

MAXIMIZING A CONVEX QUADRATIC FUNCTION OVER A HYPERCUBE

Hiroshi Konno
University of Tsukuba

(Received February 2, 1980)

Abstract This paper deals with a new algorithm for obtaining a global maximum of a convex quadratic function over a unit hypercube, which is a classical and tough combinatorial problem. The basic idea of our algorithm is to reformulate this problem as an equivalent bilinear programming problem and to apply cutting plane approach developed by the author for solving bilinear knapsack problem. It will be shown that a deep cut can be generated with a nominal amount of computation. Some of the potential advantages of this algorithm over total enumeration are that it can generate a good local maximum at the earlier stage and that it has a good chance of obtaining a global maximum by searching only a minor portion of all solutions so that it can handle larger problems.

1. Introduction

The problem to be discussed in this paper is a special kind of quadratic programming problem:

$$\begin{cases} \text{maximize} & f(w) = 2c^t w + w^t Q w \\ \text{subject to} & 0 \leq w \leq e \end{cases} \quad (1)$$

where $c \in R^n$ and $e = (1, 1, \dots, 1)^t \in R^n$. It will be assumed throughout that $Q \in R^{n \times n}$ is a symmetric positive semi-definite matrix, so that f is convex. In this case, (1) is equivalent to a well known and difficult combinatorial problem:

$$\begin{cases} \text{maximize} & f(w) = 2c^t w + w^t Q w \\ \text{subject to} & w \in \{0, 1\}^n \end{cases} \quad (2)$$

Quadratic programs with only lower and upper bound constraints on the variables are also important because a significant portion of real world applications of quadratic programs are reported to be of this type [7].

It is certainly possible to find a global maximum of (1) in finitely many steps by enumerating 2^n possible 0-1 solutions. Also, cutting plane algorithms [16, 8] for convex maximization problems can be used.

Unfortunately, however, these methods may not be practical if $n \geq 30$. In addition, the latter algorithms are not convergent unless impractically expensive cuts such as the ones developed in [11, 15] are to be used.

The basic idea of our algorithm is to reformulate (1) as an equivalent bilinear programming problem and to apply cutting plane approach developed by the author for solving bilinear knapsack problem [10]. Some of the potential advantages of this algorithm over total enumeration are that it (i) can generate a good local maximum at the earlier stage of the computation, (ii) has a good chance of obtaining a global maximum by searching only a minor portion of all solution, (iii) can handle larger problems. On the other hand, it has a disadvantage that a 0-1 integer program has to be solved as a sub-problem. However, this difficulty can be alleviated by exploiting the structure of the problem and by adding deep cuts at local maxima. These points have yet to be checked by an extensive numerical experiments on large scale problems, but promising results of the similar algorithm applied to bilinear knapsack problem [10] show some evidence of the advantage of this approach over the other.

In the next section, bilinear programming problem equivalent to (1) will be introduced and the procedure to obtain a local optimum and a semiglobal optimum will be developed. Section 3 will be devoted to the discussion of a finitely convergent cutting plane algorithm. It will be shown that a deep cut can be generated with nominal computation by virtue of the special structure of the problem. Finally, methods to obtain stronger cuts and an illustrative numerical example are given in Section 4 and 5, respectively.

2. Equivalent Bilinear Programming Problem and LG Maxima

Let us introduce a bilinear programming problem associated with problem (1):

$$\begin{cases} \text{maximize } \phi(x, y) = c^t x + c^t y + x^t Q y \\ \text{subject to } 0 \leq x \leq e \\ \phantom{\text{subject to }} 0 \leq y \leq e \end{cases} \quad (3)$$

The following theorems are crucial to the development of this paper.

Readers are referred to [8] for the proofs of these theorems.

Theorem 1. Problem (3) has an optimal solution (x^*, y^*) where both x^* and y^* are extreme points of the unit hypercube.

Theorem 2. Let (x^*, y^*) be optimal to (3), then both x^* and y^* are optimal to (1).

These two theorems show that it suffices to find an extreme optimal solution of (3) to solve (1). We will therefore concentrate on the algorithm to obtain an extreme optimal solution of (3) throughout the remainder of this paper.

Let (x^*, y^*) be the current incumbent (the best solution identified to date) of (3) and let $\phi^* = \phi(x^*, y^*)$. Also let

$$C_I = \{w \in R^n \mid w \in \{0, 1\}^n\} \quad (4)$$

and let $X \subset C_I$ satisfy

$$\max \{\phi(x, y) \mid x \in C_I - X, y \in C_I\} \leq \phi^* \quad (5)$$

i.e., X is a subset of C_I in which a solution better than the current incumbent can possibly exist.

Define subproblems

$$K_x(y): \text{ maximize } \{\phi(x, y) \mid x \in X\} \quad (6)$$

where $y \in \{0, 1\}^n$ is fixed and

$$K(x): \text{ maximize } \{\phi(x, y) \mid y \in C_I\} \quad (7)$$

where $x \in X$ is fixed.

As the algorithm proceeds, cuts will be added to C_I in x -space, so that X will be some portion of C_I in general, whereas the feasible region of $K(x)$ is always C_I .

Alternate Mountain Climbing Procedure $AMC(X; y^0)$

Step 1: Compute an optimal solution x^0 of $K_x(y^0)$.

Step 2: Compute an optimal solution y^1 of $K(x^0)$.

Step 3: Compute an optimal solution x^1 of $K_x(y^1)$.

Step 4: If $\phi(x^1, y^1) > \phi(x^0, y^0)$, then let $x^0 = x^1$, $y^0 = y^1$ and go to Step 2. Otherwise, go to Step 5.

Step 5: If $\phi(x^1, y^1) > \phi^*$, then let $(x^*, y^*) = (x^1, y^1)$, $\phi^* = \phi(x^1, y^1)$ and halt. Otherwise, halt.

Theorem 3. If $y^0 \in C_I$, then the procedure $AMC(X; y^0)$ halts in finitely many steps generating a pair of points (\hat{x}, \hat{y}) for which the following equalities hold:

$$\begin{cases} \max \{ \phi(x, \hat{y}) \mid x \in X \} = \phi(\hat{x}, \hat{y}) \\ \max \{ \phi(\hat{x}, y) \mid y \in C_I \} = \phi(\hat{x}, \hat{y}) \end{cases} \quad (8)$$

Proof. X as well as C_I contains finitely many points, whereas the value of ϕ strictly increases at each cycle, so that the procedure eventually halts at Step 5 and $\hat{x} = x^1$, $\hat{y} = y^0$ satisfy (8). ||

The pair of points (\hat{x}, \hat{y}) for which condition (8) holds will be called a stationary pair. Note that the subproblem $K(x)$ for fixed x , i.e.,

$$\begin{cases} \text{maximize } c^t x + (c + Qx)^t y \\ \text{subject to } y \in \{0, 1\}^n \end{cases} \quad (9)$$

can be solved by inspection and an optimal solution y will be given by:

$$y_i = \begin{cases} 1 & \text{if } c_i + \sum_{j=1}^n q_{ij} x_j > 0, \\ 0 & \text{otherwise} \end{cases} \quad i = 1, \dots, n \quad (10)$$

$K_x(y)$, on the other hand, is a general 0-1 integer linear program after at least one cut is added to C_I in x -space, so that it has to be solved by some version of 0-1 integer programming algorithm [3, 5]. However, we need not always solve $K_x(y)$ to the optimum. What we need in Step 3 is essentially a pair of points which is strictly better than the previous pair of points and we can go back to Step 2 as soon as a point $x^1 \in X$ for which $\phi(x^1, y^1) > \phi(x^0, y^0)$ is detected.

Proposition 4. The inequality

$$\phi(x, y) \leq \{\phi(x, x) + \phi(y, y)\}/2 \tag{11}$$

holds for all $x, y \in \mathbb{R}^n$.

Proof. Assume the contrary. Then

$$0 < 2\phi(x, y) - \phi(x, x) - \phi(y, y) = -(x - y)^t Q(x - y)$$

a contradiction to the assumption that Q is positive semi-definite. ||

Proposition 5. If y^1 obtained in Step 2 satisfies $\phi(x^0, y^1) > \phi(x^0, y^0)$, then $\phi(y^1, y^1) > \phi(x^0, y^0)$.

Proof. $\phi(x^0, y^1) = \phi(y^1, x^0) \geq \phi(x^0, x^0)$ since y^1 is optimal to $K(x^0)$. It follows from Proposition 4 that $\phi(y^1, y^1) \geq \phi(x^0, y^1)$ and thus $\phi(y^1, y^1) > \phi(x^0, y^0)$. ||

Therefore, if $y^1 \in X$ in Step 2, then a new pair of solutions (y^1, y^1) is strictly better than (x^0, y^0) and Step 3 can be skipped, which is the most time consuming part of this algorithm.

Upon reaching a stationary pair (\hat{x}, \hat{y}) , it is no longer possible to improve it by fixing the value of either x or y , so that we will switch to a local (relative to x -space) and global (relative to y -space) search procedure.

Let $x^{(i)}$ be the i -th complement of $x \in C_I$, i.e.,

$$x^{(i)} = (x_1, \dots, x_{i-1}, 1-x_i, x_{i+1}, \dots, x_n) \tag{12}$$

and let

$$I(\mathbf{x}) = \{i \mid x^{(i)} \in X\} \quad (13)$$

Definition. A stationary pair $(\hat{\mathbf{x}}, \hat{\mathbf{y}})$ will be called an LG maximum if the following inequality holds for all $i \in I(\hat{\mathbf{x}})$

$$\max \{\phi(\hat{\mathbf{x}}^{(i)}, y) \mid y \in C_I\} \leq \phi(\hat{\mathbf{x}}, \hat{\mathbf{y}}). \quad (14)$$

Note again that the left hand side of (14) can be obtained by inspection.

LG Maximization Procedure LGM(X)

Step 0: Choose $y^0 \in C_I$ arbitrarily.

Step 1: Execute $\text{AMC}(X, y^0)$ and let $(\hat{\mathbf{x}}, \hat{\mathbf{y}})$ be a stationary pair.

Step 2: If $(\hat{\mathbf{x}}, \hat{\mathbf{y}})$ is an LG maximum, then halt. Otherwise let $\tilde{y} \in C_I$ be such a point that $\phi(\hat{\mathbf{x}}^{(i)}, \tilde{y}) > \phi(\hat{\mathbf{x}}, \hat{\mathbf{y}})$ for some $i \in I(\hat{\mathbf{x}})$. Let $y^0 = \tilde{y}$ and go to Step 1.

Theorem 6. Procedure LGM(X) where $X \neq \emptyset$ generates an LG maximum in finitely many steps.

Proof. Similar to the proof of Theorem 3. ||

3. Cutting Planes from an LG Maximum

Let $(\hat{\mathbf{x}}, \hat{\mathbf{y}})$ be an LG maximum and let

$$I_0(\hat{\mathbf{x}}) = \{i \mid \hat{x}_i = 0\}, \quad I_1(\hat{\mathbf{x}}) = \{i \mid \hat{x}_i = 1\} \quad (15)$$

$$J_0(\hat{\mathbf{y}}) = \{j \mid \hat{y}_j = 0\}, \quad J_1(\hat{\mathbf{y}}) = \{j \mid \hat{y}_j = 1\} \quad (16)$$

We want to obtain a cut which eliminates $\hat{\mathbf{x}}$ in \mathbf{x} -space and yet does not eliminate a point in X which can possibly generate a better solution than the current incumbent.

Let

$$u_i = \begin{cases} x_i & i \in I_0(\hat{\mathbf{x}}) \\ 1-x_i & i \in I_1(\hat{\mathbf{x}}) \end{cases} \quad (17)$$

$$v_j = \begin{cases} y_j & j \in J_0(\hat{y}) \\ 1-y_j & j \in J_1(\hat{y}) \end{cases} \quad (18)$$

so that \hat{x} and \hat{y} are now the origin of u -space and v -space, respectively, and the unit hypercube in x -space and y -space are again unit hypercube in u -space and v -space. Let $\psi(u, v)$ be the representation of $\phi(x, y)$ relative to u and v , i.e.,

$$\psi(u, v) = \sum_{i=1}^n \gamma_i u_i + \sum_{j=1}^n \delta_j v_j + \sum_{i=1}^n \sum_{j=1}^n \xi_{ij} u_i v_j + \phi(\hat{x}, \hat{y}) \quad (19)$$

Lemma 7.

$$\delta_j \leq 0 \text{ for all } j.$$

Proof. Assume that there exists j for which $\delta_j > 0$. Then ψ can be made strictly greater than $\phi(\hat{x}, \hat{y})$ by taking $v_j = 1$, contradicting the definition of LG maximality. ||

Let us define a function $g_i: [0, \infty) \rightarrow \mathbb{R}^1$

$$g_i(\lambda) = \max_{u, v} \{ \psi(u, v) \mid 0 \leq u \leq \lambda e_i, v \in C_I \} \quad (20)$$

where λ is a nonnegative scalar and e_i is the i -th unit vector.

Lemma 8.

$$g_i \text{ is convex on } [0, \infty).$$

Proof. By linearity of ψ with respect to u ,

$$\begin{aligned} g_i(\lambda) &= \max \{ \psi(u, v) \mid 0 \leq u \leq \lambda e_i, v \in C_I \} \\ &= \max [\max \{ \psi(0, v) \mid v \in C_I \}, \max \{ \psi(\lambda e_i, v) \mid v \in C_I \}] \end{aligned} \quad (21)$$

The first term in the bracket is a constant and it now suffices to prove that

$$h_i(\lambda) = \max \{ \psi(\lambda e_i, v) \mid v \in C_I \}$$

is convex on $[0, \infty)$. Let $\lambda_1, \lambda_2 \in [0, \infty)$ and let $\theta \in [0, 1]$. Then

$$\begin{aligned} h(\theta\lambda_1 + (1-\theta)\lambda_2) &= \max \{ \psi((\theta\lambda_1 + (1-\theta)\lambda_2)e_{\mathbf{1}}, v) \mid v \in C_{\mathbf{I}} \} \\ &= \max \{ \theta\psi(\lambda_1 e_{\mathbf{1}}, v) + (1-\theta)\psi(\lambda_2 e_{\mathbf{1}}, v) \mid v \in C_{\mathbf{I}} \} \\ &\leq \theta \max \{ \psi(\lambda_1 e_{\mathbf{1}}, v) \mid v \in C_{\mathbf{I}} \} + (1-\theta) \max \{ \psi(\lambda_2 e_{\mathbf{1}}, v) \mid v \in C_{\mathbf{I}} \} \\ &= \theta h(\lambda_1) + (1-\theta)h(\lambda_2). \end{aligned}$$

as was to be proved. $\quad ||$

Let us define

$$\hat{\lambda}_{\mathbf{i}} = \sup \{ \lambda \mid g_{\mathbf{i}}(\lambda) \leq \phi^*, \lambda \geq 0 \}, \quad (22)$$

which corresponds to the farthest point we can go along the $u_{\mathbf{i}}$ -axis without generating a pair of points u and $v \in C_{\mathbf{I}}$ for which ψ is greater than ϕ^* .

Theorem 9.

$$g_{\mathbf{i}}(\lambda) \leq \phi^* \text{ for all } \lambda \in [0, 1] \text{ and for all } \mathbf{i}.$$

Proof. By Lemma 7,

$$g_{\mathbf{i}}(0) = \max \{ \psi(0, v) \mid v \in C_{\mathbf{I}} \} = \phi(\hat{x}, \hat{y}) \leq \phi^*, \quad \forall \mathbf{i}$$

It now suffices to prove that $g_{\mathbf{i}}(1) \leq \phi^*$ for all \mathbf{i} since $g_{\mathbf{i}}$ is known to be convex. We have, by (21)

$$g_{\mathbf{i}}(1) = \max [g_{\mathbf{i}}(0), \max \{ \psi(e_{\mathbf{i}}, v) \mid v \in C_{\mathbf{I}} \}]$$

If $\mathbf{i} \in I(\hat{x})$, then by the definition of LG maximality

$$\max \{ \psi(e_{\mathbf{i}}, v) \mid v \in C_{\mathbf{I}} \} \leq \phi(\hat{x}, \hat{y}) \leq \phi^*$$

Also, if $\mathbf{i} \notin I(\hat{x})$, then $e_{\mathbf{i}}$ has been cut off by previously added cuts, which implies that

$$\max \{ \psi(e_{\mathbf{i}}, v) \mid v \in C_{\mathbf{I}} \} \leq \phi^*$$

This establishes $g_{\mathbf{i}}(1) \leq \phi^*$ for all \mathbf{i} . $\quad ||$

Corollary 10.

$$\hat{\lambda}_{\mathbf{i}} \geq 1, \text{ for all } \mathbf{i}.$$

Proof. Follows from Theorem 9 and from (22). $\quad ||$

Let

$$\Delta(\hat{\lambda}) = \{u \in \mathbb{R}^n \mid \sum_{i=1}^n u_i / \hat{\lambda}_i \leq 1, u \geq 0\}. \tag{23}$$

Theorem 11. The following inequality holds.

$$\max \{\psi(u, v) \mid u \in \Delta(\hat{\lambda}), v \in C_I\} \leq \phi^* \tag{24}$$

Proof. Fix $u \in \Delta(\hat{\lambda})$. Then u can be expressed as

$$u = \sum_{i \in I} \theta_i \Lambda_i + \sum_{j \in J} u_j e_j, \quad \sum_{i \in I} \theta_i \leq 1, \quad \theta_i \geq 0, \quad \forall i \in I; \quad u_j \geq 0, \quad \forall j \in J$$

where $I = \{i \mid \hat{\lambda}_i < \infty\}$, $J = \{j \mid \hat{\lambda}_j = \infty\}$ and $\Lambda_i = (0, \dots, 0, \hat{\lambda}_i, 0, \dots, 0)$.

Hence,

$$\begin{aligned} \max \{\psi(u, v) \mid v \in C_I\} &= \max \{\psi(\sum_{i \in I} \theta_i \Lambda_i + \sum_{j \in J} u_j e_j, v) \mid v \in C_I\} \\ &\leq \sum_{i \in I} \theta_i \max \{\psi(\Lambda_i, v) \mid v \in C_I\} \\ &\quad + \sum_{j \in J} u_j \max \{\psi(e_j, v) \mid v \in C_I\} \\ &\leq \sum_{i \in I} \theta_i \phi^* \\ &\leq \phi^* \end{aligned}$$

The second inequality follows from the fact that $\max \{\psi(e_j, v) \mid v \in C_I\} \leq 0$ for $j \in J$. (Otherwise, $\hat{\lambda}_j$ would have been finite). $\mid \mid$

Theorem 11 establishes that the cut

$$\sum_{i=1}^n u_i / \hat{\lambda}_i \geq 1 \tag{25}$$

eliminates the origin of u -space (i.e., \hat{x} in x -space) and yet does not eliminate any $x \in X$ which can possibly generate a solution which is better than the current incumbent. Such a cut will be called "valid".

Let us now describe how to compute $\hat{\lambda}_i$'s. Note that

$$g_i(\lambda) = \max [\phi(\hat{x}, \hat{y}), \max \{\psi(\lambda e_i, v) \mid v \in C_I\}]$$

Also, by (19) we have

$$\begin{aligned}
& \max \{ \psi(\lambda e_{\mathbf{I}}, v) \mid v \in C_{\mathbf{I}} \} \\
&= \max \{ \gamma_{\mathbf{I}} \lambda + \sum_{j=1}^n (\delta_j + \lambda \xi_{\mathbf{I}j}) v_j \mid v \in C_{\mathbf{I}} \} + \phi(\hat{\mathbf{x}}, \hat{\mathbf{y}}) \\
&= \gamma_{\mathbf{I}} \lambda + \sum_{j \in J_{\mathbf{I}}^+} (\delta_j + \lambda \xi_{\mathbf{I}j})_+ + \phi(\hat{\mathbf{x}}, \hat{\mathbf{y}}) \tag{26}
\end{aligned}$$

where

$$(a)_+ = \begin{cases} a & \text{if } a \geq 0 \\ 0 & \text{otherwise} \end{cases} \tag{27}$$

$$J_{\mathbf{I}}^+ = \{ j \mid \xi_{\mathbf{I}j} > 0 \} \tag{28}$$

Let

$$\lambda_{\mathbf{I}}^j = -\delta_j / \xi_{\mathbf{I}j} \quad j \in J_{\mathbf{I}}^+ \tag{29}$$

and renumber the indices of $J_{\mathbf{I}}^+$ in the increasing order of $\lambda_{\mathbf{I}}^j$, so that now

$$\lambda_{\mathbf{I}}^j \geq \lambda_{\mathbf{I}}^k \quad \text{for } \forall j, \forall k \text{ such that } j > k.$$

Then we have an analytical expression for (26)

$$\begin{aligned}
& \max \{ \psi(\lambda e_{\mathbf{I}}, v) \mid v \in C_{\mathbf{I}} \} \\
&= \gamma_{\mathbf{I}} \lambda + \sum_{j=1}^k (\delta_j + \lambda \xi_{\mathbf{I}j}) + \phi(\hat{\mathbf{x}}, \hat{\mathbf{y}}); \quad \lambda_{\mathbf{I}}^k \leq \lambda \leq \lambda_{\mathbf{I}}^{k+1}
\end{aligned}$$

Therefore,

$$g_{\mathbf{I}}(\lambda) = \phi(\hat{\mathbf{x}}, \hat{\mathbf{y}}) + \max \{ 0, [\gamma_{\mathbf{I}} \lambda + \sum_{j=1}^k (\delta_j + \lambda \xi_{\mathbf{I}j})] \}, \lambda_{\mathbf{I}}^k \leq \lambda \leq \lambda_{\mathbf{I}}^{k+1} \tag{30}$$

which is a piecewise linear convex function.

Cutting Plane Algorithm

Step 0: $X = C_{\mathbf{I}}$, $\phi^* = -\infty$.

Step 1. Execute LGM(X) to obtain an LG maximum $(\hat{\mathbf{x}}, \hat{\mathbf{y}})$.

Step 2: Compute $\hat{\lambda}_{\mathbf{I}}$'s and let

$$X = X \cap \{ \mathbf{x} \in \mathbb{R}^n \mid \sum_{i \in I_0} (\hat{\mathbf{x}})_i \hat{\lambda}_{\mathbf{I}} + \sum_{i \in I_1} (\hat{\mathbf{x}})_i (1 - \mathbf{x}_i) / \hat{\lambda}_{\mathbf{I}} \geq 1 \}$$

Step 3: If $X = \emptyset$, then stop. Otherwise go to Step 1.

Theorem 12. The cutting plane algorithm defined above generates an optimal solution of (3) in finitely many steps.

Proof. X contains at least one less extreme points of a unit hypercube whenever a new cut is added, so that X becomes empty after finitely many steps. Then the incumbent (x^*, y^*) and $\phi^* = \phi(x^*, y^*)$ gives the optimal solution of (3) by the validity of added cuts. ||

4. Stronger Cuts

We will consider here several techniques to obtain stronger cuts.

(a) Iterative Improvement

Let S be the subset of C_I for which the following inequality holds.

$$\max \{ \phi(x, y) \mid x \in S, y \in C_I \} \leq \phi^* \tag{31}$$

Also let $V(S)$ be the set $C_I - S$ represented in terms of v as defined by (18) and let

$$\lambda_i(S) = \sup \{ \lambda \mid G_i^S(\lambda) \leq \phi^*, \quad \lambda \geq 0 \} \tag{32}$$

where

$$G_i^S(\lambda) = \max \{ \psi(u, v) \mid 0 \leq u \leq \lambda e_i, \quad v \in V(S) \} \tag{33}$$

Theorem 13. $\sum_{i=1}^n u_i / \lambda_i(S) \geq 1$ is a valid cut. Also $\lambda_i(S) \geq \hat{\lambda}_i$ for all i .

Proof. We have

$$\begin{aligned} \max \{ f(w) \mid w \in S \} &= \max \{ \phi(x, y) \mid x \in S, y \in S \} \\ &\leq \max \{ \phi(x, y) \mid x \in S, y \in C_I \} \\ &\leq \phi^*. \end{aligned}$$

This inequality implies that S can be ignored when we search for a point $w \in C_I$ for which $f(w) > \phi^*$. Therefore all the results in the previous section go through by changing C_I to $C_I - S$, so that $\lambda_i(S)$'s define a valid

cut. $\lambda_{\mathbf{1}}(S) \geq \hat{\lambda}_{\mathbf{1}}$ follows from the fact that $G_{\mathbf{1}}^S(\lambda) \leq g_{\mathbf{1}}(\lambda)$ for all $\lambda \in [0, \infty)$ (note $V(S) \subset C_{\mathbf{1}}$). ||

To compute $\lambda_{\mathbf{1}}(S)$, we have to solve parametric 0-1 integer linear programs, which is a difficult task. The most practical way to implement this idea is to choose either one of the previously added valid cuts $\sum_{i=1}^n \alpha_i x_i \geq \alpha_0$ and to use $V(S^0)$ in the definition of $G_{\mathbf{1}}^S(\lambda)$ instead of $V(S)$ where

$$S^0 = \{x \in \mathbb{R}^n \mid \sum_{i=1}^n \alpha_i x_i \leq \alpha_0, x \in \{0, 1\}^n\}. \quad (34)$$

$\lambda_{\mathbf{1}}(S^0)$ gives an underestimate of $\lambda_{\mathbf{1}}(S)$, and therefore defines a valid cut.

This choice of S leads us to a parametric 0-1 integer linear program with a single constraint, which can be solved by modified Newton's approach discussed in [10, 15].

Theorem 14. Iterative improvement scheme using the cut generated at \hat{x} will either generate a uniformly deeper cut than Tuy's convexity cut directly applied to (1), or else generate a point $w \in C_{\mathbf{1}}$ for which $f(w) > \phi^*$ if Q is positive definite.

Proof. This theorem is the direct consequence of the one established in [8] for convex quadratic maximization problem in continuous variables, and will be omitted. (see [8] and [16]) ||

(b) Negative Edge Extension

An even deeper cut can be obtained by using negative edge extension [10, 12, 15] when some of $\hat{\lambda}_{\mathbf{1}}$'s are infinite. Let

$$L(\hat{x}) = \{i \mid \hat{\lambda}_{\mathbf{1}} < \infty\} \quad (35)$$

$$J(\hat{x}) = \{i \mid \hat{\lambda}_{\mathbf{1}} = \infty\} \quad (36)$$

and for $i \in J(\hat{x})$ let

$$H_{\mathbf{1}}(\mu) = \min \{\psi(u, v) \mid -\mu e_{\mathbf{1}} \leq u \leq 0, v \in C_{\mathbf{1}}\} \quad (37)$$

and

$$\hat{\mu}_1 = \sup \{ \mu \mid H_1(\mu) \leq \phi^*, \mu \geq 0 \} \tag{38}$$

Theorem 15. The following inequality defines a valid cut

$$\sum_{i \in L(\hat{x})} u_i / \hat{\lambda}_i - \sum_{i \in J(\hat{x})} u_i / \hat{\mu}_i \geq 1. \tag{39}$$

Proof. The proof of this theorem is similar to the one established in [10] and will be omitted. ||

5. Illustrative Examples

Let us consider 5 × 5 problem where

$$c = (-10, -4, 2, -8, 1) \tag{40}$$

$$Q = \begin{bmatrix} 9 & -5 & 1 & 5 & -1 \\ -5 & 11 & -5 & 1 & 5 \\ 1 & -5 & 9 & 5 & -9 \\ 5 & 1 & 5 & 11 & -5 \\ -1 & 5 & -9 & -5 & 9 \end{bmatrix} \tag{41}$$

Q is a symmetric positive semi-definite matrix.

Cycle 1: Starting with $x^0 = (1, 0, 0, 1, 0)$ we have

$$y^1 = \operatorname{argmax} \{ (c^t + (x^0)^t Q) y \} = \operatorname{argmax} \{ (4, -8, 8, 8, -5) y \} = (1, 0, 1, 1, 0)$$

$$\begin{aligned} x^1 &= \operatorname{argmax} \{ (c^t + (y^1)^t Q) x \} = \operatorname{argmax} \{ (5, -13, 17, 12, -14) x \} \\ &= (1, 0, 1, 1, 0) \end{aligned}$$

$$\begin{aligned} y^2 &= \operatorname{argmax} \{ (c^t + (x^1)^t Q) y \} = \operatorname{argmax} \{ (5, -13, 17, 12, -14) y \} \\ &= (1, 0, 1, 1, 0) \end{aligned}$$

so that $\hat{x} = \hat{y} = (1, 0, 1, 1, 0)$ is a stationary pair with $\phi(\hat{x}, \hat{y}) = 19$.

Also we have $(x^*, y^*) = (\hat{x}, \hat{y})$, $\phi^* = 19$. This (\hat{x}, \hat{y}) is an LG maximum since

$$\max_{y \in C_I} \{ c^t x^{(1)} + (c^t + (\hat{x}^1)^t Q) y \} = -6 + 16 + 8 = 18 < 19.$$

$$\max_{y \in C_I} \{ c^t x^{(2)} + (c^t + (\hat{x}^2)^t Q) y \} = -20 + 12 + 14 = 6 < 19.$$

$$\max_{y \in C_I} \{c^t x^{(3)} + (c^t + (\hat{x}^3)^t Q)y\} = -18 + 4 + 8 + 8 = 2 < 19.$$

$$\max_{y \in C_I} \{c^t x^{(4)} + (c^t + (\hat{x}^4)^t Q)y\} = -8 + 12 + 2 = 6 < 19.$$

$$\max_{y \in C_I} \{c^t x^{(5)} + (c^t + (\hat{x}^5)^t Q)y\} = -15 + 4 + 8 + 8 = 5 < 19.$$

Apply the formula

$$g_i(\lambda) = \begin{cases} c\hat{x} + c_i\lambda + \max_{y \in C_I} \{c^t + \hat{x}^t Q + \lambda q_{i.}\}y, & i \in I_0(\hat{x}) \\ c\hat{x} - c_i\lambda + \max_{y \in C_I} \{c^t + \hat{x}^t Q - \lambda q_{i.}\}y, & i \in I_1(\hat{x}) \end{cases}$$

where $q_{i.}$ is the i -th row of Q to obtain

$$g_1(\lambda) = -16 - 10\lambda + \max_{y \in C_I} (5-9\lambda, -13+5\lambda, 17-\lambda, 13-5\lambda, -14+\lambda)y.$$

$$g_2(\lambda) = -16 + 4\lambda + \max_{y \in C_I} (5-5\lambda, -13+11\lambda, 17-5\lambda, 13+\lambda, -14+5\lambda)y.$$

$$g_3(\lambda) = -16 - 2\lambda + \max_{y \in C_I} (5-\lambda, -13+5\lambda, 17-9\lambda, 13-5\lambda, -14+9\lambda)y.$$

$$g_4(\lambda) = -16 + 8\lambda + \max_{y \in C_I} (5-5\lambda, -13-\lambda, 17-5\lambda, 13-11\lambda, -14+5\lambda)y.$$

$$g_5(\lambda) = -16 + \lambda + \max_{y \in C_I} (5-\lambda, -13+5\lambda, 17-9\lambda, 13-5\lambda, -14+9\lambda)y.$$

Solving equations $g_i(\lambda) = 19$ for $i = 1, \dots, 5$, we obtain

$$\hat{\lambda} = (5/4, 49/13, 31/6, 49/13, 57/14). \quad (42)$$

The cut corresponding to $\hat{\lambda}$ in x -space is:

$$\frac{4}{5}(1-x_1) + \frac{13}{49}x_2 + \frac{6}{31}(1-x_3) + \frac{13}{49}(1-x_4) + \frac{14}{57}x_5 \geq 1 \quad (43)$$

so that now we have

$$X = \{x \mid 0.8x_1 - 0.265x_2 + 0.194x_3 + 0.265x_4 - 0.246x_5 \leq 0.259$$

$$x_i \in \{0, 1\}, \quad i=1, \dots, 5\}$$

Cycle 2: Starting from a point $x^0 = (0, 0, 0, 0, 1)$

$$y^1 = \operatorname{argmax}_{y \in C_I} \{(c^t + (x^0)^t Q)y\} = \operatorname{argmax}_{y \in C_I} \{(-11, 1, -7, -13, 10)y\}$$

$$= (0, 1, 0, 0, 1)$$

$$\begin{aligned} x^1 &= \operatorname{argmax}_{y \in C_I} \{(c^t + (y^1)^t Q)x\} = \operatorname{argmax}_{y \in C_I} \{(-16, 12, -12, -12, 15)x\} \\ &= (0, 1, 0, 0, 1) \end{aligned}$$

$$\begin{aligned} y^2 &= \operatorname{argmax}_{y \in C_I} \{(c^t + (x^1)^t Q)y\} = \operatorname{argmax}_{y \in C_I} \{(-16, 12, -12, -12, 15)y\} \\ &= (0, 1, 0, 0, 1) \end{aligned}$$

which implies that $\hat{x} = \hat{y} = (0, 1, 0, 0, 1)$ is a stationary point with $\phi(\hat{x}, \hat{y}) = 24$, so that we have $x^* = y^* = (0, 1, 0, 0, 1)$, $\phi^* = 24$. (\hat{x}, \hat{y}) is an LG maximum as can be checked easily. $g_i(\lambda)$, $i = 1, \dots, 5$ are given below.

$$g_1(\lambda) = -3 - 10\lambda + \max_{y \in C_I} \{(-16+9\lambda, 12-5\lambda, -12+\lambda, -12+5\lambda, 15-\lambda)y\}$$

$$g_2(\lambda) = -3 + 4\lambda + \max_{y \in C_I} \{(-16+5\lambda, 12-11\lambda, -12+5\lambda, -12-\lambda, 15-5\lambda)y\}$$

$$g_3(\lambda) = -3 + 2\lambda + \max_{y \in C_I} \{(-16+\lambda, 12-5\lambda, -12+9\lambda, -12+5\lambda, 15-9\lambda)y\}$$

$$g_4(\lambda) = -3 - 8\lambda + \max_{y \in C_I} \{(-16+5\lambda, 12+\lambda, -12+5\lambda, -12+11\lambda, 15-5\lambda)y\}$$

$$g_5(\lambda) = -3 - \lambda + \max_{y \in C_I} \{(-16+\lambda, 12-5\lambda, -12+9\lambda, -12+5\lambda, 15-9\lambda)y\}$$

We solve equations $g_i(\lambda) = 24$ and obtain

$$\hat{\lambda} = (13, 55/14, 51/16, 55/14, 51/13) \tag{44}$$

Hence the new cut to be added is

$$\frac{x_1}{13} + \frac{14}{55}(1-x_2) + \frac{16}{51}x_3 + \frac{14}{55}x_3 + \frac{13}{51}(1-x_5) \geq 1$$

or

$$0.077x_1 - 0.255x_2 + 0.314x_3 + 0.255x_4 - 0.255x_5 \geq 0.490 \tag{45}$$

so that now

$$X = \{x \mid 0.800x_1 - 0.265x_2 + 0.194x_3 + 0.265x_4 - 0.246x_5 \leq 0.259,$$

$$0.077x_1 - 0.255x_2 + 0.314x_3 + 0.255x_4 - 0.255x_5 \geq 0.490,$$

$$x_i \in \{0, 1\}, \quad i = 1, \dots, 5\} \tag{46}$$

The second cut implies that $x_3 = x_4 = 1$ so that the rest of the variables have to satisfy

$$-0.800x_1 + 0.317x_2 + 0.246x_5 \geq 0.195,$$

$$-0.077x_1 + 0.255x_2 + 0.255x_5 \leq 0.079$$

which is obviously contradictory, i.e., $X = \emptyset$. Thus we conclude that $x^* = y^* = (0, 1, 0, 0, 1)$ is an optimal solution with $\phi(x^*, y^*) = 24$.

If we use

$$V(S) = \{y \mid 0.8y_1 - 0.265y_2 + 0.194y_3 + 0.265y_4 - 0.246y_5 \leq 0.259 \\ y_i \in \{0, 1\}, \quad i = 1, \dots, 5\} \quad (47)$$

instead of C_I to compute $\lambda_i(S)$'s, we would have

$$\lambda_1(S) = \infty, \lambda_2(S) = 43/9, \lambda_3(S) = 51/16, \lambda_4(S) = 13/3, \lambda_5(S) = 51/13,$$

resulting in a cut

$$-0.209x_2 + 0.314x_3 + 0.230x_4 - 0.255x_5 \geq 0.536. \quad (48)$$

which is much deeper than (45). Let us compare this cut with Tuy's convexity cut applied to original problem (1) at its local maximum point $w^* = (0, 1, 0, 0, 1)$. Here we will define

$$h_i(\lambda) = \begin{cases} 2c^t(\hat{x} + \lambda e_i) + (\hat{x} + \lambda e_i)^t Q(\hat{x} + \lambda e_i), & i \in I_0(\hat{x}) \\ 2c^t(\hat{x} - \lambda e_i) + (\hat{x} - \lambda e_i)^t Q(\hat{x} - \lambda e_i), & i \in I_1(\hat{x}) \end{cases}$$

and the cut coefficients $\bar{\lambda}_i$ is defined by

$$\bar{\lambda}_i = \max \{ \lambda \mid h_i(\lambda) \leq f(w^*) \}.$$

The result is shown below:

$$\bar{\lambda} = (32/9, 24/11, 24/9, 24/11, 10/3)$$

which is smaller than $\hat{\lambda}$.

References

- [1] Balas, E. and C. A. Burdet.: Maximizing a Convex Quadratic Function Subject to Linear Constraints. MSRR 299, GSIA, Carnegie-Mellon University, 1973.
- [2] Balas, E.: Intersection Cuts - A New Type of Cutting Planes for Integer Programming. *Operations Research*, Vol. 19 (1971), pp. 19-39.
- [3] Balas, E.: Additive Algorithm for Solving Linear Programs with Zero-One Variables. *Operations Research*, Vol. 13 (1965), pp. 517-546.
- [4] Falk, J. E. and K. J. Hoffmann.: A Successive Underestimation Method for Concave Minimization Problems. T-316, Institute of Management Science and Engineering, George Washington University, 1975.
- [5] Geoffrion, A. M.: An Improved Implicit Enumeration Approach for Integer Programming. *Operations Research*, Vol. 17 (1969), pp. 437-454.
- [6] Glover, F.: Polyhedral Annexation in Mixed Integer and Combinatorial Programming. *Mathematical Programming*, Vol. 9 (1975), pp. 161-188.
- [7] Kaufman, L. and J. R. Bunch.: A Numerical Algorithm for Solving Non-Positive Definite Quadratic Programming Problems. Presented at X International Symposium on Mathematical Programming, 1979.
- [8] Konno, H.: Maximization of a Convex Quadratic Function under Linear Constraints. *Mathematical Programming*, Vol. 11 (1976), pp. 117-127.
- [9] Konno, H.: A Cutting Plane Algorithm for Solving Bilinear Programs. *Mathematical Programming*, Vol. 11 (1976), pp. 14-27.
- [10] Konno, H.: An Algorithm for Solving Bilinear Knapsack Problems. EIS-TR-79-2, Institute of Information Sciences, University of Tsukuba, 1979.
- [11] Majthay, A. and A. B. Whinston.: Quasi-Concave Minimization Subjects to Linear Constraints. *Discrete Mathematics*, Vol. 9 (1974), pp. 35-59.

- [12] Owen, G.: Cutting Planes for Programs with Disjunctive Constraints. *J. of Optimization Theory and Applications*, Vol. 11 (1973), pp. 49-55.
- [13] Raghavachari, M.: On Connections Between 0-1 Integer Programming and Concave Programming. *Operations Research*, Vol. 17 (1969), pp. 680-684.
- [14] Ritter, K.: A Method for Solving Maximum Problems with a Nonconcave Quadratic Objective Function. *Z. fur Wahrscheinlichkeits theorie und verwandte Gebiete*, Vol. 4 (1966), pp. 340-351.
- [15] Shetty, C. M. and H. D. Sherali.: A Finitely Convergent Algorithm for Bilinear Programming Problems Using Polar Cuts and Disjunctive-Face Cuts. to appear in *Mathematical Programming*.
- [16] Tuy, H.: Concave Programming Under Linear Constraints. *Soviet Mathematics*, (1964), pp. 1437-1440.

Hiroshi KONNO: Institute of
Information Sciences,
University of Tsukuba
Sakuramura, Niiharigun
Ibaraki, 305, Japan