On Kalman Filtering

A study of “A New Approach to Linear Filtering and Prediction Problems” by R. E. Kalman

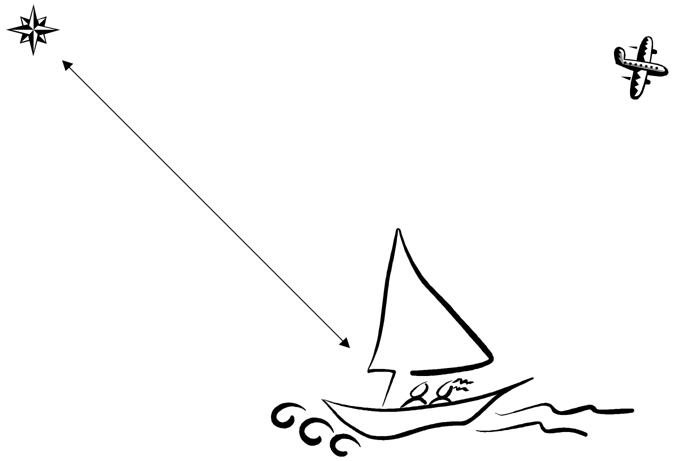
Mehul Motani

February 11, 2000

The 1960s: A Decade to Remember

- Rudolf E. Kalman in 1960
  - Research Institute for Advanced Studies (Baltimore)
  - The Discrete-time Kalman Filter
- With Richard Bucy in 1961
  - Bucy was with Johns Hopkins Applied Physics Lab
  - The Continuous-time Kalman Filter
- Kalman Filtering used widely in
  - Control Systems, Signal Processing, Communications

Some Motivation



The Static Case

- First Measurement

  z_1

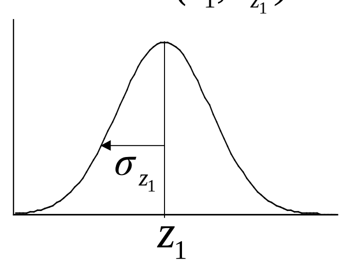
- First Estimate

$$\hat{x}_1 = z_1$$

$$\hat{\sigma}_1^2 = \sigma_{z_1}^2$$

Conditional PDF of position based on measurement z_1

$\sim Normal(z_1, \sigma_{z_1}^2)$



The Static Case (cont.)

- Second Measurement

  $z_2 \ , t_1 \cong t_2$

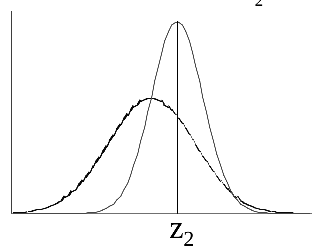
- Second Estimate

  $\hat{x}_2 = ??$

  $\hat{\sigma}_2^2 = ??$

Conditional PDF of position based on measurement z_2

$\sim Normal(z_2, \sigma_{z_2}^2)$

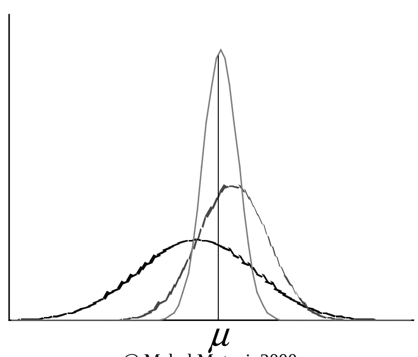


How do we optimally combine the two measurements?

The Combined Estimate

- Conditional PDF of position based on both measurements

$\sim Normal(\mu, \hat{\sigma}^2)$



The Optimum Estimate is …

$$\mu = \frac{\sigma_{z_2}^2}{\sigma_{z_1}^2 + \sigma_{z_2}^2} z_1 + \frac{\sigma_{z_1}^2}{\sigma_{z_1}^2 + \sigma_{z_2}^2} z_2$$

$$\frac{1}{\hat{\sigma}_2^{\,2}} = \frac{1}{\sigma_{z_1}^2} + \frac{1}{\sigma_{z_2}^2}$$

$$\hat{x}_2 = \mu$$

What does optimum mean?

- Unbiased (since it the conditional mean)
- Maximum Likelihood Estimate
- Least Squares Estimate
- Minimum Variance, Unbiased Estimate

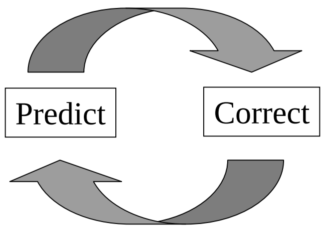
- The Kalman filter is all of the above !

Another Look

The Predictor-Corrector Structure

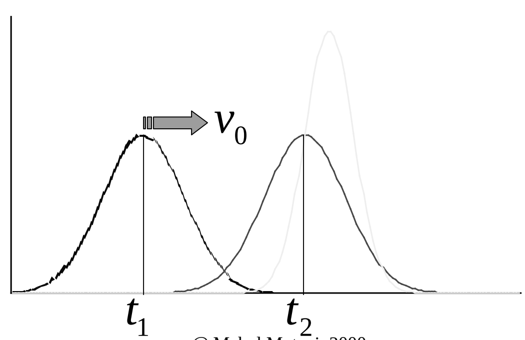
$$\hat{x}_2 = \hat{x}_1 + K_2\left[z_2 - \hat{x}_1\right]$$

$$\hat{\sigma}_2^2 = \hat{\sigma}_1^2 - K_2\hat{\sigma}_1^2$$



The Dynamic Case

- System dynamics: $\dfrac{dx}{dt} = v_0$



What is the Kalman Filter?

Optimum Recursive Data Processing Algorithm

- Optimum: Uses all available data
  - ML, Least Squares, MVUE
- Recursive: Does not store all data
  - Critical for implementation
- Data Processing
  - Incorporates discrete-time measurements

Assumptions

- Linear (Discrete) System Model
  - Admits tractable analysis
  - Linear systems theory is quite thorough.
- White Noise
  - Real systems are bandpass.
- Gaussian Noise
  - Use Central Limit Theorem arguments

General System Dynamic Model

$$x_{k+1} = \Phi_k x_k + B_k w_k + u_k$$
$$z_k = M_k x_k + v_k$$

State Vector at time $k : x_k \in \Re^n$
Input Vector : $w_k \in \Re^l$
Measurement Vector : $z_k \in \Re^m$

$n \times n$ State Transition Matrix : Φ_k
$n \times l$ Input Relation Matrix : B_k
$m \times n$ Measurement Matrix : M_k

AWG Process Noise : $u_k \sim Normal(0, Q)$
AWG Measurement Noise : $v_k \sim Normal(0, R)$

Kalman's Model

$$B_k = R = 0 \quad \Longrightarrow \quad \begin{aligned} x_{k+1} &= \Phi_k x_k + u_k \\ z_k &= M_k x_k \end{aligned}$$

- Weiner Problem

Given the observed values $\{z_k, k = 1 \cdots m\}$, find an estimate $\hat{x}_k$ which minimizes the expected loss.

Kalman's Solution

- Orthogonal Projection

$$\hat{x}_k = E[x_k \mid z_1, z_2, \cdots, z_m]$$

- Using the state representation, he derives a recursive optimum solution.
- We will not derive, but rather motivate the solution.

A Subtle Distinction

	Estimate	Estimate Error	Estimate Error Covariance
a priori	$\hat{x}_k^-$	$e_k^- = x_k - \hat{x}_k^-$	$P_k^- = E\left[e_k^- e_k^{-T}\right]$
a posteriori	$\hat{x}_k$	$e_k = x_k - \hat{x}_k$	$P_k = E\left[e_k e_k^T\right]$

Basic Operation of the Filter

- Time Update (Predict)
  - Project current state and covariance forward to the next time step, i.e. compute the next a priori estimates.
- Measurement Update (Correct)
  - Update the a priori quantities using noisy measurements, i.e. compute the a posteriori estimates.

@ Mehul Motani, 2000 17

The Optimum Solution

$$\hat{x}_k = \hat{x}_k^- + K_k\left(z_k - M_k \hat{x}_k^-\right)$$

- Choose K_k to minimize the a posteriori error covariance.
- A minimizing form for K_k is

$$K_k = P_k^- M_k^T \left(M_k P_k^- M_k^T + R_k\right)^{-1}$$

@ Mehul Motani, 2000 18

A Closer Look

$$K_k = P_k^- M_k^T \left(M_k P_k^- M_k^T + R_k \right)^{-1}$$

- Good Measurements

  As $R_k \to 0, K_k \to M_k^{-1}$

- Bad Measurements

  As $P_k^- \to 0, K_k \to 0$

@ Mehul Motani, 2000 19

Kalman Filter Algorithm

Initial estimates for P_k^- and $\hat{x}_k^-$

Predict

1. Project state ahead
$\hat{x}_{k+1}^- = \Phi_k \hat{x}_k$

2. Project error covariance
$P_{k+1}^- = \Phi_k P_k \Phi_k^T + Q_k$

Correct

1. Compute the Kalman Gain
$K_k = P_k^- M_k^T \left(M_k P_k^- M_k^T + R_k \right)^{-1}$

2. Compute a posteriori estimate
$\hat{x}_k = \hat{x}_k^- + K_k \left(z_k - M_k \hat{x}_k^- \right)$

3. Update error covariance
$P_k = \left(I - K_k M_k \right) P_k^-$

@ Mehul Motani, 2000 20

Fine Tuning the Kalman Filter

- Measurement Noise Covariance, Q
  - Can take offline samples and estimate
- Process Noise Covariance, R
  - Not so clear how to estimate
- Both quantities can be time varying!
- Choosing Q and R is actually an art.

Example: Lost in Space

- Spacecraft accelerating with random bursts from its thrusters.

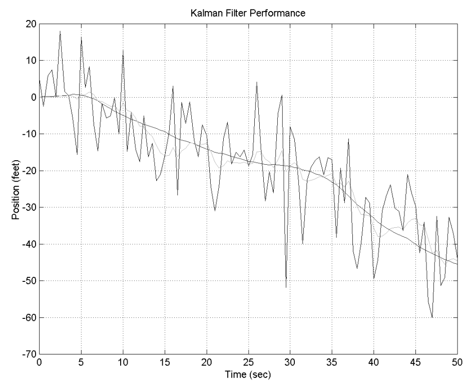
$$\begin{bmatrix} x_{k+1} \\ \dot{x}_{k+1} \end{bmatrix} = \begin{bmatrix} 1 & T \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_k \\ \dot{x}_k \end{bmatrix} + \begin{bmatrix} T^2/2 \\ T \end{bmatrix} a_k$$

$$z_k = x_k + v_k$$

$$R = \sigma_v^2 \qquad Q = \sigma_a^2 \begin{bmatrix} T^4/4 & T^3/2 \\ T^3/2 & T^2 \end{bmatrix}$$

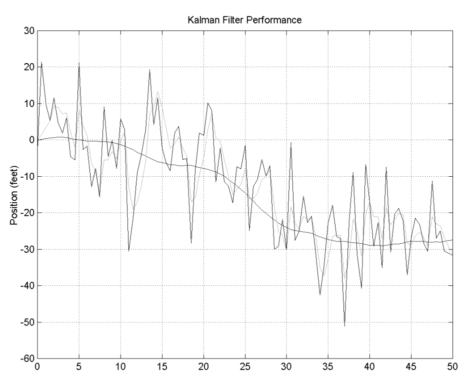
Kalman Filter Performance

$$\sigma_v = 10\,\text{ft.}, \sigma_a = 0.5\,\text{ft./sec}^2, R = \sigma_v^2$$



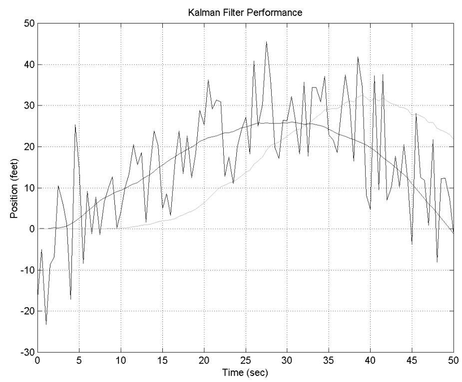
Kalman Filter Performance

$$\sigma_v = 10\,\text{ft.}, \sigma_a = 0.5\,\text{ft./sec}^2, R = 1$$



Kalman Filter Performance

$$\sigma_v = 10\,\text{ft.}, \sigma_a = 0.5\,\text{ft./sec}^2, R = 10000$$



Applications

- Navigational and Guidance Systems
- Radar tracking and Sonar ranging
- Satellite orbit computations
- Active Noise Control
- Predictive tracking for virtual reality
- MMSE receiver is Kalman filtering

Recall our Assumptions

- Linear Discrete System Model
- White Measurement and Process Noise
- Gaussian Measurement and Process Noise

The Kalman Filter is the “best” possible.

Variations on a filter

- Discrete-Discrete Kalman Filter Π
- Continuous-Discrete Kalman Filter
- Extended Kalman Filter

Continuous-Discrete Kalman

- System Model
  - Continuous Model for dynamical system
  - Discrete measurement equations
- Why?
  - Flexibility
  - Irregularly spaced measurements
  - Use numerical integration (e.g. Runge-Kutta) to project states ahead.

Non-linear Systems

- Can we relax the linearity assumption?
- Nonlinear stochastic difference equation

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

The Extended Kalman Filter

- Linearize about the current mean and covariance using Taylor Series notions.

$$x_{k+1} = \tilde{x}_{k+1} + A_k\left(x_k - \hat{x}_k\right) + W_k w_k$$
$$z_k = \tilde{z}_k + H_k\left(x_k - \tilde{x}_k\right) + U_k u_k$$

- Use Jacobians to project ahead and to relate measurement to states.

Non-Gaussian Noise

- Kalman filter is no longer "universally" optimum.
- It is still the minimum variance estimator amongst *all* linear unbiased estimators.

What did they do before Kalman?

- Weiner filter
  - Developed by Weiner at MIT in the 1940s
  - Analyzes time series in the frequency domain
  - Applies only to stationary problems
- There is much work on extending Weiner's ideas to nonstationary problems.

Kalman vs. Weiner

- Kalman Filter applies to both *stationary* and *nonstationary* problems
- Implementation Issues
  - Weiner filter operates on *all* data directly for each estimate
  - Kalman filter recursively conditions current estimate on all past measurements.

All Roads Lead From Gauss

"… since all our measurements and observations are nothing more than approximations to the truth, the same must be true of all calculations resting upon them, and the highest aim of all computations made concerning concrete phenomena must be to approximate, as nearly as practicable, to the truth. But this can be accomplished in no other way than by a suitable combination of more observations than the number absolutely requisite for the determination of the unknown quantities. This problem can only be properly undertaken when an approximate knowledge of the orbit has been already attained, which is afterwards to be corrected so as to satisfy all the observations in the most accurate manner possible."

-- From Theory of the Motion of the Heavenly Bodies Moving about the Sun in Conic Sections, Gauss, 1809

Discussion

- Non-white measurement and process noise.
- Non-independent noise
- What if the statistics of the noise are unknown or vary rapidly?
- Efficient as it is, the Kalman filter is still not practical for high dimensional systems. Can you approximate the Kalman filter for large systems?

One Last Thing to Think About

- Suppose we want to construct a time history of the states given *all* the measurements, rather than estimating the state as measurements come in.
- Can we do better than the Kalman filter?

