

Ejercicio 3

1

$$(P) \begin{cases} \min & \frac{1}{2} \|a\|^2 \\ \text{st.} & y_j (a^T x_j + b) \geq 1 \quad j=1, \dots, m \end{cases}$$

a) Se busca obtener el problema dual.

Paso 1

$$\begin{aligned} L(a, b, u) &= \frac{1}{2} \|a\|^2 - \sum_{j=1}^m u_j [y_j (a^T x_j + b) - 1] \\ &= \underbrace{\frac{1}{2} \|a\|^2 - \sum_{j=1}^m u_j (y_j a^T x_j - 1)}_{L_1(a, u)} - \underbrace{b \sum_{j=1}^m u_j y_j}_{L_2(b, u)} \\ &= L_1(a, u) + L_2(b, u) \end{aligned}$$

Busquemos ahora los $a(u)$ y $b(u)$ óptimos de $L(a, b, u)$.

$$\frac{\partial L_1}{\partial a} = a - \sum_{j=1}^m u_j y_j x_j = 0 \Leftrightarrow \boxed{a = \sum_{j=1}^m u_j y_j x_j}$$

Mientras que el $b(u)$ no se puede hallar derivando L_2 ya que es una función lineal. Por lo tanto el factor que multiplica a b deberá ser 0, en caso contrario puede haber que con $b = \pm \infty$ el $L = -\infty$

Paso 2

2

$$L_1\left(\sum_{j=1}^m \mu_j y_j x_j, \mu\right) = \frac{1}{2} \left[\sum_{j=1}^m \mu_j y_j x_j^T \sum_{i=1}^m \mu_i y_i x_i \right] - \sum_{j=1}^m \mu_j y_j \left[\sum_{i=1}^m \mu_i y_i x_i \right] x_j^T + \sum_{j=1}^m \mu_j$$

(1) (2) (3)

Si se reordenan los terminos de (2) se llega a que (2) = (1). Por lo que

$$\begin{aligned} L_1\left(\sum_{j=1}^m \mu_j y_j x_j, \mu\right) &= -\frac{1}{2} \sum_{j=1}^m \mu_j y_j x_j^T \sum_{i=1}^m \mu_i y_i x_i + \sum_{j=1}^m \mu_j \\ &= -\sum_{j=1}^m \sum_{i=1}^m \mu_j \underbrace{\frac{1}{2} y_j y_i x_j^T x_i}_{Q_{ji}} \mu_i + \sum_{j=1}^m \mu_j \end{aligned}$$

Notar que se tiene $\sum_{i,j} Q_{ji} \mu_j \mu_i$ la cual es una forma cuadrática cuya representación matricial es $\mu^T Q \mu$.

Además notar que la sumatoria de todas las componentes de un vector (μ) se puede escribir como el producto interno de un vector de 1's con μ .

Entonces:

$$L_1(a^*(\mu), \mu) = -\mu^T Q \mu + \mathbf{1}^T \mu$$

con $Q_{ij} = \frac{1}{2} y_i y_j \langle x_i, x_j \rangle$

3

Para $L_2(b^*(u), u)$ tendremos que bien es $-\infty$ o Q .

$$L_2(b^*(u), u) = \begin{cases} -\infty & \text{si } \sum_{j=1}^n u_j y_j \neq 0 \\ Q & \text{si } \sum_{j=1}^n u_j y_j = 0 \end{cases}$$

Paso 3

Se plantea $d(u)$ como $d(u) = L_1(b^*(u), u) + L_2(b^*(u), u)$

$$\Rightarrow d(u) = \begin{cases} -\infty & \text{si } y^T u \neq 0 \\ -u^T Q u + 1^T u & \text{si } y^T u = 0 \end{cases}$$

El problema dual es:

$$(D_1) \begin{cases} \max_u \begin{cases} -\infty & \text{si } y^T u \neq 0 \\ -u^T Q u + 1^T u & \text{si } y^T u = 0 \end{cases} \\ u \geq 0 \end{cases}$$

Notar que el caso de $-\infty$ puede ser escrito como una restriccion. Es decir, D_1 se puede reescribir como:

$$(D) \begin{cases} \max -u^T Q u + 1^T u \\ \text{st } y^T u = 0 \\ u \geq 0 \end{cases} \quad \text{con } Q_{ij} = y_i y_j \langle x_i, x_j \rangle$$

b) (P) es un problema convexo. Esto implica que se cumple dualidad fuerte y por lo tanto tambien se cumple complementary slackness (c.s.).

C.S. dice que se cumple ① o ②

$$\begin{cases} \gamma_j (a^T x_j + b) = 1 & \text{①} \\ \mu_j = 0 & \text{②} \end{cases}$$

La condición ① implica que x_j es un vector de soporte. Si no es vector de soporte μ_j sera 0.

Como $a = \sum_{j=1}^m \mu_j \gamma_j x_j$ y $\mu_j = 0 \quad \forall j / x_j$ no es de soporte

$\Rightarrow a$ es combinacion lineal de vectores de soporte