Class #21

Friday, April 11, 2025 10:24 AM

§4.2.2- In convex optimization (constrained or unconstrained), all local minima are global minima and there are no saddle points.

What is the optimality criterion for an x to be an optimal point for convex optimization?

Unconstrained optimization $\Rightarrow$ x^* is optimal $\Leftrightarrow$ $\nabla f(x^*) = 0$

Constrained (or unconstrained) Convex Optimization

An x^* is optimal if and only if

(i) $x^* \in C$ (C is the constraint set)

(ii) $\forall\, y \in C$, $\nabla f_0(x^*)^T (y - x^*) \geq 0$

Case I: $\nabla f_0(x^*) = 0$ and $x^* \in C$

Trivially $\nabla f_0(x^*)^T (y - x^*) = 0 \Rightarrow$ conditions (i) and (ii) are satisfied.

Case II: There does not exist any $x \in C$ s.t. $\nabla f(x) = 0$

but $\exists\, x^* \in C$ s.t. $\forall\, y \in \nabla f_0(x^*)^T (y - x^*) \geq 0$

$$-\nabla f_0(x^*)^T x^* + \nabla f_0(x^*)^T y \geq 0$$

$$-\nabla f_0(x^*)^T y \ldots \nabla f_0(x^*)^T x^* \leq 0$$

$$-\nabla f_0(x^*)^T y + \nabla f_0(x^*)^T x^* \le 0$$

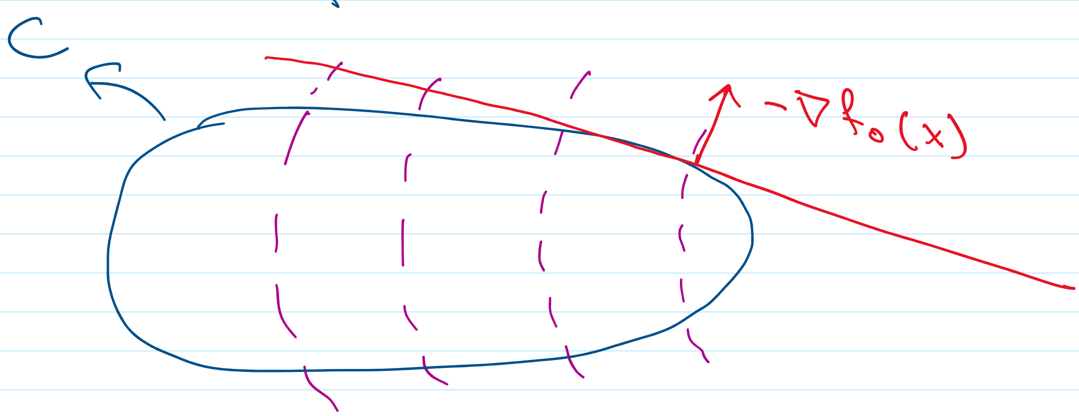
$$\underbrace{-\nabla f_0(x^*)^T y}_{a^T y} \le \underbrace{-\nabla f_0(x^*)^T x^*}_{\text{scalar} = b}$$

$$a^T y \le b \quad \forall\ y \in C$$

$$\Leftrightarrow a^T y - b \le 0 \quad \forall\ \underline{\underline{y \in C}}$$

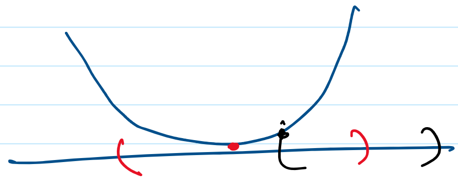
↪ Half Space

Normal vector for the half space is $-\nabla f_0(x^*)$



$-\nabla f_0(x^*)$ defines a "Supporting" hyperplane to the set C at $x^* \in C$.

The optimal point for a constrained convex optimization problem is then always on the boundary of C (unless $\exists\ x$ s.t. $\nabla f_0(x) = 0$ and $x \notin C$).



Proof : $x^* \in C$

$\Rightarrow$ Suppose $\nabla f_0(x^*)(y-x^*) \geq 0$

Linear lower bound

$\forall\, y \in \operatorname{dom} f$:

$$f_0(y) \geq f_0(x^*) + \underbrace{\nabla f_0(x^*)^T (y-x^*)}_{c\ \geq\ 0}$$

$\Rightarrow f_0(x^*) \leq f_0(y) - c \quad \forall\, y \in C$

$\Rightarrow f_0(x^*) \leq f_0(y) \quad \forall\, y \in C$

$\Leftarrow$ Suppose x^* is optimal (which means $x^* \in C$)

Show that $\nabla f_0(x^*)^T (y-x^*) \geq 0 \quad \forall\, y \in C$

Proof by Contradiction

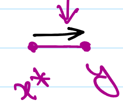
suppose $\exists\, y \in C$ s.t.

$$\nabla f_0(x^*)^T (y-x^*) < 0$$

$t \in [0,1]$

$z(t) = ty + (1-t)x^*$

$= x^* + t(y-x^*)$



Remark : $z(t) \in C$ (why?)
$\forall\, t \in [0,1]$ ↪ because C is convex

$$\frac{d}{dt} f_0(z(t))\Big|_{t=0} = \nabla f_0(x^*)^T (y - x^*) < 0$$

$\Rightarrow$ $f_0(z(t))$ is decreasing at x^* $\Rightarrow \exists\, t > 0$ s.t.

$f_0(z(t)) < f_0(x^*)$ $\Rightarrow$ Contradiction because x^* is optimal. $\blacksquare$

Special Case of Convex optimization

Linear Programming (LP)

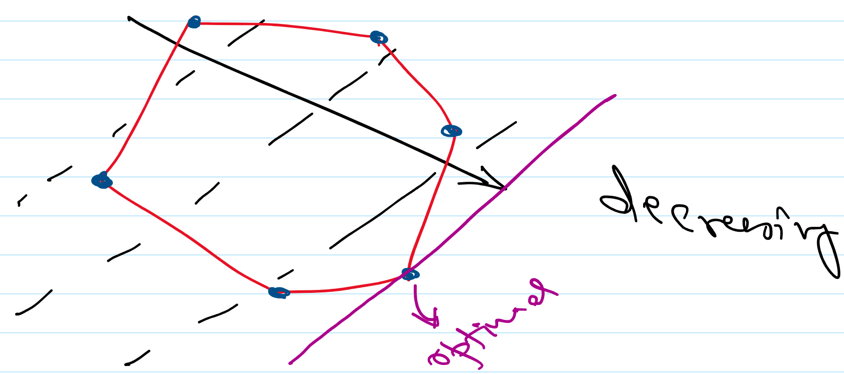
when $f_0(x)$ (objective function)

and $f_i(x)$ (Inequality constraint functions)

linear means both affine or linear

are linear then the optimization problem is called a linear program.

$\min_x \; a^T x$ $\Rightarrow$ Does not make sense without constraints



when we have linear equality and inequality constraints $\Rightarrow$ Constraint set is a polyhedron

When we have linear equality and inequality constraints $\Rightarrow$ Constraint set is a polyhedron or polytope.

$\Downarrow$ LP $\Rightarrow$ optimal is always on a vertex (if it exists)

Simplex algorithm $\Rightarrow$ Systematically traverses the vertices $\Rightarrow$ But can have exponential complexity (NP-hard algorithm).

Interior point methods $\Rightarrow$ They solve problems like LP and other convex optimization problems with constraints in guaranteed polynomial time.

Standard form LP

(i) $f_0(x)$ is linear

(ii) All equality constraints are linear

(iii) All inequality constraints are of the form:

$$x \geq 0 \quad \text{(non-negativity inequality constraints)}$$

e.g.,
$$x_1 + x_2 - 2 \leq 0$$

$\hookrightarrow$ introduce slack variable S

$$x_1 + x_2 - 2 + S = 0$$

$$s \succeq 0$$

↳ Slack variables are the ones that set inequality constraints in Standard form LP.

Equivalent convex problems

Depending upon the optimization toolbox / algorithm being used, we end up transforming convex optimization problems in many different ways:

① Sometimes we eliminate equality constraints.

→ Say x_0 satisfies $Ax = b \Rightarrow \{x : Ax = b\} = \{x_0 + \mathcal{N}(A)\}$

Say: $\boxed{Ax = b}$ is our equality constraint.

↳ $A(x_0 + v) = Ax_0 + \underbrace{Av}_{=0} = Ax_0 = b$ b/c $v \in \mathcal{N}(A)$

we change the problem to another variable z and find a matrix F s.t.

$$\text{range}(F) = \mathcal{N}(A)$$

↳ null space of A.

$$\{x : Ax = b\} = \{Fz + x_0 : z \in \mathbb{R}^k\}$$

$$k \geq \dim(\mathcal{N}(A))$$

New equivalent problem

$$\min_{z} \quad f_0(Fz + x_0)$$

$$\text{s.t.} \quad f_i(Fz + x_0) \leq 0, \quad i = 1, \ldots, m$$

→ This takes care of the equality constraint $Ax = b$

Fact: convex functions remain convex under affine maps

If $f(x)$ is convex

then $f(Ax+b)$ is convex.

② We can also bring in new equality constraints, provided they are linear.

↳ Example ⇒ Slack variables, but this only works when f_i's are linear.

③ Epigraph form of the problem

$$\min_{x} \; f_0(x)$$

$$\text{s.t.} \quad f_i(x) \le 0, \; i=1,\dots,m$$

$$Ax = b$$

Equivalent to

$$\Leftrightarrow \quad \min_{x,t} \; t$$

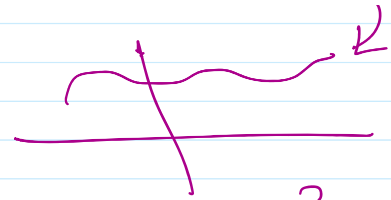
$f_0(x) \le t$

$$f_0(x) - t \le 0$$

$$f_i(x) \le 0, \; i=1,\dots,m$$

$$Ax = b$$

Graph of $f_0(x) = \{(x,t) : f_0(x) = t\} \subset \mathbb{R}^{n+1}$



Epigraph of $f_0(x) = \{(x,t) : f_0(x) \le t\} \subset \mathbb{R}^{n+1}$

A function $f_0(x)$ is convex if and only if its epigraph is a convex set.