Sebastián Pérez Márquez
https://seba-perez.github.io
July 21, 2025

Minimización del χ^2 y Selección de Modelos

¿Cuál modelo es mejor?

Usualmente comparamos modelos teóricos con datos observacionales para extraer parámetros importantes. La técnica más común es el ajuste por mínimos cuadrados, utilizando la estadística χ^2 (se pronuncia ”chi cuadrado”):

$$\chi^2 = \sum_i \frac{(O_i - M_i)^2}{\sigma_i^2} \tag{1}$$

donde:

- O_i: valor observado
- M_i: valor predicho por el modelo
- σ_i: incertidumbre asociada a O_i

χ^2 mide qué tan bien el modelo describe los datos. Un valor grande indica mal ajuste, mientras que un valor demasiado pequeño puede indicar sobreajuste.

El χ^2 reducido

Para normalizar el valor de χ^2 respecto al número de datos y parámetros del modelo, se utiliza el χ^2 **reducido**:

$$\chi^2_\nu = \frac{\chi^2}{\nu} = \frac{\chi^2}{N - p} \tag{2}$$

donde:

- N: número de datos
- p: número de parámetros ajustados
- $\nu = N - p$: grados de libertad

Interpretación de χ^2_ν:

- $\chi^2_\nu \approx 1$: buen ajuste
- $\chi^2_\nu \gg 1$: modelo incorrecto o errores subestimados
- $\chi^2_\nu \ll 1$: sobreajuste (modelo demasiado complejo)

Ejemplos típicos

- Ajuste de SEDs
- Órbitas planetarias
- Modelos 2D de imágenes

Conexión con el Maximum Likelihood Estimator (MLE)

Minimizar χ^2 es equivalente a maximizar la verosimilitud bajo el supuesto de errores gaussianos.

Función de verosimilitud

Si los errores son gaussianos e independientes, la probabilidad de observar los datos dados los parámetros del modelo es:

$$P(O|M) = \prod_{i=1}^{N} \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp\left(-\frac{(O_i - M_i)^2}{2\sigma_i^2}\right) \tag{3}$$

Tomando el logaritmo (log-verosimilitud):

$$\begin{aligned} \ln P(O|M) &= \sum_{i=1}^{N} \left[-\frac{1}{2}\ln(2\pi\sigma_i^2) - \frac{(O_i - M_i)^2}{2\sigma_i^2}\right] && (4) \\ &= -\frac{1}{2}\sum_{i=1}^{N} \left[\ln(2\pi\sigma_i^2) + \frac{(O_i - M_i)^2}{\sigma_i^2}\right] && (5) \end{aligned}$$

Como el primer término no depende del modelo, maximizar la verosimilitud es equivalente a minimizar:

$$\chi^2 = \sum_{i=1}^{N} \frac{(O_i - M_i)^2}{\sigma_i^2} \tag{6}$$

¿De dónde viene esta suposición?

La suposición de errores gaussianos proviene del Teorema Central del Límite: si repetimos un experimento muchas veces, la distribución de los errores tenderá a una gaussiana, con valores más probables cercanos al valor real.

Notas adicionales

- Los errores gaussianos son simétricos y caracterizados completamente por su desviación estándar.
- La probabilidad de error decrece rápidamente para desviaciones grandes (e^{-x^2}).
- Este marco permite estimar parámetros y sus incertidumbres de forma coherente.

Estimación de incertidumbres en los parámetros ajustados

Una vez encontrado el conjunto de parámetros $\{a_j\}$ que minimizan el valor de χ^2, es posible estimar sus incertidumbres a partir de la forma de la función χ^2 en el entorno del mínimo.

Criterio del $\Delta\chi^2 = 1$

Bajo la suposición de que el modelo es correcto y los errores son gaussianos, la incertidumbre estándar (1σ) en un parámetro ajustado a_j corresponde al cambio en a_j necesario para aumentar el valor mínimo de χ^2 en 1 unidad, manteniendo los demás parámetros fijos:

$$\Delta\chi^2 = \chi^2(a_j + \delta a_j) - \chi^2_{\text{min}} = 1 \tag{7}$$

Este método proporciona una estimación de la incertidumbre basada en la curvatura de la función χ^2 alrededor del mínimo.

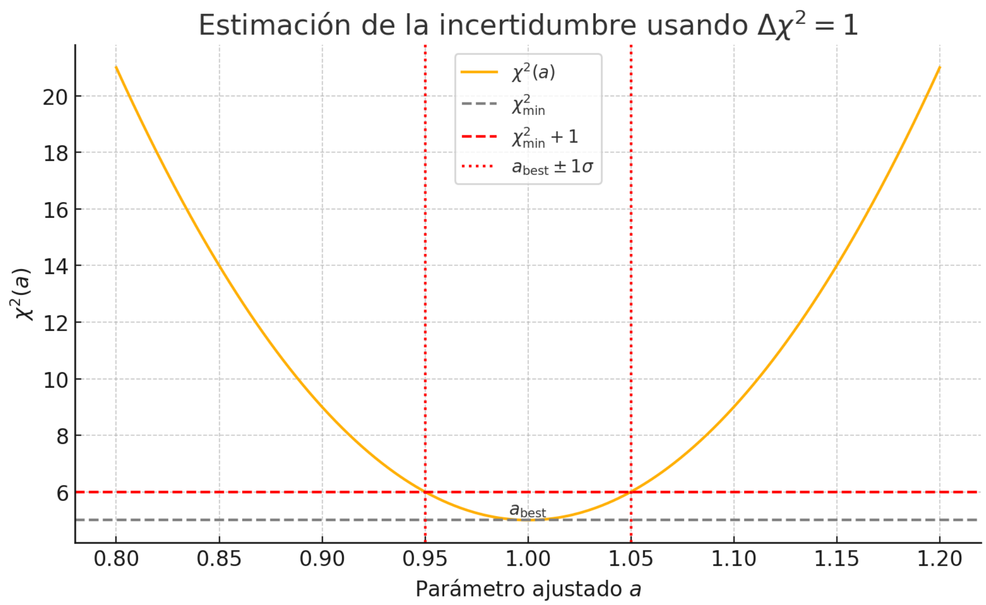

Figure 1: Estimación de la incertidumbre de un parámetro ajustado usando el criterio de $\Delta\chi^2 = 1$. La curva muestra cómo varía χ^2 con el parámetro a, y las líneas verticales marcan el intervalo de 1σ, correspondiente al rango donde $\chi^2(a) = \chi^2_{\text{min}} + 1$.

Casos generales

Este criterio se generaliza:

- $\Delta\chi^2 = 1$: intervalo de confianza del 68.3% para un solo parámetro libre
- $\Delta\chi^2 = 2.3$: para un intervalo de confianza conjunto del 68.3% en dos parámetros
- $\Delta\chi^2 = 3.5$: para tres parámetros, y así sucesivamente

Relación con χ^2_ν

En caso de que $\chi^2_\nu > 1$, se puede sospechar que los errores fueron subestimados. En ese caso, las incertidumbres estimadas deben escalarse por un factor:

$$\sigma_{\text{ajustada}} = \sqrt{\chi^2_\nu} \cdot \sigma_{\text{inicial}} \tag{8}$$

Este procedimiento corrige las barras de error para reflejar mejor la dispersión real observada en los datos.

Métodos prácticos

En la práctica, las incertidumbres también se estiman:

- A partir de la matriz de covarianza, calculada como la inversa de la matriz Hessiana de χ^2
- Mediante técnicas numéricas como `scipy.optimize.curve_fit` o métodos de Monte Carlo/MCMC