RM-75-61

A CUTTING PLANE ALGORITHM FOR SOLVING BILINEAR PROGRAMS

Hiroshi Konno

December 1975

A Cutting Plane Algorithm for Solving Bilinear Programs

Hiroshi Konno

1. Introduction

Nonconvex programs which have either a nonconvex minimand and/or a nonconvex feasible region have been considered by most mathematical programmers as a hopelessly difficult area of research. There are, however, two exceptions where considerable effort to obtain a global optimum is under way. One is integer linear programming and the other is nonconvex quadratic programming. This paper addresses itself to a special class of nonconvex quadratic program referred to as a 'bilinear program' in the literature. We will propose here a cutting plane algorithm to solve this class of problems. The algorithm is along the lines of [17] and [19] but the major difference is in its exploitation of special structure. Though the algorithm is not guaranteed at this stage to converge to a global optimum, the preliminary results are encouraging.

In Section 2, we analyze the structure of the problem and develop an algorithm to obtain an ε-locally maximum pair of basic feasible solutions. In Section 3, we will generate a cutting plane to eliminate the current pair of ε-locally maximum basic feasible solutions. For these purposes, we extensively use the simplex algorithm. Section 4 gives an illustrative example and the results of numerical experimentations. Some of the important applications of bilinear programming can be found in references [11] and [12].

2. Definitions and a Locally Maximum Pair of Basic Feasible Solutions

The bilinear program is a class of quadratic programs with the following structure:

$$\max\ \phi(x_1,x_2) = c_1^t x_1 + c_2^t x_2 + x_1^t C x_2$$

$$\text{s.t. } A_1 x_1 = b_1 \ , \quad x_1 \geq 0$$

$$A_2 x_2 = b_2 \ , \quad x_2 \geq 0 \ , \tag{2.1}$$

where c_i, $x_i \in R^{n_i}$, $b_i \in R^{m_i}$, $A_i \in R^{m_i \times n_i}$, $i = 1,2$ and $C \in R^{n_1 \times n_2}$. We will call this a bilinear program in 'standard' form.

Note that a bilinear program is a direct extension of the standard linear program: $\max\{c^t x \mid Ax = b,\ x \geq 0\}$, in which we consider c to be linearly constrained variables and maximize $c^t x$ with respect to c and x simultaneously. Let us denote

$$X_i = \{x_i \in R^{n_i} \mid A_i x_i = b_i,\ x_i \geq 0\} \ , \quad i = 1,2 \ . \tag{2.2}$$

<u>Theorem 2.1</u>. If X_i, $i = 1,2$ are non-empty and bounded, then (2.1) has an optimal solution (x_1^*, x_2^*) where x_i^* is a basic feasible solution of the constraint equations defining X_i, $i = 1,2$.

<u>Proof</u>. Let $(\hat{x}_1, \hat{x}_2)$ be an optimal solution, which clearly exists by assumption. Consider a <u>linear</u> program: $\max\{\phi(x_1,\hat{x}_2) \mid x_1 \in X_1\}$, and let x_1^* be its optimal basic solution. Then $\phi(x_1^*,\hat{x}_2) \geq \phi(\hat{x}_1,\hat{x}_2)$ since $\hat{x}_1$ is a feasible solution to the linear program considered above. Next, consider another linear program: $\max\{\phi(x_1^*,x_2) \mid x_2 \in X_2\}$, and let x_2^* be its optimal basic solution. Then by similar arguments as before, we have $\phi(x_1^*,x_2^*) \geq \phi(x_1^*,\hat{x}_2)$. Thus we conclude that $\phi(x_1^*,x_2^*) \geq \phi(\hat{x}_1,\hat{x}_2)$, which implies that (x_1^*,x_2^*) is a <u>basic</u> optimal solution of (2.1). ||

Given a feasible basis B_i of A_i, we will partition A_i as (B_i, N_i) assuming, without loss of generality, that the first m_i columns of A_i are basic. Position x_i correspondingly: $x_i = (x_{iB}, x_{iN})$. Let us introduce here a 'canonical' representation of (2.1) relative to a pair of feasible bases (B_1, B_2). Premultiplying B_i^{-1} to the constraint equation $B_i x_{iB} + N_i x_{iN} = b_i$ and suppressing the basic variables x_{iB}, we get the following system which is totally equivalent to (2.1):

$$\max \bar{\phi}(x_{1N}, x_{2N}) = \bar{c}_{1N}^t x_{1N} + \bar{c}_{2N}^t x_{2N} + x_{1N}^t \bar{C} x_{2N} + \phi(x_1^o, x_2^o)$$

$$\text{s.t. } B_1^{-1} N_1 x_{1N} \leq B_1^{-1} b_1 \quad , \quad x_{1N} \geq 0$$

$$B_2^{-1} N_2 x_{2N} \leq B_2^{-1} b_2 \quad , \quad x_{2N} \geq 0 \quad , \qquad (2.3)$$

where

$$x_i^o \equiv (x_{iB}^o, x_{iN}^o) = (B_i^{-1} b_i, 0) \quad .$$

For future reference, we will introduce the notations

$$\ell_i = n_i - m_i \quad , \quad d_i = \bar{c}_{iN} \in R^{\ell_i} \quad , \quad y_i = x_{iN} \in R^{\ell_i} \quad ,$$

$$F_i = B_i^{-1} N_i \in R^{m_i \times \ell_i} \quad , \quad f_i = B_i^{-1} b_i \in R^{m_i} \quad , \quad i = 1,2$$

$$Q = \bar{C} \in R^{\ell_1 \times \ell_2} \quad , \quad \phi_o = \phi(x_1^o, x_2^o)$$

and rewrite (2.3) as follows:

$$\max \psi(y_1, y_2) = d_1^t y_1 + d_2^t y_2 + y_1^t Q y_2$$

$$\text{s.t. } F_1 y_1 \leq f_1 \quad , \quad y_1 \geq 0$$

$$F_2 y_2 \leq f_2 \quad , \quad y_2 \geq 0 \quad . \qquad (2.4)$$

We will call (2.4) a canonical representation of (2.1) relative to (B_1, B_2) and use standard form (2.1) and canonical form (2.4) interchangeably, whichever is the more convenient for our presentation. To express the dependence of vectors in (2.4) on the pair of feasible bases (B_1, B_2), we will occasionally use the notation $d_1(B_1, B_2)$, etc.

Theorem 2.2. The origin $(y_1,y_2) = (0,0)$ of the canonical system (2.4) is

(i) a Kuhn-Tucker point if $d_i \leq 0$, $i = 1,2$;

(ii) a local maximum if (a) and (b) hold:

(a) $d_i \leq 0$, $i = 1,2$

(b) either $d_{1i} < 0$ or $d_{2j} < 0$ if $q_{ij} < 0$;

(iii) a global optimum $d_i \leq 0$, $i = 1,2$ and $Q \leq 0$.

Proof.

(i) It is straightforward to see that $y_1 = 0$, $y_2 = 0$ together with dual variables $u_1 = 0$, $u_2 = 0$ satisfy the Kuhn-Tucker condition for (2.1).

(ii) Let $y_i \in R^{\ell_i}$, $i = 1,2$ be arbitrary nonnegative vectors. Let $J_i = \{j | q_{ij} < 0\}$ and let ε be positive scalar. Then

$$\psi(\varepsilon y_1, \varepsilon y_2) = \varepsilon d_1^t y_1 + \varepsilon d_2^t y_2 + \varepsilon^2 y_1^t Q y_2 + \phi_o$$

$$\leqq \varepsilon \sum_{j \varepsilon J_1} d_{ij} y_{ij} + \varepsilon \sum_{j \varepsilon J_2} d_{2j} y_{2j} + \varepsilon^2 \sum_{\substack{i \varepsilon J_1 \\ j \varepsilon J_2}}$$

or

$$q_{ij} y_{1i} y_{2j} + \phi_o$$

because $q_{ij} \leq 0$ when $i \not\in J_1$ and $j \not\in J_2$. Obviously, the last expression is equal to ϕ_o if $J_1 = \phi$ and $J_2 = \phi$. It is less than ϕ_o for small enough ε if $J_1 \neq \phi$ or $J_2 \neq \phi$ since the linear term in ε dominates the quadratic term. This implies that $\psi(\varepsilon y_1, \varepsilon y_2) \leq \phi_o = \psi(0,0)$ for all $y_1 \geq 0$, $y_2 \geq 0$ and small enough $\varepsilon > 0$. ||

(iii) This is obviously true since $\psi(y_1,y_2) \leq \phi_o = \psi(0,0)$ for all $y_1 \geq 0$, $y_2 \geq 0$.

The proof of Theorem 1 suggests to us a vertex following algorithm to be described below:

Algorithm 1 (Mountain Climbing)

Step 1. Obtain a pair of basic feasible solutions, $x_1^o \in X_1$, $x_2^o \in X_2$. Let $k = 0$.

Step 2. Given (x_1^k, x_2^k), a pair of basic feasible solutions of X_1 and X_2, solve a subproblem: $\max\{\phi(x_1, x_2^k) | x_1 \in X_1\}$. Let x_1^{k+1} and B_1^{k+1} be its optimal basic solution and corresponding basis.

Step 3. Solve a subproblem: $\max\{\phi(x_1^{k+1}, x_2) | x_2 \in X_2\}$, and let x_2^{k+1} and B_2^{k+1} be its optimal basic solution and corresponding basis.

Step 4. Compute $d_1(B_1^{k+1}, B_2^{k+1})$, the coefficients of y_1 in the canonical representation (2.4) relative to bases B_1^{k+1}, B_2^{k+1}. If $d_1(B_1^{k+1}, B_2^{k+1}) \leq 0$, then let $B_i^* = B_i^{k+1}$ and x_i^* be the basic feasible solutions associated with B_i^*, $i = 1,2$ and HALT. Otherwise increase k by 1 and go to Step 2.

Note that the subproblems to be solved in Steps 2 and 3 are linear programs.

Proposition 2.3. If X_1 and X_2 are bounded, then Algorithm 1 halts in finitely many steps generating a Kuhn-Tucker point.

Proof. If every basis of X_1 is nondegenerate, then the value of objective function ϕ can be increased in Step 2 as long as there is a positive component in d_1. Since the number of bases of X_1 is finite and no pair of bases can be visited twice because the objective function is strictly increasing in each passage of Step 2, the algorithm will eventually terminate with the condition $d_1(B_1^{k+1}, B_2^{k+1}) \leq 0$ being satisfied. When X_1 is degenerate, then there is a chance of infinite cycling among certain pairs of basic solutions. We will show, however, that this cannot happen in the above process if we employ an appropriate tie breaking device in linear programming. Suppose that

$$\phi(x_1^{k+1},x_2^k) = \max\{\phi(x_1,x_2^k)\,|\,x_1 \in X_1\} \quad : \text{optimal basis } B_1^{k+1}$$

$$\phi(x_1^{k+1},x_2^{k+1}) = \max\{\phi(x_1^{k+1},x_2)\,|\,x_2 \in X_2\} \quad : B_2^{k+1}$$

. . . .

. . . .

$$\phi(x_1^{k+\ell},x_2^{k+\ell-1}) = \max\{\phi(x_1,x_2^{k+\ell-1})\,|\,x_1 \in X_1\} : B_1^{k+\ell}$$

$$\phi(x_1^{k+\ell},x_2^{k+\ell}) = \max\{\phi(x_1^{k+\ell},x_2)\,|\,x_2 \in X_2\} \quad : B_2^{k+\ell} \quad ,$$

where $x^{k+\ell} = x^{k+1}$, for the first time in the cycle. Since the value of objective function ϕ is nondecreasing and

$$(x_1^{k+\ell},x_2^{k+\ell}) \equiv \phi(x_1^{k+1},x_2^{k+\ell}) \leq \phi(x_1^{k+1},x_2^{k+1}) \quad ,$$

we have that

$$\phi(x_1^{k+1},x_2^{k+1}) = \phi(x_1^{k+2},x_2^{k+1}) = \ldots . = \phi(x_1^{k+\ell},x_2^{k+\ell}) \quad .$$

It is obvious that $d_2(B_1^{k+1},B_2^{k+1}) \leq 0$ by the definition of optimality of B_2^{k+1}. Suppose that the j^{th} component of $d_1(B_1^{k+1},B_2^{k+1})$ is positive. Then we could have introduced y_{ij} into the basis. However, since the objective function should not increase, y_{ij} comes into the basis at zero level. Hence the vector y_1 remains zero. We can eliminate the positive element of d_1, one by one (using tie breaking device for the degenerate LP if necessary) with no actual change in the value of y_1. Eventually, we have $d_2 \leq 0$ with $y_1 = 0$ and the corresponding basis $\tilde{B}_1^{k+1}$. Referring to the standard form, the corresponding x_1 value remains unchanged i.e., stays at x_1^{k+1} and hence $d_2(\tilde{B}_1^{k+1},B_2^{k+1}) \leq 0$, because B_2^{k+1} is the optimal basis for $x_1 = x_1^{k+1}$, and $\tilde{x}_1^{k+1} = x_1^{k+1}$. By Theorem 2 (i), the solution obtained is a Kuhn-Tucker point. ||

Let us assume in the following that a Kuhn-Tucker point has been obtained and that a canonical representation (2.4) relative to the associated pair of bases has been given.

By Theorem 2 (iii), that pair of basic feasible solutions is optimal if $Q \leq 0$. We will assume that this is not the case and let

$$K = \{(i,j) \mid q_{ij} > 0\} \quad .$$

Let us define for $(i,j) \in K$, a function $\psi_{ij} : R_+^2 \to R$,

$$\psi_{ij}(\xi,\eta) = d_{1i}\xi + d_{2j}\eta + q_{ij}\xi\eta \quad .$$

Proposition 2.4. If $\psi_{ij}(\xi_o,\eta_o) > 0$ for some $\xi_o \geq 0$, $\eta_o \geq 0$, then

$$\psi_{ij}(\xi,\eta) > \psi(\xi_o,\eta_o) \text{ for all } \xi > \xi_o, \quad \eta > \eta_o$$

Proof.

$$\begin{aligned}
\psi_{ij}(\xi,\eta) - \psi_{ij}(\xi_o,\eta_o) &= (\xi - \xi_o)(d_{1i} + q_{ij}\eta_o) \\
&\quad + (\eta - \eta_o)(d_{2j} + q_{ij}\xi_o) \\
&\quad + q_{ij}(\xi - \xi_o)(\eta - \eta_o) \\
&\geq (\xi - \xi_o)(-d_{2j}\frac{\eta_o}{\xi_o}) \\
&\quad + (\eta - \eta_o)(-d_{1i}\frac{\xi_o}{\eta_o}) \\
&\quad + q_{ij}(\xi - \xi_o)(\eta - \eta_o) > 0 \quad . \qquad ||
\end{aligned}$$

This proposition states that if the objective function increases in the directions of y_{1j} and y_{2j}, then we can increase more if we go further into this direction.

Definition 2.1. Given a basic feasible solution $x_i \in X_i$, let $N_i(x_i)$ be the set of adjacent basic feasible solutions which can be reached from x_i in one pivot step.

Definition 2.2. Let ε be a nonnegative scalar. A pair of basic feasible solutions (x_1^*, x_2^*), $x_i^* \in X_i$, $i = 1,2$ is called an ε-locally maximum pair of basic feasible solution if

(i) $d_i \leq 0$, $i = 1,2$

(ii) $\phi(x_1^*, x_2^*) \geq \phi(x_1, x_2) - \varepsilon$ for all $x_i \in N_i(x_i^*)$, $i = 1,2$.

Given a Kuhn-Tucker point (x_1^*, x_2^*), we will compute $\phi(x_1, x_2)$ for all $x_i \in N_i(x_i^*)$, $i = 1,2$ for which a potential increase of objective function ϕ is possible. Given a canonical representation, it is sufficient for this purpose to calculate $\psi_{ij}(\bar{\xi}_i, \bar{\eta}_j)$ for $(i,j) \in K$ where $\bar{\xi}_i$ and $\bar{\eta}_j$ represent the maximum level of nonbasic variables x_{1j} and x_{2j} when they are introduced into the bases without violating feasibility.

Algorithm 2 (Augmented Mountain Climbing)

Step 1. Apply Algorithm 1 and let $x_i^* \in X_i$, $i = 1,2$ be the resulting pair of basic feasible solutions.

Step 2. If (x_1^*, x_2^*) is an ε-locally maximum pair of basic feasible solutions, then HALT. Otherwise, move to the adjacent pair of basic feasible solutions $(\hat{x}_1, \hat{x}_2)$ where

$$\phi(\hat{x}_1, \hat{x}_2) = \max\{\phi(x_1, x_2) \mid x_i \in N_i(x_i^*), \quad i = 1,2\}$$

and go to Step 1.

Proposition 2.5. If X_1 and X_2 are bounded and if $\varepsilon > 0$, Algorithm 2 halts in finitely many steps generating an ε-locally maximum pair of basic feasible solutions.

Proof. It follows immediately from the following facts that:

(i) step 1 converges in finitely many steps (by Proposition 2.3),

(ii) whenever we pass Step 2, the value of the objective function is improved by at least $\varepsilon(> 0)$,

(iii) there are only finitely many basic feasible solutions for X_1 and X_2. ||

3. Cutting Planes

We will assume in this section that an ε-locally maximum pair of basic feasible solutions has been obtained and that a canonical representation relative to this pair of basic feasible solution (x_1^*, x_2^*) has been given. Since we will refer here exclusively to a canonical representation, we will reproduce it for future convenience:

$$\begin{aligned} \max\ \psi(y_1,y_2) &= d_1^t y_1 + d_2^t y_2 + y_1^t Q y_2 + \phi(x_1^*,x_2^*) \\ \text{s.t. } F_1 y_2 &\le f_1 \ , \quad y_1 \ge 0 \\ F_2 y_2 &\le f_2 \ , \quad y_2 \ge 0 \ , \end{aligned} \tag{3.1}$$

where

$$d_i \le 0 \ , \quad f_i \ge 0 \ , \quad i = 1,2 \ .$$

Let

$$Y_i = \{y_i \in R^{\ell_i} \mid F_i y_i \le f_i,\ y_i \ge 0\} \ , \quad i = 1,2 \tag{3.2}$$

$$Y_i^{(\ell)} = \{y_i \in R^{\ell_i} \mid y_{i\ell} \ge 0,\ y_{ij} = 0,\ j \ne \ell\}$$

$$\ell = 1,\dots,\ell_i,\ i = 1,2 \ , \tag{3.3}$$

i.e. $Y_i^{(\ell)}$ is the ray emanating from $y_i = 0$ in the direction $y_{i\ell}$.

<u>Lemma 3.1</u>. Let

$$\psi_1(\cdot) = \max\{\psi(\cdot, y_2) \mid y_2 \in Y_2\} \quad . \tag{3.4}$$

If $\Psi_1(u) > 0$ for some $u \in Y_1^{(\ell)}$, then $\Psi_1(v) > \Psi_1(u)$ for all $v \in Y_1^{(\ell)}$ such that $v > u$.

<u>Proof</u>. Let $u = (0,\ldots,0, u_\ell, 0,\ldots,0)$. First note that $u_\ell > 0$, since if $u_\ell = 0$, then $\Psi_1(u) = \max\{d_2^t y_2 \mid y_2 \in Y_2\} = 0$.

Let $v = (o,\ldots,0, v_\ell, 0,\ldots,0)$ where $v_\ell \geq u_\ell$. Then for all $y_2 \in Y_2$, we have

$$\begin{aligned}
\psi(v,y_2) &= \psi(u,y_2) + (v_\ell - u_\ell)\left\{d_{1\ell} + \sum_{j=1}^{\ell_2} q_{\ell j} y_{2j}\right\} \\
&\geq \psi(u,y_2) + \frac{v_\ell - u_\ell}{u_\ell}\left\{d_{1\ell}u_\ell + \sum_{j=1}^{\ell_2} (d_{2j} + q_{\ell j}u_\ell)\, y_{2j}\right\} \\
&= \frac{v_\ell}{u_\ell}\,\psi(u,y_2) \quad .
\end{aligned}$$

The inequality follows from $d_2 \leq 0$. Thus

$$\begin{aligned}
\max\{\psi(v,y_2) \mid y_2 \in Y_2\} &\geq \frac{v_\ell}{u_\ell} \max\{\psi(u,y_2) \mid y_2 \in Y_2\} \\
&\geq \max\{\psi(u,y_2) \mid y_2 \in Y_2\} \quad . \qquad ||
\end{aligned}$$

This lemma shows that the function Ψ_1 is a strictly increasing function of y_1 on $Y_1^{(\ell)}$ beyond the point where Ψ_1 first becomes positive.

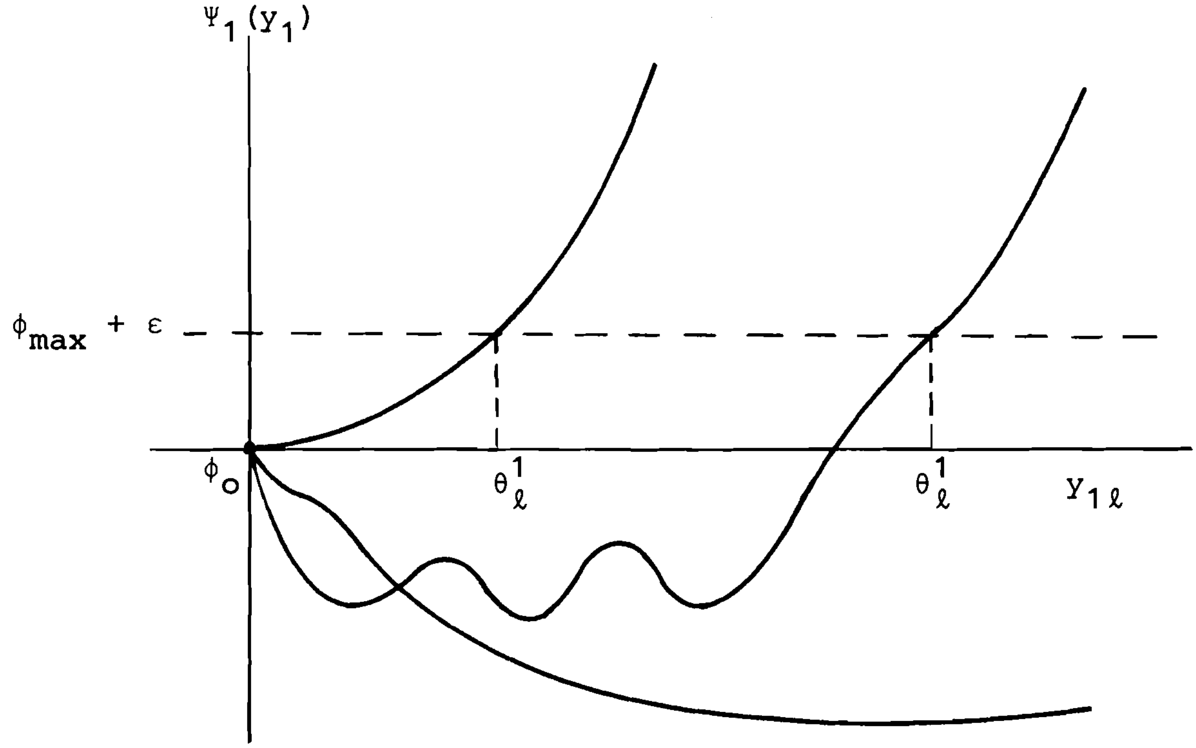


Figure 3.1. Shape of the function Ψ_1.

Let ϕ_{max} be the value of the objective function associated with the best feasible solution obtained so far by one method or another and let us define θ^1_ℓ, $\ell = 1,\ldots,\ell_1$ as follows:

$$\theta^1_\ell = \max \theta \text{ for which}$$

$$\max\{\Psi_1(y_1) \mid y_1 \in Y_1^{(\ell)},\ 0 \leq y_{1\ell} \leq \theta\} \leq \phi_{max} + \varepsilon \quad . \tag{3.5}$$

<u>Lemma 3.2</u>. $\theta^1_\ell > 0$, $\ell = 1,\ldots,\ell_1$.

<u>Proof</u>. Let $y_1 = (0,\ldots,0,\ y_{1\ell},\ 0,\ldots,0)$. Since $d_1 \leq 0$, $d_2 \leq 0$, we have

$$\Psi(y_1,y_2) = d_{1\ell}y_{1\ell} + \sum d_{2j}y_{2j} + y_{1\ell} \sum q_{\ell j}y_{2j} + \phi_o$$

$$\leq y_{1\ell} \sum q_{\ell j}y_{2j} + \phi_o \quad .$$

Letting $\alpha = \max\{\Sigma q_j y_{2j} \; y_2 \in Y_2 \geq 0$, we know from the above inequality that

$$\theta^1_\ell \begin{cases} \geq (\phi_{max} - \phi_o + \varepsilon)/\alpha > 0 & \alpha > 0 \\ = +\infty & \alpha = 0 \end{cases} . \qquad ||$$

<u>Theorem 3.3</u>. Let

$$\Delta_1(\theta^1) = \{y_1 \in R^{\ell_1} \; \sum_{j=1}^{\ell_1} y_{1j}/\theta^1_j \leq 1, \; y_1 \geq 0\} \quad . \tag{3.6}$$

Then

$$\max\{\psi(y_1,y_2) | y_1 \in \Delta_1(\theta^1), \; y_2 \in Y_2\} \leq \phi_{max} + \varepsilon \quad .$$

<u>Proof</u>. Let

$$\tilde{\theta}^1_j = \begin{cases} \theta^1_j & \text{if } \theta^1_j \text{ is finite} \\ \theta_o & \text{if } \theta^1_j = \infty \end{cases} \tag{3.7}$$

where $\theta_o > 0$ is constant. Then

$$\max\{\psi(y_1,y_2) | y_1 \in \Delta_1(\theta^1), \; y_2 \in y_2\} \quad .$$

$$= \lim_{\theta_o \to \infty} \max\{\psi(y_1,y_2) | y_1 \in \Delta_1(\tilde{\theta}^1), \; y_2 \in Y_2\} \quad .$$

The right hand term inside the limit is a bilinear program with bounded feasible region, and hence by Theorem 2.1, there exists an optimal solution among basic feasible solutions. Since the basic feasible solution for the systems of inequalities defining $\Delta(\tilde{\theta}^1)$ are $(0,\ldots,0)$ and $y_1^\ell = (0,\ldots,0, \tilde{\theta}^1_\ell, 0,\ldots,0)$, $\ell = 1,\ldots,\ell_1$, we have

$$\max\{\psi(y_1,y_2)\,|\,y_1 \in \Delta_1(\tilde{\theta}_1),\ y_2 \in Y_2\}$$

$$= \max \left[\max\{\psi(0,y_2)\,|\,y_2 \in Y_2\},\ \max_{\ell}\ \max_{y_2}\{\psi(y_1^{\ell},y_2)\,|\,y_2 \in Y_2\}\right] .$$

However, since $d_2 \leq 0$,

$$\max\{\psi(0,y_2)\,|\,y_2 \in Y_2\} = \max\{d_2^t y_2\,|\,y_2 \in Y_2\} + \phi_o \leq \phi_o \leq \phi_{max} + \varepsilon \quad .$$

Also,

$$\max_{y_2}\{\psi(y_1,y_2)\,|\,y_2 \in Y_2\} \leq \phi_{max} + \varepsilon$$

by the definition of $\tilde{\theta}_{\ell}^{1}$ (See (3.5) and (3.7)). Hence

$$\lim_{\theta_o \to \infty} \max\{\psi(y_1^{\ell},y_2)\,|\,y_2 \in Y_2\} \leq \phi_{max} + \varepsilon \quad . \qquad ||$$

This theorem shows that the value of the objective function $\phi(y_1,y_2)$ associated with the points y_1 in the region $Y_1 \cap \Delta_1(\theta^1)$ is not greater than $\phi_{max} + \varepsilon$ regardless of the choice of $y_2 \in Y_2$ and hence this region $Y_1 \cap \Delta_1(\theta^1)$ can be ignored in the succeeding process to obtain an ε-optimal solution. The cut

$$H_1(\theta^1) : \sum_{j=1}^{\ell_1} y_{1j}/\theta_j^1 \geq 1$$

is, therefore, a 'valid' cut in the sense that it:

(i) does not contain the current ε-locally maximum pair of basic feasible solutions;

(ii) contains all the candidates $y_1 \in Y_1$ for which

$$\max\{\psi(y_1,y_2)\,|\,y_2 \in Y_2\} > \phi_{max} + \varepsilon \quad .$$

Since θ^1 is dependent on the feasible region Y_2, we will occasionally use the notation $\theta^1(Y_2)$.

Since the problem is symmetric with respect to Y_1 and Y_2, we can, if we like, interchange the role of Y_1 and Y_2 to obtain another valid cutting plane relative to Y_2:

$$H_2(\theta^2) : \sum_{j=1}^{\ell_2} y_{2j}/\theta_j^2 = 1 \quad .$$

Cutting Plane Algorithm

Step 0. Set $\ell = 0$. Let $X_i^o = X_i$, $i = 1,2$.

Step 1. Apply Algorithm 2 (Augmented Mountain Climbing Algorithm) with a pair of feasible regions X_1^ℓ, X_2^ℓ.

Step 2. Compute $\theta^1(Y_2^\ell)$. Let $Y_1^{\ell+1} = Y_1^\ell \setminus \Delta_1(\theta^1(Y_2^\ell)$. If $Y_1^{\ell+1} = \phi$, stop. Otherwise proceed to the next step.

Step 2' (Optional). Compute $\theta^2(Y_1^{\ell+1})$. Let $Y_2^{\ell+1} = Y_2 \setminus \Delta_2(\theta^2(Y_1^{\ell+1}))$. If $Y_2^{\ell+1} = \phi$, stop. Otherwise proceed to the next step.

Step 3. Add 1 to ℓ. Go to Step 1.

It is now easy to prove the following theorem.

Theorem 3.4. If the cutting plane algorithm defined above stops in Step 2 or 2', with either $Y_1^{\ell+1}$ or $Y_2^{\ell+1}$ becoming empty, then ϕ_{max} and the associated pair of basic feasible solutions is an ε-optimal solution of the bilinear program.

Proof. Each cutting plane added does not eliminate any point for which the objective function is greater than $\phi_{max} + \varepsilon$. Hence if either $Y_1^{\ell+1}$ or $Y_2^{\ell+2}$ becomes empty, we can conclude that $\max\{\psi(y_1,y_2) \mid y_1 \in Y_1,\ y_2 \in Y_2\} \leq \phi_{max} + \varepsilon$.

According to our cutting plane algorithm, the number of constraints increases by 1 whenever we pass Step 2 or 2', the size of subproblem becomes bigger and the constraints are also more prone to degeneracy. From this viewpoint, we want to add a smaller number of cutting planes, particularly when the original constraints have a good structure. In such cases, we might as well omit Step 2', taking Y_2 as the constraints throughout the whole process.

Another requirement for the cut is that it should be as deep as possible, in the following sense.

Definition 3.1. Let $\theta = (\theta_j) > 0$, $\tau = (\tau_j) > 0$. Then the cut $\sum y_{1j}/\theta_j \geq 1$ is deeper than $\sum y_{1j}/\tau_j \geq 1$ if $\theta \geq \tau$, with at least one component with strict inequality.

Looking back into the definition (3.5) of θ^1, it is clear that $\theta^1(U) \geq \theta^1(V)$ when $U \subset V \subset R^{\ell_2}$ and that the cut associated with $^1(U)$ is deeper than $\theta^1(V)$. This observation leads to the following procedure.

Iterative Improvement Procedure. Let $H_1(\theta^1(Y_2))$ and $H_2(\theta^2(Y_1))$ be a pair of valid cuts and let $Y_1' = Y_1 \setminus \Delta_1(\theta'(Y_2))$, $Y_2' = Y_2 \setminus \Delta_2(\theta^2(Y_1))$ be the shrunken feasible regions. Generate cuts $H_1(\theta^1(Y_2'))$ and $H_2(\theta^2(Y_1'))$ which are generally deeper than $H_1(\theta^1(Y_2))$ and $H_2(\theta^2(Y_1))$, respectively. Iterate this process until successive cuts converge within some tolerance.

This iterative improvement scheme is very powerful when the problem is symmetric with respect to y_1 and y_2:

$$\max\{d^t t_1 + d^t y_2 + y_1^t Q y_2 \mid F y_1 \leq f,\ y_1 \geq 0,\ F y_2 \leq f,\ y_2 \geq 0\} \quad . \tag{3.8}$$

In particular, maximization of a convex quadratic function subject to linear constraints

$$\max\{2d^t x + x^t Q x \mid Fx \leq f,\ x \geq 0\}$$

is equivalent to (3.8) and the iterative process described above works remarkably well for this class of problems. The details about this, together with the comparison of our cuts with the ones proposed by Tui and Ritter, will be discussed in full in [11].

The following theorem gives us a method to compute θ^1 using the dual simplex method.

Theorem 3.5.

$$\theta^1_\ell = \min\{-d^t z + (\phi_{max} - \phi_o + \varepsilon)\, z_o\}$$

$$\text{s.t. } F_2 z - f_2 z_o \leq 0$$

$$\sum_{j=1}^{\ell_2} q_{\ell j} z_j + d_{1\ell} z_o = 1 \qquad (3.9)$$

$$z_j \geq 0,\ j = 1,\ldots,{}_2,\ell z_o \geq 0 \ .$$

Proof. Let

$$g(\theta) = \max\{d_1^t y_1 + d_2^t y_2 + y_1^t Q y_2 | F_2 y_2 \leq f_2,\ y_2 \geq 0,$$

$$0 \leq y_{1\ell} \leq \theta,\ y_{1j} = 0,\ j \neq \ell\} \ .$$

θ_ℓ is then given as the maximum of θ for which $g(\theta) \leq \phi_{max} - \phi_o + \varepsilon$. It is not difficult to observe that

$$g(\theta) = \max \left[0,\ \max\{d_{1\ell}\theta + (d_2 + \theta q_{\ell\cdot})^t y_2 | F_2 y_2 \leq f_2,\ y_2 \geq 0\}\right] \ ,$$

where $q_{\ell\cdot} = (q_{\ell 1},\ldots,q_{\ell\ell_2})^t$. Therefore, θ^1_ℓ is the maximum of θ for which

$$g_1(\theta) \equiv \max\{d_{1\ell}\theta + (d_2 + \theta q_{\ell\cdot})^t y_2 | F_2 y_2 \leq f_2,\ y_2 \geq 0\}$$

$$\leq \phi_{max} - \phi_o + \varepsilon \ .$$

The feasible region defining $g_1(\theta)$ is, by assumption, bounded and non-empty, and by art duality theorem

$$g_1(\theta) = \min\{f_2^t u + d_{1\ell}\theta \mid F_2^t u \geq d_2 + \theta q_{\ell\cdot}, \; u \geq 0\} \quad .$$

Hence θ_ℓ is the maximum of θ for which the system

$$\{f_2^t u + d_{1\ell}\theta \leq \phi_{max} - \phi_o + \varepsilon, \; -F_2^t u - q_{\ell\cdot}\theta \leq -d_2, \; u \geq 0\}$$

is feasible, i.e.,

$$\theta_\ell = \max\left\{\theta \;\middle|\; \begin{array}{l} f_2^t u + d_{1\ell}\theta \leq \phi_{max} - \phi_o + \varepsilon \\ -F_2^t u - q_{\ell\cdot}\theta \leq -d_2 \\ u \geq 0 \end{array}\right\} .$$

This problem is always feasible, and again using art duality theorem,

$$\theta_\ell = \min\left\{-d_2^t z + (\phi_{max} - \phi_o + \varepsilon)\, z_o \;\middle|\; \begin{array}{l} q_{\ell\cdot}^t z + d_{1\ell} z_o = 1 \\ f_2 z_o - Fz \geq 0 \\ z \geq 0, \; z_o \geq 0 \end{array}\right\} ,$$

with the usual understanding that $\theta_\ell = +\infty$ if the constraint set above is empty. ||

Note that $d_2 \leq 0$ and $\phi_{max} - \phi_o + \varepsilon \geq 0$ and hence $(z, z_o) = (0,0)$ is a dual feasible solution. Also the linear program defining θ_ℓ^1 is only one row different for different ℓ, so that they are expected to be solved without an excessive amount of computation. Since the value of the objective function of (3.9) approaches its minimal value monotonically from below, we can stop pivoting if we like when the value of the objective function becomes greater than some specified value. The important thing to note is that if we pivot more, we get a deeper cut, in general.

4. Numerical Examples

Let us consider the following simple two dimensional example (illustrated in Figure 4.1):

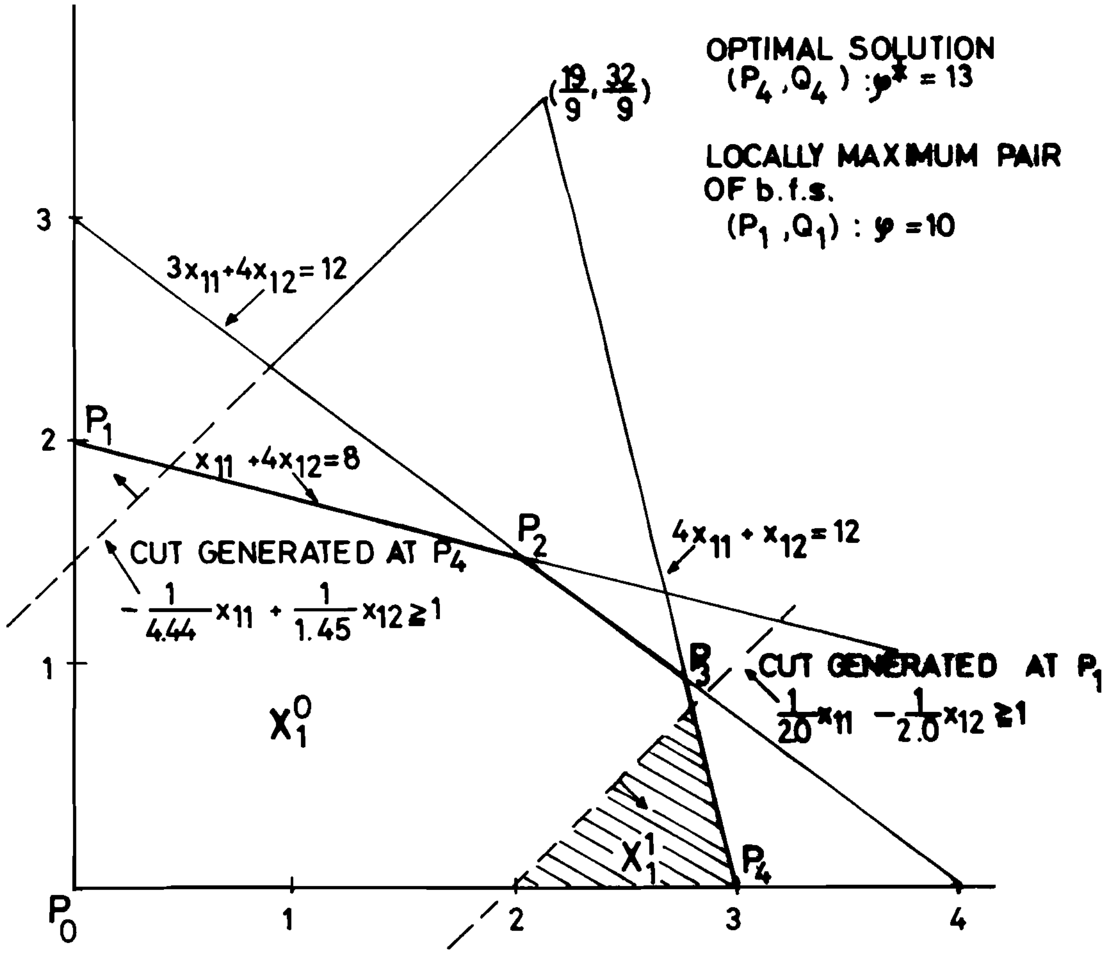
$$\text{maximize } \phi(x_1,x_2) = (-1,1)\begin{pmatrix} x_{11} \\ x_{12} \end{pmatrix} + (1,0)\begin{pmatrix} x_{21} \\ x_{22} \end{pmatrix} + (x_{11},x_{12})\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} x_{21} \\ x_{22} \end{pmatrix}$$

$$\text{s.t. } \begin{pmatrix} 1 & 4 \\ 4 & 1 \\ 3 & 4 \end{pmatrix}\begin{pmatrix} x_{11} \\ x_{12} \end{pmatrix} \leq \begin{pmatrix} 8 \\ 12 \\ 12 \end{pmatrix}, \quad \begin{pmatrix} 2 & 1 \\ 1 & 2 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} x_{21} \\ x_{22} \end{pmatrix} \leq \begin{pmatrix} 8 \\ 8 \\ 5 \end{pmatrix}$$

$$(x_{11},x_{12}) \geq 0, \qquad (x_{21},x_{22}) \geq 0 \quad .$$

There are two locally maximum pairs of basic feasible solutions i.e., (P_1,Q_1) and $P_4,Q_4)$, for which the value of the objective function is 10 and 13, respectively. We applied the algorithm omitting Step 2'. Two cuts generated at P_1 and P_4 are shown on the graph. In two steps, $X_1^2 = \phi$ and the global optimum (P_4,Q_4) has been identified.

We have coded the algorithm in FORTRAN IV for CYBER 74 at the Technische Hochschule, Vienna, and tested it for various problems of a size up to 10 × 22, 13 × 24; all of them were solved successfully.



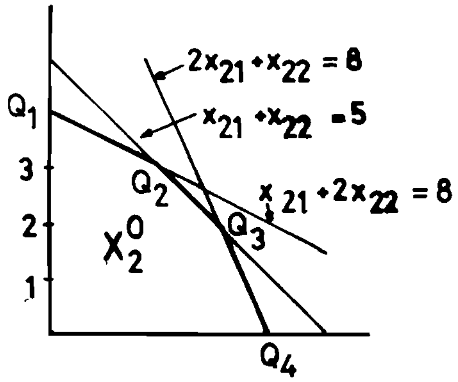


Figure 4.1. A numerical example.

Problem No.	Size of the Problem		ε/ϕ_{max}	No. of Local Maxima Identified	CPU time (sec)
	X_1	X_2			
1	2 x 4	2 x 4	0.0	1	
2	3 x 6	3 x 6	0.0	1	$\leq$ 0.5
3	2 x 5	2 x 5	0.0	1	
4	6 x 11	6 x 11	0.0	1	$\leq$ 0.5
5	3 x 5	3 x 5	0.0	2	
6	5 x 8	5 x 8	0.0	1	
7	3 x 6	3 x 6	0.0	1	1.0
8	7 x 11	7 x 11	0.0	1	
9	5 x 8	5 x 8	0.0	2	0.6
10	9 x 19	9 x 19	0.0	2	8.1
11	6 x 12	6 x 12	0.05	5	
12	6 x 12	6 x 12	0.01	6	
13	6 x 12	6 x 12	0.0	6	
14	10 x 22	13 x 24	0.05	3	20.7

Problem 2 is taken from [20] and problem 9 from [2]. 11 $\sim$ 13 are the same problems having six global maxima with equal value. These are in fact global optima. The data for this problem is given below:

$$c_1 = 0, \; c_2 = 0, \; b_1 = b_2 = (21,21,21,21,21,21)^t$$

$$C = \begin{bmatrix} 2 & -1 & 0 & 0 & 0 & 0 \\ -1 & 2 & -1 & 0 & 0 & 0 \\ 0 & -1 & 2 & -1 & 0 & 0 \\ 0 & 0 & -1 & 2 & -1 & 0 \\ 0 & 0 & 0 & -1 & 2 & -1 \\ 0 & 0 & 0 & 0 & -1 & 2 \end{bmatrix} \quad A_1 = A_2 = \left[\begin{array}{cccccc|cccccc} 1 & 2 & 3 & 4 & 5 & 6 & 1 & 0 & 0 & 0 & 0 & 0 \\ 2 & 3 & 4 & 5 & 6 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 3 & 4 & 5 & 6 & 1 & 2 & 0 & 0 & 1 & 0 & 0 & 0 \\ 4 & 5 & 6 & 1 & 2 & 3 & 0 & 0 & 0 & 1 & 0 & 0 \\ 5 & 6 & 1 & 2 & 3 & 4 & 0 & 0 & 0 & 0 & 1 & 0 \\ 6 & 1 & 2 & 3 & 4 & 5 & 0 & 0 & 0 & 0 & 0 & 1 \end{array}\right]$$

$$\qquad\qquad\qquad\qquad\qquad\qquad \uparrow A_o \qquad\qquad\qquad \uparrow I_6$$

This is the problem associated with the convex maximization problem

$$\max\{\tfrac{1}{2}x^t Cx \mid A_o x \leq b,\ x \geq 0\} \quad .$$

Data for problem 14 was generated randomly.

References

[1] Altman, M. "Bilinear Programming," Bullentin de l'Académie Polonaise des Sciences, 16, No. 9 (1968), 741-746.

[2] Balas, E. and Burdet, C.-A. "Maximizing a Convex Quadratic Function Subject to Linear Constraints," Management Science Research Report No. 299, GSIA, Carnegie-Mellon University, Pittsburgh, Pennsylvania, July 1973.

[3] Cabot, A.V. and Francis, R.L. "Solving Certain Nonconvex Quadratic Minimization Problems by Ranking Extreme Points," Operations Research, 18, No. 1 (1970), 82-86.

[4] Charnes, A. and Cooper, W.W. "Nonlinear Power of Adjacent Extreme Point Methods in Linear Programming," Econometrica, 25 (1957), 132-153.

[5] Candler, W. and Townsley, R.J. "The Maximization of a Quadratic Function of Variables Subject to Linear Inequalities," Management Science, 10, No. 3 (1964), 515-523.

[6] Cottle, R.W. and Mylander, W.C. "Ritter's Cutting Plane Method for Nonconvex Quadratic Programming," in Integer and Nonlinear Programming (J. Abadie, ed.), North Holland, Amsterdam, 1970.

[7] Dantzig, G.B. "Reduction of a 0-1 Integer Program to a Bilinear Separable Program and to a Standard Complementary Problem," Unpublished Note, July 27, 1971.

[8] Dantzig, G.B. "Solving Two-Move Games with Perfect Information," RAND Report P-1459, Santa Monica, California, 1958.

[9] Falk, J. "A Linear Max-Min Problem," Serial T-251, The George Washington University, Washington, D.C., June 1971.

[10] Gallo, G. and Ülkücü, A. "Bilinear Programming: An Exact Algorithm," Paper presented at the 8th International Symposium on Mathematical Programming, Stanford University, Stanford, California, August 1973.

[11] Konno, K. "Maximization of Convex Quadratic Function under Linear Constraints," Research Memorandum, International Institute for Applied Systems Analysis, Laxenburg, Austria, 1974. (Forthcoming in Mathematical Programming.)

[12] Konno, H. "Bilinear Programming Part II: Applications of Bilinear Programming," Technical Report No. 71-10, Department of Operations Research, Stanford University, Stanford, California, August 1971.

[13] Mangasarian, O.L. "Equilibrium Points of Bimatrix Games," SIAM Journal of Applied Mathematics, 12, No. 4 (1964), 778-780.

[14] Mangasarian, O.L. and Stone, H. "Two-Person Nonzero-Sum Games and Quadratic Programming," Journal of Mathematical Analysis and Applications, 9 (1964), 348-355.

[15] Mills, H. "Equilibrium Points in Finite Games," SIAM Journal of Applied Mathematics, 8, No. 2 (1960), 397-402.

[16] Mylander, W.C. "Nonconvex Quadratic Programming by a Modification of Lemke's Method," RAC-TP-414, Research Analysis Corporation, McLean, Virginia, 1971.

[17] Ritter, K. "A Method for Solving Maximum Problems with a Nonconcave Quadratic Objective Function," Z. Wahrscheinlichkeitstheorie, verv. Geb., 4 (1966), 340-351.

[18] Raghavachari, M. "On Connections Between Zero-One Integer Programming and Concave Programming under Linear Constraints," Operations Research, 17, No. 4 (1969), 680-684.

[19] Tui, H. "Concave Programming under Linear Constraints," Soviet Mathematics (1964), 1537-1440.

[20] Zwart, P. "Nonlinear Programming: Counterexamples to Two Global Optimization Algorithms," Operations Research, 21, No. 6 (1973), 1260-12.

[21] Zwart, P. "Computational Aspects of the Use of Cutting Planes in Global Optimization," in Proceedings of the 1971 Annual Conference of the ACM (1971), 457-465.