Advanced Empirical Finance - Assignment 1

Pablo Serrano Ascandoni, Lukas Malte Kemeter

Question 1

The purpose of the assignment is to replicate the main findings of Jobson and Korkie (1980). They show that using finite sample moments to approximate the ture moments relies on convergence of the estimates as $T \to \infty$. This way, portfolio weights obtained from rather small samples can substantially differ from the optimal portfolio weights in a simple mean variance optimisation problem. As stated in Brand (2010)[1], these asymptotic properties seem to be unreliable and using plug-in estimates can even in large finite samples lead to sub-optimal performance, especially for a large number of assets. Consider the returns of N risky assets whose mean is given by μ and variance-covariance matrix by Σ. The investor wants to optimise the portfolio by maximising the expected return and minimising the volatility:

$$\max w'\mu - \frac{\lambda}{2} w'\Sigma w \;\text{ s.t }\; w'\iota = 1 \tag{1}$$

where λ is the risk aversion. Taking the historical sample moments μ and Σ as the true moments (denoted as *"mu"* and *"sigma"* in the code) we simulate returns and obtain plug-in estimates $\hat{\mu}$ and $\hat{\Sigma}$ for smaller values of T (denoted as *"mu_hat"* and *"sigma_hat"* in the code) and optimise the portfolio weights. Once we compare the performance of the estimates against the true values, we find that the estimated value of the Certainty Equivalence is strictly lower than its true value, assuming some uncertainty in the estimation.

$$w'\hat{\mu} - \frac{\lambda}{2} w'\hat{\Sigma} w < w'\mu - \frac{\lambda}{2} w'\Sigma w \tag{2}$$

For this replication exercise, a data set from the Kenneth French website is used, which contains monthly returns for 10 different industries from July 1926 up to November 2020. Table 1 introduces the different industry classifications used in the data:

Table 1: Industry classifications

NoDur	Consumer Nondurables
Durbl	Consumer Durables
Manuf	Manufacturing
Enrgy	Oil, Gas, Coal
HiTec	Business Equipment
Telcm	Telephone, Television
Shops	Wholesale, Retail
Hlth	Healthcare, Medical Equipment, Drugs
Utils	Utilities
Other	Mines, Construction, Finance and others
Source:	*Kenneth French Homepage*

For the $(N \times 1)$ vector of returns r_t where $N = 10$, the estimated moments are:

[1] Portfolio Choice, Chapter 5.

$$\hat{\mu} = \frac{1}{T}\sum_{t=1}^{T} r_t, \text{ and } \hat{\Sigma} = \frac{1}{T-1}\sum_{t=1}^{T}(r_t - \hat{\mu})(r_t - \hat{\mu})'$$

Using the full available history of monthly returns, we estimate μ and the true variance-covariance matrix Σ using the same formulas. These are provided in Table 2 and 3.

Table 2: Average returns (considered to be the true μ)

NoDur	Durbl	Manuf	Enrgy	HiTec	Telcm	Shops	Hlth	Utils	Other
0.96	1.15	1.03	0.96	1.14	0.86	1.04	1.09	0.87	0.91

Table 3: Variance-Covariance matrix of returns (considered to be the true Σ)

	NoDur	Durbl	Manuf	Enrgy	HiTec	Telcm	Shops	Hlth	Utils	Other
NoDur	21.07	26.53	24.27	18.04	24.37	14.54	23.04	20.26	17.77	24.76
Durbl	26.53	61.98	42.33	30.82	43.70	23.03	36.20	27.52	25.79	40.45
Manuf	24.27	42.33	38.88	28.85	38.75	19.88	30.73	26.26	23.65	36.33
Enrgy	18.04	30.82	28.85	40.09	27.74	15.46	21.65	19.73	20.78	28.16
HiTec	24.37	43.70	38.75	27.74	51.97	22.72	33.33	28.83	24.29	37.04
Telcm	14.54	23.03	19.88	15.46	22.72	21.28	18.27	15.56	15.86	21.02
Shops	23.04	36.20	30.73	21.65	33.33	18.27	33.77	24.03	20.62	30.82
Hlth	20.26	27.52	26.26	19.73	28.83	15.56	24.03	30.78	18.76	26.32
Utils	17.77	25.79	23.65	20.78	24.29	15.86	20.62	18.76	30.12	25.22
Other	24.76	40.45	36.33	28.16	37.04	21.02	30.82	26.32	25.22	41.12

The highest Sharpe ratio ($SR_i := \mu_i/\sigma_i$)[2] across the whole observation period was generated by the Non-Durable (*NoDur*) industry portfolio with a value of $SR_{NoDur} = 0.21$, as reported in Table 4.

Table 4: Sharpe ratio for each industry

NoDur	Durbl	Manuf	Enrgy	HiTec	Telcm	Shops	Hlth	Utils	Other
0.21	0.15	0.16	0.15	0.16	0.19	0.18	0.2	0.16	0.14

Question 2

In question 2 we define the function: $compute\_efficient\_frontier(sigma\_hat, mu\_hat, sigma\_true, mu\_true)$ which takes $(\hat{\mu}, \hat{\Sigma}, \mu, \Sigma)$ as inputs. The function computes the minimum variance portfolio (MVP) weights (w_{mvp}) as well as the efficient portfolio (EFF) ones (w_{eff}) that delivers 2 times the expected return of the MVP weights. It then returns a tibble with the set of attainable portfolios (μ_i^*, σ_i^*) applying the Mutual Fund Theorem. The optimization problem to obtain the MVP weights is:

$$w_{mvp} = \arg\min_{w} w'\Sigma w \;\; s.t. \;\; w'\iota = 1 \;\; \rightarrow \;\; w_{mvp} = \frac{\Sigma^{-1}\iota}{\iota'\Sigma^{-1}\iota}$$

where ι is a vector of 1s, and the EFF weights for a $\bar{\mu} = 2w_{mvp}'\mu$ are obtained from:

$$w_{eff} = \arg\min_{w} w'\Sigma w \;\; s.t. \;\; w'\iota = 1, \; w'\mu \leq \bar{\mu} \;\; \rightarrow \;\; w_{eff} = w_{mvp} + \frac{\tilde{\lambda}}{2}(\Sigma^{-1}\mu - \frac{D}{C}\Sigma^{-1}\iota)$$

where $\tilde{\lambda} = 2\frac{\bar{\mu}-D/C}{E-D^2/C}$, $C = \iota'\Sigma^{-1}\iota$, $D = \mu'\Sigma^{-1}\iota$ and $E = \mu'\Sigma^{-1}\mu$.

[2]This Sharpe ratio formulation assumes a risk-free rate of $r_{rf} = 0$

According to the Mutual Fund Theorem theorem, you can characterize the optimal portfolio as defined by portfolio return and volatility by combining the MVP and the efficient portfolio as follows:

$$\mu_i^* = w_i^{*\prime}\mu, \;\; \sigma_i^* = w_i^{*\prime}\Sigma w_i^* \;\; where \;\; w_i^* = c_i w_{mvp} + (1-c_i)w_{eff}, \;\; \forall c_i \in \{-0.10, -0.11, ..., 1.19, 1.20\}$$

Question 3

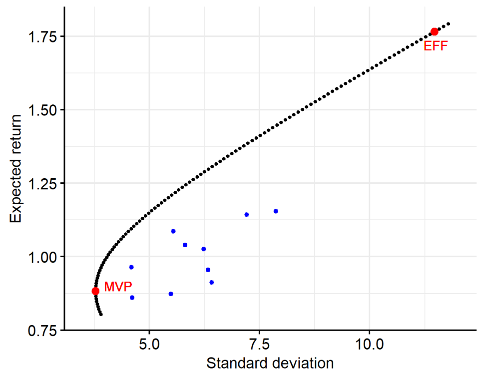


Figure 1: Full sample efficient frontier

In order to generate Figure 1, we call the function $compute\_efficient\_frontier(sigma\_true, mu\_true, sigma\_true, mu\_true)$ from above with the true parameters μ and Σ from the full sample as inputs. We then plot the resulting tibble to visualize the theoretically optimal efficient frontier. Recall that our functions returns the combinations of expected return and volatility so each point of the frontier corresponds to one row of the generated tibble. The red dots indicate the return and volatility of the minimum variance portfolio (on the left) and the efficient portfolio with twice the returns of the MVP (on the right). Blue dots indicate the volatility and return for each of the 10 industries.

Question 4

Assuming a risk-free rate ($r_{rf} = 0$), we obtain the tangent portfolio while allowing short-selling of assets. The tangent portfolio (tgc) is the one with the highest Sharpe ratio. By holding combinations on the tangent portfolio and the risk-free asset, investors with different risk aversions can attain every risk-return combination along the Capital Market Line: $r_p = r_{rf} + SR_{tgc}\sigma_p$. The tangent portfolio is:

$$w_{tgc} = \frac{\Sigma^{-1}(\mu - r_{rf})}{\iota'\Sigma^{-1}(\mu - r_{rf})}$$

Table 5 reports the tangent portfolio weights (negative weights imply short-selling).

Table 5: Tangency portfolio weights

NoDur	Durbl	Manuf	Enrgy	HiTec	Telcm	Shops	Hlth	Utils	Other
0.75	0.05	-0.14	0.18	0.04	0.37	0.01	0.26	0.05	-0.58

Figure 2 then visualizes the tangent portfolio (TGC) as a purple dot and its Sharpe ratio as the slope in that point on the previously introduced efficient frontier. The Sharpe ratio of the tangent portfolio is: $SR_{tgc} := (\mu_{tgc} - r_{rf})/\sigma_{tgc} = 0.25$. The line joining the risk-free asset (the origin (0,0) in this case) and the tangent portfolio is the Capital Market Line. In this specific case, the Capital market Line is the same as the slope in the purple point because the risk free rate is set to 0. This CML is obtained with the the possible combinations of the tangent portfolio and the risk free rate, if we allow for short-selling.

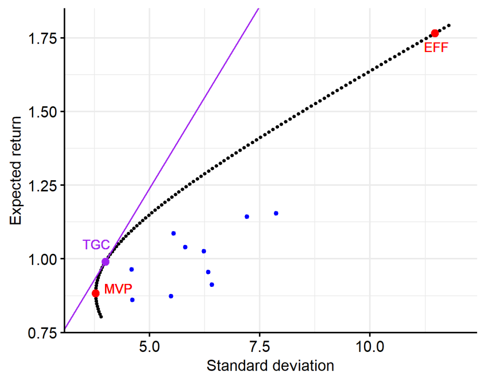


Figure 2: Full sample efficient frontier with tangent line

It can be seen that the tangent portfolio is not very balanced as it assigns a weight of 0.75 to *NoDur* while *Other* is shorted with a weight of -0.58. The tangent portfolio is also a lot closer to the minimum variance portfolio than to the efficient portfolio for a desired return that doubles that of the MVP. Every linear combination of the risk-free asset and the tangent portfolio will dominate the efficient frontier in terms of return relative to volatility as all points of the CML are strictly higher than the frontier. When implementing the tangent portfolio in reality, one might face issues such as transaction costs (which could also depend on the liquidity of certain stocks). Such transaction costs would affect the investor's ability to rebalance his portfolio towards the tangent one (see Hautsch et. al (2019). Short-selling means to borrow a share and give it back to the lender at an agreed time. This in itself has its own implementation risks, which are in reality hard to account for. If the price increases, the short-seller might get squeezed and forced to take losses when the short positions have to be closed at the agreed time[3].

Question 5

In order to generate a hypothetical return sample based on the true moments μ and Σ, we define a function $simulate\_returns(T, mu, sigma, seed)$ which draws a sample of length T from a normal distribution $N(\mu, \Sigma)$ with the true moments of the data using the function: $mvrnorm(n = T, mu, sigma)$. To replicate our results, pass the seed 2020. The function returns a data frame with returns as the only column.

Question 6

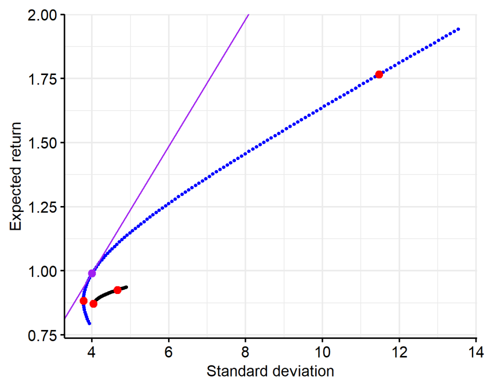


Figure 3: Efficient Frontier using full sample (blue) and simulated data (black)

Now we use the previously constructed function $simulate\_returns(T, mu, sigma, seed)$ from question 5

[3] See the Volkswagen Short Squeeze in 2008 or the Reddit driven GME price spike.

and set T to 100 in order to simulate a hypothetical return sample using the true μ and true Σ. Figure 3 displays the efficient frontier of the original data set in dark blue together with the estimated efficient frontier from the sample of size $T = 100$ in back. The red dots indicate the MVP and the efficient portfolio for each efficient frontier. The frontier was obtained using the previously explained functions to estimate $(\hat{\mu}, \hat{\Sigma})$ for the sampled data set. Both frontiers seem to differ significantly and changing the seed will change the position and shape of the simulated black line. It is clear that simulating with a relatively small sample size, the simulated efficient frontier is quite far from the optimal one.

Question 7

We then repeat the simulation 250 times by looping over 250 different seeds and calling the function $simulate\_returns(T, mu, sigma, seed)$ with $T = 100$ each time. Here we use the seeds ranging from 1 to 250 to obtain 250 samples and compute their respective frontiers. We initialize a plot containing the original efficient frontier in blue. The calculated efficient frontier for each iteration is then added to the existing plot in black. For this rather small sample size, the simulated efficient frontiers in figure 4 differ significantly from the true frontier and they all are below, meaning that the attainable return-volatility combination will always be worse than the optimum (meaning lower Sharpe ratios).

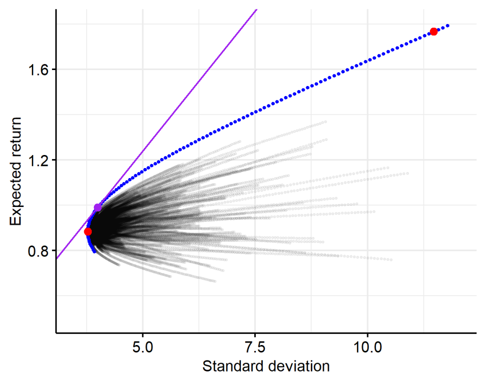
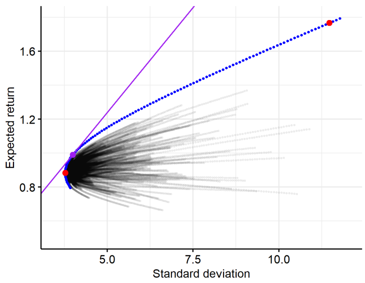


Figure 4: Efficient frontier full sample (blue), estimated frontiers T = 100(black)

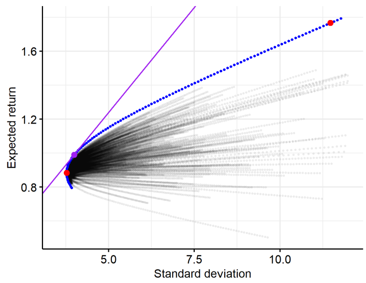
To investigate how the results depend on the sample size T, we repeat the procedure for $T \in \{100, 250, 500, 1.000\}$. For this purpose we define a function: $compute\_efficient\_frontier\_T$ $(res, mu, sigma, sample\_size, iterations)$ which takes as inputs the optimal efficient frontier, the true moments $mu = \mu$ and $sigma = \Sigma$, the sample size (T), and the number of iterations (j). res is the output of the $compute\_efficient\_frontier()$ function from question 3 (using μ and Σ) and is needed to plot the blue line. The function calculates the Sharpe ratio of the tangent portfolio for each simulation iteration, stores it into to an empty vector and additionally plots the efficient frontiers for every sample plus the original one in blue. We then call this function for each proposed sample size T while keeping the number of iterations fixed at 250 ($r_{rf} = 0$). Figure 5 shows how the simulated efficient frontiers converge to the theoretically optimal efficient frontier as T grows. There is still a lot of variation even across the simulation with 1000 observations but the trend is clear. The reason why the simulated frontiers deviate (and continue to deviate even for larger sample sizes) from the theoretically efficient frontier is that they rely on estimated sample moments. Estimators for the sample moments require larger sample sizes to converge to the true moments (as T asymptotically goes to infinity).

Question 8

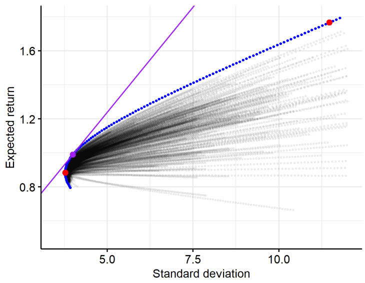
As explained in Exercise 7, the function we use to generate Figure 5 for different values of T appends



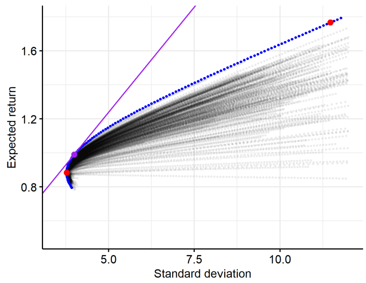
(a) T = 100



(b) = 250



(c) T = 500



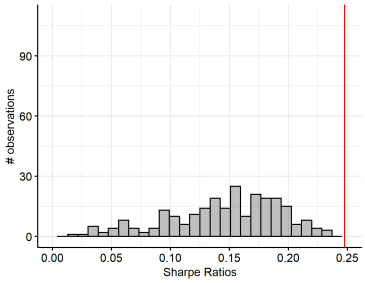
(d) T = 1.000

Figure 5: Efficient frontier using full sample (blue) compared with estimated frontiers from 250 samples of different lengths T. Convergence to theoretically efficient frontier as T grows.

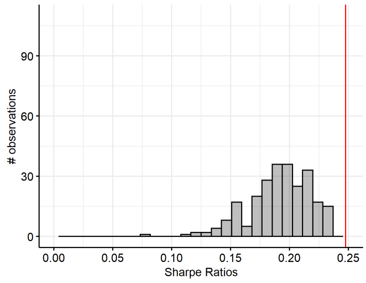
the computed Sharpe ratio for each iteration (j) into an empty vector. The resulting vector contains 250 rows and one column. This Sharpe ratio obtains the weights of the tangent portfolio using the sample moments and evaluates them at the true moments:
$SR^j_{tgc} := w^j_{tgc}{}'(\mu - r_{rf})/\sqrt{w^j_{tgc}{}'\Sigma w^j_{tgc}}$. Our code already performs what was asked in question 8 and that is explained in question 7. At the end of question 7, we create a data frame that combines the Sharpe ratio vectors for each sample size chosen (called *sharp_ratio_all*).

Question 9

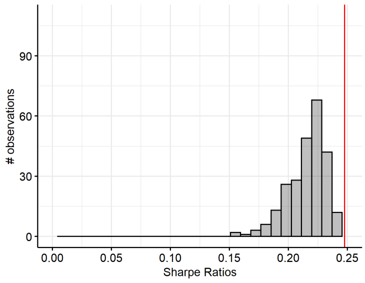
For each value of T, Figure 6 displays the histogram for the resulting vector of Sharpe ratios of each sample and compares it to the true Sharpe ratio of the tangent portfolio from question 4 (red line). The histograms show how the distribution of the Sharpe ratios converge towards the theoretically optimal Sharpe ratio from Question 4 as $T \to \infty$. There are two key insights from figure 6. First, the Sharpe ratio of the tangent portfolio is the highest Sharpe ratio you can obtain. In this case, it gives the highest possible Capital Market Line slope given a risk free rate of 0. If we would find higher Sharpe ratios for portfolios on the frontiers from our simulations, that would contradict the assumption of optimality for the tangent portfolio. Thus it makes sense that the true Sharpe ratio is always the highest observation from the distribution of Sharpe ratios. Second, using the plug-in estimated moments performs poorly for small sample sizes as they rely on asymptotic convergence towards the true moments as $T \to \infty$. This again is in line with Brand (2010) and shows the lack of precision of plug-in estimates in real life applications. Two potential alternative strategies to overcome the weakness of the plug-in estimates could be to calculate optimal portfolio weights under a short selling restriction or to simply implement the naive portfolio with weights equal to $1/N$ in each period. Inspired by the group of Niels Eriksen and Rasmus Juel we calculate the Sharpe ratios for a no-short-selling portfolio with $T = 100$. This is done by using the *solve.QP()* optimizer from the *quadprog* package. The resulting distribution of Sharpe ratios can be seen in figure 7 in sub-figure *(b)*. Figure 7 compares the Sharpe ratio distribution for $T = 100$



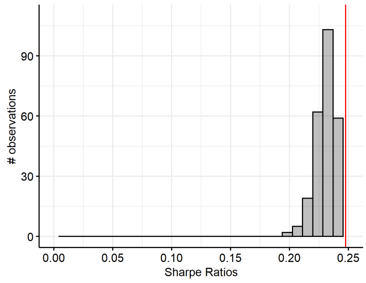
(a) T = 100



(b) = 250

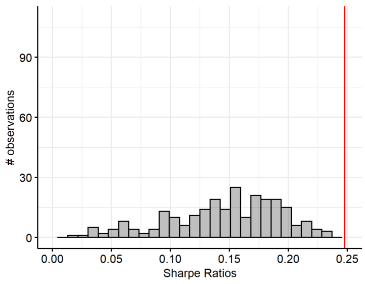


(c) T = 500

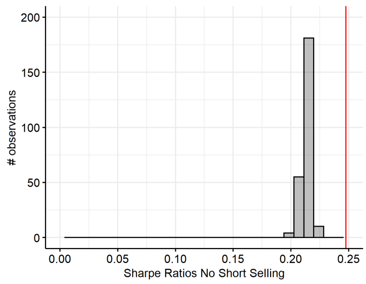


(d) T = 1.000

Figure 6: Histograms of Sharpe ratios for different T compared with the true one in red



(a) T = 100



(b) T = 100

Figure 7: Comparing Sharpe ratios for T = 100 with plug-in estimates (left side) vs. no-short-selling estimates. True Sharpe ratio in red

using the plug-in estimates on sub-figure (a) with the no-short-selling constrained results in (b). It is clear that restricting the investor to non-negative portfolio weights improves the performance - even for relatively small sample sizes as the mean in closer to the true mean (red line) and the variance is a lot smaller. As a second alternative, the naive portfolio also performs better than the plug-in estimates for small sample sizes with a estimated average Sharpe ratio of 0.19. As the weights are fixed to $1/N$ the naive portfolio's performance is independent of the sample size and does not converge. Thus, for higher sample sizes (like $T = 1000$) the naive portfolio performs worse than the plug-in estimates (this can be seen by comparing table 6 with subfigure (d) in figure 6).

	mu naive	sigma naive	Sharpe
	1	5.2	0.19

Table 6: Naive portfolio results