Chapter 4.

Portfolio Theory

This chapter of introduces Markowitz's portfolio theory (Markowitz, 1952) based on the mean variance principle. Even though the theory is simple and elegant, its implementation can be quite difficult because of, among others, the statistical estimation risk. In comparison, we also present some elementary or naive portfolio management methods. These portfolio management methods can be also be broadly viewed as tools for trading strategies.

4.1. The constant rebalanced portfolio and the naive Talmud strategy

The naive Talmud rule allegedly dates back to 200 years B.C. It is very simple: always keeping the distribution of portfolio equal weighted. An interesting example is the following. Suppose there are only two assets A and B. Asset A always stays flat, while asset B doubles on odd days and decreases by half on even days. Then, buy and hold either A or B will in the long run realize no significant profit, and it appears that there is no obvious strategy that realizes geometric growth. But the naive Talmud method can! At the beginning of any day, his portfolio is distributed half and half in assets A and B. At the end of the odd day after asset B doubled, his wealth increased by 50%, and he sells 1/4 of his asset B, and use it buy asset A. So in the beginning of the next day, an even day, his portfolio is till half and half. In the end of this day, after asset B shrank by half, his portfolio worth is only 75% of the worth in the morning, (decreased by 25%), and he sells 1/4 of his asset A and use it to buy asset B to keep the balanced half-half asset distribution... Then, within two consecutive days, the portfolio worth becomes $1.5 * 0.75 = 9/8$. In other words, it increased by $1/8 = 12.5\%$. Within a year of 250 trading days, the portfolio worth would increase $(9/8)^{125} = 2477795$ times.

In reality, there is no such perfect and deterministic scenario. However, the example serves the purpose of illustrating the idea behind the constant rebalanced portfolio, which in essense is a mean reverting strategy, as seen in the above over-simplified example. Even though it is indeed simple and naive, without optimization and/or risk control, it has been well documented in the modern finance study that the naive rule, also called the 1/N rule, can compete with the more sophisticated portfolio management rules (e.g., DeMiguel, Garlappi and Uppal, , 2005). The primary reason is: rules based on optimization generally involve the estimation risk, while the naive rules do not have the estimation step.

Suppose there are totally p assets, with prices $P_{t,1}, ..., P_{t,p}$, $t = 0, 1, ...$, and returns $\mathbf{x}_t = (x_{t1}, ...x_{tp})$, and $t = 1, 2, ...,$, where $x_{tj} = P_{t,j}/P_{t,j-1} - 1$. The constant rebalance portfolio extends the naive 1/N rule to a class of portfolios that always holds fixed proportion of each asset. Let the proportion be denoted by $\mathbf{a} = (a_1, ..., a_p)$ with $a_i \geq 0$ and $\sum_i a_i = 1$. The vector $\mathbf{a}$ can be used to indicate such a constant rebalanced portfolio. Notice that buy and hold one single asset, say asset i, is a special case of constant rebalanced portfolios, with $\mathbf{a} = (0, .., 0, 1, 0..., 0)$ where the 1 appear in the i-th entry.

It is natural to narrow the search space of the optimal portfolios to the constant rebalanced portfolios and try to find the optimal one. The universal portfolio (Cover, 1991) is aimed at this purpose, with

$$\mathbf{b}_{t+1} = \frac{\int_B \mathbf{b} W_t(\mathbf{b}) d\mathbf{b}}{\int_B W_t(\mathbf{b}) d\mathbf{b}} \tag{4.1}$$

Here $W_t(\mathbf{b}) = \prod_{j=1}^t \mathbf{b}^T(\mathbf{1} + \mathbf{x}_j)$ represent the portfolio worth of the constant rebalanced portfolio indexed by $\mathbf{b}$, and $B = \{\mathbf{b} \in R^p : \mathbf{b} \geq 0, \mathbf{b}'\mathbf{1} = 1\}$ refers to the collection of all such indexes. The universal portfolio places weights on the constant rebalanced portfoio strategy according to their past performance. Those with better performance receive higher weights. It is mathmatically proved that in the long run, the universal portfolio can do as well as the best performing single asset.

It is intersting to observe that, constant reblanced portfolio is a special case of mean reverting. While the univeral portfolio's selection of constant rebalanced portfolio based on past formance is special case of momentum or trend-following. Therefore it can be viewed as a mixture of mean reverting and momentum. Moreover, while constant rebalanced portfolio, particularly the 1/N rule, is regarded as a diversification (without optimization) over the assets, the universal portfolio can also be regarded as a diversication over a class of portfolio management methods, the constant reblanced portfolios. This view point and the practical implementation would be quite useful.

4.2. Markowitz's mean variance portfolio: simple cases.

Markowitz's mean variance portfolio uses the mean and variance of returns to calibrate the gain and risk of the assets or portfolios. For ease of presentation, we first consider two risky assets, i.e., $p = 2$. Let

$$\mu = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} \qquad \Sigma = \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{pmatrix}$$

be their mean and variances of the returns. Consider a portfolio with weights $\mathbf{w} = (w, 1-w)^T$, which places weight w on asset 1 and $1-w$ on asset 2. Then, the mean and variance of this portfolio are

$$\mu_P = \mathbf{w}^T \mu = w\mu_1 + (1-w)\mu_2 \quad \sigma_P^2 = \mathbf{w}^T \Sigma \mathbf{w} = w^2\sigma_{11} + 2w(1-w)\sigma_{12} + (1-w)^2\sigma_{22}.$$

Clearly, μ_P is a linear function of w and σ_P^2 is a quadratic function of w. Assume there is a risk free return asset F with risk free rate μ_f.

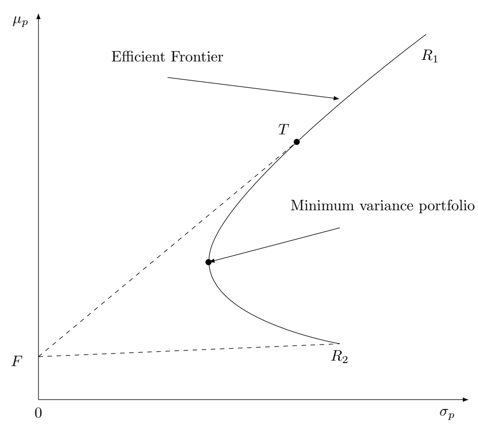

Figure 4.1

The above figure shows in the solide curve the variation of μ_P and σ_P along with w. The left most point of the curve refers to the portfolio with minimum variance, and hence called *minimum variance portfolio*. Upper portion of the curve starting from the minimum variance portfolio is called *efficient frontier*. Each point on the efficient frontier, called *efficient portfolio*, is a portfolio such that, it has higher return than any other portfolio with the same variance, or, in other words, it has smaller variance than any other portfolio with the same mean. The two ends of the curve correspond to the portfolios with $w = 0$ or $w = 1$, the portfolios with either asset 1 or asset 2. It is seen that portfolios in the lower part of the curve below minimum variance portfolio is inferior to the efficient frontier.

With the availability of risk free asset with risk free return rate μ_f, one can combine the risk free asset with any portfolio of the two risky assets to form a three assets portfolio. Recall that μ_P and σ_P^2 denote the mean and variance of a portfolio with the two risky assets. Suppose, for the new portfolio with three assets, the weight on the risk-free asset is $\tilde{w}$ and that on a portfolio of two risky asset is $1 - \tilde{w}$. Then, the mean and variance of this new portfolio is

$$\tilde{w}\mu_f + (1-\tilde{w})\mu_P \qquad (1-\tilde{w})^2\sigma_P^2$$

The mean and standard deviation are both linear function of $\tilde{w}$. Their variation along with $\tilde{w}$ is shown in the dotted straight lines. The end point of line at F is the portfolio entirely on risk free asset ($\tilde{w} = 1$), and the other end point of the line at the solid curve is the portfolio with two risky assets and no risk free asset ($\tilde{w} = 0$).

The slope of the dotted real line is

$$\frac{\tilde{w}\mu_f + (1-\tilde{w})\mu_P - \mu_f}{\sqrt{(1-\tilde{w})^2\sigma_P^2}} = \frac{\mu_P - \mu_f}{\sigma_P}.$$

It measures the amount of excess returns, over the risk free return, per unit risk. Here the risk is measured by the standard deviation. This quantity is called *Sharpe's ratio.* It is natural that the portfolio with highest Sharpe's ratio would be regarded as the best. Notice that all portfolios on each dotted line share the same Sharpe's ratio. The point on the solid curve with the highest Sharpe's ratio is called *tangency portfolio*, since it corresponds to the dotted line that is tangent to the curve and with the highest slope.

With these three assets, the efficient portfolio, those with highest Sharpe ratios, must be a combination of the tangency portfolio with the risk free asset. This is because any portfolio must be on the line connecting a point on the curve with F and its slope, the Sharpe ratio, must be lower than that of the line connecting the tangency portfolio with F. Again, the efficient portfolio, here with three assets, mean that, any portfolio with the same risk must have lower mean return, or any portfolio with same mean return must have higher risk.

That all efficient portfolios must be a combination of the risk free asset with the tangency portfolio implies that the composition of the risky assets remain the same for all efficient portfolios. The percentage of the risk free asset, which may vary, adjusts the returns and the risks of the efficient portfolios. The higher the risk, the higher the return.

4.3. Markowitz's mean variance portfolio: general cases.

Consider the genearl case with p assets. Let

$$\mu = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_p \end{pmatrix} \quad \text{and} \quad \Sigma = \begin{pmatrix} \sigma_{11} & \cdots & \sigma_{1p} \\ \vdots & \vdots & \vdots \\ \sigma_{p1} & \cdots & \sigma_{pp} \end{pmatrix}$$

be the mean and variance of the p assets. Let $\mathbf{w} = (w_1, ..., w_p)^T$ be the weights on the p-assets to form a portfolio, with $\mathbf{w}^T\mathbf{1} = \sum_{j=1}^p w_j = 1$. Then, the mean and variance of this portfolio is $\mathbf{w}^T\mu$ and $\mathbf{w}^T\Sigma\mathbf{w}$. Suppose we set a target of mean return μ_0 and wish to minimize the risk. It is an optimization problem:

$$\text{argmin}_{\mathbf{w}}\{\mathbf{w}^T\Sigma\mathbf{w}\} \qquad \text{subject to } \ \mathbf{w}^T\mu = \mu_0, \ \text{and } \mathbf{w}^T\mathbf{1} = 1. \tag{4.2}$$

A different but equivalent formulation is, we set a target risk of σ_0^2 and wish to maximize the mean return:

$$\text{argmax}_{\mathbf{w}}\{\mathbf{w}^T\mu\} \qquad \text{subject to } \ \mathbf{w}^T\Sigma\mathbf{w} = \sigma_0^2, \ \text{and } \mathbf{w}^T\mathbf{1} = 1.$$

Set $\lambda > 0$ as the risk aversion parameter. The larger the λ, the more risk averse. Then, another equivalent formulation is is

$$w^\star(\lambda) = \text{argmax}_{\mathbf{w}}\{\mathbf{w}^T\mu - \frac{\lambda}{2}\mathbf{w}^T\Sigma\mathbf{w}\} \qquad \text{subject to } \ \mathbf{w}^T\mathbf{1} = 1. \tag{4.3}$$

The solution, corresponding to the effient frontior, is, through Lagrange multiplier,

$$\mathbf{w}^\star(\lambda) = \frac{1}{\lambda}\Sigma^{-1}\mu + \frac{\lambda - d_2}{\lambda d_1}\Sigma^{-1}\mathbf{1}, \tag{4.4}$$

where $d_2 = \mathbf{1}^T\Sigma^{-1}\mu$ and $d_1 = \mathbf{1}^T\Sigma^{-1}\mathbf{1}$. The optimized mean returns and variances are, respectively

$$\frac{d_3 d_1 - d_2^2}{\lambda d_1} + \frac{d_2}{d_1} \qquad \text{and} \qquad \frac{d_3 d_1 - d_2^2}{\lambda^2 d_1} + \frac{1}{d_1} \tag{4.5}$$

where $d_3 = \mu^T\Sigma^{-1}\mu$. It follows from the quadratic optimization that all above three formulation lead to the same solutions, each depending on μ_0, σ_0 and λ.

Suppose now in addition to the p risky asset, we also have a risk free asset, asset 0, with risk free rate μ_f. Let the portfolio be $(w_0, \mathbf{w})$ with $w_0 + w_1 + ... + w_p = w_0 + \mathbf{w}^T\mathbf{1} = 1$. We can similarly consider the efficient portfolio as an optimization problem

$$\text{argmax}_{(w_0,\mathbf{w})}\{w_0\mu_f + \mathbf{w}^T\mu - \frac{\lambda}{2}\mathbf{w}^T\Sigma\mathbf{w}\} \qquad \text{subject to } \ w_0 + \mathbf{w}^T\mathbf{1} = 1. \tag{4.6}$$

It can be shown that the efficient portfolio is

$$w_0 = 1 + (1/\lambda)(\mu_f d_1 - d_2), \qquad \mathbf{w} = (1/\lambda)\Sigma^{-1}(\mu - \mu_f\mathbf{1}). \tag{4.7}$$

Here $\Sigma^{-1}(\mu - \mu_f\mathbf{1})$, with constant rescale, corresponds to the tangency portfolio. It can be computed that the mean return of the efficient portfolio is

$$\frac{1}{\lambda}(\mu - \mu_f\mathbf{1})^T\Sigma^{-1}(\mu - \mu_f\mathbf{1}) = \frac{1}{\lambda}(d_3 - 2\mu_f d_2 + \mu_f^2 d_1)$$

and the variance of the efficient portfolio is

$$\frac{1}{\lambda^2}(\mu - \mu_f\mathbf{1})^T\Sigma^{-1}(\mu - \mu_f\mathbf{1}) = \frac{1}{\lambda^2}(d_3 - 2\mu_f d_2 + \mu_f^2 d_1).$$

The sharpe's ratio of the efficient portfolio is then

$$\frac{(1/\lambda)(d_3 - 2\mu_f d_2 + \mu_f^2 d_1)}{\{(1/\lambda^2)(d_3 - 2\mu_f d_2 + \mu_f^2 d_1)\}^{1/2}} = \{(\mu - \mu_f\mathbf{1})^T\Sigma^{-1}(\mu - \mu_f\mathbf{1})\}^{1/2} = (d_3 - 2\mu_f d_2 + \mu_f^2 d_1)^{1/2}$$

Note that the Sharpe ratio of all efficient portfolios are irrelevant with the risk aversion parameter λ. The effect of λ is again adjusting the weight of the risk free asset.

Mathematically, with or without risk-free asset, the efficient portfolio problem can be formulated in the same way. In the set up of p risky assets, one can simply take the variance of one of the assets, say the first asset, as tending 0, its correlation with other assets as 0, and its return as the risk free rate μ_f. Then σ_{11} tend to 0 and $\sigma_{1j} = 0$ for $j \neq 1$. Taking limit, one gets exactly the same result.

4.4. Markowitz's mean variance portfolio with constriants.

The above general case of Markowitz's mean variance portfolio allows short sell. In other words, the weights can be negative. In reality, many assets are traded under constraints, typically the no-short-sell constraint. Under these constriants, the mean variance principle still applies. The associated optimization problem is with these constraints. In general, the analytic solutions with closed form are often not available, but with the aid of computer, numerical solutions are straightforward.

With the no-short-sell constriant, the weights must be nonnegative. And the efficient portfolio is the optimization of

$$\text{argmax}_{\mathbf{w}}\{\mathbf{w}^T\mu - \frac{\lambda}{2}\mathbf{w}^T\Sigma\mathbf{w}\} \qquad \text{subject to } \ \mathbf{w}^T\mathbf{1} = 1 \text{ and } w_i \geq 0, i = 1, ..., p.$$

A slightly more general constriant is constraining the short level, say 20%, which can be presented as

$$\text{argmax}_{\mathbf{w}}\{\mathbf{w}^T\mu - \frac{\lambda}{2}\mathbf{w}^T\Sigma\mathbf{w}\} \qquad \text{subject to } \ \mathbf{w}^T\mathbf{1} = 1 \text{ and } \sum_{j=1}^{n}|w_j| \leq 1.2.$$

In other words, with 1000 net asset you are allowed to own 1100 dollar in long position and 100 dollar in short position. The no short sell constraint is the same as $\sum_{j=1}^{p} |w_j| \le 1$.

Many funds restrict the holding of one single asset in their portfolios to be lower than a given threshold, say 5%. Then, the efficient portfolio is the optimization of

$$\text{argmax}_{\mathbf{w}}\{\mathbf{w}^T\mu - \frac{\lambda}{2}\mathbf{w}^T\Sigma\mathbf{w}\} \qquad \text{subject to } \ \mathbf{w}^T\mathbf{1} = 1 \text{ and } 0 \le w_i \le 5\%,\ i = 1, ..., p.$$

Further restrictions may be possible. For example, restricting the holding of a class of assets, such as real estate stocks, to be below a threshold, say 5%. This restriction is linear and can be presented as

$$\sum_{j\in A} w_j \le 5\%.$$

where A is the collection of all real estate stocks. Efficient portfolios under linear constraints can be solved easily using constrained quadratic optimization.

The no short sell constraint is often imposed by the market for many assets. However, in finance literature, it has been found that no short sell constraint can actually help to control risk (Jagannathan and Ma, 2003).

4.5. Statistical estimation.

The previous discusion of Markowitz mean variance portfolio is static and theoretical. In reality, the true population means and variances that we use in deriving the efficient portfolios and unknown. Worse, they are dynamic and may change with time.

When actually contructing Markowitz's efficient portfolio in practice, one faces a daunting task of estimating the future mean and variance of returns of all assets. Not surprisingly, many empirical studies show that efficient portfolios have poor performance, very often cannot beat even the naive portfolio. One of the main reason is the difficulty in accurately estimating future mean and variance of the returns of the assets. Afterall, accurate estimation of few, not all, of the assets would suffice for a successful portfolio, even without the mean variance optimazation.

The direct and obvious way of estimating the mean and variances of the p assets is using the sample mean and sample variance of assets in the past period of time, which we may call training period. Using $\mathbf{x}_j = (x_{j1}, ..., x_{jp})$, $j = t-1, t-2, ..., t-n$ as the returns for the past n periods. Assume they are iid with mean μ and varaince Σ. Then the sample mean is approximately normal with mean μ and variance Σ/n; and the sample variance also has mean Σ and is close to a Wishart distribution. Given sufficient sample size, the estimation can be quite accurate, justified easily using statistical asymptotic theory.

However, the problem in reality is that they are not iid random variables. And direct use of sample mean and sample variance are not favorable in empirical studies. Some statistical techniques can be applied to make certain improvements. One typical approach involve Stein's shrinkage: for estimating the mean returns of asset j, rather than using sample mean $\bar{x}_j$, one uses

$$\alpha\bar{x}_j + (1-\alpha)\bar{x},$$

where $\bar{x} = (1/p)\sum_{j=1}^{p} \bar{x}_j$ is the overall sample mean of all assets, and α is the shrinkage factor. The method can be heuristically understood as a conservative method: all returns are shrunk towards the overall mean. Higher/lower past return could, at least in part, resulted from certain random factors, or good/bad luck, that would not sustain in the future. Determining the shrinkage factors is an important research problem. By the same token, the variance estimator is often also shrunk towards the identity matrix.

In the literature, resampling methods or Bayesain methods are also applied with the purpose of improving the estimation of the mean and variance. The bootstrap method (poorman's Baysian method) is one of the most popular resampling methods in statistics. The reported results are mixed. Here we consider the problem of estimating the Sharpe ratio of the efficient portfolio, for the purpose to illustrate the method of bootstrap.

Suppose $\hat{\mu}$ and $\hat{\Sigma}$ are our estimates of the mean and variance of returns μ and Σ. They are not necessarily the sample mean and sample variance. Suppose $(\hat{w}_0, \hat{\mathbf{w}})$ is one estimated efficient portfolio. That is

$$(\hat{w}_0, \hat{\mathbf{w}}) = \text{argmax}_{(w_0,\mathbf{w})}\{w_0\mu_f + \mathbf{w}^T\hat{\mu} - \frac{\lambda}{2}\mathbf{w}^T\hat{\Sigma}\mathbf{w}\} \qquad \text{subject to} \;\; w_0 + \mathbf{w}^T\mathbf{1} = 1.$$

The estimated efficient portfolio is

$$\hat{w}_0 = 1 + (1/\lambda)(\mu_f\hat{d}_1 - \hat{d}_2), \qquad \hat{\mathbf{w}} = (1/\lambda)\hat{\Sigma}^{-1}(\hat{\mu} - \mu_f\mathbf{1}).$$

Here $\hat{\Sigma}^{-1}(\hat{\mu} - \mu_f\mathbf{1})$ corresponds to the estimated tangency portfolio, and $\hat{d}_1, \hat{d}_2$ and $\hat{d}_3$ are defined analogously. The estimated Sharpe's ratio is

$$\frac{\hat{w}_0\mu_f + \hat{\mathbf{w}}^T\hat{\mu} - \mu_f}{\{\hat{\mathbf{w}}^T\hat{\Sigma}\hat{\mathbf{w}}\}^{1/2}.} = (\hat{d}_3 - 2\mu_f\hat{d}_2 + \mu_f^2\hat{d}_1)^{1/2}.$$

But the actual Sharpe ratio is

$$\frac{\hat{w}_0\mu_f + \hat{\mathbf{w}}^T\mu - \mu_f}{\{\hat{\mathbf{w}}^T\Sigma\hat{\mathbf{w}}\}^{1/2}.}$$

In order to estimate the estimation error of the Sharpe ratio, one can conduct the boostrap method. Take, with replacement, n samples from the data $\{\mathbf{x}_{t-1}, ..., \mathbf{x}_{t-n}\}$, denote as $\{\mathbf{x}_1^*, ..., \mathbf{x}_n^*\}$ and compute

$$s_1^* = \frac{\hat{w}_0^*\mu_f + (\hat{\mathbf{w}}^*)^T(\hat{\mu}^* - \mu_f)}{\{(\hat{\mathbf{w}}^*)^T\hat{\Sigma}^*\hat{\mathbf{w}}^*\}^{1/2}} = (\hat{d}_3^* - 2\mu_f\hat{d}_2^* + \mu_f^2\hat{d}_1^*)^{1/2}$$

and

$$s_1 = \frac{\hat{w}_0^*\mu_f + (\hat{\mathbf{w}}^*)^T(\hat{\mu} - \mu_f)}{\{(\hat{\mathbf{w}}^*)^T\hat{\Sigma}\hat{\mathbf{w}}^*\}^{1/2}}.$$

Repeat this process M times, and obtain $s_2^*, s_2, ..., s_M^*, s_M$, with M being large. The distribution of $s_1^* - s_1, s_2^* - s_2, ..., s_M^* - s_M$ is used to approximate the distribution of the estimation error of Sharpe ratio.

4.6. Examples.

We selected 100 stocks that have been listed in the China A-stock Market since 2001. Their trade codes are shown in the following table. We will consider three trading frequencies: daily, weekly (5 trading days) and monthly (20 trading days). Two types of transaction costs, 0 and 0.004, will also be involved in the following examples.

Example 1

In this example, we show the empirical performance of the naive Talmud rule (section 4.1) for portfolio management. For comparison, we also consider the ruls of follow-the-leader and follow-the-loser:

- Equal weight: we will adjust portfolio so that every stock has equal weight at the beginning of each period.
- Follow-the-leader: buy equal weight of the 10% stocks (i.e. 10 stocks) that have the largest increment in the last period.
- Follow-the-loser: buy equal weight of the 10% stocks (i.e. 10 stocks) that have the largest decrease in the last observation period.

600005.SH	600082.SH	600123.SH	000023.SZ	000400.SZ
600006.SH	600084.SH	600125.SH	000027.SZ	000401.SZ
600008.SH	600085.SH	600127.SH	000037.SZ	000408.SZ
600016.SH	600095.SH	600128.SH	000039.SZ	000410.SZ
600019.SH	600097.SH	600130.SH	000040.SZ	000413.SZ
600038.SH	600098.SH	600132.SH	000045.SZ	000416.SZ
600054.SH	600100.SH	600133.SH	000048.SZ	000417.SZ
600056.SH	600103.SH	600135.SH	000058.SZ	000419.SZ
600059.SH	600104.SH	600136.SH	000059.SZ	000420.SZ
600061.SH	600105.SH	600138.SH	000060.SZ	000428.SZ
600063.SH	600106.SH	600145.SH	000061.SZ	000429.SZ
600064.SH	600110.SH	600146.SH	000065.SZ	000488.SZ
600066.SH	600112.SH	000002.SZ	000066.SZ	000505.SZ
600068.SH	600116.SH	000004.SZ	000069.SZ	000509.SZ
600069.SH	600117.SH	000007.SZ	000088.SZ	000511.SZ
600070.SH	600118.SH	000008.SZ	000089.SZ	000513.SZ
600073.SH	600119.SH	000014.SZ	000096.SZ	000514.SZ
600074.SH	600120.SH	000016.SZ	000150.SZ	000516.SZ
600078.SH	600121.SH	000018.SZ	000151.SZ	000519.SZ
600079.SH	600122.SH	000019.SZ	000159.SZ	000521.SZ

Table 1: Stock list

The performance of three portfolio strategies adjusted in the daily basis is shown in the figure below. Here we assume the initial account balance is 1. The dashed lines show the variation with transaction cost (for simplicity, we assume that 0.4% will be deducted when buying stocks, no more cost will be paid when selling stocks).

Figures 2 and 3 show performance with adjustments on the weekly and monthly frequency, respectively. Here are some comments for these plots:

- In the case of adjustment on daily frequency, the naive Talmud rule has s the best performance. The other two methods seem to be disastrous. This implies that pure momentum or mean-reverting (in the daily frequency) probably do not work well. Since the best or worst performing stocks usually change between different trading periods, much more transaction cost is carried with the follow-the-leader or follow-the-loser methods than with the equal weight method, which usually only need to tune the portfolio slightly each time.
- In the case of adjustment on weekly frequency, the results are similar to those on daily frequency. But this time follow-the-loser method performs a little better, although after considering transaction cost, positive returns disappear.
- In the case of adjustment on monthly frequency, follow-the-loser method performs much better than the other two methods. One possible reason is, in the long run, nearly all stocks have a trend of increasing. In the dataset, only 1 out of the 100 stocks with current price lower than 16 years ago. About 2/3 of the stocks have doubled their price during the time period. Monthly frequency can reflect this trend to some degree. If one stock is lagging behind in consecutive periods, it is likely to catch up with the market. The effect of transaction cost in the weekly and monthly transaction cases is lower, since there are fewer portfolio adjustments.

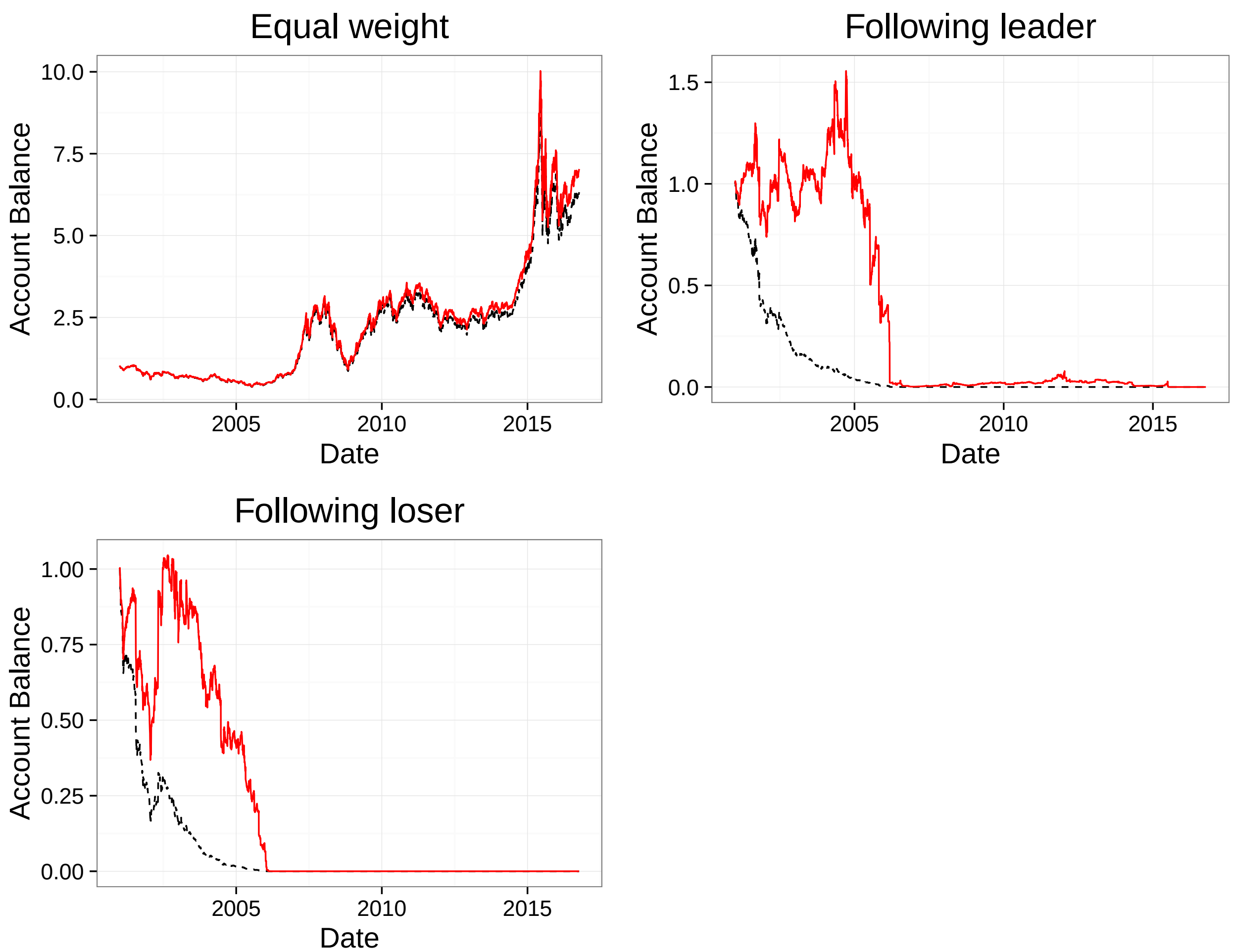

Figure 1: Naive rule application: daily frequency case

Example 2

Part 1

In this example, the same dataset as in Example 1 is used. We apply Markowitz's mean-variance principle to calculate the portfolio weights. Specifically, at each period, we will reweigh portfolios based on the sample mean and variance in previous 30 periods. Four methods with different constraints are considered:

- Method 1. Short selling allowed. (Note: to avoid unrealistic results, we constrain the short level to 20%)
- Method 2. Short selling not allowed.
- Method 3. Same as Method 1, but absolute value of weight for each stock must be less than 5%.
- Method 4. Same as Method 2, but absolute value of weight for each stock must be less than 5%.

Here we will use the function `solve.QP` in the package `quadprog` of `R` language for quadratic programming. Specifically, the quadratic objective function to be minimized is

$$\frac{1}{2}w^T(\Sigma)w - w^T\mu$$

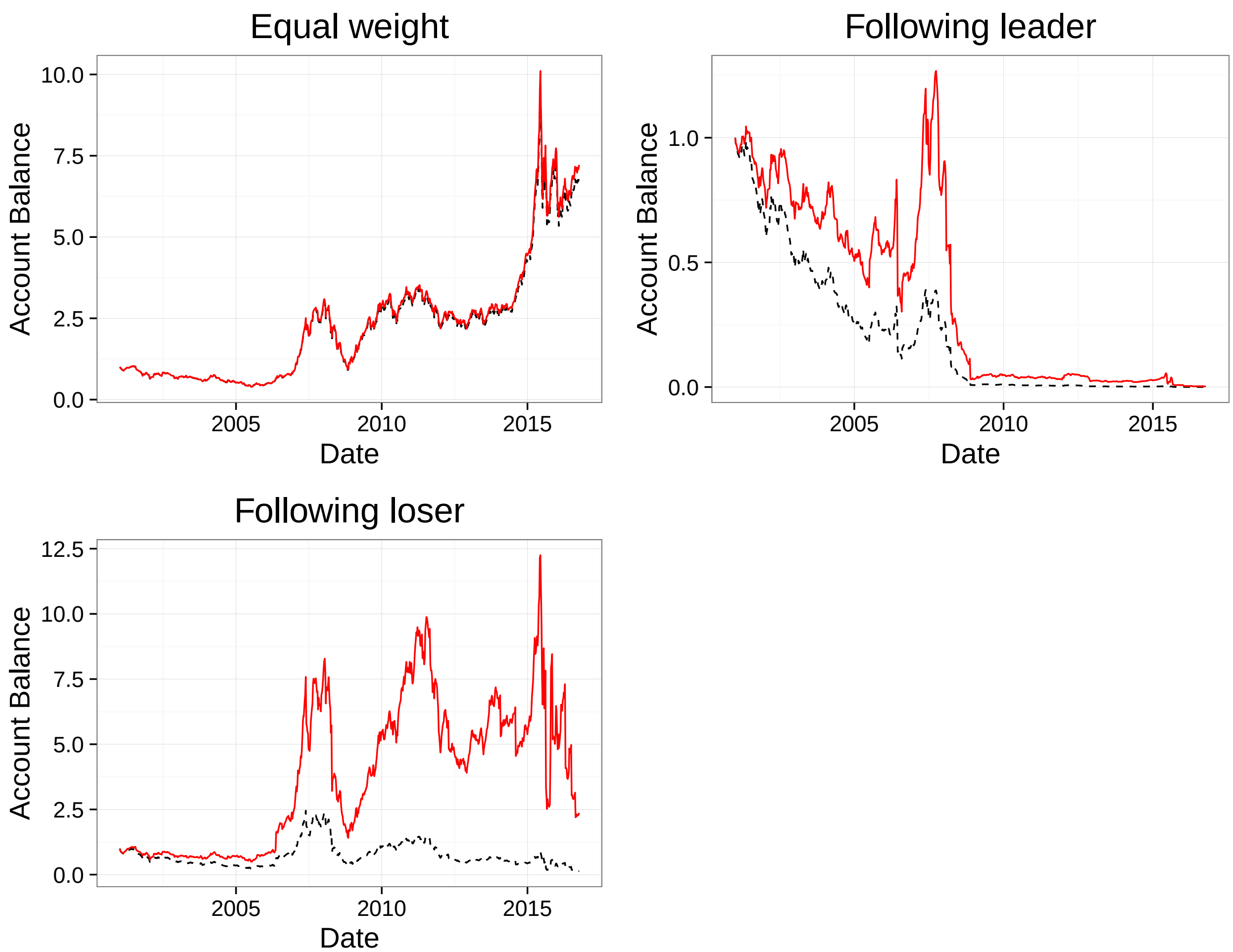

Figure 2: Naive rule application: weekly frequency case

where w is the weight vector for 100 stocks which we want to solve, Σ and μ are the sample covariance matrix and mean estimated by the training data. The results for daily/weekly/monthly frequency cases are illustrated in Figure 4, 5 and 6 respectively.

And we have some comments for the above results:

- No short sell constraint does actually help to control risk. For example, method 1 (short selling allowed) in all three cases perform worst. The account balance shrinks to nearly 0 only after a few years. Instead, method 2 (short selling not allowed) peforms much better.
- In the monthly frequency adjustments, method 4 (short selling not allowed and max weight is constrained by 5%) performs best, which implies in the long-term trading, this approach can also play a role of risk control.
- Again, the effect of transaction cost is significant for the daily frequency case, but nearly negligible for the monthly frequency case.

Part 2 The analysis in Part 1 is modified with a shrinkage used in estimation. we use shrinkage factor of $\alpha = 1/2$ to reweigh the sample means, i.e. for each stock i, we will use the following $\hat{\mu_i}$ as the estimation of sample mean in the quadratic programming. Other settings stay the same.

$$\hat{\mu_i} = \frac{1}{2}\mu_i + \frac{1}{2}\frac{1}{100}\sum_{i=1}^{100}\mu_i$$

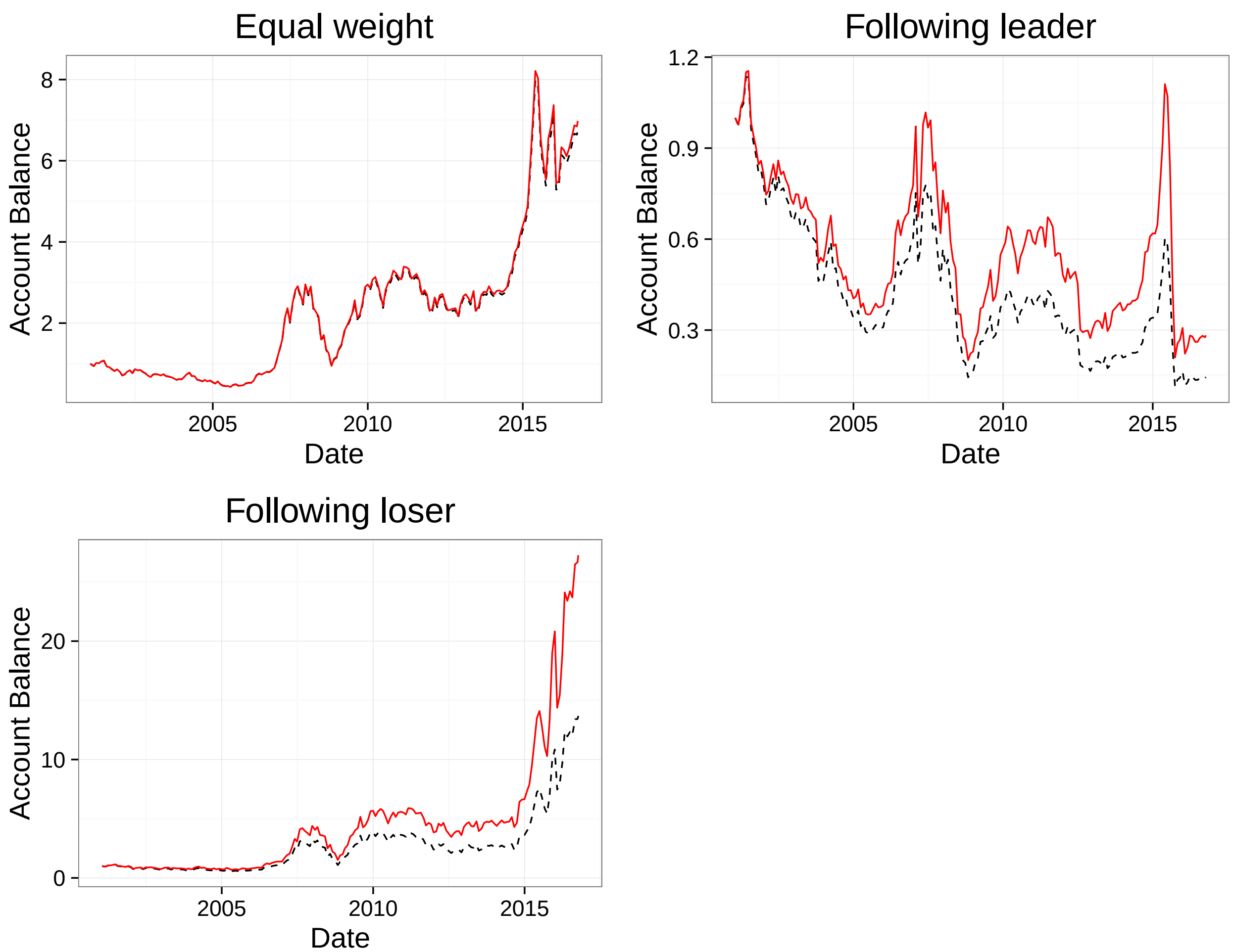

Figure 3: Naive rule application: monthly frequency case

The results for the four methods and three trading frequencies are shown in Figures 7, 8 and 9. By comparing them to Part 1, we found that the effect of shrinkage factor α is not quite significant.

Exercises.

4.1. Does $\mathbf{b}_{t+1}$ belong to the portfolio set B with no short sell.

4.2. Verify that (4.4) is the solution of (4.3).

4.3. Verify (4.5).

4.4. Write the codes of page 297 of RU into your R software and conduct a constrained quadratic optimization by designing a μ and Σ.

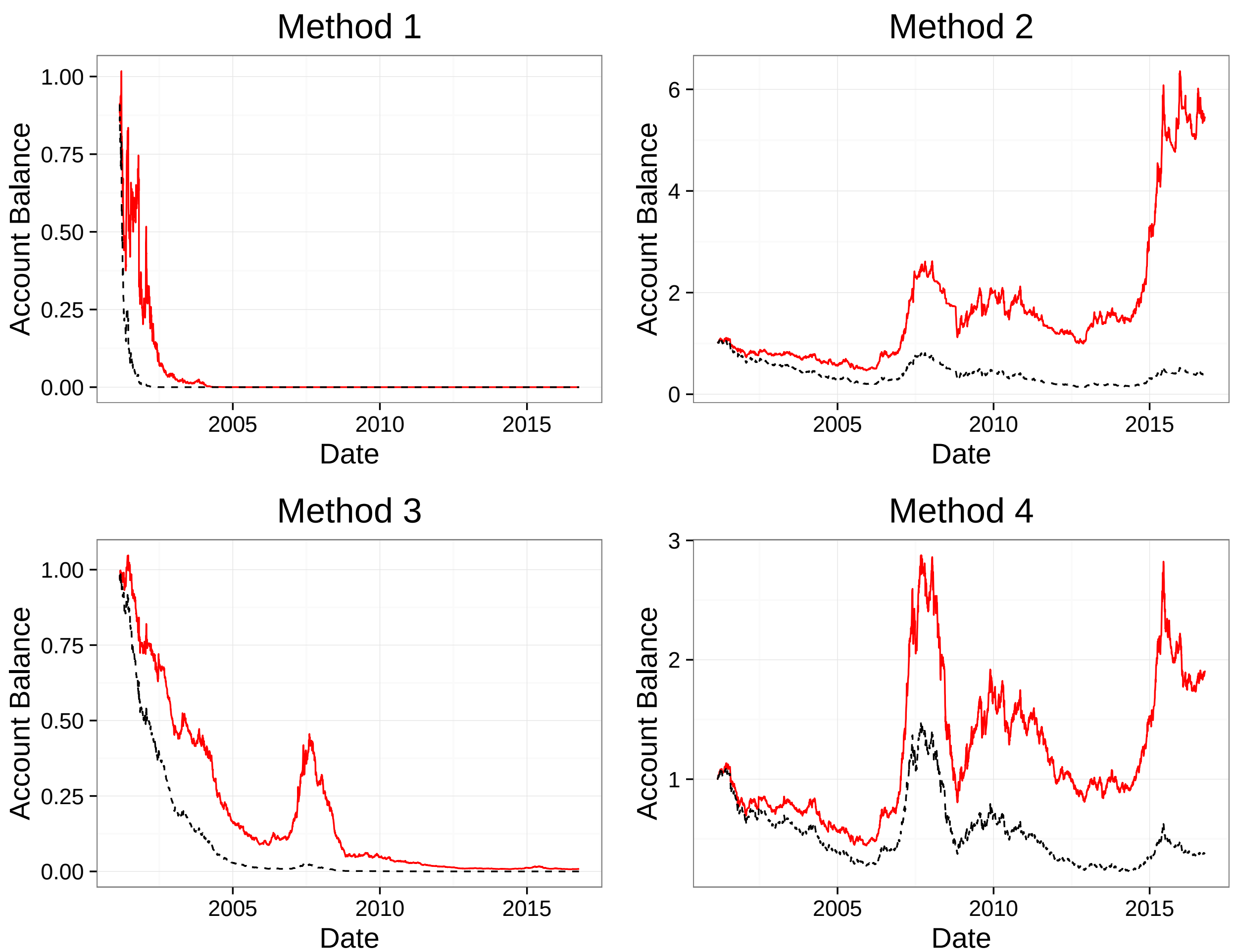

Figure 4: Markowitz's mean-variance principle: daily frequency case

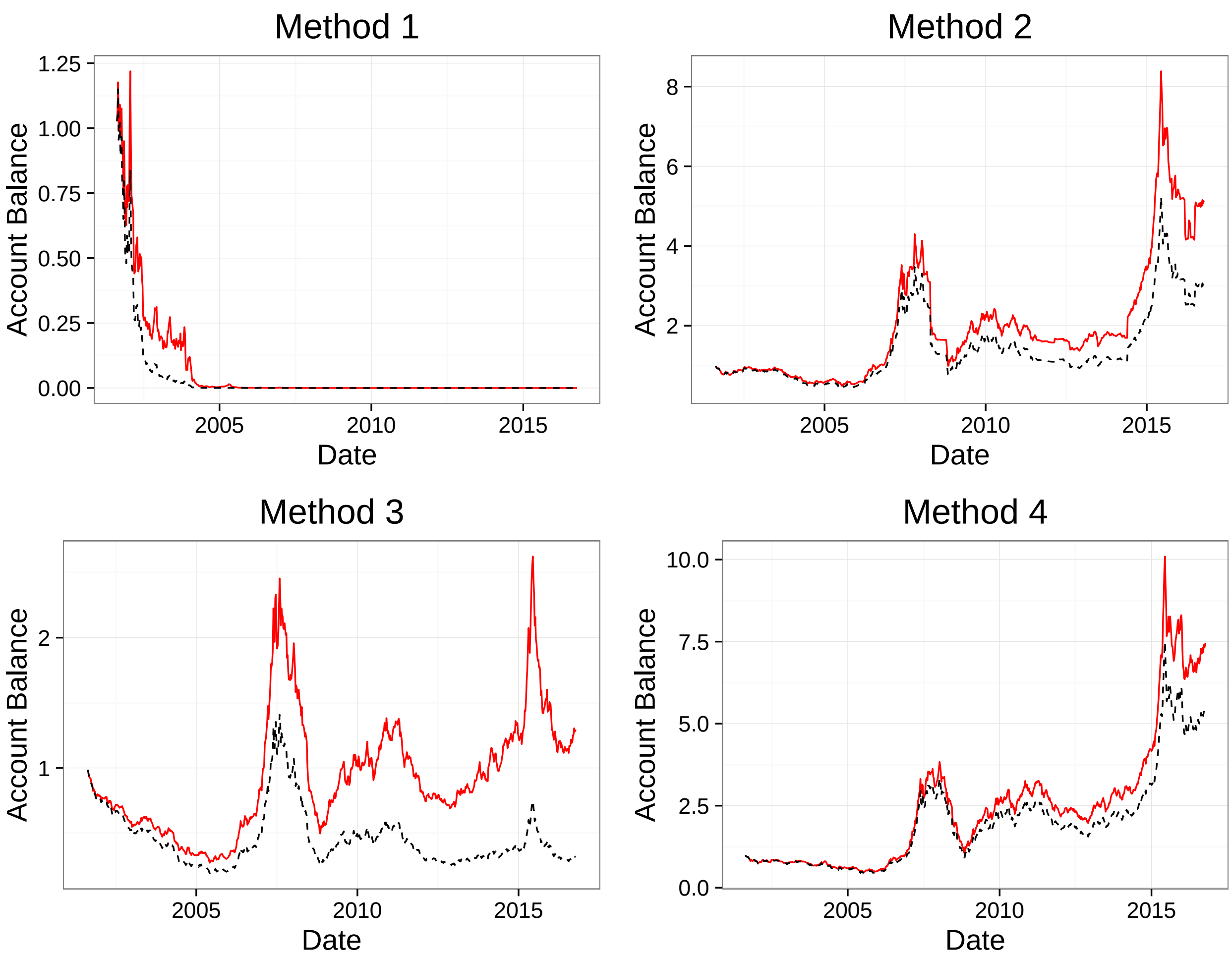

Figure 5: Markowitz's mean-variance principle: weekly frequency case

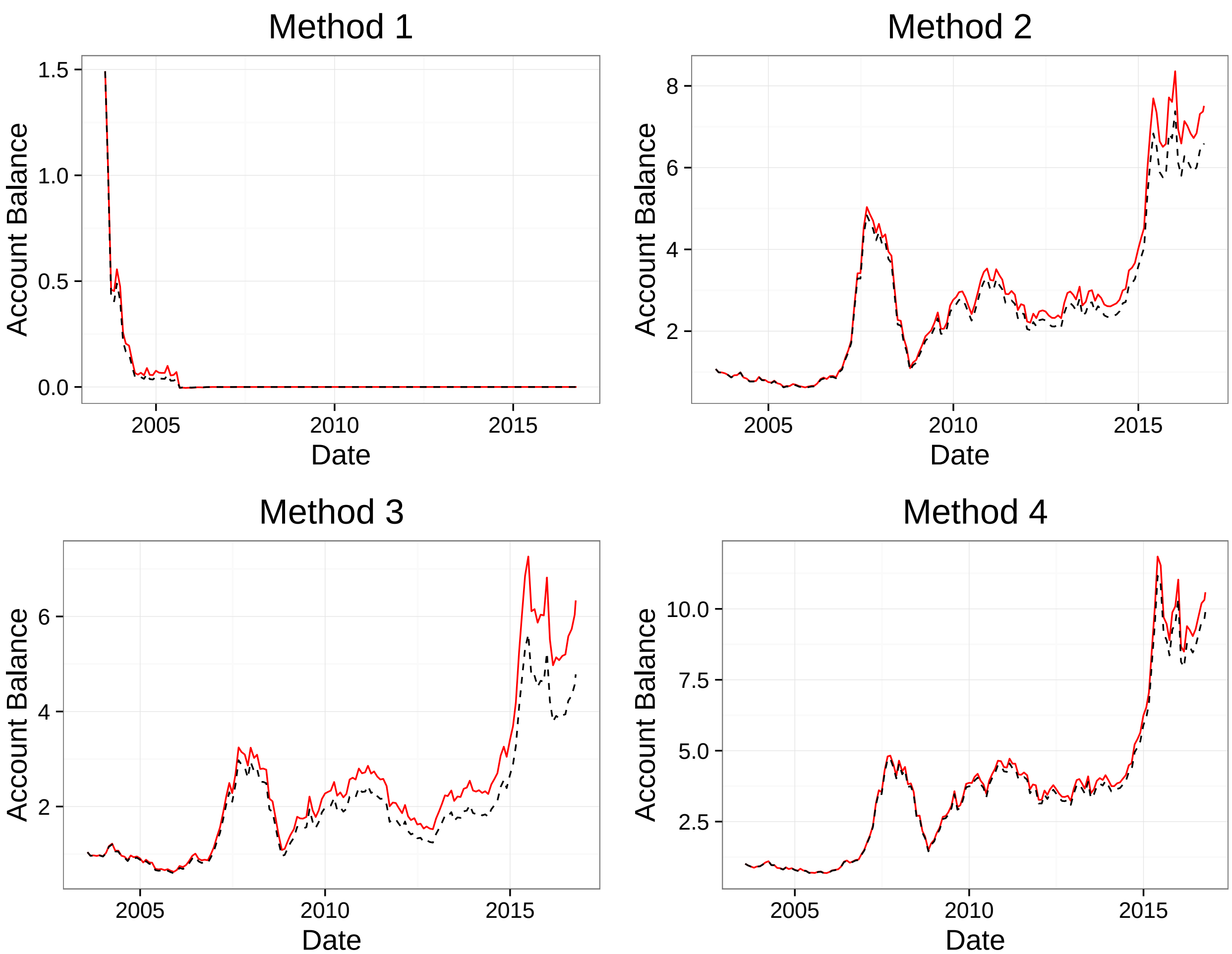

Figure 6: Markowitz's mean-variance principle: monthly frequency case

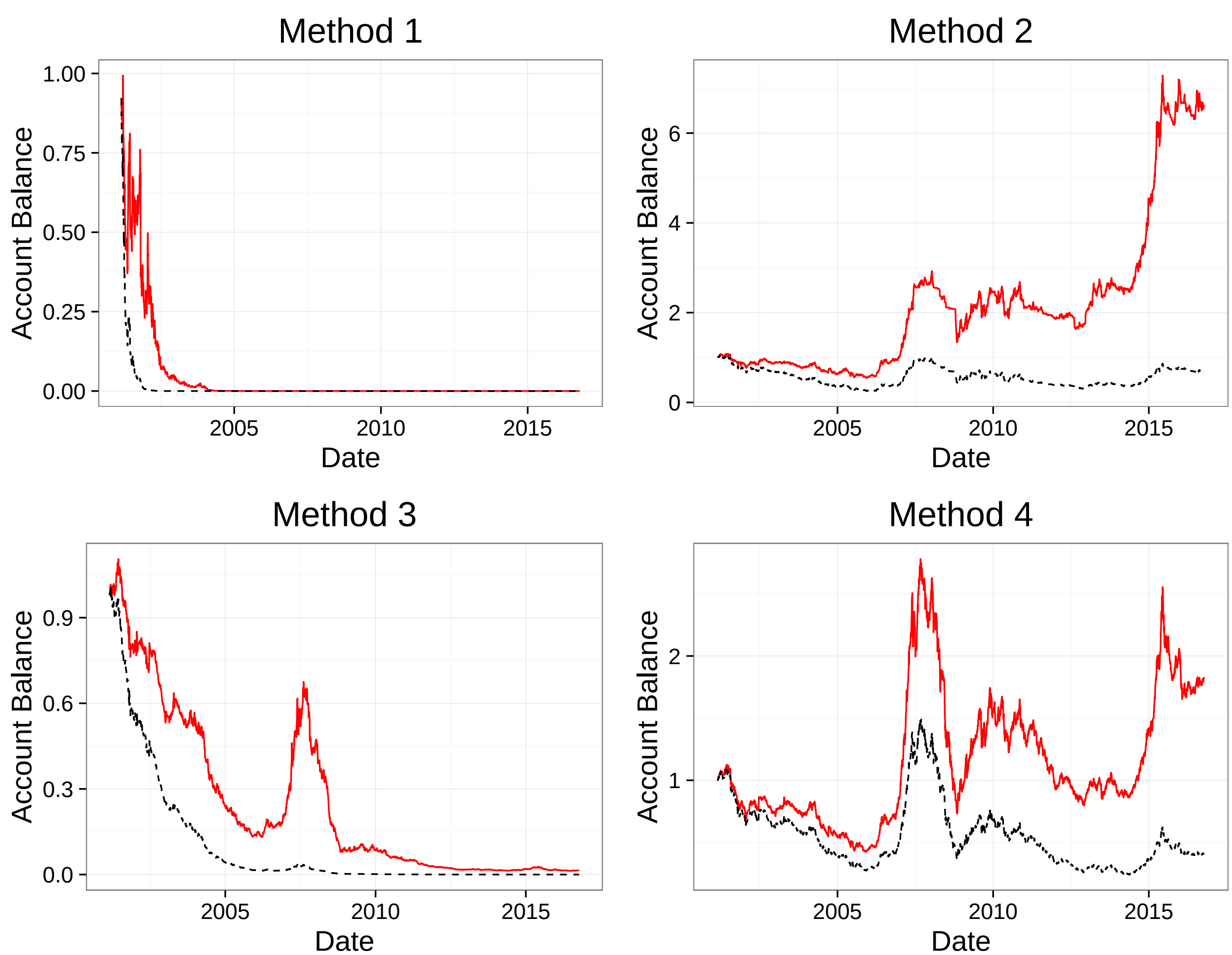

Figure 7: Markowitz's mean-variance principle: daily frequency case

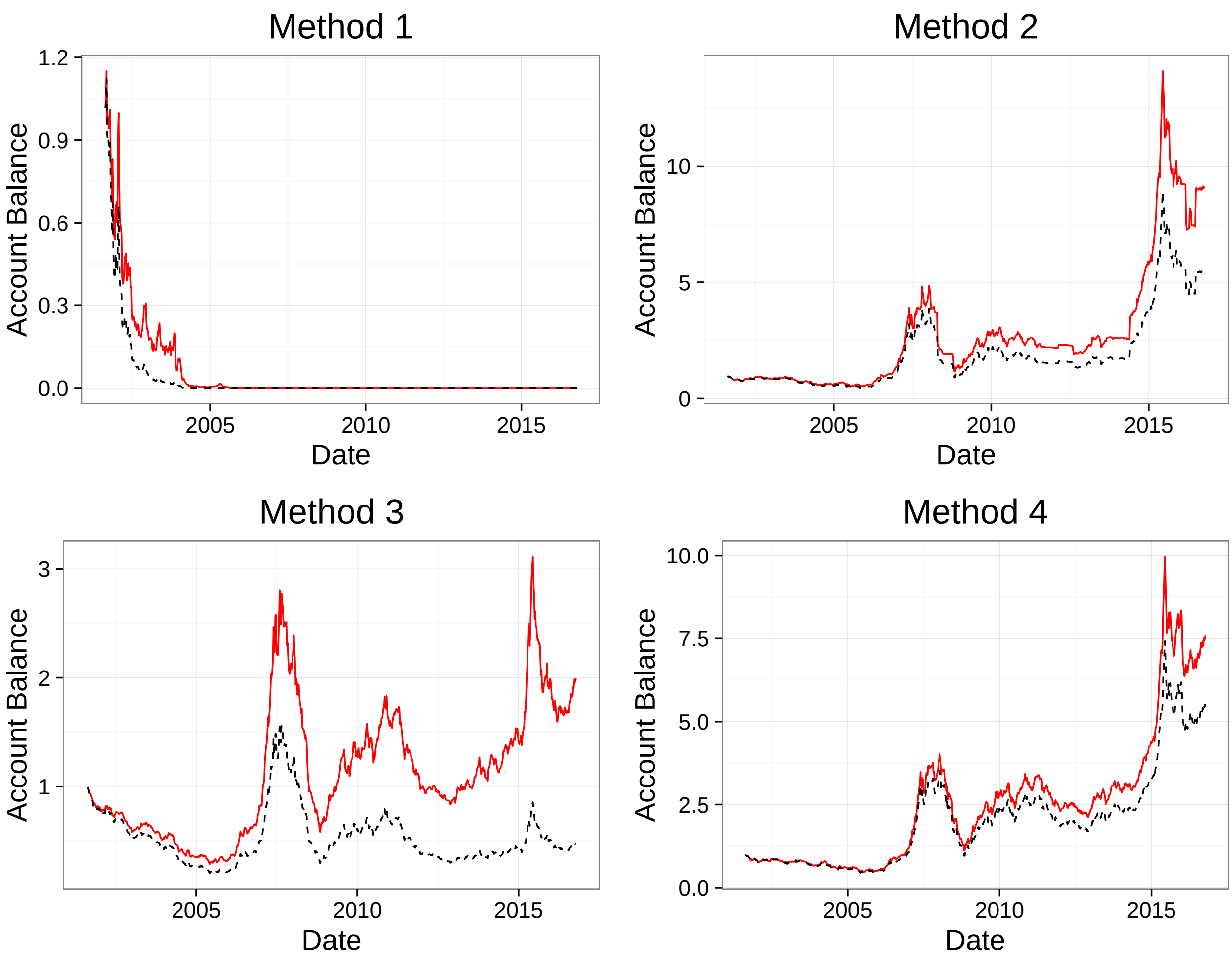

Figure 8: Markowitz's mean-variance principle: weekly frequency case

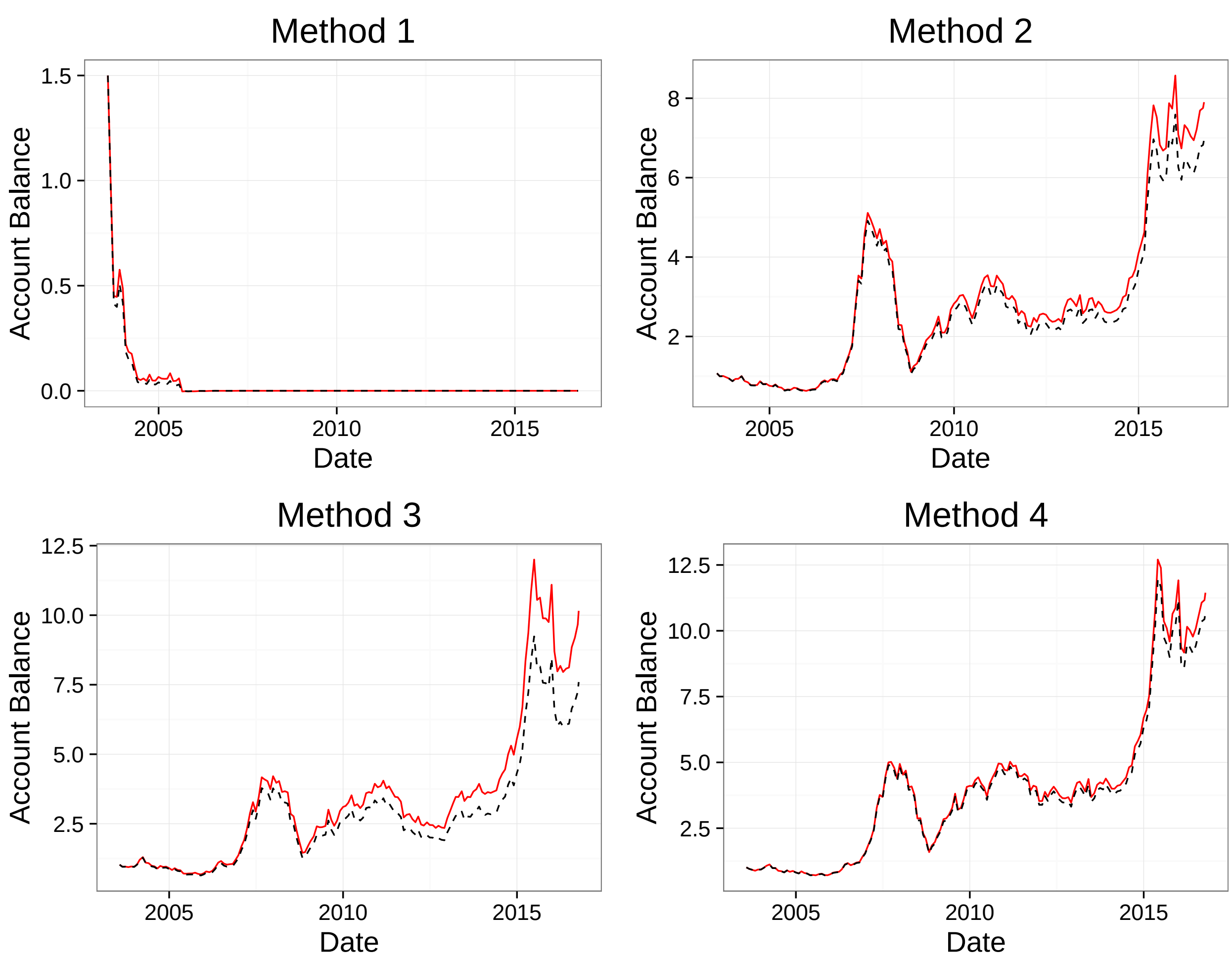

Figure 9: Markowitz's mean-variance principle: monthly frequency case