

Cause and correlation

One day, when I was a small child, walking in the woods with my father, he started clicking his fingers in an odd, rhythmic way. “What are you doing?”, I asked, to which he replied, “The noise keeps the elephants out of the trees.” I stared at him like he was insane. “Well”, he said, “you haven’t seen any elephants in the trees have you?” I could see at once that he had a point, yet that something was also perversely wrong. Correlation, my father was teaching me, does not imply causation.

Yet there is perhaps no method so prized in science as searching for correlations. We plot data and look at graphs hoping for patterns to emerge. We calculate regressions and search for combinations of variables able to ‘explain’ anything from patterns in solar flares to inflation.

And correlations often do point to real causal effects. We first learned that light had a finite speed when Ole Romer observed a correlation between variations in the period of Jupiter’s moon Io and the Earth’s motion towards or away from Jupiter. Today, in fields from econometrics to statistical physics, the search for correlations in data drives a small scientific industry. But we’re really only beginning to appreciate how we might go beyond statistics to get at cause and effect more directly.

If correlation does not imply causation, it’s also true that causation need not imply correlation. Writing in *Science*, ecologist George Sugihara and colleagues give a nice example (*Science* **338**, 496–500; 2012). Consider a simple model for the interaction of two species — foxes and rabbits, say — where the yearly change in each population follows a simple logistic map, with some weak coupling also between populations. That is, if $X(t)$ and $Y(t)$ denote rabbit and fox populations respectively, we have $X(t+1) = X(t)[r_X - r_X X(t) - \beta_{X,Y} Y(t)]$ and $Y(t+1) = Y(t)[r_Y - r_Y Y(t) - \beta_{Y,X} X(t)]$, with parameters r_X , r_Y , $\beta_{X,Y}$ and $\beta_{Y,X}$.

Despite the simplicity of these equations, the two variables in a single sample run can move up and down in strong synchrony for a time, then lapse into a period of opposing or anti-correlated behaviour, and then drift into another regime with no apparent link at all. There is no long-term statistical correlation between the two variables, despite a clear causal link.

Correlation alone isn’t informative, nor is the lack of it. Might there be more subtle



Not only does correlation not imply causality, but lack of correlation needn’t imply a lack of causality either.

patterns, beyond correlations, that do really signify causal influence?

On this question, one influential idea was introduced in 1969 by economist Clive Granger. He reasoned that if some thing X causally influences some other thing Y , then including X in a predictive scheme should improve predictions of Y . Conversely, excluding X should make predictions worse. Causal factors can be identified as those that reduce predictive accuracy when excluded.

This notion of ‘Granger causality’ makes obvious intuitive sense, and has found many applications, especially in econometrics. However, the theory applies to stochastic variables, especially in linear systems. As Granger himself noted, “The theory is, in fact, non-relevant for non-stochastic variables.” Which is unfortunate as so much of the world seems to be more suitably described by nonlinear, deterministic systems.

For this world, Sugihara and colleagues now propose an alternative, which they call ‘convergent cross mapping’. It’s an empirical method that takes data and tests for causal links and, as they demonstrate, seems to have significant power over and above Granger causality and other methods.

One problem with Granger causality, the authors point out, is that intimate connections between the parts of any nonlinear system make ‘excluding’ a variable more or less impossible. In the system of foxes and rabbits, for example, you might exclude Y and see if you can predict X . That sounds like asking if a variable can predict itself, and the answer is yes. A key result in dynamical system theory — the Takens embedding theorem — shows that one can always reconstruct the dynamical attractor for a system from data in the form of lagged samples of just one variable. In effect, $X(t)$ is always predictable from enough of its earlier values. Excluding Y doesn’t make X any less predictable. The notion of Granger causality would erroneously conclude that Y is non-causal.

To get around this problem, Sugihara and colleagues use the embedding theorem to their advantage. The reconstruction

trick can be done for both variables X and Y , reconstructing two manifolds, M_X and M_Y , each of which describes the dynamical attractor of the whole system. Now, sensibly, if X has a causal influence on Y , one should expect this influence to show up as a direct link between the dynamics on these two manifolds. Knowing states on the manifold M_Y at a certain time should make it possible to know the states on M_X at the same time.

All that sounds rather abstract, but Sugihara and colleagues demonstrate it in several examples. First, the technique works beautifully in the rabbits and foxes model, clearly demonstrating causal links between X and Y . It also works well in a second case with known answers: a classic ecological system involving the interaction of the carnivorous single-celled predator *Didinium* and its prey *Paramecium*. Again, the method shows that both organisms exert a causal influence on the other, as is known from direct study of their interactions. In contrast, tests using Granger causality fail for both these problems.

But the method really shows its value in the context of an outstanding puzzle. Ecologists have for decades debated what’s going on with two fish species, the Pacific sardine and northern anchovy, the populations of which on a global scale alternate powerfully on a decadal timescale. The data, some suggest, imply that these species must have some direct competition or other interaction, as when the numbers of one go up, those of the other go down. Failing any direct observation of such interactions, however, others have proposed that the global synchrony betrays something else — global forcing from changing sea surface temperatures, which just happen to affect the two species differently.

Strikingly, the results from convergent cross mapping seem to resolve the matter in one stroke. The analysis shows no evidence at all for a direct causal link between the two species, and clear evidence for a link from sea surface temperature to each species.

This is a beautifully simple application of the basic ideas of dynamical system theory, illustrating how much more value might be drawn from that theory even now. I find it hard to believe this technique won’t find immediate application in economics and finance as well as in ecology, neuroscience and elsewhere. Causality is as fundamental a concept as we have in science, yet we still have much to learn about recognizing it. □

MARK BUCHANAN