Chapter 6

Discovery Algorithms without Causal Sufficiency

6.1 Introduction

The preceding chapter complied with a common statistical fantasy, namely that in typical data sets it is known that no part of the statistical dependencies among measured variables are due to unmeasured common causes. We almost always fail to measure all of the causes of variables we do measure, and we often fail to measure variables that are causes of two or more measured variables. Any examination of collections of social science data gives the striking impression that variables in one study often seem relevant to those in other studies. Record keeping practices sometimes force econometricians to ignore variables in studies of one economy thought to have a causal role in studies of other economies (Klein, 1961). In many studies in psychometrics, social psychology and econometrics, the real variables of interest are unmeasured or measured only by proxies or "indicators." In epidemiological studies that claim to show that exposure to a risk factor causes disease, a burden of the argument is to show that the statistical association is not due to some common cause of risk factor and disease; since not everything imaginably relevant can be measured, the argument is radically incomplete unless a case can be made that unmeasured variables do not "confound" the association. If, as we believe, no reliable empirical study can proceed without considering whether relevant variables are unmeasured, then few published uncontrolled empirical studies are reliable.

In both experimental and non-experimental studies the unrecognized presence of unmeasured variables can lead to erroneous conclusions about the causal relations among the variables that are measured, and to erroneous predictions of the effects of policies that manipulate some of these variables. Until reliable, data-based methods are used to identify the presence or absence of unmeasured common causes, most causal inferences from observational data can be no more than guesswork at best and pseudo-science at worst. Are such methods possible? That question surely ought to be among the most important theoretical issues in statistics.
Statistical methods for detecting unmeasured common causes, or "confounding" in the terminology epidemiologists prefer, has been chiefly developed in psychometrics, where criteria for the existence and numbers of common causes have been sought since the turn of the century for special statistical models. The results include a literature on linear systems that contain criteria (e.g., Charles Spearman's (1904) vanishing tetrad differences) for latent variables that proved, however, to be neither necessary nor sufficient even assuming linearity. Criteria for two latent common causes were introduced by Kelley (1928), and related criteria are used in factor analysis, but they are not correct unless it is assumed that all statistical dependencies are due to unmeasured common causes. For problems in which the measured variables are discrete and their values a stochastic function of an unobserved continuous vector parameter \( \theta \), a number of criteria have been developed for the dimensionality of \( \theta \) (Holland and Rosenbaum, 1986). Suppes and Zanotti (1981) showed that for discrete variables there always exists a formal latent variable model in which all measured variables are effects of an unmeasured common cause and all pairs of measured variables are independent conditional on the latent variable. Their argument assumes the model must satisfy only the Markov Condition; the result does not hold if it is required that the distributions be faithful.

Among epidemiologists, (Breslow and Day, 1980; Kleinbaum, Kupper and Morgenstern, 1982) the criteria introduced by the Surgeon General's report on Smoking and Health (1964) are sometimes still advocated as a means for deciding whether a statistical dependency between exposure to risk factor \( A \) and disease \( B \) is "causal," apparently meaning that \( A \) causes \( B \) and \( A \) and \( B \) have no common causes. The criteria include (i) increase in response with dosage; (ii) that the statistical dependency between a risk factor and disease be specific to particular disease subgroups and to particular conditions of risk exposure; (iii) that the statistical association be strong; (iv) that exposure to a risk factor precede the period of increased risk; (v) lack of alternative explanations.

Even in causally sufficient systems, where all common causes of measured variables are themselves measured, such criteria do not separate causes from correlated variables. They fail even to come to grips with the problem of unmeasured "confounders." Criterion (v) is an evasion; the problem in uncontrolled studies is exactly that there are too many alternative explanations of the data. Criterion (iv) is a banality which is of no use at all in deciding whether there are measured or unmeasured common causes at work. Criterion (iii) is defended on the grounds that "If an observed association is not causal, but simply the reflection of a causal association between some other factor and disease, then this latter factor must be more strongly related to disease (in terms of relative risk) than is the former factor," (Breslow and Day, 1980).