IT 585 - End Semester Exam

End Semester Exam

April 27, 2024
Duration: 120 minutes
Maximum Marks: 30

This is an open notes exam. You can use your handwritten notes or printed copy of handwritten notes but no typed material, textbooks etc. In case you have any doubt, make an appropriate assumption, state the assumption clearly, and proceed. Notations are the ones used in class unless and otherwise specified. Proofs should be complete. Answer to the point (The paper is not lengthy if you stick to what is asked). Verbose answers without relevant content will be penalized.

1. Let $\mathcal{H} = \{h_1, h_2, \ldots, h_M\}$ with some finite M be a set of hypotheses. Prove that the VC-dimension of $\mathcal{H}$ is upper bounded by $\log_2 M$. [5 Marks]

2. In Fisher's Linear Discriminant analysis, show that maximization of the class separation criterion given by equation

$$\tilde{m}^+ - \tilde{m}^- = \mathbf{w}^T(\mathbf{m}^+ - \mathbf{m}^-)$$

with respect to $\mathbf{w}$, using a Lagrange multiplier to enforce the constraint $\mathbf{w}^T\mathbf{w} = 1$, leads to the result that $\mathbf{w} \propto (\mathbf{m}^+ - \mathbf{m}^-)$. Here $\mathbf{m}^+$ and $\mathbf{m}^-$ are the mean vectors of the points belonging to the positive and negative class respectively. Assume that the points and $\mathbf{w}$ belong to $\mathbb{R}^d$ [7 Marks]

3. Consider a mixture distribution of the form

$$p(\mathbf{x}) = \sum_{k=1}^{K} \pi_k p(\mathbf{x}|k)$$

where the elements of $\mathbf{x}$ could be discrete or continuous or a combination of these. Denote the mean and covariance of $p(\mathbf{x}|k)$ by μ_k and Σ_k, respectively. Show that the mean and covariance of the mixture distribution are given by the following equations

$$\mathbb{E}[\mathbf{x}] = \sum_{k=1}^{K} \pi_k \mu_\mathrm{k}$$

$$\text{cov}[\mathbf{x}] = \sum_{k=1}^{K} \pi_k \{\boldsymbol{\Sigma}_\mathbf{k} + \boldsymbol{\mu}_\mathbf{k}\boldsymbol{\mu}_\mathbf{k}{}^T\} - \mathbb{E}[\mathbf{x}]\mathbb{E}[\mathbf{x}]^T$$

Hint: Use the definitions and properties of expectation and covariance and their basic properties.

[3 + 4 Marks]

4. (a) For a non-degenerate quadratic function of the form

$$f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^\mathbf{T}\mathbf{M}\mathbf{x} - \mathbf{q}^\mathbf{T}\mathbf{x} + c$$

where $\mathbf{M} \in \mathbb{R}^{d\times d}$ is an invertible and symmetric matrix, $\mathbf{q} \in \mathbb{R}^d$, and $c \in \mathbb{R}$, let $\mathbf{x}^* = \mathbf{M}^{-1}\mathbf{q}$ be the unique solution of $\nabla f(\mathbf{x}) = 0$. Prove that in this case, the Newton method converges to an optimal solution in a single iteration starting from any initial point. [5 Marks]

(b) For an L-smooth and μ-strongly convex, differentiable function $f(\mathbf{x}) : \mathbb{R}^d \to \mathbb{R}$ and with $\mathbf{x}^*$ as the point of minimum, prove the following properties:

i.

$$\frac{\mu}{2}||\mathbf{x} - \mathbf{x}^*||^2 \leq f(\mathbf{x}) - f(\mathbf{x}^*) \leq \frac{L}{2}||\mathbf{x} - \mathbf{x}^*||^2$$

[1 Mark]

ii.

$$\frac{1}{2L}||\nabla f(\mathbf{x})||^2 \leq f(\mathbf{x}) - f(\mathbf{x}^*) \leq \frac{1}{2\mu}||\nabla f(\mathbf{x})||^2$$

[3 Marks]

iii.

$$\frac{1}{L^2}||\nabla f(\mathbf{x})||^2 \leq ||\mathbf{x} - \mathbf{x}^*||^2 \leq \frac{1}{\mu^2}||\nabla f(\mathbf{x})||^2$$

[2 Marks]

Hint: Use the definitions with appropriate vectors. You can state and use results proved in class directly.