

Tarea 3 de estadística multivariada

Bernardo Garcia Bulle Bueno

March 2019

- 1 Muestra que si obtenemos la distribución condicional de X_1 dado $X_2 = x_2$ en el contexto donde (X_1, X_2) tienen distribución gaussiana bivariada, corresponde a la distribución gaussiana donde la media es una regresión de X_1 en x_2**

Defino $z = x_1 + Ax_2$. En este caso $A = -\Sigma_{12}\Sigma_{22}^{-1}$. Σ_{ij} es la covarianza de x_i con x_j y μ_i es la media de x_i . Auxiliariamente, demostraré que $cov(z, x_2) = 0$.

$$\begin{aligned} cov(z, x_2) &= cov(x_1 + Ax_2, x_2) = cov(x_1, x_2) + Acov(x_2, x_2) = \\ &= \Sigma_{12} + Avar(x_2) = \Sigma_{12} + A\Sigma_{22} = \Sigma_{12} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{22} = \\ &= \Sigma_{12} - \Sigma_{12} = 0 \end{aligned}$$

Ahora saquemos $E(X_1|X_2 = x_2)$

$$\begin{aligned} E(X_1|X_2 = x_2) &= E(z - Ax_2|X_2 = x_2) = E(z|X_2 = x_2) - Ax_2 = \\ &= E(z) - Ax_2 = E(X_1) + AE(X_2) - Ax_2 = \mu_1 + A(\mu_2 - x_2) = \\ &= \mu_1 + (\Sigma_{12}\Sigma_{22}^{-1})(x_2 - \mu_2) \end{aligned}$$

- 2 Generalicen el resultado anterior considerando $X = (X_1, X_2)$, donde X es p-variado, X_1 y X_2 son p_1 , y p_2 variados ($p = p_1 + p_2$), con $X \sim N_p(x|\mu, \Sigma)$, para derivar la distribución de X_1 condicional en $X_2 = x_2$**

La forma anterior ya es general para x_1 y x_2 multivariados. Las únicas diferencias son dos. La primera es que Σ_{ij} no es la covarianza, sino es una matriz de

covarianzas con la covarianza de cada componente de x_i con cada componente de x_j , tiene dimensiones $i \times j$. Las filas corresponden a los componentes de x_i . La segunda diferencia es que μ_i es un vector con las medias de los componentes de x_i . $E(x_i)$ tiene las mismas dimensiones que μ_i .

Podemos verificar que las dimensiones funcionan para la fórmula de arriba. z debe tener las dimensiones de x_1 , es decir $p_1 \times 1$. x_1 obviamente las tiene, Ax_2 tiene dimensiones, si tomamos en cuenta cada componente: $(p_1 \times p_2) \times (p_2 \times p_2) \times (p_2 \times 1)$ lo cual resulta en $p_1 \times 1$. De ahí también vemos que A tiene dimensiones $p_1 \times p_2$. Por lo tanto la fórmula de la covarianza funciona pues al final los dos sumandos tienen la misma dimensión.

Por lo mismo, los sumandos de la esperanza final también tendrán las dimensiones $p_1 \times 1$.