Advanced Sensor Data Fusion in Distributed Systems

SS 2018
Felix Govaers,
Dr. rer. nat., Dipl.-Math.
BROWNIAN MOTION
Brownian Motion Definition

Goal: Model for a macroscopic description of microscopic random effects.
Consider the Random Walk described by:

\[ S_n = S_0 + \sum_{i=1}^{n} X_i \]

where \( X_1, X_2, \ldots \) are independent and identically distributed (iid).

The continuous limit of this walk is a Random Path, a continuous time stochastic process:

\[ \{ B(t) : t \geq 0 \} \]
1-D Examples
Continuous Time Stochastic Process

A random path $B(t)$ has the following properties:

- for all $t_n$, $t_{n-1}$, ...$t_0$, $t_1$, the following RVs are mutually independent:

$$B(t_n) - B(t_{n-1}), B(t_{n-1}) - B(t_{n-2}), \ldots, B(t_2) - B(t_1)$$

- the distribution of $B(t+h) - B(t)$ does not depend on $t$.

If the increments are Gaussian distributed, $B$ is called a Wiener process.

Well suited model for target tracking?
PDF of a Brownian Motion
Supplemental material

STOCHASTIC DIFFERENTIAL EQUATION
The Wiener process is defined by:

1. \( w_0 = 0 \)
2. \( w_k - w_{k-1} \) is independent of \( w_j - w_{j-1} \)
3. \( w_k - w_l \sim N(0, t_k - t_l) \)

The Wiener process has the following properties:

- It is a Gaussian process
- \( E[w_k] = 0 \quad \text{cov}[w_k, w_l] = \min(t_k, t_l) \)
- It is a Markov process
A Wiener process can be integrated by using a Îto integral:

$$\int_{0}^{t} B \, dw := \lim_{n \to \infty} \sum_{i=1}^{n} B(i)(w_i - w_{i-1})$$

A stochastic process therefore is defined by either the stochastic integral form or equivalently the stochastic differential equation (SDE):

$$x_k = x_0 + \int_{t_0}^{t_k} A(x_\tau) \, d\tau + \int_{t_0}^{t_k} B(x_\tau) \, dw_\tau$$

$$dx_l = A(x_l) \, dt + B(x_l) \, dw_l$$
Kinematic Stochastic Process (Example)

Consider the following state variables

\[ x_l = \begin{pmatrix} r_l \\ v_l \end{pmatrix} \]

\[ r_l = (x^1, x^2, \ldots, x^d)^\top \]

\[ v_l = (v^1, v^2, \ldots, v^d)^\top \]

Then, the process follows the following linear SDE

\[ \text{d}x_l = A x_l \text{d}t + B \text{d}w_l. \]

\[ A = \begin{pmatrix} O & I \\ O & O \end{pmatrix} \quad B > 0. \]

Then:

\[ x_1 = x_0 + A x_0 \text{d}t, \]
\[ x_2 = x_0 + A x_0 \text{d}t + A (x_0 + A x_0 \text{d}t) \text{d}t \]
\[ = x_0 + A x_0 (2 \text{d}t) + A^2 x_0 \text{d}t^2 \]

Since \( A^m = 0 \) for \( m \geq 2 \)

\[ x_l = \begin{pmatrix} x_0 + A x_0 (l \text{d}t) \\ F_{l \mid 0} x_0 \end{pmatrix} \]

where

\[ F_{l \mid 0} = I + (t_l - t_0) A. \]

w/o noise!
Time Discrete Motion Model

The general solution for the transition matrix $F$ is given by

$$F_{l|0} = \exp(A(ldt)) = \sum_{m=0}^{\infty} \frac{A^m(ldt)^m}{m!}.$$ 

Then, the time discrete difference equation is

$$x_{l+1} = F_{l+1|l}x_l + w_{l+1|l},$$

$$w_{l+1|l} := \int_{t_l}^{t_{l+1}} BF_{l+1|\tau}dw_{\tau}$$
Statistics of Time Discrete Noise

We have:

\[ E \left[ \mathbf{w}_{l+1|l} \right] = E \left[ \int_{t_l}^{t_{l+1}} B \mathbf{F}_{l+1|\tau} d\mathbf{w}_\tau \right] \]

\[ = E \left[ \lim_{m \to \infty} \sum_{j=1}^{m} B \mathbf{F}_{l+1|\tau} (\mathbf{w}_{\tau j} - \mathbf{w}_{\tau j-1}) \right] \]

\[ = \lim_{m \to \infty} B \sum_{j=1}^{m} \mathbf{F}_{l+1|\tau} E \left[ (\mathbf{w}_{\tau j} - \mathbf{w}_{\tau j-1}) \right] \]

\[ = 0 \]

\[ E \left[ \mathbf{w}_{l+1|l} \right] = E \left[ \int_{t_l}^{t_{l+1}} B \mathbf{F}_{l+1|\tau} d\mathbf{w}_\tau \right] \]

\[ = E \left[ \int_{t_l}^{t_{l+1}} B \mathbf{F}_{l+1|\tau} d\mathbf{w}_\tau \left( \mathbf{F}_{l+1|\tau} \right)^T \right] \]

\[ = B^2 \lim_{m \to \infty} \sum_{j=1}^{m} \mathbf{F}_{l+1|\tau} E \left[ (\mathbf{w}_{\tau j} - \mathbf{w}_{\tau j-1}) (\mathbf{w}_{\tau j} - \mathbf{w}_{\tau j-1})^T \right] \mathbf{F}_{l+1|\tau}^T \]

\[ = B^2 \lim_{m \to \infty} \sum_{j=1}^{m} \mathbf{F}_{l+1|\tau} \mathbf{F}_{l+1|\tau}^T \cdot \frac{1}{m} \]

\[ = B^2 \int_{t_l}^{t_{l+1}} d\tau \mathbf{F}_{l+1|\tau} \mathbf{F}_{l+1|\tau}^T \]

\[ =: Q_{l+1|l} \]

\[ p(\mathbf{x}_k|\mathbf{x}_{k-1}) = \mathcal{N}(\mathbf{x}_k; \mathbf{F}_k|k-1 \mathbf{x}_{k-1}, \mathbf{Q}_k|k-1) \]
Parameter estimation for Sensor Data Fusion using the

BAYESIAN APPROACH
Properties of Joint Densities

- Non-negative:
  \[ p(x, y) \geq 0 \]

- Normalized:
  \[ \int dx \, dy \, p(x, y) = 1 \]

- Relation between \( x \) and \( y \)
  \[ p(x) = \int dy \, p(x, y) \]
  \[ p(y) = \int dx \, p(x, y) \]

- We saw:
  \[
  p(x|y) = \frac{p(x, y)}{p(y)}
  \]

- If \( x \) and \( y \) are mutually independent, iff:
  \[ p(x, y) = p(x) \cdot p(y) \]

- then:
  \[ p(x|y) = \frac{p(x, y)}{p(y)} = \frac{p(x)p(y)}{p(y)} = p(x) \]
Objective: Target Tracking

- **Given:** Set of measurements.
- **Recursive expression:**
  \[ Z^k = \{z_k, z_{k-1}, z_{k-2}, \ldots, z_1\} = \{z_k Z^{k-1}\} \]
- **Multivariate RV \( x \): state at time \( k \)
  \[ x_k = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \]
- **Objective:** Calculate
  \[ p(x_k | Z^k) \]
  Gaussian case:
  \[ x_k|k = E \left[ x_k | Z^k \right] \]
  \[ P_{k|k} = \text{cov} \left[ x_k | Z^k \right] \]
The BAYES Formula

- Rule (incl. proof):

\[
p(x|y) = \frac{p(x, y)}{p(y)} = \frac{p(y, x)}{p(y)} = \frac{p(y|x)p(x)}{p(y)}
\]

Bayes Rule: 

\[
p(x|y) = \frac{p(y|x)p(x)}{p(y)}
\]
Simplifications of Bayes Theorem

Normalization can be solved by

- **Integration**

\[
p(x|y) = \frac{p(y|x)p(x)}{p(y)} = \frac{p(y|x)p(x)}{\int dx \ p(y, x)} = \frac{p(y|x)p(x)}{\int dx \ p(y|x)p(x)}
\]

Assume \( p(x|y) \) is a Gaussian density.

\[
p(x|y) = \frac{p(y|x)p(x)}{p(y)} \]

Gaussian is normalized.

\[
\rightarrow \text{sufficient to know:} \quad p(x|y) \propto p(y|x)p(x)
\]

**Posterior** \( \propto \) **Likelihood function** \( \cdot \) **Prior**
The core of Linear Estimation

THE KALMAN FILTER
The track of a Kalman filter is initialized by a Gaussian density

\[ p(x_0) = \mathcal{N}(x_0; x_{0|0}, P_{0|0}) \]

with
- Initialization estimate \( x_{0|0} \)
- First measurement
- Track extraction algorithm
- Init. covariance \( P_{0|0} \)
- User parameter
- “Initial ignorance” -> ‘large’ covariance
Prediction in Bayesian Estimation

Prediction is the prior target state pdf for the next time step.

\[ p(x_{k-1} | \mathcal{Z}^{k-1}) \xrightarrow{\text{dynamics model}} p(x_k | \mathcal{Z}^{k-1}) \]

\[ p(x_k | \mathcal{Z}^{k-1}) \xrightarrow{\text{measurement } z_k} p(x_k | \mathcal{Z}^k) \]
Derivation of the Prior PDF

Use marginalization:

\[
p(x_k|Z^{k-1}) = \int dx_{k-1} p(x_k, x_{k-1}|Z^{k-1})
= \int dx_{k-1} p(x_k|x_{k-1}, Z^{k-1}) \cdot p(x_{k-1}|Z^{k-1})
= \int dx_{k-1} p(x_k|x_{k-1}) \cdot p(x_{k-1}|Z^{k-1})
\]

The last equation follows from the Markov property!

But how to solve integral?
Kalman Filter Assumptions

The Kalman filter assumptions for dynamics are

- linear Gaussian transition kernel
  
  $$p(x_k | x_{k-1}) = \mathcal{N}(x_k; F_{k|k-1}x_{k-1}, Q_{k|k-1})$$

- previous posterior pdf is a Gaussian
  
  $$p(x_{k-1} | Z^{k-1}) = \mathcal{N}(x_{k-1}; x_{k-1|k-1}, P_{k-1|k-1})$$

Then, the prior for time $k$ is also a Gaussian:

$$p(x_k | Z^{k-1}) = \mathcal{N}(x_k; x_{k|k-1}, P_{k|k-1})$$
Prior Mean for Kalman Filter

The prior mean is given by

\[
\begin{align*}
x_{k|k-1} &= \mathbb{E} \left[ x_k | \mathcal{Z}^{k-1} \right] \\
&= \mathbb{E} \left[ F_{k|k-1} x_{k-1} + w_k | \mathcal{Z}^{k-1} \right] \\
&= F_{k|k-1} \mathbb{E} \left[ x_{k-1} | \mathcal{Z}^{k-1} \right] + \mathbb{E} \left[ w_k | \mathcal{Z}^{k-1} \right] \\
&= F_{k|k-1} \mathbb{E} \left[ x_{k-1} | \mathcal{Z}^{k-1} \right] + \mathbb{E} \left[ w_k \right] \\
&= F_{k|k-1} x_{k-1|k-1}
\end{align*}
\]

\[
x_k = F_{k|k-1} x_{k-1} + w_k
\]
Prior Covariance for Kalman Filter

The prior covariance is given by

$$x_k = F_{k|k-1}x_{k-1} + w_k$$

$$P_{k|k-1} = \text{cov} \left[ x_k | Z^{k-1} \right]$$

$$= E \left[ (x_k - E \left[ x_k | Z^{k-1} \right]) (x_k - E \left[ x_k | Z^{k-1} \right])^\top \mid Z^{k-1} \right]$$

$$= E \left[ (x_k - x_{k|k-1}) (x_k - x_{k|k-1})^\top \mid Z^{k-1} \right]$$

$$= E \left[ (F_{k|k-1}x_{k-1} + w_k - x_{k|k-1}) (F_{k|k-1}x_{k-1} + w_k - x_{k|k-1})^\top \mid Z^{k-1} \right]$$

$$= E \left[ (F_{k|k-1}x_{k-1} - F_{k|k-1}x_{k-1|k-1}) (F_{k|k-1}x_{k-1} - F_{k|k-1}x_{k-1|k-1})^\top \mid Z^{k-1} \right] + E \left[ ww^\top \right]$$

$$= F_{k|k-1} \text{cov} \left[ x_{k-1} | Z^{k-1} \right] F_{k|k-1}^\top + \text{cov} \left[ w_k \right]$$

$$= F_{k|k-1} P_{k-1|k-1} F_{k|k-1}^\top + Q_{k|k-1}$$
Kalman Prediction

As a conclusion, we have

\[
p(x_{k-1}|Z^{k-1}) \xrightarrow{\text{dynamics model}} p(x_k|Z^{k-1})
\]

\[
\mathcal{N}(x_{k-1}; x_{k-1|k-1}, P_{k-1|k-1}) \xrightarrow{\text{dynamics model}} \mathcal{N}(x_k; x_{k|k-1}, P_{k|k-1})
\]

\[
x_{k|k-1} = F_{k|k-1} x_{k-1|k-1}
\]

\[
P_{k|k-1} = F_{k|k-1} P_{k-1|k-1} F_{k|k-1}^\top + Q_{k|k-1}
\]
The FUNDAMENTAL EQUATIONS
Filtering is the *posterior* target state pdf after processing the current measurement $z$. 

$$
p(x_{k-1} | z^{k-1}) \xrightarrow{\text{dynamics model}} p(x_k | z^{k-1})
$$

$$
p(x_k | z^{k-1}) \xrightarrow{\text{measurement } z_k} p(x_k | z^k)
$$
Derivation of the Posterior PDF

The solution is given by BAYES theorem:

$$p(x_k|Z^k) = \frac{p(z_k|x_k, Z^{k-1}) \cdot p(x_k|Z^{k-1})}{\int dx_k p(z_k|x_k, Z^{k-1}) \cdot p(x_k|Z^{k-1})}$$

$$= \frac{p(z_k|x_k) \cdot p(x_k|Z^{k-1})}{\int dx_k p(z_k|x_k) \cdot p(x_k|Z^{k-1})}$$

Prior PDF

Sensor model / Likelihood function

Posterior \(\propto\) Likelihood function \(\cdot\) Prior
Kalman Filter Assumptions

The sensor model is given by a Gaussian linear pdf:

\[
p(z_k|x_k) = \mathcal{N}(z_k; H_k x_k, R_k)
\]

The prior pdf is also a Gaussian pdf:

\[
p(x_k|\mathcal{Z}^{k-1}) = \mathcal{N}(x_k; x_{k|k-1}, P_{k|k-1})
\]

How to calculate the *product* of two Gaussians?

\[
p(x_k|\mathcal{Z}^k) = c \cdot \mathcal{N}(z_k; H_k x_k, R_k) \cdot \mathcal{N}(x_k; x_{k|k-1}, P_{k|k-1})
\]
A small excursion to

LEAST SQUARES FILTERING
Problem Setting

Assume a random variable $z$ is given as a set of observations of $x$. We further have

$$E[z] = Hx$$
$$\text{cov}[z] = R$$

Observe:

$$\arg\max_x \mathcal{N}(z; Hx, R) = \arg\min_x (z - Hx)^T R^{-1} (z - Hx)$$

- Maximum Likelihood Estimate $\text{ML}(x|z)$
- Least Squares Estimate $\text{LS}(x|z)$
Computation of Least Squares Solution

The solution of the LS estimator has zero gradient:

\[ 0 = \nabla_x \left[ (z - Hx)^\top R^{-1} (z - Hx) \right] \]
\[ = 2 (z - Hx)^\top R^{-1} H \]

Therefore \[ H^\top R^{-1} z = H^\top R^{-1} H x \]
and finally

\[ x = (H^\top R^{-1} H)^{-1} H^\top R^{-1} z \]

Least Squares Estimate \( \hat{x} \)
Covariance of LS Solution

A short computation yields

\[
\text{cov} [\hat{x}] = E \left[ (x - \hat{x})(x - \hat{x})^\top \right] \\
= E \left[ (x - (H^\top R^{-1}H)^{-1}H^\top R^{-1}z)(x - (H^\top R^{-1}H)^{-1}H^\top R^{-1}z)^\top \right] \\
= E \left[ (H^\top R^{-1}H)((H^\top R^{-1}H)x - H^\top R^{-1}z)((H^\top R^{-1}H)x - H^\top R^{-1}z)^\top (H^\top R^{-1}H)^{-1} \right] \\
= E \left[ (H^\top R^{-1}H)H^\top R^{-1}(Hx - z)(Hx - z)^\top (H^\top R^{-1})^\top (H^\top R^{-1}H)^{-1} \right] \\
= (H^\top R^{-1}H)H^\top R^{-1}E \left[ (Hx - z)(Hx - z)^\top \right] R^{-1}H(H^\top R^{-1}H)^{-1} \\
= (H^\top R^{-1}H)^{-1}
\]
Least Squares Conclusion

Together, we have

\[ P = (H^\top R^{-1}H)^{-1} \]
\[ \hat{x} = PH^\top R^{-1}z \]

if

\[ E[z] = Hx \]
\[ \text{cov}[z] = R \]

and where

\[ P = \text{cov} [\hat{x}|z] \]
The alternative derivation

KALMAN FILTER AS LEAST SQUARES SOLUTION
The posterior estimate is obtained as the solution of the LS where the prior estimate and the measurement are the input data.

We define:

\[
    z^* = \begin{pmatrix} x_{k|k-1} \\ z_k \end{pmatrix}
\]

Therefore, it holds that

\[
    E[z^* | x_k] = \begin{pmatrix} x_k \\ H_k x_k \end{pmatrix}
\]

which also equals \( H^* x_k \).

We obtain

\[
    H^* = \begin{pmatrix} I \\ H_k \end{pmatrix}
\]

Furthermore, we have

\[
    \text{cov} [z^* | x_k] = R^* = \begin{pmatrix} P_{k|k-1} & R_k \end{pmatrix}
\]
Posterior Covariance by LS Solution

Now, we apply the LS equations:

\[
P_{k|k} = (H^* \, ^\top R^* \, ^{-1} H^*)^{-1}
\]

\[
= \left( \begin{pmatrix} I & H_k^\top \end{pmatrix} \begin{pmatrix} P_{k|k-1} & R_k \end{pmatrix}^{-1} \begin{pmatrix} I \\ H_k \end{pmatrix} \right)^{-1}
\]

\[
= \left( \begin{pmatrix} I & H_k^\top \end{pmatrix} \begin{pmatrix} P_{k|k-1}^{-1} & R_k^{-1} \end{pmatrix} \begin{pmatrix} I \\ H_k \end{pmatrix} \right)^{-1}
\]

\[
= \left( \begin{pmatrix} P_{k|k-1}^{-1} & H_k^\top R_k^{-1} \end{pmatrix} \begin{pmatrix} I \\ H_k \end{pmatrix} \right)^{-1}
\]

\[
= (P_{k|k-1}^{-1} + H_k^\top R_k^{-1} H_k)^{-1}
\]
Posterior Estimate by LS Solution

And for the state estimate, we have

\[
x_{k|k} = P_{k|k}(H^* \mathsf{T} R^*-1 z^*)
\]

\[
= P_{k|k} \begin{pmatrix} I & H^*_k \end{pmatrix} \begin{pmatrix} P_{k|k-1}^{-1} & 0 \end{pmatrix} \begin{pmatrix} x_{k|k-1} \\ z_k \end{pmatrix}
\]

\[
= P_{k|k} \begin{pmatrix} P_{k|k-1}^{-1} & H^*_k R_k^{-1} \end{pmatrix} \begin{pmatrix} x_{k|k-1} \\ z_k \end{pmatrix}
\]

\[
= P_{k|k} (P_{k|k-1}^{-1} x_{k|k-1} + H^*_k R_k^{-1} z_k)
\]

Together, we have obtained the **information filter** form of the Kalman filter:

\[
P_{k|k} = (P_{k|k-1}^{-1} + H^*_k R_k^{-1} H_k)^{-1}
\]

\[
x_{k|k} = P_{k|k} (P_{k|k-1}^{-1} x_{k|k-1} + H^*_k R_k^{-1} z_k)
\]
Information Filter Form Summary

We obtained the following formulas which are equivalent to the Kalman filter update:

\[ P_{k|k} = \left( P_{k|k-1}^{-1} - H_k^T R_k^{-1} H_k \right)^{-1} \]

\[ x_{k|k} = P_{k|k} \left( P_{k|k-1}^{-1} x_{k|k-1} + H_k^T R_k^{-1} z_k \right) \]

Note: The inverse of a covariance matrix is also called *Fisher Information Matrix*.

There is a one–one correspondence between Kalman filter – Information filter.