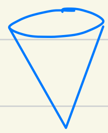
Convex functions.

$\partial u(x) = \{p \mid u(y) \ge u(x) + p\cdot(y-x)\ \forall y\}$.

• e.g. $u(x_1, x_2) = \sqrt{x_1^2 + x_2^2}$



find $\partial(\vec{x})$ at $x = \vec{0}$ we need to find $\vec{p}$ s.t.

$|y| = \sqrt{y_1^2 + y_2^2} \ge p\cdot y$, if $y = 0$, this always holds , if $y \ne 0$, we need $\frac{p\cdot y}{|y|} \le 1$.

$\forall p : \frac{p y}{\|y\|} \le \frac{\|p\|\cdot\|y\|}{\|y\|} = \|p\|$. This always holds for $\|p\| \le 1$.

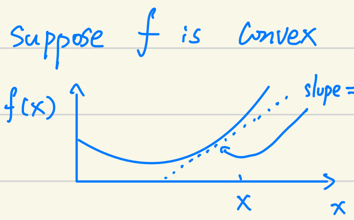
If $\hat{p} = \hat{y} \Rightarrow$ we get equality. i.e. if $\|p\| > 1$ we can find y that violate $\frac{p\cdot y}{|y|} \le 1$.

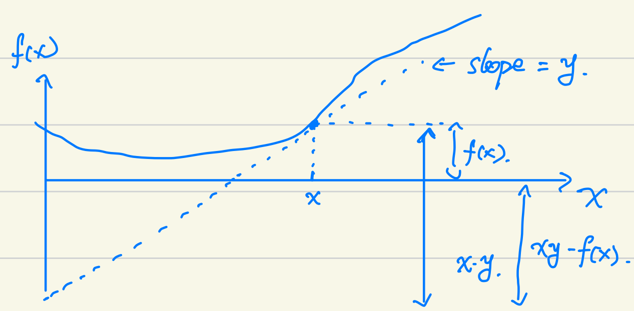
$\partial u(\vec{0})\ \{p \in \mathbb{R}^2 \mid \|p\| \le 1\} = B(\vec{0}; 1)$ ← Ball radius = 1 center at (0,0).

$$\boxed{\text{So } \partial u(\vec{x}) = \begin{cases} \overline{B(\vec{0}, 1)} & \text{if } \vec{x} = \vec{0} \\ \vec{x}/\|\vec{x}\| & \text{o.w.} \end{cases}}$$

$\Rightarrow \partial u(\mathbb{R}^2) = \overline{B(\vec{0}; 1)}$ ← sub-gradient.

Suppose f is convex on $\mathbb{R}$.





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

This extra condition also comes from diff g w.r.t. x and setting to 0.

maybe : $g(y) = \max_x / \min \{xy - f(x)\}$.

$f(x)$ is convex. $\Rightarrow xy - f(x)$ is concave make sense to maximize.

Promose that $g(y) = \max_x \{xy - f(x)\}$ may be interesting.

Promose that $g(y) = \max_x \{xy - f(x)\}$. may be interesting.

Define the <u>Legendre - Fenchel transform</u> of a f by $f^*(y) = \sup_{x \in E} \{x \cdot y - f(x)\}$.

- e.g. $f(x) = 0$ $E = \mathbb{R}$ $\Rightarrow f^*(y) = \sup_{x \in \mathbb{R}} \{xy\} = 0,\ y = 0.$ unbdd if $y \neq 0$.
  $\mathrm{Dom}(f^*) = \{0\}$.
- e.g. $f(x) = 0$ $E = [-1, 1]$. $\Rightarrow f^*(y) = \sup_{x \in [-1,1]} \{xy\} = |y|$, $\mathrm{Dom}(f^*) = \mathbb{R}$.
- e.g. $f(x) = p \cdot x$ $\Rightarrow f^{**}(x) = p(x)$.

We can take repeated L-F transforms (biconjugate of f).

- <u>Property</u>: f^* is convex Let $y_1, y_2 \in \mathrm{Dom}(f^*)$ and $\lambda \in [0,1]$.

$$f^*(\lambda y_1 + (1-\lambda)y_2) = \sup_x \{ \lambda x \cdot y_1 + (1-\lambda) x \cdot y_2 - \underbrace{f(x)}_{= \lambda f(x) + (1-\lambda) f(x)} \}$$

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

$$= \lambda f^*(y_1) + (1-\lambda) f^*(y_2). \quad (\text{Convexity.})$$

- <u>Property</u>: $\forall x \in \mathrm{Dom}(f)$ and $y \in \mathrm{Dom}(f^*)$ then $f(x) + f^*(y) \geq x \cdot y$. with equality iff $y \in \partial f(x)$.

pf: Inequality is immediate. let $y \in \partial f(x)$. $\Leftrightarrow f(z) \geq f(x) + y \cdot (z - x)$ $\forall z$.

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

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

Combine with $\underline{f(x) + f^*(y) \geq x\cdot y}$.

we get the equality. ∎

• Property: If $f \leq g$ everywhere then $g^* \leq f^*$.

• Property: If f is convex and lower-Semi-Continuous then $f^{**}(x) = f(x)$.

1st: know that $f(x) + f^*(y) \geq x\cdot y \Rightarrow \underline{f^{**}(x) = \sup_y \{x\cdot y - f^*(y)\} \leq f(x)}$.

Represent f as a supremum of hyperplanes. $f(x) = \sup_{\alpha\in\mathcal{A}} \{L^\alpha(x)\}$. ↑ hyper-planes.

choose any $\alpha \in \mathcal{A}$ $\quad f(x) \geq L^\alpha(x) \ \forall x \Rightarrow f^*(y) \leq L^{\alpha*}(y) \Rightarrow f^{**}(x) \geq \underset{\substack{\| \\ L^\alpha(x)}}{L^{\alpha**}(x)}$. Since L^α is affine.

$\Rightarrow \underbrace{f^{**}(x) \geq \sup_\alpha L^\alpha(x) = f(x)}_{\text{only hold for convex } f}$. and $\underbrace{f^{**}(x) \leq f(x)}_{\text{always hold.}}$ $\Rightarrow f^{**}(x) = f(x)$. ∎

It's now reasonable to talk about convex L-F Differential Dual functions φ, ψ s.t. $\varphi = \psi^*, \psi = \varphi^*$.

• Prop: If $\varphi(x), \psi(y)$ are L-F duals on bdd domains X, Y then they have uniform Lipschitz bdds.

pf: $\varphi(x_1) - \varphi(x_2) = \sup_{y\in Y}\{x_1\cdot y - \psi(y)\} - \sup_{y\in Y}\{x_2\cdot y - \psi(y)\}$.

$\forall \varepsilon > 0$ we can find $y_1 \in Y$ s.t. $\varphi(x_1) - \varphi(x_2) \leq \underline{x_1\cdot y_1 - \psi(y_1) + \varepsilon - x_2\cdot y_1 + \psi(y_1)}$.

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

$\leq \sup_{y\in Y} |y| \cdot |x_1 - x_2| + \varepsilon$.

Taking $\varepsilon \downarrow 0$: $\varphi(x_1) - \varphi(x_2) \leq M|x_1 - x_2|$, Similarly we get $|\varphi(x_1) - \varphi(x_2)| \leq M|x_1 - x_2|$.

$\Rightarrow \varphi, \psi$ are Uniformly Lipschitz. ∎

Let's come back to Dual pbm (DP) for $c(x,y) = \frac{1}{2}|x-y|^2$.

objective fun.

$$\max_{(u,v)\in\Phi} J[u,v] \text{ , where } J[u,v] = \int_X u(x)\,d\mu(x) + \int_Y v(y)\,d\nu(y).$$

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

$$\varphi(x) = \frac{1}{2}|x|^2 - u(x) \text{ , } \psi(y) = \frac{1}{2}|y|^2 - v(y).$$

Instead of max J we minimize $-J = -\int_X u(x)\,d\mu(x) - \int_Y v(y)\,d\nu(y)$

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

or just minimize. $\int_X \varphi(x)\,d\mu(x) + \int_Y \psi(y)\,d\nu(y) \equiv L[\varphi,\psi]$.

Constraints: $\frac{1}{2}|x-y|^2 \geq u(x) + v(y) = \frac{1}{2}|x|^2 - \varphi(x) + \frac{1}{2}|y|^2 - \psi(y)$.

$$\Rightarrow \varphi(x) + \psi(x) \geq x\cdot y.$$

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

New pbm is $(DP)^*$ $\min\limits_{(\varphi,\psi)\in\Phi^*} L[\varphi,\psi]$.

check Φ^* not empty: let $x\cdot y = 1000$. $\varphi(x) = 1000$, $\psi(y) = 1000$.

const fun   Const fun.

So the feasible set is non-empty (large Const fun).

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

Let $(\varphi, \psi) \in \Phi^* \Rightarrow \varphi(x) \geq x \cdot y - \psi(y)$. $\forall x, y$.

$\Downarrow$ take sup is L-F transform.

$\Rightarrow \underbrace{\varphi(x) \geq \psi^*(x)}_{①}$.

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

Also ② $\psi(y) \geq \psi^{**}(y)$. and ③ $\psi^*(x) + \psi^{**}(y) \geq x \cdot y$.

We have a new feasible pair (ψ^*, ψ^{**}). which are L-F dual pairs.

Let's look at objective function $L[\psi^*, \psi^{**}] = \int_X \psi^*(x)\, d\mu(x) + \int_Y \psi^{**}(y)\, d\nu(y)$.

not need to go over all possible pairs in $L[\varphi, \psi]$. fine to only look at L-F Duals $[\psi^*, \psi^{**}]$.

by ①&② $\leq \underbrace{\int_X \varphi(x)\, d\mu(x) + \int_Y \psi(y)\, d\nu(y)}$

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

- Conconclusion: We can minimize over this set $\Phi^{**} = \{(\varphi, \psi) \in \Phi^* \mid \varphi = \psi^*, \psi = \varphi^*\}$.