Lecture 8

- Some convex analysis to understand $A = A^T$ $\to A \succ 0$ $x^T A x > 0$

$f(\underline{x}) = \frac{1}{2} \underline{x}^T A \underline{x} - \underline{x}^T \underline{b}$

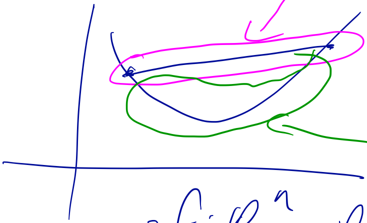
- "New method" for $A\underline{x} = \underline{b}$

↳ steepest descent

↳ gradient descent.

- Coordinate Descent.

$f(\underline{x})$



$\to f : \mathbb{R}^n \to \mathbb{R}$

$f(x)$ is convex if

$f(\alpha \underline{x} + (1-\alpha) y) \le$

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

$0 \le \alpha \le 1$

Last time

$$\frac{\partial f}{\partial \underline{x}} = A\underline{x} - \underline{b}$$

~~Proof:~~

$$\cancel{f(x_1, x_2, \ldots, x_n)}$$

Let's start 2x2

$$f(x,y) = \frac{1}{2}\begin{bmatrix} x \\ y \end{bmatrix}^T \begin{bmatrix} 3 & -1 \\ -1 & 4 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} - \begin{bmatrix} 7 \\ -6 \end{bmatrix}^T \begin{bmatrix} x \\ y \end{bmatrix}$$

$$= \frac{1}{2}\left(3x^2 - 2xy + 4y^2\right) - 7x + 6y$$

$$\frac{\partial f}{\partial x} = 3x -$$

$$\frac{\partial f}{\partial y} = 4y -$$

$$\begin{bmatrix} \partial f / \partial x \\ \partial f / \partial y \end{bmatrix} = \begin{bmatrix} 3 & \\ -1 & \end{bmatrix}$$

A

$y \quad -7$

$x \quad +6$

$$\begin{bmatrix} 1 \\ 4 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} - \begin{bmatrix} 7 \\ -6 \end{bmatrix}$$

$$\underline{x} \qquad -\underline{b}$$

$$f(x_1, \ldots, x_n) = \frac{1}{2}\underline{x}^T A \underline{x} - \underline{x}^T \underline{b}$$

$$= \frac{1}{2}\sum_{ij} A_{ij} x_i x_j - \sum_i x_i b_i$$

T

$$\frac{1}{2}\left[\begin{array}{c|c|c|c} A_{11}x_1^2 & A_{12}x_1x_2 & - & A_{1n}x_1x_n \\ \hline A_{21}x_2x_1 & A_{22}x_2^2 & - & \\ & & & \end{array}\right]$$

$$\partial f / \partial x_i = \frac{1}{2}\left(\sum_{j \neq i} A_{ji} x_j + \sum_{j \neq i} A_{ij}\right.$$

nce - David F. Gleich - 2019 CS 515 - Purdue University - Computer Science - David F. Gleich - 2019 CS 515 - Purdue University - Computer Science - David F. Gleich - 2019

by sym, cancel 2's

$$\frac{\partial f}{\partial x_i} = \underbrace{\sum_{j \neq i} (A_{ij} x_j) + A_{ii} x_i}_{} - b_i$$

$$= \sum_j A_{ij} x_j - b_i$$

$$= [Ax - b] \text{ ith row.}$$

$$x_j + 2 A_{ii} x_i) - b_i$$

A very brief

- A $\boxed{\text{local}}$

Is any poin

$f(x^*)$

st $\| x$

Some ε

Example

intro to optimization

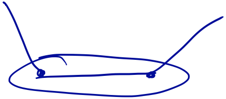
$\boxed{\text{minimizer}}$ of $f(x): \mathbb{R}^n \to \mathbb{R}$

A x^* where

$\le f(x)$ for all x

$x^* - x\|_2 \le \varepsilon$ for

> 0.



A $\boxed{\text{global minimizer}}$ of $f(x)$ is any point x^* s.t.

$f(x^*) \le f(x)$ for all x.

Convexity: for a convex function f, then local mins are global

---

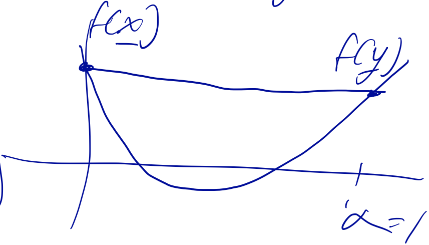
Quiz: Draw a function f s.t. f has a local min but no global min.

Pf (picture)

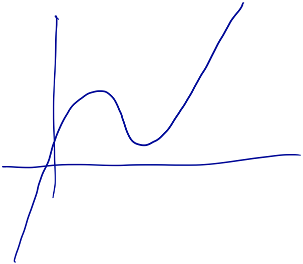
Let $\underline{x}, y$

local min

global min



mins.



Next Result

$$\frac{\partial f}{\partial \underline{x}} = 0 \iff \text{global min}$$

for

$$f(\underline{x}) = \frac{1}{2}\underline{x}^T A \underline{x} - x^T \underline{b}$$

We have

$$\underbrace{A = A^T}_{\text{sym.}}, \quad \underbrace{x^T A x > 0}_{\text{pos. def.}}$$

Let y be a pt where $Ay = \underline{b}$

$\Longleftrightarrow$

$$\frac{\partial f}{\partial x}(y) = Ay - \underline{b} = 0$$

We will show

$$f(\underline{x}) \geq f(y) \quad \forall \underline{x}$$

Let $\underline{x} = y + \alpha \underline{z}$

$$f(x) = f(y + \alpha \underline{z})$$

$$= \frac{1}{2}(y + \alpha \underline{z})^T A (y + \alpha \underline{z}) - (y + \alpha z)^T \underline{b}$$

$$= \frac{1}{2} y^T A y + \frac{1}{2}\alpha^2 \underline{z}^T A \underline{z} + \alpha \underline{z}^T A y - y^T \underline{b} - \alpha \underline{z}^T \underline{b}$$

$$= f(y) + \frac{1}{2}\alpha^2 \underline{z}^T A \underline{z} + \alpha z^T \underline{b}$$

$=$

$$\geq f(y) \quad ∎$$

$-\alpha 2^T \underline{b}$
0

Next goal:

Show gradient descent works!

$$x_{k+1} = x_k - \alpha \frac{\partial f}{\partial \underline{x}}$$

$$= x_k - \alpha(A\underline{x} - \underline{b})$$

$$= x_k + \alpha(\underline{b} - A\underline{x})$$

First, if $\partial f / \partial \underline{x} \neq 0$, then there must be an α s.t. $f(x_{k+1}) < f(x_k)$

Let $\underline{g} = A\underline{x}_k - \underline{b}$

$\underline{g} \neq 0$

$$f(x_{k+1}) = f(x_k - \alpha \underline{g})$$

$$= \frac{1}{2}(x_k - \alpha g)^T A (x_k - \alpha g) - (x_k - \alpha g)^T \underline{b}$$

$$= \boxed{\frac{1}{2} \underline{x}_k^T A x_k} + \frac{1}{2}\alpha^2 g^T A g - \alpha g^T A \underline{x}_k$$

$$= f(x_k) \quad \boxed{- x_k^T \underline{b}} + \alpha g^T \underline{b}$$

$$f(x_{k+1}) = f(x_k) + \frac{1}{2}\alpha^2 g^T A g - \alpha g^T A \underline{x}_k + \alpha g^T \underline{b}$$

$$= f(\underline{x}_k) - \alpha(g^T \boxed{A x_k}$$

$$-g^T($$

$$= f(x_k) - \alpha(g^T g -$$

$- \frac{\alpha}{2} g^T A g$

$b))$

$\frac{\alpha}{2} g^T A g)$

Goal: pick $\alpha > 0$

s.t. $f(x_{k+1}) < f(x_k) \Rightarrow$

$\alpha \geq 0$ s.t.

$$g^T g - \frac{\alpha}{2} g^T A g > 0$$

Use any α s.t.

$$g^T g \geq \frac{\alpha}{2} g^T A g \Rightarrow$$

$$2 \frac{g^T g}{g^T A g} \geq \alpha$$

$\Rightarrow$

$$\frac{g^T A g}{g^T g} \leq$$

Let $\max_{g \neq 0}$ -

Then $\rho \leq \frac{2}{\alpha}$

$$\alpha \leq \frac{2}{\rho}$$

$$\frac{\partial}{\alpha}$$

$$\frac{g^TAg}{g^Tg} = \rho \,.$$

$\alpha\sim$

$=$

$$\underset{\alpha}{\text{minimize}} \; f(x_k - \alpha\underbrace{(Ax_k - b)}_{g})$$

Quiz Not today!

$$f(x - \alpha g) = \cancel{\tfrac{1}{2} x^T A x} + \tfrac{1}{2}\alpha^2 g^T$$

$$- \alpha g^T A x -$$

$$f(\alpha) = \tfrac{1}{2}\alpha^2 \overset{g^TAg}{\lambda_2} - \alpha(\overset{(g^TAx}{\lambda_1})$$

$$\Longleftrightarrow$$

$$\min_{\alpha} \; \tfrac{1}{2}\alpha^2 \lambda_2 - \alpha \lambda_1$$

$^T A g$

$\cancel{x^T b} + \alpha g^T b$

$- g^T b)$

$\alpha = \frac{\lambda_1}{\lambda_2}$

$= \frac{g^T g}{g^T A g}$

Coordinate Des

$x_{k+1} = x_k$

As algebra

$x_{k+1} = x_k +$

Goal

pick γ

is as small

ent

$\longrightarrow$

$$\underbrace{+ \text{"change in } i\text{th coord."}}$$

$$\gamma \underline{e}_i$$

s.t. $f(x_{k+1}) = f(x_k + \gamma e_i)$

as possible!

$$\min_{\gamma} f(x_k + \gamma \underline{e}_i)$$

$\Rightarrow$

$$\min_{\gamma} \frac{1}{2} x_k^T A x_k + \frac{1}{2} \gamma$$

$$+ \gamma e_i^T A$$

$$- \underline{x}_k^T \underline{b} \; - \gamma e_i$$

$\Leftrightarrow$

$$\min_{\gamma} \frac{1}{2} \gamma^2 \underline{e}_i^T \underline{A} e_i$$

$\Leftrightarrow$

$$A_{ii} \gamma + g_i = 0$$

$^2 e_i^T A e_i$

$\underline{x}_k$

$^T \underline{b}$

$+ \gamma [e_i^T][A\underline{x} - \underline{b}]$ with $\underbrace{A\underline{x} - \underline{b}}_{g}$

$\Rightarrow \gamma = -\frac{g_i}{A_{ii}}$