Penalized Mean-Variance Portfolio Optimization using Weighted Elastic Net

516 Final Project

Jason Bohne and Jarryd Scully

Applied Mathematics and Statistics Department
Stony Brook University

December 6, 2022

Overview

1. Model Formulation
2. Parameter Estimation
3. Split Bregman Algorithm
4. Paper Results
5. Original Results
6. Conclusion

General Motivation

- Markowitz introduced mean-variance selection methods for portfolio construction
- Common issues faced are parameter estimation and robust inference methods
- Markowitz's approach can lead to overfitting and poor *out of sample* performance
- Regularizing the objective in the inference procedure can lead to better extrapolation

Objective Function

- Mean-variance objective for single period portfolio optimization

$$min_w \; w^T \Gamma w - w^T \mu$$

- Numerically solved as a convex quadratic programming problem
- Alternate formulation ignoring mean of asset returns simplifies objective

$$min_w \; w^T \Gamma w \text{ s.t. } \sum_{i=1}^{N} w_i = 1$$

- Minimum variance criterion empirically results in better out-of-sample performance

Regularization

- Elastic net regularization applies a L1 and L2 penalty to objective function
- Support for weighted elastic net regularization of adaptive penalty weights

$$\lambda_1||w||_1 = \sum_{k=1}^{N} \beta_k |w_k|$$
$$\lambda_2||w||_2 = \sum_{k=1}^{N} \alpha_k w_k^2$$

- Incorporating into the minimum variance objective; resulting objective is

$$min_w \ w^T \Gamma w - w^T \mu + ||\beta w||_1 + ||\alpha w||_2$$

Parameter Estimation

- James-Stein estimator used to shrink estimate of the mean vector

$$\hat{\mu} = (1 - \rho)\hat{\mu_s} + \rho\eta\bar{1}$$

- Coefficients defined as

$$\eta = \max\left(\frac{1}{N}\sum_{i=1}^{N} \mu_{S,i}, 0.0004\right) \text{ and } \rho = \min\left(1, \frac{N-2}{T_{train}(\mu_S - \eta\bar{1})^T \hat{\Gamma}^{-1}(\mu_s - \eta\bar{1})}\right)$$

- Estimated covariance matrix defined using

$$\hat{\Gamma} = \rho_1 \hat{\Gamma}_s + \rho_2 I$$

- Parameter ρ is found by minimizing the quadratic loss

$$E[||\hat{\Sigma} - \Sigma||^2]$$

Split Bregman Algorithm

- Solves L1 regularization problems such as absolute shrinkage and total variation
- Reformulates objective into distinct optimization problems solved iteratively
- First problem is unconstrained quadratic problem incorporating L2 penalty
- Second problem is constrained linear problem incorporating L1 penalty
- Respectively solved via closed form or gradient descent and shrinkage operator

Split Bregman Algorithm

Algorithm 1 Split Bregman Algorithm

Require: $(\hat{R}, \hat{\mu}, \lambda), k = 1, b^k = 0, w^k = 0, d^k = 0$

while $||w^{k+1} - w^k||_2 > \epsilon$ **do**

$w^{k+1} = \arg\min_w w^T \hat{R} w - w^T \hat{\mu} + \frac{\lambda}{2}||d^k - \beta w^k - b^k||_2^2$

$d^{k+1} = \arg\min_d \frac{\lambda}{2}||d - \beta w^{k+1} - b^k||_2^2 + ||d||_1$

$b^{k+1} = b^k + \beta w^{k+1} - d^{k+1}$

$k+ = 1$

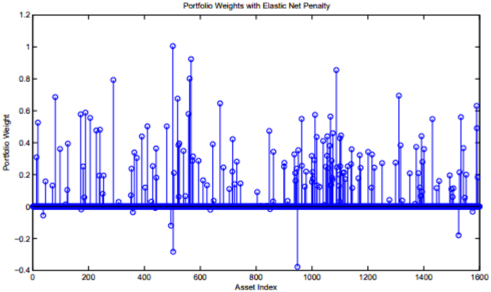
end while

Paper Experiments

- Performance of the elastic net penalty was tested by building a portfolio of stocks
- Daily returns of 630 U.S. stocks from January 1, 2001 to July 1, 2014
- Compute portfolios for ever 63 trading days, using the prior year as training data
- Evaluate results using the out of sample Sharpe Ratio

$$SR = \frac{(1/\tau)\sum_{i=1}^{\tau} w(t_i)^T r(t_i)}{((1/\tau)\sum_{i=1}^{\tau}(w(t_i)^T r(t_i) - (1/\tau)(\sum_{j=1}^{\tau}(w(t_j)^T r(t_j))))^{1/2}}$$

Paper Results



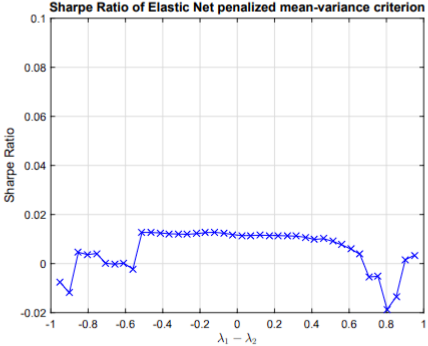


Figure: (Left) Portfolio weights (Right) Sharpe Ratio with changing penalizers.

Data Preprocessing

- Original dataset consists of S&P 500 universe between Jan. 2016 and Dec. 2019
- Daily close prices fetched from Yahoo Finance results in 1008 samples per asset
- Distribution of daily returns across portfolio universe is approximately normal
- Magnitude of singular values in sample covariance matrix supports regularization

Exploratory Data Analysis

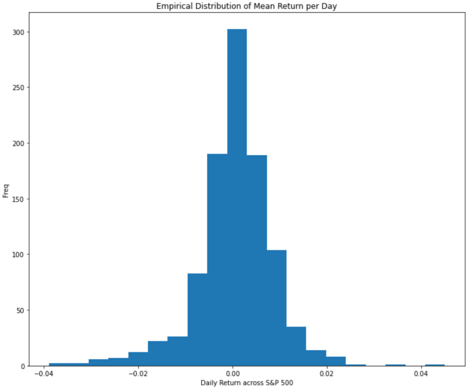


Figure: Empirical Distribution of Daily S&P Returns

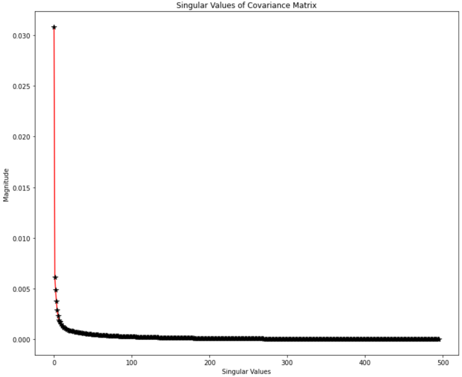


Figure: Magnitude of Covariance Singular Values

Numerical Results

- Split the dataset into training and testing set of ratio 50%-50%.
- Using regularized parameter estimates and Split-Bregman construct portfolios.
- Cross-validate hyperparameter λ_1 across grid of 12 potential values.
- Optimal portfolio associated with $\lambda_1 = 50$ with objective value of machine precision.
- Optimal portfolio has 489 components with allocation between $10e^{-5}$ and $9e^{-2}$.

Sample Portfolios

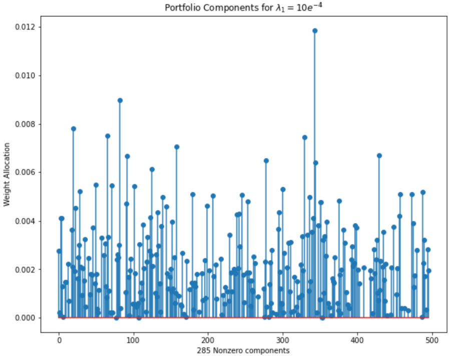


Figure: L1 Penalty $\lambda = 10e^{-4}$

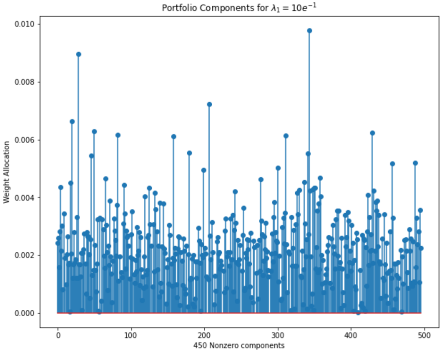


Figure: L1 Penalty $\lambda = 10e^{-1}$

Out of Sample Performance

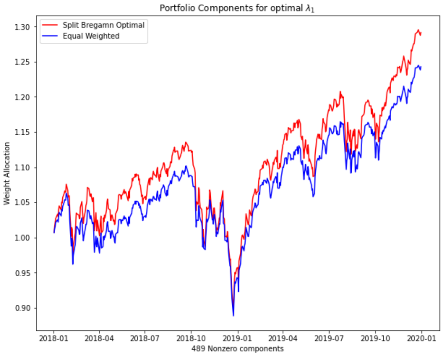


Figure: Optimal Elastic Net Penalized $\lambda = 5 * 10e^1$

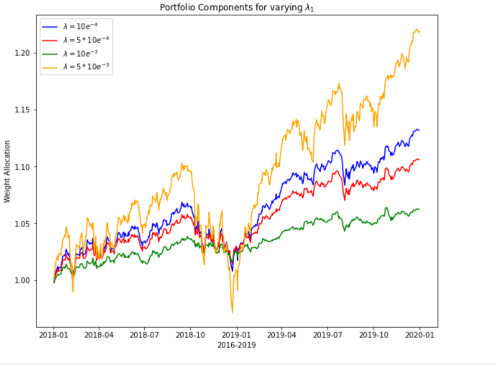


Figure: Various Penalization Values

Conclusion

- Classic mean-variance optimization overfits causing poor out-of-sample performance
- Elastic net regularization improves robustness by inducing sparse estimates
- Parameter estimates via James-Stein estimator and L2 regularized covariance matrix
- Reformulate optimization problem into two distinct objectives
- Using Split-Bregman to iteratively solve for the optimal vector w
- Hyperparameters specifying the penalty strength can be tuned via cross-validation
- Applications to historical U.S. equity data outperform an equal allocation

Code Repository

To access the data and scripts to reproduce examples, refer to the Github,
https://github.com/jasonbohne123/Split_Bregman_MPT

References

- Weighted Elastic Net Penalized Mean-Variance Portfolio Design and Computation
  Michael Ho, Zheng Sun, Jack Xin *SIAM Journal on Financial Mathematics* 2015
  https://doi.org/10.1137/15M1007872
- The Split Bregman Method for L1-Regularized Problems
  Goldstein, Tom and Osher, Stanley *SIAM Journal on Imaging Sciences* 2009
  https://doi.org/10.1137/080725891
- Bayes-Stein Estimation for Portfolio Analysis
  Jorion, Philippe Stanley *Cambridge University Press* 1986
  https://doi.org/10.2307/2331042
- A well-conditioned estimator for large-dimensional covariance matrices
  Olivier Ledoit and Michael Wolf *Journal of Multivariate Analysis* 2004
  https://doi.org/10.1016/S0047-259X(03)00096-4