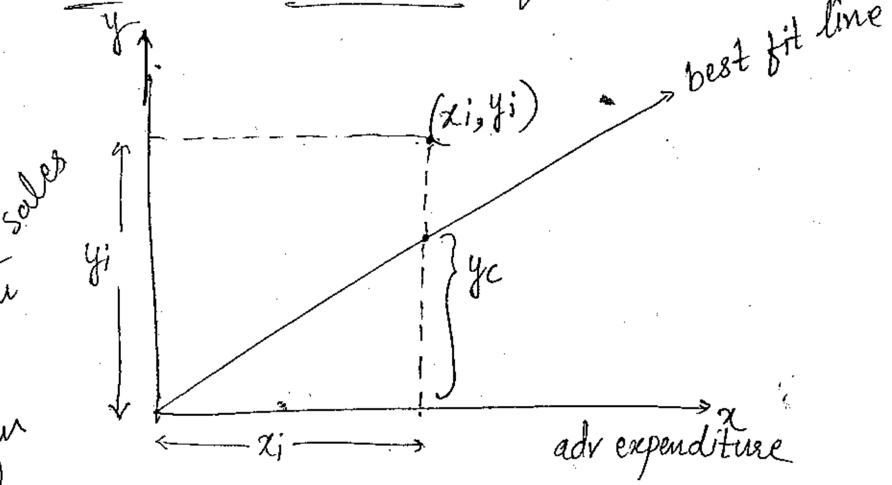
It seems reasonable that the farther away a point is from the estimating line, the more serious is the error. We would have several small absolute errors than one large absolute error. In effect we wish to penalise large absolute errors so that we can avoid them. We can accomplish this task if we square the individual error before we add them. Squaring gives weight to both +ve and -ve errors and magnifies the larger errors.

- e_i is the <u>deviation</u> of a point.



Here, we calculate y's regression on x. Because y is DV & x is IV. Reverse application of regression will not be applicable in this case, because x is not depended on y, but the other way round.

$$e_i = y_i - y_c$$
(deviation)

$$\therefore e_1 = y_1 - y_c$$

Now, $\sum_{e=i}^{n} e_i^2$ should be the least.

i.e summation of deviations of all points lying outside the line should be the least.

That would be the best fit line.