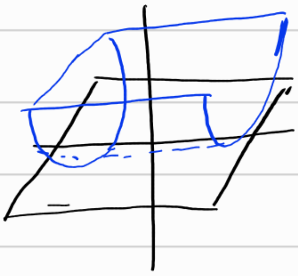
Global Optimality Conditions

- Let $f:\mathbb{R}^n \to \mathbb{R}$ be twice continuous differentiable

Def: f is called <u>convex</u> if $\nabla^2 f(\vec{x}) \succeq 0$ for all $\vec{x} \in \mathbb{R}^n$



f is called <u>strictly convex</u> if $\nabla^2 f(\vec{x}) \succ 0$ for all $\vec{x} \in \mathbb{R}^n$

" " " <u>concave</u> if $\nabla^2 f(\vec{x}) \preceq$ " " ".

" " " <u>strictly concave</u> if $\nabla^2 f(x) \prec 0$ " " "

<u>theorem:</u> Suppose f is (strictly) convex/concave then every critical point $\vec{x}^*$ of f is a strict global minimizer/maximizer

<u>Lemma:</u> (Taylor's theorem w/ Lagrange form of the remainder)

For any vectors $\vec{x}, \vec{y} \in \mathbb{R}^n$
there exists $\vec{z} \in \mathbb{R}^n$ (depending on $\vec{x}, \vec{y}$) such that

$$f(\vec{x}) = f(\vec{y}) + \nabla f(\vec{y})^T(\vec{x}-\vec{y}) + \tfrac{1}{2}(\vec{x}-\vec{y})^T \nabla^2 f(\vec{z})(\vec{x}-\vec{y})$$

(case f is convex)

<u>Proof:</u>
- Let $\vec{x}^*$ be a critical point of f $(\nabla f(\vec{x}^*) = 0)$
- Let $\vec{x} \in \mathbb{R}^n$. By the lemma, $\exists$ $\vec{z} \in \mathbb{R}^n$ (depends on $\vec{x}$)

"there exists"

such that $f(\vec{x}) - f(\vec{x}^*) = \tfrac{1}{2}(\vec{x}-\vec{x}^*)^T \underbrace{\nabla^2 f(\vec{z})}_{\text{always PSD b/c } f \text{ is convex}} (\vec{x}-\vec{x}^*) \geq 0$

$\Rightarrow f(\vec{x}) \geq f(\vec{x}^*)$ for all $\vec{x} \in \mathbb{R}^n \Rightarrow \vec{x}^*$ is a global minimizer

Ex: $f(x,y) = \frac{1}{2}x^2$

$$\nabla f(x,y) = \begin{bmatrix} x \\ 0 \end{bmatrix}$$

$$\nabla^2 f(x,y) = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \succeq 0 \quad \text{for all } (x,y)$$

$\Rightarrow$ f is convex

critical points $(0,y)$ for all $y \in \mathbb{R}$ are all minimizers

Important corollary:

- If a convex function has only one critical point, it must be the unique (strict) global minimizer

Ex: $f(x,y) = x^2 + y^4$

$$\nabla f(x,y) = \begin{bmatrix} 2x \\ 4y^3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Leftrightarrow (x,y) = (0,0)$$

$$\nabla^2 f(x,y) = \begin{bmatrix} 2 & 0 \\ 0 & 12y^2 \end{bmatrix} \succeq 0 \quad \text{for any } (x,y) \Rightarrow f \text{ is convex}$$

Since f is convex and $(0,0)$ is the only critical point $\Rightarrow (x,y) = (0,0)$ is the unique global minimizer

Ex: $f(x,y) = (x^2+y^2)^2 - 4x - 4y = x^4 + 2x^2y^2 + y^4 - 4x - 4y$

$$\nabla f(x,y) = \begin{bmatrix} 4x^3 + 4xy^2 - 4 \\ 4x^2y + 4y^4 - 4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\Rightarrow \begin{cases} x(x^2+y^2) = 1 \\ y(x^2+y^2) = 1 \end{cases}$$

$\Rightarrow$ subtract equations $\quad (x-y)(x^2+y^2) = 0$

either $x = y$

or $\cancel{x^2+y^2 = 0 \iff (x,y) = (0,0)}$

does not satisfy

if $x = y$

$$x(x^2+x^2) = 1$$
$$2x^3 = 1$$
$$x = \frac{1}{\sqrt[3]{2}} = y$$

$\Rightarrow (x,y) = \left(\frac{1}{\sqrt[3]{2}}, \frac{1}{\sqrt[3]{2}}\right)$ is the only critical point

$$\nabla^2 f(x,y) = \begin{bmatrix} 12x^2 + 4y^2 & 8xy \\ 8xy & 8xy + 16y^3 \end{bmatrix}$$

$$= \begin{bmatrix} 4(x^2+y^2) + 8x^2 & 8xy \\ 8xy & 4(x^2+y^2) + 8y^2 \end{bmatrix}$$

$$= \underbrace{\begin{bmatrix} 4(x^2+y^2) & 0 \\ 0 & 4(x^2+y^2) \end{bmatrix}}_{PSD} + 8\begin{bmatrix} x^2 & xy \\ xy & y^2 \end{bmatrix}$$

$$\begin{bmatrix} x^2 & xy \\ xy & y^2 \end{bmatrix} = \underbrace{8\begin{bmatrix} x \\ y \end{bmatrix}\begin{bmatrix} x & y \end{bmatrix}}_{PSD}$$

Fact 1 If $A = BB^T$ or $A = B^T B$ for some matrix B, then A is positive semi-definite.

(Why? $x^T B B^T x = (B^T x)^T B^T x = \|B^T x\|^2 \geq 0$

Fact 2 If A and B are positive semi-definite, then so is $A+B$

(why? $x^T(A+B)x = \underbrace{x^T A x}_{\geq 0} + \underbrace{x^T B x}_{\geq 0} \geq 0$

$\to$ $\nabla^2 f(x,y)$ is a sum of PSD matrices, and so PSD itself $\Rightarrow f$ is convex

$\therefore (x,y) = \left(\frac{1}{\sqrt[3]{2}}, \frac{1}{\sqrt[3]{2}}\right)$ is the unique global minimizer

Optima over a set

- Suppose $f: U \to \mathbb{R}$ is defined over a set $U \subseteq \mathbb{R}^n$

Ex: $f(x,y) = x+y$ st $x^2+y^2 \leq 1$

$g(x,y) = x^2+y^2$ st $x+y \geq 1$

• when does f have a global minimizer or maximizer over a set U

Weierstrauss theorem:

If $f: U \to \mathbb{R}$ is a continuous function defined over a compact set $U \subseteq \mathbb{R}^n$, then there exists a global minimizer and global maximizer of f be going to U

Ex: $f(x,y) = x+y$ st $x^2+y^2 \leq 1$

global max: $(x,y) = (1/\sqrt{2}, 1/\sqrt{2})$

" min: $(x,y) = (-1/\sqrt{2}, -1/\sqrt{2})$

• what if U is closed, but not compact?

Def: A function $f: U \to \mathbb{R}$ defined on an unbounde set $U \subseteq \mathbb{R}^n$ is called cohersive if

$\lim_{\|x\| \to \infty} f(\vec{x}) = \infty$ limit over all possible unbounded sequences of vectors

Ex: $f(x,y) = x^2+y^2$ is cohersive (bc $f(x,y) = \|[\begin{smallmatrix} x \\ y \end{smallmatrix}]\|^2$)

but $f(x,y) = x^2$ is not cohersive (bc $f(0,y) = 0$ for all y)

Theorem: Suppose $U \subseteq \mathbb{R}^n$ is a closed and unbounded set. If $f: U \to \mathbb{R}$ is continuous and cohersive then f has a global minimizer in U

Ex $f(x,y) = x^2+y^2$ st $x+y \geq 1$

global minimizer $(x,y) = (1/2, 1/2)$

Quadratic functions

$b \in \mathbb{R}^n$

$f(\vec{x}) = x^t A x + 2\vec{b}^t \vec{x} + c$ $\rightsquigarrow c \in \mathbb{R}$

symetric $n \times n$

Note: $\nabla f(\vec{x}) = 2Ax + 2b$

$\nabla^2 f(\vec{x}) = 2A$

Lemma: Let $f(x)$ be a quadratic function

a) $\vec{x}$ is a critical point iff $A\vec{x} = -\vec{b}$

(b) if $A \succeq 0$, then f is convex, and so $\vec{x}$ is a global minimizer of f iff $A\vec{x} = -\vec{b}$

(c) if $A \succ 0$, then $\vec{x} = -A^{-1}\vec{b}$ is the uniqu global minimizer

this implies A is invertible