Overview

1. Problem Statement

2. HJB Equation as Optimal Discounted Value Function PDE

3. Reducing the PDE to an ODE

4. Optimal Allocation and Consumption

5. Insights and Real-World Adaptation
You will live for (deterministic) $T$ more years

Current Wealth + PV of Future Income (less Debt) is $W_0 > 0$.

You can invest in (allocate to) $n$ risky assets and a riskless asset

Each 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 utility of consumption

Consumption Utility assumed to have constant Relative Risk-Aversion
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 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 $\gamma = 1$
- $\gamma = \text{(constant) Relative Risk-Aversion}$ $\frac{-x \cdot U''(x)}{U'(x)}$
Problem Statement

- We write $\pi_t, c_t$ 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:

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

- 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 $\epsilon$-formulation for technical reasons).

- We will solve this problem for $\gamma \neq 1$ ($\gamma = 1$ is easier, hence omitted).
Think of this as a continuous-time Stochastic Control problem
- The State is \((t, W_t)\)
- The Action is \([\pi_t, c_t]\)
- The Reward per unit time is \(U(c_t)\)
- The Return is the usual accumulated discounted Reward
- Find Policy \( (t, W_t) \rightarrow [\pi_t, c_t] \) that maximizes the Expected Return
- Note: \(c_t \geq 0\), but \(\pi_t\) is unconstrained
Optimal Discounted Value Function

- Instead of the usual Value Function (*Expected Return from a given State*), we consider the Discounted Value Function

- Discounted Value Function is simply the Value Function further discounted to time 0

- We focus on the Optimal Discounted Value Function $V^*(t, W_t)$

$$V^*(t, W_t) = \max_{\pi_t, c_t} E \left[ \int_t^T e^{-\rho s} \cdot \frac{c_s^{1-\gamma}}{1 - \gamma} \cdot ds + \frac{e^{-\rho T} \cdot \epsilon_{\gamma} \cdot W_T^{1-\gamma}}{1 - \gamma} \right]$$

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

$$V^*(t, W_t) = \max_{\pi_t, c_t} E \left[ V^*(t_1, W_{t_1}) + \int_t^{t_1} e^{-\rho s} \cdot \frac{c_s^{1-\gamma}}{1 - \gamma} \cdot ds \right]$$
Rewriting in stochastic differential form, we have the HJB formulation

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

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:

$$\max_{\pi_t, c_t} \left[ \frac{\partial V^*}{\partial t} + \frac{\partial V^*}{\partial W_t} ((\pi_t (\mu - r) + r) W_t - c_t) + \frac{\partial^2 V^*}{\partial W_t^2} \frac{\pi_t^2 \sigma^2 W_t^2}{2} + \frac{e^{-\rho t} \cdot c_t^{1-\gamma}}{1-\gamma} \right] = 0$$

Let us write the above equation more succinctly as:

$$\max_{\pi_t, c_t} \Phi(t, W_t; \pi_t, c_t) = 0$$

Note: we are working with the constraints $W_t > 0, c_t \geq 0$ for $0 \leq t < T$
Find optimal \( \pi^*_t, c^*_t \) by taking partial derivatives of \( \Phi(t, W_t; \pi_t, c_t) \) with respect to \( \pi_t \) and \( c_t \), and equate to 0 (first-order conditions for \( \Phi \)).

- Partial derivative of \( \Phi \) with respect to \( \pi_t \):

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

\[\Rightarrow \pi^*_t = -\frac{\partial V^*_t}{\partial W_t} \cdot (\mu - r) \frac{\partial^2 V^*_t}{\partial W_t^2} \cdot \sigma^2 \cdot W_t
\]

- Partial derivative of \( \Phi \) with respect to \( c_t \):

\[
-\frac{\partial V^*_t}{\partial W_t} + e^{-\rho t} \cdot (c_t^*)^{-\gamma} = 0
\]

\[\Rightarrow c^*_t = \left( \frac{\partial V^*_t}{\partial W_t} \cdot e^{\rho t} \right)^{-\frac{1}{\gamma}}
\]
Optimal Discounted Value Function PDE

Now substitute $\pi_t^*$ and $c_t^*$ in $\Phi(t, W_t; \pi_t, c_t)$ and set it to 0, which gets us the Optimal Discounted Value Function PDE:

$$\frac{\partial V^*}{\partial t} - \frac{(\mu - r)^2}{2\sigma^2} \cdot \left( \frac{\partial V^*}{\partial W_t} \right)^2 + \frac{\partial V^*}{\partial W_t} \cdot r \cdot W_t + \frac{\gamma}{1 - \gamma} \cdot e^{-\rho t} \cdot \left( \frac{\partial V^*}{\partial W_t} \right)^{\gamma - 1} = 0$$

The boundary condition is:

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

The second-order conditions for $\Phi$ 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., $\gamma > 0$. 

Ashwin Rao (Stanford)  
HJB and Merton Portfolio  
January 2, 2019 10 / 15
We surmise with a guess solution

\[ V^*(t, W_t) = f(t)\gamma \cdot e^{-\rho t} \cdot \frac{W_t^{1-\gamma}}{1-\gamma} \]

Then,

\[ \frac{\partial V^*}{\partial t} = (\gamma \cdot f(t)^{\gamma-1} \cdot f'(t) - \rho \cdot f(t)^{\gamma}) \cdot e^{-\rho t} \cdot \frac{W_t^{1-\gamma}}{1-\gamma} \]

\[ \frac{\partial V^*}{\partial W_t} = f(t)^{\gamma} \cdot e^{-\rho t} \cdot W_t^{-\gamma} \]

\[ \frac{\partial^2 V^*}{\partial W_t^2} = -f(t)^{\gamma} \cdot e^{-\rho t} \cdot \gamma \cdot W_t^{\gamma-1} \]
Substituting the guess solution in the PDE, we get the simple ODE:

\[ f'(t) = \nu \cdot f(t) - 1 \]

where

\[ \nu = \frac{\rho - (1 - \gamma) \cdot \left( \frac{(\mu - r)^2}{2\sigma^2\gamma} + r \right)}{\gamma} \]

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:

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

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

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

- $f(t) > 0$ for all $0 \leq t < T$ (for all $\nu$) 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 $\Phi$ 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:

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

- 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 $< \text{Expected Portfolio Growth}$
- Wealth Growth Volatility is constant ($= \frac{\mu - r}{\sigma \gamma}$)
Expected Portfolio Return and Fractional Consumption Rate

Annual Rate

Expected Portfolio Return

Fractional Consumption Rate

t (in years)
Porting this to Real-World Portfolio Optimization

- 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

⇒ Approximate Dynamic Programming or Reinforcement Learning

- Large Action Space points to Policy Gradient Algorithms