Example on Multiple Linear Regression

Use multiple linear regression to fit the following table of data:

x_1	**1**	**2**	**3**	**4**	**5**
x_2	**4**	**3**	**2**	**1**	**0**
y	**18**	**16**	**16**	**12**	**10**

Compute coefficients and the error of estimate.

Solution:

$y = f(x_1, x_2) = a_1 + a_2x_1 + a_3x_2$

Then, the sum of square of errors is given by:

Q = $\sum_{i=1}^{n}\{y_i - f(x_1, x_2) = \}^2 = \sum_{i=1}^{n}\{y_i - (a_1 + a_2x_1 + a_3x_2)\}^2$

Differentiation above equation w. r. t. a1, a2, a3 and equating them to zero, we will get the condition for minimum error.

$$\frac{dq}{da_1} = -2\sum\{y_i - (a_1 + a_2x_1 + a_3x_2)\} = 0$$

$$a_1\, n + a_2\sum x_{1i} + a_3\sum x_{2i} = \sum y_i$$

$$\frac{dq}{da_2} = -2\sum\{y_i - (a_1 + a_2x_{1i} + a_3x_{2i})x_{1i}\} = 0$$

$$a_1\sum x_{1i} + a_2\sum \boldsymbol{x_{2i}^2} + a_3\sum x_{1i}x_{2i} = \sum x_{1i}y_i$$

$$\frac{dq}{da_3} = -2\sum\{y_i - (a_1 + a_2x_i + cz_i)z_i\} = 0$$

$$a_1\sum x_{2i} + a_2\sum x_{1i}x_{2i} + a_3\sum {x_{2i}}^{\mathbf{2}} = \sum x_{2i}y_i$$

SO the simultaneous equations are:

$$a_1\, n + a_2\sum x_{1i} + a_3\sum x_{2i} = \sum y_i$$

$$a_1\sum x_{1i} + a_2\sum \boldsymbol{x_{1i}^2} + a_3\sum x_{1i}x_{2i} = \sum x_{1i}y_i$$

$$a_1\sum x_{2i} + a_2\sum x_{1i}x_{2i} + a_3\sum {x_{2i}}^{\mathbf{2}} = \sum x_{2i}y_i$$

x_1	x_2	$\boldsymbol{y}$	$\boldsymbol{x_1^2}$	$\boldsymbol{x_2^2}$	$\boldsymbol{x_1x_2}$	$\boldsymbol{x_1y}$	$\boldsymbol{x_2y}$
1	4	18	1	16	4	18	72
2	3	16	4	9	6	32	48
3	2	16	9	4	6	48	32
4	1	12	16	1	4	48	12
5	0	10	25	0	0	50	0
∑15	∑10	∑72	∑55	∑30	∑20	∑196	∑164

So, three simultaneous equations will be:
$5a_1 + 15a_2 + 10a_3 = 72$ ……………………………………………….. (i)
$15a_1 + 55a_2 + 20a_3 = 196$ … ………………………………………… (ii)
$10a_1 + 20a_2 + 30a_3 = 164$ … ……………………………………………..(iii)

On solving equations using Gauss Elimination we can get the values of unknowns.
$a_1 = 20.4$
$a_2 = -2$
$a_3 = 0$
Thus the regression function is: $y = 20.4 - 2x_1$