Series de tiempo univariadas - Presentación 3

Mauricio Alejandro Mazo Lopera

Universidad Nacional de Colombia Facultad de Ciencias Escuela de Estadística Medellín



Veamos qué es un proceso estocástico:

Proceso estocástico: Es un conjunto de variables aleatorias que denotatemos por Z(w,t) donde w pertenece al espacio muestral y t pertenece a un conjunto de índices. En este curso asumiremos que el conjunto de índices donde t toma valores es el conjunto de los números enteros.

Veamos qué es un proceso estocástico:

Proceso estocástico: Es un conjunto de variables aleatorias que denotatemos por Z(w,t) donde w pertenece al espacio muestral y t pertenece a un conjunto de índices. En este curso asumiremos que el conjunto de índices donde t toma valores es el conjunto de los números enteros.

Con respecto a esta definición podemos considerar lo siguiente:

- Para un valor fijo t, Z(w, t) es una variable aleatoria (v.a.).
- Para un valor dado w, Z(w, t) como una función de t, es llamada una realización del proceso estocástico. Formalmente, una serie de tiempo es una realización del proceso estocástico y corresponde a una observación del proceso.
- En adelante, Z(w, t) se escribirá simplemente como Z_t o Z(t).

Ejemplo 1: Considere que decide lanzar una moneda cuatro veces y cada lanzamiento se hace cada hora comenzando desde este instante.

Ejemplo 1: Considere que decide lanzar una moneda cuatro veces y cada lanzamiento se hace cada hora comenzando desde este instante. El espacio muestral, para cada t=1,2,3,4, denotado por \mathcal{S}_t sería:

$$S_t = \{w_t : w_t = C \text{ (cara) \'o } w_t = S \text{ (sello)}\}$$

Ejemplo 1: Considere que decide lanzar una moneda cuatro veces y cada lanzamiento se hace cada hora comenzando desde este instante. El espacio muestral, para cada t = 1, 2, 3, 4, denotado por S_t sería:

$$\mathcal{S}_t = \{ w_t : w_t = C \text{ (cara) \'o } w_t = S \text{ (sello)} \}$$

Y el espacio muestral completo sería:

$$S = S_1 \times S_2 \times S_3 \times S_4$$

= {(w_1, w_2, w_3, w_4): con $w_t = C, S, y t = 1, 2, 3, 4$ }

En este caso:

• Para cada valor fijo t, $Z(w_t, t)$ es una v.a. que puede tomar valores según:

$$Z(w_t, t): \mathcal{S}_t \longmapsto \mathbb{R}$$
 $w_t \longrightarrow Z(w_t, t) = \begin{cases} 1, & \text{si } w_t = C \\ 0, & \text{si } w_t = S \end{cases}$

• Para $w = (w_1, w_2, w_3, w_4)$ dado, Z(w, t) es una función de t:

$$Z(\mathbf{w}, t) : \{1, 2, 3, 4\} \longmapsto \mathbb{R}$$

$$t \longrightarrow Z(\mathbf{w}, t) = \begin{cases} 1, & \text{si } w_t = C \\ 0, & \text{si } w_t = S \end{cases}$$

Ejemplo 2: Considere el experimento que consiste en ver los precios de cierre diarios de cierta acción durante un mes.

Ejemplo 2: Considere el experimento que consiste en ver los precios de cierre diarios de cierta acción durante un mes. Suponiendo que en dicho mes solo hay 21 días de mercado, entonces el espacio muestral, para cada $t=1,\ldots,21$, denotado por \mathcal{S}_t sería:

$$\mathcal{S}_t = \{w_t: \ 0 \le w_t < \infty\}$$

Ejemplo 2: Considere el experimento que consiste en ver los precios de cierre diarios de cierta acción durante un mes. Suponiendo que en dicho mes solo hay 21 días de mercado, entonces el espacio muestral, para cada $t=1,\ldots,21$, denotado por \mathcal{S}_t sería:

$$\mathcal{S}_t = \{w_t: \ 0 \le w_t < \infty\}$$

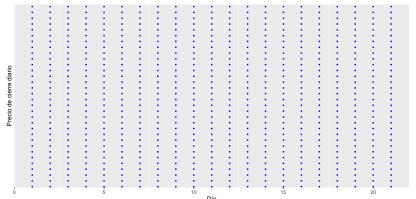
Y el espacio muestral completo sería:

$$\begin{array}{lcl} \mathcal{S} & = & \mathcal{S}_1 \times \ldots \times \mathcal{S}_{21} \\ & = & \{ \left(w_1, \ldots, w_{21} \right) : \text{ con } 0 \leq w_t < \infty, \text{ para } t = 1, \ldots, 21 \} \end{array}$$

En este caso:

• Para cada valor fijo t, $Z(w_t, t)$ es una v.a. que puede tomar valores según:

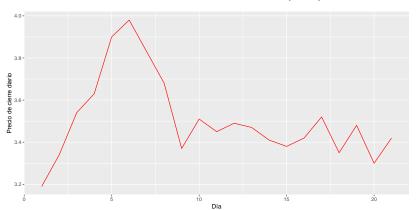
$$Z(w_t, t) : \mathcal{S}_t \longmapsto \mathbb{R}$$
 $w_t \longrightarrow Z(w_t, t) = w_t$



• Para $w = (w_1, \ldots, w_{21})$ dado, es decir, después de ver los precios diarios durante ese mes, se tiene que Z(w, t) es una función de t:

$$Z(\mathbf{w},t): \{1,\ldots,21\} \longmapsto \mathbb{R}$$

 $t \longrightarrow Z(\mathbf{w},t) = w_t$



Considere un conjunto finito de variables aleatorias

$$\{Z_{t_1}, Z_{t_2}, \dots, Z_{t_n}\}$$

de un proceso estocástico $\{Z(w,t): t=0,\pm 1,\pm 2,\ldots\}$. Su función de distribución conjunta está definida como:

$$F_{Z_{t_1}, \dots, Z_{t_n}}(z_1, z_2, \dots, z_n) = P(Z_{t_1} \leq z_1, Z_{t_2} \leq z_2, \dots, Z_{t_n} \leq z_n)$$

donde z_1, z_2, \ldots, z_n son constantes reales.

Considere un conjunto finito de variables aleatorias

$$\{Z_{t_1}, Z_{t_2}, \ldots, Z_{t_n}\}$$

de un proceso estocástico $\{Z(w,t): t=0,\pm 1,\pm 2,\ldots\}$. Su función de distribución conjunta está definida como:

$$F_{Z_{t_1}, \dots, Z_{t_n}}(z_1, z_2, \dots, z_n) = P(Z_{t_1} \leq z_1, Z_{t_2} \leq z_2, \dots, Z_{t_n} \leq z_n)$$

donde z_1, z_2, \ldots, z_n son constantes reales.

Se considera lo siguiente:

 Un proceso estocástico es llamado estacionario de orden n en distribución si la función de distribución conjunta ndimensional es invariante en el tiempo. Es decir:

$$F_{Z_{t_1},\ldots,Z_{t_n}}(\mathbf{z_1},\ldots,\mathbf{z_n}) = F_{Z_{t_1+k},\ldots,Z_{t_n+k}}(\mathbf{z_1},\ldots,\mathbf{z_n})$$
 (1)

para cualquier *n*-túpla (t_1, t_2, \ldots, t_k) y cualquier entero k.

 Un proceso estocástico es llamado estacionario de orden n en distribución si la función de distribución conjunta ndimensional es invariante en el tiempo. Es decir:

$$F_{Z_{t_1},\ldots,Z_{t_n}}(\mathbf{z_1},\ldots,\mathbf{z_n}) = F_{Z_{t_1+k},\ldots,Z_{t_n+k}}(\mathbf{z_1},\ldots,\mathbf{z_n}) \qquad (1)$$

para cualquier *n*-túpla (t_1, t_2, \dots, t_k) y cualquier entero k.

• Un proceso es llamado **estrictamente estacionario** (o **fuertemente estacionario**, o **completamente estacionario**) si la ecuación (1) es cierta para todo $n = 1, 2, \ldots$

 Un proceso estocástico es llamado estacionario de orden n en distribución si la función de distribución conjunta ndimensional es invariante en el tiempo. Es decir:

$$F_{Z_{t_1},\ldots,Z_{t_n}}(\mathbf{z_1},\ldots,\mathbf{z_n}) = F_{Z_{t_1+k},\ldots,Z_{t_n+k}}(\mathbf{z_1},\ldots,\mathbf{z_n})$$
 (1)

para cualquier *n*-túpla (t_1, t_2, \dots, t_k) y cualquier entero k.

- Un proceso es llamado **estrictamente estacionario** (o **fuertemente estacionario**, o **completamente estacionario**) si la ecuación (1) es cierta para todo $n = 1, 2, \ldots$
- Si se cumple la propiedad (1) para n, entonces se cumple para todo $m \le n$, es decir, la estacionariedad de orden alto implica siempre la estacionariedad de ordenes más bajos.

 En la práctica, generalmente es muy difícil probar si un proceso es estrictamente estacionario, y de manera alternativa se trata de caracterizar los procesos estocásticos en términos de sus momentos, cuyas propiedades se pueden verificar más fácilmente.

- En la práctica, generalmente es muy difícil probar si un proceso es estrictamente estacionario, y de manera alternativa se trata de caracterizar los procesos estocásticos en términos de sus momentos, cuyas propiedades se pueden verificar más fácilmente.
- Un proceso es llamado débilmente estacionario de orden n, si todos sus momentos conjuntos hasta de orden n son finitos e invariantes en el tiempo.

- Un proceso débilmente estacionario de segundo orden es tal que:
 - La media es constante
 - La varianza es constante
 - Las covarianzas y correlaciones no dependen del tiempo, solamente dependen del número de períodos que separan los términos del proceso.

Esta clase de proceso es también llamado proceso estacionario en sentido amplio o proceso estacionario en covarianza, o, simplemente, estacionario.

Veamos un ejemplo donde planteamos un proceso estocástico a partir de la ecuación:

$$Z_t = A \sin(\omega t + \Theta)$$

donde A y Θ son variables aleatorias independientes tales que:

$$E(A) = 0$$
, $Var(A) = 1$, $\Theta \sim Uniforme(-\pi, \pi)$

y ω es una constante.

Encontremos $E(Z_t)$, $Var(Z_t)$ y $Cov(Z_t, Z_{t+k})$.

$$E(Z_t) = E(A\sin(\omega t + \Theta))$$

$$E(Z_t) = E(A\sin(\omega t + \Theta))$$

= $E(A)E(\sin(\omega t + \Theta))$ (Por independencia)

$$E(Z_t) = E(A\sin(\omega t + \Theta))$$

$$= E(A)E(\sin(\omega t + \Theta)) \text{ (Por independencia)}$$

$$= 0 \times E(\sin(\omega t + \Theta))$$

$$E(Z_t) = E(A\sin(\omega t + \Theta))$$

$$= E(A)E(\sin(\omega t + \Theta)) \text{ (Por independencia)}$$

$$= 0 \times E(\sin(\omega t + \Theta))$$

$$= 0$$

Así,
$$E(Z_t) = 0$$
 para todo t .

$$Cov(Z_t, Z_{t+k}) = E[(Z_t - \underbrace{E(Z_t)}_{0})(Z_{t+k} - \underbrace{E(Z_{t+k})}_{0})]$$

$$Cov(Z_t, Z_{t+k}) = E[(Z_t - \underbrace{E(Z_t)}_{0})(Z_{t+k} - \underbrace{E(Z_{t+k})}_{0})]$$
$$= E[Z_t Z_{t+k}]$$

$$Cov(Z_{t}, Z_{t+k}) = E[(Z_{t} - \underbrace{E(Z_{t})}_{0})(Z_{t+k} - \underbrace{E(Z_{t+k})}_{0})]$$

$$= E[Z_{t}Z_{t+k}]$$

$$= E[A\sin(\omega t + \Theta) \times A\sin(\omega(t+k) + \Theta)]$$

$$Cov(Z_{t}, Z_{t+k}) = E[(Z_{t} - \underbrace{E(Z_{t})})(Z_{t+k} - \underbrace{E(Z_{t+k})})]$$

$$= E[Z_{t}Z_{t+k}]$$

$$= E[A\sin(\omega t + \Theta) \times A\sin(\omega(t+k) + \Theta)]$$

$$= E[A^{2}\sin(\omega t + \Theta)\sin(\omega(t+k) + \Theta)]$$

$$Cov(Z_{t}, Z_{t+k}) = E[(Z_{t} - \underbrace{E(Z_{t})})(Z_{t+k} - \underbrace{E(Z_{t+k})})]$$

$$= E[Z_{t}Z_{t+k}]$$

$$= E[A\sin(\omega t + \Theta) \times A\sin(\omega(t+k) + \Theta)]$$

$$= E[A^{2}\sin(\omega t + \Theta)\sin(\omega(t+k) + \Theta)]$$

$$= \underbrace{E[A^{2}]}_{1} E[\sin(\omega t + \Theta)\sin(\omega(t+k) + \Theta)]$$

$$Cov(Z_{t}, Z_{t+k}) = E[(Z_{t} - \underbrace{E(Z_{t})})(Z_{t+k} - \underbrace{E(Z_{t+k})})]$$

$$= E[Z_{t}Z_{t+k}]$$

$$= E[A\sin(\omega t + \Theta) \times A\sin(\omega(t+k) + \Theta)]$$

$$= E[A^{2}\sin(\omega t + \Theta)\sin(\omega(t+k) + \Theta)]$$

$$= \underbrace{E[A^{2}]}_{1} E[\sin(\omega t + \Theta)\sin(\omega(t+k) + \Theta)]$$

$$= E[\sin(\omega t + \Theta)\sin(\omega(t+k) + \Theta)]$$

$$Cov(Z_{t}, Z_{t+k}) = E[(Z_{t} - \underbrace{E(Z_{t})})(Z_{t+k} - \underbrace{E(Z_{t+k})})]$$

$$= E[Z_{t}Z_{t+k}]$$

$$= E[A\sin(\omega t + \Theta) \times A\sin(\omega(t+k) + \Theta)]$$

$$= E[A^{2}\sin(\omega t + \Theta)\sin(\omega(t+k) + \Theta)]$$

$$= \underbrace{E[A^{2}]}_{1} E[\sin(\omega t + \Theta)\sin(\omega(t+k) + \Theta)]$$

$$= E[\sin(\omega t + \Theta)\sin(\omega(t+k) + \Theta)]$$

Pero
$$sin(u) sin(v) = [cos(u - v) - cos(u + v)]/2$$
.

$$Cov(Z_{t}, Z_{t+k}) = E[(Z_{t} - \underbrace{E(Z_{t})})(Z_{t+k} - \underbrace{E(Z_{t+k})})]$$

$$= E[Z_{t}Z_{t+k}]$$

$$= E[A\sin(\omega t + \Theta) \times A\sin(\omega(t+k) + \Theta)]$$

$$= E[A^{2}\sin(\omega t + \Theta)\sin(\omega(t+k) + \Theta)]$$

$$= \underbrace{E[A^{2}]}_{1} E[\sin(\omega t + \Theta)\sin(\omega(t+k) + \Theta)]$$

$$= E[\sin(\omega t + \Theta)\sin(\omega(t+k) + \Theta)]$$

Pero
$$sin(u) sin(v) = [cos(u - v) - cos(u + v)]/2$$
. Luego,

$$Cov(Z_t, Z_{t+k}) = \frac{1}{2}E[\cos(\omega k) - \cos(\omega(2t+k) + 2\Theta)]$$

$$Cov(Z_t, Z_{t+k}) = \frac{1}{2}E[\cos(\omega k) - \cos(\omega(2t+k) + 2\Theta)]$$
$$= \frac{1}{2}E[\cos(\omega k)] - \frac{1}{2}E[\cos(\omega(2t+k) + 2\Theta)]$$

$$Cov(Z_t, Z_{t+k}) = \frac{1}{2}E[\cos(\omega k) - \cos(\omega(2t+k) + 2\Theta)]$$

$$= \frac{1}{2}E[\cos(\omega k)] - \frac{1}{2}E[\cos(\omega(2t+k) + 2\Theta)]$$

$$= \frac{1}{2}\cos(\omega k) - \frac{1}{2}\int_{-\pi}^{\pi}\cos(\omega(2t+k) + 2\Theta)\frac{1}{2\pi}d\Theta$$

$$Cov(Z_t, Z_{t+k}) = \frac{1}{2}E[\cos(\omega k) - \cos(\omega(2t+k) + 2\Theta)]$$

$$= \frac{1}{2}E[\cos(\omega k)] - \frac{1}{2}E[\cos(\omega(2t+k) + 2\Theta)]$$

$$= \frac{1}{2}\cos(\omega k) - \frac{1}{2}\int_{-\pi}^{\pi}\cos(\omega(2t+k) + 2\Theta)\frac{1}{2\pi}d\Theta$$

$$= \frac{1}{2}\cos(\omega k) - \frac{1}{4\pi}\int_{-\pi}^{\pi}\cos(\omega(2t+k) + 2\Theta)d\Theta$$

$$\begin{aligned} Cov(Z_t, Z_{t+k}) &= \frac{1}{2} E[\cos(\omega k) - \cos(\omega(2t+k) + 2\Theta)] \\ &= \frac{1}{2} E[\cos(\omega k)] - \frac{1}{2} E[\cos(\omega(2t+k) + 2\Theta)] \\ &= \frac{1}{2} \cos(\omega k) - \frac{1}{2} \int_{-\pi}^{\pi} \cos(\omega(2t+k) + 2\Theta) \frac{1}{2\pi} d\Theta \\ &= \frac{1}{2} \cos(\omega k) - \frac{1}{4\pi} \int_{-\pi}^{\pi} \cos(\omega(2t+k) + 2\Theta) d\Theta \\ &= \frac{1}{2} \cos(\omega k) - \frac{1}{4\pi} \times 0 \end{aligned}$$

$$Cov(Z_t, Z_{t+k}) = \frac{1}{2}E[\cos(\omega k) - \cos(\omega(2t+k) + 2\Theta)]$$

$$= \frac{1}{2}E[\cos(\omega k)] - \frac{1}{2}E[\cos(\omega(2t+k) + 2\Theta)]$$

$$= \frac{1}{2}\cos(\omega k) - \frac{1}{2}\int_{-\pi}^{\pi}\cos(\omega(2t+k) + 2\Theta)\frac{1}{2\pi}d\Theta$$

$$= \frac{1}{2}\cos(\omega k) - \frac{1}{4\pi}\int_{-\pi}^{\pi}\cos(\omega(2t+k) + 2\Theta)d\Theta$$

$$= \frac{1}{2}\cos(\omega k) - \frac{1}{4\pi} \times 0 = \frac{1}{2}\cos(\omega k)$$

$$Cov(Z_t, Z_{t+k}) = \frac{1}{2}E[\cos(\omega k) - \cos(\omega(2t+k) + 2\Theta)]$$

$$= \frac{1}{2}E[\cos(\omega k)] - \frac{1}{2}E[\cos(\omega(2t+k) + 2\Theta)]$$

$$= \frac{1}{2}\cos(\omega k) - \frac{1}{2}\int_{-\pi}^{\pi}\cos(\omega(2t+k) + 2\Theta)\frac{1}{2\pi}d\Theta$$

$$= \frac{1}{2}\cos(\omega k) - \frac{1}{4\pi}\int_{-\pi}^{\pi}\cos(\omega(2t+k) + 2\Theta)d\Theta$$

$$= \frac{1}{2}\cos(\omega k) - \frac{1}{4\pi} \times 0 = \frac{1}{2}\cos(\omega k)$$

$$Var(Z_t) = Cov(Z_t, Z_t)$$

$$Cov(Z_t, Z_{t+k}) = \frac{1}{2}E[\cos(\omega k) - \cos(\omega(2t+k) + 2\Theta)]$$

$$= \frac{1}{2}E[\cos(\omega k)] - \frac{1}{2}E[\cos(\omega(2t+k) + 2\Theta)]$$

$$= \frac{1}{2}\cos(\omega k) - \frac{1}{2}\int_{-\pi}^{\pi}\cos(\omega(2t+k) + 2\Theta)\frac{1}{2\pi}d\Theta$$

$$= \frac{1}{2}\cos(\omega k) - \frac{1}{4\pi}\int_{-\pi}^{\pi}\cos(\omega(2t+k) + 2\Theta)d\Theta$$

$$= \frac{1}{2}\cos(\omega k) - \frac{1}{4\pi} \times 0 = \frac{1}{2}\cos(\omega k)$$

$$Var(Z_t) = Cov(Z_t, Z_t) = Cov(Z_t, Z_{t+0})$$

$$\begin{aligned} Cov(Z_t, Z_{t+k}) &= \frac{1}{2} E[\cos(\omega k) - \cos(\omega(2t+k) + 2\Theta)] \\ &= \frac{1}{2} E[\cos(\omega k)] - \frac{1}{2} E[\cos(\omega(2t+k) + 2\Theta)] \\ &= \frac{1}{2} \cos(\omega k) - \frac{1}{2} \int_{-\pi}^{\pi} \cos(\omega(2t+k) + 2\Theta) \frac{1}{2\pi} d\Theta \\ &= \frac{1}{2} \cos(\omega k) - \frac{1}{4\pi} \int_{-\pi}^{\pi} \cos(\omega(2t+k) + 2\Theta) d\Theta \\ &= \frac{1}{2} \cos(\omega k) - \frac{1}{4\pi} \times 0 = \frac{1}{2} \cos(\omega k) \end{aligned}$$

$$Var(Z_t) = Cov(Z_t, Z_t) = Cov(Z_t, Z_{t+0})$$
$$= \frac{1}{2} cos(\omega 0)$$

$$\begin{aligned} Cov(Z_t, Z_{t+k}) &= \frac{1}{2} E[\cos(\omega k) - \cos(\omega(2t+k) + 2\Theta)] \\ &= \frac{1}{2} E[\cos(\omega k)] - \frac{1}{2} E[\cos(\omega(2t+k) + 2\Theta)] \\ &= \frac{1}{2} \cos(\omega k) - \frac{1}{2} \int_{-\pi}^{\pi} \cos(\omega(2t+k) + 2\Theta) \frac{1}{2\pi} d\Theta \\ &= \frac{1}{2} \cos(\omega k) - \frac{1}{4\pi} \int_{-\pi}^{\pi} \cos(\omega(2t+k) + 2\Theta) d\Theta \\ &= \frac{1}{2} \cos(\omega k) - \frac{1}{4\pi} \times 0 = \frac{1}{2} \cos(\omega k) \end{aligned}$$

$$Var(Z_t) = Cov(Z_t, Z_t) = Cov(Z_t, Z_{t+0})$$

= $\frac{1}{2}cos(\omega 0) = \frac{1}{2}$

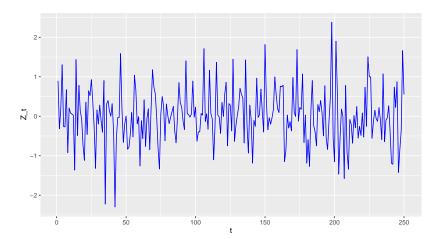
En la práctica, no se tiene la ecuación que describe el proceso estocástico, sino una realización del mismo, es decir una serie de tiempo, y a partir de este debemos llegar a algún modelo que describa el proceso.

En la práctica, no se tiene la ecuación que describe el proceso estocástico, sino una realización del mismo, es decir una serie de tiempo, y a partir de este debemos llegar a algún modelo que describa el proceso.

En R podemos simular una realización del proceso o en otras palabras una serie de tiempo:

```
n<-250 # Tamaño de la serie de tiempo
omega <- 3 # Damos un valor constante
t <- 1:n # Generamos el vector de tiempos t
# Escogemos cualquiere distribución con media 0
# y varianza 1 para A:
A <- rnorm(n, 0, 1)
# Simulamos Theta con una uniforme:
Theta <- runif(n, -pi, pi)
# Obtenemos el proceso Z_t:
Z_t <- A*sin(omega*t+Theta)
df <- data.frame(t, Z_t) # Generamos un data frame</pre>
```

```
require(tidyverse)
ggplot(df, aes(x=t, y=Z_t))+
  geom_line(col="blue")
```



 En la práctica, generalmente se trabaja con procesos estacionarios en covarianza. Este es un supuesto mucho menos restrictivo que la estacionaridad estricta y más fácil de probar en la práctica.

- En la práctica, generalmente se trabaja con procesos estacionarios en covarianza. Este es un supuesto mucho menos restrictivo que la estacionaridad estricta y más fácil de probar en la práctica.
- Un proceso estocástico es llamado Gaussiano o Normal si su función de distribución conjunta es normal. Considere el vector $\mathbf{Z} = [Z_1, Z_2, \dots, Z_n]^{\top}$ con n variables aleatorias de un proceso estocástico. Se dice que el proceso es Gaussiano si la función de densidad de probabilidad (f.d.p) del vector \mathbf{Z} es

$$f(\mathbf{z}) = f(z_1, z_2, \dots, z_n) = \frac{1}{(2\pi)^{n/2} |\Sigma|^{1/2}} e^{(\mathbf{z} - \mu)^{\top} \Sigma^{-1} (\mathbf{z} - \mu)}$$

donde $\mu = E(\mathbf{Z})$ es el vector de medias de \mathbf{Z} y Σ es la matriz de covarianzas.

- El proceso estocástico es llamado un proceso estocástico de valor real si sólo toma valores reales.
 Para un proceso estocástico de valor real definimos:
 - La función de medias del proceso: $\mu_t = E(Z_t)$.
 - La función de varianza del proceso: $\sigma_t^2 = E[(Z_t \mu_t)^2]$.
 - La función de covarianza entre Z_{t_1} y Z_{t_2} está dada por

$$\gamma(t_1, t_2) = E[(Z_{t_1} - \mu_{t_1})(Z_{t_2} - \mu_{t_2})]$$

• La función de correlación entre Z_{t_1} y Z_{t_2} está dada por

$$\rho(t_1,t_2) = \frac{\gamma(t_1,t_2)}{\sigma_{t_1}\sigma_{t_2}}$$

- Para un proceso estrictamente estacionario cuyos momentos de segundo orden existen, se cumple que:
 - Las funciones de medias del proceso $\mu_t = \mu$ es constante.
 - La función de varianzas del proceso $\sigma_t^2 = \sigma^2$ es contante.
 - Si $t_1 = t k$ y $t_2 = t$, la función de covarianza entre Z_{t_1} y Z_{t_2} es igual a

$$\gamma(t_1, t_2) = \gamma(t - k, t) = \gamma(t + k, t) = \gamma_k$$

es decir, solamente depende del número de periodos que separan a Z_{t_1} y Z_{t_2} y no del tiempo t.

• Similarmente, si $t_1 = t - k$ y $t_2 = t$, la función de correlación entre Z_{t_1} y Z_{t_2} es igual a

$$\rho(t_1, t_2) = \rho(t - k, t) = \rho(t + k, t) = \rho_k$$

y también depende solo del número de periodos que separan a Z_{t_1} y Z_{t_2} y no del tiempo t.

Función de autocovarianza

La función de autocovarianza de un proceso estacionario $\{Z_t\}$ se define como:

$$\gamma_k = Cov(Z_t, Z_{t+k}) = E\left[(Z_t - \mu)(Z_{t+k} - \mu)\right]$$

donde $\mu = E(Z_t)$, k es un número entero y γ_k depende solo de la separación entre las v.a. Z_t y Z_{t+k} .

Función de autocorrelación (ACF)

La función de autocorrelación (ACF por sus siglas en inglés) de un proceso estacionario $\{Z_t\}$ se define como:

$$\rho_k = Corr(Z_t, Z_{t+k}) = \frac{Cov(Z_t, Z_{t+k})}{\sqrt{Var(Z_t)Var(Z_{t+k})}} = \frac{\gamma_k}{\gamma_0}$$

donde $Var(Z_t) = Var(Z_{t+k}) = \gamma_0$ y k es un número entero. Al valor ρ_k lo llamaremos coeficiente de autocorrelación de orden k.

Algunas propiedades de γ_k y ρ_k :

Las funciones de autocovarianza y de autocorrelación cumplen con las siguientes propiedades:

- $\gamma_k = \gamma_{-k}$, para todo entero k, es decir, es simétrica con respecto a cero.
- $\rho_k = \rho_{-k}$, para todo entero k, es decir, es simétrica con respecto a cero. Esto implica que solo basta con calcular los valores para k > 0, ya que para valores negativos de k es simétrica.
- $\rho_0 = 1$.
- $|\rho_k| \leq 1$, para todo entero k.
- El gráfico de ρ_k es llamado **correlograma**.

Como vimos anteriormente, el proceso dado por la ecuación:

$$Z_t = A\sin(\omega t + \Theta)$$

donde A y Θ son variables aleatorias independientes tales que:

$$E(A) = 0$$
, $Var(A) = 1$, $\Theta \sim Uniforme(-\pi, \pi)$

y ω es una constante, cumple con ser estacionario, ya que:

$$E(Z_t) = 0$$
, $Var(Z_t) = \frac{1}{2}$ y $Cov(Z_t, Z_{t+k}) = \frac{1}{2}cos(\omega k)$

para todo entero t.

Luego,

$$\gamma_k = Cov(Z_t, Z_{t+k}) = \frac{1}{2}\cos(\omega k)$$

para todo entero k.

Luego,

$$\gamma_k = Cov(Z_t, Z_{t+k}) = \frac{1}{2}\cos(\omega k)$$

para todo entero k.

Además,

$$\rho_k = \frac{\gamma_k}{\gamma_0} = \frac{\cos(\omega k)/2}{1/2} = \cos(\omega k)$$

Luego,

$$\gamma_k = Cov(Z_t, Z_{t+k}) = \frac{1}{2}\cos(\omega k)$$

para todo entero k.

Además,

$$\rho_k = \frac{\gamma_k}{\gamma_0} = \frac{\cos(\omega k)/2}{1/2} = \cos(\omega k)$$

Así, tenemos que $|\rho_k| = |\cos(\omega k)| \le 1$ para todo entero k.

Luego,

$$\gamma_k = Cov(Z_t, Z_{t+k}) = \frac{1}{2}\cos(\omega k)$$

para todo entero k.

Además,

$$\rho_k = \frac{\gamma_k}{\gamma_0} = \frac{\cos(\omega k)/2}{1/2} = \cos(\omega k)$$

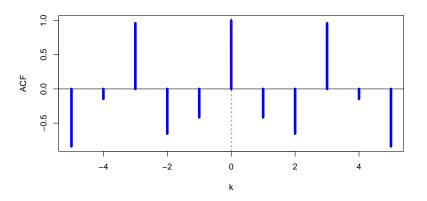
Así, tenemos que $|\rho_k| = |\cos(\omega k)| \le 1$ para todo entero k.

Además como la función coseno es una función par, entonces $\rho_{-k} = \cos(-\omega k) = \cos(\omega k) = \rho_k$.

Evaluando en algunos valores de k para un valor cualquiere de ω , por ejemplo $\omega=2$, tenemos que:

k	ACF
<u>-5</u>	-0.839
-4	-0.146
-3	0.960
-2	-0.654
-1	-0.416
0	1.000
1	-0.416
2	-0.654
3	0.960
4	-0.146
5	-0.839

El gráfico de la ACF sería:



Dada la simetría en k = 0, basta con graficarla para k > 0.

La función de autocorrelación parcial (PACF por sus siglas en inglés) de un proceso estacionario se define como la función de autocorrelación condicional:

$$\phi_{kk} = Corr(Z_t, Z_{t+k} \mid Z_{t+1}, \dots, Z_{t+k-1})$$

La función de autocorrelación parcial (PACF por sus siglas en inglés) de un proceso estacionario se define como la función de autocorrelación condicional:

$$\phi_{kk} = Corr(Z_t, Z_{t+k} \mid Z_{t+1}, \dots, Z_{t+k-1})$$

= $Corr(Z_{t+k} - \widehat{Z}_{t+k}, Z_t - \widehat{Z}_t)$

donde

$$\widehat{Z}_{t} = \beta_{1} Z_{t+1} + \beta_{2} Z_{t+2} + \dots + \beta_{k-1} Z_{t+k-1}$$

$$\widehat{Z}_{t+k} = \alpha_{1} Z_{t+k-1} + \alpha_{2} Z_{t+k-2} + \dots + \alpha_{k-1} Z_{t+1}$$

La pregunta que surge es ¿qué sentido tiene calcular la PACF?

La pregunta que surge es ¿qué sentido tiene calcular la PACF?

La ACF entre Z_t y Z_{t+k} tiene en cuenta la influencia lineal de los términos que están entre estas dos v.a.

$$Z_t$$
 Z_{t+1} Z_{t+2} \cdots Z_{t+k-1} Z_{t+k}

Para eliminar la influencia de dichas v.a. se plantea la PACF, definida como una correlación condicional entre Z_t y Z_{t+k} , teniendo como variables dadas:

$$Z_{t+1}$$
 Z_{t+2} \cdots Z_{t+k-1}

Para eliminar la influencia de dichas v.a. se plantea la PACF, definida como una correlación condicional entre Z_t y Z_{t+k} , teniendo como variables dadas:

$$Z_{t+1}$$
 Z_{t+2} \cdots Z_{t+k-1}

En el texto de William Wei (se encuentra en la carpeta de textos del curso como TEXTO 3 - TEÓRICO - Wei-Time series analysis) se encuentra una demostración paso a paso (páginas 11 a 15) que arroja la siguiente forma (o fórmulas) para la PACF:

Función de autocorrelación parcial (PACF): ϕ_{kk}

$$\phi_{kk} = \frac{\begin{vmatrix} 1 & \rho_1 & \rho_2 & \cdots & \rho_{k-2} & \rho_1 \\ \rho_1 & 1 & \rho_1 & \cdots & \rho_{k-3} & \rho_2 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ \rho_{k-1} & \rho_{k-2} & \rho_{k-3} & \cdots & \rho_1 & \rho_k \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 & \rho_2 & \cdots & \rho_{k-2} & \rho_{k-1} \\ \rho_1 & 1 & \rho_1 & \cdots & \rho_{k-3} & \rho_{k-2} \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ \rho_{k-1} & \rho_{k-2} & \rho_{k-3} & \cdots & \rho_1 & 1 \end{vmatrix}}$$
(2)

Para cada valor de k=1,2,3,... se obtiene la correlación parcial que en la mayoría de textos se denota por ϕ_{kk} .

ACF y PACF: Ejemplo 4

Considere el proceso o modelo estacionario: $X_t = c + \alpha X_{t-1} + w_t$ donde α y c son constantes y los w_t son independientes para todo t con $w_t \sim N(0, \sigma^2)$.

- Encuentre el valor de c en términos de α y $\mu = E(X_t)$.
- **2** Encuentre la función de autocovarianza $\gamma(k)$.
- Encuentre la ACF y la PACF.

ACF y PACF: Ejemplo 4 - Solución

Como el proceso es estacionario la media y la varianza son constantes, esto implica que:

$$E(X_t) = E(c + \alpha X_{t-1} + w_t)$$

ACF y PACF: Ejemplo 4 - Solución

Como el proceso es estacionario la media y la varianza son constantes, esto implica que:

$$E(X_t) = E(c + \alpha X_{t-1} + w_t) = c + \alpha E(X_{t-1}) + E(w_t)$$

ACF y PACF: Ejemplo 4 - Solución

Como el proceso es estacionario la media y la varianza son constantes, esto implica que:

$$E(X_t) = E(c + \alpha X_{t-1} + w_t) = c + \alpha E(X_{t-1}) + E(w_t)$$

Por tanto,

$$\mu = c + \alpha \mu + 0$$

Como el proceso es estacionario la media y la varianza son constantes, esto implica que:

$$E(X_t) = E(c + \alpha X_{t-1} + w_t) = c + \alpha E(X_{t-1}) + E(w_t)$$

Por tanto,

$$\mu = c + \alpha \mu + 0 \implies c = (1 - \alpha)\mu$$

Como el proceso es estacionario la media y la varianza son constantes, esto implica que:

$$E(X_t) = E(c + \alpha X_{t-1} + w_t) = c + \alpha E(X_{t-1}) + E(w_t)$$

Por tanto,

$$\mu = c + \alpha \mu + 0 \implies c = (1 - \alpha)\mu$$

Por otra parte,

$$Var(X_t) = Var(c + \alpha X_{t-1} + w_t)$$

Ocomo el proceso es estacionario la media y la varianza son constantes, esto implica que:

$$E(X_t) = E(c + \alpha X_{t-1} + w_t) = c + \alpha E(X_{t-1}) + E(w_t)$$

Por tanto,

$$\mu = c + \alpha \mu + 0 \implies c = (1 - \alpha)\mu$$

Por otra parte,

$$Var(X_t) = Var(c + \alpha X_{t-1} + w_t)$$

= $\alpha^2 Var(X_{t-1}) + Var(w_t) + 2\alpha Cov(X_{t-1}, w_t)$

Pero $Cov(X_{t-1}, w_t) = 0$. Ya que X_{t-1} depende de $w_{t-1}, w_{t-2}, w_{t-3}, \ldots$, y como los w_t son independientes entonces la covarianza entre ellos es igual a 0.

Así,
$$Var(X_t) = \alpha^2 Var(X_{t-1}) + Var(w_t)$$

Así,
$$Var(X_t) = \alpha^2 Var(X_{t-1}) + Var(w_t)$$
, de donde
$$\sigma_x^2 = \alpha^2 \sigma_x^2 + \sigma^2$$

Así,
$$Var(X_t) = \alpha^2 Var(X_{t-1}) + Var(w_t)$$
, de donde

$$\sigma_x^2 = \alpha^2 \sigma_x^2 + \sigma^2$$

Por tanto,

$$\gamma(0) = \sigma_x^2 = \frac{\sigma^2}{1 - \alpha^2}$$

- 2 La función de autocovarianza de un proceso estacionario está dada por:
 - Para k=1:

$$\gamma(1) = Cov(X_t, X_{t+1})$$

Así,
$$Var(X_t) = \alpha^2 Var(X_{t-1}) + Var(w_t)$$
, de donde

$$\sigma_x^2 = \alpha^2 \sigma_x^2 + \sigma^2$$

Por tanto,

$$\gamma(0) = \sigma_x^2 = \frac{\sigma^2}{1 - \alpha^2}$$

- 2 La función de autocovarianza de un proceso estacionario está dada por:
 - Para k = 1:

$$\gamma(1) = Cov(X_t, X_{t+1}) = Cov(X_t, c + \alpha X_t + w_{t+1})$$

Así,
$$Var(X_t) = \alpha^2 Var(X_{t-1}) + Var(w_t)$$
, de donde
$$\sigma_x^2 = \alpha^2 \sigma_x^2 + \sigma^2$$

Por tanto.

$$\gamma(0) = \sigma_x^2 = \frac{\sigma^2}{1 - \alpha^2}$$

- 2 La función de autocovarianza de un proceso estacionario está dada por:
 - Para k=1:

$$\gamma(1) = Cov(X_t, X_{t+1}) = Cov(X_t, c + \alpha X_t + w_{t+1})$$

= $Cov(X_t, c) + \alpha Cov(X_t, X_t) + Cov(X_t, w_{t+1})$

Así,
$$Var(X_t) = \alpha^2 Var(X_{t-1}) + Var(w_t)$$
, de donde
$$\sigma_x^2 = \alpha^2 \sigma_x^2 + \sigma^2$$

Por tanto.

$$\gamma(0) = \sigma_x^2 = \frac{\sigma^2}{1 - \alpha^2}$$

- 2 La función de autocovarianza de un proceso estacionario está dada por:
 - Para k=1:

$$\gamma(1) = Cov(X_t, X_{t+1}) = Cov(X_t, c + \alpha X_t + w_{t+1})$$

$$= Cov(X_t, c) + \alpha Cov(X_t, X_t) + Cov(X_t, w_{t+1})$$

$$= \alpha \gamma(0)$$

$$\gamma(2) = Cov(X_t, X_{t+2})$$

$$\gamma(2) = Cov(X_t, X_{t+2}) = Cov(X_t, c + \alpha X_{t+1} + w_{t+2})$$

$$\gamma(2) = Cov(X_t, X_{t+2}) = Cov(X_t, c + \alpha X_{t+1} + w_{t+2})
= Cov(X_t, c) + \alpha Cov(X_t, X_{t+1}) + Cov(X_t, w_{t+2})$$

$$\gamma(2) = Cov(X_t, X_{t+2}) = Cov(X_t, c + \alpha X_{t+1} + w_{t+2})
= Cov(X_t, c) + \alpha Cov(X_t, X_{t+1}) + Cov(X_t, w_{t+2})
= \alpha^2 \gamma(0)$$

• Para k = 2:

$$\gamma(2) = Cov(X_t, X_{t+2}) = Cov(X_t, c + \alpha X_{t+1} + w_{t+2})
= Cov(X_t, c) + \alpha Cov(X_t, X_{t+1}) + Cov(X_t, w_{t+2})
= \alpha^2 \gamma(0)$$

$$\gamma(3) = Cov(X_t, X_{t+3})$$

• Para k = 2:

$$\gamma(2) = Cov(X_t, X_{t+2}) = Cov(X_t, c + \alpha X_{t+1} + w_{t+2})
= Cov(X_t, c) + \alpha Cov(X_t, X_{t+1}) + Cov(X_t, w_{t+2})
= \alpha^2 \gamma(0)$$

$$\gamma(3) = Cov(X_t, X_{t+3}) = Cov(X_t, c + \alpha X_{t+2} + w_{t+3})$$

• Para k = 2:

$$\gamma(2) = Cov(X_t, X_{t+2}) = Cov(X_t, c + \alpha X_{t+1} + w_{t+2})
= Cov(X_t, c) + \alpha Cov(X_t, X_{t+1}) + Cov(X_t, w_{t+2})
= \alpha^2 \gamma(0)$$

$$\gamma(3) = Cov(X_t, X_{t+3}) = Cov(X_t, c + \alpha X_{t+2} + w_{t+3})
= Cov(X_t, c) + \alpha Cov(X_t, X_{t+2}) + Cov(X_t, w_{t+3})$$

• Para k = 2:

$$\gamma(2) = Cov(X_t, X_{t+2}) = Cov(X_t, c + \alpha X_{t+1} + w_{t+2})
= Cov(X_t, c) + \alpha Cov(X_t, X_{t+1}) + Cov(X_t, w_{t+2})
= \alpha^2 \gamma(0)$$

$$\gamma(3) = Cov(X_t, X_{t+3}) = Cov(X_t, c + \alpha X_{t+2} + w_{t+3})
= Cov(X_t, c) + \alpha Cov(X_t, X_{t+2}) + Cov(X_t, w_{t+3})
= \alpha^3 \gamma(0)$$

• Para k = 2:

$$\gamma(2) = Cov(X_t, X_{t+2}) = Cov(X_t, c + \alpha X_{t+1} + w_{t+2})
= Cov(X_t, c) + \alpha Cov(X_t, X_{t+1}) + Cov(X_t, w_{t+2})
= \alpha^2 \gamma(0)$$

• Para k = 3:

$$\gamma(3) = Cov(X_t, X_{t+3}) = Cov(X_t, c + \alpha X_{t+2} + w_{t+3})
= Cov(X_t, c) + \alpha Cov(X_t, X_{t+2}) + Cov(X_t, w_{t+3})
= \alpha^3 \gamma(0)$$

En general, teniendo en cuenta que la función de autocovarianza depende solo del la distancia,

$$\gamma(k) = \alpha^{|k|} \gamma(0), \quad \text{para} \quad k = \pm 1, \pm 2, \pm 3, \dots$$

$$\rho_k = \frac{\gamma(k)}{\gamma(0)}$$

$$\rho_k = \frac{\gamma(k)}{\gamma(0)} = \frac{\alpha^{|k|}\gamma(0)}{\gamma(0)}$$

$$\rho_k = \frac{\gamma(k)}{\gamma(0)} = \frac{\alpha^{|k|}\gamma(0)}{\gamma(0)} = \alpha^{|k|}$$

La ACF está dada por:

$$\rho_k = \frac{\gamma(k)}{\gamma(0)} = \frac{\alpha^{|k|}\gamma(0)}{\gamma(0)} = \alpha^{|k|}$$

$$\phi_{11} = \rho_1 = \alpha$$

La ACF está dada por:

$$\rho_k = \frac{\gamma(k)}{\gamma(0)} = \frac{\alpha^{|k|}\gamma(0)}{\gamma(0)} = \alpha^{|k|}$$

$$\phi_{11}=\rho_1=\alpha$$

$$\phi_{22} = \frac{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & \rho_2 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{vmatrix}}$$

La ACF está dada por:

$$\rho_k = \frac{\gamma(k)}{\gamma(0)} = \frac{\alpha^{|k|}\gamma(0)}{\gamma(0)} = \alpha^{|k|}$$

$$\phi_{11} = \rho_1 = \alpha$$

$$\phi_{22} = \frac{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & \rho_2 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{vmatrix}} = \frac{\rho_2 - \rho_1^2}{1 - \rho_1^2}$$

La ACF está dada por:

$$\rho_k = \frac{\gamma(k)}{\gamma(0)} = \frac{\alpha^{|k|}\gamma(0)}{\gamma(0)} = \alpha^{|k|}$$

$$\phi_{11} = \rho_1 = \alpha$$

$$\phi_{22} = \frac{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & \rho_2 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{vmatrix}} = \frac{\rho_2 - \rho_1^2}{1 - \rho_1^2} = \frac{\alpha^2 - \alpha^2}{1 - \alpha^2} = 0$$

La ACF está dada por:

$$\rho_k = \frac{\gamma(k)}{\gamma(0)} = \frac{\alpha^{|k|}\gamma(0)}{\gamma(0)} = \alpha^{|k|}$$

$$\phi_{11} = \rho_1 = \alpha$$

$$\phi_{22} = \frac{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & \rho_2 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{vmatrix}} = \frac{\rho_2 - \rho_1^2}{1 - \rho_1^2} = \frac{\alpha^2 - \alpha^2}{1 - \alpha^2} = 0$$

$$\phi_{33} = \begin{vmatrix} 1 & \rho_1 & \rho_1 \\ \rho_1 & 1 & \rho_2 \\ \rho_2 & \rho_1 & \rho_3 \\ \hline 1 & \rho_1 & \rho_2 \\ \rho_1 & 1 & \rho_1 \\ \rho_2 & \rho_1 & 1 \end{vmatrix}$$

$$\phi_{33} = \frac{\begin{vmatrix} 1 & \rho_1 & \rho_1 \\ \rho_1 & 1 & \rho_2 \\ \rho_2 & \rho_1 & \rho_3 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 & \rho_2 \\ \rho_1 & 1 & \rho_1 \\ \rho_2 & \rho_1 & 1 \end{vmatrix}} = \frac{1 \times (\rho_3 - \rho_1 \rho_2) - \rho_1 \times (\rho_1 \rho_3 - \rho_2^2) + \rho_1 \times (\rho_1^2 - \rho_2)}{1 \times (1 - \rho_1^2) - \rho_1 \times (\rho_1 - \rho_1 \rho_2) + \rho_2 \times (\rho_1^2 - \rho_2)}$$
$$= \frac{(\alpha^3 - \alpha^3) - \alpha \times (\alpha^4 - \alpha^4) + \alpha \times (\alpha^2 - \alpha^2)}{1 \times (1 - \alpha^2) - \alpha \times (\alpha - \alpha^3) + \alpha^2 \times (\alpha^2 - \alpha^2)}$$

$$\phi_{33} = \frac{\begin{vmatrix} 1 & \rho_1 & \rho_1 \\ \rho_1 & 1 & \rho_2 \\ \rho_2 & \rho_1 & \rho_3 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 & \rho_2 \\ \rho_1 & 1 & \rho_1 \\ \rho_2 & \rho_1 & 1 \end{vmatrix}} = \frac{1 \times (\rho_3 - \rho_1 \rho_2) - \rho_1 \times (\rho_1 \rho_3 - \rho_2^2) + \rho_1 \times (\rho_1^2 - \rho_2)}{1 \times (1 - \rho_1^2) - \rho_1 \times (\rho_1 - \rho_1 \rho_2) + \rho_2 \times (\rho_1^2 - \rho_2)}$$
$$= \frac{(\alpha^3 - \alpha^3) - \alpha \times (\alpha^4 - \alpha^4) + \alpha \times (\alpha^2 - \alpha^2)}{1 \times (1 - \alpha^2) - \alpha \times (\alpha - \alpha^3) + \alpha^2 \times (\alpha^2 - \alpha^2)} = 0$$

En general, se puede probar de manera sucesiva que $\phi_{kk}=0$ para k>1.