Chapter 9
Parametric Decomposition

A partly convex problem as formulated in Sect. 8.2 of the preceding chapter can also be viewed as a convex optimization problem depending upon a parameter $x \in \mathbb{R}^n$. The dimension n of the parameter space is referred to as the *nonconvexity rank* of the problem. Roughly speaking, this is the number of "nonconvex variables" in the problem. As we saw in Sect. 8.2, a partly convex problem of rank n can be efficiently solved by a BB decomposition algorithm with branching performed in $\mathbb{R}^n$. This chapter discusses decomposition methods for partly convex problems with small nonconvexity rank n. It turns out that these problems can be solved by streamlined decomposition methods based on parametric programming.

9.1   Functions Monotonic w.r.t. a Convex Cone

Let X be a convex set in $\mathbb{R}^n$ and K a closed convex cone in $\mathbb{R}^n$. A function $f : \mathbb{R}^n \to \mathbb{R}$ is said to be *monotonic* on X with respect to K (or *K-monotonic* on X for short) if

$$x, x' \in X, \ x' - x \in K \quad \Rightarrow \quad f(x') \geq f(x). \tag{9.1}$$

In other words, for every $x \in X$ and $u \in K$ the univariate function $\theta \mapsto f(x + \theta u)$ is nondecreasing in the interval $(0, \theta^*)$ where $\theta^* = \sup\{\theta \mid x + \theta u \in X\}$. When $X = \mathbb{R}^n$ we simply say that $f(x)$ is K-monotonic.

Let L be the lineality space of K, and Q an $r \times n$ matrix of rank r such that $L = \{x \mid Qx = 0\}$. From (9.1) it follows that

$$x, x' \in X, \ Qx = Qx' \quad \Rightarrow \quad f(x) = f(x'), \tag{9.2}$$

i.e., $f(x)$ is constant on every set $\{x \in X \mid Qx = \text{const}\}$. Denote $Q(X) = \{y \in \mathbb{R}^r \mid y = Qx, x \in X\}$.

Proposition 9.1 *Let $K \subset \operatorname{rec}X$. A function $f(x)$ is K-monotonic on X if and only if*

$$f(x) = F(Qx) \quad \forall x \in X, \tag{9.3}$$

where $F : Q(X) \to \mathbb{R}$ satisfies

$$y, y' \in Q(X), \ y' - y \in Q(K) \quad \Rightarrow \quad F(y') \geq F(y). \tag{9.4}$$

Proof If $f(x)$ is K-monotonic on X then we can define a function $F : Q(X) \to \mathbb{R}$ by setting, for every $y \in Q(X)$:

$$F(y) = f(x) \text{ for any } x \in X \text{ such that } y = Qx. \tag{9.5}$$

This definition is correct in view of (9.2). Obviously, $f(x) = F(Qx)$. For any $y, y' \in Q(X)$ such that $y' - y \in Q(K)$ we have $y = Qx, y' = Qx'$, with $Q(x' - x) = Qu$, $u \in K$, i.e., $Qx' = Q(x + u)$, $u \in K$. Here $x + u \in X$ because $u \in \operatorname{rec}X$, hence by (9.1), $f(x') = f(x + u) \geq f(x)$, i.e., $F(y') \geq F(y)$. Conversely, if $f(x) = F(Qx)$ and $F : Q(X) \to \mathbb{R}$ satisfies (9.4), then for any $x, x' \in X$ such that $x' - x \in K$ we have $y = Qx, y' = Qx'$, with $y' - y \in Q(K)$, hence $F(y') \geq F(y)$, i.e., $f(x') \geq f(x)$.   □

The above representation of K-monotonic functions explains their role in global optimization. Indeed, using this representation, any optimization problem in $\mathbb{R}^n$ of the form

$$\min\{f(x)| \quad x \in D\} \tag{9.6}$$

where $D \subset X$, and $f(x)$ is K-monotonic on X, can be rewritten as

$$\min\{F(y)| \quad y \in Q(D)\} \tag{9.7}$$

which is a problem in $\mathbb{R}^r$. Assume further, that $f(x)$ is quasiconcave on X, i.e., by (9.5), $F(y)$ is quasiconcave on $Q(X)$.

Proposition 9.2 *Let $K \subset \operatorname{rec}X$. A quasiconcave function $f : X \to \mathbb{R}$ on X is K-monotonic on X if and only if for every $x^0 \in X$ with $f(x^0) = \gamma$:*

$$x^0 + K \subset X(\gamma) := \{x \in X | f(x) \geq \gamma\}.$$

Proof If $f(x)$ is K-monotonic on X, then for any $x \in X(\gamma)$ and $u \in K$ we have $x + u \in X$ (because $K \subset \operatorname{rec}X$) and condition (9.1) implies that $f(x + u) \geq f(x) \geq \gamma$, hence $x + u \in X(\gamma)$, i.e., u is a recession direction for $X(\gamma)$. Conversely, if for any $\gamma \in f(X), K$ is contained in the recession cone of $X(\gamma)$, then for any $x', x \in X$ such that $x' - x \in K$ one has, for $\gamma = f(x) : x \in X(\gamma)$, hence $x' \in X(\gamma)$, i.e., $f(x') \geq f(x)$.
□

Proposition [9.1](#) provides the foundation for reducing the dimension of problem ([9.6](#)), whereas Proposition [9.2](#) suggests methods for restricting the global search domain. Indeed, since local information is generally insufficient to verify the global optimality of a solution, the search for a global optimal solution must be carried out, in principle, over the entire feasible set. If, however, the objective function $f(x)$ in ([9.6](#)) is K-monotonic, then, once a solution $x^0 \in D$ is known, one can ignore the whole set $D \cap (x^0 + K) \subset \{x \in D | f(x) \geq f(x^0)\}$, because no better feasible solution than x^0 can be found in this set. Such kind of information is often very helpful and may drastically simplify the problem by limiting the global search to a restricted region of the feasible domain. In the next sections, we shall discuss how efficient algorithms for K-monotonic problems can be developed based on these observations.

Below are some important classes of quasiconcave monotonic functions encountered in the applications.

Example 9.1 $f(x) = \varphi(x_1, \ldots, x_p) + d^T x$,

with $\varphi(y_1, \ldots, y_p)$ a continuous concave function of $y = (y_1, \ldots, y_p)$.

Here $F(y) = \varphi(y_1, \ldots, y_p) + y_{p+1}, Qx = (x_1, \ldots, x_p, d^T x)^T$. Monotonicity holds with respect to the cone $K = \{u | u_i = 0 (i = 1, \ldots, p), d^T u \geq 0\}$. It has long been recognized that concave minimization problems with objective functions of this form can be efficiently solved only by taking advantage of the presence of the linear part in the objective function (Rosen 1983; Tuy 1984).

Example 9.2 $f(x) = -\sum_{i=1}^{p} [\langle c^i, x \rangle]^2 + \langle c^{p+1}, x \rangle$.

Here $F(y) = -\sum_{i=1}^{p} y_i^2 + y_{p+1}, Qx = (\langle c^1, x \rangle, \ldots, \langle c^{p+1}, x \rangle)^T$. Monotonicity holds with respect to $K = \{u | \langle c^i, u \rangle = 0 (i = 1, \ldots, p), \langle c^{p+1}, u \rangle \geq 0\}$. For $p = 1$ such a function $f(x)$ cannot in general be written as a product of two affine functions and finding its minimum over a polytope is a problem known to be NP-hard (Pardalos and Vavasis 1991), however, quite practically solvable by a parametric algorithm (see Sect. [9.5](#)).

Example 9.3 $f(x) = \prod_{i=1}^{r} [\langle c^i, x \rangle + d_i]^{\alpha_i}$ with $\alpha_i > 0$, $i = 1, \ldots, r$.

This class includes functions which are products of affine functions. Since $\log f(x) = \sum_{i=1}^{r} \alpha_i \log[\langle c^i, x \rangle + d_i]$ for every x such that $\langle c^i, x \rangle + d_i > 0$, it is easily seen that $f(x)$ is quasiconcave on the set $X = \{x | \langle c^i, x \rangle + d_i \geq 0, i = 1, \ldots, r\}$. Furthermore, $f(x)$ is monotonic on X with respect to the cone $K = \{u | \langle c^i, u \rangle \geq 0, i = 1, \ldots, r\}$. Here $f(x)$ has the form ([9.3](#)) with $F(y) = \prod_{i=1}^{r} (y_i + d_i)^{\alpha_i}$. Problems of minimizing functions $f(x)$ of this form (with $\alpha_i = 1$, $i = 1, \ldots, r$) under linear constraints are termed linear multiplicative programs and will be discussed later in this chapter (Sect. [9.5](#)).

Example 9.4 $f(x) = -\sum_{i=1}^{r} \theta_i e^{\langle c^i, x \rangle}$ with $\theta_i > 0, i = 1, \ldots, r$.

Here $F(y) = -\sum_{i=1}^{r} \theta_i e^{y_i}$ is concave in y, so $f(x)$ is concave in x. Monotonicity holds with respect to the cone $K = \{x | \langle c^i, x \rangle \leq 0, i = 1, \ldots, r\}$. Functions of this type appear when dealing with geometric programs with negative coefficients ("signomial programs," cf Example 5.6).

Example 9.5 $f(x)$ quasiconcave and differentiable, with

$$\nabla f(x) = \sum_{i=1}^{r} p_i(x)c^i, \quad c^i \in \mathbb{R}^n, \; i = 1, \dots, r.$$

For every $x, x' \in X$ satisfying $\langle c^i, x' - x \rangle = 0, i = 1, \dots, r$, there is some $\theta \in (0, 1)$ such that $f(x') - f(x) = \langle \nabla f(x + \theta(x' - x)), x' - x \rangle = \sum_{i=1}^{r} p_i(x + \theta(x' - x))\langle c^i, x' - x \rangle = 0$, hence $f(x)$ is monotonic with respect to the space $L = \{x \mid \langle c^i, x \rangle = 0, \; i = 1, \dots, r\}$. If, in addition, $p_i(x) \geq 0 \;\; \forall x \; (i = 1, \dots, r)$ then monotonicity holds with respect to the cone $K = \{u \mid \langle c^i, u \rangle \geq 0, i = 1, \dots, r\}$.

Example 9.6 $f(x) = \min\{y^T Q x \mid y \in E\}$

where E is a convex subset of $\mathbb{R}^m$ and Q is an $m \times n$ matrix of rank r.

This function appears when transforming *rank r bilinear programs* into concave minimization problems (see Sect. 10.4). Here monotonicity holds with respect to the subspace $K = \{u \mid Qu = 0\}$. If $E \subset \mathbb{R}^m_+$ then $Qu \geq 0$ implies $y^T Q(x + u) \geq y^T Q x$, hence $f(x + u) \geq f(x)$, and so $f(x)$ is monotonic with respect to the cone $K = \{u \mid Qu \geq 0\}$.

Example 9.7 $f(x) = \sup\{d^T y \mid Ax + By \leq q\}$

where $A \in \mathbb{R}^{p \times n}$ with $\mathrm{rank} A = r$, $B \in \mathbb{R}^{p \times m}$, $y \in \mathbb{R}^m$ and $q \in \mathbb{R}^p$.

This function appears in bilevel linear programming, for instance, in the max-min problem (Falk 1973b):

$$\min_x \max_y \{c^T x + d^T y \mid Ax + By \leq q, x \geq 0, y \geq 0\}.$$

Obviously, $Au \leq 0$ implies $A(x + u) \leq Ax$, hence $\{y \mid Ax + By \leq q\} \subset \{y \mid A(x + u) + By \leq q\}$, hence $f(x + u) \geq f(x)$. Therefore, monotonicity holds with respect to the cone $K = \{u \mid Au \leq 0\}$.

Note that by Corollary 1.15, $\mathrm{rank} Q = n - \dim L = \dim K^\circ$. A K- monotonic function $f(x)$ on X, with $\mathrm{rank} Q = r$ is also said to possess the *rank r monotonicity property*.

9.2 Decomposition by Projection

Let D be a compact subset of a closed convex set X in $\mathbb{R}^n$ and let $f : \mathbb{R}^n \to \mathbb{R}$ be a quasiconcave function, monotonic on X with respect to a polyhedral convex cone K contained in the recession cone of X. Assume that the lineality space of K is $L = \{x \mid Qx = 0\}$, where Q is an $r \times n$ matrix of rank r. Consider the *quasiconcave monotonic problem*

$$(QCM) \qquad \min\{f(x) \mid \; x \in D\} \tag{9.8}$$

In this and the next two sections, we shall discuss decomposition methods for solving this problem, under the additional assumption that D is a polytope. The idea of decomposition is to transform the original problem into a sequence of subproblems involving each a reduced number of nonconvex variables. This can be done either by projection, or by dualization (polyhedral annexation), or by parametrization (for $r \leq 3$). We first present decomposition methods by projection.

As has been previously observed, by setting

$$F(y) = f(x) \text{ for any } x \in X \text{ satisfying } y = Qx, \tag{9.9}$$

we unambiguously define a quasiconcave function $F : Q(X) \to \mathbb{R}$ such that problem (QCM) is equivalent to

$$\min F(y) \quad \text{s.t.} \quad y = (y_1, \ldots, y_r) \in G \tag{9.10}$$

where $G = Q(D)$ is a polytope in $\mathbb{R}^r$. The algorithms presented in Chaps. 5 and 6 can of course be adapted to solve this quasiconcave minimization problem.

We first show how to compute $F(y)$. Since $\text{rank} Q = r$, by writing $Q = [Q_B, Q_N]$, $x = \begin{bmatrix} x_B \\ x_N \end{bmatrix}$, where Q_B is an $r \times r$ nonsingular matrix, the equation $Qx = y$ yields

$$x = \begin{bmatrix} Q_B^{-1} \\ 0 \end{bmatrix} y + \begin{bmatrix} -Q_B^{-1} Q_N x_N \\ x_N \end{bmatrix}$$

Setting $Z = \begin{bmatrix} Q_B^{-1} \\ 0 \end{bmatrix}$ and $u = \begin{bmatrix} -Q_B^{-1} Q_N x_N \\ x_N \end{bmatrix}$ we obtain

$$x = Zy + u \quad \text{with } Qu = -Q_N x_N + Q_N x_N = 0. \tag{9.11}$$

Since $Qx = Q(Zy)$ with $Zy = x - u \in X - L = X$, it follows that $f(x) = f(Zy)$. Thus, the objective function of (9.10) can be computed by the formula

$$F(y) = f(Zy). \tag{9.12}$$

We shall assume that $f(x)$ is continuous on some open set in $\mathbb{R}^n$ containing D, so $F(y)$ is continuous on some open set in $\mathbb{R}^r$ containing the constraint polytope $G = Q(D)$ of (9.10). The peculiar form of the latter set (image of a polytope D under the linear map Q) is a feature which requires special treatment, but does not cause any particular difficulty.

9.2.1 Outer Approximation

To solve (9.10) by outer approximation, the key point is: given a point $\bar{y} \in Q(\mathbb{R}^n) \subset \mathbb{R}^r$, determine whether $\bar{y} \in G = Q(D)$, and if $\bar{y} \notin G$ then construct a linear inequality (cut) $l(y) \leq 0$ to exclude $\bar{y}$ without excluding any point of G.

Observe that $\bar{y} \notin Q(D)$ if and only if the linear system $Qx = \bar{y}$ has no solution in D, so the existence of $l(y)$ is ensured by the separation theorem or any of its equivalent forms (such as the Farkas–Minkowski Lemma). However, since we are not so much interested in the existence as in the effective construction of $l(y)$, the best way is to use the duality theorem of linear programming (which is another form of the separation theorem). Specifically, assuming $D = \{x|\ Ax \leq b,\ x \geq 0\}$, with $A \in \mathbb{R}^{m \times n},\ b \in \mathbb{R}^m$, consider the dual pair of linear programs

$$\min\{\langle h, x\rangle|\ Qx = \bar{y}, Ax \leq b, x \geq 0\} \tag{9.13}$$

$$\max\{\langle \bar{y}, v\rangle - \langle b, w\rangle|\ Q^T v - A^T w \leq h,\ w \geq 0\}. \tag{9.14}$$

where h is any vector chosen so that (9.14) is feasible (for example, $h = 0$) and can be solved with least effort.

Proposition 9.3 *Solving (9.14) either yields a finite optimal solution or an extreme direction $(\bar{v}, \bar{w})$ of the cone $Q^T v - A^T w \leq 0,\ w \geq 0$, such that $\langle \bar{y}, \bar{v}\rangle - \langle b, \bar{w}\rangle > 0$. In the former case, $\bar{y} \in G = Q(D)$; in the latter case, the affine function*

$$l(y) = \langle y, \bar{v}\rangle - \langle b, \bar{w}\rangle \tag{9.15}$$

satisfies

$$l(\bar{y}) > 0, \quad l(y) \leq 0 \quad \forall y \in G. \tag{9.16}$$

Proof Immediate from the duality theorem of linear programming.   □

On the basis of this proposition a finite OA procedure for solving (9.10) can be carried out in the standard way, starting from an initial r-simplex $P_1 \supset G = Q(D)$ with a known or readily computable vertex set. In most cases, one can take this initial simplex to be

$$P_1 = \left\{y|\ y_i \geq \alpha_i,\ (i = 1, \ldots, r),\ \sum_{i=1}^{r}(y_i - \alpha_i) \leq \beta\right\} \tag{9.17}$$

where

$$\alpha_i = \min\{y_i|\ y = Qx, x \in D\}, i = 1, \ldots, r, \tag{9.18}$$

$$\beta = \max\left\{\sum_{i=1}^{r}(y_i - \alpha_i)|\ y = Qx, x \in D\right\}. \tag{9.19}$$

Remark 9.1 A typical case of interest is the *concave program with few nonconvex variables* (Example 7.1):

$$\min\{\varphi(y) + dz \mid Ay + Bz \le b, \ y \ge 0, \ z \ge 0\} \tag{9.20}$$

where the function $\varphi : \mathbb{R}^p \to \mathbb{R}$ is concave and $(y, z) \in \mathbb{R}^p \times \mathbb{R}^q$. By rewriting this problem in the form (9.10):

$$\min\{\varphi(y) + t \mid dz \le t, \ Ay + Bz \le b, \ y \ge 0, \ z \ge 0\}$$

we see that for checking the feasibility of a vector $(\bar{y}, \bar{t}) \in \mathbb{R}_+^p \times \mathbb{R}$ the best choice of h in (9.13) is $h = d$. This leads to the pair of dual linear programs

$$\min\{dz \mid Bz \le b - A\bar{y}, \ z \ge 0\} \tag{9.21}$$

$$\max\{\langle A\bar{y} - b, v \rangle \mid -B^T v \le d, \ v \ge 0\}. \tag{9.22}$$

We leave it to the reader to verify that:

1. If $\bar{v}$ is an optimal solution to (9.22) and $\langle A\bar{y} - b, \bar{v} \rangle \le \bar{t}$ then $(\bar{y}, \bar{t})$ is feasible;
2. If $\bar{v}$ is an optimal solution to (9.22) and $\langle A\bar{y} - b, \bar{v} \rangle > \bar{t}$ then the cut $\langle Ay - b, \bar{v} \rangle \le t$ excludes $(\bar{y}, \bar{t})$ without excluding any feasible point;
3. If (9.22) is unbounded, i.e., $\langle A\bar{y} - b, \bar{v} \rangle > 0$ for some extreme direction $\bar{v}$ of the feasible set of (9.22), then the cut $\langle Ay - b, \bar{v} \rangle \le 0$ excludes $(\bar{y}, \bar{t})$ without excluding any feasible point.

    Sometimes the matrix B has a favorable structure such that for each fixed $y \ge 0$ the problem (9.20) is solvable by efficient specialized algorithms (see Sect. 9.8). Since the above decomposition preserves this structure in the auxiliary problems, the latter problems can be solved by these algorithms.

9.2.2 Branch and Bound

A general rule in the branch and bound approach to nonconvex optimization problems is to branch upon the nonconvex variables, except for rare cases where there are obvious reasons to do otherwise. Following this rule the space to be partitioned for solving (QCM) is $Q(\mathbb{R}^n) = \mathbb{R}^r$. Given a partition set $M \subset \mathbb{R}^r$ a basic operation is to compute a lower bound for $F(y)$ over the feasible points in M. To this end, select an appropriate affine minorant $l_M(y)$ of $F(y)$ over M and solve the linear program

$$\min\{l_M(y) \mid Qx = y, x \in D, y \in M\}.$$

If $\beta(M)$ is the optimal value of this linear program then obviously

$$\beta(M) \leq \min\{F(y)| \ y \in Q(D) \cap M\}.$$

According to condition (6.4) in Sect. 6.2 this can serve as a valid lower bound, provided, in addition, $\beta(M) < +\infty$ whenever $Q(D) \cap M \neq \emptyset -$ which is the case when M is bounded as, e.g., when using simplicial or rectangular partitioning.

Simplicial Algorithm

If $M = [u^1, \ldots, u^{r+1}]$ is an r-simplex in $Q(\mathbb{R}^n)$ then $l_M(x)$ can be taken to be the affine function that agrees with $F(y)$ at $u^1, \ldots, u^{r+1}$. Since $l_M(y) = \sum_{i=1}^{r+1} t_i F(u^i)$ for $y = \sum_{i=1}^{r+1} t_i u^i$ with $t = (t_1, \ldots, t_{r+1}) \geq 0, \sum_{i=1}^{r+1} t_i = 1$, the above linear program can also be written as

$$\min \left\{ \sum_{i=1}^{r+1} t_i F(u^i)| \ x \in D, Qx = \sum_{i=1}^{r+1} t_i u^i, \sum_{i=1}^{r+1} t_i = 1, t \geq 0 \right\}.$$

Simplicial subdivision is advisable when $f(x) = f_0(x) + dx$, where $f_0(x)$ is concave and monotonic with respect to the cone $Qx \geq 0$. Although $f(x)$ is then actually monotonic with respect to the cone $Qx \geq 0, dx \geq 0$, by writing (QCM) in the form

$$\min\{F_0(y) + dx| \ Qx = y, x \in D\}$$

one can see that it can be solved by a simplicial algorithm operating in $Q(\mathbb{R}^n)$ rather than $Q(\mathbb{R}^n) \times \mathbb{R}$. As initial simplex M_1 one can take the simplex (9.17).

Rectangular Algorithm

When $F(y)$ is concave separable, i.e.,

$$F(y) = \sum_{i=1}^{r} F_i(y_i),$$

where each $F_i(\cdot), i = 1, \ldots, r$ is a concave univariate function, this structure can be best exploited by rectangular subdivision. In that case, for any rectangle $M = [p, q] = \{y| \ p \leq y \leq q\}$, $l_M(y)$ can be taken to be the uniquely defined affine function which agrees with $F(y)$ at the corners of M. This function is $l_M(y) = \sum_{i=1}^{r} l_{i,M}(y_i)$ where each $l_{i,M}(t)$ is the affine univariate function that agrees with $F_i(t)$ at the endpoints p_i, q_i of the segment $[p_i, q_i]$. The initial rectangle can be taken to be $M_1 = [\alpha_1, \beta_1] \times \ldots \times [\alpha_r, \beta_r]$, where α_i is defined as in (9.18) and

$$\beta_i = \sup\{y_i| \ y = Qx, x \in D\}.$$

Conical Algorithm

In the general case, since the global minimum of $F(y)$ is achieved on the boundary of $Q(D)$, conical subdivision is most appropriate. Nevertheless, instead of bounding, we use a selection operation, so that the algorithm is a branch and select rather than branch and bound procedure.

The key point in this algorithm is to decide, at each iteration, which cones in the current partition are 'non promising' and which one is the 'most promising' (selection rule).

Let $\bar{y}$ be the current best solution (CBS) at a given iteration and $\gamma := F(\bar{y})$. Assume that all the cones are vertexed at 0 and $F(0) > \gamma$, so that

$$0 \in \text{int}\Omega_\gamma,$$

where $\Omega_\gamma = \{y \in \Omega \mid F(y) \geq \gamma\}$. Let $l_M(y) = 1$ be the equation of the hyperplane in $\mathbb{R}^r$ passing through the intersections of the edges of M with the surface $F(y) = \gamma$. Denote by $\omega(M)$ a basic optimal solution and by $\mu(M)$ the optimal value of the linear program

$$LP(M, \gamma) \qquad \max\{l_M(y) \mid y \in Q(D) \cap M\}.$$

Note that if $u^1, \ldots, u^r$ are the directions of the edges of M and $U = (u^1, \ldots, u^r)$ is the matrix of columns $u^1, \ldots, u^r$ then $M = \{y \mid y = Ut, t = (t_1, \ldots, t_r)^T \geq 0\}$, so LP(M,$\gamma$) can also be written as

$$\max\{l_M(Ut) \mid Ax \leq b, \; Qx = Ut, \; x \geq 0, \; t \geq 0\}.$$

The deletion and selection rule are based on the following:

Proposition 9.4 *If $\mu(M) \leq 1$ then M cannot contain any feasible point better than CBS. If $\mu(M) > 1$ and $\omega(M)$ lies on an edge of M then $\omega(M)$ is a better feasible solution than CBS.*

Proof Clearly $D \cap M$ is contained in the simplex $Z := M \cap \{y \mid l_M(y) \leq \mu(M)\}$. If $\mu(M) \leq 1$, then the vertices of Z other than 0 lie inside Ω_γ, hence the values of $F(y)$ at these vertices are at least equal to γ. But by quasiconcavity of $F(y)$, its minimum over Z must be achieved at some vertex of Z, hence must be at least equal to γ. This means that $F(y) \geq \gamma$ for all $y \in Z$, and hence, for all $y \in D \cap M$. To prove the second assertion, observe that if $\omega(M)$ lies on an edge of M and $\mu(M) > 1$, then $\omega(M) \notin \Omega_\gamma$, i.e., $F(\omega(M)) < \gamma$; furthermore, $\omega(M)$ is obviously feasible. □

The algorithm is initialized from a partition of $\mathbb{R}^r$ into $r + 1$ cones vertexed at 0 (take, e.g., the partition induced by the r-simplex $[e^0, e^1, \ldots, e^r]$ where e^i, for $i = 1, \ldots, r$, is the i-th unit vector of $\mathbb{R}^r$ and $e^0 = -\frac{1}{r}(e^1 + \ldots + e^r)$). On the basis of Proposition 9.4, at iteration k a cone M will be fathomed if $\mu(M) \leq 1$, while the most promising cone M_k (the candidate for further subdivision) is the unfathomed cone that has largest $\mu(M)$. If $\omega^k = \omega(M_k)$ lies on any edge of M^k it yields, by Proposition 9.4, a better feasible solution than CBS and the algorithm can go to the next iteration with this improved CBS. Otherwise, M_k can be subdivided via the ray through ω^k.

As proved in Chap. 7 (Theorem 7.3), the above described conical algorithm converges, in the sense that any accumulation point of the sequence of CBS that

it generates is a global optimal solution. Note that no bound is needed in this branch and select algorithm. Although a bound for $F(y)$ over the feasible points in M can be easily derived from the constructions used in Proposition 9.4, it would require the computation of the values of $F(y)$ at r additional points (the intersections of the edges of M with the hyperplane $l_M(y) = \mu(M)$), which may entail a nonnegligible extra cost.

9.3   Decomposition by Polyhedral Annexation

The decomposition method by outer approximation is conceptually simple. However, a drawback of this method is its inability to be restarted. Furthermore, the construction of the initial polytope involves solving $r + 1$ linear programs. An alternative decomposition method which allows restarts and usually performs better is by dualization through polyhedral annexation .

Let $\bar{x}$ be a feasible solution and $C := \{x \in X \mid f(x) \geq f(\bar{x})\}$. By quasiconcavity and continuity of $f(x)$, C is a closed convex set. By translating if necessary, assume that 0 is a feasible point such that $f(0) > f(\bar{x})$, so $0 \in D \cap \text{int} C$. The PA method (Chap. 8, Sect. 8.1) uses as a subroutine Procedure DC* for solving the following subproblem:

Find a point $x \in D \setminus C$ (i.e., a better feasible solution than $\bar{x}$) or else prove that $D \subset C$ (i.e., $\bar{x}$ is a global optimal solution).

The essential idea of Procedure DC* is to build up, starting with a polyhedron $P_1 \supset C^\circ$ (Lemma 8.1), a nested sequence of polyhedrons $P_1 \supset P_2 \supset \ldots \supset C^\circ$ that yields eventually either a polyhedron $P_k \subset D^\circ$ (in that case $C^\circ \subset D^\circ$, hence $D \subset C$) or a point $x^k \in D \setminus C$. In order to exploit the monotonicity structure, we now try to choose the initial polyhedron P_1 so that $P_1 \subset L^\perp$ (orthogonal complement of L). If this can be achieved, then the whole procedure will operate in $L^\perp$, which allows much computational saving since $\dim L^\perp = r < n$.

Let $c^1, \ldots, c^r$ be the rows of Q (so that $L^\perp$ is the subspace generated by these vectors) and let $c^0 = -\sum_{i=1}^{r} c^i$. Let $\pi : L^\perp \to \mathbb{R}^r$ be the linear map such that $y = \sum_{i=1}^{r} t_i c^i \Leftrightarrow \pi(y) = t$. For each $i = 0, 1, \ldots, r$ since $0 \in \text{int} C$, the halfline from 0 through c^i intersects C. If this halfline meets ∂C, we define α_i to be the positive number such that $z^i = \alpha_i c^i \in \partial C$; otherwise, we set $\alpha_i = +\infty$. Let $I = \{i \mid \alpha_i < +\infty\}$, $S_1 = \text{conv}\{0, z^i, i \in I\} + \text{cone}\{c^i, i \notin I\}$.

Lemma 9.1

(i) *The polar P_1 of $S_1 + L$ is an r-simplex in $L^\perp$ defined by*

$$\pi(P_1) = \left\{ t \in \mathbb{R}^r \mid \sum_{i=1}^{r} t_i \langle c^i, c^j \rangle \leq \frac{1}{\alpha_j}, \ j = 0, 1, \ldots, r \right\}, \qquad (9.23)$$

(as usual $\frac{1}{+\infty} = 0$).

(ii) $C^\circ \subset P_1$ *and if V is the vertex set of any polytope P such that $C^\circ \subset P \subset P_1$,*
then

$$\max\{\langle v, x \rangle | \quad x \in D\} \leq 1 \quad \forall v \in V \quad \Rightarrow \quad C^\circ \subset D^\circ. \tag{9.24}$$

Proof Clearly $S_1 \subset L^\perp$ and $0 \in \mathrm{ri}S_1$, so $0 \in \mathrm{int}(S_1 + L)$, hence P_1 is compact.
Since $(S_1 + L)^\circ = S_1^\circ \cap L^\perp$, any $y \in P_1$ belongs to $L^\perp$, i.e., is of the form
$y = \sum_{i=1}^r t_i c^i$ for some $t = (t_1, \dots, t_r)$. But by Proposition 1.28, $y \in S_1^\circ$ is equivalent
to $\langle c^j, y \rangle \leq \frac{1}{\alpha_j} \forall j = 0, 1, \dots, r$. Hence P_1 is the polyhedron defined by (9.23). We
have $0 < \frac{1}{\alpha_j} \forall j = 0, 1, \dots, r$, so $\dim P_1 = r$, (Corollary 1.16) and since P_1 is
compact it must be an r-simplex. To prove (ii) observe that $S_1 + L \subset C \cap L^\perp + L = C$,
so $C^\circ \subset (S_1 + L)^\circ = P_1$. If $\max\{\langle v, x \rangle | x \in D\} \leq 1 \quad \forall v \in V$, then
$\max\{\langle y, x \rangle | x \in D\} \leq 1 \forall y \in P$, i.e., $P \subset D^\circ$, hence $C^\circ \subset D^\circ$ because $C^\circ \subset P$. $\quad\square$

Thus Procedure DC* can be applied, starting from P_1 as an initial polytope. By
incorporating Procedure DC* (initialized from P_1) into the PA Algorithm for (BCP)
(Sect. 7.1) we obtain the following decomposition algorithm for (QCM):

PA Algorithm for (QCM)
Initialization. Let $\bar{x}$ be a feasible solution (the best available), $C = \{x \in X | f(x) \geq f(\bar{x})\}$. Choose a feasible point x^0 of D such that $f(x^0) > f(\bar{x})$ and set $D \leftarrow D - x^0$, $C \leftarrow C - x^0$. Let $\tilde{P}_1 = \pi(P_1)$ be the simplex in $\mathbb{R}^r$ defined by (9.23) (or $\tilde{P}_1 = \pi(P_1)$ with P_1 being any polytope in $L^\perp$ such that $C^\circ \subset P_1$). Let V_1 be the vertex set of $\tilde{P}_1$. Set $k = 1$.

Step 1.   For every new $t = (t_1, \dots, t_r) \in V_k$ solve the linear program

$$\max \left\{ \sum_{i=1}^r t_i \langle c^i, x \rangle | \ x \in D \right\} \tag{9.25}$$

to obtain its optimal value $\mu(t)$ and a basic optimal solution $x(t)$.
Step 2.   Let $t^k \in \mathrm{argmax}\{\mu(t) | \ t \in V_k\}$. If $\mu(t^k) \leq 1$ then terminate: $\bar{x}$ is an
optimal solution of (QCM).
Step 3.   If $\mu(t^k) > 1$ and $x^k = x(t^k) \notin C$, then update the current best feasible
solution and the set C by resetting $\bar{x} = x^k$.
Step 4.   Compute

$$\theta_k = \sup\{\theta | \ f(\theta x^k) \geq f(\bar{x})\}$$

and define

$$\tilde{P}_{k+1} = \tilde{P}_k \cap \left\{ t | \ \sum_{i=1}^r t_i \langle x^k, c^i \rangle \leq \frac{1}{\theta_k} \right\}.$$

Step 5.   From V_k derive the vertex set V_{k+1} of $\tilde{P}_{k+1}$. Set $k \leftarrow k + 1$ and go back to
Step 1.

Remark 9.2 Reported computational experiments (Tuy and Tam 1995) seem to indicate that the PA method is quite practical for problems with fairly large n provided r is small (typically $r \leq 6$). For $r \leq 3$, however, specialized parametric methods can be developed which are more efficient. These will be discussed in the next section.

Case When $K = \{u|\ Qu \geq 0\}$

A case when the initial simplex P_1 can be constructed so that its vertex set is readily available when $K = \{u|\ Qu \geq 0\}$, with an $r \times n$ matrix Q of rank r. Let $c^1, \ldots, c^r$ be the rows of Q. Using formula (9.11) one can compute a point w satisfying $Qw = -e$, where $e = (1, \ldots, 1)^T$, and a value $\alpha > 0$ such that $\hat{w} = \alpha w \in C$ (for the efficiency of the algorithm, α should be taken to be as large as possible).

Lemma 9.2 *The polar P_1 of the set $S_1 = \hat{w} + K$ is an r- simplex with vertex set $\{0, -c^1/\alpha, \ldots, -c^r/\alpha\}$ and satisfying condition (ii) in Lemma 9.1.*

Proof Since $K \subset \operatorname{rec} C$ and $\hat{w} \in C$ it follows that $S_1 = \hat{w} + K \subset C$, and hence $P_1 \supset C^\circ$. Furthermore, clearly $S_1 = \{x|\ \langle c^i, x \rangle \geq -\alpha, \ i = 1, \ldots, r\}$ because $x \in S_1$ if and only if $x - \hat{w} \in K$, i.e., $\langle c^i, x - \hat{w} \rangle \geq 0 \ i = 1, \ldots, r$, i.e., $\langle c^i, x \rangle \geq \langle c^i, \hat{w} \rangle = -\alpha, \ i = 1, \ldots, r$. From Proposition 1.28 we then deduce that $S_1^\circ = \operatorname{conv}\{0, -c^1/\alpha, \ldots, -c^r/\alpha\}$. The rest is immediate.   □

Also note that in this case $K^\circ = \operatorname{cone}\{-c^1, \ldots, -c^r\}$ and so $\pi(K^\circ) = \mathbb{R}^r_-$ while $\pi(P_1) = \{t \in \mathbb{R}^r_-|\ \sum_{i=1}^r t_i \geq -1/\alpha\}$.

9.4   The Parametric Programming Approach

In the previous sections, it was shown how a quasiconcave monotonic problem can be transformed into an equivalent quasiconcave problem of reduced dimension. In this section an alternative approach will be presented in which a quasiconcave monotonic problem is solved through the analysis of an associated linear program depending on a parameter (of the same dimension as the reduced problem in the former approach).

Consider the general problem formulated in Sect. 8.2:

$$(QCM) \qquad \min\{f(x)|\quad x \in D\} \tag{9.26}$$

where D is a compact set in $\mathbb{R}^n$, not necessarily even convex, and $f(x)$ is a quasiconcave monotonic function on a closed convex set $X \supset D$ with respect to a polyhedral cone K with lineality space $L = \{x|\ Qx = 0\}$. To this problem we associate the following program where v is an arbitrary vector in K° :

$$LP(v) \qquad \min\{\langle v, x \rangle \mid \quad x \in D\}. \tag{9.27}$$

Denote by K_b° any compact set such that $\text{cone}(K_b^\circ) = K^\circ$.

Theorem 9.1 *For every $v \in -K^\circ$, let x^v be an arbitrary optimal solution of $LP(v)$. Then*

$$\min\{f(x) \mid x \in D\} = \min\{f(x^v) \mid v \in -K_b^\circ\}. \tag{9.28}$$

Proof Let $\gamma = \min\{f(x^v) \mid v \in -K_b^0\}$. Clearly $\gamma = \min\{f(x^v) \mid v \in -K^\circ \setminus \{0\}\}$. Denote by E the convex hull of the closure of the set of all $x^v, v \in -K^\circ \setminus \{0\}$. Since D is compact, so is E and it is easy to see that the convex set $G = E + K$ is closed. Indeed, for any sequence $x^v = u^v + v^v \to x^0$ with $u^v \in E, v^v \in K$, one has, by passing to subsequences if necessary, $u^v \to u^0 \in E$, hence $v^v = x^v - u^v \to x^0 - u^0 \in K$, which implies that $x^0 = u^0 + (x^0 - u^0) \in E + K = G$. Thus, G is a closed convex set. By K-monotonicity we have $f(y + u) \geq f(y)$ for all $y \in E, u \in K$, hence $f(x) \geq \gamma$ for all $x \in E + K = G$. Suppose now that $\gamma > \min\{f(x) \mid x \in D\}$, so that for some $z \in D$ we have $f(z) < \gamma$, i.e., $z \in D \setminus G$. By Proposition 1.15 on the normals to a convex set there exists $x^0 \in \partial G$ such that $v := x^0 - z$ satisfies

$$\langle v, x - x^0 \rangle \geq 0 \quad \forall x \in G, \quad \langle v, z - x^0 \rangle < 0. \tag{9.29}$$

For every $u \in K$, since $x^0 + u \in G + K = G$, we have $\langle v, u \rangle \geq 0$, hence $v \in -K^\circ \setminus \{0\}$. But, since $z \in D$, it follows from the definition of x^v that $\langle v, x^v \rangle \leq \langle v, z \rangle < \langle v, x^0 \rangle$ by (9.29), hence $\langle v, x^v - x^0 \rangle < 0$, conflicting with the fact $x^v \in E \subset G$ and (9.29). $\qquad \square$

The above theorem reduces problem (QCM) to minimizing the function $\varphi(v) := f(x^v)$ over K_b°. Though $K^\circ \subset L^\perp$, and $\dim L^\perp = r$ may be small, this is still a *difficult problem* even when D is convex, because $\varphi(v)$ is generally highly nonconvex. The point, however, is that in many important cases the problem (9.28) is more amenable to computational analysis than the original problem. For instance, as we saw in the previous section, when D is a polytope the PA Method solves (QCM) through solving a suitable sequence of linear programs (9.25) which can be recognized to be identical to (9.27) with $v = -\sum_{i=1}^{r} t_i c^i$.

An important consequence of Theorem 9.1 can be drawn when:

$$\left| \begin{array}{l} D \text{ is a polyhedron;} \\ K = \{x \mid \langle c^i, x \rangle \geq 0, i = 1, \ldots, p; \ \langle c^i, x \rangle = 0, i = p + 1, \ldots, r\} \end{array} \right. \tag{9.30}$$

$(c^1, \ldots, c^r$ being linearly independent vectors of $\mathbb{R}^n$). Let

$$\Lambda = \{\lambda \in \mathbb{R}^p_+ | \ \lambda_p = 1\},$$

$$W = \{\alpha = (\alpha_{p+1}, \ldots, \alpha_r)| \ \underline{\alpha}_i \le \alpha_i \le \overline{\alpha}_i, \ i = p+1, \ldots, r\},$$

$$\underline{\alpha}_i = \inf_{x \in D} \langle c^i, x \rangle; \quad \overline{\alpha}_i = \sup_{x \in D} \langle c^i, x \rangle \ \ i = p+1, \ldots, r.$$

Corollary 9.1 *For each $(\lambda, \alpha) \in \Lambda \times W$ let $x^{\lambda, \alpha}$ be an arbitrary basic optimal solution of the linear program*

$$LP(\lambda, \alpha) \quad \left| \begin{array}{l} \min \sum_{i=1}^{p-1} \lambda_i \langle c^i, x \rangle + \langle c^p, x \rangle \\ \text{s.t. } x \in D, \ \langle c^i, x \rangle = \alpha_i, \ i = p+1, \ldots, r. \end{array} \right.$$

Then

$$\min\{f(x)| \ x \in D\} = \min\{f(x^{\lambda, \alpha})| \ \lambda \in \Lambda, \ \alpha \in W\}. \tag{9.31}$$

Proof Clearly

$$\min\{f(x)| \ x \in D\} = \min_{\alpha} \inf\{f(x)| \ x \in D \cap H_\alpha\},$$

where $H_\alpha = \{x| \ \langle c^i, x \rangle = \alpha_i, \ i = p+1 \ldots, r\}$. Let $\tilde{K} = \{x| \ \langle c^i, x \rangle \ge 0, \ i = 1, \ldots, p\}$. Since $f(x)$ is quasiconcave $\tilde{K}$-monotonic on $X \cap H_\alpha$, we have by Theorem 9.1 :

$$\min\{f(x)| \ x \in D \cap H_\alpha\} = \min\{f(x^{v, \alpha})| \ v \in -\tilde{K}_b^\circ\}, \tag{9.32}$$

where $x^{v, \alpha}$ is any basic optimal solution of the linear program

$$\min\{\langle v, x \rangle| \ x \in D \cap H_\alpha\}. \tag{9.33}$$

Noting that $\tilde{K}^\circ = -\text{cone}\{c^1, \ldots, c^p\}$ we can take $\tilde{K}_b^\circ = -[c^1, \ldots, c^p]$. Also we can assume $v \in -\text{ri}\tilde{K}_b^\circ$ since for any $v \in -\tilde{K}_b^\circ$ there exists $v' \in -\text{ri}\tilde{K}_b^\circ$ such that $x^{v', \alpha}$ is a basic optimal solution of (9.33). So $v = \sum_{i=1}^p t_i c^i$ with $t_p > 0$ and the desired result follows by setting $\lambda_i = \frac{t_i}{t_p}$.   □

As mentioned above, by Theorem 9.1 one can solve (QCM) by finding first the vector $\overline{v} \in -K_b^\circ$ that minimizes $f(x^v)$ over all $v \in K_b^\circ$. In certain methods, as, for instance, in the PA method when D is polyhedral, an optimal solution $\overline{x} = x^{\overline{v}}$ is immediately known once $\overline{v}$ has been found. In other cases, however, $\overline{x}$ must be sought by solving the associated program $\min\{\langle \overline{v}, x \rangle| \ x \in D\}$. The question then arises as to whether any optimal solution of the latter program will solve (QCM).

Theorem 9.2 *If, in addition to the already stated assumptions, $D \subset \text{int}X$, while the function $f(x)$ is continuous and strictly quasiconcave on X, then for any optimal solution $\bar{x}$ of (QCM) there exists a $\bar{v} \in K^\circ \setminus \{0\}$ such that $\bar{x}$ is an optimal solution of $LP(\bar{v})$ and any optimal solution to $LP(\bar{v})$ will solve (QCM).*

(A function $f(x)$ is said to be strictly quasiconcave on X if for any two $x, x' \in X, f(x') < f(x)$ always implies $f(x') < f(\theta x + (1-\theta)x') \ \forall \theta \in (0,1)$)

Proof It suffices to consider the case when $f(x)$ is not constant on D. Let $\bar{x}$ be an optimal solution of (QCM) and $X(\bar{x}) = \{x \in X \mid f(x) \geq f(\bar{x})\}$. By quasiconcavity and continuity of $f(x)$, this set $X(\bar{x})$ is closed and convex. By optimality of $\bar{x}$, $D \subset X(\bar{x})$ and again by continuity of $f(x)$, $\{x \in X \mid f(x) > f(\bar{x})\} \subset \text{int}X(\bar{x})$. Furthermore, since $f(x)$ is not constant on D one can find $x' \in D$ such that $f(x') > f(\bar{x})$, so if $\bar{x} \in \text{int}X(\bar{x})$ then for $\theta > 0$ small enough, $x^\theta = \bar{x} - \theta(x' - \bar{x}) \in X(\bar{x})$, $f(x^\theta) < f(x')$, hence, by strict quasiconcavity of $f(x)$, $f(x^\theta) < f(\bar{x})$, conflicting with the definition of $X(\bar{x})$. Therefore, $\bar{x}$ lies on the boundary of $X(\bar{x})$. By Theorem 1.5 on supporting hyperplanes there exists then a vector $\bar{v} \neq 0$ such that

$$\langle \bar{v}, x - \bar{x} \rangle \geq 0 \quad \forall x \in X(\bar{x}). \tag{9.34}$$

For any $u \in K$ we have $\bar{x} + u \in D + K \subset X(\bar{x})$, so $\langle \bar{v}, -u \rangle \geq 0$, and consequently, $\bar{v} \in -K^\circ \setminus \{0\}$. If $\tilde{x}$ is any optimal solution of $LP(\bar{v})$, then $\langle \bar{v}, \tilde{x} - \bar{x} \rangle = 0$, and in view of (9.34), $\tilde{x}$ cannot be an interior point of $X(\bar{x})$ for then $\langle \bar{v}, x - \bar{x} \rangle = 0 \ \forall x \in \mathbb{R}^n$, conflicting with $\bar{v} \neq 0$. So $\tilde{x}$ is a boundary point of $X(\bar{x})$, and hence $f(\tilde{x}) = f(\bar{x})$, because $\{x \in X \mid f(x) > f(\bar{x})\} \subset \text{int}X(\bar{x})$. Thus, any optimal solution $\tilde{x}$ to $LP(\bar{v})$ is optimal to (QCM).   □

Remark 9.3 If $f(x)$ is differentiable and pseudoconcave then it is strictly quasiconcave (see, e.g., Mangasarian 1969). In that case, since $X(\bar{x}) = \{x \in X \mid f(x) \geq f(\bar{x})\}$ one must have $\langle \nabla f(\bar{x}), x - \bar{x} \rangle \geq 0 \ \forall x \in X(\bar{x})$ and since one can assume that there is $x \in X$ satisfying $f(x) > f(\bar{x})$ the strict quasiconcavity of $f(x)$ implies that $\nabla f(\bar{x}) \neq 0$ (Mangasarian 1969). So $\nabla f(\bar{x})$ satisfies (9.34) and one can take $\bar{v} = \nabla f(\bar{x})$. This result was established in Sniedovich (1986) as the foundation of the so-called C-programming.

Corollary 9.2 *Let $F : \mathbb{R}^q \to \mathbb{R}$ be a continuous, strictly quasiconcave and K-monotonic function on a closed convex set $Y \subset \mathbb{R}^q$, where $K \subset \mathbb{R}^q$ is a polyhedral convex cone contained in the recession cone of Y. Let $g : D \to \mathbb{R}^q$ be a map defined on a set $D \subset \mathbb{R}^n$ such that $g(D) \subset \text{int}Y$. If $F(g(x))$ attains a minimum on D then there exists $t \in -K^\circ$ such that any optimal solution of*

$$\min\{\langle t, g(x) \rangle \mid \ x \in D\} \tag{9.35}$$

will solve

$$\min\{F(g(x)) \mid \ x \in D\}. \tag{9.36}$$

Proof If $\bar{x}$ is a minimizer of $F(g(x))$ over D, then $\bar{y} = g(\bar{x})$ is a minimizer of $F(y)$ over $G = g(D)$, and the result follows from Theorem 9.2.   □

In particular, when $K \supset \mathbb{R}^q_+$ and D is a convex set while $g_i(x), i = 1, \ldots, q$ are convex functions, problem (9.35) is a parametric convex program ($t \geq 0$ because $-K^\circ \subset \mathbb{R}^q_+$).

Example 9.8 Consider the problem (Tanaka et al. 1991):

$$\min_x \left\{ \left(\sum_{i=1}^n f_{1,i}(x_i) \right) \cdot \left(\sum_{i=1}^n f_{2,i}(x_i) \right) \mid x_i \in X_i, \ i = 1, \ldots, n \right\} \tag{9.37}$$

where $X_i \subset \mathbb{R}_{++}$ and $f_{1,i}, f_{2,i}$ are positive-valued functions ($i = 1 \ldots, n$). Applying the Corollary for $g(x) = (g_1(x), g_2(x))$, $g_1(x) = \sum_{i=1}^n f_{1,i}(x_i)$, $g_2(x) = \sum_{i=1}^n f_{2,i}(x_i)$, $F(y) = y_1 y_2$ for every $y = (y_1, y_2) \in \mathbb{R}^2_{++}$, we see that if this problem is solvable then there exists a number $t > 0$ such that any optimal solution of

$$\min_x \left\{ \sum_{i=1}^n (f_{1,i}(x_i) + t f_{2,i}(x_i)) \mid x_i \in X_i, \ i = 1, \ldots, n \right\} \tag{9.38}$$

will solve (9.37). For fixed $t > 0$ the problem (9.38) splits into n one-dimensional subproblems

$$\min\{f_{1,i}(x_i) + t f_{2,i}(x_i) \mid x_i \in X_i\}.$$

In the particular case of the problem of optimal ordering policy for jointly replenished products as treated in the mentioned paper of Tanaka et al. $f_{1,i}(x_i) = a/n + a_i/x_i, f_{2,i}(x_i) = b_i x_i/2$ ($a, a_i, b_i > 0$) and X_i is the set of positive integers, so each of the above subproblems can easily be solved.

9.5 Linear Multiplicative Programming

In this section we discuss parametric methods for solving problem (*QCM*) under assumptions (9.30). Let Q be the matrix of rows $c^1, \ldots, c^r$. In view of Proposition 9.1, the problem can be reformulated as

$$\min\{F(\langle c^1, x \rangle, \ldots, \langle c^r, x \rangle) \mid x \in D\}, \tag{9.39}$$

where D is a polyhedron in $\mathbb{R}^n$, and $F : \mathbb{R}^r \to \mathbb{R}$ is a quasiconcave function on a closed convex set $Y \supset Q(D)$, such that for any $y, y' \in Y$:

$$\left. \begin{array}{l} y_i \geq y'_i \ (i = 1, \ldots, p) \\ y_i = y'_i \ (i = p+1, \ldots, r) \end{array} \right\} \Rightarrow F(y) \geq F(y'). \tag{9.40}$$

A typical problem of this class is the *linear multiplicative program* :

$$(LMP) \qquad \min \left\{ \prod_{j=1}^{r} \langle c^i, x \rangle \mid \quad Ax = b, x \geq 0 \right\} \qquad (9.41)$$

which corresponds to the case $F(y) = \prod_{i=1}^{r} y_i$. Aside from (LMP), a wide variety of other problems can be cast in the form (9.39), as shown by the many examples in Sect. 7.1. In spite of that, it turns out that solving (9.39) with different functions $F(y)$ reduces to comparing the objective function values on a finite set of extreme points of D dependent upon the vectors $c^1, \dots, c^r$ but not on the specific form of $F(y)$.

Due to their relevance to applications in various fields (multiobjective programming, bond portfolio optimization, VLSI design, etc...), the above class of problems, including (LMP) and its extensions, has been the subject of quite a few research. Parametric methods for solving (LMP) date back to Swarup (1966a), Swarup (1966b), Swarup (1966c), Forgo (1975), and Gabasov and Kirillova (1980). However intensive development in the framework of global optimization began only with the works of Konno and Kuno (1990, 1992). The basic idea of the parametric method can be described as follows. By Corollary 9.1 one can solve (LMP) by solving the problem

$$\min \{ f(x^{\lambda,\alpha}) \mid \lambda \in \Lambda, \alpha \in W \}, \qquad (9.42)$$

where, in the notation in Corollary 9.1, $x^{\lambda,\alpha}$ is an arbitrary optimal solution of the linear program

$$LP(\lambda, \alpha) \qquad \left| \begin{array}{l} \min \sum_{i=1}^{p} \lambda_i \langle c^i, x \rangle \\ \text{s.t.} \quad x \in D, \ \langle c^j, x \rangle = \alpha_j, \ j = p+1, \dots, r. \end{array} \right.$$

For solving (9.42) the parametric method exploits the property of $LP(\lambda, \alpha)$ that the parameter space is partitioned in finitely many polyhedrons over each of which the minimum of the function $\varphi(\lambda, \alpha) = f(x^{\lambda,\alpha})$ can be computed relatively easily. Specifically, by writing the constraints in this linear program in the form

$$Ax = b + S\alpha, \quad x \geq 0, \qquad (9.43)$$

where S is a suitable diagonal matrix, denote the collection of all basis matrices of this system by $\mathscr{B}$.

Lemma 9.3 *Each basis B of the system (9.43) determines a set $\Pi_B \times \Delta_B \subset \Lambda \times W$ which is a polyhedron whenever nonempty, and an affine map $x^B : \Delta_B \to \mathbb{R}^n$, such that $x^B(\alpha)$ is a basic optimal solution of $LP(\lambda, \alpha)$ for all $(\lambda, \alpha) \in \Pi_B \times \Delta_B$. The collection of polyhedrons $\Pi_B \times \Delta_B$ corresponding to all bases $B \in \mathscr{B}$ covers all (λ, α) such that $LP(\lambda, \alpha)$ has an optimal solution.*

Proof Let B be any basis of this system. If $A = (B, N), x = \begin{pmatrix} x_B \\ x_N \end{pmatrix}$, so that $Bx_B + Nx_N = b + S\alpha$, then

$$x_B = B^{-1}(b + S\alpha) - B^{-1}Nx_N \qquad (9.44)$$

and the associated basic solution of (9.43) is

$$x_B = B^{-1}(b + S\alpha), \ x_N = 0. \qquad (9.45)$$

This basic solution is feasible for all $\alpha \in W$ such that $B^{-1}(b + S\alpha) \geq 0$ (nonnegativity condition) and is dual feasible for all $\lambda \in \Lambda$ such that $\sum_{i=1}^{p} \lambda_i[(c_N^i)^T - (c_B^i)^T B^{-1}N] \geq 0$ (optimality criterion), where $c^i = \begin{pmatrix} c_B^i \\ c_N^i \end{pmatrix}$. If we denote the vector (9.45) by $x^B(\alpha)$ and define the polyhedrons

$$\Pi_B = \left\{ \lambda \in \Lambda | \ \sum_{i=1}^{p} \lambda_i[(c_N^i)^T - (c_B^i)^T B^{-1}N] \geq 0 \right\}$$

$$\Delta_B = \{\alpha \in W| \ B^{-1}(b + S\alpha) \geq 0\}$$

then whenever both Π_B and Λ_B are nonempty the affine map $\alpha \in \Delta_B \mapsto x^B(\alpha) \in \mathbb{R}^n$ is such that $x^B(\alpha)$ is a basic optimal solution of $\mathrm{LP}(\lambda, \alpha)$ for all $\lambda \in \Pi_B, \ \alpha \in \Delta_B$. The last assertion of the Lemma is obvious from the fact that if $\mathrm{LP}(\lambda, \alpha)$ has a basic optimal solution, then $(\lambda, \alpha) \in \Pi_B \times \Delta_B$ for the corresponding basis B. $\qquad\square$

Each polyhedron $\Pi_B \times \Delta_B$ is called a *cell*. Thus, by taking $x^{\lambda,\alpha} = x^B(\alpha)$ for every $(\lambda, \alpha) \in \Pi_B \times \Delta_B$, the function $\varphi(\lambda, \alpha) = f(x^{\lambda,\alpha})$ is constant in $\lambda \in \Pi_B$ for every fixed $\alpha \in \Delta_B$ and quasiconcave in $\alpha \in \Delta_B$ for every fixed $\lambda \in \Pi_B$. Hence its minimum over the cell $\Pi_B \times \Delta_B$ is equal to the minimum of $f(x^B(\alpha))$ over Δ_B and is achieved at a vertex of the polytope Δ_B. Therefore, if the collection of cells can be effectively computed then the problem (9.42) can be solved through scanning the vertices of the cells. Unfortunately, however, when $r > 3$ the cells are very complicated and there is in general no practical method for computing these cells, except in rare cases where additional structure (like network constraints, see Sect. 9.8) may simplify this computation. Below we briefly examine the method for problems with $r \leq 3$. For more detail the reader is referred to the cited papers of Konno and Kuno.

9.5.1 Parametric Objective Simplex Method

When $p = r = 2$ the problem (9.39) is

$$\min\{F(\langle c^1, x \rangle, \langle c^2, x \rangle)| \ x \in D\} \qquad (9.46)$$

where D is a polyhedron contained in the set $\{x \mid \langle c^1, x \rangle \geq 0, \langle c^2, x \rangle \geq 0\}$, and $F(y)$ is a quasiconcave function on the orthant $\mathbb{R}_+^2$ such that

$$y, y' \in \mathbb{R}_+^2, y \geq y' \Rightarrow F(y) \geq F(y'). \tag{9.47}$$

As seen earlier, aside from the special case when $F(y) = y_1 y_2$ (linear multiplicative program), the class (9.46) includes problems with quite different objective functions, such as

$$(\langle c^1, x \rangle)^{q_1} (\langle c^2, x \rangle)^{q_2}; \quad -q_1 e^{\langle c^1, x \rangle} - q_2 e^{\langle c^2, x \rangle} \quad (q_1 > 0, q_2 > 0)$$

corresponding to different choices of $F(y)$:

$$y_1^{q_1} y_2^{q_2}; \quad -q_1 e^{y_1} - q_2 e^{y_2}.$$

By Corollary 9.1 (with $p = r = 2$) the parametric linear program associated with (9.46) is

$$LP(\lambda) \qquad \min\{\langle c^1 + \lambda c^2, x \rangle \mid x \in D\}, \quad \lambda \geq 0. \tag{9.48}$$

Since the parameter domain Λ is the nonnegative real line, the cells of this parametric program are intervals and can easily be determined by the standard *parametric objective simplex method* (see, e.g., Murty 1983) which consists in the following.

Suppose a basis B_k is already available such that the corresponding basic solution x^k is optimal to LP(λ) for all $\lambda \in \Pi_k := [\lambda_{k-1}, \lambda_k]$ but not for λ outside this interval. The latter means that if λ is slightly greater than λ_k then at least one of the inequalities in the optimality criterion

$$(c^1 + \lambda c^2)_{N_k} - (c^1 + \lambda c^2)_{B_k} B_k^{-1} N_k \geq 0$$

becomes violated. Therefore, by performing a number of simplex pivot steps we will pass to a new basis B_{k+1} with a basic solution x^{k+1} which will be optimal to LP(λ) for all λ in a new interval $\Pi_{k+1} \subset [\lambda_k, +\infty)$. In this way, starting from the value $\lambda_0 = 0$ one can determine the interval $\Pi_1 = [\lambda_0, \lambda_1]$ then pass to the next $\Pi_2 = [\lambda_1, \lambda_2]$, and so on, until the whole halfline is covered. Note that the optimum objective value function $\xi(\lambda)$ in the linear program LP(λ) is a concave piecewise affine function of λ with Π_k as linearity intervals. The points $\lambda_0 = 0, \lambda_1, \ldots, \lambda_l$ are the *breakpoints* of $\xi(\lambda)$. An optimal solution to (9.46) is then x^{k*} where

$$k* \in \text{argmin}\{F(\langle c^1, x^k \rangle, \langle c^2, x^k \rangle) \mid k = 1, 2, \ldots, l\}. \tag{9.49}$$

Remark 9.4 Condition (9.47) is a variant of (9.40) for $p = r = 2$. Condition (9.40) with $p = 1, r = 2$ always holds since it simply means $F(y) = F(y')$ whenever

$y = y'$. Therefore, without assuming (9.47), Corollary 9.1 and hence the above method still applies. However, in this case, $\Lambda = \mathbb{R}$, i.e., the linear program $LP(\lambda)$ should be considered for all $\lambda \in (-\infty, +\infty)$.

9.5.2 Parametric Right-Hand Side Method

Consider now problem (9.39) under somewhat weaker assumptions than previously, namely: $F(y)$ is not required to be quasiconcave, while condition (9.47) is replaced by the following weaker one:

$$y_1 \geq y_1', \ y_2 = y_2' \quad \Rightarrow \quad F(y) \geq F(y'). \tag{9.50}$$

The latter condition implies that for fixed y_2 the univariate function $F(., y_2)$ is monotonic on the real line. It is then not hard to verify that Corollary 9.1 still holds for this case. In fact this is also easy to prove directly.

Lemma 9.4 *Let* $\underline{\alpha} = \min\{\langle c^2, x\rangle | \ x \in D\}$, $\overline{\alpha} = \max\{\langle c^2, x\rangle | \ x \in D\}$ *and for every* $\alpha \in [\underline{\alpha}, \overline{\alpha}]$ *let* x^α *be a basic optimal solution of the linear program*

$$LP(\alpha) \qquad\qquad \min\{\langle c^1, x\rangle | \ x \in D, \ \langle c^2, x\rangle = \alpha\} \qquad\qquad .$$

Then

$$\min\{F(\langle c^1, x\rangle, \langle c^2, x\rangle) | \ x \in D\} = \min\{F(\langle c^1, x^\alpha\rangle, \langle c^2, x^\alpha\rangle) | \ \underline{\alpha} \leq \alpha \leq \overline{\alpha}\}.$$

Proof For every feasible solution x of $LP(\alpha)$ we have $\langle c^1, x - x^\alpha\rangle \geq 0$ (because x^α is optimal to LP(α)) and $\langle c^2, x - x^\alpha\rangle = 0$ (because x^α is feasible to LP(α)). Hence, from (9.50) $F(\langle c^1, x\rangle, \langle c^2, x\rangle) \geq F(\langle c^1, x^\alpha\rangle, \langle c^2, x^\alpha\rangle)$ and the conclusion follows.
$\square$

Solving (9.39) thus reduces to solving LP(α) parametrically in α. This can be done by the following standard *parametric right-hand side method* (see, e.g., Murty 1983). Assuming that the constraints in LP(α) have been rewritten in the form

$$Ax = b + \alpha b_0, \ x \geq 0,$$

let B_k be a basis of this system satisfying the dual feasibility condition

$$c_{N_k}^1 - c_{B_k}^1 B_k^{-1} N_k \geq 0.$$

Then the basic solution defined by this basis, i.e., the vector $x^k = u^k + \alpha v^k$ such that $x_{B_k}^k = B_k^{-1}(b + \alpha b_0), x_{N_k}^k = 0$ [see (9.45)], is dual feasible. This basic solution x^k is optimal to LP(α) if and only if it is feasible, i.e., if and only if it satisfies the nonnegativity condition

$$B_k^{-1}(b + \alpha b_0) \geq 0.$$

Let $\Delta_k := [\alpha_{k-1}, \alpha_k]$ be the interval formed by all values α satisfying the above inequalities. Then x^k is optimal to LP(α) for all α in this interval, but for any value α slightly greater than α_k at least one of the above inequalities becomes violated, i.e., at least one of the components of x^k becomes negative. By performing then a number of dual simplex pivots, we can pass to a new basis B_{k+1} with a basic solution x^{k+1} which will be optimal for all α in some interval $\Delta_{k+1} = [\alpha_k, \alpha_{k+1}]$. Thus, starting from the value $\alpha_0 = \underline{\alpha}$, one can determine an initial interval $\Delta_1 = [\underline{\alpha}, \alpha_1]$, then pass to the next interval $\Delta_2 = [\alpha_1, \alpha_2]$, and so on, until the whole interval $[\underline{\alpha}, \overline{\alpha}]$ is covered. The generated points $\alpha_0 = \underline{\alpha}, \alpha_1, \ldots, \alpha_h = \overline{\alpha}$ are the breakpoints of the optimum objective value function $\eta(\alpha)$ in LP(α) which is a convex piecewise affine function. By Lemma 9.4 an optimal solution to (9.39) is then x^{k*} where

$$k^* \in \operatorname{argmin}\{F(\langle c^1, x^k \rangle, \langle c^2, x^k \rangle)| \ k = 1, 2, \ldots, h\}. \tag{9.51}$$

Remark 9.5 In both objective and right-hand side methods the sequence of points x^k on which the objective function values have to be compared depends upon the vectors c^1, c^2 but not on the specific function $F(y)$. Computational experiments (Konno and Kuno 1992), also (Tuy and Tam 1992) indicate that solving a problem (9.39) requires no more effort than solving just a few linear programs of the same size. Two particular problems of this class are worth mentioning:

$$\min\{c^1 x . c^2 x| \quad x \in D\} \tag{9.52}$$

$$\min\{-x_1^2 + c_0 x| \quad x \in D\}, \tag{9.53}$$

where D is a polyhedron [contained in the domain $c^1 x > 0, c^2 x > 0$ for (9.52)]. These correspond to functions $F(y) = y_1 y_2$ and $F(y) = -y_1^2 + y_2$, respectively and have been shown to be NP-hard (Matsui 1996; Pardalos and Vavasis 1991), although, as we saw above, they are efficiently solved by the parametric method.

9.5.3 Parametric Combined Method

When $r = 3$ the parametric linear program associated to (9.39) is difficult to handle because the parameter domain is generally two-dimensional. In some particular cases, however, a problem (9.39) with $r = 3$ may be successfully tackled by the parametric method. As an example consider the problem

$$\min\{c^0 x + c^1 x . c^2 x| \quad x \in D\} \tag{9.54}$$

which may not belong to the class (QCM) because we cannot prove the quasiconcavity of the function $F(y) = y_1 + y_2 y_3$. It is easily seen, however, that this problem is equivalent to

$$\min\{\langle c^0 + \alpha c^1, x \rangle| \quad x \in D, \ \langle c^2, x \rangle = \alpha, \ \underline{\alpha} \le \alpha \le \overline{\alpha}\} \tag{9.55}$$

where $\underline{\alpha} = \min_{x \in D}\langle c^2, x \rangle$, $\overline{\alpha} = \max_{x \in D}\langle c^2, x \rangle$. By writing the constraints $x \in D$, $\langle c^2, x \rangle = \alpha$ in the form

$$Ax = b + \alpha b_0. \tag{9.56}$$

we can prove the following:

Lemma 9.5 *Each basis B of the system (9.56) determines an interval $\Delta_B \subset [\underline{\alpha}, \overline{\alpha}]$ and an affine map $x^B : \Delta_B \to \mathbb{R}^n$ such that $x^B(\alpha)$ is a basic optimal solution of the problem*

$$P(\alpha) \qquad\qquad \min\{\langle c^0 + \alpha c^1, x \rangle | \quad x \in D, \ \langle c^2, x \rangle = \alpha\}$$

for all $\alpha \in \Delta_B$. The collection of all intervals Δ_B corresponding to all bases B such that $\Delta_B \ne \emptyset$ covers all α for which $P(\alpha)$ has an optimal solution.

Proof Let B be a basis, $A = (B, N), x = \begin{pmatrix} x_B \\ x_N \end{pmatrix}$, so that the associated basic solution is

$$x_B = B^{-1}(b + \alpha b_0), \quad x_N = 0. \tag{9.57}$$

This basic solution is feasible for all $\alpha \in [\underline{\alpha}, \overline{\alpha}]$ satisfying the nonnegativity condition

$$B^{-1}(b + \alpha b_0) \ge 0. \tag{9.58}$$

and is dual feasible for all α satisfying the dual feasibility condition (optimality criterion)

$$(c^0 + \alpha c^1)_N - (c^0 + \alpha c^1)_B B^{-1} N \ge 0. \tag{9.59}$$

Conditions (9.58) and (9.59) determine each an interval. If the intersection Δ_B of the two intervals is nonempty, then the basic solution (9.57) is optimal for all $\alpha \in \Delta_B$. The rest is obvious.   □

Based on this Lemma, one can generate all the intervals Δ_B corresponding to different bases B by a procedure combining primal simplex pivots with dual simplex pivots (see, e.g., Murty 1983). Specifically, for any given basis B_k, denote by $u^k + \alpha v^k$ the vector (9.57) which is the basic solution of $P(\alpha)$ corresponding to this basis, and by $\Delta_k = [\alpha_{k-1}, \alpha_k]$ the interval of optimality of this basic solution. When α is slightly greater than α_k either the nonnegativity condition [see (9.58)] or the dual feasibility condition [see (9.59)] becomes violated. In the latter case we perform a primal simplex pivot step, in the former case a dual simplex pivot step, and repeat as long as necessary, until a new basis B_{k+1} with a new interval Δ_{k+1} is obtained. Thus, starting from an initial interval $\Delta_1 = [\underline{\alpha}, \alpha_1]$ we can pass from one interval to

the next, until we reach $\overline{\alpha}$. Let then $\Delta_k, k = 1, \ldots, l$ be the collection of all intervals. For each interval Δ_k let $u^k + \alpha v^k$ be the associated basic solution which is optimal to P(α) for all $\alpha \in \Delta_k$ and let $x^k = u^k + \xi_k v^k$, where ξ_k is a minimizer of the quadratic function $\varphi(\alpha) := \langle c^0 + \alpha c^1, u^k + \alpha v^k \rangle$ over Δ_k. Then from (9.55) and by Lemma 9.5, an optimal solution of (9.54) is x^{k*} where

$$k^* \in \operatorname{argmin}\{c^0 x^k + (c^1 x^k)(c^2 x^k)| \ k = 1, 2, \ldots, l\}.$$

Note that the problem (9.54) is also NP-hard (Pardalos and Vavasis 1991).

9.6 Convex Constraints

So far we have been mainly concerned with solving quasiconcave monotonic minimization problems under linear constraints. Consider now the problem

$$(QCM) \qquad \min\{f(x)| \quad x \in D\}. \qquad (9.60)$$

formulated in Sect. 9.2 [see (9.8)], where the constraint set D is compact convex but not necessarily polyhedral. Recall that by Proposition 9.1 where $Qx = (c^1 x, \ldots, c^r x)^T$ there exists a continuous quasiconcave function $F : \mathbb{R}^r \to \mathbb{R}$ on a closed convex set $Y \supset Q(D)$ satisfying condition (9.40) and such that

$$f(x) = F(\langle c^1, x \rangle, \ldots, \langle c^r, x \rangle). \qquad (9.61)$$

It is easy to extend the branch and bound algorithms in Sect. 8.2 and the polyhedral annexation algorithm in Sect. 8.3 to (QCM) with a compact convex constraint set D, and in fact, to the following more general problem:

$$(GQCM) \qquad \min\{F(g(x))| \quad x \in D\} \qquad (9.62)$$

where D and $F(y)$ are as previously (with $g(D)$ replacing $Q(D)$), while $g = (g_1, \ldots, g_r)^T : \mathbb{R}^n \to \mathbb{R}^r$ is a map such that $g_i(x), i = 1, \ldots, p$, are convex on a closed convex set $X \supset D$ and $g_i(x) = \langle c^i, x \rangle + d_i, \ i = p + 1, \ldots, r$.

For the sake of simplicity we shall focus on the case $p = r$, i.e., we shall assume $F(y) \geq F(y')$ whenever $y, y' \in Y, y \geq y'$. This occurs, for instance, for the convex multiplicative programming problem

$$\min\left\{\prod_{i=1}^{r} g_i(x)| \ x \in D\right\} \qquad (9.63)$$

where $F(y) = \prod_{i=1}^{r} y_i$, $D \subset \{x| \ g_i(x) \geq 0\}$. Clearly problem (GQCM) can be rewritten as

$$\min\{F(y)| \ g(x) \leq y, x \in D\} \tag{9.64}$$

which is to minimize the quasiconcave $\mathbb{R}_+^r$-monotonic function $F(y)$ on the closed convex set

$$E = \{y|g(x) \leq y, x \in D\} = g(D) + \mathbb{R}_+^r. \tag{9.65}$$

9.6.1   Branch and Bound

If the function $F(y)$ is concave (and not just quasiconcave) then simplicial subdivision can be used. In this case, a lower bound for $F(y)$ over the feasible points y in a simplex $M = [u^1, \ldots, u^{r+1}]$ can be obtained by solving the convex program

$$\min\{l_M(y)| \ y \in E \cap M\},$$

where $l_M(y)$ is the affine function that agrees with $F(y)$ at the vertices of M (this function satisfies $l_M(y) \leq F(y) \ \forall y \in M$ in view of the concavity of $F(y)$). Using this lower bounding rule, a simplicial algorithm for (GQCM) can be formulated following the standard branch and bound scheme. It is also possible to combine branch and bound with outer approximation of the convex set E by a sequence of nested polyhedrons $P_k \supset E$.

If the function $F(y)$ is concave separable, i.e., $F(y) = \sum_{j=1}^{r} F_j(y_j)$, where each $F_j(.)$ is a concave function of one variable, then rectangular subdivision can be more convenient. However, in the general case, when $F(y)$ is quasiconcave but not concave, an affine minorant of a quasiconcave function over a simplex (or a rectangle) is in general not easy to compute. In that case, one should use conical rather than simplicial or rectangular subdivision.

9.6.2   Polyhedral Annexation

Since $F(y)$ is monotonic with respect to the orthant $\mathbb{R}_+^r$, we can apply the PA Algorithm using Lemma 9.2 for constructing the initial polytope P_1.

Let $\bar{y} \in E$ be an initial feasible point of (9.64) and $C = \{y| \ F(y) \geq F(\bar{y})\}$. By translating, we can assume that $0 \in E$ and $0 \in \text{int} C$, i.e., $F(0) > F(\bar{y})$. Note that instead of a linear program like (9.25) we now have to solve a convex program of the form

$$\max\{\langle t, y\rangle| \ y \in E\} = \max\{\langle t, g(x)\rangle| \ x \in D\} \tag{9.66}$$

where $t \in (\mathbb{R}^r_+)^\circ = -\mathbb{R}^r_+$. Therefore, to construct the initial polytope P_1 (see Lemma 9.2) we compute $\alpha > 0$ such that $F(\alpha w) = F(\bar{y})$ (where $w = (-1, \ldots, -1)^T \in \mathbb{R}^r$) and take the initial polytope P_1 to be the r-simplex in $\mathbb{R}^r$ with vertex set $V_1 = \{0, -e^1/\alpha, \ldots, -e^r/\alpha\}$. We can thus state:

PA Algorithm (for (*GQCM*))

Initialization. Let $\bar{y}$ be a feasible solution (the best available), $C = \{y| F(y) \geq F(\bar{y})\}$. Choose a point $y^0 \in E$ such that $F(y^0) > F(\bar{y})$ and set $E \leftarrow E - y^0$, $C \leftarrow C - y^0$. Let P_1 be the simplex in $\mathbb{R}^r$ with vertex set V_1. Set $k = 1$.

Step 1. For every new $t = (t_1, \ldots, t_r)^T \in V_k \backslash \{0\}$ solve the convex program (9.66), obtaining the optimal value $\mu(t)$ and an optimal solution $y(t)$.

Step 2. Let $t^k \in \operatorname{argmax}\{\mu(t)| t \in V_k\}$. If $\mu(t^k) \leq 1$, then terminate: $\bar{y}$ is an optimal solution of (*GQCM*).

Step 3. If $\mu(t^k) > 1$ and $y^k := y(t^k) \notin C$ (i.e., $F(y^k) < F(\bar{y})$), then update the current best feasible solution and the set C by resetting $\bar{y} = y^k$.

Step 4. Compute

$$\theta_k = \sup\{\theta| \ F(\theta y^k) \geq F(\bar{y})\} \tag{9.67}$$

and define

$$P_{k+1} = P_k \cap \{t| \langle t, y^k \rangle \leq \frac{1}{\theta_k}\}.$$

Step 5. From V_k derive the vertex set V_{k+1} of P_{k+1}. Set $k \leftarrow k + 1$ and go back to Step 1.

To establish the convergence of this algorithm we need two lemmas.

Lemma 9.6 $\mu(t^k) \searrow 1$ *as $k \to +\infty$.*

Proof Observe that $\mu(t) = \max\{\langle t, y \rangle| y \in g(D)\}$ is a convex function and that $y^k \in \partial\mu(t^k)$ because $\mu(t) - \mu(t^k) \geq \langle t, y^k \rangle - \langle t^k, y^k \rangle = \langle t - t^k, y^k \rangle \ \forall t$. Denote $l_k(t) = \langle t, y^k \rangle - 1$. We have $l_k(t^k) = \langle t^k, y^k \rangle - 1 = \mu(t^k) - 1 > 0$, and for $t \in [g(D)]^\circ$: $l_k(t) = \langle t, y^k \rangle - 1 \leq \mu(t) - 1 \leq 0$ because $y^k \in g(D) + \mathbb{R}^r_+$. Since $g(D)$ is compact, there exist $t^0 \in \operatorname{int}[g(D)]^\circ$ (Proposition 1.21), i.e., t^0 such that $l(t^0) = \mu(t^0) - 1 < 0$ and $s^k \in [t^0, t^k)] \setminus \operatorname{int}[g(D)]^\circ$ such that $l(s^k) = 0$. By Theorem 6.1 applied to the set $[g(D)]^\circ$, the sequence $\{t^k\}$ and the cuts $l_k(t)$, we conclude that $t^k - s^k \to 0$. Hence every accumulation point t^* of $\{t^k\}$ satisfies $\mu(t^*) = 1$, i.e., $\mu(t^k) \searrow \mu(t^*) = 1$. $\qquad\square$

Denote by C_k the set C at iteration k.

Lemma 9.7 $C^\circ_k \subset P_k$ *for every k.*

Proof The inclusion $C_1^\circ \subset P_1$ follows from the construction of P_1. Arguing by induction on k, suppose that $C_k^\circ \subset P_k$ for some $k \geq 1$. Since $\theta_k y^k \in C_{k+1}$, for all $t \in C_{k+1}^\circ$ we must have $\langle t, \theta_k y^k \rangle \leq 1$, and since $C_k \subset C_{k+1}$, i.e., $C_{k+1}^\circ \subset C_k^\circ \subset P_k$ it follows that $t \in P_k \cap \{t \mid \langle t, \theta_k y^k \rangle \leq 1\} = P_{k+1}$.   □

Proposition 9.5 *Let $\bar{y}^k$ be the incumbent at iteration k. Either the above algorithm terminates by an optimal solution of (GQCM) or it generates an infinite sequence $\{\bar{y}^k\}$ every accumulation point of which is an optimal solution.*

Proof Let y^* be any accumulation point of the sequence $\{\bar{y}^k\}$. Since by Lemma 9.6 $\max\{\mu(t) \mid t \in P_k\} \searrow 1$ it follows from Lemma 9.7 that $\max\{\mu(t) \mid t \in \cap_{k=1}^\infty C_k^\circ\} \leq 1$. This implies that $\cap_{k=1}^\infty C_k^\circ \subset E^\circ$, and hence $E \subset \cup_{k=1}^\infty C_k$. Thus, for any $y \in E$ there is k such that $y \in C_k$, i.e., $F(y) \geq F(\bar{y}^k) \geq F(\bar{y})$, proving the optimality of $\bar{y}$.   □

Incidentally, the above PA Algorithm shows that

$$\min\{F(y) \mid y \in E\} = \min\{F(g(x(t))) \mid t \in \mathbb{R}_-^r\}, \qquad (9.68)$$

where $x(t)$ is an arbitrary optimal solution of (9.66), i.e., of the convex program

$$\min\{\langle -t, g(x) \rangle \mid x \in D\}.$$

This result could also be derived from Theorem 9.1.

Remark 9.6 All the convex programs (9.66) have the same constraints. This fact should be exploited for an efficient implementation of the above algorithm. Also, in practice, these programs are usually solved approximately, so $\mu(t)$ is only an ε-optimal value, i.e., $\mu(t) \geq \max\{\langle t, y \rangle \mid y \in E\} - \varepsilon$ for some tolerance $\varepsilon > 0$.

9.6.3 Reduction to Quasiconcave Minimization

An important class of problems (GQCM) is constituted by problems of the form:

$$\text{minimize} \sum_{i=1}^p \prod_{j=1}^{q_i} g_{ij}(x) \quad \text{s.t} \quad x \in D, \qquad (9.69)$$

where D is a compact convex set and $g_{ij} : D \to \mathbb{R}_{++}, i = 1, \ldots, p, j = 1, \ldots, q_i$, are continuous convex positive-valued functions on D. This class includes *generalized convex multiplicative programming problems* (Konno and Kuno 1995; Konno et al. 1994) which can be formulated as:

$$\min\left\{g(x) + \prod_{j=1}^{q} g_j(x) \mid x \in D\right\}. \tag{9.70}$$

or, alternatively as

$$\min\left\{g(x) + \sum_{i=1}^{p} g_i(x) h_i(x) \mid x \in D\right\}, \tag{9.71}$$

where all functions $g(x), g_i(x), h_i(x)$ are convex positive-valued on D. Another special case worth mentioning is the problem of minimizing the scalar product of two vectors:

$$\min\left\{\sum_{i=1}^{n} x_i y_i \mid (x, y) \in D \subset \mathbb{R}^n_{++}\right\}. \tag{9.72}$$

Let us show that any problem (9.69) can be reduced to concave minimization under convex constraints. For this we rewrite (9.69) as

$$
\begin{aligned}
\text{minimize} \quad & \sum_{i=1}^{p} \left(\prod_{j=1}^{q_i} y_{ij}\right)^{1/q_i} \\
\text{s.t.} \quad & [g_{ij}(x)]^{q_i} \le y_{ij}, \; i = 1, \ldots, p; j = 1, \ldots, q_i \\
& y_{ij} \ge 0, \quad x \in D.
\end{aligned} \tag{9.73}
$$

Since $\varphi_j(y) = y_{ij}, j = 1, \ldots, q_i$ are linear, it follows from Proposition 2.7 that each term $(\prod_{j=1}^{q_i} y_{ij})^{1/q_i}$, $i = 1, \ldots, p$, is a concave function of $y = (y_{ij}) \in \mathbb{R}^{q_1 + \ldots q_p}$, hence their sum, i.e., the objective function $F(y)$ in (9.73), is also a concave function of y. Furthermore, $y \ge y'$ obviously implies $F(y) \ge F(y')$, so $F(y)$ is monotonic with respect to the cone $\{y \mid y \ge 0\}$. Finally, since every function $g_{ij}(x)$ is convex, so is $[g_{ij}(x)]^{q_i}$. Consequently, (9.73) is a concave programming problem. If $q = q_1 + \ldots + q_p$ is relatively small this problem can be solved by methods discussed in Chaps. 5 and 6.

An alternative way to convert (9.69) into a concave program is to use the following relation already established in the proof of Proposition 2.7 [see (2.6)]:

$$\prod_{j=1}^{q_i} g_{ij}(x) = \frac{1}{q_i} \min_{\xi^i \in T_i} \sum_{j=1}^{q_i} \xi_{ij} g_{ij}^{q_i}(x),$$

with $\xi^i = (\xi_{i1}, \ldots, \xi_{iq_i}), T_i = \{\xi^i : \prod_{j=1}^{q_i} \xi_{ij} \ge 1\}$. Hence, setting $\varphi_i(\xi^i, x) = \frac{1}{q_i} \sum_{j=1}^{q_i} \xi_{ij} g_{ij}^{q_i}(x)$, the problem (9.69) is equivalent to

$$\min_{x \in D} \sum_{i=1}^{p} \min_{\xi^i \in T_i} \varphi_i(\xi^i, x)$$

$$= \min_{x \in D} \min \{ \sum_{i=1}^{p} \varphi_i(\xi^i, x) : \ \xi^i \in T_i, i = 1, \ldots, p \}$$

$$= \min \{ \min_{x \in D} \sum_{i=1}^{p} \varphi_i(\xi^i, x) : \ \xi^i \in T_i, i = 1, \ldots, p \}$$

$$= \min \{ \varphi(\xi^1, \ldots, \xi^p) | \ \xi^i \in T_i, i = 1, \ldots, p \} \tag{9.74}$$

where

$$\varphi(\xi^1, \ldots, \xi^p) = \min_{x \in D} \sum_{i=1}^{p} \varphi_i(\xi^i, x) = \min \left\{ \sum_{i=1}^{p} \sum_{j=1}^{q_i} \frac{1}{q_i} \xi_{ij} g_{ij}^{q_i}(x) | \ x \in D \right\}.$$

Since each function $g_{ij}^{q_i}(x)$ is convex, $\varphi(\xi^1, \ldots, \xi^p)$ is the optimal value of a convex program. Furthermore, the objective function of this convex program is a linear function of $(\xi^1, \ldots, \xi^p)$ for fixed x. Therefore, $\varphi(\xi^1, \ldots, \xi^p)$ is a concave function (pointwise minimum of a family of linear functions). Finally, each $T_i, i = 1, \ldots, p$, is a closed convex set in $\mathbb{R}^{q_i}$ because

$$T_i = \left\{ \xi^i | \ \sum_{j=1}^{q_i} \log \xi_{ij} \geq 0 \right\}.$$

Thus, problem (9.69) is equivalent to (9.74) which seeks to minimize the concave function $\varphi(\xi^1, \ldots, \xi^p)$ over the convex set $\prod_{i=1}^{p} T_i$.

Example 9.9 As an illustration, consider the problem of determining a rectangle of minimal area enclosing the projection of a given compact convex set $D \subset \mathbb{R}^n (n > 2)$ onto the plane $\mathbb{R}^2$. This problem arises from applications in packing and optimal layout (Gale 1981; Haims and Freeman 1970; Maling et al. 1982), especially when two-dimensional layouts are restricted by $n - 2$ factors. When the vertices of D are known, there exists an efficient algorithm based on computational geometry (Bentley and Shamos 1978; Graham 1972; Toussaint 1978). In the general case, however, the problem is more complicated.

Assume that D has full dimension in $\mathbb{R}^n = \mathbb{R}^2 \times \mathbb{R}^{n-2}$. The projection of D onto $\mathbb{R}^2$ is the compact convex set $\mathrm{pr}D = \{ y \in \mathbb{R}^2 | \ \exists z \in \mathbb{R}^{n-2}, \ (y, z) \in D \}$. For any vector $x \in \mathbb{R}^2$ such that $\|x\| = 1$ define

$$g_1(x) = \max \{ \langle x, y \rangle | \ (y, z) \in D \}$$

$$g_2(x) = \min \{ \langle x, y \rangle | \ (y, z) \in D \}.$$

Then the width of pr$D$ in the direction x is

$$g(x) = g_1(x) - g_2(x).$$

On the other hand, since $Hx = (-x_2, x_1)^T$ is the vector orthogonal to x such that $\|H(x)\| = 1$ if $\|x\| = 1$, the width of pr$D$ in the direction orthogonal to x is $g(Hx) = g(-x_2, x_1)$. Hence the product $g(x).g(Hx)$ measures the area of the smallest rectangle containing pr$D$ and having a side parallel to x. The problem can then be formulated as

$$\min\{g(x).g(Hx)|\ x \in \mathbb{R}_+^2,\ \|x\| = 1\}. \tag{9.75}$$

Lemma 9.8 *The function $g(x)$ is convex and satisfies $g(\alpha x) = \alpha g(x)$ for any $\alpha \geq 0$.*

Proof Clearly $g_1(x)$ is convex as the pointwise maximum of a family of affine functions, while $g_2(x)$ is concave as the pointwise minimum of a family of affine functions. It is also obvious that $g_i(\alpha x) = \alpha g_i(x), i = 1, 2$. Hence, the conclusion.

□

Since $H : \mathbb{R}^2 \to \mathbb{R}^2$ is a linear map, it follows that $g(Hx)$ is also convex and that $g(H(\alpha x)) = \alpha g(Hx)$ for every $\alpha \geq 0$. Exploiting this property, a successive underestimation method for solving problem (9.75) has been proposed in Kuno (1993). In view of the above established property, problem (9.75) is actually equivalent to the convex multiplicative program

$$\min\{g(x) \cdot g(Hx)|\quad x \in \mathbb{R}_+^2, \|x\| \leq 1\}, \tag{9.76}$$

and so can be transformed into the concave minimization problem

$$\begin{aligned}
&\min\ \sqrt{t_1 t_2} \\
&\text{s.t.}\ \ g(x) \leq t_1,\ g(Hx) \leq t_2 \\
&\qquad x \in \mathbb{R}_+^2,\ \|x\| \leq 1,
\end{aligned}$$

or, alternatively,

$$\min\{\varphi(\xi_1, \xi_2)|\ \xi_1 \xi_2 \geq 1, \xi \in \mathbb{R}_+^2\}. \tag{9.77}$$

where

$$\varphi(\xi_1, \xi_2) = \min\{\xi_1 g(x) + \xi_2 g(Hx)|\ x \geq 0, \|x\| \leq 1\}. \tag{9.78}$$

In the special case when D is a *polytope* given by a system of linear inequalities $Ay + Bz \leq b$ the problem can be solved by an efficient parametric method (Kuno 1993). For any point $x \in \mathbb{R}_+^2$ with $\|x\| = 1$, let $\sigma(x) = (\lambda, 1 - \lambda)$ be the point

where the halfline from 0 through x intersects the line segment joining $(1, 0)$ and $(0, 1)$. Since obviously, $\|\sigma(x)\|^2 = \lambda^2 + (1 - \lambda)^2$, using Lemma 9.8 problem (9.76) becomes

$$\min\{F(\lambda)|\ \lambda \in [0, 1]\}, \tag{9.79}$$

where

$$F(\lambda) = \frac{g(\lambda, 1 - \lambda).g(\lambda - 1, \lambda)}{\lambda^2 + (1 - \lambda)^2}.$$

Furthermore, $g(\lambda, 1 - \lambda) = g_1(\lambda, 1 - \lambda) - g_2(\lambda, 1 - \lambda)$ with

$$g_1(\lambda, 1 - \lambda) = \max\{\lambda y_1 + (1 - \lambda)y_2|\ Ay + Bz \leq b\}$$
$$g_2(\lambda, 1 - \lambda) = \min\{\lambda y_1 + (1 - \lambda)y_2|\ Ay + Bz \leq b\}.$$

So each $g_i(\lambda, 1 - \lambda), i = 1, 2$ is the optimal value of a linear parametric program. Let $0, \lambda_1^{1i}, \dots, \lambda_{p_{1i}}^{1i}, 1$ be the sequence of breakpoints of $g_i(\lambda, 1 - \lambda), i = 1, 2$. Analogously, let $0, \lambda_1^{2i}, \dots, \lambda_{p_{2i}}^{2i}, 1$ be the sequence of breakpoints of $g_i(\lambda - 1, \lambda)$. Finally let

$$0 = \lambda_0, \lambda_1, \lambda_2, \dots, \lambda_{N-1}, 1 = \lambda_N \tag{9.80}$$

be the union of all these four sequences rearranged in the increasing order.

Proposition 9.6 *A global minimum of the function $F(\lambda)$ over $[0, 1]$ exists among the points $\lambda_s, s = 0, \dots, N$.*

Proof For any interval $[\lambda_s, \lambda_{s+1}]$, there exist two vectors $u = x^1 - x^2, v = x^3 - v^4$, where $x^i, i = 1, \dots, 4$ are vertices of D, such that for every $x \in \mathbb{R}_+^2$ of length $\|x\| = 1$, with $\sigma(x) = (\lambda, 1 - \lambda), \lambda \in (\lambda_s, \lambda_{s+1})$, we have $g(x) = \langle x, u \rangle, g(Hx) = \langle Hx, v \rangle$. Then

$$F(\lambda) = (\|u\| \cos \alpha) \times (\|v\| \cos \beta),$$

where $\alpha = \text{angle}(x, u)$ (angle between the vectors x and u) and $\beta = \text{angle}(Hx, v)$. Noting that $\text{angle}(x, u) + \text{angle}(u, v) + \text{angle}(v, y) = \text{angle}(x, y)$, we deduce

$$\beta = \alpha + \omega - \pi/2, \tag{9.81}$$

where $\omega = \text{angle}(u, v)$. Thus, for all $\lambda \in (\lambda_s, \lambda_{s+1})$:

$$F(\lambda) = (\|u\| \cos \alpha)(\|v\| \sin(\alpha + \omega)).$$

Setting $G(\alpha) = \cos \alpha \sin(\alpha + \omega)$, we have $G'(\alpha) = -\sin \alpha \sin(\alpha + \omega) + \cos \alpha \cos(\alpha + \omega) = \cos(2\alpha + \omega)$. If $F(.)$ attains a minimum at λ, i.e., $G(.)$ attains a minimum at α, then one must have

$$\cos(2\alpha + \omega) = 0, \text{ hence } 2\alpha + \omega = \pi/2 \pm \pi. \tag{9.82}$$

But from (9.81) it follows that $\alpha + \omega = \beta + \pi/2$, hence $2\alpha + \omega = \alpha + \beta + \pi/2$. In view of (9.82) this in turn implies that

$$\alpha + \beta = \pm\pi,$$

contradicting the fact that $0 < |\alpha| < \pi/2$, $0 < |\beta| < \pi/2$ (because $\langle x, u \rangle > 0$, $\langle Hx, v \rangle > 0$). Therefore, the minimum of $F(\lambda)$ over an interval $[\lambda_s, \lambda_{s+1}]$ cannot be attained at any λ such that $\lambda_s < \lambda < \lambda_{s+1}$, but must be attained at either λ_s or λ_{s+1}. Hence the global minimum over the entire segment $[0, 1]$ is attained at one point of the sequence $\lambda_0, \lambda_1, \dots, \lambda_N$.   □

Corollary 9.3 *An optimal solution of (9.79) is*

$$x^* = \frac{(\lambda_*, 1 - \lambda_*)}{\sqrt{\lambda_*^2 + (1 - \lambda_*)^2}}, \quad \lambda_* \in \operatorname{argmin}\{F(\lambda_s) : s = 0, 1, \dots, N\}.$$

Proof Indeed, $x^* = \sigma(x^*)/\|\sigma(x^*)\|$ and $\sigma(x^*) = (\lambda_*, 1 - \lambda_*)$.

9.7   Reverse Convex Constraints

In the previous sections we were concerned with low rank nonconvex problems with nonconvex variables in the objective function. We now turn to low rank nonconvex problems with nonconvex variables in the constraints. Consider the following *monotonic reverse convex problem*

$$(MRP) \qquad \min\{\langle c, x \rangle | \quad x \in D, \ h(x) \le 0\},$$

where $c \in \mathbb{R}^n, D$ is a closed convex set in $\mathbb{R}^n$, and the function $h : X \to \mathbb{R}$ is quasiconcave and K-monotonic on a closed convex set $X \supset D$. In this section we shall present decomposition methods for (MRP) by specializing the algorithms for reverse convex programs studied in Chaps. 6 and 7.

As in Sect. 9.2 assume that $K \subset \operatorname{rec}X$, with lineality $L = \{x| Qx = 0\}$, where Q is an $r \times n$ matrix with r linearly independent rows $c^1, \dots, c^r$. By quasiconcavity of $h(x)$ the set $C = \{x \in X| h(x) \ge 0\}$ is convex and by Proposition 9.2 $K \subset \operatorname{rec}C$. Assume further that $D = \{x| Ax \le b, x \ge 0\}$ is a polytope with a vertex at 0 and that:

(a) $\min\{\langle c, x \rangle | x \in D, h(x) \le 0\} > 0$, so that $h(0) > 0$;
(b) The problem is regular, i.e.,

$$D \setminus \operatorname{int}C = \operatorname{cl}(D \setminus C).$$

9.7.1   Decomposition by Projection

By writing, as in Sect. 7.2 [see formulas (9.11) and (9.12)], $Q = [Q_B, Q_N]$, $x = \begin{bmatrix} x_B \\ x_N \end{bmatrix}$, where Q_B is an $r \times r$ nonsingular matrix, and setting

$$Z = \begin{bmatrix} Q_B^{-1} \\ 0 \end{bmatrix}, \quad \varphi(y) = h(Zy)$$

we define a function $\varphi(y)$ such that $\varphi(y) = h(x)$ for all $x \in X$ satisfying $y = Qx$. The latter also implies that $\varphi(y)$ is quasiconcave on $Q(X)$.

Denote by $\psi(y)$ the optimal value of the linear program

$$\min\{\langle c, x \rangle \mid x \in D, \; Qx = y\}.$$

Then (*MRP*) can be rewritten as a problem in y :

$$\min\{\psi(y) \mid y \in Q(D), \; \varphi(y) \leq 0\}. \tag{9.83}$$

Since it is well known that $\psi(y)$ is a convex piecewise affine function, (9.83) is a reverse convex program in the r-dimensional space $Q(\mathbb{R}^n)$. If r is not too large this program can be solved in reasonable time by outer approximation or branch and bound.

9.7.2   Outer Approximation

Following the OA Algorithm for (*CDC*) (Sect. 6.3), we construct a sequence of polytopes $P_1 \supset P_2 \supset \ldots$ outer approximating the convex set

$$\Omega = \{y \in Q(D) \mid \psi(y) \leq \tilde{\gamma}\},$$

where $\tilde{\gamma}$ is the optimal value of (*MRP*). Also we construct a sequence $+\infty \geq \gamma_1 \geq \gamma_2 \geq \ldots$ such that

$$\gamma_k < +\infty \quad \Rightarrow \quad \exists \bar{y}^k \in Q(D), \; \varphi(\bar{y}^k) \leq 0, \psi(\bar{y}^k) = \gamma_k \;;$$
$$\{y \in Q(D) \mid \psi(y) \leq \gamma_k\} \subset P_k.$$

At iteration k we select (by vertex enumeration)

$$y^k \in \operatorname{argmin}\{\varphi(y) \mid y \in P_k\}$$

(note that this requires $\varphi(y)$ to be defined on P_1). If $\varphi(y^k) \geq 0$ then (MRP) is infeasible when $\gamma_k = +\infty$, or $\bar{y}^k$ is optimal when $\gamma_k < +\infty$. If $\varphi(y^k) < 0$, then $y^k \notin \Omega$ and we construct an affine inequality strictly separating y^k from Ω. The polytope P_{k+1} is defined by appending this inequality to P_k.

The key issue in this OA scheme is, given a point y^k such that $\varphi(y^k) < 0$, to construct an affine function $l(y)$ satisfying

$$l(y^k) > 0, \qquad l(y) \leq 0 \quad \forall y \in \Omega. \tag{9.84}$$

Note that we assume (a) and $0 \in D$, so

$$\varphi(0) > 0, \quad \psi(0) < \min\{\psi(y) \mid y \in Q(D), \ \varphi(y) \leq 0\}. \tag{9.85}$$

Now, since $\varphi(y^k) < 0 < \varphi(0)$ and $\gamma_k - \psi(y^k) \leq 0 < \gamma_k - \psi(0)$, one can compute $z^k \in [0, y^k]$ satisfying $\min\{\varphi(z^k), \gamma_k - \psi(z^k)\} = 0$. Consider the linear program

$$(SP(z^k)) \qquad \max\{\langle z^k, v \rangle - \langle b, w \rangle \mid Q^T v - A^T w \leq c, v \geq 0\}.$$

Proposition 9.7

(i) *If $SP(z^k)$ has a (finite) optimal solution (v^k, w^k) then $z^k \in Q(D)$ and (9.84) is satisfied by the affine function*

$$l(y) := \langle v^k, y - z^k \rangle. \tag{9.86}$$

(ii) *If $SP(z^k)$ has no finite optimal solution, so that the cone $Q^T v - A^T w \leq d$, $w \geq 0$ has an extreme direction (v^k, w^k) such that $\langle z^k, v^k \rangle - \langle b, w^k \rangle > 0$, then $z^k \notin Q(D)$ and (9.84) is satisfied by the affine function*

$$l(y) := \langle y, v^k \rangle - \langle b, w^k \rangle. \tag{9.87}$$

Proof If $SP(z^k)$ has an optimal solution (v^k, w^k) then its dual

$$(SP^*(z^k)) \qquad \min\{\langle c, x \rangle \mid Ax \leq b, \ Qx = z^k, \ x \geq 0\}$$

is feasible, hence $z^k \in Q(D)$. It is easy to see that $v^k \in \partial \psi(z^k)$. Indeed, for any y, since $\psi(y)$ is the optimal value in $SP^*(y)$, it must also be the optimal value in $SP(y)$; i.e., $\psi(y) \geq \langle y, v^k \rangle - \langle b, w^k \rangle$, hence $\psi(y) - \psi(z^k) \geq \langle y, v^k \rangle - \langle b, w^k \rangle - \langle z^k, v^k \rangle + \langle b, w^k \rangle = \langle v^k, y - z^k \rangle$, proving the claim. For every $y \in \Omega$ we have $\psi(y) \leq \tilde{\gamma} \leq \psi(z^k)$, so $l(y) := \langle v^k, y - z^k \rangle \leq \psi(y) - \psi(z^k) \leq 0$. On the other hand, $y^k = \theta z^k = \theta z^k + (1 - \theta)0$ for some $\theta \geq 1$, hence $l(y^k) = \theta l(z^k) + (1 - \theta)l(0) > 0$ because $l(0) \leq \psi(0) - \psi(z^k) < 0$, while $l(z^k) = 0$.

To prove (ii), observe that if $SP(z^k)$ is unbounded, then $SP^*(z^k)$ is infeasible, i.e., $z^k \notin Q(D)$. Since the recession cone of the feasible set of $SP(z^k)$ is the cone $Q^T v - A^T w \leq 0, v \geq 0$, $SP(z^k)$ may be unbounded only if an extreme direction

(v^k, w^k) of this cone satisfies $\langle z^k, v^k \rangle - \langle b, w^k \rangle > 0$. Then the affine function (9.87) satisfies $l(z^k) > 0$ while for any $y \in Q(D)$, SP*(y) (hence SP(y)) must have a finite optimal value, and consequently, $\langle y, v^k \rangle - \langle b, w^k \rangle \leq 0$, i.e., $l(y) \leq 0$. $\qquad \square$

The convergence of an OA method for (*MRP*) based on the above proposition follows from general theorems established in Sect. 6.3.

Remark 9.7 In the above approach we only used the constancy of $h(x)$ on every manifold $\{x| \; Qx = y\}$. In fact, since $\varphi(y') \geq \varphi(y)$ for all $y' \geq y$, problem (*MRP*) can also be written as

$$\min\{\psi(y)| \; x \in D, \; Qx \leq y, \; \varphi(y) \leq 0\}. \qquad (9.88)$$

Therefore, SP(z^k) could be replaced by

$$\max\{\langle z^k, v \rangle - \langle b, w \rangle| \; Q^T v - A^T w \geq 0, \; v \geq 0, \; w \geq 0\}.$$

9.7.3   Branch and Bound

We describe a conical branch and bound procedure. A simplicial and a rectangular algorithm for the case when $\varphi(y)$ is separable can be developed similarly.

From assumptions (a), (b) we have that

$$\begin{vmatrix} 0 \in Q(D), & \varphi(0) > 0, \\ \psi(0) < \min\{\psi(y)| \; y \in Q(D), \; \varphi(y) \leq 0\}. \end{vmatrix} \qquad (9.89)$$

A conical algorithm for solving (9.83) starts with an initial cone M_0 contained in $Q(X)$ and containing the feasible set $Q(D) \cap \{y| \; \varphi(y) \leq 0\}$.

Given any cone $M \subset Q(\mathbb{R}^n)$ of base $[u^1, \ldots, u^k]$ let $\theta_i u^i$ be the point where the i-th edge of M intersects the surface $\varphi(y) = 0$. If U denotes the matrix of columns $u^1, \ldots, u^k$ then $M = \{t \in \mathbb{R}^k| \; Ut \geq 0, t \geq 0\}$ and $\sum_{i=1}^{k} t_i/\theta_i = 1$ is the equation of the hyperplane passing through the intersections of the edges of M with the surface $\varphi(y) = 0$. Denote by $\beta(M)$ and $\bar{t}(M)$ the optimal value and a basic optimal solution of the linear program

$$LP(M) \qquad \min\left\{\langle c, x \rangle| \; x \in D, \; Qx = \sum_{i=1}^{k} t_i u^i, \; \sum_{i=1}^{k} \frac{t_i}{\theta_i} \geq 1, t \geq 0\right\}.$$

Proposition 9.8 $\beta(M)$ *gives a lower bound for* $\psi(y)$ *over the set of feasible points in* M. *If* $\omega(M) = U\bar{t}(M)$ *lies on an edge of* M *then*

$$\beta(M) = \min\{\psi(y)| \; y \in Q(D) \cap M, \varphi(y) \leq 0\}. \qquad (9.90)$$

Proof The first part of the proposition follows from the inclusion $\{y \in Q(D) \cap M | \; \varphi(y) \leq 0\} \subset \{y \in Q(D) \cap M | \; \sum_{i=1}^{k} t_i/\theta_i \geq 1\}$ and the fact that $\min\{\psi(y)| \; y \in Q(D) \cap M, \varphi(y) \leq 0\} = \min\{\langle c, x\rangle | \; x \in D, Qx = y, y \in M, \varphi(y) \leq 0\}$. On the other hand, if $\omega(M)$ lies on an edge of M it must lie on the portion of this edge outside the convex set $\varphi(y) \geq 0$, hence $\varphi(\omega(M)) \leq 0$. Consequently, $\omega(M)$ is feasible to the minimization problem in (9.90) and hence is an optimal solution of it. □

It can be proved that a conical branch and bound procedure using ω-subdivision or ω-bisection (see Sect. 7.1) with the above bounding is guaranteed to converge.

Example 9.10

$$\begin{aligned}
\min \; & 2x_1 + x_2 + 0.5x_3 \\
\text{s.t.} \; & -2x_1 + x_3 - x_4 \leq 2.5 \\
& x_1 - 3x_2 + x_4 \leq 2 \\
& x_1 + x_2 \leq 2 \\
& x_1, \, x_2, \, x_3, \, x_4 \geq 0 \\
& h(x) := -(3x_1 + 6x_2 + 8x_3)^2 - (4x_1 + 5x_2 + x_4)^2 + 154 \leq 0.
\end{aligned}$$

Here the function $h(x)$ is concave and monotonic with respect to the cone $K = \{y| \; c^1 y = 0, \; c^2 y = 0\}$, where $c^1 = (3, 6, 8, 0), c^2 = (4, 5, 0, 1)$ (cf Example 7.2). So Q is the matrix of two rows c^1, c^2 and $Q(D) = \{y = Qx, x \in D\}$ is contained in the cone $M_0 = \{y|y \geq 0\}$ (D denotes the feasible set). Since the optimal solution $0 \in \mathbb{R}^4$ of the underlying linear program satisfies $h(0) > 0$, conditions (9.89) hold, and $M_0 = \text{cone}\{e^1, e^2\}$ can be chosen as an initial cone. The conical algorithm terminates after 4 iterations, yielding an optimal solution of $(0, 0, 1.54, 2.0)$ and the optimal value 0.76. Note that branching is performed in $\mathbb{R}^2$ (t-space) rather than in the original space $\mathbb{R}^4$.

9.7.4 Decomposition by Polyhedral Annexation

From $L = \{x| \; Qx = 0\} \subset K \subset C$ we have

$$C^\circ \subset K^\circ \subset L^\perp, \tag{9.91}$$

where $L^\perp$ (the orthogonal complement of L) is the space generated by the rows $c^1, \ldots, c^r$ of Q (assuming rank$Q = r$). For every $\gamma \in \mathbb{R} \cup +\{\infty\}$ let $D_\gamma = \{x \in D| \; \langle c, x\rangle \leq \gamma\}$. By Theorem 7.6 the value $\overline{\gamma}$ is optimal if and only if

$$D_{\overline{\gamma}} \setminus \text{int}C \neq \emptyset, \quad D_{\overline{\gamma}} \subset C. \tag{9.92}$$

Following the PA Method for (*LRC*) (Sect. 8.1), to solve (*MRP*) we construct a nested sequence of polyhedrons $P_1 \supset P_2 \supset \ldots$ together with a nonincreasing

sequence of real numbers $\gamma_1 \geq \gamma_2 \geq \ldots$ such that $D_{\gamma_k} \setminus \text{int} C \neq \emptyset$, $C^\circ \subset P_k \subset L^\perp$ $\forall k$, and eventually $P_l \subset [D_{\gamma_l}]^\circ$ for some l : then $D_{\gamma_l} \subset P_l^\circ \subset C$, hence γ_l is optimal by the criterion (9.92).

To exploit monotonicity [which implies (9.91)], the initial polytope P_1 is taken to be the r-simplex constructed as in Lemma 9.1 (note that $0 \in \text{int} C$ by assumption (a), so this construction is possible). Let $\pi : L^\perp \to \mathbb{R}^r$ be the linear map $y = \sum_{i=1}^r t_i c^i \mapsto \pi(y) = t$. For any set $E \subset \mathbb{R}^r$ denote $\tilde{E} = \pi(E) \subset \mathbb{R}^r$. Then by Lemma 9.1 $\tilde{P}_1$ is the simplex in $\mathbb{R}^r$ defined by the inequalities

$$\sum_{i=1}^r t_i \langle c^i, c^j \rangle \leq \frac{1}{\alpha_j} \quad j = 0, 1, \ldots, r$$

where $c^0 = -\sum_{i=1}^r c^i$, $\alpha_j = \sup\{\alpha | \alpha c^j \in C\}$. In the algorithm below, when a better feasible solution x^k than the incumbent has been found, we can compute a feasible point $\hat{x}^k$ at least as good as x^k and lying on an edge of D (see Sect. 5.7). This can be done by carrying out a few steps of the simplex algorithm. We shall refer to $\hat{x}^k$ as a point obtained by local improvement from x^k.

PA Algorithm for (*MRP*)

Initialization. Choose a vertex x^0 of D such that $h(x^0) > 0$ and set $D \leftarrow D - x^0$, $C \leftarrow C - x^0$. Let $\bar{x}^1 =$ best feasible solution available, $\gamma_1 = \langle c, \bar{x}^1 \rangle$ ($\bar{x}^1 = \emptyset, \gamma_k = +\infty$ if no feasible solution is available). Let P_1 be the simplex constructed as in Lemma 9.1, V_1 the vertex set of $\tilde{P}_1$. Set $k = 1$.

Step 1.   For every $t = (t_1, \ldots, t_r)^T \in V_k$ solve the linear program

$$\max \left\{ \sum_{i=1}^r t_i \langle c^i, x \rangle \,|\, x \in D, \langle c, x \rangle \leq \gamma_k \right\}$$

to obtain its optimal value $\mu(t)$. (Only the new $t \in V_k$ should be considered if this step is entered from Step 3). If $\mu(t) \leq 1$ $\forall t \in V_k$, then terminate: if $\gamma_k < \infty$, $\bar{x}^k$ is an optimal solution; otherwise (MRP) is infeasible.

Step 2.   If $\mu(t^k) > 1$ for some $t^k \in V_k$, then let x^k be a basic optimal solution of the corresponding linear program. If $x^k \notin C$, then let $\hat{x}^k$ be the solution obtained by local improvement from x^k, reset $\bar{x}^k = \hat{x}^k$, $\gamma_k = \langle c, \hat{x}^k \rangle$, and return to Step 1.

Step 3.   If $x^k \in C$ then set $\bar{x}^{k+1} = \bar{x}^k$, $\gamma_{k+1} = \gamma_k$, compute $\theta_k = \sup\{\theta | \theta x^k \in C\}$ and define

$$\tilde{P}_{k+1} = \tilde{P}_k \cap \left\{ t | \sum_{i=1}^r t_i \langle x^k, c^i \rangle \leq \frac{1}{\theta_k} \right\} .$$

From V_k derive the vertex set V_{k+1} of $\tilde{P}_{k+1}$. Set $k \leftarrow k + 1$ and return to Step 1.

Proposition 9.9 *The above PA Algorithm for* (*MRP*) *is finite.*

Proof Immediate, since this is a mere specialization of the PA Algorithm for (*LRCP*) in Sect. 6.4 to (*MRP*). □

Remark 9.8 If $K = \{x | Qx \geq 0\}$ then P_1 can be constructed as in Lemma 9.2. The above algorithm assumes regularity of the problem (assumption (b)). Without this assumption, we can replace C by $C_\varepsilon = \{x \in X | h(x) \leq \varepsilon\}$. Applying the above algorithm with $C \leftarrow C_\varepsilon$ will yield an ε-approximate optimal solution (cf Sect. 7.3), i.e., a point x^ε satisfying $x^\varepsilon \in D, h(x^\varepsilon) \leq \varepsilon$, and $\langle c, x^\varepsilon \rangle \leq \min\{\langle c, x \rangle | x \in D, h(x) \leq 0\}$.

9.8 Network Constraints

Consider a special concave programming problem of the form:

$$(SCP) \quad \min g(y_1, \ldots, y_r) + \langle c, x \rangle$$

$$\text{s.t. } y \in Y \tag{9.93}$$

$$Qx = y, \ Bx = d, \ x \geq 0. \tag{9.94}$$

where $g : \mathbb{R}^r_+ \to \mathbb{R}^+$ is a concave function, Y a polytope in $\mathbb{R}^r_+$, $c, x \in \mathbb{R}^n, d \in \mathbb{R}^m, Q$ an $r \times n$ matrix of rank r, and B an $m \times n$ matrix. In an economic interpretation, $y = (y_1, \ldots, y_r)^T$ may denote a production program to be chosen from a set Y of technologically feasible production programs, x a distribution–transportation program to be determined so as to meet the requirements (9.94). The problem is then to find a production–distribution–transportation program, satisfying conditions (9.94), with a minimum cost. A variant of this problem is the classical *concave production–transportation problem*

$$(PTP) \quad \left| \begin{array}{l} \min g(y) + \sum_{i=1}^{r} \sum_{j=1}^{m} c_{ij} x_{ij} \\[2ex] \text{s.t. } \sum_{j=1}^{m} x_{ij} = y_i \quad i = 1, \ldots, r \\[2ex] \sum_{i=1}^{r} x_{ij} = d_j \quad j = 1, \ldots, m \\[1ex] x_{ij} \geq 0, \quad i = 1, \ldots, r, \ j = 1, \ldots, m \end{array} \right.$$

where y_i is the production level to be determined for factory i and x_{ij} the amount to be sent from factory i to warehouse j in order to meet the demands $d_1, \ldots, d_m$ of the warehouses. Another variant of *SCP* is the *minimum concave cost network flow problem* which can be formulated as

$$(MCCNFP) \qquad \left| \begin{array}{l} \min \sum_{i=1}^{r} g_i(x_i) + \sum_{i=r+1}^{n} c_i x_i \\ \text{s.t. } Ax = d, \ x \geq 0. \end{array} \right.$$

where A is the node-arc incidence matrix of a given directed graph with n arcs, d the vector of node demands (with supplies interpreted as negative demands), and x_i the flow value on arc i. Node j with $d_j > 0$ are the *sinks*, nodes j with $d_j < 0$ are the *sources*, and it is assumed that $\sum_j d_j = 0$ (balance condition). The cost of shipping t units along arc i is a nonnegative concave nondecreasing function $g_i(t)$ for $i = 1, \ldots, r$, and a linear function $c_i t$ (with $c_i \geq 0$) for $i = r + 1, \ldots, n$. To cast $(MCCNFP)$ into the form (SCP) it suffices to rewrite it as

$$\left| \begin{array}{l} \min \sum_{i=1}^{r} g_i(y_i) + \sum_{i=r+1}^{n} c_i x_i \\ \text{s.t. } x_i = y_i \ (i = 1, \ldots, r), \ Ax = d, \ x \geq 0. \end{array} \right. \qquad (9.95)$$

Obviously, Problem (SCP) is equivalent to

$$\min g(y) + t$$
$$\text{s.t. } \langle c, x \rangle \leq t, \ Qx = y, \ Bx = d, \ x \geq 0, \ y \in Y. \qquad (9.96)$$

Since this is a concave minimization problem with few nonconvex variables, it can be handled by the decomposition methods discussed in the previous Sects. 7.2 and 7.3.

Note that in both (PTP) and $(MCCNFP)$ the constraints (9.94) have a nice structure such that when y is fixed the problem reduces to a special linear transportation problem (for (PTP)) or a linear cost flow problem (for $(MCCNFP)$), which can be solved by very efficient specialized algorithms. Therefore, to take full advantage of the structure of (9.94) it is important to preserve this structure under the decomposition.

9.8.1 Outer Approximation

This method is initialized from a polytope $P_1 = Y \times [\alpha, \beta]$, where α, β are, respectively, a lower bound and an upper bound of $\langle c, x \rangle$ over the polytope $Bx = d, x \geq 0$. At iteration k, we have a polytope $P_k \subset P_1$ containing all (y, t) for which there exists $x \geq 0$ satisfying (9.96). Let

$$(y^k, t^k) \in \operatorname{argmin}\{t + g(y) | \ (y, t) \in P_k\}.$$

Following Remark 7.1 to check whether (y^k, t^k) is feasible to (9.96) we solve the pair of dual programs

$$\min\{\langle c, x \rangle | \ Qx = y^k, \ Bx = d, \ x \geq 0\} \qquad (9.97)$$

$$\max\{\langle y^k, u \rangle + \langle d, v \rangle \mid Q^T u + B^T v \le c\} \tag{9.98}$$

obtaining an optimal solution x^k of (9.97) and an optimal solution (u^k, v^k) of the
dual (9.98). If $\langle c, x^k \rangle \le t^k$ then (x^k, y^k) solves (SCP). Otherwise, we define

$$P_{k+1} = P_k \cap \{(y, t) \mid \langle y, u^k \rangle + \langle d, v^k \rangle \le t\} \tag{9.99}$$

and repeat the procedure with $k \leftarrow k + 1$.

   In the particular case of (PTP) the subproblem (9.97) is a linear transportation
problem obtained from (PTP) by fixing $y = y^k$, while in the case of (MCCNFP)
the subproblem (9.97) is a linear min-cost network flow problem obtained from
(MCCNFP) by fixing $x_i = y_i^k$ for $i = 1, \ldots, r$ (or equivalently, by deleting the arcs
$i = 1, \ldots, r$ and accordingly modifying the demands of the nodes which are their
endpoints).

9.8.2   Branch and Bound Method

Assume that in problem (SCP) the function $g(y)$ is concave separable, i.e., $g(y) = \sum_{i=1}^{r} g_i(y_i)$. For each rectangle $M = [a, b]$ let

$$l_M(y) = \sum_{i=1}^{R} l_{i,M}(y_i)$$

where $l_{i,M}(t)$ is the affine function that agrees with $g_i(t)$ at the endpoints a_i, b_i of the
segment $[a_i, b_i]$, i.e.,

$$l_{i,M}(t) = g_i(a_i) + \frac{g_i(b_i) - g_i(a_i)}{b_i - a_i}(t - a_i). \tag{9.100}$$

(see Proposition 7.4). Then a lower bound $\beta(M)$ for the objective function in (SCP)
over M is provided by the optimal value in the linear program

$$(LP(M)) \qquad \min \left\{ \sum_{i=1}^{r} l_{i,M}(y_i) + \langle c, x \rangle \mid Qx = y, Ax = d, x \ge 0 \right\}.$$

Using this lower bound and a rectangular ω-subdivision rule, a convergent rectan-
gular algorithm has been developed in Horst and Tuy (1996) which is a refined
version of an earlier algorithm by Soland (1974) and should be able to handle
even large scale (PTP) provided r is relatively small. Also note that if the data
are all integral, then it is well known that a basic optimal solution of the linear
transportation problem $LP(M)$ always exists with integral values. Consequently, in
this case, starting from an initial rectangle with integral vertices the algorithm will
generate only subrectangles with integral vertices, and so it will necessarily be finite.

In practice, aside from production and transportation costs there may also be shortage/holding costs due to a failure to meet the demands exactly. Then we have the following problem (cf Example 5.1) which includes a formulation of the *stochastic transportation–location problem* as a special case:

$$(GPTP) \quad \begin{vmatrix} \min \ \langle c, x \rangle + \sum_{i=1}^{r} g_i(y_i) + \sum_{j=1}^{m} h_j(z_j) \\[2mm] \text{s.t.} \ \sum_{j=1}^{m} x_{ij} = y_i \quad (i = 1, \ldots, r) \\[2mm] \phantom{\text{s.t.} \ } \sum_{i=1}^{r} x_{ij} = z_j \quad (j = 1, \ldots, m) \\[2mm] \phantom{\text{s.t.} \ } 0 \le y_i \le s_i \quad \forall i, \\[1mm] \phantom{\text{s.t.} \ } x_{ij} \ge 0 \quad \forall i, j. \end{vmatrix}$$

where $g_i(y_i)$ is the cost of producing y_i units at factory i and $h_j(z_j)$ is the shortage/holding cost to be incurred if the warehouse j receives $z_j \ne d_j$ (demand of warehouse j). It is assumed that g_i is a concave nondecreasing function satisfying $g_i(0) = 0 \le g_i(0_+)$ (so possibly $g_i(0_+) > 0$ as, e.g., for a fixed charge), while $h_j : [0, s] \to \mathbb{R}_+$ is a convex function ($s = \sum_{i=1}^{r} s_i$), such that $h_j(d_j) = 0$, and the right derivative of h_j at 0 has a finite absolute value ξ_j. The objective function is then a d.c. function, which is a new source of difficulty. However, since the problem becomes convex when y is fixed, it can be handled by a branch and bound algorithm in which branching is performed upon y (Holmberg and Tuy 1995).

For any rectangle $[a, b] \subset Y := \{y \in \mathbb{R}_+^r \mid 0 \le y_i \le s_i \ i = 1, \ldots, r\}$ let as previously $l_{i,M}(t)$ be defined by (9.100) and $l_M(y) = \sum_{i=1}^{r} l_{i,M}(y_i)$. Then a lower bound of the objective function over the feasible points in M is given by the optimal value $\beta(M)$ in the convex program

$$CP(M) \quad \begin{vmatrix} \min \ \langle c, x \rangle + l_M(y) + \sum_{j=1}^{m} h_j(z_j) \\[2mm] \text{s.t.} \ \sum_{j=1}^{m} x_{ij} = y_i \quad (i = 1, \ldots, r) \\[2mm] \phantom{\text{s.t.} \ } \sum_{i=1}^{r} x_{ij} = z_j \quad (j = 1, \ldots, m) \\[2mm] \phantom{\text{s.t.} \ } a_i \le y_i \le b_i \quad \forall i, \\[1mm] \phantom{\text{s.t.} \ } x_{ij} \ge 0 \quad \forall i, j. \end{vmatrix}$$

Note that if an optimal solution $(\bar{x}(M), \bar{y}(M), \bar{z}(M))$ of this convex program satisfies $l_M(\bar{y}(M)) = g(\bar{y}(M))$ then $\beta(M)$ equals the minimum of the objective function over the feasible solutions in M. When y is fixed, $(GPTP)$ is a convex program in x, z. If $\varphi(y)$ denotes the optimal value of this program then the problem amounts to minimizing $\varphi(y)$ over all feasible (x, y, z).

Fig. 9.1 Consequence of concavity of $g_1(t)$
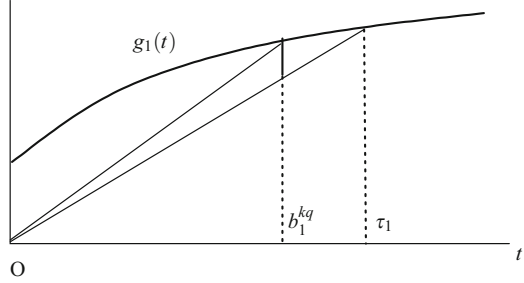
BB Algorithm for (*GPTP*)

Initialization. Start with a rectangle $M_1 \subset Y$. Let y^1 be the best feasible solution available, INV $= \varphi(y^1)$. Set $\mathscr{P}_1 = \mathscr{S}_1 = \{M_1\}, k = 1$.

Step 1. For each $M \in \mathscr{R}_1$ solve the associated convex program CP(M) obtaining its optimal value $\beta(M)$ and optimal solution $(\bar{x}(M), \bar{y}(M), \bar{z}(M))$.

Step 2. Update INV and y^k.

Step 3. Delete all $M \in \mathscr{S}_k$ such that $\beta(M) \geq INV - \varepsilon$. Let $\mathscr{R}_k$ be the collection of remaining rectangles.

Step 4. If $\mathscr{R}_k = \emptyset$, then terminate: y^k is a global ε-optimal solution of (*GPTP*).

Step 5. Select $M_k \in \arg\min\{\beta(M)| M \in \mathscr{R}_k\}$. Let $\bar{y}^k = \bar{y}(M_k)$,

$$i_k \in \arg\max_i\{g_i(\bar{y}^k) - l_{i,M_k}(\bar{y}^k)\}. \tag{9.101}$$

If $g_{i_k}(\bar{y}^k_{i_k}) - l_{i_k,M_k}(\bar{y}^k) = 0$, then terminate: $\bar{y}^k$ is a global optimal solution.

Step 6. Divide M_k via $(\bar{y}^k, i_k)$ (see Sect. 5.6). Let $\mathscr{P}_{k+1}$ be the partition of M_k and $\mathscr{S}_{k+1} = \mathscr{R}_k \setminus (\{M_k\}) \cup \mathscr{P}_{k+1}$. Set $k \leftarrow k + 1$ and return to Step 1.

To establish the convergence of this algorithm, recall that $\bar{y}^k = \bar{y}(M_k)$. Similarly denote $\bar{x}^k = \bar{x}(M_k), \bar{z}^k = \bar{z}(M_k)$, so $(\bar{x}^k, \bar{y}^k, \bar{z}^k)$ is an optimal solution of $CP(M_k)$. Suppose the algorithm is infinite and let $\bar{y}$ be any accumulation point of $\{\bar{y}^k\}$, say $\bar{y} = \lim_{q\to\infty} \bar{y}^{k_q}$. Without loss of generality we may assume $\bar{x}^{k_q} \to \bar{x}, \bar{z}^{k_q} \to \bar{z}$ and also $i_{k_q} = 1 \, \forall q$, so that

$$1 \in \arg\max_i\{g_i(\bar{y}^{k_q}) - l_{i,M_{k_q}}(\bar{y}^{k_q})\}. \tag{9.102}$$

Clearly if for some q we have $\bar{y}^{k_q}_1 = 0$, then $g_i(\bar{y}^{k_q}) = l_{1,M_{k_q}}(\bar{y}^{k_q})$, and the algorithm terminates in Step 5. Therefore, since the algorithm is infinite, we must have $\bar{y}^{k_q}_1 > 0 \, \forall q$.

Lemma 9.9 *If $g_1(0_+) > 0$ then*

$$\bar{y}_1 = \lim_{q\to\infty} \bar{y}^{k_q}_1 > 0. \tag{9.103}$$

In other words, $\bar{y}_1$ is always a continuity point of $g_1(t)$.

Proof Let $M_{k_q,1} = [a_1^q, b_1^q]$, so that $\overline{y}_1^{k_q} \in [a_1^q, b_1^q]$ and $[a_1^q, b_1^q], q = 1, 2, \ldots$ is a sequence of nested intervals. If $0 \notin [a_1^{q_0}, b_1^{q_0}]$ for some q_0, then $0 \notin [a_1^q, b_1^q] \ \forall q \geq q_0$, and $\overline{y}_1^{k_q} \geq a_1^{q_0} > 0 \ \forall q \geq q_0$, hence $\overline{y}_1 \geq a_1^{q_0} > 0$, i.e., (9.103) holds. Thus, it suffices to consider the case

$$a_1^q = 0, \ \overline{y}_1^{k_q} > 0 \quad \forall q. \tag{9.104}$$

Recall that ξ_j denotes the absolute value of the right derivative of $h_j(t)$ at 0. Let

$$\pi = \max_j \xi_j - \min_j c_{1j}. \tag{9.105}$$

Since $\sum_{j=1}^n \overline{x}_{1j}^{k_q} = \overline{y}_1^{k_q} > 0$, there exists j_0 such that $\overline{x}_{1j_0}^{k_q} > 0$. Let $(\tilde{x}^{k_q}, \tilde{y}^{k_q}, \tilde{z}^{k_q})$ be a feasible solution to $CP(M)$ obtained from $(\overline{x}^{k_q}, \overline{y}^{k_q}, \overline{z}^{k_q})$ by subtracting a very small positive amount $\eta < \overline{x}_{1j_0}^{k_q}$ from each of the components $\overline{x}_{1j_0}^{k_q}, \overline{y}_1^{k_q}, \overline{z}_{j_0}^{k_q}$, and letting unchanged all the other components. Then the production cost at factory 1 decreases by $l_{1,M_{k_q}}(\overline{y}_1^{k_q}) - l_{1,M_{k_q}}(\tilde{y}_1^{k_q}) = g_1(b_1^{k_q})\eta/b_1^{k_q} > 0$, the transportation cost in the arc $(1, j_0)$ decreases by $c_{1j_0}\eta$, while the penalty incurred at warehouse j_0 either decreases (if $\overline{z}_{j_0}^{k_q} > d_j$), or increases by $h_{j_0}(\overline{z}_{j_0}^{k_q} - \eta) - h_{j_0}(\overline{z}_{j_0}^{k_q}) \leq \xi_{j_0}\eta$. Thus the total cost decreases by at least

$$\delta = \left[\frac{g_1(b_1^{k_q})}{b_1^{k_q}} + c_{1j_0} - \xi_{j_0} \right] \eta. \tag{9.106}$$

If $\pi \leq 0$ then $c_{1j_0} - \xi_{j_0} \geq 0$, hence $\delta > 0$ and $(\tilde{x}^{k_q}, \tilde{y}^{k_q}, \tilde{z}^{k_q})$ would be a better feasible solution than $(\overline{x}^{k_q}, \overline{y}^{k_q}, \overline{z}^{k_q})$, contradicting the optimality of the latter for $CP(M_{k_q})$. Therefore, $\pi > 0$. Now suppose $g_1(0_+) > 0$. Since $g_1(t)/t \to +\infty$ as $t \to 0_+$ (in view of the fact $g_1(0_+) > 0$), there exists $\tau_1 > 0$ satisfying

$$\frac{g_1(\tau_1)}{\tau_1} > \pi.$$

Observe that since $M_{k_q,1} = [0, b_1^{k_q}]$ is divided via $\overline{y}_1^{k_q}$, we must have $[0, b_1^{k'_q}] \subset [0, \overline{y}_1^{k_q}]$ for all $q' > q$, while $[0, \overline{y}_1^{k_q}] \subset [0, b_1^{k_q}]$ for all q. With this in mind we will show that

$$b_1^{k_q} \geq \tau_1 \quad \forall q. \tag{9.107}$$

By the above observation this will imply that $\overline{y}_1^{k_q} \geq \tau_1 \forall q$, thereby completing the proof. Suppose (9.107) does not hold, i.e., $b_1^{k_q} < \tau_1$ for some q. Then $\overline{y}_1^{k_q} \leq b_1^{k_q} < \tau_1$, and since $g_1(t)$ is concave it is easily seen by Fig. 9.1 that $g_1(b_1^{k_q}) \geq b_1^{k_q} g_1(\tau_1)/\tau_1$, i.e.,

$$\frac{g_1(b_1^{k_q})}{b_1^{k_q}} \geq \frac{g_1(\tau_1)}{\tau_1}.$$

This, together with (9.106) implies that

$$\delta \geq \left(\frac{g_1(\tau_1)}{\tau_1} - \pi \right) \bar{x}_{1j_0}^{k_q} > 0$$

hence $(\tilde{x}^{k_q}, \tilde{y}^{k_q}, \tilde{z}^{k_q})$ is a better feasible solution than $(\bar{x}^{k_q}, \bar{y}^{k_q}, \bar{z}^{k_q})$. This contradiction proves (9.107). $\square$

Theorem 9.3 *The above algorithm for (GPTP) can be infinite only if $\varepsilon = 0$ and in this case it generates an infinite sequence $\bar{y}^k = \bar{y}(M_k), k = 1, 2, \dots$ every accumulation point of which is an optimal solution.*

Proof For $\varepsilon = 0$ let $\bar{y} = \lim_{q \to \infty} \bar{y}^{k_q}$ be an accumulation point of $\{\bar{y}^{k_q}\}$ with, as previously, $i_{k_q} = 1 \ \forall q$, and $M_{k_q,1} = [a_1^q, b_1^q]$. Since $\{a_1^q\}$ is nondecreasing, $\{b_1^q\}$ is nonincreasing, we have $a_1^q \to a_1^*, b_1^q \to b_1^*$. For each q either of the following situations occurs:

$$a_1^q < a^{q+1} < b_1^{q+1} < \bar{y}_1^{k_q} < b_1^q \tag{9.108}$$

$$a_1^q < \bar{y}_1^{k_q} < a_1^{q+1} < b_1^{q+1} < b_1^q \tag{9.109}$$

If (9.108) occurs for infinitely many q then $\lim_{q \to \infty} \bar{y}_1^{k_q} = \lim_{q \to \infty} b_1^q = b_1^*$. If (9.109) occurs for infinitely many q then $\lim_{q \to \infty} \bar{y}_1^{k_q} = \lim_{q \to \infty} a_1^q = a_1^*$. Thus, $\bar{y}_1 \in \{a_1^*, b_1^*\}$. By Lemma 9.9, the concave function $g_1(t)$ is continuous at $\bar{y}_1$. Suppose, for instance, that $\bar{y}_1 = a_1^*$ (the other case is similar). Since $\bar{y}_1^{k_q} \to a_1^*$ and $a_1^q \to a_1^*$, it follows from the continuity of $g_1(t)$ at a_1^* that $g_1(\bar{y}_1^{k_q}) - l_{1,M_{k_q}}(\bar{y}_1^{k_q}) \to 0$ and hence, by (9.102),

$$g_i(\bar{y}_i^{k_q}) - l_{i,M_{k_q}}(\bar{y}_i^{k_q}) \to 0 \quad i = 1, \dots, r$$

as $q \to \infty$. This implies that $g(\bar{y}^{k_q}) - l_{M_{k_q}}(\bar{y}^{k_q}) \to 0$ and hence, that

$$\beta(M_{k_q}) \to \langle c, \bar{x} \rangle + g(\bar{y}) + h(\bar{z})$$

as $q \to \infty$. Since $\beta(M_{k_q}) \leq \gamma :=$ optimal value of (GPTP) it follows that $\langle c, \bar{x} \rangle + g(\bar{y}) + h(\bar{z})$ is the sought minimum of (GPTP). $\square$

Computational experiments have shown that the above algorithm for (GPTP) can solve problems with up to $r = 100$ supply points and $n = 500$ demand points in reasonable time on a Sun SPARCstation SLC (Holmberg and Tuy 1995). To apply this algorithm to the case when $h_j(z_j) = +\infty$ for $z_j \neq d_j$ (e.g., the fixed charge problem) it suffices to redefine $h_j(z_j) = \xi_j(|d_j - x_j|)$ where ξ_j is finite but sufficiently large.

9.9   Some Polynomial-Time Solvable Problems

As mentioned previously, even quadratic programming with one negative eigenvalue is NP-hard (Pardalos and Vavasis 1991). There are, nevertheless, genuine concave programs which are polynomial-time solvable. In this section we discuss this topic, to illustrate the efficiency of the parametric approach for certain classes of low rank nonconvex problems.

9.9.1   The Special Concave Program CPL(1)

Consider the special concave program under linear constraints:

$$(CPL(1)) \quad \left| \begin{array}{l} \min\ g(y) + \langle c, x \rangle \\ \text{s.t.}\ \sum_{j=1}^{n} x_j = y, \quad 0 \le x_j \le d_j\ \forall j. \end{array} \right.$$

where $g(y)$ is a concave function, $c \in \mathbb{R}^n$ and $d \in \mathbb{R}^n_+$. Let $s = \sum_{j=1}^{n} d_j$. To solve this concave program by the parametric right-hand side method presented in Sect. 8.5, we first compute the breakpoints $y_0 = 0 \le y_1 \le \ldots \le y_N = s$ of the optimal value function of the associated parametric linear program

$$(LP(y)) \qquad \min\{\langle c, x \rangle \mid \sum_{j=1}^{n} x_j = y,\ 0 \le x_j \le d_j\ \forall j\}, \quad 0 \le y \le s. \qquad .$$

If $\bar{x}^i$ is a basic optimal solution of $LP(y_i)$ then, by Lemma 9.4, an optimal solution of CPL(1) is given by $(\bar{x}^{i*}, y_{i_*})$ where

$$i_* \in \arg\min\{g(y_i) + c\bar{x}^i \mid i = 0, 1, \ldots, N\}. \tag{9.110}$$

Thus, all is reduced to solving $LP(y)$ parametrically in $y \in [0, s]$. But this is very easy because $LP(y)$ is a continuous knapsack problem. In fact, by reindexing we can assume that the c_j are ordered by increasing values:

$$c_1\ \le\ c_2\ \le \ldots \le\ c_n.$$

Let k be an index such that $\sum_{j=1}^{k-1} d_j \le y \le \sum_{j=1}^{k} d_j$.

Lemma 9.10 *An optimal solution of $LP(y)$ is $\bar{x}$ with*

$$\left| \begin{array}{ll} \bar{x}_j = d_j & \forall j = 1, \ldots, k-1 \\ \bar{x}_k = y - \sum_{j=1}^{k-1} d_j, & \bar{x}_j = 0 \quad \forall j = k+1, \ldots, n. \end{array} \right.$$

Proof This well-known result (see, e.g., Dantzig 1963) is rather intuitive because for any feasible solution x of LP(y), by noting that $y = \sum_{j=1}^{n} x_j$ we can write

$$\langle c, \bar{x} \rangle = \sum_{j=1}^{k-1} c_j d_j + c_k \sum_{j=1}^{n} x_j - c_k \sum_{j=1}^{k-1} d_j$$

$$= \sum_{j=1}^{k-1} c_j d_j - c_k \sum_{j=1}^{k-1} (d_j - x_j) + c_k \sum_{j=k}^{n} x_j$$

$$\leq \sum_{j=1}^{k-1} c_j d_j - \sum_{j=1}^{k-1} c_j (d_j - x_j) + c_k \sum_{j=k}^{n} x_j \leq \langle c, x \rangle. \square$$

Proposition 9.10

(i) *The breakpoints of the optimal value function of LP(y), $0 \leq y \leq s$, are among the points:*

$$y_0 = 0, \ y_k = y_{k-1} + d_k \quad k = 1, \ldots, n$$

(ii) *For every $k = 0, 1, \ldots, n$ a basic optimal solution of TP(y_k) is the vector $\bar{x}^k$ such that*

$$\bar{x}_j^k = d_j, \ j = 1, \ldots, k; \quad \bar{x}_j^k = 0, \ j = k+1, \ldots n. \qquad (9.111)$$

Proof From Lemma 9.10, for any $y \in [y_k, y_{k+1}]$ a basic optimal solution of LP(y) is

$$x^y = \bar{x}^k + \theta(\bar{x}^{k+1} - \bar{x}^k) \text{ with } \theta = \frac{y - y_k}{d_{k+1}}.$$

Hence each segment $[y_k, y_{k+1}]$ is a linear piece of the optimal value function $\varphi(y)$ of LP(y), which proves (i). Part (ii) is immediate. $\square$

As a result, the concave program $CPL(1)$ can be solved by the following:

Parametric Algorithm for CPL(1)
Initialization. Order the coefficients c_j by increasing values:

$$c_1 \leq c_2 \leq \ldots \leq c_n.$$

Step 2.   For $k = 0, 1, \ldots, n$ compute

$$f_k = g \left(\sum_{j=1}^{k} d_j \right) + \sum_{j=1}^{k} c_j d_j.$$

Step 3.   Find $f_{k^*} = \min\{f_0, f_1, \ldots, f_n\}$. Then $\bar{x}^{k^*}$ is an optimal solution.

Proposition 9.11 *The above algorithm requires $O(n \log_2 n)$ elementary operations and n evaluations of $g(\cdot)$*

Proof The complexity dominant part of the algorithm is Step 1 which requires $O(n \log_2 n)$ operations, using, e.g., the procedure Heapsort (Ahuja et al. 1993). Also an evaluation of $g(\cdot)$ is needed for each number f_k.   □

Thus, assuming the values of the nonlinear function $g(\cdot)$ provided by an oracle the above algorithm for *CPL*(1) runs in *strongly polynomial time*.

9.10   Applications

9.10.1   Problem PTP(2)

For convenience we shall refer to problem (*PTP*) with r factories as PTP(r). So *PTP*(2) is the problem

$$PTP(2) \quad \left| \begin{array}{l} \min \; g(y) + \displaystyle\sum_{i=1}^{2} \sum_{j=1}^{m} c_{ij} x_{ij} \\[2ex] \text{s.t.} \; \displaystyle\sum_{j=1}^{m} x_{1j} = y \\[2ex] \quad x_{1j} + x_{2j} = d_j \quad j = 1, \ldots, m, \\ \quad 0 \le x_{ij} \le d_j \qquad i = 1, 2, \; j = 1, \ldots, m \end{array} \right.$$

Substituting $d_j - x_{1j}$ for x_{2j} and setting $x_j = x_{1j}, c_j = c_{1j} - c_{2j}$ we can convert this problem into a CPL(1):

$$\min \; g(y) + \langle c, x \rangle$$
$$\text{s.t.} \; \sum_{j=1}^{m} x_j = y, \quad 0 \le x_j \le d_j \; \forall j.$$

Therefore,

Proposition 9.12 *PTP*(2) *can be solved by an algorithm requiring $O(m \log_2 m)$ elementary operations and m evaluations of $g(\cdot)$.*

9.10.2 Problem FP(1)

We shall refer to Problem (*MCCNFP*) with a single source and r nonlinear arc costs as FP(r). So $FP(1)$ is the problem of finding a feasible flow with minimum cost in a given network **G** with m nodes (indexed by $j = 1, \ldots, m$), and n arcs (indexed by $i = 1, \ldots, n$), where node j has a demand d_j (with $d_m < 0, d_j \geq 0$ for all other j, $\sum_{j=1}^{m} d_j = 0$), arc 1 has a concave nonlinear cost function $g_1(t)$ while each arc $i = 2, \ldots, n$ has a linear cost $g_i(t) = c_i t$ (with $c_i \geq 0$). If A_j^+ and A_j^- denote the sets of arcs entering and leaving node j, respectively, then the problem is

$$
FP(1) \quad \left|
\begin{aligned}
&\min\ g_1(x_1) + \sum_{i=2}^{n} c_i x_i \\
&\text{s.t.}\ \sum_{i \in A_j^+} x_i - \sum_{i \in A_j^-} x_i = d_j, \quad j = 1, 2, \ldots, m, \\
&x \geq 0.
\end{aligned}
\right.
$$

Proposition 9.13 $FP(1)$ *can be reduced to a* PTP(2) *by means of* $O(m \log_2 m + n)$ *elementary operations.*

Proof To simplify the language we call the single arc with nonlinear cost the *black* arc. All the other arcs are called *white*, and the unit cost c_i attached to a white arc is called its length. A directed path through white arcs only is called a *white path*; its length is then the sum of the lengths of its arcs. A white path from a node j to a node j' is called *shortest* if its length is smallest among all white paths from j to j'. Denote by Q_{0j} the shortest white path *from the source to sink j*, by Q_{1j} the shortest white path *from the head of the black arc to sink j*, and by P the shortest white path *from the source to the tail of the black arc*. Let c_{0j}, c_{1j}, p be the lengths of these paths, respectively (the length being $+\infty$, or an arbitrarily large positive number, if the path does not exist). Assuming $\{j \mid d_j > 0\} = \{1, \ldots, m_1\}$, let $s = \sum_{j=1}^{m_1} d_j$. Consider now a problem *PTP*(2) defined for two factories F_0, F_1, and m_1 warehouses $W_1, \ldots, W_{m_1}$ with demands $d_1, \ldots, d_{m_1}$, where the cost of producing y units at factory F_0 and $s - y$ units at factory F_1 is $g_1(y) + py$ while the unit transportation cost from F_i to W_j is c_{ij} ($i = 0, 1$), i.e.,

$$
\left|
\begin{aligned}
&\min\ g_1(y) + py + \sum_{i=0}^{1} \sum_{j=1}^{m_1} c_{ij} x_{ij} \\
&\text{s.t.}\ \sum_{j=1}^{m_1} x_{0j} = y \\
&\qquad x_{0j} + x_{1j} = d_j \quad j = 1, \ldots, m_1 \\
&\qquad 0 \leq x_{0j} \leq d_j \quad j = 1, \ldots, m_1
\end{aligned}
\right. \tag{9.112}
$$

(the corresponding reduced network is depicted in Fig. 9.2). Clearly from an optimal solution $[\bar{x}_{ij}]$ of (9.112) we can derive an optimal solution of $FP(1)$ by setting, in $FP(1)$,

$$
x_1 = \sum_{j=1}^{m_1} \bar{x}_{0j}
$$

Fig. 9.2 Reduced network
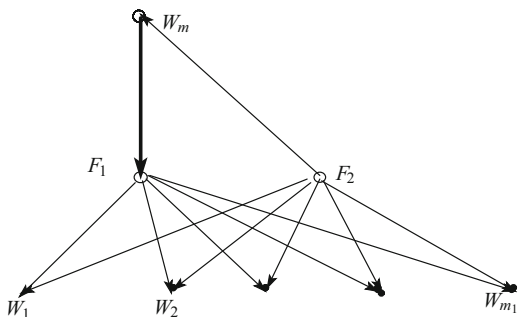
and solving the corresponding linear min-cost network flow problem. Thus, solving $FP(1)$ is reduced to solving the just defined $PTP(2)$. To complete the proof, it remains to observe that the computation of $c_{ij}, i = 0, 1, \ j = 1, \ldots, m_1$, and p needed for the above transformation, requires solving two single-source multiple-sink shortest problems with nonnegative arc costs, hence a time of $O(m \log_2 m + n)$ (using Fredman and Tarjan's (1984) implementation of Dijkstra's algorithm).

It follows from Propositions 9.11 and 9.13 that $FP(1)$ can be solved by a strongly polynomial algorithm requiring $O(m \log_2 m + m \log_2 m + n) = O(m \log_2 m + n)$ elementary operations and m evaluations of the function $g_1(.)$.

Remark 9.9 A polynomial (but not strongly polynomial) algorithm for $FP(1)$ was first given by Guisewite and Pardalos (1992). Later a strongly polynomial algorithm based on solving linear min-cost network flow problems with a parametric objective function was obtained by Klinz and Tuy (1993). The algorithms presented in this section for $PTP(1)$ and $FP(1)$ were developed by Tuy et al. (1993a) and subsequently extended in a series of papers by Tuy et al. (1993b, 1994a, 1995b, 1996a) to prove the strong polynomial solvability of a class of network flow problems including $PTP(r)$ and $FP(r)$ with fixed r.

9.10.3   Strong Polynomial Solvability of PTP(r)

The problem PTP(r) as was formulated at the beginning of Sect. 8.8 is

$$
PTP(r) \quad \left|
\begin{array}{l}
\min g(\sum_{j=1}^{m} x_{1j}), \ldots, \sum_{j=1}^{m} x_{rj}) + \sum_{i=1}^{r} \sum_{j=1}^{m} c_{ij} x_{ij} \\[2ex]
\text{s.t.} \ \sum_{i=1}^{r} x_{ij} = d_j \quad j = 1, \ldots, m \\[2ex]
\quad x_{ij} \geq 0, \quad i = 1, \ldots, r, j = 1, \ldots, m
\end{array}
\right.
$$

where $g(y') \geq g(y)$ whenever $y' \geq y$.

Following the approach in Sect. 8.5 we associate with PTP(r) the parametric problem

$$P(t) \quad \left| \begin{array}{l} \min \sum_{i=1}^{r} t_i \sum_{j=1}^{m} x_{ij} + \sum_{i=1}^{r} \sum_{j=1}^{m} c_{ij} x_{ij} \\[2mm] \text{s.t.} \ \sum_{i=1}^{r} x_{ij} = d_j \quad j = 1, \dots, m \\[2mm] \quad\quad x_{ij} \geq 0, \quad i = 1, \dots, r, j = 1, \dots, m \end{array} \right.$$

where $t \in \mathbb{R}_+^r$. The parametric domain $\mathbb{R}_+^r$ is partitioned into a finite collection $\mathscr{P}$ of polyhedrons ("cells") such that for each cell $\Pi \in \mathscr{P}$ there is a basic solution x^{Π} which is optimal to $P(t)$ for all $t \in \Pi$. Then an optimal solution of PTP(r) is x^{Π^*} where

$$\Pi^* \in \operatorname{argmin}\left\{ g\left(\sum_{j=1}^{m} x_{1j}^{\Pi}, \dots, \sum_{j=1}^{m} x_{rj}^{\Pi} \right) + \sum_{i,j} c_{ij} x_{ij}^{\Pi} \,|\, \Pi \in \mathscr{P} \right\}. \qquad (9.113)$$

Following Tuy (2000b) we now show that the collection $\mathscr{P}$ can be constructed and its cardinality is bounded by a polynomial in m.

Observe that the dual of $P(t)$ is

$$P^*(t) \quad \left| \begin{array}{l} \max \sum_{j=1}^{m} d_j u_j \\[2mm] \text{s.t.} \ u_j \leq t_i + c_{ij}, \quad i = 1, \dots, r, j = 1, \dots, m \end{array} \right.$$

Also for any fixed $t \in r_+^r$ a basic solution of $P(t)$ is a vector x^t such that for every $j = 1, \dots, m$ there is an i_j satisfying

$$x_{ij}^t = \begin{cases} d_j \ i = i_j \\ 0 \ i \neq i_j \end{cases} \qquad (9.114)$$

By the duality of linear programming x^t is a basic optimal solution of $P(t)$ if and only if there exists a feasible solution $u = (u_1, \dots, u_m)$ of $P^*(t)$ satisfying

$$u_j \begin{cases} = t_i + c_{ij} & i = i_j \\ \leq t_i + c_{ij} & i \neq i_j \end{cases}$$

or, alternatively, if and only if for every $j = 1, \dots, m$:

$$i_j \in \operatorname{argmin}_{i=1,\dots,r}\{t_i + c_{ij}\}. \qquad (9.115)$$

Now let I_*^2 be the set of all pairs (i_1, i_2) such that $i_1 < i_2 \in \{1, \dots, r\}$. Define a cell to be a polyhedron $\Pi \subset \mathbb{R}_+^r$ which is the solution set of a linear system formed by taking, for every pair $(i_1, i_2) \in I_*^2$ and every $j = 1, \dots, m$, one of the following inequalities:

$$t_{i_1} + c_{i_1 j} \leq t_{i_2} + c_{i_2 j}, \quad t_{i_1} + c_{i_1 j} \geq t_{i_2} + c_{i_2 j}. \tag{9.116}$$

Then for every $j = 1, \dots, m$ the order of magnitude of the sequence

$$t_i + c_{ij}, \quad i = 1, \dots, r$$

remains unchanged as t varies over a cell Π. Hence the index i_j satisfying $PTP(3)$ remains the same for all $t \in \Pi$; in other words, x^t (basic optimal solution of $P(t)$) equals a constant vector x^Π for all $t \in \Pi$. Let $\mathscr{P}$ be the collection of all cells defined that way. Since every $t \in \mathbb{R}_+^r$ satisfies one of the inequalities (9.116) for every $(i_1, i_2) \in I_*^2$ and every $j = 1, \dots, m$, the collection $\mathscr{P}$ covers all of $\mathbb{R}_+^r$. Let us estimate an upper bound of the number of cells in $\mathscr{P}$.

Observe that for any fixed pair $(i_1, i_2) \in I_*^2$ we have $t_{i_1} + c_{i_1 j} \leq t_{i_2} + c_{i_2 j}$ if and only if $t_{i_1} - t_{i_2} \leq c_{i_2 j} - c_{i_1 j}$. Let us sort the numbers $c_{i_2 j} - c_{i_1 j}, j = 1, \dots, m$, in increasing order

$$c_{i_2 j_1} - c_{i_1 j_1} \leq c_{i_2 j_2} - c_{i_1 j_2} \leq \dots \leq c_{i_2 j_m} - c_{i_1 j_m}. \tag{9.117}$$

and let $\nu_{i_1, i_2}(j)$ be the position of $c_{i_2 j} - c_{i_1 j}$ in this ordered sequence.

Proposition 9.14 *A cell Π is characterized by a map $k_\Pi : I_*^2 \to \{1, \dots, m, m+1\}$ such that Π is the solution set of the linear system*

$$t_{i_1} + c_{i_1 j} \leq t_2 + c_{i_2 j} \quad \forall (i_1, i_2) \in I_*^2 \text{ s.t. } \nu_{i_1, i_2}(j) \geq k_\Pi(i_1, i_2) \tag{9.118}$$

$$t_{i_1} + c_{i_1 j} \geq t_2 + c_{i_2 j} \quad \forall (i_1, i_2) \in I_*^2, \text{ s.t. } \nu_{i_1, i_2}(j) < k_\Pi(i_1, i_2) \tag{9.119}$$

Proof Let $\Pi \subset \mathbb{R}_+^r$ be a cell. For every pair (i_1, i_2) with $i_1 < i_2$ denote by $J_\Pi^{i_1, i_2}$ the set of all $j = 1, \dots, m$ such that the left inequality (9.116) holds for all $t \in \Pi$ and define

$$k_\Pi(i_1, i_2) = \begin{cases} \min\{\nu_{i_1, i_2}(j) \mid j \in J_\Pi^{i_1, i_2}\} & \text{if } J_\Pi^{i_1, i_2} \neq \emptyset \\ m+1 & \text{otherwise} \end{cases}$$

It is easy to see that Π is then the solution set of the system (9.118) and (9.119). Indeed, let $t \in \Pi$. If $\nu_{i_1, i_2}(j) \geq k_\Pi(i_1, i_2)$, then $k_\Pi(i_1, i_2) \neq m+1$, so $k_\Pi(i_1, 2) = \nu_{i_1, i_2}(l)$ for some $l \in J_\Pi^{i_1, i_2}$. Then $t_{i_1} + c_{i_1 l} \leq t_{i_1} + c_{i_2}$, hence $t_{i_1} - t_{i_2} \leq c_{i_2 l} - c_{i_1 l}$ and since the relation $\nu_{i_1, i_2}(j) \geq \nu_{i_1, i_2}(l)$ means that $c_{i_2 j} - c_{i_1 j} \geq c_{i_2 l} - c_{i_1 l}$ it follows that $t_{i_1} - t_{i_2} \leq c_{i_2 j} - c_{i_1 j}$, i.e., $t_{i_1} + c_{i_1 j} \leq t_{i_2} + c_{i_2 j}$. Therefore (9.118) holds. On the other hand, if $\nu_{i_1, i_2}(j) < k_\Pi(i_1, i_2)$ then by definition of a cell $j \notin J_\Pi^{i_1, i_2}$, hence (9.119)

holds, too [since from the definition of a cell, any $t \in \Pi$ must satisfy one of the inequalities (9.116)]. Thus every $t \in \Pi$ is a solution of the system (9.118) and (9.119). Conversely, if t satisfies (9.118) and (9.119) then for every $(i_1, i_2) \in I_*^2$, t satisfies the left inequality (9.116) for $j \in J_\Pi^{i_1, i_2}$ and the right inequality for $j \notin J_\Pi^{i_1, i_2}$, hence $t \in \Pi$. Therefore, each cell Π is determined by a map $k_\Pi : I_*^2 \to \{1, \ldots, m + 1\}$. Furthermore, it is easy seen that $k_\Pi \neq k_{\Pi'}$ for two different cells Π, Π'. Indeed, if $\Pi \neq \Pi'$ then at least for some $(i_1, i_2) \in I_*^2$ and some $j = 1, \ldots, m$, one has $j \in J_\Pi^{i_1, i_2} \setminus J_{\Pi'}^{i_1, i_2}$. Then $k_\Pi(i_1, i_2) \leq v_{i_1, i_2}(j)$ but $k_{\Pi'}(i_1, i_2) > v_{i_1, i_2}(j)$.   □

Corollary 9.4 *The total number of cells is bounded above by* $(m + 1)^{r(r-1)/2}$.

Proof The number of cells does not exceed the number of different maps $k : I_*^2 \to \{1, \ldots, m + 1\}$ and there are $(m + 1)^{r(r-1)/2}$ such maps.   □

To sum up, the proposed algorithm for solving $PTP(r)$ involves the following steps:

(1) Ordering the sequence $c_{i_2 j} - c_{i_1 j}, j = 1, \ldots, m$ for every pair $(i_1, i_2) \in I_*^2$ so as to determine $v_{i_1, i_2}(j), j = 1, \ldots, m, (i_1, i_2) \in I_*^2$.
(2) Computing the vector x^Π for every cell $\Pi \in \mathscr{P}$ ($\mathscr{P}$ is the collection of cells determinedly the maps $k_\Pi : I_*^2 \to \{1, \ldots, m + 1\}$.
(3) Computing the values $f(x^\Pi)$ and select Π^* according to (9.113).
   Steps (1) and (2) require obviously a number of elementary operations bounded by a polynomial in m, while step (3) requires $m^{r(r-1)/2}$ evaluations of $f(x)$.

9.11   Exercises

9.1 Show that a differentiable function $f(x)$ is K-monotonic on an open convex set $X \subset \mathbb{R}^n$ if and only if $-\nabla f(x) \in K^\circ$ $\forall x \in X$.

9.2 Consider the problem $\min\{f(x) \mid x \in D\}$ where D is a polyhedron in $\mathbb{R}^n$ with a vertex at 0, and $f : \mathbb{R}^n \to \mathbb{R}$ is a concave function such that for any real number $\gamma \leq f(0)$ the level set $C_\gamma = \{x \mid f(x) \geq \gamma\}$ has a polar contained in the cone generated by two linearly independent vectors $c^1, c^2 \in \mathbb{R}^n$. Show that this problem reduces to minimizing a concave function of two variables on the projection of D on $\mathbb{R}^2$. Develop an OA method for solving the problem.

9.3 Consider the problem

$$\min\{g(t) + \langle d, y \rangle \mid ta + By \leq c, t \geq 0, y \geq 0\}$$

where $t \in \mathbb{R}, y \in \mathbb{R}^n, a \in \mathbb{R}^m, B \in \mathbb{R}^{m \times n}, c \in \mathbb{R}_+^m$. Develop a parametric algorithm for solving the problem and compare with the polyhedral annexation method.

9.4 Show that a linearly constrained problem of the form

$$\min\{F(l(x))\mid\ x \in D\}$$

where D is a polyhedron in $\mathbb{R}^n$, $l(x)$ is an affine function, and $F(t)$ is an arbitrary quasiconcave real-valued function on an interval $\Delta \supset l(D)$ reduces to at most two linear programs. If $F(.)$ is monotonic on $l(D)$ then the problem is in fact equivalent to even a mere linear program.

9.5 Solve the plant location problem

$$\min \sum_{i=1}^{3} f_i(x_i) + \sum_{i=1}^{3}\sum_{j=1}^{5} d_{ij}y_{ij}$$

$$\text{s.t. } x_1 + x_2 + x_3 = 203$$

$$y_{1j} + y_{2j} + y_{3j} = b_j \quad j = 1, \ldots, 5$$

$$y_{i1} + \ldots + y_{i5} = x_i \quad i = 1, 2, 3$$

$$x_i \geq 0; \quad y_{ij} \geq 0 \quad i = 1, 2, 3; j = 1, \ldots, 5.$$

where $f_i(t) = 0$ if $t = 0$, $f_i(t) = r_i + s_i t$ if $t > 0$

$$r_1 = 1, s_1 = 1.7; \quad r_2 = 88, s_2 = 8.4; \quad r_3 = 39; s_3 = 4.7$$

$$b = (62, 65, 51, 10, 15); \quad d = [d_{ij}] = \begin{bmatrix} 6 & 66 & 68 & 81 & 4 \\ 40 & 20 & 34 & 83 & 27 \\ 90 & 22 & 82 & 17 & 8 \end{bmatrix}$$

9.6 Suppose that details of n different kinds must be manufactured. The production cost of t items of i-th kind is a concave function $f_i(t)$. The number of items of i-th kind required is $a_i > 0$, but one item of i-th kind can be replaced by one item of any j-th kind with $j > i$. Determine the number x_i of items of i-th kind, $i = 1, \ldots, n$, to be manufactured so as to satisfy the demands with minimum cost. By setting $\alpha_i = \sum_{j=1}^{i} a_j$, this problem can be formulated as:

$$\min \sum_{i=1}^{n} f_i(x_i) \quad \text{s.t.}$$

$$x_1 \leq \alpha_1, \ x_1 + x_2 \leq \alpha_2, \ldots, x_1 + \ldots + x_n \leq \alpha_n$$

$$x_1, \ldots, x_n \geq 0$$

(Zukhovitzki et al. 1968). Solve this problem.

(*Hint*: Any extreme point x of the feasible polyhedron can be characterized by the condition: for each $i = 1, \ldots, n$, either $x_i = 0$ or $x_1 + \ldots + x_i = \alpha_i$).

9.7 Develop an algorithm for solving the problem

$$\min\{c^0 x + (c^1 x)^{1/2} \times (c^2 x)^{1/3} \mid \quad x \in D\}$$

where $D \subset \mathbb{R}^n_+$ is a polytope, $c^1, c^2 \in \mathbb{R}^n_+$, $c^0 \in \mathbb{R}^n_-$ and cx denotes the inner product of c and x. Apply the algorithm to a small numerical example with $n = 2$.

9.8 Solve the numerical Example 9.10 (page 317).

9.9 Solve by polyhedral annexation:

$$\min \ (0.5x_1 + 0.75x_2 + 0.75x_3 + 1.25x_4) \quad \text{s.t.}$$

$$0.25x_2 + 0.25x_3 + 0.75x_4 \leq 2$$

$$-0.25x_2 - 0.25x_3 + 0.25x_4 \leq 0$$

$$x_1 - 0.5x_2 + 0.5x_3 + 1.5x_4 \leq 1.5$$

$$h(x) + 4x_1 - x_2 = 0$$

$$x = (x_1, \ldots, x_4) \geq 0$$

where $h(x)$ is the optimal value of the linear program

$$\min \ (-4y_1 + y_2) \quad \text{s.t.}$$

$$y_1 + y_2 \leq 1.5 - 0.5x_2 - 0.5x_3 - 1.5x_4$$

$$y_2 \leq x_1 + x_2; \quad y_1, y_2 \geq 0$$

(Hint: Exploit the monotonicity property of $h(x)$)

9.10 Solve the problem

$$\min \ \sum_{i=1}^{n} \frac{x_i}{y_i + \eta} \quad \text{s.t.}$$

$$\sum_{i=1}^{n} x_i = a, \quad \sum_{i=1}^{n} y_i = b$$

$$0 \leq x_i \leq \gamma, \quad 0 \leq y_i \leq 1 \quad i = 1, \ldots, n$$

where $0 < \eta, 0 < a \leq n\gamma, 0 < b \leq n$.

Hint: Reduce the problem to the concave minimization problem

$$\min \left\{ h(x) \mid \sum_{i=1}^{n} x_i = a, 0 \leq x_i \leq \gamma \right\}$$

where

$$h(x) = \min \left\{ \sum_{i=1}^{n} \frac{x_i}{y_i + \eta} \mid \sum_{i=1}^{n} y_i = b, 0 \leq y_i \leq 1 \quad i = 1, \ldots, n \right\}$$

and use the results on the special concave minimization CPL(1).

9.11 Solve the problem in Example 5.5 (page 142) where the constraints (5.24) are replaced by

$$\sum_{i=1}^{n} x_i = a, \quad \sum_{i=1}^{n} y_i = b.$$