2.4 Separation and Support of Sets

The notions of supporting hyperplanes and separation of disjoint convex sets are very important in optimization. Almost all optimality conditions and duality relationships use some sort of separation or support of convex sets. The results of this section are based on the following geometric fact: Given a closed convex set S and a point $\mathbf{y} \notin S$, there exists a unique point $\overline{\mathbf{x}} \in S$ with minimum distance from y and a hyperplane that separates y and S.

Minimum Distance from a Point to a Convex Set

To establish the above important result, the following parallelogram law is needed. Let $\mathbf{a}$ and $\mathbf{b}$ be two vectors in R^n. Then

$$\|a+b\|^2 = \|a\|^2 + \|b\|^2 + 2a^t b$$
$$\|a-b\|^2 = \|a\|^2 + \|b\|^2 - 2a^t b$$

By adding we get

$$\|a+b\|^2 + \|a-b\|^2 = 2\|a\|^2 + 2\|b\|^2.$$

$\|x\| = \sqrt{\langle x, x\rangle}$ より,

$$\begin{aligned}\|a+b\|^2 &= \langle a+b, a+b\rangle \\ &= \|a\|^2 + 2\langle a, b\rangle + \|b\|^2\end{aligned}$$

同様に,

$$\begin{aligned}\|a-b\|^2 &= \langle a-b, a-b\rangle \\ &= \|a\|^2 - 2\langle a, b\rangle + \|b\|^2\end{aligned}$$

This result is illustrated in Figure 2.6 and can be interpreted as follows: The sum of squared norms of the diagonals of a parallelogram is equal to the sum of squared norms of its sides.

We now state and prove the closest-point theorem. Again, the reader is encouraged to investigate how the various assumptions play a role in guaranteeing the various assertions.

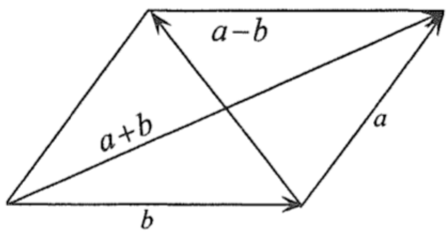

Figure 2.6 Parallelogram law.

2.4.1 Theorem

Let S be a nonempty, closed convex set in R^n and $\mathbf{y} \notin S$. Then there exists a unique point $\overline{\mathbf{x}} \in S$ with minimum distance from $\mathbf{y}$. Furthermore, $\overline{\mathbf{x}}$ is the minimizing point if and only if $(\mathbf{y} - \overline{\mathbf{x}})^t(\mathbf{x} - \overline{\mathbf{x}}) \leq 0$ for all

$\mathbf{x} \in S$.

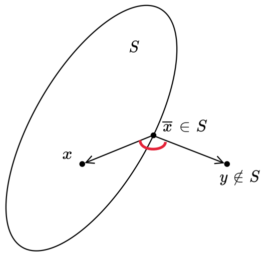

図 1　Theorem 2.4.1

Proof

First, let us establish the existence of a closest point. Since $S \neq \varnothing$, there exists a point $\hat{\mathbf{x}} \in S$, and we can confine our attention to the set $\bar{S} = S \cap \{\mathbf{x} : \|\mathbf{y} - \mathbf{x}\| \leq \|\mathbf{y} - \hat{\mathbf{x}}\|\}$ in seeking the closest point. In other words, the closest-point problem $\inf\{\|y - x\| : \mathbf{x} \in S\}$ is equivalent to $\inf\{\|\mathbf{y} - \mathbf{x}\| : \mathbf{x} \in \bar{S}\}$. But the latter problem involves finding the minimum of a continuous function over a nonempty, compact set $\bar{S}$, so by Weierstrass's theorem, Theorem 2.3.1, we know that there exists a minimizing point $\overline{\mathbf{x}}$ in S that is closest to the point y.

ノルムは連続，有限次元 $\mathbb{R}^n$ では有界閉集合 $\Leftrightarrow$ コンパクト集合

To show uniqueness, suppose that there is an $\overline{\mathbf{x}}' \in S$ such that $\|\mathbf{y} - \overline{\mathbf{x}}\| = \|\mathbf{y} - \overline{\mathbf{x}}'\| = \gamma$. By the convexity of $S, (\overline{\mathbf{x}} + \overline{\mathbf{x}}')/2 \in S$. By the triangle inequality we get

$$\left\|\mathbf{y} - \frac{\overline{\mathbf{x}} + \overline{\mathbf{x}}'}{2}\right\| \leq \frac{1}{2}\|\mathbf{y} - \overline{\mathbf{x}}\| + \frac{1}{2}\|\mathbf{y} - \overline{\mathbf{x}}'\| = \gamma.$$

If strict inequality holds, we have a contradiction to $\overline{\mathbf{x}}$ being the closest point to y. Therefore, equality holds, and we must have $\mathbf{y} - \overline{\mathbf{x}} = \lambda(\mathbf{y} - \overline{\mathbf{x}}')$ for some λ. Since $\|\mathbf{y} - \overline{\mathbf{x}}\| = \|\mathbf{y} - \overline{\mathbf{x}}'\| = \gamma$, we have $|\lambda| = 1$. Clearly, $\lambda \neq -1$, because otherwise, $\mathbf{y} = (\overline{\mathbf{x}} + \overline{\mathbf{x}}')/2 \in S$, contradicting the assumption that $\mathbf{y} \notin S$. So $\lambda = 1$, yielding $\overline{\mathbf{x}}' = \overline{\mathbf{x}}$, and uniqueness is established.

$\|\mathbf{y} - \frac{\overline{\mathbf{x}} + \overline{\mathbf{x}}'}{2}\| < \gamma$ が成り立つとすると，$\overline{\mathbf{x}}$ と $\overline{\mathbf{x}}'$ 以外に closest-point があることになるので，仮定に矛盾する．したがって，$\|\mathbf{y} - \frac{\overline{\mathbf{x}} + \overline{\mathbf{x}}'}{2}\| = \frac{1}{2}\|\mathbf{y} - \overline{\mathbf{x}}\| + \frac{1}{2}\|\mathbf{y} - \overline{\mathbf{x}}'\| = \gamma$ となるので，$\mathbf{y} - \overline{\mathbf{x}} = \lambda(\mathbf{y} - \overline{\mathbf{x}}')$ となる実数 λ が存在する．$\|\mathbf{y} - \overline{\mathbf{x}}\| = \|\mathbf{y} - \overline{\mathbf{x}}'\| = \gamma$ より，$|\lambda| = 1$ となり，$\lambda = 1$ と $\lambda = -1$ の場合が考えられる．そこで，$\lambda = -1$ とすると，$\mathbf{y} = (\overline{\mathbf{x}} + \overline{\mathbf{x}}')/2 \in S$ となり，$\mathbf{y} \notin S$ であることに矛盾する．よって $\lambda = 1$ となり，$\overline{\mathbf{x}}' = \overline{\mathbf{x}}$ が示せる．

ある $\overline{\mathbf{x}} \in S$ とすべての $\mathbf{x} \in S$ に対して，$(\mathbf{y} - \overline{\mathbf{x}})^t(\mathbf{x} - \overline{\mathbf{x}}) \leq 0$ が成り立つ $\Rightarrow$ $\overline{\mathbf{x}} \in S$ は最小点

To complete the proof, we need to show that $(\mathbf{y} - \overline{\mathbf{x}})^t(\mathbf{x} - \overline{\mathbf{x}}) \leq 0$ for all $\mathbf{x} \in S$ is both a necessary and a sufficient condition for $\bar{x}$ to be the point in S closest to $\mathbf{y}$. To prove sufficiency, let $\mathbf{x} \in S$. Then

$$\|\mathbf{y} - \mathbf{x}\|^2 = \|\mathbf{y} - \overline{\mathbf{x}} + \overline{\mathbf{x}} - \mathbf{x}\|^2 = \|\mathbf{y} - \overline{\mathbf{x}}\|^2 + \|\overline{\mathbf{x}} - \mathbf{x}\|^2 + 2(\overline{\mathbf{x}} - \mathbf{x})^t(\mathbf{y} - \overline{\mathbf{x}})$$

Since $\|\overline{\mathbf{x}} - \mathbf{x}\|^2 \geq 0$ and $(\overline{\mathbf{x}} - \mathbf{x})^t(\mathbf{y} - \overline{\mathbf{x}}) \geq 0$ by assumption, $\|\mathbf{y} - \mathbf{x}\|^2 \geq \|\mathbf{y} - \overline{\mathbf{x}}\|^2$ and $\overline{\mathbf{x}}$ is the minimizing point.

ある $\overline{\mathbf{x}} \in S$ とすべての $\mathbf{x} \in S$ に対して，$(\mathbf{y}-\overline{\mathbf{x}})^t(\mathbf{x}-\overline{\mathbf{x}}) \leq 0$ が成り立つ $\Leftarrow$ $\overline{\mathbf{x}} \in S$ は最小点

Conversely, assume that $\|\mathbf{y}-\mathbf{x}\|^2 \geq \|\mathbf{y}-\overline{\mathbf{x}}\|^2$ for all $\mathbf{x} \in S$. Let $\mathbf{x} \in S$ and note that $\overline{\mathbf{x}} + \lambda(\mathbf{x}-\overline{\mathbf{x}}) \in S$ for $0 \leq \lambda \leq 1$ by the convexity of S. Therefore,

$$\|\mathbf{y}-\overline{\mathbf{x}}-\lambda(\mathbf{x}-\overline{\mathbf{x}})\|^2 \geq \|\mathbf{y}-\overline{\mathbf{x}}\|^2 \tag{2.3}$$

Also,

$$\|\mathbf{y}-\overline{\mathbf{x}}-\lambda(\mathbf{x}-\overline{\mathbf{x}})\|^2 = \|\mathbf{y}-\overline{\mathbf{x}}\|^2 + \lambda^2\|\mathbf{x}-\overline{\mathbf{x}}\|^2 - 2\lambda(\mathbf{y}-\overline{\mathbf{x}})^t(\mathbf{x}-\overline{\mathbf{x}}) \tag{2.4}$$

From (2.3) and (2.4) we get

$$2\lambda(\mathbf{y}-\overline{\mathbf{x}})^t(\mathbf{x}-\overline{\mathbf{x}}) \leq \lambda^2\|\mathbf{x}-\overline{\mathbf{x}}\|^2 \tag{2.5}$$

for all $0 \leq \lambda \leq 1$. Dividing (2.5) by any such $\lambda > 0$ and letting $\lambda \to 0^+$, the result follows.

Theorem 2.4.1 is illustrated in Figure 2.7a. Note that the angle between $(\mathbf{y}-\overline{\mathbf{x}})$ and $(\mathbf{x}-\overline{\mathbf{x}})$ for any point $\mathbf{x}$ in S is greater than or equal to 90°, and hence $(\mathbf{y}-\overline{\mathbf{x}})^t(\mathbf{x}-\overline{\mathbf{x}}) \leq 0$ for all $\mathbf{x} \in S$. This says that the set S lies in the half-space $\alpha^t(x-\overline{\mathbf{x}}) \leq 0$ relative to the hyperplane $\alpha^t(\mathbf{x}-\overline{\mathbf{x}}) = 0$ passing through $\overline{\mathbf{x}}$ and having a normal $\alpha = (\mathbf{y}-\overline{\mathbf{x}})$. Note also by referring to Figure 2.7b that this feature does not necessarily hold even over $N_\varepsilon(\bar{x}) \cap S$ if S is not convex.

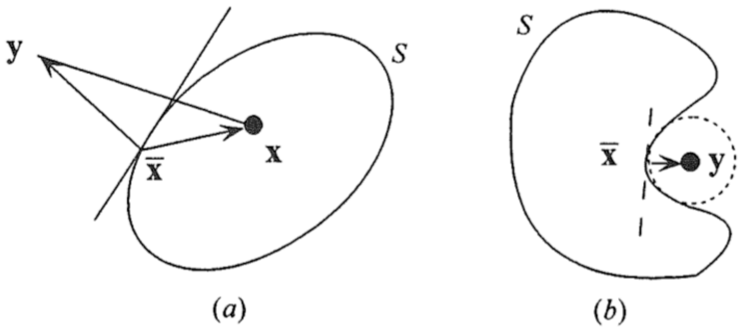

Figure 2.7 Minimum distance to a closed convex set.