

La Conjetura de Correlación Gaussiana: Demostración en el Plano

Saúl Toscano Palmerín*
Departamento de Matemáticas
Universidad de Guanajuato
toscano.saul@gmail.com

Verano del 2011

Resumen

Se analiza y se desglosa la demostración que da Loren D. Pitt para la conjetura de correlación gaussiana en el plano. En el apéndice se presentan definiciones y material para comprender el artículo.

La conjetura enuncia que si A y B son dos conjuntos convexos y simétricos en \mathbb{R}^n , X un vector aleatorio con distribución normal multivariante con vector de esperanza cero y matriz de covarianza la identidad, entonces

$$\mathbb{P}[X \in A \cap B] \geq \mathbb{P}[X \in A]\mathbb{P}[X \in B].$$

Otra formulación y su prueba se enuncian en el Teorema 3.

El siguiente resultado de análisis es fundamental para probar la conjetura en el plano.

Teorema 1. Sean $f(x)$ y $g(x)$ funciones en \mathbb{R}^2 pares, suaves y casi-cóncavas. Supongamos también que los gradientes $\nabla f(x)$ y $\nabla g(x)$ nunca se anulan si $x \neq 0$. Entonces para cualquier función no negativa $\phi(x) = \phi(|x|)$ que sea una función decreciente de $|x|$,

$$\int_{\mathbb{R}^2} \nabla f(x) \cdot \nabla g(x) \phi(x) d^2x \geq 0$$

siempre que la integral converja.

*Trabajo realizado durante el IV Verano de Probabilidad y Estadística del CIMAT bajo la dirección del Dr. Víctor Manuel Pérez Abreu Carrión.

Demostración. Consideremos el conjunto $F(\lambda) \equiv \{x : f(x) \geq \lambda\}$ donde $\lambda \in \mathbb{R} - \{f(0)\}$. El cual es simétrico pues f es par. Consideremos el conjunto $B = \{x : f(x) < \lambda\}$ que es abierto pues f es continua y por ende $F(\lambda)$ es cerrado. También $F(\lambda)$ es convexo, pues si $t \in (0, 1)$

$$f((1-t)x + ty) \geq \min(f(x), f(y)) \geq \lambda$$

por ser f casi-cóncava.

Demostremos que $\partial F(\lambda) = \{x : f(x) = \lambda\}$ o $\partial F(\lambda) = \emptyset$. Supongamos que $\partial F(\lambda) \neq \emptyset$, de aquí que

$$\partial F(\lambda) = \overline{F(\lambda)} \cap \overline{(F(\lambda))^c} = F(\lambda) \cap (F(\lambda)^c)' \neq \emptyset$$

Observemos que $\partial F(\lambda) \subset \{x : f(x) \leq \lambda\}$, pues si x es tal que $f(x) > \lambda$, como f es continua existe $\epsilon > 0$ tal que si $z \in B_\epsilon(x)$ entonces $f(z) > \lambda$, por ende $x \notin (F(\lambda)^c)'$. Usando lo anterior obtenemos que $\partial F(\lambda) \subset \{x : f(x) = \lambda\}$. Finalmente $\partial F(\lambda) = \{x : f(x) = \lambda\}$, pues si existiera x tal que $f(x) = \lambda$ y que $x \notin \partial F(\lambda)$, entonces existe $B_\epsilon(x) \subset F(\lambda)$ y por ende $\nabla f(x) = 0$. Por hipótesis $x \neq 0$, así que tenemos una contradicción. En el caso que $\partial F(\lambda) = \{x : f(x) = \lambda\}$, como $\nabla f(x) \neq 0$ si $x \neq 0$, entonces $\partial F(\lambda)$ es una curva suave, simétrica y convexa en \mathbb{R}^2 . Resultados similares se obtienen para $G(\mu) = \{x : g(x) \geq \mu\}$ y $\partial G(\mu) = \{x : g(x) = \mu\}$.

Para cada $x \neq 0$ sea $\theta(x)$, $0 \leq \theta(x) < \pi$ ¹, el ángulo entre los vectores $\nabla f(x)$ y $\nabla g(x)$. Es decir, $\cos \theta(x) = \frac{\nabla f(x) \cdot \nabla g(x)}{|\nabla f(x)| |\nabla g(x)|}$, $\nabla f(x)$ es la normal interior a la curva $\partial F(f(x))$ y $\nabla g(x)$ es la normal interior a la curva $\partial G(g(x))$. Sea $T = \{x \in \mathbb{R}^2 : \frac{\pi}{2} < \theta(x) < \pi\}$ y para cada $x \in T$ definamos $y = \alpha(x)$ el primer punto de $\partial G(g(x))$ que se alcanza desde x recorriendo la curva $\partial F(f(x))$ en sentido contrario a las manecillas del reloj. Demostremos que α está bien definida, es decir hay que probar que existe un primer punto en que la función satisface lo anterior. Como f y g son pares existe un punto. Ahora hay que probar que $\partial F(f(x))$ y $\partial G(g(x))$ no coinciden en un intervalo de puntos en dirección contraria a las manecillas del reloj. Para esto sean γ_1 y γ_2 las curvas respectivas con $\gamma_1(0) = \gamma_2(0) = x$, si sucediera lo anterior entonces tendríamos que $\gamma_1'(0) = \gamma_2'(0)$ y por tanto $\theta(x) \in \{0, \pi\}$, lo cual es una contradicción.

De la Figura 1

$$\begin{aligned} A &= 180 - \theta(x) - r - s < 90 - r - s \leq 90, \\ p &= \sqrt{a^2 + b^2 - 2ab \cos A}, \\ q &= \sqrt{a^2 + b^2 + 2ab \cos A}. \end{aligned}$$

¹Si $\theta(x) = \pi$, entonces $\nabla f(x) = -\beta \nabla g(x)$ con $\beta > 0$, pero la dirección de $\nabla f(x)$ es hacia el interior de la curva $\partial F(f(x))$ y la dirección de $\nabla g(x)$ sería hacia el exterior de la curva $\partial G(g(x))$, lo cual es imposible.

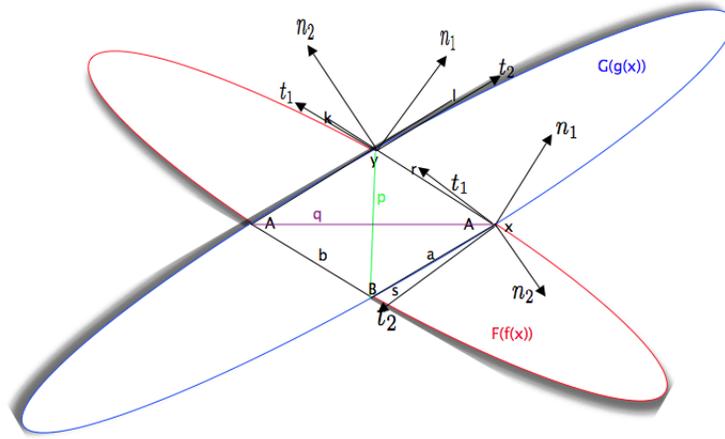


Figura 1: $n_1 = -\frac{\nabla f(x)}{|\nabla f(x)|}$, $n_2 = -\frac{\nabla g(x)}{|\nabla g(x)|}$; t_1 =tangente a la curva $\partial F(f(x))$,
 t_2 =tangente a la curva $\partial G(g(x))$.

Además

$$\begin{aligned}
 \cos A &\geq 0 \\
 \Rightarrow -\cos A &\leq \cos A \\
 \Rightarrow a^2 + b^2 - 2ab \cos A &\leq a^2 + b^2 + 2ab \cos A \\
 \Rightarrow p &\leq q \\
 \Rightarrow |y| &\leq |x|. \tag{1}
 \end{aligned}$$

Luego,

$$\begin{aligned}
 180 - \theta(y) + 2(180 - B) + B - k - l &= 360 \\
 \Rightarrow 180 - \theta(y) - k - l - B &= 0 \\
 \Rightarrow \theta(y) &= 180 - k - l - 180 + A < 90 - k - l \\
 \Rightarrow \theta(y) &\text{ es agudo.}
 \end{aligned}$$

Como $\frac{\pi}{2} < \theta(x) < \pi \Rightarrow 0 < -\cot \theta(x)$. Por lo anterior,

$$\begin{aligned}
 \theta(y) &= A - k - l = 180 - \theta(x) - m \quad (m \geq 0) \\
 \Rightarrow 0 &\leq \theta(x) + \theta(y) = 180 - m \leq 180 \\
 \Rightarrow \sin(\theta(x) + \theta(y)) &\geq 0 \\
 \Rightarrow \cos \theta(y) \sin \theta(x) + \cos \theta(x) \sin \theta(y) &\geq 0 \Rightarrow -\cos \theta(x) \sin \theta(y) \leq \cos \theta(y) \sin \theta(x) \\
 \Rightarrow -\frac{\cos \theta(x)}{\sin \theta(x)} &\leq \frac{\cos \theta(y)}{\sin \theta(y)} \text{ si } \theta(y) > 0 \\
 \Rightarrow -\cot \theta(x) &\leq \cot \theta(y) \text{ si } \theta(y) > 0. \tag{2}
 \end{aligned}$$

Definamos $T_0 = \{x \in T: \theta(\alpha(x)) \neq 0\}$ y la función $\Phi : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ por $\Phi(x) = (f(x), g(x))$. El determinante Jacobiano es

$$|J_\Phi(x)| = |f_{x_1}(x)g_{x_2}(x) - f_{x_2}(x)g_{x_1}(x)| = |\nabla f(x)| |\nabla g(x)| \sin \theta(x) \quad (3)$$

Si $x \in T$, $\Phi(\alpha(x)) = \Phi(x)$ y entonces $\Phi(T - T_0) = \Phi(\alpha(T - T_0))$.

$$\begin{aligned} \text{Si } \alpha(x) &\in \alpha(T - T_0) \\ &\Rightarrow \theta(\alpha(x)) = 0 \\ &\Rightarrow \frac{\nabla f(\alpha(x)) \cdot \nabla g(\alpha(x))}{|\nabla f(\alpha(x))| |\nabla g(\alpha(x))|} = 1 \\ &\Rightarrow |D\Phi(\alpha(x))| = 0 \\ &\Rightarrow \text{rank} D\Phi(\alpha(x)) < 2 \\ &\Rightarrow \alpha(T - T_0) \subset \{x \in \mathbb{R}^2 : \text{rank} D\Phi(x) < 2\}. \end{aligned}$$

Por el Teorema de Sard y Brown (ver [4])

$$m(\Phi(\{x : \text{rank} D\Phi(x) < 2\})) = 0 \Rightarrow m(\Phi(T - T_0)) = m(\Phi(\alpha(T - T_0))) = 0.$$

Sean

$$B_1 = T \cap \{(x, y) \in \mathbb{R}^2 : x > 0\} \text{ y } B_2 = T \cap \{(x, y) \in \mathbb{R}^2 : x < 0\}.$$

Observemos que Φ es inyectiva en B_1 , pues si

$$\begin{aligned} \Phi(x) &= \Phi(y) \\ &\Rightarrow f(x) = f(y) \text{ y } g(x) = g(y) \\ &\Rightarrow \begin{cases} y = \alpha(x) & \text{pero } \theta(y) \text{ es agudo lo cual es una contradicción,} \\ y = -\alpha(x) & \text{pero } \theta(y) \text{ es agudo lo cual es una contradicción,} \\ y = -x & \text{pero } x \notin C. \end{cases} \end{aligned}$$

Por ende Φ es inyectiva. Análogamente Φ es inyectiva en B_2 . También $\Phi \in C^1$ y $|J_\Phi(x)| \neq 0$ si $x \in T - T_0$. Enlistemos otros hechos necesarios y fáciles de ver:

- $|J_\Phi(x)|$ es medible,
- θ es continua pues $\theta(x) = \arccos(n_1(x)n_2(x))$,
- T_0 es medible pues θ es continua y α también,
- T_0 es abierto (pues ∇f y ∇g son continuas) $\Rightarrow T - T_0$ es medible,
- B_1 y B_2 son abiertos, y por ende, $\Phi(B_1)$ y $\Phi(B_2)$ son abiertos.

Sean $H_i : \Phi(B_i) \rightarrow \mathbb{R}^2$ definidas por $H_i(x) = \Phi^{-1}(x)$ con $i = 1, 2$. Como $|J_\Phi(x)| \neq 0$ si $x \in T$, entonces $H_i \in C^1$ y por ende son medibles. Además, H_i es diferenciable en el conjunto medible $\Phi(B_i \cap T_0^c)$ ² con $i = 1, 2$. Usando el hecho de que

²Son medibles pues $m(\Phi(T - T_0)) = 0$.

$m(\Phi(B_i - T_0))=0$ con $i = 1, 2$, concluimos que $m(H_i(\Phi(B_i \cap T_0^c))) = 0$ (ver [3, Cap. 15. J, p. 500]) y por ende $m(B_i \cap T_0^c) = 0^3$ con $i = 1, 2$. Luego

$$\begin{aligned} T \cap T_0^c &= (B_1 \cap T_0^c) \cup B_2 \cap T_0^c \cup (T_0^c \cap T \cap \{(0, y) \in \mathbb{R}^2\}), \\ (T_0^c \cap T \cap \{(0, y) \in \mathbb{R}^2\}) &\subset \{(0, y) \in \mathbb{R}^2\} \end{aligned}$$

de donde

$$m((T_0^c \cap T \cap \{(0, y) \in \mathbb{R}^2\})) = 0 \Rightarrow m(T \cap T_0^c) = 0.$$

Ahora demosetremos que $\alpha \in C^1(T_0)$, Pitt indica que primero hay que demostrar la continuidad y luego aplicando la regla de la cadena a $\Phi(\alpha(x)) = \Phi(x)$ se obtiene que α es diferenciable, sin embargo yo voy a dar una demostración de esto mismo usando el teorema de la función inversa.

Sea $x \in T_0$ entonces $\theta(\alpha(x)) \neq 0$, por tanto $|J_{\Phi}(x)| \neq 0$, de donde $D\Phi(\alpha(x))$ es invertible. Usando lo anterior y que $\Phi \in C^1(\mathbb{R}^2)$, por el teorema de la función inversa, existe $R : U \rightarrow V$ tal que $R \in C^1(U)$, $\Phi(x) = \Phi(\alpha(x)) \in U$, $\alpha(x) \in V$ y $(R \circ \Phi)(\alpha(x)) = \alpha(x)$. Así que

$$\alpha(x) = (R \circ \Phi)(\alpha(x)) = (R \circ \Phi)(x) \Rightarrow \alpha \in C^1(T_0).$$

Usando el hecho de que $\Phi(\alpha(x)) = \Phi(x)$, obtenemos

$$|J_{\alpha}(x)| = |J_{\Phi}(x)| |J_{\Phi}(\alpha(x))|^{-1}. \quad (4)$$

Fácilmente vemos que $T_0 \cap \alpha(T_0) = \emptyset$, pues si $x \in T_0 \cap \alpha(T_0) \Rightarrow \theta(x) \leq \frac{\pi}{2}$ lo cual es una contradicción. Notemos que si

$$x \notin T \Rightarrow 0 \leq \theta(x) \leq \frac{\pi}{2} \Rightarrow \cos \theta(x) \geq 0 \Rightarrow \nabla f(x) \nabla g(x) \geq 0.$$

También α es 1 - 1 en T_0 , dado que si

$$\begin{aligned} \alpha(x) &= \alpha(y) \\ \Rightarrow f(\alpha(x)) &= f(\alpha(y)) = f(x) = f(y) \text{ y } g(\alpha(x)) = g(x) = g(\alpha(y)) = g(y) \\ \Rightarrow \Phi(x) &= \Phi(y) \\ \Rightarrow x &= y. \end{aligned}$$

Hacemos notar los siguientes hechos:

- ϕ es medible, pues si $a \in \mathbb{R}$

$$\{x : \phi(|x|) \leq a\} = \{x : |x| \geq \phi^{-1}(a)\} \in \mathfrak{M}(m), \text{ donde } \phi^{-1}(a) = \inf\{|x| : \phi(|x|) \leq a\}.$$

- $\alpha(T_0)$ es medible, pues T_0 es abierto, $\alpha \in C^1(T_0)$ y $|J_{\alpha}(x)| \neq 0$ si $x \in T_0$.

³ $B_i \cap T_0^c \subset H_i(\Phi(B_i \cap T_0^c))$ con $i = 1, 2$.

Ahora, usando (4), (3), (2) y (1)

$$\begin{aligned}
\int_{\alpha(T_0)} \nabla f(x) \cdot \nabla g(x) \phi(x) d^2x &= \int_{T_0} \nabla f(\alpha(x)) \cdot \nabla g(\alpha(x)) \phi(\alpha(x)) |J_\alpha(x)| d^2x \\
&= \int_{T_0} \nabla f(\alpha(x)) \cdot \nabla g(\alpha(x)) \phi(\alpha(x)) |J_\Phi(x)| |J_\Phi(\alpha(x))|^{-1} d^2x \\
&= \int_{T_0} \cot(\theta(x)) \phi(\alpha(x)) |J_\Phi(x)| d^2x \\
&\geq - \int_{T_0} \cot(\theta(x)) \phi(x) |J_\Phi(x)| d^2x \\
&= - \int_{T_0} \nabla f(x) \cdot \nabla g(x) \phi(x) d^2x. \tag{5}
\end{aligned}$$

Usando que $T^c = \alpha(T_0) \cup (T^c - T_0)$ ($\theta(\alpha(T_0)) \leq \frac{\pi}{2}$), $T_0 \cap (T - T_0) \cap T^c = \emptyset$ y (5)

$$\begin{aligned}
\int_{\mathbb{R}^2} \nabla f(x) \cdot \nabla g(x) \phi(x) d^2x &= \int_{T_0} \nabla f(x) \cdot \nabla g(x) \phi(x) d^2x + \int_{T-T_0} \nabla f(x) \cdot \nabla g(x) \phi(x) d^2x \\
&\quad + \int_{T^c} \nabla f(x) \cdot \nabla g(x) \phi(x) d^2x \\
&= \int_{T_0} \nabla f(x) \cdot \nabla g(x) \phi(x) d^2x + \int_{\alpha(T_0)} \nabla f(x) \cdot \nabla g(x) \phi(x) d^2x \\
&\quad + \int_{T^c - \alpha(T_0)} \nabla f(x) \cdot \nabla g(x) \phi(x) d^2x \\
&\geq \int_{T_0} \nabla f(x) \cdot \nabla g(x) \phi(x) d^2x + \int_{\alpha(T_0)} \nabla f(x) \cdot \nabla g(x) \phi(x) d^2x \\
&\geq 0.
\end{aligned}$$

Lo cual termina la demostración⁴.

■

Ahora demostraremos otro teorema, sobre la distribución gaussiana multivariante, que hace uso del teorema anterior. El Teorema 2 nos permitirá demostrar la conjetura de correlación gaussiana en el plano.

⁴Si $T = \emptyset$ el teorema es trivialmente cierto.

Antes de enunciar el próximo teorema, consideremos el vector aleatorio $(X_1, \dots, X_n, Y_1, \dots, Y_n)$ distribuido con la densidad normal con media cero y matriz de covarianza Σ donde $\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{12}^T & \Sigma_{22} \end{pmatrix}$; donde $\Sigma_{11} = (\mathbb{E}X_i X_j)$, $\Sigma_{22} = (\mathbb{E}Y_i Y_j)$ y $\Sigma_{12} = (\mathbb{E}X_i Y_j)$. Sean $X = (X_1, \dots, X_n)$ y $Y = (Y_1, \dots, Y_n)$. Para conjuntos convexos y simétricos $A, B \subset \mathbb{R}^n$ definamos la probabilidad $p(\Sigma) = \mathbb{P}\{X \in A; Y \in B\}$. Para $\lambda, 0 \leq \lambda \leq 1$, la matriz $\Sigma_\lambda = \begin{pmatrix} \Sigma_{11} & \lambda \Sigma_{12} \\ \lambda \Sigma_{12}^T & \Sigma_{22} \end{pmatrix}$ es una matriz definida positiva. Dado que

- Si $\lambda = 0$,

$$(x, y) \begin{pmatrix} \Sigma_{11} & 0 \\ 0 & \Sigma_{22} \end{pmatrix} (x, y)^T = x \Sigma_{11} x^T + y \Sigma_{22} y^T > 0 \text{ si } (x, y) \neq 0.$$

- Si $0 < \lambda < 1$,

$$\begin{aligned} (x, y) \Sigma_\lambda (x, y)^T &= x \Sigma_{11} x^T + \lambda y \Sigma_{12}^T x^T + \lambda x \Sigma_{12} y^T + y \Sigma_{22} y^T \\ &> x \Sigma_{11} x^T - \lambda x \Sigma_{11} x^T - \lambda y \Sigma_{22} y^T + y \Sigma_{22} y^T \\ &> 0 \text{ si } (x, y) \neq 0. \end{aligned}$$

Por ende, podemos considerar la probabilidad $p(\Sigma_\lambda)$ como función de λ , con $0 \leq \lambda \leq 1$.

Teorema 2. Con la notación de arriba, si $\text{rank}(\Sigma_{12}) \leq 2$ entonces $p(\Sigma_\lambda)$ es una función creciente de λ , con $0 \leq \lambda \leq 1$.

Demostración. Sin pérdida de generalidad podemos suponer que $\text{rank} \Sigma_{11} = \text{rank} \Sigma_{22} = n$. Introduciendo variables canónicas (ver [1, Cap. 12]) también podemos asumir que $\Sigma_{11} = \Sigma_{22} = I$ y que Σ_{12} es diagonal con entradas no negativas. Por tanto, sea $\Sigma_{12} = \text{diag}(\lambda_1, \dots, \lambda_n)$ con $1 \geq \lambda_1 \geq \lambda_2 \geq \dots \geq 0$. Como $\text{rank} \Sigma_{12} \leq 2$ entonces $\lambda_3 = \lambda_4 = \dots = \lambda_n = 0$. Observemos primero que $1_A(x)$ y $1_B(y)$ son funciones pares y log-cóncavas. Son pares dado que A y B son simétricos. La log-cóncavidad es trivial revisarla. También tenemos que $1_A \in \mathfrak{L}(\mathbb{R}^n)$, respecto a la medida gaussiana. Definamos

$$\phi(x) = \begin{cases} 0 & \text{si } |x| \geq 1 \\ \frac{1}{c} \exp\left(-\frac{1}{1-|x|^2}\right) & \text{si } |x| < 1 \end{cases}$$

donde $c = \int_{B(0,1)} \exp\left(-\frac{1}{1-|x|^2}\right) dx$.

Sea $\phi_a(x) = a^{-n} \phi\left(\frac{x}{a}\right)$ con $a \in (0, \infty)$. Enlistemos los siguientes hechos (ver [3, Cap. 7.C, p. 176])

- $\phi_a \in C^\infty$,

- $\phi_a \geq 0$,
- $\phi_a > 0 \iff |x| < a$,
- $\int_{\mathbb{R}^n} \phi_a(x) dx = 1$,
- ϕ_a es log-cóncava.

Luego, $f_a = 1_A * \phi_a$ es log-cóncava (ver [2]), también $|f_a(x)| \leq 1$ y $f_a \rightarrow 1_A$ en $\mathfrak{L}(\mathbb{R}^n)$ cuando $a \rightarrow 0$; ver [3, Chapter 12.C, p. 287]. Resultados análogos, se obtienen para las funciones $g_a = 1_B * \phi_a$. Usando que $|f_a(x)g_a(y)| \leq 1$ con $(x, y) \in \mathbb{R}^{2n}$ y el Teorema de Convergencia Dominada con respecto a la medida gaussiana en \mathbb{R}^{2n}

$$\mathbb{E}(1_A(X)1_B(Y)) = \lim_{a \rightarrow 0} \mathbb{E}(f_a(X)g_a(Y))$$

Por ende, es suficiente probar que la esperanza $\mathbb{E}(\lambda) = \mathbb{E}f_a(X)g_a(Y)$ es una función creciente de λ para un a fijo.

Demostremos que $R = (X_3, \dots, X_n)$ y $S = (X_1, X_2, Y_1, Y_2)$ son independientes. Como $\Sigma_{12} = \text{diag}(\lambda_1, \lambda_2, 0, \dots, 0)$ entonces $\mathbb{E}(X_j Y_i) = 0$ con $j \geq 3$, $1 \leq i \leq 2$. Usando que $\Sigma_{11} = I$ entonces $\mathbb{E}(X_j X_i) = 0$ con $j \geq 3$, $1 \leq i \leq 2$, de lo anterior que sean independientes. Además,

$$\mathbb{E}(f_a(X)|X_1, X_2, Y_1, Y_2) = \int_{\mathbb{R}^{n-2}} f_a(x) f_{X_3, \dots, X_n}(x_3, \dots, x_n) dx_3 \cdots dx_n = f(X_1, X_2).$$

$$\text{donde } f(x_1, x_2) = (2\pi)^{-(n-2)/2} \int f_a(x) \exp(-\frac{1}{2} \sum_3^n x_j^2) dx_3 \cdots dx_n.$$

Asimismo,

$$\mathbb{E}(g_a(Y)|X_1, X_2, Y_1, Y_2) = \int_{\mathbb{R}^{n-2}} g_a(y) f_{Y_3, \dots, Y_n}(y_3, \dots, y_n) dy_3 \cdots dy_n = g(Y_1, Y_2)$$

$$\text{donde } g(y_1, y_2) = (2\pi)^{-(n-2)/2} \int g_a(y) \exp(-\frac{1}{2} \sum_3^n y_j^2) dy_3 \cdots dy_n.$$

Sean

$$\begin{aligned} \psi(X_1, X_2, Y_1, Y_2) &= \mathbb{E}(f_a(X)|X_1, X_2, Y_1, Y_2), \\ \Upsilon(X_1, X_2, Y_1, Y_2) &= \mathbb{E}(g_a(Y)|X_1, X_2, Y_1, Y_2). \end{aligned}$$

Así que,

$$\begin{aligned} \mathbb{E}(\psi(X_1, X_2, Y_1, Y_2)\Upsilon(X_1, X_2, Y_1, Y_2)) &= \mathbb{E}(f_a(X)\Upsilon(X_1, X_2, Y_1, Y_2)) \\ &= \mathbb{E}(f_a(X)g_a(Y)) \\ &= \mathbb{E}(f(X_1, X_2)g(Y_1, Y_2)) \end{aligned}$$

en donde $f(x_1, x_2)$ y $g(y_1, y_2)$ son log-cóncavas, dado que (ver Propiedad B7 en Apéndice)

$$f_a(x) \exp(-\frac{1}{2} \sum_3^n x_j^2) \text{ y } g_a(y) \exp(-\frac{1}{2} \sum_3^n y_j^2) \text{ son log-cóncavas.}$$

Haciendo cálculos se ve fácilmente que la densidad condicional de (Y_1, Y_2) dado que $(X_1, X_2) = (x_1, x_2)$ es

$$p(\lambda, x_1, x_2; y_1, y_2) = \prod_{j=1}^2 [2\pi(1 - (\lambda\lambda_j)^2)]^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} \frac{(\lambda\lambda_j x_j - y_j)^2}{1 - (\lambda\lambda_j)^2} \right\}.$$

Sean

$$\begin{aligned} n(x_1, x_2) &= (2\pi)^{-1} \exp \left(-\frac{x_1^2 + x_2^2}{2} \right), \\ g(\lambda, x_1, x_2) &= \int p(\lambda, x_1, x_2; y_1, y_2) g(y_1, y_2) dy_1 dy_2. \end{aligned}$$

Usando el hecho de que $f(x_1, x_2)$, $g(y_1, y_2)$, $f_{X_1, X_2, Y_1, Y_2}(x_1, x_2, y_1, y_2) \geq 0$ y las funciones son medibles, por el Teorema de Fubini

$$\begin{aligned} \mathbb{E}(\lambda) &= \int f(x_1, x_2) g(y_1, y_2) f_{X_1, X_2, Y_1, Y_2}(x_1, x_2, y_1, y_2) dx_1 dx_2 dy_1 dy_2 \\ &= \int f(x_1, x_2) g(y_1, y_2) p(\lambda, x_1, x_2; y_1, y_2) n(x_1, x_2) dx_1 dx_2 dy_1 dy_2 \\ &= \int f(x_1, x_2) g(\lambda, x_1, x_2) n(x_1, x_2) dx_1 dx_2. \end{aligned}$$

Sea $\lambda = e^t$, $-\infty < t \leq 0$. Haciendo cálculos se verifica que

$$\frac{\partial}{\partial t} g(e^t, x_1, x_2) = \left(\sum_{j=1}^2 \frac{\partial^2}{\partial x_j^2} - x_j \frac{\partial}{\partial x_j} \right) g(e^t, x_1, x_2)$$

y realizando una integración por partes

$$\frac{\partial}{\partial t} \mathbb{E}(\mathbf{e}^t) = \int \nabla f(x_1, x_2) \nabla g(e^t, x_1, x_2) n(x_1, x_2) dx_1 dx_2.$$

Como $p(\lambda, x_1, x_2; y_1, y_2)$ es log-cóncava, entonces $g(e^t, x_1, x_2)$ es log-cóncava. Como n es radialmente decreciente, por el Teorema 1 $\frac{\partial}{\partial t} \mathbb{E}(\mathbf{e}^t) \geq 0$ (las funciones log-cóncavas y pares son casi-cóncavas) y se sigue el resultado. ■

Con los teoremas anteriores estamos listos para demostrar la conjetura de correlación gaussiana en el plano.

Teorema 3. Sea $\gamma(x) = \frac{1}{2\pi} \exp(-\frac{|x|^2}{2})$ con $x \in \mathbb{R}^2$. Si A y B son conjuntos simétricos y convexos entonces

$$\int_{A \cap B} \gamma(x) d^2x \geq \left(\int_A \gamma(x) d^2x \right) \left(\int_B \gamma(x) d^2x \right)$$

Demostración. Es claro que $\mathbb{P}(X \in A)$ y $\mathbb{P}(Y \in B)$ no dependen de λ . Además, $\Sigma_\lambda \rightarrow \begin{pmatrix} \Sigma_{11} & 0 \\ 0 & \Sigma_{22} \end{pmatrix}$ cuando $\lambda \rightarrow 0$; de aquí que cuando $\lambda \rightarrow 0$ X y Y son independientes, por ende $\lim_{\lambda \rightarrow 0} \mathbb{P}(X \in A; Y \in B) = \mathbb{P}(X \in A)\mathbb{P}(Y \in B)$. Usando el Teorema 2 con $n = 2$ y $\Sigma_{11} = \Sigma_{12} = \Sigma_{22} = I$ obtenemos el resultado, a saber $\mathbb{P}(X \in A; Y \in B) \geq \mathbb{P}(X \in A)\mathbb{P}(Y \in B)$ y se concluye el teorema. ■

Apéndice: Funciones casi-cóncavas y log-cóncavas.

Se enlistan las principales propiedades de estas funciones; ver [2] y [6].

A) Funciones casi-cóncavas.

1. Una función $h(x)$ definida en \mathbb{R}^n se dice casi-cóncava si para cualesquiera $x_1, x_2 \in \mathbb{R}^n$ y $0 < \lambda < 1$,

$$h(\lambda x_1 + (1 - \lambda)x_2) \geq \min\{h(x_1), h(x_2)\}.$$

2. Si f es cóncava entonces f es casi-cóncava.
3. Supongamos que existe la segunda derivada de f y que es continua. Sea $T_{\nabla f(x)} = \{v : \nabla f(x) \cdot v = 0\}$. Entonces $D^2 f(x)$ es semi-definida negativa en $T_{\nabla f(x)}$ si y sólo si f es casi-cóncava.
4. Si $f : \mathbb{R} \rightarrow \mathbb{R}$ es monótona entonces es casi-cóncava.
5. Si $f \in C^2(\mathbb{R}^2)$, f es casi-cóncava si $2f_x f_y f_{xy} - f_x^2 f_{yy} - f_y^2 f_{xx} \geq 0$.

B) Funciones log-cóncavas.

1. Una función $f(x)$ no-negativa definida en \mathbb{R}^n se dice log-cóncava si a) el conjunto $E_0^f = \{x : f(x) > 0\}$ es abierto; b) $f(x)$ es continua en E_0^f ; c) para cualesquiera $x_1, x_2 \in \mathbb{R}^n$

$$f((x_1 + x_2)/2) \geq \sqrt{f(x_1)f(x_2)}.$$

2. Si f y g son log-cóncavas entonces fg también lo es.
3. Sea $E \subset \mathbb{R}^n$. E es convexo y abierto si y sólo si χ_E es log-cóncava.
4. Si f es log-cóncava para cualesquiera $x_1, x_2 \in \mathbb{R}^n$ y $0 < \lambda < 1$,

$$f(\lambda x_1 + (1 - \lambda)x_2) \geq [f(x_1)]^\lambda [f(x_2)]^{1-\lambda}.$$

5. Si f y g son log-cóncavas entonces $(f * g)(x) = \int_{\mathbb{R}^n} f(y)g(x - y)dy$ es log-cóncava.
6. Si f es log-cóncava entonces es casi-cóncava.
7. Si $H(x_1, \dots, x_k, y_1, \dots, y_l)$ es log-cóncava entonces

$$h(x_1, \dots, x_k) = \int H(x_1, \dots, x_k, y_1, \dots, y_l) dy_1 \cdots dy_l \text{ es log-cóncava.}$$

Referencias

- [1] ANDERSON, T.W. (1958). *An Introduction to Multivariate Statistical Analysis*. Wiley, New York.
- [2] DAVIDOVIC, Ju. S., KORENBLIUM, B..J. and HACET, B.I. (1969). A property of logarithmically concave functions. *Soviet Math. Dokl.* **10** 477-480.
- [3] JONES, Frank. (1993). *Lebesgue Integration on Euclidean Space*. Jones and Bartlett Publishers.
- [4] MILNOR, J. (1965). *Topology for the Differentiable Viewpoint*. Univ. Press of Virginia, Charlottesville.
- [5] PITT, L. (1977). *A Gaussian Correlation Inequality for Symmetric Convex Sets*. University of Virginia. *The Annals of Probability*, **5**, 470-474.
- [6] PRÉKOPA, A. (1973). On logarithmic concave measures and functions. *Acta Sci. Math.* **34** 335-343.