# Application of Information Geometry in Stochastic Portfolio Theory

**Dr. Rajni**
Assistant Professor in Mathematics
S.K Government College, Kanwali ,District Rewari


## Abstract

Imagine the stock market as a landscape where each point represents a different combination of stocks. This landscape changes over time as stock prices fluctuate. Market diversity tells us how spread out the investments are in the market. If everyone invests in a few popular stocks, the market is less diverse. A more diverse market means that investments are spread across many different stocks. Market volatility refers to how much stock prices change over time. A volatile market is one where prices move up and down frequently. This paper provides a mathematical framework for creating investment strategies based on these two concepts: One key idea is to create a portfolio that benefits from market volatility while minimizing the impact of changes in market diversity. This is achieved using an *exponentially concave function* to guide investment allocation. A new measure of market volatility, called the *L(α)-divergence,* is introduced. This measure is more flexible than traditional methods and enables a broader range of investment strategies. Using these tools, the paper constructs portfolios that outperform the market, even when transaction costs are considered. This research provides a sophisticated approach to portfolio construction, allowing investors to develop strategies that are robust to market fluctuations and have the potential for higher returns.


**Key Terms:** Market diversity, Market volatility, Multiplicatively Generated Portfolio, L($\alpha$)-Divergence, Generalized Functional Portfolio Generation

## 1. Mathematical Equations

- **Market Weight:** The market weight of stock $i$ is defined as:

$$\mu_i(t) := \frac{X_i(t)}{X_1(t) + \cdots + X_n(t)},$$

  where $X_i(t)$ is the market capitalization of stock $i$ at time $t$.

- **L($\alpha$)-Divergence:** The $L(\alpha)$-divergence of an $\alpha$-exponentially concave function $\phi$ is given by:

$$D^{(\alpha)}[q : p] := \frac{1}{\alpha} \log\left(1 + \alpha \nabla\phi(p) \cdot (q - p)\right) - \left(\phi(q) - \phi(p)\right),$$

  where $p$ and $q$ are market weight vectors.

- ($\alpha, C$)**-Generated Portfolio:** This portfolio is constructed by specifying the number of shares $\eta_i(t)$ of stock $i$ at time $t$ as:

$$\eta_i(t) = \alpha(C + V_\eta(t))D_{e_i-\mu(t)}\phi(\mu(t)) + V_\eta(t),$$ where $V_\eta(t)$ is the

  portfolio value at time $t$.

## 1.1 Results

- The $L(\alpha)$-divergence can construct portfolios with desirable properties, such as relative arbitrage.
- The ($\alpha, C$)-generation method creates a wider range of portfolios compared to traditional methods, allowing for more flexibility in portfolio construction.
- Empirical data demonstrates the potential of ($\alpha, C$)-generated portfolios to outperform other strategies.

## 1.2 What is Shannon Entropy?

Shannon entropy, introduced by Claude Shannon in information theory, measures the uncertainty or "information content" of a probability distribution. It is mathematically defined as:

$$H(P) = -\sum_{i=1}^{n} p_i \log(p_i)$$

where:

- $P = \{p_1, p_2, \ldots, p_n\}$ is a probability distribution over $n$ possible outcomes,
- $p_i$ is the probability of outcome $i$,
- $H(P)$ represents the expected information content or the uncertainty in the distribution.

In portfolio theory, Shannon entropy is adapted to represent the diversification of portfolio allocations. A higher entropy implies greater diversification, while lower entropy suggests concentration in specific assets.

## 1.3 Why is Shannon Entropy Used in Portfolio Theory?

### *1.3.1 Quantifying Diversification*

- In a portfolio, the weights $\{w_1, w_2, \ldots, w_n\}$ can be treated as a probability distribution, where $w_i$ is the proportion of the portfolio invested in asset $i$.
- Shannon entropy measures how evenly the investments are distributed across assets.
- A portfolio with equal weights (e.g., $w_i = \frac{1}{n}$) has maximum entropy, indicating complete
- diversification.

### *1.3.2 Risk Management*

- Portfolios with higher entropy are generally less risky because they spread exposure across multiple assets, reducing the impact of individual asset volatility.
- Shannon entropy penalizes concentration in a few assets, promoting stability.

### *1.3.3 Optimization Objective*

- Entropy can be used as an objective or constraint in portfolio optimization. For in- stance, one might maximize entropy while maintaining a target return, ensuring both high diversification and performance.

### *1.3.4 Connection to Exponential Growth*

- In long-term investment strategies, maximizing entropy is closely linked to maximizing the expected growth rate of wealth. This is because higher entropy portfolios are more resilient to fluctuations, improving compound growth over time.

## 1.4 How Shannon Entropy is Applied in Portfolio Theory

- In the paper on information geometry in portfolio theory, Shannon entropy is used to:
- **Define the Portfolio Map:** Shannon entropy plays a role in transforming portfolio weights into "dual coordinates" for geometric analysis.
- **Analyze Growth Rates:** The entropy is linked to the **excess growth rate** and other metrics, helping to quantify the trade-off between risk and return.
- **Evaluate Divergence:** By integrating entropy with divergence measures like $L(1)$- divergence, the paper provides tools for comparing and optimizing portfolios.

## 1.5 Real-Time Example: Using Shannon Entropy for Portfolio Diversification

### *1.5.1 Scenario: Constructing a Diversified Investment Portfolio*

Imagine an investor has $1,000,000 to allocate across 5 asset classes:

1. Stocks
2. Bonds
3. Real Estate
4. Commodities
5. Cash

### 1.5.2 Step 1: Portfolio Weights

The investor considers two allocation strategies:

- **Concentrated Portfolio:**
  - Stocks: 70%, Bonds: 20%, Real Estate: 5%, Commodities: 4%, Cash: 1%.
  - $W = \{0.7, 0.2, 0.05, 0.04, 0.01\}$
- **Diversified Portfolio:**
  - Stocks: 20%, Bonds: 20%, Real Estate: 20%, Commodities: 20%, Cash: 20%.
  - W = {0.2, 0.2, 0.2, 0.2, 0.2}

### 1.5.3 Step 2: Calculate Shannon Entropy

For each portfolio, entropy is calculated as:

$$H(W) = -\sum_{i=1}^{n} w_i \log(w_i)$$

**Concentrated Portfolio:**

$$H = -(0.7\log(0.7) + 0.2\log(0.2) + 0.05\log(0.05) + 0.04\log(0.04) + 0.01\log(0.01))$$

Approximation:

$$H \approx 0.249 + 0.464 + 0.216 + 0.201 + 0.067 = 0.902$$

**Diversified Portfolio:**

$$H = -(5 \times 0.2\log(0.2))$$

Approximation:

$$H \approx -5 \times 0.2 \times (-0.698) = 0.698$$

### 1.5.4 Step 3: Interpretation

- **Concentrated Portfolio:** Lower entropy ($H = 0.902$) indicates the portfolio is heavily weighted towards stocks, with minimal diversification.
- **Diversified Portfolio:** Higher entropy ($H = 0.698$) represents equal allocation across all asset classes, maximizing diversification.

### 1.5.5 Step 4: Decision-Making

- If the investor seeks higher growth but is willing to take on more risk, they might prefer the concentrated portfolio.
- If the investor prioritizes stability and risk mitigation, they would choose the diversified portfolio with higher entropy.

## 1.6 Practical Implications of Shannon Entropy in Portfolio Theory

1. **Automated Portfolio Management:** Modern portfolio management platforms use entropy to optimize allocations for robo-advisors and index funds.
2. **Balancing Diversification with Performance:** Shannon entropy ensures portfolios are neither overly diversified (which may dilute returns) nor overly concentrated (which increases risk).
3. **Dynamic Portfolio Adjustments:** Entropy can guide rebalancing strategies, suggesting shifts in weights as market conditions change.

## 1.7 Self-financed trading strategy

Definition and Framework

A self-financing trading strategy is represented by a sequence of portfolios $\eta(t)$ for $t \geq 0$, with each portfolio value defined in terms of holdings in $n$ assets. The portfolio evolves according to the self-financing identity:

$$\sum_{i=1}^{n} \eta_i(t)\mu_i(t+1) = \sum_{i=1}^{n} \eta_i(t+1)\mu_i(t+1),$$

where:

- $\eta_i(t)$ is the quantity of the $i$-th asset held at time $t$,
- $\mu_i(t)$ is the relative value (market weight) of the $i$-th asset at time $t$.

This identity ensures that changes in the portfolio value are exclusively due to price movements of the assets, with no additional capital being added or withdrawn.

## 1.8 Value Process of the Portfolio

The relative value process of the portfolio, denoted $V_\eta(t)$, tracks the portfolio's performance relative to the market and is given by:

$$V_\eta(t) = V_\eta(0) + \sum_{s=0}^{t-1} \eta(s) \cdot \mu(s+1) - \mu(s) ,$$

where:

- $\eta(s) \cdot \mu(s+1) - \mu(s)$ represents the portfolio's return at time $s$,
- $V_\eta(0)$ is the initial portfolio value.

For simplicity, the model assumes that all trading strategies are self-financed and studies portfolio value relative to the market portfolio.

## 1.9 Portfolio Weight Vector

When the portfolio value $V_\eta(t)$ is strictly positive for all $t$, the *portfolio weight vector* $\pi(t)$ can be defined as:

$$\pi_i(t) = \frac{\eta_i(t)\mu_i(t)}{V_\eta(t)}, \quad i = 1, \ldots, n.$$

## 1.10 Multiplicatively Generated Portfolio

A multiplicatively generated portfolio takes a different approach, focusing on exponentially concave generating functions that define portfolio weights through the derivatives in a few directions. This method constructs portfolios where the weight of each asset at time $t$ is derived from a function $\phi(\cdot)$ (the generating function), which is smooth and concave. The portfolio is defined by the mapping:

$$\eta_i(t) = (1 + \nabla_i\phi(t)),$$

where $i = 1, 2, \ldots, n$, $\nabla_i\phi(t)$ represents the portfolio allocation at time $t$, and $\nabla\phi(t)$ is the directional derivative along the tangent vector of a point $\phi$. This framework defines all long strategies, meaning the portfolio does not involve short positions.

The log-relative value is expressed as:

$$\log V(t) - \log V(0) = \phi(t) - \phi(0) + \sum_{s=0}^{t-1} \nabla\phi(s) \cdot \mu(s+1) - \mu(s) ,$$

where $X_i(t)$ is the capitalization of asset $i$ at time $t$. The market evolves as a path in this simplex over time.

The authors introduce measures of diversity and volatility to capture essential market characteristics:

- **Diversity:** Quantified using functions like Shannon entropy or other concave scoring functions.
- **Volatility:** Evaluated using divergences, which serve as "distances" between market states at successive time steps.

The identity is given as:

$$\sum_{i=1}^{n} \eta_i(t)\mu_i(t+1) = \sum_{i=1}^{n} \eta_i(t+1)\mu_i(t+1),$$

where:

- $\eta_i(t)$ is the quantity of the $i$-th asset held at time $t$,
- $\mu_i(t)$ is the relative value (market weight) of the $i$-th asset at time $t$.

## 1.11 Generalized Functional Portfolio Generation with Mathematical Framework

### *1.11.1 Motivations*

Portfolio construction aims to design strategies that manage wealth dynamically while adhering to self-financing constraints. The multiplicatively generated portfolio from Theorem ensures pathwise decomposition. However, alternative constructions, such as the additively generated portfolio introduced by Karatzas and Ruf in **Karatzas2011**, offer novel approaches. These portfolios are generated using a smooth, concave function $\phi$, enabling additive pathwise decompositions.[1]

**Theorem 3: Additively Generated Portfolio** Let $\phi : \mathbb{R}^n \to \mathbb{R}$ be a smooth, concave function, and let $V(0) = v_0$ be the initial portfolio value. Then, the self-financing strategy satisfies:

$$V(t) = V(0) + \phi(\pi(t)) - \phi(\pi(0)) + \sum_{s=0}^{t-1} D^{(0)}[\pi(s+1) \mid \pi(s)], \tag{5.1}$$

where $D^{(0)}[\pi(s+1) \mid \pi(s)]$ is the Bregman divergence of $\phi$ at time $s$:

$$D^{(0)}[\pi(s+1) \mid \pi(s)] = \phi(\pi(s+1)) - \phi(\pi(s)) - \nabla\phi(\pi(s))^\top(\pi(s+1) - \pi(s)). \tag{5.2}$$

The trading strategy is explicitly given by:

$$\pi_i(t) = V(t) \cdot \frac{\partial \phi}{\partial p_i}(\pi(t)), \quad i = 1, \dots, n, \tag{5.3}$$

where $\pi(t) = (\pi_1(t), \pi_2(t), \dots, \pi_n(t))$ represents portfolio weights.

**Example: Quadratic Function** Consider $\phi(p) = -\frac{1}{2}|p|^2 = -\frac{1}{2}\sum_{i=1}^{n} p_i^2$.
Its gradient and Hessian are:

$$\nabla\phi(p) = -p, \quad \text{and} \quad \mathrm{Hess}(\phi) = -I, \tag{5.4}$$

where $I$ is the identity matrix. The corresponding Bregman divergence is:

$$D^{(0)}[q \mid p] = \frac{1}{2}|q - p|^2. \tag{5.5}$$

This divergence represents the squared Euclidean distance, assigning financial meaning to geometric changes in portfolio weights.A New Functional Portfolio Generation

The $(\lambda, C)$-generation generalizes portfolio construction, incorporating parameters $\lambda > 0$ and $C \geq 0$. For a smooth $\lambda$-exponentially concave function $\phi$, the self-financing strategy satisfies:

$$\pi_i(t) = \frac{(C + V(t)) \cdot \frac{\partial \phi}{\partial p_i}(\pi(t))}{V(t)}, \quad i = 1, \dots, n. \tag{5.6}$$

The portfolio value evolves as:

$$V(t+1) = V(t) + \pi(t)^\top(\pi(t+1) - \pi(t)). \tag{5.7}$$

This ensures the pathwise decomposition:

$$\frac{1}{\lambda}\log(C + V(t)) - \frac{1}{\lambda}\log(C + V(0)) = \phi(\pi(t)) - \phi(\pi(0)) + \sum_{s=0}^{t-1} D^{(\lambda)}[\pi(s+1) \mid \pi(s)], \tag{5.8}$$

where $D^{(\lambda)}[\pi(s+1) \mid \pi(s)]$ is the $L^{(\lambda)}$-divergence:

$$D^{(\lambda)}[\pi(s+1) \mid \pi(s)] = \phi(\pi(s+1)) - \phi(\pi(s)) - \frac{1}{\lambda}\nabla\phi(\pi(s))^\top(\pi(s+1) - \pi(s)). \tag{5.9}$$

**Portfolio Weight Interpretation** The weight vector $\pi(t)$ of the $(\lambda, C)$-generated strategy can be decomposed into two components:

$$\pi_i(t) = \frac{1}{V(t)}\left[C \cdot \pi_i^{\text{market}}(t) + V(t) \cdot \pi_i^{\text{mult}}(t)\right], \tag{5.10}$$

where $\pi^{\text{market}}(t)$ represents the market portfolio weights, and $\pi^{\text{mult}}(t)$ represents the weights from the multiplicatively generated portfolio.

By increasing $C$, the portfolio becomes more aggressive, amplifying deviations from market behavior. For $C = 0$ and $\lambda = 1$, the strategy reduces to a multiplicatively generated portfolio. In the limit $\lambda \to 0$ with $C \to \infty$, the strategy approximates additive generation.

### *1.11.2 Conclusion*

This framework unifies additive and multiplicative portfolio generation methods through the $(\lambda, C)$-generation, providing a flexible tool for portfolio construction. The decomposition via Bregman divergence and its variants establishes a strong geometric and financial foundation for analyzing portfolio dynamics. These methods extend the applicability of functional portfolio theory, setting the stage for future research into generalized divergence metrics and continuous- time extensions.

## 1.12 Interpolation and Comparison: Methods and Findings

### *1.12.1* Interpolation of Exponentially Concave Functions

A key property of exponentially concave functions is their ability to be interpolated while maintaining exponential concavity. Given two exponentially concave functions, $\Phi^{(0)}$ and $\Phi^{(1)}$, defined on $\Delta_n$, the interpolation:

$$\Phi^{(\lambda)}(p) = (1-\lambda)\Phi^{(0)}(p) + \lambda\Phi^{(1)}(p), \quad 0 < \lambda < 1,$$

results in another exponentially concave function.

If $\Phi^{(0)}$ and $\Phi^{(1)}$ generate portfolio maps $\pi^{(0)}(p)$ and $\pi^{(1)}(p)$ respectively, then $\Phi^{(\lambda)}$ generates the portfolio map:

$$\pi^{(\lambda)}(p) = (1-\lambda)\pi^{(0)}(p) + \lambda\pi^{(1)}(p).$$

This implies that the space of exponentially concave functions and their associated *multiplicatively generated portfolios* is convex. Additionally, the *L-divergence* $D^{(1)}[\cdot|\cdot]$, associated with these functions, is concave in $\Phi$. This property extends to solving problems such as displacement interpolation for logarithmic optimal transport.

### *1.12.2 Dominance of L-Divergences*

For two L-divergences $D^{(1)}_{\Phi}$ and $D^{(1)}_{\Psi}$, generated by the functions $\Phi$ and $\Psi$ respectively, we say $D^{(1)}_{\Phi}$ dominates $D^{(1)}_{\Psi}$ if:

$$D^{(1)}_{\Phi}[q|p] \geq D^{(1)}_{\Psi}[q|p], \qquad \forall p, q \in \Delta_n.$$

This means that the portfolio generated by $\Phi$ captures more market volatility than the one generated by $\Psi$. Identifying maximal elements in this partial order is a significant problem in the study of portfolio generation.

Theorem 7: Maximal Generating Functions

**Theorem 7:** establishes conditions for a generating function to be *maximal* in the partial order of L-divergences:

1. If $\Phi$ is symmetric, meaning it is invariant under any permutation of the coordinates, and satisfies:

$$\Phi(p) = \int_0^1 \exp \left( (1-t)e_1 + te \right) dt,$$

where $e_1, e_2, \ldots, e_n$ are standard basis vectors, then $D^{(1)}_{\Phi}$ is maximal in the partial order.

2. A generating function $\Phi$ is maximal if it captures all possible market volatility. Examples include:
   - $\Phi(p) = \frac{1}{n}\sum_{i=1}^{n} \log p_i$: This function generates the *equal-weighted portfolio.*
   - $\Phi(p) = \log \sum_{i=1}^{n} p_i \log p_i$ : This function is related to the *logarithm of Shannon entropy*.

### *1.12.3 Global and Local Properties*

Theorem 7 focuses on global properties of generating functions and their maximality across the entire domain $\Delta_n$. However, local properties of generating functions can also be studied. For instance, constructing *short-term relative arbitrages* involves identifying locally maximal exponentially concave functions, as explored in other research.

### *1.12.4 Findings*

1. **Convexity of Portfolio Spaces:** The spaces of exponentially concave functions and their associated portfolio maps form convex sets. This enables the interpolation of portfolios and generating functions while preserving their concavity and financial properties.

2. **Dominance of L-Divergences:** A portfolio that generates a larger L-divergence captures greater market volatility, providing a method to rank portfolios based on their ability to exploit market dynamics.

3. **Maximal Generating Functions:** Symmetric functions like the logarithm of Shannon entropy are globally maximal, meaning they dominate other generating functions in terms of volatility capture.**Applications in Optimal Transport:** Interpolating generating functions provides a geometric framework for solving displacement interpolation problems in optimal transport theory.

These findings demonstrate the versatility of exponentially concave generating functions in portfolio theory and their ability to model and optimize market dynamics effectively.

## References:

[1] T. K. L. Wong, “Information geometry in portfolio theory,” in Information Geometry in Portfolio Theory, Springer, 2019, pp. 121–145. doi: 10.1007/978-3-030-02520-5_6.
[2] M. Suzuki, Information Geometry and Statistical Manifold. Springer, 2014.
[3] F. Critchley and P. Marriott, “Information geometry and its applications: An overview,” in Information Geometry and Its Applications, Springer, 2017, pp. 1–22. doi: 10.1007/978-3-319-47058-0_1.
[4] S.-I. Amari, Information Geometry and Its Applications. Springer, 2016.

[5] Cover’s universal portfolio and stochastic portfolio theory, 2017

[6] Stochastic Portfolio Theory: an Overview, 2008