3.1 Basics of Convex Optimization

Let's consider the problem

$$\min_{x \in C} f(x)$$

where $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ is a convex function and C is a convex subset of $\mathbb{R}^n$.

Definition: A point $x \in C \cap \text{dom} f$ is called a feasible point.
If there is at least one feasible point, then the problem is called feasible.
A point x^* is called a *minimum* of f over C if

$$x^* \in C \cap \text{dom} f, \quad f(x^*) = \inf_{x \in C} f(x)$$

We may write $x^* \in \arg\min_{x \in C} f(x)$ or even $x^* = \arg\min_{x \in C} f(x)$ if x^* is the unique minimizer.

Other than global minimum, we also have a weaker definition of local minimum, one that is only minimum compared to the points nearby.

Definition:(Local minimizer) We call x^* a local minimum of f over C if $x^* \in C \cap \text{dom} f$ and there exists $\epsilon > 0$ such that

$$f(x^*) \leq f(x), \ \forall x \in C \text{ with } ||x - x^*|| < \epsilon$$

In the convex setting, we have the following nice result.

Proposition: Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ be a convex function and let C be a convex set. Then a local mimimum of f over C is also a global minimum of f over C. If f is strictly convex, then there exists at most one global minimum of f over C.

Proof. Suppose x^* is a local minimum that is not global.
Then there exists x such that $f(x) < f(x^*)$. Then for $\lambda \in (0, 1)$,

$$f(\lambda x^* + (1 - \lambda)x) \leq \lambda f(x^*) + (1 - \lambda) f(x) < f(x^*)$$

Since f has smaller value on the line connecting x and x^*, this contradicts the local minimality of x^*.
Suppose f is strictly convex, let x^* be a global minimum of f over C. Let $x \in C$ such that $x \neq x^*$. Consider $y = (x + x^*)/2$. Then $y \in C$ and

$$f(y) < \frac{1}{2}(f(x) + f(x^*)) \leq f(x)$$

Since x^* is a global minimum, $f(x^*) \leq f(y)$.
Then $f(x^*) < f(x)$. Hence x^* is the unique global minimum of f over C. $\square$

3.1.1 Existence of solution

Let's consider a general optimization problem

$$\min_{x \in C} f(x)$$

where $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ and $C \subseteq \mathbb{R}^n$.
A basic question is whether a solution to the above problem exists.
Recall the famous Weierstrass theorem. **Proposition:** If f is continuous and C is compact, then there exists a global minimum.

In order to consider cases where C is not bounded (e.g. $\mathbb{R}^n$), we need a new notation.

Definition: (Coercivity) A function $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ is called *coercive* if for all sequence $\{x_k\}$ with $\|x_k\| \to \infty$, we have $\lim_{k\to\infty} f(x_k) = \infty$.

Lemma: Let $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ be a continuous function. Then the following are equivalent.

1. All level sets of f are compact, i.e. $\{x| \ f(x) \leq a\}$ is compact for all a.
2. f is coercive.

Proof. Suppose all level sets of f are compact. Suppose $\{x_k\}$ is a sequence with $||x_k|| \to \infty$. Suppose $f(x_k) \not\to \infty$. Then there exists subsequence x_{k_j} such that $f(x_{k_j})$ is bounded by α for some α. Then $\{x_{k_j}\} \subset V_\alpha$. This contradicts the compactness of V_α. Hence, f is coercive.
Conversely, suppose f is coercive. Suppose V_α is not compact for some α. Since f is continuous, V_α must be closed, this means V_α is not bounded.
Hence, there exists a sequence $\{x_k\} \subset V_\alpha$ such that $||x_k|| \to \infty$. This contradicts the coercivity of f since $f(x_k) \leq \alpha$. □

Proposition: Suppose f is lower-semicontinuous and coercive. Suppose C is non-empty and closed. Then f has a global minimum over C.

Proof. We may assume that $f(x) < \infty$ for some $x \in C$. Then $f^* = \inf_{x\in C} f(x) < \infty$.
Let $\{x_k\} \subset C$ be a sequence such that $\lim f(x_k) = f^* < \infty$. Then since f is coercive, $\{x_k\}$ is bounded. Then there exists a subsequence x_{k_j} converging to a point x^*.
Since C is closed, $x^* \in C$. Then

$$f^* = \lim_{k\to\infty} f(x_k) = \lim_{j\to\infty} f(x_{k_j}) \geq f(x^*)$$

Therefore, x^* is a global minimmum of f over C. □

3.1.2 Optimal condition

For a unconstrained problem, one has a simple optimality test, which is the 'derivative' test in calculus.

Let f be a differentiable convex function on $\mathbb{R}^n$. Then x^* solves

$$\min_{x\in\mathbb{R}^n} f(x)$$

if and only if $\nabla f(x^*) = 0$. How about a constrained problem?
Let's consider the general constrained problem

$$\min_{x\in C} f(x)$$

where C is a convex set, and f is convex.
We have the following result.

Proposition: Let C be a nonempty convex set and let $f : \mathbb{R}^n \to \mathbb{R}$ be a convex differentiable function over an open set that contains C. Then $x^* \in C$ minimizes f over C if and only if

$$\langle \nabla f(x^*), (z - x^*) \rangle \geq 0, \ \forall z \in C.$$

Proof. Suppose $\langle \nabla f(x^*), (z - x^*) \rangle \geq 0, \ \forall z \in C$, then we have,

$$f(z) - f(x^*) \geq \langle \nabla f(x^*), (z - x^*) \rangle \geq 0, \ \forall z \in C.$$

Hence x^* indeed minimizes f over C.
Conversely, suppose x^* minimizes f over C. Suppose on the contrary that $\langle \nabla f(x^*), (z - x^*) \rangle < 0$ for some $z \in C$, then

$$\lim_{\alpha\downarrow 0} \frac{f(x^* + \alpha(z - x^*)) - f(x^*)}{\alpha} = \langle \nabla f(x^*), (z - x^*) \rangle < 0.$$

Then for sufficiently small α, we have $f(x^*+\alpha(z-x^*))-f(x^*) < 0$, contradicting the optimality of x^*. □

Examples (a) Let's consider the following linear constrained problem.

$$\min_{x\in\mathbb{R}^n} f(x) \text{ subject to } Ax = b$$

where A is a $m \times n$ matrix and $b \in \mathbb{R}^m$.
Suppose we have a solution x^*, then

$$\langle \nabla f(x^*), y - x^* \rangle \geq 0, \ \forall y \text{ such that } Ay = b$$

This is the same as

$$\langle \nabla f(x^*), h \rangle \geq 0, \ \forall h \in N(A).$$

Since $-h \in N(A)$ if $h \in N(A)$, we have

$$\langle \nabla f(x^*), h \rangle = 0, \ \forall h \in N(A).$$

Hence $\nabla f(x^*) \in N(A)^\perp = R(A^T)$.
So there exists $\mu \in \mathbb{R}^m$ such

$$\nabla f(x^*) + A^T \mu = 0.$$

To conclude, x^* is a solution to the minimization problem if and only if

1. $Ax^* = b$
2. There exists $\mu^* \in \mathbb{R}^m$ such that $\nabla f(x^*) + A^T \mu = 0$.

(b) Let's consider the minimization problem

$$\min_{x \in \mathbb{R}^n} f(x), \text{ subject to } x \geq 0.$$

Suppose we have a solution x^*, then

$$\langle \nabla f(x^*), y - x^* \rangle \geq 0, \ \forall y \in \mathbb{R}^n_+.$$

In particular, $0, 2x^* \in \mathbb{R}^n_+$, so

$$\langle \nabla f(x^*), x^* \rangle = 0, \ \langle \nabla f(x^*), y \rangle \geq 0, \ \forall y \in \mathbb{R}^n_+.$$

Hence, $\nabla f(x^*) \geq 0$. This is the same as saying there exists $\lambda^* \geq 0$ such that

$$\nabla f(x^*) - \lambda^* = 0$$

To conclude, x^* is a solution if and only if

1. $x^* \geq 0$
2. There exists $\lambda^* \geq 0$ such that $\nabla f(x^*) - \lambda^* = 0$
3. $\lambda_i^* x_i^* = 0$