Fitting of a Parabola - by method of least squares

The ideas of fitting of a straight line are extended to find the least-squares Parabola that fits a set of sample points, whose equation is,

$$y = a + bx + cx^2$$

where a, b, c are constants.

Let the sample points be (x_1, y_1) (x_2, y_2) $\dots (x_n, y_n)$. Then the values of y on the least squares Parabola corresponding to $x_1, x_2 \dots x_n$ are

$$a + bx_1 + cx_1^2, \; a + bx_2 + cx_2^2, \dots a + bx_n + cx_n^2$$

The deviations from $y_1, y_2 \dots y_n$ are given by

$$\Sigma d^2 = \Sigma (a + bx + cx^2 - y)^2$$

This is a function of a, b and c.

ie $$E(a, b, c) = \Sigma (a + bx + cx^2 - y)^2$$

The necessary condition for this function to be a minimum is,

$$\frac{\partial E}{\partial a} = 0, \quad \frac{\partial E}{\partial b} = 0, \quad \frac{\partial E}{\partial c} = 0.$$

ie $$\frac{\partial E}{\partial a} = \Sigma \frac{\partial}{\partial a} (a + bx + cx^2 - y)^2 = \Sigma 2 (a + bx + cx^2 - y)$$

$$\frac{\partial E}{\partial b} = \Sigma \frac{\partial}{\partial b} (a + bx + cx^2 - y)^2 = \Sigma 2x (a + bx + cx^2 - y)$$

$$\frac{\partial E}{\partial c} = \Sigma \frac{\partial}{\partial c} (a + bx + cx^2 - y)^2 = \Sigma 2x^2 (a + bx + cx^2 - y)$$

ie $\sum 2(a+bx+cx^2-y)=0$.

$\sum 2a+\sum 2bx+\sum 2cx^2-\sum 2y=0$.

$\therefore \sum a+\sum bx+\sum cx^2-\sum y=0$,

$\therefore na+b\sum x+c\sum x^2-\sum y=0$.

$\therefore \sum y = na+b\sum x+c\sum x^2$

also $\sum 2x(a+bx+cx^2-y)=0$

$\therefore \sum 2ax+\sum 2bx^2+\sum 2cx^3-\sum 2xy=0$

$\therefore \sum ax+\sum bx^2+\sum cx^3-\sum xy=0$

$\therefore a\sum x+b\sum x^2+c\sum x^3-\sum xy=0$.

$\therefore \sum xy = a\sum x+b\sum x^2+c\sum x^3$

also $\sum 2x^2(a+bx+cx^2-y)=0$.

$\therefore \sum 2ax^2+\sum 2bx^3+\sum 2cx^4-\sum 2x^2y=0$

$\therefore \sum ax^2+\sum bx^3+\sum cx^4-\sum x^2y=0$

$\therefore a\sum x^2+b\sum x^3+c\sum x^4-\sum x^2y=0$

$\therefore \sum x^2y = a\sum x^2+b\sum x^3+c\sum x^4$

So the normal equations are,

$\sum y = na+b\sum x+c\sum x^2$

$\sum xy = a\sum x+b\sum x^2+c\sum x^3$

$\sum x^2y = a\sum x^2+b\sum x^3+c\sum x^4$

Solving these equations, the best values of a, b, & c can be determined.

Prob(1)

Fit a least squares Parabola having the form $y = a + bx + cx^2$ to the following data

x :	1.2	1.8	3.1	4.9	5.7	7.1	8.6	9.8
y :	4.5	5.9	7.0	7.8	7.2	6.8	4.5	2.7

The normal equns are,

$$\Sigma y = an + b\Sigma x + c\Sigma x^2$$

$$\Sigma xy = a\Sigma x + b\Sigma x^2 + c\Sigma x^3$$

$$\Sigma x^2 y = a\Sigma x^2 + b\Sigma x^3 + c\Sigma x^4.$$

x	y	x^2	x^3	x^4	xy	x^2y
1.2	4.5	1.44	1.73	2.08	5.40	6.48
1.8	5.9	3.24	5.83	10.49	10.62	19.12
3.1	7.0	9.61	29.79	92.35	21.70	67.27
4.9	7.8	24.01	117.65	576.48	38.22	187.28
5.7	7.2	32.49	185.19	1055.58	41.04	233.93
7.1	6.8	50.41	357.91	2541.16	48.28	342.79
8.6	4.5	73.96	636.06	5470.12	38.70	332.82
9.8	2.7	96.04	941.19	9223.66	26.46	259.31
Σx = 42.2	Σy = 46.4	Σx^2 = 291.20	Σx^3 = 2275.35	Σx^4 = 18971.92	Σxy = 230.42	Σx^2y = 1449.00

The normal equations becomes,

$$8a + 42.2b + 291.20c = 46.4$$

$$42.2a + 291.20b + 2275.35c = 230.42$$

$$291.20a + 2275.35b + 18971.92c = 1449.$$

Solving these we get

$$a = 2.588$$

$$b = 2.065$$

$$c = -0.2110.$$ Hence the required Least squares Parabola has the equation

$$y = 2.588 + 2.065x - 0.2110x^2$$

(3)

Prob: 2.

Fit the Parabola $y = a + bx + cx^2$ ~~$y = ax^2 + bx + c$~~, for the following data by the method of least squares. Estimate the value of y when $x = 10$

x :	1	2	3	4	5	6	7	8	9
y :	2	6	7	8	10	11	11	10	9

x	y	x^2	x^3	x^4	xy	x^2y
1	2	1	1	1	2	2
2	6	4	8	16	12	24
3	7	9	27	81	21	63
4	8	16	64	~~512~~ 256	32	128
5	10	25	125	625	50	250
6	11	36	216	1296	66	396
7	11	49	343	2401	77	539
8	10	64	512	4096	80	640
9	9	81	729	6561	81	729
$\Sigma x = 45$	$\Sigma y = 74$	$\Sigma x^2 = 285$	2025 (Σx^3)	15333 (Σx^4)	421 (Σxy)	2771 ($\Sigma x^2 y$)

The normal equns are $\qquad n = 9$

$$\Sigma y = an + b\Sigma x + c\Sigma x^2$$

$$\Sigma xy = a\Sigma x + b\Sigma x^2 + c\Sigma x^3$$

$$\Sigma x^2 y = a\Sigma x^2 + b\Sigma x^3 + c\Sigma x^4$$

ie

$$74 = 9a + 45b + 285c \quad \text{---} (1)$$

$$421 = 45a + 285b + 2025c \quad \text{---} (2)$$

$$2771 = 285a + 2025b + 15333c \quad \text{---} (3)$$

$(1) \times 5 \rightarrow \quad 370 = 45a + 225b + 1425c \quad — (4)$

$(2) - (4) \rightarrow \quad 51 = 60b + 600c \quad — (5)$

$(2) \times 57 \rightarrow \quad 23997 = 2565a + 16245b + 115425c \quad — (6)$

$(3) \times 9 \rightarrow \quad 24939 = 2565a + 18225b + 137997c \quad — (7)$

$(7) - (6) \rightarrow \quad 942 = 0 + 1980b + 22572c \quad — (8)$

$51 = 60b + 600c \quad — (5)$

$(5) \times 33 \rightarrow \quad 1683 = 1980b + 19800c \quad — (9)$

$(9) - (8) \rightarrow \quad 741 = {}^{-}2772c$

$$c = \frac{{}^{-}741}{2772} = {}^{-}0.267$$

Substituting the value of c in (5)

$51 = 60b + 600 \times {}^{-}0.267$

$= 60b + {}^{-}160.2$

$60b = 211.2$

$b = 211.2/60 = 3.52$

Substituting the value of b & c in (1)

$74 = 9a + 45 \times 3.52 + 285 \times {}^{-}0.267$

$= 9a + 158.4 - 76.095$

$= 9a + 82.305$

$9a = -8.305$

$a = -8.305/9 = {}^{-}0.92$

So $a = -0.92$;

$b = 3.52$

$c = -0.267$

Hence the least square Parabola has the equation,

$$y = -0.92 + 3.52x - 0.267x^2$$

when $x = 10$,

$y = -0.92 + 3.52 \times 10 - 0.267 \times 10^2$

$= -0.92 + 35.2 - 26.7$

$= 35.2 - 27.62 = 7.58$

So $y = 7.58$ when $x = 10$

Prob: 2 Fit a second degree curve of regression of y on x to the following data.

(x, y): $(1, 2)$ $(0, 0)$ $(0, 1)$ $(1, 2)$

[second degree curve can be taken as

$y = a + bx + cx^2$]

Prob(2) Fit a Parabola of the type $y = a + bx + cx^2$ to the following data.

x:	10	15	20	25	30	35	40
y:	11	13	16	20	27	34	41