week12a: Kapitel 5.3: Effizienz und Konsistenz, Teil 5
Die Form von Minimum-Varianz-Schätzern und erwartungstreue Schätzer für das Beispiel 3

In der letzten Veranstaltung hatten wir die Fisher-Informationsmatrix $I = I(\alpha, \beta, \nu)$ für das Beispiel 3 explizit berechnet und hatten dann die Cramer-Rao Abschätzungen für erwartungstreue Schätzer $(\hat{\alpha}, \hat{\beta}, \hat{\nu})$ hingeschrieben. Daraus konnten wir dann untere Schranken für die Varianzen von erwartungstreuen Schätzen $\hat{\kappa}$ und $\widehat{\sigma^2}$ herleiten, die genau im Einklang standen mit dem Verhalten, was wir in dem week7b.pdf in den Histogrammen gesehen hatten.

Etwas allgemeiner können wir sagen, dass in allen 3 Beispielen, die wir uns angeschaut haben, die Varianzen der Maximum-Likelihood-Schätzer qualitativ übereinstimmen mit den Varianzen, die durch die Cramer-Rao Lower Bound, das ist das $[I^{-1}(\theta)]_{k,k}$, gegeben sind.

Hier wollen wir jetzt noch folgendes machen: In dem week7b.pdf hatten wir ja durch Simulation gesehen, dass der Schätzer $\hat{\kappa}_{\mathrm{ML}}$ für die mean reversion speed κ nicht erwartungstreu ist. Im week10b.pdf hatten wir in der Folgerung 5.3.3 eine Formel hingeschrieben, das war

$$\hat{\theta}(x) \;=\; \theta \;+\; I^{-1}(\theta)\, \nabla_\theta \log p_\theta(x) \tag{1}$$

und gesagt, dass erwartungstreue Minimum-Varianz-Schätzer von dieser Form sein müssen. Insbesondere auf dem letzten Übungsblatt 12 hatten wir dann nochmal explizit nachgerechnet, dass für Schätzer von der Form (1) tatsächlich

$$\mathsf{E}[\,\hat{\theta}_k\,] \;=\; \theta_k \tag{2}$$

$$\mathsf{V}[\,\hat{\theta}_k\,] \;=\; [I^{-1}(\theta)]_{k,k} \tag{3}$$

gilt, die sind also immer erwartungstreu.

Jetzt wollen wir den Ausdruck auf der rechten Seite von (1) für das Beispiel 3 explizit berechnen. Das $I(\theta)$ und die Ableitungen in dem Gradienten $\nabla_\theta p_\theta$ hatten wir ja letztes Mal schon berechnet, da müssen wir das im wesentlichen nur noch hinschreiben. Insbesondere wollen wir uns das $\hat{\kappa}$ anschauen, das muss dann ja erwartungstreu sein. Der Maximum-Likelihood-Schätzer $\hat{\kappa}_{\mathrm{ML}}$ war ja nicht erwartungstreu, also, wenn wir die rechte Seite von (1) berechnen, vielleicht bekommen wir dann ja eine gute Idee, wie wir aus dem nicht erwartungstreuen $\hat{\kappa}_{\mathrm{ML}}$ ein erwartungstreues $\hat{\kappa}$ machen können.

Die rechte Seite von (1) kann man typischerweise nicht direkt benutzen, weil da eben die theoretischen Modellparameter explizit auftauchen, aber, wie wir ja in der 2. Aufgabe auf dem Übungsblatt 11 gesehen haben, kann es durchaus mal sein, dass sich die theoretischen Modellparameter am Ende rauskürzen. Also probieren wir das mal:

Für das $I(\theta) = I(\alpha, \beta, \nu)$ hatten wir letztes Mal den folgenden Ausdruck bekommen:

$$
\begin{aligned}
I(\theta) &= \begin{pmatrix} \frac{1}{\nu}\sum_{k=1}^{n} \mathsf{E}[\,x_{k-1}^2\,] & \frac{1}{\nu}\sum_{k=1}^{n} \mathsf{E}[\,x_{k-1}\,] & 0 \\ \frac{1}{\nu}\sum_{k=1}^{n} \mathsf{E}[\,x_{k-1}\,] & \frac{n}{\nu} & 0 \\ 0 & 0 & \frac{n}{2\nu^2} \end{pmatrix} \\
&= \frac{n}{\nu} \begin{pmatrix} \frac{1}{n}\sum_{k=1}^{n} \mathsf{E}[\,x_{k-1}^2\,] & \frac{1}{n}\sum_{k=1}^{n} \mathsf{E}[\,x_{k-1}\,] & 0 \\ \frac{1}{n}\sum_{k=1}^{n} \mathsf{E}[\,x_{k-1}\,] & 1 & 0 \\ 0 & 0 & \frac{1}{2\nu} \end{pmatrix} \\
&=: \frac{n}{\nu} \begin{pmatrix} \langle x^2\rangle & \langle x\rangle & 0 \\ \langle x\rangle & 1 & 0 \\ 0 & 0 & \frac{1}{2\nu} \end{pmatrix}
\end{aligned}
$$

und damit

$$
I^{-1}(\theta) = \frac{\nu}{n} \begin{pmatrix} \frac{1}{\langle x^2\rangle - \langle x\rangle^2} & -\frac{\langle x\rangle}{\langle x^2\rangle - \langle x\rangle^2} & 0 \\ -\frac{\langle x\rangle}{\langle x^2\rangle - \langle x\rangle^2} & \frac{\langle x^2\rangle}{\langle x^2\rangle - \langle x\rangle^2} & 0 \\ 0 & 0 & 2\nu \end{pmatrix}
$$

Für die Ableitungen $\nabla_\theta p = \left(\frac{\partial p}{\partial\alpha}, \frac{\partial p}{\partial\beta}, \frac{\partial p}{\partial\nu}\right)$ hatten wir die folgenden Ausdrücke gefunden:

$$
\begin{aligned}
\frac{\partial \log p_\theta}{\partial\alpha} &= +\frac{1}{\nu}\sum_{k=1}^{n}(\,x_k - \alpha x_{k-1} - \beta\,)x_{k-1} \\
\frac{\partial \log p_\theta}{\partial\beta} &= +\frac{1}{\nu}\sum_{k=1}^{n}(\,x_k - \alpha x_{k-1} - \beta\,) \\
\frac{\partial \log p_\theta}{\partial\nu} &= -\frac{n}{2\nu} + \frac{1}{2\nu^2}\sum_{k=1}^{n}(\,x_k - \alpha x_{k-1} - \beta\,)^2
\end{aligned}
$$

Die Schätzer aus (1) sehen dann also folgendermassen aus:

$$
\begin{aligned}
\begin{pmatrix} \hat\alpha \\ \hat\beta \end{pmatrix} &= \begin{pmatrix} \alpha \\ \beta \end{pmatrix} + \frac{\nu}{n} \begin{pmatrix} \frac{1}{\langle x^2\rangle - \langle x\rangle^2} & -\frac{\langle x\rangle}{\langle x^2\rangle - \langle x\rangle^2} \\ -\frac{\langle x\rangle}{\langle x^2\rangle - \langle x\rangle^2} & \frac{\langle x^2\rangle}{\langle x^2\rangle - \langle x\rangle^2} \end{pmatrix} \begin{pmatrix} \frac{1}{\nu}\sum_{k=1}^{n}(\,x_k - \alpha x_{k-1} - \beta\,)x_{k-1} \\ \frac{1}{\nu}\sum_{k=1}^{n}(\,x_k - \alpha x_{k-1} - \beta\,) \end{pmatrix} \\
&= \begin{pmatrix} \alpha \\ \beta \end{pmatrix} + \begin{pmatrix} \langle x^2\rangle & \langle x\rangle \\ \langle x\rangle & 1 \end{pmatrix}^{-1} \begin{pmatrix} \frac{1}{n}\sum_{k=1}^{n}(\,x_k - \alpha x_{k-1} - \beta\,)x_{k-1} \\ \frac{1}{n}\sum_{k=1}^{n}(\,x_k - \alpha x_{k-1} - \beta\,) \end{pmatrix} \qquad (4)
\end{aligned}
$$

und das $\hat{\nu}$ ist gegeben durch

$$\begin{aligned}
\hat{\nu} &= \nu + \frac{2\nu^2}{n}\left\{ -\frac{n}{2\nu} + \frac{1}{2\nu^2}\sum_{k=1}^{n}(x_k - \alpha x_{k-1} - \beta)^2 \right\} \\
&= \nu - \nu + \frac{1}{n}\sum_{k=1}^{n}(x_k - \alpha x_{k-1} - \beta)^2 \\
&= \frac{1}{n}\sum_{k=1}^{n}(x_k - \alpha x_{k-1} - \beta)^2 \qquad (5)
\end{aligned}$$

Betrachten wir zunächst die Gleichung (4). Den Vektor auf der rechten Seite hinter der inversen Matrix können wir auch schreiben als

$$\begin{pmatrix} \frac{1}{n}\sum_{k=1}^{n}(x_k - \alpha x_{k-1} - \beta)x_{k-1} \\ \frac{1}{n}\sum_{k=1}^{n}(x_k - \alpha x_{k-1} - \beta) \end{pmatrix} =$$

$$\begin{pmatrix} \frac{1}{n}\sum_{k=1}^{n} x_k x_{k-1} \\ \frac{1}{n}\sum_{k=1}^{n} x_k \end{pmatrix} - \begin{pmatrix} \frac{1}{n}\sum_{k=1}^{n} x_{k-1}^2 & \frac{1}{n}\sum_{k=1}^{n} x_{k-1} \\ \frac{1}{n}\sum_{k=1}^{n} x_{k-1} & 1 \end{pmatrix}\begin{pmatrix} \alpha \\ \beta \end{pmatrix}$$

Wenn wir das in (4) einsetzen, bekommen wir

$$\begin{aligned}
\begin{pmatrix} \hat{\alpha} \\ \hat{\beta} \end{pmatrix} &= \begin{pmatrix} \alpha \\ \beta \end{pmatrix} + \begin{pmatrix} \langle x^2 \rangle & \langle x \rangle \\ \langle x \rangle & 1 \end{pmatrix}^{-1} \begin{pmatrix} \frac{1}{n}\sum_{k=1}^{n}(x_k - \alpha x_{k-1} - \beta)x_{k-1} \\ \frac{1}{n}\sum_{k=1}^{n}(x_k - \alpha x_{k-1} - \beta) \end{pmatrix} \\
&= \begin{pmatrix} \alpha \\ \beta \end{pmatrix} - \begin{pmatrix} \langle x^2 \rangle & \langle x \rangle \\ \langle x \rangle & 1 \end{pmatrix}^{-1} \begin{pmatrix} \frac{1}{n}\sum_{k=1}^{n} x_{k-1}^2 & \frac{1}{n}\sum_{k=1}^{n} x_{k-1} \\ \frac{1}{n}\sum_{k=1}^{n} x_{k-1} & 1 \end{pmatrix}\begin{pmatrix} \alpha \\ \beta \end{pmatrix} \\
&\quad + \begin{pmatrix} \langle x^2 \rangle & \langle x \rangle \\ \langle x \rangle & 1 \end{pmatrix}^{-1} \begin{pmatrix} \frac{1}{n}\sum_{k=1}^{n} x_k x_{k-1} \\ \frac{1}{n}\sum_{k=1}^{n} x_k \end{pmatrix} \\
&=: \hat{\theta}_0 + \hat{\theta}_{\mathrm{main}} \qquad (6) \\
& \qquad (7)
\end{aligned}$$

mit den Schätzern

$$\hat{\theta}_0 := \begin{pmatrix} \alpha \\ \beta \end{pmatrix} - \begin{pmatrix} \langle x^2 \rangle & \langle x \rangle \\ \langle x \rangle & 1 \end{pmatrix}^{-1} \begin{pmatrix} \frac{1}{n}\sum_{k=1}^{n} x_{k-1}^2 & \frac{1}{n}\sum_{k=1}^{n} x_{k-1} \\ \frac{1}{n}\sum_{k=1}^{n} x_{k-1} & 1 \end{pmatrix}\begin{pmatrix} \alpha \\ \beta \end{pmatrix} \qquad (8)$$

$$\hat{\theta}_{\mathrm{main}} := \begin{pmatrix} \langle x^2 \rangle & \langle x \rangle \\ \langle x \rangle & 1 \end{pmatrix}^{-1} \begin{pmatrix} \frac{1}{n}\sum_{k=1}^{n} x_k x_{k-1} \\ \frac{1}{n}\sum_{k=1}^{n} x_k \end{pmatrix} \qquad (9)$$

Der Schätzer $\hat\theta_0$ hat Erwartungswert 0,

$$\mathsf{E}[\,\hat\theta_0\,] \;=\; \mathsf{E}\left[\begin{pmatrix}\alpha\\ \beta\end{pmatrix} - \begin{pmatrix}\langle x^2\rangle & \langle x\rangle\\ \langle x\rangle & 1\end{pmatrix}^{-1}\begin{pmatrix}\frac{1}{n}\sum_{k=1}^n x_{k-1}^2 & \frac{1}{n}\sum_{k=1}^n x_{k-1}\\ \frac{1}{n}\sum_{k=1}^n x_{k-1} & 1\end{pmatrix}\begin{pmatrix}\alpha\\ \beta\end{pmatrix}\right]$$

$$= \begin{pmatrix}\alpha\\ \beta\end{pmatrix} - \begin{pmatrix}\langle x^2\rangle & \langle x\rangle\\ \langle x\rangle & 1\end{pmatrix}^{-1}\begin{pmatrix}\frac{1}{n}\sum_{k=1}^n \mathsf{E}[x_{k-1}^2] & \frac{1}{n}\sum_{k=1}^n \mathsf{E}[x_{k-1}]\\ \frac{1}{n}\sum_{k=1}^n \mathsf{E}[x_{k-1}] & 1\end{pmatrix}\begin{pmatrix}\alpha\\ \beta\end{pmatrix}$$

$$= \begin{pmatrix}\alpha\\ \beta\end{pmatrix} - \begin{pmatrix}\langle x^2\rangle & \langle x\rangle\\ \langle x\rangle & 1\end{pmatrix}^{-1}\begin{pmatrix}\langle x^2\rangle & \langle x\rangle\\ \langle x\rangle & 1\end{pmatrix}\begin{pmatrix}\alpha\\ \beta\end{pmatrix} \;=\; \begin{pmatrix}0\\ 0\end{pmatrix}$$

Also ist

$$\hat\theta_{\text{main}} \;:=\; \begin{pmatrix}\langle x^2\rangle & \langle x\rangle\\ \langle x\rangle & 1\end{pmatrix}^{-1}\begin{pmatrix}\frac{1}{n}\sum_{k=1}^n x_k x_{k-1}\\ \frac{1}{n}\sum_{k=1}^n x_k\end{pmatrix}$$

immer noch erwartungstreu,

$$\mathsf{E}[\,\hat\theta_{\text{main}}\,] \;=\; \begin{pmatrix}\alpha\\ \beta\end{pmatrix}$$

Wir können schreiben

$$\hat\theta_{\text{main}} \;=\; \begin{pmatrix}\langle x^2\rangle & \langle x\rangle\\ \langle x\rangle & 1\end{pmatrix}^{-1}\left\{\begin{pmatrix}\frac{1}{n}\sum_{k=1}^n (x_k - x_{k-1})x_{k-1}\\ 0\end{pmatrix} + \begin{pmatrix}\frac{1}{n}\sum_{k=1}^n x_{k-1}^2\\ \frac{1}{n}\sum_{k=1}^n x_k\end{pmatrix}\right\}$$

$$=:\; \hat\theta_{\text{mainpart}} \;+\; \hat\theta_1$$

mit den Schätzern

$$\hat\theta_1 \;:=\; \begin{pmatrix}\langle x^2\rangle & \langle x\rangle\\ \langle x\rangle & 1\end{pmatrix}^{-1}\begin{pmatrix}\frac{1}{n}\sum_{k=1}^n x_{k-1}^2\\ \frac{1}{n}\sum_{k=1}^n x_k\end{pmatrix} \tag{10}$$

$$\hat\theta_{\text{mainpart}} \;:=\; \begin{pmatrix}\langle x^2\rangle & \langle x\rangle\\ \langle x\rangle & 1\end{pmatrix}^{-1}\begin{pmatrix}\frac{1}{n}\sum_{k=1}^n (x_k - x_{k-1})x_{k-1}\\ 0\end{pmatrix} \tag{11}$$

$\hat\theta_1$ hat Erwartungswert

$$\mathsf{E}[\,\hat\theta_1\,] \;=\; \mathsf{E}\left[\begin{pmatrix}\langle x^2\rangle & \langle x\rangle\\ \langle x\rangle & 1\end{pmatrix}^{-1}\begin{pmatrix}\frac{1}{n}\sum_{k=1}^n x_{k-1}^2\\ \frac{1}{n}\sum_{k=1}^n x_k\end{pmatrix}\right] \;=\; \begin{pmatrix}\langle x^2\rangle & \langle x\rangle\\ \langle x\rangle & 1\end{pmatrix}^{-1}\begin{pmatrix}\frac{1}{n}\sum_{k=1}^n \mathsf{E}[x_{k-1}^2]\\ \frac{1}{n}\sum_{k=1}^n \mathsf{E}[x_k]\end{pmatrix}$$

$$= \begin{pmatrix}\langle x^2\rangle & \langle x\rangle\\ \langle x\rangle & 1\end{pmatrix}^{-1}\begin{pmatrix}\langle x^2\rangle\\ \langle x\rangle\end{pmatrix} \;=\; \begin{pmatrix}1\\ 0\end{pmatrix}$$

und damit muss

$$\hat{\theta}_{\text{mainpart}} \;=\; \frac{1}{\langle x^2\rangle \,-\, \langle x\rangle^2}\begin{pmatrix} 1 & -\langle x\rangle \\ -\langle x\rangle & \langle x^2\rangle \end{pmatrix}\begin{pmatrix} \frac{1}{n}\sum_{k=1}^{n}(x_k - x_{k-1})x_{k-1} \\ 0 \end{pmatrix}$$

den Erwartungswert

$$\mathsf{E}[\,\hat{\theta}_{\text{mainpart}}\,] \;=\; \begin{pmatrix} \alpha - 1 \\ \beta \end{pmatrix} \tag{12}$$

haben. Die folgenden Schätzer für α und β sind also erwartungstreu:

$$\hat{\alpha} \;:=\; 1 \,+\, \frac{1}{\langle x^2\rangle - \langle x\rangle^2}\,\frac{1}{n}\sum_{k=1}^{n}(x_k - x_{k-1})x_{k-1} \tag{13}$$

$$\hat{\beta} \;:=\; -\,\frac{\langle x\rangle}{\langle x^2\rangle - \langle x\rangle^2}\,\frac{1}{n}\sum_{k=1}^{n}(x_k - x_{k-1})x_{k-1} \tag{14}$$

Wegen

$$\kappa \;=\; \frac{1-\alpha}{\Delta t}$$
$$\mu \;=\; \frac{\beta}{1-\alpha}$$

würden sich dann Schätzer von folgender Form anbieten:

$$\hat{\kappa} \;\sim\; -\,\frac{1}{\langle x^2\rangle - \langle x\rangle^2}\,\frac{1}{n}\sum_{k=1}^{n}\frac{x_k - x_{k-1}}{\Delta t}\,x_{k-1} \tag{15}$$

$$\hat{\mu} \;\sim\; \langle x\rangle \;=\; \frac{1}{n}\sum_{k=1}^{n}\mathsf{E}[x_{k-1}] \tag{16}$$

Wir haben $\sim$ anstatt $=$ geschrieben, weil die Ausdrücke auf der rechten Seite von (15,16) so natürlich nicht als Schätzer benutzbar sind, da ja die Grössen $\langle x^2\rangle$ und $\langle x\rangle^2$ als Erwartungswerte ja von den gesuchten Modellparametern abhängen. Durch (16) wird aber offensichtlich die folgende Wahl motiviert:

$$\hat{\mu} \;:=\; \frac{1}{n}\sum_{k=1}^{n} x_{k-1} \tag{17}$$

Insbesondere für den Fall $(\kappa, T) = (1,1)$ ist dieses $\hat{\mu}$ deutlich besser als das $\hat{\mu}_{\text{ML}}$.

$\rightarrow$ R-Simulation

Schauen wir uns die Gleichung (15) an. Wir haben die exakte Identität

$$\kappa \;=\; \mathsf{E}\left[\,-\,\frac{1}{\langle x^2\rangle - \langle x\rangle^2}\,\frac{1}{n}\sum_{k=1}^{n}\frac{x_k - x_{k-1}}{\Delta t}\,x_{k-1}\,\right]$$

$$=\; +\,\frac{1}{\langle x^2\rangle - \langle x\rangle^2}\,\mathsf{E}\left[\,\frac{1}{n}\sum_{k=1}^{n}\frac{x_{k-1} - x_k}{\Delta t}\,x_{k-1}\,\right]$$

oder

$$\big(\,\langle x^2\rangle - \langle x\rangle^2\,\big)\,\kappa \;=\; \mathsf{E}\Big[\,\frac{1}{n}\sum_{k=1}^{n}\frac{x_{k-1}-x_k}{\Delta t}\;x_{k-1}\,\Big] \tag{18}$$

Mit dem Theorem 5.3.4 vom letzten Mal können wir schreiben

$$\begin{aligned} F(\kappa) \;&:=\; \big(\,\langle x^2\rangle - \langle x\rangle^2\,\big)\,\times\,\kappa \;\stackrel{\Delta t\to 0}{\approx}\; \frac{\sigma^2}{2\kappa}\,\frac{2\kappa T - 1 + e^{-2\kappa T}}{2\kappa T}\,\times\,\kappa \\ &=\; \frac{\sigma^2}{2} \;-\; \frac{\sigma^2}{2}\,\frac{1-e^{-2\kappa T}}{2\kappa T} \end{aligned} \tag{19}$$

Eine Idee wäre nun geweseen, die Kombination

$$\frac{1}{n}\sum_{k=1}^{n}\frac{x_{k-1}-x_k}{\Delta t}\;x_{k-1} \tag{20}$$

als erwartungstreuen Schätzer für das $F(\kappa)$ zu verwenden, und dann κ durch sowas wie

$$\hat\kappa \;=\; F^{-1}\Big(\,\frac{1}{n}\sum_{k=1}^{n}\frac{x_{k-1}-x_k}{\Delta t}\;x_{k-1}\,\Big)$$

zu schätzen. Wir schreiben 'durch sowas wie' weil wegen

$$\begin{aligned} F(\kappa) \;&=\; \mathsf{E}\Big[\,\frac{1}{n}\sum_{k=1}^{n}\frac{x_{k-1}-x_k}{\Delta t}\;x_{k-1}\,\Big] \\ \Leftrightarrow \qquad \kappa \;&=\; F^{-1}\Big(\,\mathsf{E}\Big[\,\frac{1}{n}\sum_{k=1}^{n}\frac{x_{k-1}-x_k}{\Delta t}\;x_{k-1}\,\Big]\,\Big) \\ &\neq\; \mathsf{E}\Big[\,F^{-1}\Big(\,\frac{1}{n}\sum_{k=1}^{n}\frac{x_{k-1}-x_k}{\Delta t}\;x_{k-1}\,\Big)\,\Big] \end{aligned}$$

man dann noch mit dem Convexity Adjustment eine entsprechende Korrektur machen müsste, um das möglichst erwartungstreu zu machen. Allerdings ist es nun so, dass die Funktion $F(\kappa)$ gegeben durch die rechte Seite von (19) kaum von κ abhängt, sondern insbesondere für grosse $\kappa T \gg 1$ ist das im wesentlichen das $\sigma^2/2$. Mit anderen Worten, die Kombination (20) ist nicht wirklich sensitiv bezüglich Variation von κ und wird sich deshalb kaum dafür eignen, das κ zu schätzen. Ok, müsste man sich vielleicht mal ein paar Simulationen anschauen.

Die Idee, die wir hier aber gerade gehabt haben, ist vielleicht nicht so schlecht, wir suchen eine Kombination von x_{t_k}'s, wo wir den Erwartungswert explizit berechnen können und wo das Resultat möglichst sensitiv bezüglich κ sein sollte, und wenn wir den Erwartungswert dann mit $F(\kappa)$ bezeichnen, schätzen wir das κ durch die Umkehrfunktion F^{-1} und machen dann mit dem Convexity Adjustment noch eine Korrektur. Schauen wir uns mal die folgende Kombination an:

$$\begin{aligned} \widehat{s^2} \;&:=\; \frac{1}{n}\sum_{k=1}^{n}x_{t_k}^2 \;-\; \Big(\,\frac{1}{n}\sum_{k=1}^{n}x_{t_k}\Big)^2 \;=\; \frac{1}{n}\sum_{k=1}^{n}x_{t_k}^2 \;-\; \hat\mu^{\,2} \\ &=\; \frac{1}{n}\sum_{k=1}^{n}(x_{t_k}-\hat\mu)^2 \end{aligned} \tag{21}$$

Den Erwartungswert kann man berechnen und man findet

$$\mathsf{E}[\,\widehat{s^2}\,] \;=\; \frac{\sigma^2 T}{2}\;F(\kappa T) \tag{22}$$

mit der streng monotonen Funktion

$$F(x) \;=\; \frac{2x-1+e^{-2x}}{2x^2} \;+\; \frac{e^{-2x}-4e^{-x}-2x+3}{x^3} \tag{23}$$

Wir brauchen das κ. Die folgenden Gleichungen sind exakt (im Limes $\Delta t \to 0$)

$$\begin{aligned} F(\kappa T) &= \frac{2}{\sigma^2 T}\,\mathsf{E}[\,\widehat{s^2}\,] \\ \kappa &= \frac{1}{T}\,F^{-1}\!\left(\frac{2}{\sigma^2 T}\,\mathsf{E}[\,\widehat{s^2}\,]\right) \end{aligned} \tag{24}$$

Es liegt also nahe, den Schätzer

$$\hat{\kappa} \;:=\; \frac{1}{T}\,F^{-1}\!\left(\frac{2}{\sigma^2 T}\,\widehat{s^2}\right) \tag{25}$$

zu betrachten. Wegen

$$\mathsf{E}[\,\hat{\kappa}\,] \;=\; \frac{1}{T}\,\mathsf{E}\!\left[F^{-1}\!\left(\frac{2}{\sigma^2 T}\,\widehat{s^2}\right)\right] \;\neq\; \frac{1}{T}\,F^{-1}\!\left(\frac{2}{\sigma^2 T}\,\mathsf{E}[\,\widehat{s^2}\,]\right) \;\stackrel{(24)}{=}\; \kappa \tag{26}$$

können wir nicht erwarten, dass dieser Schätzer erwartungstreu ist, aber wir könnten eine Korrektur mit dem Convexity Adjustment berechnen. Dazu bräuchten wir dann auch eine Formel für die Varianz von dem $\widehat{s^2}$, also für die Grösse $\mathsf{V}[\,\widehat{s^2}\,]$.

Okay, das kann man auch berechnen, wird dann aber schon eine etwas länglichere Rechnung. Probieren wir an dieser Stelle vielleicht einfach nur nochmal aus, wie gut der Schätzer (25) so hinkommt im Vergleich mit dem Standard Maximum-Likelihood-Schätzer $\hat{\kappa}_{\mathrm{ML}}$, also ganz ohne Convexity Adjustment:

$\rightarrow$ R-Simulation

Die Umkehrfunktion $F^{-1}(y)$ in (25) bestimmen wir numerisch mit Hilfe von Newton-Iteration, indem wir zu gegebenen y die Gleichung $F(x) - y = 0$ lösen, also indem wir die Nullstelle der Funktion

$$f(x) \;:=\; F(x) - y$$

durch Newton-Iteration

$$x_{n+1} \;=\; x_n - \frac{f(x_n)}{f'(x_n)} \;=\; x_n - \frac{F(x_n)-y}{F'(x_n)}$$

bestimmen. Dazu brauchen wir offensichtlich die Ableitung von dem $F(x)$, man berechnet

$$F'(x) \;=\; \frac{-(x+1)e^{-2x}-x+1}{x^3} \;-\; \frac{(2x+3)e^{-2x}-(4x+12)e^{-x}-4x+9}{x^4}\;.$$

Betrachten wir schliesslich noch den Schätzer für das $\hat{\nu}$ gegeben durch (5),

$$\hat{\nu} \;=\; \frac{1}{n}\sum_{k=1}^{n}(\,x_k - \alpha x_{k-1} - \beta\,)^2$$

Tatsächlich können wir für kleine Δt die rechte Seite durch einen von α und β unabhängigen Ausdruck ersetzen, denn wegen

$$\begin{aligned}\alpha &= 1 - \kappa\,\Delta t \\ \beta &= \kappa\mu\,\Delta t\end{aligned}$$

können wir schreiben

$$\begin{aligned}\hat{\nu} &= \frac{1}{n}\sum_{k=1}^{n}\big[\,x_k - \alpha x_{k-1} - \beta\,\big]^2 \\ &= \frac{1}{n}\sum_{k=1}^{n}\big[\,x_k - (1-\kappa\Delta t)x_{k-1} - \kappa\mu\Delta t\,\big]^2 \\ &= \frac{1}{n}\sum_{k=1}^{n}\big[\,(x_k - x_{k-1}) \,+\, \Delta t\,\kappa(x_{k-1} - \mu)\,\big]^2 \\ &\approx \frac{1}{n}\sum_{k=1}^{n}(x_k - x_{k-1})^2\end{aligned}$$

da $x_k - x_{k-1} = O(\sqrt{\Delta t})$ und $\sqrt{\Delta t} \gg \Delta t$ im Limes $\Delta t \to 0$. Der Parameter von Interesse ist das σ oder das σ^2 mit

$$\sigma^2 \;=\; \frac{\eta^2}{\Delta t} \;=\; \frac{\nu}{\Delta t}\;,$$

das wir dann also mit

$$\widehat{\sigma^2} \;:=\; \frac{1}{n\Delta t}\sum_{k=1}^{n}(x_k - x_{k-1})^2 \tag{27}$$

schätzen können.

Die Performance der Schätzer $(\hat{\kappa}, \hat{\mu}, \widehat{\sigma^2})$ aus (25,17,27) können wir uns dann morgen noch in einer R-Simulation anschauen und mit der Performance der Maximum-Likelihood-Schätzer vergleichen.