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)

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

f_i's, h_i's ~~f_i's~~ not necessarily convex, affine

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 $h_i(x) = 0$

• constraint func. are soft now

• benefit for ~~comp~~ $f_i(x), h_i(x) \le 0$ ~~being~~ (if $\lambda_i, \nu_i \ge 0$)

Lagrange dual function

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

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

bc ① is affine in λ, ν \+ pointwise infimum of affine is concave

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

oppo. of pointwise sup.

• $\mathcal{D}$ does not have to be convex

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

• lower bound on optimal value

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

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

bc $\sum_{i=1}^{m} \lambda_i f_i(x) \le 0$ and $\sum_{i=1}^{p} \nu h_i(x) = 0$

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

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:

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

- plug ① in in L to obtain g: bc $g(\lambda,\nu) = \inf_{x \in \mathcal{D}} L(x,\lambda,\nu)$

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

a concave function of ν

↳ we know dual func. always concave

NSD

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

• LB isn't helpful for small problem w/ analytic sol'n but can help figure out when to terminate large iterative approach

Standard form LP

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

$-x \preceq 0$

dual function

- Lagrangian is

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

- L is affine in x, hence

• line has non $-\infty$ min iff slope $= 0$

$-bc$ $-b^T\nu$

$$g(\lambda, \nu) = \inf_x L(x, \lambda, \nu) = \begin{cases} -b^T \nu & A^T \nu - \lambda + c = 0 \; ① \\ -\infty & \text{otherwise} \end{cases}$$

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 $\underbrace{A^T \nu + c \succeq 0}$

• any ν s.t. $A^T\nu + c \succeq 0$ $\rightarrow$ $-b^T\nu$ is lower bound for the LP

① holds when $\lambda = A^T\nu + c$ and lower bound prop. holds for $\lambda \succeq 0$!

• any standard LP has this form lower bound prop.

Equality constrained norm minimization

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

• no square term + applies to any norm

dual function

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

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

• Hard problem

$\sum_j \sum_i x_i x_j W_{ij}$

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

~~if a point is assigned~~

• W_{ij} is an "irritation" matrix b/t all pairs of points.

- 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

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

min of quad form is 0 when PSD + $-\infty$ otherwise

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

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

dual function

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

- 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 kown

example: entropy maximization

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

Key Takeaways
1. Use original problem to form Lagrange dual problem
2. Optimize Lagrange dual func. to get the best lower bound on the original problem possible from the dual.
• Dual problem always convex.

The dual problem

Lagrange dual problem

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

this is the dual func.

highest lower bound, nothing to do w/ tightness

- 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, actually not dual but equiv. to dual (page 1–5)

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

i.e., find the best lower bound

• lower bound prop!
$p^\star \geq -b^T \nu$ if
$A^T \nu + c \succeq 0$

Weak and strong duality

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

- 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 1–7

strong duality: $d^\star = p^\star$

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

• i.e., "if primal problem is convex and (constraint qualifications), ~~then~~ then $p^\star = d^\star$"

Slater's constraint qualification

strong duality holds for a convex problem

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

if it is strictly feasible, *i.e.*, concerned w/ only inequal constraints

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

- also guarantees that the dual optimum is attained from the dual problem (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, . . .
- there exist many other types of constraint qualifications

• Holds for almost everything practical

Inequality form LP

primal problem

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

dual function

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

equiv. to
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 $\tilde{x}$ i.e., nonempty interior
- in fact, $p^\star = d^\star$ for LP except when primal and dual are infeasible

Quadratic program

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

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

dual function

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

dual problem – actual dual problem bc dom $g = \mathbb{R}^n$

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

- 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

A nonconvex problem with strong duality

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

$A \not\succeq 0$, hence nonconvex

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

- unbounded below if $A + \lambda I \not\succeq 0$ 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)

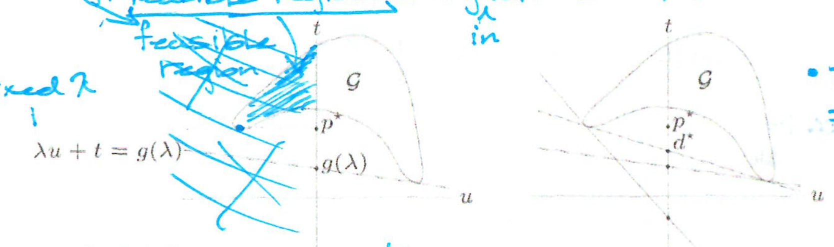
Geometric interpretation

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

interpretation of dual function:

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

actually, feasible region is only left side of t AND in G



• $p^* - d^*$ = duality gap

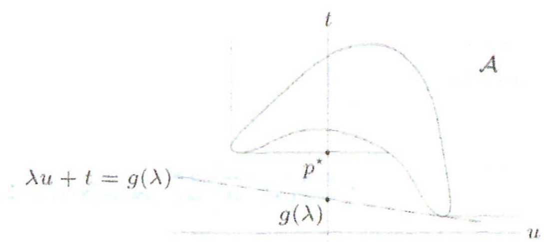
• height corresponds to objective value

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

• $g(\lambda)$ affected by infeasible part of domain too (at least here it is)

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

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



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

Complementary slackness

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

$$\begin{aligned} f_0(x^\star) = g(\lambda^\star, \nu^\star) &= \inf_x \left(f_0(x) + \sum_{i=1}^m \lambda_i^\star f_i(x) + \sum_{i=1}^p \nu_i^\star h_i(x) \right) \\ &\le f_0(x^\star) + \underbrace{\sum_{i=1}^m \underbrace{\lambda_i^\star}_{\ge 0} \underbrace{f_i(x^\star)}_{\le 0}}_{\le 0} + \underbrace{\sum_{i=1}^p \nu_i^\star h_i(x^\star)}_{=0} \\ &\le f_0(x^\star) \end{aligned}$$

certificates of optimality (that $x^\star$ is optimal)

=0 (after inf.)

≤ 0 → every term in sum must $= 0$.

hence, the two inequalities hold with equality

- $x^\star$ minimizes $L(x, \lambda^\star, \nu^\star)$
- $\lambda_i^\star f_i(x^\star) = 0$ for $i = 1, \ldots, m$ (known as complementary slackness): - can't have λ_i, $f_i(x^\star)$ non-zero for same i

$$\underbrace{\lambda_i^\star > 0 \Longrightarrow f_i(x^\star) = 0}_{\text{tight}}, \qquad \underbrace{f_i(x^\star) < 0 \Longrightarrow \lambda_i^\star = 0}_{\text{loose}}$$

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

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

$\nabla_x L = 0$

• ∃ generalization of KKT to nondifferen. cases. • Our sol'n is to transform non-diff prob. into a diff one.

① from page 1–17: if strong duality holds and x, λ, ν are optimal, then they must satisfy the KKT conditions

• For convex or nonconvex problem

• analog of: if x is a minimum of a nonconvex differentiable func., the gradient at that point vanishes (converse not true)

KKT conditions for convex problem

② if $\tilde{x}$, $\tilde{\lambda}$, $\tilde{\nu}$ satisfy KKT for a convex problem, then they are primal+dual $\tilde{x}, \tilde{\lambda}, \tilde{\nu}$ optimal:

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

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

~~• $\nabla_x L$ is convex~~

• Lagrangian is convex for a convex problem

• 4 says L differentiable and where gradient vanishes → point it vanishes must be minimizer → $\inf_x L = g$

(strictly feasible convex problem → strong duality)

③ 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 $\alpha_i > 0$)

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

using ③

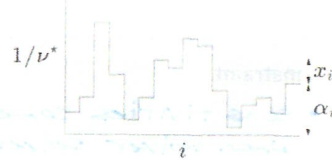
x is optimal iff $x \succeq 0$, $\mathbf{1}^T x = 1$, and there exist $\lambda \in \mathbf{R}^n$, $\nu \in \mathbf{R}$ such that

$$\lambda \succeq 0, \qquad \lambda_i x_i = 0, \qquad \underbrace{\frac{1}{x_i + \alpha_i} + \lambda_i = \nu}_{\text{from KKT condition 4}}$$

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



• one of the very few problems where the KKT conditions leads to the sol'n of the problem

Perturbation and sensitivity analysis

(unperturbed) optimization problem and its dual

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

• very useful

perturbed problem and its dual

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

• $\uparrow u_i \to$ relaxing, $\downarrow u_i \to$ tightening

• v_i doesn't have same interp. Just shifting affine constraint in inequal. constraint polyhedron.

- x is primal variable; u, v are parameters
- $p^\star(u, v)$ is optimal value as a function of u, v — of perturbed problem
- we are interested in information about $p^\star(u, v)$ that we can obtain from the solution of the unperturbed problem and its dual

~~convex~~ perturbed problem is convex in u, v

optimal value if you change constraints of problem and reoptimize

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:

LHS always worse than RHS

$$\begin{aligned} p^\star(u, v) &\ge g(\lambda^\star, \nu^\star) - u^T\lambda^\star - v^T\nu^\star \\ &= \underbrace{p^\star(0, 0) - u^T\lambda^\star - v^T\nu^\star}_{\text{affine}} \end{aligned}$$

affine $\to$ RHS is hyperplane $\le$ LHS (see next slide pic.)

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 $v_i < 0$

~~if for any wiggle of the RHS (constant) of an inequality constraint s.t. $p^\star$ is unaffected~~ $\uparrow$ tightness

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,

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

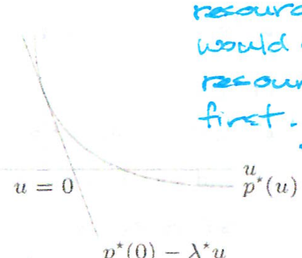
$$\frac{\partial p^\star(0,0)}{\partial u_i} = \lim_{t \nearrow 0} \frac{p^\star(te_i, 0) - p^\star(0,0)}{t} \le -\lambda_i^\star$$

hence, equality

• $f_1(x^\star) = 0$ — $\lambda_i^\star$: 0.001
$f_2(x^\star) = 0$ — 10
$f_3(x^\star) = 0$ — 0.02
$\to$ 2nd constraint is "much tighter" than others

• Want more of resource 2 and would give up resource 1 first.

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



• global sensitivity result gives lower bound on $\uparrow p^\star$ if we tighten (~~RH~~ $\downarrow$ RHS) any constraint

• sensitivities come for free when solver solves problem

• to manually verify solver solution, check 1. $\tilde{x}$ feasible, 2. $\tilde{\lambda}_i \ge 0$, 3. KKT 4th condition (if diff. $f_0(\tilde{x})$)
$\hookrightarrow$ if $\nabla_x L = 0$, then check $\tilde{x}(\tilde{x})$

Duality and problem reformulations

$P \to D$
$\downarrow$
$\tilde{P} \to \tilde{D}$

- 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

• Main idea: P + trivial formulation of P will likely have very diff. duals w/o ~~obv.~~ relationship

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

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

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

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

(see page 1–4)

dual of norm approximation problem

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

Implicit constraints

LP with box constraints: primal and dual problem

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

for same component i of λ_1 and λ_2, both cannot be simul. positive bc $\lambda_{1i} > 0 \rightarrow x_i = -1$ & $\lambda_{2i} > 0 \rightarrow x_i = 1$ (from complementary slackness)

reformulation with box constraints made implicit

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

dual function

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

~~attaches~~ pulled out from this

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

* Duality for feasibility problems is a topic not covered in lecture but is in the book. 1-2 HW problems on this

Problems with generalized inequalities

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

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

definitions are parallel to scalar case: in negative cone

- 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

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

dual cone

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

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

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

inner prod. of ~~things~~ something in + dual cone + something in cone is nonneg. Nonpos when latter in negative cone.

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

dual problem

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

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

- ~~Mon~~ Everything from b4 still works (lower bound property, Slater, etc.)

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 $Z \in \mathbf{S}^k$
- Lagrangian $L(x, Z) = c^T x + \mathbf{tr}\left(Z(x_1 F_1 + \cdots + x_n F_n - G)\right)$ — affine in x
- dual function

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

dual SDP

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

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