ECE 302: Lecture 1.4 Linear Algebra (Optional)

Prof Stanley Chan

School of Electrical and Computer Engineering
Purdue University

Outline

- 1.1 Infinite Series
  - 1.1.1. Geometric Series
  - 1.1.2. Binomial Series
- 1.2 Approximations
  - 1.2.1. Taylor Approximation
  - 1.2.2. Exponential Series
  - 1.2.3. Logarithmic Approximation
- 1.3 Integration
  - 1.3.1. Odd and Even Functions
  - 1.3.2. Fundamental Theorem of Calculus
- 1.4 Linear Algebra (Optional)
  - 1.4.1. Inner Products (Optional)
  - 1.4.2. Matrix Calculus (Optional)
  - 1.4.3. Matrix Inversion (Optional)
- 1.5 Combinatorics
  - 1.5.1. Permutation
  - 1.5.2. Combination

Basic Notation

- Vector: $\boldsymbol{x} \in \mathbb{R}^n$
- Matrix: $\boldsymbol{A} \in \mathbb{R}^{m \times n}$; Entries are a_{ij} or $[\boldsymbol{A}]_{ij}$.
- Transpose:

$$\boldsymbol{A} = \begin{bmatrix} | & | & & | \\ \boldsymbol{a}_1 & \boldsymbol{a}_2 & \dots & \boldsymbol{a}_n \\ | & | & & | \end{bmatrix}, \quad \text{and} \quad \boldsymbol{A}^T = \begin{bmatrix} — & \boldsymbol{a}_1^T & — \\ — & \boldsymbol{a}_2^T & — \\ & \vdots & \\ — & \boldsymbol{a}_n^T & — \end{bmatrix}.$$

- Column: $\boldsymbol{a}_i$ is the i-th column of $\boldsymbol{A}$
- Identity matrix $\boldsymbol{I}$
- All-one vector $\boldsymbol{1}$ and all-zero vector $\boldsymbol{0}$
- Standard basis $\boldsymbol{e}_i$.

Inner Product

Definition

Let $\boldsymbol{x} = [x_1, x_2, \ldots, x_N]^T$ and $\boldsymbol{y} = [y_1, y_2, \ldots, y_N]^T$ be two vectors. The **inner product** $\boldsymbol{x}^T\boldsymbol{y}$ is

Example. Let $\boldsymbol{x} = [x_1, x_2]^T$. The inner product
$\boldsymbol{x}^T\boldsymbol{x} =$

Inner product

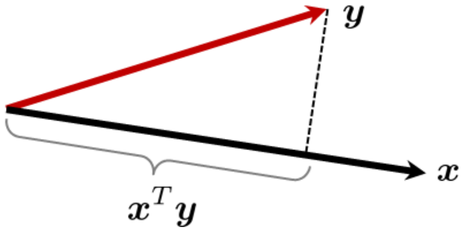

Figure: Geometric interpretation of inner product: We project one vector onto the other vector. The projected distance is the inner product.

Weighted Inner Product

Example.

Let $\boldsymbol{x} = [x_1, x_2]^T$, $\boldsymbol{\mu} = [\mu_1, \mu_2]$ and $\boldsymbol{C} = \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}$. The product $(\boldsymbol{x} - \boldsymbol{\mu})^T \boldsymbol{C} (\boldsymbol{x} - \boldsymbol{\mu})$ is

The ℓ_2-norm

Also called the **Euclidean norm**:

Definition

$$\|\boldsymbol{x}\|_2 = \sqrt{\sum_{i=1}^{n} x_i^2}. \tag{1}$$

- The set $\Omega = \{\boldsymbol{x} \mid \|\boldsymbol{x}\|_2 \leq r\}$ defines a circle:

$$\Omega = \{\boldsymbol{x} \mid \|\boldsymbol{x}\|_2 \leq r\} = \{(x_1, x_2) \mid x_1^2 + x_2^2 \leq r^2\}.$$

- $f(\boldsymbol{x}) = \|\boldsymbol{x}\|_2$ is not the same as $f(\boldsymbol{x}) = \|\boldsymbol{x}\|_2^2$.
- Triangle inequality holds:

$$\|\boldsymbol{x} + \boldsymbol{y}\|_2 \leq \|\boldsymbol{x}\|_2 + \|\boldsymbol{y}\|_2.$$

Outline

- 1.1 Infinite Series
  - 1.1.1. Geometric Series
  - 1.1.2. Binomial Series
- 1.2 Approximations
  - 1.2.1. Taylor Approximation
  - 1.2.2. Exponential Series
  - 1.2.3. Logarithmic Approximation
- 1.3 Integration
  - 1.3.1. Odd and Even Functions
  - 1.3.2. Fundamental Theorem of Calculus
- 1.4 Linear Algebra (Optional)
  - 1.4.1. Inner Products (Optional)
  - 1.4.2. Matrix Calculus (Optional)
  - 1.4.3. Matrix Inversion (Optional)
- 1.5 Combinatorics
  - 1.5.1. Permutation
  - 1.5.2. Combination

Matrix Calculus

Definition

Let $f : \mathbb{R}^n \to \mathbb{R}$ be a scalar field. The gradient of f with respect to $\boldsymbol{x} \in \mathbb{R}^n$ is defined as

$$\nabla_{\boldsymbol{x}} f(\boldsymbol{x}) = \begin{bmatrix} \frac{\partial f(\boldsymbol{x})}{\partial x_1} \\ \vdots \\ \frac{\partial f(\boldsymbol{x})}{\partial x_n} \end{bmatrix}. \tag{2}$$

Example 1. $f(\boldsymbol{x}) = \boldsymbol{a}^T \boldsymbol{x}$. In this case, the gradient is

$$\nabla_{\boldsymbol{x}} \left(\boldsymbol{a}^T \boldsymbol{x} \right) = \begin{bmatrix} \frac{\partial f(\boldsymbol{x})}{\partial x_1} \\ \vdots \\ \frac{\partial f(\boldsymbol{x})}{\partial x_n} \end{bmatrix} = \begin{bmatrix} \frac{\partial}{\partial x_1} \sum_{j=1}^n a_j x_j \\ \vdots \\ \frac{\partial}{\partial x_n} \sum_{j=1}^n a_j x_j \end{bmatrix} = \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} = \boldsymbol{a}. \tag{3}$$

More Examples

Example 2. $f(\boldsymbol{x}) = \boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x}$. Then,

$$\nabla_{\boldsymbol{x}}\left(\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x}\right) = \begin{bmatrix} \frac{\partial f(\boldsymbol{x})}{\partial x_1} \\ \vdots \\ \frac{\partial f(\boldsymbol{x})}{\partial x_n} \end{bmatrix} = \begin{bmatrix} \frac{\partial}{\partial x_1} \sum_{i,j=1}^n a_{ij}x_ix_j \\ \vdots \\ \frac{\partial}{\partial x_n} \sum_{i,j=1}^n a_{ij}x_ix_j \end{bmatrix}$$

$$= \begin{bmatrix} \sum_{j=1}^n a_{1,j}x_j \\ \vdots \\ \sum_{j=1}^n a_{n,j}x_j \end{bmatrix} + \begin{bmatrix} \sum_{i=1}^n a_{i,1}x_i \\ \vdots \\ \sum_{i=1}^n a_{i,n}x_i \end{bmatrix} = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{A}^T\boldsymbol{x}$$

If $\boldsymbol{A}$ is symmetric so that $\boldsymbol{A} = \boldsymbol{A}^T$ then $\nabla_{\boldsymbol{x}} f(\boldsymbol{x}) = 2\boldsymbol{A}\boldsymbol{x}$

More Examples

Example 3. $f(\boldsymbol{x}) = \|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{y}\|^2$. The gradient is

$$\begin{aligned}
\nabla_{\boldsymbol{x}}\left(\|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{y}\|^2\right) &= \nabla_{\boldsymbol{x}}\left(\boldsymbol{x}^T\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{x} - 2\boldsymbol{y}^T\boldsymbol{A}\boldsymbol{x} + \boldsymbol{y}^T\boldsymbol{y}\right) \\
&= \nabla_{\boldsymbol{x}}\left(\boldsymbol{x}^T\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{x}\right) - 2\nabla_{\boldsymbol{x}}\left(\boldsymbol{y}^T\boldsymbol{A}\boldsymbol{x}\right) + \nabla_{\boldsymbol{x}}\left(\boldsymbol{y}^T\boldsymbol{y}\right) \\
&= 2\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{x} - 2\boldsymbol{A}^T\boldsymbol{y} + 0 = 2\boldsymbol{A}^T(\boldsymbol{A}\boldsymbol{x} - \boldsymbol{y}).
\end{aligned}$$

Definition

The Hessian of f with respect to $\boldsymbol{x} \in \mathbb{R}^n$ is defined as

$$\nabla^2_{\boldsymbol{x}} f(\boldsymbol{x}) = \begin{bmatrix} \frac{\partial^2 f(\boldsymbol{x})}{\partial x_1^2} & \cdots & \frac{\partial^2 f(\boldsymbol{x})}{\partial x_1 \partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f(\boldsymbol{x})}{\partial x_n \partial x_1} & \cdots & \frac{\partial^2 f(\boldsymbol{x})}{\partial x_n^2} \end{bmatrix}. \tag{4}$$

Outline

- 1.1 Infinite Series
  - 1.1.1. Geometric Series
  - 1.1.2. Binomial Series
- 1.2 Approximations
  - 1.2.1. Taylor Approximation
  - 1.2.2. Exponential Series
  - 1.2.3. Logarithmic Approximation
- 1.3 Integration
  - 1.3.1. Odd and Even Functions
  - 1.3.2. Fundamental Theorem of Calculus
- 1.4 Linear Algebra (Optional)
  - 1.4.1. Inner Products (Optional)
  - 1.4.2. Matrix Calculus (Optional)
  - 1.4.3. Matrix Inversion (Optional)
- 1.5 Combinatorics
  - 1.5.1. Permutation
  - 1.5.2. Combination

Determinant and Inverse

Definition

Determinant Let $\boldsymbol{\Sigma} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, the determinant of $\boldsymbol{\Sigma}$ is

Definition (Inverse)

Let $\boldsymbol{\Sigma} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, the inverse of $\boldsymbol{\Sigma}$ is

Questions?