$$(P)\quad \min_{x \in \mathcal{X}} f(x)$$

* A Taxonomy of Solutions to (P)

- Terminology: I will not distinguish b/w "solution" and "minimizer."

* **Def.** We say that $x^* \in \mathrm{dom}(f)$ is:

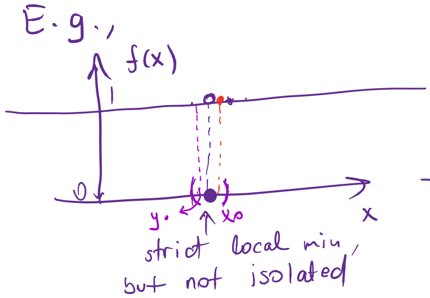
1) a local minimizer of (P) (or a local solution to (P)) if $\exists$ a neighborhood $\mathcal{N}_{x^*}$ of x^* s.t. $\forall x \in \mathcal{N}_{x^*} \cap \mathcal{X}$ we have $f(x) \geq f(x^*)$.

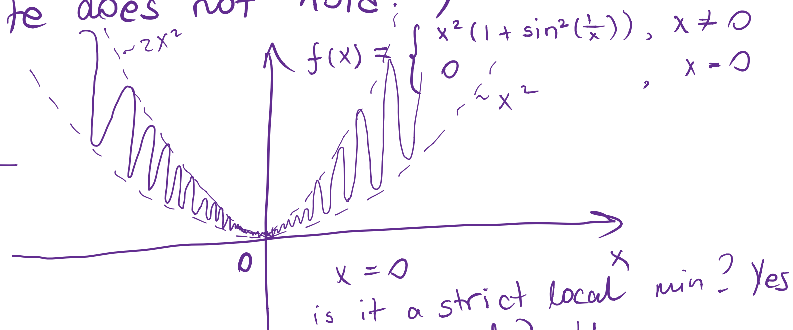
2) a strict local minimizer of (P) if it satisfies 1) but the inequality is strict: $f(x) > f(x^*)$.
3) a global minimizer of (P) if $\forall x \in \mathcal{X}: f(x) \geq f(x^*)$.
4) an isolated local minimizer of (P), if $\exists$ a neighborhood $\mathcal{N}_{x^*}$ s.t. $\forall x \in \mathcal{N}_{x^*} \cap \mathcal{X}: f(x) \geq f(x^*)$ and $\mathcal{N}_{x^*}$ does not contain any other local minimizers.

* Ex. Every isolated local minimizer is strict.

(But the opposite does not hold.)

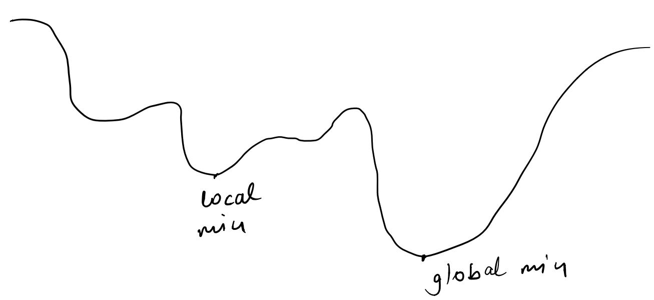
E.g.,





$x = 0$ is it a strict local min? Yes

Is it isolated? No.



* Taylor's Theorem:

- For this part and until explicitly stated otherwise, we will be assuming that f is at least once cont.'ly diff.'able
- Taylor's Theorem for 1D functions from calculus:

  $f: \mathbb{R} \to \mathbb{R}$, f k-times cont.'ly diff.'able

  $$\Rightarrow \forall x, y \in \mathbb{R}: f(y) = f(x) + \frac{1}{1!} f'(x)(y-x) + \frac{1}{2!} f''(x)(y-x)^2 + \dots + \frac{1}{k!} f^{(k)}(x)(y-x)^k + \underbrace{R_k(x)}_{\text{remainder}}$$

  Typical forms of $R_k(x)$:
  (assume that f is $(k+1)$-times cont.'ly diff.'able)

  1) Lagrange (mean-value) remainder:
  $$\exists \mu \in (0,1): R_k(x) = \frac{1}{(k+1)!} f^{(k+1)}(x + \mu(y-x))(y-x)^{k+1}$$

  2) Integral remainder:
  $$R_k(x) = \frac{1}{k!} \int_0^1 f^{(k+1)}(x + t(y-x))(y-x)^{k+1}\, dt$$

* **Theorem (2.1 in WR)** Let $f: \mathbb{R}^d \to \bar{\mathbb{R}}$ be a cont.'ly diff.'able function. Then, $\forall x, y \in \mathrm{dom}(f)$ and s.t. $\forall \alpha \in (0,1)$ $(1-\alpha)x + \alpha y \in \mathrm{dom}(f)$:

1) $f(y) = f(x) + \int_0^1 \langle \nabla f(x + t(y-x)), y - x \rangle\, dt$
2) $\exists \mu \in (0,1): f(y) = f(x) + \langle \nabla f(x + \mu(y-x)), y-x \rangle$.
   (Mean Value Thm)

If f is twice cont.'ly diff.'able:

3) $\nabla f(y) = \nabla f(x) + \int_0^1 \underbrace{\nabla^2 f(x + t(y-x))}_{\text{Hessian matrix}}(y-x)\, dt$

"second-order derivative of f"

$$\nabla^2 f(x) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2}(x) & \frac{\partial^2 f}{\partial x_1 \partial x_2}(x) & \dots & \frac{\partial^2 f}{\partial x_1 \partial x_d}(x) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_d \partial x_1}(x) & \frac{\partial^2 f}{\partial x_d \partial x_2}(x) & \dots & \frac{\partial^2 f}{\partial x_d^2}(x) \end{bmatrix}$$

4) $\exists \gamma \in (0,1)$:

$$f(y) = f(x) + \langle \nabla f(x), y-x\rangle + \frac{1}{2}\langle \underbrace{\nabla^2 f(x+\gamma(y-x))(y-x), y-x}_{(y-x)^T \nabla^2 f(x+\gamma(y-x))(y-x)}\rangle$$

Q. Can you have:

$\cancel{\exists \gamma \in (0,1): \nabla f(y) = \nabla f(x) + \nabla^2 f(x+\gamma(y-x))(y-x)}$? NO

$\left(f(x) = \sum_i f_i(x_i) \right)$

* Properties of smooth functions:

* Terminology: f is L-smooth w.r.t. $\|\cdot\|$ $\iff \forall x, y \in \mathrm{dom}(f): \|\nabla f(x) - \nabla f(y)\|_* \le L\|x-y\|$.

* Lemma 2.2. Let $f: \mathbb{R}^d \to \bar{\mathbb{R}}$ be an L-smooth function w.r.t. $\|\cdot\|$. Then, $\forall x, y \in \mathrm{dom}(f)$:

$$f(y) \boxed{\le} f(x) + \langle \nabla f(x), y-x\rangle + \frac{L}{2}\|y-x\|^2.$$

Proof: From Part 1) of TT:

$$\begin{aligned}
f(y) - f(x) - \langle \nabla f(x), y-x\rangle &= \int_0^1 \langle \nabla f(x+t(y-x)), y-x\rangle\, dt \\
&\quad - \int_0^1 \langle \nabla f(x), y-x\rangle\, dt \\
&= \int_0^1 \langle \nabla f(x+t(y-x)) - \nabla f(x), y-x\rangle\, dt \\
&\le \int_0^1 \|\nabla f(x+t(y-x)) - \nabla f(x)\|_* \|y-x\|\, dt \\
&\le \int_0^1 L\, t\, \|y-x\|^2\, dt \\
&= \frac{L}{2}\|y-x\|^2.
\end{aligned}$$

∎

* Ex. Prove that, under the same assumptions as in Lemma 2.2,

$$f(y) \ge f(x) + \langle \nabla f(x), y-x\rangle - \frac{L}{2}\|y-x\|^2.$$

* <u>Lemma 2.3</u>. Suppose that $f: \mathbb{R}^d \to \bar{\mathbb{R}}$ is twice cont.'ly diff.'able on dom(f). Then f is L-smooth w.r.t. $\|\cdot\|_2$ if and only if $-LI \boxed{\preccurlyeq} \nabla^2 f(x) \preccurlyeq LI$, $\forall x \in \text{dom}(f)$.

Löwner order

$A \succcurlyeq B$

$\Leftrightarrow A - B \succcurlyeq 0$

$\Leftrightarrow A - B$ is PSD.