Mid-Term Exam

[AIX7026] Advanced Machine Learning

Due: 04/16 (Fri), 23:59:59

1. 주어진 vector $\mathbf{x}$에 대한 함수 $f(\mathbf{x}) = \|\mathbf{x}\|^2$가 convex function임을 증명하세요. [10점]

2. Python의 *CVXOPT* library는 아래와 같은 quadratic program (QP)의 해를 구하는 library 입니다. 1–2주차 강의자료의 notation을 이용하여, hard SVM을 QP 형태로 만드세요. 즉, cost function을 위한 $\mathbf{P}, \mathbf{q}, r$과 constraint를 위한 $\mathbf{G}, \mathbf{A}, \mathbf{h}, \mathbf{b}$를 각각 정의하세요. [20점]

$$\text{minimize } \frac{1}{2}\mathbf{x}^T\mathbf{P}\mathbf{x} + \mathbf{q}^T\mathbf{x} + r \tag{1}$$

$$\text{subject to } \mathbf{G}\mathbf{x} \leq \mathbf{h} \tag{2}$$

$$\mathbf{A}\mathbf{x} = \mathbf{b} \tag{3}$$

3. Ridge regression에 대한 다음의 cost $J(\boldsymbol{\theta})$를 최소화하는 closed form solution $\hat{\boldsymbol{\theta}}$를 유도하세요. [12점]

$$J(\boldsymbol{\theta}) = \text{SSE}(\boldsymbol{\theta}) + \frac{\alpha}{2}\|\boldsymbol{\theta}\|^2 \tag{4}$$

4. 다음과 같은 multivariate Gaussian distribution을 따르는 d-dimensional random variable $\mathbf{x} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$을 두 개의 sub-dimension $\mathbf{x}_A = [\ x_1 \ \ \cdots \ \ x_r\ ]^T \in \mathbb{R}^r$과 $\mathbf{x}_B = [\ x_{r+1} \ \ \cdots \ \ x_d\ ]^T \in \mathbb{R}^{d-r}$로 partitioning하고, 그에 맞게 $\boldsymbol{\mu}$와 $\boldsymbol{\Sigma}$도 아래와 같이 partitioning 했습니다.

$$\mathbf{x} = \begin{bmatrix} \mathbf{x}_A \\ \mathbf{x}_B \end{bmatrix} \qquad \boldsymbol{\mu} = \begin{bmatrix} \boldsymbol{\mu}_A \\ \boldsymbol{\mu}_B \end{bmatrix} \qquad \boldsymbol{\Sigma} = \begin{bmatrix} \boldsymbol{\Sigma}_{AA} & \boldsymbol{\Sigma}_{AB} \\ \boldsymbol{\Sigma}_{BA} & \boldsymbol{\Sigma}_{BB} \end{bmatrix}. \tag{5}$$

   (a) 서로에 대한 conditional distribution도 Gaussian distribution을 증명하세요. [10점]

   (b) Conditional distribution의 parameter가 다음과 같이 정리되는 것을 보이세요. [15점]

$$\mathbf{x}_A \mid \mathbf{x}_B \sim \mathcal{N}(\boldsymbol{\mu}_A + \boldsymbol{\Sigma}_{AB}\boldsymbol{\Sigma}_{BB}^{-1}(\mathbf{x}_B - \boldsymbol{\mu}_B), \boldsymbol{\Sigma}_{AA} - \boldsymbol{\Sigma}_{AB}\boldsymbol{\Sigma}_{BB}^{-1}\boldsymbol{\Sigma}_{BA}) \tag{6}$$

$$\mathbf{x}_B \mid \mathbf{x}_A \sim \mathcal{N}(\boldsymbol{\mu}_B + \boldsymbol{\Sigma}_{BA}\boldsymbol{\Sigma}_{AA}^{-1}(\mathbf{x}_A - \boldsymbol{\mu}_A), \boldsymbol{\Sigma}_{BB} - \boldsymbol{\Sigma}_{BA}\boldsymbol{\Sigma}_{AA}^{-1}\boldsymbol{\Sigma}_{AB}) \tag{7}$$

5. 다음의 transition probability에 대한 stationary probability를 구하세요. [8점]

$$\mathbf{T} = \begin{bmatrix} 0.5 & 0.4 & 0.1 \\ 0.2 & 0.5 & 0.3 \\ 0.1 & 0.3 & 0.6 \end{bmatrix} \tag{8}$$

6. (객관식) 다음은 확률과 관련된 세 종류의 수렴 (convergence) 에 대한 정의입니다. 아래의 질문에 답하세요.

- **Convergence in distribution:** A sequence of random variables $X_1, X_2, ..., X_n$ converges to X in distribution, if

$$\lim_{n\to\infty} F_{X_n}(t) = F_X(t) \tag{9}$$

for all points t where the C.D.F F_X is continuous.

- **Convergence in probability**: A sequence of random variables $X_1, X_2, ..., X_n$ converges in probability to a random variable X, if for every $\epsilon > 0$, we have that

$$\lim_{n\to\infty} \mathrm{P}\left(|X_n - X| \geq \epsilon\right) = 0. \tag{10}$$

- **Almost sure convergence**: A sequence of random variables $X_1, X_2, ..., X_n$ converges to X almost surely, i.e., $X_n \overset{\text{a.s.}}{\to} X$, if

$$\mathrm{P}\left(\lim_{n\to\infty} X_n = X\right) = 1. \tag{11}$$

(a) 서로 간의 포함관계는 어떻게 되는지 답하세요. [5점]

(b) Central limit theorem은 어떤 종류의 수렴에 해당하는지 고르세요. [10점]

(c) Metropolis-Hastings method는 어떤 종류의 수렴에 해당하는지 고르세요. [10점]