Lecture 13
Duality

– Lagrangian duality
– interpretations
– examples

Lagrangian

standard form problem

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

– optimal value $p^\star$, domain D
– called **primal problem** (in context of duality)

(for now) we **don't** assume convexity

Lagrangian $L : \mathbf{R}^{n+m} \to \mathbf{R}$

$$L(x, \lambda) = f_0(x) + \lambda_1 f_1(x) + \cdots + \lambda_m f_m(x)$$

– λ_i called *Lagrange multipliers* or *dual variables*
– objective is *augmented* with weighted sum of constraint functions

Weak duality

fact:

$$\sup_{\lambda \geq 0} L(x,\lambda) = \begin{cases} f_0(x) & \text{if } f_i(x) \leq 0, \quad i=1,\dots,m \\ +\infty & \text{else} \end{cases}$$

hence original problem can be cast as **unconstrained** pb

$$p^* = \inf_x \sup_{\lambda \geq 0} L(x,\lambda)$$

interpretation: $\sup_{\lambda \geq 0} L(x,\lambda)$ acts as a barrier function

weak duality: exchange **inf** and **sup** and get **lower bound** on p^*:

$$d^* := \sup_{\lambda \geq 0} \inf_x L(x,\lambda) \leq \inf_x \sup_{\lambda \geq 0} L(x,\lambda) = p^*$$

recall: hols even if original problem is not convex!

Lagrange dual function

(Lagrange) dual function $g : \mathbf{R}^m \to \mathbf{R} \cup \{-\infty\}$

$$\begin{aligned} g(\lambda) &= \inf_x L(x,\lambda) \\ &= \inf_x \left(f_0(x) + \lambda_1 f_1(x) + \cdots + \lambda_m f_m(x)\right) \end{aligned}$$

- minimum of augmented cost as function of weights
- can be $-\infty$ for some λ
- g is concave (even if f_i not convex!)

Lagrange dual problem: find **best** lower bound on $p^\star$:

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

- called **(Lagrange) dual problem** (associated with primal problem)
- always a convex problem, even if primal isn't!
- optimal value denoted $d^\star$
- we always have $d^\star \leq p^\star$ (*weak duality*)
- $p^\star - d^\star$ is *optimal duality gap*

Strong duality

for **convex** problems, we usually (not always) have *strong duality* at the optimum, *i.e.*:

$$d^* = p^*$$

(*i.e.*, *zero duality gap*)

for **convex** problems, a sufficient conditions for zero duality gap is *Slater's* condition:

> if the primal problem is convex, and is **strictly feasible,** *i.e.*
>
> there exist x_0 such that $f_i(x_0) < 0, \quad i = 1, \dots, m$
>
> then $p^* = d^*$

condition only sufficient—many alternative, sufficient "constraint qualification" conditions

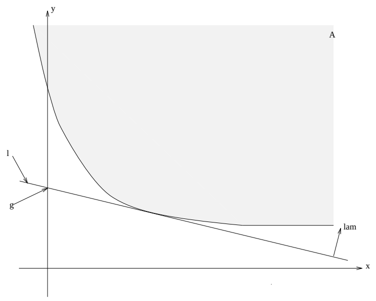
Geometric interpretation of duality

consider set

$$\mathcal{A} = \{\ (u,t) \in \mathbf{R}^{m+1} \mid \exists x\ f_i(x) \le u_i, \quad f_0(x) \le t\ \}$$

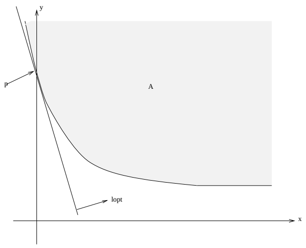
- $\mathcal{A}$ is convex if f_i are
- for $\lambda \succeq 0$,

$$g(\lambda) = \mathbf{inf}\left\{ \begin{bmatrix} \lambda \\ 1 \end{bmatrix}^T \begin{bmatrix} u \\ t \end{bmatrix} \;\middle|\; \begin{bmatrix} u \\ t \end{bmatrix} \in \mathcal{A} \right\}$$



(Idea of) proof of Slater's theorem

problem convex, strictly feasible $\Longrightarrow$ strong duality



- $(0, p^\star) \in \partial\mathcal{A} \Rightarrow \exists$ supporting hyperplane at $(0, p^\star)$:

$$(u, t) \in \mathcal{A} \Longrightarrow \mu_0(t - p^\star) + \mu^T u \geq 0$$

- $\mu_0 \geq 0$, $\mu \succeq 0$, $(\mu, \mu_0) \neq 0$
- strong duality $\Leftrightarrow \exists$ supp. hyperpl. with $\mu_0 > 0$:
  for $\lambda^\star = \mu/\mu_0$, we have

$$\begin{aligned} p^\star &\leq t + \lambda^{\star T} u \;\; \forall (t, u) \in \mathcal{A} \\ p^\star &\leq g(\lambda^*) \end{aligned}$$

- Slater's condition: all supporting hyperplanes at $(0, p^\star)$ are non-vertical ($\mu_0 > 0$)

Dual of a linear program

(primal) LP

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

- n variables, m inequality constraints

dual of LP is (after making implicit equality constraints explicit)

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

- dual of LP is also an LP (indeed, in std LP format)
- m variables, n equality constraints, m nonnegativity contraints

for LP we have strong duality except in one (pathological) case: primal and dual *both* infeasible ($p^\star = +\infty$, $d^\star = -\infty$)

Dual of quadratic program

(primal) QP

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

we assume $P \succ 0$ for simplicity

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

$\nabla_x L(x, \lambda) = 0$ yields $x = -(1/2)P^{-1}A^T\lambda$, hence dual function is

$$g(\lambda) = -(1/4)\lambda^T A P^{-1} A^T \lambda - b^T \lambda$$

- concave quadratic function
- all $\lambda \succeq 0$ are dual feasible

dual of QP is

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

. . . another QP

Duality in algorithms

many algorithms produce at iteration k

- a primal feasible $x^{(k)}$
- and a dual feasible $\lambda^{(k)}$

with $f_0(x^{(k)}) - g(\lambda^{(k)}) \to 0$ as $k \to \infty$

hence at iteration k we **know** $p^\star \in \left[g(\lambda^{(k)}), f_0(x^{(k)})\right]$

- useful for stopping criteria
- algorithms that use dual solution are often more efficient (*e.g.*, LP)

Nonheuristic stopping criteria

absolute error $= f_0(x^{(k)}) - p^\star \le \epsilon$

stopping criterion:

$$\textbf{until } \left(f_0(x^{(k)}) - g(\lambda^{(k)}) \le \epsilon\right)$$

relative error $= \dfrac{f_0(x^{(k)}) - p^\star}{|p^\star|} \le \epsilon$

stopping criterion:

$$\textbf{until } \left(g(\lambda^{(k)}) > 0 \;\;\&\;\; \frac{f_0(x^{(k)}) - g(\lambda^{(k)})}{g(\lambda^{(k)})} \le \epsilon\right)$$

$$\textbf{or } \left(f_0(x^{(k)}) < 0 \;\;\&\;\; \frac{f_0(x^{(k)}) - g(\lambda^{(k)})}{-f_0(x^{(k)})} \le \epsilon\right)$$

achieve **target value** ℓ or, prove ℓ is unachievable (*i.e.*, determine either $p^\star \le \ell$ or $p^\star > \ell$)

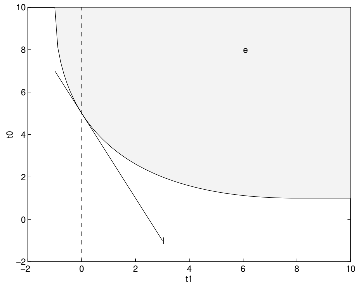
stopping criterion:

$$\textbf{until } \left(f_0(x^{(k)}) \le \ell \textbf{ or } g(\lambda^{(k)}) > \ell\right)$$

Sensitivity analysis via duality

define $p^\star(u)$ as the optimal value of

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



$\lambda^\star$ gives lower bound on $p^\star(u)$:

$$p^\star(u) \ge p^\star - \sum_{i=1}^m \lambda_i^\star u_i$$

- if $\lambda_i^\star$ large: $u_i < 0$ greatly increases $p^\star$
- if $\lambda_i^\star$ small: $u_i > 0$ does not decrease $p^\star$ too much

if $p^\star(u)$ is differentiable, $\lambda_i^\star = -\dfrac{\partial p^\star(0)}{\partial u_i}$

$\lambda_i^\star$ is sensitivity of $p^\star$ w.r.t. ith constraint

Equality constraints

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

- optimal value $p^\star$
- again assume (for now) not necessarily convex

define **Lagrangian** $L : \mathbf{R}^{n+m+p} \to \mathbf{R}$ as

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

dual function is $g(\lambda, \nu) = \mathbf{inf}_x\, L(x, \lambda, \nu)$

(λ, ν) is dual feasible if $\lambda \succeq 0$ and $g(\lambda, \nu) > -\infty$

(no sign condition on ν)

lower bound property: if x is primal feasible and (λ, ν) is dual feasible, then $g(\lambda, \nu) \leq f_0(x)$, hence

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

dual problem: find best lower bound

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

(note ν unconstrained), optimal value $d^\star$

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

strong duality: if primal is convex then (usually) $d^\star = p^\star$

Slater condition: if primal is convex (*i.e.*, f_i convex, h_i affine) and strictly feasible, *i.e.*, there exists $x \in \mathbf{relint} D$ s.t.

$$f_i(x) < 0, \quad h_i(x) = 0,$$

then $d^\star = p^\star$

Example: equality constrained least-squares

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

A is fat, full rank (solution is $x^\star = A^T(AA^T)^{-1}b$)

dual function is

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

dual problem is

$$\text{maximize} \;\; -\tfrac{1}{4}\nu^T AA^T \nu - b^T \nu$$

solution: $\nu^\star = -2(AA^T)^{-1}b$

can check $d^\star = p^\star$

Introducing equality constraints

idea: simple transformation of primal problem can lead to very different dual

example: unconstrained geometric programming

primal problem:

$$\text{minimize} \;\; \log \sum_{i=1}^m \exp(a_i^T x - b_i)$$

dual function is constant $g = p^\star$

(we have strong duality, but it's useless)

now **rewrite primal problem** as

$$\begin{array}{ll} \text{minimize} & \log \sum_{i=1}^m \exp y_i \\ \text{subject to} & y = Ax - b \end{array}$$

- introduce m new vbles $y_1, \ldots, y_m$
- introduce m new equality constraints $y = Ax - b$

dual function

$$g(\nu) = \inf_{x,y} \left(\log \sum_{i=1}^{m} \exp y_i + \nu^T (Ax - b - y) \right)$$

- infimum is $-\infty$ if $A^T \nu \neq 0$
- assuming $A^T \nu = 0$, let's minimize over y:

$$\frac{e^{y_i}}{\sum_{j=1}^{n} e^{y_j}} = \nu_i$$

  solvable iff $\nu_i > 0$, $\mathbf{1}^T \nu = 1$

$$g(\nu) = -\sum_i \nu_i \log \nu_i - b^T \nu$$

- same expression if $\nu \succeq 0$, $\mathbf{1}^T \nu = 1$ $\big(0 \log 0 = 0\big)$

dual problem

$$\begin{array}{ll} \text{maximize} & -b^T \nu - \sum_i \nu_i \log \nu_i \\ \text{subject to} & \mathbf{1}^T \nu = 1, \quad (\nu \succeq 0) \\ & A^T \nu = 0 \end{array}$$

moral: trivial reformulation can yield different dual

Generalized inequalities

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

- $\preceq_{K_i}$ are generalized inequalities on $\mathbf{R}^{m_i}$
- $f_i : \mathbf{R}^n \to \mathbf{R}^{m_i}$ are K_i-convex

Lagrangian $L : \mathbf{R}^n \times \mathbf{R}^{m_1} \times \cdots \times \mathbf{R}^{m_L} \to \mathbf{R}$,

$$L(x, \lambda_1, \ldots, \lambda_L) = f_0(x) + \lambda_1^T f_1(x) + \cdots + \lambda_L^T f_L(x)$$

dual function

$$g(\lambda_1, \ldots, \lambda_L) = \inf_x \left(f_0(x) + \lambda_1^T f_1(x) + \cdots + \lambda_L^T f_L(x) \right)$$

λ_i **dual feasible** if $\lambda_i \succeq_{K_i^\star} 0$, $g(\lambda_1, \ldots, \lambda_L) > -\infty$

lower bound property: if x primal feasible and $(\lambda_1, \ldots, \lambda_L)$ is dual feasible, then

$$g(\lambda_1, \ldots, \lambda_L) \leq f_0(x)$$

(hence, $g(\lambda_1, \ldots, \lambda_L) \leq p^\star$)

dual problem

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

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

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

Slater condition: if primal is strictly feasible, *i.e.*,

$$\exists x \in \textbf{relint} D : \quad f_i(x) \prec_{K_i} 0, \quad i = 1, \ldots, L$$

then $d^\star = p^\star$

Example: semidefinite programming

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & F_0 + x_1 F_1 + \cdots + x_n F_n \preceq 0 \end{array}$$

Lagrangian (multiplier $Z = Z^T \in \mathbf{R}^{m \times m}$)

$$L(x, Z) = c^T x + \textbf{Tr}\, Z(F_0 + x_1 F_1 + \cdots + x_n F_n)$$

dual function

$$\begin{aligned} g(Z) &= \inf_x \left(c^T x + \textbf{Tr}\, Z(F_0 + x_1 F_1 + \cdots + x_n F_n) \right) \\ &= \begin{cases} \textbf{Tr}\, F_0 Z & \text{if } \textbf{Tr}\, F_i Z + c_i = 0, \quad i = 1, \ldots, n \\ -\infty & \text{otherwise} \end{cases} \end{aligned}$$

dual problem

$$\begin{array}{ll} \text{maximize} & \textbf{Tr}\, F_0 Z \\ \text{subject to} & \textbf{Tr}\, F_i Z + c_i = 0, \quad i = 1, \ldots, n \\ & Z = Z^T \succeq 0 \end{array}$$

strong duality holds if there exists x with

$$F_0 + x_1 F_1 + \cdots + x_n F_n \prec 0$$