Ejercicio 3

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

a) Se buscará obtener el problema dual.

Paso 1

$$L(a,b,\mu) = \frac{1}{2}\|a\|^2 - \sum_{j=1}^{m} \mu_j\left[y_j(a^T x_j + b) - 1\right]$$

$$= \underbrace{\frac{1}{2}\|a\|^2 - \sum_{j=1}^{m} \mu_j(y_j a^T x_j - 1)}_{L_1(a,\mu)} \underbrace{- b\sum_{j=1}^{m} \mu_j y_j}_{L_2(b,\mu)}$$

$$= L_1(a,\mu) + L_2(b,\mu)$$

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

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

Mientras que el $b(\mu)$ 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 puedo hacer que con $b = \pm\infty$ el $L = -\infty$

Paso 2

$$L_1\Big(\overbrace{\sum_{j=1}^{m} u_j y_j x_j}^{a}, u\Big) = \overbrace{\frac{1}{2}\sum_{j=1}^{m} u_j y_j x_j^T \sum_{i=1}^{m} u_i y_i x_i}^{①} - \overbrace{\sum_{j=1}^{m} u_j y_j \Big(\sum_{i=1}^{m} u_i y_i x_i\Big) x_j^T}^{②}$$
$$+ \underbrace{\sum_{j=1}^{m} u_j}_{③}$$

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

$$L_1\Big(\sum_{j=1}^{m} u_j y_j x_j, u\Big) = -\frac{1}{2}\sum_{j=1}^{m} u_j y_j x_j^T \sum_{i=1}^{m} u_i y_i x_i + \sum_{j=1}^{m} u_j$$
$$= -\sum_{j=1}^{m}\sum_{i=1}^{m} u_j \underbrace{\frac{1}{2} y_j y_i x_j^T x_i}_{Q_{ji}} u_i + \sum_{j=1}^{m} u_j$$

Notar que se tiene $\sum_{i,j} Q_{ji} u_j u_i$ la cual es una forma cuadratica cuya representacion matricial es $u^T Q u$.

Además notar que la sumatoria de todos los componentes de un vector (u) se puede escribir como el producto interno de un vector de 1's con u.

Entonces:

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

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

Para $L_2(b^*(u),u)$ tendremos que bien es $-\infty$ o 0.

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

Paso 3

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

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

El problema dual es:

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

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

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

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

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

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

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

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