WHY FMI BACKTESTS ARE EQUIVALENT TO FORWARD TESTS (A MATHEMATICAL PROOF)

A PREPRINT

Revant Nayar
CTO
FMI Technologies
New York, NY, 10004
rnayar@fmitech.net

Douglas Stanford
School of Natural Sciences
IAS, Princeton
douglas@fmitech.net

March 12, 2019

ABSTRACT

Field Machine Intelligence (FMI) can be represented as a continuum limit of neural networks. It might then appear that it too is prone to backtest overfitting- i.e, the observations that forward tests are significantly worse than backtests as a result of the algorithm having learnt specifics of the noise that do not generalise. In this document we prove the fundamental reason why field machine intelligence eliminates all backtest over-fitting and look ahead biases, as opposed to most other quantitative algorithms including all known forms of artificial intelligence. The first reason is that there are no hyper-parameters- i.e, those parameters of the model which are not fit but have to be set by hand before the other parameters are fit to the data (in neural networks, the 'weights'). This is unlike a machine learning model which cannot be run in the first place without specifying the depth of the net or the value of some hyper-parameters. The second reason is that model selection does not come into play wherein the best models are retrospectively selected. We will make this more rigorous in Section 2. The third major kind of overfitting that is eliminated is due to regime shifts discussed in Section 3.

***K*eywords** Backtest Overfitting · Field Theory · Artificial Intelligence

1 Background

This document presents proof that backtests conducted using field machine intelligence are equivalent to forward tests unlike the majority of, if not all other trading algorithms known. We begin with some context. We have seen the same phrase repested in quantitative finance circles,"I have never seen a bad backtest." It is often seen that stellar backtests mostly translate to poor forward test results in trading. In particular, 70 percent of quantitative hedge funds and 95 percent of machine learning based hedge funds have in the last five years failed to beat the market. In fact backtest overfitting is a problem with strategies devised by established traders and hedge funds and are often caught in strategies published in quantitative finance journals. This is due to two reasons. First, backtest overfitting the second is model selection and the third is regime dependence.

2 Model Selection

Repeatedly backtesting an ML algorithm over the same data points produces 'pseudo-discoveries'- strategies with unexpectedly high Sharpe ratios. The algorithm should be allowed to access the data points a finite and limit number of times in principle; however in practice even large hedge funds often fail at this. In field theory, this problem does not arise since we are accessing the data points only as many times as to get to a certain predetermined level of nonlinearity. We neither oversample nor under sample data points. Note that an n point function of n possibly distinct variables is a function of n prices times and volumes:
$< \psi_1(P_1, V_1, t_1)...\psi_n(P_n, V_n, t_n) >= f(\{P_i, V_i, t_i\}_{i=1,n})$

Note that the function f is in general arbitrary. Let us conjecture that there exists a group G that has a predetermined action on the time series. The Lie Algebra is characterised by the rank three tensor k_{abc}.
$[G^a, G^b] = k^{abc} G^c$
Where summation over repeated indices is implied. This implies that the price-time generator of each element of the group transforms the variables in a predetermined way (for instance a trivial transformation leaves the observables $(< a_1 ... a_n >)$ invariant).
$G(\partial_t, t, \partial_P, P, V, \partial_V, ..., t^\alpha, P^\beta, ..., V^\gamma \partial_t^\theta \partial_P^\delta \partial_V^\mu) < \psi_1 ... \psi_n >= F_g(P, V)$
Demanding invariance under each generator will generally reduce the degrees of freedom by one. After combining known symmetries, one often gets new generators required to close the algebra. If the group has k generators it will usually reduce the degrees of freedom from $n-1$ to $n-1-k$. Then our observable is expressed in terms of a modified function f_{mod}:
$< \psi_1(t_1, P_1, V_1) ... \psi_n(t_n, P_n, V_n) >= f_{mod}(q_1, ... q_{n-1-k})$
Where the qs are now complicated cross ratios of the three original variables that will depend on the exact price-volume-time representation of the group H_i. Note that for each symmetry present in the operation not only do the number of degrees of freedom reduce by one, but also the number of samples leading to the signal multiply by n' for some $n' < n$.
Let us see this in a toy model suggested by de Prado. Suppose we have a sample of iid variables following a Gaussian distribution (x_i)- i.e, a market following the most vanilla efficient market hypothesis with pure white Gaussian noise and no persistence, reversal or other anomalies. Now we test I strategies on an inserting a martingale, with Sharpe ratios $y_{i=1,...I}$ such that
$E(y_i) = 0, \sigma^2(y_i) > 0$
Clearly the true Sharpe ratio is zero; yet we will expect to find one strategy with a Sharpe ratio:
$E(max(y_i)) = E(max(x_i))\sigma(y_i)$
In most ML applications, WF will imply most of the decisions are based on a small region of the dataset and the second moment is divergent. That will imply $\sigma(y_i)_N \to \infty$ as $N \to \infty$ and hence that we will artificially get an inflated Sharpe ratio. In FMI on the other hand, $\sigma(y_i)$ is nonlinearly quantified for each strategy even in cases where the second moment does not converge as proved above.

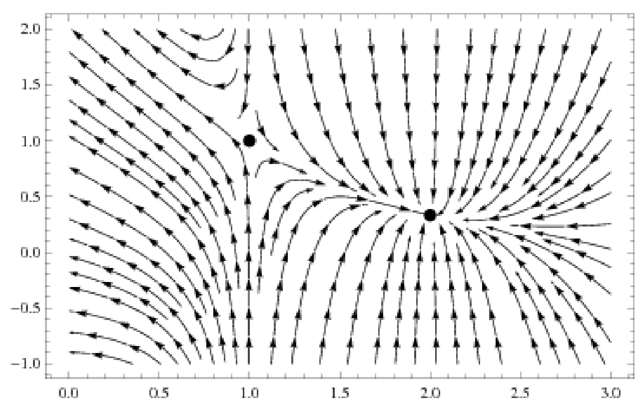

Figure 1: A simulation of renormalisation group

2.1 Backtest without Backtesting

We show that backtests conducted on FMI are free from both these problems and are thus in effect equivalent to forward tests in the specific case that G is the renormalisation group. To understand this, we have to understand that in FMI one specifies a trading horizon say τ.Say the full data is of length T. Assuming without loss of generality that T is divisible by τ by a factor n, we define the $\tau \times n$ rank two tensor:
T_{ij}
We have to introduce a renormalisation of the variables. The renormalisation group works by rescaling the time and space variables. In the continuum limit it is given by the Callan. Symanzik equation:
$(\Lambda \partial_\Lambda + \sum_i \beta \partial_{g_i} + \gamma) G = F(g_i, \Lambda)$
Clearly each iteration reduces one index of the tensor $T_{ij} \to T_i$ which is the new time variable and the process above repeats. Demanding a specific evolution under this transformation fixes the trading strategy in an EFT expansion.

3 Regime Dependence

Lastly there is the problem of regime dependence that one cannot mathematically prove that one is free of for any algorithm, by definition. It is well known that black box and gray box models overfit to particular market regimes and hence can create great losses when market regimes shift (i.e, for instance, correlations weaken due to a change in global macroeconomic variables). Note that if we had k symmetry generators we had $n_1 \times ... \times n_k$ samples. In the event of a regime shift the probability distributions for certain signals will all move. In particular we will have certain couplings destabilise from their equilibrium distributions. In fact in FMI we account for background shifts in small-price-time parameter expansions.

4 Results

Here you are being shown various three dimensional subslices of (in principle) an infinite dimensional space of all possible trading strategies. See the next page for a sample (more to come)

Time evolution of these returns are difficult to plot as there are no readily available tools for four dimensional visualisation in a static report. To learn more contact info@fmitech.net.

References

[1] Prado, Lopez de, and Marcos. “The 10 Reasons Most Machine Learning Funds Fail.” :: SSRN, 18 Jan. 2018,

[2] Makino, et al. “Gradient Flow and the Wilsonian Renormalization Group Flow.”, American Physical Society, 19 Apr. 2018, arxiv.org/abs/1802.07897.

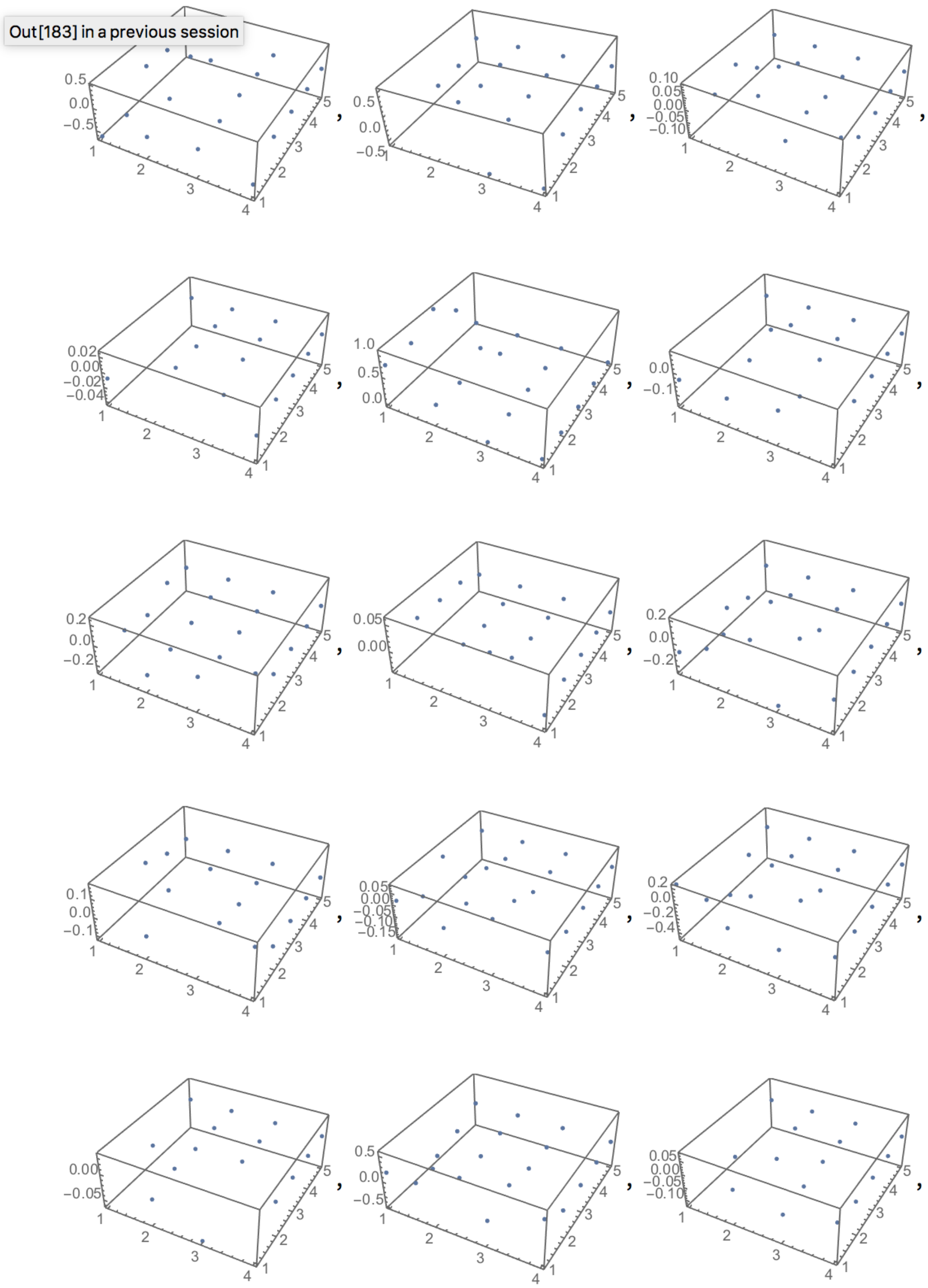

Figure 2: Three dimensional subslices of infinite dimensional space of all strategies