S&P 500: Portfolio Optimisation via Factor Analysis

Introduction: Modern Portfolio Theory

Modern Portfolio Theory (firt proposed by economist **Harry Markowitz**) draws on the logical idea that, given the choice between two portfolios with the same expected return, you would prefer the portfolio with the smallest variance (or alternatively, the highest average return for a given level or risk). In this context, the variance of a portfolio is a direct metric for the portfolio's "risk".

Definition: A *Portfolio* of financial assets is defined as a set of weights $\boldsymbol{w} = (w_1, \ldots, w_p)$ such that $\boldsymbol{w}^T \mathbf{1}_p \leq 1$, where $\mathbf{1}_p$ is the p-length vector of 1's. This definition allows for *short-selling* (where an investor hold a negative quantity of a financial asset). In some cases the additional constraint that $w_i \geq 0$, $\forall i = 1 \ldots p$ which prevents short selling.

The idea of MPT boils down to a constrianed optimisation problem, taking the form

$$\max_{\boldsymbol{w}} \boldsymbol{w}^{\boldsymbol{T}} \boldsymbol{\mu}$$

$$\text{s.t. } \boldsymbol{w}^{\boldsymbol{T}} \boldsymbol{\Sigma} \boldsymbol{w} \leq \sigma_0^2, \quad \boldsymbol{w}^T \mathbf{1}_p \leq 1$$

Where $\boldsymbol{\mu}, \boldsymbol{\Sigma}$ are the mean and covariance of the returns respectively. That is we look for a portfolio that delivers the maximum expected return such that the variance of the portfolio does not exceed σ_0^2. The choice of σ_0^2 quantifies an investors aversion to risk.

In this form (i.e. no constraints on short-selling) there is a well known analytical solution to this optimization problem *see Proposition 2.1* (Bai, Liu, and Wong 2009):

$$\text{If } \sigma_0 B \leq \sqrt{A}, \text{ then } R = \sigma_0 \sqrt{A} \text{ and } \boldsymbol{w} = \frac{\sigma_0}{\sqrt{A}} \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}$$

$$\text{If } \sigma_0 B > \sqrt{A}, \text{ then } R = \frac{B}{C} + D(A - \frac{B^2}{C}) \text{ and } \boldsymbol{w} = \frac{1}{C} \boldsymbol{\Sigma}^{-1} \mathbf{1}_p + D(\boldsymbol{\Sigma}^{-1} \boldsymbol{\mu} - \frac{B}{C} \boldsymbol{\Sigma}^{-1} \mathbf{1}_p)$$

Where R denotes the expected return of the portfolio, $A = \boldsymbol{\mu}^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}$, $B = \mathbf{1}_p^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu}$, $C = \mathbf{1}_p^T \boldsymbol{\Sigma}^{-1} \mathbf{1}_p$ and $D = \sqrt{\frac{C\sigma_0^2 - 1}{AC - B^2}}$.

Factor Analysis for Portfolio Optimisation

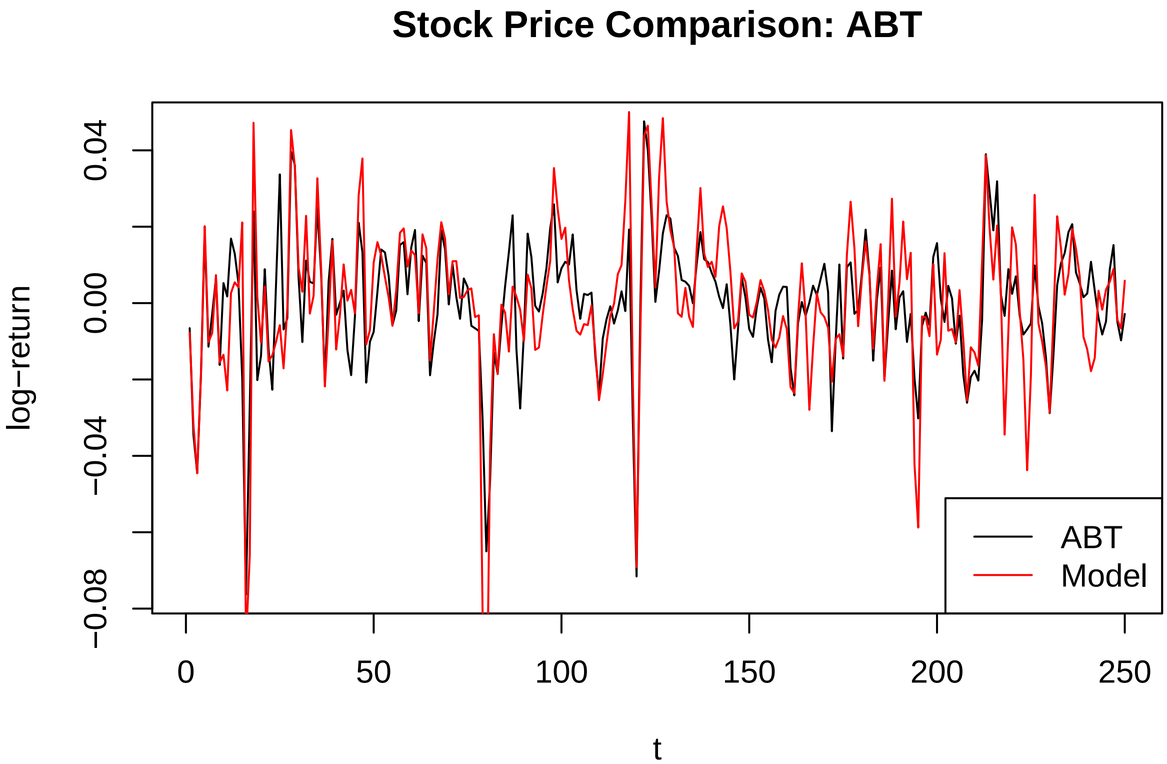
Following a similar approach to that described by (Chen, Huang, and Pan 2015), we use a *factor model* to model our data the S&P 500 stock prices. We have access to daily closing-bid prices from 2014 to 2019.

```
##   X.1       date   ABT  ABBV    ACN    ACT   ADBE  ADT   AES   AFL
## 1   1 2018-12-31 72.33 92.19 141.01 21.400 226.24 6.01 14.46 45.56
## 2   2 2018-12-28 71.09 91.12 139.82 21.359 223.13 6.11 14.27 44.95
## 3   3 2018-12-27 70.63 89.91 140.41 21.250 225.14 6.16 14.29 44.98
## 4   4 2018-12-26 69.62 89.04 139.01 21.560 222.95 6.29 14.28 44.14
## 5   5 2018-12-24 65.56 84.16 133.67 21.030 205.16 6.01 13.82 42.33
## 6   6 2018-12-21 67.27 84.92 137.20 21.420 208.80 6.21 14.45 43.23
```

The data is modelled as

$$y_{it} = \mu_i + \boldsymbol{\lambda}_i^T \boldsymbol{f}_t + u_{it} \Leftrightarrow \boldsymbol{Y} = \boldsymbol{\mu}\boldsymbol{1}_{\boldsymbol{n}}{}^T + \boldsymbol{\Lambda}^T \boldsymbol{F} + \boldsymbol{U}$$

With $\boldsymbol{Y} = (\boldsymbol{y}_1, \ldots, \boldsymbol{y}_n)$, $\boldsymbol{F} = (\boldsymbol{f}_1, \ldots, \boldsymbol{f}_n)$, $\boldsymbol{U} = (\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n)$, $\boldsymbol{\Lambda} = (\boldsymbol{\lambda}_1, \ldots, \boldsymbol{\lambda}_p)$. Taking the 2-day log-returns from 2016-01-01 to 2017-01-01 as a training set, we can now fit our model.



After fitting a model with 10 factors, we can see from this plot that the model roughly fits the first timeseries corresponding to ABT. Of course, it is important to remember that to properly assess the performance of the model we would need to compute the **testing error**. Unfortunately, since our model contains latent variables that must be estimated, we cannot evaluate our model at a point in the test set.

As a replacement metric for performance, we will construct a covariance matrix from the model and use MPT to find a portfolio. We will compare this to the simplistic strategy of dividing all capital equally amongst the assets.

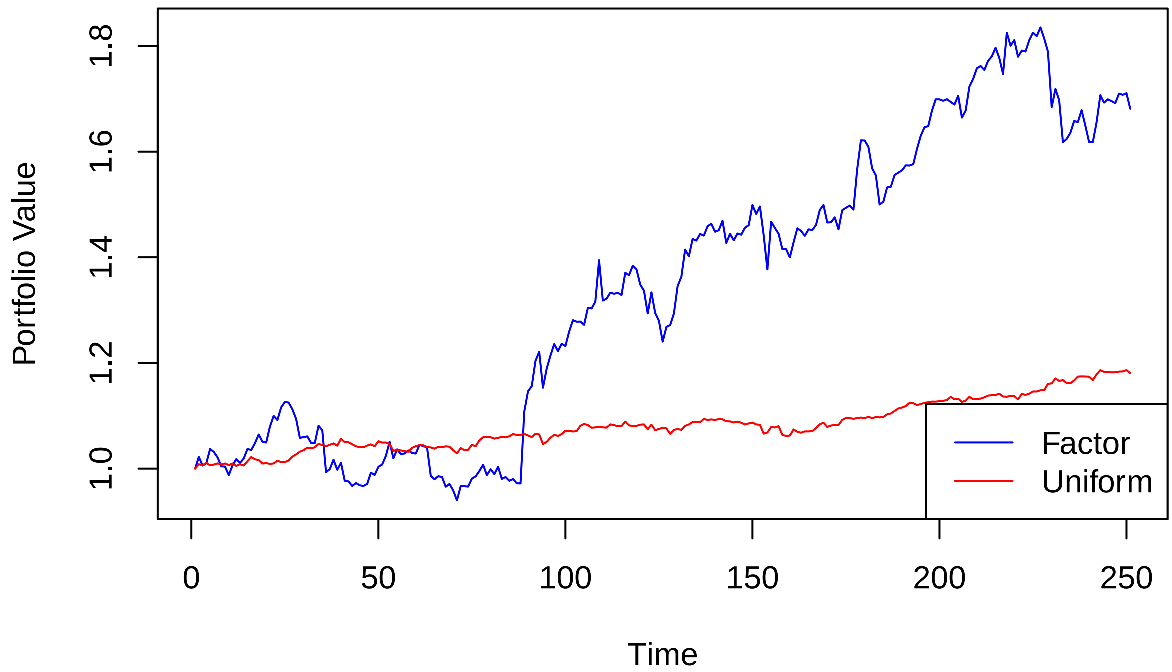
Here will will make use of a **quadratic programming** algorithm found in the `quadprog` package. This makes use of the dual method proposed by (Goldfarb and Idnani 1983) and solves the problem:

$$\min_{\boldsymbol{w}} \frac{1}{2} \boldsymbol{w}^T \boldsymbol{\Sigma} \boldsymbol{w} - \boldsymbol{\mu}^T \boldsymbol{w}$$

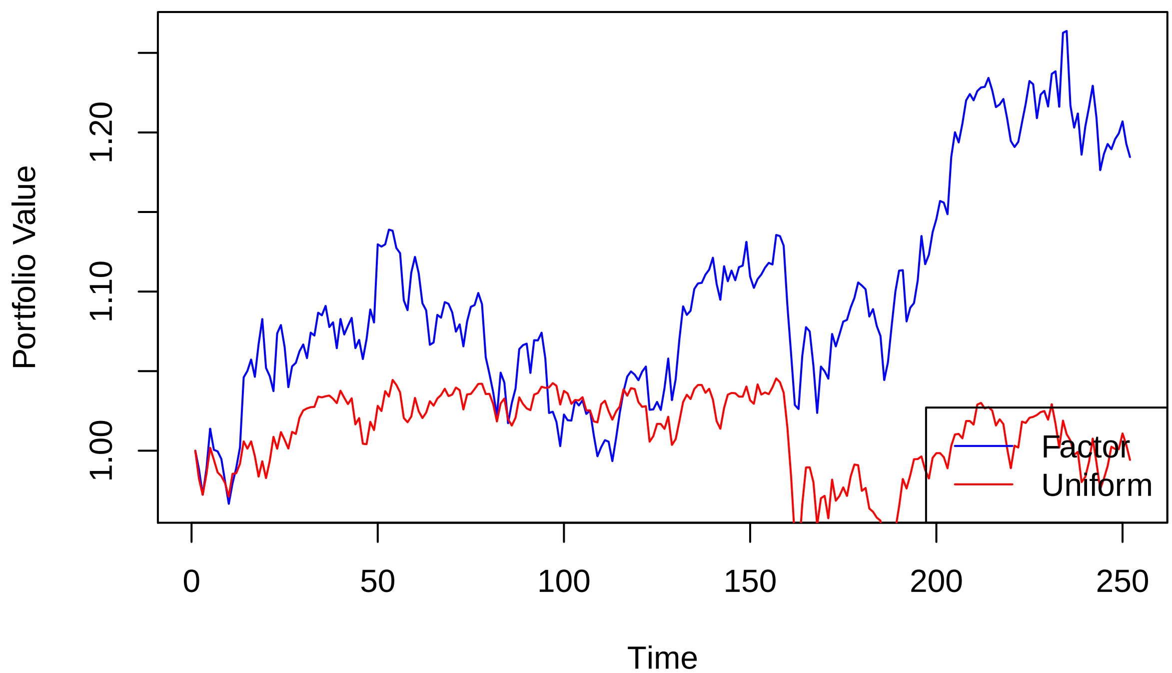
$$\text{S.t. } \boldsymbol{A}^T \boldsymbol{w} \geq \boldsymbol{b}_0$$

Instead of treating the variance as a constraint, this approach aims to minimize/maximize the variance/expected return of the portfolio directly. In the following implementation we allow the variance term to be multiplied by a positive constant $\lambda = 10$. A higher value of λ will result in a solution with a smaller variance.

Train: 2016-01-01 to 2017-01-01 Test: 2017-01-01 to 2018-01-01



Train: 2014-01-01 to 2015-01-01 Test: 2015-01-01 to 2016-01-01



Conclusion: Over seperate time periods the portfolios constructed via factor analysis seem to out-perform the uniformly distributed portfolio. This suggests there is a chance that our model contains some predictive power. **However**, without a proper metric such as the testing error, we do not have sufficient evidence to claim our model is successful at representing the stock market data.

References

Bai, Zhidong, Huixia Liu, and Wing-Keung Wong. 2009. "ENHANCEMENT of the Applicability of Markowitz'S Portfolio Optimization by Utilizing Random Matrix Theory." *Mathematical Finance* 19 (4): 639–67. https://doi.org/10.1111/j.1467-9965.2009.00383.x.

Chen, Binbin, Shih-Feng Huang, and Guangming Pan. 2015. "High Dimensional Mean–Variance Optimization Through Factor Analysis." *Journal of Multivariate Analysis* 133: 140–59. https://doi.org/https://doi.org/10.1016/j.jmva.2014.09.006.

Goldfarb, D., and A. Idnani. 1983. "A Numerically Stable Dual Method for Solving Strictly Convex Quadratic Programs." *Mathematical Programming* 27. https://doi.org/https://doi.org/10.1007/BF02591962.