

Unit 5 — Curve Fitting

Principle of least squares · Working procedure · Straight line · Parabola · Exponential curve
($y = ae^{bx}$, $y = ab^x$) · Power curve ($y = ax^b$)

Exam slot: Part-B Q10 (10 M) — OR with Unit 6

Part-A: 1i (2 M)

Written in ASD-STE100 style

PYQ TREND (2020 – 2026): WHAT THE EXAM ASKS FROM THIS UNIT

- "State the principle of least squares" is in Part A of **every** paper (2 M). Learn 3.1 word for word.
- The "cars inspected" data (x : 2, 7, 9, 1, 5, 12; y : 13, 21, 23, 14, 15, 21) was asked 3 times in a row: parabola (Apr-2026), ae^{bx} (May-2025), ax^b (Nov-2024). All three are solved below.
- Quadratic on $x = 1 \dots 6$, $y = 6, 4, 3, 5, 4, 2$: Mar/Apr-2024, Oct-2023. Write/derive normal equations: 2–5 M.

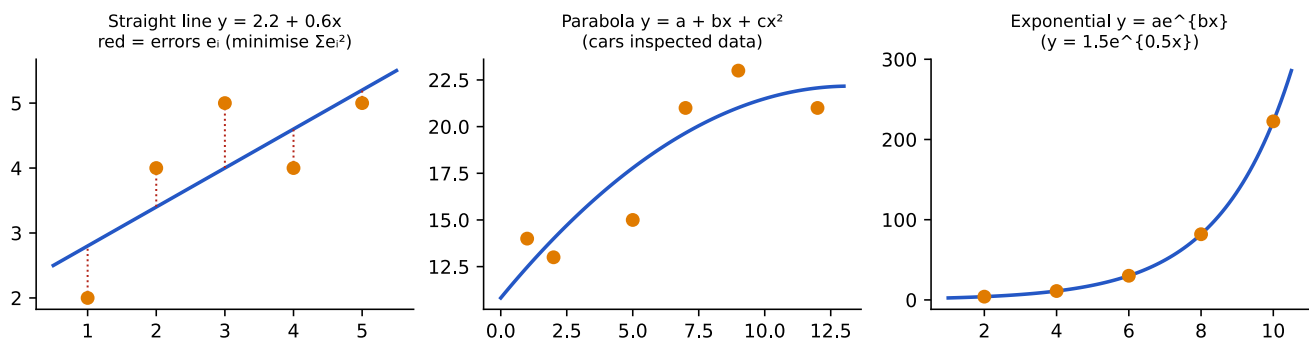


Fig 5.1 — The three curve types in this unit, fitted to PYQ data.

5.1 Principle of Least Squares

STATEMENT EVERY PAPER · 2M

Let (x_i, y_i) , $i = 1, 2, \dots, n$, be the given data points. Let $y = f(x)$ be the curve that we fit to the data. At $x = x_i$, the **observed** value is y_i and the value on the curve (**expected** value) is $f(x_i)$.

The **error (residual)** at x_i is $e_i = y_i - f(x_i)$.

Let $S = e_1^2 + e_2^2 + \dots + e_n^2 = \sum_{i=1}^n [y_i - f(x_i)]^2$.

Principle of least squares: the curve of best fit is the curve for which S , the **sum of the squares of the errors**, is minimum.

ADVANTAGES OF THE LEAST SQUARES METHOD APR/MAY-2026 · 2M

1. It gives one **unique** best-fit curve. Two persons get the same answer (no guesswork as in a freehand curve).
2. Positive and negative errors do not cancel, because we square them.
3. The method is simple: we only solve linear normal equations.
4. It applies to many curves: line, parabola, exponential and power curves.

Working procedure for fitting a curve Oct-2025 · 5M Nov-2020 · 7M

- Step 1** Write the form of the curve, e.g. $y = a + bx$. Make it linear by taking logs if necessary.
- Step 2** Write the normal equations for the unknown constants.
- Step 3** Make a table with columns x , y , x^2 , xy , ... as the normal equations need. Find the column totals.
- Step 4** Put the totals in the normal equations.
- Step 5** Solve the equations for the constants $a, b, \dots$
- Step 6** Put the constants in the curve. This is the curve of best fit.

5.2 Fitting a Straight Line $y = a + bx$

Derivation of normal equations Mar/Apr-2022 · 8M May-2025 · 2M Nov/Dec-2022 · 2M

Let $y = a + bx \dots (1)$ be the line. Then $S = \sum_{i=1}^n (y_i - a - bx_i)^2$.

For S minimum: $\frac{\partial S}{\partial a} = 0$ and $\frac{\partial S}{\partial b} = 0$.

$$\frac{\partial S}{\partial a} = 2 \sum (y_i - a - bx_i)(-1) = 0 \Rightarrow \sum y_i - na - b \sum x_i = 0 \Rightarrow \boxed{na + b \sum x_i = \sum y_i} \dots (2)$$

$$\frac{\partial S}{\partial b} = 2 \sum (y_i - a - bx_i)(-x_i) = 0 \Rightarrow \sum x_i y_i - a \sum x_i - b \sum x_i^2 = 0 \Rightarrow \boxed{a \sum x_i + b \sum x_i^2 = \sum x_i y_i} \dots (3)$$

Equations (2) and (3) are the **normal equations** of the straight line. Solve them for a, b and put the values in (1).

PYQ. Fit a straight line to the data: $x: 1, 2, 3, 4, 5$; $y: 2, 4, 5, 4, 5$. Oct-2025 · 5M Class notes

x	y	x^2	xy
1	2	1	2
2	4	4	8
3	5	9	15
4	4	16	16
5	5	25	25
15	20	55	66

Normal equations ($n = 5$): $5a + 15b = 20 \dots (i)$; $15a + 55b = 66 \dots (ii)$.

(i) $\times 3$: $15a + 45b = 60$. Subtract from (ii): $10b = 6 \Rightarrow b = 0.6$. Then $a = \frac{20-9}{5} = 2.2$. **$y = 2.2 + 0.6x$**

Example 1. Find the least squares straight line: $x: 1, 2, 3, 4, 5, 6$; $y: 6, 4, 3, 5, 4, 2$. Class notes

$\sum x = 21$, $\sum y = 24$, $\sum x^2 = 91$, $\sum xy = 6 + 8 + 9 + 20 + 20 + 12 = 75$, $n = 6$.

$6a + 21b = 24 \dots (i)$; $21a + 91b = 75 \dots (ii)$. From (i), $a = 4 - 3.5b$. Put in (ii): $84 - 73.5b + 91b = 75 \Rightarrow$

$17.5b = -9 \Rightarrow b = -0.514$. $a = 4 + 1.8 = 5.8$. **$y = 5.8 - 0.514x$**

CORRECTION TO CLASS NOTES

The class notes write $b = +\frac{18}{35}$. The correct value is $b = -\frac{18}{35} = -0.514$ (y decreases as x increases).

PYQ. By the method of least squares, find the straight line that best fits: x: 1, 2, 3, 4, 5; y: 14, 27, 40, 55, 68. **Nov-2020 · 8M**

$\sum x = 15, \sum y = 204, \sum x^2 = 55, \sum xy = 14 + 54 + 120 + 220 + 340 = 748.$
 $5a + 15b = 204; 15a + 55b = 748.$ Solve: $10b = 748 - 612 = 136 \Rightarrow b = 13.6; a = \frac{204 - 204}{5} = 0.$

$$y = 13.6x$$

PYQ. Fit a straight line: x: 1, 2, 3, 4, 5, 6; y: 3.7, 6.2, 8.7, 11.2, 13.7, 16.2. **Jul-2021 · 7M**

$\sum y = 59.7, \sum xy = 252.7.$ $6a + 21b = 59.7; 21a + 91b = 252.7.$ Solve: $b = 2.5, a = 1.2.$ **$y = 1.2 + 2.5x$**
 (The points lie exactly on this line.)

5.3 Fitting a Parabola $y = a + bx + cx^2$

NORMAL EQUATIONS**APR/MAY-2026 · 5M****OCT-2023 · 2M****APR/MAY-2023 · 2M**

$S = \sum (y_i - a - bx_i - cx_i^2)^2.$ Put $\frac{\partial S}{\partial a} = \frac{\partial S}{\partial b} = \frac{\partial S}{\partial c} = 0:$

$$\begin{aligned} na + b \sum x + c \sum x^2 &= \sum y \\ a \sum x + b \sum x^2 + c \sum x^3 &= \sum xy \\ a \sum x^2 + b \sum x^3 + c \sum x^4 &= \sum x^2 y \end{aligned}$$

For $y = ax^2 + bx + c$ (Apr/May-2023 form), the same equations hold with the letters renamed: c is the constant term, b multiplies x , a multiplies x^2 .

Memory rule: multiply $y = a + bx + cx^2$ by 1, by x , by x^2 , then add over all points.

CAUTION — THE MOST COMMON MISTAKE

The right side of the **first** equation is $\sum y$, not $\sum xy$. Check each right side against its multiplier (1, x , x^2).

PYQ. The table shows how many weeks six persons worked at an automobile inspection station and the number of cars each inspected between noon and 2 PM. Fit $y = a + bx + cx^2$. x : 2, 7, 9, 1, 5, 12; y : 13, 21, 23, 14, 15, 21. Apr/May-2026 · 5M Class notes

x	y	x^2	x^3	x^4	xy	x^2y
2	13	4	8	16	26	52
7	21	49	343	2401	147	1029
9	23	81	729	6561	207	1863
1	14	1	1	1	14	14
5	15	25	125	625	75	375
12	21	144	1728	20736	252	3024
36	107	304	2934	30340	721	6357

Normal equations ($n = 6$):

$$6a + 36b + 304c = 107 \dots (i)$$

$$36a + 304b + 2934c = 721 \dots (ii)$$

$$304a + 2934b + 30340c = 6357 \dots (iii)$$

Solve (eliminate a , then b): $a = 10.822$, $b = 1.717$, $c = -0.0650$. $y = 10.822 + 1.717x - 0.065x^2$

CORRECTION TO CLASS NOTES

The class notes write equation (i) as $6a + 36b + 304c = 721$. The right side must be $\sum y = 107$. That error gives the wrong curve $y = -15.584 + 21.704x + 0.1089x^2$. Check: at $x = 12$, it gives $y \approx 260$, but the data value is 21.

PYQ. Find a least squares quadratic curve: x : 1, 2, 3, 4, 5, 6; y : 6, 4, 3, 5, 4, 2. Mar/Apr-2024 · 8M

Oct-2023 · 8M Class notes

Sums: $\sum x = 21$, $\sum x^2 = 91$, $\sum x^3 = 441$, $\sum x^4 = 2275$, $\sum y = 24$, $\sum xy = 75$, $\sum x^2y = 301$.

$6a + 21b + 91c = 24 \dots (i)$; $21a + 91b + 441c = 75 \dots (ii)$; $91a + 441b + 2275c = 301 \dots (iii)$.

Solve: $a = 5.8$, $b = -0.514$, $c = 0$. $y = 5.8 - 0.514x (+ 0 \cdot x^2)$

For this data the best parabola has $c = 0$, so it is the same as the straight line of Example 1. This is a correct and complete answer; state it in one line.

PYQ. Fit a parabola $y = a + bx + cx^2$: x : 1, 2, 3, 4, 5, 6, 7; y : 2.3, 5.2, 9.7, 16.5, 29.4, 35.5, 54.4.

Nov/Dec-2022 · 10M

$n = 7$, $\sum x = 28$, $\sum x^2 = 140$, $\sum x^3 = 784$, $\sum x^4 = 4676$, $\sum y = 153$, $\sum xy = 848.6$, $\sum x^2y = 5053$.

$7a + 28b + 140c = 153$; $28a + 140b + 784c = 848.6$; $140a + 784b + 4676c = 5053$.

Solve: $a = 2.371$, $b = -1.093$, $c = 1.193$. $y = 2.371 - 1.093x + 1.193x^2$

SHORT-CUT

When x values are equally spaced, put $X = x - 4$ (middle value). Then $\sum X = \sum X^3 = 0$ and the equations become very easy.

PYQ. Fit a least squares parabola: X: 0, 0.2, 0.4, 0.7, 0.9, 1; Y: 1.016, 0.768, 0.648, 0.401, 0.272, 0.193.

Oct/Nov-2021 · 15M

$$\sum x = 3.2, \sum x^2 = 2.5, \sum x^3 = 2.144, \sum x^4 = 1.9234, \sum y = 3.298, \sum xy = 1.1313, \sum x^2y = 0.7442.$$

$$6a + 3.2b + 2.5c = 3.298; 3.2a + 2.5b + 2.144c = 1.1313; 2.5a + 2.144b + 1.9234c = 0.7442.$$

$$y = 0.999 - 1.009x + 0.213x^2$$

5.4 Exponential Curve $y = ae^{bx}$

METHOD (CHANGE TO A STRAIGHT LINE)

Take log (base 10) on both sides: $\log y = \log a + bx \log e$.

Put $Y = \log_{10} y$, $A = \log_{10} a$, $B = b \log_{10} e = 0.4343 b$. Then $Y = A + Bx$ (a straight line).

Normal equations: $nA + B \sum x = \sum Y$, $A \sum x + B \sum x^2 = \sum xY$.

Then $a = 10^A$ and $b = \frac{B}{0.4343}$.

If you use natural log ($Y = \ln y$), then $B = b$ directly and $a = e^A$. Both ways give the same curve.

CAUTION — CORRECTION TO CLASS NOTES

With $\log_{10}$, the slope is $B = b \log_{10} e$, not b . The class notes put $B = 0.02248$ directly in e^{bx} . The correct exponent is $b = \frac{0.02248}{0.4343} = 0.0518$. (The class answer $y = 12.75(1.0529)^x$ for the form ab^x is correct, and it is the same curve.)

PYQ. For the "cars inspected" data (x: 2, 7, 9, 1, 5, 12; y: 13, 21, 23, 14, 15, 21), fit $y = ae^{bx}$.

May-2025 · 8M

Class notes

x	y	$Y = \log_{10} y$	x^2	xY
2	13	1.1139	4	2.2278
7	21	1.3222	49	9.2554
9	23	1.3617	81	12.2553
1	14	1.1461	1	1.1461
5	15	1.1761	25	5.8805
12	21	1.3222	144	15.8664
36		7.4422	304	46.6315

$$6A + 36B = 7.4422 \dots \text{(i)}; 36A + 304B = 46.6315 \dots \text{(ii)}. \text{(i)} \times 6: 36A + 216B = 44.6532. \text{ Subtract: } 88B = 1.9783 \Rightarrow B = 0.02248. A = \frac{7.4422 - 0.8093}{6} = 1.1055.$$

$$a = 10^{1.1055} = 12.75, b = \frac{0.02248}{0.4343} = 0.0518. \quad y = 12.75 e^{0.0518x}$$

PYQ. Determine a and b by least squares such that $y = ae^{bx}$: x: 2, 4, 6, 8, 10; y: 4.077, 11.084, 30.128, 81.897, 222.62. Jul-2021 · 8M

$$Y = \log_{10} y: 0.6103, 1.0447, 1.4790, 1.9133, 2.3476. \sum Y = 7.3948, \sum xY = 53.055, \sum x = 30, \sum x^2 = 220.$$

$5A + 30B = 7.3948$; $30A + 220B = 53.055$. Solve: $B = 0.21715$, $A = 0.17606$.

$$a = 10^{0.17606} = 1.50, \quad b = \frac{0.21715}{0.4343} = 0.50. \quad \mathbf{y = 1.5 e^{0.5x}}$$

PYQ. Fit $Y = Ae^{BX}$: X : 2, 4, 6, 8, 10, 12; Y : 1.8, 1.5, 1.4, 1.1, 1.1, 0.9. Oct/Nov-2021 · 7M

$\log_{10} Y$: 0.2553, 0.1761, 0.1461, 0.0414, 0.0414, -0.0458. Sum = 0.6145; $\sum X \log Y = 2.2877$; $\sum X = 42$; $\sum X^2 = 364$.

$6a_0 + 42b_0 = 0.6145$; $42a_0 + 364b_0 = 2.2877 \rightarrow b_0 = -0.02877$, $a_0 = 0.3038$. $A = 10^{0.3038} = 2.013$, $B =$

$$\frac{-0.02877}{0.4343} = -0.0662. \quad \mathbf{Y = 2.013 e^{-0.0662X}}$$

The form $y = ab^x$

METHOD

$\log y = \log a + x \log b$. Put $Y = \log y$, $A = \log a$, $B = \log b$: $Y = A + Bx$. Same normal equations. Then $a = 10^A$, $b = 10^B$.

Example 2. Fit $y = ab^x$ to the cars data. Class notes

Same table: $A = 1.1055$, $B = 0.02248$. $a = 10^{1.1055} = 12.75$, $b = 10^{0.02248} = 1.053$. $\mathbf{y = 12.75 (1.053)^x}$

5.5 Power Curve $y = ax^b$

METHOD AND NORMAL EQUATIONS MAR/APR-2024 · 2M

$\log y = \log a + b \log x$. Put $Y = \log y$, $X = \log x$, $A = \log a$: $Y = A + bX$.

Normal equations: $nA + b \sum X = \sum Y$, $A \sum X + b \sum X^2 = \sum XY$. Then $a = 10^A$.

PYQ. For the cars data, fit $y = ax^b$. Nov-2024 · 10M Class notes

x	y	$X = \log x$	$Y = \log y$	X^2	XY
2	13	0.3010	1.1139	0.0906	0.3353
7	21	0.8451	1.3222	0.7142	1.1174
9	23	0.9542	1.3617	0.9106	1.2994
1	14	0	1.1461	0	0
5	15	0.6990	1.1761	0.4886	0.8221
12	21	1.0792	1.3222	1.1647	1.4269
		3.8785	7.4422	3.3686	5.0011

$6A + 3.8785b = 7.4422$; $3.8785A + 3.3686b = 5.0011$. Solve: $b = 0.2209$, $A = 1.0976$. $a = 10^{1.0976} = 12.52$.

$$\mathbf{y = 12.52 x^{0.2209}}$$

PYQ. Fit $y = ax^b$: X: 50, 100, 250, 500, 1000; Y: 108, 53, 24, 9, 5. Mar/Apr-2022 · 7M

$\log x$: 1.6990, 2.0000, 2.3979, 2.6990, 3.0000 (sum 11.7959; $\sum X^2 = 28.9211$).

$\log y$: 2.0334, 1.7243, 1.3802, 0.9542, 0.6990 (sum 6.7911); $\sum XY = 14.8853$.

$5A + 11.7959b = 6.7911$; $11.7959A + 28.9211b = 14.8853$. Solve: $b = -1.040$, $A = 3.8116$, $a = 10^{3.8116} = 6481$. $y = 6481 x^{-1.04}$

PYQ. Estimate y at $x = 5$ by fitting a least squares curve $y = \frac{b}{x(x-a)}$: x: 3.6, 4.8, 6.0, 7.2, 8.4, 9.6, 10.8; y: 0.83, 0.31, 0.17, 0.10, 0.07, 0.05, 0.04. Apr/May-2023 · 10M

Step 1 Make it linear: $x(x-a) = \frac{b}{y} \Rightarrow \frac{1}{xy} = \frac{x-a}{b} = -\frac{a}{b} + \frac{1}{b}x$. Put $Y = \frac{1}{xy}$, $A = -\frac{a}{b}$, $B = \frac{1}{b}$: $Y = A + Bx$.

Step 2 Y: 0.3347, 0.6720, 0.9804, 1.3889, 1.7007, 2.0833, 2.3148. $\sum Y = 9.4748$, $\sum xY = 79.599$, $\sum x = 50.4$, $\sum x^2 = 403.2$.

Step 3 $7A + 50.4B = 9.4748$; $50.4A + 403.2B = 79.599$. Solve: $B = 0.2822$, $A = -0.6786$.

Step 4 $b = \frac{1}{B} = 3.543$, $a = -Ab = 2.404$. So $y = \frac{3.543}{x(x-2.404)}$.

Step 5 At $x = 5$: $y = \frac{3.543}{5(2.596)} = 0.273$.

$y(5) \approx 0.27$

5.6 Quick Revision Sheet

Curve	Change of variable	Normal equations	Back to constants
$y = a + bx$	—	$na + b\sum x = \sum y$; $a\sum x + b\sum x^2 = \sum xy$	—
$y = a + bx + cx^2$	—	multiply by 1, x , x^2 and sum (3 equations)	—
$y = ae^{bx}$	$Y = \log y$	$nA + B\sum x = \sum Y$; $A\sum x + B\sum x^2 = \sum xY$	$a = 10^A$, $b = B/0.4343$
$y = ab^x$	$Y = \log y$	same as above	$a = 10^A$, $b = 10^B$
$y = ax^b$	$Y = \log y$, $X = \log x$	$nA + b\sum X = \sum Y$; $A\sum X + b\sum X^2 = \sum XY$	$a = 10^A$

★ Assignment — Unit 5 (5 questions)

1. State the principle of least squares. Derive the normal equations to fit $y = a + bx + cx^2$. [5M]
2. Fit a straight line to x: 0, 1, 2, 3, 4; y: 1, 1.8, 3.3, 4.5, 6.3. Estimate y at x = 5. [5M]
3. Fit a parabola $y = a + bx + cx^2$ to x: 0, 1, 2, 3, 4; y: 1, 1.8, 1.3, 2.5, 6.3. [10M]
4. Fit $y = ae^{bx}$ to x: 0, 2, 4; y: 5.012, 10, 31.62. [5M]
5. Fit $y = ax^b$ to x: 1, 2, 3, 4, 5; y: 0.5, 2, 4.5, 8, 12.5. [10M]

Answers for self-check: Q2: $y = 0.72 + 1.33x$, $y(5) = 7.37$. Q3: $y = 1.42 - 1.07x + 0.55x^2$. Q4: $y = 4.64e^{0.460x}$. Q5: $y = 0.5x^2$ (exact fit).