Filters for nonlinear systems

ISTA 410 / INFO 510: Bayesian Modeling and Inference

U. of Arizona School of Information

December 7, 2020

Outline

Last time:

- Kalman filters for linear Gaussian dynamical systems

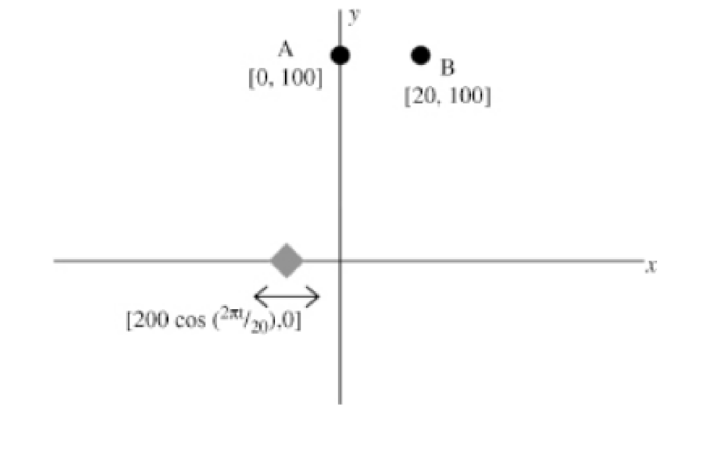
Today:

- Filtering for nonlinear systems

Extended Kalman filter

A motivating example

Here is an example[1]:



[1] from Brown & Hwang, *Introduction to Random Signals*

Assumptions for the filter

Linear dynamics:

$$x_{t+1} = Ax + N(0, Q)$$

where

$$A = \begin{pmatrix} 1 & dt \\ 0 & 1 \end{pmatrix}$$

Problem: nonlinear measurements

The challenge arises because our measurements are nonlinear:

$$z_1 = \sqrt{x^2 + 100^2}$$
$$z_2 = \sqrt{(x - 20)^2 + 100^2}$$

The approach: linearize around the current estimate.

The ad-hoc fix

Suppose we have an estimate of the state at time $k-1$, $\hat{x}_{k-1}$.
Then we can define approximations:

$$\tilde{x}_k = f(\hat{x}_{k-1}, u_{k-1}, 0)$$
$$\tilde{z}_k = h(\tilde{x}_k, 0)$$

by taking the noise value to be 0.

The approach

Then, approximate the true values x_k, z_k by linear mappings:

$$x_k = \tilde{x}_k + A(x_{k-1} - \hat{x}_{k-1}) + Ww_k$$
$$z_k = \tilde{z}_k + H(x_k - \tilde{x}_k) + Vv_k$$

where A, W, H, K are the Jacobian matrices:

- $A_{ij} = \frac{\partial f_i}{\partial x_j}$
- $W_{ij} = \frac{\partial f_i}{\partial w_j}$
- $H_{ij} = \frac{\partial h_i}{\partial x_j}$
- $V_{ij} = \frac{\partial v_i}{\partial v_j}$

The approach

The problem: in this system,

$$x_k = \tilde{x}_k + A(x_{k-1} - \hat{x}_{k-1}) + Ww_k$$
$$z_k = \tilde{z}_k + H(x_k - \tilde{x}_k) + Vv_k$$

we don't know the value of x_k (or x_{k-1}). So we can't run this directly.

Instead, define

$$\tilde{e}_{x_k} = x_k - \tilde{x}_k$$
$$\tilde{e}_{z_k} = z_k - \tilde{z}_k$$

The approach

Then, plugging the definitions of $\tilde{e}_{x_k}, \tilde{e}_{z_k}$ into the linearized equations, we get:

$$\tilde{e}_{x_k} = A(x_{k-1} - \hat{x}_{k-1}) + \varepsilon_k$$
$$\tilde{e}_{z_k} = H(x_k - \tilde{x}_k) + \eta_k$$

and this starts to look like something we can estimate with a Kalman filter.

The approach

Look ahead a couple of steps:

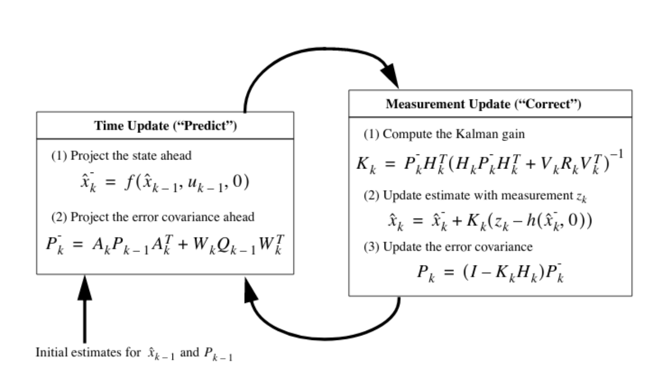
- The prior for the error process at each step should be $\hat{e}_k = 0$; our best estimate of error is zero
- So, the Kalman update step for $\hat{e}_k$ is just

$$\hat{e}_k = K_k \tilde{e}_{z_k}$$

- Plugging back into $\hat{x}_k = \tilde{x}_k + \hat{e}_k$, we get

$$\hat{x}_k = \tilde{x}_k + K_k(z_k - \tilde{z}_k)$$

Summing it all up



Applying this to our estimation problem

For our estimation problem, we assumed:

- Linear dynamics
- No actual noise in the process
- Additive noise in the measurements

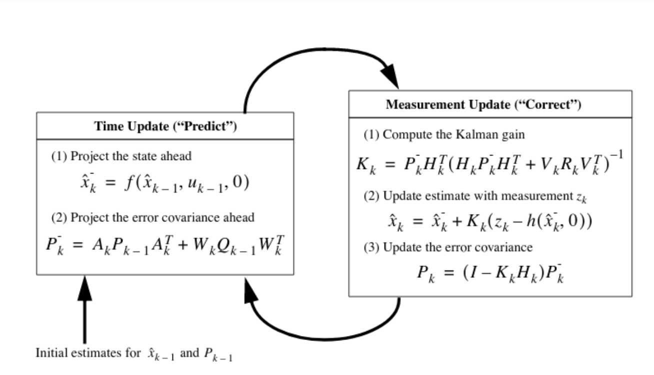
This means: $w_k = 0, \eta_k = v_k, f = id$.

An ad-hoc fix

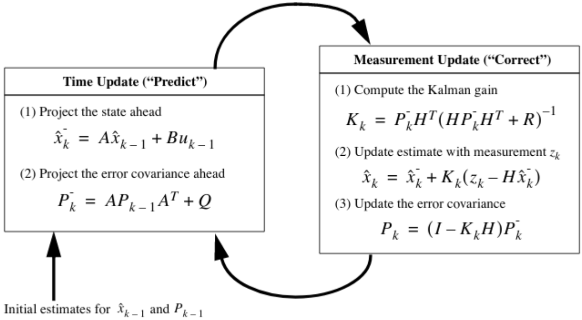
Calculate the Jacobian matrix of the measurement function:

$$H = \frac{\partial h}{\partial x_i} = \begin{pmatrix} \frac{x}{\sqrt{x^2+100^2}} & 0 \\ \frac{(x-20)}{\sqrt{(x-20)^2+100^2}} & 0 \end{pmatrix}$$

Summing it all up



Original KF


Time Update ("Predict")
(1) Project the state ahead
$\hat{x}_k^- = A\hat{x}_{k-1} + Bu_{k-1}$
(2) Project the error covariance ahead
$P_k^- = AP_{k-1}A^T + Q$
Measurement Update ("Correct")
(1) Compute the Kalman gain
$K_k = P_k^- H^T (HP_k^- H^T + R)^{-1}$
(2) Update estimate with measurement z_k
$\hat{x}_k = \hat{x}_k^- + K_k(z_k - H\hat{x}_k^-)$
(3) Update the error covariance
$P_k = (I - K_k H)P_k^-$
Initial estimates for $\hat{x}_{k-1}$ and P_{k-1}

Drawbacks

The EKF has numerous drawbacks:

- Distributions no longer normal
- Filter can diverge if errors accumulate too much
- Estimates can be poor if nonlinearity is high

Next time: sampling-based approaches