Suavizado no causal en representación interna: RTS smoother

Antonio Sala Piqueras

Control de sistemas complejos

Dept. Ing. Sistemas y Automatica (DISA)

Universitat Politècnica de València (UPV)

Video-presentación disponible en:

http://personales.upv.es/asala/YT/V/**rtsm**.html

Presentación

Motivación:

El filtro de Kalman es causal: da la mejor predicción del estado actual conocida información pasada y presente, pero información futura puede disminuir la incertidumbre. En aplicaciones fuera de linea, puede generalizarse para usar datos futuros (no causal).

Objetivos:

Comprender el problema de suavizado y las ideas generales del suavizador Rauch-Tung-Striebel.

Contenidos:

Mejor predicción lineal. Información imperfecta. Sistemas dinámicos. Ecuaciones RTS. Conclusiones.

Preliminares: mejor predicción lineal

Supongamos que matriz de vzas-covarianzas de dos grupos de variables s y z normales multidimensional, media $\bar{s}$ y $\bar{z}$, es:

$$E\left(\begin{pmatrix} x \\ s \end{pmatrix}\begin{pmatrix} s^T & z^T \end{pmatrix}\right) := \begin{pmatrix} M & N^T \\ N & Q \end{pmatrix}$$

mejor predicción de s dado z es:

$$s|z \sim \mathcal{N}(\bar{s} + N^T Q^{-1}(z - \bar{z}),\ M - N^T Q^{-1} N)$$

válido si se dispone de información adicional "perfecta", z determinista.
Equiv., definiendo $G = N^T Q^{-1}$,

$$s|z \sim \mathcal{N}(\bar{s} + G(z - \bar{z}),\ M - GQG^T)$$

Modelos en representación interna

Supongamos que tenemos modelo $x_{k+1} = Ax_k + w_k$, $y_k = Cx_k + v_k$, $E(ww^T) = W$, $E(vv^T) = V$. Consideremos:

- Una estimación del estado actual x_k, con información de "**pasado y presente**". Media: $\bar{x}_k$, Vza.: Z_k, de filtro Kalman (a posteriori, con $\{y_0, \ldots, y_k\}$).
- Una predicción del siguiente estado x_{k+1} con el "**pasado y presente**" (antes de medir y_{k+1}). Media: $\hat{x}_{k+1} = A\bar{x}_k$, Vza.: $P_{k+1} := AZ_kA^T + W$, con W varianza ruido de proceso.
- Una predicción (simulación) de salidas **futuras**, $y_F = [y_{k+1}; \ \ldots \ ; y_N]$. Media $\hat{y}_F = \mathcal{O}\hat{x}_{k+1}$, varianza $Q := \mathcal{O}P_{k+1}\mathcal{O}^T + H$, siendo H la matriz de varianza* de ruidos de proceso y medida futuros.

*no es importante su estructura explícita en función de V y W.

Matriz de vzas covarianzas conjunta

Objetivo:
Obtener la mejor pred. de x_k dadas las medidas y_F, $x_k|y_F \sim \mathcal{N}(\bar{x}_k^s, \Sigma_k^s)$.

Como las covarianzas son:
$E(x_k x_{k+1}^T) = E(x_k(x_K^T A^T + w_k^T)) = Z_k A^T$,
$E(x_{k+1} y_F) = E(x_{k+1}(\mathcal{O} x_{k+1} + h)^T) = P_{k+1}\mathcal{O}^T$,
$E(x_k y_F^T) = E(x_k(\mathcal{O}(A x_k + w_k) + h)^T) = Z_k A^T \mathcal{O}^T$,
tenemos:

$$E\left(\begin{pmatrix} x_k \\ x_{k+1} \\ y_F \end{pmatrix} \begin{pmatrix} x_k^T & x_{k+1}^T & y_F^T \end{pmatrix}\right) = \begin{pmatrix} Z_k & Z_k A^T & Z_k A^T \mathcal{O}^T \\ A Z_k & P_{k+1} & P_{k+1}\mathcal{O}^T \\ \mathcal{O} A Z_k & \mathcal{O} P_{k+1} & Q \end{pmatrix}$$

Predicciones dada información futura

- Mejor pred. lineal* de x_k dado y_F, con $G_1 := Z_k A^T \mathcal{O}^T Q^{-1}$:

$$\bar{x}_k^s := \bar{x}_k + G_1(y_F - \mathcal{O}A\bar{x}_k), \qquad \Sigma_k^s := Z_k - G_1 Q G_1^T$$

*Es nuestro objetivo, pero obtener y manipular $\mathcal{O}$ es tedioso.

Utilizaremos, para evitar manipular $\mathcal{O}$, variables intermedias*:

- Mejor pred. lineal de x_k dado x_{k+1}, con $G_2 := Z_k A^T P_{k+1}^{-1}$:

$$\bar{\xi} := \bar{x}_k + G_2(x_{k+1} - A\bar{x}_k), \qquad \Sigma_\xi := Z_k - G_2 P_{k+1} G_2^T$$

- Mejor pred. lineal de x_{k+1} dado y_F, con $G_3 := P_{k+1}\mathcal{O}^T Q^{-1}$:

$$\bar{x}_{k+1}^s := \hat{x}_{k+1} + G_3(y_F - \mathcal{O}\hat{x}_{k+1}), \qquad \Sigma_{k+1}^s := P_{k+1} - G_3 Q G_3^T$$

*Las variables "intermedias" x_{k+1} pueden usarse porque x_k e y_F son 'condicionamente independientes dado x_{k+1}', detalle omitido por brevedad.

Expresión alternativa para $\bar{x}_k^s$ y Σ_k^s

Observamos que $G_1 = G_2 G_3$. Por tanto, la media $\bar{x}_k^s$ puede expresarse:

$$\bar{x}_k^s = \bar{x}_k + G_1(y_F - \mathcal{O}A\bar{x}_k) = \bar{x}_k + G_2 G_3(y_F - \mathcal{O}\hat{x}_{k+1}) = \bar{x}_k + G_2(\bar{x}_{k+1}^s - A\bar{x}_k)$$

y la varianza Σ_k^s:

$$\Sigma_k^s = Z_k - G_1 Q G_1^T = Z_k - G_2 G_3 Q G_3^T G_2^T$$

$$= Z_k - G_2 P_{k+1} G_2^T + G_2 \underbrace{(P_{k+1} - G_3 Q G_3^T)}_{\Sigma_{k+1}^s} G_2^T = Z_k - G_2(P_{k+1} - \Sigma_{k+1}^s) G_2^T$$

Rauch-Tung-Striebel smoothing:

1. **Iteración causal:** Calcular $\bar{x}_k$, P_k, Z_k con filtro **Kalman** hacia adelante en el tiempo, desde $k = 0$ a $k = N$, iniciando con P_0, $\bar{x}_0$.
2. **Iteración anticausal:** Calcular $\bar{x}_k^s$ y Σ_k^s hacia **atrás en el tiempo**, desde $k = N - 1$ a $k = 0$, inicializando $\bar{x}_N^s = \bar{x}_N$ y $\Sigma_N^s = P_N$.

Conclusiones

- El suavizado no causal en transf. Fourier (discreta), o `filtfilt` tiene su paralelo en representación interna y no estacionario (efecto cond. iniciales de varianza).
- El desarrollo es el **Rauch-Tung-Striebel smoother**, `https://doi.org/10.2514/3.3166`, año 1965.
- Generaliza el filtro de Kalman (causal, mejor predicción dado presente y pasado) a la mejor predicción dado presente, pasado y **futuro**.
- Implementación en dos bucles: un paso hacia adelante de Kalman y una corrección hacia atrás en el tiempo.