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

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

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

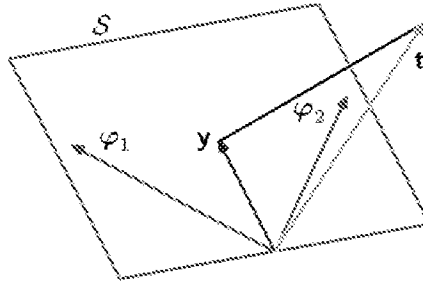


Figura 1:

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

$$x = \sum_n \alpha_n x_n$$

donde $\alpha_n \geq 0$ y $\sum_n \alpha_n = 1$. Considere un segundo conjunto de puntos y_n y su envolvente convexa. Por definición, los dos conjuntos de puntos $\{x_n\}$ y $\{y_n\}$ serán linealmente separables si existe un vector \hat{w} y un escalar w_0 tales que para todo x_n , $\hat{w}^T x_n + w_0 > 0$, y para todo y_n , $\hat{w}^T y_n + w_0 < 0$. 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 intersección nula.