MATH3027: Optimization (UK 22/23)

Week 8: Convex Sets

Prof. Richard Wilkinson
School of Mathematical Sciences
University of Nottingham, United Kingdom
Please send any comments or mistakes to
r.d.wilkinson@nottingham.ac.uk

So far we have only considered unconstrained optimization problems. This week we begin our study of **constrained optimization** by discussing one of the most important concepts in continuous optimization: convexity. Solving general constrained optimization problems is hard[1], whereas in contrast, convex problems can be solved reliably and relatively quickly. Convex problems include linear programmes[2] and constrained least-squares problems, as well as many other types of problem that naturally occur throughout science, engineering, economics etc.

This week, we focus on the properties of convex sets, establishing results that will be useful in the forthcoming weeks. In particular, we will study the properties of the convex hull, and the structure of convex polytopes.

Convex Sets

We begin discussing convexity by defining what a convex set is.

[1] NP hard for those of you who know what that means.
[2] Linear programmes are of the form

$$\begin{aligned} \max_{\mathbf{x}} \quad & \mathbf{c}^\top \mathbf{x} \\ \text{subject to } & \mathbf{A}\mathbf{x} \leq \mathbf{b} \\ & \mathbf{x} \geq \mathbf{0} \end{aligned}$$

Definition (Convex Set)**.** *A set $C \subseteq \mathbb{R}^n$ is called convex if for any $\mathbf{x}, \mathbf{y} \in C$ and $\lambda \in [0,1]$ the point $\lambda\mathbf{x} + (1-\lambda)\mathbf{y}$ belongs to C.*

In other words, the above definition is equivalent to saying that for any $\mathbf{x}, \mathbf{y} \in C$, the line segment $[\mathbf{x}, \mathbf{y}]$ is also in C. See Figure 1.

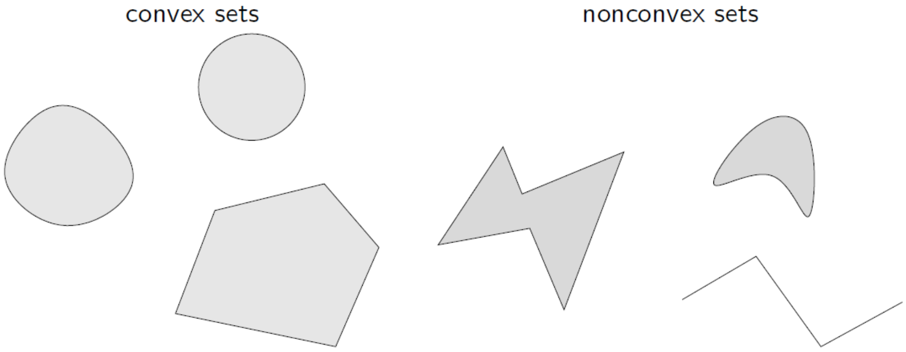

Figure 1: Examples of convex and nonconvex sets.

Important Convex Sets

- $\mathbb{R}^n$
- Lines in $\mathbb{R}^n$, which are of the form

$$L = \{\mathbf{z} + t\mathbf{d} : t \in \mathbb{R}\}$$

  where $\mathbf{z}, \mathbf{d} \in \mathbb{R}^n$ and $\mathbf{d} \neq \mathbf{0}$.

- The open and closed line segments $[\mathbf{x}, \mathbf{y}]$, $(\mathbf{x}, \mathbf{y})$ for $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n (\mathbf{x} \neq \mathbf{y})$, and more generally, the open balls

$$B(\mathbf{c}, r) = \{\mathbf{x} \in \mathbb{R}^n : \|\mathbf{x} - \mathbf{c}\| < r\} \ ,$$

  and closed balls

$$B[\mathbf{c}, r] = \{\mathbf{x} \in \mathbb{R}^n : \|\mathbf{x} - \mathbf{c}\| \leq r\} \ ,$$

  are convex. Note that this is true for an arbitrary norm defined over $\mathbb{R}^n$, and that the shape of the set depends upon the particular norm. See Figure 2.

- Hyperplanes, which are sets of the form

$$H = \left\{\mathbf{x} \in \mathbb{R}^n : \mathbf{a}^\top \mathbf{x} = b\right\} \quad (\mathbf{a} \in \mathbb{R}^n \backslash \{\mathbf{0}\}, b \in \mathbb{R}) \ .$$

- The associated half-space is the set

$$H^- = \left\{\mathbf{x} \in \mathbb{R}^n : \mathbf{a}^\top \mathbf{x} \leq b\right\} \ .$$

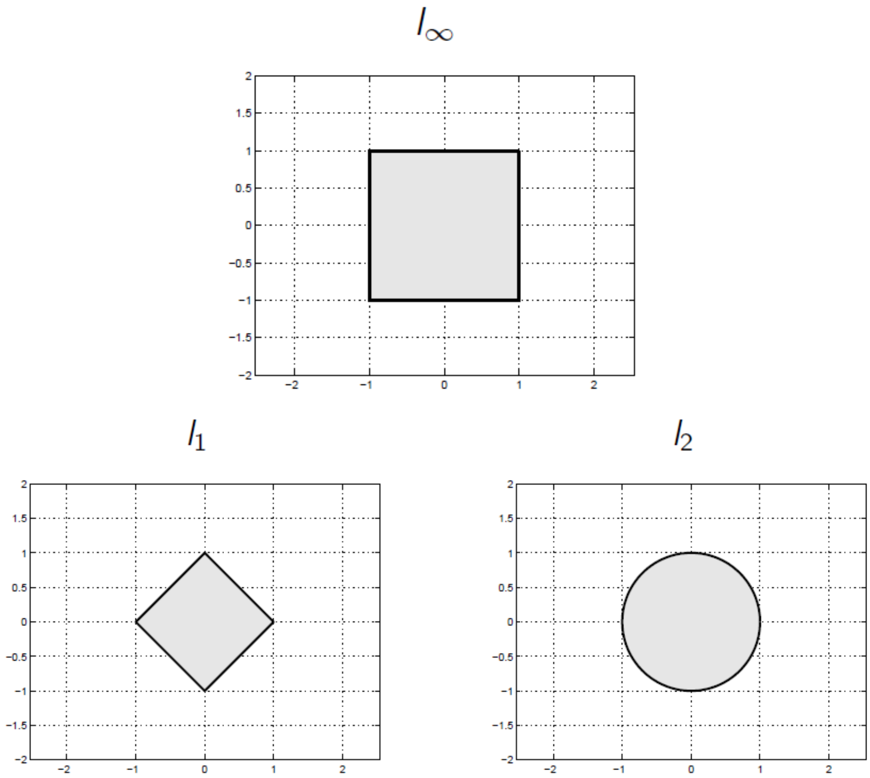

Figure 2: Different unit balls $\|\mathbf{x}\|_p \leq 1$ in $\mathbb{R}^2$.

- Ellipsoids, which can be written as

$$E = \left\{ \mathbf{x} \in \mathbb{R}^n : \mathbf{x}^\top \mathbf{Q}\mathbf{x} + 2\mathbf{b}^\top \mathbf{x} + c \leq 0 \right\} \, ,$$

  where $\mathbf{Q} \in \mathbb{R}^{n \times n}$ is positive semidefinite, $\mathbf{b} \in \mathbb{R}^n$ and $c \in \mathbb{R}$.

We'll prove the result for ellipsoids. Prove the other sets are convex. 

Proof. Write E as $E = \{\mathbf{x} \in \mathbb{R}^n : f(\mathbf{x}) \leq 0\}$ where $f(\mathbf{x}) \equiv \mathbf{x}^\top \mathbf{Q}\mathbf{x} + 2\mathbf{b}^\top \mathbf{x} + c$. Then, take $\mathbf{x}, \mathbf{y} \in E$ and $\lambda \in [0, 1]$, and $f(\mathbf{x}) \leq 0, f(\mathbf{y}) \leq 0$. The vector $z = \lambda \mathbf{x} + (1 - \lambda)\mathbf{y}$ satisfies

$$\mathbf{z}^\top \mathbf{Q}\mathbf{z} = \lambda^2 \mathbf{x}^\top \mathbf{Q}\mathbf{x} + (1-\lambda)^2 \mathbf{y}^\top \mathbf{Q}\mathbf{y} + 2\lambda(1-\lambda)\mathbf{x}^\top \mathbf{Q}\mathbf{y} \, ,$$

and using Cauchy-Schwartz it follows that

$$\begin{aligned} \mathbf{x}^\top \mathbf{Q}\mathbf{y} \leq \left\| \mathbf{Q}^{1/2}\mathbf{x} \right\| \cdot \left\| \mathbf{Q}^{1/2}\mathbf{y} \right\| &= \sqrt{\mathbf{x}^\top \mathbf{Q}\mathbf{x}} \sqrt{\mathbf{y}^\top \mathbf{Q}\mathbf{y}} \\ &\leq \frac{1}{2} \left(\mathbf{x}^\top \mathbf{Q}\mathbf{x} + \mathbf{y}^\top \mathbf{Q}\mathbf{y} \right) \, . \end{aligned}$$

Can you see why the second inequality holds? 

This implies that

$$\mathbf{z}^\top \mathbf{Q}\mathbf{z} \leq \lambda \mathbf{x}^\top \mathbf{Q}\mathbf{x} + (1-\lambda)\mathbf{y}^\top \mathbf{Q}\mathbf{y} \, .$$

Finally,

$$\begin{aligned}
f(\mathbf{z}) &= \mathbf{z}^\top \mathbf{Q}\mathbf{z} + 2\mathbf{b}^\top \mathbf{z} + c \\
&\le \lambda \mathbf{x}^\top \mathbf{Q}\mathbf{x} + (1-\lambda)\mathbf{y}^\top \mathbf{Q}\mathbf{y} + 2\lambda \mathbf{b}^\top \mathbf{x} + 2(1-\lambda)\mathbf{b}^\top \mathbf{y} + c \\
&= \lambda \left(\mathbf{x}^\top \mathbf{Q}\mathbf{x} + 2\mathbf{b}^\top \mathbf{x} + c\right) + (1-\lambda)\left(\mathbf{y}^\top \mathbf{Q}\mathbf{y} + 2\mathbf{b}^\top \mathbf{y} + c\right) \\
&= \lambda f(\mathbf{x}) + (1-\lambda) f(\mathbf{y}) \le 0
\end{aligned}$$

establishing the desired result that $\mathbf{z} \in E$. □

Algebraic Operations with Convex Sets

Convexity is preserved under addition, Cartesian products, intersections, and linear maps.

Lemma (Intersection of convex sets is convex)**.** *Let $C_i \subseteq \mathbb{R}^n$ be a convex set for any $i \in I$ where I is an index set (possibly infinite). Then the set $\bigcap_{i \in I} C_i$ is convex.*

Some important algebraic properties of convex sets are summarized in the following result:

Theorem.

1. *Let $C_1, C_2, \ldots, C_k \subseteq \mathbb{R}^n$ be convex sets and let $\mu_1, \mu_2, \ldots, \mu_k \in \mathbb{R}$. Then the set*
$$\mu_1 C_1 + \mu_2 C_2 + \ldots + \mu_k C_k = \left\{ \sum_{i=1}^{k} \mu_i \mathbf{x}_i : \mathbf{x}_i \in C_i, \mu_i \in \mathbb{R} \right\}$$
is convex.
2. *Let $C_i \subseteq \mathbb{R}^{k_i}, i = 1, \ldots, m$ be convex sets. Then the Cartesian product*
$$C_1 \times C_2 \times \cdots \times C_m = \{(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m) : \mathbf{x}_i \in C_i, i = 1, 2, \ldots, m\}$$
is convex.
3. *Let $M \subseteq \mathbb{R}^n$ be a convex set and let $\mathbf{A} \in \mathbb{R}^{m \times n}$. Then the set*
$$\mathbf{A}(M) = \{\mathbf{A}\mathbf{x} : \mathbf{x} \in M\}$$
is convex.
4. *Let $D \subseteq \mathbb{R}^m$ be convex and let $\mathbf{A} \in \mathbb{R}^{m \times n}$. Then the set*
$$\mathbf{A}^{-1}(D) = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{A}\mathbf{x} \in D\}$$
is convex.

E.3 Prove 3. and 4.

The Convex Hull

Definition (Convex Combinations). *Given m points $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m \in \mathbb{R}^n$, a convex combination of these m points is a vector of the form $\lambda_1\mathbf{x}_1 + \lambda_2\mathbf{x}_2 + \cdots + \ldots + \lambda_m\mathbf{x}_m$, where $\lambda_1, \lambda_2, \ldots, \lambda_m$ are nonnegative numbers satisfying $\lambda_1 + \lambda_2 + \ldots + \lambda_m = 1$ (i.e., $\boldsymbol{\lambda} \in \Delta_m$).*

We defined a convex set to be a set for which any convex combination of two points from the set is also in the set. We will now show that a convex combination of any number of points from a convex set is in the set.

Theorem. *Let $C \subseteq \mathbb{R}^n$ be a convex set and let $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m \in C$. Then for any $\boldsymbol{\lambda} \in \Delta_m$, the relation $\sum_{i=1}^m \lambda_i\mathbf{x}_i \in C$ holds.*

Proof. Proof by induction on m. For $m = 1$ the result is obvious. The induction hypothesis is that for any m vectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m \in C$ and any $\boldsymbol{\lambda} \in \Delta_m$, the vector $\sum_{i=1}^m \lambda_i\mathbf{x}_i$ belongs to C. We will now prove the theorem for $m+1$ vectors. Suppose that $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_{m+1} \in C$ and that $\boldsymbol{\lambda} \in \Delta_{m+1}$. We will show that $\mathbf{z} \equiv \sum_{i=1}^{m+1} \lambda_i\mathbf{x}_i \in C$. For this, if $\lambda_{m+1} = 1$, then $\mathbf{z} = \mathbf{x}_{m+1} \in C$ and the result obviously follows. Otherwise, if $\lambda_{m+1} < 1$, then

$$\begin{aligned}
\mathbf{z} &= \sum_{i=1}^{m} \lambda_i\mathbf{x}_i + \lambda_{m+1}\mathbf{x}_{m+1} \\
&= (1-\lambda_{m+1}) \sum_{i=1}^{m} \frac{\lambda_i}{1-\lambda_{m+1}}\mathbf{x}_i + \lambda_{m+1}\mathbf{x}_{m+1}\,.
\end{aligned}$$

Since $\sum_{i=1}^m \frac{\lambda_i}{1-\lambda_{m+1}} = \frac{1-\lambda_{m+1}}{1-\lambda_{m+1}} = 1$, it follows that $\mathbf{v} = \sum_{i=1}^m \frac{\lambda_i}{1-\lambda_{m+1}}\mathbf{x}_i$ is a convex combination of m points from C, and hence by the induction hypotheses we have that $\mathbf{v} \in C$. Thus, by the definition of a convex set, $\mathbf{z} = (1-\lambda_{m+1})\,\mathbf{v} + \lambda_{m+1}\mathbf{x}_{m+1} \in C$. □

Definition (The Convex Hull). *Let $S \subseteq \mathbb{R}^n$. The convex hull of S, denoted by $\operatorname{conv}(S)$, is the set comprising all the convex combinations of vectors from S:*

$$\operatorname{conv}(S) \equiv \left\{ \sum_{i=1}^{k} \lambda_i\mathbf{x}_i : \mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k \in S, \boldsymbol{\lambda} \in \Delta_k \right\}$$

The convex hull $\operatorname{conv}(S)$ is "smallest" convex set containing S, in the sense that if another convex set T contains S, then $\operatorname{conv}(S) \subset T$.

The following well-known result, called the Caretheódory theorem, states that any element in the convex hull of a subset of a given set $S \subset \mathbb{R}^n$ can be expressed as a **convex combination** of no more than $n+1$ vectors from S.

Theorem (Carathéodory). *Let $S \subseteq \mathbb{R}^n$ and let $\mathbf{x} \in \operatorname{conv}(S)$. Then, there exist $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_{n+1} \in S$ such that $\mathbf{x} \in \operatorname{conv}(\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_{n+1}\})$, that is, there exist $\lambda \in \Delta_{n+1}$ such that*

$$\mathbf{x} = \sum_{i=1}^{n+1} \lambda_i\mathbf{x}_i$$

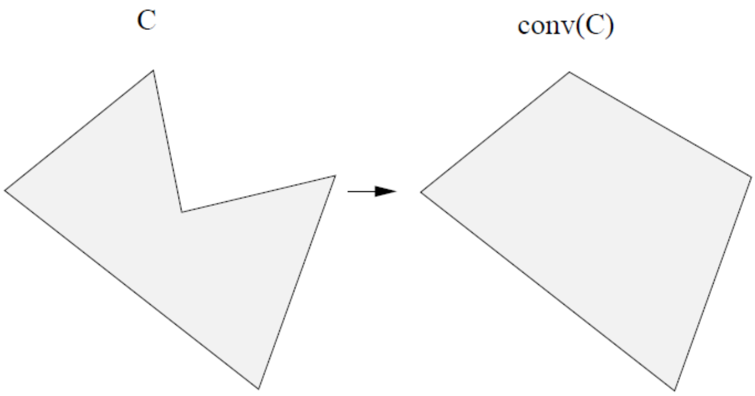

Figure 3: The convex hull of a non-convex set.

We present this proof as it provides a construction mechanism.

Proof. Let $\mathbf{x} \in \mathrm{conv}(S)$. Then, there exist vectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k \in S$ and $\boldsymbol{\lambda} \in \Delta_k$ such that

$$\mathbf{x} = \sum_{i=1}^{k} \lambda_i \mathbf{x}_i \,.$$

We can assume that $\lambda_i > 0$ for all $i = 1, 2, \ldots, k$. If $k \leq n+1$, the result is proven. Otherwise, if $k \geq n+2$, then the vectors $\mathbf{x}_2 - \mathbf{x}_1, \mathbf{x}_3 - \mathbf{x}_1, \ldots, \mathbf{x}_k - \mathbf{x}_1$, are necessarily linearly dependent (as they are in $\mathbb{R}^n$ and there are at least $n+1$ of them) implying that there exist $\mu_2, \mu_3, \ldots, \mu_k$ not all zeros such that

$$\sum_{i=2}^{k} \mu_i \left(\mathbf{x}_i - \mathbf{x}_1\right) = \mathbf{0} \,.$$

Defining $\mu_1 = -\sum_{i=2}^{k} \mu_i$, we obtain that

$$\sum_{i=1}^{k} \mu_i \mathbf{x}_i = \mathbf{0} \,.$$

Note that not all of the coefficients $\mu_1, \mu_2, \ldots, \mu_k$ are zeros and $\sum_{i=1}^{k} \mu_i = 0$. Thus, there exists an index i for which $\mu_i < 0$. Let $\alpha \in \mathbb{R}_+$. Then,

$$\mathbf{x} = \sum_{i=1}^{k} \lambda_i \mathbf{x}_i = \sum_{i=1}^{k} \lambda_i \mathbf{x}_i + \alpha \sum_{i=1}^{k} \mu_i \mathbf{x}_i = \sum_{i=1}^{k} \left(\lambda_i + \alpha \mu_i\right) \mathbf{x}_i \,.$$

We have $\sum_{i=1}^{k} \left(\lambda_i + \alpha \mu_i\right) = 1$, so the equation above is a convex combination if and only if

$$\lambda_i + \alpha \mu_i \geq 0 \text{ for all } i = 1, \ldots, k.$$

But since $\lambda_i > 0$ for all i, it follows that these inequalities are satisfied for $\alpha = \min_{i:\mu_i<0} \left\{-\frac{\lambda_i}{\mu_i}\right\}$. In addition, we must have $\lambda_j + \varepsilon \mu_j = 0$ for $j \in \underset{i:\mu_i<0}{\operatorname{argmin}} \left\{-\frac{\mu_i}{\lambda_i}\right\}$. This means that we found a representation of $\mathbf{x}$ as a convex combination of $k-1$ (or less) vectors. This process can be carried on until a representation of $\mathbf{x}$ as a convex combination of no more than $n+1$ vectors is derived. $\square$

Example E.4 For $n = 2$, consider the four vectors

$$\mathbf{x}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \mathbf{x}_2 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \mathbf{x}_3 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \mathbf{x}_4 = \begin{pmatrix} 2 \\ 2 \end{pmatrix}$$

and let $\mathbf{x} \in \operatorname{conv}(\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4\})$ be given by

$$\mathbf{x} = \frac{1}{8}\mathbf{x}_1 + \frac{1}{4}\mathbf{x}_2 + \frac{1}{2}\mathbf{x}_3 + \frac{1}{8}\mathbf{x}_4 = \begin{pmatrix} \frac{13}{8} \\ \frac{11}{8} \end{pmatrix}$$

Find a representation of $\mathbf{x}$ as a convex combination of no more than 3 vectors.

Convex Cones

A set S is called a cone if it satisfies the following property:

$$\lambda\mathbf{x} \in S \text{ for any } \mathbf{x} \in S \text{ and } \lambda \geq 0.$$

The following lemma shows that there is a very simple and elegant characterization of **convex** cones.

Lemma. *A set S is a convex cone if and only if the following properties hold:*

- $\mathbf{x}, \mathbf{y} \in S \Rightarrow \mathbf{x} + \mathbf{y} \in S.$
- $\mathbf{x} \in S, \lambda \geq 0 \Rightarrow \lambda\mathbf{x} \in S.$

E.5 Prove this lemma.

Examples of Convex Cones

- The convex **polytope**

  $$C = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{A}\mathbf{x} \leq \mathbf{0}\},$$

  where $\mathbf{A} \in \mathbb{R}^{m \times n}$.

- Lorentz Cone or ice cream cone is given by

$$L^n = \left\{ \begin{pmatrix} \mathbf{x} \\ t \end{pmatrix} \in \mathbb{R}^{n+1} : \|\mathbf{x}\| \leq t, \mathbf{x} \in \mathbb{R}^n, t \in \mathbb{R} \right\}.$$

  See Figure 4.

- Nonnegative polynomials. The set consisting of all possible coefficients of polynomials of degree $n-1$ which are nonnegative over $\mathbb{R}$:

$$K^n = \left\{ \mathbf{x} \in \mathbb{R}^n : x_1 t^{n-1} + x_2 t^{n-2} + \ldots + x_{n-1} t + x_n \geq 0\,,\ \forall t \in \mathbb{R} \right\}.$$

E.6

Prove these are convex cones.

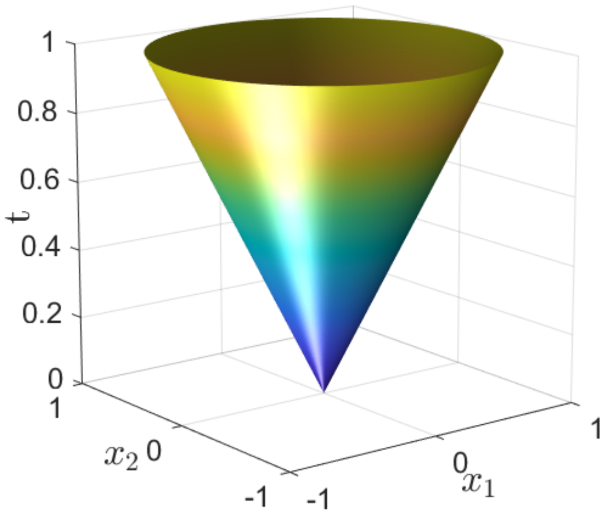

Figure 4: An example of the Lorentz cone in $2+1$ dimensions, here the z axis plays the role of t.

The Conic Hull

Given m points $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_m \in \mathbb{R}^n$, a conic[3] combination of these m points is a vector of the form $\lambda_1\mathbf{x}_1 + \lambda_2\mathbf{x}_2 + \cdots + \lambda_m\mathbf{x}_m$, where $\lambda \in \mathbb{R}^m_+$. The definition of the conic hull now follows:

Definition (Conic Hull)**.** *Let $S \subseteq \mathbb{R}^n$, then the conic hull of S, denoted by* $\operatorname{cone}(S)$ *is the set comprising all the conic combinations of vectors from S :*

$$\operatorname{cone}(S) \equiv \left\{ \sum_{i=1}^{k} \lambda_i \mathbf{x}_i : \mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k \in S, \boldsymbol{\lambda} \in \mathbb{R}^k_+ \right\}$$

[3] Conic, as in, relating to cones

Note that $\text{cone}(S)$ is a cone, and it is convex.

Similarly to the convex hull, the conic hull of a set S is the smallest cone containing S. If $S \subseteq T$ for some convex cone T, then cone $(S) \subseteq T$.

As for convex hulls, there is also a representation theorem for conic hulls. It is slightly stronger than the result for convex hulls, as we only need n vectors, whereas the Carathéodory theorem required $n+1$ vectors.

Theorem (Conic Representation Theorem). *Let $S \subseteq \mathbb{R}^n$ and let $\mathbf{x} \in \text{cone}(S)$. Then, there exist $k \leq n$ linearly independent vectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k \in S$ and $\boldsymbol{\lambda} \in \mathbb{R}^k_+$ such that*

$$\mathbf{x} = \sum_{i=1}^{k} \lambda_i \mathbf{x}_i \,.$$

Convex Polytopes and Basic Feasible Solutions

This representation theorem has an important application to convex polytopes of the form

$$P = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{A}\mathbf{x} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}\}, \quad \text{where } \mathbf{A} \in \mathbb{R}^{m \times n}, \mathbf{b} \in \mathbb{R}^m.$$

As we will see in the forthcoming weeks, this is a standard formulation of the constraints of a linear programming problem. We will assume without loss of generality that the rows of A are linearly independent[4].

An important property of nonempty convex polytopes of the form P is that they contain at least one vector with at most m nonzero elements (m being the number of constraints). An important concept related to this is the idea of a **basic feasible solution**.

Definition (Basic Feasible Solution). *$\overline{\mathbf{x}}$ is a basic feasible solution (bfs) of P if the columns of $\mathbf{A}$ corresponding to the indices of the positive values of $\overline{\mathbf{x}}$ are linearly independent.*

Example. 

Consider the linear system:

$$\begin{aligned} x_1 + x_2 + x_3 &= 6 \\ x_2 + x_4 &= 3 \\ x_1, x_2, x_3, x_4 &\geq 0. \end{aligned}$$

Find all the basic feasible solutions.

This definition seems unintuitive at first, but as we will see, if a linear programme has an optimal solution then there is an optimal bfs. So to find the optimal solution, we only

[4] If not, we can remove a constraint without changing the optimization problem

need to find the basic feasible solutions. E.8 If $\mathbf{A}$ is $m \times n$, then how many basic feasible solutions may there be?

Theorem (Existence of Basic Feasible Solutions). *Let* $P = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{A}\mathbf{x} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}\}$, *where* $\mathbf{A} \in \mathbb{R}^{m \times n}$ *and* $\mathbf{b} \in \mathbb{R}^m$. *If* $P \neq \emptyset$, *then it contains at least one* **bfs***.*

Extreme points

We can characterize basic feasible solutions as extreme points, which are points in the set that cannot be represented as a nontrivial convex combination of two different points in S.

Definition (Extreme Point). *Let* $S \subseteq \mathbb{R}^n$ *be a convex set. A point* $\mathbf{x} \in S$ *is called an extreme point of* S *if there do not exist* $\mathbf{x}_1, \mathbf{x}_2 \in S$ $(\mathbf{x}_1 \neq \mathbf{x}_2)$ *and* $\lambda \in (0, 1)$, *such that* $\mathbf{x} = \lambda \mathbf{x}_1 + (1 - \lambda)\mathbf{x}_2$. *The set of extreme point is denoted by* $\mathrm{ext}(S)$.

For example, the set of extreme points of a convex polytope consists of all its vertices - see Figure 5. There is an equivalence between basic feasible solutions and extreme points.

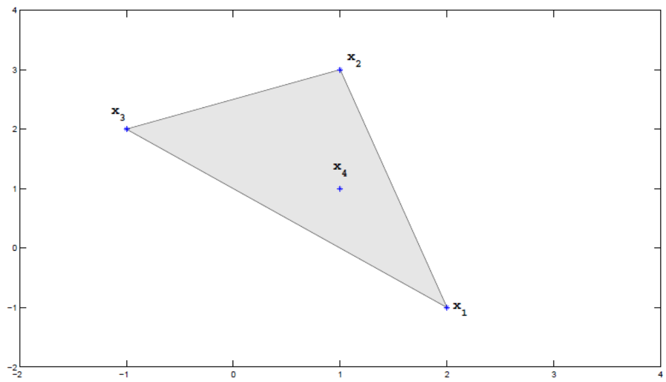

Figure 5: The extreme points of this triangle are given by its vertices $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3$. The point $\mathbf{x}_4$ is not an extreme point.

Theorem (Equivalence between Basic Feasible Solutions and Extreme Points). *Let* $P = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{A}\mathbf{x} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}\}$, *where* $\mathbf{A} \in \mathbb{R}^{m \times n}$ *has linearly independent rows and* $\mathbf{b} \in \mathbb{R}^m$. *The* $\overline{\mathbf{x}}$ *is a basic feasible solution of* P *if and only if it is an extreme point of* P.

E.9 Find the extreme points of the system

$$
\begin{aligned}
x_1 + x_2 + x_3 &= 6 \\
x_2 + x_4 &= 3 \\
x_1, x_2, x_3, x_4 &\geq 0.
\end{aligned}
$$

Finally, we state an important result which says that a compact convex set is the convex hull of its extreme points.

Theorem (The Krein-Milman Theorem)**.** *Let $S \subseteq \mathbb{R}^n$ be a compact convex set. Then*

$$S = \operatorname{conv}(\operatorname{ext}(S)) .$$

E.10 Next week, we will see that linear programmes, which are optimization problems of the form

$$
(\mathbf{LP}) : \quad \begin{array}{ll} \min_{\mathbf{x}} & \mathbf{c}^\top \mathbf{x} \\ \text{s.t.} & \mathbf{A}\mathbf{x} = \mathbf{b} \\ & \mathbf{x} \geq 0 \end{array}
$$

have at least one optimal solution which is an extreme point, or equivalently, a basic feasible solution. Note that problems of the form

$$
(\mathbf{LP}) : \quad \begin{array}{ll} \min_{\mathbf{x}} & \mathbf{c}^\top \mathbf{x} \\ \text{s.t.} & \mathbf{A}\mathbf{x} \leq \mathbf{b} \\ & \mathbf{x} \geq 0 \end{array}
$$

are essentially the same form, as we can add a slack variable $\mathbf{z} \geq 0$ to convert the inequality to an equality.

Solve the following optimization problems:

1.

$$
\begin{array}{ll} \max_{\mathbf{x}} & 2x_1 + 3x_2 - x_3 \\ & x_1 + x_2 + x_3 = 6 \\ \text{such that} & \quad x_2 + x_4 = 3 \\ & x_1, x_2, x_3, x_4 \geq 0 \end{array}
$$

What is the minimum of this problem? Is there a unique $\mathbf{x}$ that solves the problem?

2.

$$
\begin{array}{ll} \min_{\mathbf{x}} & x_1 - 3x_2 + x_3 - x_4 \\ & x_1 + x_2 + x_3 = 6 \\ \text{such that} & \quad x_2 + x_4 = 3 \\ & x_1, x_2, x_3, x_4 \geq 0 \end{array}
$$

3.

$$\begin{aligned} \min_{\mathbf{x}} \quad & x_1 + x_2 \\ \text{such that} \quad & 2x_1 + x_2 \leq 2 \\ & x_2 + x_3 \leq 3 \\ & x_1, x_2, x_3 \geq 0 \end{aligned}$$

Hint: you can introduce new variables $x_4, x_5 \geq 0$ to convert the inequality constraints into equality constraints.

Checklist

The idea of this checklist is to help you to self-evaluate your progress and understanding of the subject, and to give you some guidance on where to focus. If you can tick all the boxes it means you're doing alright, otherwise you need to study a bit more, grab a book, watch the videos, or seek help from classmates, the lecturers, or the demonstrators. Try to fill as many gaps as quickly as possible.

And remember to do the 

's!

Learning Outcome	Check
I can state the definition of a convex set and verify the examples in the notes.	
I understand the concept of a convex combination and its relation to the unit simplex presented in Week 2.	
I understand the notion of convex hull of a set S as the *smallest* convex set containing S.	
I have studied the proof of the representation theorem for convex hulls and I can use it.	
I understand what a convex polytope is.	
I understand the definition of basic feasible solution and I can verify it.	
I understand the relation between bfs and extreme points, and its relation to feasibility in optimization.	