AMATH 514 Assignment 2

Tyler Chen

Exercise 2.2

Let $C \subseteq \mathbb{R}^n$ be a convex set and let A be a $m \times n$ matrix. Show that the set $\{Ax \mid x \in C\}$ is again convex.

Any two points in C can be written as Au and Av for some $u, v, \in C$. Since C is convex, $\forall \lambda \in [0, 1]$, $\lambda u + (1 - \lambda)v \in C$. Thus,

$$\lambda Au + (1 - \lambda)Av = A(\lambda u + (1 - \lambda)v) \in \{Ax \mid x \in C\}, \qquad \forall \lambda \in [0, 1]$$

That is, $\{Ax \mid x \in C\}$ is convex. $\square$

Exercise 2.4

Show that if $z \in \text{convhull}(X)$, then there exist affinely independent vectors $x_1, ..., x_m$ in X such that $z \in \text{convhull}\{x_1, ..., x_m\}$.

Let $z \in \text{convhull}(X)$. Then for some $t \in \mathbb{N}$, $x_1, \ldots, x_t \in X$, and $c_1, \ldots, c_t \geq 0$ we can write,

$$z = c_1 x_1 + \cdots + c_t x_t \qquad \text{where} \qquad c_1 + \cdots + c_t = 1$$

WLOG assume $c_1, \ldots, c_t > 0$ since we can drop any x_j for which $c_j = 0$.

If $x_1, \ldots, x_t$ are affinely independent we are done. If not, then there are coefficients $d_1, \ldots, d_t$ not all zero such that,

$$d_1 x_1 + \cdots + d_t x_t = 0 \qquad \text{where} \qquad d_1 + \cdots d_t = 0$$

Observe that for any $\mu > 0$, since $\sum_{j=1}^{t}(c_j - \mu c_j) = \sum_{j=1}^{t} c_j - \mu \sum_{j=1}^{t} d_j = 1 - \mu(0) = 1$,

$$z = (c_1 - \mu d_1)x_1 + \cdots + (c_t - \mu d_t)x_t \qquad \text{where} \qquad \sum_{j=1}^{t}(c_j - \mu d_j) = 1$$

Note that for $\mu = 0$ we have the original combination for z, and each $(c_j - \mu d_j) = c_j > 0$.

Since the $\sum_j d_j = 0$ and not all d_j are zero, at least one $d_j > 0$. By hypothesis, $c_j > 0$. Thus, $\{c_j/d_j \bigm| 1 \leq j \leq t, c_j/d_j > 0\}$ is nonempty. Pick $\mu = c_i/d_i$ where $i = \text{argmin}\{c_j/d_j \bigm| 1 \leq j \leq t, c_j/d_j > 0\}$. Note $\mu > 0$.

If $d_j < 0$, then $\mu d_j < 0$ so $c_j - \mu d_j > 0$.

If $d_j \geq 0$, then since $(c_j - (c_i/d_i)d_j)/d_j = c_j/d_j - c_i/d_i \geq 0$,

$$c_j - \mu d_j = c_j - (c_i/d_i)d_j \geq c_i - (c_i/d_i)d_i = 1$$

So the coefficients for this combination of $\{x_1, \ldots, x_t\} \setminus \{x_i\}$ are all non-negative with sum equal to one.

Thus $z \in \text{convhull}(\{x_1, \ldots, x_t\} \setminus \{x_i\})$. Since the (finite) set generating the convex hull has fewer elements we can repeat this procedure until we have a finite affinely independent generating set. $\square$

Note: I got a bit of inspiration from stackexchange when looking up Carathéodory's theorem. However, I wrote my solution without looking back on that page.

Exercise 2.5

(i) Let C and D be two nonempty, bounded, closed, convex subsets of $\mathbb{R}^n$ such that $C \cap D = \emptyset$. Derive from Theorem 2.1 that there exists and affine hyperplane H seperating C and D.

(ii) Show that in (i) we cannot delete the boundedness condition.

(i) The function $l : C \times D \to \mathbb{R}_{\geq 0}$ defined as $l(x, y) := \|x - y\|$ is a continuous function (norm on $\mathbb{R}^{2n} \supset C \times D$ is continuous) on a compact set (Cartesian product of compact sets is compact). Therefore, $(c, d) = \operatorname{argmin}\{l(x, y) \mid x \in C, y \in D\}$ exists in $C \times D$.

Note then that d satisfies $d = \operatorname{argmin}\{\|d - c\| \mid d \in D\}$, and c satisfies $c = \operatorname{argmin}\{\|c - d\| \mid c \in C\}$.

Thus, by Theorem 2.1, d is separated from C by the hyperplane $H_d = \{x \mid (d - c)^T x = \delta\}$, and c is separated from D by the hyperplane $H_c = \{x \mid (c - d)^T x = \delta\}$, where $\delta = (\|c\|^2 - \|d\|^2)/2$ for both hyperplanes. That is, $(d - c)^T x < \delta$ for all $x \in C$, and $(c - d)^T x < \delta$ for all $x \in D$.

But then, $(d - c)^T x > \delta$ for all $x \in D$. This proves $H = H_d$ separates C from D. □

(ii) Let $C = \{(x, y) \mid y \leq 0\}$ and $D = \{(x, y) \mid y \geq 1/x, x > 0\}$. Then both sets are convex, closed, and unbounded.

Clearly any line (hyperplane in $\mathbb{R}^2$) which is not horizontal will intersect C. But any horizontal line will intersect D if it is above the x axis, or intersect C if it is on or below the x axis. Therefore, there is no line which separates C from D since such a line cannot intersect either C or D.

More rigorously, write $H = \{(x, y) \mid (c_1, c_2)^T (x, y) = \delta\} = \{(x, y) \mid c_1 x + c_2 y = \delta\}$.

Suppose $c_1 \neq 0$. Then the points $((\delta + c_2 + 1)/c_1, -1)$ and $((\delta + c_2 - 1)/c_1, -1)$ are both in C. However,

$$(c_1, c_2)^T ((\delta + c_2 + 1)/c_1, -1) = c_1(\delta + c_2 + 1)/c_1 + c_2(-1) = \delta + 1 > \delta$$
$$(c_1, c_2)^T ((\delta + c_2 - 1)/c_1, -1) = c_1(\delta + c_2 - 1)/c_1 + c_2(-1) = \delta - 1 < \delta$$

So the points $((\delta + c_2 + 1)/c_1, -1)$ and $((\delta + c_2 - 1)/c_1, -1)$ are on opposite sides of H.

Suppose $c_1 = 0$ and $\delta/c_2 \leq 0$. Then the points $(0, 0)$ and $(0, \delta/c_2 - 1)$ are both in C. However,

$$(0, c_2)^T (0, 0) = 0$$
$$(0, c_2)^T (0, \delta/c_2 - 1) = c_2(\delta/c_2 - 1) = \delta - c_2$$

If $\delta < 0$ then $c_2 > 0$ so $\delta - c_2 < \delta$. If $\delta \geq 0$ then $c_2 < 0$ so $\delta - c_2 > \delta$. Then the points $(0, 0)$ and $(\delta/c_2 - 1)$ are on opposite sides of H.

Suppose $c_1 = 0$ and $\delta/c_2 > 0$. Then the points $(2c_2/\delta, \delta/2c_2)$ and $(2c_2/\delta, 2\delta/c_2)$ are both in D. However,

$$(0, c_2)^T (2c_2/\delta, \delta/2c_2) = \delta/2$$
$$(0, c_2)^T (2c_2/\delta, 2\delta/c_2) = 2\delta$$

If $\delta < 0$ then $\delta/2 > \delta$ and $2\delta < \delta$. If $\delta > 0$ then $\delta/2 < \delta$ and $2\delta > \delta$. Then the points $(2c_2/\delta, \delta/2c_2)$ and $(2c_2/\delta, 2\delta/c_2)$ are on opposites sides of H.

This shows that there is no hyperplane H which separates C from D, as every hyperplane will split either C or D.

We therefore need the boundedness condition to have a strict separating hyperplane. □