

2 Optimality and duality

Name

Salvador Pineda Morente

Learning objectives

- Derive and apply the optimality conditions for a convex optimization problem
- Formulate the Lagrange dual function and Lagrange dual problem of an optimization problem
- Explain the characteristics and usage of the dual problem
- Describe the weak and strong duality theorems
- Use the optimal values of dual variables for sensitivity analysis
- Formulate the KKT conditions of an optimization problem
- Derive the primal-dual formulation of a convex optimization problem

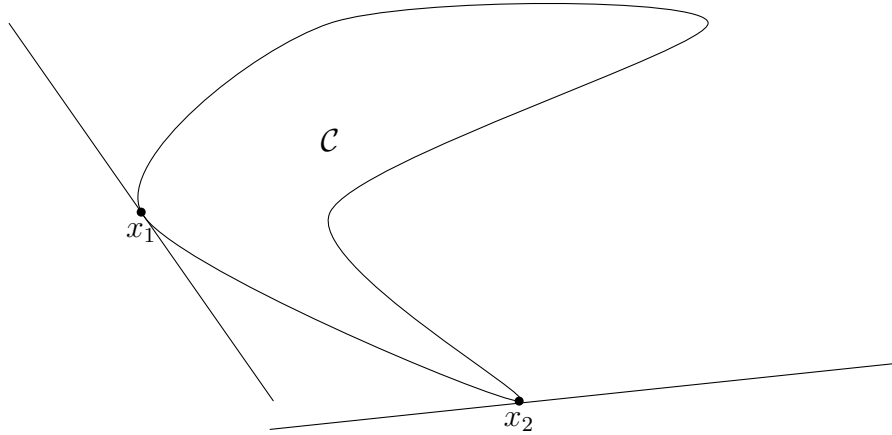
Bibliography

- Basic reading:
 - Boyd, Stephen P., and Lieven Vandenberghe. Convex optimization. Cambridge university press, 2004 (chapters 5).
- Further reading:
 - Motto, A.L.; Arroyo, J.M.; Galiana, F.D., “A Mixed-Integer LP Procedure for the Analysis of Electric Grid Security Under Disruptive Threat,” Power Systems, IEEE Transactions on , vol.20, no.3, pp.1357,1365, Aug. 2005.

Definition of a supporting hyperplane

Let $\mathcal{C}$ be a set in $\mathbb{R}^n$, $\tilde{x}$ a point in its boundary and $a \in \mathbb{R}^n, a \neq 0$. Then the hyperplane $\{x \in \mathbb{R}^n : a^T(x - \tilde{x}) = 0\}$ is a supporting hyperplane of the set $\mathcal{C}$ at the point $\tilde{x}$ if $a^T(x - \tilde{x}) \leq 0, \forall x \in \mathcal{C}$.

Draw supporting hyperplanes to set $\mathcal{C}$ at x_1 and x_2



Definition of sublevel set

The α -sublevel set of a function $f(x) : \mathbb{R}^n \rightarrow \mathbb{R}$ denoted as C_α is defined as:

$$C_\alpha = \{x \in \text{dom } f : f(x) \leq \alpha\}$$

Example 2-1: Sublevel sets of a convex function

Let $f(x) : \mathbb{R}^n \rightarrow \mathbb{R}$ be a convex function. Are its sublevel sets convex? Let us assume that

$$x_1 \in C_\alpha \implies f(x_1) \leq \alpha$$

$$x_2 \in C_\alpha \implies f(x_2) \leq \alpha$$

$$x_3 = \theta x_1 + (1 - \theta)x_2, \quad 0 \leq \theta \leq 1$$

The question is whether $x_3 \in C_\alpha$. Since $f(x)$ is convex we have

$$f(\theta x_1 + (1 - \theta)x_2) \leq \theta f(x_1) + (1 - \theta)f(x_2) \leq \alpha$$

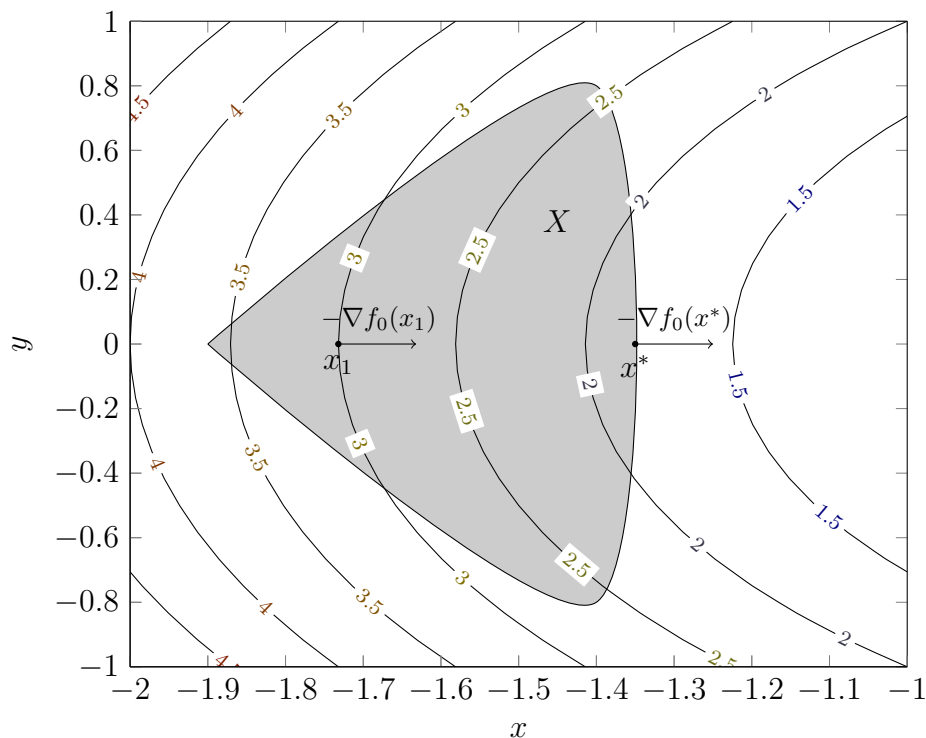
Therefore $x_3 \in C_\alpha$ which means that C_α is a convex set.

Example 2-2: Optimality conditions (I)

We have an optimization problem formulated as follows:

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & x \in X \end{array}$$

The figure below plots the convex feasibility set X and the contour lines of the convex objective function f_0 . Draw the vector $-\nabla f_0(x_1)$. Approximately indicate the optimal point x^* and draw the vector $-\nabla f_0(x^*)$.



- What is the sign of $-\nabla f_0^T(x^*)(x - x^*)$, $\forall x \in X$? $-\nabla f_0^T(x^*)(x - x^*) \leq 0$, $\forall x \in X$
- Prove that if $-\nabla f_0$ defines a supporting hyperplane of the set X at x^* , then x^* is optimal.
Tip: Use the definition of convex function.

Since $-\nabla f_0$ defines a supporting hyperplane of the set X at x^* , then we have that

$$-\nabla f_0^T(x^*)(x - x^*) \leq 0 \quad \forall x \in X$$

Since f_0 is convex, we also have that

$$f_0(x^*) \leq f_0(x) - \nabla f_0^T(x^*)(x - x^*) \quad \forall x \in X$$

Using both inequalities we obtain

$$f_0(x^*) \leq f_0(x) \quad \forall x \in X$$

which means that x^* is optimal.

📌 Example 2-3: Optimality conditions (II)

Which are the optimality conditions for an unconstrained problem, i.e., with $X = \mathbb{R}^n$?
If $X = \mathbb{R}^n$, then the optimality condition

$$f_0(x^*) \leq f_0(x) - \nabla f_0^T(x^*)(x - x^*) \quad \forall x \in X$$

is only satisfied for $\nabla f_0^T(x^*) = 0$. That is, the optimal points of unconstrained problems is achieved when the gradient vanishes.

📌 Example 2-4: Optimality conditions for generation scheduling

We consider a generating unit with whose production cost can be expressed as $\mathcal{C}(x) = ax^2 + bx + c$, where x represents the production level at MWh, and a , b and c are known positive parameters. We disregard any limit on the production and denote the electricity price by p . Formulate an optimization problem that determines the optimal generation level to maximize the profit made by the generating unit. Express the optimality condition for this unconstrained optimization problem and discuss its implications.

The optimization problem is

$$\underset{x}{\text{minimize}} \quad ax^2 + bx + c - px$$

The optimal points must thus satisfies that

$$\nabla f_0^T(x^*) = 0 \implies 2ax + b - p = 0 \implies x = \frac{p - b}{2a}$$

This means that the generating unit produces until the marginal cost equals the price.

✏ Solving Example 2-4

Following the optimality condition obtained in the previous exercise, fill in the following table with the optimal production levels for different prices for $a = 0.2$, $b = 10$ and $c = 100$. Afterwards, check your results by solving the code Scheduling.gms for very large values of $\underline{x}$ and $\bar{x}$.

p (€/MWh)	10	20	30	40	50
x (MWh)	0	25	50	75	100

The Lagrangian

Given an optimization problem:

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

where $x \in \mathbb{R}^n$ and $f_0(x), f_i(x), h_i(x) : \mathbb{R}^n \rightarrow \mathbb{R}$. We define the Lagrangian function $\mathcal{L} : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^p \rightarrow \mathbb{R}$ as

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

The vectors λ and ν are called Lagrange multipliers or dual variables.

Lagrangian of Example 1-13

Assuming $a = 0$, determine the Lagrangian corresponding to Example 1-13. Use $\underline{\eta}$ and $\bar{\eta}$ as the Lagrange multipliers corresponding to the production constraints.

$$\mathcal{L}(x, \underline{\eta}, \bar{\eta}) = bx + c + \underline{\eta}(\underline{x} - x) + \bar{\eta}(x - \bar{x}) - px$$

Lagrangian of Example 1-14

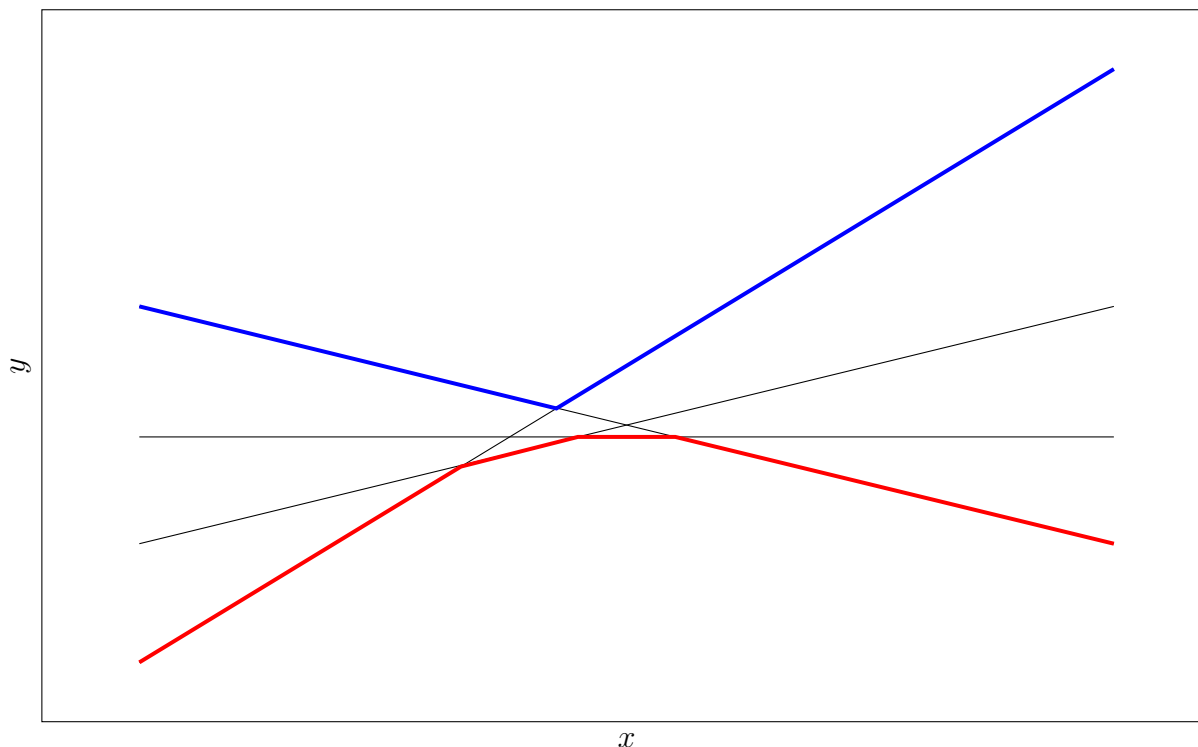
Assuming $a_g = 0$, determine the Lagrangian corresponding to Example 1-14. Use $\underline{\eta}_g$ and $\bar{\eta}_g$ as the Lagrange multipliers corresponding to production constraints, and λ as the Lagrange multiplier for the balance equation.

$$\mathcal{L}\left(x, \underline{\eta}_g, \bar{\eta}_g, \lambda\right) = \sum_g \left(b_g x_g + c_g + \underline{\eta}_g(\underline{x}_g - x_g) + \bar{\eta}_g(x_g - \bar{x}_g) \right) + \lambda \left(\sum_g x_g - d \right)$$



Pointwise maximum and infimum of affine functions

- Draw the pointwise maximum of the functions plotted below. Is that function convex?
- Draw the pointwise minimum of the functions plotted below. Is that function convex?



Example 2-5: Pointwise maximum convex functions

Let $f_1(x)$ and $f_2(x)$ be two convex functions. Let $g(x)$ be defined as the pointwise maximum of $f_1(x), f_2(x)$, i.e.,

$$g(x) = \max\{f_1(x), f_2(x)\}$$

Prove that $g(x)$ is a convex function.

$$\begin{aligned} g(\theta x_1 + (1 - \theta)x_2) &= \max\{f_1(\theta x_1 + (1 - \theta)x_2), f_2(\theta x_1 + (1 - \theta)x_2)\} \\ &\leq \max\{\theta f_1(x_1) + (1 - \theta)f_1(x_2), \theta f_2(x_1) + (1 - \theta)f_2(x_2)\} \\ &\leq \theta \max\{f_1(x_1), f_2(x_1)\} + (1 - \theta) \max\{f_1(x_2), f_2(x_2)\} \\ &= \theta g(x_1) + (1 - \theta)g(x_2) \end{aligned}$$

And therefore the function $g(x)$ is convex.

Example 2-6: Pointwise minimum of concave functions

Let $f_1(x)$ and $f_2(x)$ be two concave functions. Let $g(x)$ be defined as the pointwise minimum of $f_1(x), f_2(x)$, i.e.,

$$g(x) = \min\{f_1(x), f_2(x)\}$$

Prove that $g(x)$ is a concave function.

$$\begin{aligned} g(\theta x_1 + (1 - \theta)x_2) &= \min\{f_1(\theta x_1 + (1 - \theta)x_2), f_2(\theta x_1 + (1 - \theta)x_2)\} \\ &\geq \min\{\theta f_1(x_1) + (1 - \theta)f_1(x_2), \theta f_2(x_1) + (1 - \theta)f_2(x_2)\} \\ &\geq \theta \min\{f_1(x_1), f_2(x_1)\} + (1 - \theta) \max\{f_1(x_2), f_2(x_2)\} \\ &= \theta g(x_1) + (1 - \theta)g(x_2) \end{aligned}$$

And therefore the function $g(x)$ is concave.

The Lagrange dual function

The Lagrange dual function of an optimization problem is denoted as $g : \mathbb{R}^m \times \mathbb{R}^p \rightarrow \mathbb{R}$ and defined as the minimum value of the Lagrangian over x , that is,

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

Considering that $f_0(x), f_i(x), h_i(x)$ are neither convex nor concave. Is $g(\lambda, \nu)$ convex? And concave? Justify your answer.

For each x I have an affine function on λ and ν , which is both convex and concave. Since the function $g(\lambda, \nu)$ is the pointwise minimum of a set of concave functions, then this function is concave.

 Lagrange dual function of Example 1-13

Provide the Lagrange dual function corresponding to the optimization problem of Example 1-13 ($a = 0$)

$$\begin{aligned} g(\underline{\eta}, \bar{\eta}) &= \min_x \mathcal{L}(x, \underline{\eta}, \bar{\eta}) \\ &= \min_x bx + c + \underline{\eta}(\underline{x} - x) + \bar{\eta}(x - \bar{x}) - px \\ &= \begin{cases} c + \underline{\eta}\underline{x} - \bar{\eta}\bar{x} & \text{if } b - p - \underline{\eta} + \bar{\eta} = 0 \\ -\infty & \text{otherwise} \end{cases} \end{aligned}$$

 Lagrange dual function of Example 1-14

Provide the Lagrange dual function corresponding to the optimization problem of Example 1-14 ($a_g = 0$)

$$\begin{aligned} g(\underline{\eta}_g, \bar{\eta}_g, \lambda) &= \min_x \mathcal{L}(x, \underline{\eta}_g, \bar{\eta}_g, \lambda) \\ &= \min_x \sum_g \left(b_g x_g + c_g + \underline{\eta}_g(\underline{x}_g - x_g) + \bar{\eta}_g(x_g - \bar{x}_g) \right) + \lambda \left(\sum_g x_g - d \right) \\ &= \begin{cases} \sum_g \left(c_g + \underline{\eta}_g \underline{x}_g - \bar{\eta}_g \bar{x}_g \right) - \lambda d & \text{if } b_g - \underline{\eta}_g + \bar{\eta}_g + \lambda = 0, \forall g \\ -\infty & \text{otherwise} \end{cases} \end{aligned}$$

Lower bounds on optimal value using the dual function

Let $\tilde{x}$ be a feasible point that satisfies

$$\begin{aligned}f_i(\tilde{x}) &\leq 0, \forall i = 1, \dots, m \\h_i(\tilde{x}) &= 0, \forall i = 1, \dots, p\end{aligned}$$

For any $\lambda \succeq 0$ and ν we have that

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

Prove that $g(\lambda, \nu) \leq f_0(\tilde{x})$. Is that inequality also valid for $\tilde{x} = x^*$? Discuss the implications.

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

In particular,

$$g(\lambda, \nu) \leq f_0(x^*)$$

This means that for any $\lambda \geq 0, \nu$, $g(\lambda, \nu)$ is a valid lower bound on p^* .

Lower bound on optimal value of Example 1-13

Solve Scheduling.gms for $a = 0$ and $p = 30$. Chose any “interesting” value of the Lagrange multipliers of Example 1-13. Confirm that the value of the dual function yields a lower bound on the optimal value of the optimization problem.

$$p = 30$$

$$b = 10$$

$$c = 100$$

$$\underline{x} = 0$$

$$\bar{x} = 100$$

$$p^* = -1900$$

$$\underline{\eta} = 0$$

$$\bar{\eta} = 20$$

$$b - p - \underline{\eta} + \bar{\eta} = 0$$

$$g(\underline{\eta}, \bar{\eta}) = -1900 \leq -1900 = p^*$$

Lower bound on optimal value of Example 1-14

Solve EcoDispatch.gms for $a_g = 0$ and $d = 300$. Chose any “interesting” value of the Lagrange multipliers of Example 1-14. Confirm that the value of the dual function yields a lower bound on the optimal value of the optimization problem.

$$d = 300$$

$$p^* = 2050$$

$$\lambda = 10$$

$$\underline{\eta}_g = (15, 20, 26)$$

$$\bar{\eta}_g = (0, 0, 0)$$

$$g(\underline{\eta}_g, \bar{\eta}_g, \lambda) = -2700 \leq 2050 = p^*$$

The Lagrange dual problem

For any $\lambda \succeq 0$ and ν we have that $g(\lambda, \nu)$ is a lower bound on the optimal value. Therefore we can be interested in knowing the highest lower bound we can find, that is, to solve the following optimization problem.

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

This problem is called the Lagrange dual problem.

Example 2-7: Is the Lagrange dual problem convex?

Indicate under which conditions of f_0, f_i, h_i the Lagrange dual problem is convex.

The Lagrange dual problem is always convex since $g(\lambda, \nu)$ is concave for any f_0, f_i, h_i .

Weak duality

Let p^* and d^* denote the optimal values of the original optimization problem and its Lagrange dual problem respectively. Then it is always satisfied that

$$d^* \leq p^*$$

This inequality is known as weak duality and it is true even if the original optimization problem is not convex.

Strong duality

We say that strong duality holds if

$$d^* = p^*$$

Strong duality holds if the optimization problem is convex and the problem satisfies what are called constraint qualifications. Among them one of the easiest is the Slater's condition imposing that there must exist a strictly feasible point, i.e., a point $\tilde{x}$ that satisfies

$$\begin{aligned}f_i(\tilde{x}) &< 0, \forall i = 1, \dots, m \\ h_i(\tilde{x}) &= 0, \forall i = 1, \dots, p\end{aligned}$$

📌 Example 2-8: Solve the dual problem of Example 1-13 (SchedulingDual.gms)

Formulate the Lagrange dual problem of Example 1-13 and solve it for $a = 0$ and $p = 30$ using GAMS. Write the optimal values of the dual variables as well as the optimal value d^* .

$$\begin{aligned} & \underset{\underline{\eta}, \bar{\eta}}{\text{maximize}} && c + \underline{\eta}x - \bar{\eta}\bar{x} \\ & \text{subject to} && b - p - \underline{\eta} + \bar{\eta} = 0 \\ & && \underline{\eta}, \bar{\eta} \geq 0 \end{aligned}$$

The solution to this problem is

$$\begin{aligned} \underline{\eta}^* &= 0 \\ \bar{\eta}^* &= 20 \\ d^* &= -1900 \end{aligned}$$

📌 Example 2-9: Solve the dual problem of Example 1-14 (EcoDispatchDual.gms)

Formulate the Lagrange dual problem of Example 1-14 and solve it for $a_g = 0$ and $d = 300$ using GAMS. Write the optimal values of the dual variables as well as the optimal value d^* .

$$\begin{aligned} & \underset{\underline{\eta}, \bar{\eta}}{\text{maximize}} && \sum_g \left(c_g + \underline{\eta}_g x_g - \bar{\eta}_g \bar{x}_g \right) - \lambda d \\ & \text{subject to} && b_g - \underline{\eta}_g + \bar{\eta}_g + \lambda = 0, \quad \forall g \\ & && \underline{\eta}_g, \bar{\eta}_g \geq 0, \quad \forall g \end{aligned}$$

The solution to this problem is

$$\begin{aligned} \underline{\eta}_g^* &= (0, 0, 6) \\ \bar{\eta}_g^* &= (5, 0, 0) \\ \lambda^* &= -10 \\ d^* &= 2050 \end{aligned}$$

📌 Example 2-10: Duality as a certificate of optimality of Example 1-13

Solve Scheduling.gms with $a = 0$ and $p = 40$ and write down the optimal values of the dual variables. Check that the dual variables are feasible points of the corresponding dual problem. Check whether the strong duality condition holds. What does this mean?

The solution provided by GAMS is

$$p^* = -2900$$

$$x^* = 100$$

$$\underline{\eta}^* = 0$$

$$\overline{\eta}^* = 30$$

First we check that the dual solution is feasible, i.e.,

$$\underline{\eta}^*, \overline{\eta}^* \geq 0$$

$$b - p - \underline{\eta}^* + \overline{\eta}^* = 0$$

Then we compute the value of d^*

$$d^* = c + \underline{\eta}^* \underline{x} - \overline{\eta}^* \overline{x} = -2900$$

Since $p^* = d^*$ the solution provided by GAMS is correct and we should not care about how this solution has been obtained.

📌 Example 2-11: Duality as a certificate of optimality of Example 1-14

Solve EcoDispatch.gms with $a_g = 0$ and $d = 400$ and write down the optimal values of the dual variables. Check that the dual variables are feasible points of the corresponding dual problem. Check whether the strong duality condition holds. What does this mean?

The solution provided by GAMS is

$$p^* = 3050$$

$$x^* = (250, 150, 0)$$

$$\underline{\eta}_g^* = (0, 0, 0)$$

$$\overline{\eta}_g^* = (11, 6, 0)$$

$$\lambda^* = -16$$

First we check that the dual solution is feasible, i.e.,

$$\underline{\eta}_g^*, \overline{\eta}_g^* \geq 0$$

$$b_g + \lambda - \underline{\eta}_g^* + \overline{\eta}_g^* = 0, \forall g$$

Then we compute the value of d^*

$$d^* = \sum_g \left(c_g + \underline{\eta}_g^* \underline{x}_g - \overline{\eta}_g^* \overline{x}_g \right) - \lambda^* d = 3050$$

Since $p^* = d^*$ the solution provided by GAMS is correct.

The perturbed problem

Let us consider the following perturbed problem

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

with f_0, f_i convex functions and h_i affine functions. Let us denote its optimal value as $p^*(u, v)$. For $u, v = 0$, the Lagrange dual function of this problem is:

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

We denote the optimal Lagrange multipliers of the unperturbed problem as λ^*, ν^* . Since the unperturbed problem is convex, any $\lambda \geq 0, \nu$ provide a lower bound on the optimal value, that is,

$$\begin{aligned} \min_x \left(f_0(x) + \sum_{i=1}^m \lambda_i (f_i(x) - u_i) + \sum_{i=1}^p \nu_i (h_i(x) - v_i) \right) &\leq p^*(u, v) \\ \min_x \left(f_0(x) + \sum_{i=1}^m \lambda_i f_i(x) + \sum_{i=1}^p \nu_i h_i(x) \right) - \lambda^T u - \nu^T v &\leq p^*(u, v) \\ g(\lambda, \nu) - \lambda^T u - \nu^T v &\leq p^*(u, v) \end{aligned}$$

In particular, this is also true for λ^*, ν^* and therefore

$$g(\lambda^*, \nu^*) - \lambda^{*T} u - \nu^{*T} v \leq p^*(u, v)$$

Since strong duality holds for the unperturbed problem, we finally obtain:

$$p^*(0, 0) - \lambda^{*T} u - \nu^{*T} v \leq p^*(u, v)$$

Example 2-12: Fill in the following table

λ_i^*	u_i	ν_i^*	v_i	then $p^*(u, v)$ will
> 0	< 0	-	-	increase
> 0	> 0	-	-	-
-	-	> 0	> 0	-
-	-	> 0	< 0	increase
-	-	< 0	> 0	increase
-	-	< 0	< 0	-

Local sensitivity analysis and shadow price

Lagrange multipliers are the local sensitivities of p^* with respect to constraints perturbations, i.e.,

$$\lambda_i^* = -\frac{\partial p^*(0, 0)}{\partial u_i} \qquad \nu_i^* = -\frac{\partial p^*(0, 0)}{\partial v_i}$$

📌 Example 2-13: Sensitivity analysis

Using sensitivity theory, determine a lower bound of the objective functions of Example 1-15 if the maximum values of production levels and power flows ($\bar{x}_g$ and $\bar{f}_{nm}$) are increased by 1 MW. Contrast your results by solving the perturbed version of EcoDispatch2.gms with GAMS.

The solution provided by GAMS is

$$p^*(0, 0) = 2700$$

$$\underline{\eta}_g^* = (0, 0, 0)$$

$$\bar{\eta}_g^* = (0, 0, 0)$$

$$\lambda^* = (8, 12, 19)$$

$$\underline{\theta}^* = (0, 0, 0)$$

$$\bar{\theta}^* = (4, 11, 7)$$

Using the inequality of the perturbed problem we have

$$p^*(u, v) \geq 2700 - 4 - 11 - 7 = 2678$$

Now we solve the perturbed problem with GAMS and we obtain

$$p^*(u, v) = 2678.16 \geq 2678$$

📌 Example 2-14: Shadow prices

According to the results of EcoDispatch2.gms provided by GAMS, in which technical limit would you prefer to increase to reduce the operation cost of the system?

We should try to increase the capacity of line l_2 connecting $n_1 - n_3$ because the associated shadow price is the highest.

Complementary slackness

Let x^* and λ^*, ν^* be the optimal primal and dual solution of an optimization problem in which strong duality holds, i.e.,

$$f_0(x^*) \stackrel{(1)}{=} g(\lambda^*, \nu^*) \stackrel{(2)}{=} \min_x \left(f_0(x) + \sum_{i=1}^m \lambda_i^* f_i(x) + \sum_{i=1}^p \nu_i^* h_i(x) \right) \stackrel{(3)}{\leq} f_0(x^*) + \sum_{i=1}^m \lambda_i^* f_i(x^*) + \sum_{i=1}^p \nu_i^* h_i(x^*) \stackrel{(4)}{\leq} f_0(x^*)$$

Explain the reasoning for each equality or inequality:

- (1) Strong duality holds
- (2) Definition of Lagrange dual function
- (3) Definition of minimum operator
- (4) Feasible primal and dual points, which means $\lambda_i^* \geq 0$, $f_i(x^*) \leq 0$ and $h_i(x^*) = 0$.

Explain why this set of equalities and inequalities implies that

$$\lambda_i^* f_i(x^*) = 0, \forall i = 1, \dots, m$$

known as the complementarity slackness condition.

All terms of the inequalities have to be equal to $f_0(x^*)$ and therefore $\sum_{i=1}^m \lambda_i^* f_i(x^*) = 0$. Since this is a sum of negative number, this is only true if each term is equal to 0, i.e., $\lambda_i^* f_i(x^*) = 0, \forall i = 1, \dots, m$.

KKT conditions for a general optimization problem

Let f_0, f_i, h_i be differentiable functions. Let x^* and λ^*, ν^* be the optimal primal and dual solution of that general optimization problem with zero duality gap. The solution x^* must be feasible and therefore

$$\begin{aligned} f_i(x^*) &\leq 0, & \forall i = 1, \dots, m \\ h_i(x^*) &= 0, & \forall i = 1, \dots, p \end{aligned}$$

Likewise λ^*, ν^* are dual feasible which means

$$\lambda_i^* \geq 0, \quad \forall i = 1, \dots, m$$

Since x^* minimizes $\mathcal{L}(x, \lambda^*, \nu^*)$ we can write

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

Finally, complementarity slackness must be satisfied as well

$$\lambda_i^* f_i(x^*) = 0, \quad \forall i = 1, \dots, m$$

All these conditions together are known as the Karush-Kuhn,Tucker (KKT) conditions. If we have an optimization problem with differentiable functions and the optimal primal and dual optimal points satisfy strong duality, then these points must also satisfy the KKT condition above.

Example 2-15: KKT conditions

Write down the optimality conditions of the following optimization problem

$$\begin{aligned} \max_{x,y} \quad & x^2 + 5y^2 \\ \text{s.t.} \quad & x^2 - 7y^2 \geq 8, \quad -7x + 3y \leq 5, \quad x - y^2 = 0 \end{aligned}$$

First it is recommended to rewrite the optimization problem in standard form as follows

$$\begin{aligned} \min_{x,y} \quad & -x^2 - 5y^2 \\ \text{s.t.} \quad & -x^2 + 7y^2 \leq -8, \quad -7x + 3y \leq 5, \quad x - y^2 = 0 \end{aligned}$$

The corresponding KKT conditions are

$$\begin{aligned} -x^2 + 7y^2 &\leq -8 \\ -7x + 3y &\leq 5 \\ x - y^2 &= 0 \\ \lambda_1, \lambda_2 &\geq 0 \\ -2x - 2x\lambda_1 - 7\lambda_2 + \nu &= 0 \\ -10y + 14y\lambda_1 + 3\lambda_2 - 2y\nu &= 0 \\ (-x^2 + 7y^2 + 8) \lambda_1 &= 0 \\ (-7x + 3y - 5) \lambda_2 &= 0 \end{aligned}$$

KKT conditions for a convex optimization problem

Let assume we have a convex optimization problem. Let $\tilde{x}, \tilde{\lambda}, \tilde{\nu}$ be a point that satisfies the KKT conditions, i.e.,

$$\begin{aligned} f_i(\tilde{x}) &\leq 0, & \forall i = 1, \dots, m \\ h_i(\tilde{x}) &= 0, & \forall i = 1, \dots, p \\ \tilde{\lambda}_i &\geq 0, & \forall i = 1, \dots, m \\ \nabla f_0(\tilde{x}) + \sum_{i=1}^m \tilde{\lambda}_i \nabla f_i(\tilde{x}) + \sum_{i=1}^p \tilde{\nu}_i \nabla h_i(\tilde{x}) &= 0 \\ \tilde{\lambda}_i f_i(\tilde{x}) &= 0, & \forall i = 1, \dots, m \end{aligned}$$

This implies that the Lagrangian function

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

is convex and therefore

$$g(\tilde{\lambda}, \tilde{\nu}) \stackrel{(1)}{=} \min_x \left(\mathcal{L}(x, \tilde{\lambda}, \tilde{\nu}) \right) \stackrel{(2)}{=} \mathcal{L}(\tilde{x}, \tilde{\lambda}, \tilde{\nu}) \stackrel{(3)}{=} f_0(\tilde{x}) + \sum_{i=1}^m \tilde{\lambda}_i f_i(\tilde{x}) + \sum_{i=1}^p \tilde{\nu}_i h_i(\tilde{x}) \stackrel{(4)}{=} f_0(\tilde{x})$$

Explain the reasoning for each equality:

- (1) Definition of Lagrange dual function
- (2) The Lagrange function is convex and its gradient vanishes at $\tilde{x}$
- (3) Definition of Lagrangian
- (4) Because $h_i(\tilde{x}) = 0$ and $\tilde{\lambda}_i f_i(\tilde{x}) = 0$

Since strong duality holds, we can thus confirm that $\tilde{x}$ and $\tilde{\lambda}, \tilde{\nu}$ are primal and dual optimal, respectively. This means that if we have a convex optimization problem, the KKT conditions are sufficient for optimality, i.e., any point satisfying the KKT is guaranteed to be an optimal point. For KKT conditions to be necessary and sufficient, the convex problem must also satisfy some constraint qualification as, for example, the Slater condition.

📌 Example 2-16: KKT conditions of the scheduling problem

Determine the KKT conditions corresponding to Example 1-13. Check that the primal and dual solutions provided by GAMS satisfy these conditions. Use $p = 30$ and the original data corresponding to the generating unit.

The optimization problem is

$$\begin{aligned} & \underset{x}{\text{minimize}} && ax^2 + bx + c - px \\ & \text{subject to} && \underline{x} \leq x \leq \bar{x} : \underline{\eta}, \bar{\eta} \end{aligned}$$

The corresponding KKT conditions are

$$\begin{aligned} & \underline{x} \leq x \leq \bar{x} \\ & \underline{\eta}, \bar{\eta} \geq 0 \\ & 2ax + b - p - \underline{\eta} + \bar{\eta} = 0 \\ & (\underline{x} - x)\underline{\eta} = 0 \\ & (x - \bar{x})\bar{\eta} = 0 \end{aligned}$$

The solution provided by GAMS is

$$\begin{aligned} x^* &= 50 \\ \underline{\eta}^* &= 0 \\ \bar{\eta}^* &= 0 \end{aligned}$$

It is easy to see that the primal and dual solution satisfies the KKT conditions.

📌 Example 2-17: KKT conditions of the economic dispatch

Determine the KKT conditions corresponding to Example 1-14. Check that the primal and dual solutions provided by GAMS satisfy these conditions. Use $d = 300$ and the original data corresponding to the generating units.

The optimization problem is

$$\begin{aligned} & \underset{x_g}{\text{minimize}} && \sum_g ax_g^2 + bx_g + c \\ & \text{subject to} && \underline{x}_g \leq x_g \leq \bar{x}_g, \quad \forall g \\ & && \sum_g x_g = d \end{aligned}$$

The corresponding KKT conditions are

$$\begin{aligned} & \underline{x}_g \leq x_g \leq \bar{x}_g, \quad \forall g \\ & \sum_g x_g = d \\ & \underline{\eta}_g, \bar{\eta}_g \geq 0, \quad \forall g \\ & 2a_g x_g + b_g + \lambda - \underline{\eta}_g + \bar{\eta}_g = 0, \quad \forall g \\ & (\underline{x}_g - x_g) \underline{\eta}_g = 0 \\ & (x_g - \bar{x}_g) \bar{\eta}_g = 0 \end{aligned}$$

The solution provided by GAMS is

$$\begin{aligned} x_g^* &= (250, 50, 0) \\ \underline{\eta}_g^* &= (0, 0, 4) \\ \bar{\eta}_g^* &= (2, 0, 0) \\ \lambda^* &= -12 \end{aligned}$$

It is easy to see that the primal and dual solution satisfies the KKT conditions.

📌 Example 2-18: Primal-dual formulation

Let assume we have a convex optimization problem. Let $\tilde{x}, \tilde{\lambda}, \tilde{\nu}$ be a point that satisfies

$$\begin{aligned} f_i(\tilde{x}) &\leq 0, & \forall i = 1, \dots, m \\ h_i(\tilde{x}) &= 0, & \forall i = 1, \dots, p \\ \tilde{\lambda}_i &\geq 0, & \forall i = 1, \dots, m \\ \nabla f_0(\tilde{x}) + \sum_{i=1}^m \tilde{\lambda}_i \nabla f_i(\tilde{x}) + \sum_{i=1}^p \tilde{\nu}_i \nabla h_i(\tilde{x}) &= 0 \\ f_0(\tilde{x}) &= g(\tilde{\lambda}, \tilde{\nu}) \end{aligned}$$

Prove that imposing $f_0(\tilde{x}) = g(\tilde{\lambda}, \tilde{\nu})$ is equivalent to impose $\tilde{\lambda}_i f_i(\tilde{x}) = 0, \forall i = 1, \dots, m$.

$$\begin{aligned} g(\tilde{\lambda}, \tilde{\nu}) &= \min_x \mathcal{L}(x, \tilde{\lambda}, \tilde{\nu}) \\ &= \mathcal{L}(\tilde{x}, \tilde{\lambda}, \tilde{\nu}) \\ &= f_0(\tilde{x}) + \sum_{i=1}^m \tilde{\lambda}_i f_i(\tilde{x}) + \sum_{i=1}^p \tilde{\nu}_i h_i(\tilde{x}) \\ &= f_0(\tilde{x}) \end{aligned}$$

This implies that $\tilde{\lambda}_i f_i(\tilde{x}) = 0$.

📌 Example 2-19: Primal-dual formulation of the scheduling problem

Determine the primal-dual formulation of the optimization problem in Example 1-13 for $a = 0$. Compare it with its corresponding KKT conditions for $a = 0$.

$$\begin{aligned}\underline{x} &\leq x \leq \bar{x} \\ \underline{\eta}, \bar{\eta} &\geq 0 \\ 2ax + b - p - \underline{\eta} + \bar{\eta} &= 0 \\ bx + c - px &= c + \underline{\eta}\underline{x} - \bar{\eta}\bar{x}\end{aligned}$$

This is just a linear system of equations and inequations. The original KKT conditions include some non-linear equalities.

📌 Example 2-20: Primal-dual formulation of the economic dispatch problem

Determine the primal-dual formulation of the optimization problem in Example 1-14 for $a_g = 0$. Compare it with its corresponding KKT conditions for $a_g = 0$.

$$\begin{aligned}\underline{x}_g &\leq x_g \leq \bar{x}_g, \quad \forall g \\ \sum_g x_g &= d \\ \underline{\eta}_g, \bar{\eta}_g &\geq 0, \quad \forall g \\ 2a_g x_g + b_g + \lambda - \underline{\eta}_g + \bar{\eta}_g &= 0, \quad \forall g \\ \sum_g b_g x_g &= \sum_g \left(\underline{\eta}_g \underline{x}_g - \bar{\eta}_g \bar{x}_g \right) - \lambda d\end{aligned}$$

Again, this only includes linear equations and inequations.

Constraint qualifications (QC)

A constraint qualification (also known as regularity condition) is a set of conditions on the constraints of an optimization problem to guarantee that the KKT conditions are necessary for optimality. Those points for which a constraint qualification is satisfied are also known as regular points. Let $\tilde{x}$ be a feasible regular point that minimizes the objective function $f_0(x)$, then we can ensure that the point $\tilde{x}$ satisfies the KKT conditions. The most common constraint qualifications are:

- Non-convex problems:
 - Linearity constraint qualification (LCQ): all constraints are affine functions
 - Linear independence constraint qualification (LICQ): the gradients of the active inequality constraints and the gradients of the equality constraints are linearly independent at $\tilde{x}$.
 - Mangasarian-Fromovitz constraint qualifications (MFCQ): the gradients of the equality constraints are linearly independent at $\tilde{x}$ and there exists a vector $d \in \mathbb{R}^n$ such that $\nabla h_i(\tilde{x})^T d = 0, i = 1, \dots, p$ and $\nabla f_i(\tilde{x})^T d < 0, \forall i \in \mathcal{A}(\tilde{x})$.
- Convex problem:
 - Slater condition: a convex problem satisfies the Slater condition if there exists a strictly feasible point $\tilde{x}$ such that $f_i(\tilde{x}) < 0, \forall i$

Summary of KKT conditions and their implications

Cvx?	CQ?	Implications
NO	NO	KKT are neither necessary nor sufficient optimality conditions. The first implication means that the optimal point is not necessarily satisfying the KKT conditions. The second implication means that it is not sufficient that a point satisfies the KKT conditions to be optimal.
NO	YES	KKT are necessary but not sufficient for optimality. This means that the optimal point is necessarily satisfying the KKT conditions. However, it is not sufficient that a point satisfies the KKT conditions to be optimal.
YES	NO	KKT are sufficient for optimality but not necessary. This means that it is sufficient for a point to satisfy the KKT conditions to be optimal. However, the optimal points is not necessarily satisfying the KKT conditions.
YES	YES	KKT are necessary and sufficient optimality conditions. This means that the optimal point is necessarily satisfying the KKT conditions, and also that it is sufficient that a point satisfies the KKT conditions to be optimal.

Example 2-21: Constraint qualification and necessary optimality conditions

Let us have the following optimization problem:

$$\begin{aligned}
 \min_{x, y \geq 0} \quad & -x \\
 \text{s.t.} \quad & y - (1 - x)^3 \leq 0
 \end{aligned}$$

Is the problem convex? Plot the feasible region in $\mathbb{R}^2$ and determine the optimal point. Is the LICQ satisfied at the optimal point? Write down the KKT conditions and check whether the optimal point satisfies them.