Portfolio Optimisation: Hidden Regularisers in In-built Optimisers

By Lewis Mead

Supervised by Imre Kondor & Fabio Caccioli

LONDON MATHEMATICAL LABORATORY

Simplified Problem

- Simplified version of the problem:
  Let $w = (w_1, \dots, w_N)$ be the vector of portfolio weights and σ the covariance matrix of portfolio returns.

  $$\min \left(\sum_{i,j} w_i \sigma_{i,j} w_j\right)$$

  $$s.t \sum_{i=1}^{N} w_i = 1$$

This is a quadratic programming problem.

- The solution to this problem is easy to compute by Lagrange multipliers

$$w_i^* = \frac{\sum_{j=1}^{N} \sigma_{i,j}^{-1}}{\sum_{j,k} \sigma_{j,k}^{-1}}$$

LONDON MATHEMATICAL LABORATORY

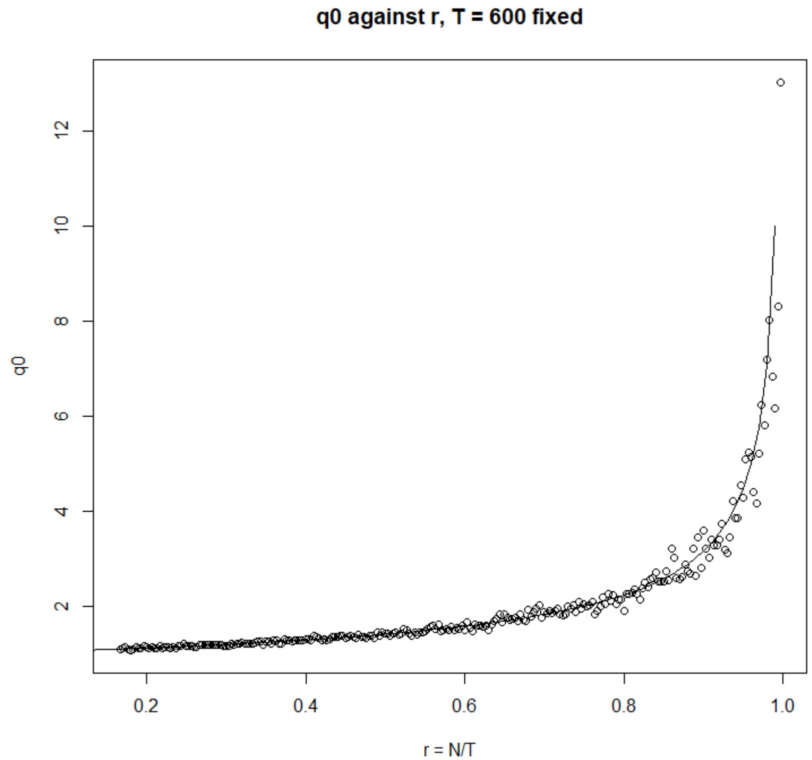
Measuring Risk In Noisy Estimates

- In reality we do not know the true covariance matrix of the returns so we have to estimate this based on sample data.
- However in toy examples we understand the true covariance matrix of the returns and we have a natural measure of the true risk of our predictions, q_0, dependent on $\sigma^{(true)}$ and $\sigma^{(est)}$.
- **Remark.** As $T \to \infty, \sigma^{(est)} \to \sigma^{(true)}$.
- **Remark.** If $N > \mathrm{T}$ then $\sigma^{(est)}$ is singular with probability 1 in which case the optimisation problem has no solution. So $N/T = 1$ is a hard cut off point – beyond here you would be dividing by 0.
- This remark tells us that the ratio $r = N/T$ plays an important role in the computation of $\sigma^{(est)}$ and hence of q_0.
- **Lemma.** For $r = N/T$ fixed and $N, T \to \infty$ we have the relation $\mathrm{q}_0 = \frac{1}{\sqrt{1-r}}$.

LONDON MATHEMATICAL LABORATORY

Experimental Framework

For simplicity I took $\sigma^{(true)} = I_N$, kept T (the length of time series) fixed and let X be populated by i.i.d $N(0,1)$ samples.



The calculated q_0 values fit the predicted analytic curve well. As we approach the point $r = 1$ the true risk diverges, and beyond this point we can make no meaningful conclusions.

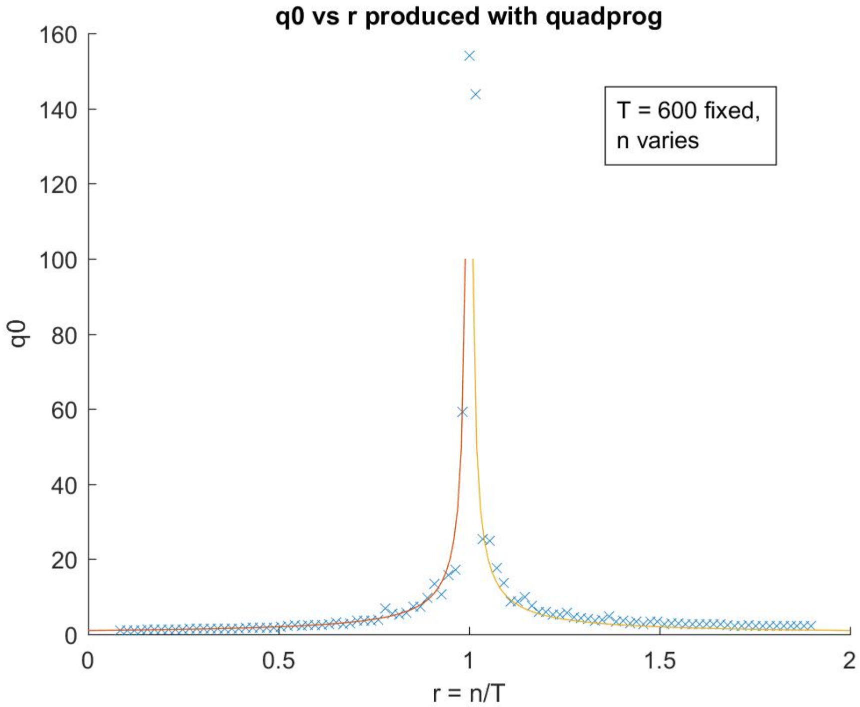
LONDON MATHEMATICAL LABORATORY

In-Built Solvers

- In reality the task is much tougher:
  - You may wish to optimise a more complex system.
  - You will not know the covariance matrix of returns.
  - Your length of time series is very limited.

In-Built Solvers

Often you will need to use some in-built function to solve this optimisation problem for you. For example quadprog in MATLAB.

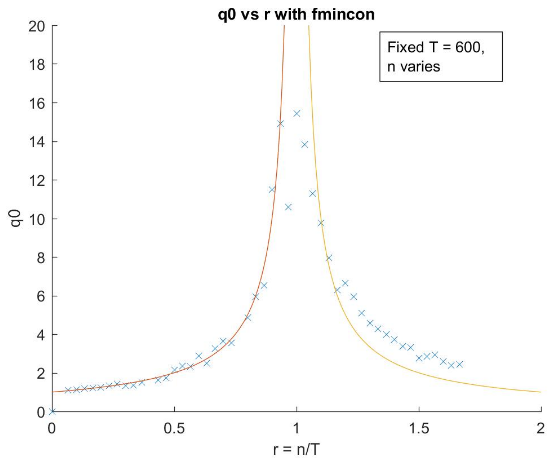


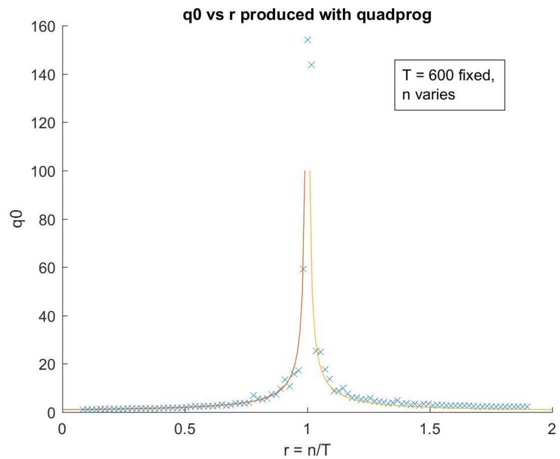
Which seems to suggest we can get meaningful conclusions beyond $r = 1$. Moreover it suggests that the further we go beyond this point the lower the true risk.

In-Built Solvers

I investigated a wide range of such solvers in a variety of languages and found that this was a problem pertinent to many.

MATLAB

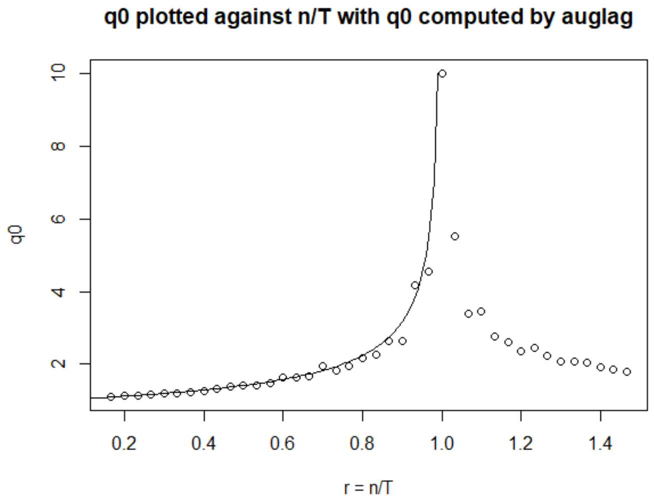


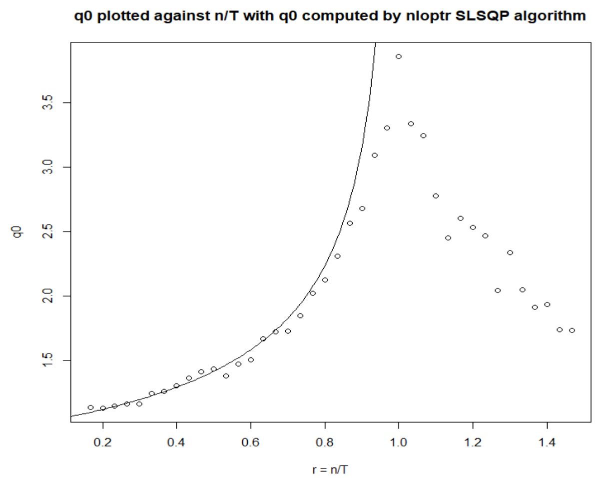


LONDON MATHEMATICAL LABORATORY

In-Built Solvers

R

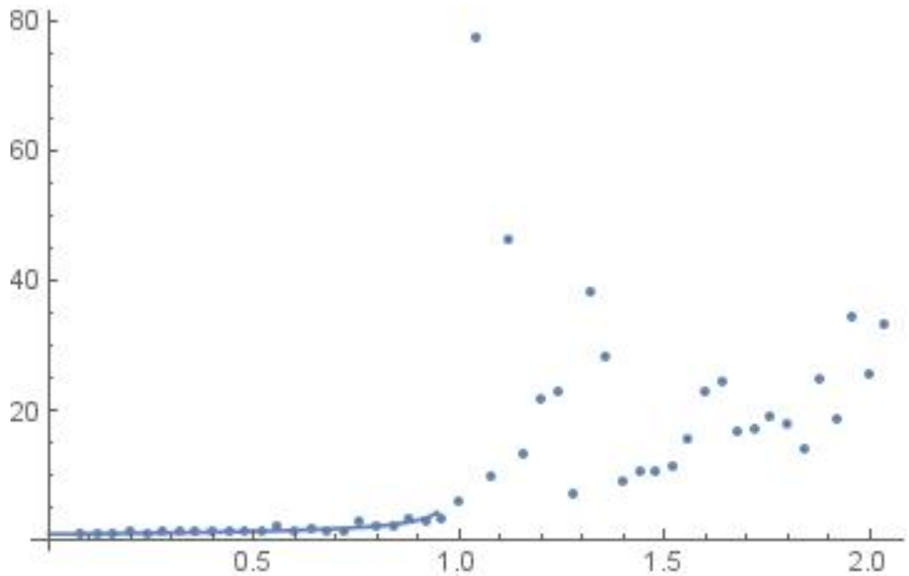




LONDON MATHEMATICAL LABORATORY

In-Built Solvers

Mathematica



Why?

- It's clear than the in-built solvers are doing *something*. Possibly to make the problem simpler to solve. Possibly by inherent problems with the algorithms.
- Not all in-built solvers exhibit this behaviour so this is not a universal issue and there seems to be an approach that can avoid this – at least sometimes.

LONDON MATHEMATICAL LABORATORY

How?

- What the algorithms are doing is not clear. In fact often times the source code is protected or obfuscated (MATLAB).
- There are two main possibilities
  - Regularisation
  - Moore-Penrose pseudoinverse

LONDON MATHEMATICAL LABORATORY

Regularisation

- Regularisation attempts to solve overfitting data in statistical models. When you measure too many variables the model may become too sensitive to new input data.
- Regularisation introduces bias to the system but you hope the trade off is worthwhile.
- This is usually done by adding some multiple of the norm of the parameters to your model. In our case this corresponds to solving:

$$\min \sum_{i,j} w_i \sigma_{i,j} w_j + \eta||w||$$

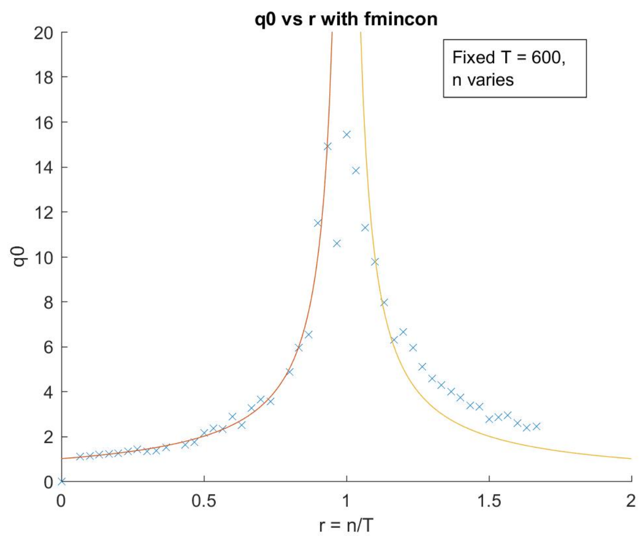
$$s.t \sum_{i=1}^{N} w_i = 1$$

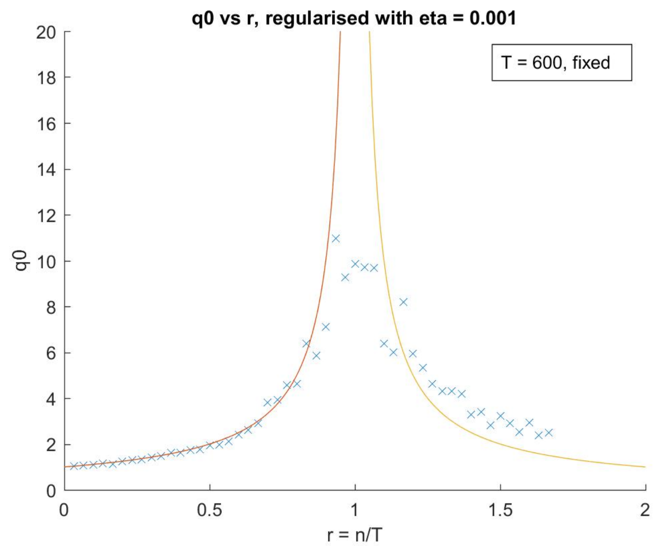
where η is some fixed chosen value.

LONDON MATHEMATICAL LABORATORY

Regularisation

fmincon and the true solution of the regularised problem display considerable similarities. The peak at $r = 1$ is flattened and they tail off similarly.



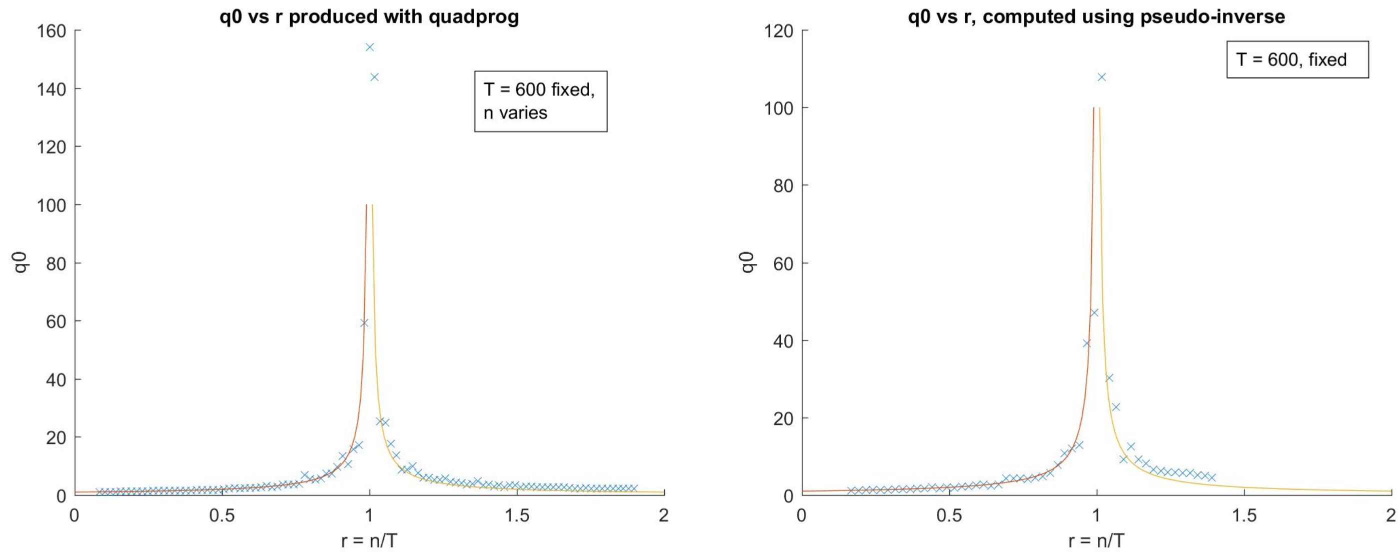


LONDON MATHEMATICAL LABORATORY

Pseudoinverse

- The Moore-Penrose pseudoinverse, A^+, of a matrix A is a generalisation of the inverse matrix. It allows a non-square or singular matrix to have some notion of an inverse.
- $A^+ = A^{-1}$ when A^{-1} exists.

quadprog and the closed solution provided by using the pseudoinverse in place of any inverses have similarities. Both shoot off to infinity as they approach $r = 1$ and tail off in a similar manner.



LONDON MATHEMATICAL LABORATORY

Broader Context

- The use of statistical and mathematical tools you do not understand is very dangerous. This is something that can also have a big impact upon the reproducibility of an experiment.
- Dimension matters. Perhaps more so than any underlying distribution as often phase transitions are universal (independent of sample distribution) but dependent on dimension.
- When non-statisticians are using statistical tools they do not fully understand there's likely to be issues of incorrect inference. This can only be compounded when such tools may provide incorrect solutions.

LONDON MATHEMATICAL LABORATORY