EE2211 Pre-Tutorial 4

Dr Feng LIN
feng_lin@nus.edu.sg

Agenda

- Recap
- Self-learning
- Tutorial 4

Today's Attendance

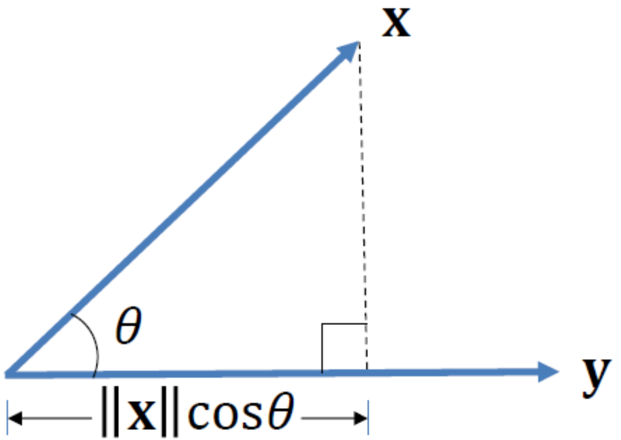
Recap

- Operations on Vectors and Matrices
  - Dot-product, matrix inverse
- System of Linear Equations
  - Matrix-vector notation, linear dependency, invertible
  - Even-, over-, under-determined linear systems
- Set and Functions
  - Inner product function
  - Linear and affine functions

Operations on Vectors and Matrices

Dot Product or Inner Product of Vectors:

$$\mathbf{x} \cdot \mathbf{y} = \mathbf{x}T\ \mathbf{y}$$
$$= [x_1 \;\; x_2] \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$
$$= x_1\, y_1 + x_2 y_2$$



Geometric definition:

$$\mathbf{x} \cdot \mathbf{y} = \|\mathbf{x}\|\ \|\mathbf{y}\| \cos\theta$$

Where θ is the angle between x and y , and $\|\mathrm{x}\| = \sqrt{\mathrm{x} \cdot \mathrm{x}}$ is the Euclidean length of vector x

Operations on Vectors and Matrices

Matrix inversion computation

$$A^{-1} = \frac{1}{\det(A)} \operatorname{adj}(A)$$

- $\det(A)$ is the determinant of A
- $\operatorname{adj}(A)$ is the adjugate or adjoint of A

Determinant computation

Example: 2x2 matrix

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

$$\det(A) = |A| = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$$

Systems of Linear Equations

Even-, over-, under-determined linear systems

$X \in R^{m} \times d$
e.g. Samples of features

$$Xw = y$$

y: Target values

w: Weightings

X is Square	Even-determined	$m = d$	One unique solution in general	$\hat{w} = X^{-1}y$
X is Tall	Over-determined	$m > d$	No exact solution in general; An approximated solution	$\hat{w} = (X^T X)^{-1} X^T y$ Left-inverse
X is Wide	Under-determined	$m < d$	Infinite number of solutions in general; Unique constrained solution	$\hat{w} = X^T (X X^T)^{-1} y$ Right-inverse

Systems of Linear Equations

$$Xw = y$$

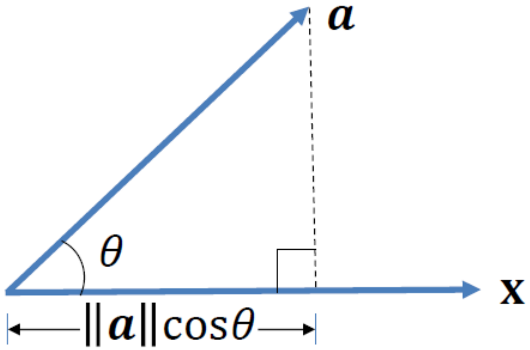
- X is Square
  - X is invertible → Unique solution
  - X is NOT invertible → No solution, or <u>infinity many solutions*</u> (Tut 4, Q2)
- X is Tall
  - $(X^T X)$ is invertible → Left-inverse → An approximate solution
  - $(X^T X)$ is NOT invertible → No solution, or <u>infinity many solutions*</u>
- X is Wide
  - (XX^T) is invertible → Right-inverse → A constrained solution
  - (XX^T) is NOT invertible → No solution (Lec 4, Example 4)

* This particular case will not be covered in this course.

Functions

The inner production function

- Suppose $\boldsymbol{a}$ is a *d*-vector. We can define a scalar valued function f of *d*-vectors, given by
$$f(\mathbf{x}) = \boldsymbol{a}^T\mathbf{x} = a_1x_1 + a_2x_2 + \cdots a_dx_d \qquad (1)$$
for any *d*-vector $\mathbf{x}$
- The inner product of its *d*-vector argument $\mathbf{x}$ with some (fixed) *d*-vector $\boldsymbol{a}$
- We can also think of f as forming a **weighted sum** of the elements of $\mathbf{x}$; the elements of $\boldsymbol{a}$ give the weights



Functions

Linear Functions

Superposition and linearity

- The inner product function $f(\mathbf{x}) = \boldsymbol{a}^T\mathbf{x}$ defined in equation (1) (slide 9) satisfies the property

$$\begin{aligned} f(\alpha\mathbf{x} + \beta\mathbf{y}) &= \boldsymbol{a}^T(\alpha\mathbf{x} + \beta\mathbf{y}) \\ &= \boldsymbol{a}^T(\alpha\mathbf{x}) + \boldsymbol{a}^T(\beta\mathbf{y}) \\ &= \alpha(\boldsymbol{a}^T\mathbf{x}) + \beta(\boldsymbol{a}^T\mathbf{y}) \\ &= \alpha f(\mathbf{x}) + \beta f(\mathbf{y}) \end{aligned}$$

  for all d-vectors $\mathbf{x}$, $\mathbf{y}$, and all scalars α, β.

- This property is called superposition, which consists of homogeneity and additivity
- A function that satisfies the superposition property is called linear

Functions

Linear and Affine Functions

A linear function plus a constant is called an affine function

A linear function $f: \mathcal{R}^d \rightarrow \mathcal{R}$ is **affine** if and only if it can be expressed as $f(\mathbf{x}) = \boldsymbol{a}^T\mathbf{x} + b$ for some d-vector $\boldsymbol{a}$ and scalar b, which is called the offset (or bias)

Example:

$$f(\mathbf{x}) = 2.3 - 2x_1 + 1.3x_2 - x_3$$

This function is affine, with $b = 2.3$, $\boldsymbol{a}^T = [-2,\ 1.3,\ -1]$.

THANK YOU