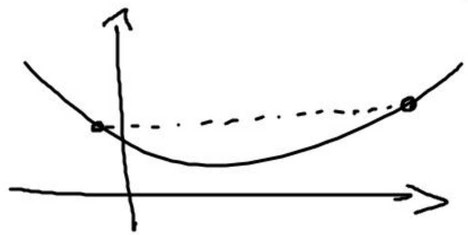
LECTURE 11: CONVEX OPTIMIZATION

RECAP: • A SET $S \subseteq \mathbb{R}^n$ IS CONVEX IF

$$\left.\begin{matrix} x^1, x^2 \in S \\ \lambda \in (0,1) \end{matrix}\right\} \Rightarrow \lambda x^1 + (1-\lambda) x^2 \in S$$

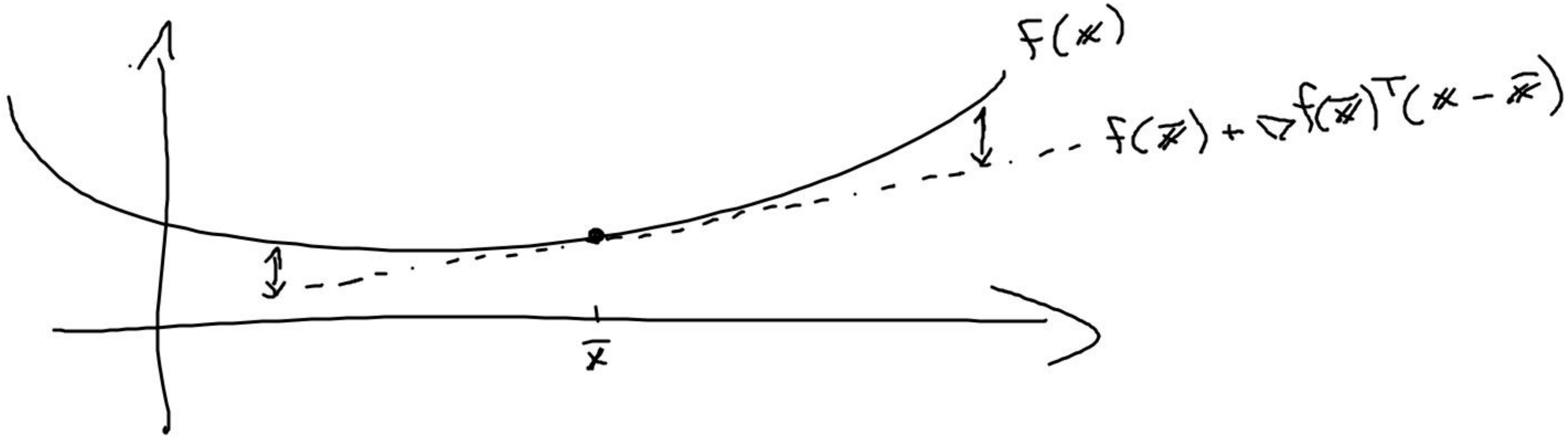
• A FUNCTION $f: \mathbb{R}^n \to \mathbb{R}$ IS CONVEX ON THE CONVEX SET S IF

$$\left.\begin{matrix} x^1, x^2 \in S \\ \lambda \in (0,1) \end{matrix}\right\} \Rightarrow f(\lambda x^1 + (1-\lambda) x^2) \leq \lambda f(x^1) + (1-\lambda) f(x^2)$$



IF $f \in C^1$ THEN f IS CONVEX IF

$$f(x) \geq f(\bar{x}) + \nabla f(\bar{x})^T (x - \bar{x}) \quad , \quad \forall x, \bar{x} \in S$$



A **CONVEX OPTIMIZATION PROBLEM** IS

$$f^* = \inf f(x) \quad \text{s.t. } x \in S$$

WHERE S IS A CONVEX SET AND f IS A CONVEX FUNCTION ON S

WE USUALLY HAVE:

$$\begin{aligned} f^* = \inf\ & f(x) \\ \text{s.t. } & g_i(x) \leq 0 \quad , i=1,\dots,m \\ & h_j(x) = 0 \quad , j=1,\dots,k \end{aligned}$$

IF
- f IS CONVEX
- g_i ARE CONVEX
- h_j ARE AFFINE

$\Rightarrow$ THE PROBLEM IS A CONVEX OPTIMIZATION PROBLEM.

NOT $\Leftarrow$. FOR EXAMPLE:

$$\begin{cases} \min\ x_1 + x_2 \\ \text{s.t. } x_1^2 + x_2^2 \leq 1 \\ \quad (x_1+4)^2 + (x_2-5)^2 \geq 1 \end{cases}$$

THEOREM: x^* LOCAL MIN TO COP $\Longrightarrow$ x^* GLOBAL MIN TO COP

CONVEX OPT. PROBLEM

- USUALLY ASSUME $f \in C^1$ AND CONSTRUCT METHODS, E.G.

$$x^{k+1} = x^k - \alpha_k \nabla f(x^k) \qquad \text{[STEEPEST DESCENT]}$$

- FOR CONVEX OPT. PROBLEMS THIS CAN BE RELAXED AND WE CAN INSTEAD CONSTRUCT METHODS USING SUBGRADIENTS

$$x^{k+1} = x^k - \alpha_k \mathbb{P}_k$$

SUBGRADIENTS

- IMPORTANT TO HANDLE $f \notin C^1$. E.G. $f(x) = \|x\|_1$ $f(x) = \|x\|_\infty$

DEFINITION: LET $S \subseteq \mathbb{R}^n$ BE A NONEMPTY CONVEX SET AND LET $f: S \to \mathbb{R}$ BE A CONVEX FUNCTION. THEN $p \in \mathbb{R}^n$ IS A **SUBGRADIENT** TO f AT $\bar{x}$ IF

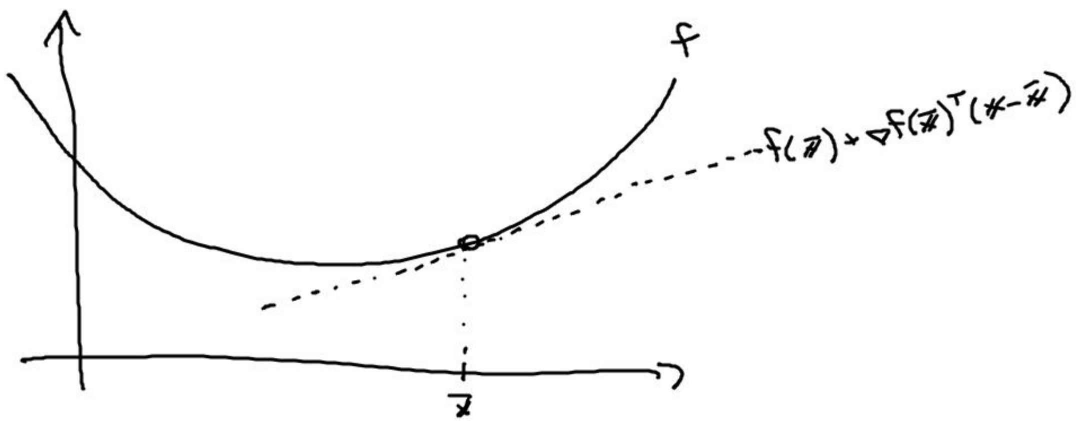
$$f(x) \geq f(\bar{x}) + p^T (x - \bar{x}) \quad \text{FOR ANY } x \in S$$

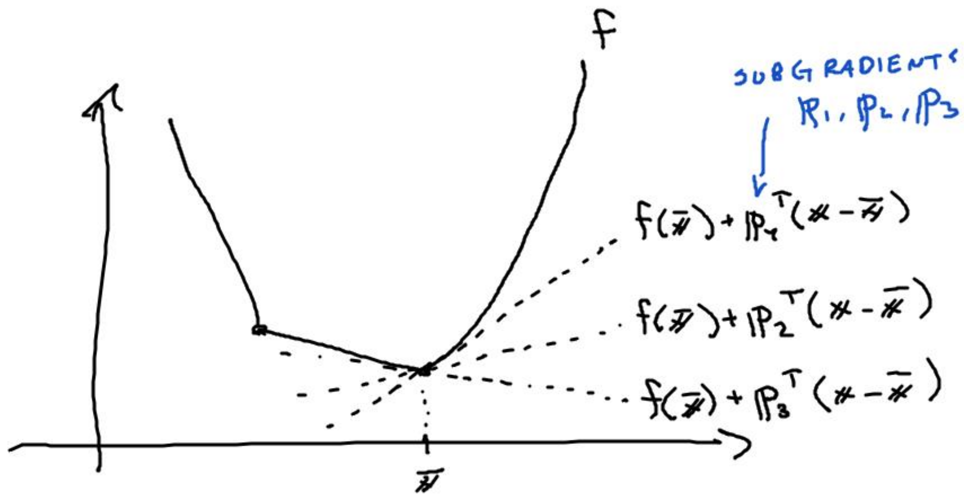
THE SET OF SUBGRADIENTS TO f AT $\bar{x}$ IS CALLED THE **SUBDIFFERENTIAL** OF f AT $\bar{x}$

$$\partial f(\bar{x}) = \left\{ p \in \mathbb{R}^n \mid f(x) \geq f(\bar{x}) + p^T (x - \bar{x}) \right\}$$

LEMMA: LET S BE NONEMPTY AND CONVEX AND $f: S \to \mathbb{R}$ A CONVEX FUNCTION. SUPPOSE THAT AT $\bar{x} \in \operatorname{int} S$, f IS DIFFERENTIABLE, MEANING THAT $\nabla f(\bar{x})$ EXISTS. THEN

$$\partial f(\bar{x}) = \{\nabla f(\bar{x})\}$$



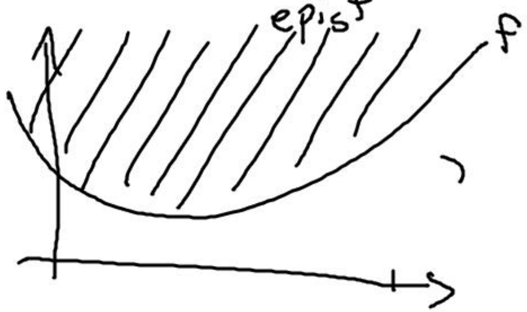


THEOREM: LET $S \subseteq \mathbb{R}^n$ BE A CONVEX SET AND $f : S \to \mathbb{R}$ A CONVEX FUNCTION. FOR EACH $\bar{x} \in \operatorname{int} S$, THERE EXISTS A VECTOR p SUCH THAT $f(x) \geq f(\bar{x}) + p^T(x - \bar{x})$.

i.e. THERE EXISTS A SUBGRADIENT

EPIGRAPH:

$$\operatorname{epi}_S f = \{(x, \alpha) \mid f(x) \leq \alpha\}$$

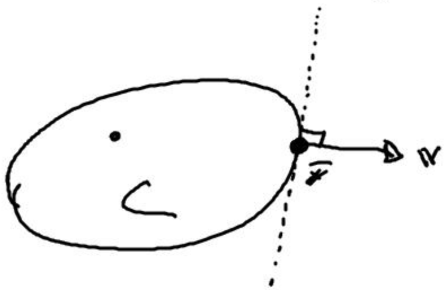


+

SUPPORTING HYPERPLANE THM

LET $\bar{x}$ BE A POINT ON THE BOUNDARY OF $C \subseteq \mathbb{R}^n$ (CONVEX).

THEN $\exists\, v \neq 0 : v^T(x - \bar{x}) \leq 0,\ \forall x \in C$ $\Longrightarrow$



HOMEWORK!

PROVE THEOREM ABOVE

- WE KNOW SUBGRADIENTS ALWAYS EXISTS!

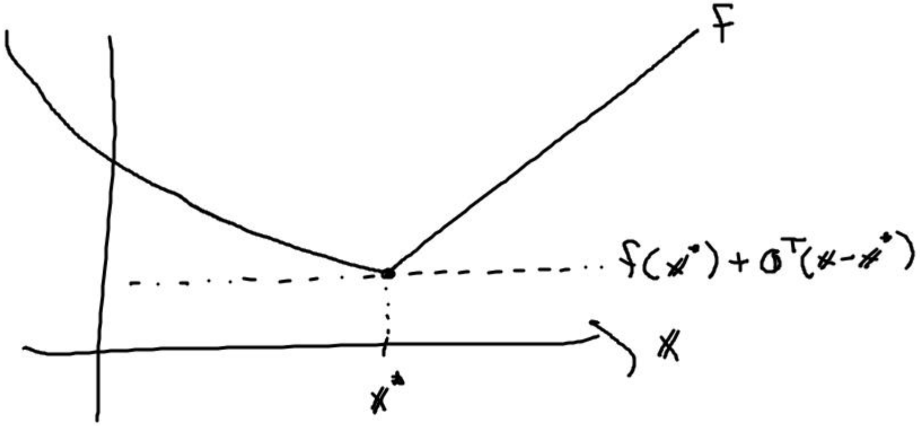
- HOW DO WE FIND THEM?
  - IN GENERAL, NO EASY WAY!

- HOW CAN USE THEM?
  - OPTIMALITY CONDITIONS
  - CONSTRUCTING OPT. METHODS.

$\min f(x)$
s.t. $x \in \mathbb{R}^n$

PROP: LET $f: \mathbb{R}^n \to \mathbb{R}$ BE CONVEX. THE FOLLOWING HOLDS:

$$f \text{ IS GLOBALLY MINIMIZED AT } x^* \in \mathbb{R}^n \iff 0 \in \partial f(x^*)$$



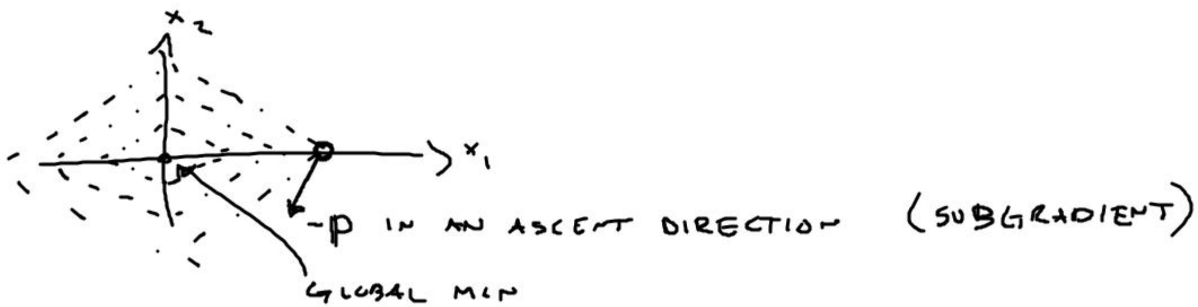
REMARK: IF $f \in C^1$.

$\Rightarrow$ PROPOSITION SAYS

x^* GLOBAL MIN $\iff \nabla f(x^*) = 0$

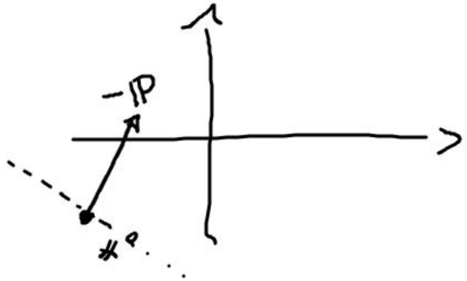
SUBGRADIENT METHODS

- ANALOGOUS TO GRADIENT METHODS. ITERATE $x^{k+1} = x^k - \alpha_k p^k$.
- <u>NOTE</u>: $-p$ NEED NOT BE A DESCENT DIRECTION

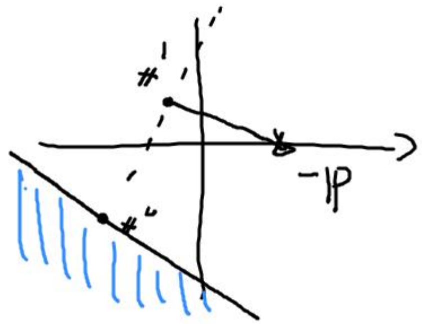


- HOWEVER, THEY STILL MOVE US "CLOSER" TO THE OPTIMAL SOLUTION.

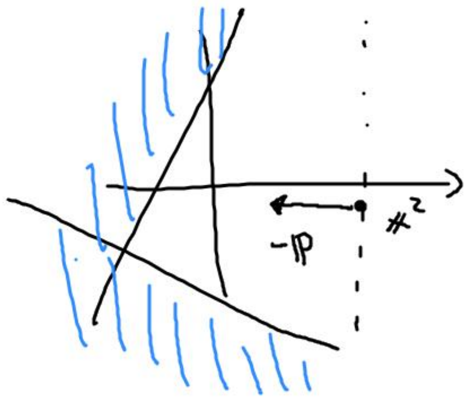
ITER 1:



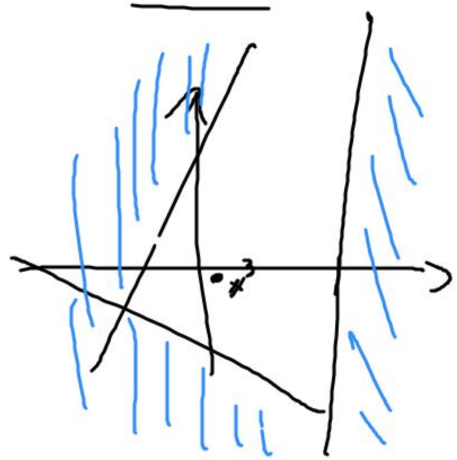
ITER 2:



ITER 3



ITER 4:



SUBGRADIENT METHOD (UNCONSTRAINED CASE $\min f(x)$ s.t. $x \in \mathbb{R}^n$)

0: INITIATE $x^0 \in \mathbb{R}^n$. $f^0_{BEST} = f(x^0)$, $k = 0$

1: FIND SUBGRADIENT p^k TO f AT x^k

2: UPDATE $x^{k+1} = x^k - \alpha_k p^k$ (α_k IS STEP LENGTH)

3: LET $f^{k+1}_{BEST} = \min \{ f^k_{BEST} , f(x^{k+1}) \}$

4: CHEK TERMINATION CRITERIA. IF NOT FULFILLED, GOTO 1.

STEP LENGTH RULES

- CONSTANT STEP SIZE.

$$\alpha_k = \alpha \quad \forall k$$

- SQUARE SUMMABLE BUT NOT SUMMABLE STEPS.

$$\sum_{k=0}^{\infty} \alpha_k^2 < \infty \quad , \quad \sum_{k=0}^{\infty} \alpha_k = \infty$$

Ex: $\alpha_k = \dfrac{a}{b+ck}$

LAGRANGIAN RELAXATION

$$f^* = \inf f(x)$$
$$\text{s.t. } g_i(x) \leq 0, \quad i=1,\dots,m$$
$$x \in X$$

DUAL FUNCTION $\quad q(\mu) = \inf_{x \in X} \left[f(x) + \mu^T g(x) \right]$

DUAL PROBLEM:

$$q^* = \sup q(\mu)$$
$$\text{s.t. } \mu \geq 0$$

- q CONCAVE FUNCTION
- $\mu \geq 0$ IS A CONVEX CONSTRAINT
- DUAL PROBLEM IS ACTUALLY A CONVEX OPTIMIZATION PROBLEM

NOT DEPENDING ON f, g_i, X

- $$\boxed{\begin{aligned} q^* = \ & \sup \ q(\mu) \\ & \text{s.t.} \ \ \mu \geq 0 \end{aligned}}$$

- HOW DO WE FIND SUBGRADIENTS TO q?

- IN ORDER TO EVALUATE $q(\mu)$ WE NEED TO SOLVE THE PROBLEM

$$q(\mu) = \inf_{x \in X} \left[f(x) + \mu^T g(x) \right]$$

LET $X(\mu)$ BE THE SOLUTION SET TO THIS PROBLEM.

PROP: ASSUME X IS NONEMPTY & COMPACT. THEN

a) LET $\mu \geq 0$. IF $x \in X(\mu)$, THEN $g(x)$ IS A SUBGRADIENT TO q AT μ

b) LET $\mu \geq 0$. THEN

$$\partial q(\mu) = \operatorname{conv} \{ g(x) \mid x \in X(\mu) \}.$$

← HOMEWORK!

- WHEN WE EVALUATE $q(\mu) = \inf_{x \in X} \left[f(x) + \mu^T g(x) \right]$ WE SOLVE THE "INNER" PROBLEM. LET $\bar{x}$ BE THE SOLUTION WE FOUND.
- THEN $g(\bar{x})$ IS A SUBGRADIENT TO q AT μ.

SUBGRADIENT METHOD FOR LAGR. DUAL PROBLEM

$$\boxed{\begin{array}{l} \sup q(\mu) \\ \text{s.t. } \mu \geq 0 \end{array}}$$

0. INITIALIZE $\mu^0 \geq 0$, $q^0_{BEST} = q(\mu^0)$. $k = 0$.

1. EVALUATE $q(\mu^k)$. i.e. SOLVE $q(\mu^k) = \inf\limits_{x \in X} \left(f(x) + \mu^{kT} g(x) \right)$

   LET x^k BE SOLUTION TO THE "INNER" PROBLEM

   THEN $g(x^k)$ IS A SUBGRADIENT TO q AT μ^k.

2. UPDATE $\mu^{k+1} = \left[\mu^k + \alpha_k \, g(x^k) \right]_+$ (PROJECTION ON NONNEGATIVE ORTHANT)

3. LET $q^{k+1}_{BEST} = \max\left(q^k_{BEST}, q(\mu^{k+1}) \right)$

4. CHECK TERMINATION CRITERIA. IF NOT FULFILLED, GOTO 1.