Chapter 18
The STATESPACE Procedure

Chapter Table of Contents

Chapter 18
The STATESPACE Procedure

Overview

The STATESPACE procedure analyzes and forecasts multivariate time series using the state space model. The STATESPACE procedure is appropriate for jointly forecasting several related time series that have dynamic interactions. By taking into account the autocorrelations among the whole set of variables, the STATESPACE procedure may give better forecasts than methods that model each series separately.

By default, the STATESPACE procedure automatically selects a state space model appropriate for the time series, making the procedure a good tool for automatic forecasting of multivariate time series. Alternatively, you can specify the state space model by giving the form of the state vector and the state transition and innovation matrices.

The methods used by the STATESPACE procedure assume that the time series are jointly stationary. Nonstationary series must be made stationary by some preliminary transformation, usually by differencing. The STATESPACE procedure allows you to specify differencing of the input data. When differencing is specified, the STATESPACE procedure automatically integrates forecasts of the differenced series to produce forecasts of the original series.

The State Space Model

The *state space model* represents a multivariate time series through auxiliary variables, some of which may not be directly observable. These auxiliary variables are called the *state vector*. The state vector summarizes all the information from the present and past values of the time series relevant to the prediction of future values of the series. The observed time series are expressed as linear combinations of the state variables. The state space model is also called a Markovian representation, or a canonical representation, of a multivariate time series process. The state space approach to modeling a multivariate stationary time series is summarized in Akaike (1976).

The state space form encompasses a very rich class of models. Any Gaussian multivariate stationary time series can be written in a state space form, provided that the dimension of the predictor space is finite. In particular, any autoregressive moving average (ARMA) process has a state space representation and, conversely, any state space process can be expressed in an ARMA form (Akaike 1974). More details on the relation of the state space and ARMA forms are given in "Relation of ARMA and State Space Forms" later in this chapter.

Let $\mathbf{x}_t$ be the $r \times 1$ vector of observed variables, after differencing (if differencing is specified) and subtracting the sample mean. Let $\mathbf{z}_t$ be the state vector of dimension s, $s \geq r$, where the first r components of $\mathbf{z}_t$ consist of $\mathbf{x}_t$. Let the notation $\mathbf{x}_{t+k|t}$

represent the conditional expectation (or prediction) of $\mathbf{x}_{t+k}$ based on the information available at time *t*. Then the last $s - r$ elements of $\mathbf{z}_t$ consist of elements of $\mathbf{x}_{t+k|t}$, where *k*>0 is specified or determined automatically by the procedure.

There are various forms of the state space model in use. The form of the state space model used by the STATESPACE procedure is based on Akaike (1976). The model is defined by the following *state transition equation*:

$$\mathbf{z}_{t+1} = \mathbf{F}\mathbf{z}_t + \mathbf{G}\mathbf{e}_{t+1}$$

In the state transition equation, the $s \times s$ coefficient matrix **F** is called the *transition matrix*; it determines the dynamic properties of the model.

The $s \times r$ coefficient matrix **G** is called the *input matrix*; it determines the variance structure of the transition equation. For model identification, the first *r* rows and columns of **G** are set to an $r \times r$ identity matrix.

The input vector $\mathbf{e}_t$ is a sequence of independent normally distributed random vectors of dimension *r* with mean **0** and covariance matrix $\Sigma_{\mathbf{ee}}$. The random error $\mathbf{e}_t$ is sometimes called the innovation vector or shock vector.

In addition to the state transition equation, state space models usually include a *measurement equation* or *observation equation* that gives the observed values $\mathbf{x}_t$ as a function of the state vector $\mathbf{z}_t$. However, since PROC STATESPACE always includes the observed values $\mathbf{x}_t$ in the state vector $\mathbf{z}_t$, the measurement equation in this case merely represents the extraction of the first *r* components of the state vector.

The measurement equation used by the STATESPACE procedure is

$$\mathbf{x}_t = [\mathbf{I}_r \mathbf{0}]\mathbf{z}_t$$

where $\mathbf{I}_r$ is an $r \times r$ identity matrix. In practice, PROC STATESPACE performs the extraction of $\mathbf{x}_t$ from $\mathbf{z}_t$ without reference to an explicit measurement equation.

In summary:

$\mathbf{x}_t$	is an observation vector of dimension *r*.
$\mathbf{z}_t$	is a state vector of dimension *s*, whose first *r* elements are $\mathbf{x}_t$ and whose last $s - r$ elements are conditional prediction of future $\mathbf{x}_t$.
F	is an $s \times s$ transition matrix.
G	is an $s \times r$ input matrix, with the identity matrix $\mathbf{I}_r$ forming the first *r* rows and columns.
$\mathbf{e}_t$	is a sequence of independent normally distributed random vectors of dimension *r* with mean **0** and covariance matrix $\Sigma_{\mathbf{ee}}$.

How PROC STATESPACE Works

The design of the STATESPACE procedure closely follows the modeling strategy proposed by Akaike (1976). This strategy employs canonical correlation analysis for the automatic identification of the state space model.

Following Akaike (1976), the procedure first fits a sequence of unrestricted vector autoregressive (VAR) models and computes Akaike's information criterion (AIC) for each model. The vector autoregressive models are estimated using the sample autocovariance matrices and the Yule-Walker equations. The order of the VAR model producing the smallest Akaike information criterion is chosen as the order (number of lags into the past) to use in the canonical correlation analysis.

The elements of the state vector are then determined via a sequence of canonical correlation analyses of the sample autocovariance matrices through the selected order. This analysis computes the sample canonical correlations of the past with an increasing number of steps into the future. Variables that yield significant correlations are added to the state vector; those that yield insignificant correlations are excluded from further consideration. The importance of the correlation is judged on the basis of another information criterion proposed by Akaike. See the section "Canonical Correlation Analysis" for details. If you specify the state vector explicitly, these model identification steps are omitted.

Once the state vector is determined, the state space model is fit to the data. The free parameters in the $\mathbf{F}$, $\mathbf{G}$, and $\mathbf{\Sigma_{ee}}$ matrices are estimated by approximate maximum likelihood. By default, the $\mathbf{F}$ and $\mathbf{G}$ matrices are unrestricted, except for identifiability requirements. Optionally, conditional least-squares estimates can be computed. You can impose restrictions on elements of the $\mathbf{F}$ and $\mathbf{G}$ matrices.

After the parameters are estimated, forecasts are produced from the fitted state space model using the Kalman filtering technique. If differencing was specified, the forecasts are integrated to produce forecasts of the original input variables.

Getting Started

The following introductory example uses simulated data for two variables X and Y. The following statements generate the X and Y series.

```
data in;
   x=10;  y=40;
   x1=0; y1=0;
   a1=0; b1=0;
   iseed=123;
   do t=-100 to 200;
      a=rannor(iseed);
      b=rannor(iseed);
      dx = 0.5*x1 + 0.3*y1 + a - 0.2*a1 - 0.1*b1;
      dy = 0.3*x1 + 0.5*y1 + b;
      x = x + dx + .25;
      y = y + dy + .25;
      if t >= 0 then output;
      x1 = dx; y1 = dy;
      a1 = a; b1 = b;
   end;
   keep t x y;
run;
```

The simulated series X and Y are shown in Figure 18.1.

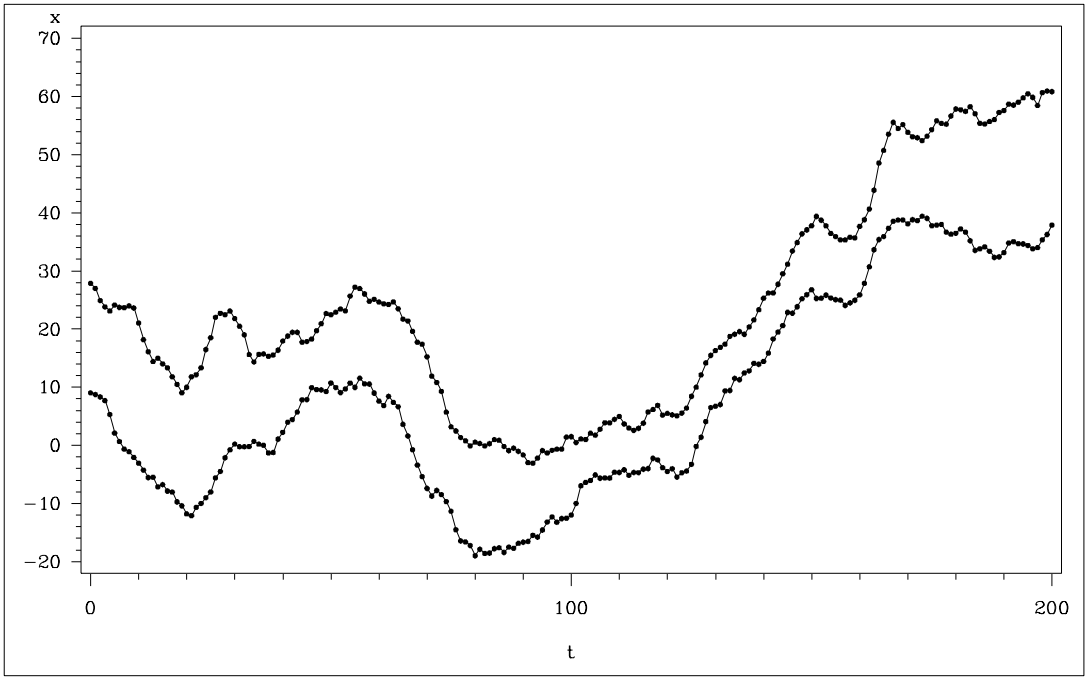

Figure 18.1. Example Series

Automatic State Space Model Selection

The STATESPACE procedure is designed to automatically select the best state space model for forecasting the series. You can specify your own model if you wish, and you can use the output from PROC STATESPACE to help you identify a state space model. However, the easiest way to use PROC STATESPACE is to let it choose the model.

Stationarity and Differencing

Although PROC STATESPACE selects the state space model automatically, it does assume that the input series are stationary. If the series are nonstationary, then the process may fail. Therefore the first step is to examine your data and test to see if differencing is required. (See the section "Stationarity and Differencing" later in this chapter for further discussion of this issue.)

The series shown in Figure 18.1 are nonstationary. In order to forecast X and Y with a state space model, you must difference them (or use some other de-trending method). If you fail to difference when needed and try to use PROC STATESPACE with nonstationary data, an inappropriate state space model may be selected, and the model estimation may fail to converge.

The following statements identify and fit a state space model for the first differences of X and Y, and forecast X and Y 10 periods ahead:

```
proc statespace data=in out=out lead=10;
   var x(1) y(1);
   id t;
run;
```

The DATA= option specifies the input data set and the OUT= option specifies the output data set for the forecasts. The LEAD= option specifies forecasting 10 observations past the end of the input data. The VAR statement specifies the variables to forecast and specifies differencing. The notation X(1) Y(1) specifies that the state space model analyzes the first differences of X and Y.

Descriptive Statistics and Preliminary Autoregressions

The first page of the printed output produced by the preceding statements is shown in Figure 18.2.

```
                          The STATESPACE Procedure

                       Number of Observations    200


                                    Standard
          Variable        Mean        Error

          x           0.144316    1.233457    Has been differenced.
                                              With period(s) = 1.
          y           0.164871    1.304358    Has been differenced.
                                              With period(s) = 1.



                          The STATESPACE Procedure

                Information Criterion for Autoregressive Models

  Lag=0    Lag=1    Lag=2    Lag=3    Lag=4    Lag=5    Lag=6    Lag=7    Lag=8

149.697 8.387786 5.517099 12.05986 15.36952 21.79538 24.00638 29.88874 33.55708

                                    Information
                                   Criterion for
                                  Autoregressive
                                     Models

                                   Lag=9      Lag=10

                                 41.17606    47.70222


                     Schematic Representation of Correlations

  Name/Lag     0     1     2     3     4     5     6     7     8     9    10

  x           ++    ++    ++    ++    ++    ++    +.    ..    +.    +.    ..
  y           ++    ++    ++    ++    ++    +.    +.    +.    +.    ..    ..

           + is > 2*std error,  - is < -2*std error,  . is between
```

Figure 18.2. Descriptive Statistics and VAR Order Selection

Descriptive statistics are printed first, giving the number of nonmissing observations after differencing, and the sample means and standard deviations of the differenced series. The sample means are subtracted before the series are modeled (unless the NOCENTER option is specified), and the sample means are added back when the forecasts are produced.

Let X_t and Y_t be the observed values of X and Y, and let x_t and y_t be the values of X and Y after differencing and subtracting the mean difference. The series $\mathbf{x}_t$ modeled by the STATEPSPACE procedure is

$$\mathbf{x}_t = \begin{bmatrix} x_t \\ y_t \end{bmatrix} = \begin{bmatrix} (1-B)X_t - 0.144316 \\ (1-B)Y_t - 0.164871 \end{bmatrix}$$

where B represents the backshift operator.

After the descriptive statistics, PROC STATESPACE prints the Akaike information criterion (AIC) values for the autoregressive models fit to the series. The smallest AIC value, in this case 5.517 at lag 2, determines the number of autocovariance matrices analyzed in the canonical correlation phase.

A schematic representation of the autocorrelations is printed next. This indicates which elements of the autocorrelation matrices at different lags are significantly greater or less than 0.

The second page of the STATESPACE printed output is shown in Figure 18.3.

```
                    The STATESPACE Procedure

          Schematic Representation of Partial Autocorrelations

Name/Lag     1     2     3     4     5     6     7     8     9    10

x           ++    +.    ..    ..    ..    ..    ..    ..    ..    ..
y           ++    ..    ..    ..    ..    ..    ..    ..    ..    ..

      + is > 2*std error,  - is < -2*std error,  . is between


               Yule-Walker Estimates for Minimum AIC

                  --------Lag=1-------    --------Lag=2-------
                         x           y           x           y

      x           0.257438    0.202237    0.170812    0.133554
      y           0.292177    0.469297    -0.00537    -0.00048
```

Figure 18.3. Partial Autocorrelations and VAR Model

Figure 18.3 shows a schematic representation of the partial autocorrelations, similar to the autocorrelations shown in Figure 18.2. The selection of a second order autoregressive model by the AIC statistic looks reasonable in this case because the partial autocorrelations for lags greater than 2 are not significant.

Next, the Yule-Walker estimates for the selected autoregressive model are printed. This output shows the coefficient matrices of the vector autoregressive model at each lag.

Selected State Space Model Form and Preliminary Estimates

After the autoregressive order selection process has determined the number of lags to consider, the canonical correlation analysis phase selects the state vector. By default, output for this process is not printed. You can use the CANCORR option to print details of the canonical correlation analysis. See the section "Canonical Correlation Analysis" later in this chapter for an explanation of this process.

Once the state vector is selected the state space model is estimated by approximate maximum likelihood. Information from the canonical correlation analysis and from the preliminary autoregression is used to form preliminary estimates of the state space model parameters. These preliminary estimates are used as starting values for the iterative estimation process.

The form of the state vector and the preliminary estimates are printed next, as shown in Figure 18.4.

```
                The STATESPACE Procedure
        Selected Statespace Form and Preliminary Estimates

                          State Vector

      x(T;T)          y(T;T)          x(T+1;T)


              Estimate of Transition Matrix

                    0               0               1
             0.291536        0.468762        -0.00411
              0.24869         0.24484        0.204257


              Input Matrix for Innovation

                            1               0
                            0               1
                     0.257438        0.202237


             Variance Matrix for Innovation

                     0.945196        0.100786
                     0.100786        1.014703
```

Figure 18.4. Preliminary Estimates of State Space Model

Figure 18.4 first prints the state vector as X[T;T] Y[T;T] X[T+1;T]. This notation indicates that the state vector is

$$\mathbf{z}_t = \begin{bmatrix} x_{t|t} \\ y_{t|t} \\ x_{t+1|t} \end{bmatrix}$$

The notation $x_{t+1|t}$ indicates the conditional expectation or prediction of x_{t+1} based on the information available at time *t*, and $x_{t|t}$ and $y_{t|t}$ are x_t and y_t respectively.

The remainder of Figure 18.4 shows the preliminary estimates of the transition matrix $\mathbf{F}$, the input matrix $\mathbf{G}$, and the covariance matrix $\mathbf{\Sigma_{ee}}$.

Estimated State Space Model

The next page of the STATESPACE output prints the final estimates of the fitted model, as shown in Figure 18.5. This output has the same form as in Figure 18.4, but shows the maximum likelihood estimates instead of the preliminary estimates.

```
          The STATESPACE Procedure
Selected Statespace Form and Fitted Model

              State Vector

x(T;T)         y(T;T)         x(T+1;T)


       Estimate of Transition Matrix

             0              0             1
      0.297273        0.47376      -0.01998
        0.2301       0.228425      0.256031


       Input Matrix for Innovation

                1              0
                0              1
         0.257284       0.202273


      Variance Matrix for Innovation

          0.945188       0.100752
          0.100752       1.014712
```

Figure 18.5. Fitted State Space Model

The estimated state space model shown in Figure 18.5 is

$$\begin{bmatrix} x_{t+1|t+1} \\ y_{t+1|t+1} \\ x_{t+2|t+1} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 \\ 0.297 & 0.474 & -0.020 \\ 0.230 & 0.228 & 0.256 \end{bmatrix} \begin{bmatrix} x_t \\ y_t \\ x_{t+1|t} \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0.257 & 0.202 \end{bmatrix} \begin{bmatrix} e_{t+1} \\ n_{t+1} \end{bmatrix}$$

$$var \begin{bmatrix} e_{t+1} \\ n_{t+1} \end{bmatrix} = \begin{bmatrix} 0.945 & 0.101 \\ 0.101 & 1.015 \end{bmatrix}$$

The next page of the STATESPACE output lists the estimates of the free parameters in the $\mathbf{F}$ and $\mathbf{G}$ matrices with standard errors and *t* statistics, as shown in Figure 18.6.

The STATESPACE Procedure

Parameter Estimates

Parameter	Estimate	Standard Error	t Value
F(2,1)	0.297273	0.129995	2.29
F(2,2)	0.473760	0.115688	4.10
F(2,3)	-0.01998	0.313025	-0.06
F(3,1)	0.230100	0.126226	1.82
F(3,2)	0.228425	0.112978	2.02
F(3,3)	0.256031	0.305256	0.84
G(3,1)	0.257284	0.071060	3.62
G(3,2)	0.202273	0.068593	2.95

Figure 18.6. Final Parameter Estimates

Convergence Failures

The maximum likelihood estimates are computed by an iterative nonlinear maximization algorithm, which may not converge. If the estimates fail to converge, warning messages are printed in the output.

If you encounter convergence problems, you should recheck the stationarity of the data and ensure that the specified differencing orders are correct. Attempting to fit state space models to nonstationary data is a common cause of convergence failure. You can also use the MAXIT= option to increase the number of iterations allowed, or experiment with the convergence tolerance options DETTOL= and PARMTOL=.

Forecast Data Set

The following statements print the output data set. The WHERE statement excludes the first 190 observations from the output, so that only the forecasts and the last 10 actual observations are printed.

```
proc print data=out;
   id t;
   where t > 190;
run;
```

The PROC PRINT output is shown in Figure 18.7.

t	x	FOR1	RES1	STD1	y	FOR2	RES2	STD2
191	34.8159	33.6299	1.18600	0.97221	58.7189	57.9916	0.72728	1.00733
192	35.0656	35.6598	-0.59419	0.97221	58.5440	59.7718	-1.22780	1.00733
193	34.7034	35.5530	-0.84962	0.97221	59.0476	58.5723	0.47522	1.00733
194	34.6626	34.7597	-0.09707	0.97221	59.7774	59.2241	0.55330	1.00733
195	34.4055	34.8322	-0.42664	0.97221	60.5118	60.1544	0.35738	1.00733
196	33.8210	34.6053	-0.78434	0.97221	59.8750	60.8260	-0.95102	1.00733
197	34.0164	33.6230	0.39333	0.97221	58.4698	59.4502	-0.98046	1.00733
198	35.3819	33.6251	1.75684	0.97221	60.6782	57.9167	2.76150	1.00733
199	36.2954	36.0528	0.24256	0.97221	60.9692	62.1637	-1.19450	1.00733
200	37.8945	37.1431	0.75142	0.97221	60.8586	61.4085	-0.54984	1.00733
201	.	38.5068	.	0.97221	.	61.3161	.	1.00733
202	.	39.0428	.	1.59125	.	61.7509	.	1.83678
203	.	39.4619	.	2.28028	.	62.1546	.	2.62366
204	.	39.8284	.	2.97824	.	62.5099	.	3.38839
205	.	40.1474	.	3.67689	.	62.8275	.	4.12805
206	.	40.4310	.	4.36299	.	63.1139	.	4.84149
207	.	40.6861	.	5.03040	.	63.3755	.	5.52744
208	.	40.9185	.	5.67548	.	63.6174	.	6.18564
209	.	41.1330	.	6.29673	.	63.8435	.	6.81655
210	.	41.3332	.	6.89383	.	64.0572	.	7.42114

Figure 18.7. OUT= Data Set Produced by PROC STATESPACE

The OUT= data set produced by PROC STATESPACE contains the VAR and ID statement variables. In addition, for each VAR statement variable, the OUT= data set contains the variables FOR*i*, RES*i*, and STD*i*. These variables contain the predicted values, residuals, and forecast standard errors for the *i*th variable in the VAR statement list. In this case, X is listed first in the VAR statement, so FOR1 contains the forecasts of X, while FOR2 contains the forecasts of Y.

The following statements plot the forecasts and actuals for the series.

```
proc gplot data=out;
   plot for1*t=1 for2*t=1 x*t=2 y*t=2 /
        overlay href=200.5;
   symbol1 v=circle i=join;
   symbol2 v=star i=none;
   where t > 150;
run;
```

The forecast plot is shown in Figure 18.8. The last 50 observations are also plotted to provide context, and a reference line is drawn between the historical and forecast periods. The actual values are plotted with asterisks.

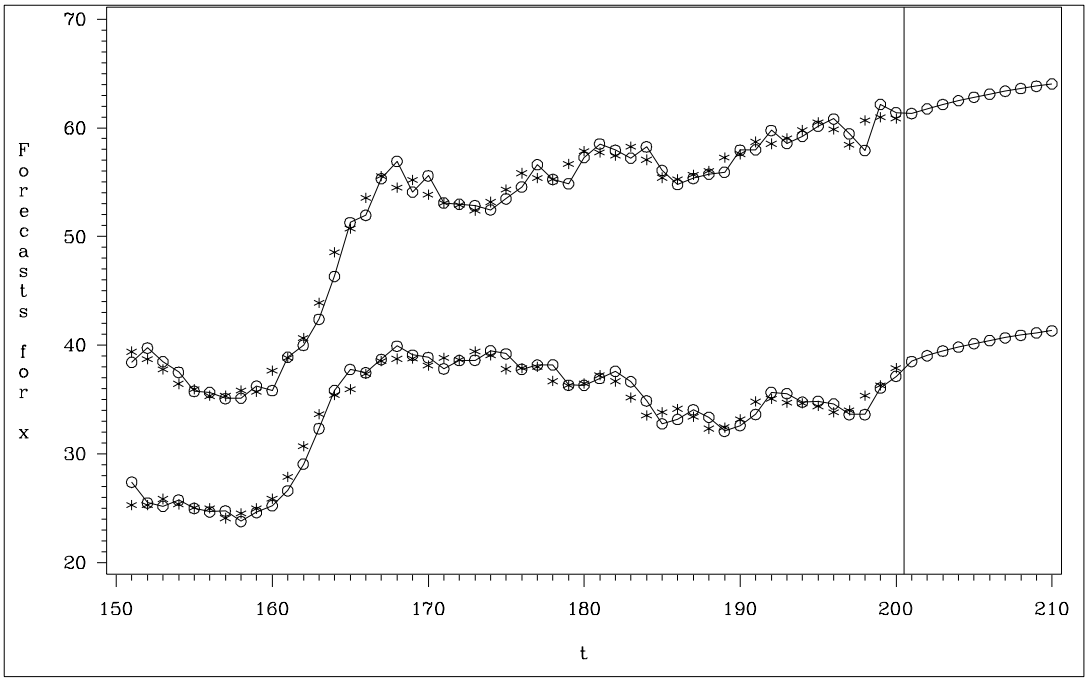

Figure 18.8. Plot of Forecasts

Controlling Printed Output

By default, the STATESPACE procedure produces a large amount of printed output. The NOPRINT option suppresses all printed output. You can suppress the printed output for the autoregressive model selection process with the PRINTOUT=NONE option. The descriptive statistics and state space model estimation output are still printed when PRINTOUT=NONE is specified. You can produce more detailed output with the PRINTOUT=LONG option and by specifying the printing control options CANCORR, COVB, and PRINT.

Specifying the State Space Model

Instead of allowing the STATESPACE procedure to select the model automatically, you can use FORM and RESTRICT statements to specify a state space model.

Specifying the State Vector

Use the FORM statement to control the form of the state vector. You can use this feature to force PROC STATESPACE to estimate and forecast a model different from the model it would select automatically. You can also use this feature to reestimate the automatically selected model (possibly with restrictions) without repeating the canonical correlation analysis.

The FORM statement specifies the number of lags of each variable to include in the state vector. For example, the statement FORM X 3; forces the state vector to include $x_{t|t}$, $x_{t+1|t}$, and $x_{t+2|t}$. The following statement specifies the state vector $(x_{t|t}, y_{t|t}, x_{t+1|t})$, which is the same state vector selected in the preceding example:

```
form x 2 y 1;
```

You can specify the form for only some of the variables and allow PROC STATESPACE to select the form for the other variables. If only some of the variables are specified in the FORM statement, canonical correlation analysis is used to determine the number of lags included in the state vector for the remaining variables not specified by the FORM statement. If the FORM statement includes specifications for all the variables listed in the VAR statement, the state vector is completely defined and the canonical correlation analysis is not performed.

Restricting the F and G matrices

After you know the form of the state vector, you can use the RESTRICT statement to fix some parameters in the **F** and **G** matrices to specified values. One use of this feature is to remove insignificant parameters by restricting them to 0.

In the introductory example shown in the preceding section, the F[2,3] parameter is not significant. (The parameters estimation output shown in Figure 18.6 gives the *t* statistic for F[2,3] as -0.06. F[3,3] and F[3,1] also have low significance with $t < 2$.)

The following statements reestimate this model with F[2,3] restricted to 0. The FORM statement is used to specify the state vector and thus bypass the canonical correlation analysis.

```
proc statespace data=in out=out lead=10;
   var x(1) y(1);
   id t;
   form x 2 y 1;
   restrict f(2,3)=0;
run;
```

The final estimates produced by these statements are shown in Figure 18.9.

The STATESPACE Procedure
Selected Statespace Form and Fitted Model

State Vector

x(T;T)	y(T;T)	x(T+1;T)

Estimate of Transition Matrix

0	0	1
0.290051	0.467468	0
0.227051	0.226139	0.26436

Input Matrix for Innovation

1	0
0	1
0.256826	0.202022

Variance Matrix for Innovation

0.945175	0.100696
0.100696	1.014733

The STATESPACE Procedure

Parameter Estimates

Parameter	Estimate	Standard Error	t Value
F(2,1)	0.290051	0.063904	4.54
F(2,2)	0.467468	0.060430	7.74
F(3,1)	0.227051	0.125221	1.81
F(3,2)	0.226139	0.111711	2.02
F(3,3)	0.264360	0.299537	0.88
G(3,1)	0.256826	0.070994	3.62
G(3,2)	0.202022	0.068507	2.95

Figure 18.9. Results using RESTRICT Statement

Syntax

The STATESPACE procedure uses the following statements:

PROC STATESPACE *options*;
BY *variable . . .* ;
FORM *variable value . . .* ;
ID *variable*;
INITIAL **F***(row,column)=value . . .* **G***(row,column)=value ...* ;
RESTRICT **F***(row,column)=value . . .* **G***(row,column)=value ...* ;
VAR *variable (difference, difference, . . .) ...* ;

Functional Summary

The statements and options used by PROC STATESPACE are summarized in the following table.

Description	Statement	Option
Input Data Set Options		
specify the input data set	PROC STATESPACE	DATA=
prevent subtraction of sample mean	PROC STATESPACE	NOCENTER
specify the ID variable	ID	
specify the observed series and differencing	VAR	
Options for Autoregressive Estimates		
specify the maximum order	PROC STATESPACE	ARMAX=
specify maximum lag for autocovariances	PROC STATESPACE	LAGMAX=
output only minimum AIC model	PROC STATESPACE	MINIC
specify the amount of detail printed	PROC STATESPACE	PRINTOUT=
write preliminary AR models to a data set	PROC STATESPACE	OUTAR=
Options for Canonical Correlation Analysis		
print the sequence of canonical correlations	PROC STATESPACE	CANCORR
specify upper limit of dimension of state vector	PROC STATESPACE	DIMMAX=
specify the minimum number of lags	PROC STATESPACE	PASTMIN=
specify the multiplier of the degrees of freedom	PROC STATESPACE	SIGCORR=
Options for State Space Model Estimation		
specify starting values	INITIAL	
print covariance matrix of parameter estimates	PROC STATESPACE	COVB
specify the convergence criterion	PROC STATESPACE	DETTOL=
specify the convergence criterion	PROC STATESPACE	PARMTOL=

Description	Statement	Option
print the details of the iterations	PROC STATESPACE	ITPRINT
specify an upper limit of the number of lags	PROC STATESPACE	KLAG=
specify maximum number of iterations allowed	PROC STATESPACE	MAXIT=
suppress the final estimation	PROC STATESPACE	NOEST
write the state space model parameter estimates to an output data set	PROC STATESPACE	OUTMODEL=
use conditional least squares for final estimates	PROC STATESPACE	RESIDEST
specify criterion for testing for singularity	PROC STATESPACE	SINGULAR=
Options for Forecasting		
start forecasting before end of the input data	PROC STATESPACE	BACK=
specify the time interval between observations	PROC STATESPACE	INTERVAL=
specify multiple periods in the time series	PROC STATESPACE	INTPER=
specify how many periods to forecast	PROC STATESPACE	LEAD=
specify the output data set for forecasts	PROC STATESPACE	OUT=
print forecasts	PROC STATESPACE	PRINT
Options to Specify the State Space Model		
specify the state vector	FORM	
specify the parameter values	RESTRICT	
BY Groups		
specify BY-group processing	BY	
Printing		
suppresses all printed output	NOPRINT	

PROC STATESPACE Statement

PROC STATESPACE *options*;

The following options can be specified in the PROC STATESPACE statement.

Printing Options

NOPRINT

suppresses all printed output.

Input Data Options

DATA= *SAS-data-set*

specifies the name of the SAS data set to be used by the procedure. If the DATA= option is omitted, the most recently created SAS data set is used.

LAGMAX= *k*

specifies the number of lags for which the sample autocovariance matrix is computed. The LAGMAX= option controls the number of lags printed in the schematic representation of the autocorrelations.

The sample autocovariance matrix of lag *i*, denoted as $\mathbf{C}_i$, is computed as

$$\mathbf{C}_i = \frac{1}{N-1} \sum_{t=1+i}^{N} \mathbf{x}_t \mathbf{x}'_{t-i}$$

where $\mathbf{x}_t$ is the differenced and centered data and N is the number of observations. (If the NOCENTER option is specified, 1 is not subtracted from N.) LAGMAX= *k* specifies that $\mathbf{C}_0$ through $\mathbf{C}_k$ are computed. The default is LAGMAX=10.

NOCENTER

prevents subtraction of the sample mean from the input series (after any specified differencing) before the analysis.

Options for Preliminary Autoregressive Models

ARMAX= *n*

specifies the maximum order of the preliminary autoregressive models. The ARMAX= option controls the autoregressive orders for which information criteria are printed, and controls the number of lags printed in the schematic representation of partial autocorrelations. The default is ARMAX=10. See "Preliminary Autoregressive Models" later in this chapter for details.

MINIC

writes to the OUTAR= data set only the preliminary Yule-Walker estimates for the VAR model producing the minimum AIC. See "OUTAR= Data Set" later in this chapter for details.

OUTAR= *SAS-data-set*

writes the Yule-Walker estimates of the preliminary autoregressive models to a SAS data set. See "OUTAR= Data Set" later in this chapter for details.

PRINTOUT= SHORT | LONG | NONE

determines the amount of detail printed. PRINTOUT=LONG prints the lagged covariance matrices, the partial autoregressive matrices, and estimates of the residual covariance matrices from the sequence of autoregressive models. PRINTOUT=NONE suppresses the output for the preliminary autoregressive models. The descriptive statistics and state space model estimation output are still printed when PRINTOUT=NONE is specified. PRINTOUT=SHORT is the default.

Canonical Correlation Analysis Options

CANCORR

prints the canonical correlations and information criterion for each candidate state vector considered. See "Canonical Correlation Analysis" later in this chapter for details.

DIMMAX= *n*

specifies the upper limit to the dimension of the state vector. The DIMMAX= option can be used to limit the size of the model selected. The default is DIMMAX=10.

PASTMIN= *n*

specifies the minimum number of lags to include in the canonical correlation analysis. The default is PASTMIN=0. See "Canonical Correlation Analysis" later in this chapter for details.

SIGCORR= *value*

specifies the multiplier of the degrees of freedom for the penalty term in the information criterion used to select the state space form. The default is SIGCORR=2. The larger the value of the SIGCORR= option, the smaller the state vector tends to be. Hence, a large value causes a simpler model to be fit. See "Canonical Correlations Analysis" later in this chapter for details.

State Space Model Estimation Options

COVB

prints the inverse of the observed information matrix for the parameter estimates. This matrix is an estimate of the covariance matrix for the parameter estimates.

DETTOL= *value*

specifies the convergence criterion. The DETTOL= and PARMTOL= option values are used together to test for convergence of the estimation process. If, during an iteration, the relative change of the parameter estimates is less than the PARMTOL= value and the relative change of the determinant of the innovation variance matrix is less than the DETTOL= value, then iteration ceases and the current estimates are accepted. The default is DETTOL=1E-5.

ITPRINT

prints the iterations during the estimation process.

KLAG= *n*

sets an upper limit for the number of lags of the sample autocovariance matrix used in computing the approximate likelihood function. If the data have a strong moving average character, a larger KLAG= value may be necessary to obtain good estimates. The default is KLAG=15. See "Parameter Estimation" later in this chapter for details.

MAXIT= *n*

sets an upper limit to the number of iterations in the maximum likelihood or conditional least-squares estimation. The default is MAXIT=50.

NOEST

suppresses the final maximum likelihood estimation of the selected model.

OUTMODEL= *SAS-data-set*

writes the parameter estimates and their standard errors to a SAS data set. See "OUTMODEL= Data Set" later in this chapter for details.

PARMTOL= *value*

specifies the convergence criterion. The DETTOL= and PARMTOL= option values are used together to test for convergence of the estimation process. If, during an

iteration, the relative change of the parameter estimates is less than the PARMTOL= value and the relative change of the determinant of the innovation variance matrix is less than the DETTOL= value, then iteration ceases and the current estimates are accepted. The default is PARMTOL=.001.

RESIDEST

computes the final estimates using conditional least squares on the raw data. This type of estimation may be more stable than the default maximum likelihood method but is usually more computationally expensive. See "Parameter Estimation" later in this chapter for details of the conditional least squares method.

SINGULAR= *value*

specifies the criterion for testing for singularity of a matrix. A matrix is declared singular if a scaled pivot is less than the SINGULAR= value when sweeping the matrix. The default is SINGULAR=1E-7.

Forecasting Options

BACK= *n*

starts forecasting *n* periods before the end of the input data. The BACK= option value must not be greater than the number of observations. The default is BACK=0.

INTERVAL= *interval*

specifies the time interval between observations. The INTERVAL= value is used in conjunction with the ID variable to check that the input data are in order and have no missing periods. The INTERVAL= option is also used to extrapolate the ID values past the end of the input data. See Chapter 3, “Date Intervals, Formats, and Functions,” for details on the INTERVAL= values allowed.

INTPER= *n*

specifies that each input observation corresponds to *n* time periods. For example, the options INTERVAL=MONTH and INTPER=2 specify bimonthly data and are equivalent to specifying INTERVAL=MONTH2. If the INTERVAL= option is not specified, the INTPER= option controls the increment used to generate ID values for the forecast observations. The default is INTPER=1.

LEAD= *n*

specifies how many forecast observations are produced. The forecasts start at the point set by the BACK= option. The default is LEAD=0, which produces no forecasts.

OUT= *SAS-data-set*

writes the residuals, actual values, forecasts, and forecast standard errors to a SAS data set. See "OUT= Data Set" later in this chapter for details.

PRINT

prints the forecasts.

BY Statement

BY *variable ...* **;**

A BY statement can be used with the STATESPACE procedure to obtain separate analyses on observations in groups defined by the BY variables.

FORM Statement

FORM *variable value ...* **;**

The FORM statement specifies the number of times a variable is included in the state vector. Values can be specified for any variable listed in the VAR statement. If a value is specified for each variable in the VAR statement, the state vector for the state space model is entirely specified, and automatic selection of the state space model is not performed.

The FORM statement forces the state vector, $\mathbf{z}_t$, to contain a specific variable a given number of times. For example, if Y is one of the variables in $\mathbf{x}_t$, then the statement

```
form y 3;
```

forces the state vector to contain $Y_t, Y_{t+1|t}$, and $Y_{t+2|t}$, possibly along with other variables.

The following statements illustrate the use of the FORM statement:

```
proc statespace data=in;
   var x y;
   form x 3 y 2;
run;
```

These statements fit a state space model with the following state vector:

$$\mathbf{z_t} = \begin{bmatrix} x_{t|t} \\ y_{t|t} \\ x_{t+1|t} \\ y_{t+1|t} \\ x_{t+2|t} \end{bmatrix}$$

ID Statement

ID *variable*;

The ID statement specifies a variable that identifies observations in the input data set. The variable specified in the ID statement is included in the OUT= data set. The values of the ID variable are extrapolated for the forecast observations based on the values of the INTERVAL= and INTPER= options.

INITIAL Statement

INITIAL F *(row,column)*= *value* **...** **G***(row, column)*= *value* **... ;**

The INITIAL statement gives initial values to the specified elements of the **F** and **G** matrices. These initial values are used as starting values for the iterative estimation.

Parts of the **F** and **G** matrices represent fixed structural identities. If an element specified is a fixed structural element instead of a free parameter, the corresponding initialization is ignored.

The following is an example of an INITIAL statement:

```
initial f(3,2)=0 g(4,1)=0 g(5,1)=0;
```

RESTRICT Statement

RESTRICT F*(row,column)*= *value* **...** **G***(row,column)*= *value* **... ;**

The RESTRICT statement restricts the specified elements of the **F** and **G** matrices to the specified values.

To use the restrict statement you need to know the form of the model. Either specify the form of the model with the FORM statement, or do a preliminary run, perhaps with the NOEST option, to find the form of the model that PROC STATESPACE selects for the data.

The following is an example of a RESTRICT statement:

```
restrict f(3,2)=0 g(4,1)=0 g(5,1)=0 ;
```

Parts of the **F** and **G** matrices represent fixed structural identities. If a restriction is specified for an element that is a fixed structural element instead of a free parameter, the restriction is ignored.

VAR Statement

VAR *variable (difference, difference, ...) ...* **;**

The VAR statement specifies the variables in the input data set to model and forecast. The VAR statement also specifies differencing of the input variables. The VAR statement is required.

Differencing is specified by following the variable name with a list of difference periods separated by commas. See the section "Stationarity and Differencing" for more information on differencing of input variables.

The order in which variables are listed in the VAR statement controls the order in which variables are included in the state vector. Usually, potential inputs should be listed before potential outputs.

For example, assuming the input data are monthly, the following VAR statement specifies modeling and forecasting of the one period and seasonal second difference of X and Y:

```
var x(1,12) y(1,12);
```

In this example, the vector time series analyzed is

$$\mathbf{x}_t = \begin{bmatrix} (1-B)(1-B^{12})X_t - \overline{x} \\ (1-B)(1-B^{12})Y_t - \overline{y} \end{bmatrix}$$

where B represents the back shift operator, and $\overline{x}$ and $\overline{y}$ represent the means of the differenced series. If the NOCENTER option is specified the mean differences are not subtracted.

Details

Missing Values

The STATESPACE procedure does not support missing values. The procedure uses the first contiguous group of observations with no missing values for any of the VAR statement variables. Observations at the beginning of the data set with missing values for any VAR statement variable are not used or included in the output data set.

Stationarity and Differencing

The state space model used by the STATESPACE procedure assumes that the time series are stationary. Hence, the data should be checked for stationarity. One way to check for stationarity is to plot the series. A graph of series over time can show a time trend or variability changes.

You can also check stationarity by using the sample autocorrelation functions displayed by the ARIMA procedure. The autocorrelation functions of nonstationary series tend to decay slowly. See Chapter 7, "The ARIMA Procedure," for more information.

Another alternative is to use the STATIONARITY= option on the IDENTIFY statement in PROC ARIMA to apply Dickey-Fuller tests for unit roots in the time series. See Chapter 7, "The ARIMA Procedure," for more information on Dickey-Fuller unit root tests.

The most popular way to transform a nonstationary series to stationarity is by differencing. Differencing of the time series is specified in the VAR statement. For example, to take a simple first difference of the series X, use this statement:

```
var x(1);
```

In this example, the change in X from one period to the next is analyzed. When the series has a seasonal pattern, differencing at a period equal to the length of the seasonal cycle may be desirable. For example, suppose the variable X is measured quarterly and shows a seasonal cycle over the year. You can use the following statement to analyze the series of changes from the same quarter in the previous year:

```
var x(4);
```

To difference twice, add another differencing period to the list. For example, the following statement analyzes the series of second differences $(X_t - X_{t-1}) - (X_{t-1} - X_{t-2}) = X_t - 2X_{t-1} + X_{t-2}$:

```
var x(1,1);
```

The following statement analyzes the seasonal second difference series.

```
var x(1,4);
```

The series modeled is the 1-period difference of the 4-period difference: $(X_t - X_{t-4}) - (X_{t-1} - X_{t-5}) = X_t - X_{t-1} - X_{t-4} + X_{t-5}$.

Another way to obtain stationary series is to use a regression on time to de-trend the data. If the time series has a deterministic linear trend, regressing the series on time produces residuals that should be stationary. The following statements write residuals of X and Y to the variable RX and RY in the output data set DETREND.

```
data a;
   set a;
   t=_n_;
run;

proc reg data=a;
   model x y = t;
   output out=detrend r=rx ry;
run;
```

You then use PROC STATESPACE to forecast the de-trended series RX and RY. A disadvantage of this method is that you need to add the trend back to the forecast series in an additional step. A more serious disadvantage of the de-trending method is that it assumes a deterministic trend. In practice, most time series appear to have a stochastic rather than a deterministic trend. Differencing is a more flexible and often more appropriate method.

There are several other methods to handle nonstationary time series. For more information and examples, refer to Brockwell and Davis (1991).

Preliminary Autoregressive Models

After computing the sample autocovariance matrices, PROC STATESPACE fits a sequence of vector autoregressive models. These preliminary autoregressive models are used to estimate the autoregressive order of the process and limit the order of the autocovariances considered in the state vector selection process.

Yule-Walker Equations for Forward and Backward Models

Unlike a univariate autoregressive model, a multivariate autoregressive model has different forms, depending on whether the present observation is being predicted from the past observations or from the future observations.

Let $\mathbf{x}_t$ be the *r*-component stationary time series given by the VAR statement after differencing and subtracting the vector of sample means. (If the NOCENTER option is specified, the mean is not subtracted.) Let *n* be the number of observations of $\mathbf{x}_t$ from the input data set.

Let $\mathbf{e}_t$ be a vector white noise sequence with mean vector $\mathbf{0}$ and variance matrix $\mathbf{\Sigma}_p$, and let $\mathbf{n}_t$ be a vector white noise sequence with mean vector $\mathbf{0}$ and variance matrix $\mathbf{\Omega}_p$. Let *p* be the order of the vector autoregressive model for $\mathbf{x}_t$.

The forward autoregressive form based on the past observations is written as follows:

$$\mathbf{x}_t = \sum_{i=1}^{p} \mathbf{\Phi}_i^p \mathbf{x}_{t-i} + \mathbf{e}_t$$

The backward autoregressive form based on the future observations is written as follows:

$$\mathbf{x}_t = \sum_{i=1}^{p} \mathbf{\Psi}_i^p \mathbf{x}_{t+i} + \mathbf{n}_t$$

Letting E denote the expected value operator, the autocovariance sequence for the $\mathbf{x}_t$ series, $\mathbf{\Gamma}_i$, is

$$\mathbf{\Gamma}_i = E\mathbf{x}_t\mathbf{x}_{t-i}'$$

The Yule-Walker equations for the autoregressive model that matches the first *p* elements of the autocovariance sequence are

$$\begin{bmatrix} \mathbf{\Gamma}_0 & \mathbf{\Gamma}_1 & \cdots & \mathbf{\Gamma}_{p-1} \\ \mathbf{\Gamma}_1' & \mathbf{\Gamma}_0 & \cdots & \mathbf{\Gamma}_{p-2} \\ \vdots & \vdots & & \vdots \\ \mathbf{\Gamma}_{p-1}' & \mathbf{\Gamma}_{p-2}' & \cdots & \mathbf{\Gamma}_0 \end{bmatrix} \begin{bmatrix} \mathbf{\Phi}_1^p \\ \mathbf{\Phi}_2^p \\ \vdots \\ \mathbf{\Phi}_p^p \end{bmatrix} = \begin{bmatrix} \mathbf{\Gamma}_1 \\ \mathbf{\Gamma}_2 \\ \vdots \\ \mathbf{\Gamma}_p \end{bmatrix}$$

and

$$\begin{bmatrix} \mathbf{\Gamma}_0 & \mathbf{\Gamma}_1' & \cdots & \mathbf{\Gamma}_{p-1}' \\ \mathbf{\Gamma}_1 & \mathbf{\Gamma}_0 & \cdots & \mathbf{\Gamma}_{p-2}' \\ \vdots & \vdots & & \vdots \\ \mathbf{\Gamma}_{p-1} & \mathbf{\Gamma}_{p-2} & \cdots & \mathbf{\Gamma}_0 \end{bmatrix} \begin{bmatrix} \mathbf{\Psi}_1^p \\ \mathbf{\Psi}_2^p \\ \vdots \\ \mathbf{\Psi}_p^p \end{bmatrix} = \begin{bmatrix} \mathbf{\Gamma}_1' \\ \mathbf{\Gamma}_2' \\ \vdots \\ \mathbf{\Gamma}_p' \end{bmatrix}$$

Here $\mathbf{\Phi}_i^p$ are the coefficient matrices for the past observation form of the vector autoregressive model, and $\mathbf{\Psi}_i^p$ are the coefficient matrices for the future observation form. More information on the Yule-Walker equations in the multivariate setting can be found in Whittle (1963) and Ansley and Newbold (1979).

The innovation variance matrices for the two forms can be written as follows:

$$\mathbf{\Sigma}_p = \mathbf{\Gamma}_0 - \sum_{i=1}^{p} \mathbf{\Phi}_i^p \mathbf{\Gamma}_i'$$

$$\mathbf{\Omega}_p = \mathbf{\Gamma}_0 - \sum_{i=1}^{p} \mathbf{\Psi}_i^p \mathbf{\Gamma}_i$$

The autoregressive models are fit to the data using the preceding Yule-Walker equations with $\mathbf{\Gamma}_i$ replaced by the sample covariance sequence $\mathbf{C_i}$. The covariance matrices are calculated as

$$\mathbf{C}_i = \frac{1}{N-1} \sum_{t=i+1}^{N} \mathbf{x}_t \mathbf{x}'_{t-i}$$

Let $\widehat{\mathbf{\Phi}}_p$, $\widehat{\mathbf{\Psi}}_p$, $\widehat{\mathbf{\Sigma}}_p$, and $\widehat{\mathbf{\Omega}}_p$ represent the Yule-Walker estimates of $\mathbf{\Phi}_p$, $\mathbf{\Psi}_p$, $\mathbf{\Sigma}_p$, and $\mathbf{\Omega}_p$ respectively. These matrices are written to an output data set when the OUTAR= option is specified.

When the PRINTOUT=LONG option is specified, the sequence of matrices $\widehat{\mathbf{\Sigma}}_p$ and the corresponding correlation matrices are printed. The sequence of matrices $\widehat{\mathbf{\Sigma}}_p$ is used to compute Akaike information criteria for selection of the autoregressive order of the process.

Akaike Information Criterion

The Akaike information criterion, or AIC, is defined as -2(*maximum of log likelihood*)+2(*number of parameters*). Since the vector autoregressive models are estimates from the Yule-Walker equations, not by maximum likelihood, the exact likelihood values are not available for computing the AIC. However, for the vector autoregressive model the maximum of the log likelihood can be approximated as

$$\ln(L) \approx -\frac{n}{2} \ln(|\widehat{\mathbf{\Sigma}}_p|)$$

Thus, the AIC for the order *p* model is computed as

$$AIC_p = n\ln(|\widehat{\mathbf{\Sigma}}_p|) + 2pr^2$$

You can use the printed AIC array to compute a likelihood ratio test of the autoregressive order. The log-likelihood ratio test statistic for testing the order *p* model against the order $p-1$ model is

$$-n\ln(|\widehat{\mathbf{\Sigma}}_p|) + n\ln(|\widehat{\mathbf{\Sigma}}_{p-1}|)$$

This quantity is asymptotically distributed as a χ^2 with r^2 degrees of freedom if the series is autoregressive of order $p-1$. It can be computed from the AIC array as

$$AIC_{p-1} - AIC_p + 2r^2$$

You can evaluate the significance of these test statistics with the PROBCHI function in a SAS DATA step, or with a χ^2 table.

Determining the Autoregressive Order

Although the autoregressive models can be used for prediction, their primary value is to aid in the selection of a suitable portion of the sample covariance matrix for use in computing canonical correlations. If the multivariate time series $\mathbf{x}_t$ is of autoregressive order *p*, then the vector of past values to lag *p* is considered to contain essentially all the information relevant for prediction of future values of the time series.

By default, PROC STATESPACE selects the order, *p*, producing the autoregressive model with the smallest AIC_p. If the value *p* for the minimum AIC_p is less than the value of the PASTMIN= option, then *p* is set to the PASTMIN= value. Alternatively, you can use the ARMAX= and PASTMIN= options to force PROC STATESPACE to use an order you select.

Significance Limits for Partial Autocorrelations

The STATESPACE procedure prints a schematic representation of the partial autocorrelation matrices indicating which partial autocorrelations are significantly greater or significantly less than 0. Figure 18.10 shows an example of this table.

```
                    The STATESPACE Procedure

          Schematic Representation of Partial Autocorrelations

Name/Lag     1     2     3     4     5     6     7     8     9    10

x           ++    +.    ..    ..    ..    ..    ..    ..    ..    ..
y           ++    ..    ..    ..    ..    ..    ..    ..    ..    ..

      + is > 2*std error,  - is < -2*std error,  . is between
```

Figure 18.10. Significant Partial Autocorrelations

The partial autocorrelations are from the sample partial autoregressive matrices $\widehat{\mathbf{\Phi}}_p^p$. The standard errors used for the significance limits of the partial autocorrelations are computed from the sequence of matrices $\mathbf{\Sigma}_p$ and $\mathbf{\Omega}_p$.

Under the assumption that the observed series arises from an autoregressive process of order $p-1$, the *p*th sample partial autoregressive matrix $\widehat{\mathbf{\Phi}}_p^p$ has an asymptotic variance matrix $\frac{1}{n}\mathbf{\Omega}_p^{-1} \otimes \mathbf{\Sigma}_p$.

The significance limits for $\widehat{\mathbf{\Phi}}_p^p$ used in the schematic plot of the sample partial autoregressive sequence are derived by replacing $\mathbf{\Omega}_p$ and $\mathbf{\Sigma}_p$ with their sample estimators to produce the variance estimate, as follows:

$$\widehat{Var}\left(\widehat{\mathbf{\Phi}}_p^p\right) = \left(\frac{1}{n-rp}\right) \widehat{\mathbf{\Omega}}_p^{-1} \otimes \widehat{\mathbf{\Sigma}}_p$$

Canonical Correlation Analysis

Given the order *p*, let $\mathbf{p}_t$ be the vector of current and past values relevant to prediction of $\mathbf{x}_{t+1}$:

$$\mathbf{p}_t = (\mathbf{x}_t', \mathbf{x}_{t-1}', \cdots, \mathbf{x}_{t-p}')'$$

Let $\mathbf{f}_t$ be the vector of current and future values:

$$\mathbf{f}_t = (\mathbf{x}_t', \mathbf{x}_{t+1}', \cdots, \mathbf{x}_{t+p}')'$$

In the canonical correlation analysis, consider submatrices of the sample covariance matrix of $\mathbf{p}_t$ and $\mathbf{f}_t$. This covariance matrix, $\mathbf{V}$, has a block Hankel form:

$$\mathbf{V} = \begin{bmatrix} \mathbf{C}_0 & \mathbf{C}_1' & \mathbf{C}_2' & \cdots & \mathbf{C}_p' \\ \mathbf{C}_1' & \mathbf{C}_2' & \mathbf{C}_3' & \cdots & \mathbf{C}_{p+1}' \\ \vdots & \vdots & \vdots & & \vdots \\ \mathbf{C}_p' & \mathbf{C}_{p+1}' & \mathbf{C}_{p+2}' & \cdots & \mathbf{C}_{2p}' \end{bmatrix}$$

State Vector Selection Process

The canonical correlation analysis forms a sequence of potential state vectors, $\mathbf{z}_t^j$. Examine a sequence, $\mathbf{f}_t^j$, of subvectors of $\mathbf{f}_t$, and form the submatrix, $\mathbf{V}^j$, consisting of the rows and columns of $\mathbf{V}$ corresponding to the components of $\mathbf{f}_t^j$, and compute its canonical correlations.

The smallest canonical correlation of $\mathbf{V}^j$ is then used in the selection of the components of the state vector. The selection process is described in the following. For more details about this process, refer to Akaike (1976).

In the following discussion, the notation $\mathbf{x}_{t+k|t}$ denotes the wide sense conditional expectation (best linear predictor) of $\mathbf{x}_{t+k}$, given all $\mathbf{x}_s$ with *s* less than or equal to *t*. In the notation $x_{i,t+1}$, the first subscript denotes the *i*th component of $\mathbf{x}_{t+1}$.

The initial state vector $\mathbf{z}_t^1$ is set to $\mathbf{x}_t$. The sequence $\mathbf{f}_t^j$ is initialized by setting

$$\mathbf{f}_t^1 = (\mathbf{z}_t^{1'}, x_{1,t+1|t})' = (\mathbf{x}_t', x_{1,t+1|t})'$$

That is, start by considering whether to add $x_{1,t+1|t}$ to the initial state vector $\mathbf{z}_t^1$.

The procedure forms the submatrix $\mathbf{V}^1$ corresponding to $\mathbf{f}_t^1$ and computes its canonical correlations. Denote the smallest canonical correlation of $\mathbf{V}^1$ as ρ_{min}. If ρ_{min} is significantly greater than 0, $x_{1,t+1|t}$ is added to the state vector.

If the smallest canonical correlation of $\mathbf{V}^1$ is not significantly greater than 0, then a linear combination of $\mathbf{f}_t^1$ is uncorrelated with the past, $\mathbf{p}_t$. Assuming that the determinant of $\mathbf{C}_0$ is not 0, (that is, no input series is a constant), you can take the coefficient

of $x_{1,t+1|t}$ in this linear combination to be 1. Denote the coefficients of $\mathbf{z}_t^1$ in this linear combination as ℓ. This gives the relationship:

$$x_{1,t+1|t} = \ell' \mathbf{x}_t$$

Therefore, the current state vector already contains all the past information useful for predicting $x_{1,t+1}$ and any greater leads of $x_{1,t}$. The variable $x_{1,t+1|t}$ is not added to the state vector, nor are any terms $x_{1,t+k|t}$ considered as possible components of the state vector. The variable x_1 is no longer active for state vector selection.

The process described for $x_{1,t+1|t}$ is repeated for the remaining elements of $\mathbf{f}_t$. The next candidate for inclusion in the state vector is the next component of $\mathbf{f}_t$ corresponding to an active variable. Components of $\mathbf{f}_t$ corresponding to inactive variables that produced a zero ρ_{min} in a previous step are skipped.

Denote the next candidate as $x_{l,t+k|t}$. The vector $\mathbf{f}_t^j$ is formed from the current state vector and $x_{l,t+k|t}$ as follows:

$$\mathbf{f}_t^j = (\mathbf{z}_t^{j'}, x_{l,t+k|t})'$$

The matrix $\mathbf{V}^j$ is formed from $\mathbf{f}_t^j$ and its canonical correlations are computed. The smallest canonical correlation of $\mathbf{V}^j$ is judged to be either greater than or equal to 0. If it is judged to be greater than 0, $x_{l,t+k|t}$ is added to the state vector. If it is judged to be 0, then a linear combination of $\mathbf{f}_t^j$ is uncorrelated with the $\mathbf{p}_t$, and the variable x_l is now inactive.

The state vector selection process continues until no active variables remain.

Testing Significance of Canonical Correlations

For each step in the canonical correlation sequence, the significance of the smallest canonical correlation, ρ_{min}, is judged by an information criterion from Akaike (1976). This information criterion is

$$-n\ln(1-\rho_{min}^2) - \lambda(r(p+1) - q + 1)$$

where q is the dimension of $\mathbf{f}_t^j$ at the current step, r is the order of the state vector, p is the order of the vector autoregressive process, and λ is the value of the SIGCORR= option. The default is SIGCORR=2. If this information criterion is less than or equal to 0, ρ_{min} is taken to be 0; otherwise, it is taken to be significantly greater than 0. (Do not confuse this information criterion with the AIC.)

Variables in $\mathbf{x}_{t+p|t}$ are not added in the model, even with positive information criterion, because of the singularity of $\mathbf{V}$. You can force the consideration of more candidate state variables by increasing the size of the $\mathbf{V}$ matrix by specifying a PASTMIN= option value larger than p.

Printing the Canonical Correlations

To print the details of the canonical correlation analysis process, specify the CANCORR option in the PROC STATESPACE statement. The CANCORR option prints

the candidate state vectors, the canonical correlations, and the information criteria for testing the significance of the smallest canonical correlation.

Bartlett's χ^2 and its degrees of freedom are also printed when the CANCORR option is specified. The formula used for Bartlett's χ^2 is

$$\chi^2 = -(n - .5(r(p+1) - q + 1))\ln(1 - \rho^2_{min})$$

with $r(p+1) - q + 1$ degrees of freedom.

Figure 18.11 shows the output of the CANCORR option for the introductory example shown in the "Getting Started" section of this chapter.

The STATESPACE Procedure
Canonical Correlations Analysis

x(T;T)	y(T;T)	x(T+1;T)	Information Criterion	Chi Square	DF
1	1	0.237045	3.566167	11.4505	4

x(T;T)	y(T;T)	x(T+1;T)	y(T+1;T)	Information Criterion	Chi Square	DF
1	1	0.238244	0.056565	-5.35906	0.636134	3

x(T;T)	y(T;T)	x(T+1;T)	x(T+2;T)	Information Criterion	Chi Square	DF
1	1	0.237602	0.087493	-4.46312	1.525353	3

Figure 18.11. Canonical Correlations Analysis

New variables are added to the state vector if the information criteria are positive. In this example, $\mathbf{y}_{t+1|t}$ and $\mathbf{x}_{t+2|t}$ are not added to the state space vector because the information criteria for these models are negative.

If the information criterion is nearly 0, then you may want to investigate models that arise if the opposite decision is made regarding ρ_{min}. This investigation can be accomplished by using a FORM statement to specify part or all of the state vector.

Preliminary Estimates of $\mathbf{F}$

When a candidate variable $x_{l,t+k|t}$ yields a zero ρ_{min} and is not added to the state vector, a linear combination of $\mathbf{f}^j_t$ is uncorrelated with the $\mathbf{p}_t$. Because of the method used to construct the $\mathbf{f}^j_t$ sequence, the coefficient of $x_{l,t+k|t}$ in $\mathbf{l}$ can be taken as 1. Denote the coefficients of $\mathbf{z}^j_t$ in this linear combination as $\mathbf{l}$.

This gives the relationship:

$$x_{l,t+k|t} = \mathbf{l}'\mathbf{z}^j_t$$

The vector $\mathbf{l}$ is used as a preliminary estimate of the first *r* columns of the row of the transition matrix $\mathbf{F}$ corresponding to $x_{l,t+k-1|t}$.

Parameter Estimation

The model is $\mathbf{z}_{t+1} = \mathbf{F}\mathbf{z}_t + \mathbf{G}\mathbf{e}_{t+1}$, where $\mathbf{e}_t$ is a sequence of independent multivariate normal innovations with mean vector **0** and variance $\Sigma_{\mathbf{ee}}$. The observed sequence, $\mathbf{x}_t$, composes the first *r* components of $\mathbf{z}_t$ and, thus, $\mathbf{x}_t = \mathbf{H}\mathbf{z}_t$, where $\mathbf{H}$ is the $r \times s$ matrix $[\mathbf{I}_r\ \mathbf{0}]$.

Let $\mathbf{E}$ be the $r \times n$ matrix of innovations:

$$\mathbf{E} = [\ \mathbf{e}_1 \quad \cdots \quad \mathbf{e}_n\ ]$$

If the number of observations, *n*, is reasonably large, the log likelihood, L, can be approximated up to an additive constant as follows:

$$L = -\frac{n}{2}\ln(|\mathbf{\Sigma_{ee}}|) - \frac{1}{2}trace(\mathbf{\Sigma_{ee}^{-1}}\mathbf{E}\mathbf{E}')$$

The elements of $\mathbf{\Sigma_{ee}}$ are taken as free parameters and are estimated as follows:

$$\mathbf{S}_0 = \frac{1}{n}\mathbf{E}\mathbf{E}'$$

Replacing $\mathbf{\Sigma_{ee}}$ by $\mathbf{S}_0$ in the likelihood equation, the log likelihood, up to an additive constant, is

$$\mathbf{L} = -\frac{n}{2}\ln(|\mathbf{S}_0|)$$

Letting B be the backshift operator, the formal relation between $\mathbf{x}_t$ and $\mathbf{e}_t$ is

$$\mathbf{x}_t = \mathbf{H}(\mathbf{I} - B\mathbf{F})^{-1}\mathbf{G}\mathbf{e}_t$$

$$\mathbf{e}_t = (\mathbf{H}(\mathbf{I} - B\mathbf{F})^{-1}\mathbf{G})^{-1}\mathbf{x}_t = \sum_{i=0}^{\infty}\mathbf{\Xi}_i\mathbf{x}_{t-i}$$

Letting $\mathbf{C}_i$ be the *i*th lagged sample covariance of $\mathbf{x}_t$, and neglecting end effects, the matrix $\mathbf{S}_0$ is

$$\mathbf{S}_0 = \sum_{i,j=0}^{\infty}\mathbf{\Xi}_i\mathbf{C}_{-i+j}\mathbf{\Xi}_j^{'}$$

For the computation of $\mathbf{S}_0$, the infinite sum is truncated at the value of the KLAG= option. The value of the KLAG= option should be large enough that the sequence $\mathbf{\Xi}_i$ is approximately 0 beyond that point.

Let θ be the vector of free parameters in the $\mathbf{F}$ and $\mathbf{G}$ matrices. The derivative of the log likelihood with respect to the parameter θ is

$$\frac{\partial L}{\partial \theta} = -\frac{n}{2} \operatorname{trace}\left(\mathbf{S}_0^{-1} \frac{\partial \mathbf{S}_0}{\partial \theta}\right)$$

The second derivative is

$$\frac{\partial^2 \mathbf{L}}{\partial \theta \partial \theta'} = \frac{n}{2}\left(\operatorname{trace}\left(\mathbf{S}_0^{-1} \frac{\partial \mathbf{S}_0}{\partial \theta'} \mathbf{S}_0^{-1} \frac{\partial \mathbf{S}_0}{\partial \theta}\right) - \operatorname{trace}\left(\mathbf{S}_0^{-1} \frac{\partial^2 \mathbf{S}_0}{\partial \theta \partial \theta'}\right)\right)$$

Near the maximum, the first term is unimportant and the second term can be approximated to give the following second derivative approximation:

$$\frac{\partial^2 L}{\partial \theta \partial \theta'} \cong -n \operatorname{trace}\left(\mathbf{S}_0^{-1} \frac{\partial \mathbf{E}}{\partial \theta} \frac{\partial \mathbf{E}'}{\partial \theta'}\right)$$

The first derivative matrix and this second derivative matrix approximation are computed from the sample covariance matrix $\mathbf{C}_0$ and the truncated sequence $\mathbf{\Xi}_i$. The approximate likelihood function is maximized by a modified Newton-Raphson algorithm employing these derivative matrices.

The matrix $\mathbf{S}_0$ is used as the estimate of the innovation covariance matrix, $\mathbf{\Sigma_{ee}}$. The negative of the inverse of the second derivative matrix at the maximum is used as an approximate covariance matrix for the parameter estimates. The standard errors of the parameter estimates printed in the parameter estimates tables are taken from the diagonal of this covariance matrix. The parameter covariance matrix is printed when the COVB option is specified.

If the data are nearly nonstationary, a better estimate of $\mathbf{\Sigma_{ee}}$ and the other parameters can sometimes be obtained by specifying the RESIDEST option. The RESIDEST option estimates the parameters using conditional least squares instead of maximum likelihood.

The residuals are computed using the state space equation and the sample mean values of the variables in the model as start-up values. The estimate of $\mathbf{S}_0$ is then computed using the residuals from the *i*th observation on, where *i* is the maximum number of times any variable occurs in the state vector. A multivariate Gauss-Marquardt algorithm is used to minimize $|\mathbf{S}_0|$. Refer to Harvey (1981a) for a further description of this method.

Forecasting

Given estimates of $\mathbf{F}$, $\mathbf{G}$, and $\mathbf{\Sigma_{ee}}$, forecasts of $\mathbf{x}_t$ are computed from the conditional expectation of $\mathbf{z}_t$.

In forecasting, the parameters $\mathbf{F}$, $\mathbf{G}$, and $\mathbf{\Sigma_{ee}}$ are replaced with the estimates or by values specified in the RESTRICT statement. One-step-ahead forecasting is performed for the observation $\mathbf{x}_t$, where $t \leq n - b$. Here n is the number of observations and *b* is the value of the BACK= option. For the observation $\mathbf{x}_t$, where $t > n - b$, *m*-step-ahead forecasting is performed for $m = t - n + b$. The forecasts are generated recursively with the initial condition $\mathbf{z}_0 = 0$.

The *m*-step-ahead forecast of $\mathbf{z}_{t+m}$ is $\mathbf{z}_{t+m|t}$, where $\mathbf{z}_{t+m|t}$ denotes the conditional expectation of $\mathbf{z}_{t+m}$ given the information available at time *t*. The *m*-step-ahead forecast of $\mathbf{x}_{t+m}$ is $\mathbf{x}_{t+m|t} = \mathbf{H}\mathbf{z}_{t+m|t}$, where the matrix $\mathbf{H} = [\mathbf{I}_r \mathbf{0}]$.

Let $\Psi_i = \mathbf{F}^i\mathbf{G}$. Note that the last $s - r$ elements of $\mathbf{z}_t$ consist of the elements of $\mathbf{x}_{u|t}$ for $u > t$.

The state vector $\mathbf{z}_{t+m}$ can be represented as

$$\mathbf{z}_{t+m} = \mathbf{F}^m\mathbf{z}_t + \sum_{i=0}^{m-1} \mathbf{\Psi}_i\mathbf{e}_{t+m-i}$$

Since $\mathbf{e}_{t+i|t} = \mathbf{0}$ for $i > 0$, the *m*-step-ahead forecast $\mathbf{z}_{t+m|t}$ is

$$\mathbf{z}_{t+m|t} = \mathbf{F}^m\mathbf{z}_t = \mathbf{F}\mathbf{z}_{t+m-1|t}$$

Therefore, the *m*-step-ahead forecast of $\mathbf{x}_{t+m}$ is

$$\mathbf{x}_{t+m|t} = \mathbf{H}\mathbf{z}_{t+m|t}$$

The *m*-step-ahead forecast error is

$$\mathbf{z}_{t+m} - \mathbf{z}_{t+m|t} = \sum_{i=0}^{m-1} \mathbf{\Psi}_i\mathbf{e}_{t+m-i}$$

The variance of the *m*-step-ahead forecast error is

$$\mathbf{V}_{z,m} = \sum_{i=0}^{m-1} \mathbf{\Psi}_i\mathbf{\Sigma_{ee}}\Psi_i'$$

Letting $\mathbf{V}_{z,0} = \mathbf{0}$, the variance of the *m*-step-ahead forecast error of $\mathbf{z}_{t+m}$, $\mathbf{V}_{z,m}$, can be computed recursively as follows:

$$\mathbf{V}_{z,m} = \mathbf{V}_{z,m-1} + \mathbf{\Psi}_{m-1}\mathbf{\Sigma_{ee}}\mathbf{\Psi}'_{m-1}$$

The variance of the *m*-step-ahead forecast error of $\mathbf{x}_{t+m}$ is the $r \times r$ left upper submatrix of $\mathbf{V}_{z,m}$; that is,

$$\mathbf{V}_{x,m} = \mathbf{H}\mathbf{V}_{z,m}\mathbf{H}'$$

Unless the NOCENTER option is specified, the sample mean vector is added to the forecast. When differencing is specified, the forecasts $\mathbf{x}_{t+m|t}$ plus the sample mean vector are integrated back to produce forecasts for the original series.

Let $\mathbf{y}_t$ be the original series specified by the VAR statement, with some 0 values appended corresponding to the unobserved past observations. Let B be the backshift operator, and let $\mathbf{\Delta}(B)$ be the $s \times s$ matrix polynomial in the backshift operator corresponding to the differencing specified by the VAR statement. The off-diagonal elements of $\mathbf{\Delta}_i$ are 0. Note that $\mathbf{\Delta}_0 = \mathbf{I}_s$, where $\mathbf{I}_s$ is the $s \times s$ identity matrix. Then $\mathbf{z}_t = \mathbf{\Delta}(B)\mathbf{y}_t$.

This gives the relationship

$$\mathbf{y}_t = \mathbf{\Delta}^{-1}(B)\mathbf{z}_t = \sum_{i=0}^{\infty} \mathbf{\Lambda}_i \mathbf{z}_{t-i}$$

where $\mathbf{\Delta}^{-1}(B) = \sum_{i=0}^{\infty} \mathbf{\Lambda}_i B^i$ and $\mathbf{\Lambda}_0 = \mathbf{I}_s$.

The *m*-step-ahead forecast of $\mathbf{y}_{t+m}$ is

$$\mathbf{y}_{t+m|t} = \sum_{i=0}^{m-1} \mathbf{\Lambda}_i \mathbf{z}_{t+m-i|t} + \sum_{i=m}^{\infty} \mathbf{\Lambda}_i \mathbf{z}_{t+m-i}$$

The *m*-step-ahead forecast error of $\mathbf{y}_{t+m}$ is

$$\sum_{i=0}^{m-1} \mathbf{\Lambda}_i \left(\mathbf{z}_{t+m-i} - \mathbf{z}_{t+m-i|t} \right) = \sum_{i=0}^{m-1} \left(\sum_{u=0}^{i} \mathbf{\Lambda}_u \mathbf{\Psi}_{i-u} \right) \mathbf{e}_{t+m-i}$$

Letting $\mathbf{V}_{y,0} = \mathbf{0}$, the variance of the *m*-step-ahead forecast error of $\mathbf{y}_{t+m}$, $\mathbf{V}_{y,m}$, is

$$\begin{aligned} \mathbf{V}_{y,m} &= \sum_{i=0}^{m-1} \left(\sum_{u=0}^{i} \mathbf{\Lambda}_u \mathbf{\Psi}_{i-u} \right) \mathbf{\Sigma}_{\mathbf{ee}} \left(\sum_{u=0}^{i} \mathbf{\Lambda}_u \mathbf{\Psi}_{i-u} \right)' \\ &= \mathbf{V}_{y,m-1} + \left(\sum_{u=0}^{m-1} \mathbf{\Lambda}_u \mathbf{\Psi}_{m-1-u} \right) \mathbf{\Sigma}_{\mathbf{ee}} \left(\sum_{u=0}^{m-1} \mathbf{\Lambda}_u \mathbf{\Psi}_{m-1-u} \right)' \end{aligned}$$

Relation of ARMA and State Space Forms

Every state space model has an ARMA representation, and conversely every ARMA model has a state space representation. This section discusses this equivalence. The following material is adapted from Akaike (1974), where there is a more complete discussion. Pham-Dinh-Tuan (1978) also contains a discussion of this material.

Suppose you are given the following ARMA model:

$$\mathbf{\Phi}(B)\mathbf{x}_t = \mathbf{\Theta}(B)\mathbf{e}_t$$

or, in more detail

$$\mathbf{x}_t - \mathbf{\Phi}_1\mathbf{x}_{t-1} - \cdots - \mathbf{\Phi}_p\mathbf{x}_{t-p} = \mathbf{e}_t + \mathbf{\Theta}_1\mathbf{e}_{t-1} + \cdots + \mathbf{\Theta}_q\mathbf{e}_{t-q} \qquad (1)$$

where $\mathbf{e}_t$ is a sequence of independent multivariate normal random vectors with mean $\mathbf{0}$ and variance matrix $\mathbf{\Sigma_{ee}}$; B is the backshift operator ($B\mathbf{x}_t = \mathbf{x}_{t-1}$); $\mathbf{\Phi}(B)$ and $\mathbf{\Theta}(B)$ are matrix polynomials in B; and $\mathbf{x}_t$ is the observed process.

If the roots of the determinantial equation $|\mathbf{\Phi}(B)| = 0$ are outside the unit circle in the complex plane, the model can also be written as

$$\mathbf{x}_t = \mathbf{\Phi}^{-1}(B)\mathbf{\Theta}(B)\mathbf{e}_t = \sum_{i=0}^{\infty}\mathbf{\Psi}_i\mathbf{e}_{t-i}$$

The $\mathbf{\Psi}_i$ matrices are known as the impulse response matrices and can be computed as $\mathbf{\Phi}^{-1}(B)\mathbf{\Theta}(B)$.

You can assume $p > q$ since, if this is not initially true, you can add more terms $\mathbf{\Phi}_i$ that are identically 0 without changing the model.

To write this set of equations in a state space form, proceed as follows. Let $\mathbf{x}_{t+i|t}$ be the conditional expectation of $\mathbf{x}_{t+i}$ given $\mathbf{x}_w$ for $w \leq t$. The following relations hold:

$$\mathbf{x}_{t+i|t} = \sum_{j=i}^{\infty}\mathbf{\Psi}_j\mathbf{e}_{t+i-j}$$

$$\mathbf{x}_{t+i|t+1} = \mathbf{x}_{t+i|t} + \mathbf{\Psi}_{i-1}\mathbf{e}_{t+1}$$

However, from equation (1) you can derive the following relationship:

$$\mathbf{x}_{t+p|t} = \mathbf{\Phi}_1\mathbf{x}_{t+p-1|t} + \cdots + \mathbf{\Phi}_p\mathbf{x}_t \qquad (2)$$

Hence, when $i = p$, you can substitute for $\mathbf{x}_{t+p|t}$ in the right-hand side of equation (2) and close the system of equations.

This substitution results in the following model in the state space form $\mathbf{z}_{t+1} = \mathbf{F}\mathbf{z}_t + \mathbf{G}\mathbf{e}_{t+1}$:

$$\begin{bmatrix} \mathbf{x}_{t+1} \\ \mathbf{x}_{t+2|t+1} \\ \vdots \\ \mathbf{x}_{t+p|t+1} \end{bmatrix} = \begin{bmatrix} 0 & \mathbf{I} & 0 & \cdots & 0 \\ 0 & 0 & \mathbf{I} & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ \mathbf{\Phi}_p & \mathbf{\Phi}_{p-1} & & \cdots & \mathbf{\Phi}_1 \end{bmatrix} \begin{bmatrix} \mathbf{x}_t \\ \mathbf{x}_{t+1|t} \\ \vdots \\ \mathbf{x}_{t+p-1|t} \end{bmatrix} + \begin{bmatrix} \mathbf{I} \\ \mathbf{\Psi}_1 \\ \vdots \\ \mathbf{\Psi}_{p-1} \end{bmatrix} \mathbf{e}_{t+1}$$

Note that the state vector $\mathbf{z}_t$ is composed of conditional expectations of $\mathbf{x}_t$ and the first *r* components of $\mathbf{z}_t$ are equal to $\mathbf{x}_t$.

The state space form can be cast into an ARMA form by solving the system of difference equations for the first *r* components.

When converting from an ARMA form to a state space form, you can generate a state vector larger than needed; that is, the state space model may not be a minimal representation. When going from a state space form to an ARMA form, you can have nontrivial common factors in the autoregressive and moving average operators that yield an ARMA model larger than necessary.

If the state space form used is not a minimal representation, some but not all components of $\mathbf{x}_{t+i|t}$ may be linearly dependent. This situation corresponds to $[\mathbf{\Phi}_p \mathbf{\Theta}_{p-1}]$ being of less than full rank when $\mathbf{\Phi}(B)$ and $\mathbf{\Theta}(B)$ have no common nontrivial left factors. In this case, $\mathbf{z}_t$ consists of a subset of the possible components of $[\mathbf{x}_{t+i|t}] \quad i = 1, 2, \cdots, p-1$. However, once a component of $\mathbf{x}_{t+i|t}$ (for example, the *j*th one) is linearly dependent on the previous conditional expectations, then all subsequent *j*th components of $\mathbf{x}_{t+k|t}$ for $k > i$ must also be linearly dependent. Note that in this case, equivalent but seemingly different structures can arise if the order of the components within $\mathbf{x}_t$ is changed.

OUT= Data Set

The forecasts are contained in the output data set specified by the OUT= option on the PROC STATESPACE statement. The OUT= data set contains the following variables:

- the BY variables
- the ID variable
- the VAR statement variables. These variables contain the actual values from the input data set.
- FOR*i*, numeric variables containing the forecasts. The variable FOR*i* contains the forecasts for the *i*th variable in the VAR statement list. Forecasts are one-step-ahead predictions until the end of the data or until the observation specified by the BACK= option.

- RES*i*, numeric variables containing the residual for the forecast of the *i*th variable in the VAR statement list. For forecast observations, the actual values are missing and the RES*i* variables contain missing values.
- STD*i*, numeric variables containing the standard deviation for the forecast of the *i*th variable in the VAR statement list. The values of the STD*i* variables can be used to construct univariate confidence limits for the corresponding forecasts. However, such confidence limits do not take into account the covariance of the forecasts.

OUTAR= Data Set

The OUTAR= data set contains the estimates of the preliminary autoregressive models. The OUTAR= data set contains the following variables:

- ORDER, a numeric variable containing the order *p* of the autoregressive model that the observation represents
- AIC, a numeric variable containing the value of the information criterion AIC_p
- SIGF*l*, numeric variables containing the estimate of the innovation covariance matrices for the forward autoregressive models. The variable SIGF*l* contains the *l*th column of $\widehat{\mathbf{\Sigma}}_p$ in the observations with ORDER=*p*.
- SIGB*l*, numeric variables containing the estimate of the innovation covariance matrices for the backward autoregressive models. The variable SIGB*l* contains the *l*th column of $\widehat{\mathbf{\Omega}}_p$ in the observations with ORDER=*p*.
- FOR*k_l*, numeric variables containing the estimates of the autoregressive parameter matrices for the forward models. The variable FOR*k_l* contains the *l*th column of the lag *k* autoregressive parameter matrix $\widehat{\mathbf{\Phi}}_k^p$ in the observations with ORDER=*p*.
- BAC*k_l*, numeric variables containing the estimates of the autoregressive parameter matrices for the backward models. The variable BAC*k_l* contains the *l*th column of the lag *k* autoregressive parameter matrix $\widehat{\mathbf{\Psi}}_k^p$ in the observations with ORDER=*p*.

The estimates for the order *p* autoregressive model can be selected as those observations with ORDER=*p*. Within these observations, the *k,l*th element of $\mathbf{\Phi}_i^p$ is given by the value of the FOR*i_l* variable in the *k*th observation. The *k,l*th element of $\mathbf{\Psi}_i^p$ is given by the value of BAC*i_l* variable in the *k*th observation. The *k,l*th element of $\mathbf{\Sigma}_p$ is given by SIGF*l* in the *k*th observation. The *k,l*th element of $\mathbf{\Omega}_p$ is given by SIGB*l* in the *k*th observation.

Table 18.1 shows an example of the OUTAR= data set, with ARMAX=3 and $\mathbf{x}_t$ of dimension 2. In Table 18.1, (i, j) indicate the *i,j*th element of the matrix.

Table 18.1. Values in the OUTAR= Data Set

Obs	ORDER	AIC	SIGF1	SIGF2	SIGB1	SIGB2	FOR1_1	FOR1_2	FOR2_1	FOR2_2	FOR3_1
1	0	AIC_0	$\mathbf{\Sigma}_{0(1,1)}$	$\mathbf{\Sigma}_{0(1,2)}$	$\mathbf{\Omega}_{0(1,1)}$	$\mathbf{\Omega}_{0(1,2)}$	.	.	.	.	.
2	0	AIC_0	$\mathbf{\Sigma}_{0(2,1)}$	$\mathbf{\Sigma}_{0(2,2)}$	$\mathbf{\Omega}_{0(2,1)}$	$\mathbf{\Omega}_{0(2,2)}$	.	.	.	.	.
3	1	AIC_1	$\mathbf{\Sigma}_{1(1,1)}$	$\mathbf{\Sigma}_{1(1,2)}$	$\mathbf{\Omega}_{1(1,1)}$	$\mathbf{\Omega}_{1(1,2)}$	$\mathbf{\Phi}^1_{1(1,1)}$	$\mathbf{\Phi}^1_{1(1,2)}$	.	.	.
4	1	AIC_1	$\mathbf{\Sigma}_{1(2,1)}$	$\mathbf{\Sigma}_{1(1,2)}$	$\mathbf{\Omega}_{1(2,1)}$	$\mathbf{\Omega}_{1(1,2)}$	$\mathbf{\Phi}^1_{1(2,1)}$	$\mathbf{\Phi}^1_{1(2,2)}$	.	.	.
5	2	AIC_2	$\mathbf{\Sigma}_{2(1,1)}$	$\mathbf{\Sigma}_{2(1,2)}$	$\mathbf{\Omega}_{2(1,1)}$	$\mathbf{\Omega}_{2(1,2)}$	$\mathbf{\Phi}^2_{1(1,1)}$	$\mathbf{\Phi}^2_{1(1,2)}$	$\mathbf{\Phi}^2_{2(1,1)}$	$\mathbf{\Phi}^2_{2(1,2)}$	.
6	2	AIC_2	$\mathbf{\Sigma}_{2(2,1)}$	$\mathbf{\Sigma}_{2(1,2)}$	$\mathbf{\Omega}_{2(2,1)}$	$\mathbf{\Omega}_{2(1,2)}$	$\mathbf{\Phi}^2_{1(2,1)}$	$\mathbf{\Phi}^2_{1(2,2)}$	$\mathbf{\Phi}^2_{2(2,1)}$	$\mathbf{\Phi}^2_{2(2,2)}$	.
7	3	AIC_3	$\mathbf{\Sigma}_{3(1,1)}$	$\mathbf{\Sigma}_{3(1,2)}$	$\mathbf{\Omega}_{3(1,1)}$	$\mathbf{\Omega}_{3(1,2)}$	$\mathbf{\Phi}^3_{1(1,1)}$	$\mathbf{\Phi}^3_{1(1,2)}$	$\mathbf{\Phi}^3_{2(1,1)}$	$\mathbf{\Phi}^3_{2(1,2)}$	$\mathbf{\Phi}^3_{3(1,1)}$
8	3	AIC_3	$\mathbf{\Sigma}_{3(2,1)}$	$\mathbf{\Sigma}_{3(1,2)}$	$\mathbf{\Omega}_{3(2,1)}$	$\mathbf{\Omega}_{3(1,2)}$	$\mathbf{\Phi}^3_{1(2,1)}$	$\mathbf{\Phi}^3_{1(2,2)}$	$\mathbf{\Phi}^3_{2(2,1)}$	$\mathbf{\Phi}^3_{2(2,2)}$	$\mathbf{\Phi}^3_{3(2,1)}$

Obs	FOR3_2	BACK1_1	BACK1_2	BACK2_1	BACK2_2	BACK3_1	BACK3_2
1	.	.	.	.	.	.	.
2	.	.	.	.	.	.	.
3	.	$\Psi^1_{1(1,1)}$	$\Psi^1_{1(1,2)}$	.	.	.	.
4	.	$\Psi^1_{1(2,1)}$	$\Psi^1_{1(2,2)}$	.	.	.	.
5	.	$\Psi^2_{1(1,1)}$	$\Psi^2_{1(1,2)}$	$\Psi^2_{2(1,1)}$	$\Psi^2_{2(1,2)}$	.	.
6	.	$\Psi^2_{1(2,1)}$	$\Psi^2_{1(2,2)}$	$\Psi^2_{2(2,1)}$	$\Psi^2_{2(2,2)}$	.	.
7	$\Phi^3_{3(1,2)}$	$\Psi^3_{1(1,1)}$	$\Psi^3_{1(1,2)}$	$\Psi^3_{2(1,1)}$	$\Psi^3_{2(1,2)}$	$\Psi^3_{3(1,1)}$	$\Psi^3_{3(1,2)}$
8	$\Phi^3_{3(2,2)}$	$\Psi^3_{1(2,1)}$	$\Psi^3_{1(2,2)}$	$\Psi^3_{2(2,1)}$	$\Psi^3_{2(2,2)}$	$\Psi^3_{3(2,1)}$	$\Psi^3_{3(2,2)}$

The estimated autoregressive parameters can be used in the IML procedure to obtain autoregressive estimates of the spectral density function or forecasts based on the autoregressive models.

OUTMODEL= Data Set

The OUTMODEL= data set contains the estimates of the $\mathbf{F}$ and $\mathbf{G}$ matrices and their standard errors, the names of the components of the state vector, and the estimates of the innovation covariance matrix. The variables contained in the OUTMODEL= data set are as follows:

- the BY variables
- STATEVEC, a character variable containing the name of the component of the state vector corresponding to the observation. The STATEVEC variable has the value STD for standard deviations observations, which contain the standard errors for the estimates given in the preceding observation.
- F_*j*, numeric variables containing the columns of the $\mathbf{F}$ matrix. The variable F_*j* contains the *j*th column of $\mathbf{F}$. The number of F_*j* variables is equal to the value of the DIMMAX= option. If the model is of smaller dimension, the extraneous variables are set to missing.
- G_*j*, numeric variables containing the columns of the $\mathbf{G}$ matrix. The variable G_*j* contains the *j*th column of $\mathbf{G}$. The number of G_*j* variables is equal to *r*, the dimension of $\mathbf{x}_t$ given by the number of variables in the VAR statement.
- SIG_*j*, numeric variables containing the columns of the innovation covariance matrix. The variable SIG_*j* contains the *j*th column of $\mathbf{\Sigma}_{ee}$. There are *r* variables SIG_*j*.

Table 18.2 shows an example of the OUTMODEL= data set, with $\mathbf{x}_t = (x_t, y_t)'$, $\mathbf{z}_t = (x_t, y_t, x_{t+1|t})'$, and DIMMAX=4. In Table 18.2, $\mathbf{F}_{i,j}$ and $\mathbf{G}_{i,j}$ are the *i,j*th elements of $\mathbf{F}$ and $\mathbf{G}$ respectively. Note that all elements for F_4 are missing because $\mathbf{F}$ is a 3×3 matrix.

Table 18.2. Value in the OUTMODEL= Data Set

Obs	STATEVEC	F_1	F_2	F_3	F_4	G_1	G_2	SIG_1	SIG_2
1	X(T;T)	0	0	1	.	1	0	$\mathbf{\Sigma}_{1,1}$	$\mathbf{\Sigma}_{1,2}$
2	STD	.	.	.	.	.	.	.	.
3	Y(T;T)	$\mathbf{F}_{2,1}$	$\mathbf{F}_{2,2}$	$\mathbf{F}_{2,3}$	.	0	1	$\mathbf{\Sigma}_{2,1}$	$\mathbf{\Sigma}_{2,2}$
4	STD	std $\mathbf{F}_{2,1}$	std $\mathbf{F}_{2,2}$	std $\mathbf{F}_{2,3}$	.	.	.	.	.
5	X(T+1;T)	$\mathbf{F}_{3,1}$	$\mathbf{F}_{3,2}$	$\mathbf{F}_{3,3}$	.	$\mathbf{G}_{3,1}$	$\mathbf{G}_{3,2}$	.	.
6	STD	std $\mathbf{F}_{3,1}$	std $\mathbf{F}_{3,2}$	std $\mathbf{F}_{3,3}$	.	std $\mathbf{G}_{3,1}$	std $\mathbf{G}_{3,2}$	.	.

Printed Output

The printed output produced by the STATESPACE procedure is described in the following:

1. descriptive statistics, which include the number of observations used, the names of the variables, their means and standard deviations (Std), and the differencing operations used.
2. the Akaike information criteria for the sequence of preliminary autoregressive models
3. if the PRINTOUT=LONG option is specified, the sample autocovariance matrices of the input series at various lags.
4. if the PRINTOUT=LONG option is specified, the sample autocorrelation matrices of the input series.
5. a schematic representation of the autocorrelation matrices, showing the significant autocorrelations.
6. if the PRINTOUT=LONG option is specified, the partial autoregressive matrices. (These are $\mathbf{\Phi}_p^p$ as described in "Preliminary Autoregressive Models" earlier in this chapter.)
7. a schematic representation of the partial autocorrelation matrices, showing the significant partial autocorrelations.
8. the Yule-Walker estimates of the autoregressive parameters for the autoregressive model with the minimum AIC.
9. if the PRINTOUT=LONG option is specified, the autocovariance matrices of the residuals of the minimum AIC model. This is the sequence of estimated innovation variance matrices for the solutions of the Yule-Walker equations.
10. if the PRINTOUT=LONG option is specified, the autocorrelation matrices of the residuals of the minimum AIC model.
11. If the CANCORR option is specified, the canonical correlations analysis for each potential state vector considered in the state vector selection process. This includes the potential state vector, the canonical correlations, the information criterion for the smallest canonical correlation, Bartlett's χ^2 statistic ("Chi Square") for the smallest canonical correlation, and the degrees of freedom of Bartlett's χ^2.

12. the components of the chosen state vector.

13. the preliminary estimate of the transition matrix, $\mathbf{F}$, the input matrix, $\mathbf{G}$, and the variance matrix for the innovations, $\mathbf{\Sigma_{ee}}$.

14. if the ITPRINT option is specified, the iteration history of the likelihood maximization. For each iteration, this shows the iteration number, the number of step halvings, the determinant of the innovation variance matrix, the damping factor Lambda, and the values of the parameters.

15. the state vector, printed again to aid interpretation of the following listing of $\mathbf{F}$ and $\mathbf{G}$.

16. the final estimate of the transition matrix, $\mathbf{F}$.

17. the final estimate of the input matrix, $\mathbf{G}$.

18. the final estimate of the variance matrix for the innovations, $\mathbf{\Sigma_{ee}}$.

19. a table listing the estimates of the free parameters in $\mathbf{F}$ and $\mathbf{G}$ and their standard errors and *t* statistics.

20. if the COVB option is specified, the covariance matrix of the parameter estimates.

21. if the COVB option is specified, the correlation matrix of the parameter estimates.

22. if the PRINT option is specified, the forecasts and their standard errors.

ODS Table Names

PROC STATESPACE assigns a name to each table it creates. You can use these names to reference the table when using the Output Delivery System (ODS) to select tables and create output data sets. These names are listed in the following table. For more information on ODS, see Chapter 6, "Using the Output Delivery System."

Table 18.3. ODS Tables Produced in PROC STATESPACE

ODS Table Name	Description	Option
NObs	Number of observations	default
Summary	Simple summary statistics table	default
InfoCriterion	Information criterion table	default
CovLags	Covariance Matrices of Input Series	PRINTOUT=LONG
CorrLags	Correlation Matrices of Input Series	PRINTOUT=LONG
PartialAR	Partial Autoregressive Matrices	PRINTOUT=LONG
YWEstimates	Yule-Walker Estimates for Minimum AIC	default
CovResiduals	Covariance of Residuals	PRINTOUT=LONG
CorrResiduals	Residual Correlations from AR Models	PRINTOUT=LONG
StateVector	State vector table	default
CorrGraph	Schematic Representation of Correlations	default
TransitionMatrix	Transition Matrix	default
InputMatrix	Input Matrix	default
VarInnov	Variance Matrix for the Innovation	default
CovB	Covariance of Parameter Estimates	COVB
CorrB	Correlation of Parameter Estimates	COVB
CanCorr	Canonical Correlation Analysis	CANCORR
IterHistory	Iterative Fitting table	ITPRINT
ParameterEstimates	Parameter Estimates Table	default
Forecasts	Forecasts Table	PRINT
ConvergenceStatus	Convergence Status Table	default

Example

Example 18.1. Series J from Box and Jenkins

This example analyzes the gas furnace data (series J) from Box and Jenkins. (The data are not shown. Refer to Box and Jenkins (1976) for the data.)

First, a model is selected and fit automatically using the following statements.

```
title1 'Gas Furnace Data';
title2 'Box & Jenkins Series J';
title3 'Automatically Selected Model';

proc statespace data=seriesj cancorr;
   var x y;
run;
```

The results for the automatically selected model are shown in Output 18.1.1.

Output 18.1.1. Results for Automatically Selected Model

```
                          Gas Furnace Data
                       Box & Jenkins Series J
                     Automatically Selected Model

                        The STATESPACE Procedure

                     Number of Observations    296


                                          Standard
                     Variable        Mean       Error

                     x           -0.05683    1.072766
                     y           53.50912    3.202121



                          Gas Furnace Data
                       Box & Jenkins Series J
                     Automatically Selected Model

                        The STATESPACE Procedure

              Information Criterion for Autoregressive Models

   Lag=0    Lag=1    Lag=2    Lag=3    Lag=4    Lag=5    Lag=6    Lag=7    Lag=8

651.3862 -1033.57 -1632.96 -1645.12 -1651.52 -1648.91 -1649.34 -1643.15 -1638.56

                                  Information
                                  Criterion for
                                  Autoregressive
                                      Models

                                  Lag=9      Lag=10

                                 -1634.8    -1633.59


                  Schematic Representation of Correlations

   Name/Lag     0     1     2     3     4     5     6     7     8     9    10

   x           +-    +-    +-    +-    +-    +-    +-    +-    +-    +-    +-
   y           -+    -+    -+    -+    -+    -+    -+    -+    -+    -+    -+

          + is > 2*std error,  - is < -2*std error,  . is between
```

Gas Furnace Data
Box & Jenkins Series J
Automatically Selected Model

The STATESPACE Procedure

Schematic Representation of Partial Autocorrelations

Name/Lag	1	2	3	4	5	6	7	8	9	10
x	+.	-.	+.	..	..	-.	..	..	..	..
y	-+	--	-.	.+	..	..	..	..	..	.+

+ is > 2*std error, - is < -2*std error, . is between

Yule-Walker Estimates for Minimum AIC

	Lag=1		Lag=2		Lag=3		Lag=4	
	x	y	x	y	x	y	x	y
x	1.925887	-0.00124	-1.20166	0.004224	0.116918	-0.00867	0.104236	0.003268
y	0.050496	1.299793	-0.02046	-0.3277	-0.71182	-0.25701	0.195411	0.133417

Gas Furnace Data
Box & Jenkins Series J
Automatically Selected Model

The STATESPACE Procedure
Canonical Correlations Analysis

x(T;T)	y(T;T)	x(T+1;T)	Information Criterion	Chi Square	DF
1	1	0.804883	292.9228	304.7481	8

x(T;T)	y(T;T)	x(T+1;T)	y(T+1;T)	Information Criterion	Chi Square	DF
1	1	0.906681	0.607529	122.3358	134.7237	7

x(T;T)	y(T;T)	x(T+1;T)	y(T+1;T)	x(T+2;T)	Information Criterion	Chi Square	DF
1	1	0.909434	0.610278	0.186274	-1.54701	10.34705	6

x(T;T)	y(T;T)	x(T+1;T)	y(T+1;T)	y(T+2;T)	Information Criterion	Chi Square	DF
1	1	0.91014	0.618937	0.206823	0.940392	12.80924	6

x(T;T)	y(T;T)	x(T+1;T)	y(T+1;T)	y(T+2;T)	y(T+3;T)	Information Criterion	Chi Square	DF
1	1	0.912963	0.628785	0.226598	0.083258	-7.94103	2.041584	5

Gas Furnace Data
Box & Jenkins Series J
Automatically Selected Model

The STATESPACE Procedure
Selected Statespace Form and Preliminary Estimates

State Vector

x(T;T)	y(T;T)	x(T+1;T)	y(T+1;T)	y(T+2;T)

Estimate of Transition Matrix

0	0	1	0	0
0	0	0	1	0
-0.84718	0.026794	1.711715	-0.05019	0
0	0	0	0	1
-0.19785	0.334274	-0.18174	-1.23557	1.787475

Input Matrix for Innovation

1	0
0	1
1.925887	-0.00124
0.050496	1.299793
0.142421	1.361696

Gas Furnace Data
Box & Jenkins Series J
Automatically Selected Model

The STATESPACE Procedure
Selected Statespace Form and Preliminary Estimates

Variance Matrix for Innovation

0.035274	-0.00734
-0.00734	0.097569

Gas Furnace Data
Box & Jenkins Series J
Automatically Selected Model

The STATESPACE Procedure
Selected Statespace Form and Fitted Model

State Vector

x(T;T)	y(T;T)	x(T+1;T)	y(T+1;T)	y(T+2;T)

Estimate of Transition Matrix

0	0	1	0	0
0	0	0	1	0
-0.86192	0.030609	1.724235	-0.05483	0
0	0	0	0	1
-0.34839	0.292124	-0.09435	-1.09823	1.671418

Input Matrix for Innovation

1	0
0	1
1.92442	-0.00416
0.015621	1.258495
0.08058	1.353204

Gas Furnace Data
Box & Jenkins Series J
Automatically Selected Model

The STATESPACE Procedure
Selected Statespace Form and Fitted Model

Variance Matrix for Innovation

0.035579	-0.00728
-0.00728	0.095577

Parameter Estimates

Parameter	Estimate	Standard Error	t Value
F(3,1)	-0.86192	0.072961	-11.81
F(3,2)	0.030609	0.026167	1.17
F(3,3)	1.724235	0.061599	27.99
F(3,4)	-0.05483	0.030169	-1.82
F(5,1)	-0.34839	0.135253	-2.58
F(5,2)	0.292124	0.046299	6.31
F(5,3)	-0.09435	0.096527	-0.98
F(5,4)	-1.09823	0.109525	-10.03
F(5,5)	1.671418	0.083737	19.96
G(3,1)	1.924420	0.058162	33.09
G(3,2)	-0.00416	0.035255	-0.12
G(4,1)	0.015621	0.095771	0.16
G(4,2)	1.258495	0.055742	22.58
G(5,1)	0.080580	0.151622	0.53
G(5,2)	1.353204	0.091388	14.81

The two series are believed to have a transfer function relation with the gas rate (variable X) as the input and the CO_2 concentration (variable Y) as the output. Since the parameter estimates shown in Output 18.1.1 support this kind of model, the model is reestimated with the feedback parameters restricted to 0. The following statements fit the transfer function (no feedback) model.

```
title3 'Transfer Function Model';
proc statespace data=seriesj printout=none;
   var x y;
   restrict f(3,2)=0 f(3,4)=0
            g(3,2)=0 g(4,1)=0 g(5,1)=0;
run;
```

The last 2 pages of the output are shown in Output 18.1.2.

Output 18.1.2. STATESPACE Output for Transfer Function Model

```
                          Gas Furnace Data
                        Box & Jenkins Series J
                       Transfer Function Model

                      The STATESPACE Procedure
                Selected Statespace Form and Fitted Model

                              State Vector

x(T;T)         y(T;T)         x(T+1;T)       y(T+1;T)       y(T+2;T)


                       Estimate of Transition Matrix

          0              0              1              0              0
          0              0              0              1              0
    -0.68882              0       1.598717              0              0
          0              0              0              0              1
    -0.35944       0.284179        -0.0963       -1.07313       1.650047


                       Input Matrix for Innovation

                                1              0
                                0              1
                         1.923446              0
                                0       1.260856
                                0       1.346332
```

Gas Furnace Data
Box & Jenkins Series J
Transfer Function Model

The STATESPACE Procedure
Selected Statespace Form and Fitted Model

Variance Matrix for Innovation

0.036995	-0.0072
-0.0072	0.095712

Parameter Estimates

Parameter	Estimate	Standard Error	t Value
F(3,1)	-0.68882	0.050549	-13.63
F(3,3)	1.598717	0.050924	31.39
F(5,1)	-0.35944	0.229044	-1.57
F(5,2)	0.284179	0.096944	2.93
F(5,3)	-0.09630	0.140876	-0.68
F(5,4)	-1.07313	0.250385	-4.29
F(5,5)	1.650047	0.188533	8.75
G(3,1)	1.923446	0.056328	34.15
G(4,2)	1.260856	0.056464	22.33
G(5,2)	1.346332	0.091086	14.78

References

Akaike, H. (1974), "Markovian Representation of Stochastic Processes and Its Application to the Analysis of Autoregressive Moving Average Processes," *Annals of the Institute of Statistical Mathematics*, 26, 363-387.

Akaike, H. (1976), "Canonical Correlations Analysis of Time Series and the Use of an Information Criterion," in *Advances and Case Studies in System Identification*, eds. R. Mehra and D.G. Lainiotis, New York: Academic Press.

Anderson, T.W. (1971), *The Statistical Analysis of Time Series*, New York: John Wiley & Sons.

Ansley, C.F. and Newbold, P. (1979), "Multivariate Partial Autocorrelations," *Proceedings of the Business and Economic Statistics Section*, American Statistical Association, 349-353.

Box, G.E.P. and Jenkins, G. (1976), *Time Series Analysis: Forecasting and Control*, San Francisco: Holden-Day.

Brockwell, P.J. and Davis, R.A. (1991), *Time Series: Theory and Methods*, 2nd Edition, Springer-Verlag.

Hannan, E.J. (1970), *Multiple Time Series*, New York: John Wiley & Sons.

Hannan, E.J. (1976), "The Identification and Parameterization of ARMAX and State Space Forms," *Econometrica*, 44, 713-722.

Harvey, A.C. (1981a), *The Econometric Analysis of Time Series*, New York: John Wiley & Sons.

Harvey, A.C. (1981b), *Time Series Models*, New York: John Wiley & Sons.

Jones, R.H. (1974), "Identification and Autoregressive Spectrum Estimation," *IEEE Transactions on Automatic Control*, AC-19, 894-897.

Pham-Dinh-Tuan (1978), "On the Fitting of Multivariate Processes of the Autoregressive-Moving Average Type," *Biometrika*, 65, 99-107.

Priestley, M.B. (1980), "System Identification, Kalman Filtering, and Stochastic Control," in *Directions in Time Series*, eds. D.R. Brillinger and G.C. Tiao, Institute of Mathematical Statistics.

Whittle, P. (1963), "On the Fitting of Multivariate Autoregressions and the Approximate Canonical Factorization of a Spectral Density Matrix," *Biometrika*, 50, 129-134.

The correct bibliographic citation for this manual is as follows: SAS Institute Inc., *SAS/ETS User's Guide, Version 8*, Cary, NC: SAS Institute Inc., 1999. 1546 pp.

SAS/ETS User's Guide, Version 8

ISBN 1–58025–489–6

SAS Institute Inc., SAS Campus Drive, Cary, North Carolina 27513.

1st printing, October 1999