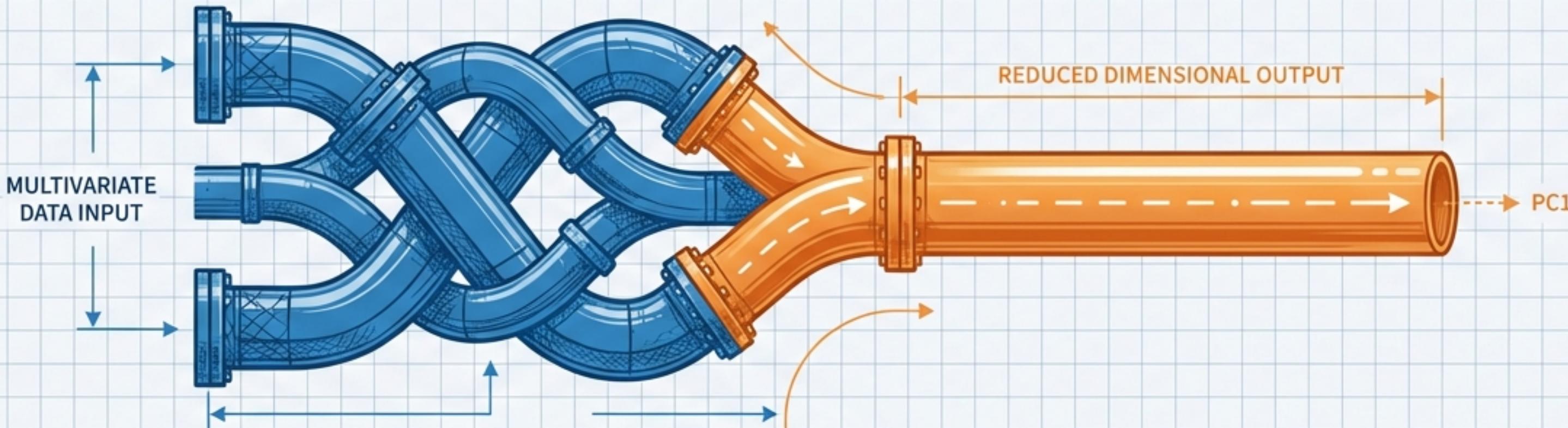


# Unit 06 | 主成分分析 (Principal Component Analysis)



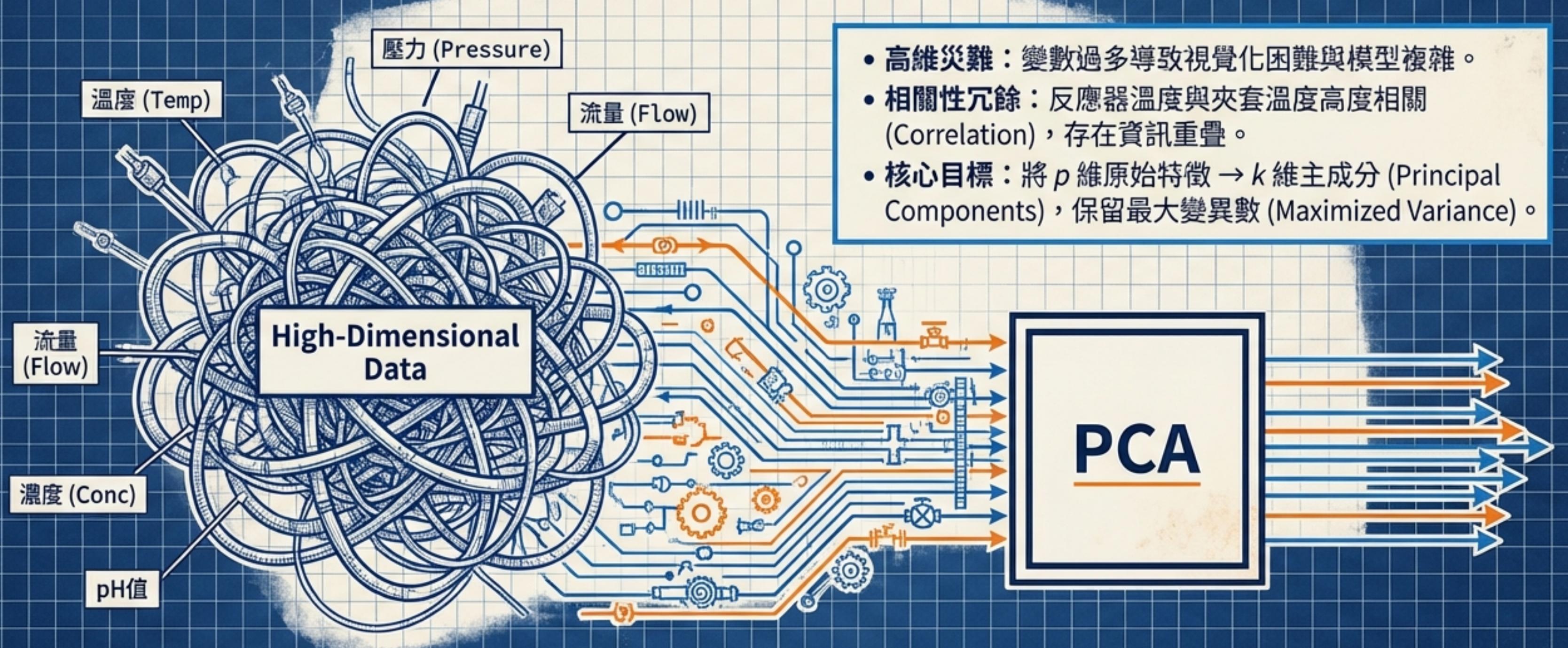
## 降維技術與化工製程監控

Course: AI在化工上之應用

Lab: 逢甲大學 智慧程序系統工程實驗室

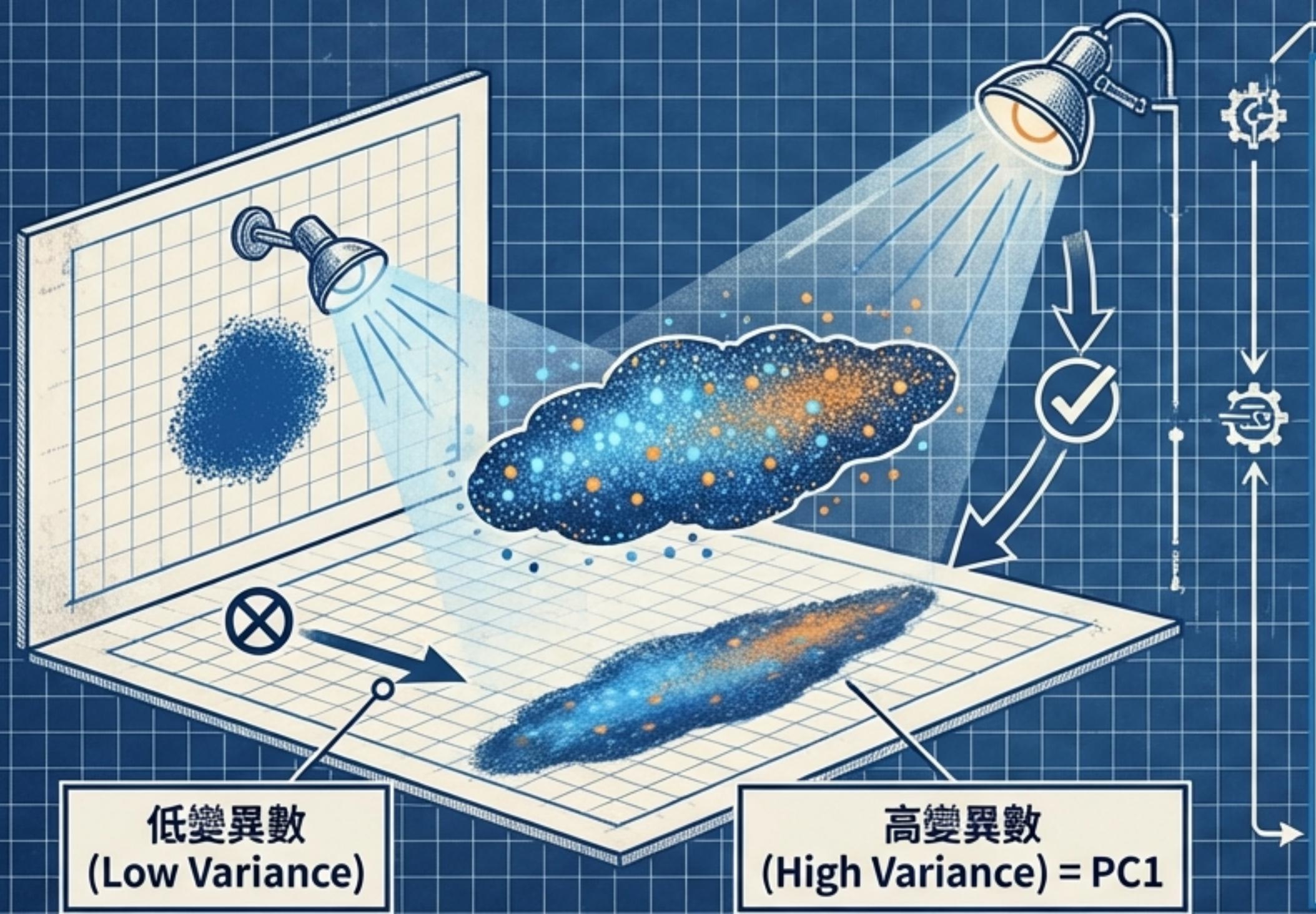
Instructor: 莊曜禎 助理教授

# 數據過載：化工製程面臨的高維度挑戰



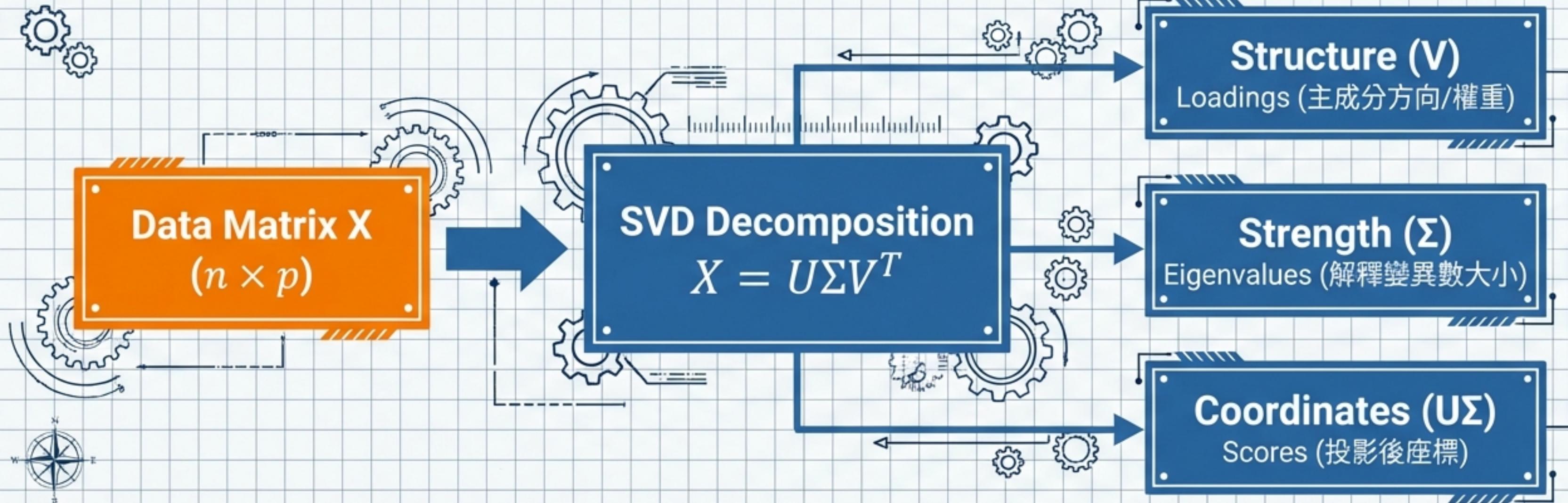
數百個感測器同時運作，難以單獨監控

# 核心原理：尋找數據的最佳投影視角



1. **投影 (Projection)**：將高維數據投射到低維平面。
2. **變異數最大化**：找到數據分布最廣的方向作為第一主成分 (PC1)。
3. **正交性 (Orthogonality)**：PC2 必須與 PC1 垂直 ( $90^\circ$ )，捕捉剩餘的變異資訊。
4. **去關聯 (Decorrelation)**：新的主成分之間互不相關，消除了共線性。

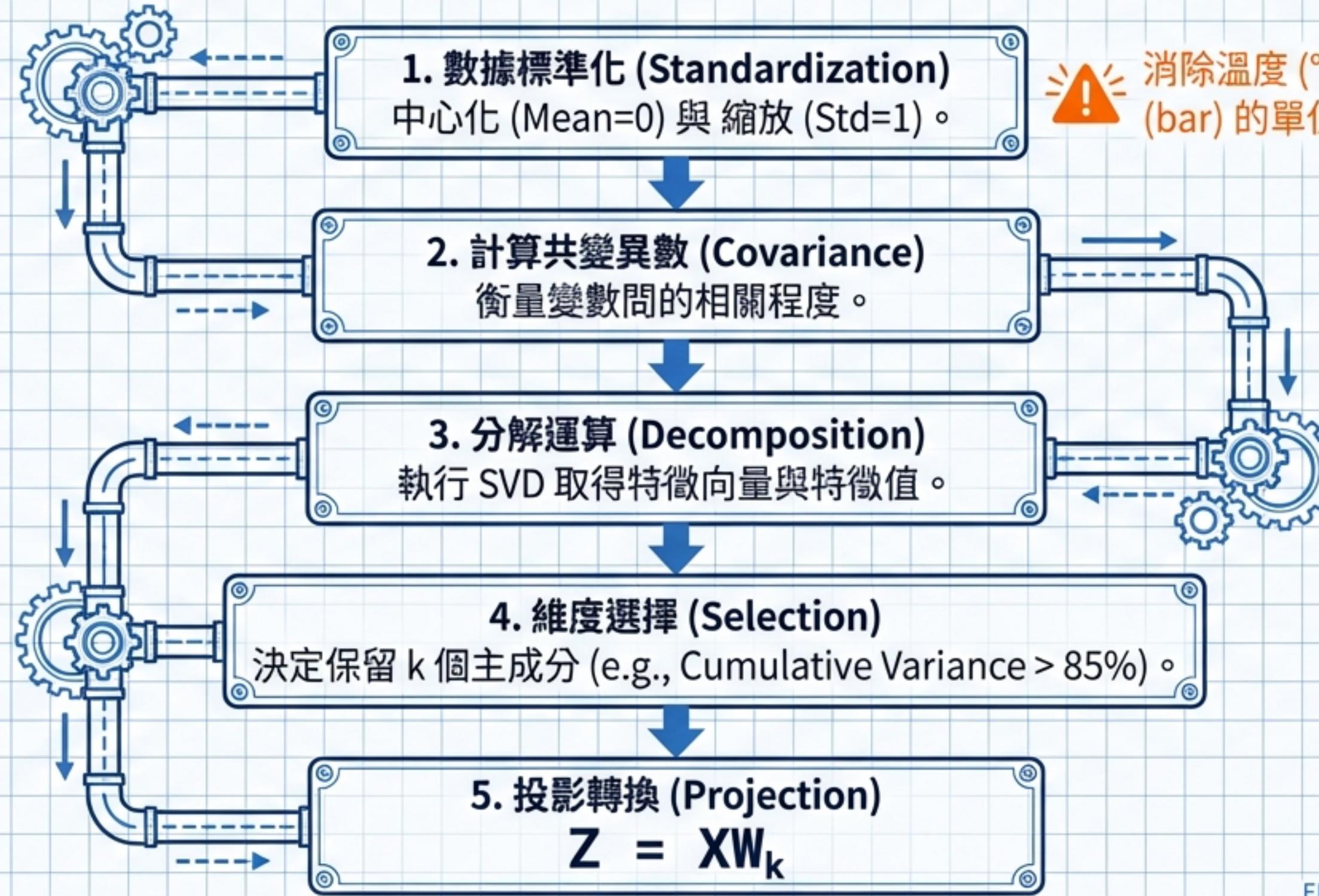
# 數學引擎：奇異值分解 (SVD)



Key Relationship:  $\lambda_i = \sigma_i^2 / (n-1)$

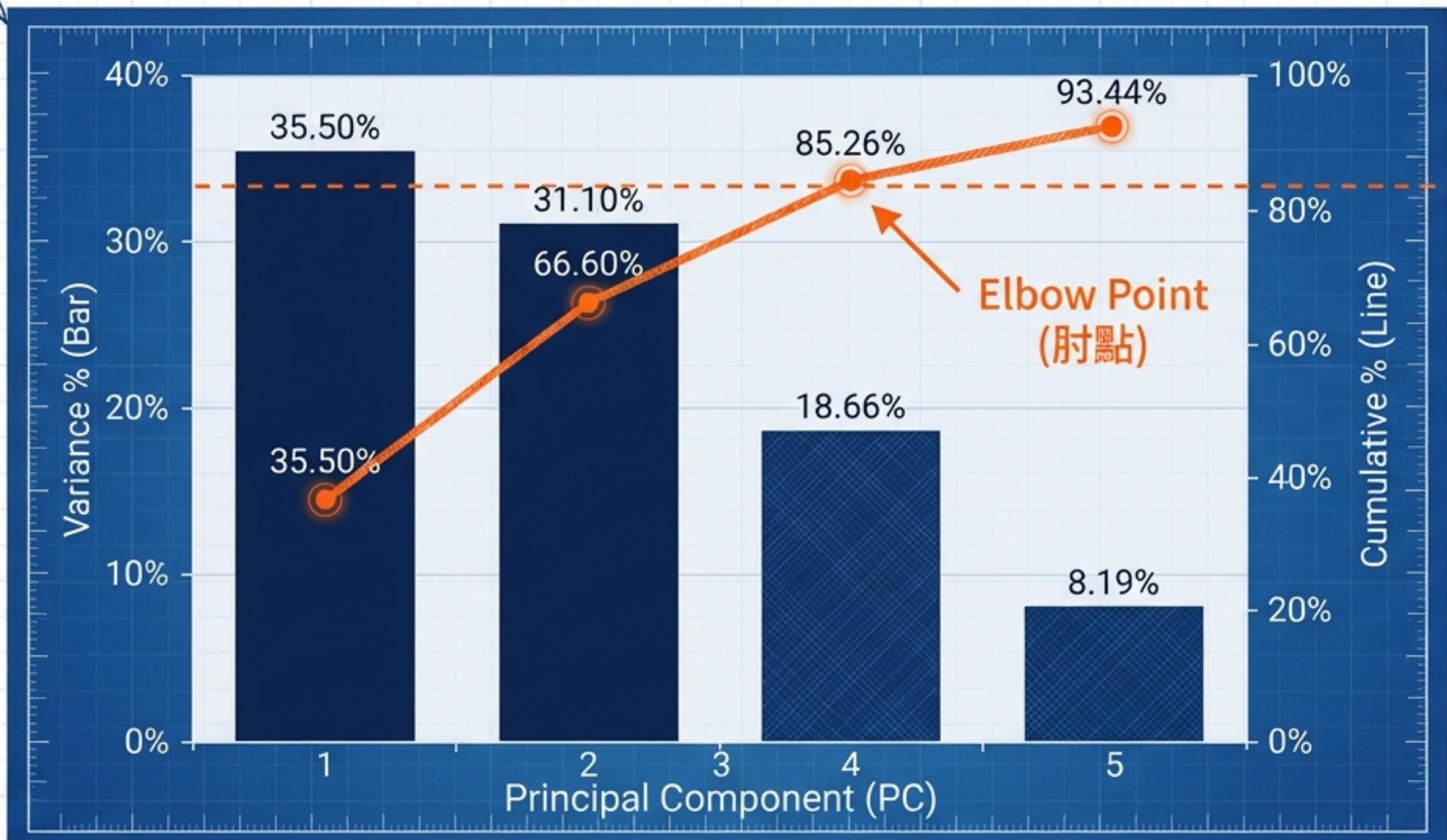
Note: SVD 是數值上最穩定且高效的解法，優於傳統特徵值分解。

# PCA 演算法標準作業流程 (SOP)



ENGINEERING SPEC

# 維度決策：陡坡圖 (Scree Plot) 分析

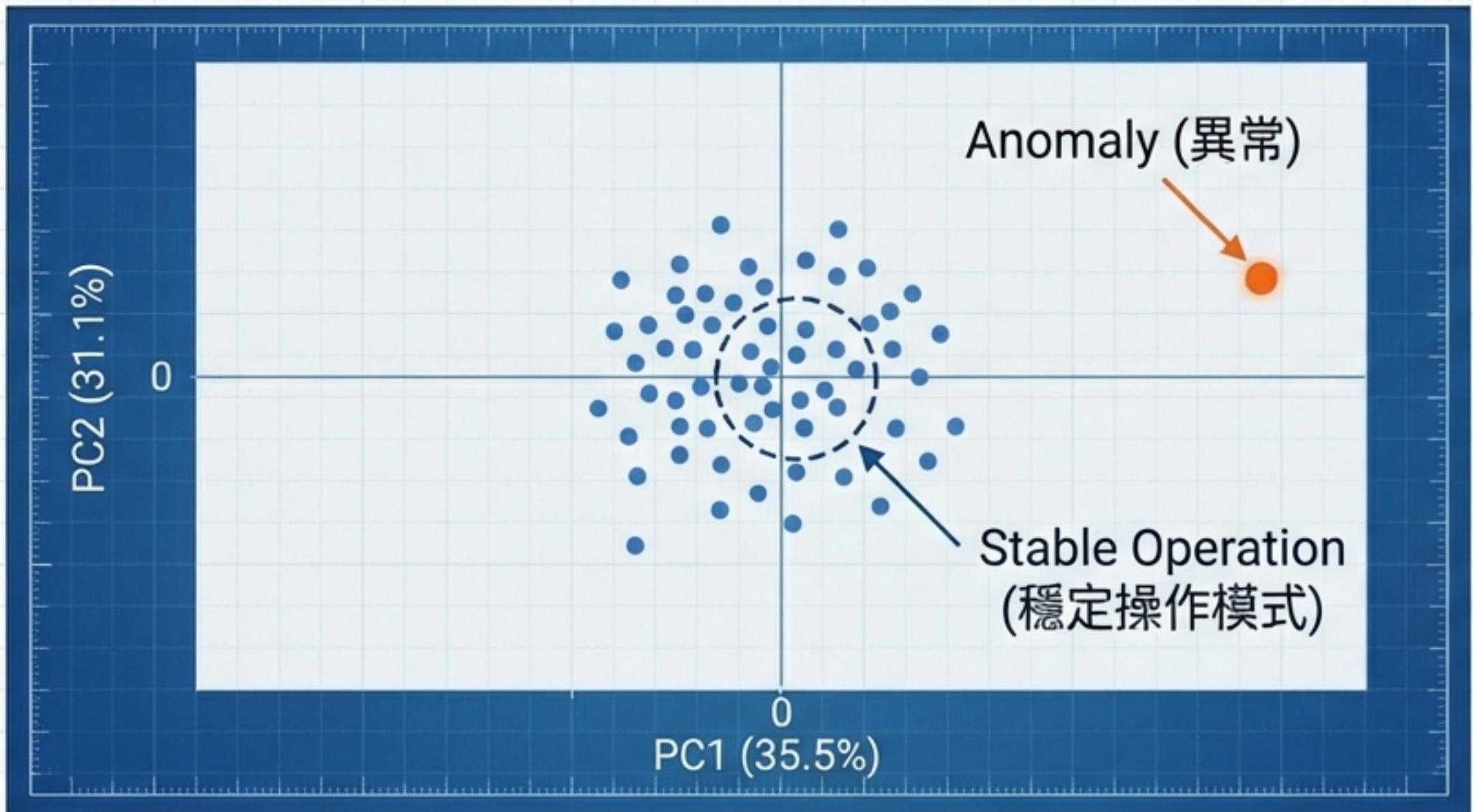


Information Retention Threshold

前三個主成分保留了 85.26% 的製程資訊。PC4 之後僅為雜訊，故選擇 k=3。

# 讀懂儀表板 (1)：得分 (Scores, Z)

原始數據在主成分空間的投影座標 (GPS Coordinates)



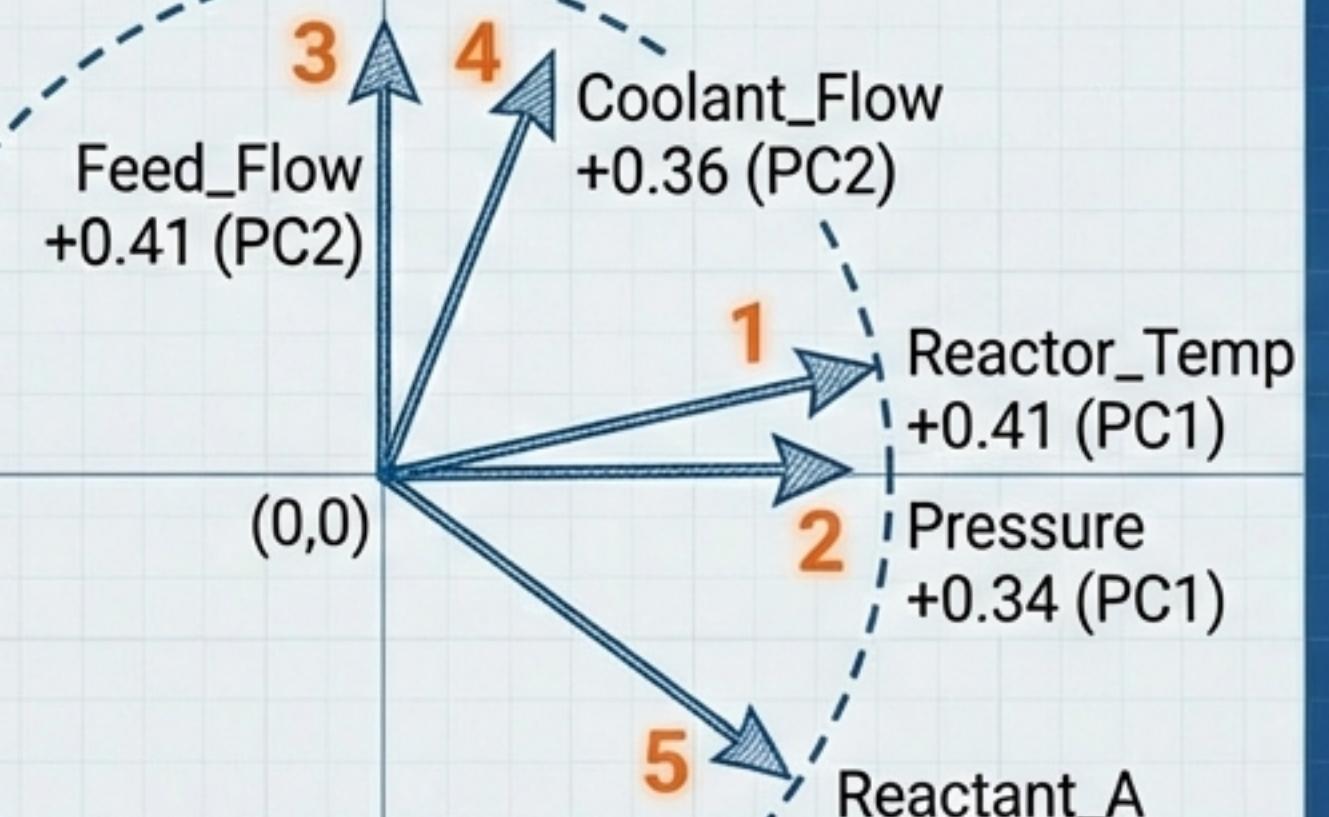
- 位置 (Location)：點的位置代表該樣本的操作狀態。
- 距離 (Distance)：點與點越近 → 操作條件越相似。
- 離群 (Outlier): 遠離中心的點代表異常。

如果 PCA 是地圖，Score 就是每一個樣本在這張地圖上的經緯度。

# 讀懂儀表表板 (2) : 載荷 (Loadings, W)

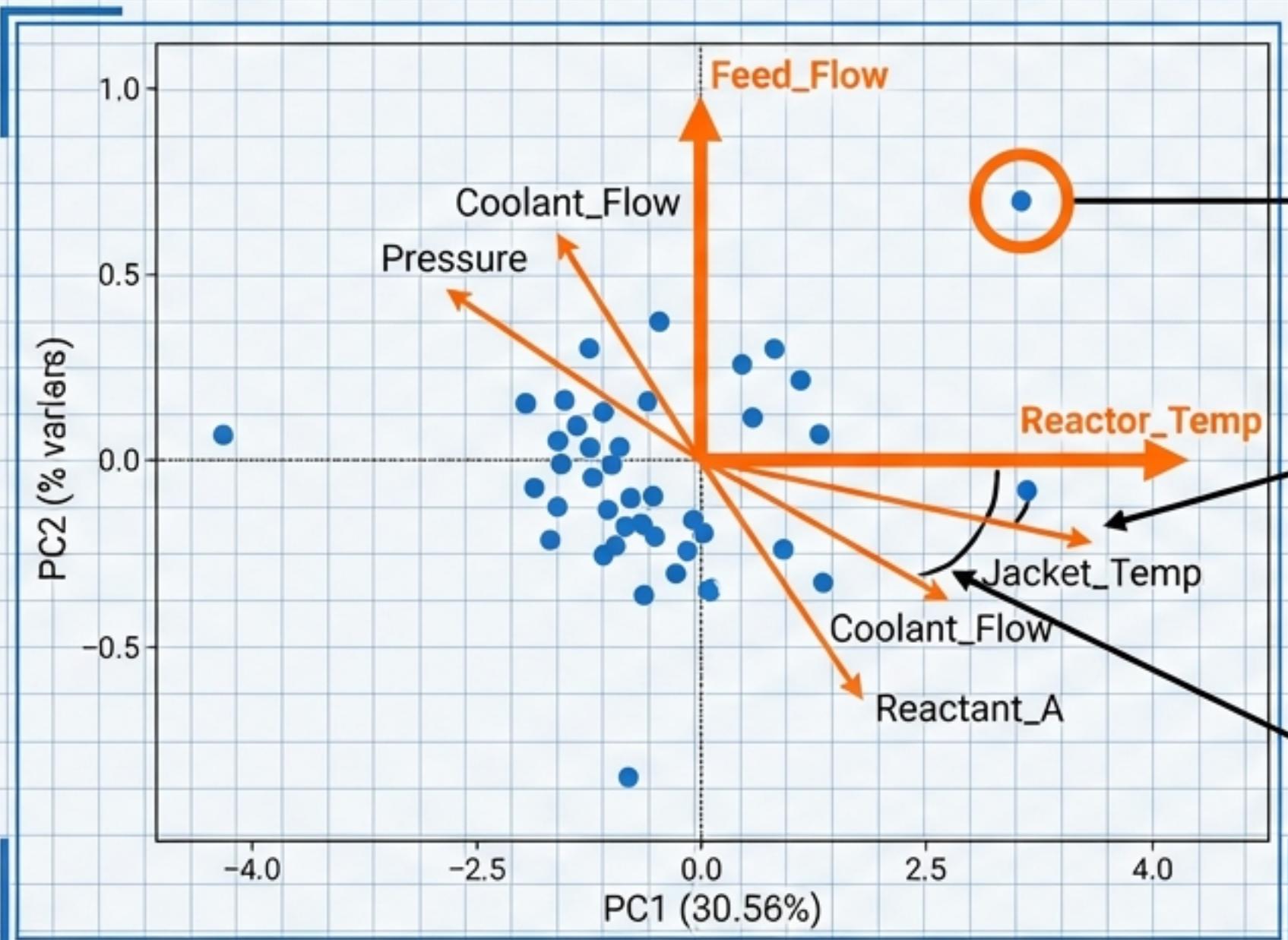
原始特徵對主成分的貢獻係數 (Contribution Weights)

Vector Plot



	PC1	PC2	
PC1	Reactor_Temp	Pressure	熱力學狀態
PC2	Feed_Flow	Coolant_Flow	流量控制
PC3	Reactant_A		濃度平衡

# | 綜合視圖 : Biplot (Scores + Loadings)



1. 樣本落在箭頭指向的方向  
→ 該變數數值偏高。
2. 箭頭夾角小 → 變數高度正相關  
(e.g., Reactor\_Temp & Jacket\_Temp)。
3. 箭頭夾角  $90^\circ$  → 變數不相關  
(e.g., Temp & Flow)。

Biplot 讓我們同時看見『樣本的分布』與『造成分布的變數原因』。

# 製程監控核心：雙軌檢測系統

## Hotelling's $T^2$



監控系統性偏移  
(Systematic Shift)

Formula:

$$T^2 = \sum \frac{\text{Score}}{(\text{Variance})^2}$$

樣本在模型內的距離 (Distance within Model)

## SPE (Q-Statistic)



監控新型態故障  
(New Fault Type)

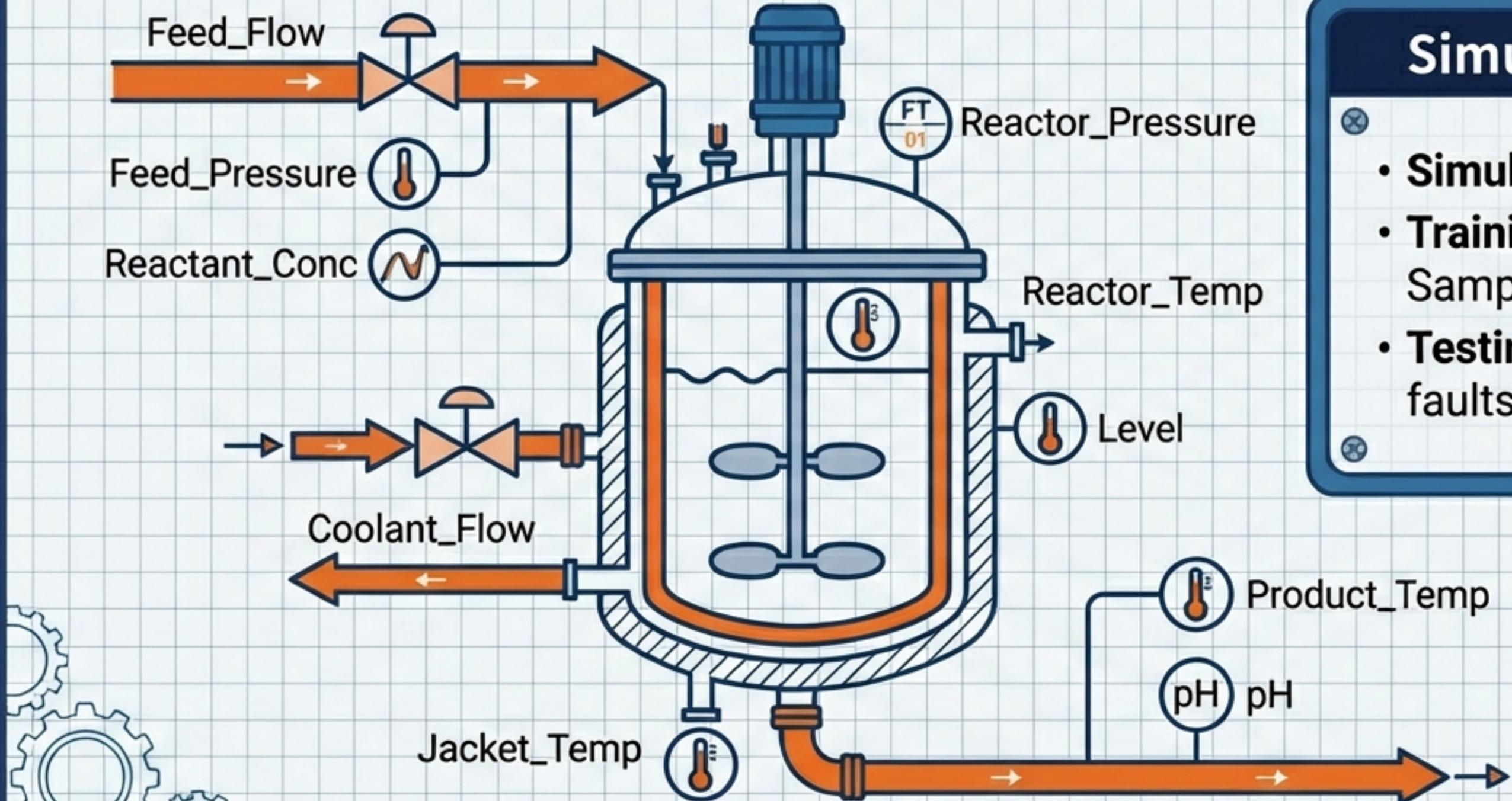
Formula:

$$\text{SPE} = \| \mathbf{x} - \hat{\mathbf{x}}_{\text{hat}} \|^2$$

樣本與模型平面的距離 (Residuals)

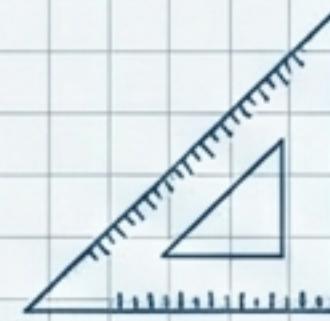
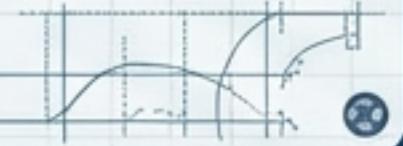
→  $T^2$  監控是否開太快 (幅度) ; SPE 監控是否開出跑道 (模式)。

# | 案例研究 : CSTR 反應器監控

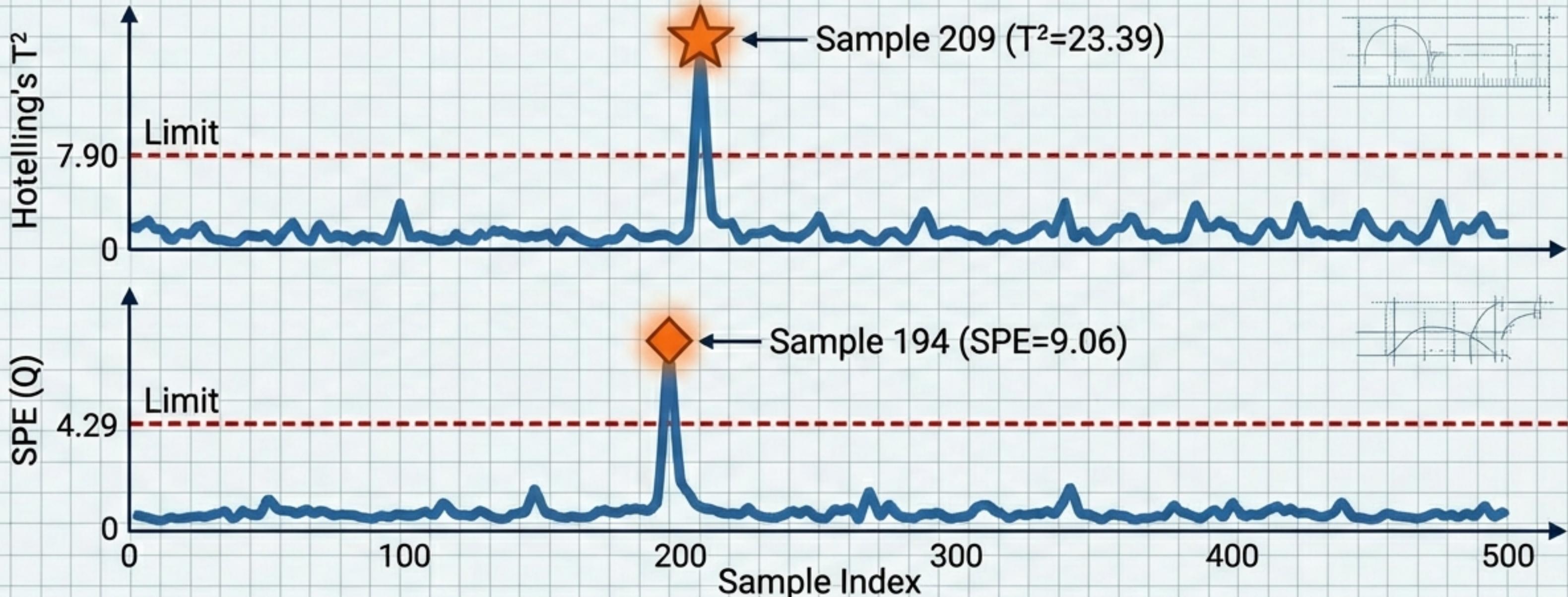


## Simulation Context

- **Simulation:** 500 Samples
- **Training:** 400 Normal Samples
- **Testing:** Includes simulated faults



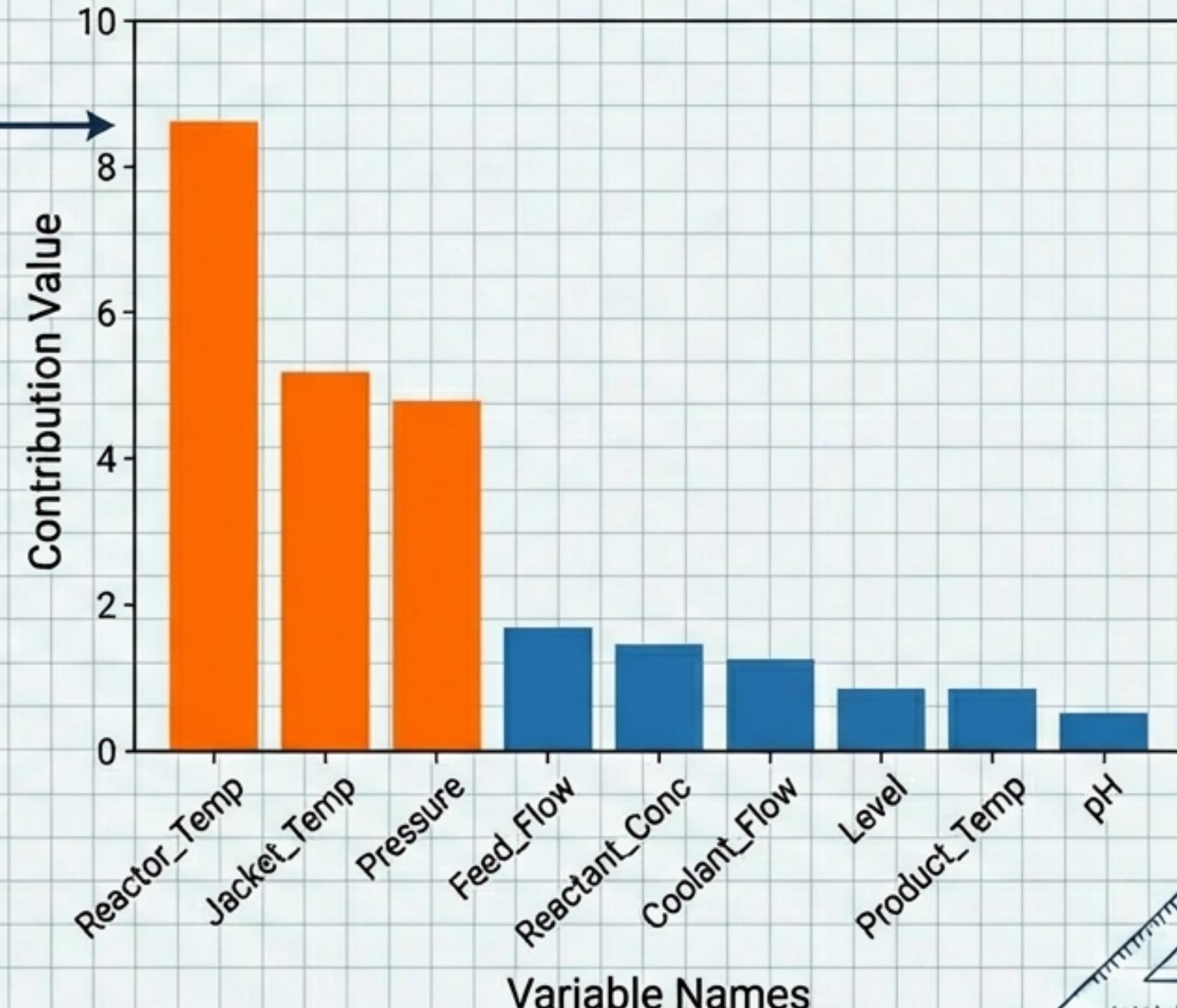
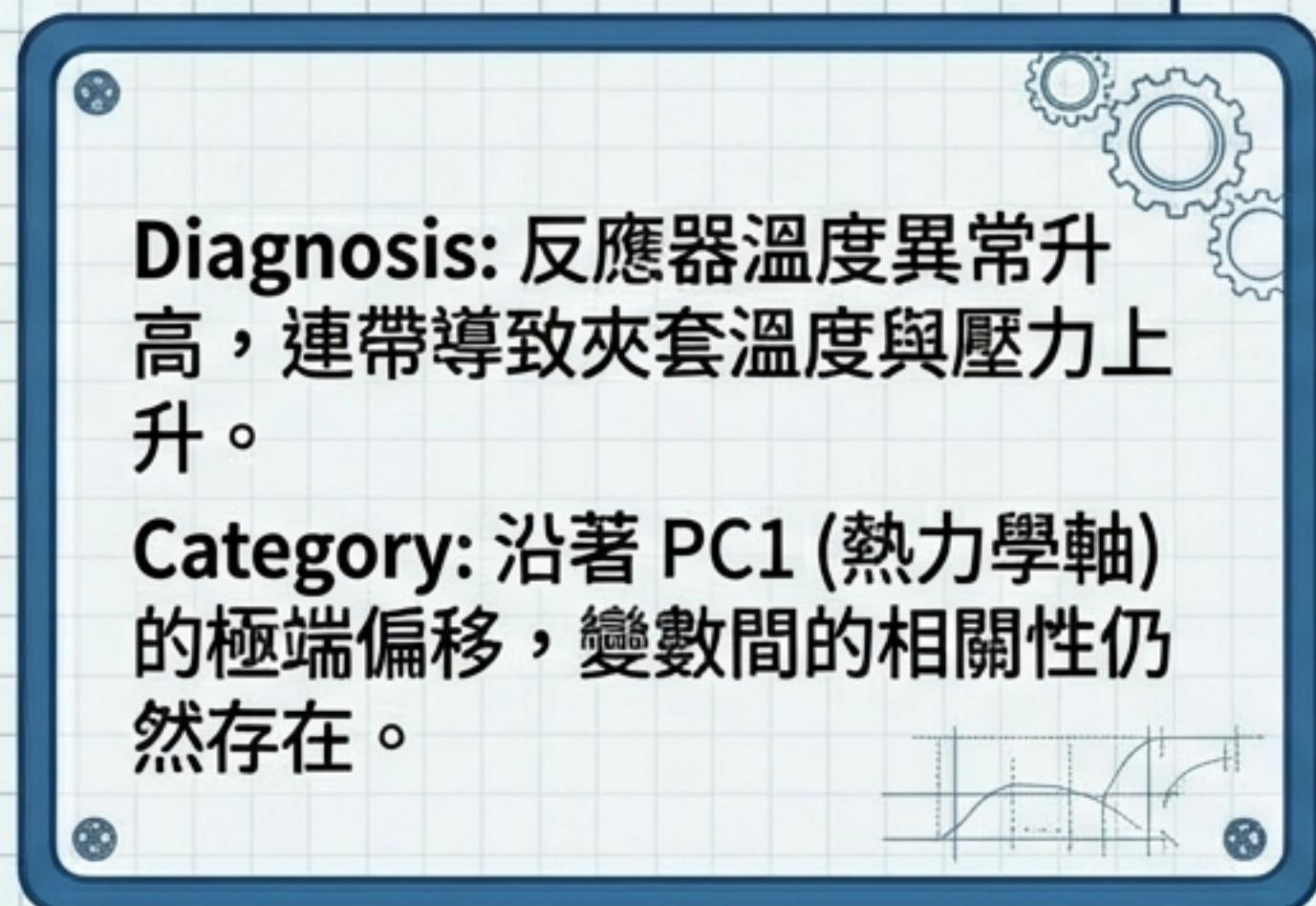
# | 異常偵測結果：控制圖 (Control Charts)



Sample 209:  $T^2$  嚴重超標 (系統性偏移)。  
Sample 194: SPE 嚴重超標 (新型態變異)。

# | 故障診斷 (1) : 系統性偏移 (Sample 209)

## Max $T^2$ Anomaly Analysis

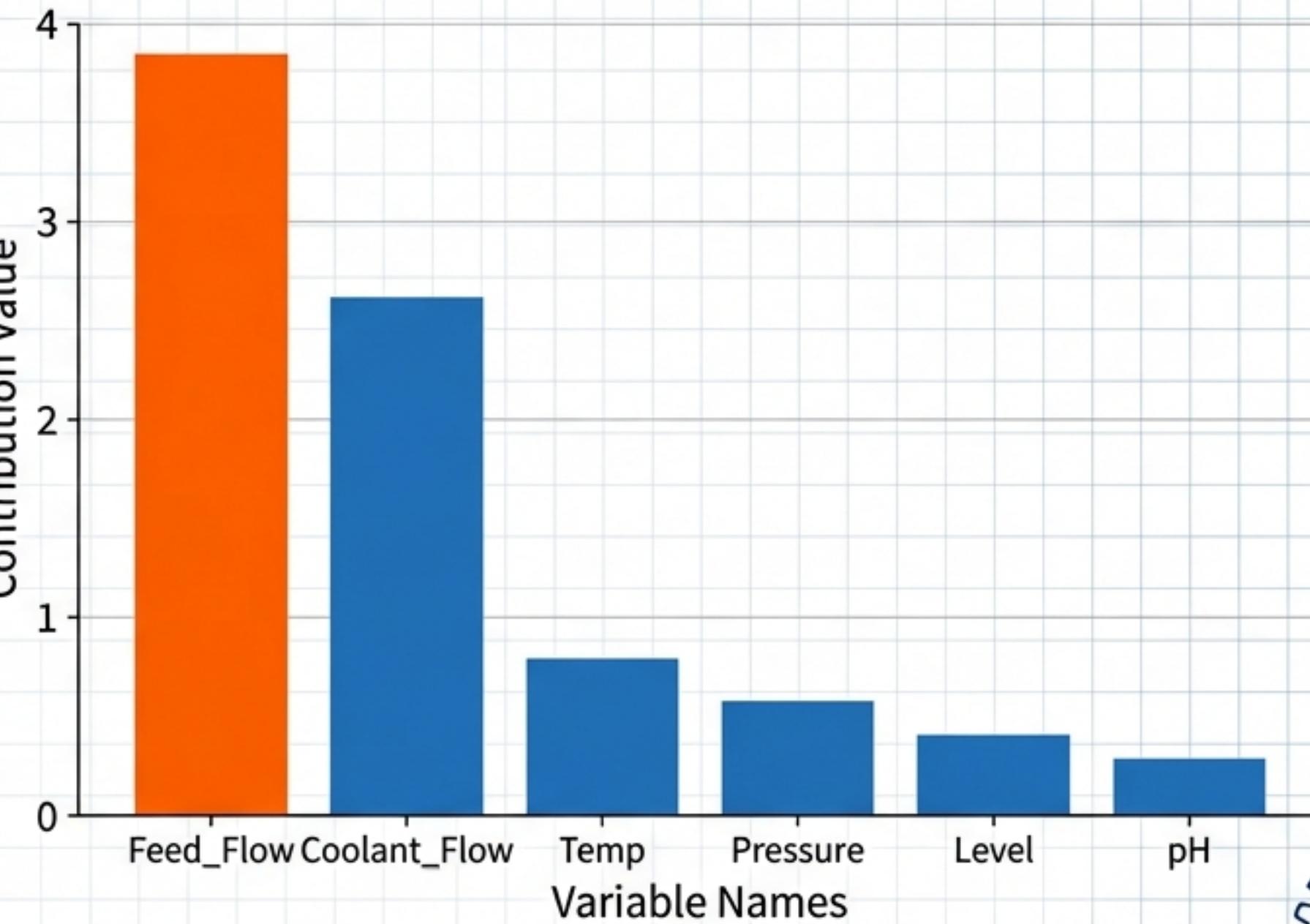


# 故障診斷 (2)：新型態變異 (Sample 194)

Diagnosis: 進料流量控制出現異常模式，且未依照正常邏輯觸發冷卻響應。

Category: PCA 模型平面外的變異 (Model Mismatch)，代表訓練數據中未曾見過的故障模式。

Max SPE Anomaly Analysis



# 實務應用總結與檢查清單



“PCA 不僅是數學公式，它是將數據轉化為決策的透鏡。”

Next Step: 進階主題：**動態 PCA (DPCA)** 與 **非線性 Kernel PCA**。