Kalman Filter

1) Vereinfachte Herleitung für ein skalares System

Es soll der **Filterschritt** des Kalman-Filters, also des linearen erwartungstreuen Schätzers mit der kleinstmöglichen Schätzfehlerkovarianz, hergeleitet werden. Wir betrachten hier ein skalares System mit der **linearen Messgleichung**

$$\boldsymbol{y} \;=\; H\boldsymbol{x} + \boldsymbol{v} \;. \tag{1}$$

$\boldsymbol{y}$ ist die Messung, $\boldsymbol{x}$ ist der aktuelle Systemzustand und $\boldsymbol{v}$ ist eine **mittelwertfreie** additive Störung mit der Varianz $\sigma_{\boldsymbol{v}}^2 = C^{\boldsymbol{v}}$. Das Messmodell, die Messmatrix $\mathbf{H}$ ist in diesem Fall also nur ein skalarer Faktor H. Man könnte es sich vorstellen als fehlerbehaftete Messung mit einem linearen Messumformer, wobei H die Proportionalitätskonstante ist, mit der der Messumformer die Messgröße, beispielsweise eine Kraft, in eine Spannung umwandelt.

Aus der Prädiktion seien **Prädiktions-Schätzwert** $\hat{x}^p$ und **Prädiktions-Kovarianz** C^p der **Prioren** $\boldsymbol{x}^p$ bereits bekannt. Nun wollen wir mit der **Messung** $\boldsymbol{y}$, sprich dem aktuell gewonnenen **Messwert** $\hat{y}$ und dem aktuellen **Messrauschen** $C^{\boldsymbol{y}}$ (welches ebenfalls bekannt sei) diese Schätzung zur **Posterioren** $\boldsymbol{x}^e$ mit **Filter-Schätzwert** $\hat{x}^e$ und (hoffentlich kleinerer) **Filter-Kovarianz** C^e verbessern. Mit anderen Worten, die Information aus $\boldsymbol{x}^p$ und $\boldsymbol{y}$ soll zu $\boldsymbol{x}^e$ fusioniert werden. Da wir ein **lineares Filter** suchen, verwenden wir als Ansatz hierfür die Linearkombination der beiden bekannten Zufallsvariablen $\boldsymbol{x}^p$ und $\boldsymbol{y}$

$$\boldsymbol{x}^e \;=\; \widetilde{K}\boldsymbol{x}^p \;+\; K\boldsymbol{y} \tag{2}$$

und werden nun die Verstärkungsfaktoren $\widetilde{K}$ und K so bestimmen, dass der Schätzer erstens erwartungstreu ist und zweitens die Filter-Kovarianz C^e möglichst klein wird.

a) Berechnen Sie aus (2) und (1) den **Erwartungswert** $\mathrm{E}\{\boldsymbol{x}^e\}$ in Abhängigkeit von $\mathrm{E}\{\boldsymbol{x}\}$, $\widetilde{K}$, K und H. Nehmen Sie dabei an, dass die Prädiktion erwartungstreu ist, d. h. $\mathrm{E}\{\boldsymbol{x}^p\} = \mathrm{E}\{\boldsymbol{x}\}$.

(1) in (2) einsetzen

$$\begin{aligned}
\boldsymbol{x}^e &= \widetilde{K}\boldsymbol{x}^p \;+\; K\cdot(H\boldsymbol{x}+\boldsymbol{v}) \\
&= \widetilde{K}\cdot\boldsymbol{x}^p \;+\; KH\boldsymbol{x} \;+\; K\boldsymbol{v}
\end{aligned}$$

Erwartungswert bilden

$$\begin{aligned}
\mathrm{E}\{\boldsymbol{x}^e\} &= \mathrm{E}\Big\{\widetilde{K}\boldsymbol{x}^p \;+\; KH\boldsymbol{x} \;+\; K\boldsymbol{v}\Big\} \\
&= \mathrm{E}\Big\{\widetilde{K}\cdot\boldsymbol{x}^p\Big\} \;+\; \mathrm{E}\{KH\boldsymbol{x}\} \;+\; \mathrm{E}\{K\boldsymbol{v}\} \\
&= \widetilde{K}\cdot\mathrm{E}\{\boldsymbol{x}^p\} \;+\; KH\cdot\mathrm{E}\{\boldsymbol{x}\} \;+\; K\cdot\mathrm{E}\{\boldsymbol{v}\} \\
&\qquad \mid \mathrm{E}\{\boldsymbol{x}^p\} = \mathrm{E}\{\boldsymbol{x}\}\,; \quad \boldsymbol{v} \text{ ist mittelwertfrei } \Rightarrow \mathrm{E}\{\boldsymbol{v}\} = 0 \\
&= \underline{\underline{\widetilde{K}\,\mathrm{E}\{\boldsymbol{x}\} \;+\; KH\,\mathrm{E}\{\boldsymbol{x}\}}}
\end{aligned}$$

b) Das Filter soll **erwartungstreu** sein, d. h. der Erwartungswert der Posterioren $\boldsymbol{x}^e$ soll dem Erwartungswert des tatsächlichen Zustands $\boldsymbol{x}$ entsprechen. Eliminieren Sie durch diese Bedingung $\widetilde{K}$ und drücken Sie die posteriore Schätzung (2) so aus, dass sie nur noch von K abhängt.

$$\begin{aligned}
\mathrm{E}\{\boldsymbol{x}^e\} &= \mathrm{E}\{\boldsymbol{x}\} \\
\widetilde{K}\,\mathrm{E}\{\boldsymbol{x}\} \;+\; KH\,\mathrm{E}\{\boldsymbol{x}\} &= \mathrm{E}\{\boldsymbol{x}\} \\
\widetilde{K} \;+\; KH &= 1 \\
\widetilde{K} &= 1 - KH \qquad \mid \text{ in (2) einsetzen} \\
\Rightarrow \quad \boldsymbol{x}^e &= \underline{\underline{(1-KH)\,\boldsymbol{x}^p \;+\; K\boldsymbol{y}}} \\
&= \underline{\underline{\boldsymbol{x}^p \;+\; K\cdot(\boldsymbol{y} - H\boldsymbol{x}^p)}}
\end{aligned}$$

$$\text{vgl.} \quad \underline{\hat{x}}_k^e \;=\; \underline{\hat{x}}_k^p + \mathbf{K}_k \left(\underline{\hat{y}}_k - \mathbf{H}_k \underline{\hat{x}}_k^p\right)$$

c) Nun betrachten wir den Schätzfehler $(\boldsymbol{x} - \boldsymbol{x}^e)$. Zeigen Sie, dass sich der Schätzfehler ergibt zu

$$(\boldsymbol{x} - \boldsymbol{x}^e) \;=\; (1 - KH) \cdot (\boldsymbol{x} - \boldsymbol{x}^p) \;-\; K\boldsymbol{v} \;. \tag{3}$$

$$\begin{aligned}
\boldsymbol{x} - \boldsymbol{x}^e \;&=\; \boldsymbol{x} \;-\; [(1-KH)\,\boldsymbol{x}^p \;+\; K\boldsymbol{y}] \\
&=\; \boldsymbol{x} \;-\; [(1-KH)\,\boldsymbol{x}^p \;+\; K \cdot (H\boldsymbol{x} + \boldsymbol{v})] \\
&=\; -\;(1-KH)\,\boldsymbol{x}^p \;+\; \boldsymbol{x} \;-\; KH\,\boldsymbol{x} \;-\; K\boldsymbol{v} \\
&=\; (1-KH)\cdot(\boldsymbol{x} - \boldsymbol{x}^p) \;-\; K\boldsymbol{v}
\end{aligned}$$

d) Den verbleibenden Freiheitsgrad K wollen wir verwenden, um den Erwartungswert des quadratischen Schätzfehlers, also die **Filter-Kovarianz** $C^e = \mathrm{Cov}\{\boldsymbol{x} - \boldsymbol{x}^e\} = \mathrm{E}\Big\{(\boldsymbol{x} - \boldsymbol{x}^e)^2\Big\}$, zu minimieren. Geben Sie zunächst

$$C^e \;=\; \mathrm{E}\Big\{(\boldsymbol{x} - \boldsymbol{x}^e)^2\Big\}$$

in Abhängigkeit von K, H, C^p und $C^{\boldsymbol{v}}$ an. Nehmen Sie dabei an, dass das $\boldsymbol{v}$ und $\boldsymbol{x}$ sowie $\boldsymbol{v}$ und $\boldsymbol{x}^p$ unkorreliert sind. Berechnen Sie dann dasjenige K, welches C^e minimiert.

Zunächst wird C^e berechnet.

$$\begin{aligned}
C^e \;&=\; \mathrm{Cov}\{(\boldsymbol{x} - \boldsymbol{x}^e)\} \\
&=\; \mathrm{E}\Big\{(\boldsymbol{x} - \boldsymbol{x}^e \;-\; \mathrm{E}\{\boldsymbol{x} - \boldsymbol{x}^e\})^2\Big\} \quad \Big|\; \text{erwartungstreu: } \hat{x} = \hat{x}^e, \quad \mathrm{E}\{\boldsymbol{x} - \boldsymbol{x}^e\} = 0 \\
&=\; \mathrm{E}\Big\{(\boldsymbol{x} - \boldsymbol{x}^e)^2\Big\} \quad \Big|\; \text{mit (3)} \\
&=\; \mathrm{E}\Big\{[(1-KH)\cdot(\boldsymbol{x} - \boldsymbol{x}^p) \;-\; K\boldsymbol{v}]^2\Big\} \quad \Big|\; (a-b)^2 = a^2 - 2ab + b^2 \\
&=\; \mathrm{E}\Big\{(1-KH)^2 \cdot (\boldsymbol{x} - \boldsymbol{x}^p)^2 \;-\; 2\,(1-KH)\,(\boldsymbol{x} - \boldsymbol{x}^p)\,K\boldsymbol{v} \;+\; (K\boldsymbol{v})^2\Big\} \\
&=\; (1-KH)^2 \cdot \underbrace{\mathrm{E}\Big\{(\boldsymbol{x} - \boldsymbol{x}^p)^2\Big\}}_{C^p} - 2\,(1-KH)\,K \cdot \left(\underbrace{\mathrm{E}\Big\{\underbrace{\boldsymbol{x}\cdot\boldsymbol{v}}_{\text{unkorr.}}\Big\}}_{0} - \underbrace{\mathrm{E}\Big\{\underbrace{\boldsymbol{x}^p\cdot\boldsymbol{v}}_{\text{unkorr.}}\Big\}}_{0}\right) \;+\; K^2 \cdot \underbrace{\mathrm{E}\{\boldsymbol{v}^2\}}_{C^{\boldsymbol{v}}} \\
&=\; \underline{\underline{(1-KH)^2\,C^p \;+\; K^2 C^{\boldsymbol{v}}}}
\end{aligned}$$

Jetzt können wir C^e nach dem Parameter K minimieren. Dazu setzen wir die Ableitung gleich Null.

$$\begin{aligned}
\frac{\partial}{\partial K} C^e \;&=\; \frac{\partial}{\partial K}\,(1-KH)^2\,C^p \;+\; K^2 C^{\boldsymbol{v}} \\
&=\; -2H\cdot(1-KH)\,C^p \;+\; 2KC^{\boldsymbol{v}} \\
\frac{\partial}{\partial K} C^e \;&\overset{!}{=}\; 0 \\
\Rightarrow \quad 2KC^{\boldsymbol{v}} \;&=\; 2H\cdot(1-KH)\,C^p \\
K\cdot\left(C^{\boldsymbol{v}} + H^2C^p\right) \;&=\; C^pH \\
K \;&=\; \frac{C^pH}{C^{\boldsymbol{v}} + H^2C^p} \\
&=\; \underline{\underline{\frac{H}{\frac{C^{\boldsymbol{v}}}{C^p} + H^2}}}
\end{aligned}$$

$$\text{vgl.} \quad \mathbf{K}_k = \mathbf{C}_k^p\mathbf{H}_k^\top \left(\mathbf{C}_k^{\boldsymbol{v}} + \mathbf{H}_k\mathbf{C}_k^p\mathbf{H}_k^\top\right)^{-1}$$

e) Berechnen Sie die sich daraus ergebende optimale (minimal mögliche) Filter-Kovarianz C^e.

$$
\begin{aligned}
C^e &= (1-KH)^2\, C^p \;+\; K^2 C^{\boldsymbol{v}} \\
&= K^2 \cdot \left(C^{\boldsymbol{v}} \;+\; H^2 C^p\right) \;-\; K \cdot 2C^P H \;+\; C^p \qquad \Bigg|\; K = \frac{C^p H}{C^{\boldsymbol{v}} + H^2 C^p} \\
&= C^p \;-\; \frac{(C^p H)^2}{C^{\boldsymbol{v}} + H^2 C^p} \\
&= \frac{C^p C^{\boldsymbol{v}}}{C^{\boldsymbol{v}} + H^2 C^p} \\
&= \underline{\underline{\frac{1}{\frac{1}{C^p} + \frac{H^2}{C^{\boldsymbol{v}}}}}}
\end{aligned}
$$

vgl. $\mathbf{C}_k^e = \mathbf{C}_k^p - \mathbf{C}_k^p \mathbf{H}_k^\top \left(\mathbf{C}_k^{\boldsymbol{v}} + \mathbf{H}_k \mathbf{C}_k^p \mathbf{H}_k^\top\right)^{-1} \mathbf{H}_k \mathbf{C}_k^p$

f) Wir betrachten jetzt den Fall (sehr) unzuverlässiger Messdaten, $C^{\boldsymbol{v}} \gg C^p$. Vereinfachen Sie K unter dieser Bedingung und erläutern Sie, wie das Filter sich verhält. [1]

$$
\begin{aligned}
K &= \frac{H}{\frac{C^{\boldsymbol{v}}}{C^p} + H^2} \\
\lim_{\frac{C^{\boldsymbol{v}}}{C^p} \to \infty} K &= \underline{\underline{0}}
\end{aligned}
$$

Der Verstärkungsfaktor K strebt gegen 0. Damit vereinfacht sich die Berechnung des Filter-Schätzwertes $\boldsymbol{x}^e$ zu

$$
\begin{aligned}
\boldsymbol{x}^e &= \boldsymbol{x}^p \;+\; K \cdot (\boldsymbol{y} - H\boldsymbol{x}^p) \qquad |\; K = 0 \\
\boldsymbol{x}^e_{K=0} &= \underline{\underline{\boldsymbol{x}^p}} \;.
\end{aligned}
$$

Für die Filter-Kovarianz gilt

$$
\begin{aligned}
C^e &= (1-KH)^2\, C^p \;+\; K^2 C^{\boldsymbol{v}} \qquad |\; K = 0 \\
C^e|_{K=0} &= \underline{\underline{C^p}} \;.
\end{aligned}
$$

Wenn die Messung absolut unzuverlässig ist, reicht das Filter also einfach den Prädiktions-Schätzwert unverändert durch. Der Zustand wird also lediglich mit Hilfe des internen Systemmodells geschätzt. Dementsprechend verbessert sich auch die Filter-Kovarianz C^e nicht mehr im Vergleich zur Prädiktions-Kovarianz C^p.

g) Nun betrachten wir den umgekehrten Fall – starkes Systemrauschen bzw. ungenaues Systemmodell, $C^p \gg C^{\boldsymbol{v}}$. Berechnen Sie K für diesen Fall und erläutern Sie das Verhalten des Filters. [1]

$$
\begin{aligned}
K &= \frac{H}{\frac{C^{\boldsymbol{v}}}{C^p} + H^2} \\
\lim_{\frac{C^{\boldsymbol{v}}}{C^p} \to 0} K &= \underline{\underline{1/H}}
\end{aligned}
$$

Die Berechnung des Filter-Schätzwertes mit diesem K lautet nun

$$
\begin{aligned}
\boldsymbol{x}^e &= \boldsymbol{x}^p \;+\; K \cdot (\boldsymbol{y} - H\boldsymbol{x}^p) \qquad |\; K = 1/H \\
\boldsymbol{x}^e|_{K=1/H} &= \boldsymbol{x}^p \;+\; \frac{1}{H}\boldsymbol{y} \;-\; \frac{1}{H} H \boldsymbol{x}^p \\
&= \underline{\underline{\frac{1}{H}\boldsymbol{y}}} \;.
\end{aligned}
$$

Durch das große Systemrauschen kann das Kalman-Filter den Systemzustand nicht zuverlässig prädizieren. Es schätzt den Zustand dann jedes Mal neu aus der Messgröße, ohne Vorwissen zu verwenden. Das entspricht der Vorgehensweise bei einem statischen System.

$$
\begin{aligned}
C^e &= (1-KH)^2 C^p + K^2 C^{\boldsymbol{v}} \quad | \, K = 1/H \\
C^e|_{K=1/H} &= \underline{\underline{\frac{1}{H^2} C^{\boldsymbol{v}}}} \ .
\end{aligned}
$$

Dementsprechend wird die Zuverlässigkeit des Ergebnisses durch die Zuverlässigkeit der Messung und die Steilheit der Kennlinie des Messumformers bestimmt.

2) Korrelierte Variablen

Betrachten Sie zwei korrelierte Zufallsvariablen $\boldsymbol{x}_1$ und $\boldsymbol{x}_2$. Das Wissen über $\boldsymbol{x}_1$ soll durch eine Messung von $\boldsymbol{x}_2$ verbessert werden. Die Messung von $\boldsymbol{x}_2$ ist verrauscht.

Gegeben ist

$$
\underline{\boldsymbol{x}} = \begin{bmatrix} \boldsymbol{x}_1 \\ \boldsymbol{x}_2 \end{bmatrix}
$$

und

$$
\boldsymbol{y} = \boldsymbol{x}_2 + \boldsymbol{v} \ ,
$$

mit $\boldsymbol{v} \sim \mathrm{N}(v - 0, \ \sigma_{\boldsymbol{v}}^2)$. Die Unsicherheit der Schätzung von $\underline{\boldsymbol{x}}$ vor der Messung wird durch die Kovarianzmatrix

$$
\mathbf{C}^p = \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{bmatrix} \ ,
$$

beschrieben. σ_1^2 ist die Varianz der Schätzung von $\boldsymbol{x}_1$ und σ_2^2 die Varianz der Schätzung von $\boldsymbol{x}_2$. Die entsprechende Kovarianz wird durch σ_{12} bezeichnet.

a) Wie lautet die Messmatrix $\mathbf{H}$?

$$
\mathbf{H} = \underline{\underline{\begin{bmatrix} 0 & 1 \end{bmatrix}}} \ .
$$

b) Berechnen Sie die Gewichtungsmatrix $\mathbf{K}$, indem Sie $\mathbf{C}^p$ und $\mathbf{H}$ in den Ausdruck für $\mathbf{K}$ aus der Vorlesung einsetzen.

$$
\begin{aligned}
\mathbf{K} &= \mathbf{C}^p \mathbf{H}^\top \left(\mathbf{C}^{\boldsymbol{v}} + \mathbf{H} \mathbf{C}^p \mathbf{H}^\top \right)^{-1} \\
\mathbf{C}^p \mathbf{H}^\top &= \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} \sigma_{12} \\ \sigma_2^2 \end{bmatrix} \\
\mathbf{H} \mathbf{C}^p \mathbf{H}^\top &= \begin{bmatrix} 0 & 1 \end{bmatrix} \begin{bmatrix} \sigma_{12} \\ \sigma_2^2 \end{bmatrix} = \sigma_2^2 \\
\mathbf{C}^{\boldsymbol{v}} + \mathbf{H} \mathbf{C}^p \mathbf{H}^\top &= \sigma_{\boldsymbol{v}}^2 + \sigma_2^2 \\
\mathbf{C}^p \mathbf{H}^\top \left(\mathbf{C}^{\boldsymbol{v}} + \mathbf{H} \mathbf{C}^p \mathbf{H}^\top \right)^{-1} &= \underline{\underline{\begin{bmatrix} \sigma_{12} \\ \sigma_2^2 \end{bmatrix} \cdot \left(\sigma_{\boldsymbol{v}}^2 + \sigma_2^2 \right)^{-1}}} = \mathbf{K}
\end{aligned}
$$

c) Berechnen Sie ebenso die Kovarianzmatrix $\mathbf{C}^e$ nach der Messung.

$$
\begin{aligned}
\mathbf{C}^e &= \mathbf{C}^p - \mathbf{C}^p \mathbf{H}^\top \left(\mathbf{C}^{\boldsymbol{v}} + \mathbf{H} \mathbf{C}^p \mathbf{H}^\top \right)^{-1} \mathbf{H} \mathbf{C}^p \\
&= \mathbf{C}^p - \mathbf{K} \mathbf{H} \mathbf{C}^p \\
&= \left(\mathbf{I} - \mathbf{K} \mathbf{H} \right) \mathbf{C}^p \\
&= \left(\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} \sigma_{12} \\ \sigma_2^2 \end{bmatrix} \cdot \left(\sigma_{\boldsymbol{v}}^2 + \sigma_2^2 \right)^{-1} \begin{bmatrix} 0 & 1 \end{bmatrix} \right) \mathbf{C}^p \\
&= \left(\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} 0 & \sigma_{12} \\ 0 & \sigma_2^2 \end{bmatrix} \cdot \left(\sigma_{\boldsymbol{v}}^2 + \sigma_2^2 \right)^{-1} \right) \mathbf{C}^p \\
&= \begin{bmatrix} 1 & -\sigma_{12} \left(\sigma_{\boldsymbol{v}}^2 + \sigma_2^2 \right)^{-1} \\ 0 & 1 - \sigma_2^2 \left(\sigma_{\boldsymbol{v}}^2 + \sigma_2^2 \right)^{-1} \end{bmatrix} \mathbf{C}^p
\end{aligned}
$$

$$
\begin{aligned}
&= \begin{bmatrix} 1 & -\sigma_{12}\left(\sigma_{\boldsymbol{v}}^2+\sigma_2^2\right)^{-1} \\ 0 & 1-\sigma_2^2\left(\sigma_{\boldsymbol{v}}^2+\sigma_2^2\right)^{-1} \end{bmatrix} \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{bmatrix} \\
&= \begin{bmatrix} \sigma_1^2 - \sigma_{12}\left(\sigma_{\boldsymbol{v}}^2+\sigma_2^2\right)^{-1}\sigma_{12} & \sigma_{12} - \sigma_{12}\left(\sigma_{\boldsymbol{v}}^2+\sigma_2^2\right)^{-1}\sigma_2^2 \\ \left(1-\sigma_2^2\left(\sigma_{\boldsymbol{v}}^2+\sigma_2^2\right)^{-1}\right)\sigma_{12} & \left(1-\sigma_2^2\left(\sigma_{\boldsymbol{v}}^2+\sigma_2^2\right)^{-1}\right)\sigma_2^2 \end{bmatrix} \\
&= \underline{\underline{\begin{bmatrix} \sigma_1^2 - \sigma_{12}^2\left(\sigma_{\boldsymbol{v}}^2+\sigma_2^2\right)^{-1} & \sigma_{12}\left(1-\sigma_2^2\left(\sigma_{\boldsymbol{v}}^2+\sigma_2^2\right)^{-1}\right) \\ \sigma_{12}\left(1-\sigma_2^2\left(\sigma_{\boldsymbol{v}}^2+\sigma_2^2\right)^{-1}\right) & \sigma_2^2 - \sigma_2^4\left(\sigma_{\boldsymbol{v}}^2+\sigma_2^2\right)^{-1} \end{bmatrix}}}
\end{aligned}
$$

d) Wie sind $\boldsymbol{x}_1$ und $\boldsymbol{x}_2$ korreliert? Geben Sie die Definition des Korrelationskoeffizienten ρ an.

$$\rho = \frac{\sigma_{12}}{\underline{\underline{\sigma_1\sigma_2}}} \ .$$

e) Bestimmen Sie $\mathbf{C}^e$ für den Fall $\sigma_{\boldsymbol{v}} = 0$.

$\left(\sigma_{\boldsymbol{v}}^2+\sigma_2^2\right)^{-1}$ vereinfacht sich dann zu σ_2^{-2}.

$$\mathbf{C}^e\big|_{\sigma_{\boldsymbol{v}}=0} = \begin{bmatrix} \sigma_1^2 - \sigma_{12}^2\sigma_2^{-2} & 0 \\ 0 & 0 \end{bmatrix} = \underline{\underline{\begin{bmatrix} \sigma_1^2\left(1-\rho^2\right) & 0 \\ 0 & 0 \end{bmatrix}}}$$

Die posteriore Unsicherheit in $\boldsymbol{x}_2$ (rechts unten in $\mathbf{C}^e$) ist somit Null, weil wir eine exakte Messung davon haben. Die posteriore Unsicherheit in $\boldsymbol{x}_1$ (links oben in $\mathbf{C}^e$) dagegen ist abhängig von der Korrelation. Sie bewegt sich von 0 bei $\rho = \pm 1$ bis zu der Unsicherheit σ_1^2, die schon in der Prioren vorhanden war, für $\rho = 0$.

f) Bestimmen Sie $\mathbf{C}^e$ für den Fall $\rho = 0$.

Hier ist $\sigma_{12} = 0$. Eingesetzt in $\mathbf{C}^e$ ergibt sich

$$
\begin{aligned}
\mathbf{C}^e\big|_{\sigma_{12}=0} &= \begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 - \sigma_2^4\left(\sigma_{\boldsymbol{v}}^2+\sigma_2^2\right)^{-1} \end{bmatrix} \\
&= \underline{\underline{\begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2\left(1-\left(1+\frac{\sigma_{\boldsymbol{v}}^2}{\sigma_2^2}\right)^{-1}\right) \end{bmatrix}}}
\end{aligned}
$$

Die Unsicherheit in $\boldsymbol{x}_1$ wird unverändert von der Prioren übernommen, da keine Information für $\boldsymbol{x}_1$ aus der Messung von $\boldsymbol{x}_2$ abgeleitet werden können. Die Unsicherheit in $\boldsymbol{x}_2$ ist gleich wie in $\mathbf{C}^e$.

g) Bestimmen Sie $\mathbf{C}^e$ für den Fall $\rho = \pm 1$.

Um $\rho = \frac{\sigma_{12}}{\sigma_1\sigma_2} = \pm 1$ zu erzwingen, setzen wir $\sigma_{12} = \pm\sigma_1\sigma_2$ in $\mathbf{C}^e$ ein.

$$
\begin{aligned}
\mathbf{C}^e\big|_{\rho=\pm 1} &= \begin{bmatrix} \sigma_1^2 - \sigma_1^2\sigma_2^2\left(\sigma_{\boldsymbol{v}}^2+\sigma_2^2\right)^{-1} & \pm\sigma_1\sigma_2\left(1-\sigma_2^2\left(\sigma_{\boldsymbol{v}}^2+\sigma_2^2\right)^{-1}\right) \\ \pm\sigma_1\sigma_2\left(1-\sigma_2^2\left(\sigma_{\boldsymbol{v}}^2+\sigma_2^2\right)^{-1}\right) & \sigma_2^2 - \sigma_2^4\left(\sigma_{\boldsymbol{v}}^2+\sigma_2^2\right)^{-1} \end{bmatrix} \\
&= \underline{\underline{\begin{bmatrix} \sigma_1^2\left(1+\frac{\sigma_2^2}{\sigma_{\boldsymbol{v}}^2}\right)^{-1} & \pm\sigma_1\sigma_2\left(1-\left(1+\frac{\sigma_{\boldsymbol{v}}^2}{\sigma_2^2}\right)^{-1}\right) \\ \pm\sigma_1\sigma_2\left(1-\left(1+\frac{\sigma_{\boldsymbol{v}}^2}{\sigma_2^2}\right)^{-1}\right) & \sigma_2^2\left(1-\left(1+\frac{\sigma_{\boldsymbol{v}}^2}{\sigma_2^2}\right)^{-1}\right) \end{bmatrix}}}
\end{aligned}
$$

Die Unsicherheit in $\boldsymbol{x}_1$ hängt hier vom Verhältnis der prioren Unsicherheit in $\boldsymbol{x}_2$ zur Unsicherheit des Messrauschens $\boldsymbol{v}$ ab. Die Unsicherheit in $\boldsymbol{x}_2$ ist auch hier gleich wie in $\mathbf{C}^e$.

3) Anwendung: Stationäres Kalman Filter

Diese Aufgabe ist eher zur Information gedacht und stellt einen Zusammenhang zum LQ-Regler (linearer quadratischer Regler, auch Riccati-Regler) her.

Gegeben sei ein lineares System

$$\underline{\boldsymbol{x}}_{k+1} = \mathbf{A}\underline{\boldsymbol{x}}_k + \underline{\boldsymbol{w}}_k \ ,$$
$$\underline{\boldsymbol{y}}_k = \mathbf{C}\underline{\boldsymbol{x}}_k + \underline{\boldsymbol{v}}_k \ ,$$

mit stationären Gauss-Rauschprozessen $\underline{\boldsymbol{w}}_k, \underline{\boldsymbol{v}}_k$ mit $\mathrm{E}\{\underline{\boldsymbol{w}}_k \underline{\boldsymbol{w}}_k^T\} = \mathbf{W}$, $\mathrm{E}\{\underline{\boldsymbol{v}}_k \underline{\boldsymbol{v}}_k^T\} = \mathbf{V}$.

a) Warum könnte es Sinn machen, in diesem Fall ein stationäres Kalman-Filter zu verwenden, d.h. ein Kalman-Filter mit einem festen Kalman-Gain $\mathbf{K}$, der vorab berechnet wird?

Das System ist zeitinvariant, die Rauschprozesse sind stationär und die Messungen sind zu jedem Zeitpunkt verfügbar. Typischerweise verwendet man ja eine „große“ Kovarianzmatrix für den initialen Zustand, jedenfalls wenn der initiale Zustandswert „aus der Luft gegriffen“ wird, z.B. als Nullvektor gesetzt wird. Aber egal was für eine Kovarianzmatrix man auch verwendet, unter den genannten Bedingungen wird sich der Kalman-Gain immer einer ganz bestimmten Matrix annähern. Daher liegt es nahe, diesen Gain einfach von Anfang an zu verwenden und dadurch Rechenzeit zu sparen. Zu Beginn, wo der „richtige“ Gain anders gewesen wäre, wird die Schätzung natürlich schlechter sein, aber nach einiger Zeit sind beide Kalman-Gains praktisch gleich, und damit wird die Schätzung nach einiger Zeit auch genauso gut.

b) Leiten Sie das stationäre Kalman Filter her (für Teilnehmer mit Regelungstechnik-Hintergrund).

Stationärer Fall:

$$\mathbf{K} = \mathbf{X}\mathbf{H}^T \left(\mathbf{C}^y + \mathbf{H}\mathbf{X}\mathbf{H}^T\right)^{-1} \ ,$$

wobei $\mathbf{X}$ die Lösung von

$$\mathbf{X} = \mathbf{A}\mathbf{X}\mathbf{A}^T - \mathbf{A}\mathbf{X}\mathbf{H}^T \left(\mathbf{C}^y + \mathbf{H}\mathbf{X}\mathbf{H}^T\right)^{-1} \mathbf{H}\mathbf{X}\mathbf{A}^T$$

ist.

4) Zeitkomplexität Kalmanfilter

Die Zeitkomplexität für Matrixmultiplikation und Matrixinversion sei gegeben durch

$$\begin{aligned} \mathbf{A}\cdot\mathbf{B} &\rightarrow \mathcal{O}(M\cdot N\cdot P) & \mathbf{A}&\in\mathbb{R}^{M\times N}, \quad \mathbf{B}\in\mathbb{R}^{N\times P}\\ \mathbf{A}+\mathbf{A} &\rightarrow \mathcal{O}(M\cdot N) & &\\ \mathbf{C}^{-1} &\rightarrow \mathcal{O}(N^3) & \mathbf{C}&\in\mathbb{R}^{N\times N} \end{aligned}$$

Der Zustandsraum sei N-dimensional, der System-Input P-dimensional und die Messung M-dimensional.

a) Bestimmen Sie die Zeitkomplexität des Kalman-Prädiktionsschrittes. Es gelte $P \leq N$.

Zunächst der Prädiktions-Schätzwert $\hat{\underline{x}}_k^{\mathrm{p}}$.

$$\mathbf{A}\in\mathbb{R}^{N\times N}, \quad \hat{\underline{x}}_{k-1}^{\mathrm{e}}\in\mathbb{R}^{N\times 1}, \quad \mathbf{B}\in\mathbb{R}^{N\times P}, \quad \hat{\underline{u}}_{k-1}\in\mathbb{R}^{P\times 1}$$
$$\hat{\underline{x}}_k^{\mathrm{p}} = \underbrace{\mathbf{A}\cdot\hat{\underline{x}}_{k-1}^{\mathrm{e}}}_{N\times 1} + \underbrace{\mathbf{B}\cdot\hat{\underline{u}}_{k-1}}_{N\times 1}$$
$$\begin{aligned} \text{Complexity} &= \mathcal{O}(N\cdot N\cdot 1 \quad + \quad N\cdot P\cdot 1 \quad + \quad N\cdot 1)\\ &= \mathcal{O}(N^2 + NP + N) \qquad | \, P\leq N\\ &= \underline{\underline{\mathcal{O}(N^2)}} \end{aligned}$$

Nun zur Prädiktions-Kovarianz $\mathbf{C}_k^{\mathrm{p}}$. Matrixmultiplikation ist assoziativ, es können also beliebig Klammern gesetzt und Zwischenergebnisse berechnet werden, wobei sich die Laufzeiten addieren.

$$\mathbf{A}\in\mathbb{R}^{N\times N}, \quad \mathbf{C}^{\mathrm{e}}\in\mathbb{R}^{N\times N}, \quad \mathbf{B}\in\mathbb{R}^{N\times P}, \quad \mathbf{C}^w\in\mathbb{R}^{P\times P}$$
$$\begin{aligned} \mathbf{C}_k^{\mathrm{p}} &= \mathbf{A}\cdot\mathbf{C}_{k-1}^{\mathrm{e}}\cdot\mathbf{A}^\top + \mathbf{B}\cdot\mathbf{C}_{k-1}^{w}\cdot\mathbf{B}^\top\\ &= \underbrace{(\mathbf{A}\cdot\mathbf{C}_{k-1}^{\mathrm{e}})}_{N\times N}\cdot\mathbf{A}^\top + \underbrace{(\mathbf{B}\cdot\mathbf{C}_{k-1}^{w})}_{N\times P}\cdot\mathbf{B}^\top \end{aligned}$$

$$
\begin{aligned}
\text{Complexity} &= \mathcal{O}(N\cdot N\cdot N + N\cdot N\cdot N \quad + \quad N\cdot P\cdot P + N\cdot P\cdot N \quad + \quad N\cdot N) \\
&= \mathcal{O}(2N^3 + N^2 + N^2P + NP^2) \qquad |P \le N \\
&= \underline{\underline{\mathcal{O}(N^3)}}
\end{aligned}
$$

b) Bestimmen Sie die Zeitkomplexität des Kalman-Gains. Vereinfachen Sie auf die Spezialfälle $M = N$ sowie $M = 1$.

Zunächst der Kalman Gain $\mathbf{K}$.

$$
\begin{aligned}
&\mathbf{C}^\mathrm{p} \in \mathbb{R}^{N\times N}, \qquad \mathbf{H} \in \mathbb{R}^{M\times N}, \qquad \mathbf{C}^v \in \mathbb{R}^{M\times M} \\
&\mathbf{K} = \mathbf{C}^\mathrm{p}\cdot\mathbf{H}^\top\cdot\left(\mathbf{C}^v + \mathbf{H}\cdot\mathbf{C}^\mathrm{p}\cdot\mathbf{H}^\top\right)^{-1} \\
&\quad = \underbrace{\left(\mathbf{C}^\mathrm{p}\cdot\mathbf{H}^\top\right)}_{N\times M}\cdot\underbrace{\left(\mathbf{C}^v + \underbrace{\left(\mathbf{H}\cdot\mathbf{C}^\mathrm{p}\right)}_{M\times N}\cdot\mathbf{H}^\top\right)^{-1}}_{M\times M}
\end{aligned}
$$

$$
\begin{aligned}
\text{Complexity} &= \mathcal{O}(N\cdot N\cdot M + M\cdot N\cdot N + M\cdot N\cdot M + M\cdot M + M^3 + N\cdot M\cdot M) \\
&= \mathcal{O}(2N^2M + 2M^2N + M^2 + M^3) \qquad |\; M \le N \\
&= \mathcal{O}(N^2M) \\
\text{Complexity}_{M=N} &= \underline{\underline{\mathcal{O}(N^3)}} \\
\text{Complexity}_{M=1} &= \underline{\underline{\mathcal{O}(N^2)}}
\end{aligned}
$$

c) Bestimmen Sie die Zeitkomplexität des Kalman-Filterschritts bei gegebenem Kalman-Gain. Es gelte $M \le N$.

Der Filter-Schätzwert $\underline{\hat{x}}^\mathrm{e}$:

$$
\begin{aligned}
&\underline{\hat{x}}^\mathrm{p} \in \mathbb{R}^{N\times 1}, \qquad \mathbf{K} \in \mathbb{R}^{N\times M}, \qquad \underline{\hat{y}} \in \mathbb{R}^{M\times 1}, \qquad \mathbf{H} \in \mathbb{R}^{M\times N} \\
&\underline{\hat{x}}^\mathrm{e} = \underline{\hat{x}}^\mathrm{p} + \mathbf{K}\cdot\underbrace{\left(\underline{\hat{y}} - \mathbf{H}\cdot\underline{\hat{x}}^\mathrm{p}\right)}_{M\times 1}
\end{aligned}
$$

$$
\begin{aligned}
\text{Complexity} &= \mathcal{O}(M\cdot N\cdot 1 + M\cdot 1 + N\cdot M\cdot 1 + N\cdot 1) \\
&= \underline{\underline{\mathcal{O}(MN)}}
\end{aligned}
$$

Die Filter-Kovarianz $\mathbf{C}^\mathrm{e}$:

$$
\begin{aligned}
&\mathbf{K} \in \mathbb{R}^{N\times M}, \qquad \mathbf{H} \in \mathbb{R}^{M\times N}, \qquad \mathbf{C}^\mathrm{p} \in \mathbb{R}^{N\times N} \\
&\mathbf{C}^\mathrm{e} = \underbrace{\left(\mathbf{I} - \mathbf{K}\cdot\mathbf{H}\right)}_{N\times N}\cdot\mathbf{C}^\mathrm{p}
\end{aligned}
$$

$$
\begin{aligned}
\text{Complexity} &= \mathcal{O}(N\cdot M\cdot N + N\cdot N + N\cdot N\cdot N) \\
&= \mathcal{O}(N^3 + N^2M + N^2) \qquad |\; M \le N \\
&= \mathcal{O}(N^3)
\end{aligned}
$$

d) Sie sollen nun für ein weltweites Wettermodell Zustandsschätzung betreiben. Der Zustandsraum besteht aus Temperatur, Luftdruck und Windgeschwindigkeit – und das jeweils für jeden Punkt auf einem weltumspannenden engmaschigen Grid. Von Zeit zu Zeit liefern einzelne Wetterstationen Messwerte. Vor welchem Problem stehen Sie?

N ist sehr groß, daher ist die Berechnung etwa der Filter-Kovarianz mit Zeitkomplexität $\mathcal{O}(N^3)$ sehr zeitaufwendig bis unmöglich. Zwar ist M klein, aber das bringt erst einmal leider keinen Vorteil.

e) Teilen Sie die fünf Berechnungsschritte des Kalman-Filters (Kalman-Gain; Schätzwert und Kovarianz jeweils im Prädiktions- und Filterschritt) danach ein, wie problematisch die Komplexität bei dem Wettermodell ist. Was fällt Ihnen dabei auf?

- *Unproblematisch* ist die Berechnung des Filter-Schätzwertes (bei gegebenem Kalman-Gain) mit einer Komplexität von $M \cdot N$.
- *Viel Zeit* benötigt die Berechnung des Prädiktions-Schätzwertes sowie des Kalman-Gains mit jeweils N^2.
- *Sehr viel Zeit bzw. unmöglich* ist die Berechnung der Prädiktions-Kovarianz sowie der Filter-Kovarianz mit jeweils N^3.

Auffällig ist, dass gerade der eigentliche Kern der Kalman-Fusion, die Berechnung des Filter-Schätzwertes, die geringste Zeitkomplexität hat.

f) Sie geben nicht auf und wollen das Problem beherrschbar machen. Statt „wie gehabt" mit N-dimensionalen Schätzwert-Vektoren plus $N \times N$ – Kovarianz-Matrizen

$$\hat{\underline{x}}_k^{\mathrm{p}} \in \mathbb{R}^{N\times 1} \qquad \hat{\underline{x}}_k^{\mathrm{e}} \in \mathbb{R}^{N\times 1}$$
$$\mathbf{C}_k^{\mathrm{p}} \in \mathbb{R}^{N\times N} \qquad \mathbf{C}_k^{\mathrm{e}} \in \mathbb{R}^{N\times N}$$

repräsentieren Sie den unsicheren Schätzwert nun per „Streuungsbreite" einer Punktwolke. Als „unsicheren Zustand" verwenden Sie also eine große Anzahl N-dimensionaler Vektoren als Samples

$$\underline{x}_{k,i}^{\mathrm{p}} \in \mathbb{R}^{N\times 1} \qquad \underline{x}_{k,j}^{\mathrm{e}} \in \mathbb{R}^{N\times 1}$$
$$i \in \{1, 2, \ldots, L^{\mathrm{p}}\} \qquad j \in \{1, 2, \ldots, L^{\mathrm{e}}\} \quad ,$$

wobei die Samples auch als Spalten einer Matrix aufgefasst werden können, welche all diese Samples kompakt zusammenfasst

$$\mathcal{X}_k^{\mathrm{p}} \in \mathbb{R}^{N\times L^{\mathrm{p}}} \qquad \mathcal{X}_k^{\mathrm{e}} \in \mathbb{R}^{N\times L^{\mathrm{e}}} \quad .$$

Versuchen Sie, alle Berechnungsschritte rein mit Hilfe dieser Samples umzusetzen.
Hinweis: $\mathbf{K} = \mathbf{C}_{xy} \cdot \mathbf{C}_{yy}^{-1}$

Ensemble Kalmanfilter!

Prädiktionsschritt Lineares Systemmodell:

$$\begin{aligned}
\underline{x}_{k,i}^{\mathrm{p}} &= \mathbf{A}_{k-1} \cdot \underline{x}_{k-1,i}^{\mathrm{e}} + \mathbf{B}_{k-1} \cdot (\underline{u}_{k-1} + \underline{w}_{k-1,i}) \\
\mathcal{X}_k^{\mathrm{p}} &= \mathbf{A}_{k-1} \cdot \mathcal{X}_{k-1}^{\mathrm{e}} + \mathbf{B}_{k-1} \cdot (\underline{u}_{k-1} + \mathcal{W}_{k-1}) \\
\text{Complexity} &= \mathcal{O}(N^2 L + NPL + NL) \\
&= \mathcal{O}(N^2 L) \qquad |\ \text{UKF:} \quad L = 2N \\
&= \mathcal{O}(N^3)
\end{aligned}$$

Nichtlineares Systemmodell:

$$\underline{x}_{k,i}^{\mathrm{p}} = \underline{a}_{k-1}(\underline{x}_{k-1,i}^{\mathrm{e}},\ \underline{u}_{k-1},\ w_{k-1,i})$$

Wichtig ist, für jedes $i = 1, 2, ..., L$ eine andere Realisierung des Rauschens $\underline{w}_{k-1,i}$ einzusetzen. Oder mehrere Rausch-Samples verwenden, dann steigt allerdings die Anzahl der Samples L (kartesisches Produkt). Unterschiedliche Methoden denkbar, siehe Vorlesung. Durch die Vorwärts-Berechnung der Samples ist schon alles getan, da die Unsicherheit / Streuungsbreite ja in der Ausdehnung der Sample-Wolke abgebildet ist. Die Komplexität ist zwar immer noch sehr groß und kubisch in der Dimension, immerhin lässt sich die Berechnung hier aber sehr gut parallelisieren.

Filterschritt Zunächst „prädizierte" Mess-Samples berechnen (erwartete Messung basierend auf Prädiktion und Messmodell, jeweils inkl. Unsicherheit), lineare Messgleichung:

$$\begin{aligned}
\underline{y}_{k,i} &= \mathbf{H}_k \cdot \underline{x}_{k,i}^{\mathrm{p}} + \underline{v}_{k,i} \\
\mathcal{Y}_k &= \mathbf{H}_k \cdot \mathcal{X}_k^{\mathrm{p}} + \mathcal{V}_k \\
\text{Complexity} &= \mathcal{O}(MNL + NL)
\end{aligned}$$

$$= \mathcal{O}(NML)$$

Für nichtlineare Messgleichung entsprechend:

$$\begin{aligned}
\underline{y}_{k,i} &= \underline{h}(\underline{x}^{\mathrm{p}}_{k,i},\ \underline{v}_{k,i}) \\
\mathcal{Y}_k &= \begin{bmatrix} \underline{y}_{k,1} & \underline{y}_{k,2} & \cdots & \underline{y}_{k,L} \end{bmatrix}
\end{aligned}$$

Filterschritt, mit der tatsächlichen Messung $\underline{\hat{y}}_k$

$$\begin{aligned}
\underline{x}^{\mathrm{e}}_{k,i} &= \underline{x}^{\mathrm{p}}_{k,i} + \mathbf{K} \cdot \left(\underline{\hat{y}}_k - \underline{y}_{k,i}\right) \\
\mathcal{X}^{\mathrm{e}}_k &= \mathcal{X}^{\mathrm{p}}_k + \mathbf{K} \cdot \left(\underline{\hat{y}}_k \cdot \underline{1}^\top - \mathcal{Y}_k\right) \\
\text{Complexity} &= \mathcal{O}(ML + NML + NL) \\
&= \mathcal{O}(NML)
\end{aligned}$$

Berechnung Kalman Gain:

$$\begin{aligned}
\mathbf{C}_{xy} &= \frac{1}{L} \sum_{i=1}^{L} \underline{x}^{\mathrm{p}}_{k,i} \cdot \underline{y}^\top_{k,i} \\
&= \frac{1}{L}\, \mathcal{X}^{\mathrm{p}}_k \cdot \mathcal{Y}^\top_k \\
\mathbf{C}_{yy} &= \frac{1}{L} \sum_{i=1}^{L} \underline{y}_{k,i} \cdot \underline{y}^\top_{k,i} \\
&= \frac{1}{L}\, \mathcal{Y}_k \cdot \mathcal{Y}^\top_k \\
\mathbf{K} &= \mathbf{C}_{xy} \cdot \mathbf{C}^{-1}_{yy} \\
\text{Complexity}_{\mathbf{C}_{yy}} &= \mathcal{O}(M^2 L) \\
\text{Complexity}_{\mathbf{C}_{xy}} &= \mathcal{O}(NML) \\
\text{Complexity}_{\mathbf{K}} &= \mathcal{O}(M^3 + NM^2)
\end{aligned}$$

Zeitkomplexität des Filterschritts ingesamt:

$$\begin{aligned}
\text{Complexity} &= \mathcal{O}(NML + NML + M^2L + NML + M^3 + NM^2) \\
&= \mathcal{O}(3NML + M^2L + M^3 + NM^2) \qquad \mid \text{UKF:} \quad L = 2N \\
&= \mathcal{O}(6N^2M + 3M^2N + M^3) \qquad \mid M \leq N \\
&= \mathcal{O}(N^2M) \\
\text{Complexity}_{M=1} &= \mathcal{O}(NL) \qquad \mid \text{UKF:} \quad L = 2N \\
&= \mathcal{O}(N^2)
\end{aligned}$$

Hier ist die Komplexität überraschenderweise kleiner als bei der Prädiktion. Und für den Fall $M < N$ entfällt die $\mathcal{O}(N^3)$ schwere Berechnung der posterioren Kovarianzmatrix.

g) Vergleichen Sie die Komplexität, wenn man eine N-dimensionale Messung (kompletter Zustandsraum wird gemessen, $M = N$) direkt in einem Filterschritt einbringt, mit der Komplexität wenn man diese Messung in N eindimensionale Messungen aufteilt und diese in N direkt nacheinander ausgeführten „Mini-Filterschritten“ einspielt. Was muss überhaupt gelten, damit beide Methoden äquivalent sind?

Die Messungen bzw. das Messrauschen muss unkorreliert sein, $\mathbf{C}^v_k$ muss also eine Diagonalmatrix sein. Dann können wir die Messungen einzeln einspielen und führen M Filterschritte der Komplexität von jeweils $\mathcal{O}(NL)$ aus, insgesamt erhalten wir also $\mathcal{O}(NML)$, bzw. bei $M = N$ haben wir $\mathcal{O}(N^2L)$. Direkt die N-dimensionale Messung zu verarbeiten hätte die Komplexität $\mathcal{O}(N^2L + N^3/2)$. Für $L = 2N$ ist das EnKF jedoch auch wieder kubisch im Aufwand, allerdings besser parallelisierbar.

Gaußdichte

5) Normierungsfaktor der Normalverteilung

Zeigen Sie, dass gilt

$$\int_{-\infty}^{\infty} e^{-x^2} \, \mathrm{d}x \;=\; \sqrt{\pi} \;.$$

Hinweis: Quadrieren Sie zunächst die Gleichung und führen Sie für die beiden miteinander multiplizierten Integrale unterschiedliche Variablen x, y ein. Dann können Sie das Integral in Polarkoordinaten R, φ lösen.

$$\begin{aligned}
\left(\int_{-\infty}^{\infty} e^{-x^2} \, \mathrm{d}x\right)^2 &= \left(\int_{-\infty}^{\infty} e^{-x^2} \, \mathrm{d}x\right) \cdot \left(\int_{-\infty}^{\infty} e^{-y^2} \, \mathrm{d}y\right) \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-x^2} e^{-y^2} \, \mathrm{d}x \, \mathrm{d}y \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-\left(x^2+y^2\right)} \, \mathrm{d}x \, \mathrm{d}y \\
& \qquad \big| \; x = R\cos(\varphi); \quad y = R\sin(\varphi) \; ; \quad \mathrm{d}x \, \mathrm{d}y = R \, \mathrm{d}\varphi \, \mathrm{d}R \\
&= \int_{0}^{\infty} \int_{0}^{2\pi} e^{-\left([R\cos(\varphi)]^2 + [R\sin(\varphi)]^2\right)} R \, \mathrm{d}\varphi \, \mathrm{d}R \\
&= 2\pi \int_{0}^{\infty} R \cdot e^{-R^2\left(\cos^2(\varphi)+\sin^2(\varphi)\right)} \, \mathrm{d}R \quad \big| \; \cos^2(\varphi) + \sin^2(\varphi) = 1 \\
&= \pi \int_{0}^{\infty} 2R \cdot e^{-R^2} \, \mathrm{d}R \quad \big| \; u = R^2; \quad u' = 2R \\
&= \pi \int_{0^2}^{\infty^2} e^{-u} \, \mathrm{d}u \;=\; \pi \int_{0}^{\infty} e^{-u} \, \mathrm{d}u \\
&= \pi \left[e^{-u}\right]_0^{\infty} \cdot (-1) \\
&= \pi \cdot \left(e^{-\infty} - e^{0}\right) \cdot (-1) \\
&= \pi \cdot (0 - 1) \cdot (-1) \\
&= \pi
\end{aligned}$$

$$\Longrightarrow \quad \int_{-\infty}^{\infty} e^{-x^2} \, \mathrm{d}x \;=\; \sqrt{\pi}$$

Literatur

[1] U. Kiencke, M. Schwarz, and T. Weickert. *Signalverarbeitung – Zeit-Frequenz-Analyse und Schätzverfahren.* Oldenbourg, 2008.