Correlation, functional relationship, causation: the different role of causation in deterministic and probabilistic models

Sergio Focardi, PhD

Articles on causal modelling tend to present causation as a more robust link than correlation and tend to emphasize that causal models are more robust than models simply based on correlation. In many papers (see for example Scholkopf and Kukelgen) it is clearly stated that the current interest for causal modelling is due to the need to make advanced ML applications more robust. Causal modelling for business, economic and financial applications is primarily used as a decision support tool, but robustness is still an important concern.

But this article claims that the dichotomy correlation vs. causation gives only a partial view of the problem of modelling physical reality as well as business, economic, financial, and social systems. Much of modern physical modelling is based on laws that are neither correlation nor causation. Much of modern science is based on descriptive laws and much of scientific explanation is logical not causal. On the contrary, as we will discuss, probabilistic models of business, economic, financial, and social systems exhibit primarily causal functional relationship. The key reason is that probabilistic functional relationships are generally noninvertible.

In many domains, no truly fundamental laws are known with certainty. Knowledge is probabilistic and acquired through learning. In most cases, the process of knowledge discovery starts with estimating correlation. In these situations, it is natural to look for functional relationships as the next step of knowledge discovery. However, probabilistic functional relationships are generally non-invertible. Non-invertible functional relationships are the signature of causation.

Correlation

Consider two random variables, X and Y that belong to a bigger system. Suppose X and Y are correlated which means they move together. Important to note that we are discussing static or atemporal relationships between variables not time-dependent processes. Correlation between two variables means that, given an independent sample of pairs (X,Y), we will find that they have a similar behavior relative to the respective means.

The relationship of correlation is observational, which means that to compute correlations we simply observe the behavior of variables. Therefore, computing correlations gives no guarantee that an independent mechanism be responsible for correlation. Correlation between X and Y might depend on the system in which X and Y are embedded. In particular, correlation might depend on confounders, that is, empirical correlation might depend on a third variable. For example, suppose X and Y are the returns of the stocks of two firms. If the two firms are suppliers of the same large firm the correlation of X and Y might depend on the third firm.

The possible existence of confounders is the main reason why correlation is considered a weak relationship. In practice, there are also other reasons. For instance, in financial econometrics, in business and social studies, only rarely we have large samples. Therefore, we cannot really trust estimation. If we are working with large systems of hundreds of variables the covariance matrix estimated on samples of reasonable size will be noisy. In addition, modern economies are evolving complex systems which means that a correlation coefficient does not describe a stable economic reality. Correlation might break down because the economy changes.

But the paucity of data and the evolving nature of economies are problems common to all econometric tools that need to be estimated. The weakness of correlation is attributed primarily to the fact that correlation does not describe a mechanism independent from the environment. There is always the possibility of spurious correlations.

There are examples often used in the literature such as the correlation between consumption of ice cream and number of drownings of swimmers. The two variables are correlated but obviously there is no direct mechanism that links drownings and ice cream consumption.

An observation is in order. The Reichenbach common cause principle (RCCP) states that two variables X and Y can be correlated only if one variable causes the other or if there is a third variable Z causing both variables. In this case the two variables are conditionally independent: $P(XY|Z) = P(X|Z)P(Y|Z)$

The RCCP principle is widely used in many algorithms for causal discovery, though it is known that it admits exceptions. The RCCP applies to observed correlation: we observe the correlation between two variables, and we conclude that either one of the two variables causes the other one or a third variable causes both variables.

Why do we consider only causal relationships? The reason is twofold. First deterministic, non-causal relationships between X and Y would be common in the domain of physical systems but are basically excluded from the probabilistic, uncertain domains that we are considering.

Correlations are computed only in probabilistic systems; computing correlations for deterministic systems has little or no interest. Second, critically important, probabilistic functional relationships are generally non-invertible. Probabilistic functional relationships are therefore primarily causal relationships. This is the main topic of this article.

Causation in physical systems

It is useful to start by studying physical systems. Studying causation is difficult for several reasons including the lack of an independent definition of causation. Causation has been studied by philosophers since antiquity. Aristotle was the first to create a systematic theory of causation. Intuitively, causation responds to the question: Why things happen? But defining the object of causation is difficult. Are we discussing causation between events? Or between facts? Or between variables? The progress made in the last thirty years in studying causation is partially due to the framing of causation in mathematical terms as a relationship between variables.

Let's therefore restrict the concept of causation to variables: we say that a variable X causes a variable Y if a change of X is followed by a change of Y. In the probabilistic case we say that after a change of X we will find a change of the probability distribution of Y. The concept of causation is inherently asymmetric: a variable X has a causal effect on a variable Y but not vice versa. This is the key for distinguishing causation from functional relationships: causation is an asymmetrical functional relationship.

Is modern science a causal theory? To answer this question let's first analyse the concept of scientific explanation in physics. Loosely speaking, scientific explanation consists in establishing fundamental laws that apply to everything. The behavior of actual physical systems can be logically inferred from fundamental laws through mathematical deductions. There is no ontological commitment in modern physics, there is no attempt to find a deep, ontological Why?. There are basic laws and logical deductions.

This is the essence of the Deductive-Nomological (DM) theory of Carl Hempel and Paul Oppenheim. Modern physical laws are typically formulated with differential equations. The DM principle claims that scientific explanations consist in applying basic laws with appropriate initial and boundary conditions that describe the system. For instance, if we want to study how temperature propagates in a piece of metal, we solve the heat equations with appropriate initial and boundary conditions.

It is meaningless to ask why temperature propagates according to the heat equations, or why material points obey Newton's laws of dynamics or why electromagnetic fields propagate according to Maxwell's equations. Fundamental laws have the status of empirical theories, that is, they are empirical hypotheses. Fundamental laws are logically interconnected. The Duhem-Quine principle states that science, physics in particular, responds globally to the empirical test.

Science evolves in time. According to Thomas Kuhn the progress of science is not smooth but happens through discrete, disruptive paradigm shifts, where old laws and descriptive frameworks are replaced by new laws and new descriptive frameworks. The new theories are often incommensurable with old theories.

Modern science questions reductionism, that is the idea that there is a set of universal laws that explain everything. In the 1972 paper *More is Different*, Philip Anderson, recipient of the 1977 Nobel Prize in Physics, claimed that physics is hierarchical, that is, there is a hierarchy of physical theories that cannot be reduced to a single theory but needs special principles. This point of view is now shared by many scientists.

Fundamental physics is not causal.[1] Differential equations are not causal. Physics exhibits also functional relationships that are not causal. For instance, in classical physics, the force of gravitation between two masses is represented by the formula

$$F = K\frac{m1m2}{r\hat{}2}$$

This is a deterministic formula that links four variables. It is an empirical hypothesis verified in the course of more than three centuries. Above all, this formula does not exist in isolation, but it belongs to the theory of classical dynamics. According to the Duhem-Quine thesis theories respond to the empirical test globally. This law, taken together with the law of dynamics

$$F = ma$$

allows to derive the motion of the planets. Therefore, the level of validation of universal gravitation is high.[2]

---

[1] Quantum mechanics would need a separate discussion beyond the scope of this paper.

[2] General Relativity has replaced classical Newtonian gravitation law with a new law of gravitation based on the curvature of the space-time. We ignore this point because Newtonian gravitation is still a good approximation and because it is not important for our present discussion.

The law of gravitation is not a causal law, it is a descriptive law. It is a descriptive law because it is symmetrical and invertible while causality means that a variable X causes a variable Y but not vice versa. We do not say that the law of gravitation causes the motion of the planets. The law of universal gravitation is a fundamental law of nature. Together with laws of classical dynamics it allows to infer the motion of any object, from planets to projectiles. And we do not say that the law of gravitation describes a mechanism. It is a basic principle, a law of nature that we have observed in many direct measurements. In addition, it implies innumerable consequences such as the motion of the planets.

However, in practice science studies not only basic laws but also systems, be they molecules of drugs or the Earth climate. In addition, "soft" systems such as biological systems are studied, at least partially, with scientific methods. These soft systems exhibit causal behavior. And there are innumerable physical systems engineered as systems to be controlled. From home appliances to planes there are physical systems that exhibit an explicitly causal behavior where input variables control output variables.

Causation is not a law of nature, but it is due to the structure of the systems. The braking system of a car is engineered to be causal; it functions according to basic non causal laws, but it is engineered so that some variables control other variables asymmetrically. The same principle applies to natural causal systems, such as many biological systems.

In summary:

1. Science is based on fundamental laws of Nature that are primitive laws which cannot be further explained. These laws are not causal
2. The behavior of physical systems is described through deductions from basic laws and initial conditions; deductions from basic laws plus initial conditions is the essence of scientific explanations.
3. Systems might be described by variables whose behavior is represented by non-causal laws.
4. Some physical systems are engineered to be controlled; they exhibit input variables that can be manipulated arbitrarily and output variables that depend on input but not vice versa. Causal systems implement a causal mechanism that can be explained through deductions from basic laws plus conditions that describe the system.

"Soft systems": economies, financial markets, social systems, firms

For most systems such as financial markets and social and economic systems it would be impossible to establish deterministic laws of nature with a reasonable level of precision. Simply put, the behavior of these systems is too complex to be reduced to a deterministic axiomatic theory. The main reason is that we cannot reduce complex systems to the behavior of elementary components.

These systems are complex evolving systems. Thus far no elementary law has been found and it is questionable whether elementary laws for economies and social systems do exist. Understanding social, economic, financial and business systems heavily relies on learning. Given the level of uncertainty learning is generally formulated in a probabilistic framework.

The simplest level of learning is learning correlations. Suppose we have defined the descriptive framework of the system we want to study. Suppose also that historical data are available so that we can compute a covariance matrix. As discussed in the previous paragraphs the covariance matrix is a weak description of the system because of confounders.

The next level of description would be to learn functional relationships or even dynamical laws. Let's point out explicitly that learning causal models is not the only way of learning functional relationships between descriptive variables. If we allow models to be dynamic a number of different modelling strategies are available.

System Dynamics, invented and developed by Jay Forrester at MIT in the 1950-1970 period is a technique for developing models that include feedback loops and delays. System Dynamics was used to create WORLD3, a model of the world economy developed by a team at MIT that produced the famous report *The limits of development*.

Linear models such as Vector Auto Regressive (VAR) models are widely used. VAR models offer a dynamic representation of systems that evolve in time. Non-linear models, such as Regime Shifting models represent systems that exhibit jumps between different states.

But discussing dynamic models and eventual notions of dynamic causality is beyond the scope of this paper. Here we want to discuss the current generation of causal models and their relationship to non-causal models.

The state-of-the-art causal model is the SCM model developed by Judea Pearl and his collaborators and by a group of philosophers at the Carnegie Mellon University: Peter Spirtes, Clark Glymour, and Richard Scheines who developed TETRAD a software for creating causal models and for discovering causal structures.

Suppose a system is described by a set of variables $Xi, i = 1, \dots, n$. A SCM is formed by a set of structural equations: $Xi = f(PAi, Ui)$ where PAi represents the variables that are direct causes of *Xi* and *Ui* is a random term. Under a number of non-statistical assumptions, a SCM can be learned from a covariance matrix. The assumptions essentially require that the SCM represent a causal structure. If we assume that the *Ui*s are independent there are no confounders, and the causal system can be estimated.

Why this emphasis on causality? While don't we search for invertible functional relationships as in the physical case?

Recall from the previous sections, that in the physical case laws of nature which are not causal, coexist with non-causal and causal functional relationships. Why in the soft cases it seems that causal systems are the most important, and perhaps the only, representation of static systems?

There are two different types of explanations. The first is that in daily life as well as in business and in economics we are primarily interested in tools to support decision making. Soft systems are primarily human systems where we are interested in making decisions such as business decisions or policy changes. Therefore, we have an obvious interest in understanding causality. In medical applications we are interested, for instance, in understanding if a treatment is effective or not. Therefore, again we have an obvious interest in causation.

The second reason is technical. It is unlikely that we can invert static functional relationships that include probabilistic terms. The intuition is the following. A functional relationship Y=f(X) is generally a one-to-one relationship between a value of X and a value of Y. Therefore, it is reasonable to assume that a deterministic functional relationship be invertible.

However, a probabilistic functional relationship Y=f(X,U) projects every X into a set of values Y. It is therefore difficult to invert such a relationship. The CMU blog 7-Causal inference (https://blog.ml.cmu.edu/2020/08/31/7-causality/), authored by Lucas, Huang and Stelmakh, offers a simple explanation. Consider a simple regression:

$$Y = f(X) + u$$

where *u* is a noise term uncorrelated with X. If we try to invert this relationship we obtain in general:

$$X = g(Y) + v$$

Where *v* is no longer independent from Y. In other words, a regression equation cannot be inverted because, except in special cases, we will obtain different regression equations.

Differences between functional relationship and causal relationships

From the above we see that the main difference between causation and functional relationship is that causation is a non-invertible functional relationship. An invertible functional relationship cannot be causal because we cannot intervene on both variables. For example, when we push the brake pedal, we want to make sure that the car stops, and we do not want to experience an opposite reaction that the car continues to go and moves the brake pedal. Non-causal functional relationships do not admit interventions.

In the deterministic case, which is prevalent in physical systems, there are both causal and non-causal relationships. Non causal relationships are typical of laws where we cannot intervene, while causal relationships are engineered to allow interventions.

In the probabilistic case, prevalent in soft applications, functional dependencies are generally non-invertible, have a direction and therefore are causal relationships.

In summary

In summary, in physics there is a sharp distinction between laws of nature and laws descriptive of physical systems. Laws of nature are primarily differential equations, that describe processes. However, there are laws of nature formulated as functional relationships.

Laws of nature are primitive and are not justified or cannot be inferred from other principles. Laws of nature are descriptive not causal. Laws describing relationships between descriptive variables of physical systems might be descriptive or causal. Both imply independent mechanisms. The difference is essentially that causal laws are asymmetric while descriptive laws are symmetric.

"Soft systems" such as economies, financial markets, social systems, and firms do not admit laws of nature. Perhaps it is safer to state that thus far no law of nature for soft systems have been discovered. There is no well-established theory, validated with a high level of precision. Much knowledge about these systems is obtained through learning.

We can recognize three levels of description. Correlations are empirical laws that might depend on the entire system. That is, correlations might be influenced by confounders. To reach a higher level of robustness, we must write functional equations. Functional equations imply the existence of independent mechanisms which, given the absence of fundamental laws have to remain unspecified.

Probabilistic causal relationships are generally non-invertible and therefore are causal relationships.