AMS 761 Nonlinear Optimization I	**FALL 2023**

Lecture 3: September 5

Lecturer: Mateo Díaz — *Scribe: Hans Harck Tønning*

3.1 Optimality conditions (Continued)

Theorem 3.1 (1st order sufficient condition) *Assume $f : \mathbb{R}^d \to \mathbb{R}$ is convex and differentiable. Then x^* is a global minimizer $\iff$ $\nabla f(\bar{x}^*) = 0$.*

Proof: "$\implies$": Follows from the 1st order necessary condition.

"$\impliedby$": Let $\bar{y} \in \mathbb{R}^d \setminus \{x^*\}$. Define $\psi(t) = f(\bar{x}^* + t(\bar{y} - \bar{x}^*))$. By the chain rule $\psi'(0) = \nabla f(\bar{x}^*)(\bar{y} - \bar{x}^*) = 0$.

For any $t \in (0, 1]$:

$$\frac{f(\bar{x}^* + t(\bar{y} - \bar{x}^*)) - f(\bar{x}^*)}{t} \le \frac{(1-t)f(\bar{x}^*) + tf(\bar{y}) - f(\bar{x}^*)}{t} = f(\bar{y}) - f(\bar{x}^*).$$

Thus, by taking the limit as t goes to zero, we get $0 = \psi'(0) \le f(\bar{y}) - f(\bar{x}^*)$, and we have that $f(\bar{x}^*) \le f(\bar{y})$.

■

Theorem 3.2 (2nd order necessary condition) *Suppose that $f : \mathbb{R}^d \to \mathbb{R}$ is twice differentiable. If $\bar{x}^*$ is a local minimizer, then $\nabla f(\bar{x}^*) = 0$ and $\nabla^2 f(\bar{x}^*) \succeq 0$.*

Note that $\nabla^2 f(\bar{x}^*) \ge 0$, means for all $\bar{s} \in \mathbb{R}^d \setminus \{0\}$ we have that $\bar{s}^\top \nabla^2 f(\bar{x}^*) \bar{s} \ge 0$.

Proof: Seeking contradiction, assume $\nabla f(\bar{x}^*) = 0$ and there exists a $bars \in \mathbb{R}^d \setminus \{0\}$ s.t. $\bar{s}^\top \nabla^2 f(\bar{x}^*) \bar{s} < 0$ and $||\bar{s}|| = 1$.

Define $\psi(t) = f(\bar{x}^* + t\bar{s})$. Then

$$0 > \frac{1}{2}\psi''(0) = \lim_{t \to 0} \frac{\psi(t) - \psi(0)}{t^2}.$$

For small enough $t > 0$, we have that $\frac{\psi(t)-\psi(0)}{t^2} \le \frac{1}{4}\psi''(0) < 0$. But this means that $\overbrace{f(\bar{x}^* + t\bar{s})}^{\phi(t)} < \overbrace{f(\bar{x}^*)}^{\phi(0)}$, which is a contradiction. ■

Theorem 3.3 (2nd order sufficient condition) *Suppose that $f : \mathbb{R}^d \to \mathbb{R}$ is twice differentiable. We have that $\bar{x}^*$ is a strict local minimizer if $\nabla f(\bar{x}^*) = 0$ and $\nabla^2 f(\bar{x}^*) \succ 0$.*

Proof:

Suppose $\bar{x}^*$ satisfies the assumptions ($\nabla f(\bar{x}^*) = 0$ and $\nabla^2 f(\bar{x}^*) \succ 0$). Take $\bar{U} \in \mathbb{R}^d$ s.t. $||\bar{|}|| = 1$. Let $\psi(t) = f(\bar{x}^* + t\bar{U})$.

By the Fundamental Theorem of Calculus (FTC), we have that:

$$\phi(s) = \phi(0) + \int_0^s \phi'(\alpha) d\alpha.$$

Applying FTC again to $\phi'(\alpha)$:

$$\phi(s) = \phi(0) + \int_0^s \phi'(\alpha) + \int_0^\alpha \phi''(\beta)\ d\beta d\alpha. \tag{3.1}$$

Since $\nabla^2 f(\bar{x})$ is continuous and $\lambda_{\min}(\nabla^2 f(\bar{x}^*)) > 0$. For all $\bar{y}$ close to $\bar{x}^*$, we have that $\lambda_{\min}(\nabla^2 f(\bar{y})) \geq \lambda > 0$ where λ is some positive constant.

From (3.1) we have that

$$\phi(s) = \phi(0) + \phi'(0)s + \int_0^s \int_0^\alpha \bar{U}^\top \nabla^2 f(\bar{x}^* + \beta\bar{U})\bar{U} d\beta\ d\alpha \geq \phi(0) + \lambda \int_0^s \int_0^\alpha 1 d\beta\ d\alpha = f(\bar{x}^*) + \frac{\lambda}{2}s^2.$$

■

Note that the above theorem is not an if and only if, as can be seen with the following counter example.

Example 3.4 *Let $f(x) = x^4$. Then $x = 0$ is clearly a global minimizer, but $f''(0) = \nabla^2 f(0) = 0$.*

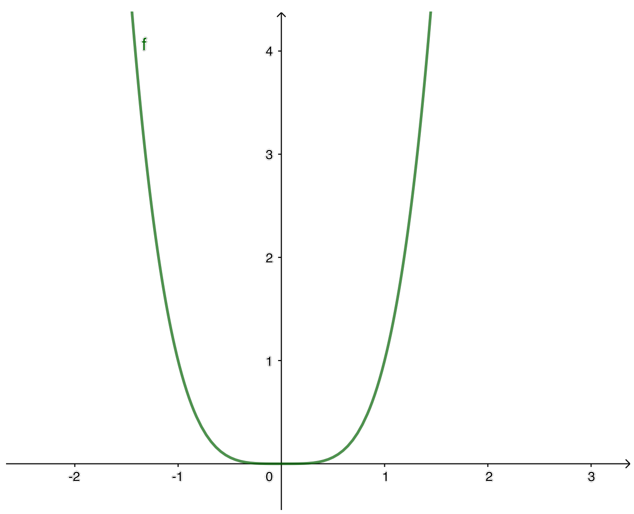

Figure 3.1: Plot of $f(x) = x^4$.

3.2 Basic Convexity

Definition 3.5 *A set $C \subseteq \mathbb{R}^d$ is convex if for all $\bar{y} \in C$ and $t \in [0, 1]$:*

$$t\bar{x} + (1 - t)\bar{y} \in C.$$

I.e. the straight line between any two points in C must be entirely within C.

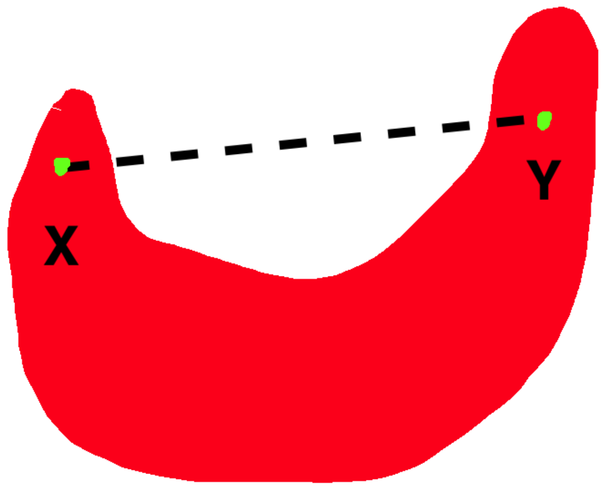

Figure 3.2: An example of a non-convex set in $\mathbb{R}^2$.

Definition 3.6 *Given any function $f : \mathbb{R}^d \to \mathbb{R}$ its epigraph is given by:*

$$\mathrm{epi} f = \{(\bar{x}, t) \mid f(\bar{x}) \leq t\}.$$

I.e. all points that are on or above the graph of $f(x)$.

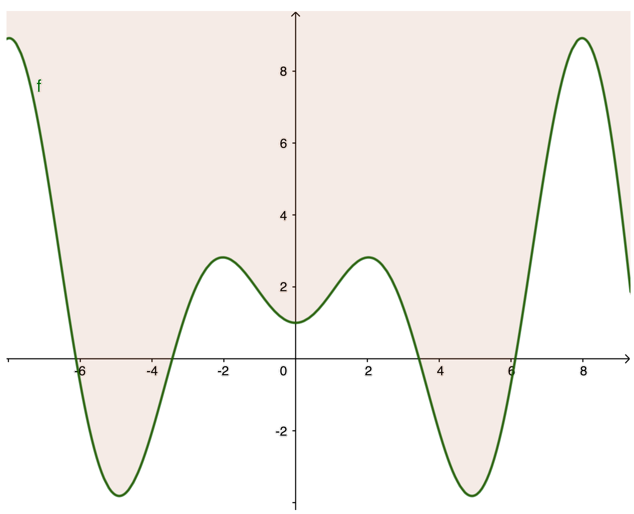

Figure 3.3: The epigraph of $f(x) = x\sin(x) + 1$.

Theorem 3.7 *A function f is convex iff the epigraph* epi$f$ *is convex.*

Proof: Homework exercise. Follows easily from the definitions. ■

Lemma 3.8 *Let $C_1, C_2 \subseteq \mathbb{R}^d$ be convex sets. Then $C_1 \cap C_2$ is also convex.*

Proof: Let $x, y \in C_1 \cap C_2$, let $t \in [0, 1]$. Since C_1 is convex $t\bar{x} + (1-t)\bar{y} \in C_1$, and since C_2 is convex $t\bar{x} + (1-t)\bar{y} \in C_2$. Thus $t\bar{x} + (1-t)\bar{y} \in C_1 \cap C_2$. ■