Problems 14–Chapter 7: Sparse Kernel Machines

Siou Lin (slin)

26 April 2007

1 Method

I used the method outlined in LIBSVM to classify three data sets. As outlined in *A Practical Guide to Support Vector Classification* by Hsu, Chang, and Lin, the steps are:

- Transform data to the format of LIBSVM
- Scale the data to be in the range -1 to 1
- Choose a kernel
- Use cross-validation to find the best parameter
- Use the best parameter C and γ to train the whole training set
- Test

2 Results

I ran LIBSVM on three different data sets: face, splice, and vowel. The prediction results worked best for the face data. It was also found that scaling and cross-validating the data did not always help achieve better results. Of the four kernels used, the linear and radial basis kernels consistently yielded the best results. The polynomial and sigmoid kernels consistently yielded the worst results.

The face data consisted of 2 classes, 200 training points, and 49 features. LIBSVM was used to predict 26 test points. For the unscaled data, $C = 2.0$, $\gamma = 3.05 * 10^{-5}$, and the cross-validation rate was 96.5%. For the scaled data, $C = 8.0$, $\gamma = 0.125$, and the cross-validation rate was 98.0%. Tables 1-4 show the resulting prediction statistics.

The splice data consisted of 2 classes, 1000 training points, and 60 features. LIBSVM was used to predict 2175 test points. Tables 5-6 show the resulting prediction statistics.

The vowel data consisted of 11 classes, 528 training points, and 10 features. LIBSVM was used to predict 462 test points. Since this data set came pre-scaled, no additional scaling had to be performed. $C = 8.0$, $\gamma = 2.0$, and the cross-validation rate was 99.43%. Tables 7-8 show the resulting prediction statistics.

The vowel and splice data came in the correct input format for LIBSVM so the face data was the only one that needed to be reformated. I reformated the face data using Matlab. The code is attached below.

```
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
% Problems 14: Formating Face Data
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
clear;
clc;

% load training data
face = load('face_test.txt');
no_face = load('no_face_test.txt');

% determine dimensions of training data
[N,D]=size(face);
N = 2*N;

% open output file
fid = fopen('test.txt','w');

% format file and interleave data
for i=1:N
    if mod(i,2) == 1
        fprintf(fid, '+1');
        for j=1:D
            fprintf(fid, ' %d:%f', j, face(ceil(i/2),j));
        end
        fprintf(fid, '\n');
    else
        fprintf(fid, '-1');
        for j=1:D
            fprintf(fid, ' %d:%f', j, no_face(ceil(i/2),j));
        end
        fprintf(fid, '\n');
    end
end

fclose(fid);
```

Kernel Type	Correct Classifications	Percentage Correct
linear	25	95.15%
polynomial	25	95.15%
radial basis	26	100.00%
sigmoid	25	95.15%

Table 1: Predictions without scaling and without cross validation for face data

Kernel Type	Correct Classifications	Percentage Correct
linear	26	100.00%
polynomial	16	61.54%
radial basis	26	100.00%
sigmoid	25	95.15%

Table 2: Predictions without scaling and with cross validation for face data

Kernel Type	Correct Classifications	Percentage Correct
linear	26	100.00%
polynomial	16	61.54%
radial basis	26	100.00%
sigmoid	25	96.15%

Table 3: Predictions with scaling and without cross validation for face data

Kernel Type	Correct Classifications	Percentage Correct
linear	26	100.00%
polynomial	25	95.15%
radial basis	25	95.15%
sigmoid	19	73.08%

Table 4: Predictions with scaling and with cross validation for face data

Kernel Type	Correct Classifications	Percentage Correct
linear	1844	84.78%
polynomial	1922	88.37%
radial basis	1962	92.21%
sigmoid	1131	52.00%

Table 5: Predictions without scaling and without cross validation for splice data

Kernel Type	Correct Classifications	Percentage Correct
linear	1207	55.49%
polynomial	1109	50.99%
radial basis	1117	51.36%
sigmoid	1131	52.00%

Table 6: Predictions without scaling and with cross validation for splice data

Kernel Type	Correct Classifications	Percentage Correct
linear	230	49.78%
polynomial	58	12.55%
radial basis	237	51.30%
sigmoid	200	43.29%

Table 7: Predictions with scaling and without cross validation for vowel data

Kernel Type	Correct Classifications	Percentage Correct
linear	245	53.03%
polynomial	237	51.30%
radial basis	285	61.69%
sigmoid	105	22.72%

Table 8: Predictions with scaling and with cross validation for vowel data