Problem 1

$$\arg\min \frac{1}{2}\|w\|^2 + \frac{C}{2}\sum_{i=1}^{n}\xi_i^2 \quad \text{s.t. } y_i(w^Tx_i+b) \geq 1-\xi_i \quad \forall i\in[n]$$

① - Lagrangian Primal Function:

- use Lagrangian Multiplier to make in terms of α

$$L_p = \frac{1}{2}\|w\|^2 + \frac{C}{2}\sum_{i=1}^{n}\xi_i^2 - \sum_{i=1}^{n}\alpha_i\left[y_i(w^Tx_i+b) - 1 + \xi_i\right]$$

② - Minimize w.r.t w, b, and ξ

solve for w

$$\frac{\partial L}{\partial w} = w + 0 - \sum_{i=1}^{n}\alpha_i y_i x_i + 0 \rightarrow \boxed{w = \sum_{i=1}^{n}\alpha_i y_i x_i}$$

$$\frac{\partial L}{\partial b} = 0 + 0 + \sum_{i=1}^{n}\alpha_i y_i + 0 \rightarrow \frac{\partial L}{\partial b} = \boxed{\sum_{i=1}^{n}\alpha_i y_i = 0}$$

New constraint

$$\frac{\partial L}{\partial \xi} = 0 + C\xi - 0 - \alpha_i \longrightarrow \frac{\partial L}{\partial \xi} = C\xi_i - \alpha_i = 0 \rightarrow \boxed{\xi_i = \frac{\alpha_i}{C}}$$

"b" gets canceled out since	$\boxed{\sum_{i=1}^{n}\alpha_i y_i = 0}$

$$\|w\|_2^2 = w^Tw = \left(\sum_{i=1}^{n}\alpha_i y_i x_i\right)^T\left(\sum_{j=1}^{n}\alpha_j y_j x_j\right) = \sum_{i=1}^{n}\sum_{j=1}^{n}\alpha_i\alpha_j y_i y_j x_i^T x_j$$

From Lecture (2/7-2/9)

③ - Re-work to Lagrangian Dual Function:

$$L_D = \frac{1}{2}\sum\sum\alpha_i\alpha_j y_i y_j x_i^T x_j - \sum_{i=1}^{n}\alpha_i\left[y_i\left(\left[\sum_{j=1}^{n}\alpha_j y_j x_j\right]^T x_i + b\right) - 1 + \frac{\alpha_i}{C}\right]$$

$$= \frac{1}{2}\sum\sum\alpha_i\alpha_j y_i y_j x_i^T x_j - \sum_{i=1}^{n}\sum_{j=1}^{n}\alpha_i\alpha_j y_i y_j x_i^T x_j - b\sum_{i=1}^{n}\alpha_i y_i + \sum_{i=1}^{n}\alpha_i - \sum_{i=1}^{n}\frac{\alpha_i^2}{C}$$

$$+ \frac{C}{2}\sum_{i=1}^{n}\frac{\alpha^2}{C^2}$$

$$\frac{1}{C}\left(1 - \frac{1}{2}\right)$$

$$= -\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\alpha_i\alpha_j y_i y_j x_i^T x_j + \sum_{i=1}^{n}\alpha_i + \underbrace{\frac{1}{C}\sum_{i=1}^{n}\alpha^2 - \frac{1}{2C}\sum_{i=1}^{n}\alpha^2}_{-\frac{1}{2C}\sum_{i=1}^{n}\alpha_i^2}$$

$$\arg\max_{\alpha} \sum_{i=1}^{n}\alpha_i - \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\alpha_i\alpha_j y_i y_j x_i^T x_j - \frac{1}{2C}\sum_{i=1}^{n}\alpha_i^2$$

$$\text{s.t. } \alpha_i \geq 0 \quad \forall i \in [n] \quad \text{and} \quad \sum_{i=1}^{n}\alpha_i y_i = 0$$

Problem 2

Problem 2 Domain Adaptation SVM

Primal optimization problem: $\arg\min \frac{1}{2}\|w_T\|_2^2 + C\sum_{i=1}^{n}\xi_i - B w_T^T w_s$

s.t.

ξ_i : Slack variable

w_T : h.p target

w_s : h.p. Source

C, B : regularization params.

constraints: ① $y_i(w_T^T x_i + b) \geq 1 - \xi_i \quad \forall i \in \{1, \dots, n\}$

② $\xi_i \geq 0 \quad \forall i \in \{1, \dots, n\}$

Find h.p w_T, to minimize the opt problem s.t constraints.

Find the dual form.

i) Using Lagrangian multipliers (α, μ) in Primal form:

$$L_P = \frac{1}{2}\|w\|_2^2 + C\sum_{i=1}^{n}\xi_i - B w_T^T w_s - \sum_{i=1}^{n}\alpha_i\left[y_i(w^T x_i + b) - 1 + \xi_i\right] - \sum_{i=1}^{n}\xi_i \mu_i$$

ii) Minimize: find $\frac{\partial L_P}{\partial w_T}, \frac{\partial L_P}{\partial b}, \frac{\partial L_P}{\partial \xi}$:

$$\frac{\partial L_P}{\partial w_T} = w_T + 0 - B w_s - \sum_{i=1}^{n}\alpha_i y_i x_i - 0 = 0$$

$$= w_T - B w_s - \sum \alpha_i y_i x_i = 0 \longrightarrow w_T = B w_s + \sum_i \alpha_i y_i x_i$$

$$\frac{\partial L_P}{\partial b} = 0 + 0 - 0 + \sum_i \alpha_i y_i - 0 = 0$$

$$\sum_i \alpha_i y_i = 0$$

$$\frac{\partial L_P}{\partial \xi} = 0 + C - 0 - \alpha_i - \mu_i = 0$$

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

New constraint: $0 \leq \alpha_i \leq C$

note: $\|w\|_2^2 = w^T w = (B w_s + \sum_i \alpha_i y_i x_i)^T (B w_T + \sum_i \alpha_i y_i x_i)$

$$= B^2 w_s^T w_s + B w_s^T \sum \alpha y x + B w_s \sum \alpha y x^T + \sum_i \sum_j \alpha_i \alpha_j y_i y_j x_i^T x_j$$

$$= \frac{1}{2}\overbrace{\left[B^2 w_s^T w_s + B w_s^T \sum_i \alpha y x + B w_s \sum_i \alpha y x^T + \sum_i \sum_j \alpha_i \alpha_j y_i y_j x_i^T x_j \right]}^{Q}$$

$$+ C \sum_i \xi_i - B\left[B w_s + \sum_i \alpha y x\right]^T w_s$$

$$- \sum_i \alpha_i \left[y_i \left(\left[B w_s + \sum_i \alpha y x\right]^T x_i - b \right) - 1 + \xi_i \right] - \sum \xi_i \mu_i$$

$$= \frac{1}{2} Q + C \sum_i \xi_i - B^2 w_s^T w_s - B w_s \sum_i \alpha y x^T$$

$$\underbrace{- \sum \alpha_i \left[B w_s^T y_i x_i + \sum_i \sum_j \alpha_i y_i y_j x_i^T x_j - b y_i - 1 + \xi_i \right]} - \sum \xi_i \mu_i$$

$$- B w_s^T \sum_i \alpha_i y_i x_i - \sum_i \sum_j \alpha_i \alpha_j y_i y_j x_i^T x_j + b \sum \alpha_i y_i + \sum \alpha_i - \sum \alpha_i \xi_i$$

$$= \frac{1}{2} B^2 w_s^T w_s + \frac{1}{2} B w_s^T \sum_i \alpha y x + \frac{1}{2} B w_s \sum \alpha y x^T + \frac{1}{2} \sum_i \sum_j \alpha_i \alpha_j y_i y_j x_i^T x_j$$

$$+ C \sum \xi_i - B^2 w_s^T w_s - B w_s \sum_i \alpha y x^T$$

$$- B w_s^T \sum_i \alpha_i y_i x_i - \sum_i \sum_j \alpha_i \alpha_j y_i y_j x_i^T x_j + \overset{0}{b \sum_i \alpha_i y_i} + \sum \alpha_i - \sum_i \alpha_i \xi_i - \sum_i \xi_i \mu_i$$

$$= -\frac{1}{2} B^2 w_s^T w_s - \frac{1}{2} B w_s^T \sum_i \alpha_i y_i x_i - \frac{1}{2} B w_s \sum \alpha_i y_i x_i^T$$

$$- \frac{1}{2} \sum_i \sum_j \alpha_i \alpha_j y_i y_j x_i^T x_j + C \sum \xi_i + \sum \alpha_i - \sum \alpha_i \xi_i - \sum \xi_i \mu_i$$

Pull out $\sum \xi_i$

$$\sum \xi_i (C - \alpha_i - \mu_i)$$

goes to zero

$$= -\frac{1}{2} B^2 w_s^T w_s + \sum \alpha_i - B w_s \sum \alpha_i y_i x_i$$

$$- \frac{1}{2} \sum_i \sum_j \alpha \alpha y y x^T x$$

$$= -\frac{B^2}{2} w_s^T w_s + \sum_i \alpha_i \left(1 - B y_i x_i^T w_s \right) - \frac{1}{2} \sum_i \sum_j \alpha_i \alpha_j y_i y_j x_i^T x_j$$

Since we are optimizing alpha, the front term is constant and can be dropped

$$-\frac{\beta^2}{2} w_s^T w_s +$$

$$\underset{\alpha}{\arg\max}\ \overbrace{\sum \alpha_i}^{f} (1 - B y_i x_i^T w_s) - \overbrace{\frac{1}{2} \sum\sum \alpha_i \alpha_j y_i y_j x_i^T x_j}^{H}$$

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

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