Sunghyun Kim
5362318
kimx5261@umn.edu

Homework 0
CSci 5525 Spring 23

2023-01-25

1 Have you read through the class syllabus, noted the important dates, and the class policies?

Yes

2

2.1 Background in ML

I took none of the courses listed. STAT 4052 - Intro. to Statistical Learning is the closest course I took.

2.2 Courses on Probability/Statistics

STAT 3032, STAT 3701, STAT 4051, STAT 4052, MATH 5651, STAT 5102, STAT 5511, STAT 5601

2.3 Courses on Linear Algebra

MATH 2243, MATH 4242

2.4 Courses on Optimization

I have not taken any courses on optimization.

3 $w^* = \arg\min_{w \in \mathbb{R}^P} \frac{1}{2}||y - Xw||^2 + \frac{\lambda}{2}||w||^2$ into closed form.

Solution:

$$\begin{aligned} w^* &= \arg\min_{w \in \mathbb{R}^P} \frac{1}{2}||y - Xw||^2 + \frac{\lambda}{2}||w||^2 \\ \frac{\partial w^*}{\partial w} &= -\frac{1}{2} * 2X^T(y - Xw) + \frac{1}{2} * 2\lambda w \\ &= -X^T(y - Xw) + \lambda w \end{aligned}$$

The closed expression is derived by

$$\begin{aligned} -X^T(y - Xw) + \lambda w &= 0 \\ X^T Xw + \lambda w &= X^T y \\ w(X^T X + \lambda I) &= X^T y \\ w &= (X^T X + \lambda I)^{-1} X^T y \end{aligned}$$

It is valid even when $n < p$. When $\lambda > 0$, the sample covariance matrix makes all of its eigenvalues to be greater than 0. Thus, it makes the sample covariance matrix to be invertible, and that invertibility produces a unique solution when $n < p$.

4 Python code for Q3

Consult the code submitted.

5 Solution to $\max_{w \in \mathbb{R}^n : w^T w = 1} w^T A w$ and $\min_{w \in \mathbb{R}^n : w^T w = 1} w^T A w$

It is the largest eigenvalue of matrix A. A could also be called as the Spectral Radius of A as A is a positive definite.