

I'll share a few:

[alo.mit.edu](https://alo.mit.edu)

[

]([https://alo.mit.edu/wp-content/uploads/2015/06/Learning\\_Connections\\_in\\_Financial\\_Time\\_Series.pdf](https://alo.mit.edu/wp-content/uploads/2015/06/Learning_Connections_in_Financial_Time_Series.pdf))

### [Learning Connections in Financial Time Series.pdf](https://alo.mit.edu/wp-content/uploads/2015/06/Learning_Connections_in_Financial_Time_Series.pdf)

477.30 KB

#### Learning Connections in Financial Time Series

To reduce risk, investors seek assets that have high expected return and are unlikely to move in tandem. Correlation measures are generally used to quantify the connections between equities. The 2008 financial crisis, and its aftermath, demonstrated the need for a better way to quantify these connections. We present a machine learning-based method to build a connectedness matrix to address the shortcomings of correlation in capturing events such as large losses. Our method uses an unconstrained optimization to learn this matrix, while ensuring that the resulting matrix is positive semi-definite. We show that this matrix can be used to build portfolios that not only “beat the market,” but also outperform optimal (i.e., minimum variance) portfolios.

Although it uses a quadratic programming approach instead of an ML-based optimizer, the sections on connectedness are useful for anyone that wants to understand a method for addressing shortcomings of covariance-based approaches when dealing with long-tail/black-swan events.

[https://www.researchgate.net/publication/228267339\\_The\\_101\\_Ways\\_to\\_Measure\\_Portfolio\\_Performance](https://www.researchgate.net/publication/228267339_The_101_Ways_to_Measure_Portfolio_Performance)

#### The 101 Ways to Measure Portfolio Performance

This paper performs a census of the 101 performance measures for portfolios that have been proposed so far in the scientific literature. We discuss their main strengths and weaknesses and provide a classification based on their objectives, properties and degree of generalization. The measures are categorized based on the general way they are computed: asset selection vs. market timing, standardized vs. individualized, absolute vs. relative and excess return vs. gain measure. We show that several categories have been exhausted while some others feature very heterogeneous ways to assess performance within the same sets of objectives

If you liked [Performance Stationarity](#), then you'll enjoy this paper, which has a healthy discussion of the pros and cons of all the different methods for measuring your performance.

[arxiv.org](https://arxiv.org)

[

](<https://arxiv.org/pdf/1907.02893.pdf>)

### [1907.02893.pdf](https://arxiv.org/pdf/1907.02893.pdf)

923.12 KB

#### Invariant Risk Minimization

We introduce Invariant Risk Minimization (IRM), a learning paradigm to estimate invariant correlations across multiple training distributions. To achieve this goal, IRM learns a data representation such that the optimal classifier, on top of that data representation, matches for all training distributions. Through theory and experiments, we show how the invariances learned by IRM relate to the causal structures governing the data and enable out-of-distribution generalization.

Sections 1 and 2 give a really good background on the issues with training ML models on “data collected in different environments”. It's not explicitly about time series modelling, but I think the issue translates well into the Numerai dataset. The main takeaway I took from this paper was “it would be great if your ML model came out the same, regardless of which rows you sample for training/validation”.