Machine Learning for Finance. Current works

Joseph MIKAEL / Webinar ILB/EDF/ACPR/London Turing Institute – Sep 21

Hedging

Hedging Strategy Computation : On the algorithm side

The PDE way : In the high dimensional cases (e.g. basket options), thanks to the pioneering work of Jentzen et. al (2017), the PDEs that are involved in the pricing and hedging of European & American style options are now solved with machine learning methods.

More recent work (Germain et.al (2021)) improve the resolution scheme and estimate the number of neurons necessary for the convergence.

The default risk can be included (Germain et. al (2020)).

Hedging Strategy Computation : On the algorithm side

The global forward way :

Regarding **european style options**, the most used algorithm is a global approach moving forward (Fécamp et. al (2020), Buelher et.al(2019)).

Dimension	Traditional Algorithms	Deep Hedging
1	< minute	Minutes on laptop
3	Hours on cluster	Minutes on laptop
>4	Not attainable	Minutes on laptop

Tab1. Computation time for a hedging of spread option with d underlyings

For **american style options**, a pure forward pass method is still difficult but some attempts exists (Deschatre et. al (2020))
In stylized cased, these forward algorithms are less precise than some backward state-of-the-art techniques but they offer a flexibility beyond competitors.

Hedging Strategy Computation : On the algorithm side

The local backward way :

Regarding **european style options**, Longstaff-Schwartz like Backward schemes benefits from proof of convergence. The regressions on basis functions are replaced by neural net regressions (Huré et. al(2020)).

For **american style options** still, the backward pass is the more precise.

Some attempts to replace the linear regression by Gradient Boosting technique have been proposed but does not seem as efficient as expected.

Hedging Strategy Computation : Global Forward VS Local Backward

	Local	Global
Resolution way	Backward	Forward
What is minimized	Sequence of loss functions	vs one global objective function
What can be minimized	Constrained Criteria (Time Consistent problems relying on DP)	Whatever empirically works
Implementation	Backward pass hard to implement and not fast to run	Easy & Fast to implement
Error control	Appropriate for error analysis: (Germain, Pham, Warin 21) + Rate of convergence w.r.t. number of layers and neurons of NN	Not a lot of available results
Theoretical results	Available	« »

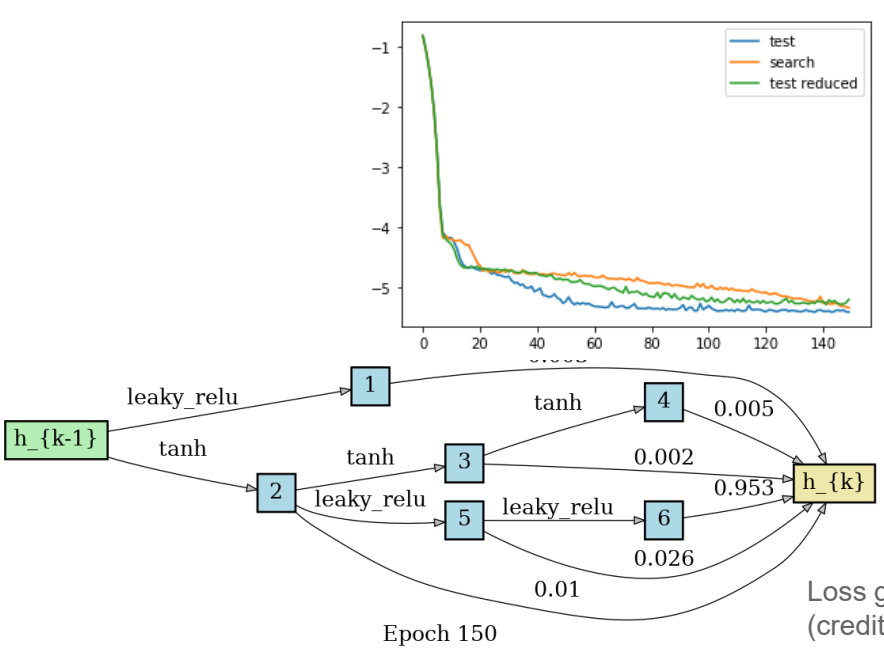
Hedging : On the Algorithm side/AutoML

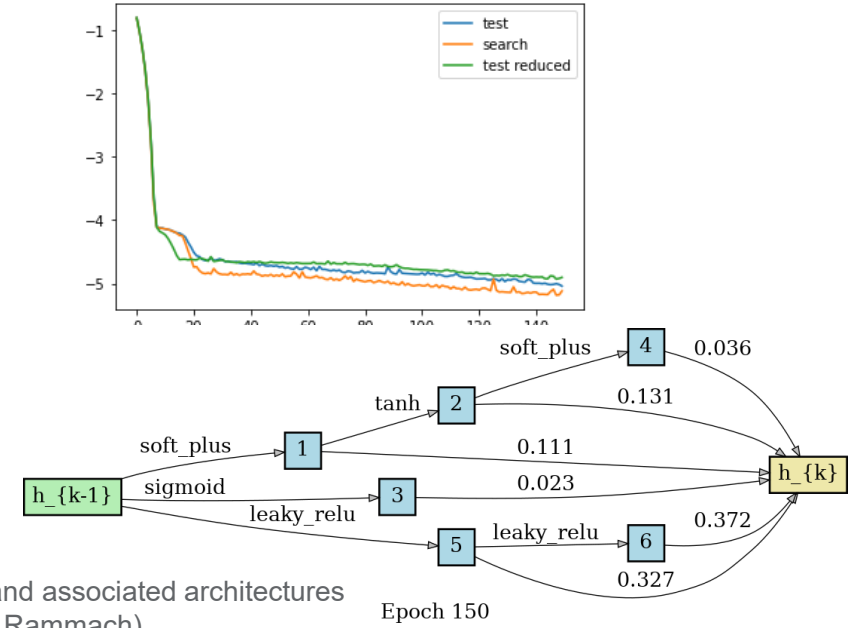
Differentiable **Ar**chitecture **S**earch : Methods exploiting the very nature of neural networks are promising. We are currently doing a lot of research towards DARTS techniques (DARTS (Liu et.al (2018)).

One problem that arise with these methods is that it does not penalize complex architectures and still need some hyperparameters to be set up.

Hedging : On the Algorithm side/AutoML

Results are not revolutionnary as we don’t need a lot of layers…





Loss graphes and associated architectures
(credits Ilyes Er Rammach)

Hedging : On the business side

The lack of dimensionality wall and the flexibility these methods bring allow us to tackle historical problems

- by including more and more sources of uncertainties
- by adding liquidity constraints, bid ask spreads, …
- by changing the performance measure

Hedging : On the business side

The lack of dimensionality wall and the flexibility these methods bring allow us to tackle historical problems

- by including more and more sources of uncertainties
- by adding liquidity constraints, bid ask spreads, …
- by changing the performance measure

A lot of study we are doing does not bother a lot with the algorithm side but more with the business side.

Hedging Strategy Computation : On the business side
Change of the performance measure

The question of the risk criteria we want to minimize is crucial.

For the hedging side, some natural criteria leads to strategy that we won't use operationally making us work even harder on the definition of the criteria we want to optimize.

We did try asymmetric risk criteria, risk reward criteria, budget target criteria, quantile hedging…

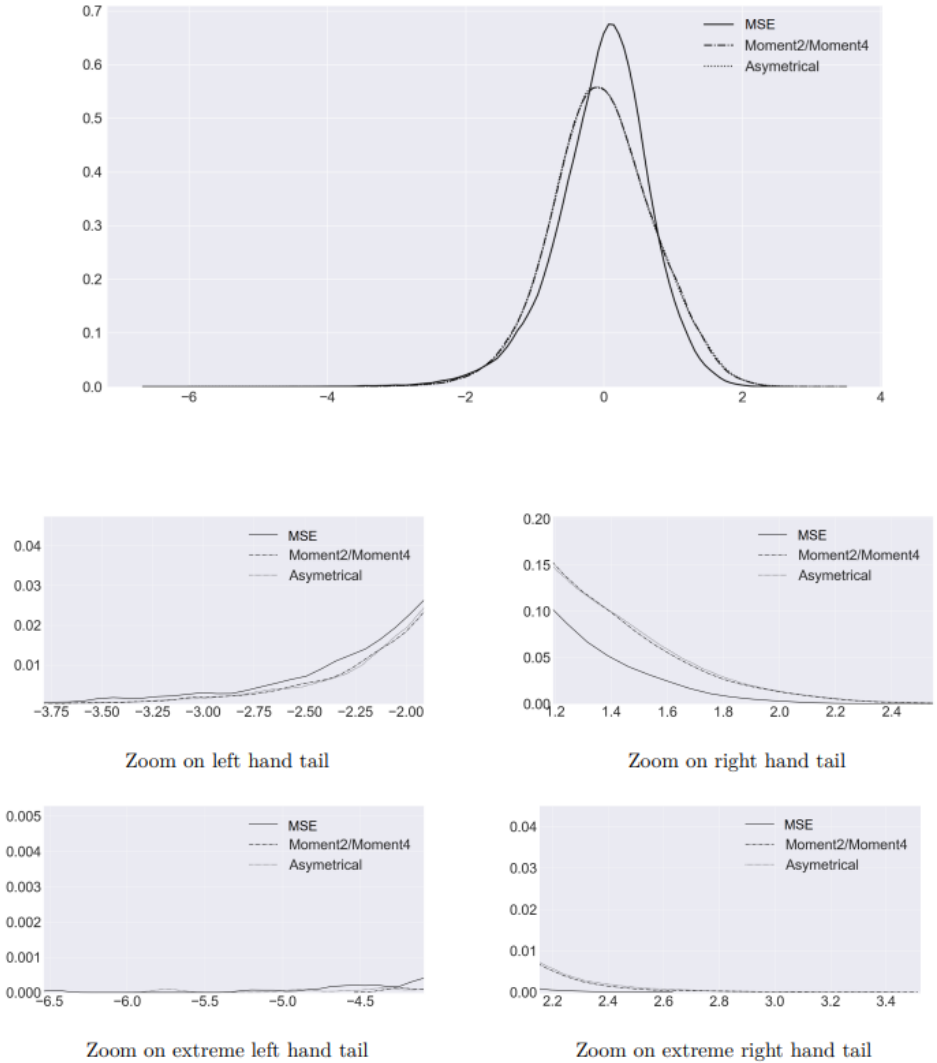


Figure. Distribution of the hedged portfolio (3 markets spread option) with symmetric and assymetric risk criteria

Hedging Strategy Computation : The confidence side/XAI

For control stuff, and strategy computation, we are at a step where it is difficult to put these neural nets algorithm in an operational way.

We still prefer deriving heuristic from algorithm outputs.

This derivation is time consuming and costly.

The initial temptation to include all available risk factors hits the wall of the explainability.

Hedging Strategy Computation : The confidence side/XAI

We have a tool called A implementing a reinforcement algorithm training on Monte Carlo scenarios of the time series of interest. Optimal controls are thus available at each time step and at each simulation.

Auditability : Our actions can be audited by the regulator. We must be able to justify taking a position on the markets in accordance with a risk policy. Finding a heuristic from the strategies found by Tool A is a long process. Can we find tools to help highlight a heuristic?

Error quantification : From what level of loss can we be confident about the controls? In some situations, very similar loss function values can give very different optimal controls.

Generative methods

From deep hedging and deep asset management to Generative methods

Deep hedging and Deep Asset Management represents a big temptation for the modelling of new kind of risk factors.

These modelling are lacking for a lot of these underlyings.

Generative methods: An expected guest

Everyone waits for Generative methods applied to time series as this would allow :

1. to rapidly take into account subit changes in market structure
2. to design a model in a short period of time

Generative methods : 2 ways to do it

Explicit Way

One way to accomplish the task is to learn *P* and to use known techniques (inversion, …) to generate the data.

Implicit Way

The task would be to directly generate data. Of course, any algorithm would learn some things related to the distribution. But the distribution estimate is not an output

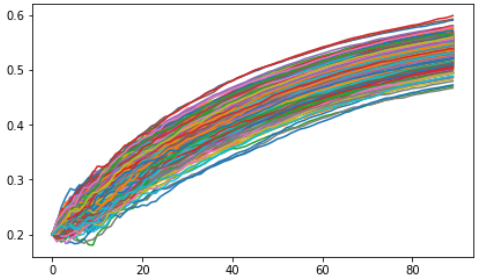
Applying GANs on time series

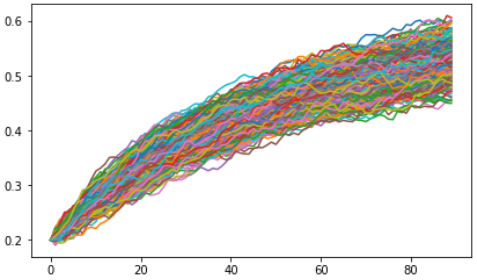
Problems

- No control on generations
- Learning marginals is not enough
- Learning joint distribution too
- Few theoretical results

+

Does not appear to work.





Generative methods applied to time series

Five methods :

- Models based on GANs (Time Series Gans (Yoon et. al (2017))
- Models based on Wasserstein & Sinkhorn (Cuturi, Genevay)
- Signature (Hao et.al, Buehler et.al (2020)
- Causal Optimal Transport (Tianlin et. al (2020))
- Model based on SDE and conditionning (Remlinger et. al (2021))

Focus on CEGEN Algorithm (Remlinger (2021))

• CEGEN is time series generator combining an Euler structure with a dedicated loss on conditional distributions

$$Y^{\theta}_{t_i+\Delta t} = Y^{\theta}_{t_i} + g^{b}_{\theta}(t_i, Y^{\theta}_{t_i})\Delta t + g^{\sigma}_{\theta}(t_i, Y^{\theta}_{t_i})Z_{t_i}$$

• A theoretical result ensuring a bound for the process parameter estimation error.
• Numerical study on synthetic and various real world data sets: accurate correlation structure up to dimension 20

Focus on CEGEN Algorithm (Remlinger (2021))

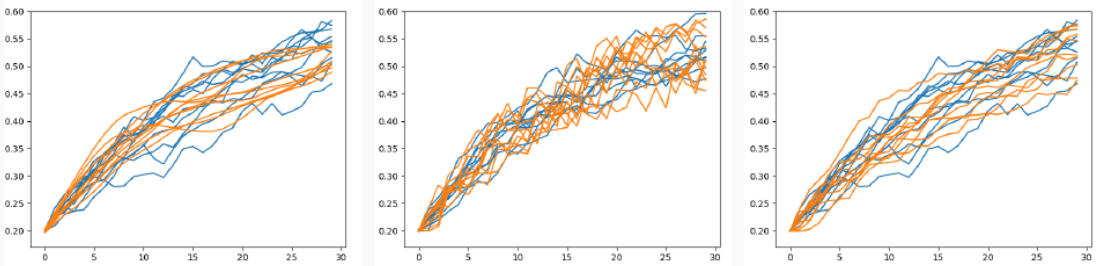
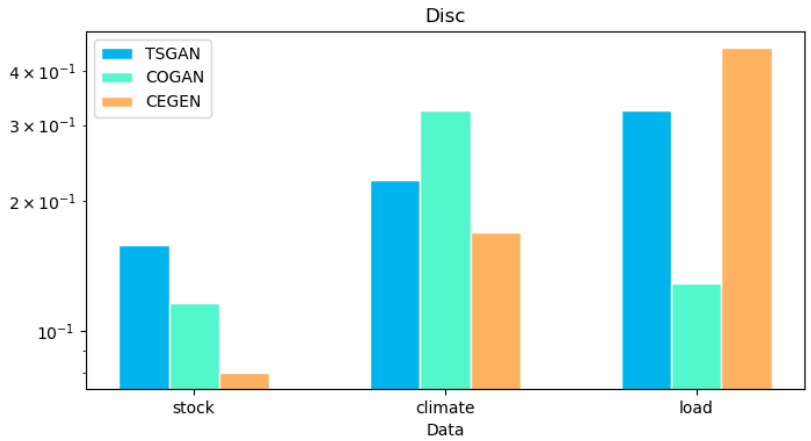


Figure 6: Ornstein-Uhlenbeck process samples (blue) and generations (orange) from COTGAN, TSGAN and CEGEN.

Table 5. **Exp. D** MSE between Monte Carlo and generated empirical correlation matrices on BS

Dim	CEGEN	EWGAN	EDGAN	TSGAN
4	**7.71e-03**	1.53e-02	5.29e-02	1.77e-01
10	**1.05e-02**	5.47e-02	2.24e-02	2.59e-01
20	**6.57e-03**	3.36e-02	1.40e-02	4.81e-01

Focus on CEGEN Algorithm (Remlinger (2021))



Discriminative scores on real data

(Parenthesis) Amount of papers

Since 1973, **42877** citations of Black-Scholes celebrated paper.

Since 2014, **34574** citations of Goodfellow et.al GAN paper.

The finance community faces an amount of research dealing with other areas far beyond what we've used to know.

Generative methods : Confidence/XAI

Generative methods is an expected guest but with it, we'd lose the control when compared to our good old MC models

- A difficulty to monitor the model behavior over time.
- A difficulty to control the model during an online training phase

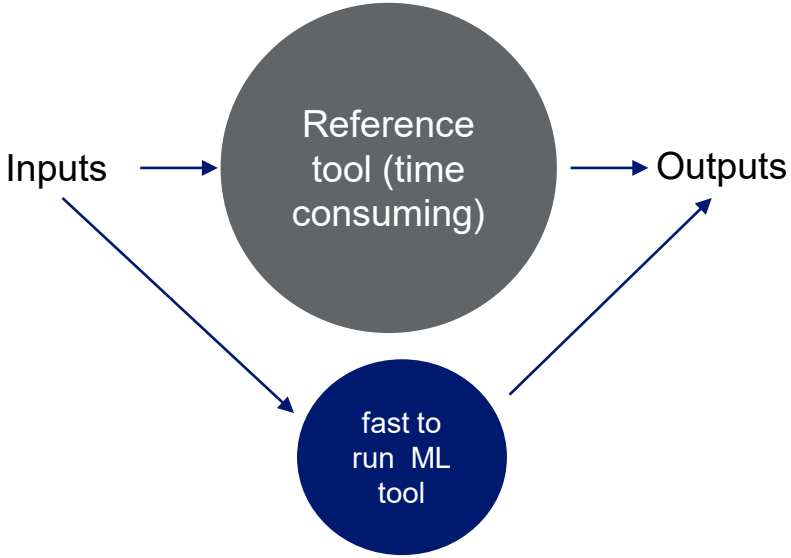
Generators should not be left free, as this could lead to damageable consequences

- Contrarily to the existing model-free approaches, it may be seen as an improvement to impose a limited set of models as done in CEGEN

Reproducing tools

Pricing & Risk evaluation : Reproducing tools

One challenge these last few years is to reproduce our traditional tools with machine learning models.



Pricing & Risk evaluation : Reproducing tools

The outcomes are

1. Much more faster pricer & Risk computation tools
2. Sensitivities according to inputs.

Questions

1. How to make historical & mirroring tool live together?
2. The design of an optimized tool may not offer as much return as it used to be.

XAI

Confidence : on the business side

How can we certify processes ?

Do the next step is to build up a reference data set?

Confidence : on the IT side

A lot of questions arise for which any answer would be welcome.

- In a tool oriented perspective, the whole unitary test, functional test framework may not be sufficient. It became even more difficult to test out things in such a world.

- In a functional oriented perspective, online training makes things even more difficult. How you testify that things continue to work?

Merci

Bibliography

Maximilien Germain, Huyên Pham, and Xavier Warin. "Approximation error analysis of some deep backward schemes for nonlinear PDEs ". In: to appear in SIAM Journal on Scientific Computing

Huyên Pham, Xavier Warin, and Maximilien Germain "Neural networks-based algorithms for stochastic control and PDEs in finance", : To appear in "Machine Learning for Financial Markets: a guide to contemporary practices, Cambridge University Press, Editors: Agostino Capponi and Charles-Albert Lehalle

BUEHLER, Hans, HORVATH, Blanka, LYONS, Terry, *et al.* Generating financial markets with signatures. *Available at SSRN*, 2020.

NI, Hao, SZPRUCH, Lukasz, WIESE, Magnus, *et al.* Conditional sig-wasserstein gans for time series generation. *arXiv preprint arXiv:2006.05421*, 2020.

Remlinger, Carl, JM, Elie.R « Conditional & adversarial generative methods for time series » (2021)
Fecamp S, JM, Warin.X « Deep learning for discrete-time hedging in incomplete markets »(Journal of computationl Finance (2020))