Project Class Asset Management

Revisiting Key Concepts & Group Project Kickoff

Tim Baumgartner

Welcome to the Final Lecture

- **Objective**: Recap key concepts from the course and discuss the group project.
- **Today's Agenda**:
  1. Review of Backtesting Fundamentals
  2. Statistical Learning in R
  3. Group Project: Task and Expectations
  4. Q&A

Backtesting Fundamentals

- **What is Backtesting?**
  - Simulating a trading strategy using historical data to evaluate performance.
  - Key components: In-sample (IS) tests and Out-of-sample (OOS) forecasts.
- **In-sample Modelling**:
  - Model: $r_{i,t+h} = a + bX_t + u_t$, where r is excess return, X is predictor(s).
  - High power but sensitive to sample period.
- **Out-of-sample Tests**:
  - Evaluate model on unseen data

Avoiding P-hacking in Backtesting

Definition

The p-value indicates how likely it is to see your results just by random chance – assuming there is actually no real effect or difference.

- **P-hacking/Data Mining**:
  - Testing multiple strategies on the same dataset increases the chance of finding false positives.
- **Solution**:
  - **Real money performance**: At least 3 years of live performance for reliability.
- **Takeaway**: Ensure your trading strategy is robust and economically meaningful.

Statistical Learning

- **Key Concepts**:
  - **Mean Squared Error (MSE)**: Measures model accuracy, $\text{MSE} = \frac{1}{n}\sum(y_i - \hat{y}_i)^2$.
  - **Bias-Variance Tradeoff**: Flexible models reduce bias but increase variance.
  - **Overfitting vs. Underfitting**: Balance model complexity to minimize test MSE.
- **Methods in R**:
  - Linear Regression (OLS): Baseline model, $\beta = (X'X)^{-1}X'y$.
  - Ridge Regression: Adds penalty $\lambda \sum \beta_j^2$ to reduce variance.
  - Lasso: Penalty $\lambda \sum |\beta_j|$ can set coefficients to zero for sparsity.
  - Elastic Net: Combines Ridge and Lasso penalties.

Resampling Methods for Model Selection

- **Purpose**: Estimate test error and select optimal model flexibility.
- **Cross Validation**:
  - **Validation Set Approach**: Split data into training and testing sets.
  - **Leave-One-Out CV (LOOCV)**: Use $n-1$ observations for training, 1 for testing.
  - **K-fold CV**: Divide data into K parts (e.g., $K = 5$ or 10), average MSE across folds.
- **Practical Tip**: Use K-fold CV ($K = 5$ or 10) for a balance of bias and computational efficiency.

Group Project: Overview

- **Task**: Develop and backtest an economically sensible trading strategy using crypto price data.
- **Data**: Full dataset available at https://cloudstore.uni-ulm.de/s/RjHJZRtWJEtTNAc.
- **Deliverables**:
  - **Written Report**: R Markdown.
  - **Presentation**: Present final results.

Group Project: Goals

- **Objective**: Create a trading strategy that makes economic sense (e.g., based on predictors like price trends, volatility, or volume).
- **Data Usage**:
  - Focus on price data.
- **Expectations**:
  - **Originality**: Do not just adapt demo code; develop creative strategies.
  - **Quality**: High-quality analysis and robust backtesting (most important: out-of-sample tests).
  - Use R for implementation (leverage Ridge, Lasso, or other methods from the course).
- **Tips**:
  - Use cross-validation to select model parameters (e.g., λ in Lasso).
  - Avoid p-hacking by testing strategies thoughtfully.
  - Document your work clearly in the report.

Resources and Support

- **Reading**: *An Introduction to Statistical Learning* (free eBook, selected chapters).
- **Videos**: www.statlearning.com, www.dataschool.io/15-hours-of-expert-machine-learning-videos.
- **Course Materials**: Available on Moodle.
- **AI Tools**: You are welcome and advised to use AI tools like ChatGPT, Cursor, Claude, Gemini (free trial month of Gemini Pro!), etc.
- The room is book every week throughout the semester – You are welcome to meet here on Wednesdays with your groups and to collaborate between groups.

Questions and Answers

- Please ask any questions :)
- Are you ready to start your group work?
- Good luck, and we look forward to your reports and presentations!