15F2

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. Calcule la divergencia de Kullback-Leibler entre dos gaussianas $p(\boldsymbol{x}) = \mathcal{N}(\boldsymbol{x}|\boldsymbol{\mu}, \boldsymbol{\Sigma})$ y $q(\boldsymbol{x}) = \mathcal{N}(\boldsymbol{x}|\boldsymbol{m}, \boldsymbol{L})$.

2. Sea $t = \boldsymbol{w}^T \boldsymbol{\phi}(\boldsymbol{x}) + \epsilon$;
   t, un variable aleatoria que toma valores en $\mathbb{R}$;
   $\boldsymbol{x}$, un vector de $\mathbb{R}^{in}$;
   $\boldsymbol{\phi} = (1, \phi_1(\boldsymbol{x}), ..., \phi_{M-1}(\boldsymbol{x}))^T$ con $\phi_i(\cdot)$ $(i = 1, ..., M-1)$ un conjunto de $M-1$ transformaciones (no necesariamente lineales) escalares sobre $\mathbb{R}^{in}$; $\boldsymbol{w} = (w_0, w_1, ..., w_{M-1})^T$, un vector de $\mathbb{R}^M$; y, finalmente, ϵ una variable aleatoria distribuida según una gaussiana $\mathcal{N}(0, \beta^{-1})$. Sea $(\boldsymbol{x}_1, ..., \boldsymbol{x}_N)$ y $(t_1, ..., t_N)$ un conjunto de valores de $\boldsymbol{x}$ y t obtenidos a partir de las distribuciones descritas por todo lo anterior.

   Defínase $\boldsymbol{\Phi}$ como la matriz de dimensiones $N \times M$ en que la fila j es igual a $(1, \phi_1(\boldsymbol{x}_j), ..., \phi_{M-1}(\boldsymbol{x}_j))$.

   Demuestre que la matriz $\Phi(\Phi^T\Phi)^{-1}\Phi^T$ proyecta cualquier vector v en el espacio generado por los vectores columna de ϕ. Utilice este resultado para probar que la solución de mínimos cuadrados dada por $w = (\Phi^T\Phi)^{-1}\Phi^T t$ corresponde a una proyección ortogonal del vector t sobre la variedad (manifold en inglés) lineal S definida por los vectores columna de $\boldsymbol{\Phi}$, $\psi_k = (\phi_k(\boldsymbol{x}_1), ..., \phi_k(\boldsymbol{x}_N))$ como se muestra en la Fig. 1.

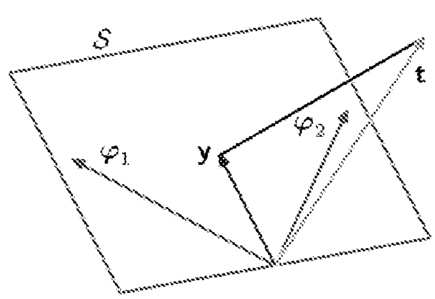

Figura 1:

3. Dado un conjunto de puntos $\{\boldsymbol{x}_n\}$, podemos definir su envolvente convexa (*convex hull*) como el conjunto de todos los puntos $\boldsymbol{x}$ tales que

$$\boldsymbol{x} = \sum_n \alpha_n \boldsymbol{x}_n$$

donde $\alpha_n \geq 0$ y $\sum_n \alpha_n = 1$. Considere un segundo conjunto de puntos $\boldsymbol{y}_n$ y su envolvente convexa. Por definición, los dos conjuntos de puntos $\{\boldsymbol{x}_n\}$ y $\{\boldsymbol{y}_n\}$ serán linealmente separables si existe un vector $\hat{\boldsymbol{w}}$ y un escalar w_0 tales que para todo $\boldsymbol{x}_n$, $\hat{\boldsymbol{w}}^T\boldsymbol{x}_n + w_0 > 0$, y para todo $\boldsymbol{y}_n$, $\hat{\boldsymbol{w}}^T\boldsymbol{y}_n + w_0 < 0$. Demuestre que si sus envolventes convexas tienen intersección no nula, entonces los dos conjuntos de puntos no pueden ser linealmente separables y *vice versa*: si son linealmente separables entonces sus envolventes convexas tienen interesección nula.