6. Recursive Estimation: Kalman and Particle Filter

6.1 Recursive LMMSE Estimation I

Consider noisy measurements

$$Y_n = nX + V_n \quad n = 0, 1, 2, \ldots,$$

where $X \in \mathbb{R}$ is a random variable with unit variance $\sigma_X^2 = 1$ and zero-mean $\mathrm{E}[X] = 0$. The measurement noise $V_n \in \mathbb{R}$ is a random sequence which is uncorrelated with X and has unit variance $\sigma_{V_n}^2 = 1$ and zero-mean $\mathrm{E}[V_n] = 0$. In addition, V_i and V_j are uncorrelated for $i \neq j$. The state-space model, where $X_n = X$ is the constant state variable, is given as

$$\begin{aligned} X_{n+1} &= X_n = X \\ Y_n &= nX + V_n \end{aligned}$$

for $n = 0, 1, 2, \ldots$. Let $\hat{x}_{n|n}$ denote the LMMSE estimate for X_n based on $\boldsymbol{Y}_{(n)} = [Y_1, \ldots, Y_n]^{\mathrm{T}} = \boldsymbol{y}_{(n)}$.

a) Determine $\boldsymbol{h}_n \in \mathbb{R}^n$ such that $\boldsymbol{Y}_{(n)}$ can be expressed as $\boldsymbol{Y}_{(n)} = \boldsymbol{h}_n X + \boldsymbol{V}_{(n)}$, where $\boldsymbol{V}_{(n)} = [V_1, \ldots, V_n]^{\mathrm{T}}$.

b) Compute the LMMSE estimate $\hat{x}_{n|n}$ for X_n based on $\boldsymbol{Y}_{(n)} = \boldsymbol{y}_{(n)}$.
Hint: Use that

$$\boldsymbol{h}_n^{\mathrm{T}} \left(\boldsymbol{h}_n \boldsymbol{h}_n^{\mathrm{T}} + \boldsymbol{I} \right)^{-1} = \frac{\boldsymbol{h}_n^{\mathrm{T}}}{1 + \|\boldsymbol{h}_n\|^2}.$$

c) Show that The LMMSE estimate $\hat{x}_{n|n}$ can be alternatively computed recursively using

$$\hat{x}_{n|n} = \alpha_n \hat{x}_{n-1|n-1} + \beta_n y_n$$

with some $\alpha_n \in \mathbb{R}$ and $\beta_n \in \mathbb{R}$.

d) State the prediction and update step of Kalman filter to estimate X_n based on $\boldsymbol{Y}_{(n)} = \boldsymbol{y}_{(n)}$.

e) Compute the Kalman gain k_n for an estimation of X_n based on $\boldsymbol{Y}_{(n)} = \boldsymbol{y}_{(n)}$.

f) What happens to k_n if $n \to \infty$? No calculations are required, but justify your answer and give an interpretation of your result.

6.2 The Extended Kalman Filter

Consider the time invariant state-space model given as

$$X_{n+1} = g(X_n) = X_n^2,$$
$$Y_n = h(X_n) + W_n = X_n + W_n,$$

where $X_n \in \mathbb{R}$, $Y_n \in \mathbb{R}$, and $W_n \in \mathbb{R}$ are scalar random variables. The initial state X_0 follows a uniform distribution over $(-1, 1)$ and the measurement noise follows a uniform distribution over $(-0.5, 0.5)$. We assume that W_n and the initial state X_0 are uncorrelated. In the following, we consider the prediction of X_1 given $Y_0 = y_0$.

The so-called extended Kalman filter (EKF) equations for the scalar case are given as

$$\hat{x}^{\text{EKF}}_{n+1|n} = g(\hat{x}^{\text{EKF}}_{n|n}),$$
$$\hat{x}^{\text{EKF}}_{n|n} = \hat{x}^{\text{EKF}}_{n|n-1} + k_n \left(y_0 - h(\hat{x}^{\text{EKF}}_{n|n-1}) \right).$$

For the computation of the Kalman gain K_n and the error covariance matrices, the EKF uses the linearizations

$$\hat{g}_n = \left. \frac{\partial g(x)}{\partial x} \right|_{x=\hat{x}_{n|n}} \quad \text{and} \quad \hat{h}_n = \left. \frac{\partial h(x)}{\partial x} \right|_{x=\hat{x}_{n|n-1}}$$

i.e., the standard formulas of the KF are used with $\hat{g}_n$ and $\hat{h}_n$. Like the KF, the EKF is initialized with $\hat{x}^{\text{EKF}}_{0|-1} = \mathrm{E}[X_0] = 0$ and $\sigma^2_{X_{0|-1}} = \mathrm{Var}[X_0] = \frac{4}{12}$. In addition, we have $\sigma^2_{W_n} = \frac{1}{12}$.

a) Compute the estimate $\hat{x}^{\text{EKF}}_{0|0}$ of X_0 given $Y_0 = y_0$ using an EKF.

b) Compute the estimate $\hat{x}^{\text{EKF}}_{1|0}$ of X_1 given $Y_0 = y_0$ using an EKF.

c) If $y_0 = -1.5$, we know that $x_0 = -1$ and hence $x_1 = 1$ and if $y_0 = 1.5$, we know that $x_0 = 1$ and thus $x_1 = 1$. Check if $\hat{x}^{\text{EKF}}_{0|0}$ and $\hat{x}^{\text{EKF}}_{1|0}$ obtain these estimates or not.

The conditional PDF of X_0 given the observation $Y_0 = y_0 \in (-1.5, 1.5)$ is given as

$$f_{X_0|Y_0}(x_0|y_0) = \begin{cases} \frac{1}{1.5+y_0} & \text{for } x_0 \in (-1, y_0 + 0.5) \text{ if } y_0 \in (-1.5, -0.5), \\ 1 & \text{for } x_0 \in (-0.5 + y_0, 0.5 + y_0) \text{ if } y_0 \in (-0.5, 0.5), \\ \frac{1}{1.5-y_0} & \text{for } x_0 \in (y_0 - 0.5, 1) \text{ if } y_0 \in (0.5, 1.5), \\ 0 & \text{otherwise.} \end{cases}$$

d) Compute the MMSE estimate $\hat{x}^{\text{MMSE}}_{0|0}$ of X_0 given Y_0.
Hint: http://www.wolframalpha.com/ helps to simply the expressions.

e) Compute the MMSE estimate $\hat{x}^{\text{MMSE}}_{1|0}$ of X_1 given Y_0.
Hint: http://www.wolframalpha.com/ helps to simply the expressions.

f) Sketch and compare $\hat{x}^{\text{EKF}}_{1|0}$ and $\hat{x}^{\text{MMSE}}_{1|0}$. Give an interpretation of your result.

6.3 Recursive LMMSE Estimation II

Consider the linear state-space model given as

$$\boldsymbol{X}_n = \boldsymbol{G}_n \boldsymbol{X}_{n-1} + \boldsymbol{V}_n,$$
$$\boldsymbol{Y}_n = \boldsymbol{H}_n \boldsymbol{X}_n + \boldsymbol{W}_n,$$

where $\boldsymbol{X}_n \in \mathbb{R}^q$, $\boldsymbol{Y}_n \in \mathbb{R}^p$, $\boldsymbol{V}_n \in \mathbb{R}^q$, and $\boldsymbol{W}_n \in \mathbb{R}^p$ are random vectors. The state transition and measurement functions are linear and given by the matrices $\boldsymbol{G}_n \in \mathbb{R}^{q \times q}$ and $\boldsymbol{H}_n \in \mathbb{R}^{p \times p}$. The sequence $\boldsymbol{V}_n$ and the sequence $\boldsymbol{W}_n$ are assumed to be uncorrelated sequences of uncorrelated random vectors. In addition, we assume that $\boldsymbol{V}_n$, $\boldsymbol{W}_n$, and the initial state $\boldsymbol{X}_0$ are pairwise uncorrelated. $\boldsymbol{V}_n$ and $\boldsymbol{W}_n$ have zero-mean and the covariance matrices $\boldsymbol{C}_{\boldsymbol{V}_n}$ and $\boldsymbol{C}_{\boldsymbol{W}_n}$. Note that $\boldsymbol{V}_n$ and $\boldsymbol{W}_n$ are not necessarily sequences of Gaussian random vectors. Finally, we assume that the initial mean $\boldsymbol{\mu}_{\boldsymbol{X}_0} = \mathrm{E}[\boldsymbol{X}_0] = \boldsymbol{0}$ and the initial covariance matrix $\boldsymbol{C}_{\boldsymbol{X}_0} = \mathrm{E}\left[\boldsymbol{X}_0 \boldsymbol{X}_0^{\mathrm{T}}\right]$ of the state vector are known. A compact notation for the correlations between the variables is given by

$$\mathrm{Cov}\left[\begin{bmatrix} \boldsymbol{V}_n \\ \boldsymbol{W}_m \\ \boldsymbol{X}_0 \end{bmatrix}, \begin{bmatrix} \boldsymbol{V}_k \\ \boldsymbol{W}_l \\ \boldsymbol{X}_0 \\ 1 \end{bmatrix}\right] = \begin{bmatrix} \boldsymbol{C}_{\boldsymbol{V}_n}\delta(n-k) & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{C}_{\boldsymbol{W}_m}\delta(m-l) & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{C}_{\boldsymbol{X}_0} & \boldsymbol{0} \end{bmatrix}.$$

In the following, we will derive a recursive LMMSE estimator for the states $\boldsymbol{X}_n$.

a) Define an innovation sequence $\Delta\boldsymbol{Y}_n$ using a measurement $\boldsymbol{Y}_n$ and a LMMSE estimate $\hat{\boldsymbol{Y}}_{n|n-1} = \mathrm{E}^L[\boldsymbol{Y}_n|\boldsymbol{Y}_{(n-1)}]$, where $\boldsymbol{Y}_{(n-1)} = [\boldsymbol{Y}_0, \ldots, \boldsymbol{Y}_{n-1}]$. Express $\Delta\boldsymbol{Y}_n$ in terms of $\boldsymbol{Y}_n$ and $\hat{x}_{n|n-1}$.

b) Using your result from a), show that the LMMSE estimate $\hat{\boldsymbol{x}}_{n|n} = \mathrm{E}^L[\boldsymbol{X}_n|\boldsymbol{Y}_{(n)}]$ for the LMMSE estimation of $\boldsymbol{X}_n$ based on $\boldsymbol{Y}_{(n)} = [\boldsymbol{Y}_0, \ldots, \boldsymbol{Y}_n]$ can be expressed as

$$\hat{\boldsymbol{x}}_{n|n} = \mathrm{E}^L[\boldsymbol{X}_n|\boldsymbol{Y}_{(n)}] = \hat{\boldsymbol{x}}_{n|n-1} + \boldsymbol{K}_n \Delta\boldsymbol{y}_n,$$

where $\hat{\boldsymbol{x}}_{n|n-1} = \mathrm{E}^L[\boldsymbol{X}_n|\boldsymbol{Y}_{(n-1)}]$ is the LMMSE estimate of of $\boldsymbol{X}_n$ based on $\boldsymbol{Y}_{(n-1)}$
Hint: For uncorrelated random variables $\boldsymbol{Z}_1$ and $\boldsymbol{Z}_2$ with zero-mean and a zero-mean random variable $\boldsymbol{X}$, it follows that the LMMSE estimate of $\boldsymbol{X}$ based on $\boldsymbol{Z}_1$ and $\boldsymbol{Z}_2$ can be expressed as

$$\mathrm{E}^L[\boldsymbol{X}|\boldsymbol{Z}_1\boldsymbol{Z}_2] = \mathrm{E}^L[\boldsymbol{X}|\boldsymbol{Z}_1] + \mathrm{E}^L[\boldsymbol{X}|\boldsymbol{Z}_2].$$

c) Show that the LMMSE estimate $\hat{\boldsymbol{x}}_{n|n-1} = \mathrm{E}^L[\boldsymbol{X}_n|\boldsymbol{Y}_{(n-1)}]$ of the state $\boldsymbol{X}_n$ based on $\boldsymbol{Y}_{(n-1)}$ can be expressed as

$$\hat{\boldsymbol{x}}_{n|n-1} = \boldsymbol{G}_n \hat{\boldsymbol{x}}_{n-1|n-1}.$$

d) Let $\Delta\boldsymbol{X}_{n|n-1} = \boldsymbol{X}_n - \hat{\boldsymbol{x}}_{n|n-1}$ be the error of the prediction of the state $\boldsymbol{X}_n$ based on $\boldsymbol{Y}_{(n-1)}$. Show that

$$\boldsymbol{C}_{\Delta\boldsymbol{Y}_{n|n-1}} = \boldsymbol{H}_n \boldsymbol{C}_{\Delta\boldsymbol{X}_{n|n-1}} \boldsymbol{H}_n^{\mathrm{T}} + \boldsymbol{C}_{\boldsymbol{W}_n},$$

$$\boldsymbol{C}_{\boldsymbol{X}_n \Delta \boldsymbol{Y}_{n|n-1}} = \boldsymbol{C}_{\Delta \boldsymbol{X}_{n|n-1}} \boldsymbol{H}_n^{\mathrm{T}}.$$

Note: In the lecture notes, we used the short notation $\boldsymbol{C}_{\Delta \boldsymbol{X}_{n|n-1}} = \boldsymbol{C}_{\boldsymbol{X}_{n|n-1}}$.

e) In order to determine $\boldsymbol{K}_n$, we need an expression for $\boldsymbol{C}_{\Delta \boldsymbol{X}_{n|n-1}}$. Let $\Delta \boldsymbol{X}_{n|n} = \boldsymbol{X}_n - \hat{\boldsymbol{x}}_{n|n}$ and show that

$$\boldsymbol{C}_{\Delta \boldsymbol{X}_{n|n-1}} = \boldsymbol{G}_n \boldsymbol{C}_{\Delta \boldsymbol{X}_{n-1|n-1}} \boldsymbol{G}_n^{\mathrm{T}} + \boldsymbol{C}_{\boldsymbol{V}_n}.$$

f) As a last step for our recursive LMMSE estimator $\hat{\boldsymbol{x}}_{n|n} = \mathrm{E}^L[\boldsymbol{X}_n | \boldsymbol{Y}_{(n)}]$ we need an expression for $\boldsymbol{C}_{\Delta \boldsymbol{X}_{n-1|n-1}}$. To this end, show that

$$\boldsymbol{C}_{\Delta \boldsymbol{X}_{n|n}} = (\boldsymbol{I} - \boldsymbol{K}_n \boldsymbol{H}_n)\, \boldsymbol{C}_{\Delta \boldsymbol{X}_{n|n-1}}.$$

g) Finally, collect all the expressions from sub-problems a) to f) into the prediction step where $\boldsymbol{X}_n$ is estimated based on $\boldsymbol{Y}_{(n-1)}$ using the LMMSE estimate $\hat{\boldsymbol{x}}_{n|n-1}$ and the correction step where $\boldsymbol{X}_n$ is estimated based on $\boldsymbol{Y}_{(n)}$ using the LMMSE estimate $\hat{\boldsymbol{x}}_{n|n}$.

h) What happens to our LMMSE estimator if the random variables $\boldsymbol{X}_0$, $\boldsymbol{V}_n$, and $\boldsymbol{W}_n$ are Gaussian?

6.4 The Particle Filter

The state equations of a hidden Markov model with real-valued and scalar states X_n and real-valued and scalar observations Y_n are

$$X_{n+1} = X_n + V_{n+1}$$
$$Y_{n+1} = X_{n+1} + W_{n+1}$$

with $X_0 \sim f_{X_0}$, $V_n \sim f_V$ for all n, and $W_n \sim f_W$ for all n. In addition, we have stochastical independence of V_n and V_m for $n \neq m$, of W_n and W_m for $n \neq m$, and of V_n and W_m for all n and m. Let $Y_{(n)}$ denote the observations Y_1 to Y_n, i.e., $Y_{(n)} = [Y_1, \ldots, Y_n]$. We are running a particle filter (PF) in order to approximate the conditional probability density function (PDF) $f_{X_n|Y_{(n)}}(x_n|y_{(n)})$ using $f_{X_n|Y_{(n)}}(x_n|y_{(n)}) \approx \sum_{i=1}^4 w_n^i \delta(x_n - x_n^i)$. In the nth step, we have the particles

$$x_n^1 = 2, \quad x_n^2 = -2, \quad x_n^3 = 1, \quad x_n^4 = 0,$$

with corresponding normalized weights

$$w_n^1 = \frac{2}{10}, \quad w_n^2 = \frac{3}{10}, \quad w_n^3 = \frac{4}{10}, \quad w_n^4 = \frac{1}{10}.$$

a) Draw the approximation of the conditional PDF $f_{X_n|Y_{(n)}}(x_n|y_{(n)})$.

b) Compute the approximation of the conditional mean estimator $\hat{x}_{n|n} = \mathrm{E}[X_n|Y_{(n)}]$ based on the PF approximation of $f_{X_n|Y_{(n)}}(x_n|y_{(n)})$.

c) Compute the approximation of the maximum a posteriori (MAP) estimator defined as

$$\hat{x}_{n|n}^{\mathrm{MAP}} = \underset{x_n}{\operatorname{argmax}} \left\{ f_{X_n|Y_{(n)}}(x_n|y_{(n)}) \right\}$$

based on the PF approximation of $f_{X_n|Y_{(n)}}(x_n|y_{(n)})$.

The PF propagates the particles using the suboptimal importance density

$$q_{X_{n+1}|X_n, Y_{n+1}}(x_{n+1}|x_n^i, y_{n+1}) = f_{X_{n+1}|X_n}(x_{n+1}|x_n^i).$$

d) State the update rule for the particle weights $\tilde{w}_{n+1}^i$.

e) State the update rule for the normalized particle weights w_{n+1}^i.

f) Express the likelihood $f_{Y_{n+1}|X_{n+1}}(y_{n+1}|x_{n+1}^i)$ of the observation y_{n+1} given a particle x_{n+1}^i in terms of the PDF $f_W(\cdot)$ of the measurement noise.

Now, consider a measurement noise $W_{n+1} \sim f_W$ with a PDF given as

$$f_W(w_{n+1}) = \begin{cases} w_{n+1} + 1 & \text{if } -1 \leq w_{n+1} \leq 0, \\ 1 - w_{n+1} & \text{if } 0 < w_{n+1} \leq 1, \\ 0 & \text{otherwise.} \end{cases}$$

You make the observation $y_{n+1} = 0.5$ and the new particles are given as

$$x_{n+1}^1 = 1.5, \quad x_{n+1}^2 = 0.5, \quad x_{n+1}^3 = 1, \quad x_{n+1}^4 = -1.5.$$

g) Compute the normalized weights w_{n+1}^i of the particles x_{n+1}^i using the update rules from d) and e).

h) Check if particle weight degeneracy is detected with a threshold $w_{\text{thr}} = 2$.