Prediction in nonlinear dynamic systems: extended Kalman filter (discrete time)

Antonio Sala

Lecture Notes on Multivariable Control Systems

Dept. Ing. Sistemas y Automatica (DISA)
Universitat Politècnica de València (UPV)

Video-presentación disponible en:

http://personales.upv.es/asala/YT/V/**ekfteoEN**.html

Outline

Motivation:

The best prediction formulae, in linear processes, allow obtaining the minimum-error-variance state estimate $\hat{x}$, given a sequence of measurements y. This estimator (optimal observer) is the **Kalman filter**, and it is widely used. However, in many cases, the model's equations are **nonlinear**.

Objectives:

Understand the derivation of the equations of the extended Kalman filter (linearised around an estimated trajectory).

Contents:

Review of linear Kalman filter. Linealirization around estimated trajectory. Extended filter. Conclusions.

Discrete-time linear stochastic processes

Let us consider a discrete-time LTI system given by a state-space representation:

$$x_{k+1} = Ax_k + Bu_k + w_k$$
$$y_k = Cx_k + v_k$$

with $x \in \mathbb{R}^n$, $u \in \mathbb{R}^m$, $y \in \mathbb{R}^p$, $v \in \mathbb{R}^n$, $w \in \mathbb{R}^p$ being:

1. Vector x is the **state** vector,
2. Vector y is a **measurable output**,
3. Vector u is a **deterministic input** (known),
4. w is a **process** noise, $\{v_0, v_1, \ldots, v_k\}$, zero mean and variance W, uncorrelated with x.
5. v is a **measurement** noise, $\{w_0, w_1, \ldots, w_k\}$, zero mean and variance V, uncorrelated with x.

Linear Kalman filter, propagation of expected value:

$$\underbrace{\Sigma_{x_k} = A \cdot \overbrace{\Sigma_{x_{k-1},posteriori}}^{\text{variance after measuring } y_{k-1}} \cdot A^T + W}_{\text{prediction (variance) before measuring } y_k}$$

- Best linear prediction of x_k given y_k (+ prediction $\hat{x}_{k-1}$, $\Sigma_{x_{k-1},p}$, recurrent):

$$\hat{x}_k = \bar{x}_k + \Sigma_{x_k y_k} \Sigma_{y_k}{}^{-1}(y_k - \bar{y}_k)$$

$$= \underbrace{(A\hat{x}_{k-1} + Bu_{k-1})}_{\text{prediction (mean)}} + \underbrace{\underbrace{\Sigma_{x_k} C^T (C\Sigma_{x_k} C^T + V)^{-1}}_{\text{observer gain, } L_k} \underbrace{(y_k - C\overbrace{(A\hat{x}_{k-1} + Bu_{k-1})}^{\bar{x}_k})}_{\text{innovation error}}}_{\text{correction}}$$

*It's an **observer** with time-varying gain, depending on the precision of the state before measuring, i.e., Σ_{x_k}.

Linear Kalman Filter, variance update:

- **A posteriori** variance of estimation error $(x - \hat{x})$:

$$\boxed{\Sigma_{x_k,posteriori}} = \Sigma_{x_k} - \Sigma_{x_k y_k}\Sigma_{y_k}^{-1}\Sigma_{x_k y_k}^T$$

$$= \boxed{\Sigma_{x_k} - \underbrace{\Sigma_{x_k} C^T (C\Sigma_{x_k} C^T + V)^{-1} C\Sigma_{x_k}}_{\text{variance reduction by sensor info}}}$$

*Using $\Sigma_{x_k} = A\Sigma_{x_{k-1},posteriori}A^T + W$.

Recursive implementation: repeat everything with $\hat{x}_k$, $\Sigma_{x_k,posteriori}$, when next measurement (y_{k+1}) becomes available.

Stationary Kalman filter: After a handful of samples, variance equations and observer gain converge in LTI case. Many implementations are a stationary filter (Matlab's `dlqe`, $\mathcal{H}_2$ control).

Nonlinear processes

Consider a discrete-time system in state-space form given by:

$$x_{k+1} = f(x_k, u_k, w_k)$$
$$y_k = h(x_k, v_k)$$

with $x \in \mathbb{R}^n$, $u \in \mathbb{R}^m$, $y \in \mathbb{R}^p$, $v \in \mathbb{R}^n$, $w \in \mathbb{R}^p$ being:

1. Vector x is the **state** vector,
2. Vector y is a **measurable output**,
3. Vector u is a **deterministic input** (known),
4. w is a **process** noise, $\{v_0, v_1, \ldots, v_k\}$, zero mean and variance W, uncorrelated with x.
5. v is a **measurement** noise, $\{w_0, w_1, \ldots, w_k\}$, zero mean and variance V, uncorrelated with x.

Linearization around estimated trajectory

If $x_k \approx \hat{x}_k$, y $w_k \approx 0$, then

$$x_{k+1} = f(x_k, u_k, w_k)$$

$$\approx \underbrace{f(\hat{x}_k, u_k, 0)}_{\bar{x}_{k+1}} + \underbrace{\left.\frac{\partial f}{\partial x}\right|_{(\hat{x}_k, u_k, 0)}}_{A(\hat{x}_k, u_k)} \cdot \underbrace{(x_k - \hat{x}_k)}_{\substack{e_k \\ \text{error after measuring } y_k}} + \underbrace{\left.\frac{\partial f}{\partial w}\right|_{(\hat{x}_k, u_k, 0)}}_{G(\hat{x}_k, u_k)} \cdot w_k$$

so, the error "before measuring y_{k+1}", $\epsilon_{k+1} := (x_{k+1} - \bar{x}_{k+1})$, will be

$$\epsilon_{k+1} \approx \underbrace{A(\hat{x}_k, u_k)}_{A_k} \cdot e_k + \underbrace{G(\hat{x}_k, u_k)}_{G_k} w_k$$

and will have a variance given by (approx.):

$$\Sigma_{k+1} \approx A_k \cdot \Sigma_{k,posteriori} \cdot A_k^T + G_k \cdot W \cdot G_k^T$$

Incorporation of info from new measurement

Regarding output equation, linearising again, if $x_{k+1} \approx \bar{x}_{k+1}$, we will approximately have:

$$y_{k+1} = h(x_{k+1}, v_{k+1})$$
$$\approx \underbrace{h(\bar{x}_{k+1}, 0)}_{\bar{y}_{k+1}} + \underbrace{\left.\frac{\partial h}{\partial x}\right|_{(\bar{x}_{k+1},0)}}_{C(\bar{x}_{k+1})} \cdot \underbrace{(x_{k+1} - \bar{x}_{k+1})}_{\substack{\epsilon_{k+1} \\ \text{error before measuring } y_{k+1}}} + \underbrace{\left.\frac{\partial h}{\partial v}\right|_{(\bar{x}_{k+1},0)}}_{R(\bar{x}_{k+1})} \cdot v_{k+1}$$

hence, $y_{k+1} - \bar{y}_{k+1}$ will be (approx.) $\sim \mathcal{N}(0\,,\, C_{k+1}\Sigma_{k+1}C_{k+1}^T + R_{k+1}VR_{k+1}^T)$ and its covariance with ϵ_{k+1} will be approximated by $\Sigma_{y_{k+1}\epsilon_{k+1}} \approx C_{k+1}\Sigma_{k+1}$.

Summary: EKF algorithm

Start with $\bar{x}_0$, Σ_0. Make $k = 0$, and iterate:

1. Measure y_k, update state (`correct`) with:

$$\hat{x}_k = \bar{x}_k + \Sigma_k C_k^T (C_k^T \Sigma_k C_k^T + R_k V R_k^T)^{-1} \cdot (y_k - h(\bar{x}_k, 0))$$
$$\Sigma_{k,posteriori} = \Sigma_k - \Sigma_k C_k^T (C_k \Sigma_k C_k^T + R_k V R_k^T)^{-1} C \Sigma_k$$

C_k, R_k stand for $C(\bar{x}_k)$, $R(\bar{x}_k)$.

2. Make an open-loop prediction (`predict`) of next state:

$$\bar{x}_{k+1} = f(\hat{x}_k, u_k, 0)$$
$$\Sigma_{k+1} = A_k \cdot \Sigma_{k,posteriori} \cdot A_k^T + G_k \cdot W \cdot G_k^T$$

A_k, G_k stand for $A(\hat{x}_k, u_k)$, $G(\hat{x}_k, u_k)$.

3. $k = k + 1$, go to step 1.

Conclusions

- A non-linear discrete-time process may be approximated by its linearization around the estimated state trajectory.
- The linearised model allows setting up an approximation of stochastic mean, variance and covariance equations.
- The approximations with A_k, G_k, C_k, R_k replace the "constant" (A,G,C,R) in non-stationary Kalman filter equations.

  *Stationary Kalman filter (`dlqe`) is meaningless in a nonlinear case (well, a time-varying approximation of it).

- The resulting expressions are called **extended** Kalman filter. There are reasonably accurate for "smooth" nonlinearities (plus small Hessian, ...) and initial estimates close to the "true" process state.
  - **Warning:** With initial estimate far from the true state and/or abruptly changing nonlinearities, EKF may be far from optimal, or even **unstable**.

    *There are alternative options (**unscented, particle**) trying to solve some of the EKF drawbacks.