15.S08 Recitation 4: Nonlinear Duality

Shuvomoy Das Gupta

Topics to discuss

- Lagrangian for a nonlinear problem
- Lagrange dual function
- The dual problem
- Weak and strong duality
- Complementary slackness
- KKT conditions
- Sensitivity analysis: marginal price interpretation

Lagrangian

standard form problem (not necessarily convex)

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

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

Lagrange dual function

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

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

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

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

proof: if $\tilde{x}$ is feasible and $\lambda \succeq 0$, 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 $p^\star \geq g(\lambda, \nu)$

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

Standard form LP and its dual

Primal

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

Dual

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

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

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

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

Complementary slackness

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

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

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):

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

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

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

Perturbation and sensitivity analysis

(unperturbed) optimization problem and its dual

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

perturbed problem and its dual

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

- 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:

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

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: