Optimality conditions

Optimization problems in standard form

$$\begin{aligned} \text{minimize} \quad & f_0(\mathbf{x}) \\ \text{subject to} \quad & f_i(\mathbf{x}) \le 0, \quad i = 1, \cdots, m \\ & h_i(\mathbf{x}) = 0, \quad i = 1, \cdots, p \end{aligned}$$

- $\mathbf{x} = (x_1, \cdots, x_n) \in \mathbb{R}^n$: optimization variables
- $f_0 : \mathbb{R}^n \to \mathbb{R}$: objective (or cost) function
- $f_i : \mathbb{R}^n \to \mathbb{R}$: inequality constraint functions
- $h_i : \mathbb{R}^n \to \mathbb{R}$: equality constraint functions
- **feasible set**:
  $X = \{x \mid f_i(x) \le 0, i = 1, \cdots, m, h_i(x) = 0, i = 1, \cdots, p\}$

Optimal value:

$$p^* = \inf \{f_0(x) \mid f_i(x) \le 0, i = 1, \cdots, m, h_i(x) = 0, i = 1, \cdots, p\}$$

- $p^* = \infty$ if problem is infeasible
- $p^* = -\infty$ if problem is unbounded below

Optimal and locally optimal points

x^* is an **optimal point** if x^* is feasible (i.e., satisfying the constraints) and $f_0(x^*) = p^*$.

The **optimal set**, denoted X_{opt}, is the set of all optimal points,

A feasible point x with $f_0(x) \leq p^* + \epsilon$ ($\epsilon > 0$) is called ϵ-suboptimal

Definition (locally optimal)

A feasible point x is **locally optimal** if $\exists R > 0$ such that $f(x) \leq f(y)$ for all feasible y that satisfies $\|y - x\|_2 \leq R$. In other words, x solves

$$
\begin{aligned}
\text{minimize} \quad & f_0(\mathbf{z}) \\
\text{subject to} \quad & f_i(\mathbf{z}) \leq 0, \quad i = 1, \cdots, m \\
& h_i(\mathbf{z}) = 0, \quad i = 1, \cdots, p \\
& \|z - x\| \leq R
\end{aligned}
$$

Optimal and locally optimal points: examples

Examples (unconstrained problems):

- $f_0(x) = 1/x$, $\mathbf{dom}\, f_0 = R_{++}$: $p^* = 0$, no optimal point
- $f_0(x) = -\log x$, $\mathbf{dom}\, f_0 = R_{++}$: $p^* = -\infty$, unbounded below
- $f_0(x) = x \log x$, $\mathbf{dom}\, f_0 = R_{++}$: $p^* = -1/e$, $x = 1/e$ is optimal
- $f_0(x) = x^3 - 3x$, $\mathbf{dom}\, f_0 = R$: $p^* = -\infty$, $x = 1$ is locally optimal

Local and global optima

Theorem

Any locally optimal point of a convex optimization problem is also (globally) optimal

Local and global optima

Theorem

Any locally optimal point of a convex optimization problem is also (globally) optimal

Proof.

Suppose x is locally optimal and $y \neq x$ is globally optimal with $f_0(y) < f_0(x)$.
x is locally optimal $\implies \exists R > 0$ such that

$$z \text{ is feasible}, \|z - x\|_2 \leq R \implies f_0(z) \geq f_0(x)$$

Now consider $z = \theta y + (1-\theta)x$ with $\theta = \frac{R}{2\|y-x\|_2}$

- $\|y - x\|_2 > R \implies \theta \in (0, 1/2)$
- z is feasible since it is a convex combination of two feasible points
- $\|z - x\|_2 = R/2$ and $f_0(z) \leq \theta f_0(x) + (1-\theta)f_0(y) < f_0(x)$, which contradicts the assumption that x is locally optimal

□

An optimality criterion for differential f_0

Theorem

Suppose that f_0 in a convex optimization problem is differentiable. Let X denote the feasible set. Then x is optimal if and only if $x \in X$ and

$$\nabla f_0(x)^T (y - x) \geq 0 \quad \forall y \in X$$

An optimality criterion for differential f_0: proof

Proof.

Suppose $x \in X$. We need to prove

$$f_0(x) \leq f_0(y) \quad \forall y \in X \iff \nabla f_0(x)^T(y - x) \geq 0 \quad \forall y \in X$$

- To prove $\Longleftarrow$, suppose $\nabla f_0(x)^T(y - x) \geq 0$ for all $y \in X$. Because f_0 is convex, for all $y \in X$,

$$f_0(y) \geq f_0(x) + \nabla f_0(x)^T(y - x) \geq f_0(x)$$

- To prove $\Longleftarrow$, suppose x is optimal, but there exists a $y \in X$ with $\nabla f_0(x)^T(y - x) < 0$. Consider the point $z(t) = ty + (1 - t)x$ with $t \in [0, 1]$. Clearly $z(t) \in X$. Because

$$\lim_{t \to 0} \frac{f_0(z(t)) - f_0(x)}{t} = \nabla f_0(x)^T(y - x) < 0$$

For sufficiently small t, $f(z) < f(x)$, which contradicts our assumption that x is optimal.

□

An optimality criterion for differential f_0: some special cases

- **unconstrained problem**: x is optimal iff

$$x \in \mathbf{dom}\, f_0, \qquad \nabla f_0(x) = 0$$

- **equality constrained problem** ($Ax = b$): x is optimal iff $\exists \nu$ such that

$$x \in \mathbf{dom}\, f_0, \quad Ax = b, \quad \nabla f_0(x) + A^T \nu = 0$$

- **minimization over nonnegative orthant** ($\min\ f_0(x)$ s.t. $x \succeq 0$): x is optimal iff

$$x \in \mathbf{dom}\, f_0, \quad x \succeq 0, \quad \nabla f_0(x) \succeq 0, \quad \nabla f_0(x)_i x_i = 0$$

First-order optimality condition

Theorem (Optimality condition)

Suppose f_0 is differentiable and the feasible set X is convex.

- *If x^* is a local minimum of f_0 over X, then*

$$\nabla f_0(x^*)^T(x - x^*) \geq 0, \qquad \forall x \in X$$

- *If f_0 is convex, then the above condition is also sufficient for x^* to minimize f_0 over X*

Projection on a convex set

Let $z \in R^n$ and $K \subseteq R^n$ closed, convex set

Project problem:

$$\begin{aligned} \text{minimize} \quad & f(x) = \|z - x\|_2^2 \\ \text{subject to} \quad & x \in K \end{aligned}$$

denoted: find $x^* = \mathrm{Pr}_K(z)$

Projection theorem:

- exists a unique $x^* \in K$ solution to the problem; denote $[z]^+ = x^*$
- x^* is the solution iff $(z - x^*)(x - x^*) \leq 0$ for all $x \in K$
- the map $g : R^n \to K$ with $g(z) = [z]^+$ is continuous, nonexpansive, i.e.,

$$\|[z_1]^+ - [z_2]^+\|_2 \leq \|z_1 - z_2\|_2$$

Projection reformulation of optimality condition

First order optimality condition:

$$\nabla f_0(x^*)^T(x - x^*) \geq 0, \quad \forall x \in X$$

is equivalent to

$$\text{find } x^* \in X: \; x^* = \Pr_K(x^* - \rho \nabla f(x^*) \quad \rho > 0$$

Necessary conditions: Fritz John

Theorem (Fritz John necessary conditions)

Let $\bar{x}$ be a feasible solution of the standard form optimization problem. If $\bar{x}$ is a local minimum, then there exists (u_0, u, v) such that

$$
\begin{aligned}
u_0 \nabla f_0(\bar{x}) + \sum_{i=1}^{m} u_i \nabla f_i(\bar{x}) + \sum_{i=1}^{p} v_i \nabla h_i(\bar{x}) &= 0 \\
(u_0, u) \succeq 0, \quad (u_0, u, v) &\neq 0 \\
u_i f_i(\bar{x}) = 0, \quad i &= 1, \cdots, m
\end{aligned}
$$

Necessary conditions: Karush-Kuhn-Tucker (KKT)

Theorem (KKT necessary conditions)

Let $\bar{x}$ be a feasible solution of the standard form optimization problem and let $I = \{i \mid f_i(\bar{x}) = 0, i = 1, \cdots, m\}$. Suppose that $\nabla f_i(\bar{x})$ for $i \in I$ and $\nabla g_i(\bar{x})$ for $i = 1, \cdots, p$ are linearly independent. If $\bar{x}$ is a local minimum, then there exists (u, v) such that

$$\nabla f_0(\bar{x}) + \sum_{i=1}^{m} u_i \nabla f_i(\bar{x}) + \sum_{i=1}^{p} v_i \nabla h_i(\bar{x}) = 0$$

$$u \succeq 0, \quad u_i f_i(\bar{x}) = 0, \quad i = 1, \cdots, m$$

Sufficient conditions for optimality

The differentiable function $f : R^n \to R$ with convex domain X is **psudoconvex** if $\forall x, y \in X$, $\nabla f(x)^T(y - x) \geq 0$ implies $f(y) \geq f(x)$. (All differentiable convex functions are psudoconvex.) Example: $x + x^3$ is pseudoconvex, but not convex

Theorem (KKT sufficient conditions)

Let $\bar{x}$ be a feasible solution of the standard form optimization problem, and suppose it satifies

$$\nabla f_0(\bar{x}) + \sum_{i=1}^{m} u_i \nabla f_i(\bar{x}) + \sum_{i=1}^{p} v_i \nabla h_i(\bar{x}) = 0$$
$$u \succeq 0, \quad u_i f_i(\bar{x}) = 0, \quad i = 1, \cdots, m$$

If f_0 is pseudoconvex, $f_i(x)$ is quasiconvex for $i = 1, \cdots, m$, and $h_i(x)$ is linear, then $\bar{x}$ is a global optimal solution.