5. Duality

Xinfu Liu

Outline

- 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)

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

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

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

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

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

Lagrange dual function

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

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

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

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

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

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

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

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:

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

- plug in L to obtain g:

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

a concave function of ν

lower bound property: $p^* \geq -\frac{1}{4} \nu^T A A^T \nu - b^T \nu$ for all ν

Standard form LP

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

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

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

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

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

Equality constrained norm minimization

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

dual function

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

where $\|v\|_* = \sup\{u^T v \mid \|\mathrm{u}\| \leq 1\}$ 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 $\|u\| \leq 1, \sup_u u^T y = \|y\|_* > 1$:

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

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

Two-way partitioning

$$\begin{array}{ll} \text{Minimize} & x^T W x \\ \text{Subject to} & x_i^2 = 1, \qquad i = 1, \dots, n \end{array}$$

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

Dual function

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

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

Lower bound property: $p^* \geq -\mathbf{1}^T \nu$ if $W + \boldsymbol{diag}(\nu) \succcurlyeq 0$

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

Lagrange dual function and conjugate function

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

dual function

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

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

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

example: entropy maximization

$$\min f_0(x) = \sum_{i=1}^{n} x_i \log x_i$$
$$\text{s.t. } Ax \preccurlyeq b$$
$$1^T x = 1$$

conjugate function of f_0: $\quad f_0^*(y) = \sum_{i=1}^{n} e^{y_i - 1}$

dual function

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

The dual problem

Lagrange dual problem

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

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

example: standard form **LP** and its dual problem

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

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**

  for example, solving the SDP

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

  gives a lower bound for the two-way partitioning problem

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 condition

strong duality holds for a convex problem

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

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

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

- also guarantees that **the dual optimum** is attained (if $p^* \succ -\infty$)
- can be refined: *e.g.*, **linear inequalities do not need to hold with strict inequality**,...
- there exist many other types of constraint qualifications

Inequality form LP

primal problem

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

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

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

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

Quadratic program

primal problem (assume $P \in \boldsymbol{S}_{++}^n$)

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

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

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

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

A nonconvex problem with strong duality

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

$A \in S^n$ and $A \nsucceq 0$, hence nonconvex

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

- unbounded below if $A + \lambda I \nsucceq 0$ or if $A + \lambda I \succcurlyeq 0$ and $b \notin \mathcal{R}(A + \lambda I)$
- minimized by $x = -(A + \lambda I)^\dagger b$ otherwise

  and $g(\lambda) = -b^T(A + \lambda I)^\dagger b - \lambda$, where $(A + \lambda I)^\dagger$ is the pseudo-inverse

dual problem and equivalent SDP:

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

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

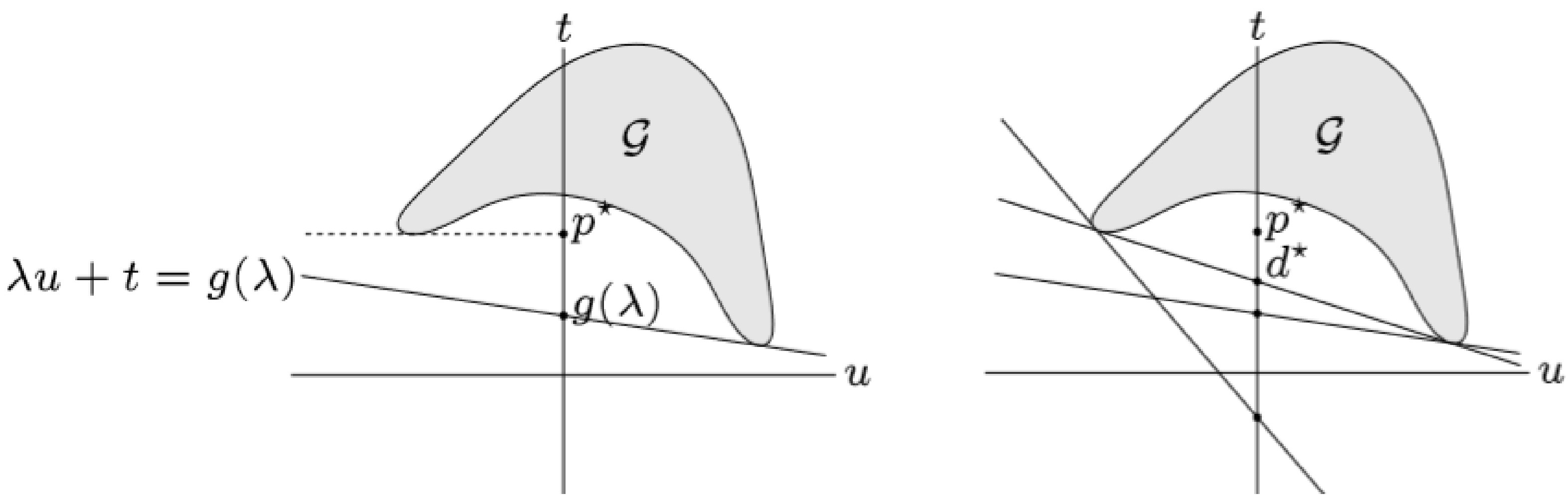
Geometric interpretation

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

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

interpretation of dual function:

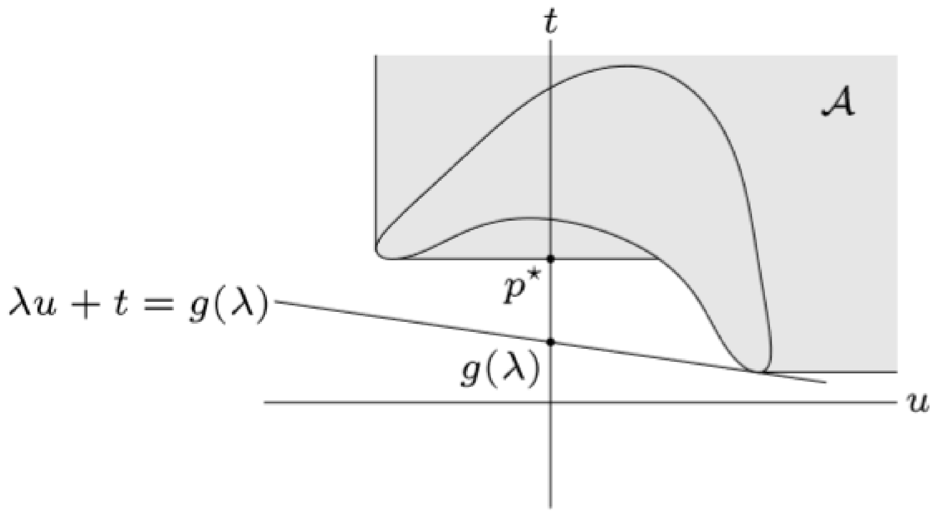
$$g(\lambda) = \inf_{(u,t) \in \mathcal{G}} (t + \lambda u), \quad \text{where} \quad \mathcal{G} = \{(f_1(x), f_0(x)) | x \in \mathcal{D}\}$$



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

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

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



strong duality

- holds if there is a non-vertical supporting hyperplane to $\mathcal{A}$ at $(0, p^*)$
- for convex problem, $\mathcal{A}$ is convex, hence has supp. Hyperplane at $(0, p^*)$

Complementary slackness

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

$$
\begin{aligned}
f_0(x^*) &= g(\lambda^*, \nu^*) \\
&= \inf_x \left(f_0(x) + \textstyle\sum_{i=1}^m \lambda_i^* f_i(x) + \sum_{i=1}^p \nu_i^* h_i(x)\right) \\
&\le f_0(x^*) + \sum_{i=1}^m \lambda_i^* f_i(x^*) + \sum_{i=1}^p \nu_i^* h_i(x^*) \\
&\le f_0(x^*)
\end{aligned}
$$

hence, the two inequalities hold with equality

- x^* minimizes $L(x, \lambda^*, \nu^*)$
- $\lambda_i^* f_i(x^*) = 0$ for $i = 1, \ldots, m$ (known as **complementary slackness**):

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

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: $f_i(x) \leq 0, i = 1, \dots, m, h_i(x) = 0, i = 1, \dots, p$
2. dual constraints: $\lambda \succcurlyeq 0$
3. complementary slackness: $\lambda_i f_i(x) = 0, i = 1, \dots, m$
4. gradient of Lagrangian with respect to x vanishes:

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

For **any optimization problem** with differentiable f_i, h_i, if **strong duality holds** and x, λ, ν are optimal, then **they must satisfy the KKT conditions**

KKT conditions for convex problem

For a convex problem, if $\tilde{x}, \tilde{\lambda}, \tilde{\nu}$ satisfy the KKT conditions, then they are optimal (KKT are also **sufficient conditions**):

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

hence, $f_0(\tilde{x}) = \mathrm{g}(\tilde{\lambda}, \tilde{\nu}) \rightarrow$ **zero duality gap** $\rightarrow$ primal and dual optimal

Necessary and sufficient conditions for optimality:

If a convex optimization problem satisfies **Slater's condition**, then x is optimal **if and only** if there exist λ, ν, that together with x, satisfy the KKT conditions

- Recall that Slater's condition implies strong duality, and dual optimum is attained
- Generalizes optimality condition $\nabla f_0(x) = 0$ for unconstrained problem

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

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

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

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

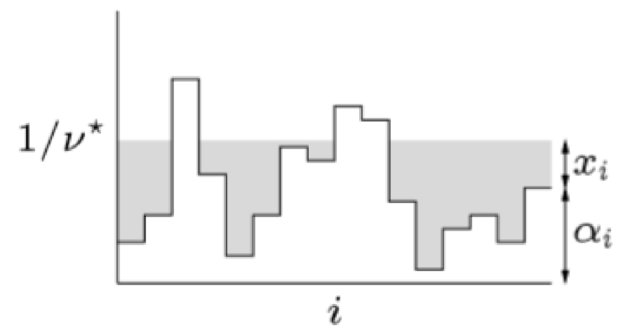
- If $\nu < 1/\alpha_i$: $\lambda_i = 0$ and $x_i = 1/\nu - \alpha_i$
- If $\nu \geq 1/\alpha_i$: $\lambda_i = \nu - 1/\alpha_i$ and $x_i = 0$

that is $x_i^* = \begin{cases} 1/\nu - \alpha_i & \nu < 1/\alpha_i \\ 0 & \nu \geq 1/\alpha_i \end{cases}$

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

Interpretation (why called water-filling)

- n patches; level of patch i is at height α_i
- Flood area with unit amount of water
- Resulting level is $1/\nu^*$



Perturbation and sensitivity analysis

(unperturbed) optimization problem and its dual

minimize	$f_0(x)$	maximize	$g(\lambda, \nu)$
subject to	$f_i(x) \leq 0, i = 1, \dots, m$	Subject to	$\lambda \succcurlyeq 0$
	$h_i(x) = 0, i = 1, \dots, p$		

perturbed problem and its dual

min.	$f_0(x)$	maximize	$g(\lambda, \nu) - \lambda^T u - \nu^T v$
s.t.	$f_i(x) \leq u_i, i = 1, \dots, m$	Subject to	$\lambda \succcurlyeq 0$
	$h_i(x) = v_i, i = 1, \dots, p$		

- x is primal variable; u, v are parameters
- $p^*(u, v)$ is optimal value as a function of u, v
- We are interested in information about $p^*(u, v)$ 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 λ^*, ν^* are dual optimal for **unperturbed problem**

apply weak duality to perturbed problem:

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

sensitivity interpretation

- if λ_i^* large: p^* increases greatly if we tighten constraint i $(u_i < 0)$
- if λ_i^* small: p^* does not decrease much if we loosen constraint i $(u_i > 0)$
- if ν_i^* large and positive: p^* increases greatly if we take $v_i < 0$

  if ν_i^* large and negative: p^* increases greatly if we take $v_i > 0$
- if ν_i^* small and positive: p^* does not decrease much if we take $v_i > 0$

  if ν_i^* small and negative: p^* does not decrease much if we take $v_i < 0$

local sensitivity

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

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

proof (for λ_i^*): from global sensitivity result,

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

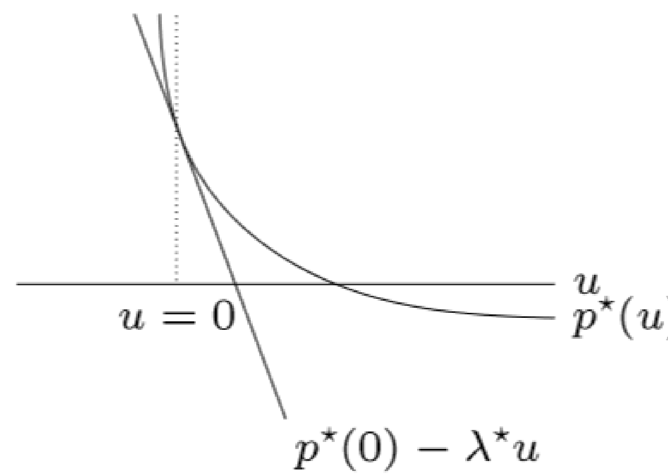
hence, equality holds

method can be used to establish

The same $\dfrac{\partial p^*(0,0)}{\partial v_i} = -v_i^*$

$p^*(u)$ for a convex problem with one (inequality

constraint is shown in the right figure



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 associated 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 and increasing

Introducing new variables and equality constraints

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

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

reformulated problem and its dual

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

dual function follows from

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

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

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

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

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

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

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

reformulation with box constraints made implicit

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

dual function

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

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

Problems with generalized inequalities

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

$\preccurlyeq_{K_i}$ is generalized inequality on $\boldsymbol{R}^{k_i}$

definitions are parallel to scalar case:

- Lagrangian multiplier for $f_i(x) \preccurlyeq_{K_i} 0$ is vector $\lambda_i \in \boldsymbol{R}^{k_i}$
- Lagrangian L: $\boldsymbol{R}^n \times \boldsymbol{R}^{k_1} \times \cdots \times \boldsymbol{R}^{k_m} \times \boldsymbol{R}^p \to \boldsymbol{R}$, is defined as

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

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

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

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

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

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

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

dual problem

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

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

Semidefinite program

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

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

- Lagrange multiplier is **matrix** $Z \in \boldsymbol{S}^k$
- Lagrangian $L(x, Z) = c^T x + \boldsymbol{tr}\big(Z(x_1 F_1 + \cdots + x_n F_n - G)\big)$
- dual function

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

dual SDP

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

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