Duality

- Lagrange dual problem
- weak and strong duality
- optimality conditions
- perturbation and sensitivity analysis
- generalized inequalities

Lagrangian

Consider the optimization problem in standard form

$$\begin{aligned} \min \quad & f_0(x) \\ \text{s. t.} \quad & f_i(x) \leq 0, \quad i = 1, \cdots, m \\ & h_i(x) = 0, \quad i = 1, \cdots, p \end{aligned}$$

variable $x \in R^n$, domain $\mathcal{D} = \bigcap_{i=0}^{m} \mathbf{dom}\, f_i \ \cap \ \bigcap_{i=1}^{p} \mathbf{dom}\, h_i$, optimal value $p*$

Lagrangian: $L : R^n \times R^m \times R^p \to R$, with **dom** $L = \mathcal{D} \times R^m \times R^p$

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

- λ_i: Lagrange multiplier associated with $f_i(x) \leq 0$
- ν_i: Lagrange multiplier associated with $h_i(x) = 0$

Lagrangian dual function

Lagrangian dual function: $g : R^m \times R^p \to R$,

$$g(\lambda, \nu) = \inf_{x \in \mathcal{D}} \; L(x, \lambda, \nu)$$
$$= \inf_{x \in \mathcal{D}} \; f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x) + \sum_{i=1}^{p} \nu_i h_i(x)$$

g is concave, can be $-\infty$ for some λ, ν

Lagrangian dual function: lower bound property

Theorem (lower bounds on optimal value)

For any $\lambda \succeq 0$ *and any* ν, *we have*

$$g(\lambda, \nu) \leq p^*$$

Proof.

Suppose $\tilde{x}$ is a feasible point. Since $\lambda \succeq 0$,

$$\sum_{i=1}^{m} \lambda_i f_i(\tilde{x}) + \sum_{i=1}^{p} \nu_i h_i(\tilde{x}) \leq 0$$

Therefore $L(\tilde{x}, \lambda, \nu) \leq f_0(\tilde{x})$. Hence

$$g(\lambda, \nu) = \inf_{x \in \mathcal{D}} L(x, \lambda, \nu) \leq L(\tilde{x}, \lambda, \nu) \leq f_0(\tilde{x})$$

which holds for every feasible $\tilde{x}$. Thus the lower bound follows. □

The dual problem

Lagrange dual problem

$$\begin{array}{ll} \text{max} & g(\lambda, \nu) \\ \text{s. t.} & \lambda \succeq 0 \end{array}$$

- find the best lower bound on p^*
- a convex optimization problem; optimal value denoted d^*
- λ, ν are dual feasible if $\lambda \succeq 0, (\lambda, \nu) \in \textbf{dom}\, g$ (means $g(\lambda, \nu) > -\infty$)
- (λ^*, ν^*): dual optimal multipliers
- $p^* - d^*$ is called the **optimal duality gap**

Weak and strong duality

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

- always true (for both convex and nonconvex problems)

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**

Slater's constraint qualification

Consider the standard convex optimization problem

$$\begin{aligned} \min \quad & f_0(x) \\ \text{s.t.} \quad & f_i(x) \leq 0, \quad i = 1, \cdots, m \\ & Ax = b \end{aligned}$$

with variable $x \in R^n$, domain $\mathcal{D} = \bigcap_{i=0}^{m} \operatorname{dom} f_i$.

Slater's condition: exists a point that is **strictly feasible**, i.e.,

$$\exists x \in \operatorname{relint} \mathcal{D} \text{ such that } f_i(x) < 0, \; i = 1, \cdots, m, \; Ax = b$$

(interior relative to affine hull) can be relaxed: affine inequalities do not need to hold with strict inequalities

Slater's theorem: The strong duality holds if the Slater's condition holds and the problem is convex.

Complementary slackness

Suppose strong duality holds; x^* is primal optimal; (λ^*, ν^*) is dual optimal

$$\begin{aligned} f_0(x^*) = g(\lambda^*, \nu^*) &= \inf_{x \in \mathcal{D}} \; f_0(x) + \sum_{i=1}^{m} \lambda_i^* f_i(x) + \sum_{i=1}^{p} \nu_i^* h_i(x) \\ &\leq f_0(x^*) + \sum_{i=1}^{m} \lambda_i^* f_i(x^*) + \sum_{i=1}^{p} \nu_i^* h_i(x^*) \\ &\leq f_0(x^*) \end{aligned}$$

Hence the inequalities must hold with equality

- x^* minimizes $L(x, \lambda^*, \nu^*)$
- $\lambda_i^* f_i(x^*) = 0$ for all $i = 1, \cdots, m$:

$$\lambda_i^* = 0 \implies f_i(x^*) = 0, \qquad f_i(x^*) < 0 \implies \lambda_i^* = 0$$

known as complementary slackness

Karush-Kuhn-Tucker (KKT) conditions

If strong duality holds, x is primal optimal, (λ, ν) is dual optimal, and f_i, h_i are differentiable, then the following four conditions (called **KKT conditions**) must hold

1. primal constraints: $f_i(x) \leq 0, i = 1, \cdots, m, h_i(x) = 0, i = 1, \cdots, p$
2. dual constraints: $\lambda \succeq 0$
3. complementary slackness: $\lambda_i^* f_i(x^*) = 0$
4. gradient of Lagrangian w.r.t. x vanishes:

$$\nabla f_0(x) + \sum_{i=1}^{m} \lambda_i \nabla f_i(x) + \sum_{i=1}^{p} \nu_i \nabla h_i(x) = 0$$

KKT conditions for convex problem

If $\tilde{x}, \tilde{\lambda}, \tilde{\nu}$ satisfy KKT for a convex problem, then they are optimal:

- from complementary slackness: $f_0(\tilde{x}) = L(\tilde{x}, \tilde{\lambda}, \tilde{\nu})$
- from the 4-th condition and convexity: $g(\tilde{\lambda}, \tilde{\nu}) = L(\tilde{x}, \tilde{\lambda}, \tilde{\nu})$

So $f_0(\tilde{x}) = g(\tilde{\lambda}, \tilde{\nu})$

If the Slater's condition is satisfied and f_i is differentiable, then x is optimal iff $\exists \lambda, \nu$ that satisfy KKT

Minimax interpretation

Given Lagrangian

$$L(x, \lambda, \nu) = f_0(x) + \lambda^T f(x) + \nu^T h(x)$$

The primal problem:

$$\text{(P)} \qquad \inf_{x \in P} \sup_{\lambda \succeq 0, \nu} L(x, \lambda, \nu)$$

The dual problem:

$$\text{(D)} \qquad \sup_{\lambda \succeq 0, \nu} \inf_{x \in P} L(x, \lambda, \nu)$$

Weak duality:

$$\sup_{\lambda \succeq 0, \nu} \inf_{x \in P} L(x, \lambda, \nu) \leq \inf_{x \in P} \sup_{\lambda \succeq 0, \nu} L(x, \lambda, \nu)$$

Saddle point implies strong duality

Strong duality:

$$\sup_{\lambda \succeq 0, \nu} \inf_{x \in P} L(x, \lambda, \nu) = \inf_{x \in P} \sup_{\lambda \succeq 0, \nu} L(x, \lambda, \nu)$$

Saddle-point interpretation: (x^*, λ^*, ν^*) is a saddle point of L if

$$L(x^*, \lambda, \nu) \leq L(x^*, \lambda^*, \nu^*) \leq L(x, \lambda^*, \nu^*)$$

for all $\lambda \succeq 0, \nu, x \in P$.

The strong duality holds if $\exists (x^*, \lambda^*, \nu^*)$ a saddle point of L

Examples

Standard form LP

$$(P) \qquad \begin{aligned} \min \quad & c^T x \\ \text{s.t.} \quad & Ax = b, \quad x \succeq 0 \end{aligned}$$

$$(D) \qquad \begin{aligned} \max \quad & -b^T \nu \\ & A^T \nu + c \succeq 0 \end{aligned}$$

Quadratic program

$$(P) \qquad \begin{aligned} \min \quad & x^T P x \\ \text{s.t.} \quad & Ax \preceq b \end{aligned}$$

(assume $P \in S^n_{++}$)

$$(D) \qquad \begin{aligned} \max \quad & -\frac{1}{4}\lambda^T A P^{-1} A^T \lambda - b^T \lambda \\ \text{s.t.} \quad & \lambda \succeq 0 \end{aligned}$$

Equality constrained norm minimization

$$(P) \qquad \begin{aligned} \min \quad & \|x\| \\ \text{s. t.} \quad & Ax = b \end{aligned}$$

$$(D) \qquad \begin{aligned} \max \quad & b^T \nu \\ & \|A^T \nu\|_* \leq 1 \end{aligned}$$

Note:

- $\|y\|_* = \sup\{x^T y \mid \|x\| \leq 1\}$ is the dual norm of $\|\cdot\|$.
- $\inf_x(\|x\| - y^T x) = 0$ if $\|y\|_* \leq 1$ and $-\infty$ otherwise.

Two-way partitioning

$$(P) \qquad \begin{array}{ll} \min & x^T W x \\ \text{s.t.} & x_i^2 = 1, \quad i = 1, \cdots, n \end{array}$$

- nonconvex; feasible set: 2^n discrete points
- partition $\{1, \cdots, n\}$ into two sets; W_{ij} cost of associating i, j to the same set; $-W_{ij}$ cost of assigning to different sets

$$(D) \qquad \begin{array}{ll} \max & -1^T \nu \\ \text{s.t.} & W + \text{diag}(\nu) \succeq 0 \end{array}$$

- lower bound example: $\nu = -\lambda_{\min}(W)1$ gives $p^* \geq n\lambda_{\min}(W)$

Perturbation and sensitivity anaysis

perturbed optimization problem

$$\begin{aligned} \min \quad & f_0(x) \\ \text{s.t.} \quad & f_i(x) \le u_i, \quad i = 1, \cdots, m \\ & h_i(x) = v_i, \quad i = 1, \cdots, p \end{aligned}$$

its dual

$$\begin{aligned} \min \quad & g(\lambda, \nu) - u^T\lambda - v^T\nu \\ \text{s.t.} \quad & \lambda \succeq 0 \end{aligned}$$

- x primal variable; u, v are parameters
- $p^*(u, v)$ is optimal value as a function of u, v

Global sensitivity anaysis

assume strong duality holds for unperturbed problem ($u = 0, v = 0$), and λ^*, ν^* are dual optimal for unperturbed problem

By weak duality on the perturbed problem:

$$\begin{aligned} p^*(u, v) &\geq g(\lambda^*, \nu^*) - u^T \lambda^* - v^T \nu^* \\ &\geq p^*(0, 0) - u^T \lambda^* - v^T \nu^* \end{aligned}$$

sensitivity interpretation:

- λ_i^* large: small $u_i \implies$ large change in p^*
- ν_i^* large and positive: $v_i < 0 \implies$ large increase in p^*

Local sensitivity analysis: LP

Consider LP

$$(P) \qquad \min \quad c^T x \qquad \text{s. t.} \quad Ax \leq y$$

$$(D) \qquad \max \quad -y^T \lambda \qquad A^T \lambda + c = 0, \quad \lambda \succeq 0$$

The optimal value: $p^* = -y^T \lambda^*$. Thus

$$\frac{\partial p^*}{\partial y_i} \mid_{y_i=0} = -\lambda_i^*$$

Local sensitivity analysis

Suppose strong duality holds for unperturbed problem ($u = 0, v = 0$), and λ^*, ν^* are dual optimal for unperturbed problem. If $p^*(u, v)$ is differentiable at $(0, 0)$, then

$$\lambda_i^* = -\frac{\partial p^*(0,0)}{\partial u_i} \quad \nu_i^* = -\frac{\partial p^*(0,0)}{\partial v_i}$$

If $p^*(u, v)$ is not differentiable at $(0, 0)$, then $(-\lambda^*, -\nu^*) \in \partial p^*(0, 0)$.

Proof.

By the weak duality on the perturbed problem:

$$\frac{\partial p^*(0,0)}{\partial u_i} = \lim_{t \downarrow 0} \frac{p^*(te_i, 0) - p^*(0,0)}{t} \geq -\lambda_i^*$$
$$\frac{\partial p^*(0,0)}{\partial u_i} = \lim_{t \uparrow 0} \frac{p^*(te_i, 0) - p^*(0,0)}{t} \leq -\lambda_i^*$$

Thus equality holds. Similar proof for ν_i^* □

Problems with generalized inequalities

$$\begin{aligned} \min \quad & f_0(x) \\ \text{s.t.} \quad & f_i(x) \succeq_{K_i} 0, \quad i = 1, \cdots, m \\ & h_i(x) = v_i, \quad i = 1, \cdots, p \end{aligned}$$

Lagrangian

- λ_i: Lagrange multiplier for $f_i(x) \preceq_{K_i} 0$
- Lagrangian L:

$$L(x, \lambda_1, \cdots, \lambda_m, \nu) = f_0(x) + \sum_{i=1}^{m} \lambda_i^T f_i(x) + \sum_{i=1}^{p} \nu_i h_i(x)$$

- Dual function g:

$$g(\lambda_1, \cdots, \lambda_m, \nu) = \inf_{x \in \mathcal{D}} L(x, \lambda_1, \cdots, \lambda_m, \nu, x)$$

Problems with generalized inequalities: dual problem

Theorem (lower bound property)

$$\lambda_i \succeq_{K_i^*} 0, \; i = 1, \cdots, m \implies g(\lambda_1, \cdots, \lambda_m, \nu) \leq p^*$$

Lagrange dual problem

$$\begin{array}{ll} \max & g(\lambda_1, \cdots, \lambda_m, \nu) \\ \text{s. t.} & \lambda_i \succeq_{K_i^*} 0, \quad i = 1, \cdots, m \end{array}$$

- weak duality: $p^* \geq d^*$
- $p^* - d^*$: optimal duality gap
- strong duality: $p^* = d^*$ for convex problem with constraint qualification. Slater's: primal problem is strictly feasible

Semidefinite program (SDP)

Primal SDP

$$\begin{aligned} \min \quad & c^T x \\ \text{s. t.} \quad & x_1 F_1 + \cdots + x_n F_n \preceq G \end{aligned}$$

where $F_i, G \in S^k$

Dual SDP

$$\begin{aligned} \max \quad & -\mathbf{tr}(GZ) \\ \text{s. t.} \quad & \mathbf{tr}(F_i Z) + c_i = 0, \quad i = 1, \cdots, n \\ & Z \succeq 0 \end{aligned}$$

where $Z \in S^k$.

Strong duality if primal SDP is strictly feasible, i.e. $\exists x$ with $x_1 F_1 + \cdots + x_n F_n \prec G$