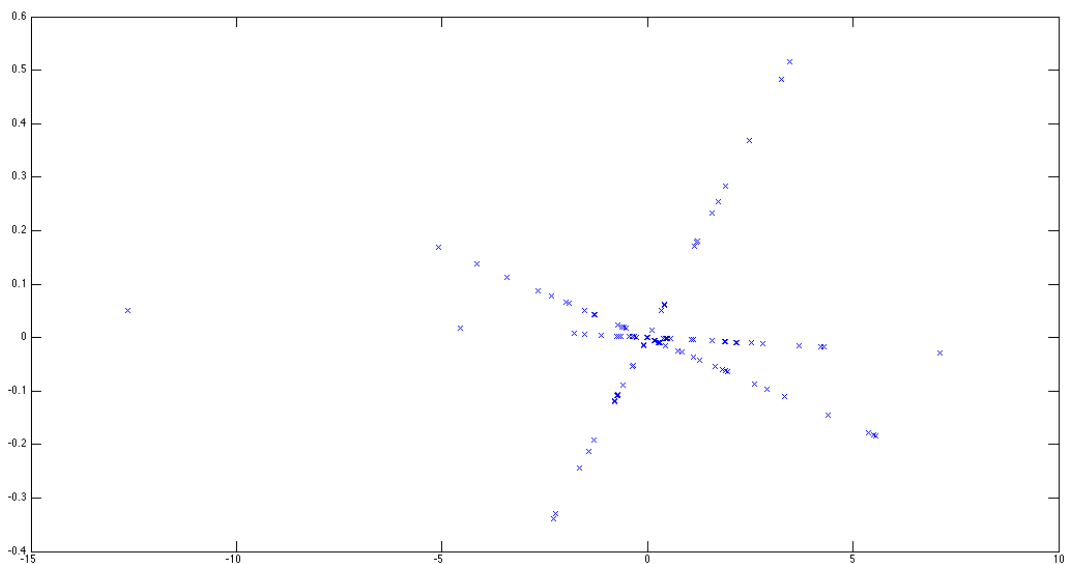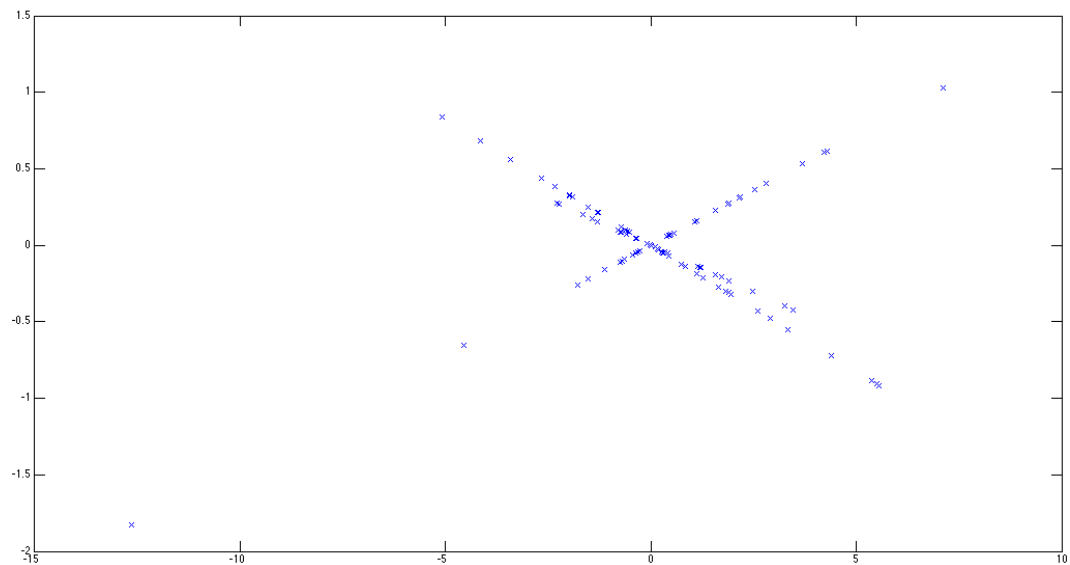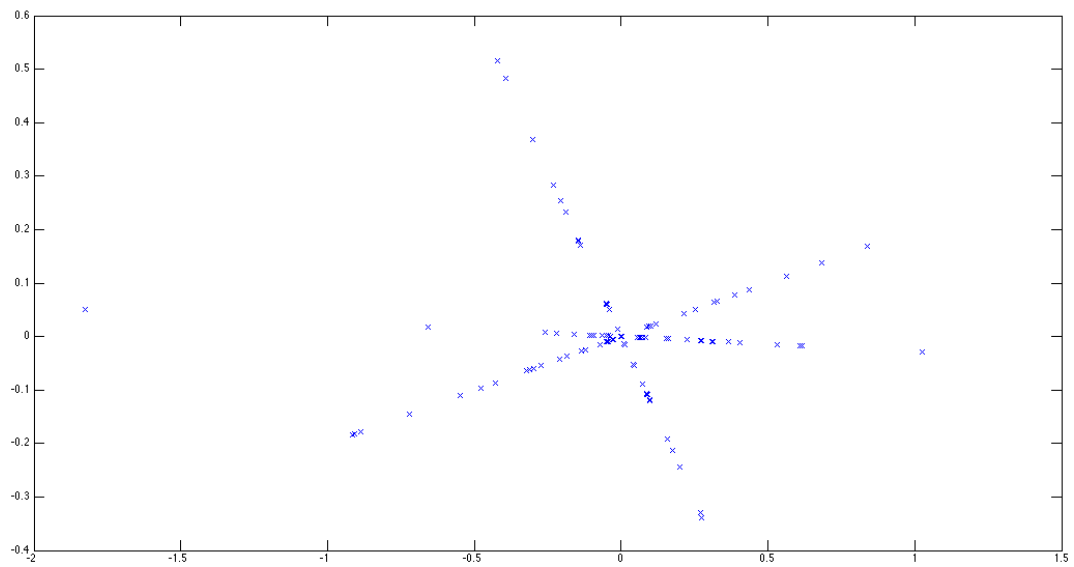# Hw3-3: ICA

15826 – Multimedia Databases and Data Mining

Fall 2013, C. Faloutsos
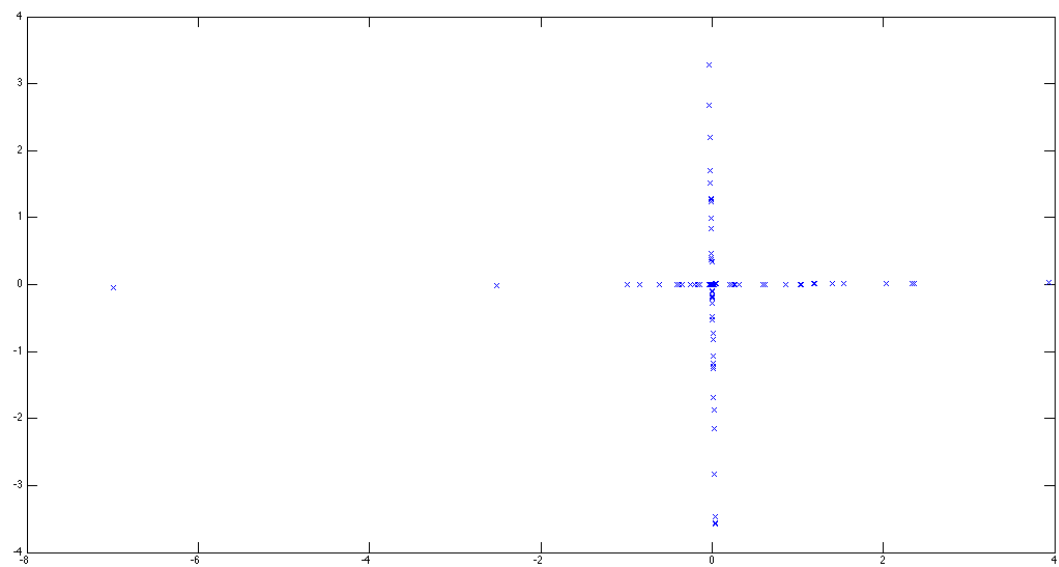
Emma R. Zhang{runyunz@andrew.cmu.edu} - November 18, 2013

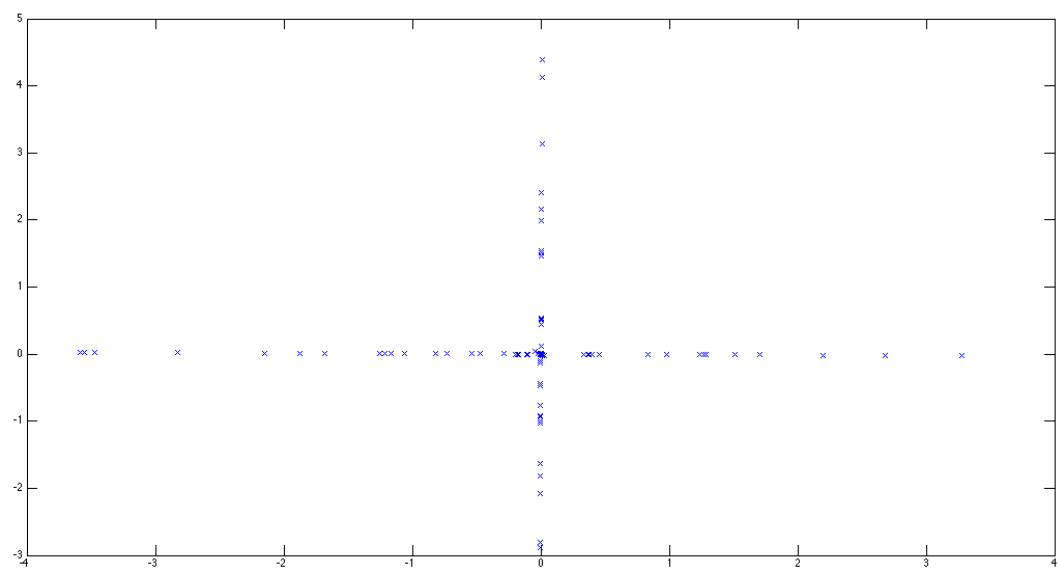## Plot

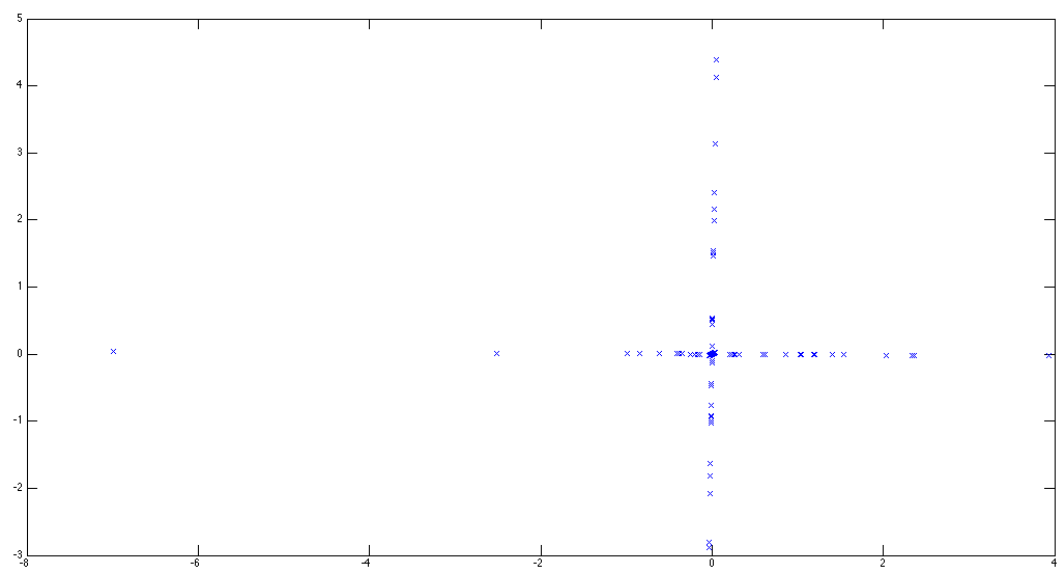*S1: PCA Plot*

S2: ICA Plot

Result

S3 - PCA Purity Score

3	8.352401e-01
5	8.557647e-01
6	8.557647e-01
8	8.515628e-01
9	8.352403e-01
10	8.515629e-01
14	8.557652e-01
15	8.557648e-01
17	8.352402e-01
18	8.352403e-01
19	8.515629e-01
20	8.352403e-01
21	8.557649e-01
22	8.515628e-01
24	8.515631e-01
25	8.557653e-01
26	8.352403e-01
28	8.352398e-01
29	8.352403e-01
31	8.352402e-01
33	8.557649e-01
34	8.515632e-01
35	8.557648e-01
36	8.352403e-01
41	8.515628e-01
43	8.352398e-01
45	8.352404e-01
48	8.352403e-01
49	8.557648e-01
50	8.352400e-01
52	8.557648e-01
53	8.352402e-01
55	8.515633e-01
56	8.557648e-01
57	8.352403e-01
58	8.557648e-01
59	8.557647e-01

61	8.515631e-01
63	8.557649e-01
64	8.557648e-01
67	8.557649e-01
68	8.352402e-01
69	8.352403e-01
70	8.557647e-01
72	8.352404e-01
73	8.515631e-01
75	8.515632e-01
77	8.515628e-01
80	8.515628e-01
82	8.515630e-01
83	8.352402e-01
84	8.557649e-01
86	8.352401e-01
87	8.352403e-01
91	8.557648e-01
94	8.352405e-01
98	8.352402e-01
99	8.352401e-01

sum(purity) for pca:4.906242e+01

S4 - ICA Score

3	9.887450e-01
5	9.932978e-01
6	9.933002e-01
8	9.886038e-01
9	9.887446e-01
10	9.886042e-01
14	9.932974e-01
15	9.932995e-01
17	9.887448e-01
18	9.887446e-01
19	9.886041e-01
20	9.887446e-01
21	9.932990e-01
22	9.886042e-01
24	9.886045e-01

25	9.933012e-01
26	9.887445e-01
28	9.887459e-01
29	9.887446e-01
31	9.887446e-01
33	9.932970e-01
34	9.886038e-01
35	9.932995e-01
36	9.887446e-01
41	9.886040e-01
43	9.887458e-01
45	9.887441e-01
48	9.887446e-01
49	9.932984e-01
50	9.887452e-01
52	9.932984e-01
53	9.887447e-01
55	9.886041e-01
56	9.932963e-01
57	9.887446e-01
58	9.932986e-01
59	9.933035e-01
61	9.886045e-01
63	9.933017e-01
64	9.933006e-01
67	9.932935e-01
68	9.887446e-01
69	9.887444e-01
70	9.932941e-01
72	9.887442e-01
73	9.886044e-01
75	9.886042e-01
77	9.886042e-01
80	9.886041e-01
82	9.886040e-01
83	9.887448e-01
84	9.932963e-01
86	9.887451e-01
87	9.887446e-01
91	9.933009e-01

94   9.887440e-01
98   9.887448e-01
99   9.887450e-01
sum(purity) for ica: 5.743175e+01

S5 - Comparison

The Result of ICA is better than PCA because it holds bigger purity score. Concretely, only one dimension in the point has a large value, and the rest of two are close to zero, which means every points are projected very close to one of the 3 axises, and thus have higher purity.

Code

q3.m

```matlab
[U,S,V] = svd(X);

U3D = U(:,1:3);
S3D = S(1:3,1:3);
X3D = U3D * S3D;

plot(X3D(:,1),X3D(:,2), 'x');
plot(X3D(:,1),X3D(:,3), 'x');
plot(X3D(:,2),X3D(:,3), 'x');


[XICA] = fastica(transpose(X), 'numOfIC', 3);
XICA = transpose(XICA);

plot(XICA(:,1),XICA(:,2), 'x');
plot(XICA(:,1),XICA(:,3), 'x');
plot(XICA(:,2),XICA(:,3), 'x');

[nrow,ncol]=size(X);

sall = 0;
for i=1:nrow,
    if norm(X(i,:)) > 1,
        tmp = sort(abs(X3D(i,:)), 'descend');
        score = 1 - tmp(2)/tmp(1);
        fprintf('%d\t%d\n', i, score);
        sall = sall  + score;
    end
end
fprintf('sum(purity) for pca:\t%d\n', sall);


sall = 0;
for i=1:nrow,
    if norm(X(i,:)) > 1,
        tmp = sort(abs(XICA(i,:)), 'descend');
        score = 1 - tmp(2)/tmp(1);
        fprintf('%d\t%d\n', i, score);
        sall = sall  + score;
    end
end
fprintf('sum(purity) for ica:\t%d\n', sall);
```