Modern Portfolio Theory and Quantitative Risk Models

From Markowitz to Advanced Quantitative Methods

Leonardo Tiditada Pedersen

October 2025

Table of Contents

Defining Risk & Return Mathematically

The building blocks of quantitative finance

Expected Return: The Mean (μ)

The expected return of an asset is the probability-weighted average of its potential returns. For a history of T returns R_t:

$$\mu = \mathbb{E}[R] = \frac{1}{T}\sum_{t=1}^{T} R_t$$

Risk: The Standard Deviation (σ)

Risk is the volatility of returns, measured as the square root of the variance.

$$\sigma^2 = \mathrm{Var}(R) = \frac{1}{T-1}\sum_{t=1}^{T}(R_t - \mu)^2 \quad \Longrightarrow \quad \sigma = \sqrt{\sigma^2}$$

Defining Risk & Return Mathematically 2

Systematic vs. Unsystematic Risk

Total Risk (σ^2) can be decomposed:

$$\sigma^2_{\text{Total}} = \underbrace{\beta^2 \sigma^2_{\text{Market}}}_{\text{Systematic}} + \underbrace{\sigma^2_{\epsilon}}_{\text{Unsystematic}}$$

Diversification aims to make the unsystematic risk term, σ^2_{ϵ}, approach zero.

The Mathematics of Diversification (N Assets)

Generalizing portfolio risk with matrix algebra

N-Asset Portfolio Expected Return

Let $\boldsymbol{w}$ be the vector of portfolio weights and $\boldsymbol{\mu}$ be the vector of expected returns.

$$\mathbb{E}[R_p] = \boldsymbol{w}^T \boldsymbol{\mu} = \sum_{i=1}^{N} w_i \mu_i$$

The Mathematics of Diversification (N Assets)

Generalizing portfolio risk with matrix algebra

N-Asset Portfolio Expected Return

Let $\boldsymbol{w}$ be the vector of portfolio weights and $\boldsymbol{\mu}$ be the vector of expected returns.

$$\mathbb{E}[R_p] = \boldsymbol{w}^T \boldsymbol{\mu} = \sum_{i=1}^{N} w_i \mu_i$$

N-Asset Portfolio Variance

Let $\boldsymbol{\Sigma}$ be the $N \times N$ covariance matrix of asset returns.

$$\sigma_p^2 = \boldsymbol{w}^T \boldsymbol{\Sigma} \boldsymbol{w} = \sum_{i=1}^{N} \sum_{j=1}^{N} w_i w_j \mathrm{Cov}(R_i, R_j)$$

Where $\mathrm{Cov}(R_i, R_j) = \rho_{ij} \sigma_i \sigma_j$. The diagonal elements are the variances σ_i^2.

Deriving the Efficient Frontier

The Markowitz Optimization Problem

The Formal Problem

The Efficient Frontier is solved by finding the portfolio weight vector $\boldsymbol{w}$ that minimizes portfolio risk for a target expected return μ_p^*. This is a constrained optimization problem.

The Lagrangian Formulation

Minimize the Lagrangian $\mathcal{L}$ with respect to the weights $\boldsymbol{w}$:

$$\min_{\boldsymbol{w}} \mathcal{L} = \frac{1}{2}\boldsymbol{w}^T \boldsymbol{\Sigma} \boldsymbol{w} - \lambda_1(\boldsymbol{w}^T \boldsymbol{\mu} - \mu_p^*) - \lambda_2(\boldsymbol{w}^T \mathbf{1} - 1)$$

- **Objective:** Minimize portfolio variance: $\frac{1}{2}\boldsymbol{w}^T \boldsymbol{\Sigma} \boldsymbol{w}$.
- **Constraint 1:** The portfolio's return must equal the target return: $\boldsymbol{w}^T \boldsymbol{\mu} = \mu_p^*$.
- **Constraint 2:** The weights must sum to 1: $\boldsymbol{w}^T \mathbf{1} = 1$.

Finding the Optimal (Tangency) Portfolio

Maximizing the Sharpe Ratio

The Optimization Problem

The tangency portfolio is the specific portfolio of risky assets that has the highest possible Sharpe Ratio. We maximize the Sharpe Ratio with respect to the weight vector $\boldsymbol{w}$.

$$\max_{\boldsymbol{w}} \quad S_p = \frac{\boldsymbol{w}^T \boldsymbol{\mu} - R_f}{(\boldsymbol{w}^T \boldsymbol{\Sigma} \boldsymbol{w})^{1/2}}$$

Subject to the constraint that $\boldsymbol{w}^T \boldsymbol{1} = 1$.

Finding the Optimal (Tangency) Portfolio

Maximizing the Sharpe Ratio

The Optimization Problem

The tangency portfolio is the specific portfolio of risky assets that has the highest possible Sharpe Ratio. We maximize the Sharpe Ratio with respect to the weight vector $\boldsymbol{w}$.

$$\max_{\boldsymbol{w}} \quad S_p = \frac{\boldsymbol{w}^T \boldsymbol{\mu} - R_f}{(\boldsymbol{w}^T \boldsymbol{\Sigma} \boldsymbol{w})^{1/2}}$$

Subject to the constraint that $\boldsymbol{w}^T \mathbf{1} = 1$.

Finding the Optimal (Tangency) Portfolio 2

The Solution

The analytical solution for the weights of the tangency portfolio, $\boldsymbol{w}_{\tan}$, is given by:

$$\boldsymbol{w}_{\tan} = \frac{\boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu} - R_f \mathbf{1})}{\mathbf{1}^T \boldsymbol{\Sigma}^{-1}(\boldsymbol{\mu} - R_f \mathbf{1})}$$

- $\boldsymbol{\Sigma}^{-1}$ is the inverse of the covariance matrix.
- $(\boldsymbol{\mu} - R_f \mathbf{1})$ is the vector of excess returns.

Value at Risk (VaR) - The Mathematics

Defining the quantile of the P&L distribution

Formal Definition

For a given confidence level $\alpha \in (0, 1)$, the VaR of a portfolio at level α is the smallest number l such that the probability of a loss L exceeding l is no larger than $(1 - \alpha)$.

$$\text{VaR}_\alpha = \inf\{l \in \mathbb{R} : P(L > l) \leq 1 - \alpha\}$$

Value at Risk (VaR) - The Mathematics

Defining the quantile of the P&L distribution

Formal Definition

For a given confidence level $\alpha \in (0,1)$, the VaR of a portfolio at level α is the smallest number l such that the probability of a loss L exceeding l is no larger than $(1-\alpha)$.

$$\mathrm{VaR}_\alpha = \inf\{l \in \mathbb{R} : P(L > l) \leq 1 - \alpha\}$$

The Parametric (Variance-Covariance) VaR Formula

If we assume portfolio returns are normally distributed, the VaR for a given time horizon T is:

$$\mathrm{VaR}_\alpha(T) = -(\mu_p T + Z_\alpha \sigma_p \sqrt{T}) V_0$$

- V_0 is the initial portfolio value.
- μ_p and σ_p are the portfolio's expected return and volatility.

Conditional VaR (CVaR) - The Mathematics

Calculating the Expected Shortfall

The Problem with VaR

VaR is a quantile. It says nothing about the severity of losses beyond that point. It is not a "coherent" risk measure as it can fail subadditivity.

Conditional VaR (CVaR) - The Mathematics

Calculating the Expected Shortfall

The Problem with VaR

VaR is a quantile. It says nothing about the severity of losses beyond that point. It is not a "coherent" risk measure as it can fail subadditivity.

Formal Definition of CVaR

Conditional VaR (or Expected Shortfall) is the expected value of the loss, L, given that the loss is greater than or equal to the VaR at confidence level α.

$$\text{CVaR}_\alpha = \mathbb{E}[L \mid L \geq \text{VaR}_\alpha]$$

Conditional VaR (CVaR) - The Mathematics 2

Calculating the Expected Shortfall

For a Normal Distribution

If returns are normally distributed, CVaR has a closed-form solution:

$$\text{CVaR}_\alpha = \mu_p - \sigma_p \frac{\phi(Z_\alpha)}{1-\alpha}$$

- $\phi(\cdot)$ is the probability density function (PDF) of the standard normal distribution.
- Z_α is the z-score of the α-quantile.

CVaR will always be greater than or equal to VaR.

The Black-Litterman Master Formula

A Bayesian framework for expected returns

The Core Idea

The model produces a posterior estimate for expected returns by blending a prior estimate (from market equilibrium) with new information (the investor's views).

The Black-Litterman Master Formula

A Bayesian framework for expected returns

The Core Idea

The model produces a posterior estimate for expected returns by blending a prior estimate (from market equilibrium) with new information (the investor's views).

The Black-Litterman Master Formula 2

A Bayesian framework for expected returns

The Master Equation

The posterior estimate of expected returns, $\mathbb{E}[R]$, is a weighted average:

$$\mathbb{E}[R] = \left[(\tau\boldsymbol{\Sigma})^{-1} + \boldsymbol{P}^T\boldsymbol{\Omega}^{-1}\boldsymbol{P}\right]^{-1}\left[(\tau\boldsymbol{\Sigma})^{-1}\boldsymbol{\Pi} + \boldsymbol{P}^T\boldsymbol{\Omega}^{-1}\boldsymbol{Q}\right]$$

- $\boldsymbol{\Pi}$: Vector of implied equilibrium returns (the prior).
- $\tau\boldsymbol{\Sigma}$: Covariance matrix of the prior returns. τ is a scaling constant.
- $\boldsymbol{P}$: The "pick" matrix, identifying the assets in the investor's views.
- $\boldsymbol{Q}$: The vector of expected returns for the views.
- $\boldsymbol{\Omega}$: The covariance matrix of the error terms in the views, representing the investor's confidence.

The Black-Litterman Master Formula 3

A Bayesian framework for expected returns

Result

This blended $\mathbb{E}[R]$ vector is then used in a standard mean-variance optimizer to produce stable, well-diversified, and intuitive portfolio weights.

Beyond Beta: Multi-Factor Models

Explaining the cross-section of returns

Jensen's Alpha (α)

From the CAPM regression, Alpha is the intercept. It measures the average return of a portfolio above or below that predicted by the model. A positive, statistically significant alpha indicates superior risk-adjusted performance.

$$R_{p,t} - R_{f,t} = \alpha_p + \beta_p(R_{m,t} - R_{f,t}) + \epsilon_{p,t}$$

Beyond Beta: Multi-Factor Models

Explaining the cross-section of returns

Jensen's Alpha (α)

From the CAPM regression, Alpha is the intercept. It measures the average return of a portfolio above or below that predicted by the model. A positive, statistically significant alpha indicates superior risk-adjusted performance.

$$R_{p,t} - R_{f,t} = \alpha_p + \beta_p(R_{m,t} - R_{f,t}) + \epsilon_{p,t}$$

Fama-French 3-Factor Model

Adds size (SMB) and value (HML) factors to the market factor.

$$R_{p,t} - R_{f,t} = \alpha_p + \beta_{mkt}(R_{m,t} - R_{f,t}) + \beta_{smb}\mathrm{SMB}_t + \beta_{hml}\mathrm{HML}_t + \epsilon_{p,t}$$

Carhart 4-Factor Model

Adds a momentum (MOM) factor to the Fama-French model.

$$R_{p,t} - R_{f,t} = \alpha_p + \beta_{mkt}(R_{m,t} - R_{f,t}) + \beta_{smb}\mathrm{SMB}_t + \beta_{hml}\mathrm{HML}_t + \beta_{mom}\mathrm{MOM}_t + \epsilon_{p,}$$

Time Series Models I: Forecasting Volatility (GARCH)

Modeling time-varying risk

The Problem: Volatility is Not Constant

Financial returns exhibit **volatility clustering**: high-volatility periods are followed by high-volatility periods, and low by low. Standard models assuming constant variance (σ^2) will be incorrect.

Time Series Models I: Forecasting Volatility (GARCH)

Modeling time-varying risk

The Problem: Volatility is Not Constant

Financial returns exhibit **volatility clustering**: high-volatility periods are followed by high-volatility periods, and low by low. Standard models assuming constant variance (σ^2) will be incorrect.

The GARCH(1,1) Model

The Generalized Autoregressive Conditional Heteroskedasticity model forecasts future variance (σ_t^2) based on past variance and past squared returns.

$$\sigma_t^2 = \omega + \alpha_1 \epsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2$$

- σ_t^2: The conditional variance for today (our forecast).
- ω: A constant term.
- ϵ_{t-1}^2: Yesterday's squared residual (the "ARCH" term - news about volatility from the previous period).
- σ_{t-1}^2: Yesterday's variance (the "GARCH" term - persistence of volatility).

Time Series Models I: Forecasting Volatility (GARCH) 3

Modeling time-varying risk

Application

GARCH models are essential for accurate VaR calculations, option pricing, and dynamic risk management.

Time Series Models II: Forecasting Returns (ARIMA)

Modeling serial correlation in data

The Concept

The Autoregressive Integrated Moving Average (ARIMA) model is a general class of models for forecasting time series data. It assumes that past values have a linear effect on the current value.

The Components: ARIMA(p,d,q)

- **AR(p): Autoregressive** → The value today is a weighted sum of the previous 'p' values.
- **I(d): Integrated** → The data has been differenced 'd' times to make it stationary (i.e., its mean and variance are constant over time).
- **MA(q): Moving Average** → The value today is a weighted sum of the previous 'q' forecast errors (shocks).

The Concept

The Autoregressive Integrated Moving Average (ARIMA) model is a general class of models for forecasting time series data. It assumes that past values have a linear effect on the current value.

The Components: ARIMA(p,d,q)

- **AR(p): Autoregressive** $\rightarrow$ The value today is a weighted sum of the previous ‘p‘ values.
- **I(d): Integrated** $\rightarrow$ The data has been differenced ‘d‘ times to make it stationary (i.e., its mean and variance are constant over time).
- **MA(q): Moving Average** $\rightarrow$ The value today is a weighted sum of the previous ‘q‘ forecast errors (shocks).

Time Series Models II: Forecasting Returns (ARIMA) 2

Modeling serial correlation in data

General Equation (for an ARMA(1,1) process)

Let Y_t be the stationary time series.

$$Y_t = c + \phi_1 Y_{t-1} + \theta_1 \epsilon_{t-1} + \epsilon_t$$

- ϕ_1: The autoregressive parameter.
- θ_1: The moving average parameter.
- ϵ_t: White noise error term.

Data Science: Natural Language Processing (NLP) for Risk

Extracting signals from unstructured text data

The Opportunity

A vast amount of financial information exists as unstructured text (news, reports, social media). NLP allows us to systematically analyze this data for risk and return signals.

Sentiment Analysis

- Automatically classify text as positive, negative, or neutral.
- **Application:** Gauge market sentiment by analyzing millions of news articles or tweets.
- **Risk Signal:** A sudden shift to negative sentiment in the news for a specific stock can be an early warning signal for risk managers.

Topic Modeling & Textual Risk Factors

- Use algorithms like Latent Dirichlet Allocation (LDA) to identify key topics or themes in large document sets.
- **Application:** Analyze a company's 10-K filings over 20 years.
- **Risk Signal:** Identify emerging risks by tracking the frequency of topics like "cybersecurity," "supply chain disruption," or "regulatory scrutiny" over time.

Summary of Key Concepts

Portfolio Theory

- Portfolio variance is calculated as $\sigma_p^2 = \boldsymbol{w}^T \Sigma \boldsymbol{w}$.
- The Efficient Frontier is the solution to a constrained quadratic optimization problem.
- The Tangency Portfolio is found by maximizing the Sharpe Ratio: $\frac{\mathbb{E}[R_p]-R_f}{\sigma_p}$.

Quantitative Models

- **VaR** is the α-quantile of the P&L distribution.
- **CVaR** is the conditional expectation in the tail: $\mathbb{E}[L \mid L \geq \mathrm{VaR}_\alpha]$.
- **Black-Litterman** uses a Bayesian approach to produce a posterior estimate of expected returns for asset allocation.

Contact:leonardo.ped@st.econ.tu.ac.th