

CS 760: Machine Learning - Fall 2020

Homework 2: Linear Regression

Due : 10/13/2020

Zijie Zhang

October 10, 2020

Problem 1

Proof. We notice that:

$$\begin{aligned}\|\mathbf{y} - \mathbf{X}\boldsymbol{\theta}\|_2^2 &= (\mathbf{y} - \mathbf{X}\boldsymbol{\theta})^T (\mathbf{y} - \mathbf{X}\boldsymbol{\theta}) \\ &= \mathbf{y}^T \mathbf{y} - 2\boldsymbol{\theta}^T \mathbf{X}^T \mathbf{y} + \boldsymbol{\theta}^T \mathbf{X}^T \mathbf{X} \boldsymbol{\theta}\end{aligned}$$

Compute the differential:

$$\begin{aligned}d \operatorname{tr} (\mathbf{y}^T \mathbf{y} - 2\boldsymbol{\theta}^T \mathbf{X}^T \mathbf{y} + \boldsymbol{\theta}^T \mathbf{X}^T \mathbf{X} \boldsymbol{\theta}) &= -2 \operatorname{tr} ((d\boldsymbol{\theta})^T \mathbf{X}^T \mathbf{y}) + \operatorname{tr} ((d\boldsymbol{\theta})^T \mathbf{X}^T \mathbf{X} \boldsymbol{\theta}) + \operatorname{tr} (\boldsymbol{\theta}^T \mathbf{X}^T \mathbf{X} (d\boldsymbol{\theta})) \\ &= 2 \operatorname{tr} ((d\boldsymbol{\theta})^T (\mathbf{X}^T \mathbf{X} \boldsymbol{\theta} - \mathbf{X}^T \mathbf{y}))\end{aligned}$$

It follows that

$$\frac{d}{d\boldsymbol{\theta}} = 2\mathbf{X}^T \mathbf{X} \boldsymbol{\theta} - 2\mathbf{X}^T \mathbf{y}$$

Let this derivative to zero, we have

$$\arg \min_{\boldsymbol{\theta}} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y}$$

□

Problem 2

Proof.

□

Problem 3

Proof.

□

Problem 4

Proof.

□

Problem 5

Proof.

□

Problem 6

Proof.

□

Problem 7

Proof.

□