$$= [2.08 \quad -0.74 \quad 0.768 \quad -0.299 \quad 1.39]\begin{bmatrix} 32 \\ 124 \\ 161 \\ 160 \\ 110 \end{bmatrix} - 6 * (5.33)^2$$

$$SSR(X_1, X_2, X_3, X_4) = 203.508 - 170.666$$

$$SSR(X_1, X_2, X_3, X_4) = 32.84$$

We now find the sum of squares for a model containing only the variables X4 and X3, as follows:

$$SSR(X_3, X_4) = \hat{\beta}' X' y - n\bar{y}^2$$

$$= [4.644 \quad -0.2595 \quad 0.6683]\begin{bmatrix} 32 \\ 160 \\ 110 \end{bmatrix} - 6 * (5.33)^2$$

$$SSR(X_3, X_4) = 180.601 - 170.66$$

$$SSR(X_3, X_4) = 9.934$$

Thus, we now find the additional sum of squares for adding the variables X1, X2 to a model containing the variables X3, X4:

$$SSR(X_1, X_2 / X_3, X_4) = SSR(X_1, X_2, X_3, X_4) - SSR(\, X_3, X_4)$$

$$SSR(X_1, X_2 / X_3, X_4) = 32.84 - 9.934$$

$$SSR(X_1, X_2 / X_3, X_4) = 22.88$$

The shortcut method:

$$SSR(X_1, X_2 / X_3, X_4) = \hat{\beta}_{12}{}' C_{12}^{-1} \hat{\beta}_{12}$$

$$= [\hat{\beta}_1 \quad \hat{\beta}_2]\begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix}^{-1} \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \end{bmatrix}$$

$$= [-0.74 \quad 0.768]\begin{bmatrix} 0.2811 & 0.0135 \\ 0.0135 & 0.03147 \end{bmatrix}^{-1} \begin{bmatrix} -0.74 \\ 0.768 \end{bmatrix}$$

$$SSR(X_1, X_2 / X_3, X_4) = 22.88$$

Choose the best regression equation

Suppose there is a regression line equation consisting of m independent variables, as shown in the model:

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_m x_{im} + e_i, i = 1,2, \dots, n$$

Since introducing large numbers of independent variables into the equation costs effort, time, and expense, it is better to choose an equation that contains the smallest possible number of independent variables.

There are several ways to choose the best regression equation, which are as follows:

1− The method of all possible regressions

2−Reverse deletion method

3− Forward selection method (direct)

4−Stepwise selection method

Three criteria will be used to compare the selected equations:

1−Coefficient of determination R^2

2−Mean Square Error Mse

3−C_p value

In order to illustrate the above methods, the following example will be applied:

Example: From the following data, determine the best regression equation using the method of all possible regressions.

Observation	y	X1	X2	X3	X4
1	78.5	7	26	6	60
2	74.3	1	29	15	52
3	104.3	11	56	8	20
4	87.6	11	31	8	47
5	95.9	7	52	6	33
6	109.2	11	55	9	22
7	102.7	3	71	17	6
8	72.5	1	31	22	44
9	93.1	2	54	18	22
10	115.9	21	47	4	26
11	83.8	1	40	23	34
12	113.3	11	66	9	12
13	109.4	10	68	8	12

Method of all possible regressions:

This method is summarized by finding all possible types of regression equations using all possible combinations of independent variables. If we had m independent variables, the total number of total equations would be 2^m. In our example above, which includes four independent variables, the total number of equations will be $2^m = 16$ they are as follows:

These equations can be put into five groups:

Group A is the equations that do not contain any independent variable.

Set B is a set of equations that contain one independent variable.

Set C is a set of equations that contain two independent variables.

Set D is a set of equations that contains three independent variables.

Set E is a set of equations that contains four three independent variables.

Thus, the estimation of the equations will be as follows:

Number in the equation	the group	The independent variables in the equation	Sum of squares		Parameter estimation				
			Regression	Error	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$
0	A	Nothing	-						
1		X1	1450.07	1226.6	81.47	11.69			
1		X2	1809.4	906.33	57.42		0.789		
1	B	X3	776.36	199.005	110.2			-1.256	
1		X4	1831.8	883.86	117.5				-0.738
2		X1X2	2657.8	557.045	52.57	1.46	0.662		
2		X1X3	14888.9	1227.07	7.34	2.31		0.94	
2		X1X4	2641.01	74.76	103.09	-1.44			-0.614
2	C	X2X1	2300.3	415.44	72.07		0.731	-1.00	
2		X2X4	1846.88	868.88	94.16		0.311		-0.457
2		X3X4	2540.02	175.73	131.28			-1.2	-0.724
3		X1X2X3	2667.65	48.11	48.19	1.69	0.657	0.25	
3	D	X1X2X4	2667.7	47.97	71.64	1.45	0.416		-0.277
3		X1X3X4	2664.92	50.83	203.6		0.923	-1.44	-1.55
3		X2X3X4	2641.94	73.81	11.68	1.05		-0.41	-0.64
4	E	X1X2X3X4	2667.8	47.86	62.40	1.55	0.510	-0.102	-0.144

The three criteria will be calculated:

1-Coefficient of determination R^2

2-Mean Square Error Mse

3-$C_p$value

The results are shown in the table below: