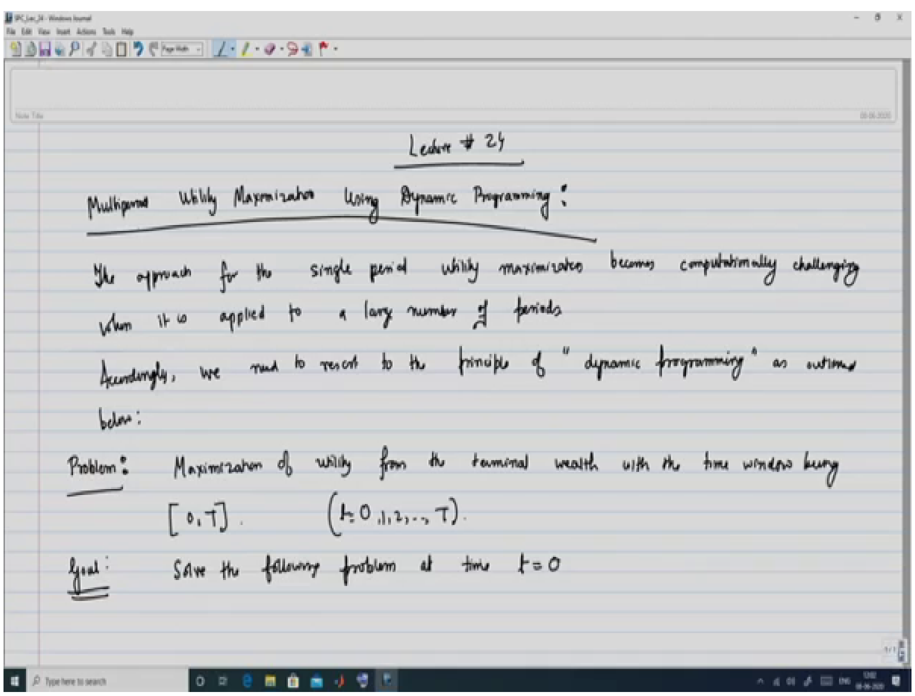
Mathematical Portfolio Theory

Prof. Siddhartha Pratim Chakrabarty
Department of Mathematics
Indian Institute of Technology Guwahati

Module 05: Optimal Portfolio and Consumption
Lecture 03: Optional portfolio for multi-period discrete time model; Discrete Dynamic Programm

Hello viewers, welcome to this lecture on the NPTEL MOOC course on Mathematical Portfolio Theory. You recall that in the previous two lectures we looked at the problem of determining an Optimal Portfolio and Consumption, and we talked about the notion of consumption in the paradigm of utility function. And then we talked about a single period utilized maximization of utilization utility of the final wealth, and we used what is known as the a dynamic programming approach to determine what is going to be the optimal portfolio in case of a single step model. So, in today's class, we will extend that notion into a multi-step model, and we are going to apply the dynamic programming principle in order to determine what is going to be the dynamically obtained optimal portfolio for a given utility function.

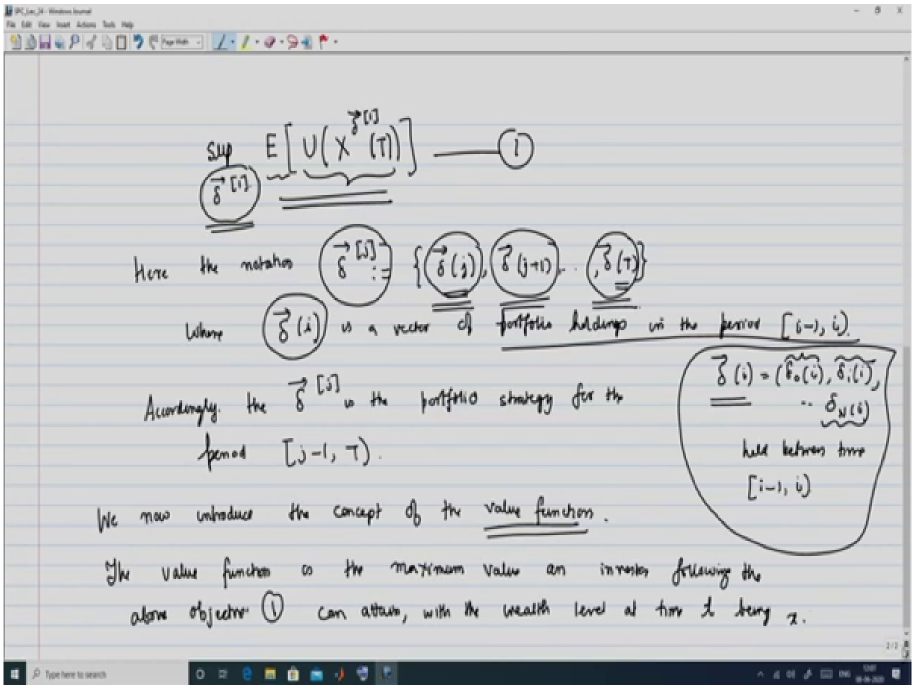
(Refer Slide Time: 01:26)



So, accordingly we start our lecture on this topic of multi-period utility maximization using dynamic programming. So, the approach for the single period utility maximization becomes computationally challenging when it is applied to a large number of periods. So, accordingly we need to resort to the principle of dynamic programming as outlined below. So, the problem statement goes as follows that we want the maximization of utility from the terminal wealth that is the wealth at the final time point capital T with the time window for the investment being $[0, T]$. And we do this maximization at time point t is equal to

$0, 1, 2, \cdots, T$. So, in a more mathematically tractable formulation, the goal is to solve the following problem at time $t = 0$.

(Refer Slide Time: 04:31)

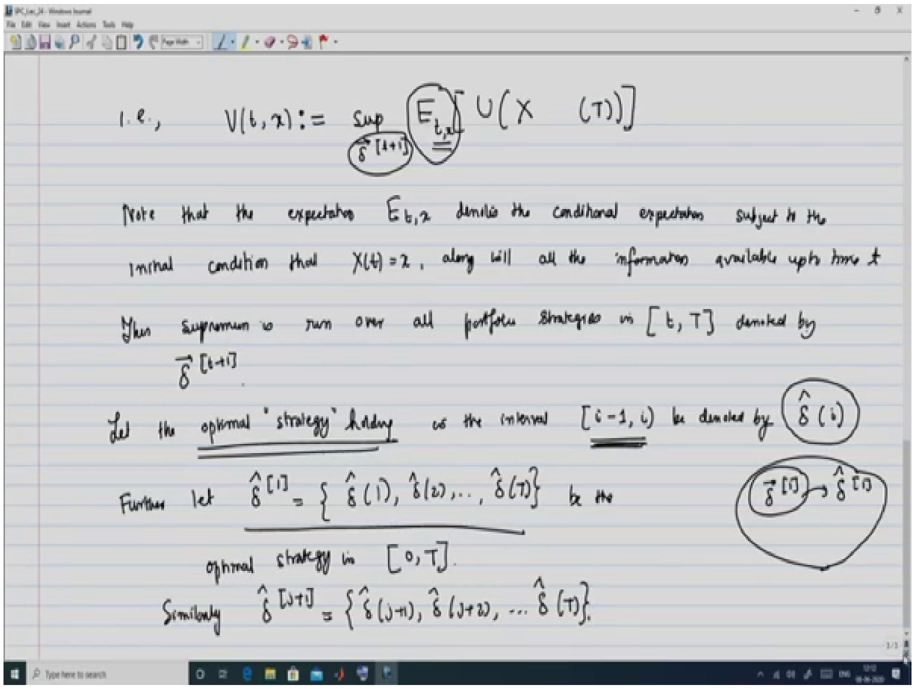


And what is the problem we want to solve? We want to find the supremum of the expected utility. So, the utility at the final time point, so the wealth of the final time point is X of T and we consider is utility which is a random variable. So, we take its expectation. And this supremum or the maximization will happen over the vector delta superscript 1 over delta superscript 1. Now, here we have introduced this new notation vector of vector delta superscript 1. So, in order to explain this, we say that the notation in general the $\vec{\delta}(j)$ this is defined as

$$\vec{\delta}^{[j]} := \left\{ \vec{\delta}(j), \vec{\delta}(j+1), \cdots, \vec{\delta}(T) \right\}.$$

Where generically this $\vec{\delta}(i)$, which is of the form $\vec{\delta}(j)$, $\vec{\delta}(j+1)$, $\cdots$ $\vec{\delta}(T)$ for each of those $\vec{\delta}(i)$ is a generic notation of a vector of portfolio holdings in the period $[i-1, i]$. So, this means that this $\vec{\delta}(i)$ is going to be nothing but some combination of $\delta_0(i), \delta_1(i), \cdots, \delta_N(i)$. And these are going to be the number of units of the bonds and capital N number of stocks held between time $i-1$ and i, so that means, this is the position you get into a time $i-1$ and hold on to until time i at which point you reshuffled your portfolio, and the reshuffle portfolio is going to be $\delta(i+1)$. So, then what you do is that, we basically carry out the maximization of the expected utility over the possible or the feasible portfolios of the form $\delta(1), \delta(2), \cdots, \delta(T)$. So, that means, that it we do this optimization over each of those periods by taking into consideration all feasible portfolios that can be held in each of those individual periods ok. So, now accordingly, so once we have this notation of the portfolio holdings in the period $[i-1, i]$, so accordingly this gives us that this notation of delta a superscript square bracket of j, this is going to be the collection of the portfolio strategies from j all the way to capital T. So, it is the portfolio strategy for the period $[j-1, T]$, alright. So, the next thing we do is we introduce the concept of what is known as the value function. So, what is the value function? So, in a qualitative way of looking at it, the value function is the maximum value an investor following the above objective say 1 can attain, with the wealth level at time small t being small x. So, this basically means that if a time t, you start off with a wealth level of x, then the value function is going to be the maximum value that the investor is going to obtain as a result of following our objective 1.

(Refer Slide Time: 09:40)



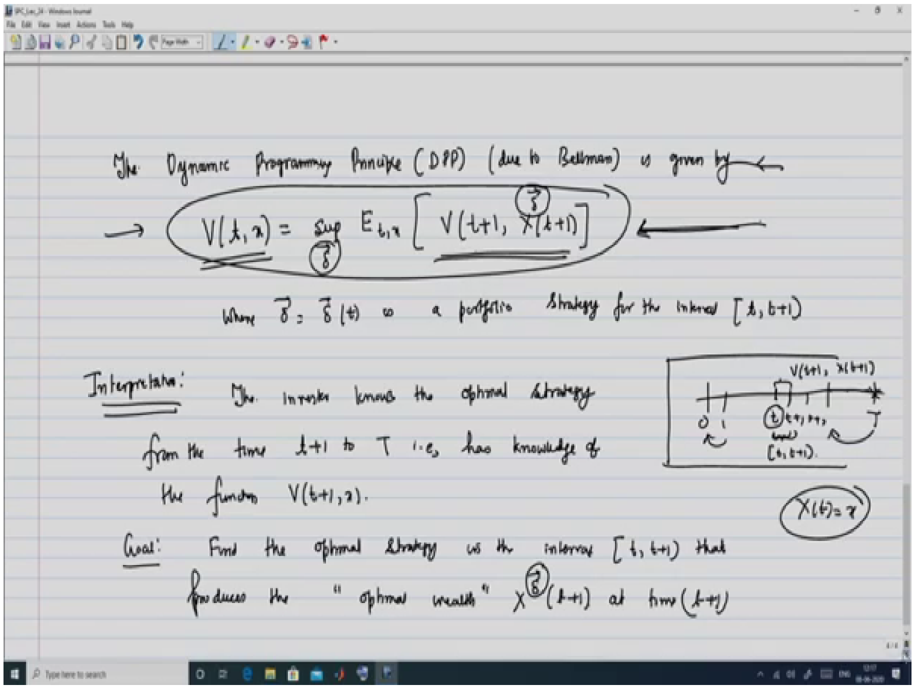
So, this means that now the mathematical formulation accordingly is going to be V of t x. So, this is the value function and this is going to be nothing but the supremum of the expected utility of $X(T)$ as before, but it is conditioned on the wealth at time $t = x$, and that means, you do this overall portfolios from the time interval at $t = T$ denoted by $\vec{\delta}^{[t+1]}$. So, now here you observe that this expectation, this expectation is a subscript t and x. So, we need to make a note about this that the expectation E subscript of t, x this will denote the conditional expectation subject to the initial condition that $X(t) = x$ along with all the information available up to time of small t. And this supremum is run over all portfolio strategies in the interval $[t, T]$ denoted by $\vec{\delta}^{[t+1]}$, alright. So, now let us introduce new notation after the optimization in then. So, accordingly let the optimal strategy holding in the interval $[i-1, i]$ be denoted by $\hat{\delta}_i$. So, before optimization this was the delta vector i, and once it has been optimized, so delta hat of i is going to be that particular portfolio which is the optimized portfolio in the interval $[i-1, i]$ driven by the objective of calculating the supremum of the expected utility in that particular time interval. So, then, so this optimal strategy holding in this interval which are denote by delta hat of i, and accordingly so further this automatically gives me the analogous delta hat of square bracket of i, this is going to be nothing, but the optimal strategy $\hat{\delta}(1), \hat{\delta}(2), \cdots, \hat{\delta}(T)$. So, what you want to find is essentially that if you have delta vector of 1, and then once you have optimized this; that means, you have ascertained that delta vector of 1 which will give you the optimized value which is delta hat of 1 as a result from maximizing the expected utility. So, this, so accordingly we let this be the optimal strategy in the entire investment window of $[0, T]$. Or similarly we can have the notation that delta hat of j plus 1, so this is the optimal strategy from the time point j all the way to time point capital T.

(Refer Slide Time: 14:18)

So, now once we have all the background work done, we are now in a position to talk about the Dynamic Programming Principle or DPP, due to Bellman, which is given by the following. That the value function at time t when the wealth level at time t is x is going to be given by the supremum of the expected value.

$$V(t, x) = \sup_{\vec{\delta}} E_{t,x}[V(t+1), X(t+1)]$$

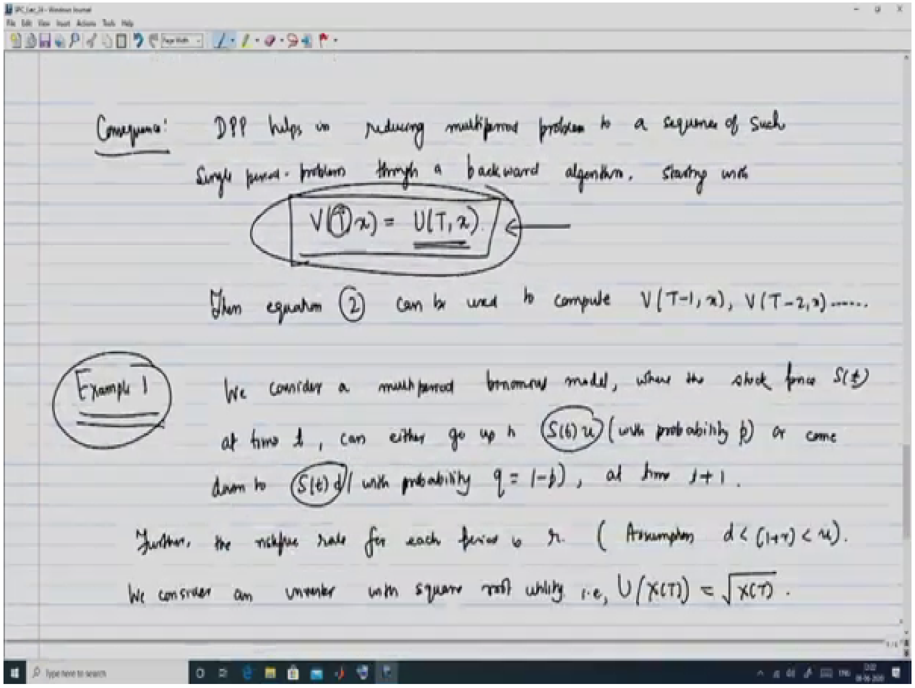
So, it is based on the assumption that you have determined what is the value function at time t plus 1 when the wealth level at time t plus 1 is denoted by X of t plus 1. And then this is over all the possible portfolio strategies of delta and you take the supremum over delta, where this delta vector is delta of t is a portfolio strategy for the interval closed t and open t plus 1. So, let me just elaborate a little more on this. So, this



means that we have this time point say a 0, 1, so on and we have this time point $t, t+1, t+2, \cdots, T$. So, the dynamic program principle does the following is that you have started off with time capital T, and suppose that you have arrived at time t plus 1 and you have evaluated what is the value function at that point which is $V(t+1)$ and $X(t+1)$. So, what do you need to do now is now the dynamic programming principle as given here, this dynamic programming principle is to determine what is going to be the optimal strategy in the interval $[t, t+1]$ amongst all the possible values of delta. So, there could be many such delta vectors which are allowable as portfolios during the period $[t, t+1]$. And what you do is the dynamic programming principle gives you that the value at (t, x), of course, you know when you have found out what is the value function at $t+1$ and $x+1$ that is going to be given by the supremum of the expected value of the value function over all possible portfolios delta vector that are allowed in the interval $[t, t+1]$, alright. So, basically this is a recursive formulation where you start off in backward in time and then eventually we arrive at time $t = 0$. So, what is, so let me just you know put down the interpretation for this. So, the interpretation of this is the following that as I have explained that you know what is happened, and from capital T to time t plus 1. So, then we can say that the investor knows the optimal strategy from the time $t+1$ to T. So, that is the investor has knowledge of the function $V(t+1, x)$. And what is the goal? The goal is to a find the optimal strategy in the interval, remember we are now looking at the interval $[t, t+1]$, so in the interval $[t, t+1]$ that produces the optimal wealth X superscript vector delta vector t plus 1 at of course, time t plus 1. So, this means that you use the principle to figure out the optimal strategy that you can adopt with at time t with a wealth level X at time t being available to you, so that this condition of dynamic programming principle is achieved.

(Refer Slide Time: 19:30)

So, what is the consequence of this? And the consequence of this is that DPP helps in reducing what is a multi-period problem to a sequence of such single period problem through a backward algorithm. Remember we start off with time capital T and go backward in time. And in this backward algorithm, we start off with the value function at capital T being equal to the same as the utility function. So, I am writing the utility function as $U(T, x)$, just to leave it in a general form in case the utility function is time dependent, but so far you have already looked at a utility function that are just dependent on the wealth level. So, then, so let me just come back to this equation let me call this equation 2. So, then the equation 2 or the DPP this can be used starting from a V at time capital T, it can be used to compute the value function at the preceding time level, that means, $V(T-1, x), V(T-2, x)$ and so on, alright. So, now, let us try to have a



better clarity on the DPP through a couple of examples. And so let me start off with the first example. So, again we considers remember that we are in the discrete time framework. So, we consider a multi-period binomial model, where the stock price S of t at time t, can either go up to S of t into some of factor u with probability p or come down to S of t into d with probability, so d is the down factor, so this will happen with probability a q which of course is 1 minus p. And this values $S(t)u$, or $S(t)d$, these are attained at time $t+1$. Further the risk free rate for each period is r. So, I am taking an identical risk free rate for each period. So, obviously, you recall that the assumption is that in order to avoid any sort of risk less profit is going to be $d < 1 + r < u$. Now, we have to specify now that we have the entire framework ready, we have to now specify what is going to be the utility function. So, accordingly we consider an investor with square root utility, so that means, that your u, so the terminal utility is going to be

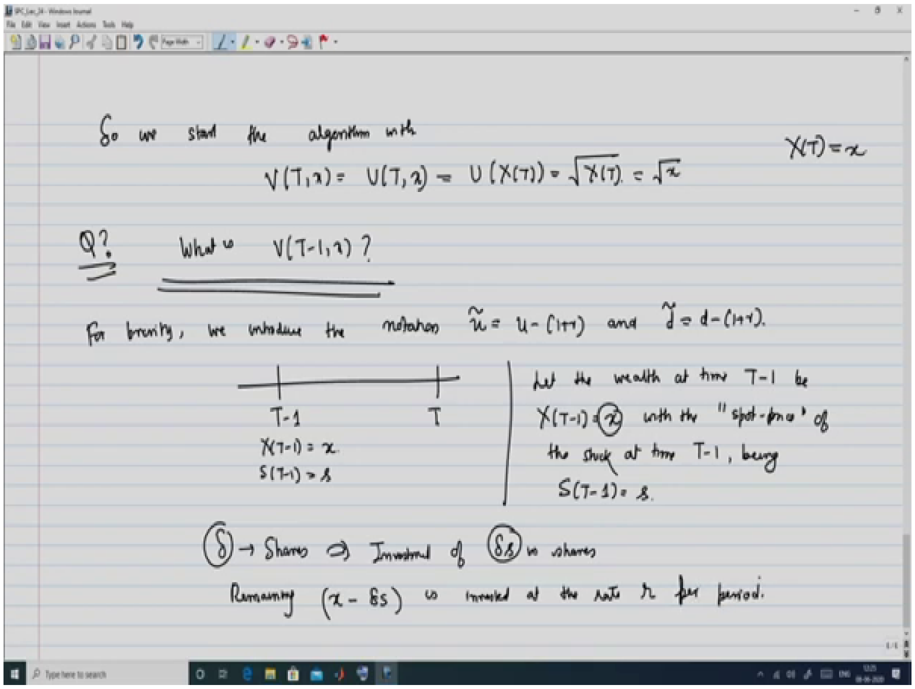
$$U(X(T)) = \sqrt{X(T)}.$$

(Refer Slide Time: 23:50)

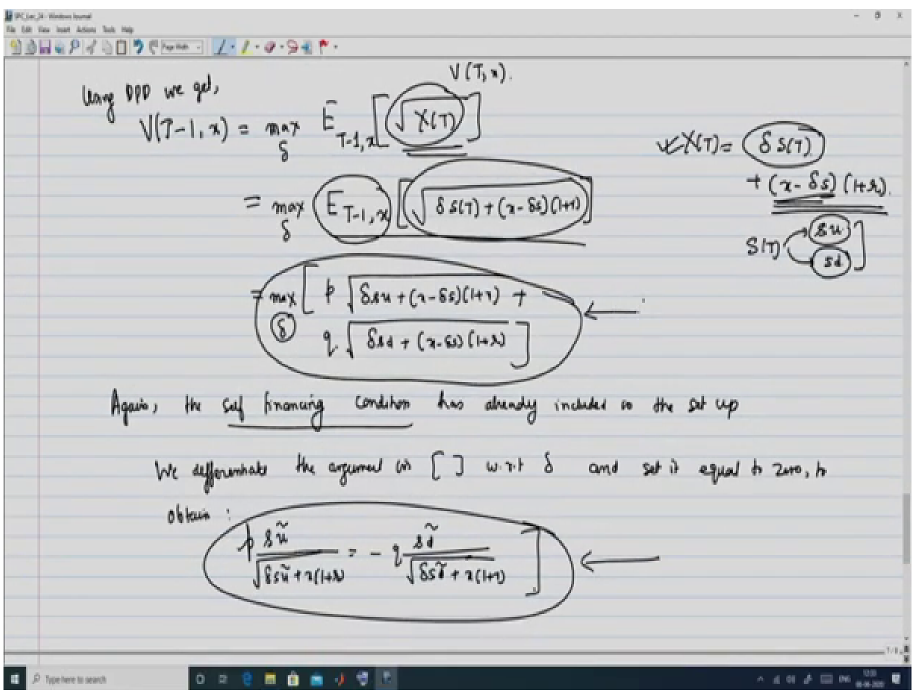
So, we start the DPP algorithm with, so remember what was the assumption that the DPP algorithm we start off with

$$V(T, x) = U(T, x) = U(X(T)) = \sqrt{X(T)} = \sqrt{x}.$$

And now suppose that we take that the wealth level at time t, X of T is equal to x that is your target wealth. So, this can then be written as square root of x. So, this is, this x is just some sort of a dummy variable. And the question that you want to answer here is what is $V(T-1, x)$, and then of course, $V - V(T-2, x)$ and so on. So, for the sake of brevity as we will see later on we introduce the notation u tilde which is $u - (1 + r)$, and d tilde which is $d - (1 + r)$. So, we are going to use these two notations later on. So, now let us look at this time period time point T minus 1 prior to capital T, and we take the wealth level at X T minus 1 again to be some generic value of x, and the spot price S of t minus 1 is equal to s. So, that is that let the wealth level for the investor at time capital T minus 1, what is this going to be this is X of capital T minus 1, and this is going to be x with the spot price of the stock at time T minus 1 being, so this is going to be s of T minus 1, let it be some small s. So, what is the strategy? So, recall that we are going to buy a delta number of shares, and this would mean an investment of an amount of delta s in stocks. And this gives that what is going to be the remaining amount? The remaining amount is you had an amount of x out of which you have invested an amount of delta s, so this remaining amount of x minus delta s is invested at the rate of r per period.



(Refer Slide Time: 26:59)



Then the dynamic programming principle, so using DPP, we get V of T minus 1 x, what is this going to be? This is going to be the maximum value of delta over delta of the expected utility at the next point which is square root of X of T subject to the condition that at time T minus 1 your wealth level is x. Now, remember that your X of T is going to be what, remember you had invested in delta number of stocks all right. So, this value at time 1 will become delta into s of capital T, and you had invested x minus delta s in the bond, so this will at time t become starting from time t minus 1 this will become x minus delta s into 1 plus r. So, accordingly I can now replace the square root of x t this is going to be, remember I have getting this, this is nothing but V of T, x. So, this is maximization over delta of expected conditional expectation of T minus 1 x. and I can replace X of T by this expression here, so this is going to be

$$\delta S(T) + (x - \delta s)(1 + r).$$

So, now you see that this part of X t is same irrespective of whether the stock goes up or down, but this term S of T is a random variable which can either take the value of small s which was the price at time capital T minus 1. So, this is going to be s into u or it is going to be s into d. So, accordingly this expected value is of this random variable which can take two values, either it is going to take

$$\delta su + (x - \delta s)(1 + r).$$

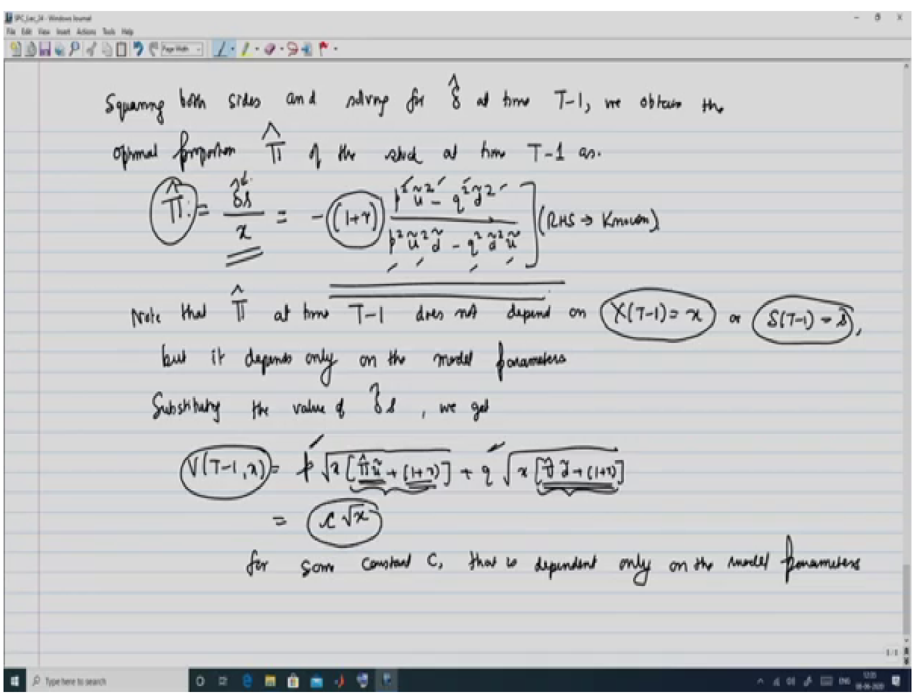
Or it is going to take the value

$$\delta sd + (x - \delta s)(1 + r).$$

Now, since it is a square root utility, so I will take that the random variable is going to be the square root of these two quantities. And since delta this s u value can be taken with probability p, and s d value will be taken with probability q, so then the expectation will be given by p into this expression plus q into the expression for the case when the stock price goes down. So, now this is the this expression here is going to be the expected value of conditional expectation of $(T - 1, x)$. And now we have to maximize this over delta. So, your eventual goal is to figure out what is going to be your optimal delta. Now, you again observe that just like the example done in the previous class, the self financing condition has already, so by the inclusion of this term $x - \delta s$ as the self financing condition has already been included in the setup. Now, let us focus on this problem of maximization of this expected value. So, for that purpose, we use single variable calculus. Since, we need to optimize only with respect to delta. So, we differentiate with the argument in the square bracket with respect to delta and set it equal to 0 to eventually obtain the following. So, after some algebra you obtain that p of s u tilde. So, this is remember my u tilde, I had defined, what is my u tilde? My $\tilde{u}$ was $u - (1 + r)$, and $\tilde{d}$ was $d - (1 + r)$. So, we are not going to make use of them. So, you obtain the expression that

$$\frac{ps\tilde{u}}{\delta s\tilde{u} + x(1 + r)} = -\frac{qs\tilde{d}}{\delta s\tilde{d} + x(1 + r)}.$$

(Refer Slide Time: 31:54)



So, we square both the sides, and solving for delta hat at time, so now, I can call this delta hat because this delta as a result of the maximization. So, I solve for delta hat at time $T - 1$. We obtain the optimal proportion which will denote by $\hat{\pi}$ of the stock at time $T - 1$ as. So, what is going to $\hat{\pi}$?

$$\hat{\pi} = \frac{\hat{\delta}s}{x} = -(1 + r)\frac{p^2\tilde{u}^2 - q^2\tilde{d}^2}{p^2\tilde{u}^2\tilde{d} - q^2\tilde{d}^2\tilde{u}}.$$

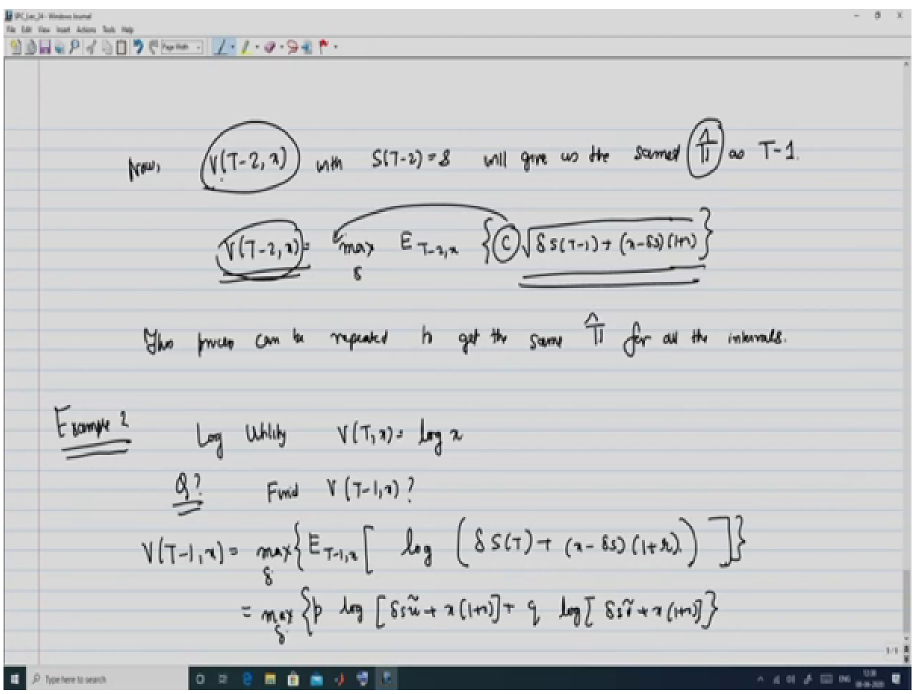
Now, observe very carefully what are the terms we have, I have a 1 plus r term which is already known p and q unknown, and since my u d and 1 plus r are known so; obviously, u tilde d tilde are both known. So, every quantity on this right hand side this is known, and these are nothing but the parameters of the model. So, accordingly we make the observation that this pi hat that we have at remember that this pi hat is at time T minus 1. So, this pi hat at time T minus 1 this does not depend on little x which was the wealth level at time T minus 1 that we started off with or the spot price of the stock at time T minus 1 that is little s. But it depends only on the model parameters. So, now that we have obtained this delta hat here. We can now substitute this to obtain the actual maximum value that means to determine what is $V(T-1, x)$.

$$V(T-1, x) = p\sqrt{x[\hat{\pi}\tilde{u} + (1+r)]} + q\sqrt{x[\hat{\pi}\tilde{d} + (1+r)]}$$

And now observe carefully that this is nothing but, so the p and q are constant, and this expression is known as well as this expression is known. So, that means, that I can collate this as some constant c. So, I can take the square root of x out and the remaining term, I will collate and call this as c. And this c is a constant which is dependent only on the model parameters, that means, it depends on p, q, and this $\hat{\pi}\tilde{u}$, these are known $1+r$, remember pi hat is only dependent on what is the model parameter. So, by the same logic, this term is also just dependent on the model parameter. So, this means that you started off with $V(T, x)$ is being $\sqrt{x}$, and now we have

$$V(T-1, x) = c\sqrt{x}.$$

(Refer Slide Time: 36:48)



So, now, if we take $V(T-2, x)$ with $S(T-2) = S$ will give us the same pi hat as time $T-1$. So, remember that here again when I do the backward algorithm, I look at $V(T-2)$. What is this going to be? This is again going to be simply, so you will get

$$V(T-2, x) = \max_{\delta} E_{T-2,x}\left\{C\sqrt{\delta s(T-1) + (x - \delta s)(1+r)}\right\}.$$

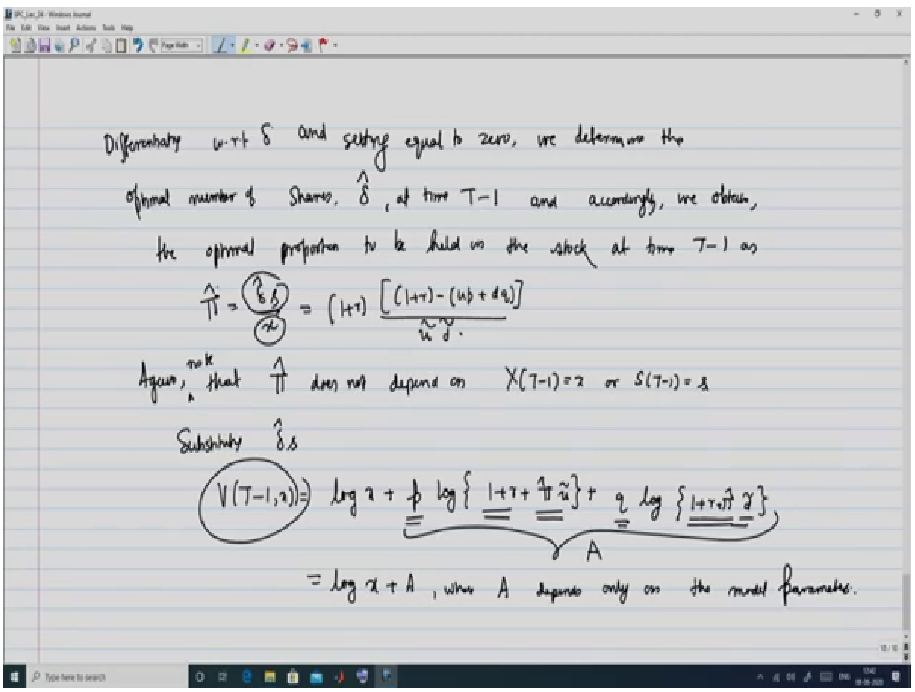
So, it is going to be again the same exercise except this C goes out. So, that means, this maximization are to determine $V(T-2, x)$ is again the same problem of maximizing the square root of this random variable, the expected value of this random variable. So, accordingly this means that even in this step the pi hat will remain the same, and this will hold true for all the preceding steps. And the only thing that will change from step to step is going to be what is going to be the value function, what is the constant in the value function. So, generically this means that, so this process can be repeated to get the same pi hat for all the intervals. So, just to have a better clarity on this, let us look at another example. So, in this case, what we did is that we now consider an investor with log utility that is $V(T, x)$ is going to be $\log(x)$. And as before what is the question, the question is I want to use the dynamic programming principle to find what is $V(T-1, x)$. The dynamic programming principle gives $V(T-1, x)$ this is going to be nothing but the maximum value of

$$V(T-1, x) = \max_{\delta} \{E_{T-1,x}[\log(\delta s(T) + (x - \delta s)(1+r))]\}$$

$$= \max_{\delta} \left\{ Ep \log[\delta s \tilde{u} + x(1+r)] + q \log[\delta s \tilde{d} + x(1+r)] \right\}$$

So, again using so again here as you observe that we have taken care of the self-financing condition, and this turns out to be nothing but the maximization of is function of a single variable namely δ.

(Refer Slide Time: 40:30)



So, accordingly, what you do is we differentiate with respect to delta and setting equal to 0, we determine the optimal number of shares which you denote by delta hat, at time capital T minus 1, and accordingly, we obtain, the optimal weight or proportion to be held in the stock at time capital T minus 1 as pi hat. So, this is again the optimal weight is equal to

$$\frac{\hat{\delta} s}{x}.$$

Remember, this is the investment in the stock over the total initial investment is equal to

$$(1+r)\frac{[(1+r) - (up + dq)]}{\tilde{u}\tilde{d}}.$$

And again in this case also for the log utility, we note that pi hat does not depend on X of T minus 1 equal to x, or S of T minus 1 equal to s. So, substituting this delta hat of s, so I will substitute this delta hat of s in the expression for V of T minus 1, so we get that

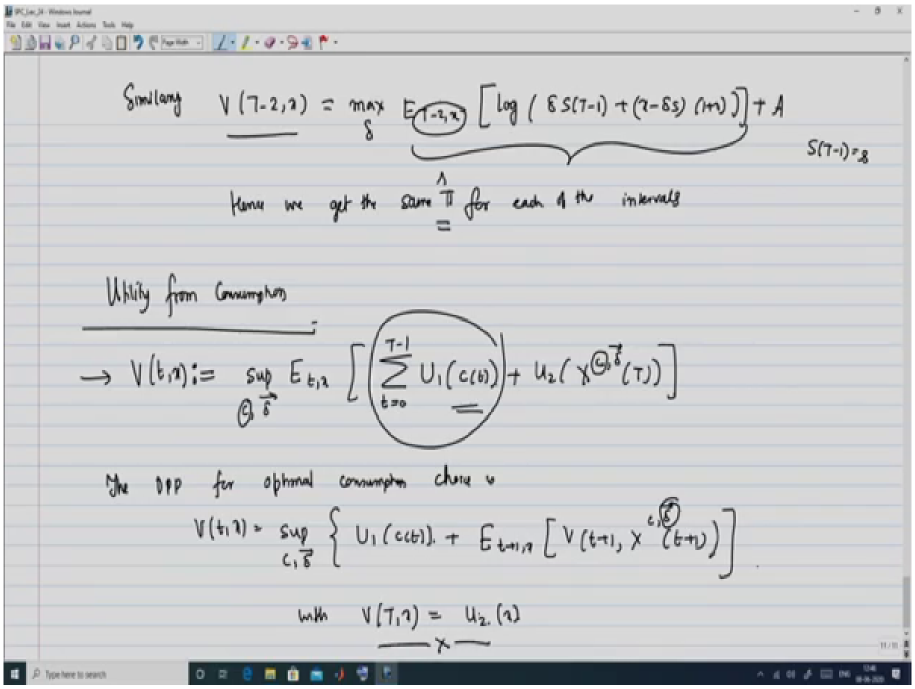
$$V(T-1, x) = \log x + p \log\{1 + r + \hat{\pi}\tilde{u}\} + q \log\{1 + r + \hat{\pi}\tilde{d}\}.$$

Now, observe carefully here $1 + r$ and $\hat{\pi}\tilde{u}$ as well as $1 + r$ and $\hat{\pi}\tilde{u}$, they are known and dependent on model parameters and so r, p and q. So, that means, this entire expression here, this is going to be some constant A. So, accordingly we get

$$V(T-1, x) = \log x + A,$$

where this A depends only on the model parameters.

(Refer Slide Time: 43:48)



So, similarly, you will get

$$V(T-2, x) = \max_{\delta} E_{T-2,x}[\log(\delta S(T-1)) + (x - \delta s)(1 + r)] + A.$$

So, here we take S of T minus 1 to be this dummy variable little s plus of course, some constant. So, you see that now this maximization problem again reduces to the same problem done at time T minus 1. So, hence we get the same pi hat for each of the intervals. And the value is going to be of the form log x plus A, except the constant A is going to change depending on what the interval is, but the pi hat is going to be identical for each of the intervals. So, we just end our discussion with the problem statement for utility from consumption. So, for the utility for consumption, I first define what is the value function $V(t, x)$. And this is going to be nothing but

$$\sup_{c,\vec{\delta}} E_{t,x}\left[\sum_{t=0}^{T-1} U_1(c(t)) + U_2(X^{c,\vec{\delta}}(T))\right].$$

And in addition I have to take the utility of the consumption process C of t from time t is equal to 0 to capital T minus 1, and accordingly the supremum is going to be taken over not only just the vector delta, but also c. So, the only change that has happened is that I have this term we just shown up, and now accordingly the optimization has to be done over all the consumption process. So, once I have defined what is the value function by taking into account the consumption process, then the dynamic programming principle for optimal consumption choice is then given by

$$V(t, x) = \sup_{c,\vec{\delta}} \left\{U_1(c(t)) + E_{t+1,x}[V(t+1, X^{c,\vec{\delta}}(t+1))]\right\},$$

with the final condition that $V(T, x) = U_2(x)$. So, this brings us to the end of this lecture. So, just to do a recap of whatever we have done on this lecture, we started off by looking at the means of extending the single period optimization that you had done in the previous class to a multi-period optimization. However, recognizing the fact that extension of the approach use for a single period into a multi-period setup will result in a significant increase in the computational consideration. So, according to this program of multi-period optimization was then reduced to a series of single period optimization by making use of what is known as the dynamic programming principle. And for the dynamic programming principle, we considered what is known as the value function. And determine the value function at any particular time t by taking into account that or the assumption that the value function at time $t + 1$ has been determined based on the optimization having been carried out between time t plus 1 with the terminal point time capital T, where the value function at time capital T is given to be the utility function. And in order to have a better clarity on the application of the dynamic programming principle, we looked at two examples of utility functions namely the square utility and the log utility wherein we considered a portfolio of a stock where the investment is being made in delta number of stocks following the binomial model, and the remaining amount is invested in a money market account or a risk free asset with the risk free rate for a single period being taken as r. And you concluded this by talking about the formulation for utility for consumption by first defining what is the value function, and then defining what is going to be the dynamic programming principle by accommodating what is the consumption choice. So, in the next class, we will move on to the framework of continuous time modeling, and then we will look at the counterpart of the dynamic programming principle namely the Hamilton-Jacobi-Bellman equation.

Thank you for watching.