

Common Task Framework

Jakob Jorsal Rasmussen

Copenhagen Business School, Department of Finance, Pension Research Center

1 Data Preparation

The empirical analysis is conducted using the three datasets provided by the competition platform: monthly stock characteristics (`ctff_chars`), the corresponding list of eligible features (`ctff_features`), and daily excess returns (`ctff_daily_ret`). All transformations and model inputs are constructed in a strictly time-consistent manner, ensuring that no information dated after portfolio formation month t is used in prediction or weight construction.

The first step in preparing the data consists of handling missing values in the characteristics. Because the modeling framework is cross-sectional, missing observations are imputed within each month rather than across time. Specifically, for each characteristic and month, missing values are replaced by the cross-sectional median of that characteristic in the same month. If all observations of a characteristic are missing in a given month, any remaining missing entries are set to zero.

After handling missing entries, all characteristics are standardized cross-sectionally using a rank-based transformation. Within each month, values are converted to percentile ranks and then linearly mapped to the interval $[-1, 1]$. This transformation reduces sensitivity to extreme values and ensures that all characteristics enter the model on a comparable scale. By relying on ranks rather than raw values, the procedure improves robustness and stabilizes the estimation.

2 Return Prediction

The code uses an elastic net to make a prediction at time t regarding which stocks will have the highest expected return at time $t + 1$. As the target of this exercise is to identify high return stocks, the model is used to predict demeaned excess returns. In other words, the model is focused on relative stock performance rather than absolute stock performance. By centering returns within each month, the regression is tasked with capturing cross-sectional dispersion rather than common shocks.

The elastic net combines Ridge and Lasso regressions. The advantage of a Lasso regression in this setting is that it creates a parsimonious model by discarding irrelevant or weakly predictive characteristics. The Ridge regression stabilizes estimation by shrinking coefficients towards zero. By shrinking coefficients and selecting only the most informative predictors, Elastic Net reduces estimation noise and (hopefully) improves out-of-sample stability, which is essential for portfolio construction.

Model estimation follows a rolling-window design. At each portfolio formation date t , only historical data strictly prior to t are used for training. A 60-month window (five years) of past observations is employed to estimate model coefficients. Hyperparameters are re-tuned at fixed intervals of 120 months (ten years), using a split of the preceding 60-month window into 48 months for training and 12 months for validation. Between retuning dates, the selected

hyperparameters are held fixed, while the model is re-estimated on the most recent rolling window of historical data.

The implementation uses ElasticNet from `sklearn.linear_model`. For each ten year period, it selects hyperparameters two hyperparameters. The hyperparameter α controls how strongly coefficients are shrunk towards zero. The algorithm selects between 10 possible alphas spaced logarithmically between 10^{-5} and 10. The hyperparameter L1 determines the weight of Lasso relative to Ridge. The algorithm selects 4 L1 values linearly separated between 0.2 and 0.8.

3 Portfolio Weighting

After having estimated demeaned excess returns for the next period, a long-only portfolio is constructed using the decile of stocks with the highest predicted return. Since the stock selection focused only on high (relative) expected returns, weights will be assigned to limit the variance of the portfolio. For each stock in the portfolio, historical daily excess returns are used to estimate the variance. Each stock is then assigned a weight in proportion to the inverse of its variance. Weights are then standardized such that they sum to 1. Under a zero correlation assumption, this weighting method would result in the global minimum-variance portfolio.