README

Alex Lostado

Description of Files

- `data_fetching.R` downloads data from the various sources used, cleans it, and stores it. The data sources are:
  - NilssonHedge CTA Database
  - Kenneth R. French 5 Factor Data
  - AQR Momentum Index Data
  - David A. Hsieh's Hedge Fund Risk Factors Data Library
- `factor_models.R` runs each factor model for each fund using all the data available and generates (and saves) graphs.
- `rolling_factors_datagen.R` runs each factor model for each fund on a 12-month rolling window. This takes about 1h 20min to run (on my computer).
- `rolling_factors_graphing.R` compiles the results from `rolling_factors_datagen.R` and generates (and saves) graphs.

Factor Models

What follows is a technical detail of the computations run by the aforementioned factor models.

Setup

Let $r_{i,t}$ represent the (monthly) return of fund $i \in \{1, \dots, N\}$ on month t. Let $\mathcal{T}_i$ be the set of months for which there is data on fund i, so that $t \in \mathcal{T}_i$. Funds are given labels or classifications depending on their characteristics. $\mathcal{I}_k$ denotes the set of funds belonging to group/classification k, with $k \in \mathcal{K}$, so that $i \in I_k$ if fund i is classified as k. Then, $\cup_{k \in \mathcal{K}} \mathcal{I}_k = \{1, \dots, N\}$.

General Factor Models

A factor model is a relation of the following kind:

$$\begin{aligned} r_{i,t} &= \alpha_i + \beta_{i,1} f_{1,t} + \cdots + \beta_{i,m} f_{m,t} + \varepsilon_{i,t} \\ &= \alpha_i + \boldsymbol{\beta}_i' \boldsymbol{f}_t + + \varepsilon_{i,t} \end{aligned}$$

The second line is in vector notation. The factors are $\boldsymbol{f}_t = (f_{1,t}, \dots, f_{m,t})'$ and the factor loadings or betas are $\boldsymbol{\beta}_i = (\beta_{i,m}, \dots, \beta_{i,m})'$. In finance, the common interpretation is that the factors are "risk" factors, and commonly include series such as excess market returns and/or momentum portfolios. Usually, excess returns are used both for the left hand side and the factors on the right hand side. In this setting, the parameter α_i would denote a fund's "unexplained" return, and would therefore indicate whether, on average, a fund outperforms or underperforms the factor model benchmark.

CAPM

The CAPM (Capital Asset Pricing Model) is

$$r_{i,t} - r_t^{(f)} = \alpha_i + \beta_{i,\text{Mkt-RF}} r_{\text{Mkt-RF},t} + \varepsilon_{i,t}$$

where $r_{\text{Mkt-RF},t}$ is the market excess return. The risk-free rate $r_t^{(f)}$ is the one-month Treasury bill rate. Details on factor portfolio construction regarding Kenneth R. French data can be found here (also applies to the FF3 and FF5 factors).

FF3

The FF3 (Fama-French 3-Factor) model is

$$r_{i,t} - r_t^{(f)} = \alpha_i + \beta_{i,\text{Mkt-RF}} r_{\text{Mkt-RF},t} + \beta_{i,\text{SMB}} r_{\text{SMB},t} + \beta_{i,\text{HML}} r_{\text{HML},t} + \varepsilon_{i,t}$$

where SMB is the "Small-Minus-Big" factor, and HML is the "High-Minus-Low" factor.

FF5

The FF5 (Fama-French 5-Factor) model is

$$r_{i,t} - r_t^{(f)} = \alpha_i + \beta_{i,\text{Mkt-RF}} r_{\text{Mkt-RF},t} + \beta_{i,\text{SMB}} r_{\text{SMB},t} + \beta_{i,\text{HML}} r_{\text{HML},t} + \beta_{i,\text{RMW}} r_{\text{RMW},t} + \beta_{i,\text{CMA}} r_{\text{CMA},t} + \varepsilon_{i,t}$$

where RMW is the "Robust-Minus-Weak" factor, and CMA is the "Conservative-Minus-Aggressive" factor.

FF5+MOM

The FF5+MOM (Fama-French 5-Factor plus Momentum) model is

$$r_{i,t} - r_t^{(f)} = \alpha_i + \beta_{i,\text{Mkt-RF}} r_{\text{Mkt-RF},t} + \beta_{i,\text{SMB}} r_{\text{SMB},t} + \beta_{i,\text{HML}} r_{\text{HML},t} + \beta_{i,\text{RMW}} r_{\text{RMW},t} + \beta_{i,\text{CMA}} r_{\text{CMA},t} + \beta_{i,\text{MOM}} r_{\text{MOM},t} + \varepsilon_{i,t}$$

where MOM is the AQR Large-Cap and Mid-Cap U.S. Equities Momentum Index.

LBS

The LBS (Lookback Straddle) model is

$$r_{i,t} - r_t^{(f)} = \alpha_i + \beta_{i,\text{PTFSBD}} r_{\text{PTFSBD},t} + \beta_{i,\text{PTFSFX}} r_{\text{PTFSFX},t} + \beta_{i,\text{PTFSCOM}} r_{\text{PTFSCOM},t} + \beta_{i,\text{PTFSIR}} r_{\text{PTFSIR},t} + \beta_{i,\text{PTFSSTK}} r_{\text{PTFSSTK}}$$

where:

- PTFSBD: return of PTFS bond lookback straddle
- PTFSFX: return of PTFS currency lookback straddle
- PTFSCOM: return of PTFS commodity lookback straddle
- PTFSIR: return of PTFS short-term interest rate lookback straddle
- PTFSSTK: return of PTFS stock index lookback straddle

More information about this model can be found here.

Analysis

For the analysis in `factor_models.R`, we use OLS to regress $r_{i,t}$ on $\boldsymbol{f}_t$ (and a constant term) to obtain fund-specific loadings $\boldsymbol{\beta}_i$. This is repeated for every model. The graphs produced show the distributions of these estimated loadings using kernel density estimations. These are grouped into results for funds of the same Type.

In `rolling_factors_datagen.R`, the same approach is taken but, this time, using a 12-month rolling window to obtain fund-specific loadings at each point in time at and after the initial 12 months. The loading estimates are grouped by Type and, at each point in time, the median, and 50%, 80%, and 95% bands around the median are taken and graphed over time. This gives a visualization of the development of the distribution of factor loadings over time.

Due to data availability constraints, not all funds are included in the analysis. Funds whose series of returns is discontinuous or completely non-overlapping with factor data are excluded from both computations. Funds with more observations than the number of factors in the model plus one (accounting for α) are *not* excluded from the former computation, the rest are. Funds with less than 12 return data points are excluded from the latter computation.