

Índice general

3. Inferencia en procesos de difusión	3
3.1. Introducción	3
3.2. Estimación máximo verosímil	3
3.2.1. Ejemplo: proceso wiener.	4
3.2.2. Proceso de Ornstein-Uhlenbeck con factores exógenos.	6

Tema 3

Inferencia en procesos de difusión

3.1. Introducción

En este tema abordaremos la estimación máximo verosímil en procesos de difusión. Para ello consideraremos el caso de muestreo discreto, esto es, supondremos que se dispone de observaciones del proceso en instantes de tiempo t_1, \dots, t_n en los cuales se observan las variables $X(t_1), \dots, X(t_n)$, cuyos valores observados constituirán la muestra base del estudio inferencial.

El procedimiento que seguiremos está basado en el método de estimación por máxima verosimilitud. Para ello será necesario conocer, excepto eventualmente para valores de parámetros desconocidos, la distribución conjunta de la muestra observada, lo cual conlleva conocer las distribuciones finito-dimensionales del proceso.

En nuestro caso, y puesto que los procesos de difusión son procesos de Markov¹, la propiedad de Markov permite que a partir de la distribución inicial del proceso y las transiciones se tenga cualquier distribución finito dimensional (ver tema 1) y, con ello, podamos aplicar la teoría de estimación máximo verosímil.

3.2. Estimación máximo verosímil

Sea $\{X(t); t \geq t_0\}$ un proceso de difusión, del cual conocemos sus distribuciones unidimensionales y transiciones, y para el cual es posible realizar observaciones del mismo en instantes de tiempo prefijados. Sean t_1, \dots, t_n dichos instantes de tiempo y llamemos x_1, \dots, x_n a los valores observados del proceso.

Notemos f_1 y f a la densidad de la variable $X(t_1)$ y a la función de densidad de transición (o funciones masa de probabilidad en el caso discreto), respectivamente, y sean θ_1 y θ los parámetros asociados a ambas. Así, la función de verosimilitud de la muestra observada $\mathbf{x} = (x_1, \dots, x_n)'$ es

$$\mathbb{L}_{\mathbf{x}}(\theta_1, \theta) = f_1(x_1) \prod_{i=2}^n f(x_i, t_i | x_{i-1}, t_{i-1}),$$

a partir de la cual se obtendrán los estimadores máximo verosímiles de θ_1 y θ . En concreto, asociadas a la muestra observada \mathbf{x} , se obtendrán las estimaciones $\hat{\theta}_1 = \hat{\theta}_1(\mathbf{x})$ y $\hat{\theta} = \hat{\theta}(\mathbf{x})$ tales que

$$\mathbb{L}_{\mathbf{x}}(\hat{\theta}_1, \hat{\theta}) = \sup_{\theta_1, \theta} \mathbb{L}_{\mathbf{x}}(\theta_1, \theta).$$

Un caso particular, pero bastante usual en la práctica, sobre todo cuando se considera una sola trayectoria del proceso, es aquél en el que la distribución de $X(t_1)$ es degenerada, esto es, $P(X(t_1) = x_1) = 1$. En dicho caso la función de verosimilitud queda como

$$\mathbb{L}_{\mathbf{x}}(\theta) = \prod_{i=2}^n f(x_i, t_i | x_{i-1}, t_{i-1}),$$

¹En realidad, el procedimiento que vamos a emplear no es válido sólo para procesos de difusión sino para procesos markovianos en general.

que depende sólo del parámetro θ .

Como hemos comentado, cuando se observa una sola trayectoria se debe considerar en t_1 una distribución degenerada ya que, en otro caso, no se dispondría de información suficiente para estimar el parámetro θ_1 . En otras ocasiones se dispone de información sobre d trayectorias, observadas en instantes de tiempo t_{ij} , ($i = 1, \dots, d$, $j = 1, \dots, n_i$). Observemos que no es necesario que los instantes de observación sean los mismos para cada trayectoria, si bien el instante inicial conviene que sí lo sea ya que hay que imponer una distribución inicial. Así pues consideraremos que $t_{i1} = t_1$, $i = 1, \dots, d$. Llamando $\{x_{ij}\}_{i=1, \dots, d; j=1, \dots, n_i}$ a los valores observados, y \mathbf{x} al vector conteniendo dichos valores, la función de verosimilitud queda en este caso en la forma

$$\mathbb{L}_{\mathbf{x}}(\theta_1, \theta) = \prod_{i=1}^d f_1(x_{i1}) \prod_{j=2}^{n_i} f(x_{ij}, t_{ij} | x_{i,j-1}, t_{i,j-1})$$

si se considera una distribución no degenerada en t_1 , mientras que en otro caso es

$$\mathbb{L}_{\mathbf{x}}(\theta) = \prod_{i=1}^d \prod_{j=2}^{n_i} f(x_{ij}, t_{ij} | x_{i,j-1}, t_{i,j-1}).$$

Como es sabido, es habitual considerar el logaritmo de la función de verosimilitud para calcular el estimador máximo verosímil. Considerando el caso de múltiples trayectorias con distribución inicial no degenerada, ello conduce a

$$\log(\mathbb{L}_{\mathbf{x}}(\theta_1, \theta)) = \sum_{i=1}^d \log(f_1(x_{i1})) + \sum_{i=1}^d \sum_{j=2}^{n_i} \log(f(x_{ij}, t_{ij} | x_{i,j-1}, t_{i,j-1})),$$

por lo que, en el caso de que θ_1 y θ sean independientes, la estimación de ambos también lo será. En tal caso, para la estimación de θ_1 sólo se considera la información del instante inicial de observación, mientras que la estimación de θ coincide en el caso de distribución inicial degenerada y no degenerada.

3.2.1. Ejemplo: proceso wiener.

Sea $\{X(t); t \geq t_0\}$ el proceso de Wiener con media infinitesimal $A_1(x, t) = \mu$ y varianza infinitesimal $A_2(x, t) = \sigma^2$, con $\mu \in \mathbb{R}$ y $\sigma > 0$.

Su función de densidad de transición (ver ejercicio 2 de la relación resuelta del tema 2) es normal unidimensional; concretamente para $s < t$:

$$\tilde{X}(t) | \tilde{X}(s) = y \rightsquigarrow N_1[y + \mu(t-s); \sigma^2(t-s)].$$

Consideremos un muestreo discreto de una trayectoria en instantes de tiempo $t_1 < \dots < t_n$ ($t_1 \geq t_0$) y sean x_1, \dots, x_n los valores observados. Además, supongamos que $P(X(t_1) = x_1) = 1$, o sea, el proceso tiene distribución degenerada en t_1 .

En consecuencia, la función de verosimilitud para la muestra observada es

$$\begin{aligned} \mathbb{L}_{\mathbf{x}}(\mu, \sigma^2) &= \prod_{i=2}^n f(x_i, t_i | x_{i-1}, t_{i-1}) = \prod_{i=2}^n \frac{1}{\sqrt{2\pi\sigma^2(t_i - t_{i-1})}} \exp\left(-\frac{[x_i - x_{i-1} - \mu(t_i - t_{i-1})]^2}{2\sigma^2(t_i - t_{i-1})}\right) = \\ &= (2\pi\sigma^2)^{-\frac{n-1}{2}} \prod_{i=2}^n (t_i - t_{i-1})^{-\frac{1}{2}} \exp\left(-\frac{1}{2\sigma^2} \sum_{i=2}^n \frac{[x_i - x_{i-1} - \mu(t_i - t_{i-1})]^2}{t_i - t_{i-1}}\right) \end{aligned}$$

y su logaritmo

$$\log(\mathbb{L}_{\mathbf{x}}(\mu, \sigma^2)) = -\frac{n-1}{2} \log(\sigma^2) - \frac{n-1}{2} \log(2\pi) - \frac{1}{2} \sum_{i=2}^n \log(t_i - t_{i-1}) - \frac{1}{2\sigma^2} \sum_{i=2}^n \frac{[x_i - x_{i-1} - \mu(t_i - t_{i-1})]^2}{t_i - t_{i-1}}.$$

A continuación derivamos esta última expresión respecto de los parámetros:

$$\frac{\partial \log (\mathbb{L}_{\mathbf{x}}(\mu, \sigma^2))}{\partial \mu} = \frac{1}{\sigma^2} \sum_{i=1}^n (x_i - x_{i-1} - \mu(t_i - t_{i-1})) = \frac{1}{\sigma^2} ((x_n - x_1) - \mu(t_n - t_1))$$

$$\frac{\partial \log (\mathbb{L}_{\mathbf{x}}(\mu, \sigma^2))}{\partial \sigma^2} = -\frac{n-1}{2\sigma^2} + \frac{1}{2\sigma^4} \sum_{i=2}^n \frac{[x_i - x_{i-1} - \mu(t_i - t_{i-1})]^2}{t_i - t_{i-1}}.$$

Igualando las derivadas a cero se obtienen las ecuaciones de verosimilitud

$$\frac{\partial \log (\mathbb{L}_{\mathbf{x}}(\mu, \sigma^2))}{\partial \mu} = 0 \Rightarrow x_n - x_1 = \mu(t_n - t_1)$$

$$\frac{\partial \log (\mathbb{L}_{\mathbf{x}}(\mu, \sigma^2))}{\partial \sigma^2} = 0 \Rightarrow \sum_{i=2}^n \frac{[x_i - x_{i-1} - \mu(t_i - t_{i-1})]^2}{t_i - t_{i-1}} = (n-1)\sigma^2,$$

cuya solución es

$$\hat{\mu} = \frac{x_n - x_1}{t_n - t_1} \quad \text{y} \quad \widehat{\sigma^2} = \frac{1}{n-1} \sum_{i=2}^n \frac{[x_i - x_{i-1} - \hat{\mu}(t_i - t_{i-1})]^2}{t_i - t_{i-1}}.$$

No obstante hemos de verificar que la solución obtenida conduce al supremo que se busca para la función de verosimilitud. En primer lugar calculamos la matriz hessiana de $\log (\mathbb{L}_{\mathbf{x}})$, para lo cual calculamos las derivadas segundas:

$$\frac{\partial^2 \log (\mathbb{L}_{\mathbf{x}}(\mu, \sigma^2))}{\partial \mu^2} = -\frac{t_n - t_1}{\sigma^2}$$

$$\frac{\partial^2 \log (\mathbb{L}_{\mathbf{x}}(\mu, \sigma^2))}{\partial \mu \partial \sigma^2} = -\frac{(x_n - x_1) - \mu(t_n - t_1)}{\sigma^4}$$

$$\frac{\partial^2 \log (\mathbb{L}_{\mathbf{x}}(\mu, \sigma^2))}{\partial (\sigma^2)^2} = \frac{n-1}{2\sigma^4} - \frac{1}{\sigma^6} \sum_{i=2}^n \frac{[x_i - x_{i-1} - \mu(t_i - t_{i-1})]^2}{t_i - t_{i-1}}.$$

Así la matriz hessiana en cualquier punto (μ, σ^2) es

$$H(\mu, \sigma^2) = \begin{pmatrix} -\frac{t_n - t_1}{\sigma^2} & -\frac{(x_n - x_1) - \mu(t_n - t_1)}{\sigma^4} \\ -\frac{(x_n - x_1) - \mu(t_n - t_1)}{\sigma^4} & \frac{n-1}{2\sigma^4} - \frac{1}{\sigma^6} \sum_{i=2}^n \frac{[x_i - x_{i-1} - \mu(t_i - t_{i-1})]^2}{t_i - t_{i-1}} \end{pmatrix}$$

por lo que en $(\hat{\mu}, \widehat{\sigma^2})$ es

$$H(\hat{\mu}, \widehat{\sigma^2}) = -\frac{1}{\widehat{\sigma^2}} \begin{pmatrix} t_n - t_1 & 0 \\ 0 & \frac{n-1}{2\widehat{\sigma^2}} \end{pmatrix},$$

matriz que es definida negativa (el determinante es positivo y al menos uno de sus elementos de la diagonal es negativo). En consecuencia el punto $(\hat{\mu}, \widehat{\sigma^2})$ es un máximo relativo.

Para verificar que ese máximo relativo es absoluto, y con ello el supremo de la función de verosimilitud, observamos que dicha función, definida en $\Theta = \mathbb{R} \times \mathbb{R}^+$, es continua en todo punto y que los límites de la misma

en la frontera de Θ (esto es, cuando μ tiende a más y menos infinito y cuando σ^2 tiende a cero y a más infinito) valen cero. Por lo tanto el máximo relativo obtenido es absoluto.

En definitiva, observado el valor \mathbf{x} de la muestra se tienen las siguientes estimaciones para los parámetros

$$\hat{\mu} = \frac{x_n - x_1}{t_n - t_1} \quad \text{y} \quad \hat{\sigma}^2 = \frac{1}{n-1} \sum_{i=2}^n \frac{[x_i - x_{i-1} - \hat{\mu}(t_i - t_{i-1}))]^2}{t_i - t_{i-1}}.$$

A partir de la estimación máximo verosímil de los parámetros es posible, mediante la aplicación del teorema de Zehna, estimar funciones paramétricas asociadas al proceso como, por ejemplo, la tendencia (función media), la varianza o la covarianza. Así, aplicando las propiedades del condicionamiento

$$E[X^k(t)] = E[E[X^k(t)|X(t_1)]] , \quad t \geq t_1; \quad k \geq 1$$

y como $E[X^k(t)|X(t_1)]$ es un variable aleatoria cuyos valores son $E[X^k(t)|X(t_1) = y]$, entonces, teniendo en cuenta la expresión de las densidades de transición, se concluye

$$E[X(t)] = E[E[X(t)|X(t_1)]] = E[X(t_1) + \mu(t - t_1)] = E[X(t_1)] + \mu(t - t_1)$$

$$\begin{aligned} E[X^2(t)] &= E[E[X^2(t)|X(t_1)]] = E[\sigma^2(t - t_1) + (X(t_1) + \mu(t - t_1))^2] = \\ &= \sigma^2(t - t_1) + E[X^2(t_1)] + \mu^2(t - t_1)^2 + 2E[X(t_1)]\mu(t - t_1) \end{aligned}$$

$$\text{Var}[X(t)] = E[X^2(t)] - (E[X(t)])^2 = \sigma^2(t - t_1) + E[X^2(t_1)] - (E[X(t_1)])^2 = \sigma^2(t - t_1) + \text{Var}[X(t_1)].$$

Observemos que la tendencia y la varianza, a partir del primer instante de observación, dependen de los momentos de $X(t_1)$. En este caso concreto, y dado que hemos supuesto una distribución degenerada en t_1 , esto es, $P(X(t_1) = x_1) = 1$, entonces

$$\begin{aligned} E[X(t)] &= x_1 + \mu(t - t_1) \\ \text{Var}[X(t)] &= \sigma^2(t - t_1) \end{aligned}$$

por lo que, a partir de una muestra observada, sus estimaciones máximo verosímiles son

$$\begin{aligned} \widehat{E[X(t)]} &= x_1 + \hat{\mu}(t - t_1) \\ \widehat{\text{Var}[X(t)]} &= \hat{\sigma}^2(t - t_1) \end{aligned}$$

3.2.2. Proceso de Ornstein-Uhlenbeck con factores exógenos.

Planteamiento general

A continuación vamos a considerar el proceso de Ornstein-Uhlenbeck con factores exógenos, que generaliza al considerado en el tema anterior (ejercicios resueltos 3 y 6). Este es un proceso de difusión $\{X(t); t \geq t_0\}$ con momentos infinitesimales

$$A_1(x, t) = -g(t)x \quad \text{y} \quad A_2(x, t) = \sigma^2$$

donde $\sigma > 0$ y g es una función continua ($g(t) = \alpha$ en el caso homogéneo de los ejercicios anteriormente citados).

Este es un proceso transformado del proceso Wiener, por lo que existe una transformación del tipo considerado en el teorema 2.6.1 del tema 2. Concretamente (¡compruébes!) las funciones C_1 y C_2 que allí aparecían son

$$C_1(t) = -\frac{2g(t)}{\sigma}z \quad \text{y} \quad C_2(t) = -2g(t)$$

por lo que la transformación es

$$x' = \psi(x, t) = \frac{k_1^{1/2}}{\sigma}x \exp\left(\int_{t_0}^t g(s) ds\right) - \frac{k_1^{1/2}}{\sigma}z \exp\left(\int_{t_0}^{t_2} g(s) ds\right) + k_2$$

$$t' = \phi(t) = k_1 \int_{t_1}^t \exp\left(2 \int_{t_0}^s g(\theta) d\theta\right) ds + k_3.$$

Por lo tanto, la densidad de transición de este proceso, $f(x, t|y, \tau)$, está relacionada ² con la del proceso Wiener estándar, $f'(x', t'|y', \tau')$, mediante la expresión

$$f(x, t|y, \tau) = \frac{\partial \psi(x, t)}{\partial x} f'(x', t'|y', \tau')$$

siendo

$$f'(x', t'|y', \tau') = \frac{1}{\sqrt{2\pi(t' - \tau')}} \exp\left(-\frac{(x' - y')^2}{2(t' - \tau')}\right)$$

con $x' = \psi(x, t)$, $y' = \psi(y, \tau)$, $t' = \phi(t)$ y $\tau' = \phi(\tau)$. Por lo tanto,

$$\begin{aligned} f(x, t|y, \tau) &= \frac{\partial \psi(x, t)}{\partial x} \frac{1}{\sqrt{2\pi(\phi(t) - \phi(\tau))}} \exp\left(-\frac{(\psi(x, t) - \psi(y, \tau))^2}{2(\phi(t) - \phi(\tau))}\right) = \\ &= \frac{k_1^{1/2} e^{\int_{\tau}^t g(\theta) d\theta}}{\sigma} \frac{1}{\sqrt{2\pi k_1 \int_{\tau}^t e^{2 \int_{\tau}^s g(\theta) d\theta} ds}} \exp\left(-\frac{k_1 \left(x - y e^{-\int_{\tau}^t g(\theta) d\theta}\right)^2}{2\sigma^2 k_1 e^{-2 \int_{\tau}^t g(\theta) d\theta} \int_{\tau}^t e^{2 \int_{\tau}^s g(\theta) d\theta} ds}\right) = \\ &= \frac{1}{\sqrt{2\pi \sigma^2 e^{-2 \int_{\tau}^t g(\theta) d\theta} \int_{\tau}^t e^{2 \int_{\tau}^s g(\theta) d\theta} ds}} \exp\left(-\frac{\left(x - y e^{-\int_{\tau}^t g(\theta) d\theta}\right)^2}{2\sigma^2 e^{-2 \int_{\tau}^t g(\theta) d\theta} \int_{\tau}^t e^{2 \int_{\tau}^s g(\theta) d\theta} ds}\right). \end{aligned}$$

Consideremos ahora una muestra discreta del proceso $\mathbf{x} = (x_1, \dots, x_n)'$ tomada en los instantes de tiempo $\{t_1, \dots, t_n\}$. Asimismo consideremos en este caso que $P[X(t_1) = x_1] = 1$, con lo cual se tiene la verosimilitud

$$\begin{aligned} \mathbb{L}_{\mathbf{x}}(\theta, \sigma^2) &= (2\pi)^{-\frac{n-1}{2}} (\sigma^2)^{-\frac{n-1}{2}} \prod_{\alpha=2}^n \left[e^{-2 \int_{t_{\alpha-1}}^{t_{\alpha}} g(v) dv} \int_{t_{\alpha-1}}^{t_{\alpha}} e^{2 \int_{t_{\alpha-1}}^u g(v) dv} du \right]^{-\frac{1}{2}} \times \\ &\times \exp\left(-\frac{1}{2\sigma^2} \sum_{\alpha=2}^n \frac{\left[x_{\alpha} - x_{\alpha-1} e^{-\int_{t_{\alpha-1}}^{t_{\alpha}} g(v) dv}\right]^2}{e^{-2 \int_{t_{\alpha-1}}^{t_{\alpha}} g(v) dv} \int_{t_{\alpha-1}}^{t_{\alpha}} e^{2 \int_{t_{\alpha-1}}^u g(v) dv} du}\right) \end{aligned}$$

donde θ es un parámetro (posiblemente multidimensional) incluido en la función g , mientras que su logaritmo es

$$\begin{aligned} \log(\mathbb{L}_{\mathbf{x}}(\theta, \sigma^2)) &= -\frac{n-1}{2} \log(2\pi) - \frac{n-1}{2} \log(\sigma^2) - \frac{1}{2} \sum_{\alpha=2}^n \log\left(e^{-2 \int_{t_{\alpha-1}}^{t_{\alpha}} g(v) dv} \int_{t_{\alpha-1}}^{t_{\alpha}} e^{2 \int_{t_{\alpha-1}}^u g(v) dv} du\right) - \\ &- \frac{1}{2\sigma^2} \sum_{\alpha=2}^n \frac{\left[x_{\alpha} - x_{\alpha-1} e^{-\int_{t_{\alpha-1}}^{t_{\alpha}} g(v) dv}\right]^2}{e^{-2 \int_{t_{\alpha-1}}^{t_{\alpha}} g(v) dv} \int_{t_{\alpha-1}}^{t_{\alpha}} e^{2 \int_{t_{\alpha-1}}^u g(v) dv} du} \end{aligned}$$

Observemos cómo la forma que adopta la varianza de la transición hace prácticamente imposible abordar el problema de forma genérica. Por ello vamos a mostrar un caso concreto.

²Hemos cambiado la notación $f(x, t|x_0, t_0)$ por $f(x, t|y, \tau)$ ya que la expresión del teorema no alude a un instante inicial sino a un instante anterior a t .

Caso particular

Tomemos la función $g(t) = \frac{\theta}{t+1}$, con $\theta > -1/2$. Con ello se tiene que los momentos de la transición son

$$\begin{aligned} E[X(t)|X(s) = x_s] &= x_s \left(\frac{s+1}{t+1} \right)^\theta \\ \text{Var}[X(t)|X(s) = x_s] &= \frac{\sigma^2}{2\theta+1} \left[(t+1) - (s+1) \left(\frac{s+1}{t+1} \right)^{2\theta} \right] \end{aligned}$$

a partir de los cuales el logaritmo de la verosimilitud queda de la forma

$$\begin{aligned} \log(\mathbb{L}_{\mathbf{x}}(\theta, \sigma^2)) &= -\frac{n-1}{2} \log(2\pi) - \frac{n-1}{2} \log(\sigma^2) + \frac{n-1}{2} \log(2\theta+1) - \\ &\quad - \frac{1}{2} \sum_{\alpha=2}^n \log \left((t_\alpha+1) - (t_{\alpha-1}+1) \left(\frac{t_{\alpha-1}+1}{t_\alpha+1} \right)^{2\theta} \right) - \\ &\quad - \frac{2\theta+1}{2\sigma^2} \sum_{\alpha=2}^n \frac{\left[x_\alpha - x_{\alpha-1} \left(\frac{t_{\alpha-1}+1}{t_\alpha+1} \right)^\theta \right]^2}{(t_\alpha+1) - (t_{\alpha-1}+1) \left(\frac{t_{\alpha-1}+1}{t_\alpha+1} \right)^{2\theta}} \end{aligned}$$

En cuanto a las parciales de la verosimilitud se tiene

$$\begin{aligned} \frac{\partial \log(\mathbb{L}_{\mathbf{x}}(\theta, \sigma^2))}{\partial \sigma^2} &= -\frac{n-1}{2\sigma^2} + \frac{2\theta+1}{2\sigma^4} \sum_{\alpha=2}^n \frac{\left[x_\alpha - x_{\alpha-1} \left(\frac{t_{\alpha-1}+1}{t_\alpha+1} \right)^\theta \right]^2}{(t_\alpha+1) - (t_{\alpha-1}+1) \left(\frac{t_{\alpha-1}+1}{t_\alpha+1} \right)^{2\theta}} \\ \frac{\partial \log(\mathbb{L}_{\mathbf{x}}(\theta, \sigma^2))}{\partial \theta} &= \frac{n-1}{2\theta+1} + \sum_{\alpha=2}^n \frac{(t_{\alpha-1}+1) \log \left(\frac{t_{\alpha-1}+1}{t_\alpha+1} \right) \left(\frac{t_{\alpha-1}+1}{t_\alpha+1} \right)^{2\theta}}{(t_\alpha+1) - (t_{\alpha-1}+1) \left(\frac{t_{\alpha-1}+1}{t_\alpha+1} \right)^{2\theta}} - \\ &\quad - \frac{1}{\sigma^2} \sum_{\alpha=2}^n \frac{\left[x_\alpha - x_{\alpha-1} \left(\frac{t_{\alpha-1}+1}{t_\alpha+1} \right)^\theta \right]^2}{(t_\alpha+1) - (t_{\alpha-1}+1) \left(\frac{t_{\alpha-1}+1}{t_\alpha+1} \right)^{2\theta}} + \\ &\quad + \frac{2\theta+1}{\sigma^2} \sum_{\alpha=2}^n \frac{\log \left(\frac{t_{\alpha-1}+1}{t_\alpha+1} \right) \left(\frac{t_{\alpha-1}+1}{t_\alpha+1} \right)^\theta \left[x_\alpha - x_{\alpha-1} \left(\frac{t_{\alpha-1}+1}{t_\alpha+1} \right)^\theta \right]}{\left[(t_\alpha+1) - (t_{\alpha-1}+1) \left(\frac{t_{\alpha-1}+1}{t_\alpha+1} \right)^{2\theta} \right]^2} \times \\ &\quad \times \left[x_{\alpha-1}(t_\alpha+1) - x_\alpha(t_{\alpha-1}+1) \left(\frac{t_{\alpha-1}+1}{t_\alpha+1} \right)^\theta \right] \end{aligned}$$

Planteamos las ecuaciones de verosimilitud

$$\frac{\partial \log(\mathbb{L}_{\mathbf{x}}(\theta, \sigma^2))}{\partial \sigma^2} = 0 \quad \text{y} \quad \frac{\partial \log(\mathbb{L}_{\mathbf{x}}(\theta, \sigma^2))}{\partial \theta} = 0,$$

obteniéndose, de la primera:

$$\hat{\sigma}^2 = \frac{2\hat{\theta}+1}{n-1} \sum_{\alpha=2}^n \frac{\left[x_\alpha - x_{\alpha-1} \left(\frac{t_{\alpha-1}+1}{t_\alpha+1} \right)^{\hat{\theta}} \right]^2}{(t_\alpha+1) - (t_{\alpha-1}+1) \left(\frac{t_{\alpha-1}+1}{t_\alpha+1} \right)^{2\hat{\theta}}}$$

mientras que de la segunda se tiene

$$\sum_{\alpha=2}^n \frac{(t_{\alpha-1} + 1) \log \left(\frac{t_{\alpha-1}+1}{t_{\alpha}+1} \right) \left(\frac{t_{\alpha-1}+1}{t_{\alpha}+1} \right)^{2\hat{\theta}}}{(t_{\alpha} + 1) - (t_{\alpha-1} + 1) \left(\frac{t_{\alpha-1}+1}{t_{\alpha}+1} \right)^{2\hat{\theta}}} + \frac{2\hat{\theta} + 1}{\hat{\sigma}^2} \sum_{\alpha=2}^n \frac{\log \left(\frac{t_{\alpha-1}+1}{t_{\alpha}+1} \right) \left(\frac{t_{\alpha-1}+1}{t_{\alpha}+1} \right)^{\hat{\theta}} \left[x_{\alpha} - x_{\alpha-1} \left(\frac{t_{\alpha-1}+1}{t_{\alpha}+1} \right)^{\hat{\theta}} \right]}{\left[(t_{\alpha} + 1) - (t_{\alpha-1} + 1) \left(\frac{t_{\alpha-1}+1}{t_{\alpha}+1} \right)^{2\hat{\theta}} \right]^2} \times$$

$$\times \left[x_{\alpha-1}(t_{\alpha} + 1) - x_{\alpha}(t_{\alpha-1} + 1) \left(\frac{t_{\alpha-1}+1}{t_{\alpha}+1} \right)^{\hat{\theta}} \right] = 0.$$

Sustituyendo el valor de $\hat{\sigma}^2$ en la segunda expresión, tenemos una ecuación bastante compleja para calcular el valor de $\hat{\theta}$, por lo que será necesario el empleo de métodos numéricos para su resolución³

³Dada la complejidad, obviamos el estudio de que la solución del sistema conduce al máximo, si bien es posible hacerlo en el mismo sentido que el expuesto anteriormente.