Duality

Yuqian Zhang

Rutgers University

yqz.zhang@rutgers.edu

February 11, 2021

Lagrangian

Standard form problem

$$\begin{aligned} \text{minimize} \quad & f(x) \\ \text{subject to} \quad & f_i(x) \leq 0, \quad i = 1, \cdots, m \\ & h_i(x) = 0, \quad i = 1, \cdots, p \end{aligned}$$

- no convex assumption
- make the optimization problem unconstrained
- trick: move the constraints into the objective

Lagrangian

Lagrangian $L : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^p \to R$ with $\text{dom}(L) = \mathcal{D} \times \mathbb{R}^m \times \mathbb{R}^p$

$$L(x, \lambda, v) = f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x) + \sum_{i=1}^{p} v_i f_i(x)$$

- weighted sum of objective and constraint functions
- λ_i is Lagrange multiplier associated with $f_i(x) \leq 0$
- v_i is Lagrange multiplier associated with $h_i(x) = 0$

Lagrange dual function

Lagrange dual function $g : \mathbb{R}^m \times \mathbb{R}^p \to R$

$$
\begin{aligned}
g(\lambda, v) &= \inf_{x \in \mathcal{D}} L(x, \lambda, v) \\
&= \inf_{x \in \mathcal{D}} f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x) + \sum_{i=1}^{p} v_i f_i(x)
\end{aligned}
$$

Hence, independent of the convexity of f_i's, we have

- $L(x, \lambda, v)$ is linear and convex in λ and v
- $-g(\lambda, v)$ is convex in λ and v
- $g(\lambda, v)$ is concave independent of the convexity of f_i's

Convex function

- **Pointwise maximum**: If $f_1, \cdots, f_m$ are convex, then

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

  is convex.

- **Pointwise supremum**: If $f(x, y)$ is convex in x for each $y \in \mathcal{A}$, then

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

  is convex.

Lower bound property

Lower bound property

If $\lambda_i \geq 0$, then $g(\lambda, v) \leq p^*$.

If $\tilde{x}$ is feasible and $\lambda_i \geq 0$, then

$$f_i(\tilde{x}) \leq 0, \quad h_i(\tilde{x}) = 0.$$

minimizing over all feasible $\tilde{x}$

$$\begin{aligned} g(\lambda, v) &= \inf_{x \in \mathcal{D}} L(x, \lambda, v) \\ &\leq L(\tilde{x}, \lambda, v) \\ &= f_0(\tilde{x}) + \sum \lambda_i f_i(\tilde{x}) + \sum v_i h_i(\tilde{x}) \\ &\leq f_0(\tilde{x}) \end{aligned}$$

Lower bound property

Lower bound property

If $\lambda_i \geq 0$, then $g(\lambda, v) \leq p^*$.

- Primal feasible solution x
- Dual feasible solution (λ, v)
- Every dual feasible point provides a lower bound on the solution of the original problem if $g(\cdot, \cdot)$ is known

Least-norm solution of linear equations

$$
\begin{array}{ll}
\text{minimize} & x^T x \\
\text{subject to} & Ax = b
\end{array}
$$

Dual function

- Lagrangian $L(x, v) = x^T x + v^T (Ax - b)$
- to minimize L over x, set gradient equal to zero

$$\nabla_x L(x, v) = 2x + A^T v = 0 \quad \Longrightarrow \quad x = -\frac{1}{2} A^T v$$

- plug in L to obtain g, a concave function of v

$$g(v) = L((-1/2)A^T v, v) = -\frac{1}{4} v^T A A^T v - b^T v,$$

Least-norm solution of linear equations

$$
\begin{array}{rl}
\text{minimize} & x^T x \\
\text{subject to} & Ax = b
\end{array}
$$

- Lower bound property

$$
-\frac{1}{4} v^T A A^T v - b^T v \leq \inf \left\{ x^T x \mid Ax = b \right\}, \forall v \in \mathbb{R}^m
$$

- pick any v and find a lower bound without solving any optimization

Standard form LP

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & x \geq 0 \\ & Ax = b \end{array}$$

- Lagrangian is affine in x,

$$\begin{aligned} L(x, \lambda, x) &= c^T x + v^T(Ax - b) - \lambda^T x, \\ &= -b^T v + (c + A^T v - \lambda)^T x. \end{aligned}$$

- g is linear on affine domain $\left\{(\lambda, v) \mid A^T v - \lambda + c = 0\right\}$

$$g(\lambda, v) = \inf_x L(x, \lambda, v) = \begin{cases} -b^T v & A^T v - \lambda + c = 0 \\ -\infty & \text{otherwise} \end{cases}$$

- Lower bound property: $p^* \geq -b^T v$ if $A^T v + c \geq 0$

Equality constrained norm minimization

$$\begin{array}{rl} \text{minimize} & \|x\| \\ \text{subject to} & Ax = b \end{array}$$

Dual function

$$\begin{aligned} g(v) &= \inf_x \left(\|x\| - v^T A x + b^T v \right) \\ &= \begin{cases} b^T v & \left\|A^T v\right\|_\star \leq 1 \\ -\infty & \text{otherwise} \end{cases}, \end{aligned}$$

where the dual norm of $\|\cdot\|$ is defined as $\|v\|_\star = \sup_{\|u\| \leq 1} u^T v$, hence

$$\left\|A^T v\right\|_\star \leq 1 \quad \Longrightarrow \quad \sup_x \quad \frac{x^T}{\|x^T\|} A^T v \leq 1$$

Equality constrained norm minimization

$$g(v) = \inf_x \left(\|x\| - v^T A x + b^T v \right) = \begin{cases} b^T v & \left\|A^T v\right\|_\star \leq 1 \\ -\infty & \text{otherwise} \end{cases},$$

Follows from $\inf_x \left(\|x\| - y^T x \right) = 0$ is $\|y\|_\star \leq 1$, $-\infty$ otherwise.

- $\|y\|_\star \leq 1$, then $\|x\| - y^T x \geq 0$ for all x, with equality if $x = 0$
- $\|y\|_\star > 1$, choose $x = tu$ where $\|u\| \leq 1$, $u^T y = \|y\|_\star > 1$:

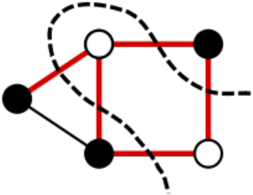
$$\|x\| - y^T x = t \left(\|u\| - \|y\|_\star \right) \rightarrow -\infty, \text{ as} t \rightarrow \infty$$

Lower bound property: $p^\star \geq b^T v$ if $\left\|A^T v\right\|_\star \leq 1$

Max cut

Given a graph G, find a vertix subset S such that the number of edges between S and its complement is large as possible.

$$W_{ij} = \begin{cases} 1 & (i,j) \in G \\ 0 & (i,j) \notin G \end{cases}, \quad x_i = \begin{cases} 1 & i \in S \\ -1 & i \notin S \end{cases}$$



Max cut

Nonconvex formulation

$$\begin{array}{ll} \text{maximize} & \sum_i \sum_j (1 - x_i x_j) W_{ij} \\ \text{subject to} & x_i^2 = 1, \quad i = 1, \cdots, n \end{array}$$

General Formulation

$$\begin{array}{ll} \text{minimize} & x^T W x \\ \text{subject to} & x_i^2 = 1, \quad i = 1, \cdots, n \end{array}$$

Max cut

$$\begin{array}{ll} \text{minimize} & x^T W x \\ \text{subject to} & x_i^2 = 1, \quad i = 1, \cdots, n \end{array}$$

Dual function

$$\begin{aligned} g(v) &= \inf_x \left(x^T W x + \sum_i v_i (x_i^2 - 1) \right) \\ &= \inf_x x^T \left(W + \operatorname{diag}(v) \right) x - \mathbf{1}^T v \\ &= \begin{cases} -\mathbf{1}^T v & W + \operatorname{diag}(v) \succeq 0 \\ -\infty & \text{otherwise} \end{cases} \end{aligned}$$

Lower bound property: $p^\star \geq -\mathbf{1}^T v$ if $W + \operatorname{diag}(v) \succeq 0$

Lagrange dual and conjugate function

$$
\begin{aligned}
\text{minimize} \quad & f_0(x) \\
\text{subject to} \quad & Ax \leq b \\
& Cx = d
\end{aligned}
$$

Dual function

$$
\begin{aligned}
g(\lambda, v) &= \inf_{x \in \mathrm{dom}(f_0)} \left(f_0(x) + (A^T\lambda + C^T v)^T x - b^T\lambda - d^T v \right) \\
&= -f_0^*(-A^T\lambda - C^T v) - b^T\lambda - d^T v
\end{aligned}
$$

- conjugate $f^*(y) = \sup_{x \in \mathrm{dom}(f)} (y^T x - f(x))$
- simplifies derivation of dual if conjugate of f_0 is known

Lagrange dual and conjugate function

Dual problem

$$\begin{array}{ll} \text{maximize} & g(\lambda, v) \\ \text{subject to} & \lambda \geq 0 \end{array}$$

- finds best lower bound on p^*, obtained from Lagrange dual function
- a convex optimization problem; optimal value denoted d^*
- λ, v are dual feasible if $\lambda \geq 0$
- often simplified by making implicit constraint $(\lambda, v) \in \mathrm{dom}(g)$ explicit

Weak and strong duality

Weak duality $d^* \leq p^*$

- always holds (for convex and nonconvex problems)
- can be used to find nontrivial lower bounds for difficult problems
- example, solving the SDP

$$\begin{array}{ll} \text{maximize} & -\mathbf{1}^T v \\ \text{subject to} & W + \mathrm{diag}(v) \succeq 0 \end{array}$$

gives a lower bound for the max cut partitioning problem

Weak and strong duality

Strong duality $d^* = p^*$

- does not hold in general
- (usually) holds for convex problems
- conditions that guarantee strong duality in convex problems are called **constraint qualifications**

Theorem

Strong duality holds for a convex optimization if slater's condition holds.

Slater's constraint qualification

Strong duality holds for a convex problem

$$
\begin{aligned}
\text{minimize} \quad & f_0(x) \\
\text{subject to} \quad & f_i(x) \leq 0, \quad i = 1, \cdots, m \\
& Ax = b
\end{aligned}
$$

if it is strictly feasible

$$
\begin{aligned}
\exists x \in \operatorname{int} \mathcal{D} : \quad & f_i(x) < 0 \quad i = 1, \cdots, m, \\
& Ax = b
\end{aligned}
$$

- also guarantees that the dual optimum is attained (if $p^* > -\infty$)
- there exist many other types of constraint qualifications

Inequality form LP

Primal problems

$$\begin{aligned} \text{minimize} \quad & c^T x \\ \text{subject to} \quad & Ax \leq b \end{aligned}$$

Dual function

$$\begin{aligned} g(\lambda) &= \inf \left((c + A^T\lambda)^T x - b^T\lambda \right) \\ &= \begin{cases} -b^T\lambda & A^T\lambda + c = 0 \\ -\infty & \text{otherwise} \end{cases} \end{aligned}$$

Inequality form LP

Primal problems

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax \leq b \end{array}$$

Dual problem

$$\begin{array}{ll} \text{maximize} & -b^T \lambda \\ \text{subject to} & A^T \lambda + c = 0 \end{array}$$

- from Slater's condition: $p^* = d^*$ if $A\tilde{x} < b$ for some $\tilde{x}$
- in fact, $p^* = d^*$ except when primal and dual are infeasible

Quadratic program

Primal problem, assume $P \in \mathbb{R}^n_{++}$

$$\begin{array}{rl} \text{minimize} & x^T P x \\ \text{subject to} & Ax \leq b \end{array}$$

Dual function

$$\begin{aligned} g(\lambda) &= \inf_x \left(x^T P x + \lambda^T (Ax - b) \right) \\ &= -\frac{1}{4} \lambda^T A P^{-1} A^T \lambda - b^T \lambda \end{aligned}$$

Quadratic program

Primal problem, assume $P \in \mathbb{R}^n_{++}$

$$\begin{aligned} \text{minimize} \quad & x^T P x \\ \text{subject to} \quad & Ax \leq b \end{aligned}$$

Dual problem

$$\begin{aligned} \text{maximize} \quad & -\frac{1}{4}\lambda^T A P^{-1} A^T \lambda - b^T \lambda \\ \text{subject to} \quad & \lambda \geq 0 \end{aligned}$$

- from Slater's condition: $p^* = d^*$ if $A\tilde{x} < b$ for some $\tilde{x}$
- in fact, $p^* = d^*$ always

Next Lecture

Duality