Prediction & Estimation

Recursive Minimum Mean Square Error Estimation:

Kalman Filtering

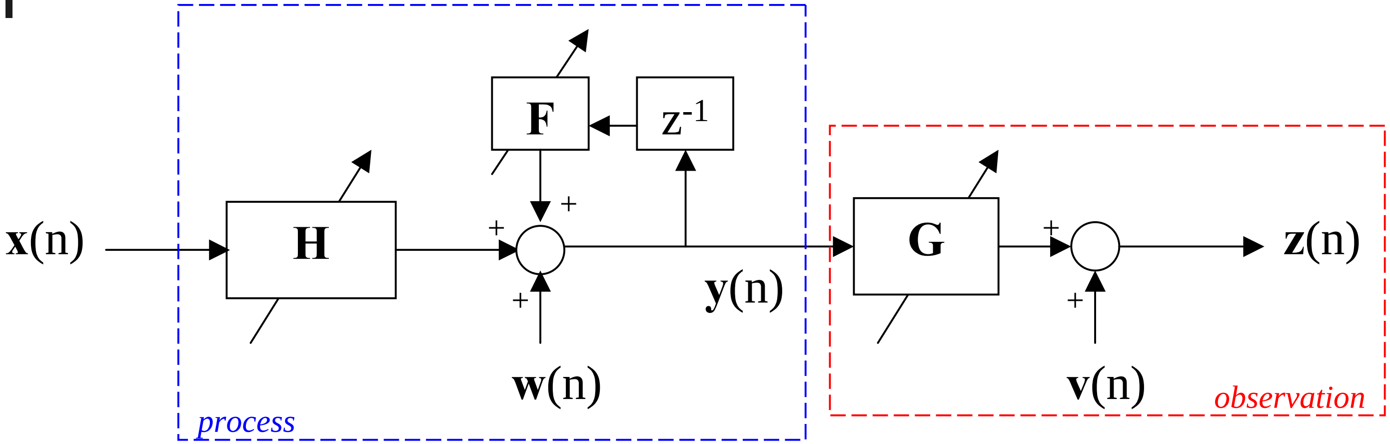
Problem Statement

- Given:
  - Discrete series of nonobservable analogue outputs (internal states) $y(n)$ of a LTI/LTV system with optional control inputs $x(n)$ and additive white Gaussian process noise $w(n)$
  - External observations $z(n)$ of LTI/LTV filtered $y(n)$ with additive white Gaussian observation noise $v(n)$
- Find:
  - Optimal recursive filter $K_{\mathrm{opt}}(n)$ that transforms observations such that MMSE estimate of $y(n)$ is obtained
- Exploit availability of noisy measurement of current value to improve on forecast based on previous ones

Problem Statement

- Strategy:
  - If discrepancy between predicted and measured values is large: attach little importance to evolution of $y(n)$
  - If discrepancy between predicted and measured values is small: attach much importance to evolution of $y(n)$
- Closely related to “Hidden Markov model” (except here: $y(n)$ continuous and $w(n)$ Gaussian)

LTV State-Space Model



- Process (system) model (1st-order Markov):

$$\mathbf{y}(n) = \mathbf{F}(n)\cdot\mathbf{y}(n-1) + \mathbf{H}(n)\cdot\mathbf{x}(n) + \mathbf{w}(n) \qquad \mathbf{w}(n) \sim N(\mathbf{0},\mathbf{Q}(n))$$

- Observation (measurement) model:

$$\mathbf{z}(n) = \mathbf{G}(n)\cdot\mathbf{y}(n) + \mathbf{v}(n), \qquad \mathbf{v}(n) \sim N(\mathbf{0},\mathbf{R}(n))$$

- Assume $\{\mathbf{y}(0),\{\mathbf{w}(n)\},\{\mathbf{v}(n)\}\}$ statistically independent

Definitions

- Prior-state based estimate $\mathbf{y}^{(-)}(n) = \hat{\mathbf{y}}(n|n-1)$ at *n*:
  - Use observations estimate prior to *n* (i.e., up to and including ones for step *n*-1) to estimate output *y* at *n*
- Posterior-state based estimate $\mathbf{y}^{(+)}(n) = \hat{\mathbf{y}}(n|n)$ at *n*:
  - Use observations estimate prior to *n* <u>and</u> current observation at *n* to estimate output *y* at *n*
- Prior-state based covariance of $\hat{\mathbf{y}}(n)$: $\mathbf{P}(n|n-1)$
  - Measures the estimation accuracy for output, based on info prior to *n*, compared to true output obtained at *n*
- Posterior-state based covariance of $\hat{\mathbf{y}}(n)$: $\mathbf{P}(n|n)$

Kalman Filter Algorithm

- Three stages per iteration step:
  - (I) Predict
    - (a) *Prior estimate of state:* use present input state and previous posterior estimated output (i.e., all available info) to estimate present prior output state:

$$\hat{\mathbf{y}}(n|n-1)=\mathbf{F}(n)\cdot\hat{\mathbf{y}}(n-1|n-1)+\mathbf{H}(n)\cdot\mathbf{x}(n)$$

    - (b) *Prior covariance:* use covariance of previous process noise and previous posterior covariance to estimate present prior covariance:

$$\mathbf{P}(n|n-1)=\mathbf{F}(n)\cdot\mathbf{P}(n-1|n-1)\cdot\mathbf{F}^{\mathrm{T}}(n)+\mathbf{Q}(n-1),$$

where $\mathbf{P}(n|n-1)=\mathrm{cov}\left[\mathbf{y}(n)-\hat{\mathbf{y}}(n|n-1)\right]$, $\mathbf{P}(n|n)=\mathrm{cov}\left[\mathbf{y}(n)-\hat{\mathbf{y}}(n|n)\right]$

Kalman Filter Algorithm

- (II) Evaluate
  - (a) *Observation deviation:* use present observation and present prior estimated output to calculate residual of present observation (estimated measurement error):

$$\tilde{\mathbf{u}}(n) = \mathbf{z}(n) - \mathbf{G}(n) \cdot \hat{\mathbf{y}}(n \mid n-1)$$

  - (b) *Kalman gain:* calculate optimal Kalman gain (cf. infra for derivation) from covariance of current residual and from prior covariance of current state:

$$\mathbf{K}(n) = \mathbf{P}(n \mid n-1) \cdot \mathbf{G}^{\mathrm{T}}(n) \cdot \mathbf{S}^{-1}(n),$$

where

$$\mathbf{S}(n) = \mathrm{cov}\left[\tilde{\mathbf{u}}(n)\right] = \mathbf{G}(n) \cdot \mathbf{P}(n \mid n-1) \cdot \mathbf{G}^{\mathrm{T}}(n) + \mathbf{R}(n)$$

Kalman Filter Algorithm

- (III) Update
  - (a) *Posterior estimate of state:* use current residual with current Kalman gain to adjust prior estimate of present state:

$$\hat{\mathbf{y}}(n \mid n) = \hat{\mathbf{y}}(n \mid n-1) + \mathbf{K}(n) \cdot \tilde{\mathbf{u}}(n)$$

  - (b) *Posterior covariance:* use optimal Kalman gain and observation transfer function to adjust prior covariance of estimated present state:

$$\mathbf{P}(n \mid n) = \left[\mathbf{I} - \mathbf{K}(n) \cdot \mathbf{G}(n)\right] \cdot \mathbf{P}(n \mid n-1)$$

Determination of $\mathbf{K}_{\mathrm{opt}}$

$$
\begin{aligned}
\mathbf{P}(n\,|\,n) &= \operatorname{cov}\left[\mathbf{y}(n) - \hat{\mathbf{y}}(n\,|\,n)\right] \\
&= \operatorname{cov}\left[\mathbf{y}(n) - \left(\hat{\mathbf{y}}(n\,|\,n-1) + \mathbf{K}(n)\cdot\tilde{\mathbf{u}}(n)\right)\right] \\
&= \operatorname{cov}\left[\mathbf{y}(n) - \left(\hat{\mathbf{y}}(n\,|\,n-1) + \mathbf{K}(n)\cdot\left(\mathbf{z}(n) - \mathbf{G}(n)\cdot\hat{\mathbf{y}}(n\,|\,n-1)\right)\right)\right] \\
&= \operatorname{cov}\left[\mathbf{y}(n) - \left(\hat{\mathbf{y}}(n\,|\,n-1) + \mathbf{K}(n)\cdot\left(\left(\mathbf{G}(n)\cdot\mathbf{y}(n) + \mathbf{v}(n)\right) - \mathbf{G}(n)\cdot\hat{\mathbf{y}}(n\,|\,n-1)\right)\right)\right] \\
&= \operatorname{cov}\left[\left(\mathbf{I} - \mathbf{K}(n)\cdot\mathbf{G}(n)\right)\left(\mathbf{y}(n) - \hat{\mathbf{y}}(n\,|\,n-1)\right) - \mathbf{K}(n)\cdot\mathbf{v}(n)\right]
\end{aligned}
$$

Since $\{\mathbf{y}(0),\{\mathbf{v}(n)\}\}$ are independent:

$$
\begin{aligned}
\mathbf{P}(n\,|\,n) &= \operatorname{cov}\left[\left(\mathbf{I} - \mathbf{K}(n)\cdot\mathbf{G}(n)\right)\left(\mathbf{y}(n) - \hat{\mathbf{y}}(n\,|\,n-1)\right)\right] + \operatorname{cov}\left[\mathbf{K}(n)\cdot\mathbf{v}(n)\right] \\
&= \left(\mathbf{I} - \mathbf{K}(n)\cdot\mathbf{G}(n)\right)\cdot\operatorname{cov}\left(\mathbf{y}(n) - \hat{\mathbf{y}}(n\,|\,n-1)\right)\cdot\left(\mathbf{I} - \mathbf{K}(n)\cdot\mathbf{G}(n)\right)^{\mathrm{T}} + \mathbf{K}(n)\cdot\operatorname{cov}\left(\mathbf{v}(n)\right)\cdot\mathbf{K}^{\mathrm{T}}(n \\
&= \left(\mathbf{I} - \mathbf{K}(n)\cdot\mathbf{G}(n)\right)\cdot\mathbf{P}(n\,|\,n-1)\cdot\left(\mathbf{I} - \mathbf{K}(n)\cdot\mathbf{G}(n)\right)^{\mathrm{T}} + \mathbf{K}(n)\cdot\mathbf{R}(n)\cdot\mathbf{K}^{\mathrm{T}}(n)
\end{aligned}
$$

(general expression)

Determination of $\mathbf{K}_{\text{opt}}$

$$\begin{aligned}\mathbf{P}(n|n) &= \left(\mathbf{I}-\mathbf{K}(n)\cdot\mathbf{G}(n)\right)\cdot\mathbf{P}(n|n-1)\cdot\left(\mathbf{I}-\mathbf{K}(n)\cdot\mathbf{G}(n)\right)^{\mathrm{T}}+\mathbf{K}(n)\cdot\mathbf{R}(n)\cdot\mathbf{K}^{\mathrm{T}}(n)\\ &= \mathbf{P}(n|n-1)-\mathbf{K}(n)\cdot\mathbf{G}(n)\cdot\mathbf{P}(n|n-1)-\mathbf{P}(n|n-1)\cdot\mathbf{G}^{\mathrm{T}}(n)\cdot\mathbf{K}^{\mathrm{T}}(n)\\ &\quad+\mathbf{K}(n)\cdot\left[\mathbf{G}(n)\cdot\mathbf{P}(n|n-1)\cdot\mathbf{G}^{\mathrm{T}}(n)+\mathbf{R}(n)\right]\cdot\mathbf{K}^{\mathrm{T}}(n)\\ &= \mathbf{P}(n|n-1)-\mathbf{K}(n)\cdot\mathbf{G}(n)\cdot\mathbf{P}(n|n-1)-\mathbf{P}(n|n-1)\cdot\mathbf{G}^{\mathrm{T}}(n)\cdot\mathbf{K}^{\mathrm{T}}(n)+\mathbf{K}(n)\cdot\mathbf{S}(n)\cdot\mathbf{K}^{\mathrm{T}}(n)\end{aligned}$$

- MMSE: minimum value of $E\left(\left|\mathbf{y}(n)-\hat{\mathbf{y}}(n|n)\right|^2\right)$ is reached when trace of $\mathbf{P}(n|n)$ is minimum:

$$\frac{\partial \mathrm{Tr}\left(\mathbf{P}(n|n)\right)}{\partial \mathbf{K}(n)} = -2\left[\mathbf{G}(n)\cdot\mathbf{P}(n|n-1)\right]^{\mathrm{T}}+2\mathbf{K}(n)\cdot\mathbf{S}(n)=0$$

i.e., when

$$\boxed{\mathbf{K}_{\mathrm{opt}}(n)=\mathbf{P}(n|n-1)\cdot\mathbf{G}^{\mathrm{T}}(n)\cdot\mathbf{S}^{-1}(n)}$$

(optimal Kalman gain)

Posterior Covariance for $\mathbf{K}_{\text{opt}}$

- In this case, $\mathbf{K}_{\text{opt}}(n)\cdot\mathbf{S}(n)\cdot\mathbf{K}_{\text{opt}}^{\mathrm{T}}(n)=\mathbf{P}(n\,|\,n-1)\cdot\mathbf{G}^{\mathrm{T}}(n)\cdot\mathbf{K}_{\text{opt}}^{\mathrm{T}}(n)$ so that then

$$\begin{aligned}\mathbf{P}(n\,|\,n)&=\mathbf{P}(n\,|\,n-1)-\mathbf{K}_{\text{opt}}(n)\cdot\mathbf{G}(n)\cdot\mathbf{P}(n\,|\,n-1)\\&\quad-\mathbf{P}(n\,|\,n-1)\cdot\mathbf{G}^{\mathrm{T}}(n)\cdot\mathbf{K}_{\text{opt}}^{\mathrm{T}}(n)+\mathbf{K}_{\text{opt}}(n)\cdot\mathbf{S}(n)\cdot\mathbf{K}_{\text{opt}}^{\mathrm{T}}(n)\\&=\left[\mathbf{I}-\mathbf{K}_{\text{opt}}(n)\cdot\mathbf{G}(n)\right]\cdot\mathbf{P}(n\,|\,n-1)\end{aligned}$$

which is more efficiently computed than the general expression (p. 9) valid for arbitrary $\mathbf{K}(n)$

Nonlinear Kalman Filters

- Schmidt's extended Kalman filter
  - If process and/or observation dynamics are nonlinear
  - Process model must be differentiable
  - Equivalent to linearization of nonlinear function around current estimate
  - Covariance matrix is replaced by Jacobian
- Unscented Kalman filter
  - For strongly nonlinear dynamics
  - Avoids calculation of Jacobian
  - Uses unscented transform: selects sample points around mean value and propagates states & observations (Monte Carlo)
- Stratonovich-Kalman-Bucy filter
  - Same as simple Kalman filter, but for continuous time I/O
  - Differential equations instead of difference equations

Kalman Filter: Probabilistic Characterisation

- 1st-order Markov process:
  - Process state: $p(\mathbf{y}(n) \mid \mathbf{y}(0),\ldots,\mathbf{y}(n-1)) = p(\mathbf{y}(n) \mid \mathbf{y}(n-1))$
  - Measured state: $p(\mathbf{z}(n) \mid \mathbf{y}(0),\ldots,\mathbf{y}(n)) = p(\mathbf{z}(n) \mid \mathbf{y}(n))$

⇒Probability density for all process & observation states, by iteration and substitution,

$$p(\mathbf{z}(n),\ldots,\mathbf{z}(1);\mathbf{x}(0),\ldots,\mathbf{x}(n)) = p(\mathbf{y}(0)\prod_{k=1}^{n} p(\mathbf{z}(k) \mid \mathbf{y}(k))\, p(\mathbf{y}(k) \mid \mathbf{y}(k-1))$$

- Kalman filter: requires measured states up to present time as basis for estimating probability of present state: $p(\mathbf{y}(n) \mid [\mathbf{z}(n)]),$ $\quad [\mathbf{z}(n)] = \{\mathbf{z}(1),\ldots,\mathbf{z}(n)\}$

Kalman Filter: Probabilistic Characterisation

- Chapman-Kolmogorov relation for conditional probabilities:

$$p(\mathbf{y}(n)\,|\,[\mathbf{z}(n-1)]) = \int p(\mathbf{y}(n)\,|\,\mathbf{y}(n-1))\; p(\mathbf{y}(n-1)\,|\,[\mathbf{z}(n-1)]\,d\mathbf{y}(n-1)$$

- Posterior probability: via Bayes' relation

$$\boxed{p(\mathbf{y}(n)\,|\,[\mathbf{z}(n)]) = \frac{p(\mathbf{z}(n)\,|\,\mathbf{y}(n))\; p(\mathbf{y}(n)\,|\,[\mathbf{z}(n-1)])}{p(\mathbf{z}(n)\,|\,[\mathbf{z}(n-1)])}}$$

with

$$p\big(\mathbf{z}(n)\,|\,[\mathbf{z}(n-1)]\big) = \int p(\mathbf{z}(n)\,|\,\mathbf{y}(n))\; p(\mathbf{y}(n)\,|\,[\mathbf{z}(n-1)])\,d\mathbf{y}(n)$$

Higher-Order Kalman Filters

- Our formulation: 1-step prediction (n-1→n)
- Can be extended to predict farther in time (n-1→n, n+1, n+2)
  - Then: larger prediction error in general

Applications of Kalman Filters

- Navigation & tracking
  - Apollo, missile, GNSS, autopilot, radar
- Medicine (neurology, cardiology)
- Geology (climate / weather forecasting)
- Macroeconomics
- Chaotic signals
- Mapping
- Phase-locked loop (PLL)
- Etc.