Appendix B
Basics of Optimization

In this appendix, we review some basics of optimization. In Sect. B.1, we introduce the mathematical formulation for both general and linear optimization problems. Duality theory in linear programming is briefly presented in Sect. B.2. The Karush–Kuhn–Tucker (KKT) optimality conditions are presented in Sect. B.3. Finally, Mathematical Programs with Equilibrium Constraints (MPECs) are introduced in Sect. B.4.

B.1 Formulation of an Optimization Problem

The general mathematical formulation of an optimization problem is:

$$\underset{x}{\text{Min.}}\, f(x) \tag{B.1a}$$

$$\text{s.t. } h(x) = 0, \tag{B.1b}$$

$$g(x) \leq 0. \tag{B.1c}$$

Problem (B.1) includes the following elements:

- $x \in \mathbb{R}^n$ is a vector including the n decision variables.
- $f(\,\cdot\,) : \mathbb{R}^n \rightarrow \mathbb{R}$ is the objective function of the optimization problem. It maps values of the decision vector x to a real value representing the desirability of this solution to the decision-maker. Typically the objective function represents a cost in minimization problems or a benefit in maximization ones.
- $h(\,\cdot\,) : \mathbb{R}^n \rightarrow \mathbb{R}^m$ and $g(\,\cdot\,) : \mathbb{R}^n \rightarrow \mathbb{R}^l$ are vector-valued functions of the decision vector x. They define m equality and l inequality constraints through (B.1b) and (B.1c), respectively. Note that we assume that the zero-valued vectors on the right-hand side of (B.1b) and (B.1c) are properly sized to match the dimension of the vectors on the left-hand side.

The joint enforcement of equalities (B.1b) and inequalities (B.1c) defines the feasibility region of the optimization problem. A decision x is called *feasible* if it satisfies (B.1b)–(B.1c).

J. M. Morales et al., *Integrating Renewables in Electricity Markets,*
International Series in Operations Research & Management Science 205,
DOI 10.1007/978-1-4614-9411-9, © Springer Science+Business Media New York 2014

The aim of problem (B.1) is to determine, among the set of feasible decisions, the one that yields the lowest value of the objective function (B.1a).

The simplest instance of an optimization problem is a linear programming problem (LP). This is obtained when the functions $f(\cdot)$, $h(\cdot)$, and $g(\cdot)$ in (B.1) are linear. We can formulate a linear programming problem as:

$$\underset{x}{\text{Min.}}\ c^\top x \tag{B.2a}$$

$$\text{s.t. } A_\text{E}x = b_\text{E}, \tag{B.2b}$$

$$A_\text{I}x \geq b_\text{I}. \tag{B.2c}$$

Note that the general functions $f(\cdot)$, $h(\cdot)$, and $g(\cdot)$ are replaced by affine expressions involving the following vectors and matrices:

- $c \in \mathbb{R}^n$ is the cost coefficient of the decision vector x.
- $A_\text{E} \in \mathbb{R}^{m\times n}$, and $b_\text{E} \in \mathbb{R}^m$ define the m equality constraints (B.2b).
- $A_I \in \mathbb{R}^{l\times n}$ and $b_I \in \mathbb{R}^l$ define the l linear inequality constraints (B.2c). Note that the sign of the constraints is changed with respect to (B.1c). This is to simplify the representation of the dual problem in the next section.

LPs model a wide variety of real-world problems, also within the area of electricity markets. Very large LPs can be solved using commercially available software.

B.2 Duality in Linear Programming

Let us associate the vector $\lambda \in \mathbb{R}^m$ to the equalities (B.2b) and the vector $\mu \in \mathbb{R}^l$ to the inequalities (B.2c). The following linear maximization problem is the *dual* of LP (B.2), which is referred to as the *primal* problem:

$$\underset{\lambda,\mu}{\text{Max.}}\ b_\text{E}^\top \lambda + b_\text{I}^\top \mu \tag{B.3a}$$

$$\text{s.t. } A_\text{E}^\top \lambda + A_\text{I}^\top \mu = c, \tag{B.3b}$$

$$\mu \geq 0. \tag{B.3c}$$

The dual problem (B.3) can be considered a transposed version of the primal problem (B.2). Indeed, the following relationships hold:

- While the primal problem (B.2) has n decision variables and $m + l$ constraints, the dual problem has $m + l$ decision variables (λ and μ) and n constraints.
- The constraints (B.3b) of the dual problem involve the transposed of the matrices A_E and A_I defining the constraints (B.2b)–(B.2c) of the primal problem.
- The constant vectors b_E and b_I on the right-hand side of the primal constraints (B.2b)–(B.2c) form the cost coefficients of the dual linear objective function (B.3a). Viceversa, the cost coefficient vectors c of the primal objective function (B.2a) appear on the right-hand side of the dual constraints (B.3b).

Table B.1 Relationships between direction of the optimization problem, sign of the constraints, and bounds on the optimization variables in the primal and dual problems

Problem	Primal		Dual
Objective	Minimization	Objective	Maximization
	≥ 0		≥ 0
Constraint type	$= 0$	Variable bound	free
	≤ 0		≤ 0
	≥ 0		≤ 0
Variable bound	free	Constraint type	$= 0$
	≤ 0		≥ 0

The direction of optimization (minimization or maximization), the sign of the constraints ($\geq$, $=$, or $\leq$) and the bounds on the variables ($\geq$ 0, free, or $\leq$ 0) for the primal and the dual problem are linked. Specifically, the direction of the dual optimization problem is opposite to the one of the primal one. Furthermore, the signs of the primal constraints set the bounds on the associated dual variables and, conversely, the bounds on the primal variables set the signs of the dual constraints. Table B.1 includes all the possible combinations of constraint types and variable bounds for a primal minimization problem, and the corresponding ones for the associated dual maximization problem. Finally, note that the dual of the dual problem is the primal problem, see [2]. This implies that the headers in Table B.1 can be swapped, so that the right column pertains to a primal maximization problem and the left one to its dual minimization one.

The objective function values of the primal and dual problems are related to each other through the so-called *weak* and *strong* duality theorems. Such theorems are of particular importance. In what follows, we shall present them without proof. The interested reader is referred to [5] for further details and proofs.

Theorem B.1 (Weak Duality) *If* x *is feasible for* (B.2), *and* λ, μ *are feasible for* (B.3), *then* $c^\top x \geq b_{\mathrm{E}}^\top \lambda + b_{\mathrm{I}}^\top \mu$.

Theorem B.2 (Strong Duality) *If the primal problem has a finite optimal solution* x^**, so does the dual problem and at optimality it holds that* $c^\top x^* = b_{\mathrm{E}}^\top \lambda^* + b_{\mathrm{I}}^\top \mu^*$.

Since the dual of the dual problem is again the primal problem, the converse of the previous theorems holds trivially.

Note that the dual variables λ and μ have an important economic interpretation, as they are marginal costs. Indeed, they represent the per-unit change (increase) in the optimal value of the objective function (B.2a) if the right-hand side of the associated constraint is increased marginally. Naturally, $\mu \geq 0$. Indeed, a marginal increase of any element of the vector b_{I} would result in a smaller feasible space for (B.2), and hence in a larger, i.e., worse, optimal value of the objective function.

B.3 Karush–Kuhn–Tucker Conditions

In this appendix, we only deal with KKT optimality conditions for convex problems, and refer to [1] for a general introduction to duality theory in nonlinear programming.

Let us consider the general formulation (B.1), and suppose that $f(\cdot)$, $g(\cdot)$ are continuously differentiable and convex, and $h(\cdot)$ is affine. Furthermore, we assume that a constraint qualification holds. For example, we may require that $g(\cdot)$ be affine (linearity constraint qualification). Another common constraint qualification requires linear independence of the gradients of active inequality constraints and of equality constraints. We refer the reader to specialized books on optimization, for instance [1], for a detailed treatment of constraint qualifications.

We can define the Lagrangian function for problem (B.1) as follows:

$$\mathcal{L}(x, \lambda, \mu) = f(x) + \lambda^\top h(x) + \mu^\top g(x). \tag{B.4}$$

Under the assumptions above, the following KKT conditions are necessary and sufficient for optimality for problem (B.1):

$$\nabla_x f(x) + \lambda^\top \nabla_x h(x) + \mu^\top \nabla_x g(x) = 0, \tag{B.5a}$$

$$h(x) = 0, \tag{B.5b}$$

$$g(x) \leq 0, \tag{B.5c}$$

$$\mu \geq 0, \tag{B.5d}$$

$$\mu^\top g(x) = 0. \tag{B.5e}$$

Equations (B.5a) are stationarity conditions. Constraints (B.5b) and (B.5c) enforce feasibility of the primal problem, while (B.5d) is a feasibility condition of the dual problem. Finally, (B.5e) enforces complementary slackness. Note that in view of (B.5c) and (B.5d), the scalar product on the left-hand side of (B.5e) is actually the sum of non-positive terms only. As a result, (B.5e) implies that the element-by-element product between μ_i and $g_i(x)$ is equal to 0.

Note that constraint qualifications are needed for ensuring that KKT conditions are necessary for optimality, while convexity is needed to ensure their sufficiency.

The dual vectors λ and μ retain the interpretation of marginal costs discussed in Sect. B.2.

Finally, the notation for constraints (B.5c)–(B.5e) can be compacted into the following nonlinear constraint:

$$0 \geq g(x) \perp \mu \geq 0, \tag{B.6}$$

where the $\perp$ (*perpendicular*) operator enforces the perpendicularity condition between the vectors on the left- and right-hand sides, i.e., that their element-by-element product is equal to zero.

B.4 Mathematical Programs with Equilibrium Constraints

MPECs is a relatively recent area of optimization, which has been applied to study electricity markets with increasing success in the recent years. In this section, we briefly introduce the concept of MPEC and present how it can be used to model bilevel programs, i.e., optimization problems constrained by other optimization problems. The reader is referred to [4] and [6] for an in-depth treatment of the subject.

The general formulation of a bilevel optimization problem is the following:

$$\underset{x,y}{\text{Min.}}\ f^{\text{U}}(x,y) \tag{B.7a}$$

$$\text{s.t. } g^{\text{U}}(x,y) \leq 0, \tag{B.7b}$$

$$h^{\text{U}}(x,y) = 0, \tag{B.7c}$$

$$y \in \underset{z}{\arg\min} \left\{ f^{\text{L}}(x,z) \text{ s.t. } h^{\text{L}}(x,z) = 0, g^{\text{L}}(x,z) \leq 0 \right\}. \tag{B.7d}$$

The fundamental difference between the MPEC (B.7) and the general optimization problem (B.1) is the enforcement of conditions (B.7d). These conditions ensure that at any feasible point (x, y) of problem (B.7), the choice of variable y is optimal for the minimization problem within the braces in (B.7d).

Note that formulation (B.7) includes two optimization problems: an upper-level one that aims at the minimization of $f^{\text{U}}(\,\cdot\,)$, and a lower-level one consisting in the minimization of $f^{\text{L}}(\,\cdot\,)$.

The two problems are interdependent, since in general the upper-level objective function (B.7a) and constraints (B.7b)–(B.7c) depend on the lower-level decision variables y. Conversely, the objective function and the constraints of the lower-level problem (B.7d) depend on the upper-level variable vector x.

There is a hierarchical relationship between the two problems. Indeed, the lower-level problem is solved assuming that the upper-level decision x is fixed. On the contrary, the upper-level problem is solved accounting for the response of the lower-level problem to decision vector x.

Moreover, it should be emphasized that model (B.7) can accommodate several lower-level optimization problems, simply by concatenating multiple optimality conditions of the type of (B.7d).

Under the assumption that KKT conditions are necessary and sufficient for optimality in the lower-level problem, we can employ them to replace condition (B.7d). This results in the following formulation for the bilevel problem:

$$\underset{x,y,\lambda,\mu}{\text{Min.}}\ f^{\text{U}}(x,y) \tag{B.8a}$$

$$\text{s.t. } h^{\text{U}}(x,y) = 0, \tag{B.8b}$$

$$g^{\text{U}}(x,y) \leq 0, \tag{B.8c}$$

$$\nabla_y f^{\text{L}}(x,y) + \lambda^{\top} \nabla_y h^{\text{L}}(x,y) + \mu^{\top} \nabla_y g^{\text{L}}(x) = 0, \tag{B.8d}$$

$$h^{\text{L}}(x,y) = 0, \tag{B.8e}$$

$$g^{\text{L}}(x,y) \leq 0, \tag{B.8f}$$

$$\mu \geq 0, \tag{B.8g}$$

$$\mu^{\top} g^{\text{L}}(x,y) = 0, \tag{B.8h}$$

where λ and μ represent the dual variables associated to constraints $h^{\mathrm{L}}(x,z) = 0$ and $g^{\mathrm{L}}(x,z) \leq 0$, respectively, in the lower-level problem (B.7d).

The advantage of formulation (B.8) is the replacement of the nested lower-level problem with the set (B.8d)–(B.8h) of equations and inequalities, which results in a single-level optimization problem that fits the general formulation (B.1). However, note that solving the single-level program (B.8) is far from trivial. Indeed, KKT conditions are in general nonlinear and non convex, as they involve cross products between variables in the complementarity condition (B.8h).

A number of approaches for solving MPECs have been proposed in the literature. Among these, the method presented in [3] deserves to be mentioned because of its simplicity and its wide use in the literature on MPEC. This approach is based on the so-called *big M* reformulation of the complementarity conditions (B.8h) employing binary variables. In practice, we can replace the conditions:

$$\mu_i g_i^{\mathrm{L}}(x,y) = 0, \qquad \forall i, \tag{B.9}$$

with the following ones:

$$g_i^{\mathrm{L}}(x,y) \geq -z_i M_{1i}, \qquad \forall i, \tag{B.10a}$$

$$\mu_i \leq (1 - z_i) M_{2i}, \qquad \forall i, \tag{B.10b}$$

$$z_i \in \{0,1\}, \qquad \forall i. \tag{B.10c}$$

The use of binary variable z_i forces one of the right-hand sides of (B.10a) and (B.10b) to be equal to 0. In combination with (B.8f) and (B.8g), this implies that $g_i^{\mathrm{L}}(x,y)$ and/or μ_i must be equal to 0, as required by (B.8h).

For reformulation (B.10) to be valid within a bilevel problem, the constants M_{1i} and M_{2i} must be large enough so as not to leave the optimal solution out of the feasible space of (B.10). In practice, the choice of the *big M* constants is a rather challenging issue, as too large values for the constants result in computational inefficiencies in the solution of the resulting mixed-integer optimization problems.

As a final remark, it should be pointed out that if the feasible space of the lower-level problem(s) is defined by affine equality and inequality constraints, and if the partial derivatives of its objective function with respect to the lower-level decision variables are also affine, reformulation (B.10) of the complementarity conditions results in a mixed-integer linear program (MILP). Problems of this type can be efficiently tackled by specialized software.

We refer the interested reader to [4] and [6] for a more detailed presentation of the MPEC framework and of alternative solution techniques.

References

1. Bazaraa, M.S., Sherali, H.D., Shetty, C.M.: Nonlinear Programming: Theory and Algorithms, 3rd edn. Wiley, Hoboken, NJ (2013)

2. Castillo, E., Conejo, A.J., Pedregal, P., Garcia, R., Alguacil, N.: Building and Solving Mathematical Programming Models in Engineering and Science. Pure and Applied Mathematics. Wiley, New York (2001)
3. Fortuny-Amat, J., McCarl, B.: A representation and economic interpretation of a two-level programming problem. J. Oper. Res. Soc. **32**(9), 783–792 (1981)
4. Gabriel, S.A., Conejo, A.J., Fuller, J.D., Hobbs, B.F., Ruiz, C.: Complementarity Modeling in Energy Markets. International Series in Operations Research & Management Science. Springer, New York (2013)
5. Luenberger, D., Ye, Y.: Linear and Nonlinear Programming, 3rd edn. International Series in Operations Research & Management Science. Springer, New York (2008)
6. Luo, Z.Q., Pang, J.S., Ralph, D.: Mathematical Programs with Equilibrium Constraints. Cambridge University Press, Cambridge, UK (1996)