A Continuous Method for Computing Bounds in Integer Quadratic Optimization Problems

A. KAMATH and N. KARMARKAR
AT&T Bell Laboratories, Murray Hill, NJ 07974, U.S.A.

(Received: 11 November 1991; accepted: 21 November 1991)

Abstract. In the graph partitioning problem, as in other NP-hard problems, the problem of proving the existence of a cut of given size is easy and can be accomplished by exhibiting a solution with the correct value. On the other hand proving the non-existence of a cut better than a given value is very difficult. We consider the problem of maximizing a quadratic function $\mathbf{x}^T Q\mathbf{x}$ where Q is an $n \times n$ real symmetric matrix with $\mathbf{x}$ an n-dimensional vector constrained to be an element of $\{-1, 1\}^n$. We had proposed a technique for obtaining upper bounds on solutions to the problem using a continuous approach in [4]. In this paper, we extend this method by using techniques of differential geometry.

Key words. Bounds, interior-point, integer quadratic optimization, Riemannian geometry.

0. Outline

In the first section we introduce the quadratic optimization problem and give a motivation for the concepts underlying the development of the interior point approach for solving the problem. In the second section, we transform the problem into a continuous optimization problem. We also formulate a Riemannian metric defined on the feasible region of the optimization problem. We then describe a method of generating a sequence of decreasing upper bounds for the quadratic maximization problem. Conclusions and directions for future work are presented in the last section.

1. Introduction

Let S be the set of n-dimensional vectors defined as follows

$$S = \{\mathbf{x} = (x_1, \ldots, x_n) \,|\, x_i \in \{-1, 1\}\,, \quad i = 1, \ldots, n\}\,.$$

Consider the quadratic optimization problem

$$\begin{aligned} &\max \quad f(\mathbf{x}) \\ &\text{s.t.} \quad \mathbf{x} \in S \end{aligned}$$

where $f(\mathbf{x}) = \mathbf{x}^T Q\mathbf{x}\,, \quad Q \in R^{n \times n}$ is a symmetric matrix. (1)

Let $f_{\max}$ denote the maximum value of $f(\mathbf{x})$ in problem (1).

Since Q is a real symmetric matrix all its eigenvalues are real. If all the eigenvalues are negative the problem is easy [3], [11]. Otherwise problem (1) is

© 1992 *Kluwer Academic Publishers. Printed in the Netherlands.*

NP-hard [2] and obtaining an upper bound on $f_{\max}$ by finding the optimum solution to (1) is difficult. We consider only the NP-hard case in which at least one eigenvalue of Q is positive [12] and we shall present in this paper a technique for finding a good upper bound on $f_{\max}$.

Evaluating $f(\mathbf{x})$ at any $\mathbf{x} \in S$ gives us a lower bound on the solution to this problem. However, to show that the optimal value of $f(\mathbf{x})$ in S, can be no more than some real number μ using a combinatorial approach, one may have to evaluate $f(\mathbf{x})$ for all $\mathbf{x}$ in S. In contrast we present here an efficient technique based on the interior point approach, for obtaining upper bounds on solutions to this maximization problem.

The main concept in an interior point approach to such problems is to embed the discrete set (S in the current case) in a continuous set T where $S \subseteq T$. Maximum of a function $f(\mathbf{x})$ over T gives an upper bound on $f(\mathbf{x})$ in S. We would like to choose set T so that not only can the maximum be found in a computationally efficient manner but also the bound obtained would be as good as possible.

A commonly used approach is to choose the continuous set to be the box X defined as follows

$$X = \{\mathbf{x} \in R^n \mid -1 \leq x_i \leq 1\,, \quad i = 1, \ldots, n\}\,.$$

But optimizing a quadratic function $f(\mathbf{x})$ over X is still NP-hard [12] so instead we enclose the box in a ball B where

$$B = \{\mathbf{x} \in R^n \mid \mathbf{x}^T\mathbf{x} \leq n\}\,.$$

For maximizing $f(\mathbf{x})$ on B, we need to just find the maximum eigenvalue $\lambda_{\max}$ of Q. Then since

$$\frac{\mathbf{x}^T Q \mathbf{x}}{\mathbf{x}^T \mathbf{x}} \leq \lambda_{\max}\,, \quad \forall \mathbf{x} \in R^n - \{\emptyset\}\,,$$

$$\mathbf{x}^T Q \mathbf{x} \leq n\lambda_{\max}\,, \quad \forall \mathbf{x} \in B$$

and therefore we get $n\lambda_{\max}$ to be an upper bound on $f_{\max}$.

Instead of the ball B we may choose any ellipsoid E enclosing the box X and still retain the computational ease of the resulting maximization problem. In this paper we shall limit ourselves to ellipsoids $E(\mathbf{w})$ whose axes are along the coordinate axes. A technique for ellipsoids with some other given set of axes can be derived along similar lines.

$$\text{Let } U\left\{\mathbf{w} = (w_1, w_2, \ldots, w_n) \in R^n \,\middle|\, \sum_{i=1}^{n} w_i = 1,\ w_i \geq 0,\ i = 1, \ldots, n\right\},$$
$$\text{and } E(\mathbf{w}) = \{\mathbf{x} \in R^n \mid \mathbf{x}^T W \mathbf{x} \leq 1 \text{ where } W = \operatorname{diag}(\mathbf{w}),\ \mathbf{w} \in U\}\,.$$

It can easily be verified that the ellipsoid E encloses the box X.

If λ is the maximum eigenvalue of the matrix $W^{-1/2} Q W^{-1/2}$ then

$$\max_x \frac{\mathbf{x}^T Q \mathbf{x}}{\mathbf{x}^T W \mathbf{x}} = \max_y \frac{\mathbf{y}^T W^{-1/2} Q W^{-1/2} \mathbf{y}}{\mathbf{y}^T \mathbf{y}} = \lambda \tag{2}$$

implying

$$\mathbf{x}^T Q\mathbf{x} \leq \lambda\,, \quad \forall \mathbf{x} \in E(\mathbf{w})\,.$$

This gives us a simple technique for maximizing $f(\mathbf{x})$ on $E(\mathbf{w})$.

Since $E(\mathbf{w})$ contains the set S, λ is an upper bound on $f_{\max}$. We also note that λ is a function of the weights $\mathbf{w}$. This suggests that the value of λ may be reducible by suitably modifying the weights $\mathbf{w}$. Hence to obtain a better bound on $f_{\max}$, we need to find the right values for the weights $\mathbf{w}$ that minimize λ. In the following section we shall use this idea to construct a method for computing better bounds on $f_{\max}$.

2. Interior Point Approach to the Problem

Our interior point approach to the problem is an iterative method which can be roughly described as follows. We start with some weights $\mathbf{w}^{(0)} = (w_1^0, w_2^0, \ldots, w_n^0) \in U$ and iteratively modify them, obtaining in the process a sequence of ellipsoids $E^{(1)}, \ldots, E^{(k)}$ such that if $\mu^{(i)}$ is maximum of $f(\mathbf{x})$ over $E^{(i)}$ then

$$\mu^{(1)} > \mu^{(2)} > \cdots > \mu^{(k)}\,.$$

The condition that μ is the maximum of $f(\mathbf{x})$ over $E(\mathbf{w})$ can be written as

$$\frac{\mathbf{x}^T Q\mathbf{x}}{\mathbf{x}^T W\mathbf{x}} \leq \mu\,, \quad \forall x \in R^n - \{\mathbf{0}\}\,,$$

which can be rephrased as

$$\mathbf{x}^T(\mu W - Q)\mathbf{x} \geq 0\,, \quad \forall \mathbf{x} \in R^n \tag{3}$$

implying that the matrix $\mu W - Q$ is positive semidefinite.

The problem to be solved can now be posed as

$$\begin{aligned} \min \quad & \mu \\ \text{s.t.} \quad & \mathbf{w} \in U \\ & \mathbf{x}^T(\mu W - Q)\mathbf{x} \geq 0\,, \quad \forall \mathbf{x} \in R^n\,. \end{aligned} \tag{4}$$

It can be proved that the level sets of the above optimization problem are convex.

We shall now modify the problem space to simplify the minimization problem.

$$\text{Let } V = \left\{ \mathbf{d} = (d_1, d_2, \ldots, d_n) \in R^n \,\middle|\, \sum_{i=1}^n d_i = 0 \right\}.$$

LEMMA 1. *Let* $\mathbf{w} \in U$ *and* $\mathbf{d} \in V$, *then given*

$$W = \text{diag}(\mathbf{w}) \text{ and } D = \text{diag}(\mathbf{d})\,,$$

$$\frac{\mathbf{x}^T(Q - D)\mathbf{x}}{\mathbf{x}^T W\mathbf{x}} \leq \mu\,, \quad \forall \mathbf{x} \in R^n - \{0\}$$

$$\Rightarrow \mathbf{x}^T Q\mathbf{x} \leq 0\,, \quad \forall \mathbf{x} \in S\,. \tag{5}$$

Proof. Since $\forall \mathbf{x} \in S$, $\mathbf{x}_i \in \{-1, 1\}$, $i = 1, \ldots, n$, we have

$$\mathbf{x}^T W \mathbf{x} = \sum w_i = 1$$

and since $\mathbf{d} \in V$ we have

$$\mathbf{x}^T D \mathbf{x} = \sum d_i = 0 .$$

Hence

$$\mathbf{x}^T Q \mathbf{x} \leq \mu , \quad \forall \mathbf{x} \in S . \qquad \blacksquare$$

This suggests that we have an additional degree of freedom in minimizing the bound $f_{\max}$ on S. So we may define the optimization problem over $(\mu, \mathbf{w}, \mathbf{d})$ space. But before doing so let us consider the inequality (5) which is the only constraint that has both the $\mathbf{w}$ and $\mathbf{d}$ variables. This constraint may be rephrased as

$$\mathbf{x}^T(\mu W + D - Q)\mathbf{x} \geq 0 , \quad \forall \mathbf{x} \in R^n .$$

Let us define for $\mathbf{w} \in U$ and $\mathbf{d} \in V$

$$M(\mu, \mathbf{w}, \mathbf{d}) = \mu W + D - Q .$$

For any given $(\mu, \mathbf{w}, \mathbf{d})$, we can define $(\mu, \mathbf{w}', \mathbf{d}')$ where $\mathbf{w}' = \frac{1}{n}\mathbf{e}$ and $\mathbf{d}' = \mu \mathbf{w} + \mathbf{d} - \frac{\mu}{n}\,\mathbf{e}$.

It is obvious that $\mathbf{w}' \in U$ and $\mathbf{d}' \in V$ and that

$$\begin{aligned} M\Big(\mu, \frac{1}{n}\,\mathbf{e}, \mathbf{d}'\Big) &= \frac{\mu}{n}\, I + \mu W + D - \frac{\mu}{n}\, I - Q \\ &= M(\mu, \mathbf{w}, \mathbf{d}) . \end{aligned}$$

Hence it suffices to treat M as a function of only μ and $\mathbf{d}$, keeping $\mathbf{w}$ constant at $\frac{1}{n}\,\mathbf{e}$ and varying $\mathbf{d}$ iteratively to reduce the upper bound μ. Better still we define a new vector $\mathbf{z} = \mu \mathbf{w} + \mathbf{d}$. We note that $\mathbf{e}^T \mathbf{z} = \mu$ and the optimization problem gets simplified to the following

$$\begin{aligned} &\min \quad \mathbf{e}^T \mathbf{z} \\ &\text{s.t.} \quad \mathbf{x}^T(Z - Q)\mathbf{x} \geq 0 , \quad \forall \mathbf{x} \in R^n , \\ &\text{where } Z = \operatorname{diag}(\mathbf{z}) . \end{aligned} \tag{6}$$

Let $M = Z - Q$. It is apparent from above, that the optimization has to be done over a region in which $\mathbf{M}$ is positive semidefinite. If we were to enforce the positive semidefiniteness condition for each $\mathbf{x}$ in R^n, it would require an infinite number of inequalities. The same effect can be achieved by imposing conditions on eigenvalues of M.

We shall use the notation $\lambda_i(M)$ to denote the ith eigenvalue of a real symmetric matrix M. Since M is a real and symmetric matrix, it has n real eigenvalues $\lambda_i(M)$, $i = 1, \ldots, n$. For M to be positive semi-definite, its eigenvalues must be non-negative which give us the inequality constraints

$$\lambda_i(M) \geqslant 0\,, \quad i = 1, \ldots, n\,. \tag{7}$$

The optimization problem may now be finally stated as follows:

$$\begin{aligned} &\min \quad \mathbf{e}^T\mathbf{z} \\ &\text{s.t.} \quad \lambda_i(M) \geqslant 0\,, \quad i = 1, \ldots, n\,. \end{aligned} \tag{8}$$

RIEMANNIAN METRIC

In the projective algorithm for linear programming, the optimization is done over a simplex and at each iteration of the method, we bring the current interior point to the center of the simplex by making a suitable projective transformation.

We want to adopt a similar approach towards solving this problem. We consider the set of real $n \times n$ symmetric positive definite matrices.

$$P = \{A \in R^{n \times n} \mid A = A^T\,, \quad \mathbf{x}^T A \mathbf{x} > 0\,, \quad \forall \mathbf{x} \in R^n - \{\mathbf{0}\}\}\,.$$

This set is in fact the interior of the set of real $n \times n$ symmetric positive semidefinite matrices, which is defined similarly except that $\mathbf{x}^T A \mathbf{x}$ is constrained to be non-negative rather than strictly positive. Another property to be noted, is that a matrix A in P has a unique square-root in P which we shall denote by $A^{1/2}$.

We define a transformation L_S: $P \to P$, where S is an $n \times n$ invertible matrix. The transformation is defined as

$$L_S(X) = SXS^T\,.$$

The transformation can be easily shown to be one-one and onto. By varying S over the set of $n \times n$ invertible matrices, we get a family of transformations which we denote by G. Composition of transformations in G are given as follows

$$\begin{aligned} L_S \circ L_T(X) &= STXT^TS^T \\ &= L_{ST}(X)\,. \end{aligned}$$

Also, the inverse of the transformation L_S is given as follows

$$\begin{aligned} L_S^{-1}(X) &= S^{-1}X(S^{-1})^T \\ &= L_{S^{-1}}(X)\,. \end{aligned}$$

It can be shown that these transformations form a Lie group with respect to the composition operation '$\circ$'. The identity of the Lie group is of course L_I which maps all positive definite matrices to themselves.

For any given matrix in P, there exists a unique transformation in G which maps the matrix to identity or vice versa. So if you need to map a positive definite matrix A to identity or the other way around then we need to use just the transforms $L_{A^{-1/2}}$ or $L_{A^{1/2}}$ as

$$L_{A^{-1/2}}(A) = I$$
$$\text{and } L_{A^{1/2}}(I) = A\,.$$

As in linear programming, we would like to define a centering transformation. Since the constraints in the current problem are $\lambda_i(M) \geq 0$, $i = 1, \ldots, n$. A centering transform would be one that made all the eigenvalues of M equal. Hence the centering transform would be one that mapped the current matrix M to identity I.

One can construct a Riemannian metric $g_{ij}(x)\,\mathrm{d}x^i\,\mathrm{d}x^j$ on the original space which is invariant with respect to the family of transformations G. Once we define the metric at identity the transformation $L_{M^{1/2}}$ would give us a unique way of extending it to all positive definite matrices M, since the derivative of the transformation gives a linear map between the tangent space at I and the tangent space at M. To define the distance between M and $M + \Delta M \in P$, we proceed as follows.

We choose the transform to be L_T where $T = M^{-1/2}$ which maps M to identity. We then compute the image of $M + \Delta M$ under this transform

$$L_{M^{-1/2}}(M) = I$$
$$L_{M^{-1/2}}(M + \Delta M) = I + \Delta M'$$
$$\text{where } \Delta M' = M^{-1/2}\,\Delta M M^{-1/2}\,.$$

We choose to use the uniform spherical metric to define the distances at I. Hence the distance can be written as

$$\begin{aligned} d^2(I, I + \Delta M') &= \sum_i \sum_j \Delta m'^2_{ij} \\ &= tr(\Delta M'\,\Delta M'^T) \\ &= tr(M^{-1/2}\,\Delta M M^{-1}\,\Delta M M^{-1/2}) \\ &= tr(\Delta M M^{-1}\,\Delta M M^{-1})\,. \end{aligned}$$

By requiring that that the distances be invariant with respect to the transformations in G we need that

$$\begin{aligned} d^2(M, M + \Delta M) = d^2(L_T(M)\,,\quad L_T(M + \Delta M)) &= d^2(I, I + \Delta M') \\ = tr(\Delta M M^{-1}\,\Delta M M^{-1})\,. \end{aligned}$$

Since only the diagonals of M are variable, ΔM is a diagonal and can be written simply as

$$\Delta M = \sum_i \Delta z_i \mathbf{e}_i \mathbf{e}_i^T\,.$$

Also, the Riemannian metric will be denoted as $g_{ij}(M)\,\mathrm{d}z^i\,\mathrm{d}z^j$.

Hence the distance in the Riemannian metric would be given as follows

$$\begin{aligned}
g_{ij}(M)\Delta z^i \Delta z^j &= tr(M^{-1}\Delta Z M^{-1}\Delta Z) \\
&= tr(M^{-1}\sum_i \Delta z_i \mathbf{e}_i\mathbf{e}_i^T M^{-1}\sum_j \Delta z_j \mathbf{e}_j\mathbf{e}_j^T) \\
&= \sum_i \sum_j tr(M^{-1}\mathbf{e}_i\mathbf{e}_i^T M^{-1}\mathbf{e}_j\mathbf{e}_j^T)\Delta z_i \Delta z_j \\
&= \sum_i \sum_j (\mathbf{e}_i^T M^{-1}\mathbf{e}_j)^2 \Delta z_i \Delta z_j .
\end{aligned}$$

From this we conclude that

$$g_{ij}(M) = (\mathbf{e}_i^T M^{-1}\mathbf{e}_j)^2 .$$

In these formulae $g_{ij}(M)$ is a covariant symmetric tensor of degree 2. The corresponding contravariant tensor is denoted by $g^{ij}(M)$ and satisfies

$$g^{ij}g_{jk} = \delta^i_k . \tag{9}$$

If $f: P \to R$ is a C^1-function defined on P, then its direction of descent $\mathbf{d}$ with respect to the Riemannian metric is the contravariant vector obtained from the gradient of f, which is a co-variant vector, by raising indices

$$d^i = \sum_j g^{ij} \frac{\delta f}{\delta z_j} .$$

In the present problem, the function to be optimized is given to be $f(\mathbf{z}) = \mathbf{e}^T\mathbf{z}$ and hence the descent direction can be obtained easily. If we were to follow the descent directions infinitesimally, we would get a continuous trajectory which is given by the differential equation

$$\frac{\mathrm{d}z_i}{\mathrm{d}t} = \sum_j g^{ij} \frac{\delta f}{\delta z_j} . \tag{10}$$

Later, we shall given an alternate interpretation for this continuous trajectory.

Corresponding to the constraints in (7), we introduce a potential function [5], [8], [9] which is invariant up to a constant with respect to the transformations in G. It is given as

$$\begin{aligned}
\phi(M) &= -\ln \Pi \lambda_i(M) \\
&= -\ln \det(M) .
\end{aligned} \tag{11}$$

The potential function is defined only for points that are in the interior of the constrained region, i.e., $\lambda_i(M) > 0$, $i = 1, \ldots, n$. We note that unlike the eigenvalues, the determinant of a matrix is a smooth function which can be expressed as a polynomial in the components of the matrix [13], [14]. The potential differences can be shown to be invariant with respect to the transformations in G as follows.

Let L_C be a transform in G and let $A, B \in P$. Let $A' = L_C(A) = CAC^T$ be the image of A under the transform L_C and similarly let $B' = CBC^T$. Then the difference in the values of the potential function at A' and B' can be given by

$$\begin{aligned}
&\ln \det(A') - \ln \det(B') \\
&= \ln\left(\frac{(\det(CAC)}{\det(CBC)} \right) \\
&= \ln\left(\frac{(\det C)^2 \det A}{(\det C)^2 \det B)} \right) \\
&= \ln \det A - \ln \det B .
\end{aligned}$$

This demonstrates the invariance of potential differences under the transformations in G.

We shall now compute the partial derivatives of the potential function with respect to the variables in $\mathbf{z}$. In this context, we prove the following lemmas.

LEMMA 2. *Let* $\mathbf{x} = (x_1, \ldots, x_m) \in R^m$ *and* $M(\mathbf{x})$ *be a real symmetric positive-definite matrix of size* $n \times n$ *whose elements depend on* $\mathbf{x}$. *Let* $\phi(\mathbf{x}) = -\ln \det M(\mathbf{x})$ *and suppose for a pair of indices i and j*

$$\frac{\delta M}{\delta x_i} = B_i \text{ and } \frac{\delta M}{\delta x_j} = B_j ,$$

where B_i, $B_j \in R^{n \times n}$ *are symmetric matrices independent of* $\mathbf{x}$. *Also, let* $tr(M)$ *denote the trace of the matrix M.*

Then the first and second order derivatives of $\phi(\mathbf{x})$ *are given by*

$$\frac{\delta \phi}{\delta x_i} = -tr(M^{-1} B_i) , \quad \frac{\delta \phi}{\delta x_j} = -tr(M^{-1} B_j)$$

and

$$\frac{\delta^2 \phi}{\delta x_i \, \delta x_j} = tr(M^{-1} B_i M^{-1} B_j) .$$

Proof. For very small ϵ it can be shown that

$$\det(I + \epsilon B) = 1 + \epsilon tr(B) + 0(\epsilon^2) .$$

We shall use this result later in the proof.

In order to differentiate with respect to x_i we keep all variables except x_i constant. For analysis, let us use the notation $M_i(x_i)$ for $M(\mathbf{x})$ when all variables except x_i are unchanged.

$$\begin{aligned}
\frac{\delta \phi}{\delta x_i} &= \frac{-1}{\det(M(\mathbf{x}))} \frac{\delta}{\delta x_i} \det(M(\mathbf{x})) \\
&= \frac{-1}{\det(M(\mathbf{x}))} \lim_{\Delta x_i \to 0} \frac{\det(M_i(x_i + \Delta x_i)) - \det(M_i(x_i))}{\Delta x_i}
\end{aligned}$$

$$= \frac{-1}{\det(M(\mathbf{x}))} \lim_{\Delta x_i \to 0} \frac{\det(M_i(x_i + \Delta x_i B_i)) - \det(M_i(x_i))}{\Delta x_i}$$

$$= \frac{-1}{\det(M(\mathbf{x}))} \lim_{\Delta x_i \to 0} \frac{\det(M_i(x_i))(1 + \Delta x_i tr(M^{-1}B_i)) - \det(M_i(x_i))}{\Delta x_i}$$

$$= -tr(M^{-1}B_i) .$$

Similarly, we can show that

$$\frac{\delta\phi}{\delta x_j} = -tr(M^{-1}B_j) .$$

On differentiating $MM^{-1} = I$ with respect to x_j we get

$$\frac{\delta M^{-1}}{\delta x_j} = -M^{-1} \frac{\delta M}{\delta x_j} M^{-1} = -M^{-1}B_j M^{-1} .$$

The result for the second order derivatives now follows easily as

$$\frac{\delta^2\phi}{\delta x_i \, \delta x_j} = -tr\left(\frac{\delta M^{-1}}{\delta x_j} B_i\right)$$

$$= tr(M^{-1}B_j M^{-1} B_i) . \qquad \blacksquare$$

LEMMA 3. *Let* $\mathbf{h}$ *be the gradient and* H *be the Hessian of the potential function* $\phi(\mathbf{z})$ *in* (11). *Let the inverse of* $M(\mathbf{z}) = Z - Q$ *be denoted by* $M^{-1} = [r_{ij}]$ *then*

$$h_i = -r_{ii} \text{ and } H_{ij} = r_{ij}^2 , \quad i, j = 1, \ldots, n .$$

Proof. Let $\mathbf{e}_i$ be an n-dimensional vector whose ith coordinate is 1 and all other coordinates are 0. Then we get the results,

$$\frac{\delta M(\mathbf{z})}{\delta z_i} = \mathbf{e}_i \mathbf{e}_i^T , \quad i = 1, \ldots, n$$

hence

$$h_i = \frac{\delta\phi(\mathbf{z})}{\delta z_i} = -tr(M^{-1}\mathbf{e}_i\mathbf{e}_i^T) = -r_{ii}$$

and

$$H_{ij} = \frac{\delta^2\phi(\mathbf{z})}{\delta z_i \, \delta z_j} = tr(M^{-1}\mathbf{e}_j\mathbf{e}_j^T M^{-1}\mathbf{e}_i\mathbf{e}_i^T) = r_{ij}^2 . \qquad \blacksquare$$

It is clear from the above lemmas that g_{ij} is in fact equal to H_{ij}.

It is obvious that minimizing the potential function corresponds to centering with respect to the constraints. We shall now show that the curve defined in (10) and starting from the center has the property that for each value of the objective function, the point on the curve minimizes the potential function for that objective value. To prove this, let us consider the first order optimality condition at the point that minimizes the potential function while keeping a linear objective function $\mathbf{c}^T\mathbf{z}$ constant (where $\mathbf{c} \in R^n$). It is given by

$$\frac{\delta\phi}{\partial z_i} = s\,\frac{\delta(\mathbf{c}^T\mathbf{z})}{\delta z_i} = sc_i$$

where s is a scalar parameter.

On differentiating the above differential equation with respect to s and using the chain rule, we get

$$\sum_j \frac{\delta^2\phi}{\delta z_i\,\delta z_j}\,\frac{\mathrm{d}z_j}{\mathrm{d}s} = c_i\,.$$

But since we have already shown that $(\delta^2\phi)/(\delta z_i\,\delta z_j) = g_{ij}$, the differential equation becomes

$$\sum_j g_{ij}\,\frac{\mathrm{d}z_j}{\mathrm{d}s} = c_i$$

and hence using (9) we get

$$\frac{\mathrm{d}z_j}{\mathrm{d}s} = \sum_i g^{ij}c_i\,.$$

This is the same differential equation as given in (10) with $f = \mathbf{c}^T\mathbf{z}$. This proves that the points obtained from minimizing the potential function for different values of the objective, form a curve which is identical to the continuous trajectory defined in (10). In our present problem the objective is $\mathbf{e}^T\mathbf{z}$ and the differential equation corresponding to the continuous trajectory can be obtained by substituting $\mathbf{c}$ by $\mathbf{e}$.

This suggests a method for solving problem (8) by alternating between two steps – a potential step in which we minimize the potential function $\phi(\mathbf{z})$ and an objective step in which we reduce the objective $\mathbf{e}^T\mathbf{z}$.

We now propose our iterative approach for solving problem (8). We have to start at a point in the interior of the region of optimization which means that the initial values $\mathbf{z}^{(0)}$ of the vector $\mathbf{z}$ should be such that the matrix $M^{(0)} = Z^{(0)} - Q$ is positive-definite. One such value for $\mathbf{z}^{(0)}$ can be obtained as follows. We compute $\lambda_{\max}(Q)$ and take $\mu > \lambda_{\max}(Q)$. Set $\mathbf{z}^{(0)} = \mu\mathbf{e}$. The matrix $M^{(0)}$ can be shown to be positive definite. Let $\mathbf{z}^{(k)}$ be the state of the z-vector at the kth iteration. We define $M^{(k)} = Z^{(k)} - Q$.

As mentioned earlier, we shall alternate between the potential and the objective steps. In the potential step we try to minimize the potential function while keeping the value of the objective constant at its current value. The problem to be solved in the potential step can be defined as follows,

$$\begin{aligned} \min \quad & \phi(\mathbf{z}) \\ \text{s.t.} \quad & \mathbf{e}^T\mathbf{z} = \text{const.} \\ & \lambda_i(M) \geq 0\,, \quad i = 1, \ldots, n\,. \end{aligned}$$

This is a constrained optimization problem and we shall solve it approximately as follows.

Since $\mathbf{z}^{(k)}$ is an interior point, the potential function $\phi(\mathbf{z})$ is defined for $\mathbf{z}^{(k)}$ and we can make a quadratic Taylor series approximation of $\phi^{(k)}(\mathbf{z})$ around $\mathbf{z}^{(k)}$ which is given by

$$T^{(k)}(\Delta\mathbf{z}_p) = \phi(\mathbf{z}^{(k)}) + \mathbf{h}^T\,\Delta\mathbf{z}_p + \frac{1}{2}\,\Delta\mathbf{z}_p^T H\,\Delta\mathbf{z}_p \text{ where } \Delta\mathbf{z}_p = \mathbf{z} - \mathbf{z}^{(k)}$$

and $\mathbf{h}$ is the gradient and H is the Hessian of $\phi(\mathbf{z})$ at $\mathbf{z}^{(k)}$.

The Hessian H and gradient $\mathbf{h}$ can be computed using Lemma 3.

To obtain the descent direction $\Delta\mathbf{z}_p$ that minimizes $\phi(\mathbf{z})$, while maintaining the objective value constant, we solve the problem

$$\begin{aligned} \min \quad & T^{(k)}(\Delta\mathbf{z}_p) \\ \text{s.t.} \quad & \mathbf{e}^T\,\Delta\mathbf{z}_p = 0\,. \end{aligned}$$

On applying the first-order Kuhn–Tucker optimality conditions, we get the linear system

$$\begin{aligned} H\,\Delta\mathbf{z}_p + \mathbf{h}^T &= \alpha\mathbf{e}^T \\ \mathbf{e}^T\,\Delta\mathbf{z}_p &= 0\,. \end{aligned}$$

We solve this linear system to obtain $\Delta\mathbf{z}_p$ which gives a new set of diagonals

$$\mathbf{z}_p^{(k)} = \mathbf{z}^{(k)} + \Delta\mathbf{z}_p\,.$$

After the potential step, we need to take a step that will improve the objective $\mathbf{e}^T\mathbf{z}$. As we have shown earlier, if $H_0 = [g_{ij}(Z_p^{(k)} - Q)]$ then the steepest direction of descent with respect to the Riemannian metric is given by

$$\Delta\mathbf{z}_0 = \alpha H_0^{-1}\mathbf{e}\,.$$

This gives us the direction for the objective step. Before determining the step length, we must however ensure that after taking the step the resulting matrix M continues to remain strictly positive definite. We can start with a value of α so that the length of $\Delta\mathbf{z}_0$ in the Riemannian metric is given by $\|\,\Delta\mathbf{z}_0\,\|_R = n^{1/2}$ and then halve it successively if necessary until the corresponding M becomes positive definite. The method for adjusting alpha is outlined below.

procedure adjust(α,$\mathbf{d}$,$\mathbf{z}$,Q,H)
{α-scaling factor,
$\mathbf{d}$-descent direction,
$\mathbf{z}$-state of z-vector,
H-Hessian of $\phi(\mathbf{z})$ at $\mathbf{z}$.}
Begin

1. Set $\alpha = \left(\frac{n}{\mathbf{d}^T H \mathbf{d}}\right)^{1/2}$.
2. While ($(Z - Q + \alpha D)$ is not positive-definite) do

$$\alpha = \alpha/2$$

   endwhile
3. return $(\alpha/2)$

End

Our algorithm for minimizing the bound $f_{\max}$ on (1) can hence be described as follows:

Algorithm
Begin

1. Compute maximum eigenvalue of Q and take $\mu > \lambda_{\max}(Q)$
2. Initialise $\mathbf{z}^{(0)} = \mu\mathbf{e}$ and $K = 0$.
3. Define the potential function

$$\phi(\mathbf{z}) = \ln\det(Z - Q).$$

4. Repeat
   begin loop
5. Construct

$$M^{(k)} = Z^{(k)} - Q$$

6. Compute Hessian H and gradient $\mathbf{h}$ for $\phi(\mathbf{z})$ at $\mathbf{z}^{(K)}$
7. Solve

$$\begin{bmatrix} \mathrm{H} & -\mathbf{e}^T \\ -\mathbf{e}^T & 0 \end{bmatrix}\begin{bmatrix} \Delta\mathbf{z}_p \\ \alpha \end{bmatrix} = \begin{bmatrix} -\mathbf{h}^T \\ 0 \end{bmatrix}$$

8. $\mathbf{z}_p^{(K)} = \mathbf{z}^{(K)} + \Delta\mathbf{z}_p$
9. Compute new Hessian H' of $\phi(\mathbf{z})$ at $\mathbf{z}_p^{(K)}$
10. Solve

$$H'\,\Delta\mathbf{z}_0 = \mathbf{e}$$

11. call adjst(α,$\Delta\mathbf{z}_0$,$\mathbf{z}_p^{(K)}$,Q,H')
12. $\mathbf{z}^{(K+1)} = \mathbf{z}_p^{(k)} + \alpha\,\Delta\mathbf{z}_0$
13. $K = K + 1$
    end loop
14. Until $\|\mathbf{z}^{(K)} - \mathbf{z}^{(K-1)}\| < \delta$
15. Return $tr(M^{(k)})$.

End

3. Conclusion and Directions for Future Work

In this paper, we have addressed the problem of finding an upper bound in a quadratic maximization problem with integer constraints. Combinatorial techniques usually have a hard time on such problems. We had proposed a continuous method for solving the problem [4]. We have extended our earlier approach [4] by using techniques of differential geometry. Further computational experiments on this method are underway. Extensions and improvements to this method as also the complexity analysis of the technique is a subject of our current research. We believe that this method would be useful in several related applications.

References

1. Gantmakher, F. (1959), *Theory of Matrices*, Chelsea.
2. Garey, M. and Johnson, D. (1979), *Computers and Intractability: A Guide to the Theory of NP-Completeness*, W. H. Freeman, San Francisco.
3. Han, C., Pardalos, P. and Ye, Y. (1989), Computational Aspects of an Interior Point Algorithm for Quadratic Problems with Box Constraints, *Proc. of Conf. on Large-Scale Numerical Optimization*, Cornell Univ.
4. Kamath, A. and Karmarkar, N. (1991), A Continuous Approach to Compute Upper Bounds in Quadratic Maximization Problems with Integer Constraints, *Proc. of Conf. on Recent Advances in Global Optimization*, eds. Floudas and Pardalos, pp. 125–140.
5. Kamath, A., Karmarkar, N., Ramakrishnan, K. and Resende, M. (1990), Computational Experience with an Interior Point Algorithm on the Satisfiability Problem, *Annals of Opns. Res.* **25**, 43–58.
6. Karmarkar, N. (1984), A New Polynomial Time Algorithm for Linear Programming, *Combinatorica* **4**, 373–395.
7. Karmarkar, N. (1990), Riemannian Geometry Underlying Interior-Point Methods for Linear Programming, *Contemporary Mathematics* **114**, 51–75.
8. Karmarkar, N. (1991), An Interior-Point Approach to NP-Complete Problems, Math. Develop. arising from Linear Prog., Proc. of a Joint Summer Research Conf., Bowdoin College, 1988, *American Math. Soc. – Contemporary Math.* **114**, 297–308.
9. Karmarkar, N., Resende, M. and Ramakrishnan, K. (1991), An Interior Point Algorithm to Solve Computationally Difficult Set Covering Problems, *Mathematical Programming*, Special Issue on "Interior Point Methods: Theory and Practice", eds. Roos and Vial. **52**, 597–618.
10. Luenberger, D. (1973), *Linear and Nonlinear Programming*, Addison Wesley.
11. Monteiro, R. and Adler, I. (1989), Interior Path Following Primal-Dual Algorithms, Part II: Convex Quadratic Programming, *Mathematical Programming* **44**, 43–66.
12. Pardalos, P. and Vavasis, S. (1991), Quadratic Programming with One Negative Eigenvalue Is NP-hard, *Journal of Global Optimization* **1**, 15–22.
13. Rellich, F. (1969), *Perturbation Theory of Eigenvalue Problems*, Gordon and Breach.
14. Wilkinson, J. (1965), *The Algebraic Eigenvalue Problem*, Clarendon Press, Oxford.