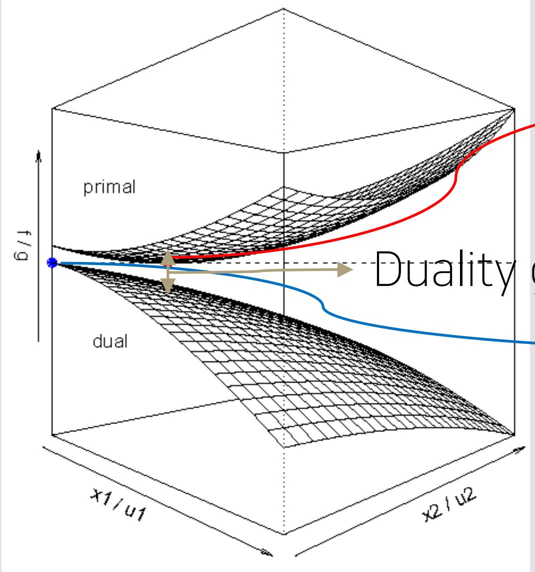
# Duality

# Duality(optimization)

https://en.wikipedia.org/wiki/Duality_(optimization)

- 수학에서 최적화 이론에서 이중성 혹은 이중성 원리는 최적화 문제를 primal problem 혹은 dual problem 의 두 관점에서 볼 수 있다는 원칙.
- Dual problem 의 솔루션은 primal problem 즉, 최소화 문제의 솔루션의 lower bound 를 제공함.
- 하지만, 일반적으로 primal 과 dual problem 의 최적화된 값은 동치가 아니며 이를 duality gap 이라고 부름.
- convex optimization problem 에서, duality gap 은 constraint qualification 조건하에서 0 임.




Duality gap

최적의 primal value

Primal 을 dual 로 바꾸어 풀 수 있음.

최적의 dual value

목표는 조건이 주어진 최소화 문제에서 최대화 문제로 근사하여 최솟값을 찾겠다는 것으로 핵심은 주어진 조건을 포함하는 하나의 식으로 만들어 주는 것임. 최소화 문제에서 최적의 최소값을 찾도록 하는 강력한 lower bound 를 가지는 $g(u)$ 찾는 것이 목표임.

- Dual 함수, $g(u)$ 는 0 이상의 모든 u 에 대하여 $f^*$ 즉, primal 의 최적 값에 대한 lower bound 를 제공함.

<Example of quadratic program in 2D>

# Dual problem(Lagrange dual problem) 의 개요

- 다음과 같이 주어진 선형 조건을 만족시키면서 선형인 목적함수 $c^T x$ 를 최소화 하는 문제(primal problem) 가 주어져 있다고 가정.

라그랑주 승주법에서 만족해야 하는 조건 2가지

$$\min_x \quad c^T x$$

$$Ax = b$$
$$Gx \le h$$

Primal problem

Dual problem

$$\max_{u,v} \quad -ub^T - vh^T$$

$$\text{subject to} \quad (-uA^T - vG^T)^T = c$$
$$v \ge 0$$

양변에 $u^T$와 $v^T$를 곱하여줌. $(v \ge 0)$

$$u^T Ax + v^T Gx \le u^T b + v^T h$$
$$(u^T A + v^T G)x \le u^T b + v^T h$$
$$(uA^T + vG^T)^T x \le u^T b + v^T h$$
$$(-uA^T - vG^T)^T x \ge -u^T b - v^T h$$
$$(-uA^T - vG^T)^T = c \text{ 일 때},$$
$$\therefore c^T x \ge -u^T b - v^T h$$

$c^T x$ Lower bound(하한) 이 $-u^T b - v^T h$ 임.

따라서, 을 최대화 하는 dual problem 으로 primal problem 이 치환 될 수 있음.

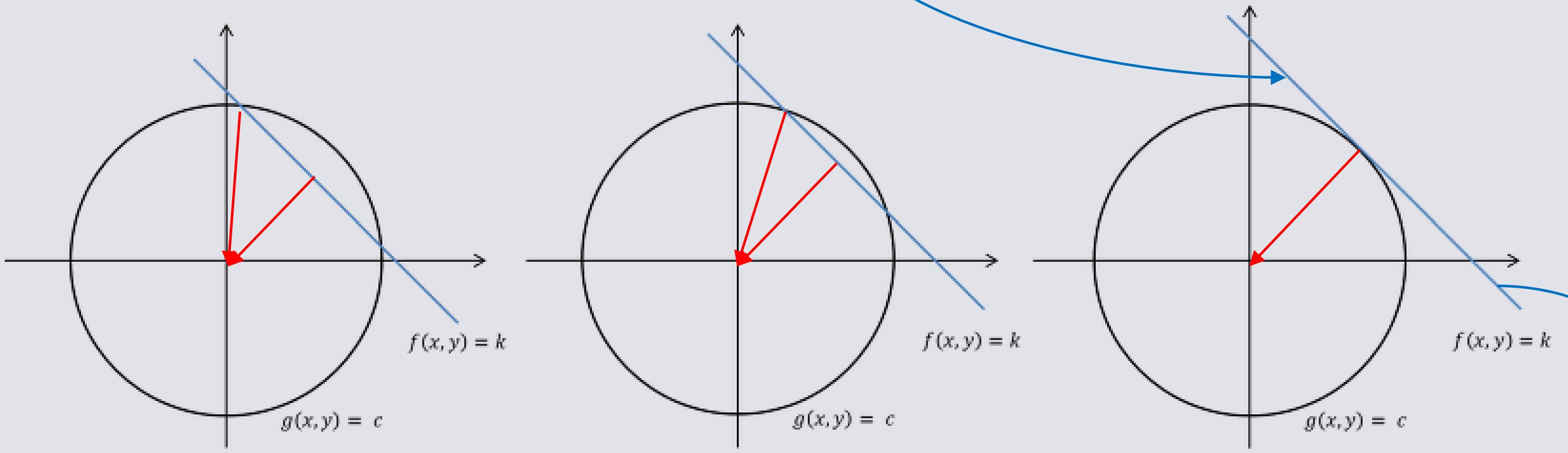
# 라그랑주 승주법

- 라그랑주 승수법이란?
- 프랑스의 수학자 조세프루이 라그랑주 (Joseph-Louis Lagrange)가 **제약 조건이 있는 최적화 문제를 풀기위해 고안한 방법**임.
- 라그랑주 승수법은 어떠한 문제의 최적점을 찾는 것이 아닌 최적점이 되기 위한 조건을 찾는 방법으로 즉, **최적해의 필요조건을 찾는 방법**임.
- 핵심 아이디어
- 최적화 하려는 값에 형식적인 라그랑주 승수를 더하여, 제약된 문제를 제약이 없는 문제로 문제를 접근 하였음.

https://untitledtblog.tistory.com/96

# 라그랑주 승주법의 기하학적인 접근 방식

- 두 함수의 공통 접선이 최대, 최소 문제를 해결할 수 있는 핵심 아이디어가 될 수 있다는 것에서 시작함.
- 즉, 라그랑주 승수법의 기본 가정은 다음과 같음.
- *"제약 조건 g를 만족하는 f 의 최솟값 혹은 최댓값은 f 와 g가 접하는 곳에 존재할 수도 있다."*




불일치 불일치 일치

- 제약조건 $g(x, y) = c$ 와 $f(x, y)$가 접할때 제약 조건을 만족하는 $f(x, y)$ 는 최대가 됨.

이렇게 접하는 지점의 특징은 제약 조건 $g(x, y) = c$ 와 $f(x, y) = k$ 의 접선이 같다는 점임.

즉, 서로의 normal vector 가 평행한 point 임!

https://untitledtblog.tistory.com/96

# 라그랑주 승주법의 기하학적인 접근 방식

- 이러한 관찰을 바탕으로 수식화를 해보면 다음과 같음.

$$\nabla f = \lambda \nabla g$$

두 함수의 gradient vector 가 평행하다!
즉, 내적을 0 이며 scaling factor, $\lambda$ 만을 통해서 서로를 표현 할 수 있음을 의미함.

$$\nabla f - \lambda \nabla g = \nabla L = 0$$ ← (내가 수식화 한 부분이라 맞는지 잘 모르겠음,,,)

$$L(x, y, \lambda) = f(x, y) - \lambda(g(x, y) - c)$$

➡ 정의된 함수 $L$ 의 gradient vector 가 영벡터가 되는 지점을 찾는 문제로 주어진 식 $\nabla f = \lambda \nabla g$ 을 수식화 할 수 있음.

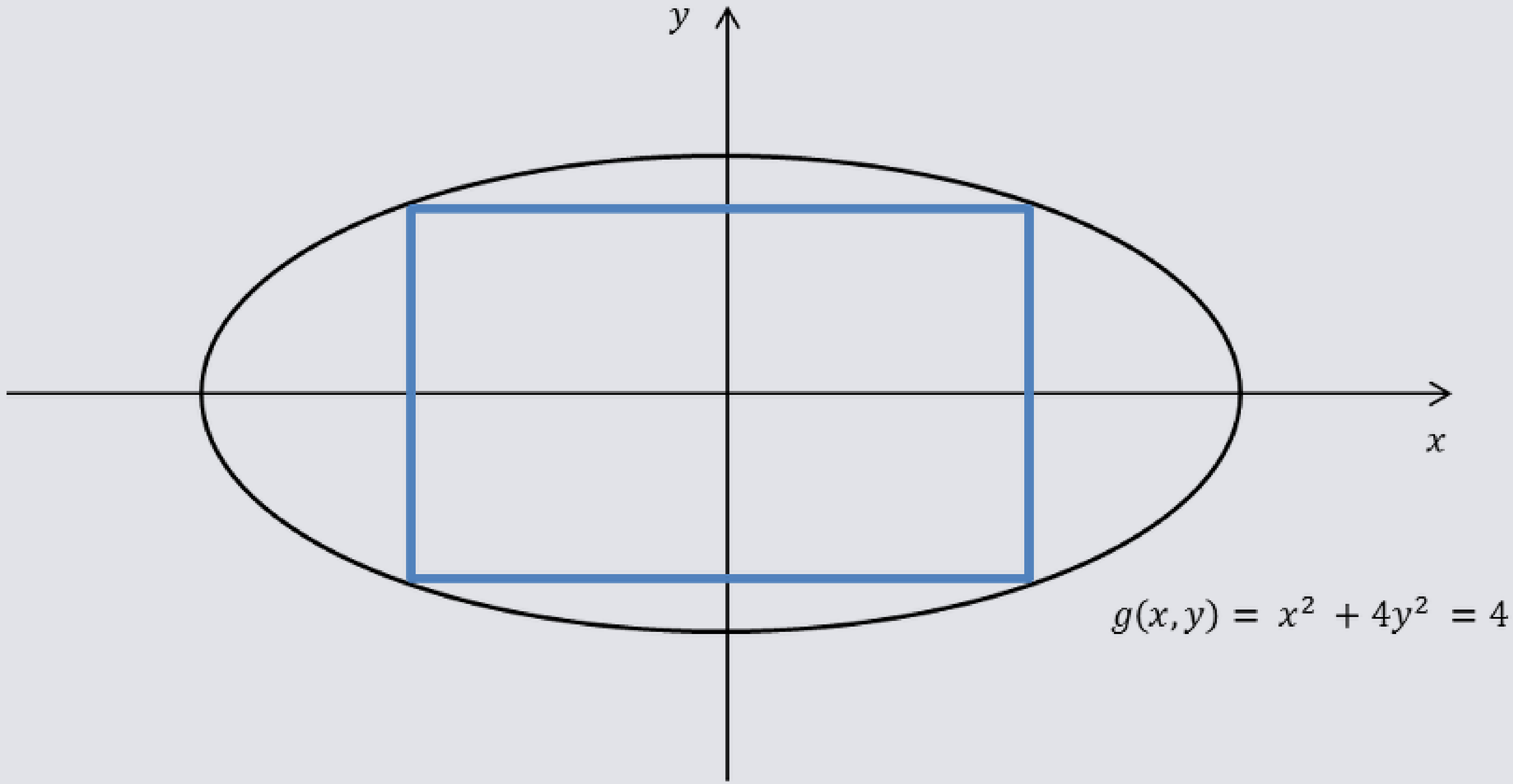
이제, $\boxed{\frac{dL}{dx}, \frac{dL}{dy}}, g(x, y) = c$ 총 제약조건 3개로 함수 $L(x, y, \lambda)$ 를 해결 할 수 있음.

각 $x, y$ 에 대한
편 미분

https://untitledtblog.tistory.com/96 - 라그랑주 승수법의 전미분을 이용한 해석 정리

# 라그랑주 승수법의 예제

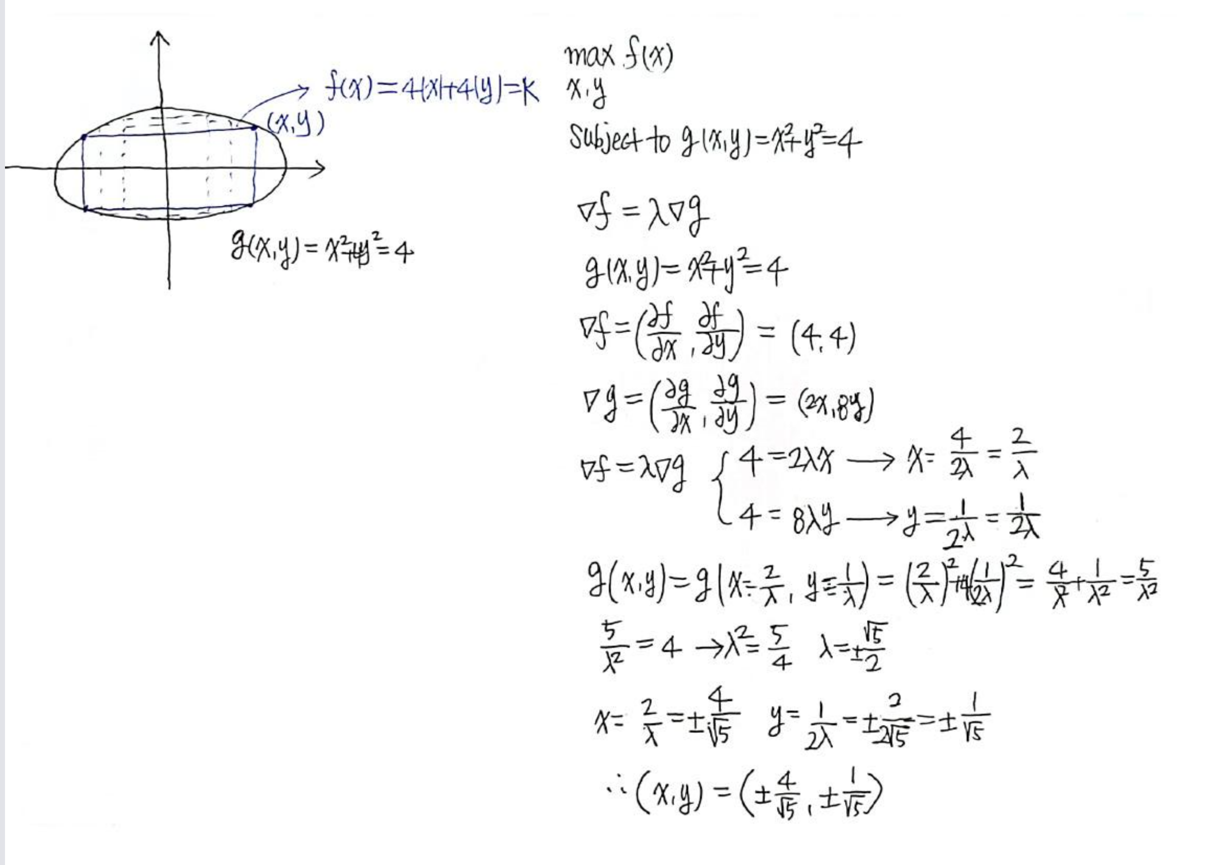
- 그림과 같이 제약 조건 $g(x, y) = x^2 + 4y^2 = 4$ 를 만족하면서, 사각형의 네 변의 합을 최대화 $x, y$ 를 구하라.

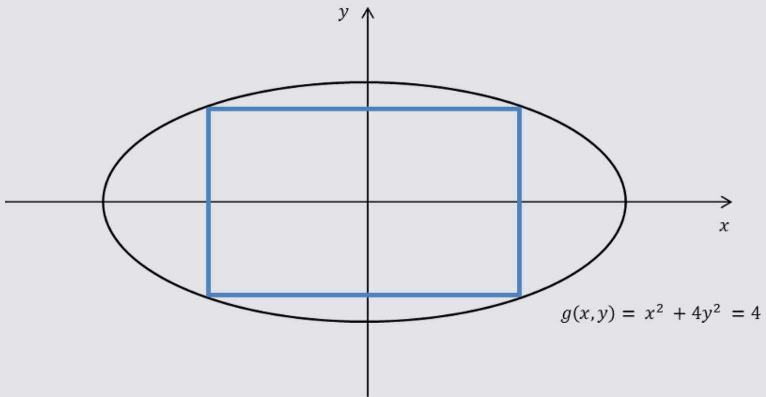



https://untitledtblog.tistory.com/96

# 라그랑주 승수법의 예제

- 그림과 같이 제약 조건 $g(x,y) = x^2 + 4y^2 = 4$ 를 만족하면서, 사각형의 네 변의 합을 최대화 $x, y$ 를 구하라.




https://m.blog.naver.com/mindo1103/90154212128 - 더 많은 예제 확인

# Duality 의 Lagrangian 승주법을 적용한 접근 방식

- 앞에서 언급한 primal problem 을 다시 작성해 보면,

$$\min_x \quad c^T x \quad \longleftarrow \text{ 최소화 하고자 하는 함수 } f$$

$$Ax = b \quad \longleftarrow \text{ 제약조건 } g$$

$$Gx \le h$$

$f^* \ge g(u,v)$ 의 관계로 인하여 primal 의 최소값을 찾는 것은 $g(u,v)$ 즉, 라그랑지 듀얼 함수를 최대화 하는 문제와 동일하게 됨.

$$\max_{u,v} \quad g(u,v)$$

Lagrange dual problem 의 의미 $\quad$ subject to $\quad u \ge 0$

- 라그랑지 듀얼 함수는 $f^*$(primal 함수의 최적의 해) 에 대하여 모든 $v \ge 0, u$ 에 대한 강력한 lower bound $f^* \ge g(u,v)$ 를 제공하며 이는 제약식이 있는 문제를 제약식이 없는 문제로 바꾸어 해당 식의 값을 더 작게 만들 수 있다는 의미가 있음.
- 따라서, 가능한 모든 $u$와 $v$ 에 대하여 $g(u,v)$ 를 최대화 함으로써 가장 좋은 lower bound 를 얻을 수 있음.

$$\nabla f - \lambda \nabla g = \nabla L = 0$$

$$L(x,y,\lambda) = f(x,y) - \lambda(g(x,y) - c)$$

$$= c^T x + u^T(Ax - b) + v^T(Gx - h)$$

$Ax - b = 0(Ax = b$에 의하여)이며 $Gx - h \le 0(Gx \le h)$으로

$u^T(Ax-b) + v^T(Gx-h) \le 0$이 되기 때문에

$$= L(x,y,\lambda) = \boxed{c^T x + u^T(Ax-b) + v^T(Gx-h)} \le \boxed{c^T x}$$

C 를 원초 문제의 제약식을 만족하는 $x$의 집합,

$f^*$ 를 찾으려고 하는 최적의 값이라고 할때,

이를 다시 풀어 수식화하면,

$\therefore f^* \ge \boxed{\min_{x \in C} L(x,u,v)} \ge \boxed{\min_x L(x,u,v)}$ 식이 성립하게 되고

$\min_x L(x,u,v) = g(u,v)$라고 하며

$g(u,v)$ 는 라그랑지 듀얼 함수가 됨.

# 일반적인 최적화 문제에 Lagrangian 접근을 적용

- 아래와 같이 일반적인 primal problem 이 적용되어 있다고 가정.

$$\begin{aligned} &\min_x && f(x) \\ &subject \quad to && h_i(x) \le 0, \quad i = 1, \dots, m \\ & && l_j(x) = 0, \quad j = 1, \dots, r \end{aligned}$$

Dual problem 으로 치환

$$L(x, u, v) = f(x) + \sum_{i=1}^{m} u_i h_i(x) + \sum_{j=1}^{r} v_j l_j(x)$$

$$g(u, v) = \min_x L(x, u, v)$$

라그랑지 함수 $L$

Dual 함수 $g$

<Dual 함수>

- 이때, **dual problem 은** primal problem 이 convex 가 아니어도 **항상 convex optimization problem** 이 됨.
  - Why? Dual 함수 $g(u, v)$ 는 항상 convexity 를 보존하는 연산만 수행되기 때문.
  - 이것이 바로 primal problem 을 dual problem 으로 치환하여 문제를 해결해야 하는 이유라고 이해함.

https://wikidocs.net/20584

$$\begin{aligned} &\max_{u,v} && g(u, v) \\ &subject \quad to && u \ge 0 \end{aligned}$$

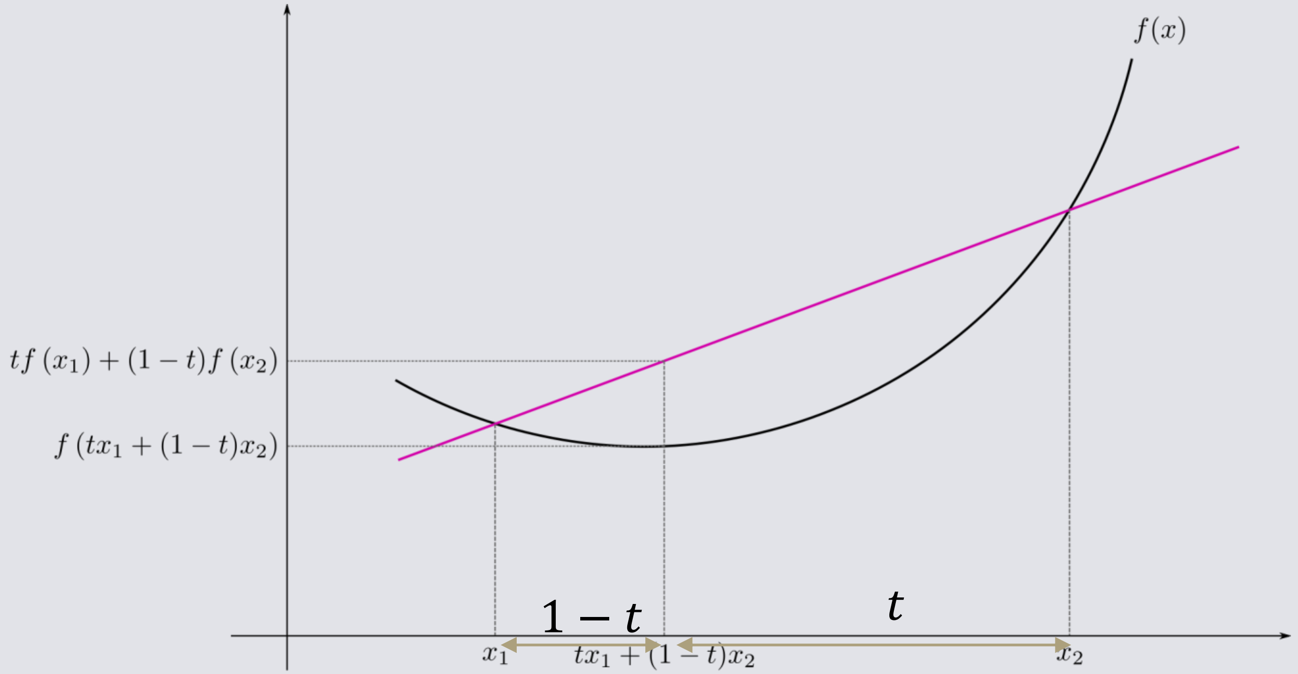
<Dual problem>

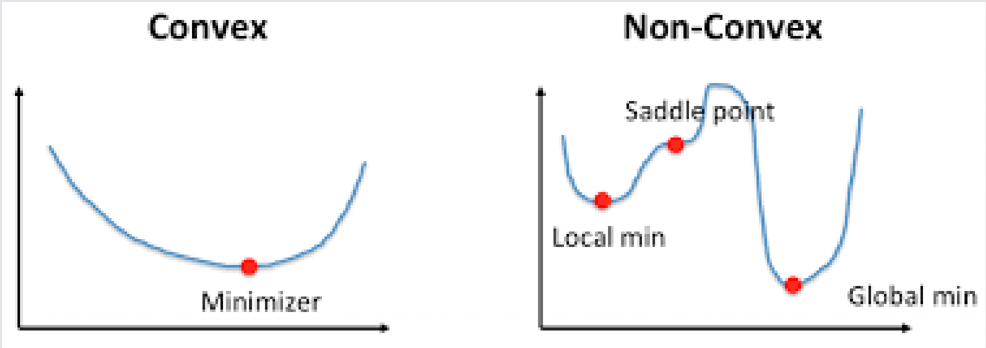
https://ratsgo.github.io/convex%20optimization/2018/01/25/duality/

# convexity

- Convex function 의 정의
- 임의의 두 점 $x, y$ 와 $[0,1]$ 사이의 값 $t$ 에 대해 다음이 항상 성립하는 함수 $f$임.

$$f(tx+(1-t)y) \leq (1-t)f(y)+tf(x)$$




<convexity 가 중요한 이유>

- $t$ 는 $x_2$ 로 부터 의 distance 임.

https://ratsgo.github.io/convex%20optimization/2017/12/26/convexfunction/

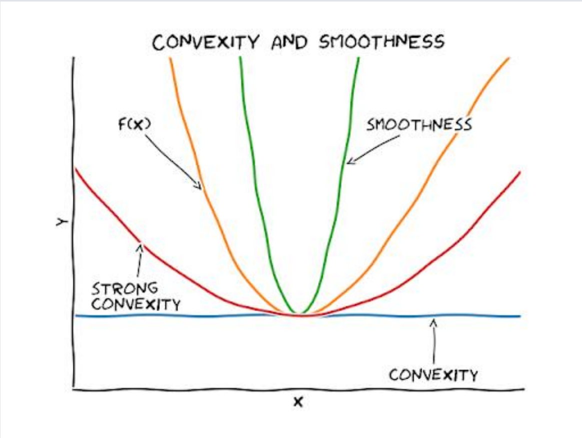
https://wikidocs.net/17495

# Convex function 의 예시

https://ratsgo.github.io/convex%20optimization/2017/12/26/convexfunction/

# Convexity 의 special case

- Convexity 의 대표적인 부분집합(special case) 는 strong convexity 와 strict convexity 가 있음.



# Dual 함수가 convexity 성질을 보존하는 이유 – pointwise maximization

- **convexity**를 보존하는 연산
- *convex function*에 대해 다음 연산은 *convexity*를 보존합니다.

Pointwise maximization: if $f_s$ is convex for any $s \in S$, then $f(x) = \max_{s \in S} \; f_s(x)$ is convex. Note that the set $S$ here (number of functions $f_s$) can be infinite

Another key property: the dual problem is a convex optimization problem (as written, it is a concave maximization problem)

Again, this is always true (even when primal problem is not convex)

By definition:

$$
\begin{aligned}
g(u,v) &= \min_x \left\{ f(x) + \sum_{i=1}^{m} u_i h_i(x) + \sum_{j=1}^{r} v_j \ell_j(x) \right\} \\
&= \underbrace{\max_x \left\{ -f(x) - \sum_{i=1}^{m} u_i h_i(x) - \sum_{j=1}^{r} v_j \ell_j(x) \right\}}_{\text{pointwise maximum of convex functions in } (u,v)}
\end{aligned}
$$

https://ratsgo.github.io/convex%20optimization/2018/01/25/duality/

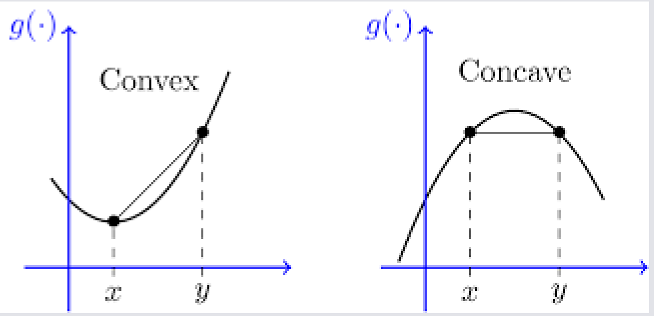
# Dual 함수가 convexity 성질을 보존하는 이유

$$g(u,v) = \min_x \{f(x) + \sum_{i=1}^{m} u_i h_i(x) + \sum_{j=1}^{r} v_j l_j(x)\}$$

$$= \underbrace{-\max_x \{-f(x) - \sum_{i=1}^{m} u_i h_i(x) - \sum_{j=1}^{r} v_j l_j(x)\}}_{\text{pointwise maximum of convex functions in } (u,v)}$$

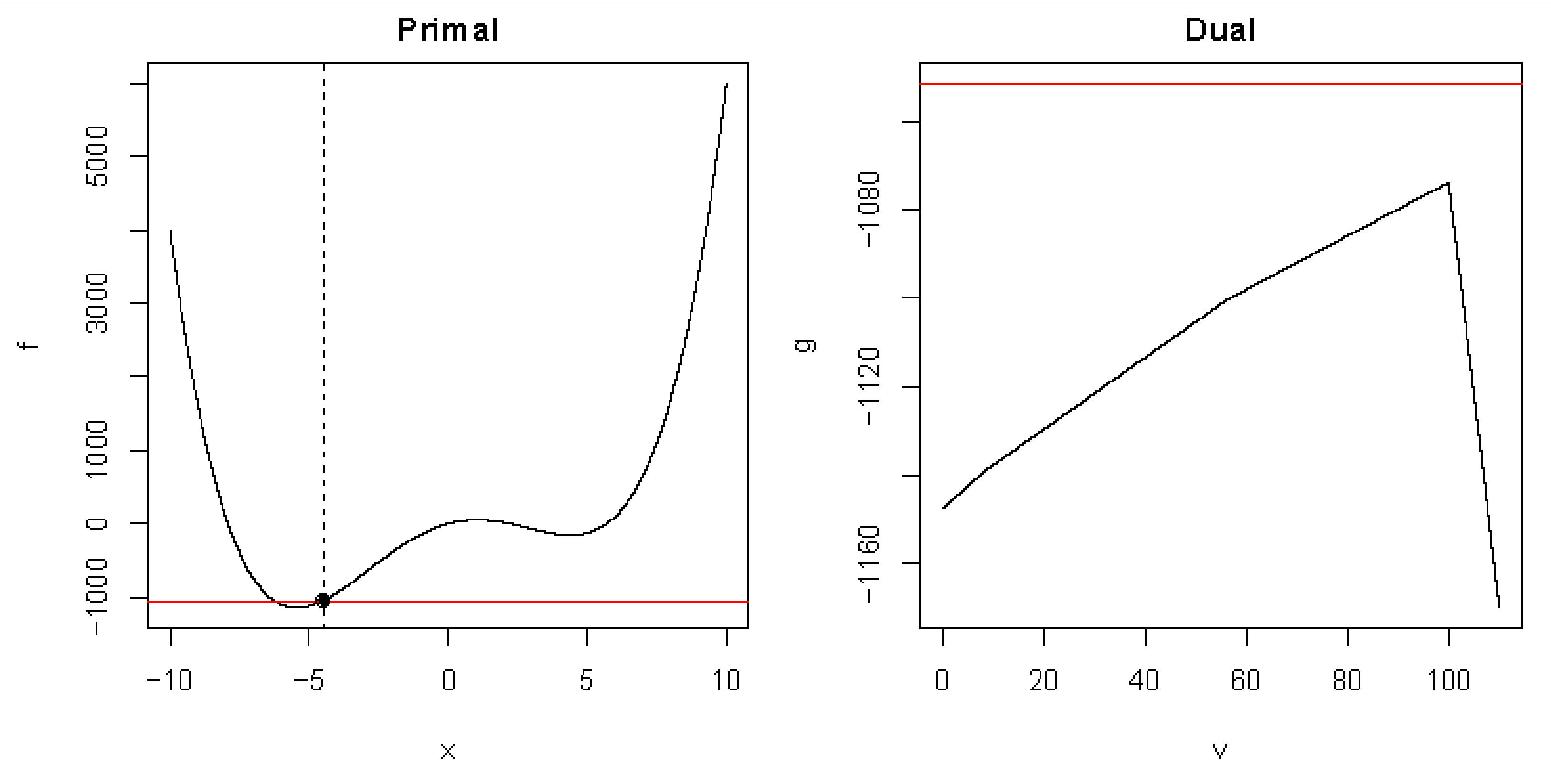
- **concave function(오목함수)**
  란 *convex function*에 음수를 취한 함수를 지칭하는 것으로*convex function*에 집중해서 분석할 수 있음.
- 이로 인하여 dual 문제는 concave maximization 문제에 해당 됨.

Pointwise maximization: if $f_s$ is convex for any $s \in S$, then $f(x) = \max_{s \in S} f_s(x)$ is convex. Note that the set $S$ here (number of functions $f_s$) can be infinite




https://ratsgo.github.io/convex%20optimization/2018/01/25/duality/

# 예제 : Primal problem 이 nonconvex 일 경우




- Nonconvex 를 dual problem 으로 치환 concave 가 되도록 하여 문제를 쉽게 해결 할 수 있음.

https://ratsgo.github.io/convex%20optimization/2018/01/25/duality/

# Duality gap

- Duality 의 목적함수 $g(u, v)$ 를 최대화 하는 문제의 해와 primal problem 을 최소화 하는 문제의 해가 반드시 같지는 않다는 개념을 duality gap 이라고 함.
- $f^* \geq \min_x L(x, u, v) = g(u, v)$

$f^*$ 의 최적의 값과 $g(u, v)$ 의 최적의 값 사이에는 차이가 존재함.

- 이때, 특정 조건을 만족하면 $f^*$ 의 최적의 값과 $g(u, v)$ 의 최적의 값이 동치가 되는데 이 조건이 *Slater's condition 임.*

Given primal problem

$$\begin{aligned} \min_x \quad & f(x) \\ \text{subject to} \quad & h_i(x) \leq 0, \; i = 1, \dots m \\ & \ell_j(x) = 0, \; j = 1, \dots r \end{aligned}$$

Slater's condition: if the primal is a convex problem (i.e., $f$ and $h_1, \dots h_m$ are convex, $\ell_1, \dots \ell_r$ are affine), and there exists at least one strictly feasible $x \in \mathbb{R}^n$, meaning

$$\boxed{h_1(x) < 0, \dots h_m(x) < 0} \quad \text{and} \quad \ell_1(x) = 0, \dots \ell_r(x) = 0$$

then strong duality holds

https://ratsgo.github.io/convex%20optimization/2018/01/25/duality/
https://wikidocs.net/20583
https://pasus.tistory.com/91