Rock Engineering
Prof. Priti Maheshwari
Department of Civil Engineering
Indian Institute of Technology-Roorkee

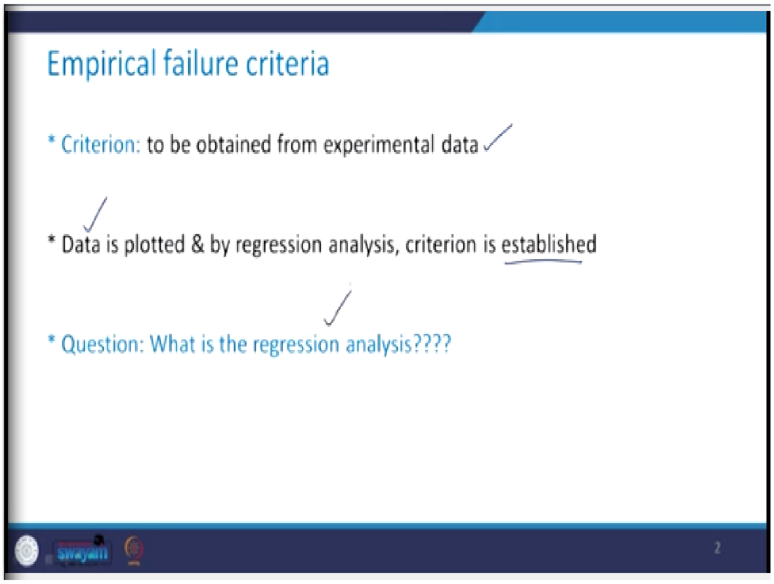
Lecture-31
Empirical Failure Criteria: Basics of Regression Analysis

Hello everyone. In the previous class we discussed about Mohr Coulomb failure criterion and Coulomb-Navier failure criteria. I told you about the derivation of these criteria, today we will take up the empirical failure criterion. I have been telling you that in case of the rock mechanics many a times we have to rely on the test results. And then we try to fit some kind of relationship between the two quantities which we get from the experimental data.

So, when we get such type of any equation or any kind of fitted curve, fitted mathematical model that is what we call as empirical correlation. So, today we will discuss few aspects related to empirical failure criterion. Before I tell you about these empirical failure criteria which are applicable in case of rocks and rock masses, we will learn some basics of the regression analysis. Because most of the time you have some data on y and you have the data on x axis, you just plot it in excel and then you try to see that which curve is the best fit curve between that data.

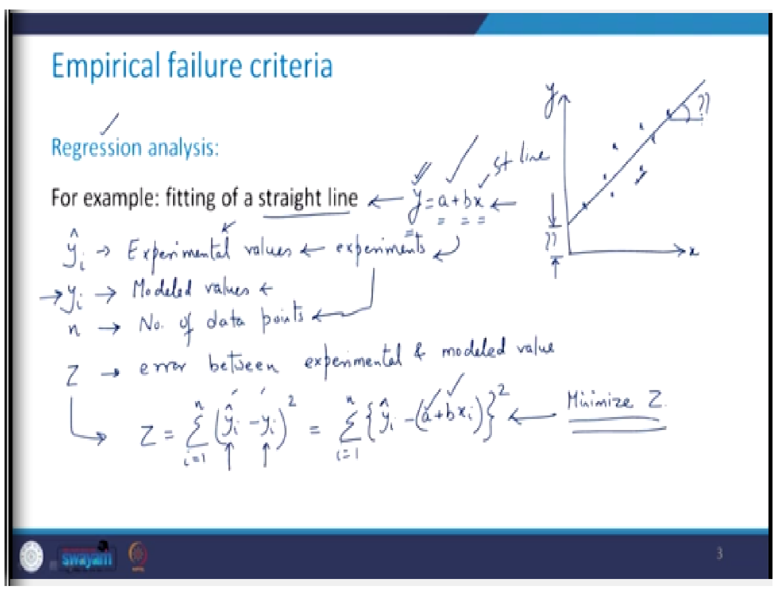
What happens behind that? How excel gives you the best fit data? Whether it is a straight line or a polynomial of 2-degree, 3 degree or whatever it is. So, today we will first discuss about the basics of the regression analysis and then we will learn about some of the failure criteria. And finally, I will be introducing you to the most commonly adopted empirical failure criteria for rocks which is Hoek and Brown criterion.

(Refer Slide Time: 02:50)



So, to start with the criterion as I mentioned it is to be obtained from the experimental data. This data is plotted and then in the regression analysis is carried out in order to establish the criterion. Question in front of us is what is this regression analysis?

(Refer Slide Time: 03:13)



Let us try to have a look with the help of a simple example of fitting a straight line between y and x with the help of some experimental data. So, please see here that I have here on x axis and here this is on y axis. So, we had some data say, that is say these are some data points which we already have. Now we want to fit a straight line to this data that means this kind of situation where it will have some intercept and it will have some slope, I really do not know what are these?

So, let us try to see mathematically that how can we fit a straight line into the data set. So, let us say that the fitting of the straight line and this equation is $y = a + bx$, this y and x they can be anything. Right now, I am just telling you the basics of the regression analysis. Then I will show you that how this analysis can be applicable in case of the rocks and rock masses. So, say I represent here $\hat{y}_i$, these are the experimental values.

And I represent another term y_i for the modeled values, what does that mean that? This is what that we are getting from the experiments that I am going to conduct in the lab. And since I want to fit this straight line which has $y = a + bx$. So, this is the model that I want to fit into this experimental data. So, whatever is the value of y that I will get from this expression they will fall under this category which are the modeled values.

Now another quantity that I want to define is n which is say we have number of data points, can be anything 10, 20, 30 anything. It can be anything, the number of the experimental values**,** this is what that we are going to get. Now when can I say that this model will be perfectly fitting these experimental values when the error between this experimental value and the model value will be minimum.

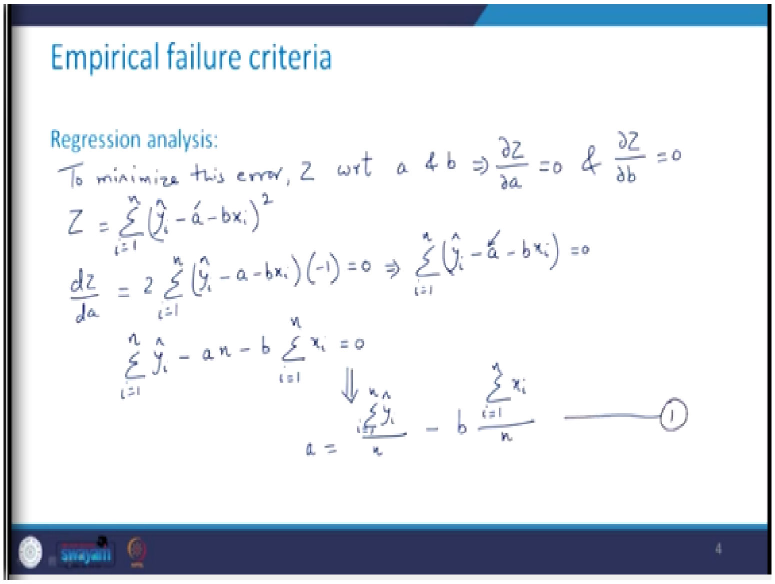
So, I define another term z, that is the error between experimental and the modeled value, which is defined as

$$z = \sum_{i=1}^{n} (\hat{y}_i - y_i)^2 = \sum_{i=1}^{n} (\hat{y}_i - (a + bx_i))^2$$

n is the number of data points. That here we have the experimental value, here we have the modeled value and the difference between the two, and take the square of that. So, that we are not bothered about the relative sign of this difference.

Either of the two can be larger or smaller than the other one. Now just substitute this modeled value. Now in order to minimize this error, our objective is to minimize this z, a and b these are the parameters of this model and in this case, it is the straight line. So, in order to minimize this from the simple mathematics you know that we have to take the differentiation of this z with respect to a and b and make them equal to 0. And that way we will be able to determine the parameters of this model which are a and b.

(Refer Slide Time: 08:58)



To minimize this error which is Z with respect to a and b, what is that? We need to do is

$$\frac{\partial z}{\partial a} = 0 \; and \; \frac{\partial z}{\partial b} = 0$$

So, what was the expression for Z?

$$z = \sum_{i=1}^{n} (\widehat{y_i} - (a + bx_i))^2$$

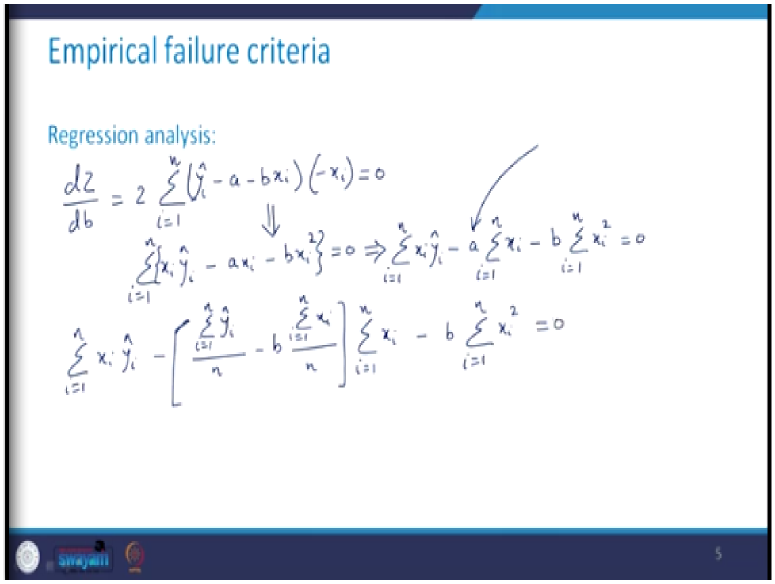
. Now just differentiate it with respect to a, so what we are going to get is,

$$\frac{dz}{da} = 2\sum_{i=1}^{n} (\widehat{y_i} - (a + bx_i))\,(-1) = 0 \Rightarrow \sum_{i=1}^{n} (\widehat{y_i} - a - bx_i) = 0$$

$$\sum_{i=1}^{n} \widehat{y_i} - an - b\sum_{i=1}^{n} x_i = 0$$

$$a = \frac{\sum_{i=1}^{n} \widehat{y_i}}{n} - b\,\frac{\sum_{i=1}^{n} x_i}{n}$$

(Refer Slide Time: 12:08)

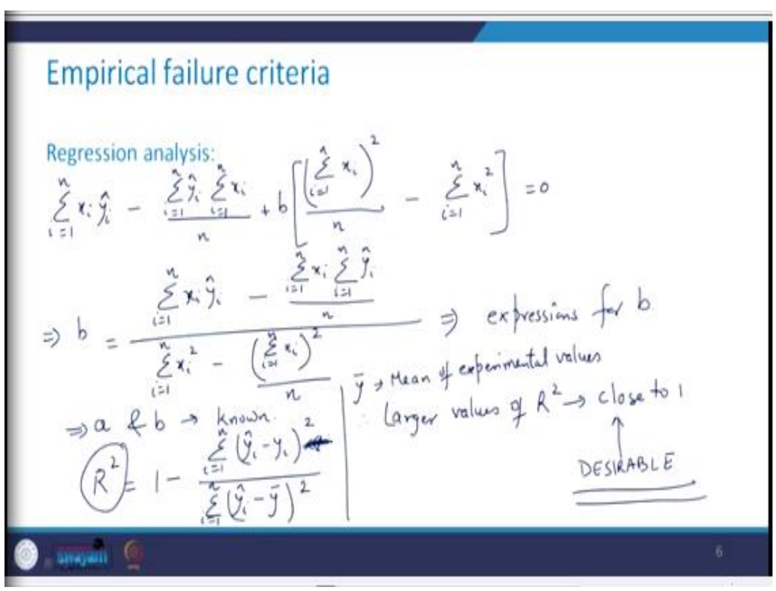


$$\frac{dz}{da} = 2\sum_{i=1}^{n}(\hat{y}_i - (a + bx_i))\,(-x_i) = 0$$

$$\sum_{i=1}^{n}\{x_i\hat{y}_i - ax_i - bx_i{}^2\} = 0 \Rightarrow \sum_{i=1}^{n}x_i\hat{y}_i - a\sum_{i=1}^{n}x_i - b\sum_{i=1}^{n}x_i{}^2 = 0$$

$$\sum_{i=1}^{n}x_i\hat{y}_i - \left[\frac{\sum_{i=1}^{n}\hat{y}_i}{n} - b\,\frac{\sum_{i=1}^{n}x_i}{n}\right]\sum_{i=1}^{n}x_i - b\sum_{i=1}^{n}x_i{}^2 = 0$$

(Refer Slide Time: 14:55)



$$\sum_{i=1}^{n}x_i\hat{y}_i - \frac{\sum_{i=1}^{n}\hat{y}_i\sum_{i=1}^{n}x_i}{n} + b\left[\frac{(\sum_{i=1}^{n}x_i)^2}{n} - \sum_{i=1}^{n}x_i{}^2\right] = 0$$

$$b = \frac{\sum_{i=1}^{n} x_i \hat{y}_i - \frac{\sum_{i=1}^{n} \hat{y}_i \sum_{i=1}^{n} x_i}{n}}{\frac{(\sum_{i=1}^{n} x_i)^2}{n} - \sum_{i=1}^{n} {x_i}^2}$$

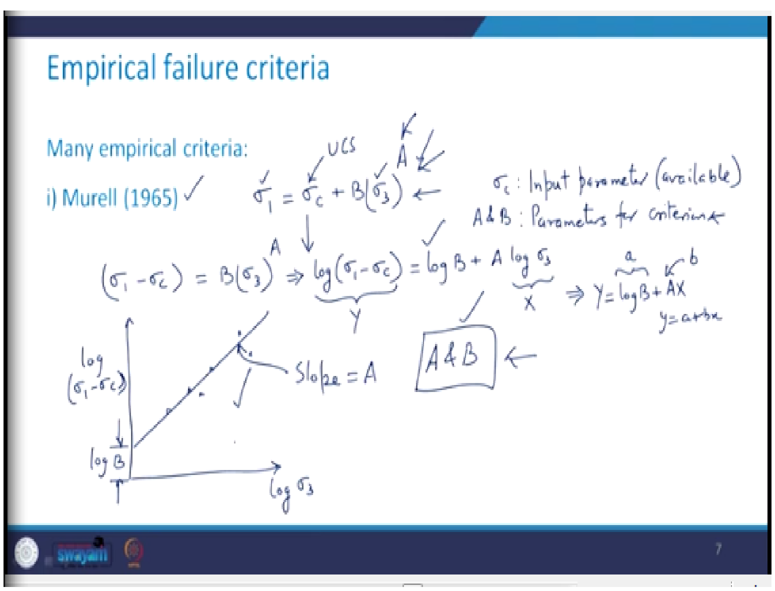
So, this is how we can obtain the expression for b. And for a, we have already obtained, so now we know a and b, they are known to us. Once it is known I would like to see that what exactly is the accuracy of that? So, for that I would define an index,

Now what is this $\bar{y}$? So, this $\bar{y}$ is the mean of the experimental values which you can just obtain by summing and dividing it by the number of data points. So, now what will happen? If this quantity is small, see if the numerator of this term is small then what will happen? This R^2 is going to be larger; smaller of the difference signifies that it is the better representation of the data by the model.

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

Therefore, we can say that larger values of R^2, what does that mean? That it will be close to 1, this is what is desirable. Because ultimately, we want this numerator to be minimum and accordingly, we have defined this R^2.

(Refer Slide Time: 19:21)



Now there are many empirical criteria which are applicable in case of the rocks and rock masses. Once we have learnt about the basic of the regression analysis, we can go ahead with empirical correlations or empirical failure criteria that are applicable in case of rocks and rock masses. I took a very simple example and try to show you that what do we mean by the regression analysis with the help of a straight line.

Most of the time the model that is being fitted into x and y, it may not be a straight line, how to handle such type of the model? Let us try to have a look, so the first one in this series which I am going to discuss with you is the criteria which was given by Murell in 1965, and he suggested that

$$\sigma_1 = \sigma_c + B\sigma_3^A$$

So, you see here the data which you have maybe the result of the triaxial data.

And now σ_c is given as the UCS of the intact rock that means that σ_c is going to be an input parameter, which is already known to me, so this is available to me. So, what are unknown parameters? These are A and B which are the parameters for the criterion. So, I need to obtain this, the question is earlier it was very simple y = a + bx and we could fit very nicely in a simplified manner that data into this straight line.

Now how to tackle this? This is not a straight-line equation, so we have to do some kind of mathematical jugglery, see how? I can write this equation as

$$\sigma_1 - \sigma_c = B\sigma_3^A$$

Now you take the log of this equation, that is on both the side of the equation you take the log and see what you will get,

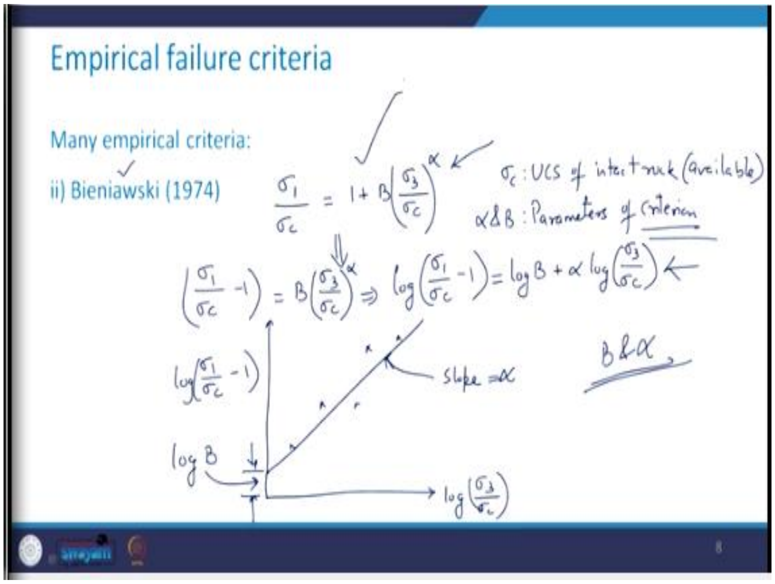
$$\log \sigma_1 - \sigma_c = \log B + A \log \sigma_3$$

Now just imagine that this quantity is Y and this quantity is X. So, what I will get from here that is going to be $Y = \log B + AX$. If I just compare it with the straight-line equation just that we discussed that was $y = a + bx$. So, if we try to plot the experimental data in this space, that is on x axis I plot $\log \sigma_3$ and on y axis I plot $\log \sigma_1 - \sigma_c$. And if I fit, say these are the points that we have and if I try to fit a straight line to this what this intercept is going to give me is $\log B$.

And it is slope will be giving me this parameter *A*. So, you see that this is how by converting this non-linear equation into a straight line in this fashion we could get the parameters *A* and *B*. Once I know these parameters, I can just substitute it here and we can get that for this data, this is the failure criterion. Again, here you can just convert it like let us say that you take this to be as *a* and this you can take as *b*.

And whatever are these steps that we followed earlier in order to fit a straight line, you follow the same steps for the regression analysis and get *A* and *B* or you can do that graphically in this particular fashion.

(Refer Slide Time: 25:09)



Let us move to the next empirical criteria which was given by Bieniawski in 1974 and this was given as

$$\frac{\sigma_1}{\sigma_c} = 1 + B\left(\frac{\sigma_3}{\sigma_c}\right)^{\alpha}$$

Again, here σ_c is the UCS of intact rock which is available to us, we can just give it as an input parameter. And here α and *B*, these are the parameters of criterion. So, again this is a nonlinear equation, what to do with this?

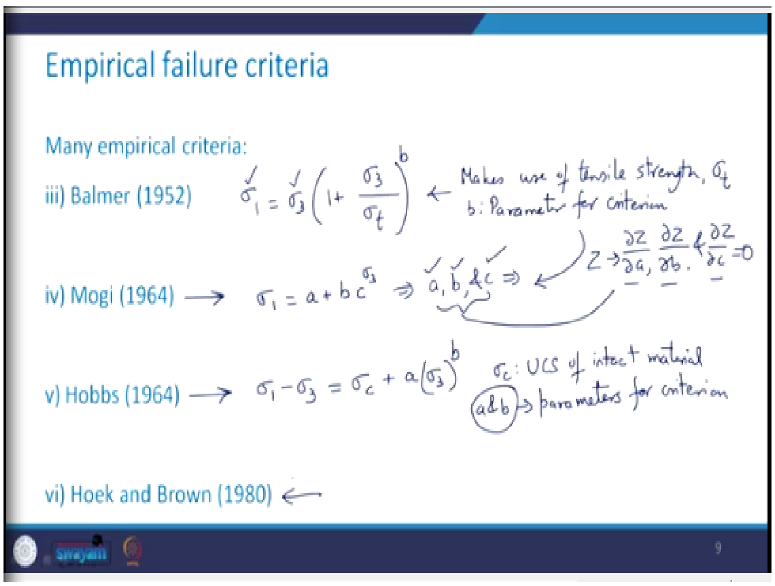
So, again we will have some kind of mathematical jugglery, take a look how?

$$\frac{\sigma_1}{\sigma_c} - 1 = B\left(\frac{\sigma_3}{\sigma_c}\right)^{\alpha} \Rightarrow \log\frac{\sigma_1}{\sigma_c} - 1 = \log B + \alpha \log\frac{\sigma_3}{\sigma_c}$$

So, now if I try to plot the experimental data in this space, that is on x axis we have $\log\frac{\sigma_3}{\sigma_c}$, and on y axis we have $\log\frac{\sigma_1}{\sigma_c} - 1$

And say these are the some of the data points; you just try to fit a curve to this, whatever is this intercept. That will give me the *log B* and slope of this straight line is going to give me α. So, once I know this B and α for the best fit line for this expression, this is going to give me the parameters for the criterion, substitute it here and you will get the empirical criteria.

(Refer Slide Time: 28:01)



Likewise, we have many empirical criteria, let us take a look at them one by one. So, the first one is the Balmer in this series, and it was given as

$$\sigma_1 = \sigma_3\left(1 + \frac{\sigma_3}{\sigma_t}\right)^b$$

So, you see that this criterion makes use of the tensile strength, that is σ_t. Again σ_1, σ_3, they are major and minor principal stresses, and in this case, b is the parameter for this criterion.

Next one is given by Mogi in 1964, and this is defined by the equation

$$\sigma_1 = a + bc^{\sigma_3}$$

So, here in this case a, b and c these are the parameters for criterion. So, in this case whenever you define the error, so in this case you will get $\partial z/\partial a$, and $\partial z/\partial b$ and $\partial z/\partial c$, all these three you have to equate to 0. So, you will have three equations, and there are these three unknowns, so you will be able to obtain these three parameters a, b and c.

Similarly, another one was given by Hobbs, that was given by this equation

$$\sigma_1 - \sigma_3 = \sigma_c + a(\sigma_3)^b$$

where this σ_c is again the UCS of intact material or intact rock, a and b these are the parameters for criterion. And we need to again do the mathematical jugglery in order to carry out the regression analysis.

And then we will be able to obtain these parameters a and b, in such a manner that the difference between experimental value and the modeled value is minimum for that dataset. This error maybe minimum but based upon the value of R^2 only, you need to decide whether this equation is fitting that data or not. It may happen that for a set of data, none of these equations may fit well R^2 may come out to be very low.

So, in that case you have to search for such a mathematical model which fits the data in the best possible manner, that means for which you get maximum R^2. The next empirical criteria in this category are the Hoek and Brown criterion which was given in 1980. And this is the criteria which is most commonly adopted in order to represent the failure criteria for rocks and rock masses.

So, in the next class we are going to take up this criterion in detail. It is equation, derivation, how to determine it is parameter? So, today we discussed about the basics of the regression analysis and with the help of a simple example of a straight line. I told you how to go ahead for the regression analysis, and how to minimize the error between experimental value and the modeled value?

Then with the help of some of these empirical criteria, I gave you the idea that how the regression analysis can be applicable in order to get the empirical criteria for the rocks and rock masses. So, the next one which is commonly adopted in case of these materials rocks and rock masses, we will discuss in the next class. Thank you very much.