MAT3007 Optimization
Lecture 15 Convexity
Lagrangian Relaxation

Yuang Chen

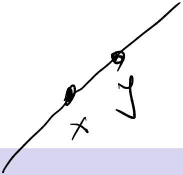
School of Data Scienc
The Chinese University of Hong Kong, Shenzhen

July 9, 2025

Outline

Outline

Affine Set

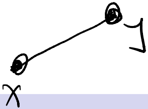


Definition

A set $\mathcal{X}$ is affine if $\forall x, y \in \mathcal{X}$ such that $\lambda x + (1-\lambda)y \in \mathcal{X}, \forall \lambda$.

- Affine set contains the line through any two distinct points in the set.
- Example: solution set of linear equations $\{x | Ax = b\}$. Conversely, every affine set can be expressed as solution set of linear equations

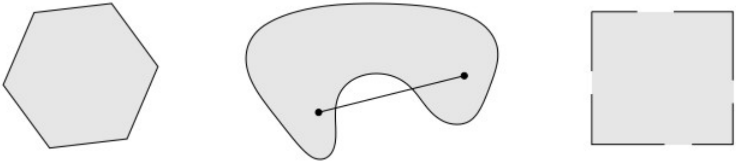
Convex Set



Definition

A set $\mathcal{X}$ is **convex** if $\forall x, y \in \mathcal{X}$ such that $\lambda x + (1-\lambda)y \in \mathcal{X}, \forall \lambda \in [0,1]$.

- Convex set contains line segment between any two points in the set
- Example:

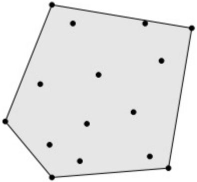


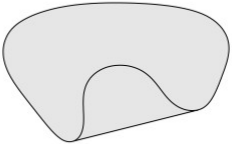
Convex Hull

Definition

Convex hull: set of all convex combinations of elements in the set

$$conv(\mathcal{X}) = \{\sum_i \lambda_i x_i, x_i \in \mathcal{X}, \lambda_i \in [0,1], \sum_i \lambda_i = 1\}$$





The convex hull of the set X is the smallest convex set containing all elements in X.

Convex Set Examples

- **Hyperplane:** set of the form $\{x \mid a^T x = b\}$ where $a \neq 0$
- **Halfspace:** set of the form $\{x \mid a^T x \leq b\}$ where $a \neq 0$
- **(Euclidean) ball** with center x_c and radius r:

$$B(x_c, r) = \{x \mid \|x - x_c\|_2 \leq r\}$$

  where $\|\cdot\|_2$ denotes the Euclidean norm
- **Polyhedron**: set of the form $\{x \mid a^T x \leq b\}$ where $a \neq 0$

Operations that Preserve Convexity I

- **Intersection:** Given a family of convex sets $\{X_a\}_{a \in A}$, the set

$$X = \bigcap_{a \in A} X_a$$

  is convex.

- **Sum:** Given a family of convex sets $\{X_a\}_{a \in A}$, the set

$$X = \left\{ \sum_{a \in A} x_a : x_a \in X_a \ \forall a \in A \right\}$$

  is convex.

- **Product:** If $X_i \subseteq \mathbb{R}^{n_i}$ is a convex set for $i = 1, \ldots, k$, then the set

$$X_1 \times X_2 \times \cdots \times X_k = \left\{ (x_1, \ldots, x_k) \in \mathbb{R}^{n_1 + \cdots + n_k} : x_i \in X_i \ \forall i \right\}$$

  is convex.

Proof of Produc Rule

Select $x \in X_1 \times X_2 \times \cdots \times X_k$

$y \in X_1 \times X_2 \times \cdots \times X_k$

For any $\lambda \in [0, 1]$

$$\lambda x + (1-\lambda) y = \begin{bmatrix} \lambda x_1 + (1-\lambda) y_1 \\ \lambda x_2 + (1-\lambda) y_2 \\ \vdots \\ \lambda x_k + (1-\lambda) y_k \end{bmatrix} \begin{array}{l} \in X_1 \\ \in X_2 \\ \vdots \\ \in X_k \end{array}$$

$$\in X_1 \times X_2 \times \cdots \times X_k$$

Operations that Preserve Convexity II

- **Affine image:** If $X \subseteq \mathbb{R}^n$ is a convex set, $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$, then
$$AX + b = \{Ax + b : x \in X\}$$
is a convex set in $\mathbb{R}^m$.
- **Inverse affine image:** If $Y \subseteq \mathbb{R}^m$ is a convex set, $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$, then
$$S = \{x \in \mathbb{R}^n : \underline{Ax + b \in Y}\}$$
is a convex set. Prove in hw

Proof of Affine Image

select $x, y \in X \rightarrow$ $\begin{matrix} Ax+b \\ Ay+b \end{matrix} \in AX+b$

$\forall \lambda \in [0,1]$

$$\lambda(Ax+b) + (1-\lambda)(Ay+b)$$

$$= \underbrace{\lambda Ax + (1-\lambda)Ay} + \lambda b + (1-\lambda) b$$

$$= A \underbrace{(\lambda x + (1-\lambda) y)}_{\in X} + b$$

$$\in AX + b$$

Outline

Convex Function

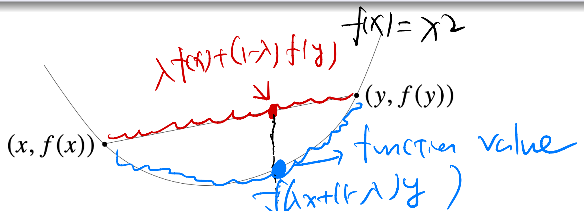
Definition

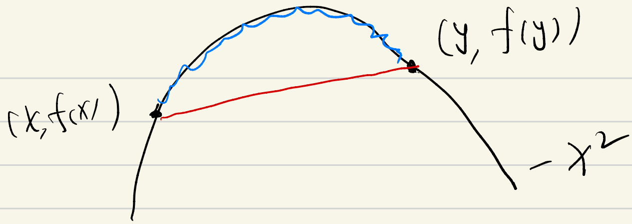
A function $f : \mathbb{R}^n \to \mathbb{R}$ is a **convex** function iff the $\text{dom}(f)$ is a convex set and

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda) f(y)$$

for all $x, y \in \text{dom}(f)$ and $\lambda \in [0,1]$.
f is **concave** if $-f$ is convex.
f is **strictly** convex iff the $\text{dom}(f)$ is a convex set and

$$f(\lambda x + (1-\lambda)y) < \lambda f(x) + (1-\lambda) f(y)$$





$$f(\lambda x + (1-\lambda) y) \geq \lambda f(x) + (1-\lambda) f(y)$$

$\Rightarrow$ Concave.

Thm: $f(\lambda x + (1-\lambda) y) = \lambda f(x) + (1-\lambda) f(y)$

↳ linear function

A Function is both convex and concave

$\Updownarrow$

Linear Function

Remark:

A Function is either convex or concave or both.

↳ WRONG!

Non-Convex Function: not convex and not concave

A Set is either convex or non convex. → T

Convex Function Examples

- affine: $ax + b$ on $\mathbb{R}$, for any $a, b \in \mathbb{R}$
- exponential: e^{ax}, for any $a \in \mathbb{R}$
- powers: x^α on $x > 0$, for $\alpha \geq 1$ or $\alpha \leq 0$
- powers of absolute value: $|x|^p$ on $\mathbb{R}$, for $p \geq 1$
- negative entropy: $x \log x$ on $x > 0$
- norms: $\|x\|_p = (\sum_{i=1}^n |x_i|^p)^{1/p}$ for $p \geq 1$; $\|x\|_\infty = \max_k |x_k|$

Concave

- affine: $ax + b$ on $\mathbb{R}$, for any $a, b \in \mathbb{R}$
- powers: x^α on $x > 0$, for $0 \leq \alpha \leq 1$ $\sqrt{x}$
- logarithm: $\log x$ on $x > 0$

Statement: $\|x\|_p$ is a convex function if $p \geq 1$

Proof for $1 \leq p < \infty$, $f(x) = \|x\|_p$

$\forall x, y, \quad \forall \lambda \in [0, 1]$

$$f(\lambda x + (1-\lambda) y) = \|\lambda x + (1-\lambda) y\|_p$$

$$\leq \|\lambda x\|_p + \|(1-\lambda) y\|_p \quad \text{triangle inequality}$$

$$= \lambda \|x\|_p + (1-\lambda) \|y\|_p$$

$$= \lambda f(x) + (1-\lambda) f(y)$$

$$\|x\|_\infty = \max_i |x_i|$$

$x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ Prove $f(x) = \max_i \{x_i\}$ is convex

$\forall x, y, \quad \forall \lambda \in [0, 1]$

$$f(\lambda x + (1-\lambda) y) = \max_i \{\lambda x_i + (1-\lambda) y_i\}$$

$$\leq \lambda \max_i \{x_i\} + (1-\lambda) \max \{y_i\}$$

$$= \lambda f(x) + (1-\lambda) f(y)$$

$\lambda \max_i \{x_i\} + (1-\lambda) \max_i \{y_i\}$

$\geq \lambda x_i + (1-\lambda) y_i$

select i s.t. $\lambda x_i + (1-\lambda) y_i$ is max

Q: $f(x) = \|x\|_0 = $ # of nonzero elements in x is convex $\rightarrow$ WRONG

Quiz: $f(x) = [x]_2$

$= 2^{nd}$ largest element of x

$f(x)$ is nonconvex.

How to prove a function is not convex?

$\exists x, y, \exists \lambda \in [0,1]$ s.t.

$f(\lambda x + (1-\lambda) y) > \lambda f(x) + (1-\lambda) f(y)$

$x = \begin{bmatrix} 10 \\ 2 \end{bmatrix} \quad y = \begin{bmatrix} 2 \\ 10 \end{bmatrix} \quad \lambda = \frac{1}{2}$

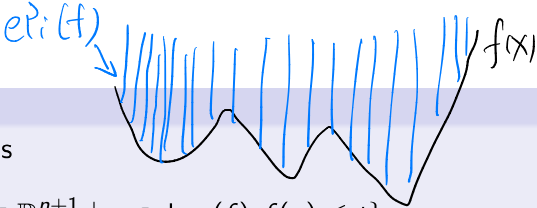
$f(x) = 2 \quad f(y) = 2$

$f(\lambda x + (1-\lambda) y) = f\left(\frac{1}{2}\begin{bmatrix} 10 \\ 2 \end{bmatrix} + \frac{1}{2}\begin{bmatrix} 2 \\ 10 \end{bmatrix}\right) = f\left(\begin{bmatrix} 6 \\ 6 \end{bmatrix}\right) = 6$

$\lambda f(x) + (1-\lambda) f(y) = \frac{1}{2} \times 2 + \frac{1}{2} \times 2 = 2$

So $f(x)$ is not convex.

Remark: $f(x)$ is not concave.



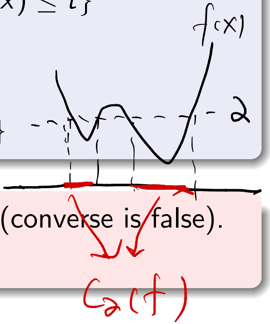
Definition

The **epigraph** of $f : \mathbb{R}^n \to \mathbb{R}$ is

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

The α-**sublevel set** of $f : \mathbb{R}^n \to \mathbb{R}$ is

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



- The sublevel sets of convex functions are convex (converse is false).
- f is convex iff $epi(f)$ is a convex set.

function   set

i) f is convex $\Rightarrow$ epi(f) is convex

Proof by contradiction

Suppose epi(f) is not a convex set.

$\exists \ (x, t_x)$ and $(y, t_y) \in$ epi(f)

$\exists \ \lambda \in [0,1]$

s.t. $(\lambda x + (1-\lambda) y, \ \lambda t_x + (1-\lambda) t_y) \notin$ epi(f)

$$\Rightarrow f(\lambda x + (1-\lambda) y) > \lambda t_x + (1-\lambda) t_y$$

$$\geq \lambda f(x) + (1-\lambda) f(y)$$

$\Rightarrow f(x)$ is not convex, contradiction!

ii) epi(f) is convex $\Rightarrow f$ is convex

Suppose f is not convex function

$\exists \ x, y, \ \exists \ \lambda \in [0,1]$

s.t. $f(\lambda x + (1-\lambda) y) > \lambda f(x) + (1-\lambda) f(y)$ ⓐ

Since $(x, f(x))$ and $(y, f(y)) \in$ epi(f)

and epi(f) is convex, then we have

$(\lambda x + (1-\lambda) y, \ \lambda f(x) + (1-\lambda) f(y)) \in$ epi(f) ⓑ

ⓐ & ⓑ are contradiction!

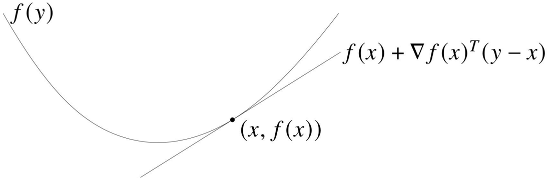
First-order Condition of Convex Functions

First-order condition of convex functions

A differentiable function $f : \mathbb{R}^n \to \mathbb{R}$ with convex domain is convex iff

$$f(y) \geq \underbrace{f(x) + \nabla f(x)^\mathsf{T}(y - x)}_{\text{tangent line}}$$

for all $x, y \in \mathrm{dom}(f)$.



Second-order Condition of Convex Functions

Second-order condition of convex functions

A twice differentiable function $f : \mathbb{R}^n \to \mathbb{R}$ with convex domain is convex iff

$$\nabla^2 f(x) \succcurlyeq 0 \quad \text{(the Hessian is positive semi-definite)}$$

$f''(x) \geq 0 \ , \forall x$

for all $x \in \mathrm{dom}(f)$.

Positive Semi-definite

A $n \times n$ matrix A is positive semi-definite if $x^T Ax \geq 0$ for all $x \in \mathbb{R}^n$. Equivalently, A is positive semi-definite if all its eigenvalues are nonnegative.

Remark: Concave function has a negative semi-definite Hessian.

Examples of Using First and Second-order Conditions

$\nabla f(x) = Px + q$

$\nabla^2 f(x) = P$

$f(x)$ is convex if $P \succeq 0$

- Quadratic function: $f(x) = (1/2)x^\intercal Px + q^\intercal x + r$
- Least squares objective: $f(x) = ||Ax - b||_2^2$

$\nabla f(x) = 2A^T(Ax-b)$

$\nabla^2 f(x) = 2A^TA \succeq 0$

$f(x)$ is convex for any A.

Lemma: Sum Rule nonnegative weighted sum

If $a_1, \ldots, a_m \geq 0$, and $f_1, \ldots, f_m$ are convex (concave) functions, then $f = a_1 f_1 + \cdots + a_m f_m$ is a convex (concave) function.

- Nonnegative linear combination preserves convexity (concavity).
- Examples: $x_1^2 + x_2^2$, $e^x + |x|$.

Lemma: Composition with Linear Functions

If $g : \mathbb{R}^m \to \mathbb{R}$ is convex (concave) and $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$ are given, then $f(x) := g(Ax + b)$ is convex (concave).

- Examples: e^{2x+3} (convex), $(x_1 - x_2)^2 + (x_2 + x_3)^2$ (convex), $\|Ax - b\|$ (convex), $\log(-2x_1 + 3x_2 + 5)$ (concave).

Proof: $\forall x, y$, $\forall \lambda \in [0,1]$

$$f(\lambda x + (1-\lambda) y) \underset{\text{def of } f}{=} g(A(\lambda x + (1-\lambda) y) + b)$$

$$= g(\lambda A x + (1-\lambda) A y + b)$$

$b = \lambda b + (1-\lambda) b$

$$= g(\lambda A x + \lambda b + (1-\lambda) A y + (1-\lambda) b)$$

$$= g(\lambda (Ax+b) + (1-\lambda)(Ay+b))$$

$$\overset{g \text{ convex}}{\leq} \lambda g(Ax+b) + (1-\lambda) g(Ay+b)$$

$$= \lambda f(x) + (1-\lambda) f(y)$$

Convex Function Calculus II

Lemma: Pointwise Maximum

If $f_1, \ldots, f_m$ are convex functions, then

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

is a convex function (this can be extended to uncountably many).

- Examples: $|x| = \max\{-x, x\}$, $\max_i\{a_i^\top \mathbf{x} + b_i\}$, $\|\mathbf{x}\|_1 = \sum_{i=1}^n |x_i|$.

Lemma: Pointwise Minimum

If $f_1, \ldots, f_m$ are concave functions, then

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

is a concave function (this can be extended to uncountably many).

- Examples: $\underline{-|x|} = \min\{-x, x\}$, $\min_i\{a_i^\top \mathbf{x} + b_i\}$.

$g(x_1, \cdots, x_m)$

$x_1 \downarrow \to g \uparrow$

$x_2 \downarrow \to g \uparrow$

$\vdots$

Lemma: Composition with Convex and Nondecreasing Function

If $g : \mathbb{R}^m \to \mathbb{R}$ is convex and componentwise nondecreasing, and each $f_i : \mathbb{R}^n \to \mathbb{R}$ is convex, then

$$f(x) = g(f_1(x), \ldots, f_m(x))$$

is a convex function.

- Examples: $f(x) = e^{\sum_{i=1}^m |a_i x - b_i|}$.

$$f(x) = e^{\sum_i |a_i x - b_i|}$$

$l_i(x) = a_i x - b_i$ → linear → convex

$h_i(x) = |a_i x - b_i|$ → convex by Max Rule

$= |l_i(x)|$

$g(x) = \sum_i h_i(x) = \sum_i |a_i x - b_i|$ → convex by Sum Rule

$f(x) = e^{g(x)} = m(g(x))$ → Convex by Composition Rule

$m(x) = e^x$ convex & non-decreasing

Remark:

$f(x,y)$ is convex (in x and y)

e.g. $f(x,y) = x^2 + y^2$

$g(x,y)$ is convex in x and concave in y.

e.g. $g(x,y) = x^2 - y^2$

Convex Function in ML Examples

1 - Linear Regression

$$\min_{\beta \in \mathbb{R}^n} \quad \sum_{i=1}^{m} \left(x_i^\top \beta - y_i \right)^2 = \|X\beta - \mathbf{y}\|_2^2$$

2 - Robust Regression

$$\min_{\beta \in \mathbb{R}^n} \quad \sum_{i=1}^{m} \left| x_i^\top \beta - y_i \right| = \|X\beta - \mathbf{y}\|_1$$

3 - Logistic Regression

$$\min_{w \in \mathbb{R}^n} \quad \frac{1}{m} \sum_{i=1}^{m} \log \left(1 + \exp(\underbrace{-y_i \cdot \mathbf{w}^\top x_i}_{\text{linear}}) \right)$$

1. $f(\beta) = \|X\beta - y\|_2^2$

$$\nabla f(\beta) = 2X^T(X\beta - y)$$

$$\nabla^2 f(\beta) = 2X^T X \succeq 0$$

2. $f(\beta) = \|X\beta - y\|_1$

$$= \underbrace{\sum_i \underbrace{| \underbrace{x_i^T \beta - y_i}_{\text{linear}} |}_{\text{convex}}}_{\text{convex}}$$

3. $g(z) = \log(1 + \exp(z))$

is convex z

$$g''(z) = \dots \succeq 0 \text{ in HW}$$

Outline

Convex optimization

max. $g(x)$ → Concave
s.t. $x \in X$ → Convex set

$$\begin{aligned} \min . \quad & f(x) \to \text{Convex function} \\ \text{s.t.} \quad & x \in X \to \text{Convex set} \end{aligned}$$

If f is a convex function and X is a convex set, then the problem is a convex optimization problem. (Minimize a convex function is equivalent to maximize a concave function.) Otherwise, we call it non-convex optimization problem

Examples:

- o $\min\{x^2 : -1 \leq x \leq 1\}$ is a convex opt. problem.
- o $\min\{-x^2 : -1 \leq x \leq 1\}$ is not a convex opt. problem.
- o $\min\{x^2 : -5 \leq x \leq 5, x \in \mathbb{Z}\}$ is not a convex opt. problem.

Constraint Types

What constraints would make the feasible region convex?

$l(x)=0$

$l(x)\geq 0$ $l(x)\leq 0$

both convex and concave

$l(x)$ linear

Convex Constraint

Let f be a convex (concave) function. Then for any number c, the set $\Gamma_c = \{x : f(x) \leq (\geq) c\}$ is a convex set.

Therefore,

- If we have constraint $g(x) \leq 0$, and $g(x)$ is convex, then this is a convex constraint.
- If we have constraint $g(x) \geq 0$, and $g(x)$ is concave, then this is a convex constraint.
- Linear constraints are always convex constraints.
- Sometimes, even if a constraint doesn't appear to be in the above form, it still could be a convex constraint.

Being able to identify convex problem is an important skill.

Recognizing Convex Optimization Problem

$$\begin{aligned} \min. \quad & f(x) \\ \text{s.t.} \quad & x \in X \end{aligned}$$

Suppose X is in the following format

$$X = \{\underbrace{g_i(x) \le b_i \ \ \forall i \in I, \quad h_j(x) = d_j \ \ \forall j \in J\}},$$

if f is convex, g_i is convex for all $i \in I$ and h_j is affine for all $j \in J$, then the problem is a convex optimization problem.

Checking Convexity

- Check that all variables are continuous
- Check that the objective function is convex
- Check each equality constraint to see if it is linear
- Write each constraint as an inequality constraint in $\leq$ form with a constant on the right-hand-side, and check the convexity of the function on the left-hand-side

If it passes all the checks then you have a convex optimization problem. Otherwise, it may or may not be convex (the conditions are sufficient, not necessary).

Convex Optimization Example

$$
\begin{aligned}
\min. \quad & x_1^2 + x_2^2 \\
\text{s.t.} \quad & \frac{x_1}{1 + x_2^2} \leq 0 \\
& (x_1 + x_2)^2 = 0
\end{aligned}
$$

the problem is equivalent to

$$
\begin{aligned}
\min. \quad & x_1^2 + x_2^2 \\
\text{s.t.} \quad & x_1 \leq 0 \\
& x_1 + x_2 = 0
\end{aligned}
$$

Example

$$\begin{aligned} \max. \quad & xyz \\ \text{s.t.} \quad & z^2 - xy \le 0 \\ & x, y, z \ge 0 \end{aligned}$$

Convex Optimization: Local is Global

$$\text{(P)}: \quad \min.\{f(x) : \; x \in X\}$$

Theorem

If (P) is a convex optimization problem then a solution $x^* \in X$ is a local optimal solution of (P) iff it is a global optimal solution.

The ‘Easy’ and ‘Difficult’ Optimization Problems

- Linear v.s. nonlinear?
- Differentiable v.s. nondifferentiable?

Classify whether a problem is hard or easy: **Convex (easy)** v.s. **nonconvex (hard)**.

- **Convex optimization:** Benign global geometry and hence reasonable algorithms can almost always find the global minimum.
- **Nonconvex optimization:** Can be anything (e.g., the *saddle point*). In general, finding a stationary point is NP-hard.

Outline

Relaxation

$$(P): \min_x \{f(x) : x \in X\} \qquad (Q): \min_x \{g(x) : x \in Y\}$$

Problem (Q) is a relaxation of (P) if:

- $X \subseteq Y$
- $f(x) \geq g(x) \quad \forall x \in X$

Relaxation Properties

- The relaxation of an optimization problem should be easier to solve.
- The optimal value of the relaxation provides a lower bound on the original problem.
- If the relaxation is infeasible, then clearly the original problem is also infeasible.
- Suppose only the constraints are relaxed. If a solution to the relaxation is feasible for the original problem, then it must be an optimal solution to the original problem.

Lagrangian Relaxation

$$\begin{aligned} \min \quad & f(x) \\ \text{s.t.} \quad & g_i(x) \leq b_i, \quad \forall i \in I \\ & h_j(x) = d_j, \quad \forall j \in J \end{aligned}$$

The Lagrangian function is

$$L(x, \lambda, \mu) = f(x) + \sum_{i \in I} \lambda_i \left[g_i(x) - b_i\right] + \sum_{j \in J} \mu_j \left[h_j(x) - d_j\right]$$

where $\lambda_i \geq 0$ for all $i \in I$ and μ_j for $j \in J$ are Lagrangian multipliers. Then the following problem is the Lagrangian relaxation of (P):

$$(Q) : \min_x L(x, \lambda, \mu)$$

Lagrangian Relaxation

$$v_P = \min\{f(x) : g_i(x) \le b_i,\ \forall i \in I,\ h_j(x) = d_j,\ \forall j \in J\}$$

For $\lambda_i \geq 0$, let:

$$\begin{aligned}
\mathcal{L}(\lambda, \mu) &= \min_x L(x, \lambda, \mu) \\
&= \min_x \left\{ f(x) + \sum_{i \in I} \lambda_i [g_i(x) - b_i] + \sum_{j \in J} \mu_j [h_j(x) - d_j] \right\}.
\end{aligned}$$

Then:

$$\mathcal{L}(\lambda, \mu) \le v_P \quad \forall \lambda \geq 0.$$

Weak duality

Original Problem:

$$(P): \quad v_P = \min_x \{f(x) : g_i(x) \le b_i,\ \forall i \in I,\ h_j(x) = d_j,\ \forall j \in J\}$$

(Lagrangian) Dual Problem:

$$(D): \quad v_D = \max_{\lambda,\mu} \{\mathcal{L}(\lambda, \mu) : \lambda \ge 0\}$$

where

$$\mathcal{L}(\lambda, \mu) = \min_x \left\{ f(x) + \sum_{i \in I} \lambda_i [g_i(x) - b_i] + \sum_{j \in J} \mu_j [h_j(x) - d_j] \right\}.$$

Weak duality: $v_D \le v_P$

Example: two-way partitioning problem

$$\begin{aligned} \min \quad & x^\top W x \\ \text{s.t.} \quad & x_i^2 = 1, \forall i = 1, ..., n \end{aligned}$$

- Feasible region $\{-1, 1\}^n$ contains 2^n discrete points
- Interpretation: partition in two sets, $x_i \in \{-1, 1\}$ is an assignment for item i.