Conditions for Local Minimizers

Global minimizer x^* : $f(x) \geqslant f(x^*) \quad \forall\, x \in \Omega \setminus \{x^*\}$

Strict global minimizer x^* : $f(x) > f(x^*) \quad \forall\, x \in \Omega \setminus \{x^*\}$

Local minimizer : $f(x) \geqslant f(x^*)$ when x is "near" x^*.

Mathematically, "near" can be characterize as $\|x - x^*\| < \varepsilon$.
What is ε? A small number. What is the value of ε? Doesn't matter

x^* is a local minimizer if $\exists\, \varepsilon > 0$, s.t.

$$f(x) \geqslant f(x^*) \quad \forall\, x \in \Omega \setminus \{x^*\} \ \&\ \|x - x^*\| < \varepsilon.$$

$\exists$: exists $\quad \forall$: for all $\quad$ s.t. : such that $\quad \varepsilon$: epsilon

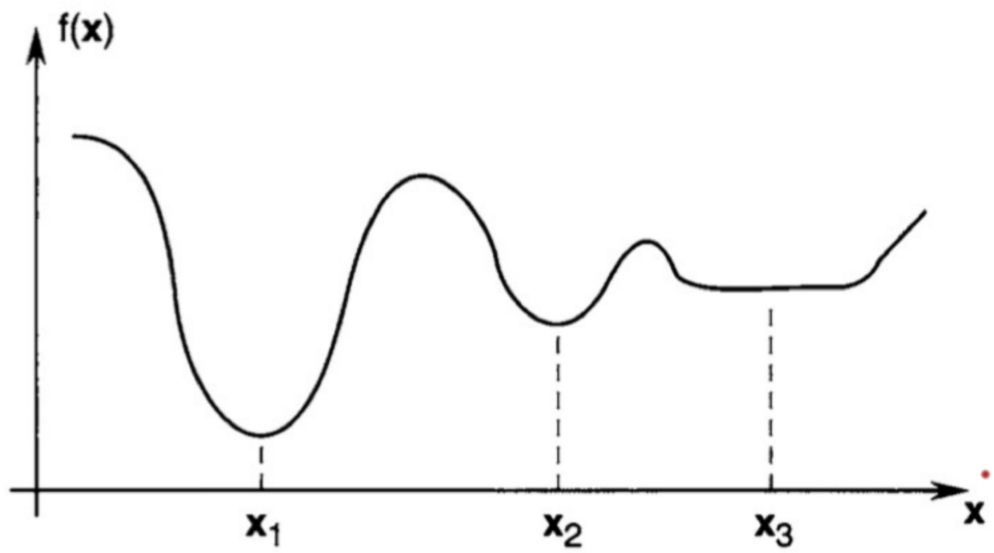


Figure 6.1 Examples of minimizers: x_1: strict global minimizer; x_2: strict local minimizer; x_3: local (not strict) minimizer.

First Order Necessary Condition

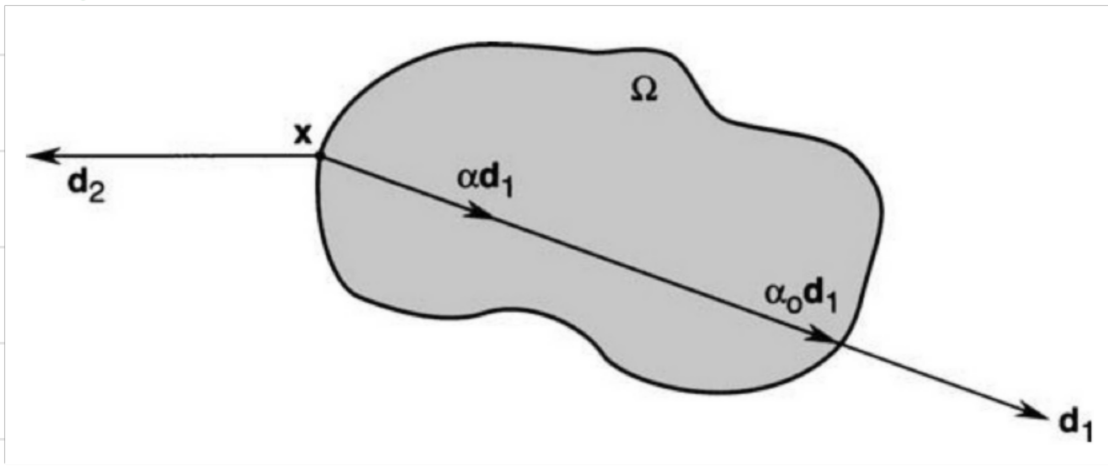
Theorem. If x^* is a local minimizer of f over Ω, then for any feasible direction d at x^*, we have

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

Explanation. ① feasible direction d at a point $x \in \Omega$ is a direction so that: starting from x and moving towards d remains in Ω.

Math language: $\exists\, \alpha_0 > 0$ s.t. $x + \alpha d \in \Omega$, $\forall\, \alpha \in [0, \alpha_0]$.

Figure:



② $\nabla f(x^*) \cdot d$: inner product of two vectors. Also write as

$d^T \nabla f(x^*)$ or $(\nabla f(x^*), d)$, $\langle \nabla f(x^*), d \rangle$

$\frac{\partial f}{\partial d} \triangleq \nabla f \cdot d$ is the directional derivative when $\|d\| = 1$.

③ Define $\phi(\alpha) = f(x^* + \alpha d)$ for $\alpha \in [0, \alpha_0]$

Then
$$\phi'(0) = \begin{cases} \lim\limits_{\alpha \to 0^+} \dfrac{\phi(\alpha) - \phi(0)}{\alpha} = \lim\limits_{\alpha \to 0^+} \dfrac{f(x^* + \alpha d) - f(x^*)}{\alpha} & \text{Def} \\ \nabla f(x^*) \cdot d & \text{Chain rule} \end{cases}$$

Proof. Let d be any feasible direction at x^*. Define $\phi(\alpha) = f(x^* + \alpha d)$

Then $f(x^* + \alpha d) - f(x^*) = \phi(\alpha) - \phi(0) = \phi'(0)\alpha + o(\alpha)$

$$= (\nabla f(x^*) \cdot d)\,\alpha + o(\alpha)$$

If x^* is a local minizer ($\exists\, \varepsilon$, s.t. $f(x) \geqslant f(x^*)$ $\forall\, x \in \Omega \setminus \{x^*\}$ & $\|x - x^*\| < \varepsilon$)

for sufficiently small α (e.g. $\|\alpha d\| < \varepsilon$), $f(x^* + \alpha d) - f(x^*) \geqslant 0$,

then $\phi'(0) = \nabla f(x^*) \cdot d \geqslant 0$. #.

FONC Two possibilities for a given feasible direction d.

$$\begin{cases} \nabla f(x^*) \cdot d > 0 & \text{then } f(x^* + \alpha d) > f(x^*) \text{ for all sufficiently small } \alpha > 0 \\ \nabla f(x^*) \cdot d = 0. & \text{Check second-order derivative} \end{cases}$$

Second-Order Necessary Condition

Theorem. If x^* is a local minimizer of f over Ω, and there exists a feasible direction d at x^* s.t. $\nabla f(x^*) \cdot d = 0$, then

$$d^T \nabla^2 f(x^*)\, d \geqslant 0.$$

Proof. Consider $\phi(\alpha) = f(x^* + \alpha d)$ and its Taylor series at $\alpha = 0$.

$$\phi(\alpha) = \phi(0) + \underbrace{\cancel{\phi'(0)\alpha}}_{0} + \frac{1}{2}\phi''(0)\alpha^2 + o(\alpha^2)$$

as $\phi'(0) = \nabla f(x^*) \cdot d = 0$

So we have $\phi(\alpha) - \phi(0) = \frac{1}{2}\phi''(0)\alpha^2 + o(\alpha^2)$. Written in terms of f is $f(x^* + \alpha d) - f(x^*) = \frac{\alpha^2}{2} d^T \nabla^2 f(x^*) d + o(\alpha^2)$.

If $d^T \nabla^2 f(x^*) d < 0$, then for sufficiently small α (how small?)

$$f(x^* + \alpha d) - f(x^*) = \frac{\alpha^2}{2} d^T \nabla^2 f(x^*) d + o(\alpha^2) < 0$$

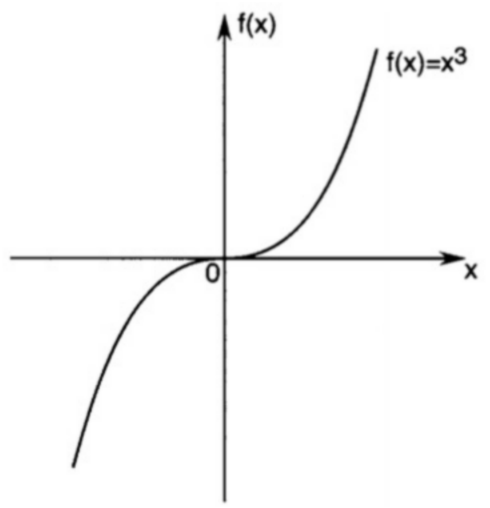
which contradicts that x^* is a local minimizer. #.

Corollary. x^* is an interior local minimizer of f, then

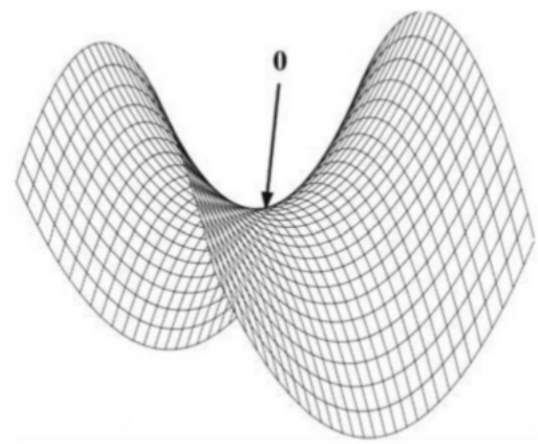
- **FONC.** $\nabla f(x^*) = 0$.
- **SONC.** $d^T \nabla^2 f(x^*) d \geq 0, \ \forall d \in \mathbb{R}^n$.

Examples. 6.3 (page 86), 6.5 (page 89).

Necessary conditions are not sufficient



$f(x) = x^3$, $f'(x) = 3x^2$, $f''(x) = 6x$



$f(\boldsymbol{x}) = x_1^2 - x_2^2$

$\mathbf{0}$ is a saddle point: $\nabla f(\mathbf{0}) = \mathbf{0}$ but neither a local minimizer nor maximizer

By SONC, $\mathbf{0}$ is not a local minimizer!

Second-Order Sufficient Condition

Th 6.3 (SOSC) $f \in C^2(\Omega)$, $x^* \in \Omega$ is an interior point.

Suppose that ① $\nabla f(x^*) = 0$; ② $\nabla^2 f(x^*) > 0$.

Then x^* is a strict local minimizer of f.

Pf. $\nabla^2 f(x^*) > 0 \Leftrightarrow \lambda_{min}(\nabla^2 f(x^*)) > 0$

(Prove by diagonalization of $\nabla^2 f(x^*) = Q^T \Lambda Q$)

For a feasible direction $d \neq 0$, define $\phi(\alpha) = f(x^* + \alpha d)$.

Then $\phi(\alpha) = \phi(0) + \phi'(0)\alpha + \frac{1}{2}\phi''(0)\alpha^2 + o(\alpha^2)$.

$$
\begin{aligned}
f(x^* + \alpha d) - f(x^*) = \phi(\alpha) - \phi(0) &= \frac{1}{2}\phi''(0)\alpha^2 + o(\alpha^2) \\
&= \frac{1}{2} d^T \nabla^2 f(x^*) d \ \alpha^2 + o(\alpha^2) \\
&\geq \frac{1}{2} \lambda_{min} \|d\|^2 \alpha^2 + o(\alpha^2) > 0
\end{aligned}
$$

if α is sufficiently small. #

$\boxed{x^* \text{ is a local minimizer of } f}$

Necessary Sufficient

interior pt $\nabla f(x^*) = 0,\ \nabla^2 f(x^*) \geq 0$ $\nabla f(x^*) = 0,\ \nabla f(x^*) > 0$

general pt $\nabla f(x^*) \cdot d \geq 0 \quad \forall d$

If $\nabla f(x^*) \cdot d = 0$ for some d, $d^T \nabla^2 f(x^*) d \geq 0$