

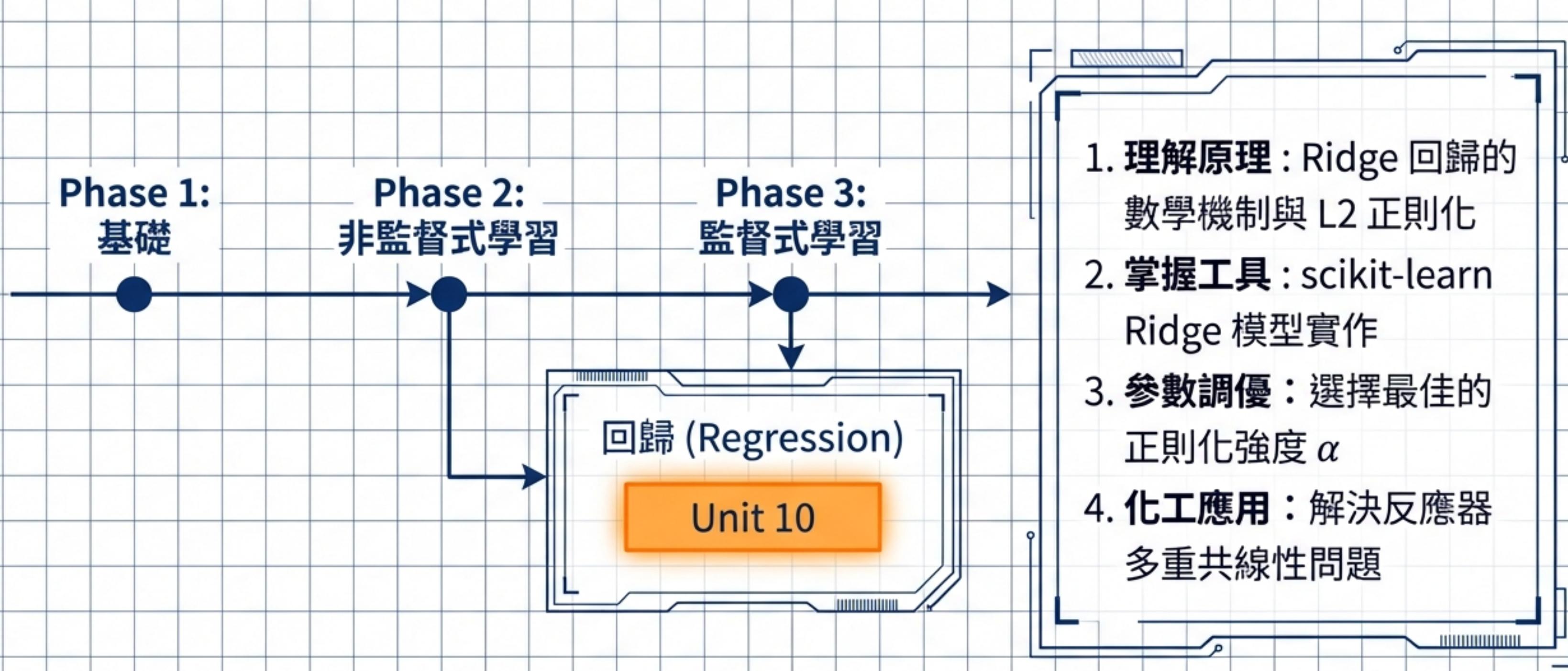
Unit 10 Ridge 回歸 (Ridge Regression)

L2 正則化與多重共線性解決方案



授課教師：莊曜禎 助理教授
課程：AI在化工上之應用
更新日期：2026-01-28

課程地圖與學習路徑

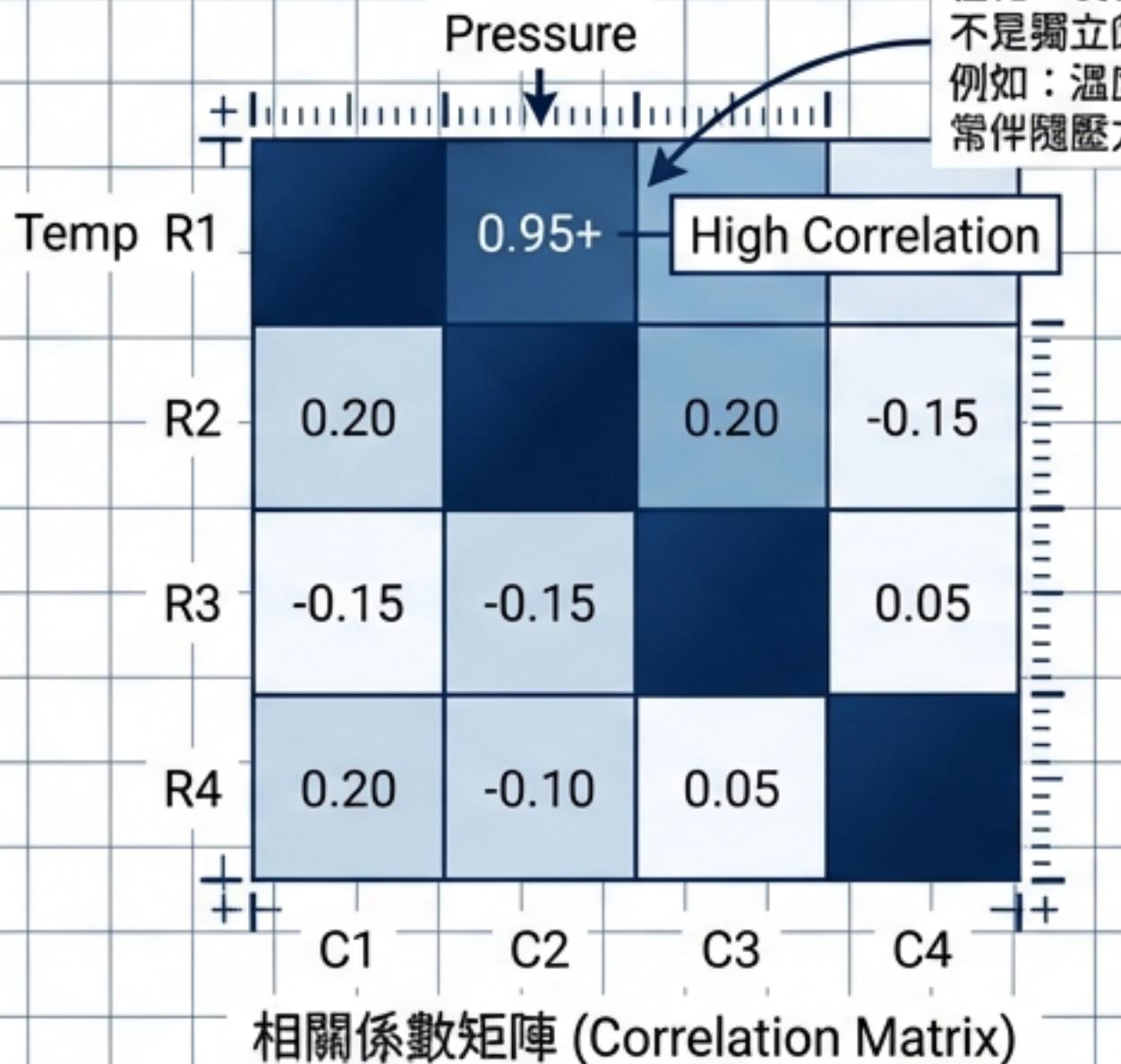


問題：當普通最小平方法失效

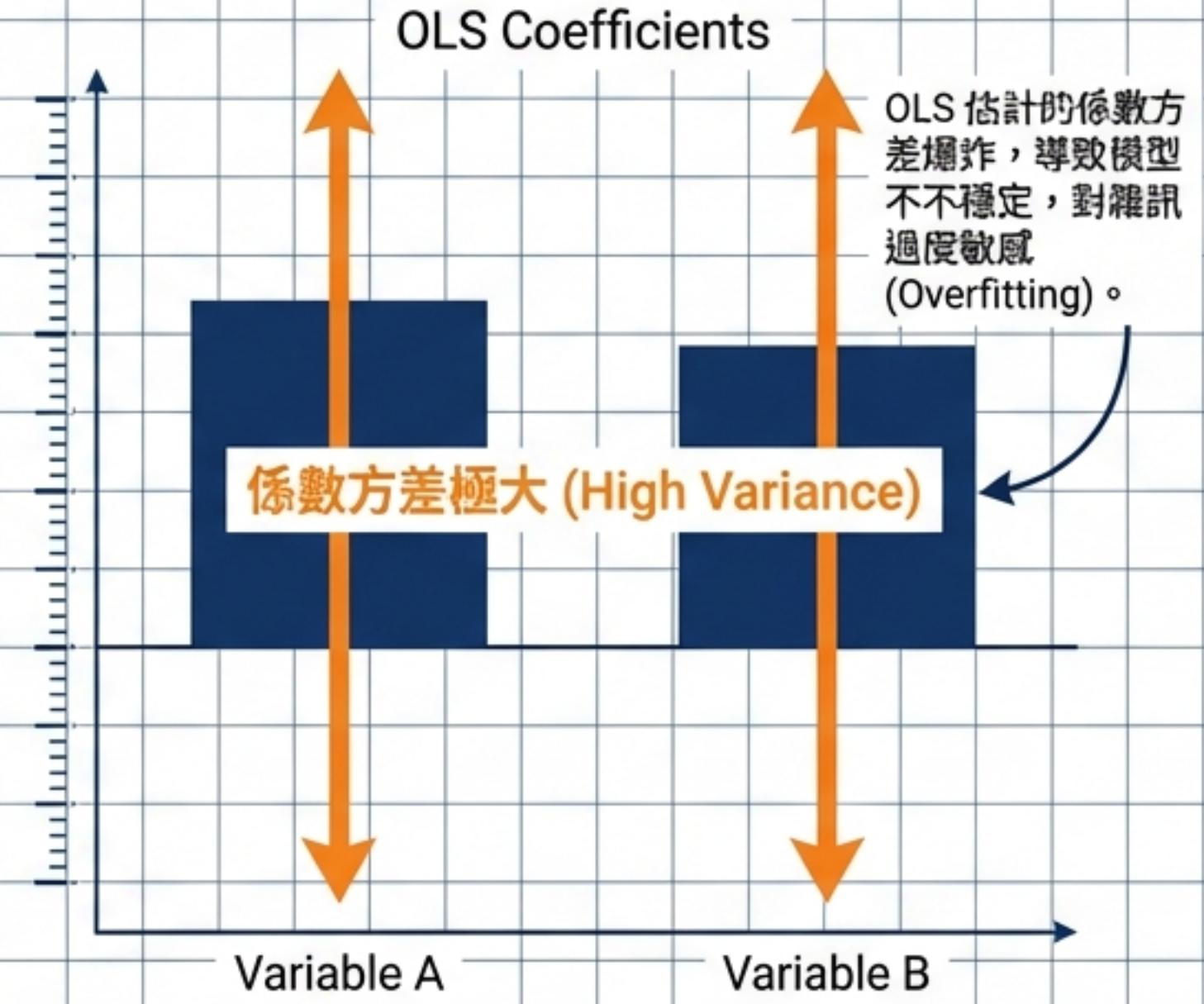
現實世界的挑戰：多重共線性



THE PHENOMENON (現象)



THE CONSEQUENCE (後果)



核心概念：什麼是 Ridge 回歸？

$$\min_{\beta} \left\{ \|y - X\beta\|_2^2 + \alpha \|\beta\|_2^2 \right\}$$

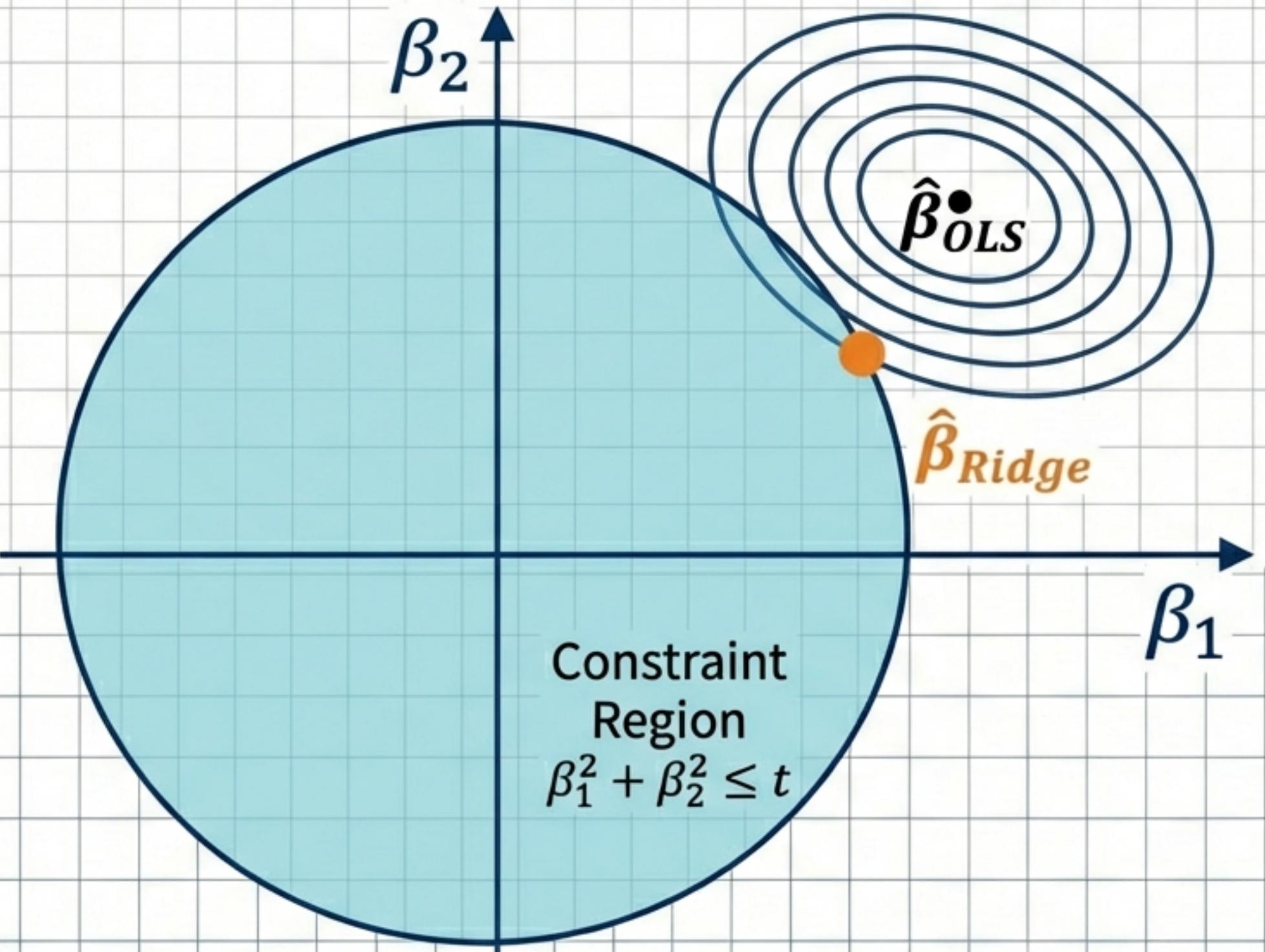
OLS Loss (擬合誤差)
最小化預測誤差

正則化強度 (Strength)
控制「懲罰」的力道

L2 Penalty (懲罰項)
懲罰係數的平方和

OLS + 限制模型複雜度 = Ridge 回歸

幾何意義：將係數限制在圓形區域



Ridge 限制係數向量長度不能超過半徑 t 。

係數被「拉」向原點 (Shrinkage)，變得更小且更穩定。

注意：係數會趨近於 0，但不會完全等於 0 (Non-sparse)。

數學機制：解析解與奇異矩陣修復

OLS Solution (可能失效)：

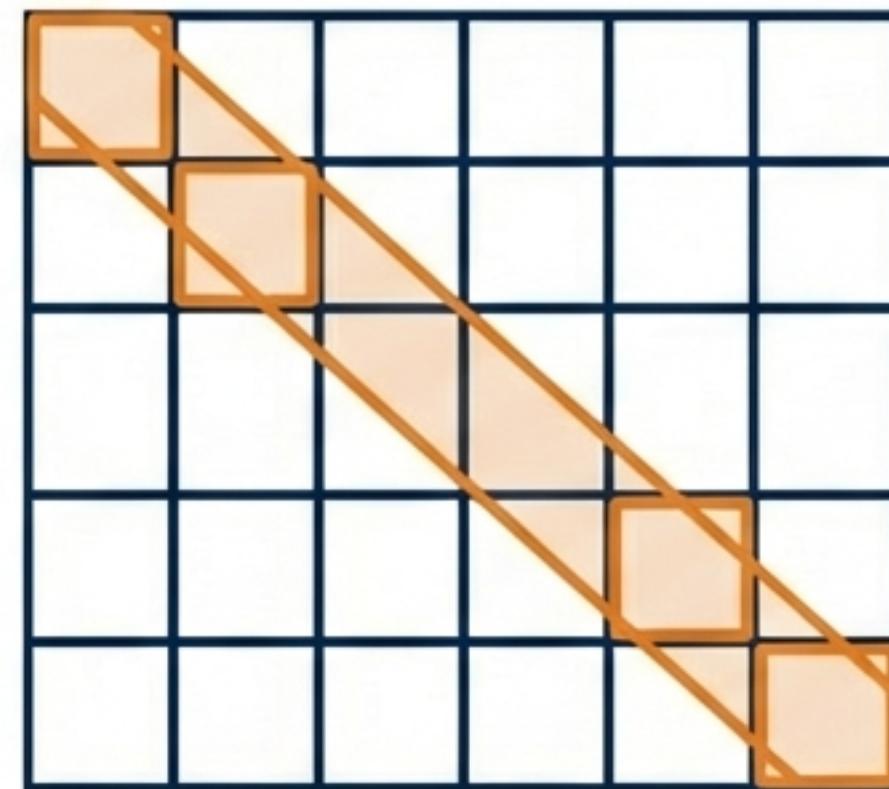
$$\hat{\beta}_{OLS} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y}$$

若 $\mathbf{X}^T \mathbf{X}$ 為奇異矩陣 (Singular)，則無法求逆。

Ridge Solution (數值穩定)：

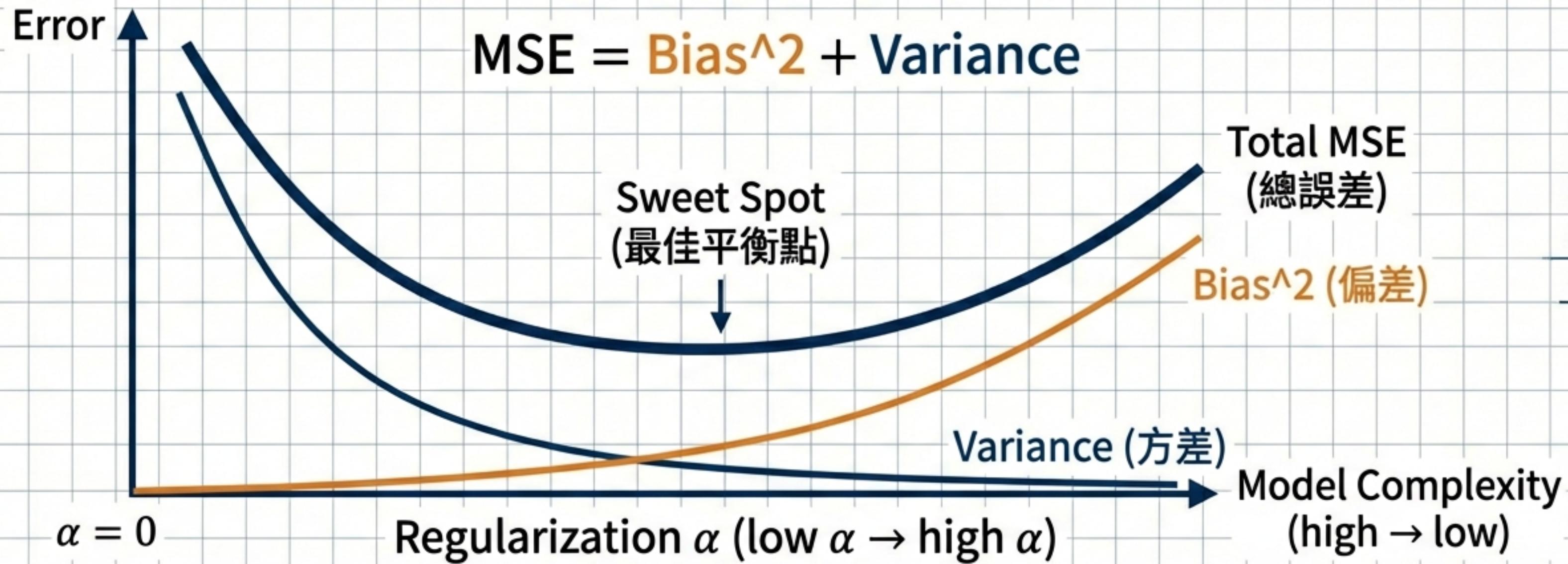
$$\hat{\beta}_{Ridge} = (\mathbf{X}^T \mathbf{X} + \alpha \mathbf{I})^{-1} \mathbf{X}^T \mathbf{y}$$

加上 $\alpha \mathbf{I}$ (對角線增量)



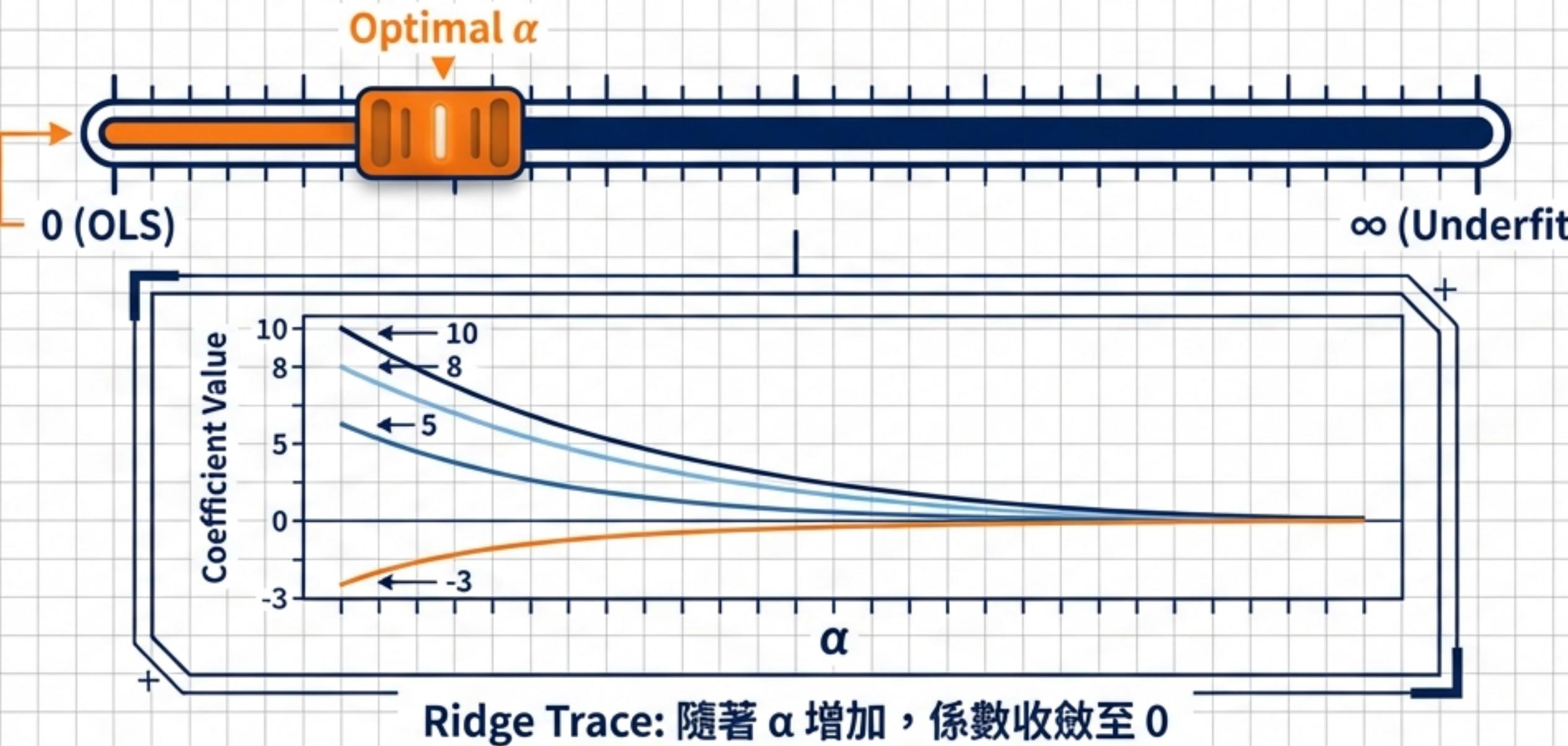
即使原矩陣不可逆，加上微小正數 α 後保證矩陣可逆 (Invertible)。

偏差與方差權衡 (Bias-Variance Tradeoff)



Ridge 引入少量的偏差 (Orange)，換取方差 (Blue) 的大幅降低，從而降低總誤差。

關鍵參數：正則化強度 α



經驗法則 (Rule of Thumb): 化工過程建模通常從 $\alpha \approx 1$ 開始嘗試。

Python 實作 : scikit-learn 工具箱

```
from sklearn.linear_model import Ridge, RidgeCV  
from sklearn.preprocessing import StandardScaler  
  
# 1. 數據標準化 (必須!)  
scaler = StandardScaler()  
X_scaled = scaler.fit_transform(X)  
  
# 2. 建立模型 (自動選擇 alpha)  
# alphas: 測試範圍 (e.g., 0.1, 1.0, 10.0)  
model = RidgeCV(alphas=[0.1, 1.0, 10.0])  
  
# 3. 訓練  
model.fit(X_scaled, y)  
print(f"Best Alpha: {model.alpha_}")
```

Crucial Pre-processing

Cross-Validation
built-in

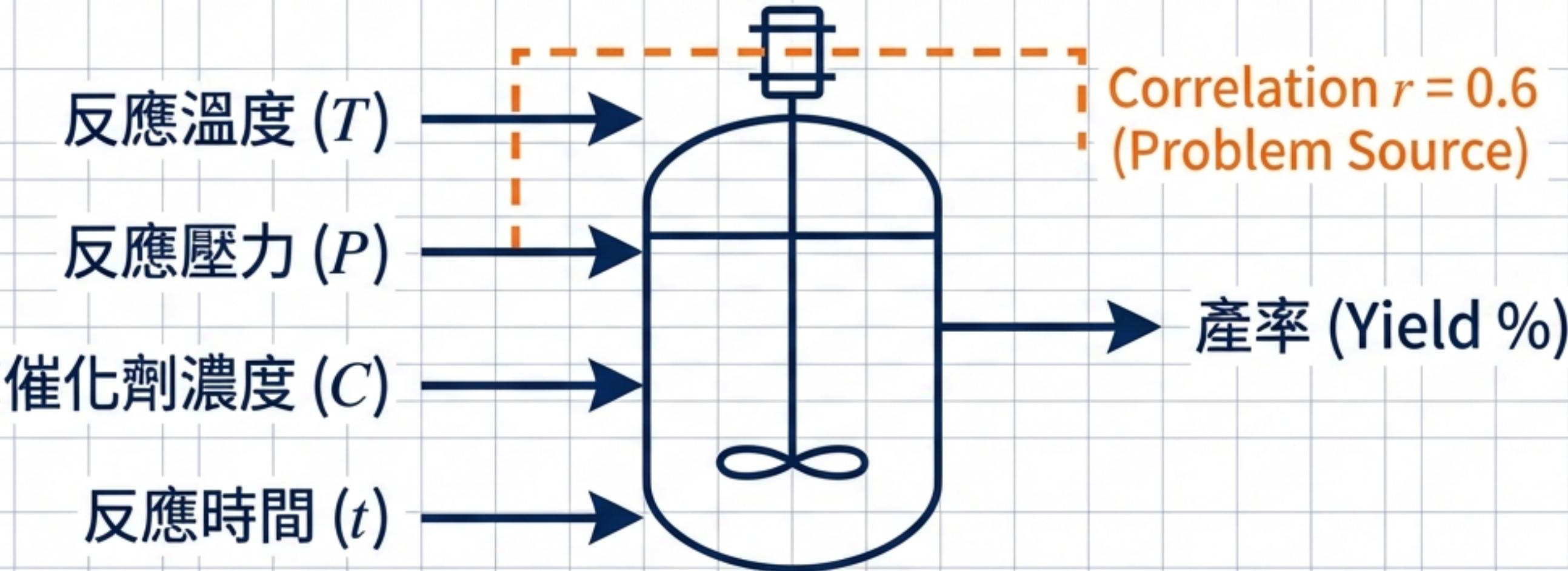
模型對決：何時使用 Ridge？

Feature	OLS	Ridge (L2)	Lasso (L1)
多重共線性 (Multicollinearity)	✗ 不穩定 (Unstable)	✓ 穩健 (Robust)	✓ 穩健 (Robust)
特徵保留 (Feature Retention)	✓ 全部保留	✓ 全部保留 (Shrinkage)	✗ 稀疏解 (Selection)
可解釋性 (Interpretability)	High	Medium	Medium/High

Decision Logic: Use Ridge When...

- 存在多重共線性，且認為**所有特徵都有貢獻** (Keep all features)。
- Use Lasso if you need to eliminate variables.

化工實戰：催化反應產率預測

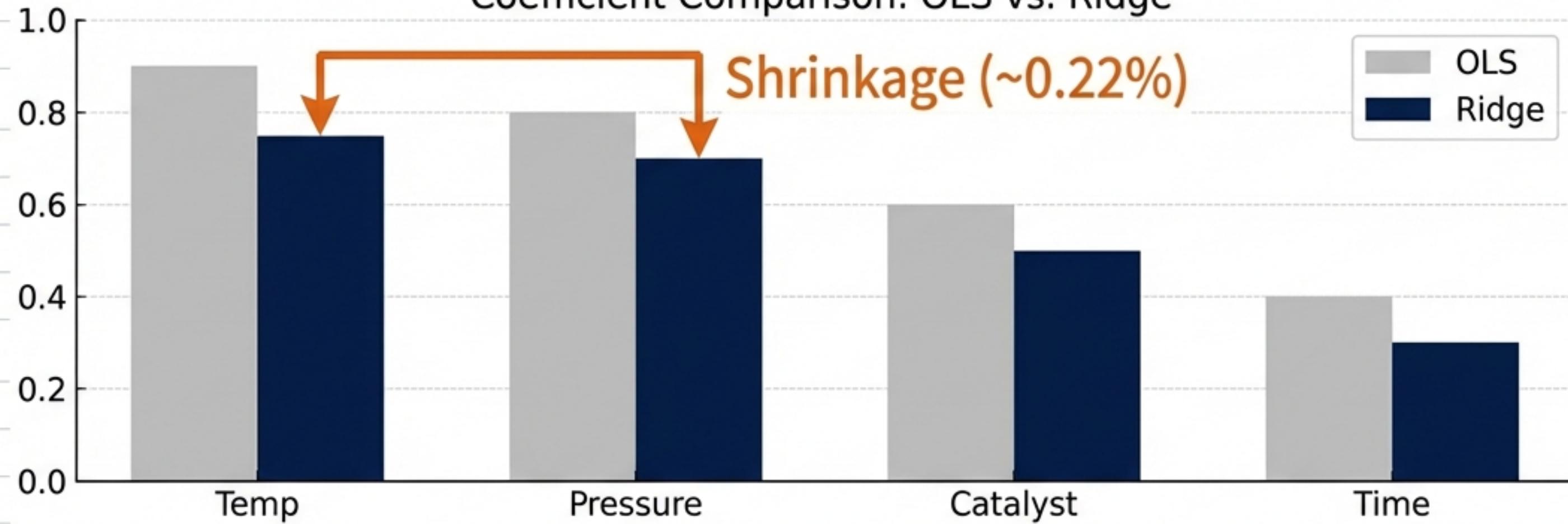


目標：建立穩健的預測模型，避免過度擬合 (Robust Prediction)

結果分析：參數調優與係數比較

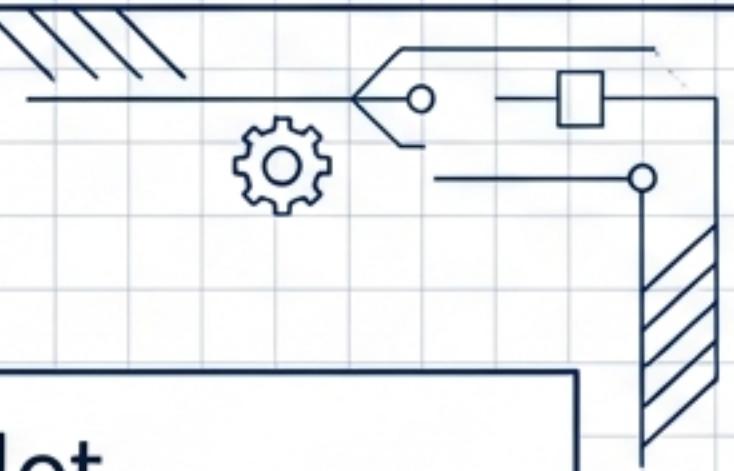
Best Alpha found by CV: 1.1514

Coefficient Comparison: OLS vs. Ridge



即使共線性輕微，Ridge 仍能提供更穩定的估計 (More stable estimation)。

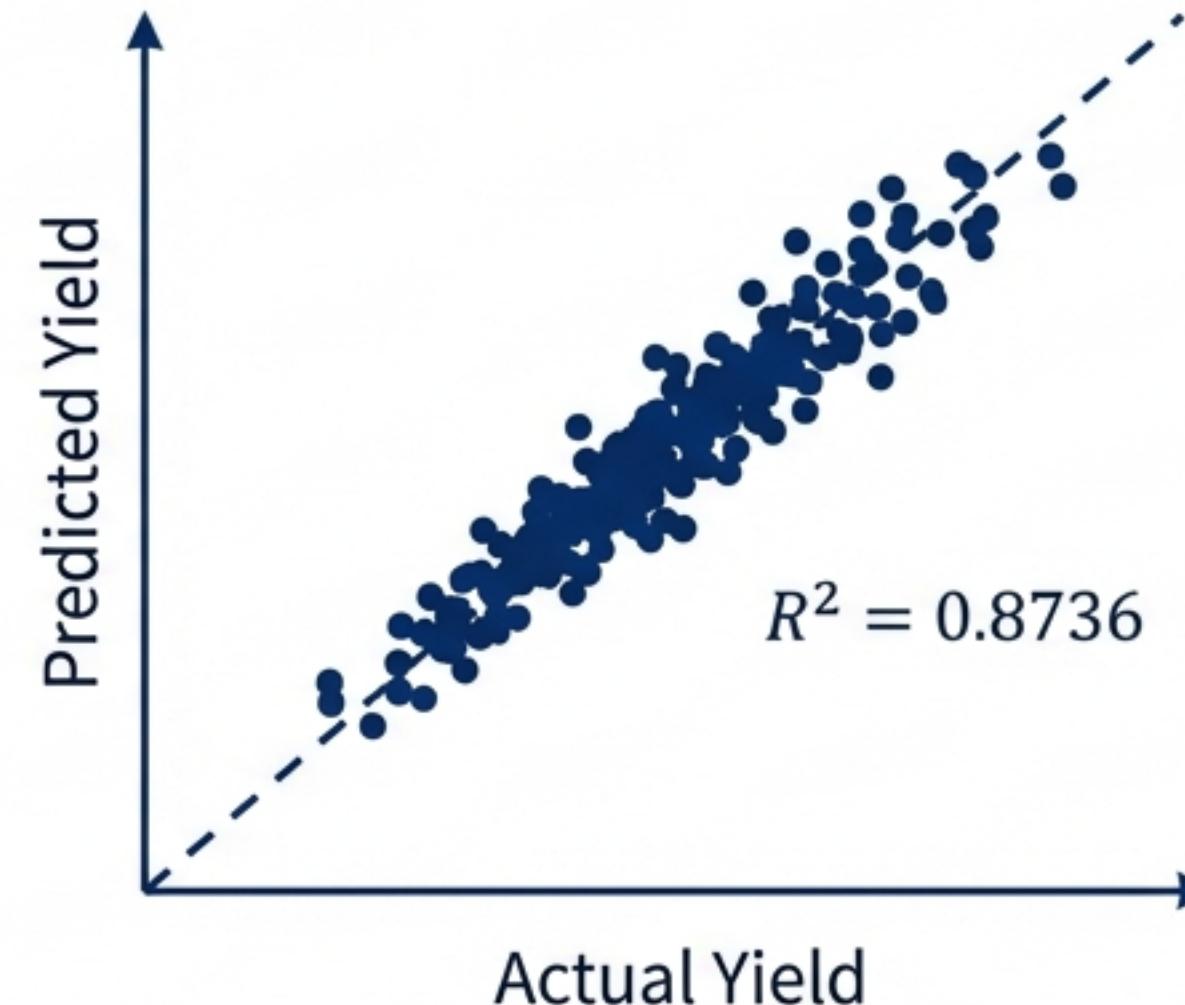
預測與優化：新操作條件模擬



Scenario Table

Scenario	Inputs / Outputs	
Low Intensity	90°C / 2.0 bar	→ Yield 62.38%
Med Intensity	100°C / 3.5 bar	→ Yield 78.45%
High Intensity	110°C / 4.5 bar	→ Yield 95.65%

Parity Plot



產率隨操作強度線性增長，符合物理預期 (Physically consistent).



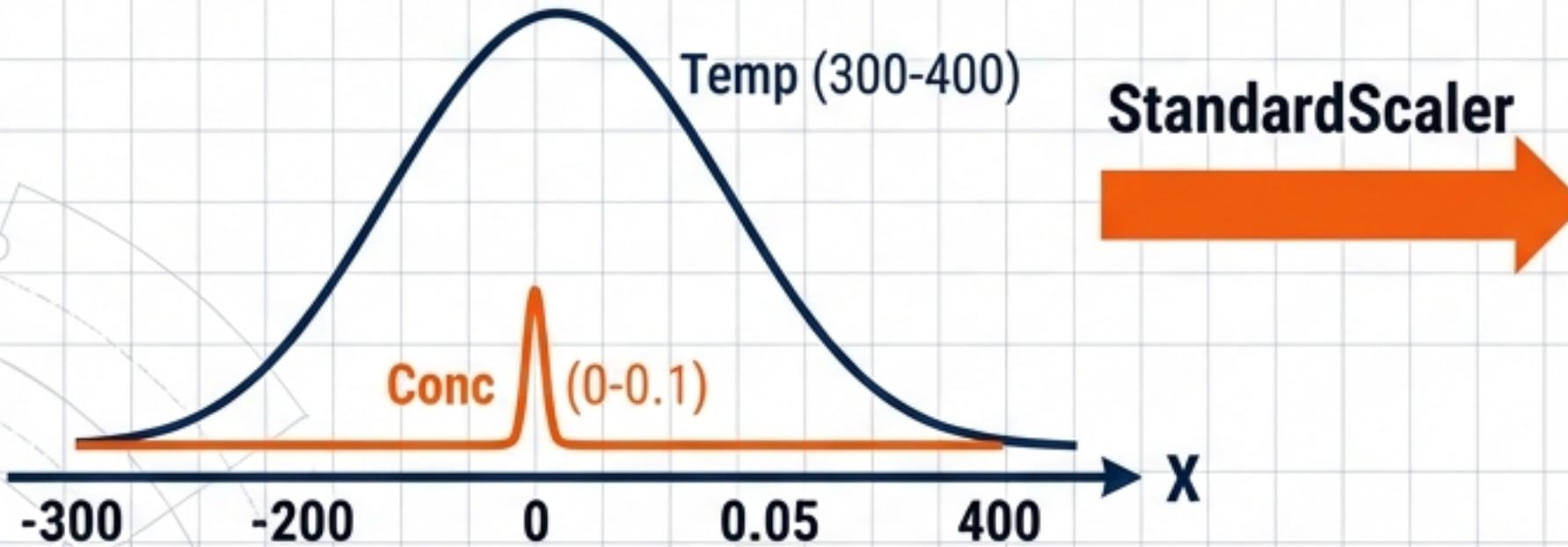
關鍵流程：數據標準化 (Critical Workflow)



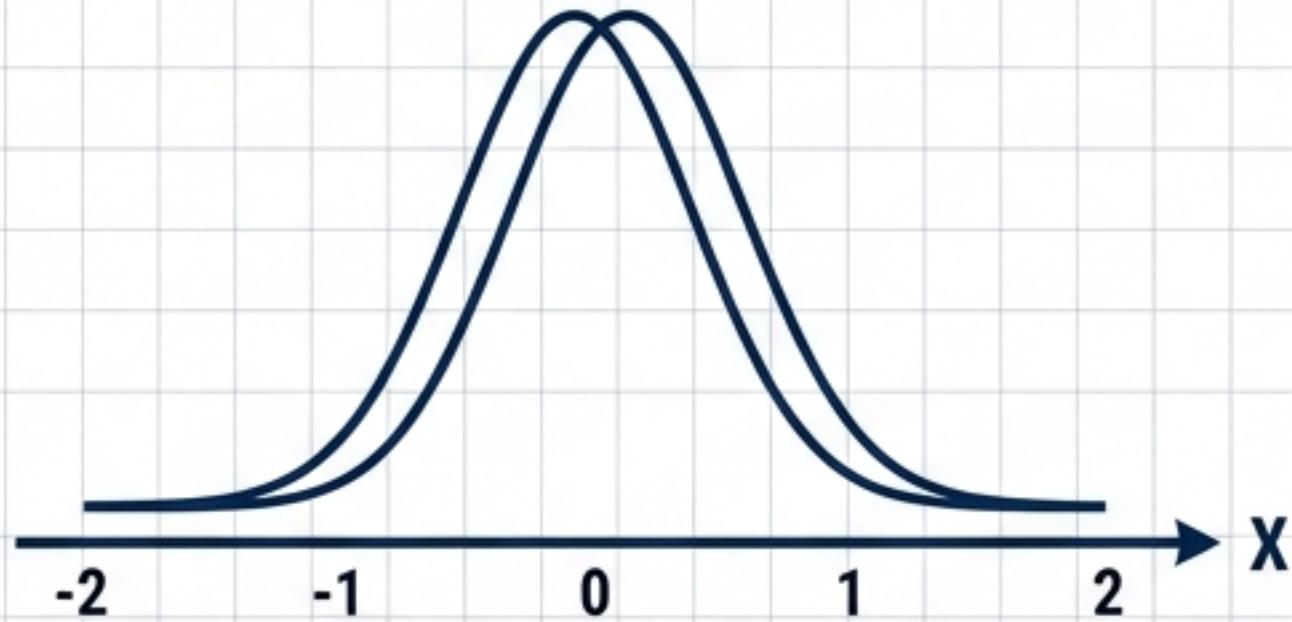
Garbage In, Garbage Out

應用 Ridge 回歸前，**必須進行標準化 (Standardization is MANDATORY)**。

BEFORE: Unstandardized Data



AFTER: Standardized Data
(Unit Variance, Mean 0)



防止大數值特徵被過度懲罰 ($x \rightarrow \frac{x-\mu}{\sigma}$)。

總結與下一步 (Summary & Next Steps)

- ✓ L2 正則化：解決多重共線性與過擬合
- ✓ 保留特徵：係數變小但不為零
- ✓ 調參關鍵：Alpha 決定模型的偏差
(All features kept)
- ✓ 調參關鍵：Alpha 決定模型的偏差與方差

下一步 (Next Step)：
前往 Unit 11 Lasso Regression

學習如何進行「自動特徵選擇」
(Automatic Feature Selection)