Contents


```
% Endsem CH5440
% Ojas Phadake - CH22B007

clc;
clear all;
close all;

load steamdata.mat
```

A

```
nsamples = size(Ftrue, 1);
nvar = size(Ftrue, 2);

Z = Ftrue;
Zs = Z - mean(Z);
Sz = Zs'*Zs/nsamples;

[u,s,v] = svd(Zs/sqrt(nsamples), "econ");
D = s.*s;

fprintf("Using the values which I obtain in D, there is a noticeable difference between the values, as " + ...
    "it goes from 10^-28 to 10^0 which is a very big difference. \nHence, as there are 8 eigenvalues which " + ...
    "are equal, there are 8 constraints.\n")
fprintf("**********************************************************************************************")
```

```
Using the values which I obtain in D, there is a noticeable difference between the values, as it goes from 10^-28 to 10^0 which is a very big difference.
Hence, as there are 8 eigenvalues which are equal, there are 8 constraints.
**********************************************************************************************
```

B

```
dest = 8;

nfact = nvar - dest;
for i = nfact+1:nvar
    Amat(i-nfact,:) = v(:, i)';% This matrix contains values of constraints which are in total 8
end

% Say the 1st 17 are selected to be independent and the last 8 as dependent
% variables

Aesti = Amat(:, 8:10);
Aesti = [Aesti Amat(:, 12:25)];
Aestd = Amat(:, 1:7);
Aestd = [Aestd Amat(:, 11)];

cond(Aestd);
fprintf("The minimized condition number is: %0.4f", cond(Aestd));

Restb = -inv(Aestd)*Aesti;

fprintf("\nWe have selected a good set of 17 variables as independent and the other 8 as dependent.\n")
```

```
The minimized condition number is: 3.8413
We have selected a good set of 17 variables as independent and the other 8 as dependent.
```

C

```
% Finding out the best set of independent variables so that the condition
% number is minimized

[u1, s1, v1] = svd((Fmeas - mean(Fmeas))/sqrt(nsamples), "econ");
D1= s1.*s1;
```

```
for i = nfact+1:nvar
    Amat1(i-nfact,:) = v1(:, i)';% This matrix contains values of constraints which are in total 8
end

index = 1:1:25;
set_independent = nchoosek(index, 8);
cond_num = 10^5; % Intialized as a high value

for i=1:size(set_independent, 1)
    Aestd1 = Amat1(:, set_independent(i, :));
    if cond(Aestd) < cond_num
        min = i;
        cond_num = cond(Aestd1);
    end
    if cond_num < 5 % Set an arbitrary limit for condition number
        break;
    end
end

index(set_independent(min, :)) = [];
Aesti1 = Amat1(:, index);
Restc = -inv(Aestd1)*Aesti1;

fprintf("The true regression matrix is given by Restc: ");
disp(Restc);


% Hypothesis testing begins
lambda = diag(D1);
alpha = 0.05;

Test_statistic = zeros(24,1);
Test_criterion = zeros(24,1);
Degrees_of_freedom = zeros(24,1);
k_value = (2:1:25)';
n = 25;

for d = (n-1):-1:2
   Degrees_of_freedom(d-1) = (d-1)*(d+2)/2;
   lbar = mean(lambda(n-d+1:end));
   Test_statistic(d-1) = (nvar-1)*(d*log(lbar)-sum(log(lambda(n-d+1:end))));
   Test_criterion(d-1) = chi2inv(1-alpha,Degrees_of_freedom(d-1));
end
flag = 1;
d = n-2;
dest = 1;
while flag
    if ( Test_statistic(d) > Test_criterion(d) )
        d = d-1;
        if ( d < 2)
           flag = 0;
        end
    else
        dest = d + 1;
        flag = 0;
    end
end

disp(table(k_value,Degrees_of_freedom,Test_statistic,Test_criterion));
nfact = nvar - dest;

fprintf("We get dest as 9 which is the true number of constraints by implementing PCA\n");
disp(dest);

L = diag(D1);
err_var = mean(L(end - d + 1:end));

fprintf("The mean error variances are given by: %0.4f\n", err_var)
```

```
The true regression matrix is given by Restc:   Columns 1 through 7

   -0.0091    0.0103   -1.0477    0.0318   -0.0135    1.0486    1.0092
   -1.0092   -0.0014   -1.0847    0.0160   -0.0699   -0.0134    1.0057
   -0.0207    1.0253    0.0637    0.0110    0.0589    0.0115   -0.0479
    0.0004   -0.0037   -1.0463   -0.9511   -0.0577    1.0147    1.0068
   -0.0065    0.9859    0.0191   -0.0005    0.0449   -0.0447    0.0037
   -0.0129    0.0073    0.0392    1.0000    0.0824   -0.0190   -0.0181
    0.0316   -1.0315    0.0019    0.0089    0.9633   -0.0908    0.0752
    0.0078    0.0192   -1.0338   -0.9630   -0.0786    1.0194    1.0008
```

```
  Columns 8 through 14

  -0.9386    0.0274   -0.9952    1.0202    0.9794   -0.0554   -0.9999
   0.0183   -1.0304   -0.0334    0.0502    0.0653   -0.0764   -0.0333
   1.0103    0.0072    0.0402   -0.0165   -0.0025   -0.9753    0.0059
  -0.9419    0.0267   -1.0113    0.0511    1.0108   -0.0734   -1.0195
   0.9770   -0.0333    0.0187   -0.0227   -0.0077    0.0287   -0.0110
   1.0376   -0.0212    0.0146    0.0107   -0.0020   -0.0074    0.0120
  -0.9675   -0.0128    0.0117    0.0047   -0.0491    0.9973   -0.0026
  -0.9594    0.0307   -1.0157    0.0423    1.0032   -0.0678   -0.0309

  Columns 15 through 17

   0.9674    0.9784    0.0118
   0.9850   -0.0555    0.0150
  -0.0135    0.0129    1.0131
   0.9241    0.9760   -0.0032
   0.0036   -0.0115   -0.0043
   0.0057   -0.0100    0.0521
   0.9698    0.9520   -1.0084
   0.9303    0.9714    0.0107

   k_value    Degrees_of_freedom    Test_statistic    Test_criterion
   _______    __________________    ______________    ______________

      2                2              0.0033983           5.9915
      3                5               0.087587           11.07
      4                9                0.20527           16.919
      5               14                0.33439           23.685
      6               20                 0.4617            31.41
      7               27                 0.6407           40.113
      8               35                 1.0052           49.802
      9               44                 39.988           60.481
     10               54                 76.365           72.153
     11               65                 144.33           84.821
     12               77                 186.84           98.484
     13               90                 214.73           113.15
     14              104                 235.65            128.8
     15              119                 253.71           145.46
     16              135                  268.8           163.12
     17              152                 287.78           181.77
     18              170                 308.49           201.42
     19              189                 335.16           222.08
     20              209                 362.58           243.73
     21              230                 401.98           266.38
     22              252                 449.83           290.03
     23              275                 499.25           314.68
     24              299                 601.58           340.33
     25                0                      0                0

We get dest as 9 which is the true number of constraints by implementing PCA
     9

The mean error variances are given by: 1.3226
```

D

```
maxdiff = max(max(abs(Restc - Restb)));
fprintf("The value of maximum difference comes out as: %0.4f\n", maxdiff)
```

```
The value of maximum difference comes out as: 2.0092
```

E

```
theta_pca = subspace(Restb', Restc');
fprintf("The value of theta PCA is: %0.4f", theta_pca)

fprintf("The true regression matrix is given by Restb: ");
disp(Restb);

fprintf("\nAs the value of theta PCA is small, we can say that the solution is correct.")
fprintf("\nAlong with that, even maxdiff is small, which means that in the same way as theta subspace, " + ...
    "we are getting a correct answer. \n ")
```

```
The value of theta PCA is: 1.5501The true regression matrix is given by Restb:   Columns 1 through 7

    1.0000    0.0000    0.0000    1.0000   -0.0000   -0.0000    0.0000
```

```
    1.0000   -1.0000   -0.0000    1.0000    0.0000   -1.0000   -0.0000
    0.0000   -0.0000    1.0000    0.0000   -0.0000   -0.0000   -0.0000
    1.0000   -0.0000    0.0000   -0.0000    0.0000   -0.0000    0.0000
    0.0000    0.0000    1.0000    0.0000   -0.0000    0.0000   -0.0000
    0.0000   -0.0000    0.0000    1.0000   -0.0000   -0.0000   -0.0000
   -0.0000   -0.0000   -1.0000   -0.0000    1.0000    0.0000    0.0000
   -1.0000   -0.0000    0.0000   -1.0000   -0.0000    1.0000    1.0000

  Columns 8 through 14

    0.0000    0.0000   -0.0000    1.0000    0.0000   -0.0000   -1.0000
    1.0000   -1.0000    1.0000    0.0000   -1.0000    0.0000    0.0000
    1.0000    0.0000    0.0000    0.0000   -0.0000   -1.0000   -0.0000
    0.0000   -0.0000    0.0000   -0.0000    0.0000   -0.0000   -1.0000
    1.0000    0.0000    0.0000   -0.0000    0.0000   -0.0000   -0.0000
    1.0000   -0.0000    0.0000    0.0000   -0.0000   -0.0000    0.0000
   -1.0000   -0.0000   -0.0000    0.0000    0.0000    1.0000   -0.0000
   -1.0000   -0.0000   -1.0000   -0.0000    1.0000   -0.0000   -0.0000

  Columns 15 through 17

   -0.0000   -0.0000    0.0000
   -0.0000   -1.0000   -0.0000
   -0.0000   -0.0000    1.0000
   -0.0000    0.0000    0.0000
   -0.0000   -0.0000    0.0000
   -0.0000   -0.0000    0.0000
    1.0000    1.0000   -1.0000
    1.0000    1.0000    0.0000
```

As the value of theta PCA is small, we can say that the solution is correct.
Along with that, even maxdiff is small, which means that in the same way as theta subspace, we are getting a correct answer.

Published with MATLAB® R2023a