Kalman-Bucy Filter in stetiger Zeit

Achille Mbope und Dominic Arnold

9. Februar 2011

Lemma (Allgemeine Itô-Formel)

Sei $dX_t(t) = udt + vdB(t)$ ein n-dimensionaler Itô-Prozeß sowie $g(t,x) = (g_1(t,x), ..., g_p(t,x)) \in C^2([0,\infty) \times \mathbb{R}^p)$.
Dann ist der Prozeß $Y(t,w) = g(t,X_t)$ wieder ein Itô-Prozeß und seine $k-te$ Komponente Y_k ist gegeben durch

$$dY_k = \frac{\partial g_k}{\partial t}(t,X)dt + \sum_i \frac{\partial g_k}{\partial X_i}(t,X)dX_i + \frac{1}{2}\sum_{i,j} \frac{\partial^2 g_k}{\partial x_i \partial x_j}(t,X)dX_i dX_j$$

wobei $dB_i dB_j = \delta_{ij}dt$, $dB_i dt = dtdB_i = 0$

Wichtige Begriffe

Lemma (Allgemeine Itô-Formel)

Sei $dX_t(t) = udt + vdB(t)$ ein n-dimensionaler Itô-Prozeß sowie $g(t,x) = (g_1(t,x), ..., g_p(t,x)) \in C^2([0,\infty) \times \mathbb{R}^p)$.
Dann ist der Prozeß $Y(t,w) = g(t,X_t)$ wieder ein Itô-Prozeß und seine $k-te$ Komponente Y_k ist gegeben durch

$$dY_k = \frac{\partial g_k}{\partial t}(t,X)dt + \sum_i \frac{\partial g_k}{\partial X_i}(t,X)dX_i + \frac{1}{2}\sum_{i,j} \frac{\partial^2 g_k}{\partial x_i \partial x_j}(t,X)dX_i dX_j$$

wobei $dB_i dB_j = \delta_{ij} dt$, $dB_i dt = dt dB_i = 0$

Lemma (Itô-Isometrie)

Wichtige Begriffe

Lemma (Allgemeine Itô-Formel)

Sei $dX_t(t) = udt + vdB(t)$ *ein n-dimensionaler Itô-Prozeß sowie* $g(t,x) = (g_1(t,x), ..., g_p(t,x)) \in C^2([0,\infty) \times \mathbb{R}^p)$.
Dann ist der Prozeß $Y(t,w) = g(t, X_t)$ *wieder ein Itô-Prozeß und seine* $k-te$ *Komponente* Y_k *ist gegeben durch*

$$dY_k = \frac{\partial g_k}{\partial t}(t,X)dt + \sum_i \frac{\partial g_k}{\partial X_i}(t,X)dX_i + \frac{1}{2}\sum_{i,j}\frac{\partial^2 g_k}{\partial x_i \partial x_j}(t,X)dX_i dX_j$$

wobei $dB_i dB_j = \delta_{ij}dt$, $dB_i dt = dtdB_i = 0$

Lemma (Itô-Isometrie)

Sei W *ein Wiener Prozeß und* X_t *ein 1-dimensionaler stochastischer Prozeß, welcher adaptiert ist bezüglich der Filtration des Wiener Prozeß. Dann gilt*

Wichtige Begriffe

Lemma (Allgemeine Itô-Formel)

Sei $dX_t(t) = udt + vdB(t)$ ein n-dimensionaler Itô-Prozeß sowie $g(t,x) = (g_1(t,x), ..., g_p(t,x)) \in C^2([0,\infty) \times \mathbb{R}^p)$.
Dann ist der Prozeß $Y(t,w) = g(t,X_t)$ wieder ein Itô-Prozeß und seine $k-te$ Komponente Y_k ist gegeben durch

$$dY_k = \frac{\partial g_k}{\partial t}(t,X)dt + \sum_i \frac{\partial g_k}{\partial X_i}(t,X)dX_i + \frac{1}{2}\sum_{i,j} \frac{\partial^2 g_k}{\partial x_i \partial x_j}(t,X)dX_i dX_j$$

wobei $dB_i dB_j = \delta_{ij} dt$, $dB_i dt = dt dB_i = 0$

Lemma (Itô-Isometrie)

Sei W ein Wiener Prozeß und X_t ein 1-dimensionaler stochastischer Prozeß, welcher adaptiert ist bezüglich der Filtration des Wiener Prozeß. Dann gilt

$$\mathbb{E}\left[\left(\int_0^T X_t dW_t\right)^2\right] = \mathbb{E}\left[\int_0^T X_t^2 dt\right]$$

Wichtige Begriffe

Ähnliche Aussagen gelten auch für d-dimensionale Brownsche Bewegungen $(B_t, t \geq 0)$ und (nicht zwangsweise deterministische) Funktionen $f, g : \mathbb{R}^+ \to \mathbb{R}^{n \times d}$:

$$\begin{aligned}
\mathbb{E}\left[\left(\int_0^t g(s)dB_s\right)^T \left(\int_0^t f(s)dB_s\right)\right] &= \mathbb{E}\left[\int_0^t spur\left(g(s)^T f(s)\right) ds\right] \\
\mathbb{E}\left[\left(\int_0^t g(s)dB_s\right) \left(\int_0^t f(s)dB_s\right)^T\right] &= \mathbb{E}\left[\int_0^t g(s) f(s)^T ds\right]
\end{aligned}$$

Wichtige Begriffe

Theorem

Sei $X_j : \Omega \longrightarrow \mathbb{R}$ *eine Zufallsvariable;* $0 \leq j \leq n$. *Dann ist*

$$X = (X_1, ..., X_n)$$

normalverteilt genau dann wenn

$$Y = \lambda_1 X_1 + ... + \lambda_n X_n$$

normalverteilt ist für alle $\lambda_1, ...\lambda_n \in \mathbb{R}$

Wichtige Begriffe

Theorem

Sei $X_j : \Omega \longrightarrow \mathbb{R}$ eine Zufallsvariable; $0 \leq j \leq n$. Dann ist

$$X = (X_1, ..., X_n)$$

normalverteilt genau dann wenn

$$Y = \lambda_1 X_1 + ... + \lambda_n X_n$$

normalverteilt ist für alle $\lambda_1, ...\lambda_n \in \mathbb{R}$

Theorem

Annahme: $X_k : \Omega \longrightarrow \mathbb{R}^n$ ist normalverteilt für alle k und $X_k \longrightarrow X$ in $L^2(\Omega)$, dh

$$\mathbb{E}\left[\|X_k - X\|^2\right] \longrightarrow 0 \qquad k \longrightarrow \infty$$

Dann ist X normalverteilt

Problemstellung

Gegeben sei ein n-dimensionaler Prozess $(X_t)_{t\geq 0}$, beschrieben durch folgende stochastische Differentialgleichung:

$$dX_t = F(t)X_t dt + C(t)dU_t, \tag{1}$$

Problemstellung

Gegeben sei ein n-dimensionaler Prozess $(X_t)_{t\geq 0}$, beschrieben durch folgende stochastische Differentialgleichung:

$$dX_t = F(t)X_t dt + C(t)dU_t, \tag{1}$$

wobei $F : \mathbb{R}_0^+ \longrightarrow \mathbb{R}^{n\times n}$ und $C : \mathbb{R}_0^+ \longrightarrow \mathbb{R}^{n\times p}$ matrixwertige Funktionen sind, die auf beschränkten Intervallen $[0, T]$ beschränkt sein sollen. Der Prozeß $(U_t)_{t\geq 0}$ sei eine p-dimensionale Brownsche Bewegung.

Problemstellung

Gegeben sei ein n-dimensionaler Prozess $(X_t)_{t\geq 0}$, beschrieben durch folgende stochastische Differentialgleichung:

$$dX_t = F(t)X_t dt + C(t)dU_t, \tag{1}$$

wobei $F : \mathbb{R}_0^+ \longrightarrow \mathbb{R}^{n\times n}$ und $C : \mathbb{R}_0^+ \longrightarrow \mathbb{R}^{n\times p}$ matrixwertige Funktionen sind, die auf beschränkten Intervallen $[0, T]$ beschränkt sein sollen. Der Prozeß $(U_t)_{t\geq 0}$ sei eine p-dimensionale Brownsche Bewegung.
Dieser kann nicht direkt beobachtet werden, sondern nur verrauscht über einen m-dimensionalen Beobachtungsprozess $(Z_t)_{t\geq 0}$, beschrieben durch die stochastische Differentialgleichung

$$dZ_t = G(t)X_t dt + D(t)dV_t. \tag{2}$$

Problemstellung

Gegeben sei ein n-dimensionaler Prozess $(X_t)_{t\geq 0}$, beschrieben durch folgende stochastische Differentialgleichung:

$$dX_t = F(t)X_t dt + C(t)dU_t, \tag{1}$$

wobei $F : \mathbb{R}_0^+ \longrightarrow \mathbb{R}^{n\times n}$ und $C : \mathbb{R}_0^+ \longrightarrow \mathbb{R}^{n\times p}$ matrixwertige Funktionen sind, die auf beschränkten Intervallen $[0, T]$ beschränkt sein sollen. Der Prozeß $(U_t)_{t\geq 0}$ sei eine p-dimensionale Brownsche Bewegung.
Dieser kann nicht direkt beobachtet werden, sondern nur verrauscht über einen m-dimensionalen Beobachtungsprozess $(Z_t)_{t\geq 0}$, beschrieben durch die stochastische Differentialgleichung

$$dZ_t = G(t)X_t dt + D(t)dV_t. \tag{2}$$

Wir nehmen erneut an, dass die matrixwertigen Funktionen $G : \mathbb{R}_0^+ \to \mathbb{R}^{m\times n}$ und $D : \mathbb{R}_0^+ \to \mathbb{R}^{m\times r}$ beschränkt sind auf beschränkten Intervallen $[0, T]$ und der Prozeß $(V_t, t \geq 0)$ eine von U unabhängige, r-dimensionale Brownsche Bewegung sei. Weiterhin sei X_0 normalverteilt und unabhängig von U sowie V, $Z_0 = 0$ und $D(t)D(t)^T$ sei invertierbar.

Problemstellung

Ferner sei $\mathcal{G}_t = \sigma\left(\{Z_s : s \leq t\}\right)$, die vom Prozeß $(Z_t, t \geq 0)$ erzeugte Filtration.

Problemstellung

Ferner sei $\mathcal{G}_t = \sigma\left(\{Z_s : s \leq t\}\right)$, die vom Prozeß $(Z_t, t \geq 0)$ erzeugte Filtration. Das Filterproblem besteht darin, einen bester Schätzer $\hat{X}_t$ für X_t zu finden, gegeben die Beobachtung Z_s, $s \leq t$, dh. wir fordern das folgende für $\hat{X}_t$

Problemstellung

Ferner sei $\mathcal{G}_t = \sigma\left(\{Z_s : s \leq t\}\right)$, die vom Prozeß $(Z_t, t \geq 0)$ erzeugte Filtration. Das Filterproblem besteht darin, einen bester Schätzer $\hat{X}_t$ für X_t zu finden, gegeben die Beobachtung Z_s, $s \leq t$, dh. wir fordern das folgende für $\hat{X}_t$

- $\hat{X}_t$ soll $\mathcal{G}_t$-messbar sein

Problemstellung

Ferner sei $\mathcal{G}_t = \sigma\left(\{Z_s : s \leq t\}\right)$, die vom Prozeß $(Z_t, t \geq 0)$ erzeugte Filtration. Das Filterproblem besteht darin, einen bester Schätzer $\hat{X}_t$ für X_t zu finden, gegeben die Beobachtung Z_s, $s \leq t$, dh. wir fordern das folgende für $\hat{X}_t$

- $\hat{X}_t$ soll $\mathcal{G}_t$-messbar sein
- $\hat{X}_t$ soll optimal im L^2-Sinne sein, dh.
  $\mathbb{E}\left[\|X_t - \hat{X}_t\|^2\right] = \inf\left\{\mathbb{E}\left[\|X_t - Y\|^2\right] : Y \in L^2(\Omega, \mathcal{G}_t, \mathbb{P})\right\}$

Problemstellung

Ferner sei $\mathcal{G}_t = \sigma(\{Z_s : s \leq t\})$, die vom Prozeß $(Z_t, t \geq 0)$ erzeugte Filtration. Das Filterproblem besteht darin, einen bester Schätzer $\hat{X}_t$ für X_t zu finden, gegeben die Beobachtung Z_s, $s \leq t$, dh. wir fordern das folgende für $\hat{X}_t$

- $\hat{X}_t$ soll $\mathcal{G}_t$-messbar sein
- $\hat{X}_t$ soll optimal im L^2-Sinne sein, dh.
  $\mathbb{E}\left[\|X_t - \hat{X}_t\|^2\right] = \inf\left\{\mathbb{E}\left[\|X_t - Y\|^2\right] : Y \in L^2(\Omega, \mathcal{G}_t, \mathbb{P})\right\}$

d.h. der gesuchte Prozeß zum Zeitpunkt t $\hat{X}_t$ ist die Projektion von X_t auf $L^2(\Omega, \mathcal{G}_t, \mathbb{P})$. Dann wissen wir:

Problemstellung

Ferner sei $\mathcal{G}_t = \sigma\left(\{Z_s : s \leq t\}\right)$, die vom Prozeß $(Z_t, t \geq 0)$ erzeugte Filtration. Das Filterproblem besteht darin, einen bester Schätzer $\hat{X}_t$ für X_t zu finden, gegeben die Beobachtung Z_s, $s \leq t$, dh. wir fordern das folgende für $\hat{X}_t$

- $\hat{X}_t$ soll $\mathcal{G}_t$-messbar sein
- $\hat{X}_t$ soll optimal im L^2-Sinne sein, dh.
  $\mathbb{E}\left[\|X_t - \hat{X}_t\|^2\right] = \inf\left\{\mathbb{E}\left[\|X_t - Y\|^2\right] : Y \in L^2(\Omega, \mathcal{G}_t, \mathbb{P})\right\}$

d.h. der gesuchte Prozeß zum Zeitpunkt t $\hat{X}_t$ ist die Projektion von X_t auf $L^2(\Omega, \mathcal{G}_t, \mathbb{P})$. Dann wissen wir:

Satz

$$\hat{X}_t = \mathcal{P}_{\mathcal{K}(Z,t)}(X_t) = \mathbb{E}\left[X_t \,|\, \mathcal{G}_t\right]$$

Problemstellung

Ferner sei $\mathcal{G}_t = \sigma(\{Z_s : s \leq t\})$, die vom Prozeß $(Z_t, t \geq 0)$ erzeugte Filtration. Das Filterproblem besteht darin, einen bester Schätzer $\hat{X}_t$ für X_t zu finden, gegeben die Beobachtung Z_s, $s \leq t$, dh. wir fordern das folgende für $\hat{X}_t$

- $\hat{X}_t$ soll $\mathcal{G}_t$-messbar sein
- $\hat{X}_t$ soll optimal im L^2-Sinne sein, dh.
  $\mathbb{E}\left[\|X_t - \hat{X}_t\|^2\right] = \inf\left\{\mathbb{E}\left[\|X_t - Y\|^2\right] : Y \in L^2(\Omega, \mathcal{G}_t, \mathbb{P})\right\}$

d.h. der gesuchte Prozeß zum Zeitpunkt t $\hat{X}_t$ ist die Projektion von X_t auf $L^2(\Omega, \mathcal{G}_t, \mathbb{P})$. Dann wissen wir:

Satz

$$\hat{X}_t = \mathcal{P}_{\mathcal{K}(Z,t)}(X_t) = \mathbb{E}\left[X_t \,|\, \mathcal{G}_t\right]$$

$$\mathcal{K}(Z,t) := \left\{Y : \Omega \longrightarrow \mathbb{R}^n;\, Y \in L^2(\mathbb{P}) \textit{ und } Y\ \mathcal{G}_t\textit{-meßbar}\right\}$$

Problemstellung

Ferner sei $\mathcal{G}_t = \sigma\left(\{Z_s : s \leq t\}\right)$, die vom Prozeß $(Z_t, t \geq 0)$ erzeugte Filtration. Das Filterproblem besteht darin, einen bester Schätzer $\hat{X}_t$ für X_t zu finden, gegeben die Beobachtung Z_s, $s \leq t$, dh. wir fordern das folgende für $\hat{X}_t$

- $\hat{X}_t$ soll $\mathcal{G}_t$-messbar sein
- $\hat{X}_t$ soll optimal im L^2-Sinne sein, dh.
  $\mathbb{E}\left[\|X_t - \hat{X}_t\|^2\right] = \inf\left\{\mathbb{E}\left[\|X_t - Y\|^2\right] : Y \in L^2(\Omega, \mathcal{G}_t, \mathbb{P})\right\}$

d.h. der gesuchte Prozeß zum Zeitpunkt t $\hat{X}_t$ ist die Projektion von X_t auf $L^2(\Omega, \mathcal{G}_t, \mathbb{P})$. Dann wissen wir:

Satz

$$\hat{X}_t = \mathcal{P}_{\mathcal{K}(Z,t)}(X_t) = \mathbb{E}\left[X_t \mid \mathcal{G}_t\right]$$

$$\mathcal{K}(Z,t) := \left\{Y : \Omega \longrightarrow \mathbb{R}^n ; Y \in L^2(\mathbb{P}) \textit{ und } Y\ \mathcal{G}_t\textit{-meßbar}\right\}$$

Der gesuchte Prozeß wird auch **Kalman-Bucy-Filter** in stetiger Zeit genannt.

Problemstellung

Ferner sei $\mathcal{G}_t = \sigma\left(\{Z_s : s \leq t\}\right)$, die vom Prozeß $(Z_t, t \geq 0)$ erzeugte Filtration. Das Filterproblem besteht darin, einen bester Schätzer $\hat{X}_t$ für X_t zu finden, gegeben die Beobachtung Z_s, $s \leq t$, dh. wir fordern das folgende für $\hat{X}_t$

- $\hat{X}_t$ soll $\mathcal{G}_t$-messbar sein
- $\hat{X}_t$ soll optimal im L^2-Sinne sein, dh.
  $\mathbb{E}\left[\|X_t - \hat{X}_t\|^2\right] = \inf\left\{\mathbb{E}\left[\|X_t - Y\|^2\right] : Y \in L^2(\Omega, \mathcal{G}_t, \mathbb{P})\right\}$

d.h. der gesuchte Prozeß zum Zeitpunkt t $\hat{X}_t$ ist die Projektion von X_t auf $L^2(\Omega, \mathcal{G}_t, \mathbb{P})$. Dann wissen wir:

Satz

$$\hat{X}_t = \mathcal{P}_{\mathcal{K}(Z,t)}(X_t) = \mathbb{E}\left[X_t \,|\, \mathcal{G}_t\right]$$

$$\mathcal{K}(Z,t) := \left\{Y : \Omega \longrightarrow \mathbb{R}^n;\, Y \in L^2(\mathbb{P}) \textit{ und } Y\ \mathcal{G}_t\textit{ -meßbar}\right\}$$

Der gesuchte Prozeß wird auch **Kalman-Bucy-Filter** in stetiger Zeit genannt. Das Ziel dieses Vortrags ist die Berechnung einer SDE, die $\hat{X}_t$ beschreibt.

Berechnung des Filters

Die Konstruktion des Schätzers gliedert sich in 5 Schritte:

Berechnung des Filters

Die Konstruktion des Schätzers gliedert sich in 5 Schritte:

- **Schritt 1**
  Der Z-linearer Schätzer entspricht dem besten Z-messbaren Schätzer

Die Konstruktion des Schätzers gliedert sich in 5 Schritte:

- **Schritt 1**
  Der Z-linearer Schätzer entspricht dem besten Z-messbaren Schätzer
- **Schritt 2**
  Der Innovationsprozeß

Berechnung des Filters

Die Konstruktion des Schätzers gliedert sich in 5 Schritte:

- **Schritt 1**
  Der Z-linearer Schätzer entspricht dem besten Z-messbaren Schätzer
- **Schritt 2**
  Der Innovationsprozeß
- **Schritt 3**
  Innovationsprozeß und Brownsche Bewegung

Berechnung des Filters

Die Konstruktion des Schätzers gliedert sich in 5 Schritte:

- **Schritt 1**
  Der Z-linearer Schätzer entspricht dem besten Z-messbaren Schätzer
- **Schritt 2**
  Der Innovationsprozeß
- **Schritt 3**
  Innovationsprozeß und Brownsche Bewegung
- **Schritt 4**
  Explizite Formel für X_t

Berechnung des Filters

Die Konstruktion des Schätzers gliedert sich in 5 Schritte:

- **Schritt 1**
  Der Z-linearer Schätzer entspricht dem besten Z-messbaren Schätzer
- **Schritt 2**
  Der Innovationsprozeß
- **Schritt 3**
  Innovationsprozeß und Brownsche Bewegung
- **Schritt 4**
  Explizite Formel für X_t
- **Schritt 5**
  Stochastische Differentialgleichung für $\hat{X}_t$

Im Gaußschen Fall ist der lineare Schätzer gleich dem besten Schätzer

Wir wissen aus Überlegungen zum bedingten Erwartungswert, daß ein linearer Schätzer nicht optimal zu sein braucht. Wir werden aber sehen, dass im Gaußschen Fall der lineare Schätzer gleich dem besten Schätzer ist

Lemma

Im Gaußschen Fall ist der lineare Schätzer gleich dem besten Schätzer

Wir wissen aus Überlegungen zum bedingten Erwartungswert, daß ein linearer Schätzer nicht optimal zu sein braucht. Wir werden aber sehen, dass im Gaußschen Fall der lineare Schätzer gleich dem besten Schätzer ist

Lemma

Seien $X, Z_s; s \leq t$ Zuallsvektoren in $L^2(\mathbb{P})$.

Im Gaußschen Fall ist der lineare Schätzer gleich dem besten Schätzer

Wir wissen aus Überlegungen zum bedingten Erwartungswert, daß ein linearer Schätzer nicht optimal zu sein braucht. Wir werden aber sehen, dass im Gaußschen Fall der lineare Schätzer gleich dem besten Schätzer ist

Lemma

Seien $X, Z_s; s \leq t$ Zuallsvektoren in $L^2(\mathbb{P})$.
Wir nehmen an dass $(X, Z_{s_1}, Z_{s_2}, ..., Z_{s_n})$ normalverteilt ist für alle $s_1, s_2, ..., s_n \leq t,, n \geq 1$.

Im Gaußschen Fall ist der lineare Schätzer gleich dem besten Schätzer

Wir wissen aus Überlegungen zum bedingten Erwartungswert, daß ein linearer Schätzer nicht optimal zu sein braucht. Wir werden aber sehen, dass im Gaußschen Fall der lineare Schätzer gleich dem besten Schätzer ist

Lemma

Seien $X, Z_s; s \leq t$ Zuallsvektoren in $L^2(\mathbb{P})$.
Wir nehmen an dass $(X, Z_{s_1}, Z_{s_2}, ..., Z_{s_n})$ normalverteilt ist für alle $s_1, s_2, ..., s_n \leq t,, n \geq 1$.
Wir betrachten den folgenden Raum: $\mathcal{L}(Z, t) :=$
$cl\{C_0 + C_1 Z_{s_1} + \cdots + C_n Z_{s_n}; C_0 \in \mathbb{R}^n, C_i \in \mathbb{R}^{n \times m}$ mit $s_1, \ldots, s_n \leq t$ und $n \in \mathbb{N}\}$
Der Abschluss ist im L^2-Sinn gemeint.

Im Gaußschen Fall ist der lineare Schätzer gleich dem besten Schätzer

Wir wissen aus Überlegungen zum bedingten Erwartungswert, daß ein linearer Schätzer nicht optimal zu sein braucht. Wir werden aber sehen, dass im Gaußschen Fall der lineare Schätzer gleich dem besten Schätzer ist

Lemma

Seien $X, Z_s; s \leq t$ Zuallsvektoren in $L^2(\mathbb{P})$.
Wir nehmen an dass $(X, Z_{s_1}, Z_{s_2}, ..., Z_{s_n})$ normalverteilt ist für alle $s_1, s_2, ..., s_n \leq t,, n \geq 1$.
Wir betrachten den folgenden Raum: $\mathcal{L}(Z,t) :=$
$cl\{C_0 + C_1 Z_{s_1} + \cdots + C_n Z_{s_n}; C_0 \in \mathbb{R}^n, C_i \in \mathbb{R}^{n\times m}$ mit $s_1, \ldots, s_n \leq t$ und $n \in \mathbb{N}\}$
Der Abschluss ist im L^2-Sinn gemeint. Dann gilt:

$$\hat{X}_t = \mathcal{P}_{\mathcal{K}(Z,t)}(X_t) = \mathbb{E}\left[X_t \,|\, \mathcal{G}_t\right] = \mathcal{P}_{\mathcal{L}(Z,t)}(X_t)$$

Im Gaußschen Fall ist der lineare Schätzer gleich dem besten Schätzer

Wir wissen aus Überlegungen zum bedingten Erwartungswert, daß ein linearer Schätzer nicht optimal zu sein braucht. Wir werden aber sehen, dass im Gaußschen Fall der lineare Schätzer gleich dem besten Schätzer ist

Lemma

Seien $X, Z_s; s \leq t$ Zuallsvektoren in $L^2(\mathbb{P})$.
Wir nehmen an dass $(X, Z_{s_1}, Z_{s_2}, ..., Z_{s_n})$ normalverteilt ist für alle $s_1, s_2, ..., s_n \leq t,, n \geq 1$.
Wir betrachten den folgenden Raum: $\mathcal{L}(Z,t) := cl\{C_0 + C_1 Z_{s_1} + \cdots + C_n Z_{s_n}; C_0 \in \mathbb{R}^n, C_i \in \mathbb{R}^{n\times m}$ mit $s_1, \ldots, s_n \leq t$ und $n \in \mathbb{N}\}$
Der Abschluss ist im L^2-Sinn gemeint. Dann gilt:

$$\hat{X}_t = \mathcal{P}_{\mathcal{K}(Z,t)}(X_t) = \mathbb{E}\left[X_t \,|\, \mathcal{G}_t\right] = \mathcal{P}_{\mathcal{L}(Z,t)}(X_t)$$

Beweis:

Im Gaußschen Fall ist der lineare Schätzer gleich dem besten schätzer

Um das vorherige Lemma auf unser Filterproblem anwenden zu können brauchen wir noch:

Im Gaußschen Fall ist der lineare Schätzer gleich dem besten schätzer

Um das vorherige Lemma auf unser Filterproblem anwenden zu können brauchen wir noch:

Lemma

$$M_t = \begin{pmatrix} X_t \\ Z_t \end{pmatrix}$$

ist ein Gaußscher Prozeß

Im Gaußschen Fall ist der lineare Schätzer gleich dem besten schätzer

Um das vorherige Lemma auf unser Filterproblem anwenden zu können brauchen wir noch:

Lemma

$$M_t = \begin{pmatrix} X_t \\ Z_t \end{pmatrix}$$

ist ein Gaußscher Prozeß

Beweis:

Sei $\mathcal{L}(Z,T)$ der im vorherigen Lemma definierte Raum. Man kann zeigen, dass gilt:

Lemma

$\mathcal{L}(Z,T) = \{C_0 + \int_0^T f(t)dZ_t \; mit \; C_0 \in \mathbb{R}^n, \; f : [0,T] \longrightarrow \mathbb{R}^{n\times m}\}$ *wobei für die Funktion f gelten soll:* $f_{ij} \in L^2([0,T])$.

Sei $\mathcal{L}(Z,T)$ der im vorherigen Lemma definierte Raum. Man kann zeigen, dass gilt:

Lemma

$\mathcal{L}(Z,T) = \{C_0 + \int_0^T f(t)dZ_t \text{ mit } C_0 \in \mathbb{R}^n,\ f : [0,T] \longrightarrow \mathbb{R}^{n\times m}\}$ wobei für die Funktion f gelten soll: $f_{ij} \in L^2([0,T])$.

Definition

Der Innovationsprozeß ist wie folgt definiert

$$N_t = Z_t - \int_0^t \mathcal{G}_t(s)\hat{X}_{t\,s}ds$$

bzw.

$$dN_t = G(t)(X_t - \hat{X}_{t\,t})dt + D(t)dV_t$$

Der Innovationsprozeß

Lemma

- *N_t hat orthogonale Zuwächse*
- $\mathbb{E}\left[N_t N_t^T\right] = \int_0^t D(s)D^T(s)ds$
- $\mathcal{L}(N,t) = \mathcal{L}(Z,t)$
- $(N_t, t \geq 0)$ *ist ein Gaußscher Prozeß.*

Aus diesem Lemma lässt sich sofort schlußfolgern:

$$\hat{X}_t = \mathbb{E}\left[X_t | \mathcal{G}_t\right]$$

Vorbemerkung zu Schritt 3

Offensichtlich ist die Matix $DD^T \in \mathbb{R}^{m \times m}$ symmetrisch und positiv definit. Dann wissen wir, dass eine Matrix $A \in \mathbb{R}^{m \times m}$ existiert, so dass gilt:

Vorbemerkung zu Schritt 3

Offensichtlich ist die Matix $DD^T \in \mathbb{R}^{m \times m}$ symmetrisch und positiv definit. Dann wissen wir, dass eine Matrix $A \in \mathbb{R}^{m \times m}$ existiert, so dass gilt:

$$A^2 = DD^T \Longrightarrow \left(DD^T\right)^{\frac{1}{2}} := A \text{ ist wohldefiniert.}$$

Vorbemerkung zu Schritt 3

Offensichtlich ist die Matix $DD^T \in \mathbb{R}^{m \times m}$ symmetrisch und positiv definit. Dann wissen wir, dass eine Matrix $A \in \mathbb{R}^{m \times m}$ existiert, so dass gilt:

$$A^2 = DD^T \Longrightarrow \left(DD^T\right)^{\frac{1}{2}} := A \text{ ist wohldefiniert.}$$

Desweiteren gilt aufgrund der Invertierbarkeit von DD^T, dass $\left(DD^T\right)^{\frac{1}{2}}$ ebenfalls invertierbar ist und dass gilt:

Offensichtlich ist die Matix $DD^T \in \mathbb{R}^{m\times m}$ symmetrisch und positiv definit. Dann wissen wir, dass eine Matrix $A \in \mathbb{R}^{m\times m}$ existiert, so dass gilt:

$$A^2 = DD^T \Longrightarrow \left(DD^T\right)^{\frac{1}{2}} := A \text{ ist wohldefiniert.}$$

Desweiteren gilt aufgrund der Invertierbarkeit von DD^T, dass $\left(DD^T\right)^{\frac{1}{2}}$ ebenfalls invertierbar ist und dass gilt:

$$\left((DD^T)^{-\frac{1}{2}}\right)^2 = (DD^T)^{-1}, \text{ da } \left((DD^T)^{-\frac{1}{2}}\right)^2 (DD^T) = A^{-1}A^{-1}(DD^T) = I_m$$

Schritt 3

Es ist daher möglich einen Prozeß $(R_t, t \geq 0)$ folgendermaßen zu definieren:

Schritt 3

Es ist daher möglich einen Prozeß $(R_t, t \geq 0)$ folgendermaßen zu definieren:

$$dR_t = (DD^T)^{-\frac{1}{2}} dN_t \Rightarrow R_t = (DD^T)^{-\frac{1}{2}} N_t$$

Schritt 3

Es ist daher möglich einen Prozeß $(R_t, t \geq 0)$ folgendermaßen zu definieren:

$$dR_t = (DD^T)^{-\frac{1}{2}} dN_t \Rightarrow R_t = (DD^T)^{-\frac{1}{2}} N_t$$

Lemma

Der Prozeß $(R_t, t \geq 0)$ ist eine m-dimensionale Brownsche Bewegung.

Schritt 3

Es ist daher möglich einen Prozeß $(R_t, t \geq 0)$ folgendermaßen zu definieren:

$$dR_t = (DD^T)^{-\frac{1}{2}} dN_t \Rightarrow R_t = (DD^T)^{-\frac{1}{2}} N_t$$

Lemma

Der Prozeß $(R_t, t \geq 0)$ ist eine m-dimensionale Brownsche Bewegung.

Der Beweis folgt sofort aus der Darstellung

$$dR_t = (DD^T)^{-\frac{1}{2}} G \left(X_t - \widehat{X}_t \right) dt + (DD^T)^{-\frac{1}{2}} D dV_t$$

und dem folgendem Theorem:

Schritt 3

Es ist daher möglich einen Prozeß $(R_t, t \geq 0)$ folgendermaßen zu definieren:

$$dR_t = (DD^T)^{-\frac{1}{2}} dN_t \Rightarrow R_t = (DD^T)^{-\frac{1}{2}} N_t$$

Lemma

Der Prozeß $(R_t, t \geq 0)$ ist eine m-dimensionale Brownsche Bewegung.

Der Beweis folgt sofort aus der Darstellung

$$dR_t = (DD^T)^{-\frac{1}{2}} G \left(X_t - \widehat{X}_t \right) dt + (DD^T)^{-\frac{1}{2}} D dV_t$$

und dem folgendem Theorem:
Sei $dY_t = u(t, \omega)dt + v(t, \omega)dW_t$ ein Itô-Prozeß in $\mathbb{R}^n$. Dann sind die folgenden Aussagen äquivalent:

Schritt 3

Es ist daher möglich einen Prozeß $(R_t, t \geq 0)$ folgendermaßen zu definieren:

$$dR_t = (DD^T)^{-\frac{1}{2}} dN_t \Rightarrow R_t = (DD^T)^{-\frac{1}{2}} N_t$$

Lemma

Der Prozeß $(R_t, t \geq 0)$ ist eine m-dimensionale Brownsche Bewegung.

Der Beweis folgt sofort aus der Darstellung

$$dR_t = (DD^T)^{-\frac{1}{2}} G \left(X_t - \widehat{X}_t \right) dt + (DD^T)^{-\frac{1}{2}} D dV_t$$

und dem folgendem Theorem:
Sei $dY_t = u(t, \omega)dt + v(t, \omega)dW_t$ ein Itô-Prozeß in $\mathbb{R}^n$. Dann sind die folgenden Aussagen äquivalent:

- $(Y_t, t \geq 0)$ ist eine n-dimensionale Brownsche Bewegung

Schritt 3

Es ist daher möglich einen Prozeß $(R_t, t \geq 0)$ folgendermaßen zu definieren:

$$dR_t = (DD^T)^{-\frac{1}{2}} dN_t \Rightarrow R_t = (DD^T)^{-\frac{1}{2}} N_t$$

Lemma

Der Prozeß $(R_t, t \geq 0)$ ist eine m-dimensionale Brownsche Bewegung.

Der Beweis folgt sofort aus der Darstellung

$$dR_t = (DD^T)^{-\frac{1}{2}} G \left(X_t - \widehat{X}_t \right) dt + (DD^T)^{-\frac{1}{2}} D dV_t$$

und dem folgendem Theorem:
Sei $dY_t = u(t, \omega) dt + v(t, \omega) dW_t$ ein Itô-Prozeß in $\mathbb{R}^n$. Dann sind die folgenden Aussagen äquivalent:

- $(Y_t, t \geq 0)$ ist eine n-dimensionale Brownsche Bewegung
- $\mathbb{E}\left[u(t, \cdot) | \mathcal{N}_t\right] = 0$ sowie $vv^T(t, \omega) = I_n$ für f.a. (t, ω) wobei $\mathcal{N}_t$ die von Y erzeugte σ-Algebra sei.

Schritt 3

Es ist offensichtlich, dass gilt:

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

Schritt 3

Es ist offensichtlich, dass gilt:

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

und somit wissen wir, dass der gesuchte Prozeß zum Zeitpunkt t die Projektion von X_t auf den Raum $\mathcal{L}(R, t)$ ist, d.h.

$$\widehat{X}_t = \mathcal{P}_{\mathcal{L}(R,t)}(X_t).$$

Schritt 3

Es ist offensichtlich, dass gilt:

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

und somit wissen wir, dass der gesuchte Prozeß zum Zeitpunkt t die Projektion von X_t auf den Raum $\mathcal{L}(R,t)$ ist, d.h.

$$\widehat{X}_t = \mathcal{P}_{\mathcal{L}(R,t)}(X_t).$$

Es lässt sich sogar zeigen, dass $\widehat{X}_t$ dann die folgende Gestalt haben muss:

Es ist offensichtlich, dass gilt:

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

und somit wissen wir, dass der gesuchte Prozeß zum Zeitpunkt t die Projektion von X_t auf den Raum $\mathcal{L}(R,t)$ ist, d.h.

$$\widehat{X}_t = \mathcal{P}_{\mathcal{L}(R,t)}(X_t).$$

Es lässt sich sogar zeigen, dass $\widehat{X}_t$ dann die folgende Gestalt haben muss:

Lemma

$$\widehat{X}_t = \mathbb{E}\, X_t + \int_0^t \frac{\partial}{\partial s}\, \mathbb{E}[X_t R_s] dR_s. \tag{3}$$

Schritt 4

Nach der Itô-Formel gilt für $r \leq t$:

$$\begin{aligned} & d\left(\exp\left(-(t-r)F\right)X_t\right) \\ = \quad & \exp\left(-(t-r)F\right)dX_t - \exp\left(-(t-r)F\right)FX_t dt. \end{aligned} \tag{4}$$

Schritt 4

Nach der Itô-Formel gilt für $r \leq t$:

$$\begin{aligned} & d\left(\exp\left(-(t-r)F\right)X_t\right) \\ = \quad & \exp\left(-(t-r)F\right)dX_t - \exp\left(-(t-r)F\right)FX_t dt. \end{aligned}$$

Damit ist es möglich eine explizite Lösung der SDE zu berechnen, man erhält:

$$dX_t \quad = \quad FX_t dt + CdU_t \tag{4}$$

Schritt 4

Nach der Itô-Formel gilt für $r \leq t$:

$$\begin{aligned} & d\left(\exp\left(-(t-r)F\right)X_t\right) \\ = \quad & \exp\left(-(t-r)F\right)dX_t - \exp\left(-(t-r)F\right)FX_t dt. \end{aligned}$$

Damit ist es möglich eine explizite Lösung der SDE zu berechnen, man erhält:

$$\begin{aligned} dX_t &= FX_t dt + C dU_t \\ \exp\left(-(t-r)F\right)\left(dX_t - FX_t dt\right) &= \exp\left(-(t-r)F\right)FX_t dU_t \end{aligned} \tag{4}$$

Schritt 4

Nach der Itô-Formel gilt für $r \leq t$:

$$\begin{aligned} & d\left(\exp\left(-(t-r)F\right)X_t\right) \\ = \quad & \exp\left(-(t-r)F\right)dX_t - \exp\left(-(t-r)F\right)FX_t dt. \end{aligned}$$

Damit ist es möglich eine explizite Lösung der SDE zu berechnen, man erhält:

$$\begin{aligned} dX_t &= FX_t dt + CdU_t \\ \exp\left(-(t-r)F\right)\left(dX_t - FX_t dt\right) &= \exp\left(-(t-r)F\right)FX_t dU_t \\ d\left(\exp\left(-(t-r)F\right)X_t\right) &= \exp\left(-(t-r)F\right)FX_t dU_t \end{aligned} \tag{4}$$

Schritt 4

Nach der Itô-Formel gilt für $r \leq t$:

$$\begin{aligned} & d\left(\exp\left(-(t-r)F\right)X_t\right) \\ = \quad & \exp\left(-(t-r)F\right)dX_t - \exp\left(-(t-r)F\right)FX_t dt. \end{aligned}$$

Damit ist es möglich eine explizite Lösung der SDE zu berechnen, man erhält:

$$\begin{aligned} dX_t &= FX_t dt + CdU_t \\ \exp\left(-(t-r)F\right)\left(dX_t - FX_t dt\right) &= \exp\left(-(t-r)F\right)FX_t dU_t \\ d\left(\exp\left(-(t-r)F\right)X_t\right) &= \exp\left(-(t-r)F\right)FX_t dU_t \\ \exp\left(-(t-r)F\right)X_t - X_r &= \int_r^t \exp\left(-(s-r)F\right)CdU_s \end{aligned} \tag{4}$$

Schritt 4

Nach der Itô-Formel gilt für $r \leq t$:

$$
\begin{aligned}
& d\left(\exp\left(-(t-r)F\right)X_t\right) \\
= \quad & \exp\left(-(t-r)F\right)dX_t - \exp\left(-(t-r)F\right)FX_t dt.
\end{aligned}
$$

Damit ist es möglich eine explizite Lösung der SDE zu berechnen, man erhält:

$$
\begin{aligned}
dX_t &= FX_t dt + CdU_t \\
\exp\left(-(t-r)F\right)\left(dX_t - FX_t dt\right) &= \exp\left(-(t-r)F\right)FX_t dU_t \\
d\left(\exp\left(-(t-r)F\right)X_t\right) &= \exp\left(-(t-r)F\right)FX_t dU_t \\
\exp\left(-(t-r)F\right)X_t - X_r &= \int_r^t \exp\left(-(s-r)F\right)CdU_s \\
X_t &= \exp\left(t-r\right)FX_r + \int_r^t \exp\left(t-s\right)F\,CdU_s. \qquad (4)
\end{aligned}
$$

Schritt 4

Nach der Itô-Formel gilt für $r \leq t$:

$$
\begin{aligned}
& d\left(\exp\left(-(t-r)F\right)X_t\right) \\
= \quad & \exp\left(-(t-r)F\right)dX_t - \exp\left(-(t-r)F\right)FX_t dt.
\end{aligned}
$$

Damit ist es möglich eine explizite Lösung der SDE zu berechnen, man erhält:

$$
\begin{aligned}
dX_t &= FX_t dt + CdU_t \\
\exp\left(-(t-r)F\right)\left(dX_t - FX_t dt\right) &= \exp\left(-(t-r)F\right)FX_t dU_t \\
d\left(\exp\left(-(t-r)F\right)X_t\right) &= \exp\left(-(t-r)F\right)FX_t dU_t \\
\exp\left(-(t-r)F\right)X_t - X_r &= \int_r^t \exp\left(-(s-r)F\right)CdU_s \\
X_t &= \exp\left(t-r\right)FX_r + \int_r^t \exp\left(t-s\right)F\,CdU_s
\end{aligned}
\tag{4}
$$

Im Besonderem gilt für den Erwartungswert von X_t:

$$\mathbb{E}\,X_t = \exp\left(tF\right)\mathbb{E}\,X_0. \tag{5}$$

Schritt 5

Wir wissen bereits nach Schritt 4, dass unser Schätzer $\widehat{X}_t$ die folgende Form hat:

$$\widehat{X}_t = \mathbb{E}\, X_t + \int_0^t \underbrace{\frac{\partial}{\partial s}\, \mathbb{E}[X_t R_s^T]}_{=: f(s,t)} \, dR_s.$$

Schritt 5

Wir wissen bereits nach Schritt 4, dass unser Schätzer $\widehat{X}_t$ die folgende Form hat:

$$\widehat{X}_t = \mathbb{E}\, X_t + \int_0^t \underbrace{\frac{\partial}{\partial s}\, \mathbb{E}[X_t R_s^T]}_{=:f(s,t)}\, dR_s.$$

Desweiteren lässt sich der Prozeß R zur Zeit t darstellen als

$$R_t = (DD^T)^{-\frac{1}{2}} G \int_0^t (X_s - \widehat{X}_s) ds + (DD^T)^{-\frac{1}{2}} DV_t$$

Schritt 5

Wir wissen bereits nach Schritt 4, dass unser Schätzer $\widehat{X}_t$ die folgende Form hat:

$$\widehat{X}_t = \mathbb{E}\, X_t + \int_0^t \underbrace{\frac{\partial}{\partial s}\, \mathbb{E}[X_t R_s^T]}_{=:f(s,t)}\, dR_s.$$

Desweiteren lässt sich der Prozeß R zur Zeit t darstellen als

$$R_t = (DD^T)^{-\frac{1}{2}} G \int_0^t (X_s - \widehat{X}_s) ds + (DD^T)^{-\frac{1}{2}} DV_t$$

und man erhält somit

$$\mathbb{E}[X_t R_s^T] = \int_0^s \mathbb{E}[X_t (\underbrace{X_r - \widehat{X}_r}_{=:\widetilde{X}_r})^T] dr G^T (DD^T)^{-\frac{1}{2}}.$$

Wir wissen bereits nach Schritt 4, dass unser Schätzer $\widehat{X}_t$ die folgende Form hat:

$$\widehat{X}_t = \mathbb{E}\, X_t + \int_0^t \underbrace{\frac{\partial}{\partial s}\, \mathbb{E}[X_t R_s^T]}_{=:f(s,t)} \, dR_s.$$

Desweiteren lässt sich der Prozeß R zur Zeit t darstellen als

$$R_t = (DD^T)^{-\frac{1}{2}} G \int_0^t (X_s - \widehat{X}_s) ds + (DD^T)^{-\frac{1}{2}} DV_t$$

und man erhält somit

$$\mathbb{E}[X_t R_s^T] = \int_0^s \mathbb{E}[X_t (\underbrace{X_r - \widehat{X}_r}_{=:\widetilde{X}_r})^T] dr G^T (DD^T)^{-\frac{1}{2}}.$$

Unter Benutzung der Darstellung 4 und der Unabhängigkeit von $\widetilde{X}_r$ und der Inkremente des Prozeßes $(U_s, r \leq s \leq t)$ sieht man, dass gilt

$$\mathbb{E}[X_t \widetilde{X}_r]^T \quad = \quad \exp\left((t-r)F\right) \mathbb{E}[X_r \widetilde{X}_r]$$

Schritt 5

Wir wissen bereits nach Schritt 4, dass unser Schätzer $\widehat{X}_t$ die folgende Form hat:

$$\widehat{X}_t = \mathbb{E}\, X_t + \int_0^t \underbrace{\frac{\partial}{\partial s}\,\mathbb{E}[X_t R_s^T]}_{=:f(s,t)}\, dR_s.$$

Desweiteren lässt sich der Prozeß R zur Zeit t darstellen als

$$R_t = (DD^T)^{-\frac{1}{2}} G \int_0^t (X_s - \widehat{X}_s) ds + (DD^T)^{-\frac{1}{2}} DV_t$$

und man erhält somit

$$\mathbb{E}[X_t R_s^T] = \int_0^s \mathbb{E}[X_t(\underbrace{X_r - \widehat{X}_r}_{=:\widetilde{X}_r})^T] dr G^T (DD^T)^{-\frac{1}{2}}.$$

Unter Benutzung der Darstellung 4 und der Unabhängigkeit von $\widetilde{X}_r$ und der Inkremente des Prozeßes $(U_s, r \leq s \leq t)$ sieht man, dass gilt

$$\begin{aligned} \mathbb{E}[X_t \widetilde{X}_r]^T &= \exp\left((t-r)F\right) \mathbb{E}[X_r \widetilde{X}_r] \\ &= \exp\left((t-r)F\right) \mathbb{E}[\widetilde{X}_r \widetilde{X}_r]. \end{aligned}$$

Schritt 5

Man erhält also zusammenfassend

$$\mathbb{E}[X_t R_s^T] = \int_0^s \exp\left((t-r)F\right) \mathbb{E}[\widetilde{X}_r \widetilde{X}_r] dr G^T (DD^T)^{-\frac{1}{2}}. \tag{6}$$

Schritt 5

Man erhält also zusammenfassend

$$\mathbb{E}[X_t R_s^T] = \int_0^s \exp\left((t-r)F\right) \mathbb{E}[\widetilde{X}_r \widetilde{X}_r] dr G^T (DD^T)^{-\frac{1}{2}}.$$

sowie für die Funktion f die folgende Darstellung:

$$f(s,t) = \exp\left((t-s)F\right) \mathbb{E}[\widetilde{X}_s \widetilde{X}_s] G^T (DD^T)^{-\frac{1}{2}}.$$

(6)

Schritt 5

Man erhält also zusammenfassend

$$\mathbb{E}[X_t R_s^T] = \int_0^s \exp\left((t-r)F\right) \mathbb{E}[\widetilde{X}_r \widetilde{X}_r] dr G^T (DD^T)^{-\frac{1}{2}}.$$

sowie für die Funktion f die folgende Darstellung:

$$f(s,t) = \exp\left((t-s)F\right) \mathbb{E}[\widetilde{X}_s \widetilde{X}_s] G^T (DD^T)^{-\frac{1}{2}}.$$

Lemma

Der matrixwertige Funktion $S(t) := \mathbb{E}[\widetilde{X}_t \widetilde{X}_t^T] = \mathbb{E}[(X_t - \widehat{X}_t)(X_t - \widehat{X}_t)^T]$ *erfüllt die folgende Matrix-Riccati-Differentialgleichung:*

(6)

Schritt 5

Man erhält also zusammenfassend

$$\mathbb{E}[X_t R_s^T] = \int_0^s \exp\left((t-r)F\right) \mathbb{E}[\widetilde{X}_r \widetilde{X}_r] dr G^T (DD^T)^{-\frac{1}{2}}.$$

sowie für die Funktion f die folgende Darstellung:

$$f(s,t) = \exp\left((t-s)F\right) \mathbb{E}[\widetilde{X}_s \widetilde{X}_s] G^T (DD^T)^{-\frac{1}{2}}.$$

Lemma

Der matrixwertige Funktion $S(t) := \mathbb{E}[\widetilde{X}_t \widetilde{X}_t^T] = \mathbb{E}[(X_t - \widehat{X}_t)(X_t - \widehat{X}_t)^T]$ *erfüllt die folgende Matrix-Riccati-Differentialgleichung:*

$$\frac{dS}{dt} = FS + SF^T - SG^T(DD^T)^{-1}GS + CC^T \tag{6}$$

Schritt 5

Man erhält also zusammenfassend

$$\mathbb{E}[X_t R_s^T] = \int_0^s \exp\left((t-r)F\right) \mathbb{E}[\widetilde{X}_r \widetilde{X}_r] dr G^T (DD^T)^{-\frac{1}{2}}.$$

sowie für die Funktion f die folgende Darstellung:

$$f(s,t) = \exp\left((t-s)F\right) \mathbb{E}[\widetilde{X}_s \widetilde{X}_s] G^T (DD^T)^{-\frac{1}{2}}.$$

Lemma

Der matrixwertige Funktion $S(t) := \mathbb{E}[\widetilde{X}_t \widetilde{X}_t^T] = \mathbb{E}[(X_t - \widehat{X}_t)(X_t - \widehat{X}_t)^T]$ *erfüllt die folgende Matrix-Riccati-Differentialgleichung:*

$$\frac{dS}{dt} = FS + SF^T - SG^T(DD^T)^{-1}GS + CC^T \tag{6}$$

mit der Anfangsbedingung: $S(0) = \mathbb{E}\left[(X_0 - \mathbb{E}\,X_0)(X_0 - \mathbb{E}\,X_0)^T\right].$

Schritt 5

Wir kennen bereits die folgende Darstellung von $\widehat{X}_t$:

$$\widehat{X}_t = \mathbb{E}\,X_t + \int_0^t f(s,t)dR_s. \tag{7}$$

Schritt 5

Wir kennen bereits die folgende Darstellung von $\widehat{X}_t$:

$$\widehat{X}_t = \mathbb{E}\,X_t + \int_0^t f(s,t)dR_s. \tag{7}$$

Durch Ausnutzen der Definitionen von R_t bzw. N_t erhalten wir eine Charakterisierung von $\widehat{X}_t$ mittels einer SDE:

Schritt 5

Wir kennen bereits die folgende Darstellung von $\widehat{X}_t$:

$$\widehat{X}_t = \mathbb{E}\, X_t + \int_0^t f(s,t)dR_s. \tag{7}$$

Durch Ausnutzen der Definitionen von R_t bzw. N_t erhalten wir eine Charakterisierung von $\widehat{X}_t$ mittels einer SDE:

$$d\widehat{X}_t \quad = \quad F\,\mathbb{E}\, X_t dt + f(t,t)dR_t + \left(\int_0^t \frac{\partial}{\partial t} f(s,t)dR_s \right) dt$$

Schritt 5

Wir kennen bereits die folgende Darstellung von $\widehat{X}_t$:

$$\widehat{X}_t = \mathbb{E}\, X_t + \int_0^t f(s,t) dR_s. \tag{7}$$

Durch Ausnutzen der Definitionen von R_t bzw. N_t erhalten wir eine Charakterisierung von $\widehat{X}_t$ mittels einer SDE:

$$\begin{aligned} d\widehat{X}_t &= F\,\mathbb{E}\, X_t dt + f(t,t) dR_t + \left(\int_0^t \frac{\partial}{\partial t} f(s,t) dR_s \right) dt \\ &= F\,\mathbb{E}\, X_t dt + SG^T (DD^T)^{-\frac{1}{2}} dR_t + F \left(\int_0^t f(s,t) dR_s \right) dt \end{aligned}$$

Schritt 5

Wir kennen bereits die folgende Darstellung von $\widehat{X}_t$:

$$\widehat{X}_t = \mathbb{E}\, X_t + \int_0^t f(s,t) dR_s. \tag{7}$$

Durch Ausnutzen der Definitionen von R_t bzw. N_t erhalten wir eine Charakterisierung von $\widehat{X}_t$ mittels einer SDE:

$$\begin{aligned} d\widehat{X}_t &= F\,\mathbb{E}\, X_t dt + f(t,t) dR_t + \left(\int_0^t \frac{\partial}{\partial t} f(s,t) dR_s \right) dt \\ &= F\,\mathbb{E}\, X_t dt + SG^T (DD^T)^{-\frac{1}{2}} dR_t + F \left(\int_0^t f(s,t) dR_s \right) dt \\ &\overset{7}{=} F\widehat{X}_t dt + SG^T (DD^T)^{-1} dN_t \end{aligned}$$

Schritt 5

Wir kennen bereits die folgende Darstellung von $\widehat{X}_t$:

$$\widehat{X}_t = \mathbb{E}\, X_t + \int_0^t f(s,t) dR_s. \tag{7}$$

Durch Ausnutzen der Definitionen von R_t bzw. N_t erhalten wir eine Charakterisierung von $\widehat{X}_t$ mittels einer SDE:

$$\begin{aligned}
d\widehat{X}_t &= F\,\mathbb{E}\, X_t dt + f(t,t) dR_t + \left(\int_0^t \frac{\partial}{\partial t} f(s,t) dR_s \right) dt \\
&= F\,\mathbb{E}\, X_t dt + SG^T (DD^T)^{-\frac{1}{2}} dR_t + F \left(\int_0^t f(s,t) dR_s \right) dt \\
&\overset{7}{=} F\widehat{X}_t dt + SG^T (DD^T)^{-1} dN_t \\
&= F\widehat{X}_t dt + SG^T (DD^T)^{-1} (dZ_t - G\widehat{X}_t dt)
\end{aligned}$$

Schritt 5

Wir kennen bereits die folgende Darstellung von $\widehat{X}_t$:

$$\widehat{X}_t = \mathbb{E}\, X_t + \int_0^t f(s,t)\,dR_s. \tag{7}$$

Durch Ausnutzen der Definitionen von R_t bzw. N_t erhalten wir eine Charakterisierung von $\widehat{X}_t$ mittels einer SDE:

$$\begin{aligned}
d\widehat{X}_t &= F\,\mathbb{E}\,X_t dt + f(t,t)dR_t + \left(\int_0^t \frac{\partial}{\partial t} f(s,t)dR_s\right) dt \\
&= F\,\mathbb{E}\,X_t dt + SG^T(DD^T)^{-\frac{1}{2}} dR_t + F\left(\int_0^t f(s,t)dR_s\right) dt \\
&\stackrel{7}{=} F\widehat{X}_t dt + SG^T(DD^T)^{-1} dN_t \\
&= F\widehat{X}_t dt + SG^T(DD^T)^{-1}(dZ_t - G\widehat{X}_t dt) \\
&= \left(F - SG^T(DD^T)^{-1}G\right)\widehat{X}_t dt + SG^T(DD^T)^{-1} dZ_t.
\end{aligned}$$

Auch für nicht konstante Matrixfunktionen erhält man die selben Aussagen:

Mehrdimensionaler Kalman-Bucy Filter

Auch für nicht konstante Matrixfunktionen erhält man die selben Aussagen:

Theorem

Der beste Schätzer $\widehat{X}_t$ *eines mehrdimensionalen Filterproblems, das durch die beiden SDE's*

Mehrdimensionaler Kalman-Bucy Filter

Auch für nicht konstante Matrixfunktionen erhält man die selben Aussagen:

Theorem

Der beste Schätzer $\widehat{X}_t$ eines mehrdimensionalen Filterproblems, das durch die beiden SDE's

$$\begin{aligned} dX_t &= F(t)X_t dt + C(t)dU_t \\ dZ_t &= G(t)X_t dt + D(t)dV_t \end{aligned}$$

Mehrdimensionaler Kalman-Bucy Filter

Auch für nicht konstante Matrixfunktionen erhält man die selben Aussagen:

Theorem

Der beste Schätzer $\widehat{X}_t$ eines mehrdimensionalen Filterproblems, das durch die beiden SDE's

$$\begin{aligned} dX_t &= F(t)X_t dt + C(t)dU_t \\ dZ_t &= G(t)X_t dt + D(t)dV_t \end{aligned}$$

beschrieben wird, ist eine Lösung der folgenden SDE:

Mehrdimensionaler Kalman-Bucy Filter

Auch für nicht konstante Matrixfunktionen erhält man die selben Aussagen:

Theorem

Der beste Schätzer $\widehat{X}_t$ eines mehrdimensionalen Filterproblems, das durch die beiden SDE's

$$\begin{aligned} dX_t &= F(t)X_t dt + C(t)dU_t \\ dZ_t &= G(t)X_t dt + D(t)dV_t \end{aligned}$$

beschrieben wird, ist eine Lösung der folgenden SDE:

$$d\widehat{X}_t = \left(F - SG^T(DD^T)^{-1}G\right)\widehat{X}_t dt + SG^T(DD^T)^{-1}dZ_t \tag{8}$$

Mehrdimensionaler Kalman-Bucy Filter

Auch für nicht konstante Matrixfunktionen erhält man die selben Aussagen:

Theorem

Der beste Schätzer $\widehat{X}_t$ eines mehrdimensionalen Filterproblems, das durch die beiden SDE's

$$\begin{aligned} dX_t &= F(t)X_t dt + C(t)dU_t \\ dZ_t &= G(t)X_t dt + D(t)dV_t \end{aligned}$$

beschrieben wird, ist eine Lösung der folgenden SDE:

$$d\widehat{X}_t = \left(F - SG^T(DD^T)^{-1}G\right)\widehat{X}_t dt + SG^T(DD^T)^{-1}dZ_t \tag{8}$$

mit dem Startwert $\widehat{X}_0 = \mathbb{E}\, X_0$.

Mehrdimensionaler Kalman-Bucy Filter

Auch für nicht konstante Matrixfunktionen erhält man die selben Aussagen:

Theorem

Der beste Schätzer $\widehat{X}_t$ eines mehrdimensionalen Filterproblems, das durch die beiden SDE's

$$\begin{aligned} dX_t &= F(t)X_t dt + C(t)dU_t \\ dZ_t &= G(t)X_t dt + D(t)dV_t \end{aligned}$$

beschrieben wird, ist eine Lösung der folgenden SDE:

$$d\widehat{X}_t = \left(F - SG^T(DD^T)^{-1}G\right)\widehat{X}_t dt + SG^T(DD^T)^{-1}dZ_t \tag{8}$$

mit dem Startwert $\widehat{X}_0 = \mathbb{E}\, X_0$.
Die in 8 auftauchende Matrixfunktion ist desweiteren eine Lösung der Matrix-Riccati-Differentialgleichung

Mehrdimensionaler Kalman-Bucy Filter

Auch für nicht konstante Matrixfunktionen erhält man die selben Aussagen:

Theorem

Der beste Schätzer $\widehat{X}_t$ eines mehrdimensionalen Filterproblems, das durch die beiden SDE's

$$\begin{aligned} dX_t &= F(t)X_t dt + C(t)dU_t \\ dZ_t &= G(t)X_t dt + D(t)dV_t \end{aligned}$$

beschrieben wird, ist eine Lösung der folgenden SDE:

$$d\widehat{X}_t = \left(F - SG^T(DD^T)^{-1}G\right)\widehat{X}_t dt + SG^T(DD^T)^{-1}dZ_t \tag{8}$$

mit dem Startwert $\widehat{X}_0 = \mathbb{E}\,X_0$.
Die in 8 auftauchende Matrixfunktion ist desweiteren eine Lösung der Matrix-Riccati-Differentialgleichung

$$\frac{dS}{dt} = FS + SF^T - SG^T(DD^T)^{-1}GS + CC^T$$

Mehrdimensionaler Kalman-Bucy Filter

Auch für nicht konstante Matrixfunktionen erhält man die selben Aussagen:

Theorem

Der beste Schätzer $\widehat{X}_t$ eines mehrdimensionalen Filterproblems, das durch die beiden SDE's

$$\begin{aligned} dX_t &= F(t)X_t dt + C(t)dU_t \\ dZ_t &= G(t)X_t dt + D(t)dV_t \end{aligned}$$

beschrieben wird, ist eine Lösung der folgenden SDE:

$$d\widehat{X}_t = \left(F - SG^T(DD^T)^{-1}G\right)\widehat{X}_t dt + SG^T(DD^T)^{-1}dZ_t \tag{8}$$

mit dem Startwert $\widehat{X}_0 = \mathbb{E}\, X_0$.
Die in 8 auftauchende Matrixfunktion ist desweiteren eine Lösung der Matrix-Riccati-Differentialgleichung

$$\frac{dS}{dt} = FS + SF^T - SG^T(DD^T)^{-1}GS + CC^T$$

mit der Anfangsbedingung: $S(0) = \mathbb{E}\left[(X_0 - \mathbb{E}\, X_0)(X_0 - \mathbb{E}\, X_0)^T\right]$.

Gestörtes Bevölkerungswachstumsmodell mit gestörten Beobachtungen

Wir betrachten nun das folgende Modell, welches den Bevölkerungszuwachs in einem Land darstellen soll. Wir nehmen an, dass die Bevölkerung aufgrund von beispielsweise Seuchen bzw. neuer Krebsmedikamente nicht konstant wachse. Desweiteren sei die beobachtete Bevölkerungszahl ebenfalls gestört. Wir haben also die folgenden SDE's:

Gestörtes Bevölkerungswachstumsmodell mit gestörten Beobachtungen

Wir betrachten nun das folgende Modell, welches den Bevölkerungszuwachs in einem Land darstellen soll. Wir nehmen an, dass die Bevölkerung aufgrund von beispielsweise Seuchen bzw. neuer Krebsmedikamente nicht konstant wachse. Desweiteren sei die beobachtete Bevölkerungszahl ebenfalls gestört. Wir haben also die folgenden SDE's:

$$\begin{aligned} dX_t &= rX_t dt + m_1 dU_t \qquad r, m_1 \equiv const \\ dZ_t &= X_t dt + m_2 dV_t \qquad m_2 \equiv const \end{aligned}$$

Gestörtes Bevölkerungswachstumsmodell mit gestörten Beobachtungen

Wir betrachten nun das folgende Modell, welches den Bevölkerungszuwachs in einem Land darstellen soll. Wir nehmen an, dass die Bevölkerung aufgrund von beispielsweise Seuchen bzw. neuer Krebsmedikamente nicht konstant wachse. Desweiteren sei die beobachtete Bevölkerungszahl ebenfalls gestört. Wir haben also die folgenden SDE's:

$$\begin{aligned} dX_t &= rX_t dt + m_1 dU_t \qquad r, m_1 \equiv const \\ dZ_t &= X_t dt + m_2 dV_t \qquad m_2 \equiv const \end{aligned}$$

Es ergibt sich für die Riccati-Differentialgleichung:

$$\frac{dS}{dt} = 2rS - \frac{S^2}{m_2^2} + m_1^2$$

Gestörtes Bevölkerungswachstumsmodell mit gestörten Beobachtungen

Wir betrachten nun das folgende Modell, welches den Bevölkerungszuwachs in einem Land darstellen soll. Wir nehmen an, dass die Bevölkerung aufgrund von beispielsweise Seuchen bzw. neuer Krebsmedikamente nicht konstant wachse. Desweiteren sei die beobachtete Bevölkerungszahl ebenfalls gestört. Wir haben also die folgenden SDE's:

$$\begin{aligned} dX_t &= rX_t dt + m_1 dU_t \qquad r, m_1 \equiv const \\ dZ_t &= X_t dt + m_2 dV_t \qquad m_2 \equiv const \end{aligned}$$

Es ergibt sich für die Riccati-Differentialgleichung:

$$\frac{dS}{dt} = 2rS - \frac{S^2}{m_2^2} + m_1^2$$

Eine Lösung der obigen Differentialgleichung ist gegeben durch:

$$S(t) := 2rm_2^2 \left(1 + K \exp\left(-2rt\right)\right)^{-1} + m_1^2 t \text{ mit } K := \frac{2rm_2^2}{\mathbb{V}X_0} - 1,$$

Wir betrachten nun das folgende Modell, welches den Bevölkerungszuwachs in einem Land darstellen soll. Wir nehmen an, dass die Bevölkerung aufgrund von beispielsweise Seuchen bzw. neuer Krebsmedikamente nicht konstant wachse. Desweiteren sei die beobachtete Bevölkerungszahl ebenfalls gestört. Wir haben also die folgenden SDE's:

$$\begin{aligned} dX_t &= rX_t dt + m_1 dU_t \qquad r, m_1 \equiv const \\ dZ_t &= X_t dt + m_2 dV_t \qquad m_2 \equiv const \end{aligned}$$

Es ergibt sich für die Riccati-Differentialgleichung:

$$\frac{dS}{dt} = 2rS - \frac{S^2}{m_2^2} + m_1^2$$

Eine Lösung der obigen Differentialgleichung ist gegeben durch:

$$S(t) := 2rm_2^2 \left(1 + K \exp\left(-2rt\right)\right)^{-1} + m_1^2 t \text{ mit } K := \frac{2rm_2^2}{\mathbb{V}X_0} - 1,$$

d.h. der gesuchte Schätzer muss die SDE

$$d\widehat{X}_t = \left(r - \frac{S}{m_2^2}\right)\widehat{X}_t dt + \frac{S}{m_2^2} dZ_t$$

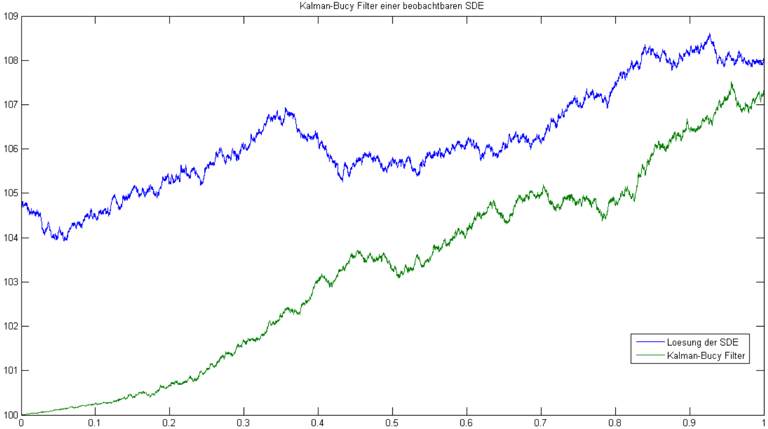
erfüllen.

Numerische Lösung des Filterproblems

Wenn wir versuchen das vorige Beispiel numerisch zu lösen, so erhält man für die Parameter $r := \ln 1.02$, $m_1 := m_2 := 1$, $\mathbb{E}\, X_0 = 100$ sowie $\mathbb{V} X_0 = 5$ beispielsweise die folgenden Plots:

Numerische Lösung des Filterproblems

Wenn wir versuchen das vorige Beispiel numerisch zu lösen, so erhält man für die Parameter $r := \ln 1.02$, $m_1 := m_2 := 1$, $\mathbb{E}\, X_0 = 100$ sowie $\mathbb{V} X_0 = 5$ beispielsweise die folgenden Plots:



sowie über einen längeren Zeithorizont