SIMPLE OPTIMALITY CONDITIONS FOR CONSTRAINED OPTIMIZATION

1. Optimality Conditions: Smooth Constrained

By using the tangent cones in Definition **??**, as well as Exercises **??** and **??**, simple optimality conditions for C^1-smooth convexly constrained problems are easily obtained.

Theorem 1.1. *(First-order necessary conditions) Suppose U is an open set in $\mathbf{E}$ and that $\bar{x}$ is a local minimizer of a function $f : U \to \mathbf{R}$ over the nonempty closed set $\Omega \subset U$. If f is differentiable at $\bar{x}$, then $\langle \nabla f(\bar{x}),\, v \rangle \geq 0$ for all $v \in T\,(\bar{x} \,|\, \Omega)$.*

Proof. Let $v \in T\,(\bar{x} \,|\, S)$. With no loss in generality, we may assume that $v \neq 0$. Then, by Exercise **??**, there is a $t > 0$ and a sequence $\{x^k\} \subset \Omega \setminus \{\bar{x}\}$ such that $x^k \to \bar{x}$, $\left\|x^k - \bar{x}\right\|^{-1}(x^k - \bar{x}) \to u$, and $v = tu$. Since $\bar{x}$ is a local minimizer of f on Ω, we may assume that $f(\bar{x}) \leq f(x^k)$ for all k. Then, for all k,

$$f(\bar{x}) \leq f(\bar{x}) + \left\langle \nabla f(\bar{x}),\, x^k - \bar{x} \right\rangle + o\left(\left\|x^k - \bar{x}\right\|\right),$$

and so

$$0 \leq \left\langle \nabla f(\bar{x}),\, x^k - \bar{x} \right\rangle + o\left(\left\|x^k - \bar{x}\right\|\right) \quad \forall\, k.$$

Dividing through by $\left\|x^k - \bar{x}\right\|$ and taking the limit in k yields the result. ☐

Notice that the first-order necessary condition above works for arbitrary nonempty closed sets Ω. However, to obtain a second-order conditions, we assume that Ω is convex, and make use of the tangent cone characterizations in Exercises **??** and **??**.

Theorem 1.2. *(Second-order conditions)*
Consider a C^2-smooth function $f : U \to \mathbf{R}$, where $U \subset \mathbf{E}$ is open. Fix a point $\bar{x} \in \Omega \subset U$, where Ω is a nonempty close convex set. Then the following are true.

(1) *(Necessary conditions) Assume that Ω is a convex polyhedron. If $\bar{x}$ is a local minimizer of f on Ω, then*

$$\langle \nabla f(\bar{x}),\, v \rangle \geq 0 \qquad \forall\ v \in T\,(\bar{x} \,|\, \Omega)$$

and

$$v^T \nabla^2 f(x) v \geq 0 \qquad \forall\ v \in T\,(\bar{x} \,|\, \Omega) \cap \operatorname{span}\left(\nabla f(\bar{x})\right)^{\perp}.$$

(2) (Sufficient conditions) If the relations

$$\langle \nabla f(\bar{x}),\, v\rangle \geq 0 \qquad \forall\ v \in T\,(\bar{x}\,|\,\Omega)$$

and

$$v^T\nabla^2 f(\bar{x})v > 0 \qquad \forall\ v \in \Big(T\,(\bar{x}\,|\,\Omega) \cap \mathrm{span}\,(\nabla f(\bar{x}))^{\perp}\Big) \setminus \{0\}.$$

hold, then there is an $\epsilon > 0$ *and* $\beta > 0$ *such that*

$$f(x) \geq f(\bar{x}) + \frac{\beta}{2}\,\|x-\bar{x}\|^2 \qquad \forall\ x \in B_\epsilon(\bar{x}) \cap \Omega. \tag{1.1}$$

Proof. Theorem 1.1 tells us that $\langle \nabla f(\bar{x}),\, v\rangle \geq 0$ for all $v \in T\,(\bar{x}\,|\,\Omega)$. Next let $v \in T\,(\bar{x}\,|\,\Omega) \cap \mathrm{span}\,(\nabla f(\bar{x}))^{\perp}$. With no loss in generality, we may assume that $\|v\| = 1$. By Exercise **??** there exists $\bar{t} > 0$ such that $\bar{x} + tv \in \Omega$ for all $t \in (0, \bar{t})$. Since $\bar{x}$ is a local solution, we may take $\bar{t}$ so small that that $f(\bar{x}) \leq f(\bar{x} + tv)$ for all $t \in (0, \bar{t})$. Then, for all $t \in (0, \bar{t})$,

$$f(\bar{x}) \leq f(\bar{x}) + t\langle \nabla f(\bar{x}),\, v\rangle + \frac{t^2}{2}\langle \nabla f(\bar{x})v,\, v\rangle + o\left(t^2\right)$$

and so

$$0 \leq \frac{1}{2}\langle \nabla f(\bar{x})v,\, v\rangle + \frac{o(t^2)}{t^2} \quad \forall\, t \in (0, \bar{t}).$$

Letting $t \to 0$ yields the second-order necessary condition.

To see the second-order sufficient condition, we suppose that the result is false so that there exists a sequences $\beta_k \downarrow 0$ and $x^k \to \bar{x}$ such that

$$f(x^k) < f(\bar{x}) + \frac{\beta_k}{2}\left\|x^k - \bar{x}\right\|^2 \qquad \forall\, k,$$

or equivalently,

$$f(\bar{x}) + \left\langle \nabla f(\bar{x}),\, x^k - \bar{x}\right\rangle + \frac{1}{2}\left\langle \nabla^2 f(\bar{x})(x^k - \bar{x}),\, (x^k - \bar{x})\right\rangle + o(\left\|x^k - \bar{x}\right\|^2) \\ \leq f(\bar{x}) + \frac{\beta_k}{2}\left\|x^k - \bar{x}\right\|^2 \qquad \forall\, k. \tag{1.2}$$

With no loss in generality, we may assume that there is a unit vector u such that $\left\|x^k - \bar{x}\right\|^{-1}(x^k - \bar{x}) \to u \in T\,(\bar{x}\,|\,\Omega)$. Dividing by $\left\|x^k - \bar{x}\right\|$ and letting $k \uparrow \infty$ yields $0 \leq \langle \nabla f(\bar{x}),\, u\rangle \leq 0$ so that $u \in (T\,(\bar{x}\,|\,\Omega) \cap \mathrm{span}\,(\nabla f(\bar{x}))^{\perp}) \setminus \{0\}$. Further note that by Exercise **??**, $(x^k - \bar{x}) \in T\,(\bar{x}\,|\,\Omega)$ for all k so that $\left\langle \nabla f(\bar{x}),\, x^k - \bar{x}\right\rangle \geq 0$ for all k. Hence, (1.2) tells us that

$$\frac{1}{2}\left\langle \nabla^2 f(\bar{x})(x^k - \bar{x}),\, (x^k - \bar{x})\right\rangle + o(\left\|x^k - \bar{x}\right\|^2) \leq \frac{\beta_k}{2}\left\|x^k - \bar{x}\right\|^2 \qquad \forall\, k.$$

Dividing by $\left\|x^k - \bar{x}\right\|^2$ and taking the limit as $k \uparrow \infty$ gives the contradiction $\left\langle \nabla^2 f(\bar{x})u,\, u\right\rangle \leq 0$ which proves the result. $\square$

In later sections we will improve on the second-order conditions in this theorem by delving deeper into the curvature properties of the set Ω. These later results will not only allow us to remove the convexity hypotheses, but will also be stronger even in the convex case. As a first illustration of the limitations of Theorem 1.2, the following example shows that the polyhedrality hypothesis used in the necessary condition cannot be weakened.

Example 1.3. Consider the problem

$$\begin{array}{ll} \min & \frac{1}{2}(x_2 - x_1^2) \\ \text{subject to} & 0 \leq x_2,\ x_1^3 \leq x_2^2. \end{array}$$

Observe that the constraint region in this problem can be written as $\Omega := \{(x_1, x_2) : |x_1|^{\frac{3}{2}} \leq x_2\}$, therefore

$$\begin{aligned} f(x) &= \frac{1}{2}(x_2 - x_1^2) \\ &\geq \frac{1}{2}(|x_1|^{\frac{3}{2}} - |x_1|^2) \\ &= \frac{1}{2}|x_1|^{\frac{3}{2}}(1 - |x_1|^{\frac{1}{2}}) > 0 \end{aligned}$$

whenever $0 < |x_1| \leq 1$. Consequently, the origin is a strict local solution for this problem. Nonetheless,

$$T(0 \,|\, \Omega) \cap [\nabla f(0)]^{\perp} = \{(\delta, 0) : \delta \in \mathbf{R}\},$$

while

$$\nabla^2 f(0) = \begin{bmatrix} -1 & 0 \\ 0 & 0 \end{bmatrix}.$$

That is, even though the origin is a strict local solution, the Hessian of f is negative definite on $T(0 \,|\, \Omega) \cap [\nabla f(0)]^{\perp}$.

The second-order sufficiency condition in Theorem 1.2 is also lacking since, as is shown in the next example, the quadratic growth condition (1.1) can be satisfied even if the hessian is not positive definite on the set $(T(\bar{x} \,|\, \Omega) \cap \operatorname{span}(\nabla f(\bar{x}))^{\perp}) \setminus \{0\}$.

Example 1.4. Consider the problem

$$\begin{array}{ll} \min & x_2 \\ \text{subject to} & x_1^2 \leq x_2. \end{array}$$

Clearly, $\bar{x} = 0$ is the unique global solution to this convex program. Moreover,

$$\begin{aligned} f(\bar{x}) + \frac{1}{2}\|x - \bar{x}\|^2 &= \frac{1}{2}(x_1^2 + x_2^2) \\ &\leq \frac{1}{2}(x_2 + x_2^2) \\ &\leq x_2 = f_0(x) \end{aligned}$$

for all x in the constraint region Ω with $\|x - \bar{x}\| \leq 1$. However, $\nabla^2 f(\bar{x}) = 0$.