Dynamic Portfolio Choice II

Dynamic Programming

Leonid Kogan

MIT, Sloan

15.450, Fall 2010

Outline

Overview

- When all state-contingent claims are redundant, i.e., can be replicated by trading in available assets (e.g., stocks and bonds), dynamic portfolio choice reduces to a static problem.
- There are many practical problems in which derivatives are not redundant, e.g., problems with constraints, transaction costs, unspanned risks (stochastic volatility).
- Such problems can be tackled using Dynamic Programming (DP).
- DP applies much more generally than the static approach, but it has practical limitations: when the closed-form solution is not available, one must use numerical methods which suffer from the curse of dimensionality.

IID Returns

Formulation

- Consider the discrete-time market model.
- There is a risk-free bond, paying gross interest rate $R_f = 1 + r$.
- There is a risky asset, stock, paying no dividends, with gross return R_t, IID over time.
- The objective is to maximize the terminal expected utility

$$\max \mathrm{E}_0 \left[U(W_T) \right]$$

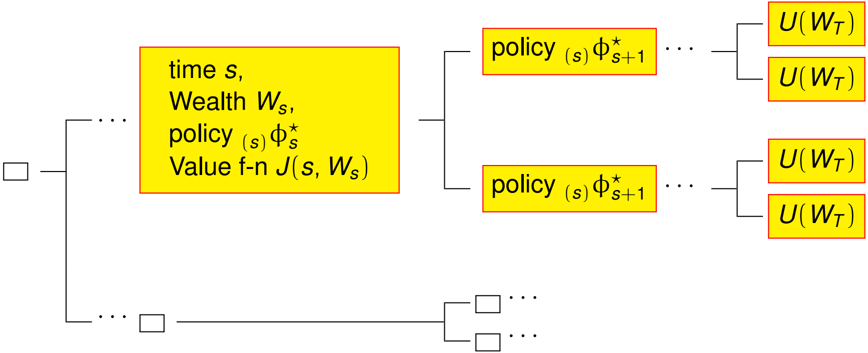
where portfolio value W_t results from a self-financing trading strategy

$$W_t = W_{t-1} \left[\phi_{t-1} R_t + (1 - \phi_{t-1}) R_f \right]$$

ϕ_t denotes the share of the stock in the portfolio.

Principle of Optimality

- Suppose we have solved the problem, and found the optimal policy $\phi_t^\star$.
- Consider a tail subproblem of maximizing $\mathrm{E}_s\left[U(W_T)\right]$ starting at some point in time s with wealth W_s.



Principle of Optimality

- Let

$$_{(s)}\phi^\star_s,\ _{(s)}\phi^\star_{s+1},\ \ldots,\ _{(s)}\phi^\star_{T-1}$$

  denote the optimal policy of the subproblem.
- The Principle of Optimality states that the optimal policy of the tail subproblem coincides with the corresponding portion of the solution of the original problem.
- The reason is simple: if policy $(_{(s)}\phi^\star_{\ldots})$ could outperform the original policy on the tail subproblem, the original problem could be improved by replacing the corresponding portion with $(_{(s)}\phi^\star_{\ldots})$.

IID Returns

DP

- Suppose that the time-t conditional expectation of terminal utility under the optimal policy depends only on the portfolio value W_t at time t, and nothing else. This conjecture needs to be verified later.

$$\mathrm{E}_t\left[U(W_T)|_{(t)}\phi^{\star}_{t,\ldots,T-1}\right] = J(t, W_t)$$

- We call $J(t, W_t)$ the indirect utility of wealth.
- Then we can compute the optimal portfolio policy at $t-1$ and the time-$(t-1)$ expected terminal utility as

$$J(t-1, W_{t-1}) = \max_{\phi_{t-1}} \mathrm{E}_{t-1}\left[J(t, W_t)\right] \qquad \text{(Bellman equation)}$$

$$W_t = W_{t-1}\left[\phi_{t-1}R_t + (1-\phi_{t-1})R_f\right]$$

- $J(t, W_t)$ is called the value function of the dynamic program.

IID Returns

DP

- DP is easy to apply.
- Compute the optimal policy one period at a time using backward induction.
- At each step, the optimal portfolio policy maximizes the conditional expectation of the next-period value function.
- The value function can be computed recursively.
- Optimal portfolio policy is dynamically consistent: the state-contingent policy optimal at time 0 remains optimal at any future date t. Principle of Optimality is a statement of dynamic consistency.

IID Returns

Binomial tree

- Stock price

$$S_t = S_{t-1} \times \begin{cases} u, & \text{with probability } p \\ d, & \text{with probability } 1-p \end{cases}$$

- Start at time $T-1$ and compute the value function

$$J(T-1, W_{T-1}) = \max_{\phi_{T-1}} \mathrm{E}_{T-1}\left[U(W_T)|\phi_{T-1}\right] =$$

$$\max_{\phi_{T-1}} \left\{ \begin{array}{l} pU\left[W_{T-1}\left(\phi_{T-1}u + (1-\phi_{t-1})R_f\right)\right] + \\ \\ (1-p)U\left[W_{T-1}\left(\phi_{T-1}d + (1-\phi_{T-1})R_f\right)\right] \end{array} \right\}$$

- Note that value function at $T-1$ depends on W_{T-1} only, due to the IID return distribution.

IID Returns

Binomial tree

- Backward induction. Suppose that at t, $t+1$, ..., $T-1$ the value function has been derived, and is of the form $J(s, W_s)$.
- Compute the value function at $t-1$ and verify that it still depends only on portfolio value:

$$J(t-1, W_{t-1}) = \max_{\phi_{t-1}} \mathrm{E}_{t-1}\left[J(t, W_t)|\phi_{t-1}\right] =$$

$$\max_{\phi_{t-1}} \left\{ \begin{array}{l} pJ\left[t, W_{t-1}\left(\phi_{t-1}u + (1-\phi_{t-1})R_f\right)\right] + \\ \\ (1-p)J\left[t, W_{t-1}\left(\phi_{t-1}d + (1-\phi_{t-1})R_f\right)\right] \end{array} \right\}$$

- Optimal portfolio policy $\phi^\star_{t-1}$ depends on time and the current portfolio value:

$$\phi^\star_{t-1} = \phi^\star(t-1, W_{t-1})$$

IID Returns, CRRA Utility

Binomial tree

- Simplify the portfolio policy under CRRA utility $U(W_T) = \frac{1}{1-\gamma} W_T^{1-\gamma}$

$$J(T-1, W_{T-1}) = \max_{\phi_{T-1}} \mathrm{E}_{T-1}\left[\frac{1}{1-\gamma} W_T^{1-\gamma} | \phi_{T-1}\right] =$$

$$\max_{\phi_{T-1}} \left\{ \begin{array}{l} p \frac{1}{1-\gamma} W_{T-1}^{1-\gamma} \left(\phi_{T-1} u + (1-\phi_{T-1}) R_f\right)^{1-\gamma} + \\ \\ (1-p) \frac{1}{1-\gamma} W_{T-1}^{1-\gamma} \left(\phi_{T-1} d + (1-\phi_{T-1}) R_f\right)^{1-\gamma} \end{array} \right\}$$

$$= A(T-1)\, W_{T-1}^{1-\gamma}$$

where $A(T-1)$ is a constant given by

$$A(T-1) = \max_{\phi_{T-1}} \frac{1}{1-\gamma} \left\{ \begin{array}{l} p \left(\phi_{T-1} u + (1-\phi_{T-1}) R_f\right)^{1-\gamma} + \\ \\ (1-p) \left(\phi_{T-1} d + (1-\phi_{T-1}) R_f\right)^{1-\gamma} \end{array} \right\}$$

IID Returns, CRRA Utility

Binomial tree

- Backward induction

$$J(t-1, W_{t-1}) = \max_{\phi_{t-1}} \mathrm{E}_{t-1}\left[A(t)\,W_t^{1-\gamma} | \phi_{t-1}\right] =$$

$$\max_{\phi_{t-1}} \left\{ \begin{array}{l} pA(t)\,W_{t-1}^{1-\gamma}\;(\phi_{t-1}u + (1-\phi_{t-1})R_f)^{1-\gamma} + \\ \\ (1-p)A(t)\,W_{t-1}^{1-\gamma}\;(\phi_{t-1}d + (1-\phi_{t-1})R_f)^{1-\gamma} \end{array} \right\}$$

$$= A(t-1)\;W_{t-1}^{1-\gamma}$$

where $A(t-1)$ is a constant given by

$$A(t-1) = \max_{\phi_{t-1}} A(t) \left\{ \begin{array}{l} p\;(\phi_{t-1}u + (1-\phi_{t-1})R_f)^{1-\gamma} + \\ \\ (1-p)\;(\phi_{t-1}d + (1-\phi_{t-1})R_f)^{1-\gamma} \end{array} \right\}$$

Black-Scholes Model, CRRA Utility

Limit of binomial tree

- Parameterize the binomial tree so the stock price process converges to the Geometric Brownian motion with parameters μ and σ: $p = 1/2$,

$$u = \exp\left(\left(\mu - \frac{\sigma^2}{2}\right)\Delta t + \sigma\sqrt{\Delta t}\right), \quad d = \exp\left(\left(\mu - \frac{\sigma^2}{2}\right)\Delta t - \sigma\sqrt{\Delta t}\right)$$

- Let $R_f = \exp(r\,\Delta t)$. Time step is now Δt instead of 1.
- Take a limit of the optimal portfolio policy as $\Delta t \to 0$:

$$\phi_t^\star = \arg\max_{\phi_t} A(t+\Delta t)\left\{ \begin{array}{l} p\,\left(\phi_t u + (1-\phi_t)R_f\right)^{1-\gamma} + \\ (1-p)\,\left(\phi_t d + (1-\phi_t)R_f\right)^{1-\gamma} \end{array} \right\}$$

$$\approx \arg\max_{\phi_t} A(t+\Delta t)\left\{ \begin{array}{l} 1 + (1-\gamma)(r + \phi_t(\mu - r))\,\Delta t \\ -(1/2)(1-\gamma)\gamma\phi_t^2\sigma^2\,\Delta t \end{array} \right\}$$

- Optimal portfolio policy

$$\phi_t^\star = \frac{\mu - r}{\gamma\sigma^2}$$

Black-Scholes Model, CRRA Utility

- Optimal portfolio policy

$$\phi_t^{\star} = \frac{\mu - r}{\gamma \sigma^2}$$

- We have recovered the Merton's solution using DP. Merton's original derivation was very similar, using DP in continuous time.
- The optimal portfolio policy is myopic, does not depend on the problem horizon.
- The value function has the same functional form as the utility function: indirect utility of wealth is CRRA with the same coefficient of relative risk aversion as the original utility. That is why the optimal portfolio policy is myopic.
- If return distribution was not IID, the portfolio policy would be more complex. The value function would depend on additional variables, thus the optimal portfolio policy would not be myopic.

General Formulation

- Consider a discrete-time stochastic process $Y_t = (Y_t^1, ..., Y_t^N)$.
- Assume that the time-t conditional distribution of Y_{t+1} depends on time, its own value and a control vector ϕ_t:

$$pdf_t(Y_{t+1}) = p(Y_{t+1}, Y_t, \phi_t, t)$$

- For example, vector Y_t could include the stock price and the portfolio value, $Y_t = (S_t, W_t)$, and the transition density of Y would depend on the portfolio holdings ϕ_t.
- The objective is to maximize the expectation

$$\mathrm{E}_0 \left[\sum_{t=0}^{T-1} u(t, Y_t, \phi_t) + u(T, Y_T) \right]$$

- For example, in the IID+CRRA case above, $Y_t = W_t$, $u(t, Y_t, \phi_t) = 0$, $t = 0, ..., T-1$ and $u(T, Y_T) = (1-\gamma)^{-1} (Y_T)^{1-\gamma}$.
- We call Y_t a controlled Markov process.

Formulation

State augmentation

- Many dynamic optimization problems of practical interest can be stated in the above form, using controlled Markov processes. Sometimes one needs to be creative with definitions.
- State augmentation is a common trick used to state problems as above.
- Suppose, for example, that the terminal objective function depends on the average of portfolio value between 1 and T.
- Even in the IID case, the problem does not immediately fit the above framework: if the state vector is $Y_t = (W_t)$, the terminal objective

$$\frac{1}{1-\gamma}\left(\frac{1}{T}\sum_{t=1}^{T} W_t\right)^{1-\gamma}$$

cannot be expressed as

$$\sum_{t=0}^{T-1} u(t, Y_t, \phi_t) + u(T, Y_T)$$

Formulation

State augmentation

- Continue with the previous example. Define an additional state variable A_t:

$$A_t = \frac{1}{t} \sum_{s=1}^{t} W_s$$

- Now the state vector becomes

$$Y_t = (W_t, A_t)$$

  Is this a controlled Markov process?

- The distribution of W_{t+1} depends only on W_t and ϕ_t.
- Verify that the distribution of (W_{t+1}, A_{t+1}) depends only on (W_t, A_t):

$$A_{t+1} = \frac{1}{t+1} \sum_{s=1}^{t+1} W_s = \frac{1}{t+1} (tA_t + W_{t+1})$$

  Y_t is indeed a controlled Markov process.

Formulation

Optimal stopping

- Optimal stopping is a special case of dynamic optimization, and can be formulated using the above framework.
- Consider the problem of pricing an American option on a binomial tree. Interest rate is r and the option payoff at the exercise date τ is $H(S_\tau)$.
- The objective is to find the optimal exercise policy $\tau^\star$, which solves

$$\max_{\tau} \mathrm{E}_0^{\mathbf{Q}} \left[(1+r)^{-\tau} H(S_\tau) \right]$$

  The exercise decision at τ can depend only on information available at τ.
- Define the state vector

$$(S_t, X_t)$$

  where S_t is the stock price and X_t is the status of the option

$$X_t \in \{0, 1\}$$

  If $X_t = 1$, the option has not been exercised yet.

Formulation

Optimal stopping

- Let the control be of the form $\phi_t \in \{0, 1\}$. If $\phi_t = 1$, the option is exercised at time t, otherwise it is not.
- The stock price itself follows a Markov process: distribution of S_{t+1} depends only on S_t.
- The option status X_t follows a controlled Markov process:

$$X_{t+1} = X_t(1 - \phi_t)$$

  Note that once X_t becomes zero, it stays zero forever. Status of the option can switch from $X_t = 1$ to $X_{t+1} = 0$ provided $\phi_t = 1$.
- The objective takes form

$$\max_{\phi_t} \mathrm{E}_0^{\mathbf{Q}} \left[\sum_{t=0}^{T-1} (1 + r)^{-t} H(S_t) X_t \phi_t \right]$$

Bellman Equation

- The value function and the optimal policy solve the Bellman equation

$$J(t-1, Y_{t-1}) = \max_{\phi_{t-1}} \mathrm{E}_{t-1}\left[u(t-1, Y_{t-1}, \phi_{t-1}) + J(t, Y_t) | \phi_{t-1}\right]$$
$$J(T, Y_T) = u(T, Y_T)$$

American Option Pricing

- Consider the problem of pricing an American option on a binomial tree. Interest rate is r and the option payoff at the exercise date τ is $H(S_\tau)$.
- The objective is to find the optimal exercise policy $\tau^\star$, which solves

$$\max_{\tau} \mathrm{E}_0^{\mathbf{Q}} \left[(1+r)^{-\tau} H(S_\tau) \right]$$

  The exercise decision at τ can depend only on information available at τ.
- The objective takes form

$$\max_{\phi_t \in \{0,1\}} \mathrm{E}_0^{\mathbf{Q}} \left[\sum_{t=0}^{T-1} (1+r)^{-t} H(S_t) X_t \phi_t \right]$$

  If $X_t = 1$, the option has not been exercised yet.
- Option price $P(t, S_t, X=0) = 0$ and $P(t, S_t, X=1)$ satisfies

$$P(t, S_t, X=1) = \max \Big(H(S_t), (1+r)^{-1} \mathrm{E}_t^{\mathbf{Q}} [P(t+1, S_{t+1}, X=1)] \Big)$$

Asset Allocation with Return Predictability

Formulation

- Suppose stock returns have a binomial distribution: $p = 1/2$,

$$u_t = \exp\left(\left(\mu_t - \frac{\sigma^2}{2}\right)\Delta t + \sigma\sqrt{\Delta t}\right), \quad d_t = \exp\left(\left(\mu_t - \frac{\sigma^2}{2}\right)\Delta t - \sigma\sqrt{\Delta t}\right)$$

  where the conditional expected return μ_t is stochastic and follows a Markov process with transition density

$$f(\mu_t|\mu_{t-1})$$

- Conditionally on μ_{t-1}, μ_t is independent of R_t.
- Let $R_f = \exp(r\,\Delta t)$.
- The objective is to maximize expected CRRA utility of terminal portfolio value

$$\max \mathrm{E}_0\left[\frac{1}{1-\gamma} W_T^{1-\gamma}\right]$$

Asset Allocation with Return Predictability

Bellman equation

- We conjecture that the value function is of the form

$$J(t, W_t, \mu_t) = A(t, \mu_t) W_t^{1-\gamma}$$

- The Bellman equation takes form

$$A(t-1, \mu_{t-1}) W_{t-1}^{1-\gamma} = \max_{\phi_{t-1}} \mathrm{E}_{t-1} \left[A(t, \mu_t) \left(W_{t-1}(\phi_{t-1}(R_t - R_f) + R_f) \right)^{1-\gamma} \right]$$

- The initial condition for the Bellman equation implies

$$A(T, \mu_T) = \frac{1}{1-\gamma}$$

- We verify that the conjectured value function satisfies the Bellman equation if

$$A(t-1, \mu_{t-1}) = \max_{\phi_{t-1}} \mathrm{E}_{t-1} \left[A(t, \mu_t) \left(\phi_{t-1}(R_t - R_f) + R_f \right)^{1-\gamma} \right]$$

  Note that the RHS depends only on μ_{t-1}.

Asset Allocation with Return Predictability

Optimal portfolio policy

- The optimal portfolio policy satisfies

$$\phi^\star_{t-1} = \arg\max_{\phi_{t-1}} \mathrm{E}_{t-1}\left[A(t,\mu_t)\left(\phi_{t-1}(R_t - R_f) + R_f\right)^{1-\gamma}\right]$$
$$= \arg\max_{\phi_{t-1}} \mathrm{E}_{t-1}\left[A(t,\mu_t)\right]\,\mathrm{E}_{t-1}\left[\left(\phi_{t-1}(R_t - R_f) + R_f\right)^{1-\gamma}\right]$$

  because, conditionally on μ_{t-1}, μ_t is independent of R_t.
- Optimal portfolio policy is myopic, does not depend on the problem horizon. This is due to the independence assumption.
- Can find $\phi^\star_t$ numerically.
- In the continuous-time limit of $\Delta t \to 0$,

$$\phi^\star_t = \frac{\mu_t - r}{\gamma\sigma^2}$$

Asset Allocation with Return Predictability

Hedging demand

- Assume now that the dynamics of conditional expected returns is correlated with stock returns, i.e., the distribution of μ_t given μ_{t-1} is no longer independent of R_t.
- The value function has the same functional form as before,

$$J(t, W_t, \mu_t) = A(t, \mu_t) W_t^{1-\gamma}$$

- The optimal portfolio policy satisfies

$$\phi_{t-1}^{\star} = \arg\max_{\phi_{t-1}} \mathrm{E}_{t-1}\left[A(t, \mu_t)\left(\phi_{t-1}(R_t - R_f) + R_f\right)^{1-\gamma}\right]$$

- Optimal portfolio policy is no longer myopic: dependence between μ_t and R_t affects the optimal policy.
- The deviation from the myopic policy is called hedging demand. It is non-zero due to the fact that the investment opportunities (μ_t) change stochastically, and the stock can be used to hedge that risk.

Optimal Control of Execution Costs

Formulation

- Suppose we need to buy $\overline{b}$ shares of the stock in no more than T periods.
- Our objective is to minimize the expected cost of acquiring the $\overline{b}$ shares.
- Let b_t denote the number of shares bought at time t.
- Suppose the price of the stock is S_t.
- The objective is

$$\min_{b_{0,\dots,T-1}} \mathrm{E}_0 \left[\sum_{t=0}^{T} S_t b_t \right]$$

- What makes this problem interesting is the assumption that trading affects the price of the stock. This is called *price impact*.

Optimal Control of Execution Costs

Formulation

- Assume that the stock price follows

$$S_t = S_{t-1} + \theta b_t + \varepsilon_t, \quad \theta > 0$$

- Assume that ε_t has zero mean conditional on S_{t-1} and b_t:

$$\mathrm{E}[\varepsilon_t | b_t, S_{t-1}] = 0$$

- Define an additional state variable W_t denoting the number of shares left to purchase:

$$W_t = W_{t-1} - b_{t-1}, \quad W_0 = \overline{b}$$

- The constraint that $\overline{b}$ shares must be bought at the end of period T can be formalized as

$$b_T = W_T$$

Optimal Control of Execution Costs

Solution

- We can capture the dynamics of the problem using a state vector

$$Y_t = (S_{t-1}, W_t)$$

which clearly is a controlled Markov process.

- Start with period T and compute the value function

$$J(T, S_{T-1}, W_T) = \mathrm{E}_T[S_T \, W_T] = (S_{T-1} + \theta W_T)\, W_T$$

- Apply the Bellman equation once to compute

$$J(T-1, S_{T-2}, W_{T-1}) = \min_{b_{T-1}} \mathrm{E}_{T-1}\left[S_{T-1} b_{T-1} + J(T, S_{T-1}, W_T)\right]$$

$$= \min_{b_{T-1}} \mathrm{E}_{T-1}\left[\begin{array}{l}(S_{T-2} + \theta b_{T-1} + \varepsilon_{T-1}) b_{T-1} + \\ J(T, S_{T-2} + \theta b_{T-1} + \varepsilon_{T-1}, W_{T-1} - b_{T-1})\end{array}\right]$$

- Find

$$b^{\star}_{T-1} = \frac{W_{T-1}}{2}$$

$$J(T-1, S_{T-2}, W_{T-1}) = W_{T-1}\left(S_{T-2} + \frac{3}{4}\theta W_{T-1}\right)$$

Optimal Control of Execution Costs

Solution

- Continue with backward induction to find

$$b^\star_{T-k} = \frac{W_{T-k}}{k+1}$$

$$J(T-k, S_{T-k-1}, W_{T-k}) = W_{T-k}\left(S_{T-k-1} + \frac{k+2}{2(k+1)}\theta W_{T-k}\right)$$

- Conclude that the optimal policy is deterministic

$$b^\star_0 = b^\star_1 = \cdots = b^\star_T = \frac{\overline{b}}{T+1}$$

Key Points

- Principle of Optimality for Dynamic Programming.
- Bellman equation.
- Formulate dynamic portfolio choice using controlled Markov processes.
- Merton's solution.
- Myopic policy and hedging demand.

References

- Bertsimas, D., A. Lo, 1998, "Optimal control of execution costs," *Journal of Financial Markets* 1, 1-50.

15.450 Analytics of Finance

Fall 2010