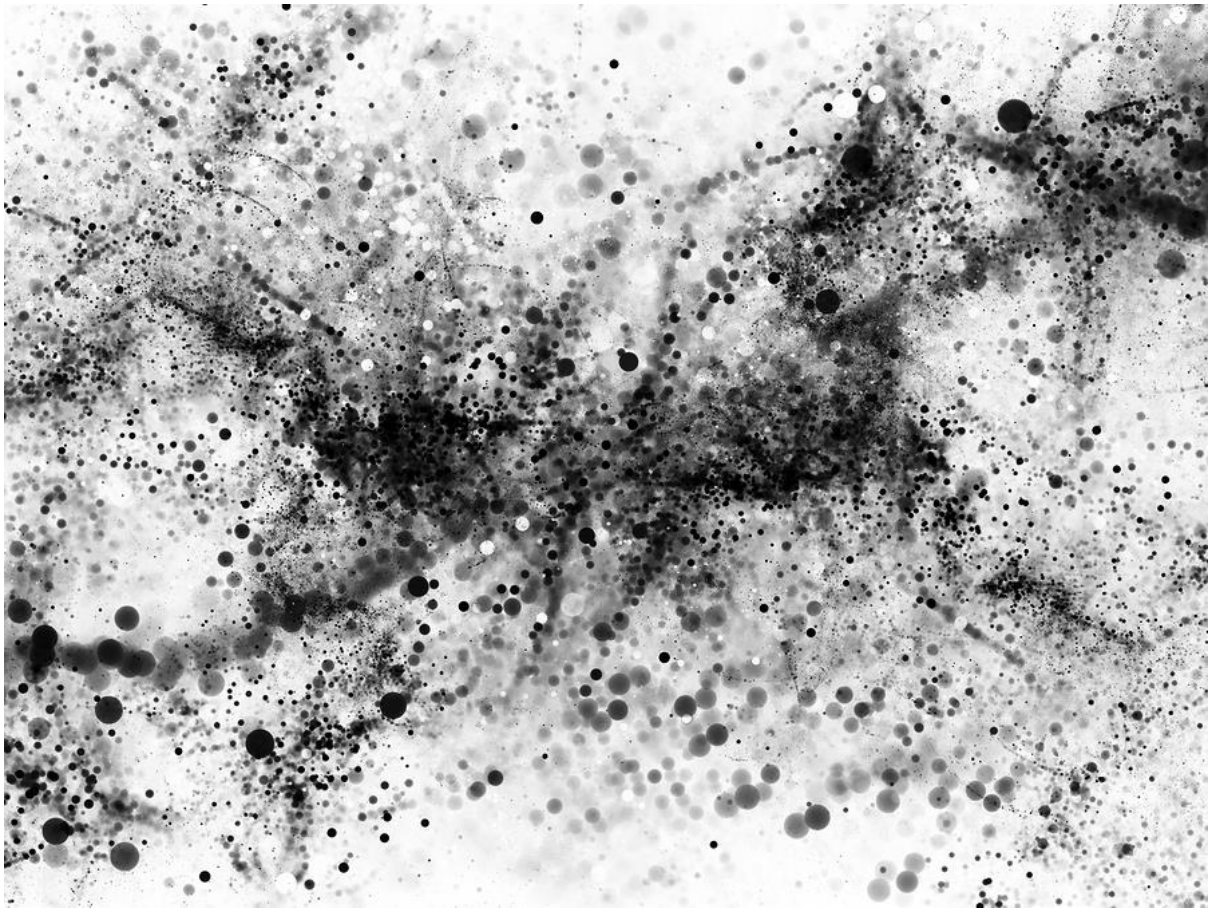
The Dangers of Linear Correlation

April 21st, 2017

"The difficulty lies not so much in developing new ideas as in escaping from old ones" – John Maynard Keynes



Correlations are present everywhere. The concept of correlation is one of the key constructs of statistics, modelling, simulation. It is used to design portfolios, to estimate risks, to perform VaR analysis, compute Probabilities of Default, etc. A correlation expresses how strongly two variables are interdependent. It is therefore of paramount importance to measure correlations correctly.

Instead of computing correlations based on traditional approaches based on variance, we use a proprietary entropy-based measure of correlation, called generalized correlation. Unlike variance which measures concentration only around the mean, the entropy-based generalized correlation takes into account the actual distribution of data in the entire domain spanned by the data in question. This allows us to provide a more realistic and correct measure of correlation.

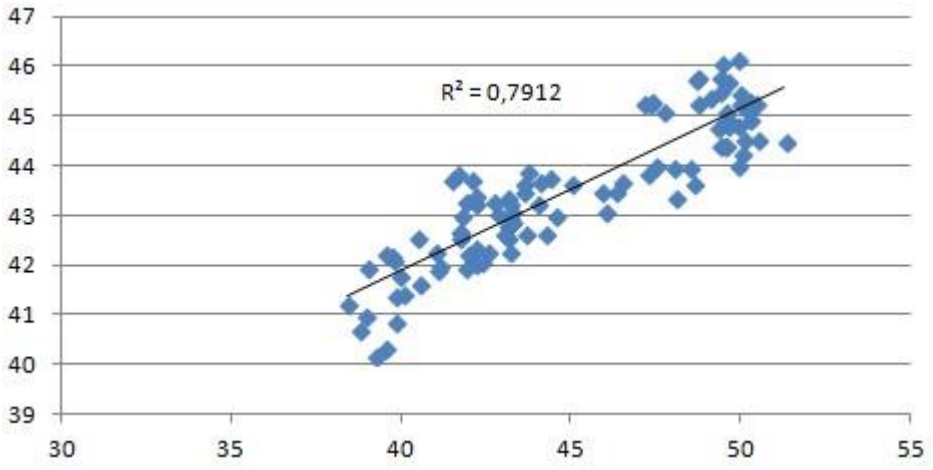
A popular and conventional correlation is the linear correlation by Pearson. Pearson's correlations function properly when applied to data which is linear in character. In cases which include data

concentrations, clustering, bifurcations or other forms of discontinuity, applying linear correlations is outright wrong. The results of linear correlation analysis may in fact provide outcomes which can induce unjustified optimism and distort significantly any risk-type calculations.

The surprising fact is that this shortcoming of linear correlations is widely known and yet neglected by the mainstream of fund managers and analysts. Traditional models which are used to compute risk or degree of asset diversification in investment portfolios may be easily proved to be incorrect.

Two simple examples are illustrated below.

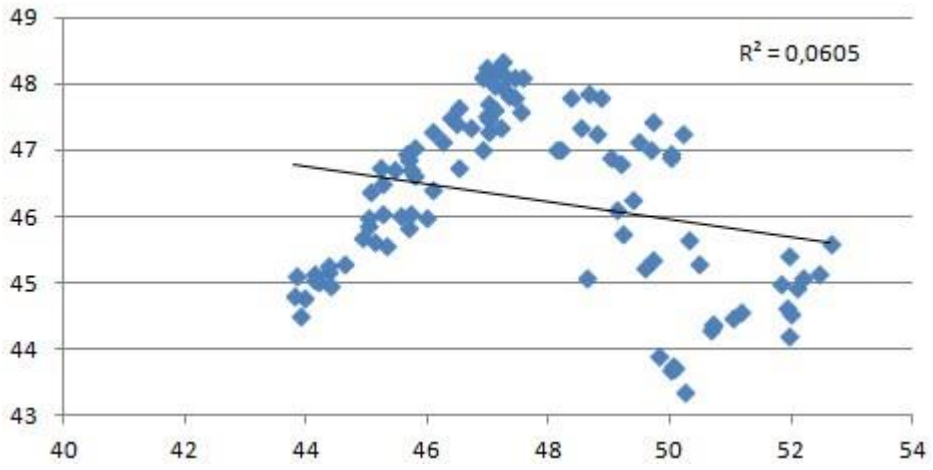
Figure 1. Stock 1 vs Stock 2



Stock 1 vs. Stock 2 Generalized Correlation is 76%, while linear is 89%.

Source: Optimum Complexity

Figure 2. Stock 3 vs Stock 4.



Stock 1 vs. Stock 2 Generalized Correlation is 76%, while linear is 24%.

Source: Optimum Complexity

The plot shows a strongly non-linear situation. Linear correlations in this case are not applicable. In this case, linear correlations would suggest that the variables are independent. Difference is 52%!

Even in moderate size investment portfolios there are thousands of interdependencies. If these are analysed using conventional correlations to estimate exposure, expected returns or degree of diversification there exists a concrete possibility that these analyses are simply wrong. Imagine what a 10% error in correlation can do in a covariance matrix-based portfolio design.

"The difficulty lies not so much in developing new ideas as in escaping from old ones" – John Maynard Keynes

---

Optimum Complexity Ltd.
1 Mayfair Pl, W1J 8AJ
London, United Kingdom

E-mail: dc@optimumcomplexity.com

Website: www.optimumcomplexity.com

Phone: +44 7808 763 348