ME384Q.3 / ORI 390R.3: Time-Series Analysis

Prof. Dragan Djurdjanovic

Homework 1 - Solutions

Problem 1:

Part (a)

The code used to solve this problem is below.

```
%This is the m-code used to solve problem 2 in Homework 1 of the
%time-series class

%x is the variable containing all the data samples and should be
available
%in your workspace in order for this code to work. Variable x should be
a
%column vector of dimenstion 50 (i.e. of dimensionality 50 by 1).


x=x-mean(x); %removing the mean from the time series. Not too important
in
%this problem since mean of this time series is 0 any way.

N=length(x); %This is 50 actually)
Y=x(2:N);
X=[x(1:N-1)];
Phi=inv(X'*X)*X'*Y; %This gives AR(1) coefficient \phi1
Residuals=Y-Phi*X; %This gives residuals
SigmaAtSquare=(Residuals'*Residuals)/(N-2); %I estimated the mean and
the
%coefficient phi1, so there are 2 degrees of freedom lost.
%SigmaAtSquare determines the variance of the residuals.

VarPhi=SigmaAtSquare*inv(X'*X);
StDevPhi=sqrt(VarPhi);
```

Results that came out gave an AR(1) model with estimate of the autoregressive coefficient ϕ_1 being

$$\hat{\phi}_1 = -0.0832$$

and standard deviation of that estimate being

$$\sigma_{\hat{\phi}_1} = 0.1425$$

Part (b)

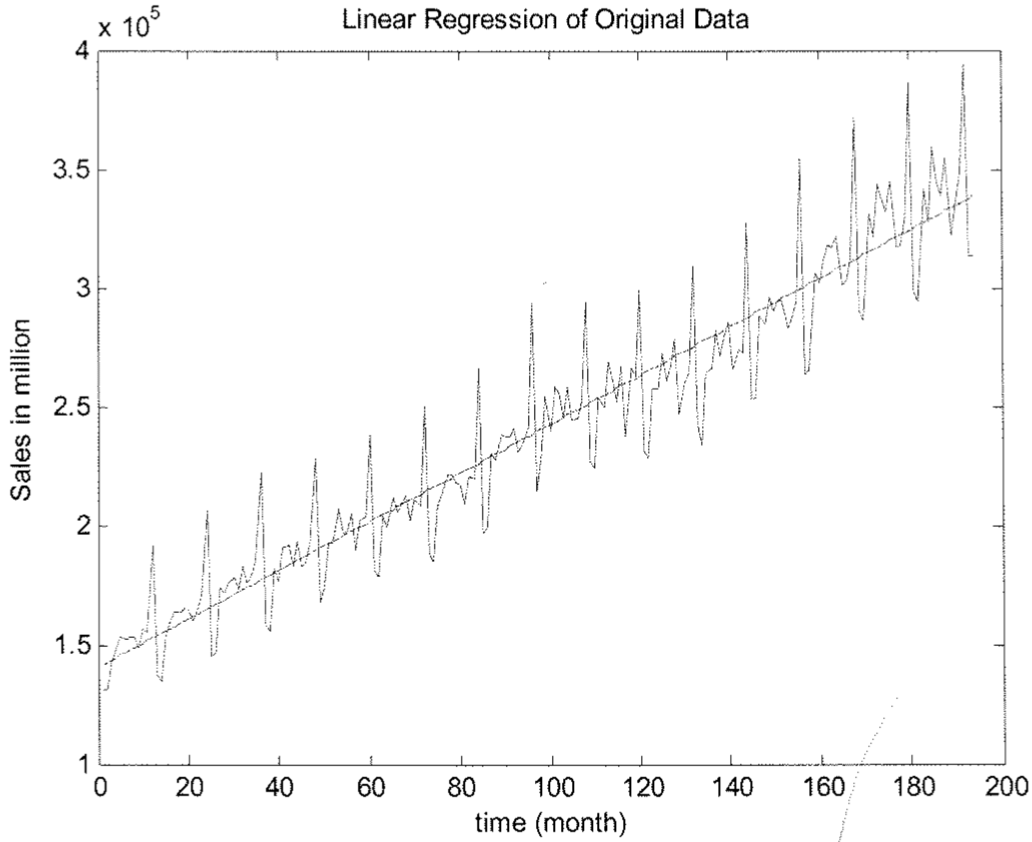
Obviously, zero is well within the plus/minus 2 sigma (or even 1 sigma) interval of the estimate of the AR(1) coefficient. Seems that simple model the AR(0) form

$$x_t = a_t$$

where a_t is a Gaussian white noise, describes this data well. In other words, seems that the time series is simple white noise – and it actually is. I generated it by asking the computer to make 50 Gaussian, independent numbers and this is what we got.

Problem 2:

a)



$\hat{\beta}_0$=b0 = 1.4053e+005

$\hat{\beta}_1$=b1 = 1.0210e+003

$\rightarrow \hat{y}_t = \hat{\beta}_0 + \hat{\beta}_1 z_t$

$$
\begin{array}{c}
y_t \\
\begin{bmatrix} y_1 \\ \vdots \\ y_{194} \end{bmatrix} \\
\underline{Y}
\end{array}
=
\begin{array}{c}
\begin{array}{cc} 1 & z_t \end{array} \\
\begin{bmatrix} 1 & z_1 \\ 1 & z_2 \\ \vdots & \vdots \\ 1 & z_{194} \end{bmatrix} \\
\underline{X}
\end{array}
\begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix}
+
\begin{array}{c}
\begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_{194} \end{bmatrix} \\
\underline{\varepsilon}
\end{array}
$$

$$
\begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \end{bmatrix} = (\underline{X}^T \underline{X})^{-1} (\underline{X}^T \underline{Y})
$$

```
clear all;
DATA_Original = xlsread('DATA.xls');

%Linear Regression
Y=DATA_Original(:,3);
N=length(Y);
X=[ones(N,1),[1:N]'];
beta=inv(X'*X)*(X'*Y);
b0=beta(1);
b1=beta(2);
plot([1:N]',Y,[1:N]',(b0+b1*[1:N]'));
ylabel('Sales in million'); xlabel('time (month)');
title('Linear Regression of Original Data');
```

(b)

Residual

$$\dot{x}_t = y_t - \hat{y}_t = y_t - \left(\hat{\beta}_0 + \hat{\beta}_1 z_t\right) \text{ (before normalization)}$$

$$x_t = \dot{x}_t - \bar{x} \text{ (after normalization)}$$

AR(2) Model

$$x_t = \phi_1 x_{t-1} + \phi_2 x_{t-2} + a_t \qquad t = 1 \sim N = 194 \qquad a_t \sim NID\ N(0, \sigma_a^2)$$

$\hat{\phi}_1$ =phi1 = -0.0520

$\hat{\phi}_2$ =phi2 = -0.2449

$$\hat{x}_t = \hat{\phi}_1 x_{t-1} + \hat{\phi}_2 x_{t-2}$$

(c)

```
clear all;
DATA_Original = xlsread('DATA.xls');

%Linear Regression
Y=DATA_Original(:,3);
N=length(Y);
X=[ones(N,1),[1:N]'];
beta=inv(X'*X)*(X'*Y);
b0=beta(1);
b1=beta(2);
plot([1:N]',Y,[1:N]',(b0+b1*[1:N]'));
ylabel('Sales in million'); xlabel('time (month)');
title('Linear Regression of Original Data');
```

RSS_AR2=6.5742e+010;
Var_AR2_at=3.4601e+008;

Nice Matlab code you can use to get everything.

```
Data_original=xlsread('Data.xls');
%Finding the linear regression model
Y=Data_original(:,3);
N=length(Y);
X=[ones(N,1),[1:N]'];
beta=inv(X'*X)*(X'*Y);
b0=beta(1);
b1=beta(2);
Data_residual=(Y-(b0+b1*[1:N]'));
xtdot=Data_residual;
xtbar=mean(xtdot);
xt=xtdot-xtbar;

%Fit the residual of linear regression model with AR(2)
X=[xt(2:N-1,1),xt(1:N-2,1)];
Y=[xt(3:N,1)];
phi_hat=inv(X'*X)*(X'*Y);
phi1=phi_hat(1);
phi2=phi_hat(2);
% Calculate the residual of the AR(2) model
Residual_AR2=Y-phi1*X(:,1)-phi2*X(:,2);
plot(Residual_AR2);
%RSS of the AR(2)
RSS_AR2=sum(Residual_AR2.^2);
%This is variance a_t^2
Var_AR2_at=1/(N-2-2)*RSS_AR2;
```

(d)

Figure 2.1 shows residual sums of squares (RSS) obtained for increasing orders of AR models fit to the residuals obtained after the linear fit you made to the data (residuals after you drew a lone through wholesales data when you plot them versus time). You can see how RSS drops significantly until AR(14), after which the drop becomes very small. Seems like we do not need to go much beyond AR(14).

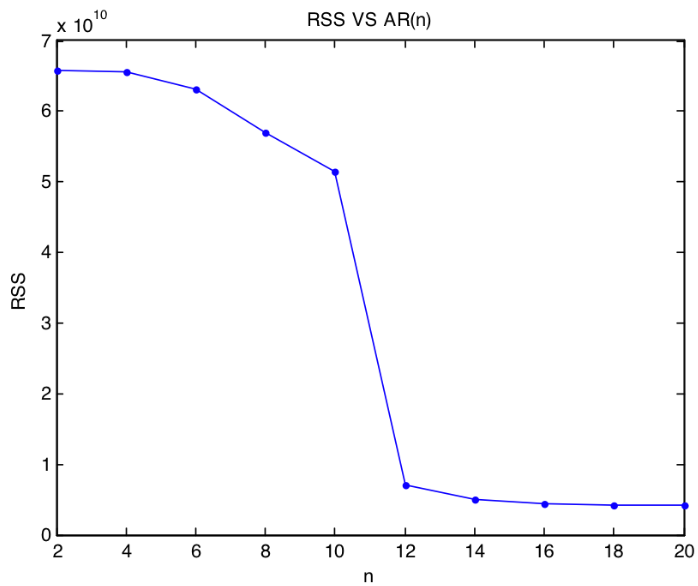


Figure 2.1: Residual sums of squares for AR(n) models fit to the residuals obtained after a linear trend was removed from the wholesales data.