Advanced Portfolio Theory: Mathematical Foundations and Modern Developments

A Rigorous Treatment for Mathematicians and Statisticians

Portfolio Theory Course

Advanced Mathematics Department

June 12, 2025

Table of Contents

Table of Contents

Course Objectives

Primary Goals

- Rigorous mathematical treatment of portfolio theory
- Bridge classical results with modern developments
- Emphasis on measure-theoretic foundations
- Applications of advanced optimization and statistical methods

Prerequisites

- Measure theory and functional analysis
- Stochastic processes and martingale theory
- Convex optimization
- Statistical inference and asymptotic theory

Historical Context and Modern Relevance

Timeline of Portfolio Theory:

- **1952**: Markowitz mean-variance theory
- **1964**: CAPM (Sharpe, Lintner, Mossin)
- **1973**: Merton's continuous-time theory
- **1979**: Harrison-Kreps fundamental theorems
- **1990s**: Coherent risk measures
- **2000s**: High-dimensional methods
- **2010s**: Machine learning integration

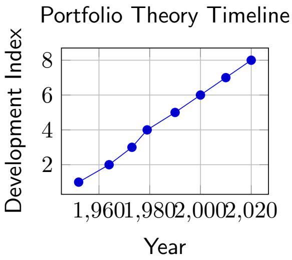


Table of Contents

Probability Spaces and Filtrations

Definition (Filtered Probability Space)

A **filtered probability space** is a tuple $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P})$ where:

- $(\Omega, \mathcal{F}, \mathbb{P})$ is a complete probability space
- $\{\mathcal{F}_t\}_{t\geq 0}$ is a right-continuous filtration
- $\mathcal{F}_s \subseteq \mathcal{F}_t \subseteq \mathcal{F}$ for $0 \leq s \leq t$

Usual Conditions

1. **Right-continuity**: $\mathcal{F}_t = \bigcap_{s>t} \mathcal{F}_s$
2. **Completeness**: $\mathcal{F}_0$ contains all $\mathbb{P}$-null sets

Stochastic Processes: Definitions

Definition (Adapted Process)

A stochastic process $X = \{X_t\}_{t \geq 0}$ is **adapted** to $\{\mathcal{F}_t\}$ if X_t is $\mathcal{F}_t$-measurable for all $t \geq 0$.

Definition (Predictable Process)

A process X is **predictable** if it is measurable with respect to the predictable σ-algebra on $\Omega \times [0, \infty)$.

Theorem (Doob-Meyer Decomposition)

Every right-continuous adapted submartingale Y can be uniquely decomposed as:

$$Y_t = M_t + A_t$$

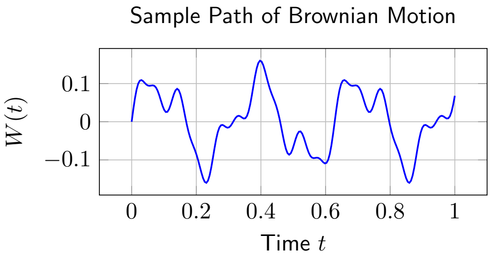
where M is a right-continuous martingale and A is a predictable increasing process.

Brownian Motion

Definition (Standard Brownian Motion)

A process $W = \{W_t\}_{t \geq 0}$ is standard Brownian motion if:

1. $W_0 = 0$ almost surely
2. W has independent increments
3. $W_t - W_s \sim \mathcal{N}(0, t - s)$ for $0 \leq s < t$
4. W has continuous paths



Itô Calculus

Theorem (Itô's Lemma)

Let X_t *satisfy* $dX_t = \mu_t dt + \sigma_t dW_t$ *and* $f \in C^{1,2}([0,T] \times \mathbb{R})$. *Then:*

$$df(t, X_t) = \left(\frac{\partial f}{\partial t} + \mu_t \frac{\partial f}{\partial x} + \frac{1}{2}\sigma_t^2 \frac{\partial^2 f}{\partial x^2} \right) dt + \sigma_t \frac{\partial f}{\partial x} dW_t$$

Itô Isometry

For predictable H with $\mathbb{E}\left[\int_0^T H_s^2 ds\right] < \infty$:

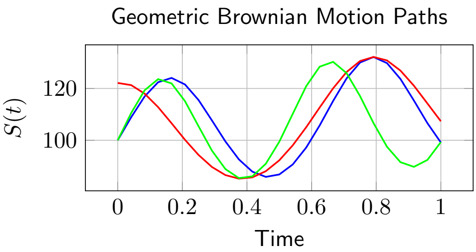
$$\mathbb{E}\left[\left(\int_0^T H_s dW_s\right)^2\right] = \mathbb{E}\left[\int_0^T H_s^2 ds\right]$$

Geometric Brownian Motion

Definition (Geometric Brownian Motion)

The process $S_t = S_0 e^{(\mu - \frac{1}{2}\sigma^2)t + \sigma W_t}$ satisfies:

$$dS_t = \mu S_t dt + \sigma S_t dW_t$$



Martingale Theory

Definition (Martingale)

A process $M = \{M_t\}_{t\geq 0}$ is a **martingale** if:

1. M is adapted and integrable
2. $\mathbb{E}[M_t|\mathcal{F}_s] = M_s$ for all $0 \leq s \leq t$

Theorem (Optional Stopping Theorem)

If M is a uniformly integrable martingale and τ is a bounded stopping time, then:

$$\mathbb{E}[M_\tau] = \mathbb{E}[M_0]$$

Theorem (Martingale Representation Theorem)

Every square-integrable $\mathcal{F}_T$-measurable random variable H can be represented as:

$$H = \mathbb{E}[H] + \int_0^T \phi_s dW_s$$

Table of Contents

The Markowitz Problem

Portfolio Optimization Problem

$$\min_{\mathbf{w}} \frac{1}{2}\mathbf{w}^T\boldsymbol{\Sigma}\mathbf{w}$$

subject to:

$$\boldsymbol{\mu}^T\mathbf{w} = \mu_p \tag{1}$$

$$\mathbf{1}^T\mathbf{w} = 1 \tag{2}$$

where:

- $\mathbf{w} \in \mathbb{R}^n$ is the portfolio weight vector
- $\boldsymbol{\Sigma} \in \mathbb{R}^{n\times n}$ is the covariance matrix
- $\boldsymbol{\mu} \in \mathbb{R}^n$ is the expected return vector

Lagrangian Solution

The Lagrangian:

$$\mathcal{L}(\mathbf{w}, \lambda, \gamma) = \frac{1}{2}\mathbf{w}^T \boldsymbol{\Sigma} \mathbf{w} - \lambda(\boldsymbol{\mu}^T \mathbf{w} - \mu_p) - \gamma(\mathbf{1}^T \mathbf{w} - 1)$$

First-order conditions yield:

$$\mathbf{w} = \lambda \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu} + \gamma \boldsymbol{\Sigma}^{-1} \mathbf{1}$$

Let:

- $A = \mathbf{1}^T \boldsymbol{\Sigma}^{-1} \mathbf{1}$
- $B = \mathbf{1}^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}$
- $C = \boldsymbol{\mu}^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}$
- $D = AC - B^2$

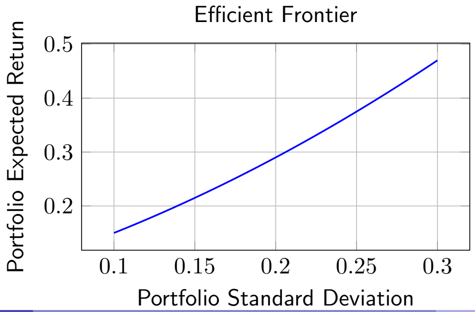
Efficient Frontier

Theorem (Efficient Frontier Equation)

The minimum variance for expected return μ_p is:

$$\sigma_p^2(\mu_p) = \frac{1}{D}(A\mu_p^2 - 2B\mu_p + C)$$

This is a hyperbola in (σ_p^2, μ_p) space.



Two-Fund Separation Theorem

Theorem (Two-Fund Separation)

Every efficient portfolio can be expressed as a linear combination of any two efficient portfolios.

Proof.

Let $\mathbf{w}_1$ and $\mathbf{w}_2$ be two efficient portfolios with expected returns μ_1 and μ_2. Any portfolio with expected return μ_p can be written as:

$$\mathbf{w}_p = \alpha \mathbf{w}_1 + (1 - \alpha)\mathbf{w}_2$$

where $\alpha = \frac{\mu_p - \mu_2}{\mu_1 - \mu_2}$.

Since both $\mathbf{w}_1$ and $\mathbf{w}_2$ lie on the efficient frontier, their linear combination also lies on the efficient frontier. $\square$

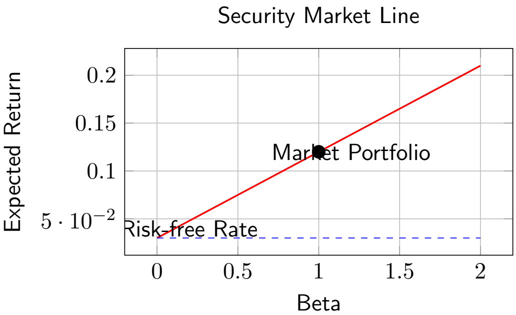
CAPM Derivation

Theorem (Capital Asset Pricing Model)

In equilibrium, the expected return of any asset i satisfies:

$$\mathbb{E}[R_i] = r_f + \beta_i(\mathbb{E}[R_M] - r_f)$$

where $\beta_i = \frac{\text{Cov}(R_i, R_M)}{\text{Var}(R_M)}$.



Factor Models

Definition (K-Factor Model)

$$R_i = \alpha_i + \sum_{k=1}^{K} \beta_{ik} F_k + \epsilon_i$$

where:

- R_i: return on asset i
- F_k: factor k return
- β_{ik}: loading of asset i on factor k
- ϵ_i: idiosyncratic error

In matrix form: $\mathbf{R} = \boldsymbol{\alpha} + \mathbf{B}\mathbf{F} + \boldsymbol{\epsilon}$

Theorem (Factor Model Covariance)

$$\boldsymbol{\Sigma} = \mathbf{B}\boldsymbol{\Sigma}_f\mathbf{B}^T + \boldsymbol{\Psi}$$

where $\boldsymbol{\Psi}$ is diagonal (idiosyncratic risk)

Table of Contents

Von Neumann-Morgenstern Axioms

Axioms of Expected Utility

1. **Completeness**: For any lotteries L_1, L_2: either $L_1 \succeq L_2$, $L_2 \succeq L_1$, or both
2. **Transitivity**: If $L_1 \succeq L_2$ and $L_2 \succeq L_3$, then $L_1 \succeq L_3$
3. **Continuity**: If $L_1 \succ L_2 \succ L_3$, then $\exists \alpha \in (0,1)$ such that $L_2 \sim \alpha L_1 + (1-\alpha)L_3$
4. **Independence**: If $L_1 \succeq L_2$, then $\alpha L_1 + (1-\alpha)L_3 \succeq \alpha L_2 + (1-\alpha)L_3$

Theorem (VNM Representation)

The axioms imply existence of a utility function u such that:

$$L_1 \succeq L_2 \iff \mathbb{E}[u(L_1)] \geq \mathbb{E}[u(L_2)]$$

Risk Aversion Measures

Definition (Absolute Risk Aversion)

$$A(x) = -\frac{u''(x)}{u'(x)}$$

Definition (Relative Risk Aversion)

$$R(x) = -\frac{xu''(x)}{u'(x)} = xA(x)$$

Common Utility Functions

- **CARA**: $u(x) = -\frac{1}{a}e^{-ax}$, $A(x) = a$
- **CRRA**: $u(x) = \frac{x^{1-\gamma}}{1-\gamma}$, $R(x) = \gamma$
- **HARA**: $u(x) = \frac{1-\gamma}{\gamma}\left(\frac{ax}{1-\gamma} + b\right)^{\gamma}$

Optimal Portfolio Choice: CARA Utility

Problem Setup

Maximize: $\mathbb{E}[u(W_T)]$ where $u(w) = -\frac{1}{a}e^{-aw}$
Terminal wealth: $W_T = W_0(1 + r_f) + \mathbf{w}^T(\mathbf{R} - r_f\mathbf{1})W_0$

Theorem (CARA Optimal Portfolio)

Under multivariate normal returns $\mathbf{R} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$*:*

$$\mathbf{w}^* = \frac{1}{aW_0}\boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu} - r_f\mathbf{1})$$

Note: Portfolio weights are independent of wealth level (no wealth effect).

Stochastic Dominance

Definition (First-Order Stochastic Dominance)

F dominates G (written $F \succeq_{FSD} G$) if:

$$F(x) \leq G(x) \quad \forall x$$

with strict inequality for some x.

Definition (Second-Order Stochastic Dominance)

F dominates G (written $F \succeq_{SSD} G$) if:

$$\int_{-\infty}^{x} [G(t) - F(t)]dt \geq 0 \quad \forall x$$

Theorem (SSD and Risk Aversion)

$F \succeq_{SSD} G$ if and only if $\mathbb{E}_F[u] \geq \mathbb{E}_G[u]$ for all concave u.

Table of Contents

Merton's Continuous-Time Model

Setup

- Asset prices: $dS_i(t) = S_i(t)[\mu_i dt + \sigma_i dW_i(t)]$
- Risk-free asset: $dB(t) = rB(t)dt$
- Wealth dynamics:

$$dW(t) = [rW(t) + \boldsymbol{\pi}^T(\boldsymbol{\mu} - r\mathbf{1})W(t) - c(t)]dt + W(t)\boldsymbol{\pi}^T\boldsymbol{\sigma}d\mathbf{W}(t)$$

where $\boldsymbol{\pi}$ is the vector of portfolio weights and $c(t)$ is consumption.

Hamilton-Jacobi-Bellman Equation

Definition (Value Function)

$$J(W,t) = \max_{\{\boldsymbol{\pi},c\}} \mathbb{E}\left[\int_t^T u(c(s),s)ds + B(W(T),T) \middle| W(t) = W\right]$$

Theorem (HJB Equation)

The value function satisfies:

$$0 = \max_{\boldsymbol{\pi},c} \Big\{ u(c,t) + J_t + J_W[rW + \boldsymbol{\pi}^T(\boldsymbol{\mu} - r\mathbf{1})W - c] \tag{3}$$

$$+ \frac{1}{2} J_{WW} W^2 \boldsymbol{\pi}^T \boldsymbol{\Sigma} \boldsymbol{\pi} \Big\} \tag{4}$$

with boundary condition $J(W,T) = B(W,T)$.

First-Order Conditions

From the HJB equation:

$$\frac{\partial u}{\partial c} = J_W \quad \text{(consumption FOC)} \tag{5}$$

$$J_W(\boldsymbol{\mu} - r\mathbf{1})W + J_{WW}W^2\boldsymbol{\Sigma}\boldsymbol{\pi} = \mathbf{0} \quad \text{(portfolio FOC)} \tag{6}$$

The optimal portfolio is:

$$\boldsymbol{\pi}^* = -\frac{J_W}{J_{WW}W}\boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu} - r\mathbf{1})$$

CRRA Case: Explicit Solution

For CRRA utility $u(c) = \frac{c^{1-\gamma}}{1-\gamma}$, conjecture:

$$J(W,t) = \frac{W^{1-\gamma}}{1-\gamma} f(t)$$

Theorem (Merton's Solution)

The optimal consumption and portfolio rules are:

$$c^*(t) = \delta^{1/\gamma} W(t) f(t)^{-1/\gamma} \tag{7}$$

$$\boldsymbol{\pi}^*(t) = \frac{1}{\gamma} \boldsymbol{\Sigma}^{-1} (\boldsymbol{\mu} - r\mathbf{1}) \tag{8}$$

where $f(t)$ satisfies an ODE.

Note: The portfolio rule is independent of wealth and time (myopic property).

Finite vs. Infinite Horizon

Infinite Horizon ($T \to \infty$)

- Steady-state consumption: $c^* = \delta W$
- Constant portfolio weights
- No hedging demands

Finite Horizon

- Time-varying consumption rate
- Potential hedging demands
- Horizon effects in portfolio choice

Theorem (Turnpike Property)

In long-horizon problems, optimal paths converge to a steady-state "turnpike" regardless of initial conditions.

Table of Contents

Fundamental Theorems of Asset Pricing

Definition (Arbitrage)

An arbitrage opportunity is a self-financing strategy with:

- Initial cost: $V_0 = 0$
- Final payoff: $V_T \geq 0$ a.s., $\mathbb{P}(V_T > 0) > 0$

Theorem (First FTAP)

No arbitrage exists if and only if there exists an equivalent martingale measure $\mathbb{Q} \sim \mathbb{P}$ such that discounted price processes are $\mathbb{Q}$-martingales.

Theorem (Second FTAP)

The market is complete if and only if the equivalent martingale measure is unique.

Transaction Costs: No-Trade Region

Bid-Ask Spread Model

- Bid price: $S_t^b = S_t(1 - \lambda)$
- Ask price: $S_t^a = S_t(1 + \lambda)$
- Transaction cost parameter: $\lambda > 0$

Theorem (No-Trade Region)

With proportional transaction costs, the optimal strategy exhibits a no-trade region around the frictionless optimum.

The boundaries satisfy:

$$\frac{\partial V}{\partial \pi_i} = \pm \lambda W \frac{\partial V}{\partial W}$$

No-Trade Region Illustration

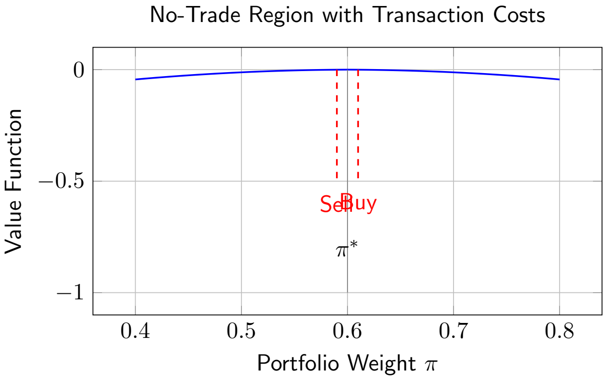


Table of Contents

Convex Optimization Foundations

Definition (Convex Function)

$f : \mathbb{R}^n \to \mathbb{R}$ is convex if: $f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)f(y)$ for all $x, y \in \mathbb{R}^n$ and $\lambda \in [0, 1]$.

Theorem (Jensen's Inequality)

If f is convex and X is a random variable: $f(\mathbb{E}[X]) \leq \mathbb{E}[f(X)]$

KKT Conditions

Theorem (Karush-Kuhn-Tucker Conditions)

If x^* is optimal and constraint qualification holds, then $\exists \lambda^*, \nu^*$ such that:

$$\nabla f_0(x^*) + \sum_{i=1}^{m} \lambda_i^* \nabla f_i(x^*) + \sum_{j=1}^{p} \nu_j^* \nabla h_j(x^*) = 0 \tag{9}$$

$$f_i(x^*) \leq 0, \quad i = 1, \ldots, m \tag{10}$$

$$h_j(x^*) = 0, \quad j = 1, \ldots, p \tag{11}$$

$$\lambda_i^* \geq 0, \quad i = 1, \ldots, m \tag{12}$$

$$\lambda_i^* f_i(x^*) = 0, \quad i = 1, \ldots, m \tag{13}$$

Robust Portfolio Optimization

Definition (Uncertainty Set)

Let $\mathcal{U}$ be a convex, compact set containing possible values of $(\boldsymbol{\mu}, \boldsymbol{\Sigma})$.

Definition (Robust Optimization Problem)

$\min_{\mathbf{w}} \max_{(\boldsymbol{\mu},\boldsymbol{\Sigma})\in\mathcal{U}} \mathbf{w}^T\boldsymbol{\Sigma}\mathbf{w}$ subject to: $\min_{(\boldsymbol{\mu},\boldsymbol{\Sigma})\in\mathcal{U}} \boldsymbol{\mu}^T\mathbf{w} \geq \mu_{\min}$

Ellipsoidal Uncertainty Set

$$\mathcal{U} = \left\{(\boldsymbol{\mu}, \boldsymbol{\Sigma}) : \|\boldsymbol{\mu} - \hat{\boldsymbol{\mu}}\|_{\mathbf{Q}^{-1}} \leq \kappa, \boldsymbol{\Sigma} = \hat{\boldsymbol{\Sigma}}\right\}$$

Table of Contents

Coherent Risk Measures

Definition (Coherent Risk Measure)

ρ is coherent if it satisfies:

1. **Monotonicity**: $X \leq Y \Rightarrow \rho(X) \geq \rho(Y)$
2. **Translation Invariance**: $\rho(X + c) = \rho(X) - c$
3. **Positive Homogeneity**: $\rho(\lambda X) = \lambda \rho(X)$ for $\lambda \geq 0$
4. **Subadditivity**: $\rho(X + Y) \leq \rho(X) + \rho(Y)$

Definition (Value-at-Risk)

$\text{VaR}_\alpha(X) = -\inf\{x : P(X \leq x) \geq \alpha\} = -q_\alpha(X)$

Problem: VaR fails subadditivity, making it non-coherent.

Conditional Value-at-Risk

Definition (CVaR/Expected Shortfall)

$\text{CVaR}_\alpha(X) = \mathbb{E}[X|X \leq q_\alpha(X)]$
Equivalently: $\text{CVaR}_\alpha(X) = \frac{1}{\alpha}\int_0^\alpha q_u(X)du$

Theorem (CVaR Coherence)

CVaR is a coherent risk measure.

Theorem (CVaR Dual Representation)

$CVaR_\alpha(X) = \min_{t\in\mathbb{R}} \left\{t + \frac{1}{\alpha}\mathbb{E}[(X-t)^+]\right\}$

This representation makes CVaR optimization tractable via linear programming.

VaR vs CVaR Illustration

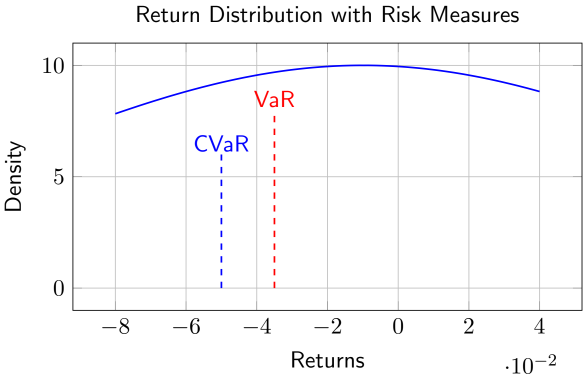


Table of Contents

Curse of Dimensionality

The Problem

With n assets and T observations:

- Sample covariance matrix: $\hat{\mathbf{\Sigma}} = \frac{1}{T-1}\sum_{t=1}^{T}(\mathbf{r}_t - \bar{\mathbf{r}})(\mathbf{r}_t - \bar{\mathbf{r}})^T$
- When $n/T \to c > 0$: $\hat{\mathbf{\Sigma}}$ is ill-conditioned
- Traditional Markowitz fails catastrophically

Theorem (Marchenko-Pastur Law)

As $n, T \to \infty$ with $n/T \to c \in (0,1)$, the eigenvalues of $\hat{\mathbf{\Sigma}}$ converge to:
$\rho_{MP}(\lambda) = \frac{1}{2\pi c\sigma^2}\frac{\sqrt{(\lambda_+ - \lambda)(\lambda - \lambda_-)}}{\lambda}$ *where* $\lambda_\pm = \sigma^2(1 \pm \sqrt{c})^2$*.*

Shrinkage Estimators

Definition (Ledoit-Wolf Shrinkage)

$\hat{\mathbf{\Sigma}}_{LW} = \alpha \mathbf{F} + (1-\alpha)\hat{\mathbf{\Sigma}}$ where $\mathbf{F}$ is a structured target and α^* minimizes expected loss.

Theorem (Optimal Shrinkage Intensity)

The optimal shrinkage intensity is: $\alpha^* = \frac{\sum_{i,j} \mathit{Var}(\hat{\Sigma}_{ij})}{\sum_{i,j}(\hat{\Sigma}_{ij}-F_{ij})^2}$

Theorem (Stein's Paradox)

For $p \geq 3$, the James-Stein estimator dominates the sample mean under squared error loss.

Factor Models and PCA

Definition (Strict Factor Model)

$\mathbf{r}_t = \boldsymbol{\alpha} + \mathbf{B}\mathbf{f}_t + \boldsymbol{\varepsilon}_t$

Theorem (Factor Model Covariance)

$\mathbf{\Sigma} = \mathbf{B}\mathbf{\Sigma}_f\mathbf{B}^T + \mathbf{\Psi}$ *where* $\mathbf{\Psi}$ *is diagonal.*

Theorem (Principal Components Consistency)

Under regularity conditions, the first k *principal components consistently estimate the factor space as* $n, T \to \infty$.

Regularization Methods

Definition (LASSO Objective)

$\min_{\boldsymbol{\beta}} \frac{1}{2}\|\mathbf{y} - \mathbf{X}\boldsymbol{\beta}\|_2^2 + \lambda\|\boldsymbol{\beta}\|_1$

ℓ_1-Regularized Portfolio

$\min_{\mathbf{w}} \mathbf{w}^T\hat{\boldsymbol{\Sigma}}\mathbf{w} + \lambda\|\mathbf{w}\|_1$ subject to: $\boldsymbol{\mu}^T\mathbf{w} = \mu_p$, $\mathbf{1}^T\mathbf{w} = 1$

This promotes sparsity in portfolio weights.

Theorem (Large Portfolio Asymptotics)

As $n \to \infty$ with bounded factor structure: Var$(\mathbf{w}^T\mathbf{r}) \to \mathbf{w}^T\mathbf{B}\boldsymbol{\Sigma}_f\mathbf{B}^T\mathbf{w}$

Table of Contents

Prospect Theory

Definition (Value Function)

$v(x) = \begin{cases} x^\alpha & \text{if } x \geq 0 \\ -\lambda(-x)^\beta & \text{if } x < 0 \end{cases}$ where $\alpha, \beta \in (0, 1)$ (diminishing sensitivity) and $\lambda > 1$ (loss aversion).

Definition (Probability Weighting)

$w(p) = \frac{p^\gamma}{(p^\gamma + (1-p)^\gamma)^{1/\gamma}}$

Theorem (Prospect Theory Preferences)

$V = \sum_i w(p_i)v(x_i - r)$ *where r is the reference point.*

Ambiguity Aversion Models

Definition (Max-Min Expected Utility)

$V = \min_{P \in \mathcal{P}} \mathbb{E}_P[u(W)]$ where $\mathcal{P}$ is a set of probability measures.

Definition (Smooth Ambiguity Model)

$V = \phi^{-1}\left(\mathbb{E}_\mu[\phi(\mathbb{E}_P[u(W)])]\right)$ where ϕ captures ambiguity attitude and μ is second-order belief.

Definition (Habit Formation)

$u(C_t, X_t) = \frac{(C_t - X_t)^{1-\gamma}}{1-\gamma}$ where X_t is the habit level.

Table of Contents

Monte Carlo Methods

Definition (Monte Carlo Estimator)

For $\theta = \mathbb{E}[g(X)]$: $\hat{\theta}_n = \frac{1}{n}\sum_{i=1}^{n} g(X_i)$ where X_i are iid samples.

Theorem (Central Limit Theorem)

$\sqrt{n}(\hat{\theta}_n - \theta) \xrightarrow{d} \mathcal{N}(0, \sigma^2)$ where $\sigma^2 = Var(g(X))$.

Variance Reduction Techniques

- **Antithetic Variables**: Use pairs $(U, 1-U)$
- **Control Variates**: $\hat{\theta}_{CV} = \hat{\theta} - c(\hat{\beta} - \beta)$
- **Importance Sampling**: Change of measure

Finite Difference Methods

HJB Equation Discretization

$$\frac{\partial J}{\partial t} + \max_\pi \{\mathcal{L}^\pi J\} = 0$$

Definition (Upwind Scheme)

$$\frac{J_i^{n+1} - J_i^n}{\Delta t} + \max_\pi \left\{ \mu(\pi) \frac{J_{i+1}^n - J_i^n}{\Delta W} + \frac{\sigma^2(\pi)}{2} \frac{J_{i+1}^n - 2J_i^n + J_{i-1}^n}{(\Delta W)^2} \right\} = 0$$

Algorithm 1 Value Iteration

Initialize V^0

repeat

$V^{k+1}(x) = \max_\pi \{u(x,\pi) + \gamma \mathbb{E}[V^k(f(x,\pi,\xi))]\}$

until $\|V^{k+1} - V^k\| < \epsilon$

Table of Contents

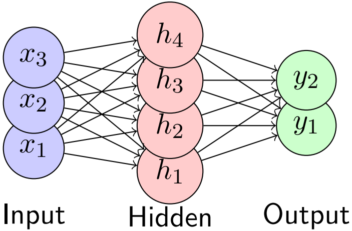
Deep Learning for Asset Pricing

Definition (Neural Network)

$f(\mathbf{x}) = W^{(L)}\sigma(W^{(L-1)}\sigma(\cdots\sigma(W^{(1)}\mathbf{x} + b^{(1)})\cdots) + b^{(L-1)}) + b^{(L)}$ where σ is an activation function (e.g., ReLU, tanh).

Theorem (Universal Approximation)

Neural networks with one hidden layer can approximate any continuous function on compact sets arbitrarily well.



Reinforcement Learning

Definition (Markov Decision Process)

A tuple $(\mathcal{S}, \mathcal{A}, P, R, \gamma)$ where:

- $\mathcal{S}$: state space
- $\mathcal{A}$: action space
- $P(s'|s,a)$: transition probabilities
- $R(s,a)$: reward function
- γ: discount factor

Definition (Q-Function)

$$Q^\pi(s,a) = \mathbb{E}^\pi \left[\sum_{t=0}^\infty \gamma^t R(S_t, A_t) \middle| S_0 = s, A_0 = a \right]$$

Theorem (Bellman Equation)

$$Q^*(s,a) = R(s,a) + \gamma \sum_{s'} P(s'|s,a) \max_{a'} Q^*(s',a')$$

Table of Contents

ESG Integration

Definition (ESG-Constrained Optimization)

$\min_{\mathbf{w}} \mathbf{w}^T \boldsymbol{\Sigma} \mathbf{w}$ subject to:

$$\boldsymbol{\mu}^T \mathbf{w} = \mu_p \tag{14}$$

$$\mathbf{s}^T \mathbf{w} \geq s_{\min} \quad \text{(ESG constraint)} \tag{15}$$

$$\mathbf{1}^T \mathbf{w} = 1 \tag{16}$$

where $\mathbf{s}$ is the vector of ESG scores.

Theorem (ESG Premium)

The ESG premium is: $\pi_{ESG} = \frac{\lambda_{ESG}}{\mathbf{1}^T \boldsymbol{\Sigma}^{-1} \mathbf{1}}$ *where* λ_{ESG} *is the Lagrange multiplier on the ESG constraint.*

Cryptocurrency Portfolio Theory

Definition (Crypto Return Model)

$r_{i,t} = \alpha_i + \sum_{j=1}^{k} \beta_{ij} F_{j,t} + \gamma_i M_t + \epsilon_{i,t}$ where:

- $F_{j,t}$: traditional risk factors
- M_t: crypto market factor
- γ_i: crypto beta

Unique Considerations

- Extreme volatility and fat tails
- 24/7 trading (no overnight risk)
- Limited fundamental anchors
- Regulatory uncertainty
- Network effects and adoption dynamics

Future Research Directions

Emerging Areas

- **Quantum Computing**: Portfolio optimization with quantum algorithms
- **Climate Risk**: Integration of physical and transition risks
- **Alternative Data**: Satellite imagery, social media sentiment
- **Federated Learning**: Privacy-preserving collaborative modeling
- **Explainable AI**: Interpretable machine learning for finance

Mathematical Challenges

- Non-convex optimization in deep learning
- High-dimensional statistical inference
- Robustness to distributional shifts
- Real-time adaptive algorithms
- Integration of discrete and continuous methods

Table of Contents

Key Mathematical Themes

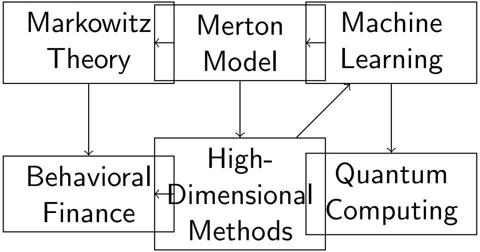
Foundational Mathematics

- **Measure Theory**: Rigorous probability foundations
- **Functional Analysis**: Optimization in infinite dimensions
- **Stochastic Calculus**: Continuous-time modeling
- **Convex Analysis**: Optimization theory and duality

Applied Methods

- **Numerical Analysis**: Computational implementation
- **Statistics**: Parameter estimation and testing
- **Machine Learning**: Pattern recognition and prediction
- **Game Theory**: Multi-agent interactions

Integration of Classical and Modern Methods



Practical Implementation Guidelines

Model Selection Hierarchy

1. Start with mean-variance framework
2. Add realistic constraints and transaction costs
3. Incorporate robust optimization techniques
4. Consider behavioral factors for retail investors
5. Apply machine learning for alpha generation
6. Use advanced risk measures for risk management

Key Takeaways

- Rigorous foundations enable robust applications
- Interdisciplinary approach combining finance, economics, psychology
- Leverage modern computing for previously intractable problems
- Always validate theoretical insights with real data

Final Remarks

Portfolio Theory Evolution

From Markowitz's mean-variance foundation to modern machine learning applications, portfolio theory continues to evolve with:

- More sophisticated mathematical tools
- Increased computational power
- Richer datasets and alternative data sources
- Growing awareness of behavioral and systemic factors

Thank you for your attention!

Questions and Discussion