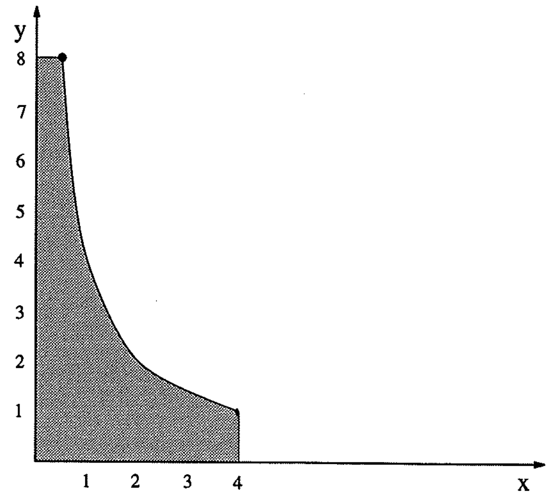

Figure 6.4: Feasible region of illustration 6.3.4

Remark 4 If we substitute x_2 from the equality constraint, then we obtain the equivalent formulation:

$$\begin{array}{ll} \min & -12x_1 - 14(1 - x_1^4) + 4(1 - x_1^4)^2 \\ \text{s.t.} & 0 \leq x_1 \leq 2 \end{array}$$

Eigenvalue analysis on this formulation shows that it is a convex problem. As a result, the projected $v(y)$ problem in our example is convex, and hence the **v3–GBD** converges to the global solution from any point.

Illustration 6.3.4 This example is taken from Floudas and Visweswaran (1990) and served as a motivating example for global optimization.

$$\begin{array}{rrcl} \min & & -x - y & \\ \text{s.t.} & xy & \leq & 4 \\ & 0 \leq & x & \leq 4 \\ & 0 \leq & y & \leq 8 \end{array}$$

The feasible region is depicted in Figure 6.4 and is nonconvex due to the bilinear inequality constraint. This problem exhibits a strong local minimum at $(x, y) = (4, 1)$ with objective equal to -5, and a global minimum at $(x, y) = (0.5, 8)$ with objective equal to -8.5.

The projection onto the y space, $v(y)$, is depicted in Figure 6.5.

Note that $v(y)$ is nonconvex and if we select as a starting point in **v3–GBD** $y^1 = 2$, then the algorithm will terminate with this as the solution, which is in fact not even a local solution. This is due to the common assumption of **v2–GBD** and **v3–GBD**.

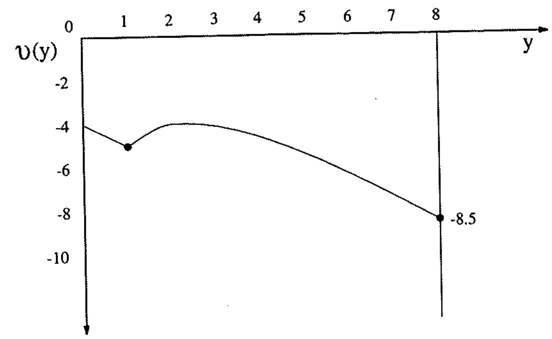

Figure 6.5: Projection onto the y–space for illustration 6.3.4

Remark 5 The global optimization approach (**GOP**) (Floudas and Visweswaran, 1990; Floudas and Visweswaran, 1993) overcomes this fundamental difficulty and guarantees ϵ–global optimality for several classes of nonconvex problems. The area of global optimization has received significant attention during the last decade, and the reader interested in global optimization theories, algorithms, applications, and test problems is directed to the books of Horst and Tuy (1990), Neumaier (1990), Floudas and Pardalos (1990), Floudas and Pardalos (1992), Hansen (1992), Horst and Pardalos (1995), and Floudas and Pardalos (1995).

6.3.6 GBD in Continuous and Discrete–Continuous Optimization

We mentioned in remark 1 of the formulation section (i.e., 6.3.1), that problem (6.2) represents a subclass of the problems for which the Generalized Benders Decomposition **GBD** can be applied. This is because in problem (6.2) we considered the $y \in Y$ set to consist of 0–1 variables, while Geoffrion (1972) proposed an analysis for Y being a continuous, discrete or continuous–discrete set.

The main objective in this section is to present the modifications needed to carry on the analysis presented in sections 6.3.1–6.3.5 for continuous Y and discrete-continuous Y set.

The analysis presented for the primal problem (see section 6.3.3.1) remains the same. The analysis though for the master problem changes only in the dual representation of the projection of problem (6.2) (i.e., $v(y)$) on the y-space. In fact, theorem 3 is satisfied if in addition to the two conditions mentioned in C3 we have that

(iii) For each fixed y, $v(y)$ is finite, $h(x,y)$, $g(x,y)$, and $f(x,y)$ are continuous on X, X is closed and the ε-optimal solution of the primal problem $P(y)$ is nonempty and bounded for some $\varepsilon \geq 0$.

Hence, theorem 3 has as assumptions: C1 and C3, which now has (i), (ii), and (iii). The algorithmic procedure discussed in section 6.3.4.2 remains the same, while the theorem for the finite convergence becomes finite ε- convergence and requires additional conditions, which are described in the following theorem:

Theorem 6.3.6 (Finite ε-convergence)
Let

(i) *Y be a nonempty subset of V,*

(ii) *X be a nonempty convex set,*

(iii) *f, g be convex on X for each fixed $y \in Y$,*

(iv) *h be linear on X for each fixed $y \in Y$,*

(v) *f, g, h be continuous on $X \times Y$,*

(vi) *The set of optimal multiplier vectors for the primal problem be nonempty for all $y \in Y$, and uniformly bounded in some neighborhood of each such point.*

Then, for any given $\epsilon > 0$ the **GBD** *terminates in a finite number of iterations.*

Remark 1 Assumption (i) (i.e., $Y \subseteq V$) eliminates the possibility of step 3b, and there are many applications in which $Y \subseteq V$ holds (e.g., variable factor programming). If, however, $Y \not\subseteq V$, then we may need to solve step 3b infinitely many successive times. In such a case, to preserve finite ϵ–convergence, we can modify the procedure so as to finitely truncate any excessively long sequence of successive executions of step 3b and return to step 3a with $\hat{y}$ equal to the extrapolated limit point which is assumed to belong to $Y \cap V$. If we do not make the assumption $Y \subseteq V$, then the key property to seek is that V has a representation in terms of a finite collection of constraints because if this is the case then step 3b can occur at most a finite number of times. Note that if in addition to C1, we have that X represents bounds on the x–variables or X is given by linear constraints, and h, g satisfy the separability condition, then V can be represented in terms of a finite collection of constraints.

Remark 2 Assumption (vi) requires that for all $y \in Y$ there exist optimal multiplier vectors and that these multiplier vectors do not go to infinity, that is they are uniformly bounded in some neighborhood of each such point. Geoffrion (1972) provided the following condition to check the uniform boundedness:

If X is nonempty, compact, convex set and there exists a point $\bar{x} \in X$ such that

$$h(\bar{x}, \bar{y}) = 0,$$

$$g(\bar{x}, \bar{y}) < 0,$$

then the set of optimal multiplier vectors is uniformly bounded in some open neighborhood of $\bar{y}$.

Illustration 6.3.5 This example is taken from Kocis and Grossmann (1988) and can be stated as

$$\begin{aligned} \min_{x,y} \quad & 2x + y \\ \text{s.t.} \quad 1.25 - x^2 - y &\leq 0 \\ x + y &\leq 1.6 \\ x &\geq 0 \\ y &= 0, 1 \end{aligned}$$

Note that the first constraint is concave is x and hence this is a nonconvex problem. Also note that the 0–1 y variable appears linearly and separably from the x variable.

If we use the **v3–GBD** approach, then we introduce one new variable x_1 and one additional equality constraint

$$x_1 - x = 0,$$

and the problem can be written in the following equivalent form:

$$\begin{aligned} \min_{x,x_1,y} \quad & 2x + y \\ \text{s.t.} \quad 1.25 - x_1 x - y &\le 0 \\ x_1 - x &= 0 \\ x + y &\le 1.6 \\ x,\ x_1 &\ge 0 \\ y &= 0,1 \end{aligned}$$

Careful analysis of the set of constraints that define the bounds reveals the following:

(i) For $y = 0$ we have $x^2 \ge 1.25$, and for $y = 1$ we have $x^2 \ge 0.25$ (see first inequality constraint).

(ii) From the second inequality constraint we similarly have that, for $y = 0$, $x \le 1.6$, while for $y = 1$, $x \le 0.6$.

Using the aforementioned observations we have a new lower bound for x (and hence x_1);

$$0.5 \le x \le 1.6.$$

The set of complicating variables for this example is defined as

$$\mathrm{y} = (x_1, y),$$

and it is a mixed set of a continuous and 0–1 variable. The primal problem takes the form:

$$\begin{aligned} \min_{x} \quad & 2x + y^k \\ \text{s.t.} \quad 1.25 - x_1^k x - y^k &\le 0 \\ x_1^k - x &= 0 \\ 0.5 \le x &\le 1.6 \end{aligned}$$

The relaxed master problem of the **v3–GBD** approach takes the form (assuming feasible primal):

$$\begin{aligned} \min \quad & \mu_B \\ \text{s.t.} \quad \mu_B &\ge L(x^k, \mathrm{y}, \lambda^k, \mu^k) \\ x_1 + y &\le 1.6 \\ 0.5 \le x_1 &\le 1.6 \\ y &= 0,1 \end{aligned}$$

where $L(x^k, y, \lambda^k, \mu^k) = 2x^k + y + \lambda^k(x_1 - x^k) + \mu^k(1.25 - x_1 x^k - y)$.

Note that the constraint $x + y \leq 1.6$ has been written as $x_1 + y \leq 1.6$ and since both x_1 and y are complicating variables it is moved directed to the relaxed master problem.

Note also that in this case the primal problem is a linear programming problem, while the relaxed master problem is a mixed-integer linear programming problem. The **v3–GBD** was applied to this problem from several starting points (see Floudas *et al.* (1989)) and the global solution was obtained in two iterations, even though the theoretical conditions for determining the global solution were not satisfied.

6.4 Outer Approximation, OA

6.4.1 Formulation

Duran and Grossmann (1986a; 1986b) proposed an Outer Approximation **OA** algorithm for the following class of MINLP problems:

$$
\begin{aligned}
& \min_{xy} \quad c^T y + f(x) && (6.13)\\
& \quad s.t. \quad g(x) + By \leq 0 \\
x \in X \quad &= \quad \{x : x \in \Re^n, A_1 x \leq a_1\} \subseteq \Re^n \\
y \in Y \quad &= \quad \{y : y \in \{0,1\}^q, A_2 y \leq a_2\}
\end{aligned}
$$

under the following conditions:

C1: X is a nonempty, compact, convex set and the functions

$$f : \Re^n \longrightarrow \Re,$$

$$g : \Re^n \longrightarrow \Re^p,$$

are convex in x.

C2: f and g are once continuously differentiable.

C3: A constraint qualification (e.g., Slater's) holds at the solution of every nonlinear programming problem resulting from (6.13) by fixing y.

Remark 1 Note that formulation (6.13) corresponds to a subclass of problem (6.2) that the **GBD** can address. This is due to the inherent assumptions of

(i) Separability in x and y; and

(ii) Linearity in y.

Also, note that problem (6.13) does not feature any nonlinear equality constraints. Hence, the implicit assumption in the **OA** algorithm is that

(iii) Nonlinear equalities can be eliminated algebraically or numerically.

Remark 2 Under the aforementioned assumptions (i) and (ii), problem (6.13) satisfies property (P) of Geoffrion (1972), and hence the **OA** corresponds to a subclass of **v1–GBD** (see sections 6.3.5.1) Furthermore, as we have seen in section 6.3.5.2, assumptions (i) and (ii) make the assumption imposed in **v2–GBD** valid (see remark of section 6.3.5.2) and therefore the **OA** can be considered as equivalent to **v2–GBD** with separability in x and y and linearity in y. Note though that the **v1–GBD** can handle nonlinear equality constraints.

Remark 3 Note that the set of constraints

$$g(x) + By \le 0,$$

can be written in the following equivalent form:

$$\begin{aligned} g(x) - Cx' &\le 0, \\ Cx' + By &\le 0, \end{aligned}$$

by introducing a new set of variables x', and therefore augmenting x to (x, x'), and a new set of constraints. Now, if we define the x variables as (x, x'), and the first constraints as $G(x) \le 0$, we have

$$\begin{aligned} G(x) &\le 0, \\ Cx + By &\le 0, \end{aligned}$$

where the first set of constraints is nonlinear in the x–type variables, while the second set of constraints in linear in both x and y variables. The penalty that we pay with the above transformation is the introduction of the x' and their associated constraints.

6.4.2 Basic Idea

The basic idea in **OA** is similar to the one in **GBD** that is, at each iteration we generate an upper bound and a lower bound on the MINLP solution. The upper bound results from the solution of the problem which is problem (6.13) with fixed y variables (e.g., $y = y^k$). The lower bound results from the solution of the master problem. The master problem is derived using primal information which consists of the solution point x^k of the primal and is based upon an outer approximation (linearization) of the nonlinear objective and constraints around the primal solution x^k. The solution of the master problem, in addition to the lower bound, provides information on the next set of fixed y variables (i.e., $y = y^{k+1}$) to be used in the next primal problem. As the iterations proceed, two sequences of updated upper bounds and lower bounds are generated which are shown to be nonincreasing and nondecreasing respectively. Then, it is shown that these two sequences converge within ϵ in a finite number of iterations.

Remark 1 Note that the distinct feature of **OA** versus **GBD** is that the master problem is formulated based upon primal information and outer linearization.

6.4.3 Theoretical Development

6.4.3.1 The Primal Problem

The primal problem corresponds to fixing the y variables in (6.13) to a 0–1 combination, which is denoted as y^k, and its formulation is

$$\begin{aligned} \min_{x} \quad & c^T y^k + f(x) \\ s.t. \quad & g(x) + By^k \le 0 \\ & x \in X = \{x : x \in \Re^n, A_1 x \le a_1\} \end{aligned}$$

Depending on the fixation point y^k, the primal problem can be feasible or infeasible, and these two cases are analyzed in the following:

Case (i): Feasible Primal

If the primal is feasible at iteration k, then its solution provides information on the optimal x^k, $f(x^k)$, and hence the current upper bound $UBD = c^T y^k + f(x^k)$. Using information on x^k, we can subsequently linearize around x^k the convex functions $f(x)$ and $g(x)$ and have the following relationships satisfied:

$$\begin{aligned} f(x) &\geq f(x^k) + \nabla f(x^k)(x - x^k), \quad \forall x^k \in X, \\ g(x) &\geq g(x^k) + \nabla g(x^k)(x - x^k), \quad \forall x^k \in X, \end{aligned}$$

due to convexity of $f(x)$ and $g(x)$.

Case (ii): Infeasible Primal

If the primal is infeasible at iteration k, then we need to consider the identification of a feasible point by looking at the constraint set:

$$g(x) + By^k \leq 0,$$

and formulating a feasibility problem in a similar way as we did for the **GBD** (see case (ii) of section 6.3.3.1).

For instance, if we make use of the l_1–minimization, we have

$$\begin{aligned} \min_{x \in X} \quad & \sum_{j=1}^{p} a_j \\ s.t. \quad g_j(x) + By^k &\leq a_j, \quad j = 1, 2, \ldots, p \\ a_j &\geq 0 \end{aligned}$$

Its solution will provide the corresponding x^l point based upon we can linearize the constraints:

$$g(x) \geq g(x^l) + \nabla g(x^l)(x - x^l), \quad \forall x^l,$$

where the right-hand side is a valid linear support.

6.4.3.2 The Master Problem

The derivation of the master problem in the **OA** approach involves the following two key ideas:

(i) Projection of (6.13) onto the y–space; and

(ii) Outer approximation of the objective function and the feasible region.

(i) Projection of (6.13) onto the y–space
Problem (6.13) can be written as

$$\begin{aligned} \min_{y} \quad & \inf_{x} c^T y + f(x) \\ \text{s.t.} \quad g(x) + By & \leq 0 \\ x & \in X \\ y & \in Y \end{aligned} \tag{6.14}$$

Note that the inner problem is written as infimum with respect to x to cover the case of having unbounded solution for a fixed y. Note also that $c^T y$ can be taken outside of the infimum since it is independent of x.

Let us define $v(y)$:

$$\begin{aligned} v(y) &= c^T y + \inf_{x} f(x) \\ \text{s.t.} \quad g(x) + By & \leq 0 \\ bx & \in X \end{aligned} \tag{6.15}$$

Remark 1 $v(y)$ is parametric in the y–variables, and it corresponds to the optimal value of problem (6.13) for fixed y (i.e., the primal problem $P(y^k)$).

Let us also define the set $V$of y's for which exist feasible solutions in the x variables, as

$$V = \{y : g(x) + By \leq 0, \quad \text{for some} \quad x \in X\}, \tag{6.16}$$

Then, problem (6.13) can be written as

$$\begin{aligned} \min_{y} \quad & v(y) \\ \text{s.t.} \quad & y \in Y \cap V \end{aligned} \tag{6.17}$$

Remark 2 Problem (6.17) is the projection of (6.13) onto the y–space. The projection needs to satisfy feasibility requirements, and this is represented in (6.17) by imposing $y \in Y \cap V$.

Remark 3 Note that we can replace the infimum with respect to $x \in X$ with the minimum with respect to $x \in X$, since for $y \in Y \cap V$ existence of solution x holds true due to the compactness assumption of X. This excludes the possibility for unbounded solution of the inner problem for fixed $y \in Y \cap V$.

Remark 4 The difficulty with solving (6.17) arises because both V and $v(y)$ are known implicitly. To overcome this difficulty, Duran and Grossmann (1986a) considered outer linearization of $v(y)$ and a particular representation of V.

(ii) Outer Approximation of $v(y)$

The outer approximation of $v(y)$ will be in terms of the intersection of an infinite set of supporting functions. These supporting functions correspond to linearizations of $f(x)$ and $g(x)$ at all $x^k \in X$. Then, the following conditions are satisfied:

$$\begin{aligned} f(x) &\geq f(x^k) + \nabla f(x^l)\,(x - x^k), \quad \forall x^k \in X, \\ g(x) &\geq g(x^k) + \nabla g(x^l)\,(x - x^k), \quad \forall x^k \in X, \end{aligned}$$

due to the assumption of convexity and once continuous differentiability. $\nabla f(x^k)$ represents the n–gradient vector of the $f(x)$ and $\nabla g(x^k)$ is the $(n \times p)$ Jacobian matrix evaluated at $x^k \in X$.

Remark 5 Note that the support functions are linear in x, and as a result $v(y)$ will be a mixed integer linear programming MILP problem.

The constraint qualification assumption, which holds at the solution of every primal problem for fixed $y \in Y \cap V$, coupled with the convexity of $f(x)$ and $g(x)$, imply the following Lemma:

Lemma 6.4.1

$$v(y) = \begin{bmatrix} \min\limits_{x} c^T y + f(x) & & \\ \text{s.t.} \quad g(x) + By & \leq & 0 \\ x & \in & X \end{bmatrix}$$

$$= \begin{bmatrix} \min\limits_{x} & & c^T y + f(x^k) + \nabla f(x^k)\,(x - x^k) \\ \text{s.t.} & 0 \geq & g(x^k) + \nabla g(x^k)\,(x - x^k) + By \\ & x \in & X \end{bmatrix} \quad \forall x^k \in X.$$

Remark 6 It suffices to include those linearizations of the constraints that are active at (x^k, y^k). This implies that fewer constraints are needed in the master problem.

Then, by substituting $v(y)$ of the above Lemma to the projection problem (6.17) we have

$$\begin{aligned} \min_{x} \min_{y} \quad & c^T y + f(x^k) + \nabla f(x^k)\,(x - x^k) \\ \text{s.t.} \quad 0 &\geq g(x^k) + \nabla g(x^k)\,(x - x^k) + By \\ x &\in X \\ y &\in Y \cap V \end{aligned} \tag{6.18}$$

By combining the min operators and introducing a scalar μ_{OA}, problem (6.18) can be written in the following equivalent form:

$$
\begin{array}{llll}
\min\limits_{x,y,\mu_{OA}} & c^T y + \mu_{OA} & & (6.19) \\
\text{s.t.} & \left.\begin{array}{rcl} \mu_{OA} & \geq & f(x^k) + \nabla f(x^k)(x - x^k), \forall k \in \mathbf{F} \\ 0 & \geq & g(x^k) + \nabla g(x^k)(x - x^k) + By \end{array}\right\}, \forall k \in \mathbf{F} & & \\
& x \in X & & \\
& y \in Y \cap V & &
\end{array}
$$

where $\mathbf{F} = \{k : x^k$ is a feasible solution to the primal $P(y^k)\}$.

Duran and Grossmann (1986a) made the additional assumption that we can replace $y \in Y \cap V$ with $y \in Y$ using as argument that a representation of the constraints $y \in Y \cap V$ is included in the linearizations of problem (6.19) provided that the appropriate integer cuts that exclude the possibility of generation of the same integer combinations are introduced. Subsequently, they defined the master problem of the **OA** as

$$
\begin{array}{llll}
\min\limits_{x,y,\mu_{OA}} & c^T y + \mu_{OA} & & (6.20) \\
\text{s.t.} & \left.\begin{array}{rcl} \mu_{OA} & \geq & f(x^k) + \nabla f(x^k)(x - x^k), \forall k \in \mathbf{F} \\ 0 & \geq & g(x^k) + \nabla g(x^k)(x - x^k) + By \end{array}\right\}, \forall k \in \mathbf{F} & & \\
& x \in X & & \\
& y \in Y & & \\
& \sum\limits_{i \in \mathbf{B}^k} y_i^k - \sum\limits_{i \in \mathbf{NB}^k} y_i^k \leq |\mathbf{B}^k| - 1, \quad k \in \mathbf{F} & &
\end{array}
$$

Remark 7 Note that the master problem (6.20) is a mixed-integer linear programming MILP problem since it has linear objective and constraints, continuous variables (x, μ_{OA}) and 0-1 variables (y). Hence, it can be solved with standard branch and bound algorithms.

Remark 8 The master problem consists of valid linear supports, and hence relaxations of the nonlinear functions, for all points x^k that result from fixing $y = y^k \in Y$ as stated by (6.20). As a result it represents a relaxation of the original MINLP model (6.13), and hence it is a lower bound on its solution, and it is identical to its solution if all supports are included.

Remark 9 It is not efficient to solve the master problem (6.20) directly, since we need to know all feasible x^k points which implies that we have to solve all the primal problems $P(y^k)$, $y \in Y$ (i.e., exhaustive enumeration of 0–1 alternatives). Instead, Duran and Grossmann (1986a) proposed a relaxation of the master problem which will be discussed in the next section.

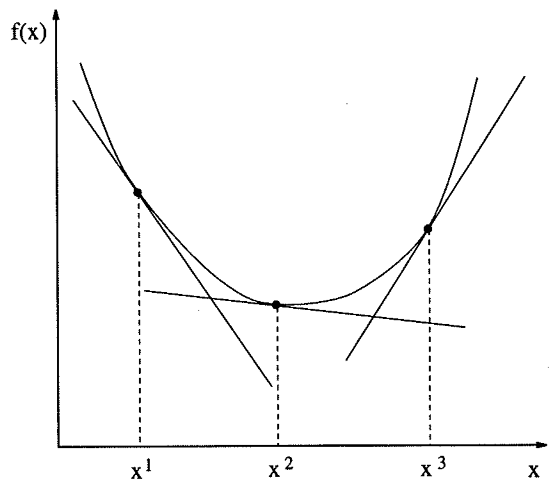

Figure 6.6: Linear support functions.

Remark 10 (Geometrical Interpretation of Master in OA) The master problem of the **OA** can be interpreted geometrically by examining the effect of the linear support function (i.e., outer linearizations) on the objective function and the constraints. Figure 6.6 shows the linear supports of the objective function $f(\boldsymbol{x})$ taken at $\boldsymbol{x}^1, \boldsymbol{x}^2$, and $\boldsymbol{x}^3$.

Note that the linear supports are the tangents of $f(\boldsymbol{x})$ at $\boldsymbol{x}^1, \boldsymbol{x}^2$, and $\boldsymbol{x}^3$ and that they underestimate the objective function. Also, note that accumulation of these linear underestimating supports results in a better approximation of the objective function $f(\boldsymbol{x})$ from the outside (i.e., outer approximation). Notice also that the linear underestimators are valid because $f(\boldsymbol{x})$ is convex in $\boldsymbol{x}$.

Figure 6.7 shows the outer approximation of the feasible region consisting of two linear (i.e., g_3, g_4) and two nonlinear (g_1, g_2) inequalities at a point $\boldsymbol{x}^1$.

u_{11} is the linear support of g_1, while u_{12} is the linear support of g_2, and they both result from linearizing g_1 and g_2 around the point $\boldsymbol{x}^1$. Note that the feasible region defined by the linear supports and g_3, g_4 includes the original feasible region, and hence the relaxation of the nonlinear constraints by taking their linearization corresponds to an overestimation of the feasible region.

Note also that the linear supports provide valid overestimators of the feasible region only because the functions $\boldsymbol{g}(\boldsymbol{x})$ are convex.

With the above in mind, then the master problem of the **OA** can be interpreted geometrically as a relaxation of the original MINLP (in the limit of taking all $\boldsymbol{x}^k$ points it is equivalent) defined as

(i) underestimating the objective function; and

(ii) overestimating the feasible region.

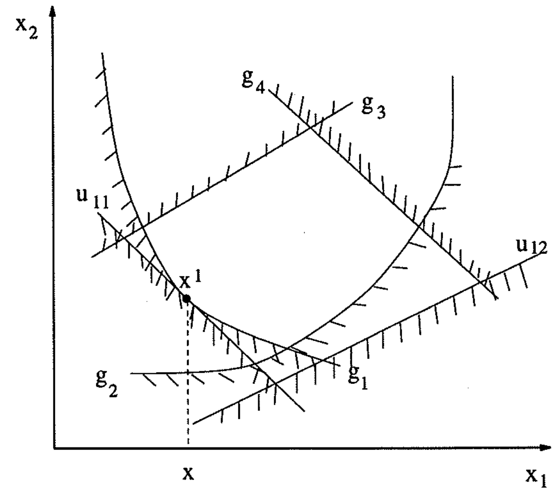

Figure 6.7: Geometric interpretation of outer approximation OA

As such, its solution provides a lower bound on the solution of the original MINLP problem.

Remark 11 Note that at every iteration k we need to add in the master problem of **OA** the linearizations of the active constraints. This implies that the master problem may involve a large number of constraints as the iterations proceed. For instance, if we have $Y = \{0,1\}^q$ and (6.13) has m inequalities, then the master problem should have

$$2^q\,(m+1)$$

constraints so as to be exactly equivalent to (6.13). This provides the motivation for solving the master problem via a relaxation strategy.

6.4.4 Algorithmic Development

The central point in the algorithmic development of **OA** is the solution of the master problem since it is quite straightforward to solve the primal problem which is a convex nonlinear problem.

The natural approach of solving the master problem is relaxation; that is, consider at each iteration the linear supports of the objective and constraints around all previously linearization points. This way, at each iteration a new set of linear support constraints are added which improve the relaxation and therefore the lower bound.

6.4.4.1 Algorithmic Statement of OA

Step 1: Let an initial point $y^1 \in Y$ or $Y \cap V$ if available. Solve the resulting primal problem $P(y^1)$ and obtain an optimal solution x^1. Set the iteration counter $k = 1$. Set the current upper bound $UBD = P(y^1) = v(y^1)$.

Step 2: Solve the relaxed master (**RM**) problem

$$
(\mathbf{RM})\quad
\left[
\begin{array}{lll}
\min\limits_{\boldsymbol{x},\boldsymbol{y},\mu_{OA}} & \boldsymbol{c}^T\boldsymbol{y} + \mu_{OA} & \\
\text{s.t.} & \left.\begin{array}{rcl} \mu_{OA} & \geq & f(\boldsymbol{x}^k) + \nabla f(\boldsymbol{x}^k)\,(\boldsymbol{x} - \boldsymbol{x}^k) \\ 0 & \geq & \boldsymbol{g}(\boldsymbol{x}^k) + \nabla \boldsymbol{g}(\boldsymbol{x}^k)\,(\boldsymbol{x} - \boldsymbol{x}^k) + B\boldsymbol{y} \end{array}\right\}, & \forall k \in \mathbf{F} \\
& \boldsymbol{x} \in X & \\
& \boldsymbol{y} \in Y & \\
& \sum\limits_{i\in\mathbf{B}^k} y_i^k - \sum\limits_{i\in\mathbf{NB}^k} y_i^k \leq |\mathbf{B}^k| - 1, & k = 1, 2, \ldots, K-1
\end{array}
\right.
$$

Let $(\boldsymbol{y}^{k+1}, \mu_{OA}^k)$ be the optimal solution of the relaxed master problem, where $\mu_{OA}^k + \boldsymbol{c}^T\boldsymbol{g}^{k+1}$ is the new current lower bound on (6.13), $LBD = \mu_{OA}^k + \boldsymbol{c}^T\boldsymbol{y}^{k+1}$ and $\boldsymbol{y}^{k+1}$ is the next point to be considered in the primal problem $P(\boldsymbol{y}^{k+1})$. If $UBD - LBD \leq \epsilon$, then terminate. Otherwise, return to step 1.

If the master problem does not have a feasible solution, then terminate. The optimal solution is given by the current upper and its associated optimal vectors $(\boldsymbol{x}, \boldsymbol{y})$.

Remark 1 Note that in step 1 instead of selecting a $\boldsymbol{y}^1 \in Y$ or $Y \cap V$ we can solve the continuous relaxation of (6.13) (i.e., treat $0 \leq \boldsymbol{y} \leq 1$) and set the $\boldsymbol{y}$–variables to their closest integer value. This way, it is still possible that the resulting primal is infeasible which implies that, according to the **OA** algorithm proposed by Duran and Grossmann (1986a), we eliminate this infeasible combination in the relaxed master problem of the first iteration and continue until another $\boldsymbol{y}$ is found that belong to $Y \cap V$. It is clear though at this point that in the relaxed master no information is provided that can be used as linearization point for such a case, apart from the integer cut.

As we have discussed in the outer approximation of $v(\boldsymbol{y})$, Duran and Grossmann (1986a) made the assumption that $\boldsymbol{y} \in Y$ instead of $\boldsymbol{y} \in Y \cap V$ and introduced the integer cut constraints which they claimed make their assumption valid. Fletcher and Leyffer (1994), however, presented a counter example which can be stated as follows:

$$
\begin{array}{rrl}
& \min\limits_{x,\,y} & -2y - x \\
s.t. & x^2 + y & \leq \; 0 \\
& y & \in \; \{-1, 1\}
\end{array}
$$

which has as solution $(x^*, y^*) = (1, -1)$, and objective equal to 1. Let us start at $y^1 = -1$, for which $x^1 = 2$. The master problem according to Duran and Grossmann (1986a) can be written as

$$
\begin{array}{rrl}
& \min & -2y + \mu_{OA} \\
\text{s.t.} & \mu_{OA} & \geq \; -x \\
& 0 & \geq \; 1 + 2\,(x - 1) + y \\
& y & \in \; \{-1, 1\} \cap \{y = -1\}
\end{array}
$$

which has as solution

$$
(y^2, -2y^2 + \mu_{OA}^1) = (1, -2).
$$

Continuing to the next primal $P(y^2)$ we find that

$$\begin{array}{ll} \min\limits_{x} & -2 - x \\ \text{s.t.} & x^2 + 1 \leq 0 \end{array}$$

is infeasible.

Therefore, the integer cut did not succeed in eliminating infeasible combinations.

This example shows clearly that information from infeasible primal problems needs to be included in the master problems and that a representation of the constraint

$$y \in Y \cap V$$

in terms of linear supports has to be incorporated. In other words, we would like to ensure that integer combinations which produce infeasible primal problems are also infeasible in the master problem, and hence they are not generated. Fletcher and Leyffer (1994) showed that for the case of infeasible primal problems the following constraints, which are the linear supports for this case,

$$0 \geq g(x^l) + \nabla g(x^l)(x - x^l) + By, \quad l \in \bar{\mathbf{F}},$$

$$\text{where} \quad \bar{\mathbf{F}} = \{j : P(y^l) \text{ is infeasible and } x^l \text{ solves the feasibility problem}\}$$

need to be incorporated in the master problem. Then, the master problem of **OA** takes the form:

$$\left[\begin{array}{ll} \min\limits_{x,y,\mu_{OA}} & c^T y + \mu_{OA} \\ \text{s.t.} & \left.\begin{array}{l} \mu_{OA} \geq f(x^k) + \nabla f(x^k)(x - x^k), \forall k \in \mathbf{F} \\ 0 \geq g(x^k) + \nabla g(x^k)(x - x^k) + By \end{array}\right\}, \forall k \in \mathbf{F} \\ & \left. 0 \geq g(x^l) + \nabla g(x^l)(x - x^l) + By \right\}, \forall l \in \bar{\mathbf{F}} \\ & x \in X \\ & y \in Y \end{array}\right. \tag{6.21}$$

Appropriate changes need also to be made in the relaxed master problem described in the algorithmic procedure section. Note also that the correct formulation of the master problem via (6.21) increases the number of constraints to be included significantly.

6.4.4.2 Finite Convergence of OA

For problem (6.13), Duran and Grossmann (1986a) proved finite convergence of the **OA** stated in 6.4.4.1 which is as follows:

> " *If conditions C1, C2, C3 hold and Y is a discrete set, then the* **OA** *terminates in a finite number of iterations.*"

Duran and Grossmann (1986a) also proved that a comparison of **OA** with **v2–GBD** yields always:

$$LBD_{\mathbf{OA}} \geq LBD_{\mathbf{v2\text{–}GBD}}$$

where $LBD_{\mathbf{OA}}, LBD_{\mathbf{v2\text{–}GBD}}$ are the lower bounds at corresponding iterations of the **OA** and **v2–GBD**.

Remark 1 The aforementioned property of the bounds is quite important because it implies that the **OA** for convex problems will terminate in fewer iterations that the **v2–GBD**. Note though that this does not necessarily imply that the **OA** will terminate faster, since the relaxed master of **OA** has many constraints as the iterations proceed, while the relaxed master of **v2–GBD** adds only one constraint per iteration.

6.5 Outer Approximation with Equality Relaxation, OA/ER

6.5.1 Formulation

To handle explicitly nonlinear equality constraints of the form $\boldsymbol{h}(\boldsymbol{x}) = 0$, Kocis and Grossmann (1987) proposed the outer approximation with equality relaxation **OA/ER** algorithm for the following class of MINLP problems:

$$
\begin{aligned}
\min_{\boldsymbol{x},\boldsymbol{y}} \quad & c^T\boldsymbol{y} + f(\boldsymbol{x}) \\
s.t. \quad & \boldsymbol{h}(\boldsymbol{x}) = \boldsymbol{0} \\
& \boldsymbol{g}(\boldsymbol{x}) \leq \boldsymbol{0} \\
& C\boldsymbol{x} + B\boldsymbol{y} \leq \boldsymbol{0} \\
& \boldsymbol{x} \in \boldsymbol{X} = \{\boldsymbol{x} : \boldsymbol{x} \in \Re^n, \Lambda_1\boldsymbol{x} \leq \boldsymbol{a}_1\} \subseteq \Re^n \\
& \boldsymbol{y} \in \boldsymbol{Y} = \{\boldsymbol{y} : \boldsymbol{y} \in \{0,1\}^q, A_2\boldsymbol{y} \leq \boldsymbol{a}_2\}
\end{aligned}
\tag{6.22}
$$

under the following conditions:

C1: $\boldsymbol{X}$ is a nonempty, compact, convex set, and the functions satisfy the conditions:

$f(\boldsymbol{x})$ is convex in $\boldsymbol{x}$,
$g_i(\boldsymbol{x})$ $i \in \boldsymbol{I}_{IN} = \{i : g_i(\boldsymbol{x}) < 0\}$ are convex in $\boldsymbol{x}$,
$g_i(\boldsymbol{x})$ $i \in \boldsymbol{I}_{EQ} = \{i : g_i(\boldsymbol{x}) = 0\}$ are quasi–convex in $\boldsymbol{x}$, and
$T\boldsymbol{h}(\boldsymbol{x})$ are quasi–convex in $\boldsymbol{x}$,

where T is a diagonal matrix $(m \times m)$ with elements t_{ii}

$$
t_{ii} = \left\{ \begin{array}{rl} -1 & \text{if} \quad \lambda_i < 0 \\ +1 & \text{if} \quad \lambda_i > 0, \\ 0 & \text{if} \quad \lambda_i = 0 \end{array} \right\} \qquad i = 1, 2, \ldots, m
$$

and λ_i are the Lagrange multipliers associated with the m equality constraints.

C2: $f, \boldsymbol{h}$, and $\boldsymbol{g}$ are continuously differentiable.

C3: A constraint qualification holds at the solution of every nonlinear programming problem resulting from (6.21) by fixing $\boldsymbol{y}$.

Remark 1 The nonlinear equalities $\boldsymbol{h}(\boldsymbol{x}) = \boldsymbol{0}$ and the set of linear equalities which are included in $\boldsymbol{h}(\boldsymbol{x}) = \boldsymbol{0}$, correspond to mass and energy balances and design equations for chemical process systems, and they can be large. Since the nonlinear equality constraints cannot be treated explicitly by the **OA** algorithm, some of the possible alternatives would be to perform:

(i) Algebraic elimination of the nonlinear equalities;

(ii) Numerical elimination of the nonlinear equalities; and

(iii) Relaxation of the nonlinear equalities to inequalities.

Alternative (i) can be applied successfully to certain classes of problems (e.g., design of batch processes Vaselenak *et al.* (1987); synthesis of gas pipelines (e.g., Duran and Grossmann (1986a)). However, if the number of nonlinear equality constraints is large , then the use of algebraic elimination is not a practical alternative.

Alternative (ii) involves the numerical elimination of the nonlinear equality constraints at each iteration of the **OA** algorithm through their linearizations. Note though that these linearizations may cause computational difficulties since they may result in singularities depending on the selection of decision variables. In addition to the potential problem of singularities, the numerical elimination of the nonlinear equality constraints may result in an increase of the nonzero elements, and hence loss of sparsity, as shown by Kocis and Grossmann (1987).

The aforementioned limitations of alternatives (i) and (ii) motivated the investigation of alternative (iii) which forms the basis for the **OA/ER** algorithm.

Remark 2 Note that condition C1 of the **OA/ER** involves additional conditions of quasi–convexity in x for

$$g_i(x), \quad i \in I_{EQ}, \text{ and } \mathrm{T}h(x).$$

Also note that the diagonal matrix T defines the direction for the relaxation of the equalities into inequalities, and it is expected that such a relaxation can only be valid under certain conditions.

6.5.2 Basic Idea

The basic idea in **OA/ER** is to relax the nonlinear equality constraints into inequalities and subsequently apply the **OA** algorithm. The relaxation of the nonlinear equalities is based upon the sign of the Lagrange multipliers associated with them when the primal (problem (6.21) with fixed y) is solved. If a multiplier λ_i is positive then the corresponding nonlinear equality $h_i(x) = 0$ is relaxed as $h_i(x) \leq 0$. If a multiplier λ_i is negative, then the nonlinear equality is relaxed as $-h_i(x) \leq 0$. If, however, $\lambda_i = 0$, then the associated nonlinear equality constraint is written as $0 \cdot h_i(x) = 0$, which implies that we can eliminate from consideration this constraint. Having transformed the nonlinear equalities into inequalities, in the sequel we formulate the master problem based on the principles of the **OA** approach discussed in section 6.4.

6.5.3 Theoretical Development

Since the only difference between the **OA** and the **OA/ER** lies on the relaxation of the nonlinear equality constraints into inequalities, we will present in this section the key result of the relaxation and the master problem of the **OA/ER**.

Property 6.5.1
If conditions C1, C3 are satisfied, then for fixed $y = y^k$ *problem (6.22) is equivalent to (6.23):*

$$\min_{x} \quad c^T y^k + f(x) \tag{6.23}$$

$$\begin{aligned} \text{s.t.} \quad & T^k \boldsymbol{h}(\boldsymbol{x}) \le \boldsymbol{0} \\ & \boldsymbol{g}(\boldsymbol{x}) \le \boldsymbol{0} \\ & C\boldsymbol{x} + B\boldsymbol{y}^k \le \mathbf{d} \\ & \boldsymbol{x} \in \boldsymbol{X} = \{\boldsymbol{x} : \boldsymbol{x} \in \Re^n, A_1\boldsymbol{x} \le a_1\} \subseteq \Re^n \end{aligned}$$

where T *is a diagonal matrix* $(m \times m)$ *with elements* t_{ii}^k *defined as*

$$t_{ii}^k = \left\{ \begin{array}{rl} -1 & \textit{if} \;\; \lambda_i^k < 0 \\ +1 & \textit{if} \;\; \lambda_i^k > 0, \\ 0 & \textit{if} \;\; \lambda_i^k = 0 \end{array} \right\} \qquad i = 1, 2, \ldots, m$$

Remark 1 Under the aforementioned conditions, the primal problems of (6.22) and (6.23) have unique local solutions which are in fact their respective global solutions since the **KKT** conditions are both necessary and sufficient.

Remark 2 If condition C1 is not satisfied, then unique solutions are not theoretically guaranteed for the NLP primal problems, and as a result the resulting master problems may not provide valid lower bounds. This is due to potential failure of maintaining the equivalence of (6.22) and (6.23). It should be noted that the equivalence between the primal problems of (6.22) and (6.23) needs to be maintained at every iteration. This certainly occurs when the Lagrange multipliers are invariant in sign throughout the iterations k and they satisfy condition C1 at the first iteration.

Remark 3 After the relaxation of the nonlinear equalities we deal with an augmented set of inequalities:

$$\begin{aligned} T^k \boldsymbol{h}(\boldsymbol{x}) &\le \boldsymbol{0}, \\ \boldsymbol{g}(\boldsymbol{x}) &\le \boldsymbol{0}, \end{aligned}$$

for which the principles of **OA** described in section 6.4 can be applied.

Illustration 6.5.1 To illustrate the relaxation of equalities into inequalities we consider the following simple example taken from Kocis and Grossmann (1987):

$$\begin{aligned} \min \quad & -y + 2x_1 + x_2 \\ \text{s.t.} \quad & x_1 - 2e^{-x_2} = 0 \\ & -x_1 + x_2 + y \le 0 \\ & 0.5 \le x_1 \le 1.4 \\ & y = 0, 1 \end{aligned} \tag{6.24}$$

By fixing $y^1 = 0$ and solving the resulting primal problem, which is nonlinear, we have

$$OBJ = 2.558, \tag{6.25}$$

$$(x_1, x_2) = (0.853, 0.853), \tag{6.26}$$

$$\lambda = -1.619 < 0. \tag{6.27}$$

We have only one equality constraint, and since its Lagrange multiplier λ is negative, then the T^1 matrix (1×1) is

$$t_{11} = -1.$$

Then, the equality is relaxed as:

$$\begin{aligned} T^1 h(x) &\leq 0, \\ -\left(x_1 - 2\exp\left(-x_2\right)\right) &\leq 0, \\ 2\exp\left(-x_2\right) - x_1 &\leq 0. \end{aligned}$$

Remark 4 Note that the relaxed equality

$$2\exp\left(-x_2\right) - x_1 \leq 0,$$

is convex in x_1 and x_2, and therefore condition C1 is satisfied.

Remark 5 If instead of the above presented relaxation procedure we had selected the alternative of algebraic elimination of x_1 from the equality constraint, then the resulting MINLP is

$$\begin{aligned} \min \quad & -y + 4\exp\left(-x_2\right) + x_2 \\ s.t. \quad & -2\exp -x_2 + x_2 + y \leq 0 \\ & 0.357 \leq x_2 \leq 1.386 \\ & y = 0, 1 \end{aligned} \tag{6.28}$$

By selecting $y = 0$ as in our illustration, we notice that the inequality constraint in (6.28) is in fact nonconvex in x_2. Hence, application application of **OA** to (6.28) cannot guarantee global optimality due to having nonconvexity in x_2.

This simple example clearly demonstrates the potential drawback of the algebraic elimination alternative.

Remark 6 If we had selected x_2 to be algebraically eliminated using the equality constraint, then the resulting MINLP would be

$$\begin{aligned} \min \quad & -y + 2x_1 - 2\ln x_1 + \ln x_2 \\ s.t. \quad & -x_1 - \ln x_1 + \ln 2 + y \leq 0 \\ & 0.5 \leq x_1 \leq 1.4 \\ & y = 0, 1 \end{aligned}$$

Note that in this case the inequality constraint is convex in x_1 and so is the objective function. Hence, application of **OA** can guarantee its global solution.

As a final comment on this example, we should note that when algebraic elimination is to be applied, care must be taken so as to maintain the convexity and in certain cases sparsity characteristics if possible.

6.5.3.1 The Master Problem

The master problem of the **OA/ER** algorithm is essentially the same as problem (6.20) described in section 6.4.3.2, with the difference being that the vector of inequality constraints will be augmented by the addition of the relaxed equalities:

$$T^k h(x) \leq 0.$$

The general form of the relaxed master problem for the **OA/ER** algorithm is

$$\left[\begin{array}{ll}
\min\limits_{x,y,\mu} & c^T y + \mu_{OA} \\
\text{s.t.} & \left.\begin{array}{rcl} \mu & \geq & f(x^k) + \nabla f(x^k)\,(x - x^k) \\ 0 & \geq & g(x^k) + \nabla g(x^k)\,(x - x^k) \\ 0 & \geq & T^k \left[h(x^k) + \nabla h(x^k)^T \left(x - x^k\right)\right] \end{array}\right\}, \ \forall k = 1, 2, \ldots, K \\
& Cx + By \leq \mathbf{d} \\
& x \in X = \{x : x \in \Re^n, A_1 x \leq a_1\} \subseteq R^n \\
& y \in Y = \{y : y \in \{0,1\}^q, A_2 y \leq a_2\} \\
& \sum\limits_{i \in \mathbf{B}^k} y_i^k - \sum\limits_{i \in \mathbf{NB}^k} y_i^k \leq |B^k| - 1, k = 1, 2, \ldots, K \\
& Z_L^{k-1} \leq c^T y + \mu \leq Z_U
\end{array}\right. \tag{6.29}$$

where Z_L^{k-1} is the lower bound at iteration $k-1$, and Z_U is the current upper bound and are used to expedite the solution of (6.29) and to have infeasibility as the termination criterion.

Remark 1 The right-hand sides of the first three sets of constraints in (6.29) can be written in the following form:

$$\begin{array}{rcl}
\nabla f(x^k)^T x - [\nabla f(x^k)^T x^k - f(x^k)] & = & (w^k)^T x - w_o^k \\
\nabla g(x^k)^T x - [\nabla g(x^k)^T x^k - g(x^k)] & = & (S^k) x - s^k \\
\nabla h(x^k)^T x - [\nabla h(x^k)^T x^k] & = & (R^k) x - r^k
\end{array}$$

$$\text{where} \quad \begin{array}{rclcrcl}
w^k & = & \nabla f(x^k) & , & w_o^k & = & \nabla f(x^k)^T x^k - f(x^k), \\
S^k & = & \nabla g(x^k) & , & s^k & = & \nabla g(x^k)^T x^k - g(x^k), \\
R^k & = & \nabla h(x^k) & , & r^k & = & \nabla h(x^k)^T x^k.
\end{array}$$

Note that $h(x^k) = 0$.

It should be noted that $w^k, w_o^k, S^k, s^k, R^k, r^k$ can be calculated using the above expressions immediately after the primal problem solution since already information is made available at the solution of the primal problem.

Remark 2 The right-hand sides of the first three sets of constraints are the support functions that are represented as outer approximations (or linearizations) at the current solution point x^k of the primal problem. If condition C1 is satisfied then these supports are valid underestimators and as a result the relaxed master problem provides a valid lower bound on the global solution of the MINLP problem.

Remark 3 If linear equality constraints in x exist in the MINLP formulation, then these are treated as a subset of the $h(x) = 0$ with the difference that we do not need to compute their corresponding T matrix but simply incorporate them as linear equality constraints in the relaxed master problem directly.

Remark 4 The relaxed master problem is a mixed integer linear programming MILP problem which can be solved for its global solution with standard branch and bound codes. Note also that if $f(x), h(x), g(x)$ are linear in x, then we have an MILP problem. As a result, since the relaxed master is also an MILP problem, the **OA/ER** should terminate in 2 iterations.

6.5.4 Algorithmic Development

The **OA/ER** algorithm can be stated as follows:

Step 1: Let an initial point $y^1 \in Y$, or $y^1 \in Y \cap V$ if available. Solve the resulting primal problem $P(y^1)$ and obtain an optimal solution x^1 and an optimal multiplier vector λ^1 for the equality constraints $h(x) = 0$. Set the current upper bound $UBD = P(y^1) = v(y^1)$.

Step 2: Define the $(m \times m)$ matrix T^k. Calculate $w^k, w_o^k, s^k, S^k, R^k, r^k$.

Step 3: Solve the relaxed master problem **(RM)**:

$$
\begin{aligned}
Z_L^K = \min_{x,y,\mu} \quad & c^T y + \mu \qquad\qquad \textbf{(RM)} \\
\text{s.t.} \quad & \left.\begin{aligned} \mu &\geq (w^k)^T x - w_o^k \\ 0 &\geq (S^k)^T x - s^k \\ 0 &\geq T^k R^k x - T^k r^k \end{aligned}\right\} , k = 1, 2, \ldots, K \\
& Cx + By \leq \mathrm{d} \\
& x \in X = \{x : x \in \Re^n, A_1 x \leq a_1\} \subseteq R^n \\
& y \in Y = \{y : y \in \{0,1\}^q, A_2 y \leq a_2\} \\
& \sum_{i \in \mathbf{B}^k} y_i^k - \sum_{i \in \mathbf{NB}^k} y_i^k \leq |B^k| - 1, \quad k = 1, 2, \ldots, K-1 \\
& Z_L^{k-1} \leq c^T y + \mu \leq Z_U
\end{aligned}
$$

If the relaxed master problem is feasible then, let (y^{k+1}, μ^k) be the optimal solution, where $Z_L^k = c^T y^{k+1} + \mu^K$ is the new current lower bound on (6.22), and y^{k+1} the next point to be considered in the primal problem $P(y^{k+1})$.

If $UBD - Z_L^k \leq \epsilon$, then terminate. Otherwise, return to step 1.

If the master problem is infeasible, then terminate.

Remark 1 In the case of infeasible primal problem, we need to solve a feasibility problem. One way of formulating this feasibility problem is the following:

$$\begin{aligned} \min_{\boldsymbol{x}} \quad & \alpha \\ s.t. \quad & \boldsymbol{h}(\boldsymbol{x}) = \boldsymbol{0} \\ & \boldsymbol{g}(\boldsymbol{x}) \leq \boldsymbol{0} \\ & C\boldsymbol{x} + B\boldsymbol{y}^K - \mathbf{d} \leq \alpha \\ & \boldsymbol{x} \in X = \{\boldsymbol{x} : \boldsymbol{x} \in \Re^n, A_1\boldsymbol{x} \leq a_1\} \\ & \alpha \geq 0 \end{aligned} \tag{6.30}$$

Note that in (6.30) we maintain $\boldsymbol{h}(\boldsymbol{x}) = \boldsymbol{0}$ and $\boldsymbol{g}(\boldsymbol{x}) \leq \boldsymbol{0}$ while we allow the $C\boldsymbol{x} + B\boldsymbol{y}^K - \mathbf{d} \leq \boldsymbol{0}$ constraint to be relaxed by α, which we minimize in the objective function. Kocis and Grossmann (1987) suggested the feasibility problem (6.30) and they also proposed the use of the Lagrange multipliers associated with the equality constraints of (6.30) so as to identify the matrix T^k. In case however, that problem (6.30) has no feasible solution, then Kocis and Grossmann (1987) suggested exclusion of this integer combination by an integer cut and solution of the relaxed master problem. This, however, may not be correct using the arguments of Fletcher and Leyffer (1994).

Remark 2 Following similar arguments to those of **OA**, the **OA/ER** algorithm attains finite convergence to the global minimum as long as C1, C2, and C3 are satisfied.

Remark 3 Kocis and Grossmann (1989a) suggested another alternative formulation of the feasibility problem, in which a penalty-type contribution is added to the objective function; that is,

$$\begin{aligned} \min_{\boldsymbol{x}} \quad & c^T\boldsymbol{y} + f(\boldsymbol{x}) + p\,\alpha \\ \text{s.t.} \quad & \boldsymbol{h}(\boldsymbol{x}) = \boldsymbol{0} \\ & \boldsymbol{g}(\boldsymbol{x}) \leq \alpha \\ & C\boldsymbol{x} + B\boldsymbol{y}^K \leq \alpha \\ & \boldsymbol{x} \in X = \{\boldsymbol{x} : \boldsymbol{x} \in \Re^n, A_1\boldsymbol{x} \leq a_1\} \subseteq \Re^n \end{aligned} \tag{6.31}$$

A feasible solution to the primal problem exists when the penalty term is driven to zero. If the primal does not have a feasible solution, then the solution of problem (6.31) corresponds to minimizing the maximum violation of the inequality constraints (nonlinear and linear in $\boldsymbol{x}$). A general analysis of the different types of feasibility problem is presented is section 6.3.3.1.

6.5.5 Illustration

This example is a slightly modified version of the small planning problem considered by Kocis and Grossmann (1987). A representation of alternatives is shown in Figure 6.8 for producing product

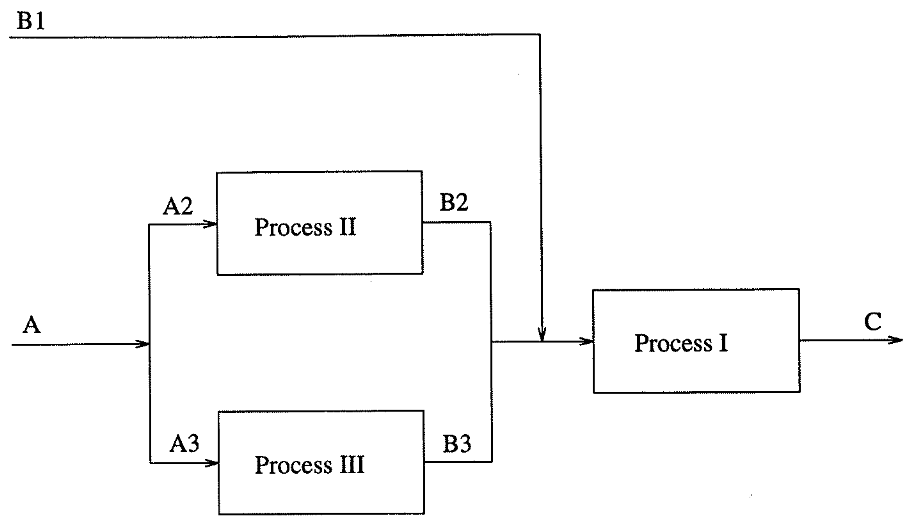

Figure 6.8: Planning problem

C from raw materials A, B via processes I, II, III. Product C can be produced through process I only, process I and III, process I and II. Processes II and III cannot take place simultaneously.

The investment costs of the processes are

$$\begin{array}{ll} \text{Process I:} & 3.5y_1 + 2C, \\ \text{Process II:} & 1.0y_2 + 1.0B_2, \\ \text{Process III:} & 1.5y_3 + 1.2B_3, \end{array}$$

where y_1, y_2, y_3 are binary variables denoting existence or nonexistence of the processes; B_2, B_3, and C are the flows of products of the processes.

The revenue expressed as a difference of selling the product C minus the cost of raw materials A,B is

$$Revenue = 13C - 1.8\,(A_2 + A_3) - 7B_1,$$

The objective is to minimize the costs minus the revenue, which takes the form:

$$\begin{aligned} \text{Objective} &= \text{Costs} - \text{Revenue} \\ &= (3.5y_1 + 2C) + (y_2 + B_2) + (1.5y_3 + 1.2B_3) \\ &\quad -13C + 1.8\,(A_2 + A_3) + 7B_1 \\ &= -11C + 7B_1 + B_2 + 1.2B_3 + 1.8A_2 + 1.8A_3 \\ &\quad +3.5y_1 + y_2 + 1.5y_3. \end{aligned}$$

The mass balances for the representation of alternatives shown in Figure 8 are

$$\begin{array}{ll} \text{Process I:} & C - 0.9\,(B_1 + B_2 + B_3) = 0, \\ \text{Process II:} & B_2 - \ln(1 + A_2) = 0, \\ \text{Process III:} & B_3 - 1.2\ln(1 + A_3) = 0. \end{array}$$

The bounds on the outputs of each process are

$$\text{Process I:} \quad C \leq 1,$$

$$\text{Process II:} \quad B_2 \leq \tfrac{1}{0.9},$$

$$\text{Process III:} \quad B_3 \leq \tfrac{1}{0.9}.$$

Note that imposing the upper bound of 1 on C automatically sets the upper bounds on B_2 and B_3.

The logical constraints for the processes are

$$\begin{array}{ll} \text{Process I:} & C \leq 1y_1, \\ \text{Process II:} & B_2 \leq \tfrac{1}{0.9}y_2, \\ \text{Process III:} & B_3 \leq \tfrac{1}{0.9}y_3. \end{array}$$

We also have the additional integrality constraint:

$$y_2 + y_3 \leq 1.$$

Then, the complete mathematical model can be stated as follows:

$$\begin{array}{rl} \min & -11C + 7B_1 + B_2 + 1.2B_3 + 1.8A_2 \\ & +1.8A_3 + 3.5y_1 + y_2 + 1.5y_3 \\ \text{s.t.} & B_2 - \ln(1 + A_2) = 0 \\ & B_3 - 1.2\ln(1 + A_3) = 0 \\ & C - 0.9\,(B_1 + B_2 + B_3) = 0 \\ & C - 1y_1 \leq 0 \\ & B_2 - \dfrac{1}{0.9}y_2 \leq 0 \\ & B_3 - \dfrac{1}{0.9}y_3 \leq 0 \\ & y_2 + y_3 \leq 1 \\ & C, B_1, B_2, B_3, A_2, A_3 \geq 0 \\ & y_1, y_2, y_3 = 0, 1 \end{array} \tag{6.32}$$

Let us start with $y^1 = (1, 1, 0)$. The primal problem then becomes

$$\begin{array}{rl} \min & -11C + 7B_1 + B_2 + 1.2B_3 + 1.8A_2 \\ & +1.8A_3 + 4.5 \\ \text{s.t.} & B_2 - \ln(1 + A_2) = 0 \\ & B_3 - 1.2\ln(1 + A_3) = 0 \\ & C - 0.9\,(B_1 + B_2 + B_3) = 0 \\ & C - 1 \leq 0 \\ & B_2 - \dfrac{1}{0.9} \leq 0 \\ & B_3 \leq 0 \\ & C, B_1, B_2, B_3, A_2, A_3 \geq 0 \end{array}$$

and has as solution:

$$
\begin{aligned}
C &= 1, \\
B_1 &= 0, \\
B_2 &= (1/0.9) = 1.11111, \\
B_3 &= 0, \\
A_2 &= 2.037732, \\
A_3 &= 0, \\
Obj &= -1.72097.
\end{aligned}
$$

The multipliers for the two nonlinear equality constraints are

$$
\begin{aligned}
\lambda_1 &= 5.46792, \\
\lambda_2 &= 0,
\end{aligned}
$$

since both λ_1 and λ_2 are nonnegative, then the nonlinear equalities can be relaxed in the form:

$$
\begin{aligned}
B_2 - \ln(1 + A_2) &\leq 0, \\
B_3 - 1.2\ln(1 + A_3) &\leq 0.
\end{aligned}
$$

Note that they are both convex and hence condition C1 is satisfied for $T^1 \boldsymbol{h}(\boldsymbol{x}) \leq 0$.

To derive the linearizations around the solution x^1 of the primal problem, we only need to consider the two nonlinear relaxed equalities which become

$$
\begin{aligned}
B_2 - \left[\ln(1 + 2.03773) + \frac{1}{1 + 2.03773}(A_2 - 2.03773)\right] &\leq 0, \\
B_3 - 1.2\left[\ln(1 + 0) + \frac{1}{1 + 0}(A_3 - 0)\right] &\leq 0,
\end{aligned}
$$

and take the form:

$$
\begin{aligned}
B_2 - 0.329193 A_2 - 0.440303 &\leq 0, \\
B_3 - 1.2 A_3 &\leq 0.
\end{aligned}
$$

Then, the relaxed master problem is of the form:

$$
\begin{aligned}
\min \quad & 3.5y_1 + y_2 + 1.5y_3 + \mu \\
\text{s.t.} \quad & \mu \geq -11C + 7B_1 + B_2 + 1.2B_3 + 1.8A_2 + 1.8A_3 \\
& 0 \geq B_2 - 0.329193A_2 - 0.440303 \\
& 0 \geq B_3 - 1.2A_3 \\
& C - 0.9\,(B_1 + B_2 + B_3) = 0 \\
& C - 1y_1 \leq 0
\end{aligned}
$$

$$
\begin{aligned}
& B_2 - \frac{1}{0.9}y_2 \leq 0 \\
& B_3 - \frac{1}{0.9}y_3 \leq 0 \\
& y_2 + y_3 \leq 1 \\
& y_1 + y_2 - y_3 \leq 1 \\
& y_1, y_2, y_3 = 0, 1
\end{aligned}
$$

which has as solution:

$$
\begin{aligned}
y_1 &= 1, \\
y_2 &= 0, \\
y_3 &= 1, \\
C &= 1, \\
B_1 &= 0, \\
B_2 &= 0, \\
B_3 &= (1/0.9) = 1.1111111, \\
A_2 &= 0, \\
A_3 &= 0.925926, \\
\mu &= -8, \\
OBJ &= -3.
\end{aligned}
$$

Thus, after the first iteration we have

$$
\begin{aligned}
UBD &= -1.72097, \\
LBD &= -3,
\end{aligned}
$$

and $y^2 = (1, 0, 1)$.

Solving the primal problem with $y^2 = (1, 0, 1)$ we have the following formulation:

$$
\begin{aligned}
\min \quad & -11C + 7B_1 + B_2 + 1.2B_3 + 1.8A_2 \\
& +1.8A_3 + 5.0 \\
\text{s.t.} \quad & B_2 - \ln(1 + A_2) = 0 \\
& B_3 - 1.2\ln(1 + A_3) = 0 \\
& C - 0.9(B_1 + B_2 + B_3) = 0 \\
& C - 1 \leq 0 \\
& B_2 - 0 \leq 0 \\
& B_3 \leq 0 \\
& C, B_1, B_2, B_3, A_2, A_3 \geq 0
\end{aligned}
$$

which has as solution:

$$
\begin{aligned}
C &= 1, \\
B_1 &= 0, \\
B_2 &= 0, \\
B_3 &= (1/0.9) = 1.1111111, \\
A_2 &= 0, \\
A_3 &= 1.5242, \\
OBJ &= -1.9231 \text{ (New Upper Bound).}
\end{aligned}
$$

The multipliers for the two nonlinear equality constraints are

$$
\begin{aligned}
\lambda_1 &= 0, \\
\lambda_2 &= 3.7863.
\end{aligned}
$$

Since both λ_1 and λ_2 are nonnegative, then we can relax the nonlinear equalities into

$$
\begin{aligned}
B_2 - \ln(1 + A_2) &\leq 0, \\
B_3 - 1.2\ln(1 + A_3) &\leq 0,
\end{aligned}
$$

which are convex and hence satisfy condition C1.

The linearizations around the solution of the primal of the second iteration are

$$
\begin{aligned}
B_2 - A_2 &\leq 0, \\
B_3 - 0.475398A_3 - 0.386507 &\leq 0.
\end{aligned}
$$

Then, the relaxed master problem of the second iteration takes the form:

$$
\begin{aligned}
\min \quad & 3.5y_1 + y_2 + 1.5y_3 + \mu \\
\text{s.t.} \quad & \mu \geq -11C + 7B_1 + B_2 + 1.2B_3 + 1.8A_2 + 1.8A_3 \\
& 0 \geq B_2 - 0.329193A_2 - 0.44303 \\
& 0 \geq B_3 - 1.2A_3 \\
& 0 \geq B_2 - A_2 \\
& 0 \geq B_3 - 0.475398A_3 - 0.386507 \\
& C - 0.9\,(B_1 + B_2 + B_3) = 0 \\
& C - 1y_1 \leq 0 \\
& B_2 - \frac{1}{0.9}y_2 \leq 0 \\
& B_3 - \frac{1}{0.9}y_3 \leq 0 \\
& y_2 + y_3 \leq 1 \\
& y_1 + y_2 - y_3 \leq 1 \\
& y_1 + y_3 - y_2 \leq 1 \\
& y_1, y_2, y_3 = 0, 1 \\
& -3 \leq 3.5y_1 + y_2 + 1.5y_3 + \mu \leq -1.9231
\end{aligned}
$$

which has no feasible solution, and hence termination has been obtained with optimal solution the one of the primal problem of the second iteration

$$OBJ = -1.9231, \quad y = (1, 0, 1).$$