y, z >$
3. $\langle x, y \rangle = \langle y, x \rangle$

Inner products define a (hilbert) norm: $||x|| = \sqrt{\langle x, x \rangle}$
Example $\langle x, y \rangle = x^\top y$ then $||x||$ = Euclid norm.

$x^\top y = ||x||_2 ||y||_2 \cos(\theta)$ Regardless of dim
Cosine similarity used for RAG
Important Case $\theta = \pm 90^{\text{o}} \Rightarrow x^\top y = 0$
More generally, if $\mathcal{X}$ is an inner product space, x, y is orthogonal if $\langle x, y \rangle = 0$. $x \perp y$

Consequence **CBS**:
$\langle x, y\rangle \leq ||x||||y||$
Remark: if $\frac{1}{p}+\frac{1}{q}=1, p \geq 1$ then $x^{\top} y \leq\|x\|_p\|y\|_q$
Holder's inequality.

3. Vectors as a function

A function $f: \mathcal{X} \rightarrow \mathbb{R}$ is linear if
$f(\alpha x+\beta y)=\alpha f(x)+\beta f(y) \forall \alpha, \beta \in \mathbb{R}\ x, y \in \mathcal{X}$
(a vector space) Dual space of $\mathcal{X}$:

$$\mathcal{X}^*=\{f: \mathcal{X} \rightarrow \mathbb{R}, f \text { is linear}\}$$

$\mathcal{X}^* \equiv \mathbb{R}^n$
$f(x)=f(\Sigma x_i e_i)=\Sigma x_i f(e_i)$
Let $f_i=f(e_i)$

When $||\cdot||$ is a norm on $\mathcal{X}$, we can define "operator norm" on $\mathcal{X}^*$ via

$$||f||_{op}=\max_{x:||x|| \leq 1}|f(x)|$$

Fact:

$$\|\cdot\|_p^*=\|\cdot\|_q$$

where $\frac{1}{p}+\frac{1}{q}=1,\ p \geq 1$