III.3. Variable Aleatoria Multidimensional

May 13, 2019

1 Variables Aleatorias N-Dimensionales

Al introducir el concepto de variable aleatoria conjunta, ya se mencionó que los vectores aleatorios pueden tener un número arbitrario M de componentes:

$$\mathbf{X} = (X_1, \dots X_M)^{\mathrm{T}}$$

Ya hemos visto en profundidad la caracterización conjunta y marginal para el caso de dos variables aleatorias (vector aleatorio de dimesión 2). La extensión a M componentes es inmediata:

La **Función de Distribución Conjunta** $F_{\mathbf{X}}(\mathbf{x}) \equiv F_{X_1 \dots X_M}(x_1, \dots x_M)$ es

$$F_{\mathbf{X}}(\mathbf{x}) \equiv F_{X_1 \dots X_M}(x_1, \dots x_M) = P(X_1 \leq x_1, \dots X_M \leq x_M)$$

Las **funciones de distribución marginales** se obtienen sin más que evaluar en ∞ las variables que se desean descartar. Por ejemplo:

$$F_{X_1 X_3}(x_1, x_3) = F_{X_1 \dots X_N}(x_1, \infty, x_3, \infty, \dots \infty)$$

$$F_{X_3}(x_3) = F_{X_1 \dots X_N}(\infty, \infty, x_3, \infty, \dots \infty)$$

Por supuesto, $F_{\mathbf{X}}(\boldsymbol{\infty}) \equiv F_{X_1 \dots X_M}(\infty, \dots \infty) = 1$

Suele resultar oportuno definir el vector aleatorio $\mathbf{V}$ con dos componentes que son, a su vez, vectores aleatorios, $\mathbf{X} = (X_1, \dots X_M)^{\mathrm{T}}$ e $\mathbf{X} = (X_1, \dots X_M)^{\mathrm{T}}$:

$$\mathbf{V} = (\mathbf{X}^{\mathrm{T}}, \mathbf{Y}^{\mathrm{T}})^{\mathrm{T}} = (X_1, \dots X_M, Y_1 \dots Y_M)^{\mathrm{T}}$$

Ello permite aplicar de forma prácticamente directa las expresiones vistas para la caracterización del vector bidimensional (X, Y) al vector $M + N$-dimensional $\mathbf{V} = (\mathbf{X}^{\mathrm{T}}, \mathbf{Y}^{\mathrm{T}})^{\mathrm{T}}$, recurriendo a ciertos *abusos de notación* muy sencillos de entender. Por ejemplo:

$$F_{\mathbf{V}}(\mathbf{v}) \equiv F_{\mathbf{XY}}(\mathbf{x}, \mathbf{y}) = P(X_1 \leq x_1, \dots X_M \leq x_M, Y_1 \leq y_1 \dots Y_M \leq y_M) \equiv P(\mathbf{X} \leq \mathbf{x}, \mathbf{Y} \leq \mathbf{y})$$

$$F_{\mathbf{X}}(\mathbf{x}) = F_{\mathbf{XY}}(\mathbf{x}, \boldsymbol{\infty}) \equiv F_{X_1 \dots X_M Y_1 \dots Y_N}(x_1, \dots x_M, \infty, \dots \infty) = F_{X_1 \dots X_M}(x_1, \dots x_M)$$

$$F_{\mathbf{Y}}(\mathbf{y}) = F_{\mathbf{XY}}(\boldsymbol{\infty}, \mathbf{y}) \equiv F_{X_1 \dots X_M Y_1 \dots Y_N}(\infty, \dots \infty, y_1, \dots y_N) = F_{Y_1 \dots Y_N}(y_1, \dots y_N)$$

1.0.1 Función de densidad conjunta multidimensional

La **función de densidad conjunta** del vector aleatorio **X**, suponiendo que está compuesto por **variables aleatorias continuas**, es:

$$f_{\mathbf{X}}(\mathbf{x}) \equiv f_{X_1 \ldots X_M}(x_1, \ldots x_M) = \frac{\partial^M F_{X_1 \ldots X_M}(x_1, \ldots, x_M)}{\partial x_1, \ldots, \partial x_M} \equiv \frac{\boldsymbol{\partial} F_{\mathbf{X}}(\mathbf{x})}{\boldsymbol{\partial}\mathbf{x}}$$

Por tanto, puede obtenerse la función de distribución a partir de la de densidad conjunta:

$$F_{\mathbf{X}}(\mathbf{x}) \equiv F_{X_1 \ldots X_M}(x_1, \ldots x_M) = \int_{-\infty}^{x_1} \cdots \int_{-\infty}^{x_M} f_{X_1 \ldots X_M}(x_1, \ldots x_M) dx_1 \ldots dx_M \equiv \int_{-\infty}^{\mathbf{x}} f_{\mathbf{X}}(\mathbf{x}) \mathbf{dx}$$

La masa total de probabilidad es uno:

$$\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_{X_1 \ldots X_M}(x_1, \ldots x_M) dx_1 \ldots dx_M = F_{X_1 \ldots X_M}(\infty, \ldots \infty) = 1 \equiv \ \equiv \int_{-\infty}^{\infty} f_{\mathbf{X}}(\mathbf{x}) \mathbf{dx} = F_{\mathbf{X}}(\boldsymbol{\infty}) = 1$$

Las **funciones de densidad marginales** se obtienen sin más que integrar las variables que se desean descartar. Por ejemplo:

$$f_{X_1 X_3}(x_1, x_3) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_{X_1 \ldots X_M}(x_1, \ldots x_M) dx_2 dx_4 \ldots dx_M$$

$$f_{X_1}(x_1) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_{X_1 \ldots X_M}(x_1, \ldots x_M) dx_2 \ldots dx_M$$

Considerando ahora el vector aleatorio **V** definido como antes:

$$f_{\mathbf{V}}(\mathbf{v}) \equiv f_{\mathbf{XY}}(\mathbf{x}, \mathbf{y}) = \frac{\boldsymbol{\partial}^2 F_{\mathbf{XY}}(\mathbf{x}, \mathbf{y})}{\boldsymbol{\partial}\mathbf{x} \boldsymbol{\partial}\mathbf{y}}$$

$$F_{\mathbf{V}}(\mathbf{v}) \equiv F_{\mathbf{XY}}(\mathbf{x}, \mathbf{y}) = \int_{-\infty}^{\mathbf{x}} \int_{-\infty}^{\mathbf{y}} f_{\mathbf{XY}}(\mathbf{x}, \mathbf{y}) \mathbf{dx dy} \equiv \int_{-\infty}^{\mathbf{v}} f_{\mathbf{V}}(\mathbf{v}) \mathbf{dv}$$

Y las funciones de densidad marginales de ambos vectores componentes son:

$$f_{\mathbf{X}}(\mathbf{x}) = \int_{-\infty}^{\infty} f_{\mathbf{XY}}(\mathbf{x}, \mathbf{y}) \mathbf{dy}$$

$$f_{\mathbf{Y}}(\mathbf{y}) = \int_{-\infty}^{\infty} f_{\mathbf{XY}}(\mathbf{x}, \mathbf{y}) \mathbf{dx}$$

Por supuesto, $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{\mathbf{XY}}(\mathbf{x}, \mathbf{y}) \mathbf{dx dy} = F_{\mathbf{XY}}(\infty, \infty) = 1$

Como ya sabemos, las funciones marginales no permiten conocer la conjunta salvo que sean independientes. La **independencia conjunta** requiere que todas las combinaciones posibles de variables aleatorias componentes sean independientes. En tal caso:

$$F_{X_1 \ldots X_M}(x_1 \ldots x_M) = F_{X_1}(x_1) \ldots F_{X_M}(x_M)$$

$$f_{X_1 \ldots X_M}(x_1 \ldots x_M) = f_{X_1}(x_1) \ldots f_{X_M}(x_M)$$

En el caso de que tengamos dos vectores aleatorios **X**, **Y** independientes entre sí (pero no necesariamente las componentes de cada uno de ellos por separado):

$$F_{\mathbf{XY}}(\mathbf{x},\mathbf{y}) = F_{\mathbf{X}}(\mathbf{x})F_{\mathbf{Y}}(\mathbf{y})$$

$$f_{\mathbf{XY}}(\mathbf{x},\mathbf{y}) = f_{\mathbf{X}}(\mathbf{x})f_{\mathbf{Y}}(\mathbf{y})$$

1.0.2 Función de masa de probabilidad conjunta

Si las componentes del vector aleatorio son **variables aleatorias discretas**, la función de distribución resulta discontinua y no es derivable. Como ya hemos visto, en tal caso recurrimos a la **función de masa de probabilidad conjunta**

$$p_{\mathbf{X}}(\mathbf{x_i}) \equiv p_{X_1 \ldots X_M}(x_{i_1}, \ldots x_{i_M}) = P(X_1 = x_{i_1}, \ldots X_M = x_{i_M})$$

$$p_{\mathbf{V}}(\mathbf{v_k}) \equiv p_{\mathbf{XY}}(\mathbf{x_i}, \mathbf{y_j}) = P(\mathbf{X} = \mathbf{x_i}, \mathbf{Y} = \mathbf{y_j})$$

donde $\mathbf{i} = (i_1 \ldots i_M)$, $\mathbf{j} = (j_1 \ldots j_N)$ y **k** es la concatenación de ambos. La **función de distribución conjunta** se obtiene acumulando la función de masa de probabilidad conjunta:

$$F_{\mathbf{X}}(\mathbf{x_i}) \equiv F_{X_1 \ldots X_M}(x_{i_1}, \ldots x_{i_M}) = \sum_{p_1=-\infty}^{i_1} \cdots \sum_{p_M=-\infty}^{i_M} f_{X_1 \ldots X_M}(x_{p_1}, \ldots x_{p_M}) \equiv \textstyle\sum_{\mathbf{p}=-\infty}^{\mathbf{i}} f_{\mathbf{X}}(\mathbf{x_p})$$

$$F_{\mathbf{V}}(\mathbf{v_k}) \equiv F_{\mathbf{XY}}(\mathbf{x_i}, \mathbf{y_j}) = \sum_{\mathbf{p}=-\infty}^{\mathbf{i}} \sum_{\mathbf{q}=-\infty}^{\mathbf{j}} p_{\mathbf{XY}}(\mathbf{x_p}, \mathbf{y_q}) \equiv \textstyle\sum_{\mathbf{r}=-\infty}^{\mathbf{k}} f_{\mathbf{V}}(\mathbf{v_r})$$

1.0.3 Momentos conjuntos: Vector de medias y matrices de covarianza

El **vector de medias** se obtiene a partir de la esperanza matemática de las funciones marginales de cada una de las variables aleatorias, como ya se vio en el caso bidimensional:

$$E(\mathbf{X}) \equiv \boldsymbol{\mu}_{\mathbf{X}} = (E(X_1), \ldots E(X_M))^{\mathrm{T}}$$

$$E(\mathbf{V}) \equiv \boldsymbol{\mu}_{\mathbf{V}} = \left(\boldsymbol{\mu}_{\mathbf{X}}^{\mathrm{T}}, \boldsymbol{\mu}_{\mathbf{Y}}^{\mathrm{T}}\right)^{\mathrm{T}} = \left(E(\mathbf{X})^{\mathrm{T}}, E(\mathbf{Y})^{\mathrm{T}}\right)^{\mathrm{T}} = (E(X_1), \ldots E(X_M), E(Y_1), \ldots E(Y_N))^{\mathrm{T}}$$

La **matriz de autocorrelación** del vector aleatorio **X** es:

$$\mathbf{R_{XX}} = E(\mathbf{X}\mathbf{X}^{\mathrm{T}}) = \begin{pmatrix} E(X_1^2) & \ldots & R_{X_1 X_M} \\ \vdots & \ddots & \vdots \\ R_{X_M X_1} & \ldots & E(X_M^2) \end{pmatrix}$$

que resulta **simétrica y semidefinida positiva**. Si todas las componentes son ortogonales entre sí, la matriz resulta diagonal.

La **matriz de correlación cruzada** de los vectores aleatorios X e Y es:

$$\mathbf{R_{XY}} = E(\mathbf{X}\mathbf{Y}^{\mathrm{T}}) = \begin{pmatrix} R_{X_1Y_1} & \dots & R_{X_1Y_N} \\ \vdots & \ddots & \vdots \\ R_{X_MY_1} & \dots & R_{X_MY_N} \end{pmatrix}$$

y se cumple que $\mathbf{R_{YX}} = \mathbf{R_{XY}}^{\mathrm{T}}$. Por tanto, la matriz de autocorrelación del vector $\mathbf{V} = (\mathbf{X}^{\mathrm{T}}, \mathbf{Y}^{\mathrm{T}})^{\mathrm{T}}$ es

$$\mathbf{R_V} = E(\mathbf{V}\mathbf{V}^{\mathrm{T}}) = \begin{pmatrix} \mathbf{R_{XX}} & \mathbf{R_{XY}} \\ \mathbf{R_{YX}} & \mathbf{R_{YY}} \end{pmatrix}$$

que también resulta simétrica y semidefinida positiva.

La **matriz de autocovarianza** del vector aleatorio **X** es:

$$\mathbf{C_{XX}} = E((\mathbf{X} - \text{¯}_{\mathbf{X}})(\mathbf{X} - \text{¯}_{\mathbf{X}})^{\mathrm{T}}) = \begin{pmatrix} \sigma^2_{X_1} & \dots & C_{X_1X_M} \\ \vdots & \ddots & \vdots \\ C_{X_MX_1} & \dots & \sigma^2_{X_M} \end{pmatrix} = \mathbf{R_{XX}} - \boldsymbol{\mu}_{\mathbf{X}}\boldsymbol{\mu}_{\mathbf{X}}^{\mathrm{T}}$$

que resulta **simétrica y semidefinida positiva**. Si todas las componentes son **incorreladas** entre sí, la matriz resulta **diagonal**: $\mathbf{C_{XX}} = \mathbf{diag}[\sigma^2_{X_1} \dots \sigma^2_{X_M}]$

Las covarianzas de las componentes (elementos fuera de la diagonal) pueden ponerse en función de los coeficientes de correlación:

$$C_{X_iX_j} = \rho_{X_iX_j}\sigma_{X_i}\sigma_{X_j}$$

La **matriz de covarianza cruzada** de los vectores aleatorios X e Y es:

$$\mathbf{C_{XY}} = E((\mathbf{X} - \text{¯}_{\mathbf{X}})(\mathbf{Y} - \text{¯}_{\mathbf{Y}})^{\mathrm{T}}) = \begin{pmatrix} C_{X_1Y_1} & \dots & C_{X_1Y_N} \\ \vdots & \ddots & \vdots \\ C_{X_MY_1} & \dots & C_{X_MY_N} \end{pmatrix} = \mathbf{R_{XX}} - \boldsymbol{\mu}_{\mathbf{X}}\boldsymbol{\mu}_{\mathbf{Y}}^{\mathrm{T}}$$

y se cumple que $\mathbf{C_{YX}} = \mathbf{C_{XY}}^{\mathrm{T}}$. Por tanto, la matriz de autocovarianza del vector $\mathbf{V} = (\mathbf{X}^{\mathrm{T}}, \mathbf{Y}^{\mathrm{T}})^{\mathrm{T}}$ es

$$\mathbf{C_V} = E((\mathbf{V} - \text{¯}_{\mathbf{V}})(\mathbf{V} - \text{¯}_{\mathbf{V}})^{\mathrm{T}}) = \begin{pmatrix} \mathbf{C_{XX}} & \mathbf{C_{XY}} \\ \mathbf{C_{YX}} & \mathbf{C_{YY}} \end{pmatrix}$$

que también resulta simétrica y semidefinida positiva.

1.0.4 Distribuciones condicionadas. Teoremas de Bayes y de la Probabilidad Total

Consideremos el vector aleatorio $\mathbf{X} = (X_1, \dots X_{p-1}, X_p, X_{p+1} \dots X_M)^{\mathrm{T}}$. Podemos obtener la **función de distribución condicionada** de unas componentes por otras:

$$F_{X_1 \dots X_{p-1}|X_p \dots X_M}(x_1, \dots x_{p-1}|x_p, \dots x_M) = \frac{F_{X_1 \dots X_M}(x_1, \dots x_M)}{F_{X_p \dots X_M}(x_p, \dots x_M)}$$

Considerando que las variables aleatorias son continuas, la **función de densidad condicionada** es:

$$f_{X_1 \dots X_{p-1}|X_p \dots X_M}(x_1, \dots x_{p-1}|x_p, \dots x_M) = \frac{f_{X_1 \dots X_M}(x_1, \dots x_M)}{f_{X_p \dots X_M}(x_p, \dots x_M)}$$

Si las variables aleatorias son discretas, la **función de masa de probabilidad** resulta:

$$p_{X_1\ldots X_{p-1}|X_p\ldots X_M}(x_{i_1},\ldots x_{i_{p-1}}|x_{i_p},\ldots x_{i_M}) = \frac{p_{X_1\ldots X_M}(x_{i_1},\ldots x_{i_M})}{p_{X_p\ldots X_M}(x_{i_p},\ldots x_{i_M})}$$

Regla de la cadena Aplicando sucesiva y ordenadamente la definición de función de densidad condicionada:

$$\begin{aligned} f_{X_1\ldots X_M}(x_1\ldots x_M) &= f_{X_1|X_2\ldots X_M}(x_1|x_2\ldots x_M)\ldots f_{X_{M-1}|X_M}(x_{M-1}|x_M)f_{X_M}(x_M) \quad (1)\\ &= f_{X_M|X_{M-1}\ldots X_1}(x_M|x_{M-1}\ldots x_1)\ldots f_{X_2|X_1}(x_2|x_1)f_{X_1}(x_1) \quad (2)\end{aligned}$$

Adviértase que no hay un orden establecido por el cual haya que condicionar las funciones de densidad. Además de estas dos formas de construir la cadena de productos, podrían haberse elaborado otras extrayendo en distinto orden las variables condicionantes. Cuando veamos secuencias o procesos temporales, la causalidad establecerá un ordenamiento que sí habrá que respetar.

Si las variables son discretas, puede obtenerse una descomposición idéntica utilizando funciones de masa de probabilidad.

Ecuación de Chapman-Kolmogorov Se obtiene combinando la definición de función de densidad o de masa de probbailidad con la obtención de funciones marginales. Supone una *extensión del Teorema de la Probabilidad Total*. Veámoslo con unos ejemplos:

$$f_{X_3|X_1}(x_3|x_1) = \int_{-\infty}^{\infty} f_{X_3X_2|X_1}(x_3,x_2|x_1)dx_2 = \int_{-\infty}^{\infty} f_{X_3|X_2X_1}(x_3|x_2,x_1)f_{X_2}(x_2|x_1)dx_2$$

$$f_{X_1|X_4}(x_1|x_4) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X_1X_2X_3|X_4}(x_1,x_2,x_3|x_4)dx_2dx_3 = = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X_1|X_2X_3X_4}(x_1|x_2,x_3,x_4)f_{X_2X_3|X_4}(x_2,x_3|x_4)dx$$

La ecuación puede aplicarse al caso discreto sin más que sustituir las funciones de densidad por funciones de masa de probabilidad y las integrales por sumatorios

La **función de distribución de un vector aleatorio, Y, condicionada por otro**, **X**, o viceversa, se define extendiendo la definición ya vista para dos variables alatorias

$$F_{\mathbf{Y}|\mathbf{X}}(\mathbf{y}|\mathbf{x}) = \frac{F_{\mathbf{XY}}(\mathbf{x},\mathbf{y})}{F_{\mathbf{X}}(\mathbf{x})} \qquad F_{\mathbf{X}|\mathbf{Y}}(\mathbf{x}|\mathbf{y}) = \frac{F_{\mathbf{XY}}(\mathbf{x},\mathbf{y})}{F_{\mathbf{Y}}(\mathbf{y})}$$

A partir de tal definición es inmediato obtener, para el **caso continuo**, la **función de densidad condicionada**:

$$f_{\mathbf{Y}|\mathbf{X}}(\mathbf{y}|\mathbf{x}) = \frac{f_{\mathbf{XY}}(\mathbf{x},\mathbf{y})}{f_{\mathbf{X}}(\mathbf{x})} \qquad f_{\mathbf{X}|\mathbf{Y}}(\mathbf{x}|\mathbf{y}) = \frac{f_{\mathbf{XY}}(\mathbf{x},\mathbf{y})}{f_{\mathbf{Y}}(\mathbf{y})}$$

Y, para el **caso discreto**, la **función de masa de probabilidad condicionada**:

$$p_{\mathbf{Y}|\mathbf{X}}(\mathbf{y_j}|\mathbf{x_i}) = \frac{p_{\mathbf{XY}}(\mathbf{x_i},\mathbf{y_j})}{p_{\mathbf{X}}(\mathbf{x_i})} \qquad p_{\mathbf{X}|\mathbf{Y}}(\mathbf{x_i}|\mathbf{y_j}) = \frac{p_{\mathbf{XY}}(\mathbf{x_i},\mathbf{y_j})}{p_{\mathbf{Y}}(\mathbf{y_j})}$$

1.0.5 Transformaciones Lineales y Teorema del Límite Central

Es posible definir transformaciones $\mathbf{Y} = \mathbf{g}(\mathbf{X})$ que convierten el vector aleatorio $\mathbf{X}$ en otro $\mathbf{Y}$, cuya caracterización conjunta es posible obtener. Entre todas ellas, vamos a prestarle atención a las **transformaciones lineales**. Por ejemplo:

$$Y = k_1 X_1 + \ldots k_M X_M = \mathbf{k}^{\mathrm{T}}\mathbf{X}$$

La media resulta:

$$\mu_Y \equiv E(Y) = \mathbf{k}^{\mathrm{T}}\boldsymbol{\mu}_{\mathbf{X}}$$

Y la varianza:

$$\sigma_Y^2 = E\left(\mathbf{k}^{\mathrm{T}}(\mathbf{X} - \boldsymbol{\mu}_{\mathbf{X}})(\mathbf{X} - \boldsymbol{\mu}_{\mathbf{X}})^{\mathrm{T}}\mathbf{k}\right) = \mathbf{k}^{\mathrm{T}}\mathbf{C}_{\mathbf{XX}}\mathbf{k}$$

Si el vector aleatorio $\mathbf{X}$ tiene sus componentes incorreladas $\mathbf{C}_{\mathbf{XX}} = \mathbf{diag}[\sigma_{X_1}^2 \ldots \sigma_{X_M}^2]$ y:

$$\sigma_Y^2 = k_1^2 \sigma_{X_1}^2 + \ldots + k_M \sigma_{X_M}^2$$

En el caso que las componentes del vector $\mathbf{X}$ sean independientes, y que los coeficiones de la combinación lineal sean todos la unidad, $\mathbf{k} = (1 \ldots 1)^{\mathrm{T}}$, puede demostrarse que la función de densidad de probabilidad de Y es la **convolución** de las funciones de densidad de probabilidad (en el caso discreto, se trataría de la **convolución discreta** de las funciones de masa de probabilidad):

$$f_Y(y) = f_{X_1}(y) * \ldots * f_{X_M}(y)$$

En el caso de las variables aleatorias $X_1 \ldots X_M$ sean Gaussianas, su convolución resulta también Gaussiana, con media $\mu_Y = \mu_1 + \ldots + \mu_M$. Como la independencia implica incorrelación, por la varianza es $\sigma_Y^2 = \sigma_{X_1}^2 + \ldots + \sigma_{X_M}^2$. Por tanto $f_Y(y) = N(\mu_Y, \sigma_Y^2)$.

El **Teorema del Límite Central** establece que, cuando M es muy grande, la función de densidad de la superposición de variables aleatorias independientes tiende a una Gaussiana, independientemente de la caracterización de cada una de ellas. Este resultado justifica el abundante uso de Gaussianas en la práctica, pues frecuentemente nos encontraremos con superposiciones aleatorias de muchos elementos independientes.

La transformación lineal vista anteriormente puede extenderse como sigue:

$$\mathbf{Y} = \begin{pmatrix} Y_1 \\ \vdots \\ Y_N \end{pmatrix} = \begin{bmatrix} k_{1,1} & \ldots & k_{1,M} \\ \vdots & \ddots & \vdots \\ k_{N,1} & \ldots & k_{N,M} \end{bmatrix} \begin{pmatrix} X_1 \\ \vdots \\ X_M \end{pmatrix} + \begin{pmatrix} m_1 \\ \vdots \\ m_N \end{pmatrix} = \mathbf{K}^{\mathrm{T}}\mathbf{X} + \mathbf{m}$$

Por tanto:

$$\boldsymbol{\mu}_{\mathbf{Y}} \equiv E(\mathbf{Y}) = \mathbf{K}^{\mathrm{T}}\boldsymbol{\mu}_{\mathbf{X}} + \mathbf{m}$$

$$\mathbf{C}_{\mathbf{YY}} = \mathbf{K}^{\mathrm{T}}\mathbf{C}_{\mathbf{XX}}\mathbf{K}$$

Podemos también obtener la matriz de covarianza cruzada entre el resultado y los operandos:

$$\mathbf{C}_{\mathbf{YX}} = \mathbf{K}^{\mathrm{T}}\mathbf{C}_{\mathbf{XX}}$$