Convexity Theory I — Class Notes

Section 1. Linear Algebra and Topology

In this section, we will cover some of the basic ideas of linear algebra and topology which will be used in the remainder of the course. We will work in Euclidean space $\mathbb{E}^n$. We denote the *inner product* of $x, y \in \mathbb{E}^n$ by $\langle x, y\rangle$. In terms of coordinates

$$\langle x, y\rangle = x_1 y_1 + \ldots + x_n y_n$$

where $x = (x_1, \ldots, x_n)$ and $y = (y_1, \ldots, y_n)$. The *norm* of x is denoted by $\|x\|$, and is given by

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

Geometrically, it is the distance from the point x to the origin. We recall that $\langle x, y\rangle = \|x\|\,\|y\| \cos\gamma$ where γ is the angle between the vectors and x and y. In particular, $\langle x, y\rangle = 0$ if x and y are orthogonal vectors. The following is a list of simple, but important, facts about inner products and norms.

a) $\langle x, x\rangle \geq 0$ and $\langle x, x\rangle = 0 \Leftrightarrow x = 0$
b) $\langle x, y\rangle = \langle y, x\rangle$
c) $\langle x + y, z\rangle = \langle x, z\rangle + \langle y, z\rangle$
d) $\langle \alpha x, y\rangle = \alpha\langle x, y\rangle$, if $\alpha \in \mathbb{R}$
e) $\|x\| > 0$ if $x \neq 0$
f) $\|\alpha x\| = |\alpha|\,\|x\|$, if $\alpha \in \mathbb{R}$

There is one further property, known as the *triangle inequality*

$$\|x + y\| \leq \|x\| + \|y\| .$$

The key to proving this is Schwarz' inequality

$$|\langle x, y\rangle| \leq \|x\|\,\|y\|$$

which follows from the fact that $|\cos\alpha| \leq 1$. Schwarz' inequality gives

$$\begin{aligned}\|x + y\|^2 &= \langle x + y, x + y\rangle = \|x\|^2 + \|y\|^2 + 2\langle x, y\rangle \\ &\leq \|x\|^2 + \|y\|^2 + 2\,\|x\|\,\|y\| = (\|x\| + \|y\|)^2.\end{aligned}$$

We already mentioned that the norm measures distances from the origin. In fact, it can be used to measure any distance. The *distance* $d(x, y)$ is defined by

$$d(x, y) = \|x \Leftrightarrow y\| \qquad\qquad x, y \in \mathbb{E}^n.$$

In terms of coordinates, this is

$$d(x,y) = \sqrt{\sum_{i=1}^{n}(x_i \Leftrightarrow y_i)^2}.$$

The distance function satisfies

a) $d(x,y) > 0$ if $x \neq y$
b) $d(x,y) = d(y,x)$
c) $d(x,y) \leq d(x,z) + d(z,y)$ (triangle inequality)
d) $d(\lambda x, \lambda y) = |\lambda| d(x,y) \qquad \lambda \in \mathbb{R}$
e) $d(x+z, y+z) = d(x,y)$

We now turn to some topological considerations. These will all make use of the distance function. If $x \in \mathbb{E}^n$ and $\delta > 0$, the *open ball* $B(x,\delta)$ is defined by

$$B(x,\delta) = \{y \in \mathbb{E}^n : d(x,y) < \delta\}.$$

The point x of a set S in $\mathbb{E}^n$ is an *interior point* if there is a $\delta > 0$ such that $B(x,\delta) \subset S$. A set S is *open* if each of its points is an interior point. Note that an open ball is open. (Why?) The collection of all open sets as defined above, is referred to as the *usual topology* on $\mathbb{E}^n$. If S is a subset of $\mathbb{E}^n$, then

$$\{U \subset S : U = V \cap S, \text{ where } V \text{ is open in } \mathbb{E}^n\}$$

is called the *relative topology* on S. For example, if $S \subset \mathbb{E}^2$ is defined by $S = \{(x_1, x_2) : x_1 \geq 0\}$, then

$$G = \{(x_1, x_2) : 0 \leq x_1 < 1, 0 < x_2 < 1\}$$

is open in the relative topology of S, but G is not an open set in $\mathbb{E}^2$.

It is important to note the following:

a) the empty set ϕ, and the whole space $\mathbb{E}^n$ are open;
b) any union of open sets is open;
c) any finite intersection of open sets is open.

Can you give an example of an infinite collection of open sets whose intersection is not open? Here is one:

$$\bigcap_{n=1}^{\infty} B(o, \frac{1}{n}) \qquad n = 1, 2, \ldots,$$

which is simply $\{o\}$.

A set S in $\mathbb{E}^n$ is *closed* if its complement $\sim S = \mathbb{E}^n \setminus S$ is open in $\mathbb{E}^n$. Note that

a) ϕ and $\mathbb{E}^n$ are closed;
b) all intersections of closed sets are closed;
c) any finite union of closed sets is closed.

We have therefore seen that ϕ and $\mathbb{E}^n$ are both open and closed. In fact, these are the only possibilities in $\mathbb{E}^n$, but we won't prove that. But there are many examples of sets which are neither open nor closed, for example the set G mentioned previously.

The set S is said to be *bounded* if there is an $x \in \mathbb{E}^n$ and a $\delta > 0$ such that $S \subset B(x, \delta)$. The *interior* of a set S is the union of all open sets contained in S; it is denoted by $\operatorname{int} S$. The *closure* of a set S is the intersection of all closed sets that contain S; it is denoted by $\operatorname{cl} S$. For example, if $S \subset \mathbb{E}^2$ is the set of points with rational coordinates then

$$\operatorname{int} S = \phi \qquad\qquad \operatorname{cl} S = \mathbb{E}^2$$

A function $f : \mathbb{E}^n \to \mathbb{E}^m$ is *continuous* if and only if $f^{-1}(U)$ is open in $\mathbb{E}^n$ whenever U is open in $\mathbb{E}^m$. This definition is equivalent to the familiar ε, δ definition of continuity. In the latter, we say that f is continuous at $x \in \mathbb{E}^n$ if, for every $\varepsilon > 0$, there is a $\delta > 0$ such that

$$f(B(x, \delta)) \subset B(f(x), \varepsilon).$$

Then $f : \mathbb{E}^n \to \mathbb{E}^m$ is said to be continuous if it is continuous at each point $x \in \mathbb{E}^n$. To see this equivalence, first assume $f : \mathbb{E}^n \to \mathbb{E}^m$ is continuous (inverse image definition) and let $x \in \mathbb{E}^n$ and $\varepsilon > 0$. Note that $B(f(x), \varepsilon)$ is open in $\mathbb{E}^m$ and therefore $f^{-1}(B(f(x), \varepsilon)$ is open in $\mathbb{E}^n$. Of course $x \in f^{-1}(B(f(x), \varepsilon))$. The definition of open set now shows that there is a $\delta > 0$ such that

$$B(x, \delta) \subset f^{-1}(B(f(x), \varepsilon)),$$

and this is just the statement

$$f(B(x, \delta)) \subset B(f(x), \varepsilon).$$

So we have proved that if f is continuous then f is continuous at each $x \in \mathbb{E}^n$.

For the converse, we assume $f : \mathbb{E}^n \to \mathbb{E}^m$ is continuous at each $x \in \mathbb{E}^n$ (ε, δ definition) and let U be open in $\mathbb{E}^m$. We assume that $x \in f^{-1}(U)$ and aim to find a $\delta > 0$ such that

$$B(x, \delta) \subset f^{-1}(U).$$

We have $f(x) \in U$ and so there is an $\varepsilon > 0$ such that $B(f(x), \varepsilon) \subset U$. The continuity of f at x implies there is a $\delta > 0$ with

$$f(B(x, \delta)) \subset B(f(x), \varepsilon).$$

In particular,

$$B(x, \delta) \subset f^{-1}(U)$$

and so $f^{-1}(U)$ is open.

There is a further equivalent definition of continuity at x which we will find helpful: f is continuous at $x \in \mathbb{E}^n$ if the sequence $(f(x_i))_{i=1}^{\infty}$ converges to $f(x)$ in $\mathbb{E}^m$ whenever the sequence $(x_i)_{i=1}^{\infty}$ converges to $x \in \mathbb{E}^n$.

It is easy to show that the following functions are continuous

a) $f : \mathbb{E}^n \times \mathbb{E}^n \to \mathbb{E}^n$ defined by $f(x, y) = x + y$
b) if $a \in \mathbb{E}^n, f_a : \mathbb{E}^n \to \mathbb{E}^n$, defined by $f_a(x) = a + x$
c) if $\lambda \in \mathbb{R}, f_\lambda : \mathbb{E}^n \to \mathbb{E}^n$, defined by $f_\lambda(x) = \lambda x$
d) if $x, y \in \mathbb{E}^n, f : \mathbb{R} \to \mathbb{E}^n$, defined by $f(\lambda) = \lambda x + (1 \Leftrightarrow \lambda)y$

If $A, B \subset \mathbb{E}^n$ and $\lambda \in \mathbb{R}$, we define

$$A + B = \{x + y : x \in A, y \in B\}$$

and

$$\lambda A = \{\lambda x : x \in A\}.$$

It is unfortunate that, in general,

$$A + A \neq 2A.$$

For example, if $A = \{x, y\} \subset \mathbb{E}^2$, then $A + A = \{2x, x + y, 2y\}$, whereas $2A = \{2x, 2y\}$. If $A = \{x\}$ we write $x + B$ for $A + B$. This is a *translate* of B obtained by translating each point of B by the vector x. It is convenient to note that

$$A + B = \bigcup_{x \in A} (x + B) = \bigcup_{y \in B} (A + y).$$

Sets of the form $x + \lambda A$ with $\lambda \neq 0$ are said to be *homothetic* to A or *homothets* of A.

The *boundary* of A, $\operatorname{bd} A$, is defined by

$$\operatorname{bd} A = \operatorname{cl} A \cap \operatorname{cl}(\sim A)$$

A set K in $\mathbb{E}^n$ is said to be *compact* if it is closed and bounded.

The following theorem (one of the most significant in mathematics) gives us an equivalent notion of compactness. It uses the idea of "open cover." A collection of open sets is said to form an *open cover* of S if S is contained in the union of the sets of the collection. For example

$$\{B(x, 1 \Leftrightarrow 1/n) : n = 2, 3, \ldots\}$$

is an open cover of $B(x, 1)$.

Heine-Borel Theorem. *The set $K \subset \mathbb{E}^n$ is compact, if and only if every open cover has a finite subcover.*

Note that the above open cover of $B(x, 1)$ has no finite subcover; so $B(x, 1)$ is not compact.

Theorem 1.22. *Let $f : \mathbb{E}^n \to \mathbb{E}^m$ be continuous and assume $K \subset \mathbb{E}^n$ is compact. Then $f(K)$ is compact in $\mathbb{E}^m$.*

Proof. We assume

$$\mathcal{F} = \{\mathcal{F}_\alpha : \alpha \in A\}$$

is an open cover of $f(K)$, that is

$$f(K) \subset \bigcup_{\alpha \in A} \mathcal{F}_\alpha.$$

Each $\mathcal{F}_\alpha$ is open in $\mathbb{E}^m$ and so, by continuity of f, each $f^{-1}(\mathcal{F}_\alpha)$ is open in $\mathbb{E}^n$. Clearly

$$K \subset \bigcup_{\alpha \in A} f^{-1}(\mathcal{F}_\alpha)$$

and so we have an open cover of K. We may therefore choose $\alpha_1, \ldots, \alpha_t \in A$ such that

$$K \subset \bigcup_{i=1}^{t} f^{-1}(\mathcal{F}_{\alpha_i}).$$

It follows immediately that

$$f(K) \subset \bigcup_{i=1}^{t} \mathcal{F}_{\alpha_i}$$

and so we have found the required finite subcover. $\square$

Corollary 1.23. *Let K be compact in $\mathbb{E}^n$ and assume $f : \mathbb{E}^n \to \mathbb{R}$ is continuous. Then there is an $m \in \mathbb{R}$ such that $|f(x)| \leq m$ for all $x \in K$. Furthermore, there are points $x_1, x_2 \in K$ such that*

$$f(x_1) = \inf_{x \in K} f(x) \qquad f(x_2) = \sup_{x \in K} f(x).$$

Proof. Note that $f(K)$ is a compact subset of $\mathbb{R}$. It is therefore closed and bounded. The boundedness implies there is an interval $[a, b]$ such that $f(K) \subset [a, b]$. If we put $m = \max\{|a|, |b|\}$ we have $|f(x)| \leq m$ for all $x \in K$. The boundedness of $f(K)$ also implies that $\inf_{x \in K} f(K)$ and $\sup_{x \in K} f(K)$ both exist. The fact that $f(K)$ is closed now implies that both the infimum and supremum are members of $f(K)$. $\square$