→ SVM (Support vector Machine).

การเลือกใช้ระนาบสร้าง hyperplane ที่แบ่ง X_+, X_- ออกจากกัน นอกจาก Line search algorithm เราสามารถพิจารณา ค่าน้อยสุดของ $\|w\|$ (หรือ $\frac{1}{2}\|w\|^2$) ภายใต้เงื่อนไข.

$$\langle w, z\rangle \geq c > 0 \quad \text{ทุก } z \in C = conv(X_+ - X_-)$$

เกี่ยวข้องกับ θ_w.

เมื่อพิจารณา C จากการพิสูจน์ก่อนหน้านี้ เราจะเห็นได้ว่า สามารถเลือกได้

$$\exists \theta : \begin{cases} \langle w, x\rangle - \theta \geq 1, & \forall x \in X_+ \\ -1(\langle w, y\rangle - \theta \leq -1), & \forall y \in X_- \end{cases}$$

$$\langle w, -y\rangle + \theta \geq 1$$

จะได้ $\langle w, \underbrace{x - y}_{=z}\rangle \geq 2 \quad \forall z \in conv(X_+ - X_-)$

โดยสามารถเขียนอยู่ในรูปของ quadratic optimization ได้ดังนี้.

$$\begin{cases} \text{Minimize} & \frac{1}{2}\|w\|^2 \\ \text{subject to} & y_i(\langle w, x_i\rangle - \theta) \geq 1, \quad i = 1, \dots, N. \end{cases}$$

โดยที่ $y_i = \pm 1$, $x_i \in X_\pm$

Thm: (Kuhn-Tucker: generalized Lagrange Multiplier.)

ให้ $f : \mathbb{R}^m \to \mathbb{R}$ และ $g_i : \mathbb{R}^m \to \mathbb{R}$, $i = 1, \dots, N$.
เป็นฟังก์ชัน ต่อเนื่องและสามารถหาอนุพันธ์ได้. แล้ว ปัญหาค่าต่ำสุดของ f ในรูป.

$$\boxed{\begin{array}{l} \text{Minimize } f(x) \\ \text{subject to } g_1(x) \le 0, \ldots, g_N(x) \le 0. \end{array}}$$

จะสามารถเขียนได้ในรูป.

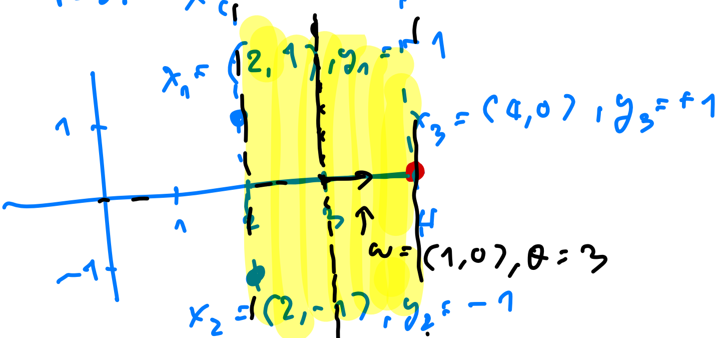
$$\boxed{\begin{array}{l} \text{Maximize } F(x,\lambda) := f(x) + \sum_{i=1}^{N} \lambda_i g_i(x) \\ \text{subject to } G(x,\lambda) := \nabla f(x) + \sum_{i=1}^{N} \lambda_i \nabla g_i(x) = 0 \\ \text{และ } \lambda_1, \ldots, \lambda_N \ge 0. \end{array}}$$

เมื่อ $\lambda = (\lambda_1, \ldots, \lambda_N)^T$

Ex: $x = \begin{bmatrix} w \\ \theta \end{bmatrix}$, $f(x) = f(w,\theta) = \frac{1}{2}\|w\|^2$

$$g_i(x) = g_i(w,\theta) = 1 - y_i(\langle w, x_i \rangle - \theta)$$

เมื่อ $x_i \in X = X_+ \cup X_-$



$X_+ = \{x_3\}$

$X_- = \{x_1, x_2\}$

$X = \{x_1, x_2, x_3\}$

$$\text{Maximize } F(\underset{(w_1,w_2)}{w}, \underset{\theta}{\theta}, \underset{(\lambda_1,\lambda_2,\lambda_3)}{\lambda}) := \frac{1}{2}\|w\|^2 + \sum_{i=1}^{3} \lambda_i g_i(w,\theta)$$

เมื่อ
$$\begin{aligned} g_1(w,\theta) &= 1 - y_1 \langle w, x_1 \rangle - \theta) \\ &= 1 - (-1)\left(\left\langle \begin{pmatrix} w_1 \\ w_2 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \end{pmatrix} \right\rangle - \theta\right) \\ &= 1 + (2w_1 + w_2 - \theta) \end{aligned}$$

$$g_2(\omega,\theta) = 1 - y_2 \left(\langle \omega, x_2 \rangle - \theta\right)$$

$$= 1 - (-1) \left(\left\langle \begin{pmatrix} \omega_1 \\ \omega_2 \end{pmatrix}, \begin{pmatrix} 2 \\ -1 \end{pmatrix} \right\rangle - \theta\right)$$

$$= 1 + \left(2\omega_1 - \omega_2 - \theta\right)$$

$$g_3(\omega,\theta) = 1 - y_3 \left(\langle \omega, x_3 \rangle - \theta\right)$$

$$= 1 - (+1) \left(\left\langle \begin{pmatrix} \omega_1 \\ \omega_2 \end{pmatrix}, \begin{pmatrix} 4 \\ 0 \end{pmatrix} \right\rangle - \theta\right)$$

$$= 1 - \left(4\omega_1 - \theta\right)$$

$$\Rightarrow F(\omega,\theta,\lambda) = \frac{1}{2}\left(\omega_1^2 + \omega_2^2\right) \leftarrow f(\omega,\theta) = \frac{1}{2}\|\omega\|^2$$

$$+ \lambda_1\left(1 + \left(2\omega_1 + \omega_2 - \theta\right)\right) \leftarrow g_1(\omega,\theta)$$

$$+ \lambda_2\left(1 + \left(2\omega_1 - \omega_2 - \theta\right)\right) \leftarrow g_2(\omega,\theta)$$

$$+ \lambda_3\left(1 - \left(4\omega_1 - \theta\right)\right) \leftarrow g_3(\omega,\theta)$$

$$\Rightarrow G(\omega,\theta,\lambda) = \nabla f(\omega,\theta) + \sum_{i=1}^{3} \lambda_i \nabla g_i(\omega,\theta) = 0$$

Dazu $\nabla f = \begin{pmatrix} \partial\omega_1 \\ \partial\omega_2 \\ \partial\theta \end{pmatrix} f(\omega,\theta) = \quad \leftarrow \frac{1}{2}\|\omega\|^2 = \frac{1}{2}\left(\omega_1^2 + \omega_2^2\right)$

$$= \begin{pmatrix} \omega_1 \\ \omega_2 \\ 0 \end{pmatrix}$$

$$\nabla g_1 = \begin{pmatrix} \partial\omega_1 \\ \partial\omega_2 \\ \partial\theta \end{pmatrix} \left(1 + \left(2\omega_1 + \omega_2 - \theta\right)\right) = \begin{pmatrix} 2 \\ 1 \\ -1 \end{pmatrix}$$

$\nabla g_2 =$ - -

- - - - .

หรือ สามารถเขียนได้ในรูป.

$$x = \begin{pmatrix} w \\ \theta \end{pmatrix}, \quad f(w,\theta) = \frac{1}{2}\|w\|^2, \quad g_i(w,\theta) = 1 - y_i(\langle w, x_i\rangle - \theta)$$

$$\Rightarrow \nabla f(w,\theta) = \begin{pmatrix} w_1 \\ w_2 \\ 0 \end{pmatrix} = \begin{pmatrix} w \\ 0 \end{pmatrix}$$

$$(w_1 x_{i1} + w_2 x_{i2} - \theta)$$

$$\Rightarrow \nabla g_i(w,\theta) = 1 - y_i(\langle w, x_i\rangle - \theta)$$

$$\begin{pmatrix} x_{i1} \\ x_{i2} \end{pmatrix}$$

$$= \begin{pmatrix} -y_i x_{i1} \\ -y_i x_{i2} \\ y_i \end{pmatrix} = \begin{pmatrix} -y_i x_i \\ y_i \end{pmatrix}$$

จะได้ว่า $G(w,\theta,\lambda) = 0$ เขียนได้.

$$G(w,\theta,\lambda) = \begin{pmatrix} w \\ 0 \end{pmatrix} + \lambda_1 \begin{pmatrix} -y_1 x_1 \\ y_1 \end{pmatrix} + \lambda_2 \begin{pmatrix} -y_2 x_2 \\ y_2 \end{pmatrix} + \lambda_3 \begin{pmatrix} -y_3 x_3 \\ y_3 \end{pmatrix}$$

$$= \begin{pmatrix} w - \sum_{i=1}^{3} \lambda_i y_i x_i \\ 0 + \sum_{i=1}^{3} \lambda_i y_i \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

ดังนั้น. $\quad w = \sum_{i=1}^{3} \lambda_i y_i x_i, \quad \sum_{i=1}^{3} \lambda_i y_i = 0$

นำไปแทนค่าใน $F(w,\theta,\lambda)$

นั่นคือ $\quad F(w,\theta,\lambda) = \frac{1}{2}\|w\|^2 + \sum_{i=1}^{3} \lambda_i (1 - y_i(\langle w, x_i\rangle - \theta))$

$$= \frac{1}{2}\|w\|^2 + \sum_{i=1}^{3} \lambda_i - \sum_{i=1}^{3} \lambda_i y_i \langle w, x_i\rangle + \underbrace{\left(\sum_{i=1}^{3} \lambda_i y_i\right)}_{=0} \theta$$

$$= \frac{1}{2}\|w\|^2 + \sum_{i=1}^{3} \lambda_i - \langle w, \underbrace{\sum_{i=1}^{3} \lambda_i y_i x_i}_{=w} \rangle$$

$$\boxed{F(w,\theta,\lambda) = -\frac{1}{2}\|w\|^2 + \sum_{i=1}^{3} \lambda_i}$$

เมื่อแทน $w = \sum_{i=1}^{3} \lambda_i y_i x_i$ จะได้

$$F(w,\theta,\lambda) = -\frac{1}{2}\sum_{i,j=1}^{3} \lambda_i \lambda_j y_i y_j \langle x_i, x_j \rangle + \sum_{i=1}^{3} \lambda_i =: \tilde{F}(\lambda)$$

สังเกต. เราสามารถใช้ Kuhn-Tucker thm เขียนปัญหา ของ SVM ได้ดังนี้

Thm! เราสามารถหา w^* ที่เป็นคำตอบของปัญหา.

$$\begin{cases} \text{Minimize} & \frac{1}{2}\|w\|^2 \\ \text{subject to} & y_i(\langle w, x_i \rangle - \theta) \geq 1 \quad , \quad i = 1, \dots, N. \end{cases}$$

โดยที่ $y_i = \pm 1$, $x_i \in X_{\pm}$

ได้ในรูปของ.

$$w^* = \sum_{i=1}^{N} \lambda_i^* y_i x_i .$$

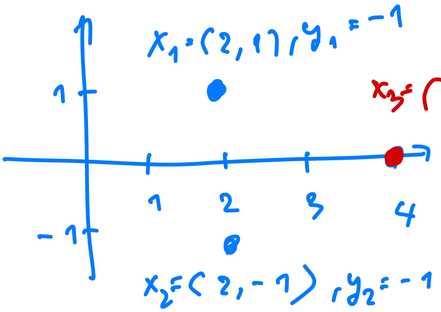
โดยที่ $\lambda^* = (\lambda_1^*, \lambda_2^*, \dots, \lambda_N^*)$ เป็นคำตอบของ

$$\begin{cases} \text{Maximize} & \tilde{F}(\lambda) = -\frac{1}{2}\sum_{i,j=1}^{N} \lambda_i \lambda_j y_i y_j \langle x_i, x_j \rangle + \sum_{i=1}^{N} \lambda_i \\ \text{subject to} & \sum_{i=1}^{N} \lambda_i y_i = 0 \quad \text{and} \quad \lambda_1, \dots, \lambda_N \geq 0 . \end{cases}$$

และ จะสามารถหา θ^* ได้เช่นกัน

$$\theta^* = \frac{1}{2}\left(\min_{x \in X_+} \sum_{i=1}^{N} \lambda_i^* y_i \langle x_i, x \rangle + \max_{y \in X_-} \sum_{i=1}^{N} \lambda_i^* y_i \langle x_i, x \rangle \right)$$

נשים לב שרק עבור x_i ש- $\lambda_i^* \neq 0$ הם נקראים support vectors.



פונק' המטרה

$$\underset{\nearrow\ (\lambda_1,\lambda_2,\lambda_3)}{\tilde{F}(\lambda)} = -\frac{1}{2} \sum_{i,j=1}^{3} \lambda_i \lambda_j y_i y_j \langle x_i, x_j \rangle + \sum_{i=1}^{3} \lambda_i$$

$$\tilde{F}(\lambda) = -\frac{1}{2} \Bigg[\lambda_1^2 (-1)^2 \overbrace{\left\langle \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \end{pmatrix} \right\rangle}^{=5} + \lambda_1 \lambda_2 (-1)(-1) \overbrace{\left\langle \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ -1 \end{pmatrix} \right\rangle}^{3} + \lambda_1 \lambda_3 (-1)(+1) \overbrace{\left\langle \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 4 \\ 0 \end{pmatrix} \right\rangle}^{8}$$

$$+ \lambda_2 \lambda_1 (-1)(-1) \overbrace{\left\langle \begin{pmatrix} 2 \\ -1 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \end{pmatrix} \right\rangle}^{3} + \lambda_2^2 (-1)^2 \overbrace{\left\langle \begin{pmatrix} 2 \\ -1 \end{pmatrix}, \begin{pmatrix} 2 \\ -1 \end{pmatrix} \right\rangle}^{5} + \lambda_2 \lambda_3 (-1)(1) \overbrace{\left\langle \begin{pmatrix} 2 \\ -1 \end{pmatrix}, \begin{pmatrix} 4 \\ 0 \end{pmatrix} \right\rangle}^{8}$$

$$+ \lambda_3 \lambda_1 (1)(-1) \overbrace{\left\langle \begin{pmatrix} 4 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \end{pmatrix} \right\rangle}^{=8} + \lambda_3 \lambda_2 (+1)(-1) \cdot \overbrace{\left\langle \begin{pmatrix} 4 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ -1 \end{pmatrix} \right\rangle}^{=8}$$

$$+ \lambda_3^2 (1)^2 \overbrace{\left\langle \begin{pmatrix} 4 \\ 0 \end{pmatrix}, \begin{pmatrix} 4 \\ 0 \end{pmatrix} \right\rangle}^{=16} \Bigg] + (\lambda_1 + \lambda_2 + \lambda_3)$$

$$\tilde{F}(\lambda) = -\frac{1}{2}\left[5\lambda_1^2 + 5\lambda_2^2 + 16\lambda_3^2 + 6\lambda_1\lambda_2 - 16\lambda_1\lambda_3 - 16\lambda_2\lambda_3\right]$$

$$+ \lambda_1 + \lambda_2 + \lambda_3$$

ภายใต้เงื่อนไข $\sum_{i=1}^{3} \lambda_i y_i = 0 \Rightarrow -\lambda_1 - \lambda_2 + \lambda_3 = 0$

$\Rightarrow \lambda_3 = \lambda_1 + \lambda_2$ แทนค่า

จะได้ว่า

$$\tilde{F}(\lambda) = -\frac{1}{2}\left[5\lambda_1^2 + 5\lambda_2^2 + 16\cancel{(\lambda_1+\lambda_2)}^2 + 6\lambda_1\lambda_2 - 16\cancel{(\lambda_1+\lambda_2)}^2\right]$$

$$+ 2(\lambda_1+\lambda_2)$$

พิจารณา $\nabla \tilde{F}(\lambda) = \begin{pmatrix} \partial\lambda_1 \\ \partial\lambda_2 \end{pmatrix} \tilde{F}(\lambda) = \begin{pmatrix} -\frac{1}{2}(10\lambda_1 + 6\lambda_2) + 2 \\ -\frac{1}{2}(10\lambda_2 + 6\lambda_1) + 2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$

จะได้ระบบสมการ

$$-5\lambda_1 - 3\lambda_2 + 2 = 0 \quad \text{—— } ①$$

$$-5\lambda_2 - 3\lambda_1 + 2 = 0 \quad \text{—— } ②$$

$$\left.\begin{array}{l} ① \times 3 \Rightarrow -15\lambda_1 - 9\lambda_2 + 6 = 0 \\ ② \times 5 \Rightarrow -25\lambda_2 - 15\lambda_1 + 10 = 0 \end{array}\right\} \Rightarrow 16\lambda_2 - 4 = 0$$

$$\Rightarrow \lambda_2 = \frac{4}{16} = \frac{1}{4}$$

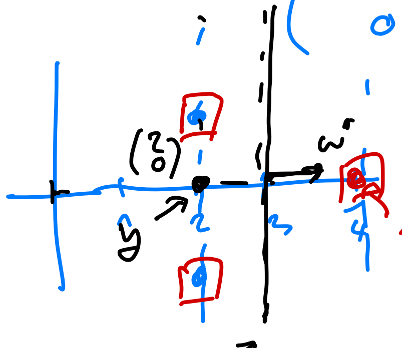
จาก ① $\Rightarrow \lambda_1 = \frac{2 - 5\lambda_2}{3} = \frac{2 - 5\left(\frac{1}{4}\right)}{3} = \frac{1}{4} \Rightarrow \lambda_1 = \frac{1}{4}$

จาก $\lambda_3 = \lambda_1 + \lambda_2 = \frac{2}{4}$

$\therefore$ จะได้ $\lambda_1 = \lambda_2 = \frac{1}{4}$, $\lambda_3 = \frac{2}{4}$

ดังนั้น $\omega^* = \sum_{i=1}^{3} \lambda_i y_i x_i = \frac{1}{4}(-1)\begin{pmatrix} 2 \\ 1 \end{pmatrix} + \frac{1}{4}(-1)\begin{pmatrix} 2 \\ -1 \end{pmatrix} + \frac{2}{4}(1)\begin{pmatrix} 4 \\ 0 \end{pmatrix}$

$$= \begin{pmatrix}-1+2\\0\end{pmatrix} = \begin{pmatrix}1\\0\end{pmatrix}$$



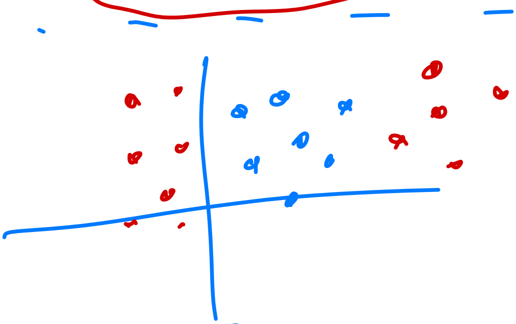
$w = \begin{pmatrix}1\\0\end{pmatrix}, \theta^* = 3$

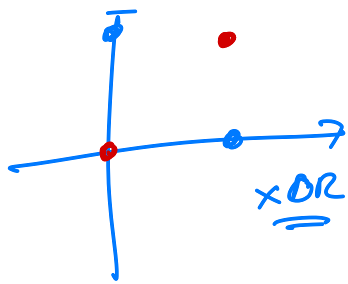
หา $\theta^* = \frac{1}{2}\left(\min_{x \in X_+} \sum_{i=1}^{l} \lambda_i^* y_i \langle x_i, x \rangle + \max_{y \in X_-} \sum \lambda_i^* y_i \langle x_i, y \rangle \right)$

$$= \frac{1}{2}\left(\underbrace{\left(\frac{2}{4}\right)\cdot(+1)\left\langle \begin{pmatrix}4\\0\end{pmatrix}, \begin{pmatrix}4\\0\end{pmatrix} \right\rangle}_{1} + \underbrace{\left(\frac{1}{4}\right)(-1)\left\langle \begin{pmatrix}2\\1\end{pmatrix}, \begin{pmatrix}2\\0\end{pmatrix} \right\rangle + \left(\frac{1}{4}\right)(-1)\left\langle \begin{pmatrix}2\\-1\end{pmatrix}, \begin{pmatrix}2\\0\end{pmatrix} \right\rangle}_{2} \right)$$

ไม่ต้องใช้สูตรนี้ก็ได้ครับ !

$$= \frac{1}{2}\left(\frac{\cancel{16}^{\,8}}{\cancel{2}} + (-1) + (-1) \right) = \frac{6}{2} = 3$$





ตัวอย่าง: จงใช้กระบวนการตามทฤษฎีบท (SVM) หา w^* และ θ^* ของจุด x_1, x_2, x_3 ใน $\mathbb{R}^2$.

โดยที่: $x_1 = \begin{pmatrix}2\\1\end{pmatrix}, \; x_2 = \begin{pmatrix}3\\1\end{pmatrix}, \; x_3 = \begin{pmatrix}4\\0\end{pmatrix}$

ข้อ 7. (แก้ไขส่วนที่ผิดของการคำนวณ θ^*)

$$\theta^* = \frac{1}{2}\left(\underbrace{\min_{x\in X_+}\sum_{i=1}^{N}\lambda_i^* y_i\langle x_i, x\rangle}_{①} + \underbrace{\max_{y\in X_-}\sum_{i=1}^{N}\lambda_i^* y_i\langle x_i, x\rangle}_{②}\right)$$

①: จาก $X_+ = \left\{\begin{pmatrix}4\\0\end{pmatrix}\right\}$, $x_1 = \begin{pmatrix}2\\1\end{pmatrix}$, $x_2 = \begin{pmatrix}2\\-1\end{pmatrix}$, $x_3 = \begin{pmatrix}4\\0\end{pmatrix}$

$\lambda_1^* = \lambda_2^* = \frac{1}{4}$, $\lambda_3^* = \frac{2}{4}$

จะได้

$$\min_{x\in X_+}\left\{\frac{1}{4}(-1)\left\langle\begin{pmatrix}2\\1\end{pmatrix},\begin{pmatrix}4\\0\end{pmatrix}\right\rangle + \frac{1}{4}(-1)\left\langle\begin{pmatrix}2\\-1\end{pmatrix},\begin{pmatrix}4\\0\end{pmatrix}\right\rangle + \frac{2}{4}(1)\left\langle\begin{pmatrix}4\\0\end{pmatrix},\begin{pmatrix}4\\0\end{pmatrix}\right\rangle\right\}$$

$$= \min_{x\in X_+}\left\{\frac{-8}{4} + \frac{-8}{4} + \frac{2\cdot 16}{4}\right\} = 4$$

②: จาก $X_- = \left\{\begin{pmatrix}2\\1\end{pmatrix},\begin{pmatrix}2\\-1\end{pmatrix}\right\}$

จะได้

$$\max_{x\in X_-}\left\{\left[\sum_{i=1}^{3}\lambda_i^* y_i\left\langle x_i,\begin{pmatrix}2\\1\end{pmatrix}\right\rangle\right], \left[\sum_{i=1}^{3}\lambda_i^* y_i\left\langle x_i,\begin{pmatrix}2\\-1\end{pmatrix}\right\rangle\right]\right\}$$

$$= \max\left\{\left[\frac{-5}{4} + \frac{-3}{4} + \frac{2\cdot 8}{4}\right], \left[\frac{-3}{4} + \frac{-5}{4} + \frac{2\cdot 8}{4}\right]\right\}$$

$$= \max\left\{ \frac{8}{4}, \frac{8}{4} \right\} = 2.$$

ดังนั้น
หาได้ $\theta^* = \frac{1}{2}\left(4 + 2\right) = 3$ $\blacksquare$