The EF and the estimation error

Section 1

The EF as a random variable

A biased estimator

There are two main problems, from a statistical point of view, with the portfolio frontier.
First, the portfolio frontier is a biased estimator: the estimated frontier is usually more optimistic than the frontier computed with the true, unknown, values. The true frontier is in its turn dominant with respect with the out of sample frontier, which is obtained from the estimated weights and the true parameters.
Second, the portfolio frontier is highly affected by the estimation error of the parameters, especially the means, resulting in low efficiency. This is the most relevant problem.

Resampling

Resampling allows to take into account the statistical error: this leads to smoother changes in composition and at some extent it improves the out of sample performance.

1. Estimate the parameters from the data and compute N portfolio along the efficient frontier
2. Resample the data and compute again the portfolios
3. After a suitable number of iterations, average the weights of the portfolios. Portfolios are associated by rank (alternatively the λ coefficient in $\mu_p - \lambda\sigma_p^2$ can be used).

- The sample size can be adjusted to reflect the confidence in the estimates
- The resampled frontier can be non monotonic
- Certain constraints can be violated by the resampled frontier (e.g. a constraint on the number of stocks)

Resampled frontier

The technique can be employed to determine whether a portfolio is significantly different from another, thus avoiding costly and unnecessary rebalancing. First of all, define the distance between two portfolios $\boldsymbol{w}$ and $\boldsymbol{w}_0$ as $(\boldsymbol{w} - \boldsymbol{w}_0)'\boldsymbol{\Sigma}(\boldsymbol{w} - \boldsymbol{w}_0)$.

If you want to compare two portfolios from the classical frontier, compute the resampled portfolios associated to $\boldsymbol{w}$, then compute the distance between each of them and $\boldsymbol{w}_0$. Finally compute the quantile corresponding to the distance between $\boldsymbol{w}$ and $\boldsymbol{w}_0$.

If you want to compare two resampled portfolios, perform a nested resampling for each of the frontier you averaged to obtain the new resampled portfolio. Then compute the distance between the new portfolio and each of the associated portfolios of the two times resampled frontiers.

Section 2

Improving the estimates

Stein Estimator

The sample mean is unbiased but highly inefficient. As a consequence the estimated expected returns are noisy and variable, inducing variability on the frontier. James-Stein is an alternative estimator. Let's assume we have a sensible view on the expected returns, say $\boldsymbol{\mu}_{tgt}$. Then the James-Stein estimator is given by:

$$\bar{\boldsymbol{\mu}} = (1-\gamma)\hat{\boldsymbol{\mu}} + \gamma \boldsymbol{\mu}_{tgt}$$

where

$$\gamma = \max\left\{0, \min\left\{1, \frac{1}{T} \frac{\text{tr}(\boldsymbol{\Sigma}) - 2\rho_1^2}{(\hat{\boldsymbol{\mu}} - \boldsymbol{\mu}_{tgt})'(\hat{\boldsymbol{\mu}} - \boldsymbol{\mu}_{tgt})}\right\}\right\}$$

and ρ_1^2 represents the biggest eigenvalue of $\boldsymbol{\Sigma}$, T is the sample size.

Ledoit & Wolf estimator

The sorted eigenvalues of the estimated covariance matrix are steeper than the true unknown eigenvalues. Defining the condition number as $\frac{\rho_1}{\rho_N}$ we observe higher condition number for the estimated covariance matrix than for the true one. To reduce the problem one can resort to Ledoit & Wolf estimator for the covariance matrix. Given a normally distributed time series ϵ_t define $\mathbf{\Sigma}_{tgt} = \frac{1}{N}\mathrm{tr}(\hat{\mathbf{\Sigma}})\mathbb{I}$. The estimator reads:

$$\bar{\mathbf{\Sigma}} = (1-\gamma)\hat{\mathbf{\Sigma}} + \gamma\mathbf{\Sigma}_{tgt}$$

and

$$\gamma = \max\left\{0, \min\left\{1, \frac{1}{T}\frac{\frac{1}{T}\sum_{t=1}^{T}\mathrm{tr}[(\epsilon_t\epsilon_t' - \hat{\mathbf{\Sigma}})^2]}{\mathrm{tr}[(\hat{\mathbf{\Sigma}} - \frac{1}{N}\mathrm{tr}(\hat{\mathbf{\Sigma}})\mathbb{I})^2]}\right\}\right\}$$

References

Richard O. Michaud, Robert O. Michaud, Efficient asset management,Oxford university press, 2008