

|  |  |
| --- | --- |
| Average Daily Log Returns: | 0.00567383199315 |
| Standard Dev. Daily Log Returns: | **0.0126189200237** |
| Annualized Sharpe Ratio | **6.9963789661** |
| Skewness | **5.58448554934** |
| Kurtosis | **89.0234959108** |
| Max Drawdown Duration (days) | **6** |
| Max Drawdown Loss (% from peak) | **-4.20960993532%** |
| Equal-Weight Long Portfolio Correlation | **0.0121560966988** |

Our performance using this strategy is significantly improved from the strategy in Part 1. The annualized Sharpe ratio of 6.99 that we achieved is quite high – however, it’s likely our parameters have overfit to the data (more on this below). Nonetheless, the consistent upward trend in our cumulative returns is an indicator that these parameters work well throughout the data history and therefore that our results would be reproducible in a test sample. The maximum drawdown is solid and not concerning. Our portfolio returns are right-skewed and we have high kurtosis, but this is not really an issue since the mean returns are high.

This strategy is based on the idea of mean reversion for stocks, combined with other metrics such as volatility and volume. The parameter values that we use are as follows, for {a1, … , a12} respectively:

[-0.118, -3.057, 3.798, -2.059, 0.391, 3.857, -4.329, -3.933, -0.468, -2.175, 2.487, 2.485]

We obtained these parameters using a Nelder-Mead quasi-Newton optimization method for gradient descent. The main challenge was finding an appropriate starting point, as these methods are sensitive to parameter initialization, and twelve-parameter models are reasonably complex.

To determine a good initialization point, we first ran a simplified model with 3 parameters (a, b, c), defined below:

1. The average of ROO(t), RCC(t-1), ROC(t-1) (to replace a1 through a4)
2. The first parameter multiplied by trading volume differences (to replace a5 through a8)
3. The first parameter multiplied by volatility differences (to replace a9 through a12)

Our intuition was that ROO, RCO, ROC, and RCO are all indicators of very recent performance, and draw statistics from overlapping time frames. Therefore, they should be highly correlated, and we can combine them into one variable in a simplified model.

Optimization of this three-parameter model was fairly quick, so we were able to run it many times. Upon examining our results, we noticed that about two-thirds of them converged to the same Sharpe ratio

(talk about why we removed RCO?)

(talk about patterns we noticed in optimal 3 parameter results?)

We then performed a Nelder-Mead quasi-Newton optimization method on these 3 parameters (setting initial values as random uniform between -1 and 1 multiple times) to see a pattern of convergence towards the 3 parameters that maximize Sharpe Ratio. After obtaining these values, we plugged them into the full 12-parameter model as the initial values:

{a,a,a,a,b,b,b,b,c,c,c,c}

Running the same optimization method on these parameters then allowed us to solve for the optimal 12 unique parameters defined above to maximize our Sharpe Ratio. Note that using the average of ROO(t), RCC(t-1) and ROC(t-1) to replace each of the first 4 parameters was determined to be a valid process due when we tested for the high correlation between them.