12.540 Principles of the Global Positioning System
Lecture 13

Prof. Thomas Herring

Estimation

- Summary
  - First-order Gauss Markov Processes
  - Kalman filters – Estimation in which the parameters to be estimated are changing with time

Specific common processes

- White-noise: Autocorrelation is Dirac-delta function; PSD is flat; integral of power under PSD is variance of process (true in general)
- First-order Gauss-Markov process (one of most models common in Kalman filtering)

$$\varphi_{xx}(\tau) = \sigma^2 e^{-\beta|\tau|}$$

$$\Phi_{xx}(\omega) = \frac{2\beta\sigma^2}{\omega^2 + \beta^2} \qquad \frac{1}{\beta} \text{ is correlation time}$$

Other characteristics of FOGM

Excitation function $\frac{dx}{dt} = -\beta x(t) + w(t)$

Solution $x(t+\Delta t) = e^{-\beta\Delta t} x(t) + \underbrace{e^{-\beta\Delta t} \int_0^{\Delta t} e^{u\beta} w(t+u)\,du}_{\text{White Noise Excitation}}$

Variance over interval T $\sigma_T^2(T) = \sigma^2\left[1 + \frac{(1 - e^{-2T\beta})}{2T\beta}\right]$

Variance of change in τ $D(\tau) = 2\sigma^2\left(1 - e^{-\beta|\tau|}\right)$

White noise, $w(t)$, variance $\sigma^2 = q/(2\beta)$

Characteristics of FOGM

- This process noise model is very useful because as □, inverse correlation time, goes to infinity (zero correlation time), the process is white noise
- When the correlation time goes to infinity (□–>0), process becomes random walk (ie, sum of white noise).
- NOTE: Random walk is not a stationary process because its variance tends to infinity as time goes to infinity
- In the FOGM solution equation, note the damping term $e^{-□t□}x$ which keeps the process bounded

Formulation of Kalman filter

- A Kalman filter is an implementation of a Bayes estimator.
- Basic concept behind filter is that some of the parameters being estimated are random processes and as data are added to the filter, the parameter estimates depend on new data and the changes in the process noise between measurements.
- Parameters with no process noise are called deterministic.

Formulation

- For a Kalman filter, you have measurements y(t) with noise v(t) and a state vector (parameter list) which have specified statistical properties.

$$y_t = \mathbf{A_t} x_t + v_t \quad \text{Observation equation at time t}$$

$$x_{t+1} = \mathbf{S_t} x_t + w_t \quad \text{State transition equation}$$

$$< v_t v_t^T >= \mathbf{V_t} \quad < w_t w_t^T >= \mathbf{W_t} \quad \text{Covariance matrices}$$

Basic Kalman filter steps

- Kalman filter can be broken into three basic steps
- Prediction: Using process noise model, "predict" parameters at next data epoch
  - Subscript is time quantity refers to, superscript is data

$$\hat{x}_{t+1}^t = \mathbf{S}_t \hat{x}_t^t \quad \mathbf{S}_t \text{ is state transition matrix}$$

$$\mathbf{C_{t+1}^t} = \mathbf{S_t C_t^t S_t^T} + \mathbf{W_t} \quad \mathbf{W_t} \text{ is process noise covariance matrix}$$

Prediction step

- The state transition matrix **S** projects state vector (parameters) forward to next time.
  - For random walks: S=1
  - For rate terms: S is matrix [1 Δt][0 1]
  - For FOGM: S=e $^{-\Delta t/\tau}$
  - For white noise S=0
- The second equation projects the covariance matrix of the state vector , **C**, forward in time. Contributions from state transition and process noise (**W** matrix). **W** elements are 0 for deterministic parameters

Kalman Gain

- The Kalman Gain is the matrix that allocates the differences between the observations at time t+1 and their predicted value at this time based on the current values of the state vector according to the noise in the measurements and the state vector noise

$$\boxed{\mathbf{K} = \mathbf{C}^{t}_{t+1}\mathbf{A}^{T}_{t+1}\left(\mathbf{V}_{t+1} + \mathbf{A}_{t+1}\mathbf{C}^{t}_{t+1}\mathbf{A}^{T}_{t+1}\right)^{-1}}$$

Update step

- Step in which the new observations are “blended” into the filter and the covariance matrix of the state vector is updated.

$$\hat{x}_{t+1}^{t+1} = \hat{x}_{t+1}^{t} + \mathbf{K}(y_{t+1} - \mathbf{A}_{\mathbf{t+1}}\hat{x}_{t+1}^{t})$$
$$\mathbf{C}_{\mathbf{t+1}}^{\mathbf{t+1}} = \mathbf{C}_{\mathbf{t+1}}^{\mathbf{t}} - \mathbf{K}\mathbf{A}_{\mathbf{t+1}}\mathbf{C}_{\mathbf{t+1}}^{\mathbf{t}}$$

- The filter has now been updated to time t+1 and measurements from t+2 can added and so on until all the observations have been added.

Aspects to note about Kalman Filters

- How is the filter started? Need to start with an apriori state vector covariance matrix (basically at time 0)
- Notice in updating the state covariance matrix. **C**, that at each step the matrix is decremented. If the initial covariances are too large, then significant rounding error in calculation eg. If position assumed ±100 m (variance 10^{10} mm apriori and data determines to 1 mm, then **C** is decremented by 10 orders of magnitude (double precision has on 12 significant digits).
- Square-root-information filters overcome this problem but usually take longer to run than a standard Kalman filter.

"Smoothing" filters

- In a standard Kalman filters, the stochastic parameters obtained during the filter run are not optimum because they do not contain information about the deterministic parameters obtained from future data.
- A smoothing Kalman filter, runs the filter forwards (FRF) and backwards in time (BRF), taking the full average of the forward filter at the update step with the backwards filter at the prediction step.

Smoothing filters

- The derivation of the full average can be derived from the filter equations.
- The smoothing filter is

Smoothing Kalman Gain

$\mathbf{B} = \mathbf{C}_+(\mathbf{C}_+ + \mathbf{C}_-)^{-1}$ $\quad \mathbf{C}_+$ from FRF, $\mathbf{C}_-$ from BRF

$\hat{x}_t^s = \hat{x}_+ + \mathbf{B}(\hat{x}_- - \hat{x}_+)$ Smoothed state vector estimate

$\mathbf{C}_t^s = \mathbf{C}_+ - \mathbf{B}\mathbf{C}_-$ Smoothed estimate covariance matrix

Properties of smoothing filter

- Deterministic parameters (ie., no process noise) should remain constant with constant variance in smoothed results.
- Solution takes about 2.5 times longer to run than just a forward filter
- If deterministic parameters are of interest only, then just FRF needed.

Note on apriori constraints

- In Kalman filter, apriori covariances must be applied to all parameters, but cannot be too large or else large rounding errors (non-positive definite covariance matrices).
- Error due to apriori constraints given approximately by (derived from filter equations).
- Approximate formulas assuming uncorrelated parameter estimates and the apriori variance is large compared to intrinsic variance with which parameter can be determined.

Errors due to apriori constraints

$\Delta\hat{x} \approx \frac{\sigma^2}{\sigma_0^2}(x_0 - \hat{x})$ $\Delta\hat{x}$ is error in estimates due to error in

apriori $(x_0 - \hat{x})$; σ_0^2 is apriori variance, σ^2 is variance of estimate.

$\frac{\sigma^2}{\sigma_0^2}$ is assumed $<< 1$

Error in variance estimate is $\approx \frac{\sigma^4}{\sigma_0^2}$

Note: Error depends on ratio of aposteriori to apriori variance rather than absolute magnitude of error in apriori to apriori variance

Contrast between WLS and Kalman Filter

- In Kalman filters, apriori constraints must be given for all parameters; not needed in weighted least squares (although can be done).
- Kalman filters allow zero variance parameters; can not be done is WLS since inverse of constraint matrix needed
- Kalman filters allow zero variance data; can not be done in WLS again due to inverse of data covariance matrix.
- Kalman filters allow method for applying absolute constraints; can only be tightly constrained in WLS
- In general, Kalman filters are more prone to numerical stability problems and take longer to run (strictly many more parameters).
- Process noise models can be implemented in WLS but very slow.

Applications in GPS

- Most handheld GPS receivers use Kalman filters to estimate velocity and position as function of time.
- Clock behaviors are “white noise” and can be treated with Kalman filter
- Atmospheric delay variations ideal for filter application
- Stochastic variations in satellite orbits