

數學建模講義

第 3 講：模型擬合與殘差分析

Kenneth Sok Kin Cheng

修訂日期：2026 年 8 月 20 日

學習目標

完成本課後，讀者應能：

1. 區分模型擬合與插值；
2. 計算並解釋殘差；
3. 以最小二乘法擬合正比模型與直線模型；
4. 利用變量變換，把簡單的指數或幕次關係線性化；
5. 利用圖像、殘差模式、誤差指標及留出數據比較擬合模型；
6. 說明為何數值上擬合良好的模型仍可能不適用。

1 從擬議關係到擬合模型

第 2 講利用物理推理提出以下關係：

$$d_b \propto v^2$$

用以描述制動距離與速度。這類陳述描述了模型的形式，卻尚未確定比例常數；真實觀測亦不大可能完全落在數學曲線上。因此，我們需要系統方法估計參數，並判斷擬議關係是否可信。

引入活動：三條直線、三個故事

[10 分鐘]

圖中有六個數據點，以及不同小組繪畫的三條候選擬合直線。

二人一組討論：

1. 「最佳擬合」可以有甚麼含義？
2. 正誤差應否與負誤差互相抵消？
3. 一個誤差很大的點，是否應比多個小誤差更重要？
4. 選擇直線前還需要甚麼資料？

建模要點：不能只憑外觀或單一數值選擇擬合模型；準則必須符合模型用途。

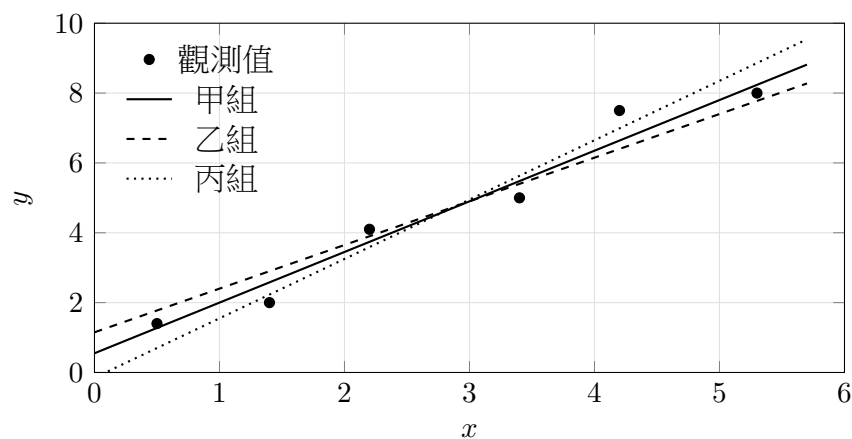


圖 1: 不同擬合準則可以偏好不同模型，因此必須說明擬合目的。

擬合、插值與預測

- 在模型擬合中，先由理論或先前推理提出模型族，再用數據估計參數；曲線毋須通過每個點。
- 在插值中，所建曲線通過指定數據點，通常用於估計點與點之間的數值；這是第 4 講的主題。
- 擬合模型可用於預測，但在觀測範圍外預測屬於外推，必須格外謹慎。

2 殘差與差異來源

假設 x_i 的觀測值為 y_i ，模型預測 $\hat{y}_i = f(x_i)$ 。殘差定義為

$$r_i = y_i - \hat{y}_i.$$

正殘差表示模型低估觀測值；負殘差表示模型高估觀測值。

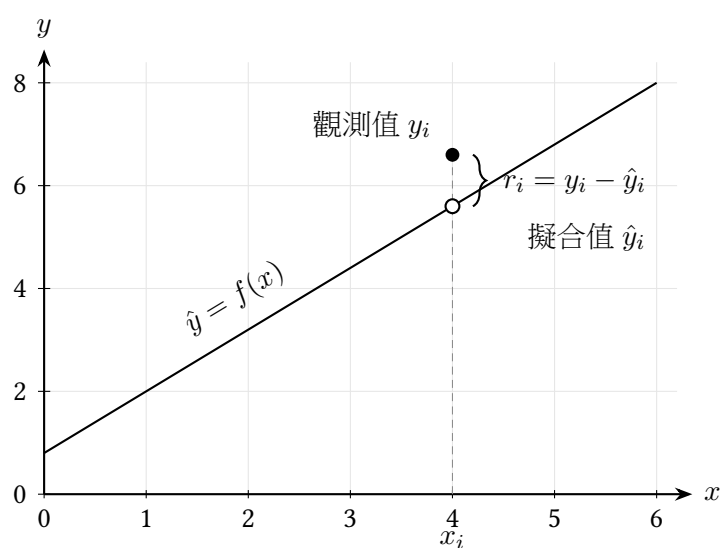


圖 2: 在 x_i 處，殘差是觀測值 y_i 與擬合值 $\hat{y}_i = f(x_i)$ 之間帶正負號的垂直差。

差異不一定源於擬合失敗

數據與模型的差異可以來自多種來源：

1. 測量誤差：精度有限、記錄錯誤或自然變異；
2. 建模誤差：遺漏變量或函數形式不合適；
3. 參數不確定性：現有數據不能精確確定參數；
4. 數值誤差：計算過程中的捨入或近似。

再精確的擬合算法，也不能補救不良模型結構或不可靠數據。

檢查點：正負號與單位

某模型預測煞停距離為 42.3 m，觀測距離則為 45.0 m。

1. 計算殘差。
2. 模型是高估還是低估？
3. 殘差的單位是甚麼？

3 擬合通過原點的正比模型

若理論指出

$$y = kx,$$

便不應自動擬合一般直線 $y = ax + b$ ，因為非零截距會與擬議機制矛盾。對數據 (x_i, y_i) ， k 的最小二乘估計使下式最小：

$$S(k) = \sum_{i=1}^m (y_i - kx_i)^2.$$

微分並令 $S'(k) = 0$ ，得到

$$k = \frac{\sum x_i y_i}{\sum x_i^2}.$$

例題 1：校準制動距離模型

第 2 講提出 $d_b = kv^2$ ，其中 v 是以 m/s 計的速度， d_b 是以 m 計的制動距離。令

$$X = v^2.$$

模型便成為 $d_b = kX$ ，即關於變換後變量 X 的正比關係。

v (m/s)	5	10	15	20	25	30
$X = v^2$	25	100	225	400	625	900
d_b (m)	2.0	7.6	17.9	30.1	49.2	68.5

利用最小二乘公式，

$$k = \frac{\sum X_i d_i}{\sum X_i^2} = \frac{109277.5}{1421875} \approx 0.07685.$$

因此

$$d_b \approx 0.07685v^2.$$

在減速度固定時， $d_b = v^2/(2a)$ ，故擬合係數對應於

$$a \approx \frac{1}{2(0.07685)} \approx 6.51 \text{ m/s}^2.$$

因此，擬合參數具有物理解釋，而非只是一個數值常數。

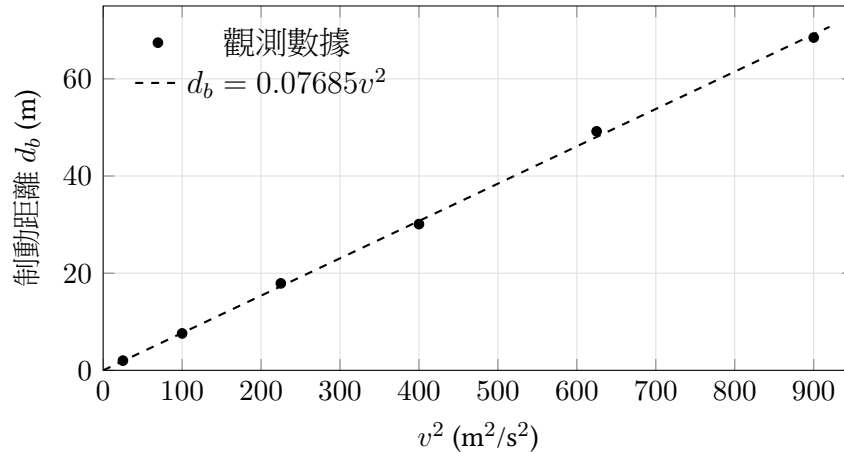


圖 3: 以 v^2 為橫軸繪畫 d_b ，可把擬議的二次關係化為通過原點的直線。

模型檢視：應否強制截距為零？

當 $v = 0$ 時，物理推理要求 $d_b = 0$ ，故強制直線通過原點有充分理據。在其他問題中，非零基線可能是真實的：儀器可以有偏移，送貨費亦可包含固定收費。擬合約束必須源於情境，而非只為方便。

殘差檢查

[8 分鐘]

當 $v = 25 \text{ m/s}$ 時，擬合模型預測

$$\hat{d}_b = 0.07685(25)^2 \approx 48.03 \text{ m}.$$

觀測值為 49.2 m 。

1. 計算殘差。
2. 說明模型低估還是高估。
3. 提出差異的一項物理原因。

4 擬合一般直線

當非零截距合理時，擬合

$$y = ax + b.$$

最小二乘直線使下式最小：

$$S(a, b) = \sum_{i=1}^m (y_i - ax_i - b)^2.$$

解的一個方便形式是

$$a = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (x_i - \bar{x})^2}, \quad b = \bar{y} - a\bar{x}.$$

例題 2：校準力感測器

把以 kg 計的已知質量 x 放在感測器上，產生輸出電壓 y 。

x (kg)	0	1	2	3	4	5
y (V)	0.18	1.72	3.19	4.76	6.25	7.81

此處 $\bar{x} = 2.5$ 、 $\bar{y} = 3.985$ ，

$$\sum (x_i - \bar{x})(y_i - \bar{y}) = 26.655, \quad \sum (x_i - \bar{x})^2 = 17.5.$$

因此

$$a \approx 1.523, \quad b \approx 0.177,$$

校準模型為

$$y \approx 1.523x + 0.177.$$

截距有實際意義：感測器空載時讀數約為 0.18 V，而非零。強制模型通過原點會掩蓋此偏移。

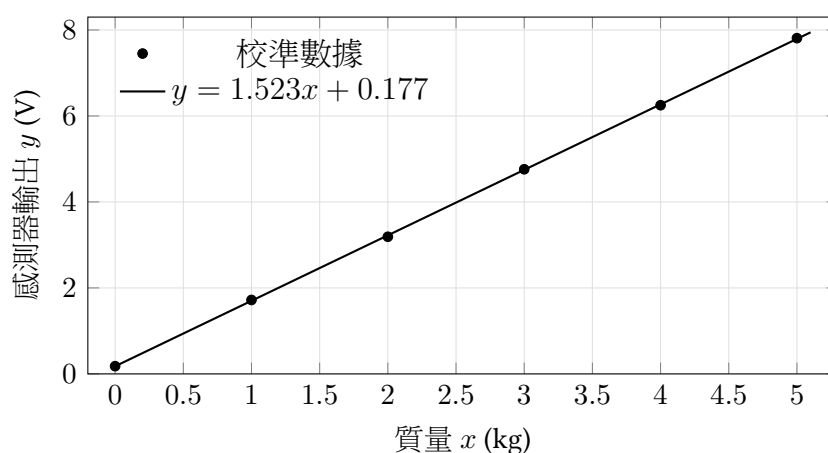


圖 4: 擬合截距表示感測器偏移；參數應按原情境解釋。

為何使用平方？

最小二乘法使下式最小：

$$\text{SSE} = \sum r_i^2,$$

即誤差平方和。平方可防止正、負殘差互相抵消，亦會更強烈懲罰大殘差。這令方法在數學上方便，卻也對離群值敏感。

5 誤差指標與殘差圖

沒有單一誤差指標能回答所有問題。三種常用摘要為

$$\text{MAE} = \frac{1}{m} \sum |r_i|, \quad \text{RMSE} = \sqrt{\frac{1}{m} \sum r_i^2}, \quad \max |r_i|.$$

平均絕對誤差容易解釋；均方根誤差對大誤差懲罰更強；最大誤差則聚焦最壞情況。三者的單位都與響應變量相同。

對制動距離擬合，

$$\text{MAE} \approx 0.54 \text{ m}, \quad \text{RMSE} \approx 0.66 \text{ m}, \quad \max |r_i| \approx 1.17 \text{ m}.$$

這些數值只描述已觀測的校準範圍，不能證明模型在更高速度或不同道路上仍然準確。

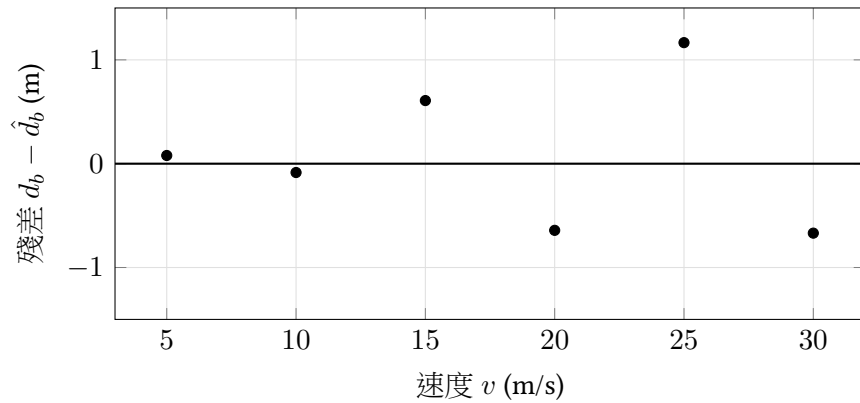


圖 5: 應檢查殘差有否系統性模式，而非只把它們平均為一項統計量。

檢查點：閱讀殘差圖

假設殘差在低 x 時為正、中段為負、高 x 時再次為正。

1. 散佈是否隨機？
2. 彎曲模式對擬合直線有何啟示？
3. 記錄更多小數位能否解決主要問題？

6 把指數模型線性化

某些模型族經變換後會成為線性。對指數衰減模型

$$C = Ae^{-kt},$$

取自然對數可得

$$\ln C = \ln A - kt.$$

因此，以 t 為橫軸、 $\ln C$ 為縱軸的圖像應近似線性，斜率為 $-k$ ，截距為 $\ln A$ 。

例題 3：電池容量衰減

在相同負載下重複測試一枚電池，並把容量 C 記錄為初始值的百分比。

循環區段 t	0	1	2	3	4	5
容量 C (%)	100	82	68	55	46	37

對 $(t, \ln C)$ 擬合所得最小二乘直線約為

$$\ln C = 4.607 - 0.19765t.$$

取指數後得到

$$C \approx 100.2e^{-0.19765t}.$$

擬合模型預測每個循環區段後，容量乘以

$$e^{-0.19765} \approx 0.821$$

即每個區段平均減少當時剩餘容量的約 17.9%。

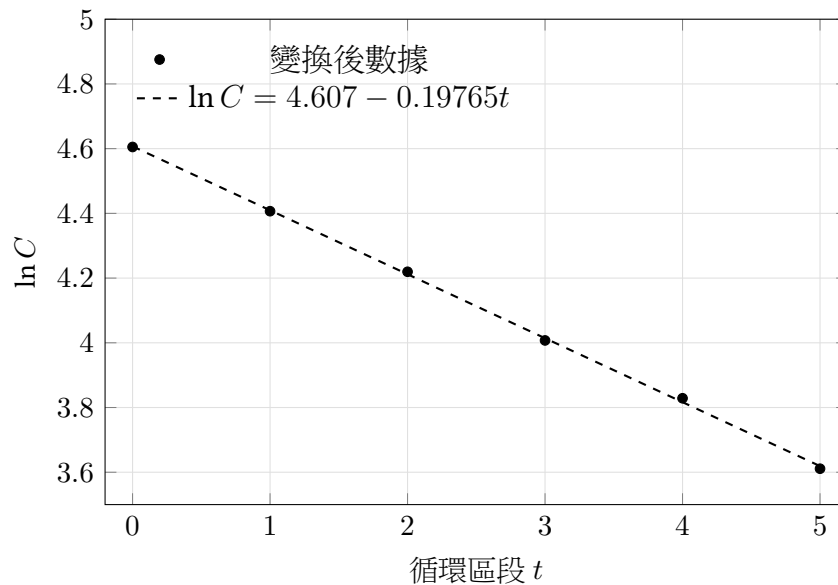


圖 6: 線性化使參數估計更方便，但仍須在原容量尺度上檢查所得模型。

模型檢視：變換會改變擬合準則

在 $\ln C$ 上使用最小二乘法，是把對數之間的平方差而非容量之間的平方差最小化，因而相對重視較小的數值。因此：

1. 只在模型形式與數據容許時才作變換；
2. 把擬合模型轉回原變量；
3. 在原尺度上計算殘差與誤差指標；
4. 使用同一響應變量比較候選模型。

選擇變換

[7分鐘]

說明應繪畫哪些變量，才能檢驗每個擬議模型。

1. $y = Ae^{bx}$
2. $y = Ax^n$
3. $y = A/x$

對每種情況，指出哪個斜率或截距具有參數解釋。

7 選擇模型：擬合、結構與模型效度驗證

不能只因某模型的校準誤差最小便選擇它。我們還要判斷它是否符合機制、殘差是否呈現結構，以及能否預測未用於擬合參數的數據。

例題 4：線性還是指數衰減？

利用 $t = 0$ 至 $t = 5$ 的六個電池觀測值：

- 指數擬合為 $C \approx 100.2e^{-0.19765t}$ ，校準 RMSE 約為 0.39 個百分點；
- 最小二乘直線為 $C \approx 95.81 - 12.46t$ ，校準 RMSE 約為 2.93 個百分點。

擬合時留起 $t = 6$ 的數值 $C = 30.5\%$ ，用作模型效度驗證。

$$\hat{C}_{\text{exp}}(6) \approx 30.61, \quad \hat{C}_{\text{lin}}(6) \approx 21.07.$$

指數模型在留出點的表現好得多，亦不會預測負容量。在此範圍內，數值證據與機制都支持指數衰減。

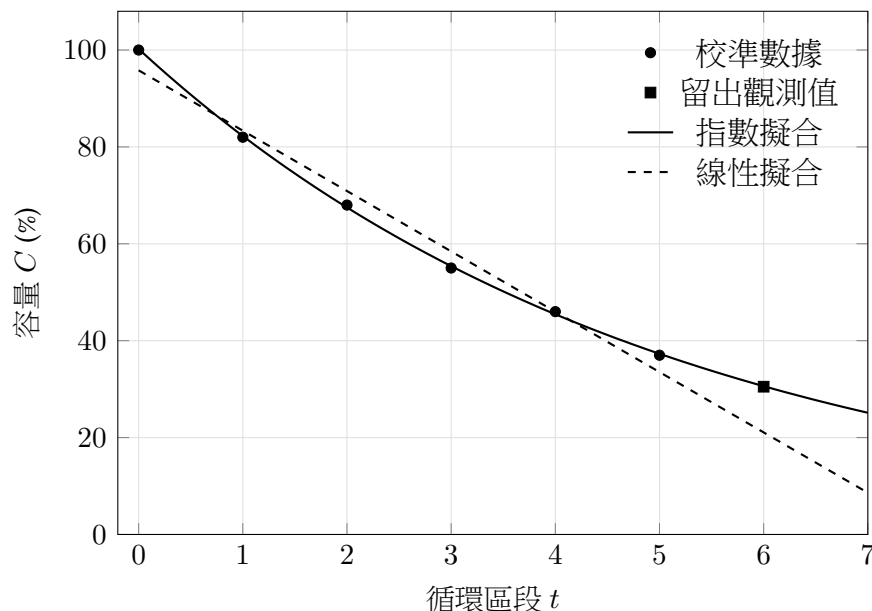


圖 7: 校準誤差、模型效度驗證表現與長期合理性都支持指數模型。

實用模型選擇檢核表

接受擬合模型前，應追問：

1. 機制：函數形式是否符合假設或支配關係？
2. 尺度：是否在原響應變量上評估誤差？
3. 殘差：殘差是否細小且沒有系統性模式？
4. 模型效度驗證：模型能否預測未用於擬合的觀測？
5. 範圍：預定預測位於觀測範圍內還是範圍外？
6. 可解釋性：參數與結論在情境中是否合理？

競賽式表達

[10分鐘]

為電池例子撰寫一段四句的模型選擇說明：

1. 指出所比較的兩個模型；
 2. 報告一項校準指標；
 3. 解釋留出數據的結果；
 4. 陳述結論的一項限制或條件。
- 避免聲稱高質素擬合「證明」模型正確。

選修延伸：現代模型審核：測量不確定性與可識別性

殘差細小不代表每個擬合參數都已充分確定。若 x_i 含測量誤差，卻被視為精確值，普通最小二乘斜率可能產生偏差；此時殘差混合了響應雜訊、輸入不確定性、模型形式誤差與參數不確定性。應記錄儀器分辨率與重複測量的變異，再把合理擾動傳播至擬合結果。

對模型 $f(x; \theta)$ ，只有當現有觀測能區分各參數的效應時，參數才具實際可識別性。一項實用局部診斷是敏感度矩陣

$$J_{ij} = \frac{\partial f(x_i; \theta)}{\partial \theta_j}.$$

若 J 的兩欄近乎成比例，不同參數組合便會產生幾乎相同的擬合曲線。增加小數位或降低訓練誤差都不能消除這種含糊性。應報告參數區間或剖面、在已述測量不確定性下測試擬合，並在相互競爭的參數效應差異最大之處收集觀測。即使個別參數不可識別，預測仍可能穩定，因此須分開審核參數主張與預測主張。

8 課堂總結

應掌握的七個觀念

1. 模型擬合在擬議模型族內估計參數；插值是另一項任務。
2. 殘差等於觀測值減預測值，單位與響應變量相同。
3. 通過原點或容許截距等擬合約束，必須由情境支持。
4. 最小二乘法把殘差平方和最小化，因而對大誤差與離群值敏感。
5. 殘差圖可揭示單一誤差統計量掩蓋的系統性結構。
6. 變換可簡化擬合，卻會改變誤差準則；必須返回原變量檢查。
7. 可信模型結合機制、校準、模型效度驗證、解釋及已陳述限制。

檢查點：課末小結

[2分鐘]

完成以下兩句：

1. 「RMSE 細小仍不足以接受模型，因為_____。」
2. 「只有當_____時，我才會強制擬合直線通過原點。」

9 練習題

核心練習

1. 某模型預測用水需求為 82、91 及 106 L，相應觀測值為 85、88 及 110 L。
 - (a) 計算三個殘差。
 - (b) 求 MAE 與 RMSE。
 - (c) 哪個觀測對 SSE 的貢獻最大？
2. 以下數據描述彈簧的力 F 與伸長量 x ：

x (cm)	1	2	3	4	5
F (N)	1.9	4.2	6.0	8.1	10.3

以最小二乘法擬合 $F = kx$ 。解釋 k 的單位，並計算 $x = 4 \text{ cm}$ 時的殘差。

3. 某溫度感測器給出以下讀數：

x (實際 $^{\circ}\text{C}$)	0	10	20	30	40
y (感測器 $^{\circ}\text{C}$)	1.2	10.8	20.5	30.6	40.1

解釋為何應容許截距。利用試算表或計算機擬合直線，並解釋兩個參數。

4. 說明每個模型所需的變換，並解釋變換後呈直線可支持甚麼：

- (a) $y = Ae^{bx}$;
 (b) $y = Ax^n$;
 (c) $y = A/(B + x)$.

5. 兩個模型有以下以分鐘計的殘差：

$$A : (-1.0, -0.8, 0.2, 0.4, 1.2), \quad B : (-0.2, -0.1, 0.0, 0.1, 3.0).$$

計算並比較兩者的 MAE、RMSE 與最大誤差。MAE、RMSE 及避免嚴重最壞情況各自偏好哪個模型？解釋此處為何必須先說明損失量度，才可談「平均表現」。

6. 某擬合二次曲線的殘差在零附近隨機正負交替，而擬合直線的殘差呈明顯 U 形。解釋這有何啟示。為何單憑「二次曲線有較多參數」不足以以下結論？
7. 某人口模型用 2018–2024 年數據擬合，再用於預測 2045 年。解釋為何訓練 RMSE 細小不能證明 2045 年預測可信，並提出兩項額外檢查。

練習核對與提示

- 殘差為 $(3, -3, 4) \text{ L}$ ； $\text{MAE} = 3.33 \text{ L}$ ， $\text{RMSE} \approx 3.37 \text{ L}$ ，而第三項觀測對 SSE 的貢獻最大。
- $k = 2.04 \text{ N/cm}$ ；在 $x = 4 \text{ cm}$ 時，殘差為 $8.1 - 8.16 = -0.06 \text{ N}$ 。
- 容許截距的擬合為 $y \approx 0.976x + 1.12$ ；截距表示零負載偏移。
- 應分別使用 $(x, \ln y)$ 、 $(\ln x, \ln y)$ 及 $(x, 1/y)$ ；直線關係支持相應模型族，但不能單獨完成模型效度驗證。
- 對 A ， $\text{MAE} = 0.72$ 、 $\text{RMSE} \approx 0.81$ 、最大誤差 = 1.2；對 B ，三者分別為 0.68、1.35 及 3.0。故 MAE 偏好 B ，RMSE 與最壞情況控制則偏好 A 。
- 直線的 U 形殘差顯示遺漏曲率；接受二次曲線前，須比較留出資料表現及殘差結構。
- 可信的長期預測檢查包括時間留出測試，並至少加入機制審視、情境敏感度或替代模型族比較其中一項。

建模挑戰

選擇一組小組可收集或取得的數據：冷卻溫度、輪候時間、植物生長、裝置電池衰減、交通行程時間，或其他獲准情境。

撰寫一頁擬合簡報，內容包括：

- 問題、變量、單位與擬議模型族；
- 容許或排除截距的理由；
- 擬合參數值，以及一幅以原變量繪畫的圖；

4. 殘差圖及至少兩項誤差指標；
 5. 一項模型效度驗證策略及一項限制；
 6. 為非專業決策者撰寫的簡短結論。
- 目標並非透過增加項數取得盡可能低的誤差，而是建立有理據的模型。

A 選讀推導：最小二乘直線

本附錄不屬於兩小時核心課堂內容，可供熟習偏微分的學生選讀。

對 $y = ax + b$ ，把下式最小化：

$$S(a, b) = \sum_{i=1}^m (y_i - ax_i - b)^2.$$

令兩個偏導數皆等於零，得到

$$\begin{aligned}\frac{\partial S}{\partial a} &= -2 \sum x_i (y_i - ax_i - b) = 0, \\ \frac{\partial S}{\partial b} &= -2 \sum (y_i - ax_i - b) = 0.\end{aligned}$$

因此

$$\begin{aligned}a \sum x_i^2 + b \sum x_i &= \sum x_i y_i, \\ a \sum x_i + mb &= \sum y_i.\end{aligned}$$

這些是正規方程。求解它們等同於使用課堂中的中心化平均公式。

B 選讀延伸：最小二乘法的替代準則

最小二乘法並非唯一有理據的擬合準則。

- L^1 準則把 $\sum |r_i|$ 最小化，對單一大型離群值較不敏感。
- L^∞ 或極小化極大準則把 $\max |r_i|$ 最小化，適用於最壞誤差最重要的情況。

合適準則取決於決策。醫療安全模型、預測模型與工程校準問題，未必以相同方式衡量誤差。

模型審核卡

在信任或傳達模型之前先作審核

目的	模型要回答甚麼問題或支援甚麼決策？	資料	哪些輸入屬觀測、推導或模擬資料？
假設	哪些簡化可能改變答案？	單位	量綱與尺度是否一致？
運算 / 程式核證	方程、程式、限制與極限情況是否已核對？	模型效度驗證	是否用獨立證據檢驗預定用途下的表現？
不確定性	測量、參數、求解器及模型結構誤差如何影響結論？	倫理	誰受益、誰承擔風險，公平準則是甚麼？
可重現性	是否記錄資料來源、版本、種子與轉換？	溝通	主張是否有條件、可追溯且與決策相關？

延伸閱讀

- 《工程統計手冊》（NIST / SEMATECH）
- 《基礎統計學》第二版（OpenStax）

來源、資料脈絡與授權

本講義依循 SIAM / COMAP GAIMME 報告對反覆建模與評核的觀點。競賽相關說明於 2026 年 8 月 20 日依據官方 COMAP HiMCM 規則核對；提交前仍須查閱當時最新規則。除非另有明確來源，課堂資料與數值情境均屬合成或模擬教學例子，不構成任何真實機構的實證主張。第三方參考資料沿用各自條款。Kenneth Sok Kin Cheng 的原創教材採用 CC BY-NC-SA 4.0 授權。

學生作答指引與檢核準則

封閉式計算題請列出定義、代入、單位及至少一項獨立檢查；本課的完整例題可作方法提示，但須自行完成數值。開放建模題按五項檢核：目的與輸出是否明確、資料與來源是否交代、假設與適用範圍是否合理、運算／程式核證與模型效度驗證是否可追蹤，以及不確定性、限制與建議是否互相一致。每項須附具體證據，而非只寫「已完成」。