Lecture 4

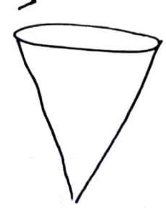
Convex Functions

TODAY REVIEW CONVEX ANALYSIS IN ORDER TO APPLY IT TO DUAL FORMULATION OF KANTOROVICH PROBLEM

Recall subgradient

$$\partial u(x) = \{ p \mid u(y) \geq \underbrace{u(x) + p(y-x)}_{\text{tangent plane}} \quad \forall y \}$$

Eg: $u(x_1, x_2) = \sqrt{x_1^2 + x_2^2}$



Find subgradient of u, $\partial u(\vec{x})$ (at the origin)

@ $x = \vec{0}$ we need $\vec{p}$ s.th

$|y| = \sqrt{y_1^2 + y_2^2} \geq p \cdot y$

magnitude of y

- if $y = 0$ this always holds
- @ $y \neq 0$ we need $\frac{p \cdot y}{|y|} \leq 1$

· $\forall p$: $\frac{py}{\|y\|} \leq \frac{\|p\| \, \|y\|}{\|y\|} = \|p\| \Rightarrow$ as long as p is less than one our condition is satisfied

· If $\hat{p} = \hat{y}$ we get equality

· If $\|p\| > 1$ we can find y to violate $\frac{py}{|y|} \leq 1$

$\therefore$ $\partial u(\vec{0}) = \{ p \in \mathbb{R}^2 \mid \|p\| \leq 1 \}$ $\rightarrow$ Now we conclude $\partial u(\vec{x}) = \begin{cases} B(\vec{0}, 1) & \text{if } \vec{x} = \vec{0} \\ \vec{x}/\|\vec{x}\| & \text{otherwise} \end{cases}$

$= B(\vec{0}; 1)$

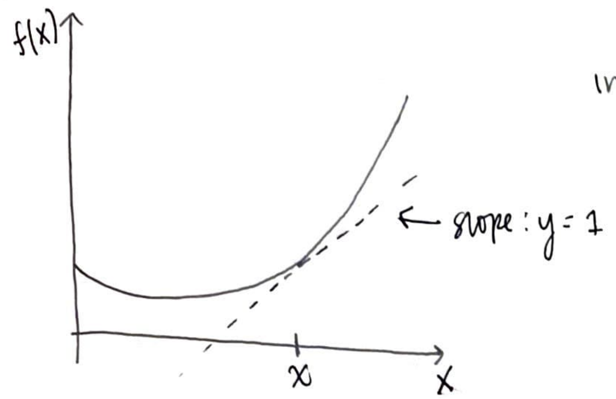
a ball of radius 1 centered @ origin

We could talk about subgradient @ a pt. or over a set.

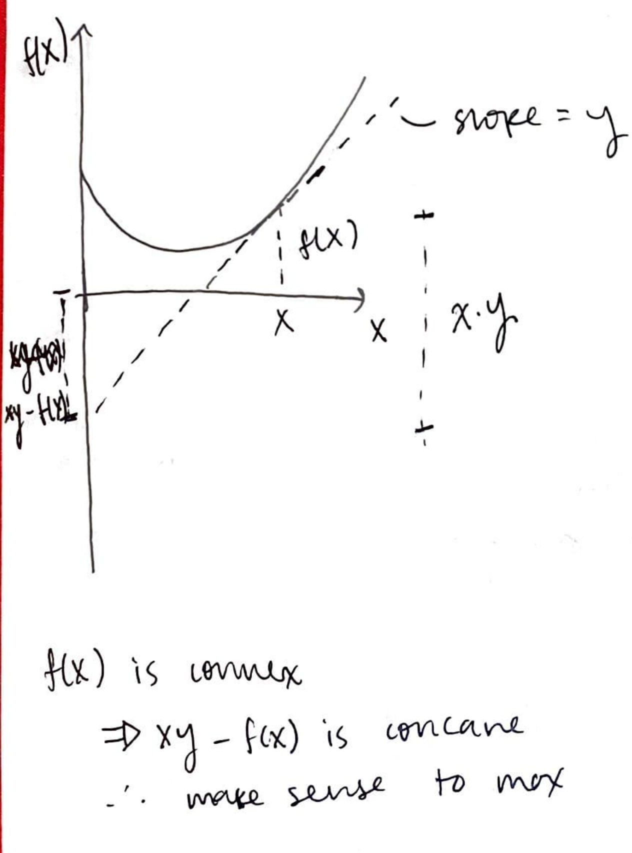
e.g. subgradient over all of $\mathbb{R}^2$

$$\partial U(\mathbb{R}^2) = \overline{B(\vec{0}, 1)}$$

Suppose F is convex on $\mathbb{R}$



interested in slope
→ can think of as indep. variable



represents y-intercept of tangent line

for fixed x define $g(y) = xy - f(x)$
where $y = f'(x)$

- This extra condition as comes from differentiating g w.r.t x and setting equal to 0
- maybe : $g(y) = \max_x/\min \{xy - f(x)\}$

f(x) is convex
$\Rightarrow xy - f(x)$ is concave
∴ make sense to max

• NOTE usually we seek to minimize a convex func., f(x). Since we have the negative of f(x) in $\{xy - f(x)\}$ we'll maximize whole expression.

Propose:

$g(y) = \max_{x} \{xy - f(x)\}$ ← Legendre transform background

may be interesting

Def (Legendre-Fenchel transform of f by

dot product b/c in $\mathbb{R}^n$ and not real line

$$\boxed{f^*(y) = \sup_{x \in E} \{x \overset{\downarrow}{\bullet} y - f(x)\}}$$

EX $f(x) = 0$ $E = \mathbb{R}$ ← domain is everything

$\Rightarrow f^*(y) = \sup_{x \in \mathbb{R}} \{xy\} = 0, \ y = 0$

domain $\text{dom}(f^*) = \{0\}$

EX $f(x) = 0$ $E = [-1, 1]$

$f^*(y) = \sup_{x \in [-1,1]} \{xy\} = |y|$

$\text{dom}(f^*) = \mathbb{R}$

EX $f(x) = p \cdot x \Rightarrow f^{**}(x) = p \cdot x$

*We can take repeated Legendre transform (AKA biconjugate of f)

*Property: f^* is convex

Let $y_1, y_2 \in \text{dom}(f^*)$ and $\lambda \in [0,1]$

$f^*(\lambda y_1 + (1-\lambda) y_2)$

$= \sup_x \{\lambda x \cdot y_1 + (1-\lambda) x \cdot y_2 - f(x)\} \leq \sup_x \{\lambda x \cdot y_1 - \lambda f(x)\} + \sup_x \{(1-\lambda) x y_2 - (1-\lambda) f(x)\}$

$= \lambda f^*(y_1) + (1-\lambda) f^*(y_2)$

ALGEBRA

??? $-\lambda f(x) - (1-\lambda) f(x)$

$\Rightarrow \cancel{-\lambda f(x)} - f(x) + \cancel{\lambda f(x)}$

$= -f(x)$ ✓

→ what we started w/

*NOTE whatever we start w/, once we take the Legendre transform we end up w/ a convex func.

Property: ~~∄~~ $\forall x \in dom(f)$ and $y \in dom(f^*)$ then

$$\boxed{-f(x) + f^*(y) \geq x \cdot y}$$

y is element of subgradient of $f(x)$

→ with the equality IFF $y \in \partial f(x)$

~~PROOF~~

PROOF
- inequality is immediate

- Let $y \in \partial f(x)$

$\iff f(z) \geq f(x) + y \cdot (z - x) \quad \forall z$

want to rearrange so I can get terms that look like legendre transform

$\iff x \cdot y - f(x) \geq z \cdot y - f(z) \quad \forall z$

If this is true for all z, it's also true when we take the supremum over all z

legendre transform

$\iff x \cdot y - f(x) \geq \sup_z \{ z \cdot y - f(z) \} = f^*(y)$

$\iff f(x) + f^*(y) \leq x \cdot y$

combined w/ $f(x) + f^*(y) \geq x \cdot y \quad \forall xy$

we get the equality

Property: If $f \leq g$ everywhere in the domain then $g^* \leq f^*$

⇒ i.e taking Legendre transform ~~preserves~~ or actually reverses ordering properties

Property: If f is convex and lower semi continuous then taking legendre transform twice gets us back where we started

$\Rightarrow f^{**}(x) = f(x)$

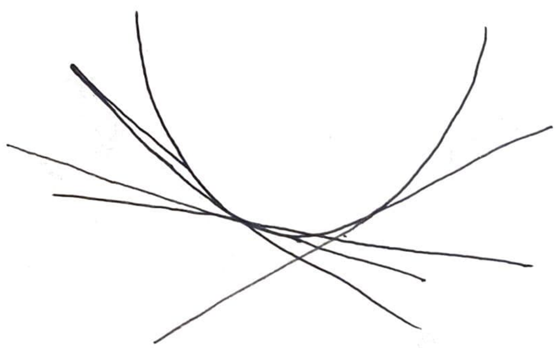
Property (continued):

- 1st: know that $f(x) + f^*(y) \geq x \cdot y$

$$\Rightarrow f^{**}(x) = \sup_y \{ x \cdot y - f^*(y) \} \leq f(x)$$

We aim to get equality here

gonna take function, f, & say ~~which~~ it's convex and: I can represent it as the supremum of a bunch of hyperplanes



$$f(x) = \sup_{\alpha \in \mathcal{A}} \{ L^\alpha(x) \}$$

hyperplanes

*this representation is specific to the fact that we have a convex func.

- Choose any $\alpha \in \mathcal{A}$

$$f(x) \geq L^\alpha(x)$$

$$\Rightarrow f^*(y) \leq L^{\alpha *}(y)$$

$$\Rightarrow f^{**}(x) \geq L^{\alpha **}(x)$$

$$= L^\alpha(x) \text{ since } L^\alpha \text{ is affine ???}$$

$$\Rightarrow f^{**}(x) \geq \sup_\alpha L^\alpha(x) = f(x)$$

this is how we represented convex func. originally

$$\therefore f^{**} = f$$

with this property it's now reasonable to talk about convex Legendre-Fenchel

- dual functions

$$\varphi, \psi \text{ s.th } \varphi = \psi^* \text{ \& } \psi = \varphi^*$$

Property if $\phi(x)$ & $\psi(y)$ are L-F duals on bounded domains X, Y then they have uniform Lipschitz bounds

PROOF:

$$\phi(x_1) - \phi(x_2) = \sup_{y \in Y} \{x_1 \cdot y - \psi(y)\} - \sup_{y \in Y} \{x_2 \cdot y - \psi(y)\}$$

$\forall \varepsilon > 0$ we can find a $y_1 \in \Psi$ s.th

$$\phi(x_1) - \phi(x_2) \le x_1 \cdot y_1 - \psi(y_1) + \varepsilon - x_2 \cdot y_1 + \psi(y_1)$$

$$= (x_1 - x_2) \cdot y_1 + \varepsilon$$

$$\le \sup_{y \in Y} |x_1 - x_2| + \varepsilon$$

if we take $\varepsilon \to 0$: some constant

$$\phi(x_1) - \phi(x_2) \le M |x_1 - x_2|$$

similarly we get bounds on the magnitude

$$|\phi(x_1) - \phi(x_2)| \le M |x_1 - x_2|$$

ϕ, ψ are uniformly Lipschitz ???

BACK TO KANTOROVICH DUALITY FOR QUADRATIC COST $\mathbb{E}^c$

$[c(x,y) = \frac{1}{2}|x-y|^2]$

Recall our dual problem

$$\max_{(u,v) \in \Phi} J[u,v]$$

where $J[u,v] = \int_X u(x)\, d\mu(x) + \int_Y v(y)\, d\nu(y)$

Since we want to use tools from convex analysis let's transform:

$$\phi(x) = \frac{1}{2}|x|^2 - u(x), \quad \psi(y) = \frac{1}{2}|y|^2 - v(y)$$

*are these arbitrary choices for ϕ & ψ ???

Instead of maximizing J we minimize $-J$,

$$-J = -\int_X u(x)\,d\mu(x) - \int_Y v(y)\,d\nu(y)$$

$$= \int_X \left(\varphi(x) - \tfrac{1}{2}|x|^2\right) d\mu(x) + \int_Y \left(\psi(y) - \tfrac{1}{2}|y|^2\right) d\nu(y)$$

OR just minimize ← how can we do this???

$$\int_X \varphi(x)\,d\mu(x) + \int_Y \psi(y)\,d\nu(y)$$

$$\equiv L[\varphi, \psi]$$ ← new obj. function

constraints:

$$\tfrac{1}{2}|x-y|^2 \geq u(x) + v(y)$$ ← quadratic cost provides an upper bound on $u+v$

$$= \tfrac{1}{2}|x|^2 - \varphi(x) + \tfrac{1}{2}|y|^2 - \psi(y)$$

① $\varphi(x) + \psi(y) \geq x \cdot y$

new constraint set: $\Phi^* = \{(\varphi, \psi) \in C^0(X) \times C^0(Y) \mid \varphi(x) + \psi(y) \geq x \cdot y \quad \forall x \in X, y \in Y\}$

New problem is DP^*

$$\min_{(\varphi,\psi) \in \Phi^*} L[\varphi, \psi]$$

The feasible set is non-empty → (e.g. large const. func's)

Let's try to understand what feasible pairs actually look like.

Let's start w/ the feasible pair $(\varphi, \psi) \in \Phi$

$\Rightarrow \varphi(x) \geq x \cdot y - \psi(y) \quad \forall\, x, y$

$\Rightarrow \varphi(x) \geq \psi^*(x)$

we have a new feasible pair (ψ^*, ψ^{**})

which are L-F duals

RECALL

$-\psi(y) \geq \psi^{**}(y)$

$-\psi^*(x) + \psi^{**}(y) \geq x \cdot y$

*NOTE If I start w/ a feasible pair I can build a new feasible pair

Let's look at obj. func.:

$$L[\psi^*, \psi^{**}] = \int_X \psi^*(x)\, d\mu(x) + \int_Y \psi^{**}(y)\, d\nu(y)$$

$$\leq \int_X \varphi(x)\, d\mu(x) + \int_Y \psi(y)\, d\nu(y)$$

$$= L[\varphi, \psi]$$

Conclusion: we can minimize over this set

$$\Phi^{**} = \{(\varphi, \psi) \in \Phi^* \mid \varphi = \psi^*, \ \psi = \varphi^*\}$$

*Idea: we don't need to look over entire feasible set we can ~~look @ the~~ focus on sets of Legendre duals b/c they are guaranteed to achieve min if it exists

Let's start w/ the feasible p

$\Rightarrow \varphi(x) \geq x \cdot y - \psi(y) \quad \forall$ x

$\Rightarrow \varphi(x) \geq \psi^*(x)$

we have a new feasible

(ψ^*, ψ^{**})

which are L-F duals

Let's look at obj. func.:

$L[\psi^*, \psi^{**}] = \int_X \psi^*(x)\, d\mu(x) + \int_Y$

$\leq \int_X \varphi(x)\, d\mu(x) + \int_Y \psi($

$= L[\varphi, \psi]$

Conclusion: we can mi

$\Phi^{**} = \{(\varphi, \psi) \in \Phi^* \mid$

SUMMARY

This lecture peeks into convex analysis so we can apply it to the dual formulation of the kantorovich problem. We start w/ an introduction to the subgradient which gives us background for Legendre-Fenchel transforms. We examine some properties of Legendre transforms. Two important properties are that legendre transforms are convex and performing Legendre transform twice results in the original function. This sets us up to discuss Legendre duals and rewrite our dual problem in terms of them. We go on to show that we don't need to minimize over entire feasible set, and instead can focus on Legendre pairs b/c they're guaranteed to achieve a min.

$\cdot y$

a

n

ime

LEC OUTLINE

- SUBGRADIENT
- LEGENDRE-FENCHEL TRANSFORMATION
- PROPERTIES OF LEGENDRE TRANSFORMS
- LEGENDRE DUALS
- KANTOROVICH DUALITY W/ QUADRATIC COST

on't need to look

the feasible set

~~look @ the~~ focus

of Legendre

/c they are

eed to achieve

if it exists