How the Scalar function works

the main problem of the data generated by the model is the scale of the data itself. As we can see in the following picture the series seams realistic if we just look at the trend but the magnitude of the data is totally wrong. To avoid this I've created an algorithm to force the series to the real data range.

Introduction

The scaler aims to scale the time series without changing the trend. The idea is that when we have a peack in the time series of the return of a stock the role of the peak depends by the distance of the peak itself from the second highest peak in the series. For exemple let's take the following series:

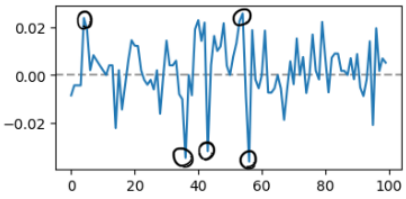


Figure 1

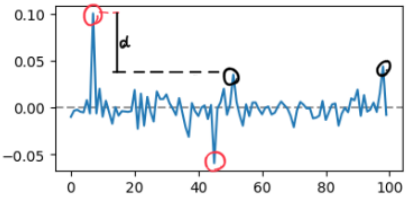


Figure 2

As we can see the in (**Figure 1**) the 2 highest peaks are not very far from each-other and that makes them lower in terms of magnitude compared to the peak we can see in the (**Figure 2**). This because a peak can have 2 different meanings in a time series:

- a strong movement in the price but still not that unusual
- a shock, due to news or particural events.

is very unlikely to have two big shoks in in the same window. Following this logic when the 1^{st} and 2^{nd} peaks in a series are close they are probably just strong movements in the price. On the other side when the 1^{st} peak in terms of magnitude is way bigger than the 2^{nd} one probably It's a shock so these types of shoks will have bigger values.

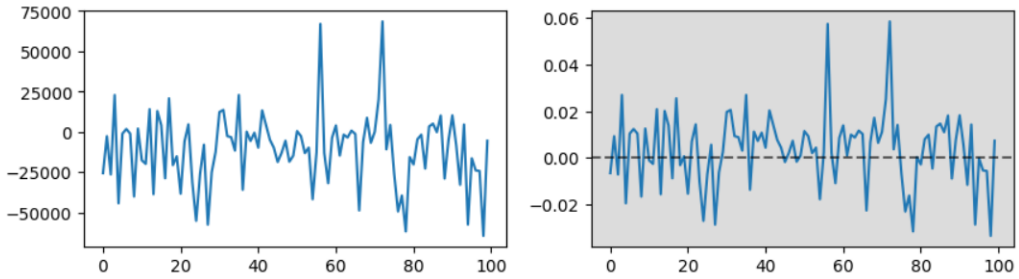
Scalar Logic

Baset on the above reasonment the Scalar aims assign to "lonley" peaks in the generated series the value that usually shocks have in the real data and assign to close peacks value that are typical of strong market movements.

Scalar Implementation

To do It the scalar compute the quantiles of the max values distribution and same thing for the distribution of the difference between the 1^{st} and 2^{nd} peaks. Then the algorith take the sample and see where the distance between the 1^{st} and 2^{nd} peaks are among the quantiles of the differences for real data. Then pick a random value from a uniform distribution where the lower and upper bonudaries are the corresponding quantiles in the max distribution. Same thing for the negative peaks.

Exemple:



In the above exemple we can see on the left the series generated by our model and on the rigth the series after the we apply the Scaler. As we can see the scaler preserves the pattern generated by our model and just rescale the data so we have realistic data.