
Para la realización del examen se permite un único libro (sea cual sea, salvo el texto **PATTERN RECOGNITION AND MACHINE LEARNING - SOLUTIONS TO EXERCISES - TUTORS' EDITION** de M. Svensén y C.M. Bishop) sin anotaciones de ningún tipo. Si considera que el enunciado contiene algún erratum, hágalo constar en su respuesta y, **de confirmarse**, el equipo docente compensará el tiempo invertido por el estudiante. En los enunciados, asuma el empleo de la notación habitual del texto base.

1. Considere la distribución de probabilidad conjunta $p(\mathbf{x}, \mathbf{y})$ definida por las distribuciones marginal y condicional siguientes:

$$\begin{aligned} p(\mathbf{x}) &= \mathcal{N}(\mathbf{x} | \boldsymbol{\mu}, \boldsymbol{\Lambda}^{-1}) \\ p(\mathbf{y} | \mathbf{x}) &= \mathcal{N}(\mathbf{y} | \mathbf{A}\mathbf{x} + \mathbf{b}, \mathbf{L}^{-1}) \end{aligned}$$

donde $\boldsymbol{\mu}$ y $\mathbf{A}\mathbf{x} + \mathbf{b}$ son las medias de la distribución gaussiana, y $\boldsymbol{\Lambda}$ y \mathbf{L} son las matrices de precisión. Utilice la técnica de completar el cuadrado para obtener una expresión analítica de la media y covarianza de $p(\mathbf{x} | \mathbf{y})$.

2. Utilice la identidad matricial

$$(\mathbf{M} + \mathbf{v}\mathbf{v}^T)^{-1} = \mathbf{M}^{-1} - \frac{(\mathbf{M}^{-1}\mathbf{v})(\mathbf{v}^T\mathbf{M}^{-1})}{1 + \mathbf{v}^T\mathbf{M}^{-1}\mathbf{v}}$$

y suponga que

$$\mathbf{S}_{N+1}^{-1} = \mathbf{S}_N^{-1} + \beta \boldsymbol{\phi}_{N+1} \boldsymbol{\phi}_{N+1}^T \quad (1)$$

para demostrar que la incertidumbre de un modelo predictivo bayesiano construido a partir de N observaciones, y dada por

$$\sigma_N^2(\mathbf{x}) = \frac{1}{\beta} + \boldsymbol{\phi}(\mathbf{x})^T \mathbf{S}_N \boldsymbol{\phi}(\mathbf{x}) \quad (2)$$

(con la interpretación habitual de los símbolos utilizados en el texto base) satisface que $\sigma_{N+1}^2(\mathbf{x}) \leq \sigma_N^2(\mathbf{x})$.

3. Sabiendo que en un problema de clasificación con dos clases

$$p(\mathcal{C}_1 | \mathbf{x}) = \frac{1}{1 + \exp(-\ln \frac{p(\mathbf{x} | \mathcal{C}_1)p(\mathcal{C}_1)}{p(\mathbf{x} | \mathcal{C}_2)p(\mathcal{C}_2)})} = \frac{1}{1 + \exp(-a)} = \sigma(a) \quad (3)$$

y suponiendo un modelo generativo en el que las verosimilitudes (likelihoods) de las dos clases vienen dadas por dos gaussianas de medias $\boldsymbol{\mu}_1$ y $\boldsymbol{\mu}_2$, pero la misma varianza $\boldsymbol{\Sigma}$, demuestre que

$$p(C_1|\mathbf{x}) = \sigma(\mathbf{w}^T \mathbf{x} + w_0) \quad (4)$$

con $\mathbf{w} = \boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)$ y

$$w_0 = -\frac{1}{2}\boldsymbol{\mu}_1^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}_1 + \frac{1}{2}\boldsymbol{\mu}_2^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}_2 + \ln \frac{p(C_1)}{p(C_2)}. \quad (5)$$