Duality

Operations Research

Anthony Papavasiliou

Table of Contents

Lagrangian Function

Standard form problem (not necessarily convex):

$$\begin{aligned} &\min f_0(x) \\ &\text{s.t. } f_i(x) \leq 0, i = 1, \ldots, m \\ &h_i(x) = 0, i = 1, \ldots, p \end{aligned}$$

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

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

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

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

Dual Function

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

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

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

Dual Function is Concave

Consider any (λ_1, ν_1), (λ_2, ν_2) with $\lambda_1, \lambda_2 \geq 0$ and $\alpha \in [0, 1]$

$$
\begin{aligned}
& g(\alpha\lambda_1 + (1-\alpha)\lambda_2, \alpha\nu_1 + (1-\alpha)\nu_2) \\
&= \inf_{x\in\mathcal{D}} (f_0(x) + \sum_{i=1}^{m} (\alpha\lambda_{1,i} f_i(x) + (1-\alpha)\lambda_{2,i} f_i(x)) \\
&+ \sum_{i=1}^{p} (\alpha\nu_{1,i} h_i(x) + (1-\alpha)\nu_{2,i} h_i(x))) \\
&\geq \alpha \inf_{x\in\mathcal{D}} (f_0(x) + \sum_{i=1}^{m} \lambda_{1,i} f_i(x) + \sum_{i=1}^{p} \nu_{1,i} h_i(x)) \\
&+ (1-\alpha) \inf_{x\in\mathcal{D}} (f_0(x) + \sum_{i=1}^{m} \lambda_{2,i} f_i(x) + \sum_{i=1}^{p} \nu_{2,i} h_i(x)) \\
&= \alpha g(\lambda_1, \nu_1) + (1-\alpha) g(\lambda_2, \nu_2)
\end{aligned}
$$

Dual Function is a Lower Bound

If $\lambda \geq 0$ then $g(\lambda, \nu) \leq p^\star$

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

Lagrange Relaxation of Stochastic Programs

Consider 2-stage stochastic program:

$$\begin{aligned} &\min f1(x) + \mathbb{E}_\omega[f2(y(\omega),\omega)] \\ &\text{s.t. } h1_i(x) \leq 0, i = 1,\ldots,m_1, \\ &h2_i(x, y(\omega),\omega) \leq 0, i = 1,\ldots,m_2 \end{aligned}$$

Introduce **non-anticipativity constraint** $x(\omega) = x$ and reformulate problem as

$$\begin{aligned} &\min f1(x) + \mathbb{E}_\omega[f2(y(\omega),\omega)] \\ &\text{s.t. } h1_i(x) \leq 0, i = 1,\ldots,m_1, \\ &h2_i(x(\omega), y(\omega),\omega) \leq 0, i = 1,\ldots,m_2 \\ (\nu(\omega)) : \quad & x(\omega) = x \end{aligned}$$

Dual Function of Stochastic Program

$$g(\nu) = g1(\nu) + \mathbb{E}_\omega g2(\nu(\omega), \omega)$$

where

$$g1(\nu) = \inf f1(x) + (\sum_{\omega \in \Omega} \nu(\omega))^T x$$
$$\text{s.t. } h1_i(x) \leq 0, i = 1, \ldots, m_1,$$

and

$$g2(\nu, \omega) = \inf f2(y(\omega), \omega) - \nu x(\omega)$$
$$\text{s.t. } h2_i(x(\omega), y(\omega), \omega) \leq 0, i = 1, \ldots, m_2$$

Table of Contents

The Dual Problem

Lagrange dual problem:

$$d^\star = \max g(\lambda, \nu)$$
$$\text{s.t. } \lambda \geq 0$$

- Finds best lower bound on $p^\star$ from Lagrangian dual function
- Convex optimization problem with optimal value $d^\star$
- λ, ν are dual feasible if $\lambda \geq 0$, $(\lambda, \nu) \in \text{dom } g$

Weak and Strong Duality

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

- Always holds (for convex and non-convex problems)
- Can be used for finding non-trivial bounds to difficult problems

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

- Does not hold in general
- Usually holds for convex problems
- Conditions that guarantee strong duality in convex problems are called constraint qualifications

Linear Programming Duality Mnemonic Table

Primal	Minimize	Maximize	Dual
Constraints	$\geq b_i$	≥ 0	Variables
	$\leq b_i$	≤ 0	
	$= b_i$	Free	
Variables	≥ 0	$\leq c_j$	Constraints
	≤ 0	$\geq c_j$	
	Free	$= c_j$	

Prove the mnemonic table using Lagrangian relaxation

Table of Contents

Complementary Slackness

If strong duality holds, $x^\star$ primal optimal, $\lambda^\star, \nu^\star$ dual optimal

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

Therefore, the two inequalities above hold with equality and

- $x^\star$ minimizes $L(x, \lambda^\star, \nu^\star)$
- $\lambda_i^\star f_i(x^\star) = 0$ for $i = 1, \ldots, m$

This is known as **complementary slackness**:

$$
\lambda_i^\star > 0 \Rightarrow f_i(x^\star) = 0 \quad f_i(x^\star) < 0 \Rightarrow \lambda_i^\star = 0
$$

KKT Conditions

KKT conditions for a problem with differentiable f_i, h_i:

- Primal constraints: $f_i(x) \le 0$, $i = 1, \dots, m$, $h_i(x) = 0, i = 1, \dots, p$
- Dual constraints: $\lambda \ge 0$
- Complementary slackness: $\lambda_i f_i(x) = 0$, $i = 1, \dots, m$
- Gradient of the Lagrangian function with respect to x vanishes:

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

From previous slide, if strong duality holds and x, λ, ν are optimal, then they must satisfy the KKT conditions

KKT Conditions of Maximization with Linear Constraints

Consider a maximization problem with linear constraints:

$$
\begin{aligned}
& \max \quad f(x, y) \\
(\lambda): \quad & Ax + By \leq b \\
(\mu): \quad & Cx + Dy = d \\
(\lambda_2): \quad & x \geq 0
\end{aligned}
$$

Then the KKT conditions have the following form:

$$
\begin{aligned}
& Cx + Dy - d = 0 \\
& 0 \leq \lambda \perp Ax + By - b \leq 0 \\
& 0 \leq x \perp \lambda^T A + \mu^T C - \nabla_x f(x, y)^T \geq 0 \\
& \lambda^T B + \mu^T D - \nabla_y f(x, y)^T = 0
\end{aligned}
$$

Example: The Diet Problem - KKT Conditions

Consider the diet problem with $b_1 = 1$ and $b_2 = 2$:

$$\begin{aligned} & \min x_1 + 2x_2 + x_3 \\ (\pi_1): \quad & 0.5x_1 + 4x_2 + x_3 = 1 \\ (\pi_2): \quad & 2x_1 + x_2 + 2x_3 = 2 \\ & x_1, x_2, x_3 \geq 0 \end{aligned}$$

KKT conditions:

$$0.5x_1 + 4x_2 + x_3 = 1 \tag{1}$$

$$2x_1 + x_2 + 2x_3 = 2 \tag{2}$$

$$0 \leq x_1 \perp 0.5\pi_1 + 2\pi_2 + 1 \geq 0 \tag{3}$$

$$0 \leq x_2 \perp 4\pi_1 + \pi_2 + 2 \geq 0 \tag{4}$$

$$0 \leq x_3 \perp \pi_1 + 2\pi_2 + 1 \geq 0 \tag{5}$$

Example: The Diet Problem - Equivalent KKT Systems

Note: Since $h(x) = 0 \Leftrightarrow -h(x) = 0$, we can get a different KKT system with a different solution $\pi^\star$:

$$0.5x_1 + 4x_2 + x_3 = 1 \tag{6}$$

$$2x_1 + x_2 + 2x_3 = 2 \tag{7}$$

$$0 \leq x_1 \perp -0.5\pi_1 - 2\pi_2 + 1 \geq 0 \tag{8}$$

$$0 \leq x_2 \perp -4\pi_1 - \pi_2 + 2 \geq 0 \tag{9}$$

$$0 \leq x_3 \perp -\pi_1 - 2\pi_2 + 1 \geq 0 \tag{10}$$

Claim:

- Primal optimal solution: $(x^\star)^T = (0, 0.1429, 0.4286, 0)$
- Dual optimal solution: $(\pi^\star)^T = (0.4286, 0.2857)$

Proof: verify that $x^\star$ and $\pi^\star$ satisfy KKT conditions

Table of Contents

Non-Uniqueness of KKT Conditions

1. The KKT conditions of a problem depend on how we define the Lagrangian function
2. The sign of dual multipliers depends on the KKT conditions (therefore, how we define the Lagrangian function)
3. The sensitivity interpretation of dual multipliers depends on the KKT conditions (therefore, how we define the Lagrangian function)
4. Different software interprets user syntax differently!

Dual Multipliers in AMPL

In order to be able to anticipate the sign of multipliers that AMPL will assign to constraints, note that:

- A constraint of the form $f_1(x) \leq, =, \geq f_2(x)$ is equivalently expressed as $f_1(x) - f_2(x) \leq, =, \geq 0$,
- the constraints are relaxed by *subtracting* their product with their corresponding multiplier from the Lagrangian function,
- the sign of the dual multiplier is such that the Lagrangian function provides a bound to the optimization problem,
- the primal-dual optimal pair is such that the KKT conditions corresponding to this Lagrangian function are satisfied.
- In this way, the dual multipliers reported by AMPL can always be interpreted as sensitivities.

Example

$$\{\min x + 2y \text{ s.t. } 0 \leq x, (\lambda_1), x \leq 2, (\lambda_2), y = 1, (\mu)\}$$

Objective function $f(x, y) = x + 2y$, inequality constraints $f_1(x, y) = -x \leq 0$ (i.e., a $\leq$ constraint), $f_2(x, y) = x - 2$, $h(x, y) = y - 1$

- AMPL Lagrangian:
  $L(x, y) = (x + 2y) - \lambda_1(-x) - \lambda_2(x - 2) - \mu(y - 1)$

AMPL KKT Conditions

KKT conditions:

- Primal feasibility: $g_1(x, y) \leq 0$, $g_2(x, y) \leq 0$, $h(x, y) = 0$
- Dual feasibility: $\lambda_1 \leq 0$, $\lambda_2 \leq 0$
- Complementarity: $\lambda_1 \perp g_1(x, y)$, $\lambda_2 \perp g_2(x, y)$
- Stationarity:

$$\nabla f(x, y) - \lambda_1 \nabla g_1(x, y) - \lambda_2 \nabla g_2(x, y) - \mu \nabla h(x, y) = 0$$

Solution: $x = 0, y = 1, \lambda_1 = -1, \lambda_2 = 0, \mu = 2$