REVISTA PROYECCIONES N° 6: 103-114.
DICIEMBRE 1983.

SOBRE PREDICCIÓN EN UNA SERIE DE TIEMPO CON COMPONENTES PERIÓDICOS

ENRIQUE RAMOS M.*

RESUMEN:

Se dispone de una serie cronológica observada en tiempo discreto y se desea predecir valores futuros de ésta. Bajo sospecha de que la serie presenta componentes periódicos a frecuencias conocidas, es natural sugerir un modelo sinusoidal. Esto indica que dicha serie no debe presentar ningún tipo de tendencia y que las observaciones sean hechas en tiempos equiespaciadas.

* Profesor Depto. Matemáticas, Facultad de Ciencias, Universidad del Norte.

1. INTRODUCCION.

La importancia del problema de predicción se manifiesta en diferentes áreas de la ciencia y consiste en conocer valores futuros, a partir de una secuencia de valores observados, de algún fenómeno físico que sucede en el tiempo.

La predicción puede realizarse subjetivamente o a través de procedimientos matemáticos. Es hecha subjetivamente cuando se utiliza el juicio, la intuición y en general cualquier otra información relevante. En cambio son procedimientos matemáticos los que se apoyan en una teoría de esa naturaleza. Una teoría matemática permite describir fenómenos físicos a través de modelos matemáticos, los cuales a su vez serían empleados para algún objetivo previamente propuesto. En el caso de la predicción, el modelo elegido resultará bueno si ajustado a partir de su pasado se revela preciso y los valores a predecir dependerán de este modelo.

2. ALGUNOS CONCEPTOS A USAR.

2.1. Series de tiempo:

Una serie de tiempo es una sucesión de observaciones tomadas secuencialmente en el tiempo, y se puede considerar como una muestra del conjunto infinito de series que podrían haberse observado.

Un concepto más amplio es el de proceso estocástico, que se puede describir como un fenómeno estadístico que se desarrolla en el tiempo de acuerdo a leyes probabilísticas dadas. Bajo este contexto, se considera a cada miembro del conjunto infinito de series de tiempo como una posible realización de un proceso estocástico. Más aún, la serie de tiempo observada puede pensarse como una realización particular del proceso.

2.2. Series de Fourier:

El análisis de Fourier trata básicamente sobre la aproximación de una función, continua o discreta, por una suma de términos seno y coseno, llamada representación en series de Fourier. Esto es, si $f(t)$ está definida en el intervalo $(-\pi, \pi)$ entonces bajo las condiciones de Dirichlet^(*), puede aproximarse por la serie de Fourier:

$$\frac{a_0}{2} + \sum_{r=1}^m (a_r \cos rt + b_r \sin rt)$$

donde:

$$a_0 = \frac{1}{\pi} \int_{-\pi}^{\pi} f(t) dt$$

$$a_r = \frac{1}{\pi} \int_{-\pi}^{\pi} f(t) \cos rt dt$$

$$b_r = \frac{1}{\pi} \int_{-\pi}^{\pi} f(t) \sin rt dt \quad r = 1, 2, \dots$$

son los coeficientes de Fourier. Se demuestra que esta serie converge a $f(t)$ cuando $m \rightarrow \infty$, excepto en los puntos de discontinuidades donde converge a:

$$\frac{1}{2} [f(t^-) + f(t^+)]$$

2.3. Modelo sinusoidal:

La ecuación trigonométrica $f(t) = R \cos (wt + \theta)$ donde: R , $\frac{2\pi}{w}$, θ y w son amplitud, período, fase y frecuencia de f respectivamente, permite definir un modelo matemático para el estudio de ciertas series de tiempo: "Aquellas que solo presenten componentes periódicas y no posean ningún tipo de tendencias". Si $y_t = f(t)$ es una función discreta; o sea, definida solamente en los enteros $t = 1, 2, \dots, n$; que describe una de estas series, entonces:

(*) $f(t)$ es absolutamente integrable en $(-\pi, \pi)$, tiene un número finito de discontinuidades y un número finito de máximos y mínimos.

$$y_t = u + R \cos (wt + \theta) + e_t ; \quad t = 1, 2, \dots, n$$

será el modelo apropiado para su estudio y se llama "modelo sinusoidal". Aquí e_t es el error traído por las observaciones y_t (aleatorio). La generalización de este modelo es:

$$y_t = u + \sum_{r=1}^m R_r \cos (w_r t + \theta_r) + e_t ; \quad t = 1, 2, \dots, n$$

y será aquí que centraremos nuestra atención, ya que el modelo de regresión que aquí se expone emerge naturalmente de este, y no es mas que una serie de Fourier discreta que ajusta a y_t .

3. PRESENTACION DEL MODELO.

Supongamos una serie de tiempo $\{y_t/t \in T\}$ observada en tiempo discreto ($T = \mathbb{Z}$) que presente solo componentes periódicas y no tenga ningún tipo de tendencias. Denotemos por $y_1, \dots, y_n$ una secuencia de valores observados (medibles) de esta serie, los que se suponen tomados en tiempos equiespaciados. Bajo este contexto, nuestro problema consistirá en ajustar a las observaciones una serie de Fourier discreta de la forma:

$$y_t = u + \sum_{r=1}^m (a_r \cos \frac{2\pi}{n} t_r + b_r \sin \frac{2\pi}{n} t_r) + e_t^{(*)} \quad (1)$$

en los tiempos $t = 1, 2, \dots, n$. Esta ecuación es llamada "Modelo de regresión armónica" y está determinado por los $2m+1$ parámetros: $u, a_1, \dots, a_m, b_1, \dots, b_m$.

La teoría precedente (especialmente el análisis de varianza) se desarrolla bajo la suposición de que los errores e_t , traídos por las observaciones, son aleatorios e independientes en el sentido estocástico y distribuidos normalmente con esperanza cero y varianza σ^2 .

Con el fin de simplificar el trabajo, las n -ecuaciones generadas por la fórmula (1) se pueden escribir matricialmente como:

(*) Es una expresión equivalente del modelo sinusoidal.

$$\vec{Y} = X\vec{A} + \vec{e}, \quad (2)$$

donde $\vec{Y}$, $\vec{A}$ y $\vec{e}$ son vectores de observaciones, parámetro y error respectivamente, es decir:

$$\begin{aligned} \vec{Y} &= (y_1, \dots, y_n)^T \\ \vec{A} &= (u, a_1, \dots, a_m, b_1, \dots, b_m)^T \\ \vec{e} &= (e_1, \dots, e_n)^T \end{aligned} \quad (3)$$

y X matriz de diseño que caracteriza al modelo, cuyas entradas son de la forma:

$$x_{tj} = \begin{cases} 1 & \text{si } j = 1 \\ \cos \frac{2\pi}{n} t \frac{j}{2} & \text{si } j \text{ par} \\ \sin \frac{2\pi}{n} t \frac{j-1}{2} & \text{si } j \text{ impar, } j > 1 \end{cases} \quad (4)$$

con $t = 1, 2, \dots, n$ y $j = 1, 2, \dots, 2m+1$. Suponemos que en (1) se tienen más ecuaciones que parámetros, es decir, el rango de la matriz de diseño será igual al número de sus columnas. Esto equivale a decir que el modelo (2) o su equivalente (1), es un "modelo lineal de rango máximo".

4. ESTIMACION DE LOS PARAMETROS.

El primer objetivo frente al modelo será estimar los $2m+1$ parámetros: $u, a_1, \dots, a_m, b_1, \dots, b_m$, para lo cual será indispensable tener en cuenta toda la información disponible. Para tal efecto será suficiente recurrir al método de mínimos cuadrados.

4.1. Principio de Mínimos Cuadrados:

Representemos por $\mathbb{R}^n$ el espacio de observaciones y $\mathbb{R}^{2m+1}$ el espacio de parámetros (no olvidemos que estamos bajo el supuesto $2m+1 < n$). Los vectores columnas de la matriz de diseño son linealmente independiente por lo tanto generan un subespacio S , de dimensión $2m+1$, de $\mathbb{R}^n$. El principio de los mínimos cuadrados consiste en encontrar un vector

$\vec{\alpha} = (\hat{u}, \hat{a}_1, \dots, \hat{a}_m, \hat{b}_1, \dots, \hat{b}_m)^T$ en $\mathbb{R}^{2m+1}$, tal que $X\vec{\alpha}$ ajuste $\vec{y}$, ésto es, tomar $X\vec{\alpha}$ en S lo más cercano de $\vec{y}$ (en este caso nos conformaremos con la métrica usual). Haciendo $X\vec{\alpha}$ igual al "vector proyección ortogonal de $\vec{y}$ sobre S ", se obtiene lo querido, esto es:

$$\min_{\vec{z} \in S} \|\vec{y} - \vec{z}\|^2 = \|\vec{y} - X\vec{\alpha}\|^2 \quad (5)$$

o sea,

$$\langle \vec{y} - X\vec{\alpha}, X\vec{\alpha} \rangle = 0 \quad (5)$$

cuyo desarrollo conduce a las conocidas "ecuaciones normales":

$$X^T X \vec{\alpha} = X^T \vec{y}. \quad (7)$$

Puesto que X es de rango máximo, existe la inversa de $X^T X$, la que permite resolver (7), obteniendo:

$$\vec{\alpha} = (X^T X)^{-1} X^T \vec{y} \quad (8)$$

que será fácil de computar si $X^T X$ fuese una matriz diagonal.

4.2. Ortogonalidad en la Matriz de Diseño.

La matriz $X^T X$ es diagonal si X es ortogonal, o sea, si tiene sus vectores columnas ortogonales entre sí.

En vista de los siguientes, bien conocidos, resultados trigonométricos:

$$\sum_{t=1}^n \cos \frac{2\pi}{n} t j = \begin{cases} n & \text{si } j = 0 \\ 0 & \text{si } j = 1, 2, \dots \end{cases}$$

$$\sum_{t=1}^n \sin \frac{2\pi}{n} t j = 0 \quad \text{si } j = 0, 1, 2, \dots$$

$$\sum_{t=1}^n \cos \frac{2\pi}{n} tj \cos \frac{2\pi}{n} tk = \begin{cases} n & \text{si } j = k = 0 \\ \frac{n}{2} & \text{si } j = k \neq 0 \\ 0 & \text{si } j \neq k \end{cases}$$

$$\sum_{t=1}^n \sin \frac{2\pi}{n} tj \sin \frac{2\pi}{n} tk = \begin{cases} 0 & \text{si } j = k = 0 \\ \frac{n}{2} & \text{si } j = k \neq 0 \\ 0 & \text{si } j \neq k \end{cases}$$

$$\sum_{t=1}^n \cos \frac{2\pi}{n} tj \sin \frac{2\pi}{n} tk = 0$$

Podemos concluir fácilmente que los vectores columnas de la matriz de diseño son ortogonales, o sea, para $j \neq k$ tenemos:

$$\langle \vec{X}^{(j)}, \vec{X}^{(k)} \rangle = \begin{cases} \sum_{t=1}^n \cos \frac{2\pi}{n} t \cdot \frac{k}{2} = 0 & \text{si } j=1, k \text{ par} \\ \sum_{t=1}^n \sin \frac{2\pi}{n} t \cdot \frac{k-1}{2} = 0 & \text{si } j=1, k \text{ impar } (>1) \\ \sum_{t=1}^n \cos \frac{2\pi}{n} \cdot \frac{j}{2} \cos \frac{2\pi}{n} \cdot \frac{k}{2} & \text{si } j, k \text{ pares} \\ \sum_{t=1}^n \sin \frac{2\pi}{n} \cdot \frac{j-1}{2} \cos \frac{2\pi}{n} \cdot \frac{k-1}{2} = 0 & \text{si } j, k \text{ impares} \\ \sum_{t=1}^n \cos \frac{2\pi}{n} \cdot \frac{j}{2} \sin \frac{2\pi}{n} \cdot \frac{k-1}{2} = 0 & \text{si } j \text{ par, } k \text{ impar } (>1) \end{cases} \quad (10)$$

donde $\vec{X}^{(i)}$ es el i -ésimo vector columna de la matriz de diseño. Ahora si $j = k$ de (9) obtenemos:

$$||\vec{X}^{(j)}||^2 = \begin{cases} n & \text{si } j = 1 \\ \sum_{t=1}^n \cos^2 \frac{2\pi}{n} t \cdot \frac{j}{2} = \frac{n}{2} & \text{si } j \text{ es par} \\ \sum_{t=1}^n \sin^2 \frac{2\pi}{n} t \cdot \frac{j-1}{2} = \frac{n}{2} & \text{si } j \text{ es impar} \end{cases} \quad (11)$$

resultando:

$$X^T X = \text{diag} \left\{ n, \frac{n}{2}, \dots, \frac{n}{2} \right\}, \quad (12)$$

cuya inversa, muy fácil de calcular, será:

$$(X^T X)^{-1} = \text{diag} \left\{ \frac{1}{n}, \frac{2}{n}, \dots, \frac{2}{n} \right\} \quad (13)$$

lo que junto con:

$$X^T Y = \begin{bmatrix} \sum_{t=1}^n y_t \\ \sum_{t=1}^n y_t \cos \frac{2\pi}{n} t \cdot 1 \\ \sum_{t=1}^n y_t \sin \frac{2\pi}{n} t \cdot 1 \\ \dots \dots \dots \sum_{t=1}^n y_t \cos \frac{2\pi}{n} t \cdot m \\ \sum_{t=1}^n y_t \sin \frac{2\pi}{n} t \cdot m \end{bmatrix} \quad (14)$$

y en virtud de (8) arroja las siguientes estimaciones para los parámetros del modelo:

$$\begin{aligned}
\hat{u} &= \frac{1}{n} \sum_{t=1}^n y_t \\
\hat{a}_r &= \frac{2}{n} \sum_{t=1}^n y_t \cos \frac{2\pi}{n} t \cdot r \\
\hat{b}_r &= \frac{2}{n} \sum_{t=1}^n y_t \sin \frac{2\pi}{n} t \cdot r \quad r = 1, 2, \dots, n
\end{aligned} \tag{15}$$

5. ECUACION DE PREDICCIÓN.

Estimados todos los parámetros del modelo, contamos con la ecuación de regresión:

$$\hat{y}_t = \hat{u} + \sum_{r=1}^m (\hat{a}_r \cos \frac{2\pi}{n} t \cdot r + \hat{b}_r \sin \frac{2\pi}{n} t \cdot r) \tag{16}$$

la que se puede usar con el fin de predecir valores futuros de la serie en estudio. El uso de esta ecuación, como ecuación de predicción, en muchas ocasiones resulta caro por la cantidad de información de que se dispone. Surge entonces la siguiente pregunta: ¿podemos obtener una ecuación equivalente a (16) con el fin de predecir y con un costo mínimo? La respuesta es SI. Efectivamente, en el modelo lineal de rango máximo, cuando la matriz de diseño posee la propiedad de ortogonalidad entre sus vectores columnas, podemos utilizar el método de eliminar las variables (partiendo de la regresión completa) cuyos parámetros no son significativos para el modelo. Esto es, si $\vec{\epsilon} = \vec{Y} - X\vec{\alpha}$ es el vector de errores estimados y $\vec{\alpha} = (\alpha_1, \dots, \alpha_{2m+1})$ el vector de parámetros estimados, entonces bajo el supuesto de normalidad para $\vec{\epsilon}$ podemos obtener el indicador:

$$F_j = q \frac{||x^{(j)}||^2 \alpha_j^2}{||\vec{\epsilon}||^2}, \quad j = 1, 2, \dots, 2m+1 \tag{17}$$

donde q es la dimensión del subespacio de $\mathbb{R}^n$, "complemento ortogonal de S ", definido por el núcleo de la transformación lineal generada por la matriz de diseño transpuesta; este indicador permite decir que parámetro es significativo para el modelo, a un nivel de significación β , al

compararlo con una cuantila $F(1, q)$ de la distribución EFE.

Si $S^\perp$ denota el complemento ortogonal de S en $\mathbb{R}^n$, entonces $\mathbb{R}^n$ se puede descomponer en suma directa de estos dos subespacios, lo que garantiza:

$$||\vec{Y}||^2 = ||\vec{\varepsilon}||^2 + ||X\vec{\alpha}||^2 \quad (18)$$

(teorema de Pitágoras en $\mathbb{R}^n$). Bajo este contexto, el indicador (17) para el modelo de regresión armónica será:

$$F_j = \begin{cases} \frac{n(n-2m-1) \hat{u}^2}{||\vec{\varepsilon}||^2} & \text{si } j = 1 \\ \frac{n(n-2m-1) \hat{a}_r^2}{2||\vec{\varepsilon}||^2} & \text{si } j = 2r \\ \frac{n(n-2m-1) \hat{b}_r^2}{2||\vec{\varepsilon}||^2} & \text{si } j = 2r+1 \end{cases} \quad (19)$$

$r = 1, 2, \dots, m$

donde:

$$||\vec{\varepsilon}||^2 = \sum_{t=1}^n y_t^2 - n\hat{u}^2 - \frac{n}{2} \sum_{r=1}^m (\hat{a}_r^2 + \hat{b}_r^2) \quad (20)$$

Pero en este tipo de regresión, interesa más determinar la significación de la "amplitud" de cada senoide que compone el modelo, que determinar la significación de cada parámetro por separado; tenemos entonces:

5.1. Significación de las Amplitudes.

Consideremos m componentes cíclicas. Del modelo en estudio es fácil obtener las amplitudes R_r de cada una de estas componentes, las que asumen la forma:

$$R_r = \sqrt{a_r^2 + b_r^2}, \quad r = 1, 2, \dots, m \quad (21)$$

y por lo tanto sus estimadores:

$$\hat{R}_r^2 = \hat{a}_r^2 + \hat{b}_r^2, \quad r = 1, 2, \dots, m \quad (22)$$

Como cada R_r contiene información existente en dos subespacios de dimensión 1 correspondiente a a_r y b_r , le asignamos dos grados de libertad a cada una de estas amplitudes. Si $p_r = \frac{n}{r}$, $r = 1, 2, \dots, m$ son los períodos de las sinusoides, entonces "Rechazamos la posibilidad de que la componente cíclica de período p_r no es significativa, al nivel de significación β , si el indicador:

$$F_r = \frac{\frac{n}{4} (n-2m-1) \hat{R}_r^2}{\sum_{t=1}^n y_t^2 - n\hat{u}^2 - \frac{n}{2} \sum_{r=1}^m R_r^2} \quad r = 1, 2, \dots, m \quad (23)$$

es mayor que la cuantila $F_{\beta}(2, n-2m-1)$ de la distribución EFE".

Con estos resultados ya podemos contar con una buena ecuación de predicción para el estudio de series de tiempo que presenten componentes periódicas y no posean ningún tipo de tendencias.

BIBLIOGRAFIA.

- [1] CHATFIELD, C: The Analysis of Time Serie: Theory and Practice, John Wiley and Son. New York, 1975.
- [2] RAMOS, E. : Ideas elementales sobre el modelo lineal general de rango máximo.
2^a Jornadas de Matemáticas,
Universidad del Norte, Agosto 1983.