CANONICAL PORTFOLIOS: OPTIMAL ASSET AND SIGNAL COMBINATION

Nick Firoozye[1], **Vincent Tan**[2], Stefan Zohren[2]

October 2021

[1]University College London
[2]University of Oxford and Oxford-Man Institute

Journal of Finance (1952)

PORTFOLIO SELECTION*

HARRY MARKOWITZ
The Rand Corporation

THE PROCESS OF SELECTING a portfolio may be divided into two stages. The first stage starts with observation and experience and ends with beliefs about the future performances of available securities. The second stage starts with the relevant beliefs about future performances and ends with the choice of portfolio. This paper is concerned with the second stage. We first consider the rule that the investor does (or should) maximize discounted expected, or anticipated, returns. This rule is rejected both as a hypothesis to explain, and as a maximum to guide investment behavior. We next consider the rule that the investor does (or should) consider expected return a desirable thing *and* variance of return an undesirable thing. This rule has many sound points, both as a maxim for, and hypothesis about, investment behavior. We illustrate geometrically relations between beliefs and choice of portfolio according to the "expected returns—variance of returns" rule.

Key Deliverables

- Provide a novel framework that generalizes Markowitz (1952) to include cross-sectional information from (1) the signals and (2) their cross-relationships with the returns
- Simplify a large-dimensional problem using a key tool from statistics and machine learning: Canonical Correlation Analysis of Hotelling (1936)
- Operationalize our framework to bridge the gap between theory and practice using covariance estimation techniques

Outline

Problem Setting

Environment

Notation:

- $N \times 1$ vector of returns $\mathbf{r}_t := (r_{t,1}, \ldots, r_{t,N})$
- $N \times 1$ vector of signals $\mathbf{x}_t := (x_{t,1}, \ldots, x_{t,N})$
  - Proxy for conditional expected returns
- Assumption on the returns and signals: jointly Gaussisan with zero mean
- The composite vector $(\mathbf{r}_{t+1}, \mathbf{x}_t)$ has a joint covariance matrix

$$\boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\Sigma}_r & \boldsymbol{\Sigma}_{rx} \\ \boldsymbol{\Sigma}'_{rx} & \boldsymbol{\Sigma}_x \end{pmatrix}$$

- Covariance of returns $\mathrm{Var}(\mathbf{r}_{t+1}) := \mathrm{Cov}(\mathbf{r}_{t+1}, \mathbf{r}_{t+1}) = \boldsymbol{\Sigma}_r$
- Covariance of signals $\mathrm{Var}(\mathbf{x}_t) := \mathrm{Cov}(\mathbf{x}_t, \mathbf{x}_t) = \boldsymbol{\Sigma}_x$
- Cross-covariance of returns and signals $\mathrm{Cov}(\mathbf{r}_{t+1}, \mathbf{x}_t) = \boldsymbol{\Sigma}_{rx}$
- Population moments - Greek letters; Sample-based moments - Latin letters
- Estimated moments have a hat ($\hat{}$) accent

Dynamic Strategies

- Linear portfolio policies, à la Brandt and Santa Clara (2006):

$$\mathbf{w}_t := \mathbf{A}'\mathbf{x}_t \text{ where } \mathbf{A} \text{ is a } N \times N \text{ matrix of coefficients}$$

Dynamic Strategies

- Linear portfolio policies, à la Brandt and Santa Clara (2006):

$$\mathbf{w}_t := \mathbf{A}'\mathbf{x}_t \text{ where } \mathbf{A} \text{ is a } N \times N \text{ matrix of coefficients}$$

- Strategy returns (or Profit-and-Loss) at time $t+1$:

$$\text{PnL}_{t+1} := \mathbf{x}_t'\mathbf{A}\mathbf{r}_{t+1}$$

Dynamic Strategies

- Linear portfolio policies, à la Brandt and Santa Clara (2006):

$$\mathbf{w}_t := \mathbf{A}'\mathbf{x}_t \ \text{ where } \mathbf{A} \text{ is a } N \times N \text{ matrix of coefficients}$$

- Strategy returns (or Profit-and-Loss) at time $t+1$:

$$\mathrm{PnL}_{t+1} := \mathbf{x}_t'\mathbf{A}\mathbf{r}_{t+1}$$

- Investor's conditional optimization problem:

$$\min_{\mathbf{A}} \frac{1}{2}\mathrm{Var}_t[\mathrm{PnL}_{t+1}]$$
$$\text{subject to } \mathbb{E}_t[\mathrm{PnL}_{t+1}] \geq \mathcal{G}$$

where $\mathcal{G}$ is a target return

Dynamic Strategies

- Linear portfolio policies, à la Brandt and Santa Clara (2006):

$$\mathbf{w}_t := \mathbf{A}'\mathbf{x}_t \text{ where } \mathbf{A} \text{ is a } N \times N \text{ matrix of coefficients}$$

- Strategy returns (or Profit-and-Loss) at time $t+1$:

$$\mathsf{PnL}_{t+1} := \mathbf{x}_t'\mathbf{A}\mathbf{r}_{t+1}$$

- Investor's conditional optimization problem:

$$\min_{\mathbf{A}} \frac{1}{2}\mathsf{Var}_t[\mathsf{PnL}_{t+1}]$$
$$\text{subject to } \mathbb{E}_t[\mathsf{PnL}_{t+1}] \geq \mathcal{G}$$

  where $\mathcal{G}$ is a target return

- Since $\mathbf{A}$ constant over time, the investor's conditional problem is equal to investor's unconditional problem

Dynamic Strategies

- Linear portfolio policies, à la Brandt and Santa Clara (2006):

$$\mathbf{w}_t := \mathbf{A}'\mathbf{x}_t \text{ where } \mathbf{A} \text{ is a } N \times N \text{ matrix of coefficients}$$

- Strategy returns (or Profit-and-Loss) at time $t+1$:

$$\text{PnL}_{t+1} := \mathbf{x}_t'\mathbf{A}\mathbf{r}_{t+1}$$

- Investor's conditional optimization problem:

$$\min_{\mathbf{A}} \frac{1}{2}\text{Var}_t[\text{PnL}_{t+1}]$$

$$\text{subject to } \mathbb{E}_t[\text{PnL}_{t+1}] \geq \mathcal{G}$$

where $\mathcal{G}$ is a target return

- Since $\mathbf{A}$ constant over time, the investor's conditional problem is equal to investor's unconditional problem
- Using Isserlis (1918) or Wick (1950) yields:

$$\min_{\mathbf{A}} \frac{1}{2}\text{Tr}(\boldsymbol{\Sigma}_x\mathbf{A}\boldsymbol{\Sigma}_r\mathbf{A}') + \cancel{\frac{1}{2}\text{Tr}(\boldsymbol{\Sigma}_{rx}\mathbf{A}\boldsymbol{\Sigma}_{rx}\mathbf{A})}$$

$$\text{subject to } \text{Tr}(\mathbf{A}\boldsymbol{\Sigma}_{rx}) \geq \mathcal{G}$$

Solution to Optimization Problem

Solution to optimization:

$$\mathbf{w}_t = \underbrace{\lambda \times \mathbf{\Sigma}_r^{-1}\mathbf{\Sigma}_{rx}\mathbf{\Sigma}_x^{-1}}_{\mathbf{A}'}\mathbf{x}_t, \;\; \text{where } \lambda := \frac{\mathcal{G}}{\mathrm{Tr}(\mathbf{\Sigma}_{rx}\mathbf{\Sigma}_x^{-1}\mathbf{\Sigma}_{rx}'\mathbf{\Sigma}_r^{-1})}$$

Solution to Optimization Problem

Solution to optimization:

$$\mathbf{w}_t = \underbrace{\lambda \times \boldsymbol{\Sigma}_r^{-1} \boldsymbol{\Sigma}_{rx} \boldsymbol{\Sigma}_x^{-1}}_{\mathbf{A}'} \mathbf{x}_t, \quad \text{where } \lambda := \frac{\mathcal{G}}{\mathrm{Tr}(\boldsymbol{\Sigma}_{rx} \boldsymbol{\Sigma}_x^{-1} \boldsymbol{\Sigma}_{rx}' \boldsymbol{\Sigma}_r^{-1})}$$

Observations:

- Implied two-stage process
  - Stage 1: Multivariate linear regression of the returns on signals
  - Stage 2: Asset allocation
- However, the two-stage idea is silent about the fact that the optimal asset and signal combination problem is a joint selection process

Solution to Optimization Problem

Solution to optimization:

$$\mathbf{w}_t = \underbrace{\lambda \times \boldsymbol{\Sigma}_r^{-1} \boldsymbol{\Sigma}_{rx} \boldsymbol{\Sigma}_x^{-1}}_{\mathbf{A}'} \mathbf{x}_t, \quad \text{where } \lambda := \frac{\mathcal{G}}{\mathrm{Tr}(\boldsymbol{\Sigma}_{rx} \boldsymbol{\Sigma}_x^{-1} \boldsymbol{\Sigma}_{rx}' \boldsymbol{\Sigma}_r^{-1})}$$

Observations:

- Implied two-stage process
    - Stage 1: Multivariate linear regression of the returns on signals
    - Stage 2: Asset allocation
- However, the two-stage idea is silent about the fact that the optimal asset and signal combination problem is a joint selection process

Note: Portfolio is not scale invariant in $\mathcal{G}$, $\boldsymbol{\Sigma}_{rx}$ and $\mathbf{x}_t$:

- Set $\mathcal{G} := \mathrm{Tr}(\boldsymbol{\Sigma}_{rx})/N$ to remove scale dependency

Portfolio Constraints

Commonplace in quantitative equity investing to impose portfolio constraints

- Fully-invested (sum-to-one) solution:

$$\mathbf{w}_t^{\mathrm{FI}} = (1-\lambda^{\mathrm{FI}})\frac{\boldsymbol{\Sigma}_r^{-1}\mathbf{1}}{\mathbf{1}'\boldsymbol{\Sigma}_r^{-1}\mathbf{1}} + \lambda^{\mathrm{FI}}\frac{\boldsymbol{\Sigma}_r^{-1}\boldsymbol{\Sigma}_{rx}\boldsymbol{\Sigma}_x^{-1}\mathbf{x}_t}{\mathbf{1}'\boldsymbol{\Sigma}_r^{-1}\boldsymbol{\Sigma}_{rx}\boldsymbol{\Sigma}_x^{-1}\mathbf{x}_t},$$

$$\text{where } \lambda^{\mathrm{FI}} := \frac{\mathcal{G}ab - b^2}{ac - b^2},$$

$$\text{with } a := (\mathbf{1}'\boldsymbol{\Sigma}_r^{-1}\mathbf{1})(\mathbf{x}_t'\boldsymbol{\Sigma}_x^{-1}\mathbf{x}_t),$$

$$b := \mathbf{1}'\boldsymbol{\Sigma}_r^{-1}\boldsymbol{\Sigma}_{rx}\boldsymbol{\Sigma}_x^{-1}\mathbf{x}_t, \text{ and}$$

$$c := \mathrm{Tr}(\boldsymbol{\Sigma}_x^{-1}\boldsymbol{\Sigma}_{rx}'\boldsymbol{\Sigma}_r^{-1}\boldsymbol{\Sigma}_{rx}).$$

- Dollar-neutral (sum-to-zero) solution:

$$\mathbf{w}_t^{\mathrm{ZI}} = \lambda^{\mathrm{ZI}}\left[\frac{\boldsymbol{\Sigma}_r^{-1}\boldsymbol{\Sigma}_{rx}\boldsymbol{\Sigma}_x^{-1}\mathbf{x}_t}{\mathbf{1}'\boldsymbol{\Sigma}_r^{-1}\boldsymbol{\Sigma}_{rx}\boldsymbol{\Sigma}_x^{-1}\mathbf{x}_t} - \frac{\boldsymbol{\Sigma}_r^{-1}\mathbf{1}}{\mathbf{1}'\boldsymbol{\Sigma}_r^{-1}\mathbf{1}}\right],$$

$$\text{where } \lambda^{\mathrm{ZI}} := \mathcal{G}\left[\frac{\mathrm{Tr}(\boldsymbol{\Sigma}_{rx}\boldsymbol{\Sigma}_x^{-1}\boldsymbol{\Sigma}_{rx}'\boldsymbol{\Sigma}_r^{-1})}{\mathbf{1}'\boldsymbol{\Sigma}_r^{-1}\boldsymbol{\Sigma}_{rx}\boldsymbol{\Sigma}_x^{-1}\mathbf{x}_t} - \frac{\mathbf{1}'\boldsymbol{\Sigma}_r^{-1}\boldsymbol{\Sigma}_{rx}\boldsymbol{\Sigma}_x^{-1}\mathbf{x}_t}{(\mathbf{1}'\boldsymbol{\Sigma}_r^{-1}\mathbf{1})(\mathbf{x}_t'\boldsymbol{\Sigma}_x^{-1}\mathbf{x}_t)}\right]^{-1}$$

Relation to Existing Literature

Markowitz (1952, JF)

- Solution assumes signals are the conditional expected returns
  $\implies \boldsymbol{\Sigma}_{rx} = \boldsymbol{\Sigma}_x$

Brandt and Santa-Clara (2006, JF)

- Solution: $\lambda \times \mathbb{E}[(\mathbf{x}_t \mathbf{x}_t') \otimes (\mathbf{r}_{t+1}\mathbf{r}_{t+1}')]^{-1}\mathbb{E}[\mathbf{x}_t \otimes \mathbf{r}_{t+1}]$
- Requires inverting an $N^2 \times N^2$ matrix $\rightarrow$ Less computationally tractable in large dimensions
- Less clarity on the structure of the optimal solution

Kelly, Malamud, and Pedersen (2022, JF forthcoming)

- Exploits cross-covariance matrix but ignores the cross-sectional information from the returns and signals
- Different objective function
- Empirical analysis is based on small universe and monthly rebalancing

Strategy Diversification

Canonical Correlation Analysis (CCA)

Goal of CCA:

- Dimension reduction on two different data sets while simultaneously retaining as much of the correlation present in both of them

Many different approaches to solve the CCA problem:

- Standard eigenvalue problem (Hotelling, 1936)
- Singular value decomposition (SVD) (Healy, 1957) $\leftarrow$ our focus!
- Generalized eigenvalue problem (Bach and Jordan, 2002)

Performing CCA

Step 1: Form cross-correlation matrix

- Build synthetic returns $\tilde{\mathbf{r}} := \mathbf{\Sigma}_r^{-1/2}\mathbf{r}$ and synthetic signals $\tilde{\mathbf{x}} := \mathbf{\Sigma}_x^{-1/2}\mathbf{x}$
- Build matrix $\mathbf{\Sigma}_{\tilde{r}\tilde{x}} := \mathbb{E}[\tilde{\mathbf{r}}\tilde{\mathbf{x}}'] = \mathbf{\Sigma}_r^{-1/2}\mathbf{\Sigma}_{rx}\mathbf{\Sigma}_x^{-1/2}$

Performing CCA

Step 1: Form cross-correlation matrix

- Build synthetic returns $\tilde{\mathbf{r}} := \mathbf{\Sigma}_r^{-1/2}\mathbf{r}$ and synthetic signals $\tilde{\mathbf{x}} := \mathbf{\Sigma}_x^{-1/2}\mathbf{x}$
- Build matrix $\mathbf{\Sigma}_{\tilde{r}\tilde{x}} := \mathbb{E}[\tilde{\mathbf{r}}\tilde{\mathbf{x}}'] = \mathbf{\Sigma}_r^{-1/2}\mathbf{\Sigma}_{rx}\mathbf{\Sigma}_x^{-1/2}$

Step 2: Perform SVD on $\mathbf{\Sigma}_{\tilde{r}\tilde{x}}$

- Singular values and singular vectors ordered from smallest to largest $((s_1, \ldots, s_N); (\mathbf{u}_1, \ldots, \mathbf{u}_N); (\mathbf{v}_1, \ldots, \mathbf{v}_N))$

Performing CCA

Step 1: Form cross-correlation matrix

- Build synthetic returns $\tilde{\mathbf{r}} := \boldsymbol{\Sigma}_r^{-1/2}\mathbf{r}$ and synthetic signals $\tilde{\mathbf{x}} := \boldsymbol{\Sigma}_x^{-1/2}\mathbf{x}$
- Build matrix $\boldsymbol{\Sigma}_{\tilde{r}\tilde{x}} := \mathbb{E}[\tilde{\mathbf{r}}\tilde{\mathbf{x}}'] = \boldsymbol{\Sigma}_r^{-1/2}\boldsymbol{\Sigma}_{rx}\boldsymbol{\Sigma}_x^{-1/2}$

Step 2: Perform SVD on $\boldsymbol{\Sigma}_{\tilde{r}\tilde{x}}$

- Singular values and singular vectors ordered from smallest to largest $((s_1, \ldots, s_N); (\mathbf{u}_1, \ldots, \mathbf{u}_N); (\mathbf{v}_1, \ldots, \mathbf{v}_N))$

Terminologies:

- Canonical directions: $((\boldsymbol{\Sigma}_r^{-1/2}\mathbf{u}_i, \boldsymbol{\Sigma}_x^{-1/2}\mathbf{v}_i))_{i=1}^N$
- Canonical variates: $((\mathbf{u}_i'\boldsymbol{\Sigma}_r^{-1/2}\mathbf{r}, \mathbf{v}_i'\boldsymbol{\Sigma}_x^{-1/2}\mathbf{x}))_{i=1}^N$
- Canonical correlations: $(s_i)_{i=1}^N$

Performing CCA

Step 1: Form cross-correlation matrix

- Build synthetic returns $\tilde{\mathbf{r}} := \boldsymbol{\Sigma}_r^{-1/2}\mathbf{r}$ and synthetic signals $\tilde{\mathbf{x}} := \boldsymbol{\Sigma}_x^{-1/2}\mathbf{x}$
- Build matrix $\boldsymbol{\Sigma}_{\tilde{r}\tilde{x}} := \mathbb{E}[\tilde{\mathbf{r}}\tilde{\mathbf{x}}'] = \boldsymbol{\Sigma}_r^{-1/2}\boldsymbol{\Sigma}_{rx}\boldsymbol{\Sigma}_x^{-1/2}$

Step 2: Perform SVD on $\boldsymbol{\Sigma}_{\tilde{r}\tilde{x}}$

- Singular values and singular vectors ordered from smallest to largest $((s_1, \ldots, s_N); (\mathbf{u}_1, \ldots, \mathbf{u}_N); (\mathbf{v}_1, \ldots, \mathbf{v}_N))$

Terminologies:

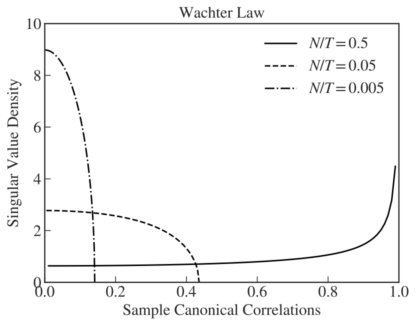
- Canonical directions: $((\boldsymbol{\Sigma}_r^{-1/2}\mathbf{u}_i, \boldsymbol{\Sigma}_x^{-1/2}\mathbf{v}_i))_{i=1}^N$
- Canonical variates: $((\mathbf{u}_i'\boldsymbol{\Sigma}_r^{-1/2}\mathbf{r}, \mathbf{v}_i'\boldsymbol{\Sigma}_x^{-1/2}\mathbf{x}))_{i=1}^N$
- Canonical correlations: $(s_i)_{i=1}^N$

In practice: Replace population moments with their sample counterparts

Spectrum of Sample Cross-Correlation Matrix

A restricted Wachter (1980) setting:

- Two independent N-dimensional variable sets $\mathbf{x}$ and $\mathbf{y}$
- Each set consists of i.i.d Gaussian random variables



Sample canonical correlations becomes more dispersed for larger values of N/T

Decomposition of Strategy Returns

Apply CCA to the strategy returns:

- Re-express strategy returns in terms of synthetic variables

$$\mathbf{x}_t'\mathbf{A}\mathbf{r}_{t+1} = \lambda \times \mathbf{x}_t'\mathbf{\Sigma}_x^{-1}\mathbf{\Sigma}_{rx}'\mathbf{\Sigma}_r^{-1}\mathbf{r}_{t+1} = \lambda \times \tilde{\mathbf{x}}_t'\mathbf{\Sigma}_{\tilde{r}\tilde{x}}'\tilde{\mathbf{r}}_{t+1} = \tilde{\mathbf{x}}_t'\mathbf{B}\tilde{\mathbf{r}}_{t+1},$$

where $\mathbf{B} := \mathbf{\Sigma}_x^{1/2}\mathbf{A}\mathbf{\Sigma}_r^{1/2}$

Decomposition of Strategy Returns

Apply CCA to the strategy returns:

- Re-express strategy returns in terms of synthetic variables

$$\mathbf{x}_t'\mathbf{A}\mathbf{r}_{t+1} = \lambda \times \mathbf{x}_t'\mathbf{\Sigma}_x^{-1}\mathbf{\Sigma}_{rx}'\mathbf{\Sigma}_r^{-1}\mathbf{r}_{t+1} = \lambda \times \tilde{\mathbf{x}}_t'\mathbf{\Sigma}_{\tilde{r}\tilde{x}}'\tilde{\mathbf{r}}_{t+1} = \tilde{\mathbf{x}}_t'\mathbf{B}\tilde{\mathbf{r}}_{t+1},$$

  where $\mathbf{B} := \mathbf{\Sigma}_x^{1/2}\mathbf{A}\mathbf{\Sigma}_r^{1/2}$

- Apply SVD on the cross-correlation matrix

$$\mathbf{x}_t'\mathbf{A}\mathbf{r}_{t+1} = \lambda \sum_{i=1}^{N} s_i(\mathbf{v}_i'\tilde{\mathbf{x}}_t)(\mathbf{u}_i'\tilde{\mathbf{r}}_{t+1})$$

Decomposition of Strategy Returns

Apply CCA to the strategy returns:

- Re-express strategy returns in terms of synthetic variables

$$\mathbf{x}_t'\mathbf{A}\mathbf{r}_{t+1} = \lambda \times \mathbf{x}_t'\mathbf{\Sigma}_x^{-1}\mathbf{\Sigma}_{rx}'\mathbf{\Sigma}_r^{-1}\mathbf{r}_{t+1} = \lambda \times \tilde{\mathbf{x}}_t'\mathbf{\Sigma}_{\tilde{r}\tilde{x}}'\tilde{\mathbf{r}}_{t+1} = \tilde{\mathbf{x}}_t'\mathbf{B}\tilde{\mathbf{r}}_{t+1},$$

  where $\mathbf{B} := \mathbf{\Sigma}_x^{1/2}\mathbf{A}\mathbf{\Sigma}_r^{1/2}$

- Apply SVD on the cross-correlation matrix

$$\mathbf{x}_t'\mathbf{A}\mathbf{r}_{t+1} = \lambda \sum_{i=1}^{N} s_i(\mathbf{v}_i'\tilde{\mathbf{x}}_t)(\mathbf{u}_i'\tilde{\mathbf{r}}_{t+1})$$

Two views of the strategy return:

- Combination of original assets and signals
- Sum of N uncorrelated long-short portfolios weighted by their canonical correlations

Financial Interpretation of CCA

Financial intuition of SVD:

- The change of basis reorganizes the set of N synthetic assets and N synthetic signals into a set of N uncorrelated strategies

Financial Interpretation of CCA

Financial intuition of SVD:

- The change of basis reorganizes the set of N synthetic assets and N synthetic signals into a set of N uncorrelated strategies
- Each of these strategies is expressed as a certain weighted combination of the synthetic assets and synthetic signals

Financial Interpretation of CCA

Financial intuition of SVD:

- The change of basis reorganizes the set of N synthetic assets and N synthetic signals into a set of N uncorrelated strategies
- Each of these strategies is expressed as a certain weighted combination of the synthetic assets and synthetic signals
- These combinations are the singular vectors, which we refer to as canonical portfolios

Financial Interpretation of CCA

Financial intuition of SVD:

- The change of basis reorganizes the set of N synthetic assets and N synthetic signals into a set of N uncorrelated strategies
- Each of these strategies is expressed as a certain weighted combination of the synthetic assets and synthetic signals
- These combinations are the singular vectors, which we refer to as canonical portfolios
- Ordering of canonical portfolios:

Financial Interpretation of CCA

Financial intuition of SVD:

- The change of basis reorganizes the set of N synthetic assets and N synthetic signals into a set of N uncorrelated strategies
- Each of these strategies is expressed as a certain weighted combination of the synthetic assets and synthetic signals
- These combinations are the singular vectors, which we refer to as canonical portfolios
- Ordering of canonical portfolios:
    - Highest canonical correlation $\rightarrow$ *most linearly predictable canonical portfolio*

Financial Interpretation of CCA

Financial intuition of SVD:

- The change of basis reorganizes the set of N synthetic assets and N synthetic signals into a set of N uncorrelated strategies
- Each of these strategies is expressed as a certain weighted combination of the synthetic assets and synthetic signals
- These combinations are the singular vectors, which we refer to as canonical portfolios
- Ordering of canonical portfolios:
  - Highest canonical correlation $\rightarrow$ *most linearly predictable canonical portfolio*
  - Second highest canonical correlation $\rightarrow$ *second most linearly predictable canonical portfolio*, etc.

Financial Interpretation of CCA

Financial intuition of SVD:

- The change of basis reorganizes the set of N synthetic assets and N synthetic signals into a set of N uncorrelated strategies
- Each of these strategies is expressed as a certain weighted combination of the synthetic assets and synthetic signals
- These combinations are the singular vectors, which we refer to as canonical portfolios
- Ordering of canonical portfolios:
    - Highest canonical correlation $\rightarrow$ *most linearly predictable canonical portfolio*
    - Second highest canonical correlation $\rightarrow$ *second most linearly predictable canonical portfolio*, etc.

Economically sensible:

- Optimizer is assigning capital to strategies that have high correlations and are also orthogonal to each other $\rightarrow$ diversification!

Expected Performance of Canonical Portfolios

- Return of the ith canonical portfolio:

$$\widetilde{\mathsf{PnL}}_{t+1,i} := (\mathbf{v}_i'\tilde{\mathbf{x}}_t)(\mathbf{u}_i'\tilde{\mathbf{r}}_{t+1})$$

- Assuming $\mathbf{r}_{t+1}$ and $\mathbf{x}_t$ are N-dimensional Gaussian variables, then using a result from Firoozye and Koshiyama (2020):

$$\mathbb{E}[\widetilde{\mathsf{PnL}}_{t+1,i}] = \mathbf{v}_i'\boldsymbol{\Sigma}_{\tilde{r}\tilde{x}}'\mathbf{u}_i = s_i, \text{ and } \mathsf{IR}_i := \frac{\mathbb{E}[\widetilde{\mathsf{PnL}}_{t+1,i}]}{\sqrt{\mathsf{Var}[\widetilde{\mathsf{PnL}}_{t+1,i}]}} = \frac{s_i}{\sqrt{1+s_i^2}}$$

- Return and Information Ratio of these canonical portfolios can be ranked according to the canonical correlation s_i.

Information Ratio and Canonical Correlations

- True, in-sample, and out-of-sample Information Ratio of strategy return:

$$\mathsf{IR} := \frac{\mathrm{Tr}(\mathbf{B}\boldsymbol{\Sigma}_{\tilde{r}\tilde{x}})}{\sqrt{\mathrm{Tr}(\mathbf{B}\mathbf{B}')}}, \quad \widehat{\mathsf{IR}} := \frac{\mathbb{E}[\mathrm{Tr}(\hat{\mathbf{B}}\mathbf{S}_{\tilde{r}\tilde{x}})]}{\sqrt{\mathbb{E}[\mathrm{Tr}(\hat{\mathbf{B}}\hat{\mathbf{B}}')]}}, \quad \text{and} \quad \mathsf{IR}^{\circ} := \frac{\mathbb{E}[\mathrm{Tr}(\hat{\mathbf{B}}\boldsymbol{\Sigma}_{\tilde{r}\tilde{x}})]}{\sqrt{\mathbb{E}[\mathrm{Tr}(\hat{\mathbf{B}}\hat{\mathbf{B}}')]}}$$

Information Ratio and Canonical Correlations

- True, in-sample, and out-of-sample Information Ratio of strategy return:

$$\mathrm{IR} := \frac{\mathrm{Tr}(\mathbf{B}\boldsymbol{\Sigma}_{\tilde{r}\tilde{x}})}{\sqrt{\mathrm{Tr}(\mathbf{B}\mathbf{B}')}}, \quad \widehat{\mathrm{IR}} := \frac{\mathbb{E}[\mathrm{Tr}(\hat{\mathbf{B}}\mathbf{S}_{\tilde{r}\tilde{x}})]}{\sqrt{\mathbb{E}[\mathrm{Tr}(\hat{\mathbf{B}}\hat{\mathbf{B}}')]}}, \quad \text{and} \quad \mathrm{IR}^\circ := \frac{\mathbb{E}[\mathrm{Tr}(\hat{\mathbf{B}}\boldsymbol{\Sigma}_{\tilde{r}\tilde{x}})]}{\sqrt{\mathbb{E}[\mathrm{Tr}(\hat{\mathbf{B}}\hat{\mathbf{B}}')]}}$$

- Express in terms of canonical correlations:

$$\mathrm{IR} = \sqrt{\sum_{i=1}^{N} s_i^2}, \quad \widehat{\mathrm{IR}} = \sqrt{\sum_{i=1}^{N} \mathbb{E}[\hat{s}_i^2]}, \quad \text{and} \quad \mathrm{IR}^\circ = \frac{\sum_{i=1}^{N} \mathbb{E}[\hat{s}_i s_i^\circ]}{\sqrt{\sum_{i=1}^{N} \mathbb{E}[\hat{s}_i^2]}},$$

where $s_i^\circ := \hat{\mathbf{u}}_i' \boldsymbol{\Sigma}_{\tilde{r}\tilde{x}} \hat{\mathbf{v}}_i$ is the out-of-sample canonical correlation

Information Ratio and Canonical Correlations

- True, in-sample, and out-of-sample Information Ratio of strategy return:

$$\mathrm{IR} := \frac{\mathrm{Tr}(\mathbf{B}\boldsymbol{\Sigma}_{\tilde{r}\tilde{x}})}{\sqrt{\mathrm{Tr}(\mathbf{B}\mathbf{B}')}}, \quad \widehat{\mathrm{IR}} := \frac{\mathbb{E}[\mathrm{Tr}(\hat{\mathbf{B}}\mathbf{S}_{\tilde{r}\tilde{x}})]}{\sqrt{\mathbb{E}[\mathrm{Tr}(\hat{\mathbf{B}}\hat{\mathbf{B}}')]}}, \quad \text{and} \quad \mathrm{IR}^{\circ} := \frac{\mathbb{E}[\mathrm{Tr}(\hat{\mathbf{B}}\boldsymbol{\Sigma}_{\tilde{r}\tilde{x}})]}{\sqrt{\mathbb{E}[\mathrm{Tr}(\hat{\mathbf{B}}\hat{\mathbf{B}}')]}}$$

- Express in terms of canonical correlations:

$$\mathrm{IR} = \sqrt{\sum_{i=1}^{N} s_i^2}, \quad \widehat{\mathrm{IR}} = \sqrt{\sum_{i=1}^{N} \mathbb{E}[\hat{s}_i^2]}, \quad \text{and} \quad \mathrm{IR}^{\circ} = \frac{\sum_{i=1}^{N} \mathbb{E}[\hat{s}_i s_i^{\circ}]}{\sqrt{\sum_{i=1}^{N} \mathbb{E}[\hat{s}_i^2]}},$$

  where $s_i^{\circ} := \hat{\mathbf{u}}_i' \boldsymbol{\Sigma}_{\tilde{r}\tilde{x}} \hat{\mathbf{v}}_i$ is the out-of-sample canonical correlation

- **Point**: We have a link between performance and canonical correlations

Information Ratio and Canonical Correlations

- True, in-sample, and out-of-sample Information Ratio of strategy return:

$$\mathrm{IR} := \frac{\mathrm{Tr}(\mathbf{B}\boldsymbol{\Sigma}_{\tilde{r}\tilde{x}})}{\sqrt{\mathrm{Tr}(\mathbf{B}\mathbf{B}')}}, \quad \widehat{\mathrm{IR}} := \frac{\mathbb{E}[\mathrm{Tr}(\hat{\mathbf{B}}\mathbf{S}_{\tilde{r}\tilde{x}})]}{\sqrt{\mathbb{E}[\mathrm{Tr}(\hat{\mathbf{B}}\hat{\mathbf{B}}')]}}, \text{ and } \mathrm{IR}^{\circ} := \frac{\mathbb{E}[\mathrm{Tr}(\hat{\mathbf{B}}\boldsymbol{\Sigma}_{\tilde{r}\tilde{x}})]}{\sqrt{\mathbb{E}[\mathrm{Tr}(\hat{\mathbf{B}}\hat{\mathbf{B}}')]}}$$

- Express in terms of canonical correlations:

$$\mathrm{IR} = \sqrt{\sum_{i=1}^{N} s_i^2}, \quad \widehat{\mathrm{IR}} = \sqrt{\sum_{i=1}^{N} \mathbb{E}[\hat{s}_i^2]}, \text{ and } \mathrm{IR}^{\circ} = \frac{\sum_{i=1}^{N} \mathbb{E}[\hat{s}_i s_i^{\circ}]}{\sqrt{\sum_{i=1}^{N} \mathbb{E}[\hat{s}_i^2]}},$$

  where $s_i^{\circ} := \hat{\mathbf{u}}_i' \boldsymbol{\Sigma}_{\tilde{r}\tilde{x}} \hat{\mathbf{v}}_i$ is the out-of-sample canonical correlation
- **Point**: We have a link between performance and canonical correlations
- Application: Using convexity arguments, one can show $\widehat{\mathrm{IR}} \geq \mathrm{IR}^{\circ}$

Information Ratio and Canonical Correlations

- True, in-sample, and out-of-sample Information Ratio of strategy return:

$$\text{IR} := \frac{\text{Tr}(\mathbf{B}\boldsymbol{\Sigma}_{\tilde{r}\tilde{x}})}{\sqrt{\text{Tr}(\mathbf{B}\mathbf{B}')}}, \quad \widehat{\text{IR}} := \frac{\mathbb{E}[\text{Tr}(\hat{\mathbf{B}}\mathbf{S}_{\tilde{r}\tilde{x}})]}{\sqrt{\mathbb{E}[\text{Tr}(\hat{\mathbf{B}}\hat{\mathbf{B}}')]}}, \quad \text{and} \quad \text{IR}^{\circ} := \frac{\mathbb{E}[\text{Tr}(\hat{\mathbf{B}}\boldsymbol{\Sigma}_{\tilde{r}\tilde{x}})]}{\sqrt{\mathbb{E}[\text{Tr}(\hat{\mathbf{B}}\hat{\mathbf{B}}')]}}$$

- Express in terms of canonical correlations:

$$\text{IR} = \sqrt{\sum_{i=1}^{N} s_i^2}, \quad \widehat{\text{IR}} = \sqrt{\sum_{i=1}^{N} \mathbb{E}[\hat{s}_i^2]}, \quad \text{and} \quad \text{IR}^{\circ} = \frac{\sum_{i=1}^{N} \mathbb{E}[\hat{s}_i s_i^{\circ}]}{\sqrt{\sum_{i=1}^{N} \mathbb{E}[\hat{s}_i^2]}},$$

  where $s_i^{\circ} := \hat{\mathbf{u}}_i' \boldsymbol{\Sigma}_{\tilde{r}\tilde{x}} \hat{\mathbf{v}}_i$ is the out-of-sample canonical correlation
- **Point**: We have a link between performance and canonical correlations
- Application: Using convexity arguments, one can show $\widehat{\text{IR}} \geq \text{IR}^{\circ}$
  - Overfitting: in-sample IR is always optimistic

Information Ratio and Canonical Correlations

- True, in-sample, and out-of-sample Information Ratio of strategy return:

$$\mathrm{IR} := \frac{\mathrm{Tr}(\mathbf{B}\boldsymbol{\Sigma}_{\tilde{r}\tilde{x}})}{\sqrt{\mathrm{Tr}(\mathbf{B}\mathbf{B}')}}, \quad \widehat{\mathrm{IR}} := \frac{\mathbb{E}[\mathrm{Tr}(\hat{\mathbf{B}}\mathbf{S}_{\tilde{r}\tilde{x}})]}{\sqrt{\mathbb{E}[\mathrm{Tr}(\hat{\mathbf{B}}\hat{\mathbf{B}}')]}}, \quad \text{and} \quad \mathrm{IR}^{\circ} := \frac{\mathbb{E}[\mathrm{Tr}(\hat{\mathbf{B}}\boldsymbol{\Sigma}_{\tilde{r}\tilde{x}})]}{\sqrt{\mathbb{E}[\mathrm{Tr}(\hat{\mathbf{B}}\hat{\mathbf{B}}')]}}$$

- Express in terms of canonical correlations:

$$\mathrm{IR} = \sqrt{\sum_{i=1}^{N} s_i^2}, \quad \widehat{\mathrm{IR}} = \sqrt{\sum_{i=1}^{N} \mathbb{E}[\hat{s}_i^2]}, \quad \text{and} \quad \mathrm{IR}^{\circ} = \frac{\sum_{i=1}^{N} \mathbb{E}[\hat{s}_i s_i^{\circ}]}{\sqrt{\sum_{i=1}^{N} \mathbb{E}[\hat{s}_i^2]}},$$

  where $s_i^{\circ} := \hat{\mathbf{u}}_i' \boldsymbol{\Sigma}_{\tilde{r}\tilde{x}} \hat{\mathbf{v}}_i$ is the out-of-sample canonical correlation

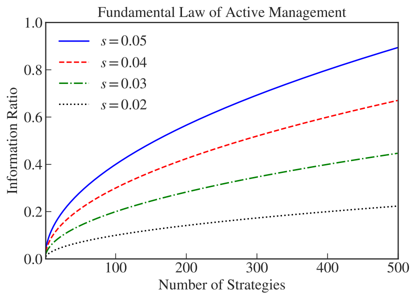
- **Point**: We have a link between performance and canonical correlations
- Application: Using convexity arguments, one can show $\widehat{\mathrm{IR}} \geq \mathrm{IR}^{\circ}$
  - Overfitting: in-sample IR is always optimistic
  - Need to shrink the canonical correlations and align singular vectors towards to the truth

Relation to Grinold's Law

- If $s_i = s$ for all i,

$$\mathrm{IR} = s\sqrt{N}$$

- *True* performance improves with
  - Strength of the *true* canonical correlation, s
  - Number of strategies, N
- Generalized **Fundamental Law of Active Management** of Grinold (1989)



Estimation

Theory to Practice

Most of the analysis so far takes place in the population

- In *practice*, need to replace the population objects with estimated ones

Three objects require estimation:

- Covariance of returns, $\boldsymbol{\Sigma}_r$
- Covariance of signals, $\boldsymbol{\Sigma}_x$
- Cross-covariance of returns and signals, $\boldsymbol{\Sigma}_{rx}$

Challenge: Last two of the above is less widely discussed in the literature

Covariance Estimation

Structured Estimators

- Factor models
  - Economic factors: Sharpe (1963), Fama and French (1994), etc.
  - Statistical factors: Connor, Goldberg, and Korajczyk (2010), Fan, Liao, and Mincheva (2014), etc.

Structure-Free Estimators

- Rotational equivariant estimators
  - Linear shrinkage to (scaled) identity matrix: Haff (1980), Ledoit and Wolf (2004)
  - Non-linear shrinkage: Stein (1975), El Karoui (2008), Ledoit and Wolf (2012, 2015, 2020), Abadir et al. (2014), Lam (2016), Bun, Bouchaud, and Potters (2017), and Tan and Zohren (2021, work-in-progress), etc.

We rely on structure-free estimators for $\hat{\boldsymbol{\Sigma}}_r$ and $\hat{\boldsymbol{\Sigma}}_x$

Cross-Covariance Estimation

Idea: Regularize the cross-correlation matrix as guided by CCA

Cross-Covariance Estimation

Idea: Regularize the cross-correlation matrix as guided by CCA

Proposed approach:

- Consider the estimator $\hat{\mathbf{\Sigma}}_{\tilde{r}\tilde{x}} := \hat{\varphi}\mathbb{I}_N$ where $\hat{\varphi} = \mathrm{Tr}(\mathbf{S}_{\tilde{r}\tilde{x}})/N$
- Since the sample cross-correlation matrix $\mathbf{S}_{\tilde{r}\tilde{x}}$ is not defined for $N > T$, we consider the following regularized version instead

$$\hat{\mathbf{\Sigma}}_r^{-1/2}\mathbf{S}_{rx}\hat{\mathbf{\Sigma}}_x^{-1/2}$$

- Build cross-covariance according to

$$\hat{\mathbf{\Sigma}}_{rx} := \hat{\mathbf{\Sigma}}_r^{1/2}\hat{\mathbf{\Sigma}}_{\tilde{r}\tilde{x}}\hat{\mathbf{\Sigma}}_x^{1/2} = \hat{\varphi} \cdot \hat{\mathbf{\Sigma}}_r^{1/2}\hat{\mathbf{\Sigma}}_x^{1/2}$$

Cross-Covariance Estimation

Idea: Regularize the cross-correlation matrix as guided by CCA

Proposed approach:

- Consider the estimator $\hat{\mathbf{\Sigma}}_{\tilde{r}\tilde{x}} := \hat{\varphi}\mathbb{I}_N$ where $\hat{\varphi} = \mathrm{Tr}(\mathbf{S}_{\tilde{r}\tilde{x}})/N$
- Since the sample cross-correlation matrix $\mathbf{S}_{\tilde{r}\tilde{x}}$ is not defined for $N > T$, we consider the following regularized version instead

$$\hat{\mathbf{\Sigma}}_r^{-1/2}\mathbf{S}_{rx}\hat{\mathbf{\Sigma}}_x^{-1/2}$$

- Build cross-covariance according to

$$\hat{\mathbf{\Sigma}}_{rx} := \hat{\mathbf{\Sigma}}_r^{1/2}\hat{\mathbf{\Sigma}}_{\tilde{r}\tilde{x}}\hat{\mathbf{\Sigma}}_x^{1/2} = \hat{\varphi}\cdot\hat{\mathbf{\Sigma}}_r^{1/2}\hat{\mathbf{\Sigma}}_x^{1/2}$$

Cross-covariations are allowed but when they are appropriately normalized, the cross relationships become insignificant

Canonical Strategy Parity

- Putting it all together, yields the portfolio

$$\hat{\mathbf{w}}_t = \text{scalar} \times \hat{\boldsymbol{\Sigma}}_r^{-1/2} \hat{\boldsymbol{\Sigma}}_x^{-1/2} \mathbf{x}_t$$

Canonical Strategy Parity

- Putting it all together, yields the portfolio

$$\hat{\mathbf{w}}_t = \text{scalar} \times \hat{\boldsymbol{\Sigma}}_r^{-1/2} \hat{\boldsymbol{\Sigma}}_x^{-1/2} \mathbf{x}_t$$

- Estimated Information Ratios of all canonical portfolios is the same with weights that are the standard basis
  $\implies$ Canonical strategy parity portfolio

Canonical Strategy Parity

- Putting it all together, yields the portfolio

$$\hat{\mathbf{w}}_t = \text{scalar} \times \hat{\boldsymbol{\Sigma}}_r^{-1/2} \hat{\boldsymbol{\Sigma}}_x^{-1/2} \mathbf{x}_t$$

- Estimated Information Ratios of all canonical portfolios is the same with weights that are the standard basis
  $\implies$ Canonical strategy parity portfolio
- The strategy return is a dot product between the synthetic returns and synthetic signals

$$\widehat{\text{PnL}}_{t+1} = \text{scalar} \times \mathbf{x}_t' \hat{\boldsymbol{\Sigma}}_x^{-1/2} \hat{\boldsymbol{\Sigma}}_r^{-1/2} \mathbf{r}_{t+1} = \text{scalar} \times \tilde{\mathbf{x}}_t \cdot \tilde{\mathbf{r}}_{t+1}$$

Canonical Strategy Parity

- Putting it all together, yields the portfolio

$$\hat{\mathbf{w}}_t = \text{scalar} \times \hat{\boldsymbol{\Sigma}}_r^{-1/2} \hat{\boldsymbol{\Sigma}}_x^{-1/2} \mathbf{x}_t$$

- Estimated Information Ratios of all canonical portfolios is the same with weights that are the standard basis
  $\implies$ Canonical strategy parity portfolio
- The strategy return is a dot product between the synthetic returns and synthetic signals

$$\widehat{\text{PnL}}_{t+1} = \text{scalar} \times \mathbf{x}_t' \hat{\boldsymbol{\Sigma}}_x^{-1/2} \hat{\boldsymbol{\Sigma}}_r^{-1/2} \mathbf{r}_{t+1} = \text{scalar} \times \tilde{\mathbf{x}}_t \cdot \tilde{\mathbf{r}}_{t+1}$$

- **Interesting property**: PnL is invariant under rotations of the assets and the signals c.f. Benichou et. al. (2016)

Canonical Strategy Parity

- Putting it all together, yields the portfolio

$$\hat{\mathbf{w}}_t = \text{scalar} \times \hat{\boldsymbol{\Sigma}}_r^{-1/2} \hat{\boldsymbol{\Sigma}}_x^{-1/2} \mathbf{x}_t$$

- Estimated Information Ratios of all canonical portfolios is the same with weights that are the standard basis
  $\implies$ Canonical strategy parity portfolio
- The strategy return is a dot product between the synthetic returns and synthetic signals

$$\widehat{\text{PnL}}_{t+1} = \text{scalar} \times \mathbf{x}_t' \hat{\boldsymbol{\Sigma}}_x^{-1/2} \hat{\boldsymbol{\Sigma}}_r^{-1/2} \mathbf{r}_{t+1} = \text{scalar} \times \tilde{\mathbf{x}}_t \cdot \tilde{\mathbf{r}}_{t+1}$$

- **Interesting property**: PnL is invariant under rotations of the assets and the signals c.f. Benichou et. al. (2016)
  - Strategy capacity reduced

Canonical Strategy Parity

- Putting it all together, yields the portfolio

$$\hat{\mathbf{w}}_t = \text{scalar} \times \hat{\boldsymbol{\Sigma}}_r^{-1/2} \hat{\boldsymbol{\Sigma}}_x^{-1/2} \mathbf{x}_t$$

- Estimated Information Ratios of all canonical portfolios is the same with weights that are the standard basis
  $\implies$ Canonical strategy parity portfolio
- The strategy return is a dot product between the synthetic returns and synthetic signals

$$\widehat{\text{PnL}}_{t+1} = \text{scalar} \times \mathbf{x}_t' \hat{\boldsymbol{\Sigma}}_x^{-1/2} \hat{\boldsymbol{\Sigma}}_r^{-1/2} \mathbf{r}_{t+1} = \text{scalar} \times \tilde{\mathbf{x}}_t \cdot \tilde{\mathbf{r}}_{t+1}$$

- **Interesting property**: PnL is invariant under rotations of the assets and the signals c.f. Benichou et. al. (2016)
  - Strategy capacity reduced
  - Increase robustness and protect against overfitting

Backtest Analysis

Data and Portfolio Construction Rules

Data:

- Get daily stock return data from CRSP
- Period: 01/01/1998 – 31/12/2019

Updating:

- Portfolios are rebalanced on a daily basis

Out-of-sample period:

- Begins on 29/01/1999
- This gives us 5265 daily returns

Data and Portfolio Construction Rules

Universe size:

- $N \in \{30, 50, 100, 250, 500\}$

Universe construction:

- Obtain new stocks at the beginning of each day
- Identify pairs of highly correlated stocks ($r > 0.95$), and remove the one with lower volume
- Select the N largest stocks by market-capitalization such that they have
    1. a nearly complete 252-day (~ 1 year) return history
    2. a complete 1-day return future

Estimation:

- Use past $T = 252$ days to estimate the covariances

Signal Construction

For the signal $\mathbf{x}_t$, we use the reversal of Lehmann (1990):

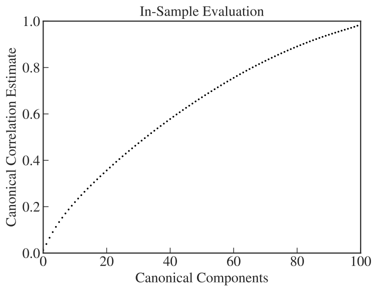
- Simple average over the past 21 days multiplied with a negative sign
- Economic rationale: Liquidity-provision of returns e.g. see Campbell et al. (1993), Pástor and Stambaugh (2003), and Nagel (2012)
- Can be computed from observed return data alone

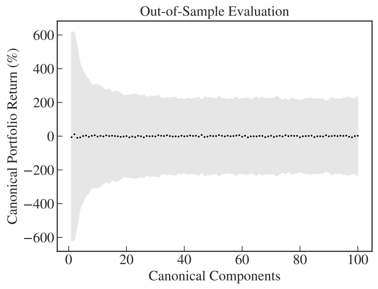
Pre-processing:

- Cross-sectionally de-mean, standardize, and winsorize

Analysis of Sample Cross-Correlation Matrix

- In-sample and out-of-sample returns of the canonical portfolios for $N = 100$, and $T = 252$; each dot is a value averaged over time
- Computations are based on $\mathbf{S}_r^{-1/2}\mathbf{S}_{rx}\mathbf{S}_x^{-1/2}$

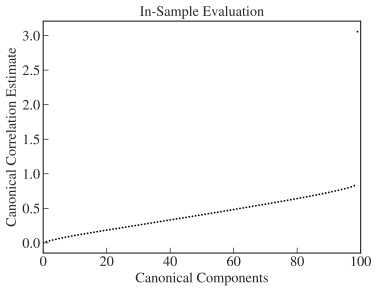


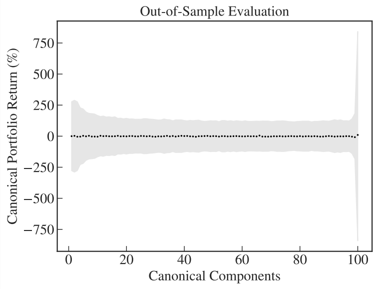


- No relationship between in-sample and out-of-sample canonical correlations
- Clear degradation of realized returns

Analysis of Regularized Sample Cross-Correlation Matrix

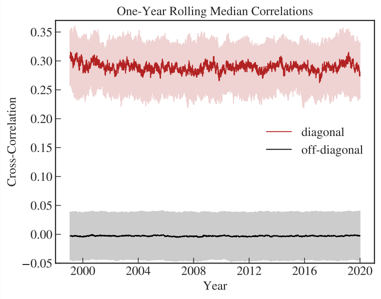
- In-sample and out-of-sample returns of the canonical portfolios for $N = 100$, and $T = 252$; each dot is a value averaged over time
- Computations are based on $\hat{\boldsymbol{\Sigma}}_r^{-1/2} \mathbf{S}_{rx} \hat{\boldsymbol{\Sigma}}_x^{-1/2}$



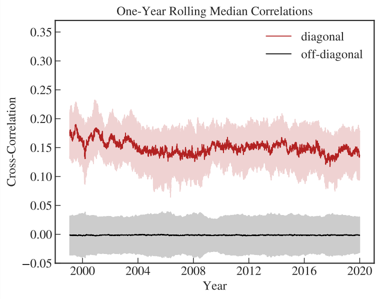


- No relationship between in-sample and out-of-sample canonical correlations
- Clear degradation of realized returns

Cross-Correlation Instabilities



(a) Sample Cross-Correlation Matrix



(b) Regularized Sample Cross-Correlation Matrix

- Off-diagonal elements consistently close to zero
- Diagonal elements are positive and are about 100 times larger in magnitude compared to the off-diagonal elements
- Makes sense: The ith reversal signal was constructed from the ith asset return for all $i \implies$ self-predictability is enforced by design!

Candidate Portfolios

Competing portfolios based on optimal matrix of coefficients $\mathbf{A}$:

- **EW**: Equally-weighted portfolio according to $\mathbf{x}_t$
- **MVO-SC**: sample covariance for $\hat{\mathbf{\Sigma}}_r$ and $\hat{\mathbf{\Sigma}}_x$ with $\hat{\mathbf{\Sigma}}_{rx} = \hat{\mathbf{\Sigma}}_x$
- **MVO-LS**: linear shrinkage for $\hat{\mathbf{\Sigma}}_r$ and $\hat{\mathbf{\Sigma}}_x$ with $\hat{\mathbf{\Sigma}}_{rx} = \hat{\mathbf{\Sigma}}_x$
- **MVO-NL**: nonlinear shrinkage for $\hat{\mathbf{\Sigma}}_r$ and $\hat{\mathbf{\Sigma}}_x$ with $\hat{\mathbf{\Sigma}}_{rx} = \hat{\mathbf{\Sigma}}_x$
- **CSP-SC**: sample covariances for $\hat{\mathbf{\Sigma}}_r$ and $\hat{\mathbf{\Sigma}}_x$ and (scaled) identity for $\hat{\mathbf{\Sigma}}_{\tilde{r}\tilde{x}}$
- **CSP-LS**: linear shrinkage for $\hat{\mathbf{\Sigma}}_r$ and $\hat{\mathbf{\Sigma}}_x$ and (scaled) identity for $\hat{\mathbf{\Sigma}}_{\tilde{r}\tilde{x}}$
- **CSP-NL**: nonlinear shrinkage for $\hat{\mathbf{\Sigma}}_r$ and $\hat{\mathbf{\Sigma}}_x$ and (scaled) identity for $\hat{\mathbf{\Sigma}}_{\tilde{r}\tilde{x}}$

Performance Measures

All measures are computed with 5265 out-of-sample returns

Annualized performance statistics:

- AV: Average Return
- SD: Standard Deviation
- IR: Information Ratio $\leftarrow$ our focus!

Average weight statistics:

- TO: Turnover
- GL: Gross Leverage
- N_{eff}: Number of effective positions (inverse of Herfindahl index)

Results: Performance Statistics

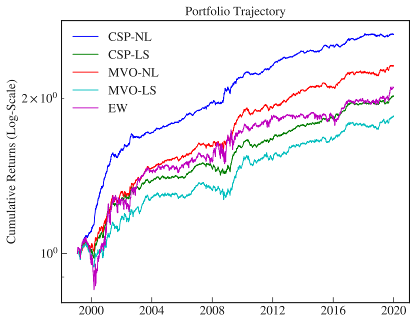
	EW	MVO-SC	MVO-LS	MVO-NL	CSP-SC	CSP-LS	CSP-NL
Out-of-Sample Period: 29/01/1999 to 31/12/2019							
				$N = 30$			
AV	4.04	5.43	5.15	4.98	8.86	7.16	7.52
SD	8.21	5.45	5.21	5.27	6.47	5.73	5.70
IR	0.49	1.00	0.99	0.95	**1.37**	1.25	1.32
				$N = 50$			
AV	4.36	4.76	4.44	4.40	10.03	7.03	7.84
SD	7.43	4.64	4.34	4.44	5.68	4.60	4.60
IR	0.59	1.03	1.02	0.99	**1.76**	1.53	1.70
				$N = 100$			
AV	5.68	6.26	5.68	5.86	11.37	7.08	7.67
SD	7.00	4.22	3.78	3.86	5.85	3.72	3.62
IR	0.81	1.48	1.50	1.52	1.94	1.90	**2.12**
				$N = 250$			
AV	4.87	4.15	4.84	5.87	177.80	5.61	7.27
SD	6.90	47.56	3.00	3.73	1925.77	2.68	3.69
IR	0.71	0.09	1.62	1.57	0.09	**2.09**	1.97
				$N = 500$			
AV	3.76	NaN	2.98	4.04	NaN	3.39	4.70
SD	6.53	NaN	3.08	2.52	NaN	2.30	1.92
IR	0.58	NaN	0.97	1.61	NaN	1.47	**2.44**

Most competitive IR value is highlighted in **blue**

Results: Weight Statistics

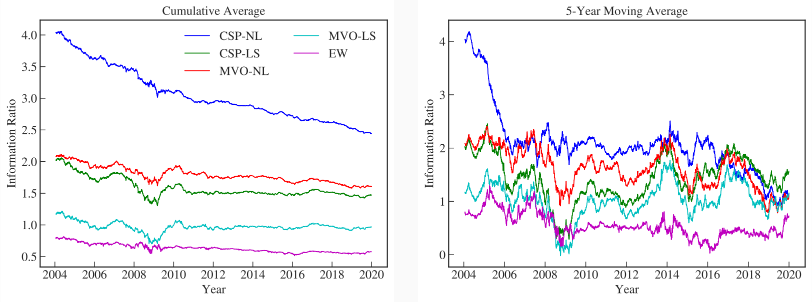
Out-of-Sample Period: 29/01/1999 to 31/12/2019							
	EW	MVO-SC	MVO-LS	MVO-NL	CSP-SC	CSP-LS	CSP-NL
				$N = 30$			
TO	0.23	0.38	0.33	0.33	0.77	0.54	0.62
GL	0.76	1.12	1.00	1.01	1.39	1.12	1.16
N_{eff}	31.69	16.75	21.09	20.79	10.86	16.84	15.54
				$N = 50$			
TO	0.23	0.41	0.34	0.34	1.10	0.61	0.75
GL	0.75	1.21	1.03	1.01	1.57	1.10	1.15
N_{eff}	53.01	22.34	31.66	32.75	13.22	27.87	25.06
				$N = 100$			
TO	0.23	0.53	0.38	0.34	2.12	0.69	0.91
GL	0.74	1.48	1.13	1.02	2.19	1.06	1.08
N_{eff}	107.06	29.14	52.75	63.76	12.91	59.36	55.43
				$N = 250$			
TO	0.22	23.89	0.43	0.87	1331.03	0.74	2.10
GL	0.73	22.42	1.29	1.28	867.26	1.02	1.61
N_{eff}	271.89	4.34	106.98	137.93	0.00	159.36	93.60
				$N = 500$			
TO	0.22	NaN	0.42	0.31	NaN	0.67	1.13
GL	0.72	NaN	1.36	0.96	NaN	1.00	0.93
N_{eff}	554.50	NaN	190.28	359.42	NaN	325.49	381.70

Portfolio Return Trajectory



Portfolio return trajectory of the EW, MVO and CSP schemes for different estimated covariances with universe size $N = 500$

Robustness Check: Subperiod Analysis



Annualized information ratio computed on a 5-year expanding (left panel) and rolling (right panel) basis. The universe size is $N = 500$

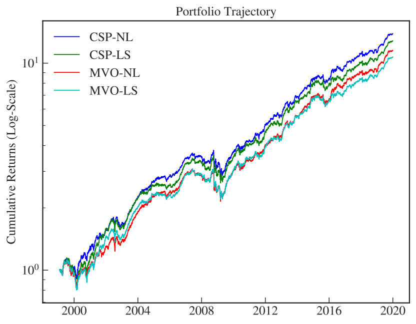
Fully-Invested Results: Performance Statistics

Out-of-Sample Period: 29/01/1999 to 31/12/2019							
	EW	FI-SC	FI-LS	FI-NL	CSP-SC	CSP-LS	CSP-NL
				$N = 30$			
AV	4.04	15.96	15.20	14.90	19.37	17.16	17.40
SD	8.21	14.92	14.48	14.56	15.26	14.64	14.66
IR	0.49	1.07	1.05	1.02	**1.27**	1.17	1.19
				$N = 50$			
AV	4.36	14.12	13.71	13.42	19.16	16.26	16.68
SD	7.43	14.23	13.67	13.62	14.66	13.77	13.68
IR	0.59	0.99	1.00	0.99	**1.31**	1.18	1.22
				$N = 100$			
AV	5.68	11.70	12.12	12.76	16.64	13.42	14.47
SD	7.00	13.60	12.27	11.95	14.25	12.22	11.85
IR	0.81	0.86	0.99	1.07	1.17	1.10	**1.22**
				$N = 250$			
AV	4.87	23.61	13.85	15.19	-195.75	14.70	16.76
SD	6.90	82.04	11.20	15.07	1402.58	11.12	15.10
IR	0.71	0.29	1.24	1.01	-0.14	**1.32**	1.11
				$N = 500$			
AV	3.76	NaN	11.80	12.11	NaN	12.65	12.98
SD	6.53	NaN	9.39	8.93	NaN	9.28	8.72
IR	0.58	NaN	1.26	1.36	NaN	1.36	**1.49**

Fully-Invested Results: Weight Statistics

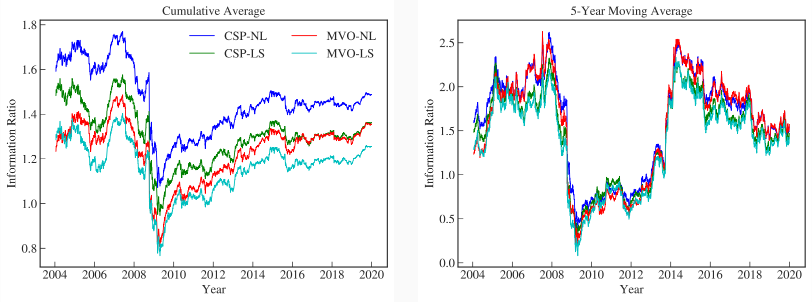
Out-of-Sample Period: 29/01/1999 to 31/12/2019							
	EW	FI-SC	FI-LS	FI-NL	CSP-SC	CSP-LS	CSP-NL
				$N = 30$			
TO	0.23	0.40	0.35	0.35	0.78	0.55	0.63
GL	0.76	2.42	2.12	2.14	2.57	2.19	2.22
N_{eff}	31.69	3.00	4.02	3.97	2.72	3.81	3.70
				$N = 50$			
TO	0.23	0.50	0.40	0.38	1.15	0.64	0.78
GL	0.75	3.11	2.61	2.51	3.28	2.65	2.57
N_{eff}	53.01	2.96	4.51	4.80	2.70	4.40	4.57
				$N = 100$			
TO	0.23	0.82	0.52	0.43	2.23	0.78	0.95
GL	0.74	4.75	3.52	2.97	5.02	3.50	2.99
N_{eff}	107.06	2.45	5.05	6.92	2.20	5.12	6.82
				$N = 250$			
TO	0.22	52.04	0.83	5.99	1002.81	1.05	6.50
GL	0.73	54.67	5.16	6.76	659.95	5.10	6.87
N_{eff}	271.89	0.08	6.36	6.76	0.00	6.54	6.42
				$N = 500$			
TO	0.22	NaN	0.74	0.40	NaN	0.93	1.17
GL	0.72	NaN	4.87	3.35	NaN	4.80	3.34
N_{eff}	554.50	NaN	13.49	28.63	NaN	13.90	28.77

Fully-Invested Portfolio Return Trajectory



Portfolio return trajectory of the EW, MVO and CSP schemes for different estimated covariances with universe size $N = 500$

Fully-Invested Robustness Check: Subperiod Analysis



Annualized information ratio computed on a 5-year expanding (left panel) and rolling (right panel) basis. The universe size is $N = 500$

Conclusion

Conclusion

Summary:

- Presented a simple framework that generalizes Markowitz (1952) by solving the optimal asset and signal combination problem in one stage
- Provided a novel decomposition of the solution using CCA
  - Allows for greater transparency for our proposed solution
- Employed covariance estimation techniques and innovations of our own to deal with the large-dimensional instabilities

Further work:

- Multiple heterogeneous signals and nonlinear extensions
- Transaction costs and signal decay (Firoozye, Tan, Zohren)
- Application to different datasets e.g. futures, fixed income, etc.

Thank You!