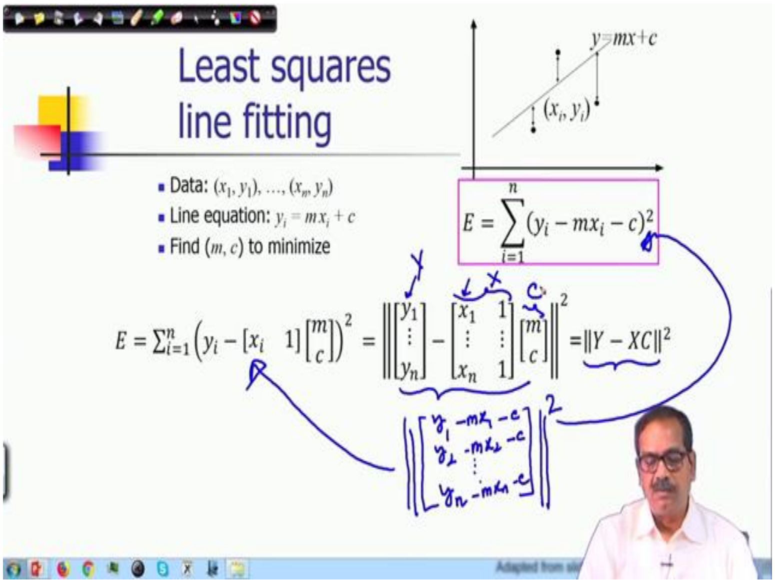
Computer Vision
Prof. Jayanta Mukhopadhyay
Department of Computer Science and Engineering
Indian Institute of Technology, Kharagpur

Lecture – 32
Feature Matching and Model Fitting Part – IV

(Refer Slide Time: 00:24)



We are discussing about least squares line fitting and as we have considered that we have been given a set of data points and the form of the model in this case is a line, straight line given in the form of equation y=mx+c and the error term is defined as the sum of square of vertical divisions of the observed value from the predicted value, which is been given by the equation

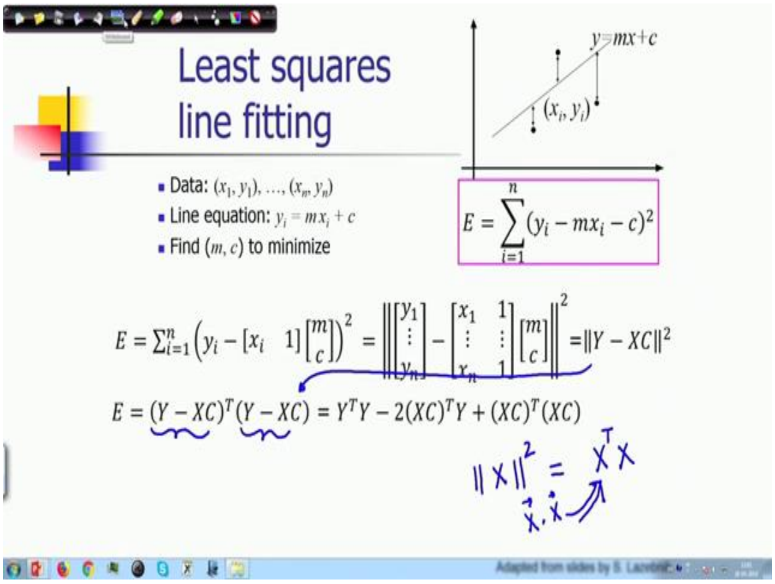
$$E = \sum_{i=1}^{n} (y_i - mx_i - c)^2$$

So, we have to minimize this error with appropriate values of m and c. So, that is a problem, find m and c to minimize this error. So, we can write the expression of error in this form, as you can see here, that we have converted this expression in the short form of matrix notation of your data, just to elaborate that how this representation

$$E = \sum_{i=1}^{n} (y_i - [x_i \; 1][\begin{matrix} m \\ c \end{matrix}])^2 = \left\| \begin{bmatrix} y_1 \\ y_n \end{bmatrix} - \begin{bmatrix} x_1 & 1 \\ x_n & 1 \end{bmatrix} \begin{bmatrix} m \\ c \end{bmatrix} \right\|^2 = ||Y - XC||^2$$

Similarly, in the second row we will get y 2 minus mx 2, minus c and in this way, you will get y n minus mx n, minus c. So, this is the column vector and if you would like to take the norm of this vector and take the square of the norm, that exactly we will give you this equation or that is what we will give you also this form, you can verify that. So, this is nothing but, represented in this short form. So, this error E is represented in a shorter matrix notation as capital Y minus XC where, Y is the this matrix, X is this matrix and C is this matrix. So, this is how this notation has been derived, let me wipe out this writings and proceed further.

(Refer Slide Time: 03:05)



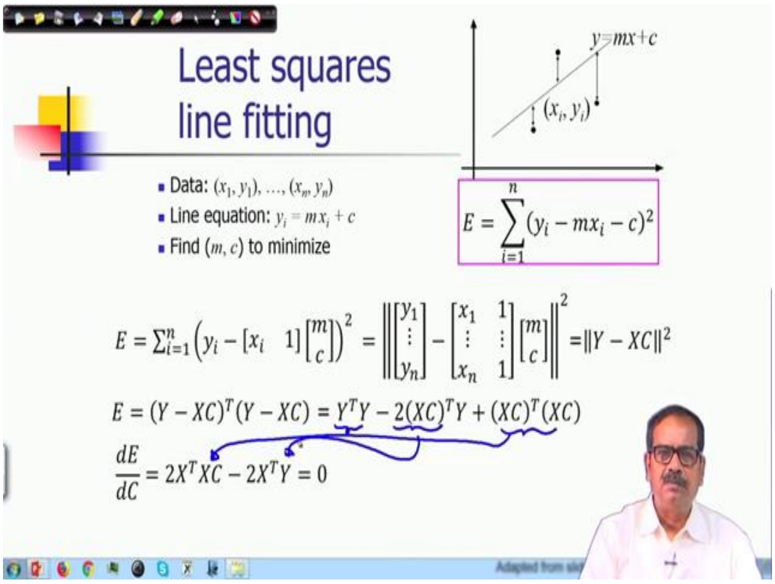
So, where E can be error can be expressed now, we have a matrix notation and as the norm of a matrix can be written in this form

$$E = (Y - XC)^T(Y - XC) = Y^TY - 2(XC)^TY + (XC)^T(XC)$$

(Y-XC) just to explain once again suppose, you have a matrix X or a vector, which is represent in the form of a column vector, square of the magnitude, that is nothing but, it is a you understand it is a scalar amount.

And now you expand each one using matrix algebra and then you can perform the multiplications also. Since, matrix is a linear operation. So, you can perform derivatives everything in the matrix form itself and you can extinct the analysis, what we know for ordinary single dimensional variable, single dimensional function cases.

(Refer Slide Time: 04:45)



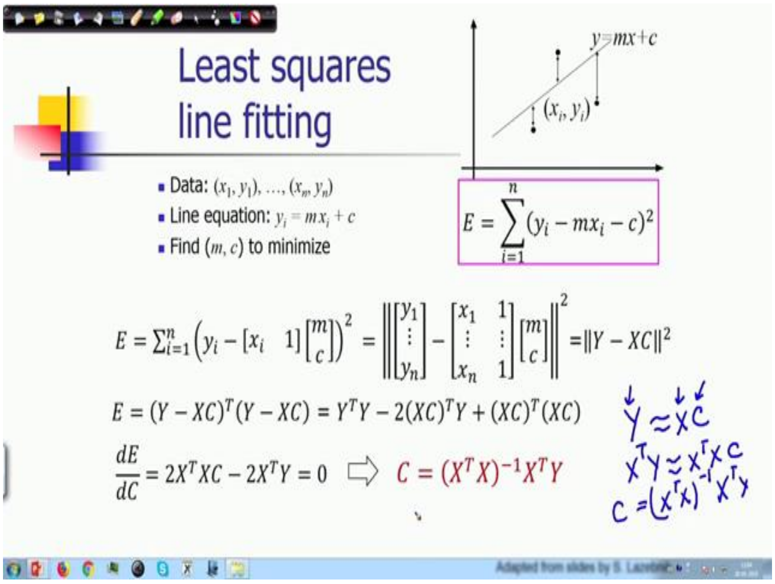
So, in this case, if I perform the derivative with respect to C then, we will get these equations, which means at independently you are taking derivative of E with respect to each component of C,

$$\frac{dE}{dC} = 2X^TXC - 2X^TY = 0$$

So, this relationships from here, which is established or which can be derived using matrix algebra and their derivative that in short, we have written in this form, as you can see here this is not related to see, this is a kind of constant term, these are the term which are related to C.

You can intuitively extent your knowledge of differential calculus, applying single dimensional, you know variable or applying those or in ordinary single dimensional functional space, you can just extinct them using the matrix notation, you need to know practice on that part, when you are deriving this in a very short and faster way. Otherwise you can do it component wise, we will find they have a one to one relationships with this kind of expression. So, once you obtain this then you perform once again, the algebraic manipulations with this relation.

(Refer Slide Time: 06:40)

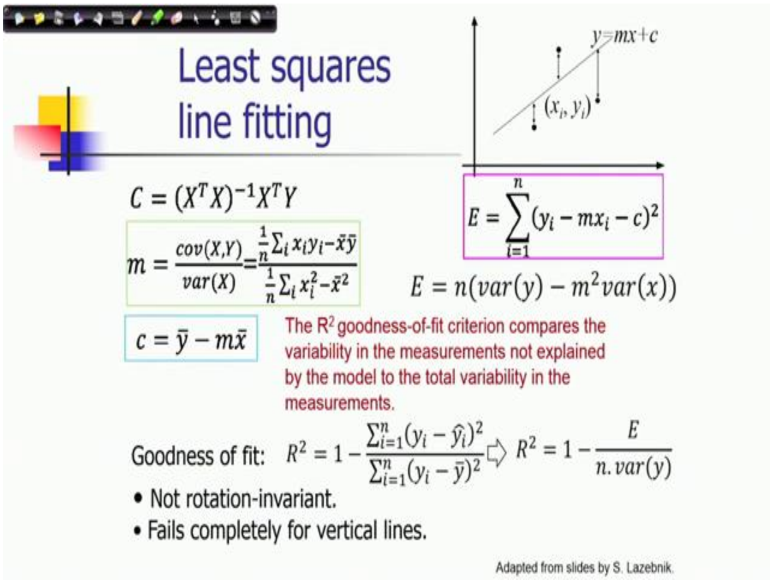


$$C = (X^T X)^{-1} X^T Y$$

This is the famous pseudo inverse relationships what we previously also discussed, just to give you the same picture, that I am trying to fit a model like Y=XC. So, I am trying to get C, given Y and X. So, these are the given data.

That is why what exactly we have derived here, it is the result of minimization of the error.

(Refer Slide Time: 07:41)



So, just to expand this relationships in a more granular level of data elements; that means, to get a solution of m and c, we can get this form, m can be expressed as ratio

$$m = \frac{cov(X,Y)}{var(X)} = \frac{\frac{1}{n}\sum_i \ x_i y_i - \hat{x}\hat{y}}{\frac{1}{n}\sum_i \ {x_i}^2 - \hat{x}^2}$$

Anyway, there are simpler way of deriving this relationships by performing partial derivative of E with respect to m and with respect to c and following 2 equations and solve for m and c.

So, these are the equations we will get here, y bar and x bar is the mean of x is and mean of. So, y bar is mean of y is and x bar is mean of x is. So, error can be estimated by replacing the value of m and c in that expression, eventually you can find that this would come like this

$$E = n(var(y) - m^2 var(x))$$

So, goodness of fit as I mentioned, that any module fitting you should observe this error, it is related to this error and you can express as a quantity called

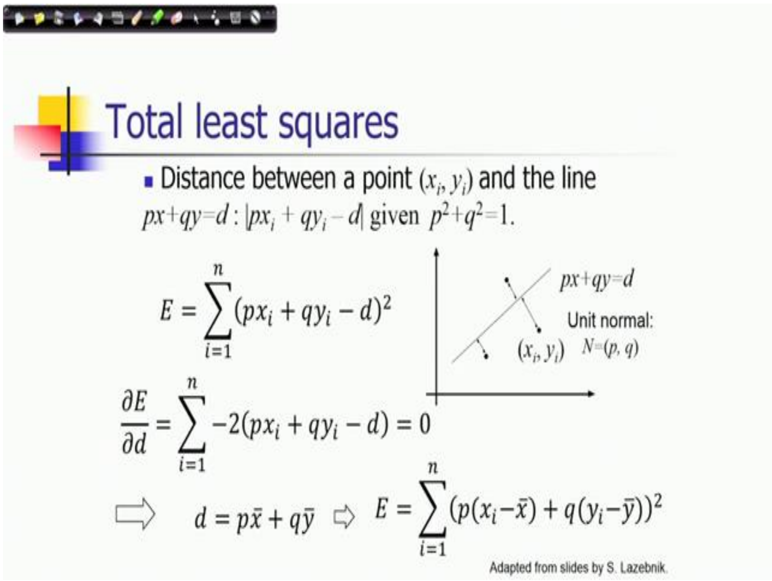
$$R^2 = 1 - \frac{\sum_{i=1}^{n} {y_i}^2 - \hat{y}^2}{\sum_{i=1}^{n} {y_i}^2 - \bar{y}^2} \rightarrow R^2 = 1 - \frac{E}{n.var(y)}$$

So, it is trying to explain that what part of this variance is explained by this data fitting. So, this is what it compares the variability in the measurements, not explained by the model to the total variability in the measurements.

So, if it is R square value if it is very high, then your model fit is good. So, this value should lie between 0 to 1 and you can also express in this way, this R square is related to E, sometimes it is called coefficients of liberation. So, once you fit a model you should find out also R square and the value should be quite high, say near about 0.8 or more than 0.8 that would be a good fit, we consider that as a good fit in that case.

You should also note that this technique is not rotational invariant and it fails completely for vertical lines because, if it is really a vertical line then, as you can understand your m is going to be almost it is m is. So, if it is vertical line, it is x = C. So, it should be almost like infinity y is coming now. So, it is very difficult to get term. So, variance of x is training to very very small and m is going to be infinity so, it will fail in that case.

(Refer Slide Time: 11:17)



So, there are techniques which takes care of the situations and also here the objective criteria for fitting model exposed in a different way and this technique is called total least squares. So, it is suitable for any such situation whether, it is a vertical line or not. Now, in this particular diagram it has been shown that, what kind of error measurement we are considering here instead of vertical deviation, rather we are considering the deviation of the point from that line itself. So, deviation means the perpendicular distance between the point and line.

So, those are the deviations and sum of square of those deviations that defined the error, we assume here the line is given in this form. So, px + qy = d. So, that the normal direction of this line is given p q that is the property as you can see. So, the distance between a point and the line as I mentioned, you have to consider the perpendicular distance which is algebraically given in this form,

$$|px_i + qy_i - d| given\ p^2 + q^2 = 1$$

So, you have to consider that representation. That you can make always in your straight line equation by choosing appropriate d or scaling them.

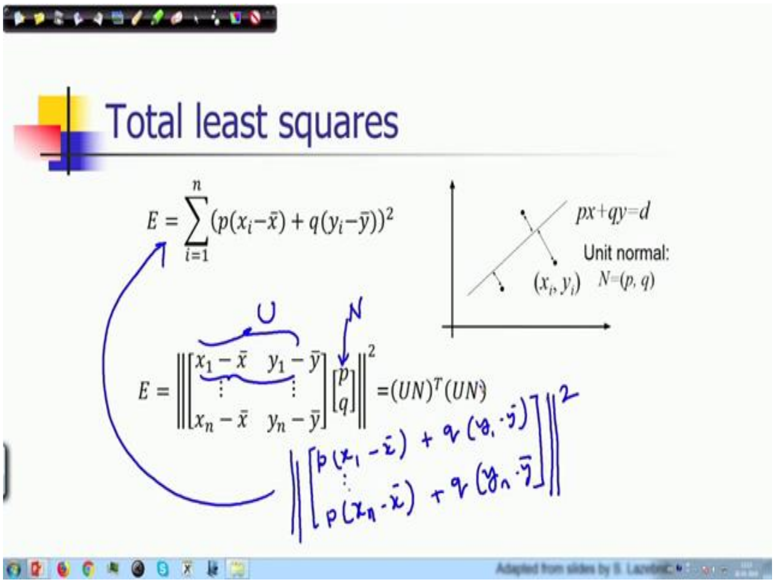
So, the error is given in this form, $px_i + qy_i - d$ is whole square and if I perform the derivative with respect to d then,

$$\frac{dE}{dd} = \sum_{i=1}^{n} -2(px_i + qy_i - d) = 0 \rightarrow d = p\bar{x} + q\bar{y}$$

$$E = \sum_{i=1}^{n} (p(x_i - \bar{x}) + q(y_i - \bar{y}))$$

Now, it is basically a function of p and q by replacing d there and the constant is that p square plus q square should be equal to 1. So, you have to choose that p and q, which will minimize this error E, with the constraint that $p^2 + q^2 = 1$, that is the problem that you need to solve.

(Refer Slide Time: 13:34)



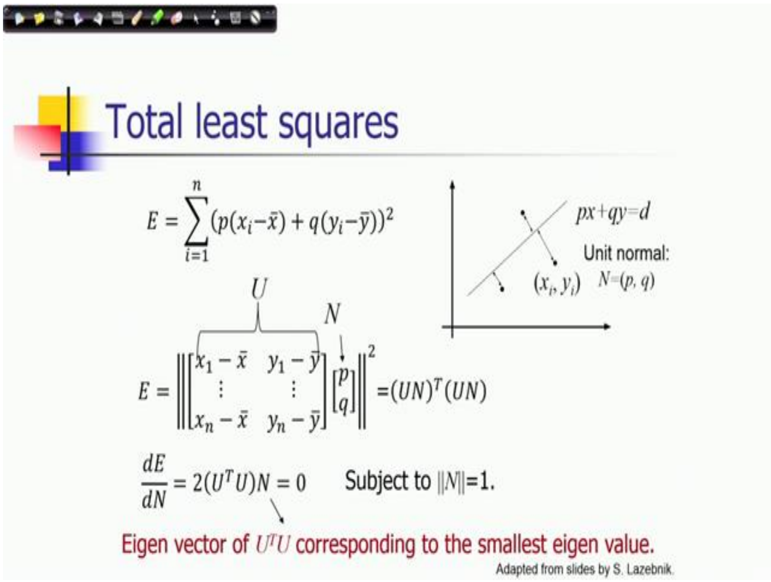
So, once again we can use the matrix notation to represent the square of error, sum of squares in this short form; that means, it is norm of this particular vector. So, just to elaborate once again just that you can see that, if you considered the first row, it is giving you, p into x 1 minus x bar plus, q into y 1 minus y bar and in this way, you consider each row is formed. So, the last row is p x n minus x bar plus q, y n minus y bar. So, this column vector and then if you take it is norm of square. So, you get this expression.

So, this is how it is representing in this form and you would like to make E = 0, you would like to get a solution but, this can be expressed also this is this form is

$$E = (UN)^T(UN)$$

So, this is how this short form is explained and let us continue with this representation.

(Refer Slide Time: 15:16)



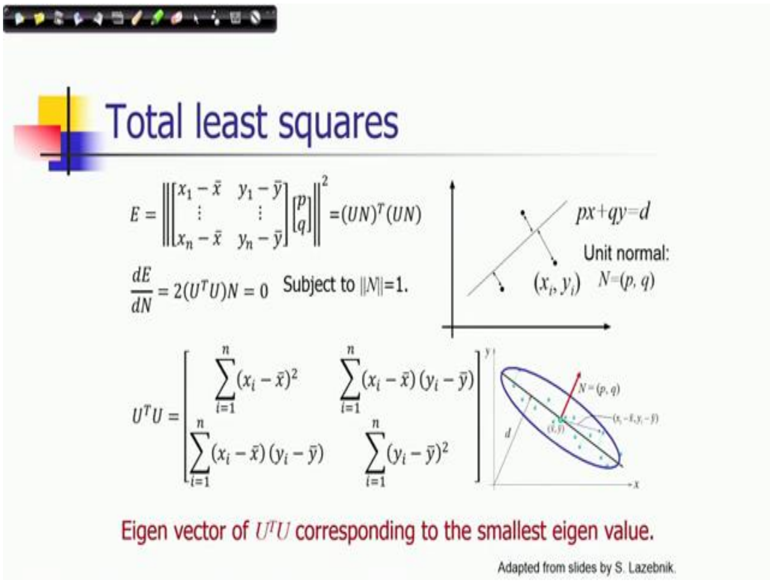
So, as I mentioned this is U this is N. So, we need to solve this problem, to minimize this error by choosing appropriate N. So, if I derive E with respect to N, we get this relationships

$$\frac{dE}{dN} = 2(U^TU)N = 0 \; subject \; to \; ||N|| = 1$$

which is p and q should be equal to 1 because, that is the constraint we have put, as you can see now, that actually this is giving a set of homogeneous equation.

So, you have to find out the 0 vector of U^TU, you can do it in various ways. So, one of the way could be that if you perform the, if you compute the eigenvectors eigenvalues, there will be one value which should give you the 0 but, ideally it should give 0 but, because of the measurements because of noise etcetera, you should consider that eigenvector which corresponds to the smallest eigenvalue of the matrix U^TU. So, that should be your solution in this case. So, this is what eigenvector of $U^T U$ corresponding to the smallest eigenvalue, that you should consider.

(Refer Slide Time: 16:53)

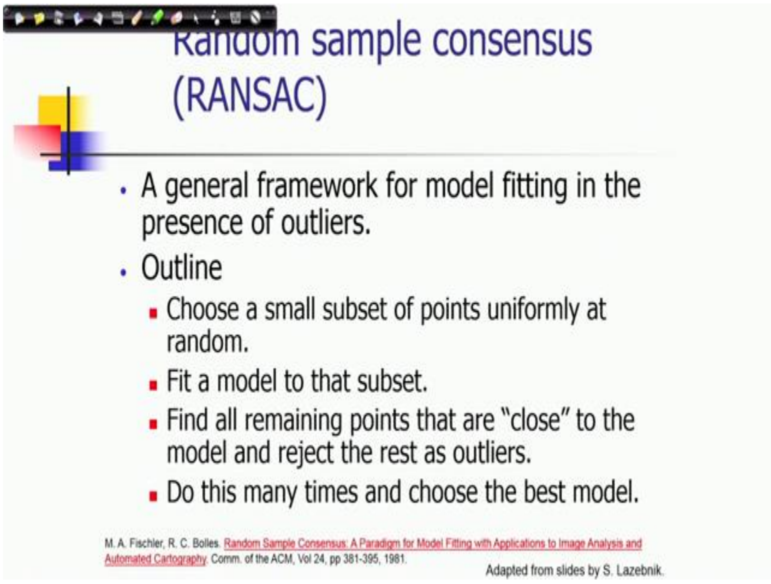


We can understand this result in a more by applying on notion of geometry, as I have summarised the analysis and you can see that the direction of normal is given that N is actually they in a perpendicular direction of the line, what you have fitted here and in the form of px + qy = d, d is an interpretation of parameter d as it is a perpendicular distance from origin to that line.

The structure of $U^T U$ is also interesting to note, you can see that it is a 2 X 2 matrix and whose eigenvalues there will be 2 eigenvalues and you have to choose the minimum eigenvalue and the vectors would be also of dimension 2 X 1 and the diagonal elements and the variances of x coordinates and y coordinates whereas, of diagonal elements and covariance of x and y and this is a symmetric matrix. So, this is what once again eigenvector of $U^T U$ corresponding to the smallest eigenvalue, that is the solution of this problem.

$$U^T U = \begin{bmatrix} \sum_{i=1}^{n} (x_i - \bar{x})^2 & \sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y}) \\ \sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y}) & \sum_{i=1}^{n} (y_i - \bar{y})^2 \end{bmatrix}$$

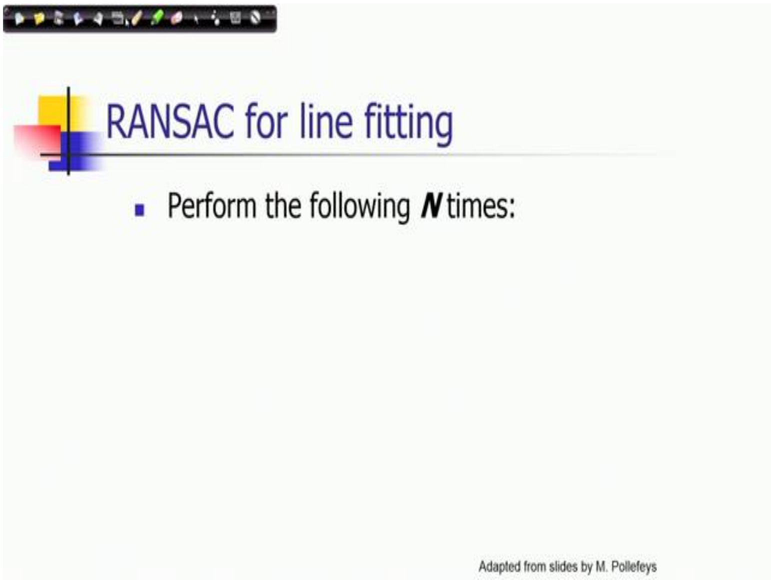
(Refer Slide Time: 18:18)



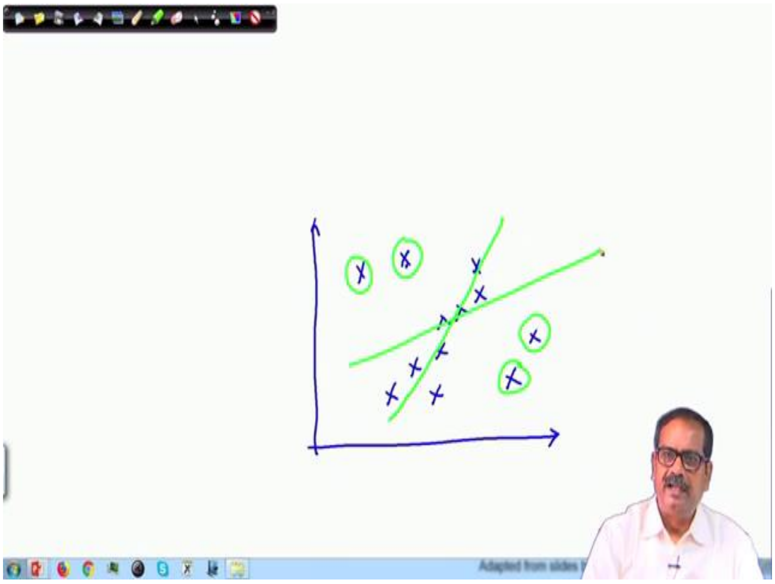
Let us now consider the problem of line fitting, when you have a situation where there are clotted observations or when you there are some outliers in your observations. In fact, this technique though we will be discussing with respect to line fitting, it provides you a general framework for model fitting in the presence outliers. You can apply this technique in our previous cases of finding homograph matrix or fundamental matrix or projection matrix, those we discussed in our previous topics.

So, outline of this technique is that, you should choose a small subset of points uniformly at random and then should fit a model to that subset then, find all remaining points that are close to the model and reject the rest as outliers and do this many times and choose the best model. So, this is what is this technique will be considering and this is what is used or general line of approach will be discussing with respect to a straight line for solving this. So, let me give you a general outline first here let me explain it.

(Refer Slide Time: 19:26)

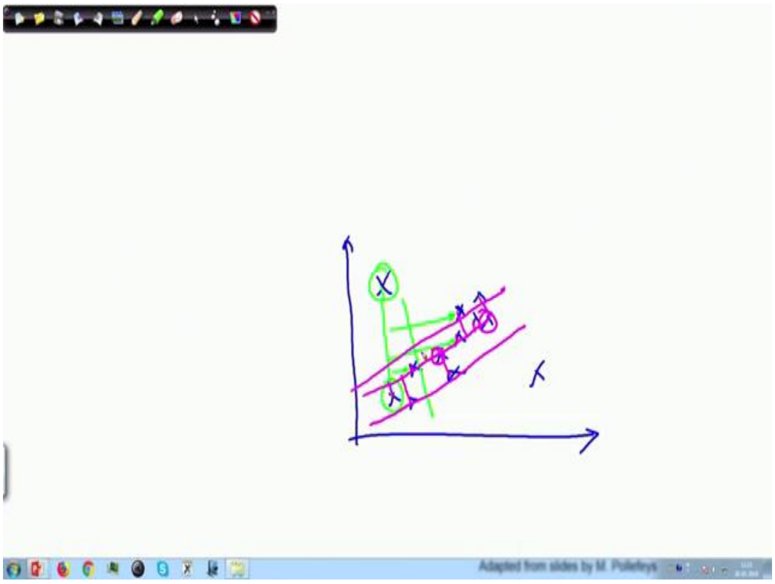


(Refer Slide Time: 19:33)



So, the approach what I mentioned earlier that considered, you have a set of points, which you required to fit in a straight line. There would be some points which are also deviated and which we considered their outline. So, you may consider say this could be a good fit of a straight line among this points but, if I apply the least square error method, the outlier will cause providing you a straight line which is not really fit in this one, it may come something like that, which will minimize the least square error, which is not desirable. So, your objective is to first find out a set of reliable points which we called inlier points and then apply model fitting on those points only.

(Refer Slide Time: 20:35)



So, in that case what we are doing once again, use that examples of some other points, which forms a good inlier points for fitting a straight line and see there are 2 outlier points and as I mentioned. So, what you consider here that you may choose, it may some take some initial setup points, arbitrarily choose any set up points, say initially you have chosen this point and see this point minimally and then you can fit a straight line and then you find out how many points which are lying in vicinity to the straight lines by looking at the perpendicular distance and give a threshold to declare that, they are lying in the outlier, there lying as an inlier point to that model.

Now, if this number is quite high, then you can say that you know this is a good model and then again use all those additional inlier points to refine the fitting of your model. But, as you can see in this scenario, there are only say 2 points, additionally 2 points only 4 points and I mean you can set your parameters in such a way which can decide that this is not a very good fit. So, we will try another setup points.

So, let us consider now that you have chosen, say this point and this point and this is your randomly again you are choosing and this is giving initial setup inlier points and then again you perform this test; that means, you are finding out which pointer line closer to this line and by configuring the distances and you consider all the points, which are outside this lines, that is outside this threshold there outliers.

So, now you can see at least, you can get a good mini points within this straight line and if you are with this number is high, when you may choose this setup points as inlier points and refine your model.

So, unless you get a good setup inlier points, you can go on doing this trial, go on doing this operation N number of times; that means, more number of times and maybe at after certain, if you do not get in any trial there is a good setup inlier points you may drop the idea of fitting this model because, it may happen the data is not good enough for fitting the model otherwise, if you get a good set of inlier points, then you can simply use them and refine your model. So, we will elaborate this process in the next lecture, for the time being let us stop here and.

Thank you for your attention.

Keywords: Least squares, line fitting, RANSAC.