Portfolio replication with sparse regression

Akshay Kothkari, Albert Lai and Jason Morton *

December 12, 2008

Suppose an investor (such as a hedge fund or fund-of-fund) holds a secret portfolio of assets, which may change over time. Suppose further that the investor makes public the overall returns on this portfolio on a regular basis, as is the case wth hedge funds, funds of hedge funds, and indexes of hedge funds. There are many reasons why we might want to replicate the performance of this portfolio. One is to gain exposure to the same sources of risk and return while avoiding the high fees and liquidity restrictions typically imposed by these investment vehicles. Alternatively, we may already be invested and wish to hedge the risk of a significant downturn in the fund or fund of funds. As we cannot know its holdings, we are forced to estimate them statistically.

Most practical strategies for replication available today make use of a rolling regression against fewer than ten factors. These factors are selected in advance by hand. We instead investigate the use of L_1-regularized regression to choose the regressors from a large universe of securities automatically. We consider four simple types of strategies to replicate, one static and three dynamic, all based on one or more true weight vectors h^P.

1. The *buy-and-hold* portfolio. Given a total initial portfolio value of V dollars and initial weight vector h^P, this strategy buys Vh_i^P of asset i at price S_i and simply holds the assets for the duration. A good replication strategy for a buy-and-hold portfolio should quickly converge to the true weights n^P in terms of the *number of shares* Vh_I^P/S_i initially bought of each asset and perform minimal or no further reweightings or trades once the true portfolio is found. The weights in terms of value of this strategy will change over time.

2. The *fixed constantly rebalanced* portfolio. Given an initial weight vector h^P, this stratgy rebalances every period to hold asset i in proportion h_i^P of the total value of the portfolio, selling and buying as necessary. A replication strategy for a fixed constantly rebalanced strategy will try to find h^P and thereby match the rebalancing trades of the true strategy each period.

3. The *fixed periodically rebalanced* portfolio, a hybrid of the first two. This strategy, common in practice (e.g. it is used by many ETFs and mutual funds) rebalances back to the initial weights h^P every n periods. A replication strategy must implicitly find both h^P and n.

4. The *fixed periodically rebalanced with change points* portfolio. Here the target weight vector h^P changes at several points over the course of the observed period, where new securities may be added and existing ones dropped at the change point.

*Not in the class, but helped out and provided advice.

We use real return series from the S&P 500, and synthetic portfolios for types 1-3. For type 4, we look at the performance of our method in replicating a few ETFs which follow this type of strategy.

1 Algorithms

We experimented with a variety of algorithms for rolling regularized regression using lasso regressors, and clipped stochastic gradient descent. We report on the most successful variant. This algorithm uses a rolling lasso estimator on the residuals to select which portfolio weights to consider updating, and an unconstrained regression on the winners to fix the final weight changes themselves.

The winners were deemed to be assets with coefficients greater than one standard deviation from than the mean, which is very close to zero due to the sparsity. Because we would like to have long only portfolios for simplicity, all coefficients are constrained to be positive.

Input: Return series of investible assets (r_t) and benchmark asset (b_t). A window w, relative sparsity parameter α, and unconstrained regression inclusion threshold ϵ.
Output: Vector of (continuosly rebalanced) portfolio weights β_t at each time point after the initial w.

Initialize: Parameter vector $\beta_{w-1} = 0$
for each time point t:

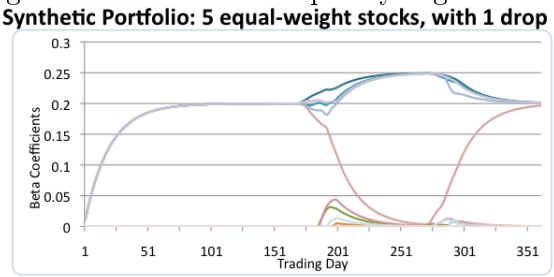
- Compute residuals $e_{t-w,\ldots,t} = r_{t-w,\ldots,t}\beta_{t-1} - b_{t-w,\ldots,t}$, where $r_{t-w,\ldots,t}$ is the $w \times p$ matrix of returns on the p investible assets during w periods in the rolling window
- Let ξ be the L_1-regularized coefficients of $e_{t-w,\ldots,t}$ regressed on $r_{t-w,\ldots,t}$ with sparsity parameter α.
- Set $\gamma = \beta_{t-1} + \xi$
- Clip all elements of $\gamma < 0$ to 0
- Let $I = \{i \in \{1, \ldots, p\} \mid \gamma_i > \epsilon\}$, the set of assets appearing with nonzero coefficients in γ
- Let η be the coefficients of the unconstrained regression of $e_{t-w,\ldots,t}$ on $r_{(t-w,\ldots,t),I}$, the assets in I.
- Set $\beta_t = \beta_{t-1} + \eta$

A stochastic version of this algorithm was also shown to be effective; however it did not converge as quickly and had noisier results. The cost of extra computation in the rolling window algorithm is not as serious an that of requiring more time. That is, in a real world situation, one would rather spend more computing power than wait more days for the regression to converge, so the rolling window algorithm was chosen.

2 Experiments

The first goal was to replicate a synthetic portfolio of five equally weighted, constantly rebalanced assets drawn from the S&P 500. Then using Least Angle Regression as the lasso regressor in our algorithm, determine if it is able to pick out the correct stocks. To see the algorithm's response to a varying portfolio, it was run against synthetic portfolios where the assets change weights, are added, and get dropped. It was seen that the longer the window, the smoother the results and fewer false positives. However, longer windows also resulted in slower convergence and becuase time is valuable, slow convergence rates are very detrimental to algorithm performance. Thus a rolling window of 30 days was used.

Figure 1: Coefficients Output by Algorithm on Synthetic Portfolio



As seen in Figure 1, the change point was at day 180 where an asset was dropped and then picked up again at day 280. Here, the algorithm selects some extra coefficients before damping them back to zero. Raising the shrink parameter results in faster convergenge, but creates more noise and instability. In these experiments, a shrink of 0.05 was used.

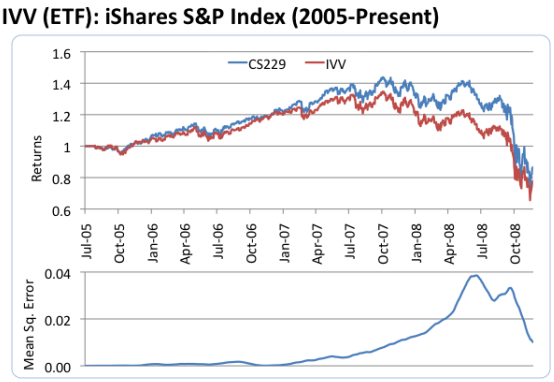
Once it was verified that the algorithm can identify the correct assets in a toy problem, the algorithm was run on real-world data, which we chose to be several Exchange Traded Funds. The funds were regressed against the S&P 500 from June 2005 to November 2008. The method to judge accuracy was to take one dollar initially invested in the ETF, and one dollar invested in the algorithm's CS229 Fund both at time t_0 and compare the subsequent returns.

Because it is not desirable to rebalance the portfolio daily due to the costs of trading, the CS229 Fund samples the output of the algorithm every four weeks and rebalances the portfolio to the new coefficients. The cost of trading was not factored directly into the algorithm and should be involved in future work, but monthly rebalancing suited the purposes of the experiment. Additionally, one wishes to avoid having to trade a very large number of stocks of miniscule weight. Thus only the assets with coefficients greater than 0.02 were included in the portfolio. These coefficients do not necessarily sum to one, so this subset of coefficients is normalized to prevent gaining or losing extra leverage. These coefficients are the portfolio of the CS229 Fund.

Recent events in the economy provided the unlucky opportunity to run the algorithm on bullish as well as bearish markets. As shown in the Figure 2, the CS229 Fund tracks fairly well for the two ETF's. The error is most apparent when the ETF is experiencing a high rate of change, in particular mid-2008 when the markets began falling.

The standard error $\frac{1}{w}(\sum_{t-w}^{t}(actual - predicted)^2)^{1/2}$ for a window is also plotted in Figure 2. The error rates are well below 0.01 until the sudden recent downturn beginning

Figure 2: Plots for Replicating Two ETF's



in 2008 where error rates climbed to 0.05. However, the error slowly shrank in recent months. Figures 3 and 4 indicate that greater number of coefficients, or assets present in the CS229 Fund, are associated with greater errors. Determining whether it is a causal relationship is worth investigating. It could be the case that a large error causes the algorithm to select more assets in an attempt to make up for the error, or having too many assets results in overfitting.

Figure 3: Error vs. number of assets when replicating the IVV iShares ETF

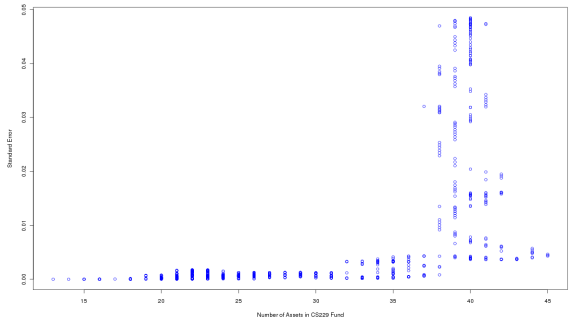
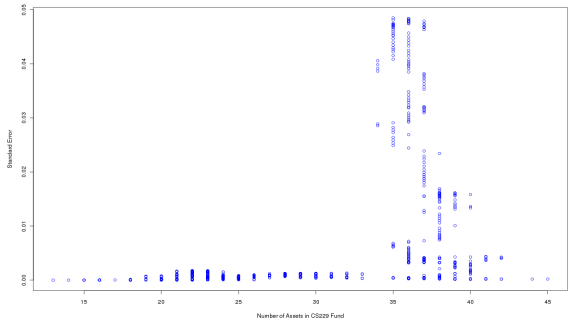


Figure 4: Error vs. number of assets when replicating the IVE iShares ETF



3 Conclusion

As shown, the L1 regularized regression with rolling windows works well with ETFs. The LARS regression was very good at picking out the proper assets out of a large universe. However to obtain the coefficients, drastic improvement were obtained by regressing the target again against the reduced universe consisting only of the assets with larger coefficients. In bullish as well as bearish years the tracking error was fairly small, however sharp declines yielded large errrors. With monthly rebalancing, the CS 229 Fund replicates ETF returns with a maximum mean square error of 0.05 and shows much promise for further investigation.

4 Future Work

We would like to test the algorithm with existing hedge funds using the returns available from TASS database. Also, we would like to model the costs associated with rebalancing to find the optimal rebalancing frequency, a metric that we overlooked in this project.

References

[1] B. Efron, T. Hastie, I Johnstone and R. Tibshirani, Least Angle Regression. Annals of Statistics, 32(2), 2004, pp. 407-499.

[2] A. Lo and J. Hasanhodzic, Can hedge-fund returns be replicated?: the linear case. Journal of investment management, 5(2), 2007, pp. 545.

[3] A. Ng. Feature selection, L1 vs. L2 regularization, and rotational invariance. In Proceedings of the Twenty-first International Conference on Machine Learning, 2004

[4] W. Fung and D. Hsieh, Empirical characteristics of dynamic trading strategies: the case of hedge funds. *Review of Financial Studies,* **10**, 275-302.