5. Duality

- Lagrange dual problem

- weak and strong duality

- geometric interpretation

- optimality conditions

- perturbation and sensitivity analysis

- examples

- generalized inequalities

Lagrangian

standard form problem (not necessarily convex)

$$
\begin{array}{ll}
\text{minimize} & f_0(x) \\
\text{subject to} & f_i(x) \le 0, \quad i = 1, \ldots, m \\
& h_i(x) = 0, \quad i = 1, \ldots, p
\end{array}
$$

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

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

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

- weighted sum of objective and constraint functions

- λ_i is Lagrange multiplier associated with $f_i(x) \le 0$

- ν_i is Lagrange multiplier associated with $h_i(x) = 0$

Lagrange dual function

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

对偶变量

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

g 是 concave 的 关于 λ, ν

$$= \inf_{x \in \mathcal{D}} \left(f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x) + \sum_{i=1}^{p} \nu_i h_i(x) \right)$$

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

lower bound property: if $\lambda \succeq 0$, then $g(\lambda, \nu) \leq p^\star$

proof: if $\tilde{x}$ is feasible and $\lambda \succeq 0$, then

$\Rightarrow \quad \max_{\lambda \geq 0} g(\lambda, \nu) \leq p^*$

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

minimizing over all feasible $\tilde{x}$ gives $p^\star \geq g(\lambda, \nu)$

Least-norm solution of linear equations

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

dual function

- Lagrangian is $L(x, \nu) = x^T x + \nu^T (Ax - b)$

- to minimize L over x, set gradient equal to zero:

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

- plug in in L to obtain g:

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

a concave function of ν

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

Standard form LP

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

dual function

- Lagrangian is

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

- L is affine in x, hence

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

g is linear on affine domain $\{(\lambda, \nu) \mid A^T \nu - \lambda + c = 0\}$, hence concave

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

Equality constrained norm minimization

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

dual function

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

where $\|v\|_* = \sup_{\|u\| \leq 1} u^T v$ is dual norm of $\|\cdot\|$

proof: follows from $\inf_x (\|x\| - y^T x) = 0$ if $\|y\|_* \leq 1$, $-\infty$ otherwise

- if $\|y\|_* \leq 1$, then $\|x\| - y^T x \geq 0$ for all x, with equality if $x = 0$

- if $\|y\|_* > 1$, choose $x = tu$ where $\|u\| \leq 1$, $u^T y = \|y\|_* > 1$:

$$\|x\| - y^T x = t(\|u\| - \|y\|_*) \to -\infty \quad \text{as } t \to \infty$$

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

Two-way partitioning

$$\text{minimize} \quad x^T W x$$
$$\text{subject to} \quad x_i^2 = 1, \quad i = 1, \ldots, n$$

- a nonconvex problem; feasible set contains 2^n discrete points

- interpretation: partition $\{1, \ldots, n\}$ in two sets; W_{ij} is cost of assigning i, j to the same set; $-W_{ij}$ is cost of assigning to different sets

dual function

$$g(\nu) = \inf_x (x^T W x + \sum_i \nu_i(x_i^2 - 1)) \quad = \quad \inf_x x^T(W + \mathbf{diag}(\nu))x - \mathbf{1}^T\nu$$

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

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

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

Lagrange dual and conjugate function

$$\text{minimize} \quad f_0(x)$$
$$\text{subject to} \quad Ax \preceq b, \quad Cx = d$$

dual function

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

$$= -f_0^*(-A^T \lambda - C^T \nu) - b^T \lambda - d^T \nu$$

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

example: entropy maximization

$$f_0(x) = \sum_{i=1}^{n} x_i \log x_i, \qquad f_0^*(y) = \sum_{i=1}^{n} e^{y_i - 1}$$

Example:
$$\min \quad f(x) + g(y)$$
$$\text{s.t} \quad Ax - y = 0$$

求对偶: $L(x, y, v) = f(x) + g(y) + \langle v, Ax - y \rangle$

$$h(v) = \inf_{x, y} \{ f(x) + g(y) + \langle v, Ax - y \rangle \}$$

$$= \inf_{x} \{ f(x) + \langle A^T v, x \rangle \} + \inf_{y} \{ g(y) - \langle v, y \rangle \}$$

$$= - f^*(-A^T v) - g^*(v)$$

用共轭函数写出其对偶

The dual problem

Lagrange dual problem

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

- finds best lower bound on $p^\star$, obtained from Lagrange dual function

- a convex optimization problem; optimal value denoted $d^\star$

- λ, ν are dual feasible if $\lambda \succeq 0$, $(\lambda, \nu) \in \mathbf{dom}\, g$

- often simplified by making implicit constraint $(\lambda, \nu) \in \mathbf{dom}\, g$ explicit

example: standard form LP and its dual (page 5–5)

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

Weak and strong duality

weak duality: $d^\star \le p^\star$

- always holds (for convex and nonconvex problems)
- can be used to find nontrivial lower bounds for difficult problems

  for example, solving the 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 the two-way partitioning problem on page 5–7

strong duality: $d^\star = p^\star$ ← 目标 寻找 相等条件

$$\Longleftrightarrow \sup_{x} \inf_{\lambda \ge 0, \nu} L(x, \lambda, \nu) = \inf_{x} \sup_{\lambda, \nu} L(x, \lambda, \nu)$$

$$\min_{\lambda \ge 0, \nu} g(\lambda, \nu)$$

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

$$\Longleftarrow (\tilde{x}, \tilde{\lambda}, \tilde{\nu}) \ 为 L(x, \lambda, \nu) \ 的 鞍点.$$

Slater's constraint qualification

strong duality holds for a convex problem

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

if it is strictly feasible, *i.e.*,

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

内点   或 相对内点

- also guarantees that the dual optimum is attained (if $p^\star > -\infty$)

- can be sharpened: *e.g.*, can replace $\mathbf{int}\,\mathcal{D}$ with $\mathbf{relint}\,\mathcal{D}$ (interior relative to affine hull); linear inequalities do not need to hold with strict inequality, . . .

  线性约束可 弱化 相对内点条件

- there exist many other types of constraint qualifications

Inequality form LP

primal problem

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

dual function

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

dual problem

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

- from Slater's condition: $p^\star = d^\star$ if $A\tilde{x} \prec b$ for some $\tilde{x}$

- in fact, $p^\star = d^\star$ except when primal and dual are infeasible

Quadratic program

primal problem (assume $P \in \mathbf{S}^n_{++}$) 平格正定

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

dual function

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

dual problem

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

- from Slater's condition: $p^\star = d^\star$ if $A\tilde{x} \prec b$ for some $\tilde{x}$

- in fact, $p^\star = d^\star$ always

  但由原约束是线性的

  以以只求 $A\tilde{x} \leq b$ 即可

A nonconvex problem with strong duality

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

信賴域方法

$A \not\succeq 0$, hence nonconvex $A \in S_n$

dual function: $g(\lambda) = \inf_x (x^T (A + \lambda I) x + 2b^T x - \lambda)$

取 $x \in \ker(A + \lambda I)$

- unbounded below if $A + \lambda I \not\succeq 0$ or if $A + \lambda I \succeq 0$ and $b \notin \mathcal{R}(A + \lambda I)$

- minimized by $x = -(A + \lambda I)^\dagger b$ otherwise: $g(\lambda) = -b^T (A + \lambda I)^\dagger b - \lambda$

dual problem and equivalent SDP:

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

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

对 $X = \begin{bmatrix} A & B \\ B^T & C \end{bmatrix}$ $\det A \neq 0$

$S = C - B^T A^{-1} B$

则 ① $X > 0 \iff A > 0, S > 0$

② 若 $A > 0$ 则 $X \geq 0 \iff S \geq 0$

③ $X \geq 0 \iff A \geq 0, \quad (I - AA^+)B = 0$

$S = C - B^T A^+ B \geq 0$

$(I - AA^T)Bv = 0 \quad \forall v \quad \Big| \quad AA^+ = U\Sigma V^T V \Sigma^+ U^T$

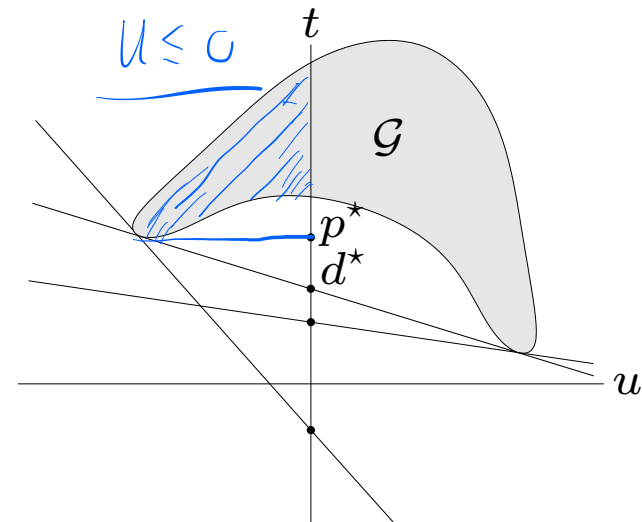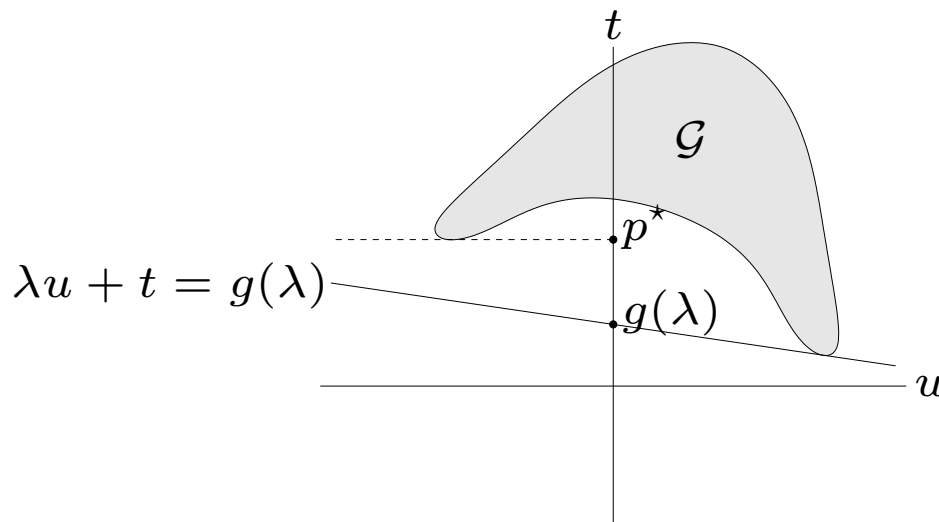$\iff Bv \in Rang(A) \qquad \qquad = UU^T$

Geometric interpretation

for simplicity, consider problem with one constraint $f_1(x) \leq 0$

interpretation of dual function:

$$g(\lambda) = \inf_{(u,t)\in\mathcal{G}} (t + \lambda u), \qquad \text{where} \quad \mathcal{G} = \{(f_1(x), f_0(x)) \mid x \in \mathcal{D}\}$$

- $\lambda u + t = g(\lambda)$ is (non-vertical) supporting hyperplane to $\mathcal{G}$
- hyperplane intersects t-axis at $t = g(\lambda)$

类似 考虑 $\min\limits_{x} f_0(x)$, s.t. $f(x)\leq 0$, $h(x)=0$

记 $G = \{ (f(x), h(x), f_0(x)) \in \mathbb{R}^m \times \mathbb{R}^p \times \mathbb{R} \mid x \in D\}$

$g(\lambda,\mu) = \inf\limits_{x} L(x,\lambda,\mu) = f_0(x) + \lambda^T f(x) + \mu^T f_0(x)$

$$= (\lambda,\mu,1)^T (u,v,t) , \quad (u,v,t)\in G$$

$\Rightarrow \underline{(\lambda,\mu,1)^T} (u,v,t) \geq g(\lambda,\mu) \quad \forall (u,v,t)\in G$

$\lambda \geq 0$

$\downarrow$

定义了 $\mathbb{R}^m \times \mathbb{R}^p \times \mathbb{R}$ 中 G 的一个支撑超平面的方向

而 $p^* = \inf \{t \mid (u,v,t)\in G, \ u\leq 0, \ v=0\}$

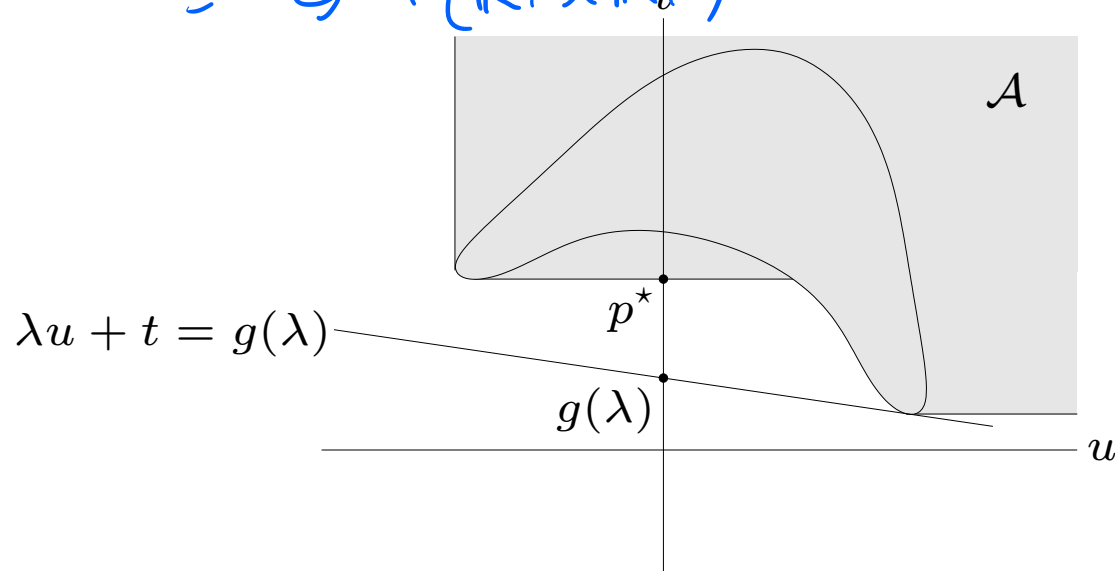$\geq \inf \{(\lambda,\mu,1)^T(u,v,t) \mid (u,v,t)\in G, \ u\leq 0, v=0\}$

$\geq \inf \{ \qquad\qquad\qquad \mid (u,v,t)\in G\} = g(\lambda,\mu)$

epigraph variation: same interpretation if $\mathcal{G}$ is replaced with

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

$$= \mathcal{G} + (\mathbb{R}_+ \times \mathbb{R}_t)$$

strong duality

- holds if there is a non-vertical supporting hyperplane to $\mathcal{A}$ at $(0, p^\star)$
- for convex problem, $\mathcal{A}$ is convex, hence has supp. hyperplane at $(0, p^\star)$
- Slater's condition: if there exist $(\tilde{u}, \tilde{t}) \in \mathcal{A}$ with $\tilde{u} < 0$, then supporting hyperplanes at $(0, p^\star)$ must be non-vertical

类似考虑 $A = G + \{R_+^m \times \{0\} \times R_+\}$

$= \{(u, v, t) \mid \exists x \in D \text{ s.t } f(x) \le u, h(x) = v, f_0(x) \le t\}$

则 A 为凸集 f, h, f_0 为凸)

$p^* = \{t \mid (0, 0, t) \in A\}$　　所以对 $\lambda \ge 0$

有 $g(\lambda, \mu) = \inf \{(\lambda, \mu, 1)^T (u, v, t) \mid (u, v, t) \in A\}$

$\|$

$\lambda^T f(x) + \mu^T h(x) + t f(x) + \lambda^T u + ts$

$\Rightarrow (\lambda, \mu, 1)^T (u, v, t) \ge g(\lambda, \mu) \quad \forall (u, v, t) \in A.$

$\underline{\text{Proof of (SCQ)}}$ 若 ① $\hat{x} \in \text{int } D$

② $\text{rank}(A) = p$

$f(\hat{x}) < 0.$

对 $\min \ f_0(x)$

$\text{s.t. } f(x) \le 0$

$Ax = b, A \in \mathbb{R}^{p \times n}.$

考虑 $B = \{(0, 0, s) \mid s < p^*\}$ 凸　　则 $A \cap B = \phi$ ∴ $\exists (\tilde{\lambda}, \tilde{v}, \mu) \ne 0$ 及 $\alpha.$

使 $\forall (u, v, t) \in A$ 有 $\tilde{\lambda}^T u + \tilde{v}v + \mu t \ge \alpha.$ ⊛

$\forall (u, v, t) \in B$ 有 $\cdots \cdots \cdots \le \alpha.$ ⊗

由 ⊛ 知 $\tilde{\lambda} \ge 0, \mu \ge 0$ 再由 ⊗ 知 $\mu t \le \alpha \quad \forall t < p^* \Rightarrow \mu p^* \le \alpha.$

于是 $\tilde{\lambda}^T u + \tilde{v}v + \mu t \ge \alpha \ge \mu p^*$

$$\therefore \quad \tilde{\lambda}^T f(x) + \tilde{\upsilon}^T (Ax - b) + \mu f_0(x) \geq \mu p^*$$

① 若 $\mu > 0 \quad \Rightarrow \quad p^* \leq \frac{1}{\mu} (\qquad)$

$$= L(x, \frac{\tilde{\lambda}}{\mu}, \frac{\tilde{\upsilon}}{\mu})$$

$$\Rightarrow p^* \leq \inf_x L(x, \frac{\tilde{\lambda}}{\mu}, \frac{\tilde{\upsilon}}{\mu}) = g(\tilde{\lambda}/\mu, \tilde{\upsilon}/\mu) \leq d^*$$

$\Rightarrow p^* \leq d^*$ 又 强对偶成立 $p^* = d^*$

② 若 $\mu = 0 \quad \Rightarrow \quad \tilde{\lambda}^T f(x) + \tilde{\upsilon}^T(Ax - b) \geq 0 \qquad \forall x \in D$

$\Rightarrow$ 取 $\tilde{x}$ 由 $A\tilde{x} = b \Rightarrow f(\tilde{x}) < 0$

而有 $\tilde{\lambda}^T f(\tilde{x}) \geq 0$ 又 $\tilde{\lambda}^T \geq 0 \Rightarrow \tilde{\lambda} = 0$

又 $(\tilde{\lambda}, \tilde{\upsilon}, \mu) \neq 0 \quad \therefore \tilde{\upsilon} \neq 0 \Rightarrow \tilde{\upsilon}^T(Ax - b) \geq 0$

$$\tilde{\upsilon}^T(A\tilde{x} - b) = 0$$

$\Rightarrow \tilde{\upsilon}^T A(x - \tilde{x}) \geq 0 \quad \forall x \in D$. 由 $\tilde{x}$ 在内点中

$\Rightarrow A^T \upsilon = 0 \Rightarrow \tilde{\upsilon} = 0$ 由 rank$(A) = P$ 满秩 矛盾.

$\therefore$ 只能为情形①

一般地有 $p^* \geq d^*$

$\therefore f_0(x) - p^* \leq f_0(x) - d^* \leq \underbrace{f_0(x) - g(\lambda, \nu)}_{对偶间隙}$

一般 用算法 时 用 $\| f_0(x) - g(\lambda, \nu) \| \leq \xi$ 作为终止条件

Complementary slackness

assume strong duality holds, $x^\star$ is primal optimal, $(\lambda^\star, \nu^\star)$ is dual optimal

$$f_0(x^\star) = g(\lambda^\star, \nu^\star) \quad = \quad \inf_x \left(f_0(x) + \sum_{i=1}^{m} \lambda_i^\star f_i(x) + \sum_{i=1}^{p} \nu_i^\star h_i(x) \right)$$

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

$$\leq \quad f_0(x^\star)$$

hence, the two inequalities hold with equality

- $x^\star$ minimizes $L(x, \lambda^\star, \nu^\star)$ $\Rightarrow$ $\nabla_x L(x^*, \lambda^*, \nu^*) = 0$ （若凸 则是等价

- $\lambda_i^\star f_i(x^\star) = 0$ for $i = 1, \ldots, m$ (known as complementary slackness):

互补条件

$$\lambda_i^\star > 0 \Longrightarrow f_i(x^\star) = 0, \qquad f_i(x^\star) < 0 \Longrightarrow \lambda_i^\star = 0$$

Karush-Kuhn-Tucker (KKT) conditions

the following four conditions are called KKT conditions (for a problem with differentiable f_i, h_i):

1. primal constraints: $f_i(x) \leq 0$, $i = 1, \ldots, m$, $h_i(x) = 0$, $i = 1, \ldots, p$

   可行性

2. dual constraints: $\lambda \succeq 0$

3. complementary slackness: $\lambda_i f_i(x) = 0$, $i = 1, \ldots, m$

4. gradient of Lagrangian with respect to x vanishes:

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

强对偶下   最优 则 KKT.

from page 5–17: if strong duality holds and x, λ, ν are optimal, then they must satisfy the KKT conditions

必要条件

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 4th condition (and convexity): $g(\tilde{\lambda}, \tilde{\nu}) = L(\tilde{x}, \tilde{\lambda}, \tilde{\nu})$

hence, $f_0(\tilde{x}) = g(\tilde{\lambda}, \tilde{\nu})$ 凸 + KKT ⟹ 最优化

Slater + 凸 + KKT ⟺ 凸 + Slater + 最优

if **Slater's condition** is satisfied:

x is optimal if and only if there exist λ, ν that satisfy KKT conditions

- recall that Slater implies strong duality, and dual optimum is attained

- generalizes optimality condition $\nabla f_0(x) = 0$ for unconstrained problem

example: water-filling (assume $\alpha_i > 0$)

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

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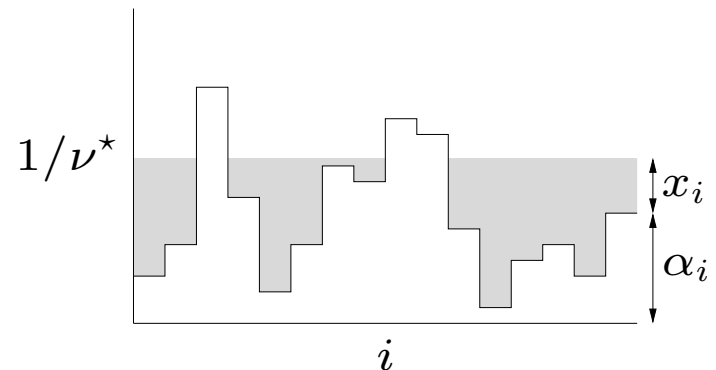
KKT 条件

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

- if $\nu < 1/\alpha_i$: $\lambda_i = 0$ and $x_i = 1/\nu - \alpha_i$

- if $\nu \geq 1/\alpha_i$: $\lambda_i = \nu - 1/\alpha_i$ and $x_i = 0$

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

interpretation

- n patches; level of patch i is at height α_i

- flood area with unit amount of water

- resulting level is $1/\nu^\star$

Perturbation and sensitivity analysis

(unperturbed) optimization problem and its dual

$$
\begin{array}{lll}
\text{minimize} & f_0(x) \\
\text{subject to} & f_i(x) \le 0, & i = 1, \dots, m \\
& h_i(x) = 0, & i = 1, \dots, p
\end{array}
\qquad
\begin{array}{ll}
\text{maximize} & g(\lambda, \nu) \\
\text{subject to} & \lambda \succeq 0
\end{array}
$$

perturbed problem and its dual

$$
\begin{array}{lll}
\text{min.} & f_0(x) \\
\text{s.t.} & f_i(x) \le u_i, & i = 1, \dots, m \\
& h_i(x) = v_i, & i = 1, \dots, p
\end{array}
\qquad
\begin{array}{ll}
\text{max.} & g(\lambda, \nu) - u^T \lambda - v^T \nu \\
\text{s.t.} & \lambda \succeq 0
\end{array}
$$

由弱对偶. $P^*(u, v) \geqslant g(\lambda^*, \mu^*) - u^T \lambda^* - v^T \mu^*$

$= P^*(0, 0) - u^T \lambda^* - v^T \mu^*$

- x is primal variable; u, v are parameters

- $p^\star(u, v)$ is optimal value as a function of u, v

- we are interested in information about $p^\star(u, v)$ that we can obtain from the solution of the unperturbed problem and its dual

global sensitivity result

assume strong duality holds for unperturbed problem, and that $\lambda^\star$, $\nu^\star$ are dual optimal for unperturbed problem

apply weak duality to perturbed problem:

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

sensitivity interpretation

- if $\lambda_i^\star$ large: $p^\star$ increases greatly if we tighten constraint i $(u_i < 0)$

- if $\lambda_i^\star$ small: $p^\star$ does not decrease much if we loosen constraint i $(u_i > 0)$

- if $\nu_i^\star$ large and positive: $p^\star$ increases greatly if we take $v_i < 0$;
  if $\nu_i^\star$ large and negative: $p^\star$ increases greatly if we take $v_i > 0$

- if $\nu_i^\star$ small and positive: $p^\star$ does not decrease much if we take $v_i > 0$;
  if $\nu_i^\star$ small and negative: $p^\star$ does not decrease much if we take $v_i < 0$

local sensitivity: if (in addition) $p^\star(u, v)$ is differentiable at $(0, 0)$, then

$$\lambda_i^\star = -\frac{\partial p^\star(0, 0)}{\partial u_i}, \qquad \nu_i^\star = -\frac{\partial p^\star(0, 0)}{\partial v_i}$$
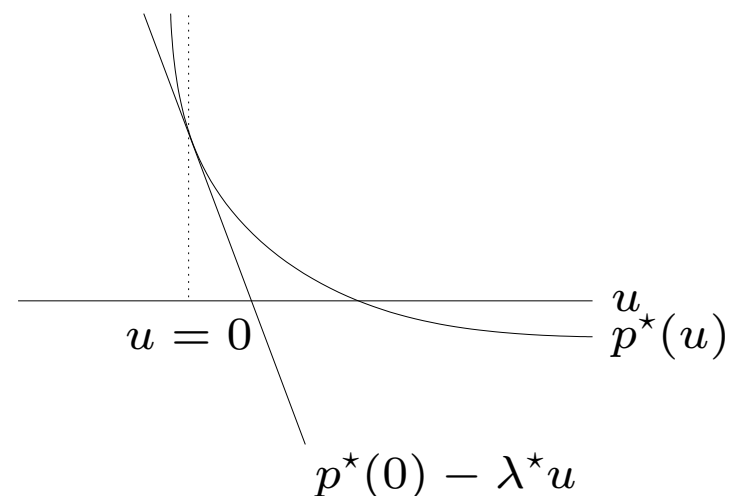
proof (for $\lambda_i^\star$): from global sensitivity result,

$$\frac{\partial p^\star(0, 0)}{\partial u_i} = \lim_{t \searrow 0} \frac{p^\star(te_i, 0) - p^\star(0, 0)}{t} \geq -\lambda_i^\star$$

$$\frac{\partial p^\star(0, 0)}{\partial u_i} = \lim_{t \nearrow 0} \frac{p^\star(te_i, 0) - p^\star(0, 0)}{t} \leq -\lambda_i^\star$$

hence, equality

$p^\star(u)$ for a problem with one (inequality) constraint:

$$u = 0$$

$$\frac{u}{p^\star(u)}$$

$$p^\star(0) - \lambda^\star u$$

Duality and problem reformulations

- equivalent formulations of a problem can lead to very different duals

- reformulating the primal problem can be useful when the dual is difficult to derive, or uninteresting

common reformulations

- introduce new variables and equality constraints

- make explicit constraints implicit or vice-versa

- transform objective or constraint functions

  e.g., replace $f_0(x)$ by $\phi(f_0(x))$ with ϕ convex, increasing

Introducing new variables and equality constraints

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

- dual function is constant: $g = \inf_x L(x) = \inf_x f_0(Ax + b) = p^\star$
- we have strong duality, but dual is quite useless

reformulated problem and its dual

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

dual function follows from

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

norm approximation problem: minimize $\|Ax - b\|$

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

can look up conjugate of $\|\cdot\|$, or derive dual directly

$$
\begin{aligned}
g(\nu) &= \inf_{x,y}(\|y\| + \nu^T y - \nu^T Ax + b^T \nu) \\
&= \begin{cases} b^T\nu + \inf_y(\|y\| + \nu^T y) & A^T\nu = 0 \\ -\infty & \text{otherwise} \end{cases} \\
&= \begin{cases} b^T\nu & A^T\nu = 0, \quad \|\nu\|_* \leq 1 \\ -\infty & \text{otherwise} \end{cases}
\end{aligned}
$$

(see page 5–4)

dual of norm approximation problem

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

Implicit constraints

LP with box constraints: primal and dual problem

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

加入部分约束

reformulation with box constraints made implicit

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

dual function

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

dual problem: maximize $-b^T \nu - \|A^T \nu + c\|_1$

Problems with generalized inequalities

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

$\preceq_{K_i}$ is generalized inequality on $\mathbf{R}^{k_i}$

definitions are parallel to scalar case:

- Lagrange multiplier for $f_i(x) \preceq_{K_i} 0$ is vector $\lambda_i \in \mathbf{R}^{k_i}$

- Lagrangian $L : \mathbf{R}^n \times \mathbf{R}^{k_1} \times \cdots \times \mathbf{R}^{k_m} \times \mathbf{R}^p \to \mathbf{R}$, is defined as

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

- dual function $g : \mathbf{R}^{k_1} \times \cdots \times \mathbf{R}^{k_m} \times \mathbf{R}^p \to \mathbf{R}$, is defined as

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

lower bound property: if $\lambda_i \succeq_{K_i^*} 0$, then $g(\lambda_1, \ldots, \lambda_m, \nu) \leq p^\star$

proof: if $\tilde{x}$ is feasible and $\lambda \succeq_{K_i^*} 0$, then

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

minimizing over all feasible $\tilde{x}$ gives $p^\star \geq g(\lambda_1, \ldots, \lambda_m, \nu)$

dual problem

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

- weak duality: $p^\star \geq d^\star$ always
- strong duality: $p^\star = d^\star$ for convex problem with constraint qualification (for example, Slater's: primal problem is strictly feasible)

Semidefinite program

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

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

- Lagrange multiplier is matrix $Z \in \mathbf{S}^k$

矩阵内积

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

- dual function

$$g(Z) = \inf_x L(x, Z) = \begin{cases} -\mathbf{tr}(GZ) & \mathbf{tr}(F_i Z) + c_i = 0, \quad i = 1, \ldots, n \\ -\infty & \text{otherwise} \end{cases}$$

dual SDP

$$\begin{array}{ll} \text{maximize} & -\mathbf{tr}(GZ) \\ \text{subject to} & Z \succeq 0, \quad \mathbf{tr}(F_i Z) + c_i = 0, \quad i = 1, \ldots, n \end{array}$$

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