Financial Data Analysis Using The Informational Energy Unilateral Dependency Measure

Angel Caţaron
Electronics and Computers Department
Transilvania University of Braşov
Romania
Email: cataron@unitbv.ro

Răzvan Andonie
Computer Science Department
Central Washington University
Ellensburg, USA
and
Electronics and Computers Department
Transilvania University of Braşov
Romania
Email: andonie@cwu.edu

Yvonne Chueh
Department of Mathematics
Central Washington University
Ellensburg, USA
Email: chueh@cwu.edu

***Abstract*—Our research area is the unilateral dependency (UD) analysis of non-linear relationships within pairs of simultaneous data. The application is in financial analysis, using the data reported by Kodak and Apple for the period of 1999–2014. We compute and analyze the UD between Kodak's and Apple's financial time series in order to understand how they influence each other over their company assets and liabilities. We also analyze within each of the two companies the UD between assets and liabilities. Our formal approach is based on the informational energy UD measure derived by us in previous work. This measure is estimated here from available sample data, using a non-parametric asymptotically unbiased and consistent kNN estimator.**

I. Introduction

Most of us intuitively recognize causal relationships in our daily lives. We make statements like "X causes Y", "Y depends on X", or "X and Y are correlated". We have to observe from the very beginning that there are nuances in these statements. A common error is the confusion between statistical correlation and causation. An occurrence can cause another occurrence (such as smoking causes lung cancer), or it can correlate to another occurrence (such as smoking is correlated to alcoholism). If one occurrence causes another, then they are most certainly correlated. The converse is not necessarily true.

Causal analysis is not merely a search for correlations, but an investigation of cause and effect relationships. Judea Pearl makes a harsh distinction between the causal and statistical relationships. According to him, causal analysis goes one step further than statistical analysis, since it aims to infer not only the likelihood of events under static conditions, but also the dynamics of events under *changing conditions* [1].

Practically, it is very difficult to establish causality between two correlated events. In contrast, it is relatively easy to establish a statistically significant correlation. To prove causation, a controlled experiment must be performed. The standard way to generate causal evidence is to use randomized experiments, but this is often not feasible in real-life applications.

The logic of causal analysis and the problems involved in establishing causal linkages are discussed at length in [2]. These authors, continuing the work of epistemologists such as John Stuart Mill, identify three key criteria for inferring a cause and effect relationship: (1) the cause preceded the effect, (2) the cause was related to the effect, and (3) we can find no plausible alternative explanation for the effect other than the cause.

Application areas of causal relationships include finance, medicine, and reliability engineering. For example, in econometrics, we are interested in the correlation (and beyond that, causality as well) between two time series such as a market/bench index and an individual stock/ETF products. Most investors in the stock market consider various indexes to be important sources of basic information that can be used to analyze and predict the market perspectives.

The Granger[1] causality test [3] is a statistical hypothesis test for determining whether one time series is useful in forecasting another. According to Granger, causality in economics could be reflected by measuring the ability of predicting the future values of a time series using past values of another time series. Some econometricians assert that the Granger test finds only "predictive causality" [4]. The Granger test is based on linear regression modeling of stochastic processes. More complex extensions to nonlinear cases exist, but these extensions are more difficult to apply in practice. A comparison between different causality tests is provided in [5], where the causality between Private Consumption and GDP in the USA and Mexico during the period 1960-2002 is determined.

In contrast to causality, the statistical correlation is a well-established concept. When studying the correlation between two financial indexes, we may use a standard bilateral statistical measure, like the mutual information (MI). However, we may also use a unilateral measure, like the transfer entropy, which measures the directionality of a variable with respect to time was introduced by Schreiber [6]. Based on the transfer entropy concept, Kwon and Yang [7] found that the amount of information flow from index to stock is larger than from stock to index. It indicates that the market index plays a role of major driving force to individual stock. Interestingly, such asymmetry

[1]Clive Granger, recipient of the 2003 Nobel Prize in Economics.

occurs with identical direction to every market from mature to emerging market. However, the strength of the asymmetry in mature market is higher than it in emerging market. Hence, there must be asymmetric information flow between an index and a stock.

The focus of our work is the inference of non-linear causal relationships within pairs of simultaneous data. Here is a simple example to demonstrate the simultaneity problem [8]: Hiring more police-officers (X) should reduce crime (Y). However, it is also possible too that when crime goes up, cities hire more police officers. The original Granger causality test cannot be used here, since it does not capture simultaneous causal relationships. Therefore, we have to find a different approach. Let us also observe that we do not comply here with the epistemologic time asynchronicity criterion "the cause preceded the effect".

Our attempt is to not only to calculate unilateral dependencies between the analyzed time-series, but to go beyond this and create the premises for a causal interpretation. Although, in general, statistical analysis cannot distinguish genuine causation from spurious covariation in every conceivable case, this is still possible in many cases [1]. Actually, according to the results of the ChaLearn cause-effect pair challenge [9], causal inference can be successfully addressed as a supervised machine learning approach. Recently, causal relationships were inferred using the Markov Blankets of two variables causally connected, in the context of supervised learning [10].

We present a machine learning application in financial data analysis, where we attempt an inference of causal relationships. We analyze the UD between Kodak's and Apple's assets and liabilities, reported by the two companies for the period of 1999–2014, in order to understand how they would influence each other. We use a UD measure based on Onicescu's information energy (IE) [11]. The IE can be interpreted as a measure of average certainty. In previous work, we have introduced a non-parametric asymptotically unbiased and consistent estimator of the IE. Our method can be applied to both continuous and discrete data, meaning that we can use it both in classification and regression algorithms. Based on the IE, we have introduced [12], [13] a UD measure between random variables and we showed how to estimate this UD measure from an available sample set, using the kNN estimator.

The paper is organized as follows. First, we will review (Section II) the properties of IE and the kNN method. Section III describes the approximation method for our unilateral information measure. The financial application is presented in Section IV. Section V has the final remarks.

II. Background and Previous Work

A. Onicescu's Informational Energy and The Unilateral Dependency Measure

Generally, information measures refer to uncertainty. However, information measures can also refer to certainty, and probability can be considered as a measure of certainty. More general, any monotonically growing and continuous probability function can be considered as a measure of certainty. For instance, Onicescu's IE was interpreted by several authors as a measure of expected commonness, a measure of average certainty, or as a measure of concentration.

For a continuous random variable X with probability density function $f(x)$, the IE is [14], [15]:

$$IE(X) = \int_{-\infty}^{+\infty} f^2(x)dx \tag{1}$$

A UD measure between random variables X and Y was defined in [11]:

$$o(X,Y) = IE(X|Y) - IE(X) \tag{2}$$

where:

$$IE(X|Y) = \int_{-\infty}^{+\infty} f(y)IE(X|y)dy$$
$$= \int_{-\infty}^{+\infty} f(y) \int_{-\infty}^{+\infty} f(x|y)^2 dx\, dy.$$

The conditional probability density function can be written as the ratio between the joint density function and the marginal density function: $f(x|y) = f(x,y)/f(y)$. Then:

$$IE(X|Y) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x,y)f(x|y)dy\, dx$$

and

$$o(X,Y) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x,y)f(x|y)dy\, dx - \int_{-\infty}^{+\infty} f^2(x)dx.$$

The measure $o(X,Y)$ quantifies the UD characterizing X with respect to Y and corresponds to the amount of information detained by Y about X. There is an obvious analogy between $o(X,Y)$ and $MI(X,Y)$, since both measure the same phenomenon. However, the MI is a symmetric, not a unilateral measure: $MI(X,Y) = MI(Y,X)$.

Our goal is to approximate $o(X,Y)$ from the available data samples, using the kNN method. As a a first step, we will approximate the IE.

B. The nearest neighbor method

The kNN estimators represent an attempt to adapt the amount of smoothing to the "local" density of data. The degree of smoothing is controlled by an integer k, chosen to be considerably smaller than the sample size. We define the distance $d_j(x_i)$ between two points on the line as $|x_i - x_j|$, and for each x_i we define the distances $d_1(x_i) \leq d_2(x_i) \leq \ldots \leq d_n(x_i)$, arranged in ascending order, from x_i to the points of the sample.

The kNN density estimate $f(x_i)$ is defined by [16]:

$$\hat{f}(x_i) = \frac{k}{2nd_k(x_i)}$$

Leonenko *et al.* [17] introduced an asymptotic unbiased and consistent estimator of the entropy in a multidimensional space. When the sample points are very close one to each other, small fluctuations in their distances produce high fluctuations of the entropy estimator. To overcome this problem, Singh *et al.* [18] defined a kNN non-parametrical estimator of the entropy, based on the k-th nearest neighbor distance between n points in a sample, where k is a fixed parameter and $k \leq n-1$. Later, other authors have applied kNN estimators to approximate the MI from data samples [19], [20].

According to [21], the kNN MI estimation outperforms histogram methods. kNN works well if the value of k is optimally chosen. However, there is no model selection method for determining the number of nearest neighbors k. This is a limitation of the kNN estimation.

We are now ready to introduce our kNN method for the IE approximation.

C. Estimation of the Informational Energy

The IE can be easily computed if the data sample is extracted from known distributions. When the underlying distribution of data sample is unknown, the IE has to be estimated. More formally, our goal is to estimate formula (1) from a random sample x_1, x_2, ..., x_n of n d-dimensional observations from a distribution with the unknown probability density $f(x)$. This problem is even more difficult if the number of available points is small.

The IE is the average of $f(x)$, therefore we have to estimate $f(x)$. The n observations from our samples have the same probability $\frac{1}{n}$. A convenient estimator of the IE is:

$$\hat{IE}_k^{(n)}(X) = \frac{1}{n}\sum_{i=1}^{n}\hat{f}(x_i). \tag{3}$$

We will determine first the probability density $P_{ik}(\epsilon)$ of the random distance $R_{i,k,n}$ between a fixed point x_i and its k^{th} nearest neighbor from the remaining $n-1$ points. The probability $P_{ik}(\epsilon)d\epsilon$ of the k^{th} nearest neighbor to be within distance $R_{i,k,n} \in [\epsilon, \epsilon + d\epsilon]$ from x_i, $k-1$ points at a smaller distance, and $n-k-1$ at a larger distance, can be expressed in terms of the trinomial formula [20]:

$$P_{ik}(\epsilon)d\epsilon = \frac{(n-1)!}{1!(k-1)!(n-k-1)!}dp_i(\epsilon)p_i^{k-1}(1-p_i)^{n-k-1},$$

where $p_i(\epsilon) = \int_{\|x-x_i\|<\epsilon} f(x)dx$ is the mass of the ϵ-ball centered at x_i and $\int P_{ik}(\epsilon)d\epsilon = 1$.

We can express the expected value of $p_i(\epsilon)$ using the probability mass function of the trinomial distribution:

$$E_{P_{ik}(\epsilon)}(p_i(\epsilon)) = \int_0^{\infty} P_{ik}(\epsilon)p_i(\epsilon)d\epsilon$$

$$= k\binom{n-1}{k}\int_0^1 p^{k-1}(1-p)^{n-k-1}pdp$$

$$= k\binom{n-1}{k}\int_0^1 p^{(k+1)-1}(1-p)^{(n-k)-1}dp.$$

This equality can be reformulated using the *Beta* function:

$$B(m,n) = \int_0^1 x^{m-1}(1-x)^{n-1} = \frac{\Gamma(m)\Gamma(n)}{\Gamma(m+n)}.$$

We obtain:

$$E_{P_{ik}(\epsilon)}(p_i(\epsilon)) = k\binom{n-1}{k}\frac{\Gamma(k+1)\Gamma(n-k)}{\Gamma(n+1)}$$

$$= k\frac{(n-1)!}{(n-k-1)!k!}\frac{k!(n-k-1)!}{n!},$$

which can be rewritten as:

$$E_{P_{ik}(\epsilon)}(p_i(\epsilon)) = \frac{k}{n}. \tag{4}$$

On the other hand, assuming that $f(x)$ is almost constant in the entire ϵ-ball around x_i, we have:

$$p_i(\epsilon) \approx V_1 R_{i,k,n}^d f(x_i),$$

where we denote the volume of the ball of radius $\rho_{r,n}$ in a d-dimensional space by:

$$V_{\rho_{r,n}} = V_1\rho_{r,n}^d = \frac{\pi^{\frac{p}{2}}\rho_{r,n}^d}{\Gamma(\frac{p}{2}+1)}.$$

V_1 is the volume of the unit ball and $R_{i,k,n}$ is the Euclidean distance between the reference point x_i and its k^{th} nearest neighbor. This means that $V_1 R_{i,k,n}^d$ is the volume of the d-dimensional ball of radius $R_{i,k,n}$.

We obtain the expected value of $p_i(\epsilon)$:

$$E(p_i(\epsilon)) = E(V_1 R_{i,k,n}^d f(x_i)) = V_1 R_{i,k,n}^d \hat{f}(x_i). \tag{5}$$

Equations (4) and (5) both estimate $E(p_i(\epsilon))$. Their results are approximately equal:

$$V_1 R_{i,k,n}^d \hat{f}(x_i) = \frac{k}{n},$$

That is:

$$\hat{f}(x_i) = \frac{k}{nV_1 R_{i,k,n}^d}, i = 1...n.$$

This is the estimate of the probability density function. By substituting $\hat{f}(x_i)$ in formula (3), we finally obtain the following IE approximation:

$$\hat{IE}_k^{(n)}(f) = \frac{1}{n}\sum_{i=1}^{n}\frac{k}{nV_1 R_{i,k,n}^d}. \tag{6}$$

Consistency of an estimator means that as the sample size gets large the estimate gets closer and closer to the true value of the parameter. Unbiasedness is a statement about the expected value of the sampling distribution of the estimator. The ideal situation, of course, is to have an unbiased consistent estimator. This may be very difficult to achieve.

Yet unbiasedness is desirable but just a little bias is permitted, as long as the estimator converges to unbiased. Therefore, an asymptotically unbiased consistent estimator may be acceptable. Our estimator has the following important property, given here without proof.

Theorem 1. *The informational energy estimator $\hat{IE}_k^{(n)}(f)$ is asymptotically unbiased and consistent.*

Therefore, the probability of the estimator being arbitrarily close to its true value converges to one, as the sample size increases. The $\hat{IE}_k^{(n)}(f)$ estimator is computationally intensive, but it is a "good" estimator because it is asymptotically unbiased and consistent. These are nice properties that other approximation methods (for instance, the histogram method) do not share.

III. The kNN $o(X,Y)$ estimator

Our goal is to infer $o(X,Y)$ from the random sample x_1, x_2, ..., x_n. We will deduct the kNN estimator for $o(X,Y)$.

First, we substitute $\widehat{IE}_k^{(n)}(X)$ from eq. (6) in eq. (2):

$$\hat{o}(X,Y) = \widehat{IE}_k^{(n)}(X|Y) - \widehat{IE}_k^{(n)}(X)$$

where:

$$\widehat{IE}_k^{(n)}(X|Y) = \sum_{j=1}^{m} \hat{f}(y_j)\widehat{IE}_k^{(n)}(X|y_j) \tag{7}$$

and

$$\widehat{IE}_k^{(n)}(X) = \frac{1}{n}\sum_{i=1}^{n} \frac{k_1}{nV_{1(X)}R_i^{d_1}}$$

is an adaptation of eq. (6).

We have:

$$\widehat{IE}_k^{(n)}(X|y_j) = \frac{1}{n}\sum_{i=1}^{n} \hat{f}(x_i|y_j) = \frac{1}{n}\sum_{i=1}^{n} \frac{\hat{f}(x_i,y_j)}{\hat{f}(y_j)} \tag{8}$$

and from (7) and (8) we can write:

$$\begin{aligned}\widehat{IE}_k^{(n)}(X|Y) &= \sum_{j=1}^{m} \hat{f}(y_j)\frac{1}{n}\sum_{i=1}^{n} \frac{\hat{f}(x_i,y_j)}{\hat{f}(y_j)} \\ &= \frac{1}{n}\sum_{j=1}^{m}\sum_{i=1}^{n} \hat{f}(x_i,y_j).\end{aligned}$$

The estimate of $f(x_i,y_j)$ can be obtained as

$$\hat{f}(x_i,y_j) = \frac{k_2}{pV_{1(X,Y)}R_{i,j}^{d_2}},$$

where p is the number of (x_i,y_j) pairs.

Now we can re-write eq. (7):

$$\widehat{IE}_k^{(n)}(X|Y) = \frac{k_2}{npV_{1(X,Y)}}\sum_{j=1}^{m}\sum_{i=1}^{n} \frac{1}{R_{i,j}^{d_2}}.$$

where R_i is the Euclidean distance between the reference point x_i and its k_1^{th} nearest neighbor, when the points are drawn from the one-dimensional probability distribution $f(x)$: $R_i = \|x_i - x_{i,k_1}\|$. Similarly, R_j is the Euclidean distance between the reference point y_j and its k_1^{th} nearest neighbor, when the points are drawn from the one-dimensional probability distribution $f(Y)$: $R_j = \|y_j - y_{j,k_1}\|$. Then, R_{ij} is the Euclidean distance between the reference point (x_i,y_j) and its k_2^{th} nearest neighbor, when the points are drawn from the joint probability distribution $f(X,Y)$: $R_{ij} = \sqrt{(x_{ij} - x_{ij,k_2})^2 + (y_{ij} - y_{ij,k_2})^2}$.

The estimate of $o(X,Y)$ is:

$$\begin{aligned}\hat{o}(X,Y) = &\frac{k_2}{npV_{1(X,Y)}}\sum_{j=1}^{m}\sum_{i=1}^{n} \frac{1}{R_{i,j}^{d_2}} \\ &-\frac{k_1}{n^2V_{1(X)}}\sum_{i=1}^{n} \frac{1}{R_i^{d_1}}.\end{aligned} \tag{9}$$

Although we do not have a general method to set the nearest neighbor parameter, Silverman [16] suggests that an optimal choice of k is proportional to $n^{4/(d+4)}$. In our case, the optimal values of k_1 and k_2 may not be equal, because the these two parameters refer to different samples.

IV. An Application in Financial Data Analysis

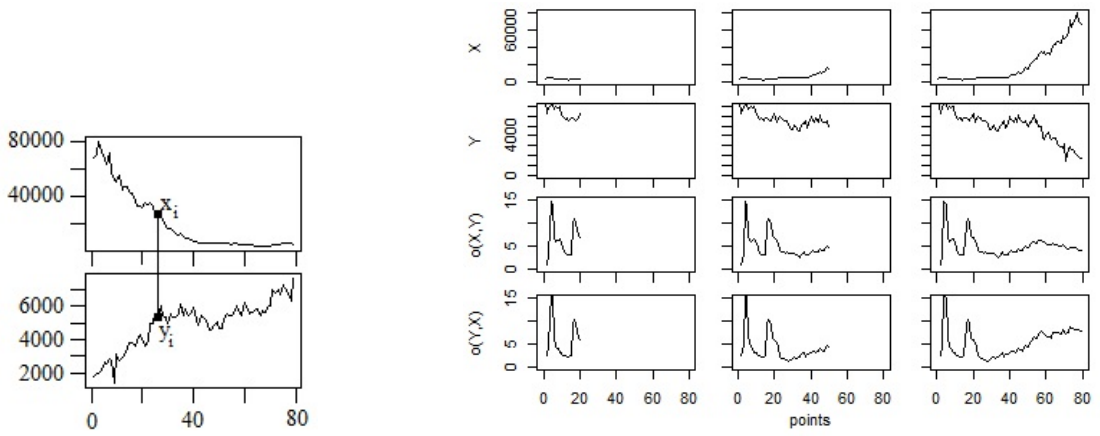
In the following, we will apply these general concepts to a real-world financial analysis. We choose assets and liabilities for their intricate interaction and fundamental importance.

A. The financial data

Apple[2] is the largest publicly traded corporation in the world by market capitalization, with an estimated market capitalization of $446 billion by January 2014. As of June 2014, Apple maintains 425 retail stores in fourteen countries as well as the online Apple Store and iTunes Store, the latter of which is the world's largest music retailer.

Eastman Kodak Company[3], commonly known as Kodak, is a company focused on imaging solutions and services for businesses. In January 2012, Kodak filed for Chapter 11 bankruptcy protection in the United States District Court. In August 2012, Kodak announced the intention to sell its photographic film (excluding motion picture film), commercial scanners and kiosk operations as a measure to emerge from bankruptcy. In January 2013, the Court approved financing for the company to emerge from bankruptcy by mid-2013. On September 3, 2013, Kodak emerged from bankruptcy having shed its large legacy liabilities and exited several businesses.

[2] http://en.m.wikipedia.org/wiki/Apple_Computer
[3] http://en.m.wikipedia.org/wiki/Eastman_Kodak



(a) (b)

Fig. 1. (a) Synchronous values from two time series denoted by X and Y produce a joint observation (x_i, y_i). On the horizontal axis we represented the indexes of data points, while the vertical axis displays the scale of the time series values; (b) The evolution over time of the dependency measure shows how the graphs of $o(X, Y)$ and $o(Y, X)$ progress when the sample size increases. Apple current assets are represented by X and Kodak current assets are represented by Y. The horizontal axis denotes the sample size and the vertical axis of the series X and Y correspond to the value of Apple and Kodak current assets in US dollars.

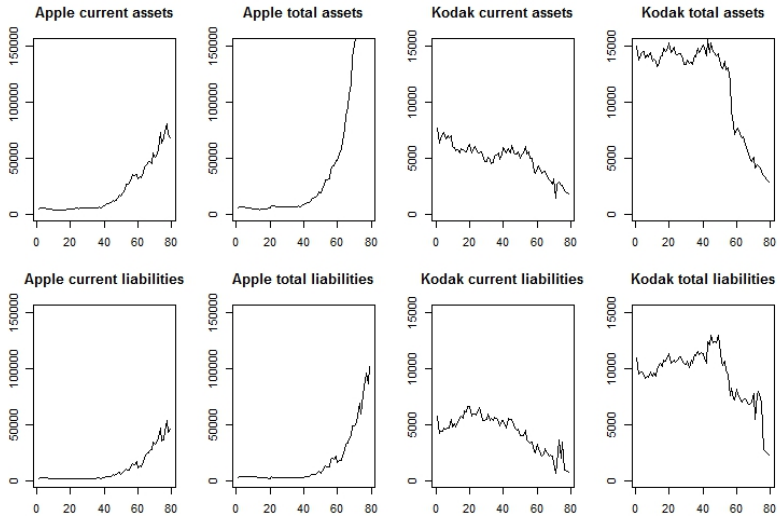


Fig. 2. Apple and Kodak assets and liabilities. Time series plot shows the data points indexes on the horizontal axis and the value in US dollars on the vertical axis. Rising company shows rising assets and liabilities as well. Falling company demonstrates the falling trend of assets and liabilities.

Personalized Imaging and Document Imaging are now part of Kodak Alaris, a separate company owned by the U.K.-based Kodak Pension Plan.

Our application interest lies on the financial insight behind these two rising and ever-falling legacy companies in the US. Before looking into their financial analytics, we aim to compute and analyze the dynamics of their unilateral dependencies with respect to liabilities and assets. This way, we may discover some similarities and differences.

We first define the basic terms (i.e., assets and liabilities) for the non-financial readers, using standard definitions from Investopedia[4].

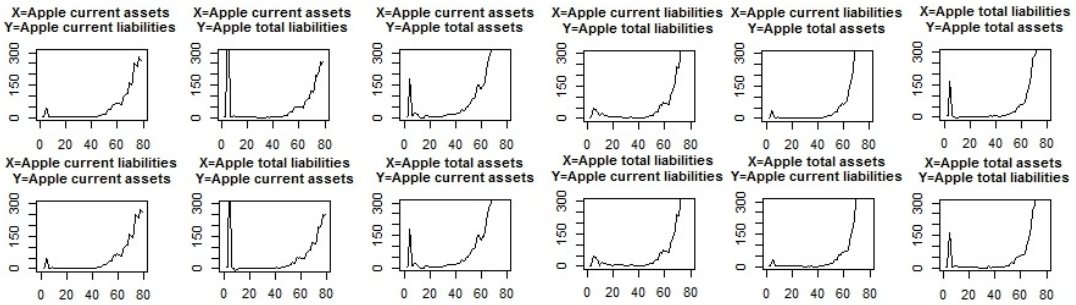
Assets are defined as:

1) A resource with economic value that an individual, corporation or country owns or controls with the expectation that it will provide future benefit.
2) A balance sheet item representing what a firm owns.

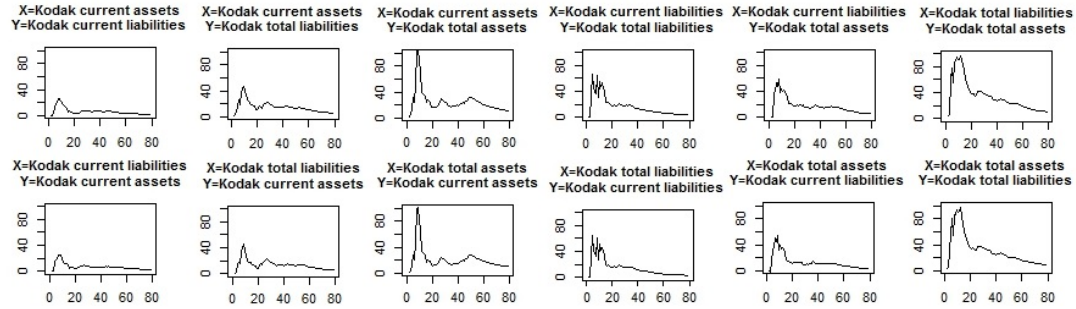
To help understand assets, they can be explained further as:

1) Assets are bought to increase the value of a firm or benefit the firm's operations. You can think of an asset as something that can generate cash flow, regardless of whether it's a company's manufacturing equipment or an individual's rental apartment.
2) In the context of accounting, assets are either current or fixed (non-current). Current means that the asset will be consumed within one year. Generally, this includes things like cash, accounts receivable and inventory. Fixed assets

[4]http://www.investopedia.com/



(a) Apple assets and liabilities.



(b) Kodak assets and liabilities.

Fig. 3. Unilateral dependency computed from the financial data of Apple and Kodak. The graphs represent the value of $o(X,Y)$, where X and Y were specified above each graph. The horizontal axis display the number of historical sample points used to obtain the corresponding value of $o(X,Y)$. (a) Apple assets and liabilities showed strong dependence and increasing coordination in management. (b) In contrast, it seems like Kodak's assets and liabilities diverge; they do not depend on each other proven by the values which decrease over time.

are those that are expected to keep providing benefit for more than one year, such as equipment, buildings and real estate.

As an important counter part of assets, *liabilities* are defined as "A company's legal debts or obligations that arise during the course of business operations. Liabilities are settled over time through the transfer of economic benefits including money, goods or services", and explained further as follows: "Recorded on the balance sheet (right side), liabilities include loans, accounts' payable, mortgages, deferred revenues and accrued expenses. Liabilities are a vital aspect of a company's operations because they are used to finance operations and pay for large expansions. They can also make transactions between businesses more efficient. For example, the outstanding money that a company owes to its suppliers would be considered a liability".

Outside of accounting and finance this term simply refers to any money or service that is currently owed to another party. One form of liability, for example, would be the property taxes that a homeowner owes to the municipal government.

Current liabilities are debts payable within one year, while long-term liabilities are debts payable over a longer period.

Asset/Liability Management, as known as *surplus management*, is a technique companies employ in coordinating the management of assets and liabilities so that an adequate return may be earned. By managing a company's assets and liabilities, executives are able to influence net earnings, which may translate into increased stock prices.

B. Unilateral interaction between financial data series

When studying the interaction between two financial random variables, why is a UD measure useful? Let us consider two experiments characterized respectively by the random variables X and Y. The experiments run simultaneously and interact probabilistically. Our question lies on whether X variable influences Y probabilistically more than the vice versa. Thus, we would like to examine the difference in the dependence of $X|Y$ (i.e., X depends on Y) and $Y|X$ (i.e., Y depends on X). While the correlation quantifies linear dependency and MI describes the degree of interdependence between two random variables, the UD measure $o(X,Y)$ helps us understand which random variable, either X or Y, has a stronger influence on the other one.

When the data is acquired from the real-world, our kNN estimator of $o(X,Y)$ adapts incrementally to the incoming X and Y values, which are considered here as data streams or time-series. We are interested in the evolution of the $o(X,Y)$ over time because this may tell us how the influence of X over Y and vice versa changes when the amount of historical data is higher. In this case, we consider the interaction between X and Y as a dynamic process described by unilateral dependencies.

In our application, we compute the UD of assets and liabilities for the Apple Computers and Eastman Kodak Company, respectively, using all their company's officially filed data in 10Q, in order to find the quantitative evidence of assets *drive/influence* company liabilities, or vice versa. We hope to determine whether a financial quantify influence another, and to what extent using the UD measure.

The IE and UD measures both help understand the two

random variables. The estimation becomes more precise when the number of observations continue to increase. When we analyze two time series, we consider each of them as sets of n observations x_i and y_i, $i = 1...n$, of the random variables X and Y. Our goal is to find synchronous relationships between the two time series which have been considered as two samples of the two random values X and Y. A point x_i from the first series is paired with a point y_i from the second one, producing a joint observation (Fig. 1a). Therefore, from two time series with n points each, we obtain n pairs (x_i, y_i), $i = 1...n$. The $o(X, Y)$ measure between X and Y can be estimated with formula (9), based on X and (X, Y), while $o(Y, X)$ can be estimated with the same formula based on Y and (Y, X). In general, the accuracy of the $o(X, Y)$ estimator increases when more values of $X|Y$ are available for each value of Y. We note that in our experiment this condition cannot be fulfilled. Each observation x_i corresponds to exactly one observation y_i, and the set of pairs (X, Y) corresponds to the set of pairs (Y, X), since for each observation x_i exactly one observation y_i is used.

We study the evolution in time of the UD measures. The values $o(X, Y)$ and $o(Y, X)$ can be estimated at the moment t_m using the entire history of the past m observations of x_i and y_i between the initial moment t_0 and t_m. A new set of observations x_{m+1} and y_{m+1} allows us to re-estimate $o(X, Y)$ and $o(Y, X)$, as illustrated by Fig. 1b. In the three columns of this figure, we present the raw signals X and Y after reading 20, 50 and 80 data points, as well as the evolution of $o(X, Y)$ and $o(Y, X)$ as it is computed from the history.

C. Experiments and results

We conduct two series of analyses. First, we analyze the UD between each company's assets and liabilities. Second, we look at the unilateral dependencies between the two companies in terms of assets and liabilities.

1) Unilateral dependencies between assets and liabilities: The series of Apple and Kodak assets and liabilities (Fig. 2) were used to study the evolution over time of the unilateral dependencies (Figs. 3a and 3b).

We can observe a slightly higher influence of liabilities over assets than vice versa. This is in accordance to the *liability-driven* investment strategy (LDI), which is based on the cash flows needed to fund future liabilities. LDI differs from a *benchmark-driven* strategy, which is based on achieving better returns than an external index such as the S&P 500 or a combination of indices that invest in the same types of asset classes. LDI is designed for situations where future liabilities can be predicted with some degree of accuracy [22].

2) Unilateral dependencies between Kodak's and Apple's financial reports: Having the series *Current assets* and *Total assets* from Apple and Kodak at hand, we analyze how they influence each other by taking all possible pairs, as represented in Fig. 4. Comparing $o(X, Y)$ with $o(Y, X)$, we could find an interesting evolution of the dependency measure when the two time series are represented by *Kodak current assets* and *Apple total assets*. Although, in general, for the studied four series, the graphs representing the dependency measure are almost similar, when we look at $o(X, Y)$ versus $o(Y, X)$, *Kodak current assets* seems to be strongly depending on *Apple total assets* when recent data is taken into consideration, while the reverse dependency is weak. When looking at *Kodak current assets* and *Apple current assets*, one can deduce that the trend is that Kodak increases its dependency on Apple.

Kodak sold many of its patents for approximately $525,000,000 to a group of companies (including Apple, Google, Facebook, Amazon, Microsoft, Samsung, Adobe Systems and HTC) under the name Intellectual Ventures and RPX Corporation[5]. This sale announced in the end of 2012 was a step toward emerging from bankruptcy. Looking at the Kodak financial data we notice a decrease of the amount of total liabilities during 2013 and 2014, while Apple's figures show an increase. This divergent evolution was captured in the graphs representing the dependency measure over time (Fig. 4) by the change of the trend from increase to decrease around data point 75 on the horizontal axes.

V. Conclusion and Open Problems

In our study, we find it evident that a growing company like Apple rides on consistent rising $o(X, Y)$ measures both from assets to liabilities and vice versa. Also, during the period a company is falling, as in Kodak's case, the $o(X, Y)$ measures are declining or flattened curves. It requires further research to determine if the trend of the $o(X, Y)$ measures can be used to predict a company's financial strength or the effectiveness of the assets and liabilities management.

In general, it is very tempting to search for causality relationships and explanations. But without clear reasons or logic to accept causality, we should only accept correlation. We tried to find causal relationships and explanations which go beyond simple statistical correlations. Having the evidence of an event, we may try to find its cause. The UD gives us a hint in explaining the influence of a time series over another one and furthermore finding the possible reason of the event. In our method, the $o(X, Y)$ is a time dependent UD measure, able to adjust incrementally to the incoming data streams. We have traced evident historical events of Apple and Kodak.

We are aware that this paper is far from being able to establish rigorous causal relationships. However, the time diagram of $o(X, Y)$ seems to be a great diagnostics tool, backed by rigorous probability theory. We consider our approach and experiments as a first step and we plan to continue this research.

References

[1] J. Pearl, *Causality: Models, Reasoning and Inference*, 2nd ed. New York, NY, USA: Cambridge University Press, 2009.

[2] W. R. Shadish, T. D. Cook, and D. T. Campbell, *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*, 7th ed. Boston: Houghton Mifflin, 2001.

[3] C. W. J. Granger, "Investigating causal relations by econometric models and cross-spectral methods," *Econometrica*, vol. 37, no. 3, pp. pp. 424–438, 1969.

[4] F. X. Diebold, *Elements of Forecasting*, 7th ed. Cincinnati: South Western, 2001.

[5] M. C. Guisan, "A comparison of causality tests applied to the bilateral relationship between consumption and gdp in the usa and mexico," *International journal of applied econometrics and quantitative studies : IJAEQS*, vol. 1, no. 1, (1/3), pp. 115–130, 2004.

[5] http://en.m.wikipedia.org/wiki/Eastman_Kodak

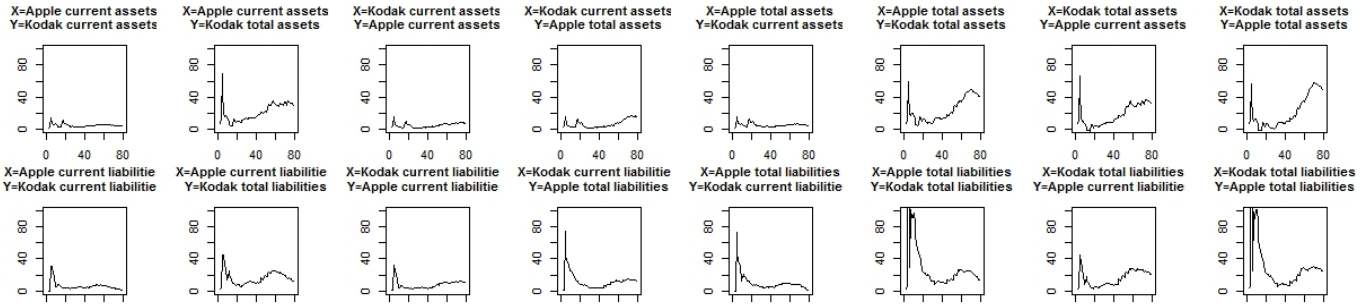


Fig. 4. Unilateral dependency of Apple versus Kodak assets and liabilities calculated as $o(X, Y)$, where X and Y were specified above each graph. The horizontal axis display the number of historical sample points used to obtain the corresponding value of $o(X, Y)$. When we compare the graph showing Apple current assets/liabilities with Apple total assets/liabilities we can identify similar patterns but magnified. The same remark is available for Kodak current assets/liabilities versus Kodak total assets/liabilities.

[6] T. Schreiber, "Measuring information transfer," *Phys. Rev. Lett.*, vol. 85, pp. 461–464, Jul 2000.

[7] O. Kwon and G. Oh, "Asymmetric information flow between market index and individual stocks in several stock markets," *EPL (Europhysics Letters)*, vol. 97, no. 2, p. 28007, 2012.

[8] J. Antonakis, S. Bendahan, P. Jacquart, and R. Lalive, "On making causal claims: A review and recommendations," *The Leadership Quarterly*, vol. 21, no. 6, pp. 1086 – 1120, 2010, leadership Quarterly Yearly Review.

[9] I. Guyon, "Results and analysis of the 2013 ChaLearn cause-effect pair challenge," in *NIPS 2013 Workshop on Causality*, 2013. [Online]. Available: http://videolectures.net/nipsworkshops2013_guyon_pair_challenge/

[10] G. Bontempi and M. Flauder, "From dependency to causality: a machine learning approach," *CoRR*, vol. abs/1412.6285, 2014. [Online]. Available: http://arxiv.org/abs/1412.6285

[11] R. Andonie and F. Petrescu., "Interacting systems and informational energy," *Foundation of Control Engineering*, vol. 11, pp. 53–59, 1986.

[12] A. Caţaron, R. Andonie, and Y. Chueh, "Asymptotically unbiased estimator of the informational energy with kNN," *International Journal of Computers, Communications and Control*, vol. 8, pp. 689–698, 2013.

[13] ——, "knn estimation of the unilateral dependency measure between random variables," in *Proceedings of the 2014 IEEE Symposium on Computational Intelligence and Data Mining ($CIDM$ 2014), IEEE Symposium Series on Computational Intelligence ($SSCI$)*, December 2014, pp. 471–478.

[14] O. Onicescu, "Theorie de l'information. energie informationelle," *C. R. Acad. Sci. Paris, Ser. A–B*, vol. 263, pp. 841–842, 1966.

[15] S. Guiasu, *Information theory with applications*. McGraw Hill New York, 1977.

[16] B. Silverman, *Density Estimation for Statistics and Data Analysis (Chapman & Hall/CRC Monographs on Statistics & Applied Probability)*. Chapman and Hall/CRC, 1986.

[17] L. F. Kozachenko and N. N. Leonenko, "Sample estimate of the entropy of a random vector," *Probl. Peredachi Inf.*, vol. 23, no. 2, pp. 9–16, 1987.

[18] H. Singh, N. Misra, V. Hnizdo, A. Fedorowicz, and E. Demchuk, "Nearest neightboor estimates of entropy," *American Journal of Mathematical and Management Sciences*, vol. 23, pp. 301–321, 2003.

[19] Q. Wang, S. R. Kulkarni, and S. Verdu, "A nearest-neighbor approach to estimating divergence between continuous random vectors," in *Proc. of the IEEE International Symposium on Information Theory*, Seattle, WA, 2006.

[20] L. Faivishevsky and J. Goldberger, "Ica based on a smooth estimation of the differential entropy," in *Advances in Neural Information Processing Systems 21*, D. Koller, D. Schuurmans, Y. Bengio, and L. Bottou, Eds. Curran Associates, Inc., 2009, pp. 433–440. [Online]. Available: http://papers.nips.cc/paper/3500-ica-based-on-a-smooth-estimation-of-the-differential-entropy.pdf

[21] J. Walters-Williams and Y. Li, "Estimation of mutual information: A survey," in *Proceedings of the 4th International Conference on Rough Sets and Knowledge Technology*. Berlin, Heidelberg: Springer-Verlag, 2009, pp. 389–396.

[22] T. Lemke and G. Lins, *ERISA for Money Managers*. Thomson Reuters, 2014.