Alien Stock Market Intelligence — Numerai's True Contribution

Richard Craib

Follow

Numerai

--

6

Listen

Share

Numerai has released a new signal evaluation method called True Contribution. True Contribution is computed by treating Numerai as an end-to-end artificial intelligence system. By computing the gradient of optimized portfolio returns with respect to the NMR staked on a signal using

[differentiable convex optimization layers

](https://github.com/cvxgrp/cvxpylayers), Numerai can now surface and incentivize signals making the largest intelligence contributions to our hedge fund.

[Numerai's performance is already good

](https://numer.ai)(our Sharpe Ratio* is 2.54 over the last 12 months) but True Contribution has the potential to make Numerai the first "Type IV" hedge fund.

The Kardashev Scale

The Kardashev scale proposed by Soviet astrophysicist Nikolai Kardashev is a way to grade the technological advancement civilizations based on the energy they harness. A Type I civilization can harness all energy on its planet. A Type II can harness all the energy in its solar system whereas a Type III civilization can harness all energy in its galaxy.

The Kardashev scale contains no practical blueprint for creating these civilizations but it does make for fertile ground for thought experiments in considering the technologies the different types of civilizations would require. For example, in order to become a Type II civilization we would have to invent something like a Dyson Sphere to capture all the energy from our sun. In order to become a Type III civilization, we would require the technology for self-replicating robots in order to span the galaxy.

A Kardashev Scale For Hedge Funds

So is there a Kardashev scale for hedge funds? What would meaningfully separate different levels of technological advancement in hedge funds?

We can think of every hedge fund as trading some signal of stock market predictions which should be correlated with future returns in stocks. If the signal is predictive of future returns, the best stocks to buy in a signal would have the highest signal values and the best stocks to short would have the lowest. For quantitative hedge funds, their signal comes directly out of their mathematical models.

I want to propose the following Kardashev Scale for hedge funds which grades hedge funds based on the quality of their signal:

A Type I hedge fund has a signal which is predictive of subsequent returns before trading costs.

A Type II hedge fund has a signal which is predictive of subsequent returns after trading costs.

A Type III hedge fund has a signal which can be used to make profitable alterations to the best hedge fund in the world's signal.

A Type IV hedge fund has a signal which is so good that no other signal can be used to make profitable alterations to it.

A Type I hedge fund has the ability to predict future price movements; however, those predictions cannot be taken advantage of profitably because trading costs such as market impact costs would eliminate the gains. Type II hedge funds

can make money after costs and therefore violate the efficient-market hypothesis.

Type III is where things get interesting so let me explain a Type III hedge fund with an example. Suppose we could identify the best hedge fund in the world and let's just say it is Renaissance Technologies. Suppose in some alternate universe, the CEO of Two Sigma (another hedge fund) decided the only thing he wanted to do was improve

Renaissance's signal.

He decides to give Renaissance Two Sigma's best signal — their highest quality stock predictions. He uploads Two Sigmas best signal to an FTP server that Renaissance can now see and use. Now, if Renaissance decides their own signal is so good already that there is no way to use Two Sigma's signal to make profitable alterations to their trades

, then Two Sigma is not a Type III hedge fund. But if there is a way for Renaissance to use Two Sigma's signal to make profitable alterations to their trades, then Two Sigma is a Type III hedge fund. Even if Renaissance still makes 96% of the same trades they would normally make without Two Sigma's signal, the 4% of profitable trade alterations which improve Renaissance's signal would be enough to make Two Sigma a Type III hedge fund.

Being a Type III hedge fund is a lot harder than being a Type II hedge fund because a Type II hedge fund only has to make profitable trades in the market whereas a Type III hedge fund has to make profitable alterations to a signal that is already making very profitable trades in the market.

A Type IV hedge fund has a signal which no one can make profitable alterations to. A Type IV hedge fund wouldn't just be the best hedge fund in the world, it would be a hedge fund where no other known signals anywhere in the universe could be combined to its signal to improve it. (In the example above, Renaissance would not be a Type IV hedge fund if Two Sigma is a Type III hedge fund under Renaissance trade alterations.) A Type IV hedge fund does not necessarily trade a perfect signal with perfect stock market prediction accuracy, it just means the signal it does trade is maximally good for all that is currently known.

It has integrated all known signals perfectly.

A Type IV hedge fund would be like an alien super intelligence for the stock market. It would be a bit like the best possible version of DeepMind's AlphaZero playing Go where no alterations to its game by humans (or older versions of AlphaZero or AlphaGo) could improve it.

I don't think a Type IV hedge fund exists yet. But I'm the founder of a new kind of hedge fund called Numerai and I've spent the last few years pursuing a thought experiment: what are the necessary properties of a Type IV hedge fund and can it be built?

Necessary Properties Of A Type IV Hedge Fund

Clearly Type IV hedge fund would need to be able to rapidly and automatically onboard any new signal which is keeping them from being a Type IV hedge fund. If there is a Type IV hedge fund, it would be the best hedge fund therefore there cannot be any Type III hedge funds whose signals are not assimilated by the Type IV hedge fund. A Type IV hedge fund must be able to assimilate any Type III signals instantaneously or they would not be a Type IV hedge fund. Because of this, a Type IV hedge fund would need to be an open system

where new signals could be uploaded by anyone to make all the alterations to trades that they are able to profitably make.

Renaissance, Two Sigma or any other hedge fund for that matter have closed pre-internet, pre-blockchain organizational designs which cannot create a Type IV hedge fund just like Citibank's organizational design could not create Bitcoin. A Type IV hedge fund would be quite a new kind of thing indeed. It might look and feel a bit more like Bitcoin than Two Sigma.

Of course, an obvious property of a Type IV hedge fund is that it would need to be rich enough to offer enough money to buy up all the Type III signals.

If Two Sigma can make profitable alterations to a candidate Type IV hedge fund's trades, then that Type IV hedge fund needs to have the capital to incentivize Two Sigma to shut down its trading business altogether and instead sell their signals to the Type IV hedge fund

. Getting Two Sigma to turn over their signal would be very expensive indeed. Perhaps even Renaissance couldn't afford to become a Type IV hedge fund even if they really wanted to be one.

A Type IV hedge fund would have to be an open market place for buying signals.Numerai is such a system already. Anyone can submit signals to Numerai using our free obfuscated data or Numerai Signals using their own data. Numerai doesn't have that much money but we do have about $150m in our own cryptocurrency, NMR, which makes Numerai already by far the highest paying data science competition in the world and the largest buyer of stock market signals on the internet.

With an open way to onboard signals and cryptocurrency to incentivize people to submit new signals, it may appear as though Numerai has the right properties to become a Type IV hedge fund.

A Hidden Assumption

But there is a lingering, pernicious assumption in this line of argument which is hiding from view. It's the kind of assumption that is missed in thought experiments but shows up in practical reality. And here it is: that Numerai would have any idea how to evaluate whether some new signal will improve our existing good signal.

Numerai already combines thousands of signals submitted by our data scientists into what we call the Meta Model. So given a new signal, how do we know we can include it in this already large ensemble of signals and have that inclusion generate profitable alterations to the trading strategy? Without a good technological solution to this signal evaluation question it is impossible to know whether a signal is of Type III and therefore reaching Type IV is impossible.

Over the years at Numerai, we have had to learn how to get good at signal evaluation. We have made many improvements to signal evaluation since Numerai was founded. For example, we launched MMC and staking where Numerai data scientists stake NMR on their models to prove they believe their models will work (generalize) out of sample. Both MMC and staking improved the quality of signals on Numerai.

But we have never solved the signal evaluation problem. We could not determine whether in our end to end system whether a signal was Type III with respect to our existing Meta Model. But today we are announcing a new system we have spent years building towards called True Contribution, which solves the signal evaluation problem.

Signal Evaluation With Information Coefficient

It seems natural to assume that signals with high correlation with future stock returns are likely to be the most helpful in Numerai's Meta Model. Because of this, Numerai has spent many years making payouts to data scientists based on how good they are at generating signals with strong correlation with Numerai's targets (which are like residual returns). But this incentivizes Type I signals, and rewarding signals based on their correlation with targets isn't an accurate representation of how much that signal truly contributed to Numerai's post-optimization portfolio return. Rewarding users based on their correlation with targets simply ignores too many very important details such as: the correlation of the signal with other signals, its interaction effects with the existing Meta Model or the hundreds of portfolio optimization parameters that Numerai uses to turn the Meta Model signal into a balanced portfolio of a few hundred stocks.

Clearly we need a signal evaluation method which accounts for every detail in the system to be able to evaluate whether a signal can make profitable alterations to Numerai.

End-To-End Signal Evaluation

To be end-to-end

we need to map out every aspect of how a signal and the NMR stake associated with that signal affects the final portfolios constructed by Numerai's optimizer.

As you can see in the diagram above, Numerai first combines the signals generated by data scientists' machine learning models. We do this by computing the stake-weighted average of every signal to create the Stake-Weighted Meta Model. A data scientist who stakes a large amount of NMR on their model will have a larger weight in the Stake-Weighted Meta Model.

The Stake-Weighted Meta Model is still just a signal of predictions on ~5000 global equities. It still needs to be turned into a realistic portfolio with hundreds of risk constraints (such as market, country and sector risk neutralization), and that's what the optimization step does. Once the optimizer creates a realistic hypothetical portfolio that satisfies all risk constraints, Numerai can observe the subsequent returns of the portfolio.

To evaluate a signal properly, we must consider the signal's effect on the whole system above from signal to Stake-Weighted Meta Model to returns of a post-optimization portfolio. And that's what True Contribution does.

True Contribution

Put simply, True Contribution is the answer to the question: if a data scientist staked slightly more on their model (thereby increasing their weight in the Stake-Weighted Meta Model), what would the change be to post-optimization portfolio returns?

To someone in quantitative finance, you could think of True Contribution as a sophisticated signal attribution

.

To someone in machine learning, the diagram above of how Numerai works may look reminscent of a neural network architecture. And if you've ever built neural networks, you might be wondering if it is possible to take the gradient of the optimized portfolio return with respect to the stake

. That is exactly what True Contribution is.

But how is the gradient computed through the portfolio optimization layer

?

As it turns out, this is possible using new techniques developed by Stephen Boyd of Stanford University, Brandon Amos from Facebook AI et al (see their paper: Differentiable Convex Optimization Layers)

By using cvxpylayers, we can include a cvxpy defined convex portfolio optimization as a layer in a PyTorch model.

This lets us efficiently compute the gradient of the optimized portfolio return with respect to the stake values and determine the True Contribution of every signal submitted to Numerai.

Consequences of True Contribution

With True Contribution portfolio construction, size of stake, originality of the model and the strength of the signal are all taken into account in the exact proportions in which they actually matter for producing returns in realistic portfolios that Numerai could actually trade.

Signals which are original and help Numerai to make different, and more profitable trades to what we otherwise would have will now receive the highest NMR rewards on their stakes, and will therefore tend to have larger and larger weights in the Stake-Weighted Meta Model. This reward feedback is important to every data scientist on Numerai too because they can now improve their models to maximize their True Contribution.

With True Contribution, Numerai is creating a feedback loop designed to continuously incentivize the creation and submission of signals

which make profitable alterations to our hedge fund and disincentivize all other signals. Every round of Numerai will become like another pass of backpropagation on the overall cybernetic system of Numerai. Feedback and error correction propagate through a layer of distributed AI models, a blockchain staking layer, a Meta Model, and a convex optimization. In other words, with every round of Numerai we take a step closer towards becoming a Type IV hedge fund.

Early Indications of True Contribution

Staking on True Contribution will begin on April 9th but in the meantime we have backfilled True Contribution for every user on Numerai for the last ~2 years.

The results show large potential for True Contribution as a new signal evaluation metric on Numerai.

For example, there are many data scientists such asLANCE_A_LOT with very high True Contribution ranks but much lower ranks on other metrics such as correlation with the target. It's clear that at least over the recent period, LANCE_A_LOT had a model that was helping Numerai the most

but they were not rewarded properly for their contribution.

HB is an engineer at NASA Jet Propulsion Lab working on theEuropa Clipper Mission, and a long time Numerai data scientist. HB has multiple models submitted to Numerai but the model he is staking the most NMR on (765 NMR worth $22,000) has much lower

True Contribution than his other models. Some of his best models in terms of True Contribution have no stake on them at all meaning these excellent models have zero weight in the Meta Model

and are not being rewarded at all.

With the ability to start staking on True Contribution, data scientists will begin moving their stakes to models with the highest expected True Contribution. In this dynamical system, as data scientists like HB and LANCE_A_LOT adjust their stakes to earn more True Contribution, there is no reason to believe Numerai's Stake-Weighted Meta Model could not become significantly more intelligent.

Better With Time

Of course, we don't know whether we will reach Type IV. But what type of things would you expect to see if we were on the path to become Type IV? I think the key thing to look for is risk adjusted performance increasing with time and AUM

.

Hedge funds tend to get worse with time

. How sad. This happens because of signal decay (the trading signal a hedge fund started with gets worse with time as the market gets more efficient) and capacity constraints (the trading strategy that worked great on $10m doesn't work at all on $100m).

But Numerai so far isn't getting worse with time at all — it's getting better, and that's a very good sign.

Over the period in the graph, Numerai's hedge fund AUM has almost 10xed from about $7m to $64m (still early). The decay on our signal over time and the AUM growth should hurt performance over time but our risk adjusted returns (Sharpe Ratio) continue to increase with time. And that's because over this same period, Numerai has gone from 300 staked models to over 4000 staked models in the Meta Model. Numerai does not have signal decay, we have constant signal rejuvenation — the Meta Model is rebuilt with the latest signals every week.

What's remarkable about this continued increase in our risk adjusted Sharpe happened without Numerai users receiving the correct feedback on their models

i.e. without True Contribution…