an appropriate variable, so as to suggest that it is the line corresponding to that variable being zero. For example, the equality $x_1 + x_2 = 13$ implies that the slack variable $x_3 = 0$. So the line $x_1 + x_2 = 13$ is the same as the line $x_3 = 0$, illustrating why that line is labelled x_3.

The point $\mathbf{x}^{(1)} = (13, 0)^\mathsf{T}$ is the intersection of the lines $x_2 = 0$ and $x_3 = 0$. Thus if we take x_2 and x_3 as parameters, this leaves x_1, x_4, x_5, x_6, and x_7 for the basis. That is, the point $(13, 0)^\mathsf{T}$ corresponds to a basic solution (which incidentally arises from our first pivot operation).

Likewise, pivot once again and consider the resulting basic solution. It has parameters x_3 and x_4. Sure enough, the lines $x_3 = 0$ and $x_4 = 0$ intersect at $\mathbf{x}^{(2)} = (9, 4)^\mathsf{T}$, just as the resulting tableau predicts. This particular tableau is feasible, so its corresponding basic solution is a extreme point of the feasible region. One more pivot takes us to the basis $\{1, 2, 3, 5, 7\}$. That is, parameters x_4 and x_6 are zero; these lines intersect at $\mathbf{x}^{(3)} = (16, 11)^\mathsf{T}$.

The point $X = (80, 204)^\mathsf{T}/11$ corresponds to the basis $\{1, 2, 3, 4, 7\}$. Since this point is on the infeasible side of the line $x_7 = 0$, it must be that $x_7 < 0$. Thus the pivot required to travel from $(16, 11)^\mathsf{T}$ to $(80/11, 204/11)^\mathsf{T}$ must be wrong according to our Simplex rules. Indeed, you can check that the pivot $4 \mapsto 5$ violates the b-ratio rule. The correct pivot of $4 \mapsto 7$ yields the extreme point $\mathbf{x}^{(4)} = (12, 14)^\mathsf{T}$, which is optimal.

Workout 3.1.3 *It seems clear that in two dimensions (i.e. two problem variables), there is a one-to-one correspondence between basic solutions and intersection points of two constraint lines. However, this is not quite true — why? What is true, and what is the general argument in $\mathbb{R}^n$? [HINT: Recall Workouts 2.1.5 and 2.8.1.]*

Surely, this is precisely what would please Descartes so. Knowing from geometry that the optimum must occur at a extreme point, it is quite satisfying to discover that the Simplex algorithm spends its time jumping from one extreme point to another.

3.2 Convexity

Let's discuss the shapes of our feasible regions in greater detail. You might have noticed that the 2-dimensional problems considered up to this point all have had feasible regions sharing an interesting property. If you randomly pick any feasible point and your friend arbitrarily picks any other feasible point, then the straight line segment which joins these two points is contained entirely in the feasible region. (Check them and see that this is so!)

A region S is called **convex** if the line segment joining any two points of S is contained entirely in S. So the statement above can be rephrased to say that the feasible regions so far all have been convex. The question then arises, is every feasible region convex? We claim that the answer is yes. convex region

Workout 3.2.1 *Consider the half-space R defined by $3x_1 - 7x_2 \le -12$.*

- **(a)** *Prove that R is convex.*
- **(b)** *More generally, prove that the half-space defined by $\Sigma_{j=1}^{n} a_j x_j \le b$ is convex.*

Workout 3.2.2 *Let S and T be two arbitrarily chosen convex regions.*

- **(a)** *Prove by example that $S \cup T$ is not always convex.*
- **(b)** *Prove that $S \cap T$ is always convex.*
- **(c)** *Use induction to prove that the intersection of an arbitrary number (finite or infinite) of convex regions is convex.*

(d) *Prove that all polyhedra are convex.*

Theorem 3.2.3 *Let* $\mathbf{A} \in \mathbb{R}^{m \times n}$, $\mathbf{b} \in \mathbb{R}^m$, *and let* $S = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{Ax} \le \mathbf{b}\}$. *Then S is convex.*

Note that Theorem 3.2.3 implies that the feasible region of any LOP in standard form is convex. Indeed, the standard form $\{\mathbf{Ax} \le \mathbf{b}, \mathbf{x} \ge \mathbf{0}\}$ is a special instance of the form $\{\mathbf{Ax} \le \mathbf{b}\}$. In fact, the forms are equivalent since nonnegativity constraints can be considered as part of the problem constraints.

Workout 3.2.4

(a) *Use Workout 3.2.1 and 3.2.2 to prove Theorem 3.2.3.*

(b) *Does the same hold for LOPs not in standard form?*

Let $L = L(\mathbf{x}^1, \mathbf{x}^2)$ be the line segment that joins the two points $\mathbf{x}^1$ and $\mathbf{x}^2$ in $\mathbb{R}^n$ (here, the superscripts do not denote exponents, but rather merely first and second points). Can we describe L algebraically? If $\mathbf{x}^1 = (3,5)^\mathsf{T}$ and $\mathbf{x}^2 = (8,-1)^\mathsf{T}$ then the line that passes through both points is defined by the set of all points $(x_1, x_2)^\mathsf{T}$ which satisfy $6x_1 + 5x_2 = 43$. A parametric description can be given by those points of the form $(3,5)^\mathsf{T} + t(5,-6)^\mathsf{T}$, which can be written instead as $(1-t)(3,5)^\mathsf{T} + t(8,-1)^\mathsf{T}$. Then the substitutions $t_1 = 1-t$ and $t_2 = t$ yield the form $t_1\mathbf{x}^1 + t_2\mathbf{x}^2$, with $t_1 + t_2 = 1$. Notice that the point $\mathbf{x}^1$ corresponds to the value $t = 0$ ($(t_1, t_2) = (1, 0)$), and the point $\mathbf{x}^2$ corresponds to the value $t = 1$ ($(t_1, t_2) = (0, 1)$). So it seems that the points on the line segment between $\mathbf{x}^1$ and $\mathbf{x}^2$ must be defined as those which arise when both $t_1, t_2 \ge 0$ (for example, the midpoint $(5.5, 2)^\mathsf{T}$ has $(t_1, t_2) = (.5, .5)$). Therefore, we equivalently and algebraically can re-define a region (set) S to be convex if $t_1\mathbf{x}^1 + t_2\mathbf{x}^2 \in S$ whenever $\mathbf{x}^1$, $\mathbf{x}^2 \in S$, $t_1 + t_2 = 1$, and $t_1, t_2 \ge 0$. Of course, since $\mathbf{x}^1$ and $\mathbf{x}^2$ are linearly independent, we know from linear algebra that the region of all points of the form $t_1\mathbf{x}^1 + t_2\mathbf{x}^2$ is all of $\mathbb{R}^2$.

Workout 3.2.5 *With* $\mathbf{x}^1$ *and* $\mathbf{x}^2$ *as above, consider the four points* $A = (2,6)^\mathsf{T}$, $B = (-2,11)^\mathsf{T}$, $C = (10,-1)^\mathsf{T}$ *and* $D = (\frac{25}{6}, \frac{18}{5})^\mathsf{T}$.

(a) *For each of the points* $(s_1, s_2)^\mathsf{T}$ *above, find the values of* t_1 *and* t_2 *that yield* $t_1\mathbf{x}^1 + t_2\mathbf{x}^2 = (s_1, s_2)^\mathsf{T}$.

(b) *Which of the above points lie on the line containing* $\mathbf{x}^1$ *and* $\mathbf{x}^2$*?*

(c) *Which of the above points lie between the rays from the origin through* $\mathbf{x}^1$ *and* $\mathbf{x}^2$*?*

(d) *Which of the above points lie on the line segment between* $\mathbf{x}^1$ *and* $\mathbf{x}^2$*?*

linear/affine/conic/convex combination

For a set $X = \{\mathbf{x}^1, \ldots, \mathbf{x}^k\}$ of k points in $\mathbb{R}^n$, we call the sum $\Sigma_{j=1}^k t_j\mathbf{x}^j$ a **linear combination** of X. If it is the case that $\Sigma_{j=1}^k t_j = 1$, then we call the sum an **affine combination** of X. If instead it is the case that each $t_j \ge 0$, then we call the sum a **conic combination** of X. Finally, if the linear combination is both affine and conic, then we call it a **convex combination** of X. (Note that for a LOP in standard form, the dual variables are used to form conic combinations of its constraints.)

Workout 3.2.6 *Suppose that α and β are each convex combinations of X, and that γ is a convex combination of α and β. Prove that γ is a convex combination of X.*

Before proceeding, let's get more comfortable with these definitions. If we can think of pictures in $\mathbb{R}^3$ for a moment, imagine three points $\mathbf{x}^1$, $\mathbf{x}^2$, and $\mathbf{x}^3$ which are linearly independent. The points $(1,0,0)^\mathsf{T}$, $(0,1,0)^\mathsf{T}$, and $(0,0,1)^\mathsf{T}$ will do fine, but you can think of another three if you wish. (In fact, these points are completely general, as a change of basis matrix converts any other example to this one.) The region of all linear combinations of these three points consists of all of $\mathbb{R}^3$. The region of all affine combinations consists of the plane that contains all three points. (A 2-dimensional linear space is a plane going through the origin. A 2-dimensional affine space is a plane shifted way from the origin, like this one.) The region of all conic combinations consists of all the nonnegative points of $\mathbb{R}^3$. The region of all convex combinations consists of the points on the affine plane which lie on the triangle whose extreme points are $\mathbf{x}^1$, $\mathbf{x}^2$, and $\mathbf{x}^3$: the intersection of the affine and conic spaces. Let's test that what we have said here is true.

Workout 3.2.7 *Choose three points* $\mathbf{x}^1, \mathbf{x}^2, \mathbf{x}^3 \in \mathbb{R}^3$. *Let* P *be the plane containing* $\mathbf{x}^1, \mathbf{x}^2$ *and* $\mathbf{x}^3$*, and let* T *be the triangle having corners* $\mathbf{x}^1, \mathbf{x}^2$ *and* $\mathbf{x}^3$*. (Since every three noncolinear points determine a plane, for the purposes of this workout it may be simplest to first pick an equation* $\sum_{j=1}^{3} a_j x_j = c$ *and then pick* $\mathbf{x}^1, \mathbf{x}^2$ *and* $\mathbf{x}^3$ *satisfying it — what some call 'reverse engineering'.)*

- **(a)** *Pick a point* γ *that you know is on* P *but not* T.
  - **[i]** *Verify that* $t_1 + t_2 + t_3 = 1$.
  - **[ii]** *Verify that one of the* $t_j s$ *is negative.*
- **(b)** *Pick a point* γ *that you know is on both* P *and* T.
  - **[i]** *Verify that* $t_1 + t_2 + t_3 = 1$.
  - **[ii]** *Verify that each* $t_j \geq 0$.
- **(c)** *Pick a point* τ *that you know is not on* P *and verify that* $t_1 + t_2 + t_3 \neq 1$.

Workout 3.2.8 *Draw the affine, conic and convex spaces determined by* $\mathbf{x}^1 = (1,5,6)^\mathsf{T}$, $\mathbf{x}^2 = (7,1,5)^\mathsf{T}$ *and* $\mathbf{x}^3 = (6,7,1)^\mathsf{T}$.

Why is it that all the points on the interior of the triangle T formed by $\mathbf{x}^1$, $\mathbf{x}^2$, and $\mathbf{x}^3$ are convex combinations of the $\mathbf{x}^i s$? Suppose $\gamma \in T$ and consider the line through γ and $\mathbf{x}^3$. That line intersects the line segment which joins $\mathbf{x}^1$ and $\mathbf{x}^2$ at some point which we can call δ. We know that δ is a convex combination of $\mathbf{x}^1$ and $\mathbf{x}^2$, and that γ is a convex combination of δ and $\mathbf{x}^3$. So indeed, by Workout 3.2.6, γ is a convex combination of $\mathbf{x}^1$, $\mathbf{x}^2$, and $\mathbf{x}^3$.

Define the **convex hull** of a set $X \subseteq \mathbb{R}^n$, denoted $\text{vhull}(X)$, to be the smallest convex region containing X. Denote the set of all convex combinations of finitely many points in X by $\text{vcomb}(X)$. convex hull

Workout 3.2.9 *Use the definition of* vhull *and Workout 3.2.6 to prove that, for all* $X \subseteq \mathbb{R}^n$*, every* $\gamma \in \text{vhull}(X)$ *is in* $\text{vcomb}(X)$*; that is,* $\text{vhull}(X) \subseteq \text{vcomb}(X)$. *(Note that* X *may be finite or infinite.)*

An obvious and interesting question to pursue is whether or not $\text{vhull}(X) = \text{vcomb}(X)$; in other words, whether or not $\text{vcomb}(X) \subseteq \text{vhull}(X)$ also holds. To this end, define $\text{vcomb}_k(X)$ to be the set of all convex combinations involving exactly k of the points in X. More precisely, $\mathbf{x} \in \text{vcomb}_k(X)$ if and only if there exist $X' = \{\mathbf{x}^1, \ldots, \mathbf{x}^k\} \subseteq X$ such that $\mathbf{x} \in \text{vcomb}(X')$. In particular, $\text{vcomb}_1(X) = X$ and $\cup_{k=1}^{\infty} \text{vcomb}_k(X) = \text{vcomb}(X)$.

Workout 3.2.10 *Let* C *be any convex region containing the set* $X \subseteq \mathbb{R}^n$. *Use induction on* k *to prove that* $\text{vcomb}_k(X) \subseteq C$ *for all* $k \geq 1$. *[HINT: The technique is not too different from that used preceding Workout 3.2.9.]*

Theorem 3.2.11 *For every set* $X \subseteq \mathbb{R}^n$ *we have* $\text{vhull}(X) = \text{vcomb}(X)$.

Workout 3.2.12 *Use Workouts 3.2.9 and 3.2.10 to prove Theorem 3.2.11.*

3.3 A Little Practice

Workout 3.3.1 *Consider the set* $X = \{(5,14)^\mathsf{T}, (9,4)^\mathsf{T}, (16,11)^\mathsf{T}, (12,14)^\mathsf{T}, (7,8)^\mathsf{T}, (3,10)^\mathsf{T}\} \subset \mathbb{R}^2$.

- **(a)** *Draw* $\mathrm{vhull}(X)$.
- **(b)** *Which points of* X *can be thrown away without altering* $\mathrm{vhull}(X)$*?*
- **(c)** *For each such point* γ*, find the smallest* k *for which* $\gamma \in \mathrm{vcomb}_k(X)$.
- **(d)** *Can you partition* $\mathrm{vhull}(X)$ *into* t *triangles for*
  - **[i]** $t = 2$*?*
  - **[ii]** $t = 3$*?*
  - **[iii]** $t = 4$*?*
  - **[iv]** $t = 5$*?*
- **(e)** *Draw* $\mathrm{vcomb}_k(X)$ *for each of the values*
  - **[i]** $k = 1$.
  - **[ii]** $k = 2$.
  - **[iii]** $k = 3$.
  - **[iv]** $k = 4$.
- **(f)** *Which of the* $\mathrm{vcomb}_k(X)$ *above are convex?*
- **(g)** *Find the inequalities that define the half-spaces whose intersection equals* $\mathrm{vhull}(X)$.

Workout 3.3.2 *Let* X *be a finite set of points in* $\mathbb{R}^n$*. Prove that* $\mathrm{vhull}(X)$ *is a polytope.*

3.4 Carathéodory's Theorem

Note that $\mathrm{vcomb}_{k-1}(X) \subseteq \mathrm{vcomb}_k(X)$ for every $X \subseteq \mathbb{R}^n$, because including an extra point of X with coefficient zero does not alter the convex combination requirements. Thus we have that $\mathrm{vhull}(X) = \cup_{k=1}^{\infty}\mathrm{vcomb}_k(X) = \lim_{k\to\infty} \mathrm{vcomb}_k(X)$. It is natural then to ask if there is some finite k for which $\mathrm{vhull}(X) = \mathrm{vcomb}_k(X)$. For example, the value $k = |X|$ suffices in the case that X is finite. But what about the case for infinite X? In either case, what is the smallest such k? Does the dimension n matter? In 1907 Carathéodory answered these questions in the following theorem.

Carathéodory's Theorem

Theorem 3.4.1 *Let* $X \subseteq \mathbb{R}^n$*. Then* $\mathrm{vhull}(X) = \mathrm{vcomb}_{n+1}(X)$.

The result is, of course, best possible. For example, let X be any n linearly independent points with entirely nonnegative coordinates, plus the origin. Then the average of those $n+1$ points is in $\mathrm{vcomb}_{n+1}(X)$ but not $\mathrm{vcomb}_n(X)$. The proof is not too tricky, and uses the ideas from Workout 3.3.1c, as well as the Fundamental Theorem 2.9.1.

Proof. We already know that $\mathrm{vcomb}_{n+1}(X) \subseteq \mathrm{vhull}(X)$. Therefore we need only show that $\mathrm{vhull}(X) \subseteq \mathrm{vcomb}_{n+1}(X)$. Let $\boldsymbol{\gamma} \in \mathrm{vhull}(X)$. Then for some k and some $\mathbf{x}^1, \ldots, \mathbf{x}^k \in X$ we can write

$$\boldsymbol{\gamma} = \sum_{j=1}^{k} t_j \mathbf{x}^j \quad \text{with} \quad \sum_{j=1}^{k} t_j = 1 \tag{3.1}$$