Sequential Minimum Optimization.

$$\min_{\alpha} \frac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{m}\alpha_i\alpha_j y_i y_j k(x_i, x_j) - \sum_{i=1}^{m}\alpha_i$$

$$\text{s.t} \quad 0 \le \alpha_i \le C$$

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

By KKT conditions:

$$\alpha_i = 0 \Rightarrow y^{(i)}(\omega^T x^{(i)} + b) \ge 1$$

$$\alpha_i = C \Rightarrow y^{(i)}(\omega^T x^{(i)} + b) \le 1$$

$$0 < \alpha_i < C \Rightarrow y^{(i)}(\omega^T x^{(i)} + b) = 1$$

1. Choosing α_i and α_j to optimize

First we want to find bounds L and H such that $L \le \alpha_j \le H$ must hold in order for α_j to satisfy the constraint that $0 \le \alpha_j \le C$.

If $y^{(i)} \ne y^{(j)}$ $L = \max(0, \alpha_j - \alpha_i)$, $H = \min(C, C + \alpha_j - \alpha_i)$ ☆

If $y^{(i)} = y^{(j)}$ $L = \max(0, \alpha_i + \alpha_j - C)$ $H = \min(C, \alpha_i + \alpha_j)$ ☆

Now we want to find α_j so as to minimize the objective function.

$$\alpha_j := \alpha_j - \frac{y^{(j)}(E_i - E_j)}{\eta} \quad ☆$$

P.S

$$\boxed{\begin{array}{l}\text{In cornel version, it is } \gamma = 2(w_{t-1}^T x - y) \\ \alpha_t = \alpha_{t-1} - s\gamma_{t-1} \text{, where } s \text{ is step size.}\end{array}}$$

where $E_k = f(x^{(k)}) - y^{(k)}$

P.S < > kernel, $\eta = 2\langle x^{(i)}, x^{(j)}\rangle - \langle x^{(i)}, x^{(i)}\rangle - \langle x^{(j)}, x^{(j)}\rangle$ ☆

Next, clip α_j to lie within range $[L, H]$

$$\alpha_j \begin{cases} H & \text{if } \alpha_j > H \\ \alpha_j & \text{if } L \le \alpha_j \le H \\ L & \text{if } \alpha_j < L. \end{cases}$$

Finally, solve $\alpha_i \Rightarrow \alpha_i = \alpha_i + y^{(i)} y^{(j)} (\alpha_j^{(old)} - \alpha_j)$ ☆

Then, Computing the b threshold.

$$b_1 = b - E_i - y^{(i)}(\alpha_i - \alpha_i^{(old)})\langle x^{(i)}, x^{(i)}\rangle - y^{(j)}(\alpha_j - \alpha_j^{(old)})\langle x^{(i)}, x^{(j)}\rangle$$

$$b_2 = b - E_j - y^{(i)}(\alpha_i - \alpha_i^{(old)})\langle x^{(i)}, x^{(j)}\rangle - y^{(j)}(\alpha_j - \alpha_j^{(old)})\langle x^{(j)}, x^{(j)}\rangle$$

then $b := \begin{cases} b_1 & \text{if } 0 < \alpha_i < C \\ b_2 & \text{if } 0 < \alpha_j < C \\ (b1+b2)/2 & \text{otherwise} \end{cases}$ $\rightarrow$ if both, ~~both~~ threshold valid

---

Pseudo-code for simplified SMO

- Input: C - regularization parameter

  tol: numerical tolerance

  max-passes : max # of times to iterate over α's

  $(x^{(1)}, y^{(1)}), \ldots, (x^{(m)}, y^{(m)})$: training data

- Output: $\alpha \in R$: Lagrange multipliers for solution

$b \in R$: threshold for solution

- Initialize $\alpha_i = 0$, $\forall i$, $b = 0$
- Initialize passes $= 0$
- while(passes < max-passes):
  - num_changed_alphas $= 0$
  - for $i = 1, \dots, m$:
    - Calculate $E_i = f(x^{(i)}) - y^{(i)}$ using

      $\left(f(x) = \sum_{j=1}^{m} \alpha_j y^{(j)} k\langle x^{(i)}, x^{(j)} \rangle + b \right)$ ①
    - If $((y^{(i)} E_i < -tol \ \&\& \ \alpha_i < C) \ ||$

      $(y^{(i)} E_i > tol \ \&\& \ \alpha_i > 0))$
      - Select $j \neq i$ randomly
      - Calculate $E_j = f(x^{(j)}) - y^{(j)}$ using ①
      - Save old α's : $\alpha_i^{(old)} = \alpha_i$, $\alpha_j^{(old)} = \alpha_j$
      - Comput L and H by

        ② If $y^{(i)} \neq y^{(j)}$, $L = \max(0, \alpha_j - \alpha_i)$, $H = \min(C, C + \alpha_j - \alpha_i)$

        ③ If $y^{(i)} = y^{(j)}$, $L = \max(0, \alpha_i + \alpha_j - C)$, $H = \min(C, \alpha_i + \alpha_j)$

- If $(L == H)$:

  continue to next i

- Compute η by

(4) $\eta = 2\langle x^{(i)}, x^{(j)}\rangle - \langle x^{(i)}, x^{(i)}\rangle - \langle x^{(j)}, x^{(j)}\rangle$

- If $(\eta \geq 0)$:

  continue to next i

- Compute and clip new value for α_j using

(5) $\alpha_j := \alpha_j - \dfrac{y^{(j)}(E_i - E_j)}{\eta}$

(6) $\alpha_j := \begin{cases} H & \text{if } \alpha_j > H \\ \alpha_j & \text{if } L \leq \alpha_j \leq H \\ L & \text{if } \alpha_j < L \end{cases}$

- If $(|\alpha_j - \alpha_j^{(old)}| < 10^{-5})$:

  continue to next i

- Determine value for α_i using

(7) $\alpha_i := \alpha_i + y^{(i)} y^{(j)} (\alpha_j^{(old)} - \alpha_j)$

- Compute b_1 and b_2 using

⑧ $b_1 = b - E_i - y^{(i)}(\alpha_i - \alpha_i^{(old)})\langle x^{(i)}, x^{(i)}\rangle -$

$y^{(j)}(\alpha_j - \alpha_j^{(old)})\langle x^{(i)}, x^{(j)}\rangle$

⑨ $b_2 = b - E_j - y^{(i)}(\alpha_i - \alpha_i^{(old)})\langle x^{(i)}, x^{(j)}\rangle -$

$y^{(j)}(\alpha_j - \alpha_j^{(old)})\langle x^{(j)}, x^{(j)}\rangle$

- Compute b using

$$b := \begin{cases} b_1 & \text{if } 0 < \alpha_i < C \\ b_2 & \text{if } 0 < \alpha_j < C \\ (b_1 + b_2)/2 & \text{otherwise} \end{cases}$$

- num_changed_alphas := num_changed_alphas +1

End If

End for

- If (num_changed_alphas == 0):

  passes := passes + 1

- else:

  passes := 0

- End While