Lecture

On-Line Bayesian Tracking

EE 570: Location and Navigation

Lecture Notes Update on April 11, 2011

Aly El-Osery, Electrical Engineering Dept., New Mexico Tech

.1

Objective

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

.2

1 Problem Statement

Given State-Space Equations

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

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

where $\vec{x}_k$ 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 $\vec{w}_k$ and $\vec{v}_k$ i.i.d state noise. The state process is Markov chain, i.e., $p(\vec{x}_k|\vec{x}_1, \ldots, \vec{x}_{k-1}) = p(\vec{x}_k|\vec{x}_{k-1})$ and the distribution of $\vec{z}_k$ conditional on the state $\vec{x}_k$ is independent of previous state and measurement values, i.e., $p(\vec{z}_k|\vec{x}_{1:k}, \vec{z}_{1:k-1}) = p(\vec{z}_k|\vec{x}_k)$ .3

Objective

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

Optimal MMSE Estimate

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

in other words find the conditional mean

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

.4

2 Recursive Bayesian Estimation

Prediction Stage

$$p(\vec{x}_k|\vec{z}_{1:k-1}) = \int \; p(\vec{x}_k|\vec{x}_{k-1}) \;\; p(\vec{x}_{k-1}|\vec{z}_{1:k-1}) \; d\vec{x}_{k-1} \tag{5}$$

$p(\vec{x}_k|\vec{x}_{k-1})$ is defined using the state equation, and $p(\vec{x}_{k-1}|\vec{z}_{1:k-1})$ is assumed available from previous iteration. .5

Update Stage

Using Bayes' Rule

$$p(\vec{x}_k|\vec{z}_{1:k}) \;\; = \frac{p(\vec{z}_k|\vec{x}_k) \;\; p(\vec{x}_k|\vec{z}_{1:k-1})}{p(\vec{z}_k|\vec{z}_{1:k-1})} \tag{6}$$

where $p(\vec{x}_k|\vec{z}_{1:k})$ is the posterior distribution, $p(\vec{z}_k|\vec{x}_k)$ is the likelihood defined by the measurement equation and the statistics of the measurement noise, $p(\vec{x}_k|\vec{z}_{1:k-1})$ is prior distribution defined by the state equation and the statistics of the state noise, and $p(\vec{z}_k|\vec{z}_{1:k-1})$ is the evidence defined as $p(\vec{x}_k|\vec{z}_{1:k-1})d\vec{x}_k$ and depends on the likelihood function $p(\vec{z}_k|\vec{x}_k)$

The following is useful in deriving the above equation:

$$p(\vec{x}|\vec{y}) = \frac{p(\vec{y}|\vec{x})p(\vec{x})}{\int_X p(\vec{y}|\vec{x})p(\vec{x})d\vec{x}} \tag{7}$$

$$p(\vec{x}|\vec{y}) = \int_Z p(\vec{x},\vec{z}|\vec{y})d\vec{z} \tag{8}$$

$$\mathbb{E}_{p(\vec{x}|\vec{y})}\{g(\vec{x})\} = \int_X p(\vec{x}|\vec{y})g(\vec{x})d\vec{x} \tag{9}$$

.6

Limitations

1. Need to keep track of all previous states.
2. Generally can't be determined analytically. .7

3 Kalman Filter

Assumptions

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

$$\mathbb{E}\{\vec{w}_k\vec{w}_i^T\} = \begin{cases} \mathbf{Q}_k & i = k \\ 0 & i \neq k \end{cases} \tag{10}$$

$$\mathbb{E}\{\vec{v}_k\vec{v}_i^T\} = \begin{cases} \mathbf{R}_k & i = k \\ 0 & i \neq k \end{cases} \tag{11}$$

$$\mathbb{E}\{\vec{w}_k\vec{v}_i^T\} = \begin{cases} 0 & \forall i,k \end{cases} \tag{12}$$

.8

Assumptions

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

$$\vec{x}_k = \;\; \mathbf{\Phi}_{k-1} \; \vec{x}_{k-1} + \vec{w}_{k-1} \tag{13}$$

$$\vec{y}_k = \;\; \mathbf{H}_k \; \vec{x}_k + \vec{v}_k \tag{14}$$

where $\mathbf{\Phi}_{k-1}$ is $(n \times n)$ transition matrix relating $\vec{x}_{k-1}$ to $\vec{x}_k$, $\mathbf{H}_k$ is $(m \times n)$ matrix provides noiseless connection between measurement and state vectors. .9

pdf notation

$$p(\vec{x}_{k-1}|\vec{z}_{1:k-1}) = \mathcal{N}(\vec{x}_{k-1}; \vec{m}_{k-1|k-1},\ \mathbf{P}_{k-1|k-1}\ ) \tag{15}$$

$$p(\vec{x}_{k}|\vec{z}_{1:k-1}) = \mathcal{N}(\vec{x}_{k}; \vec{m}_{k|k-1}, \mathbf{P}_{k|k-1}) \tag{16}$$

$$p(\vec{x}_{k}|\vec{z}_{1:k}) = \mathcal{N}(\vec{x}_{k}; \vec{m}_{k|k},\ \mathbf{P}_{k|k}\ ) \tag{17}$$

where

$$\vec{m}_{k|k-1} = \mathbf{\Phi}_{k-1}\vec{m}_{k-1|k-1} \tag{18}$$

$$\mathbf{P}_{k|k-1} = \mathbf{Q}_{k-1} + \mathbf{\Phi}_{k-1}\mathbf{P}_{k-1|k-1}\mathbf{\Phi}_{k-1}^T \tag{19}$$

$$\vec{m}_{k|k} = \vec{m}_{k|k-1} + \mathbf{K}_k(\vec{z}_k - \mathbf{H}_k\vec{m}_{k|k-1}) \tag{20}$$

$$\mathbf{P}_{k|k} = \mathbf{P}_{k|k-1} - \mathbf{K}_k\mathbf{H}_k\mathbf{P}_{k|k-1} \tag{21}$$

where $\mathbf{P}_{k-1|k-1}$ is a priori error covariance ($\mathbb{E}\{(\hat{\vec{x}}_{k-1|k-1} - \vec{x}_{k-1})(\hat{\vec{x}}_{k-1|k-1} - \vec{x}_{k-1})^T\}$), diagonal terms are the variances in the state estimates off-diagonal show correlation between the errors in the different states, $\mathbf{P}_{k|k}$ is posterior error covariance ($\mathbb{E}\{(\hat{\vec{x}}_{k|k} - \vec{x}_k)(\hat{\vec{x}}_{k|k} - \vec{x}_k)^T\}$). .10

State-Space Equations

$$\hat{\vec{x}}_{k|k-1} = \mathbf{\Phi}_{k-1}\hat{\vec{x}}_{k-1|k-1} \tag{22}$$

$$\mathbf{P}_{k|k-1} = \mathbf{Q}_{k-1} + \mathbf{\Phi}_{k-1}\mathbf{P}_{k-1|k-1}\mathbf{\Phi}_{k-1}^T \tag{23}$$

$$\hat{\vec{x}}_{k|k} = \hat{\vec{x}}_{k|k-1} +\ \mathbf{K}_k\ \ (\vec{z}_k - \mathbf{H}_k\hat{\vec{x}}_{k|k-1}) \tag{24}$$

$$\mathbf{P}_{k|k} = \mathbf{P}_{k|k-1}(\mathbf{I} - \mathbf{K}_k\mathbf{H}_k\mathbf{P}_{k|k-1}) \tag{25}$$

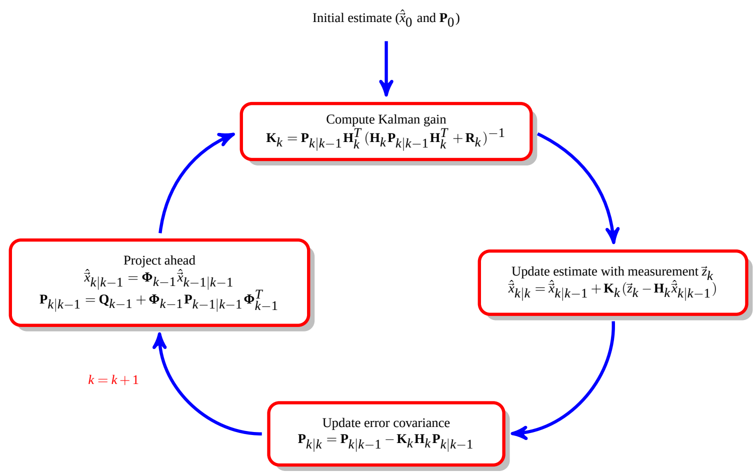
where $\mathbf{K}_k$ is $(n \times m)$ Kalman gain, and $(\vec{z}_k - \mathbf{H}_k\hat{\vec{x}}_{k|k-1})$ is the measurement innovation. .11

Kalman Gain

$$\mathbf{K}_k = \mathbf{P}_{k|k-1}\mathbf{H}_k^T(\ \mathbf{H}_k\mathbf{P}_{k|k-1}\mathbf{H}_k^T + \mathbf{R}_k\ )^{-1} \tag{26}$$

.12

Kalman filter data flow



.13

Observability

The system is observable if the observability matrix

$$\mathscr{O}(k) = \begin{bmatrix} H(k-n+1) \\ H(k-n-2)\mathbf{\Phi}(k-n+1) \\ \vdots \\ H(k)\mathbf{\Phi}(k-1)\ldots\mathbf{\Phi}(k-n+1) \end{bmatrix} \tag{27}$$

where n is the number of states, has a rank of n. The rank of $\mathscr{O}$ 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. .14

A Better Observability Measure

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

$$\mathscr{O} = U\Sigma V^* \tag{28}$$

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. .15

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.

$$P_{k|k} = (I - K_k H_k)\,P_{k|k-1}\,(I - K_k H_k)^T + K_k R_k K_k^T \tag{29}$$

.16

System Model

$$\dot{\vec{x}}(t) = \mathbf{F}(t)\vec{x}(t) + \mathbf{G}(t)\vec{w}(t) \tag{30}$$

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

$$\mathbb{E}\{\dot{\vec{x}}(t)\} = \frac{\partial}{\partial t}\hat{\vec{x}}(t) = \mathbf{F}(t)\hat{\vec{x}}(t) \tag{31}$$

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

$$\hat{\vec{x}}(t) = e^{\left(\int_{t-\tau_s}^{t}\mathbf{F}(t')dt'\right)}\hat{\vec{x}}(t-\tau_s) \tag{32}$$

where $\mathbf{F}_{k-1}$ is the average of $\mathbf{F}$ at times t and $t-\tau_s$. .17

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,

$$\mathbf{\Phi}_{k-1} = e^{\mathbf{F}_{k-1}\tau_s} \approx \mathbf{I} + \mathbf{F}_{k-1}\tau_s \tag{33}$$

where $\mathbf{F}_{k-1}$ is the average of $\mathbf{F}$ at times t and $t-\tau_s$, and first order approximation is used. .18

Discrete Covariance Matrix $\mathbf{Q}_k$

The solution to (30) is

$$\vec{x}_k = \mathbf{\Phi}_{k-1}\vec{x}_{k-1} + \int_{t-\tau_s}^{t} e^{\mathbf{F}_{k-1}(t-\eta)}\mathbf{G}_{k-1}\vec{w}(\eta)d\eta \tag{34}$$

where $\mathbf{G}_{k-1}$ is the average of $\mathbf{G}$ at times t and $t-\tau_s$. Now let's look at the error covariance matrix

$$\mathbf{P}_{k|k-1} = \mathbb{E}\{(\hat{\vec{x}}_{k|k-1} - \vec{x}_k)(\hat{\vec{x}}_{k|k-1} - \vec{x}_k)^T\} \tag{35}$$

where

$$\hat{\vec{x}}_{k|k-1} - \vec{x}_k = \mathbf{\Phi}_{k-1}(\hat{\vec{x}}_{k-1|k-1} - \vec{x}_k) - \int_{t-\tau_s}^{t} e^{\mathbf{F}_{k-1}(t-\eta)}\mathbf{G}_{k-1}\vec{w}(\eta)d\eta \tag{36}$$

.19

Discrete Covariance Matrix $\mathbf{Q}_k$ (cont.)

Since the state estimate errors and the system noise $\vec{w}(t)$ are uncorrelated

$$\begin{aligned}\mathbf{P}_{k|k-1} =& \mathbf{\Phi}_{k-1}\mathbf{P}_{k-1|k-1}\mathbf{\Phi}_{k-1}^T + \\ & \mathbb{E}\left\{\iint_{t-\tau_s}^{t} e^{F_{k-1}(t-\eta)}\mathbf{G}_{k-1}\vec{w}(\eta)\vec{w}^T(\zeta)\mathbf{G}_{k-1}^T e^{\mathbf{F}_{k-1}^T(t-\zeta)}d\eta d\zeta\right\} \\ =& \mathbf{\Phi}_{k-1}\mathbf{P}_{k-1|k-1}\mathbf{\Phi}_{k-1}^T + \\ & \iint_{t-\tau_s}^{t} e^{F_{k-1}(t-\eta)}\mathbf{G}_{k-1}\mathbb{E}\{\vec{w}(\eta)\vec{w}^T(\zeta)\}\mathbf{G}_{k-1}^T e^{\mathbf{F}_{k-1}^T(t-\zeta)}d\eta d\zeta \\ =& \mathbf{\Phi}_{k-1}\mathbf{P}_{k-1|k-1}\mathbf{\Phi}_{k-1}^T + \mathbf{Q}_{k-1}\end{aligned} \tag{37}$$

.20

Discrete Covariance Matrix $\mathbf{Q}_k$ (cont.)

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

$$\mathbf{Q}_{k-1} \approx \frac{1}{2}\left[\mathbf{\Phi}_{k-1}\mathbf{G}_{k-1}\mathbf{Q}(t_{k-1})\mathbf{G}_{k-1}^T\mathbf{\Phi}_{k-1}^T + \mathbf{G}_{k-1}\mathbf{Q}(t_{k-1})\mathbf{G}_{k-1}^T\right]\tau_s \tag{38}$$

where

$$\mathbb{E}\{\vec{w}(\eta)\vec{w}^T(\zeta)\} = \mathbf{Q}(\eta)\delta(\eta-\zeta) \tag{39}$$

.21

4 Extended Kalman Filter (EKF)

Linearized System

$$\mathbf{F}_k = \left.\frac{\partial\mathbf{f}(\vec{x})}{\partial\vec{x}}\right|_{\vec{x}=\hat{\vec{x}}_{k|k-1}}, \qquad \mathbf{H}_k = \left.\frac{\partial\mathbf{h}(\vec{x})}{\partial\vec{x}}\right|_{\vec{x}=\hat{\vec{x}}_{k|k-1}} \tag{40}$$

where

$$\frac{\partial\mathbf{f}(\vec{x})}{\partial\vec{x}} = \begin{pmatrix}\frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots \\ \vdots & \ddots & \vdots\end{pmatrix}, \qquad \frac{\partial\mathbf{h}(\vec{x})}{\partial\vec{x}} = \begin{pmatrix}\frac{\partial h_1}{\partial x_1} & \frac{\partial h_1}{\partial x_2} & \cdots \\ \frac{\partial h_2}{\partial x_1} & \frac{\partial h_2}{\partial x_2} & \cdots \\ \vdots & \ddots & \vdots\end{pmatrix} \tag{41}$$

.22

5 Example

First Order Markov Noise

State Equation

$$\dot{n}(t) = -\frac{1}{T_c} n(t) + w(t) \tag{42}$$

Autocorrelation Function

$$\mathbb{E}\{n(t)(n(t+\tau)\} = \sigma^2 e^{-|\tau|/T_c} \tag{43}$$

where

$$\mathbb{E}\{w(t)w(t+\tau)\} = Q(t)\delta(t-\tau) \tag{44}$$

$$Q(t) = \frac{2\sigma^2}{T_c} \tag{45}$$

and T_c is the correlation time. .23

Discrete First Order Markov Noise

State Equation

$$n_k = e^{-\frac{1}{T_c}\tau_s} n_{k-1} + w_{k-1} \tag{46}$$

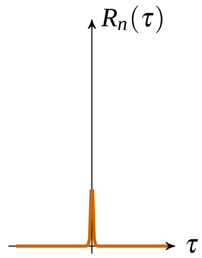
System Covariance Matrix

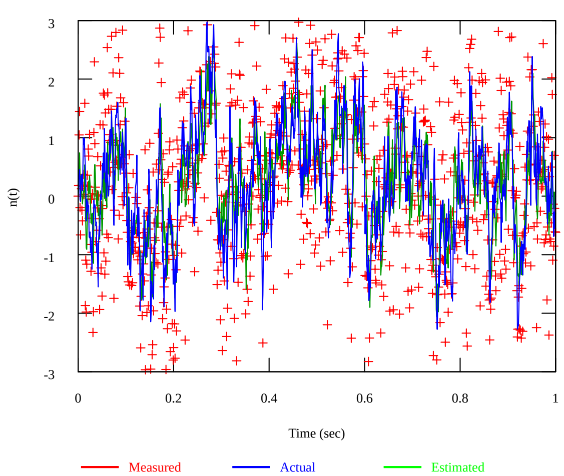
$$Q = \sigma^2 [1 - e^{-\frac{2}{T_c}\tau_s}] \tag{47}$$

.24

Autocorrelation of 1st order Markov .25

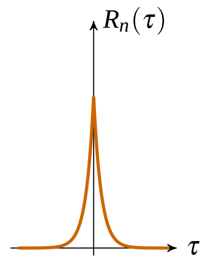
Small Correlation Time $T_c = 0.01$

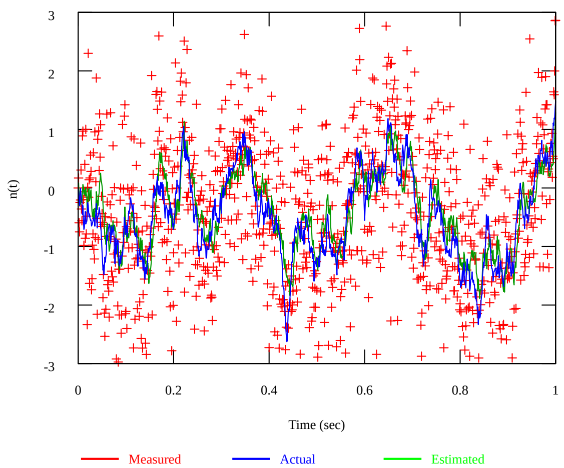




.26

Larger Correlation Time $T_c = 0.1$





.27

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. .28

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. .29

7 References

Bayesian Filtering: From Kalman Filters to Particle Filters, and Beyond, by Zhe Chen .30