HW 0

Arnab Dey
Student ID: 5563169
Email: dey00011@umn.edu

Solution 1

Yes.

Solution 2.i

I have taken CSCI 5521: Introduction to Machine Learning

Solution 2.ii

During my masters, I took 'Statistical Signal Processing' at Indian Institute of Technology, Kharagpur, India. At UMN, I am taking 'EE 5531: Probability and Stochastic Process' in this fall 2020 semester.

Solution 2.iii

In the fall 2019 semester, I took 'EE 5231: Linear Systems and Optimal Control', which covered Linear Algebra.

Solution 2.iv

I have not taken any specific course on optimization.

Solution 3

The objective is to

$$\min_{w \in \mathbb{R}^p} \frac{1}{2}\|y - Xw\|^2 + \frac{c}{2}\|w\|^2,$$

where $X \in \mathbb{R}^{n \times p}, y \in \mathbb{R}^n$ and $c > 0$ is a constant.

The cost function is

$$J(w) = \frac{1}{2}(y - Xw)^T(y - Xw) + \frac{c}{2}w^T w.$$

Therefore,

$$\frac{\partial J(w)}{\partial w} = -X^T(y - Xw) + cw. \tag{1}$$

Equating (1) to 0, we obtain:

$$\begin{aligned} & -X^T y + X^T X w^* + c w^* = 0 \\ \implies & (X^T X + cI)w^* = X^T y \\ \implies & w^* = (X^T X + cI)^{-1} X^T y, \end{aligned}$$

where $I \in \mathbb{R}^{p \times p}$ is an identity matrix.

Solution 4

The solution to

$$\max_{w \in \mathbb{R}^n : w^T w = 1} w^T A w$$

is the largest eigen-value of A and the solution to

$$\min_{w \in \mathbb{R}^n : w^T w = 1} w^T A w$$

is the smallest eigen-value of A.

Solution 5

The probability density function $p(x; \mu, \Sigma), x \in \mathbb{R}^d$ of a multivariate Gaussian distribution with mean $\mu \in \mathbb{R}^d$ and covariance $\Sigma \in \mathbb{R}^{d \times d}$ is

$$p(x; \mu, \Sigma) = \frac{1}{(2\pi)^{\frac{d}{2}} |\Sigma|^{\frac{1}{2}}} e^{-\frac{1}{2}(x-\mu)^T \Sigma^{-1} (x-\mu)},$$

where $|\Sigma|$ denotes the determinant of Σ.

If the precision matrix is denoted by $\Theta = \Sigma^{-1}$, then

$$p(x; \mu, \Theta^{-1}) = \frac{1}{(2\pi)^{\frac{d}{2}}} |\Theta|^{\frac{1}{2}} e^{-\frac{1}{2}(x-\mu)^T \Theta (x-\mu)},$$

where $|\Theta|$ is the determinant of the precision matrix Θ.