Math 142 Midterm Project Proposal

Abhi Uppal

September 13, 2021

1 Background

Financial machine learning is a notoriously difficult field – feature engineering, non-stationarity, violation of IID assumptions, high probabilities of overfitting, low signal-to-noise ratios, and so many more factors make it a daunting task. Finding even a small signal-turned-feature that trumps your competitors could translate into a massive advantage in live trading performance.

In this sense, analyzing one ticker at a time leaves much to be desired. The stock market, for example, is complex not only because each individual company engages in unpredictable activities, but because there is an implicit structure behind which everything operates. If one can understand this structure, one may be able to make better predictions and therefore better decisions. In essence, a graph structure on the market could provide additional signal to be used in predictive models (for example, using node embeddings of a market graph as a feature in predictive models to aggregate neighborhood information).

We note that, once a graph has been constructed, we can start applying some differential geometric techniques. A graph is, in essence, a one-dimensional approximation of a manifold. The core of this project is to see if we can extract any useful geometric features from the graph itself.

2 Approach

The first question that comes to mind is how might one construct such a graph? A survey by Marti et al. done in a chapter of the book "Progress in Information Geometry" [1] details many classical approaches that allow us to construct hierarchical graphs that are more complicated and information-rich than simple linear correlation networks via a variety of distance measures and tree-construction algorithms. Methods will be adopted from this paper to construct the initial network.

Once the network is constructed, what do we do with it? This question is, in part, answered by a paper by Sandhu et al. [3] which states that the Oliver-Ricci curvature of a stock market graph (a discrete approximation of the Ricci curvature of a manifold) is an indicator for a market crash. However, the paper only adopted a simple network generated by a minimum spanning tree (MST) where distances where computed via angular separation. That is, the authors imposed a metric between time series of the form

$$d(i, j) = \sqrt{1 - \rho(i, j)}, \tag{1}$$

where $\rho(i, j)$ is the Pearson (linear) correlation between two series, and constructed an MST using this metric. The survey above [1] indicates many of the issues with this approach, including but not limited to extreme sensitivity to noise and the limitations of the linear correlation coefficient. Sandhu et al.'s approach may then be able to be improved upon by constructing a network using a stronger construction algorithm (for example, a Planar Maximally Filtered Graph (PMFG) or clustering via maximum likelihood estimation).

Once this has been done, the analysis can be extended by applying more methods in information geometry. While it is currently out of the scope of my knowledge, the survey by Nielsen [2] provides an introduction to the field and methods that may be used.

The essence of this project is to find and extract features to be used in financial machine learning. That is, we want to find features that either help us incorporate market structure into machine learning pipelines (graph structure and embeddings), or geometric features that may help us predict a crash.

References

[1] Gautier Marti et al. "Progress in Information Geometry". In: *Signals and Communication Technology* (2021). ISSN: 1860-4870. DOI: `10.1007/978-3-030-65459-7`. URL: `https://arxiv.org/abs/1703.00485`.

[2] Frank Nielsen. "An Elementary Introduction to Information Geometry". In: *Entropy* 22.10 (Sept. 2020), p. 1100. ISSN: 1099-4300. DOI: `10.3390/e22101100`. URL: `http://dx.doi.org/10.3390/e22101100`.

[3] Romeil Sandhu, Tryphon Georgiou, and Allen Tannenbaum. *Market Fragility, Systemic Risk, and Ricci Curvature*. 2015. arXiv: `1505.05182 [q-fin.RM]`.