Contents

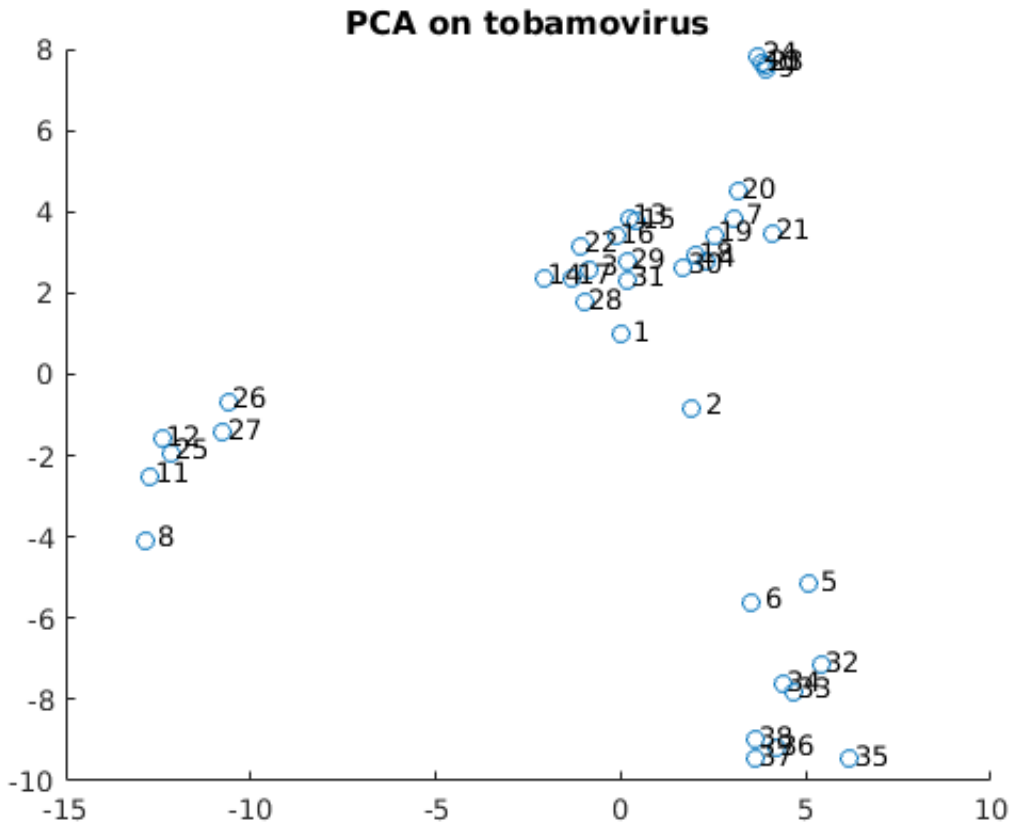
Implementation of PCA . PPCA - EM and general

load the data

```
load('../../Data/virus3.dat');
Y = virus3';
```

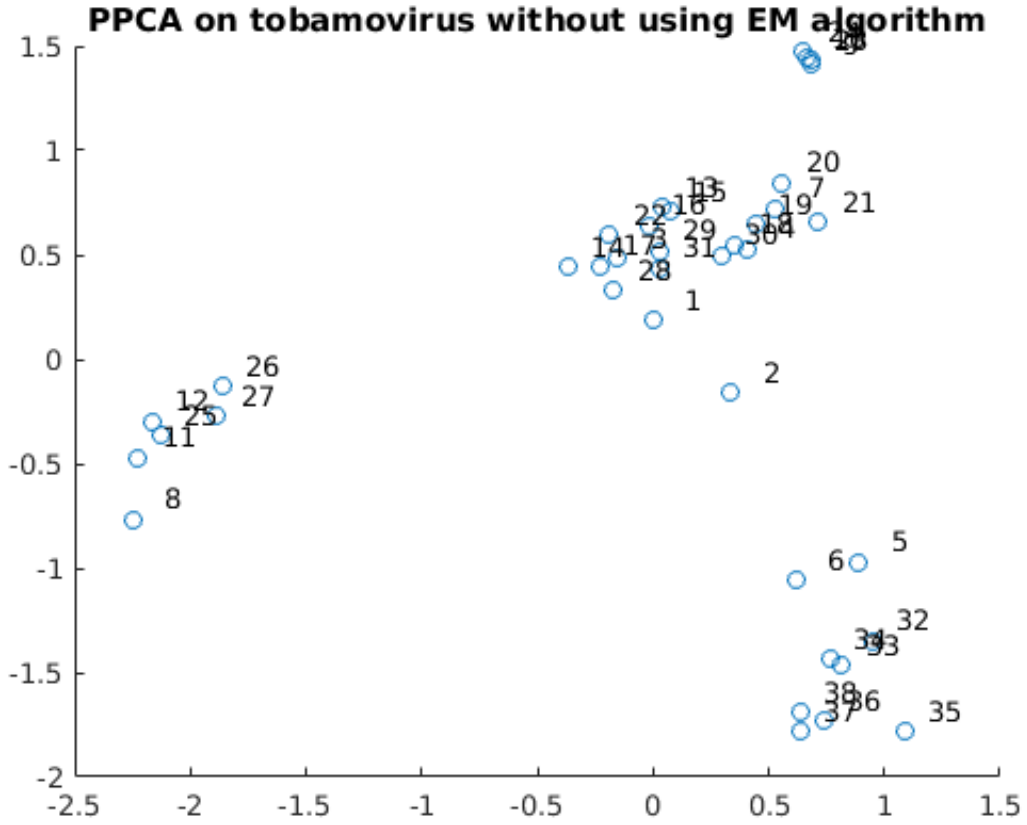
PCA implementation on entire data

```
[W, X] = PCA(Y, 2);
figure;
scatter(X(1, :), X(2, :));
T = cellstr(num2str([1:size(X, 2)]'));
text(X(1, :) + 0.1, X(2, :) + 0.1, T);
title('PCA on tobamovirus');
```



PPCA implementation without EM

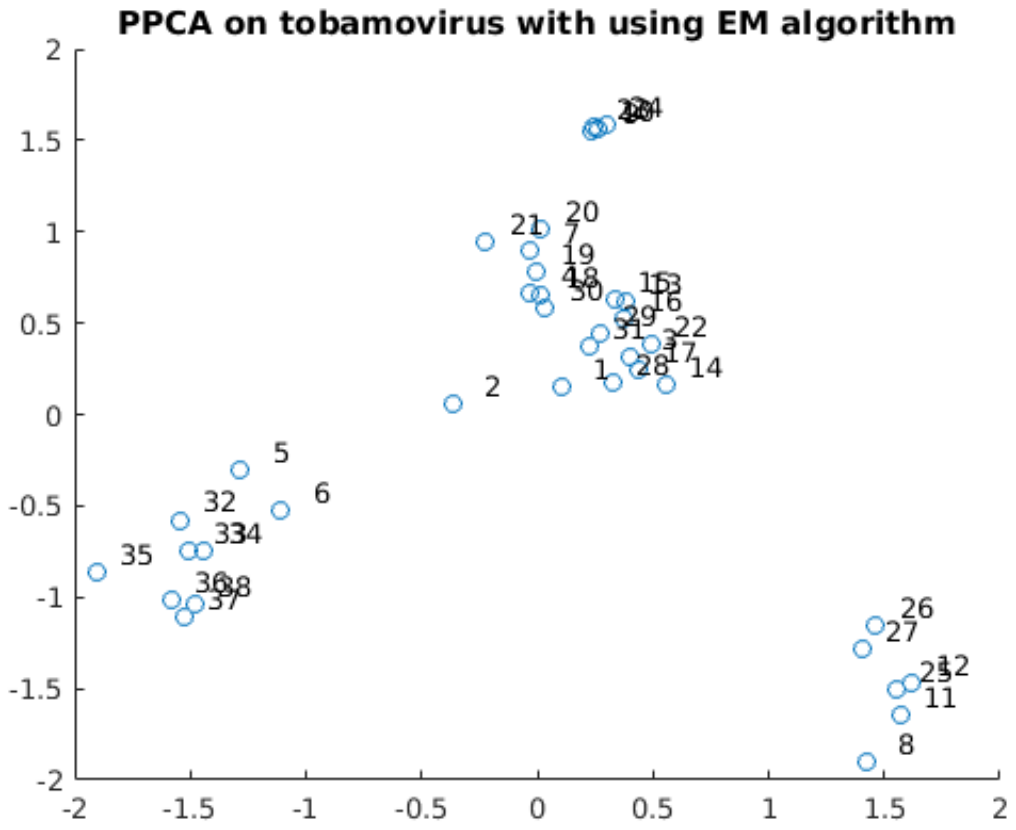
```
[W, var, X] = PPCAWithoutEM(Y, 2);
figure;
scatter(X(1, :), X(2, :));
T = cellstr(num2str([1:size(X, 2)]'));
text(X(1, :) + 0.1, X(2, :) + 0.1, T);
title('PPCA on tobamovirus without using EM algorithm');
```



PPCA with EM

```
[W, var, X] = PPCAWithEM(Y, 2);
figure;
scatter(X(1, :), X(2, :));
T = cellstr(num2str([1:size(X, 2)]'));
text(X(1, :) + 0.1, X(2, :) + 0.1, T);
title('PPCA on tobamovirus with using EM algorithm');
```

```
no of iterations = 87
```

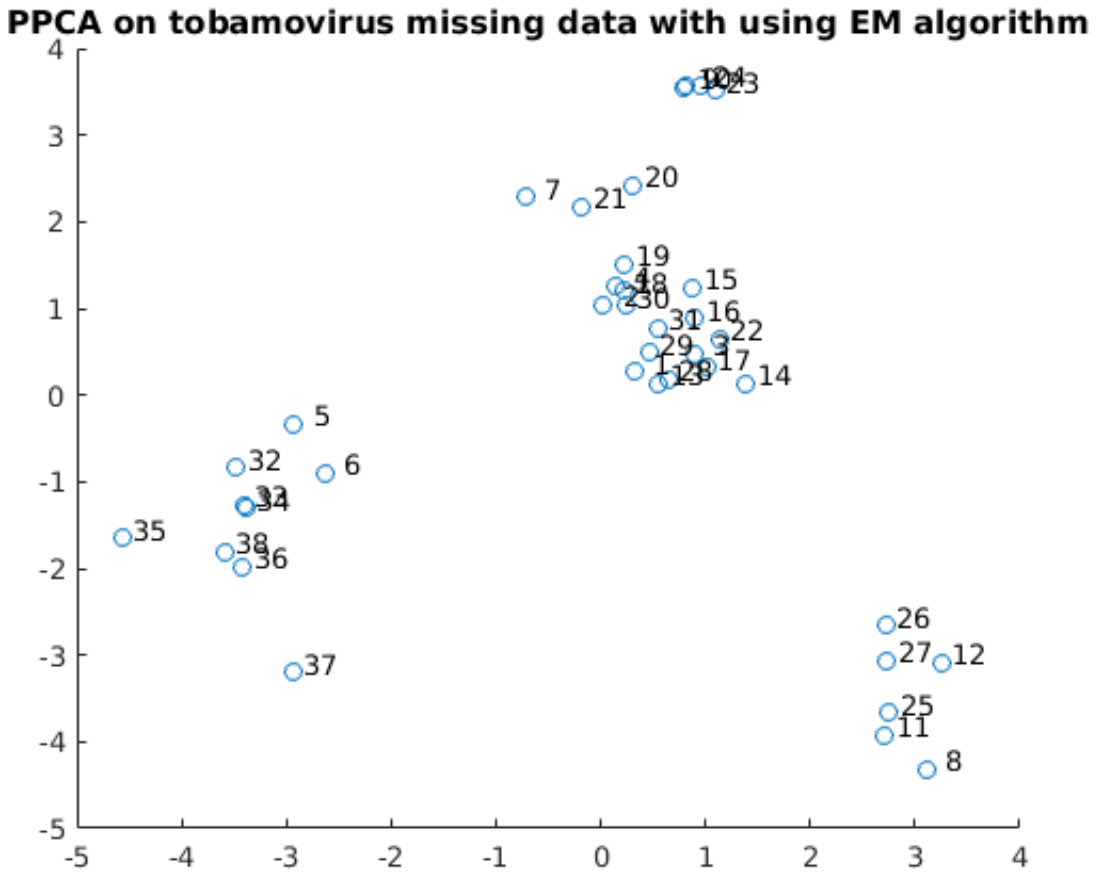


PPCA with missing data and EM

```
M = rand(size(Y)) > 0.9;
fprintf('Missing values count = %d\n', sum(sum(M)));
disp(sum(M));
fprintf('Data points with no missing elements - ');
MInstance = sum(M);

for i = 1 : size(Y, 2)
    if MInstance(i) == 0
        fprintf('%d ', i);
    end
end

[W, var, X] = PPCAMissingDataWithEM(Y, 2, M);
figure;
scatter(X(1, :), X(2, :));
T = cellstr(num2str([1:size(X, 2)]'));
text(X(1, :) + 0.1, X(2, :) + 0.1, T);
title('PPCA on tobamovirus missing data with using EM algorithm');
```

```
Missing values count = 70
  Columns 1 through 13

     0     3     4     0     2     0     4     3     1     3     0     2     3

  Columns 14 through 26

     2     4     2     2     0     0     1     1     1     2     1     1     2

  Columns 27 through 38
```

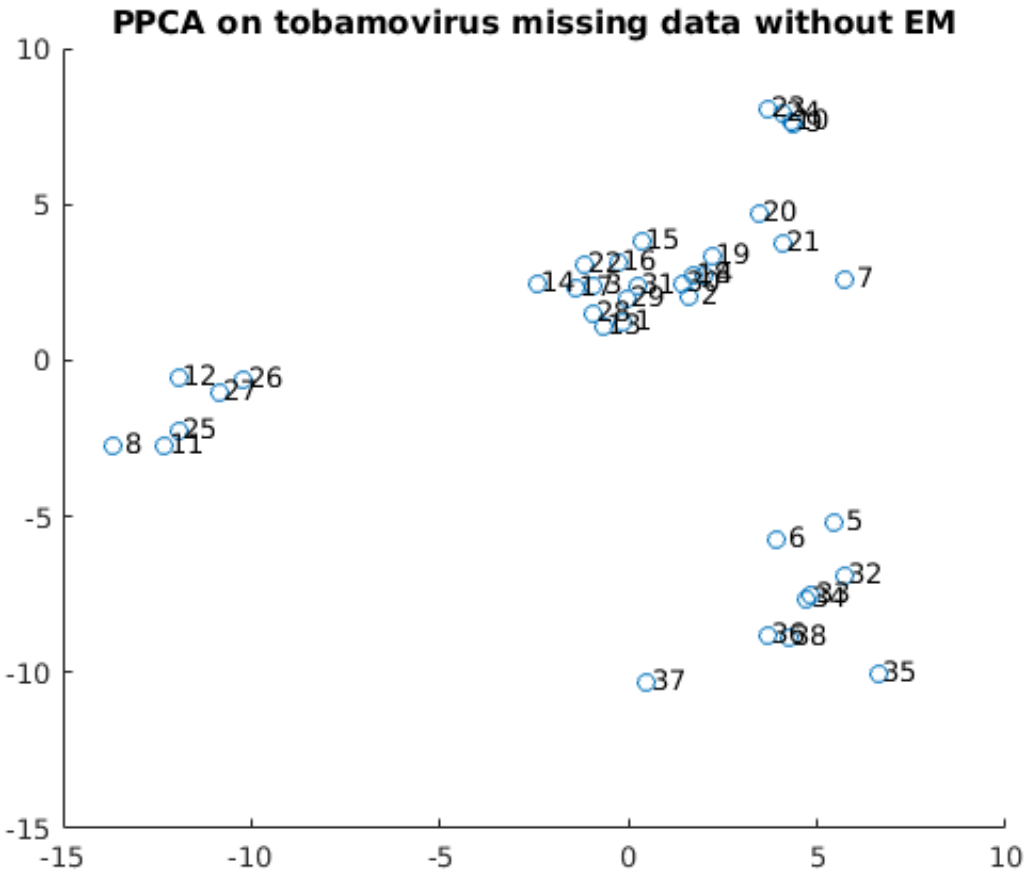
```
    2     4     2     3     1     3     2     0     3     3     2     1

Data points with no missing elements - 1 4 6 11 18 19 34 no of iterations = 124
```



PCA with missing data and no EM

```
[W, X] = PCAWithMissingData(Y, 2, M);
figure;
scatter(X(1, :), X(2, :));
T = cellstr(num2str([1:size(X, 2)]'));
text(X(1, :) + 0.1, X(2, :) + 0.1, T);
title('PPCA on tobamovirus missing data without EM');
```

```
Iteration count = 10
```



Published with MATLAB® R2015b