NONLINEAR SIMULTANEOUS EQUATION MODELS IN DISCRETE TIME

John Sall, SAS Institute

Models to be addressed

The model is a system of equations, possibly nonlinear, that are solved simultaneously for a set of variables at a given point in time. The model can be pictured:

---

**Diagram:**

- Input variables → **MODEL** → Output variables
- Exogenous
- Endogenous
- Lagged endogenous (feedback)

---

Given values for the exogenous variables, the system determines values for the endogenous variables that uniquely solve the equations.

For dynamic models (across time), there is a series of equally-spaced points in time for which the model is solved. This is usually monthly, quarterly, yearly, daily, or hourly. Sometimes the results from one point in time are used in the system after a delay of one or more periods. These are written in the system as "lagged endogenous" variables.

Applications of Models

1. Econometric Models
   - These are usually macro-models of the entire economy of a country, region, or market segment.

2. Corporate Financial Models
   - These include the usual balance-sheet and income statement accounting equations, plus some financial planning equations and policy variables that financial specialists use to simulate the behavior of a company's finances under various conditions.

3. Physical, Biological, and Ecological Models
   - Physical like a mechanical, electronic, hydraulic or hydrologic system
   - Biological and chemical models of living processes
   - Ecological models of systems in nature

Discrete Time Models Contrasted with Other Types of Models in Time

Discrete time models are often called "difference system" models because they can be written as a system of difference equations instead of lag terms. The model is simpler in many ways than other types of dynamic models. This type of model usually corresponds to the way data is collected in batches at intervals.

Another type of dynamic model is run in continuous time. These models can be described by a system of differential equations. CSHP is one software product that handles these types of models. The basic technique is to compute integrals across time using variable stepsize methods that guarantee certain accuracy.

Another type of dynamic model is event-oriented. These models run in jumps of time that are not equally spaced. GPSS and SIMSCRIPT are popular software products to handle these models. The basic technique is to generate events that will happen at some time in the future, and manage these events in a next-event queue.

Some processes are inherently one type or another. Sometimes the process may be one type, but be modeled by another type. For example, consider raindrops falling into a bucket. The raindrops themselves are event-oriented. However, since so many of these events happen, the process of the bucket filling with water might be considered continuous in time. However, if you measured the filling of the bucket at equally-spaced intervals of time, you might want to approximate the continuous model with a discrete time model.

Translating between Discrete Time Models and Other Types

1. Continuous models may be approximated as accurately as you want by discrete time models. The trick is to pick a small enough time interval to capture the behavior of the system across time. Using a discrete time model for continuous systems is equivalent to using Euler's method for approximating an integral:
   \[ Y = \int X(T) \, DT \quad \text{Approximated by} \quad Y = \Delta X(T) \times DT \]
   I.E. \[ \frac{DY}{DT} = X(T) \quad \text{Approximated by} \quad \Delta \frac{Y}{Y} = X(T) \]
   Calculated by \[ Y = \text{LAG}(Y) + X \times DT \]

2. Event-oriented models may be approximated accurately by discrete time models if a lot of events happen in each discrete time interval, and if the behavior within a time interval can be described by a system of equations. For example, the economy of the United States operates through millions of transactions (events), however these events can be aggregated and the behavior modeled in the aggregate.

Steps in Modeling

1. FORMULATION of the model.
2. ESTIMATION and HYPOTHESIS TESTING.

You give it data from the model, you want parameter estimates.

This is the scientific side of modeling—often you want to test hypotheses concerning the
parameters. Or you may just be "calibrating" the model for later use in policy simulation.

3. PREDICTION and SIMULATION.

You supply the parameters (and exogenous data); you want to find out the endogenous data that fits the model.

The running of the model may be for several reasons:

a. Evaluation of the model by comparing simulated results with actual results, MODEL VALIDITY, statistics of fit.

b. Predicting the future, FORECASTING.

e. Evaluating the consequences of changing some parameters, performing POLICY EXPERIMENTS, doing "what if" analysis.

Estimation and Hypothesis Testing

Least squares regression of each equation in the model produces inconsistent estimators!

This is because the endogenous variables intertwine through the equations to produce statistically incestuous relationships between carrier variables (on the R.H.S. of the equations) and the error terms.

The standard cure to produce consistent estimators is to clean all R.H.S. endogenous variables by the method of projecting them through the space of the exogenous and lagged endogenous variables, or other instruments. This purgative can be effected by using a preliminary (instrumental) regression to predict the R.H.S. endogenous variables, then performing a regression for the structural equation using these predicted endogenous variables. This is the method of

TWO STAGE LEAST SQUARES.

You get estimates that have the same asymptotic properties as estimates in a classical regression model—though the small sample properties might be somewhat different. Another method of estimation that is slightly more efficient than 2SLS is

THREE STAGE LEAST SQUARES.

Prediction and Simulation

The only basic equipment you need is some method of solving a system of simultaneous equations.

1) NEWTON's method (uses derivatives, inverses)

2) GAUSS-SEIDEL (just execute the program repeatedly)

Newton's method is best for models that:

A) Are small or can be made small.

B) Are very simultaneous.

C) Contain implicit equations.

Gauss-Seidel is best for models that:

A) Are large.

B) Are Sparse (not very simultaneous).

C) Can be written explicitly for each endogenous variable.

The simulation is characterized by the values that are used for lagged endogenous variables.

A) STATIC simulation uses actual values for the lagged endogenous variables. It is just predicting each point in time.

B) DYNAMIC simulation uses results from previous simulation periods to feed the lagged endogenous values. It is predicting through time without knowing anything but the exogenous variables.

MONTE-CARLO brute force can be used to get distributions of the results given probability distributions on inputs.

SAS Procedures for Modeling

With SAS, the modeling process is divided into three parts, and implemented as three separate procedures that communicate with each other. These procedures were written last year, but are still in an experimental stage—more work needs to be done to ready them for commercial release. The three procedures are:

1. PROC MODEL is used to specify the model.

2. PROC SYSLIN is used to estimate parameters using data.

3. PROC SIMLIN is used to predict and simulate with the system.

These procedures have been designed to incorporate the best software engineering to make them both efficient and easily used. Several features in the product that are rarely seen on other modeling systems are:

1. The model is compiled into machine code rather than just being interpreted. This makes it very efficient.

2. Derivative methods are used. This means that equations can be specified implicitly. Derivative methods are more successful on small dense systems.

3. SAS automatically solves for the derivatives analytically, and places the derivative code into the model.

4. You only have to specify the model once for many estimations and simulations.
5. SAS provides advanced estimation methods for simultaneous models.

6. The modeling system interfaces with all the other features already in SAS.

The MODEL procedure

With the MODEL procedure, the user enters a series of equations and other statements that form a program. The equations can use lag and difference functions as well as the full range of mathematical functions. Programming constructs available include the IF-THEN-ELSE-DO-END structured programming control features similar to PL/I. With PROC MODEL, the user also identifies which variables are endogenous, exogenous, parameters, and instruments. If parameters are to be estimated later, MODEL solves out the analytic derivatives of the endogenous variables with respect to the parameters. If a derivative-based solution for simulation is to be used, PROC MODEL solves out the derivatives of the endogenous variables with respect to the other endogenous variables. The result of PROC MODEL is an intermediate file containing the parsed statements.

Example: Klein's Model I

This famous model of the economy of the U.S. has eight endogenous variables as described by eight equations. The first three equations have parameters that need to be estimated. The remaining equations are identities. Three endogenous variables are present in lags. To include them in the list of instruments for the estimation process, we needed some programming statements. This model happens to be linear in both the parameters and the endogenous variables, thus you can compare its specification here with its specification for the SYSREG procedure.

```
PROC MODEL;
ENDOGENOUS C P W I X WSUM K Y;
EXOGENOUS WP G T YEAR;
INSTRUMENTS WP G T YEAR PLAG KLAG XLAG;
PARAMETERS C0 C1 C2 C3 I0 I1 I2 I3 W0 W1 W2 W3;
PLAG=LAG(P); KLAG=LAG(K); XLAG=LAG(X);
C = C0 + C1*P + C2*LAG(P) + C3*WSUM;
I = I0 + I1*P + I2*LAG(P) + I3*LAG(K);
W = W0 + W1*X + W2*LAG(X) + W3*YEAR;
X = C + I + G;
Y = X - T;
P = X - W - T;
K = LAG(K) + I;
WSUM = W + WP;
```

The variables in the model are described:

- **ENDOGENOUS**
  - C=consumption
  - P=profits
  - I=investment
  - K=capital stock
  - Y=national income
  - X=private production
  - W=private wage bill
  - WSUM=total wage bill

- **EXOGENOUS**
  - WP=gov't wage bill
  - G=government demand
  - T=taxes

Further details on this classic model can be found in almost any textbook in econometrics.

The SYSNLIN Procedure

The SYSNLIN procedure is used to estimate the parameters in the system. The method used is either nonlinear two-stage least squares or three-stage least squares. The iteration method for solving the estimates is Gauss-Newton, which uses the derivatives of the endogenous variable with respect to the parameters. The estimates are printed out with standard errors, significance tests, and statistics of fit. The parameter estimates are stored back in the model, and can be used directly by the SIMNLIN procedure in the next step. The specification in SAS for the Klein model would be:

```
PROC SYSNLIN DATA=KDATA;
```

where KDATA is a SAS data set containing data for the Klein model.

The SIMNLIN Procedure

The SIMNLIN procedure is used to solve the system repeatedly—generating predicted values or simulation results for the system in a time series. If Newton's method is to be used the procedure automatically solves out the block-recursive form so that many small systems are solved rather than one big one.

The Gauss-Seidel method of solving the system is also available, and is preferred in large sparse systems. Simulation results can be output to a SAS data set for plots, reports, or further analysis. A series of simulation experiments can be set up easily and run through the procedure using the existing BY-group facility in SAS. If historical data is supplied to the procedure, several statistics of fit are printed.

The specification in SAS for the Klein model would be:

```
PROC SIMNLIN DATA=KDATA;
```

where KDATA supplies the time series data for the Klein model.