Bayes-Stein Estimation for Portfolio Analysis

Philippe Jorion, JFQA, Sep 1986

Krishnan Raman

Table of Contents

Prelude: Stein's Paradox

Stein's Paradox in Statistics, Efron & Morris, Scientific American(1977)

- There are estimators better than the arithmetic average.

Prelude: Stein's Paradox

Stein's Paradox in Statistics, Efron & Morris, Scientific American(1977)

- There are estimators better than the arithmetic average.
- Charles Stein at Stanford discovered such an estimator in 1955

Prelude: Stein's Paradox

Stein's Paradox in Statistics, Efron & Morris, Scientific American(1977)

- There are estimators better than the arithmetic average.
- Charles Stein at Stanford discovered such an estimator in 1955
- Shrinkage Estimator undermined *150 years* of estimation theory !!!

Prelude: Stein's Paradox

Stein's Paradox in Statistics, Efron & Morris, Scientific American(1977)

- There are estimators better than the arithmetic average.
- Charles Stein at Stanford discovered such an estimator in 1955
- Shrinkage Estimator undermined *150 years* of estimation theory !!!
- Led to long period of frequent, angry debate

Prelude: Stein's Paradox

Stein's Paradox in Statistics, Efron & Morris, Scientific American(1977)

- There are estimators better than the arithmetic average.
- Charles Stein at Stanford discovered such an estimator in 1955
- Shrinkage Estimator undermined *150 years* of estimation theory !!!
- Led to long period of frequent, angry debate
- Shrinkage estimators now routinely used in applied statistics

Prelude: Stein's Paradox

Stein's Paradox in Statistics, Efron & Morris, Scientific American(1977)

- There are estimators better than the arithmetic average.
- Charles Stein at Stanford discovered such an estimator in 1955
- Shrinkage Estimator undermined *150 years* of estimation theory !!!
- Led to long period of frequent, angry debate
- Shrinkage estimators now routinely used in applied statistics
- Stein's estimate contradicts most elementary law of statistical theory: individual average is the best estimator

Prelude: Stein's Paradox

Stein's Paradox in Statistics, Efron & Morris, Scientific American(1977)

- There are estimators better than the arithmetic average.
- Charles Stein at Stanford discovered such an estimator in 1955
- Shrinkage Estimator undermined *150 years* of estimation theory !!!
- Led to long period of frequent, angry debate
- Shrinkage estimators now routinely used in applied statistics
- Stein's estimate contradicts most elementary law of statistical theory: individual average is the best estimator
- Individual averages that exceed grand mean shrink, those that are smaller inflate.

Stein's Paradox, Dr. Richard Samworth , Statslab Cambridge

- **Data, Distribution, Parameter**
  $X_i \sim N(\theta_i, 1)$, $i = 1..p$, $p > 2$, $\theta = (\theta_1, \theta_2, \ldots, \theta_p)'$

Stein's Paradox, Dr. Richard Samworth, Statslab Cambridge

- **Data, Distribution, Parameter**
  $X_i \sim N(\theta_i, 1), i = 1..p, p > 2, \theta = (\theta_1, \theta_2, \ldots, \theta_p)'$
- **Estimator**
  MLE & UMVUE: $\hat{\theta}^0(X) = (X_1, X_2, \ldots, X_p)'$
  James-Stein: $\hat{\theta}^{JS}(X) = (1 - \frac{p-2}{\|X\|^2})X$

Stein's Paradox, Dr. Richard Samworth , Statslab Cambridge

- **Data, Distribution, Parameter**
  $X_i \sim N(\theta_i, 1), i = 1..p, p > 2, \theta = (\theta_1, \theta_2, \ldots, \theta_p)'$
- **Estimator**
  MLE & UMVUE: $\hat{\theta}^0(X) = (X_1, X_2, \ldots, X_p)'$
  James-Stein: $\hat{\theta}^{JS}(X) = (1 - \frac{p-2}{\|X\|^2})X$
- **Admissibility**
  Squared Error Loss: $L(\hat{\theta}, \theta) = \left|\left|\hat{\theta} - \theta\right|\right|^2$
  Risk Function: $R(\hat{\theta}, \theta) = E\{L(\hat{\theta}, \theta)\}$
  If $R(\hat{\theta}^1, \theta) \leq R(\hat{\theta}^2, \theta)\ \forall \theta$, we say $\hat{\theta}^2$ is **inadmissible**

Stein's Paradox, Dr. Richard Samworth , Statslab Cambridge

- **Data, Distribution, Parameter**
  $X_i \sim N(\theta_i, 1), i = 1..p, p > 2, \theta = (\theta_1, \theta_2, \ldots, \theta_p)'$
- **Estimator**
  MLE & UMVUE: $\hat{\theta}^0(X) = (X_1, X_2, \ldots, X_p)'$
  James-Stein: $\hat{\theta}^{JS}(X) = (1 - \frac{p-2}{\|X\|^2})X$
- **Admissibility**
  Squared Error Loss: $L(\hat{\theta}, \theta) = \left\|\hat{\theta} - \theta\right\|^2$
  Risk Function: $R(\hat{\theta}, \theta) = E\{L(\hat{\theta}, \theta)\}$
  If $R(\hat{\theta}^1, \theta) \leq R(\hat{\theta}^2, \theta)\ \forall\theta$, we say $\hat{\theta}^2$ is **inadmissible**
- **Stein's Paradox**
  $R(\hat{\theta}^{JS}, \theta) \leq R(\hat{\theta}^0, \theta)\ \forall\theta => \hat{\theta}^0(X)$ is inadmissible!

Prelude: Markowitz Portfolio Analysis

Markowitz: A 50 Year Retrospective, Rubinstein (2002)

- Title: Portfolio Selection
  Author: Harry Markowitz
  Journal: The Journal of Finance
  Date: March 1952

Prelude: Markowitz Portfolio Analysis

Markowitz: A 50 Year Retrospective, Rubinstein (2002)

- Title: Portfolio Selection
  Author: Harry Markowitz
  Journal: The Journal of Finance
  Date: March 1952
- Markowitz Paper was the **birth of Modern Financial Economics**

Prelude: Markowitz Portfolio Analysis

Markowitz: A 50 Year Retrospective, Rubinstein (2002)

- Title: Portfolio Selection
  Author: Harry Markowitz
  Journal: The Journal of Finance
  Date: March 1952
- Markowitz Paper was the **birth of Modern Financial Economics**
- Markowitz Paper in Finance $\sim$ Isaac Newton's contribution in Math & Physics

Prelude: Markowitz Portfolio Analysis

Markowitz: A 50 Year Retrospective, Rubinstein (2002)

- Title: Portfolio Selection
  Author: Harry Markowitz
  Journal: The Journal of Finance
  Date: March 1952
- Markowitz Paper was the **birth of Modern Financial Economics**
- Markowitz Paper in Finance $\sim$ Isaac Newton's contribution in Math & Physics
- Main Message: Diversification is good

Prelude: Markowitz Portfolio Analysis

Markowitz: A 50 Year Retrospective, Rubinstein (2002)

- Title: Portfolio Selection
  Author: Harry Markowitz
  Journal: The Journal of Finance
  Date: March 1952
- Markowitz Paper was the **birth of Modern Financial Economics**
- Markowitz Paper in Finance $\sim$ Isaac Newton's contribution in Math & Physics
- Main Message: Diversification is good
- "But divide your investments among many places, for you do not know what risks might lie ahead" - Book of Ecclesiastes, Bible

Prelude: Markowitz Portfolio Analysis

Markowitz: A 50 Year Retrospective, Rubinstein (2002)

- Title: Portfolio Selection
  Author: Harry Markowitz
  Journal: The Journal of Finance
  Date: March 1952
- Markowitz Paper was the **birth of Modern Financial Economics**
- Markowitz Paper in Finance $\sim$ Isaac Newton's contribution in Math & Physics
- Main Message: Diversification is good
- "But divide your investments among many places, for you do not know what risks might lie ahead" - Book of Ecclesiastes, Bible
- "Split one's assets into thirds: one third in business, one third gold, one third in land" - Talmud

Prelude: Markowitz Portfolio Analysis

Markowitz: A 50 Year Retrospective, Rubinstein (2002)

- Title: Portfolio Selection
  Author: Harry Markowitz
  Journal: The Journal of Finance
  Date: March 1952
- Markowitz Paper was the **birth of Modern Financial Economics**
- Markowitz Paper in Finance $\sim$ Isaac Newton's contribution in Math & Physics
- Main Message: Diversification is good
- "But divide your investments among many places, for you do not know what risks might lie ahead" - Book of Ecclesiastes, Bible
- "Split one's assets into thirds: one third in business, one third gold, one third in land" - Talmud
- ". . . I thank my fortune for it,
  My ventures are not in one bottom trusted,
  Nor to one place; nor is my whole estate
  Upon the fortune of this present year . . ."
  Antonio in *The Merchant of Venice*, William Shakespeare

Prelude: Markowitz Portfolio Analysis

Markowitz: A 50 Year Retrospective, Rubinstein (2002)

- Title: Portfolio Selection
  Author: Harry Markowitz
  Journal: The Journal of Finance
  Date: March 1952
- Markowitz Paper was the **birth of Modern Financial Economics**
- Markowitz Paper in Finance $\sim$ Isaac Newton's contribution in Math & Physics
- Main Message: Diversification is good
- "But divide your investments among many places, for you do not know what risks might lie ahead" - Book of Ecclesiastes, Bible
- "Split one's assets into thirds: one third in business, one third gold, one third in land" - Talmud
- "... I thank my fortune for it,
  My ventures are not in one bottom trusted,
  Nor to one place; nor is my whole estate
  Upon the fortune of this present year . . ."
  Antonio in *The Merchant of Venice*, William Shakespeare
- "...it is advisable to divide goods which are exposed to some small danger into several portions rather than to risk them all together" - Daniel Bernoulli, 1738 (LLN, 1715)

Prelude: Markowitz Portfolio Analysis

Markowitz: A 50 Year Retrospective, Rubinstein (2002)

- Title: Portfolio Selection
  Author: Harry Markowitz
  Journal: The Journal of Finance
  Date: March 1952
- Markowitz Paper was the **birth of Modern Financial Economics**
- Markowitz Paper in Finance $\sim$ Isaac Newton's contribution in Math & Physics
- Main Message: Diversification is good
- "But divide your investments among many places, for you do not know what risks might lie ahead" - Book of Ecclesiastes, Bible
- "Split one's assets into thirds: one third in business, one third gold, one third in land" - Talmud
- ". . . I thank my fortune for it,
  My ventures are not in one bottom trusted,
  Nor to one place; nor is my whole estate
  Upon the fortune of this present year . . ."
  Antonio in *The Merchant of Venice*, William Shakespeare
- ". . . it is advisable to divide goods which are exposed to some small danger into several portions rather than to risk them all together" - Daniel Bernoulli, 1738 (LLN, 1715)
- Keynes, 1920s "If possible,hold assets with opposed risks since they are likely to move in opposite directions when there are general fluctuation"

Prelude: Markowitz Portfolio Analysis

Markowitz: A 50 Year Retrospective, Rubinstein (2002)

- Title: Portfolio Selection
  Author: Harry Markowitz
  Journal: The Journal of Finance
  Date: March 1952
- Markowitz Paper was the **birth of Modern Financial Economics**
- Markowitz Paper in Finance $\sim$ Isaac Newton's contribution in Math & Physics
- Main Message: Diversification is good
- "But divide your investments among many places, for you do not know what risks might lie ahead" - Book of Ecclesiastes, Bible
- "Split one's assets into thirds: one third in business, one third gold, one third in land" - Talmud
- "... I thank my fortune for it,
  My ventures are not in one bottom trusted,
  Nor to one place; nor is my whole estate
  Upon the fortune of this present year . . ."
  Antonio in *The Merchant of Venice*, William Shakespeare
- "...it is advisable to divide goods which are exposed to some small danger into several portions rather than to risk them all together" - Daniel Bernoulli, 1738 (LLN, 1715)
- Keynes, 1920s "If possible,hold assets with opposed risks since they are likely to move in opposite directions when there are general fluctuation"
- But how to quantify diversification ?

Prelude: Markowitz Portfolio Analysis

Markowitz: A 50 Year Retrospective, Rubinstein (2002)

- Title: Portfolio Selection
  Author: Harry Markowitz
  Journal: The Journal of Finance
  Date: March 1952
- Markowitz Paper was the **birth of Modern Financial Economics**
- Markowitz Paper in Finance $\sim$ Isaac Newton's contribution in Math & Physics
- Main Message: Diversification is good
- "But divide your investments among many places, for you do not know what risks might lie ahead" - Book of Ecclesiastes, Bible
- "Split one's assets into thirds: one third in business, one third gold, one third in land" - Talmud
- ". . . I thank my fortune for it,
  My ventures are not in one bottom trusted,
  Nor to one place; nor is my whole estate
  Upon the fortune of this present year . . ."
  Antonio in *The Merchant of Venice*, William Shakespeare
- ". . . it is advisable to divide goods which are exposed to some small danger into several portions rather than to risk them all together" - Daniel Bernoulli, 1738 (LLN, 1715)
- Keynes, 1920s "If possible,hold assets with opposed risks since they are likely to move in opposite directions when there are general fluctuation"
- But how to quantify diversification ?
- Jacob Marschak(Money & the theory of assets, 1938, Econometrica)

Prelude: Markowitz Portfolio Analysis

Markowitz: A 50 Year Retrospective, Rubinstein (2002)

- Title: Portfolio Selection
  Author: Harry Markowitz
  Journal: The Journal of Finance
  Date: March 1952
- Markowitz Paper was the **birth of Modern Financial Economics**
- Markowitz Paper in Finance $\sim$ Isaac Newton's contribution in Math & Physics
- Main Message: Diversification is good
- "But divide your investments among many places, for you do not know what risks might lie ahead" - Book of Ecclesiastes, Bible
- "Split one's assets into thirds: one third in business, one third gold, one third in land" - Talmud
- "... I thank my fortune for it,
  My ventures are not in one bottom trusted,
  Nor to one place; nor is my whole estate
  Upon the fortune of this present year . . ."
  Antonio in *The Merchant of Venice*, William Shakespeare
- "...it is advisable to divide goods which are exposed to some small danger into several portions rather than to risk them all together" - Daniel Bernoulli, 1738 (LLN, 1715)
- Keynes, 1920s "If possible,hold assets with opposed risks since they are likely to move in opposite directions when there are general fluctuation"
- But how to quantify diversification ?
- Jacob Marschak(Money & the theory of assets, 1938, Econometrica)
- Marschak's idea: Use the means and the covariance matrix of commodities

Prelude: Markowitz Portfolio Analysis

Markowitz: A 50 Year Retrospective, Rubinstein (2002)

- Title: Portfolio Selection
  Author: Harry Markowitz
  Journal: The Journal of Finance
  Date: March 1952
- Markowitz Paper was the **birth of Modern Financial Economics**
- Markowitz Paper in Finance $\sim$ Isaac Newton's contribution in Math & Physics
- Main Message: Diversification is good
- "But divide your investments among many places, for you do not know what risks might lie ahead" - Book of Ecclesiastes, Bible
- "Split one's assets into thirds: one third in business, one third gold, one third in land" - Talmud
- "... I thank my fortune for it,
  My ventures are not in one bottom trusted,
  Nor to one place; nor is my whole estate
  Upon the fortune of this present year . . ."
  Antonio in *The Merchant of Venice*, William Shakespeare
- "...it is advisable to divide goods which are exposed to some small danger into several portions rather than to risk them all together" - Daniel Bernoulli, 1738 (LLN, 1715)
- Keynes, 1920s "If possible,hold assets with opposed risks since they are likely to move in opposite directions when there are general fluctuation"
- But how to quantify diversification ?
- Jacob Marschak(Money & the theory of assets, 1938, Econometrica)
- Marschak's idea: Use the means and the covariance matrix of commodities
- A.D. Roy (Safety first & the holding of assets, 1952, Econometrica)

Prelude: Markowitz Portfolio Analysis

Markowitz: A 50 Year Retrospective, Rubinstein (2002)

- Title: Portfolio Selection
  Author: Harry Markowitz
  Journal: The Journal of Finance
  Date: March 1952
- Markowitz Paper was the **birth of Modern Financial Economics**
- Markowitz Paper in Finance $\sim$ Isaac Newton's contribution in Math & Physics
- Main Message: Diversification is good
- "But divide your investments among many places, for you do not know what risks might lie ahead" - Book of Ecclesiastes, Bible
- "Split one's assets into thirds: one third in business, one third gold, one third in land" - Talmud
- "... I thank my fortune for it,
  My ventures are not in one bottom trusted,
  Nor to one place; nor is my whole estate
  Upon the fortune of this present year . . ."
  Antonio in *The Merchant of Venice*, William Shakespeare
- "...it is advisable to divide goods which are exposed to some small danger into several portions rather than to risk them all together" - Daniel Bernoulli, 1738 (LLN, 1715)
- Keynes, 1920s "If possible,hold assets with opposed risks since they are likely to move in opposite directions when there are general fluctuation"
- But how to quantify diversification ?
- Jacob Marschak(Money & the theory of assets, 1938, Econometrica)
- Marschak's idea: Use the means and the covariance matrix of commodities
- A.D. Roy (Safety first & the holding of assets, 1952, Econometrica)
- Roy's idea: Maximize $\frac{\mu_p - d}{\sigma_p^2}$, d = disaster return

Prelude: Markowitz Portfolio Analysis

Markowitz: A 50 Year Retrospective, Rubinstein (2002)

- Markowitz' most important contribution was the formulation.

Prelude: Markowitz Portfolio Analysis

Markowitz: A 50 Year Retrospective, Rubinstein (2002)

- Markowitz' most important contribution was the formulation.
- Maximize expected portfolio return μ_p

Prelude: Markowitz Portfolio Analysis

Markowitz: A 50 Year Retrospective, Rubinstein (2002)

- Markowitz' most important contribution was the formulation.
- Maximize expected portfolio return μ_p
- while Minimizing portfolio variance of return σ_p^2

Prelude: Markowitz Portfolio Analysis

Markowitz: A 50 Year Retrospective, Rubinstein (2002)

- Markowitz' most important contribution was the formulation.
- Maximize expected portfolio return μ_p
- while Minimizing portfolio variance of return σ_p^2
- $\mu_p = \sum x_j \mu_j$ *subject to* $\sum x_j = 1$

Prelude: Markowitz Portfolio Analysis

Markowitz: A 50 Year Retrospective, Rubinstein (2002)

- Markowitz' most important contribution was the formulation.
- Maximize expected portfolio return μ_p
- while Minimizing portfolio variance of return σ_p^2
- $\mu_p = \sum x_j \mu_j \textit{ subject to } \sum x_j = 1$
- $\sigma_p^2 = \sum_j x_j^2 \sigma_j^2 + \sum_j \sum_{k \neq j} x_j x_k \rho_{jk} \sigma_j \sigma_k$

Prelude: Markowitz Portfolio Analysis

Markowitz: A 50 Year Retrospective, Rubinstein (2002)

- Markowitz' most important contribution was the formulation.
- Maximize expected portfolio return μ_p
- while Minimizing portfolio variance of return σ_p^2
- $\mu_p = \sum x_j \mu_j \textit{ subject to } \sum x_j = 1$
- $\sigma_p^2 = \sum_j x_j^2 \sigma_j^2 + \sum_j \sum_{k \neq j} x_j x_k \rho_{jk} \sigma_j \sigma_k$
- Inputs to the Markowitz Problem: The means μ_j, the variances σ_j^2, the correlations ρ_{jk}

Prelude: Markowitz Portfolio Analysis

Markowitz: A 50 Year Retrospective, Rubinstein (2002)

- Markowitz' most important contribution was the formulation.
- Maximize expected portfolio return μ_p
- while Minimizing portfolio variance of return σ_p^2
- $\mu_p = \sum x_j \mu_j \textit{ subject to } \sum x_j = 1$
- $\sigma_p^2 = \sum_j x_j^2 \sigma_j^2 + \sum_j \sum_{k \neq j} x_j x_k \rho_{jk} \sigma_j \sigma_k$
- Inputs to the Markowitz Problem: The means μ_j, the variances σ_j^2, the correlations ρ_{jk}
- Output of the Markowitz Problem: The x_j, optimal proportion of weights that sum to 1.

Introduction

Jorion: Section I

- Markowitz is a 2 step procedure.

Introduction

Jorion: Section I

- Markowitz is a 2 step procedure.
- Step 0: Pick a bunch of stocks
  Grab time series of their historical returns

Introduction

Jorion: Section I

- Markowitz is a 2 step procedure.
- Step 0: Pick a bunch of stocks
  Grab time series of their historical returns
- Step 1: Estimate moments (mean, variance) from time series
  Certainity Equivalence : Treat estimates as population parameters !!!
  Certainity Equivalence bad idea: Barry,Brown,Klein, Dickenson et al.
  Markowitz procedure is inherently multivariate
  Cannot estimate each expected return independently!
  Sample Mean inadmissible if portfolio has >2 stocks.
  Portfolio with multiple stocks must use Stein Estimator

Introduction

Jorion: Section I

- Markowitz is a 2 step procedure.
- Step 0: Pick a bunch of stocks
  Grab time series of their historical returns
- Step 1: Estimate moments (mean, variance) from time series
  Certainity Equivalence : Treat estimates as population parameters !!!
  Certainity Equivalence bad idea: Barry,Brown,Klein, Dickenson et al.
  Markowitz procedure is inherently multivariate
  Cannot estimate each expected return independently!
  Sample Mean inadmissible if portfolio has $>$2 stocks.
  Portfolio with multiple stocks must use Stein Estimator
- Step 2: Solve *Markowitz problem* using moments computed above.

Estimation Risk

Jorion: Section II

- Given a vector of returns r, choose set of weights q

Estimation Risk

Jorion: Section II

- Given a vector of returns r, choose set of weights q
- Return on portfolio z = q'r

Estimation Risk

Jorion: Section II

- Given a vector of returns r, choose set of weights q
- Return on portfolio z = q'r
- Maximize the expected utility of the return
  $EU(z) = \int U(z)p(z|\theta)dz$
  $U(z)$ a utility function
  $p(z|\theta)$ Distribution of return, conditioned on unknown θ

Estimation Risk

Jorion: Section II

- Given a vector of returns r, choose set of weights q
- Return on portfolio z = q'r
- Maximize the expected utility of the return
  $EU(z) = \int U(z)p(z|\theta)dz$
  $U(z)$ a utility function
  $p(z|\theta)$ Distribution of return, conditioned on unknown θ
- Certainty equivalence: Assumes $\theta = \hat{\theta}$

Estimation Risk

Jorion: Section II

- Given a vector of returns r, choose set of weights q
- Return on portfolio z = q'r
- Maximize the expected utility of the return
  $EU(z) = \int U(z)p(z|\theta)dz$
  $U(z)$ a utility function
  $p(z|\theta)$ Distribution of return, conditioned on unknown θ
- Certainty equivalence: Assumes $\theta = \hat{\theta}$
- Certainty equivalence obviously ignores parameter uncertainty.

Estimation Risk

Jorion: Section II

- Given a vector of returns r, choose set of weights q
- Return on portfolio z = q'r
- Maximize the expected utility of the return
  $EU(z) = \int U(z)p(z|\theta)dz$
  $U(z)$ a utility function
  $p(z|\theta)$ Distribution of return, conditioned on unknown θ
- Certainty equivalence: Assumes $\theta = \hat{\theta}$
- Certainty equivalence obviously ignores parameter uncertainty.
- Bayesian solution (Zellner/Chetty):
  Optimal portfolio choice via Predictive Density function
  Integrate out the unknown parameter θ
  Integrating out will explicitly takes into account uncertainty about θ

Estimation Risk

Jorion: Section II

- Given a vector of returns r, choose set of weights q
- Return on portfolio z = q'r
- Maximize the expected utility of the return
  $EU(z) = \int U(z)p(z|\theta)dz$
  $U(z)$ a utility function
  $p(z|\theta)$ Distribution of return, conditioned on unknown θ
- Certainty equivalence: Assumes $\theta = \hat{\theta}$
- Certainty equivalence obviously ignores parameter uncertainty.
- Bayesian solution (Zellner/Chetty):
  Optimal portfolio choice via Predictive Density function
  Integrate out the unknown parameter θ
  Integrating out will explicitly takes into account uncertainty about θ
- $max\ E_\theta[E_{y|\theta}[U(z)|\theta]]$
  $= \int \int U(z)p(z|\theta)dz\ p(\theta|y, I_0)d\theta$
  $= \int U(z)[\ \int p(z|\theta)\ p(\theta|y, I_0)d\theta\ ]dz$

Estimation Risk

Jorion: Section II

- Given a vector of returns r, choose set of weights q
- Return on portfolio z = q'r
- Maximize the expected utility of the return
  $EU(z) = \int U(z)p(z|\theta)dz$
  $U(z)$ a utility function
  $p(z|\theta)$ Distribution of return, conditioned on unknown θ
- Certainty equivalence: Assumes $\theta = \hat{\theta}$
- Certainty equivalence obviously ignores parameter uncertainty.
- Bayesian solution (Zellner/Chetty):
  Optimal portfolio choice via Predictive Density function
  Integrate out the unknown parameter θ
  Integrating out will explicitly takes into account uncertainty about θ
- $max\ E_\theta[E_{y|\theta}[U(z)|\theta]]$
  $= \int \int U(z)p(z|\theta)dz\ p(\theta|y, I_0)d\theta$
  $= \int U(z)[\ \int p(z|\theta)\ p(\theta|y, I_0)d\theta\ ]dz$
- Predictive Density Function of $z = q'r = p(z|\theta)\ p(\theta|y, I_0)d\theta$

Estimation Risk

Jorion: Section II

- Given a vector of returns r, choose set of weights q
- Return on portfolio z = q'r
- Maximize the expected utility of the return
  $EU(z) = \int U(z)p(z|\theta)dz$
  $U(z)$ a utility function
  $p(z|\theta)$ Distribution of return, conditioned on unknown θ
- Certainty equivalence: Assumes $\theta = \hat{\theta}$
- Certainty equivalence obviously ignores parameter uncertainty.
- Bayesian solution (Zellner/Chetty):
  Optimal portfolio choice via Predictive Density function
  Integrate out the unknown parameter θ
  Integrating out will explicitly takes into account uncertainty about θ
- $max\ E_\theta[E_{y|\theta}[U(z)|\theta]]$
  $= \int\int U(z)p(z|\theta)dz\ p(\theta|y, I_0)d\theta$
  $= \int U(z)[\ \int p(z|\theta)\ p(\theta|y, I_0)d\theta\ ]dz$
- Predictive Density Function of $z = q'r = p(z|\theta)\ p(\theta|y, I_0)d\theta$
- Posterior Density of $\theta = \ p(\theta|y, I_0)$

Estimation Risk

Jorion: Section II

- Utility function $EU(z) = F(\mu_z, \sigma_z^2)$

Estimation Risk

Jorion: Section II

- Utility function $EU(z) = F(\mu_z, \sigma_z^2)$
- Expected Portfolio Return: $\mu_z = q'\mu$

Estimation Risk

Jorion: Section II

- Utility function $EU(z) = F(\mu_z, \sigma_z^2)$
- Expected Portfolio Return: $\mu_z = q'\mu$
- Variance of Portfolio: $\sigma_z^2 = q'\Sigma q$

Estimation Risk

Jorion: Section II

- Utility function $EU(z) = F(\mu_z, \sigma_z^2)$
- Expected Portfolio Return: $\mu_z = q'\mu$
- Variance of Portfolio: $\sigma_z^2 = q'\Sigma q$
- q = The vector of investment *proportions*

Estimation Risk

Jorion: Section II

- Utility function $EU(z) = F(\mu_z, \sigma_z^2)$
- Expected Portfolio Return: $\mu_z = q'\mu$
- Variance of Portfolio: $\sigma_z^2 = q'\Sigma q$
- q = The vector of investment *proportions*
- Choose q to maximize expected utility
  Subject to the constraint investment proportions sum to unity i.e. $\Sigma q_i = 1$

Estimation Risk

Jorion: Section II

- Utility function $EU(z) = F(\mu_z, \sigma_z^2)$
- Expected Portfolio Return: $\mu_z = q'\mu$
- Variance of Portfolio: $\sigma_z^2 = q'\Sigma q$
- q = The vector of investment *proportions*
- Choose q to maximize expected utility
  Subject to the constraint investment proportions sum to unity i.e. $\Sigma q_i = 1$
- Suppose we know $\theta = (\mu, \Sigma)$

Estimation Risk

Jorion: Section II

- Utility function $EU(z) = F(\mu_z, \sigma_z^2)$
- Expected Portfolio Return: $\mu_z = q'\mu$
- Variance of Portfolio: $\sigma_z^2 = q'\Sigma q$
- q = The vector of investment *proportions*
- Choose q to maximize expected utility
  Subject to the constraint investment proportions sum to unity i.e. $\Sigma q_i = 1$
- Suppose we know $\theta = (\mu, \Sigma)$
- Optimal portfolio is $F_{MAX} = F(q_*'\mu, q_*'\Sigma q_*)$
  But we don't know θ

Estimation Risk

Jorion: Section II

- Utility function $EU(z) = F(\mu_z, \sigma_z^2)$
- Expected Portfolio Return: $\mu_z = q'\mu$
- Variance of Portfolio: $\sigma_z^2 = q'\Sigma q$
- q = The vector of investment *proportions*
- Choose q to maximize expected utility
  Subject to the constraint investment proportions sum to unity i.e. $\Sigma q_i = 1$
- Suppose we know $\theta = (\mu, \Sigma)$
- Optimal portfolio is $F_{MAX} = F(q_*'\mu, q_*'\Sigma q_*)$
  But we don't know θ
- So we choose q based on sample estimate $\hat{\theta}(y)$

Estimation Risk

Jorion: Section II

- Utility function $EU(z) = F(\mu_z, \sigma_z^2)$
- Expected Portfolio Return: $\mu_z = q'\mu$
- Variance of Portfolio: $\sigma_z^2 = q'\Sigma q$
- q = The vector of investment *proportions*
- Choose q to maximize expected utility
  Subject to the constraint investment proportions sum to unity i.e. $\Sigma q_i = 1$
- Suppose we know $\theta = (\mu, \Sigma)$
- Optimal portfolio is $F_{MAX} = F(q_*'\mu, q_*'\Sigma q_*)$
  But we don't know θ
- So we choose q based on sample estimate $\hat{\theta}(y)$
- The expected utility will now necessarily be lower than before

Estimation Risk

Jorion: Section II

- Utility function $EU(z) = F(\mu_z, \sigma_z^2)$
- Expected Portfolio Return: $\mu_z = q'\mu$
- Variance of Portfolio: $\sigma_z^2 = q'\Sigma q$
- q = The vector of investment *proportions*
- Choose q to maximize expected utility
  Subject to the constraint investment proportions sum to unity i.e. $\Sigma q_i = 1$
- Suppose we know $\theta = (\mu, \Sigma)$
- Optimal portfolio is $F_{MAX} = F(q_*'\mu, q_*'\Sigma q_*)$
  But we don't know θ
- So we choose q based on sample estimate $\hat{\theta}(y)$
- The expected utility will now necessarily be lower than before
- $F(\hat{q}(\hat{\theta}(y)|\mu, \Sigma) = F(\hat{q}'\mu, \hat{q}'\Sigma\hat{q}) \leq F_{MAX}$

Estimation Risk

Jorion: Section II

- Utility function $EU(z) = F(\mu_z, \sigma_z^2)$
- Expected Portfolio Return: $\mu_z = q'\mu$
- Variance of Portfolio: $\sigma_z^2 = q'\Sigma q$
- q = The vector of investment *proportions*
- Choose q to maximize expected utility
  Subject to the constraint investment proportions sum to unity i.e. $\Sigma q_i = 1$
- Suppose we know $\theta = (\mu, \Sigma)$
- Optimal portfolio is $F_{MAX} = F(q_*'\mu, q_*'\Sigma q_*)$
  But we don't know θ
- So we choose q based on sample estimate $\hat{\theta}(y)$
- The expected utility will now necessarily be lower than before
- $F(\hat{q}(\hat{\theta}(y)|\mu, \Sigma) = F(\hat{q}'\mu, \hat{q}'\Sigma\hat{q}) \leq F_{MAX}$
- The loss due to estimation risk can be measured

Estimation Risk

Jorion: Section II

- Utility function $EU(z) = F(\mu_z, \sigma_z^2)$
- Expected Portfolio Return: $\mu_z = q'\mu$
- Variance of Portfolio: $\sigma_z^2 = q'\Sigma q$
- q = The vector of investment *proportions*
- Choose q to maximize expected utility
  Subject to the constraint investment proportions sum to unity i.e. $\Sigma q_i = 1$
- Suppose we know $\theta = (\mu, \Sigma)$
- Optimal portfolio is $F_{MAX} = F(q_*'\mu, q_*'\Sigma q_*)$
  But we don't know θ
- So we choose q based on sample estimate $\hat{\theta}(y)$
- The expected utility will now necessarily be lower than before
- $F(\hat{q}(\hat{\theta}(y)|\mu, \Sigma) = F(\hat{q}'\mu, \hat{q}'\Sigma\hat{q}) \leq F_{MAX}$
- The loss due to estimation risk can be measured
- $L(q^*, \hat{q}) = \frac{F_{MAX} - F(\hat{q})}{F_{MAX}} > 0$

Estimation Risk

Jorion: Section II

- Utility function $EU(z) = F(\mu_z, \sigma_z^2)$
- Expected Portfolio Return: $\mu_z = q'\mu$
- Variance of Portfolio: $\sigma_z^2 = q'\Sigma q$
- q = The vector of investment *proportions*
- Choose q to maximize expected utility
  Subject to the constraint investment proportions sum to unity i.e. $\Sigma q_i = 1$
- Suppose we know $\theta = (\mu, \Sigma)$
- Optimal portfolio is $F_{MAX} = F(q'_*\mu, q'_*\Sigma q_*)$
  But we don't know θ
- So we choose q based on sample estimate $\hat{\theta}(y)$
- The expected utility will now necessarily be lower than before
- $F(\hat{q}(\hat{\theta}(y)|\mu, \Sigma) = F(\hat{q}'\mu, \hat{q}'\Sigma\hat{q}) \leq F_{MAX}$
- The loss due to estimation risk can be measured
- $L(q^*, \hat{q}) = \frac{F_{MAX} - F(\hat{q})}{F_{MAX}} > 0$
- Linear mapping from q to θ

Estimation Risk

Jorion: Section II

- Utility function $EU(z) = F(\mu_z, \sigma_z^2)$
- Expected Portfolio Return: $\mu_z = q' \mu$
- Variance of Portfolio: $\sigma_z^2 = q' \Sigma q$
- q = The vector of investment *proportions*
- Choose q to maximize expected utility
  Subject to the constraint investment proportions sum to unity i.e. $\Sigma q_i = 1$
- Suppose we know $\theta = (\mu, \Sigma)$
- Optimal portfolio is $F_{MAX} = F(q_*' \mu, q_*' \Sigma q_*)$
  But we don't know θ
- So we choose q based on sample estimate $\hat{\theta}(y)$
- The expected utility will now necessarily be lower than before
- $F(\hat{q}(\hat{\theta}(y)|\mu, \Sigma) = F(\hat{q}' \mu, \hat{q}' \Sigma \hat{q}) \leq F_{MAX}$
- The loss due to estimation risk can be measured
- $L(q^*, \hat{q}) = \frac{F_{MAX} - F(\hat{q})}{F_{MAX}} > 0$
- Linear mapping from q to θ
- $R(\theta) = \int L(\theta, \hat{\theta}) f(y|\theta) dy > 0$

Estimation Risk

Jorion: Section II

- Utility function $EU(z) = F(\mu_z, \sigma_z^2)$
- Expected Portfolio Return: $\mu_z = q'\mu$
- Variance of Portfolio: $\sigma_z^2 = q'\Sigma q$
- q = The vector of investment *proportions*
- Choose q to maximize expected utility
  Subject to the constraint investment proportions sum to unity i.e. $\Sigma q_i = 1$
- Suppose we know $\theta = (\mu, \Sigma)$
- Optimal portfolio is $F_{MAX} = F(q_*'\mu, q_*'\Sigma q_*)$
  But we don't know θ
- So we choose q based on sample estimate $\hat{\theta}(y)$
- The expected utility will now necessarily be lower than before
- $F(\hat{q}(\hat{\theta}(y)|\mu, \Sigma) = F(\hat{q}'\mu, \hat{q}'\Sigma\hat{q}) \leq F_{MAX}$
- The loss due to estimation risk can be measured
- $L(q^*, \hat{q}) = \frac{F_{MAX} - F(\hat{q})}{F_{MAX}} > 0$
- Linear mapping from q to θ
- $R(\theta) = \int L(\theta, \hat{\theta}) f(y|\theta) dy > 0$
- A reasonable minimum requirement for any estimator is admissibility

Estimation Risk

Jorion: Section II

- Utility function $EU(z) = F(\mu_z, \sigma_z^2)$
- Expected Portfolio Return: $\mu_z = q' \mu$
- Variance of Portfolio: $\sigma_z^2 = q' \Sigma q$
- q = The vector of investment *proportions*
- Choose q to maximize expected utility
  Subject to the constraint investment proportions sum to unity i.e. $\Sigma q_i = 1$
- Suppose we know $\theta = (\mu, \Sigma)$
- Optimal portfolio is $F_{MAX} = F(q_*' \mu, q_*' \Sigma q_*)$
  But we don't know θ
- So we choose q based on sample estimate $\hat{\theta}(y)$
- The expected utility will now necessarily be lower than before
- $F(\hat{q}(\hat{\theta}(y)|\mu, \Sigma) = F(\hat{q}' \mu, \hat{q}' \Sigma \hat{q}) \leq F_{MAX}$
- The loss due to estimation risk can be measured
- $L(q^*, \hat{q}) = \frac{F_{MAX} - F(\hat{q})}{F_{MAX}} > 0$
- Linear mapping from q to θ
- $R(\theta) = \int L(\theta, \hat{\theta}) f(y|\theta) dy > 0$
- A reasonable minimum requirement for any estimator is admissibility
- *Central thesis of this paper is that the usual sample mean is not admissible for portfolio estimation*

Stein Estimation

Jorion: Section III

- Let $\hat{\mu} =$ vector of N normal means, with covariance matrix assumed known.

Stein Estimation

Jorion: Section III

- Let $\hat{\mu} =$ vector of N normal means, with covariance matrix assumed known.
- $y_t \sim N(\hat{\mu}, \Sigma),\ \ t = 1..T$ are the stock returns.

Stein Estimation

Jorion: Section III

- Let $\hat{\mu} =$ vector of N normal means, with covariance matrix assumed known.
- $y_t \sim N(\hat{\mu}, \Sigma),\ \ t = 1..T$ are the stock returns.
- Stein's Paradox: If N $>$ 2, MLE $=$ vector of sample means $=$ is inadmissible.

Stein Estimation

Jorion: Section III

- Let $\hat{\mu}$ = vector of N normal means, with covariance matrix assumed known.
- $y_t \sim N(\hat{\mu}, \Sigma),\ \ t = 1..T$ are the stock returns.
- Stein's Paradox: If N $>$ 2, MLE = vector of sample means = is inadmissible.
- Inadmissibility is wrt quadratic loss $L(\mu, \hat{\mu}(y)) = (\mu - \hat{\mu}(y))'\sigma^{-1}(\mu - \hat{\mu}(y))$

Stein Estimation

Jorion: Section III

- Let $\hat{\mu}$ = vector of N normal means, with covariance matrix assumed known.
- $y_t \sim N(\hat{\mu}, \Sigma), \ t = 1..T$ are the stock returns.
- Stein's Paradox: If N $>$ 2, MLE = vector of sample means = is inadmissible.
- Inadmissibility is wrt quadratic loss $L(\mu, \hat{\mu}(y)) = (\mu - \hat{\mu}(y))'\sigma^{-1}(\mu - \hat{\mu}(y))$
- If $N \leq 2$, MLE is admissible & the best estimator.

Stein Estimation

Jorion: Section III

- Let $\hat{\mu}$ = vector of N normal means, with covariance matrix assumed known.
- $y_t \sim N(\hat{\mu}, \Sigma),\ \ t = 1..T$ are the stock returns.
- Stein's Paradox: If N $>$ 2, MLE = vector of sample means = is inadmissible.
- Inadmissibility is wrt quadratic loss $L(\mu, \hat{\mu}(y)) = (\mu - \hat{\mu}(y))'\sigma^{-1}(\mu - \hat{\mu}(y))$
- If $N \leq 2$, MLE is admissible & the best estimator.
- Shrinkage Estimator: $\hat{\mu}_{JS}(y) = (1 - \hat{w})Y + \hat{w}Y_0 1$, Y_0 is any point.

Stein Estimation

Jorion: Section III

- Let $\hat{\mu}$ = vector of N normal means, with covariance matrix assumed known.
- $y_t \sim N(\hat{\mu}, \Sigma),\ \ t = 1..T$ are the stock returns.
- Stein's Paradox: If N > 2, MLE = vector of sample means = is inadmissible.
- Inadmissibility is wrt quadratic loss $L(\mu, \hat{\mu}(y)) = (\mu - \hat{\mu}(y))'\sigma^{-1}(\mu - \hat{\mu}(y))$
- If $N \leq 2$, MLE is admissible & the best estimator.
- Shrinkage Estimator: $\hat{\mu}_{JS}(y) = (1 - \hat{w})Y + \hat{w}Y_0 1$, Y_0 is any point.
- $\hat{w} = min(1, \frac{(N-2)/T}{(Y - Y_0 1)'\Sigma^{-1}(Y - Y_0 1)})$

Stein Estimation

Jorion: Section III

- Let $\hat{\mu}$ = vector of N normal means, with covariance matrix assumed known.
- $y_t \sim N(\hat{\mu}, \Sigma),\ \ t = 1..T$ are the stock returns.
- Stein's Paradox: If N $>$ 2, MLE = vector of sample means = is inadmissible.
- Inadmissibility is wrt quadratic loss $L(\mu, \hat{\mu}(y)) = (\mu - \hat{\mu}(y))'\sigma^{-1}(\mu - \hat{\mu}(y))$
- If $N \leq 2$, MLE is admissible & the best estimator.
- Shrinkage Estimator: $\hat{\mu}_{JS}(y) = (1 - \hat{w})Y + \hat{w}Y_0 1$, Y_0 is any point.
- $\hat{w} = min(1, \frac{(N-2)/T}{(Y-Y_0 1)'\Sigma^{-1}(Y-Y_0 1)})$
- Maximum benefit when Y_0 closest to true value, though $\hat{\mu}_{JS}(y)$ always has lower risk than MLE

Stein Estimation

Jorion: Section III

- Let $\hat{\mu}$ = vector of N normal means, with covariance matrix assumed known.
- $y_t \sim N(\hat{\mu}, \Sigma),\ \ t = 1..T$ are the stock returns.
- Stein's Paradox: If N > 2, MLE = vector of sample means = is inadmissible.
- Inadmissibility is wrt quadratic loss $L(\mu, \hat{\mu}(y)) = (\mu - \hat{\mu}(y))'\sigma^{-1}(\mu - \hat{\mu}(y))$
- If $N \leq 2$, MLE is admissible & the best estimator.
- Shrinkage Estimator: $\hat{\mu}_{JS}(y) = (1 - \hat{w})Y + \hat{w}Y_0 1$, Y_0 is any point.
- $\hat{w} = min(1, \frac{(N-2)/T}{(Y-Y_0 1)'\Sigma^{-1}(Y-Y_0 1)})$
- Maximum benefit when Y_0 closest to true value, though $\hat{\mu}_{JS}(y)$ always has lower risk than MLE
- Stein Estimator has nonzero bias. It is nonlinear since $\hat{w}$ is itself a function of data.

Stein Estimation

Jorion: Section III

- Let $\hat{\mu}$ = vector of N normal means, with covariance matrix assumed known.
- $y_t \sim N(\hat{\mu}, \Sigma),\ \ t = 1..T$ are the stock returns.
- Stein's Paradox: If N $>$ 2, MLE = vector of sample means = is inadmissible.
- Inadmissibility is wrt quadratic loss $L(\mu, \hat{\mu}(y)) = (\mu - \hat{\mu}(y))'\sigma^{-1}(\mu - \hat{\mu}(y))$
- If $N \leq 2$, MLE is admissible & the best estimator.
- Shrinkage Estimator: $\hat{\mu}_{JS}(y) = (1 - \hat{w})Y + \hat{w}Y_0 1$, Y_0 is any point.
- $\hat{w} = min(1, \frac{(N-2)/T}{(Y-Y_0 1)'\Sigma^{-1}(Y-Y_0 1)})$
- Maximum benefit when Y_0 closest to true value, though $\hat{\mu}_{JS}(y)$ always has lower risk than MLE
- Stein Estimator has nonzero bias. It is nonlinear since $\hat{w}$ is itself a function of data.
- Superiority of Stein Estimator result is a *startling* result

Stein Estimation

Jorion: Section III

- Let $\hat{\mu}$ = vector of N normal means, with covariance matrix assumed known.
- $y_t \sim N(\hat{\mu}, \Sigma), \;\; t = 1..T$ are the stock returns.
- Stein's Paradox: If N > 2, MLE = vector of sample means = is inadmissible.
- Inadmissibility is wrt quadratic loss $L(\mu, \hat{\mu}(y)) = (\mu - \hat{\mu}(y))'\sigma^{-1}(\mu - \hat{\mu}(y))$
- If $N \leq 2$, MLE is admissible & the best estimator.
- Shrinkage Estimator: $\hat{\mu}_{JS}(y) = (1 - \hat{w})Y + \hat{w}Y_0 1$, Y_0 is any point.
- $\hat{w} = min(1, \frac{(N-2)/T}{(Y - Y_0 1)'\Sigma^{-1}(Y - Y_0 1)})$
- Maximum benefit when Y_0 closest to true value, though $\hat{\mu}_{JS}(y)$ always has lower risk than MLE
- Stein Estimator has nonzero bias. It is nonlinear since $\hat{w}$ is itself a function of data.
- Superiority of Stein Estimator result is a *startling* result
- Stein Estimator is one of the *most important statistical idea of the decade* : Dennis Lindley

Stein Estimation

Jorion: Section III

- Let $\hat{\mu}$ = vector of N normal means, with covariance matrix assumed known.
- $y_t \sim N(\hat{\mu}, \Sigma),\ \ t = 1..T$ are the stock returns.
- Stein's Paradox: If N > 2, MLE = vector of sample means = is inadmissible.
- Inadmissibility is wrt quadratic loss $L(\mu, \hat{\mu}(y)) = (\mu - \hat{\mu}(y))'\sigma^{-1}(\mu - \hat{\mu}(y))$
- If $N \leq 2$, MLE is admissible & the best estimator.
- Shrinkage Estimator: $\hat{\mu}_{JS}(y) = (1 - \hat{w})Y + \hat{w}Y_0 1$, Y_0 is any point.
- $\hat{w} = min(1, \frac{(N-2)/T}{(Y-Y_0 1)'\Sigma^{-1}(Y-Y_0 1)})$
- Maximum benefit when Y_0 closest to true value, though $\hat{\mu}_{JS}(y)$ always has lower risk than MLE
- Stein Estimator has nonzero bias. It is nonlinear since $\hat{w}$ is itself a function of data.
- Superiority of Stein Estimator result is a *startling* result
- Stein Estimator is one of the *most important statistical idea of the decade* : Dennis Lindley
- What about Normality assumption ?

Stein Estimation

Jorion: Section III

- Let $\hat{\mu}$ = vector of N normal means, with covariance matrix assumed known.
- $y_t \sim N(\hat{\mu}, \Sigma),\ \ t = 1..T$ are the stock returns.
- Stein's Paradox: If N > 2, MLE = vector of sample means = is inadmissible.
- Inadmissibility is wrt quadratic loss $L(\mu, \hat{\mu}(y)) = (\mu - \hat{\mu}(y))'\sigma^{-1}(\mu - \hat{\mu}(y))$
- If $N \leq 2$, MLE is admissible & the best estimator.
- Shrinkage Estimator: $\hat{\mu}_{JS}(y) = (1 - \hat{w})Y + \hat{w}Y_0 1$, Y_0 is any point.
- $\hat{w} = min(1, \frac{(N-2)/T}{(Y-Y_0 1)'\Sigma^{-1}(Y-Y_0 1)})$
- Maximum benefit when Y_0 closest to true value, though $\hat{\mu}_{JS}(y)$ always has lower risk than MLE
- Stein Estimator has nonzero bias. It is nonlinear since $\hat{w}$ is itself a function of data.
- Superiority of Stein Estimator result is a *startling* result
- Stein Estimator is one of the *most important statistical idea of the decade* : Dennis Lindley
- What about Normality assumption ?
- Normality is NOT critical to Stein estimate

Stein Estimation

Jorion: Section III

- Let $\hat{\mu}$ = vector of N normal means, with covariance matrix assumed known.
- $y_t \sim N(\hat{\mu}, \Sigma), \; t = 1..T$ are the stock returns.
- Stein's Paradox: If N > 2, MLE = vector of sample means = is inadmissible.
- Inadmissibility is wrt quadratic loss $L(\mu, \hat{\mu}(y)) = (\mu - \hat{\mu}(y))' \sigma^{-1} (\mu - \hat{\mu}(y))$
- If $N \leq 2$, MLE is admissible & the best estimator.
- Shrinkage Estimator: $\hat{\mu}_{JS}(y) = (1 - \hat{w})Y + \hat{w} Y_0 1$, Y_0 is any point.
- $\hat{w} = min(1, \frac{(N-2)/T}{(Y - Y_0 1)' \Sigma^{-1} (Y - Y_0 1)})$
- Maximum benefit when Y_0 closest to true value, though $\hat{\mu}_{JS}(y)$ always has lower risk than MLE
- Stein Estimator has nonzero bias. It is nonlinear since $\hat{w}$ is itself a function of data.
- Superiority of Stein Estimator result is a *startling* result
- Stein Estimator is one of the *most important statistical idea of the decade* : Dennis Lindley
- What about Normality assumption ?
- Normality is NOT critical to Stein estimate
- What about assumption of known σ^2 ?

Stein Estimation

Jorion: Section III

- Let $\hat{\mu}$ = vector of N normal means, with covariance matrix assumed known.
- $y_t \sim N(\hat{\mu}, \Sigma), \ \ t = 1..T$ are the stock returns.
- Stein's Paradox: If N > 2, MLE = vector of sample means = is inadmissible.
- Inadmissibility is wrt quadratic loss $L(\mu, \hat{\mu}(y)) = (\mu - \hat{\mu}(y))' \sigma^{-1} (\mu - \hat{\mu}(y))$
- If $N \leq 2$, MLE is admissible & the best estimator.
- Shrinkage Estimator: $\hat{\mu}_{JS}(y) = (1 - \hat{w})Y + \hat{w} Y_0 1$, Y_0 is any point.
- $\hat{w} = min(1, \frac{(N-2)/T}{(Y - Y_0 1)' \Sigma^{-1} (Y - Y_0 1)})$
- Maximum benefit when Y_0 closest to true value, though $\hat{\mu}_{JS}(y)$ always has lower risk than MLE
- Stein Estimator has nonzero bias. It is nonlinear since $\hat{w}$ is itself a function of data.
- Superiority of Stein Estimator result is a *startling* result
- Stein Estimator is one of the *most important statistical idea of the decade* : Dennis Lindley
- What about Normality assumption ?
- Normality is NOT critical to Stein estimate
- What about assumption of known σ^2 ?
- In real-life (finance), standard deviations of stocks tend to not change as much (Merton)

Stein Estimation

Jorion: Section III

- Let $\hat{\mu}$ = vector of N normal means, with covariance matrix assumed known.
- $y_t \sim N(\hat{\mu}, \Sigma)$, $t = 1..T$ are the stock returns.
- Stein's Paradox: If N $>$ 2, MLE = vector of sample means = is inadmissible.
- Inadmissibility is wrt quadratic loss $L(\mu, \hat{\mu}(y)) = (\mu - \hat{\mu}(y))'\sigma^{-1}(\mu - \hat{\mu}(y))$
- If $N \leq 2$, MLE is admissible & the best estimator.
- Shrinkage Estimator: $\hat{\mu}_{JS}(y) = (1 - \hat{w})Y + \hat{w}Y_0 1$, Y_0 is any point.
- $\hat{w} = min(1, \frac{(N-2)/T}{(Y - Y_0 1)'\Sigma^{-1}(Y - Y_0 1)})$
- Maximum benefit when Y_0 closest to true value, though $\hat{\mu}_{JS}(y)$ always has lower risk than MLE
- Stein Estimator has nonzero bias. It is nonlinear since $\hat{w}$ is itself a function of data.
- Superiority of Stein Estimator result is a *startling* result
- Stein Estimator is one of the *most important statistical idea of the decade* : Dennis Lindley
- What about Normality assumption ?
- Normality is NOT critical to Stein estimate
- What about assumption of known σ^2 ?
- In real-life (finance), standard deviations of stocks tend to not change as much (Merton)
- Standard deviations of stocks tend to be larger than the means

Stein Estimation

Jorion: Section III

- Let $\hat{\mu}$ = vector of N normal means, with covariance matrix assumed known.
- $y_t \sim N(\hat{\mu}, \Sigma), \;\; t = 1..T$ are the stock returns.
- Stein's Paradox: If N > 2, MLE = vector of sample means = is inadmissible.
- Inadmissibility is wrt quadratic loss $L(\mu, \hat{\mu}(y)) = (\mu - \hat{\mu}(y))' \sigma^{-1} (\mu - \hat{\mu}(y))$
- If $N \leq 2$, MLE is admissible & the best estimator.
- Shrinkage Estimator: $\hat{\mu}_{JS}(y) = (1 - \hat{w})Y + \hat{w} Y_0 1$, Y_0 is any point.
- $\hat{w} = min(1, \frac{(N-2)/T}{(Y - Y_0 1)' \Sigma^{-1} (Y - Y_0 1)})$
- Maximum benefit when Y_0 closest to true value, though $\hat{\mu}_{JS}(y)$ always has lower risk than MLE
- Stein Estimator has nonzero bias. It is nonlinear since $\hat{w}$ is itself a function of data.
- Superiority of Stein Estimator result is a *startling* result
- Stein Estimator is one of the *most important statistical idea of the decade* : Dennis Lindley
- What about Normality assumption ?
- Normality is NOT critical to Stein estimate
- What about assumption of known σ^2 ?
- In real-life (finance), standard deviations of stocks tend to not change as much (Merton)
- Standard deviations of stocks tend to be larger than the means
- Stein Estimator instead of sample mean is already a significant improvement

Empirical Bayes

Jorion: Section IV

- Goal: Find the predictive distribution of future returns, conditioned on the prior, data, covariance matrix & scale factor

Empirical Bayes

Jorion: Section IV

- Goal: Find the predictive distribution of future returns, conditioned on the prior, data, covariance matrix & scale factor
- Data $y = (y_1, y_2, \ldots y_t)$

Empirical Bayes

Jorion: Section IV

- Goal: Find the predictive distribution of future returns, conditioned on the prior, data, covariance matrix & scale factor
- Data $y = (y_1, y_2, \ldots y_t)$
- Covariance Matrix Σ

Empirical Bayes

Jorion: Section IV

- Goal: Find the predictive distribution of future returns, conditioned on the prior, data, covariance matrix & scale factor
- Data $y = (y_1, y_2, \ldots y_t)$
- Covariance Matrix Σ
- Scale Factor λ

Empirical Bayes

Jorion: Section IV

- Goal: Find the predictive distribution of future returns, conditioned on the prior, data, covariance matrix & scale factor
- Data $y = (y_1, y_2, \ldots y_t)$
- Covariance Matrix Σ
- Scale Factor λ
- Empirical Bayes, $\eta, \lambda =$ prior grand mean & prior precision , are both estimated from the data (could alternately use prior knowledge)

Empirical Bayes

Jorion: Section IV

- Goal: Find the predictive distribution of future returns, conditioned on the prior, data, covariance matrix & scale factor
- Data $y = (y_1, y_2, \ldots y_t)$
- Covariance Matrix Σ
- Scale Factor λ
- Empirical Bayes, $\eta, \lambda =$ prior grand mean & prior precision , are both estimated from the data (could alternately use prior knowledge)
- $p(r|y, \Sigma, \lambda) = \int \int p(r, \mu, \eta | y, \Sigma, \lambda) d\mu d\eta$

Empirical Bayes

Jorion: Section IV

- Goal: Find the predictive distribution of future returns, conditioned on the prior, data, covariance matrix & scale factor
- Data $y = (y_1, y_2, \ldots y_t)$
- Covariance Matrix Σ
- Scale Factor λ
- Empirical Bayes, $\eta, \lambda =$ prior grand mean & prior precision , are both estimated from the data (could alternately use prior knowledge)
- $p(r|y, \Sigma, \lambda) = \int \int p(r, \mu, \eta|y, \Sigma, \lambda) d\mu d\eta$
- The joint density of r,μ and η is
  $p(r, \mu, \eta|y, \Sigma, \lambda) = p(r|\mu, \eta, \Sigma, \lambda) \cdot p(\mu, \eta|y, \Sigma, \lambda)$
  $\propto p(r|\mu, \eta) \cdot f(y|\mu, \eta) \cdot p(\mu|\eta, \lambda) \cdot p(\eta)$

EMPIRICAL BAYES

JORION: SECTION IV

- Goal: Find the predictive distribution of future returns, conditioned on the prior, data, covariance matrix & scale factor
- Data $y = (y_1, y_2, \dots y_t)$
- Covariance Matrix Σ
- Scale Factor λ
- Empirical Bayes, $\eta, \lambda =$ prior grand mean & prior precision , are both estimated from the data (could alternately use prior knowledge)
- $p(r|y, \Sigma, \lambda) = \int \int p(r, \mu, \eta|y, \Sigma, \lambda) d\mu d\eta$
- The joint density of r,μ and η is
  $p(r, \mu, \eta|y, \Sigma, \lambda) = p(r|\mu, \eta, \Sigma, \lambda) \cdot p(\mu, \eta|y, \Sigma, \lambda)$
  $\propto p(r|\mu, \eta) \cdot f(y|\mu, \eta) \cdot p(\mu|\eta, \lambda) \cdot p(\eta)$
- We assume the returns are normally distributed. The conditional likelihood of y is
  $f(y_t|\mu, \Sigma) \propto exp((-1/2)(y_t - \mu)'\Sigma^{-1}(y_t - \mu))$

Empirical Bayes

Jorion: Section IV

- Goal: Find the predictive distribution of future returns, conditioned on the prior, data, covariance matrix & scale factor
- Data $y = (y_1, y_2, \ldots y_t)$
- Covariance Matrix Σ
- Scale Factor λ
- Empirical Bayes, $\eta, \lambda =$ prior grand mean & prior precision , are both estimated from the data (could alternately use prior knowledge)
- $p(r|y, \Sigma, \lambda) = \int\int p(r, \mu, \eta|y, \Sigma, \lambda) d\mu d\eta$
- The joint density of r,μ and η is
  $p(r, \mu, \eta|y, \Sigma, \lambda) = p(r|\mu, \eta, \Sigma, \lambda) \cdot p(\mu, \eta|y, \Sigma, \lambda)$
  $\propto p(r|\mu, \eta) \cdot f(y|\mu, \eta) \cdot p(\mu|\eta, \lambda) \cdot p(\eta)$
- We assume the returns are normally distributed. The conditional likelihood of y is
  $f(y_t|\mu, \Sigma) \propto exp((-1/2)(y_t - \mu)'\Sigma^{-1}(y_t - \mu))$
- Density of μ is given by the conjugate Prior $p(\mu|\eta, \lambda) \propto \; exp((-1/2)(\mu - 1\eta)'\lambda\Sigma^{-1}(\mu - 1\eta))$

Empirical Bayes

Jorion: Section IV

- Goal: Find the predictive distribution of future returns, conditioned on the prior, data, covariance matrix & scale factor
- Data $y = (y_1, y_2, \ldots y_t)$
- Covariance Matrix Σ
- Scale Factor λ
- Empirical Bayes, $\eta, \lambda =$ prior grand mean & prior precision , are both estimated from the data (could alternately use prior knowledge)
- $p(r|y, \Sigma, \lambda) = \int \int p(r, \mu, \eta|y, \Sigma, \lambda) d\mu d\eta$
- The joint density of r,μ and η is
  $p(r, \mu, \eta|y, \Sigma, \lambda) = p(r|\mu, \eta, \Sigma, \lambda) \cdot p(\mu, \eta|y, \Sigma, \lambda)$
  $\propto p(r|\mu, \eta) \cdot f(y|\mu, \eta) \cdot p(\mu|\eta, \lambda) \cdot p(\eta)$
- We assume the returns are normally distributed. The conditional likelihood of y is
  $f(y_t|\mu, \Sigma) \propto exp((-1/2)(y_t - \mu)'\Sigma^{-1}(y_t - \mu))$
- Density of μ is given by the conjugate Prior $p(\mu|\eta, \lambda) \propto \ exp((-1/2)(\mu - 1\eta)'\lambda\Sigma^{-1}(\mu - 1\eta))$
- The λ parameter measures the tightness of the prior

Empirical Bayes

Jorion: Section IV

- Goal: Find the predictive distribution of future returns, conditioned on the prior, data, covariance matrix & scale factor
- Data $y = (y_1, y_2, \ldots y_t)$
- Covariance Matrix Σ
- Scale Factor λ
- Empirical Bayes, $\eta, \lambda =$ prior grand mean & prior precision , are both estimated from the data (could alternately use prior knowledge)
- $p(r|y, \Sigma, \lambda) = \int \int p(r, \mu, \eta|y, \Sigma, \lambda) d\mu d\eta$
- The joint density of r,μ and η is
  $p(r, \mu, \eta|y, \Sigma, \lambda) = p(r|\mu, \eta, \Sigma, \lambda) \cdot p(\mu, \eta|y, \Sigma, \lambda)$
  $\propto p(r|\mu, \eta) \cdot f(y|\mu, \eta) \cdot p(\mu|\eta, \lambda) \cdot p(\eta)$
- We assume the returns are normally distributed. The conditional likelihood of y is
  $f(y_t|\mu, \Sigma) \propto exp((-1/2)(y_t - \mu)'\Sigma^{-1}(y_t - \mu))$
- Density of μ is given by the conjugate Prior $p(\mu|\eta, \lambda) \propto \; exp((-1/2)(\mu - 1\eta)'\lambda\Sigma^{-1}(\mu - 1\eta))$
- The λ parameter measures the tightness of the prior
- Why doesn't Sample mean work ?
  Special Case $\lambda = 0$, corresponds to Diffuse Prior
  Improper Prior, won't integrate to 1
  In this special case Bayes rule won't be admissible (Berger: Statistical Decision Theory)

Empirical Bayes

Jorion: Section IV

- Goal: Find the predictive distribution of future returns, conditioned on the prior, data, covariance matrix & scale factor
- Data $y = (y_1, y_2, \ldots y_t)$
- Covariance Matrix Σ
- Scale Factor λ
- Empirical Bayes, $\eta, \lambda =$ prior grand mean & prior precision , are both estimated from the data (could alternately use prior knowledge)
- $p(r|y, \Sigma, \lambda) = \int \int p(r, \mu, \eta|y, \Sigma, \lambda) d\mu d\eta$
- The joint density of r,μ and η is
  $p(r, \mu, \eta|y, \Sigma, \lambda) = p(r|\mu, \eta, \Sigma, \lambda) \cdot p(\mu, \eta|y, \Sigma, \lambda)$
  $\propto p(r|\mu, \eta) \cdot f(y|\mu, \eta) \cdot p(\mu|\eta, \lambda) \cdot p(\eta)$
- We assume the returns are normally distributed. The conditional likelihood of y is
  $f(y_t|\mu, \Sigma) \propto exp((-1/2)(y_t - \mu)'\Sigma^{-1}(y_t - \mu))$
- Density of μ is given by the conjugate Prior $p(\mu|\eta, \lambda) \propto \; exp((-1/2)(\mu - 1\eta)'\lambda\Sigma^{-1}(\mu - 1\eta))$
- The λ parameter measures the tightness of the prior
- Why doesn't Sample mean work ?
  Special Case $\lambda = 0$, corresponds to Diffuse Prior
  Improper Prior, won't integrate to 1
  In this special case Bayes rule won't be admissible (Berger: Statistical Decision Theory)
- Another interpretation of above prior: Means μ vary around the common grand mean η

Empirical Bayes

Jorion: Section IV: Derivation

- $p(r, \mu, \eta | y, \Sigma, \lambda) \propto exp((-1/2)G)$
  $G = (r-\mu)'\Sigma^{-1}(r-\mu) + \sum_{t=1}^{T}(y_t-\mu)\Sigma^{-1}(r-\mu) + (\mu - 1\eta)'\lambda\Sigma^{-1}(\mu - 1\eta))$

Empirical Bayes

Jorion: Section IV: Derivation

- $p(r, \mu, \eta | y, \Sigma, \lambda) \propto exp((-1/2)G)$
  $G = (r-\mu)'\Sigma^{-1}(r-\mu) + \sum_{t=1}^{T}(y_t - \mu)\Sigma^{-1}(r-\mu) + (\mu - 1\eta)'\lambda\Sigma^{-1}(\mu - 1\eta))$
- After integration over η, μ, we find the predictive density is Normal with mean vector & covariance matrix
  Mean Vector $E(r) = (1-w)Y + w1Y_0$
  w = shrinkage factor = $\frac{\lambda}{T+\lambda}$
  Y = vector of sample means = $\frac{\sum_{t=1}^{T} y_t}{T}$
  Y_0 = grand mean = x'Y
  x' = weights of the minimum variance portfolio = $\frac{1'\Sigma^{-1}}{1'\Sigma^{-1}1}$
  Covariance matrix $V = \Sigma(1 + \frac{1}{T+\lambda}) + \frac{\lambda}{T(T+1+\lambda)}\frac{11'}{1'\Sigma^{-1}1}$

Empirical Bayes

Jorion: Section IV: Derivation

- $p(r, \mu, \eta | y, \Sigma, \lambda) \propto exp((-1/2)G)$
  $G = (r-\mu)'\Sigma^{-1}(r-\mu) + \sum_{t=1}^{T}(y_t-\mu)\Sigma^{-1}(r-\mu) + (\mu - 1\eta)'\lambda\Sigma^{-1}(\mu - 1\eta))$
- After integration over η, μ, we find the predictive density is Normal with mean vector & covariance matrix
  Mean Vector $E(r) = (1-w)Y + w1Y_0$
  w = shrinkage factor = $\frac{\lambda}{T+\lambda}$
  Y = vector of sample means = $\frac{\sum_{t=1}^{T} y_t}{T}$
  Y_0 = grand mean = x'Y
  x' = weights of the minimum variance portfolio = $\frac{1'\Sigma^{-1}}{1'\Sigma^{-1}1}$
  Covariance matrix $V = \Sigma(1 + \frac{1}{T+\lambda}) + \frac{\lambda}{T(T+1+\lambda)} \frac{11'}{1'\Sigma^{-1}1}$
- Covariance Matrix Interpretation:
  Σ : variation of y_t around the mean μ
  $\frac{\Sigma}{T+\lambda}$: uncertainity of the sample average Y
  Third term is the uncertainity in the common factor

Empirical Bayes

Jorion: Section IV: Derivation

- $p(r, \mu, \eta | y, \Sigma, \lambda) \propto exp((-1/2)G)$
  $G = (r-\mu)'\Sigma^{-1}(r-\mu) + \sum_{t=1}^{T}(y_t - \mu)\Sigma^{-1}(r-\mu) + (\mu - 1\eta)'\lambda\Sigma^{-1}(\mu - 1\eta))$
- After integration over η, μ, we find the predictive density is Normal with mean vector & covariance matrix
  Mean Vector $E(r) = (1-w)Y + w1Y_0$
  w = shrinkage factor = $\frac{\lambda}{T+\lambda}$
  Y = vector of sample means = $\frac{\sum_{t=1}^{T} y_t}{T}$
  Y_0 = grand mean = x'Y
  x' = weights of the minimum variance portfolio = $\frac{1'\Sigma^{-1}}{1'\Sigma^{-1}1}$
  Covariance matrix $V = \Sigma(1 + \frac{1}{T+\lambda}) + \frac{\lambda}{T(T+1+\lambda)} \frac{11'}{1'\Sigma^{-1}1}$
- Covariance Matrix Interpretation:
  Σ : variation of y_t around the mean μ
  $\frac{\Sigma}{T+\lambda}$: uncertainity of the sample average Y
  Third term is the uncertainity in the common factor
- Large T: $w \to 0$, $r \to Y$, $V \to \Sigma$

Empirical Bayes

Jorion: Section IV: Derivation

- $p(r, \mu, \eta | y, \Sigma, \lambda) \propto exp((-1/2)G)$
  $G = (r-\mu)'\Sigma^{-1}(r-\mu) + \sum_{t=1}^{T}(y_t - \mu)\Sigma^{-1}(r-\mu) + (\mu - 1\eta)'\lambda\Sigma^{-1}(\mu - 1\eta))$
- After integration over η, μ, we find the predictive density is Normal with mean vector & covariance matrix
  Mean Vector $E(r) = (1-w)Y + w1Y_0$
  w = shrinkage factor = $\frac{\lambda}{T+\lambda}$
  Y = vector of sample means = $\frac{\sum_{t=1}^{T} y_t}{T}$
  Y_0 = grand mean = x'Y
  x' = weights of the minimum variance portfolio = $\frac{1'\Sigma^{-1}}{1'\Sigma^{-1}1}$
  Covariance matrix $V = \Sigma(1 + \frac{1}{T+\lambda}) + \frac{\lambda}{T(T+1+\lambda)} \frac{11'}{1'\Sigma^{-1}1}$
- Covariance Matrix Interpretation:
  Σ : variation of y_t around the mean μ
  $\frac{\Sigma}{T+\lambda}$: uncertainity of the sample average Y
  Third term is the uncertainity in the common factor
- Large T: $w \to 0, r \to Y, V \to \Sigma$
- Small λ : $w \to 0, E(r) \to Y, V \to \Sigma(1 + (1/T))$

Empirical Bayes

Jorion: Section IV: Derivation

- $p(r, \mu, \eta | y, \Sigma, \lambda) \propto exp((-1/2)G)$
  $G = (r - \mu)'\Sigma^{-1}(r - \mu) + \sum_{t=1}^{T}(y_t - \mu)\Sigma^{-1}(r - \mu) + (\mu - 1\eta)'\lambda\Sigma^{-1}(\mu - 1\eta))$
- After integration over η, μ, we find the predictive density is Normal with mean vector & covariance matrix
  Mean Vector $E(r) = (1 - w)Y + w1Y_0$
  w = shrinkage factor = $\frac{\lambda}{T+\lambda}$
  Y = vector of sample means = $\frac{\sum_{t=1}^{T} y_t}{T}$
  Y_0 = grand mean = x'Y
  x' = weights of the minimum variance portfolio = $\frac{1'\Sigma^{-1}}{1'\Sigma^{-1}1}$
  Covariance matrix $V = \Sigma(1 + \frac{1}{T+\lambda}) + \frac{\lambda}{T(T+1+\lambda)} \frac{11'}{1'\Sigma^{-1}1}$
- Covariance Matrix Interpretation:
  Σ : variation of y_t around the mean μ
  $\frac{\Sigma}{T+\lambda}$: uncertainity of the sample average Y
  Third term is the uncertainity in the common factor
- Large T: $w \to 0, r \to Y, V \to \Sigma$
- Small λ : $w \to 0, E(r) \to Y, V \to \Sigma(1 + (1/T))$
- Large λ : $w \to 1, E(r) \to 1Y_0, V \to \frac{11'}{T(1'\Sigma^{-1}1)}$

Empirical Bayes

Jorion: Section IV: Simplified Working Formulas

- $Y =$ vector of sample means of returns

Empirical Bayes

Jorion: Section IV: Simplified Working Formulas

- Y = vector of sample means of returns
- S = sample covariance matrix

Empirical Bayes

Jorion: Section IV: Simplified Working Formulas

- Y = vector of sample means of returns
- S = sample covariance matrix
- $\hat{\Sigma} = \frac{T-1}{T-N-2}S$

Empirical Bayes

Jorion: Section IV: Simplified Working Formulas

- Y = vector of sample means of returns
- S = sample covariance matrix
- $\hat{\Sigma} = \frac{T-1}{T-N-2} S$
- $\hat{w} = \frac{N+2}{(N+2)+(Y-Y_0 1)' T \hat{\Sigma}^{-1} (Y-Y_0 1)}$

Empirical Bayes

Jorion: Section IV: Simplified Working Formulas

- Y = vector of sample means of returns
- S = sample covariance matrix
- $\hat{\Sigma} = \frac{T-1}{T-N-2} S$
- $\hat{w} = \frac{N+2}{(N+2)+(Y-Y_0 1)' T \hat{\Sigma}^{-1} (Y-Y_0 1)}$
- $Y_0 = \frac{1' \hat{\Sigma}^{-1}}{1' \hat{\Sigma}^{-1} 1} Y$

Empirical Bayes

Jorion: Section IV: Simplified Working Formulas

- Y = vector of sample means of returns
- S = sample covariance matrix
- $\hat{\Sigma} = \frac{T-1}{T-N-2} S$
- $\hat{w} = \frac{N+2}{(N+2)+(Y-Y_0 1)' T \hat{\Sigma}^{-1}(Y-Y_0 1)}$
- $Y_0 = \frac{1' \hat{\Sigma}^{-1}}{1' \hat{\Sigma}^{-1} 1} Y$
- $\hat{\lambda} = \frac{\hat{w} T}{1-\hat{w}}$

Empirical Bayes

Jorion: Section IV: Simplified Working Formulas

- Y = vector of sample means of returns
- S = sample covariance matrix
- $\hat{\Sigma} = \frac{T-1}{T-N-2} S$
- $\hat{w} = \frac{N+2}{(N+2)+(Y-Y_0 1)' T \hat{\Sigma}^{-1} (Y-Y_0 1)}$
- $Y_0 = \frac{1' \hat{\Sigma}^{-1}}{1' \hat{\Sigma}^{-1} 1} Y$
- $\hat{\lambda} = \frac{\hat{w} T}{1-\hat{w}}$
- $E(r) = (1-\hat{w})Y + \hat{w} 1 Y_0$

Empirical Bayes

Jorion: Section IV: Simplified Working Formulas

- Y = vector of sample means of returns
- S = sample covariance matrix
- $\hat{\Sigma} = \frac{T-1}{T-N-2} S$
- $\hat{w} = \frac{N+2}{(N+2)+(Y-Y_0 1)' T \hat{\Sigma}^{-1} (Y-Y_0 1)}$
- $Y_0 = \frac{1' \hat{\Sigma}^{-1}}{1' \hat{\Sigma}^{-1} 1} Y$
- $\hat{\lambda} = \frac{\hat{w} T}{1-\hat{w}}$
- $E(r) = (1-\hat{w})Y + \hat{w} 1 Y_0$
- $\hat{V}(r) = \hat{\Sigma} \cdot (1 + \frac{1}{T+\hat{\lambda}}) + \frac{\hat{\lambda}}{T(T+1+\hat{\lambda})} \frac{11'}{1' \Sigma^{-1} 1}$

My Project

Krishnan: Markowitz for FAANG stocks

- Consider Daily returns for Facebook, Amazon, Apple, Netflix & Google for the year 2020

My Project

Krishnan: Markowitz for FAANG stocks

- Consider Daily returns for Facebook, Amazon, Apple, Netflix & Google for the year 2020
- 253 Trading Days in 2020, using 223 Days of returns

My Project

Krishnan: Markowitz for FAANG stocks

- Consider Daily returns for Facebook, Amazon, Apple, Netflix & Google for the year 2020
- 253 Trading Days in 2020, using 223 Days of returns
- We construct a Markowitz portfolio allocation with Certainity Equivalence

My Project

Krishnan: Markowitz for FAANG stocks

- Consider Daily returns for Facebook, Amazon, Apple, Netflix & Google for the year 2020
- 253 Trading Days in 2020, using 223 Days of returns
- We construct a Markowitz portfolio allocation with Certainity Equivalence
- We construct a Markowitz portfolio allocation with Bayes-Stein estimates.

My Project

Krishnan: Markowitz for FAANG stocks

- Consider Daily returns for Facebook, Amazon, Apple, Netflix & Google for the year 2020
- 253 Trading Days in 2020, using 223 Days of returns
- We construct a Markowitz portfolio allocation with Certainity Equivalence
- We construct a Markowitz portfolio allocation with Bayes-Stein estimates.
- Compare the allocations for the remaining 30 Days

Conclusion

Future Directions

- We looked at 2 different estimators. There are many more...

Conclusion

Future Directions

- We looked at 2 different estimators. There are many more...
- Certainity Equivalence(Markowitz): Y, S - use sample mean, sample cov

Conclusion

Future Directions

- We looked at 2 different estimators. There are many more...
- Certainity Equivalence(Markowitz): Y, S - use sample mean, sample cov
- Bayes Diffuse Prior(Brown): $\lambda = 0$, Y, $\hat{V} = \hat{\Sigma} \cdot (1 + 1/T)$

Conclusion

Future Directions

- We looked at 2 different estimators. There are many more...
- Certainity Equivalence(Markowitz): Y, S - use sample mean, sample cov
- Bayes Diffuse Prior(Brown): $\lambda = 0$, Y, $\hat{V} = \hat{\Sigma} \cdot (1 + 1/T)$
- Minimum Variance(Jobson): $w = 1, 1Y_0$, $\hat{V} = \Sigma + (1/T)11'(1/1'\Sigma^{-1}1)$
  Based entirely on grand mean
  Matrix added to Σ reflects uncertainity in this grand mean.

Conclusion

Future Directions

- We looked at 2 different estimators. There are many more...
- Certainity Equivalence(Markowitz): Y, S - use sample mean, sample cov
- Bayes Diffuse Prior(Brown): $\lambda = 0$, Y, $\hat{V} = \hat{\Sigma} \cdot (1 + 1/T)$
- Minimum Variance(Jobson): $w = 1, 1Y_0$, $\hat{V} = \Sigma + (1/T)11'(1/1'\Sigma^{-1}1)$
  Based entirely on grand mean
  Matrix added to Σ reflects uncertainity in this grand mean.
- Bayes-Stein Estimate(Jorion): compromise between grand mean & sample mean

Conclusion

Future Directions

- We looked at 2 different estimators. There are many more...
- Certainity Equivalence(Markowitz): Y, S - use sample mean, sample cov
- Bayes Diffuse Prior(Brown): $\lambda = 0$, Y, $\hat{V} = \hat{\Sigma} \cdot (1 + 1/T)$
- Minimum Variance(Jobson): $w = 1, 1Y_0$, $\hat{V} = \Sigma + (1/T)11'(1/1'\Sigma^{-1}1)$
  Based entirely on grand mean
  Matrix added to Σ reflects uncertainity in this grand mean.
- Bayes-Stein Estimate(Jorion): compromise between grand mean & sample mean
- Unified View of Shrinkage: Stein Estimator if dim(response) > 2
  vs Ridge/Lasso/ElasticNet if dim(predictors) $>$ p (model selection)

Conclusion

Future Directions

- We looked at 2 different estimators. There are many more...
- Certainity Equivalence(Markowitz): Y, S - use sample mean, sample cov
- Bayes Diffuse Prior(Brown): $\lambda = 0$, Y, $\hat{V} = \hat{\Sigma} \cdot (1 + 1/T)$
- Minimum Variance(Jobson): $w = 1, 1Y_0$, $\hat{V} = \Sigma + (1/T)11'(1/1'\Sigma^{-1}1)$
  Based entirely on grand mean
  Matrix added to Σ reflects uncertainity in this grand mean.
- Bayes-Stein Estimate(Jorion): compromise between grand mean & sample mean
- Unified View of Shrinkage: Stein Estimator if dim(response) > 2
  vs Ridge/Lasso/ElasticNet if dim(predictors) $>$ p (model selection)
- Ledger (Peer Reviewed Blockchain Journal) & BTC Assets