Lecture

On-Line Bayesian Tracking

EE 565: Position, Navigation, and Timing

Lecture Notes Update on April 4, 2018

Aly El-Osery and Kevin Wedeward, Electrical Engineering Dept., New Mexico Tech
In collaboration with
Stephen Bruder, Electrical & Computer Engineering, Embry-Riddle Aeronautical University

.1

Objective

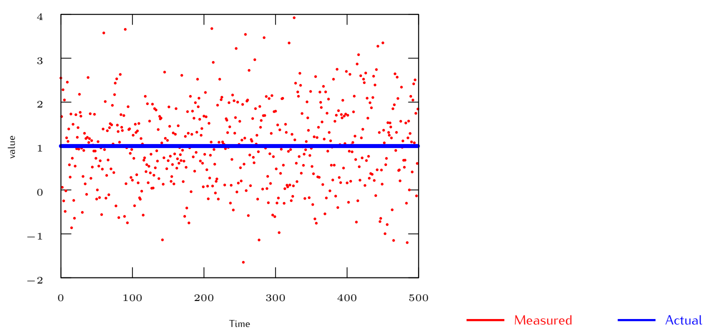
Sequentially estimate on-line the states of a system as it changes over time using observations that are corrupted with noise.

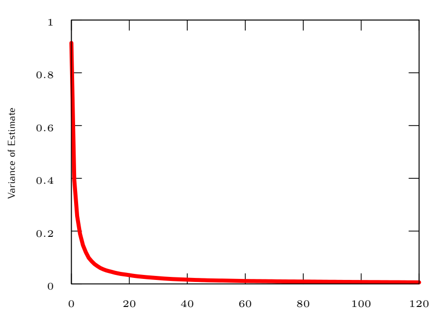
- *Filtering*: the time of the estimate coincides with the last measurement.
- *Smoothing*: the time of the estimate is within the span of the measurements.
- *Prediction*: the time of the estimate occurs after the last available measurement.

.2

Example: random constant

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.

.3

Remarks and Questions

- 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?

.4

1 Problem

Given State-Space Equations

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

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

where $\vec{\boldsymbol{x}}_k$ is $(n \times 1)$ state vector at time k, $\mathbf{f}_k$ and $\mathbf{h}_k$ are possibly non-linear function $\mathbf{f}_k : \Re^n \times \Re^{n_w} \mapsto \Re^n$ and $\mathbf{h}_k : \Re^m \times \Re^{n_v} \mapsto \Re^m$, respectively, and $\vec{\boldsymbol{w}}_k$ and $\vec{\boldsymbol{v}}_k$ i.i.d state noise. 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., $p(\vec{\boldsymbol{z}}_k|\vec{\boldsymbol{x}}_{1:k}, \vec{\boldsymbol{z}}_{1:k-1}) = p(\vec{\boldsymbol{z}}_k|\vec{\boldsymbol{x}}_k)$ .5

Objective

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

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

.6

2 Bayesian Estimation

3 Kalman Filter

Assumptions

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

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

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

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

.7

Assumptions

- $\mathbf{f}_k$ and $\mathbf{h}_k$ are both linear, e.g., the state-space system equations may be written as

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

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

where $\boldsymbol{\Phi}_{k-1}$ is $(n \times n)$ transition matrix relating $\vec{\boldsymbol{x}}_{k-1}$ to $\vec{\boldsymbol{x}}_k$, $\mathbf{H}_k$ is $(m \times n)$ matrix provides noiseless connection between measurement and state vectors. .8

State-Space Equations

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

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

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

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

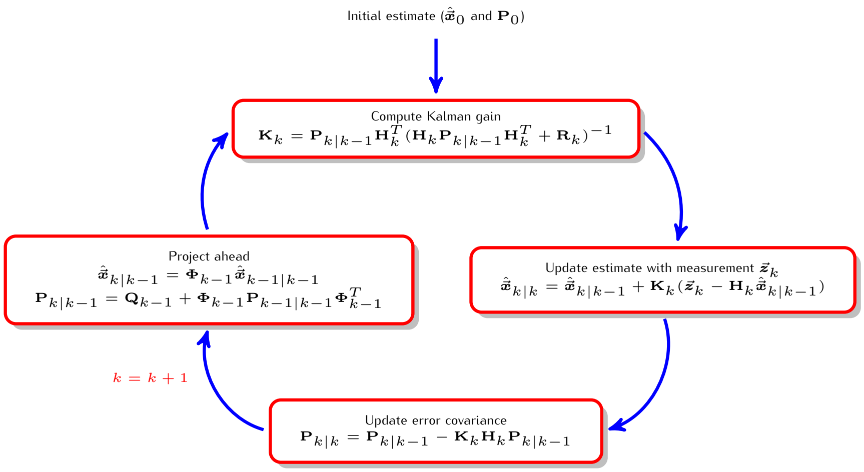
where $\mathbf{K}_k$ is $(n \times m)$ Kalman gain, and $(\vec{\boldsymbol{z}}_k - \mathbf{H}_k\hat{\vec{\boldsymbol{x}}}_{k|k-1})$ is the measurement innovation. .9

Kalman Gain

$$\mathbf{K}_k = \mathbf{P}_{k|k-1}\mathbf{H}_k^T(\ \mathbf{H}_k\mathbf{P}_{k|k-1}\mathbf{H}_k^T + \mathbf{R}_k\ )^{-1} \tag{14}$$

.10

Kalman filter data flow



.11

System Model

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

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

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

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

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

where $\mathbf{F}_{k-1}$ is the average of $\mathbf{F}$ at times t and $t - \tau_s$. .12

System Model Discretization

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

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

Therefore,

$$\mathbf{\Phi}_{k-1} = e^{\mathbf{F}_{k-1}\tau_s} \approx \mathbf{I} + \mathbf{F}_{k-1}\tau_s \tag{18}$$

where $\mathbf{F}_{k-1}$ is the average of $\mathbf{F}$ at times t and $t - \tau_s$, and first order approximation is used. .13

Discrete Covariance Matrix $\mathbf{Q}_k$

Assuming white noise, small time step, $\mathbf{G}$ is constant over the integration period, and the trapezoidal integration

$$\mathbf{Q}_{k-1} \approx \frac{1}{2}\left[\mathbf{\Phi}_{k-1}\mathbf{G}_{k-1}\mathbf{Q}(t_{k-1})\mathbf{G}_{k-1}^T\mathbf{\Phi}_{k-1}^T + \mathbf{G}_{k-1}\mathbf{Q}(t_{k-1})\mathbf{G}_{k-1}^T\right]\tau_s \tag{19}$$

where

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

.14

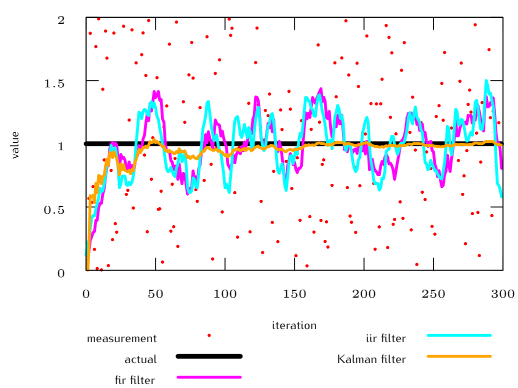
4 Example

Example: random constant

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

Design a Kalman filter to estimate x_k

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



.15

Example: first order Markov noise

State Equation

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

Autocorrelation Function

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

where

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

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

and T_c is the correlation time. .16

Discrete First Order Markov Noise

State Equation

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

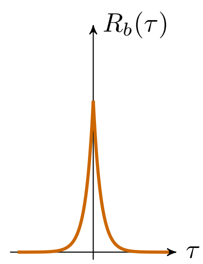
System Covariance Matrix

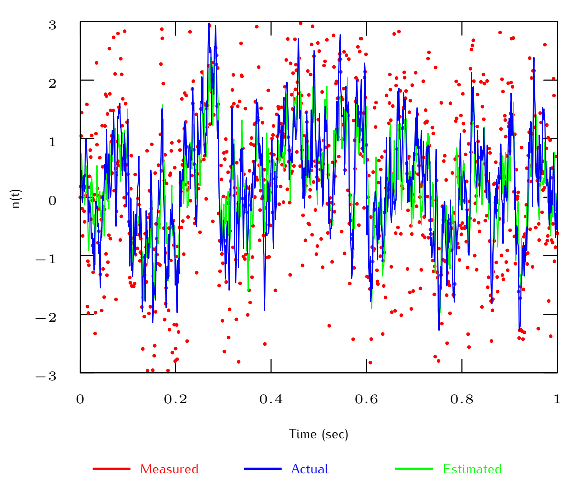
$$Q = \sigma_{BI}^2 [1 - e^{-\frac{2}{T_c}\tau_s}] \qquad (26)$$

.17

Autocorrelation of 1st order Markov .18

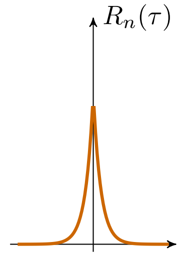
Small Correlation Time $T_c = 0.01$

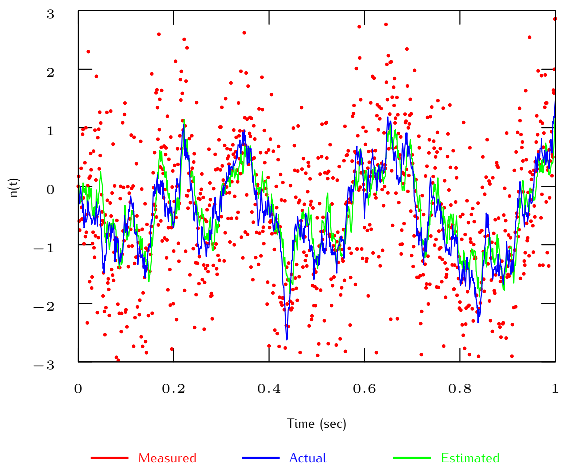




.19

Larger Correlation Time $T_c = 0.1$





.20

5 EKF

Linearized System

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

where

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

.21

Sequential Processing

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 $i = 1, 2, \ldots, r$

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

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

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

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

.22

Observability

The system is observable if the observability matrix

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

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. .23

A Better Observability Measure

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. .24

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.

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

.25

6 Other Solutions

Unscented Kalman Filter (UKF)

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. .26

Particle Filter

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. .27

7 References

Bayesian Filtering: From Kalman Filters to Particle Filters, and Beyond, by Zhe Chen .28