Chapter 5 Duality

Last update on 2025-02-28 17:47:21+08:00

Table of contents

Lagrange dual problem

Weak and strong duality

Geometric interpretation

Optimality conditions

Perturbation and sensitivity analysis

Examples

Theorems of alternatives

Generalized inequalities

Lagrangian

$$\begin{array}{lll} \text{minimize} & f_0(x) & \\ \text{subject to} & f_i(x) \leq 0, & i = 1, \cdots, m \\ & h_i(x) = 0, & i = 1, \cdots, p \end{array}$$

variable $x \in \mathbb{R}^n$, domain $\mathcal{D}$, optimal value p^*

Lagrangian $L\colon \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^p \to \mathbb{R}$ with $\mathbf{dom}\, L = \mathcal{D} \times \mathbb{R}^m \times \mathbb{R}^p$

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

- weighted sum of objective and constraint functions
- λ_i and ν_i are Lagrange multipliers

Lagrange dual function

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

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

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

Lower bound property $g(\lambda, \nu) \leq p^*$ for any $\lambda \succeq 0$

Proof for any feasible $\bar{x}$ and $\lambda \succeq 0$

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

minimizing over all feasible $\bar{x}$ gives $g(\lambda, \nu) \leq p^*$

Least-norm solution of linear equations

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

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

$$\nabla_x L(x, \nu) = 2x + A^T \nu = 0 \qquad \Longrightarrow \qquad x = -(1/2) A^T \nu$$

- dual function (concave in ν)

$$g(\nu) = L\left((-1/2)A^T \nu, \nu\right) = -(1/4)\nu^T A A^T \nu - b^T \nu$$

- lower bound property $\quad p^* \geq -(1/4)\nu^T A A^T \nu - b^T \nu \quad$ for all ν

Standard form LP

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

- Lagrangian

$$L(x, \lambda, \nu) = c^T x + \nu^T (Ax - b) - \lambda^T x = -b^T \nu + (c + A^T \nu - \lambda)^T x$$

- dual function (linear on affine domain hence concave)

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

- lower bound property $\quad p^* \geq -b^T \nu \quad$ if $A^T \nu + c \succeq 0$

Equality constrained norm minimization

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

- Lagrangian $\quad L(x, \nu) = \|x\| - \nu^T(Ax - b) = \|x\| - \nu^T Ax + b^T \nu$
- dual function

$$g(\nu) = \inf_x L(x, \nu) = \begin{cases} b^T \nu & \|A^T \nu\|_* \leq 1 \\ -\infty & \text{otherwise} \end{cases}$$

where $\|v\|_* = \mathbf{sup}_{\|u\| \leq 1} u^T v$ is the dual norm (proof on next page)

- lower bound property $\quad p^* \geq b^T \nu \quad$ if $\|A^T \nu\|_* \leq 1$

Proof

observe that

$$\inf_x \left(\|x\| - y^T x\right) = \begin{cases} 0 & \|y\|_* \le 1 \\ -\infty & \text{otherwise} \end{cases}$$

- if $\|y\|_* \le 1$, then $y^T x \le \|x\|\|y\|_* \le \|x\|$ for all x, with equality if $x = 0$
- if $\|y\|_* > 1$, choose $x = tu$ such that $\|u\| \le 1$ and $y^T u > 1$, then

$$\lim_{t\to\infty} \left(\|x\| - y^T x\right) = t\left(\|u\| - \|y\|_*\right) = -\infty$$

Two-way partitioning problem

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

- nonconvex problem, feasible set contains 2^n discrete points
- $W \in \mathbb{S}^n$, W_{ij} is cost of assigning i and j to the same set
- interpretation: find the most harmonies way to divide $\{1, \cdots, n\}$ in two sets

- Lagrangian

$$L = x^T W x + \sum_{i=1}^{n} \nu_i (x_i^2 - 1)$$

- dual function

$$g(\nu) = \inf_x \left(x^T (W + \mathbf{diag}(\nu)) x - \mathbf{1}^T \nu \right) = \begin{cases} -\mathbf{1}^T \nu & W + \mathbf{diag}(\nu) \succeq 0 \\ -\infty & \text{otherwise} \end{cases}$$

- lower bound property

$$p^\star \geq -\mathbf{1}^T \nu \qquad \text{if } W + \mathbf{diag}(\nu) \succeq 0$$

- example

$$\nu = -\lambda_{\min}(W)\mathbf{1} \quad \text{gives bound} \quad p^\star \geq n\lambda_{\min}(W)$$

Lagrange dual & conjugate function

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & Ax \preceq b \\ & Cx = d \end{array}$$

dual function

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

- recall definition of conjugate $\qquad f^*(y) = \sup_{x \in \mathbf{dom}\, f} \left(y^T x - f(x)\right)$
- simplifies derivation of dual if conjugate of f_0 is known

Entropy maximization

$$\begin{array}{ll} \text{minimize} & f_0(x) = \sum_{i=1}^n x_i \log x_i \\ \text{subject to} & Ax \preceq b \\ & \mathbf{1}^T x = 1 \end{array}$$

- conjugate of $f_0(x)$

$$f_0^*(y) = \sum_{i=1}^n e^{y_i - 1}$$

- dual function

$$g(\lambda, \nu) = -\sum_{i=1}^n e^{-a_i^T \lambda - \nu - 1} - b^T \lambda - \nu = -e^{-\nu - 1} \sum_{i=1}^n e^{-a_i^T \lambda} - b^T \lambda - \nu$$

Lagrange dual problem

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

- finds best lower bound on p^*, obtained from Lagrange dual function
- convex optimization problem, optimal value denoted d^*
- λ and ν are dual feasible if $\lambda \succeq 0$ and $(\lambda, \nu) \in \mathbf{dom}\, g$
- often simplified by making implicit constraint $(\lambda, \nu) \in \mathbf{dom}\, g$ explicit
- original problem is called primal problem

Standard form LP

primal problem

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

dual problem

$$\begin{array}{ll} \text{maximize} & g(\lambda, \nu) = \begin{cases} -b^T \nu & A^T \nu - \lambda + c = 0 \\ -\infty & \text{otherwise} \end{cases} \\ \text{subject to} & \lambda \succeq 0 \end{array}$$

equivalent form, the Lagrange dual of the standard form LP (primal problem)

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

Inequality form LP

primal problem

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

dual problem

$$\begin{array}{ll} \text{maximize} & g(\lambda) = \begin{cases} -b^T\lambda & A^T\lambda + c = 0 \\ -\infty & \text{otherwise} \end{cases} \\ \text{subject to} & \lambda \succeq 0 \end{array}$$

equivalent form

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

Remark: the interesting symmetry between the standard and inequality form LPs and their duals: the dual of a standard form LP is an LP with only inequality constraints, and vice versa.

Two-way partition problem

primal problem

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

dual problem

$$\text{maximize} \quad g(\nu) = \begin{cases} -\mathbf{1}^T \nu & W + \mathbf{diag}(\nu) \succeq 0 \\ -\infty & \text{otherwise} \end{cases}$$

equivalent form

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

Weak duality

Statement

$$d^* \leq p^*$$

- the above weak duality inequality always holds (regardless of convexity)
- can be used to find nontrivial lower bounds for difficult problem
- if the primal problem is unbounded below ($p^* = -\infty$), we must have the infeasible Lagrange dual problem ($d^* = -\infty$). Conversely, it holds true.

Example

solving SDP

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

gives a lower bound for two-way partitioning problem

Strong duality

Statement

$$d^* = p^*$$

- does not hold in general
- usually holds for convex problems

Constraint qualifications

- conditions that guarantee strong duality for convex problems
- there exist many types, example below

Slater's constraint qualification

If a convex problem

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

is strictly feasible, namely

$$\exists\, x \in \mathbf{int}\, \mathcal{D} \qquad \text{such that} \qquad f_i(x) < 0, \quad i = 1, \cdots, m, \quad Ax = b,$$

then strong duality holds. If moreover $p^* > -\infty$, then the dual optimum is attained.

- $\mathbf{int}\, \mathcal{D}$ can be replaced with $\mathbf{relint}\, \mathcal{D}$ (interior relative to affine hull)
- linear inequalities do not need to hold with strict inequality
- strong duality holds for LP unless both primal and dual are infeasible
  (for LP, dual of dual is primal, Slater's condition and feasibility agree)

Quadratic program

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

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

dual function

$$g(\lambda) = \inf_x \left(x^T P x + \lambda^T (Ax - b) \right) = -(1/4)\lambda^T A P^{-1} A^T \lambda - b^T \lambda$$

dual problem

$$\begin{array}{ll} \text{maximize} & -(1/4)\lambda^T A P^{-1} A^T \lambda - b^T \lambda \\ \text{subject to} & \lambda \succeq 0 \end{array}$$

- by Slater's condition $p^\star = d^\star$ holds if primal problem is feasible
- in fact $p^\star = d^\star$ always holds (dual of dual is primal, dual always satisfies Slater)

A nonconvex problem with strong duality

primal problem (nonconvex if $A \not\succeq 0$)

$$\begin{array}{ll} \text{minimize} & x^T A x + 2b^T x \\ \text{subject to} & x^T x \le 1 \end{array}$$

dual problem

$$\begin{array}{ll} \text{maximize} & -b^T (A + \lambda I)^\dagger b - \lambda \\ \text{subject to} & A + \lambda I \succeq 0 \\ & b \in \mathcal{R}(A + \lambda I) \end{array}$$

equivalent SDP

$$\begin{array}{ll} \text{maximize} & -t - \lambda \\ \text{subject to} & \begin{bmatrix} A + \lambda I & b \\ b^T & t \end{bmatrix} \succeq 0 \end{array}$$

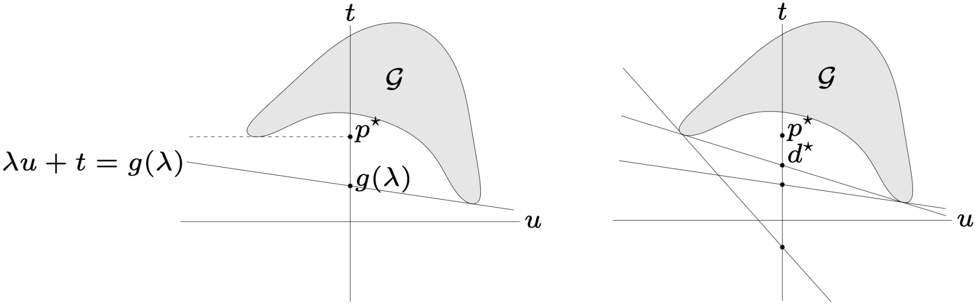
strong duality holds although primal problem is nonconvex (not easy to show)

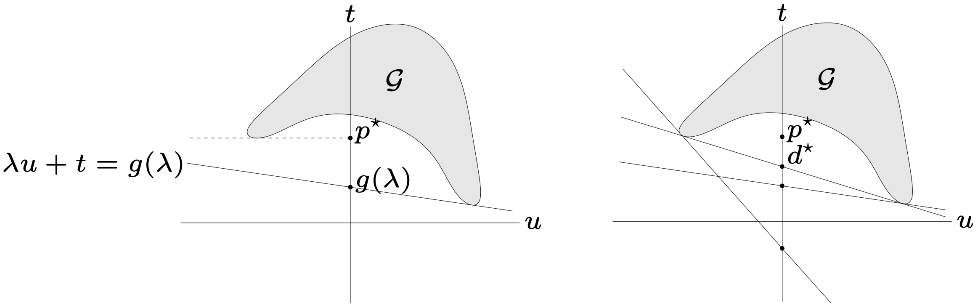
geometric description consider problem with one constraint

$$
\begin{array}{ll}
\text{minimize} & t = f_0(x) \\
\text{subject to} & u = f_1(x) \leq 0
\end{array}
$$

set of value pairs

$$
\mathcal{G} = \{(u, t) \mid u = f_1(x), t = f_0(x) \text{ for some } x \in \mathcal{D}\}
$$





interpretation of primal optimal value

$$p^* = \mathbf{inf}\{t \mid (u,t) \in \mathcal{G} \text{ and } u \leq 0\}$$

interpretation of dual objective value

$$g(\lambda) = \mathbf{inf}\{t + \lambda u \mid (u,t) \in \mathcal{G}\} = \mathbf{inf}\{(\lambda, 1)^T(u,t) \mid (u,t) \in \mathcal{G}\}$$

t-intercept of the (non-vertical) supporting hyperplane to $\mathcal{G}$ with normal vector $(\lambda, 1)^T$

interpretation of weak duality fix $\lambda \geq 0$ we have

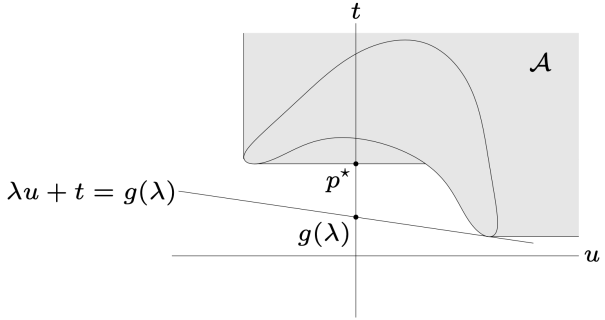
$$t + \lambda u \leq t \qquad \text{for any} \qquad (u,t) \in \mathcal{G} \text{ with } u \leq 0$$

therefore we obtain

$$\begin{array}{ccc} \mathbf{inf}\{t + \lambda u \mid (u,t) \in \mathcal{G} \text{ with } u \leq 0\} & \leq & \mathbf{inf}\{t \mid (u,t) \in \mathcal{G} \text{ with } u \leq 0\} \\ \rotatebox{90}{$\leq$} & & \| \\ g(\lambda) = \mathbf{inf}\{t + \lambda u \mid (u,t) \in \mathcal{G}\} & & p^\star \end{array}$$

epigraph variation we still assume $\lambda \geq 0$, and replace $\mathcal{G}$ by

$$\mathcal{A} = \{(u,t) \mid u \geq f_1(x) \text{ and } t \geq f_0(x) \text{ for some } x \in \mathcal{D}\}$$

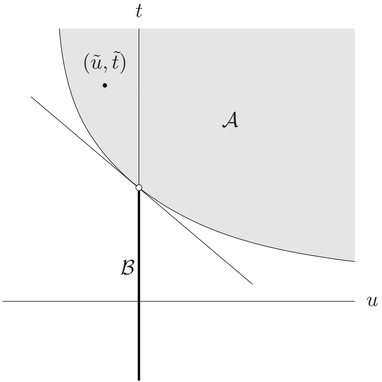


$$g(\lambda) = \mathbf{inf}\{(\lambda,1)^T(u,t) \mid (u,t) \in \mathcal{A}\} \quad \text{and} \quad p^* = \mathbf{inf}\{t \mid (0,t) \in \mathcal{A}\}$$

therefore we obtain $$g(\lambda) \leq (\lambda,1)^T(0,t) = p^*$$

strong duality holds $\iff$ $\exists$ nonvertical supporting hyperplane to $\mathcal{A}$ at $(0,p^*)$

Slater's condition for convex problems implies strong duality



- convex problems $\implies$ $\mathcal{A}$ is convex $\implies$ supporting hyperplane H at $(0, p^*)$ exists
- Slater's condition $\implies$ $\exists\ (\tilde{u}, \tilde{t}) \in \mathcal{A}$ with $\tilde{u} < 0$ $\implies$ H cannot be vertical

Slater's constraint qualification if a convex problem

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & F(x) \preceq 0 \\ & Ax = b \end{array}$$

is strictly feasible, namely

$$\exists\, x \in \mathbf{int}\, \mathcal{D} \qquad \text{such that} \qquad F(x) \prec 0 \;\; \text{and} \;\; Ax = b,$$

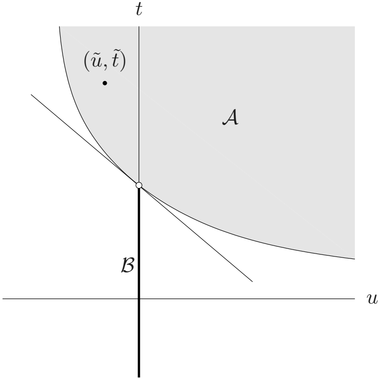
then strong duality holds. If moreover $p^* > -\infty$, then the dual optimum is attained.

Proof Without loss of generality, we assume

- p^* is finite (otherwise the result follows immediately from weak duality)
- A has full row rank (achieved by removing redundant equations)

Step 1. Consider sets $\mathcal{A},\ \mathcal{B} \subseteq \mathbb{R}^m \times \mathbb{R}^p \times \mathbb{R}$ defined as

$$\mathcal{A} = \{(u, v, t) \mid u \succeq F(x),\ v = Ax - b,\ t \geq f_0(x) \text{ for some } x \in \mathcal{D}\}$$
$$\mathcal{B} = \{(0, 0, s) \mid s < p^*\}$$



Observe that $\mathcal{A}$ and $\mathcal{B}$ are disjoint and both convex. (Prove it yourself!)

Step 2. By separating hyperplane theorem, $\exists\ (\tilde{\lambda}, \tilde{\nu}, \mu) \neq 0$ and $\alpha \in \mathbb{R}$ such that

$$(u, v, t) \in \mathcal{A} \quad \Longrightarrow \quad \tilde{\lambda}^T u + \tilde{\nu}^T v + \mu t \geq \alpha; \tag{1}$$

$$(u, v, t) \in \mathcal{B} \quad \Longrightarrow \quad \tilde{\lambda}^T u + \tilde{\nu}^T v + \mu t \leq \alpha. \tag{2}$$

(1) implies $\tilde{\lambda} \succeq 0$ and $\mu \geq 0$ (otherwise LHS is unbounded below over $\mathcal{A}$).

(2) implies $\mu p^* \leq \alpha$.

Combining them to obtain

$$\tilde{\lambda}^T F(x) + \tilde{\nu}^T (Ax - b) + \mu f_0(x) \geq \alpha \geq \mu p^* \quad \text{for all} \quad x \in \mathcal{D} \tag{3}$$

Step 3. We show that $\mu > 0$ by contradiction. If $\mu = 0$, then (3) implies

$$\tilde{\lambda}^T F(x) + \tilde{\nu}^T (Ax - b) \geq 0 \quad \text{for all} \quad x \in \mathcal{D}.$$

Assume $\tilde{x}$ is a strictly feasible point, then

$$\tilde{\lambda}^T F(\tilde{x}) \geq 0.$$

However $\tilde{\lambda} \succeq 0$ and $F(\tilde{x}) \prec 0$, hence $\tilde{\lambda} = 0$. It follows that $\tilde{\nu} \neq 0$ and

$$\tilde{\nu}^T (Ax - b) \geq 0 \quad \text{for all} \quad x \in \mathcal{D}.$$

But $A\tilde{x} - b = 0$ and $\tilde{x} \in \mathbf{int}\, \mathcal{D}$ imply $\tilde{\nu}^T (Ax - b) < 0$ for some $x \in \mathcal{D}$, unless $\tilde{\nu}^T A = 0$. The assumption of A having full row rank implies $\tilde{\nu} = 0$, contradiction.

Step 4. By Step 3 we can divide both sides of (3) by μ to obtain

$$L\left(x, \tilde{\lambda}/\mu, \tilde{\nu}/\mu\right) \geq p^* \quad \text{for all} \quad x \in \mathcal{D}.$$

Therefore

$$g\left(\tilde{\lambda}/\mu, \tilde{\nu}/\mu\right) = \inf_{x \in \mathcal{D}} L\left(x, \tilde{\lambda}/\mu, \tilde{\nu}/\mu\right) \geq p^*.$$

By weak duality we also have

$$p^* \geq d^* \geq g\left(\tilde{\lambda}/\mu, \tilde{\nu}/\mu\right).$$

Hence all of them are equal – strong duality holds and dual optimum is attained. □

Certificate of suboptimality and stopping criteria

We do not assume the primal problem is convex, unless explicitly stated

- Without knowing the exact value of p^*, we can bound how suboptimal a given feasible point is
$$f_0(x) - p^* \leq f_0(x) - g(\lambda, \nu),$$
where $\epsilon = f_0(x) - g(\lambda, \nu)$ is called the **duality gap**.
- $p^* \in [g(\lambda, \nu), f_0(x)], \quad d^* \in [g(\lambda, \nu), f_0(x)]$
- if the duality gap is zero, then x is primal optimal and (λ, ν) is dual optimal
- the stopping criterion (the condition for terminating the algorithm)
$$f_0(x^{(k)}) - g(\lambda^{(k)}, \nu^{(k)}) \leq \epsilon_{\text{abs}},$$
guarantees that when the algorithm terminates, $x^{(k)}$ is ϵ_{abs}-suboptimal.

Conditions for achieving optimality, complementary slackness

assume x^* is primal optimal, (λ^*, ν^*) is dual optimal

$$
\begin{aligned}
g(\lambda^*, \nu^*) &= \inf_{x \in \mathcal{D}} \left(f_0(x) + \sum_{i=1}^{m} \lambda_i^* f_i(x) + \sum_{i=1}^{p} \nu_i^* h_i(x) \right) \\
&\leq f_0(x^*) + \sum_{i=1}^{m} \lambda_i^* f_i(x^*) + \sum_{i=1}^{p} \nu_i^* h_i(x^*) \leq f_0(x^*) \qquad (4)
\end{aligned}
$$

assume strong duality holds, then both inequalities hold with equality

- x^* minimizes $L(x, \lambda^*, \nu^*)$
- $\lambda_i^* f_i(x^*) = 0$ for each $i = 1, \cdots, m$, namely, for each pair of inequalities

$$\lambda_i^* \geq 0 \qquad \text{and} \qquad f_i(x^*) \leq 0$$

  at least one of them achieves equality (complementary slackness)

KKT conditions

assume $f_0, f_1, \cdots, f_m$ and $h_1, \cdots, h_p$ are all differentiable (hence with open domains)

Karush-Kuhn-Tucker conditions

1. primal constraints $\quad f_i(x) \leq 0,\ i = 1, \cdots, m; \qquad h_i(x) = 0,\ i = 1, \cdots, p$
2. dual constraints $\quad \lambda \succeq 0$
3. complementary slackness $\quad \lambda_i f_i(x) = 0,\ i = 1, \cdots, m$
4. gradient of Lagrangian with respect to x vanishes

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

necessity if strong duality holds **(the primal may be nonconvex)**

$$(x^*, \lambda^*, \nu^*) \text{ are optimal} \quad \Longrightarrow \quad (x^*, \lambda^*, \nu^*) \text{ satisfy KKT}$$

sufficiency if primal problem is **convex**

$$(x^*, \lambda^*, \nu^*) \text{ satisfy KKT} \quad \Longrightarrow \quad (x^*, \lambda^*, \nu^*) \text{ are optimal}$$

proof

- conditions 1 & 2 imply primal and dual feasibility
- condition 3 (complementary slackness) is responsible for the equality of the last step in (4), $f_0(\tilde{x}) = L(\tilde{x}, \tilde{\lambda}, \tilde{\nu})$ where $(\tilde{x}, \tilde{\lambda}, \tilde{\nu})$ is any point satisfying the KKT
- condition 4 (and convexity) is responsible for the equality of the middle step in (4), $g(\tilde{\lambda}, \tilde{\nu}) = L(\tilde{x}, \tilde{\lambda}, \tilde{\nu})$

necessity + sufficiency assume differentiability + convexity + Slater then

$$x^* \text{ is optimal} \quad \Longleftrightarrow \quad (x^*, \lambda^*, \nu^*) \text{ satisfy KKT for some } \lambda^* \text{ and } \nu^*$$

Example

Equality constrained convex quadratic minimization

$$\begin{array}{ll} \text{minimize} & (1/2)x^T P x + q^T x + r \\ \text{subject to} & Ax = b \end{array}$$

where $P \in \mathbb{S}^n_+$. The KKT conditions are

$$Ax^* = b, \qquad Px^* + q + A^T \nu^* = 0,$$

which can be written as

$$\begin{bmatrix} P & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} x^* \\ \nu^* \end{bmatrix} = \begin{bmatrix} -q \\ b \end{bmatrix}$$

Example

Water-filling

assume $\alpha_i > 0$ for $i = 1, \cdots, n$

$$\begin{array}{ll} \text{minimize} & -\sum_{i=1}^{n} \log(x_i + \alpha_i) \\ \text{subject to} & x \succeq 0 \\ & \mathbf{1}^T x = 1 \end{array}$$

x is optimal $\iff$ $x \succeq 0$, $\mathbf{1}^T x = 1$, there exists $\lambda \in \mathbb{R}^n$ and $\nu \in \mathbb{R}$ such that

$$\lambda \succeq 0, \qquad \lambda_i x_i = 0, \qquad \frac{1}{x_i + \alpha_i} + \lambda_i = \nu, \qquad i = 1, \cdots, n$$

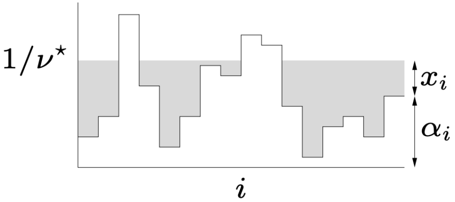
- if $\nu \leq 1/\alpha_i$, then $\lambda_i = 0$ and $x_i = 1/\nu - \alpha_i$
- if $\nu \geq 1/\alpha_i$, then $\lambda_i = \nu - 1/\alpha_i$ and $x_i = 0$

determine ν from

$$\mathbf{1}^T x = \sum_{i=1}^{n} \mathbf{max}\{0, 1/\nu - \alpha_i\} = 1$$

water-filling algorithm

- left-hand side is a piecewise linear increasing function in $1/\nu$
- n patches, level of patch i is at height α_i
- flood area with unit amount of water, resulting level is $1/\nu^*$



Perturbed problem

perturbed primal problem

$$\begin{array}{lll} \text{minimize} & f_0(x) & \\ \text{subject to} & f_i(x) \leq u_i, & i = 1, \cdots, m \\ & h_i(x) = v_i, & i = 1, \cdots, p \end{array}$$

perturbed dual problem

$$\begin{array}{ll} \text{maximize} & g(\lambda, \nu) - u^T\lambda - v^T\nu \\ \text{subject to} & \lambda \succeq 0 \end{array}$$

- u and v are parameters
- original primal & dual problems are recovered when $u = 0$ and $v = 0$
- $p^*(u, v)$ is optimal value as a function of u and v
- need to understand $p^*(u, v)$ from solution to unperturbed problem
- When the original problem is convex, $p^*(u, v)$ is a convex function of u and v

Global sensitivity

assume for the unperturbed problem that

- strong duality holds (e.g. convex + Slater)
- λ^* and ν^* are dual optimal

then weak duality for the perturbed problem implies

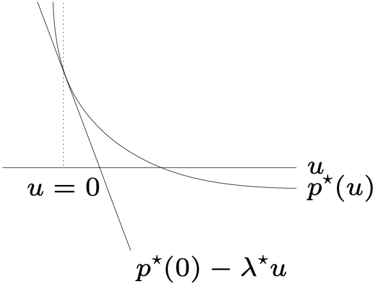
$$\begin{aligned} p^*(u,v) &\geq g(\lambda^*,\nu^*) - u^T\lambda^* - v^T\nu^* \\ &= p^*(0,0) - u^T\lambda^* - v^T\nu^* \end{aligned}$$

- λ_i^* large $\implies$ p^* increases greatly if $u_i < 0$ (tighten constraint)
- λ_i^* small $\implies$ p^* does not decrease much if $u_i > 0$ (loosen constraint)
- $\nu_i^* > 0$ large $\implies$ p^* increases greatly if $v_i < 0$
- $\nu_i^* > 0$ small $\implies$ p^* does not decrease much if $v_i > 0$
- $\nu_i^* < 0$ large $\implies$ p^* increases greatly if $v_i > 0$
- $\nu_i^* < 0$ small $\implies$ p^* does not decrease much if $v_i < 0$

Local sensitivity

assume in addition that $p^*(u, v)$ is differentiable at $(0, 0)$ then

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



(above picture exhibits $p^*(u)$ for a problem with one inequality constraint)

Duality and problem reformulations

principle

- equivalent formulations of a problem can lead to very different duals
- reformulation can be useful when dual is difficult to derive or uninteresting

common reformulations

- introduce new variables and equality constraints
- make explicit constraints implicit or vice-versa
- apply an increasing function to objective or constraint functions

Introducing new variables and equality constraints

unconstrained problem

primal problem

$$\text{minimize} \quad f_0(Ax+b)$$

dual problem

$$g = \inf_x f_0(Ax+b) = p^*$$

- no dual variable, hence dual function is constant
- strong duality holds, but dual is useless

reformulated primal problem

$$\begin{array}{ll}\text{minimize} & f_0(y) \\ \text{subject to} & Ax + b - y = 0\end{array}$$

dual of reformulated problem

$$\begin{array}{ll}\text{maximize} & b^T\nu - f_0^*(\nu) \\ \text{subject to} & A^T\nu = 0\end{array}$$

it follows from

$$g(\nu) = \inf_{x,y} \left(f_0(y) - \nu^T y + \nu^T Ax + b^T\nu \right) = \begin{cases} -f_0^*(\nu) + b^T\nu & A^T\nu = 0 \\ -\infty & \text{otherwise} \end{cases}$$

norm approximation problem

$$\text{minimize} \quad \|Ax - b\|$$

reformulated problem

$$\begin{array}{ll} \text{minimize} & \|y\| \\ \text{subject to} & y = Ax - b \end{array}$$

dual of the reformulated problem

$$\begin{array}{ll} \text{maximize} & b^T \nu \\ \text{subject to} & A^T \nu = 0 \\ & \|\nu\|_* \le 1 \end{array}$$

introducing new variables to the constraint functions

$$\begin{array}{ll} \text{minimize} & f_0(A_0x + b_0) \\ \text{subject to} & f_i(A_0x + b_0) \leq 0, \quad i = 1, \ldots, m \end{array}$$

reformulated problem

$$\begin{array}{ll} \text{minimize} & f_0(y_0) \\ \text{subject to} & f_i(y_i) \leq 0, \quad i = 1, \ldots, m \\ & A_ix + b_i = y_i, \quad i = 0, \ldots, m \end{array}$$

dual of the reformulated problem

$$\begin{array}{ll} \text{maximize} & \sum_{i=0}^{m} b_i^T \nu_i - f_0^*(\nu_0) - \sum_{i=0}^{m} \lambda_i f_i^*(\nu_i/\lambda_i) \\ \text{subject to} & \lambda \succeq 0 \\ & \sum_{i=0}^{m} A_i^T \nu_i = 0 \end{array}$$

Transforming the objective

If we replace the objective f_0 by an increasing function of f_0, the resulting problem is clearly equivalent. However, their duals can be quite different

Again: norm approximation problem

$$\text{minimize} \quad \|Ax - b\|$$

reformulated problem

$$\begin{array}{ll} \text{minimize} & \frac{1}{2}\|y\|^2 \\ \text{subject to} & y = Ax - b \end{array}$$

dual of the reformulated problem

$$\begin{array}{ll} \text{maximize} & -\frac{1}{2}\|\nu\|_*^2 + b^T\nu \\ \text{subject to} & A^T\nu = 0 \end{array}$$

Implicit constraints

LP with box constraints

primal problem

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax = b \\ & -\mathbf{1} \preceq x \preceq \mathbf{1} \end{array}$$

dual problem

$$\begin{array}{ll} \text{maximize} & -b^T \nu - \mathbf{1}^T \lambda_1 - \mathbf{1}^T \lambda_2 \\ \text{subject to} & c + A^T \nu + \lambda_1 - \lambda_2 = 0 \\ & \lambda_1 \succeq 0, \qquad \lambda_2 \succeq 0 \end{array}$$

reformulated primal problem

$$\begin{array}{ll} \text{minimize} & f_0(x) = \begin{cases} c^T x & -\mathbf{1} \preceq x \preceq \mathbf{1} \\ \infty & \text{otherwise} \end{cases} \\ \text{subject to} & Ax = b \end{array}$$

dual function

$$\begin{aligned} g(\nu) &= \inf_{-\mathbf{1} \preceq x \preceq \mathbf{1}} \left(c^T x + \nu^T (Ax - b) \right) \\ &= -b^T \nu - \|A^T \nu + c\|_1 \end{aligned}$$

dual of the reformulated problem

$$\begin{array}{ll} \text{maximize} & -b^T \nu - \|A^T \nu + c\|_1 \end{array}$$

Theorems of alternatives

- consider two systems of inequality and equality constraints
- called **weak alternatives** if no more than one system is feasible
- called **strong alternatives** if exactly one of them is feasible
- examples: for any $a \in \mathbb{R}$, with variable $x \in \mathbb{R}$
  - $x > a$ and $x \leq a - 1$ are weak alternatives
  - $x > a$ and $x \leq a$ are strong alternatives

- a **theorem of alternatives** states that two inequality systems are (weak or strong) alternatives
- can be considered the extension of duality to feasibility problems

Feasibility problems

- consider system of (not necessarily convex) inequalities and equalities

$$f_i(x) \leq 0, \quad i = 1, \ldots, m, \qquad h_i(x) = 0, \quad i = 1, \cdots, p$$

- express as **feasibility problem**

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

- if system is feasible, $p^* = 0$; if not, $p^* = \infty$

Duality for feasibility problems

- dual function of feasibility problem is

$$g(\lambda, \nu) = \inf_x \left(\sum_{i=1}^{m} \lambda_i f_i(x) + \sum_{i=1}^{p} \nu_i h_i(x) \right)$$

- for $\lambda \succeq 0$, we have $g(\lambda, \nu) \leq p^*$
- it follows that feasibility of the inequality system

$$\lambda \succeq 0, \qquad g(\lambda, \nu) > 0$$

  implies the original system is infeasible
- so this is a weak alternative to original system
- it is strong if f_i convex, h_i affine, and a constraint qualification holds
- g is positive homogeneous so we can write alternative system as

$$\lambda \succeq 0, \qquad g(\lambda, \nu) \geq 1$$

Example: Nonnegative solution of linear equations

- consider system

$$Ax = b, \qquad x \succeq 0$$

- dual function is $g(\lambda, \nu) = \begin{cases} -\nu^T b & A^T\nu = \lambda \\ -\infty & \text{otherwise} \end{cases}$

- can express strong alternative of $Ax = b, x \succeq 0$ as

$$A^T\nu \succeq 0, \qquad \nu^T b \le -1$$

(we can replace $\nu^T b \le -1$ with $\nu^T b = -1$)

Farkas' lemma

- Farkas' lemma:

$$Ax \preceq 0, \quad c^T x < 0 \;\text{ and }\; A^T y + c = 0, \quad y \succeq 0$$

  are strong alternatives

- proof: use (strong) duality for (feasible) LP

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

- the dual is

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

Problems with generalized inequalities

primal problem (proper cone $K_i \subseteq \mathbb{R}^{k_i}$ for $i = 1, \cdots, m$)

$$\begin{array}{lll} \text{minimize} & f_0(x) & \\ \text{subject to} & f_i(x) \preceq_{K_i} 0, & i = 1, \cdots, m \\ & h_i(x) = 0, & i = 1, \cdots, p \end{array}$$

- Lagrange multiplier for $f_i(x) \preceq_{K_i} 0$ is vector $\lambda_i \in \mathbb{R}^{k_i}$, for $h_i(x) = 0$ scalar $\nu_i \in \mathbb{R}$
- Lagrangian $L\colon \mathbb{R}^n \times \mathbb{R}^{k_1} \times \cdots \times \mathbb{R}^{k_m} \times \mathbb{R}^p \longrightarrow \mathbb{R}$

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

- dual function $g\colon \mathbb{R}^{k_1} \times \cdots \times \mathbb{R}^{k_m} \times \mathbb{R}^p \longrightarrow \mathbb{R}$

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

Lower bound property if $\lambda_i \succeq_{K_i^*} 0$, then $g(\lambda_1, \cdots, \lambda_m, \nu) \leq p^*$

Proof For any feasible $\tilde{x}$ we have

$$\begin{aligned} f_0(\tilde{x}) &\geq f_0(\tilde{x}) + \sum_{i=1}^{m} \lambda_i^T f_i(\tilde{x}) + \sum_{i=1}^{p} \nu_i h_i(\tilde{x}) \\ &\geq \inf_{x \in \mathcal{D}} L(x, \lambda_1, \cdots, \lambda_m, \nu) \\ &= g(\lambda_1, \cdots, \lambda_m, \nu) \end{aligned}$$

We conclude by minimizing over all feasible $\tilde{x}$.

dual problem

$$
\begin{array}{ll}
\text{maximize} & g(\lambda_1, \cdots, \lambda_m, \nu) \\
\text{subject to} & \lambda_i \succeq_{K_i^*} 0, \qquad i = 1, \cdots, m
\end{array}
$$

weak duality (always holds)

$$p^* \geq d^*$$

strong duality (holds for convex problem with constraint qualification)

$$p^* = d^*$$

Slater's condition: primal problem is strictly feasible

Semidefinite program

primal SDP (assume $F_i, G \in \mathbb{S}^k$)

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & x_1 F_1 + \cdots + x_n F_n \preceq G \end{array}$$

Lagrange multiplier

$$Z \in \mathbb{S}^k$$

Lagrangian

$$L(x, Z) = c^T x + \mathbf{tr}\left(Z(x_1 F_1 + \cdots + x_n F_n - G)\right)$$

dual function

$$g(Z) = \inf_x L(x, Z) = \begin{cases} -\mathbf{tr}(ZG) & c_i + \mathbf{tr}(ZF_i) = 0 \text{ for all } i = 1, \cdots, n \\ -\infty & \text{otherwise} \end{cases}$$

dual SDP

$$\begin{array}{ll} \text{maximize} & -\mathbf{tr}(ZG) \\ \text{subject to} & Z \succeq 0 \\ & c_i + \mathbf{tr}(ZF_i) = 0, \qquad i = 1, \cdots, n \end{array}$$

strong duality

$p^* = d^*$ holds if primal SDP is strictly feasible ($\exists\ x$ such that $x_1 F_1 + \cdots + x_n F_n \prec G$)