

GENERADOR DE PROCESOS NORMALES MULTIVARIADOS

RESUMEN

En este artículo se presenta un método para la generación de variables aleatorias que provienen de un proceso normal multivariado, de alta aplicación en la simulación de sistemas dinámicos.

PALABRAS CLAVES: Proceso normal multivariado.

ABSTRACT

In this article a method for the generation of variates appears that come from a multivariate normal process, of high application in the simulation of dynamic systems.

KEYWORDS: multivariate normal process.

ALVARO TREJOS CARPINTERO

Profesor Auxiliar
Facultad de Ingeniería Industrial
Universidad Tecnológica de Pereira
alvarot@utp.edu.co

PATRICIA CARVAJAL OLAYA

Estudiante Tesista
Maestría en Investigación Operativa y Estadística
Universidad Tecnológica de Pereira
pacarva@utp.edu.co

MAURICIO BARRERA

Estudiante décimo semestre
Facultad de Ingeniería Industrial
Universidad Tecnológica de Pereira
voneumann@hotmail.com

1. INTRODUCCIÓN

Para implementar algunos modelos de simulación es necesario generar vectores aleatorios $\bar{X} = (x_1, x_2, \dots, x_n)$ los cuales provienen de una distribución de probabilidad multivariada, donde los componentes individuales del vector aleatorio podrían no ser independientes, ello implica que las correlaciones, r_{ij} , entre las variables X_i , X_j del vector aleatorio deben ser especificadas por el modelador.

En este artículo, la atención se centra en el estudio de la generación de vectores aleatorios normales dada la importancia que la distribución tiene para describir muchos fenómenos naturales y sociales. Existe además, una justificación matemática para el empleo de la distribución normal, proporcionada por el teorema central del límite.

Aunque las formulas para la generación de variables aleatorias normales son muy conocidas en la actualidad, la deducción de ellas no es tan común por lo que se realizará un desarrollo un tanto riguroso de estas, lo que contribuirá a una mejor comprensión del método y de los supuestos bajo los cuales funcionan.

2. DISTRIBUCIÓN NORMAL MULTIVARIADA

Primero se recuerda la distribución normal univariada estándar y a partir de allí se definirá la distribución normal multivariada.

Definición 2.1

Una variable aleatoria Z tiene una distribución normal estándar univariada si y solo si la función de densidad de probabilidad de Z esta dada por:

$$n(Z;0,1) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right) \quad -\infty < z < \infty$$

donde $n(Z;0,1)$ significa que la variable Z está distribuida normalmente con media cero y varianza uno, y \exp representa la función exponencial. [4]

Definición 2.2

La variable aleatoria X tiene una distribución normal univariada con media \mathbf{m} y varianza \mathbf{S}^2 si y solo si la función de densidad de probabilidad de x esta dada por:

$$n(x; \mathbf{m}, \mathbf{S}^2) = \frac{1}{\mathbf{S} \sqrt{2\pi}} e^{-(x-\mathbf{m})^2 / 2\mathbf{S}^2} \quad -\infty < x < \infty$$

Ahora en el conjunto de variables Z_i , donde cada i -ésima Z está distribuida $n(Z;0,1)$. Se puede decir, que si las Z_i variables son mutuamente independientes, la distribución conjunta aleatoria esta dada por:

$$\prod_{i=1}^k \frac{1}{\sqrt{2\pi}} e^{-\frac{z_i^2}{2}} = \frac{1}{(2\pi)^{k/2}} e^{-\frac{1}{2} \sum_{i=1}^k z_i^2} \quad (1)$$

$$-\infty < z_i < \infty \quad i = 1, 2, \dots, k$$

Siendo Z_i el i -ésimo elemento del vector \vec{Z} (kx1) se puede escribir la ecuación 1 en forma más compacta así:

$$n(\vec{Z} : 0, I) = \frac{1}{(2\pi)^{k/2}} e^{-1/2\vec{Z}^T \vec{Z}} \quad 2.$$

La ecuación 2, puede ser vista como una generalización multidimensional de la función de densidad de probabilidad de la normal univariada estándar. [4]

Definición 2.3

Sea \vec{X} un vector aleatorio con p componentes y con vector de medias \vec{m} y matriz de covarianzas Σ , donde el rango de $\Sigma = p > 0$. Sea $\Gamma^T \Gamma = \Sigma$, esto es, $\Gamma^T \Gamma$ es una factorización no única de Σ , donde Γ^T es una matriz de p x p. \vec{X} tiene una distribución normal p-variada si y solo si la distribución de \vec{X} tiene la misma distribución del vector aleatorio $\Gamma^T \vec{Z} + \vec{m}$, donde la distribución del vector aleatorio \vec{Z} es $n(\vec{Z} : 0, I)$, es decir, \vec{Z} tiene una distribución normal estándar p-variada.

De la definición anterior se tiene que $E(\vec{Z}) = \vec{0}$ y $cov(\vec{Z}) = I$, entonces para que \vec{X} tenga la misma distribución de $\Gamma^T \vec{Z} + \vec{m}$, la esperanza de $\Gamma^T \vec{Z} + \vec{m}$ debe ser igual a \vec{m} al igual la covarianza de $\Gamma^T \vec{Z} + \vec{m}$ debe ser igual a Σ .

Prueba:

$$\begin{aligned} E(\Gamma^T \vec{Z} + \vec{m}) &= \Gamma^T E(\vec{Z}) + \vec{m} = \vec{m} = E(\vec{X}) \\ \text{debida que } \vec{Z} &\approx n(0, I) \\ COV(\Gamma^T \vec{Z} + \vec{m}) &= E\left[(\Gamma^T \vec{Z} + \vec{m} - E(\Gamma^T \vec{Z} + \vec{m})) (\Gamma^T \vec{Z} + \vec{m} - E(\Gamma^T \vec{Z} + \vec{m}))^T \right] \\ &= E\left[(\Gamma^T \vec{Z} + \vec{m} - \vec{m}) (\Gamma^T \vec{Z} + \vec{m} - \vec{m})^T \right] \\ &= E\left[(\Gamma^T \vec{Z}) (\Gamma^T \vec{Z})^T \right] = \Gamma^T E(\vec{Z} \vec{Z}^T) \Gamma = \Gamma^T \Gamma = \Sigma \end{aligned}$$

Teorema 2.1

Si \vec{X} esta distribuida $n(\vec{X} : \vec{m}, \Sigma)$ donde \vec{X} es un vector (p x 1) y si el rango de Σ es p, entonces \vec{X} tiene una función de densidad de probabilidad y está dada por:

$$n(\vec{X} : \vec{m}, \Sigma) = \frac{1}{(2\pi)^{p/2} |\Sigma|^{1/2}} e^{-1/2(\vec{X} - \vec{m})^T \Sigma^{-1} (\vec{X} - \vec{m})} \quad 3.$$

$$X \in Ep$$

Prueba:

La transformación que pasa de (Z_1, \dots, Z_p) a (X_1, \dots, X_p) es lineal.

$\vec{x} = \Gamma^T \vec{Z} + \vec{m}$ Por lo cual el jacobiano es $J = (\Gamma^T)^{-1}$ y debido a que $\Sigma = \Gamma^T \Gamma$ (definición 2.3).

$\Sigma^{-1} = \Gamma^{-1} (\Gamma^T)^{-1}$ de donde $|\Sigma^{-1}|^{1/2} = |(\Gamma^T)^{-1}|$ y además

$$\begin{aligned} \vec{Z}^T \vec{Z} &= (\Gamma^{T^{-1}} (\vec{X} - \vec{m}))^T (\Gamma^{T^{-1}} (\vec{X} - \vec{m})) \\ &= (\vec{X} - \vec{m})^T (\Gamma^{T^{-1}})^T \Gamma^{T^{-1}} (\vec{X} - \vec{m}) \\ &= (\vec{X} - \vec{m})^T (\Gamma^T \Gamma)^{-1} (\vec{X} - \vec{m}) \\ &= (\vec{X} - \vec{m})^T \Sigma^{-1} (\vec{X} - \vec{m}) \end{aligned}$$

Quedando así demostrado el teorema (2.1)

Antes de generar los vectores aleatorios es necesario abordar el tema de la factorización de la matriz Σ (covarianzas) por lo mencionado en la definición 2.3.

Aunque existen diversas factorizaciones de esta matriz la técnica de descomposición de Cholesky proporciona un algoritmo para descomponer la matriz Σ en dos matrices triangulares superior e inferior, y además garantiza que estas dos matrices sean únicas.

3. TÉCNICA DE DESCOMPOSICIÓN DE CHOLESKY [1]

A continuación se justifica la descomposición de la matriz Σ , en matrices triangulares superior e inferior debido a sus características especiales.

Teorema 3.1

Sea $S_{p \times p}$ una matriz definida positiva. Existe una matriz triangular superior T de rango P, con $T_{ii} > 0$ para $i=1,2,\dots,p$, tal que $S = T^T T$. Esta descomposición es única.

Prueba:

La prueba se hará por inducción matemática. Primero, sea $P=1$; como S^* es definida positiva entonces podemos escribir $S^* = [s_{11}^2]$. Por lo tanto $t_{11} = |s_{11}|$, con lo cual $T^* = [t_{11}]$ es una matriz triangular superior con $t_{11} > 0$ y además, t_{11} es único.

Supóngase el teorema verdadero para $P=K$. Mostraremos que es verdadero para $P=K+1$. Si el teorema es verdadero para $P=K$, entonces para cualquier matriz definida positiva de $k \times k$, S^*_{11} , existe una única matriz triangular superior T^*_{11} , con $t_{ii} > 0$ para $i=1,2,\dots,K$, tal que $S^*_{11} = T^{*T}_{11} T^*_{11}$. Sea S^* cualquier matriz definida positiva de tamaño $(k+1) \times (k+1)$, entonces podemos escribir S^* como:

$$S^* = \begin{bmatrix} S^*_{11} & \vec{S}_{12} \\ \vec{S}_{21} & S_{22} \end{bmatrix}$$

Donde S^*_{11} es una matriz definida positiva de $k \times k$ componentes pero por la hipótesis de inducción $S^*_{11} = T^{*T}_{11} T^*_{11}$, donde T^*_{11} es una matriz triangular superior con elementos positivos en la diagonal. ($\vec{S}_{12} = \vec{S}_{21}$ ya que S es simétrica) y por tanto queda:

$$S^* = \begin{bmatrix} S^*_{11} & \vec{S}_{12} \\ \vec{S}_{21} & S_{22} \end{bmatrix} = \begin{bmatrix} T^{*T}_{11} T^*_{11} & \vec{S}_{12} \\ \vec{S}_{21} & S_{22} \end{bmatrix} =$$

Y por inspección

$$\begin{bmatrix} T^{*T}_{11} T^*_{11} & \vec{S}_{12} \\ \vec{S}_{21} & S_{22} \end{bmatrix} = \begin{bmatrix} T^{*T}_{11} & \vec{O} \\ \vec{S}_{12} T^{*-1}_{11} & b \end{bmatrix} \begin{bmatrix} T^*_{11} & (T^*_{11})^{-1} \vec{S}_{12} \\ \vec{O} & b \end{bmatrix} = T^T T$$

Donde $b = (S_{22} - \vec{S}_{21} S^{-1}_{11} \vec{S}_{12})^{1/2}$, si podemos mostrar que $S_{22} - \vec{S}_{21} S^{-1}_{11} \vec{S}_{12} > 0$, ya que b es real, entonces T es triangular superior con elementos en la diagonal positivos y además T es único.

Para mostrar esto, se observa que el determinante de S^* , $|S^*|$, es igual a $|S^*_{11}| |S_{22} - \vec{S}_{21} S^{-1}_{11} \vec{S}_{12}|$. Pero

S^* y S^*_{11} son definidas positivas, entonces $|S^*| > 0, |S^*_{11}| > 0$.

Esto significa que: $|S_{22} - \vec{S}_{21} S^{-1}_{11} \vec{S}_{12}| > 0$, pero $S_{22} - \vec{S}_{21} S^{-1}_{11} \vec{S}_{12}$ es una escalar y por lo tanto igual a su determinante. Esto implica que $S_{22} - \vec{S}_{21} S^{-1}_{11} \vec{S}_{12} > 0$

Entonces Σ se puede descomponer en 2 matrices $\Gamma^T \Gamma$ así:

$$\begin{bmatrix} s_{11} & s_{12} & \dots & s_{1p} \\ s_{21} & s_{22} & \dots & s_{2p} \\ \dots & \dots & \dots & \dots \\ s_{p1} & s_{p2} & \dots & s_{pp} \end{bmatrix} = \begin{bmatrix} a_{11} & O & O \\ a_{12} & a_{22} & \dots & O \\ \dots & \dots & \dots & \dots \\ a_{1p} & a_{2p} & \dots & a_{pp} \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1p} \\ O & a_{22} & & a_{2p} \\ \dots & \dots & \dots & \dots \\ O & O & \dots & a_{pp} \end{bmatrix}$$

donde

$$\begin{aligned} s_{11} &= a_{11}^2 \\ s_{12} &= a_{11} a_{12} \\ \dots & \dots \\ s_{1p} &= a_{11} a_{1p} \\ s_{22} &= a_{12}^2 + a_{22}^2 \\ \dots & \dots \\ s_{pp} &= a_{1p}^2 + a_{2p}^2 + \dots + a_{pp}^2 \end{aligned} \quad [1]$$

4. GENERADOR DE PROCESOS MULTIVARIADOS

Para generar un vector aleatorio con componentes independientes, simplemente se crean n variables aleatorias independientes y se multiplican, es decir, sea $\vec{x} = (x_1, x_2, \dots, x_n)^T$ cuyas componentes son independientes entre sí, entonces la función de distribución simultánea de \vec{x} sería:

$F(\vec{X}) = \prod_{i=1}^n Fi(X_i)$ donde $Fi(X_i)$ representa la función de distribución marginal de X_i . Por tanto para generar el vector \vec{X} , se genera la componente X_i de esa distribución Fi independientemente.

Ejemplo: Para generar una normal multivariada de K variables aleatorias independientes Z_i , cada una es

$n(Z:0,1)$, y si las variables son mutuamente independientes. Su distribución conjunta sería:

$$F(Z_1, Z_2, \dots, Z_k) = \frac{1}{(2\pi)^{k/2}} e^{-\frac{1}{2} \sum_{i=1}^k Z_i^2} \quad (4)$$

y simplemente se generarían K variables aleatorias independientes normales univariadas.

El problema puede volverse más complejo cuando las variables son dependientes. En tal caso la función de distribución conjunta puede ser expresada por:

$$F(\vec{X}) = F_1(X_1)F_2(X_2 | X_1) \dots F_n(X_n | X_1, \dots, X_{n-1})$$

Entonces para generar el vector \vec{X} , se pueden utilizar n números aleatorios independientes u_1, u_2, \dots, u_n y resolver el siguiente conjunto recursivo de ecuaciones:

$$F_1(X_1) = U_1$$

$$F_2(X_2 | X_1) = U_2$$

.

.

$$F_n(X_n | X_1, X_2, \dots, X_{n-1}) = U_n$$

Ejemplo: Para generar una normal multivariada de p variables aleatorias X_i dependientes entre si, su distribución conjunta sería:

$$F(X_1, X_2, \dots, X_p) = \frac{1}{(2\pi)^{p/2} |\Sigma|^{1/2}} e^{-1/2(\vec{X}-\vec{m})^T \Sigma^{-1}(\vec{X}-\vec{m})}$$

y lo que se podría hacer es utilizar el hecho de que, \vec{X} tiene una distribución normal p-variada si y solo si \vec{X} tiene la misma distribución del vector aleatorio $\Gamma^T \vec{Z} + \vec{m}$ donde \vec{Z} está distribuida normal estándar y $\Gamma^T \Gamma$ es una factorización de Σ (definición 2.3), y aunque esta factorización no es única puede utilizarse la descomposición de Cholesky.

En otras palabras se ha definido un vector aleatorio normal multivariado en términos de funciones lineales de variables aleatorias normales univariadas estándar. Así que si \vec{X} tiene una distribución normal p-variada con media \vec{m} y covarianza Σ ($\Gamma^T \Gamma = \Sigma$), entonces se puede usar $\vec{X} = \Gamma^T \vec{Z} + \vec{m}$ para transformar \vec{Z} , la cual

está distribuida normal estándar, en \vec{X} con distribución normal con media \vec{m} y covarianza Σ . Quedando así:

$$X_1 = a_{11}Z_1 + U_1$$

$$X_2 = a_{21}Z_1 + a_{22}Z_2 + U_2$$

.

.

$$X_p = a_{p1}Z_1 + a_{p2}Z_2 + \dots + a_{pp}Z_p + U_p$$

Lo cual articula también con el conjunto recursivo de ecuaciones donde los a_{ij} son los términos de la matriz de la factorización de Cholesky.

El procedimiento para generar procesos normales multivariados es el siguiente:

- 1) Proporcionar \vec{m} y $\Sigma = [s_{ij}]$
- 2) Calcular la matriz Γ^T a partir de Σ utilizando la técnica de descomposición de Cholesky
- 3) Generar $\vec{Z} = (Z_1, Z_2, \dots, Z_p)^T$ variables aleatorias estándar mutuamente independientes
- 4) Calcular $\vec{X} = \Gamma^T \vec{Z} + \vec{m}$

Ahora el problema en consideración, sería como generar variables aleatorias normales estándar mutuamente independientes y para ello puede utilizarse el método de la transformación inversa para generar variables aleatorias continuas así:

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \quad -\infty < X < \infty \quad X \approx n(0,1)$$

Y lo que puede hacerse es calcular su función de distribución acumulada e igualarla a una variable aleatoria uniforme $U(0,1)$ así:

$$F(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = U \quad (5)$$

Y despejar a X en términos de U, lo cual no se puede hacer directamente debido a la complejidad del cálculo de F(x), por tanto se utilizarán las coordenadas polares para generar 2 variables aleatorias normales estándar independientes así:

$$F(x, x) = \left(\frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \right) \left(\frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} \right) \quad (6)$$

La que es una distribución normal bivariada estándar con variables independientes entre si, arreglando un poco la ecuación 6 nos daría

$$F(x, y) = \frac{1}{2\Pi} e^{-\frac{x^2+y^2}{2}} \quad (7)$$

Y utilizando la siguiente transformación

$$x^2 + y^2 = R, \quad y = \frac{y}{x} \tan \mathbf{f} \quad \text{o de esta manera:}$$

a) $X = \sqrt{R} \cos \mathbf{f}$ y calculando el determinante

b) $Y = \sqrt{R} \text{sen} \mathbf{f}$
jacobiano

$$|J| = \begin{vmatrix} \frac{\partial x}{\partial R} & \frac{\partial x}{\partial \mathbf{f}} \\ \frac{\partial y}{\partial R} & \frac{\partial y}{\partial \mathbf{f}} \end{vmatrix} = \begin{vmatrix} \frac{\cos \mathbf{f}}{2\sqrt{R}} & -\sqrt{R} \text{sen} \mathbf{f} \\ \frac{\text{sen} \mathbf{f}}{2\sqrt{R}} & \sqrt{R} \cos \mathbf{f} \end{vmatrix} = \frac{1}{2}$$

Reemplazando en 7 e integrando

$$F(R, \mathbf{f}) = \int_0^{\mathbf{f}} \int_0^R \frac{1}{2\Pi} \frac{1}{2} e^{-R/2} dR d\mathbf{f} \quad (8)$$

$$f(\phi) = \frac{1}{2\pi}, \quad \text{que es la función de densidad de una}$$

distribución uniforme en $(0, 2\Pi)$ y $g(R) = \frac{1}{2} e^{-R/2}$

que es una función de densidad exponencial negativa con media 2, se puede escribir la ecuación 8 así:

$$F(R, \phi) = \int_0^{\phi} \int_0^R f(\phi) g(R) dR d\phi \quad (9)$$

Reemplazando $f(R)$ y $f(\phi)$ en 9 se tiene:

$$F(R, \phi) = \int_0^{\phi} \frac{1}{2\Pi} d\phi \int_0^R \frac{1}{2} e^{-R/2} dR \quad (10)$$

Entonces, simplemente se debe generar una distribución uniforme entre $(0, 2\Pi)$ y una exponencial negativa con media 2 y las multiplicamos

$$\int_0^{\mathbf{f}} \frac{1}{2\Pi} d\mathbf{f} = U_1 \quad \mathbf{f} = 2\Pi U_1 \quad (11)$$

$$\int_0^R \frac{1}{2} e^{-R/2} dR = U_2 \quad 1 - e^{-R/2} = U_2 \quad \therefore \quad (12)$$

$$R = -2\ln(1 - U_2)$$

De la ecuación 12 tenemos que $1 - U_2$ es un número aleatorio quedando entonces $R = -2\ln U_2$

Utilizando 11 y 12, y reemplazando en las ecuaciones de transformación a) y b) nos queda:

$$x = \sqrt{-2 \ln U_2} \cos(2\Pi U_1) \quad (13)$$

$$y = \sqrt{-2 \ln U_2} \text{sen}(2\Pi U_1) \quad (14)$$

Para generar P variables aleatorias normales independientes simplemente se generan P secuencias de números aleatorios independientes y reemplazando en 12 y 13 obtenemos P variables aleatorias normales independientes.

Estas transformaciones 12 y 13 son conocidas como las transformaciones de Box-Muller (1958).

5. CONCLUSIONES

Aunque las transformaciones de Box-Muller se pueden hacer más eficientes computacionalmente a través de las transformaciones matemáticas basadas en relaciones trigonométricas, el método desarrollado en este trabajo proporciona un buen generador de procesos multivariados, si se toma en cuenta el avance de la tecnología computacional.

6. BIBLIOGRAFÍA

- [1]. Graybill A. Franklin. Theory and Application of the Linear Model. Duxbury Press PWS publishers. 1976.
- [2]. Thompson James R., Wiley Jhon and sons. A modelers Approach. 2000
- [3]. Law - Kelton. Simulation Modeling and Analysis. Third edition. Mc Graw Hill. 2000.
- [4]. Cuadras C.M. Métodos de Análisis Multivariado. Editorial Eunibar. Barcelona. 1981
- Ross M. Sheldon. Simulación. Segunda edición. Prentice may. 1999
- [5]. C. Radhakrishna Rao. Linear Statistical Inference and its applications. Second Edition. Wiley series in probability and statistics. 2002.