

Lista de Ejercicios (Parte 1 – **ERRATA**)

Prof.: Juan Carlos Martínez-Ovando

19 de febrero de 2016

1. Considera el modelo normal multivariado normal para $\mathbf{x}_1, \dots, \mathbf{x}_n$ vectores p -dimensionales, con

$$\mathbf{x}_i | \boldsymbol{\mu}, \lambda \sim N_p(\mathbf{x} | \boldsymbol{\mu}, \lambda \mathbf{I}), \quad (1.1)$$

donde $\boldsymbol{\mu}$ es un vector p -dimensional, λ es un escalar positivo y \mathbf{I} es la matriz identidad de dimensión $p \times p$. Interpreta $\lambda \mathbf{I}$ como la matriz de precisión del modelo.

Considera adicionalmente que la distribución inicial para $\boldsymbol{\mu}$ y λ está dada por

$$\pi(\boldsymbol{\mu}, \lambda) = N_p(\boldsymbol{\mu} | m_0, \lambda S_0) \text{Ga}(\lambda | a_0, b_0), \quad (1.2)$$

con m_0 un vector p -dimensional, S_0 una matriz simétrica positivo definida de dimensión $p \times p$, a_0 y b_0 escalares positivos, todos conocidos.

- a) Muestra que la distribución (1.2) es conjugada para (1.1).
- b) Calcula la distribución inicial predictiva el modelo, dada por

$$p(\mathbf{x}) = \int p(\mathbf{x} | \boldsymbol{\mu}, \lambda) \pi(\boldsymbol{\mu}, \lambda) d\boldsymbol{\mu} d\lambda. \quad (1.3)$$

- c) Calcula la distribución posterior para $\boldsymbol{\mu}$ y λ dado $\mathbf{x}_1, \dots, \mathbf{x}_n$.
- d) Calcula la distribución predictiva para \mathbf{x}_f , un valor futuro no observado aun, dado $\mathbf{x}_1, \dots, \mathbf{x}_n$.

2. Considera el modelo lineal de regresión para $(y_1, \mathbf{x}_1), \dots, (y_n, \mathbf{x}_n)$, para $\mathbf{x}_1, \dots, \mathbf{x}_n$ vectores p -dimensionales, expresado en forma matricial como

$$\mathbf{y} | \mathbf{X}, \boldsymbol{\beta}, \lambda \sim N_n(\mathbf{y} | \mathbf{X}\boldsymbol{\beta}, \lambda \mathbf{I}), \quad (1.4)$$

donde $\mathbf{y} = (y_1, \dots, y_n)'$ es un vector n -dimensional, $\mathbf{X} = (\mathbf{x}'_1, \dots, \mathbf{x}'_n)'$ es una matriz de dimensión $n \times p$, $\boldsymbol{\beta}$ es un vector p -dimensional, λ es un escalar y \mathbf{I} es la matriz identidad de dimensión $n \times n$. De nuevo, interpreta $\lambda\mathbf{I}$ como la matriz de precisión del modelo.

Considera adicionalmente que la distribución inicial para $\boldsymbol{\beta}$ y λ está dada por

$$\pi(\boldsymbol{\beta}, \lambda) = N_p(\boldsymbol{\beta}|b_0, \lambda S_0) \text{Ga}(\lambda|c_0, d_0), \quad (1.5)$$

con b_0 un vector p -dimensional, S_0 una matriz simétrica positivo definida de dimensión $p \times p$, c_0 y d_0 escalares positivos, todos conocidos.

- a) Muestra que la distribución (1.5) es conjugada para (1.4).
- b) Calcula la distribución inicial predictiva el modelo, dada por

$$p(\mathbf{y}|\mathbf{X}) = \int p(\mathbf{y}|\mathbf{X}, \boldsymbol{\beta}, \lambda) \pi(\boldsymbol{\beta}, \lambda) d\boldsymbol{\beta} d\lambda. \quad (1.6)$$

- c) Calcula la distribución posterior para $\boldsymbol{\beta}$ y λ dado $(y_1, \mathbf{x}_1), \dots, (y_n, \mathbf{x}_n)$.
 - d) Calcula la distribución predictiva para y_f , un valor futuro no observado aun de y , dado $(y_1, \mathbf{x}_1), \dots, (y_n, \mathbf{x}_n)$ y \mathbf{x}_f observado.
3. Considera el modelo lineal de regresión para $(y_1, \mathbf{x}_1), \dots, (y_n, \mathbf{x}_n)$, para $\mathbf{x}_1, \dots, \mathbf{x}_n$ vectores p -dimensionales, expresado en forma matricial como

$$\mathbf{y}|\mathbf{X}, \boldsymbol{\beta}, \lambda \sim N_n(\mathbf{y}|\mathbf{X}\boldsymbol{\beta}, \lambda\mathbf{I}), \quad (1.7)$$

donde $\mathbf{y} = (y_1, \dots, y_n)'$ es un vector n -dimensional, $\mathbf{X} = (\mathbf{x}'_1, \dots, \mathbf{x}'_n)'$ es una matriz de dimensión $n \times p$, $\boldsymbol{\beta}$ es un vector p -dimensional, λ es un escalar y \mathbf{I} es la matriz identidad de dimensión $n \times n$. De nuevo, interpreta $\lambda\mathbf{I}$ como la matriz de precisión del modelo.

Considera adicionalmente que la distribución inicial para $\boldsymbol{\beta}$ y λ está dada por

$$\pi(\boldsymbol{\beta}, \lambda) = N_p(\boldsymbol{\beta}|b_0, S_0) \text{Ga}(\lambda|c_0, d_0), \quad (1.8)$$

con b_0 un vector p -dimensional, S_0 una matriz simétrica positivo definida de dimensión $p \times p$, c_0 y d_0 escalares positivos, todos conocidos.

- a) Responde argumentando si la distribución (1.8) es conjugada para (1.7).

- b) Calcula la distribución inicial predictiva el modelo, dada por

$$p(\mathbf{y}|\mathbf{X}) = \int p(\mathbf{y}|\mathbf{X}, \boldsymbol{\beta}, \lambda) \pi(\boldsymbol{\beta}, \lambda) d\boldsymbol{\beta} d\lambda. \quad (1.9)$$

- c) Calcula la distribución posterior para $\boldsymbol{\beta}$ y λ dado $(y_1, \mathbf{x}_1), \dots, (y_n, \mathbf{x}_n)$.
- d) Calcula la distribución predictiva para y_f , un valor futuro no observado aun de y , dado $(y_1, \mathbf{x}_1), \dots, (y_n, \mathbf{x}_n)$ y \mathbf{x}_f observado.
4. Para el modelo de regresión de la pregunta 1 se sabe que tanto la distribución final para $(\boldsymbol{\beta}, \lambda)$ como la distribución predictiva de y_f en \mathbf{x}_f , dado los datos, son analíticamente cerradas. Aun así, diseña lo siguiente:
- a) Un algoritmo muestreador de Gibbs que actualice $\boldsymbol{\beta}$ y λ en dos bloques.
- b) Un algoritmo muestreador de Gibbs que actualice los componentes de $\boldsymbol{\beta} = (\beta_1, \dots, \beta_p)$ y λ uno a la vez.