4. CONSTRAINED MINIMIZATION

The minimization of a function f_0 over a set $C \subset I\!R^n$ can be much harder than the minimization of f_0 over all of $I\!R^n$, and it raises a host of issues. Some of these concern the ways a numerical method might be able to maneuver around in C or near to it, while others come up in characterizing optimality itself.

Feasible directions: A vector $w \neq 0$ is said to give a *feasible direction* into C at a point $x \in C$ if there's an $\varepsilon > 0$ such that the line segment $\{x + \tau w \mid 0 \leq \tau \leq \varepsilon\}$ lies in C.

Descent methods with constraints: The general class of descent methods described for unconstrained optimization could be adapted to the minimization of f_0 over C if the nature of C is such that feasible directions can readily be found. The rough idea is this. Starting from a point $x^0 \in C$, a sequence of points is generated by the scheme that, when at $x^\nu \in C$, a descent vector w^ν is chosen for f_0 which at the same time gives a feasible direction into C at x^ν. (If no such w^ν exists, the method terminates.) Next, some sort of line search is executed to produce a value $\tau^\nu > 0$ such that both $x^\nu + \tau^\nu w^\nu \in C$ and $f_0(x^\nu + \tau^\nu w^\nu) < f_0(x^\nu)$. The next point is taken then to be $x^{\nu+1} := x^\nu + \tau^\nu w^\nu$. In particular, one can imagine choosing τ^ν to be a value that minimizes $\varphi(\tau) = f_0(x^\nu + \tau w^\nu)$ over the set $\{\tau \geq 0 \mid x^\nu + \tau w^\nu \in C\}$; this would be the analogue of exact line search.

Pitfalls: Many troubles can plague this scheme, unless the situation is safeguarded by rather special features. Finding a descent vector w^ν that gives a feasible direction may be no easy matter, and even if one is found, there may be difficulties in using it effectively because of the need to keep within the confines of C. A phenomenon called *jamming* is possible, where progress is stymied by frequent collisions with the boundary of C and the method "gets stuck in a corner" of C or makes too many little zigzagging steps.

Lack of feasible directions at all: Of course, this kind of approach doesn't make much sense unless one is content to regard, as a "quasi-solution" to the problem, any point $x \in C$ at which there is no descent vector for f_0 giving a feasible direction into C. That may be acceptable for some sets C such as boxes, but not for sets C in which curvature dominates. For example, if C is a curved surface there may be *no* point $x \in C$ at which there's a feasible direction into C, because feasible directions refer only to movement along straight line segments embedded in C. Then there would be no point of C from which progress could be made by a modified descent method in minimizing f_0.

Variational geometry: Whether or not some form of descent method might be made to work, it's essential to have a solid grip on the geometry of the feasible set C. Classical geometry doesn't meet the needs, so new concepts have to be brought in. To get rolling, we need a notion of tangency which includes the tangent spaces long associated with "linearization" of curved structures, but at the same time covers vectors giving feasible directions.

Tangent vectors: For a closed set $C \subset I\!R^n$ and a point $\bar{x} \in C$, a vector w is said to be *tangent* to C at $\bar{x}$ if there is a sequence of vectors $x^\nu \to \bar{x}$ in C along with scalars $\tau^\nu \searrow 0$ such that the vectors $w^\nu = (x^\nu - \bar{x})/\tau^\nu$ converge to w.

Interpretation: The tangent vectors w to C at $\bar{x}$, apart from $w = 0$ (which corresponds in the definition to $x^\nu \equiv \bar{x}$, $w^\nu \equiv 0$), are the vectors pointing in a possibly asymptotic direction from which a sequence of points $x^\nu \in C$ can converge to $\bar{x}$. The direction of x^ν as seen from $\bar{x}$, which is the direction of w^ν, is not necessarily that of w, but gets closer and closer to it as $\nu \to \infty$.

Relation to feasible directions: Every vector $w \neq 0$ giving a feasible direction into C at $\bar{x}$ is a tangent vector to C at $\bar{x}$. Indeed, for such a vector w one can take $x^\nu = \bar{x} + \tau^\nu w$ for *any* sequence of values $\tau^\nu \searrow 0$ sufficiently small and have for $w^\nu = (x^\nu - \bar{x})/\tau^\nu$ that $w^\nu \equiv w$, hence trivially $w^\nu \to w$. Note from this that tangent vectors, in the sense defined here, can well point right into the interior of C, if that is nonempty; they don't have to lie along the boundary of C.

Relation to classical tangent spaces: When C is a "nice two-dimensional surface" in $I\!R^3$, the tangent vectors w to C at a point $\bar{x}$ form a two-dimensional linear subspace of $I\!R^3$, which gives the usual *tangent space* to C at $\bar{x}$. When C is a "nice one-dimensional curve," a one-dimensional linear subspace is obtained instead. Generalization can be made to tangent spaces to "nice curvilinear manifolds" of various dimensions d in $I\!R^n$, with $0 < d < n$. These comments are offered here on a heuristic, motivational level, but they will latter be supplied with rigor once such manifolds are identified with sets defined by equality constraints satisfying a constraint qualification that ensures a robust representation.

Tangent cone at a point: In general, the set of all vectors w that are tangent to C at a point $\bar{x} \in C$ is called the *tangent cone* to C at $\bar{x}$ and is denoted by $T_C(\bar{x})$.

Basic properties: Always, the set $T_C(\bar{x})$ contains the vector 0. Further, for every vector $w \in T_C(\bar{x})$ and every scalar $\lambda > 0$, the vector λw is again in $T_C(\bar{x})$ (because of the arbitrary size of the scaling factors in the definition). But in many situations one can have $w \in T_C(\bar{x})$, yet $-w \notin T_C(\bar{x})$. In particular, $T_C(\bar{x})$ can well be something other than a linear subspace of $I\!R^n$.

Cones: A subset of $I\!R^n$ is called a *cone* if it contains the zero vector and contains with each of its vectors all positive multiples of that vector. Geometrically, this means that a cone, unless it consists of 0 alone, is a "bundle of rays."

Limits of tangent vectors: The limit of any sequence of vectors $w^\nu \in T_C(\bar{x})$ is another vector $w \in T_C(\bar{x})$. In other words, $T_C(\bar{x})$ is always a closed set. (This can readily be gleaned from the definition of $T_C(\bar{x})$ and the general properties of convergent sequences.) In particular, if the vectors w^ν give feasible directions to C at $\bar{x}$ and $w^\nu \to w$, then $w \in T_C(\bar{x})$.

Tangents in some extreme cases: If $C = I\!R^n$, then $T_C(\bar{x}) = I\!R^n$. Indeed, any time $\bar{x}$ lies in the interior of C, one has $T_C(\bar{x}) = I\!R^n$. If C is a one-element set $\{a\}$ and $\bar{x} = a$, then $T_C(\bar{x}) = \{0\}$. More generally, the latter holds whenever $\bar{x}$ is an isolated point of C in the sense that there is no sequence of points $x^\nu \to \bar{x}$ in C apart from the constant sequence $x^\nu \equiv \bar{x}$.

Tangents to boxes: If $X = I_1 \times \cdots \times I_n$ in $I\!R^n$ with I_j a closed interval in $I\!R$, then at any point $\bar{x} = (\bar{x}_1, \ldots, \bar{x}_n) \in X$ (the component $\bar{x}_j$ lying in I_j) one has

$$T_X(\bar{x}) = T_{I_1}(\bar{x}_1) \times \cdots \times T_{I_n}(\bar{x}_n), \text{ where}$$
$$T_{I_j}(\bar{x}_j) = \begin{cases} [0,\infty) & \text{if } \bar{x}_j \text{ is the left endpoint (only) of } I_j, \\ (-\infty,0] & \text{if } \bar{x}_j \text{ is the right endpoint (only) of } I_j, \\ (-\infty,\infty) & \text{if } \bar{x}_j \text{ lies in the interior of } I_j, \\ [0,0] & \text{if } I_j \text{ is a one-point interval, consisting just of } \bar{x}_j. \end{cases}$$

In other words, the condition $w \in T_X(\bar{x})$ for $w = (w_1, \ldots, w_n)$ amounts to restricting w_j to lie in a one of the intervals $(-\infty,0]$, $[0,\infty)$, $(-\infty,\infty)$, or $[0,0]$ (the one-point interval consisting just of 0). The particular interval for an index j depends on the location of $\bar{x}_j$ relative to the endpoints of I_j.

Tangents to the nonnegative orthant: The set $X = I\!R^n_+$ is a box, the product of the intervals $I_j = [0,\infty)$. In this case one has $w \in T_X(\bar{x})$ if and only if $w_j \geq 0$ for all indices j with $\bar{x}_j = 0$. (For indices j with $\bar{x}_j > 0$, w_j can be any real number.)

Feasible directions in a box: When X is a box, not only does every vector $w \neq 0$ giving a feasible direction in X at $\bar{x}$ belong to $T_X(\bar{x})$, but conversely, every vector $w \neq 0$ in $T_X(\bar{x})$ gives a feasible direction in X at $\bar{x}$. This important property generalizes as follows to *any* set specified by finitely many linear constraints.

Tangents to polyhedral sets: For a polyhedral set C, like a box, simple formulas describe all the tangent cones. As a specific case ripe for many applications, suppose

$$x \in C \quad \Longleftrightarrow \quad \begin{cases} a_i{\cdot}x \leq b_i & \text{for } i \in [1,s], \\ a_i{\cdot}x = b_i & \text{for } i \in [s+1,m], \\ x \in X, & \text{with } X \text{ a box.} \end{cases}$$

Let $\bar{x} \in C$ and denote by $I(\bar{x})$ the set of $i \in [1,s]$ for which actually $a_i{\cdot}x = b_i$. Then

$$w \in T_C(\bar{x}) \quad \Longleftrightarrow \quad \begin{cases} a_i{\cdot}w \leq 0 & \text{for } i \in I(\bar{x}), \\ a_i{\cdot}w = 0 & \text{for } i \in [s+1,m], \\ w \in T_X(\bar{x}). \end{cases}$$

Argument: Let K denote the set of vectors w described by the final conditions (on the right). The question is whether $K = T_C(\bar{x})$. For a vector w and scalar $\tau > 0$, one has $\bar{x} + \tau w \in C$ if and only if $\bar{x} + \tau w \in X$ and $a_i{\cdot}w \leq [b_i - a_i{\cdot}\bar{x}]/\tau$ for $i \in [1,s]$, whereas $a_i{\cdot}w = 0$ for $i \in [s+1,m]$. Here for the indices $i \in [1,s]$ we have $[b_i - a_i{\cdot}\bar{x}] = 0$ if $i \in I(\bar{x})$, but $[b_i - a_i{\cdot}\bar{x}] > 0$ if $i \notin I(\bar{x})$; for the latter, $[b_i - a_i{\cdot}\bar{x}]/\tau^\nu \nearrow \infty$ whenever $\tau^\nu \searrow 0$. In light of the previously derived formula for the tangent cone $T_X(\bar{x})$ to the box X, it's clear then that the vectors of form those expressible as $w = \lim_\nu w^\nu$ with $\bar{x} + \tau^\nu w^\nu \in C$, $\tau^\nu \searrow 0$ (i.e., the vectors $w \in T_C(\bar{x})$) are none other than the vectors $w \in K$, as claimed.

Special tangent cone properties in the polyhedral case: When C is polyhedral, the tangent cone $T_C(\bar{x})$ at any point $\bar{x} \in C$ is polyhedral too. Moreover the vectors $w \neq 0$ in $T_C(\bar{x})$ all give feasible directions into C at $\bar{x}$, and they fully describe the geometry of C around $\bar{x}$ in the sense that there is a $\rho > 0$ for which

$$\big\{ x - \bar{x} \,\big|\, x \in C,\ |x - \bar{x}| \leq \rho \big\} \;=\; \big\{ w \in T_C(\bar{x}) \,\big|\, |w| \leq \rho \big\}.$$

Reason: In accordance with the definition of C being polyhedral, there's no loss of generality in taking C to have a representation of the sort just examined. The corresponding representation for $T_C(\bar{x})$ then immediately supports these conclusions. Since $a_i{\cdot}w \leq |a_i||w|$, the claimed equation holds for any $\rho > 0$ small enough that $\rho \leq [b_i - a_i\bar{x}]/|d_i|$ for every $i \notin I(\bar{x})$.

Characterizations of optimality: To what extent can the necessary and sufficient conditions for local optimality in unconstrained minimization in Theorem 3 be extended to minimization over a set C? This is a complex matter, because not only the "curvature" of f_0 as embodied in its Hessian matrices, but also that of the boundary of C can be crucial, yet we don't have any handle so far on analyzing the latter. Nonetheless, a substantial extension can already be stated for the case where C is essentially without curvature because of being polyhedral.

THEOREM 6 (local optimality conditions on a polyhedral set). *Consider the problem of minimizing f_0 over a polyhedral set C, with f_0 of class $\mathcal{C}^2$. Let $\bar{x} \in C$.*

(a) (necessary). *If $\bar{x}$ is a locally optimal solution, then*

$$\nabla f_0(\bar{x}){\cdot}w \geq 0 \text{ for all } w \in T_C(\bar{x}),$$
$$w{\cdot}\nabla^2 f_0(\bar{x})w \geq 0 \text{ for all } w \in T_C(\bar{x}) \text{ with } \nabla f_0(\bar{x}){\cdot}w = 0.$$

(b) (sufficient). *If $\bar{x}$ has the property that*

$$\nabla f_0(\bar{x}){\cdot}w \geq 0 \text{ for all } w \in T_C(\bar{x}),$$
$$w{\cdot}\nabla^2 f_0(\bar{x})w > 0 \text{ for all } w \in T_C(\bar{x}) \text{ with } \nabla f_0(\bar{x}){\cdot}w = 0,\ w \neq 0,$$

then $\bar{x}$ is a locally optimal solution. In fact there is a $\delta > 0$ such that

$$f_0(x) > f_0(\bar{x}) \text{ for all points } x \in C \text{ with } 0 < |x - x_0| < \delta.$$

Proof. The argument is an adaptation of the one for Theorem 3. To set the stage, we invoke the polyhedral nature of C to get the existence of $\rho > 0$ such that the points $x \in C$ with $0 < |x - \bar{x}| \leq \rho$ are the points expressible as $\bar{x} + \tau w$ for some vector $w \in T_C(\bar{x})$ with $|w| = 1$ and scalar $\tau \in (0, \rho]$. Then too, for any $\delta \in (0, \rho)$, the points $x \in C$ with $0 < |x - \bar{x}| \leq \delta$ are the points expressible as $\bar{x} + \tau w$ for some $w \in T_C(\bar{x})$ with $|w| = 1$ and some $\tau \in (0, \delta]$. Next we use the twice differentiability of f_0 to get second-order estimates around $\bar{x}$ in this notation: for any $\varepsilon > 0$ there is a $\delta > 0$ such that

$$\left| f_0(\bar{x} + \tau w) - f_0(\bar{x}) - \tau \nabla f_0(\bar{x}){\cdot}w - \frac{\tau^2}{2} w{\cdot}\nabla^2 f_0(\bar{x})w \right| \leq \varepsilon \tau^2 \qquad \text{for all } \tau \in [0, \delta] \text{ when } |w| = 1.$$

In (a), the local optimality of $\bar{x}$ gives in this setting the existence of $\bar{\delta} > 0$ such that $f_0(\bar{x} + \tau w) - f_0(\bar{x}) \geq 0$ for $\tau \in [0, \bar{\delta}]$ when $w \in T_C(\bar{x})$, $|w| = 1$. For such w and any $\varepsilon > 0$ we then have through second-order expansion the existence of $\delta > 0$ such that

$$\tau \nabla f_0(\bar{x}){\cdot}w + \frac{\tau^2}{2}\Big[w{\cdot}\nabla^2 f_0(\bar{x})w + 2\varepsilon \Big] \geq 0 \text{ for all } \tau \in [0, \delta].$$

This condition implies that $\nabla f_0(\bar{x}){\cdot}w \geq 0$, and if actually $\nabla f_0(\bar{x}){\cdot}w = 0$ then also that $w{\cdot}\nabla^2 f_0(\bar{x})w + 2\varepsilon \geq 0$. Since ε can be chosen arbitrarily, it must be true in the latter case that $w{\cdot}\nabla^2 f_0(\bar{x})w \geq 0$. Thus, the claim in (a) is valid for all $w \in T_C(\bar{x})$ with $|w| = 1$. It is also valid then for positive multiples of such vectors w, and hence for all $w \in T_C(\bar{x})$.

In (b), the desired conclusion corresponds to the existence of $\delta > 0$ such that $f_0(\bar{x} + \tau w) - f_0(\bar{x}) > 0$ for $\tau \in (0, \delta]$ when $w \in T_C(\bar{x})$, $|w| = 1$. Through the second-order expansion it suffices to demonstrate the existence of $\varepsilon > 0$ and $\delta' > 0$ such that

$$\tau \nabla f_0(\bar{x}){\cdot}w + \frac{\tau^2}{2}\Big[w{\cdot}\nabla^2 f_0(\bar{x})w - 2\varepsilon\Big] > 0 \qquad \text{for all } \tau \in (0, \delta'] \text{ when } w \in T_C(\bar{x}),\ |w| = 1.$$

Pursuing an argument by contradiction, let's suppose that such ε and δ' don't exist. Then, for any sequence $\varepsilon^\nu \searrow 0$ there must be sequences $\tau^\nu \searrow 0$ and $w^\nu \in T_C(\bar{x})$ with $|w^\nu| = 1$ and

$$\tau^\nu \nabla f_0(\bar{x}){\cdot}w^\nu + \frac{\tau^{\nu 2}}{2}\Big[w^\nu{\cdot}\nabla^2 f_0(\bar{x})w^\nu - 2\varepsilon^\nu\Big] \leq 0.$$

Because the sequence of vectors w^ν is bounded, it has a cluster point w; there is a subsequence $w^{\nu_\kappa} \to w$ as $\kappa \to \infty$. Then $|w| = 1$ (because the norm is a continuous function), and $w \in T_C(\bar{x})$ (because the tangent cone is a closed set). Rewriting our inequality as

$$w^{\nu_\kappa}{\cdot}\nabla^2 f_0(\bar{x})w^{\nu_\kappa} - 2\varepsilon^{\nu_\kappa} \ \leq\ -2\nabla f_0(\bar{x}){\cdot}w^{\nu_\kappa}/\tau^{\nu_\kappa},$$

where $\nabla f_0(\bar{x}){\cdot}w^{\nu_\kappa} \geq 0$ under the assumption of (b), we see when $\kappa \to \infty$ with

$$\nabla f_0(\bar{x}){\cdot}w^{\nu_\kappa} \to \nabla f_0(\bar{x}){\cdot}w, \qquad w^{\nu_\kappa}{\cdot}\nabla^2 f_0(\bar{x})w^{\nu_\kappa} + 2\varepsilon^{\nu_\kappa} \to w{\cdot}\nabla^2 f_0(\bar{x})w,$$

that $w{\cdot}\nabla^2 f_0(\bar{x})w \leq 0$, yet also $\nabla f_0(\bar{x}){\cdot}w = 0$ (for if $\nabla f_0(\bar{x}){\cdot}w > 0$ the right side of the inequality would go to $-\infty$). This mix of properties of w is impossible under the assumption of (b). The contradiction finishes the proof. □

Remark: In the case where $C = I\!R^n$, so that $T_C(\bar{x}) = I\!R^n$, the assertions in Theorem 6 turn precisely into the ones for unconstrained minimization in Theorem 3. Thus, the version of optimality conditions just obtained subsumes the earlier one.

Minimization subject to linear constraints: When the polyhedral set C has a constraint represent in the pattern considered just prior to Theorem 6, the condition $w \in T_C(\bar{x})$ has the associated constraint representation of like sort that was described there, and this can be used to augment the statement of the optimality conditions just derived.

Optimality over a box: In particular Theorem 6 applies to a conventional problem $(\mathcal{P})$ without constraint functions f_i, but just an abstract constraint $x \in X$ with X a box—specifying upper and/or lower bounds (or nonnegativity conditions) on the variables x_j. Then $C = X$.

Example: Consider a problem in which $f_0(x_1, x_2)$ is minimized over the quadrant $X = \{(x_1, x_2) \,|\, x_1 \geq 0,\, x_2 \geq 0\}$. What can be said about the possible attainment of the minimum at $\bar{x} = (0, 1)$? The tangent cone $T_X(\bar{x})$ at this point consists of all vectors $w = (w_1, w_2)$ with $w_1 \geq 0$. In both (a) and (b) of Theorem 6, the first-order condition on $\nabla f_0(\bar{x})$ comes down to the requirement that $(\partial f_0/\partial x_1)(\bar{x}) \geq 0$ but $(\partial f_0/\partial x_2)(\bar{x}) = 0$. In the case where $(\partial f_0/\partial x_1)(\bar{x}) > 0$, the second-order necessary condition in (a) is $(\partial^2 f_0/\partial x_2^2)(\bar{x}) \geq 0$, while the second-order sufficient condition in (b) is $(\partial^2 f_0/\partial x_2^2)(\bar{x}) > 0$. When $(\partial f_0/\partial x_1)(\bar{x}) > 0$, however, the condition in (a) is $w{\cdot}\nabla^2 f_0(\bar{x})w \geq 0$ for all $w = (w_1, w_2)$ with $w_1 \geq 0$, while the condition in (b) is $w{\cdot}\nabla^2 f_0(\bar{x})w > 0$ for all such $w \neq (0, 0)$.

Convexity in optimization: In constrained as well as in unconstrained minimization, convexity is a watershed concept. The distinction between problems of "convex" and "nonconvex" type is much more significant in optimization than that between problems of "linear" and "nonlinear" type.

Convex sets: A set $C \subset I\!R^n$ is *convex* if for every choice of $x_0 \in C$ and $x_1 \in C$ with $x_0 \neq x_1$ and every $\tau \in (0, 1)$ the point $(1 - \tau)x_0 + \tau x_1$ belongs to C.

Interpretation: This means that C contains with every pair of points the line segment joining them. Although "convex" in English ordinarily refers to a "bulging" appearance, the mathematical meaning is that there are no dents, gaps or holes.

Elementary rules:

Intersections: If C_i is a convex set in $I\!R^n$ for $i = 1, \ldots, r$, then $C_1 \cap \cdots \cap C_r$ is a convex set in $I\!R^n$. (This is true in fact not just for a finite intersection but the intersection of an arbitrary infinite family of convex sets.)

Products: If C_i is a convex set in $I\!R^{n_i}$ for $i = 1, \ldots, r$, then $C_1 \times \cdots \times C_r$ is a convex set in the space $I\!R^{n_1} \times \cdots \times I\!R^{n_r} = I\!R^{n_1 + \cdots + n_r}$.

Images: If C is a convex set in $I\!R^n$, A is a matrix in $I\!R^{m \times n}$ and $b \in I\!R^m$, then the set $D = \{Ax + b \,|\, x \in C\}$ is convex in $I\!R^m$.

Inverse images: If D is a convex set in $I\!R^m$, A is a matrix in $I\!R^{m \times n}$ and $b \in I\!R^m$, then the set $C = \{x \,|\, Ax + b \in D\}$ is convex in $I\!R^n$.

Basic examples of convex sets:

Extremes: The whole space $C = I\!\!R^n$ is convex. On the other hand, the empty set $C = \emptyset$ is convex. (It satisfies the definition "vacuously.") Likewise, sets $C = \{a\}$ consisting of just a single point (*singleton* sets) are convex sets.

Linear subspaces: Any subspace C of $I\!\!R^n$ (as in linear algebra) is closed and convex.

Intervals: The convex subsets of the real line $I\!\!R$ are the various *intervals*, whether bounded, unbounded, open, closed, or a mixture.

Boxes and orthants: As a product of closed intervals, any box is a closed, convex set. For instance the nonnegative orthant $I\!\!R^n_+$, being a box, is closed and convex. So too is the nonpositive orthant $I\!\!R^n_-$, which is defined analogously.

Hyperplanes and half-spaces: All such sets are closed and convex—as an elementary consequence of their definition.

Polyhedral sets: As the intersection of a family of hyperplanes or closed half-spaces, any polyhedral set is closed and convex.

Euclidean balls: For any point $\bar{x}$ and radius value $\rho \in (0, \infty)$, the *closed ball* of radius ρ around $\bar{x}$ consists of the points x with $|x - \bar{x}| \leq \rho$; the corresponding *open ball* or radius ρ around $\bar{x}$ is defined in the same way, but with strict inequality. Both kinds of balls are examples of convex sets.

Argument. For the case of $C = \{x \,|\, |x - \bar{x}| \leq \rho\}$, consider x_0 and x_1 in C and $\tau \in (0, 1)$. For the point $x = (1 - \tau)x_0 + \tau x_1$, we can use the fact that $\bar{x} = (1 - \tau)\bar{x} + \tau\bar{x}$ to write

$$\begin{aligned} |x - \bar{x}| &= \big|(1 - \tau)(x_0 - \bar{x}) + \tau(x_1 - \bar{x})\big| \\ &\leq (1 - \tau)|x_0 - \bar{x}| + \tau|x_1 - \bar{x}| \leq (1 - \tau)\rho + \tau\rho = \rho, \end{aligned}$$

from which we conclude that $x \in C$.

Tangents to convex sets: For a convex set $C \subset I\!\!R^n$, the tangent cone $T_C(\bar{x})$ at any point $\bar{x} \in C$ consists of the zero vector and all vectors $w \neq 0$ expressible as limits of sequences of vectors $w^\nu \neq 0$ giving feasible directions into C at $\bar{x}$.

Argument. This follows from the observation that when C is convex, the presence of $\bar{x} + \tau w$ in C entails that the entire line segment from $\bar{x}$ to $\bar{x} + \tau w$ lies in C. Likewise, if $\bar{x} + \tau^\nu w^\nu \in C$ then $\bar{x} + \tau w^\nu \in C$ for all $\tau \in [0, \tau^\nu]$.

THEOREM 7 (optimality over a convex set). *Consider a problem of minimizing a function f_0 of class $\mathcal{C}^1$ over a convex set $C \subset \mathbb{R}^n$. Let $\bar{x}$ be a point of C.*

(a) (necessary). *If $\bar{x}$ is locally optimal, then $\nabla f_0(\bar{x}) \cdot w \geq 0$ for all $w \in T_C(\bar{x})$, and this condition is equivalent in fact to having*

$$\nabla f_0(\bar{x}) \cdot [x - \bar{x}] \geq 0 \text{ for all } x \in C.$$

(b) (sufficient). *If this holds and f_0 is convex on C, then $\bar{x}$ is globally optimal.*

Proof. In (a), consider first any point $x \in C$ different from $\bar{x}$. The line segment joining $\bar{x}$ with x lies in C by convexity; the vector $w = x - \bar{x}$ gives a feasible direction into C at $\bar{x}$. The function $\varphi(\tau) = f(\bar{x} + \tau w)$ then has a local minimum at $\tau = 0$ relative to $0 \leq \tau \leq 1$. Hence $0 \leq \varphi'(0) = \nabla f_0(\bar{x}) \cdot w = \nabla f_0(\bar{x}) \cdot [x - \bar{x}]$. From this we see further that $\nabla f_0(\bar{x}) \cdot w \geq 0$ for *all* vectors w giving feasible directions into C at $\bar{x}$, inasmuch as these are positive multiples of vectors of the form $x - \bar{x}$ with $x \in C$. As noted above, any vector $w \neq 0$ in $T_C(\bar{x})$ is a limit of vectors w^ν giving feasible directions (which are in $T_C(\bar{x})$ as well). From having $\nabla f_0(\bar{x}) \cdot w^\nu \geq 0$ for all ν we get in the limit as $w^\nu \to w$ that $\nabla f_0(\bar{x}) \cdot w \geq 0$. Therefore, $\nabla f_0(\bar{x}) \cdot w \geq 0$ for all $w \in T_C(\bar{x})$, and this is equivalent to having $\nabla f_0(\bar{x}) \cdot [x - \bar{x}] \geq 0$ for all $x \in C$.

In (b), we have $f_0(x) - f_0(\bar{x}) \geq \nabla f_0(\bar{x}) \cdot x - \bar{x}$ for all $x \in \mathbb{R}^n$ by the convexity of f_0 (Theorem 4). If also $\nabla f_0(\bar{x}) \cdot (x - \bar{x}) \geq 0$ for all $x \in C$, we get $f_0(x) - f_0(\bar{x}) \geq 0$ for all $x \in C$, which means $\bar{x}$ is globally optimal in the minimization of f_0 over C. □

Variational inequalities: The condition in Theorem 7 that $\nabla f_0(\bar{x}) \cdot [x - \bar{x}] \geq 0$ for all $x \in C$ is known as the *variational inequality* for the mapping ∇f_0 and the convex set C, the point $\bar{x}$ being a *solution* to it. Variational inequalities can be studied also with ∇f_0 replaced by some other vector-valued mapping from C into $\mathbb{R}^n$. They have an interesting and significant place in optimization theory beyond merely the characterization of points at which a minimum is attained.

Interpretation via linearization: In terms of the function $l(x) = f_0(\bar{x}) + \nabla f_0(\bar{x}) \cdot [x - \bar{x}]$ giving the first-order expansion of f_0 at $\bar{x}$, the optimality condition in Theorem 7 is equivalent to saying that $l(x) \geq l(\bar{x})$ for all $x \in C$, or in other words, that l attains its global minimum over C at $\bar{x}$.

Convex functions on convex sets: The convexity of a function f on $I\!R^n$ has already been defined in terms of the inequality $f\big((1-\tau)x_0+\tau x_1\big) \leq (1-\tau)f(x_0)+\tau f(x_1)$ holding for all x_0 and x_1 in $I\!R^n$ and $\tau \in (0,1)$. The concept can be generalized now to the convexity of f *on a convex set* $C \subset I\!R^n$: the same inequality is used, but x_0 and x_0 are restricted to C. Similarly one speaks of f being strictly convex, or concave, or strictly concave on C. (Note that for these concepts to make sense f only has to be defined on C itself; values of f outside of C have no effect, because the convexity of C ensures that the relevant points $(1-\tau)x_0+\tau x_1$ all belong to C.)

Derivative tests: The tests in Theorem 4 apply equally well to the convexity or strict convexity of a differentiable function f on any *open* convex set $O \subset I\!R^n$.

Convexity-preserving operations: Derivative tests are by no means the only route to verifying convexity or strict convexity. Often it's easier to show that a given function is convex because it is constructed by convexity-preserving operations from other functions, already known to be convex (or, as a special case, affine). The following operations are convexity-preserving on the basis of elementary arguments using the definition of a convexity. Here C denotes a general convex set.

Sums: If f_1 and f_2 are convex functions on C, then so is f_1+f_2. (This can be extended to a sum of any number of convex functions.) Moreover, if one of the functions in the sum is *strictly* convex, then the resulting function is strictly convex.

Multiples: If f is convex on C and $\lambda \geq 0$, then λf is convex on C. (In combination with the preceding, this implies that any linear combination $\lambda_1 f_1+\cdots+\lambda_r f_r$ of convex functions with coefficients $\lambda_i \geq 0$ is convex.) Again, if one of the functions in the sum is *strictly* convex, and the associated coefficient is positive, then the function expressed by the sum is strictly convex.

Compositions I: If f is convex on C, then any function of the form $g(x)=\theta\big(f(x)\big)$ is convex on C, provided that the function θ on $I\!R^1$ is convex *and nondecreasing*.

Example: If f is convex in $I\!R^n$ *and* $f \geq 0$ *everywhere*, then the function $g(x) := f(x)^2$ is convex on $I\!R^n$, because $g(x)=\theta\big(f(x)\big)$ for θ defined by $\theta(t)=t^2$ when $t \geq 0$, but $\theta(t)=0$ when $t \leq 0$. This follows because θ is convex and nondecreasing (the convexity can be verified from the fact that $\theta'(t) = \max\{0, 2t\}$, which is nondecreasing). The tricky point is that unless $f \geq 0$ everywhere it would not be possible to write f as composed from this θ. Composition with $\theta(t) := t^2$ for all t wouldn't do, because this θ, although convex, isn't nondecreasing as a function on all of $I\!R^1$.

Compositions II: If f is convex on C, then $g(x) := f(Ax+b)$ is convex on $D := \{x \mid Ax - b \in C\}$ for any matrix $A \in I\!R^{n\times n}$ and vector $b \in I\!R^n$.

Pointwise max: If f_i is convex on C for $i = 1, \ldots, r$, then so too is the function f given by $f(x) = \max\{f_1(x), \ldots, f_r(x)\}$. If every f_i is strictly convex on C, then f is strictly convex on C.

Level sets of convex functions: If f is a convex function on $I\!R^n$, then for any $\alpha \in I\!R$ the sets $\{x \mid f(x) \leq \alpha\}$ and $\{x \mid f(x) < \alpha\}$ are convex. (Similarly, if f is a concave function the sets $\{x \mid f(x) \geq \alpha\}$ and $\{x \mid f(x) > \alpha\}$ are convex.)

Argument: Let C be such a level set for f and α, and suppose x_0 and x_1 are any two different points in C. Let $\lambda \in (0,1)$. From the convexity of f and the assumption that $f(x_0) \leq \alpha$ and $f(x_1) \leq \alpha$, we have

$$f\big((1-\tau)x_0 + \lambda x_1\big) \leq (1-\tau)f(x_0) + \tau f(x_1) \leq (1-\tau)\alpha + \tau\alpha = \alpha$$

and therefore $(1-\tau)x_0 + \tau x_1 \in C$. The case of strict inequality is similar.

Convex constraints: A *convex* constraint is a condition of the form $f_i(x) \leq c_i$ with f_i convex, or $f_i(x) \geq c_i$ with f_i concave, or $f_i(x) = c_i$ with f_i affine. Also, a condition of the form $x \in X$ is called a convex constraint if X is convex. Thus, a system of the form

$$x \in X \;\text{ and }\; f_i(x) \begin{cases} \leq 0 & \text{for } i = 1, \ldots, s, \\ = 0 & \text{for } i = s+1, \ldots, m, \end{cases}$$

is a system of convex constraints when X is convex, f_i is convex for $i = 1, \ldots, s$, and f_i is affine for $i = s+1, \ldots, m$. *Any set C defined by a system of convex constraints is a convex set*, because each separate constraint requires x to belong to a certain convex set, and C is the intersection of these sets.

Convex programming: An optimization problem in conventional format is called a *convex programming problem* if the constraints are convex, as just described, and also the objective function f_0 is convex.

Extension: This term is also used even if the objective and inequality constraint functions aren't convex all of $I\!R^n$, as long as they are convex on the convex set X. The feasible set C is still convex in that case.

Linear programming: This has been defined already, but we can now interpret it as the case of convex programming where the objective function and all the constraint functions are actually affine, and the set X is a box.

Quadratic programming: This too is a special case of convex programming; it is just like linear programming, except that the objective function can include a positive semidefinite quadratic term.

THEOREM 8 (special characteristics of convex optimization). *In a problem of minimizing a convex function f_0 over a convex set $C \subset I\!R^n$ (and thus any problem of convex programming) the following properties hold.*

(a) (local is global) *Any locally optimal solution is a globally optimal solution. Moreover, the set of all optimal solutions (if any) is convex.*

(b) (uniqueness criterion) *Strict convexity of the objective function f_0 implies there cannot be more than one optimal solution.*

Proof. (a) Suppose the point $\bar{x} \in C$ is locally optimal, i.e., there is an $\varepsilon > 0$ such that $f_0(x) \geq f_0(\bar{x})$ for all $x \in C$ satisfying $|x - \bar{x}| < \varepsilon$. Suppose also that $\tilde{x} \in C$, $\tilde{x} \neq \bar{x}$. Our aim is to show that $f_0(\tilde{x}) \geq f_0(\bar{x})$, thereby establishing the global optimality of $\bar{x}$ relative to C. For any $\tau \in (0,1)$ we know that $f_0\big((1-\tau)\bar{x} + \tau\tilde{x}\big) \leq (1-\tau)f_0(\bar{x}) + \tau f_0(\tilde{x})$. By choosing τ small enough, we can arrange that the point $x_\tau := (1-\tau)\bar{x} + \tau\tilde{x}$ (which still belongs to C by the convexity of C) satisfies $|x_\tau - \bar{x}| < \varepsilon$. (It suffices to take $\tau < \varepsilon/|\tilde{x} - \bar{x}|$.) Then the left side of the convexity inequality, which is $f_0(x_\tau)$, cannot be less than $f_0(\bar{x})$ by the local optimality of $\bar{x}$. We deduce that $f_0(\bar{x}) \leq f_0(x_\tau) \leq (1-\tau)f_0(\bar{x}) + \tau f_0(\tilde{x})$, which from the outer terms, after rearrangement, tells us that $f_0(\bar{x}) \leq f_0(\tilde{x})$, as needed.

Having determined that $\bar{x}$ is globally optimal, we can apply the same argument for arbitrary $\tau \in (0,1)$, without worrying about any ε. If $\tilde{x}$ is another optimal solution, of course, we have $f_0(\tilde{x}) = f_0(\bar{x})$, so that the right side of the double inequality $f_0(\bar{x}) \leq f_0(x_\tau) \leq (1-\tau)f_0(\bar{x}) + \tau f_0(\tilde{x})$ reduces to $f_0(\bar{x})$ and we can conclude that $f_0(x_\tau) = f_0(\bar{x})$ for all $\tau \in (0,1)$. In other words, the entire line segment joining the two optimal solutions $\bar{x}$ and $\tilde{x}$ must consist of optimal solutions; the optimal set is convex.

(b) Looking at the displayed inequality in the first part of the proof of (a) in the case where f_0 is strictly convex, and $\tilde{x}$ is again just any point of C different from the optimal solution $\bar{x}$, we get strict inequality. This leads to the conclusion that $f_0(\bar{x}) < f_0(\tilde{x})$. It's impossible, therefore, for $\tilde{x}$ to be optimal as well as $\bar{x}$. □

Convexity in estimating progress toward optimality: Another distinguishing feature of optimization problems of convex type is that in numerical methods for solving such problems it's usually possible to devise tests of how close one is getting to optimality—global optimality—as the method progresses. By contrast, for most other kinds of optimization problems one has hardly any handle on this important issue, and the question of a *stopping criterion* for an iterative procedure can only be answered in an ad hoc manner.

Upper and lower bounds on the optimal value: A simple example of the kind of estimate that can be built into a stopping criterion can be derived from the linearization inequality for convex functions in Theorem 4. Consider a problem of minimizing a differentiable convex function f_0 over a nonempty, closed set $C \subset I\!R^n$ that's also bounded, and imagine that a numerical method has generated in iteration ν a point $x^\nu \in C$. The affine function $l^\nu(x) = f_0(x^\nu) + \langle \nabla f_0(x^\nu), x - x^\nu \rangle$ has the property that $l^\nu(x) \leq f_0(x)$ for all x, and $l^\nu(x^\nu) = f_0(x^\nu)$. It follows that

$$\min_{x \in C} l^\nu(x) \;\leq\; \min_{x \in C} f_0(x) \;\leq\; f_0(x^\nu),$$

where the middle expression is the optimal value $\bar{\alpha}$ in the given problem, but the left expression, let's denote it by β^ν, is the optimal value in the possibly very easy problem of minimizing l^ν instead of f_0 over C. If C were a box, for instance, β^ν could instantly be calculated. While β^ν furnishes a current lower bound to $\bar{\alpha}$, the objective value $\alpha^\nu = f_0(x^\nu)$ furnishes a current upper bound. The difference $\alpha^\nu - \beta^\nu$ provides a measure of how far the point x^ν is from being optimal.

Duality: Optimization in a context of convexity is distinguished further by a pervasive phenomenon of "duality," in which a given problem of minimization ends up being paired with some problem of maximization in entirely different "dual" variables. Many important schemes of computation are based on this curious fact, or other aspects of convexity. In particular, almost all the known methods for breaking a large-scale problem down iteratively into small-scale problems, which perhaps could be solved in parallel, require convexity in their justification. This topic will be taken up later, after Lagrange multipliers for constraints have been introduced.

Minimization of nonsmooth convex functions: Not every convex function of interest in optimization is continuously differentiable.

Piecewise linear costs: Cost functions of a single variable often take the form of piecewise linear functions with increasing slope values. Such functions are convex. Specifically, suppose that a closed interval $C \subset I\!R^1$ is partitioned into a finite sequence of closed subintervals $C_1, C_2, \ldots, C_r$, and that the function $f : C \to I\!R$ is given on these subintervals by expressions

$$f(x) = a_i x + b_i \text{ when } x \in C_i,$$

where the formulas agree at the joins of the consecutive intervals (so that f is continuous), and $a_1 \leq a_2 \leq \cdots \leq a_r$. Then f is convex on C by the criterion for "pointwise max." In fact it can be seen in the given circumstances that

$$f(x) = \max\left\{a_1 x + b_1, \ldots, a_r x + b_r\right\} \text{ for all } x \in C.$$

Piecewise linear approximations: A smooth convex function f on $I\!R^n$ can be approximated from below by a nonsmooth convex function in a special way. We've already noted in connection with obtaining lower bounds on the optimal value in a problem of minimizing a convex function that the linearization (first-order Taylor expansion) of f at any point provides an affine lower approximation which is exact at the point in question. That degree of approximation is crude in itself, but imagine now what might be gained by linearizing f at more than one point. Specifically, consider a collection of finitely many points x_k, $k = 1, \ldots, r$, and at each such point the corresponding affine function obtained by linearization, namely

$$l_k(x) = f(x_k) + \left\langle \nabla f(x_k),\, x - x_k \right\rangle.$$

The function $g(x) = \max\{l_1(x), \ldots, l_r(x)\}$, which is convex on $I\!R^n$ (because the pointwise max of finitely many convex functions is always convex), and it satisfies

$$g(x) \leq f(x) \text{ for all } x, \text{ with } g(x_k) = f(x_k) \text{ for } k = 1, \ldots, r.$$

The convex function g is termed "piecewise linear" because its epigraph, as the intersection of the epigraphs of the l_k's, each of which is an upper closed half-space in $I\!R^{n+1}$, is a polyhedral subset of $I\!R^{n+1}$.

Cutting plane methods: An interesting class of numerical methods in convex programming relies on replacing the objective function and the inequality constraint functions, to the extent that they aren't merely affine, by such piecewise linear approximations. The finite collection of points in $\mathbb{R}^n$ on which the approximations are based is generated as the iterations proceed. These methods are called *cutting plane* methods because each new affine function entering one of the approximations cuts away part of the epigraph from the proceeding approximation.

Remark: Cutting plane methods tend to be quite slow in comparison with typical descent methods, but they are useful nonetheless in a number of situations where for some reason it's tedious or expensive to generate function values and derivatives, and approaches requiring line search are thereby precluded.

Minimizing a max of convex functions: In problems where a function of the form $f_0(x) = \max\{g_1(x), \dots, g_r(x)\}$ is to be minimized over a set C specified by convex constraints, the case where each function g_k is convex and smooth is especially amenable to treatment. Then f_0 is convex, and although it isn't smooth itself the usual device of passing to an epigraphical formulation retains convexity while bringing the smoothness of the g_k's to the surface. When an extra variable u is added, and the problem is viewed as one of minimizing the value of u over all choices of $(x_1, \dots, x_n, u) \in C \times \mathbb{R}$ such that $g_k(x_1, \dots, x_n) - u \leq 0$ for $k = 1, \dots, r$, it is seen that all constraints are convex.

Minimizing a max of affine functions over a polyhedral set: As a special case, if C is polyhedral and the functions g_k affine in the foregoing, the set $C \times \mathbb{R}$ will be polyhedral and the constraints $g_k(x_1, \dots, x_n) - u \leq 0$ are linear. In expressing C itself by a system of linear constraints, one sees that reformulated problem isn't just one of convex programming, but of *linear* programming.

Application to cutting plane methods: The subproblems generated in the cutting plane scheme of piecewise linear approximation, as described above, can, after epigraphical reformulation, be solved as linear programming problems if the set over which f_0 is to be minimized is specified by linear constraints. More generally such a reduction to a solving a sequence of linear programming problems is possible even if C is specified just by convex constraints over a box X, as long as the convex functions giving inequality constraints are smooth. The extension to this case involves generating piecewise linear approximations to those functions f_i along with the one to f_0 as computations proceed.

Norms: As another reminder that derivative tests aren't the only route to verifying convexity, consider any *norm* on $I\!R^n$, that is, a real-valued expression $\|x\|$ with the following properties, which generalize those of the Euclidean norm $|x|$.

(a) $\|x\| > 0$ for all $x \neq 0$,

(b) $\|\lambda x\| = |\lambda|\|x\|$ for all x and all λ,

(c) $\|x + y\| \leq \|x\| + \|y\|$ for all x and y.

The function $f(x) = \|x\|$ is convex, because for $0 < \tau < 1$ we have

$$f\Big((1-\tau)x_0 + \tau x_1\Big) = \big\|(1-\tau)x_0 + \tau x_1\big\| \\ \leq (1-\tau)\|x_0\| + \tau\|x_1\| = (1-\tau)f(x_0) + \tau f(x_1).$$

Commonly seen in problems of approximation are the l^p-*norms* for $p \in [1, \infty]$: for a point $x = (x_1, \ldots, x_n) \in I\!R^n$, one defines

$$\|x\|_1 := |x_1| + \cdots + |x_n|,$$
$$\|x\|_p := \Big(|x_1|^p + \cdots + |x_n|^p\Big)^{1/p} \text{ with } 1 < p < \infty \text{ (where } \|x\|_2 = |x|),$$
$$\|x\|_\infty := \max\Big\{|x_1|, \ldots, |x_n|\Big\}.$$

We won't try here to verify that these expressions do indeed give norms, i.e., that they satisfy the three axioms above, although this is elementary for $p = 1$ and $p = \infty$.

Example: Consider a parameter identification problem in which we wish to minimize an error expression of the form

$$E\big(r(a,b)\big) = \Big\|\big(r_1(a,b), \ldots, r_N(a,b)\big)\Big\|_p$$

with $r_k(a,b) := y_k - (ax_k + b)$ for $k = 1, \ldots, N$. (The unknowns here are the two parameter values a and b.) This expression is convex as a function of a and b, because it is obtained by composing an l^p-norm with an affine mapping.

Piecewise linear norms: The functions $f(x) = \|x\|_1$ and $f(x) = \|x\|_\infty$ are piecewise linear, in that each can be expressed as the pointwise max of a finite collection of affine (in fact linear) functions. Specifically, in terms of the vectors $e_j \in I\!R^n$ having coordinate 1 in jth position but 0 in all other positions, $\|x\|_1$ is the maximum of the n^2 linear functions $\big\langle \pm e_1 \pm e_2 \cdots \pm e_n, x\big\rangle$, whereas $\|x\|_\infty$ is the maximum of the $2n$ linear functions $\big\langle \pm e_j, x\big\rangle$. In contrast, the norm function $f(x) = \|x\|_2 = |x|$ can usually be treated in terms of its square, which is a simple quadratic convex function.