Kalman-Bucy-Filter

Vortrag im Rahmen des Seminars Markovprozesse
(Bernhardt)

Vorbemerkung

Vorbemerkung

Folgende Dinge werden nachfolgend vorausgesetzt:

- grundlegende Begriffe und Sätze der Stochastik,
- alle nachfolgenden $\mathcal{ZV}$ leben auf dem Wahrscheinlichkeitsraum $(\Omega, F, \mathbb{P})$,
- für vektorwertige $\mathcal{ZV}$ sind Erwartungswerte komponentenweise definiert,
- seien X, Y $\mathbb{R}^d$-, $\mathbb{R}^k$-$\mathcal{ZV}$:

  $$Cov(X, Y) = (Cov(X_i, Y_j))_{ij} = E[(X - E[X])(Y - E[Y])^T],$$

  die letzte Gleichheit impliziert Linearitätseigenschaften, die ohne Bemerkung benutzt werden,

Vorbemerkung

Folgende Dinge werden nachfolgend vorausgesetzt:

- sei X $\mathbb{R}^d$-$\mathcal{ZV}$: $Var(X) = Cov(X, X)$, ist symmetrisch und positiv semidefinit,
- Sätze über Matrizen:

  symmetrisch positiv semidefinit $\Longrightarrow$ [positiv definit $\Longleftrightarrow$ invertierbar],

  A positiv semidefinit und B positiv definit $\Longrightarrow$ $A + B$ positiv definit.

Motivation

Motivation

Bewegung eines Objektes im Raum:

$$X'(t) = f(X(t), t) \quad \text{mit} \quad X(0) = X_0$$

Diskretisierung der Zeit und Approximation der Integralversion:

$$\begin{aligned} X_{n+1} &= X_n + (t_{n+1} - t_n) f(X_n, t_n) + a_{n+1} \\ X_{n+1} &= A_{n+1} X_n + a_{n+1}, \end{aligned}$$

wobei a_{n+1} „Abweichfaktor“ (Diskretisierung, unbekannte Einflussgrößen), wird als zufällig angesehen - zentriert normal.

Motivation

Objekt wird mittels Messgeräten beobachtet:

$$Y_{n+1} = B_{n+1} X_{n+1} + b_{n+1},$$

wobei B_{n+1} häufig die gemessenen Komponenten von X_{n+1} „herausfiltert" (Bsp. X_{n+1} Position und Geschwindigkeit, Y_{n+1} nur Position)

und b_{n+1} den Messfehlern entspricht, wird als zufällig angesehen - zentriert normal.

Motivation

Haben Messungen $(y_i)_{i=1}^{n+1}$ als Realisierung der (Y_i) ,

wollen „gute" Schätzung von x_{n+1} Realisierung von X_{n+1} ,

im Allgemeinen existieren verschiedene mögliche Realisierungen von X_{n+1} gegeben (y_i) , dies entspricht einer Verteilung bzw. regulär bedingten Verteilung von X_{n+1} gegeben $(Y_i)_{i=1}^{n+1}$.

Motivation

Definition 1

$(X_n), (a_n)$ $\mathbb{R}^d$-$\mathcal{ZV}$ und $(Y_n), (b_n)$ $\mathbb{R}^k$-$\mathcal{ZV}$; (A_n) aus $\mathbb{R}^d \times \mathbb{R}^d$ und (B_n) aus $\mathbb{R}^k \times \mathbb{R}^d$; mit:

$$\begin{aligned} X_{n+1} &= A_{n+1}X_n + a_{n+1} \quad \text{für} \quad n \in \mathbb{N}_0 \\ Y_{n+1} &= B_{n+1}X_{n+1} + b_{n+1} \quad \text{für} \quad n \in \mathbb{N}_0, \end{aligned}$$

wobei $X_0, (a_n), (b_n)$ unabhängig; $X_0 \sim \mathcal{N}$ und $a_n \sim \mathcal{N}(0, Var(a))$ und $b_n \sim \mathcal{N}(0, Var(b))$; $Var(b)$ invertierbar.

Obige Definitionen werden zusammengefasst in $\mathcal{F}ilter(A, B, X, Y, a, b)$.

Motivation

Definition 1

Das Berechnen von $\mathbb{P}[X_n|(Y_i)_{i=1}^n](\omega)$ für $\mathbb{P}$-fa $\omega \in \Omega$ heißt Filterproblem.

$E[X_n|(Y_i)_{i=1}^n](\omega)$ heißt Schätzung und $Var[X_n|(Y_i)_{i=1}^n](\omega)$ heißt Güte der Schätzung für $\omega \in \Omega$ (falls existent, wird sich auch als dieses herausstellen).

Bem: (X_n bzw. Y_n Linearkomb. aus $X_0, (a_i)_{i=1}^n$ bzw. b_n)

- (X_n) normalverteilt, und $(X_i)_{i=1}^n$ unabhängig von $(a_i)_{i>n}$, (b_i),
- (Y_n) normalverteilt, und $(Y_i)_{i=1}^n$ unabhängig von $(a_i)_{i>n}$, $(b_i)_{i>n}$,

Markov Eigenschaft

Markov Eigenschaft

Definition 2

X_n aus $(S_n, \mathcal{S}_n)$-$\mathcal{ZV}$ $\forall n \in \mathbb{N}_0$; mit Übergangskernen $K_n^m = \mathbb{P}[X_n | (X_i)_{i=m}^{n-1})]$ für $m \in \{0, n-1\}$ und $n \in \mathbb{N}_0$;

(X_n) heißt Markovkette, falls: $K_0^n = K_{n-1}^n$ $\mathbb{P}$-fs.

Werden zeigen, dass (X_n) aus $\mathcal{F}ilter(A, B, X, Y, a, b)$ Markovkette ist.

Markov Eigenschaft

Lemma 1

Z_i $(S_i, \mathcal{S}_i)$-$\mathcal{ZV}$ für $i \in \{0, 1\}$ unabhängig; $f \geq 0$ $\mathcal{S}_1 \otimes \mathcal{S}_0$-messbar; dann gilt für $\mathbb{P}$-fa $\omega \in \Omega$:

$$E[f(Z_1, Z_0)|Z_0](\omega) = E[f(Z_1, Z_0(\omega))].$$

Beweis

Behauptung wird durch Eigenschaftsprüfung bewiesen, $g : S_0 \to \mathbb{R}$ via $g(s_0) := \int_{S_1} f(Z_1, s_0)\, d\mathbb{P}$,

- nach Messbarkeitsvoraussetzungen und Fubini ist g $\mathcal{S}_0$-messbar, also $g(Z_0)$ $\sigma(Z_0)$-messbar,

Markov Eigenschaft

Lemma 1 Z_0, Z_1 unabhängig; $f \geq 0$:

$$E[f(Z_1, Z_0)|Z_0](\omega) = E[f(Z_1, Z_0(\omega))].$$

Beweis

- sei $h \geq 0$ $\mathcal{S}_0$-messbar:

  $$E[h(Z_0) \cdot f(Z_1, Z_0)] \overset{\text{u.a.}}{\underset{\text{Fub}}{=}} \int_{S_0} \int_{S_1} h(s_0) \cdot f(s_1, s_0)\, d\mathbb{P}_{Z_1}(s_1)\, d\mathbb{P}_{Z_0}(s_0) = \int_{S_0} h(s_0) \cdot g(s_0)\, d\mathbb{P}_{Z_0}(s_0) = E[h(Z_0) \cdot g(Z_0)],$$

und damit folgt die Behauptung.

Markov Eigenschaft

Satz

$\mathcal{F}ilter(A, B, X, Y, a, b) \implies (X_n)$ Markovkette.

Beweis

sei $n \in \mathbb{N}$, sei $m \in \mathbb{B}^d$, sei $\forall (x, a) \in (\mathbb{R}^d)^2$
$f(x, a) := \mathbb{1}_m(A_{n+1}x + a)$, dann gilt für $\mathbb{P}$-fa $\omega \in \Omega$:

$$K_0^n(m, \omega) = \mathbb{P}[X_n \in m | (X_i)_{i=0}^{n-1}](\omega) =$$
$$E[f(X_{n-1}, a_n) | (X_i)_{i=0}^{n-1}](\omega) \overset{\text{Lemma 1}}{=} E[f(X_{n-1}(\omega), a_n)] \overset{\text{Lemma 1}}{=}$$
$$E[f(X_{n-1}, a_n | X_{n-1})](\omega) = K_{n-1}^n(m, \omega),$$

und damit folgt die Behauptung.

Lösung des Filterproblems (I)

Lösung des Filterproblems (II)

Lösung des Filterproblems (II)

Satz (Kalman-Bucy-Filter)

$\mathcal{F}ilter(A, B, X, Y, a, b)$; dann gilt für $\mathbb{P}$-fa $\omega \in \Omega$:

$$\mathbb{P}[X_n|(Y_i)_{i=1}^n](\omega) = \mathcal{N}(\hat{X}_n(\omega), \hat{V}_n),$$

wobei $\hat{X}_n$ und $\hat{V}_n$ rekursiv gegeben sind durch:

$$\hat{X}_0 = E[X_0], \quad \hat{V}_0 = Var(X_0), \quad H_{n+1} = A_{n+1}\hat{V}_n A_{n+1}^T + Var(a),$$

schreibe statt A_{n+1} A, statt B_{n+1} B, statt H_{n+1} H,

$$\begin{aligned} \hat{X}_{n+1} &= A\hat{X}_n + HB^T \cdot (BHB^T + Var(b))^{-1} \cdot (Y_{n+1} - BA\hat{X}_n) \\ \hat{V}_{n+1} &= H - HB^T \cdot (BHB^T + Var(b))^{-1} \cdot BH. \end{aligned}$$

Lösung des Filterproblems (II)

Satz(Kalman-Bucy-Filter) $\mathcal{F}ilter(A, B, X, Y, a, b)$:

$\mathbb{P}[X_n|(Y_i)_{i=1}^n](\omega) = \mathcal{N}(\hat{X}_n(\omega), \hat{V}_n)$.

Beweis

$(X_n, (Y_i)_{i=1}^n)$ normalverteilt (Linearkombination von $X_0, (a_n), (b_n)$),

$Var((Y_i)_{i=1}^n) = Var((B_iX_i + b_i)) \overset{\text{u.a.}}{=} Var((B_iX_i)) + Var((b_i))$,

linker Term ist als Varianzmatrix positiv semidefinit,
da $Var(b_i) = Var(b)$ und $Var(b)$ als invertierbare Varianzmatrix positiv definit, ist der rechte Term positiv definit, also ist obige Matrix positiv definit, insbesondere invertierbar,

nach Lemma 2 gilt für $\mathbb{P}$-fa $\omega \in \Omega$:
$\mathbb{P}[X_n|(Y_i)_{i=1}^n](\omega) = \mathcal{N}(\hat{X}_n(\omega), \hat{V}_n)$.

Lösung des Filterproblems (II)

Satz(Kalman-Bucy-Filter) $\mathcal{F}ilter(A, B, X, Y, a, b)$:

Konstruktion u.a. $\mathcal{ZV}$ für Lemma 3 (Innovationsschritt).

Beweis

$I_{n+1} := Y_{n+1} - E[Y_{n+1}|(Y_i)_{i=1}^n]$
es gilt:

- $(X_{n+1}, (Y_i)_{i+1}^n, I_{n+1})$ normal und I_{n+1} zentriert, denn: $E[Y_{n+1}|(Y_i)]$ nach Lemma 2 Linearkombination der (Y_i),
- I_{n+1} u.a. von $(Y_i)_{i=1}^n$, denn:
  I_{n+1} orthogonal zu $L^2(\Omega, (Y_i), \mathbb{P})$,
  insbesondere $Cov(I_{n+1}, (Y_i)) = 0$,
  also wegen Normalität I_{n+1} unabhängig von $(Y_i)_{i=1}^n$,

Lösung des Filterproblems (II)

Satz(Kalman-Bucy-Filter) $\mathcal{F}ilter(A, B, X, Y, a, b)$:

Konstruktion u.a. $\mathcal{ZV}$ für Lemma 3 (Innovationsschritt).

Beweis

$I_{n+1} := Y_{n+1} - E[Y_{n+1}|(Y_i)_{i=1}^n]$

es gilt:

- $I_{n+1} = Y_{n+1} - BA\hat{X}_n = BA(X_n - \hat{X}_n) + Ba_{n+1} + b_{n+1}$, denn:

  $E[Y_{n+1}|(Y_i)_{i=1}^n] = E[BX_{n+1} + b_{n+1}|(Y_i)_{i=1}^n] \overset{\text{u.a./zent.}}{=} B \cdot E[X_{n+1}|(Y_i)_{i=1}^n] = B \cdot E[AX_n + a_{n+1}|(Y_i)_{i=1}^n] \underset{\text{Lemma 2}}{\overset{\text{u.a./zent.}}{=}} BA\hat{X}_n$,

Lösung des Filterproblems (II)

Satz(Kalman-Bucy-Filter) $\mathcal{F}ilter(A, B, X, Y, a, b)$:

Konstruktion u.a. $\mathcal{ZV}$ für Lemma 3 (Innovationsschritt).

Beweis

$I_{n+1} = BA(X_n - \hat{X}_n) + Ba_{n+1} + b_{n+1}$

es gilt:

- $Var(I_{n+1})$ invertierbar, denn:

  $$Var(I_{n+1}) \overset{\text{u.a.}}{=} Var(BA(X_n - \hat{X}_n)) + Var(Ba_{n+1}) + Var(b) = BA \cdot Var(X_n - \hat{X}_n) \cdot A^T B^T + B \cdot Var(a) \cdot B^T + Var(b) \overset{\text{Lemma 2}}{=} BA\hat{V}_n A^T B^T + B \cdot Var(a) \cdot B^T + Var(b) = B(A\hat{V}_n A^T + Var(b))B^T + Var(b),$$

  also $Var(I_{n+1})$ positiv definit, insbesondere invertierbar.

Lösung des Filterproblems (II)

Satz(Kalman-Bucy-Filter) $\mathcal{F}ilter(A, B, X, Y, a, b)$:

Gleichung für $\hat{X}_{n+1}$.

Beweis

Anwendung von Lemma 3 auf X_{n+1}, $(Y_i)_{i=1}^n$, I_{n+1} :

$$\hat{X}_{n+1} = E[X_{n+1}|(Y_i)_{i=1}^n] + Cov(X_{n+1}, I_{n+1}) \cdot Var^{-1}(I_{n+1}) \cdot (I_{n+1} - E[I_{n+1}]),$$

einziger noch zu berechnender Term $Cov(X_{n+1}, I_{n+1})$:

$$Cov(X_{n+1}, I_{n+1}) = Cov((AX_n + a_{n+1}), BA(X_n - \hat{X}_n) + Ba_{n+1} + b_{n+1}) \overset{\text{u.a.}}{=} Cov(AX_n, BA(X_n - \hat{X}_n)) + Cov(a_{n+1}, Ba_{n+1}) = A \cdot Cov(X_n, X_n - \hat{X}_n) \cdot A^T B^T + Var(a) \cdot B^T$$

Satz(Kalman-Bucy-Filter) $\mathcal{F}ilter(A, B, X, Y, a, b)$:

Gleichung für $\hat{X}_{n+1}$.

Beweis

$Cov(X_n, X_n - \hat{X}_n) \overset{\text{Lemma 2}}{=}$
$Cov(X_n, X_n - Cov(X_n, (Y_i)) \cdot Var^{-1}((Y_i)) \cdot (Y_i)) =$
$Var(X_n) - Cov(X_n, (Y_i)) \cdot Var^{-1}((Y_i)) \cdot Cov((Y_i), X_n) \overset{\text{Lemma 2}}{=} \hat{V}_n,$

also:

$$Cov(X_{n+1}, I_{n+1}) = (A\hat{V}_n A^T + Var(a))B^T,$$

nach Einsetzen aller Gleichungen in die Formel für $\hat{X}_{n+1}$ ergibt sich die Behauptung für $\hat{X}_{n+1}$.

Lösung des Filterproblems (II)

Satz(Kalman-Bucy-Filter) $\mathcal{F}ilter(A, B, X, Y, a, b)$:

Gleichung für $\hat{V}_{n+1}$.

Beweis

Anwendung von Lemma 3 auf X_{n+1}, $(Y_i)_{i=1}^n$, I_{n+1} :

$$\hat{V}_{n+1} = Var(X_{n+1} - E[X_{n+1}|(Y_i)_{i=1}^n]) - Cov(X_{n+1}, I_{n+1}) \cdot Var^{-1}(I_{n+1}) \cdot Cov(I_{n+1}, X_{n+1}),$$

einziger noch zu berechnender Term $Var(X_{n+1} - A\hat{X}_n)$:

$$Var(X_{n+1} - A\hat{X}_n) = Var(A(X_n - \hat{X}_n) + a_{n+1}) \overset{\text{u.a.}}{=} Var(A(X_n - \hat{X}_n)) + Var(a_{n+1}) = A\hat{V}_nA^T + Var(a),$$

nach Einsetzen aller Gleichungen in die Formel für $\hat{V}_{n+1}$ ergibt sich die Behauptung für $\hat{V}_{n+1}$.

Beispiel

Beispiel

Bewegung eines Objektes auf einer Geraden mit konstanter Geschwindigkeit und zufälliger Beschleunigung:

$$\begin{pmatrix} x_{t+1} \\ v_{t+1} \end{pmatrix} = \begin{pmatrix} x_t + v_t \cdot dt + dt^2/2 \cdot a \\ v_t + dt \cdot a \end{pmatrix}$$

$$X_{t+1} = \begin{pmatrix} 1 & dt \\ 0 & 1 \end{pmatrix} X_t + \begin{pmatrix} dt^2/2 \cdot a \\ dt \cdot a \end{pmatrix},$$

ist dt eine Konstante dann gilt:

$$Var(\begin{pmatrix} dt^2/2 \cdot a \\ dt \cdot a \end{pmatrix}) = \begin{pmatrix} dt^4/4 \cdot Var(a) & dt^3/2 \cdot Var(a) \\ dt^3/2 \cdot Var(a) & dt^2 \cdot Cov(a) \end{pmatrix},$$

Beispiel

Objekt kann nur bzgl. des Ortes mit Messfehlern beobachtet werden:

$$Y_{n+1} = \begin{pmatrix} 1 & 0 \end{pmatrix} X_{n+1} + b,$$

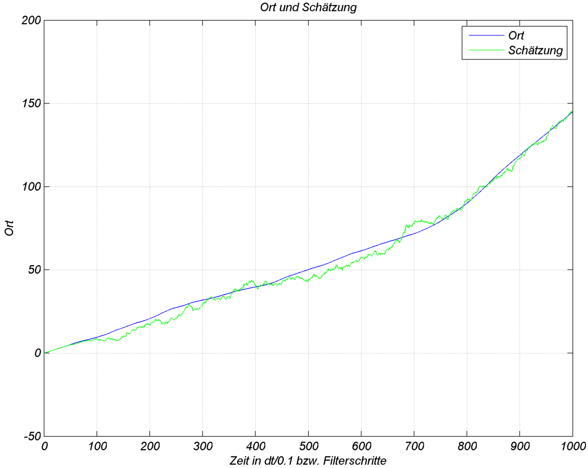
im Nachfolgenden wird von ungenauem Messgerät ausgegangen, d.h. $Var(a) \ll Var(b)$,

auf Grund der Einfachheit des Problems (gleichförmige, gleichmäßige Bewegung) interessieren wir uns nur für die Schätzung und nicht für die Güte der Schätzung (konvergiert Norm gegen einen konstanten Wert, je kleiner die Varianzen desto kleiner der Wert),

Beispiel

Startwerte:

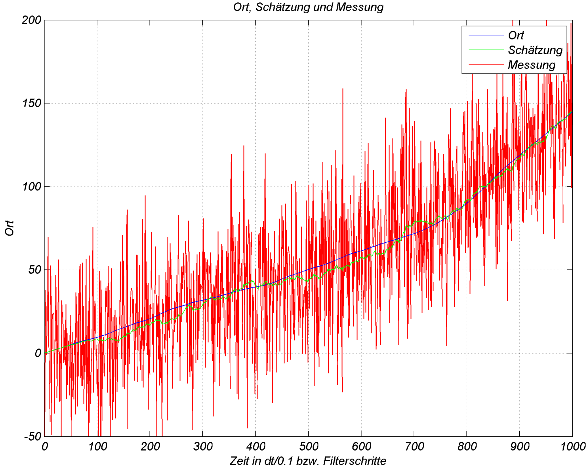
$$X_0 = (0 \;\; 1)^T, \;\; dt = 0.1, \;\; Var(a) = 0.3, \;\; Var(b) = 1000,$$



Beispiel

Startwerte:

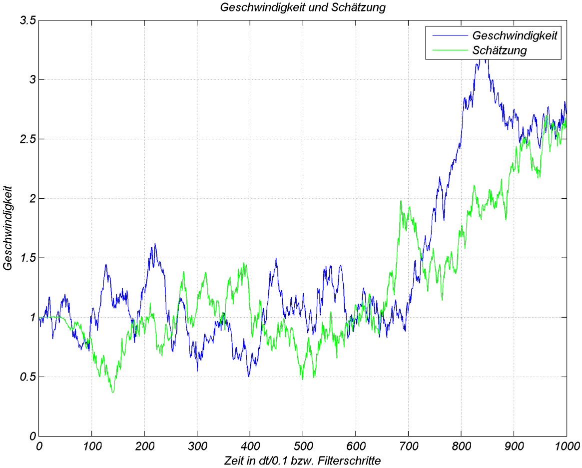
$$X_0 = (0 \;\; 1)^T, \;\; dt = 0.1, \;\; Var(a) = 0.3, \;\; Var(b) = 1000,$$



Beispiel

Startwerte:

$$X_0 = (0 \;\; 1)^T, \;\; dt = 0.1, \;\; Var(a) = 0.3, \;\; Var(b) = 1000,$$



Vielen Dank für Ihre Aufmerksamkeit