

# Herramientas Computacionales para la Astroinformática

Cristian A. Vega-Martínez  
Facundo A. Gómez



# Introducción a los métodos numéricos

Aproximación, interpolación y ajustes

# ¿Qué son los métodos numéricos?

El análisis numérico, que es la base de los métodos numéricos, es la disciplina que trata con algoritmos para la **manipulación** y **solución** de problemas matemáticos.

- Nos permite modelar y resolver problemas matemáticos complejos utilizando la computación
- Este campo de estudio es esencial para una variedad de disciplinas, incluyendo la física, la ingeniería, la economía y la ciencia de datos.
- Entre las problemáticas más importantes abordadas por estos métodos, encontramos: interpolación y aproximación, diferenciación e integración (EDOs y EDPs), y optimización.



<https://scipy.org/>

Muchas implementaciones de estas técnicas se pueden encontrar en la biblioteca SciPy, la cual implementa un gran número de algoritmos de alto nivel utilizando métodos avanzados y eficientes. Por ejemplo, incluye:

- Integración Numérica para funciones y EDO (**scipy.integrate**)
- Optimización: minimización de funciones, ajustes de curvas, la búsqueda de raíces (**scipy.optimize**)
- Interpolación de datos uni- y multi-dimensionales (**scipy.interpolate**)
- Procesamiento de Señales (**scipy.signal**)
- Álgebra Lineal más avanzada que NumPy (**scipy.linalg**)
- Estadísticas: distribuciones de probabilidad, funciones estadísticas y tests (**scipy.stats**)
- Procesamiento de Imágenes multidimensionales (**scipy.ndimage**)
- y más...

# Interpolación y aproximación

Ambas técnicas se utilizan para construir una función que se aproxima a un conjunto dado de puntos de datos, pero siguiendo estrategias distintas:

**Interpolación:** es un proceso que define una función que pase **exactamente** por un conjunto dado de puntos, "rellenando" los vacíos entre datos conocidos. Por ejemplo, la interpolación polinomial: se busca un polinomio que pase por todos los puntos definidos por los datos.

Interpolación unidimensional con SciPy: `scipy.interpolate.interp1d`

**Aproximación:** proceso que busca una función que se aproxime a los puntos de datos, sin necesariamente pasar por ellos. Este método apunta a encontrar una **coincidencia cercana que minimice el error total** (e.j. para representar datos con ruido o errores). Un ejemplo común es la regresión lineal, donde se busca una recta que minimice la distancia cuadrada total de cada punto a la recta.

Regresión lineal con SciPy: `scipy.stats.linregress`

Ejemplos de estas técnicas son desarrollados en jupyter notebook

# Ajuste de curvas

Contrario a las técnicas de aproximación, el ajuste de curvas busca **modelar** un conjunto de datos con una función, apuntando a encontrar el **mejor ajuste de los parámetros del modelo**.

- El proceso de ajuste de curvas generalmente se realiza mediante el método de los **mínimos cuadrados**, donde se busca la función que minimiza la suma de los cuadrados de las diferencias entre los datos y la función de ajuste:

Dado un conjunto de  $n$  puntos  $(x_i, y_i)$ ,  $i=1..n$ , y una función  $f(x, \mathbf{P})$  donde  $\mathbf{P}$  representa un conjunto de  $m$  parámetros  $\mathbf{P} = (p_1, p_2, \dots, p_m)$ , se busca el conjunto de valores de  $\mathbf{P}$  tal que:

$$S(P) = \sum_{i=1}^n (y_i - f(x_i, P))^2$$

alcance el valor mínimo posible.

# Ajuste de curvas

**Mínimos cuadrados:**  $S(P) = \sum_{i=1}^n (y_i - f(x_i, P))^2$

Esto se puede **resolver analíticamente** considerando que el mínimo ocurre cuando las derivadas de la función con respecto a cada parámetro valen cero, es decir:

$$\frac{\partial S}{\partial p_k} = 0, \quad k = 1..m$$

lo cual define un sistema de  $m$  ecuaciones en **cuadratura**.

En caso que el sistema no sea solucionable analíticamente, se puede realizar la minimización mediante métodos más avanzados de análisis numérico (por ejemplo: Gauss-Newton, Descenso de gradiente, Levenberg-Marquardt, etc.).

Ajuste de curvas con SciPy: `scipy.optimize.curve_fit`

# Ajuste de curvas para modelado de datos

Relacionada con los mínimos cuadrados, la técnica estadística de **reducción (o prueba) de  $\chi^2$  (chi cuadrado)**, es ampliamente utilizada para determinar qué tan bien un **modelo teórico** se ajusta a los **datos observados**.

La función a minimizar es similar a los mínimos cuadrados, pero cada diferencia se pesa por la varianza teórica (o error) asociado a cada dato. Es decir:

$$\chi^2(P) = \sum_{i=1}^n \frac{(y_i - f(x_i, P))^2}{\sigma_i}$$

Si se reemplaza  $\sigma_i$  por el valor teórico  $f(x_i, P)$ , se obtiene la **prueba chi-cuadrado de Pearson**.



# Ajuste de curvas para modelado de datos

También es muy utilizado el **chi-cuadrado reducido** para evaluar la bondad del ajuste y comparar distintos modelos. Esta función divide el valor de  $\chi^2$  por el número de grados de libertad del problema:

$$\chi_r^2(P) = \chi^2 / \nu,$$

$$\nu = n - m$$

donde los grados de libertad se calculan como el número de datos observados ( $n$ ) menos el número de parámetros del modelo ( $m$ ).

En esta definición, los valores de  $\chi^2$  cercanos a uno indican un buen ajuste del modelo, mientras que un valor muy superior a 1 o muy inferior a 1 indica un ajuste pobre.