Measuring Asset Similarity

Luosho Investment Shanghai

Motivation

Measuring similarity between financial assets is central in the financial industry for long term portfolio selection, risk assessment, marginal contributions or dimension reduction. This is particularly relevant in China where the structure of the market and price movements depend on factors that are different than in western markets.

Correlation is overwhelmingly used in finance as a proxy for such a measure of relationship between financial assets. This is ubiquitous to the fact that normal distributions are the basis of most financial approaches and correlation fully describe those distributions. However this reliance on correlation to measure dissimilarity faces several drawbacks:

- It is a first order appreciation of dependence motivated by normal distribution and i.i.d. Returns.
- It tracks very badly regime changes, tail dependence and models that are beyond the normal distribution world.
- In terms of estimation, the accuracy tends to be very bad as the dimension increases or the number of observation in a similar regime is relatively low.
- It is an approach that fundamentally ignores any underlying economic relationship between assets and specific to Chinese markets, ignore the influence of large investors, words of mouth and structure of investors behind each listed company.

It is furthermore quite important in China as several factors are not available for assets analysis -- option prices, implied volatility, etc. Therefore, several methods have to be used to filter informations that are relevant for the evolution of financial assets.

Project

This project intends to approach and test alternative measure of similarity and measure their performance against classical benchmark such as Barra investment model. Some directions that can be explored -- but are not bounded to -- includes

- Using covariance estimation methods based on version of multidimensional GARCH models or related but including additional information to detect regime switch -- quite present in China -- as
well as large dimension. Several external factors can be used to enhance the model such as shareholder data, earning calls reports, abnormal trading activities with broker ID, etc. Another approach under study are denoising methods with random matrix.

- In order to handle the dimension problem and provide more efficient investment decision, we want to implement clustering methods alternative to usual branch clustering provided by regulators and diverse private company. This could involve unsupervised clustering methods such as K-Mean, mixture models or Wasserstein distance. The goal is to use it previous to similarity analysis. The challenging problems in that regards is the choice of features that are relevant in China -- analysis of network of shareholders and their behavior -- to provide a robust clustering methods. Furthermore, the stability of clusters over time is a very important. Finally, after clustering, and similarity analysis, how to choose the optimal strategy in selecting stocks from clusters -- monkey approach, index based, largest equity, etc.
- Assess the performance of this similarity methods against standard methods towards portfolio selection benchmark, risk measures, or any meaningful performance instrument.

Data

Towards this project, the company will provide a set of cleaned data ready to implement in any platform of your choice (Python, R, Matlab, etc.). Aside the typical market datas, we compile the relevant data that may explain the behavior of markets in China.

1. Intra-day data of 3000 stocks traded in China (price volume, etc.) over 10 years (clean) and 20 years (non clean but different market behavior before)
2. 10 years of financial reports of all the companies (quarterly earning calls). We provide them in two forms:
   a. A quantitative one with hard numbers
   b. The complete text report that contains information between the companies and some of the shareholders position. Could be used with text analysis methods.
3. Sector classification, from regulators, and from different Brokers by industry, region, and different levels of granularity.
4. Major shareholder structure of the 3000 companies over time.
5. Macro economic data
6. Future index data (only relevant for a subset of companies)
7. Analyst prediction (anonymized)
8. Broker ID of abnormal trades -- China specificities reporting abnormal price changes with the list of the largest position on that trade.

About the company
Luoshu Investment was established in 2015 and its headquarters are located in Pudong, Shanghai. With a total asset under investment over 1.2B rmb, the fund is focused on quantitative trading with most of its employes performing analysis and research with an educational background in mathematics and related fields as well as computer sciences. Luoshu Investment's trading strategies are model based with holding periods ranging from high frequency to long term strategies. Its investment underlyings includes equities, commodities futures and options. Luoshu Investment will continue to increase research investment in the future, aiming to provide investors with sustained and excellent return on investment with advanced financial technology and robust quantitative models.