Optimal Asset Allocation Practitioner’s Perspective

Andrzej Palczewski
University of Warsaw

EMS School ”Risk Theory and Related Topics”
Bedlewo, October 2008

Programme

1. Optimal asset allocation - an overview.
2. Alternative approaches.
3. Black-Litterman model.
4. Improved estimation of the covariance matrix.
5. Performance analysis.

Optimal Asset Allocation

An Overview

Outline

1. Dynamic portfolio theory.
2. Markowitz model.
3. Model errors.
4. Estimation errors.
5. Pitfalls of the Markowitz model.

Dynamic portfolio optimization

The market:

- One ”risk-free” asset with a constant rate of return r.
- n risky assets with the return vector $\mathbf{R}$.
- Returns of risky assets follow a multivariate normal distribution with expected returns $E[\mathbf{R}]$ and covariance matrix Σ.

Dynamic portfolio optimization - cont.

The problem:

Given initial wealth at time t, W_t, find the investment strategy which secures optimal consumption in subsequent time moments $C_t, C_{t+1}, \ldots, C_T$.

The investment strategy is characterized by the vector of portfolio weights, $\mathbf{w}$, with n elements as the weights of n risky assets. Then $(1 - \mathbf{w}\mathbf{1})$ denotes the weight of the risk-free asset ($\mathbf{1}$ is a vector of ones).

The consumption optimality is characterized by an investor's utility function U – the objective of the investor is to maximize the expected utility of consumption up to time T

$$\max_{C,\mathbf{w}} E[U(C_t, C_{t+1}, \ldots, C_T)]$$

Dynamic portfolio optimization - cont.

Under self-financing assumption and for a large class of utility functions the solution is:

$$\mathbf{w}^* = A\mathbf{M} + B\mathbf{H}$$

where
$A\mathbf{M}$ is the myopic component and is the demand of the risky asset due to its risk premium,
$B\mathbf{H}$ is the intertemporal hedging component and represents the hedge against future changes in the investment due to the stochastic character of the market.

Dynamic optimization - conclusions

- Dynamic optimization is the superior asset allocation technique.
- But dynamic optimization . . . is too complex.
- No known large commercial applications.

Dynamic optimization - conclusions

Simple example – stock-bond-cash mix in G7 countries, state variables affect only expected returns.

- Stock returns in each country depend on 3 state variables.
- Bond returns in each country depends on 3 state variables.
- In addition, international equity market can be described by 5 state variables and the same is true about international bond market.

Conclusion: we have 15 risky assets and approx. 50 state variables !

Assets allocation in practice

In the investment industry asset allocation is essentially a single-period strategy.

One-period strategy corresponds to the myopic part $A\mathbf{M}$ of the dynamic strategy.

The result:

In practice, we operate with suboptimal portfolios (the intertemporal hedging component is neglected).

Assets allocation in practice, cont.

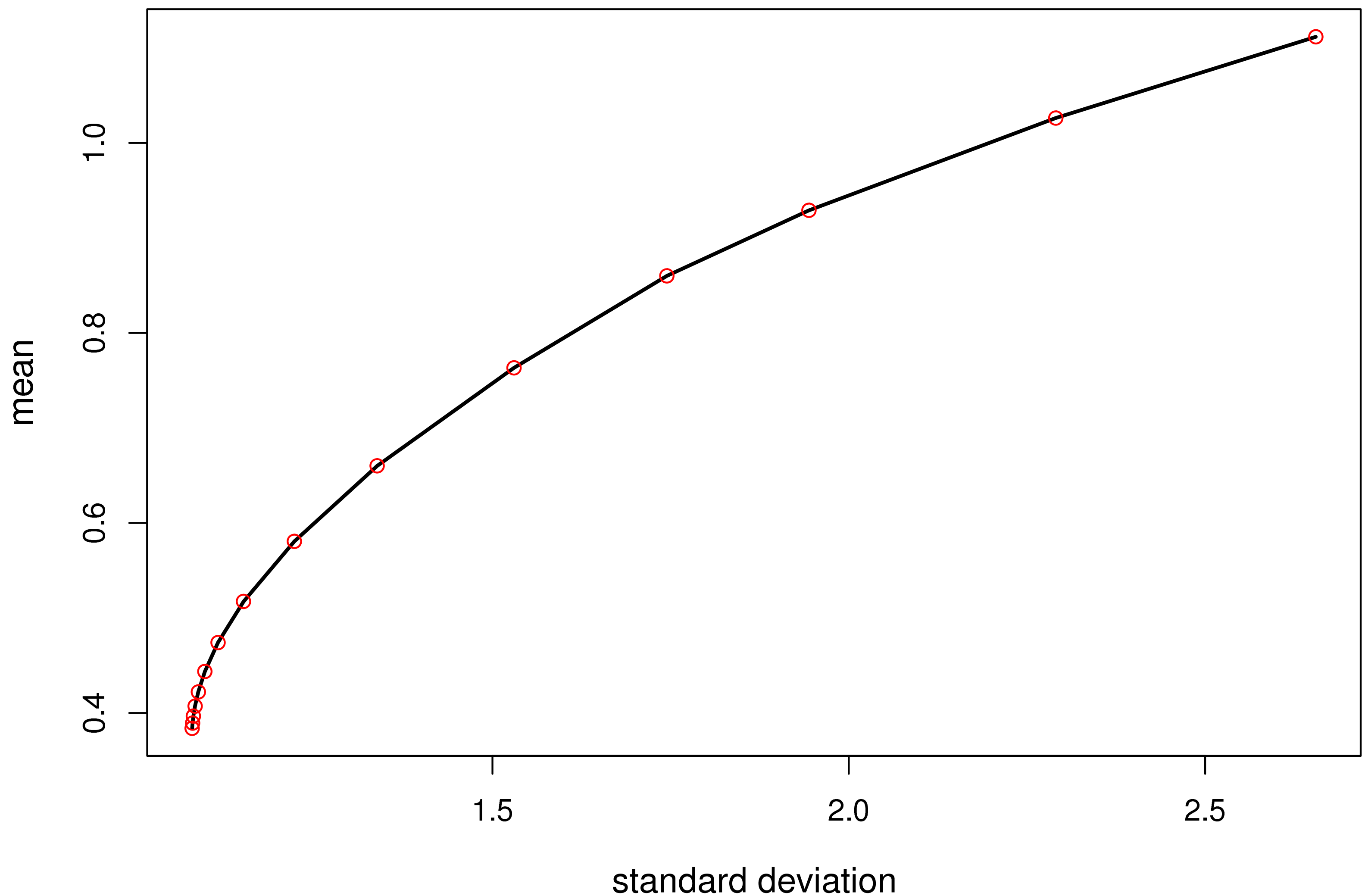
What is one-period optimization ?

This is the Markowitz return/risk (mean/variance) optimization

$$\max_{w_i} \left(\sum_i w_i \mu_i - \frac{\gamma}{2} \sum_{i,j} w_i \Sigma_{ij} w_j \right)$$

subject to $\sum_i w_i = 1$,
where
w_i are asset weights in portfolio,
γ is the investor's risk aversion.

Markowitz model – the solution



Efficient frontier

Markowitz model in use

To get ”correct” optimal portfolios from the Markowitz model we have to feed the model with good data:

- Expected (future) mean returns.
- Covariance of expected (future) returns.

In practice, we estimate future returns from historical (past) data !

Markowitz model in use – cont.

Where come the errors from ?

- Past returns are not good predictions of future returns.
- We estimate the mean and the covariance matrix from past data under the assumption that the distribution of returns is normal and constant in time (model error).
- We estimate the moments of the return distribution from a finite sample (estimation error).

Markowitz model in use – cont.

In finance, the past is not a good forecast of the (near) future.

- Historic mean returns do not forecast future mean returns.
- Historic covariance matrix predicts quite well future covariance matrix.

Explanation: like in the Black-Scholes model.

Conclusions:

- We have to find the way to forecast future mean returns (more in subsequent lectures).
- We can retain historic covariance matrix as good data for optimization.

Model errors

Stylized facts:

1. Multivariate return series show little auto-correlation and cross-correlation, but are not i.i.d. variables.
2. Series of squares of returns show profound evidence of cross-correlation and auto-correlation.
3. Conditional mean returns are close to zero.
4. Volatility and correlation between series vary over time.
5. Return distributions show high kurtosis and ”heavy tails”.

Model errors – cont.

How to minimize model errors ?

- Enlarge the class of admissible distributions (elliptic distributions).
- Use technique of nonlinear analysis of time series (GARCH).
- Use adequate estimators and long time series.

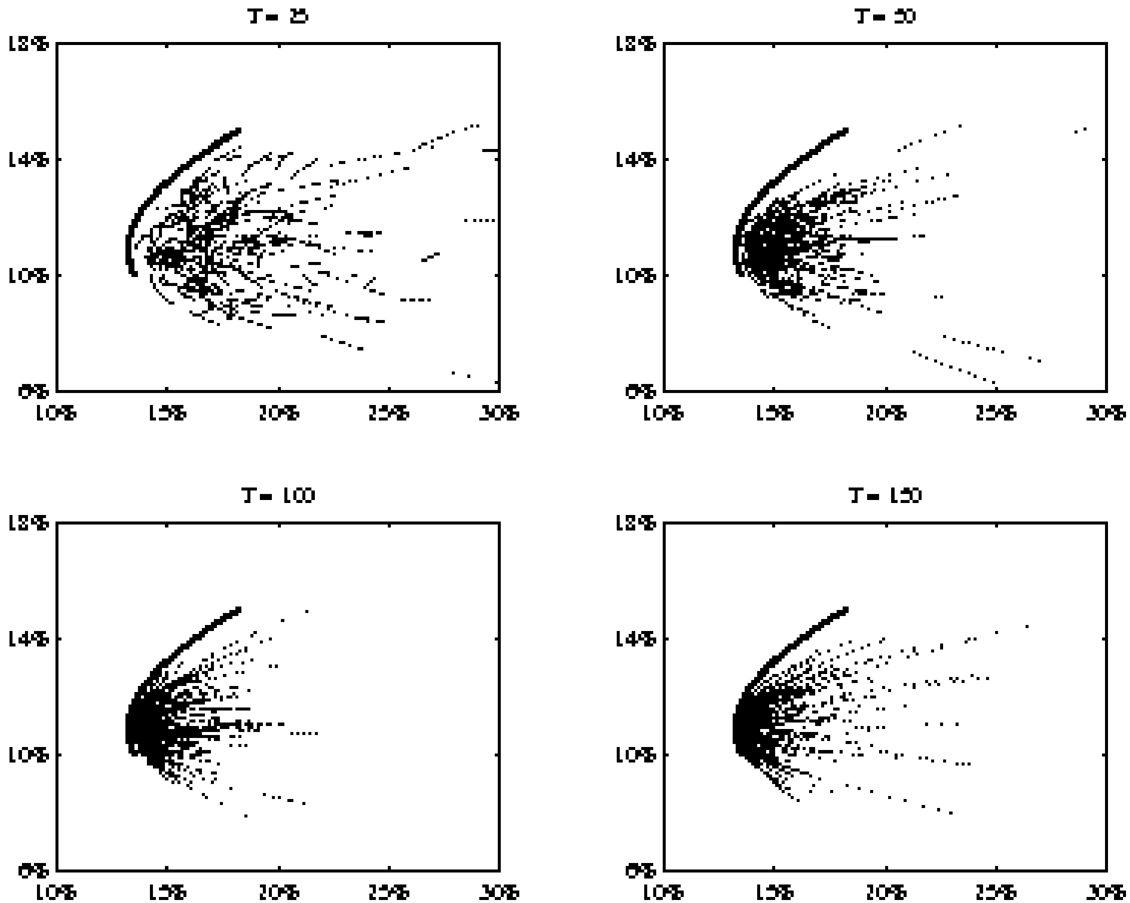
Drawback: to predict correctly mean return, you need 50 years of monthly data (Merton) !

Estimation errors

Jobson and Korkie experiment:

- 20 assets with multivariate normal distribution with mean and covariance derived from the real data of New York Stock Exchange during the period December 1949 to December 1975.
- These mean and covariance are taken to be the true moments of the distribution.
- From this distribution independent sets of hypothetical returns are simulated.
- From every set of simulated returns one computes estimates of the mean-variance efficient frontier.
- Obtained frontiers are compared to the true frontier.

Estimation errors – cont.



Brandt(2005)

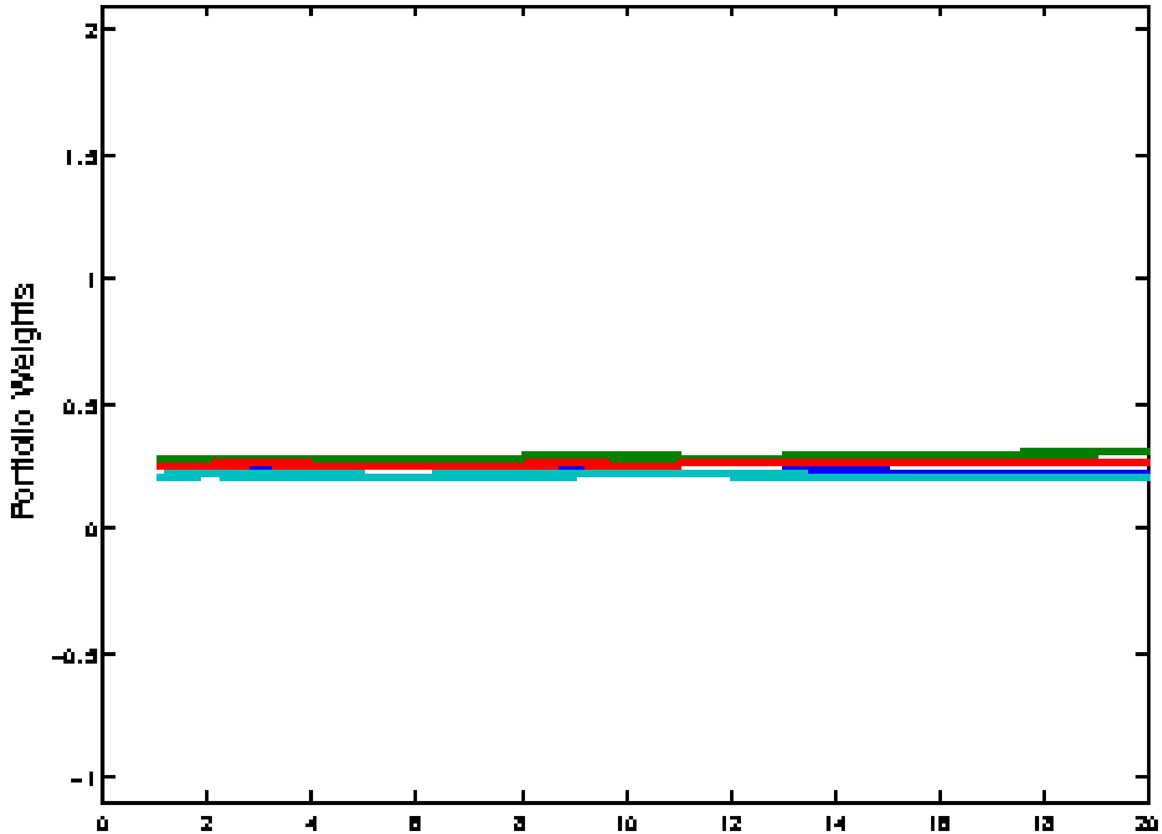
Efficient frontiers for 250 independent estimates – T is the number of monthly simulated returns. Solid line is the true frontier.

Estimation errors – cont.

DeMiguel and Nogales experiment:

- 4 assets with independent normal returns with mean return 12% and standard deviation 16%.
- From this distribution 140 hypothetical monthly returns are simulated.
- First 120 returns (i.e. returns with number 1,2,. . . , 120) are used to estimate the mean-variance optimal portfolio.
- Then the sample window is shifted by 1 (i.e. returns with number 2,3,. . . , 121 are used) and the optimal portfolio calculated.
- Repeating this procedure we obtain 20 different portfolios.

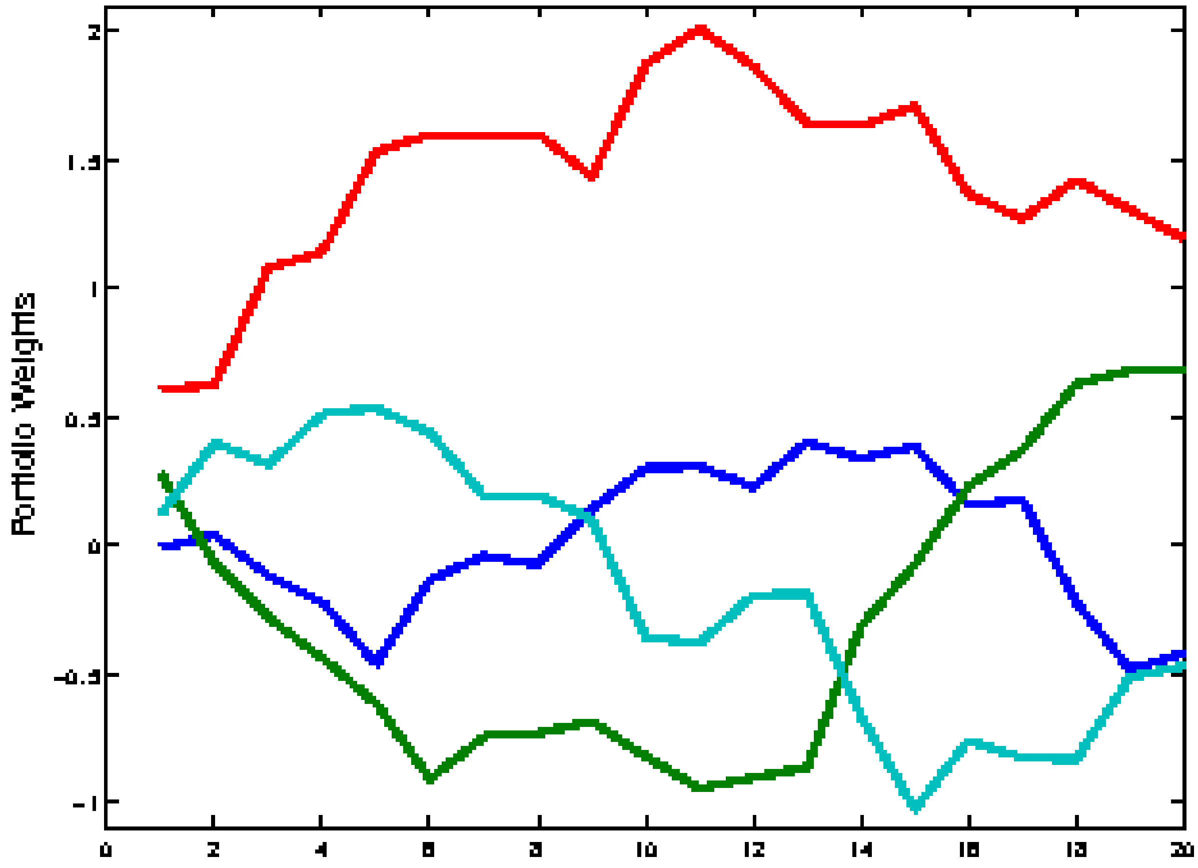
Estimation errors – cont.



DeMiguel&Nogales(2007)

Portfolio weights for minimal variance portfolio.

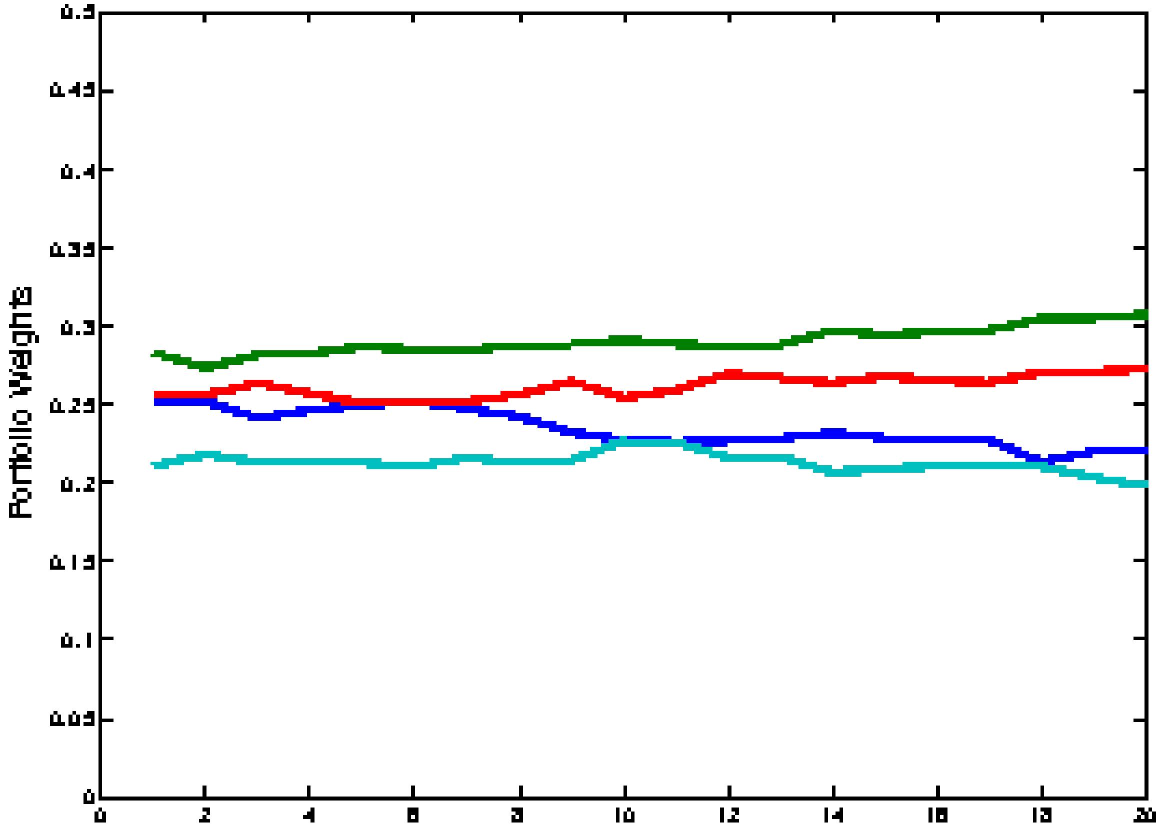
Estimation errors – cont.



DeMiguel&Nogales(2007)

Portfolio weights for investor’s risk aversion $\gamma = 1$.

Estimation errors – cont.



DeMiguel&Nogales(2007)

Portfolio weights for minimal variance portfolio (different scaling).

Pitfalls of the Markowitz model

- Portfolios are not well diversified.
- When investors impose no constraints, asset weights in the optimized portfolios almost always ordain large short positions in many assets.
- When constraints rule out short positions, the models often prescribe ”corner” solutions with zero weights in many assets, as well as unreasonably large weights in the assets of small capitalization.

Conclusion

Mean-variance optimizers are,
in a fundamental sense,
”estimation-error maximizers”.

(Michaud)

References

Chopra, V.K., Ziemba, W.T. – The effects of errors in the means, variances, and covariances, *J. Portfolio Management*, **19** (1993), 6–11.

DeMiguel, V., Nogales, F. J. – Portfolio selection with robust estimation. Preprint 2007.

Jobson, J.D., Korkie, B. – Estimation of Markowitz efficient portfolios, *J. Amer. Stat. Assoc.*, **75** (1980), 544–554.

Markowitz, H. M. – Mean-variance analysis in portfolio choice and capital markets, *J. Finance*, **7** (1952), 77–91.

Merton, R. C. – On estimating the expected return on the market: An exploratory investigation, *J. Fin. Econom.*, **8** (1980), 323–361.

Michaud, R.O. – The Markowitz optimization enigma: Is optimized optimal?, *Financial Analyst J.*, **45** (1989), 31–42.

Improving Markowitz Optimization

Outline

1. Resampling.
2. Robust estimators.
3. Robust optimization.

Introduction

The Markowitz mean-variance optimization is by far the most common formulation of the portfolio choice problem, despite all inconveniences mentioned earlier.

The Markowitz model yields two important economic insights:

- the effect of diversification,
- the fact that higher returns can only be achieved by taking on more risk.

These are the reasons why investors keep with the model trying to diminish its drawbacks.

Resampling

Resampling has been introduced by Michaud (1998) and is the subject of U.S. Patent.

The goal of resampling is the reduction of efficient portfolios sensitivity to data.

Resampling – practical receipt

1. Take a sample P_0 of historic returns. Estimate from this sample mean returns μ_0 and covariance matrix Σ_0.
2. Assume that the sample P_0 is the realization of a set of i.i.d. random variables. Make the assumption on the distribution of these variables, for instance assuming normality, and set the estimated parameters as the true parameters that determine the distribution of the returns
$$N(\mu_0, \Sigma_0).$$
3. From the above distribution generate a new sample P_1 (of same length as P_0), estimate its mean returns μ_1 and covariance matrix Σ_1 and solve the optimization problem to obtain efficient frontier.

Resampling – practical receipt

4. On the given efficient frontier choose m efficient portfolios of equally spaced target expected returns starting from minimal variance portfolio to portfolio with maximal return.
5. For every of m portfolios find asset's weights w_{1i}, $i = 1, \ldots, m$, where index 1 indicates that w_{1i} are weights for sample P_1.
6. Generate a new sample and repeat steps 3 to 5 for a large number Q of Monte Carlo simulations (Q should be of order 1000–10000).

Resampling – practical receipt

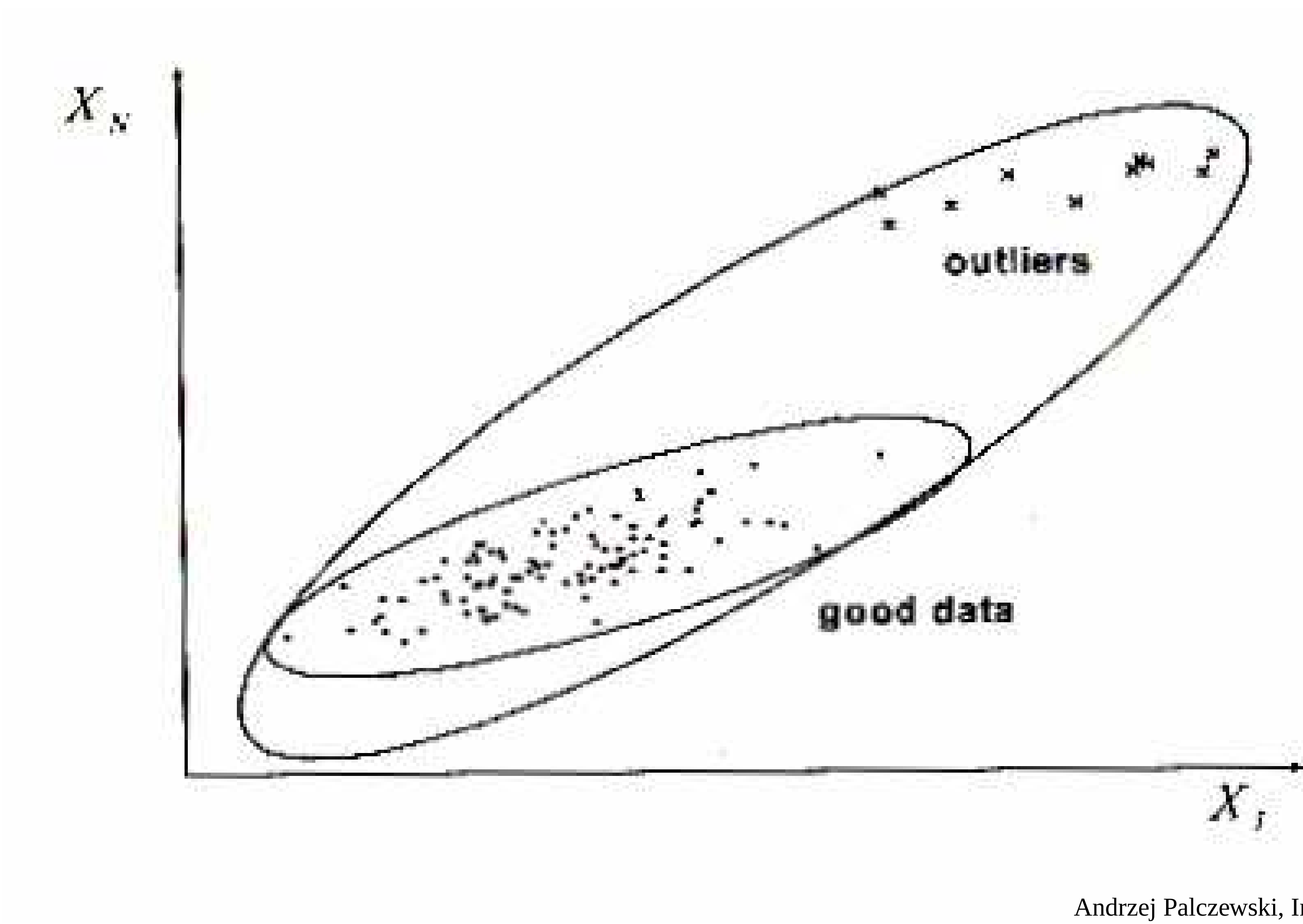
7. Use w_{qi} to calculate resampled weights

$$w_i = \frac{1}{Q} \sum_{q=1}^{Q} w_{qi}, i = 1, \ldots, m.$$

8. Recover the resampled efficient frontier from weights w_i.
9. Choose the efficient allocation on the resampled efficient frontier.

Robust estimators

Robust estimators, known for about 40 years, are estimators which eliminate outliers and estimate correct values for *location* and *dispersion* of true data.



Robust estimators

Popular robust estimators:

- M-estimators (Huber, Maronna).
- Stahel-Donoho estimators (Donoho, Stahel).
- MVE and MCD estimators (Rousseeuw).

Main drawback
Huge numerical complexity, particularly acute for financial data (long time series of multidimensional data).

Breakthrough
FastMCD algorithm of Rousseeuw and Van Driessen (1999).

Robust estimators

State of the art:

- Robust estimators are successfully applied in finance.
- Robust estimators improve stability of optimal portfolios.
- For simulated data estimated parameters are not very close to true values.

See for instance the mentioned paper of DeMiguel and Nogales.

Robust optimization

Robust optimization aims at determining a portfolio allocation $\mathbf{w}$ such that the opportunity cost is minimal for all values of parameters (mean and covariance for the Markowitz model) in a given uncertainty range.

For the standard Markowitz problem

$$\mathbf{w} = \operatorname*{argmax}_{\mathbf{w}} \left\{ \mathbf{w}^T \mu \quad | \quad \mathbf{w}^T \Sigma \mathbf{w} \le v \right\},$$

the robust optimization counterpart reads

$$\mathbf{w} = \operatorname*{argmax}_{\mathbf{w}} \left\{ \min_{\mu \in \Theta_\mu} \mathbf{w}^T \mu \quad | \quad \max_{\Sigma \in \Theta_\Sigma} \mathbf{w}^T \Sigma \mathbf{w} \le v, \right\}$$

where Θ_μ and Θ_Σ are uncertainty sets for μ and Σ.

Robust optimization

Possible specifications for the uncertainty sets:

- Elliptical set for expected values, known covariance (Ceria&Stubbs).
- Box set for expected values, elliptical set for covariance (Goldfarb&Iyengar).
- Box set for expected values, box set for covariance (Tütüncü&Koenig).

Instead of summary

DeMiguel, Garlappi and Uppal ”How inefficient is the 1/N portfolio strategy ?” (2007):

We have compared the performance of fourteen models of optimal asset allocation relative to that of the benchmark 1/N policy . . . we find that of the various optimizing models in the literature there is no single model that consistently delivers a Sharpe ratio or a certainty-equivalent return that is higher than that of the 1/N portfolio, . . .

):

References

Ben-Tal, A., Nemirovski, A. – Robust convex optimization, *Math. Oper. Research*, **23** (1998), 769–805.

Ceria, S., Stubbs, R. – Incorporating estimation errors into portfolio selection: Robust portfolio construction, *J. Asset Manag.*, **7**, No 2 (2006), 109–127.

Donoho, D. L. – Breakdown properties of multivariate location estimators. Ph.D. thesis, Harvard University 1982.

Goldfarb, D., Iyengar, G. – Robust portfolio selection problems, *Math. Oper. Research*, **28** (2003), 1–38.

Huber, P. J. – Robust estimation of a location parameter, *Ann. Math. Stoat.*, **35** (1964), 73–101.

Maronna, R. A. – Robust M-estimators of multivariate location and scatter, *Ann. Statist.*, **1** (1976), 51–67.

References, cont.

Michaud, R. O. – *Efficient Asset Management: A Practical Guide to Stock Portfolio Optimization and Asset Management*, HBS Press, 1998.

Rousseeuw, P. J. – Multivariate estimation with high breakdown point, in *Mathematical Statistics and Applications* (Eds. W. Grossmann, G. Pflug, I. Vincze and W. Wertz). Reidel 1985, pp. 283–297.

Rousseeuw, P. J., Van Driessen,K. – A fast Algorithm for the minimum covariance determinant estimator, *Technometrics*, **41** (1999), 212–223.

Stahel, W. A. – Robust estimation: Infinitesimal optimality and covariance matrix estimators. Ph.D. thesis, ETH, Zurich, 1981.

Tütüncü, R., Koenig, M. – Robust Asset Allocation, *Annals Operations Research*, **132** (2004), 157–187.

The Black-Litterman Model

Outline

1. Shrinkage estimators.
2. Black-Litterman assumptions and data.
3. CAPM (equilibrium returns).
4. Black-Litterman formula and its derivation.
5. Understanding the Black-Litterman formula.
6. Black-Litterman model in practice.

Shrinkage

Shrinkage estimators, called also Bayes-Stein estimators, are derived within Bayesian statistics. They are constructed as a convex combination of sample estimator $\bar{X}$ and a given reference point X_0

$$\hat{X} = \delta X_0 + (1 - \delta)\bar{X}.$$

Here X_0 can be thought as a Bayesian *prior* and $\bar{X}$ plays the role of an observation.

Shrinkage estimators are known in portfolio analysis quite a long time (Jobson&Korkie, Frost&Savarino, Jorion).
As has been shown by Jorion using shrinkage estimators can improve optimization procedure. But the result depends very much on the choice of X_0.

Black-Litterman assumptions

- Use the Markowitz mean-variance optimization.
- Feed the model with ”good data”.
- Historic covariances are appropriate for estimating portfolio risk.
- Historic means are ”bad data”.

BL assumptions – cont.

- To get correct ”expected mean returns” use
  - market equilibrium returns,
  - investor’s views on future market behavior,
- The core of the BL model is the method of joining the above data to obtain ”good” expected returns – *modified expected returns*.
- This is achieved by the BL formula.

BL assumptions – cont.

To obtain modified expected returns Black and Litterman have made the following assumptions:

- There exists ”market equilibrium”. In equilibrium all rational investors chose the same portfolio (market portfolio).
- Applying CAPM to the market portfolio we can calculate ”equilibrium returns”.
- Equilibrium returns are the first order approximations to true expected returns.
- Equilibrium returns are modified by investor’s views to get modified expected returns.

Black-Litterman data

To use the BL model we need the following data:

- market portfolio, weights w_i of assets in the portfolio,
- covariance matrix Σ of historic returns,
- vector of investor's views q,
- array P of views allocation (details will be given),
- matrix Ω of views covariances.

CAPM

- CAPM is used to obtain equilibrium returns for assets in market portfolio.
- In principle we can think of every asset as being in the market portfolio (eventually with zero weight).

CAPM – assumptions

1. Investors are risk averse with same risk aversion.
2. Investors optimize their portfolios using MV optimization (Markowitz).
3. Investors have same investment horizon.
4. The market is ideal: no constrains on borrowing and lending, no taxes, no transaction costs.
5. Risk-free rate is well defined.

Capital Asset Pricing Model

Let $\hat{R}$ be a random variable describing returns of a certain optimal portfolio.

(**CAPM**) Let random variable ζ_i describe the returns of asset i, then

$$\boxed{\mathbb{E}(\zeta_i) - r = \beta_i(\mathbb{E}(\hat{R}) - r),}$$

where

$$\beta_i = \frac{Cov(\zeta_i, \hat{R})}{Var(\hat{R})}$$

and r is risk-free rate.

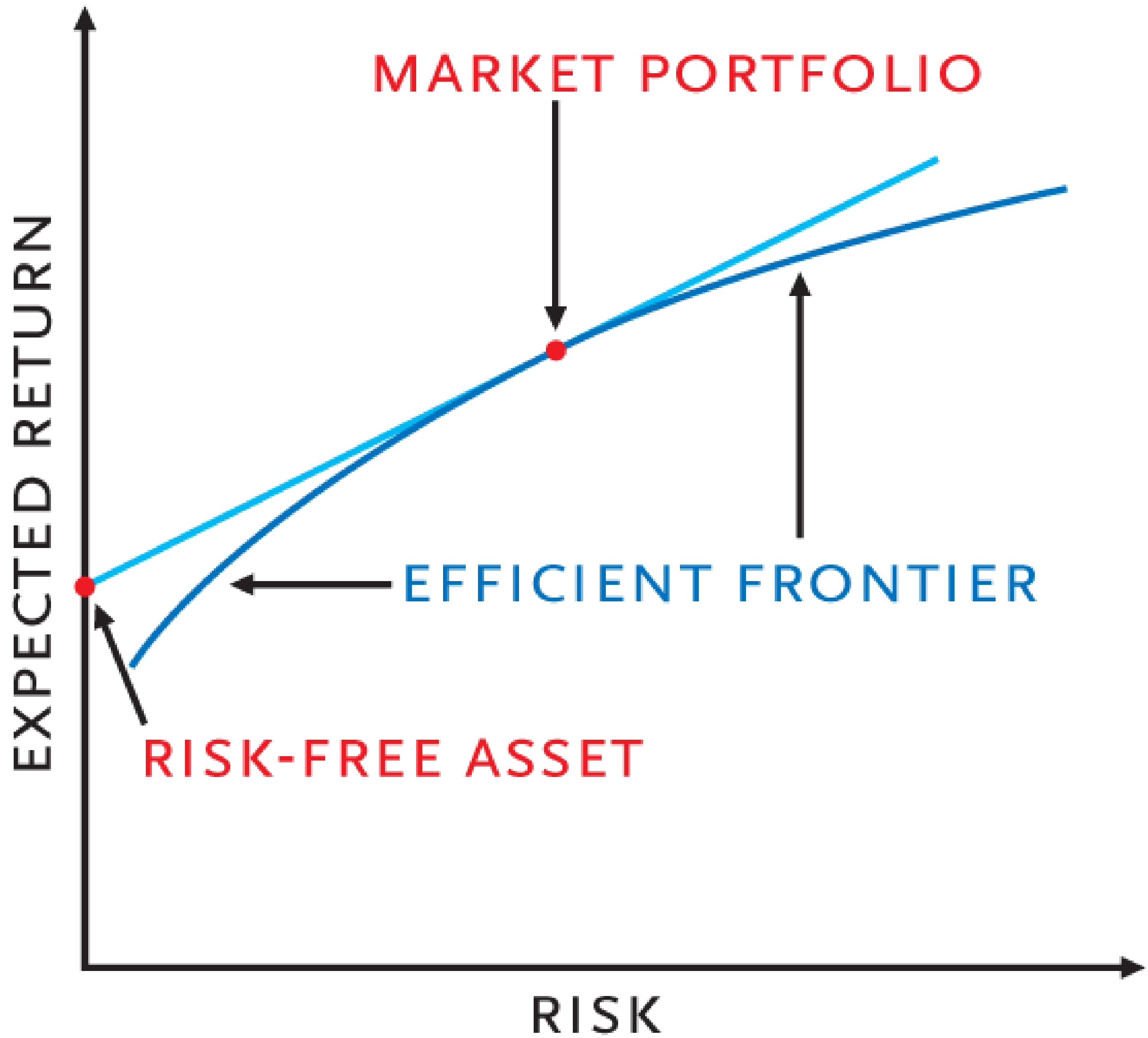
CAPM – conclusions

$$\text{Excess return asset } i \quad = \quad \beta_i * \text{excess return portfolio } \hat{R}$$

To obtain asset's excess return(μ_i) we have to find portfolio $\hat{R}$, such that:

- we can estimate coefficient β_i, i.e. covariance of portfolio $\hat{R}$ with asset i,
- we can estimate excess return of portfolio $\hat{R}$,
- portfolio $\hat{R}$ is from the efficient frontier, i.e. is optimal for certain risk aversion coefficient.

Market portfolio



Market portfolio – cont.

When we know the picture from the previous slide all is done.
But we don't.
What we know:

- risk-free rate r,
- market portfolio R_M,
- we know how to calculate β_i

$$\beta_i = \frac{Cov(\zeta_i, R_M)}{Var(R_M)}$$

From the CAPM formula we can obtain excess return μ_i for asset i

$$\mu_i - r \equiv \mathbb{E}(\zeta_i) - r = \beta_i(\mathbb{E}(R_M) - r)$$

But we don't know $\mathbb{E}(R_M)$!

Calculating market return $\mathbb{E}(R_M)$

Market portfolio $\mathbf{w}_M$ is the solution of the optimization problem

$$\mathbf{w}^T\mu - \frac{\gamma}{2}\mathbf{w}^T\Sigma\,\mathbf{w} \longrightarrow \max\,.$$

We know $\mathbf{w}_M$ and Σ but we don't know μ and γ.

When we know one of these values we can calculate the other one by reverse optimization.

γ is the market risk aversion which can be estimated from the Sharpe ratio.

Sharpe ratio

Portfolio mean return: μ_p
Portfolio variance: σ_p^2
Sharpe ratio:

$$SR = \frac{\text{mean return}}{\text{standard deviation}} = \frac{\mu_p}{\sigma_p}$$

When the portfolio is the solution to the optimization problem with risk aversion γ then

$$\gamma = \frac{SR}{\sigma_p}$$

Calculating $\mathbb{E}(R_M)$ – cont.

We find γ by knowledgable guess:

- Market risk aversion can be obtain from the Sharpe ratio:

$$\gamma = \frac{SR}{\text{market portfolio std dev.}}$$

  - to estimate Sharpe ratio we can use the rule of thumb: for stock $SR = 0.5$, for bond $SR = 1$,
  - this rule requires an empirical correction (experience).
- Market risk aversion can be estimated comparing implied asset returns with historical returns and investor's views.

Equilibrium returns

- choose assets and the market portfolio of these assets $\mathbf{w}_M$,
- calculate covariance matrix Σ and variance of the market portfolio,
- select the value of market risk aversion γ,
- find the market portfolio mean return μ_M,
- calculate coefficients β_i,
- using CAPM calculate equilibrium asset returns

$$\mu_i - r = \beta_i(\mu_M - r)$$

In the vector form

$$\boxed{\mu = \gamma \Sigma \, \mathbf{w}_M}$$

International market

Calculating equilibrium returns for investments on international market is much more involved:

- we have to take into account currency hedging,
- covariance matrix Σ is calculated for hedged excess returns,
- to calculate equilibrium returns we have to use International Capital Asset Pricing Model ICAPM (Black, Solnik).

Investor’s views

- Investor expresses views on assets mean returns.
- Investor’s views are of two types:
  - absolute view,
  - relative view.
- Absolute views specify the assets expected returns.
- Relative views specify the differences between expected returns of two or more assets.

Investor's views cont.

From investor's views we construct 3 objects:

- vector of investor's views q – each entry is the expected excess return of an asset or the difference of expected excess returns of two or more assets,
- "pick" matrix P of investor's views – each row corresponds to one view, each column to one asset,
- covariance matrix of views Ω.

Matrix P

Matrix P is build according to the following rules:

- number of rows is equal to the number of views,
- number of columns is equal to the number of assets,
- when the view is an absolute view on asset X, then in its row in column corresponding to asset X we put 1, all other columns contain 0,
- when the view is a relative view on instruments X and Y, then in its row in columns corresponding to assets X and Y we put numbers between -1 and 1 which sum is equal zero.

Matrix P cont.

Which numbers should appear in a row corresponding to a relative view:

- Black and Litterman suggests
  - 1 and -1 (1 for asset with higher return and -1 for asset with lower return),
  - for 2 pairs of instruments suggested weights are $(0.5, 0.5, -0.5, -0.5)$. This means that we put equal weights for all instruments.
- Drobetz proposes for group of instruments weights which are proportional to instrument's capitalization (still keeping sum equal to zero)

Matrix P example

We have 3 assets A, B, C. Investor's views:

1. Asset A will have an absolute excess return of 3.5%.
2. Asset B will outperform asset C by 1.5%.

For these views we construct ”pick” matrix P

$$P = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -1 \end{bmatrix},$$

and vector q

$$q = \begin{bmatrix} 3.5 \\ 1.5 \end{bmatrix}.$$

Matrix Ω

Matrix Ω expresses the investor confidence in views. Black and Litterman suggests diagonal matrix which for n views has the form

$$\Omega = \begin{bmatrix} \sigma_1^2 & 0 & \cdots & 0 \\ \vdots & \ddots & & \vdots \\ 0 & \cdots & 0 & \sigma_n^2 \end{bmatrix}.$$

The diagonal elements are calculated from the formula

$$\sigma_i^2 = p_i \Sigma p_i^T,$$

where

- p_i – a row of matrix P for i-th view,
- Σ – asset's covariance matrix.

Matrix Ω cont.

- View's confidence is proportional to the variance of asset to which the view corresponds.
- A diagonal Ω corresponds to the assumption that the deviations of expected returns from the means representing each view are independent.
- We can additionally quantify views confidence level by specifying parameter τ which measures views strength.

Parameter τ

τ measures overall confidence level of views. Its value is a subject of constant discussion among experts

- For Black and Litterman τ should be close to zero (they used $\tau = 0.025$).
- Bevan and Winkelmann suggest to choose τ such that *the information ratio* will be smaller than 2.
- Satchell and Scowcroft claim that τ should be close to 1.

Matrix Ω cont.

Another suggested choice of Ω (Meucci)

$$\Omega = P\Sigma P^T.$$

Black-Litterman formula

$$\mu = \left((\tau\Sigma)^{-1} + P^T \Omega^{-1} P \right)^{-1} \left((\tau\Sigma)^{-1} \mu_{eq} + P^T \Omega^{-1} q \right),$$

where

μ – vector of modified expected returns,

μ_{eq} – vector of equilibrium returns,

Σ – covariance matrix of equilibrium returns,

P, Ω and τ – defined as before.

Black-Litterman formula cont.

Alternative expression

$$\mu = \mu_{eq} + \Sigma P^T (\Omega/\tau + P\Sigma P^T)^{-1}(q - P\mu_{eq}).$$

The Black-Litterman formula for $\Omega = P\Sigma P^T$ and matrix P of full range

$$\mu = (1+\tau)^{-1}\left(\mu_{eq} + \tau(P^T P)^{-1} P^T q\right),$$

This is the shrinkage operator for mean.

BL formula derivation

Let

- f_μ – probability distribution *a priori* of expected returns,
- $f_{q|\mu}$ – conditional probability distribution of investor's views given expected returns,
- $f_{\mu|q}$ – probability distribution *a posteriori* of expected returns given investor's views.

Then the Black-Litterman formula is derived from the Bayes formula

$$f_{\mu|q}(\mu) = \frac{f_{q|\mu}(q) f_\mu(\mu)}{\int f_{q|\mu}(q) f_\mu(\mu) d\mu}.$$

BL formula derivation cont.

What we know

$$f_\mu(\mu) = N(\mu_{eq}, \tau\Sigma),$$

$$f_{q|\mu} \sim N(P\mu, \Omega).$$

From the Bayes formula we obtain

$$f_{\mu|q} \sim N(\hat{\mu}, \Psi),$$

where $\hat{\mu}$ is given by the Black-Litterman formula (we can also derive the formula for covariance matrix Ψ).

BL in use

- select the market (assets, time series of assets returns),
- choose the market portfolio (strategic benchmark),
- estimate covariance matrix,
- calculate equilibrium returns (choose market price of risk γ),
- collect views
- calculate modified returns from BL formula (choose confidence level for views τ),
- solve the optimization problem and find the investment portfolio.

Choosing market γ

Equilibrium risk premium for different γ

Sector	Weights	$\gamma = 1$	$\gamma = 2.5$	$\gamma = 5$	$\gamma = 7.5$	Hist.
S_L	2.89%	3.07%	7.69%	15.37%	23.06%	5.61%
S_M	3.89%	2.21%	5.52%	11.03%	16.55%	12.75%
S_H	2.21%	2.04%	5.11%	10.22%	15.33%	14.36%
B_L	59.07%	2.62%	6.55%	13.10%	19.64%	9.72%
B_M	23.26%	2.18%	5.44%	10.88%	16.32%	10.59%
B_H	8.60%	1.97%	4.91%	9.83%	14.74%	10.44%

Brandt(2005)

”S” and ”B” refer to stock capitalization: small and big.
”L”, ”M” and ”H” refer to book-to-market ratio: low, medium and high.

Choosing market γ

Risk premia for different Sharpe ratio.

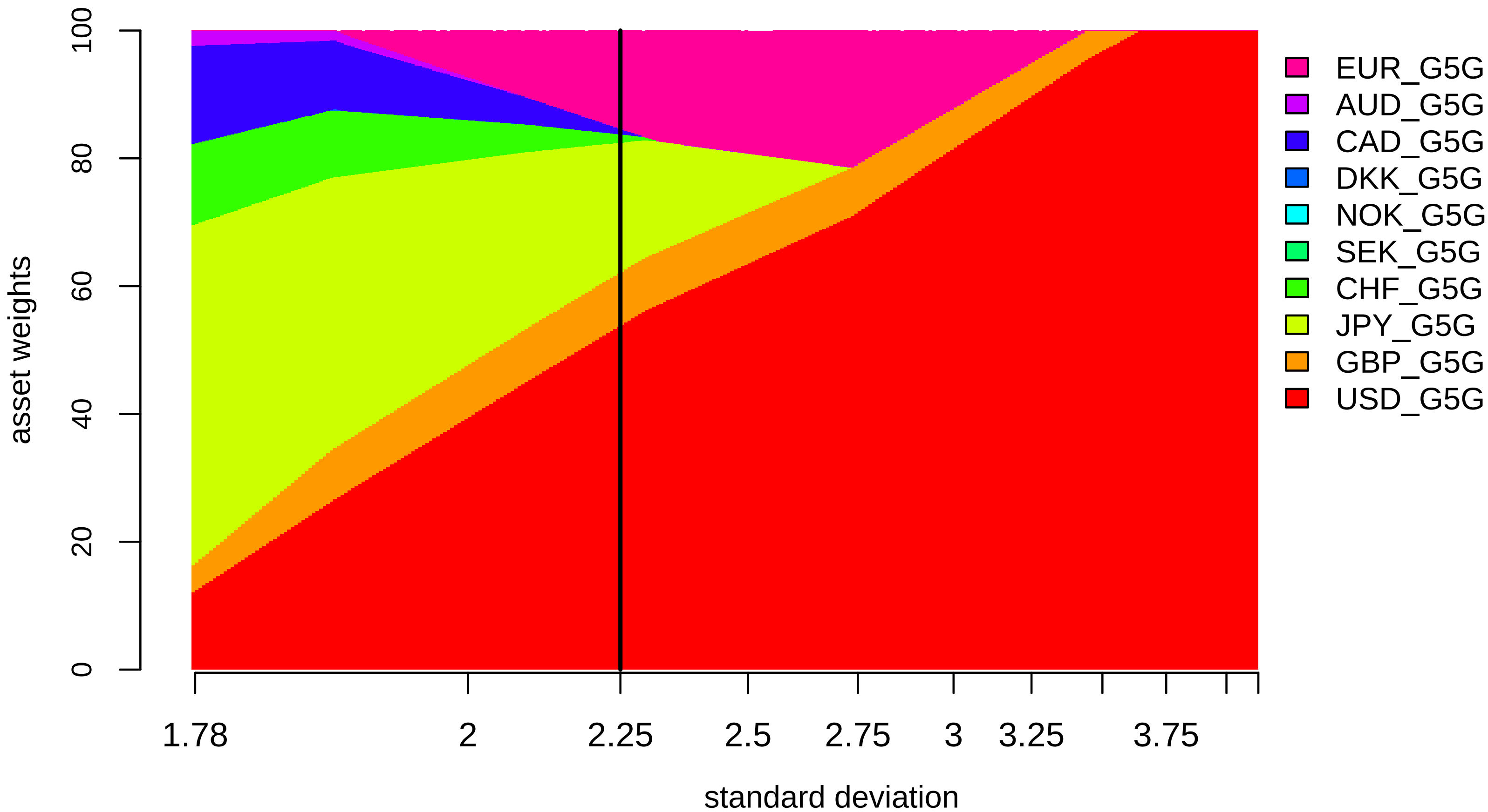
Assets(bond)	Views	$\gamma = 20$	$\gamma = 30$	$\gamma = 35$	$\gamma = 40$	$\gamma = 50$
USD	1.89	1.14	1.71	1.99	2.28	2.85
GBP	3.06	0.84	1.26	1.47	1.69	2.11
CHF	2.89	0.62	0.93	1.08	1.24	1.55
AUD	-1.47	0.33	0.5	0.58	0.66	0.83
NOK	-2.08	0.64	0.96	1.12	1.28	1.59
JPY	-0.46	0.23	0.35	0.41	0.47	0.59
EUR	1.39	0.9	1.36	1.58	1.81	2.26
SR		0.43	0.65	0.76	0.87	1.08

Choosing τ

Portfolio weights as the function of τ

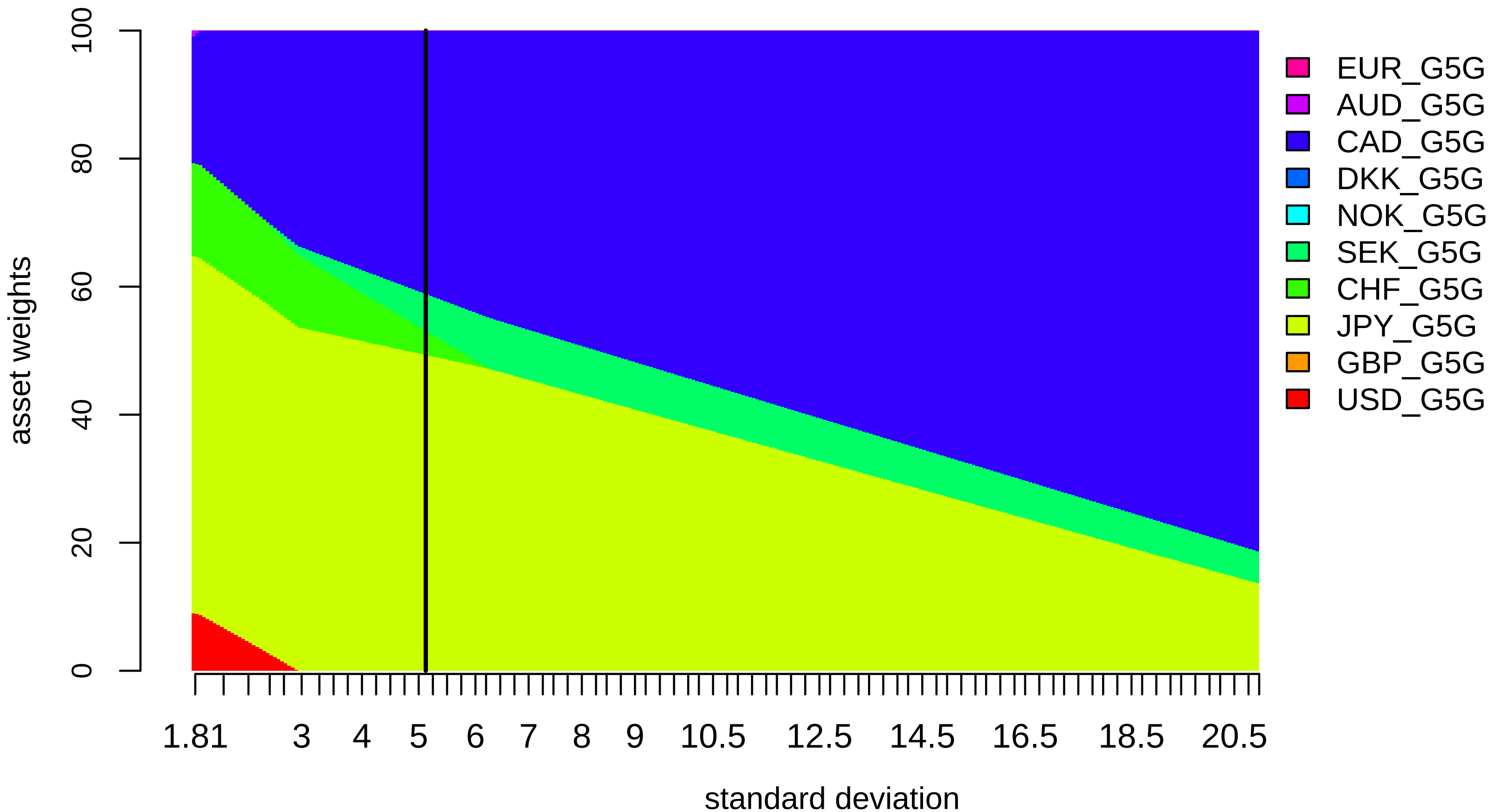
assets (bond)	portfolio weights					
	$\tau = 0$	$\tau = 0.05$	$\tau = 0.1$	$\tau = 0.15$	$\tau = 0.2$	$\tau = 0.3$
USD	45.74	46.25	46.7	47.06	47.36	47.78
GBP	5.52	8.49	10.98	13.37	15.68	20.06
CHF	4.1	7.91	10.74	13.45	16.05	20.96
AUD	1.28	0	0	0	0	0
NOK	2.13	0	0	0	0	0
JPY	16.34	14.29	11.62	9.18	6.92	2.88
EUR	24.89	23.07	19.97	16.94	14	8.32

Black-Litterman model – the solution



Portfolio composition

Markowitz model – the solution



Portfolio composition

References

Bevan, A., Winkelmann, K. – Using Black-Litterman global asset allocation model: three years of practical experience, Goldman Sachs, Fixed Income Research, June 1998.

Black, F. – Equilibrium exchange rate hedging, *J. Finance*, **45** (1990), 899–907.

Black, F., Litterman, R. – Global portfolio optimization, *Financial Analysts J.*, **48** (1992), 28–43.

Drobetz, W. – How to avoid the pitfalls in portfolio optimization? Putting the Black-Litterman approach at work, *Finan. Markets Portfolio Managm.*, **15** (2001), 59–75.

References, cont.

He, G., Litterman, R. – The intuition behind Black-Litterman model portfolios, Goldman Sachs, Investment Management Division, December 1999.

Meucci, A. – *Risk and Asset Allocation*, Springer 2005.

Satchell, S., Scowcroft, A. – A demystification of the Black-Litterman model: managing quantitative and traditional construction, *J. Asset Managm.*, **1** (2000), 138–150.

Sharpe, W. F. – An equilibrium model of the international capital market, *J. Econom. Theory*, **8** (1964), 500–524.

Solnik, B. H. – *International Investments*, Addison-Wesley 2000.

Improved Estimation of the Covariance Matrix

Outline

1. Covariance matrix estimation errors.
2. Understanding covariance matrix.
3. Improving covariance estimates:
    (a) Factor models;
    (b) PCA analysis.
4. Shrinkage estimators.
5. GARCH.

Estimating covariance

We have claimed in the previous lecture that covariance estimators produce correct data and the main problem is with risk premia.

This is true only to a certain extend.

To obtain good estimates of covariance we need

- clean data (no outliers),
- long time series,
- data from stationary distribution (normal at best).

None of these conditions is fulfilled for financial data !

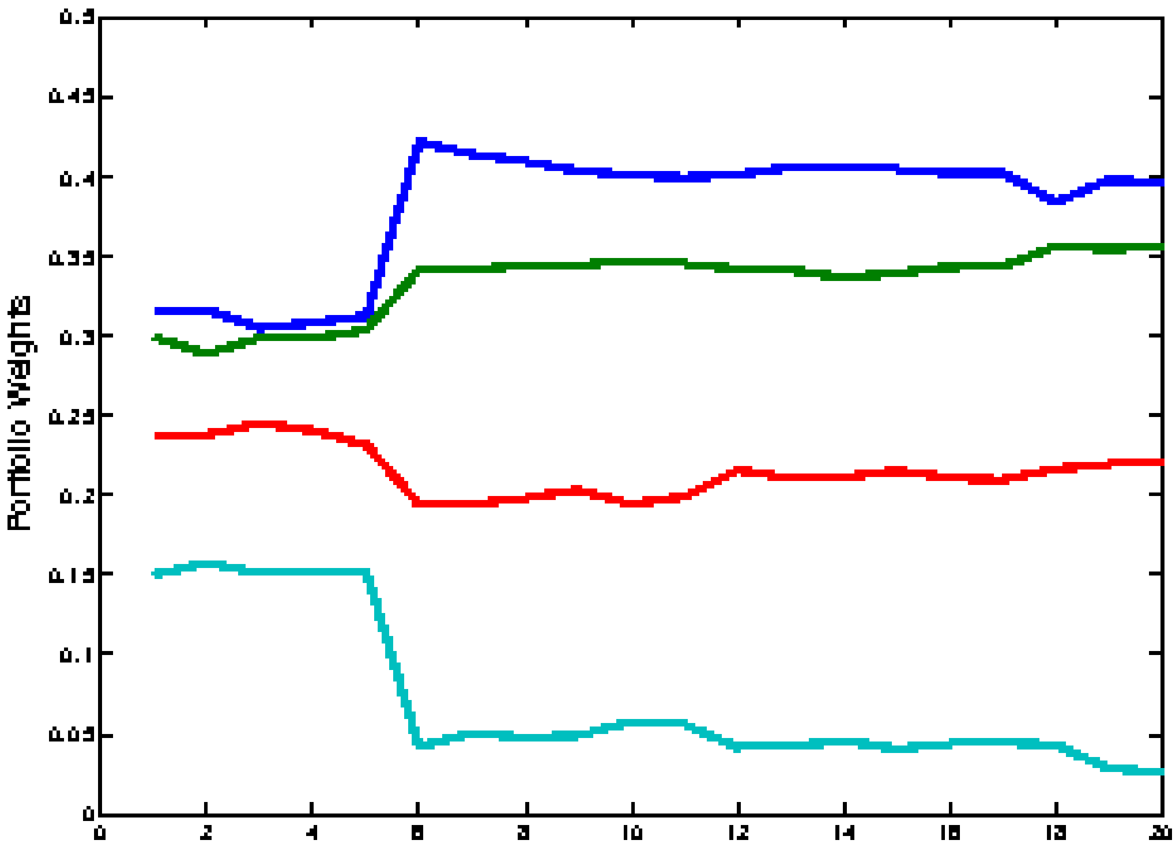
Estimating covariance cont.

DeMiguel and Nogales experiment again:

- 4 assets with independent normal returns with mean return 12% and standard deviation 16%.
- From this distribution 140 hypothetical monthly returns are simulated, 2 outliers have been artificially inserted.
- First 120 returns (i.e. returns with number 1,2,. . . , 120) are used to estimate the mean-variance optimal portfolio.
- Then the sample window is shifted by 1 (i.e. returns with number 2,3,. . . , 121 are used) and the optimal portfolio calculated.
- Repeating this procedure we obtain 20 different portfolios.

Estimating covariance cont.

The effect of outliers



DeMiguel&Nogales(2007)

Portfolio weights for minimal variance portfolio (only covariance matrix influence the result).

Estimating covariance cont.

Why small errors in covariance estimates produce big changes in portfolio composition?

Simplest MV optimization

$$\max_{\mathbf{w}} \left\{ \mathbf{w}^T \mu - \frac{\gamma}{2} \mathbf{w}^T \Sigma \, \mathbf{w} \quad | \quad \mathbf{w}^T \mathbf{1} = 1 \right\},$$

gives the solution

$$\mathbf{w}^* = \frac{1}{c} \Sigma^{-1} \mathbf{1} + \frac{1}{\gamma} \Sigma^{-1} \left(\mu - \frac{a}{c} \mathbf{1} \right),$$

where

$$a = \mathbf{1}^T \Sigma^{-1} \mu, \quad c = \mathbf{1}^T \Sigma^{-1} \mathbf{1}.$$

$\mathbf{w}^*$ depends on Σ^{-1} !

Understanding covariance matrix

Spectral decomposition of Σ:

- Λ – diagonal matrix of eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ of Σ.
- Q – orthogonal matrix with columns being eigenvectors $\mathrm{e}_1, \mathrm{e}_2, \ldots, \mathrm{e}_n$ of Σ.

$$\Sigma = Q\Lambda Q^T$$

- Small eigenvalues of Σ create troubles.
- Even small errors can give large changes in small eigenvalues.
- Small eigenvalues of Σ correspond to large eigenvalues of Σ^{-1}.
- For financial data eigenvalues of covariance matrix can differ by several orders of magnitude.

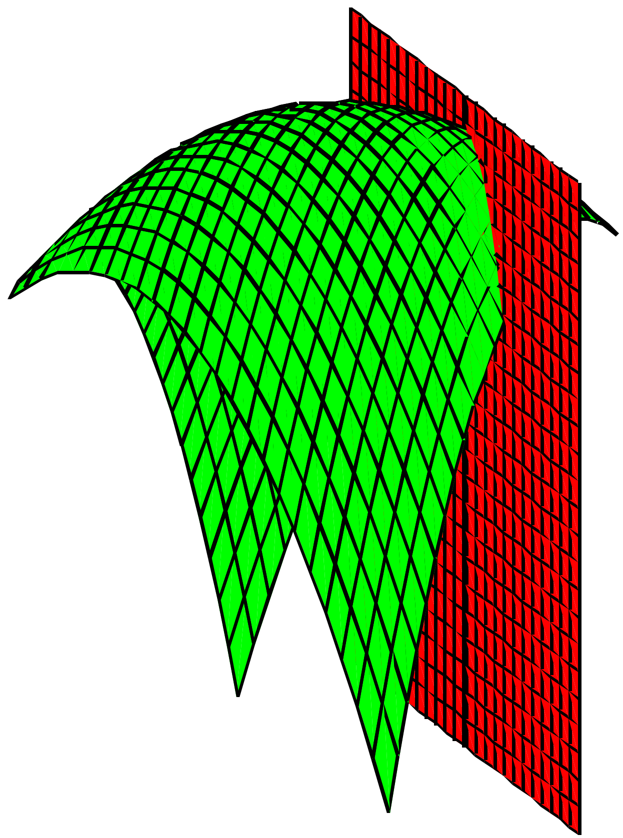
Covariance matrix – geometric picture

Optimization utility function

$$f(\mathbf{w}) = \mathbf{w}^T \mu - \frac{\gamma}{2} \mathbf{w}^t \Sigma\, \mathbf{w}.$$

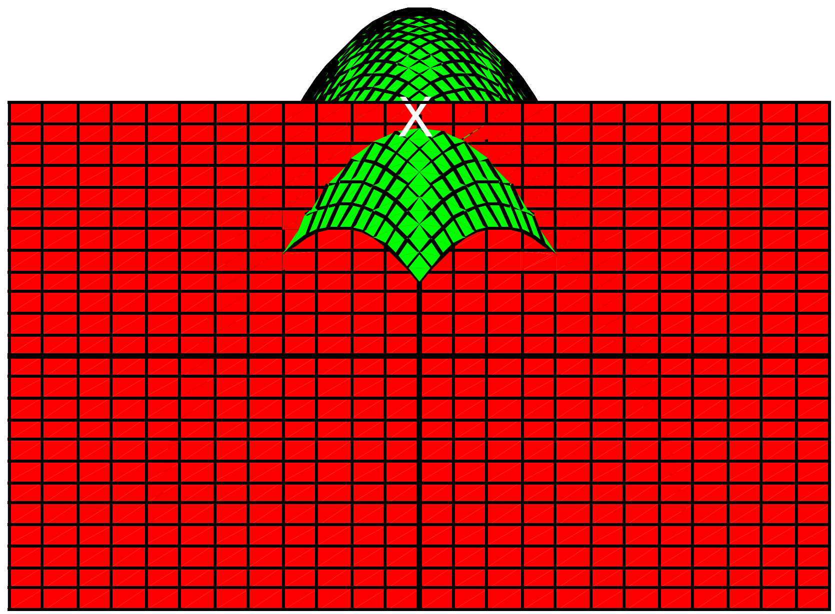
Geometric picture for 2 assets.

Utility surface



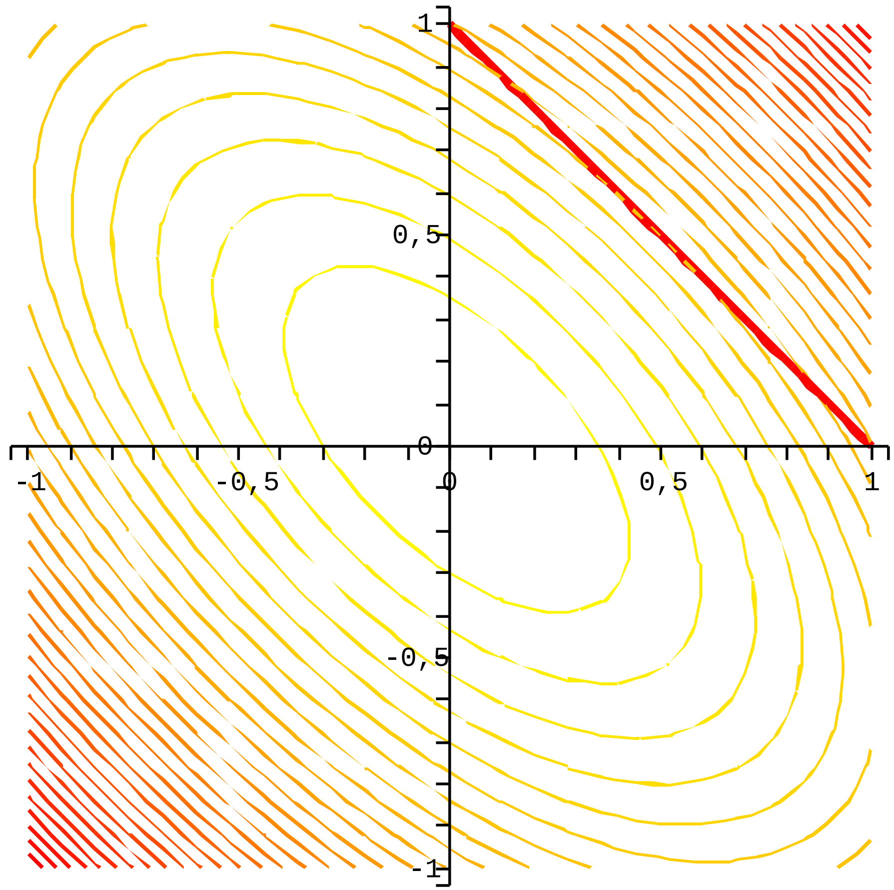
Utility surface (green) is an elliptic paraboloid.
Red is the plane $w_1 + w_2 = 1$.

Optimal solution



X is the optimal solution.
It is obtained by the intersection of the utility surface with the plane $w_1 + w_2 = 1$.

Contours (levels) of utility function



Levels are ellipses.

The point of tangency between an ellipse and the line $w_1 + w_2 = 1$ is the optimal solution.

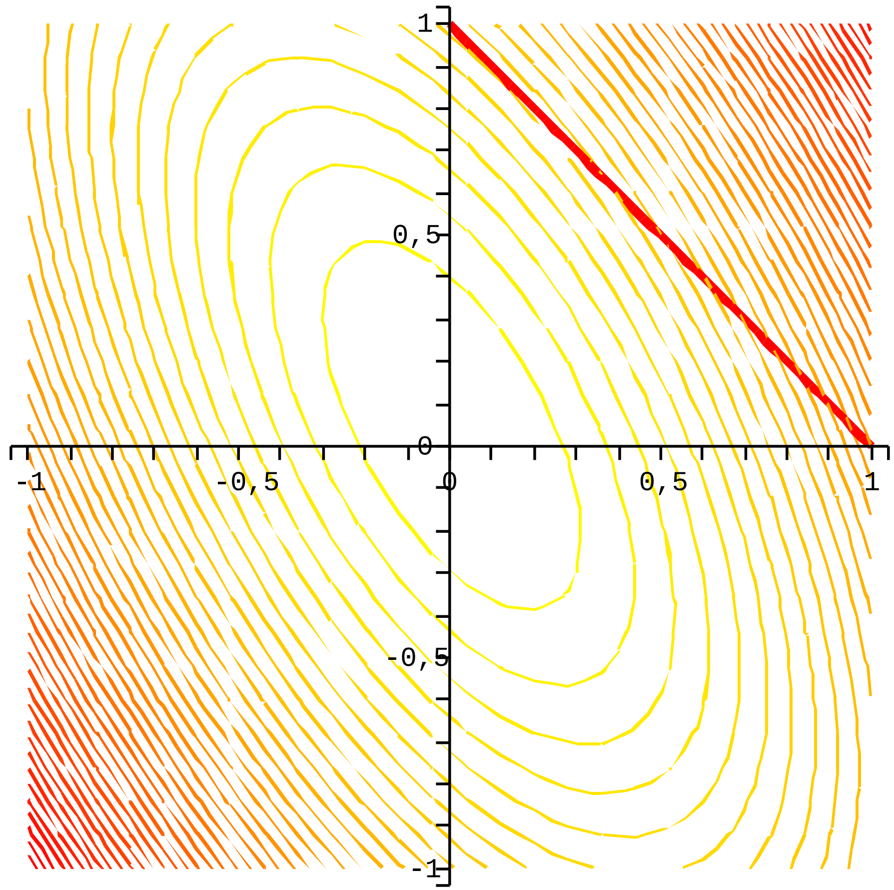
Levels of utility function

An ellipse is described by the length and direction of semiaxis.

Relation between covariance matrix Σ and the levels (ellipses) of the utility function:

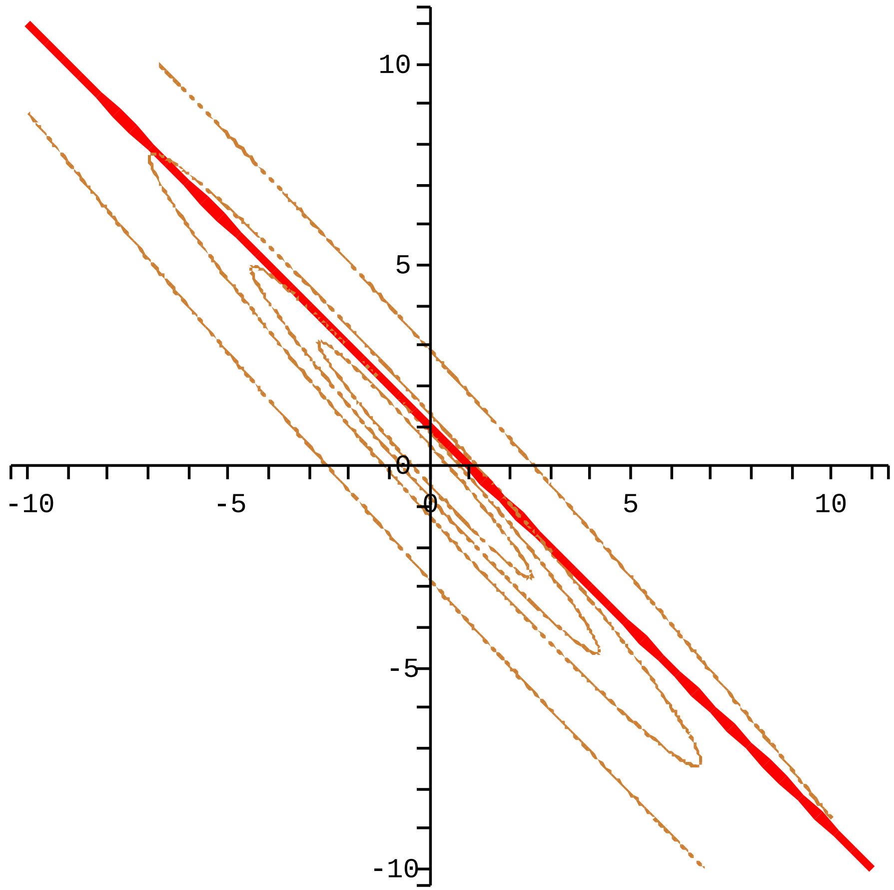
- Eigenvectors are directions of semiaxis.
- Eigenvalues are inverse-proportional to the length of semiaxis.

Levels of utility function



Small change in eigenvectors gives small change in portfolio.

Levels of utility function



Very small eigenvalue gives a long cigar.
In this case portfolio is very sensitive to small changes in eigenvectors.

Covariance matrix estimates – conclusions

- When covariance matrix eigenvalues are of similar magnitude small estimation errors don't produce substantial changes in the optimal portfolio composition.
- When covariance matrix eigenvalues differ substantially even small errors in estimation can change optimal portfolios completely.
- A small eigenvalue means the small variance of the corresponding asset (assets mix). Adding or subtracting even large amount of this asset doesn't change the value of the utility function. This explains one of the sources of portfolio instability.
- Putting ”no short sale” constrain improves stability.
- When the vortex of the utility surface is far away from the origin this also improves stability (this is in particular the case of the Black-Litterman model).

Factor models

Factor models assume that the returns are generated by specific factors.
The most important single factor is the market. The single factor model is then CAPM:

$$\mathbb{E}(\mathbf{R}) = \alpha + \beta R_M + \epsilon.$$

The covariance matrix for this model is given by the expression

$$\Sigma_{\mathbf{R}} = \sigma_M^2 \beta \beta^T + \Sigma_\epsilon$$

where σ_M^2 is the market variance and Σ_ϵ is the diagonal covariance matrix of uncorrelated residuals.
For the market of N assets we have to estimate $2N+1$ parameters, which is much less than for the sample covariance matrix ($N(N+1)/2$ parameters).

Principal Component Analysis (PCA)

- Make the spectral decomposition of the covariance matrix

$$\Sigma = \lambda_1 \mathbf{e}_1 \mathbf{e}_1^T + \cdots + \lambda_n \mathbf{e}_n \mathbf{e}_n^T.$$

- Neglect small eigenvalues in the decomposition

$$\Sigma \approx S = \lambda_1 \mathbf{e}_1 \mathbf{e}_1^T + \cdots + \lambda_k \mathbf{e}_k \mathbf{e}_k^T, \quad k < n.$$

- Perform optimization with covariance matrix Σ replaced by S (beware, S is singular).
- Choice of k based on stochastic matrix theory (Bengsson&Holst).

Shrinkage

Shrinkage estimators are constructed as a convex combination of sample covariance matrix Σ and a given matrix F

$$\delta F + (1 - \delta)\Sigma.$$

Here F can be thought as a Bayesian *prior* and Σ plays the role of an observation.

The choice of F is the fundamental difficulty of this approach. The following possibilities are popular:

1. Identity matrix.
2. Principal Component Model.
3. Main diagonal of the sample covariance matrix.
4. Factor models (in particular one-factor CAPM).

Shrinkage cont.

Shrinkage intensity δ (results of Ledoit&Wolf):

- Shrinkage intensity tends to zero as sample size T goes to infinity.
- Asymptotically optimal choice is given by $\delta \approx \kappa/T$ with constant κ.
- The paper of Ledoit&Wolf provides a consistent estimator for κ.

Which *prior* ?

All earlier mentioned possibilities has been used. Also combinations of some or all of them.

Known empirical tests:

- Shrinkage to identity (Ledoit&Wolf).
- Shrinkage to market (Ledoit&Wolf, Bengsson&Holst).
- Shrinkage to single factor PC (Bengsson&Holst).
- Shrinkage to diagonal (AP).
- Mix of sample, diagonal and market (Bengsson&Holst).
- Mix of sample, diagonal and PC (Bengsson&Holst, AP).

Estimation of covariance – conclusions

There is no single estimator eliminating all sources of errors.

Use several estimators which you trust and understand.

The portfolio selection process should take into account the optimization results for different estimators.

GARCH

GARCH is the must.

When you have to eliminate volatility clustering you have to use GARCH in some form.

GARCH

GARCH is the must.

When you have to eliminate volatility clustering you have to use GARCH in some form.

To learn ”how to do” we need another series of lectures.

GARCH

GARCH is the must.

When you have to eliminate volatility clustering you have to use GARCH in some form.

To learn ”how to do” we need another series of lectures.

Sorry

References

Bengtsson, C., Holst, J. – On portfolio selection: improved covariance matrix estimation for Swedish asset returns, Working paper, 2002.

Jaganathan, R., Ma, T. – Risk reduction in large portfolios: Why imposing the wrong constrains help, *J. Finance*, **58** (2003), 1651–1683.

Ledoit, O., Wolf, M. – Honey, I shrunk the sample covariance matrix, *J. Portfolio Manag.*, **30** No 4 (2004), 110-119.

Ledoit, O., Wolf, M. – Improved estimation of the covariance matrix of stock returns with an application to portfolio selection, *J. Empirical Fin.*, **10** (2003), 603–621.

Performance Analysis

Outline

1. Tactical asset allocation.
2. Performance evaluation.
3. Performance measures.
4. Performance attribution.
5. Performance appraisal.

Tactical asset allocation

Tactical asset allocation refers to active strategies which seek to enhance performance by shifting the asset mix in response to the changing patterns reward in the capital markets.

We speak about tactical allocation, when we have already decided about strategic asset allocation for a long period of time (benchmark or benchmarks) and are now looking for short-time opportunities to add value.

Tactical asset allocation, cont.

Tactical asset allocation is a three-step process:

- forecasting assets expected returns (making bets),
- building optimal portfolios,
- testing their performance.

Performance of TAA means always comparison of the tactical (short-living) portfolio with the benchmark.

Efficient market hypothesis

- Strong hypothesis states that all currently known information is already reflected in security prices.

There is no additional information available to active managers to use in generating exceptional returns.
Active returns are completely random.

- Semistrong hypothesis states that all publicly known information is already reflected in security prices.

Active management skill is insider trading.

- Weak hypothesis states that only previous price-based information is reflected in security prices.

This rules out technical analysis but would allow for skillful active management based on fundamental and economic analysis.

Performance goals

- The goal of performance analysis is to distinguish skilled from unskilled investment management (separate skill from luck).
- Return-based performance analysis is the simplest method for analyzing return and risk.
- Portfolio-based performance analysis is a more sophisticated way to distinguish skill and luck.
- Performance analysis is most valuable to the client when there is an *ex ante* agreement on the manager's goals.
- Performance analysis is valuable to the manager in that it lets the manager see which management decisions are compensated and which are not.

Performance evaluation

Performance evaluation can be separated into three components:

- Performance measurement.

The core of performance evaluation is the measurement of portfolio performance in comparison to the benchmark. Various measures of performance have been introduced and used. They focus on measuring portfolio return and risk and/or the trade-off between these two characteristics.

- Performance attribution.

Managers and clients need to understand how the total performance was reached. Performance attribution attributes major investment decision taken by the manager to the portfolio performance. To quantify performance attribution, detailed calculations need to be performed.

Performance evaluation, cont.

- Performance appraisal.

Clients wish to know if their managers possess true management skills. Performance appraisal formulates some judgment on the investment manager's skills. It requires to look at performance over longer horizons, taking into account the risk borne.

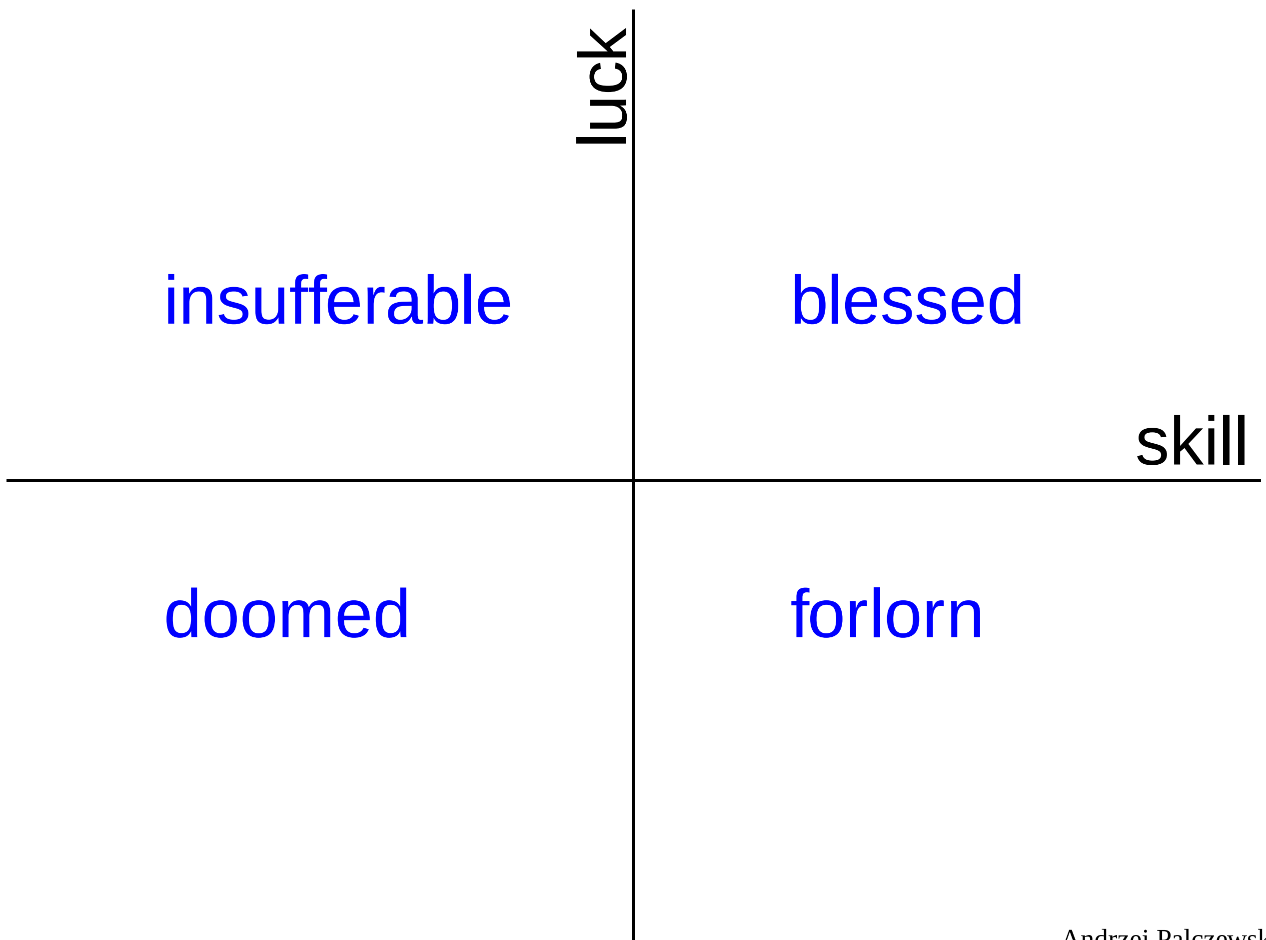
Performance evaluation, cont.

The performance appraisal should answer the following questions:

- Has the manager provided a good risk-adjusted performance over a long-run horizon?
- How does the manager compare with a peer group (universe of managers)?
- Is the performance due to luck, higher risk taken, or true investment skills?
- Is there evidence of unusual expertise and added value in a particular market?

Skill and luck

The fundamental goal of performance analysis is to separate skill from luck.



Methods of performance analysis

- Return-based analysis.

Return-based analysis is a top-down approach to attribute returns to components. This analysis is performed *ex post* based on realized returns of the portfolio and benchmark. The approach is based on CAPM and statistical analysis of the manager's added value.

- Portfolio-based analysis.

Portfolio-based performance analysis is a bottom-up approach, attributing returns to many components based on the *ex ante* portfolio holdings and giving managers credit for returns along these components. This allows the analysis not only of whether the manager has added value, but also of whether she has added value along dimensions agreed upon *ex ante*. In addition, portfolio-based analysis gives tools for analyzing *ex ante* separate bets.

Performance measures

Many performance measures are developed from the CAPM relation

$$\hat{\mu}_i = \beta_i \mathbb{E}(\hat{R}_M),$$

where

- $\hat{\mu}_i$ is the excess return of asset i,
- $\mathbb{E}(\hat{R}_M)$ is the excess return of market portfolio (benchmark).

Return

- Return is the simplest and most obvious performance measure.
- Return means always excess return

$$\text{excess return} = \text{return} - \text{risk-free rate}$$

- Return can be calculated using different formulas: arithmetic average, geometric average, average log return etc.
- As a performance measure we can take:
  - total return of the portfolio,
  - active return: return of the portfolio over return of the benchmark.

Sharpe ratio

Let **B** be the benchmark portfolio and $\hat{\mu}_{B,t}$ its excess return at period t.
Let **TAA** be a tactical portfolio with excess return $\hat{\mu}_{TAA,t}$.

$$SR = \text{Sharpe Ratio} = \frac{\hat{\mu}_{TAA,t}}{\hat{\sigma}_{TAA,t}}$$

where
$\hat{\sigma}_{TAA,t}$ is the standard deviation of excess returns of portfolio **TAA**.

Alpha (Jensen's alpha)

Regression of $\hat{\mu}_{TAA,t}$ against $\hat{\mu}_{B,t}$ gives

$$\hat{\mu}_{TAA,t} = \alpha_t + \beta_t \hat{\mu}_{B,t} + \epsilon_t.$$

The intercept in this relation which is the excess return due to active decisions is known as alpha.

This alpha is called the *realized* or *historical* alpha and is used to evaluate menager's skills.

Alpha, cont.

For short time series of alphas a good approximation is

$$\alpha_t = \hat{\mu}_{TAA,t} - \hat{\mu}_{B,t}.$$

This alpha can be used also in *ex ante* estimates.

Alpha calculated for a single period is to volatile for revealing the skill of the manager.

We have to smooth alpha by averaging over several periods and annualize it.

Tracking error

Tracking error is the standard deviation of portfolio excess returns over benchmark excess returns

$$TE_t = \sqrt{\frac{1}{N-1}\sum_{t=1}^{N}\left((\hat{\mu}_{TAA,t} - \hat{\mu}_{B,t}) - \frac{1}{N}\sum_{t=1}^{N}(\hat{\mu}_{TAA,t} - \hat{\mu}_{B,t})\right)^2}.$$

The above definition of tracking error can be applied for manager's skills evaluation.

Tracking error, cont.

In *ex ante* evaluation tracking error is replaced by *active risk*.

Let $\mathbf{w}_{B,t}$ be asset's weights in the benchmark portfolio **B** and $\mathbf{w}_{TAA,t}$ asset's weights in tactical portfolio **TAA**.

Let Σ be the covariance matrix of assets in both portfolios. Then the active risk (tracking error) is defined as follows

$$TE_t = (\mathbf{w}_{TAA,t} - \mathbf{w}_{B,t})^T \Sigma (\mathbf{w}_{TAA,t} - \mathbf{w}_{B,t}).$$

Information ratio

$$IR = \text{Information Ratio} = \frac{\text{alpha}}{\text{tracking error}}$$

$$IR_Y = \text{Annualized } IR = \frac{\text{annualized alpha}}{\text{annualized tracking error}}$$

Alpha and tracking error depend on the aggressiveness of the manager.

Information ratio is more or less aggressiveness independent.

Hit ratio

Hit ratio is a parameter which can be evaluated only *ex post* after collecting the long time series of manager's results.

$$HR = \text{Hit Ratio} = \frac{\text{number of periods the manager adds value}}{\text{number of all periods}}$$

In this measure the degree of success is ignored and only the frequency of success is measured.

Hit ratio, cont.

60% is already a very good result for the hit ratio.
On the other hand only 100% hit ratio can guarantee that the manager adds value.

If alpha is normally distributed with the mean being the arithmetic average alpha and standard deviation being the tracking error then

$$HR = 1 - \Phi(-IR),$$

where $\Phi(x)$ is a cumulative distribution function of standard normal distribution $N(0, 1)$.

Performance measures – summary

- Return measures:
  - Return.
  - Jensen's alpha.
- Risk measures:
  - Tracking error.
  - Active risk.
- Risk adjusted measures:
  - Sharpe ratio.
  - Information ratio.

Performance attribution

Performance attribution looks at the portfolio return over a single period and attributes it to factors.

We are not limited in choosing factors. They can be sector, industry or market indexes, but also investment themes such as value or momentum. In portfolios of limited size factors can be just single assets.

For large portfolios we can distinguish the following steps in performance attribution:

Performance attribution, cont.

- Security selection.

This is a manager's ability to isolate returns of the various segments (factors).

- Asset allocation.

This is a manager's ability to pick individual assets, after controlling for the segments (factors).

- Market (benchmark) timing.

Benchmark timing is an active management decision to vary the managed portfolio's beta with respect to the benchmark. If we believe that the benchmark will do better than usual, then beta is increased. If we believe the benchmark will do worst than usual, then the beta should be decreased.

Risk allocation

The total risk can be decomposed into the various risk exposures. This leads to a better understanding of the total risk borne.

The total risk is the result of decisions at two levels:

- The absolute risk allocation to each asset class.

This is the asset allocation approach, in which the risk of each asset class is measured using benchmarks.

- The active risk allocation in each asset class.

This is the risk budgeted to generate alphas in each asset class.

A risk decomposition can be quite complex, because risks are correlated and, hence, non additive.

Performance attribution – example

Alpha and tracking error are characteristics of the whole tactical portfolio.

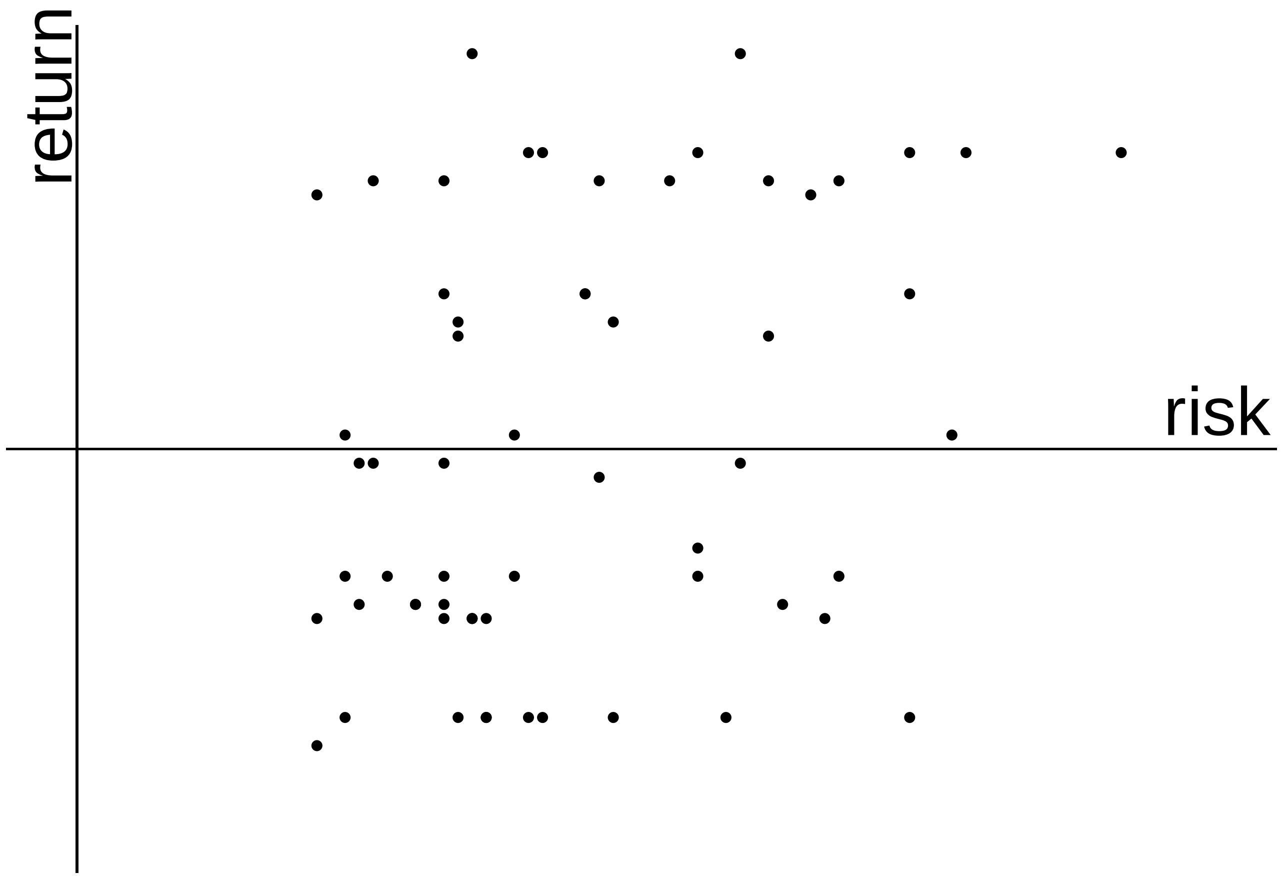
For medium size portfolio we can split both these measures into contributions coming from single assets.

On the next slide you will see an example of such split (all values are expressed in bp).

Performance attribution – example

Assets	active risk	attr. return
AS1	0.48	1.81
AS2	1.74	6.5
AS3	2.4	8.79
AS4	7.55	27.62
AS5	2.53	7.54
AS6	9.54	28.19
AS7	-19.89	-50.17
AS8	17.07	41.42
AS9	84.17	111.04
AS10	-7.51	-4.86
AS11	-0.86	-3.04
Portfolio	129.58	184.13

Performance appraisal



Realized alpha may result from the level of risk taken by the manager, rather than from true investment skills.

Performance appraisal, cont.

Because short-term results can be due to luck, rather then skills, a long horizon must be used.

For normal market, variance of IR has the following approximation

$$Var(IR) \approx \frac{1}{Y},$$

where Y is the number of years of observations.

It implies that to determine with high confidence (95%) that a manager belongs in the top quantile ($IR = 0.5$) requires

16 years of observations !

Performance appraisal, cont.

How to distinguish skills from luck?

Probability of positive alpha in a given period of time for a manager with an information ratio $IR = 0.5$.

No of years	probability (in %)
0.1	56
1	70
2	75
3	80
5	87
10	94

Hence, even for top managers there is 20% chance that in a 3-year horizon they will have negative realized alpha.

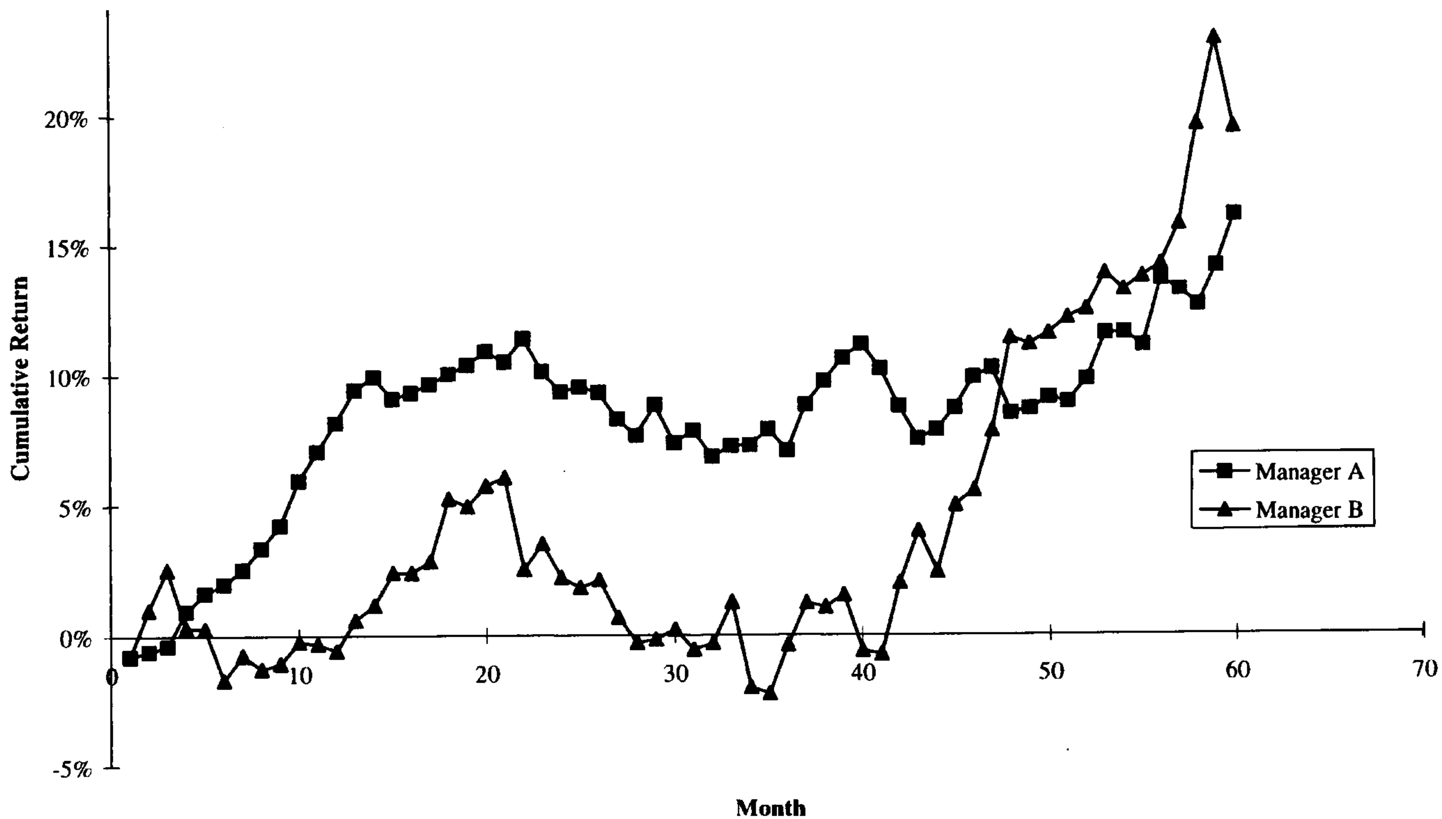
Performance appraisal, cont.

Compare two managers, **A** and **B**.

Over a 5-years period they achieved the following cumulative return (measured over benchmark):
manager **A** – 16%
manager **B** – 20%

Based on this limited set of information, whom would you prefer?

Performance appraisal, cont.



Grinold&Kahn(2000)

References

Grinold, R., Kahn, R. – *Active Portfolio Management*, McGraw-Hill 2000.

Jensen, M. C.– The performance of mutual funds in the period 1945–1964, *J. Finance*, **23** (1968),389–416.

Lee, W. – *Advanced Theory and Methodology of Tactical Asset Allocation*, Wiley 2000.

Prigent, J-L. – *Portfolio Optimization and Performance Analysis*, Chapman&Hall 2007.

Sharpe, W. – Mutual funds performance, *J. Business*, **39** (1966), 119–138.