MCDM Technique using R

Prof. Gaurav Dixit
Department of Management Studies
Indian Institute of Technology – Roorkee

Lecture – 13
Introduction of TOPSIS Part-1

Welcome to the course MCDM Techniques using R. So in previous few lectures we have talked about ELECTRE before that we have been able to cover AHP. So now in this particular lecture we will start with another technique which is called TOPSIS. So let us start.

(Refer Slide Time: 00:42)

TOPSIS

- Stands for
  - 'Technique of Order Preference Similarity to the Ideal Solution'
  - Requires a minimal number of inputs
    - Subjective parameters
      - Criteria weights
  - Output is easy to understand

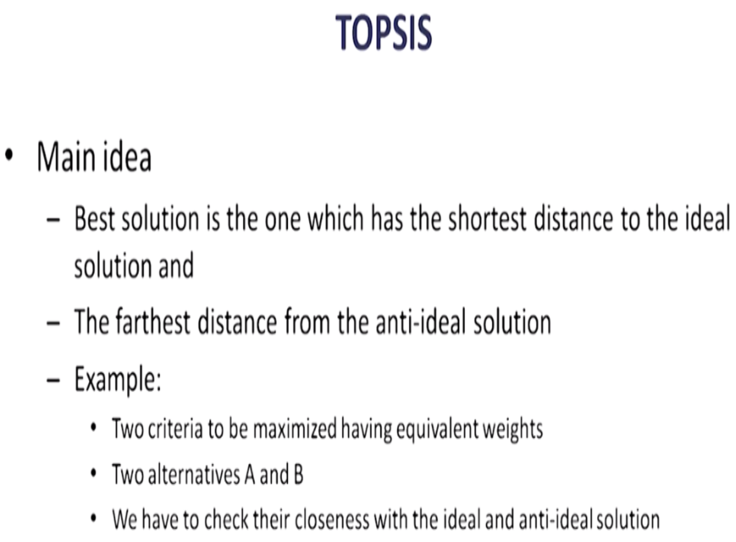
So TOPSIS actually stands for Technique of Order Preference Similarity to the Ideal Solutions. So the kind of method that it is there is some indication in the title in the full form itself. In this particular techniques TOPSIS in comparison to ELECTRE we require minimal number of inputs. So subjective parameters that we require is typically criteria weights and in some scenarios there could be other things as well.

However mainly it is you know criteria weights. If we compare it with the ELECTRE method then there apart from the criteria weights the decision-makers who were supposed provide their inputs for different parameters. So if we compare with the complexity of ELECTRE that was there the TOPSIS is quite a simple technique straight forward method. So that we will understand in more detail in this particular lecture.

So output is easy to understand so the way the whole procedure is done, the way the steps of

TOPSIS are executed that that we see is quite a straight forward easy to understand which was not the case with the ELECTRE. So let us move forward.

(Refer Slide Time: 02:05)

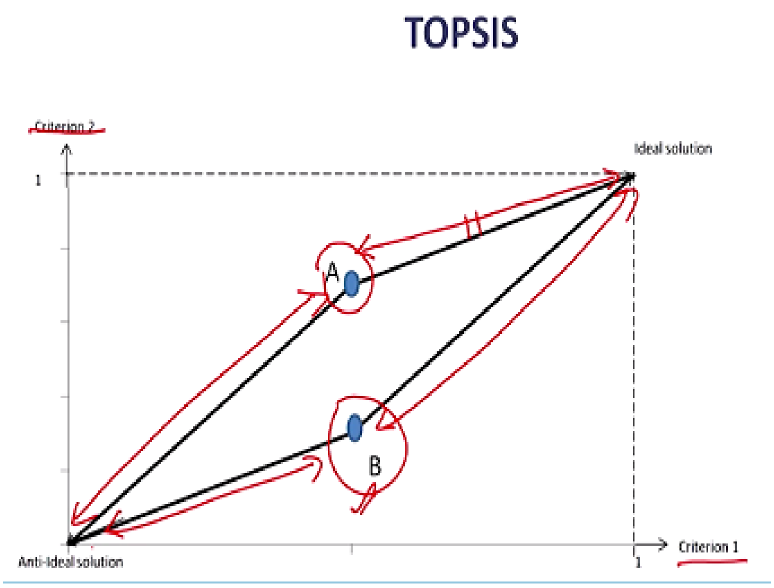


So main idea behind TOPSIS is that best solution is the one which has the shortest distance from the ideal solution. So as you would understand as we go along this lecture that this particular method TOPSIS is a distance based method. We will see those particular aspect in more detail later in this lecture, but the main idea is that you know we are looking for a solution the best solution which has the shortest distance to the ideal solution.

And farthest distance from the anti-ideal solution. So we will be identifying the ideal solution and anti-ideal solution and our best solution among all the alternatives that are going to be there is going to be closest to the ideal solution and going to be farthest from the anti-ideal solution. So that being the main idea and because of this kind of approach the TOPSIS is also referred as reference level approach.

So let us understand through example and some graphs, some charts. So let us take this example so we have 2 criteria to be maximized and they are having equivalent weights and we have 2 alternatives A and B. Now we have to check their closeness with respect to the ideal and anti-ideal solution.

(Refer Slide Time: 03:31)



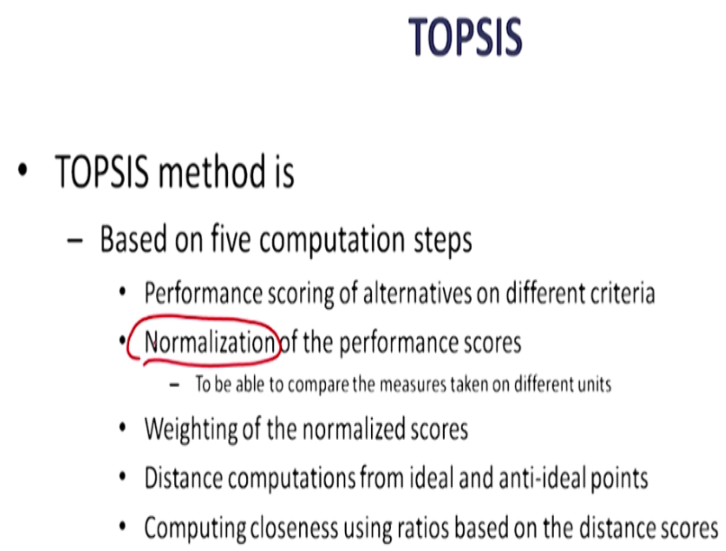
So we describe this particular example through a graphic. So you can see here we have 2 criterion one along the X axis criterion 1 and the another one along the Y axis that is criterion 2. So given these 2 criteria and 2 alternatives that we have so one alternative is this alternative A and the second one is alternative is alternative B and to find out which one of these 2 alternatives which one is the best one.

We have to compute the distance of you know each of these alternatives from the ideal solution as well as the anti-ideal solution. So you see this is the point alternative A and we have to see the distance from the ideal solution and then similarly for this one. So similarly we have to see the distance of these 2 alternatives from the anti-ideal solution so you can see from here.

So you would see that you know we are going to prefer the alternative which is going to be closer to the ideal solutions it seems to be alternative A. Just by looking at this particular graphics just by visual inspection we can see that the way these alternatives and everything else has been depicted or we can see that alternative A is closer to ideal solution and farthest from the anti-ideal solution.

While alternative B it is you know closer to anti-ideal solution and a little farther from ideal solution. So this is how the whole TOPSIS you know method is all about. So we will discuss other aspects of TOPSIS in more detail. So you know there are 5 main computation step that one has to perform to implement TOPSIS.

(Refer Slide Time: 05:26)



So first one is about performance scoring of alternatives on different criteria. So we need to have those scores so that we can you know later on perform the other steps. So performance scoring of alternatives with respect to each criteria so we are doing ranking of cars and there are certain criteria that are being considered for example price, for example efficiency, for example you know space.

So with respect to each of those criterion we have to get the scoring of alternative with respect to each of those criterion. So once that is there then we can move forward. Then the next step is normalization of the performance scores. So because these scores might have been taken might have been measured on you know different criteria different criteria which might be on different types of units.

So range of those you know range of those scales could be very different one could be in kilometers then another one could be in meters. So different kinds of units can be used and because of the numbers which are taken you know across those units could range from different low value to high value. So therefore to be able to compare the measures taken on different units we need to perform the normalization.

So this is the second step that is to be done then once the normalization has been completed then we do the weighting of the normalized scores. Now as we talked about that you know one of this subjective input that we require in TOPSIS is the weights of the criteria. So the decision makers they are required to specify the criteria weights that they would like to use for TOPSIS modeling.

So using those criteria weights will we try to compute, we are going to compute the normalized score weighted normalized scores. So that is a third step where we compute the weighted normalized scores. Then in the fourth step the distance computation from ideal and anti-ideal points are performance. So each of the alternatives you know if we have 3 alternatives so for each of those alternatives 1 alternative 2 and alternative 3.

So we will take them one-by-one and for the alternative 1 we will try to compute its distance from the ideal and anti-ideal solutions then we will take the alternative 2 and perform the same process. So these distance computations are to be performed for each of their alternative then once it is done then we can go ahead and compute the closeness using some ratio based expression formulas on the distance scores.

So once that this closeness ratio has been computed then we can move ahead and do our ranking of alternatives. So these are basic 5 steps. Now underlying Mathematics how exactly these computations are actually performed. So let us understand that particular aspect also.

(Refer Slide Time: 08:56)

TOPSIS

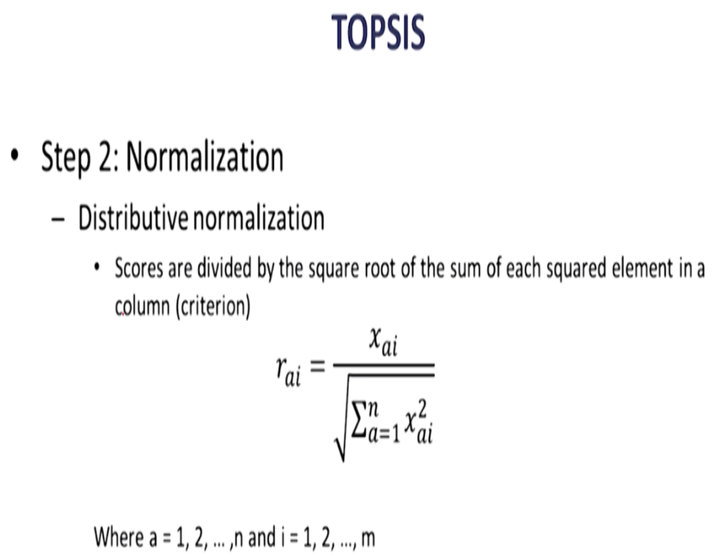
- Step 1:
  - Alternatives: a = 1, 2, ..., n
  - Criteria: i = 1, 2, ..., m
  - Decision matrix : $X = (x_{ai})$
    - Alternatives are shown on n rows and criteria are shown on m columns
    - Rank: n x m

So let us start with the step 1. So step 1 we assume that we have information on these 3 things just like other techniques AHP and ELECTRE that we have discussed before. We will be having alternatives so in this case you know to explain the Mathematics underlying TOPSIS will take alternatives as there are an alternative so we are indicating that as a= 1 to n. So there are n alternative and there are m criteria.

So i indicating that range 1 to 2 m then we should also have the decision matrix so that is being denoted using capital X and the elements of this particular matrix is being denoted as small x ai a is for alternatives which are going to be associated with rows and then i which is indicating the criteria. So this is going to be associated with the column side. So on the rows side we are going to have the alternatives in the matrix.

On the column side we are going to have the criteria in the decision matrix. So alternative are shown on n rows because we are considering n alternatives and criteria are shown on m columns because we are considering m criteria. So the rank is going to n cross m. So this is the information that we need to have to start with before we could actually apply TOPSIS modeling.

(Refer Slide Time: 10:31)



So as we talked about in the next step that is normalization. So there are different ways to perform the different procedures to perform the normalization steps. So we will talk about a few of them. So first one is distributive normalization so in this mode of normalization what we do is the scores that we already have the performance score that we already have in the decision matrix they are divided by the square root of the sum of each squared elements in a column.

So column is representing the criteria. So for each of the column therefore with respect to each of the criterion we do certain computations. So along one column that is along one criterion we will have the scores for each of those alternatives. So all those elements all those scores are going to be divided by this square root of sum of each squared elements. So

Mathematically it is expressed in the slide also as you can see.

Small r ai a is standing for alternative i is standing for criteria and column wise. So rai that is the normalized score that we will get is=in the numerator part we have the actual score actual performance score that is small xai and in the denominator we have the square root of summation of squared elements of all the scores along one particular column that is one particular criterion.

So as you can see here in the denominator we have square root then summation is from a=1 to n right. So because we are having n alternative. So all the scores for all the alternative they are being summated and squares of those scores are actually being summated and then square root is taken. So this is how we compute this distributive normalization and this is going to be used later on in the next step number 3.

Now there is another way to perform this normalization. So this is referred as ideal normalization. So again there are 2 scenario within ideal normalization. So if the criterion is to be maximized so we have talked about in previous techniques that there is typically going to be preference direction for a particular criterion. So whether the criterion is to be maximized or whether the criterion is to be minimized right.

So you know for example we talked about the ranking of car. So efficiency is something that we would like to maximize right, but you know price of the cars is something that we would like to minimize. So different criterion might have different preference direction so considering that our computation in this ideal normalization is going to be different.

(Refer Slide Time: 13:38)

TOPSIS

- Step 2: Normalization
  - Ideal normalization
    - If the criterion is to be maximized
      - Scores are divided by the highest value in each column (criterion)

$$r_{ai} = \frac{x_{ai}}{\max\limits_{a=1,2,\ldots,n} x_{ai}}$$

    - If the criterion is to be minimized
      - Scores are divided by the lowest value in each column (criterion)

$$r_{ai} = \frac{x_{ai}}{\min\limits_{a=1,2,\ldots,n} x_{ai}}$$

Where a = 1, 2, ... ,n and i = 1, 2, ..., m

So if the criterion is to be maximized then the scores are divided by the highest value in each column that is with respect to you know each criterion. So in this case the normalized score rai is going to be= in the numerator side xai/ maximum of xai where a is 1 to n. So out of n alternatives you know along a particular column we have to see the highest we have to pick the highest value and that highest value is going to be used as the denominator.

So and then we will get this ratio which is going to be used as this you know normalized score. Similarly, in the other scenario if the criterion is to be minimized then what we do we do the scores are divided by the lowest value in each column. So in each column that is with respect to each criterion we take all n alternatives and each of this scores with respect to belonging to each you know alternative they are divided by the minimum value out of all those n scores.

So as you can see here in the Mathematical expression here that normalized score small rai is being= xai in the numerator and then divided by the minimum value of xai where a is 1 to n. So in this fashion ideal normalization can be performed. So we talked about 2 types of normalization procedure. One we talked about distributive normalization where square root of summation of squared elements is done.

And in the ideal normalization where we divide by either the maximum score or the minimum score depending on the preference direction of the criterion. So there are other types of normalization procedures as well so that you can refer in your own time. So these are the 2 popular ones which are typically used. However, at this point I would like to tell you that the theory part that we are discussing here in the slides and the way we talked about how the normalization is done.

What is the exact expression Mathematical expression that can be used to you know compute to normalize this course. However, the way you know normalization is going to be implemented in R and the functions and the packages that are going to be available to us could be slightly different. So that depends on the author who has actually written those functions and the normalization procedure that they might have implemented there. So there could be slight variations of what we are discussing here.

(Refer Slide Time: 16:39)

TOPSIS

- Step 3: Weighted normalized decision matrix
  - Given: criteria weights, w_i and normalized scores r_{ai}

    Where a = 1, 2, ... ,n and i = 1, 2, ..., m

  - Obtain the weighted scores, v_{ai} as below:

So now we move to a step 3. So in this particular step you know previous step we have already computed the normalized score. Now in this step we are going to compute the weighted normalized score. So at the end of this step we will get the weighted normalized decision matrix. So we have the criteria weights wi so that are to be provided by the decision makers as subjective inputs for this particular model TOPSIS model.

And then normalized score rai that we have already computed. So weighted scores are nothing but simple multiplication of these 2 values. So criteria weights are going to be multiplied by these normalized scores. So vai that is the weighted normalized score is going to be wi multiplied with rai so we will get the weighted normalized score. So let us move forward.

(Refer Slide Time: 17:49)

TOPSIS

- Step 4: Distance computations from ideal and anti-ideal points
  - Weighted scores obtained in the previous step are used
    - To compare each alternative with the virtual ideal alternative
    - To compare each alternative with the virtual anti-ideal alternative
  - Define virtual alternatives: three ways
    - First Method
      - Construct virtual ideal alternative: $(v_1^b, v_2^b, ..., v_n^b)$
        - » Using best scores on each criterion
          - $max_a(v_{ai})$ if criterion i is to be maximized
          - $min_a(v_{ai})$ if criterion i is to be minimized
      - Construct virtual anti-ideal alternative $(v_1^w, v_2^w, ..., v_n^w)$
        - » Using worst scores on each criterion

Now we come to step 4 so there are few more details in this particular step. So what we do in this step is that distance computation from ideal and anti-ideal points so that is done. So main computation that we do is actually part of this particular step. So you know in this step weighted scores obtained in the previous step are used so weighted normalized scores we have already obtained in the step 3.

Now you know we are going to compare each alternative having these weighted normalized scores with the virtual ideal alternative. So there is no real ideal alternative so we are going to compute a we are going to define a virtual ideal alternative. Similarly, we are going to define a virtual anti-ideal alternative and each of the alternative that we have and the weighted normalized score that we have.

Those scores are going to be used to compare the alternatives with these virtual ideal alternative and virtual anti-ideal alternative. So how do we define these virtual alternatives. So we are going to cover 3 ways which can be used to define these virtual alternatives. So let us talk about the first method. So in first method you know constructing virtual ideal alternative would be like this.

So coordinates can be expressed in this form. So v1 and you can see you know superscript as b that is standing for the best value. And similarly v2b and up to vnb. So given the n alternatives that we have right. So we are going to pick the best value given the criteria that we might have we are going to pick the best value along all those criteria. So you can see here using best scores on each criterion.

So again there could be 2 scenarios so that is max a right so a is for alternatives so alternatives 1 to n. So max a so we are assuming that there are n alternative so a ranging from 1 to n and max a of vai. So v which we have already computed this is the weighted normalized scores. So if criterion i so is for the criterion i criterion i is to be maximized then we will take the max value.

Similarly, if the criterion i is to be minimized then we will take the minimum so that value is going to be minimum of a vai. So among all the alternatives so among all the scores for each of this alternative and with respect to a particular criterion. So we just look at one particular column so that would correspond to one particular criteria. So for that criterion we will see which alternative is having the maximum value or the minimum value.

So the criterion is to be maximized then we will pick the maximum value. If the criterion is to be minimized, then we will pick the minimum value. So once this is done then we use these best scores on each criterion to you know construct virtual ideal alternative. So remember if the criterion is to be minimized the best score is going to be the highest value. If the criterion is to be minimized if the criterion you know is to be minimized, then the best score is going to be the minimum value.

So therefore that has to be taken care of while constructing virtual ideal alternative. So once this is done then we will get this virtual ideal alternative denoted as v1b v2b up to vnb. Similarly, to construct virtual anti-ideal alternative we have to pick the worst value. So again 2 scenarios if the criterion i is to be maximized and since we are picking the worst value so we will take the minimum value.

And in the second scenario if criterion i is to be you know minimized ever since we are looking for the worst value we will take the maximum value. So it is just reverse of what we did in the construction of virtual ideal alternative. So in the virtual anti-ideal alternative will take the worst score and therefore we are denoting this virtual anti-ideal alternative as v1w and v2w up to vnw. So now let us move forward.

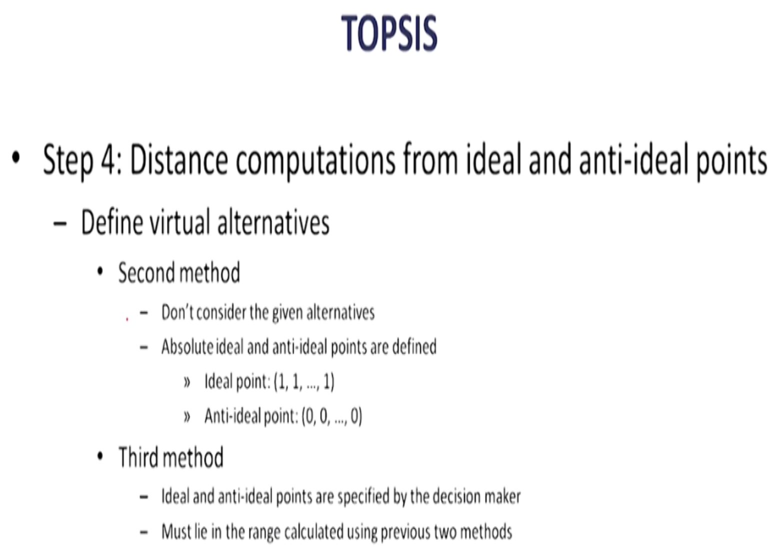
So that was you know one way to actually define the virtual alternative then there could be other ways to define alternatives. For examples you know one approach could be not considering the; these given alternatives. And the values and the scores that are there for each of those alternatives rather take a absolute position wherein absolute anti-ideal points are defined.

So in this case because you know later on as we will see that you know values that we compute is actually they are going to lie between 0 and 1. So why not take ideal points as 1, 1, 1 and the anti-ideal point as a 0, 0, 0. So irrespective of the actual values of the alternatives we will have the absolute ideal points and will have the absolute anti-ideal point and any alternatives that we might have.

They can actually be compared with these absolute ideal and absolute anti-ideal points. So this is another way to define virtual alternatives. Then what could be the third way. So in the third way you know as we talked about that criteria weights are anyway they are required to be specified as subjective inputs from decision makers. Similarly, for the virtual alternatives also we can take them from the decision make as subjective inputs.

However, we have to take care that these numbers that we take from you know decision maker this would lie within the range which are to be computed by the previous methods.

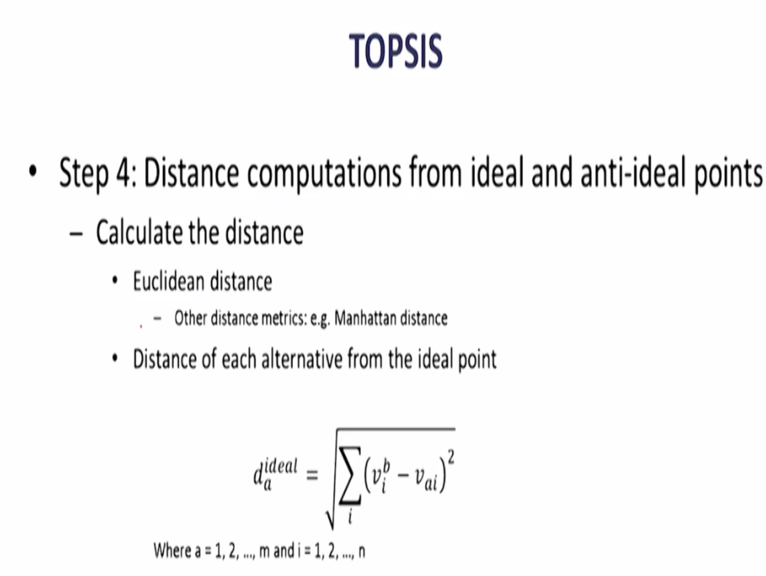
(Refer Slide Time: 24:44)



So if you see the range is typically going to be 0 and 1. So in that sense you know input provided by decision makers they should actually be lying in this particular range. So let us move forward so in this step 4 we talked about distance computation computations from ideal and anti-ideal points. So till now what we have talked about is how to define this ideal and anti-points we talked about 3 ways.

Now once we know these ideal and anti-ideal points we already have the information on the alternatives. So now we can go ahead and calculate the distance for each of these alternatives and between each of these alternatives and the ideal and anti-ideal points. So how do we compute the distance what is the matrix that should be used to perform this computation.

(Refer Slide Time: 25:40)



So the most popular distance matrix that is typically is the distance matrix Euclicdean distance matrix. There are other distance matrix Manhattan distance this is also used. So because Euclicdean distance being more popular so we will talk about this. So how do we use Euclidean distance to compute the distance between alternatives and ideal and anti-ideal point we can look at this situation here in the slide.

So small d and in the subscript a that is for alternatives that is 1, 2 to m and then ideal. So we are trying to compute the distance from the ideal point. So this is nothing but square root of summation from summation over on i where i is from 1 to n given that n alternative that we have and then within parenthesis we are having vib- vai and square of that. So this is very typical of you know Euclidean distance formula where you might remember that it is typically x1-x2 is square+ y1-y2 square and a square root of all this.

So similarly here also the similar kind of expressions have been used here. So with respect to ideal and anti-ideal points so we can use these 2 expression.

(Refer Slide Time: 27:08)

- Step 4: Distance computations from ideal and anti-ideal points
  - Calculate the distance
    - Distance of each alternative from the anti-ideal point

$$d_a^{anti} = \sqrt{\sum_i \left(v_i^w - v_{ai}\right)^2}$$

Where a = 1, 2, ..., m and i = 1, 2, ..., n

So for the anti-ideal point this is the expression da anti is going to be computed using this part square root of summation on i in parenthesis we have viw-via square. So using these 2 expressions we can always compute the distance of each alternative from ideal and anti-ideal points. So let us move forward. So at the end of step 4 we will have the distances from all the alternatives between all the alternatives and ideal and anti-ideal points. So once that is there we can go ahead and compute the closeness ratio for each alternative.

(Refer Slide Time: 28:00)

TOPSIS

- Step 5: Computing closeness ratio for each alternative

$$C_a = \frac{d_a^{anti}}{d_a^{ideal} + d_a^{anti}}$$

- C_a lies between 0 and 1, where a value approaching 1 indicates the preferred alternative

So how do we do this so you can see an expression written in the slides. So this closely ratio that is capital Ca is going to be computed for each of the alternatives. So Ca= in the numerator da anti so in the numerator we keep the distance that was computed with respect to you know anti-ideal points and in the denominator the summation of both the distance that is distance from the ideal points and you know distance from the anti-ideal point.

So you see here that you know given that we would like to pick the best alternative based on you know what we talked about that it is going to be closer to ideal points and it is going to be farthest from anti-ideal points. So if a alternative which is going to be the best alternative if it is closer to the ideal points. So therefore this distance da ideal is going to have a small value right. So you know because of this da ideal is going to be smaller value.

So once we keep in denominator and compute these ratio so this for this best alternative this Ca value is going to be closer to 1 right and you know the worst alternative this value is going to be closer to 0 because distance from the ideal point is going to be on the higher side and therefore denominator value would be on the higher side and therefore you know the overall value of closeness ratio will be closer to 0.

So this closeness ratio Ca the value is going to lie between 0 and 1 as you can see from the expression itself where as I talked a value approaching 1 indicates the preferred alternatives. So for the preferred alternative this value is going to be closer to 1. So for each of the alternative we are going to compute this closeness ratio and once we have these values for closeness ratio for each alternative then we can easily see which one is the best preferred.

And we can also produce a ranking right. So these are the 5 steps that we are supposed to perform to complete to do our TOPSIS modeling. So what we will do now is we will do an exercise in R to actually understand whatever we have discussed till now. So the steps we talked about the 5 important steps for this step was about getting all the information that we have you know the alternatives, the criteria and the performance table.

Then the second step was about you know doing normalization. So the scores are to be normalized because we would like to compare the performance on each alternative. So after that normalization we talked about the weighted normalization scores because different criteria might have different weights you know given the preferences of decision makers. So that is to be incorporated in the procedures itself so that is done in step number 3.

Then in the step 4 we define the ideal and anti-ideal points and once that is done we do the computation you know for each alternatives. So we do distance computations for each alternative with the ideal and anti-ideal points. So given you know that in the fifth step we actually compute the closeness ratio.

(Refer Slide Time: 31:44)

TOPSIS

- Open RStudio
  - Example: Ranking of cars
  - Three alternatives: "Corsa","Clio","Fiesta"
  - Four criteria: "Purchase Price","Economy", "Aesthetics","Boot Capacity"

So now in RStudio we will take this example you know rankings of cars and we are going to have 3 alternatives and 4 alternatives. So in this lecture we like to stop here and in the next lecture we will do our exercise in RStudio. Thank you.