Topics in Applied Optimization

Optimization for ML and Data Sciences

Dr. Pawan Kumar

September 30, 2020

IIIT, Hyderabad

Convex Optimization Problem: Local Optima = Global Optima

Fact: For a convex optimization problem, any local optima is a global optima

Ans: Proof on chalkboard!

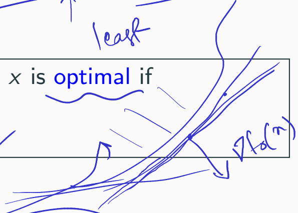
Convex Optimization Problem: Optimilaty Criteria

$f(y) \geq f(x^*) \quad \forall y$

least

Fact: If f_0 in a convex optimization problem is differentiable, then the point x is optimal if

$$\nabla f_0(x)^T (y - x) \geq 0 \quad \text{for all} \quad y \in X$$



Equivalent Cases

Proof: On chalkboard!

$f_0(x)$, cvx, diff.

x optimal $f(y) \geq f(x)$ $\forall y$ $\iff$ $\nabla f_0(x)^T (y - x) \geq 0$ $\forall y$ $\iff$ $\nabla f_0(x) = 0$

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

Convex Optimization Problem: Optimality for Unconstrained problems

Fact: For an unconstrained problem, $(m = p = 0,)$, the optimality condition $f(y) \le f(x^*)$

$$\nabla f_0(x)^T (y - x) \geq 0 \quad \text{for all} \quad y \in X$$

reduces to the well known necessary and sufficient condition

$$\nabla f_0(x) = 0$$

Proof on chalkboard!

Suppose x is optimal $\Rightarrow$ $x \in \operatorname{dom} f_0$ $\forall$ for all feasible y we have

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

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

~~Let us take~~

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

If $\nabla f_0(x^*) = 0$ $\Rightarrow$ $f_0(y) \geq f_0(x^*)$ $\forall y$

$\Rightarrow$ x^* is optimal

Scratch Space

So we prove that if $\exists\, x^*$ s.t. $\boxed{\nabla f(x^*) = 0} \Rightarrow x^*$ is optimal

~~Again if~~ $\Rightarrow$ $\underbrace{\nabla f_0(x)^T (y-x) \geq 0}$

Claim: Assume that $\nabla f_0(x)^T (y-x) \geq 0 \quad \forall\, y$, then

$\cancel{\nabla f_0(x) \geq 0}$

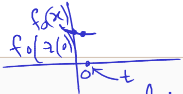
$\Rightarrow f_0(y) \geq f_0(x) + \underbrace{\nabla f_0(x)^T (y-x)}_{\geq 0} \leftarrow$ (Convexity)

$\Rightarrow \boxed{f_0(y) \geq f_0(x)} \Rightarrow x$ is optimal

Conversely: ~~Sps~~ that x is optimal and on contrary assume

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

Scratch Space



Consider that point: $z(t) = ty + (1-t)x$, $t \in [0,1]$

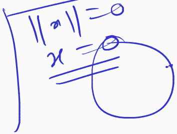
$z(t)$ is on the line joining x & y & $x, y \in$ feasible set which is cvx.

$\Rightarrow z(t) \in$ feasible set.

$$\frac{d}{dt} f_0(z(t))\Big|_{t=0} = \nabla f_0(x)^T z'(t)\Big|_{t=0} = \nabla f_0(x)^T (y-x) < 0$$

$\Rightarrow f_0(z(t))$ decreases as t increases. $\Rightarrow \exists$ small positive t for which $f_0(z(t)) < f_0(z(0)) = f_0(x)$

a contradiction to the fact that x was optimal



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

Take $y = \underbrace{x - t\nabla f_0(x)}$. For small t, point y is feasible

$$\nabla f_0(x)^T (y-x) = \nabla f_0(x)^T \left(-t\nabla f_0(x)\right) = \underbrace{-t}_{\leq 0} \underbrace{\|\nabla f_0(x)\|_2^2}_{\geq 0} \geq 0$$

$\boxed{\nabla f_0(x) = 0}$

Unconstrained Quadratic Optimization

Unconstrained Quadratic Optimization

Consider the problem of minimizing

$$f_0(x) = (1/2)x^T Px + q^T x + r,$$

Unconstrained Quadratic Optimization

Consider the problem of minimizing

$$f_0(x) = (1/2)x^T P x + q^T x + r,$$

where $P \in S^n_+$.

Unconstrained Quadratic Optimization

Consider the problem of minimizing

$$f_0(x) = (1/2)x^T Px + q^T x + r,$$

where $P \in S^n_+$. The necessary and sufficient condition for x to be minimizer of f_0 is

$$\nabla f_0(x) = Px + q = 0. \quad (\textit{Why?})$$

∵ f_0 is diff. (poly in x)

① Is f_0 convex?

$\nabla f_0(x) = Px + q$

$\nabla^2 f_0(x) = P \succeq 0$

$\Downarrow$

f_0 is cvx.

$\boxed{Px = -q}$

Unconstrained Quadratic Optimization

Consider the problem of minimizing

$$f_0(x) = (1/2)x^T Px + q^T x + r,$$

where $P \in S_+^n$. The necessary and sufficient condition for x to be minimizer of f_0 is

$$\nabla f_0(x) = Px + q = 0. \quad (\textit{Why?})$$

Several cases may occur:

Unconstrained Quadratic Optimization

Consider the problem of minimizing

$$f_0(x) = (1/2)x^T P x + q^T x + r,$$

where $P \in S_+^n$. The necessary and sufficient condition for x to be minimizer of f_0 is

$$\nabla f_0(x) = Px + q = 0. \quad (Why?)$$

Several cases may occur:

1. If $q \neq \mathcal{R}(P)$, then there is no solution. f_0 is unbounded below

Can never be satisfied

$Px = -q$

$$\begin{bmatrix} p_1 & p_2 & \cdots & p_n \end{bmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = -q$$

$= x_1 p_1 + x_2 p_2 + \cdots + x_n p_n$

Is q a L.C of p_i's

Is $q \in \mathcal{R}(P)$

Unconstrained Quadratic Optimization

Consider the problem of minimizing

$$f_0(x) = (1/2)x^T Px + q^T x + r,$$

where $P \in S^n_+$. The necessary and sufficient condition for x to be minimizer of f_0 is

$$\nabla f_0(x) = Px + q = 0. \quad (Why?)$$

Several cases may occur:

1. If $q \neq \mathcal{R}(P)$, then there is no solution. f_0 is unbounded below
2. If $P \succ 0$, then there is a unique minimizer, $x^* = -P^{-1}q$

$\Rightarrow P \in S^n_{++} \leftarrow$ $\underbrace{\text{strictly P.D.}}$ $\Rightarrow$ non-singularity.

Unconstrained Quadratic Optimization

Consider the problem of minimizing

$$f_0(x) = (1/2)x^T Px + q^T x + r,$$

where $P \in S^n_+$. The necessary and sufficient condition for x to be minimizer of f_0 is

$$\nabla f_0(x) = Px + q = 0. \quad (\textit{Why?})$$

Several cases may occur:

1. If $q \neq \mathcal{R}(P)$, then there is no solution. f_0 is unbounded below
2. If $P > 0$, then there is a unique minimizer, $x^* = -P^{-1}q$
3. If P is singular, but $q \in \mathcal{R}(P)$.

Unconstrained Quadratic Optimization

$x_1^2 - 2x_1x_2 + x_2^2 = (x_1 - x_2)^2 \geq 0$

Q $q = 0$, $r = 0$?

$$\underbrace{[x_1 \;\; x_2]}_{x^T} \underbrace{\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}}_{P} \underbrace{\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}}_{x}$$

$x^T P x \geq 0 \Rightarrow P \in S_+^2$

Consider the problem of minimizing

$$f_0(x) = (1/2)x^T P x + q^T x + r,$$

where $P \in S_+^n$. The necessary and sufficient condition for x to be minimizer of f_0 is

$$\boxed{\nabla f_0(x) = Px + q = 0.} \quad (\textit{Why?})$$

$\boxed{\nabla^2 f_0(x) = P}$

Several cases may occur:

no soln

unique

infinitely

1. If $q \neq \mathcal{R}(P)$, then there is no solution. f_0 is unbounded below
2. If $P > 0$, then there is a unique minimizer, $x^* = -P^{-1}q$
3. If P is singular, but $q \in \mathcal{R}(P)$.
   - set of optimal points is $X_{opt} = -P^+q + \mathcal{N}(P)$, where P^+ denotes the pseudoinverse of P

$x + v$

Problems with Equality Constraints

Problems with Equality Constraints

Consider the following optimization problem with equality constraints only

$$\begin{aligned} \text{minimize} \quad & f_0(x) \\ \text{subject to} \quad & Ax = b \end{aligned}$$

Problems with Equality Constraints

Consider the following optimization problem with equality constraints only

$$
\begin{aligned}
\text{minimize} \quad & f_0(x) \\
\text{subject to} \quad & Ax = b
\end{aligned}
$$

$f(y) \geq f(x) + \ldots$

- Feasible set is affine, recall that x is feasible if it satisfies

$$\nabla f_0(x)^T (y - x) \geq 0, \quad \text{for all } y \text{ such that } Ay = b$$

always

Problems with Equality Constraints

Consider the following optimization problem with equality constraints only

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & Ax = b \end{array}$$

- Feasible set is affine, recall that x is feasible if it satisfies

$$\nabla f_0(x)^T (y - x) \geq 0, \quad \text{for all } y \text{ such that } Ay = b$$

- Every feasible y has the form: $y = x + v$, (Why?) for some $v \in \mathcal{N}(A)$

Problems with Equality Constraints

Consider the following optimization problem with equality constraints only

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & Ax = b \end{array}$$

- Feasible set is affine, recall that x is feasible if it satisfies

$$\nabla f_0(x)^T (y - x) \geq 0, \quad \text{for all } y \text{ such that } Ay = b$$

- Every feasible y has the form: $y = x + v$, (Why?) for some $v \in \mathcal{N}(A)$
- Optimiality condition is

Problems with Equality Constraints

Consider the following optimization problem with equality constraints only

$$\begin{aligned} \text{minimize} \quad & f_0(x) \\ \text{subject to} \quad & Ax = b \end{aligned}$$

- Feasible set is affine, recall that x is feasible if it satisfies

$$\nabla f_0(x)^T (y - x) \geq 0, \quad \text{for all } y \text{ such that } Ay = b$$

- Every feasible y has the form: $y = x + v$, (Why?) for some $v \in \mathcal{N}(A)$
- Optimiality condition is

$$\boxed{\nabla f_0(x)^T v \geq 0 \quad \text{for all } v \in \mathcal{N}(A)}$$

Problems with Equality Constraints

$g($

Consider the following optimization problem with equality constraints only

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & Ax = b \end{array}$$

$g(v) = cv \geq 0 \quad \forall v \in S$

- Feasible set is affine, recall that x is feasible if it satisfies

$$\nabla f_0(x)^T (y - x) \geq 0, \quad \text{for all } y \text{ such that } Ay = b$$

$g(v) = \nabla f_0(x)^T v$

- Every feasible y has the form: $y = x + v$, (Why?) for some $v \in \mathcal{N}(A)$
- Optimiality condition is

$$\nabla f_0(x)^T v \geq 0 \quad \text{for all } v \in \underbrace{\mathcal{N}(A)}_{= S}$$

- $\nabla f_o(x)^T v$ is linear in v, and nonnegative on a subspace, hence

linear in v

$\boxed{g(v) \geq 0 \quad \forall v \in S}$

$-v \in S$

$\Rightarrow g(-v) = -g(v) \geq 0$

$\Rightarrow g(v) \leq 0$

$\Rightarrow \boxed{g(v) = 0}$

Problems with Equality Constraints

Consider the following optimization problem with equality constraints only

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & Ax = b \end{array}$$

- Feasible set is affine, recall that x is feasible if it satisfies

$$\nabla f_0(x)^T (y - x) \geq 0, \quad \text{for all } y \text{ such that } Ay = b$$

- Every feasible y has the form: $y = x + v$, (Why?) for some $v \in \mathcal{N}(A)$
- Optimiality condition is

$$\nabla f_0(x)^T v \geq 0 \quad \text{for all } v \in \mathcal{N}(A)$$

- $\nabla f_o(x)^T v$ is linear in v, and nonnegative on a subspace, hence
  - $\nabla f_0(x)^T v = 0$ for all $v \in \mathcal{N}(A)$. (Why?)

Problems with Equality Constraints

Consider the following optimization problem with equality constraints only

$$\begin{aligned} \text{minimize} \quad & f_0(x) \\ \text{subject to} \quad & Ax = b \end{aligned}$$

$x \in S = \{y \mid Ay = b\}$

- Feasible set is affine, recall that x is feasible if it satisfies

$$\nabla f_0(x)^T (y - x) \geq 0, \quad \text{for all } y \text{ such that } Ay = b$$

$\nabla f_0(x)^T v = 0 \quad \forall v \in \mathcal{N}(A)$

- Every feasible y has the form: $y = x + v$, (Why?) for some $v \in \mathcal{N}(A)$

$\Rightarrow \nabla f_0(x) \perp \mathcal{N}(A)$

- Optimiality condition is

$$\nabla f_0(x)^T v \geq 0 \quad \text{for all } v \in \mathcal{N}(A)$$

$\mathcal{N}(A) \oplus \mathcal{R}(A^T)$

- $\nabla f_o(x)^T v$ is linear in v, and nonnegative on a subspace, hence
  - $\nabla f_0(x)^T v = 0$ for all $v \in \mathcal{N}(A)$. (Why?)
  - $\nabla f_0(x) \perp \mathcal{N}(A) \implies \nabla f_0(x) \in \mathcal{R}(A^T)$, (Why?) i.e., $\exists v$ such that

$x = v + v^T$

Problems with Equality Constraints

$x^T P x + q^T x + r$

$$\begin{bmatrix} P & A^T \\ A & 0 \end{bmatrix}\begin{bmatrix} x \\ \nu \end{bmatrix} = \begin{bmatrix} -c \\ b \end{bmatrix}$$

Consider the following optimization problem with equality constraints only

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & Ax = b \end{array}$$

$\nabla f_0(x) = Px + c$

$Px + A^T\nu = -c$

$Ax = b$

- Feasible set is affine, recall that x is feasible if it satisfies

$$\nabla f_0(x)^T(y - x) \geq 0, \quad \text{for all } y \text{ such that } Ay = b$$

- Every feasible y has the form: $y = x + v$, (Why?) for some $v \in \mathcal{N}(A)$
- Optimiality condition is

$$\nabla f_0(x)^T v \geq 0 \quad \text{for all } v \in \mathcal{N}(A)$$

- $\nabla f_o(x)^T v$ is linear in v, and nonnegative on a subspace, hence
  - $\nabla f_0(x)^T v = 0$ for all $v \in \mathcal{N}(A)$. (Why?)
  - $\nabla f_0(x) \perp \mathcal{N}(A) \implies \nabla f_0(x) \in \mathcal{R}(A^T)$, (Why?) i.e., $\exists \nu$ such that

$$\nabla f_0(x) + A^T\nu = 0$$

  - The above together with $Ax = b$. Lagrange condition.

$\nabla f_0(x) \in \mathcal{R}(A^T)$

$\Rightarrow A^T v = \nabla f_0(x)$

$\boxed{\nabla f_0(x) + A^T u = 0}$

$\boxed{\nabla f_0(x)} + A^T\nu = 0$

$Ax = b$