Mathematical Background

Anthony Papavasiliou, National Technical University of Athens (NTUA)

Source: chapter 2 Papavasiliou [1], chapter 5 Boyd [2]

Contents

- Lagrange dual problem
- Weak and strong duality
- Optimality conditions
- Sensitivity
- Dual multipliers in AMPL

Lagrange dual problem

Lagrange function

Standard form problem (not necessarily convex):

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

$x \in \mathbb{R}^n$, D is the domain of f_0, optimal value p^*

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

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

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

Dual function

Lagrange dual function: $g\colon\ \mathbb{R}^m \times \mathbb{R}^p \rightarrow \mathbb{R}$

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

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

Dual function is a lower bound

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

Proof: If $\bar{x}$ is feasible and $\lambda \geq 0$ then:

$$f_0(\bar{x}) \geq L(\bar{x}, \lambda, \nu) \geq \min_{x \in D} L(x, \lambda, \nu) = g(\lambda, \nu)$$

Minimizing over all feasible $\bar{x}$ gives $p^* \geq g(\lambda, \nu)$

Dual function is concave

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

$$g(\alpha\lambda_1 + (1-\alpha)\lambda_2, \alpha\nu_1 + (1-\alpha)\nu_2)$$

$$= \min_{x \in \mathrm{dom}\, f_0} \left(f_0(x) + \sum_{i=1}^{m} \alpha\lambda_{1,i} f_i(x) + (1-\alpha)\lambda_{2,i} f_i(x) + \sum_{i=1}^{p} \alpha\nu_{1,i} h_i(x) + (1-\alpha)\nu_{2,i} h_i(x) \right)$$

$$\geq \alpha \min_{x \in \mathrm{dom}\, f_0} \left(f_0(x) + \sum_{i=1}^{m} \lambda_{1,i} f_i(x) + \sum_{i=1}^{p} \nu_{1,i} h_i(x) \right)$$
$$+(1-\alpha) \min_{x \in \mathrm{dom}\, f_0} \left(f_0(x) + \sum_{i=1}^{m} \lambda_{2,i} f_i(x) + \sum_{i=1}^{p} \nu_{2,i} h_i(x) \right)$$

$$= \alpha g(\lambda_1, \nu_1) + (1-\alpha) g(\lambda_2, \nu_2)$$

Example 2.1: coordinating agents

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

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

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

$$L(x, \lambda) = \sum_{g \in G} \left(f_g(x_g) + \lambda^T h1_g(x_g) \right)$$
$$g(\lambda) = \sum_{g \in G} inf_{x_g : h2_g(x_g) \leq 0} \left(\left(f_g(x_g) + \lambda^T h1_g(x_g) \right) \right)$$

Weak and strong duality

The dual problem

Lagrange dual problem:

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

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

Weak and string duality

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

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

Strong duality: $p^* = d^*$

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

Example 2.2: linear programming duality

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 Lagrange relaxation

Example 2.3: dual problem of unit commitment

Satisfy demand of 200 MW using the following technologies

Generator	Activation cost ($/h)	Marginal cost ($/MWh)	Capacity (MW)
Low cost	500	0	20
Moderate cost	1000	10	100
High cost	2000	80	100

Example 2.3: 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

$$\min_{p,u} 500 \cdot u_1 + 1000 \cdot u_2 + 10 \cdot p_2 + 2000 \cdot u_3 + 80 \cdot p_3$$
$$(\lambda): p_1 + p_2 + p_3 = 200 \quad (1)$$
$$0 \leq p_1 \leq 20 \cdot u_1$$
$$0 \leq p_2 \leq 100 \cdot u_2$$
$$0 \leq p_3 \leq 100 \cdot u_3$$
$$u_i \in \{0,1\}$$

- Which constraint makes generator decisions depend on each other?

Example 2.3: dual problem of unit commitment

- Dual function obtained by relaxing constraint (1):

$$g(\lambda) = \min_{p,u} 500 \cdot u_1 + 1000 \cdot u_2 + 10 \cdot p_2 + 2000 \cdot u_3 + 80 \cdot p_3 - \lambda \cdot (p_1 + p_2 + p_3 - 200)$$

$$\text{s.t.}\ p_1 \leq 20 \cdot u_1, p_2 \leq 100 \cdot u_2, p_3 \leq 100 \cdot u_3$$

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

- Thus,

$$g(\lambda) = g_1(\lambda) + g_2(\lambda) + g_3(\lambda) + 200 \cdot \lambda$$

where

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

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

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

Example 2.3: dual of unit commitment problem

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

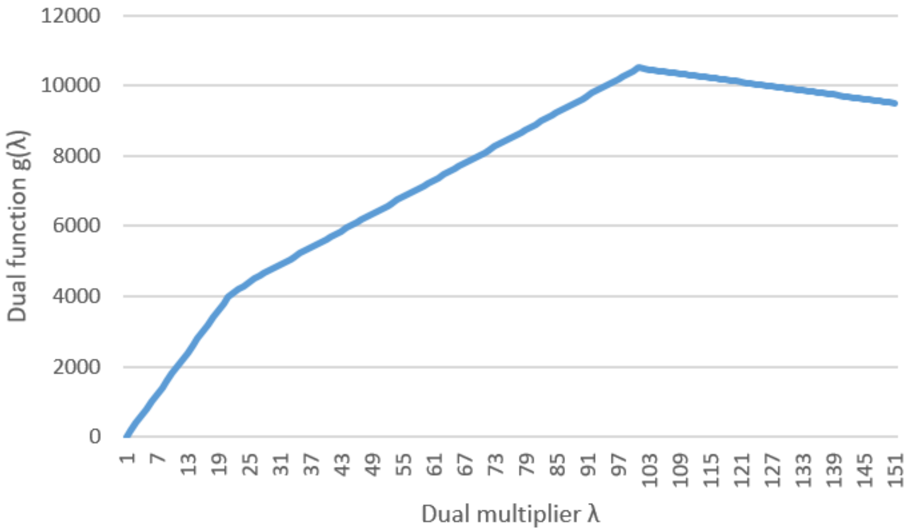
$$\lambda \geq 25 \rightarrow u_1^* = 1, p_1^* = 20$$
$$\lambda < 25 \rightarrow u_1^* = 0, p_1^* = 0$$
$$g_1(\lambda) = \begin{cases} 0, & \lambda \leq 25 \\ 500 - 20 \cdot \lambda, & \lambda > 25 \end{cases}$$

- Finally:

$$g(\lambda) = \begin{cases} 200 \cdot \lambda, & \lambda \leq 20 \\ 2000 + 100 \cdot \lambda, & 20 < \lambda \leq 25 \\ 2500 + 80 \cdot \lambda, & 25 < \lambda \leq 100 \\ 12500 - 20 \cdot \lambda, & 100 < \lambda \end{cases}$$

Example 2.3: dual problem of unit commitment

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



- Primal optimal solution: $u^* = (1,1,1)$ and $p^* = (20,100,80) \Rightarrow$ primal optimal equal to 10900
- Dual optimal equal to $10500 < 12000 \Rightarrow$ strong duality does not hold

Optimality conditions

Complementary slackness

- If strong duality holds, x^* primal optimal, λ^*, ν^* dual optimal

$$f_0(x^*) = g(\lambda^*, \nu^*) = \min_x \left(f_0(x) + \sum_{i=1}^{m} \lambda_i^* f_i(x) + \sum_{i=1}^{p} \nu_i^* h_i(x) \right)$$

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

$$\leq f_0(x^*)$$

Therefore, the two inequalities above hold with equality and

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

This is known as **complementary slackness**:

$$\lambda_i^* > 0 \Rightarrow f_i(x^*) = 0 \qquad f_i(x^*) < 0 \Rightarrow \lambda_i^* = 0$$

KKT conditions

KKT conditions for a problem with differentiable f_i, h_i:

- Primal constraints: $f_i(x) \leq 0, i = 1, \dots, m, h_i(x) = 0, i = 1, \dots, p$
- Dual constraints: $\lambda_i \geq 0, i = 1, \dots, m$
- Complementary slackness: $\lambda_i \cdot f_i(x) = 0, i = 1, \dots, m$
- 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 for convex problems

- Strong duality usually holds for convex problems (but not always)
- Conditions that ensure strong duality are called **constraint qualifications**
- If (i) constraints are linear equalities and inequalities and (ii) $\mathrm{dom}\ f_0$ is open, then strong duality holds

KKT conditions of maximization with linear constraints

- Consider a maximization problem with linear constraints:

$$\max_{x,y} c_x^T x + c_y^T y$$
$$\text{s.t. } (\lambda): Ax + By \leq b$$
$$(\mu): Cx + Dy = d$$
$$x \geq 0$$

- Then the KKT conditions have the following form:

$$Cx + Dy - d = 0$$
$$0 \leq \lambda \perp Ax + By - b \leq 0$$
$$0 \leq x \perp \lambda^T A + \mu^T C - c_x^T \geq 0$$
$$\lambda^T B + \mu^T D - c_y^T = 0$$

and are necessary and sufficient for an optimal solution

KKT conditions of minimization with linear constraints

- Consider a maximization problem with linear constraints:

$$\min_{x,y} c_x^T x + c_y^T y$$
$$\text{s.t. } (\lambda): Ax + By \leq b$$
$$(\mu): Cx + Dy = d$$
$$x \geq 0$$

- Then the KKT conditions have the following form:

$$Cx + Dy - d = 0$$
$$0 \leq \lambda \perp Ax + By - b \leq 0$$
$$0 \leq x \perp \lambda^T A + \mu^T C + c_x^T \geq 0$$
$$\lambda^T B + \mu^T D + c_y^T = 0$$

and are necessary and sufficient for an optimal solution

Example 2.4: KKT conditions for dispatch problem

Consider previous example, without activation costs

Generator	Marginal cost (€/MWh)	Capacity (MW)
Low cost	0	20
Moderate cost	10	100
High cost	80	100

$$\min 10 \cdot p_2 + 80 \cdot p_3$$
$$(\lambda): p_1 + p_2 + p_3 = 200$$
$$(\mu_1):\ p_1 \le 20$$
$$(\mu_2): p_2 \le 100$$
$$(\mu_3): p_3 \le 100$$
$$p_i \ge 0$$

Example 2.4: KKT conditions for dispatch problem

KKT conditions:

- Primal equality constraints
- Primal inequality constraints ⊥ (complementary to) non-negative dual variables
- Primal non-negative variables ⊥ (complementary to) dual inequality constraints

$$p_1 + p_2 + p_3 = 200 \quad (2)$$
$$0 \leq \mu_1 \perp 20 - p_1 \geq 0 \quad (3)$$
$$0 \leq \mu_2 \perp 100 - p_2 \geq 0 \quad (4)$$
$$0 \leq \mu_3 \perp 100 - p_3 \geq 0 \quad (5)$$
$$0 \leq p_1 \perp \lambda + \mu_1 \geq 0 \quad (6)$$
$$0 \leq p_2 \perp 10 + \lambda + \mu_2 \geq 0 \quad (7)$$
$$0 \leq p_3 \perp 80 + \lambda + \mu_3 \geq 0 \quad (8)$$

Example 2.4: KKT conditions for dispatch problem

$$p_1 + p_2 + p_3 = 200 \iff -p_1 - p_2 - p_3 = -200$$

- Therefore, three last conditions can be replaced by:

$$0 \leq p_1 \perp -\lambda + \mu_1 \geq 0 \quad (9)$$
$$0 \leq p_2 \perp 10 - \lambda + \mu_2 \geq 0 \quad (10)$$
$$0 \leq p_3 \perp 80 - \lambda + \mu_3 \geq 0 \quad (11)$$

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

KKT conditions for non-differentiable optimization problems

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

If strong duality holds, then:

- $f_i(x) \leq 0, i = 1, \ldots, m, h_i(x) = 0, i = 1, \ldots, p$
- $\lambda \geq 0$
- $\lambda_i f_i(x) = 0, i = 1, \ldots, m$
- 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

Sensitivity

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

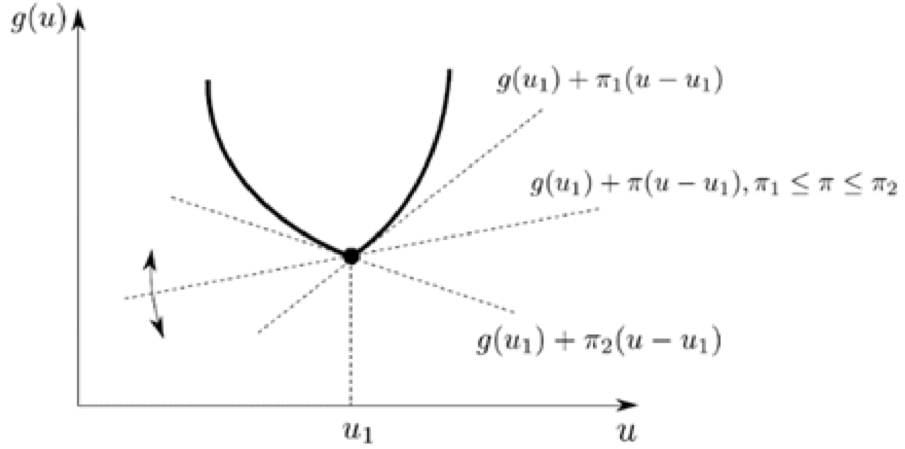
Subdifferential $\partial g(u)$: set of all subgradients at u

Subgradients are useful for:

- Generalizing KKT conditions to non-differentiable optimization problems
- Deriving sensitivity results

Geometric interpretation of subgradients

Subgradient determines linear under-estimator of a function



- π_1 and π_2: subgradients at u_1

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

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

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

Sensitivity result

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

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

and suppose that $\mathrm{dom}\, f_0$ is a convex set and f_0, f_i are convex functions

Then:

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

- 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 2.5: convexity of $c(u)$

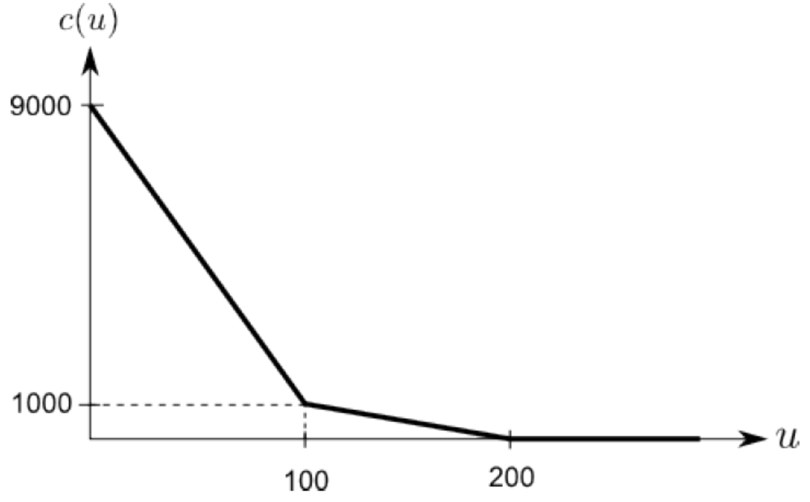
Generator	Marginal cost ($/MWh)	Capacity (MW)
Low cost	0	20
Moderate cost	10	100
High cost	80	100

- We return to example 2.4
- 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 $0 \leq u \leq 100$, $c(u) = 10 \cdot 100 + 80 \cdot (100 - u)$

Example 2.5: convexity of $c(u)$

Following the 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 2.6: slope of $c(u)$

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

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

Sensitivity interpretation of λ^*:

Right-hand side of $p_1 + p_2 + p_3 = 200$ increases by one unit ⇒ generator 3 increases output by 1 MW ⇒ additional cost of $80

Example 2.6: sensitivity

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

Solution of the KKT system is:

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

Note the change in the sign of λ^*!

Dual multipliers in AMPL

Μη-μοναδικότητα των συνθηκών KKT

- The KKT conditions of a problem depend on how we define the Lagrangian function
- The sign of dual multipliers depends on the KKT conditions (therefore, how we define the Lagrangian function)
- The sensitivity interpretation of dual multipliers depends on the KKT conditions (therefore, how we define the Lagrangian function)
- 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

$$\left\{\min_{x,y} x + 2y \; s.t. 0 \leq x, (\lambda_1), x \leq 2, (\lambda_2), y = 1, (\mu)\right\}$$

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

AMPL Lagrangian:

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

KKT conditions in AMPL

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:

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

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

References

[1] A. Papavasiliou, Optimization Models in Electricity Markets, Cambridge University Press

https://www.cambridge.org/highereducation/books/optimization-models-in-electricity-markets/0D2D36891FB5EB6AAC3A4EFC78A8F1D3#overview

[2] Boyd, Stephen P., and Lieven Vandenberghe. Convex optimization. Cambridge university press, 2004.