Convex optimization problems (Chap. 21)

- Consider the general problem

$$\begin{array}{ll} \text{minimize} & f(\boldsymbol{x}) \\ \text{subject to} & \boldsymbol{x} \in \Omega. \end{array}$$

- We have seen several types of FONC.
- When is a FONC sufficient for global optimality?
- Answer: In a *convex* optimization problem.

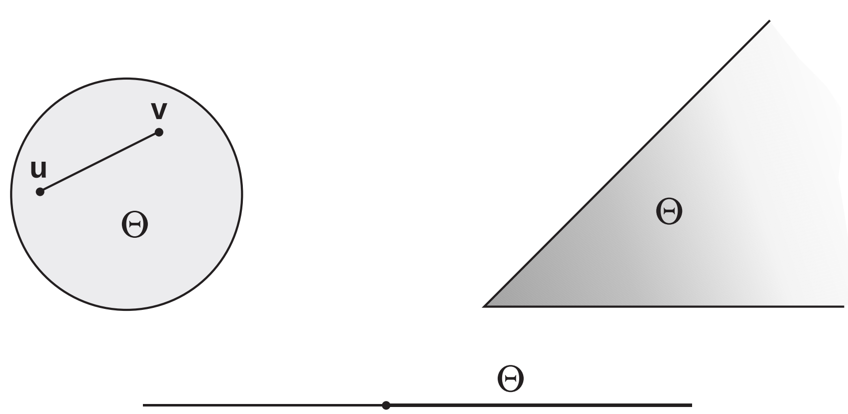
Summary of FONCs

- Set constraint: $\boldsymbol{d}^T \nabla f(\boldsymbol{x}^*) \geq 0$ for all feasible directions $\boldsymbol{d}$;
- Interior: $\nabla f(\boldsymbol{x}^*) = \mathbf{0}$;
- $\Omega = \{\boldsymbol{x} : \boldsymbol{h}(\boldsymbol{x}) = \mathbf{0}\}$: Lagrange conditions;
- $\Omega = \{\boldsymbol{x} : \boldsymbol{h}(\boldsymbol{x}) = \mathbf{0}, \boldsymbol{g}(\boldsymbol{x}) \leq \mathbf{0}\}$: KKT conditions.

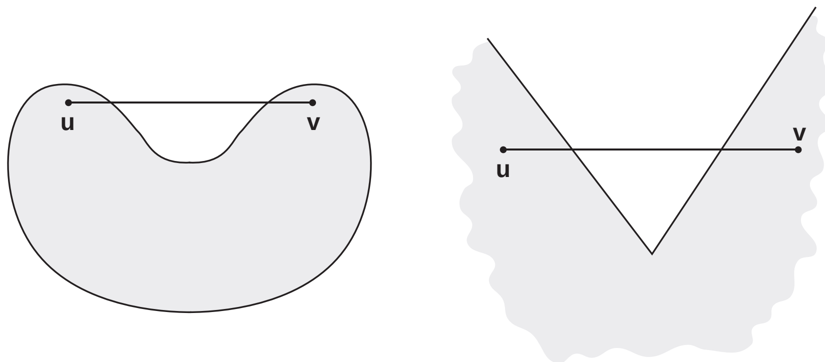
Set convexity (§4.3)

- Given: $\Omega \subset \mathbb{R}^n$.
- Definition: Ω is a convex set if, for any distinct $\boldsymbol{y}, \boldsymbol{z} \in \Omega$ and $\alpha \in (0,1)$, we have $\alpha\boldsymbol{y} + (1-\alpha)\boldsymbol{z} \in \Omega$.
- Convex set: the line segment joining any two points in the set lies completely inside the set.

Convex sets:



Nonconvex sets:



Examples of convex sets

- The empty set
- A set consisting of a single point
- A line or a line segment
- A subspace
- A hyperplane
- A linear variety
- A half-space
- $\mathbb{R}^n$

Example: Prove that $\Omega = \{\boldsymbol{x} : \boldsymbol{x} \geq \boldsymbol{0}\}$ is convex.

- Let $\boldsymbol{y}, \boldsymbol{z} \in \Omega$, and $\alpha \in (0, 1)$.
- Want to show that $\alpha \boldsymbol{y} + (1 - \alpha)\boldsymbol{z} \in \Omega$.
- Consider $\boldsymbol{x} = \alpha \boldsymbol{y} + (1 - \alpha)\boldsymbol{z}$. What does $\boldsymbol{x} \in \Omega$ mean?
- To qualify as a member of Ω, each of its component must be ≥ 0.
- Hence, we must show that each component of $\boldsymbol{x}$ is ≥ 0.
- Each component of $\boldsymbol{x} = [x_1, \ldots, x_n]^T$ satisfies $x_i = \alpha y_i + (1 - \alpha) z_i$.
- Note that we have $y_i, z_i, \alpha, 1 - \alpha \geq 0$.
- Hence, $x_i \geq 0$; i.e., $\boldsymbol{x} \geq \boldsymbol{0}$, which means that $\boldsymbol{x} \in \Omega$.

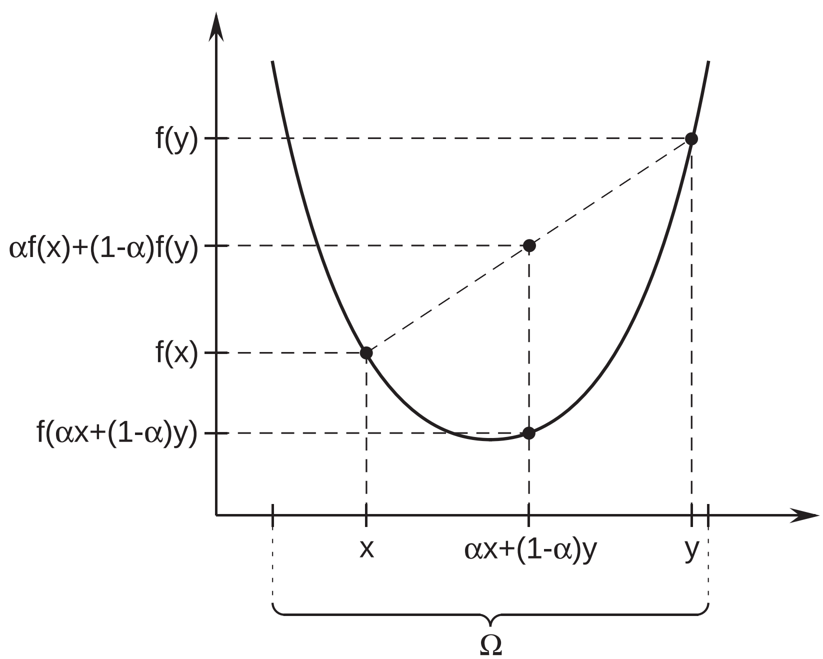
- Therefore, Ω is convex.

Exercise: Prove that $\Omega = \{\boldsymbol{x} : \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}\}$ is convex.

Exercise: How do we prove that a set is not convex?

Function convexity

- Given: a function $f : \Omega \to \mathbb{R}$, where Ω is convex.
- Definition: f is a convex function on Ω if, for any distinct $\boldsymbol{y}, \boldsymbol{z} \in \Omega$ and $\alpha \in (0, 1)$, $f(\alpha\boldsymbol{y} + (1-\alpha)\boldsymbol{z}) \le \alpha f(\boldsymbol{y}) + (1-\alpha) f(\boldsymbol{z})$.
- f is strictly convex if $\le$ is replaced by $<$.
- Convex function: line segment joining two points on the graph lies above the graph.
- f is said to be (strictly) concave if $-f$ is (strictly) convex.



Example: (21.4)

- Consider the function $f(\boldsymbol{x}) = x_1 x_2$. Is f convex over $\Omega = \{\boldsymbol{x} : x_1 \ge 0,\ x_2 \ge 0\}$?
- Answer: No.
- Consider $\boldsymbol{y} = [2, 1]^T \in \Omega$, $\boldsymbol{z} = [1, 2]^T \in \Omega$, and $\alpha = 1/2$.
- We have $\boldsymbol{x} = \alpha\boldsymbol{y} + (1-\alpha)\boldsymbol{z} = [3/2, 3/2]^T$. Hence, $f(\boldsymbol{x}) = 9/4$.

- On the other hand,

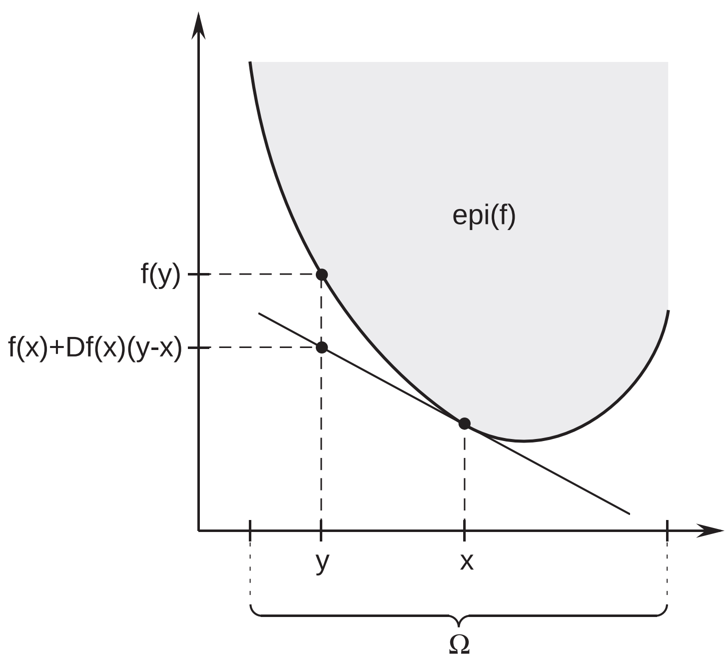
$$\alpha f(\boldsymbol{y}) + (1-\alpha) f(\boldsymbol{z}) = 2 < f(\alpha \boldsymbol{y} + (1-\alpha)\boldsymbol{z}).$$

Alternative way of interpreting function convexity

- Suppose $f : \Omega \to \mathbb{R}$, Ω convex and open, and $f \in \mathcal{C}^1$.
- Theorem (21.3): f is convex iff for all distinct $\boldsymbol{x}, \boldsymbol{y} \in \Omega$,

$$f(\boldsymbol{y}) \geq f(\boldsymbol{x}) + Df(\boldsymbol{x})(\boldsymbol{y} - \boldsymbol{x}).$$

- Interpretation: f convex means that it lies above any linear approximation of it.
- For strict convexity, replace $\geq$ by $>$.



- Suppose $f : \Omega \to \mathbb{R}$, Ω convex and open, and $f \in \mathcal{C}^2$. Let $\boldsymbol{F}(\boldsymbol{x})$ be the Hessian of f at $\boldsymbol{x}$.
- Theorem (21.4): f is convex if and only if $\boldsymbol{F}(\boldsymbol{x}) \geq 0$ for all $\boldsymbol{x} \in \Omega$.
- For strict convexity, $\boldsymbol{F}(\boldsymbol{x}) > 0$ is sufficient, but not necessary (e.g., $f(x) = x^4$ is strictly convex but $f''(0) = 0$).
- If Ω is not open, but $\boldsymbol{F}(\boldsymbol{x}) \geq 0$ for all $\boldsymbol{x}$ in an open set that contains Ω, then we conclude that f is convex.

Examples

- $f(x) = x^3$, $\Omega = (0,1)$. We have $f''(x) = 6x \geq 0$ on Ω. Hence, f is convex on Ω.
- $f(x) = -x^2$, $\Omega = \mathbb{R}$. We have $f''(x) = -2 < 0$. Hence, f is strictly concave on Ω.
- For a quadratic with Hessian $\boldsymbol{Q}$, convexity on $\mathbb{R}^n$ is equivalent to $\boldsymbol{Q} \geq 0$. Strict convexity is equivalent to $\boldsymbol{Q} > 0$.

Checking convexity for quadratics

- Proposition (21.1): Consider the quadratic function $f(\boldsymbol{x}) = \boldsymbol{x}^T\boldsymbol{Q}\boldsymbol{x}$, where $\boldsymbol{Q} = \boldsymbol{Q}^T$. Suppose Ω is a convex set. Then, the f is a convex function on Ω iff

$$(\boldsymbol{x} - \boldsymbol{y})^T\boldsymbol{Q}(\boldsymbol{x} - \boldsymbol{y}) \geq 0$$

  for all $\boldsymbol{x}, \boldsymbol{y} \in \Omega$.
- See Example 21.5.

Convex optimization problems

- Consider

$$\begin{array}{ll} \text{minimize} & f(\boldsymbol{x}) \\ \text{subject to} & \boldsymbol{x} \in \Omega, \end{array}$$

  where Ω is a convex set, and f is a convex function on Ω.
- Name: *Convex programming problem*, or *convex optimization problem*.
- Examples: LP, QP.

Theorem (21.5): In a convex programming problem, a point is a global minimizer if and only if it is a local minimizer.

Proof: The direction $\Rightarrow$ is obvious. Hence, it remains to prove $\Leftarrow$:

- Suppose $\boldsymbol{x}^* \in \Omega$ is not a global minimizer. Hence, there is a $\boldsymbol{y} \in \Omega$, $\boldsymbol{y} \neq \boldsymbol{x}^*$, such that $f(\boldsymbol{y}) < f(\boldsymbol{x}^*)$.
- Draw a line between $f(\boldsymbol{y})$ and $f(\boldsymbol{x}^*)$. Every point on that line is $< f(\boldsymbol{x}^*)$.
- By convexity of f, the actual graph of f lies *below* the line above.
- By convexity of Ω, all points on the line segment joining $\boldsymbol{y}$ and $\boldsymbol{x}^*$ are in Ω.

- Moreover, all points on the line segment above (apart from the endpoint) have objective function value $< f(\boldsymbol{x}^*)$.
- Hence, $\boldsymbol{x}^*$ cannot be a local minimizer.

Lemma (21.1): Let $g : \Omega \to \mathbb{R}$ be a convex function defined on a convex set $\Omega \subset \mathbb{R}^n$. Then, for each $c \in \mathbb{R}$, the set

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

is a convex set.

Proof:

- Let $\boldsymbol{x}, \boldsymbol{y} \in \Gamma_c$; i.e., $g(\boldsymbol{x}), g(\boldsymbol{y}) \leq c$.
- Since g is convex, for all $\alpha \in (0, 1)$,

$$g(\alpha \boldsymbol{x} + (1 - \alpha)\boldsymbol{y}) \leq \alpha g(\boldsymbol{x}) + (1 - \alpha) g(\boldsymbol{y}) \leq c$$

- Hence, $\alpha \boldsymbol{x} + (1 - \alpha)\boldsymbol{y} \in \Gamma_c$, which implies that Γ_c is convex.

Corollary (21.1): In a convex programming problem, the set of all global minimizers is convex.

Proof:

- If a global minimizer does not exist, the result is trivial.
- If a global minimizer exists, the result follows immediately from the previous lemma by setting

$$c = \min_{\boldsymbol{x} \in \Omega} f(\boldsymbol{x}).$$

- We are now ready to prove that the FONC type conditions we have seen before are sufficient for global optimality.

Summary of FONCs

- Set constraint: $\boldsymbol{d}^T \nabla f(\boldsymbol{x}^*) \geq 0$ for all feasible directions $\boldsymbol{d}$;
- Interior: $\nabla f(\boldsymbol{x}^*) = \boldsymbol{0}$;
- $\Omega = \{\boldsymbol{x} : \boldsymbol{h}(\boldsymbol{x}) = \boldsymbol{0}\}$: Lagrange conditions;
- $\Omega = \{\boldsymbol{x} : \boldsymbol{h}(\boldsymbol{x}) = \boldsymbol{0}, \boldsymbol{g}(\boldsymbol{x}) \leq \boldsymbol{0}\}$: KKT conditions.

Set constraints

- Theorem (21.6): Consider the convex programming problem

$$\begin{array}{ll} \text{minimize} & f(\boldsymbol{x}) \\ \text{subject to} & \boldsymbol{x} \in \Omega, \end{array}$$

  where $f \in \mathcal{C}^1$ on an open covex set that contains Ω. Suppose the point $\boldsymbol{x}^* \in \Omega$ satisfies

$$\boldsymbol{d}^T \nabla f(\boldsymbol{x}^*) \geq 0$$

  for any feasible direction $\boldsymbol{d}$ at $\boldsymbol{x}^*$. Then, $\boldsymbol{x}^*$ is a global minimizer.

- Corollary (21.2): If the point $\boldsymbol{x}^*$ above satisfies $\nabla f(\boldsymbol{x}^*) = \mathbf{0}$, then $\boldsymbol{x}^*$ is a global minimizer.

Proof of Theorem:

- Consider any $\boldsymbol{x} \in \Omega$. We want to show that $f(\boldsymbol{x}) \geq f(\boldsymbol{x}^*)$.
- By convexity and previous theorem,

$$f(\boldsymbol{x}) \geq f(\boldsymbol{x}^*) + Df(\boldsymbol{x}^*)(\boldsymbol{x} - \boldsymbol{x}^*).$$

- Note that $\boldsymbol{d} = \boldsymbol{x} - \boldsymbol{x}^*$ is a feasible direction (because Ω is convex). Hence, by assumption,

$$Df(\boldsymbol{x}^*)(\boldsymbol{x} - \boldsymbol{x}^*) = \boldsymbol{d}^T \nabla f(\boldsymbol{x}^*) \geq 0.$$

- Combining the above two inequalities, we have

$$f(\boldsymbol{x}) \geq f(\boldsymbol{x}^*).$$

Equality constraints

- Let us now consider problems with equality constraint $\boldsymbol{h}(\boldsymbol{x}) = \mathbf{0}$.
- Assume that the constraint set $\Omega = \{\boldsymbol{x} : \boldsymbol{h}(\boldsymbol{x}) = \mathbf{0}\}$ is convex.
- Example: $\boldsymbol{h}(\boldsymbol{x}) = \boldsymbol{b} - \boldsymbol{A}\boldsymbol{x}$.
- Further assume that f is a convex function, so that the problem is a convex programming problem.

- Theorem (21.7): Consider the convex programming problem

$$\begin{array}{ll} \text{minimize} & f(\boldsymbol{x}) \\ \text{subject to} & \boldsymbol{h}(\boldsymbol{x}) = \mathbf{0}. \end{array}$$

  Suppose there exists a feasible point $\boldsymbol{x}^*$ and a vector $\boldsymbol{\lambda}^*$ such that

$$Df(\boldsymbol{x}^*) + \boldsymbol{\lambda}^{*T} D\boldsymbol{h}(\boldsymbol{x}^*) = \mathbf{0}^T.$$

  Then, $\boldsymbol{x}^*$ is a global minimizer.

Equality and inequality constraints

- Now consider problems with both equality and inequality constraints:

$$\boldsymbol{h}(\boldsymbol{x}) = \mathbf{0}, \qquad \boldsymbol{g}(\boldsymbol{x}) \leq \mathbf{0}.$$

- The constraint set is

$$\begin{aligned} \Omega &= \{\boldsymbol{x} : \boldsymbol{h}(\boldsymbol{x}) = \mathbf{0}, \boldsymbol{g}(\boldsymbol{x}) \leq \mathbf{0}\} \\ &= \{\boldsymbol{x} : \boldsymbol{h}(\boldsymbol{x}) = \mathbf{0}\} \bigcap \{\boldsymbol{x} : \boldsymbol{g}(\boldsymbol{x}) \leq \mathbf{0}\}. \end{aligned}$$

- Note that the intersection of convex sets is convex (exercise: prove).
- Hence, Ω is convex if both the above sets are convex.

- We have already seen an example where the set $\{\boldsymbol{x} : \boldsymbol{h}(\boldsymbol{x}) = \mathbf{0}\}$ is convex.
- When is $\{\boldsymbol{x} : \boldsymbol{g}(\boldsymbol{x}) \leq \mathbf{0}\}$ convex?
- Note that

$$\{\boldsymbol{x} : \boldsymbol{g}(\boldsymbol{x}) \leq \mathbf{0}\} = \bigcap_{i=1}^{p} \{\boldsymbol{x} : g_i(\boldsymbol{x}) \leq 0\}.$$

- Therefore, if each g_i is convex, then by Lemma 21.1 we conclude that each $\{\boldsymbol{x} : g_i(\boldsymbol{x}) \leq 0\}$ is convex, and hence $\{\boldsymbol{x} : \boldsymbol{g}(\boldsymbol{x}) \leq \mathbf{0}\}$ is convex.

- Theorem (21.8): Consider the convex programming problem

$$\begin{aligned} \text{minimize} \quad & f(\boldsymbol{x}) \\ \text{subject to} \quad & \boldsymbol{h}(\boldsymbol{x}) = \mathbf{0} \\ & \boldsymbol{g}(\boldsymbol{x}) \leq \mathbf{0}. \end{aligned}$$

  Suppose there exists a feasible point $\boldsymbol{x}^*$ and vectors $\boldsymbol{\lambda}^*$ and $\boldsymbol{\mu}^*$ such that

  1. $\boldsymbol{\mu}^* \geq \mathbf{0}$;
  2. $Df(\boldsymbol{x}^*) + \boldsymbol{\lambda}^{*T} D\boldsymbol{h}(\boldsymbol{x}^*) + \boldsymbol{\mu}^{*T} D\boldsymbol{g}(\boldsymbol{x}^*) = \mathbf{0}^T$; and
  3. $\boldsymbol{\mu}^{*T}\boldsymbol{g}(\boldsymbol{x}^*) = 0$.

  Then, $\boldsymbol{x}^*$ is a global minimizer.

