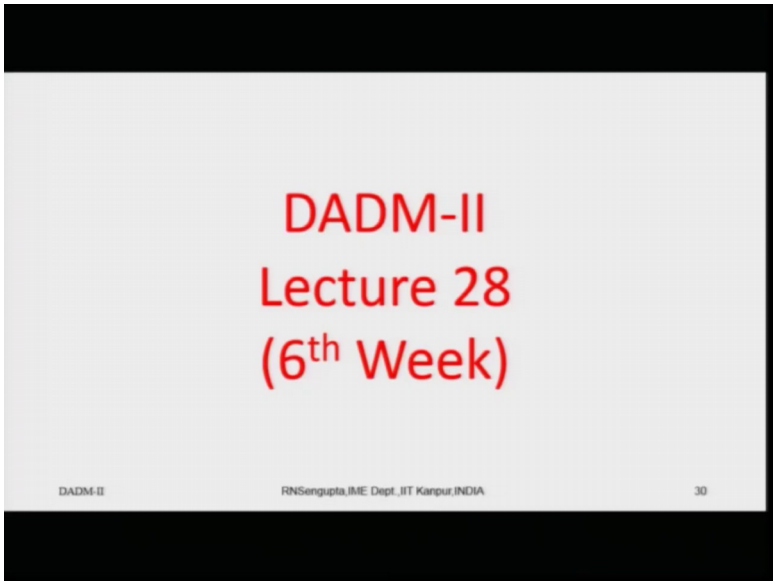
Data Analysis and Decision Making - II
Prof. Raghu Nandan Sengupta
Department of Industrial & Management Engineering
Indian Institute of Technology, Kanpur

Lecture - 28
ELECTRE

Welcome back my dear friends and dear students, a very good morning, good afternoon, good evening to all of you, wherever you are in part of this world. And this is the DADM II lecture under the NPTEL MOOC series. And as you know this total course duration hours wise is 30 which is for 12 weeks and the total number of lectures is 60, because each lecture is for half an hour.

(Refer Slide Time: 00:51)



And as you can see this DADM II which is Data Analysis and Decision Making II, we are in the 6th week which is the 28th lecture, and that means we have completed two lectures in this 6th week. and we are going to the 3rd lecture for this week. And my good name is Raghu Nandan Sengupta from IME department at IIT, Kanpur in India.

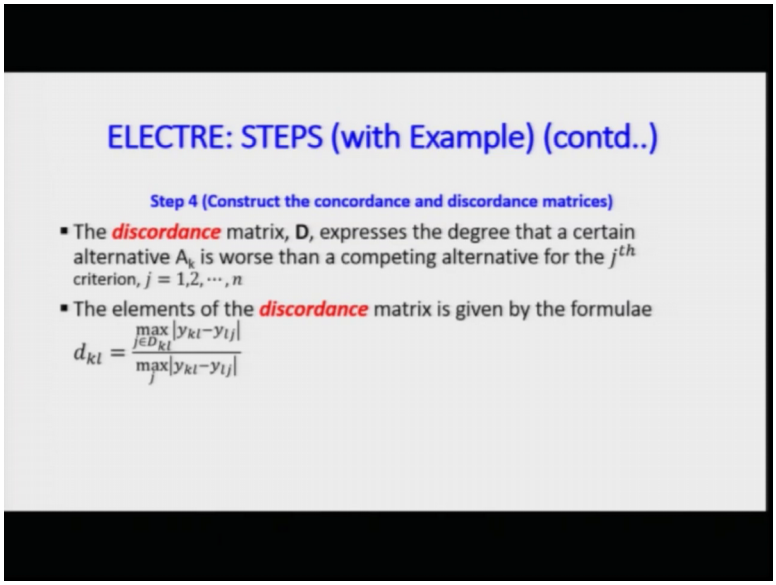
So, if you remember in the last class, we were just doing brief study of how ELECTRE process could be utilized without going to the logic of how the derivations work done. And if you remember, you are discussing that trying to collect the information of the concordance index and put them in the concordance set or in the concordance matrix

based on the fact, when we compare each and every alternative with the others based on each and every criteria, so we have already done that.

Also remember one important fact that the normalization initially which you have done was along the rows or the columns runs the whole steps would be exactly the same. And you will do now normalization based on the fact of the utility function, these points I keep repeating, so but please bear with me.

And the concordance fact on the discordance facts which you have done was based on the relative distance also; that means, on the blue line and the red line on the right hand side liking, hence in concordance set. In the left, it should be in the discordance set, but will have a different formula for that. Also remember the final C matrix is a asymmetric one, but the important fact is that the of the diagonal elements, when you add them up is basically one.

(Refer Slide Time: 02:46)



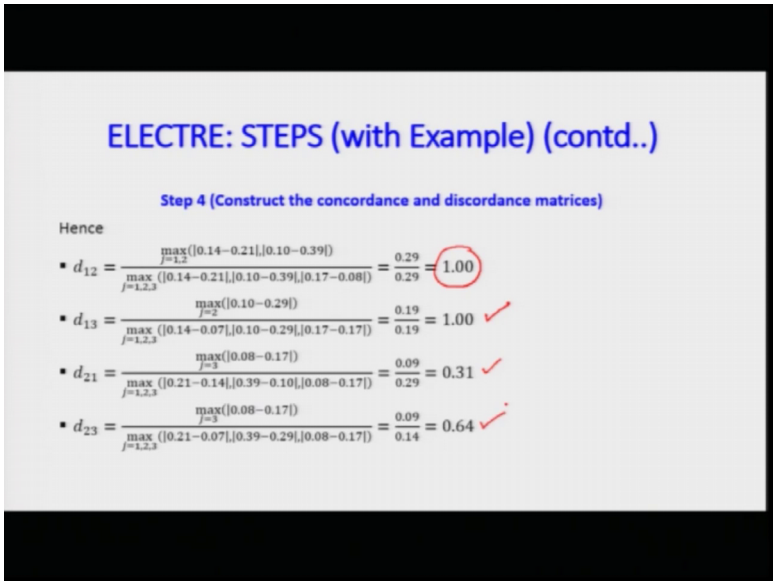
Now, continuing the step-4, we have already completed finding out the C matrix and now we will try to find out the D matrix based on the discordance indices or index values. The discordance matrix, D, expresses the degree that a certain alternative A k is worse than in the competing alternative; whatever the competing alternative is it can be A l based on the criteria's j is equal to 1 to n.

So, I will compare A 1 with A 1 based on all the criteria, then I will compare A 1 with respect to A 2 with for all the criteria, then I compare A 3, A 4, A 5 individually till A m with respect to A 1 based on each of this criteria taken as groups.

So, the elements of the discordance matrix, elements means small d 1 1, small d 1 2 so on and so forth along the first row. Similarly for the second row, third row and so on and so forth, would be given by the maximum values of levels which you have.

So, what I will do is that I will basically have a set, so d kl now, remember you are trying to basically compare the k and the l th one, based on k l and l th one the alternative based on the criteria and this j value will change from 1 to n. So, you will basically find out the maximum distance on the maximum value, for all the sets which you are comparing and then divide by the maximum so called combined of the whole, so called population which you have. I am using the over population in a very general sense. So, this ratios would definitely give you the values of the discordance set or matrix elements.

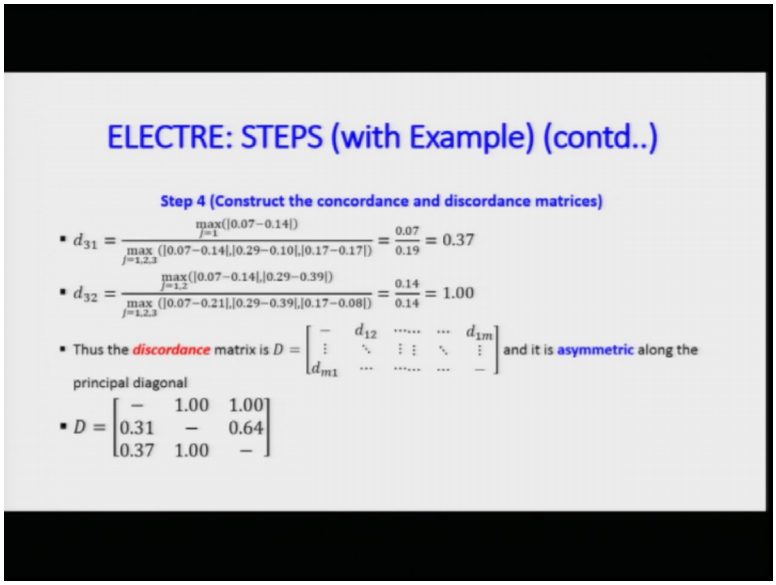
(Refer Slide Time: 04:21)



Now, again I will request if you remember; if you have your notes, when you are studying it, you have that y matrix, all those values were there. The 1 comma 1, 1 comma 2, 1 comma 3, similarly 2 comma 1, 2 comma 2, 2 comma 3, 3 comma 1, 3 comma 2, 3 comma 3, those values were there, you will keep repeating that utilizing that for finding out the C matrix and the D matrix.

So, the values of the d's are to be found out, so d 1 2 would be given when you are comparing the first and the second values and you will basically keep repeating it. And try to find out the maximum score in the numerator in the denominator, the value will come out as 1 that means, the maximum ratios is 1 similarly, for d 1 3 is 1. So, I am again repeating please have that set y and based on that you do the calculation. So, if you remember d 1 2, d 1 3 are the comparison which you are doing along the rows and the columns at each step. d 21 is 0.31, d 2 3 is 0.64, just pause have a look at these values.

(Refer Slide Time: 05:43)



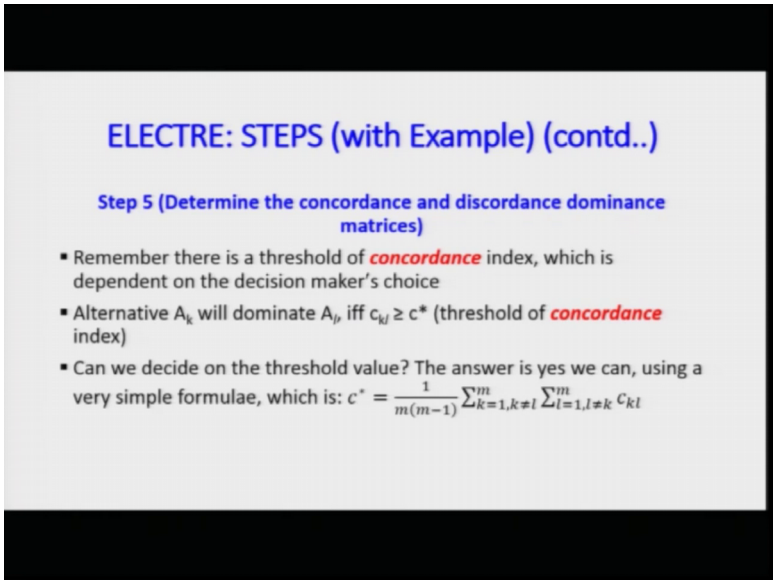
When I continue, I have d 3 1 and d 3 2. So, again in a similar way the values of C's which I have considered, the principal diagonal values were of no consequence. So, they were dash dash dash. In the similar way, when I do that for D so, I had done it for d 3 2 and d sorry, d 1 2 and d 1 3, then I had done for d 2 1 and d 2 3, similar I have done for d 3 1 and d 3 2 the values are given.

So, thus the discordance matrix D is also an asymmetric one along the principal diagonals and the values are given. Now, remember here the values need not add up to one, so if you have to seen that. So, it means when I am comparing the first with the second, the level of disliking, I am getting a score of one. But, if I am comparing 2 with respect to 1 that means, the first case 1 with respect to 2 is this means, I am taking 1 hence my level of so called dissatisfaction is 1. When I am comparing 2 to 1 and I am taking the decision 2, my level of disassociation is 3 1; 0.31; which technically means, if

I am trying to compare level of dissatisfaction, I am less dissatisfied by taking 2 than with respect to 1.

In the similar way, when I am taking the level of satisfaction, it was also taking one decision at a time trying to take the positive sense and then trying to compare. So, the discordance matrix D is given by the values, I will only read the values which is here dash, dash is basically the first 1 comma 1 would obviously be dash 1 1 0.31 dash 0.64 dash being for 2 comma 2 and the third row is 0.371 and dash and that dash being for 3 comma 3 cell.

(Refer Slide Time: 07:39)



Now, in the in the next step 5th, you have already found out the concordance indices based on that capital C, then already found out the discordance indices based on that capital D, now you want to compare them. Remember that there is a threshold or concordance index, which is dependent on decision makers choice, how good or bad or how strong the liking is.

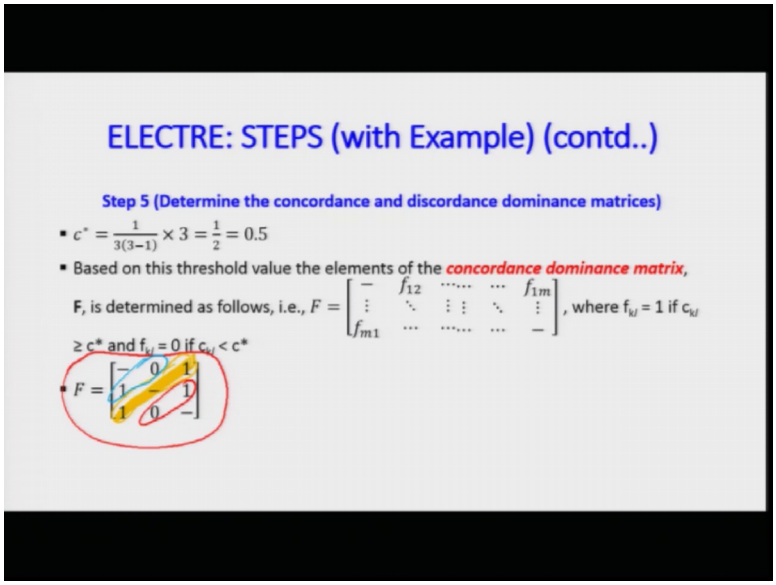
So, alternative A k will dominate A l if and only if, that value of A kl that means, when I am comparing these values of C, in that is capital C matrix are greater than some c star value which I have that means, if they cross that of the threshold, I am positive, I will take that decision.

Now, c star is calculated in the similar way as we do in the case of say for example, statistics we have the value of sigma hat square, hat is basically the estimated value. When you take the hat value you remember, that you will try to find out the best estimate for the standard deviation for the population using the standard error from the sample. Provided the sample mean is not known, hence you will replace mu by x bar n, I am talking about the normal distribution only.

And the overall formula if you remember, it will be the summation of the squares. So, what are the squared values? It will be x i minus x bar n whole square for each and every term as I changes from 1 to n. And in the denominator, you have divided by n minus 1, because that minus 1 is coming because you have lost 1 degree of freedom based on the fact that you have utilize the sample set from x 1 to x n to at least 1 time to find out the best estimate for the population mean which is the sample mean. Similarly, you will lose 2 degrees of freedom, 3 degrees of freedom, depending on the problem.

So, this is something of that sort, you will basically multiply by m into m minus 1 and add up all the values of C kl values of which are there, comparing the A k th and the A l th one based on each and every criteria, so that is c star. Greater than c star, I gave assigned the values, less than c star, I do not assign the values.

(Refer Slide Time: 09:56)



So, c star value comes out to be 0.5 of half based on that, I will basically have the concordance dominant matrix based on c, the values are if it is greater than c star or

assign a value of 1; it is less than c star, I assign a value of 0. So, the final dominant concordance mu matrix is dash, obviously it will remain dash for the 1 comma 1 cell, 2 comma 2 cell, 3 comma 3 cell. So, the values in the other 2 cells, when I am reading along the rows would be 0 1, 1 1 and 1 0.

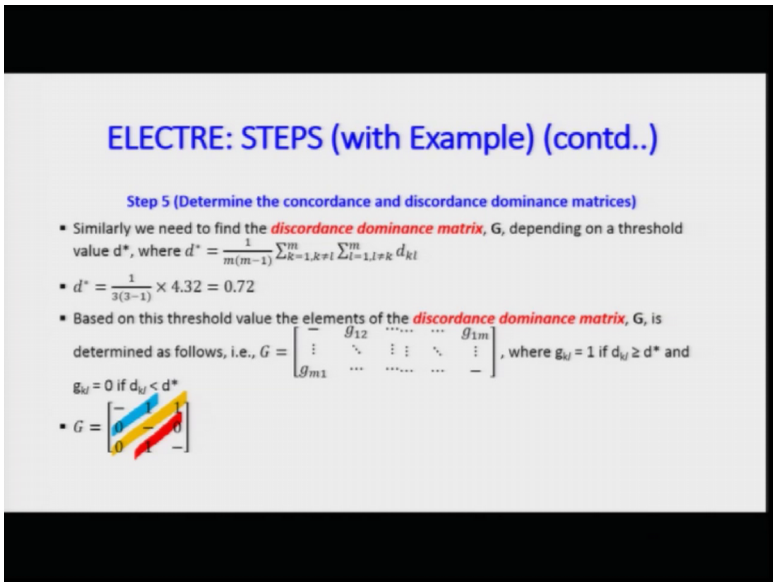
Now, what is interesting is the here? If you compare these or if you compare this or if you compare this, so now this is a concordance set liking. So, if I compare the first, so along the rows are the alternatives which I have, so there are three alternatives. So, this is alternative A 1, A 2, A 3 and along the top which is not written again the alternatives A 1, A 2, A 3.

So, when I am comparing A 1 2 with A 2, if I take A 1 I am positive, but my liking is not to that level, where I am very happy; I am happy, but not to that extent, because it does not cross that threshold value, threshold value will depend on the decision maker. But, when I am comparing C 2 to 1 sorry A 2 to A 1 that means, I am taking the reverse decision going to the other side of the fence. Then my overall satisfaction for taking A 2 is definitely positive, hence I gave n is half of the threshold value. And hence, I give a value of 1; that means, given the concordance concept without going to the discordance concept, I would always prefer 2 with respect to 1 based on all the criteria's which I have.

When I go to the yellow hashed line the orange one, so let me highlight this. So, first I had already discussed this, now I will discuss the orange one or dark yellow one. So, takings A 1 with respect to A 3 or A 3 with respect to A 1, I am always happy. In both the cases, I am always happy, hence I have a score of 1 1.

When I go to the case of comparing A 2 A 3, if I take A 2 other alternatives with respect to A 3, I am happy with A 2. If I take A 3 with respect to A 2 I am happy, but not to that x n well, it has come I am happy it does not give me any much extra value. So, this would give you on the level of concordance, dominance matrix, what are the values.

(Refer Slide Time: 13:08)



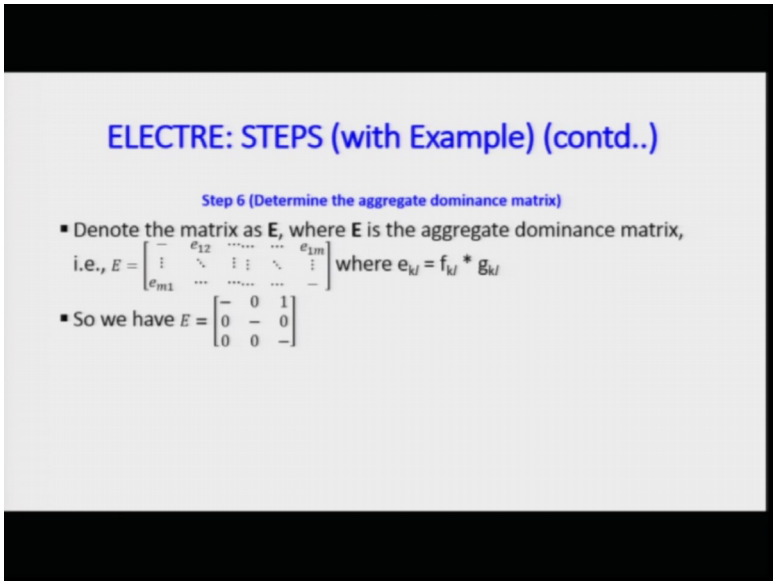
Now, similarly we should have for the discordance values also. Similarly, we find out the discordance threshold value d star same formula, summation of all the d k l values based on the fact that you are summing up for all the criteria. And again, the division value is m into m minus 1 exactly as I had mentioned few minutes back. d star value comes out to be 0.75, now I basically have the discordance dominance matrix, the concordance dominance matrix was F, discordance dominance matrix again I formulate based on the threshold.

Now, you see the values, I will just highlight them with the color. So, when I compare A 1 to A 2 and A 2 to A 1, A 1 values of discordance disliking, I am very sad for taking the A 1 decision. But, if I take the other side I go to A 2, I am sad not to that level as I was by taking A 1. When I go to compare say for example, A 3 and A 1, my level of unhappiness discordance or disliking is much higher if I take A 1 and not A 3. But, if I take A 3 and not A 1, I am unhappy, but my level of unhappiness is not to that degree.

When I go to A 2 A 3, my level on unhappiness is 0; that means, ok I have taken that I cannot do anything, when I compare A 2 to A 3. But, when I compare A 3 to A 2, I am definitely very unhappy, because I have discordance dominance matrix value as in that cell is 1. So, I have formulated F which is the concordance dominance matrix based on the threshold values being crossed. And similarly, I have the discordance dominance

matrix g based on the threshold value d star being crossed, in that case it was c star, now it is d star.

(Refer Slide Time: 15:20)



Now, when I compare, so when I am comparing and basically multiplying the values of corresponding cell values of f and g. Once I have f and g multiplying means, I am trying to find out, where I gain in both the sense. In the fact that the level of satisfaction is also positive cross that threshold value, less of dissatisfaction has also crossed that level value. So that means, if I am taking say for example, A k with respect to A l, it should be positive that I have taken A k, hence I am happy.

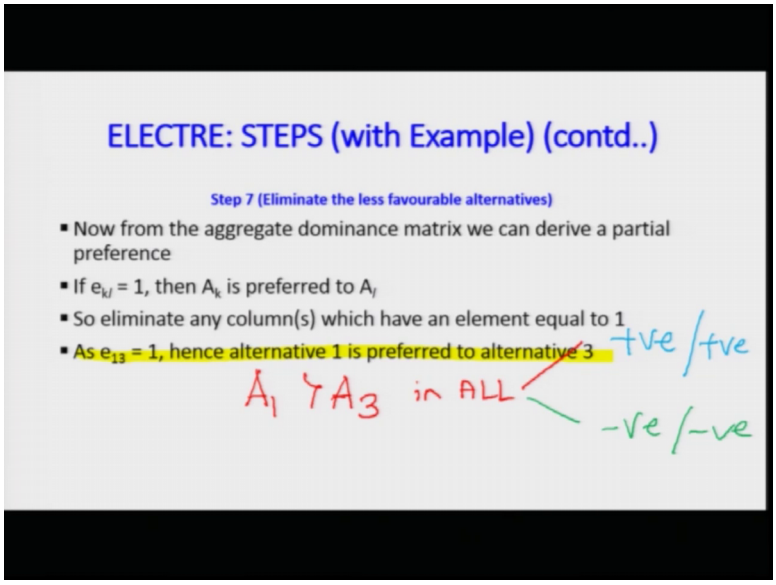
And if I take A l not A k, I am unhappy that means, in both the cases I am both positively positive and positively negative. Negative means in the negative sense, negatively negative, let me use that word negatively negative not the minus minus making plus negatively negative for taking the decision so hence, and because in this case they have crossed that thresholds hence, I am 1 1 multiplying 1 1 gives me a score of 1 that means, I am taking the right decision.

But, in some case, it may be that I am positive by taking the positive decision, but I am indecisive that I have taken the negative decision is not that negatives hence is 0. So, when I multiply 1 by 0, it is 0 that means, I am not doing the right justice to that. Similarly I reverse that, I have not taken that decision, I am positively 0 and negatively 1.

Then trying to multiply 0 and 1 also gives me a value of 0 that means, I am taking that decision is it is not accruing much value to me.

And obviously, if it is 0 0 in both the case, multiplication is 0. So, utilizing the concordance dominance matrix f and discordance dominance matrix g, when I find that value, the final cell is e which is the aggregated score. So, the aggregated score basically comes out to be only one, when I compare A 1 and A 3; that means, if I am comparing A 1 and A 3 in all the sense, I am making the best judgment. In other case, I am indecisive because there are conflicting criteria's, which gives me positive and negative benefits.

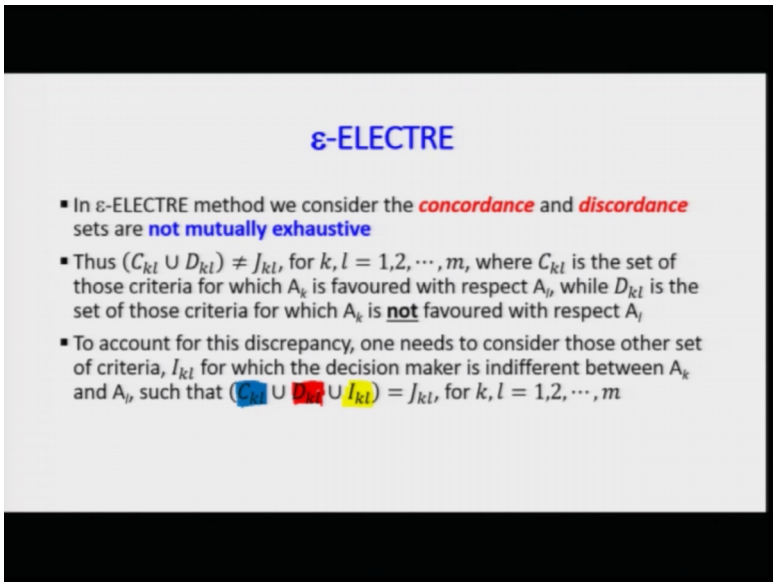
(Refer Slide Time: 17:38)



Now, for the aggregate dominance matrix, we can derive a partial preference concept. So, if that value of cell is 1, then A k would always be preferred to A l. So, we will eliminate any columns which have an; so eliminate any columns which have an element equal to 1. So, those would not be considered, because they are already positive. So, in case A e 13 is 1, hence alternative 1 is preferred to alternative 3 which I have mention.

So as, so in this case A 1 is preferred with respect to A 3 in all sense. In the sense, I will use the terms in the sense that positive was positive, positive was good hence I took. In the sense, I will write, I am using the green color. So, negative was also negative, so in score that means, it is not minus is plus. So, I take in the wrong decision, it is really good in the negative sense. I have taken the positive decision is really good in the positive sense.

(Refer Slide Time: 18:57)



Now, we will go into the concept of the epsilon ELECTRE method, so it is there is a slight change. And here I will I will I will talk about the story which I have already discussed, but I will still discuss in a brief detail. So, if we remember in the asymmetric loss function, when we are doing the utility function we said that, anyway I gave three examples, I repeat that to freshen up the concept, which I will be again doing.

So, if you remember, I have the asymmetric loss function giving me two weights; weights or the concept. One was the asymmetric loss function and one of the linear function loss function, when I combine them. So, it is e to the power a lambda minus a lambda minus 1, lambda will be is basically the difference between the predicted value and the actual value.

Now, if a is positive, the graphs I have already drawn. So, if required, I will draw it, but I will build up the story. So, if a is positive and the difference between the estimated value and the actual parameter value is also positive in that sense, you will find out that over estimation will dominate under estimation.

When I go into the second quadrant; obviously, in that case for the values of a positive, for negative values of the estimated value and the parameter value, you will find the linear part will dominate the exponential part. So, hence in the positive coordinate, exponential part is more positive and in the second coordinate, linear part is more positive; one means dominating sorry I should not use the word of a positive.

Now, when you take a as negative in that case in the first quadrant, the as the difference between the estimated value and the parameter value is positive, but in that case the linear part in the first quadrant will dominate the exponential part. When I go for a negative in the second quadrant, for the values of estimated value minus the parameter value is negative, in that case the exponential part will dominate the linear part.

So, the examples for the first case when a was positive. If you remember that overestimation would dominate underestimation was the case, when and in the case in the dominant in the sense loss function ways would be the case. When you are trying to basically formulate the problem for the electrical circuit, we are trying to basically stop the mission after the warranty time would be catastrophic.

So, if 6 months was there, rather than stopping it at the 5th month, if you have stopped at the 7th month, then the loss of life accident would be much more catastrophic if you had stopped it at the end of 5th month. So, hence it is better to underestimate.

In other case, for building the dam, when you basically have any less than 120 meters, if you or 120 feet if you remember, I had given this example 120 feet being the actual height. In one case, you build it to 122 and another case you build it to 118. In the case of 118 initially you use less manpower less material, obviously the cost is less.

But, when the flood comes, the propensity of the flood to bridge the dam is much higher, so hence they would be catastrophic environmental loss, manpower loss, cattle's would be lost, everything would be inundated. So, in that and but in the case when it if it is built to 122 feet, the overall cost initially would be high, but the propensity of the of the flood to breach the dam would be almost negligible, so in that case loss was would definitely will not there.

So, in this case underestimation will be more penalized than overestimation. And in the third and so in the overestimation is more penalized for the electrical one, underestimation more penalized for the civil engineering example. And in other case for the marketing one, when you are trying to float a product with the warranty life, whether if you basically make a warranty more or less than your competitors, obviously your loss will be calculated accordingly depending on the case that how over estimation or how underestimation have been penalized. So, it will definitely depend on the value of a is positive or negative depending on the problem which you have.

So, we are trying to utilize the same concept here. In epsilon ELECTRE method, we consider the concordance and discordance sets are not mutually exclusive. So, if you remember in the concordance and discordance sets that values of 1, 2, 3 which were coming out which is basically the alternatives, we are either clubbing them in the C set or in the D set. There were no case, where any one of the alternative was, you were un indecisive that whether you will go in to club it in the C set or in the D set, we are going to consider that now. So, hence we are considering that they may not be mutually exhaustive that means, addition should technically be the null set.

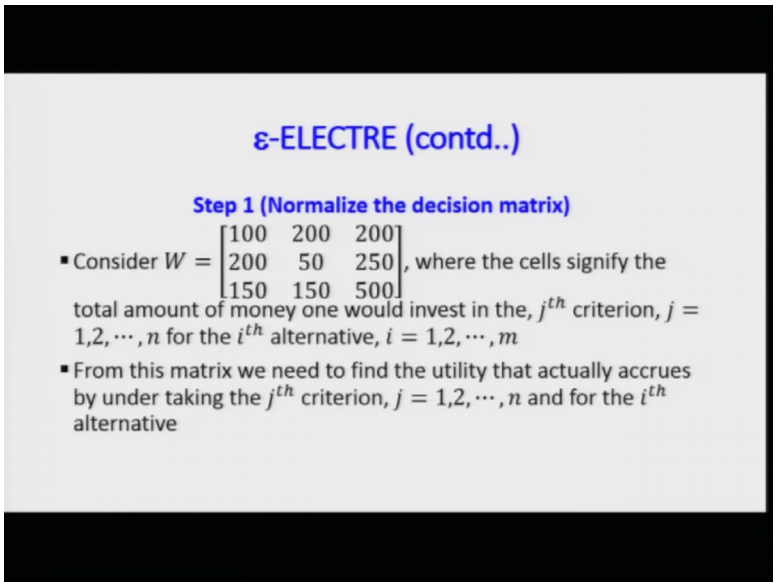
So, in this case, if you have C. C is the concordance factor, D is the discordance factor, if that is not equal to the total j set, so obviously now you will have three different sets, which are what the positive one I will mark it is blue. So, this is so concordance is C which is positive, I am marking with blue. Discordance is a red, because I do not like it hence it is red. And the other factor which I am indecisive, I am using the yellow color.

So now, you will basically have three sets C, D and I such that the totals union of D, C and I would basically with the total universal set which you have for all the sets of decisions. So, they again you have the luxury in a sense that you are able to classify them into three different sense. Again, the sum of all the elements in C, D and I would basically be the union universal set. And none of the element would be there, which basically can be in C or D or can be in C and I or can be in D and I. So, hence the intersection of C with D, C with I, and D with I would always be a null set.

So, I am going to consider each and every decisions accordingly and consider that that for all the sets of criteria's one at a time. Then again that means, A 1 will be compared with A 1 with respect to all the decisions or the criteria's, then A 1 will be compared to A 2 is based on all the criteria's, then I will compare A 3, A 4, A 5, A 6 till A m with respect to A 1 for all the criteria's.

Then I considered A 2, I will compare A 1 with A 2, A 2 to A 2, A 3 to A 2 till A m to A 2 based on all the criteria's. Then I go to this 3rd row, when I take A 3 and compare all the alternatives with A 3 with respect to all the criteria's. Then I do it for the 4th row, 5th row, 6th row till the m th row, when I take each and every alternative in that particular row and compare all the alternatives accordingly.

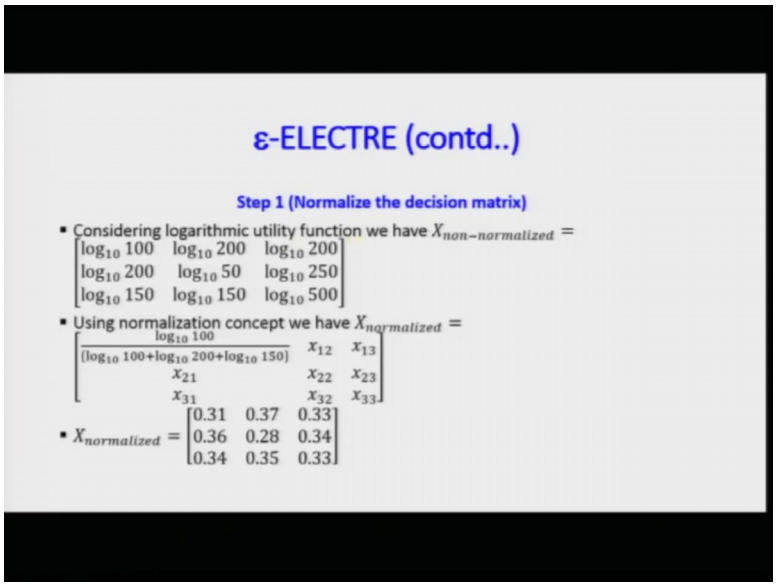
(Refer Slide Time: 26:13)



Consider this, I will change the problems accordingly with bring some twist in order to make you understand. So, consider the overall W is not the weightage, overall matrix which you have based on there is wealth, W is the wealth. So, if it was small W bold, it will be the weights. So, these values of the weights or the wealth which you have or the values, you are accruing for any particular decisions are given.

If I read along the rows are 100 200 200, second being 200 50 250, third being 150 150 500, these are thus values only without any units. Where the cells signify the total amount of money one would invest in the j th criteria based on whatever decisions which you have. From this matrix, we need to find the utility, then actually accrues by undertaking that decisions.

(Refer Slide Time: 27:26)



So now, we will basically consider some sort of a utility function in order to bring that the conclusion that utility factors do matter. Now, if you remember that in the first example, you have considered the utility function as quadratic. Now, I will just change the utility function as logarithmic. So, consider the logarithmic utility function, where we have I am for a simplicity sake, I am not going to consider the logarithmic with base e, I will basically consider base 10.

So, x being the non-normalize values, it would be log of 100, log of 200, log of 200, I am reading the first row. Then similarly, log are the values of 200 50 250, 150 150 to 500 will give me all the cell values for the second row and the third row. So, when I normalize, I am using the concept of some of the values being undertaken either for the column or the row and then trying to normalize either through the column or the row during the sum should be 1.

If I have that, the values which I have for the X normalized values are 0.31 0.37 0.33, similar 0.36 0.28 0.34, 0.34 0.35, 0.33. So, just pause do the calculations you will understand that I am using the logarithmic utility function. In order to convert the values in the logarithmic scale, scale means utility function and then normalizing in order to proceed for the next step.

So, with this I will end the 28th lecture and continue more discussion about this epsilon ELECTRE and try to highlight, how the epsilon ELECTRE would be a little bit different than this simple ELECTRE one method which you have considered.

So, have a nice day and thank you very much for your attention.