Basic Calculus Review

CBMM Summer Course, Day 2 - Machine Learning

Vector Spaces

Functionals and Operators (Matrices)

Vector Space

- A **vector space** is a set V with binary operations

$$+ : V \times V \to V \quad \text{and} \quad \cdot : \mathbb{R} \times V \to V$$

such that for all $a, b \in \mathbb{R}$ and $v, w, x \in V$:

1. $v + w = w + v$
2. $(v + w) + x = v + (w + x)$
3. There exists $0 \in V$ such that $v + 0 = v$ for all $v \in V$
4. For every $v \in V$ there exists $-v \in V$ such that $v + (-v) = 0$
5. $a(bv) = (ab)v$
6. $1v = v$
7. $(a + b)v = av + bv$
8. $a(v + w) = av + aw$

Vector Space

- A **vector space** is a set V with binary operations

$$+ : V \times V \to V \quad \text{and} \quad \cdot : \mathbb{R} \times V \to V$$

such that for all $a, b \in \mathbb{R}$ and $v, w, x \in V$:

1. $v + w = w + v$
2. $(v + w) + x = v + (w + x)$
3. There exists $0 \in V$ such that $v + 0 = v$ for all $v \in V$
4. For every $v \in V$ there exists $-v \in V$ such that $v + (-v) = 0$
5. $a(bv) = (ab)v$
6. $1v = v$
7. $(a + b)v = av + bv$
8. $a(v + w) = av + aw$

- Example: $\mathbb{R}^n$, space of polynomials, space of functions.

Inner Product

- An **inner product** is a function $\langle \cdot, \cdot \rangle \colon V \times V \to \mathbb{R}$ such that for all $a, b \in \mathbb{R}$ and $v, w, x \in V$:

Inner Product

- An **inner product** is a function $\langle \cdot, \cdot \rangle \colon V \times V \to \mathbb{R}$ such that for all $a, b \in \mathbb{R}$ and $v, w, x \in V$:
  1. $\langle v, w \rangle = \langle w, v \rangle$
  2. $\langle av + bw, x \rangle = a\langle v, x \rangle + b\langle w, x \rangle$
  3. $\langle v, v \rangle \geqslant 0$ and $\langle v, v \rangle = 0$ if and only if $v = 0$.

Inner Product

- An **inner product** is a function $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{R}$ such that for all $a, b \in \mathbb{R}$ and $v, w, x \in V$:
  1. $\langle v, w \rangle = \langle w, v \rangle$
  2. $\langle av + bw, x \rangle = a\langle v, x \rangle + b\langle w, x \rangle$
  3. $\langle v, v \rangle \geqslant 0$ and $\langle v, v \rangle = 0$ if and only if $v = 0$.
- $v, w \in V$ are orthogonal if $\langle v, w \rangle = 0$.

Inner Product

- An **inner product** is a function $\langle \cdot, \cdot \rangle \colon V \times V \to \mathbb{R}$ such that for all $a, b \in \mathbb{R}$ and $v, w, x \in V$:
  1. $\langle v, w \rangle = \langle w, v \rangle$
  2. $\langle av + bw, x \rangle = a\langle v, x \rangle + b\langle w, x \rangle$
  3. $\langle v, v \rangle \geqslant 0$ and $\langle v, v \rangle = 0$ if and only if $v = 0$.
- $v, w \in V$ are orthogonal if $\langle v, w \rangle = 0$.
- Given $W \subseteq V$, we have $V = W \oplus W^{\perp}$, where $W^{\perp} = \{\, v \in V \mid \langle v, w \rangle = 0 \ \text{ for all } w \in W \,\}$.

Inner Product

- An **inner product** is a function $\langle \cdot, \cdot \rangle \colon V \times V \to \mathbb{R}$ such that for all $a, b \in \mathbb{R}$ and $v, w, x \in V$:
  1. $\langle v, w \rangle = \langle w, v \rangle$
  2. $\langle av + bw, x \rangle = a\langle v, x \rangle + b\langle w, x \rangle$
  3. $\langle v, v \rangle \geqslant 0$ and $\langle v, v \rangle = 0$ if and only if $v = 0$.
- $v, w \in V$ are orthogonal if $\langle v, w \rangle = 0$.
- Given $W \subseteq V$, we have $V = W \oplus W^{\perp}$, where $W^{\perp} = \{\, v \in V \mid \langle v, w \rangle = 0 \text{ for all } w \in W \,\}$.
- Cauchy-Schwarz inequality: $\langle v, w \rangle \leqslant \langle v, v \rangle^{1/2} \langle w, w \rangle^{1/2}$.

Norm

- Can define norm from inner product: $\|v\| = \langle v, v \rangle^{1/2}$.

Norm

- A **norm** is a function $\|\cdot\|\colon V \to \mathbb{R}$ such that for all $a \in \mathbb{R}$ and $v, w \in V$:
  1. $\|v\| \geqslant 0$, and $\|v\| = 0$ if and only if $v = 0$
  2. $\|av\| = |a|\,\|v\|$
  3. $\|v + w\| \leqslant \|v\| + \|w\|$
- Can define norm from inner product: $\|v\| = \langle v, v\rangle^{1/2}$.

Metric

- Can define metric from norm: $d(v, w) = \|v - w\|$.

Metric

- A **metric** is a function $\mathrm{d}\colon V \times V \to \mathbb{R}$ such that for all $v, w, x \in V$:
  1. $\mathrm{d}(v, w) \geqslant 0$, and $\mathrm{d}(v, w) = 0$ if and only if $v = w$
  2. $\mathrm{d}(v, w) = \mathrm{d}(w, v)$
  3. $\mathrm{d}(v, w) \leqslant \mathrm{d}(v, x) + \mathrm{d}(x, w)$
- Can define metric from norm: $\mathrm{d}(v, w) = \|v - w\|$.

Basis

- $B = \{v_1, \ldots, v_n\}$ is a **basis** of V if every $v \in V$ can be uniquely decomposed as

$$v = a_1 v_1 + \cdots + a_n v_n$$

for some $a_1, \ldots, a_n \in \mathbb{R}$.

Basis

- $B = \{v_1, \ldots, v_n\}$ is a **basis** of V if every $v \in V$ can be uniquely decomposed as

$$v = a_1 v_1 + \cdots + a_n v_n$$

  for some $a_1, \ldots, a_n \in \mathbb{R}$.

- An orthonormal basis is a basis that is orthogonal ($\langle v_i, v_j \rangle = 0$ for $i \neq j$) and normalized ($\|v_i\| = 1$).

Vector Spaces

Functionals and Operators (Matrices)

Maps

Next we are going to review basic properties of maps on a Hilbert space.

- functionals: $\Psi : \mathcal{H} \to \mathbb{R}$
- linear operators $A : \mathcal{H} \to \mathcal{H}$, such that $A(af + bg) = aAf + bAg$, with $a, b \in \mathbb{R}$ and $f, g \in \mathcal{H}$.

Representation of Continuous Functionals

Let $\mathcal{H}$ be a Hilbert space and $g \in \mathcal{H}$, then

$$\Psi_g(f) = \langle f, g \rangle, \qquad f \in \mathcal{H}$$

is a continuous linear functional.

Riesz representation theorem

The theorem states that every continuous linear functional Ψ can be written uniquely in the form,

$$\Psi(f) = \langle f, g \rangle$$

for some appropriate element $g \in \mathcal{H}$.

Matrix

- Every linear operator $L\colon \mathbb{R}^m \to \mathbb{R}^n$ can be represented by an $m \times n$ matrix A.

Matrix

- Every linear operator $L: \mathbb{R}^m \to \mathbb{R}^n$ can be represented by an $m \times n$ matrix A.
- If $A \in \mathbb{R}^{m \times n}$, the transpose of A is $A^\top \in \mathbb{R}^{n \times m}$ satisfying

$$\langle Ax, y \rangle_{\mathbb{R}^m} = (Ax)^\top y = x^\top A^\top y = \langle x, A^\top y \rangle_{\mathbb{R}^n}$$

  for every $x \in \mathbb{R}^n$ and $y \in \mathbb{R}^m$.

Matrix

- Every linear operator $L\colon \mathbb{R}^m \to \mathbb{R}^n$ can be represented by an $m \times n$ matrix A.
- If $A \in \mathbb{R}^{m \times n}$, the transpose of A is $A^\top \in \mathbb{R}^{n \times m}$ satisfying
$$\langle Ax, y\rangle_{\mathbb{R}^m} = (Ax)^\top y = x^\top A^\top y = \langle x, A^\top y\rangle_{\mathbb{R}^n}$$
for every $x \in \mathbb{R}^n$ and $y \in \mathbb{R}^m$.
- A is symmetric if $A^\top = A$.

Eigenvalues and Eigenvectors

- Let $A \in \mathbb{R}^{n \times n}$. A nonzero vector $v \in \mathbb{R}^n$ is an eigenvector of A with corresponding eigenvalue $\lambda \in \mathbb{R}$ if $Av = \lambda v$.

Eigenvalues and Eigenvectors

- Let $A \in \mathbb{R}^{n \times n}$. A nonzero vector $v \in \mathbb{R}^n$ is an eigenvector of A with corresponding eigenvalue $\lambda \in \mathbb{R}$ if $Av = \lambda v$.
- Symmetric matrices have real eigenvalues.

Eigenvalues and Eigenvectors

- Let $A \in \mathbb{R}^{n \times n}$. A nonzero vector $v \in \mathbb{R}^n$ is an eigenvector of A with corresponding eigenvalue $\lambda \in \mathbb{R}$ if $Av = \lambda v$.
- Symmetric matrices have real eigenvalues.
- **Spectral Theorem:** Let A be a symmetric $n \times n$ matrix. Then there is an orthonormal basis of $\mathbb{R}^n$ consisting of the eigenvectors of A.

Eigenvalues and Eigenvectors

- Let $A \in \mathbb{R}^{n \times n}$. A nonzero vector $v \in \mathbb{R}^n$ is an eigenvector of A with corresponding eigenvalue $\lambda \in \mathbb{R}$ if $Av = \lambda v$.
- Symmetric matrices have real eigenvalues.
- **Spectral Theorem:** Let A be a symmetric $n \times n$ matrix. Then there is an orthonormal basis of $\mathbb{R}^n$ consisting of the eigenvectors of A.
- Eigendecomposition: $A = V \Lambda V^\top$, or equivalently,

$$A = \sum_{i=1}^{n} \lambda_i v_i v_i^\top .$$

Singular Value Decomposition

- Every $A \in \mathbb{R}^{m \times n}$ can be written as
$$A = U\Sigma V^{\top},$$
where $U \in \mathbb{R}^{m \times m}$ is orthogonal, $\Sigma \in \mathbb{R}^{m \times n}$ is diagonal, and $V \in \mathbb{R}^{n \times n}$ is orthogonal.

Singular Value Decomposition

- Every $A \in \mathbb{R}^{m \times n}$ can be written as
$$A = U\Sigma V^\top,$$
where $U \in \mathbb{R}^{m \times m}$ is orthogonal, $\Sigma \in \mathbb{R}^{m \times n}$ is diagonal, and $V \in \mathbb{R}^{n \times n}$ is orthogonal.
- Singular system:

$$\begin{aligned} Av_i &= \sigma_i u_i & \qquad AA^\top u_i &= \sigma_i^2 u_i \\ A^\top u_i &= \sigma_i v_i & \qquad A^\top A v_i &= \sigma_i^2 v_i \end{aligned}$$

Matrix Norm

- The spectral norm of $A \in \mathbb{R}^{m \times n}$ is

$$\|A\|_{\text{spec}} = \sigma_{\max}(A) = \sqrt{\lambda_{\max}(AA^\top)} = \sqrt{\lambda_{\max}(A^\top A)}.$$

Matrix Norm

- The spectral norm of $A \in \mathbb{R}^{m \times n}$ is

$$\|A\|_{\text{spec}} = \sigma_{\max}(A) = \sqrt{\lambda_{\max}(AA^\top)} = \sqrt{\lambda_{\max}(A^\top A)}.$$

- The Frobenius norm of $A \in \mathbb{R}^{m \times n}$ is

$$\|A\|_F = \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}^2} = \sqrt{\sum_{i=1}^{\min\{m,n\}} \sigma_i^2}.$$

Positive Definite Matrix

A real symmetric matrix $A \in \mathbb{R}^{m \times m}$ is positive definite if

$$x^t A x > 0, \quad \forall x \in \mathbb{R}^m.$$

A positive definite matrix has positive eigenvalues.

Note: for positive semi-definite matrices $>$ is replaced by $\geqslant$.