EE2211 Pre-Tutorial 4

Dr Feng LIN
feng_lin@nus.edu.sg

© Copyright National University of Singapore. All Rights Reserved.

Agenda

- Recap
- Self-learning
- Tutorial 4

© Copyright National University of Singapore. All Rights Reserved.

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

© Copyright National University of Singapore. All Rights Reserved.

Recap on Notations, Vectors, Matrices

Scalar	Numerical value	15, -3.5
Variable	Take scalar values	*x* or *a*
Vector	An ordered list of scalar values	$\mathbf{x}$ or $\mathbf{a}$
	Attributes of a vector	$\mathbf{a} = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$
Matrix	A rectangular array of numbers arranged in rows and columns	$\mathbf{X} = \begin{bmatrix} 2 & 4 \\ 21 & -6 \end{bmatrix}$

Capital Sigma $\sum_{i=1}^{m} x_i = x_1 + x_2 + \ldots + x_{m-1} + x_m$

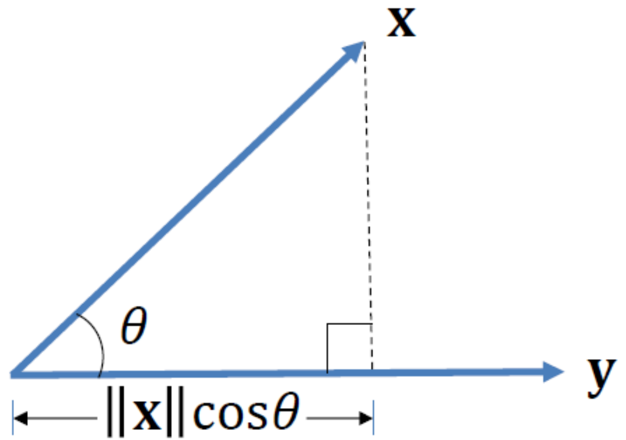
Capital Pi $\prod_{i=1}^{m} x_i = x_1 \cdot x_2 \cdot \ldots \cdot x_{m-1} \cdot x_m$

© Copyright National University of Singapore. All Rights Reserved.

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

© Copyright National University of Singapore. All Rights Reserved.

Operations on Vectors and Matrices

Matrix-Matrix Product

$$\mathbf{XW} = \begin{bmatrix} x_{1,1} & \cdots & x_{1,d} \\ \vdots & \ddots & \vdots \\ x_{m,1} & \cdots & x_{m,d} \end{bmatrix} \begin{bmatrix} w_{1,1} & \cdots & w_{1,h} \\ \vdots & \ddots & \vdots \\ w_{d,1} & \cdots & w_{d,h} \end{bmatrix}$$

$$= \begin{bmatrix} (x_{1,1}w_{1,1} + \cdots + x_{1,d}w_{d,1}) & \cdots & (x_{1,1}w_{1,h} + \cdots + x_{1,d}w_{d,h}) \\ \vdots & \ddots & \vdots \\ (x_{m,1}w_{1,1} + \cdots + x_{m,d}w_{d,1}) & \cdots & (x_{m,1}w_{1,h} + \cdots + x_{m,d}w_{d,h}) \end{bmatrix}$$

$$= \begin{bmatrix} \sum_{i=1}^{d} x_{1,i}w_{i,1} & \cdots & \sum_{i=1}^{d} x_{1,i}w_{i,h} \\ \vdots & \ddots & \vdots \\ \sum_{i=1}^{d} x_{m,i}w_{i,1} & \cdots & \sum_{i=1}^{d} x_{m,i}w_{i,h} \end{bmatrix}$$

If **X** is *m* x *d* and **W** is *d* x *h*, *then* the outcome is a *m* x *h* matrix

© Copyright National University of Singapore. All Rights Reserved.

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 $A, \operatorname{adj}(A) = C^T$
- C is the cofactor matrix

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

Minor of a_{12} is $M_{12} = \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix}$

- The (i, j) entry of the **cofactor matrix** $\mathbf{C}$ is the minor of (i, j) element times a sign factor

Cofactor $C_{ij} = (-1)^{i+j} M_{ij}$

- The **determinant** of $\mathbf{A}$ can also be defined by minors as

$$\det(\mathbf{A}) = \sum_{j=1}^{k} = a_{ij} C_{ij} = (-1)^{i+j} a_{ij} M_{ij}$$

© Copyright National University o

Python Functions

```
import numpy as np
```

Declare a matrix X	np.array($[[x_1, x_2], [x_3, x_4]]$)
Inverse of X	np.linalg.inv(X)
Transpose of X	X.transpose() or X.T
Determinant of X	np.linalg.det(X)
Dot product (1D Arrays)	X.dot(y) or X@y
Matrix multiplication (2D Matrices)	X.dot(y) or X@y (@ operator is introduced in Python 3.5)

© Copyright National University of Singapore. All Rights Reserved.

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

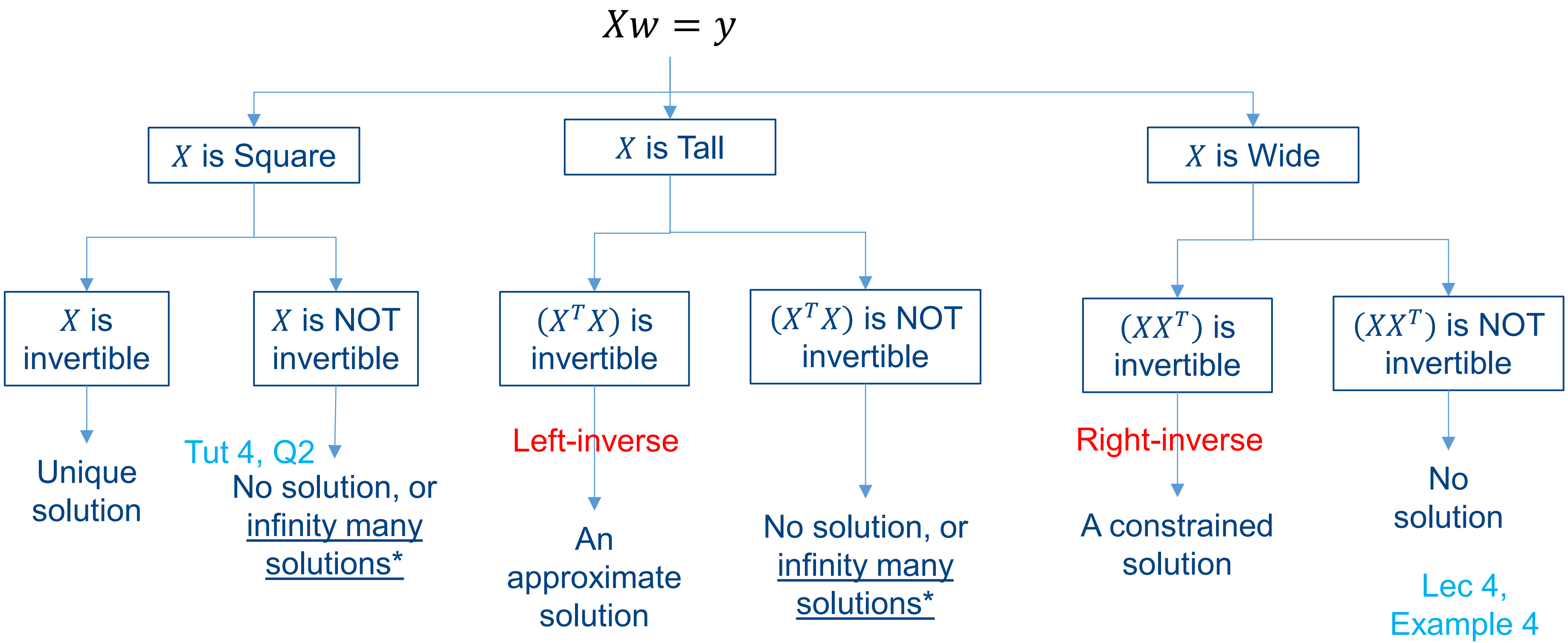
w: Weightings

$$\mathbf{X} = \begin{bmatrix} x_{1,1} & x_{1,2} & \cdots & x_{1,d} \\ \vdots & \vdots & \ddots & \vdots \\ x_{m,1} & x_{m,2} & \cdots & x_{m,d} \end{bmatrix}, \quad \mathbf{w} = \begin{bmatrix} w_1 \\ \vdots \\ w_d \end{bmatrix}, \quad \mathbf{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix}.$$

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	$\hat{w} = (X^TX)^{-1}X^Ty$ Left-inverse
X is Wide	Under-determined	$m < d$	Infinite number of solutions in general; Unique constrained solution	$\hat{w} = X^T(XX^T)^{-1}y$ Right-inverse

© Copyright National University of Singapore. All Rights Reserved.

Systems of Linear Equations



© Copyright National University of Singapore. All Rights Reserved.

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

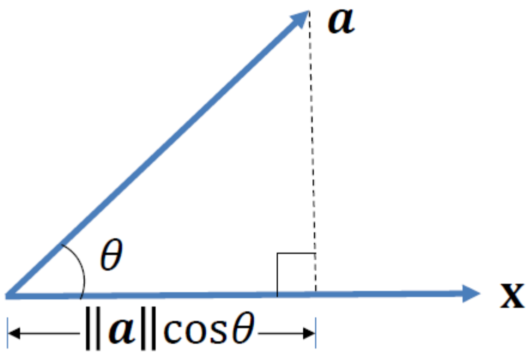
Functions

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



© Copyright National University of Singapore. All Rights Reserved.

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

© Copyright National University of Singapore. All Rights Reserved.

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 $b = 2.3$, $\boldsymbol{a}^T = [-2,\ 1.3,\ -1]$.

© Copyright National University of Singapore. All Rights Reserved.

Summary

- Operations on Vectors and Matrices
  - Dot-product, matrix inverse
- Systems of Linear Equations $\mathbf{X}\mathbf{w} = \mathbf{y}$
  - Matrix-vector notation, linear dependency, invertible
  - Even-, over-, under-determined linear systems
- Set and Functions

Assignment 1 (week 6 Fri)
Tutorial 4

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

- Scalar and vector functions
- Inner product function
- Linear and affine functions

python package *numpy*
Inverse: *numpy.linalg.inv(X)*
Transpose: *X.T*

© Copyright National University

THANK YOU

© Copyright National University of Singapore. All Rights Reserved.