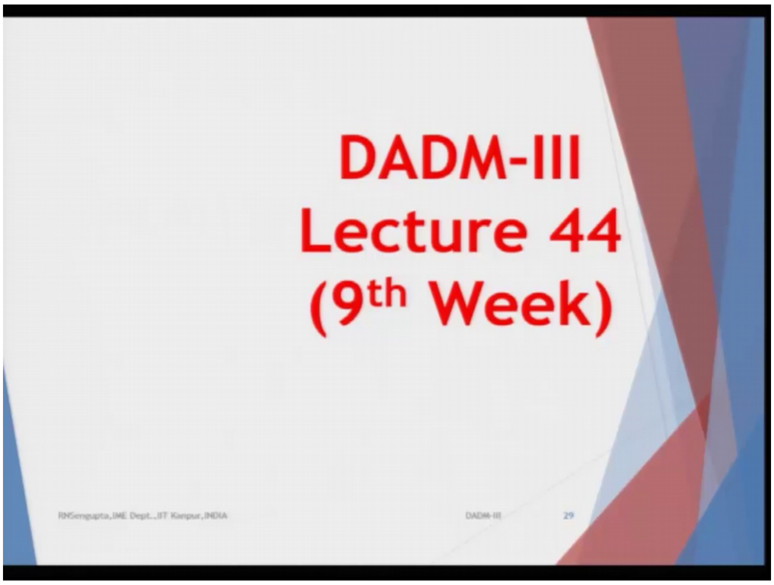
Data Analysis and Decision Making -3
Professor Raghu Nandan Sengupta
Department of Industrial and Management Engineering
Indian Institute of Technology, Kanpur
Lecture 44

Welcome back my dear friends, very good morning, good afternoon, good evening to all of you wherever you are in this part of the globe. And as you know this is DADM- 3 which is Data Analysis and Decision Making 3 courses under NPTEL mock series. And this total course duration is for 30 hours, contact hours, played over 12 weeks and number of lectures are 60 because each lecture is for half an hour.

(Refer Slide Time: 00:41)



And as you can see from the slide, this is the 9th week that means you have already completed 8 weeks, taken each week with 5 lectures and after the 8 weeks we have already taken 8 assignments. So in totality you will complete 12 assignments and then take the final examination.

Now, if you remember, we have gone into the realm of quadratic programming and I was describing an example, where there were many assumptions which you have skipped in order to make our story line a little bit simple. That means you have n number of stocks, what I have written I am just repeating it or whatever the investments are, this example being in the portfolio management only. And you have the returns, average returns that is found out using the concept of continuous compounding interest rate concepts.

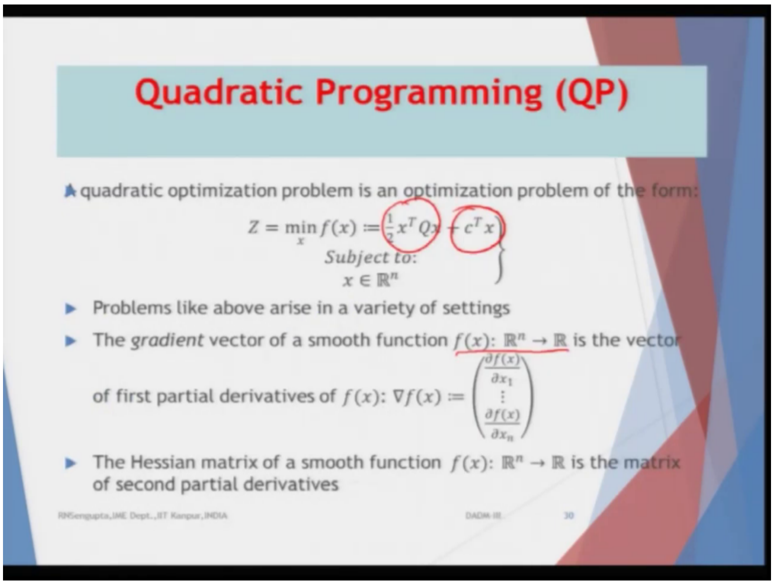
Those are not very important for our discussions here. Then you basically take the average return and multiply it by the corresponding total percentage of investment which you do in every stock. So that would submission of Wi into Ri bar. And if you remember this exactly matches by the linear component part which is Ct means the transpose function into x where x is the decision variable in our portfolio example case, Wi's are the decision variable. And the quadratic part was half of x into q x.

So I am not mentioning whether x is a transpose initial part, whether it is basically a vector one, a vector of values or row of values, which is basically a column vector or a row vector and q is basically a matrix of size m cross m. In our example for the portfolio management it would be n cross n. Now I also mention that how you will find out the quadratic component in the portfolio management. It was in the fag end of the 43rd lecture. Where you had basically the principle diagonals was basically given corresponding to the fact that you are trying to find out the quadratic component would be Wi square into sigma i square.

You add them up n number of times and the twice of the component is coming and you will divide by half is because the mirror image is along opposite sides of the principle diagonal, what basically the co-variances. Which was basically Wi into Wj into the value of sigma suffix Ij, where sigma suffix Ij it is estimated part obviously. They would be basically the co-variances between the ith and the jth stocks and if you are basically write those values of sigma ij for all the values of i and j from 1 to capital N then you will have the q matrix.

So let us basically further proceed in this direction, give you an idea and then come into the basic simple concept of mathematic. And if you remember, I did mention before when we started the quadratic programming, the concept about differentiation and why dy, dx would be 0 at the maximum and minimum point. We know that and what consequence whether if it is d2y dx2 was greater than 0 or less than 0, what consequence it has and what implication it has in the scalar concept and when it goes into this matrix concept. We have seen the concept over Hessian matrix, I am going to come to that within few minutes.

(Refer Slide Time: 4:29)



So this was basically the setup, you had the quadratic component part and the scalar part. So the problems like above arise in a variety of settings and I have given the example. Now the gradient vector, if it is gradient means dy dx function of a smooth function, when you have n number of such decision variables trying to find the f of x, where x is a vector and trying to basically take the function form and map into the scalar point, which is here.

So in that case the delta of the function, of dy dx as the function would be, where you divide that function, partial derivation remembers that because when you are finding out not divide, when you find out the derivatives it is partial derivative, because you are considering that the rate of change of that function with respect to, first with x1, keeping x2 to xn and being the number of variables as constant.

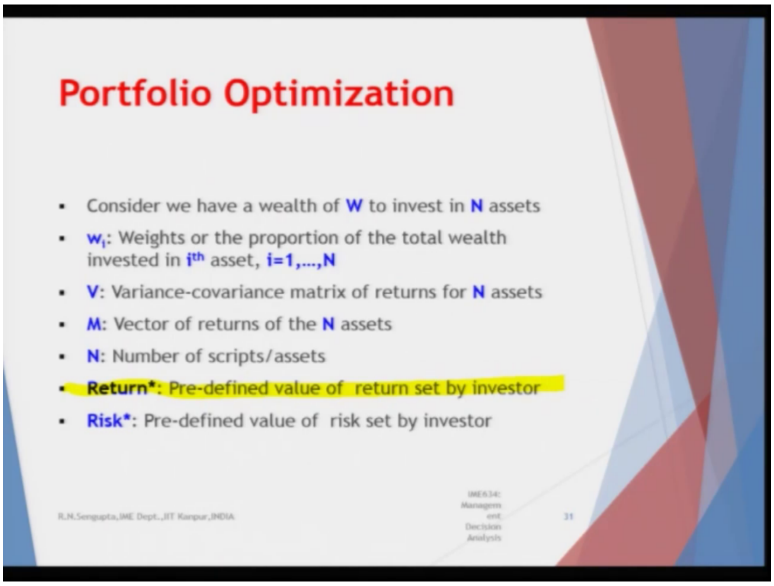
In the second element, which will be the second element in the column vector would be del f of del x2 considering x1, x3, x4 till xn as constant and finally the last value. In the column 1 would be del f by del xn, considering (x2) x1 to xn minus 1 as constant. Now, if you remember just few minutes back I mentioned about the Hessian matrix, the Hessian matrix of this function would basically be the mapping, from the n dimension to the r dimension 1, the one dimension where you are trying to basically find out the second derivate of this functions.

So obviously the second derivate would be the second partial derivative of the matrix and here it would be very simple like this. If you consider the second derivative, the corresponding principle diagonal would be del 2f del x1, the first element. The second one will be del 2f del x2 and the last element in the principle diagonal would be del 2f del xn.

While the off the diagonal elements, they would basically will consider they are similar way in the sense that the partial derivative of that function del 2f del x1 and then del x2 would be exactly equal to del 2f del x2 then del x1. That means trying out basically find out the partial derivative of that function first with x1 then 2 x2 would give you the same values that means the functions would basically be smooth and they would be symmetric in the sense the partial derivate, second derivate are same.

If it is not the obviously then you would face problems, so I am not going to nitty-gritties of the mathematics.

(Refer Slide Time: 07:22)



Now coming back to the example, first let me discuss the example, so we will consider a wealth of w total amount of wealth, so do not confuse this wealth with the concept of utility function wealth, obviously it would in the general scheme of things we go in depth and analyse that problem obviously that would be following the realm of utility function where you have the linear function first would be with respect to the expected value.

And then the second part which is second derivate which we for which we were discussing the quadratic concept would be coming from the concept of the variance of the utility functions, so we will skip that for our discussion. So consider we have a wealth of w and we want to basically invest in n assets, n assets I have already mentioned and we will consider the weights of these assets are given, weights means the amount of investment I am going to do, so out of the total amount of wealth capital W.

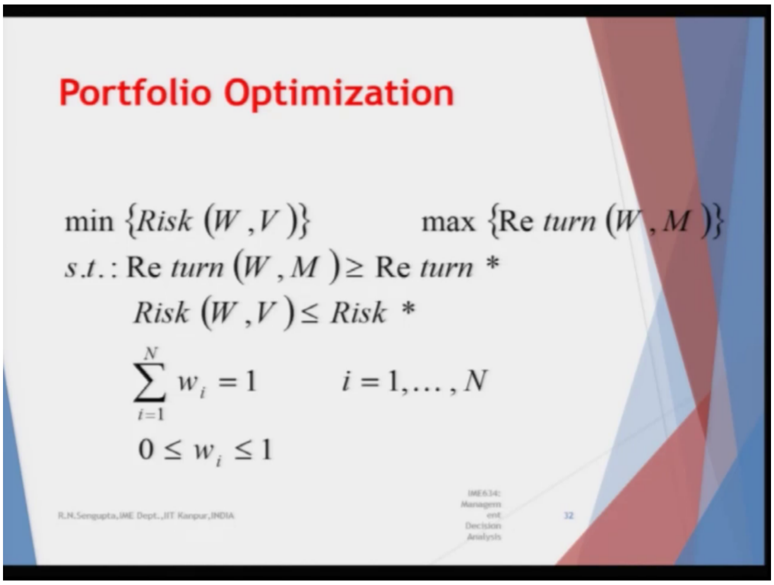
So, if it is 200, if I invest say for example 10 rupees, in the first asset it would 10 by 200 is the weight which is W1. Similarly, if I invest 50 rupees in the second one, it will be 50 by 200 will be W2. So obviously it would mean the sum of W1 to W capital N, would obviously be 1. So Wi the weights of the proportions of the total wealth invested in the ith asset and I is equal to 1 to N and capital V is the variance, co-variance matrix which I have just mentioned few minutes back.

Variance, co-variance matrix means the principle diagonal, I am repeating it please bear with mean, the principle diagonal are the variances and the off diagonal element are the co-variances. So V would be variance co-variance matrix of the returns of these n assets. While M, we will discuss as the vector returns of the n asset. So if you remember the vector I have mentioning as r1 bar, r2 bar, till rn bar and they want the best estimate corresponding to the mean values which was Mu 1 to Mu n. So this basically vector M corresponds to r1 bar to r-bar n. And n as I mention again, I am again mentioning is the number of scripts.

Now, what we want to do in our problem this is interesting. What we do want to do in our problem is that, we want to restrict the returns of the total portfolio to be greater than equal to some return star value, which is fixed beforehand. Let me number it, it will be easier or highlight it. Which will this one, so it is predefined by the investor so greater than type would basically give you some idea, about the slack and the surplus, which we would consider.

Obviously we would not be going through that discussion again but if we remember that if we had only taken C into X the linear part, then obviously it would be linear programming problem we could have solved it using the linear concept and then the concept of slack and surplus and the artificial variables would have been considered. And in case, if there is a predefined risk value, obviously we will want the risk value of the total portfolio to be less that equal to risk star. So again here the concept of the slack and surplus would be added.

(Refer Slide Time: 10:41)



So our problem now I have written both the parts, I have not given a plus sigh in between as it was trying to basically minimize half of x into q into x plus c into x, that is the quadratic part and the linear part. So the first part, the minimization part is the risk which is the quadratic function, it is the square term. Which is something to do with the variance and the second part is the return which is basically the linear part, which is with respect to the returns.

So you will basically have some function from corresponding to the weights and the number of stocks. In the second case, the risk would be corresponding to the weights and the variances (sorry) and the returns would be something to do with the functional form of the weights and the vector of the returns, M.

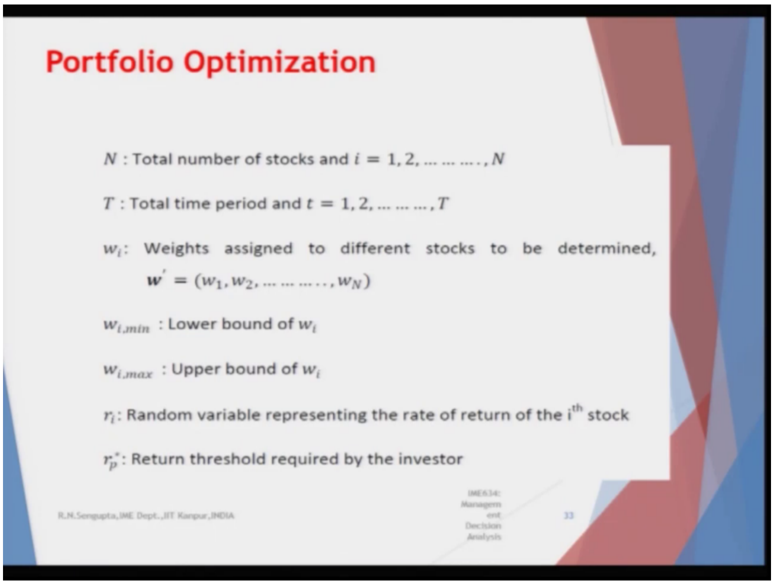
Now in this case as I mentioned if there are number of constrains, but one set of constraints would be greater than equal to the return star which is fixed beforehand, another set of constraints would be or 1 constraint would be that the risk is less than equal to risk star, which is predefined. And obviously you will have the ways restricted between 0 and 1. Now remember one thing, this 0 and 1 means that I am basically allowing all the values to be applicable between 0 and 1 which is basically the continuous case. It can be also converted into a discrete case very simply.

And finally and very logically the sum of the weight should be equal to 1, that is for Wi for all I is equal to 1 to capital is 1. Now you could basically formulate the problem in this way, rather than considering the wealth, why do not you consider say for example number of stocks you are going to buy. So considering the price of the stocks and the number of stocks you will buy, that will give you the total amount of investment which you are going to do for each and every asset.

And the total amount of investment which you will do in the total portfolio would be the sum of all the investment which you are going to in each and every asset. Which means, I would not write but I will just explain. Say for example, if n suffix i is the number of assets which you are going to buy for the ith stock. And if say for example the corresponding price is s suffix i, then the amount of investment which you are going to do for the first asset would be n1 into s1.

Similarly for the second one would be n into s2, so and so forth till the last one which will be, at this n which I am using, consider is the small n, small n suffix capital n, which is the number of stocks you are going to buy for the nth stock and s suffix capital N is basically the price of the nth stock. And sum of that is basically the total wealth W which you have.

(Refer Slide time: 13:43)



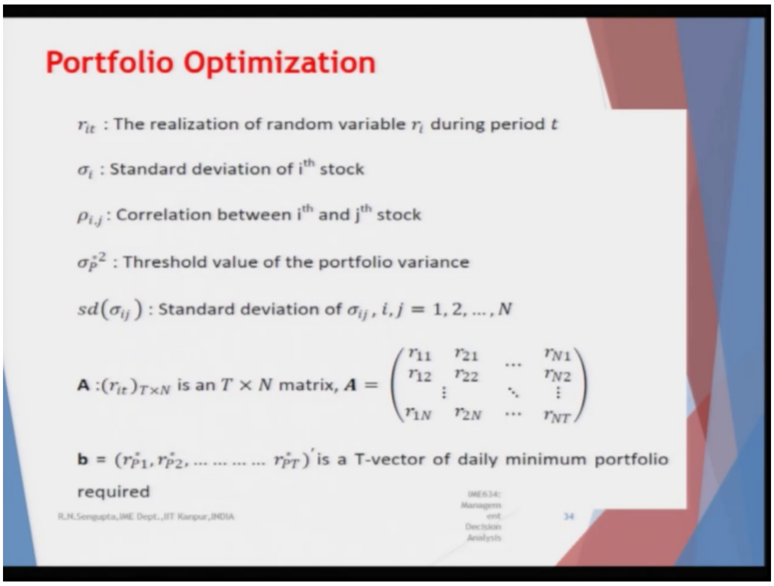
Here you could also have and then this explanation is given, the total number of stocks are given, i is equal to 1 to n, T is the total time period, so why I am bringing the total time period I am going to come within 2 minutes and Wi are the weights assigned to the different stocks which had to be determined.

So the vector, which is the vector which I mentioned that x value in the quadratic programming that x is basically now W transpose or W it is a row vector or column vector depending on how you have formulated the problem. But remember the size of v, which you have, V is the variance co-variance matrix is capital N cross capital N. So, obviously the matrix multiplication along with the column and the row vector should be applicable.

Now, these are some simple variation which I have brought into the problem. As I mentioned that the total weight of each and every stock is between 0 and 1, it may so happen that you want to restrict the weights between some maximum value, and it has to be greater than some minimum value. So in that case it will be W1, would be restricted by W1 minimum and W1 maximum, similarly for W2, W3 till W capital N.

And as I mentioned Ri's are the random variables pertaining to the returns from the ith stock. And this Rp star which was basically the return star which we have just discussed few minutes back is basically given by the investor himself or depending on the market conditions.

(Refer Slide Time: 15:18)



The other values which were important for us to understand is because if you remember I have brought the concept of time. Why I have brought the concept of time? I am going to come to that within few minutes. So, we will have rit the realization of the random variable which is the return of the stock of the Tth time and sigma suffix i is the standard duration corresponding to the ith stock or the ith script.

Rho Ij as I mentioned is basically the co-relation coefficient existing between ith and the jth stock or the jth script and obviously if we have rho ij, then row ij multiplied by sigma i multiplied by sigma j will give you the covariance of the ith and the jth one which are the off the diagonal elements in that matrix V.

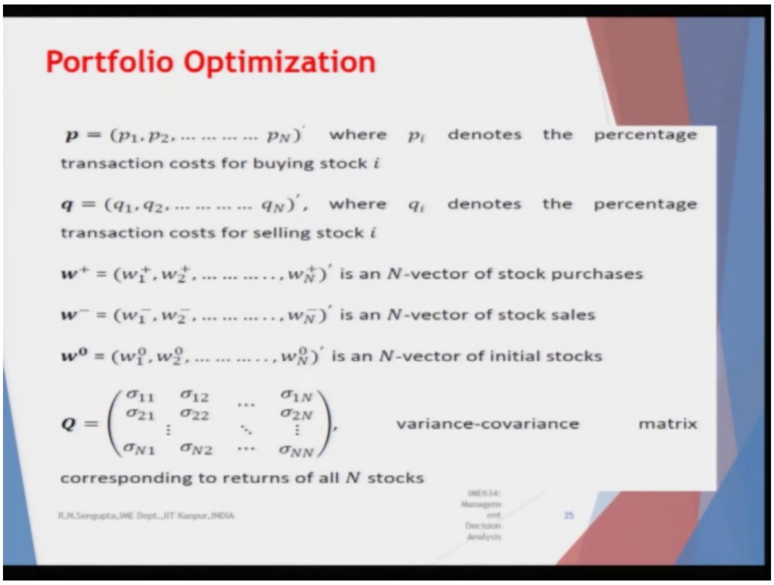
Here also if you remember we have said that there is the risk star, which is determined by the investor, so here it is what I am writing sigma square star suffix p, p is basically for the portfolio, the threshold value of the portfolio variance which is given and (here) another thing which is important for us to note, that in many of the cases we would be interested to find out how does the variance standard division of the variance also fluctuate?

I will come to that later on, but that has nothing to do with the quadratic programming concept and A would basically be matrix corresponding to rit, that means for r1, which is the return for the first stock, there would be values r11, I am talking of when I say 11, 12, 13 all numbers as I have already mentioned are the suffix values. So you have r11, r12, r13 corresponding to the first stock and different time periods of 1, 2, 3, 4, the time periods can be either seconds, can be minutes whatever it is.

And B would basically be the vector corresponding to the return of the portfolios, the star values which have been fixed by the investor at different points of time. The first period, second period, third period, fourth period based on the fact that when you are trying to basically find out the returns for the ith stock, at each and every moment.

So if you are trying to find out the return of the stock and as well as the return of the portfolio you want to fix up, it can be say for example, after the end of the day trading for Monday, Tuesday, Wednesday, Thursday, Friday then Saturday, Sunday is closed again you start from Monday.

(Refer Slide Time: 17:49)

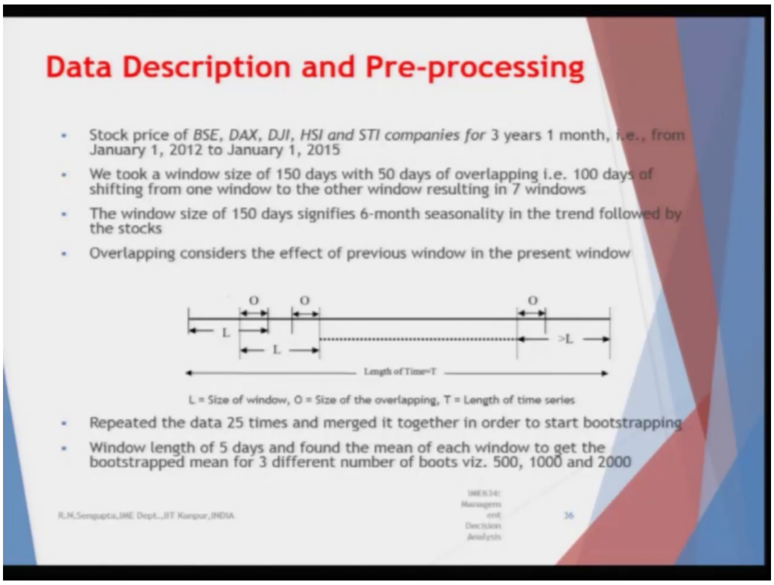


Here also few other important concepts which would be important, later on we will see would be p is the Pi who denotes the percentage transaction cost for buying and selling. So these are the transaction cost. So how the transaction cost is coming, it has something to do with the concept if you would remember, the cost of trying to transport goods per unit from city 1 to city 2 that means from the warehouses to the distributers.

So when you are buying and selling stocks, that means when you are buying dollars from the banks, selling dollars from the bank, or when you go for a loan, take a loan or basically deposit money in the bank, obviously there are different interest rates. So we are going to bring that into the picture by this concept of p, so this p q are all bold which means they are vectors. So p is the vector where pi denotes the percentage transaction cost for buying a stock, while q corresponding to all the values of q1 to capital qn is which denotes the percentage transaction cost for selling the stocks.

And if I have the n vectors or stock (purchases) purchased, so obviously any increase and decrease in the percentage terms, from W1 changes by, say for example 0.01 percentage, so obviously it would increase or decrease depending on whether the investment is changing from time period to time period. And here if you remember, the value of p is basically q, the variance covariance matrix, I am repeating time again please bear with me. So, q is variance covariance matrix corresponding to the return of the n stocks.

(Refer Slide Time: 19:37)



So rather than immediately going into the data I will come to solutions later on. So, why I have considered the rit's was basically I am taking whole time frame of readings, so say for example am taking day wise, so it is 1st of January 2019, 2nd of January 2019 so on and so forth, I have considered 240 days trading days, so I would consider that there is per week or 2 week I consider as a length of my time frame of consideration.

I find out the return of the stocks each day for the 7 days or 15 days or 2 weeks and then I find out the average, so that average I am going to consider as r1 bar, r1 bar,1 for the first seven number of days trading. So again when I do the second window of averages so it will be r1 bar, 2 so r1 basically means the first stock only, the next suffix 1, 2, 3 ,4 considering the bar values are there, they would be for the time periods.

So in general what we will do is that we will consider there are some overlaps, so this symbol over means overlap, so if I consider the time frame of length l then again a length l later on but there is an overlap of o. So obviously it would mean that I am trying to basically normalize the readings in such a way what the overlap values would try to, it is not going to eliminate, but it is trying to going to decrease the fluctuations in the prices and the returns which are there.

So L is the size of the window, O is the size of the overlap and T is the length of the time series which we have taken. So with this let me read what we have done, we take stock prices BSE, DAX Dow Jones industrial average, DAX mean from Germany, Hong Kong stock exchange and Singapore stock exchange. For all the companies who are trading in stock that not all the companies mean, that is how the index is defined.

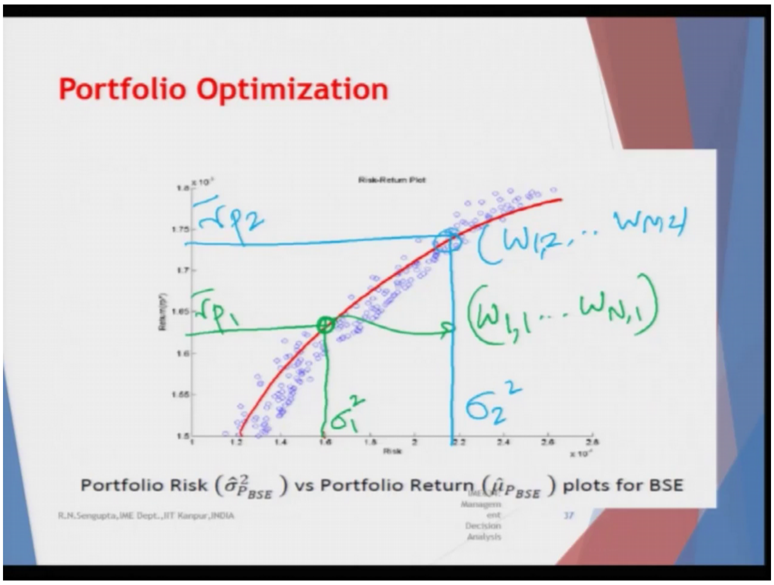
For 3 years and 1 month that is from January 1^{st}2012 to January 1^{st}2015 so total numbers of years and one month considering all the overlaps which are there. We took a window size of 150 days with 50, so this length L is 150, so what I was mentioning few minutes back, that one week or two weeks is basically a length, which is 150 days. And we take an overlap which is O of 50 days of overlapping, that is 100 days of 15 windows is console, you take 150 shift it by 100, again take that 150 shift it by 100 and go out accordingly.

So, window size of 150 signifies 6 months seasonal in the time trade, so 6 into 30 is 180 considering that Saturday, Sunday and the holidays we generally have 150 days, not exactly 150 but it is an approximation. The window size of 150 signifies 6 months seasonal in the trade followed by the stocks and the overlapping considers the effect of previous windows and the present window. So there is some spill over, not the effect, but in the sense that they overlap in order to basically continue the trend whatever you have for the previous number of days.

You repeat the data 25 times and merge it together in order to start the bootstrapping. Bootstrapping concept I would not be going into details, bootstrapping is basically you take the data and increase it proportionally in order to basically get the probability distribution as such. So window length of 5 days and we also find out the mean for each window depending on so we can take 150 and the overlaps can also be changed.

And we do three different type of bootstrap that mean we increase on concatenate the data, 500 times, 1000 and 2000 number of times to get, try to basically expand the sample and make it more nearer to the population in order to find out. The sample statistics is the best replica of the population parameters which you want to find.

(Refer Slide Time: 24:14)



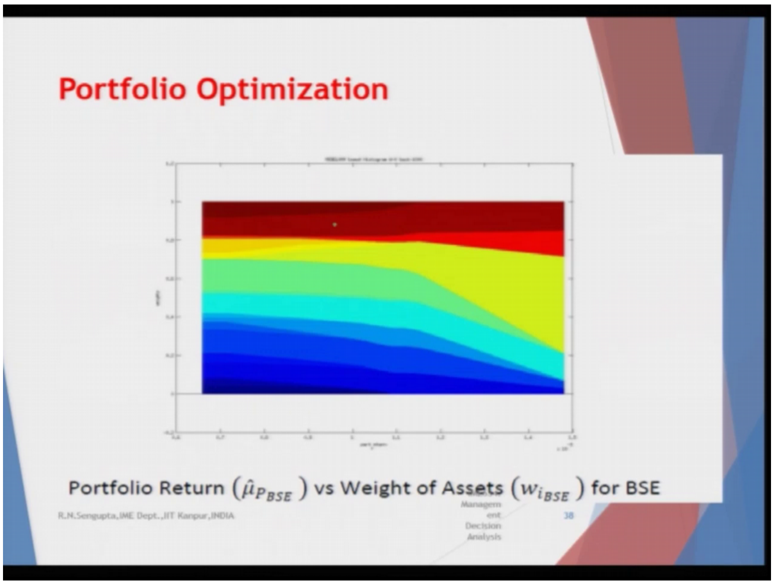
So once before coming to the actual solution techniques. So, here if you solve the quadratic programming, considering that both of the objective functions are applicable and there is a multi-objective case, now remember the concept of multi-objective we had discussed the mythologies of genetic algorithm, artificial immune system then calling the optimisation all those things in DADM-2.

So if you consider the simulation runs of about 200 in number. So you want to plot the quadratic function which was the risk and the linear function which was the return, linear function we basically plot along the Y-axis and quadratic function we basically plot along the X-axis. So it will give you such, if there are straight lines, there is no variation it will give you some straight lines like this. So what you do is that you basically go along the vertical axis, join there this will give you the portfolio corresponding to risk value of, say for example sigma 1 square, return of rp1 and from here you will get W1,1 means for the first case, similarly Wn,1.

Then again if I do it for (blue colour can I do, yes) and do it for the second case, so this is sigma 2, 2 square, this is rp2, this is bar remember, so we have W1,2 Wn2. You can do it accordingly, so portfolio risk versus portfolio return are plotted for BSE and you can find it out accordingly and do the runs as required. So if you have only plotted the returns then you would basically have the returns and the portfolio values. So you are basically trying to maximize the return that will give you some set of weights.

Which would not exactly match here then if we find out the values of the risk only trying to minimize in the quadratic case you will have only the risks only and they would different sets of weights corresponding to that.

(Refer Slide Time: 27:16)



Now, if I want to, I change, so as I am traversing the graph, I am basically changing the values of rp star which is the investor define and the sigma square p star, as I change the values, I slowly build up the frontier as I plotted and for each and every point the weights are changing, if I plot the weights for each and every point.

So you will basically have for all the stocks, so these colouring schemes would be for all the stocks. So you see the yellow one for the BSE and the price is change, prices obviously change which means the percent investment for each and every stock will change. So with this I will end the 44th lecture and discuss more about the quadratic programming in the 45th lecture. Thank you very much and have a nice day.