Open and closed sets – elementary topology in $\mathbb{R}^n$

Definitions and facts, a bit in excess of what needs to be known for Opt 2.

- An *open ball* $B_r(\boldsymbol{x}^0)$ in $\mathbb{R}^n$ (centered at $\boldsymbol{x}^0$, of radius r) is a set $\{\boldsymbol{x} : \|\boldsymbol{x} - \boldsymbol{x}^0\| < r\}$, where from now on $\|\boldsymbol{x}\| = \sqrt{x_1^2 + \ldots + x_n^2}$ is the Euclidean distance. The case $r = 0$ corresponds to the empty set, which is also open. By default, B_r (without specifying the center) means a ball centered at the origin.

- An *open set* in $\mathbb{R}^n$ is any union of open balls, in particular $\mathbb{R}^n$ itself. Therefore if X is open, then for any $\boldsymbol{x} \in X$, there exists a ball $B_r(\boldsymbol{x}) \subset X$, for some r. So, the union of any family of open sets is open. Also, the intersection of a *finite number* of open sets is open. (E.g. the family of open intervals $(-1 - 1/n, 1 + 1/n)$, $n = 1, 2, \ldots, 100$ is finite; if $n = 1, 2, \ldots$, this family is countable; the family of open sets $(-1 - \alpha, 1 + \alpha)$, $\alpha \in (0, 1)$ is uncountable.)

- A set $X \subset \mathbb{R}^n$ is *closed* if its complement $X^c = \mathbb{R}^n \setminus X$ is open. Hence, both $\mathbb{R}^n$ and $\varnothing$ are at the same time open and closed, these are the only sets of this type. Furthermore, the intersection of *any family* or union of *finitely many* closed sets is closed.

  Note: there are many sets which are neither open, nor closed.

- For any set X, its *closure* $\bar{X}$ is the smallest closed set containing X. Its *interior* $\underline{\mathrm{X}}$ is the largest open set contained in X. Its *boundary* ∂X is by definition $\bar{X} \setminus \underline{\mathrm{X}}$. Clearly, if X is closed, then $X = \bar{X}$ and if X is open, then $X = \underline{\mathrm{X}}$. Also, if $X = \{p\}$, a single point, then $X = \bar{X} = \partial X$.

- A set X is *bounded* if there exists a ball B_R such that $X \subset B_R$ for some R. A set, which is closed and bounded is called *compact*.

- A *sequence* in $\mathbb{R}^n$ is a countable collection of points $\{\boldsymbol{x}_n\}_{n=1,2,\ldots} = \{\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots\}$ (countable collection means an infinite array which can be put in one-to one correspondence with positive integers; however $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots$ are not necessarily distinct). A sequence $\{\boldsymbol{x}_n\}$ *converges* to $\boldsymbol{x}$ (i.e. $\boldsymbol{x} = \lim_{n\to\infty} \boldsymbol{x}_n$, the subscript $_{n\to\infty}$ often being omitted) if for any ball $B_\epsilon(\boldsymbol{x})$ centered at $\boldsymbol{x}$, all members of the sequence, starting from some $n = N(\epsilon)$ find themselves inside the ball $B_\epsilon(\boldsymbol{x})$. A subsequence $\{\boldsymbol{x}_{n_k}\}$ of a sequence $\{\boldsymbol{x}_n\}$ is a countable sub-colleciton of $\{\boldsymbol{x}_n\}$.

- x is a *limit point* of a sequence $\{\boldsymbol{x}_n\}$ if there is a subsequence $\{\boldsymbol{x}_{n_k}\}$ of $\{\boldsymbol{x}_n\}$, converging to $\boldsymbol{x}$. E.g. for the sequence $\{1, 1, 2, 1, 2, 3, 1, 2, 3, 4, 1, 2, 3, 4, 5, \ldots\} \subset \mathbb{R}$ every integer is a limit point. If a sequence converges, then any subsequence converges to the same limit, which is the limit of the sequence. For a set $X \subseteq \mathbb{R}^n$, $\boldsymbol{x}$ is a *limit point* if it is a limit point for some sequence $\{\boldsymbol{x}_n\} \subseteq X$. So, for any $r > 0$, there are infinitely many $\boldsymbol{x}_n \in B_r(\boldsymbol{x})$.

- The closure $\bar{X}$ of a set X is the union of all the limit points of X. Above, we've defined closure as the smallest closed set containing X. To prove equivalence of the two definitions, let us first show that the set $\hat{X}$ of all the limit points of X is closed and contains X, and then that it is the smallest such set, so $\hat{X} = \bar{X}$. First off, if $\boldsymbol{x} \in X$, then take a sequence $\{\boldsymbol{x}, \boldsymbol{x}, \boldsymbol{x}, \ldots\}$, it clearly converges to $\boldsymbol{x}$, so $\boldsymbol{x}$ is a limit point of X. So $X \subseteq \hat{X}$. Furthermore, $\hat{X}$ is closed, because its complement $\hat{X}^c$ is open. Indeed, if it's not, there is some $\boldsymbol{x}' \in \hat{X}^c$ such that any ball centered at $\boldsymbol{x}'$ would intersect $\hat{X}$. This means that any ball centered at $\boldsymbol{x}'$ will contain points of X, too. But then there is a sequence $\{\boldsymbol{x}_n\}$ of points of X, converging to $\boldsymbol{x}'$, which is a contradiction. So $\hat{X}$ is closed. Let us show that $\bar{X} = \hat{X}$ (so far it follows only that $\bar{X} \subseteq \hat{X}$, because $\bar{X}$ is the smallest closed set containing X). Consider an open set $\bar{X}^c$, let us show that it has no elements of $\hat{X}$. Indeed, if it does contain some $\boldsymbol{x}' \in \hat{X}$, then it contains come ball centered therein alongside. This ball does not intersect X (because it lies outside $\bar{X}$) and therefore its center $\boldsymbol{x}'$, although it belongs to $\hat{X}$ cannot be a limit point of X. Contradiction, unless $\bar{X} = \hat{X}$.

- This enables one to easily prove that some sets are closed, e.g. level sets $f(\boldsymbol{x}) = c$ of continuous functions or their *sublevel sets* $\{\boldsymbol{x} : f(\boldsymbol{x}) \leq c\}$. Indeed, if f is continuous, and $f(\boldsymbol{x}_n) = c$, then $f(\boldsymbol{x}) = c$ for $\boldsymbol{x} = \lim \boldsymbol{x}_n$. Same for the sublevel set case and would not be true if there was $<$ instead of $\leq$.

- A (vector-) function $f : \mathbb{R}^n \to \mathbb{R}^m$ is *continuous* if the pre-image $X \subseteq \mathbb{R}^n$ of any open set $Y \subseteq \mathbb{R}^m$ in the range of f (i.e. $X = \{\boldsymbol{x} : f(\boldsymbol{x}) = \boldsymbol{y},$ for some $\boldsymbol{y} \in Y\}$) is open. Equivalently, for any $\boldsymbol{y}$ in the range of f and any ball $B_\epsilon(\boldsymbol{y}) \subset \mathbb{R}^m$ there exists a ball $B_\delta(\boldsymbol{x}) \subset \mathbb{R}^n$, such that any $\boldsymbol{x} \in B_\delta(\boldsymbol{x})$ is taken into the ball $B_\epsilon(\boldsymbol{y})$ by f. Equivalently, if $\boldsymbol{x} = \lim \boldsymbol{x}_n$, then $f(\boldsymbol{x}) = \lim f(\boldsymbol{x}_n)$ (provided that $\boldsymbol{x}$ is also in the domain of f; this can always be achieved by assuming that the domain of f is a closed set). Note that if $n = m = 1$, then balls are intervals, e.g. $B_\epsilon(\boldsymbol{y}) = (\boldsymbol{y} - \epsilon, \boldsymbol{y} + \epsilon)$, and the above reduces to usual $\epsilon - \delta$ definitions.

- Finally, Bolzano-Weierstrass theorem, which is routinely applied in non-linear optimisation to ensure existence of optimisers. *A continuous real-valued ($m = 1$ in the above definition) function f on a compact set X reaches on X is* `supremum` *and* `infimum`. Recall that $\sup_X f$ is the *least upper bound* of the set of values $\{f(\boldsymbol{x}), \boldsymbol{x} \in X\}$. And $\inf_X f$ is the greatest lower bound. Let's prove the theorem for the supremum (as $\inf f = -\sup -f$.) First off, if $M = \sup_X f$, then there is a sequence $\{\boldsymbol{x}_k\} \subseteq X$ such that $M = \lim f(\boldsymbol{x}_k)$. Then, as long as there is a limit point $\boldsymbol{x}$ for the sequence $\{\boldsymbol{x}_k\}$ (if $\boldsymbol{x}$ exists, it is in $\in X$, as X is closed), then M is finite, equal to $f(\boldsymbol{x})$. I.e. by continuity of f, $M = \lim f(\boldsymbol{x}_k) = f(\boldsymbol{x}) < \infty$. We can also assume that all the members of the sequence $\{x_k\}$ are different. Otherwise, if any $\boldsymbol{x}$ is repeated in the sequence infinitely many times, then $M = f(\boldsymbol{x})$, and there is nothing left to prove.

  So, what is left to prove is that any sequence $\{\boldsymbol{x}_k\}$ in a compact (closed and bounded) set X has a limit point, i.e. a convergent subsequence. To do this, a bit heuristically, enclose X in some n-dimensional cube (a cube in two dimensions is a square) Q_0, this is OK since X is bounded. Now divide Q_0 into 2^n congruent cubes by dissecting every edge. Let Q_1 be one of those, which (together with its boundary) contains infinitely many members of the sequence.

  Do the same thing now with Q_1, and so on. We obtain a sequence $Q_0, Q_1, Q_2, \ldots$ of nested cubes with edge length vanishing geometrically, so that each of these cubes contains infinitely many members of the sequence $\{\boldsymbol{x}_k\}$. The real space is complete: there is a unique point $\boldsymbol{x}$ which belongs to all these cubes. And for every ball $B_r(\boldsymbol{x})$ centered in $\boldsymbol{x}$ there will be infinitely many $\boldsymbol{x}_k \in B_r(\boldsymbol{x})$. So $\boldsymbol{x}$ is a limit point of the sequence $\{\boldsymbol{x}_k\}$. Since X is a closed set, $x \in X$.

  And once again, since f is continuous, now $f(\boldsymbol{x}) = M < \infty$.

- For optimisation, this theorem has an important corollary. *An optimisation problem* `Max`$f(\boldsymbol{x})$, *on a compact feasible set F, with a continuous objective function f always has a solution (alias optimizer).* In particular, in unbounded LPs the feasible set ought to be unbounded, so Bolzano-Weierstrass theorem does not apply. When the feasible set is given in terms of constraints $g_i(\boldsymbol{x}) \leq b_i$, $i = 1, \ldots, m$, where all g_i are continuous functions, the feasible set is closed, yet not necessarily bounded.

Convex sets

Definitions and facts.

- A set $X \subseteq \mathbb{R}^n$ is *convex* if for any distinct $\boldsymbol{x}^1, \boldsymbol{x}^2 \in X$, the whole line segment $\boldsymbol{x}^\theta = \theta \boldsymbol{x}^1 + (1-\theta)\boldsymbol{x}^2$, $0 \le \theta \le 1$ between $\boldsymbol{x}^1$ and $\boldsymbol{x}^2$ is contained in X. Note that changing the condition $0 \le \theta \le 1$ to $\theta \in \mathbb{R}$ would result in $\boldsymbol{x}^\theta$ describing the straight line passing through the points $\boldsymbol{x}^1$ and $\boldsymbol{x}^2$. The empty set and a set containing a single point are also regarded as convex.

- The intersection of any family X_i of convex sets is convex. Indeed, if points $\boldsymbol{x}^{1,2} \in \bigcap_i X_i$, they belong to each set X_i, then so does the line segment $\boldsymbol{x}^\theta$, so it belongs to the intersection $\bigcap_i X_i$ as well.

- A unit vector $\boldsymbol{d} \in \mathbb{R}^n$, $\|\boldsymbol{d}\| = 1$ is a *direction* for a convex set X at a point $\boldsymbol{x}^0$, if for some small $t > 0$, the point $\boldsymbol{x}^t = \boldsymbol{x}^0 + t\boldsymbol{d}$ lies in X as well. A point $\boldsymbol{x}^0 \in X$ is an *extreme point* of the convex set X is there is no $\boldsymbol{d} \in \mathbb{R}^n$, $\|\boldsymbol{d}\| = 1$, such that both $\boldsymbol{d}$ and $-\boldsymbol{d}$ are directions at $\boldsymbol{x}^0$. Equivalently, $\boldsymbol{x}^0$ is an extreme point, if there is no line segment with endpoints $\boldsymbol{x}^{1,2} \neq \boldsymbol{x}^0$, contained in X, such that $\boldsymbol{x}^0$ lies inside this line segment. Note: open convex sets have no extreme points, as for any $\boldsymbol{x} \in X$ there would be a small ball $B_r(\boldsymbol{x}) \subset X$, in which case any $\boldsymbol{d}$ is a direction, at any $\boldsymbol{x}$.

- A *hyperplane* $H_{\boldsymbol{c},\alpha}$ in $\mathbb{R}^n$ is a set $\{\boldsymbol{x} : \mathbf{c}^T\boldsymbol{x} + \alpha = 0\}$. It's easy to verify (using the definitions only) that a hyperplane is a closed convex set. A *halfspace* $H^+_{\mathbf{c},\alpha}$ in $\mathbb{R}^n$ is a set $\{\boldsymbol{x} : \mathbf{c}^T\boldsymbol{x} + \alpha \ge 0\}$; it is also a closed convex set.

- If $\boldsymbol{x}^0$ is an extreme point of a closed convex set X, a hyperplane $H_{\boldsymbol{c},\alpha}$ is called *supporting* hyperplane to X at $\boldsymbol{x}^0$ if $\boldsymbol{x}^0 \in H_{\boldsymbol{c},\alpha}$ and $X \subseteq H^+_{\mathbf{c},\alpha}$. I.e. $\mathbf{c}^T\boldsymbol{x} + \alpha \ge 0$ for any $\boldsymbol{x} \in X$, with the equality if $\boldsymbol{x} = \boldsymbol{x}^0$.

- Important theorem on convex sets. *Given two disjoint closed convex sets X_1, X_2, there exists a* `separating hyperplane`*, namely a hyperplane $H_{\boldsymbol{c},\alpha}$, such that $\mathbf{c}^T\boldsymbol{x} + \alpha \ge 0$ for any $\boldsymbol{x} \in X_1$ and $\mathbf{c}^T\boldsymbol{x} + \alpha < 0$ for any $\boldsymbol{x} \in X_2$. If besides one of the sets $X_{1,2}$ is bounded, there exists a hyperplane $H_{\boldsymbol{c},\alpha}$ which strictly separates the sets, $\mathbf{c}^T\boldsymbol{x} + \alpha > 0$ for any $\boldsymbol{x} \in X_1$ and $\mathbf{c}^T\boldsymbol{x} + \alpha < 0$ for any $\boldsymbol{x} \in X_2$.* This fact will underlie the proof of the *Farkas alternative*, to come up soon in the course.

  The proof – schematically, only when one of the sets is bounded, details are omitted – is based on the fact that as the sets are closed, and if one of the sets $X_{1,2}$ is bounded, then by Bolzano-Weierstrass theorem, the quantity
  $$\inf_{\boldsymbol{x}^1 \in X_1,\, \boldsymbol{x}^2 \in X_2} \|\boldsymbol{x}^1 - \boldsymbol{x}^2\|$$
  (the minimum Euclidean distance between the sets X_1 and X_2) is well defined and achieved for some $\boldsymbol{x}^1 \in X_1$ and $\boldsymbol{x}^2 \in X_2$. If so, let $\boldsymbol{c} = \boldsymbol{x}^1 - \boldsymbol{x}^2$ and draw a hyperplane through the midpoint of the segment $[\boldsymbol{x}^1, \boldsymbol{x}^2]$, with the normal vector $\boldsymbol{c}$. If this plane intersected X_1 or X_2, say X_1 at some point $\boldsymbol{x}$, then by convexity the line segment $[\boldsymbol{x}\boldsymbol{x}^1]$ lies in X_1. The one can drop a perpendicular from $\boldsymbol{x}^2$ to $[\boldsymbol{x}\boldsymbol{x}^1]$ and get the intersection point $\boldsymbol{x}' \in X_1$, which is closer to $\boldsymbol{x}^2$ than $\boldsymbol{x}^1$ – contradiction.

- Theorem regarding linear optimisation. *Consider a canonical LP $A\boldsymbol{x} = \boldsymbol{b}$, $\boldsymbol{x} \ge 0$. Then the feasible set F for this LP is convex and closed. Besides, basic feasible solutions (BFS) are in one-to-one correspondence with extreme points (EP) of F.*

  `Proof:` First notice that $x_j \ge 0$ defines a half-space in $\mathbb{R}^n$, while each equation of the system $A\boldsymbol{x} = \boldsymbol{b}$ (i.e $a_{i1}x_1 + a_{i2}x_2 + \ldots + x_{in}x_n = b_i$, $i = 1, \ldots, m$) determines a hyperplane. So, the feasible set is the intersection of a family of closed and convex sets, which itself is closed and convex.

Now, let us prove that if $\boldsymbol{x}^0$ is a BFS, it is an EP. If $\boldsymbol{x}^0$ is a BFS, either $\boldsymbol{x}^0 = 0$ or

$$x^0_\alpha \boldsymbol{a}^\alpha + x^0_\beta \boldsymbol{a}^\beta + \ldots = \boldsymbol{b}, \tag{1}$$

for some columns $\boldsymbol{a}^\alpha, \boldsymbol{a}^\beta, \ldots$ of A, with x^0_α, x^0_β being strictly positive. If $\boldsymbol{x}^0 = 0$, it is an EP. Indeed, for any nonzero vector $\boldsymbol{d}$ at the origin, both $\boldsymbol{d}$ and $-\boldsymbol{d}$ cannot have all non-negative components. So suppose (1) holds, $\boldsymbol{x}^0$ is a BFS (so the columns $\boldsymbol{a}^\alpha, \boldsymbol{a}^\beta, \ldots$ are linearly independent) and $\boldsymbol{x}^0$ is not an EP. Then for some $\boldsymbol{x}^1, \boldsymbol{x}^2 \in F$ and some $\theta \in (0,1)$, $\boldsymbol{x}^0 = \theta \boldsymbol{x}^1 + (1-\theta)\boldsymbol{x}^2$. Thus both $\boldsymbol{x}^1$ and $\boldsymbol{x}^2$ cannot have any positive components other than those of $\boldsymbol{x}^0$. And they are feasible solutions. So equation (1) is satisfied by $(x^1_\alpha, x^1_\beta, \ldots)$ as well as by $(x^2_\alpha, x^2_\beta, \ldots)$. Subtraction yields

$$(x^1_\alpha - x^2_\alpha)\boldsymbol{a}^\alpha + (x^1_\beta - x^2_\beta)\boldsymbol{a}^\beta + \ldots = 0,$$

which implies that the columns $\boldsymbol{a}^\alpha, \boldsymbol{a}^\beta, \ldots$ are linearly dependent – in contradiction with the fact that $\boldsymbol{x}^0$ is a BFS.

Conversely, suppose $\boldsymbol{x}^0$ is an EP, let us show that it is a BFS. If $\boldsymbol{x}^0 = 0$, then it is basic by definition. Otherwise it satisfies (1), with positive $x^0_\alpha, x^0_\beta, \ldots$. Suppose, $\boldsymbol{x}^0$ is not a BFS, then the columns $\boldsymbol{a}^\alpha, \boldsymbol{a}^\beta, \ldots$ must be linearly dependent. That is for some array of numbers $(\lambda_\alpha, \lambda_\beta, \ldots)$, which are not all zero,

$$\lambda_\alpha \boldsymbol{a}^\alpha + \lambda_\beta \boldsymbol{a}^\beta + \ldots = 0. \tag{2}$$

Multiply equation (2) by $\pm\delta$, for some sufficiently small positive δ and add to equation (1). Get

$$(x^0_\alpha \pm \delta\lambda_\alpha)\boldsymbol{a}^\alpha + (x^0_\beta \pm \delta\lambda_\beta)\boldsymbol{a}^\beta + \ldots = b.$$

That is for a small enough δ (so all the expressions in parentheses remain positive) there is a straight line segment of feasible solutions, with endpoints $\boldsymbol{x}^0 \pm \delta\boldsymbol{\lambda}$, (where the array $(\lambda_\alpha, \lambda_\beta, \ldots)$ extends to a vector $\boldsymbol{\lambda} \in \mathbb{R}^n$ by rendering its free, i.e. not listed by $(\alpha, \beta, \ldots)$, components as zero) such that $\boldsymbol{x}^0$ is the middle thereof, so $\boldsymbol{x}^0$ is not an extreme point. Contradiction.

Q.E.D.