Mínimos Cuadrados Recursivos con Olvido

Antonio Sala Piqueras

Dept. Ing. Sistemas y Automatica (DISA)
Universitat Politècnica de València (UPV)

Presentación vídeo en:
http://personales.upv.es/asala/YT/V/mcr2.html

Introducción, Motivación,Objetivos

Motivación: Aplicaciones donde es necesario identificar unos parámetros cambiantes (detección de fallos, control adaptativo).

Objetivos: Comprender las situaciones prácticas donde los algoritmos recursivos con olvido son necesarias; entender la derivación de las fórmulas correspondientes.

Contenidos: Fórmulas con factor de olvido; Filtro de Kalman; motivación en aplicaciones.

Mínimos cuadrados recursivos con olvido: Motivación

- Fórmulas sin olvido: refinar un estimado previo en media e intervalo de confianza (p. ej., modelo teórico + tolerancias) con datos.
  - Siempre se reduce el intervalo de confianza (aumenta precisión).
  - Inspirado en fórmula de Bayes (incorporando probabilidad *a priori*).
- La aplicación a sistemas *variantes en el tiempo* requiere un "olvido" de información antigua.
  - El intervalo de confianza no siempre se reduce, depende de cuánta información pasada se olvida y cómo de informativos son los nuevos datos.
  - Nota: Ello podría resultar problemático si los últimos datos son poco informativos.

Revisión: Mínimos cuadrados recursivos sin olvido

Supongamos que de datos anteriores (o tolerancias de componentes) disponemos de un estimado previo (media y varianza):

$$\theta = \mathcal{N}(\hat{\theta}_0, \Sigma_{\theta,0})$$

Objetivo: conocida la distribución **a priori** de θ, calcular la mejor predicción de θ a partir de dicha información a priori y una nueva serie de datos (y, X), supuesto modelo $y = X\theta + v$.

Mejor predicción incorporando nueva información:

*media: $$\hat{\theta}_1 = \hat{\theta}_0 + \Sigma_{\theta,0} X^T (X\Sigma_{\theta,0}X^T + V)^{-1}(y - X\hat{\theta}_0)$$

*varianza: $$\Sigma_{\theta,1} = \Sigma_{\theta,0} - \Sigma_{\theta,0}X^T(X\Sigma_{\theta,0}X^T + V)^{-1}X\Sigma_{\theta,0}$$

Interpretación alternativa: Filtro de Kalman para $\theta_{k+1} = \theta_k$ (ec. estado) junto con $y_k = X_k\theta_k + v$ (ec. salida).

Mínimos cuadrados recursivos con olvido

En uso "en linea" (control adaptativo), conveniente olvidar datos pasados. Alternativas (entre otras):

1. Hacer mínimos cuadrados sin olvido, pero con una "**ventana**" fija de datos pasados.
2. Factor de olvido $\lambda < 1$, multiplicar $\Sigma_{\theta,i}$ por λ^{-1} para la siguiente iteración (p.ej., $\lambda \approx 0.999$), **subiendo** la varianza a priori.
3. Suponer **dinámica** de los parámetros y hacer filtro de Kalman **completo**:

$$\theta_{k+1} = H\theta_k + w_k$$
$$y_k = X_k\theta_k + v_k$$

- Por ejemplo, $H = I$ (paseo aleatorio), $H = 0.998I$ parámetros tendiendo a un pto. funcionamiento.

Detalles en siguiente transparencia...

Fórmula recursivas con "olvido"

Fórmula sin olvido original:

*media:	$\hat\theta_1 = \hat\theta_0 + \Sigma_{\theta,0}X^T(X\Sigma_{\theta,0}X^T+V)^{-1}(y-X\hat\theta_0)$
*varianza:	$\Sigma_{\theta,1} = \Sigma_{\theta,0} - \Sigma_{\theta,0}X^T(X\Sigma_{\theta,0}X^T+V)^{-1}X\Sigma_{\theta,0}$

[1] **Kalman:** Si suponemos que θ en el siguiente instante verifica $\theta_1 = H\theta_0 + w$, con matriz de varianzas $E(ww^T) = W$, entonces debe sustituirse la "varianza antes de medir" $\Sigma_{\theta,0}$ por $\widehat{S}_{\theta,1} := H\Sigma_{\theta,0}H^T + W$

$E[\theta_1\theta_1^T] = E(H\theta_0\theta_0^T H^T + H\theta_0 w^T + w\theta_0^T H^T + w^T w)$

Sin olvido, caso particular $W = 0$, $H = I \;\Rightarrow\; \widehat{S}_{\theta,1} = \Sigma_{\theta,0}$

[2] **Factor de olvido $\lambda \le 1$:** Aumentar la incertidumbre del siguiente estado (conservar sólo parte de la información): $\widehat{S}_{\theta,1} = \lambda^{-1}\cdot\Sigma_{\theta,0} \ge \Sigma_{\theta,0}$.

*Variación parámetros: $\widehat{S}_{\theta,1} = \overbrace{H\Sigma_{\theta,0}H^T + W}^{\text{Kalman}}$ / $\widehat{S}_{\theta,1} = \overbrace{\lambda^{-1}\Sigma_{\theta,0}}^{\text{Olvido exponencial}}$ $H = I$

*media: $\hat\theta_1 = H\hat\theta_0 + \widehat{S}_{\theta,1}X^T(X\widehat{S}_{\theta,1}X^T+V)^{-1}(y - XH\hat\theta_0)$

*varianza: $\Sigma_{\theta,1} = \widehat{S}_{\theta,1} - \widehat{S}_{\theta,1}X^T(X\widehat{S}_{\theta,1}X^T+V)^{-1}X\widehat{S}_{\theta,1}$

Conclusiones

- Los algoritmos recursivos de mínimos cuadrados incorporan información de nuevos datos junto a a información "a priori".
- Fórmulas con olvido /ruido de parámetros para uso "en línea" en sistemas variantes con el tiempo.
- Aplicaciones a monitorización, diagnóstico, control adaptativo.