SMO

Repeat until convergence

- pick α_i, α_j $\quad i \neq j$
- optimize $W(\alpha)$ w.r.t. α_i, α_j, holding other parameters fixed and respecting box constraints $\alpha \in [0, C]$

- How to pick α_i & α_j? (heuristics)

(1) Find α_i that violates KKT;
favor support vectors, i.e., $\alpha_i \in (0, c)$

(2) Pick α_j : $\max_{\alpha_j} |E_i - E_j|$ where $E = f(x) - y$

- x_i & x_j are badly predicted in opposite directions < one too high / one too low

$\Rightarrow$ induces biggest changes in α_i & α_j

(from math: change is a function of $E_i - E_j$)

(3) Use constraint $\sum_{i=1}^{m} \alpha_i y_i = 0$ to solve for α_i as a function of α_j

$$\alpha_i y_i + \alpha_j y_j + \sum_{k \neq i,j} \alpha_k y_k = 0$$

$$y_i \Big(\alpha_i y_i + \alpha_j y_j + \sum_{k \neq i,j} \alpha_k y_k = 0 \Big)$$

note: $y_i^2 = 1$ always ($y_i \in \{-1, 1\}$)

$$\alpha_i + \alpha_j y_i y_j + y_i \sum_{k \neq i,j} \alpha_k y_k) = 0$$

$$\alpha_i = - \alpha_j y_i y_j - y_i \sum_{k \neq i,j} \alpha_k y_k$$

The Stata Center, Building 32

32 Vassar Street, Cambridge, Massachusetts 02139-4307

www.csail.mit.edu

MIT COMPUTER SCIENCE AND ARTIFICIAL INTELLIGENCE LABORATORY

CSAIL

4. Plug α_i back into $W(\alpha)$;
   get functn of α_j (and other <u>fixed</u> α_s)

5. Take derivative w.r.t. α_j to optimize.

$$\frac{d\,W(\alpha)}{d\,\alpha_j} = 0$$

   <u>Solve for</u> α_j

6. Plug new α_j in constraint $\sum_{i=1}^{m} \alpha_i y_i = 0$ to <u>solve for α_i</u>

7. If α_i or $\alpha_j \notin [0,C]$,
   <u>clip</u> α_j as appropriate to ensure $\alpha_i, \alpha_j \in [0,C]$

The Stata Center, Building 32
32 Vassar Street, Cambridge, Massachusetts 02139-4307
www.csail.mit.edu