

Series Temporales

Semana 2

1. Procesos Estacionarios

Un proceso aleatorio es estacionario si sus propiedades estadísticas permanecen constantes en el tiempo. La hipótesis de estacionaridad resulta fundamental en muchos casos pues, con frecuencia, sólo tenemos una realización del proceso a partir de la cual queremos hacer inferencia sobre las propiedades del fenómeno que estamos estudiando, y la estacionaridad nos permite aprender sobre las propiedades del fenómeno, observándolo por un periodo suficientemente largo de tiempo.

Definición 1 Sea $X_t, t \in T$ un proceso aleatorio. Decimos que es (estrictamente) estacionario si dado cualquier $n \in \mathbb{N}$, $h \in \mathbb{R}$ y cualesquiera t_1, t_2, \dots, t_n en T tales que $t_1 + h, t_2 + h, \dots, t_n + h$ también están en T , los vectores

$$(X_{t_1}, X_{t_2}, \dots, X_{t_n}); \quad y \quad (X_{t_1+h}, X_{t_2+h}, \dots, X_{t_n+h})$$

tienen la misma distribución.

En la práctica esta condición es imposible de verificar, pues requeriría conocer todas las distribuciones finito-dimensionales del proceso. Por ello se usa una condición más débil, que sólo pone condiciones en los dos primeros momentos de las variables que forman el proceso.

1.1. Procesos Débilmente Estacionarios

Sea $\{X(t), t \in T\}$ un proceso con segundo momento finito. Decimos que $X(t)$ es estacionario en sentido amplio o débilmente estacionario si la función de media es constante:

$$m(t) = E(X(t)) = m, \quad \forall t \in T,$$

y la función de (auto)covarianza $\gamma(s, t)$ sólo depende de la diferencia $s - t$:

$$\gamma(s, t) = \gamma(s - t).$$

Como $\gamma(s, t) = \gamma(t, s)$, en el caso estacionario tenemos $\gamma(s - t) = \gamma(t - s)$ y por lo tanto $\gamma(h) = \gamma(-h)$, es decir, la función de covarianza es par.

Si $\{X(t), t \in T\}$ un proceso estacionario, la función de autocovarianza a una distancia o lapso h es

$$\gamma_X(h) = \text{Cov}(X(t+h), X(t)) = E((X(t+h) - m)(X(t) - m)).$$

La función de autocorrelación (ACF) a un lapso h es

$$\rho_X(h) = \frac{\gamma_X(h)}{\gamma_X(0)}$$

Recordamos la propiedad de linealidad para la covarianza: Si X, Y, Z son v.a. con segundo momento finito y $a, b, c \in \mathbb{R}$,

$$\text{Cov}(aX + bY + c, Z) = a \text{Cov}(X, Z) + b \text{Cov}(Y, Z)$$

1.1.1. Ejemplos

1. Ruido i.i.d.: $X_n, n \geq 1$ es una sucesión de v.a.i.i.d. con $E(X_n) = 0, \text{Var}(X_n) = \sigma^2$. En este caso $m = E(X_n) = 0$ y

$$\gamma_X(t, t+h) = \begin{cases} \sigma^2 & \text{si } h = 0 \\ 0 & \text{si } h \neq 0. \end{cases}$$

2. Ruido Blanco. En este caso las variables $X_n, n \geq 1$ no son independientes pero no están correlacionadas. El resultado en cuanto a las funciones de media y autocovarianza es el mismo.
3. Promedio Móvil.

Consideremos el proceso

$$v_t = \frac{1}{3}(w_{t-1} + w_t + w_{t+1})$$

Podemos ver que para este proceso

$$m_v(t) = E(v_t) = \frac{1}{3}(E(w_{t-1}) + E(w_t) + E(w_{t+1})) = 0$$

En cuanto a la función de autocovarianza tenemos

$$\gamma_v(h) = \begin{cases} \frac{3}{9}\sigma_w^2 & \text{para } h = 0, \\ \frac{2}{9}\sigma_w^2 & \text{para } |h| = 1, \\ \frac{1}{9}\sigma_w^2 & \text{para } |h| = 2, \\ 0 & \text{para } |h| > 2, \end{cases}$$

4. Paseo Aleatorio con Deriva. Este proceso se define por

$$X(t) = \delta t + \sum_{j=1}^t w_j$$

En consecuencia

$$m_X(t) = E(X(t)) = \delta t + \sum_{j=1}^t E(w_j) = \delta t$$

y como esta función no es constante, el paseo al azar no es un proceso estacionario. Calculemos su covarianza

$$\gamma_X(s, t) = \text{Cov}(X_s, X_t) = \text{Cov}\left(\sum_{j=1}^s w_j, \sum_{k=1}^t w_k\right) = \min\{s, t\}\sigma_w^2$$

y en particular

$$\text{Var}(X(t)) = t\sigma_w^2$$

de modo que la varianza del paseo al azar aumenta con el tiempo.

1.1.2. Propiedades de γ

- $\gamma(0) \geq 0$, $|\gamma(h)| \leq \gamma(0)$, $\forall h$, $\gamma(h) = \gamma(-h)$.
- Una función real definida en \mathbb{Z} es la autocovarianza de una ST estacionaria si y sólo si es par y definida no-negativa.

Veamos que una función de autocovarianza es definida no-negativa:

$$\begin{aligned} \sum_{i,j=1}^n a_i \gamma(i-j) a_j &= \sum_{i=1}^n \sum_{j=1}^n a_i a_j E((X_i - m)(X_j - m)) \\ &= E\left(\sum_{i=1}^n a_i (X_i - m) \left(\sum_{j=1}^n a_j X_j - m\right)\right) \\ &= E\left(\left(\sum_{i=1}^n a_i (X_i - m)\right)^2\right) \geq 0 \end{aligned}$$

La autocorrelación ρ tiene las mismas propiedades que hemos descrito para la autocovarianza y la propiedad adicional $\rho(0) = 1$.

Ejemplo 1

La función $K(h) = \cos(\omega h)$ para $\omega \neq 0$ es definida no-negativa, ya que es la función de covarianza del proceso estacionario

$$X_t = A \cos(\omega t) + B \sen(\omega t)$$

con A, B v.a. centradas de varianza 1 e independientes. Tenemos

$$\begin{aligned} X_t X_s &= (A \cos(\omega t) + B \sen(\omega t))(A \cos(\omega s) + B \sen(\omega s)) \\ &= A^2 \cos \omega t \cos \omega s + B^2 \sen \omega t \sen \omega s \\ &\quad + AB(\cos \omega t \sen \omega s + \sen \omega t \cos \omega s) \end{aligned}$$

y

$$E(X_t X_s) = \cos \omega t \cos \omega s + \sen \omega t \sen \omega s = \cos \omega(t - s)$$

Observamos que X_t se puede representar como

$$X_t = C \sen(\omega t + \phi)$$

con $C^2 = A^2 + B^2$ y $\phi = \arctan(A/b)$.

Definición 2 Si tenemos dos series de tiempo X_t, Y_t definimos la (función de) covarianza cruzada o croskovarianza por

$$\gamma_{XY}(s, t) = \text{Cov}(X_s, Y_t) = E[(X_s - \mu_X(s))(Y_t - \mu_Y(t))]$$

La función de correlación cruzada o crosacorrelación es

$$\rho_{XY}(s, t) = \frac{\gamma_{XY}(s, t)}{(\gamma_X(s, s)\gamma_Y(t, t))^{1/2}}$$

Definición 3 Dos series de tiempo X, Y son conjuntamente estacionarias si cada una es estacionaria y su función de croskovarianza sólo depende de $s - t$. La función de cros correlación de un par de series de tiempo conjuntamente estacionarias X_t, Y_t se define por

$$\rho_{XY}(h) = \frac{\gamma_{XY}(h)}{(\gamma_X(0)\gamma_Y(0))^{1/2}}$$

Ejemplo 2

Sean $X_t = w_t + w_{t-1}$, $Y_t = w_t - w_{t-1}$ con $w_t \sim iid(0, \sigma^2)$. Tenemos

$$\begin{aligned} \gamma_X(0) &= E((w_t + w_{t-1})^2) = E(w_t^2 + w_{t-1}^2 + 2w_t w_{t-1}) = 2\sigma^2 \\ \gamma_Y(0) &= \gamma_X(0) = 2\sigma^2 \\ \gamma_X(1) &= E((w_{t+1} + w_t)(w_t + w_{t-1})) = E(w_t^2) = \sigma^2 = \gamma_Y(-1) \\ \gamma_X(-1) &= \gamma_Y(1) = -\sigma^2 \\ \gamma_{XY}(1) &= E((w_{t+1} + w_t)(w_t - w_{t-1})) = E(w_t^2) = \sigma^2 \\ \gamma_{XY}(0) &= E((w_t + w_{t-1})(w_t - w_{t-1})) = E(w_t^2 - w_{t-1}^2) = 0 \\ \gamma_{XY}(-1) &= -E(w_t^2) = -\sigma^2 \end{aligned}$$

y en consecuencia

$$\rho_{XY}(h) = \begin{cases} 0 & h = 0, \\ 1/2 & h = 1 \\ -1/2 & h = -1 \\ 0 & |h| \geq 2. \end{cases}$$

Ejemplo 3

Veamos que la función $\Gamma(h)$ definida en los enteros por

$$\Gamma(h) = \begin{cases} 1 & \text{si } h = 0, \\ \rho & \text{si } h = \pm 1, \\ 0 & \text{en otro caso,} \end{cases}$$

es la función de autocovarianza de un proceso estacionario ssi $|\rho| \leq 1/2$. Sea $w_k \sim WN(0, \sigma^2)$ y $\theta \in \mathbb{R}$. Definimos

$$X_t = w_t + \theta w_{t-1}, \quad t = 0, \pm 1, \dots$$

En este caso,

$$E(X_t) = 0, \quad E(X_t^2) = E(w_t^2 + \theta^2 w_{t-1}^2 + 2\theta w_{t-1} w_t) = \sigma^2(1 + \theta^2)$$

$$\begin{aligned} \gamma_X(t+h, t) &= E(X_{t+h} X_t) = E((w_{t+h} + \theta w_{t+h-1})(w_t + \theta w_{t-1})) \\ &= \begin{cases} \sigma^2(1 + \theta^2) & \text{si } h = 0, \\ \sigma^2\theta & \text{si } h = \pm 1, \\ 0 & \text{si } |h| > 1. \end{cases} \end{aligned}$$

y la función de autocorrelación es

$$\rho_X(h) = \begin{cases} 1 & \text{si } h = 0, \\ \theta/(1 + \theta^2) & \text{si } h = \pm 1, \\ 0 & \text{si } |h| > 1. \end{cases}$$

Ponemos

$$\frac{\theta}{1 + \theta^2} = \rho; \quad \theta = \rho + \rho\theta^2; \quad \rho\theta^2 - \theta + \rho = 0$$

que tiene solución

$$\theta = \frac{1 \pm \sqrt{1 - 4\rho^2}}{2\rho}$$

Para tener una solución real es necesario que $4\rho^2 \leq 1 \Rightarrow |\rho| \leq 1/2$.

Si $|\rho| > 1/2$ no hay un proceso de este tipo que tenga covarianza Γ . Veamos que en este caso Γ no es definida no-negativa: Consideremos la matriz de covarianza para tres valores de h y $\rho > 1/2$:

$$\Gamma_3 = \begin{pmatrix} \Gamma(0) & \Gamma(1) & \Gamma(2) \\ \Gamma(1) & \Gamma(0) & \Gamma(1) \\ \Gamma(2) & \Gamma(1) & \Gamma(0) \end{pmatrix} = \begin{pmatrix} 1 & \rho & 0 \\ \rho & 1 & \rho \\ 0 & \rho & 1 \end{pmatrix}$$

Para ver que esta matriz no es definida no-negativa tomamos el vector $a = (1, -1, 1)$:

$$(1, -1, 1) \begin{pmatrix} 1 & \rho & 0 \\ \rho & 1 & \rho \\ 0 & \rho & 1 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} = 1 - \rho + 1 - 2\rho + 1 - \rho = 3 - 4\rho$$

y si $\rho > 3/4$ esta expresión es negativa. En general, si tomamos Γ_n y $a = (1, -1, 1, -1, \dots)$ entonces

$$a' \Gamma_n a = n - 2(n-1)\rho < 0$$

para $2\rho/(2\rho - 1) < n$. Para $\rho < -1/2$ usamos el mismo argumento con $a = (1, 1, 1, \dots)$

2. Proceso Lineal

El *proceso lineal* se define como una combinación lineal infinita de ruido blanco w_t :

$$X_t = \mu + \sum_{j=-\infty}^{\infty} \psi_j w_{t-j}, \quad \sum_{j=-\infty}^{\infty} |\psi_j| < \infty.$$

La segunda condición garantiza que la suma infinita que define al proceso converge con probabilidad uno. Recordemos que, usando la desigualdad de Jensen,

$$(\mathbb{E} |w_t|)^2 \leq \mathbb{E}(|w_t|^2) = \sigma^2$$

y en consecuencia $\mathbb{E} |w_t| \leq \sigma$. Tenemos

$$\begin{aligned} \mathbb{E} |X_t| &= \mathbb{E} \left| \sum_{j=-\infty}^{\infty} \psi_j w_{t-j} \right| \leq \mathbb{E} \sum_{j=-\infty}^{\infty} |\psi_j| |w_{t-j}| \\ &= \sum_{j=-\infty}^{\infty} |\psi_j| \mathbb{E}(|w_{t-j}|) \leq \sum_{j=-\infty}^{\infty} |\psi_j| \sigma < \infty. \end{aligned}$$

Para el proceso lineal tenemos $E(X_t) = \mu$ y

$$\begin{aligned} E((X_{t-h} - \mu)(X_t - \mu)) &= E\left(\sum_{j=-\infty}^{\infty} \psi_j w_{t+h-j} \sum_{k=-\infty}^{\infty} \psi_k w_{t-k}\right) \\ &= \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} \psi_j \psi_k E(w_{t+h-j} w_{t-k}) \\ &= \sigma^2 \sum_{j=-\infty}^{\infty} \psi_{j+h} \psi_j \end{aligned}$$

para $h \geq 0$. En general es posible representar los procesos ARMA de esta forma.

Introducimos el operador de retardo o *backward shift* B definido por $BX_t = X_{t-1}$. Usando este operador podemos escribir la serie que define al proceso lineal de manera más compacta como

$$X_t = \psi(B)w_t$$

donde $\psi(B) = \sum_{j=-\infty}^{\infty} \psi_j B^j$. Podemos pensar que el operador $\psi(B)$ es un filtro lineal, que al aplicarlo al proceso de ruido blanco w_t produce como resultado el proceso X_t .

Decimos que un proceso lineal es un proceso de promedio móvil o $MA(\infty)$ si $\mu = 0$ y $\psi_j = 0$ para $j < 0$, es decir, si

$$X_t = \sum_{j=0}^{\infty} \psi_j w_{t-j},$$

El siguiente resultado muestra que si aplicamos un filtro lineal a cualquier proceso estacionario, obtenemos un proceso estacionario

Proposición 1 *Sea $\{Y_t\}$ una serie de tiempo estacionaria con media 0 y función de covarianza γ_Y . Si $\sum_{j=-\infty}^{\infty} |\psi_j| < \infty$ entonces la serie de tiempo*

$$X_t = \sum_{j=-\infty}^{\infty} \psi_j Y_{t-j} = \psi(B)Y_t \quad (1)$$

es estacionaria con media 0 y función de autocovarianza

$$\gamma_X(h) = \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} \psi_j \psi_k \gamma_Y(h+k-j) \quad (2)$$

En el caso particular cuando $\{X_t\}$ es un proceso lineal,

$$\gamma_X(h) = \sum_{j=-\infty}^{\infty} \psi_j \psi_{j+h} \sigma^2.$$

Demostración. El mismo argumento que usamos en la definición del proceso lineal muestra que la serie 1 es convergente con probabilidad 1. Como $E(Y_t) = 0$ tenemos

$$E(X_i) = E\left(\sum_{j=-\infty}^{\infty} \psi_j Y_{t-j}\right) = \sum_{j=-\infty}^{\infty} \psi_j E(Y_{t-j}) = 0$$

y

$$\begin{aligned}
\mathbb{E}(X_{t+h}X_t) &= \mathbb{E} \left[\left(\sum_{j=-\infty}^{\infty} \psi_j Y_{t+h-j} \right) \left(\sum_{k=-\infty}^{\infty} \psi_k Y_{t-k} \right) \right] \\
&= \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} \psi_j \psi_k \mathbb{E}(Y_{t+h-j} Y_{t-k}) \\
&= \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} \psi_j \psi_k \gamma_Y(h-j+k)
\end{aligned}$$

Esto muestra que $\{X_t\}$ es un proceso estacionario con función de covarianza (2). Finalmente, si $\{Y_t\}$ es un ruido blanco, entonces $\gamma_Y(h-j+k) = \sigma^2$ si $k = j-h$ y es 0 en otro caso. ■

3. Estimación.

En el caso de series de tiempo usualmente tenemos una realización x_1, x_2, \dots, x_n a partir de la cual deseamos estimar la media, la autocovarianza y la autocorrelación. No disponemos, como sucede en el caso de la estadística clásica, de varias copias independientes del proceso a partir de las cuales podemos hacer la estimación. Para poder hacer la estimación de la media, por ejemplo, en el caso de un proceso estacionario, requerimos un resultado de convergencia tipo LGN: Queremos estimar

$$\mathbb{E}(X_t) = \int X_t(\omega) dP(\omega) = m$$

por

$$\frac{1}{n} \sum_{i=1}^n X_i.$$

Si las variables X_i son i.i.d. entonces sabemos que

$$\frac{1}{n} \sum_{i=1}^n X_i \rightarrow m \quad c.s.$$

pero en nuestro caso hay una estructura de correlación en el proceso y las variables X_i no son i.i.d.

Un teorema que garantice la convergencia del estimador en el caso dependiente se conoce como un teorema ergódico y esta propiedad del proceso se conoce como ergodicidad. Vamos a suponer que los procesos que consideramos tienen esta propiedad. Resultados de este tipo se pueden ver en el libro de Karlin y Taylor.

3.1. Estimación de la Media

Usamos como estimador

$$\bar{X}_n = \frac{1}{n}(X_1 + \dots + X_n).$$

\bar{X}_n es un estimador insesgado porque $E(\bar{X}_n) = \frac{1}{n} \sum_{i=1}^n E(X_i) = \mu$. El error medio cuadrático de este estimador es

$$\begin{aligned} E(\bar{X}_n - \mu)^2 &= \text{Var}(\bar{X}_n) = \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n \text{Cov}(X_i, X_j) \\ &= \frac{1}{n^2} \sum_{i-j=-n}^n (n - |i-j|) \gamma(i-j) \\ &= \frac{1}{n} \sum_{h=-n}^n \left(1 - \frac{|h|}{n}\right) \gamma(h) \end{aligned}$$

Si $\gamma(h) \rightarrow 0$ cuando $h \rightarrow \infty$, \bar{X}_n converge a μ en media cuadrática. Para demostrar esto a partir de la expresión anterior observamos que dado $\varepsilon > 0$ existe H t.q. $|h| > H \Rightarrow \gamma(h) < \varepsilon/2$. Tenemos

$$\begin{aligned} \frac{1}{n} \sum_{h=-n}^n \left(1 - \frac{|h|}{n}\right) \gamma(h) &\leq \frac{2}{n} \sum_{h=0}^n \left(1 - \frac{h}{n}\right) \gamma(h) \\ &= \frac{2}{n} \left(\sum_{h=0}^H + \sum_{h=H+1}^n \right) \left(1 - \frac{h}{n}\right) \gamma(h) \\ &\leq \frac{2H}{n} \gamma(0) + \frac{2}{n} \sum_{H+1}^n \left(1 - \frac{h}{n}\right) \frac{\varepsilon}{2} \\ &\leq \frac{C}{n} + \varepsilon \end{aligned}$$

donde C es una constante. Esto es suficiente para ver la convergencia.

Si

$$\sum_{h=-\infty}^{\infty} \gamma(h) < \infty$$

entonces

$$\lim_{n \rightarrow \infty} n \text{Var}(\bar{X}_n) = \sum_{|h| < \infty} \gamma(h)$$

Resumimos este resultado en la siguiente proposición.

Proposición 2 Si X_t es una S.T. estacionaria con media μ y función de autocovarianza γ , entonces, cuando $n \rightarrow \infty$,

$$\begin{array}{ll} \text{Var}(\bar{X}_n) = E(\bar{X}_n - \mu)^2 \rightarrow 0 & \text{si } \gamma(h) \rightarrow 0 \\ n E(\bar{X}_n - \mu)^2 \rightarrow \sum_{|h| < \infty} \gamma(h) & \text{si } \sum_h |\gamma(h)| < \infty. \end{array}$$

Para hacer inferencia sobre μ usando \bar{X}_n es útil conocer la distribución de \bar{X}_n o una aproximación. Si la S.T. es gaussiana entonces

$$\sqrt{n}(\bar{X}_n - \mu) \sim \mathcal{N}\left(0, \sum_{|h| < n} \left(1 - \frac{|h|}{n}\right) \gamma(h)\right)$$

Para muchas ST y en particular para muchos procesos lineales y modelos ARMA, \bar{X}_n es aproximadamente normal con media μ y varianza $\frac{1}{n} \sum_{|h|<\infty} \gamma(h)$ para n grande. Un intervalo de confianza aproximado al 95 % para μ es

$$\left(\bar{X}_n - \frac{1,96\nu^{1/2}}{\sqrt{n}}, \bar{X}_n + \frac{1,96\nu^{1/2}}{\sqrt{n}} \right) \quad \text{con } \nu = \sum_{|h|<\infty} \gamma(h)$$

No conocemos ν pero lo podemos estimar por

$$\hat{\nu} = \sum_{|h|<\sqrt{n}} \left(1 - \frac{|h|}{n}\right) \hat{\gamma}(h)$$

3.2. Estimación de la autocovarianza y autocorrelación

La autocovarianza muestral se define como

$$\hat{\gamma}(h) = \frac{1}{n} \sum_{t=1}^{n-|h|} (X_{t+|h|} - \bar{X}_n)(X_t - \bar{X}_n)$$

La autocorrelación muestral se define de manera análoga:

$$\hat{\rho}(h) = \frac{\hat{\gamma}(h)}{\hat{\gamma}(0)}$$

Ambos estimadores tienen sesgo (y tampoco serían insesgados si reemplazamos n por $n - h$ en el denominador). La ventaja de usar esta expresión (con n) es que la función que resulta $\hat{\gamma}(h)$ es positiva definida, al igual que la covarianza $\gamma(h)$. Esta propiedad asegura que la varianza de cualquier combinación lineal de variables X_t nunca es negativa, y es natural pedir esta misma condición al estimador de la autocovarianza.

La propiedad de ser positiva definida dice que si

$$\hat{\Gamma}_k = \begin{pmatrix} \hat{\gamma}(0) & \hat{\gamma}(1) & \cdots & \hat{\gamma}(k-1) \\ \hat{\gamma}(1) & \hat{\gamma}(0) & \cdots & \hat{\gamma}(k-2) \\ \hat{\gamma}(2) & \hat{\gamma}(1) & \cdots & \hat{\gamma}(k-3) \\ \vdots & \vdots & \vdots & \vdots \\ \hat{\gamma}(k-1) & \hat{\gamma}(k-2) & \cdots & \hat{\gamma}(0) \end{pmatrix}$$

entonces para cualquier vector \mathbf{a} de dimensión k , $\mathbf{a}'\hat{\Gamma}_k\mathbf{a} \geq 0$.

Observamos que si $\hat{\Gamma}_m$ es no-negativa definida entonces $\hat{\Gamma}_k$ también lo es para todo $k < m$.

La ACF muestral juega un papel importante en la selección de modelos adecuados para los datos. Aunque no es posible calcular la distribución de la estadística $\hat{\rho}(h)$ en la mayoría de los casos, en general se puede aproximar adecuadamente por una distribución normal si el tamaño de la muestra es grande.

Proposición 3 Si X_t es i.i.d. con cuarto momento finito entonces, para n grande, la ACF muestral $\hat{\rho}_X(h)$ para $h = 1, 2, \dots, H$ con H arbitrario pero fijo, es aproximadamente normal centrada con desviación típica dada por

$$\sigma_{\hat{\rho}_X(h)} = \frac{1}{\sqrt{H}}$$

Con base en este resultado tenemos un método para evaluar si los valores altos de $\hat{\rho}(h)$ son significativos, al determinar si salen del intervalo $\pm 2/\sqrt{N}$. Si la sucesión corresponde a un ruido blanco, aproximadamente 95 % de la ACF muestral debe caer dentro de este intervalo.

3.3. Estimación de la Covarianza Cruzada

Covarianza Cruzada:

$$\hat{\gamma}_{XY}(h) = \frac{1}{n} \sum_{t=1}^{n-h} (x_{t+h} - \bar{x})(y_t - \bar{y})$$

con $\hat{\gamma}_{XY}(-h) = \hat{\gamma}_{XY}(h)$.

Correlación Cruzada:

$$\hat{\rho}_{XY}(h) = \frac{\hat{\gamma}_{XY}(h)}{\sqrt{\gamma_X(0)\gamma_Y(0)}}$$

Proposición 4 Para muestras grandes $\hat{\rho}_{XY}(h)$ tiene distribución aproximadamente normal centrada con desviación típica

$$\sigma_{\hat{\rho}_{XY}(h)} = \frac{1}{\sqrt{n}}$$

si al menos uno de los procesos es un ruido blanco.

4. Series Vectoriales

Con frecuencia encontramos situaciones en las cuales es de interés estudiar las relaciones entre series temporales que han sido medidas conjuntamente. Por lo tanto será útil considerar la noción de una serie temporal vectorial $\mathbf{X}_t = (X_1(t), X_2(t), \dots, X_p(t))$, cuyas componentes son p series temporales univariadas. En el caso estacionario, el vector de medias μ está dado por

$$\mu = E(\mathbf{X}(t))$$

de la forma $\mu = (\mu_1, \mu_2, \dots, \mu_p)'$ y la matriz de autocovarianza de orden $p \times p$

$$\Gamma(h) = E((\mathbf{X}_i(t+h) - \mu)(\mathbf{X}_i(t) - \mu)')$$

donde los elementos de la matriz $\Gamma(h)$ son las funciones de covarianza cruzada

$$\gamma_{ij}(h) = E(((X_i(t+h) - \mu_i)(X_j(t) - \mu_j)))$$

para $i = 1, \dots, p$. Como $\gamma_{ij}(h) = \gamma_{ji}(-h)$, tenemos que

$$\Gamma(-h) = \Gamma'(h).$$

La función de autocovarianza muestral de la serie vectorial \mathbf{x}_t es la matriz $p \times p$ de croskovarianzas

$$\hat{\Gamma}(h) = \frac{1}{n} \sum_{t=1}^{n-h} (\mathbf{x}_{t+h} - \bar{\mathbf{x}})(\mathbf{x}_t - \bar{\mathbf{x}})'$$

donde

$$\bar{\mathbf{x}} = \frac{1}{n} \sum_{t=1}^n \mathbf{x}_t$$

es el vector de media muestral. La propiedad de simetría también es cierta para para la autocovarianza muestral:

$$\hat{\Gamma}(-h) = \hat{\Gamma}'(h)$$

En muchos problemas aplicados el índice de una serie puede ser también multidimensional. Por ejemplo, la posición de una unidad experimental puede describirse por dos coordenadas, s_1 y s_2 .

La función de autocovarianza para un proceso estacionario multidimensional $x_{\mathbf{s}}$ se define en términos de un vector de lapsos $\mathbf{h} = (h_1, \dots, h_r)'$, por

$$\gamma(\mathbf{h}) = \text{E}((x_{\mathbf{s}+\mathbf{h}} - \mu)(x_{\mathbf{s}} - \mu))$$

donde

$$\mu = \text{E}(s_{\mathbf{s}})$$

que no depende de la coordenada espacial \mathbf{s} . Para el proceso bidimensional de temperaturas tenemos

$$\gamma(h_1, h_2) = \text{E}((x_{s_1+h_1, s_2+h_2} - \mu)(x_{s_1, s_2} - \mu))$$

que es función del lapso, tanto en las filas (h_1) como en las columnas (h_2).

La función de autocovarianza muestral multidimensional se define como

$$\hat{\gamma}(\mathbf{h}) = (S_1 S_2 \cdots S_r)^{-1} \sum_{s_1} \sum_{s_2} \cdots \sum_{s_r} (x_{\mathbf{s}+\mathbf{h}} - \bar{x})(x_{\mathbf{s}} - \bar{x})$$

donde $\mathbf{s} = (s_1, s_2, \dots, s_r)'$ y el rango de la suma para cada índice es $1 \leq s_i \leq S_i - h_i$, para $i = 1, \dots, r$. La media se calcula para el arreglo r -dimensional:

$$\bar{x} = (S_1 S_2 \cdots S_r)^{-1} \sum_{s_1} \sum_{s_2} \cdots \sum_{s_r} x_{s_1, s_2, \dots, s_r},$$

donde los argumentos se suman sobre $1 \leq s_i \leq S_i$. La función de autocorrelación muestral mutidimensional está dada por

$$\hat{\rho}(\mathbf{h}) = \frac{\hat{\gamma}(\mathbf{h})}{\hat{\gamma}(\mathbf{0})}.$$