Mgmt 237M2 **Statistical Arbitrage**

Lecture 04: Covariance Matrices

Professor Olivier Ledoit

University of California Los Angeles
Anderson School of Management
Master of Financial Engineering

Fall 2016

Inputs into Markowitz Optimization

- Expected Returns
  - Do not use the sample mean!
  - *Use external information*: Analysts, anomalies...
- Covariance Matrix
  - Do not use the sample covariance matrix!
  - Data-driven solution: no external info needed
  - More elegant

Sample Covariance Matrix

- Obvious problem:

When more assets than observations: not invertible!

$$w = (1-\lambda)\frac{\Sigma^{-1}1}{1'\Sigma^{-1}1} + \lambda\frac{\Sigma^{-1}\mu}{1'\Sigma^{-1}\mu}$$

Russell 3000, daily data $\rightarrow$ 12 years of history!

Many companies not in the Russell 3000 in 1999

Covariance Matrix Not Invertible

- When stocks outnumber observations, sample covariance matrix thinks some portfolios of stocks are 100% safe!
- This cannot be true
- It is dangerous to believe that

Example: Uncorrelated Stocks

0	0	0	0
0	0.77	0	0
0	0	0.94	0
0	0	0	3.29

Stock 1 appears to have zero volatility

Invert this matrix $\Rightarrow$ Division-by-Zero Error!

General Case: Correlated Stocks

	Swatch	ABB	Novartis	Nestle
Swatch	2.06	0.91	0.40	0.82
ABB	0.91	1.50	0.49	0.39
Novartis	0.40	0.49	0.94	0.54
Nestle	0.82	0.39	0.54	0.51

This sample covariance matrix looks OK...

But is it really OK?

It Is Not Invertible

- This portfolio:

Swatch	ABB	Novartis	Nestle
-0.30	0.10	-0.42	0.85

- Has zero volatility!
- Not true
- Not safe
- Division-by-Zero Error

Even with Enough Observations

- Division-by-zero error is just an extreme case
- Even when observations outnumber assets, similar problem arises
- Stocks (portfolios) that appear the safest are in fact much less safe than they appear
- Stocks (portfolios) that appear the riskiest are in fact much less risky than they appear

Systematic Bias

- Markowitz portfolio optimization will overweight the stocks (portfolios) that appear the safest – but are in fact less safe than they appear
- Markowitz portfolio optimization will underweight the stocks (portfolios) that appear the riskiest – but are in fact less risky than they appear
- Michaud (1989): Error maximization

How Big Is this Problem?

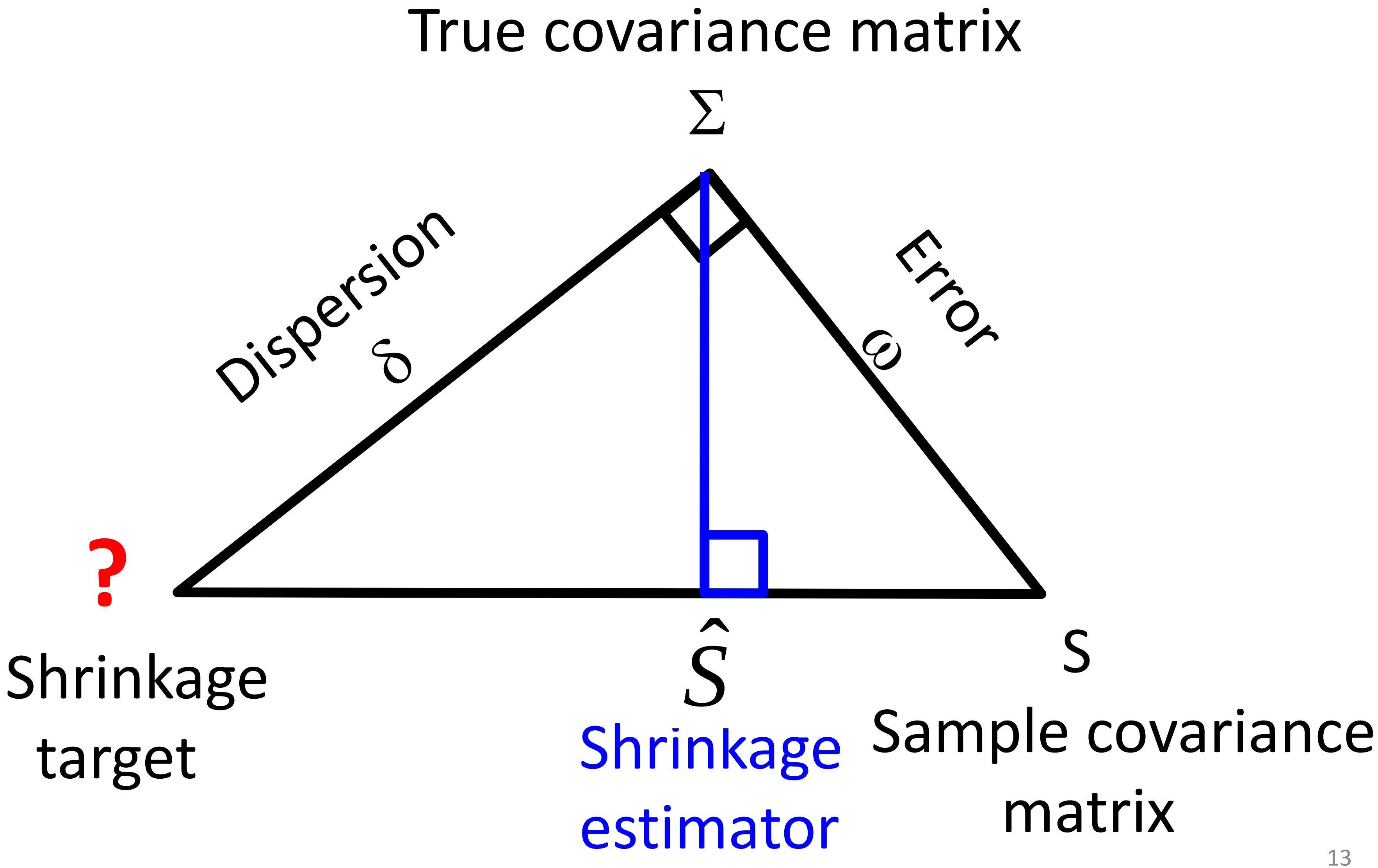
- Marčenko and Pastur (1967) proved the relative bias is of order: $2 \times \sqrt{n/T}$
  where n = number of stocks
  T = number of observations
- 30 stocks (e.g., Dow Jones)
- Daily data
- Tolerate 5% under/overweight
- How far back do you need to go?

1821

Solution


WALT DISNEY PICTURES PRESENTS
RICK MORANIS
IN
HONEY, I
SHRUNK
THE KIDS

Geometric Interpretation



Shrinkage Target for Covariance Matrix

- "Neutral" matrix
- Zero matrix or identity matrix?

$$w = (1-\lambda)\frac{\Sigma^{-1}\mathbf{1}}{\mathbf{1}'\Sigma^{-1}\mathbf{1}} + \lambda\frac{\Sigma^{-1}\mu}{\mathbf{1}'\Sigma^{-1}\mu}$$

- Identity matrix
- Properly scaled

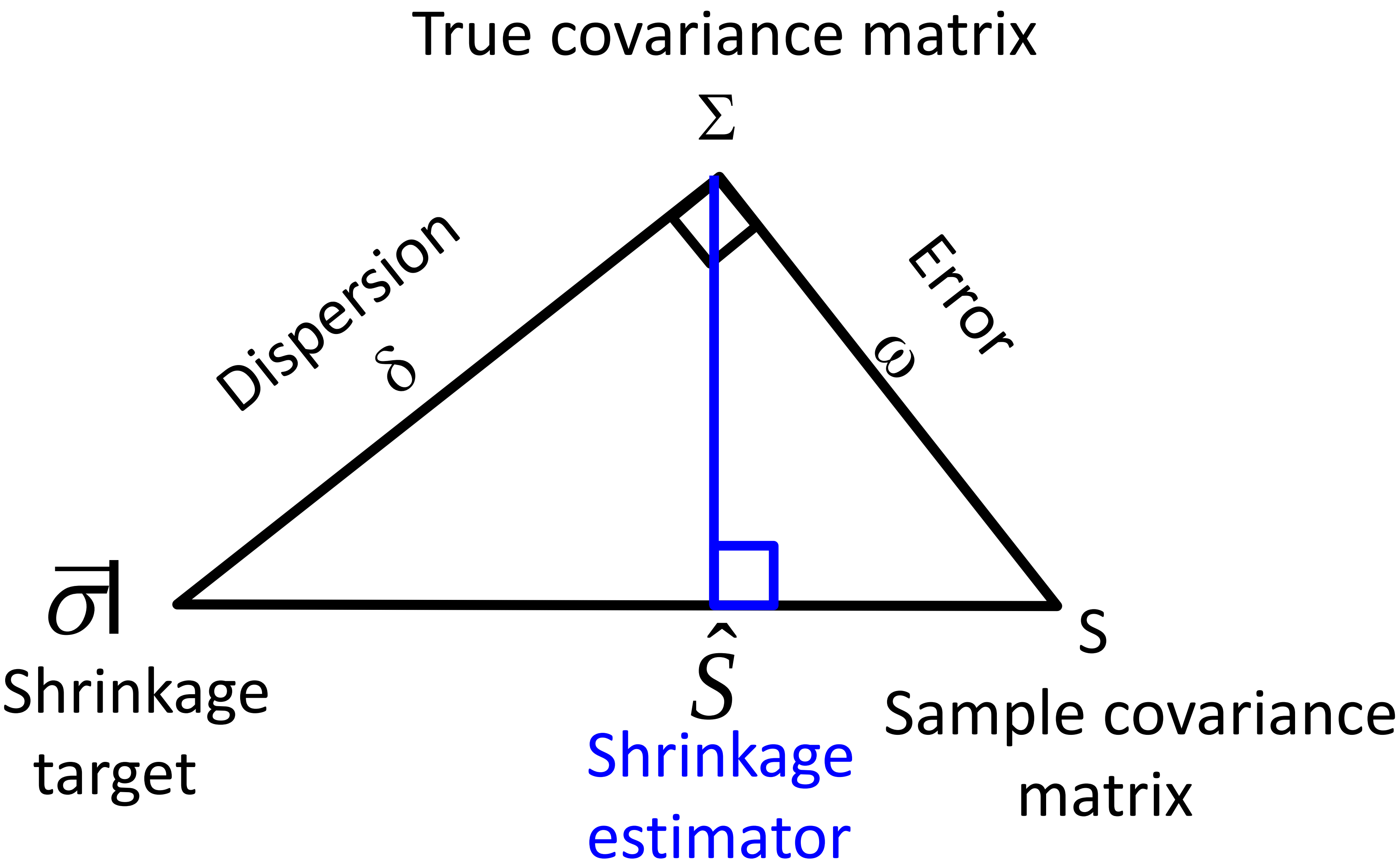
Scaling the Identity Matrix

- Identity matrix I has 1 on the diagonal (variances) and 0 off the diagonal
- Scaled identity matrix $\bar{\sigma}\mathrm{I}$ has $\bar{\sigma}$ on the diagonal (variances) and 0 elsewhere

$$\Rightarrow \bar{\sigma} = \frac{1}{n}\sum_{i=1}^{n}\sigma_{ii}$$

- Scaling factor = average variance

Geometric Interpretation



Shrinking the Covariance Matrix

$$\hat{S} = (1-\beta)\overline{\sigma I} + \beta S$$

Sample Covariance Matrix

Shrinkage Target

Shrinkage Slope

$$\beta = \frac{\delta^2}{\omega^2 + \delta^2}$$

Dispersion

Estimation Error

Distance between Matrices

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \quad B = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}$$

Squared distance between A and B:

$||A\text{-}B||^2 = (a_{11}\text{-}b_{11})^2+(a_{21}\text{-}b_{21})^2+(a_{12}\text{-}b_{12})^2+(a_{22}\text{-}b_{22})^2$

Same for any size matrices

Estimation Error ω^2

- ω^2 is squared distance between sample covariance matrix and true covariance matrix:

$$\omega^2 = E[||S - \Sigma||^2]$$

- We do not know the truth, but we can estimate how far from the truth is the sample covariance matrix

How to Estimate ω^2

- Vector containing returns on all stocks at date t:

$$X_t = \begin{bmatrix} x_{t1} \\ x_{t2} \\ \vdots \\ x_{tn} \end{bmatrix}$$

$$S = \frac{1}{T}\sum_{t=1}^{T} X_t X_t'$$

$$\hat{\omega}^2 = \frac{1}{T(T-1)}\sum_{t=1}^{T}\left\| X_t X_t' - S \right\|^2$$

Comparison with Standard Error

$$\hat{\omega}^2 = \frac{1}{T(T-1)} \sum_{t=1}^{T} \left\| X_t X_t' - S \right\|^2$$

To get confidence interval around sample mean:

$$\hat{\sigma}^2 = \frac{1}{T(T-1)} \sum_{t=1}^{T} \left(x_t - m \right)^2$$

Confidence interval: $m \pm 2\hat{\sigma}$

Dispersion δ^2

- δ^2 measures dispersion
- How far is true covariance matrix Σ away from "neutral" matrix: shrinkage target $\bar{\sigma}I$
- High dispersion: some stocks (portfolios) have much higher risk than others
- Low dispersion: all stocks (portfolios) have pretty much the same risk as one another

Example with Uncorrelated Assets

- High dispersion: $\begin{bmatrix} 0.1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1.9 \end{bmatrix} \rightarrow \delta^2 = 1.62$

- Low dispersion: $\begin{bmatrix} 0.9 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1.1 \end{bmatrix} \rightarrow \delta^2 = 0.02$

How to Estimate Dispersion δ^2

- Do the usual decomposition:

$$E[\|S - \bar{\sigma}\mathbf{I}\|^2] = E[\|S - \Sigma\|^2] + \|\Sigma - \bar{\sigma}\mathbf{I}\|^2$$

$$= \omega^2 + \delta^2$$

$E[\|S - \bar{\sigma}\mathbf{I}\|^2]$ can be estimated by $\|S - \bar{\sigma}\mathbf{I}\|^2$

- Therefore: $\hat{\delta}^2 = \|S - \bar{\sigma}\mathbf{I}\|^2 - \hat{\omega}^2$

Shrinkage Estimator of the Covariance Matrix

$$\hat{\beta} = \frac{\hat{\delta}^2}{\hat{\omega}^2 + \hat{\delta}^2}$$

Time-series

$$= 1 - \frac{1}{T(T-1)} \cdot \frac{\sum_{t=1}^{T} \left\| X_t X_t' - S \right\|^2}{\left\| S - \bar{\sigma} I \right\|^2}$$

Cross-section

$$\hat{S} = (1 - \hat{\beta})\bar{\sigma} I + \hat{\beta} S$$

What It Does

- Brings in the extreme stock (portfolio) variances
- Below-average stock (portfolio) variances are pushed upwards
- Above-average stock (portfolio) variances are pulled downwards

$\Rightarrow$ Solves the problem of error-maximization noted by Michaud (1989)

Non-Linear Effect

$$w = (1-\lambda)\frac{\left[(1-\hat{\beta})\bar{\sigma}\mathbf{I} + \hat{\beta}S\right]^{-1}\mathbf{1}}{\mathbf{1}'\Sigma^{-1}\mathbf{1}}$$

$$+\lambda\frac{\left[(1-\hat{\beta})\bar{\sigma}\mathbf{I} + \hat{\beta}S\right]^{-1}\mu}{\mathbf{1}'\Sigma^{-1}\mu}$$

$\Rightarrow$ Gives completely different mean-variance efficient frontier (unlike shrinking the mean vector)

Invertible

- No division-by-zero error even if T << n

- Shrinkage estimator is guaranteed to be always invertible

- n=300 stocks, t=24 months: no problem!

Magic?

Widely Applicable

- Many quant trading outfits use it
- Richard Michaud at New Frontier Advisors
- To find the determinants of default swap premia
- For radar detection of incoming missiles
- To decode the human genome
- To cure cancer
- To save the planet from global warming
- For mobile phones to communicate with masts

"Swiss Army Knife of covariance matrix estimation"

Alternative

- Factor models with economically meaningful factors: inflation, GDP growth, size, value, momentum, industry, etc
- Example: MSCI/BARRA
- Tricky: What are the factors?
- Problem: Very inaccurate

Other Alternative

- Factor models with statistical factors

- Tricky: How many factors are there?

- Too few factors: miss extra-factor covariance
- Too many factors: same problems as sample covariance matrix

Yet Another Alternative

- Use the sample covariance matrix but impose enough constraints on the portfolio optimization to prevent bad behavior

- Tricky: What types of constraints? At what level? Will they work? How can you tell?

Pick One Reading for Next Class

1. Acceleration Strategies (2006) by Gettleman & Marks
2. Reviving Momentum (2011) by Deutsche Bank
3. Predicting stock price movements from past returns (2004) by Grinblatt &Moskowitz
4. Is momentum really momentum? (2012), Novy-Marx
5. Do Industries Explain Momentum? (1999) by Moskowitz & Grinblatt
6. Style momentum within the S&P-500 index (2004) by Chen & De Bondt