Data modeling: CSCI E-106

Applied Linear Statistical Models

Chapter 12 – Autocorrelation in Time Series Data

Autocorrelation

many regression applications involve time series data:

- the assumption of uncorrelated or independent error terms is often not appropriate;
- the error terms are frequently correlated positively over time

Error terms correlated over time are said to be autocorrelated or serially correlated.

Omission of the key variables could cause positive autocorrelation

- Regression of annual sales of a product on yearly price of the product over a period of 30 years.
  - Population size is not on the model ⇒ may lead to the error terms being positively autocorrelated

The presence of systematic coverage errors in the response variable time series ⇒ positively correlated over time.

Problems of Autocorrelation

1. The estimated regression coefficients are still unbiased, but they no longer have the minimum variance property and may be quite inefficient.
2. MSE may seriously underestimate the variance of the error terms.
3. $s\{b_k\}$ calculated according to ordinary least squares procedures may seriously underestimate the true standard deviation of the estimated regression coefficient.
4. Confidence intervals and tests using the t and F distributions, discussed earlier, are no longer strictly applicable.

Problems of Autocorrelation, cont'd

The simple linear regression model with time series data:

$$Y_t = \beta_0 + \beta_1 X_t + \varepsilon_t$$

Y_t and X_t are observations for period t. Let us assume that ε_t are positively autocorrelated as follows:

$$\varepsilon_t = \varepsilon_{t-1} + u_t$$

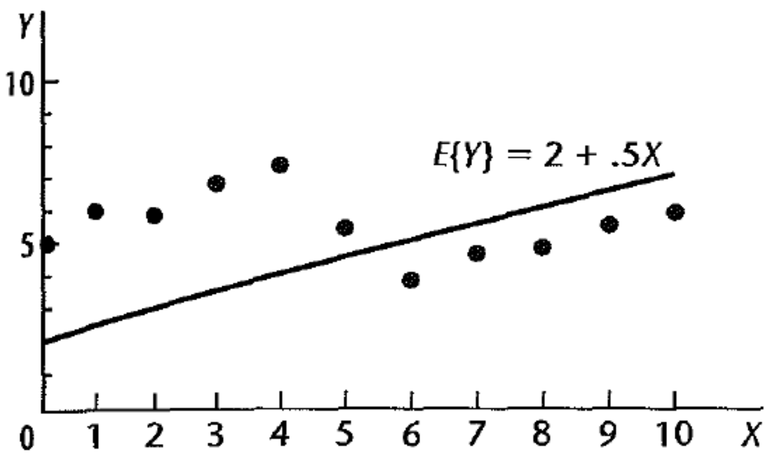
The u_t called disturbances and are i.i.d N(0,1) variables

If $\varepsilon_0 = 3$ $\Rightarrow$

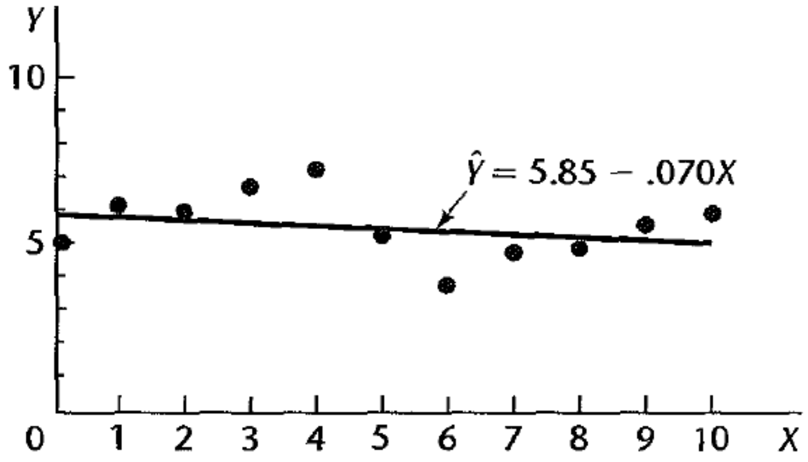
t	(1) u_t	(2) $\varepsilon_{t-1} + u_t = \varepsilon_t$	(3) $Y_t = 2 + .5X_t + \varepsilon_t$
0	—	3.0	5.0
1	.5	3.0 + .5 = 3.5	6.0
2	−.7	3.5 − .7 = 2.8	5.8
3	.3	2.8 + .3 = 3.1	6.6
4	0	3.1 + 0 = 3.1	7.1
5	−2.3	3.1 − 2.3 = .8	5.3
6	−1.9	.8 − 1.9 = −1.1	3.9
7	.2	−1.1 + .2 = −.9	4.6
8	−.3	− .9 − .3 = −1.2	4.8
9	.2	−1.2 + .2 = −1.0	5.5
10	−.1	−1.0 − .1 = −1.1	5.9

Problems of Autocorrelation, cont’d

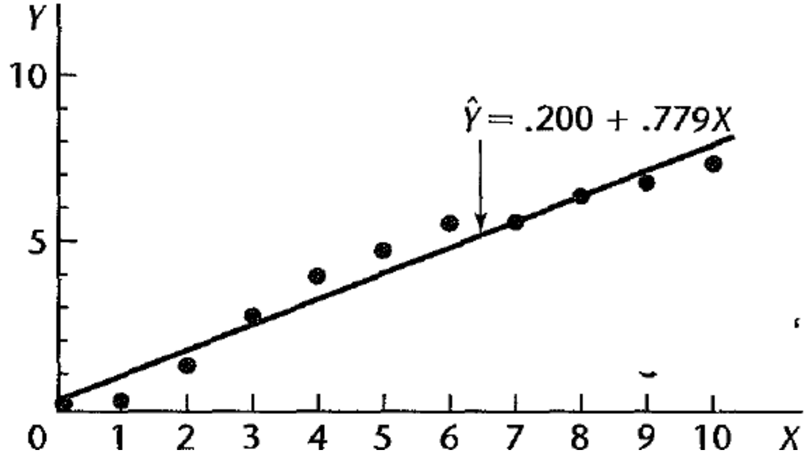
(a) True Regression Line and Observation when $\varepsilon_0 = 3$



(b) Fitted Regression Line and Observations when $\varepsilon_0 = 3$



(c) Fitted Regression Line and Observations with $\varepsilon_0 = -.2$ and Different Disturbances



- The fitted regression line differs sharply from the true regression line because the initial ε_0 value was large and the succeeding positively autocorrelated error terms tended to be large for some time. This persistency pattern in the positively autocorrelated error terms leads to a fitted regression line far from the true one.
- Had the initial ε_0 value been small, say, ε_0 = -.2, and the disturbances different, a sharply different fitted regression line might have been obtained because of the persistency pattern, as shown in c.

First-Order Autoregressive Error Model

Simple linear Regression

- The generalized simple linear regression model for one predictor variable when the random error terms follow a first-order autoregressive, or AR(1), process is:

$$Y_t = \beta_0 + \beta_1 X_t + \varepsilon_t$$
$$\varepsilon_t = \rho\varepsilon_{t-1} + u_t$$

where ρ is a parameter $|\rho| < 1$, and u_t are i.i.d $N(0, \sigma^2)$

- AR(1) is identical to the simple linear regression model except for the structure of the error terms.
- The parameter ρ is called the autocorrelation parameter.

First-Order Autoregressive Error Model, cont'd

Multiple linear Regression

- The generalized multiple linear regression model with AR(1) process:

$$Y_t = \beta_0 + \beta_1 X_{t1} + \beta_2 X_{t2} + \cdots + \beta_{p-1} X_{tp-1} + \varepsilon_t$$
$$\varepsilon_t = \rho \varepsilon_{t-1} + u_t$$

where $|\rho| < 1$, and u_t are i.i.d $N(0, \sigma^2)$

- AR(1) is identical to the multiple linear regression model except for the structure of the error terms.

First-Order Autoregressive Error Model, cont'd

Properties of Error Terms

The error terms still have mean zero and constant variance:

$$E\{\varepsilon_t\} = 0$$

$$\sigma^2\{\varepsilon_t\} = \frac{\sigma^2}{1-\rho^2}$$

- The covariance between adjacent error terms ε_t and ε_{t-1} is:

$$\sigma\{\varepsilon_t, \varepsilon_{t-1}\} = \rho\left(\frac{\sigma^2}{1-\rho^2}\right)$$

- The correlation between adjacent error terms ε_t and ε_{t-1} is

$$\rho\{\varepsilon_t, \varepsilon_{t-1}\} = \frac{\sigma\{\varepsilon_t, \varepsilon_{t-1}\}}{\sigma\{\varepsilon_t\}\sigma\{\varepsilon_{t-1}\}} = \frac{\rho\left(\frac{\sigma^2}{1-\rho^2}\right)}{\sqrt{\frac{\sigma^2}{1-\rho^2}}\sqrt{\frac{\sigma^2}{1-\rho^2}}} = \rho$$

First-Order Autoregressive Error Model, cont'd

Properties of Error Terms

The covariance between error terms that are s periods apart can be shown to be:

$\sigma\{\varepsilon_t, \varepsilon_{t-s}\} = \rho^s \left(\frac{\sigma^2}{1-\rho^2}\right)$, s$\neq 0$ is called autocovariance function.

The coefficient of correlation between ε_t and ε_{t-s} therefore is:

$\rho\{\varepsilon_t, \varepsilon_{t-s}\} = \rho^s$, s$\neq 0$

the variance-covariance matrix of the error terms for the AR(1) model is

$$\underset{n\times n}{\boldsymbol{\sigma}^2\{\boldsymbol{\varepsilon}\}} = \begin{bmatrix} \kappa & \kappa\rho & \kappa\rho^2 & \cdots & \kappa\rho^{n-1} \\ \kappa\rho & \kappa & \kappa\rho & \cdots & \kappa\rho^{n-2} \\ \vdots & \vdots & \vdots & & \vdots \\ \kappa\rho^{n-1} & \kappa\rho^{n-2} & \kappa\rho^{n-3} & \cdots & \kappa \end{bmatrix}$$

where, $\kappa = \dfrac{\sigma^2}{1-\rho^2}$

First-Order Autoregressive Error Model, cont'd

Comments:

$\varepsilon_t = \rho\varepsilon_{t-1} + u_t$, expand this definition for all t:

$$\varepsilon_{t-1} = \rho\varepsilon_{t-2} + u_{t-1}$$
$$\varepsilon_t = \rho(\rho\varepsilon_{t-2} + u_{t-1}) + u_t$$
$$\varepsilon_t = \rho^2\varepsilon_{t-2} + \rho u_{t-1} + u_t$$
$$\varepsilon_t = \rho^2(\rho\varepsilon_{t-3} + u_{t-2}) + \rho u_{t-1} + u_t$$
$$\varepsilon_t = \rho^3\varepsilon_{t-3} + \rho^2 u_{t-2} + \rho u_{t-1} + u_t$$

and so on:

$$\varepsilon_t = \sum_{s=0}^{\infty} \rho^s u_{t-s}$$

the ε_t is a linear combination of the current and preceding disturbance terms.

First-Order Autoregressive Error Model, cont'd

Comments:

Mean:

$$\varepsilon_t = \sum_{s=0}^{\infty} \rho^s u_{t-s}$$

$$E\{\varepsilon_t\} = \sum_{s=0}^{\infty} \rho^s E\{u_{t-s}\} = 0$$

Variance:

$$\sigma^2\{\varepsilon_t\} = \sum_{s=0}^{\infty} \rho^{2s} \sigma^2\{u_{t-s}\} = \sum_{s=0}^{\infty} \rho^{2s} \sigma^2 = \sigma^2 \sum_{s=0}^{\infty} \rho^{2s} = \sigma^2 \frac{1}{1-\rho^2}$$

Durbin-Watson Test for Autocorrelation for AR(1)

$$H_0: \rho = 0$$
$$H_a: \rho > 0$$

The Durbin-Watson test statistics D is obtained by using ordinary least squares to fit the regression function, calculating the ordinary residuals:

$$e_t = Y_t - \hat{Y}_t$$

Then D is:

$$D = \frac{\sum_{t=2}^{n}(e_t - e_{t-1})^2}{\sum_{t=1}^{n} e_t^2}$$

Exact critical values are difficult to obtain, but Durbin and Watson have obtained lower and upper bounds d_L and d_U such that a value of D outside these bounds leads to a definite decision. The decision rule for the test statistics

If $D > d_U$, conclude H_0

If $D < d_L$, conclude H_a

If $d_L \leq D \leq d_U$, the test is inconclusive

Durbin-Watson Test for Autocorrelation for AR(1)

Small values of D lead to the conclusion that $\rho > 0$ because the adjacent error terms ε_t and ε_{t-1} tend to be of the same magnitude when they are positively autocorrelated. Hence, the differences in the residuals, $e_t - e_{t-1}$ would tend to be small when $\rho > 0$, leading to small numerator in D and hence to a small test statistic D.

Table B.7 contains the bounds $\mathrm{d_L}$ and $\mathrm{d_U}$ for various sample sizes (n), for two levels of significance (.05 and .01), and for various numbers of X variables (p - 1) in the regression model.

TABLE B.7
Durbin-Watson Test Bounds.

	Level of Significance $\alpha = .05$									
	$p-1=1$		$p-1=2$		$p-1=3$		$p-1=4$		$p-1=5$	
n	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U
15	1.08	1.36	0.95	1.54	0.82	1.75	0.69	1.97	0.56	2.21
16	1.10	1.37	0.98	1.54	0.86	1.73	0.74	1.93	0.62	2.15
17	1.13	1.38	1.02	1.54	0.90	1.71	0.78	1.90	0.67	2.10
18	1.16	1.39	1.05	1.53	0.93	1.69	0.82	1.87	0.71	2.06
19	1.18	1.40	1.08	1.53	0.97	1.68	0.86	1.85	0.75	2.02
20	1.20	1.41	1.10	1.54	1.00	1.68	0.90	1.83	0.79	1.99
21	1.22	1.42	1.13	1.54	1.03	1.67	0.93	1.81	0.83	1.96
22	1.24	1.43	1.15	1.54	1.05	1.66	0.96	1.80	0.86	1.94
23	1.26	1.44	1.17	1.54	1.08	1.66	0.99	1.79	0.90	1.92
24	1.27	1.45	1.19	1.55	1.10	1.66	1.01	1.78	0.93	1.90
25	1.29	1.45	1.21	1.55	1.12	1.66	1.04	1.77	0.95	1.89
26	1.30	1.46	1.22	1.55	1.14	1.65	1.06	1.76	0.98	1.88
27	1.32	1.47	1.24	1.56	1.16	1.65	1.08	1.76	1.01	1.86
28	1.33	1.48	1.26	1.56	1.18	1.65	1.10	1.75	1.03	1.85
29	1.34	1.48	1.27	1.56	1.20	1.65	1.12	1.74	1.05	1.84
30	1.35	1.49	1.28	1.57	1.21	1.65	1.14	1.74	1.07	1.83
31	1.36	1.50	1.30	1.57	1.23	1.65	1.16	1.74	1.09	1.83
32	1.37	1.50	1.31	1.57	1.24	1.65	1.18	1.73	1.11	1.82
33	1.38	1.51	1.32	1.58	1.26	1.65	1.19	1.73	1.13	1.81
34	1.39	1.51	1.33	1.58	1.27	1.65	1.21	1.73	1.15	1.81
35	1.40	1.52	1.34	1.58	1.28	1.65	1.22	1.73	1.16	1.80
36	1.41	1.52	1.35	1.59	1.29	1.65	1.24	1.73	1.18	1.80
37	1.42	1.53	1.36	1.59	1.31	1.66	1.25	1.72	1.19	1.80
38	1.43	1.54	1.37	1.59	1.32	1.66	1.26	1.72	1.21	1.79
39	1.43	1.54	1.38	1.60	1.33	1.66	1.27	1.72	1.22	1.79
40	1.44	1.54	1.39	1.60	1.34	1.66	1.29	1.72	1.23	1.79
45	1.48	1.57	1.43	1.62	1.38	1.67	1.34	1.72	1.29	1.78
50	1.50	1.59	1.46	1.63	1.42	1.67	1.38	1.72	1.34	1.77
55	1.53	1.60	1.49	1.64	1.45	1.68	1.41	1.72	1.38	1.77
60	1.55	1.62	1.51	1.65	1.48	1.69	1.44	1.73	1.41	1.77
65	1.57	1.63	1.54	1.66	1.50	1.70	1.47	1.73	1.44	1.77
70	1.58	1.64	1.55	1.67	1.52	1.70	1.49	1.74	1.46	1.77
75	1.60	1.65	1.57	1.68	1.54	1.71	1.51	1.74	1.49	1.77
80	1.61	1.66	1.59	1.69	1.56	1.72	1.53	1.74	1.51	1.77
85	1.62	1.67	1.60	1.70	1.57	1.72	1.55	1.75	1.52	1.77
90	1.63	1.68	1.61	1.70	1.59	1.73	1.57	1.75	1.54	1.78
95	1.64	1.69	1.62	1.71	1.60	1.73	1.58	1.75	1.56	1.78
100	1.65	1.69	1.63	1.72	1.61	1.74	1.59	1.76	1.57	1.78

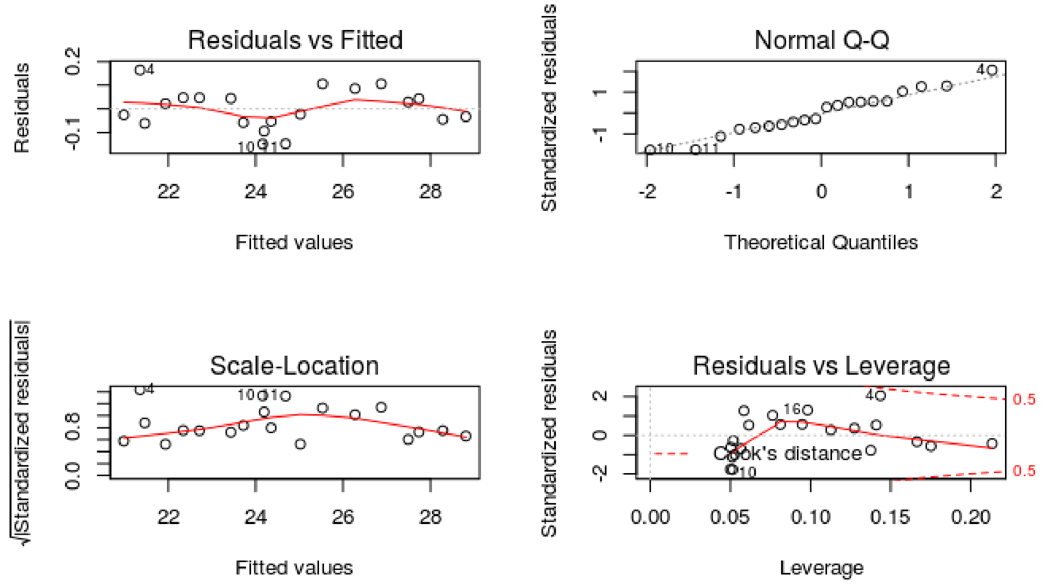
Example: Blaisdell Company

Year and Quarter	t	(1) Company Sales ($ millions) Y_t	(2) Industry Sales ($ millions) X_t	(3) Residual e_t	(4) $e_t - e_{t-1}$	(5) $(e_t - e_{t-1})^2$	(6) e_t^2
1998: 1	1	20.96	127.3	−.026052	—	—	.0006787
2	2	21.40	130.0	−.062015	−.035963	.0012933	.0038459
3	3	21.96	132.7	.022021	.084036	.0070620	.0004849
4	4	21.52	129.4	.163754	.141733	.0200882	.0268154
...	...	...	...	...	...	...	...
2002: 1	17	27.52	164.2	.029112	−.076990	.0059275	.0008475
2	18	27.78	165.6	.042316	.013204	.0001743	.0017906
3	19	28.24	168.7	−.044160	−.086476	.0074781	.0019501
4	20	28.78	171.7	−.033009	.011151	.0001243	.0010896
Total						.0979400	.1333018

$$\hat{Y} = -1.4548 + .17628X$$

$$s\{b_0\} = .21415 \qquad s\{b_1\} = .00144$$

$$MSE = .00741$$

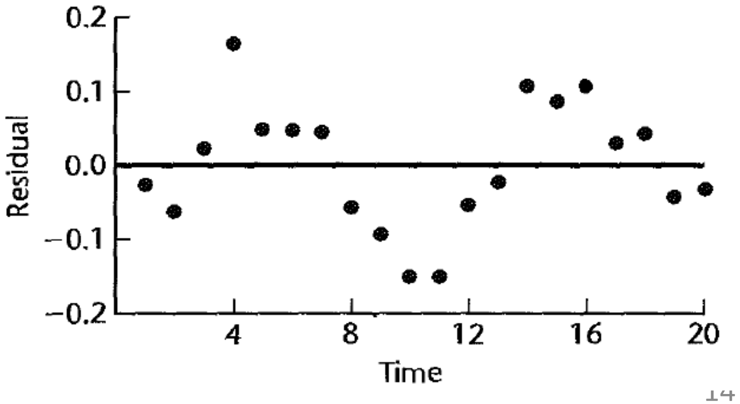


For level of significance of .01, we find in Table B.7 for n = 20 and p - 1 = 1: d_L = .95 d_u = 1.15.

D= 0.09794/0.133 = 0.735

Since D = .735 falls below d_L = .95, reject Ho: the error terms are positively autocorrelated.

FIGURE 12.3 Residuals Plotted against Time—Blaisdell Company Example.



Example: Blaisdell Company – Rcode

```
> f<-lm(Company.Sales~Industry.Sales)
> summary(f)

Call:
lm(formula = Company.Sales ~ Industry.Sales)

Residuals:
      Min        1Q    Median        3Q       Max 
-0.149142 -0.054399 -0.000454  0.046425  0.163754 

Coefficients:
                Estimate Std. Error t value Pr(>|t|)    
(Intercept)    -1.454750   0.214146  -6.793 2.31e-06 ***
Industry.Sales  0.176283   0.001445 122.017  < 2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.08606 on 18 degrees of freedom
Multiple R-squared:  0.9988,  Adjusted R-squared:  0.9987 
F-statistic: 1.489e+04 on 1 and 18 DF,  p-value: < 2.2e-16

> anova(f)
Analysis of Variance Table

Response: Company.Sales
               Df  Sum Sq Mean Sq F value    Pr(>F)    
Industry.Sales  1 110.257 110.257   14888 < 2.2e-16 ***
Residuals      18   0.133   0.007                      
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
>
```

Example: Blaisdell Company – Rcode ,cont’d

```
library(Hmisc)
et1<-Lag(et, shift = 1)
> et1
          1           2           3           4           5           6           7           8 
         NA -0.02605186 -0.06201545  0.02202096  0.16375424  0.04657049  0.04637659  0.04361706 
          9          10          11          12          13          14          15          16 
-0.05843543 -0.09439903 -0.14914246 -0.14799090 -0.05305356 -0.02292824  0.10585161  0.08546380 
         17          18          19          20 
 0.10610224  0.02911240  0.04231646 -0.04416025 

d1<-na.omit(et-et1)
> sum(d1^2)
[1] 0.09794062
> sum(et^2)
[1] 0.1333023
> 0.09794062/0.1333023
[1] 0.7347257
```

Example: Blaisdell Company – Rcode ,cont'd

1. If a test for negative autocorrelation is required, the test statistic to be used is 4 - *D,* where *D* is defined as above. The test is then conducted in the same manner described for testing for positive autocorrelation. That is, if the quantity 4 - *D* falls below d_L , we conclude $\rho < 0$, that negative autocorrelation exists, and so on.

2. A two-sided test for Ho: $\rho = 0$ versus Ha: $\rho \neq 0$ can be made by employing both one-sided tests separately. The Type I risk with the two-sided test is 2α. where α is the Type I risk for each one-sided test.

3. When the Durbin-Watson test employing the bounds dL and dU gives indeterminate results, in principle more cases are required. Of course, with time series data it may be impossible to obtain more cases, or additional cases may lie in the future and be obtainable only with great delay. A reasonable procedure is to treat indeterminate results as suggesting the presence of autocorrelated errors and employ one of the remedial actions to be discussed next.
4. The Durbin-Watson test is not robust against misspecifications of the model.

5. The Durbin-Watson test is widely used; however, other tests for autocorrelation are available. One such test, due to Theil and Nagar, is found in Reference 12.3.

Remedial Measures for Autocorrelation

1. Addition of Predictor Variables
2. Use of Transformed Variables

Consider the transformed dependent variable:

$Y_t' = Y_t - \rho Y_{t-1}$

$Y_t' = \beta_0 + \beta_1 X_t + \varepsilon_t - \rho(\beta_0 + \beta_1 X_{t-1} + \varepsilon_{t-1})$

$Y_t' = \beta_0(1-\rho) + \beta_1(X_t - \rho X_{t-1}) + \varepsilon_t + \rho\varepsilon_{t-1}$

$Y_t' = \beta_0(1-\rho) + \beta_1(X_t - \rho X_{t-1}) + u_t$

Let $X_t' = X_t - \rho X_{t-1}$ *and* $\beta_0' = \beta_0(1-\rho)$, *and* $\beta_1' = \beta_1$

$\Rightarrow Y_t' = \beta_0' + \beta_1' X_t' + u_t$

Remedial Measures for Autocorrelation, cont'd

$$Y_t' = \beta_0' + \beta_1' X_t' + u_t$$

one generally needs to estimate the autocorrelation parameter ρ since its value is usually unknown. Once an estimate of ρ has been obtained, to be denoted by r, transformed variables are obtained using this estimate of ρ:

$$Y_t' = Y_t - rY_{t-1}$$
$$X_t' = X_t - rX_{t-1}$$
$$Y_t' = b_0' + b_1' X_t'$$

If this fitted regression function has eliminated the autocorrelation in the error terms, we can transform back to a fitted regression model in the original variables as follows:

$$\hat{Y} = b_0 + b_1 X$$

Where, $b_0 = \frac{b_0'}{1-r}$ and $b_0 = b_1' \Rightarrow s\{b_0\} = \frac{s\{b_0'\}}{1-r}$ and $s\{b_0\} = \mathrm{s}\{b_1'\}$

Cochrane-Orcutt Procedure

The Cochrane-Orcutt procedure involves an iteration of three steps:

1. estimate of ρ :

$$\varepsilon_t = \rho\varepsilon_{t-1} + u_t$$

Since ε_t and ε_{t-1} are unknown, we use the residuals e_t, and e_{t-1} obtained by OLS by fitting a straight line through the origin ($e_t = re_{t-1}$). The r is an estimate of ρ :

$$r = \frac{\sum_{t=2}^{n} e_{t-1}e_t}{\sum_{t=2}^{n} e_{t-1}^2}$$

2. Fitting of transformed model: Using the estimate r, we next obtain the transformed variables and fit the regression model

$$Y_t' = Y_t - rY_{t-1}$$
$$X_t' = X_t - rX_{t-1}$$
$$Y_t' = b_0' + b_1'X_t'$$

Cochrane-Orcutt Procedure, cont'd

The Cochrane-Orcutt procedure involves an iteration of three steps:

3. Test for need to iterate. The Durbin-Watson (DW) test is then employed to test whether the error terms for the transformed model are uncorrelated. If the test indicates that they are uncorrelated, the procedure terminates. The fitted regression model in the original variables is then obtained by transforming the regression coefficients back according to (12.20).

If the DW test indicates that autocorrelation is still present after the first iteration, the parameter ρ is reestimated from the new residuals for the fitted regression model (12.20) with the original variables, which was derived from the fitted regression model (12.19) with the transformed variables. A new set of transformed variables is then obtained with the new r. This process may be continued for another iteration or two until the Durbin Watson test suggests that the error terms in the transformed model are uncorrelated. If the process does not terminate after one or two iterations, a different procedure should be employed.

Example: Blaisdell

TABLE 12.4 Transformed Variables and Regression Results for First Iteration with Cochrane-Orcutt Procedure—Blaisdell Company Example.

t	(1) Y_t	(2) X_t	(3) $Y'_t = Y_t - .631166Y_{t-1}$	(4) $X'_t = X_t - .631166X_{t-1}$
1	20.96	127.3	—	—
2	21.40	130.0	8.1708	49.653
3	21.96	132.7	8.4530	50.648
4	21.52	129.4	7.6596	45.644
...	...	...	...	...
17	27.52	164.2	10.4911	62.772
18	27.78	165.6	10.4103	61.963
19	28.24	168.7	10.7062	64.179
20	28.78	171.7	10.9559	65.222

$$\hat{Y}' = -.3941 + .17376X'$$

$$s\{b'_0\} = .1672 \qquad s\{b'_1\} = .002957$$

$$MSE = .00451$$

Table 12.2. Column 2 contains the residuals e_{t-1}, and columns 3 and 4 contain the necessary calculations. Hence, we estimate:

$$r = \frac{.0834478}{.1322122} = .631166$$

We now obtain the transformed variables Y'_t and X'_t in (12.18):

$$Y'_t = Y_t - .631166Y_{t-1}$$

$$X'_t = X_t - .631166X_{t-1}$$

Example: Blaisdell, cont’d

OLS is now used with these transformed variables based on the n – 1 cases remaining after the transformations. The fitted regression line in the transform:
variables is:

$$\hat{Y}' = -.3941 + .17376X'$$

$$Y_t' = Y_t - .631166Y_{t-1}$$

$$X_t' = X_t - .631166X_{t-1}$$

Durbin-Watson is statistic calculated. D = 1.65. From Table B.7, we find for α = .01, p-1 = 1 and n= 19: d_L*=0.93 and* d_U *=1.13*
Since D = 1.65 > d_U = 1.13, we conclude that the autocorrelation coefficient for the error term is zero. Having successfully handled the problem of auto correlated error terms, we now transform the fitted model in (12.23) back to the original variables, using (12.20):

$$b_0 = \frac{b_0'}{1-r} = \frac{-.3941}{1-.631166} = -1.0685$$

$$b_1 = b_1' = .17376$$

leading to the fitted regression function in the original variables:

$$\hat{Y} = -1.0685 + .17376X$$

Example: Blaisdell, cont'd

1. The Cochrane-Orcutt approach does not always work properly. The estimate r tends to underestimate the autocorrelation parameter ρ. When this bias is serious, it can significantly reduce the effectiveness of the Cochrane-Orcutt approach.
2. There exists an approximate relation between the Durbin-Watson test statistic D and the estimated autocorrelation parameter r

$$D \approx 2(1-r)$$

This relation indicates that the Durbin-Watson statistic ranges approximately between 0 and 4 since,. takes on values between -1 and 1, and that D is approximately 2 when r = 0.

Example: Blaisdell, cont'd

```
> f<-lm(Company.Sales~Industry.Sales)
> summary(f)
Call:
lm(formula = Company.Sales ~ Industry.Sales)

Residuals:
      Min        1Q    Median        3Q       Max
-0.149142 -0.054399 -0.000454  0.046425  0.163754

Coefficients:
               Estimate Std. Error t value Pr(>|t|)
(Intercept)    -1.454750   0.214146  -6.793 2.31e-06 ***
Industry.Sales  0.176283   0.001445 122.017  < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Residual standard error: 0.08606 on 18 degrees of freedom
Multiple R-squared:  0.9988,  Adjusted R-squared:  0.9987
F-statistic: 1.489e+04 on 1 and 18 DF,  p-value: < 2.2e-16
> anova(f)
Analysis of Variance Table

Response: Company.Sales
               Df  Sum Sq Mean Sq F value    Pr(>F)
Industry.Sales  1 110.257 110.257   14888 < 2.2e-16 ***
Residuals      18   0.133   0.007
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
m1<-sum(na.omit(et1*et))
m2<-sum(na.omit(et1^2))
> r<-m1/m2
> r
[1] 0.6311636
```

Example: Blaisdell, cont’d

```
> Yt1<-Lag(Company.Sales, shift = 1)
> Yt<-Company.Sales
> Xt<-Industry.Sales
> Xt1<-Lag(Xt, shift = 1)
> Y.new<-na.omit(Yt-r*Yt1)
X.new<-na.omit(Xt-r*Xt1)
> f2<-lm(Y.new~X.new)
> summary(f2)
Call:
lm(formula = Y.new ~ X.new)
Residuals:
      Min        1Q    Median        3Q       Max
-0.097039 -0.056815  0.009902  0.034553  0.125048
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -0.394111   0.167230  -2.357   0.0307 *
X.new        0.173758   0.002957  58.767   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
Residual standard error: 0.06715 on 17 degrees of freedom
Multiple R-squared:  0.9951,  Adjusted R-squared:  0.9948
F-statistic:  3454 on 1 and 17 DF,  p-value: < 2.2e-16
```

Hildreth-Lu Procedure

The Hildreth-Lu procedure for estimating the autocorrelation parameter ρ that minimizes the error sum of squares for the transformed regression model:

$$SSE = \sum \left(Y_t' - \hat{Y}_t'\right)^2 = \sum (Y_t' - b_0' - b_1' X_t')^2$$

Use numerical search to find the optimal value of ρ.

Hildreth-Lu Procedure, example

$$\hat{Y}' = .07117 + .16045X'$$

TABLE 12.5 Hildreth-Lu Results—Blaisdell Company Example.

ρ	SSE	ρ	SSE
.10	.1170	.94	.0718
.30	.0938	.95	.07171
.50	.0805	.96	.07167
.70	.0758	.97	.07175
.90	.0728	.98	.07197
.92	.0723		

For $\rho = .96$: $\hat{Y}' = .07117 + .16045X'$

$$s\{b_0'\} = .05798 \qquad s\{b_1'\} = .006840$$

$$MSE = .00422$$

where:

$$Y_t' = Y_t - .96Y_{t-1}$$
$$X_t' = X_t - .96X_{t-1}$$

The Durbin-Watson test statistic for this fitted model is D = 1.73. Since for n =19 p - 1 =1 , and α = .01, du = 1.13, we conclude that no autocorrelation remains in the transformed model.

Therefore, we shall transform regression function back to the original variables.

$$\hat{Y} = 1.7793 + .16045X$$

The estimated standard deviations of these regression coefficients are:

$$s\{b_0\} = 1.450 \qquad s\{b_1\} = .006840$$

First Difference Procedure

Since the autocorrelation parameter ρ is frequently large and *SSE* as a function of ρ often is quite flat for large values of ρ up to 1.0, as in the Blaisdell Company example, some economists and statisticians have suggested use of ρ = 1.0 in the transformed model.

If ρ =1 , β_0' = β_0 *(1 - ρ)* = 0, and the transformed model becomes:

$$Y_t' = \beta_1' X_t' + u_t$$

Where,

$$Y_t' = Y_t - Y_{t-1}$$
$$X_t' = X_t - X_{t-1}$$

The fitted regression function in the transformed variables:

$$\hat{Y}_t' = b_1' X_t'$$

First Difference Procedure

Since the autocorrelation parameter ρ is frequently large and *SSE* as a function of ρ often is quite flat for large values of ρ up to 1.0, as in the Blaisdell Company example, some economists and statisticians have suggested use of ρ = 1.0 in the transformed model.

If ρ =1 , β_0' = β_0 *(1 - ρ)* = 0, and the transformed model becomes:

$$Y_t' = \beta_1' X_t' + u_t$$

Where,

$$Y_t' = Y_t - Y_{t-1}$$
$$X_t' = X_t - X_{t-1}$$

The fitted regression function in the transformed variables:

$$\hat{Y}_t' = b_1' X_t'$$

First Difference Procedure, cont'd

The fitted regression function can be transformed back to the original variables as follows:

$$\hat{Y} = b_0 + b_1 X$$

Where,

$$b_0 = \bar{Y} - b_1' \bar{X}$$
$$b_1 = b_1'$$

Use double sided DW test statistics for the first difference.

Example: Blaisdell

```
> Y.diff<-Yt-Yt1
> X.diff<-Xt-Xt1
> f.diff<-lm(Y.diff ~ X.diff - 1)
> summary(f.diff)

Call:
lm(formula = Y.diff ~ X.diff - 1)

Residuals:
     Min       1Q   Median       3Q      Max 
-0.08958 -0.03231  0.02412  0.05344  0.15139 

Coefficients:
       Estimate Std. Error t value Pr(>|t|)    
X.diff 0.168488   0.005096   33.06   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.06939 on 18 degrees of freedom
  (1 observation deleted due to missingness)
Multiple R-squared:  0.9838,	Adjusted R-squared:  0.9829 
F-statistic:  1093 on 1 and 18 DF,  p-value: < 2.2e-16
```

$\Rightarrow$

$$\hat{Y}' = .16849X'$$

$$Y'_t = Y_t - Y_{t-1}$$
$$X'_t = X_t - X_{t-1}$$

Example, cont'd

```
> f3<-lm(Y.diff~X.diff)
> et<-f3$residuals
> et1<-Lag(et,shift=1)
> num1<-na.omit(et-et1)
> num2<-sum(num1^2)
> num2
[1] 0.127541
>
> den<-sum(et^2)
>
> den
[1] 0.07292917
> 0.127541/0.07292917
[1] 1.748834
```

$$\hat{Y} = -.30349 + .16849X$$

t	(1) Y_t	(2) X_t	(3) $Y'_t = Y_t - Y_{t-1}$	(4) $X'_t = X_t - X_{t-1}$
1	20.96	127.3	—	—
2	21.40	130.0	.44	2.7
3	21.96	132.7	.56	2.7
4	21.52	129.4	−.44	−3.3
...	...	...	...	...
17	27.52	164.2	.54	3.5
18	27.78	165.6	.26	1.4
19	28.24	168.7	.46	3.1
20	28.78	171.7	.54	3.0

$$\hat{Y}' = .16849X'$$

$$s\{b'_1\} = .005096 \qquad MSE = .00482$$

D = 1.75. This indicates uncorrelated error terms for either a one-sided test (with α = .01) or a two-sided test (with α = .02).

Example, cont'd

Procedure	b_1	$s\{b_1\}$	r	Estimate of σ^2 (MSE)
Cochrane-Orcutt	.1738	.0030	.63	.0045
Hildreth-Lu	.1605	.0068	.96	.0042
First differences	.1685	.0051	1.0	.0048
Ordinary least squares	.1763	.0014	—	—

1. All of the estimates of β are quite close to each other.
2. The estimated standard deviations of b_1 based on Hildreth-Lu and first differences transformation methods are quite close to each other; that with the Cochrane-Orcutt procedure is somewhat smaller. The estimated standard deviation of b_1 based on ordinary least squares regression with the original variables is still smaller. This is as expected, since the estimated standard deviations s{ bk } may be seriously underestimated
3. All three transformation methods provide essentially the same estimate of MSE

Forecasting with Autocorrelated Error Terms

$$Y_t = \beta_0 + \beta_1 X_t + \varepsilon_t$$
$$\varepsilon_t = \rho \varepsilon_{t-1} + u_t$$

We obtain:

$$Y_t = \beta_0 + \beta_1 X_t + \rho \varepsilon_{t-1} + u_t$$

For period n+1, we obtain:

$$Y_{n+1} = \beta_0 + \beta_1 X_{n+1} + \rho \varepsilon_n + u_{n+1}$$

Thus Y_{n+1} has three components:

1. The expected value $\beta_0 + \beta_1 X_{n+1}$.
2. A multiple ρ of the preceding error term ε_n.
3. An independent, random disturbance term with E{u_{n+1}} = 0.

Forecasting with Autocorrelated Error Terms, cont'd

The forecast for next period $n + 1$, to be denoted by F_{n+1} , is constructed by dealing with each of the three components:

1. Given X_{n+1}, we estimate the expected value $\beta_0 + \beta_1 X_{n+1}$ as usual from the fitted regression function:

$$\hat{Y}_{n+1} = b_0 + b_1 X_{n+1}$$

where b_0 and b_1 are the estimated regression coefficients for the original variables obtained from b_0' and b_1' for the transformed variables according to (12.20).

2. ρ is estimated by r in (12.22), and ε_n is estimated by the residual e_n:

$$e_n = Y_n - (b_0 + b_1 X_n) = Y_n - \hat{Y}_n$$

Thus, $\rho\varepsilon_n$ is estimated by re_n

3. The disturbance term U_{n+1} has expected value zero and is independent of earlier information. Hence, we use its expected value of zero in the forecast.

Forecasting with Autocorrelated Error Terms, cont'd

Thus, the forecast for period n + 1 is:

$$F_{n+1} = \hat{Y}_{n+1} + re_n$$

An approximate I - α prediction interval for $Y_{n+1(new)}$ may be obtained by employing the usual prediction limits for a new observation in (2.36), but based on the transformed observations. Thus, Y_i and X_i in formula (2.38a) for the estimated variance s^2\{pred\} are replaced by Y_t' and X_t';

An approximate I - α prediction interval for $Y_{n+1(new)}$ are:

$$F_{n+1} \pm t(1 - \frac{\alpha}{2}; n - 3)\text{s\{pred\}}$$

Note that there are only n - 1 transform cases and two degrees of freedom are lost for estimating the two parameters in the OLS.

Example

For the Blaisdell Company example, the trade association has projected that deseasonalized industry sales in the first quarter of 2003 (i.e., quarter 21) will be X_{21} = \$175.3 million. To forecast Blaisdell Company sales for quarter 21, we shall use the Cochrane-Orcutt fitted regression function (12.24):

$$\hat{Y} = -1.0685 + .17376X$$

First, we need to obtain the residual e_{20}:

$$e_{20} = Y_{20} - \hat{Y}_{20} = 28.78 - [-1.0685 + .17376(171.7)] = .0139$$

The fitted value when $X_{21} = 175.3$ is:

$$\hat{Y}_{21} = -1.0685 + .17376(175.3) = 29.392$$

The forecast for period 21 then is:

$$F_{21} = \hat{Y}_{21} + re_{20} = 29.392 + .631166(.0139) = 29.40$$

Note how the fact that company sales in quarter 20 were slightly above their estimated mean has a small positive influence on the forecast for company sales for quarter 21.

We wish to set up a 95 percent prediction interval for $Y_{21(\text{new})}$. Using the data for the transformed variables in Table 12.4, we calculate $s\{\text{pred}\}$ by (2.38) for:

$$X'_{n+1} = X_{n+1} - .631166X_n = 175.3 - .631166(171.7) = 66.929$$

We obtain $s\{\text{pred}\} = .0757$ (calculations not shown). We require $t(.975; 17) = 2.110$. We therefore obtain the prediction limits $29.40 \pm 2.110(.0757)$ and the prediction interval:

$$29.24 \leq Y_{21(\text{new})} \leq 29.56$$