Homework 4
CPSC 8420
Fall 2024
Mohammad Alshurbaji

Problem 1: Soft margin SVM

$$\min \frac{1}{2}\|w\|_2^2 + C\sum_{i=1}^{m}\xi_i$$

$$\text{s.t. } y_i(w^T x_i + b) \geq 1 - \xi_i \quad (i=1,2,\dots,m)$$
$$\xi_i \geq 0 \quad (i=1,2,\dots,m)$$

another formulation:

$$\min \frac{1}{2}\|w\|_2^2 + \frac{C}{2}\sum_{i=1}^{m}\xi_i$$

$$\text{s.t. } y_i(w^T x_i + b) \geq 1 - \xi_i \quad (i=1,2,\dots,m)$$

1) let's assume: $\xi_i = 2\theta_i^2$, where $\theta_i \in R$.
→ since $\xi_i \geq 0$ means there is always exists $\theta_i \in R$ ($\xi_i = 2\theta_i^2$)

Then, by replacing every ξ_i:

$$\min \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{m}\theta_i^2$$

s.t.

$$y_i(w^T x_i + b) \geq 1 - 2\theta_i^2$$

Note that:
$\theta_i \in R$ means $1 - 2\theta_i^2$ has the same range as $1 - \xi_i$ with $\xi_i \geq 0$ because squaring ensure non-negativity.

$$\text{s.t. } y_i(w^T x_i + b) \geq 1 - \theta_i, \ \forall i$$

Now: Sub. $\theta_i = \frac{\xi_i^2}{2}$

$$\min \frac{1}{2}\|w\|^2 + \frac{C}{2}\sum_{i=1}^{m}\xi_i^2, \text{ s.t. } y_i(w^T x_i + b) \geq 1 - \xi_i \ \#$$

Also, if we think it in another way using derivative:

$$\frac{\partial}{\partial \xi_i}\left(\frac{1}{2}\|w\|^2 + \frac{C}{2}\sum_{i=1}^{m} \xi_i^2\right) = C\,\xi_i$$

when, $\xi_i < 0 \rightarrow$ the derivative $C\xi_i$ is negative, means increasing the obj. value.

when, $\xi_i > 0 \rightarrow$ the derivative $C\xi_i$ is positive, penalizing over larger values of ξ_i.

meaning the derivative $C\xi_i$ drives ξ_i to $\xi_i = 0$, or some small positive value depending on the constraints.

Problem 1, Section 2: The generalized Lagrangian of the new soft SVM opt. problem:

$$\boxed{L(w, b, \xi, \alpha, \mu) = \frac{1}{2}\|w\|_2^2 + \frac{C}{2}\sum_{i=1}^{m}\xi_i^2 - \sum_{i=1}^{m}\alpha_i\left(y_i[w^T\Phi(x_i) + b) - 1 + \xi_i\right] - \sum_{i=1}^{m}\mu_i\,\xi_i^2}$$

Problem 1, section 3: Minimization of $L(w,b,\xi_i)$ with respect to w, b, and ξ.

1) $$\frac{\partial L}{\partial w} = w - \sum_{i=1}^{m}\alpha_i y_i \Phi(x_i) = 0$$

2) $$\frac{\partial L}{\partial b} = -\sum_{i=1}^{m}\alpha_i y_i = 0 = \sum_{i=1}^{m}\alpha_i y_i = 0$$

3) $$\frac{\partial L}{\partial \xi_i} = \sum_{i=1}^{m} C\,\xi_i - \alpha_i - 2\mu_i\,\xi_i = 0$$
$$= \sum_{i=1}^{m}\xi_i(C\,2\mu_i) - \alpha_i = 0$$
$$= \sum_{i=1}^{m}\xi_i(C - 2\mu_i) - \alpha_i = 0$$

Now, by subs. the parameter values in the 1, 2, and 3 you can find each of w, b, and ξ_i.

Note, following to Section 4, there is only one dual parameter: and by this: 3 becomes

$C\,\xi_i - \alpha_i = 0$ # (Just to follow the slides procedure).

Problem 1, Section 4: what's the dual of this version soft margin SVM opt. problem?

Eq (9):

$$\max_{\alpha} -\frac{1}{2} \sum_{i=1}^{m} \sum_{j=1}^{m} \alpha_i \alpha_j y_i y_j \langle x_i, x_j \rangle + \sum_{i=1}^{m} \alpha_i$$

which is equ. to: Eq. (10):

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

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

$$\alpha_i \geq 0, \quad i = 1, 2, \dots m$$

Sol.: by plug in the min. results in the soft lagrangian:

1) $w = \sum_{i=1}^{m} \alpha_i y_i x_i$

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

3) $\xi_i = \frac{\alpha_i}{C}$

$$L(\alpha) = \frac{1}{2} \|w\|^2 + \frac{C}{2} \sum_{i=1}^{m} \xi_i^2 - \sum_{i=1}^{m} \alpha_i [y_i(w^T x_i + b) - 1 + \xi_i]$$

First: Subs. w: into $\|w\|^2$:

$$\frac{1}{2}\|w\|^2 = \frac{1}{2} \left(\sum_{i=1}^{m} \alpha_i y_i x_i \right)^T \left(\sum_{j=1}^{m} \alpha_i y_i x_i \right)$$

$$= \frac{1}{2} \sum_{i=1}^{m} \sum_{j=1}^{m} \alpha_i \alpha_j y_i y_j \langle x_i, x_j \rangle$$

Second: Subs. ξ_i into $\frac{C}{2}\sum_{i=1}^{m}\xi_i^2$

$$\frac{C}{2}\sum_{i=1}^{m}\xi_i^2 = \frac{C}{2}\sum_{i=1}^{m}\left(\frac{\alpha_i}{C}\right)^2 = \frac{1}{2C}\sum_{i=1}^{m}\alpha_i^2$$

Final: all terms:

$$L(\alpha) = \frac{-1}{2}\sum_{i=1}^{m}\sum_{j=1}^{m}\alpha_i\alpha_j y_i y_j (x_i^T x_j) + \sum_{i=1}^{m}\alpha_i - \frac{1}{2C}\sum_{i=1}^{m}\alpha_i^2$$

Now: Duality:

$$\max_{\alpha} \sum_{i=1}^{m}\alpha_i - \frac{1}{2}\sum_{i=1}^{m}\sum_{j=1}^{m}\alpha_i\alpha_j y_i y_j \langle x_i, x_j\rangle - \frac{1}{2C}\sum_{i=1}^{m}\alpha_i^2$$

which is equal:

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

$$\text{s.t.} \quad \sum_{i=1}^{m}\alpha_i y_i = 0$$

$$\alpha_i \geq 0, \quad i = 1, 2, \dots, m$$

#.