Completed by 张晏杭

- For the optimization problem

$$min\frac{1}{2}\left||\beta|\right|^2 + C\sum_{i=1}^{n}\xi_i$$

subject to $y_i(\beta_0 + x^T\beta) + \xi_i \geq 1, \xi_i \geq 0, i = 1, \dots, n$

derive the dual maximization problem:

$$F_D(\alpha) = \sum_{i=1}^{n}\alpha_i - \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\alpha_i\alpha_j y_i y_j x_i^T x_j$$

subject to: $0 \leq \alpha_i \leq C, \sum_{i=1}^{n}\alpha_i y_i = 0$

Proof:

the lagrange function of the optimization problem is :

$$L(\beta, \beta_0, \xi, \alpha, \mu) \equiv \frac{1}{2}\left||\beta|\right|^2 + C\sum_{i=1}^{n}\xi_i - \sum_{i=1}^{n}\alpha_i(y_i(\beta_0 + x^T\beta) - 1 + \xi_i) - \sum_{i=1}^{n}\mu_i\xi_i$$

Where, $\alpha_i \geq 0$, $\mu_i \geq 0$.

First, take the derivative of $L(\beta, \beta_0, \xi, \alpha, \mu)$ with respect to ξ, α, μ.

$$\nabla_\xi L(\beta, \beta_0, \xi, \alpha, \mu) = \beta - \sum_{i=1}^{n}\alpha_i y_i x_i = 0$$

$$\nabla_\alpha L(\beta, \beta_0, \xi, \alpha, \mu) = -\sum_{i=1}^{n}\alpha_i y_i = 0$$

$$\nabla_\mu L(\beta, \beta_0, \xi, \alpha, \mu) = C - \alpha_i - \mu_i = 0$$

We get that

$$\beta = \sum_{i=1}^{n}\alpha_i y_i x_i \quad (1)$$

$$\sum_{i=1}^{n}\alpha_i y_i = 0 \quad (2)$$

$$C - \alpha_i - \mu_i = 0 \quad (3)$$

Substitute equation (1)~(3) into $L(\beta, \beta_0, \xi, \alpha, \mu)$:

$$\min\ L(\beta, \beta_0, \xi, \alpha, \mu) = \sum_{i=1}^{n}\alpha_i - \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\alpha_i\alpha_j y_i y_j x_i^T x_j$$

Then, find the maximum of α for $\min\ L(\beta, \beta_0, \xi, \alpha, \mu)$ and derive the dual problem:

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

s.t.

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

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

$$\alpha_i \geq 0,\ \mu_i \geq 0 \qquad i = 1,2,3 \dots N$$

We can eliminate the μ_i in the equality constraint $C - \alpha_i - \mu_i = 0$

$$\because \quad \mu_i \geq 0$$
$$\therefore \quad 0 \leq \alpha_i \leq C$$

So we can rewrite the condition as:

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

□