Convex Function

2023-01-19, 김종민, 전상후

Part 1.

Basic properties and examples

I. Definition

1. Convex function

- 함수 $\boldsymbol{f}: \mathbb{R}^{\boldsymbol{n}} \longrightarrow \mathbb{R}$의 정의역이 $convex\ set$, 임의의 두 점 $x, y \in dom\ \boldsymbol{f}$를 잇는 선분 위의 모든 점들이 함수 $\boldsymbol{f}$ 위의 점들보다 위에 있다면 그 함수 $\boldsymbol{f}$는 $\boldsymbol{convex}$

- $f(\theta x + (1-\theta)y) \leq \theta f(x) + (1-\theta)f(y)$, $with\ 0 \leq \theta \leq 1, for\ all\ x, y \in dom\ f$

2. Strictly convex

- $f(\theta x + (1-\theta)y) < \theta f(x) + (1-\theta)f(y)$, $with\ x, y \in dom\ f, x \neq y, 0 < \theta < 1$

3. Concave function

- f가 convex이면 –f는 concave

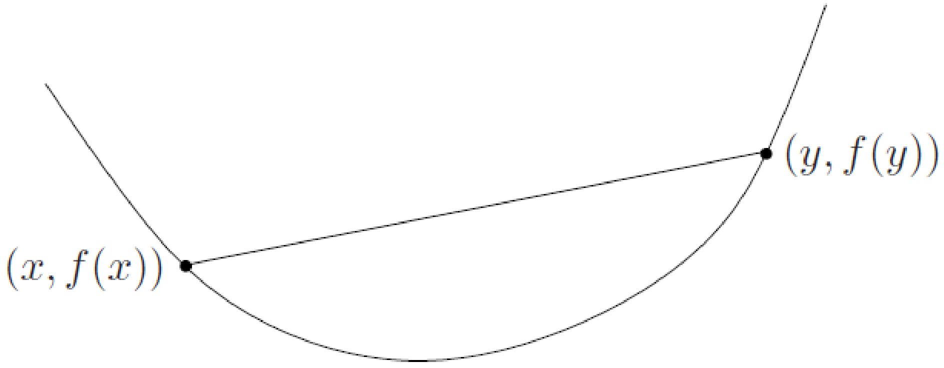
I. Definition

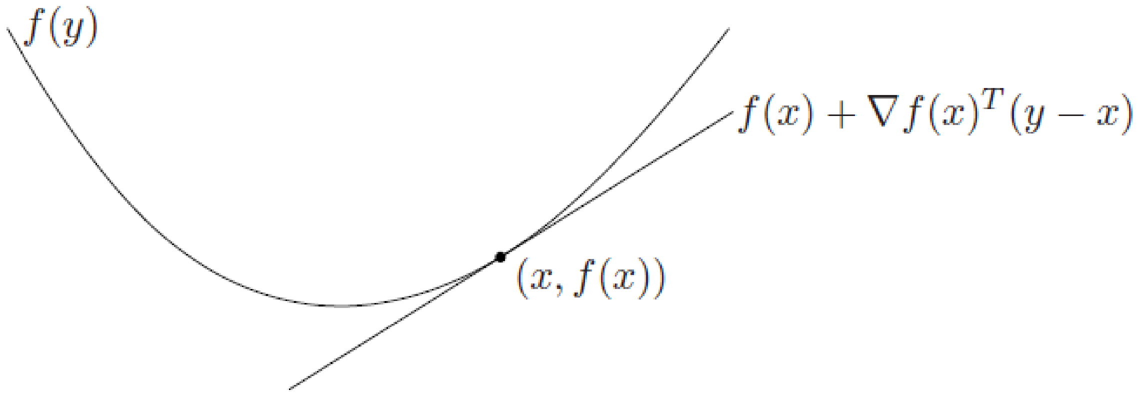
4. Affine 함수 ($f(x) = a^T x + b$)

$$
\begin{aligned}
f(\theta x + (1-\theta)y) &= a^T(\theta x + (1-\theta)y) + b \\
&= \theta a^T x + (1-\theta)a^T y + \theta b + (1-\theta)b \\
&= \theta f(x) + (1-\theta)f(y)
\end{aligned}
$$

for all $x, y \in \mathrm{dom}\, f$, with $0 \le \theta \le 1$

식이 성립하므로 affine 함수는 항상 convex이면서 동시에 concave 이다





II. Conditions

1. First-order conditions

- *Suppose f* is differentiable. Then f is convex if and only if dom *f* is convex and

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

- 위 식을 First order approximation으로 생각해볼 수 있으며 이는 global under estimator

2. Second-order conditions

- assume that f is twice differentiable. Then f is convex if and only if dom f is convex and its Hessian is positive semidefinite

- For a function on R, this reduces to the simple condition $f''(x) \geq 0$, which means that the derivative is nondecreasing.

$$\nabla^2 f(x) \geq 0$$

- It can be interpreted geometrically as the requirement that the graph of the function have positive (upward) curvature at x.

III. Examples

1. R 상의 convex function

- e^{ax} (지수 함수)
- $xlogx$ (negative entropy)

2. R^n 상의 convex function

- Norm (subadditivity 등 특성 활용한 오른쪽 식에 의해)
- Affine function : $a^T x + b \ on \ R$

$$\begin{aligned} \| \theta x + (1-\theta) y \| &\leq \| \theta x \| + \| (1-\theta) y \| \\ &= \theta \| x \| + (1-\theta) \| y \| \end{aligned}$$

3. $R^{n \times m}$ 상의 convex function

- Affine function : $tr(A^T X) + b = \sum_{i=1}^{m} \sum_{j=1}^{n} A_{ij} X_{ij} + b$
- Spectral norm : $f(X) = \| X \|_2 = \sigma_{max}(X) = (\lambda_{max}(X^T X))^{1/2}$

- Max function, Quadratic-over-linear function, Geometric mean 등등 (책 pg73~72 참고)

Log-sum-exponential 을 활용한 증명

$$f(x) = \log\left(e^{x_1} + \cdots + e^{x_n}\right)$$

Log-sum-exp. The Hessian of the log-sum-exp function is

$$\nabla^2 f(x) = \frac{1}{(\mathbf{1}^T z)^2}\left((\mathbf{1}^T z)\,\mathbf{diag}(z) - zz^T\right),$$

where $z = (e^{x_1}, \ldots, e^{x_n})$. To verify that $\nabla^2 f(x) \succeq 0$ we must show that for all v, $v^T \nabla^2 f(x) v \geq 0$, *i.e.*,

$$v^T \nabla^2 f(x) v = \frac{1}{(\mathbf{1}^T z)^2}\left(\left(\sum_{i=1}^n z_i\right)\left(\sum_{i=1}^n v_i^2 z_i\right) - \left(\sum_{i=1}^n v_i z_i\right)^2\right) \geq 0.$$

But this follows from the Cauchy-Schwarz inequality $(a^T a)(b^T b) \geq (a^T b)^2$ applied to the vectors with components $a_i = v_i\sqrt{z_i}$, $b_i = \sqrt{z_i}$.

- Log-sum-exp. 함수를 미분하면 softmax 함수와 같습니다
- Softmax는 입력받은 값을 출력으로 0~1사이의 값으로 모두 정규화하며 출력 값들의 총합은 항상 1이 되는 특성을 가졌습니다.
- 딥러닝의 다중 클래스 분류 모델의 활성화함수로 사용됩니다.

Log-sum-exponential 을 활용한 증명

정의 활용

$$f(x) = \log\left(e^{x_1} + \cdots + e^{x_n}\right)$$

Let $u_i = e^{x_i}, v_i = e^{y_i}$. So $f(\theta x + (1-\theta)y) = log(\sum_{i=1}^{n} e^{\theta x_i + (1-\theta)y_i}) = log(\sum_{i=1}^{n} u_i^{\theta} v_i^{(1-\theta)})$

From Hölder's inequality:

$$\sum_{i=1}^{n} x_i y_i \leq (\sum_{i=1}^{n} |x_i|^p)^{\frac{1}{p}} \cdot (\sum_{i=1}^{n} |y_i|^q)^{\frac{1}{q}}$$

where $1/p + 1/q = 1$.

Applying this inequality to $f(\theta x + (1-\theta)y)$:

$$log(\sum_{i=1}^{n} u_i^{\theta} v_i^{(1-\theta)}) \leq log[(\sum_{i=1}^{n} u_i^{\theta \cdot \frac{1}{\theta}})^{\theta} \cdot (\sum_{i=1}^{n} v_i^{1-\theta \cdot \frac{1}{1-\theta}})^{1-\theta}]$$

The right formula can be reduced to:

$$\theta log \sum_{i=1}^{n} u_i + (1-\theta) log \sum_{i=1}^{n} v_i$$

Here I regard θ as $\frac{1}{p}$ and $1-\theta$ as $\frac{1}{q}$.

So I achieve that $f(\theta x + (1-\theta)y) \leq \theta f(x) + (1-\theta)f(y)$.

Log-sum-exponential 을 활용한 증명

Second-order condition 활용

$$f(x) = \log\left(e^{x_1} + \cdots + e^{x_n}\right)$$

Let $f(z) = \log \sum_{i=1}^{n} z_i = \log 1^T z.$

$$\frac{\partial}{\partial x_j} f(x) = \frac{\partial}{\partial z_j} \log 1^t z \cdot \frac{\partial z_j}{\partial x_j} = \frac{1}{1^t z} z_j$$

$$\begin{aligned}
\frac{\partial^2 f}{\partial x_i \partial x_j} &= \frac{\partial}{\partial z_i}\left(\frac{z_j}{1^t z}\right) \cdot \frac{\partial z_i}{\partial x_i} \\
&= \frac{\delta_{ij} 1^t z - z_j}{(1^t z)^2} \cdot \exp x_i \\
&= \frac{\delta_{ij} z_i \cdot 1^t z - z_i z_j}{(1^t z)^2} \\
&= \frac{\delta_{ij} z_i}{1^t z} - \frac{z_i z_j}{(1^t z)^2} \\
&= \left(\frac{1}{1^t z} \operatorname{diag}(z) - \frac{1}{(1^t z)^2} z z^t\right)_{i,j}
\end{aligned}$$

$\delta_{ij} = 1\ if\ i = j, 0\ otherwise$

IV. Extended value extensions

- 특정 convex function을 모든 실수 공간으로 확장하려 할 때, 간단하게 기존 함수의 domain이 아닌 영역을 +무한대로 설정함

- Extended value extension $\tilde{f}$ $of\ f\ is$

$$\tilde{f}(x) = f(x)\ if\ x \in dom\ f,$$
$$\tilde{f}(x) = \ \infty,\ \ if\ x \notin dom\ f$$

- Often simplifies notation ; for example, the condition

$$0 \le \theta \le 1 \Longrightarrow \ \tilde{f}(\theta x + (1-\theta)y) \le \ \theta\tilde{f}(x) + (1-\theta)\tilde{f}(y)$$

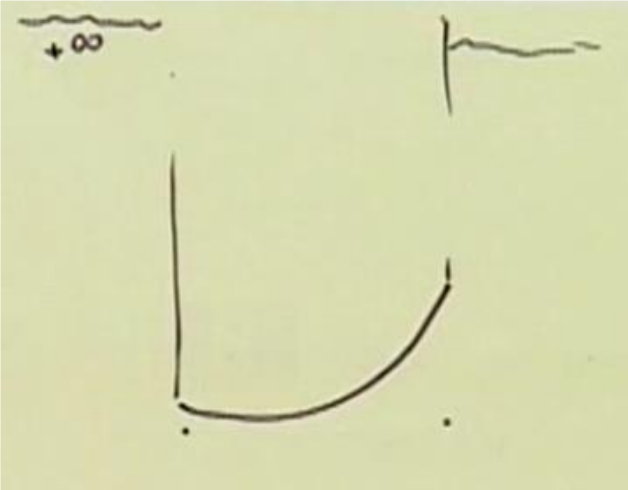


사진 출처:https://strutive07.github.io/2020/02/08/Lecture.3-logistics.html

cf. Restriction of a convex function to a line

Restriction of a convex function to a line

$f : \mathbf{R}^n \to \mathbf{R}$ is convex if and only if the function $g : \mathbf{R} \to \mathbf{R}$,

$$g(t) = f(x + tv), \qquad \mathbf{dom}\, g = \{t \mid x + tv \in \mathbf{dom}\, f\}$$

is convex (in t) for any $x \in \mathbf{dom}\, f$, $v \in \mathbf{R}^n$

can check convexity of f by checking convexity of functions of one variable

example. $f : \mathbf{S}^n \to \mathbf{R}$ with $f(X) = \log\det X$, $\mathbf{dom}\, f = \mathbf{S}^n_{++}$

$$\begin{aligned} g(t) = \log\det(X + tV) &= \log\det X + \log\det(I + tX^{-1/2}VX^{-1/2}) \\ &= \log\det X + \sum_{i=1}^{n} \log(1 + t\lambda_i) \end{aligned}$$

where λ_i are the eigenvalues of $X^{-1/2}VX^{-1/2}$

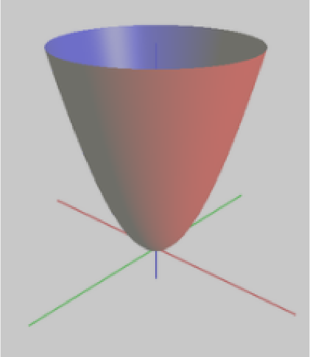
g is concave in t (for any choice of $X \succ 0$, V); hence f is concave

특정 함수가 하나 있을 때 domain 에서 하나의 점을 찍고, 그 점에서 하나의 방향으로 쭉 선을 그었다고 생각해보자.
함수의 단면이라고 생각하면 될듯하다.
그 단면이 convex 한가? 를 측정하는 것이다.
여기서 domain에 있는 모든 x 에 대하여, 어떤 방향 v 으로 이 동작을 하던 간에, 항상 그 단면이 convex 하다면, 그 함수는 convex function 이다.

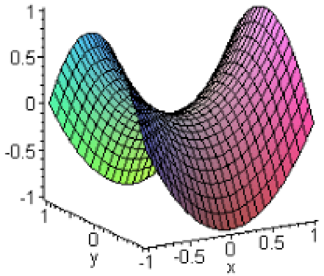
cf. Restriction of a convex function to a line

1] By restricting it to a line means, basically, you draw line in the domain of the function; then you evaluate your function only along that line.

2] Imagine a paraboloid $f : \mathbb{R} \times \mathbb{R} \mapsto \mathbb{R}$ defined by $f(x, y) = \frac{x^2}{a^2} + \frac{y^2}{b^2}$.



Now, if you draw a line in the domain and evaluate this paraboloid only along that line, it would look like a parabola. Analytically, if you want to check how the function would be along the x-axis, then substitute y = 0 in the equation above and you get $f(x, y) = \frac{x^2}{a^2}$ which you might know is the equation for the parabola. Now, a parabola is convex and since every line in the domain here would give you a parabola, a paraboloid is convex. On the other hand, if you take a hyperbolic paraboloid:



You draw a line in the domain in one direction, it would look like a parabola and you draw a line in the domain in another direction, it would look like an inverted parabola. Now, inverted parabolas are concave and not convex. Therefore, hyperbolic paraboloids are not convex.

- Images have been borrowed from the internet.

공간 상의 다면체 → x, y값에 대한 z값으로 나타남

여기서 line restriction → x, y의 모든 영역이 아닌 한 직선에 대해서만 z값을 보는 것

왼쪽 위 그림에서 z값의 범위를 line으로 제한해서 보면 포물선이 나타남, 다른 직선으로 제한해서 봐도 포물선이 나타남 (convex함)

대부분의 직선에 대해서 restriction의 결과가 convex하면, 왼쪽 위 함수는 convex함을 귀납적으로 유추 가능하다는 의의가 있음

V. Sublevel sets, epigraph

1. Sublevel sets

$\alpha - sublevel\ set\ of\ f: R^n \rightarrow R$:

$$C_\alpha = \{x \in dom\ f \mid f(x) \leq \alpha\}$$

Sublevel sets of a convex function are convex, for any value of α.

2. Epigraph

$a\ epigraph\ of\ f: R^n \rightarrow R$:

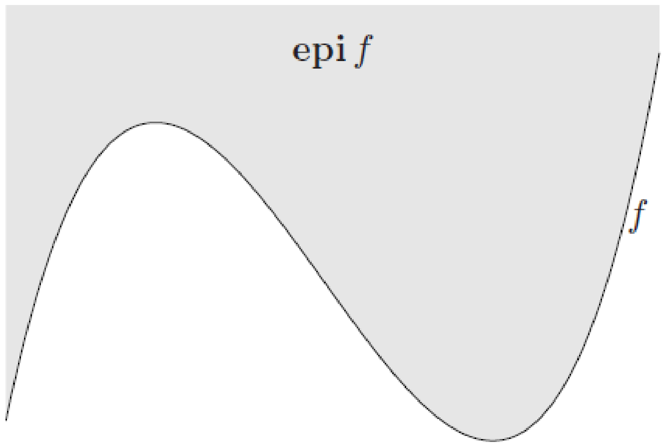
$$\text{epi } f = \{(x, t) \in R^{n+1} \mid x \in dom\ f, f(x) \leq t\}$$

f is convex if and only if epi f is a convex set. (epi means above, so epigraph means above graph)

You need to show that $f(x) \leq \alpha$, where x is chosen as convex combination of the points x_1 and x_2, i.e. $x = (1-\lambda)x_1 + \lambda x_2$. Now,

$$\begin{aligned} f(x) &= f((1-\lambda)x_1 + \lambda x_2) \\ &\leq (1-\lambda)f(x_1) + \lambda f(x_2) \quad \text{(using convexity of } f(\cdot)) \\ &\leq (1-\lambda)\alpha + \lambda\alpha \quad \text{(using epigraph definition)} \\ &= \alpha \end{aligned}$$

Thus, epi f is convex. q.e.d



VI. Jensen's inequality

- 함수 f 가 convex이고, n개의 양수 $w_1, \ldots, w_n$ 에 대하여 $\sum_{i=1}^{n} w_i = 1$일 때 아래 식 성립

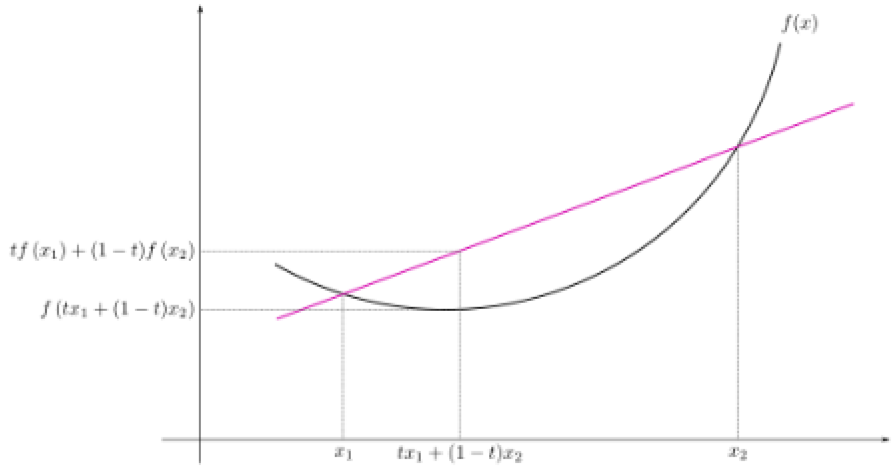
$$\sum_{i=1}^{n} w_i f(x_i) \geq f\left(\sum_{i=1}^{n} w_i x_i\right)$$

- f가 convex이면 다음 부등식 만족

$$f(tx_1 + (1-t)x_2) \leq tf(x_1) + (1-t)f(x_2) \; for \; 0 \leq t \leq 1$$

Extension:

X is a random variable supported on $\mathbf{dom}\, f$, then $f(E[X]) \leq E[f(X)]$



Part 2.

Operations that preserve convexity

(Convex function의 convexity를 유지하기 위한 연산)

0. 함수의 convexity 판단하는 방법

1. 정의를 확인한다 (보통 위의 line으로 restriction하는 것으로 확인)

2. second-order condition 체크

3. Simple한 convex function을 다음 operation에 의해 변형시켜도 여전히 convex이다.

   3-1. nonnegative weighted sum
   3-2. composition with affine function
   3-3. pointwise maximum and supremum
   3-4. composition
   3-5. minimization
   3-6. perspective

I. Nonnegative weighted sums & composition with affine function

- **Nonnegative multiple**: α f is convex if f is convex, $\alpha \geq 0$
- **Sum** : $f_1 + f_2$ $convex\ if\ f_1, f_2\ convex\ (extends\ to\ infinite\ sums, integral$
- **Composition with affine function**: $f(Ax+b)\ is\ convex\ if\ f\ is\ convex$

E.g.)

● Log barrier for linear inequalities

$$f(x) = -\sum_{i=1}^{m} \log(b_i - a_i^T x), dom\ f = \{x \mid a_i^T x < b_i, i = 1, \dots, m\}$$

● (any) norm of affine function

$$f(x) = \| Ax + b \|$$

Summary

- convex function에 non-negative 상수를 곱해도 convex function 이다
- 두 convex function을 합하여도 convex function 이다
- domain 에 affine function을 적용해도, convex function 이다

II. Pointwise maximum & supremum

If f_1 and f_2 are convex functions then their *pointwise maximum* f, defined by

$$f(x) = \max\{f_1(x), f_2(x)\},$$

with $\mathbf{dom}\, f = \mathbf{dom}\, f_1 \cap \mathbf{dom}\, f_2$, is also convex. This property is easily verified: if $0 \le \theta \le 1$ and $x,\ y \in \mathbf{dom}\, f$, then

$$\begin{array}{rcl} f(\theta x + (1-\theta)y) & = & \max\{f_1(\theta x + (1-\theta)y), f_2(\theta x + (1-\theta)y)\} \\ & \le & \max\{\theta f_1(x) + (1-\theta) f_1(y), \theta f_2(x) + (1-\theta) f_2(y)\} \\ & \le & \theta \max\{f_1(x), f_2(x)\} + (1-\theta)\max\{f_1(y), f_2(y)\} \\ & = & \theta f(x) + (1-\theta) f(y), \end{array}$$

which establishes convexity of f. It is easily shown that if $f_1, \ldots, f_m$ are convex, then their pointwise maximum

$$f(x) = \max\{f_1(x), \ldots, f_m(x)\}$$

is also convex.

Example 3.6 *Sum of r largest components.* For $x \in \mathbf{R}^n$ we denote by $x_{[i]}$ the ith largest component of x, *i.e.*,

$$x_{[1]} \ge x_{[2]} \ge \cdots \ge x_{[n]}$$

are the components of x sorted in nonincreasing order. Then the function

$$f(x) = \sum_{i=1}^{r} x_{[i]},$$

i.e., the sum of the r largest elements of x, is a convex function. This can be seen by writing it as

$$f(x) = \sum_{i=1}^{r} x_{[i]} = \max\{x_{i_1} + \cdots + x_{i_r} \mid 1 \le i_1 < i_2 < \cdots < i_r \le n\},$$

i.e., the maximum of all possible sums of r different components of x. Since it is the pointwise maximum of $n!/(r!(n-r)!)$ linear functions, it is convex.

As an extension it can be shown that the function $\sum_{i=1}^{r} w_i x_{[i]}$ is convex, provided $w_1 \ge w_2 \ge \cdots \ge w_r \ge 0$. (See exercise 3.19.)

두 convex function max 만 뽑은 function 또한 convex function 이다.

(교재 pg 80, Example 3.6 참고)

II. Pointwise maximum & supremum

Supremum에 대한 설명 및 예시

The pointwise maximum property extends to the pointwise supremum over an infinite set of convex functions. If for each $y \in \mathcal{A}$, $f(x, y)$ is convex in x, then the function g, defined as

$$g(x) = \sup_{y \in \mathcal{A}} f(x, y) \tag{3.7}$$

is convex in x. Here the domain of g is

$$\mathbf{dom}\, g = \{x \mid (x, y) \in \mathbf{dom}\, f \text{ for all } y \in \mathcal{A}, \ \sup_{y \in \mathcal{A}} f(x, y) |< \infty\}.$$

Example 3.8 *Distance to farthest point of a set.* Let $C \subseteq \mathbf{R}^n$. The distance (in any norm) to the farthest point of C,

$$f(x) = \sup_{y \in C} \|x - y\|,$$

is convex. To see this, note that for any y, the function $\|x - y\|$ is convex in x. Since f is the pointwise supremum of a family of convex functions (indexed by $y \in C$), it is a convex function of x.

show that the sup of this family is convex.

Let $(g_i)_{i \in I}$ be a family of convex functions on a convex compact set $\Omega \subseteq \mathbb{R}^d$.

Let $g := \sup_{i \in I} g_i$.

Take $x, y \in \Omega$ and $t \in [0, 1]$.

Fix $i \in I$. Since g_i is convex and bounded above by g, we have

$$g_i(tx + (1-t)y) \le t g_i(x) + (1-t) g_i(y) \le t g(x) + (1-t) g(y).$$

Since the latter holds for every $i \in I$, we can take the sup and find

$$g(tx + (1-t)y) \le t g(x) + (1-t) g(y).$$

This holds for every $x, y \in \Omega$ and every $t \in [0, 1]$. So g is convex.

Now every affine function f_i is convex, so the result follows from the general case above.

A function is convex if its epigraph is convex.
It is clear that the epigraph of *sup gi* is the intersection of all the *gi*.
Now the intersection of convex sets is convex, which yields a more geometric proof of the statement above.

III. composition

Scalar composition

composition of $g : \mathbf{R}^n \to \mathbf{R}$ and $h : \mathbf{R} \to \mathbf{R}$:

$$f(x) = h(g(x))$$

f is convex if g convex, h convex, $\tilde{h}$ nondecreasing; g concave, h convex, $\tilde{h}$ nonincreasing

vector composition

composition of $g : \mathbf{R}^n \to \mathbf{R}^k$ and $h : \mathbf{R}^k \to \mathbf{R}$:

$$f(x) = h(g(x)) = h(g_1(x), g_2(x), \ldots, g_k(x))$$

f is convex if g_i convex, h convex, $\tilde{h}$ nondecreasing in each argument; g_i concave, h convex, $\tilde{h}$ nonincreasing in each argument

e.g.

$\exp g(x)$ is convex if g is convex

$1/g(x)$ is convex if g is concave and positive

- The function $h(z) = \log(\sum_{i=1}^k e^{z_i})$ is convex and nondecreasing in each argument, so $\log(\sum_{i=1}^k e^{g_i})$ is convex whenever g_i are.

IV. minimization

Convex function의 minimum과 infimum은 convex function이다

if $f(x,y)$ is convex in (x,y) and C is a convex set, then

$$g(x) = \inf_{y \in C} f(x,y)$$

is convex

e.g.

Example 3.16 *Distance to a set.* The distance of a point x to a set $S \subseteq \mathbf{R}^n$, in the norm $\|\cdot\|$, is defined as

$$\mathbf{dist}(x,S) = \inf_{y \in S} \|x - y\|.$$

The function $\|x-y\|$ is convex in (x,y), so if the set S is convex, the distance function $\mathbf{dist}(x,S)$ is a convex function of x.

V. Perspective

convex function의 perspective function 는 convex function 이다.

Perspective

the **perspective** of a function $f : \mathbf{R}^n \to \mathbf{R}$ is the function $g : \mathbf{R}^n \times \mathbf{R} \to \mathbf{R}$,

$$g(x,t) = tf(x/t), \qquad \mathbf{dom}\, g = \{(x,t) \mid x/t \in \mathbf{dom}\, f,\ t > 0\}$$

g is convex if f is convex

e.g.

$f(x) = x^T x$ is convex; hence $g(x,t) = x^T x/t$ is convex for $t > 0$

과제

Exercise 3.1
(pg. 113)

3.1 Suppose $f : \mathbf{R} \to \mathbf{R}$ is convex, and $a,\ b \in \mathbf{dom}\, f$ with $a < b$.

(a) Show that

$$f(x) \leq \frac{b-x}{b-a} f(a) + \frac{x-a}{b-a} f(b)$$

for all $x \in [a, b]$.

(b) Show that

$$\frac{f(x)-f(a)}{x-a} \leq \frac{f(b)-f(a)}{b-a} \leq \frac{f(b)-f(x)}{b-x}$$

for all $x \in (a, b)$. Draw a sketch that illustrates this inequality.

(c) Suppose f is differentiable. Use the result in (b) to show that

$$f'(a) \leq \frac{f(b)-f(a)}{b-a} \leq f'(b).$$

Note that these inequalities also follow from (3.2):

$$f(b) \geq f(a) + f'(a)(b-a), \qquad f(a) \geq f(b) + f'(b)(a-b).$$

(d) Suppose f is twice differentiable. Use the result in (c) to show that $f''(a) \geq 0$ and $f''(b) \geq 0$.

Conjugate Function

ESC

Optimization for Machine Learning

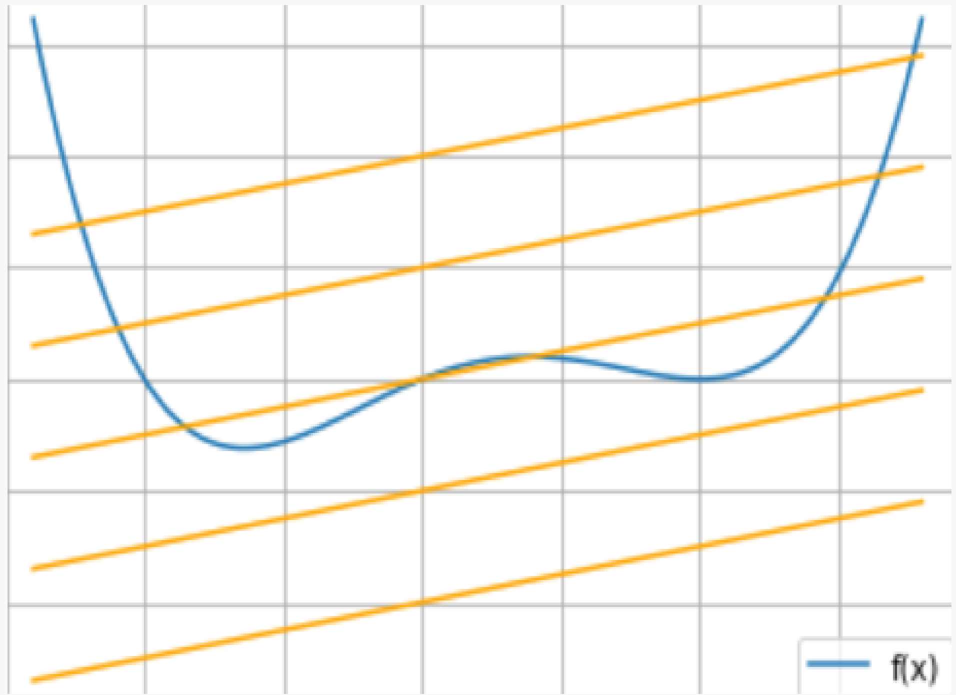
Conjugate Function 이란?

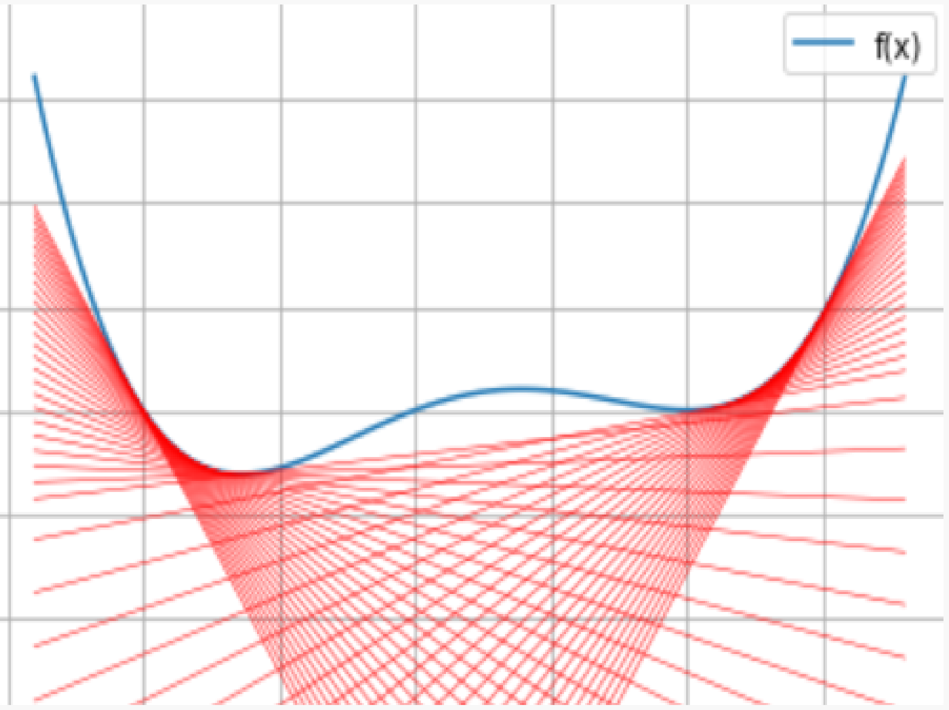
$$f^*(y) = \sup_{x \in \mathbf{dom}\, f} \left(y^T x - f(x) \right)$$

- sup? Supremum : Upper bound 중 가장 작은 값
- 중요한 것은 꺾이지 않는 Convexity (f와 관계 없이 convex)

  → recall affine function

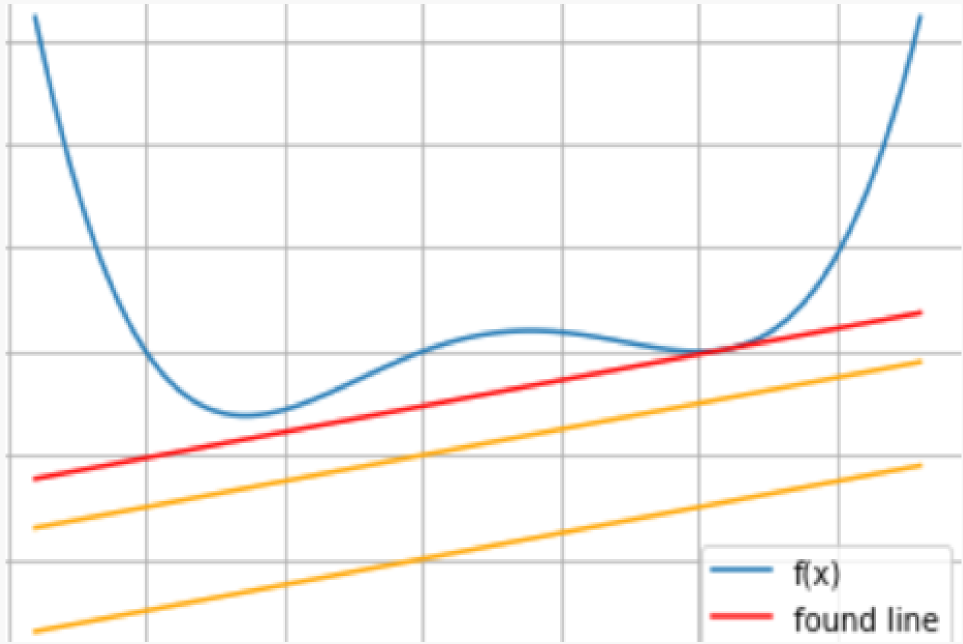
그래서 Conjugate Function 이란..???

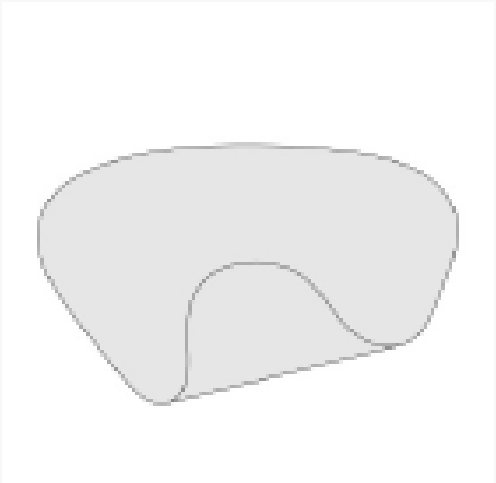




→ f와 관계없이
convex한 이유..?

To be Continued..





Convex set

https://github.com/bikestra/bikestra.github.com/blob/master/notebooks/Convex%20Conjugates.ipynb

Examples

Example 3.21

- *Affine function.* $f(x) = ax + b$. As a function of x, $yx - ax - b$ is bounded if and only if $y = a$, in which case it is constant. Therefore the domain of the conjugate function f^* is the singleton $\{a\}$, and $f^*(a) = -b$.

Examples

Example 3.21

- *Negative logarithm.* $f(x) = -\log x$, with $\mathbf{dom}\, f = \mathbf{R}_{++}$. The function $xy + \log x$ is unbounded above if $y \geq 0$ and reaches its maximum at $x = -1/y$ otherwise. Therefore, $\mathbf{dom}\, f^* = \{y \mid y < 0\} = -\mathbf{R}_{++}$ and $f^*(y) = -\log(-y) - 1$ for $y < 0$.

Examples

Example 3.21

- *Inverse.* $f(x) = 1/x$ on $\mathbf{R}_{++}$. For $y > 0$, $yx - 1/x$ is unbounded above. For $y = 0$ this function has supremum 0; for $y < 0$ the supremum is attained at $x = (-y)^{-1/2}$. Therefore we have $f^*(y) = -2(-y)^{1/2}$, with $\mathbf{dom}\, f^* = -\mathbf{R}_+$.

Examples

Example 3.22

Example 3.22 *Strictly convex quadratic function.* Consider $f(x) = \frac{1}{2}x^T Qx$, with $Q \in \mathbf{S}^n_{++}$. The function $y^T x - \frac{1}{2}x^T Qx$ is bounded above as a function of x for all y. It attains its maximum at $x = Q^{-1}y$, so

$$f^*(y) = \frac{1}{2}y^T Q^{-1}y.$$

Conjugate function 관련 용어

- Fenchel's inequality

$$f(x) + f^*(y) \geq x^T y$$

- Conjugate of the conjugate

f가 convex+closed일 때, f** = f

Log-concave and Log-convex functions

ESC

Optimization for Machine Learning

Log-concave (<-> Log-convex) function

- 단순하게, log f가 concave = log-concave (log-convex는 그 반대)

  \+ if $f(x) > 0$ for all $x \in \mathrm{dom}\, f$ and $\log f$ is concave

- Log-convex function은 convex할까?

  YES with option: function = nonnegative (exponential)

Log-concave (<-> Log-convex) function

$$f(\theta x + (1-\theta)y) \geq f(x)^{\theta} f(y)^{1-\theta}.$$

로그 기호 없이 Log-concavity를 보이는 부등식

부등식에 로그를 취하면 log f가 jensen inequality를 만족함을 알 수 있다.
(concave임을 증명)

Example 3.39

- *Affine function.* $f(x) = a^T x + b$ is log-concave on $\{x \mid a^T x + b > 0\}$.

-> Affine composition (3. 2)

- *Powers.* $f(x) = x^a$, on $\mathbf{R}_{++}$, is log-convex for $a \leq 0$, and log-concave for $a \geq 0$.

- *Exponentials.* $f(x) = e^{ax}$ is log-convex and log-concave.

-> Affine function is both convex and concave!

Example 3.40

- MVN Distribution

$$f(x) = \frac{1}{\sqrt{(2\pi)^n \det \Sigma}} e^{-\frac{1}{2}(x-\bar{x})^T \Sigma^{-1} (x-\bar{x})}$$

- 많은 probability distribution의 density function이 log-concave

`Normal, Uniform, Exponential, Beta, Wishart, Weibull..`

Properties : Twice-differentiable

f가 twice differentiable 하고 dom f가 convex 할 때,

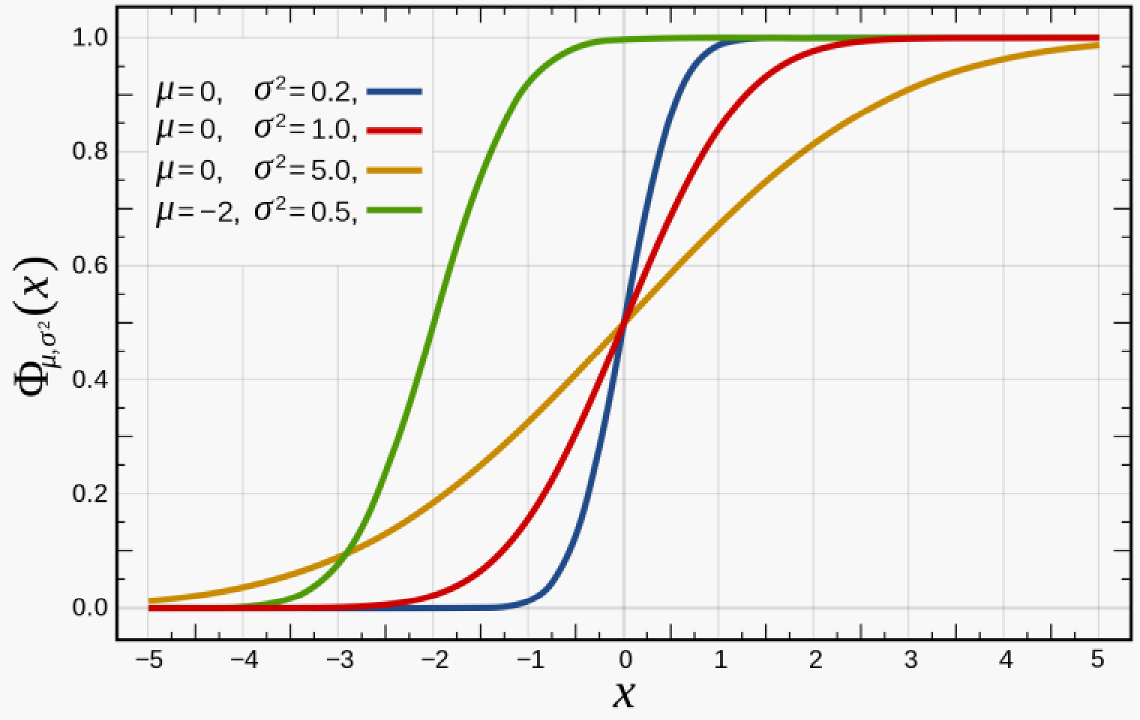
$$\nabla^2 \log f(x) = \frac{1}{f(x)}\nabla^2 f(x) - \frac{1}{f(x)^2}\nabla f(x)\nabla f(x)^T \quad \rightarrow \quad f(x)\nabla^2 f(x) \preceq \nabla f(x)\nabla f(x)^T$$

log-concave

- The cumulative distribution function of a Gaussian density,

$$\Phi(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-u^2/2}\, du,$$

is log-concave (see exercise 3.54).



Properties : Integration

- integration: if $f : \mathbf{R}^n \times \mathbf{R}^m \to \mathbf{R}$ is log-concave, then

$$g(x) = \int f(x, y)\, dy$$

is log-concave (not easy to show)

Properties : Integration

- if $C \subseteq \mathbf{R}^n$ convex and y is a random variable with log-concave pdf then

$$f(x) = \mathbf{prob}(x + y \in C)$$

is log-concave

proof: write $f(x)$ as integral of product of log-concave functions

$$f(x) = \int g(x+y)p(y)\, dy, \qquad g(u) = \begin{cases} 1 & u \in C \\ 0 & u \notin C, \end{cases}$$

Convexity with respect to generalized inequalities

ESC

Optimization for Machine Learning

복습! Generalized inequality

Proper cone K:
cone $K \in R^n$이 다음 성질들을 만족하면 proper cone이라고 부른다.

1. k is convex.

2. k is closed. (경계를 포함하는 집합)

3. k is solid. (interior is not empty)

4. k is pointed (직선을 포함하지 않는다.)

이 proper cone을 이용해 정의된 generalized inequality는 다음과 같다.

$$x \preceq_K y \iff y - x \in K, \quad x \prec_K y \iff y - x \in \mathbf{int}\, K$$

K-Convexity

- K-convex

$$f(\theta x + (1-\theta)y) \preceq_K \theta f(x) + (1-\theta)f(y).$$

부등호 부분만 빼고 기존 convex 부등식과 모두 동일

K-Monotonicity

- $$x \preceq_K y \Longrightarrow f(x) \leq f(y)$$ K-nondecreasing

- $$x \preceq_K y,\ x \neq y \Longrightarrow f(x) < f(y).$$ K-increasing

과제: Conjugate function

3.38 *Young's inequality.* Let $f : \mathbf{R} \to \mathbf{R}$ be an increasing function, with $f(0) = 0$, and let g be its inverse. Define F and G as

$$F(x) = \int_0^x f(a)\,da, \qquad G(y) = \int_0^y g(a)\,da.$$

Show that F and G are conjugates. Give a simple graphical interpretation of Young's inequality,

$$xy \le F(x) + G(y).$$

힌트: 일단 그래프를 그려보세요