3. Convexity

Takuma Habu*
takumahabu@cornell.edu

17th September 2024

1 Convex sets

Loosely speaking convexity is the idea that sets (in linear spaces) do not have "holes".

Definition 1. A *convex combination* of elements $\mathbf{x}, \mathbf{y} \in \mathbb{R}^d$ is an element $\mathbf{z} = \alpha\mathbf{x} + (1-\alpha)\mathbf{y}$ for some $\alpha \in [0, 1]$.

Definition 2. A subset $S \subseteq \mathbb{R}^d$ is *convex* if it contains all convex combinations of pairs of elements of S; i.e., $\alpha\mathbf{x} + (1-\alpha)\mathbf{y} \in S$ for any $\mathbf{x}, \mathbf{y} \in S$ and any $\alpha \in [0, 1]$.

Proposition 1. *Let $S \subseteq \mathbb{R}^d$ be convex, $\{\alpha_1, \ldots, \alpha_k\}$ be a set of $k \in \mathbb{N}$ numbers in $[0, 1]$ such that $\sum_{i=1}^k \alpha_i = 1$, and $\{\mathbf{x}_1, \ldots, \mathbf{x}_k\} \subseteq S$. Then, $\sum_{i=1}^k \alpha_i \mathbf{x}_i \in S$.*

Exercise 1. Prove this.[1]

Corollary 1. *S is convex if and only if S contains all (finite) convex combinations of elements in S.*

Proposition 2. *Let $\mathcal{C}$ be an arbitrary collection of convex sets. Then $\bigcap_{S \in \mathcal{C}} S$ is convex.*

Exercise 2. Prove Proposition 2.

Definition 3. The *convex hull of* a set $S \subseteq \mathbb{R}^d$, denoted $\text{co}(S)$, is the smallest convex set containing S; i.e.,

$$\text{co}\,(S) \coloneqq \bigcap \left\{T \subseteq \mathbb{R}^d : S \subseteq T,\, T \text{ is convex}\right\}.$$

Remark 1. Both of these definitions are well posed because: $\mathbb{R}^d$ itself is convex (and closed) and contains S; hence the intersection of *all* convex (and closed) sets that contain S is well-defined. This intersection clearly contains S and is convex [and closed] because arbitrary intersections of convex [closed] sets are convex [closed]. And this intersection must be the smallest convex [and closed] set containing S.

Proposition 3. *Let S be a subset of $\mathbb{R}^d$. Then, $\text{co}(S)$ is the collection of all finite convex combinations of elements in S; i.e.,*

$$\text{co}\,(S) = \left\{\mathbf{x} \in \mathbb{R}^d : \exists n \in \mathbb{N},\ (\mathbf{y}_i, \alpha_i)_{i=1}^n \in S^n \times [0,1]^n,\ \sum_{i=1}^n \alpha_i = 1,\ \mathbf{x} = \sum_{i=1}^n \alpha_i \mathbf{y}_i\right\}.$$

*Thanks to Giorgio Martini, Nadia Kotova and Suraj Malladi for sharing their lecture notes, on which these notes are heavily based.

[1] *Hint:* Use induction on m.

Exercise 3 (PS5)**.** Prove Proposition 3.[2]

Theorem 1 (Carathéodory)**.** *Let S be a subset of $\mathbb{R}^d$, then any point in $\operatorname{co}(S)$ can be expressed as a convex combination of no more than $d+1$ elements of S.*

Proof. Let S be a subset of $\mathbb{R}^d$. Take any point $\mathbf{y} \in \operatorname{co}(S)$. From Proposition 3, we know that y can be expressed as a convex combination of some finite elements in S. There might be different ways to do this, so fix a convex combination with the *least* number of elements from S, say, k elements. That is, we have that $\mathbf{y} = \sum_{i=1}^{k} \alpha_i \mathbf{x}_i$ with $\alpha_i > 0$ for each $i = 1, \ldots, k$ (why?).

Towards a contradiction, suppose that $k > d+1$ (remember that d is a dimension of our underlying space). Observe that the set

$$\{\mathbf{x}_1 - \mathbf{x}_k, \ldots, \mathbf{x}_{k-1} - \mathbf{x}_k\}$$

is linearly dependent because there are at least $d+1$ elements in the set but S is a subset of $\mathbb{R}^d$. This means that there exists $(\beta_1, \ldots, \beta_{k-1}) \neq \mathbf{0}$ such that

$$\sum_{i=1}^{k-1} \beta_i \left(\mathbf{x}_i - \mathbf{x}_k\right) = 0.$$

Define $\beta_k = -\sum_{i=1}^{k-1} \beta_i$. Observe that

$$\sum_{i=1}^{k} \beta_i \mathbf{x}_i = \beta_k \mathbf{x}_k + \sum_{i=1}^{k-1} \beta_i \mathbf{x}_i = \beta_k \mathbf{x}_k + \underbrace{\sum_{i=1}^{k-1} \beta_i \left(\mathbf{x}_i - \mathbf{x}_k\right)}_{=0} + \sum_{i=1}^{k-1} \beta_i \mathbf{x}_k = \mathbf{x}_k \underbrace{\left(\sum_{i=1}^{k} \beta_i\right)}_{=0} = \mathbf{0}.$$

Hence, for any $t > 0$, we can write

$$\mathbf{y} = \sum_{i=1}^{k} \alpha_i \mathbf{x}_i + t \sum_{i=1}^{k} \beta_i \mathbf{x}_i = \sum_{i=1}^{k} \left(\alpha_i - t\beta_i\right) \mathbf{x}_i.$$

Since k is finite and there exists at least one $\beta_i > 0$, the following is well defined:

$$\bar{t} \coloneqq \min \left\{ \frac{\alpha_i}{\beta_i} : \beta_i > 0,\ i \in \{1, \ldots, k\} \right\}.$$

Let $j \in \{1, \ldots, k\}$ be such that $\frac{\alpha_j}{\beta_j} = \bar{t}$ and define $\lambda_i \coloneqq \alpha_i - \bar{t}\beta_i$ for all $i \in \{1, \ldots, k\}$. Since $\bar{t} \leq \frac{\alpha_i}{\beta_i}$ for all $i \in \{1 \ldots, k\}$ with $\beta_i > 0$,

$$\lambda_i = \alpha_i - \bar{t}\beta_i \geq 0 \ \forall i \in \{1, \ldots, k\}.$$

We also have that $\sum_{i=1}^{k} \lambda_i = \sum_{i=1}^{k} \alpha_i - \bar{t} \sum_{i=1}^{k} \beta_i = 1$ and $\lambda_j = \alpha_j - \bar{t}\beta_j = 0$. Therefore,

$$\mathbf{y} = \sum_{i=1}^{k} \lambda_i \mathbf{x}_i = \sum_{i \in \{1, \ldots k\} \setminus \{j\}} \lambda_i \mathbf{x}_i.$$

[2]**Hint:** One way to prove Proposition 3 is to show that (a) the set on the right-hand side is a subset of $\operatorname{co}(S)$ and that (b) $\operatorname{co}(S)$ is a subset of the set on the right-hand side, and, hence, it has to be that the two are equal.)

But this means that $\mathbf{y}$ can be expressed as a convex combination of $k-1$ vectors which contradicts our choice of k. ■

Exercise 4. TFU: If $S \subseteq \mathbb{R}^d$ is open, then $\text{co}(S)$ is open.

Exercise 5. TFU: If $S \subseteq \mathbb{R}^d$ is closed, $\text{co}(S)$ is closed.

Proposition 4. *If $S \subseteq \mathbb{R}^d$ is a compact set, then $\text{co}(S)$ is compact.*

Exercise 6. Prove Proposition 4 using Carathéodory theorem.

Definition 4. The *closure* of a set $S \subseteq \mathbb{R}^d$, denoted $\text{cl}(S)$, is the smallest closed set that contains S; i.e.,

$$\text{cl}\,(S) \coloneqq \bigcap \left\{T \subseteq \mathbb{R}^d : S \subseteq T,\, T \text{ is closed}\right\}.$$

Remark 2. In the example above, we saw that $S = \{(x, y) \in \mathbb{R}^2 : yx^2 = 1\}$ is closed. Since S is closed, its closure is S itself. However, we saw that the convex hull of S is not closed. Hence, we realise that the closure of a convex hull of X (which must be closed) need not be the convex hull of its closure. So what is the closure of the convex hull of S?

Definition 5. The *closed convex hull* of a set $S \subseteq \mathbb{R}^d$, denoted $\overline{\text{co}}(S)$, is the smallest closed and convex set containing S; i.e.,

$$\overline{\text{co}}\,(S) \coloneqq \bigcap \left\{T \subseteq \mathbb{R}^d : S \subseteq T,\, T \text{ is convex and closed}\right\}.$$

Proposition 5. *The closed convex hull of a set S is the closure of the convex hull of S; i.e., $\overline{\text{co}}(S) = \text{cl}(\text{co}(S))$.*

Exercise 7 (PS5)**.** Prove Proposition 5.

Definition 6. Suppose $S \subseteq \mathbb{R}^d$ is convex. Then, $\mathbf{x} \in S$ is an *extreme point* of S if $\mathbf{x}$ cannot be written as a convex combination of two distinct points in S.

Theorem 2 (Krein-Milman)**.** *Suppose $S \subseteq \mathbb{R}^d$ is nonempty, convex and compact. Then, S is the closed convex hull of its extreme points.*

Remark 3. Thus, Krein-Milman tells us that convex and compact sets in $\mathbb{R}^d$ can be (almost) characterised by the extreme points of the set. Together with Carathéodory theorem, we realised that any point in a convex and compact subset in $\mathbb{R}^d$ can be written as a convex combination of at most $d+1$ of its extreme points.

Remark 4. Extreme points are also useful in characterising solutions of some optimisation problems (e.g., Bauer's Maximum Principle tells us sufficient conditions that ensures that optimum is attained at extreme points of the set of maximisers).

2 Hyperplanes

Definition 7. A hyperplane in $\mathbb{R}^d$ is given by

$$H\,(\mathbf{p}, a) \coloneqq \left\{\mathbf{x} \in \mathbb{R}^d : \mathbf{p} \cdot \mathbf{x} = a\right\}$$

for some $(\mathbf{p}, a) \in \mathbb{R}^d \times \mathbb{R}$. A *closed half-space above and below the hyperplane* $H(\mathbf{p}, a)$ are given, respectively, by

$$\overline{H}^+(\mathbf{p}, a) \coloneqq \left\{\mathbf{x} \in \mathbb{R}^d : \mathbf{p} \cdot \mathbf{x} \geq a\right\},$$
$$\overline{H}^-(\mathbf{p}, a) \coloneqq \left\{\mathbf{x} \in \mathbb{R}^d : \mathbf{p} \cdot \mathbf{x} \leq a\right\}.$$

The sets $\overline{H}^+(\mathbf{p}, a)$ and $\overline{H}^-(\mathbf{p}, a)$ are the two closed half-spaces *generated* by the hyperplane $H(\mathbf{p}, a)$. Open half-space above and below the hyperplane $H(\mathbf{p}, a)$ is defined by replacing the weak inequalities with strict inequalities.

In $\mathbb{R}^2$, any line can be characterised by two points that lie on the line: say $\mathbf{v} = (v_1, v_2)$ and $\mathbf{w} = (w_1, w_2)$. We can describe a line that connects these two points as

$$\{(v_1, v_2) + t(w_1 - v_1, w_2 - v_2) : t \in \mathbb{R}\}.$$

That is, we start from the point $(v_1, v_2) \in \mathbb{R}^2$ and move in the direction of $(w_1 - v_1, w_2 - v_2) \in \mathbb{R}^2$. Notice that above is almost equivalent to the usual formula for a line:

$$\begin{aligned}
&\left\{(z_1, z_2) \in \mathbb{R}^2 : z_1 - v_1 = t(w_1 - v_1),\ z_2 - v_2 = t(w_2 - v_2)\right\} \\
&= \left\{(z_1, z_2) \in \mathbb{R}^2 : z_1 - v_1 = t(w_1 - v_1),\ z_2 - v_2 = \frac{w_2 - v_2}{w_1 - v_1} t(w_1 - v_1)\right\} \\
&= \left\{(z_1, z_2) \in \mathbb{R}^2 : z_2 = v_2 + \frac{w_2 - v_2}{w_1 - v_1}(z_1 - v_1)\right\}.
\end{aligned}$$

We say almost because using the expression above would not allow us to consider direction along the first coordinate (we would be dividing by zero!). In an analogous way, we can think of a line in the linear space $\mathbb{R}^d$ as being described by a point $\mathbf{v} \in \mathbb{R}^d$ and a direction $\mathbf{w} \in \mathbb{R}^d$:

$$\left\{\mathbf{v} + t\mathbf{w} \in \mathbb{R}^d : t \in \mathbb{R}\right\} = \left\{\mathbf{x} \in \mathbb{R}^d : \exists t \in \mathbb{R},\ \mathbf{x} = \mathbf{v} + t\mathbf{w}\right\}.$$

Now let us think about a plane in $\mathbb{R}^3$. A point in $\mathbb{R}^3$ is described by (x, y, z), where each component corresponds to the coordinate on the x-, y- and z-axes. The easiest plane to think about is the plane that "runs" along two axes, say the x-axis and z-axis. Any points along this plane has y-coordinate of zero; however, the x and z coordinates can be anything; i.e.,

$$\left\{(x, y, z) \in \mathbb{R}^3 : y = 0\right\}.$$

Another way to describe this plane is to take any point on the plane, say the origin, and allow the points to move in the direction of the x- and z-axes but not in the direction of y-axis. Now, take any point on this plane, $(x, 0, z)$ for some $x, z \in \mathbb{R}$ and the standard basis representing the y-axis $e_2 = (0, 1, 0)$, and notice that $(x, 0, z) \cdot e_2 = 0$. That is, any point on this plane can be described using the dot/inner product.[3] Thus, we can now describe the plane that runs along the x-axis and z-axis as

$$\left\{\mathbf{0} + \mathbf{w} \in \mathbb{R}^3 : \mathbf{w} \cdot \mathbf{e}_2 = 0,\ \mathbf{w} \in \mathbb{R}^3\right\} = \left\{\mathbf{x} \in \mathbb{R}^3 : \mathbf{x} \cdot \mathbf{e}_2 = 0\right\}.$$

[3] Formally, we are using the fact that $\mathbb{R}^3$ is an inner product space so that $\mathbf{x}, \mathbf{y} \in \mathbb{R}^3$ are orthogonal to one another if $\mathbf{x} \cdot \mathbf{y} = 0$.

Of course, we can think about shifting the plane up and down the y-axis, which is the same as altering the starting point from $\mathbf{0}$. In this manner, we can think of a plane in $\mathbb{R}^d$, called a *hyperplane*, as being described by a starting point $\mathbf{v} \in \mathbb{R}^d$ and a direction $\mathbf{p} \in \mathbb{R}^d \backslash \{\mathbf{0}\}$:

$$\left\{\mathbf{v} + \mathbf{w} \in \mathbb{R}^d : \mathbf{w} \cdot \mathbf{p} = 0,\, \mathbf{w} \in \mathbb{R}^d\right\} = \left\{\mathbf{x} \in \mathbb{R}^d : \mathbf{p} \cdot (\mathbf{x} - \mathbf{v}) = 0\right\}.$$

Since $\mathbf{p} \cdot \mathbf{v}$ is a constant, letting $a \coloneqq \mathbf{p} \cdot \mathbf{v}$ gives the definition of hyperplanes above.

Remark 5. A plane may not always contain the origin. Similarly, a line may not contain the origin. Thus, in general, lines and hyperplanes are not subspaces of $\mathbb{R}^d$. But, we can make them into subspaces by translation; i.e., by adding a constant vector to all its elements. We call such sets *linear manifolds*. Formally, a *linear manifold* of $\mathbb{R}^d$ is a set $X \subseteq \mathbb{R}^d$ such that there is a subspace $S \subseteq \mathbb{R}^d$ and $\mathbf{x}_0 \in \mathbb{R}^d$ such that

$$X = S + \{\mathbf{x}_0\} \equiv \{\mathbf{x} + \mathbf{x}_0 : \mathbf{x} \in S\}.$$

Lines and hyperplanes are linear manifolds. To see this, consider a hyperplane described by a pair $(\mathbf{p}, \mathbf{x}_0)$:

$$\begin{aligned}\left\{\mathbf{z} \in \mathbb{R}^d : \mathbf{p} \cdot \mathbf{z} = \mathbf{p} \cdot \mathbf{x}_0\right\} &= \left\{\mathbf{z} \in \mathbb{R}^d : \mathbf{p} \cdot (\mathbf{z} - \mathbf{x}_0) = 0\right\} = \left\{\mathbf{x} + \mathbf{x}_0 \in \mathbb{R}^d : \mathbf{p} \cdot \mathbf{x} = 0\right\} \\ &= \left\{\mathbf{x} \in \mathbb{R}^d : \mathbf{p} \cdot \mathbf{x} = 0\right\} + \{\mathbf{x}_0\}\end{aligned}$$

where we used change of variable with $\mathbf{x} \coloneqq \mathbf{z} - \mathbf{x}_0$.

Remark 6. Observe that $\{\mathbf{x} \in \mathbb{R}^d : \mathbf{p} \cdot \mathbf{x} = 0\}$ is the null space of $f_{\mathbf{p}} : V \to \mathbb{R}$ defined as $f_{\mathbf{p}}(\mathbf{x}) \coloneqq \mathbf{p} \cdot \mathbf{x}$ and, as such, is a subspace of $\mathbb{R}^n$. One can define hyperspace in more general linear spaces using this notion.

2.1 Separating hyperplane theorems

Separating hyperplane theorems essentially are extensions of the fact that two disjoint convex sets in $\mathbb{R}^2$ can be separated by a line.

Remark 7. Separating hyperplane theorems are one of the most important results for economic theorists. They allow us to prove a lot of results in classical demand/supply theory and also are very useful in general equilibrium theory. Moreover, they are very important in statistics and lots of other fields. Here we just talk about the simplest versions of these theorems (without proofs) but you should know that far more general results are available (you can check out Hahn-Banach theorem in your free time).

Definition 8. Let $X, Y \subseteq \mathbb{R}^d$. The sets X and Y are *separated by the hyperplane* $H(\mathbf{p}, a)$ in $\mathbb{R}^d$ if X and Y lie in different closed half spaces generated by $H(\mathbf{p}, a)$.

Remark 8. For example, X and Y are *separated by the hyperplane* $H(\mathbf{p}, a)$ if

$$Y \subseteq \overline{H}^{+}(\mathbf{p}, a) \text{ and } X \subseteq \overline{H}^{-}(\mathbf{p}, a).$$

The condition above is equivalent to requiring

$$\mathbf{p} \cdot \mathbf{y} \geq a \ \forall \mathbf{y} \in Y \text{ and } \mathbf{p} \cdot \mathbf{x} \leq a \ \forall \mathbf{x} \in X. \tag{1}$$

Definition 9. Suppose X and Y are separated by the hyperplane $H(\mathbf{p}, a)$. The separation is said to be: (i) *proper* if at least one of the set is not contained in the hyperplane (i.e., there exists $\mathbf{x} \in X$ and $\mathbf{y} \in Y$ such that $\mathbf{p} \cdot \mathbf{x} \neq \mathbf{p} \cdot \mathbf{y}$); (ii) *strict* if (1) holds with strict inequalities;[4] and (iii) *strong* if X and Y are in disjoint closed half spaces.

Remark 9. Proper separation ensures that that $X \cup Y$ are not subsets of the hyperplane. Strict and strong separation mean that none of the elements in X and Y are in the separating hyperplane. Strong separation means that there is "gap" between the sets; i.e.,

$$\inf_{\mathbf{y} \in Y} \mathbf{p} \cdot \mathbf{y} > \sup_{\mathbf{x} \in X} \mathbf{p} \cdot \mathbf{x}.$$

Theorem 3 (Strong separating hyperplane theorem)**.** *Suppose X and Y are two nonempty, disjoint and convex subsets of $\mathbb{R}^d$. If X is compact and Y is closed, then there exists a hyperplane that strongly separates them.*

Proof. Define $f : X \to \mathbb{R}$ by $f(\mathbf{x}) = \inf\{\|\mathbf{x} - \mathbf{y}\| : \mathbf{y} \in Y\}$; i.e., $f(\mathbf{x})$ is the "distance" from a point $\mathbf{x}$ to the set X. The function is continuous.[5] Since X is compact, by the Extreme Value Theorem, f achieves a minimum on X at some point $\mathbf{x}^*$; i.e., $f(\mathbf{x}^*) = \inf f(X)$. Since $f(\mathbf{x}^*) = \inf\{\|\mathbf{x}^* - \mathbf{y}\| : \mathbf{y} \in Y\}$, there exists a sequence $(\mathbf{y}_n)_n$ in Y such that $\|\mathbf{x}^* - \mathbf{y}_n\| \to f(\mathbf{x}^*)$. And since Y is closed, $\mathbf{y}_n$ converges to some $\mathbf{y}^* \in Y$. Hence, $f(\mathbf{x}^*) = \|\mathbf{x}^* - \mathbf{y}^*\|$. Now define $\mathbf{p} = \mathbf{x}^* - \mathbf{y}^*$. Since X and Y are disjoint, $\mathbf{p} \neq \mathbf{0}$. Then,

$$0 < \|\mathbf{p}\|^2 = \mathbf{p} \cdot \mathbf{p} = \mathbf{p} \cdot (\mathbf{x}^* - \mathbf{y}^*) \Rightarrow \mathbf{p} \cdot \mathbf{x}^* > \mathbf{p} \cdot \mathbf{y}^*.$$

What remains to show is that $\mathbf{p} \cdot \mathbf{y}^* \geq \mathbf{p} \cdot \mathbf{y}$ for all $\mathbf{y} \in Y$ and $\mathbf{p} \cdot \mathbf{x}^* \leq \mathbf{p} \cdot \mathbf{x}$ for all $\mathbf{x} \in X$. So fix $\mathbf{y} \in Y$. Since $\mathbf{y}^*$ minimises the distance to $\mathbf{x}^*$ over Y, for any point $\mathbf{z} = \lambda \mathbf{y} + (1 - \lambda)\mathbf{y}^* = \mathbf{y}^* + \lambda(\mathbf{y} - \mathbf{y}^*)$ with $\lambda \in (0, 1]$ on the line segment between $\mathbf{y}$ and $\mathbf{y}^*$,

$$\begin{aligned}
(\mathbf{x}^* - \mathbf{z}) \cdot (\mathbf{x}^* - \mathbf{z}) &\geq (\mathbf{x}^* - \mathbf{y}^*) \cdot (\mathbf{x}^* - \mathbf{y}^*) \\
&\Leftrightarrow 0 \geq (\mathbf{x}^* - \mathbf{y}^*) \cdot (\mathbf{x}^* - \mathbf{y}^*) - (\mathbf{x}^* - \mathbf{z}) \cdot (\mathbf{x}^* - \mathbf{z}) \\
&\qquad = (\mathbf{x}^* - \mathbf{y}^*) \cdot (\mathbf{x}^* - \mathbf{y}^*) - (\mathbf{x}^* - (\mathbf{y}^* + \lambda (\mathbf{y} - \mathbf{y}^*))) \cdot (\mathbf{x}^* - (\mathbf{y}^* + \lambda (\mathbf{y} - \mathbf{y}^*))) \\
&\qquad = (\mathbf{x}^* - \mathbf{y}^*) \cdot (\mathbf{x}^* - \mathbf{y}^*) - (\mathbf{x}^* - \mathbf{y}^*) \cdot (\mathbf{x}^* - \mathbf{y}^*) + 2\lambda (\mathbf{x}^* - \mathbf{y}^*) \cdot (\mathbf{y} - \mathbf{y}^*) \\
&\qquad\qquad - \lambda^2 (\mathbf{y} - \mathbf{y}^*) \cdot (\mathbf{y} - \mathbf{y}^*) \\
&\Leftrightarrow 0 \geq 2\mathbf{p} \cdot (\mathbf{y} - \mathbf{y}^*) - \lambda (\mathbf{y} - \mathbf{y}^*) \cdot (\mathbf{y} - \mathbf{y}^*) .
\end{aligned}$$

Letting $\lambda \to 0$ from above, we conclude $\mathbf{p} \cdot \mathbf{y}^* \geq \mathbf{p} \cdot \mathbf{y}$. A similar argument for $\mathbf{x} \in X$ completes the proof. ■

[4] Some people define strict separation as meaning that X and Y are in disjoint open half-spaces generated by a hyperplane. Although this implies that (1) holds with strict inequalities, the converse is not true.

[5] To see this, for any $\mathbf{y}$, the Triangle inequality gives $\|\mathbf{x}' - \mathbf{y}\| \leq \|\mathbf{x}' - \mathbf{x}\| + \|\mathbf{x} - \mathbf{y}\|$ and $\|\mathbf{x} - \mathbf{y}\| \leq \|\mathbf{x} - \mathbf{x}'\| + \|\mathbf{x}' - \mathbf{y}\|$. Hence, $|\|\mathbf{x} - \mathbf{y}\| - \|\mathbf{x}' - \mathbf{y}\|| \leq \|\mathbf{x} - \mathbf{x}'\|$ so $|f(\mathbf{x}) - f(\mathbf{x}')| \leq \|\mathbf{x} - \mathbf{x}'\|$. Hence, f is actually Lipschitz continuous.

Exercise 8. Show by example that it is not enough for both sets to be closed for a strongly separating hyperplane to exist.

Corollary 2. *Suppose Y is a nonempty, closed and convex subset of $\mathbb{R}^d$. If $\mathbf{x} \in \mathbb{R}^d \backslash Y$, then $\{\mathbf{x}\}$ and Y are strongly separated by a hyperplane.*

Exercise 9. Prove the Corollary above.

Corollary 3. *Suppose Y is nonempty, closed and convex subset of $\mathbb{R}^d$. Then, exactly one of the following is true.*

(i) either, $\mathbf{x} \in Y$;

(ii) or, $\{\mathbf{x}\}$ and Y are strongly separated by a hyperplane; i.e., there exists $\mathbf{p} \in \mathbb{R}^d \backslash \{\mathbf{0}\}$ such that $\mathbf{p} \cdot \mathbf{x} > \sup_{\mathbf{y} \in X} \mathbf{p} \cdot \mathbf{y}$.

Theorem 4 (Separating hyperplane theorem)**.** *Suppose X and Y are two nonempty, disjoint and convex subsets of $\mathbb{R}^d$. Then, X and Y are separated by a hyperplane.*

Remark 10. Theorem 4 does not guarantee proper, strict or strong separation.

Remark 11. If both sets are open, then there exists a hyperplane that strictly separates them.

Exercise 10. Show by example that, for general nonempty convex disjoint sets X and Y, it may be necessary that both inequalities are allowed to be weak.

2.2 Supporting hyperplane theorems

From a separating hyperplane theorem, we can establish that, given a convex set and a point on its "border", there exists a hyperplane that is tangent to the set at that point. This latter result, called the *supporting hyperplane theorem* is useful in thinking about the idea of domination and it also pays a crucial role in many mathematical results (including the KKT theorem).

Definition 10. Given a set $S \subseteq \mathbb{R}^d$, a point $\mathbf{x} \in S$ is an *interior point* if there exists an open ball centred at $\mathbf{x}$ that is completely contained in S. Let $\text{int}(S)$ denote the set of all interior points of S. A point $\mathbf{x} \in S$ is a *boundary point* if $\mathbf{x} \in \text{cl}(S) \backslash \text{int}(S)$.

Definition 11. A set $X \subseteq \mathbb{R}^d$ is *bounded by a hyperplane* $H(\mathbf{p}, a)$ if X is entirely contained in one of the closed half spaces generated by $H(\mathbf{p}, a)$. A hyperplane $H(\mathbf{p}, a)$ is a *supporting hyperplane* for X if X is bounded by $H(\mathbf{p}, a)$ and $X \cap H(\mathbf{p}, a) \neq \varnothing$. A supporting hyperplane $H(\mathbf{p}, a)$ for X is *proper* if $\mathbf{p} \cdot \mathbf{x} \neq a$ for some $\mathbf{x} \in X$.

Remark 12. If $H(\mathbf{p}, a)$ is a supporting hyperplane for $X \subseteq \mathbb{R}^d$, then

$$\mathbf{p} \cdot \mathbf{x} \leq a \ \forall \mathbf{x} \in X$$

and there exists $\mathbf{x}_0 \in X$ such that

$$\mathbf{p} \cdot \mathbf{x}_0 = a.$$

Theorem 5 (Supporting hyperplane theorem)**.** *Suppose X is a nonempty, convex subset of $\mathbb{R}^d$, and let $\mathbf{x}_0 \in X$ be a boundary point of X; i.e., $\mathbf{x}_0 \in X \backslash \text{int}(X)$. Then, there exists a supporting hyperplane at $\mathbf{x}_0$. If, in addition, X has a nonempty interior, then the supporting hyperplane is proper.*

Proof. See *Minkowski Supporting Hyperplane Theorem* in Ok section G.3.2. ■

Corollary 4. *Suppose $X \subseteq \mathbb{R}^d$ is a nonempty, convex subset of $\mathbb{R}^d$ with a nonempty interior. Suppose that $\mathbf{y} \in \mathbb{R}^d$. Then exactly one of the following is true.*

(i) either, $\mathbf{y} \in \operatorname{int}(X)$;

(ii) or, we can find a supporting hyperplane; i.e., there exists $\mathbf{p} \in \mathbb{R}^d \backslash \{\mathbf{0}\}$ such that $\mathbf{p} \cdot \mathbf{y} \geq \sup_{\mathbf{x} \in X} \mathbf{p} \cdot \mathbf{x}$.

Proposition 6. *A closed and convex set $S \subseteq \mathbb{R}^d$ is the intersection of all closed half-spaces that contain it.*

Proof. Suppose $X \subseteq \mathbb{R}^d$ is closed and convex. Let $\mathcal{H}$ be the collection of closed half-spaces that contain X. We want to show that $X = \bigcap_{H \in \mathcal{H}} H$.

We first show that $X \subseteq \bigcap_{H \in \mathcal{H}} H$. For any $\mathbf{x} \in X$, using the supporting hyperplane theorem, we can find a hyperplane that generates a closed half space that contains $\mathbf{x}$ by picking any boundary point $\mathbf{x}_0 \in X$ and apply the supporting hyperplane theorem.

It remains to show that $\bigcap_{H \in \mathcal{H}} H \subseteq X$. We will prove the contrapositive statement; i.e., we will show that if $\mathbf{x} \notin X$, then $\mathbf{x} \notin \bigcap_{H \in \mathcal{H}} H$. Fix some $\mathbf{x} \notin X$. Since X is closed and convex, there is a hyperplane that strongly separates $\mathbf{x}$ from X. This hyperplane defines a closed half space H containing X. Hence, $\mathbf{x} \notin H$ implying that $\mathbf{x} \notin \bigcap_{H \in \mathcal{H}} H$. ■

Remark 13. The notion of *domination* captures the idea that one alternative is better than another under all circumstances. It turns out that separation by hyperplane is a useful way to mathematically capture the idea of domination. One example relates to the idea of dominant strategies in games. Here, we focus on a single decision-maker. Let $A \subseteq \mathbb{R}^d$ be a finite set of actions and Θ be a finite set of states. Let $u : A \times \Theta \to \mathbb{R}$ denote the decision-maker's utility. The goal is to show that a "pure" action is not dominated by some randomised decision rule if and only if it is optimal for some belief over state. Let $\Delta\Theta$ denote the set of all probability distributions over Θ; i.e., $\Delta\Theta \coloneqq \{p \in [0,1]^\Theta : \sum_{\theta \in \Theta} p(\theta) = 1\}$. Let ΔA denote the set of all probability distribution over A. Note that ΔA represents the set of all randomised decision rules. We now let X be the set of payoffs for each state that yields weakly lower payoffs than some randomisation of decision rules.

$$X \coloneqq \left\{ w \in \mathbb{R}^\Theta : \exists \sigma \in \Delta A \, w(\theta) \leq \sum_{a \in A} \sigma(a) \, u(a, \theta) \;\; \forall \theta \in \Theta \right\}.$$

In other words, X is the set of dominated payoffs. Note that X is nonempty and convex. Moreover, if $a \in A$ is dominated by some randomised decision rule, then

$$\exists \sigma \in \Delta A \, u(a, \theta) < \sum_{a \in A} \sigma(a) \, u(a, \theta) \;\; \forall \theta \in \Theta.$$

Observe that $a \in A$ being dominated means that the vector $u(a) = (u(a, \theta))_{\theta \in \Theta}$ lies in the interior of X. Thus, by Corollary 4, a is not dominated if and only if there is a supporting hyperplane at $u(a)$; i.e., there exists $\mathbf{v} \in \mathbb{R}^d \backslash \{\mathbf{0}\}$ such that

$$\mathbf{v} \cdot u(a) \geq \max \{\mathbf{v} \cdot \mathbf{w} : \mathbf{w} \in X\} \Leftrightarrow \sum_{\theta \in \Theta} v(\theta) \, u(a, \theta) \geq \max_{\mathbf{w} \in X} \sum_{\theta \in \Theta} v(\theta) \, w(\theta).$$

Because X is not bounded below and $\mathbf{v}$ is a nonzero vector, $\sum_{\theta\in\Theta} v(\theta) > 0$. So we can define $p(\theta) := \frac{v(\theta)}{\sum_{\theta\in\Theta} v(\theta)}$ for all $\theta \in \Theta$ so that $\mathbf{p} = (p(\theta))_{\theta\in\Theta} \in \Delta\Theta$. The maximum of the linear function $\mathbf{p}\cdot\mathbf{w}$ over $\mathbf{w} \in X$ is achieved at one of the extreme points in A, and the extreme points are all the vectors $u(a) = (u(a,\theta))_{\theta\in\Theta}$ for all $a \in A$. Hence,

$$\sum_{\theta\in\Theta} p(\theta)\, u(x,\theta) \geq \max_{\mathbf{w}\in X} \sum_{\theta\in\Theta} v(\theta)\, w(\theta) = \max_{a\in A} \sum_{\theta\in\Theta} p(\theta)\, u(a,\theta).$$

What we have now argued is that action $a^* \in A$ is not dominated if and only if there exists $\mathbf{p} \in \Delta\Theta$ such that

$$\sum_{\theta\in\Theta} p(\theta)\, u(a^*,\theta) \geq \sum_{\theta\in\Theta} p(\theta)\, u(a,\theta) \;\; \forall a \in A;$$

i.e., a^* is optimal for some belief $\mathbf{p}$ over Θ.

3 Convex and quasiconvex functions

Definition 12. Let $X \subseteq \mathbb{R}^d$ be convex. A function $f : X \to \mathbb{R}$ is...

- ▷ *concave* if $f(\alpha\mathbf{x} + (1-\alpha)\mathbf{y}) \geq \alpha f(\mathbf{x}) + (1-\alpha)f(\mathbf{y})$ for any $\mathbf{x}, \mathbf{y} \in X$ and $\alpha \in [0,1]$.
- ▷ *convex* if $f(\alpha\mathbf{x} + (1-\alpha)\mathbf{y}) \leq \alpha f(\mathbf{x}) + (1-\alpha)f(\mathbf{y})$ for any $\mathbf{x}, \mathbf{y} \in X$ and $\alpha \in [0,1]$.
- ▷ *strictly concave* if $f(\alpha\mathbf{x} + (1-\alpha)\mathbf{y}) > \alpha f(\mathbf{x}) + (1-\alpha)f(\mathbf{y})$ for any two distinct $\mathbf{x}, \mathbf{y} \in X$ and $\alpha \in (0,1)$.
- ▷ *strictly convex* if $f(\alpha\mathbf{x} + (1-\alpha)\mathbf{y}) < M\alpha f(\mathbf{x}) + (1-\alpha)f(\mathbf{y})$ for any two distinct $\mathbf{x}, \mathbf{y} \in X$ and $\alpha \in (0,1)$.

Remark 14. *A function is convex* (*resp. strictly convex*) if and only if $-f$ is concave (resp. strictly concave).

Define an epigraph and subgraph of a function $f : X \subseteq \mathbb{R}^d \to \mathbb{R}$ respectively as

$$\begin{aligned} \operatorname{epi}(f) &:= \{(\mathbf{x}, y) \in X \times \mathbb{R} : f(\mathbf{x}) \leq y\}, \\ \operatorname{sub}(f) &:= \{(\mathbf{x}, y) \in X \times \mathbb{R} : f(\mathbf{x}) \geq y\}. \end{aligned}$$

Proposition 7. *A function is concave (resp. convex) if and only if its subgraph (resp. epigraph) is convex.*

Exercise 11 (PS5)**.** Prove Proposition 7.

Corollary 5. *If $f : X \to \mathbb{R}$ is concave (resp. convex), then the set $\{\mathbf{x} \in X : f(\mathbf{x}) \geq r\}$ (resp. $\{\mathbf{x} \in X : f(\mathbf{x}) \leq r\}$) is convex for any $r \in \mathbb{R}$.*

Proof. By Proposition 7, if $f : X \to \mathbb{R}$ is concave, then we know that $\operatorname{sub}(f)$ is convex. Fix $r \in \mathbb{R}$ and take any $\mathbf{x}, \mathbf{x}' \in \{\mathbf{x} \in X : f(\mathbf{x}) \geq r\}$. Then, $(\mathbf{x}, r), (\mathbf{x}', r) \in \operatorname{sub}(f)$. Since $\operatorname{sub}(f)$ is convex,

$$(\alpha\mathbf{x} + (1-\alpha)\mathbf{x}', r) \in \operatorname{sub}(f) \;\; \forall \alpha \in [0,1].$$

That is,

$$f\left(\alpha\mathbf{x}+(1-\alpha)\mathbf{x}'\right)\geq r\ \forall\alpha\in[0,1].$$

Hence, $\alpha\mathbf{x}+(1-\alpha)\mathbf{x}'\in\{\mathbf{x}\in X: f(\mathbf{x})\geq r\}$ for all $\alpha\in[0,1]$. ∎

Definition. Let $X\subseteq\mathbb{R}^d$ be convex. A function $f: X\to\mathbb{R}$ is *affine* if $f(\mathbf{x})=\mathbf{a}\cdot\mathbf{x}+b$ for some $\mathbf{a}\in\mathbb{R}^d$ and $b\in\mathbb{R}$.

Exercise 12 (PS5)**.** Prove that an affine function is both convex and concave.

Proposition 8. *Let $X\subseteq\mathbb{R}^d$ be convex. If $f: X\to\mathbb{R}$ is concave or convex, then f is continuous on the interior of its domain.*

Proof. It suffices to prove the case for when X is open (because interior of X is always open and concavity of f on X implies concavity on the interior of X too). So suppose X is open and take $\mathbf{x}_0\in X$. Consider a sequence $(\mathbf{x}_n)_n$ in X that converges to $\mathbf{x}_0$. Since X is open, we can find an open ball centred at $\mathbf{x}_0$ with radius $\epsilon>0$; i.e., there exists $B_\epsilon(\mathbf{x}_0)\subseteq X$. Pick $\alpha\in(0,\epsilon)$ and let $A\subseteq B_\epsilon(\mathbf{x}_0)$ be defined by

$$A:=\left\{\mathbf{x}\in\mathbb{R}^d:\|\mathbf{x}-\mathbf{x}_0\|=\alpha\right\}.$$

There exists $N\in\mathbb{N}$ large enough (why?) so that $\|\mathbf{x}_n-\mathbf{x}_0\|<\alpha$ for all $n>N$. Then, for all $n>N$, there is $\mathbf{z}_n\in A$ such that $\mathbf{x}_n=\alpha_n\mathbf{x}_0+(1-\alpha_n)\mathbf{z}_n$ for some $\alpha_n\in(0,1)$. Because $\mathbf{x}_n\to\mathbf{x}_0$ and $\|\mathbf{z}_n-\mathbf{x}_0\|=\alpha>0$ for all $n\in\mathbb{N}$, it must be that $\alpha_n\to 1$. Therefore, by concavity of f,

$$f\left(\mathbf{x}_n\right)=f\left(\alpha_n\mathbf{x}_0+\left(1-\alpha_n\right)\mathbf{z}_n\right)\geq\alpha_n f\left(\mathbf{x}_0\right)+\left(1-\alpha_n\right)f\left(\mathbf{z}_n\right)\ \forall n>N.$$

Taking limits, we have

$$\liminf_{n\to\infty}f\left(\mathbf{x}_n\right)\geq f\left(\mathbf{x}_0\right).$$

Now, it is also true that, for all $n>\mathbb{N}$, there is $\mathbf{w}_n\in A$ and $\beta_n\in(0,1)$ such that $\mathbf{x}_0=\beta_n\mathbf{x}_n+(1-\beta_n)\mathbf{w}_n$. Then, by concavity of f,

$$\begin{aligned}
f\left(\mathbf{x}_0\right)=f\left(\beta_n\mathbf{x}_n+\left(1-\beta_n\right)\mathbf{w}_n\right)&\geq\beta_n f\left(\mathbf{x}_n\right)+\left(1-\beta_n\right)f\left(\mathbf{w}_n\right)\\
\Leftrightarrow\frac{1}{\beta_n}f\left(\mathbf{x}_0\right)-\frac{1-\beta_n}{\beta_n}f\left(\mathbf{w}_n\right)&\geq f\left(\mathbf{x}_n\right)
\end{aligned}$$

Because $\beta_n\to 1$, by taking limits, we obtain

$$f\left(\mathbf{x}_0\right)\geq\limsup_{n\to\infty}f\left(\mathbf{x}_n\right).$$

Hence, we have shown that

$$\liminf_{n\to\infty}f\left(\mathbf{x}_n\right)\geq f\left(\mathbf{x}_0\right)\geq\limsup_{n\to\infty}f\left(\mathbf{x}_n\right).$$

Recalling that $\limsup_{n\to\infty}f(\mathbf{x}_n)\geq\liminf_{n\to\infty}f(\mathbf{x}_n)$, above implies that above inequalities must in fact be equalities. Hence, $\lim_{n\to\infty}f(\mathbf{x}_n)=f(\mathbf{x}_0)$. ∎

Definition 13. Given a function $f: X\subseteq\mathbb{R}^d\to\mathbb{R}$, an element $\mathbf{x}\in X$ is a *global maximum* of f if

$$f\left(\mathbf{x}\right)\geq f\left(\mathbf{y}\right)\ \forall\mathbf{y}\in X\backslash\left\{\mathbf{x}\right\}.$$

The element $\mathbf{x} \in X$ is a strict global maximum if the inequality holds strictly above (and if it exists, it must be unique). An element $\mathbf{x} \in X$ is a *local maximum* of f if, for some $\epsilon > 0$,

$$f(\mathbf{x}) \geq f(\mathbf{y}) \;\; \forall \mathbf{y} \in B_\epsilon(\mathbf{x}) \setminus \{\mathbf{x}\}.$$

The element $\mathbf{x} \in X$ is a strict local maximum if the inequality holds strictly above. Local minima are defined analogously with the inequalities reversed.

Proposition 9. *Let $X \subseteq \mathbb{R}^d$ be convex. If $f : X \to \mathbb{R}$ is concave, then any local maximum of f is a global maximum of f. Moreover, the set of maximisers*

$$\{\mathbf{x} \in X : f(\mathbf{x}) = \sup f(X)\}$$

is convex.

Proof. Suppose f admits a local maximum at $\mathbf{x} \in X$ that is not a global maximum. By definition of a local maximum, there is $\epsilon > 0$ such that $f(\mathbf{x}) \geq f(\mathbf{y})$ for all $\mathbf{y} \in B_\epsilon(\mathbf{x}) \cap X$. Since $\mathbf{x}$ is not a global maximum, there is $\mathbf{z} \in X$ such that $f(\mathbf{z}) > f(\mathbf{x})$. Since X ix convex, $\alpha\mathbf{x} + (1-\alpha)\mathbf{z} \in X$ for all $\alpha \in [0,1]$. We may pick α sufficiently close to one (but not one) so that $\alpha\mathbf{x} + (1-\alpha)\mathbf{z} \in B_\epsilon(\mathbf{x})$. By concavity of f,

$$f(\alpha\mathbf{x} + (1-\alpha)\mathbf{z}) \geq \alpha f(\mathbf{x}) + (1-\alpha) f(\mathbf{z}) > f(\mathbf{x}).$$

But this is a contradiction because $\alpha\mathbf{x} + (1-\alpha)\mathbf{z} \in B_\epsilon(\mathbf{x})$ so that $f(\mathbf{x}) \geq f(\alpha\mathbf{x} + (1-\alpha)\mathbf{z})$.

To see the second part, suppose $\mathbf{x}_1$ and $\mathbf{x}_2$ are both maximisers of f on X. It must be that $f(\mathbf{x}_1) = f(\mathbf{x}_2)$ and concavity of f implies that, for any $\alpha \in [0,1]$,

$$f(\alpha\mathbf{x}_1 + (1-\alpha)\mathbf{x}_2) \geq \alpha f(\mathbf{x}_1) + (1-\alpha) f(\mathbf{x}_2) = f(\mathbf{x}_1).$$

Above must hold with equality because $\mathbf{x}_1$ and $\mathbf{x}_2$ are maximisers. It follows that the set of maximisers must be convex. Note also that an empty set is trivially convex so that the statement is true even if the set of maximisers is empty. ■

Remark 15. Analogous result hold when f is convex with respect to minimisers.

Definition 14. Let $X \subseteq \mathbb{R}^d$ be convex. A function $f : X \to \mathbb{R}$ is...

- ▷ *quasi-concave* if $f(\alpha x + (1-\alpha)y) \geq \min\{f(x), f(y)\}$ for any $x, y \in X$ and $\alpha \in [0,1]$.
- ▷ *quasi-convex* if $f(\alpha x + (1-\alpha)y) \leq \max\{f(x), f(y)\}$ for any $x, y \in X$ and $\alpha \in [0,1]$.
- ▷ *strictly* quasi-concave if $f(\alpha x + (1-\alpha)y) > \min\{f(x), f(y)\}$ for all distinct $x, y \in X$ and $\alpha \in (0,1)$.
- ▷ *strictly* quasi-convex if $f(\alpha x + (1-\alpha)y) < \max\{f(x), f(y)\}$ for all distinct $x, y \in X$ and $\alpha \in (0,1)$.

Remark 16. A function f is *(strictly) quasi-convex* if $-f$ is (strictly) quasi-concave.

Exercise 13 (PS5)**.** Prove the following: Let $X \subseteq \mathbb{R}^d$ be convex and let $f : X \to \mathbb{R}$. Then, f is quasiconcave (resp. quasi-convex) if and only if the upper (resp. lower) contour sets are convex; i.e., $\{\mathbf{x} \in X : f(\mathbf{x}) \geq r\}$ (resp. $\{\mathbf{x} \in X : f(\mathbf{x}) \leq r\}$) are convex for any $r \in \mathbb{R}$.

Exercise 14 (PS5)**.** TFU: If f is a quasi-concave function and $h : \mathbb{R} \to \mathbb{R}$ is nondecreasing function then $h \circ f$ is quasi-concave. Do the same replacing quasi-concave with concave (in both places).

Proposition 10. *Let $X \subseteq \mathbb{R}^d$ be convex, If $f : X \to \mathbb{R}$ is quasi-concave, the set of maximisers*

$$\{\mathbf{x} \in X : f(\mathbf{x}) = \sup f(X)\}$$

is convex. If f is strictly quasi-concave, any local maximum f is a global maximum of f on X and the set of maximisers contains at most one element.

Proof. Suppose $\mathbf{x}_1$ and $\mathbf{x}_2$ are both maximisers of f on X. It must be that $f(\mathbf{x}_1) = f(\mathbf{x}_2)$ and quasi-concavity of f implies that, for any $\alpha \in [0, 1]$,

$$f(\alpha \mathbf{x}_1 + (1 - \alpha) \mathbf{x}_2) \geq \min \{f(\mathbf{x}_1), f(\mathbf{x}_2)\} = f(\mathbf{x}_1).$$

Above must hold with equality because $\mathbf{x}_1$ and $\mathbf{x}_2$ are maximisers. It follows that the set of maximisers must be convex. Note also that an empty set is trivially convex so that the statement is true even if the set of maximisers is empty.

Now suppose f is strictly quasiconcave and the $\mathbf{x} \in X$ is a local maximum of f on X. Thus, there exists $\epsilon > 0$ such that $f(\mathbf{x}) \geq f(\mathbf{y})$ for all $\mathbf{y} \in B_\epsilon(\mathbf{x})$. Toward a contradiction, suppose that $\mathbf{x}$ is not a global maximum; i.e., there exists $\mathbf{z} \in X$ such that $f(\mathbf{z}) > f(\mathbf{x})$. But then since X is convex, for any $\alpha \in [0, 1]$, $\alpha\mathbf{x} + (1 - \alpha)\mathbf{z} \in X$ and by the quasiconcavity of f, we have

$$f(\alpha \mathbf{x} + (1 - \alpha) \mathbf{z}) > \min \{f(\mathbf{x}), f(\mathbf{z})\} = f(\mathbf{x}) \ \ \forall \alpha \in (0, 1).$$

In particular, for sufficiently large $\alpha > 0$, $\alpha\mathbf{x} + (1 - \alpha)\mathbf{z} \in B_r(\mathbf{x})$ and hence we must have $f(\mathbf{x}) \geq f(\alpha\mathbf{x} + (1 - \alpha)\mathbf{z})$; a contradiction.

Suppose that there are two global maximisers, $\mathbf{x}, \mathbf{y} \in X$. Pick any $\alpha \in (0, 1)$ and define $\mathbf{z} \coloneqq \alpha\mathbf{x} + (1 - \alpha)\mathbf{y} \in X$. By strict quasiconcavity,

$$f(\mathbf{z}) > \min \{f(\mathbf{x}), f(\mathbf{y})\} = f(\mathbf{x}) = f(\mathbf{y}).$$

But this contradicts the fact that $\mathbf{x}$ and $\mathbf{y}$ were global maximisers. ∎

Exercise 15. TFU: Let $X \subseteq \mathbb{R}^d$ be convex, If $f : X \to \mathbb{R}$ is quasi-concave, then any local maximum f is a global maximum of f on X .