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

Ashwin Rao

ICME, Stanford University

January 26, 2022
Dynamic Asset-Allocation and Consumption

The broad topic is Investment Management. It applies to corporations as well as individuals. The two considerations are:

1. How to allocate money across assets in one's investment portfolio.
2. How much to consume for one's needs, operations, or pleasures.

We consider the dynamic version of these dual considerations, focusing on 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.
The broad topic is Investment Management
The broad topic is Investment Management

Applies to Corporations as well as Individuals
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
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
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
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
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
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
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
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
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
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
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
Broadly speaking, Personal Finance involves the following aspects:

- Receiving Money: Salary, Bonus, Rental income, Asset Liquidation etc.

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
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.

- 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
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.
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
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
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.
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
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
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
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.
Merton’s Frictionless Continuous-Time Formulation

- Assume: Current wealth is $W_0 > 0$, and you’ll live for $T$ more years
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
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
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
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
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.
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
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
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.
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) for $\mu > r > 0$, $\sigma > 0$ (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 $U(x) = x^{1-\gamma}$ for $0 < \gamma \neq 1$

Utility of Consumption function $U(x) = \log(x)$ for $\gamma = 1$

$\gamma$ (Constant) Relative Risk-Aversion

$-x \cdot U''(x) \cdot U'(x)$
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)

\( \mu > r > 0 \) (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 \( U(x) = x^{1-\gamma} \) for \( 0 < \gamma \neq 1 \)

\( U(x) = \log(x) \) for \( \gamma = 1 \)

\( \gamma \) = (Constant) Relative Risk-Aversion

\(- x \cdot U''(x) \)

\( U'(x) \)
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)
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 \( W_t > 0 \)
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 \( \pi(t, W_t) \)
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 $1 - \pi(t, W_t)$
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 \( c(t, W_t) \geq 0 \)
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 $0 < \gamma \neq 1$
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 $\gamma = 1$
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 $\pi_t, c_t$ instead of $\pi(t, W_t), c(t, W_t)$ to lighten notation.

Balance constraint implies the following process for Wealth $W_t$:

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

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

$$E\left[\int_t^T e^{-\rho(s-t)} \cdot c_1 - \gamma s_1 - \gamma \cdot ds + e^{-\rho(T-t)} \cdot B(T) \cdot W_1 - \gamma T_1 - \gamma \right]$$

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).
We write $\pi_t, c_t$ instead of $\pi(t, W_t), c(t, W_t)$ to lighten notation.
Formal 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 $W_t$

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

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

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

- where $\rho \geq 0$ is the utility discount rate, $B(T)$ is the bequest function.
Formal 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 $W_t$

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

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

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

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

  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 $\epsilon$-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
Think of this as a continuous-time Stochastic Control problem

- The *State* at time $t$ is $(t, W_t)$
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 $[\pi_t, c_t]$
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 $U(c_t) = \frac{c_t^{1-\gamma}}{1-\gamma}$
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:

$$\int_t^T e^{-\rho(s-t)} \cdot \frac{c_s^{1-\gamma}}{1-\gamma} \cdot ds$$
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
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 $\pi_t$ is unconstrained.
Value Function for a *State* (under a given policy) is the *Expected Return* from the *State* (when following the given policy)
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 \( V^*(t, W_t) \)

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

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

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

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

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

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

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

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

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

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

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

$$\max_{\pi_t, c_t} E_t \left[ d\left( e^{-\rho t} \cdot V^* (t, W_t) \right) + e^{-\rho t} \cdot c_1 - \gamma \cdot dt \right] = 0$$

$$\Rightarrow \max_{\pi_t, c_t} E_t \left[ dV^* (t, W_t) + c_1 - \gamma \cdot dt \right] = \rho \cdot V^* (t, W_t)$$

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:

$$\max_{\pi_t, c_t} \left[ \frac{\partial V^*}{\partial t} + \frac{\partial V^*}{\partial W_t} \left( \pi_t (\mu - \nu + \nu) W_t - c_t \right) + \frac{\partial^2 V^*}{\partial W_t^2} \cdot \pi_t^2 \sigma^2 W_t^2 + c_1 - \gamma \right] = \rho \cdot V^* (t, W_t)$$

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 $0 \leq t < T$.
Rewriting in stochastic differential form, we have the HJB formulation

\[
\max_{\pi_t, c_t} \mathbb{E}_t [d(e^{-\rho t} \cdot V^*(t, W_t)) + \frac{e^{-\rho t} \cdot c_t^{1-\gamma}}{1-\gamma} \cdot dt] = 0
\]
Rewriting in stochastic differential form, we have the HJB formulation

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

\[
\Rightarrow \max_{\pi_t, c_t} \mathbb{E}_t \left[ dV^*(t, W_t) + \frac{c_t^{1-\gamma}}{1-\gamma} \cdot dt \right] = \rho \cdot V^*(t, W_t) \cdot dt
\]
HJB Equation for Optimal Value Function

Rewriting in stochastic differential form, we have the HJB formulation

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

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

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:

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

\[= \rho \cdot V^*(t, W_t)\]
Rewriting in stochastic differential form, we have the HJB formulation

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

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

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:

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

Let us write the above equation more succinctly as:

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

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

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 \(0 \leq t < T\)
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 $\Phi$).
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 $\pi_t$:

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

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

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

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

- Partial derivative of $\Phi$ with respect to $c_t$:

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

$$
\Rightarrow c^*_t = (\frac{\partial V^*}{\partial W_t})^{-\frac{1}{\gamma}}
$$
Now substitute $\pi^* t$ 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:

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

The boundary condition is:

$$V^*(T, W_T) = \epsilon \gamma W_1 - T \gamma$$

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)  
Asset-Allocation Chapter  
January 26, 2022
Now substitute $\pi_t^*$ 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:
Now substitute $\pi_t^*$ 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:

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

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

The boundary condition is:

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

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

The boundary condition is:

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

The second-order conditions for $\Phi$ 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., $\gamma > 0$.
Solving the PDE with a guess solution

We surmise with a guess solution

\[ V^* (t, W_t) = f(t) \gamma \cdot W_1 - \gamma t \]

Then,

\[ \frac{\partial V^*}{\partial t} = \gamma \cdot f(t) - \gamma^{-1} f'(t) \cdot W_1 - \gamma t \]

\[ \frac{\partial V^*}{\partial W_t} = f(t) \gamma \cdot W - \gamma t \frac{\partial^2 V^*}{\partial W_t^2} = -f(t) \gamma \cdot \gamma \cdot W - \gamma^{-1} t \]
Solving the PDE with a guess solution

We surmise with a guess solution

\[ V^*(t, W_t) = f(t)^\gamma \cdot \frac{W_t^{1-\gamma}}{1 - \gamma} \]
We surmise with a guess solution

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

Then,

\[ \frac{\partial V^*}{\partial t} = \gamma \cdot f(t)^{\gamma-1} \cdot f'(t) \cdot \frac{W_t^{1-\gamma}}{1 - \gamma} \]
We surmise with a guess solution

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

Then,

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

\[ \frac{\partial V^*}{\partial W_t} = f(t)^\gamma \cdot W_t^{-\gamma} \]
We surmise with a guess solution

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

Then,

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

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

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

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

where

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

with boundary condition

$$f(T) = \epsilon.$$
Substituting the guess solution in the PDE, we get the simple ODE:

$$f'(t) = \nu \cdot f(t) - 1$$
Substituting the guess solution in the PDE, we get the simple ODE:

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

where

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

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

with boundary condition \( f(T) = \epsilon \).

The solution to this ODE is:

\[ f(t) = \begin{cases} 
\frac{1+(\nu\epsilon-1)e^{-\nu(T-t)}}{\nu} & \text{for } \nu \neq 0 \\
T - t + \epsilon & \text{for } \nu = 0 
\end{cases} \]
Putting it all together (substituting the solution for \( f(t) \)), we get:

\[
\pi^* (t, W_t) = \mu - r \sigma^2 \gamma \nu^* (t, W_t) = W_t f(t) = \begin{cases} 
\nu \cdot W_t^{1+\left(\nu \epsilon - 1\right) \cdot e^{-\nu \cdot (T-t)}} & \text{for } \nu \neq 0 \\
T - t + \epsilon & \text{for } \nu = 0
\end{cases}
\]

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

\( f(t) > 0 \) for all \( 0 \leq t < T \) (for all \( \nu \)) ensures \( W_t, c^* t > 0 \), \( \frac{\partial^2 V^*}{\partial W^2} t < 0 \).

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

$$\pi^*(t, W_t) = \frac{\mu - r}{\sigma^2 \gamma}$$
Putting it all together (substituting the solution for \( f(t) \)), we get:

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

\[
c^*(t, W_t) = \frac{W_t}{f(t)} = \begin{cases} 
\frac{\nu \cdot W_t}{1 + (\nu \epsilon - 1) e^{-\nu (T-t)}} & \text{for } \nu \neq 0 \\
\frac{W_t}{T-t+\epsilon} & \text{for } \nu = 0
\end{cases}
\]
Putting it all together (substituting the solution for $f(t)$), we get:

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

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

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

\[ c^* (t, W_t) W_t \] depends only on \( t = 1 / f(t) \)

With Optimal Allocation & Consumption, the Wealth process is:

\[ dW^* t W^* t = \left( r + \left( \mu - r \right)^2 \sigma^2 \gamma - 1 / f(t) \right) \cdot dt + \mu - r \sigma \gamma \cdot dz \]

Expected Portfolio Return is constant over time (\( = r + \left( \mu - r \right)^2 \sigma^2 \gamma \))

Assuming \( \epsilon < 1 / \nu \), Fractional Consumption \( 1 / f(t) \) increases over time

Expected Rate of Wealth Growth \( r + \left( \mu - r \right)^2 \sigma^2 \gamma - 1 / f(t) \) decreases over time

If \( r + \left( \mu - r \right)^2 \sigma^2 \gamma > 1 / f(0) \), we start by Consuming < Expected Portfolio Growth and over time, we Consume > Expected Portfolio Growth

Wealth Growth Volatility is constant (\( = \mu - r \sigma \gamma \))
Gaining Insights into the Solution

- Optimal Allocation $\pi^*(t, W_t)$ is constant (independent of $t$ and $W_t$)
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 $t$ ($= \frac{1}{f(t)}$)
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:

\[
\frac{dW_t^*}{W_t^*} = \left( r + \frac{(\mu - r)^2}{\sigma^2 \gamma} - \frac{1}{f(t)} \right) \cdot dt + \frac{\mu - r}{\sigma \gamma} \cdot dz_t
\]
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 ($= r + \frac{(\mu - r)^2}{\sigma^2 \gamma}$)
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
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
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 $t \left(= \frac{1}{f(t)} \right)$
- 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 $<$ Expected Portfolio Growth and over time, we Consume $>$ Expected Portfolio Growth
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 < Expected Portfolio Growth and over time, we Consume > Expected Portfolio Growth
- Wealth Growth Volatility is constant ($= \frac{\mu - r}{\sigma \gamma}$)
Discrete-Time Asset-Allocation Example

At time steps $t = 0, 1, \ldots, T-1$, we can asset-allocate wealth $W_t$ to a risky asset, unconstrained allocation, no transaction costs. The risky asset return for each time step $\sim N(\mu, \sigma^2)$. The 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 a risky asset, $W_t - x_t$ in a riskless asset.

Utility of Wealth $W_T$ at time $T$ is given by CARA function:

$$U(W_T) = 1 - e^{-aW_T}$$

for some fixed $a \neq 0$.

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

$$E[-e^{-aW_T} | (t, W_t)]$$
Discrete-Time Asset-Allocation Example

- At time steps $t = 0, 1, \ldots, T - 1$, we can asset-allocate wealth $W_t$
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
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 $\sim \mathcal{N}(\mu, \sigma^2)$

---

Ashwin Rao  (Stanford)  
Asset-Allocation Chapter  
January 26, 2022  18 / 27
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
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 $t < T$
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 $T$
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 $T$
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
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:

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

Ashwin Rao (Stanford)  Asset-Allocation Chapter  January 26, 2022  18 / 27
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:
  \[
  U(W_T) = \frac{1 - e^{-aW_T}}{a} \quad \text{for some fixed } a \neq 0
  \]
- So we maximize, for each \( t = 0, 1, \ldots, T - 1 \), over choices of \( x_t \in \mathbb{R} \):
  \[
  \mathbb{E}\left[ \frac{-e^{-aW_T}}{a} \mid (t, W_t) \right]
  \]
MDP for Discrete-Time Asset-Allocation

All states at time $T$ are terminal states.

State $s_t \in S$ is the wealth $W_t$.

Action $a_t \in A_t$ is risky investment $x_t$.

Deterministic policy at time $t$ denoted as $\pi_t$, so $\pi_t(W_t) = x_t$.

Optimal deterministic policy at time $t$ denoted as $\pi^*_t$, so $\pi^*_t(W_t) = x^*_t$.

Single-time-step return of risky asset from $t$ to $t+1$ is $Y_t \sim N(\mu, \sigma^2)$.

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

MDP reward is 0 for all $t = 0, 1, \ldots, T-1$.

MDP discount factor $\gamma = 1$. 
MDP for Discrete-Time Asset-Allocation

- Continuous-States/Actions, Discrete-Time, Finite-Horizon MDP
Continuous-States/Actions, Discrete-Time, Finite-Horizon MDP
All states at time $T$ are terminal states
Continuous-States/Actions, Discrete-Time, Finite-Horizon MDP

All states at time $T$ are terminal states

$State \ s_t \in S_t$ is the wealth $W_t$, $Action \ a_t \in A_t$ is risky investment $x_t$
Continuous-States/Actions, Discrete-Time, Finite-Horizon MDP

All states at time $T$ are terminal states

State $s_t \in S_t$ is the wealth $W_t$, Action $a_t \in A_t$ is risky investment $x_t$

Deterministic policy at time $t$ denoted as $\pi_t$, so $\pi_t(W_t) = x_t$
MDP for Discrete-Time Asset-Allocation

- Continuous-States/Actions, Discrete-Time, Finite-Horizon MDP
- All states at time $T$ are terminal states
- State $s_t \in S_t$ is the wealth $W_t$, Action $a_t \in A_t$ is risky investment $x_t$
- Deterministic policy at time $t$ denoted as $\pi_t$, so $\pi_t(W_t) = x_t$
- Optimal deterministic policy at time $t$ denoted as $\pi^*_t$, so $\pi^*_t(W_t) = x^*_t$
Continuous-States/Actions, Discrete-Time, Finite-Horizon MDP

All states at time $T$ are terminal states

*State $s_t \in S_t$ is the wealth $W_t$, Action $a_t \in A_t$ is risky investment $x_t$*

Deterministic policy at time $t$ denoted as $\pi_t$, so $\pi_t(W_t) = x_t$

Optimal deterministic policy at time $t$ denoted as $\pi^*_t$, 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 for Discrete-Time Asset-Allocation

- Continuous-States/Actions, Discrete-Time, Finite-Horizon MDP
- All states at time $T$ are terminal states
- *State* $s_t \in S_t$ is the wealth $W_t$, *Action* $a_t \in A_t$ is risky investment $x_t$
- Deterministic policy at time $t$ denoted as $\pi_t$, so $\pi_t(W_t) = x_t$
- Optimal deterministic policy at time $t$ denoted as $\pi_t^*$, 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 $t = 0, 1, \ldots, T - 1$
Continuous-States/Actions, Discrete-Time, Finite-Horizon MDP

All states at time $T$ are terminal states

*State* $s_t \in S_t$ is the wealth $W_t$, *Action* $a_t \in A_t$ is risky investment $x_t$

Deterministic policy at time $t$ denoted as $\pi_t$, so $\pi_t(W_t) = x_t$

Optimal deterministic policy at time $t$ denoted as $\pi_t^*$, 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 for Discrete-Time Asset-Allocation

- Continuous-States/Actions, Discrete-Time, Finite-Horizon MDP
- All states at time $T$ are terminal states
- *State* $s_t \in S_t$ is the wealth $W_t$, *Action* $a_t \in A_t$ is risky investment $x_t$
- Deterministic policy at time $t$ denoted as $\pi_t$, so $\pi_t(W_t) = x_t$
- Optimal deterministic policy at time $t$ denoted as $\pi^*_t$, 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:

$$V_{\pi t}(W_t) = E_{\pi}\left[-e^{-a W_{t+1}} \mid (t, W_t)\right]$$

Denote Optimal Value Function at time $t$ as:

$$V^*_{t}(W_t) = \max_{\pi} V_{\pi t}(W_t) = \max_{\pi}\{E_{\pi}\left[-e^{-a W_{t+1}} \mid (t, W_t)\right]\}$$

Bellman Optimality Equation is:

$$V^*_{t}(W_t) = \max_x\{E_{Y_t \sim N(\mu, \sigma^2)}\left[V^*_{t+1}(W_{t+1})\right]\}$$

$$V^*_{T-1}(W_{T-1}) = \max_x\{E_{Y_{T-1} \sim N(\mu, \sigma^2)}\left[-e^{-a W_{T-1}}\right]\}$$

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

$$V^\pi_t(W_t) = \mathbb{E}_\pi[\frac{-e^{-aW_T}}{a}|(t, W_t)]$$
Denote Value Function at time $t$ for policy $\pi = (\pi_0, \pi_1, \ldots, \pi_{T-1})$ as:

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

Denote Optimal Value Function at time $t$ as:

$$V^*_t(W_t) = \max_\pi V_\pi^t(W_t) = \max_\pi \{\mathbb{E}_\pi[-e^{-aW_T}/a | (t, W_t)]\}$$
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:

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

- Denote Optimal Value Function at time $t$ as:

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

- Bellman Optimality Equation is:

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

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

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

- Denote Optimal Value Function at time $t$ as:
  $$V_t^*(W_t) = \max_{\pi} V_t^{\pi}(W_t) = \max_{\pi}\left\{\mathbb{E}_\pi\left[\frac{-e^{-aW_T}}{a}|(t, W_t)\right]\right\}$$

- Bellman Optimality Equation is:
  $$V_t^*(W_t) = \max_{x_t} \left\{\mathbb{E}_{Y_t \sim \mathcal{N}(\mu, \sigma^2)}[V_{t+1}^*(W_{t+1})]\right\}$$

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

- 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:

\[ V^*_{t}(Y_t) = \max_{x_t} \{ E[Y_t \sim N(\mu, \sigma^2)] \cdot \left[ -b_{t+1} \cdot e^{-c_{t+1} \cdot (x_t \cdot (Y_t - r) + W_t \cdot (1 + r))} \right] \} \]

The partial derivative of term inside the max with respect to \( x_t \) is 0:

\[-c_{t+1} \cdot (\mu - r) + \sigma^2 \cdot c_{t+1} \cdot x^*_{t} = 0 \]

\[ \Rightarrow x^*_{t} = \frac{\mu - r}{\sigma^2 \cdot c_{t+1}} \]
We express Bellman Optimality Equation using this functional form:

\[
V^*_t(W_t) = \max_{x_t} \{ \mathbb{E}_{Y_t \sim \mathcal{N}(\mu, \sigma^2)} \left[ -b_{t+1} \cdot e^{-c_{t+1} \cdot W_{t+1}} \right] \}
\]

\[
= \max_{x_t} \{ \mathbb{E}_{Y_t \sim \mathcal{N}(\mu, \sigma^2)} \left[ -b_{t+1} \cdot e^{-c_{t+1} \cdot (x_t \cdot (Y_t - r) + W_t \cdot (1+r))} \right] \}
\]

\[
= \max_{x_t} \{ -b_{t+1} \cdot e^{-c_{t+1} \cdot (1+r) \cdot W_t - c_{t+1} \cdot (\mu - r) \cdot x_t + c_{t+1}^2 \cdot \frac{\sigma^2}{2} \cdot x_t^2} \}
\]
We express Bellman Optimality Equation using this functional form:

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

\[ = \max_{x_t} \{ \mathbb{E}_{Y_t \sim N(\mu, \sigma^2)} [-b_{t+1} \cdot e^{-c_{t+1} \cdot (x_t \cdot (Y_t - r) + W_t \cdot (1+r))] \} \]

\[ = \max_{x_t} \{-b_{t+1} \cdot e^{-c_{t+1} \cdot (1+r) \cdot W_t - c_{t+1} \cdot (\mu - r) \cdot x_t + c_{t+1}^2 \cdot \frac{\sigma^2}{2} \cdot x_t^2} \} \]

The partial derivative of term inside the max with respect to \( x_t \) is 0:

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

\[ \Rightarrow x_t^* = \frac{\mu - r}{\sigma^2 \cdot c_{t+1}} \] (1)
Next we substitute maximizing $x^*$ in Bellman Optimality Equation:

$$V^*(t)(W_t) = -b_t \cdot e^{-c_t} \cdot W_t - (\mu - r)^2 \sigma^2,$$

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

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

$$c_t = c_{t+1} \cdot (1 + r).$$

We can calculate $b_{T-1}$ and $c_{T-1}$ from Reward:

$$V^*_T(W_{T-1}) = \max_{x_{T-1}} \{ E_{Y_{T-1} \sim N(\mu, \sigma^2)}[-e^{-a(W_{T-1} \cdot Y_{T-1} - r)} + W_{T-1} \cdot (1 + r)] \}.$$
Solving the Optimal Value Function

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

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

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

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

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

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

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

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

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

- Substituting for $W_T$, we get:

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

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

\[ V^* T - 1 \left( W^T - 1 \right) = e^{-\frac{(\mu - r)^2}{2\sigma^2} - a \cdot (1 + r) \cdot W^T - 1} \]

This gives us \( b^T - 1 \) and \( c^T - 1 \) as follows:

\[ b^T - 1 = e^{-\frac{(\mu - r)^2}{2\sigma^2} \cdot (T - t)} \]

\[ c^T - 1 = a \cdot (1 + r) \cdot (T - t) \]

Now we can unroll recursions for \( b_t \) and \( c_t \):

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

\[ c_t = a \cdot (1 + r) \cdot (T - t) \]
The expectation of this exponential (under normal distribution) is:

\[ V^*_{T-1}(W_{T-1}) = -\frac{e^{-\frac{(\mu - r)^2}{2\sigma^2}} - a(1+r)W_{T-1}}{a} \]
Solving the Optimal Value Function

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

\[ V_{T-1}^*(W_{T-1}) = \frac{-e^{-(\mu - r)^2/2\sigma^2} - a \cdot (1+r) \cdot W_{T-1}}{a} \]

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

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

\[ c_{T-1} = a \cdot (1 + r) \]
The expectation of this exponential (under normal distribution) is:

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

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:

$$\pi^*_t(W_t) = x^*_t = \mu - r \sigma^2 \cdot a \cdot (1 + r)^{T - t - 1}$$

Note optimal action at time $t$ does not depend on state $W_t$.

Hence, optimal policy $\pi^*_t(W_t)$ is a constant deterministic policy function.

Substituting for $b_t$ and $c_t$ gives us the Optimal Value Function:

$$V^*_t(W_t) = -e^{-\frac{(\mu - r)^2}{\sigma^2}(T - t)} \cdot e^{-a \cdot (1 + r)^{T - t} \cdot W_t}$$
Solving the Optimal Value Function

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

$$\pi^*_t(W_t) = x^*_t = \frac{\mu - r}{\sigma^2 \cdot a \cdot (1 + r)^{T-t-1}}$$
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 $W_t$
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 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:

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

- Analytical tractability in Merton’s formulation was due to:
  - Normal distribution of asset returns
Real-World

Analytical tractability in Merton’s formulation was due to:
- Normal distribution of asset returns
- Constant Relative Risk-Aversion

However, real-world situation involves:
- Discrete amounts of assets to hold and discrete quantities of trades
- Transaction costs
- Locked-out days for trading
- Time-Heterogeneous/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
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
- Time-Heterogeneous/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
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
- Time-Heterogeneous/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
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

⇒ Approximate Dynamic Programming or Reinforcement Learning
Large Action Space points to Policy Gradient Algorithms
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

⇒ Approximate Dynamic Programming or Reinforcement Learning
Large Action Space points to Policy Gradient Algorithms
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
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
  - Time-Heterogeneous/arbitrary/correlated processes of multiple assets
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
  - Time-Heterogeneous/arbitrary/correlated processes of multiple assets
  - Changing/uncertain risk-free rate

⇒ Approximate Dynamic Programming or Reinforcement Learning
Large Action Space points to Policy Gradient Algorithms
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
- Time-Heterogeneous/arbitrary/correlated processes of multiple assets
- Changing/uncertain risk-free rate
- Consumption constraints
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
  - Time-Heterogeneous/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
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
- Time-Heterogeneous/arbitrary/correlated processes of multiple assets
- Changing/uncertain risk-free rate
- Consumption constraints
- Arbitrary Risk-Aversion/Utility specification

⇒ Approximate Dynamic Programming or Reinforcement Learning
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
  - Time-Heterogeneous/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
Code for Discrete-time Dynamic Asset-Allocation

class AssetAllocDiscrete:
    risky returns: Sequence [ Distribution[float] ]
    riskless returns: Sequence [ float ]
    utility fun: Callable[[float, float]]
    risky choices: Sequence [ float ]
    feature functions: Sequence [ Callable[[Tuple[float, float]], float] ]
    dnn spec: DNNSpec
    initial wealth distribution: Distribution[float]

Ashwin Rao (Stanford)
@dataclass(frozen=True)
class AssetAllocDiscrete:
    risky_return_distributions: \n        Sequence[Distribution[float]]
riskless_returns: Sequence[float]
utility_func: Callable[[float], float]
risky_alloc_choices: Sequence[float]
feature_functions: \n    Sequence[Callable[[Tuple[float, float]], float]]
dnn_spec: DNNSpec
initial_wealth_distribution: Distribution[float]
Key Takeaways from this Chapter

Merton, in his landmark 1969 paper, provided an elegant closed-form solution under simplifying assumptions in continuous-time. In real-world, we need to model this problem as an MDP (capturing various frictions/constraints), and solve with ADP/RL.
Key Takeaways from this Chapter

- Fundamental problem of jointly deciding on optimal investment allocation and optimal consumption
Key Takeaways from this Chapter

- Fundamental problem of jointly deciding on optimal investment allocation and optimal consumption
- Merton, in his landmark 1969 paper, provided an elegant closed-form solution under simplifying assumptions in continuous-time
Key Takeaways from this Chapter

- Fundamental problem of jointly deciding on optimal investment allocation and optimal consumption
- Merton, in his landmark 1969 paper, provided an elegant closed-form solution under simplifying assumptions in continuous-time
- In real-world, we need to model this problem as an MDP (capturing various frictions/constraints), and solve with ADP/RL