Foundations of Machine Learning

Module 5: Support Vector Machine
Part F: SVM – Solution to the Dual Problem

Sudeshna Sarkar
IIT Kharagpur

The SMO algorithm

The SMO algorithm can efficiently solve the dual problem.
First we discuss Coordinate Ascent.

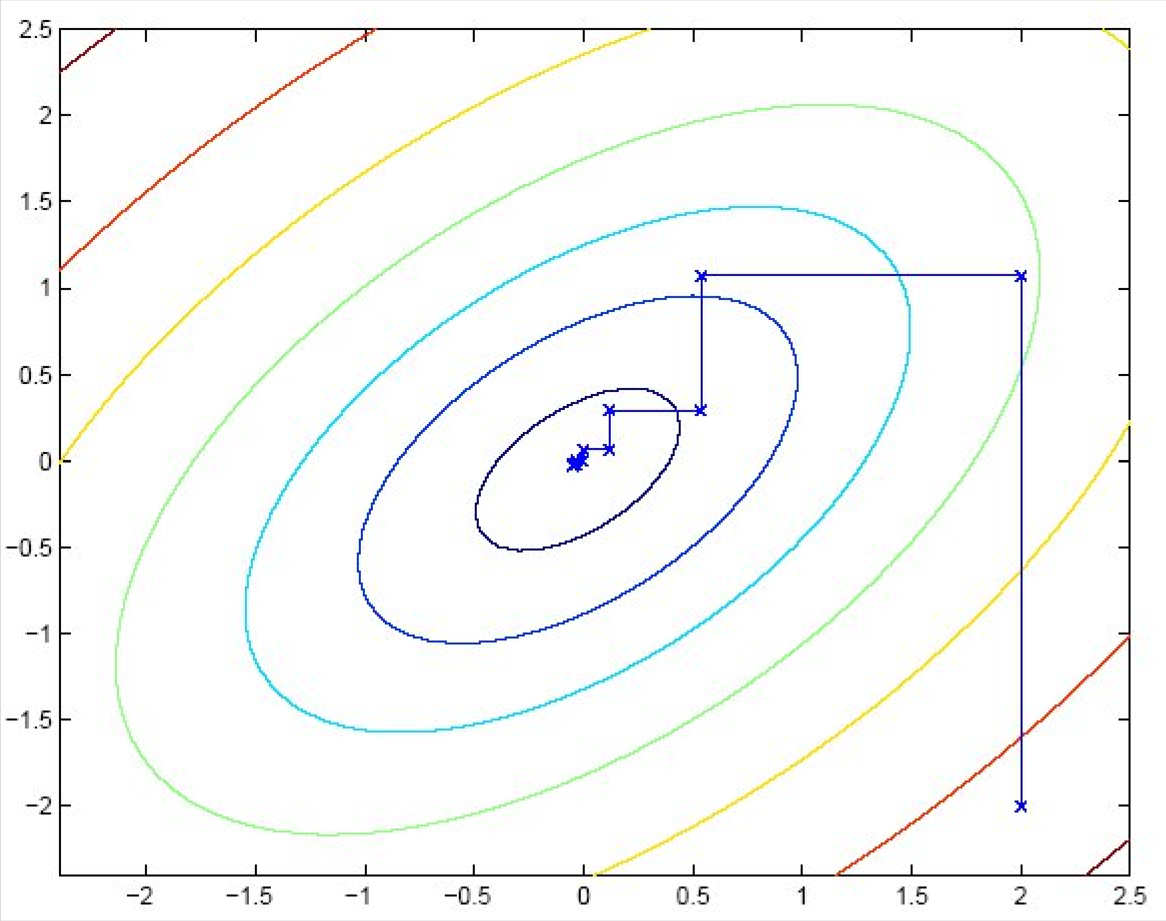
Coordinate Ascent

- Consider solving the unconstrained optimization problem:

$$\max_{\alpha} W(\alpha_1, \alpha_2, \ldots, \alpha_n)$$

Loop until convergence: {
    for $i = 1\ to\ n$ {
        $\alpha_i = arg \max_{\widehat{\alpha_i}} W(\alpha_1, \ldots, \widehat{\alpha_i}, \ldots, \alpha_n)$;
    }
}

Coordinate ascent



- Ellipses are the contours of the function.
- At each step, the path is parallel to one of the axes.

Sequential minimal optimization

- Constrained optimization:

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

$$\text{s.t.} \quad 0 \le \alpha_i \le C, \quad i = 1, \ldots, m$$

$$\sum_{i=1}^{m} \alpha_i y_i = 0.$$

- Question: can we do coordinate along one direction at a time (i.e., hold all $\alpha_{[-i]}$ fixed, and update α_i?)

The SMO algorithm

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

$$\text{s.t.} \quad 0 \le \alpha_i \le C, \quad i = 1, \ldots, m$$

$$\sum_{i=1}^{m} \alpha_i y_i = 0.$$

- Choose a set of α_1's satisfying the constraints.
- α_1 is exactly determined by the other α's.
- We have to update at least two of them simultaneously to keep satisfying the constraints.

The SMO algorithm

Repeat till convergence {

1. Select some pair α_i and α_j to update next (using a heuristic that tries to pick the two that will allow us to make the biggest progress towards the global maximum).
2. Re-optimize W(α) with respect to α_i and α_j, while holding all the other α_k 's ($k \neq i; j$) fixed.

}

- The update to α_i and α_j can be computed very efficiently.

Thank You