Jamie Sian — 27775585
js17g15@soton.ac.uk

Coursework I

Question 1

Given two assets whose expected returns and co-variances on returns are:

$$\boldsymbol{\mu} = \begin{bmatrix} 0.1 \\ 0.1 \end{bmatrix} \quad \boldsymbol{\Sigma} = \begin{bmatrix} 0.005 & 0.0 \\ 0.0 & 0.005 \end{bmatrix}$$

We can derive the Markowitz efficient frontier as follows:

Consider two portfolios α and β, with weightings

$$\pi_\alpha = \begin{bmatrix} 1 & 0 \end{bmatrix} \quad \pi_\beta = \begin{bmatrix} 0.5 & 0.5 \end{bmatrix}$$

For returns distributions R_α and R_β, in the portfolios, denoted by

$$R_x \sim \mathcal{N}(\pi_x^T \mu,\, \pi_x^T \Sigma \pi_x)$$

The expected return $E[R_\alpha] = E[R_\alpha] = 0.1$, this is trivially equal $\forall\ \pi_x$ of size $n = 2$ where $\sum_{i=0}^{n} w_i = 1$; this shows that on the y axis we can plot 'mean of portfolio returns' as 0.1.

We consider α and β as special cases as they represent portfolios of least and greatest diversity, with the standard deviation of expected returns calculated as follows:

$$Var(R_\alpha) = \pi_\alpha^T \Sigma \pi_\alpha = 0.005 \quad \therefore \quad \sigma_\alpha = 0.0707$$
$$Var(R_\beta) = \pi_\beta^T \Sigma \pi_\beta = 0.0025 \quad \therefore \quad \sigma_\beta = 0.05$$

Therefore, the domain of the frontier 'standard deviation of portfolio returns' is such that $0.05 \leq x \leq 0.0707$.

As π_β provides the most diverse portfolio for no reduction in expected returns, a risk-averse investor should prefer this portfolio over any other π_x.

Question 2

Given three securities with the following expected returns and co-variances

$$\boldsymbol{\mu}_{ABC} = \begin{bmatrix} 0.10 \\ 0.20 \\ 0.15 \end{bmatrix} \quad \boldsymbol{\Sigma}_{ABC} = \begin{bmatrix} 0.005 & -0.010 & 0.004 \\ -0.010 & 0.040 & -0.002 \\ 0.004 & -0.002 & 0.023 \end{bmatrix}$$

In this question we aim to plot the efficient frontier of all security combinations, excluding those encompassing investment in a single security. As notation we label the securities in lexicographic order i.e. $\mu_A = 0.10$.

Using `MATLAB`'s financial toolbox, we can display the expected return on security investment for a given standard deviation (can be likened to risk level). To apply pair-wise security combinations and use them with the `setAssetMoments(portfolio,`μ`,`Σ`)` method of the financial toolbox, we can reconstruct μ_{XY} and Σ_{XY} by eliminating the column and rows relating of the excluded asset from the vector/matrix, for example:

$$\boldsymbol{\mu}_{BC} = \begin{bmatrix} 0.20 \\ 0.15 \end{bmatrix} \quad \boldsymbol{\Sigma}_{BC} = \begin{bmatrix} 0.040 & -0.002 \\ -0.002 & 0.023 \end{bmatrix}$$

As Figure 1 shows, portfolios along the ABC frontier display the highest expected return for any standard deviation of portfolio returns, this is because a portfolio on this frontier has the option of investing in at least one security that the other frontiers cannot; as expected greater portfolio choice offers increased opportunity for increasing expected returns at a given risk level. Furthermore achieving an expected portfolio return greater than 0.177 is only obtainable with the inclusion of B, however doing so increases risk, standard deviation of returns.

We can generate 10,000 random portfolios, composed of these securities and plot them onto a graph, as shown in Figure 1, to display the efficient frontier, in addition to sub-optimal portfolios that lay within it.

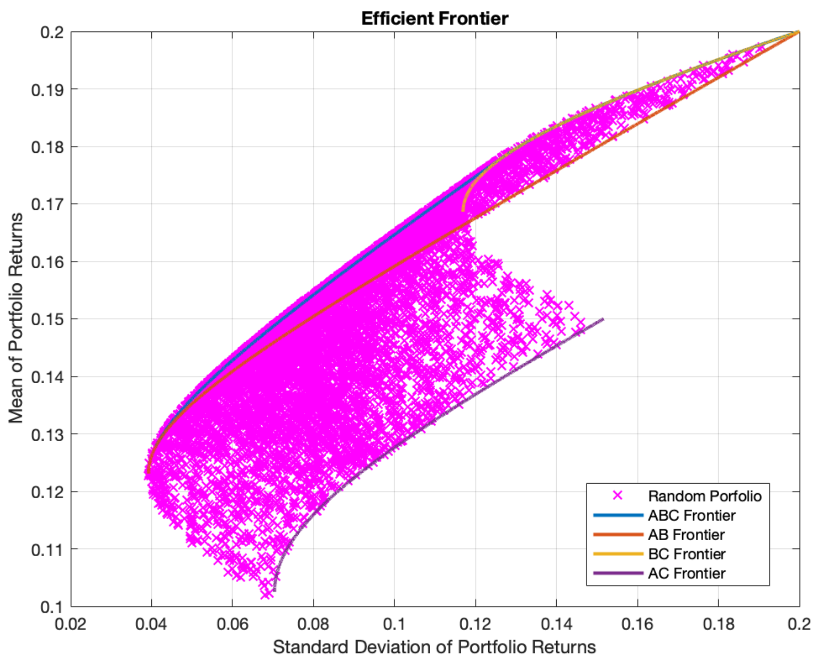

Figure 1: Plotting the efficient frontier for different security combinations

Question 3

The Convex Programming (`CVX`) toolbox for `MATLAB` provides us with a capability that can be used to model convex problems and apply constraints to them.

In the `NaiveMV` method there exists two calls to `quadprog` and one to `linprog`. These calls determine the weights associated with the maximum return and the minimum variance. As a result of this, we can iterate over n points using `quadprog` to trace the frontier.

To find the portfolio weights that maximise the expected return of the portfolio we use Equation 1

$$\max_{\pi} \ \boldsymbol{\pi}^{\boldsymbol{T}}\boldsymbol{\mu} \text{ where } \sum_{i=0}^{n} \boldsymbol{\pi}_i = 1 \ \wedge \ \forall i \in \boldsymbol{\pi}, \ i \geq 0 \tag{1}$$

Similarly we can calculate the minimum variance weights irrespective of return using Equation 2

$$\min_{\pi} \ \boldsymbol{\pi}\boldsymbol{\Sigma}\boldsymbol{\pi}^{\boldsymbol{T}} \text{ where } \sum_{i=0}^{n} \boldsymbol{\pi}_i = 1 \ \wedge \ \forall i \in \boldsymbol{\pi}, \ i \geq 0 \tag{2}$$

Both of these convex optimisations can be expressed using the `CVX` toolkit. To replace the call to `linprog` we can maximise the expected return irrespective of risk, then return the associated portfolio weights using the following code:

```
cvx_begin
variable MaxReturnWeights(NAssets)
    maximize( MaxReturnWeights' * ERet )
    subject to
        V1 * MaxReturnWeights == 1;
        MaxReturnWeights >= 0;
cvx_end
```

In a similar manner we can find the lower variance bound of the frontier irrespective of return, by minimizing $\boldsymbol{\pi\Sigma\pi^T}$ as in Equation 2.

Whilst the `CVX` toolkit provides additional control that can be leveraged to manage constraints, such as enabling of short selling, it lacks the speed of `NaiveMV`; for Question 4 we revert to the implementation using `quadprog` and `linprog`.

Question 4

For this section we consider historic daily data of the `FTSE 100` between the dates of 23/02/2016 - 22/02/2019. Of the `FTSE 100` we fetch the top 29[1] listed on Yahoo Finance[2] for which closing daily stock data was available.

In this task three stocks were selected at random from the 29 and we attempted to generate the most efficient portfolio. In this instance we study the case of *TUI*, *RMV* and *BA*.

The sample mean and co-variance matrix for these stocks from the first 18 months (P_1) are as follows:

$$\hat{\boldsymbol{\mu}}_{TRB} = 10^3 * \begin{bmatrix} 1.0851 \\ 0.4034 \\ 0.5705 \end{bmatrix} \qquad \hat{\boldsymbol{\Sigma}}_{TRB} = 10^3 * \begin{bmatrix} 6.8455 & 0.8220 & 3.1441 \\ 0.8220 & 0.3948 & 0.2940 \\ 3.1441 & 0.2940 & 3.3519 \end{bmatrix}$$

In this scenario the goal of the optimisation is to maximise profit, or at least minimise the loss of a given portfolio. To translate findings into real returns we consider a 10,000 GBX (£100) investment across these stocks, and for the purposes of this assignment assume that the stocks can be split into $\mathbb{R}$.

$$P_{1,\,end} = 10^3 * \begin{bmatrix} 1.3300 \\ 0.4028 \\ 0.5920 \end{bmatrix} \qquad P_{2,\,end} = \begin{bmatrix} 833.2000 \\ 478.0500 \\ 470.8000 \end{bmatrix}$$

Assuming purchase at the end of P_1, and sale at the end of P_2 (where only P_1 is known for sampling), when the efficient frontier is generated, the following weights represent the maximum value for which each asset x was allocated for the set of $\boldsymbol{\pi}_x$ on the efficient frontier. In addition to this, after scaling profit against

[1] We consider 29 of the 30 listed on Yahoo Finance as the 30$^{\text{th}}$ (RDSA) was missing a significant proportion of data for 2018
[2] https://uk.finance.yahoo.com

investment we can determine $E[R_X]$ in GBX.

$$\boldsymbol{\pi}_T = \begin{bmatrix}0.0000 & 0.9684 & 0.0316\end{bmatrix} \quad E[R_T] = 1.9$$
$$\boldsymbol{\pi}_R = \begin{bmatrix}0.0000 & 1.0000 & 0.0000\end{bmatrix} \quad E[R_R] = 13.9$$
$$\boldsymbol{\pi}_B = \begin{bmatrix}0.0000 & 0.9687 & 0.0313\end{bmatrix} \quad E[R_B] = 1.9$$

Furthermore we consider the Naive 1/N allocation, described as follows:

$$\boldsymbol{\pi}_{1/N} = \begin{bmatrix}\frac{1}{3} & \frac{1}{3} & \frac{1}{3}\end{bmatrix} \quad E[R_{1/N}] = -730$$

As shown in Figure 2 heavy investment in RMV proved the most logical choice in terms of the associated Gaussian distributions; if history was a good indicator of future performance, over the period P_2, the stock values would likely tend towards $\hat{\boldsymbol{\mu}}_{TRB}$. For TUI and BA, if this was the case these components of the portfolio would result in a loss; whilst according to this model RMV should provide a slight profit.

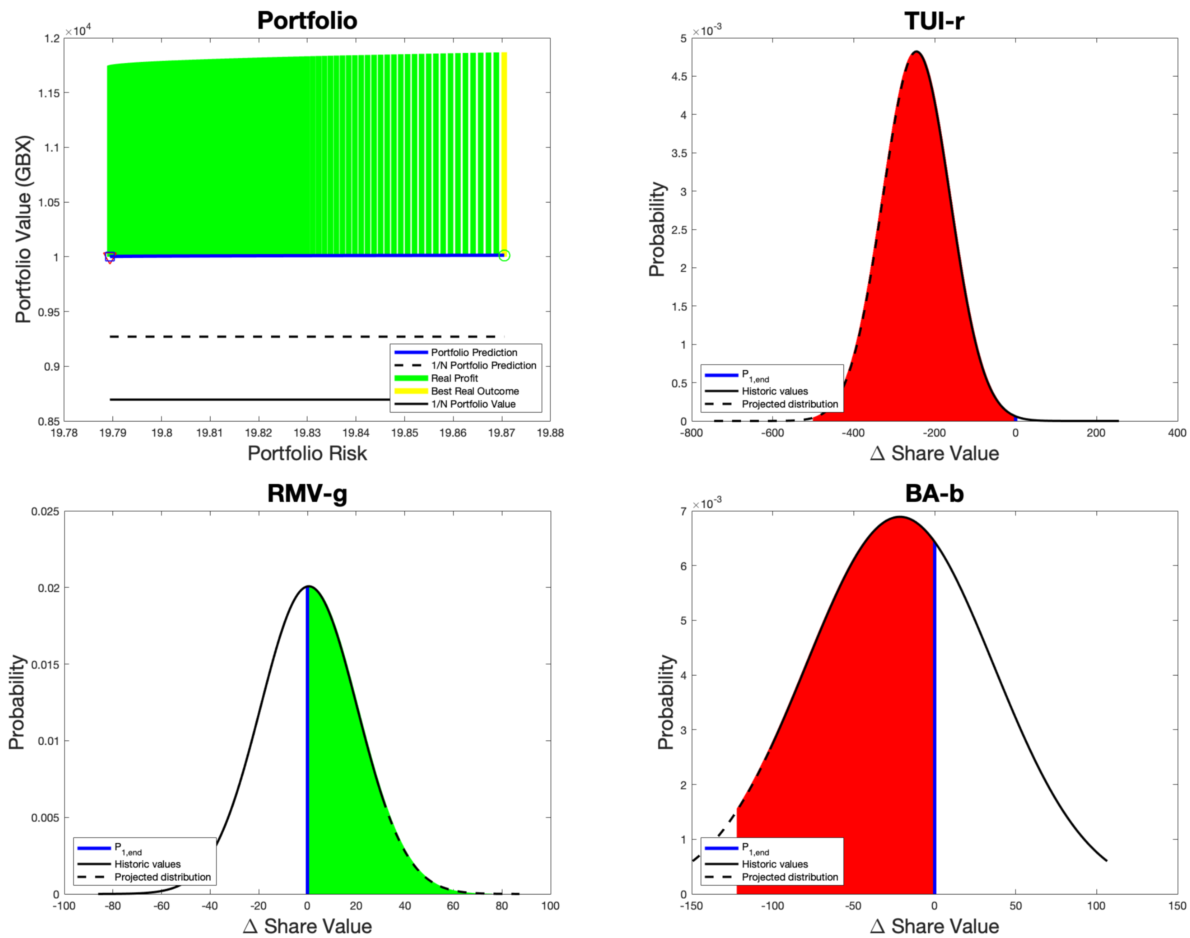

Figure 2: Graphical representation of creating a portfolio with TUI, RMV and BA. The green areas on the RMV Gaussian shows the increase in share value from $P_{1,end}$ to $P_{2,end}$, where the green area stops. Conversely the red areas on the Gaussians show a loss from $P_{1,end}$ to $P_{2,end}$, where the red area stops.

Upon analysis ex-post, whilst RMV presents the only non-bearish stock, the increase in value of RMV, given $\hat{\boldsymbol{\mu}}_{TRB}$ was very large according to the Gaussian model, so much so that RMV had not seen values this high in

P_1 (denoted by the dashed line rather than a solid black distribution boundary). Furthermore, with respect to allocation, in allocating a heavy weight to RMV, portfolios $\boldsymbol{\pi}_T$, $\boldsymbol{\pi}_R$, and $\boldsymbol{\pi}_B$ were able to avoid the, larger than expected, losses associated with TUI and BA.

Interestingly this correlation between airline losses is by no means a fluke; as fuel prices and wage prices increase in particular in this sector recently[3] it is no surprise that in both cases record lows in stock price were achieved in P_2 as the model was trained before this period of increased volatility.

This shows us that, whilst modelling based on history is a useful tool, stock prices can swing due to as a result of unpredictable short term changes.

Question 5

In this section we aim to implement the MacKinlay & Pastor model [1] to improve the estimation of expected returns.

The idea presented links incorrect pricing, due to a missing factor in the factor model, to the residual co-variance matrix. If this is the case an asset will be able to be leveraged into asymptotic arbitrage opportunities; enabling an investor to make an instantaneous profit in the market.

In order to derive the appropriate parameters required to estimate the co-variance matrix, we can format their derivation as the following multivariate constrained optimisation problem.

$$\max\ \mathcal{L}(\alpha, \theta_h, \sigma^2 | z_1, ..., z_n) \propto |\alpha\alpha'\theta_h + \sigma^2 \boldsymbol{I}|^{-T/2} \exp\{-\frac{1}{2}\sum_{t=1}^{T}(z_t - \alpha)'(\alpha\alpha'\theta_h + \sigma^2\boldsymbol{I})^{-1}(z_t - \alpha)\} \tag{3}$$

$$\max\ l(\alpha, \theta_h, \sigma^2 | z_1, ..., z_n) \propto -\frac{T}{2} \cdot \ln(|\alpha\alpha'\theta_h + \sigma^2\boldsymbol{I}|)\ -\frac{1}{2}\sum_{t=1}^{T}(z_t - \alpha)'(\alpha\alpha'\theta_h + \sigma^2\boldsymbol{I})^{-1}(z_t - \alpha) \tag{4}$$

$$\therefore \quad \Sigma = \alpha\alpha'\theta_h + \sigma^2\boldsymbol{I} \tag{5}$$

The key variables in the above equations are α, the expected mean of the distribution; θ_h the inverse of the Sharpe ratio squared; σ^2, the expected variance of the data and the vectors z_t representing the returns on the portfolio at time step t. Interestingly as θ_h is the inverse of the Sharpe ratio squared ($\frac{\sigma^2}{\mu^2}$) this can provide a view of model skepticism as if $\theta_h \to \infty$ the reliability of the model is seen to be perfect, whereas as $\theta_h \to 0$ we have no confidence in the model; furthermore implying for constrained optimisation it is always the case that $\theta_h \geq 0$. In a similar manner, it is always the case that $\sigma^2 \geq 0$ as a distribution can never have an imaginary standard deviation.

In MacKinlay & Pastor's paper they provide the formula for Likelihood $\mathcal{L}$, as shown in Equation 3. However upon implementation in `MATLAB` the program would crash and return x_0, the starting point specified in the `fminsearchcon` method, used for constrained multivariate search. To fix this we can take the natural log of $\mathcal{L}$, to form Equation 4. Notice that in this equation, the exponents $\frac{-T}{2}$ in addition to the general exp term have been cancelled; reducing the computational complexity of running this function in a constrained optimisation problem.

[3] https://uk.reuters.com/article/us-airlines-wages-inflation-analysis/higher-wages-fuel-prices-turn-up-cost-pressure-on-airlines-idUKKCN1FY292

Whilst this did help in terms of computational complexity, the mean vector produced did not accurately reflect the stocks in question; as a result, applying these values to Equation 5 produced equally anomalous expected returns. To rectify this, and further optimise expected tracking for an environment such as a trading floor, I would replace the call to `fminsearchcon` with a Newton/Quasi-Newton numerical optimisation procedure, dependent upon what level of speed/accuracy was required as Newton is more accurate but more computationally expensive. Furthermore I would look to condense Equation 4 into an equation that performs the same functionality at reduced computational cost.

Question 6

To complete the Index tracking task, in addition to the 29 `FTSE 100` components, data for the `FTSE 100` as an index was required for the same period, we denote this vector as $\boldsymbol{y}$. As with previous sections, to maintain a better view of the results, we create all portfolios with an investment of 10,000 GBX (£100)

Question 6a

To greedily track the `FTSE 100` we minimise the following function:

$$\min_{\pi} ||\boldsymbol{y} - \boldsymbol{R}\boldsymbol{\pi}||_2^2 \quad \text{where } ||\boldsymbol{\pi}||_0 = \boldsymbol{\pi}_0 \ \wedge \ \boldsymbol{\pi}^T \mathbf{1}_N = 1 \tag{6}$$

To achieve this in implementation, `fmincon` was used with the non linear constraint $||\boldsymbol{\pi}||_0 = \boldsymbol{\pi}_0$, as this cannot be calculated within the `CVX` environment as the 0^{th} norm is not convex. Upon reaching a solution this implementation takes the top 6 stocks and redistributes any remaining wealth among these top 6 according to their weight; leaving other indexes with weights of zero.

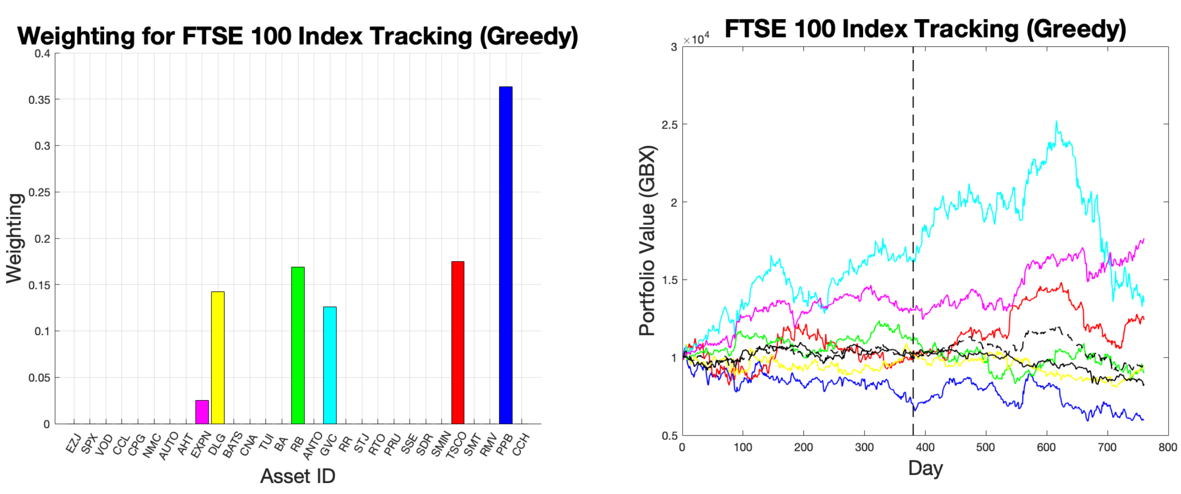

Figure 3: Index tracking of the `FTSE 100` using a Greedy approach, assuming investment of 10,000 GBX. In this graph the solid black series represents the real performance of the `FTSE 100` given investment. The dashed black series represents the index performance. The vertical dashed black line distinguishes between P_1 (train) and P_2 (test). The coloured series represent singleton portfolio performance.

Question 6b

In their paper Brodie et al. [2] introduce the notion of implementing a penalty term proportional to the sum of the absolute values of the portfolio weights, a form of lasso regression, as shown below.

$$\min_{\pi} ||\boldsymbol{y} - \boldsymbol{R}\boldsymbol{\pi}||_2^2 + \tau||\boldsymbol{\pi}||_1 \tag{7}$$

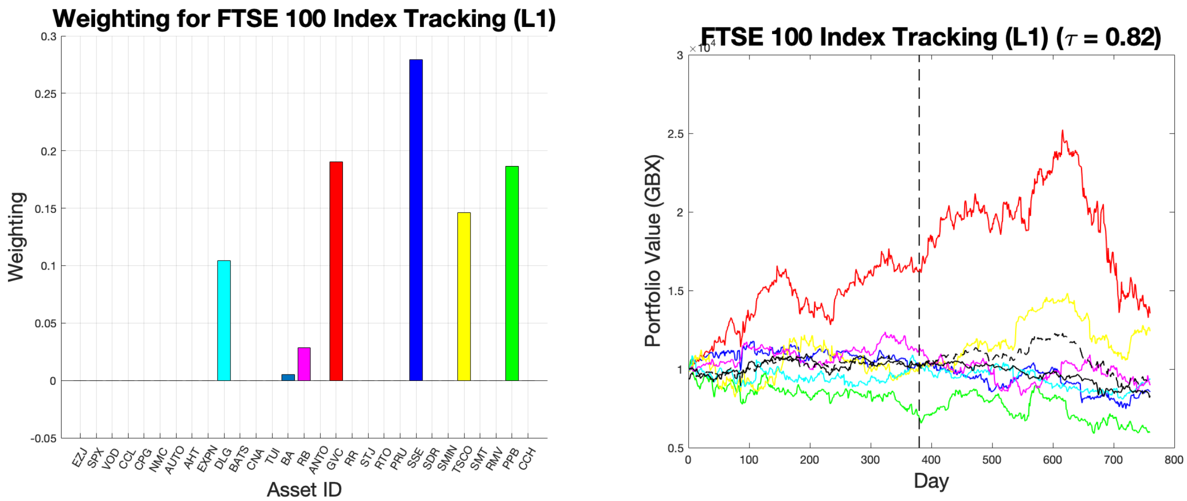

Figure 4: Index tracking of the `FTSE 100` using L1 regularization ($\tau = 0.82$), assuming investment of 10,000 GBX. In this graph the solid black series represents the real performance of the `FTSE 100` given investment. The dashed black series represents the index performance. The vertical dashed black line distinguishes between P_1 (train) and P_2 (test). The coloured series represent singleton portfolio performance.

The behaviour of ensuring the sum of the weights is 1 causes the portfolio to minimise solely based on the initial l_2 norm; negating the use of a penalty term at all. This implication is mentioned in their paper. In increasing the value of τ past the critical value τ_0 (where negative weights cease to exist) we can increase the sparsity of π; creating a less diverse portfolio that best reflects $\boldsymbol{y}$.

In this scenario, we find that $\tau_0 = 0.81$. The weights of the portfolio can be found in Table 1

Greedy			L_1		
Stock	Weight	Colour	Stock	Weight	Colour
PPB	0.36337	Blue	SSE	0.27913	Blue
TSCO	0.17467	Red	GVC	0.19062	Red
RB	0.16876	Green	PPB	0.18660	Green
DLG	0.14220	Yellow	TSCO	0.14615	Yellow
GVC	0.12600	Cyan	DLG	0.10429	Cyan
EXPN	0.02500	Magenta	RB	0.02835	Magenta

Table 1: Weights allocated via Greedy and L_1 index tracking

As shown in Figures 3, 4 and Table 2, the group of shares encompassed by the lasso technique tracked the `FTSE 100` better in training, however in testing the Greedy strategy provided a lower mean square error; a better tracking of the `FTSE 100`. Interestingly, as shown in Table 1 the stocks chosen were the same for one or two exceptions.

	Greedy	L_1
Training	9.1368e+04	8.3135e+04
Testing	1.4746e+06	1.7669e+06

Table 2: Mean Square error of index tracking against the `FTSE 100` data

Question 7

When modelling a market it is necessary to include transaction costs to best describe reality; thus modelling expected returns as accurately as possible in order to maximise profit. Transaction costs can take many different forms, and model a variety of external factors. Cases such as tax on a high value trade may be something to consider for an investor to shift their risk perspective based on returns. A further example would be the expectation of additional transaction cost across a bid-ask spread; if a seller and a buyer cannot decide on a price (reach a market equilibrium) their perspective on demand or supply may have to shift less favourably and accept the additional cost as a transaction cost [3].

Costs can also be considered as fixed or linear in nature. Linear transaction costs are those where the transaction cost per unit is the same regardless of quantity ordered; there is no discount or subsidy for bulk purchases. An example of a fixed cost would be annual salaries paid to employees of a company. There also exist interactions between these, for example break points where shares are bought; buying a quantity of shares associated with a quantity band may provide a discounted transaction cost per unit compared to a lower volume purchase.

In Figure 2 of their paper Lobo et al. [3] , display a cumulative distribution function for a normal distribution, composed of the following problem:

$$\begin{aligned}
\text{maximize} \quad & \bar{\alpha}^T(w + x^+ - x^-) \\
\text{subject to} \quad & \mathbf{1}^T(x^+_{101} - x^-_{101}) + \sum_{i=1}^{100}(\alpha_i^+ x_i^+ + \alpha_i^- x_i^-) \leq 0 \quad \wedge\, w_i + x_i^+ - x_i^- \geq s_i \quad i = 1, ..., 100 \\
& x_i^+ \geq 0, \quad x_i^- \geq 0 \quad i = 1, ..., 100 \qquad \wedge\, \Phi^{-1}(\eta_j)\,||\Sigma^{1/2}(w + x^+ - x^-)|| \quad ... \\
& \qquad \leq \bar{\alpha}^T(w + x^+ - x^-) - W_j^{Low} \; j = \{1, 2\}
\end{aligned} \tag{8}$$

The prime objective of this problem is to maximise the wealth of the investor given transactions. In addition to this we add five constraints, firstly the problem must consider that in this space there exists one risk-free asset and 100 risk-prone ones, an example of a risk-free asset could be U.S Treasury Bill, whereas a risk-prone one could be shares.

In addition to this the constraints enforce a short selling constraint s_i for all assets.

Thirdly, the condition $x_i^+ \geq 0, x_i^- \geq 0 \quad i = 1, ..., 100$ specifies that transaction costs cannot be negative.

Finally the two shortfall constraints $\Phi^{-1}(\eta_{\{1,2\}})$ provide bad and worst case scenarios for which one criteria could be used as an alert, and the other as an automatic sell before the investor makes a loss on the the asset.

Using the cumulative distribution in Figure 2 [3] we can be calculate expected returns for a given risk level risk by reading the corresponding z value on the x axis. When reading off a probability, we can consider a probability p as confidence η that the expected return z will be made where $\eta = 1 - p$.

In order to implement this exact specification into this project we would have to ascertain the following values appropriate to the FTSE 100:

$$\eta_{\{1,2\}} \quad x_i^+ \quad x_i^- \quad s_i$$

After these constraints were found we could implement this structure in `CVX` as all of the portfolio constraints described are convex; this is a second-order cone problem.

References

[1] A. Craig MacKinlay and L. Pástor, "Asset pricing models: Implications for expected returns and portfolio selection," *The Review of financial studies*, vol. 13, no. 4, pp. 883–916, 2000.

[2] J. Brodie, I. Daubechies, C. De Mol, D. Giannone, and I. Loris, "Sparse and stable markowitz portfolios," *Proceedings of the National Academy of Sciences*, vol. 106, no. 30, pp. 12 267–12 272, 2009.

[3] M. S. Lobo, M. Fazel, and S. Boyd, "Portfolio optimization with linear and fixed transaction costs," *Annals of Operations Research*, vol. 152, no. 1, pp. 341–365, 2007.