Ministério do Desenvolvimento, Indústria e Comércio Exterior

GOVERNO FEDERAL BRASIL PAÍS RICO É PAÍS SEM POBREZA

A Simplified Derivation of Scalar Kalman Filter using Bayesian Probability Theory

Gregory Kyriazis
Inmetro - Electrical Metrology Division

1. Introduction

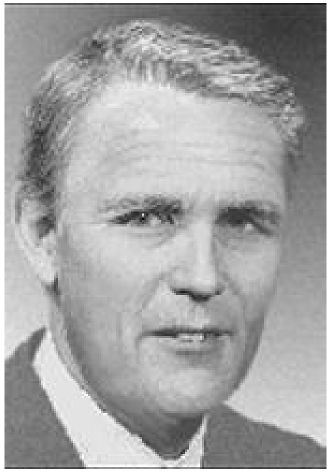
2. Linear dynamical models

3. Estimation

4. Conclusions

1. Introduction

1960 – R. E. Kalman - **“**A new approach to linear filtering and prediction problems**”**
1963 – R. E. Kalman - **“**New methods in Wiener filtering theory**”**



Kalman receives National Medal of Science

October 7, 2009

In many estimation problems, especially those involving dynamical systems, observations are made sequentially in time and up-to-date parameter estimates are required. The recursive solution to the discrete-time linear estimation problem was first published by Kalman. The estimation algorithm is called *Kalman filter*.

Kalman filter

The Kalman filtering algorithm made it possible to navigate precisely over long distances and time spans. Kalman´s algorithm is used extensively in all navigation systems for deep-space exploration.
It has also been applied to forecasting (P. J. Harrison and C. F. Stevens, 1976).

Historically, this signal estimation problem was viewed as *filtering* narrowband signals from wideband noise; hence the name “filtering” for signal estimation.

Kalman used orthogonality to derive his filtering algorithm (R. E. Kalman, 1960).

Kalman´s equations have often been derived using innovations. The concept of innovations, or the unpredictable part, of observations was introduced by Kailath (1968).

Kalman´s equations were derived using a Bayesian approach for the first time in NASA TR R-135 (1962). A similar approach was used by Y. C. Ho and R. C. K. Lee (1964). A simplified derivation is presented here.

2. Linear dynamical models

Linear dynamical models are state-space models whose state unpredictable vagaries with time are described probabilistically,

$$\boxed{X_n = b_n Z_n + \varepsilon_n, \quad \varepsilon_n \sim N\left(0, \sigma_n^2\right)}$$ Observation equation

$$\boxed{Z_n = a_{n,n-1} Z_{n-1} + w_n, \quad w_n \sim N\left(0, \omega_n^2\right)}$$ System equation

$X_n \longrightarrow$ observation sequence

$Z_n \longrightarrow$ Gauss-Markov sequence of unknown process states

$b_n \longrightarrow$ known series of constants (linear relationship)

$a_{n,n\text{-}1} \longrightarrow$ known series of constants (first-order difference equation)

$\varepsilon_n \longrightarrow$ noise sequence

$w_n \longrightarrow$ system noise driving function

Both w_n and ε_n are white and mutually uncorrelated.

2. Estimation

$X_n, Z_n \longrightarrow$ unknown values of the corresponding quantities

$\xi_n, \zeta_n \longrightarrow$ possible values of those unknowns

Let $d_n = \{d_{n-1}, X_n = x_n\}$, with d_0 describing the initial available information, including the values of $a_{n,n-1}$, b_n, ω_n and σ_n, $\forall n$.

All the information d_{n-1} about the unknown state is encoded by the posterior PDF at $n-1$ and used to derive the new posterior once the data sample $X_n = x_n$ is received at n. It is shown in the sequel how to evolve from the posterior PDF at $n-1$ to the posterior at n.

It is assumed that, initially at time $n = 1$, information concerning the state Z_0 was described as

$$\boxed{Z_0 \sim N\left(\hat{\zeta}_0, v_0^2\right)}$$ (mean and variance known)

Posterior PDF for the process state at $n-1$

$$p_{Z_{n-1}}\left(\zeta_{n-1}\middle|d_{n-1}\right)\propto\exp\left\{-\frac{1}{2\nu_{n-1}^{2}}\left(\zeta_{n-1}-\hat{\zeta}_{n-1}\right)^{2}\right\}$$

Prior PDF for the process state at n

$$p_{Z_{n}}\left(\zeta_{n}\middle|d_{n}\right)\propto\exp\left\{-\frac{1}{2\rho_{n}^{2}}\left(\zeta_{n}-\tilde{\zeta}_{n}\right)^{2}\right\}$$

$$\tilde{\zeta}_{n}=a_{n,n-1}\hat{\zeta}_{n-1}$$

$$\rho_{n}^{2}=a_{n,n-1}^{2}\nu_{n-1}^{2}+\omega_{n}^{2}$$

This follows immediately from the system equation and the properties of the Gaussian distribution.

A sampling distribution with unknown location and scale parameters is assigned that describes the prior knowledge about the noise.

As a function of those parameters the sampling distribution is then termed the *likelihood* of the parameters given the observed data.

Likelihood for Z_n, given that $X_n = x_n$,

$$\boxed{l(\zeta_n ; x_n) \propto \exp\left\{-\frac{b_n^2}{2\sigma_n^2}\left(\zeta_n - x_n / b_n\right)^2\right\}}$$

Noise and prior information are combined at each time with Bayes' theorem, i.e., *posterior* α *prior* × *likelihood.*

On combining the prior and the likelihood, completing the square, and lumping the terms that do not depend on ζ_n into the proportionality constant,

Posterior PDF for the process state at n

$$p_{Z_n}\left(\zeta_n \middle| d_n\right) \propto \exp\left\{-\frac{1}{2\nu_n^2}\left(\zeta_n - \hat{\zeta}_n\right)^2\right\}$$

$$\frac{\hat{\zeta}_n}{\nu_n^2} = \frac{\tilde{\zeta}_n}{\rho_n^2} + \frac{b_n x_n}{\sigma_n^2} \qquad \frac{1}{\nu_n^2} = \frac{1}{\rho_n^2} + \frac{b_n^2}{\sigma_n^2}$$

The posterior mean is the weighted average of the prior mean and x_n/b_n with the weights being ρ_n^2 and σ_n^2/b_n, respectively.

For fast recursive estimation one is just interested in the estimate and the associated uncertainty. In dynamical problems, each state can be estimated from the last previous estimate and the new data sample received. Thus only the last estimate and associated uncertainty need to be stored.

$\hat{\sigma}^2$

Recursive equations

We now have all the equations required to recursively generate the solution to the estimation problem:

$$\rho_n^2 = a_{n,n-1}^2 \nu_{n-1}^2 + \omega_n^2$$

$$\hat{\zeta}_n = \frac{\sigma_n^2 a_{n,n-1} \hat{\zeta}_{n-1} + \rho_n^2 b_n x_n}{\sigma_n^2 + \rho_n^2 b_n^2}$$

´best´ estimate of the state at n

$$\nu_n^2 = \frac{\sigma_n^2 \rho_n^2}{\sigma_n^2 + \rho_n^2 b_n^2}$$

uncertainty associated with the estimate

It is assumed that the state estimate at $n-1$, $\nu_{n-1}{}^2$, b_n, $a_{n,n-1}$, σ_n, ω_n and x_n are all known at n.

Scalar Kalman filter

$$\boxed{\rho_n^2 = a_{n,n-1}^2 \nu_{n-1}^2 + \omega_n^2}$$

By defining the *Kalman gain* as

$$\boxed{\kappa_n = \frac{\rho_n^2 b_n}{\rho_n^2 b_n^2 + \sigma_n^2}}$$

the state estimate at n and the associated uncertainty become

$$\boxed{\hat{\zeta}_n = a_{n,n-1}\hat{\zeta}_{n-1} + \kappa_n\left(x_n - b_n a_{n,n-1}\hat{\zeta}_{n-1}\right)} \qquad \boxed{\nu_n^2 = (1-\kappa_n)\rho_n^2}$$

If $a_{n,n-1}$ does not vary with n, and w_n and ε_n are both stationary, that is, ω_n and σ_n are both constants, then both κ_n and ν_n will approach limits as n approaches infinity.

4. Conclusion

The Bayesian approach presented here provides a simpler and more direct derivation of the solution to the problem of recursive estimation of the state of a first-order linear dynamical system.

It was shown that as expected the solution is Kalman´s filtering algorithm.