Quantitative Causal Reasoning in Stock Price Index Prediction Model

MyoungJong Kim · Ingoo Han

Graduate School of Management, KAIST

Abstract

Artificial Intelligence literatures have recognized that stock market is a highly unstructured and complex domain so that it is difficult to find knowledge that belongs to that domain. This paper demonstrates that the proposed QCOM can derive global knowledge about stock market on the basis of a set of local knowledge and express it as a digraph representation. In addition, inference mechanism using quantitative causal reasoning can describe the qualitative and quantitative effects of exogenous variables on stock market.

1. Introduction

There are obviously a variety of mathematical models in business area. However, these techniques may not be completely applied to business problems having ambiguity and imprecision because of the difficulty in formulating these problems and interpreting the results of analysis via common sense terms, and thus business problems would probably be suitable application areas for qualitative modeling.

The theories of causal ordering have been used to build qualitative economic and engineering models, and investigate causal structure of these models. The first work on theories of causal ordering was causal ordering ([5], [6], [8] and [9]), an asymmetric relation among the variables and equations of a set of simultaneous equations. Establishing a causal ordering involves finding the minimal complete sets whose values can be computed independently of the remaining variables of different orders. The essentially equivalent, but more intuitive, component-based confluence analysis of [2], [3] does not order variables in accordance with causality. Instead, the model structure is interpreted by propagating a given change through causes and effects. The causal-ordering graph [1], unifying the existing theories of causal ordering, is a graphical representation of causal ordering. It derives the signed causal ordering graph, representing global causal structure, on the basis of locally imposed directions of causality derived from the underlying theory.

The purpose of causal reasoning is to explain and predict phenomena based on causal notions embedded in a model. Causal reasoning has been traditionally used to understand the structure and behavior of qualitative models. This reasoning generates model-based explanations described in the causal terms ("A caused B") by using the causal propagation via cause/effect relations. However, qualitative causal reasoning based on theories of causal ordering is not appropriate for the area where precise quantitative information is required because the concepts of causal ordering techniques are qualitative.

Our main concern is not to test the validity of a model and elucidate the causality among variables in a model, but to demonstrate a new kinds of causal reasoning, that is, quantitative causal reasoning to propose a framework that is capable of combining qualitative and quantitative reasoning. In contrast to qualitative causal reasoning, the proposed quantitative causal reasoning is operated in a quantitative model, Korea stock price index (KOSPI) prediction model derived by the quantitative causal ordering map (QCOM). The proposed reasoning can offer both qualitative and quantitative explanations about the model via static and dynamic propagation. Static analysis is achieved by propagating a disturbance of one or more exogenous variables while dynamic behavior of model is investigated by introducing causal relations between time periods. Therefore, the proposed reasoning can results in multipliers as well as model-based explanations of these multipliers so that it can be a framework to implement qualitative and quantitative reasoning simultaneously.

The structure of this paper is as followings. Section 2 shows the process of the QCOM for KOSPI prediction model. Quantitative causal reasoning in static and dynamic model will be demonstrated with an illustrative KOSPI prediction model in section 3. In section 4, conclusion and directions for future research are presented.

2. Quantitative Causal Ordering Maps

The causal-ordering graph of [1] is proposed to represent the causal ordering graphically, equivalent to the signed digraph representation, on basis of locally imposed causal direction. The underlying idea is that each relation in the model determines exactly one variable, corresponding to a dependent variable in a mathematical model. The causal-ordering graph is developed according to the following two phases: (1) The model graph and (2) The causal-ordering graph. Refer to [1] for details.

Phase 1: *The model graph*

The model graph is a graphical representation showing the matching of relations and all variables that appear in relations. The perfect matching means that every variable matches exactly one relation.

Three kinds of criteria are used to determine the perfect matching. They are definitional relations – the direction of causality imposed is from the other variables to the dependent variable that is defined in that relation, institutional relation – the direction of causality imposed is from other variables to variable that is determined on the basis of other variables, and behavioral relation – the direction of causality is derived from underlying economic theory.

Phase 2: *The Causal-ordering graph*

Given a model graph and the perfect matching, the causal-ordering graph is derived by applying the following three steps of causal ordering procedure:

(1) Direct every arc $(v_i, r_i) \in W$ from r_i to v_j

(2) Direct every arc $(r_i, v_j) \in (A - W)$ from v_j to r_i

(3) Contract all nodes $r_i \in R$

where v_j, r_i, W, A and R indicate jth variable, ith relation, the perfect matching, the arc set and the set of relations respectively. The signed causal-ordering graph is obtained from the causal-ordering graph by assigning the appropriate signs to the links in accordance with the underlying theory.

The QCOM, an improved version of causal-ordering graph, is a weighted digraph representation. It is developed to encompass quantitative as well as qualitative reasoning so that causal reasoning based on this mechanism generates numerical multipliers and symbolistic explanations in static and dynamic analysis. The QCOM is obtained from the causal-ordering graph by applying path analysis to historical instances.

The QCOM is applied to KOSPI prediction model reflecting the effects of macroeconomic variables on the composite stock price index and to predict the composite stock price index behavior of next month based on such a model. The QCOM for KOSPI prediction model is based on an economic theory of [4] that is designed for illustrating causality among macroeconomic variables. A set of monthly growth rates of nine economic factors of current month are used in the prediction of the monthly growth rate of composite stock price index of next month ($CSPI_{+1}$) as endogenous variables, including BCA (balance of current account), EPI (export price index), GNP (gross nation production), CPI (consumer price index), WPI (wholesale price index), W (monthly average wage), ER (exchange rate), IPI (export price index), IR (interest rate). The monthly growth rates of exogenous variables are M2 (M2 average of current month), ER_{-1} (exchange rate of last month), IR_{-1} (interest rate of last month) and GNP_{-1} (GNP of last month).

The set of variables and relations are described in Table 1.

Variables	Set of Relations
$CSPI_{+1}$	r1 ($CSPI_{+1}$, GNP, IR)
BCA	r2 (BCA, EPI, IPI)
IR	r3 (IR, GNP, GNP_{-1}, M2, IR_{-1})
EPI	r4 (EPI, GNP, ER, IPI, W)
GNP	r5 (GNP, W, IPI, ER, ER_{-1}, M2)
CPI	r6 (CPI, GNP, GNP_{-1})
WPI	r7 (WPI, CPI)
W	r8 (W, WPI)
ER	r9 (ER, BCA, ER_{-1})
IPI	r10 (IPI, ER)
M2	r11 (M2)
ER_{-1}	r12 (ER_{-1})
IR_{-1}	r13 (IR_{-1})
GNP_{-1}	r14 (GNP_{-1})

Table 1. The set of variables and relations

Based on relations described in table 1, Figure 1 shows the model graph and the perfect matching, where each link of the perfect matching is denoted as a bold line.

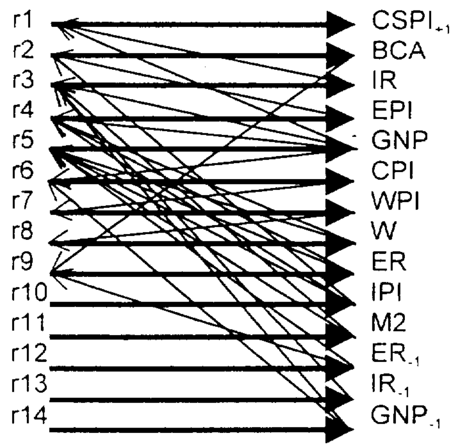

Figure 1. The Model Graph

The causal-ordering graph are derived as shown in Figure 2. The causal-ordering graph is generated by tracing the links of Figure 1, for example, there are links from r13 to IR_{-1} (r13 $\rightarrow$ IR_{-1}) and IR_{-1} $\rightarrow$ r3 $\rightarrow$ IR $\rightarrow$ r1 $\rightarrow$ $CSPI_{+1}$ in sequence. With these links, the causal ordering among them is determined as followings: IR_{-1} $\rightarrow$ IR $\rightarrow$ $CSPI_{+1}$ as shown in Figure 2, which mean that $CSPI_{+1}$ is directly causally dependent on IR while indirectly on IR_{-1} via IR.

The QCOM is obtained from the causal-ordering graph by applying path analysis to historical instances. We collect the monthly data set for 72 months from January 1990 to December 1995. Based on the Two-Stage Least Squares parameter estimates of the QCOM, the QCOM with numerical causal coefficients are represented in Figure 3.

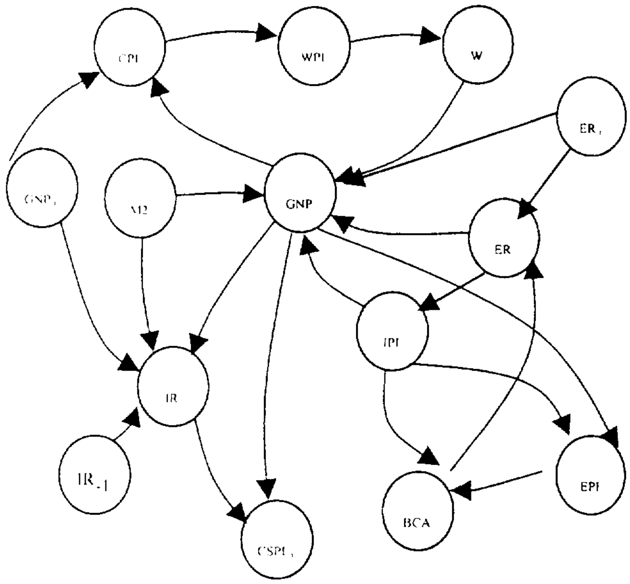

Figure 2. The Causal-ordering graph

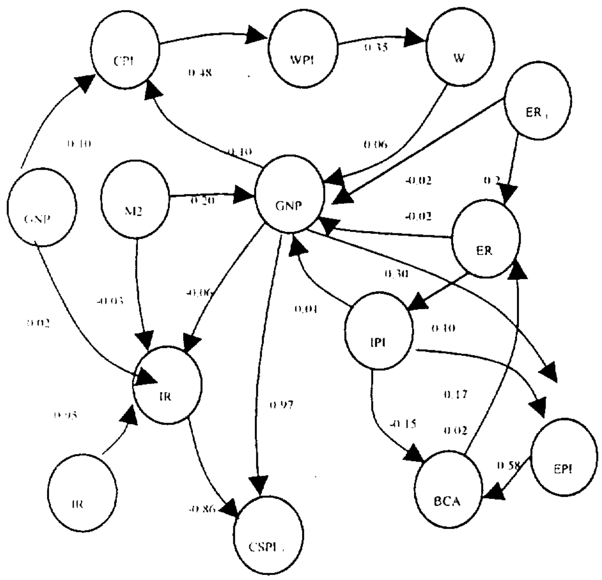

Figure 3. QCOM for KOSPI Model

3. Causal Reasoning

Propagating changes is an intuitive way to perceive and predict behaviors of qualitative systems. Causal propagation or simulation with the QCOM for KOSPI prediction model presents the behavior and structure of systems. Causal reasoning contains two kinds of propagation method: static and dynamic propagation. Causal reasoning in static model starts with the disturbance of exogenous variables and this disturbance propagates to endogenous variables through cause/effect relations while dynamic simulation is achieved by introducing causal relations between time periods.

In static analysis on KOSPI prediction model, such changes propagate along the cycles. At first, consider causal propagation in a cycle circulating GNP, CPI, WPI, W, and GNP (GNP → CPI → WPI → W → GNP). The changes in GNP affects CPI, WPI, W and GNP in sequence, yielding the changes in the corresponding variables. This causal propagation process is continued until a convergence threshold (herein, 0.0001) is attained, accumulating the changes for the variables involved each time around the cycles. This process is graphically represented as economy level causal tree in Figure 4, wherein the path of this repetitive causal propagation is denoted as bold lines. After completing causal propagation in bold-lined branch, the propagation process is continued in the neighboring branch GNP → EPI → BCA → ER → GNP presented as the dot lines in Figure 4 until a convergence threshold is realized. In a similar manner described above, this propagation process is applied to all branches of causal tree.

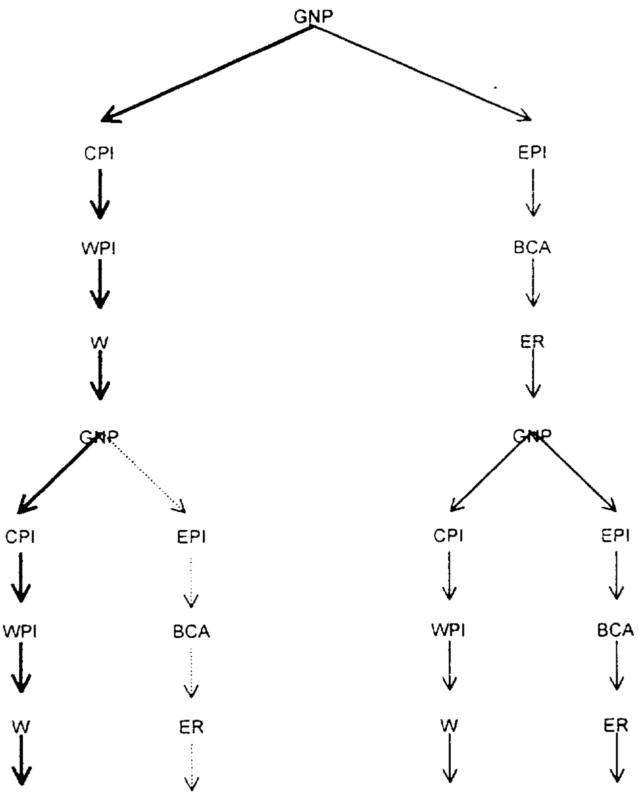

Figure 4. Causal Tree

The changes in endogenous variables IR and $CSPI_{+1}$, which is not involved in cycles, is computed once after completing causal propagation in all of the branches.

Dynamic simulation is obtained by introducing causal relations between time periods. Changes due to each of the lagged variables are propagated in a manner as done in the static model. In KOSPI prediction model, dynamic simulation is achieved by propagating the changes in lagged endogenous variables.

Table 2 presents the summarized results of causal reasoning about the increase in M2 by 20 units. The computed multipliers are close to those obtained from the matrix solution method. This indicates that the proposed reasoning generates results consistent with those generated from traditional numerical model.

Variable	Time period			
	0	1	2	3
GNP	4.0038	0.0037	-0.0001	-0.0000
EPI	1.2014	0.0012	-0.0000	-0.0000
BCA	0.6966	0.0006	-0.0000	-0.0000
FR	0.0139	0.0036	0.0009	0.0002
IPI	0.0014	0.0004	0.0001	0.0000
CPI	0.4004	0.4007	0.0004	-0.0000
WPI	0.1922	0.1924	0.0002	-0.0000
W	0.0673	0.0673	0.0001	-0.0000
IR	-0.8402	-0.7184	-0.6824	-0.6482
$CSPI_{-1}$	4.6063	0.6214	0.5867	0.5575

Table 2. The Results of Propagation

4. Concluding Remarks

Artificial Intelligence literatures have recognized that stock market is a highly unstructured and complex domain so that it is difficult to find knowledge that belongs to that domain [7]. This paper demonstrates that the proposed QCOM can derive global knowledge about stock market on the basis of a set of local knowledge and express it as a digraph representation. In addition, inference mechanism using quantitative causal reasoning can describe the qualitative and quantitative effects of exogenous variables on stock market.

The proposed mechanisms can positively contribute to existing mathematical models and qualitative causal reasoning by providing the following advantages:

(1) Mathematical models formulated in differential equations assume certainty, however it is quite difficult to represent real-world problems having ambiguity as a set of equations. The QCOM, considering the structure of system as a weighted digraph form, can be interpreted as a mathematical model by using direct effects of Table 2. Quantitative causal reasoning can support the qualitative interpretation of mathematical models by using causal propagation as shown in previous section.

(2) Quantitative causal reasoning can add the precision to qualitative causal reasoning by generating qualitative multiplier in static and dynamic propagation.

Reference

[1] Berdendsen, R., Causal ordering in economic models, Decision Support Systems 15, 1995, pp. 157-165.

[2] De Kleer, J. and Brown, J. S., A qualitative Physics based on confluences, Artificial Intelligence 24, 1984, pp. 7-83.

[3] De Kleer, J. and Brown, J. S., Theories of causal ordering, Artificial Intelligence 29, 1986, pp. 33-61.

[4] Han, S. S., and Seo, S. H., Econometric analysis of Korean economy, Seoul Economy Research Institute. 1995

[5] Iwasaki, Y., Causal ordering in a Mixed Structure, Proceedings of the Seventh National Conference on Artificial Intelligence (AAAI-88), St Paul, Minnesoth, 1988, 313-318.

[6] Iwasaki, Y., and Simon, H. A., Causality in device Behavior, Artificial Intelligence 29, 1986, pp. 3-32.

[7] Kim, M. J., Han, I. G., and Lee, K.C., The integration of machine and human knowledge by fuzzy logic for the prediction of stock price index, 1998, Accepted in 5th pacific Rim Internation Conference on Artificial Intelligence.

[8] Simon, H. A., Causal ordering and identifiability, in Hood and Koopmans, eds., Studies in Econometric Methods, Cowles Commission research 14, John Wiley and Sons, New York, 1953, pp. 49-74.

[9] Simon, H. A., and Iwasaki, Y., Causal ordering, Comparative statistic, and Near Decomposability, Journal of Econometrics 39, 1988, pp 149-173.