EE 570: Location and Navigation

On-Line Bayesian Tracking

Aly El-Osery   Kevin Wedeward

Electrical Engineering Department, New Mexico Tech
Socorro, New Mexico, USA

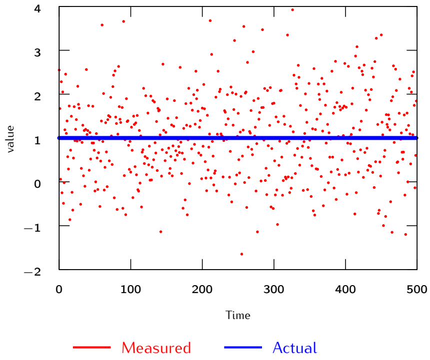
In Collaboration with
Stephen Bruder
Electrical and Computer Engineering Department
Embry-Riddle Aeronautical Univesity
Prescott, Arizona, USA

April 6, 2016

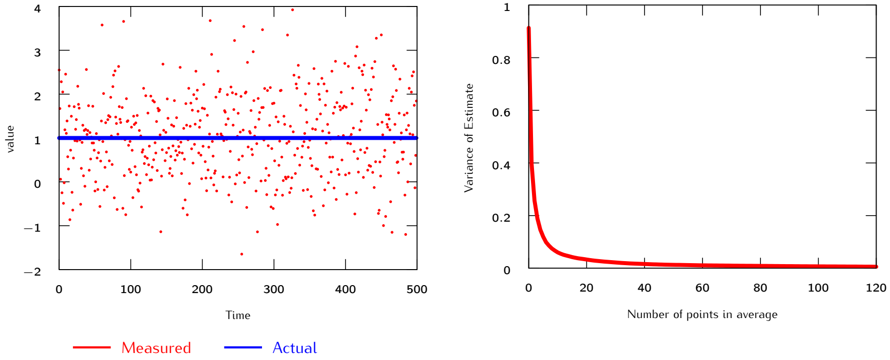
Sequentially estimate on-line the states of a system as it changes over time using observations that are corrupted with noise.

- *Filtering*: the time of the estimate coincides with the last measurement.
- *Smoothing*: the time of the estimate is within the span of the measurements.
- *Prediction*: the time of the estimate occurs after the last available measurement.

Estimate the value of a random constant. How many points do you need?



Estimate the value of a random constant. How many points do you need?



- The best estimate is the mean.
- Variance of the estimate decreases as 1/N.

- For a stationary process that represents a random constant, averaging over more points results in an improved estimate.
- What will happen if the same is applied to a non-constant?
- If we have a measurement corrupted with noise, can we use the statistical properties of the noise, and compute an estimate that maximizes the probability that this measurement actually occurred?
- For real-time applications, can we solve the estimation problem recursively?

$$\vec{\boldsymbol{x}}_k = \mathrm{f}_k(\vec{\boldsymbol{x}}_{k-1}, \vec{\boldsymbol{w}}_{k-1}) \quad (1)$$

$$\vec{\boldsymbol{z}}_k = \mathrm{h}_k(\vec{\boldsymbol{x}}_k, \vec{\boldsymbol{v}}_k) \quad (2)$$

$(n \times 1)$ state vector at time k

$$\vec{\boldsymbol{x}}_k = \mathrm{f}_k(\vec{\boldsymbol{x}}_{k-1}, \vec{\boldsymbol{w}}_{k-1}) \quad (1)$$

$$\vec{\boldsymbol{z}}_k = \mathrm{h}_k(\vec{\boldsymbol{x}}_k, \vec{\boldsymbol{v}}_k) \quad (2)$$

$(m \times 1)$ measurement vector at time k

Possibly non-linear function,
$f_k : \mathfrak{R}^n \times \mathfrak{R}^{n_w} \mapsto \mathfrak{R}^n$

$$\vec{\boldsymbol{x}}_k = f_k(\vec{\boldsymbol{x}}_{k-1}, \vec{\boldsymbol{w}}_{k-1}) \quad (1)$$

$$\vec{\boldsymbol{z}}_k = h_k(\vec{\boldsymbol{x}}_k, \vec{\boldsymbol{v}}_k) \quad (2)$$

Possibly non-linear function,
$h_k : \mathfrak{R}^m \times \mathfrak{R}^{n_v} \mapsto \mathfrak{R}^m$

i.i.d state noise

$$\vec{\boldsymbol{x}}_k = \mathrm{f}_k(\vec{\boldsymbol{x}}_{k-1}, \vec{\boldsymbol{w}}_{k-1}) \quad (1)$$

$$\vec{\boldsymbol{z}}_k = \mathrm{h}_k(\vec{\boldsymbol{x}}_k, \vec{\boldsymbol{v}}_k) \quad (2)$$

i.i.d measurement noise

$$\vec{\boldsymbol{x}}_k = \mathrm{f}_k(\vec{\boldsymbol{x}}_{k-1}, \vec{\boldsymbol{w}}_{k-1}) \tag{1}$$

$$\vec{\boldsymbol{z}}_k = \mathrm{h}_k(\vec{\boldsymbol{x}}_k, \vec{\boldsymbol{v}}_k) \tag{2}$$

The state process is Markov chain, i.e., $p(\vec{\boldsymbol{x}}_k|\vec{\boldsymbol{x}}_1, \ldots, \vec{\boldsymbol{x}}_{k-1}) = p(\vec{\boldsymbol{x}}_k|\vec{\boldsymbol{x}}_{k-1})$ and the distribution of $\vec{\boldsymbol{z}}_k$ conditional on the state $\vec{\boldsymbol{x}}_k$ is independent of previous state and measurement values, i.e., $p(\vec{\boldsymbol{z}}_k|\vec{\boldsymbol{x}}_{1:k}, \vec{\boldsymbol{z}}_{1:k-1}) = p(\vec{\boldsymbol{z}}_k|\vec{\boldsymbol{x}}_k)$

Probabilistically estimate $\vec{\boldsymbol{x}}_k$ using previous measurement $\vec{\boldsymbol{z}}_{1:k}$. In other words, construct the pdf $p(\vec{\boldsymbol{x}}_k|\vec{\boldsymbol{z}}_{1:k})$.

Probabilistically estimate $\vec{\boldsymbol{x}}_k$ using previous measurement $\vec{\boldsymbol{z}}_{1:k}$. In other words, construct the pdf $p(\vec{\boldsymbol{x}}_k|\vec{\boldsymbol{z}}_{1:k})$.

Optimal MMSE Estimate

$$\mathbb{E}\{\|\vec{\boldsymbol{x}}_k - \hat{\vec{\boldsymbol{x}}}_k\|^2|\vec{\boldsymbol{z}}_{1:k}\} = \int \|\vec{\boldsymbol{x}}_k - \hat{\vec{\boldsymbol{x}}}_k\|^2 p(\vec{\boldsymbol{x}}_k|\vec{\boldsymbol{z}}_{1:k})d\vec{\boldsymbol{x}}_k \tag{3}$$

in other words find the conditional mean

$$\hat{\vec{\boldsymbol{x}}}_k = \mathbb{E}\{\vec{\boldsymbol{x}}_k|\vec{\boldsymbol{z}}_{1:k}\} = \int \vec{\boldsymbol{x}}_k p(\vec{\boldsymbol{x}}_k|\vec{\boldsymbol{z}}_{1:k})d\vec{\boldsymbol{x}}_k \tag{4}$$

- $\vec{w}_k$ and $\vec{v}_k$ are drawn from a Gaussian distribution, uncorrelated have zero mean and statistically independent.

$$\mathbb{E}\{\vec{w}_k \vec{w}_i^T\} = \begin{cases} Q_k & i = k \\ 0 & i \neq k \end{cases} \tag{5}$$

$$\mathbb{E}\{\vec{v}_k \vec{v}_i^T\} = \begin{cases} R_k & i = k \\ 0 & i \neq k \end{cases} \tag{6}$$

$$\mathbb{E}\{\vec{w}_k \vec{v}_i^T\} = \begin{cases} 0 & \forall i, k \end{cases} \tag{7}$$

- f_k and h_k are both linear, e.g., the state-space system equations may be written as

$$\vec{x}_k = \Phi_{k-1} \vec{x}_{k-1} + \vec{w}_{k-1} \tag{8}$$

$$\vec{y}_k = H_k \vec{x}_k + \vec{v}_k \tag{9}$$

- f_k and h_k are both linear, e.g., the state-space system equations may be written as

$$\vec{\boldsymbol{x}}_k = \Phi_{k-1}\, \vec{\boldsymbol{x}}_{k-1} + \vec{\boldsymbol{w}}_{k-1} \tag{8}$$

$$\vec{\boldsymbol{y}}_k = \text{H}_k\, \vec{\boldsymbol{x}}_k + \vec{\boldsymbol{v}}_k \tag{9}$$

$(n \times n)$ transition matrix relating $\vec{\boldsymbol{x}}_{k-1}$ to $\vec{\boldsymbol{x}}_k$

- f_k and h_k are both linear, e.g., the state-space system equations may be written as

$$\vec{\boldsymbol{x}}_k = \boldsymbol{\Phi}_{k-1}\, \vec{\boldsymbol{x}}_{k-1} + \vec{\boldsymbol{w}}_{k-1} \tag{8}$$

$$\vec{\boldsymbol{y}}_k = \mathrm{H}_k\, \vec{\boldsymbol{x}}_k + \vec{\boldsymbol{v}}_k \tag{9}$$

($m \times n$) matrix provides noiseless connection between measurement and state vectors

$$\hat{\vec{\boldsymbol{x}}}_{k|k-1} = \boldsymbol{\Phi}_{k-1}\hat{\vec{\boldsymbol{x}}}_{k-1|k-1} \tag{10}$$

$$\mathsf{P}_{k|k-1} = \mathsf{Q}_{k-1} + \boldsymbol{\Phi}_{k-1}\mathsf{P}_{k-1|k-1}\boldsymbol{\Phi}_{k-1}^{T} \tag{11}$$

$$\hat{\vec{\boldsymbol{x}}}_{k|k} = \hat{\vec{\boldsymbol{x}}}_{k|k-1} + \mathsf{K}_k \left(\vec{\boldsymbol{z}}_k - \mathsf{H}_k\hat{\vec{\boldsymbol{x}}}_{k|k-1}\right) \tag{12}$$

$$\mathsf{P}_{k|k} = (\mathsf{I} - \mathsf{K}_k\mathsf{H}_k)\mathsf{P}_{k|k-1} \tag{13}$$

$$\hat{\vec{\boldsymbol{x}}}_{k|k-1} = \boldsymbol{\Phi}_{k-1}\hat{\vec{\boldsymbol{x}}}_{k-1|k-1} \tag{10}$$

$$\mathsf{P}_{k|k-1} = \mathsf{Q}_{k-1} + \boldsymbol{\Phi}_{k-1}\mathsf{P}_{k-1|k-1}\boldsymbol{\Phi}_{k-1}^{T} \tag{11}$$

$$\hat{\vec{\boldsymbol{x}}}_{k|k} = \hat{\vec{\boldsymbol{x}}}_{k|k-1} + \mathsf{K}_k(\vec{\boldsymbol{z}}_k - \mathsf{H}_k\hat{\vec{\boldsymbol{x}}}_{k|k-1}) \tag{12}$$

$$\mathsf{P}_{k|k} = (\mathsf{I} - \mathsf{K}_k\mathsf{H}_k)\mathsf{P}_{k|k-1} \tag{13}$$

$(n \times m)$ Kalman gain

$$\hat{\vec{x}}_{k|k-1} = \mathbf{\Phi}_{k-1}\hat{\vec{x}}_{k-1|k-1} \tag{10}$$

$$P_{k|k-1} = Q_{k-1} + \mathbf{\Phi}_{k-1}P_{k-1|k-1}\mathbf{\Phi}_{k-1}^T \tag{11}$$

$$\hat{\vec{x}}_{k|k} = \hat{\vec{x}}_{k|k-1} + K_k\left(\vec{z}_k - H_k\hat{\vec{x}}_{k|k-1}\right) \tag{12}$$

$$P_{k|k} = (I - K_kH_k)P_{k|k-1} \tag{13}$$

Measurement innovation

Kalman Gain

$$K_k = P_{k|k-1} H_k^T \left(H_k P_{k|k-1} H_k^T + R_k \right)^{-1} \tag{14}$$

$$K_k = P_{k|k-1} H_k^T \left(H_k P_{k|k-1} H_k^T + R_k \right)^{-1} \tag{14}$$

Covariance of the innovation term

Initial estimate ($\hat{\vec{x}}_0$ and $\mathbf{P_0}$)

Initial estimate ($\hat{\vec{x}}_0$ and $\mathbf{P_0}$)

↓

Compute Kalman gain

$$\mathbf{K}_k = \mathbf{P}_{k|k-1}\mathbf{H}_k^T(\mathbf{H}_k\mathbf{P}_{k|k-1}\mathbf{H}_k^T + \mathbf{R}_k)^{-1}$$

Kalman filter data flow

Initial estimate ($\hat{\vec{x}}_0$ and $\mathbf{P}_0$)

Compute Kalman gain

$$\mathsf{K}_k = \mathsf{P}_{k|k-1}\mathsf{H}_k^T(\mathsf{H}_k\mathsf{P}_{k|k-1}\mathsf{H}_k^T + \mathsf{R}_k)^{-1}$$

Update estimate with measurement $\vec{z}_k$

$$\hat{\vec{x}}_{k|k} = \hat{\vec{x}}_{k|k-1} + \mathsf{K}_k(\vec{z}_k - \mathsf{H}_k\hat{\vec{x}}_{k|k-1})$$

Initial estimate ($\hat{\vec{x}}_0$ and P_0)

Compute Kalman gain
$$K_k = P_{k|k-1} H_k^T (H_k P_{k|k-1} H_k^T + R_k)^{-1}$$

Update estimate with measurement $\vec{z}_k$
$$\hat{\vec{x}}_{k|k} = \hat{\vec{x}}_{k|k-1} + K_k(\vec{z}_k - H_k \hat{\vec{x}}_{k|k-1})$$

Update error covariance
$$P_{k|k} = P_{k|k-1} - K_k H_k P_{k|k-1}$$

Kalman filter data flow

Initial estimate ($\hat{\vec{x}}_0$ and $\mathbf{P}_0$)

Compute Kalman gain

$$\mathbf{K}_k = \mathbf{P}_{k|k-1}\mathbf{H}_k^T(\mathbf{H}_k\mathbf{P}_{k|k-1}\mathbf{H}_k^T + \mathbf{R}_k)^{-1}$$

Update estimate with measurement $\vec{z}_k$

$$\hat{\vec{x}}_{k|k} = \hat{\vec{x}}_{k|k-1} + \mathbf{K}_k(\vec{z}_k - \mathbf{H}_k\hat{\vec{x}}_{k|k-1})$$

Update error covariance

$$\mathbf{P}_{k|k} = \mathbf{P}_{k|k-1} - \mathbf{K}_k\mathbf{H}_k\mathbf{P}_{k|k-1}$$

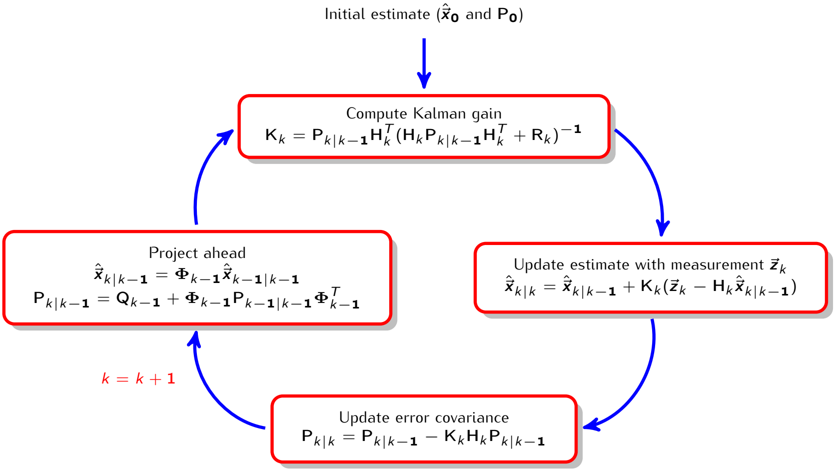
$k = k + 1$

Project ahead

$$\hat{\vec{x}}_{k|k-1} = \mathbf{\Phi}_{k-1}\hat{\vec{x}}_{k-1|k-1}$$

$$\mathbf{P}_{k|k-1} = \mathbf{Q}_{k-1} + \mathbf{\Phi}_{k-1}\mathbf{P}_{k-1|k-1}\mathbf{\Phi}_{k-1}^T$$

Kalman filter data flow



$$\dot{\vec{\boldsymbol{x}}}(t) = \mathsf{F}(t)\vec{\boldsymbol{x}}(t) + \mathsf{G}(t)\vec{\boldsymbol{w}}(t) \tag{15}$$

To obtain the state vector estimate $\hat{\vec{\boldsymbol{x}}}(t)$

$$\mathbb{E}\{\dot{\vec{\boldsymbol{x}}}(t)\} = \frac{\partial}{\partial t}\hat{\vec{\boldsymbol{x}}}(t) = \mathsf{F}(t)\hat{\vec{\boldsymbol{x}}}(t) \tag{16}$$

Solving the above equation over the interval $t - \tau_s, t$

$$\hat{\vec{\boldsymbol{x}}}(t) = e^{\left(\int_{t-\tau_s}^{t} \mathsf{F}(t')dt'\right)}\hat{\vec{\boldsymbol{x}}}(t - \tau_s) \tag{17}$$

where F_{k-1} is the average of F at times t and $t - \tau_s$.

As shown in the Kalman filter equations the state vector estimate is given by

$$\hat{\vec{\boldsymbol{x}}}_{k|k-1} = \boldsymbol{\Phi}_{k-1}\hat{\vec{\boldsymbol{x}}}_{k-1|k-1}$$

Therefore,

As shown in the Kalman filter equations the state vector estimate is given by

$$\hat{\vec{\boldsymbol{x}}}_{k|k-1} = \boldsymbol{\Phi}_{k-1}\hat{\vec{\boldsymbol{x}}}_{k-1|k-1}$$

Therefore,

$$\boldsymbol{\Phi}_{k-1} = e^{\mathsf{F}_{k-1}\tau_s} \approx \mathsf{I} + \mathsf{F}_{k-1}\tau_s \qquad (18)$$

where F_{k-1} is the average of F at times t and $t - \tau_s$, and first order approximation is used.

Assuming white noise, small time step, G is constant over the integration period, and the trapezoidal integration

$$Q_{k-1} \approx \frac{1}{2}\left[\boldsymbol{\Phi}_{k-1} G_{k-1} Q(t_{k-1}) G_{k-1}^T \boldsymbol{\Phi}_{k-1}^T + G_{k-1} Q(t_{k-1}) G_{k-1}^T\right] \tau_s \qquad (19)$$

where

$$\mathbb{E}\{\vec{\boldsymbol{w}}(\eta)\vec{\boldsymbol{w}}^T(\zeta)\} = Q(\eta)\delta(\eta - \zeta) \qquad (20)$$

$$\dot{x}(t) = 0, \qquad y_k = x_k + v_k$$

Design a Kalman filter to estimate x_k

$$\dot{x}(t) = 0, \qquad y_k = x_k + v_k$$

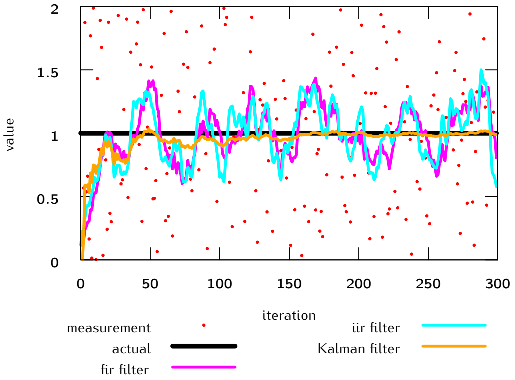
Design a Kalman filter to estimate x_k

- What is the discretized system?
- What is ϕ, Q, H, R and P?

$$\dot{x}(t) = 0, \qquad y_k = x_k + v_k$$

Design a Kalman filter to estimate x_k

- What is the discretized system?
- What is ϕ, Q, H, R and P?



State Equation

$$\dot{b}(t) = -\frac{1}{T_c} b(t) + w(t) \tag{21}$$

Autocorrelation Function

$$\mathbb{E}\{b(t)b(t+\tau)\} = \sigma_{BI}^2 e^{-|\tau|/T_c} \tag{22}$$

where

$$\mathbb{E}\{w(t)w(t+\tau)\} = Q(t)\delta(t-\tau) \tag{23}$$

$$Q(t) = \frac{2\sigma_{BI}^2}{T_c} \tag{24}$$

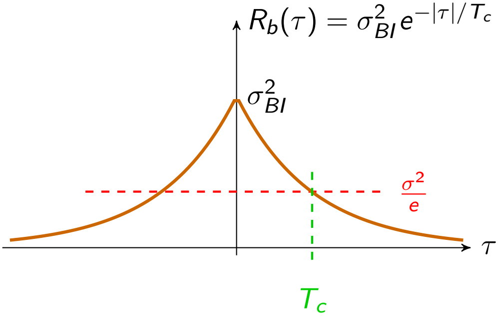
and T_c is the correlation time.

State Equation

$$b_k = e^{-\frac{1}{T_c}\tau_s} b_{k-1} + w_{k-1} \tag{25}$$

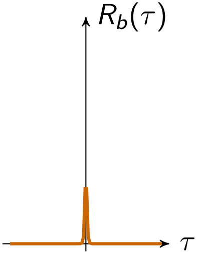
System Covariance Matrix

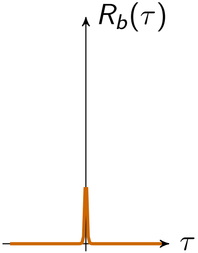
$$Q = \sigma_{BI}^2 [1 - e^{-\frac{2}{T_c}\tau_s}] \tag{26}$$

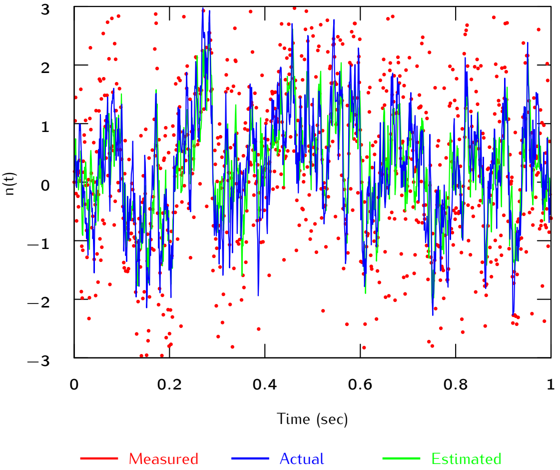


$R_b(\tau) = \sigma_{BI}^2 e^{-|\tau|/T_c}$
σ_{BI}^2
$\frac{\sigma^2}{e}$
τ
T_c

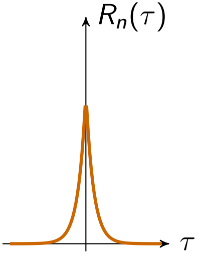
Small Correlation Time $T_c = 0.01$

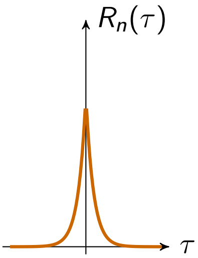


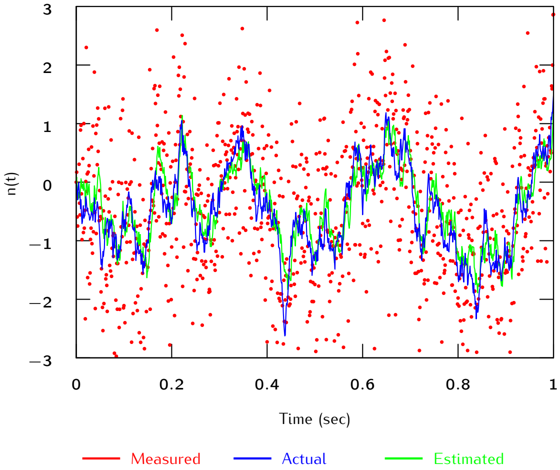

$R_b(\tau)$
τ


3
2
1
0
−1
−2
−3
n(t)
0
0.2
0.4
0.6
0.8
1
Time (sec)
Measured
Actual
Estimated

Larger Correlation Time $T_c = 0.1$




$R_n(\tau)$
τ


n(t)
Time (sec)
Measured
Actual
Estimated

$$\mathrm{F}_k = \left.\frac{\partial \mathrm{f}(\vec{\boldsymbol{x}})}{\partial \vec{\boldsymbol{x}}}\right|_{\vec{\boldsymbol{x}}=\hat{\vec{\boldsymbol{x}}}_{k|k-1}}, \qquad \mathrm{H}_k = \left.\frac{\partial \mathrm{h}(\vec{\boldsymbol{x}})}{\partial \vec{\boldsymbol{x}}}\right|_{\vec{\boldsymbol{x}}=\hat{\vec{\boldsymbol{x}}}_{k|k-1}} \tag{27}$$

where

$$\frac{\partial \mathrm{f}(\vec{\boldsymbol{x}})}{\partial \vec{\boldsymbol{x}}} = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots \\ \vdots & \ddots & \vdots \end{pmatrix}, \qquad \frac{\partial \mathrm{h}(\vec{\boldsymbol{x}})}{\partial \vec{\boldsymbol{x}}} = \begin{pmatrix} \frac{\partial h_1}{\partial x_1} & \frac{\partial h_1}{\partial x_2} & \cdots \\ \frac{\partial h_2}{\partial x_1} & \frac{\partial h_2}{\partial x_2} & \cdots \\ \vdots & \ddots & \vdots \end{pmatrix} \tag{28}$$

If R is a block matrix, i.e., $R = diag(R^1, R^2, \ldots, R^r)$. The R^i has dimensions $p^i \times p^i$. Then, we can sequentially process the measurements as:
For $i = 1, 2, \ldots, r$

$$\mathrm{K}^i = \mathrm{P}^{i-1}(\mathrm{H}^i)^T(\mathrm{H}^i\mathrm{P}^{i-1}(\mathrm{H}^i)^T + \mathrm{R}^i)^{-1} \tag{29}$$

$$\hat{\vec{\boldsymbol{x}}}^i_{k|k} = \hat{\vec{\boldsymbol{x}}}^i_{k|k} + \mathrm{K}^i(\vec{\boldsymbol{z}}^i_k - \mathrm{H}^i\hat{\vec{\boldsymbol{x}}}^{i-1}_{k|k}) \tag{30}$$

$$\mathrm{P}^i = (\mathrm{I} - \mathrm{K}^i\mathrm{H}^i)\mathrm{P}^{i-1} \tag{31}$$

where $\hat{\vec{\boldsymbol{x}}}^0_{k|k} = \hat{\vec{\boldsymbol{x}}}_{k|k-1}$, $\mathrm{P}^0 = \mathrm{P}^0_{k|k-1}$ and H^i is $p^i \times n$ corresponding to the rows of H corresponding the measurement being processed.

The system is observable if the observability matrix

$$\mathcal{O}(k) = \begin{bmatrix} \boldsymbol{H}(k-n+1) \\ \boldsymbol{H}(k-n-2)\boldsymbol{\Phi}(k-n+1) \\ \vdots \\ \boldsymbol{H}(k)\boldsymbol{\Phi}(k-1)\ldots\boldsymbol{\Phi}(k-n+1) \end{bmatrix} \quad (32)$$

where n is the number of states, has a rank of n. The rank of $\mathcal{O}$ is a binary indicator and does **not** provide a measure of how close the system is to being unobservable, hence, is prone to numerical ill-conditioning.

In addition to the computation of the rank of $\mathcal{O}(k)$, compute the Singular Value Decomposition (SVD) of $\mathcal{O}(k)$ as

$$\mathcal{O} = U\Sigma V^* \tag{33}$$

and observe the diagonal values of the matrix Σ. Using this approach it is possible to monitor the variations in the system observability due to changes in system dynamics.

Remarks

- Kalman filter is optimal under the aforementioned assumptions,
- and it is also an unbiased and minimum variance estimate.
- If the Gaussian assumptions is not true, Kalman filter is biased and not minimum variance.
- Observability is dynamics dependent.
- The error covariance update may be implemented using the *Joseph form* which provides a more stable solution due to the guaranteed symmetry.

$$\boldsymbol{P}_{k|k} = (\boldsymbol{I} - \boldsymbol{K}_k \boldsymbol{H}_k) \boldsymbol{P}_{k|k-1} (\boldsymbol{I} - \boldsymbol{K}_k \boldsymbol{H}_k)^T + \boldsymbol{K}_k \boldsymbol{R}_k \boldsymbol{K}_k^T \tag{34}$$

Propagates carefully chosen sample points (using unscented transformation) through the true non-linear system, and therefore captures the posterior mean and covariance accurately to the second order.

A Monte Carlo based method. It allows for a complete representation of the state distribution function. Unlike EKF and UKF, particle filters do not require the Gaussian assumptions.

Bayesian Filtering: From Kalman Filters to Particle Filters, and Beyond, by Zhe Chen