The Effect of Estimation in High–dimensional Portfolios

Luitgard A. M. Veraart

Karlsruhe Institute of Technology

Joint work with Axel Gandy, Imperial College London

6th World Congress of the Bachelier Finance Society
Toronto, June 2010

Outline

1. Classical Portfolio Optimisation
2. Plug-In Strategies with Estimated Parameters
3. James-Stein-Shrinkage Applied to Strategies
4. L_1–Constrained Strategies - LASSO
5. Other Strategies
6. Application to Empirical Data

Optimal Portfolio Selection

- The asset prices: 1 bond $S_0(t) = e^{rt}$, d risky assets

$$dS_i(t) = S_i(t)[\mu_i dt + \sum_{j=1}^{d} \sigma_{ij} dW_j(t)], \quad S_i(0) > 0, \quad i = 1, \ldots, d,$$

  $r > 0$ interest rate, $\mu \in \mathbb{R}^d$ drift, $\sigma \in \mathbb{R}^{d \times d}$ volatility matrix of full rank (all constant), W d-variate Brownian motion.
- Investor has $T > 0$ fixed time horizon, $X_0 > 0$ constant initial wealth, chooses $\pi_i(t)$ fraction of the wealth invested in the ith asset at time t, resulting in time-t-wealth X_t with $dX_t = \sum_{i=0}^{d} \pi_i(t) X_t \frac{dS_i(t)}{S_i(t)}$.
- Investor seeks π to maximise $V(\pi) := \mathbb{E}[\log(X_T)]$.
- Optimal solution

$$\pi^* = \Sigma^{-1}(\mu - r1),$$

  where $\pi_0(t) = 1 - \sum_{i=1}^{d} \pi_i(t)$, $\pi = (\pi_1, \ldots, \pi_d)^T$, $\Sigma = \sigma\sigma^T$.

Optimal Portfolio Selection

- The asset prices: 1 bond $S_0(t) = e^{rt}$, d risky assets

$$dS_i(t) = S_i(t)[\mu_i dt + \sum_{j=1}^{d} \sigma_{ij} dW_j(t)], \quad S_i(0) > 0, \quad i = 1, \ldots, d,$$

  $r > 0$ interest rate, $\mu \in \mathbb{R}^d$ drift, $\sigma \in \mathbb{R}^{d \times d}$ volatility matrix of full rank (all constant), W d-variate Brownian motion.
- Investor has $T > 0$ fixed time horizon, $X_0 > 0$ constant initial wealth, chooses $\pi_i(t)$ fraction of the wealth invested in the ith asset at time t, resulting in time-t-wealth X_t with $dX_t = \sum_{i=0}^{d} \pi_i(t) X_t \frac{dS_i(t)}{S_i(t)}$.
- Investor seeks π to maximise $V(\pi) := \mathbb{E}[\log(X_T)]$.
- Optimal solution

$$\pi^* = \Sigma^{-1}(\mu - r1),$$

  where $\pi_0(t) = 1 - \sum_{i=1}^{d} \pi_i(t)$, $\pi = (\pi_1, \ldots, \pi_d)^T$, $\Sigma = \sigma\sigma^T$.

The Problem

- What if we need to estimate μ?
- What if the number of risky assets $d \to \infty$?

Plug-in Merton Strategy with Estimated μ

General Unbiased Plug-in Estimator

- Estimate μ by $\hat{\mu}$ and define the plug-in strategy $\hat{\pi} = \Sigma^{-1}(\hat{\mu} - r1)$.
- Assume $\hat{\pi} \sim N(\Sigma^{-1}(\mu - r1), V_0^2)$, $V_0 \in \mathbb{R}^{d \times d}$, then

$$V(\hat{\pi}) = V(\pi^*) - \frac{T}{2} \operatorname{trace}(\Sigma V_0^2).$$

Specific Plug-in Estimator

- Observation period $[-t_{est}, 0]$ for $t_{est} > 0$.
- Set

$$\hat{\mu}_i = \frac{\log(S_i(0)) - \log(S_i(-t_{est}))}{t_{est}} + \frac{1}{2} \sum_{j=1}^{d} \sigma_{ij}^2.$$

- Then $\hat{\pi} \sim N(\Sigma^{-1}(\mu - r1), \Sigma^{-1}/t_{est})$.
- $V(\hat{\pi}) = V(\pi^*) - d\frac{T}{2t_{est}}$.
- There are realistic scenarios in which even $V(\hat{\pi}) \to -\infty$ as $d \to \infty$.

James-Stein-Type Shrinkage of the Strategy

The James-Stein-Strategy

Let $\hat{\pi} = \Sigma^{-1}(\hat{\mu} - r1)$, $\pi^0 \in \mathbb{R}^d$, $a > 0$ fixed constants. Consider

$$\hat{\pi}^{JS,\pi^0} = \left(1 - \frac{a}{(\hat{\pi} - \pi^0)^T \Sigma (\hat{\pi} - \pi^0)}\right)(\hat{\pi} - \pi^0) + \pi^0.$$

The Expected Utility for the JS-Strategy

Let $\hat{\mu} \sim N(\mu, \Sigma/t_{est})$, $K \sim \text{Poisson}(\lambda)$, $\lambda = (\pi^* - \pi^0)^T \Sigma (\pi^* - \pi^0)/2$:

$$V(\hat{\pi}^{JS,\pi^0}) = V(\hat{\pi}) + \frac{T}{2} a \left[2 \frac{d-2}{t_{est}} - a\right] \mathbb{E}\left[\frac{t_{est}}{d - 2 + 2K}\right].$$

$\hat{\pi}^{JS,\pi^0}$ dominates $\hat{\pi}$ for $0 < a < 2(d-2)/t_{est}$; optimal $a = (d-2)/t_{est}$.

Special Choices for π^0 and Optimal a

- $\pi^0 = \pi^*$: $V(\hat{\pi}^{JS,\pi^0}) = V(\pi^*) - \frac{T}{t_{est}}$.
- $\pi^0 = \frac{\beta}{d} 1$, $\beta \in \mathbb{R}$: In some situations $V(\hat{\pi}^{JS,\pi^0}) \to \infty$ as $d \to \infty$.

L_1–constrained Strategies - LASSO

General Idea

- Require that π satisfies $\|\pi\|_1 = \sum_{i=1}^{d} |\pi_i| \leq c$ for a constant $c \geq 0$.
- $V(\pi) \geq \log(X_0) + rT - T\mathbb{E}\left\{c \max_i |\mu_i - r| + \frac{c^2}{2} \max_{i,j} |\Sigma_{ij}|\right\}$.
- If $\max_i |\mu_i - r|$, $\max_{i,j} |\Sigma_{ij}|$ bounded, $V(\pi) \not\to -\infty$ as $d \to \infty$.

Specific Results

For $\Sigma = \eta^2(\rho 11^T + (1-\rho)I)$, $\eta > 0, 0 \leq \rho \leq 1$ analytic results for

- the optimal L_1-constrained strategies, if μ known.
- for the L_1-constrained plug-in strategy as $d \to \infty$:
  - the distribution of $\#\{i : \pi_i^* \neq 0\}$, if $\rho = 0$,
  - an upper bound on $\lim_{d\to\infty} \mathbb{P}(\#\{i : \pi_i^* \neq 0\} > k)$, if $\rho > 0$.

Other Strategies and Performance for $d \to \infty$

Other Norm Constraints

- L_0-restricted strategies: no degeneration of expected utility as $d \to \infty$.
- L_2-restricted strategies: degeneration possible.

Special L_1-Constraints

- $1/d$-strategy: Strategy that invests the same amount into all stocks, i.e. $\pi^{c/d} = \frac{c}{d}1$ for some $c > 0$.
- Equal Weighting of the most Extreme stocks (EWE):

$$\pi_{k_i}^{\text{EWE}} = \frac{c}{\beta d} \operatorname{sign}(\hat{a}_{k_i}) \mathbb{I}(i \leq \beta d), \quad i = 1, \ldots, d,$$

where $\hat{a}_i = \frac{\hat{\mu}_i - r}{\Sigma_{ii}}$, $c > 0$, $\beta \in (0,1)$ constants, k_i are such that $|\hat{a}_{k_1}| > |\hat{a}_{k_2}| > \cdots > |\hat{a}_{k_d}|$.

Example - Trading S&P500

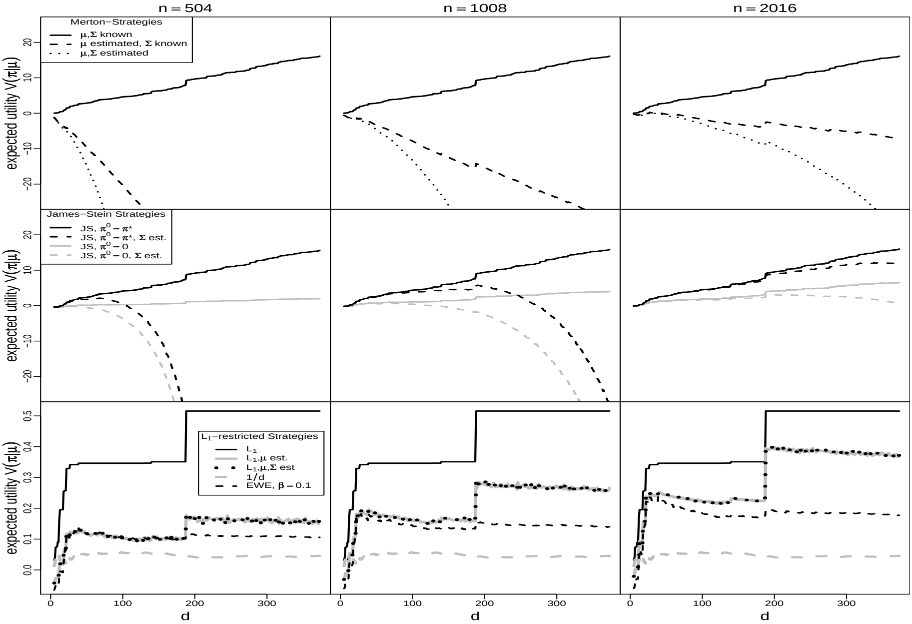
- Stocks in S&P 500 index on 01/01/2006 having daily returns for all trading days between 2001 and 2008 (373 stocks, n=2011 trading days, daily returns).
- Specific random ordering of stocks. Allow the strategies to invest in the first d stocks of this ordering.
- $X_0 = 1$, $r = 0.02$, roughly $T = 1$.
- Use of unbiased estimators based on observed stock prices at time points $0, \Delta, 2\Delta, \ldots, (n-1)\Delta$:

$$\hat{\mu}^{\text{data}} = \frac{1}{\Delta}\hat{\xi} + \frac{1}{2}\operatorname{diag}(\widehat{\Sigma}^{\text{data}}),$$

$$\widehat{\Sigma}^{\text{data}}_{\mu,\nu} = \frac{1}{\Delta(n-2)}\sum_{i=0}^{n-2}\left[R_\mu(i) - \hat{\xi}_\mu\right]\left[R_\nu(i) - \hat{\xi}_\nu\right]$$

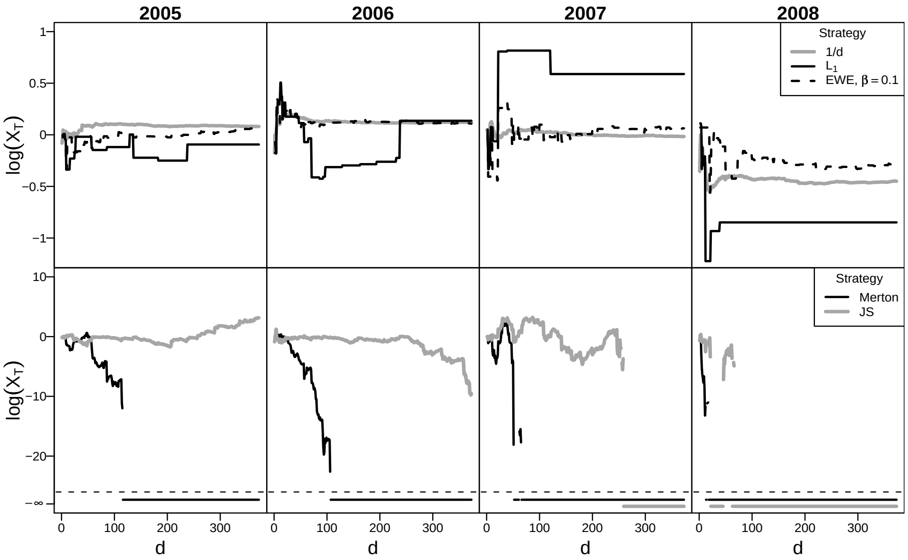
for $\mu, \nu = 1, \ldots, d$, where $R_\mu(i) = \log\left(\frac{S_\mu((i+1)\Delta)}{S_\mu(i\Delta)}\right)$,
$\hat{\xi}_\mu = \frac{1}{n-1}\sum_{i=0}^{n-2} R_\mu(i)$.

Analytic and Simulation Results



Expected utility plotted against the number d of available stocks.

Out-of-Sample Performance



$\log(X_T)$ with $T = 1$ year plotted against the number d of available stocks.

Summary

Main Contributions

- Quantification of the effect of estimation in vast portfolios (unknown μ and large d).
- Analysis of strategies which are less affected by estimation.
- Analytic formulae for James-Stein and optimal L_1-constrained strategies.

Specific Conclusions

- Estimation effects must not be ignored in vast portfolios!
- Simple plug in strategies have a loss through estimation linear in d.
- James-Stein shrinkage performs better than simple plug in strategies.
- L_1-constrained strategies cannot degenerate.
- L_1-constrained strategies and particularly the EWE-strategy and $1/d$ strategy perform well also in out-of-sample tests.

References

Brodie, J., Daubechies, I., De Mol, C., Giannone, D. & Loris, I. (2009). Sparse and stable Markowitz portfolios. *P. Natl. Acad. Sci.* **106**, 12267–12272.

Fan, J., Zhang, J. & Yu, K. (2009). Asset allocation and risk assessment with gross exposure constraints on vast portfolios. Preprint, Princeton University.

James, W. & Stein, C. (1961). Estimation with quadratic loss. In *Proc. Fourth Berkeley Symp. on Math. Statist. and Prob., Vol. 1 (Univ. of Calif. Press)*, 361–379.

Merton, R. (1971). Optimum consumption and portfolio rules in a continuous-time model. *J. Econ. Theory* **3**, 373–413.

Stein, C. M. (1981). Estimation of the mean of a multivariate normal distribution. *Ann. Statist.* **9**, 1135–1151.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *J. Roy. Stat. Soc. B* **58**, 267–288.

Analytic Results for LASSO with $\Sigma = \eta^2 I$

Optimal L_1-constrained strategy with known μ

Suppose $|\mu_1 - r| > |\mu_2 - r| > \ldots > |\mu_d - r|$. Then $\pi^\dagger = \frac{1}{\eta^2}(\mu - r1)$ is a solution to the L_1-constrained optimisation problem if $\|\pi^\dagger\|_1 \leq c$. Otherwise, the unique solution is

$$\pi^* := \frac{1}{\eta^2}(\text{sign}(\mu_1 - r)(|\mu_1 - r| - a), \ldots, \text{sign}(\mu_k - r)(|\mu_k - r| - a), 0, \ldots, 0)^T,$$

$k = \min\left\{l \in \{1, \ldots, d\} : c \leq \frac{1}{\eta^2} \sum_{i=1}^{l} (|\mu_i - r| - |\mu_{l+1} - r|)\right\}$,
$a = \frac{1}{k}\left[\eta^2 c - \sum_{i=1}^{k} |\mu_i - r|\right]$ and $\mu_{d+1} = r$.

Plug-in strategy with iid normally distributed estimators $\hat{\mu}_1, \ldots, \hat{\mu}_d$

Let $c = \alpha c_d$, $c_d > 0$ norming constants from extreme value theory for folded normal distribution $FN(\mathbb{E}(\hat{\mu}_1 - r), \mathbb{V}ar(\hat{\mu}_1))$. Then

$$\#\{i : \pi_i^* \neq 0\} \xrightarrow{\mathcal{L}} K + 1 \quad (d \to \infty),$$

where K is a Poisson distribution with expected value $\alpha\eta^2$.