1 Basic concepts of convex optimization

In convex optimization, we consider the problem

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

where $f : \mathbb{R}^n \to (-\infty, \infty]$ is a convex function and C is a convex subset of $\mathbb{R}^n$.
If $x \in C \cap \operatorname{dom}(f)$, then x is called feasible. If there is at least one feasible point, then the problem is called feasible.
x^* is called a minimum of f over C if

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

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 mimimum, one that is only minimum compared to the point nearby. We call x^* a local minimum of f over C if $x^* \in C \cap \operatorname{dom}(f)$ 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 $f : \mathbb{R}^n \to (-\infty, \infty]$ 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.

Existence of solution
Consider the problem

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

where f is convex.
Suppose the level sets $V_a = \{x | \ f(x) \leq a\}$ are also compact. Then we can consider the problem

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

for some V_a that is nonempty. Then there exist at least one global minimizer.
Remark: We can also show that f is coercive, which is equivalent to the level sets of f are compact.

1.1 Optimal conditions

In 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^*. □

1.2 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

$$\langle \nabla f(x^*), h \rangle \geq 0, \ \forall h \in \text{Null}(A).$$

Since $-h \in \text{Null}(A)$ if $h \in \text{Null}(A)$, we have

$$\langle \nabla f(x^*), h \rangle = 0, \ \forall h \in \text{Null}(A).$$

Hence $\nabla f(x^*) \in \text{Null}(A)^\perp = \text{Ran}(A^T)$.
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. $\lambda_i^* x_i^* = 0$