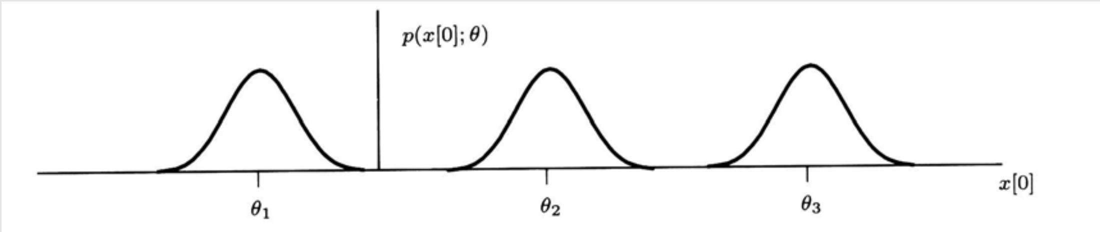
Das Schätzproblem

Geg.: N-elementiger Datensatz $X = \{ x[0], x[1], .. , x[N-1] \}$, der von einem unbekannten Parameter θ abhängt.

Ges.: Bestimmung von θ mit Hilfe eines Schätzers, d.h. einer Funktion g mit

$$\hat{\theta} = g\left(x[0], x[1], .., x[N-1]\right)$$



Grundidee: Modellierung der Wahrscheinlichkeitsdichte-funktion (WDF) der Daten in Abhängigkeit von θ

Klassische Schätztheorie: θ ist unbekannt, aber deterministisch, d.h. θ parametrisiert die WDF von *x*

Bayes'sche Schätztheorie: θ ist wie *x* die Realisierung einer Zufallsvariable => Modellierung über eine gemeinsame WDF $p(X, \theta) = p(X|\theta)\, p(\theta)$

Optimale Schätzer werden über Optimalitätskriterien gesucht, z.B. Minimierung der Varianz ohne systematische Fehler (MVU), kleinste Fehlerquadrate oder Minimierung des mittleren quadratischen Fehlers.

Mittlerer quadratischer Fehler

Ges.: Schätzer, der den (Bayes‘schen) mittleren quadratischen Fehler

$$e(\hat{\theta}) = \iint (\theta - \hat{\theta})^2 \, p(X, \theta) \, dX d\theta$$

minimiert.

Bayes‘sches Theorem: $p(X, \theta) = p(X|\theta)\, p(\theta)$ ← a-priori-WDF: vor Beobachtung der Daten (Vorwissen)

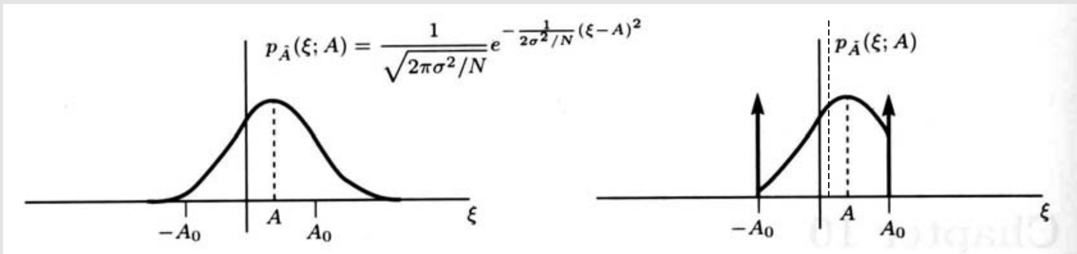
$$\Rightarrow \quad e(\hat{\theta}) = \int \left[\int (\theta - \hat{\theta})^2 \, p(\theta \mid X) \, d\theta \right] p(x)\, dX$$

Wähle $\hat{\theta}$ so, daß e minimiert wird :

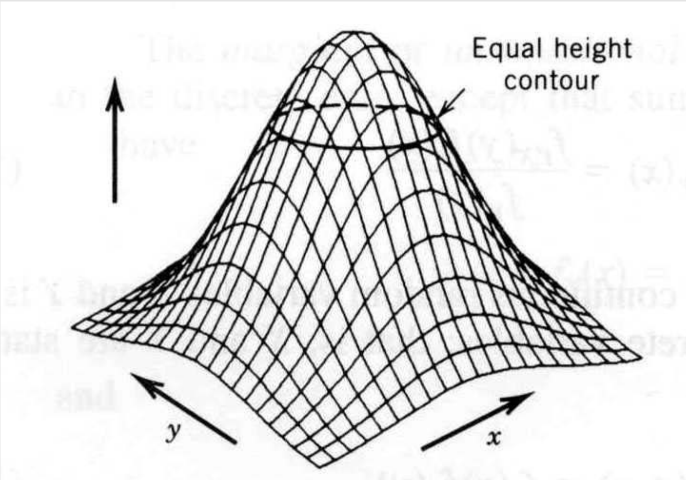
$$\frac{\partial}{\partial \hat{\theta}} \int (\theta - \hat{\theta})^2 \, p(\theta \mid X) \, d\theta = 0$$

Ergebnis: $\hat{\theta} = \int \theta \; p(\theta \mid X) \, d\theta = E(\theta \mid X)$

d.h. der optimale Schätzer, der den mittleren quadratischen Fehler minimiert (MMQF-Schätzer), ist der *Mittelwert der a-posteriori-WDF* (a posteriori : nach Beobachtung der Daten).



Multivariate Gaußverteilung



$$p(x) = \frac{1}{\sqrt{(2\pi)^N \det C_x}} \exp\left[-\frac{1}{2}(x-\mu_x)^T C_x^{-1}(x-\mu_x)\right]$$

C_x: Kovarianzmatrix (N x N, symmetrisch, positiv definit => invertierbar) mit $[C_x]_{ij} = E((x_i - E(x_i))(x_j - E(x_j)))$, μ_x: Mittelwert

- Ist C diagonal, dann sind die x_i unkorreliert und statistisch unabhängig, p faktorisiert in ein Produkt von 1D-Gaußfunktionen.
- Jede lineare Transformation von x mit $y = Ax + b$ ergibt wieder eine Gaußverteilung mit $\mu_y = A\mu_x + b$ und $C_y = AC_xA^T$.

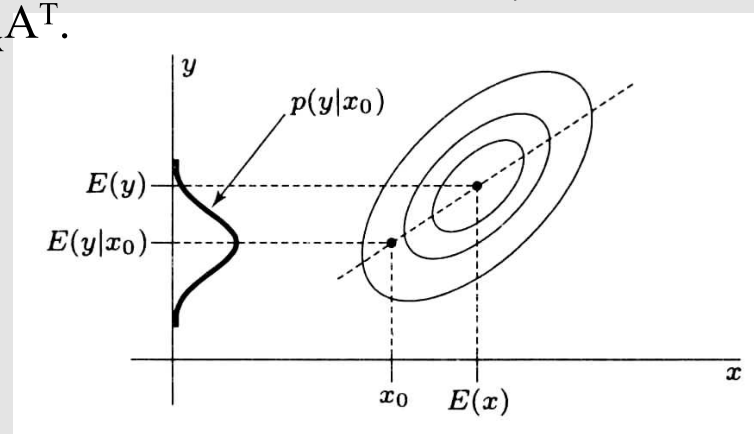
- Konturen konstanter Wahrscheinlichkeit sind Ellipsen
- Bedingte Wahrscheinlichkeiten sind wieder gaußverteilt mit

$$E(y \mid x) = E(y) + C_{xy}C_{xx}^{-1}(x - E(x))$$

$$C_{y|x} = C_{yy} - C_{yx}C_{xx}^{-1}C_{xy}$$

- Sind x_1 und x_1 unkorreliert, so gilt

$$E(y|x_1, x_2) = E(y|x_1) + E(y|x_2)$$



MMQF-Schätzer für y

Orthogonalitätsprinzip (gilt für jeden linearen Schätzer)

MMQF-Schätzer für Gaußverteilung: $\hat{\theta} = E(\theta) + C_{x\theta} C_{xx}^{-1} (x - E(x))$

Annahme $E(x) = E(\theta) = 0$ => $\hat{\theta} = C_{x\theta} C_{xx}^{-1} x$

d.h. Schätzer ist eine Linearkombination der Daten

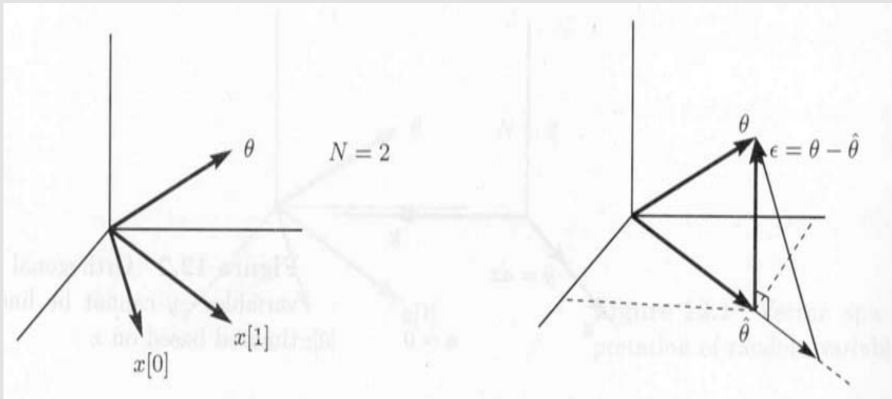
Anwendungsfall: Geg. seien Daten *x*[1] bis *x*[N-1], geschätzt werden soll *x*[N].

Innovation: Schätzfehler $\tilde{x}_N = \theta - \hat{\theta} = \theta - C_{x\theta} C_{xx}^{-1} x$

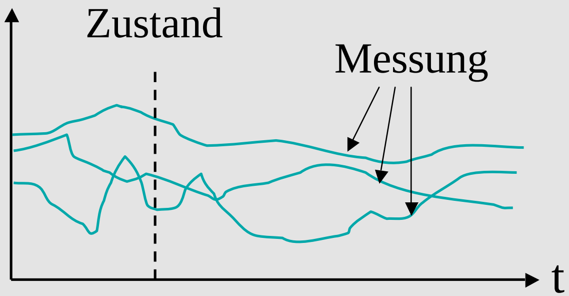
Korrelation: $E(\tilde{x}_n \cdot x) = E\left(\left(\theta - C_{x\theta} C_{xx}^{-1} x\right) \cdot x\right) = C_{x\theta} - C_{x\theta} C_{xx}^{-1} E(xx)$

$$= C_{x\theta} - C_{x\theta} C_{xx}^{-1} C_{xx} = C_{x\theta} - C_{x\theta} = 0$$

- Schätzfehler bzw. Innovation ist nicht mit den Daten korreliert, d.h. Innovation und Daten sind zueinander orthogonal
- MMQF-Schätzer nutzt die Korrelation der Daten untereinander aus bzw. „dekorreliert“ die Zeitreihe



Stochastische Prozesse



Zufallsvariable: Regel, die jedem Ausgang eines Zufallexperimentes eine *Zahl* zuweist.
Stochastischer Prozeß: Regel, die jedem Ausgang eines Zufallexperimentes eine *Funktion* bzw. eine *Zeitreihe* zuweist (d.h. ein Ensemble von Funktionen mit Wahrscheinlichkeitsdichte *p(x,t)*).

Charakterisierung:

- Mittelwert: $\mu_x(t) = E(x(t))$
- Autokorrelation: $R(t_1,t_2) = E(x(t_1)x(t_2))$ *R(t,t):* Durchschnittsleistung
- Autokovarianz: $C(t_1,t_2) = R(t_1,t_2) - \mu(t_1)\mu(t_2)$ *C(t,t)*: Varianz

Stationärer Prozeß (im weiteren Sinne, WSS): Mittelwert konstant, Autokorrelation hängt nur von Zeitdifferenz ab.

Diskreter WSS-Prozeß: $R(t_1,t_2) \to r_{xx}[k] = E(x[n]x[n+1])$

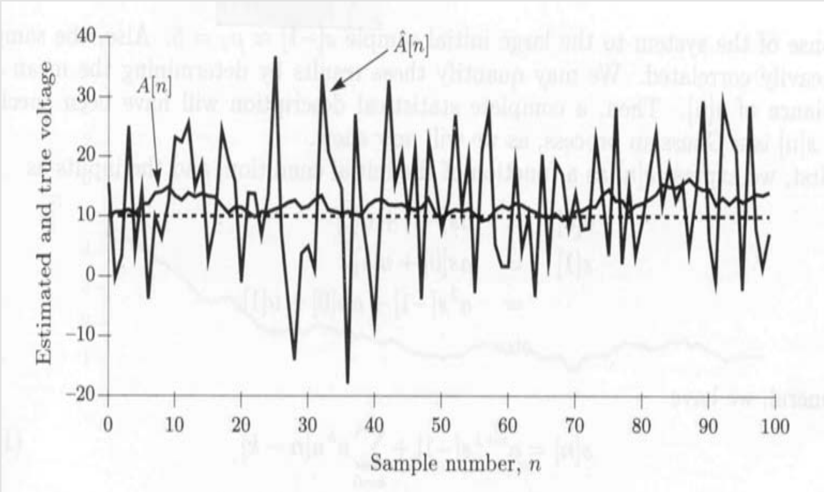
$$C(t_1,t_2) \to c_{xx}[k] = r_{xx}[k] - \mu_x^2$$

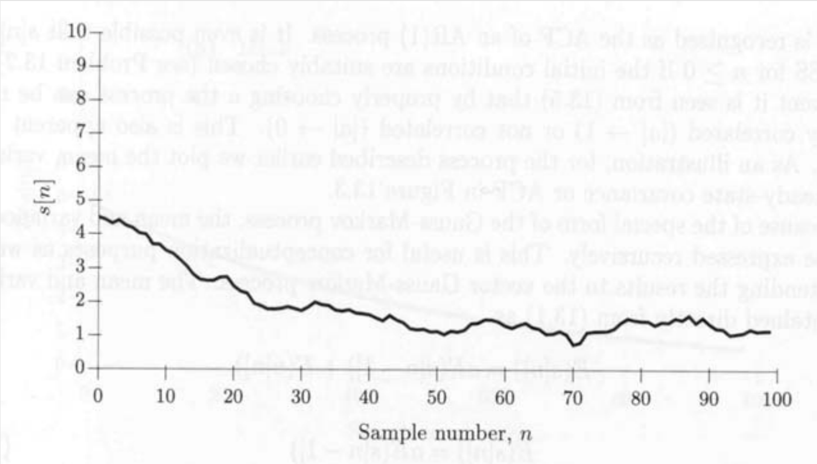
Skalarer Gauß-Markov-Prozeß

Statistische Modellierung eines Signals als korrelierter stochastischer Prozeß mit dem *Zustandsmodell*

$$s[n] = as[n-1] + u[n], \qquad n >= 0$$

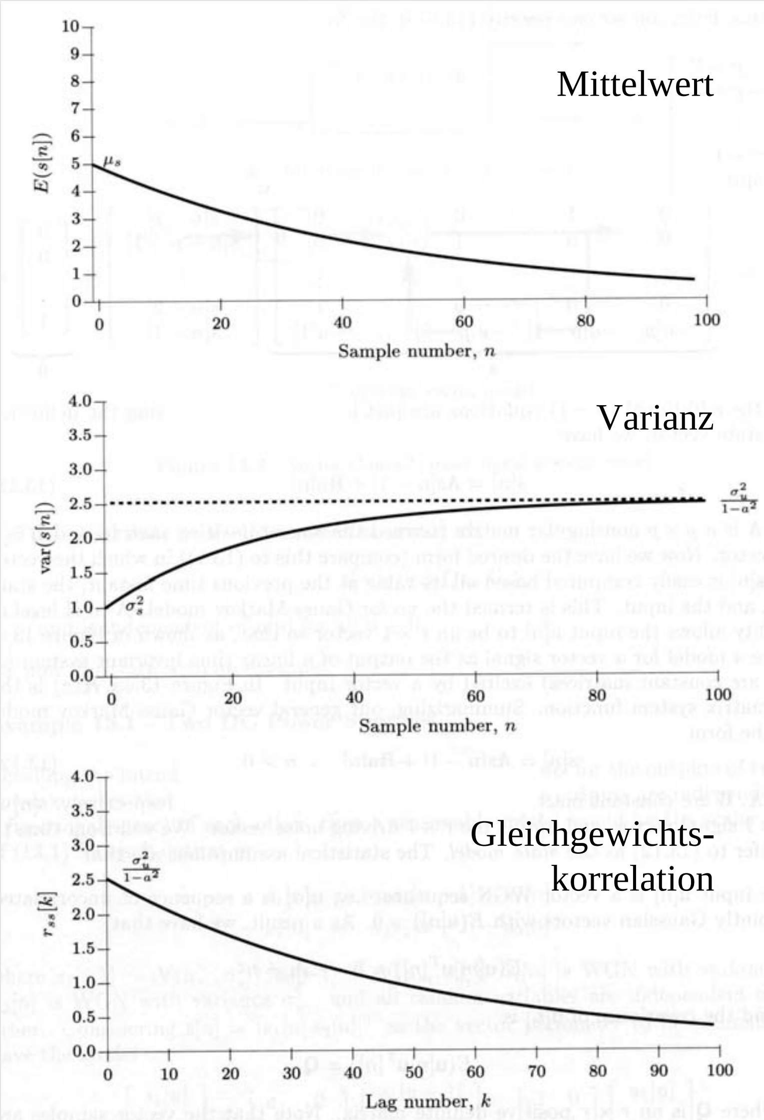
mit Zustand $s[n]$, $u \sim N(0, \sigma^2_u)$, $s[-1] \sim N(\mu_s, \sigma^2_s)$ und s[-1] unabhängig vom Anregungsrauschen u (Gauß-Markov-Prozeß 1. Ordnung).





- Zustand des Systems hängt nur vom vorherigen Zustand und vom Eingangswert u ab.
- Für $|a| < 1$ ist der Prozeß stabil, d.h. $s[n]$ bleibt beschränkt für alle n.
- Für $|a|$ -> 1 ist der Prozeß stark, für $|a|$ -> 0 schwach korreliert.
- Prozeß ist nicht stationär, wird es aber für unendliches n.
- $s[n]$ bleibt gaußverteilt für alle n.

Statistik des Gauß-Markov-Prozesses



Fortpflanzungsgleichungen:

$$E(s[n])$$
$$= aE(s[n-1]) + E(u[n])$$
$$= aE(s[n-1])$$
$$\Rightarrow E(s[n]) \xrightarrow{n\to\infty} 0$$

$$\mathrm{var}(s[n]) = E\left[(s[n] - E(s[n]))^2\right]$$
$$= E\left[(as[n-1] + u[n] - aE(s[n-1]))^2\right]$$
$$= a^2\,\mathrm{var}(s[n-1]) + \sigma_u^2$$
$$\Rightarrow \mathrm{var}(s[n]) \xrightarrow{n\to\infty} \frac{\sigma_u^2}{1-a^2}$$

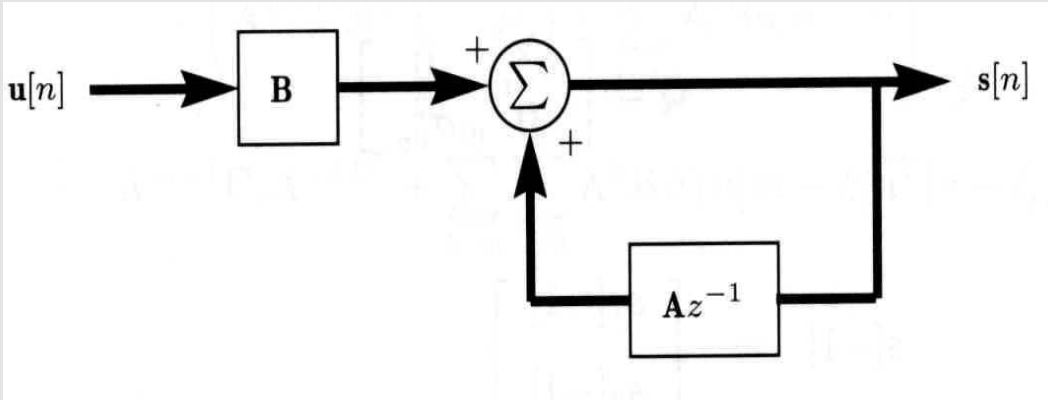
=> Für unendliches n wird der Prozeß stationär mit Autokorrelationsfunktion

$$r_{ss}[k] = \frac{\sigma_u^2}{1-a^2} a^k$$

Vektorielles Gauß-Markov-Modell

Erweiterung des skalaren Modells auf p x 1- Signalvektoren **s**[n] und r x 1-dimensionales Anregungsrauschen **u**[n]:

$$\mathbf{s}[n] = A\mathbf{s}[n-1] + B\mathbf{u}[n]$$



mit p x p - Matrix A und p x r - Matrix B, $\mathbf{u} \sim N(0, Q)$, $\mathbf{s}[-1] \sim N(\boldsymbol{\mu}_s, C_s)$ und **s**[-1] unabhängig von der Eingangsgröße **u**. A und B dürfen zeitlich variieren.

Fortpflanzungsgleichungen:

$$E(\mathbf{s}[n]) = AE(\mathbf{s}[n-1]) \qquad E(\mathbf{s}[n]) \xrightarrow{n\to\infty} 0$$

$$C[n] = AC[n-1]A^T + BQB^T \qquad C[n] \xrightarrow{n\to\infty} \sum_{k=0}^{\infty} A^k BQB^T A^{k^T}$$

- für konstantes A und B wird der Prozeß asymptotisch stationär.
- für ein stabiles System müssen die Eigenwertbeträge von A kleiner als 1 sein
- andere Eigenschaften analog zum skalaren Modell

Skalarer Kalmanfilter - Voraussetzungen

1. „Beobachtung“: Schätzung des Zustands ohne Berücksichtigung des Meßrauschens
„Filterung“: Schätzung des Zustands unter Berücksichtigung des Meßrauschens
=> Kalman-“Filter“

2. Skalares Gauß-Markov-Signalmodell: $s[n] = as[n - 1] + u[n]$, $n >= 0$
Meßmodell (additives Gauß-Rauschen): $x[n] = s[n] + w[n]$ mit $w \sim N(0, \sigma^2_n[n])$
unabh. von u und $s[-1]$

3. Vereinfachung: $\mu_s = 0$

4. A-priori-Wissen über den im Signalmodell gegebenen zeitlichen Zusammenhang soll verwendet werden => Bayes‘scher Ansatz.

5. Optimalitätskriterium: Minimierung des mittleren quadratischen Fehlers

6. Ges.: MMQF-Schätzer für $s[n]$ mit sequenzieller Korrektur, d.h. neue Schätzung soll sich aus der alten Schätzung $s[n - 1]$ und der aktuellen Beobachtung $x[n]$ berechnen.

Skalarer Kalmanfilter*

Vorhersage des Signals:

$$\hat{s}[n\,|\,n-1] = a\hat{s}[n-1\,|\,n-1]$$

Vorhersage des MMQF:

$$M[n\,|\,n-1] = a^2 M[n-1\,|\,n-1] + \sigma_u^2$$

Kalman-Gain:

$$K[n] = \frac{M[n\,|\,n-1]}{\sigma_n^2 + M[n\,|\,n-1]}$$

Korrektur der Signalschätzung:

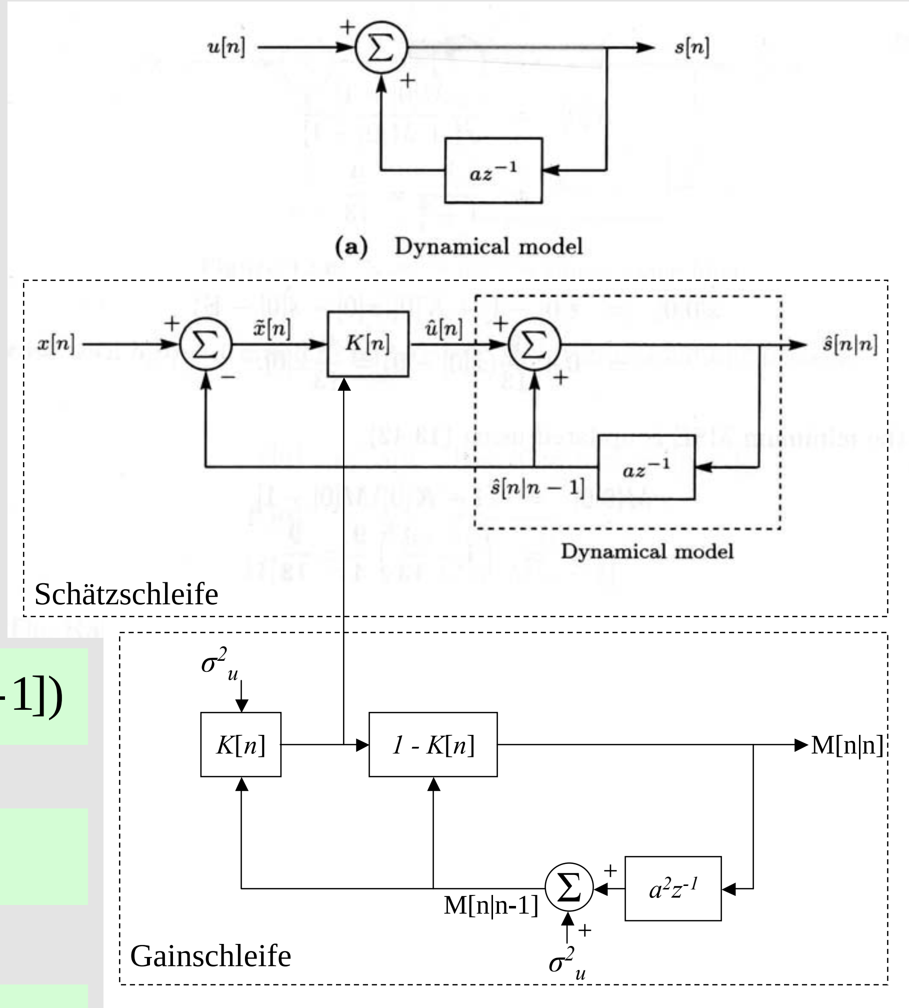
$$\hat{s}[n\,|\,n] = \hat{s}[n\,|\,n-1] + K[n](x[n] - \hat{s}[n\,|\,n-1])$$

Korrektur des MMQF:

$$M[n\,|\,n] = (1 - K[n])M[n\,|\,n-1]$$

Initialisierung:

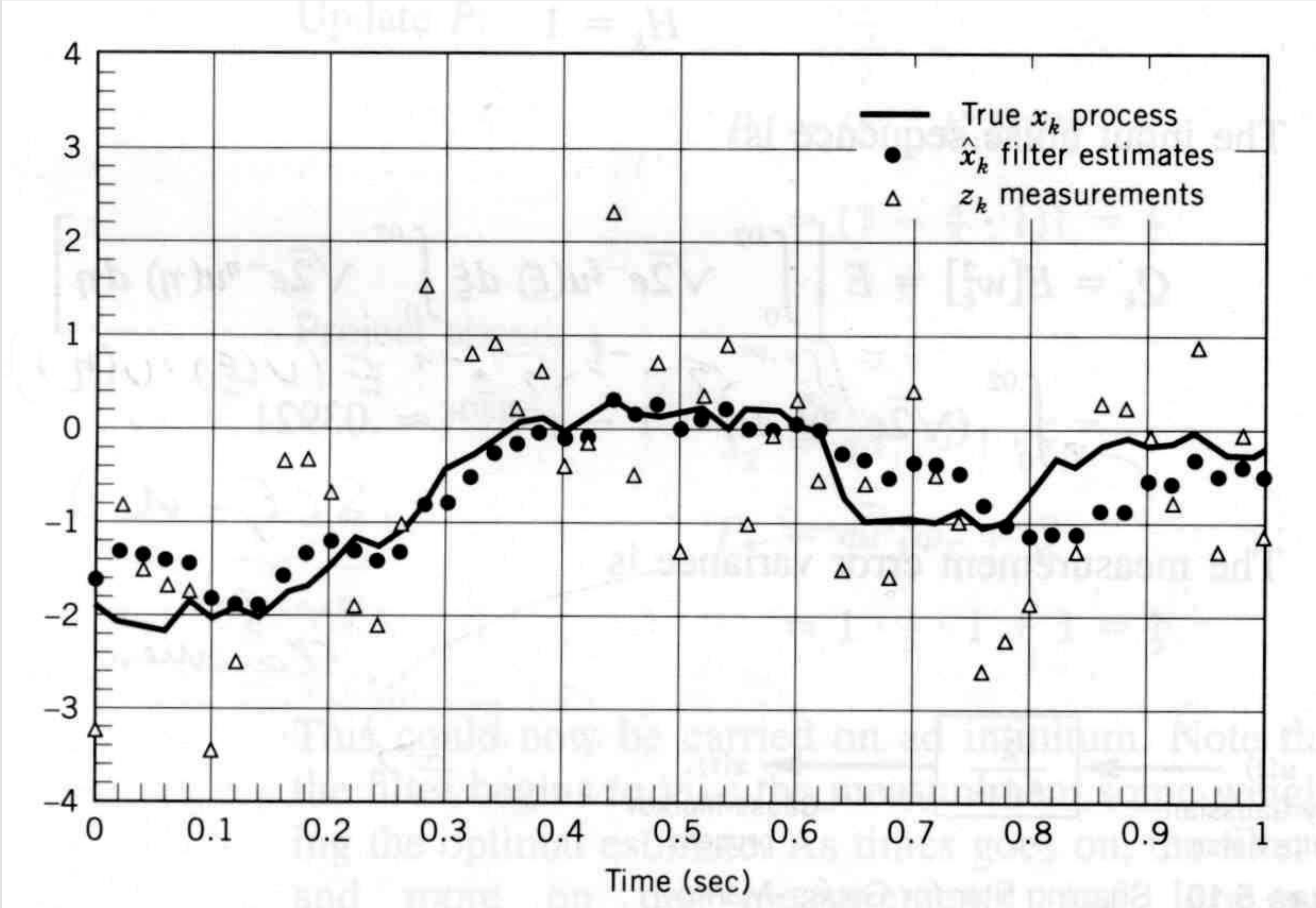
$$\hat{s}[-1\,|\,-1] = \mu_s, \quad M[-1\,|\,-1] = \sigma_s^2$$



*identische Gleichungen für $\mu_s \neq 0$

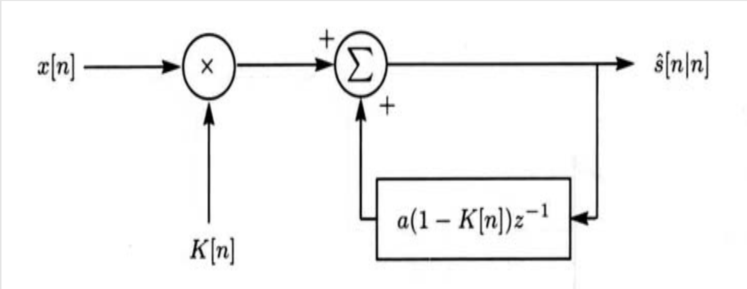
Beispiel

$a = 0.9802$
$\sigma^2_u = 0.039$
$\sigma^2_n = 1$
$\mu_s = 0$
$\sigma^2_s = 1$

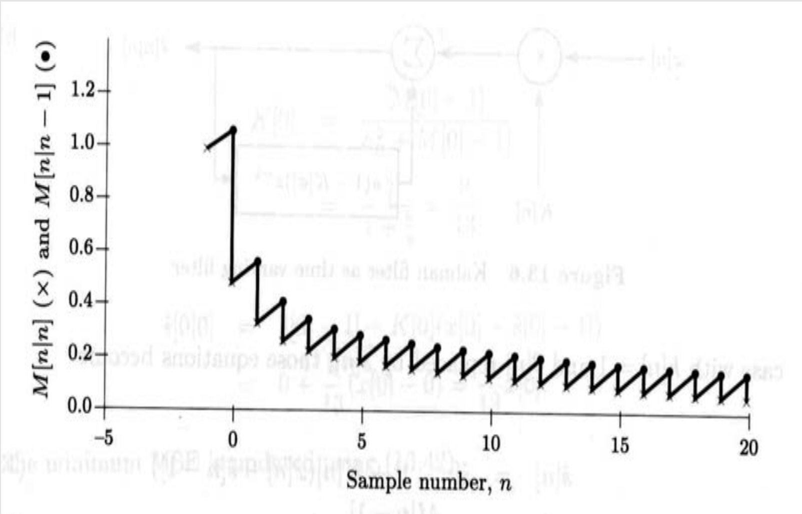


Kalmanfilter - Eigenschaften

- Gainberechnung läuft ohne Rückkopplung der Daten in einer unabhängigen Schleife
- Gain kann schon im voraus berechnet werden
- Gainschleife kann instabil werden
- Kalmanfilter ist ein rekursiver linearer Filter erster Ordnung mit zeitvarianten Koeffizienten
- Keine Matrixinversion notwendig
- Kalmanfilter liefert gleichzeitig sein eigenes Konfidenzmaß, den MMQF M[n|n]
- MMQF steigt im Vorhersageschritt und sinkt im Korrekturschritt.
- Kalmanfilter ist der bestmögliche lineare Filter, selbst wenn Gauß'sche Annahme nicht zutrifft.
- kann auf Vektorzustände und -messungen verallgemeinert werden.



$$\hat{s}[n \mid n] = a\hat{s}[n-1 \mid n-1] +$$
$$K[n](x[n] - a\hat{s}[n-1 \mid n-1])$$
$$= a(1-K[n])\hat{s}[n-1 \mid n-1] + K[n]x[n]$$



Vektorieller Zustand - skalare Messung

Signalmodell:
Vektorielles Gauß-Markov-Modell

$\mathbf{s}[n] = A\mathbf{s}[n-1] + B\mathbf{u}[n]$

mit Zusatzannahme Q diagonal.

Meßmodell:

$x[n] = \mathbf{h}^{\mathrm{T}}[n]\mathbf{s}[n] + w[n]$

mit $w \sim N(0, \sigma^2_n[n])$ unabh. von $\mathbf{u}$ und $\mathbf{s}[-1]$

- Kalman-Gain ist jetzt ein Vektor, MMQF eine Matrix
- Es müssen genügend Messungen vorhanden sein
- A und **h** müssen bestimmten Bedingungen genügen (Beobachtbarkeit)

Vorhersage des Signals:

$$\hat{\mathbf{s}}[n\,|\,n-1] = A\hat{\mathbf{s}}[n-1\,|\,n-1]$$

Vorhersage des MMQF (p x p - Matrix):

$$M[n\,|\,n-1] = AM[n-1\,|\,n-1]A^T + BQB^T$$

Kalman-Gain-Vektor (p x 1):

$$\mathbf{K}[n] = \frac{M[n\,|\,n-1]\mathbf{h}[n]}{\sigma_n^2 + \mathbf{h}^T[n]M[n\,|\,n-1]\mathbf{h}[n]}$$

Korrektur der Signalschätzung:

$$\hat{\mathbf{s}}[n\,|\,n] = \hat{\mathbf{s}}[n\,|\,n-1] + \mathbf{K}[n](x[n] - \mathbf{h}^T[n]\hat{\mathbf{s}}[n\,|\,n-1])$$

Korrektur des MMQF:

$$M[n\,|\,n] = (1 - \mathbf{K}[n]\mathbf{h}^T[n])M[n\,|\,n-1]$$

Initialisierung:

$$\hat{\mathbf{s}}[-1\,|\,-1] = \boldsymbol{\mu}_s, \quad M[-1\,|\,-1] = C_s$$

Vektorieller Zustand - vektorielle Messung

Signalmodell: wie vorher.

Lineares Meßmodell:Messungen sind m x 1- Vektoren

$\mathbf{x}[n] = \mathrm{H}[n]\mathbf{s}[n] + \mathbf{w}[n]$

mit **w** ~ *N(0, C*[*n*]) unabh. von **u** und **s**[-1]

- Kalman-Gain und MMQF sind Matrizen
- Achtung: Matrixinversion für Berechnung der Gainmatrix erforderlich.
- Erweiterungen:

- *A*, *B*, *Q* zeitvariant => identische Gleichungen
- zeitliche Korrelationen in **w**
- deterministisches **u**

Vorhersage des Signals:

$$\hat{\mathbf{s}}[n|n-1] = A\hat{\mathbf{s}}[n-1|n-1]$$

Vorhersage des MMQF (p x p - Matrix):

$$M[n|n-1] = AM[n-1|n-1]A^T + BQB^T$$

Kalman-Gain-Matrix (p x m):

$$K[n] = M[n|n-1]H^T[n]\left(C[n] + H[n]M[n|n-1]H^T[n]\right)^{-1}$$

Korrektur der Signalschätzung:

$$\hat{\mathbf{s}}[n|n] = \hat{\mathbf{s}}[n|n-1] + K[n](x[n] - H[n]\hat{\mathbf{s}}[n|n-1])$$

Korrektur des MMQF:

$$M[n|n] = (1 - K[n]H[n])M[n|n-1]$$

Initialisierung:

$$\hat{\mathbf{s}}[-1|-1] = \mu_s, \quad M[-1|-1] = C_s$$

Erweiterter Kalmanfilter (EKF)

Erweiterung auf nichtlineares Signal- und Meßmodell mit additiven Rauschtermen:

$$\mathbf{s}[n] = \mathbf{a}(\mathbf{s}[n-1]) + \mathrm{B}\mathbf{u}[n] \qquad \text{und} \qquad x[n] = \mathbf{h}(\mathbf{s}[n]) + \mathbf{w}[n]$$

mit nichtlinearen Funktionen **a** und **h**. Geschlossene Lösung für MMQF-Schätzer ist nicht bekannt => Näherungslösung durch Linearisierung um $\hat{\mathbf{s}}[n-1|n-1]$ und $\hat{\mathbf{s}}[n|n-1]$

$$\mathbf{a}(\mathbf{s}[n-1]) \approx \mathbf{a}(\hat{\mathbf{s}}[n-1|n-1]) + \left.\frac{\partial \mathbf{a}}{\partial \mathbf{s}[n-1]}\right|_{\substack{\mathbf{s}[n-1]= \\ \hat{\mathbf{s}}[n-1|n-1]}} (\mathbf{s}[n-1] - \hat{\mathbf{s}}[n-1|n-1]) \qquad \Rightarrow A[n-1] = \left.\frac{\partial \mathbf{a}}{\partial \mathbf{s}[n-1]}\right|_{\substack{\mathbf{s}[n-1]= \\ \hat{\mathbf{s}}[n-1|n-1]}}$$

$$\mathbf{h}(\mathbf{s}[n]) \approx \mathbf{h}(\hat{\mathbf{s}}[n|n-1]) + \left.\frac{\partial \mathbf{h}}{\partial \mathbf{s}[n]}\right|_{\substack{\mathbf{s}[n]= \\ \hat{\mathbf{s}}[n|n-1]}} (\mathbf{s}[n] - \hat{\mathbf{s}}[n|n-1]) \qquad \Rightarrow H[n] = \left.\frac{\partial \mathbf{h}}{\partial \mathbf{s}[n]}\right|_{\substack{\mathbf{s}[n]= \\ \hat{\mathbf{s}}[n|n-1]}}$$

Filtergleichungen ändern sich nicht, bis auf

Vorhersage: $\hat{\mathbf{s}}[n|n-1] = \mathbf{a}(\hat{\mathbf{s}}[n-1|n-1])$

und Korrektur:

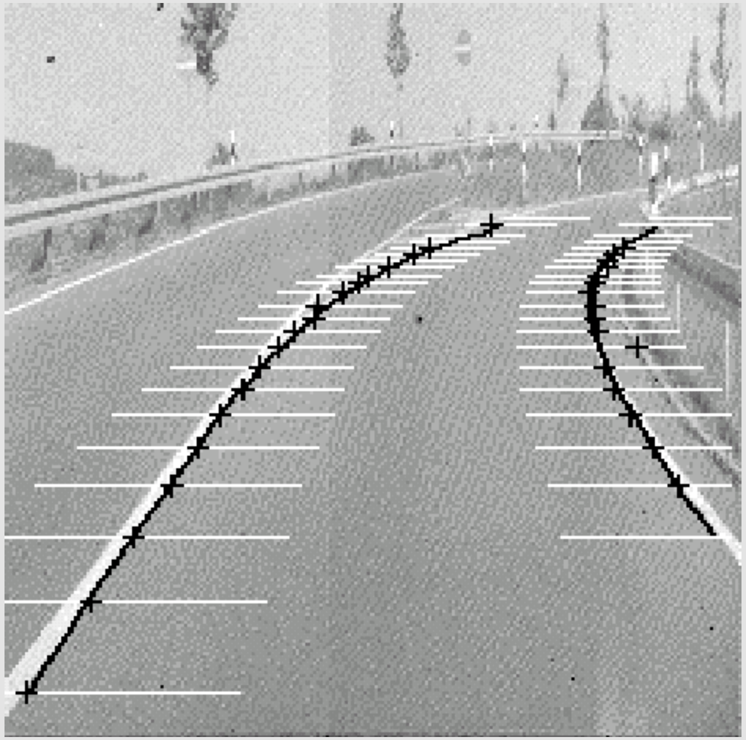
$$\hat{\mathbf{s}}[n|n] = \hat{\mathbf{s}}[n|n-1] + K[n](x[n] - \mathbf{h}(\hat{\mathbf{s}}[n|n-1]))$$

Achtung:
- keine Optimalitätseigenschften
- M ist hier kein verläßliches Konfidenzmaß
- A[n-1] H[n] müssen in jedem Zeitschritt neu berechnet werden.
- K und M können nicht im voraus berechnet werden.

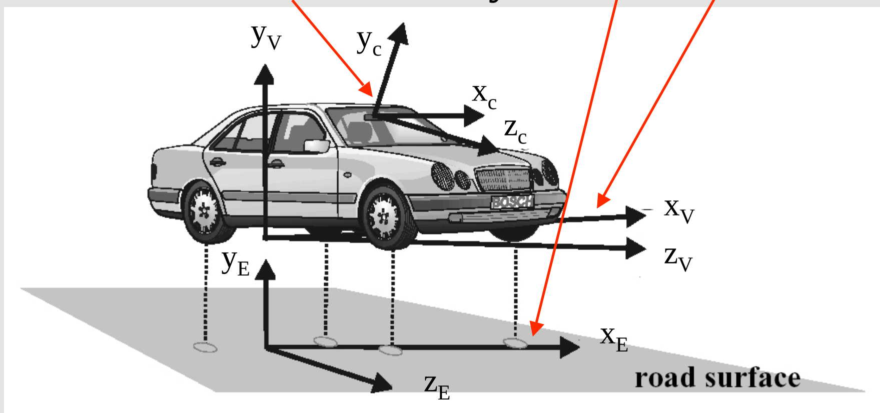
Anwendungsbeispiel Spurwarner

Aufgabe: Detektion der Fahrbahnrandmarkierungen im Kamerabild, daraus Schätzung der Fahrzeugposition innerhalb des Spurverlaufs mit einem erweiterten Kalmanfilter und ggf. Warnung bei Verlassen der Spur.

Kameras



3 Koordinatensysteme: Fahrbahn-, Fahrzeug- und Kamerakoordinatensystem:



Modellierung eines Zeitschritts

- Spurverlauf wird als Polynom 3. Grades beschrieben:

$$x_E(z_E) = \Delta x + \frac{c_0}{2} z_E^2 + \frac{c_1}{6} z_E^3$$

mit Krümmung

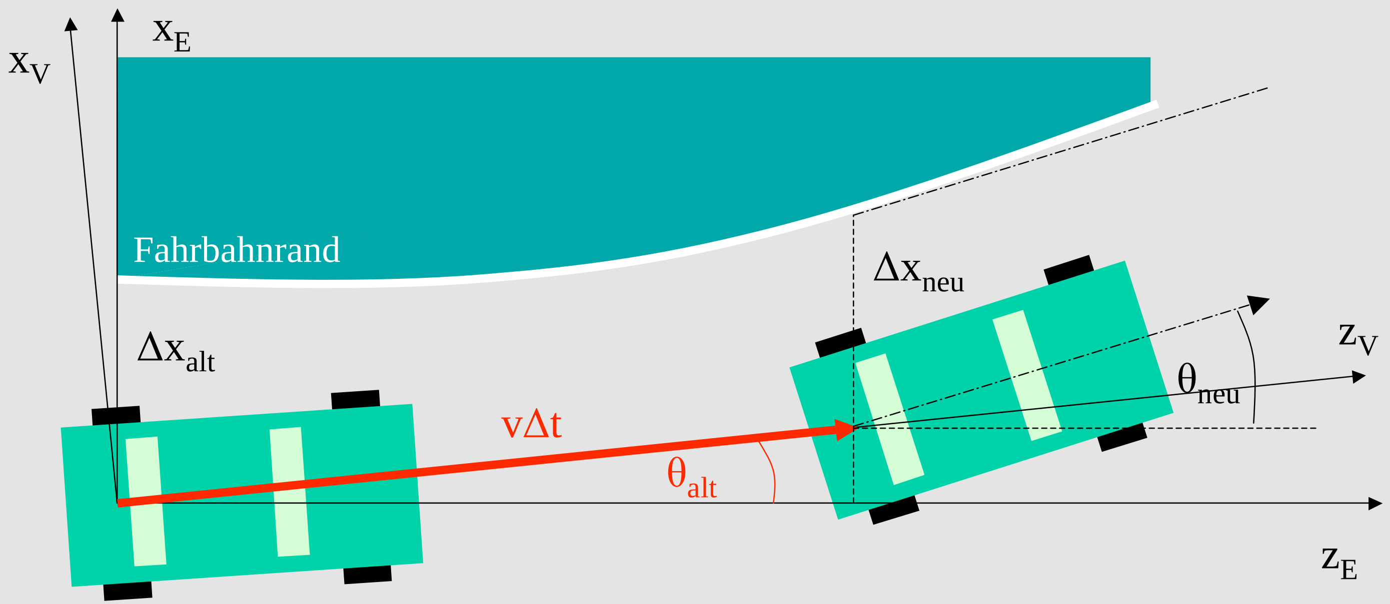
$$c_0 = \frac{d^2 x_E}{dz_E^2}$$

mit Krümmungsänderung

$$c_1 = \frac{d^3 x_E}{dz_E^3}$$

und Abstand zum Fahrbahnrand Δx

- Bewegung im Zeitschritt Δt sei stückweise konstant, d.h. erst Translation um v Δt in Richtung θ, dann Drehung parallel zum Spur-verlauf.
- Straßen-KS ist tangential zur Spur.
- v und Δt sei bekannt.



Aufstellen des Signalmodells

1. Spurposition:

$$\Delta x_{neu} = \Delta x_{alt} + \frac{c_0}{2} z_E^2 + \frac{c_1}{6} z_E^3 - v\Delta t \tan\theta$$

$$\xrightarrow{\theta<<1} z_E \approx v\Delta t, \quad \tan\theta \approx \theta \quad \Rightarrow$$

Annahme: Kleine Gierwinkel und Krümmungen

$$\Delta x_{neu} = \Delta x_{alt} + \frac{c_0^{alt}}{2}(v\Delta t)^2 + \frac{c_1^{alt}}{6}(v\Delta t)^3 - \theta v_{alt}\Delta t$$

2. Krümmung und Krümmungsänderung:

$$c_0^{neu} = c_0^{alt} + \left.\frac{\partial c_0}{\partial z_E}\right|_{z_E=0} \Delta z_E \quad \Rightarrow$$

$$c_0^{neu} = c_0^{alt} + c_1^{alt} v\Delta t$$

$$c_1^{neu} = c_1^{alt}$$

3. Gier-winkel:

$$\theta_{neu} = \theta_{alt} + \tan\frac{dx_E}{dz_E} \approx \theta_{alt} + \frac{dx_E}{dz_E} =$$

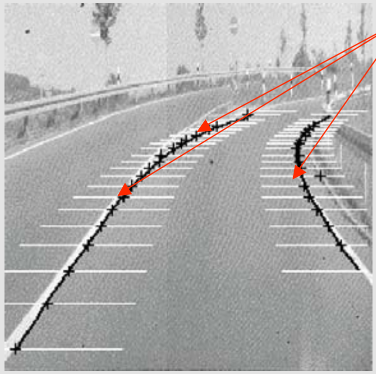
$$\theta_{alt} + c_0 z_E + \frac{c_1}{2} z_E^2 \quad \Rightarrow$$

$$\theta_{neu} = \theta_{alt} + c_0^{alt} v\Delta t + \frac{c_1^{alt}}{2}(v\Delta t)^2$$

=> Signalmodell: $\mathbf{s}[n] = A\mathbf{s}[n-1] + \mathbf{u}[n]$ mit $\mathbf{s}[n] = (\Delta x, c_0, c_1, \theta)^T$ und

$$A = \begin{pmatrix} 1 & \frac{1}{2}(v\Delta t)^2 & \frac{1}{6}(v\Delta t)^3 & -v\Delta t \\ 0 & 1 & v\Delta t & 0 \\ 0 & 0 & 1 & 0 \\ 0 & v\Delta t & \frac{1}{2}(v\Delta t)^2 & 1 \end{pmatrix}, \quad Q = \begin{pmatrix} \sigma_{\Delta x}^2 & 0 & 0 & 0 \\ 0 & \sigma_{c_0}^2 & 0 & 0 \\ 0 & 0 & \sigma_{c_1}^2 & 0 \\ 0 & 0 & 0 & \sigma_\theta^2 \end{pmatrix}$$

Aufstellen des Meßmodells



1. N Beobachtungen pro Frame mit

$$\mathbf{x}[n] = \begin{pmatrix} x_b(\mathbf{s}[n], y_{b,1}) + w_1 \\ x_b(\mathbf{s}[n], y_{b,2}) + w_2 \\ \vdots \\ x_b(\mathbf{s}[n], y_{b,N}) + w_N \end{pmatrix}$$

2. Transformation Welt-Kamera:

$$x_C = \cos\theta x_E - \sin\theta z_E$$
$$y_C = y_E - h$$
$$z_C = \sin\theta x_E + \cos\theta z_E - d$$

mit Montagehöhe h und Entfernung zur Hinterachse d

3. Perspektivische Abbildung für $\theta << 1$:

$$x_b = f\frac{x_E - \theta z_E}{\theta\, x_E + z_E - d}$$

$$y_b = f\frac{y_E - h}{\theta\, x_E + z_E - d}$$

4. Auflösen nach z_E für $\theta\, z_E << z_E$:

$$z_E \approx d - \frac{fh}{y_b} \overset{\text{einsetzen}}{\Rightarrow}$$

5. Meßgleichung (nichtlinear):

$$x_b = f\frac{\Delta x + \frac{c_0}{2}z_E^2 + \frac{c_1}{6}z_E^3 - \theta z_E}{\theta\left(\Delta x + \frac{c_0}{2}z_E^2 + \frac{c_1}{6}z_E^3\right) + z_E - d}$$

=> erweiterter Kalmanfilter benötigt linearisierte Meßgleichung mit den Ableitungen $\frac{\partial x_b}{\partial(\Delta x)}, \frac{\partial x_b}{\partial \theta}, \frac{\partial x_b}{\partial c_0}, \frac{\partial x_b}{\partial c_1}$

Beispiele