

MATH 4211/6211 – Optimization

Convex Optimization Problems

Xiaojing Ye

Department of Mathematics & Statistics

Georgia State University

Definition. A set $\Omega \subset \mathbb{R}^n$ is called *convex* if for any $\mathbf{x}, \mathbf{y} \in \Omega$, there is $\alpha \mathbf{x} + (1 - \alpha) \mathbf{y} \in \Omega$ for all $\alpha \in [0, 1]$

Definition. A function $f : \Omega \rightarrow \mathbb{R}$, where Ω is a convex set, is called *convex* if for any $\mathbf{x}, \mathbf{y} \in \Omega$ and $\alpha \in [0, 1]$, there is

$$f(\alpha \mathbf{x} + (1 - \alpha) \mathbf{y}) \leq \alpha f(\mathbf{x}) + (1 - \alpha) f(\mathbf{y}).$$

Moreover, f is called *strictly convex* if for any distinct $\mathbf{x}, \mathbf{y} \in \Omega$ and $\alpha \in (0, 1)$, there is

$$f(\alpha \mathbf{x} + (1 - \alpha) \mathbf{y}) < \alpha f(\mathbf{x}) + (1 - \alpha) f(\mathbf{y})$$

A function f is called (strictly) concave if $-f$ is (strictly) convex.

There is an alternative definition based on the convexity of the epigraph of f .

Definition. The *graph* of $f : \Omega \rightarrow \mathbb{R}$ is defined by

$$\{[\mathbf{x}; f(\mathbf{x})] \in \mathbb{R}^{n+1} : \mathbf{x} \in \Omega\}$$

Definition. The *epigraph* of $f : \Omega \rightarrow \mathbb{R}$ is defined by

$$\text{epi}(f) := \{[\mathbf{x}; \beta] \in \mathbb{R}^{n+1} : \mathbf{x} \in \Omega, \beta \geq f(\mathbf{x})\}$$

Definition. A function $f : \Omega \rightarrow \mathbb{R}$, where Ω is a convex set, is called *convex* if $\text{epi}(f)$ is a convex set.

Example. Let $f(x) = x_1x_2$ be defined on $\Omega := \{x : x \geq 0\}$. Is f convex?

Solution. f is *not* convex. The set $\Omega \subset \mathbb{R}^2$ is convex. But if we choose $x = [1; 2]$ and $y = [2; 1]$, then

$$\alpha x + (1 - \alpha)y = [2 - \alpha; 1 + \alpha].$$

On the one hand

$$f(\alpha x + (1 - \alpha)y) = 2 + \alpha - \alpha^2.$$

On the other hand,

$$\alpha f(x) + (1 - \alpha)f(y) = 2.$$

Choosing $\alpha = 1/2$ yields

$$f(\alpha x + (1 - \alpha)y) > \alpha f(x) + (1 - \alpha)f(y)$$

which means that f is not convex.

There are several *sufficient and necessary* conditions for the convexity of f .

Theorem. If $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is \mathcal{C}^1 and Ω is convex, then f is convex on Ω iff for all $\mathbf{x}, \mathbf{y} \in \Omega$,

$$f(\mathbf{y}) \geq f(\mathbf{x}) + \nabla f(\mathbf{x})^\top (\mathbf{y} - \mathbf{x})$$

Proof. (\Rightarrow) Suppose f is convex, then for any $\mathbf{x}, \mathbf{y} \in \Omega$ and $\alpha \in (0, 1]$,

$$f((1 - \alpha)\mathbf{x} + \alpha\mathbf{y}) \leq (1 - \alpha)f(\mathbf{x}) + \alpha f(\mathbf{y})$$

Rearrange terms to obtain

$$\frac{f(\mathbf{x} + \alpha(\mathbf{y} - \mathbf{x})) - f(\mathbf{x})}{\alpha} \leq f(\mathbf{y}) - f(\mathbf{x})$$

Taking the limit as $\alpha \rightarrow 0$ yields

$$f(\mathbf{y}) \geq f(\mathbf{x}) + \nabla f(\mathbf{x})^\top (\mathbf{y} - \mathbf{x})$$

Proof (cont.) (\Leftarrow) For any $\mathbf{x}, \mathbf{y} \in \Omega$ and $\alpha \in [0, 1]$, define $\mathbf{x}_\alpha = \alpha\mathbf{x} + (1 - \alpha)\mathbf{y}$. Then

$$\begin{aligned} f(\mathbf{x}) &\geq f(\mathbf{x}_\alpha) + \nabla f(\mathbf{x}_\alpha)^\top (\mathbf{x}_\alpha - \mathbf{x}) \\ f(\mathbf{y}) &\geq f(\mathbf{x}_\alpha) + \nabla f(\mathbf{x}_\alpha)^\top (\mathbf{x}_\alpha - \mathbf{y}) \end{aligned}$$

Multiplying the two inequalities by α and $1 - \alpha$ respectively, and adding together yields

$$f(\alpha\mathbf{x} + (1 - \alpha)\mathbf{y}) \leq \alpha f(\mathbf{x}) + (1 - \alpha)f(\mathbf{y})$$

Theorem. Let $f : \mathbb{R}^n \rightarrow \mathbb{R}$ be \mathcal{C}^2 and Ω be convex, then f is convex on Ω iff $\nabla^2 f(\mathbf{x}) \succeq \mathbf{0}$ for all $\mathbf{x} \in \Omega$.

Proof. (\Rightarrow) If not, then exist $\mathbf{x} \in \Omega$ and $\mathbf{d} \in \mathbb{R}^n$, such that

$$\mathbf{d}^\top \nabla^2 f(\mathbf{x}) \mathbf{d} < 0$$

Since $\nabla^2 f(\mathbf{x})$ is continuous, there exists $s > 0$ sufficiently small, such that for $\mathbf{y} = \mathbf{x} + s\mathbf{d} \in \Omega$, there is

$$\begin{aligned} f(\mathbf{y}) &= f(\mathbf{x}) + \nabla f(\mathbf{x})^\top (\mathbf{y} - \mathbf{x}) + \frac{1}{2}(\mathbf{y} - \mathbf{x})^\top \nabla^2 f(\mathbf{x} + t\mathbf{d})(\mathbf{y} - \mathbf{x}) \\ &< f(\mathbf{x}) + \nabla f(\mathbf{x})^\top (\mathbf{y} - \mathbf{x}) \end{aligned}$$

for some $t \in (0, s)$ since $(\mathbf{y} - \mathbf{x})^\top \nabla^2 f(\mathbf{x} + t\mathbf{d})(\mathbf{y} - \mathbf{x}) = s^2 \mathbf{d}^\top \nabla^2 f(\mathbf{x} + t\mathbf{d}) \mathbf{d} < 0$. Hence f is not convex, a contradiction.

Proof (cont.) (\Leftarrow) For any $\mathbf{x}, \mathbf{y} \in \Omega$, there is

$$\begin{aligned} f(\mathbf{y}) &= f(\mathbf{x}) + \nabla f(\mathbf{x})^\top (\mathbf{y} - \mathbf{x}) + \frac{1}{2}(\mathbf{y} - \mathbf{x})^\top \nabla^2 f(\mathbf{x} + t\mathbf{d})(\mathbf{y} - \mathbf{x}) \\ &\geq f(\mathbf{x}) + \nabla f(\mathbf{x})^\top (\mathbf{y} - \mathbf{x}) \end{aligned}$$

where $\mathbf{d} := \mathbf{y} - \mathbf{x}$ and $t \in (0, 1)$. Note that we used the fact that $\nabla^2 f(\mathbf{x} + t\mathbf{d}) \succeq \mathbf{0}$. Hence f is convex.

Examples. Determine if any of the following functions is convex.

$$f_1(x) = -8x^2$$

$$f_2(x) = 4x_1^2 + 3x_2^2 + 5x_3^2 + 6x_1x_2 + x_1x_3 - 3x_1 - 2x_2 + 15$$

$$f_3(x) = 2x_1x_2 - x_1^2 - x_2^2$$

Solution. $f_1''(x) = -16 < 0$, so f_1 is concave.

For f_2 , we have

$$\nabla^2 f_2 = \begin{bmatrix} 8 & 6 & 1 \\ 6 & 6 & 0 \\ 1 & 0 & 10 \end{bmatrix}$$

whose leading principal minors are 8, 12, 114. Hence f_2 is convex.

For f_3 , we have

$$\nabla^2 f_3 = \begin{bmatrix} -2 & 2 \\ 2 & -2 \end{bmatrix}$$

whose eigenvalues are -4 and 0 , hence f_3 is negative semidefinite.

Theorem. Suppose $f : \Omega \rightarrow \mathbb{R}$ is convex. Then x is a global minimizer of f on Ω iff it is a local minimizer of f .

Proof. The necessity is trivial. Suppose x is a local minimizer, then $\exists r > 0$ such that $f(x) \leq f(z)$ for all $z \in B(x, r)$. If $\exists y$, such that $f(x) > f(y)$, then let $\alpha = \frac{r}{\|y-x\|}$ and

$$x_\alpha = (1 - \alpha)x + \alpha y = x + \frac{r}{\|y-x\|}(y-x).$$

Then $x_\alpha \in B(x, r)$ and

$$f(x_\alpha) \geq f(x) > (1 - \alpha)f(x) + \alpha f(y),$$

which is a contradiction. Hence x must be a global minimizer.

Lemma. Suppose $f : \Omega \rightarrow \mathbb{R}$ is convex. Then the *sub-level set* of f

$$\Gamma_c = \{x \in \Omega : f(x) \leq c\}$$

is empty or convex for any $c \in \mathbb{R}$.

Proof. If $x, y \in \Gamma_c$, then $f(x), f(y) \leq c$. Since f is convex, we have

$$f(\alpha x + (1 - \alpha)y) \leq \alpha f(x) + (1 - \alpha)f(y) \leq c$$

i.e., $\alpha x + (1 - \alpha)y \in \Gamma_c$ for all $\alpha \in [0, 1]$. Hence Γ_c is a convex set.

Corollary. Suppose $f : \Omega \rightarrow \mathbb{R}$ is convex. Then the set of all global minimizers of f over Ω is convex.

Proof. Let $f^* = \min_{x \in \Omega} f(x)$. Then Γ_{f^*} is the set of all global minimizers. By the lemma above, we know Γ_c is a convex set.

Lemma. Suppose $f : \Omega \rightarrow \mathbb{R}$ is convex and \mathcal{C}^1 . Then x^* is a global minimizer of f over Ω iff

$$\nabla f(x^*)^\top (x - x^*) \geq 0, \quad \forall x \in \Omega.$$

Proof. (\Rightarrow) If not, then $\exists x \in \Omega$, s.t.

$$\nabla f(x^*)^\top (x - x^*) < 0$$

Denote $x_\alpha = (1 - \alpha)x^* + \alpha x = x^* + \alpha(x - x^*)$ for $\alpha \in (0, 1)$. Since $f \in \mathcal{C}^1$, we know there exists α small enough, s.t.

$$\nabla f(x_{\alpha'})^\top (x - x^*) < 0, \quad \forall \alpha' \in (0, \alpha)$$

Proof (cont.) Moreover, there exists $\alpha' \in (0, \alpha)$ s.t.

$$\begin{aligned} f(\mathbf{x}_\alpha) &= f(\mathbf{x}^*) + \nabla f(\mathbf{x}_{\alpha'})^\top (\mathbf{x}_\alpha - \mathbf{x}^*) \\ &= f(\mathbf{x}^*) + \alpha \nabla f(\mathbf{x}_{\alpha'})^\top (\mathbf{x} - \mathbf{x}^*) \\ &< f(\mathbf{x}^*) \end{aligned}$$

which contradicts to \mathbf{x}^* being a global minimizer.

(\Leftarrow) For all $\mathbf{x} \in \Omega$, there is

$$f(\mathbf{x}) \geq f(\mathbf{x}^*) + \nabla f(\mathbf{x}^*)^\top (\mathbf{x} - \mathbf{x}^*) \geq f(\mathbf{x}^*)$$

Hence \mathbf{x}^* is a global minimizer.

Theorem. Suppose $f : \Omega \rightarrow \mathbb{R}$ is convex and \mathcal{C}^1 . Then x^* is a global minimizer of f over Ω iff for any feasible direction d at x^* there is

$$d^\top \nabla f(x^*) \geq 0.$$

Proof. (\Rightarrow) Let d be feasible, then $\exists x \in \Omega$ s.t. $x - x^* = \alpha d$ for some $\alpha > 0$. Hence by the Lemma above, we have

$$\nabla f(x^*)^\top (x - x^*) = \alpha \nabla f(x^*)^\top d \geq 0.$$

So $\nabla f(x^*)^\top d \geq 0$.

(\Leftarrow) For any $x \in \Omega$, we know $x_\alpha = (1 - \alpha)x^* + \alpha x \in \Omega$ for all $\alpha \in (0, 1)$. Hence $d = x - x^* = (x_\alpha - x^*)/\alpha$ is a feasible direction. Therefore

$$\nabla f(x^*)^\top (x - x^*) = \nabla f(x^*)^\top d \geq 0.$$

As $x \in \Omega$ is arbitrary, we know x^* is a global minimizer.

Corollary. Suppose $f : \Omega \rightarrow \mathbb{R}$ is convex and \mathcal{C}^1 . If $\mathbf{x}^* \in \Omega$ is such that

$$\nabla f(\mathbf{x}^*) = \mathbf{0},$$

then \mathbf{x}^* is a global minimizer of f .

Proof. For any feasible \mathbf{d} there is $\nabla f(\mathbf{x}^*)^\top \mathbf{d} = 0$. Hence \mathbf{x}^* is a global minimizer.

Theorem. Let $f : \mathbb{R}^n \rightarrow \mathbb{R}$ and $f \in \mathcal{C}^1$ be convex, and $\Omega = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{h}(\mathbf{x}) = \mathbf{0}\}$ where $\mathbf{h} : \mathbb{R}^n \rightarrow \mathbb{R}^m$ such that Ω is convex. Then $\mathbf{x}^* \in \Omega$ is a global minimizer of f over Ω iff there exists $\boldsymbol{\lambda}^* \in \mathbb{R}^m$ such that

$$\nabla f(\mathbf{x}^*) + D\mathbf{h}(\mathbf{x}^*)^\top \boldsymbol{\lambda}^* = \mathbf{0}.$$

Proof. (\Rightarrow) By the KKT condition.

(\Leftarrow) Note that f being convex implies

$$f(\mathbf{x}) \geq f(\mathbf{x}^*) + \nabla f(\mathbf{x}^*)^\top (\mathbf{x} - \mathbf{x}^*), \quad \forall \mathbf{x} \in \Omega$$

Also note that $\nabla f(\mathbf{x}^*) = -D\mathbf{h}(\mathbf{x}^*)^\top \boldsymbol{\lambda}^*$, we know

$$f(\mathbf{x}) \geq f(\mathbf{x}^*) - \boldsymbol{\lambda}^{*\top} D\mathbf{h}(\mathbf{x}^*)(\mathbf{x} - \mathbf{x}^*)$$

Proof (cont.) For any $x \in \Omega$, we know $x^* + \alpha(x - x^*) \in \Omega$ for all $\alpha \in (0, 1)$. Hence $h(x^* + \alpha(x - x^*)) = 0$ and

$$Dh(x^*)(x - x^*) = \lim_{\alpha \rightarrow 0} \frac{h(x^* + \alpha(x - x^*)) - h(x^*)}{\alpha} = 0$$

Hence $f(x) \geq f(x^*)$ for all $x \in \Omega$. Therefore x^* is a global minimizer.

Theorem. Let $f : \mathbb{R}^n \rightarrow \mathbb{R}$ and $f \in \mathcal{C}^1$ be convex, and

$$\Omega = \{x \in \mathbb{R}^n : h(x) = 0, g(x) \leq 0\}$$

where $h : \mathbb{R}^n \rightarrow \mathbb{R}^m$ and $g : \mathbb{R}^n \rightarrow \mathbb{R}^p$ are \mathcal{C}^1 and such that Ω is convex. Then $x^* \in \Omega$ is a global minimizer of f over Ω iff there exist $\lambda^* \in \mathbb{R}^m, \mu^* \in \mathbb{R}_+^p$ such that

$$\begin{aligned} \nabla f(x^*)^\top + \lambda^{*\top} Dh(x^*) + \mu^{*\top} Dg(x^*) &= \mathbf{0}^\top, \\ g(x^*)^\top \mu^* &= 0. \end{aligned}$$

Proof. (\Rightarrow) By the KKT condition.

Proof (cont.) (\Leftarrow) Note that f being convex implies

$$f(\mathbf{x}) \geq f(\mathbf{x}^*) + \nabla f(\mathbf{x}^*)^\top (\mathbf{x} - \mathbf{x}^*), \quad \forall \mathbf{x} \in \Omega.$$

Also note that $\nabla f(\mathbf{x}^*) = -D\mathbf{h}(\mathbf{x}^*)^\top \boldsymbol{\lambda}^* - D\mathbf{g}(\mathbf{x}^*)^\top \boldsymbol{\mu}^*$, we know

$$f(\mathbf{x}) \geq f(\mathbf{x}^*) - \boldsymbol{\lambda}^{*\top} D\mathbf{h}(\mathbf{x}^*)(\mathbf{x} - \mathbf{x}^*) - \boldsymbol{\mu}^{*\top} D\mathbf{g}(\mathbf{x}^*)(\mathbf{x} - \mathbf{x}^*).$$

For any $\mathbf{x} \in \Omega$, we know $\mathbf{x}^* + \alpha(\mathbf{x} - \mathbf{x}^*) \in \Omega$ for all $\alpha \in (0, 1)$. Hence $\mathbf{h}(\mathbf{x}^* + \alpha(\mathbf{x} - \mathbf{x}^*)) = 0$ and

$$D\mathbf{h}(\mathbf{x}^*)(\mathbf{x} - \mathbf{x}^*) = \lim_{\alpha \rightarrow 0} \frac{\mathbf{h}(\mathbf{x}^* + \alpha(\mathbf{x} - \mathbf{x}^*)) - \mathbf{h}(\mathbf{x}^*)}{\alpha} = 0.$$

Proof (cont.) Moreover $g(x^* + \alpha(x - x^*)) \leq 0$, and hence $\mu^* \geq 0$ implies

$$\mu^{*\top} g(x^* + \alpha(x - x^*)) \leq 0.$$

Therefore, we have

$$\mu^{*\top} Dg(x^*)(x - x^*) = \lim_{\alpha \rightarrow 0} \frac{\mu^{*\top} g(x^* + \alpha(x - x^*)) - \mu^{*\top} g(x^*)}{\alpha} \leq 0$$

Hence we obtain

$$f(x) \geq f(x^*), \quad \forall x \in \Omega.$$

Therefore x^* is a global minimizer.

Example. Suppose we can deposit $x_i \geq 0$ amount of money into a bank account (with initial balance 0) at the beginning of the i th month for $i = 1, \dots, n$. The monthly interest rate is $r > 0$. If the total amount we can deposit is D , then find the way to maximize the total balance including the interests at the end of the n th month.

Intuition. We should deposit all money D in the first month.

Solution. Define $c = -[(1 + r)^n; \dots; (1 + r)] \in \mathbb{R}^n$. Then this is an LP (which is a convex program):

$$\begin{aligned} & \text{minimize} && c^\top x \\ & \text{subject to} && e^\top x = D \\ & && x \geq 0 \end{aligned}$$

where $e = [1; \dots; 1] \in \mathbb{R}^n$.

Solution (cont.) To show our intuition is correct, it suffices to show that $x^* = [D; 0; \dots; 0] \in \mathbb{R}^n$ satisfies the KKT condition:

$$\begin{aligned}c + \lambda^* e - \mu^* &= 0 \\ e^\top x^* &= D \\ x^* &\geq 0 \\ \mu^* &\geq 0 \\ \mu^{*\top} x^* &= 0\end{aligned}$$

for some $\lambda \in \mathbb{R}$ and $\mu^* \in \mathbb{R}^n$.

Let $\lambda^* = (1 + r)^n$ and $\mu^* = c + \lambda^* e$, then it is easy to verify that (x^*, λ^*, μ^*) satisfies the KKT condition. This implies that x^* is a global minimizer.