MA3 - Advanced Empirical Finance

7 June 2021

Contents

Introduction

Large-scale portfolio optimization has long been a field of interest. A large part of the literature focus on different ways of coping with the challenges due to noise in parameter estimates, model uncertainty and time variations. Further the presence of market imperfections and the illiquidity of certain assets in financial markets means that assets do not trade at their fundamental value but are subject to spreads and other transaction costs. These transaction costs mean that rebalancing portfolios is costly and therefore theoretical strategies, which only account for these ex post, may be sub-optimal when accounting for transaction costs.

Based on the findings in Hautsch et. al. (2019) we show how incorporating transaction cost changes the optimal portfolio by performing back-testing strategies on a sample of 40 stocks. We specifically focus on illustrating and testing the finding of Hautsch et al (2019) that

"...quadratic transaction costs can be interpreted as shrinkage of the variance-covariance matrix towards a diagonal matrix and a shift of the mean that is proportional to transaction costs and current holdings. Transaction costs proportional to the amount of rebalancing imply a regularization of the covariance matrix, acting similarly as the least absolute shrinkage and selection operator (Lasso) by Tibshirani (1996) in a regression problem and imply to put more weight on a buy-and-hold strategy."

We start by deriving the closed form solution to the portfolio choice certainty equivalent maximization problem adjusted for asset specific quadratic transaction costs. Using the derived solution we then illustrate the convergence towards the mean-variance efficient portfolio and how the convergence is affected by illiquidity. When measuring the illiquidity of the assets we use the Amihud (2002) measure, which uses the effect of trading volume on the magnitude of the returns to calculate the illiquidity of a stock. The amihud measure may not be the most accurate however it has the advantage that it is only calculated based on data of daily return and volume as opposed to other measures based on intraday data.

Next we investigate the effect of different transaction costs parameters on the out-of-sample performance

Finally we perform a backtesting of three different portfolio optimization strategies assuming proportional L_1 transaction costs. We compare the out-sample-performance of a naive portfolio, a theoretical portfolio that is optimally adjusted for the transaction costs ex-ante and a minimum variance portfolio with short sell constraint as suggested by Jagannathan and Ma (2003).

Exercise 1

We consider the portfolio maximization problem given by

$$\omega_{t+1}^* = \arg \max_{\omega_{t+1} \in \mathbb{R}^N, \iota'\omega_{t+1}=1} \omega_{t+1}'\mu - \nu_t\left(\omega_{t+1}, \beta\right) - \frac{\gamma}{2}\omega_{t+1}'\Sigma\omega_{t+1}$$

where μ is the estimator of returns, Σ is the estimator of the variance-covariance matrix and γ is the risk aversion parameter. The transaction costs are assumed to be quadratic in rebalancing and proportional to the stock illiquidity.

$$\nu_t\left(\omega_{t+1}, \beta\right) = \frac{\beta}{2}\left(\omega_{t+1} - \omega_{t+}\right)' B \left(\omega_{t+1} - \omega_{t+}\right)$$

The quadratic transaction costs function includes the amihud measure as a measure of illiquidity of the assets expressed as the vector B. As Hautsch et. al. (2019), shows the inclusion of transaction costs regularizes the covariance matrix in a similar way to Lasso and Ridge penalization

of turnover. This improves portfolio allocation as the regularization effects improves stability of the covariance estimates and reduce the frequency of rebalancing implying a reduction in turnover costs. In addition the transaction cost are proportional to the illiquidity measure, and hence we would expect transaction costs to be larger, the less liquid the given asset is.

We plug it into the maximization problem and rearrange terms:

$$\omega_{t+1}^* = \arg \max_{\omega_{t+1} \in \mathbb{R}^N, \iota' \omega_{t+1}=1} \omega_{t+1}' \mu - \frac{\beta}{2} (\omega_{t+1} - \omega_{t+})' B (\omega_{t+1} - \omega_{t+}) - \frac{\gamma}{2} \omega_{t+1}' \Sigma \omega_{t+1}$$

$$= \arg \max_{\omega_{t+1} \in \mathbb{R}^N, \iota' \omega_{t+1}=1} \omega_{t+1}' \mu - \frac{\beta}{2} \omega_{t+1}' B \omega_{t+1} + \beta \omega_{t+1} B \omega_{t+} - \frac{\beta}{2} \omega_{t+} B \omega_{t+} - \frac{\gamma}{2} \omega_{t+1}' \Sigma \omega_{t+1}$$

$$= \arg \max_{\omega_{t+1} \in \mathbb{R}^N, \iota' \omega_{t+1}=1} \omega_{t+1}' (\mu - \beta B \omega_{t+}) - \frac{\gamma}{2} \omega_{t+1}' \left(\Sigma + \frac{\beta}{\gamma} B \right) \omega_{t+1}$$

$$= \arg \max_{\omega_{t+1} \in \mathbb{R}^N, \iota' \omega_{t+1}=1} \omega_{t+1}' \mu_B^* - \frac{\gamma}{2} \omega_{t+1}' \Sigma_B^* \omega_{t+1}$$

where $\mu_B^* = \mu - \beta B \omega_{t+}$ and $\Sigma_B^* = \Sigma + \frac{\beta}{\gamma} B$. From the rearrangement of terms it follows that the maximization problem with incoorporated ex-ante transaction costs boils down to the Markowitz mean-variance problem without transaction costs, if we adjust the expressions for μ and Σ. We find the optimal portfolio weights be expressing and solving the Lagrangian for the maximization problem:

$$\mathcal{L}(\omega_{t+1}) = \omega_{t+1}' \mu_B^* - \frac{\gamma}{2} \omega_{t+1}' \Sigma_B^* \omega_{t+1} - \lambda(\iota' \omega_{t+1} - 1)$$

$$\frac{\partial \mathcal{L}(\omega_{t+1})}{\partial \omega_{t+1}} = \mu_B^* - \gamma \Sigma_B^* \omega_{t+1} - \lambda \iota = 0 \quad \Leftrightarrow \omega_{t+1} = \frac{1}{\gamma} \Sigma_B^{*-1} (\mu_B^* - \lambda \iota)$$

$$\frac{\partial \mathcal{L}(\omega_{t+1})}{\partial \lambda} = \iota' \omega_{t+1} - 1 = 0 \quad \Leftrightarrow \iota' \omega_{t+1} = 1$$

Combining the FOC's yields the following expression for λ which can be reinserted into the first FOC and rearranged in order to find the optimal portfolio weights:

$$1 = \frac{1}{\gamma} (\iota' \Sigma_B^{*-1} \mu_B^* - \lambda \iota' \Sigma_B^{*-1} \iota)$$

$$\lambda = \frac{1}{\iota' \Sigma_B^{*-1} \iota} (\iota' \Sigma_B^{*-1} \mu_B^* - \gamma)$$

$$\omega_{t+1} = \frac{1}{\gamma} \Sigma_B^{*-1} \left(\mu_B^* - \left(\frac{1}{\iota' \Sigma_B^{*-1} \iota} (\iota' \Sigma_B^{*-1} \mu_B^* - \gamma) \iota \right) \right)$$

$$\omega_{t+1} = \frac{\Sigma_B^{*-1} \iota}{\iota' \Sigma_B^{*-1} \iota} + \frac{1}{\gamma} \Sigma_B^{*-1} \iota \left(\mu_B^* - \frac{1}{\iota' \Sigma_B^{*-1} \iota} \iota' \Sigma_B^{*-1} \mu_B^* \right)$$

$$\omega_{t+1}^* = \frac{1}{\gamma} \left(\Sigma_B^{*-1} - \frac{1}{\iota' \Sigma_B^{*-1} \iota} \Sigma_B^{*-1} \iota \iota' \Sigma_B^{*-1} \right) \mu_B^* + \frac{1}{\iota' \Sigma_B^{*-1} \iota} \Sigma_B^{*-1} \iota$$

Illiquidity of the assets is in general interpreted as the extra cost or risk of capital tied up in less liqiud assets as these in general are harder to liquidate in case of a price drop on the asset. Therefore the measure of illiquidity can be interpreted as a risk premium and the more illiquid an asset is the higher the transaction costs will be. If comparing weights to a myopic investor disregarding transaction cost the weight on an asset with a higher illiquidity measure (low liquidity of the given asset) will be higher for a investor who takes transaction costs into account if we consider a price drop of the asset. This is due to the fact that is becomes more expensive to rebalance portfolio weights the less liquid and in turn higher transaction cost are for the given asset.

Exercise 2

As in Hautsch et al (2019), if we denote the initial wealth allocation as ω_0, we are able to study the long run effect, by examining the sequential rebalancing:

$$\omega_T = \sum_{i=0}^{T-1} \left(\frac{\beta}{\gamma} A(\Sigma_B^*) \right)^i \omega(\mu, \Sigma_B^*) + \left(\frac{\beta}{\gamma} A(\Sigma_B^*) \right)^T \omega_0$$

where $A(\Sigma_B^*) = \left(\Sigma_B^{*-1} - \frac{1}{\iota' \Sigma_B^{*-1} \iota} \Sigma_B^{*-1} \iota' \iota' \Sigma_B^{*-1} \right)$. The equation shows that ω_T can be interpreted as weighted average of the mean variance efficient portfolio weights and the initial allocation. Proposition 2 of Hautsch et al (2019) states that for infinitely large transaction costs the long run weights will be equal to the iniital allocation. However proposition 3 states that there exist a range for β below a certain threshold β^* where the initial allocation can be ignored in the long run. Using $T \to \infty$ and $\beta < \beta^*$ the series of sequential rebalancing converges to a unique fix point given by the weights computed with no regard to the transaction costs as finally stated in Proposition 4 of hautsch et al (2019).

$$\omega_\infty = \left(I - \frac{\beta}{\gamma} \Sigma_B^* \right)^{-1} \omega(\mu, \Sigma_B^*) = \omega(\mu, \Sigma)$$

As stated in the article this shows that for large transaction costs above the threshold, the investor will not reallocate to the efficient portfolio. However if the transaction costs are moderate we can obtain convergence towards the efficient portfolio of a setup where transaction costs are ignored.

In this exercise we show the convergence of the portfolio towards the efficient frontier as derived in proposition 4 when $T \to \infty$. The weights are computed using the function defined in exercise 1, and are computed using a loop over 10^4 values. μ is computed as the mean of each stock return using the full sample while Σ is computed with the Ledoit-Wolf shrinkage covariance estimator. We use the Ledoit-Wolf covariance estimator to estimate the covariance among the assets as proposed by Ledoit and Wolf (2003). As stated by Ledoit and Wolf (2003) the sample covariance matrix is unbiased but unfortunately often associated with a lot of estimation error when the number of observation is comparable of smaller than the number of individual stocks. In our case we have 40 assets and 500 observations. However we would argue that estimation error may still be present and follow the advice of Ledoit and Wolf (2003) and avoid using the sample covariance matrix.
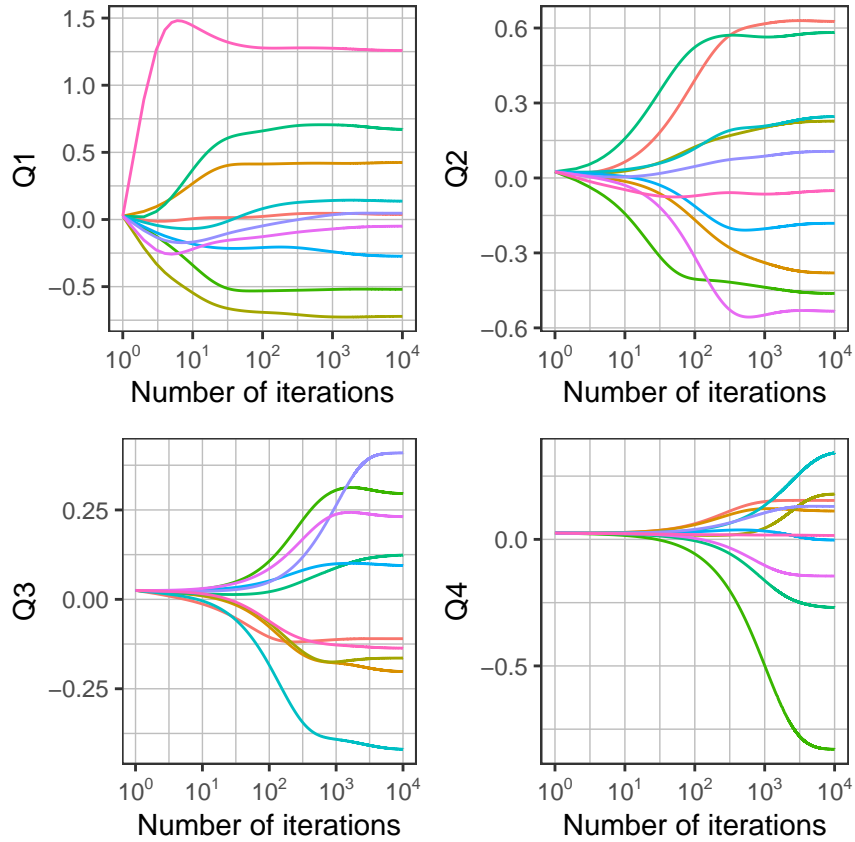
Below the convergence towards the efficient frontier is shown. The figure shows the convergence starting from the naive portfolio towards the theoretically efficient portfolio. The reallocation is notably larger in initial iterations, and the adjusting of portfolio weights decrease in speed as we loop through the iterations.

Figure 1: Convergence towards efficient portfolio
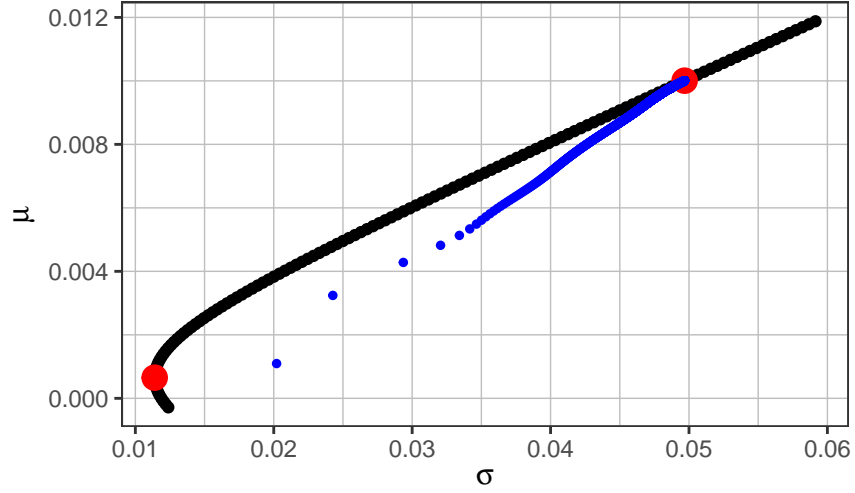
In order to illustrate the weight dynamics towards the efficient portfolio we divide the stocks into quartiles based on the measure of illiquidity. This enables os to show the difference in convergence speed towards the efficient portfolio weights for the different groups of assets. The first quartile ($Q1$) are the most liquid assets, while the last quartile ($Q4$) are the assets with the highest illiquidity measure, hence the least liquid assets. The x-axis follows logarithmic scale for a better overview of the difference in convergence speeds.

Figure 2: Effect of illiquidity measure on weight dynamics

The plot shows are clear trend of decreasing speed in convergence towards the efficient weights as the illiquidity measure rises. The most liquid assets of the first quartile $Q1$ seem to reach

the efficient weights almost immediately, while the less liquid the assets are the slower the convergence is as seen from the latter quartiles.

Exercise 3

In this exercise we illustrate the effect of different values of the transaction cost parameter β on the annualized out-of-sample Sharpe-ratio. We consider values of β between 0 and 100. However, when we evaluate the performance of the portfolio, this is done using a transaction cost parameter equal to the actual transaction cost. The true transaction cost parameter is assumed to be $\beta_{true} = 1$.

We consider two ways of estimating the variance-covariance matrix, $\hat{\Sigma}$, which is the sample covariance estimator and the shrinkage estimator proposed by Ledoit and Wolf (2003). However, we do not estimate the mean vector and simply set it to zero.

More specifically, for each considered value of β we perform a backtest using a rolling window with a length of 250 days. We use the naive portfolio as the initial allocation and then each day re-estimate Σ and implement the optimal portfolio taking transaction cost into account. Since our data set consist of 500 observations for each asset and we set the rolling window length to 250, it follows that we for each beta estimate a time series of optimal portfolios ω_t^β of length $500 - 250 = 250$. For each of the periods we calculate the realized gross portfolio return as:

$$r_t^\beta = r_t'\omega_t^\beta$$

and the portfolio return net of transaction costs as, where we use the "actual" transactions cost parameter:

$$xr_t^{\beta,nTC} = r_t^\beta - \beta_{true}\|\omega_{t+1}^\beta - \omega_{t+}^\beta\|_2$$

where the L_2-norm is calculated as $\|\omega_{t+1}^\beta - \omega_{t+}^\beta\|_2 = (\omega_{t+1} - \omega_{t+})'B(\omega_{t+1} - \omega_{t+})$ such that it takes into account the asset-specific transaction costs.

For each β we can then calculate the average net return and standard deviation, $\hat{\mu}^{\beta,pf}$ and $\hat{\sigma}^{\beta,pf}$ based on the time series of realized net returns. The Sharpe-ratio is then calculated as the ratio between these.

Results for both estimators of the the variance-covaraince matrix is reported below:

Figure 3: Annualized Sharpe ratios for different transaction costs parameter values
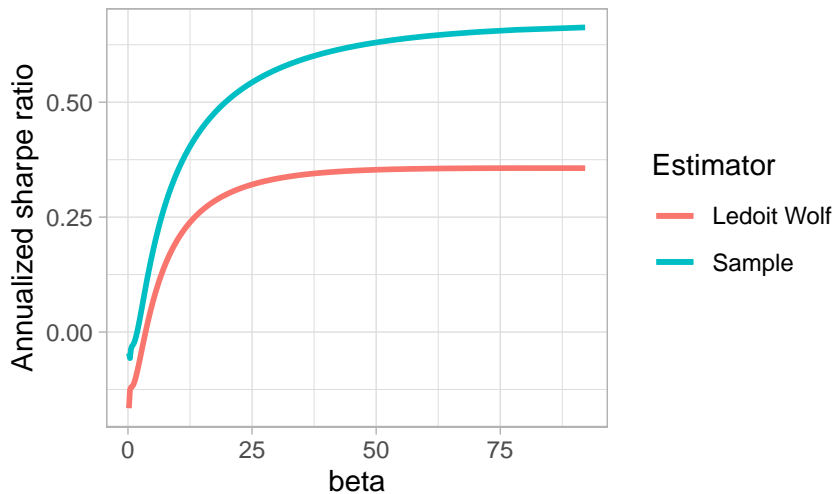
From the figure above we note, that the sample estimator of the variance-covaraince matrix clearly outperforms the Ledoit and Wolf shrinkage estimator, since it delivers higher out-of-sample Sharpe ratios for all β-values between 0 and 100.

On the one hand the transaction cost parameter improve the conditioning of the covariance matrix, but it also reduce the mean portfolio return. A change in the transaction cost parameter, β, therefore has an ambiguous effect of the Sharpe ratio. As evident from the figure, the annualized Sharpe ratio is increasing in the transaction cost parameter, β, for both estimators over the interval from 0 to 100. This means that positive effect from conditioning the covariance matrix dominates the negative effect from decreasing the mean. However, this overall positive effect on the Sharpe ratio seem to be largest for increases in small values of β, so the effect of further penalizing turnover vanishes as β goes to 100.

We assumed that the β-value reflecting the actual transaction costs is $\beta_{true} = 1$. When we use this value in the optimization problem it does not deliver the highest possible out-of-sample annualized Sharpe ratio. This means that it can be optimal to choose theoretically sub-optimal portfolios based on transaction cost parameter values that do not reflect the actual transaction costs. If we instead define our objective as maximizing the annualized out-of-sample Sharpe-ratio it is therefore tempting to view β as a parameter to be optimized, since our results indicate that it increases the Sharpe ratio if we choose a β-value which is higher than the true one.

Exercise 4

In the last exercise we perform a full-fledged backtesting strategy with proportional L_1 transactions costs. We compare the out-of-sample performance of a naive portfolio where the 40 assets are equally weighted; a portfolio which computes the theoretical optimal portfolio weights with optimal ex-ante adjustment for the L_1 transaction costs in the spirit of Hautsch et al. (2019) and a minimum variance portfolio with short sell constraint as suggested by Jagannathan and Ma (2003).

$$\nu_t\left(\omega_{t+1}, \omega_{t+}, \beta\right) = \beta\left\|\omega_{t+1} - \omega_{t+}\right\|_1 = \beta\sum_{i=1}^{N}\left|\omega_{i,t+1} - \omega_{i,t+}\right|$$

With L_1 transactions costs the cost is proportional to the sum of absolute re-balancing and thereby represent a more realistic proxy than the quadratic transaction costs we use in the exercises above. As is common in the literature we use a penalization term (β) of 50 bp.

To perform the portfolio backtesting strategy we start by defining functions that compute the optimal weights for each of the three portfolios.

The naive portfolio places equal weights on each of the assets.

$$\omega = \frac{1}{N}\iota$$

In our cases with 40 assets each assets has a weight of 2,5%.

As suggested by Hautsch et al. (2019) we compute the theoretical optimal portfolio weights with optimal ex-ante adjustment for the transaction costs.

$$\omega_{t+1}^* = \max_{\omega\in\mathbb{R}^N, \iota\omega=1} \omega'\mu - \beta\left\|\omega_{t+1} - \omega_{t+}\right\|_1 - \frac{\gamma}{2}\omega'\Sigma\omega \text{ s.t. } \iota'\omega = 1$$

It should be noted that with L_1 transaction cost there does not exist a closed form solution, therefore solve the optimization problem using a non-linear constraint optimizer.

Jagannathan and Ma (2003) show that constructing a minimum variance portfolio with short-selling constraints is equivalent to shrinking the larger elements of the covariance matrix. When computing the the portfolio weights of the short-sell constrained minimum variance portfolio we face the following optimization problem:

$$\omega_{t+1}^{\text{mvp no s.}} = \min_{\omega \in \mathbb{R}^N} \omega' \Sigma \omega \text{ s.t. } \iota' \omega = 1 \text{ and } \omega_i \geq 0 \forall i = 1, \dots N$$

As explained by Jagannathan and Ma (2003) introducing short-selling constraint may as other shrinking methods lead to a more precise estimate of true population covariance given that a large covariance is due to sampling error. However if the true population covariance is indeed large, short sell constraints and shrinkage will result in specification error.

We then use these functions to perform a out-of-sample backtest and evaluate the performance of the three strategies. In our backtest we use a window-length of 250. The mean is computed as the sample mean of the returns, while the covariance is estimated using a Ledoit-Wolf shrinkage estimation. To compare the performance of the three strategies we compute the annualized out-of-sample means, SD, Sharpe ratios and turnover. The returns are the raw returns net of trading costs.

Table 1: Performance of portfolio strategies

strategy	Mean	SD	Sharpe	Turnover
Naive	0.3217	0.4236	0.7594	2.0390
MV (TC)	0.2695	0.4021	0.6702	0.0000
MV short-sell constrained	0.0768	0.2944	0.2607	4.3333

As shown in the table above the theoretical optimal portfolio that incorporate transaction costs ex-ante has the lowest turnover and second lowest SD. The low turnover compared to the other portfolios is expected since this problem incorporates the transactions costs ex-ante. When the turnover is zero it might be an indication of the transaction cost parameter β being so large that reallocation is not profitable and the weights will be equal to the initial allocation as stated in proposition 2 by Hautsch et al. (2019). On the other hand the the minimum variance portfolio with short-sell constraint had the lowest SD but also the highest turnover and and lowest annual net return with a annual return of just 7.7 pct. The lower return is partly due to the high turnover and the costs associated with re-balancing the portfolio, but can also be contributed to it being a minimum variance portfolio. Opposed to the optimal portfolio with transaction costs it disregards the return and only aims to minimize the volatility.

We further note that the naive portfolio has a higher mean and the highest Sharpe-ratio. If we evaluate the strategies performance based on the Sharpe-Ratio's then an investor would be better of choosing the naive portfolio based on our analysis. This might be an indication that even though incorporating transaction costs ex ante or using short sell constraint as can regularize the underlying covariance matrix it might not always be sufficient to outperform a simple naive portfolio.