A Guided Tour of Chapter 6: Dynamic Asset-Allocation and Consumption

Ashwin Rao

ICME, Stanford University

January 30, 2021

Dynamic Asset-Allocation and Consumption

- The broad topic is Investment Management
- Applies to Corporations as well as Individuals
- The two considerations are:
    - How to allocate money across assets in one's investment portfolio
    - How much to consume for one's needs/operations/pleasures
- We consider the dynamic version of these dual considerations
- Asset-Allocation and Consumption decisions at each time step
- Asset-Allocation decisions typically deal with Risk-Reward tradeoffs
- Consumption decisions are about spending now or later
- Objective: Horizon-Aggregated Expected Utility of Consumption

Let's consider the simple example of Personal Finance

- Broadly speaking, Personal Finance involves the following aspects:
  - Receiving Money: Salary, Bonus, Rental income, Asset Liquidation etc.
  - Consuming Money: Food, Clothes, Rent/Mortgage, Car, Vacations etc.
  - Investing Money: Savings account, Stocks, Real-estate, Gold etc.
- Goal: Maximize lifetime-aggregated Expected Utility of Consumption
- This can be modeled as a Markov Decision Process
- *State:* Age, Asset Holdings, Asset Valuation, Career situation etc.
- *Action:* Changes in Asset Holdings, Optional Consumption
- *Reward:* Utility of Consumption of Money
- *Model:* Career uncertainties, Asset market uncertainties

Merton's Frictionless Continuous-Time Formulation

- Assume: Current wealth is $W_0 > 0$, and you'll live for T more years
- You can invest in (allocate to) n risky assets and a riskless asset
- Each risky asset has known normal distribution of returns
- Allowed to long or short any fractional quantities of assets
- Trading in continuous time $0 \leq t < T$, with no transaction costs
- You can consume any fractional amount of wealth at any time
- Dynamic Decision: Optimal Allocation and Consumption at each time
- To maximize lifetime-aggregated Expected Utility of Consumption
- Consumption Utility assumed to have Constant Relative Risk-Aversion

Problem Notation

For simplicity, we state and solve the problem for 1 risky asset but the solution generalizes easily to n risky assets.

- Riskless asset: $dR_t = r \cdot R_t \cdot dt$
- Risky asset: $dS_t = \mu \cdot S_t \cdot dt + \sigma \cdot S_t \cdot dz_t$ (i.e. Geometric Brownian)
- $\mu > r > 0, \sigma > 0$ (for n assets, we work with a covariance matrix)
- Wealth at time t is denoted by $W_t > 0$
- Fraction of wealth allocated to risky asset denoted by $\pi(t, W_t)$
- Fraction of wealth in riskless asset will then be $1 - \pi(t, W_t)$
- Wealth consumption per unit time denoted by $c(t, W_t) \geq 0$
- Utility of Consumption function $U(x) = \frac{x^{1-\gamma}}{1-\gamma}$ for $0 < \gamma \neq 1$
- Utility of Consumption function $U(x) = \log(x)$ for $\gamma = 1$
- $\gamma =$ (Constant) Relative Risk-Aversion $\frac{-x \cdot U''(x)}{U'(x)}$

Formal Problem Statement

- We write π_t, c_t instead of $\pi(t, W_t), c(t, W_t)$ to lighten notation
- Balance constraint implies the following process for Wealth W_t

$$dW_t = ((\pi_t \cdot (\mu - r) + r) \cdot W_t - c_t) \cdot dt + \pi_t \cdot \sigma \cdot W_t \cdot dz_t$$

- At any time t, determine optimal $[\pi(t, W_t), c(t, W_t)]$ to maximize:

$$\mathbb{E}[\int_t^T \frac{e^{-\rho(s-t)} \cdot c_s^{1-\gamma}}{1-\gamma} \cdot ds + \frac{e^{-\rho(T-t)} \cdot B(T) \cdot W_T^{1-\gamma}}{1-\gamma} \mid W_t]$$

- where $\rho \geq 0$ is the utility discount rate, $B(T)$ is the bequest function
- We can solve this problem for arbitrary bequest $B(T)$ but for simplicity, will consider $B(T) = \epsilon^\gamma$ where $0 < \epsilon \ll 1$, meaning "no bequest" (we need this ϵ-formulation for technical reasons).
- We will solve this problem for $\gamma \neq 1$ ($\gamma = 1$ is easier, hence omitted)

Continuous-Time Stochastic Control

- Think of this as a continuous-time Stochastic Control problem
- The *State* at time t is (t, W_t)
- The *Action* at time t is $[\pi_t, c_t]$
- The *Reward* per unit time at time t is $U(c_t) = \frac{c_t^{1-\gamma}}{1-\gamma}$
- The *Return* at time t is the accumulated discounted *Reward*:

$$\int_t^T e^{-\rho(s-t)} \cdot \frac{c_s^{1-\gamma}}{1-\gamma} \cdot ds$$

- Find *Policy* : $(t, W_t) \rightarrow [\pi_t, c_t]$ that maximizes the *Expected Return*
- Note: $c_t \geq 0$, but π_t is unconstrained

Optimal Value Function

- Value Function for a *State* (under a given policy) is the *Expected Return* from the *State* (when following the given policy)
- We focus on the Optimal Value Function $V^*(t, W_t)$

$$V^*(t, W_t) = \max_{\pi, c} \mathbb{E}_t[\int_t^T \frac{e^{-\rho(s-t)} \cdot c_s^{1-\gamma}}{1-\gamma} \cdot ds + \frac{e^{-\rho(T-t)} \cdot \epsilon^\gamma \cdot W_T^{1-\gamma}}{1-\gamma}]$$

- $V^*(t, W_t)$ satisfies a simple recursive formulation for $0 \leq t < t_1 < T$

$$V^*(t, W_t) = \max_{\pi, c} \mathbb{E}_t[\int_t^{t_1} \frac{e^{-\rho(s-t)} \cdot c_s^{1-\gamma}}{1-\gamma} \cdot ds + e^{-\rho(t_1-t)} \cdot V^*(t_1, W_{t_1})]$$

$$\Rightarrow e^{-\rho t} \cdot V^*(t, W_t) = \max_{\pi, c} \mathbb{E}_t[\int_t^{t_1} \frac{e^{-\rho s} \cdot c_s^{1-\gamma}}{1-\gamma} \cdot ds + e^{-\rho t_1} \cdot V^*(t_1, W_{t_1})]$$

HJB Equation for Optimal Value Function

Rewriting in stochastic differential form, we have the HJB formulation

$$\max_{\pi_t, c_t} \mathbb{E}_t[d(e^{-\rho t} \cdot V^*(t, W_t)) + \frac{e^{-\rho t} \cdot c_t^{1-\gamma}}{1-\gamma} \cdot dt] = 0$$

$$\Rightarrow \max_{\pi_t, c_t} \mathbb{E}_t[dV^*(t, W_t) + \frac{c_t^{1-\gamma}}{1-\gamma} \cdot dt] = \rho \cdot V^*(t, W_t) \cdot dt$$

Use Ito's Lemma on dV^*, remove the dz_t term since it's a martingale, and divide throughout by dt to produce the HJB Equation in PDE form:

$$\max_{\pi_t, c_t}[\frac{\partial V^*}{\partial t} + \frac{\partial V^*}{\partial W}((\pi_t(\mu - r) + r)W_t - c_t) + \frac{\partial^2 V^*}{\partial W^2} \cdot \frac{\pi_t^2 \sigma^2 W_t^2}{2} + \frac{c_t^{1-\gamma}}{1-\gamma}]$$
$$= \rho \cdot V^*(t, W_t)$$

Let us write the above equation more succinctly as:

$$\max_{\pi_t, c_t} \Phi(t, W_t; \pi_t, c_t) = \rho \cdot V^*(t, W_t)$$

Note: we are working with the constraints $W_t > 0, c_t \geq 0$ for $0 \leq t < T$

Optimal Allocation and Consumption

Find optimal π_t^*, c_t^* by taking partial derivatives of $\Phi(t, W_t; \pi_t, c_t)$ with respect to π_t and c_t, and equate to 0 (first-order conditions for Φ).

- Partial derivative of Φ with respect to π_t:

$$(\mu - r) \cdot \frac{\partial V^*}{\partial W_t} + \frac{\partial^2 V^*}{\partial W_t^2} \cdot \pi_t \cdot \sigma^2 \cdot W_t = 0$$

$$\Rightarrow \pi_t^* = \frac{-\frac{\partial V^*}{\partial W_t} \cdot (\mu - r)}{\frac{\partial^2 V^*}{\partial W_t^2} \cdot \sigma^2 \cdot W_t}$$

- Partial derivative of Φ with respect to c_t:

$$-\frac{\partial V^*}{\partial W_t} + (c_t^*)^{-\gamma} = 0$$

$$\Rightarrow c_t^* = (\frac{\partial V^*}{\partial W_t})^{\frac{-1}{\gamma}}$$

Optimal Value Function PDE

Now substitute π_t^* and c_t^* in $\Phi(t, W_t; \pi_t, c_t)$ and equate to $\rho V^*(t, W_t)$, which gets us the Optimal Value Function PDE:

$$\frac{\partial V^*}{\partial t} - \frac{(\mu - r)^2}{2\sigma^2} \cdot \frac{(\frac{\partial V^*}{\partial W_t})^2}{\frac{\partial^2 V^*}{\partial W_t^2}} + \frac{\partial V^*}{\partial W_t} \cdot r \cdot W_t + \frac{\gamma}{1-\gamma} \cdot (\frac{\partial V^*}{\partial W_t})^{\frac{\gamma-1}{\gamma}} = \rho V^*$$

The boundary condition is:

$$V^*(T, W_T) = \epsilon^\gamma \cdot \frac{W_T^{1-\gamma}}{1-\gamma}$$

The second-order conditions for Φ are satisfied **under the assumptions** $c_t^* > 0, W_t > 0, \frac{\partial^2 V^*}{\partial W_t^2} < 0$ for all $0 \leq t < T$ (we will later show that these are all satisfied in the solution we derive), and for concave $U(\cdot)$, i.e., $\gamma > 0$

Solving the PDE with a guess solution

We surmise with a guess solution

$$V^*(t, W_t) = f(t)^\gamma \cdot \frac{W_t^{1-\gamma}}{1-\gamma}$$

Then,

$$\frac{\partial V^*}{\partial t} = \gamma \cdot f(t)^{\gamma-1} \cdot f'(t) \cdot \frac{W_t^{1-\gamma}}{1-\gamma}$$

$$\frac{\partial V^*}{\partial W_t} = f(t)^\gamma \cdot W_t^{-\gamma}$$

$$\frac{\partial^2 V^*}{\partial W_t^2} = -f(t)^\gamma \cdot \gamma \cdot W_t^{-\gamma-1}$$

PDE reduced to an ODE

Substituting the guess solution in the PDE, we get the simple ODE:

$$f'(t) = \nu \cdot f(t) - 1$$

where

$$\nu = \frac{\rho - (1-\gamma) \cdot (\frac{(\mu-r)^2}{2\sigma^2\gamma} + r)}{\gamma}$$

with boundary condition $f(T) = \epsilon$.
The solution to this ODE is:

$$f(t) = \begin{cases} \frac{1+(\nu\epsilon-1)\cdot e^{-\nu(T-t)}}{\nu} & \text{for } \nu \neq 0 \\ T - t + \epsilon & \text{for } \nu = 0 \end{cases}$$

Optimal Allocation and Consumption

Putting it all together (substituting the solution for $f(t)$), we get:

$$\pi^*(t, W_t) = \frac{\mu - r}{\sigma^2 \gamma}$$

$$c^*(t, W_t) = \frac{W_t}{f(t)} = \begin{cases} \frac{\nu \cdot W_t}{1+(\nu\epsilon - 1)\cdot e^{-\nu(T-t)}} & \text{for } \nu \neq 0 \\ \frac{W_t}{T-t+\epsilon} & \text{for } \nu = 0 \end{cases}$$

$$V^*(t, W_t) = \begin{cases} \frac{(1+(\nu\epsilon-1)\cdot e^{-\nu(T-t)})^\gamma}{\nu^\gamma} \cdot \frac{W_t^{1-\gamma}}{1-\gamma} & \text{for } \nu \neq 0 \\ \frac{(T-t+\epsilon)^\gamma \cdot W_t^{1-\gamma}}{1-\gamma} & \text{for } \nu = 0 \end{cases}$$

- $f(t) > 0$ for all $0 \leq t < T$ (for all ν) ensures $W_t, c_t^* > 0, \frac{\partial^2 V^*}{\partial W_t^2} < 0$. This ensures the constraints $W_t > 0$ and $c_t \geq 0$ are satisfied and the second-order conditions for Φ are also satisfied.
- The HJB Formulation was key and this solution approach provides a template for similar continuous-time stochastic control problems.

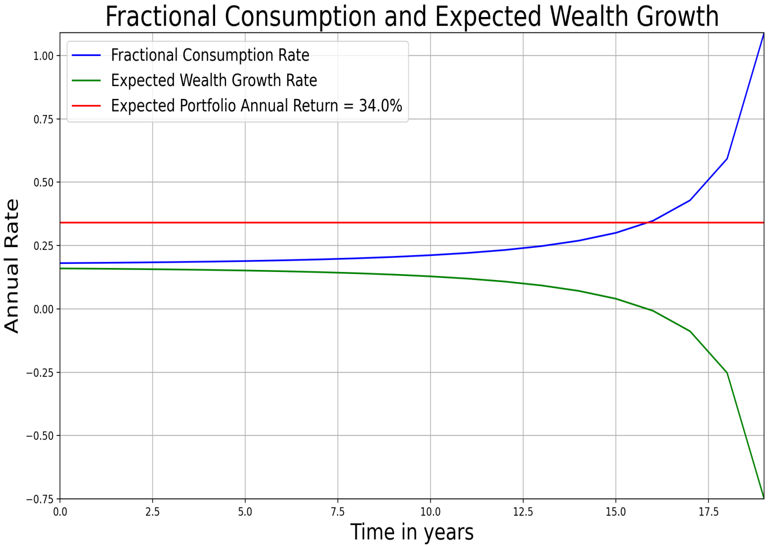
Gaining Insights into the Solution

- Optimal Allocation $\pi^*(t, W_t)$ is constant (independent of t and W_t)
- Optimal Fractional Consumption $\frac{c^*(t,W_t)}{W_t}$ depends only on t ($= \frac{1}{f(t)}$)
- With Optimal Allocation & Consumption, the Wealth process is:

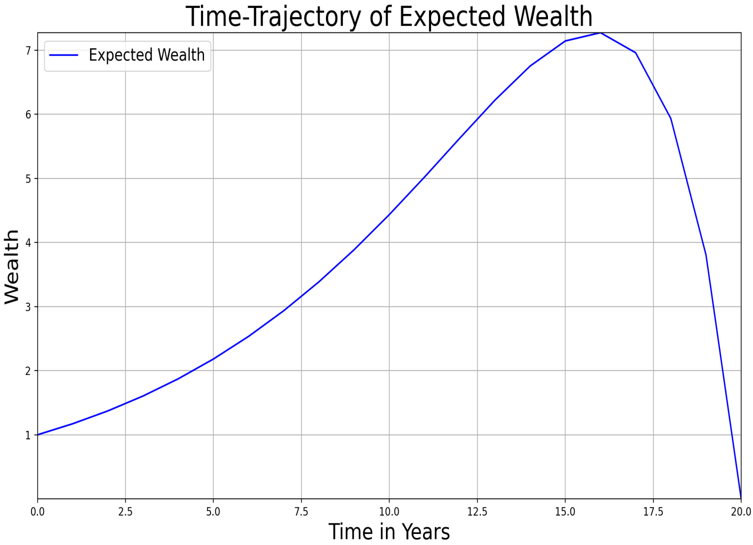
$$\frac{dW_t}{W_t} = (r + \frac{(\mu - r)^2}{\sigma^2 \gamma} - \frac{1}{f(t)}) \cdot dt + \frac{\mu - r}{\sigma \gamma} \cdot dz_t$$

- Expected Portfolio Return is constant over time ($= r + \frac{(\mu-r)^2}{\sigma^2\gamma}$)
- Assuming $\epsilon < \frac{1}{\nu}$, Fractional Consumption $\frac{1}{f(t)}$ increases over time
- Expected Rate of Wealth Growth $r + \frac{(\mu-r)^2}{\sigma^2\gamma} - \frac{1}{f(t)}$ decreases over time
- If $r + \frac{(\mu-r)^2}{\sigma^2\gamma} > \frac{1}{f(0)}$, we start by Consuming $<$ Expected Portfolio Growth and over time, we Consume $>$ Expected Portfolio Growth
- Wealth Growth Volatility is constant ($= \frac{\mu-r}{\sigma\gamma}$)

Fractional Consumption and Expected Wealth Growth



Time-Trajectory of Expected Wealth



Discrete-Time Asset-Allocation Example

- At time steps $t = 0, 1, \ldots, T-1$, we can asset-allocate wealth W_t
- 1 risky asset, unconstrained allocation, no transaction costs
- Risky asset return for each time step $\sim \mathcal{N}(\mu, \sigma^2)$
- Riskless asset has constant return r for each time step
- Assume no wealth consumption for any time $t < T$
- We liquidate and consume wealth W_T at time T
- Goal: Maximize Expected Utility of Wealth W_T at time T
- Dynamic allocation $x_t \in \mathbb{R}$ in risky asset, $W_t - x_t$ in riskless asset
- Utility of Wealth W_T at time T is given by CARA function:

$$U(W_T) = \frac{1 - e^{-aW_T}}{a} \text{ for some fixed } a \neq 0$$

- So we maximize, for each $t = 0, 1, \ldots, T-1$, over choices of $x_t \in \mathbb{R}$:

$$\mathbb{E}[\frac{-e^{-aW_T}}{a} | (t, W_t)]$$

MDP for Discrete-Time Asset-Allocation

- Continuous-States/Actions, Discrete-Time, Finite-Horizon MDP
- All states at time T are terminal states
- *State* $s_t \in \mathcal{S}_t$ is the wealth W_t, *Action* $a_t \in \mathcal{A}_t$ is risky investment x_t
- Deterministic policy at time t denoted as π_t, so $\pi_t(W_t) = x_t$
- Optimal deterministic policy at time t denoted as π_t^*, so $\pi_t^*(W_t) = x_t^*$
- Single-time-step return of risky asset from t to $t+1$ is $Y_t \sim \mathcal{N}(\mu, \sigma^2)$

$$W_{t+1} = x_t \cdot (1 + Y_t) + (W_t - x_t) \cdot (1 + r) = x_t \cdot (Y_t - r) + W_t \cdot (1 + r)$$

- MDP *Reward* is 0 for all $t = 0, 1, \ldots, T-1$
- MDP *Reward* at time T: $\frac{-e^{-aW_T}}{a}$
- MDP discount factor $\gamma = 1$

Optimal Value Function and Bellman Optimality Equation

- Denote Value Function at time t for policy $\pi = (\pi_0, \pi_1, \ldots, \pi_{T-1})$ as:

$$V_t^{\pi}(W_t) = \mathbb{E}_{\pi}[\frac{-e^{-aW_T}}{a} | (t, W_t)]$$

- Denote Optimal Value Function at time t as:

$$V_t^*(W_t) = \max_{\pi} V_t^{\pi}(W_t) = \max_{\pi}\{\mathbb{E}_{\pi}[\frac{-e^{-aW_T}}{a} | (t, W_t)]\}$$

- Bellman Optimality Equation is:

$$V_t^*(W_t) = \max_{x_t}\{\mathbb{E}_{Y_t \sim \mathcal{N}(\mu, \sigma^2)}[V_{t+1}^*(W_{t+1})]\}$$

$$V_{T-1}^*(W_{T-1}) = \max_{x_{T-1}}\{\mathbb{E}_{Y_{T-1} \sim \mathcal{N}(\mu, \sigma^2)}[\frac{-e^{-aW_T}}{a}]\}$$

- Make an educated guess for the functional form of the $V_t^*(W_t)$:

$$V_t^*(W_t) = -b_t \cdot e^{-c_t \cdot W_t}$$

where b_t, c_t are independent of the wealth W_t

Solving the Optimal Value Function

- We express Bellman Optimality Equation using this functional form:

$$
\begin{aligned}
V_t^*(W_t) &= \max_{x_t} \{\mathbb{E}_{Y_t \sim \mathcal{N}(\mu,\sigma^2)}[-b_{t+1} \cdot e^{-c_{t+1} \cdot W_{t+1}}]\} \\
&= \max_{x_t} \{\mathbb{E}_{Y_t \sim \mathcal{N}(\mu,\sigma^2)}[-b_{t+1} \cdot e^{-c_{t+1} \cdot (x_t \cdot (Y_t - r) + W_t \cdot (1+r))}]\} \\
&= \max_{x_t} \{-b_{t+1} \cdot e^{-c_{t+1} \cdot (1+r) \cdot W_t - c_{t+1} \cdot (\mu - r) \cdot x_t + c_{t+1}^2 \cdot \frac{\sigma^2}{2} \cdot x_t^2}\}
\end{aligned}
$$

- The partial derivative of term inside the max with respect to x_t is 0:

$$
-c_{t+1} \cdot (\mu - r) + \sigma^2 \cdot c_{t+1}^2 \cdot x_t^* = 0
$$

$$
\Rightarrow x_t^* = \frac{\mu - r}{\sigma^2 \cdot c_{t+1}} \tag{1}
$$

Solving the Optimal Value Function

- Next we substitute maximizing x_t^* in Bellman Optimality Equation:

$$V_t^*(W_t) = -b_{t+1} \cdot e^{-c_{t+1} \cdot (1+r) \cdot W_t - \frac{(\mu - r)^2}{2\sigma^2}}$$

- But since $V_t^*(W_t) = -b_t \cdot e^{-c_t \cdot W_t}$, we can write:

$$b_t = b_{t+1} \cdot e^{-\frac{(\mu - r)^2}{2\sigma^2}}, c_t = c_{t+1} \cdot (1 + r)$$

- We can calculate b_{T-1} and c_{T-1} from *Reward* $\frac{-e^{-aW_T}}{a}$

$$V_{T-1}^*(W_{T-1}) = \max_{x_{T-1}} \{\mathbb{E}_{Y_{T-1} \sim \mathcal{N}(\mu, \sigma^2)}[\frac{-e^{-aW_T}}{a}]\}$$

- Substituting for W_T, we get:

$$V_{T-1}^*(W_{T-1}) = \max_{x_{T-1}} \{\mathbb{E}_{Y_{T-1} \sim \mathcal{N}(\mu, \sigma^2)}[\frac{-e^{-a(x_{T-1} \cdot (Y_{T-1} - r) + W_{T-1} \cdot (1+r))}}{a}]\}$$

Solving the Optimal Value Function

- The expectation of this exponential (under normal distribution) is:

$$V^*_{T-1}(W_{T-1}) = \frac{-e^{-\frac{(\mu-r)^2}{2\sigma^2} - a \cdot (1+r) \cdot W_{T-1}}}{a}$$

- This gives us b_{T-1} and c_{T-1} as follows:

$$b_{T-1} = \frac{e^{-\frac{(\mu-r)^2}{2\sigma^2}}}{a}$$

$$c_{T-1} = a \cdot (1+r)$$

- Now we can unroll recursions for b_t and c_t:

$$b_t = \frac{e^{-\frac{(\mu-r)^2 \cdot (T-t)}{2\sigma^2}}}{a}$$

$$c_t = a \cdot (1+r)^{T-t}$$

Solving the Optimal Value Function

- Substituting the solution for c_{t+1} in (1) gives the Optimal Policy:

$$\pi_t^*(W_t) = x_t^* = \frac{\mu - r}{\sigma^2 \cdot a \cdot (1+r)^{T-t-1}}$$

- Note optimal action at time t does not depend on state W_t
- Hence, optimal policy $\pi_t^*(\cdot)$ is a constant deterministic policy function
- Substituting for b_t and c_t gives us the Optimal Value Function:

$$V_t^*(W_t) = \frac{-e^{-\frac{(\mu-r)^2(T-t)}{2\sigma^2}}}{a} \cdot e^{-a(1+r)^{T-t} \cdot W_t}$$

Real-World

- Analytical tractability in Merton's formulation was due to:
  - Normal distribution of asset returns
  - Constant Relative Risk-Aversion
  - Frictionless, continuous trading
- However, real-world situation involves:
  - Discrete amounts of assets to hold and discrete quantities of trades
  - Transaction costs
  - Locked-out days for trading
  - Non-stationary/arbitrary/correlated processes of multiple assets
  - Changing/uncertain risk-free rate
  - Consumption constraints
  - Arbitrary Risk-Aversion/Utility specification
- $\Rightarrow$ Approximate Dynamic Programming or Reinforcement Learning
- Large Action Space points to Policy Gradient Algorithms