18F2

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 $p(\boldsymbol{x}, \boldsymbol{y})$ definida por las distribuciones marginal y condicional siguientes:

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

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

2. Utilice la identidad matricial

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

y suponga que

$$\boldsymbol{S}_{N+1}^{-1} = \boldsymbol{S}_N^{-1} + \beta\boldsymbol{\phi}_{N+1}\boldsymbol{\phi}_{N+1}^T \tag{1}$$

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

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

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

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

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

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

$$p(\mathcal{C}_1|\boldsymbol{x}) = \sigma(\boldsymbol{w}^T\boldsymbol{x} + w_0) \tag{4}$$

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

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