Optimal separation (Cont.):

ทบ: (optimal separation margin).

สมมติ $X_+, X_- \subset \mathbb{R}^n$ เป็น non-empty finite sets

ที่ $\mathrm{conv}(X_+) \cap \mathrm{conv}(X_-) = \emptyset$. เราเรียกว่า optimal separation margin

$$\rho := \sup\{\rho(\omega,\theta) : (\omega,\theta) \in \mathbb{R}^{n+1},\ \omega \neq \vec{0}\}$$

มีค่าจำกัด ($<\infty$) ที่บรรลุที่ (ω^*, θ^*) และ

$$\rho = \frac{1}{2}\,\mathrm{dist}(C_+, C_-) = \frac{1}{2}\,\mathrm{dist}(C, 0)$$

โดยที่ $C_\pm = \mathrm{conv}(X_\pm)$, $C = \mathrm{conv}(X_+ - X_-)$ 1

และไฮเปอร์ $H = \{x \in \mathbb{R}^n : \langle \omega^*, x\rangle - \theta^* = 0\}$

เป็น unique และเรียกว่าเป็น optimal hyperplane. 2

บทตั้ง 1: ให้ $\emptyset \neq X_+, X_- \subset \mathbb{R}^n$ เป็น finite sets ที่

$C_+ \cap C_- = \emptyset$ และ $C := \mathrm{conv}(X_+ - X_-) = C_+ - C_-$

และ $0 \notin C$ จะได้ว่า

$$\mathrm{dist}(C_+, C_-) = \mathrm{dist}(C, 0) = \max_{\|v\|=1} \min_{z \in C} \langle v, z\rangle$$

บทตั้ง 2: ให้ $X_+, X_- \subset \mathbb{R}^n$ มีสมบัติเช่นเดียวกับ ทบ.

ให้ Γ เป็น weight polyhedron นั่นคือ

$$\Gamma = \{(\omega,\theta) \in \mathbb{R}^{n+1} : y_i(\langle \omega, x_i\rangle - \theta) > 0,\ i = 1, \dots, N\}$$

พิจารณา proj. บน $\mathbb{R}^n$.

$$W = \{ w \in \mathbb{R}^n : \exists \theta \in \mathbb{R} : (w,\theta) \in \Gamma \}$$

สามารถเขียนในรูป

$$W = \{ w \in \mathbb{R}^n : \langle w, z \rangle > 0 \quad \forall z \in C \}$$

จะได้ว่า ถ้าเลือก $w \in W$ แล้ว

$$\theta_w := \frac{1}{2}\left(\min_{x \in C_+} \langle w, x \rangle + \max_{y \in C_-} \langle w, y \rangle \right)$$

ซึ่ง $(w, \theta_w) \in \Gamma$ นอกจากนี้จะได้ว่า ถ้าเลือก $w \neq \vec{0}$

แล้ว $$\rho(w, \theta_w) = \frac{1}{2} \min_{z \in C} \left\langle \frac{w}{\|w\|}, z \right\rangle$$

<u>ทฤษฎีบท 1</u>: (สำคัญที่สุดในบท.)

สมมติ $\emptyset \neq X_+, X_- \subset \mathbb{R}^n$ เป็น finite sets ที่

$\mathrm{conv}(X_+) \cap \mathrm{conv}(X_-) = \emptyset$ แล้วจะได้ว่า optimal sep-margin

$$\rho := \sup \{ \rho(w,\theta) : (w,\theta) \in \mathbb{R}^{n+1}, w \neq 0 \}$$

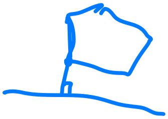
มีค่าจำกัด $(< \infty)$ และ

$$\rho = \frac{1}{2} \mathrm{dist}(C_+, C_-) = \frac{1}{2} \mathrm{dist}(C, 0).$$

โดยที่ $C_\pm := \mathrm{conv}(X_\pm)$, $C := \mathrm{conv}(X_+ - X_-) = C_+ - C_-$

<u>proof</u>: เลือก $w \neq 0 \in \mathbb{R}^n$ จะได้ว่า

$$\rho(w,\theta)\|w\| = \min\left(\min_{x \in X_+} (\langle w, x \rangle - \theta), \min_{y \in X_-} (-\langle w, y \rangle + \theta) \right)$$



$$= \min\left(\min_{x\in C_+} (\langle w,x\rangle - \theta),\ \min_{y\in C_-} (-\langle w,y\rangle + \theta) \right)$$

(จากสมบัติของ convex set $\langle w,\cdot\rangle$ จะมี min ที่จุด มุม.)

และ สำหรับ $a,b \in \mathbb{R}$ ทราบว่า $\min(a,b) \le \frac{1}{2}(a+b)$

จะเท่ากันก็ต่อเมื่อ $a=b$.

ดังนั้น. $$\rho(w,\theta)\|w\| \le \frac{1}{2} \min_{x\in C_+,\, y\in C_-} \langle w, x-y\rangle$$

ดังนั้น. $$\rho(w,\theta) \le \frac{1}{2} \min_{z\in C} \left\langle \frac{w}{\|w\|}, z\right\rangle.$$

จาก บทตั้ง 2 ทราบว่า เงื่อนไขนี้จะเท่ากัน ก็ต่อเมื่อ $\theta = \theta_w$

ดังนั้น. (*)

$$\sup_{\theta\in\mathbb{R}} \rho(w,\theta) = \rho(w,\theta_w) = \frac{1}{2} \min_{z\in C} \left\langle \frac{w}{\|w\|}, z\right\rangle.$$

ทราบว่า

$$\rho = \sup\{\rho(w,\theta) : (w,\theta)\in\mathbb{R}^{n+1}, w\ne 0\}$$

$$= \sup_{w\in\mathbb{R}^n} \sup_{\theta\in\mathbb{R}} \rho(w,\theta)$$

(*) $$= \sup_{w\in\mathbb{R}^n} \left(\frac{1}{2} \min_{z\in C} \left\langle \frac{w}{\|w\|}, z\right\rangle \right)$$

$$= \frac{1}{2} \max_{\|v\|=1} \min_{z\in C} \langle v, z\rangle$$

บทตั้ง 1 $$= \frac{1}{2}\operatorname{dist}(C,0) = \frac{1}{2}\operatorname{dist}(C_+, C_-)$$ ■

อีกส่วนหนึ่ง ต้องพิสูจน์ว่า มี (w^*, θ^*) เพียงค่าเดียว ที่ทำให้

$$\rho = \rho(w^*, \theta^*) = \frac{1}{2}\operatorname{dist}(C, 0) = \frac{1}{2}\operatorname{dist}(C_+, C_-)$$

และจะมี optimal separation $H(w^*, \theta^*)$ เพียงเส้นเดียว. ②

<u>บทตั้ง 3:</u> ให้ $\emptyset \neq C \subset \mathbb{R}^n$ เป็น closed & convex set

ที่ $0 \in C$ แล้วจะมี $z^* \in C$ เพียงค่าเดียว. ที่ทำให้

(unique & exist)

$$\|z^*\| = \min_{z \in C} \|z\| = \operatorname{dist}(C, 0)$$

ให้ $C^{\#} = \{ w \in \mathbb{R}^n : \langle w, z \rangle \geq c \quad \forall z \in C \}$.

โดยที่ $c > 0$ แล้วว่า $C^{\#} \neq \emptyset$ และเป็น closed & convex set

ที่ $0 \notin C^{\#}$ แล้วจะมี $w^* \in C^{\#}$ ที่

$$\|w^*\| = \min_{w \in C^{\#}} \|w\| = \operatorname{dist}(C^{\#}, 0)$$

และ $w^* = \dfrac{c z^*}{\|z^*\|^2}$, $z^* = \dfrac{c w^*}{\|w^*\|^2}$

<u>Proof:</u> ให้ $u \in C$ เป็นสมาชิกใดๆ ให้

$$B := \{ x \in \mathbb{R}^n : \|x\| \leq \|u\| \} \quad \text{(compact.)}$$

จะได้ว่า

$$\inf_{z \in C} \|z\| = \inf_{z \in B \cap C} \|z\|$$

เนื่องจาก B compact ดังนั้น $\exists z^* \in C$ ที่ทำให้

$$\|z^*\| = \inf_{z\in C} \|z\| \qquad \text{(existence.)}$$

ต่อไปพิสูจน์ uniqueness สมมติ z_1^*, z_2^* ที่

$$\|z_1^*\| = \|z_2^*\| = \inf_{z\in C}\|z\|$$

จากสมบัติ convex ของ C จะได้ว่าจุด $\frac{1}{2}(z_1^*+z_2^*) \in C$.

ถ้า $z_1^* \neq z_2^*$ แล้ว

$(z_1^* \neq z_2^*)$

$$\left\|\frac{1}{2}(z_1^*+z_2^*)\right\| < \frac{1}{2}\|z_1^*\| + \|z_2^*\|$$

$$= \frac{1}{2}\left(2\inf_{z\in C}\|z\|\right)$$

$$= \inf_{z\in C}\|z\| \quad \lightning$$

ดังนั้น $z_1^* = z_2^* = z^*$.

จากบทพิสูจน์ของบทตั้ง 1: จะได้ว่า

$$\|z^*\| \leq \left\langle \frac{z^*}{\|z^*\|}, z \right\rangle \quad \forall z \in C.$$

คูณ $c > 0$ ทั้งสองข้าง จะได้

$$\left\langle \frac{cz^*}{\|z^*\|^2}, z \right\rangle \geq c \quad \forall z \in C.$$

ทำให้ได้ว่า $C^\# \neq \emptyset$ เนื่องจาก.

$$z^* \in C^\# = \{w \in \mathbb{R}^n : \langle w, z\rangle \geq c \quad \forall z \in C\}$$

เพื่อแสดงว่า w^* unique จะพิจารณา $w \in C^\#$ และ $z \in C$

พิสูจน์. (Cauchy Schwarz)

$$\|w\|\|z\| \geq \langle w, z\rangle \geq c$$

ดังนั้นสำหรับ z^* พิสูจน์

$$\|w\| \geq c\|z^*\|^{-1} \quad \text{สำหรับทุก} \quad w \in C^\#$$

ถ้าให้ $w^* := \dfrac{c z^*}{\|z^*\|^2}$ พิสูจน์

$$\|w^*\| = c\|z^*\|^{-1} \quad \text{จาก. } \|z^*\| = \inf_{z \in C} \|z\|$$

พิสูจน์ว่า $\|w^*\|$ เป็น global minimum ■

และ unique เพราะว่า z^* unique.

<u>บทแทรก 2</u>: ให้ $\emptyset \neq X_+, X_- \subset \mathbb{R}^n$ เป็น finite sets และ $C = \mathrm{conv}(X_+ - X_-)$ ที่ $0 \notin C$ ให้ $C^\#$ เป็น

$$C^\# = \{w \in \mathbb{R}^n : \langle w, z\rangle \geq c \quad \forall z \in C\}$$

และให้ w^* เป็นตัวเลขที่ unique ของ

$$\|w^*\| = \min_{w \in C^\#} \|w\|$$

ให้

$$\theta_{w^*} := \frac{1}{2}\left(\min_{x \in C_+} \langle w^*, x\rangle + \max_{y \in C_-} \langle w^*, y\rangle\right)$$

แล้ว.

$$H^* = \{x \in \mathbb{R}^n : \langle w^*, x\rangle - \theta_{w^*} = 0\}$$

เป็น unique optimal separating hyperplane.

proof: จากบทตั้งที่ 2 จะได้ว่า

$(w^*, \theta_{w^*}) \in \Gamma$ จะได้ว่า H^* เป็น separating hyperplane

และเป็น optimal sep. hyp. เนื่องจาก

$$\langle w^*, z \rangle \geq c \qquad \forall z \in C$$

และจะได้ว่า

(*)

$$\langle w^*, z^* \rangle = c, \quad z^* = \frac{c\, w^*}{\|w^*\|^2} \qquad \left(\Rightarrow \|z^*\| = \frac{c\, \cancel{\|w^*\|}}{\|w^*\|^{\cancel{2}}} \right)$$

ดังนั้น

$$\rho(w^*, \theta_{w^*}) = \frac{1}{2}\left(\min_{z \in C} \left\langle \frac{w^*}{\|w^*\|}, z \right\rangle \right) = \frac{1}{2} \left\langle \frac{w^*}{\|w^*\|}, z^* \right\rangle$$

$$= \frac{1}{2} \left\langle \frac{w^*}{\|w^*\|}, \frac{c\, w^*}{\|w^*\|^2} \right\rangle = \frac{1\, c}{2\|w^*\|} \underbrace{\left\langle \frac{w^*}{\|w^*\|}, \frac{w^*}{\|w^*\|} \right\rangle}_{=1}$$

$$= \frac{1}{2} \|z^*\| = \frac{1}{2} \operatorname{dist}(C, 0)$$

สำหรับ uniqueness ให้ (w, θ) เป็นค่า สำหรับ opt. sep. hyp. ซึ่งคุณสมบัติ ทำให้ $\rho(w, \theta) = \rho$

จะได้ว่า

$$\rho(w, \theta) \leq \rho(w, \theta_w) \qquad \forall \theta \in \mathbb{R}$$

จะเท่ากัน ก็ต่อเมื่อ $\theta = \theta_w$ ดังนั้นจะได้ว่า

(**)

$$\rho(w, \theta_w) = \frac{1}{2} \min_{z \in C} \left\langle \frac{w}{\|w\|}, z \right\rangle = \rho = \frac{1}{2} \min_{z \in C} \|z\| = \frac{1}{2} \|z^*\|$$

ดังนั้น. (**)

$$\|z^*\| \le \left\langle \frac{w}{\|w\|}, z^* \right\rangle \le \underbrace{\left\|\frac{w}{\|w\|}\right\|}_{=1} \|z^*\| \le \|z^*\|$$

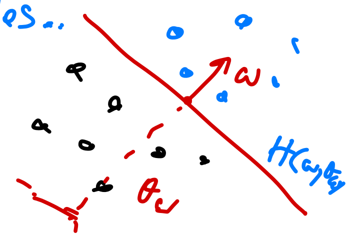
แสดงว่า $\left\langle \frac{w}{\|w\|}, z^* \right\rangle = \|z^*\|$

จะเกิดขึ้นได้ก็ต่อเมื่อ $w = \alpha z^* \Rightarrow \frac{w}{\|w\|} = \frac{\cancel{\alpha} z^*}{\|\cancel{\alpha} z^*\|}$

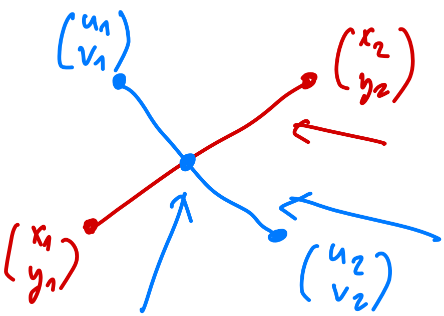
$$\left(\Rightarrow \left\langle \frac{z^*}{\|z^*\|}, z^* \right\rangle = \frac{\langle z^*, z^* \rangle}{\|z^*\|} = \frac{\|z^*\|^{\cancel{2}}}{\cancel{\|z^*\|}} = \|z^*\| \right)$$

ซึ่งพบ $\alpha > 0$.

ดังนั้น (w, θ) ที่ให้ค่า optimal separation margin จะ unique up to scalar positive multiples..



หาจุดตัดของส่วนของเส้นตรง:



$$r_1(t): \binom{x_1}{y_1} + \left(\binom{x_2}{y_2} - \binom{x_1}{y_1} \right) t$$

$$r_2(t): \binom{u_1}{v_1} + \left(\binom{u_2}{v_2} - \binom{u_1}{v_1} \right) t$$

จุดตัดอยู่บน $r_1(t)$ และ $r_2(t)$ ที่ $t_1, t_2 \in [0,1]$

ที่จุด $r_1(t_1) = r_2(t_2)$

ปัญหา

$$\begin{pmatrix} x_1 \\ y_1 \end{pmatrix} + \left(\begin{pmatrix} x_2 \\ y_2 \end{pmatrix} - \begin{pmatrix} x_1 \\ y_1 \end{pmatrix} \right) t_1 = \begin{pmatrix} u_1 \\ v_1 \end{pmatrix} + \left(\begin{pmatrix} u_2 \\ v_2 \end{pmatrix} - \begin{pmatrix} u_1 \\ v_1 \end{pmatrix} \right) t_2$$

จัดรูป.

$\Rightarrow$

$$(x_2 - x_1) t_1 - (u_2 - u_1) t_2 = u_1 - x_1$$

$$(y_2 - y_1) t_1 - (v_2 - v_1) t_2 = v_1 - y_1$$

$$\Rightarrow \quad \begin{bmatrix} (x_2 - x_1) & -(u_2 - u_1) \\ (y_2 - y_1) & -(v_2 - v_1) \end{bmatrix} \begin{pmatrix} t_1 \\ t_2 \end{pmatrix} = \begin{pmatrix} u_1 - x_1 \\ v_1 - y_1 \end{pmatrix}$$

หาคำตอบ ถ้ามีคำตอบ ก็ต่อเมื่อ $t_1, t_2 \in [0, 1]$ ■