

# Toy / Solvable Model：雜訊下自動發現「簡諧振子定律」與守恆量

## 1. 問題設計（為何這是第一步）

為了邁向「AI 能在大規模、含雜訊的資料中自動提出、化簡並驗證人類可理解的物理定律」的長程目標，本步選擇一個可解析、可量化、可重現的玩具系統：一維簡諧振子。觀測資料為

$$y_i = x(t_i) + \varepsilon_i, \quad x(t) = A \cos(\omega t + \phi), \quad \varepsilon_i \sim \mathcal{N}(0, \sigma^2).$$

目標是在 **未知模型與參數** 的前提下，從時間序列自動提出並化簡出

$$x'' + \omega^2 x = 0,$$

並驗證對應的守恆量（機械能）是否穩定；這對應最終專題中「從行星軌跡發現定律與守恆量」的縮小版流程。

## 2. 模型與方法

(1) **平滑與導數估計** 以 Savitzky–Golay 或樣條方法在固定窗長下估計  $x(t), x'(t), x''(t)$ ，兼顧抗噪與邊界效應。將離散取樣記為向量  $\mathbf{x}, \mathbf{x}', \mathbf{x}''$ 。

(2) **參數一致化（封閉解）** 在理想無雜訊時， $\min_{\omega^2} \|\mathbf{x}'' + \omega^2 \mathbf{x}\|_2^2$  的封閉解為

$$\widehat{\omega^2} = -\frac{\mathbf{x}^\top \mathbf{x}''}{\mathbf{x}^\top \mathbf{x}}.$$

有雜訊時，此估計在多週期覆蓋與適度訊雜比下相合；可用自助法（bootstrap）估計  $\widehat{\omega}$  的置信區間。

(3) 符號化化簡 (SINDy) 建構函數庫  $\Theta = [1, x, x', x^3]$ ，以稀疏回歸擬合

$$x'' \approx \xi_0 + \xi_1 x + \xi_2 x' + \xi_3 x^3.$$

期望在適當正則化下得到  $(\xi_0, \xi_1, \xi_2, \xi_3) \approx (0, -\omega^2, 0, 0)$ ，即自動「提出並化簡」最小充分的方程形式。

(4) 守恆量驗證（可解釋性）以  $\hat{\omega}$  構造

$$\hat{E}(t) = \frac{1}{2}(x'(t))^2 + \frac{1}{2}\hat{\omega}^2(x(t))^2.$$

對無阻尼系統，時間內變異  $\text{Var}_t[\hat{E}(t)]$  應趨近於 0；若刻意加入小阻尼， $\hat{E}(t)$  應呈單調遞減，且 SINDy 會自動保留  $x'$  的線性項。

### 3. 實作與展示（最小可行）

- 以 Python 產生三組資料： $\sigma \in \{0, 0.05, 0.1\}$ ，取樣頻率覆蓋多個週期。
- 圖 1：平滑前後的  $x, x', x''$ （同一圖像分三條曲線）。
- 圖 2：SINDy 係數條形圖（理想情況僅  $\xi_1$  非零且  $\xi_1 \approx -\hat{\omega}^2$ ）。
- 圖 3： $\hat{E}(t)$  的時間序列與其變異值（圖題標註  $\text{Var}_t[\hat{E}]$ ）。
- 表 1（指標）： $|\hat{\omega} - \omega|$ 、 $\|x'' + \hat{\omega}^2 x\|_2 / \|x''\|_2$ 、非零係數數量（模型簡潔度）、 $\text{Var}_t[\hat{E}]$ ，於三種噪聲下的結果。

### 4. 討論與通往最終目標

本玩具模型同時涵蓋：**抗噪流程**（平滑與導數估計）、**提出與化簡**（SINDy 稀疏回歸）、**可解釋驗證**（守恆量），且皆具可重現的量化指標；這與「從行星軌跡資料中自動發現中心力與反平方律、推得角動量守恆」的關鍵步驟高度對齊。下一步可將  $x \in \mathbb{R}$  擴展為  $\mathbf{r} \in \mathbb{R}^2$ ，把庫函數擴充為  $[\mathbf{r}, \mathbf{v}, 1/\|\mathbf{r}\|^2]$ ，驗證角動量  $L = \mathbf{r} \times \mathbf{v}$  的守恆，並以交叉驗證或最小描述長度（MDL）比較反平方律相對其他候選力律的優勢。

**重現性附註：**建議 repo 包含 /src（資料生成、平滑與導數、SINDy、作圖）、/figs（三張圖）、report.md 與 requirements.txt；README 提供三條指令即可重現所有圖表。