Table of Contents

Lab 2: CAPM modelling and analysis

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%    IMPORT DATA SETS FIRST    %%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

% Import "FF_Research_Data_Factors.txt" as a NUMERIC MATRIX and name
 'FFResearchDataFactors'
% Import "5_IndustryPortfolios.txt" as a NUMERIC MATRIX and name
 'IndustryPortfolios'
% Import only the MONTHLY from July, 1926 to December, 2011 from both
 files
% From "5_IndustryPortfolios.txt" only import the "Value-weighted"
 data at
% the top of the file.

% You may need to play around a see whether 'fixed width' or
 'Delimited'
% method works best with each file. With delimited you will need to
 choose
% the correct delimiter as space, comma, tab etc
% Also check the end of the txt file in case there is additional text
 which
% you do not want to include

% Now shorten data matrix names for ease of programming,
% and select only relevant rows (1:1026)
indmat = IndustryPortfolios(1:1026,:);
FFmat = FFResearchDataFactors(1:1026,:);

% remove old matrices (to save space)
% clear IndustryPortfolios FFResearchDataFactors

% Alternatively, simply copy and paste the monthly data into
% Matlab into the variable names 'FFmat' and 'indmat', i.e.

% FFmat = [ <copy and paste data from FF_Research_Data_Factors.txt
 here> ];
% indmat=[ <copy and paste data from 5_Industry_Portfolios.txt
 here> ];
```

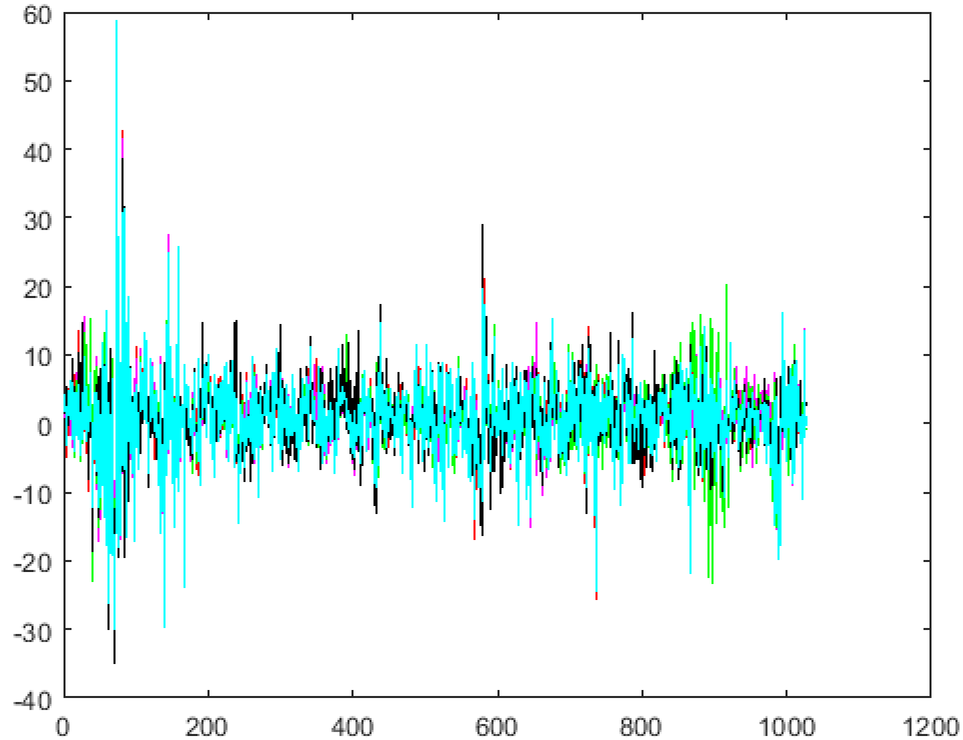
```
% NOTE: I used the VALUE weighted data here

%%%%%%%%%%%% SAVING DATA AS MATLAB WORKSPACE OBJECT %%%%%%%%%%%%%%

% You could now save as a matlab data object 'name.mat' using
% Home --> Save Workspace as -->  . I saved it as "lab2.mat"
% Alternatively you could type the command: "save lab2.mat"
%
% Saving will easily allow you to load the data later without using
 the Import Data tool
% or the copy/paste method
% To load the data again use Home --> Open -->
% Alternatively you could type the command: "load lab2.mat"
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%% BEGIN PROGRAMME  %%%%%%%%%%%%%%%%%%%%%%%%
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
```

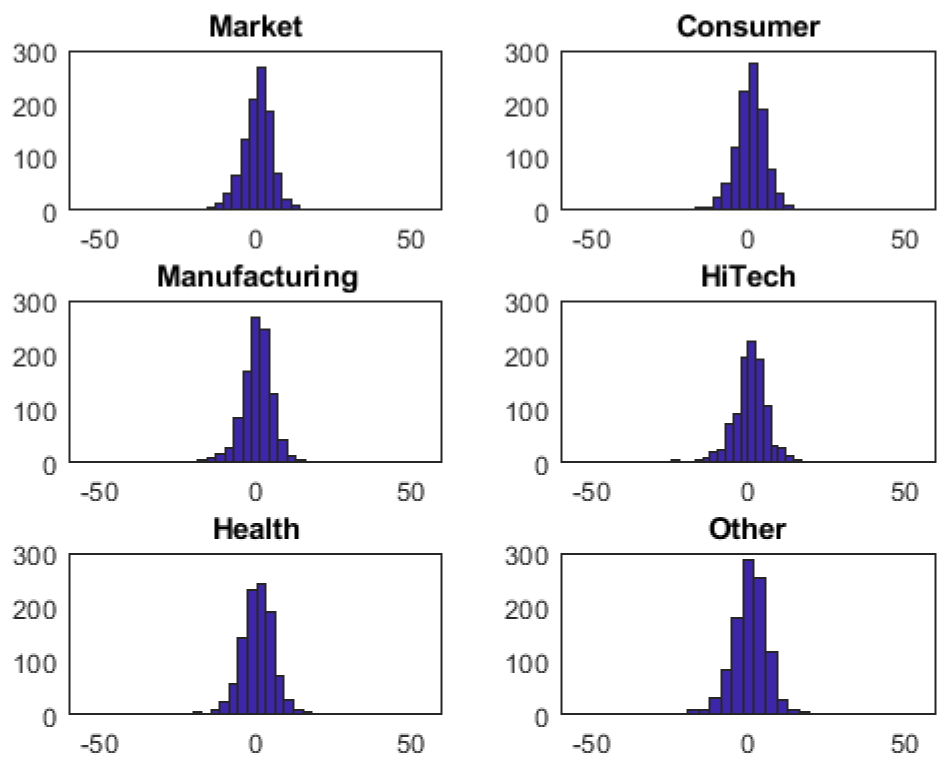
(a) Calculate excess returns, plot them, and provide summary stats

load lab2.mat

```
rf=FFmat(:,5);   % risk-free rate  无风险利率
ex_mark_ret=FFmat(:,2);  % excess market return  超额市场回报
cnsmr_ret=indmat(:,2);  % return on consumer portfolio
cnsmr_ex_ret=cnsmr_ret-rf;  % excess return  超额收益
manuf_ret=indmat(:,3);manuf_ex_ret=manuf_ret-rf;
hitech_ret=indmat(:,4);hitech_ex_ret=hitech_ret-rf;
health_ret=indmat(:,5);health_ex_ret=health_ret-rf;
other_ret=indmat(:,6);other_ex_ret=other_ret-rf;

% plots of industry excess returns against market excess returns
figure;plot(ex_mark_ret);hold on;plot(cnsmr_ex_ret,'r')
plot(manuf_ex_ret,'m');plot(hitech_ex_ret,'g')
plot(health_ex_ret,'k');plot(other_ex_ret,'c')

% summary statistics table
summ = [mean(ex_mark_ret) median(ex_mark_ret) std(ex_mark_ret)
 min(ex_mark_ret) max(ex_mark_ret) skewness(ex_mark_ret)
 kurtosis(ex_mark_ret);
mean(cnsmr_ex_ret) median(cnsmr_ex_ret) std(cnsmr_ex_ret)
 min(cnsmr_ex_ret) max(cnsmr_ex_ret) skewness(cnsmr_ex_ret)
 kurtosis(cnsmr_ex_ret);
mean(manuf_ex_ret) median(manuf_ex_ret) std(manuf_ex_ret)
 min(manuf_ex_ret) max(manuf_ex_ret) skewness(manuf_ex_ret)
 kurtosis(manuf_ex_ret);
mean(hitech_ex_ret) median(hitech_ex_ret) std(hitech_ex_ret)
 min(hitech_ex_ret) max(hitech_ex_ret) skewness(hitech_ex_ret)
 kurtosis(hitech_ex_ret);
```

```
mean(health_ex_ret) median(health_ex_ret) std(health_ex_ret)
 min(health_ex_ret) max(health_ex_ret) skewness(health_ex_ret)
 kurtosis(health_ex_ret);
mean(other_ex_ret) median(other_ex_ret) std(other_ex_ret)
 min(other_ex_ret) max(other_ex_ret) skewness(other_ex_ret)
 kurtosis(other_ex_ret)]

% you could also use histograms or boxplots here e.g.:
figure;axis([-60 60 0 300]);subplot(3,2,1);axis([-60 60 0
 300]);hist(ex_mark_ret,25);title('Market');axis([-60 60 0 300]);
subplot(3,2,2);hist(cnsmr_ex_ret,25);title('Consumer');axis([-60 60 0
 300]);
subplot(3,2,3);hist(manuf_ex_ret,25);;title('Manufacturing');axis([-60
 60 0 300]);
subplot(3,2,4);hist(hitech_ex_ret, 25);title('HiTech');axis([-60 60 0
 300]);
subplot(3,2,5);hist(health_ex_ret, 25);title('Health');axis([-60 60 0
 300]);
subplot(3,2,6);hist(other_ex_ret, 25);title('Other');axis([-60 60 0
 300]);
% note: I put all axes on the same scale for ease of comparison


summm =

    0.6175    0.9550    5.4572  -29.0400   38.2700    0.1685   10.3983
    0.6940    0.9100    5.3856  -28.2000   42.5900    0.1153   10.1566
    0.6820    0.9100    5.5966  -29.8400   41.5300    0.3579   11.0530
    0.6257    0.8900    5.6908  -26.7700   33.8000   -0.1695    6.5058
    0.7735    0.7350    5.7350  -34.8000   38.5600    0.1798   10.1047
    0.5860    0.9350    6.5587  -30.0500   58.7100    0.8929   15.8557
```

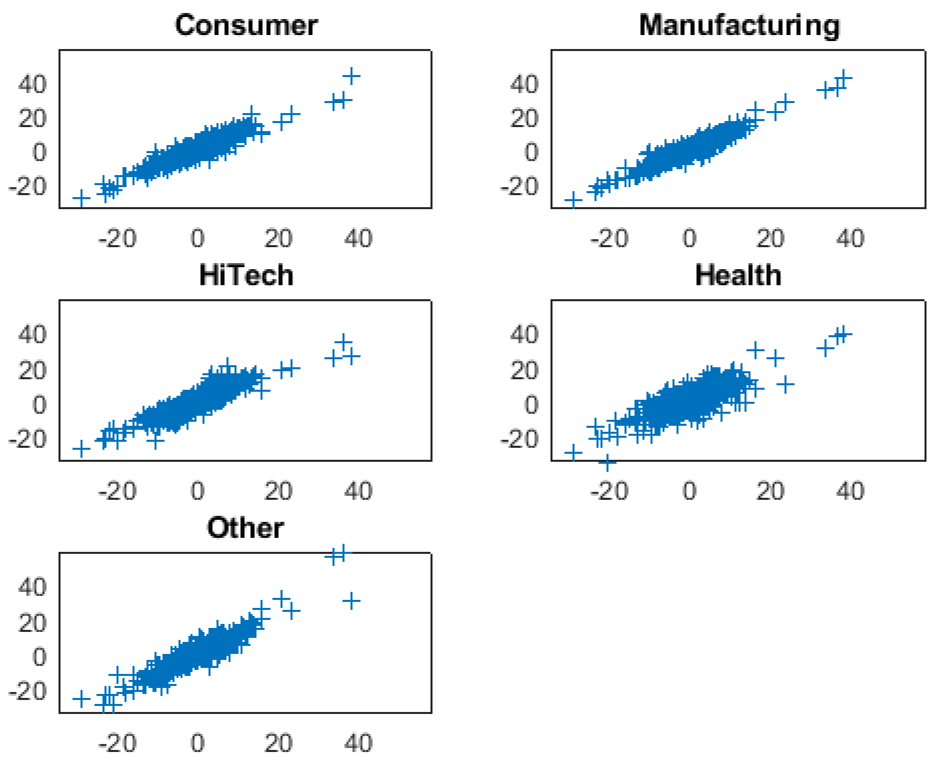

60
50
40
30
20
10
0
-10
-20
-30
-40
0
200
400
600
800
1000
1200


Market
Consumer
Manufacturing
HiTech
Health
Other
300
200
100
0
-50
0
50

(b) Scatterplots of industry excess returns against market excess returns

find min and max values for xy-axes, note that we have already calculated the min and max ex returns

```
xymin = min(summ(:,4));
xymax = max(summ(:,5));
% scatterplots of industry excess returns against market excess
 returns
figure;subplot(3,2,1);plot(ex_mark_ret,cnsmr_ex_ret,'+');axis([xymin
 xymax xymin xymax]);title('Consumer');
subplot(3,2,2);plot(ex_mark_ret,manuf_ex_ret,'+');axis([xymin xymax
 xymin xymax]);title('Manufacturing');
subplot(3,2,3);plot(ex_mark_ret,hitech_ex_ret,'+');axis([xymin xymax
 xymin xymax]);title('HiTech');
subplot(3,2,4);plot(ex_mark_ret,health_ex_ret,'+');axis([xymin xymax
 xymin xymax]);title('Health');
subplot(3,2,5);plot(ex_mark_ret,other_ex_ret,'+');axis([xymin xymax
 xymin xymax]);title('Other');
```



(c) Calculate correlations of industry excess returns against market excess returns and test

[r,p]=corrcoef(x,y) command variables 'x' and 'y' and calculates the correlation 'r' between them and also reports 'p' which is the p-value from test that rho=0

```
[rc,pc]=corrcoef(ex_mark_ret,cnsmr_ex_ret);
[rm,pm]=corrcoef(ex_mark_ret,manuf_ex_ret);
[rhi,phi]=corrcoef(ex_mark_ret,hitech_ex_ret);
[rhe,phe]=corrcoef(ex_mark_ret,health_ex_ret);
[ro,po]=corrcoef(ex_mark_ret,other_ex_ret);

[rc(2,1) rm(2,1) rhi(2,1) rhe(2,1) ro(2,1); pc(2,1) pm(2,1) phi(2,1)
 phe(2,1) po(2,1)]
% r and p here are given as 2 by 2 matrices. The (2,1) and (1,2)
 elements
% are the same and are the values we want.


ans =

    0.9352    0.9612    0.9070    0.8036    0.9346
         0         0         0    0.0000         0
```

(d) CAPM Simple Linear regressions

create X matrix for regression

```
xmat=[ones(length(ex_mark_ret),1) ex_mark_ret];

%fit regression models
% use 'doc regress; to get help on this command
% For first regression below, Bc contains slope coefficients, BINTc
 has 95% confidence intervals Rc has the residuals,
% RINTc contains 95% confidence intervals for the standardised
 residuals if these intervals do not contain zero then data point
 might be an outlier
% STATSc contains in order the following stats: R^2 statistic, the F
 statistic and its p value, and an estimate of the error variance
% In my labels 'c' refers to the consumer sector, 'm' for manufacuring
 and so on
[Bc,BINTc,Rc,RINTc,STATSc] = regress(cnsmr_ex_ret,xmat);   % runs OLS
 regression
cnsmr_est=xmat*Bc;                                          % estimates
 of average return from regression
[Bm,BINTm,Rm,RINTm,STATSm] = regress(manuf_ex_ret,xmat);   % runs OLS
 regression
manuf_est=xmat*Bm;
[Bhi,BINThi,Rhi,RINThi,STATShi] = regress(hitech_ex_ret,xmat);   %
 runs OLS regression
```

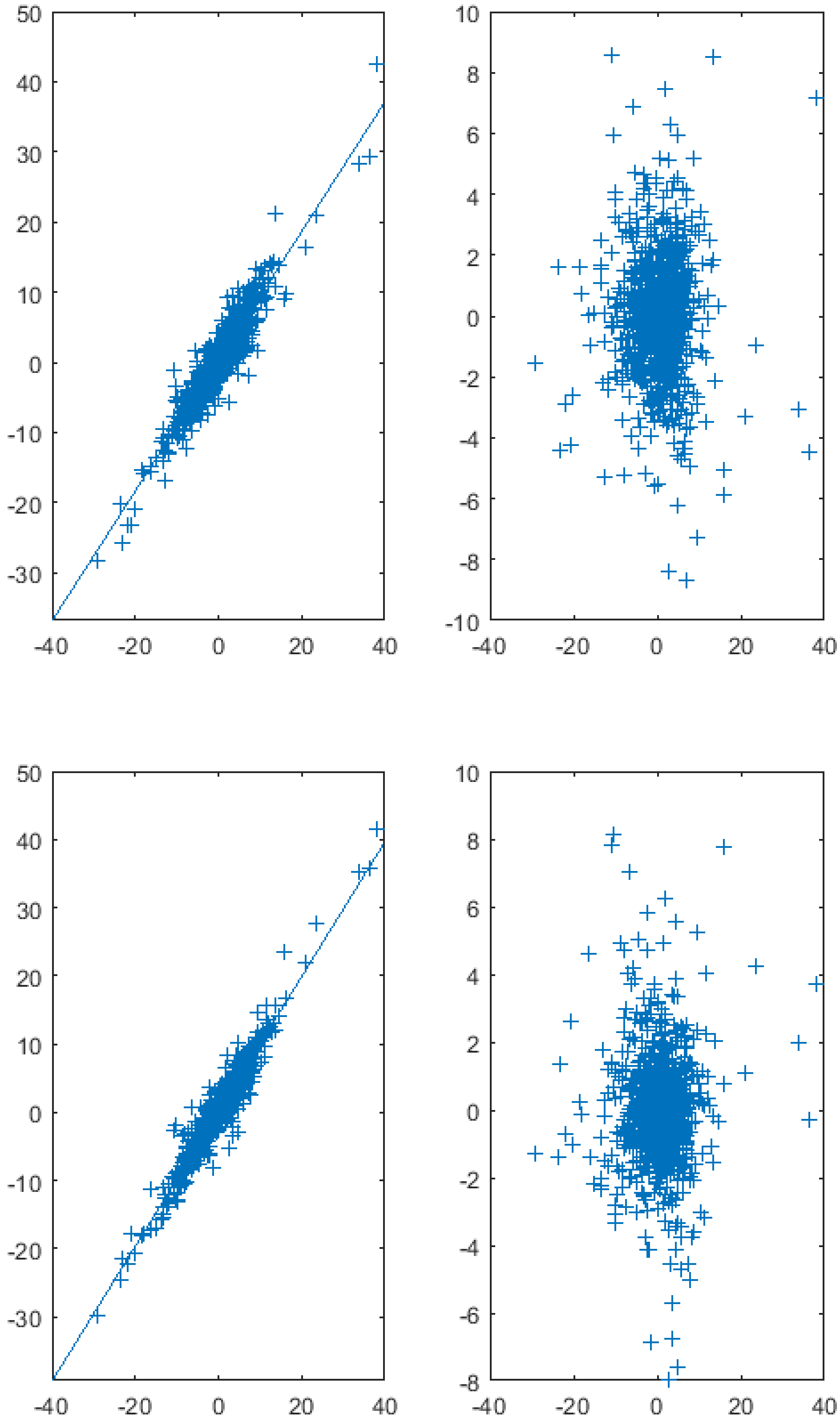
```
hitech_est=xmat*Bhi;
[Bhe,BINThe,Rhe,RINThe,STATShe] = regress(health_ex_ret,xmat);   %
 runs OLS regression
health_est=xmat*Bhe;
[Bo,BINTo,Ro,RINTo,STATSo] = regress(other_ex_ret,xmat);   % runs OLS
 regression
other_est=xmat*Bo;

% make a table containing intercept, followed by intercept's 95% CI,
 followed by beta, followed by beta's 95% CI for each sector
[Bc(1) BINTc(1,:) Bc(2) BINTc(2,:);Bm(1) BINTm(1,:) Bm(2)
 BINTm(2,:);Bhi(1) BINThi(1,:) Bhi(2) BINThi(2,:);
    Bhe(1) BINThe(1,:) Bhe(2) BINThe(2,:);Bo(1) BINTo(1,:) Bo(2)
 BINTo(2,:);]


ans =

    0.1241    0.0065    0.2418    0.9229    0.9015    0.9443
    0.0733   -0.0219    0.1685    0.9858    0.9684    1.0031
    0.0417   -0.1061    0.1895    0.9458    0.9189    0.9728
    0.2520    0.0415    0.4625    0.8445    0.8062    0.8829
   -0.1076   -0.2515    0.0363    1.1232    1.0970    1.1495
```

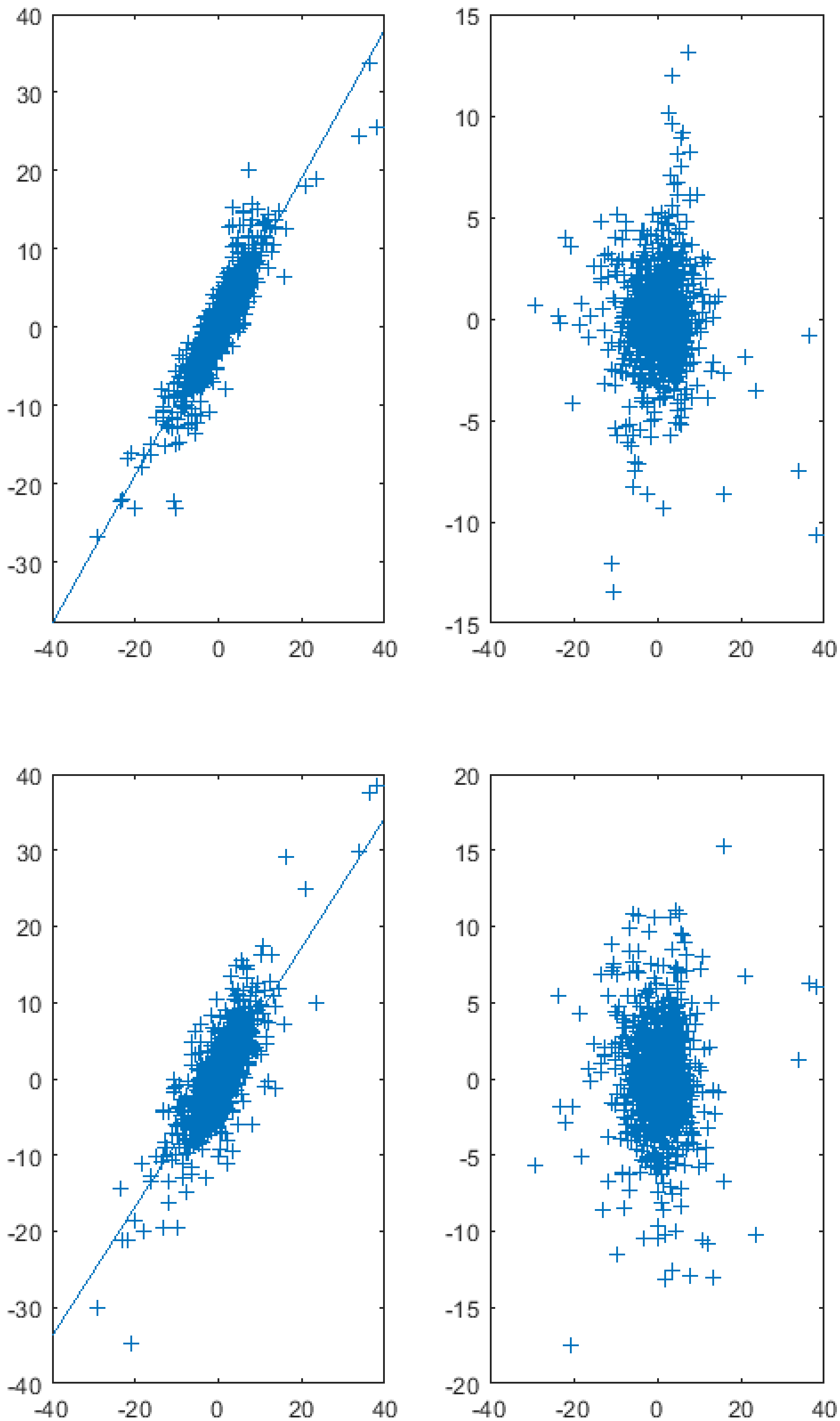
离群值(outlier)，也称逸出值，是指在数据中有一个或几个数值与其他数值相比差异较大。
chanwennt准则规定，如果一个数值偏离观测平均值的概率小于等于1/(2n)，则该数据应当舍弃

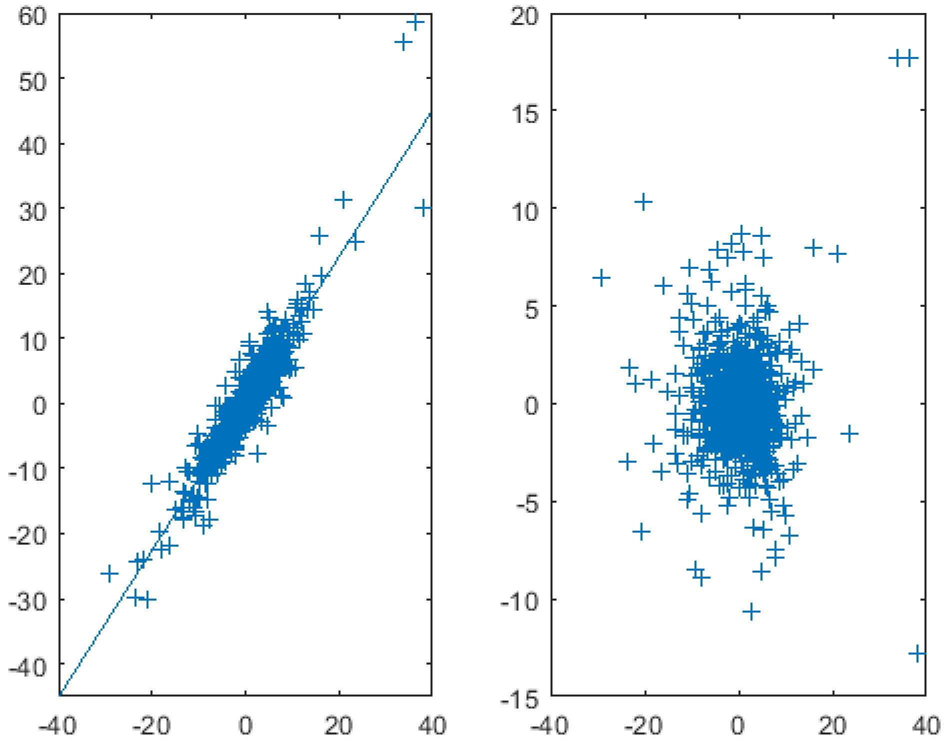
(e) Check for outliers using plots

plot of data and line and residuals, assumption 3

```
figure;subplot(1,2,1);plot(ex_mark_ret,cnsmr_ex_ret,'+');lsline
subplot(1,2,2);plot(ex_mark_ret,cnsmr_ex_ret-cnsmr_est,'+')

figure;subplot(1,2,1);plot(ex_mark_ret,manuf_ex_ret,'+');lsline
subplot(1,2,2);plot(ex_mark_ret,manuf_ex_ret-manuf_est,'+')

figure;subplot(1,2,1);plot(ex_mark_ret,hitech_ex_ret,'+');lsline
subplot(1,2,2);plot(ex_mark_ret,hitech_ex_ret-hitech_est,'+')

figure;subplot(1,2,1);plot(ex_mark_ret,health_ex_ret,'+');lsline
subplot(1,2,2);plot(ex_mark_ret,health_ex_ret-health_est,'+')

figure;subplot(1,2,1);plot(ex_mark_ret,other_ex_ret,'+');lsline
subplot(1,2,2);plot(ex_mark_ret,other_ex_ret-other_est,'+')
```


50
40
30
20
10
0
-10
-20
-30
-40
-20
0
20
40
10
8
6
4
2
0
-2
-4
-6
-8
-10
-40
-20
0
20
40
50
40
30
20
10
0
-10
-20
-30
-40
-20
0
20
40
10
8
6
4
2
0
-2
-4
-6
-8
-40
-20
0
20
40


40
30
20
10
0
-10
-20
-30
-40
-20
0
20
40
15
10
5
0
-5
-10
-15
-40
-20
0
20
40
40
30
20
10
0
-10
-20
-30
-40
-40
-20
0
20
40
20
15
10
5
0
-5
-10
-15
-20
-40
-20
0
20
40



(f) Strength of fit

R^2 and SER per sector regression

```
[STATSc(1) sqrt(STATSc(4));STATSm(1) sqrt(STATSm(4));STATShi(1)
 sqrt(STATShi(4));
    STATShe(1) sqrt(STATShe(4));STATSo(1) sqrt(STATSo(4))]
```

```
ans =

    0.8745    1.9087
    0.9239    1.5443
    0.8227    2.3976
    0.6458    3.4148
    0.8735    2.3340
```

(g) Assess industry risk (High/Medium/Low)

The command regstats(y,x) will provide regression statistics from a regression of y against x. A dialog box will allow you to choose which stats to record

```
regstats(other_ex_ret,ex_mark_ret)
% choose 'Coefficients' (saved as 'beta'), 'coefficient
 covariances' (saved as 'covb')
```

```
% and any others you feel like

% t-test of beta>1 for Other
ts = (beta(2)-1)/sqrt(covb(2,2))
pval = 1-normcdf(ts)
```

```
Variables have been created in the base workspace.

ts =

    9.2253


pval =

     0
```

Published with MATLAB® R2017b