Solution to Project 1

This project deals with the winter weather in South Bend: Just how cold was it last year compared to historical values over the last 30 years, and can the Southern Oscillation Index provide a useful prediction for winter weather.

Contents

- Parts 1 - 3
- Part 4
- Part 5
- Part 6
- Part 7

Parts 1 - 3

We import the data files for the temperature and the southern oscillation index. We have converted the temperature file to an xls spreadsheet and trimmed it out so that only the data from 1983-2014 are present. Likewise, we have trimmed the SOI data so that only 1982-2013 are present. In both cases the data files are two column, the first being the date and the second being the temperature or soi data.

```
clc;
clear;

temps=xlsread('tempdata.xls',1,'','basic');
soi=xlsread('soidata.xls',1,'','basic');

%We reshape the datafiles:
dates=reshape(temps(:,1),12,32);
mntemps=reshape(temps(:,2),12,32);
soi=reshape(soi(:,2),12,32);

%We also convert to Farenheit:
mntemps=mntemps/10*1.8+32;

%We split off the old temperatures:
tempold=mntemps(:,1:31);
temp2014=mntemps(:,32);

%We get the monthly averages over the period from 1983-2013:
avgtemp=mean(tempold')'; %Note that mean averages over columns, so we use '
```

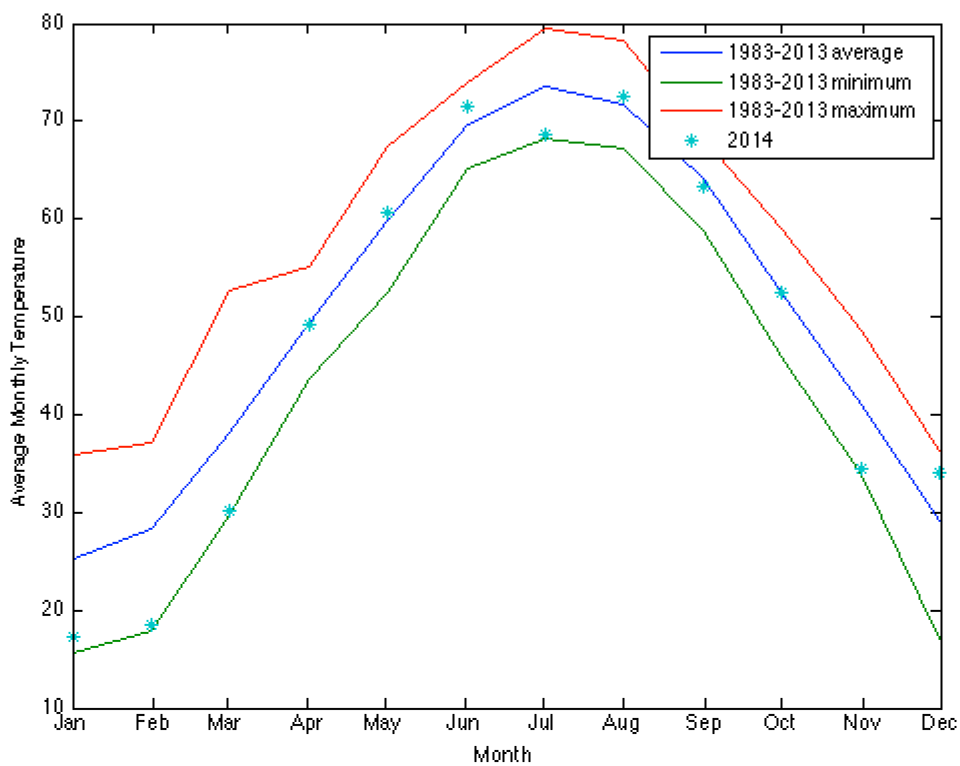
```
Warning: XLSREAD has limited import functionality on in basic mode.  Refer to HELP
XLSREAD for more information.
Warning: XLSREAD has limited import functionality on in basic mode.  Refer to HELP
XLSREAD for more information.
```

Part 4

We do a little analysis and graphics.

```
%First, we want to plot up the average, minimum, maximum and 2014 temps for
%each month.  We need to construct a vector of "dates" to plot against.  By
```

```
%default matlab deals in days, so we get a reasonable result from the
%vector:
days=[1:365/12:365]; %Note: you can also get this using the datenum command
figure(1)
plot(days,avgtemp,days,min(tempold'),days,max(tempold'),days,temp2014,'*')
datetick('x','mmm')
xlabel('Month')
ylabel('Average Monthly Temperature')
legend('1983-2013 average','1983-2013 minimum','1983-2013 maximum','2014')

%From the plot we see that in 2014 February, March, July, and November
%matched the coldest months over a 31 year period.  It was certainly a cold
%year!
```



Part 5

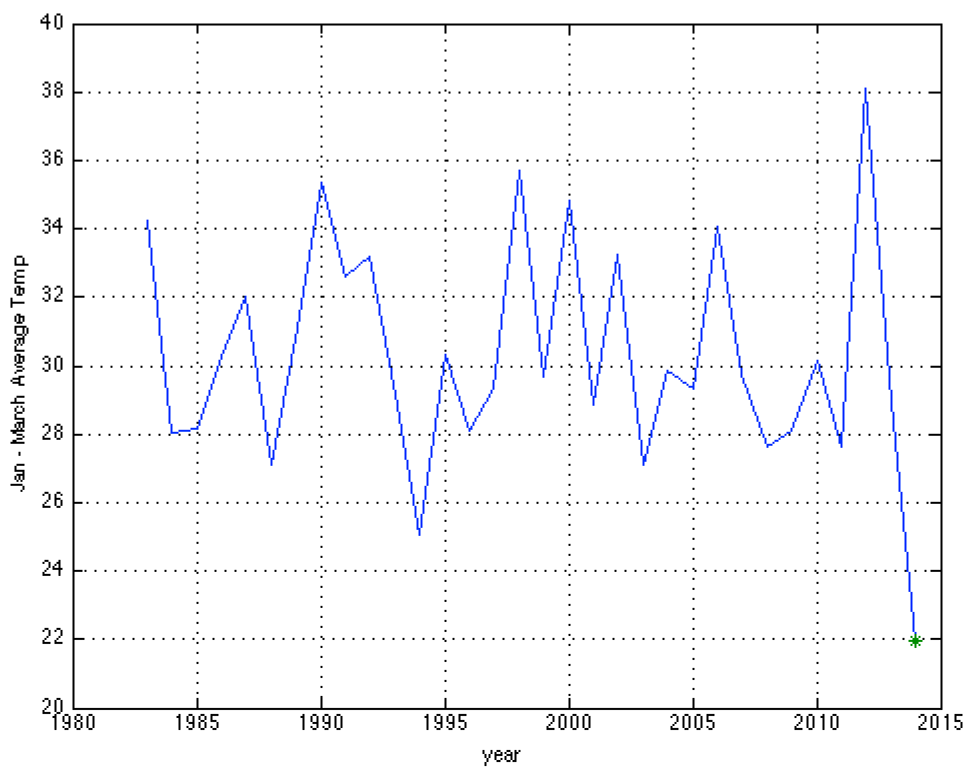
We also want to see how the winter averages stack up:

```
winters=mean(mntemps(1:3,:)); %This is all the winters including 2014
%We define a vector of years:
years=[1983:2014];
figure(2)
plot(years,winters,2014,winters(end),'*')
xlabel('year')
ylabel('Jan - March Average Temp')
grid on

echo on
```

```
%As you can see from this plot, the 2014 winter was way below any other
%winter for the last 30 years!  With an average temperature of 22 degrees,
%it was 3 degrees colder than 1994 and more than 8 degrees below the
%average winter.

%This is clearer if we set up the data in tabular form.  Tables are a bit
%of a pain in Matlab (particularly if you don't have the function
%"table.m" in your version), but they can be constructed using the num2str
%and str2mat commands.
echo off

%We want to have 4 columns: January, February, March, and Winter.  We want
%to have 6 rows of data.  Let's construct this matrix:
[cold icold]=min(tempold(1:3,:)');
[coldwin icoldwin]=min(winters(1:31)); %We exclude the 2014 data
row1=[cold,coldwin];
row2=[icold,icoldwin]+1982;

[hot ihot]=max(tempold(1:3,:)');
[hotwin ihotwin]=max(winters(1:31));
row3=[hot,hotwin];
row4=[ihot,ihotwin]+1982;

row5=round([avgtemp(1:3)',mean(avgtemp(1:3))]*100)/100; %We keep to .01

row6=[temp2014(1:3)',mean(temp2014(1:3))];

tabledata=num2str([row1;row2;row3;row4;row5;row6]);

%Now we type in the left side:
left=str2mat('Coldest Temp   ','Coldest Year   ','Warmest Temp   ', ...
    'Warmest Year   ','Average   ','2014 Data   ');
tabledata=[left,tabledata];

%And finally:
top=['             January       February          March        Average'];
tabledata=str2mat(top,tabledata);
%
%
disp(tabledata)
echo on
%Looking at this table, you can see that while none of the winter months
%were the coldest over the 31 preceding years, all of them were very close
%to the minimum and the three month average was really low: 3 degrees
%colder than the previous coldest winter over that period!
echo off
```

```
%As you can see from this plot, the 2014 winter was way below any other
%winter for the last 30 years!  With an average temperature of 22 degrees,
%it was 3 degrees colder than 1994 and more than 8 degrees below the
%average winter.
%This is clearer if we set up the data in tabular form.  Tables are a bit
%of a pain in Matlab (particularly if you don't have the function
%"table.m" in your version), but they can be constructed using the num2str
%and str2mat commands.
```

```
echo off
                January        February          March         Average
Coldest Temp   15.62          17.78          29.84          25.04
Coldest Year    2009           2007           1984           1994
Warmest Temp   35.78          37.04          52.7           38.12
Warmest Year    2006           1998           2012           2012
Average        25.27          28.25          38.09          30.53
2014 Data      17.24           18.5           30.2          21.98
%Looking at this table, you can see that while none of the winter months
%were the coldest over the 31 preceding years, all of them were very close
%to the minimum and the three month average was really low: 3 degrees
%colder than the previous coldest winter over that period!
echo off
```



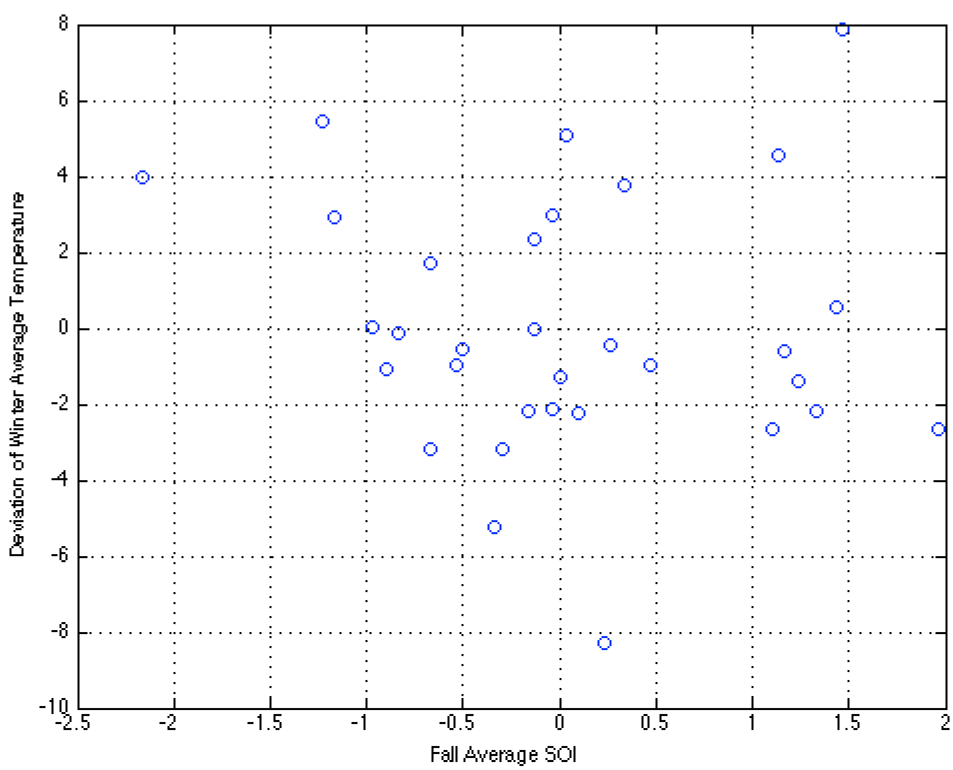
Part 6

Now we want to look at the SOI and see if there is any relationship between that and how cold the winters are.

```
%First we determine the average SOI for each fall:
soimeans=mean(soi(10:12,:))'; %We make this a column vector

%and the winter deviations:
winterdeviation=winters'-mean(winters); %We make this a column vector too.

%Now we plot it up:
figure(3)
plot(soimeans,winterdeviation,'o')
```

```
grid on
xlabel('Fall Average SOI')
ylabel('Deviation of Winter Average Temperature')

echo on
%As you can see from the plot there is a lot of scatter!  The weather
%newscasters sometimes say that an el Nino (negative SOI) leads to a warmer
%winter around here.  While there may be a negative correlation (slope), it
%is obviously very weak and much smaller than the scatter in the
%temperature data.  In particular, the warmest winter had a large SOI the
%previous fall, and the cold winter of 2014 had an SOI in fall of '13 that
%was close to zero.  Thus, it isn't a very good predictor!
echo off
```


```
%As you can see from the plot there is a lot of scatter!  The weather
%newscasters sometimes say that an el Nino (negative SOI) leads to a warmer
%winter around here.  While there may be a negative correlation (slope), it
%is obviously very weak and much smaller than the scatter in the
%temperature data.  In particular, the warmest winter had a large SOI the
%previous fall, and the cold winter of 2014 had an SOI in fall of '13 that
%was close to zero.  Thus, it isn't a very good predictor!
echo off
```

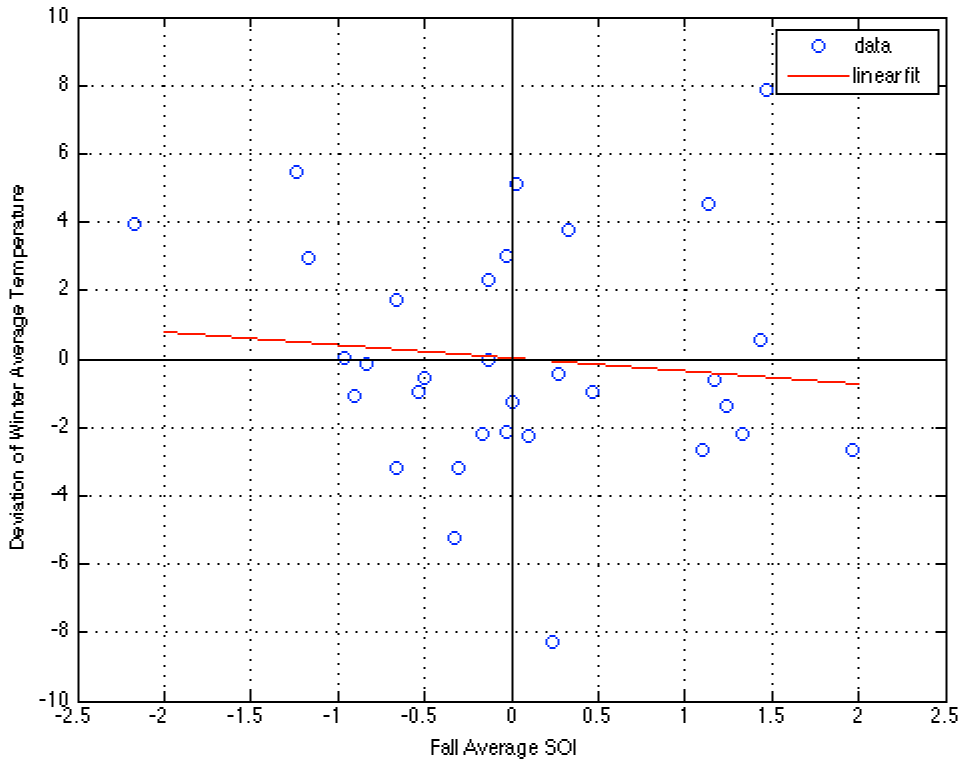
Part 7

We can do some statistical analysis to get a quantitative measure of the SOI – winter temperature relationship. Let's try to fit a line to the data:

```
a=[soimeans,ones(size(soimeans))]; %This is the matrix of modeling functions
b=winterdeviation; %These are the data points

k=inv(a'*a)*a'; %This is the matrix form of the linear regression problem
x=k*b

%The slope (x(1)) is negative, so there is some negative SOI - winter
%correlation.  Let's plot this up:
figure(4)
plot(soimeans,winterdeviation,'o')
xlabel('Fall Average SOI')
ylabel('Deviation of Winter Average Temperature')
hold on
plot([-2,2],[-2 1;2 1]*x,'r',[0,0],[-10 10],'k',[-2.5,2.5],[0,0],'k')
grid on
hold off
legend('data','linear fit')

echo on
%Now we want to determine if it is statistically significant.  First we get
%the residual (the scatter in the temperature data):
r=a*x-b;
sigb=((r'*r)/(length(b)-2))^.5

%So the winter temperature bounces around with standard deviation of 3.4
%degrees.  Now we want to see how this propagates into uncertainty in the
%regression coefficients:

varx=k*k'*sigb^2;

%In particular, we are interested in the standard deviation in the slope:
sigx1=varx(1,1)^.5

%Which is actually twice the magnitude of the slope!  This means that there
%is no statistical significance to the correlation: not a big surprise.

%Finally, we can calculate the probability that the SOI correlation is
%negative using the t-test cumulative distribution function:
probneg=tcdf(abs(x(1))/sigx1,(length(soimeans)-2))

%This means that while there is a 72% chance that there is the (expected)
%negative correlation, there is a 28% chance that the correlation was
%actually positive.  In practice, one would reject the hypothesis at the
%72% confidence level, far short the normal 95% confidence level.  Also,
%even if there is a correlation, the magnitude of the effect (based on the
%best fit line) is peanuts compared to the normal variability in winter
%temperatures - you'd never be able to feel it.

%So does the SOI matter?  From the data, it doesn't seem to have any
%relation to the winter temperature for this area.  It might affect
%precipitation, but that would require another study...
echo off
```

```
x =
```

```
   -0.3835
    0.0180

%Now we want to determine if it is statistically significant.  First we get
%the residual (the scatter in the temperature data):
r=a*x-b;
sigb=((r'*r)/(length(b)-2))^.5

sigb =

    3.3999


%So the winter temperature bounces around with standard deviation of 3.4
%degrees.  Now we want to see how this propagates into uncertainty in the
%regression coefficients:
varx=k*k'*sigb^2;

%In particular, we are interested in the standard deviation in the slope:
sigx1=varx(1,1)^.5

sigx1 =

    0.6504


%Which is actually twice the magnitude of the slope!  This means that there
%is no statistical significance to the correlation: not a big surprise.
%Finally, we can calculate the probability that the SOI correlation is
%negative using the t-test cumulative distribution function:
probneg=tcdf(abs(x(1))/sigx1,(length(soimeans)-2))

probneg =

    0.7201


%This means that while there is a 72% chance that there is the (expected)
%negative correlation, there is a 28% chance that the correlation was
%actually positive.  In practice, one would reject the hypothesis at the
%72% confidence level, far short the normal 95% confidence level.  Also,
%even if there is a correlation, the magnitude of the effect (based on the
%best fit line) is peanuts compared to the normal variability in winter
%temperatures - you'd never be able to feel it.
%So does the SOI matter?  From the data, it doesn't seem to have any
%relation to the winter temperature for this area.  It might affect
%precipitation, but that would require another study...
echo off
```



Published with MATLAB® 7.12