

**Métodos robustos para “reconciliación de datos” en Ingeniería
Química y Biotecnología
Ricardo Maronna
Universidad de La Plata y C.I.C.P.B.A**

En el monitoreo de una planta química, y en el análisis de experimentos biotecnológicos, se tienen n magnitudes desconocidas $\xi = (\xi_1, \dots, \xi_n)$ que cumplen q restricciones lineales o no lineales dadas por $G(\xi) = 0$ donde $G : R^n \rightarrow R^q$; y mediciones con error de las mismas dadas por un vector de observaciones $\mathbf{y} = (y_1, \dots, y_m)$ donde $m \leq n$ (algunas magnitudes pueden no estar medidas) con $y_i = \xi_i + e_i$ donde los e_i son errores aleatorios. Las restricciones provienen de ecuaciones de balance (“lo que entra es igual a lo que sale”) y también de las características físico-químicas de la situación. Suele llamarse “reconciliación” al problema de estimar ξ en base a \mathbf{y} . Los e_i se suponen independientes con medias nulas y varianzas σ_i^2 conocidas.

El método clásico de estimación es el de mínimos cuadrados (EMC) ponderado: el estimador es

$$\hat{\xi} = \arg \min_{\xi} \left\{ \sum_{i=1}^m \left(\frac{y_i - \xi_i}{\sigma_i} \right)^2 : G(\xi) = 0 \right\}.$$

En el caso de G lineal, esto es equivalente a un problema de regresión lineal.

Si hay valores atípicos, el EMC es poco confiable, lo que justifica el uso de métodos “robustos”, es decir, poco sensibles a dichos valores. Un enfoque de uso general son los “M-estimadores” dados por

$$\hat{\xi} = \arg \min_{\xi} \left\{ \sum_{i=1}^m \rho \left(\frac{y_i - \xi_i}{\sigma_i} \right) : G(\xi) = 0 \right\},$$

donde ρ es una función par de crecimiento más lento que el cuadrado.

En esta conferencia se tratarán las propiedades de los M-estimadores y sus limitaciones. La principal está dada por la poca redundancia, es decir, que la cantidad de observaciones m , si bien grande, no es suficientemente grande comparada con la cantidad de parámetros desconocidos $n - q$. Se tratarán en particular casos en que la redundancia puede obtenerse reuniendo mediciones realizadas bajo condiciones diferentes.