Geometric Hahn-Banach theorem (Hahn-Banach separation theorem)

2023年10月5日 13:50

- Geometric Hahn-Banach theorem

Def: Let D be a convex set in TVS (Topological Vector Space) $\mathfrak{X}$ containing 0.
then we define the Minkowski functional.

$$\mu_D(x) := \inf\{t>0 : x\in tD\},\ \forall x\in D.$$

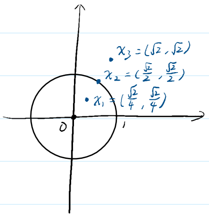
If $\{t>0 : x\in tD\}=\phi$, then we set $\mu_D(x)=\infty$.

The condition "$\{t>0 : x\in tD\}\neq\phi$" is related to the concept "Absorbing".

To exclude the case where $\{t>0 : x\in tD\}$, we can assume D is open

Intuitively, "Minkowski functional" $\approx$ "distance".

Consider $\mathfrak{X}=\mathbb{R}^2$, $D=\{(x,y) : x^2+y^2<1\}$, then



$$\mu_D(x_1)=\|x_1\|=\tfrac{1}{2} \iff \mu_D(x)<1 \quad \text{if } x\in D$$
$$\mu_D(x_2)=\|x_2\|=1 \iff \mu_D(x)=1 \quad \text{if } x\in \partial D$$
$$\mu_D(x_3)=\|x_3\|=2 \iff \mu_D(x)>1 \quad \text{if } x\in \mathfrak{X}\setminus D$$

Indeed, Minkowski functional is used to construct the topology of LCS (Locally Convex Space) which becomes a LCTVS (Locally Convex Topological Space).

Lemma: μ_D is sublinear and $\{x : \mu_D(x)<1\}\subset D\subset\{x : \mu_D(x)\le 1\}$

Proof:

Clearly, $\mu_D : \mathfrak{X}\to[0,\infty]$ is a non-negative functional.

<u>Positively homogeneous:</u>

For $\alpha\ge 0$, $x\in\mathfrak{X}$,

$$\{t>0 : \alpha x\in tD\}=\{t>0 : x\in \tfrac{t}{\alpha}D\} \overset{t'=\frac{t}{\alpha}}{=\!=} \{\alpha t' : x\in t'D\}=\alpha\{t' : x\in t'D\}$$

which implies

$$\mu_D(\alpha x)=\alpha\,\mu_D(x).$$

<u>Subadditive:</u>

For $x, y\in\mathfrak{X}$, and arbitrary $\varepsilon>0$, there exist $s, t>0$ such that

$$\mu_D(x) < s < \mu_D(x)+\varepsilon$$
$$\mu_D(y) < t < \mu_D(y)+\varepsilon$$

Moreover, $x\in tD$ and $y\in sD$ which imply there exist $d_1, d_2\in D$ such that

$$x=td_1$$
$$y=sd_2$$

Then

$$\mu(x+y)=\mu\Big((s+t)\underbrace{\big(\tfrac{s}{s+t}d_1+\tfrac{t}{s+t}d_2\big)}_{\text{convex combination of } d_1 \text{ and } d_2}\Big)$$
$$\le s+t$$
$$< \mu_D(x)+\mu_D(y)+2\varepsilon$$

therefore

$$\mu_D(x+y)\le\mu_D(x)+\mu_D(y)$$

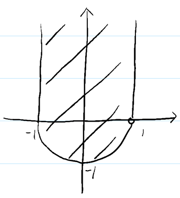
The following theorems are called the geometric Hahn-Banach theorem, also known as the Hahn-Banach separation theorem.

Thm: Let $\mathcal{U}$ be an open convex subset of TVS $\mathfrak{X}$ and $x_0 \in \mathfrak{X} \setminus \mathcal{U}$,
then there exists a hyperplane H containing x_0 such that H does not intersect with $\mathcal{U}$.

Proof:

WLOG, assume $0 \in \mathcal{U}$, otherwise we do the translation.
Then $x_0 \neq 0$. Consider
$$Y := \{\lambda x_0 : \lambda \in \mathbb{R}\}$$
Define linear functional on Y,
$$f(\lambda x_0) := \lambda$$
Note
$$\mu_{\mathcal{U}}(x_0) \geq 1.$$
Then
$$f(\lambda x_0) \leq \lambda \mu_{\mathcal{U}}(x_0) = \mu_{\mathcal{U}}(\lambda x_0)$$
Since $\mu_{\mathcal{U}}$ is sublinear, by Hahn-Banach theorem, there exists a linear functional on $\mathfrak{X}$,
$$\tilde{f}|_{\mathcal{U}} = f$$
$$\tilde{f}(x) \leq \mu_{\mathcal{U}}(x)$$
Define the hyperplane
$$H := \{x \in \tilde{f}(x) = 1\}$$
clearly, $x_0 \in H$. Moreover,
$$\tilde{f}(x) < 1, \ \forall x \in \mathcal{U} \quad (\text{If } \mathcal{U} \text{ is open convex, then } \tilde{f}(\mathcal{U}) \text{ is open})$$
therefore
$$H \cap \mathcal{U} = \phi.$$

The condition for $\mathcal{U}$ to be open is neccessary. Indeed, consider
$$\mathcal{U} := \{(x,y) : -1 \leq x \leq 1,\ y \geq -\sqrt{1-x^2}\} \setminus (1,0)\}$$
then there does not exist a straight line containing $(1,0)$ while not intersects $\mathcal{U}$.



The above result can be extended to the case where $\mathcal{U}$ is closed.

Thm: Let C be a closed convex subset of TVS $\mathfrak{X}$ and $x_0 \in \mathfrak{X} \setminus C$,
then there exists a hyperplane H containing x_0 does not intersect C.
Moreover, if $\mathfrak{X}$ is a normed space and C contains the origin,
for $0 < d < \operatorname{dist}(x_0, C)$ and $\|x_0\|_{\mathfrak{X}} \leq A$ with $A > d$, then there exists $\tilde{f} \in \mathfrak{X}^*$
such that
$$\tilde{f}(x) + \alpha < \tilde{f}(x_0), \ \forall x \in C.$$
where $\alpha = \frac{d}{2}\left[\left(1 - \frac{d}{A}\right)^{-1} - 1\right]$.

Furthermore, we can obtain the result for two disjoint sets.

Thm: Let V, W be disjoint open convex subsets of TVS $\mathfrak{X}$,
then there exis a hyperplane H separates V and W.

Thm: Let K be a compact convex subsets of TVS $\mathfrak{X}$ and C be a closed convex subset of TVS $\mathfrak{X}$,
then there exists a hyperplane H separates K and C.

Thm: Let K be a compact convex subsets of TVS $\mathfrak{X}$ and C be a closed convex subset of TVS $\mathfrak{X}$, then there exists a hyperplane H separates K and C.

In $\mathbb{R}^n$, for two disjoint convex set A and B, there exists $v \in \mathbb{R}^n$ and $c \in \mathbb{R}$ such that

$$(x, v) \geq c, \ \forall x \in A,$$
$$(x, v) \leq c, \ \forall x \in B.$$

Which means the hyperplane

$$H := \{x : (x, v) = c\}$$

separates A and B. This result is closely related to "SVM (Supported Vector Machine)" in Machine Learning.