Topiarism: The Kernel Embedding of Distributions Applied to Modern Portfolio Theory

Stephen Cook

California Polytechnic State University, San Luis Obispo

June 10, 2025

Motivation

What is a Portfolio?

What is a Portfolio?

Definition

Given a set of assets/stocks $X_1, X_2, \cdots X_n$, a ***portfolio*** is a series of weights assigned to each asset that corresponds to what proportion of your total funds you wish to invest in it.

What is a Portfolio?

Definition

Given a set of assets/stocks $X_1, X_2, \cdots X_n$, a ***portfolio*** is a series of weights assigned to each asset that corresponds to what proportion of your total funds you wish to invest in it.

Question: What portfolio SHOULD you invest in?

Modern Portfolio Theory

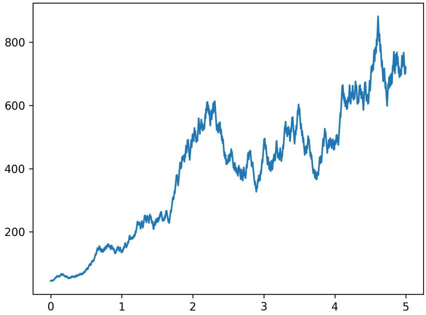
Modern Portfolio Theory

$$\lambda \psi^\intercal w - \frac{1}{2} w^\intercal C w$$

Modern Portfolio Theory

$$\lambda \psi^\top w - \frac{1}{2} w^\top C w$$

This is what stocks look like



Consequences of Modern Portfolio Theory

- Diversification

Consequences of Modern Portfolio Theory

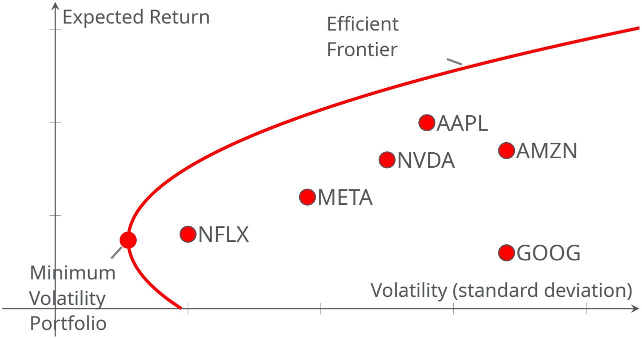
- Diversification
- Easy(ish) to compute

Consequences of Modern Portfolio Theory

- Diversification
- Easy(ish) to compute
- The efficient frontier

Consequences of Modern Portfolio Theory

- Diversification
- Easy(ish) to compute
- The efficient frontier



4 Problems with Modern Portfolio Theory That We Care About

4 Problems with Modern Portfolio Theory That We Care About

- It allows for negative weights in portfolios

4 Problems with Modern Portfolio Theory That We Care About

- It allows for negative weights in portfolios
- It is sometimes hard to find the best portfolio analytically

4 Problems with Modern Portfolio Theory That We Care About

- It allows for negative weights in portfolios
- It is sometimes hard to find the best portfolio analytically
- True optimizers are impossible to observe

4 Problems with Modern Portfolio Theory That We Care About

- It allows for negative weights in portfolios
- It is sometimes hard to find the best portfolio analytically
- True optimizers are impossible to observe
- The best portfolios use too many assets

Our Goals

Our Goals

- Optimizing should be done over positive portfolios.

Our Goals

- Optimizing should be done over positive portfolios.
- Approximation over subsets of stocks should get close to the ”true” maximum.

Our Goals

- Optimizing should be done over positive portfolios.
- Approximation over subsets of stocks should get close to the ”true” maximum.
- Finding/estimating the best portfolios should be efficient.

Our Goals

- Optimizing should be done over positive portfolios.
- Approximation over subsets of stocks should get close to the ”true” maximum.
- Finding/estimating the best portfolios should be efficient.
- The model should match the trend of the market while using less stocks.

Our Goals

- ♥ Optimizing should be done over positive portfolios.
- ♧ Approximation over subsets of stocks should get close to the ”true” maximum.
- ♤ Finding/estimating the best portfolios should be efficient.
- ♦ The model should match the trend of the market while using less stocks.
- $ The model should preserve some of the efficient frontier’s structure.

Measure Theory

What is a Measure?

Informal Definition

*A **measure** is a way of measuring weighted length and area under curves.*

What is a Measure?

Informal Definition

*A **measure** is a way of measuring weighted length and area under curves.*

For example: dx

What is a Measure?

Informal Definition

*A **measure** is a way of measuring weighted length and area under curves.*

For example: dx

- How do you measure the length of the interval $[a, b]$:

$$\int_a^b dx = b - a$$

What is a Measure?

Informal Definition

*A **measure** is a way of measuring weighted length and area under curves.*

For example: dx

- How do you measure the length of the interval $[a, b]$:

$$\int_a^b dx = b - a$$

- How do you measure the area under a curve:

$$\int_a^b f(x)\, dx$$

What is a Measure?

Informal Definition

*A **measure** is a way of measuring weighted length and area under curves.*

For example: dx

- How do you measure the length of the interval $[a, b]$:

$$\int_E dx = \int_a^b dx = b - a$$

- How do you measure the area under a curve:

$$\int_a^b f(x)\, dx$$

What is a Measure?

Informal Definition

*A **measure** is a way of measuring weighted length and area under curves.*

For example: dx

- How do you measure the length of the interval $[a, b]$:

$$\int_E dx = \int_a^b dx = b - a$$

- How do you measure the area under a curve:

$$\int_E f\,dx = \int_a^b f(x)\,dx$$

Other Examples of Measures

The ***Dirac*** measure

$$d\delta_a = \delta_a(x)dx$$

- How do you measure length:
- How do you measure area under a curve:

The ***Gaussian*** measure

$$d\mu = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} dx$$

- How do you measure length:
- How do you measure area under a curve:

Other Examples of Measures

The ***Dirac*** measure

$$d\delta_a = \delta_a(x)dx$$

- How do you measure length:

$$\int_E d\delta_a = \int_E \delta_a(x)\, dx = \begin{cases} 1, & a \in E, \\ 0, & a \notin E. \end{cases}$$

- How do you measure area under a curve:

The ***Gaussian*** measure

$$d\mu = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} dx$$

- How do you measure length:

$$\int_E d\mu = \int_E \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}\, dx$$

- How do you measure area under a curve:

Other Examples of Measures

The ***Dirac*** measure

$$d\delta_a = \delta_a(x)dx$$

- How do you measure length:

$$\int_E d\delta_a = \int_E \delta_a(x)\,dx = \begin{cases} 1, & a \in E, \\ 0, & a \notin E. \end{cases}$$

- How do you measure area under a curve:

$$\int_E f\,d\delta_a = \int_E f(x)\delta_a(x)\,dx = \begin{cases} f(a), & a \in E, \\ 0, & a \notin E. \end{cases}$$

The ***Gaussian*** measure

$$d\mu = \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}dx$$

- How do you measure length:

$$\int_E d\mu = \int_E \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}\,dx$$

- How do you measure area under a curve:

$$\int_E f\,d\mu = \int_E f(x)\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}\,dx$$

Other Examples of Measures

The ***Dirac*** measure

$$d\delta_a = \delta_a(x)dx$$

- How do you measure length:

$$\delta_a(E) = \int_E d\delta_a = \int_E \delta_a(x)\,dx = \begin{cases} 1, & a \in E, \\ 0, & a \notin E. \end{cases}$$

- How do you measure area under a curve:

$$\int_E f\,d\delta_a = \int_E f(x)\delta_a(x)\,dx = \begin{cases} f(a), & a \in E, \\ 0, & a \notin E. \end{cases}$$

The ***Gaussian*** measure

$$d\mu = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}\,dx$$

- How do you measure length:

$$\mu(E) = \int_E d\mu = \int_E \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}\,dx$$

- How do you measure area under a curve:

$$\int_E f\,d\mu = \int_E f(x)\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}\,dx$$

Cool Properties of Measures

Let $\mathcal{M}(\Omega)$ be the set of measures on some set Ω.

Cool Properties of Measures

Let $\mathcal{M}(\Omega)$ be the set of measures on some set Ω.

- $\mathcal{M}(\Omega)$ is closed under addition:

$$\int f d(\mu + \nu) = \int f d\mu + \int f d\nu$$

Cool Properties of Measures

Let $\mathcal{M}(\Omega)$ be the set of measures on some set Ω.

- $\mathcal{M}(\Omega)$ is closed under addition:

$$\int f d(\mu + \nu) = \int f d\mu + \int f d\nu$$

- $\mathcal{M}(\Omega)$ is closed under scalar multiplication:

$$\int f d(\alpha\mu) = \alpha \int f d\mu$$

Cool Properties of Measures

Let $\mathcal{M}(\Omega)$ be the set of measures on some set Ω.

- $\mathcal{M}(\Omega)$ is closed under addition:

$$\int f\,d(\mu+\nu) = \int f\,d\mu + \int f\,d\nu$$

- $\mathcal{M}(\Omega)$ is closed under scalar multiplication:

$$\int f\,d(\alpha\mu) = \alpha \int f\,d\mu$$

- If $\mu, \nu \in \mathcal{M}(\Omega)$ are probability measures, then $(1-t)\mu + t\nu$ is a probability measure:

$$\int d((1-t)\mu + t\nu) = (1-t)\int d\mu + t\int d\nu = (1-t) + t = 1$$

Cool Properties of Measures

Let $\mathcal{M}(\Omega)$ be the set of measures on some set Ω.

- $\mathcal{M}(\Omega)$ is closed under addition:

$$\int f d(\mu + \nu) = \int f d\mu + \int f d\nu$$

- $\mathcal{M}(\Omega)$ is closed under scalar multiplication:

$$\int f d(\alpha\mu) = \alpha \int f d\mu$$

$\mathcal{M}(\Omega)$ is a vector space!

- If $\mu, \nu \in \mathcal{M}(\Omega)$ are probability measures, then $(1-t)\mu + t\nu$ is a probability measure:

$$\int d((1-t)\mu + t\nu) = (1-t)\int d\mu + t\int d\nu = (1-t) + t = 1$$

Cool Properties of Measures

Let $\mathcal{M}(\Omega)$ be the set of measures on some set Ω.

- $\mathcal{M}(\Omega)$ is closed under addition:

$$\int f d(\mu + \nu) = \int f d\mu + \int f d\nu$$

- $\mathcal{M}(\Omega)$ is closed under scalar multiplication:

$$\int f d(\alpha\mu) = \alpha \int f d\mu$$

$\mathcal{M}(\Omega)$ is a vector space!

- If $\mu, \nu \in \mathcal{M}(\Omega)$ are probability measures, then $(1-t)\mu + t\nu$ is a probability measure:

$$\int d((1-t)\mu + t\nu) = (1-t)\int d\mu + t\int d\nu = (1-t) + t = 1$$

What other structure does $\mathcal{M}(\Omega)$ have?

Weak* Topology

Weak* Topology

Definition

Let V be a vector space a functions on a set Ω, and let $\mathcal{M}(\Omega)$ be the vector space of measures on Ω. We say a sequence of measures $\mu_n \subseteq \mathcal{M}(\Omega)$ approaches a measure μ ***weak*-ly*** or ***vaguely*** if

$$\int f\,d\mu_n \to \int f\,d\mu$$

for all $f \in V$.

Weak* Topology

Definition

Let V be a vector space a functions on a set Ω, and let $\mathcal{M}(\Omega)$ be the vector space of measures on Ω. We say a sequence of measures $\mu_n \subseteq \mathcal{M}(\Omega)$ approaches a measure μ ***weak*-ly*** or ***vaguely*** if

$$\int f \, d\mu_n \to \int f \, d\mu$$

for all $f \in V$.

- Under this topology, the space of nonnegative probability measures $\mathcal{M}_+^1(\Omega)$ is compact

Why Do We Care?

Why Do We Care?

Portfolio $w \iff$ Probability Measure μ_w

Why Do We Care?

$$\text{Portfolio } w \iff \text{Probability Measure } \mu_w$$

$$\text{Portfolio Payout } \psi^\top w \iff \text{Measure Payout } \int \psi \, d\mu_w$$

Why Do We Care?

Individual Assets $X \iff$ Dirac Measures δ_X

Portfolio $w \iff$ Probability Measure μ_w

Portfolio Payout $\psi^\intercal w \iff$ Measure Payout $\int \psi \, d\mu_w$

Why Do We Care?

Individual Assets $X \iff$ Dirac Measures δ_X

Portfolio $w \iff$ Probability Measure μ_w

Portfolio Payout $\psi^\top w \iff$ Measure Payout $\int \psi \, d\mu_w$

$$\Omega \xrightarrow{X \mapsto \delta_X} \mathcal{M}(\Omega)$$

Reproducing Kernel Hilbert Spaces

Hilbert Spaces

Hilbert Spaces

Definition

Let $\mathcal{H}$ be a vector space over $\mathbb{F}$. We call $\mathcal{H}$ a ***Hilbert space*** if

1. $\mathcal{H}$ is an inner product space with inner product $\langle \cdot, \cdot \rangle$, and
2. Cauchy sequences are convergent.

Hilbert Spaces

Definition

Let $\mathcal{H}$ be a vector space over $\mathbb{F}$. We call $\mathcal{H}$ a ***Hilbert space*** if

1. $\mathcal{H}$ is an inner product space with inner product $\langle \cdot, \cdot \rangle$, and
2. Cauchy sequences are convergent.

Examples include:

- $\mathbb{C}^n$ with inner product $\left\langle \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}, \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} \right\rangle = \sum_{k=1}^{n} a_k \overline{b_k}$

Hilbert Spaces

Definition

Let $\mathcal{H}$ be a vector space over $\mathbb{F}$. We call $\mathcal{H}$ a ***Hilbert space*** if

1. $\mathcal{H}$ is an inner product space with inner product $\langle \cdot, \cdot \rangle$, and
2. Cauchy sequences are convergent.

Examples include:

- $\mathbb{C}^n$ with inner product $\left\langle \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}, \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} \right\rangle = \sum_{k=1}^{n} a_k \overline{b_k}$
- $\ell^2 = \left\{ (x_n)_{n \in \mathbb{N}} \subseteq \mathbb{C} : \sum_{n=1}^{\infty} |x_n|^2 < \infty \right\}$ with inner product $\left\langle \begin{bmatrix} a_1 \\ a_2 \\ \vdots \end{bmatrix}, \begin{bmatrix} b_1 \\ b_2 \\ \vdots \end{bmatrix} \right\rangle_{\ell^2} = \sum_{k=1}^{\infty} a_k \overline{b_k}.$

The Hardy Space

The **Hardy space** is the space of analytic power series on the complex unit disk, and is formally defined by

$$H^2 = \left\{ f(z) = \sum_{n=0}^{\infty} a_n z^n : \sum_{n=1}^{\infty} |a_n|^2 < \infty \right\}.$$

The Hardy Space

The **Hardy space** is the space of analytic power series on the complex unit disk, and is formally defined by

$$H^2 = \left\{ f(z) = \sum_{n=0}^{\infty} a_n z^n : \sum_{n=1}^{\infty} |a_n|^2 < \infty \right\}.$$

$$\begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_{n-1} \end{bmatrix} \in \mathbb{C}^n$$

$$\Updownarrow$$

$$a_0 + a_1 z + \cdots + a_{n-1} z^{n-1} \in P_{n-1}(\mathbb{C})$$

The Hardy Space

The **Hardy space** is the space of analytic power series on the complex unit disk, and is formally defined by

$$H^2 = \left\{ f(z) = \sum_{n=0}^{\infty} a_n z^n : \sum_{n=1}^{\infty} |a_n|^2 < \infty \right\}.$$

$$\begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_{n-1} \end{bmatrix} \in \mathbb{C}^n \qquad \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ \vdots \end{bmatrix} \in \ell^2$$

$$\Updownarrow \qquad \Updownarrow$$

$$a_0 + a_1 z + \cdots + a_{n-1} z^{n-1} \in P_{n-1}(\mathbb{C}) \qquad a_0 + a_1 z + a_2 z^2 + \cdots \in H^2$$

The Hardy Space

The **Hardy space** is the space of analytic power series on the complex unit disk, and is formally defined by

$$H^2 = \left\{ f(z) = \sum_{n=0}^{\infty} a_n z^n : \sum_{n=1}^{\infty} |a_n|^2 < \infty \right\}.$$

$$\begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_{n-1} \end{bmatrix} \in \mathbb{C}^n \qquad \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ \vdots \end{bmatrix} \in \ell^2$$

$$\Updownarrow \qquad \Updownarrow$$

$$a_0 + a_1 z + \cdots + a_{n-1} z^{n-1} \in P_{n-1}(\mathbb{C}) \qquad a_0 + a_1 z + a_2 z^2 + \cdots \in H^2$$

- **Fact**: H^2 is a Hilbert space with inner product

$$\left\langle \sum_{n=0}^{\infty} a_n z^n, \sum_{n=0}^{\infty} b_n z^n \right\rangle_{H^2} = \sum_{n=0}^{\infty} a_n \overline{b_n}$$

The Hardy Space

The **Hardy space** is the space of analytic power series on the complex unit disk, and is formally defined by

$$H^2 = \left\{ f(z) = \sum_{n=0}^{\infty} a_n z^n : \sum_{n=1}^{\infty} |a_n|^2 < \infty \right\}.$$

$$\begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_{n-1} \end{bmatrix} \in \mathbb{C}^n \qquad\qquad \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ \vdots \end{bmatrix} \in \ell^2$$

$$\Updownarrow \qquad\qquad \Updownarrow$$

$$a_0 + a_1 z + \cdots + a_{n-1} z^{n-1} \in P_{n-1}(\mathbb{C}) \qquad\qquad a_0 + a_1 z + a_2 z^2 + \cdots \in H^2$$

- **Fact**: H^2 is a Hilbert space with inner product

$$\left\langle \sum_{n=0}^{\infty} a_n z^n, \sum_{n=0}^{\infty} b_n z^n \right\rangle_{H^2} = \sum_{n=0}^{\infty} a_n \overline{b_n} = \left\langle \begin{bmatrix} a_0 \\ a_1 \\ \vdots \end{bmatrix}, \begin{bmatrix} b_0 \\ b_1 \\ \vdots \end{bmatrix} \right\rangle_{\ell^2}$$

The Szegő Function in the Hardy Space

For each $|w| < 1$, define the **Szegő function** at w to be $k_w(z) = \dfrac{1}{1 - \overline{w}z}$ for $|z| < 1$.

The Szegő Function in the Hardy Space

For each $|w| < 1$, define the **Szegő function** at w to be $k_w(z) = \dfrac{1}{1 - \overline{w}z}$ for $|z| < 1$.

Question: What is $\langle f, k_w \rangle$?

The Szegő Function in the Hardy Space

For each $|w| < 1$, define the **Szegő function** at w to be $k_w(z) = \dfrac{1}{1 - \overline{w}z}$ for $|z| < 1$.

Question: What is $\langle f, k_w \rangle$?

- For all $|w| < 1$ and for all $f \in H^2$,

  $$f(w) = \langle f, k_w \rangle \,,$$

  so we say k_w has a ***reproducing property*** for w and call it the **reproducing kernel** at the point w.

The Szegő Function in the Hardy Space

For each $|w| < 1$, define the **Szegő function** at w to be $k_w(z) = \dfrac{1}{1 - \overline{w}z}$ for $|z| < 1$.

Question: What is $\langle f, k_w \rangle$?

- For all $|w| < 1$ and for all $f \in H^2$,

  $$f(w) = \langle f, k_w \rangle ,$$

  so we say k_w has a ***reproducing property*** for w and call it the **reproducing kernel** at the point w.

- We call the function $K(z, w) = k_w(z)$ a ***kernel function*** for our Hilbert space.

Reproducing Kernel Hilbert Spaces

Reproducing Kernel Hilbert Spaces

Definition

Let Ω be a set and let $\mathcal{H} \subseteq \mathbb{F}^{\Omega}$ be a set of functions from Ω to $\mathbb{F}$. We call $\mathcal{H}$ a ***reproducing kernel Hilbert space*** (often shortened to be ***RKHS***) on Ω if

1. $\mathcal{H}$ is a Hilbert space, and
2. there exists a kernel function $K\colon \Omega \times \Omega \to \mathbb{F}$ for $\mathcal{H}$.

Reproducing Kernel Hilbert Spaces

Definition

Let Ω be a set and let $\mathcal{H} \subseteq \mathbb{F}^\Omega$ be a set of functions from Ω to $\mathbb{F}$. We call $\mathcal{H}$ a ***reproducing kernel Hilbert space*** (often shortened to be ***RKHS***) on Ω if

1. $\mathcal{H}$ is a Hilbert space, and
2. there exists a kernel function $K\colon \Omega \times \Omega \to \mathbb{F}$ for $\mathcal{H}$.

- Example: H^2 with kernel function $K(z, w) = \dfrac{1}{1 - \overline{w}z}$

Importance of RKHSs

They are used in a variety of different fields:

Importance of RKHSs

They are used in a variety of different fields:

- Operator theory
- Probability theory
- Image and signal processing
- Machine learning
- Support vector machines
- Quantum mechanics

Importance of RKHSs

They are used in a variety of different fields:

- Operator theory
- Probability theory
- Image and signal processing
- Machine learning
- Support vector machines
- Quantum mechanics
- They're just cool!

A Few Cool Properties of RKHSs

A Few Cool Properties of RKHSs

Theorem

Let $\mathcal{H}$ be a reproducing kernel Hilbert space and let K be its kernel function. If K is continuous on $\Omega \times \Omega$, then every function $f \in \mathcal{H}$ is continuous on Ω.

A Few Cool Properties of RKHSs

Theorem

Let $\mathcal{H}$ be a reproducing kernel Hilbert space and let K be its kernel function. If K is continuous on $\Omega \times \Omega$, then every function $f \in \mathcal{H}$ is continuous on Ω.

- All continuous functions are integrable

A Few Cool Properties of RKHSs

Theorem

Let $\mathcal{H}$ be a reproducing kernel Hilbert space and let K be its kernel function. If K is continuous on $\Omega \times \Omega$, then every function $f \in \mathcal{H}$ is continuous on Ω.

- All continuous functions are integrable

Theorem

Let $\mathcal{H}$ be a reproducing kernel Hilbert space and let $A = \{k_z \in \mathcal{H} : z \in \Omega\}$ be the set of kernel functions in $\mathcal{H}$. Then the linear span of A is dense in $\mathcal{H}$.

A Few Cool Properties of RKHSs

Theorem

Let $\mathcal{H}$ be a reproducing kernel Hilbert space and let K be its kernel function. If K is continuous on $\Omega \times \Omega$, then every function $f \in \mathcal{H}$ is continuous on Ω.

- All continuous functions are integrable

Theorem

Let $\mathcal{H}$ be a reproducing kernel Hilbert space and let $A = \{k_z \in \mathcal{H} : z \in \Omega\}$ be the set of kernel functions in $\mathcal{H}$. Then the linear span of A is dense in $\mathcal{H}$.

- We can approximate elements of $\mathcal{H}$ with kernel functions
- Any cool properties of kernel functions will translate over to all elements of $\mathcal{H}$

Positive Semidefinite Kernel Functions and Aronszajn's Theorem

Positive Semidefinite Kernel Functions and Aronszajn's Theorem

- Question: If we have a function $K\colon \Omega \times \Omega \to \mathbb{F}$, can we know if it is a kernel function for some RKHS?

Positive Semidefinite Kernel Functions and Aronszajn's Theorem

- Question: If we have a function $K\colon \Omega \times \Omega \to \mathbb{F}$, can we know if it is a kernel function for some RKHS?
- Answer: Yes!

Positive Semidefinite Kernel Functions and Aronszajn's Theorem

- Question: If we have a function $K \colon \Omega \times \Omega \longrightarrow \mathbb{F}$, can we know if it is a kernel function for some RKHS?
- Answer: Yes!

Definition

Let Ω be a set and let $K \colon \Omega \times \Omega \longrightarrow \mathbb{F}$. We call K a ***positive semidefinite kernel function*** if K is conjugate symmetric and for all $x_1, \cdots, x_n \in \Omega$, the matrix $\big(K(x_i, x_j)\big)$ is positive semidefinite.

Positive Semidefinite Kernel Functions and Aronszajn's Theorem

- Question: If we have a function $K \colon \Omega \times \Omega \longrightarrow \mathbb{F}$, can we know if it is a kernel function for some RKHS?
- Answer: Yes!

Definition

Let Ω be a set and let $K \colon \Omega \times \Omega \longrightarrow \mathbb{F}$. We call K a ***positive semidefinite kernel function*** if K is conjugate symmetric and for all $x_1, \cdots, x_n \in \Omega$, the matrix $\big(K(x_i, x_j)\big)$ is positive semidefinite.

Theorem (Moore-Aronszajn)

Let Ω be a set and let $K \colon \Omega \times \Omega \longrightarrow \mathbb{F}$. Then the following are equivalent.

1. *K is a positive semidefinite kernel function.*
2. *There exists a RKHS $\mathcal{H}$ such that K is its reproducing kernel.*

Covariance is a Positive Semidefinite Kernel Function

Let $\Omega = \{X_1, \cdots X_n\}$ be a set of assets, and let $r_1, \cdots, r_n$ be the rates of return.

Covariance is a Positive Semidefinite Kernel Function

Let $\Omega = \{X_1, \cdots X_n\}$ be a set of assets, and let $r_1, \cdots, r_n$ be the rates of return.

$K(X_i, X_j) = \mathrm{Cov}(r_i, r_j)$ is a positive semidefinite kernel function!

Covariance is a Positive Semidefinite Kernel Function

Let $\Omega = \{X_1, \cdots X_n\}$ be a set of assets, and let $r_1, \cdots, r_n$ be the rates of return.

$$K(X_i, X_j) = \mathrm{Cov}(r_i, r_j) \text{ is a positive semidefinite kernel function!}$$

This gives us a unique RKHS $\mathcal{H}$ on Ω with the following properties:

- $\left\langle k_{X_i}, k_{X_j} \right\rangle = \mathrm{Cov}(r_i, r_j)$
- $\left\| k_{X_i} \right\|^2 = \mathrm{Var}(r_i)$

Covariance is a Positive Semidefinite Kernel Function

Let $\Omega = \{X_1, \cdots X_n\}$ be a set of assets, and let $r_1, \cdots, r_n$ be the rates of return.

$K(X_i, X_j) = \mathrm{Cov}(r_i, r_j)$ is a positive semidefinite kernel function!

This gives us a unique RKHS $\mathcal{H}$ on Ω with the following properties:

- $\left\langle k_{X_i}, k_{X_j} \right\rangle = \mathrm{Cov}(r_i, r_j)$
- $\left\| k_{X_i} \right\|^2 = \mathrm{Var}(r_i)$

$$\Omega \xrightarrow{x \longmapsto k_x} \mathcal{H}$$

The Kernel Embedding of Distributions

$$\Omega \xrightarrow{x \mapsto k_x} \mathcal{H}$$

The Kernel Embedding of Distributions

$$\begin{tikzcd}
\Omega \arrow[r, "x \mapsto \delta_x"] \arrow[dr, "x \mapsto k_x"'] & \mathcal{M}(\Omega) \\
& \mathcal{H}
\end{tikzcd}$$

The Kernel Embedding of Distributions

$$
\begin{tikzcd}
\Omega \arrow[r, "x \mapsto \delta_x"] \arrow[dr, "x \mapsto k_x"'] & \mathcal{M}(\Omega) \arrow[d] \\
& \mathcal{H}
\end{tikzcd}
$$

The Kernel Embedding of Distributions

$$\begin{tikzcd}
\Omega \arrow[r, "x \mapsto \delta_x"] \arrow[dr, "x \mapsto k_x"'] & \mathcal{M}(\Omega) \arrow[d] \\
& \mathcal{H}
\end{tikzcd}$$

Definition

Let Ω be a set, let K be a continuous positive semidefinite kernel function on Ω, and let $\mathcal{H}$ be the induced RKHS. Then the ***kernel embedding of distributions*** is the map $\hat{\phi} : \mathcal{M}(\Omega) \to \mathcal{H}$ defined by $\hat{\phi}(\mu) = \hat{\mu}$ where

$$\hat{\mu}(x) = \int k_x(z)\, d\mu(z).$$

The Kernel Embedding of Distributions

$$\begin{tikzcd}
\Omega \arrow[r, "x \mapsto \delta_x"] \arrow[dr, "x \mapsto k_x"'] & \mathcal{M}(\Omega) \arrow[d, "\mu \mapsto \hat{\mu}"] \\
& \mathcal{H}
\end{tikzcd}$$

Definition

Let Ω be a set, let K be a continuous positive semidefinite kernel function on Ω, and let $\mathcal{H}$ be the induced RKHS. Then the ***kernel embedding of distributions*** is the map $\hat{\phi} : \mathcal{M}(\Omega) \to \mathcal{H}$ defined by $\hat{\phi}(\mu) = \hat{\mu}$ where

$$\hat{\mu}(x) = \int k_x(z)\, d\mu(z).$$

Connection to Modern Portfolio Theory

Connection to Modern Portfolio Theory

$$w \iff \mu_w$$

$$\psi^\top w \iff \int \psi \, d\mu_w$$

Connection to Modern Portfolio Theory

$$
\begin{aligned}
w &\iff \mu_w \\
\psi^\top w &\iff \int \psi \, d\mu_w \\
w^\top C w &\iff ||\hat{\mu}_w||^2
\end{aligned}
$$

Connection to Modern Portfolio Theory

$$\begin{aligned} w &\iff \mu_w \\ \psi^\top w &\iff \int \psi \, d\mu_w \\ w^\top C w &\iff ||\hat{\mu}_w||^2 \end{aligned}$$

Modern Portfolio Theory

For a given $\lambda \in \mathbb{R}$, find a portfolio w that maximizes

$$\lambda \psi^\top w - \frac{1}{2} w^\top C w$$

Connection to Modern Portfolio Theory

$$
\begin{aligned}
w &\iff \mu_w \\
\psi^{\top} w &\iff \int \psi \, d\mu_w \\
w^{\top} C w &\iff ||\hat{\mu_w}||^2
\end{aligned}
$$

Modern Portfolio Theory

For a given $\lambda \in \mathbb{R}$, find a portfolio w that maximizes

$$\lambda \psi^{\top} w - \frac{1}{2} w^{\top} C w$$

What Our Abstract Stuff Got Us

For a given $\lambda \in \mathbb{R}$, find a probability measure μ_w that maximizes

$$\int \lambda \psi \, d\mu_w - \frac{1}{2} ||\hat{\mu_w}||^2$$

Topiarism

Reminder of Our Goals

- ♥ Optimizing should be done over positive portfolios.
- ♧ Approximation over subsets of stocks should get close to the ”true” maximum.
- ♤ Finding/estimating the best portfolios should be efficient.
- ♦ The model should match the trend of the market while using less stocks.
- $ The model should preserve some of the efficient frontier’s structure.

What is Topiarism?

Topiarism is just the application of measure theory and reproducing kernel Hilbert spaces to modern portfolio theory.

What is Topiarism?

Topiarism is just the application of measure theory and reproducing kernel Hilbert spaces to modern portfolio theory.

What is Topiarism?

Topiarism is just the application of measure theory and reproducing kernel Hilbert spaces to modern portfolio theory.

We will assume Ω is a sufficiently "nice" space, and $\mathcal{H}$ is some RKHS on this space with a continuous kernel function K.

The Objective Function

The Objective Function

Definition

Let ψ be a continuous function on Ω (not necessarily in $\mathcal{H}$). Define $\eta : \mathcal{M}_+^1(\Omega) \longrightarrow \mathbb{R}$ as

$$\eta(\mu) = \int \psi \, d\mu - \frac{1}{2} ||\hat{\mu}||^2$$

where $\hat{\mu}$ is the kernel embedding of μ. We call ψ the ***payout function*** and η the ***objective function***.

The Objective Function

Definition

Let ψ be a continuous function on Ω (not necessarily in $\mathcal{H}$). Define $\eta : \mathcal{M}_+^1(\Omega) \to \mathbb{R}$ as

$$\eta(\mu) = \int \psi \, d\mu - \frac{1}{2} ||\hat{\mu}||^2$$

where $\hat{\mu}$ is the kernel embedding of μ. We call ψ the ***payout function*** and η the ***objective function***.

Our goal will be to, find a measure $\mu \in \mathcal{M}_+^1(\Omega)$ that maximizes the objective function.

The Objective Function

Definition

Let ψ be a continuous function on Ω (not necessarily in $\mathcal{H}$). Define $\eta : \mathcal{M}_+^1(\Omega) \to \mathbb{R}$ as

$$\eta(\mu) = \int \psi \, d\mu - \frac{1}{2} ||\hat{\mu}||^2$$

where $\hat{\mu}$ is the kernel embedding of μ. We call ψ the ***payout function*** and η the ***objective function***.

Our goal will be to, find a measure $\mu \in \mathcal{M}_+^1(\Omega)$ that maximizes the objective function.

A Few Relevant Definitions

Definition

A nonnegative probability measure $\mu \in \mathcal{M}_+^1(\Omega)$ is called ***topiaric*** (with respect to Ω) if

$$\sup_{\nu \in \mathcal{M}_+^1(\Omega)} \eta(\nu) = \eta(\mu).$$

A Few Relevant Definitions

Definition

A nonnegative probability measure $\mu \in \mathcal{M}_+^1(\Omega)$ is called ***topiaric*** (with respect to Ω) if

$$\sup_{\nu \in \mathcal{M}_+^1(\Omega)} \eta(\nu) = \eta(\mu).$$

Definition

A sequence of nonnegative probability measures $\mu_n \subseteq \mathcal{M}_+^1(\Omega)$ is called ***topiaric*** (with respect to Ω) if

$$\sup_{\nu \in \mathcal{M}_+^1(\Omega)} \eta(\nu) = \lim_{n \to \infty} \eta(\mu_n).$$

A Few Relevant Definitions

Definition

A nonnegative probability measure $\mu \in \mathcal{M}_+^1(\Omega)$ is called ***topiaric*** (with respect to Ω) if

$$\sup_{\nu \in \mathcal{M}_+^1(\Omega)} \eta(\nu) = \eta(\mu).$$

Definition

A sequence of nonnegative probability measures $\mu_n \subseteq \mathcal{M}_+^1(\Omega)$ is called ***topiaric*** (with respect to Ω) if

$$\sup_{\nu \in \mathcal{M}_+^1(\Omega)} \eta(\nu) = \lim_{n \to \infty} \eta(\mu_n).$$

Theorem

There exists a topiaric measure.

A Few Relevant Definitions

Definition

A nonnegative probability measure $\mu \in \mathcal{M}_+^1(\Omega)$ is called ***topiaric*** (with respect to Ω) if

$$\sup_{\nu \in \mathcal{M}_+^1(\Omega)} \eta(\nu) = \eta(\mu).$$

Definition

A sequence of nonnegative probability measures $\mu_n \subseteq \mathcal{M}_+^1(\Omega)$ is called ***topiaric*** (with respect to Ω) if

$$\sup_{\nu \in \mathcal{M}_+^1(\Omega)} \eta(\nu) = \lim_{n\to\infty} \eta(\mu_n).$$

Theorem

There exists a topiaric measure.

Proof: Recall that $\mathcal{M}_+^1(\Omega)$ is a compact set. Additionally, it can be shown that η is continuous. Thus η exhibits a maximizer! $\square$

The Invisible Index Theorem

Theorem (The Invisible Index Theorem)

Suppose $\Omega \subseteq \mathcal{W}$. Let ν be topiaric with respect to $\mathcal{W}$. Then any topiaric measure with respect to Ω minimizes

$$||\hat{\nu} - \hat{\mu}||$$

over Ω.

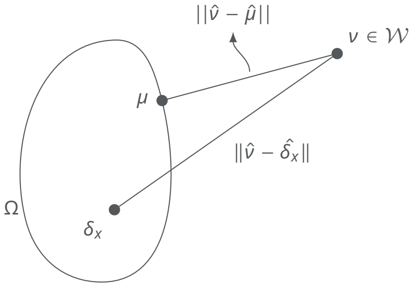
The Invisible Index Theorem

Theorem (The Invisible Index Theorem)

Suppose $\Omega \subseteq \mathcal{W}$. Let ν be topiaric with respect to $\mathcal{W}$. Then any topiaric measure with respect to Ω minimizes

$$||\hat{\nu} - \hat{\mu}||$$

over Ω.



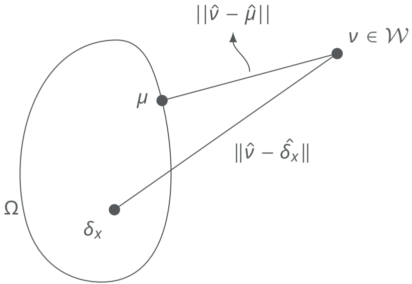
The Invisible Index Theorem

Theorem (The Invisible Index Theorem)

Suppose $\Omega \subseteq \mathcal{W}$. Let ν be topiaric with respect to $\mathcal{W}$. Then any topiaric measure with respect to Ω minimizes

$$||\hat{\nu} - \hat{\mu}|| \longleftarrow \clubsuit$$

over Ω.



Neat Theorems We Get

Theorem

Let $\mu, \nu \in \mathcal{M}_+^1(\Omega)$ both be topiaric measures. Then $\hat{\mu} = \hat{\nu}$ as embedded measures in $\mathcal{H}$.

Neat Theorems We Get

Theorem

Let $\mu, \nu \in \mathcal{M}_+^1(\Omega)$ both be topiaric measures. Then $\hat{\mu} = \hat{\nu}$ as embedded measures in $\mathcal{H}$.

Theorem

*Let μ be topiaric and let μ_n be a topiaric sequence. Then $\hat{\mu_n} \longrightarrow \hat{\mu}$ in $\mathcal{H}$ and $\mu_n \longrightarrow \mu$ weak*ly as measures.*

Neat Theorems We Get

Theorem

Let $\mu, \nu \in \mathcal{M}_+^1(\Omega)$ both be topiaric measures. Then $\hat{\mu} = \hat{\nu}$ as embedded measures in $\mathcal{H}$.

Theorem

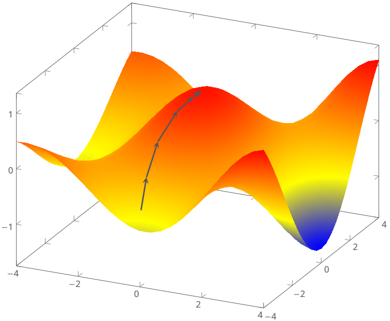
*Let μ be topiaric and let μ_n be a topiaric sequence. Then $\hat{\mu_n} \longrightarrow \hat{\mu}$ in $\mathcal{H}$ and $\mu_n \longrightarrow \mu$ weak*ly as measures.*

Theorem

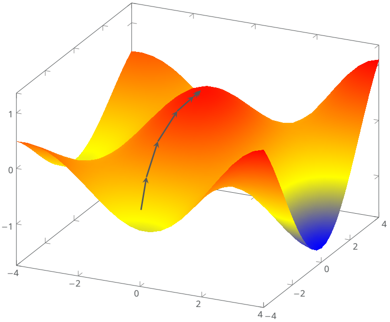
Let $\mu \in \mathcal{M}_+^1(\Omega)$. Then μ is topiaric if and only if $D_\eta(\mu, \delta_x - \mu) \leq 0$ for all $x \in \mathcal{K}$.

Approximating Using Gradient Ascent

Approximating Using Gradient Ascent

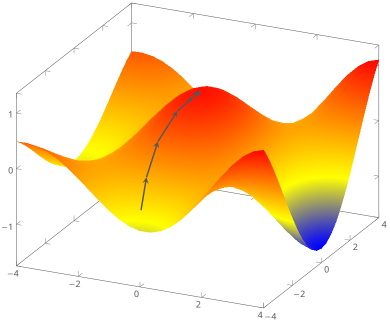


Approximating Using Gradient Ascent



Recall that that η has no local maximums, only global ones. This makes it perfect for gradient ascent!

Approximating Using Gradient Ascent



Recall that that η has no local maximums, only global ones. This makes it perfect for gradient ascent!

Theorem

The sequence of measures obtained by performing gradient ascent on the objective function is a topiaric sequence.

Testing our Algorithm

		Time Elapsed (s)		η		**# of assets used**	
# of available assets	λ	Topiary	SciPy	Topiary	SciPy	Topiary	SciPy
10	1	$5.740 \cdot 10^{-5}$	$1.388 \cdot 10^{-2}$	$5.021 \cdot 10^{-4}$	$5.021 \cdot 10^{-4}$	1	1
	.1	$9.150 \cdot 10^{-3}$	$1.607 \cdot 10^{-2}$	$2.969 \cdot 10^{-5}$	$2.968 \cdot 10^{-5}$	1	1
	.03	$6.492 \cdot 10^{-3}$	$1.774 \cdot 10^{-2}$	$-1.764 \cdot 10^{-6}$	$-1.764 \cdot 10^{-6}$	4	4
	0	$1.408 \cdot 10^{-2}$	$1.849 \cdot 10^{-2}$	$-9.936 \cdot 10^{-6}$	$-9.936 \cdot 10^{-6}$	5	5
100	1	$1.021 \cdot 10^{-3}$	.5298	$1.034 \cdot 10^{-3}$	$1.034 \cdot 10^{-3}$	1	1
	.1	$2.681 \cdot 10^{-2}$	.9956	$4.3973 \cdot 10^{-5}$	$3.870 \cdot 10^{-5}$	5	19
	.03	$9.981 \cdot 10^{-2}$	.9639	$5.280 \cdot 10^{-6}$	$1.477 \cdot 10^{-6}$	6	41
	0	$8.213 \cdot 10^{-2}$	.9534	$-3.638 \cdot 10^{-6}$	$-6.013 \cdot 10^{-6}$	7	45
488	1	$2.807 \cdot 10^{-4}$	33.82	$1.968 \cdot 10^{-3}$	$1.968 \cdot 10^{-3}$	3	3
	.1	$2.262 \cdot 10^{-2}$	44.34	$1.443 \cdot 10^{-4}$	$1.441 \cdot 10^{-4}$	4	4
	.03	$2.739 \cdot 10^{-2}$	46.87	$1.990 \cdot 10^{-5}$	$1.803 \cdot 10^{-5}$	11	55
	0	.1360	42.44	$-3.222 \cdot 10^{-6}$	$-5.029 \cdot 10^{-6}$	12	133

Testing our Algorithm

		Time Elapsed (s)		η		# of assets used	
# of available assets	λ	Topiary	SciPy	Topiary	SciPy	Topiary	SciPy
10	1	$5.740 \cdot 10^{-5}$	$1.388 \cdot 10^{-2}$	$5.021 \cdot 10^{-4}$	$5.021 \cdot 10^{-4}$	1	1
	.1	$9.150 \cdot 10^{-3}$	$1.607 \cdot 10^{-2}$	$2.969 \cdot 10^{-5}$	$2.968 \cdot 10^{-5}$	1	1
	.03	$6.492 \cdot 10^{-3}$	$1.774 \cdot 10^{-2}$	$-1.764 \cdot 10^{-6}$	$-1.764 \cdot 10^{-6}$	4	4
	0	$1.408 \cdot 10^{-2}$	$1.849 \cdot 10^{-2}$	$-9.936 \cdot 10^{-6}$	$-9.936 \cdot 10^{-6}$	5	5
100	1	$1.021 \cdot 10^{-3}$	.5298	$1.034 \cdot 10^{-3}$	$1.034 \cdot 10^{-3}$	1	1
	.1	$2.681 \cdot 10^{-2}$	.9956	$4.3973 \cdot 10^{-5}$	$3.870 \cdot 10^{-5}$	5	19
	.03	$9.981 \cdot 10^{-2}$	.9639	$5.280 \cdot 10^{-6}$	$1.477 \cdot 10^{-6}$	6	41
	0	$8.213 \cdot 10^{-2}$	.9534	$-3.638 \cdot 10^{-6}$	$-6.013 \cdot 10^{-6}$	7	45
488	1	$2.807 \cdot 10^{-4}$	33.82	$1.968 \cdot 10^{-3}$	$1.968 \cdot 10^{-3}$	3	3
	.1	$2.262 \cdot 10^{-2}$	44.34	$1.443 \cdot 10^{-4}$	$1.441 \cdot 10^{-4}$	4	4
	.03	$2.739 \cdot 10^{-2}$	46.87	$1.990 \cdot 10^{-5}$	$1.803 \cdot 10^{-5}$	11	55
	0	.1360	42.44	$-3.222 \cdot 10^{-6}$	$-5.029 \cdot 10^{-6}$	12	133

Testing our Algorithm

		Time Elapsed (s)		η		# of assets used	
# of available assets	λ	Topiary	SciPy	Topiary	SciPy	Topiary	SciPy
10	1	$5.740 \cdot 10^{-5}$	$1.388 \cdot 10^{-2}$	$5.021 \cdot 10^{-4}$	$5.021 \cdot 10^{-4}$	1	1
	.1	$9.150 \cdot 10^{-3}$	$1.607 \cdot 10^{-2}$	$2.969 \cdot 10^{-5}$	$2.968 \cdot 10^{-5}$	1	1
	.03	$6.492 \cdot 10^{-3}$	$1.774 \cdot 10^{-2}$	$-1.764 \cdot 10^{-6}$	$-1.764 \cdot 10^{-6}$	4	4
	0	$1.408 \cdot 10^{-2}$	$1.849 \cdot 10^{-2}$	$-9.936 \cdot 10^{-6}$	$-9.936 \cdot 10^{-6}$	5	5
100	1	$1.021 \cdot 10^{-3}$	.5298	$1.034 \cdot 10^{-3}$	$1.034 \cdot 10^{-3}$	1	1
	.1	$2.681 \cdot 10^{-2}$	.9956	$4.3973 \cdot 10^{-5}$	$3.870 \cdot 10^{-5}$	5	19
	.03	$9.981 \cdot 10^{-2}$	.9639	$5.280 \cdot 10^{-6}$	$1.477 \cdot 10^{-6}$	6	41
	0	$8.213 \cdot 10^{-2}$	.9534	$-3.638 \cdot 10^{-6}$	$-6.013 \cdot 10^{-6}$	7	45
488	1	$2.807 \cdot 10^{-4}$	33.82	$1.968 \cdot 10^{-3}$	$1.968 \cdot 10^{-3}$	3	3
	.1	$2.262 \cdot 10^{-2}$	44.34	$1.443 \cdot 10^{-4}$	$1.441 \cdot 10^{-4}$	4	4
	.03	$2.739 \cdot 10^{-2}$	46.87	$1.990 \cdot 10^{-5}$	$1.803 \cdot 10^{-5}$	11	55
	0	.1360	42.44	$-3.222 \cdot 10^{-6}$	$-5.029 \cdot 10^{-6}$	12	133

Check Back on our Objectives

- ♥ Optimizing should be done over positive portfolios.
- ♧ Approximation over subsets of stocks should get close to the ”true” maximum.
- ♤ Finding/estimating the best portfolios should be efficient.
- ♦ The model should match the trend of the market while using less stocks.
- $ The model should preserve some of the efficient frontier’s structure.

Check Back on our Objectives

- ✓ ♥ Optimizing should be done over positive portfolios.
- ♧ Approximation over subsets of stocks should get close to the ”true” maximum.
- ♤ Finding/estimating the best portfolios should be efficient.
- ♦ The model should match the trend of the market while using less stocks.
- $ The model should preserve some of the efficient frontier’s structure.

Check Back on our Objectives

- ✓ ♥ Optimizing should be done over positive portfolios.
- ✓ ♣ Approximation over subsets of stocks should get close to the ”true” maximum.
- ♠ Finding/estimating the best portfolios should be efficient.
- ♦ The model should match the trend of the market while using less stocks.
- $ The model should preserve some of the efficient frontier’s structure.

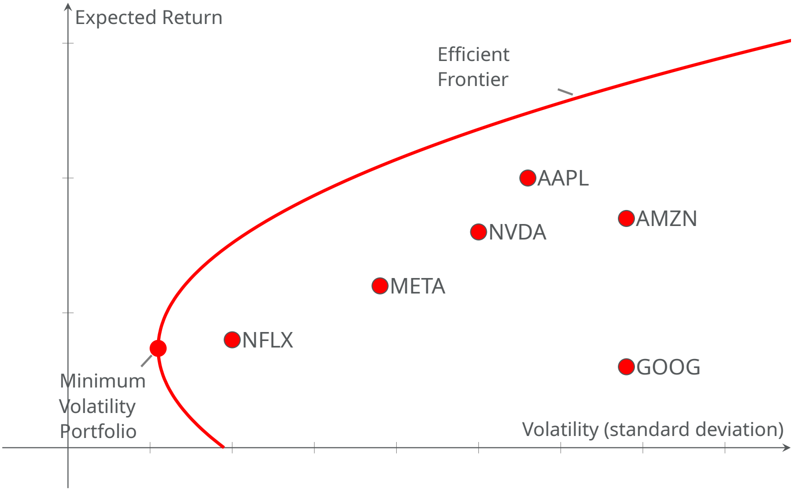
Check Back on our Objectives

- ✓ ♥ Optimizing should be done over positive portfolios.
- ✓ ♣ Approximation over subsets of stocks should get close to the ”true” maximum.
- ✓ ♠ Finding/estimating the best portfolios should be efficient.
- ♦ The model should match the trend of the market while using less stocks.
- $ The model should preserve some of the efficient frontier’s structure.

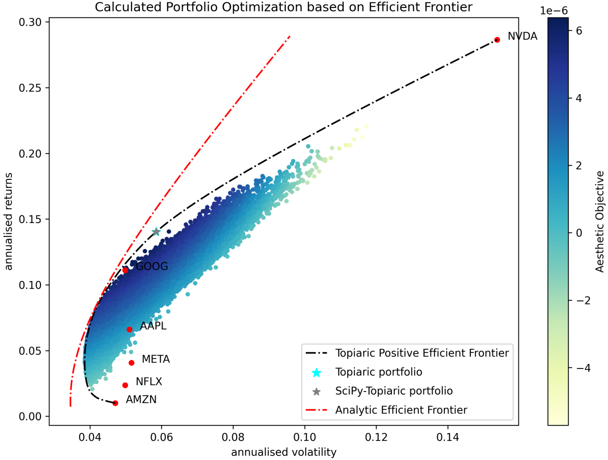
Check Back on our Objectives

- ✓ ♥ Optimizing should be done over positive portfolios.
- ✓ ♧ Approximation over subsets of stocks should get close to the ”true” maximum.
- ✓ ♤ Finding/estimating the best portfolios should be efficient.
- ✓ ♦ The model should match the trend of the market while using less stocks.
- \$ The model should preserve some of the efficient frontier’s structure.

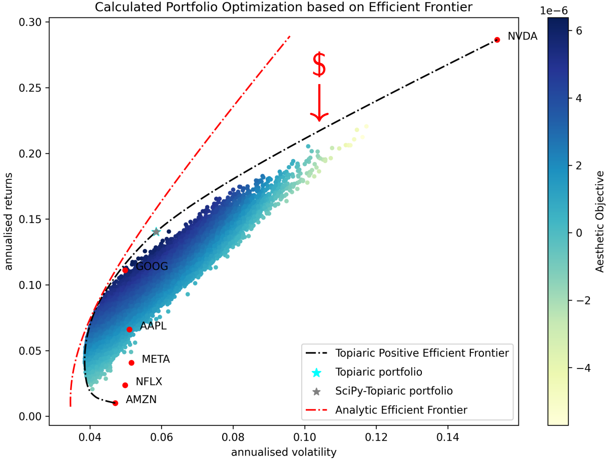
Testing the Efficient Frontier



Testing the Efficient Frontier



Testing the Efficient Frontier



Questions?

Appendix

Problems with Modern Portfolio Theory

Problems with Modern Portfolio Theory

- The model is entirely based upon historical data, which is well known to not be a good indicator of future asset performance.
- Using variance as a stand-in for risk does not differentiate between upside and downside risk.
- The model allows for short selling assets, but does not treat positive weights different than negative weights. In reality, short selling is inherently more risky because there is no cap to losses, whereas when buying longs, one can only lose as much as they invest.
- The market portfolio must be a portfolio that has investment in every asset possible, and it has been proved to be unobservable. This implies the model is unverifiable, as any empirical tests of the model can only use substitutes for the market portfolio. This is known as **Roll's critique**.
- The model also relies on the **efficient market hypothesis**, which states that asset prices reflect all available information instantaneously and that all investors act completely rationally.
- Many investors believe that it is near impossible to beat an index in the long run since if one was able to, there would exist an arbitrage opportunity. Such investors go back to methods used before the invention of modern portfolio theory, in particular focusing on dedicated research into a smaller subset of "growth stocks", i.e. assets that seem to be "on the rise".
- The model assumes that potential returns follow a normal distribution, when in

Duality

Definition

Let V be a vector space. Then the ***dual space*** of V, denoted V^*, is

$$V^* = \{L\colon V \longrightarrow \mathbb{F} : L \text{ is a continuous linear functional}\}.$$

Duality

Definition

Let V be a vector space. Then the ***dual space*** of V, denoted V^*, is

$$V^* = \{L \colon V \longrightarrow \mathbb{F} : L \text{ is a continuous linear functional}\}.$$

- If V is n dimensional, then row vectors are elements of the dual space.
- If V is a vector space of continuous functions on some set Ω, then for each $\mu \in \mathcal{M}(\Omega)$, the linear functional defined by $U(f) = \int f \, d\mu$ is an element of V^*.

Duality

Definition

Let V be a vector space. Then the ***dual space*** of V, denoted V^*, is

$$V^* = \{L\colon V \longrightarrow \mathbb{F} : L \text{ is a continuous linear functional}\}.$$

- If V is n dimensional, then row vectors are elements of the dual space.
- If V is a vector space of continuous functions on some set Ω, then for each $\mu \in \mathcal{M}(\Omega)$, the linear functional defined by $U(f) = \int f\,d\mu$ is an element of V^*.

$$U(f+g) = \int f+g\,d\mu = \int f\,d\mu + \int g\,d\mu,$$
$$U(\alpha f) = \int \alpha f\,d\mu = \alpha \int f\,d\mu$$

Duality

Definition

Let V be a vector space. Then the ***dual space*** of V, denoted V^*, is

$$V^* = \{L\colon V \longrightarrow \mathbb{F} : L \text{ is a continuous linear functional}\}.$$

- If V is n dimensional, then row vectors are elements of the dual space.
- If V is a vector space of continuous functions on some set Ω, then for each $\mu \in \mathcal{M}(\Omega)$, the linear functional defined by $U(f) = \int f\,d\mu$ is an element of V^*.

$$U(f+g) = \int f+g\,d\mu = \int f\,d\mu + \int g\,d\mu,$$

$$U(\alpha f) = \int \alpha f\,d\mu = \alpha \int f\,d\mu$$

This allows us to think of measures as a "subspace" of V^*.

Double Dual

If V is a vector space, then so is V^*. So we can define linear operators on V^*. In particular, define $\Psi_x : V^* \longrightarrow \mathbb{F}$ by

$$\Psi_x(L) = L(x).$$

Double Dual

If V is a vector space, then so is V^*. So we can define linear operators on V^*. In particular, define $\Psi_x : V^* \longrightarrow \mathbb{F}$ by

$$\Psi_x(L) = L(x).$$

This is a linear operator on V^*, as

$$\Psi_x(L+T) = L(x) + T(x) \qquad \text{and} \qquad \Psi_x(\alpha L) = \alpha L(x).$$

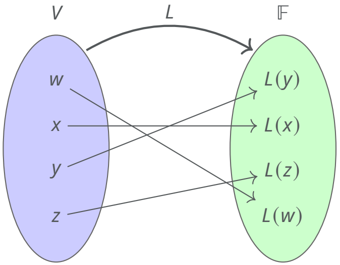
Double Dual

If V is a vector space, then so is V^*. So we can define linear operators on V^*. In particular, define $\Psi_x : V^* \longrightarrow \mathbb{F}$ by

$$\Psi_x(L) = L(x).$$

This is a linear operator on V^*, as

$$\Psi_x(L+T) = L(x) + T(x) \qquad \text{and} \qquad \Psi_x(\alpha L) = \alpha L(x).$$



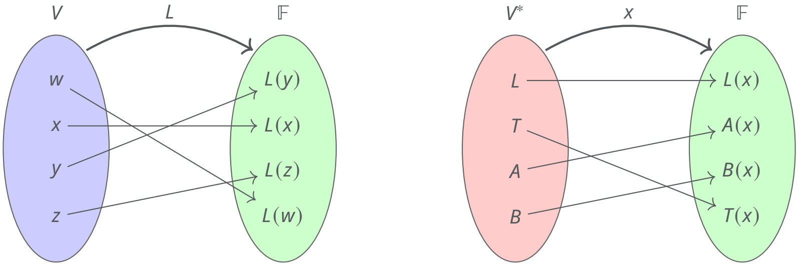
Double Dual

If V is a vector space, then so is V^*. So we can define linear operators on V^*. In particular, define $\Psi_x : V^* \longrightarrow \mathbb{F}$ by

$$\Psi_x(L) = L(x).$$

This is a linear operator on V^*, as

$$\Psi_x(L+T) = L(x) + T(x) \qquad \text{and} \qquad \Psi_x(\alpha L) = \alpha L(x).$$



Weak* Topology

Weak* Topology

Definition

Let L_n be a sequence of linear operators in V^*. Then we say that $L_n \longrightarrow L$ ***weak*-ly*** if

$$\Psi_x(L_n) \longrightarrow \Psi_x(L)$$

for all $x \in V$.

Weak* Topology

Definition

Let L_n be a sequence of linear operators in V^*. Then we say that $L_n \longrightarrow L$ ***weak*-ly*** if

$$\Psi_x(L_n) \longrightarrow \Psi_x(L)$$

for all $x \in V$.

Recall that we can think of $\mathcal{M}(\Omega)$ as a subspace of V^*, so to put structure on it, we steal this notion of convergence from that of the weak* topology!

Weak* Topology

Definition

Let L_n be a sequence of linear operators in V^*. Then we say that $L_n \longrightarrow L$ ***weak*-ly*** if

$$\Psi_x(L_n) \longrightarrow \Psi_x(L)$$

for all $x \in V$.

Recall that we can think of $\mathcal{M}(\Omega)$ as a subspace of V^*, so to put structure on it, we steal this notion of convergence from that of the weak* topology!

Definition

Let V be a vector space a functions on a set Ω, and let $\mathcal{M}(\Omega)$ be the vector space of measures on Ω. We say a sequence of measures $\mu_n \subseteq \mathcal{M}(\Omega)$ approaches a measure μ ***weak*-ly*** or ***vaguely*** if

$$\int f \, d\mu_n \rightarrow \int f \, d\mu$$

for all $f \in V$.

Weak* Topology

Definition

Let L_n be a sequence of linear operators in V^*. Then we say that $L_n \to L$ ***weak*-ly*** if

$$\Psi_x(L_n) \to \Psi_x(L)$$

for all $x \in V$.

Recall that we can think of $\mathcal{M}(\Omega)$ as a subspace of V^*, so to put structure on it, we steal this notion of convergence from that of the weak* topology!

Definition

Let V be a vector space a functions on a set Ω, and let $\mathcal{M}(\Omega)$ be the vector space of measures on Ω. We say a sequence of measures $\mu_n \subseteq \mathcal{M}(\Omega)$ approaches a measure μ ***weak*-ly*** or ***vaguely*** if

$$\int f d\mu_n \to \int f d\mu$$

for all $f \in V$.

- Under this topology, the space of nonnegative probability measures $\mathcal{M}_+^1(\Omega)$ is compact

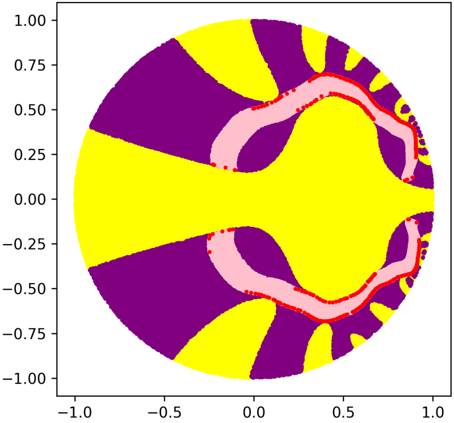
Further Applications of Topiarism

Further Applications of Topiarism

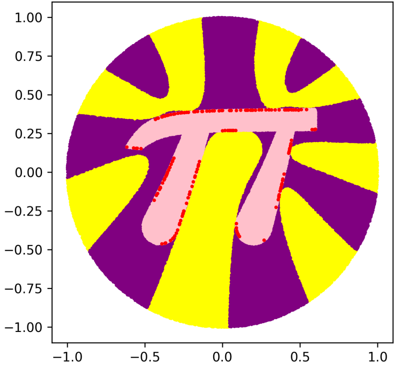
$$h^2_{\mathbb{R}}(\mathbb{D}) = \left\{ f \in h^2(\mathbb{D}) : f(re^{i\theta}) = \sum_{n=-\infty}^{\infty} c_n r^{|n|} e^{in\theta},\ \sup_{0 \le r < 1} \left(\int_{\partial \mathbb{D}} |f(rw)|\ dw \right)^{\frac{1}{2}} < \infty,\ c_n = \overline{c_{-n}} \right\},$$

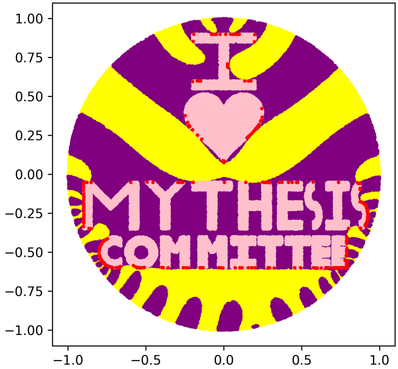
also known as the ***real harmonic Hardy space*** on the complex unit disk.

Example of Topiarism on the Harmonic Hardy Space

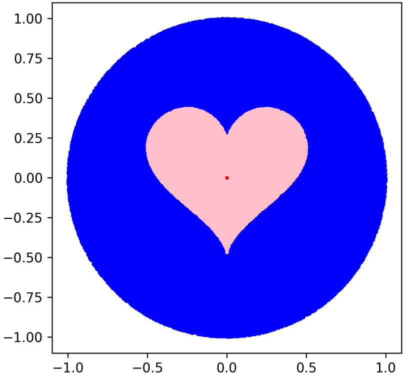


More Examples of Topiarism on the Harminc Hardy Space





Weird Edge Case for Topiary Algorithm on Harmonic Hardy Space



Questions?