Optimization

- State the problem.

$$f(\vec{x}) : \mathbb{R}$$

$$\uparrow$$

$$\mathbb{R}^n$$

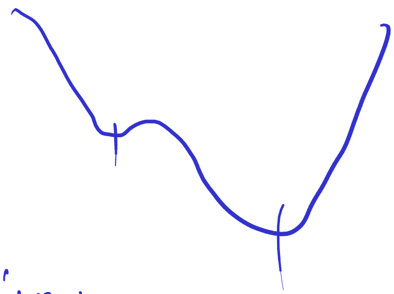
set S of feasible points

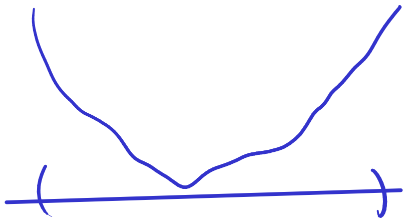
$$g(\vec{x}) = \vec{0}$$

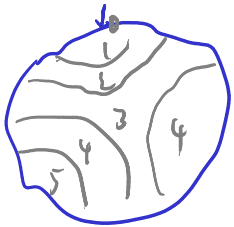
$$\vec{h}(\vec{x}) \leq \vec{0}$$

Existence/Uniqueness

- Under what conditions on f can we say something about existence/uniqueness?





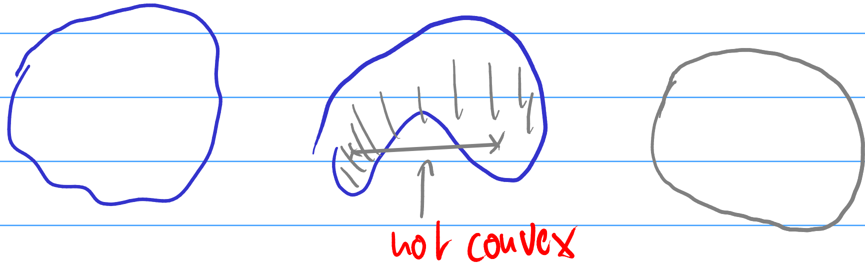


continuous function on closed domain; has a minimum

If f coercive:

$$\lim_{\|x\| \to \infty} f(x) = +\infty$$

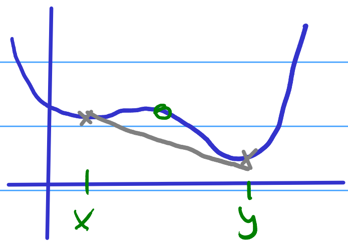
then f has a global minimum



set S is convex if for any two pts $x, y \in S$:

$$\text{for all } \alpha \in [0,1] \quad \alpha x + (1-\alpha) y \in S.$$

function f is convex on the convex set S if

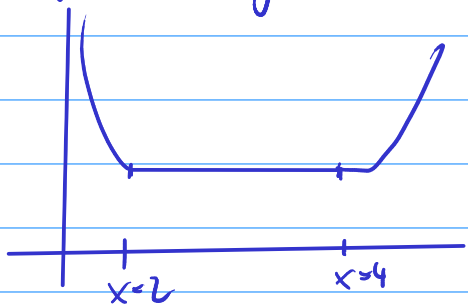


for any two pts $x, y \in S$:

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

If f convex, then local $\Rightarrow$ global

f strictly convex if $f(\alpha x + (1-\alpha)y) < \alpha f(x) + (1-\alpha) f(y)$



f strictly convex:

minimum is unique.

Optimality Conditions

- If we have found a candidate $\boldsymbol{x}^*$ for a minimum, how do we know it actually is one?
  To make this doable, assume f is smooth–i.e. has as many derivatives as needed.

In 1D $\quad f: \mathbb{R} \to \mathbb{R}$

Necessary: $\quad f'(x^*) = 0$

Sufficient: $\quad f'(x^*) = 0$ and $f''(x^*) > 0$

In nD: $\quad f: \mathbb{R}^n \to \mathbb{R}$

Necessary $\quad \nabla f(x^*) = 0$

Sufficient: $\quad \nabla f(x^*) = 0$ and $H_f(x)$ p.d.

$$H_f = \begin{pmatrix} \frac{\partial^2}{\partial x_1 \partial x_1} f & \frac{\partial^2}{\partial x_1 \partial x_n} f \\ \frac{\partial^2}{\partial x_n \partial x_1} f & \frac{\partial^2}{\partial x_n \partial x_n} f \end{pmatrix}$$

Symmetric

"Schwarz's theorem"

$\|\vec{s}\| = 1$

$$f(\vec{x} + h\vec{s}) = f(\vec{x}) + h \underbrace{\nabla f(\vec{x}) \cdot \vec{s}}_{\text{at min } = 0} + \frac{h^2}{2} \vec{s}^T H_f(\vec{x}) \vec{s} + O(h^3)$$

$A = LL^T \leftarrow$ p.d.'ness testable by Cholesky

Sensitivity and Conditioning

- How does optimization react to a slight perturbation of the minimum?

1D:

$$|f(\tilde{x}) - f(x^*)| < tol$$

min.

$$f(\underbrace{x^* + h}_{\hat{x}}) = f(x^*) + \underbrace{h f'(x^*)}_{=0} + \frac{h^2}{2} f''(x^*) + O(h^3)$$

$$tol > \frac{h^2}{2} f''(x^*) \quad \rightsquigarrow \quad h \leq \sqrt{\frac{2\,tol}{f''(x^*)}}$$

nD:

$$f(\vec{x} + h\vec{s}) = f(\vec{x}) + h\nabla(\vec{x}) \cdot \vec{s} + \frac{h^2}{2} \vec{s}^T H_f(\vec{x}) \vec{s} + O(h^3)$$

$$|h|^2 \leq \frac{2\,tol}{\lambda_{min}(H_f(x^*))}$$

6.1 Methods for unconstrained opt. in one dimension

$\vec{s}^T H_f(\vec{x}) \vec{s}$

$s^T H s$

$$f\begin{pmatrix} x_{01} \\ x_{02} \\ x_{11} \\ x_{12} \\ \\ \end{pmatrix} = \{ \quad \cdots \quad + \{$$

$\dfrac{\partial f}{\partial x_{01}}$

$f(x, y, z) = x^2 + 3xy + 5z^2$

$H_{21} = \dfrac{\partial f}{\partial y \, \partial x} = \dfrac{\partial}{\partial y}\left(2x + 3y\right)$

$= 3$

Golden Section Search

- Would like a method like bisection, but for optimization.
  In general: No invariant that can be preserved.
  Need *extra assumption.*

Demo: Golden Section Search Proportions

Newton's Method

- Reuse the Taylor approximation idea, but for optimization.