

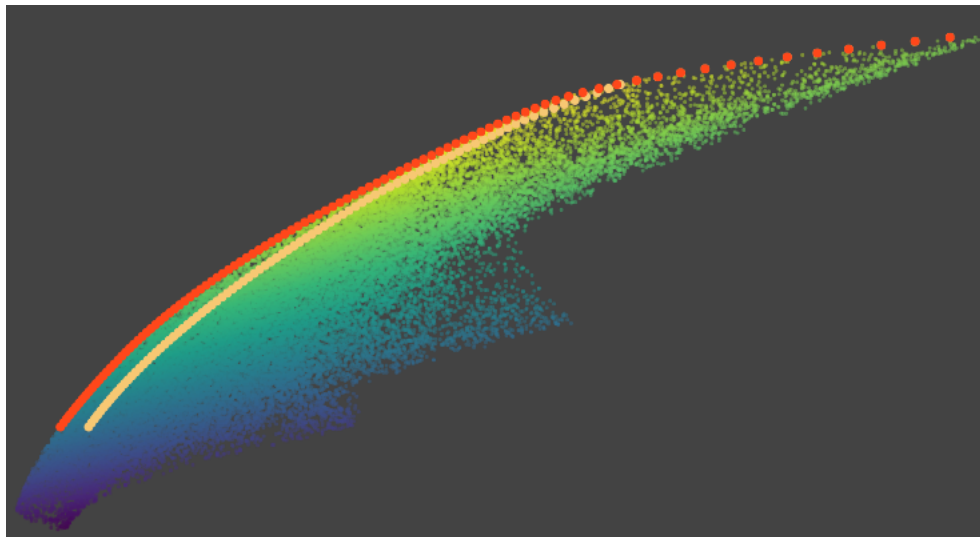
[More](#)

Yet Another Math Programming Consultant

I am a full-time consultant and provide services related to the design, implementation and deployment of mathematical programming, optimization and data-science applications. I also about projects I am doing, but there are many technical notes I'd like to share. Not in the least so I have an easy way to search and find them again myself. You can reach me at erwin@yetanothermathprogrammingconsultant.com

Thursday, April 19, 2018

Covariance Matrix not positive definite in portfolio models



When solving mean-variance portfolio optimization models, the majority of solvers will require the variance-covariance matrix is positive definite. In theory, any covariance matrix is positive semi-definite. But sometimes in practice it is not, something I saw today.

Theory

The sample covariance matrix is

$$q_{i,j} = \frac{1}{T-1} \sum_t (r_{t,i} - \mu_i)(r_{t,j} - \mu_j)$$

where T are the number of time periods, and $r_{t,i} - \mu_i$ are the mean adjusted returns. This matrix is obviously symmetric. To show it is positive semi-definite, we can prove $x^T Q x \geq 0$ for any x . We have

$$\begin{aligned} z &= x^T Q x \\ &= \sum_{i,j} x_i x_j q_{i,j} \\ &= \sum_{i,j} x_i x_j \left[\frac{1}{T-1} \sum_t (r_{t,i} - \mu_i)(r_{t,j} - \mu_j) \right] \\ &= \frac{1}{T-1} \sum_t \left(\sum_i (r_{t,i} - \mu_i) x_i \right) \left(\sum_j (r_{t,j} - \mu_j) x_j \right) \\ &= \frac{1}{T-1} \sum_t w_t^2 \geq 0 \end{aligned}$$

where

$$w_t = \sum_i (r_{t,i} - \mu_i) x_i = \sum_i \tilde{r}_{i,t} x_i$$

From this, we know the covariance matrix is in theory positive semi-definite. In practice things are more complicated.

Practice

In practice covariance matrices can become indefinite. For instance, if the number of instruments n is larger than the number of observations T we high probability we will have some slightly negative eigenvalues for Q . In the case these eigen values become more negative, you can expect to see messages like:

```
*** Objective Q not PSD (diagonal adjustment of 1.0e-05 would be required)

The quadratic coefficient matrix in the objective is not positive semidefinite as expected for a minimization problem.

The quadratic coefficient matrix in the objective is not PSD.
```

Sometimes the solver can fix things up a bit:

```
Warning: diagonal perturbation of 6.6e-06 required to create PSD Q.

Warning: diagonal adjustment of 4.0e-08 performed to make Q PSD
```

There are different fixes suggested:

- If $T < n$ the matrix is rank-deficient we can expect some issues. With this property, some of eigenvalues will be zero (i.e. the matrix Q is not strictly positive definite but rather positive semi-definite). It does not take much – just some small inaccuracies in floating point point computations – to make Q indefinite. A standard fix is: add more data (or reduce the number of instruments n).
- Fix Q by adding small numbers to the diagonal. You can see from the messages, some solvers try this automatically.
- Fix Q by performing an eigen decomposition, replace negative eigenvalues and reassemble the Q matrix [4].
- Check the data. Sometimes outliers caused by typo's, transcription errors, dropped decimal points etc., can make things worse.
- Use some advanced method to get better covariances, e.g. shrinkage [1]

An alternative model

A different approach is to use a different optimization model, inspired by the piece of theory above. The standard mean variance portfolio optimization problem looks like

$$\begin{aligned} \min \quad & x^T Q x \\ \text{subject to} \quad & \sum_i x_i = 1 \\ & \sum_i \mu_i x_i \geq R \\ & x_i \geq 0 \end{aligned}$$

Using the original mean adjusted returns $\tilde{r}_{t,i} = r_{t,i} - \mu_i$, we can develop an alternative formulation:

$$\begin{aligned} \min \quad & \sum_t w_t^2 \\ \text{subject to} \quad & w_t = \sum_i \tilde{r}_{t,i} x_i \\ & \sum_i x_i = 1 \\ & \sum_i \mu_i x_i \geq R \\ & x_i \geq 0 \end{aligned}$$

Now we have a totally benign diagonal Q matrix. With this formulation, you should never see any messages complaining about Q not being positive

definite.

If the model has $T \ll n$ this model has another advantage: much less data is moved around. The original Q matrix is $(n \times n)$ which, in this case, is larger than the $(T \times n)$ matrix $\tilde{r}_{t,i}$.

The reformulated model should provide the same solution in normal cases (where the calculated Q matrix is numerically positive semi-definite). The objective is different as I dropped the $1/(T-1)$ factor, but the optimal values for the x_i variable should be identical (up to some tolerance).

Related: least squares

This reformulation resembles how we can work around convexity issues when dealing with least squares objectives

$$\min \|Ax - b\|^2$$

This model is also known to occasionally produce errors relating to the quadratic coefficient matrix not being positive semi-definite. In [2] this is solved with the following reformulation:

$$\begin{aligned} \min \sum_i v_i^2 \\ v = Ax - b \end{aligned}$$

This reformulated model has more variables and equations, but the "easy" quadratic objective can make the difference between being able to solve problem or not. There can also be performance advantages.

References

1. Ledoit, Olivier and Wolf, Michael, *Honey, I Shrunk the Sample Covariance Matrix* (June 2003). UPF Economics and Business Working Paper No. <http://dx.doi.org/10.2139/ssrn.433840>
2. Least Squares as QP: convexity issues, <http://yetanothermathprogrammingconsultant.blogspot.com/2018/01/least-squares-as-qp-convexity-issues.html>
3. The picture is from <https://moderndata.plot.ly/portfolio-optimization-using-r-and-plotly/>
4. Eigen decomposition example. *SDP Model: Imputing a covariance matrix*, <https://yetanothermathprogrammingconsultant.blogspot.com/2019/11/sdp-model-imputing-covariance-matrix.html>

Posted by [Erwin Kalvelagen](#) at 11:39PM



No comments:

Post a Comment



Enter Comment

[Newer Post](#)

[Home](#)

[Older Post](#)

Subscribe to: [Post Comments \(Atom\)](#)

Awesome Inc. theme. Powered by [Blogger](#).