2. Convexity

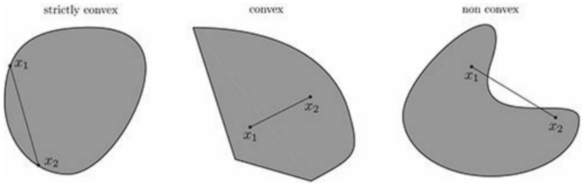
Convex Functions, Convex Problems and Lagrangian Basics

Sun Woo Lim

January 20, 2021

Convex Set

Definition of a Convex Set



$K \subseteq \mathbb{R}^n$ is a **convex set** if $\forall x_1, x_2 \in K$, $\forall \lambda \in [0,1]$, $\lambda x_1 + (1-\lambda)x_2 \in K$
"The Convex Combination also is in the set".

✓ Which of the following is / are convex?

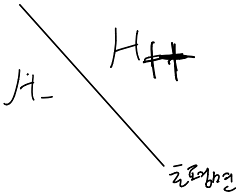
1. **Empty Set** ϕ O
2. **A set of single point** $\{x_0\}$ O
3. $\{z \in R^n : ||z - z_0||_2 \leq \epsilon\}$ **for some** $\epsilon > 0$ O
4. $\{z \in R^n : ||z - z_0||_2 = \epsilon\}$ **for some** $\epsilon > 0$ X
5. $[-2, -1] \cup [1, 2]$ X

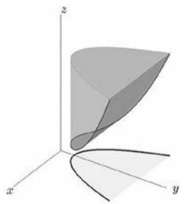
Operations preserving Convexity of a set

ex) Vectorspace $\mathbb{R}^2$에 있는 초평면과 두개의 반공간

Operations preserving Convexity of a set

1. **Intersection of convex sets**
2. **Hyperplane** $\{x|a^Tx - b = 0\}$ **and Half Spaces** $\{x|a^Tx - b \leq 0\}$ **and** $\{x|a^Tx - b > 0\}$
3. **Projection of a convex set onto a hyperplane**



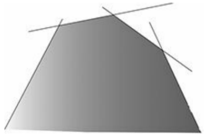


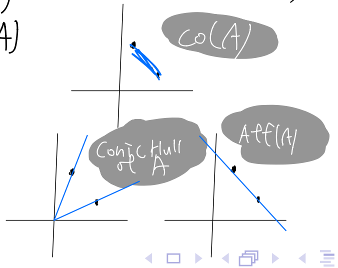
4. **"Convex Hull of A"** $Co(A) := \{\sum_{i=1}^{m} \lambda_i x_i | x_i \in A, \lambda_i \geq 0, \forall i, \sum \lambda_i = 1\}$: Smallest convex set containing A
5. **"Conic Hull of A"** $:= \{\sum_{i=1}^{m} \lambda_i x_i | x_i \in A, \lambda_i \geq 0, \forall i\}$
6. **"Affine Hull of A"** $:= \{\sum_{i=1}^{m} \lambda_i x_i | x_i \in A, \sum \lambda_i = 1\}$

$Co(A) \subseteq Aff(A) \subseteq Conic(A)$

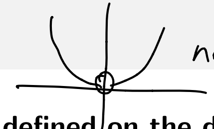
$A = \{(3,1), (1,3)\}$

Q) Why is a polyhedron convex?





Convex Functions



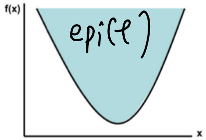
not cvx, why?

Convex / Concave Functions defined on the domain of a convex set

For $f : R^n \to R$ defined for $x \in dom(f)$ and assume f takes ∞ outside the domain. For f defined on convex domain, f is convex if $\lambda f(x_1) + (1-\lambda)f(x_2) \geq f(\lambda x_1 + (1-\lambda)x_2)$ and concave if $\lambda f(x_1) + (1-\lambda)f(x_2) \leq f(\lambda x_1 + (1-\lambda)x_2)$

✓ **Properties of Convex Functions**

1. **Pointwise Supremum of convex sets is a convex function**
2. **Nonnegative linear combination of Convex Functions is a convex function** f, g cvx → 2f+3g cvx
3. $f : \mathbb{R}^n \to \mathbb{R}$ **is a convex function** $\leftrightarrow$ $epi(f)$ **is a convex set : "Epigraph Characterization of a Convex Function"**
   $epi(f) := \{(x,t) \in \mathbb{R}^n \times \mathbb{R} : x \in dom(f), t \in \mathbb{R}, f(x) \leq t\}$



4. Jensen's Inequality : For a convex function f and a random variable X, $E(f(X)) \geq f(E(X))$

✓ **Iff conditions for differentiable convex functions**

For f which has an open domain and differentiable on $dom(f)$,

1. **First order (gradient) condition for convexity**
   f convex $\leftrightarrow f(y) \geq f(x) + \nabla f(x)^T(y-x), \forall x, y \in dom(f)$.
2. **Second order (Hessian) condition for convexity**
   f convex $\leftrightarrow \nabla^2 f(x) \succeq 0, \forall x \in dom(f)$. "Hessian is PSD".

ex 여러번미분가능한 cvx ftn이 entity
grad = 0 으로 solution을 구하고
Hess $\succeq 0$ 으로 그게 saddle pt나
local max 가 아님을 보증하는지에 대한 답

General Convex Optimization Problem

linear ftn: $y = a^T x$
affine ftn: $y = a^T x + b$

$min_x f_0(x)$: Objective Function
subject to (s.t.) $f_i(x) \leq 0, i = 1, 2, ...m,$: m inequality constraints
$h_i(x) = 0, i = 1, 2, ..., p$: p equality constraints

$\rightarrow$ This is a **Convex Optimization Problem** if $f_0, f_1, ..., f_m$ are convex functions and $h_1, ..., h_p$ are affine functions.

Terminologies

$\{x \in \{\cap_{i=0}^{m} dom(f_i) \cap \cap_{j=1}^{p} dom(h_i)\} | f_i(x) \leq 0, \forall i = 1, 2, ..., m, h_j(x) = 0, \forall j = 1, 2, ..., p\}$ is called a **Feasible Set**

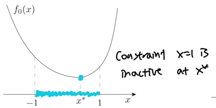
The **infimum** of the objective function over the **feasible set** is called the **primal optimal value**, denoted as $p*$.

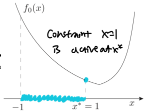
If $\exists$ feasible x satisfying $f_0(x) = p*$, say x **attains** the optimum and x^* is called **primal optimal point**.

The set of feasible points at which the optimum is attained is called an **Optimal Set**

Constraints f_i or h_j is(are) **active** at feasible point x if $f_i(x) = 0$ or $h_j(x) = 0$ respectively. Else, they are **inactive** at x.

Let $f : \mathbb{R}^n \rightarrow \mathbb{R}$ be a convex function with $dom(f)$. For $\alpha \in \mathbb{R}$, $S_\alpha := \{x \in \mathbb{R}^n | f(x) \leq \alpha\}$ is called a **Sublevel set** of f.





Dual Norm, Indicator Functions

Dual Norm of an arbitrary Norm $||x||* := sup_{\{z \in \mathbb{R}^n : ||z|| \leq 1\}} z^T x$

✓ Note that a **dual norm** is a **norm** and **every norm is a convex function**.
✓ For $p = 1, 2, 3, ..., \infty$, dual of l_p norm is l_q norm for p, q satisfying $\frac{1}{p} + \frac{1}{q} = 1 \because$ using Holder's Inequality.

- dual of l_1 norm is l_∞ norm and dual of l_∞ norm is l_1 norm
- dual of l_2 norm is l_2 norm

Indicator Function

$$I_S(x) = \begin{cases} 0 & x \in S \\ \infty & x \notin S \end{cases}$$

$\mathbb{R}_+ : \{r \mid r \geq 0\}$
$\mathbb{R}_- : \{r \mid r \leq 0\}$

$$I_{\mathbb{R}\_}(x) = \begin{cases} 0 & x \leq 0 \\ \infty & x > 0 \end{cases}$$

$$I_{\{0\}}(x) = \begin{cases} 0 & x = 0 \\ \infty & x \neq 0 \end{cases}$$

Totally different from the indicator function in probability theory (1 if in the set, 0 if not in the set).

Convex Conjugate of a Function

Let $f : \mathbb{R}^n \to \mathbb{R}$ **(need not be cvx ftn)** having a nonempty domain **(need not be cvx set)**.

$f^*(z) := sup_{x \in \mathbb{R}^n}(z^T x - f(x))$: **Convex Conjugate (Fenchel Conjugate)**

✓ Property 1) f^* is always **convex ftn** and **lower semicontinuous** (epigraph is a closed set)
✓ Property 2) If f is convex and **lower semicontinuous** them $f^{**} = f$

Examples

1. $f(x) = e^x \to f^*(z) = zlnz - z$
2. $f(x) = a^T x + b \to$
$$f^*(z) = \begin{cases} -b & z = a \\ \infty & z \neq a \end{cases}$$
3. $f(x) = ||x|| \to$ conjugate of a norm is "the indicator function of unit dual norm ball"
$$f^*(z) = \begin{cases} 0 & ||z||_* \leq 1 \\ \infty & ||z||_* > 1 \end{cases}$$
4. $I^*_{\mathbb{R}_-}(z) = I_{\mathbb{R}_+}(z)$. Also, $I_{\mathbb{R}_-}(x) = sup_{z \geq 0} zx$: either by direct calculation or applying dual of dual
5. $I^*_{\{0\}}(z) = 0$. Also, $I_{\{0\}}(x) = sup_{z \in \mathbb{R}} zx$: either by direct calculation or applying dual of dual

The Lagrangian

✓ Find a **Convex optimization** primal probelm to another **dual** problem, wanting that the **Dual** is easier to solve! Actually, can dualize non-convex problems to make a convex dual, but not dealt here.

✓ Formulate a function called **The Lagrangian** that integrates all the **constraints** into an **unconstrained problem**.

$L : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^p$ having $dom(L) = D \times \mathbb{R}^m \times \mathbb{R}^p$, $L(x, \lambda, \nu) := f_0(x) + \sum_{i=1}^m \lambda_i f_i(x) + \sum_{i=1}^p \nu_j h_j(x)$ is convex in x.

Key Idea behind the Lagrangian

$p* = min_x[f_0(x) + \sum_{i=1}^m I_{\mathbb{R}_-}(f_i(x)) + \sum_{j=1}^p I_{\{0\}}(h_j(x))]$: **Pay infinite price for disobeying the constrints**

Then, use indicator functions : $I_{\mathbb{R}_-}(f_i(x)) = sup_{\lambda_i \geq 0}\{\lambda_i f_i(x)\}$ and $I_{\{0\}}(h_j(x)) = sup_{\nu_j \in \mathbb{R}}\{\nu_j h_j(x)\}$

$\rightarrow p* = min_x max_{\lambda \geq 0, \nu} L(x, \lambda, \nu)$

Lagrangian Duality

$p* = min_x max_{\lambda \geq 0, \nu} [f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x) + \sum_{j=1}^{p} \nu_j h_j(x)]$

The Lagrange Dual Function

$g : \mathbb{R}^m \times \mathbb{R}^p \to \mathbb{R}, g(\lambda, \nu) := inf_{x \in D} L(x, \lambda, \nu)$

g is a **Concave Extended Real valued Function** possibly taking $-\infty$ as a function value or a function that is ∞ everywhere.

Dual Optimization Problem

If g is not an everywhere ∞ function, the **Dual Problem**

$d^* := sup_{\lambda \succeq 0, \nu} g(\lambda, \nu) = max_{\lambda \succeq 0, \nu} min_x L(x, \lambda, \nu)$ $= \max_{\lambda \geq 0, \nu} \min_{x} \left[f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x) + \sum_{j=1}^{p} \nu_j h_j(x) \right.$

A real convex problem is : $-d^* := inf_{\lambda \succeq 0, \nu} \{-g(\lambda, \nu)\}$ a.k.a $inf_{\lambda, \nu} \{-g(\lambda, \nu)\}$ s.t. $\lambda \succeq 0$

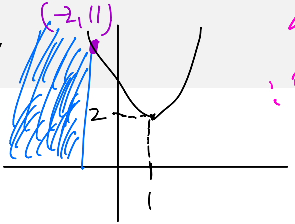
✓ Always, $d^* \leq p^*$: **weak duality**

✓ Under "good" conditions, $d^* = p^*$: **strong duality**

If $d^* = p^*$: **strong duality** and let x^* be a **primal optimal point** and let (λ^*, ν^*) be a **dual optimal point**.

Then, $f_0(x^*) = g(\lambda^*, \nu^*) = inf_{x \in D} L(x, \lambda^*, \nu^*)$

Example of Lagrangian, Lagrangian Dual, Duality



실생활 : 이딴 함수 다루려고 이 어려운 과목을 만든건 아닐꺼다. 보통 다루는 건 고차원의 어려운 함수

$p^* = min_x[(x-1)^2 + 2]$ s.t. $x + 2 \leq 0$

Surely, $(x^*, p^*) = (-2, 11)$

↳ 그냥 풀면 앎

$*$ Indicator function $I_{\mathbb{R}-}(x+2) = \sup_{\lambda \geq 0} \lambda(x+2)$

$\Rightarrow \mathcal{L} :$ dom$(f) = \mathbb{R} \times \mathbb{R} \to \mathbb{R} : \mathcal{L}(x, \lambda) = (x-1)^2 + 2 + \lambda(x+2)$

$p^* = \min_x \max_{\lambda \geq 0} [(x-1)^2 + 2 + \lambda(x+2)]$ & $d^* = \max_{\lambda \geq 0} \min_x [(x-1)^2 + 2 + \lambda(x+2)]$

$g(\lambda) = \min_x \mathcal{L}(x, \lambda) = -\frac{\lambda^2}{4} + 3\lambda + 2$, $d^* = \sup_{\lambda \geq 0} \left(-\frac{\lambda^2}{4} + 3\lambda + 2\right) = \sup_{\lambda \geq 0} \left(-\frac{1}{4}(\lambda - 6)^2 + 11\right)$

$\therefore (\lambda^*, d^*) = (6, 11)$. **From Calculation**, you know **that Strong Duality Holds**.

Next time :

✓ How to easily check **if Strong Duality holds** without calculation like this : **Slater's Condition**

✓ How to easily find the solution using the **KKT conditions**

✓ You'll see how useful **Convex Conjugate** are in deriving a dual.