Quadratic Programming Project

Jianlin Cheng, PhD

Computer Science Department

University of Missouri, Columbia

Fall, 2013

Project – Implement a Small Support Vector Machine

```
-1  1:6.000000 2:148.000000 3:72.000000 4:35.000000 5:0.000000 6:33.599998 7:0.627000 8:50.000000   a1
+1  1:1.000000 2:85.000000 3:66.000000 4:29.000000 5:0.000000 6:26.600000 7:0.351000 8:31.000000    a2
-1  1:8.000000 2:183.000000 3:64.000000 4:0.000000 5:0.000000 6:23.299999 7:0.672000 8:32.000000    a3
+1  1:1.000000 2:89.000000 3:66.000000 4:23.000000 5:94.000000 6:28.100000 7:0.167000 8:21.000000   a4
-1  1:0.000000 2:137.000000 3:40.000000 4:35.000000 5:168.000000 6:43.099998 7:2.288000 8:33.000000 a5
+1  1:5.000000 2:116.000000 3:74.000000 4:0.000000 5:0.000000 6:25.600000 7:0.201000 8:30.000000    a6
```

$$W(\alpha) = \sum_{i=1}^{m} \alpha_i - \frac{1}{2} \sum_{i,j=1}^{m} y_i y_j \alpha_i \alpha_j K(\mathbf{x}_i, \mathbf{x}_j)$$

$$\mathbf{w} - \sum_i \alpha_i y_i \mathbf{x}_i = 0$$

$$\sum_i \alpha_i y_i = 0$$

$$C - \alpha_i - \mu_i = 0$$

$$y_i(\mathbf{w}^T \mathbf{x}_i - b) - 1 + \xi_i \geq 0$$

Project – Implement a Small Support Vector Machine

- Linear kernel (x.y or x.y + 1)
- SMO algorithm (Solution 1)
- Use an existing quadratic programming package given known constraints (Solution 2)
- Gradient descent (Solution 3)
- **Implement at least two of the three solutions**
- Compare results with a SVM method (e.g., SVM light)
- Report weights of data points, weight vector, support vectors (on boundary or slack variable is not 0), and b
- Report the accuracy on the training data
- Report the accuracy on some withheld test data

Data Set

- http://www.csie.ntu.edu.tw/~cjlin/libsvmtools/datasets/binary/diabetes
- First 10, 20, 30, …, 50, all data points
- Visualize how weights change during SMO optimization if you would like
- You may use other data sets too if the diabetes data does not work. (http://archive.ics.uci.edu/ml/)

Timeline of this project and the remaining of the semester

- **Nov. 13, discussion of Project 5**
- **Nov. 15, presentation of plan of Project 5**
- **Nov. 20, deep learning lecture**
- **Nov. 22, presentation of results of Project 5**
- **Dec. 4, reading assignment of deep learning due, class is rescheduled to Dec. 11**
- **Dec. 11 (Wed), final presentation of all projects**