Box & Jenkins (1976) have devised a three stage process for the identification, parameter estimation, and verification of fit of a time series model. Using their notation, a stationary stochastic process $z_t$ is to be described by the appropriate autoregressive-moving average model of finite order $(p,q)$:

$$z_t = \phi_1 z_{t-1} + \ldots + \phi_p z_{t-p} + \epsilon_t - \theta_1 \epsilon_{t-1} - $$

$$\cdots - \theta_q \epsilon_{t-q}$$

where $\epsilon_t = z_t - \mu$, and the $\theta_j$'s (or white noise) are assumed to behave as normally and independently distributed random variables with mean zero and variance $\sigma^2$. Given $N$ observations of the infinite series, the appropriate model is identified; estimates of the model parameters are computed in the estimation stage, and diagnostic tests in the verification stage check adequacy of model fit. In recent years Box & Jenkins modelling techniques have been improved and refined to yield more exact parameter estimates, more powerful tests of fit. This paper acquaints the SAS user with improved techniques; the program, written in PROC MATRIX, offers alternatives to the estimation procedure and diagnostic test used by PROC ARIMA (which are based on Box & Jenkins modelling techniques).

Initial estimates for the time series parameters (the $\theta$'s & $\phi$'s) are input into the program, along with a SAS data set containing the appropriately differenced series. In order to ascertain that the estimates satisfy the requirements for stationarity and/or invertibility, an algorithm which approximates the ratio test described by Box & Jenkins (p. 50) provides a rough check on the parameters. In the event the test fails, the program is terminated.

Given the likelihood function for the $N$ observations of the series:

$$p(\theta, \phi, \epsilon) = (2\pi)^{n/2} |\Sigma|^{1/2} \exp(-0.5 \epsilon^T \Sigma^{-1} \epsilon),$$

the $(p+q+1)$ maximum likelihood estimates of the parameters are sought. PROC ARIMA uses a nonlinear conditional least squares technique outlined in Box & Jenkins (1976). The condition imposed in computing the estimates is that the log of the likelihood function is evaluated ignoring the contribution of the determinant of the $(m \times m)$ covariance matrix, $R$. The covariance matrix, $R$, is dependent upon the current values of the $\theta$'s and $\phi$'s. Box and Jenkins suggest ignoring the determinant $|R|$ in the log likelihood expression since for moderate to large $N$, the term is mainly dominated by the exponential term. The estimation problem is thus reduced to a least squares problem, and Box & Jenkins state that the resulting least squares estimates should provide close approximations to the conditional maximum likelihood estimates of the parameters. Recent studies have shown, however, that failing to include the determinant of the covariance matrix in the evaluation of the log likelihood function may result in inferior estimators. In light of these findings, Ansley (1979), expanding on previous work by Pradke & Kedem (1976), has derived an efficient method for computing the exact likelihood, and thus exact unconditional, rather than conditional, maximum likelihood estimates. These residuals can be used in the verification stage to check adequacy of fit of the model.

The algorithm is exact in that it does not ignore the contribution of the determinant $|R|$ in the log likelihood function. It is efficient in that the entire covariance matrix is never computed, nor is space required to store the $m \times m$ matrix. The algorithm involves a transformation of the original input series $\{\epsilon_t\}$ in such a way that the covariance matrix of the transformed series is a band matrix with a maximum bandwidth of $m$, where $m = \max(p,q)$. Using a Cholesky decomposition algorithm, the covariance matrix $\Sigma$ is decomposed so that $\Sigma = LL^T$, where $L$ is a lower triangular band matrix. Next, Ansley minimizes a set of residuals, $\{e_t\}$, which are assumed to be approximately normally and independently distributed, as $\epsilon_t = \sigma \epsilon_t$. Using this expression and the fact that the determinant $|R| = |L|^T \sigma$, the likelihood function is then simplified to:

$$p(\theta, \phi, \epsilon) = (2\pi)^{n/2} |L|^T \exp(-\Sigma e^T \sigma^{-1} e).$$

By taking the log of the likelihood, finding the maximum likelihood estimates of $\epsilon$, $\sigma$, $\Sigma$, becomes a least squares problem that of minimizing the function value $f = \frac{1}{2} \epsilon^T \sigma^{-1} \epsilon$ where $\epsilon = |L|^T e$. So to maximize the log of the likelihood function, one need only compute values for the Cholesky factor, $L$ and $|L|^T$, and the residuals, $\epsilon$. Computation of $L$ and hence $|L|^T$ is simplified because each row of $L$ depends directly on only one row of the previous rows of $L$. Therefore, the elements of $L$ and $\sigma$ can be determined recursively and at the same time. Since only $m$ previous rows of $L$ and $\sigma$ are needed to compute the current row of $L$, storage space is required for an $(m+1) \times (m+1)$ matrix. In essence, if the band covariance matrix $\Sigma$ were to be set up, the $(m+1) \times (m+1)$ matrix would begin in the left hand corner and move along the
moving average parameters. Respectively.

appropriate Chi-square table value suggest

portmanteau statistic). The Box-Pierce

greater than the theoretical mean of zero

values of

model should behave approximately as in-

where \( l \).

portmanteau statistic is approximately

distributed as chi-square with \( k-p-q \)

degrees of freedom, where \( k \) is the number

of autocorrelations computed, \( p \) and \( q \)

is the order of the autoregressive and

moving average parameters, respectively.

Values of \( Q(\hat{r}) \) greater than the appro-

priate chi-square table value suggest

model misspecification.

Recent empirical studies (Davies, Triggs, & Newbold, 1977; Ljung & Box, 1978) have questioned the use of the Box-Pierce portmanteau statistic because \( i \) and \( h \) has been found to consistently produce values lower than asymptotic theory would suggest. This means that the statistic has low power—the conservative value may not detect a poorly identified model.

A refined version of the Box-Pierce statistic is included in this program. The modified portmanteau statistic

\[
Q(\hat{r}) = \frac{(N/\sum(N+2))}{\sum(N-1)^{r} \hat{r}^2}
\]

is also approximated as chi-square with \( k-p-q \) degrees of freedom. It has been shown to be a vast improvement over the Box-Pierce portmanteau statistic. Monte Carlo studies by Godfrey & Newbold (1979) have shown that for small to moderate \( N \), the modified portmanteau statistic more closely approximates the asymptotic chi-square distribution than does the Box-Pierce portmanteau statistic. Power studies indicate that the modified portmanteau statistic better controls Type I error rate, and thus is more sensitive to model misspecification than is the Box-Pierce statistic.

An alternative procedure for diagnostic checking included in the program is the Lagrange Multiplier test statistic (Godfrey, 1979). The LM statistic differs from the portmanteau statistic in that it does not require computation of residual autocorrelations. It does, however, require specification of an alternative hypothesis. The LM statistic tests the null hypothesis that the model is ARMA \( p,q \) against an alternative hypothesis that specifies an ARMA model either of order \( (p+r,q+r) \) or of order \( (p,q) \). The user specifies a value for \( r \) at the output of the program. The LM statistic is computed as \( N \) times the coefficient of determination of the regression of a set of carriers computed using the maximum likelihood estimates, the original series \( \{ z_t \} \), and the residuals \( \{ \hat{e}_t \} \). The LM statistic is approximately distributed as chi-square with \( r \) degrees of freedom. In Monte Carlo studies comparing both portmanteau statistics and the LM statistic, Godfrey found that for small \( N \) \( (N=50) \), the LM statistic comes closer to nominal significance levels than does the Box-Pierce statistic or the modified portmanteau statistic. However, for large \( N \) both the LM statistic and the modified portmanteau statistic are closely approximate the asymptotic chi-square distribution for the Box-Pierce portmanteau statistic.

More detailed information about Ansley's...
exact likelihood procedure, the modified portmanteau statistic, and the Lagrange multiplier test statistic are obtainable in the current literature. The interested reader should be aware of the limitations of the estimation and diagnostic procedures used in PROC ARIMA, and algorithms available to correct these discrepancies.


For more information:
Carol M. Reilly
c/o The Times Leader
15 N. Main St.
Wilkes-Barre, PA 18711