Duality

Quantitative Energy Economics

Anthony Papavasiliou

Contents

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: $\mathcal{L} : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^p \to \mathbb{R}$,

dom $\mathcal{L} = \mathcal{D} \times \mathbb{R}^m \times \mathbb{R}^p$:

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

- 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: $g : \mathbb{R}^m \times \mathbb{R}^p \to \mathbb{R}$,

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

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

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

Dual Function is Concave

Consider any (λ_1, ν_1), (λ_2, ν_2) and $\alpha \in [0, 1]$:

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

Agent Coordination

Consider set of agents G with private cost $f_g(x_g)$, private constraints $h2_g(x_g) \leq 0$

$$
\begin{aligned}
&\min \sum_{g \in G} f_g(x_g) \\
&\text{s.t.} \sum_{g \in G} h1_g(x_g) = 0 \\
&h2_g(x_g) \leq 0
\end{aligned}
$$

Relax coordination constraints $\sum_{g \in G} h1_g(x_g) = 0$:

$$L(x, \lambda) = \sum_{g \in G} (f_g(p_g) + \lambda^T h1_g(x_g))$$

$$g(\lambda) = \sum_{g \in G} \inf_{h2_g(x_g) \leq 0} (f_g(p_g) + \lambda^T h1_g(x_g))$$

Table of Contents

The Dual Problem

Lagrange dual problem:

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

- 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

Example: Dual Problem of Unit Commitment

Satisfy demand of 200 MW using the following technologies:

Generator	Activation cost ($/h)	Marg. cost ($/MWh)	Capacity (MW)
Cheap	500	0	20
Moderate	1000	10	100
Expensive	2000	80	100

Example: Dual Problem of Unit Commitment

Introduce the following variables:

- p_i: power production of unit i
- u_i (binary): indicator variable for activation of unit i

$$
\begin{aligned}
& \min 500 \cdot u_1 + 1000 \cdot u_2 + 10 \cdot p_2 + 2000 \cdot u_3 + 80 \cdot p_3 \\
(\lambda): \quad & p_1 + p_2 + p_3 = 200 \qquad (1) \\
& 0 \leq p_1 \leq 20 \cdot u_1 \\
& 0 \leq p_2 \leq 100 \cdot u_2 \\
& 0 \leq p_3 \leq 100 \cdot u_3 \\
& u_i \in \{0, 1\}
\end{aligned}
$$

Which constraint makes generator decisions depend on each other?

Example: Dual Problem of Unit Commitment

Dual function obtained by relaxing constraint (1):

$$g(\lambda) = \min 500 \cdot u_1 + 1000 \cdot u_2 + 10 \cdot p_2 + 2000 \cdot u_3 + 80 \cdot p_3$$

$$-\lambda \cdot (p_1 + p_2 + p_3)$$

$$p_1 \leq 20 \cdot u_1, p_2 \leq 100 \cdot u_2, p_3 \leq 100 \cdot u_3$$

$$p_i, \geq 0, u_i \in \{0, 1\}$$

Thus,

$$g(\lambda) = g_1(\lambda) + g_2(\lambda) + g_3(\lambda),$$

where

$$g_1(\lambda) = \min_{u_1 \in \{0,1\}} \{500 \cdot u_1 - \lambda \cdot p_1, 0 \leq p_1 \leq 20 \cdot u_1\}$$

$$g_2(\lambda) = \min_{u_2 \in \{0,1\}} \{1000 \cdot u_2 + (10 - \lambda) \cdot p_2, 0 \leq p_2 \leq 100 \cdot u_2\}$$

$$g_3(\lambda) = \min_{u_3 \in \{0,1\}} \{2000 \cdot u_3 + (80 - \lambda) \cdot p_3, 0 \leq p_3 \leq 100 \cdot u_3\}$$

Example: Dual Problem of Unit Commitment

Computing $g_1(\lambda)$ (similarly for $g_2(\lambda), g_3(\lambda)$):

- $\lambda \geq 25 \Rightarrow u_1^\star = 1$, $p_1^\star = 20$
- $\lambda < 25 \Rightarrow u^\star = 0$, $p^\star = 0$

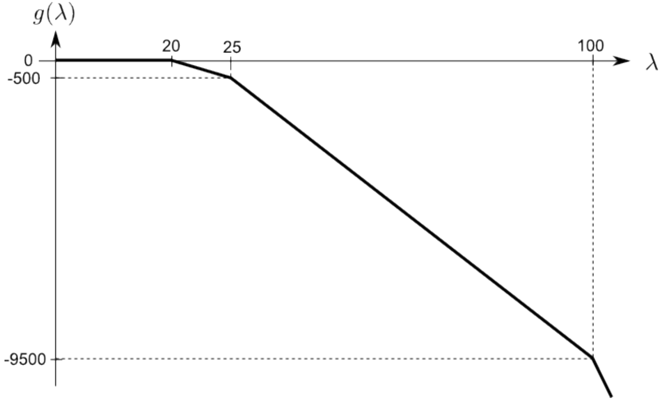
$$g_1(\lambda) = \begin{cases} 0, & \lambda \leq 25 \\ 500 - 20 \cdot \lambda, & \lambda > 25 \end{cases}$$

Finally:

$$g(\lambda) = \begin{cases} 0, & \lambda \leq 20 \\ 2000 - 100 \cdot \lambda, & 20 < \lambda \leq 25 \\ 2500 - 120 \cdot \lambda, & 25 < \lambda \leq 100 \\ 12500 - 220 \cdot \lambda, & 100 < \lambda \end{cases}$$

Dual Problem of Unit Commitment

Sanity check: $g(\lambda)$ is concave



- Primal optimal solution: $u^\star = (0, 1, 1)$ and $p^\star = (0, 100, 100) \Rightarrow p^\star = 12000$
- $d^\star = 0 < p^\star \Rightarrow$ strong duality does not hold

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 $\mathcal{L}(x, \lambda^\star, \nu^\star)$
- $\lambda_i^\star f_i(x^\star) = 0$ 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: $f_i(x) \leq 0$, $i = 1, \ldots, m$, $h_i(x) = 0, i = 1, \ldots, p$
- Dual constraints: $\lambda \geq 0$
- Complementary slackness: $\lambda_i f_i(x) = 0$, $i = 1, \ldots, m$
- 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 for Convex Problem

- Strong duality usually holds for convex problems (but not always)
- Conditions that ensure strong duality are called **constraint qualifications**
- If (i) constraints of an optimization problem are all linear equalities inequalities and (ii) and dom f_0 is open, then strong duality holds.

KKT Conditions of Maximization with Linear Constraints

Consider a maximization problem with linear constraints:

$$\max \quad f(x, y)$$
$$Ax + By \leq b, \quad (\lambda)$$
$$Cx + Dy = d, \quad (\mu)$$
$$x \geq 0, \quad (\lambda_2)$$

Then the KKT conditions have the following form:

$$Cx - d = 0$$
$$0 \leq \lambda \perp Ax - b \leq 0$$
$$0 \leq x \perp \lambda^T A + \mu^T C - \nabla_x f(x, y)^T \geq 0$$
$$\lambda^T B + \mu^T D - \nabla_y f(x, y)^T = 0$$

Example: KKT Conditions for Dispatch Problem

Consider previous example, without activation costs

Generator	Marg. cost ($/MWh)	Capacity (MW)
Cheap	0	20
Moderate	10	100
Expensive	80	100

$$
\begin{aligned}
& \min 10 \cdot p_2 + 80 \cdot p_3 \\
(\lambda): \quad & p_1 + p_2 + p_3 = 200 \\
(\mu_1): \quad & p_1 \leq 20 \\
(\mu_2): \quad & p_2 \leq 100 \\
(\mu_3): \quad & p_3 \leq 100 \\
& p_i \geq 0
\end{aligned}
$$

Example: KKT Conditions for Dispatch Problem

KKT conditions:

- Primal equality constraint
- Primal inequality constraint $\perp$ complementary non-negative dual variable
- Primal non-negative variable $\perp$ dual inequality constraint

$$p_1 + p_2 + p_3 = 200 \tag{2}$$

$$0 \leq \mu_1 \perp 20 - p_1 \geq 0 \tag{3}$$

$$0 \leq \mu_2 \perp 100 - p_2 \geq 0 \tag{4}$$

$$0 \leq \mu_3 \perp 100 - p_3 \geq 0 \tag{5}$$

$$0 \leq p_1 \perp \lambda + \mu_1 \geq 0 \tag{6}$$

$$0 \leq p_2 \perp 10 + \lambda + \mu_2 \geq 0 \tag{7}$$

$$0 \leq p_3 \perp 80 + \lambda + \mu_3 \geq 0 \tag{8}$$

Example: KKT Conditions for Dispatch Problem

$p_1 + p_2 = p_3 = 200 \Leftrightarrow -p_1 - p_2 - p_3 = -200$
Therefore, three last conditions can be replaced by:

$$0 \leq p_1 \quad \perp \quad -\lambda + \mu_1 \geq 0 \tag{9}$$

$$0 \leq p_2 \quad \perp \quad 10 - \lambda + \mu_2 \geq 0 \tag{10}$$

$$0 \leq p_3 \quad \perp \quad 80 - \lambda + \mu_3 \geq 0 \tag{11}$$

- Easy to see that $(p^\star)^T = (20, 100, 80)$ is primal optimal
- Claim: $\lambda^\star = 80$ and $(\mu^\star)^T = (80, 70, 0)$ are dual optimal
- Proof: verify that $p^\star, \lambda^\star$ and $\mu^\star$ satisfy equations (2) - (5) and (9) - (11)

KKT Conditions for Non-Differentiable Optimization

What if f_0, f_i, h_i are convex but non-differentiable?

If strong duality holds,

- $f_i(x) \leq 0$, $i = 1, \ldots, m$, $h_i(x) = 0, i = 1, \ldots, p$
- $\lambda \geq 0$
- $\lambda_i f_i(x) = 0$, $i = 1, \ldots, m$
- Subgradient of the Lagrangian function with respect to x vanishes:

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

where $\partial f(x)$ denotes a subgradient of f at x

Table of Contents

Subgradients

Consider a function g, π is a **subgradient** of g at u if

$$g(w) \geq g(u) + \pi^T(w - u) \text{ for all } w$$

Subgradients generalize gradients for non-differentiable functions

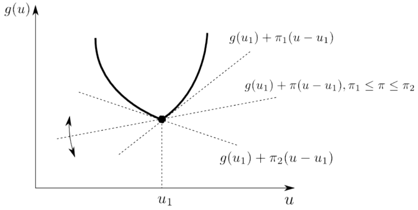
Subdifferential $\partial g(u)$: set of all subgradients at u

Subgradients are useful for

- generalizing KKT conditions to non-differentiable optimization problems
- deriving sensitivity results

Geometric Interpretation of Subgradeints

Subgradient determines linear under-estimator of a function



- π_1: unique subgradient at u_1
- π_2 and π_3: both subgradients at u_2

Subgradient Calculus

Suppose g is convex, then:

- $\partial g(u) = \{\nabla g(u)\}$ if g is differentiable at u
- Conversely, if $\partial g(u) = \{\pi\}$, then g is differentiable at u and $\pi = \nabla g(u)$
- $\partial(ag) = a\partial g$.
- $\partial(g_1 + g_2) = \partial g_1 + \partial g_2$, where the right hand side corresponds to addition of sets
- If $f(u) = g(Au + b)$, then $\partial f(u) = A^T \partial g(Au + b)$
- If $g = \max_{i=1,\ldots,m} g_i$, then

$$\partial g(u) = \text{Co}(\cup\{\partial g_i(u) | g_i(u) = g(u)\}),$$

where $\text{Co}(\cdot)$ is the convex hull

Example: Subgradient Calculus

Consider the following function:

$$g(u) = \max\{g_1(u), g_2(u), g_3(u), g_4(u)\}$$

where

$$\begin{aligned} g_1(u) &= 0, \\ g_2(u) &= 100 \cdot u - 2000, \\ g_3(u) &= 120 \cdot u - 2500, \\ g_4(u) &= 220 \cdot u - 12500. \end{aligned}$$

- At $u = 25$, $g_2(u)$ and $g_3(u)$ are the only active inequalities
- According to the last result of the previous slide, $\partial g(25) = [100, 120]$

Example: Subgradient Calculus

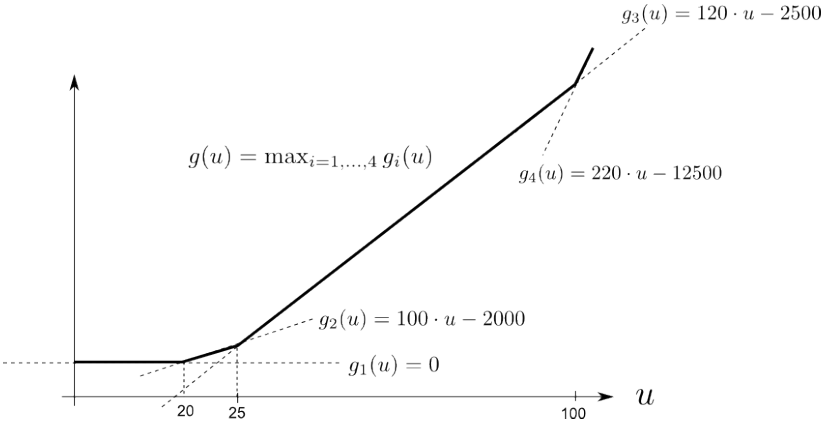


Figure: The subdifferential of g at $u = 25$ is $\partial g(25) = [100, 120]$ since g_2 and g_3 are the only active constraints.

Sensitivity Result

Define $c(u)$ as the optimal value of the following mathematical program:

$$c(u) = \min f_0(x)$$
$$f_i(x) \leq u_i, i = 1, \ldots, m$$
$$x \in \text{dom } f_0$$

and suppose that dom f_0 is a convex set and and f_0, f_i are convex functions

Then,

- $c(u)$ is a convex function
- If strong duality holds and $\lambda^\star$ maximizes the dual function $\inf_{x \in \text{dom } f_0}(f_0(x) - \lambda^T(f(x) - u))$ for $\lambda \leq 0$, then $\lambda^\star \in \partial c(u)$

If $c(u)$ is differentiable at a certain point u, then for a given constraint i:

$$\lambda_i = \frac{\partial c(u)}{\partial u_i}$$

Conclusion: λ_i is equal to the *sensitivity* of the objective function $c(u)$ to a marginal change in the right-hand-side of the constraint corresponding to λ_i

Example: Convexity of $c(u)$

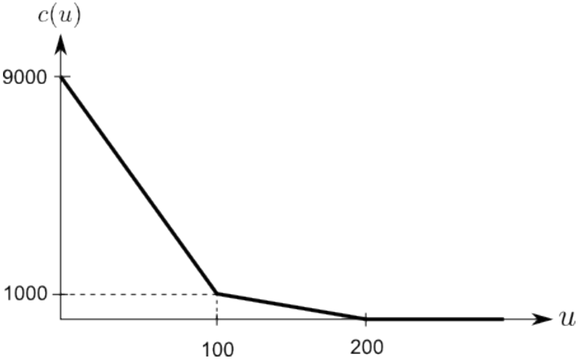
Generator	Marg. cost cost ($/MWh)	Capacity (MW)
Cheap	0	20
Moderate	10	100
Expensive	80	100

- Denote u as the capacity of generator 1
- Generally, generator 1 will be used to the greatest possible extent, followed by generator 2, followed by generator 3
- For $0 \leq u \leq 100$, $c(u) = 10 \cdot 1000 + 80 \cdot (100 - u)$

Example: Convexity of $c(u)$

Following same reasoning for $u \geq 100$:

$$c(u) = \begin{cases} 9000 - 80 \cdot u, & 0 \leq u < 100 \\ 2000 - 10 \cdot u, & 100 \leq u < 200 \\ 0, & 200 \leq u \end{cases}$$



Example: Sensitivity

Recall the solution of the KKT conditions (equations (2) - ((5) and (9) - (11)):

$$(p^\star)^T = (20, 100, 80), \lambda^\star = 80, (\mu^\star)^T = (80, 70, 0)$$

Sensitivity interpretation of $\lambda^\star$:
Right-hand-side of $p_1 + p_2 + p_3 = 200$, increases by one unit $\Rightarrow$ generator 3 increases output by 1 unit $\Rightarrow$ additional cost of 80 $

Example: Sensitivity

KKT conditions can also be expressed using equations (2) - (8)

Solution of the KKT system is

$$(p^{\star})^T = (20, 100, 80), \lambda^{\star} = -80, (\mu^{\star})^T = (80, 70, 0)$$

Note the change in the sign of $\lambda^{\star}$!

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