Index-tracking portfolio optimization model

Guillermo Navas-Palencia*

October 18, 2016

Abstract

In the present tutorial report we examine the theory and computational aspects behind the index-tracking portfolio optimization model. This model is compared with Markowitz mean-variance model. This report is distributed with an example in C using the NAG C Library.

1 Introduction

The index-tracking is a form of passive fund management. The index-tracking problem is the problem of reproducing the performance of a stock market index by considering a portfolio of assets comprised on the index. This approach differs from the full replication strategy, where a fund purchases all the stocks that make up a particular market index. A passively managed fund whose objective is to reproduce the return on an index is known as an index fund, or a tracker fund. The classical index-tracking approach represents the problem in a least-square framework with errors computed using a sample of historical data. A new approach described in [2], which this report is based on, replaces the classical risk measure of portfolio variance by the variance of tracking error between stochastic index return and the return on the portfolio selection. In this report we formulate the index-tracking portfolio optimization model and present an illustrative example where we compare the presented model with the classical Markowitz mean-variance portfolio optimization model.

2 Index-tracking optimization model

Consider a market index $M$ and select a portfolio $P$ with $n$ assets. The portfolio $P$ reproduces the performance of the market index if the return on $P$, denoted by $r_P$, follows very closely the return of the index, denoted by $r_M$, at every unit time period. Thus, the goal is to obtain the portfolio $P$ that minimizes the variance of the difference between both returns, $r_P$ and $r_M$. Denote the return on asset $i$ by $r_i$ and its mean by $\mu_i$, for $i = 1, \ldots, n$. The variance-covariance matrix of assets returns is denoted by $V$, where $\sigma_{ij}$ denotes the covariance between the returns of two assets, and $\sigma_{ii}$ the variance of $r_i$. The covariance between asset returns and market index returns is given by $\sigma_{iM} = Cov(r_i, r_M)$. The beta of asset $i$ relative to the market index $M$ is given by

$$\beta_i = \frac{\sigma_{iM}}{\sigma^2_M}$$

(1)

where $\sigma^2_M$ denotes the variance of the market index returns. As previously stated, the index-tracking model aims to minimize $\text{Var}(r_P - r_M) = \sigma_P^2 + \sigma^2_M - 2\sigma^2_M \beta_P$, the so-called measure of goodness of the index-tracking portfolio. Thus, the optimization problems is formulated as

---

* Numerical Algorithms Group (NAG) Ltd., Oxford, UK. Email: guillermo.navas@nag.co.uk
follows

\[
\begin{align*}
\text{minimize} \quad & \frac{1}{2} x^T V x - \sigma_M^2 \beta^T x \\
\text{subject to} \quad & \sum_{i=1}^n \mu_i x_i = m \\
& \sum_{i=1}^n x_i = 1
\end{align*}
\]  

(2)

and \(-1 \leq x_i \leq 1, \ i \in \{1, \ldots, n\}\). Therefore, we need to solve a constrained convex quadratic programming problem (QP). Given that the matrix \(V\) is symmetric positive semi-definitive and dense, we select the solver \texttt{e04ncc}. The following measures are used to determine the characteristics of the optimal portfolio

- Portfolio variance: \(\sigma_P^2 = x^T V x\).
- Portfolio beta: \(\beta_P = \sum_{i=1}^n \beta_i x_i\).
- Portfolio \(P\) (measure of goodness): \(\text{Var}(r_p - r_M) = \sigma_P^2 + \sigma_M^2 - 2 \sigma_M^2 \beta_P\).

Note that the tracking-efficient (TE) portfolio is equivalent to Markowitz mean-variance (MV) portfolio when the linear term \(-\sigma_M^2 \beta^T x\) is removed from the objective function. A review of the basic Markowitz portfolio optimization theory and solution using the NAG Library can be found in [3].

3 Example

We present a simple example to illustrate the described methodology. We select 7 technology stocks in order to track the S&P 500 index. The data for this example was downloaded from Yahoo Finance [4] using a \texttt{Python} script with the module \texttt{urllib2}. For this example we downloaded monthly prices from 01/2009 to 01/2016, see Figure 1. From a sample of historical data we can compute the one-period returns for each stock by using the well-known formula

\[
\begin{align*}
\text{rate}_t = \frac{\text{price}_t - \text{price}_{t-1}}{\text{price}_{t-1}}
\end{align*}
\]  

(3)

Having the returns we can easily determine the mean vector, standard deviation vector, covariance matrix and beta vector, see Table 1 and 2.

![Figure 1: Monthly prices for seven technology stocks from January 2009 to January 2016.](image)
<table>
<thead>
<tr>
<th>Asset</th>
<th>Mean ($\mu$)(%)</th>
<th>Std. dev ($\sigma$)(%)</th>
<th>Beta ($\beta$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>GSPC</td>
<td>1.11</td>
<td>4.15</td>
<td>-</td>
</tr>
<tr>
<td>AAPL</td>
<td>2.82</td>
<td>7.44</td>
<td>1.026</td>
</tr>
<tr>
<td>CSCO</td>
<td>1.08</td>
<td>7.8</td>
<td>1.250</td>
</tr>
<tr>
<td>GOOG</td>
<td>2.00</td>
<td>7.3</td>
<td>0.975</td>
</tr>
<tr>
<td>IBM</td>
<td>0.72</td>
<td>4.57</td>
<td>0.595</td>
</tr>
<tr>
<td>MSFT</td>
<td>1.79</td>
<td>6.78</td>
<td>0.994</td>
</tr>
<tr>
<td>ORCL</td>
<td>1.21</td>
<td>6.96</td>
<td>1.199</td>
</tr>
<tr>
<td>YHOO</td>
<td>1.49</td>
<td>8.47</td>
<td>0.941</td>
</tr>
</tbody>
</table>

Table 1: Mean, standard deviation and beta.

<table>
<thead>
<tr>
<th>AAPL</th>
<th>CSCO</th>
<th>GOOG</th>
<th>IBM</th>
<th>MSFT</th>
<th>ORCL</th>
<th>YHOO</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.005528</td>
<td>0.002689</td>
<td>0.001983</td>
<td>0.001417</td>
<td>0.001996</td>
<td>0.002167</td>
<td>0.001418</td>
</tr>
<tr>
<td>0.002689</td>
<td>0.006082</td>
<td>0.002637</td>
<td>0.001873</td>
<td>0.002812</td>
<td>0.003305</td>
<td>0.002146</td>
</tr>
<tr>
<td>0.001983</td>
<td>0.002637</td>
<td>0.005324</td>
<td>0.001132</td>
<td>0.002193</td>
<td>0.001718</td>
<td>0.001765</td>
</tr>
<tr>
<td>0.001417</td>
<td>0.001873</td>
<td>0.001132</td>
<td>0.002084</td>
<td>0.001021</td>
<td>0.001571</td>
<td>0.000677</td>
</tr>
<tr>
<td>0.001996</td>
<td>0.002812</td>
<td>0.002193</td>
<td>0.001021</td>
<td>0.004599</td>
<td>0.002502</td>
<td>0.001350</td>
</tr>
<tr>
<td>0.002167</td>
<td>0.003305</td>
<td>0.001718</td>
<td>0.001571</td>
<td>0.002502</td>
<td>0.004843</td>
<td>0.002146</td>
</tr>
<tr>
<td>0.001418</td>
<td>0.002167</td>
<td>0.001765</td>
<td>0.000677</td>
<td>0.001350</td>
<td>0.002146</td>
<td>0.007173</td>
</tr>
</tbody>
</table>

Table 2: Variance-covariance matrix $V$.

Following the example in [2], we set the desired portfolio mean $m = \mu_M$, where $\mu_M$ represents S&P 500 index mean return. Table 3 shows that the index-tracking model, using the market index mean, reduces the number of short positions in the portfolio with respect to Markowitz portfolio. Markowitz portfolio has 2 short positions (CSCO and ORCL) whereas the index-tracking portfolio only one short position (AAPL). In addition, the index-tracking portfolio tends to reduce extreme position, for example IBM changes the long position from 72.1% to a more conservative 44.9%.

<table>
<thead>
<tr>
<th>Ticker</th>
<th>MV</th>
<th>TE</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAPL</td>
<td>0.019969</td>
<td>-0.023608</td>
</tr>
<tr>
<td>CSCO</td>
<td>-0.123901</td>
<td>0.072067</td>
</tr>
<tr>
<td>GOOG</td>
<td>0.076037</td>
<td>0.076785</td>
</tr>
<tr>
<td>IBM</td>
<td>0.721647</td>
<td>0.449256</td>
</tr>
<tr>
<td>MSFT</td>
<td>0.171989</td>
<td>0.115741</td>
</tr>
<tr>
<td>ORCL</td>
<td>-0.001755</td>
<td>0.193798</td>
</tr>
<tr>
<td>YHOO</td>
<td>0.136014</td>
<td>0.115961</td>
</tr>
</tbody>
</table>

Table 3: Mean-variance and index-tracking optimal portfolios.

Table 4 shows the main portfolio measures. The value of the measure of goodness indicates that the TE portfolio tracks the S&P 500 index more closely with $Var(r_P - r_M) = 0.000707$ compared with the MV portfolio, $Var(r_P - r_M) = 0.001049$.

<table>
<thead>
<tr>
<th>Measure</th>
<th>MV</th>
<th>TE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Portfolio variance</td>
<td>0.001620</td>
<td>0.001962</td>
</tr>
<tr>
<td>Portfolio beta</td>
<td>0.666135</td>
<td>0.864691</td>
</tr>
<tr>
<td>Measure of goodness</td>
<td>0.001049</td>
<td>0.000707</td>
</tr>
</tbody>
</table>

Table 4: Portfolio measures.
3.1 Compilation
The presented problem has been solved using the NAG C Library. For example to compile it with NAG C Library for Linux Mark 25, CLL6I25DCL, using the gcc compiler follow the instructions below. For other platforms please refer to the appropriate User's notes of your NAG C Library, see [https://www.nag.co.uk/doc/inun/cl25.html](https://www.nag.co.uk/doc/inun/cl25.html)

Listing 1: Linux instructions.
```
gcc -O3 index_tracking.c -I ../CL25/CLl6i25dcl/include/ ../CL25/CLl6i25dcl/lib/libnagc_nag.so -lpthread -lm -o index_tracking.exe
```

Listing 2: To run the example we pass the downloaded data as standard input.
```
./index_tracking.exe < index_tracking.txt
```

Listing 3: Example data.
```
7 85
GSPC AAPL CSCO GOOG IBM MSFT ORCL YHOO
1.9E+03 9.8E+01 2.5E+01 7.3E+02 1.3E+02 5.2E+01 3.5E+01 3.1E+01
```

Listing 4: Example output.
```
Optimal portfolio selection:
----------------------------
GSPC: market index
AAPL: -2.36%
CSCO: 7.21%
GOOG: 7.68%
IBM: 44.93%
MSFT: 11.57%
ORCL: 19.38%
YHOO: 11.60%

Portfolio measures:
----------------------------
Portfolio variance : 0.001962
Portfolio beta : 0.864691
Measure of goodness : 0.000707
```

4 Conclusions
We described one of the available methods for index-tracking optimal portfolio selection. Finally, a different approach for this problem based on heuristics is presented in [1].

References