

Capítulo 1

Análisis descriptivo de una serie

Métodos clásicos para describir y analizar una serie temporal:

1. *Análisis de tendencias deterministas*: se basa en métodos de regresión; es poco flexible y no incluye correlaciones entre observaciones consecutivas. Los errores de predicción de valores futuros suelen ser grandes.
2. *Métodos de alisado*: dan mayor peso a las observaciones más recientes. Esto es útil para la predicción.
3. *Métodos de descomposición*: modelizan la tendencia y la estacionalidad por separado.
4. *Métodos armónicos*: representan una serie como una suma de funciones periódicas; el periodograma permite detectar periodos deterministas.

Estos son los métodos más simples para el Análisis de Series Temporales. Aunque no sean óptimos, ayudan a describir el comportamiento de la serie y son la base de procedimientos más avanzados.

1.1. Análisis de tendencias deterministas

- ✓ Para series con tendencia determinista;
- ✓ Estos métodos consisten en ajustar una función de regresión a la serie que explique la tendencia.

Modelo: Suponemos que la serie se puede descomponer de la forma:

$$y_t = \mu_t + a_t, \quad t = 1, \dots, n.$$

- ✓ μ_t **nivel** determinista de la serie en el instante t .
- ✓ a_t **error** aleatorio (o **innovación**) en el instante t : recoge los efectos de todos los factores desconocidos o no tenidos en cuenta que afectan a la serie.

Hipótesis básicas: En esta sección asumimos:

- ✓ $\mu_t = g(t, \boldsymbol{\beta})$: El nivel μ_t es una función conocida del tiempo y de un vector de parámetros desconocidos $\boldsymbol{\beta}$;
- ✓ $a_t \sim \text{iid } N(0, \sigma^2)$: La distribución de los errores a_t no varía con el tiempo.

Estimación: Habitualmente $\boldsymbol{\beta}$ se obtiene utilizando el método de mínimos cuadrados (**MMCC**); es decir, resolviendo el problema:

$$\min_{\boldsymbol{\beta}} \sum_{t=0}^n [y_t - g(t, \boldsymbol{\beta})]^2.$$

Derivando la función $f(\boldsymbol{\beta}) = \sum_{t=0}^n [y_t - g(t, \boldsymbol{\beta})]^2$ respecto de los elementos de $\boldsymbol{\beta}$ e igualando a cero, se obtienen las **ecuaciones normales**. El estimador mínimo cuadrático $\hat{\boldsymbol{\beta}}$ es la solución a dichas ecuaciones.

La tendencia estimada es $\hat{\mu}_t = g(t, \hat{\boldsymbol{\beta}})$, donde $\hat{\boldsymbol{\beta}}$ es un estimador de $\boldsymbol{\beta}$.

Predicción:

- ✓ Notación: $\hat{y}_n(h)$ predicción de la serie en el instante $n + h$, y_{n+h} , a partir de los datos hasta n , (y_1, \dots, y_n) .

- ✓ Predicción: Habitualmente se hace extrapolando el nivel estimado de la serie,

$$\hat{y}_n(h) = \hat{\mu}_{n+h} = g(n+h, \hat{\beta}), \quad h = 1, 2, \dots$$

- ✓ Supongamos que los parámetros β y σ^2 son conocidos, es decir, $\mu_t = g(t, \beta)$ es conocido. Entonces, como los errores a_1, \dots, a_n son independientes y el nivel μ_t no depende de los instantes anteriores, los valores futuros y_{n+1}, y_{n+2}, \dots no dependen de los valores históricos de la serie y_1, \dots, y_n .

1.1.1. Modelo con nivel constante (o sin tendencia)

Modelo: Suponemos que la serie varía alrededor de un nivel constante:

$$\mu_t = \mu \quad \Rightarrow \quad \dots$$

Esto es equivalente a asumir

$$y_t \sim \text{iid } N(\dots, \dots), \quad t = 1, \dots, n.$$

Estimación:

$$\hat{\mu} = \dots, \quad \hat{\sigma}^2 = \dots$$

Predicción:

$$\hat{y}_n(k) = \hat{\mu}_{n+k} = \dots, \quad \forall k = 1, 2, \dots$$

Ejemplo 1 *Modelo de nivel constante para la serie 1) del crecimiento del alquiler:*

Nivel estimado:

$$\dots,$$

Predicción del crecimiento en el próximo mes:

$$\dots,$$

Varianza estimada:

$$\dots,$$

1.1.2. Modelo de tendencia lineal

Modelo: Se supone que la serie varía a lo largo del tiempo con tendencia constante, bien creciente o bien decreciente:

$$\mu_t = \beta_0 + \beta_1 t \quad \Rightarrow \quad \dots$$

- ✓ β_0 nivel inicial (en $t = 0$) de la serie.
- ✓ β_1 pendiente de la recta que describe el nivel de la serie

Serie de incrementos: Al calcular incrementos desaparece la tendencia

$$\Delta y_t = y_t - y_{t-1} = \dots \quad .$$

Llamamos $z_t = \Delta y_t$ a la serie de incrementos y $a_t^* = \Delta a_t$ a sus errores:

- ✓ La serie de incrementos z_t no presenta tendencia:
 $z_t = \beta_1 + a_t^*$;
- ✓ La serie de incrementos z_t tiene errores con el doble de varianza,
 $a_t^* \sim \text{iid } N(\dots, \dots)$;
- ✓ Incremento esperado de la serie entre dos instantes consecutivos:
 $E[\Delta y_t] = \beta_1$.

Estimación: Por MMCC se obtiene

$$\hat{\beta}_1 = \dots \quad , \quad \hat{\beta}_0 = \dots \quad .$$

Residuos:

$$\hat{a}_t = y_t - \hat{\mu}_t = y_t - \hat{\beta}_0 - \hat{\beta}_1 t, \quad t = 1, \dots, n.$$

Estimación insesgada de σ^2 :

$$\hat{\sigma}^2 = \frac{1}{n-2} \sum_{t=1}^n \hat{a}_t^2 = \frac{1}{n-2} \sum_{t=1}^n (y_t - \hat{\beta}_0 - \hat{\beta}_1 t)^2.$$

Predicción:

$$\hat{y}_n(k) = \dots$$

Ejemplo 2 *Correlación entre tiempo y el índice del precio del alquiler: 0,997.*

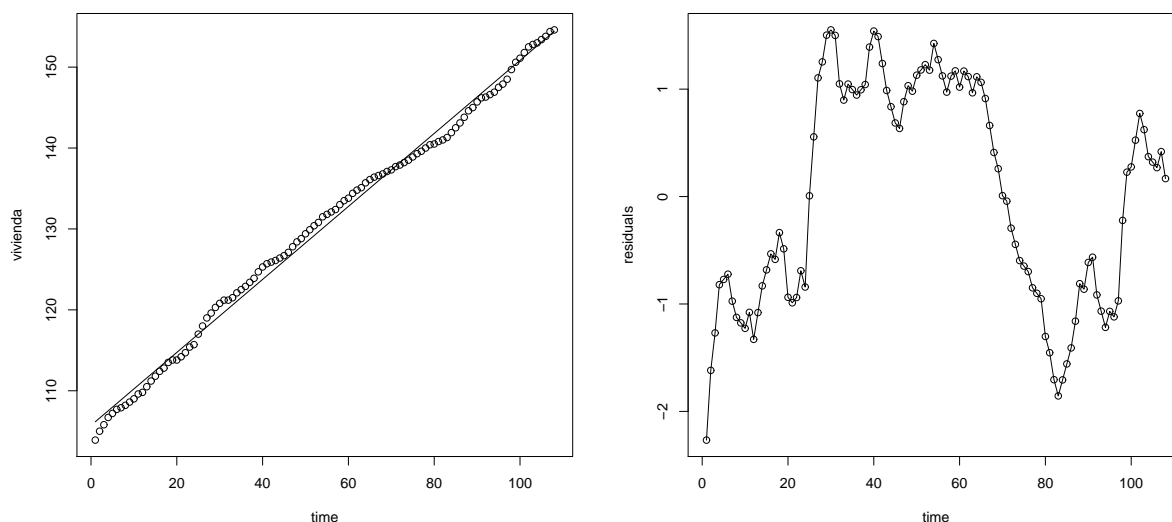
Recta ajustada:

$$y_t = 130,3 + 0,451 t, \quad t = -53,5, \dots, 53,5, \quad \hat{\sigma}^2 = 1,046.$$

El índice del alquiler se ha incrementado en promedio en 0,451 cada mes.

El gráfico de residuos revela defectos del modelo, ya que tienen cierta estructura. Aunque la tendencia parece aproximadamente lineal, las predicciones de este modelo pueden ser malas por la curvatura de la serie. Encontrar una tendencia paramétrica más apropiada parece difícil.

Figura 1.1: Precio del alquiler con tendencia lineal ajustada (izq.) y residuos (der.)



1.1.3. Modelo de tendencia polinómica

Modelo: El nivel de la serie evoluciona según un polinomio de orden r :

$$y_t = \beta_0 + \beta_1 t + \dots + \beta_r t^r + a_t,$$

Estimación: Método de MMCC. Llamamos

$$X = \begin{pmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & 2 & 4 & \cdots & 2^r \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & n & n^2 & \cdots & n^r \end{pmatrix}, \quad \mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$$

Estimador de β :

...

Estimador insesgado de σ^2 :

...

Predicción:

...

1.1.4. Limitaciones de los modelos con tendencia determinística

- ✓ Las series con nivel constante son frecuentes, pero con tendencia lineal o polinómica exacta son raras, ya que la mayoría de series reales evolucionan como resultado de un conjunto complejo de causas. Estos modelos se deben utilizar solamente para predecir muy a corto plazo.
- ✓ Se puede mejorar el ajuste utilizando tendencias lineales por tramos: cortar la serie en tramos que tengan aproximadamente tendencia o nivel constante y ajustar en cada tramo un modelo lineal o constante.
- ✓ La predicción del crecimiento $y_{t+1} - y_t$ en el modelo de tendencia lineal es la pendiente estimada $\hat{\beta}_1$. Dicha estimación es una media ponderada de los crecimientos en el periodo muestral con los pesos mayores en el centro y los pesos menores en los extremos, cosa que parece poco razonable.

Ejemplo 3 $n = 5$ observaciones en los instantes $t = -2, -1, 0, 1, 2$. En este caso $\bar{t} = 0$ y entonces el estimador de la pendiente es:

...

Calculamos el numerador y el denominador de dicho estimador:

$$\dots = \dots$$

$$\dots = \dots$$

Reemplazando $y_{-1} = 3y_{-1} - 2y_{-1}$ e $y_1 = 3y_1 - 2y_1$, obtenemos

$$\sum_{t=-2}^2 t y_t = 2(y_{-1} - y_{-2}) + 3(y_0 - y_{-1}) + 3(y_1 - y_0) + 2(y_2 - y_1).$$

Llamando $\Delta y_t = y_t - y_{t-1}$ al crecimiento en el instante t , obtenemos

$$\hat{\beta}_1 = \sum_{t=-1}^2 w_t \Delta y_t,$$

donde los pesos en el centro son mayores que en los extremos. Para una serie de cien años, el resultado sería análogo: el crecimiento más reciente recibiría el mismo peso que el crecimiento hace cien años.

En general, $\hat{\beta}_1$ es una media ponderada de los incrementos de la serie, es decir,

$$\hat{\beta}_1 = \sum_{t=2}^n w_t \Delta y_t, \quad w_t > 0 \quad \forall t, \quad \sum_{t=2}^n w_t = 1,$$

donde los pesos w_t toman los valores máximos en el centro y van disminuyendo simétricamente a ambos lados hasta alcanzar el valor mínimo en los extremos.

1.2. Predicción por métodos de alisado

Los métodos de alisado son más flexibles que los métodos de ajuste de tendencias determinísticas pues:

- ✓ los parámetros pueden variar en el tiempo;
- ✓ las observaciones más recientes tienen mayor peso en la predicción.

1.2.1. Método de alisado simple (o alisado exponencial)

Es un método recursivo de predicción muy popular por su simplicidad.

Supongamos que estamos en el instante t . La predicción $\hat{y}_t(1)$ de la serie en $t + 1$ con la información hasta t es una combinación lineal de la predicción del instante t desde $t - 1$, $\hat{y}_{t-1}(1)$, y de la observación actual y_t :

$$\hat{y}_t(1) = \theta \hat{y}_{t-1}(1) + (1 - \theta)y_t, \quad (1.1)$$

donde $0 < \theta < 1$ determina el peso de cada componente.

Fijamos un valor para la predicción del valor inicial y_1 con 0 datos; por ejemplo, $\hat{y}_0(1) = y_1$ o $\hat{y}_0(1) = \bar{y}$. Veremos que este valor inicial casi no influye.

Predicción de y_2 con y_1 :	...
Predicción de y_3 con y_1, y_2 :	...
Predicción de y_4 con y_1, y_2, y_3 :	...
\vdots	\vdots
Predicción de y_{t+1} con y_1, \dots, y_t :	...
\vdots	\vdots

Sustituyendo $\hat{y}_1(1) = \theta \hat{y}_0(1) + (1 - \theta)y_1$ en la predicción siguiente $\hat{y}_2(1)$ obtenemos:

$$\begin{aligned} \hat{y}_2(1) &= \dots \\ &= \dots \end{aligned}$$

Reemplazando esto en $\hat{y}_3(1)$, obtenemos:

$$\begin{aligned} \hat{y}_3(1) &= \dots \\ &= \dots \end{aligned}$$

Y repitiendo este proceso para $t = 4, 5, \dots, n$, obtenemos

$$\hat{y}_n(1) = \dots$$

Si n es grande, entonces $\theta^n \cong 0$, con lo cual

$$\hat{y}_n(1) \cong \dots, \quad (1.2)$$

que es una media ponderada de las observaciones previas con pesos que decrecen al alejarse en el pasado. En efecto, $\hat{y}_n(1)$ se puede expresar como

$$\hat{y}_n(1) \cong \sum_{t=1}^n w_t y_t, \quad \text{con } w_t = \dots$$

Además, la suma de los pesos es:

$$\sum_{t=1}^n w_t = \dots$$

1.

Modelo: Las predicciones por el alisado simple coinciden con las obtenidas al asumir el modelo

$$y_t = \mu_t + a_t, \quad t = 1, \dots, n,$$

donde el nivel μ_t puede evolucionar con el tiempo.

Predicción: $\hat{y}_n(1) = \hat{\mu}_{n+1}$, donde $\hat{\mu}_{n+1}$ es el valor más cercano a todos los datos con menor peso para los datos más antiguos; es decir, μ_{n+1} es la solución del problema

$$\min_{\mu_{n+1}} \sum_{t=1}^n w_t (y_t - \mu_{n+1})^2$$

donde w_1, \dots, w_n decrecen al alejarse en el pasado, para darle mayor importancia a los datos más recientes. Derivando $f(\mu_{n+1}) = \sum_{t=1}^n w_t (y_t - \mu_{n+1})^2$ respecto de μ_{n+1} e igualando a cero, se obtiene

$$\frac{\partial f(\mu_{n+1})}{\partial \mu_{n+1}} = \dots \Rightarrow \hat{\mu}_{n+1} = \dots$$

Tomando $w_t = (1 - \theta) \theta^{n-t}$, se tiene $\sum_{t=1}^n w_t \cong 1$ y entonces la estimación de $\hat{\mu}_{n+1}$ coincide con la predicción $\hat{y}_n(1)$ dada en (1.2).

Nota 1 La ecuación de predicción recursiva (1.1) Se puede reescribir como

$$\begin{aligned}
 \hat{y}_t(1) &= \theta \hat{y}_{t-1}(1) + (1 - \theta)y_t \\
 &= \hat{y}_{t-1}(1) - \hat{y}_{t-1}(1) + \theta \hat{y}_{t-1}(1) + (1 - \theta)y_t \\
 &= \hat{y}_{t-1}(1) - (1 - \theta)\hat{y}_{t-1}(1) + (1 - \theta)y_t \\
 &= \hat{y}_{t-1}(1) + (1 - \theta)[y_t - \hat{y}_{t-1}(1)].
 \end{aligned}$$

La predicción $\hat{y}_t(1)$ es igual a la predicción anterior $\hat{y}_{t-1}(1)$ modificada por una fracción $1 - \theta$ del último error de predicción.

Nota 2 ¿qué ocurre si se toma $\theta \cong 1$?...

¿Y si se toma $\theta \cong 0$?...

Estimación de θ : Se minimiza la suma de los errores cuadráticos de predicción, es decir, θ es la solución del problema

$$\min_{\theta} \sum_{t=1}^n \hat{a}_t^2 = \sum_{t=1}^n [y_t - \hat{y}_{t-1}(1)]^2.$$

Calcular la suma para una rejilla de valores de θ : 0,1; 0,2; ...; 0,9, y tomar el valor θ_0 que minimice dicha suma. La estimación adaptativa consiste en probar después con otra rejilla más fina alrededor del óptimo anterior, $\theta_0 - 0,05$; ...; $\theta_0 - 0,01$; θ_0 ; $\theta_0 + 0,01$; $\theta_0 + 0,02$; ...; $\theta_0 + 0,05$. Normalmente esta precisión es suficiente en la práctica. Muchos programas como SPSS tienen la opción de la determinación adaptativa.

Ejemplo 4 Serie 1) del crecimiento del alquiler.

Determinamos de forma adaptativa el mejor valor de θ calculando las sumas de errores cuadráticos de predicción $\sum \hat{a}_t^2$:

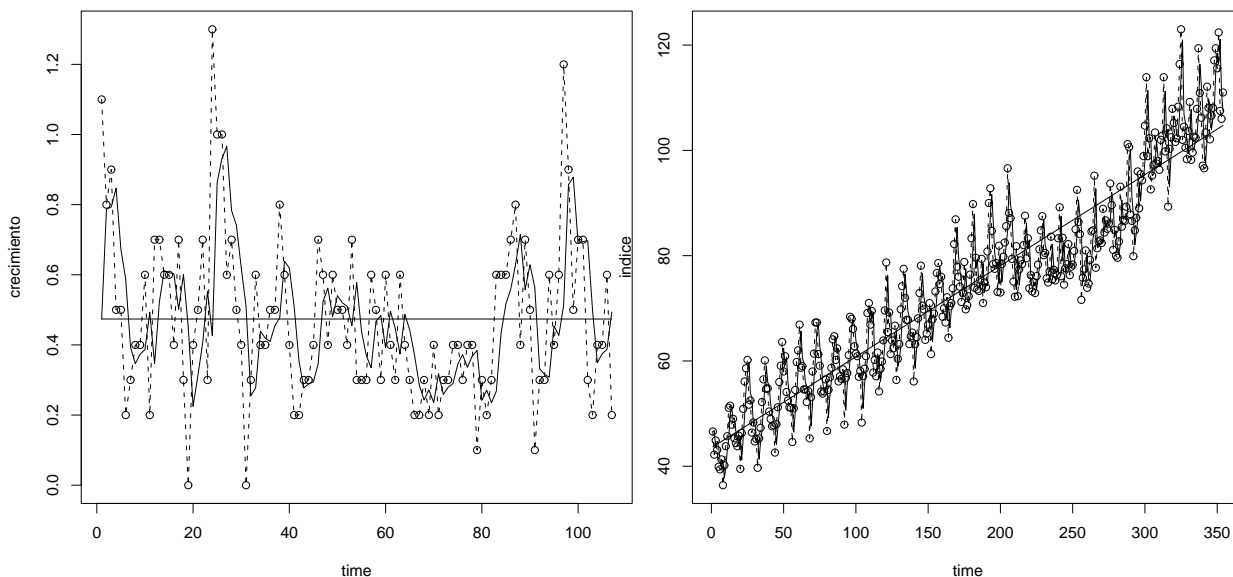
θ	1.0	0.9	0.8	0.7	0.6	0.5	0.4	0.3	0.2	0.1
$\sum \hat{a}_t^2$	5.987	5.978	5.947	5.877	5.800	5.766	5.805	5.934	6.156	6.474
θ	0.55	0.54	0.53	0.52	0.51	0.50	0.49	0.48	0.47	0.46
$\sum \hat{a}_t^2$	5.775	5.772	5.769	5.767	5.766	5.7655	5.7657	5.767	5.769	5.771

Según la tabla, dar el mismo peso a la predicción anterior y a la última observación $\hat{\theta} = 1 - \hat{\theta} = 0,5$ minimiza el error de predicción. Obsérvese en la Figura 1.2 (izqda.) que el alisado simple consigue reproducir una parte de la estructura no lineal en los datos. Esto es difícil de conseguir con una tendencia paramétrica.

Figura 1.2: Métodos de alisado

Alquiler (discont.) y alisado simple (sólida)

Energía (o) y alisado doble (sólida)



1.2.2. Método de alisado doble de Holt

Es una adaptación de las ideas anteriores a series con tendencia lineal.

Modelo: Se supone que el nivel evoluciona de forma lineal, pero la pendiente no es constante, sino que varía a lo largo del tiempo.

$$y_t = \mu_t + a_t, \quad \mu_t = \mu_{t-1} + \beta_{t-1}, \quad t = 1, \dots, n,$$

donde $\beta_{t-1} = \mu_t - \mu_{t-1}$ es la pendiente en el instante $t - 1$.

Nota 3 El modelo con tendencia lineal constante corresponde al caso $\beta_{t-1} = \beta$. Este modelo es más flexible ya que permite que la pendiente sea variable.

Predicción recursiva:

✓ *Observamos y_1 :*

Predicción del nivel μ_1 con y_1 : $\hat{\mu}_1(0)$ se fija a un valor inicial.

Predicción de la pendiente β_1 con y_1 : $\hat{\beta}_1(0)$ se fija a un valor inicial.

Predicción de y_2 con y_1 : $\hat{y}_1(1) = \hat{\mu}_1(1) = \hat{\mu}_1(0) + \hat{\beta}_1(0)$

✓ *Observamos y_2 :*

Predicción de μ_2 con y_1, y_2 : $\hat{\mu}_2(0) = \hat{\mu}_1(1) + (1 - \theta)(y_2 - \hat{\mu}_1(1))$
 (como alisado simple) $= \hat{\mu}_1(0) + \hat{\beta}_1(0) + (1 - \theta)[y_2 - \hat{\mu}_1(0) - \hat{\beta}_1(0)]$.

Predicción de β_2 con y_1, y_2 : $\hat{\beta}_2(0) = \hat{\beta}_1(0) + (1 - \gamma)[\hat{\mu}_2(0) - \hat{\mu}_1(0) - \hat{\beta}_1(0)]$

Predicción de y_3 con y_1, y_2 : $\hat{y}_2(1) = \hat{\mu}_2(1) = \hat{\mu}_2(0) + \hat{\beta}_2(0)$

✓ *Observamos y_t :*

Predicción de μ_t : $\hat{\mu}_t(0) = \hat{\mu}_{t-1}(1) + (1 - \theta)(y_t - \hat{\mu}_{t-1}(1))$
 $= \hat{\mu}_{t-1}(0) + \hat{\beta}_{t-1}(0) + (1 - \theta)[y_t - \hat{\mu}_{t-1}(0) - \hat{\beta}_{t-1}(0)]$.

Predicción de β_t : $\hat{\beta}_t(0) = \hat{\beta}_{t-1}(0) + (1 - \gamma)[\hat{\mu}_t(0) - \hat{\mu}_{t-1}(0) - \hat{\beta}_{t-1}(0)]$

Predicción de y_{t+1} : $\hat{y}_t(1) = \hat{\mu}_t(1) = \hat{\mu}_t(0) + \hat{\beta}_t(0)$

Estimación de los parámetros: El alisado doble utiliza dos parámetros, $\theta \in [0, 1]$ para el nivel y $\gamma \in [0, 1]$ para la pendiente. Ambos se determinan minimizando la suma de errores cuadráticos de predicción

$$\min_{\theta, \gamma} \sum_{t=1}^n [y_t - \hat{y}_{t-1}(1)]^2.$$

Nota 4 Si tomamos $\gamma = 1$, entonces las pendientes son constantes. Si además tomamos como valor inicial $\hat{\beta}_1(0) = 0$, obtenemos el alisado simple.

Ejemplo 5 Serie 6) del alquiler.

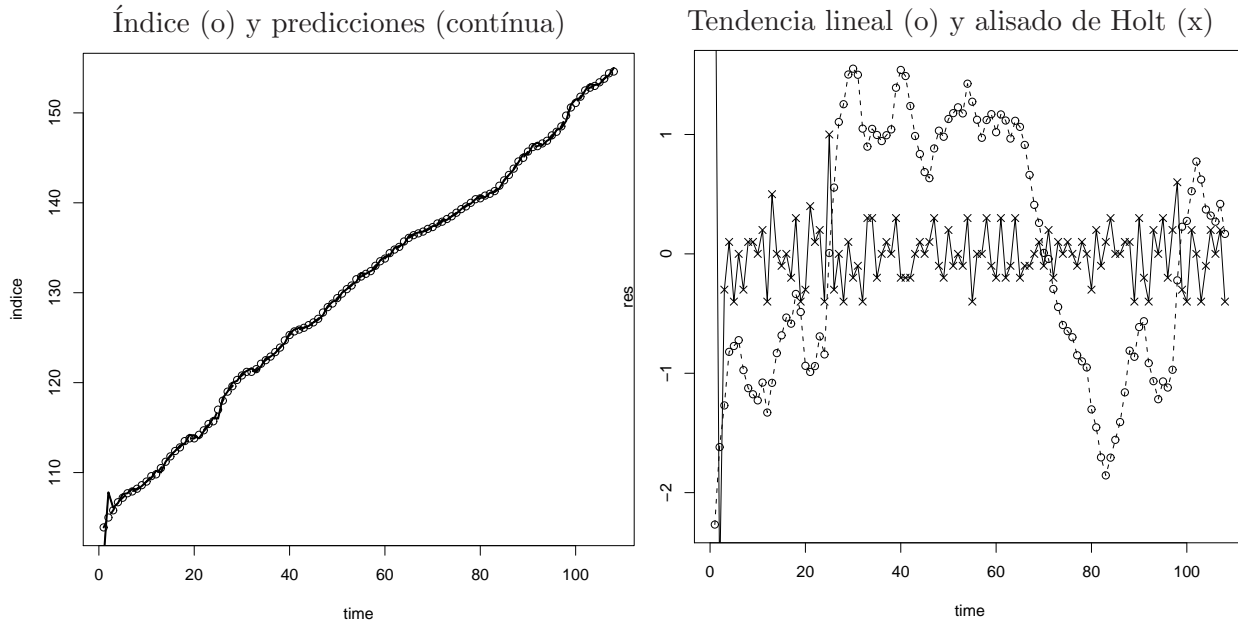
Seleccionamos (θ, γ) que minimicen la suma de los residuos cuadrados.

Cuadro 1.1: Alisado doble de Holt, serie de alquiler.

θ	1.0	0.9	0.8	0.7	0.6	0.5	0.4	0.3	0.2	0.1	0.0
γ	0.0	0.82	1.0	1.0	0.0	0.0	0.0	0.21	0.43	1.0	1.0
SSR	113.0	128.6	74.3	50.99	34.3	26.2	22.9	21.4	20.2	18.7	17.6

Se obtiene $(\hat{\theta}, \hat{\gamma}) = (0, 1)$. En la Figura 1.3 vemos que las predicciones parecen buenas y los residuos parecen aleatorios. Se mejora bastante el análisis con el modelo de tendencia lineal. ¿Qué significa que $\theta = 0$ y $\gamma = 1$?

Figura 1.3: Alquiler: predicción (izq.) y residuos (der.)



Ejemplo 6 Para la serie 15) del índice de la producción de energía, los mejores valores son $\hat{\theta} = 0,25$ y $\hat{\gamma} = 1$. Éstos reducen la desviación típica de los residuos de 123.2 (modelo de tendencia lineal) a 103.9.

Cuadro 1.2: Alisado doble de Holt, índice de energía.

θ	1.0	0.9	0.8	0.7	0.6	0.5	0.4	0.3	0.2	0.1	0.0	0.25
γ	0.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
SSR ($\times 10^3$)	15.1	12.4	12.3	12.0	11.7	11.3	11.0	10.8	10.8	11.0	11.4	10.7

1.3. Métodos de descomposición

Modelo: Se considera un *modelo con componentes aditivos*

$$y_t = \mu_t + s_t + a_t, \quad t = 1, \dots, n, \quad (1.3)$$

- ✓ μ_t *tendencia* determinista (no aleatoria);
- ✓ s_t *componente estacional* determinista: Es una función *periódica*, es decir, que se repite cada cierto número de observaciones (el *periodo*). Si el periodo es ℓ ,

$$s_t = s_{t-\ell}$$

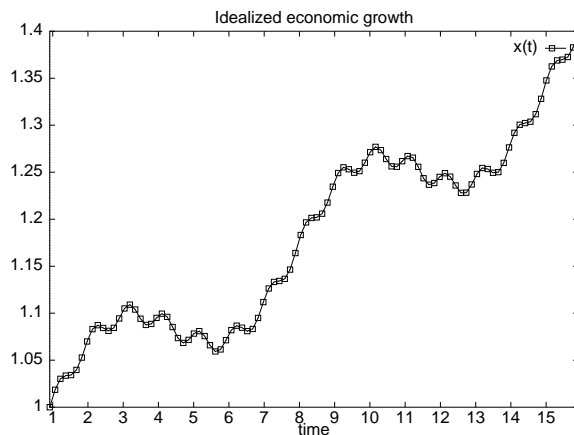
Para series mensuales (observadas cada mes) con estacionalidad anual, $\ell = 12$.

Para series diarias (observadas cada día) con estacionalidad semanal, $\ell = 7$.

Para series observadas cada hora con estacionalidad diaria, $\ell = 24$.

- ✓ a_t *componente aleatorio*: sucesión de variables incorreladas con media cero y varianza constante.

Figura 1.4: Modelo con componentes aditivos



1.3.1. Método de descomposición básico

1. *Estimación de la tendencia $\hat{\mu}_t$ y extracción de serie sin tendencia:* La tendencia se puede estimar mediante un modelo de tendencias deterministas. Después se elimina la tendencia de la forma:

$$\hat{\mu}_t \Rightarrow e_t = y_t - \hat{\mu}_t \quad \text{serie sin tendencia.}$$

2. *Estimación de la estacionalidad de la serie sin tendencia:* Una forma sencilla de estimar el efecto de las ℓ distintas estaciones es:

- ✓ Calcular la media de la serie sin tendencia: $\bar{e} = n^{-1} \sum_{t=1}^n e_t$.
- ✓ Calcular las medias de cada estación: Suponemos que hemos observado m periodos completos. Entonces el número de datos es $n = m\ell$ y los datos de la serie sin tendencia son

$$e_{(k-1)\ell+j}, \quad k = 1, \dots, m, \quad j = 1, \dots, \ell.$$

Por ejemplo, con datos mensuales de periodicidad anual, si observamos m años, $n = 12m$ y los datos son $e_{12(k-1)+j}$, $k = 1, \dots, m$, $j = 1, \dots, 12$. Si se empieza a observar la serie en Enero, los datos con $j = 1$ corresponden a Enero, $j = 2$ a Febrero, $j = 3$ a Marzo, etc.

La media de la estación j es

$$\bar{e}_j = m^{-1} \sum_{k=1}^m e_{(k-1)\ell+j}, \quad j = 1, \dots, \ell.$$

- ✓ Los componentes estacionales se estiman mediante:

$$\hat{s}_j = \bar{e}_j - \bar{e}, \quad j = 1, \dots, \ell$$

y verifican $\sum_{j=1}^{\ell} \hat{s}_j = 0$.

3. *Predicción de la serie:* Suma de la tendencia y la estacionalidad estimada,

$$\hat{y}_{i(k-1)\ell+j} = \hat{\mu}_{(k-1)\ell+j} + \hat{s}_j, \quad k = 1, \dots, m, \quad j = 1, \dots, \ell.$$

4. *Residuos:* Se obtienen restando el componente estacional a la serie sin tendencia,

$$\hat{a}_{(k-1)\ell+j} = e_{(k-1)\ell+j} - \hat{s}_j, \quad k = 1, \dots, m, \quad j = 1, \dots, \ell.$$

Nota 5 Si la variabilidad de los datos parece crecer con su nivel, como ocurre por ejemplo en la serie 19 (viajeros), podemos suponer un **modelo multiplicativo**:

$$y_t = \mu_t s_t a_t, \quad t = 1, \dots, n.$$

Tomando logaritmos y renombrando las componentes se obtiene un modelo aditivo.

Ejemplo 7 *Ajustamos una tendencia lineal a la serie 9) del índice de precios del vestido,*

$$y_t = 102,97 + 0,21t + e_t, \quad t = 1, \dots, n.$$

Restamos esta recta a la serie \rightarrow Serie sin tendencia

Calculamos las medias de cada mes del calendario \rightarrow Factores estacionales

Restamos a la serie la tendencia y la estacionalidad estimadas \rightarrow Residuos.

1.3.2. Estimación por medias móviles

Muchas series no tienen tendencia constante. En ese caso, es necesario estimar una tendencia que vaya variando en el tiempo.

Media móvil de orden m : Habitualmente se toma m impar; es decir, $m = 2k + 1$ para un cierto k . Consiste en estimar la tendencia en el instante t utilizando la media de los m datos observados en los instantes $(t - k, \dots, t - 1, t, t + 1, \dots, t + k)$,

$$\hat{\mu}_t = \frac{1}{m} (y_{t-k} + \dots + y_{t-1} + y_t + y_{t+1} + \dots + y_{t+k}), \quad t = k + 1, \dots, n - k. \quad (1.4)$$

Obsérvese que se toman y_t , los k datos anteriores y los k datos posteriores a y_t .

✓ No se puede estimar la tendencia para los primeros k y los últimos k datos.

Media móvil para series con estacionalidad aditiva:

1. *Estimación de la tendencia $\hat{\mu}_t$:* Se estima la tendencia de forma local con una *media móvil* de orden igual al periodo de la serie.

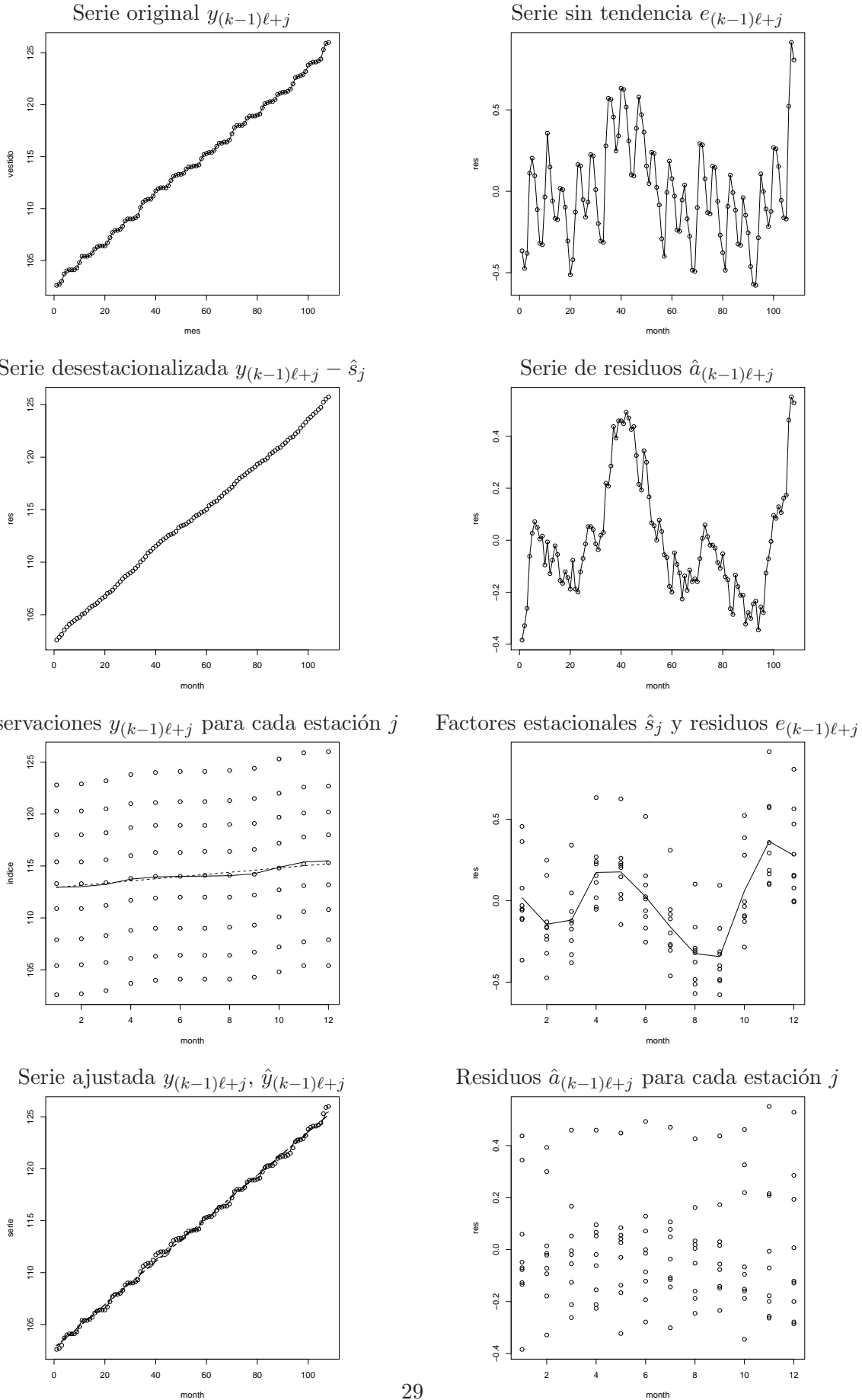
✓ Si el periodo ℓ es impar ($\ell = 2k + 1$), se estima la tendencia en t con una media móvil de orden ℓ como en (1.4).

✓ Si ℓ es par ($\ell = 2k$), hacemos una media de $2k + 1$ datos dando peso 0.5 a los datos de los extremos, es decir,

$$\hat{\mu}_t = \frac{1}{\ell} (0,5y_{t-k} + y_{t-k+1} + \dots + y_{t+k-1} + 0,5y_{t+k}), \quad t = k + 1, \dots, n - k.$$

Se completa la media móvil en $t = 1, \dots, k$ y $t = n - k + 1, \dots, n$. Dos formas:

Figura 1.5: Modelo de componentes aditivos: Índice del precio de vestido 1/1993–12/2001.



- ✓ Añadir la primera observación y_1 k veces al principio de la serie para obtener $\hat{\mu}_1, \dots, \hat{\mu}_k$, y añadir y_n al final k veces para completar $\hat{\mu}_{n-k+1}, \dots, \hat{\mu}_n$.
- ✓ Ajustar una recta a los k primeros y a los k últimos valores de la serie.

Serie sin tendencia: $e_t = y_t - \hat{\mu}_t$.

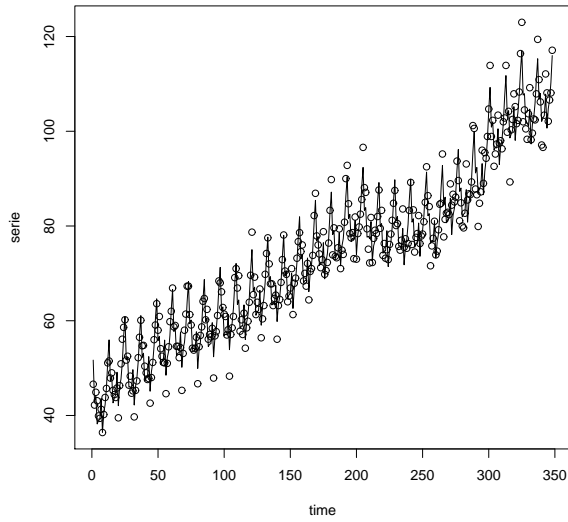
2. *Estimación de los componentes estacionales*: Se estiman promediando los datos de la misma estación y centrando estas medias para que sumen cero

$$\tilde{s}_j = \frac{1}{m} \sum_{k=1}^m e_{(k-1)\ell+j}, \quad \hat{s}_j = \tilde{s}_j - \frac{1}{\ell} \sum_{i=1}^{\ell} \tilde{s}_i, \quad j = 1, \dots, \ell.$$

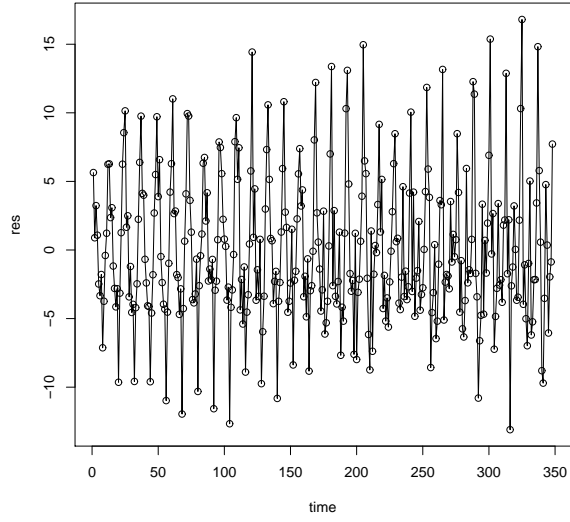
Ejemplo 8 *Serie de la producción de energía 11). Aplicamos el método de media móvil para la serie con $\ell = 12$ observaciones ($k=6$) y estimamos la estacionalidad de la serie sin tendencia.*

Figura 1.6: Método de medias móviles: Producción de energía 1/1975–12/2003.

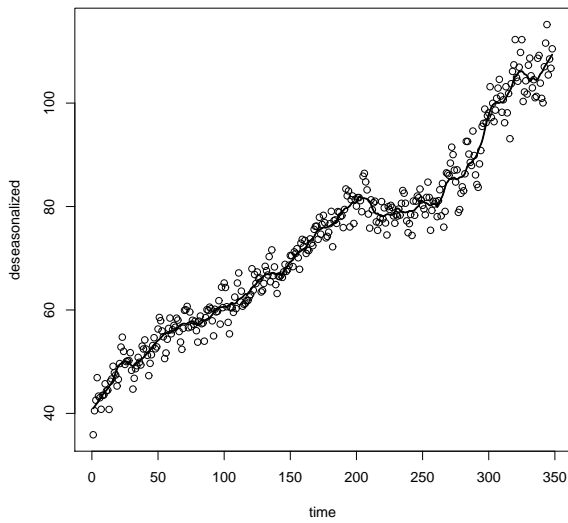
Serie original y ajuste



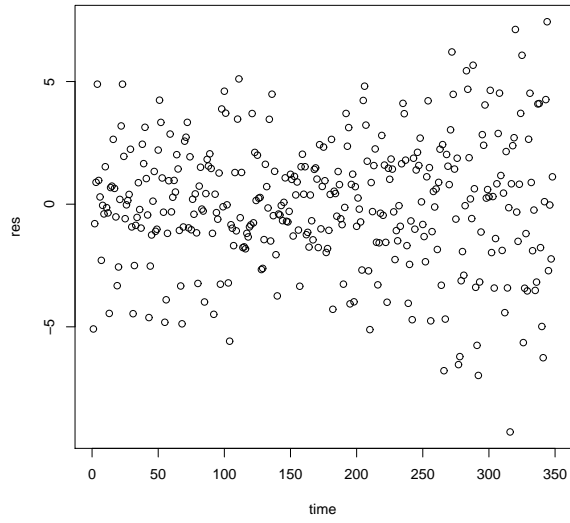
Serie sin tendencia



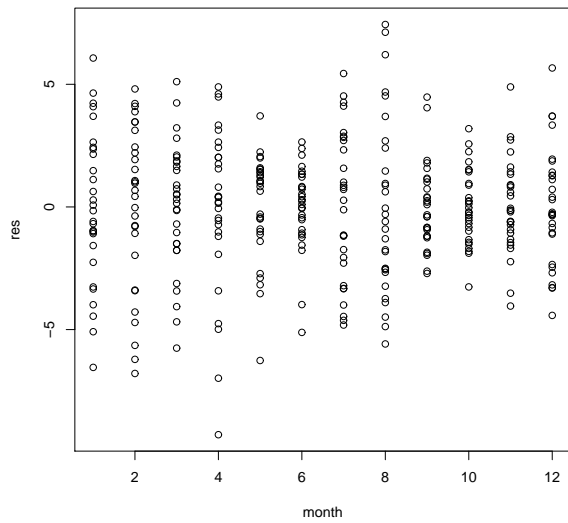
Serie desestacionalizada



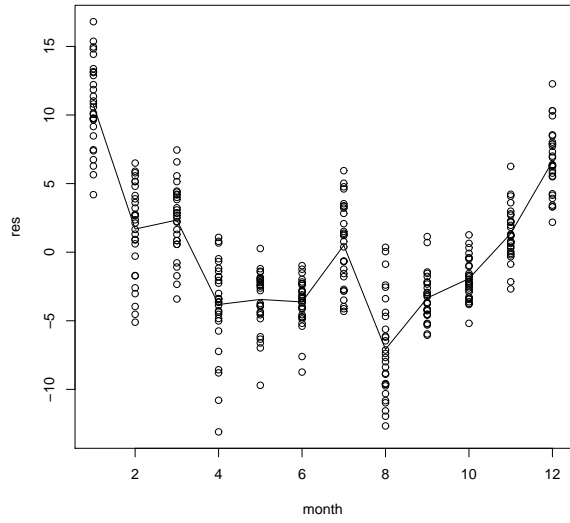
Serie de residuos



Residuos por estación



Factores estacionales y residuos



1.4. Estacionalidad y ajuste de ciclos

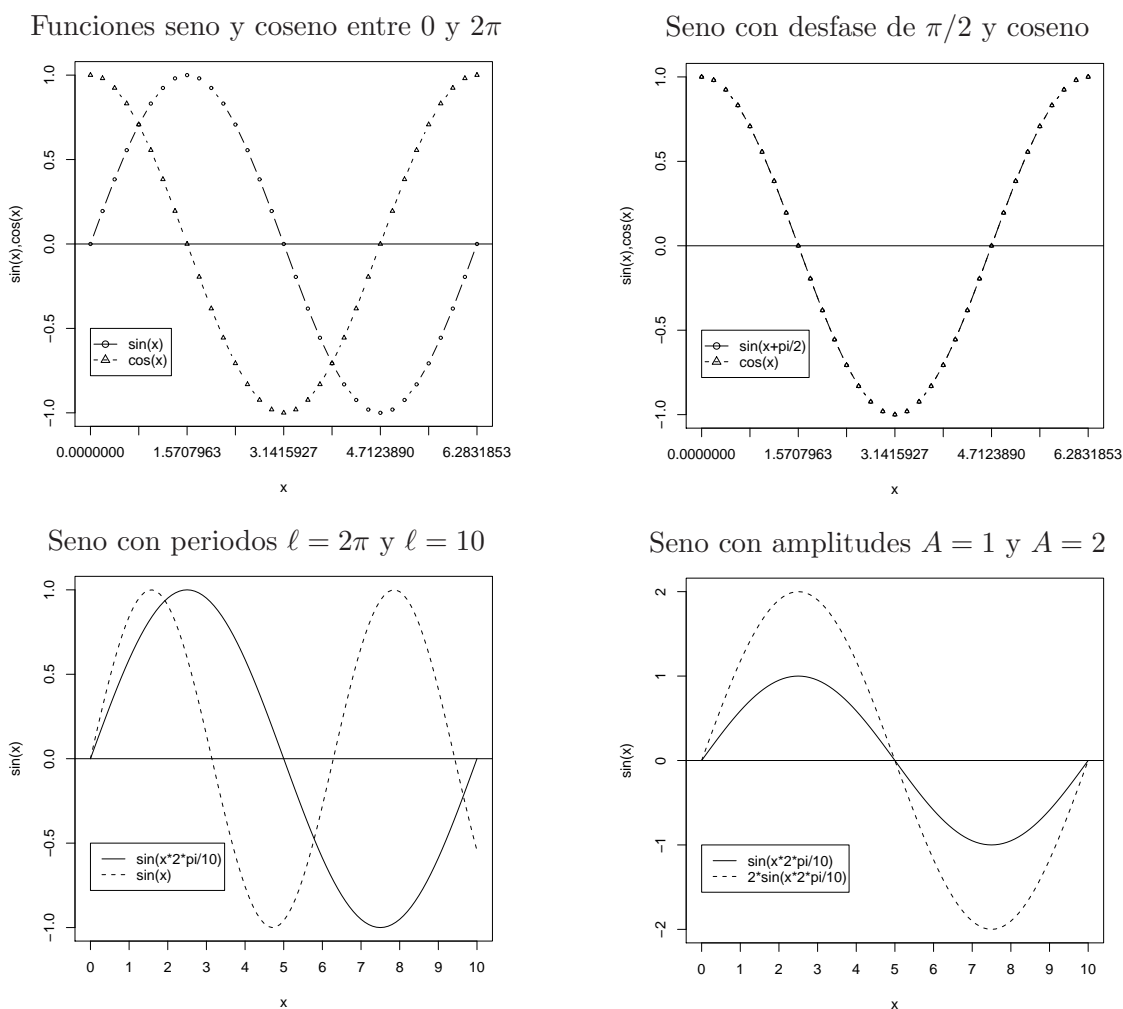
En este apartado se supone que se ha estimado y eliminado anteriormente la tendencia, y observamos estacionalidad de periodo ℓ ,

$$y_t = s_t + a_t, \quad t = 1, \dots, n,$$

donde s_t se repite cada ℓ instantes, es decir, $s_{t+\ell} = s_t$.

Una alternativa a la estimación de ℓ factores estacionales consiste en modelizar s_t mediante una función **armónica** como el seno o el coseno.

Figura 1.7: Funciones armónicas seno y coseno:



Fórmulas trigonométricas:

$$\text{sen}^2(x) + \text{cos}^2(x) = 1$$

$$\text{sen}(x \pm y) = \text{sen}(x) \text{cos}(y) \pm \text{cos}(x) \text{sen}(y)$$

$$\text{cos}(x \pm y) = \text{cos}(x) \text{cos}(y) \mp \text{sen}(x) \text{sen}(y)$$

$$\text{sen}^2(x) = \frac{1 - \text{cos}(2x)}{2}$$

$$\text{cos}^2(x) = \frac{1 + \text{cos}(2x)}{2}$$

1.4.1. Modelización con ciclo único

Para simplificar, suponemos que se han observado m periodos completos de la serie. Entonces n es un múltiplo entero de ℓ , es decir, $n = m\ell$. La estacionalidad se puede modelizar con una función seno o coseno, de la forma:

$$s_t = A \operatorname{sen}((2\pi/\ell)t + \theta)$$

- ✓ ℓ es el *periodo*: número de observaciones hasta que la serie se repite. En general, cualquier función de la forma $\beta_1 \operatorname{sen}(2\pi t/\ell) + \beta_2 \operatorname{cos}(2\pi t/\ell)$ tiene periodo ℓ .
- ✓ A es la *amplitud* de la oscilación;
- ✓ θ es el *ángulo de desfase* con relación al comienzo del ciclo.
- ✓ $f = 1/\ell$ se denomina *frecuencia* de la serie. Es la fracción de ciclo completado entre dos observaciones de la serie. Si $\ell < 1$, entonces $f > 1$ es el número de ciclos que pasan entre dos observaciones. Para una serie diaria con periodo anual ($\ell = 365$), la frecuencia es $f = 1/365 = 0,00274$. Si la periodicidad es mensual ($\ell = 30$), $f = 1/30 = 0,0333$, y si el periodo es semanal, ($\ell = 7$), $f = 0,143$.
- ✓ $\omega = 2\pi/\ell = 2\pi f$ es la *frecuencia angular*, o ángulo (en radianes) recorrido entre dos observaciones de la serie. Para $\ell = 365$, $\ell = 30$ y $\ell = 7$, se tiene $\omega = 0,0172$, $\omega = 0,2094$ y $\omega = 0,8976$ respectivamente.

Ejemplo 9 La estacionalidad de la serie 21) con periodo $\ell = 24$ se puede representar por $s_t = \operatorname{sen}(2\pi t/24)$. Esta serie tiene $n = 100$ observaciones, ¿Cuántos ciclos observamos?

Modelo:

$$y_t = \mu + A \operatorname{sen}(\omega t + \theta) + a_t, \quad t = 1, \dots, n, \quad (1.5)$$

donde μ es el nivel promedio de la serie y a_t son las innovaciones iid $N(0, \sigma^2)$.

Parámetros desconocidos: μ , A y θ .

Obsérvese que la serie no es lineal en θ . Desarrollando el seno de una suma,

$$\operatorname{sen}(\omega t + \theta) = \dots,$$

y definiendo los nuevos parámetros β y γ

$$\beta = A \operatorname{cos}(\theta), \quad \gamma = A \operatorname{sen}(\theta),$$

encontramos una parametrización del modelo más conveniente para el ajuste:

$$y_t = \mu + \beta \cos(\omega t) + \gamma \sin(\omega t) + a_t. \quad (1.6)$$

Los parámetros originales se pueden obtener a partir de los nuevos de la forma

$$A = \sqrt{\beta^2 + \gamma^2}, \quad \theta = \arctan(\gamma/\beta).$$

Llamando $x_{1t} = \sin(\omega t)$ y $x_{2t} = \cos(\omega t)$, el modelo resultante

$$y_t = \mu + \beta x_{1t} + \gamma x_{2t} + a_t, \quad t = 1, \dots, n,$$

es un modelo de regresión lineal múltiple con parámetros μ , β y γ .

Ajuste del modelo:

Ajustamos el modelo lineal por mínimos cuadrados. Es decir, encontramos los estimadores de μ, β, γ resolviendo el problema

$$\text{Min}_{\hat{\mu}, \hat{\beta}, \hat{\gamma}} \sum_{t=1}^n \left(y_t - \hat{\mu} - \hat{\beta} \sin(\omega t) - \hat{\gamma} \cos(\omega t) \right)^2 = \sum_{t=1}^n \hat{a}_t^2.$$

Derivando e igualando a cero, obtenemos las ecuaciones normales

$$\begin{aligned} \sum_{t=1}^n \hat{a}_t &= 0 \Leftrightarrow \hat{\mu} = \bar{y}; \\ \sum_{t=1}^n \hat{a}_t \sin(\omega t) &= 0 \Leftrightarrow \hat{\beta} = \frac{2}{n} \sum_{t=1}^n y_t \sin(\omega t); \\ \sum_{t=1}^n \hat{a}_t \cos(\omega t) &= 0 \Leftrightarrow \hat{\gamma} = \frac{2}{n} \sum_{t=1}^n y_t \cos(\omega t). \end{aligned}$$

Solución:

$$\begin{aligned} \hat{\mu} &= \frac{1}{n} \sum_{t=1}^n y_t, & \hat{\beta} &= \frac{2}{n} \sum_{t=1}^n y_t \sin(\omega t), & \hat{\gamma} &= \frac{2}{n} \sum_{t=1}^n y_t \cos(\omega t), \\ \hat{A}^2 &= \hat{\beta}^2 + \hat{\gamma}^2, & \hat{\theta} &= \arctan(\hat{\gamma}/\hat{\beta}). \end{aligned}$$

Estimador de σ^2 :

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{t=1}^n \hat{a}_t^2.$$

Descomposición de la variabilidad:

La variabilidad total se puede descomponer de la forma

$$VT = \sum_{t=1}^n (y_t - \bar{Y})^2 = n \left(\frac{\hat{A}^2}{2} + \hat{\sigma}^2 \right), \quad (1.7)$$

Prueba:...

Proporción de variabilidad explicada:

$$R^2 = \frac{n\hat{A}^2}{2VT} = \frac{\hat{A}^2}{2s_y^2}.$$

Ejemplo 10 *Ajustamos el modelo (1.6) a las primeras $n = 13 \times 12$ observaciones de la serie de la temperatura en Düsseldorf y obtenemos*

$$\hat{y}_t = 10,824 - 4,251 \sin(\omega t) - 6,782 \cos(\omega t).$$

La amplitud estimada es $\hat{A} = 8,0$ y el ángulo de desfase $\hat{\theta} = -0,562$ radianes. Dos veces la amplitud estima la diferencia máxima promedio entre el mes más frío (enero) y el más caliente (julio). La proporción de variabilidad explicada es $R^2 = 0,902$.

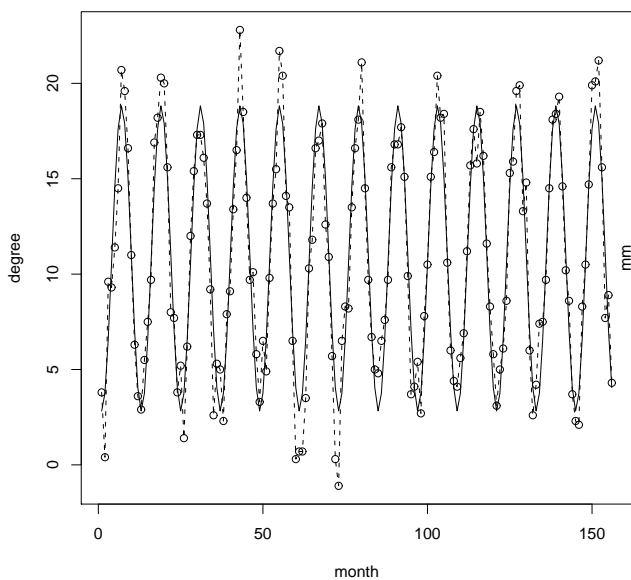
Ejemplo 11 *Idem para la serie de la lluvia en Düsseldorf:*

$$\hat{y}_t = 65,49 - 6,967 \sin(\omega t) - 0,4949 \cos(\omega t).$$

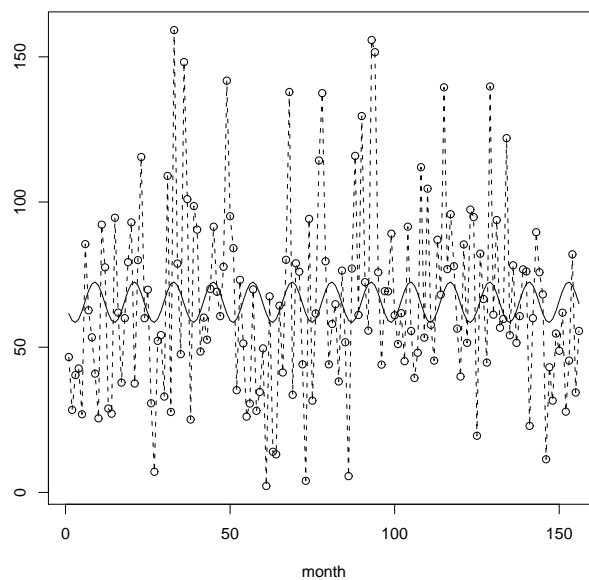
La amplitud es $\hat{A} = 6,985$ y el desfase $\hat{\theta} = -1,499$. La proporción de variabilidad explicada es ahora muy baja, $R^2 = 0,023$.

Figura 1.8: Ajuste de un ciclo: Tiempo en Düsseldorf.

Temperatura y ajuste



Lluvia y ajuste



1.4.2. Modelización de ciclos múltiples: el periodograma

El modelo (1.6) sólo es útil si s_t es exactamente una función senoidal. Sin embargo, según la representación de Fourier, cualquier función de periodo ℓ es suma de ℓ funciones senoidales con distintas amplitudes y frecuencias.

Periodos básicos o de Fourier: periodos que son fracciones exactas de n . Si n es par, los periodos básicos son

$$\ell_j = \frac{n}{j}, \quad j = 1, 2, 3, \dots, n/2.$$

Máximo periodo básico: n (para $j = 1$), observamos solo un ciclo.

Mínimo periodo básico: 2 (para $j = n/2$), no podemos observar periodos que duren menos de dos observaciones. Por ejemplo, en datos diarios no podemos detectar ciclos de periodo menor que 48 horas.

Frecuencias básicas o de Fourier: inversos de los periodos básicos,

$$f_j = \frac{j}{n}, \quad j = 1, 2, 3, \dots, n/2,$$

con lo que $1/n \leq f_j \leq 1/2$. No podemos observar frecuencias más altas.

La frecuencia máxima, $f_{n/2} = 0,5$, se llama *frecuencia de Nyquist*.

Frecuencias angulares básicas: $\omega_j = 2\pi f_j$, $j = 1, 2, \dots, n/2$.

Representación de Fourier: Cualquier serie se puede representar como suma de n ondas con frecuencias básicas:

$$y_t = \mu + \sum_{j=1}^{n/2} \beta_j \sin(\omega_j t) + \sum_{j=1}^{n/2} \gamma_j \cos(\omega_j t) + a_t \quad (1.8)$$

Como $\omega_{n/2} = 2\pi/2 = \pi$ y $\sin(\pi t) = 0$, $t = 0, 1, \dots$, podemos eliminar $\beta_{n/2} \sin(\pi t)$, con lo cual tenemos n parámetros: μ , $\beta_1, \dots, \beta_{n/2-1}$ y $\gamma_1, \dots, \gamma_{n/2}$.

Ajuste del modelo: Se verifica que todas las variables explicativas $\sin(\omega_j t)$ y $\cos(\omega_j t)$ de (1.8) son ortogonales. Por tanto, podemos ajustar las regresiones para cada frecuencia básica y obtener los coeficientes estimados de cada regresión β_j y γ_j .

El periodograma es una herramienta que permitirá decidir qué frecuencias básicas debemos incluir. Se basa en la descomposición de la variabilidad de la serie en la variabilidad debida a cada frecuencia básica.

Descomposición de la variabilidad total de la serie:

$$VT = \sum_{t=1}^n (y_t - \bar{y})^2 = \sum_{j=1}^{n/2} \frac{n}{2} \hat{A}_j^2 + n\hat{\sigma}^2$$

Obsérvese que la contribución de la onda con frecuencia f_j a la varianza de la serie es la amplitud al cuadrado dividido por 2, $\hat{A}_j^2/2$. Así, ondas de alta amplitud son importantes para explicar la serie mientras que las ondas de baja amplitud explican poco. La contribución $\hat{A}_j^2/2$ de cada onda se puede calcular a partir de los coeficientes estimados β_j y γ_j , mediante $\hat{A}_j = \sqrt{\hat{\beta}_j^2 + \hat{\gamma}_j^2}$.

Periodograma: Representación gráfica de la contribución de cada frecuencia básica en función de f_j , es decir,

$$I(f_j) = \frac{n\hat{A}_j^2}{2}, \quad f_j = \frac{1}{n}, \frac{2}{n}, \dots, \frac{1}{2}. \quad (1.9)$$

Notas:

- ✓ Se puede ajustar un modelo para cualquier frecuencia $0 \leq f \leq 0,5$ aunque no sea una fracción exacta, pero los senos y cosenos correspondientes pierden la simetría y ortogonalidad.

- ✓ Puede ser que nuestros datos sean, por ejemplo, mensuales ($f = 1/12$), pero que no exista ninguna frecuencia básica igual a $1/2$. Sin embargo, para cualquier frecuencia no básica $0 \leq f \leq 0,5$ existe una frecuencia básica f_j que es muy próxima a f , ya que f estará entre comprendida entre dos frecuencias básicas que se diferencian en $f_j - f_{j-1} = 1/n$, que es muy pequeño para n grande. Así, podemos obtener una idea de la amplitud de la onda con frecuencia f utilizando la frecuencia básica más cercada f_j .

Periodograma suavizado: En lugar de líneas, construimos rectángulos con altura $I(f_j)$, centro de la base en f_j y longitud de la base $1/n$. El periodograma se puede suavizar al igual que se hace con el histograma para aproximar una función de densidad. Area total del periodograma suavizado: s_y^2 .

Ejemplos: Ajustamos el modelo para múltiples ciclos y obtenemos el periodograma para la serie formada por las primeras 156 observaciones de las series de la temperatura y la lluvia en Dusseldorf.

Figura 1.9: Detección de ciclos múltiples: Periodograma.

