Kalman-Bucy-Filter

Niklas von Moers

17. Juni 2021

1 Einführung

- Im zeitdiskreten Modell:

$$X_{t+1} = AX_t + U_t$$

$$Z_t = GX_t + V_t$$

Oder äquivalent

$$\Delta X_t = (A - I)X_t + U_t$$

- Im zeitstetigen Modell:

$$\begin{aligned} dX_t &= F(t)X_t dt + C(t)dU_t && \text{tatsächlicher Prozess (Werte unbekannt)} \\ dZ_t &= G(t)X_t dt + D(t)dV_t && \text{Observierungen (Werte bekannt)} \end{aligned}$$

- Obige „Ito-Schreibweise“ bedeutet:

$$F(t)X_t dt = \int_0^t F(s)X_s ds$$

$$dX_t = \int_0^t 1 dX_s = X_t - X_0$$

- (Prozesse leben auf $(\Omega, \mathcal{F}, \mathbb{P})$.) Nur eindimensionaler Fall (lässt sich verallgemeinern).

- Anfangsbedingungen:

$$X_0 \sim \mathcal{N}(\mu, \sigma^2), \ Z_0 = 0$$

- $Z_0 = 0$ können wir annehmen, da wir andernfalls zum Prozess $(\int_0^t Z_s ds)_t$ übergehen können (mit neuem G und D). Etwas dubios an dieser Stelle, da man bei der Rückrichtung ableiten müsste. In der Anwendung ist dies jedoch durch Diskretisierung in Ordnung (keine Ableitung, sondern Differenzen).

- Regularitätsbedingungen: Die (bekannten) Funktionen

$$F, C, G, D, \frac{1}{D} : \mathbb{R} \to \mathbb{R}$$

  sind beschränkt auf beschränkten Intervallen.
  X_0, U und V seien unabhängig voneinander.

  Wir fordern sogar, anders als die Quelle Øksendal, dass die Funktionen stetig sind.

- Diese Regularitätsbedingungen sind hinreichend, dass obige stochastische Differentialgleichungen (SDEs) auf kompaktem Intervall $[0, T]$ eindeutige Lösungen besitzen. Der Beweis (à la Picard-Lindelöf) der Lösung ist nicht (wirklich) konstruktiv; wir werden aber eine explizite Darstellung von X berechnen.

- Gesucht: Bester „Z-messbarer Schätzer $\hat{X}_t$ für X_t mit Informationen bis zum Zeitpunkt t“ bzgl. des quadratischen Fehlers.

- Formal:

  - $S_{Y_t}(t) := \mathbb{E}[(X_t - Y_t)^2]$
  - $\mathcal{K}(Z, t) := \{Y : \Omega \to \mathbb{R} \mid Y \in L^2(\mathbb{P}), Y \text{ ist } \sigma(Z_s \mid s \leq t) \text{ mb.}\}$
  - $\hat{X}_t = \underset{Y \in \mathcal{K}(Z,t)}{\arg\min}\, S_Y(t)$

- Ist erfüllt von
$$\hat{X}_t := \mathbb{E}[X_t \mid \sigma(Z_s \mid s \leq t)].$$
Ziel ist es also, diesen bedingten Erwartungswert zu berechnen.

- Das Ergebnis ist nicht explizit, sondern durch eine stochastische Differentialgleichung (SDE) gegeben:

  Satz 1.0.1. *Der (im obigen Sinne) beste Filter $\hat{X}_t$ ist gegeben durch*
$$d\hat{X}_t = (F(t) - \frac{G^2(t)S(t)}{D^2(t)})\hat{X}_t dt + \frac{G(t)S(t)}{D^2(t)} dZ_t$$
  mit Anfangsbedingung $\hat{X}_0 = \mathbb{E}[X_0]$, wobei $S(t) := S_{\hat{X}_t}(t)$ der mittlere quadratische Fehler ist. Dieser erfüllt die (deterministische) Ricatti-Differentialgleichung
$$\frac{dS}{dt}(t) = 2F(t)S(t) - \frac{G^2(t)}{D^2(t)} S^2(t) + C^2(t).$$

Zur Konstruktion Brownscher Bewegung:

Satz 1.0.2 (Donskersches Invarianzprinzip)**.** *$(X_n)_n$ Folge normierter iid ZV (d.h. $\mathbb{E}[X_n] = 0,\ Var(X_n) = 1$). Definiere Random Walk $S_n := \sum_i^n X_i$. Interpoliere diesen linear:*
$$S(t) := S_{\lfloor t \rfloor} + (t - \lfloor t \rfloor)(S_{\lceil t \rceil} - S_{\lfloor t \rfloor}).$$
$$S_n^*(t) := \frac{1}{\sqrt{n}} S(nt),\ t \in [0,1].$$
Nach CLT $S_n^* \to \mathcal{N}(0,t)$ in Verteilung. $(S_n^*(t))_n$ konvergiert in Verteilung gegen Brownsche Bewegung in $(C[0,1], \|\cdot\|_\infty)$.

2 Beweis

2.1 Formel für X

Satz 2.1.1 (Ito-Formel)**.** *Sei $X = (X_t)_t$ ein Itō-Prozess gegeben durch*

$$dX_t(\omega) = u(t,\omega)dt + v(t,\omega)dB_t,$$

sei $g(t,x) \in C^2([0,\infty) \times \mathbb{R})$. Dann ist auch Y, definiert durch $Y_t \coloneqq g(t, X_t)$, ein Itō-Prozess und es gilt

$$dY_t = \frac{\partial g}{\partial x_1}(t,X_t)dt + \frac{\partial g}{\partial x_2}(t,X_t)dX_t + \frac{1}{2}\frac{\partial^2 g}{\partial x_2^2}(t,X_t) \cdot (dX_t)^2,$$

wobei $(dX_t)^2 = (dX_t) \cdot (dX_t)$ berechnet wird durch

$$d_t \cdot d_t = dt \cdot dB_t = dB_t \cdot d_t = 0, dB_t \cdot dB_t = dt.$$

Term $(dX_t)^2$ irrelevant, weil dieser nicht auftreten wird:

Satz 2.1.2.

$$X_t = \exp\left(\int\limits_0^t F(s)ds\right)\left(X_0 + \int\limits_0^t \exp\left(-\int\limits_0^s F(u)du\right)C(s)dU_s\right)$$

Beweis. Setze

$$g(t,x) \coloneqq \exp\left(-\int\limits_0^t F(s)ds\right)x$$

und

$$Y_t \coloneqq g(t,X_t) = \exp\left(-\int\limits_0^t F(s)ds\right)X_t.$$

Itō-Formel:

$$Y_t - Y_0 = \int\limits_0^t \frac{\partial g}{\partial x_1}(s,X_s)ds + \int\limits_0^t \frac{\partial g}{\partial x_2}(s,X_s)dX_s$$

$$
\begin{aligned}
&= \int_0^t -F(s) \exp\left(-\int_0^s F(u)du\right) X_s ds \\
&+ \int_0^t \exp\left(-\int_0^s F(u)du\right) dX_s \\
&= \int_0^t -F(s) \exp\left(-\int_0^s F(u)du\right) X_s ds \\
&+ \int_0^t \exp\left(-\int_0^s F(u)du\right) F(s) X_s ds + \int_0^t \exp\left(-\int_0^s F(u)du\right) C(s) dU_s \\
&= \int_0^t \exp\left(-\int_0^s F(u)du\right) C(s) dU_s
\end{aligned}
$$

Teilen durch $\exp\left(-\int_0^t F(s)ds\right)$ liefert das gewünschte Ergebnis. $\square$

Man kann außerdem für $0 \leq r \leq t$ die Gleichung

$$
X_t = \exp\left(\int_r^t F(s)ds\right) X_r + \int_r^t \exp\left(\int_s^t F(u)du\right) C(s) dU_s
$$

zeigen.

Insbesondere

$$
\mathbb{E}[X_t] = \exp\left(\int_0^t F(s)ds\right) X_0.
$$

Also falls $X_0 = 0$, so hat der Prozess immer Erwartungswert 0.

2.2 Weiteres Vorgehen

Beweis wird so ablaufen:

(i)

$$
\hat{X}_t = P_{\mathcal{K}(Z,t)}(X_t) = P_{\mathcal{L}(Z,t)}(X_t) = P_{\mathcal{L}(N,t)}(X_t) = P_{\mathcal{L}(R,t)}(X_t),
$$

mit $\mathcal{L}$ = Projektion nur auf Linearkombinationen von Z; N und R zwei neue Prozesse, welche schönere Eigenschaften als Z haben. N ist der sogenannte Innovationsprozess und R ist Brownsche Bewegung.

(ii) Finde für $\mathcal{L}(Z,t)$ und für $P_{\mathcal{L}(R,t)}(X_t)$ eine explizite Darstellung (welche noch von X abhängt).

(iii) Setze zuvor berechnete Formel von X ein, um SDE von $\hat{X}$ zu erhalten (welche nur noch von Z, nicht mehr von X abhängt).

2.3 Z-lineare und Z-messbare Schätzer

Es sei $\mathcal{L} = \mathcal{L}(Z,t)$ der Abschluss der Menge

$$\{c_0 + c_1 Z_{s_1} + ... + c_k Z_{s_k} \mid s_j \leq t, c_j \in \mathbb{R}\}$$

in $L^2(\mathbb{P})$. $P_{\mathcal{L}(Z,t)}$ sei die orthogonale Projektion von $L^2(\mathbb{P})$ auf $\mathcal{L}(Z,t)$.

Man kann $\hat{X}_t = \mathbb{E}[X_t \mid \sigma(Z_s \mid s \leq t)] = \mathcal{P}_{K(Z,t)}(X_t)$ zeigen. (Auf Sub-σ-Algebra bedingter Erwartungswert einer ZV ist orthogonale Projektion auf L^2 Teilraum aller messbaren Abbildungen dieser Sub-σ-Algebra.)

Lemma 2.3.1. *X und Z sind Gauß-Prozesse, d.h. jeder endliche Teilvektor von Zufallsvariablen ist mehrdimensional normalverteilt.*

Beweis. Idee: Man kann $(X,Z)^t$ als Lösung einer SDE ansehen, deren Lösung durch einen Grenzwert von Gauß-Prozessen gegeben ist (Stichwort Picarditeration). $\square$

Lemma 2.3.2.

$$\mathcal{P}_{\mathcal{L}}(X_t) = \mathbb{E}[X_t \mid \sigma(Z_s, s \leq t)] = \mathcal{P}_{\mathcal{K}}(X_t).$$

Beweis. $(Y_1, ..., Y_k)$ ist normalverteilt g.d.w. $\sum_{i=1}^k c_i Y_i$ ist normal f.a. c_i. Außerdem ist ein L^2 Grenzwert normalverteilter ZV normal. Also ist $\mathcal{P}_{\mathcal{L}}(X_t)$ normalverteilt.

Def. $\tilde{X}_t := X_t - \mathcal{P}_{\mathcal{L}}(X_t)$.

$\tilde{X}_t$ ist als Differenz von X_t und der Orthogonalprojektion von X_t auf einen Raum, welcher die Z_s enthält, zu den Z_s orthogonal:

$$\mathbb{E}[\mathcal{P}_{\mathcal{L}}(X_t) Z_s] = \mathbb{E}[X_t \mathcal{P}_{\mathcal{L}}(Z_s)] = \mathbb{E}[X_t Z_s],$$

da orthogonale Projektionen selbstadjungiert sind (was diese sogar charakterisiert).
Also $0 = \mathbb{E}[X_t Z_s] - \mathbb{E}[\mathcal{P}_{\mathcal{L}}(X_t) Z_s] = \mathbb{E}[\tilde{X}_t Z_s]$.
Somit sind $\tilde{X}_t$ und Z_s unkorreliert. Da sie normalverteilt sind, sind sie sogar unabhängig. Es folgt für jedes $G \in \sigma(Z_s \mid s \leq t)$

$$\mathbb{E}[1_G(X_t - \mathcal{P}_{\mathcal{L}}(X_t))] = \mathbb{E}[1_G \tilde{X}_t] = \mathbb{E}[1_G]\mathbb{E}[\tilde{X}_t] = 0.$$

$(\mathbb{E}[\mathcal{P}_{\mathcal{L}}(X)] = \mathbb{E}[X \mathcal{P}_{\mathcal{L}}(1)] = \mathbb{E}[X]$, da orthogonale Projektion selbstadjungiert sind.)

Da außerdem $\mathcal{P}_{\mathcal{L}}(X_t)$ $\sigma(Z_s, s \leq t)$-m.b. ist, folgt

$$\mathcal{P}_{\mathcal{L}}(X_t) = \mathbb{E}[X_t \mid \sigma(Z_s, s \leq t)].$$

□

2.4 Explizite Darstellung für $\mathcal{L}(Z, t)$

Lemma 2.4.1. *Es gibt Konstanten A_0, A_1, sodass für jedes $f \in L^2[0,t]$ gilt:*

$$A_0 \int_0^t f^2(s) ds \leq \mathbb{E}[\left(\int_0^t f(s) dZ_s\right)^2] \leq A_1 \int_0^t f^2(s) ds.$$

Wir erhalten also zwei äquivalente Normen.

Beweis auslassen.

$$\mathbb{E}[\left(\int_0^t f(s) dZ_s\right)^2] = \mathbb{E}[\left(\int_0^t f(s) G(s) X_s ds\right)^2]$$

$$+\mathbb{E}[\left(\int_0^t f(s) D(s) dV_s\right)^2] + 2\mathbb{E}[\left(\int_0^t f(s) G(s) X_s ds\right)\left(\int_0^t f(s) D(s) dV_s\right)].$$

$\mathbb{E}[\left(\int_0^t f(s) G(s) X_s ds\right)^2] \leq c \int_0^t f^2(s) ds$ nach Cauchy-Schwartz

$\mathbb{E}[\left(\int\limits_0^t f(s)D(s)dV_s\right)^2] = \int\limits_0^t f^2(s)D^2(s)ds$ nach der Ito-Isometrie □

Lemma 2.4.2. *Sei* $N_0(Z,t) := \{\int_0^t f(s)dZ_s \mid f \in L^2[0,t]\}$*. Es gilt*

$$\mathcal{L}(Z,t) = \mathbb{R} + N_0(Z,t) =: N(Z,t)$$

Beweis. $\supseteq$: Stetige Funktionen sind dicht in $L^2[0,t]$.
Für stetige Funktionen $f \in L^2[0,t] \cap C[0,t]$ gilt

$$\int_0^t f(s)dZ_s = \lim_{n\to\infty} \sum_j f(j2^{-n})(Z_{(j+1)2^{-n}} - Z_{j2^{-n}}).$$

$\subseteq$: Wir zeigen:

(i) $N(Z,t)$ enthält endliche Linearkombinationen der Z_s.

(ii) $N(Z,t)$ ist abgeschlossen in $L^2[0,t]$.

(i) Seien $0 \leq t_1 < ... < t_k \leq t$. Wir können schreiben

$$\begin{aligned}\sum_{i=1}^k c_i Z_{t_i} &= \sum_{i=1}^{k-1} c_i' \Delta Z_{t_i} \\ &= \sum_{i=1}^{k-1} \int\limits_{t_i}^{t_{i+1}} c_i' dZ_s \\ &= \int\limits_0^t \sum_{i=1}^{k-1} c_i' \mathbb{1}_{[t_i,t_{i+1})}(s)dZ_s \in N_0(Z,t)\end{aligned}$$

mit $\Delta Z_{s_i} := Z_{s_{i+1}} - Z_{s_i}$.

(ii) $L^2[0,t] \supseteq N_0(Z,t)$ ist ein Banachraum. $N_0(Z,t)$ ist normierter Vektorraum mit der Norm $Y \mapsto \sqrt{\mathbb{E}[Y^2]}$, welche laut vorigem Lemma äquivalent zu $\|\cdot\|_{L^2[0,t]}$ ist. Somit ist $N_0(Z,t)$ und damit auch $N(Z,t)$ ein abgeschlossener Teilraum.

□

2.5 Der Innovationsprozess

Definiere Innovationsprozess N durch $dN_t = G(t)(X_t - \hat{X}_t)dt + D(t)dV_t$. (Innovation ist in der Signalverarbeitung meist die Differenz der Beobachtung und des optimalen Schätzers.)

Lemma 2.5.1 (Produktregel von Ito)**.** $d(X_tY_t) = X_tdY_t + Y_tdX_t + dX_tdY_t$

Lemma 2.5.2. (i) *N hat orthogonale Inkremente, d.h.*

$$\mathbb{E}[(N_{t_1} - N_{t_0})(N_{s_1} - N_{s_0})] = 0 \ \forall s_0 \leq s_1 \leq t_0 \leq t_1,$$

(ii) $\mathbb{E}[N_t^2] = \int_0^t D^2(s)ds,$

(iii) $\mathcal{L}(N,t) = \mathcal{L}(Z,t),$

(iv) *N ist Gauß'scher Prozess.*

Beweis. (i) Wir zeigen sogar, dass $N_{t_1} - N_{t_0}$ mit jedem $Y \in \mathcal{L}(Z, t_0)$ unkorreliert ist:

$$\begin{aligned}\mathbb{E}[(N_{t_1} - N_{t_0})Y] &= \mathbb{E}[\left(\int_{t_0}^{t_1} G(s)(X_s - \hat{X}_s)ds + \int_{t_0}^{t_1} D(s)dV_s\right) Y] \\ &= \int_{t_0}^{t_1} G(s)\mathbb{E}[(X_s - \hat{X}_s)Y]ds + \mathbb{E}[\int_{t_0}^{t_1} YD(s)dV_s]\end{aligned}$$

Nach Fubini, Linearität des EW und Itos Produktregel (angewendet auf V und den konstanten Prozess Y). Da $(X_s - \hat{X}_s) \perp Y$ und der rechte Prozess ein Martingal ist, sind beide Summanden 0.

(ii) Wir rechnen mit der Ito-Formel für $g(t,x) := x^2$ nach:

$$d(N_t^2) = 2N_tdN_t + \frac{1}{2}2(dN_t)^2 = 2N_tdN_t + D^2(t)dt.$$

Letzterer Term ist deterministisch. Für den Erwartungswert des ersten gilt:

$$\mathbb{E}[N_tdN_t] = \mathbb{E}[\lim_{\Delta t_j \to 0} \sum_j N_{t_j}(N_{t_{j+1}} - N_{t_j})]$$

$$= \lim_{\Delta t_j \to 0} \sum_j \mathbb{E}[(N_{t_j} - N_0)(N_{t_{j+1}} - N_{t_j})]$$
$$= 0.$$

Wir haben hier für das Vertauschen des Grenzwertes und des Erwartungswertes majorisierte Konvergenz mit Majorante $\sup_{r\in[0,t]} |N_r|^2$ verwendet. (Für die Existenz und Integrierbarkeit dieses Supremums nutzen wir, dass N stationäre und unabhängige Inkremente hat, siehe ON THE DISTRIBUTION OF THE SUPREMUM FUNCTIONAL FOR PROCESSES WITH STATIONARY INDEPENDENT INCREMENTS BY GLEN BAXTER AND M. D. DONSKER).

(iii) $\subseteq$: Es ist $N_t = Z_t - \int_0^t G(s)\hat{X}_s ds = Z_t - \int_0^t \mathcal{P}_{\mathcal{L}(Z,s)}(G(s)X_s)ds \in \mathcal{L}(Z,t)$.

$\supseteq$: Idee: Nutze die explizite Darstellung von $\mathcal{L}(Z,t)$ und schreibe

$$Z_t = \int_0^t dZ_s = \int_0^t f(s)c(s)ds + \int_0^t f(s)dN_s$$

für geeignete f, c.

(iv) Hierbei nutzt man wieder, dass Linearkombinationen und Grenzwerte normalverteilter ZV normalverteilt sind.

□

2.6 Der Innovationsprozess und Brownsche Bewegung

Definiere R durch $dR_t := \frac{1}{D(t)} dN_t$.

Lemma 2.6.1. *R ist eine Brownsche Bewegung.*

Beweis. Benutze folgende Charakterisierung Brownscher Bewegung:

(i) R hat stetige Pfade,

(ii) R hat orthogonale Inkremente,

(iii) R ist Gaußscher Prozess,

(iv) $\mathbb{E}[R_t] = 0$, $\mathbb{E}[R_t R_s] = \min(t,s)$.

(i) Als Integral eines stetigen Prozesses hat R stetige Pfade.

(ii)

$$\begin{aligned}
\mathbb{E}[(R_{t_1} - R_{t_0})(R_{s_1} - R_{s_0})] &= \mathbb{E}[\int_{t_0}^{t_1} \frac{1}{D(s)} dN_s \int_{s_0}^{s_1} \frac{1}{D(s)} dN_s] \\
&= \mathbb{E}[\lim \lim \sum \sum (N_{t_{i+1}} - N_{t_i})(N_{s_{j+1}} - N_{s_j})] \\
&= \lim \lim \sum \sum \mathbb{E}[(N_{t_{i+1}} - N_{t_i})(N_{s_{j+1}} - N_{s_j})] \\
&= 0.
\end{aligned}$$

(iii) Als Integral bzgl. eines Gaußschen Prozesses ist R Gaußsch (da Linearkombinationen und L^2-Grenzwerte normalverteilter ZV normalverteilt sind).

(iv) Nach der Ito-Formel

$$d(R_t^2) = 2R_t dR_t + (dR_t)^2 = 2R_t dR_t + dt.$$

Da R orthogonale Inkremente hat, verschwindet im Erwartungswert der erste Term und es gilt $\mathbb{E}[R_t^2] = t$. Für $s < t$ folgt also

$$\mathbb{E}[R_t R_s] = \mathbb{E}[(R_t - R_s)R_s] + \mathbb{E}[R_s^2] = \mathbb{E}[R_s^2] = s.$$

□

Lemma 2.6.2. $\mathcal{L}(N, t) = \mathcal{L}(R, t)$.

Beweis. Es reicht, N_t als Grenzwert von Linearkombinationen von R_s und andersrum schreiben zu können.

$\subseteq$: $N_t = \int_0^t D(s) \frac{1}{D(s)} dN_s = \int_0^t D(s) dR_s = \lim_{\Delta t_j \to 0} \sum_j D(t_j)(R_{t_{j+1}} - R_{t_j})$

$\supseteq$: $R_t = \int_0^t \frac{1}{D(s)} dN_s = \lim_{\Delta t_j \to 0} \sum_j \frac{1}{D(t_j)} (N_{t_{j+1}} - N_{t_j})$

□

Lemma 2.6.3.

$$\hat{X}_t = \mathbb{E}[X_t] + \int_0^t \frac{\partial}{\partial s}\mathbb{E}[X_t R_s] dR_s.$$

Beweis. Aufgrund der Darstellung von $\mathcal{L}(R,t)$ finden wir $c_0(t) \in \mathbb{R}$, $g \in L^2[0,t]$ mit

$$\hat{X}_t = \mathcal{P}_{\mathcal{L}(R,t)}(X_t) = c_0(t) + \int_0^t g(s) dR_s.$$

Das Integral bzgl. R hat EW 0, also gilt $c_0(t) = \mathbb{E}[\hat{X}_t] = \mathbb{E}[X]$. Es ist

$$X_t - \hat{X}_t \perp \int_0^t f(s) dR_s \in \mathcal{L}(R,t) \ \forall f \in L^2[0,t],$$

also

$$\begin{aligned}
\mathbb{E}[X_t \int_0^t f(s) dR_s] &= \mathbb{E}[\hat{X}_t \int_0^t f(s) dR_s] \\
&= \mathbb{E}[\left(c_0(t) + \int_0^t g(s) dR_s\right) \int_0^t f(s) dR_s] \\
&= \mathbb{E}[\int_0^t g(s) dR_s \int_0^t f(s) dR_s] \\
&\overset{\text{Ito-Isometrie}}{=} \mathbb{E}[\int_0^t g(s) f(s) ds] \\
&= \int_0^t g(s) f(s) ds.
\end{aligned}$$

Mit der Wahl $f = \mathbb{1}_{[0,r]}$ erhalten wir $\mathbb{E}[X_t R_r] = \int_0^r g(s) ds$ oder äquivalent $\frac{\partial}{\partial r}\mathbb{E}[X_t R_r] = g(r)$. $\square$

2.7 Stochastische Differentialgleichung für $\hat{X}_t$

Die zuvor berechnete Darstellung von $\hat{X}_t$ ist

$$\hat{X}_t = \mathbb{E}[X_t] + \int_0^t f(s,t) dR_s$$

mit $f(s,t) = \frac{\partial}{\partial s}\mathbb{E}[X_t R_s]$. Wir wollen nun also diesen Erwartungswert berechnen. Es gilt

$$R_s = \int_0^s \frac{G(r)}{D(r)}(X_r - \hat{X}_r)dr + V_s$$

per Definition von R und N:

$$\begin{aligned} R_s &= \int_0^s \frac{1}{D(r)} dN_r \\ &= \int_0^s \frac{1}{D(r)}(G(r)(X_r - \hat{X}_r)dr + D(r)dV_r) \\ &= \int_0^s \frac{G(r)}{D(r)}(X_r - \hat{X}_r)dr + V_s. \end{aligned}$$

Einsetzen ergibt

$$\begin{aligned} \mathbb{E}[X_t R_s] &= \mathbb{E}[X_t \left(\int_0^s \frac{G(r)}{D(r)}(X_r - \hat{X}_r)dr + V_s \right)] \\ &= \int_0^s \frac{G(r)}{D(r)} \mathbb{E}[X_t \tilde{X}_r] dr \end{aligned}$$

mit $\tilde{X}_r \coloneqq X_r - \hat{X}_r$.

Wir haben am Anfang die Lösung der SDE für X gelöst:

$$X_t = \exp\left(\int_0^t F(s)ds \right) \left(X_0 + \int_0^t \exp\left(-\int_0^s F(u)du \right) C(s)dU_s \right).$$

Man kann zeigen, dass dies gleich

$$\exp\left(\int_r^t F(s)ds\right) X_r + \int_r^t \exp\left(\int_s^t F(u)du\right) C(s)dU_s$$

ist für $0 \leq r \leq t$. Einsetzen ergibt (der dU_s-Teil fällt weg)

$$\begin{aligned}
\mathbb{E}[X_t\tilde{X}_r] &= \exp\left(\int_r^t F(s)ds\right) \mathbb{E}[X_r\tilde{X}_r] \\
&= \exp\left(\int_r^t F(s)ds\right) \mathbb{E}[X_r\mathcal{P}_{\mathcal{L}(R,r)}(\mathcal{P}_{\mathcal{L}(R,r)}(X_r))] \\
&= \exp\left(\int_r^t F(s)ds\right) \mathbb{E}[\mathcal{P}_{\mathcal{L}(R,r)}(X_r)\mathcal{P}_{\mathcal{L}(R,r)}(X_r)] \\
&= \exp\left(\int_r^t F(s)ds\right) S(r),
\end{aligned}$$

also

$$\begin{aligned}
f(s,t) &= \frac{\partial}{\partial s}\mathbb{E}[X_tR_s] \\
&= \frac{\partial}{\partial s}\int_0^s \frac{G(r)}{D(r)} \exp\left(\int_s^t F(u)du\right) S(r)dr \\
&= \frac{G(s)}{D(s)} \exp\left(\int_s^t F(u)du\right) S(s).
\end{aligned}$$

Lemma 2.7.1 (Riccati-Differentialgleichung)**.** *Der quadratische Fehler*

$$S(t) = \mathbb{E}[(X_t - \hat{X}_t)^2]$$

erfüllt die Differentialgleichung

$$\frac{dS}{dt}(t) = 2F(t)S(t) - \frac{G^2(t)}{D^2(t)}S^2(t) + C^2(t).$$

Beweis. Es gilt

$$\begin{aligned}\mathbb{E}[(X_t - \hat{X}_t)^2] &= \mathbb{E}[X_t^2] - 2\mathbb{E}[X_t \hat{X}_t] + \mathbb{E}[\hat{X}_t^2] \\ &= \mathbb{E}[X_t^2] - \mathbb{E}[\hat{X}_t^2].\end{aligned}$$

Mit der zuvor berechneten Darstellung von $\hat{X}_t$ folgt

$$\begin{aligned}\mathbb{E}[\hat{X}_t^2] &= \mathbb{E}[(\mathbb{E}[X_t] + \int_0^t f(s,t) dR_s)^2] \\ &= \mathbb{E}[X_t]^2 + 2\mathbb{E}[X_t]\mathbb{E}[\int_0^t f(s,t) dR_s] + \mathbb{E}[(\int_0^t f(s,t) dR_s)^2] \\ &= \mathbb{E}[X_t]^2 + \mathbb{E}[\int_0^t f^2(s,t) ds],\end{aligned}$$

wobei wir im letzten Schritt die Ito-Isometrie verwendet haben. Sei $T(t) \coloneqq \mathbb{E}[X_t^2]$. Es gilt

$$S(t) = T(t) - \int_0^t f^2(s,t) ds - \mathbb{E}[X_t]^2.$$

Erneutes Einsetzen der Darstellung von X_t, Unabhängigkeit von X_0 und U und die Ito-Isometrie ergibt

$$\begin{aligned}T(t) &= \mathbb{E}[X_t^2] \\ &= \mathbb{E}[\left(\exp\left(\int_0^t F(s) ds\right) X_0 + \int_0^t \exp\left(\int_s^t F(u) du\right) C(s) dU_s\right)^2] \\ &= \exp\left(2\int_0^t F(s) ds\right) \mathbb{E}[X_0^2] \\ &+ 2\exp\left(\int_0^t F(s) ds\right) \mathbb{E}[X_0 \int_0^t \exp\left(\int_s^t F(u) du\right) C(s) dU_s] \\ &+ \mathbb{E}[\left(\int_0^t \exp\left(\int_s^t F(u) du\right) C(s) dU_s\right)^2]\end{aligned}$$

$$= \exp\left(2\int_0^t F(s)ds\right)\mathbb{E}[X_0^2] + \int_0^t \exp\left(2\int_s^t F(u)du\right)C^2(s)ds.$$

Da T differenzierbar ist, erhalten wir

$$\begin{aligned}\frac{dT}{dt}(t) &= 2F(t)\exp\left(2\int_0^t F(s)ds\right)\mathbb{E}[X_0^2] + C^2(t)\\ &\quad + \int_0^t 2F(t)\exp\left(2\int_s^t F(u)du\right)C^2(s)ds\\ &= 2F(t)T(t) + C^2(t).\end{aligned}$$

Wir haben hier die Leibnizregel für Parameterintegrale

$$\frac{d}{dt}\int_0^t g(s,t)ds = g(t,t) + \int_0^t \frac{d}{dt}g(s,t)ds$$

ausgenutzt.

Nun leiten wir obige Gleichung für S ab und setzen die Differentialgleichung für T ein:

$$\begin{aligned}\frac{dS}{dt}(t) &= \frac{d}{dt}\left(T(t) - \int_0^t f^2(s,t)ds - \mathbb{E}[X_t]^2\right)\\ &= 2F(t)T(t) + C^2(t) - f^2(t,t) - \int_0^t 2f(s,t)\frac{\partial}{\partial t}f(s,t)ds - 2F(t)\mathbb{E}[X_t].\end{aligned}$$

Hierbei haben wir

$$\begin{aligned}\frac{d}{dt}\mathbb{E}[X_t]^2 &= 2\mathbb{E}[X_t](\frac{d}{dt}\mathbb{E}[X_t])\\ &= 2\mathbb{E}[X_t](\frac{d}{dt}\mathbb{E}[\int_0^t F(s)X_sds + X_0])\\ &\overset{\text{Fubini}}{=} 2\mathbb{E}[X_t](\frac{d}{dt}\int_0^t F(s)\mathbb{E}[X_s]ds)\\ &= 2F(t)\mathbb{E}[X_t]^2\end{aligned}$$

verwendet. Nun ist $f^2(t,t) = \frac{G^2(t)}{D^2(t)} S^2(t)$. Wir rechnen weiter:

$$
\begin{aligned}
&\int\limits_0^t 2f(s,t)\frac{\partial}{\partial t} f(s,t)ds \\
=&\int\limits_0^t 2\frac{G(s)}{D(s)} \exp\left(\int_s^t F(u)du\right) S(s)\frac{\partial}{\partial t}\left(\frac{G(s)}{D(s)} \exp\left(\int_s^t F(u)du\right) S(s)\right) ds \\
=&2F(t)\int\limits_0^t f^2(s,t)ds.
\end{aligned}
$$

Einsetzen ergibt

$$
\frac{dS}{dt}(t) = 2F(t)S(t) + C^2(t) - \frac{G^2(t)}{D^2(t)} S^2(t),
$$

was zu zeigen war. □

Wir können nun endlich die stochastische Differentialgleichung für $\hat{X}_t$ aufstellen: Wir wollen im Folgenden die Formel $\hat{X}_t = c_0(t) + \int\limits_0^t f(s,t)dR_s$ mit $c_0(t) = \mathbb{E}[X_t]$ nutzen. Mit einer stochastischen Leibnizformel erhalten wir

$$
\begin{aligned}
d\hat{X}_t &= c_0'(t)dt + f(t,t)dR_t + \left(\int\limits_0^t \frac{\partial}{\partial t} f(s,t)dR_s\right) dt \\
&= c_0'(t)dt + \frac{G(t)S(t)}{D(t)} dR_t + \left(\int\limits_0^t f(s,t)dR_s\right) F(t)dt \\
&= F(t)c_0(t)dt + \frac{G(t)S(t)}{D(t)} dR_t + (\hat{X}_t - c_0(t))F(t)dt \\
&= \hat{X}_t F(t)dt + \frac{G(t)S(t)}{D(t)} dR_t.
\end{aligned}
$$

Nun setzen wir die Definition für R ein

$$
dR_t = \frac{1}{D(t)} dN_t = \frac{1}{D(t)} \left(dZ_t - G(t)\hat{X}_t dt\right)
$$

und erhalten

$$
\begin{aligned}
d\hat{X}_t &= \hat{X}_t F(t)dt + \frac{G(t)S(t)}{D(t)} dR_t \\
&= \hat{X}_t F(t)dt + \frac{G(t)S(t)}{D(t)} \frac{1}{D(t)} \left(dZ_t - G(t)\hat{X}_t dt\right) \\
&= \left(F(t) - \frac{G^2(t)S(t)}{D^2(t)}\right) \hat{X}_t dt + \frac{G(t)S(t)}{D^2(t)} dZ_t.
\end{aligned}
$$

Dies schließt den Beweis ab.

3 Beispiele

3.1 Konstanter Prozess

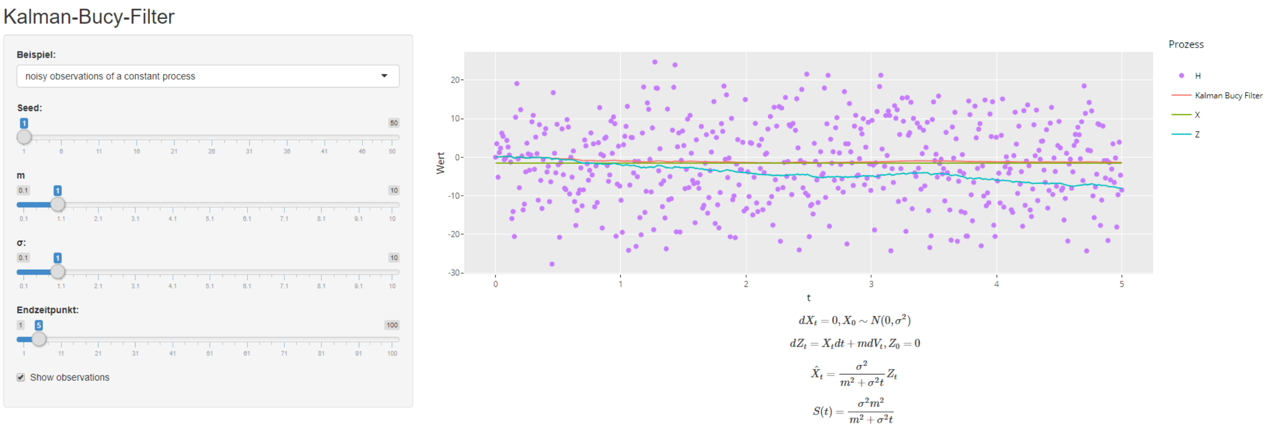

Abbildung 1: Lineares Modell mit Observierungen

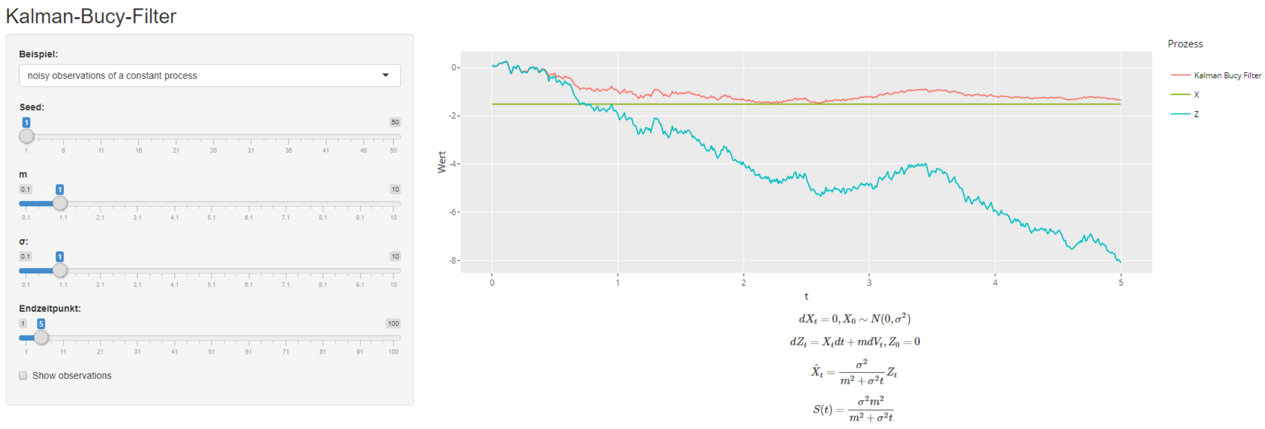

Abbildung 2: Lineares Modell ohne Observierungen

3.2 Brownsche Bewegung

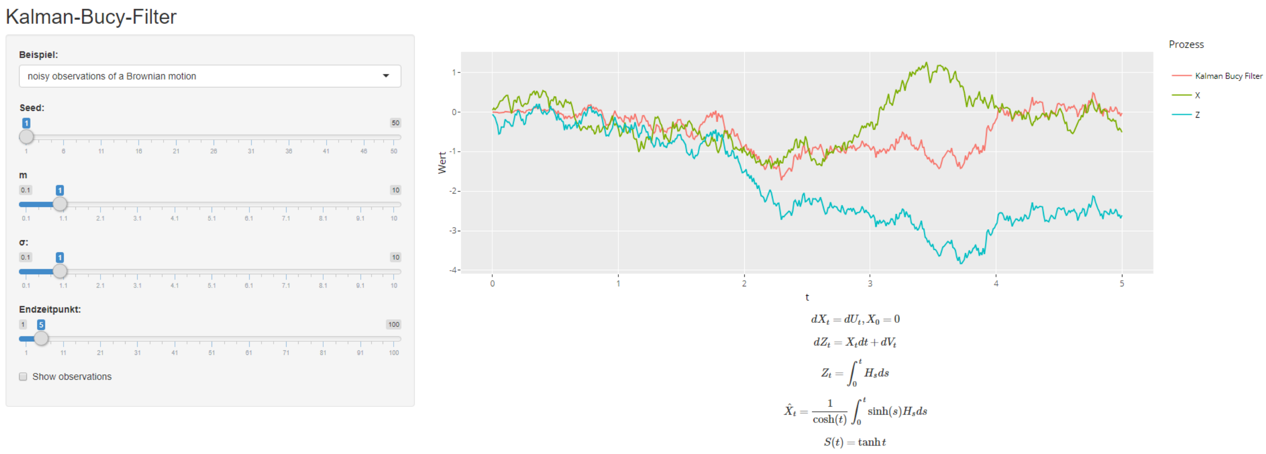

Abbildung 3: Brownsche Bewegung

Fehler geht nicht gegen 0 (er wächst sogar). D.h. i.A. gibt es keinen konsistenten Schätzer.

3.3 Bevölkerungswachstum

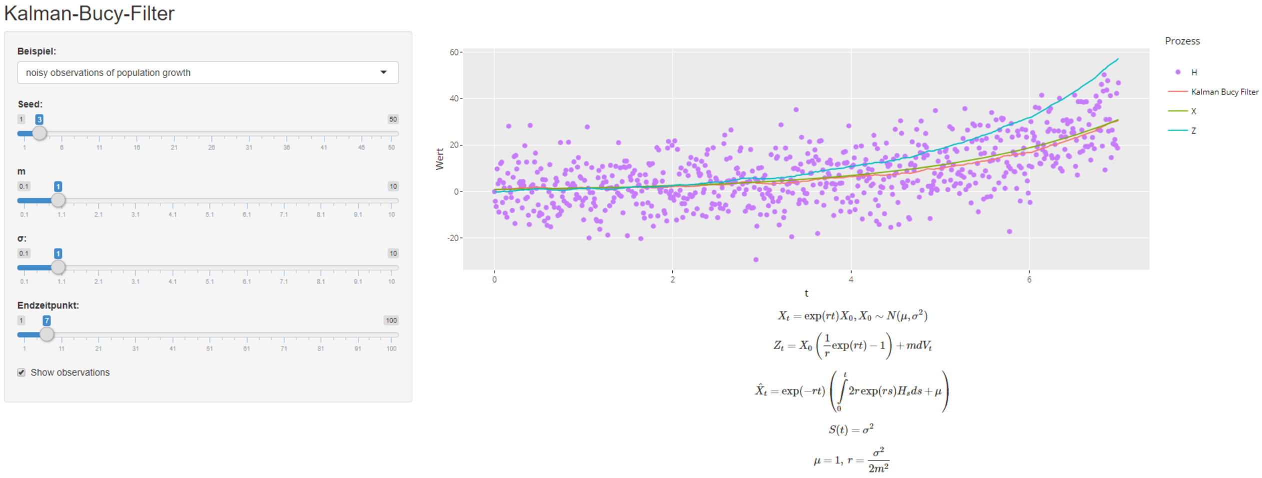

Abbildung 4: Populationswachstum

3.4 Falsches Modell

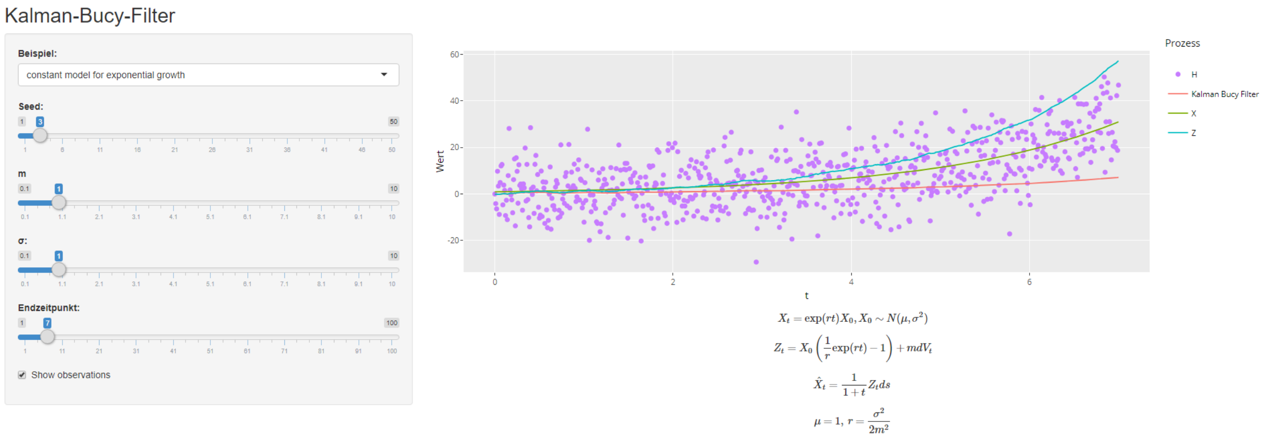

Abbildung 5: Filter für das konstante Modell angewendet auf Populationswachstum