Mathematical Programming 53 (1992) 63-78
North-Holland

Applications of parametric programming and eigenvalue maximization to the quadratic assignment problem

Franz Rendl*

Technische Universität Graz, Institut für Mathematik, Kopernikusgasse 24, A-8010 Graz, Austria

Henry Wolkowicz**

Department of Combinatorics and Optimization, University of Waterloo, Waterloo, Ontario, Canada N2L 3G1

Received 28 March 1988
Revised manuscript received 4 August 1989

We investigate new bounding strategies based on different relaxations of the quadratic assignment problem. In particular, we improve the lower bound found by using an eigenvalue decomposition of the quadratic part and by solving a linear program for the linear part. The improvement is accomplished by applying a steepest ascent algorithm to the sum of the two bounds.

Key words: Quadratic assignment problem, relaxation, lower bounds, eigenvalue decomposition, steepest ascent.

1. Introduction

The *quadratic assignment problem* (QAP) can be described as follows: given the set $N=\{1, 2, \ldots, n\}$ and three $n \times n$ matrices $A=(a_{ik})$, $B=(b_{jl})$, and $C=(c_{ij})$, find a permutation π of the set N which minimizes

$$\sum_{i=1}^{n} c_{i\pi(i)} + \sum_{i=1}^{n} \sum_{k=1}^{n} a_{ik} b_{\pi(i)\pi(k)}. \tag{1.1}$$

Equivalently, find an $n \times n$ permutation matrix X which minimizes the trace

$$\min_{X \in \Pi} \operatorname{tr}(C + AXB^{t})X^{t}, \tag{1.2}$$

where t denotes transpose, Π denotes the set of permutation matrices, and tr is the trace (see, e.g., [4]).

Both authors would like to thank the Natural Sciences and Engineering Research Council Canada and the Austrian Government for their support.

* This author would like to acknowledge partial support from the Department of Combinatorics and Optimization at the University of Waterloo.

** Research partially supported by Natural Sciences and Engineering Research Council Canada.

This model arises in, e.g., location problems, where N describes the set of sites on which plants are to be built; a_{ik} is the distance between sites i and k; b_{jl} describes the flow between plants j and l; while c_{ij} is the building or running cost for plant j in location i (see, e.g., [4] for further references and applications for this model).

The QAP is an NP-hard problem. (It contains, as a special case, the Travelling Salesman Problem.) Moreover, even the problem of finding an ε-approximation of the optimal solution is NP-hard (see [18]). Thus from a worst case point of view, QAPs are extremely difficult to solve. Recently, it has been shown [3] that solving the average case takes exponential time, when the QAPs are taken from some simple sample space of random problems. In [17], a parallel branch and bound technique failed to solve problems of dimension larger than 15. The main reason for this seems to be the lack of strong lower bounds that can be computed efficiently.

The present theory and solution techniques for QAP are surveyed in [4]. Solution techniques often require lower bounds. In [4] an "optimal" reduction scheme is presented which reduces the magnitude of the quadratic part, and thus augments the influence of the linear part. Then a lower bound for the quadratic part is found using eigenvalue decompositions, while a lower bound for the linear part is found by solving the corresponding linear assignment problem as a linear program.

In this paper we consider approximate solution techniques which result in lower bounds for QAP. In Section 3 we continue the approach in [4], and find a lower bound by using an eigenvalue decomposition for the quadratic part and solving a linear program for the linear part. Theorem 3.1 finds the exact minimum (and maximum) for the quadratic part when the relevant constraints are relaxed to include all orthogonal matrices, rather than just the permutation matrices. This yields the lower (and upper) bounds obtained in [4]. We then apply a steepest ascent algorithm to increase the sum of the two bounds. This requires some differential calculus for eigenvalue perturbations and subdifferential calculus for a quadratically perturbed linear program. We present some numerical experiments in Section 4.

2. Preliminaries

We can reformulate the QAP in (1.1) using a permutation matrix $X=(x_{ij})$ to get the form (1.2), i.e., X satisfies the transportation and 0, 1 constraints

$$\begin{aligned} &\sum_{j=1}^{n} x_{ij}=1, \quad i=1,\ldots,n,\\ &\sum_{i=1}^{n} x_{ij}=1, \quad j=1,\ldots,n,\\ &x_{ij}\in\{0,1\}, \quad i,j=1,\ldots,n. \end{aligned} \tag{2.1}$$

Then (1.1) becomes (see [4])

$$\operatorname{tr}(C+AXB^{t})X^{t}=\sum_{i}\sum_{j} c_{ij}x_{ij}+\sum_{i}\sum_{j}\sum_{k}\sum_{l} a_{ik}b_{jl}x_{ij}x_{kl}. \tag{2.2}$$

We let O denote the set of *orthogonal matrices*, i.e., $X \in O$ if $X^t X = I$, the identity. Further, we let Ω denote the set of doubly stochastic matrices, i.e., nonnegative matrices with row and column sums 1. It is well known that

$$\Pi = O \cap \Omega. \tag{2.3}$$

We assume that A and B are symmetric with eigenvalues $\lambda = (\lambda_i)$ and $\mu = (\mu_i)$, respectively. The matrices A and B can be orthogonally diagonalized, i.e., $A = P_1 \Lambda_1 P_1^t$ and $B = P_2 \Lambda_2 P_2^t$ with orthogonal matrices P_1, P_2 and diagonal matrices $\Lambda_1 = \text{diag}(\lambda)$, $\Lambda_2 = \text{diag}(\mu)$.

3. Eigenvalue approach

3.1. Orthogonal relaxation

We now consider bounds for the QAP which are found by a particular relaxation of the constraint that $X \in \Pi$, a permutation matrix. Since $\Pi = O \cap \Omega$, i.e., since it is the intersection of the orthogonal and the doubly stochastic matrices, we can relax the QAP by considering only the orthogonal matrices. This allows us to derive bounds for the pure quadratic assignment problem. We then combine this with a reduction mechanism to bound QAP.

Recall that A and B are symmetric matrices with eigenvalues λ and μ, respectively, and with orthogonal diagonalizations $A = P_1 \Lambda_1 P_1^t$ and $B = P_2 \Lambda_2 P_2^t$. Permuting the columns of P_i is equivalent to permuting the order of appearance of the eigenvalues. We let $\bar{\lambda}$ denote the ordered vector of eigenvalues with $\bar{\lambda}_1 \geq \cdots \geq \bar{\lambda}_n$, while $\underline{\lambda}$ denotes the reverse order so that $\underline{\lambda}_1 \leq \cdots \leq \underline{\lambda}_n$. We similarly define $\bar{\mu}$ and $\underline{\mu}$. We then get the *maximal scalar product*

$$\langle \lambda, \mu \rangle_+ := \langle \bar{\lambda}, \bar{\mu} \rangle \tag{3.1}$$

and the *minimal scalar product*

$$\langle \lambda, \mu \rangle_- := \langle \bar{\lambda}, \underline{\mu} \rangle. \tag{3.2}$$

The following theorem finds the optimal solutions for the relaxed pure quadratic assignment problem. The corollary following the theorem yields the maximal and minimal scalar product bounds obtained in [4]. This theorem shows that the bounds in [4] are actually attained. The proof of the theorem uses Lagrange multipliers to find the optimal solutions where the bounds are attained. The theorem can also be proved using double stochastic matrices as in [4], or using the Hoffman-Wielandt Theorem [10]. (See, e.g., [5].) We include the Lagrange multiplier proof for completeness and interest. In particular, the commutativity condition (for $X^t AX$ and B) derived in the proof seems to appear in these types of best approximation problems, and might be of independent interest. Moreover, this proof has led to an extension for the case when A and B are general matrices, not necessarily symmetric. (This extension is to appear in a forthcoming study.)

Theorem 3.1. *Suppose that the orders of the columns of P_i, $i = 1, 2$, are chosen so that $\langle \lambda, \mu \rangle = \langle \lambda, \mu \rangle_-$. Then*

$$\min_{X \in O} \operatorname{tr} AXBX^t = \operatorname{tr} AP_1P_2^tBP_2P_1^t = \langle \lambda, \mu \rangle_-. \tag{3.3}$$

On the other hand, if we choose P_i, $i = 1, 2$, so that $\langle \lambda, \mu \rangle = \langle \lambda, \mu \rangle_+$, then

$$\max_{X \in O} \operatorname{tr} AXBX^t = \operatorname{tr} AP_1P_2^tBP_2P_1^t = \langle \lambda, \mu \rangle_+. \tag{3.4}$$

Proof. Let

$$k(X) := XBX^t, \qquad g(X) := X^tX - I, \qquad f(X) := \operatorname{tr} AXBX^t. \tag{3.5}$$

Then the corresponding differentials are

$$\begin{aligned} &\mathrm{d}k(X; h) = XBh^t + hBX^t; \\ &\mathrm{d}g(X; h) = X^th + h^tX; \\ &\mathrm{d}f(X; h) = \operatorname{tr} A(\mathrm{d}k(X; h)) = \operatorname{tr} A(XBh^t + hBX^t); \end{aligned} \tag{3.6}$$

and the minimization problem can be stated as

$$\min\{f(X): g(X) = 0\}. \tag{3.7}$$

Since the feasible set is compact, the optimum exists. We can now apply the theory of Lagrange multipliers. Note that $g(X)$ is a symmetric matrix so that we can assume that the Lagrange multiplier, call it S, is a symmetric matrix. Also, for given $Q = Q^t$ let $h = \frac{1}{2}XQ$. Then $\mathrm{d}g(X; h) = Q$, which implies that the derivative, g', is onto or full rank, and so a constraint qualification holds at the optimum. Thus, the Lagrange multiplier S exists.

The Lagrangian is

$$f(X) + \operatorname{tr} Sg(X).$$

Therefore,

$$0 = \mathrm{d}f(X; h) + \operatorname{tr} S\, \mathrm{d}g(X; h) \tag{3.8}$$

for all matrices h. This yields, for all h,

$$\begin{aligned} 0 &= \operatorname{tr}[A(XBh^t + hBX^t) + S(X^th + h^tX)] \\ &= \operatorname{tr}(BX^tA + SX^t)h + \operatorname{tr} h^t(AXB + XS). \end{aligned}$$

Thus

$$AXB + XS = 0 \tag{3.9}$$

or

$$X^tAXB = -S.$$

Since S is symmetric, this implies that X^tAX and B commute, and so are mutually orthogonally diagonalizable, by P_2 say. Therefore

$$\operatorname{tr} AXBX^t = \operatorname{tr} P_2X^tAXP_2^tP_2BP_2^t = \operatorname{tr} \Lambda_1\Lambda_2. \tag{3.10}$$

The minimum value is then $\langle\lambda, \mu\rangle_-$ and is attained with $XP_2^t = P_1^t$, i.e., $X = P_1^tP_2$. The corresponding proof for the maximum is similar. □

The above theorem yields the following bounds obtained in [4, Theorem 3]. Note that one can use these bounds, in conjunction with a solution of the linear part, to obtain bounds for QAP.

Corollary 3.1. *For all* $X \in \Pi$,

$$\langle\lambda, \mu\rangle_- \leq \operatorname{tr} AXBX^t \leq \langle\lambda, \mu\rangle_+. \qquad \square \tag{3.11}$$

3.2. Reduction scheme

In the above we have obtained upper and lower bounds on the quadratic term in QAP. In [4], the following two reduction schemes are presented in order to reduce the magnitude of the quadratic term and augment the influence of the linear term, which can be solved exactly in polynomial time. In the following, let $E = e1^t + 1e^t$, and $G = g1^t + 1g^t$, where 1 denotes the n-vector of ones and e and g are n-vectors.

(Rd1) Set $A = \bar{A} + E$ and $B = \bar{B} + G$. Then for every $X \in \Pi$, $\operatorname{tr}(AXB + C)X^t = \operatorname{tr}(\bar{A}X\bar{B} + \bar{C})X^t$ with $\bar{C} = 2\bar{A}(1g^t) + 2(e1^t)B + C$.

(Rd2) Set $A = \bar{A} + R$ and $B = \bar{B} + S$ with diagonal matrices $R = \operatorname{diag}(r_1, \ldots, r_n)$ and $S = \operatorname{diag}(s_1, \ldots, s_n)$. Then for every $X \in \Pi$, $\operatorname{tr}(AXB + C)X^t = \operatorname{tr}(\bar{A}X\bar{B} + \bar{C})X^t$ with $\bar{C} = (\bar{c}_{ij}) = (c_{ij} + \bar{a}_{ii}s_j + r_i\bar{b}_{jj} + r_is_j)$.

Thus we may reduce a symmetric matrix A to the form $A = \bar{A} + E + R$. We could then apply Theorem 3.1 to get new lower and upper bounds $\langle\lambda, \mu\rangle_-$ and $\langle\lambda, \mu\rangle_+$. Improvements would result from minimizing the distance between the two bounds.

The approach in [4] attempts to minimize the fluctuation of the eigenvalues of both matrices A and B. This is done by attempting to reduce the *spread*

$$\operatorname{sp}(A) := \max_{i,j} |\lambda_i - \lambda_j|.$$

It is mentioned in [4] that since there are no simple formulas for the spread, they instead minimize the upper bound for $\operatorname{sp}(A)$ given by Mirsky [12],

$$\operatorname{sp}(A) \leq m(A) := \left[2\sum_i\sum_j a_{ij}^2 - \frac{2}{n}(\operatorname{tr} A)^2\right]^{1/2}. \tag{3.12}$$

Note that (see [12, 20]),

$$m(A) = (2n)^{1/2}s_\lambda,$$

where the variance of λ,

$$s_\lambda^2 = \frac{\operatorname{tr} A^2}{n} - \left(\frac{\operatorname{tr} A}{n}\right)^2.$$

Equality holds in (3.12) if and only if

$$\lambda_2 = \lambda_3 = \cdots = \lambda_{n-1} = \tfrac{1}{2}(\lambda_1 + \lambda_n). \tag{3.13}$$

Thus in [4] the reduced matrices $\bar{A} = A - E - R$ and $\bar{B} = B - G - S$ are found which minimize the variances s_λ^2 and s_μ^2 over all vectors e, g, r, s. A calculus argument finds the reduced matrices explicitly.

The computational scheme for the reduction of A is

$$\begin{aligned} z &:= \frac{1}{2(n-1)}\left(\left(\sum_i \sum_j a_{ij}\right) - \operatorname{tr} A\right), \\ e_k &:= \frac{1}{n-2}\left(\left(\sum_j a_{kj}\right) - a_{kk} - z\right), \quad k = 1, \ldots, n, \\ r_k &:= a_{kk} - 2e_k, \quad k = 1, \ldots, n. \end{aligned} \tag{3.14}$$

B is reduced similarly. The reduced matrices $\bar{A}$ and $\bar{B}$ then have row and column sums equal to zero as well as zero diagonals.

Suppose that the vectors λ and μ are arranged in non-increasing order. The range of values, call it v, satisfies

$$\begin{aligned} v &= \langle \lambda, \mu \rangle_+ - \langle \lambda, \mu \rangle_- = \sum_k \lambda_k(\mu_k - \mu_{n-k+1}) \\ &= \sum_k \lambda_k(\mu_k - \mu_{n-k+1}) - \left(\sum_k \lambda_k\right)\left(\sum_k (\mu_k - \mu_{n-k+1})\right) \Big/ n \\ &= n \operatorname{Cov}(\lambda, (\bar{\mu} - \underline{\mu})), \end{aligned} \tag{3.15}$$

where Cov denotes the covariance. Thus we would like to minimize this covariance. Note that

$$v \leq n s_\lambda^2 s_{(\bar{\mu}-\underline{\mu})}^2.$$

Similarly,

$$v \leq n s_\mu^2 s_{(\bar{\lambda}-\underline{\lambda})}^2.$$

Moreover,

$$s_{(\bar{\mu}-\underline{\mu})}^2 = \sum_k (\mu_k - \mu_{n-k+1})^2/n \leq (\mu_1 - \mu_n)^2 \leq 2n s_\mu^2$$

by (3.12). We conclude that the range of values

$$v \leq 2n^2 \operatorname{Var}(\lambda) \operatorname{Var}(\mu) \tag{3.16}$$

with equality conditions determined by the equality in the Cauchy-Schwarz inequality applied to Cov, and equality in Mirsky's bound for the spread. Thus equality holds if $\bar{\mu} - \underline{\mu}$ is a multiple of $\bar{\lambda}$, and both $\bar{\mu}$ and $\bar{\lambda}$ satisfy (3.13). Minimizing the upper bound for the spread is equivalent to minimizing the upper bound to v. In [4], it is shown that $v = 0$ for 3×3 matrices. This is not true in general.

Though the intention in [4] was to reduce the spread, the compromise was to reduce the upper bound given in (3.12). This is equivalent to reducing the standard deviation, or variance, of the eigenvalues. In fact, reducing the standard deviation seems to be the correct choice as seen by the upper bound of v. This is borne out by numerical tests we have done, i.e., we used a minimization routine to reduce the spread (see Overton [15]), and noted that the reduction (3.14) consistently gave better bounds.

Example 3.1 (Test example from Nugent et al. [14]).

$$A = \begin{bmatrix} 0 & 1 & 1 & 2 & 3 \\ 1 & 0 & 2 & 1 & 2 \\ 1 & 2 & 0 & 1 & 2 \\ 2 & 1 & 1 & 0 & 1 \\ 3 & 2 & 2 & 1 & 0 \end{bmatrix}, \qquad B = \begin{bmatrix} 0 & 5 & 2 & 4 & 1 \\ 5 & 0 & 3 & 0 & 2 \\ 2 & 3 & 0 & 0 & 0 \\ 4 & 0 & 0 & 0 & 5 \\ 1 & 2 & 0 & 5 & 0 \end{bmatrix}.$$

The optimal reduction scheme (3.14) yielded spreads of 2.9814 and 10.0812, while the minimization procedure from [15] gave 2.4722 and 9.9330 for the matrices A and B, respectively. However, this resulted in lower and upper bounds -16.4915 and 16.4915 for the optimal reduction scheme, while the minimization gave the weaker bounds -23.1276 and 21.7713. Note that the standard deviation of the eigenvalues for A and B from the optimal reduction were 1.0328, 3.2660 while the minimization yielded 1.2111, 4.1086, respectively. This coincides with the fact that a lower standard deviation provides better bounds.

3.3. *Improving the reduction*

The reduced matrices $\bar{A}$ and $\bar{B}$ above are found independently. Moreover, the reduction is found using upper bounds for the spread, and $\bar{C}$ is not taken into consideration. This raises the question of whether one could improve the reduction. We will try and improve the lower bound for QAP, and so will need the derivative of the minimal scalar product

$$m(\lambda, \mu) := \sum_k \lambda_k \mu_{n-k+1}, \tag{3.17}$$

as well as the subdifferential of the lower bound for the linear part. (Note that the eigenvalues λ, μ are arranged in non-increasing order, and are functions of the perturbation vector $d = (e, g, r, s)$.) We use the fact that the derivative of a simple eigenvalue λ_i at 0 is (see, e.g., [7, 11])

$$u_i^t A'(0) u_i, \tag{3.18}$$

where u_i is the corresponding normalized eigenvector of A, and $A'(0)$ denotes the derivative of A at 0. In general, there exist differentiable functions λ_i which represent the eigenvalues and have derivatives given by (3.18), with u_i chosen from appropriate eigenspaces (see [7, Theorem 5.1]). Then

$$\nabla_e m = -2 \sum_k \mu_{n-k+1}(\text{sum}(u_k))u_k, \tag{3.19}$$

where u_k is the normalized eigenvector corresponding to λ_k and sum (u_k) denotes the sum of the components of u_k; and,

$$\nabla_r m = -\sum_k \mu_{n-k+1}\text{sq}(u_k), \tag{3.20}$$

where $\text{sq}(u_k)$ denotes the vector with components being the square of the components of u_k; $\nabla_g m$ and $\nabla_s m$ are defined symmetrically.

The corresponding derivatives of the range of values

$$v = \sum \lambda_k(\mu_k - \mu_{n-k+1}) \tag{3.21}$$

are

$$\nabla_e = 2 \sum_k \mu_{n-k+1}[u_k \text{ sum}(u_k) - u_{n-k+1} \text{ sum}(u_{n-k+1})]$$

and

$$\nabla_r v = \sum_k \mu_{n-k+1}[\text{sq}(u_k) - \text{sq}(u_{n-k+1})].$$

Since the diagonal elements and the row sums of the reduced matrices A and B are 0, we see that 0 is an eigenvalue with corresponding eigenvector 1. Since the eigenvectors are mutually orthogonal, we conclude that $\text{sum}(u_k) = 0$, for all but one k, which implies that all but one component of $\nabla_e v$ is 0. Moreover, if $0 = \mu_k$ and $2k = n+1$, then $\nabla_e v = 0$. The equivalent statement holds for $\nabla_g v$. Thus we cannot hope to get much improvement from adding E (or G) alone to the reduced matrix A (resp. B).

We can use the above derivatives to try and improve the bounds for the quadratic form $\text{tr } AXBX^t$. However, this affects the linear part of QAP according to the reduction rules (Rd1) and (Rd2). To find a lower bound for the linear part, we can solve the *linear sum assignment problem* (LSAP)

$$z := \min_{X \in \Omega, X \geq 0} \text{tr } CX^t. \tag{3.22}$$

This can be solved as an ordinary linear assignment problem. The constraint $X \in \Pi$ is relaxed to $X \in \Omega$ and $X \geq 0$. As is well known, every basic feasible solution of LSAP is a 0, 1-matrix and so is in Π. To improve the lower bound, we want to apply a steepest ascent algorithm to the sum of the two bounds $m + z$.

Now if A and B are reduced by (Rd1), then the matrix $\bar{C} = (\bar{c}_{ij})$ is defined by

$$\bar{c}_{ij} = c_{ij} + 2g_j \sum_k (a_{ik} - e_i - e_k) + 2e_i \sum_k b_{kj}.$$

If the matrices are further reduced by (Rd2), then $\bar{C} = (\bar{c}_{ij})$ becomes

$$\begin{aligned}\bar{c}_{ij} &= c_{ij} + 2g_j \sum_k (a_{ik} - e_i - e_k) + 2e_i \sum_k b_{kj} + (a_{ii} - 2e_i - r_i)s_j \\ &\quad + r_i(b_{jj} - 2g_j - s_j) + r_i s_j \\ &= c_{ij} + 2g_j \sum_k a_{ik} + 2e_i \sum_k b_{kj} - 2g_j \sum_k e_k - 2ne_i g_j \\ &\quad + (a_{ii} - 2e_i - r_i)s_j + r_i(b_{jj} - 2g_j - s_j) + r_i s_j. \end{aligned} \tag{3.23}$$

Thus the derivatives are

$$\begin{aligned} &\frac{\partial \bar{c}_{ij}}{\partial e_l} = -2g_j + 2\delta_{il}\left(\sum_k b_{kj} - ng_j - s_j\right), \\ &\frac{\partial^2 \bar{c}_{ij}}{\partial e_l\, \partial g_j} = -2 - 2\delta_{il}n, \qquad \frac{\partial^2 \bar{c}_{ij}}{\partial e_l\, \partial s_j} = -2\delta_{il}, \\ &\frac{\partial \bar{c}_{ij}}{\partial g_l} = 2\delta_{jl}\left(\sum_k a_{ik} - \sum_k e_k - ne_i - r_i\right), \qquad \frac{\partial^2 \bar{c}_{ij}}{\partial g_j\, \partial r_i} = -2, \\ &\frac{\partial \bar{c}_{ij}}{\partial r_l} = \delta_{il}(-s_j + b_{jj} - 2g_j), \qquad \frac{\partial^2 \bar{c}_{ij}}{\partial r_i\, \partial s_j} = -1, \\ &\frac{\partial \bar{c}_{ij}}{\partial s_l} = \delta_{jl}(-r_i + a_{ii} - 2e_i), \end{aligned} \tag{3.24}$$

where δ_{jl} is the Kronecker delta. The missing second derivatives are 0. The first derivatives evaluated at $e = g = r = s = 0$ are

$$\begin{aligned} &\frac{\partial \bar{c}_{ij}}{\partial e_l} = 2\delta_{il} \sum_k b_{kj}, \qquad \frac{\partial \bar{c}_{ij}}{\partial g_l} = 2\delta_{jl} \sum_k a_{ik}, \\ &\frac{\partial \bar{c}_{ij}}{\partial r_l} = \delta_{il} b_{jj}, \qquad \frac{\partial \bar{c}_{ij}}{\partial s_l} = \delta_{jl} a_{ii}. \end{aligned} \tag{3.25}$$

If X^* is a unique optimum of LSAP which does not change for small perturbations (e, g, r, s), then the gradient of the bound for the linear part exists and

$$\nabla z = \sum_{x_{ij}^* > 0} \nabla \bar{c}_{ij}, \tag{3.26}$$

where $\nabla \bar{c}_{ij}$ is defined by (3.25). We can now state the following about the direction of steepest ascent of the lower bound of QAP.

Proposition 3.1. *Suppose that the matrices A, B and C for QAP are given. Moreover, suppose that A and B have simple eigenvalues and that X^* is a unique optimum of LSAP. Then the direction of steepest ascent of the lower bound for QAP is given by $\nabla m + \nabla z$, where ∇m is defined as in* (3.19) *and* (3.20) *and ∇z is defined as in* (3.26).

Proof. Since the eigenvalues of A and B are simple, we have seen that ∇m exists. Since X^* is a unique solution, it does not change under small perturbations of C and so, ∇z is given by (3.26). □

Note that, since the diagonals and the row and column sums of both A and B are zero after the optimal reduction (3.14), we see that the derivatives in (3.25) are all zero. This also shows that ∇z, if it exists, is zero.

Now suppose that X^* is an optimal basic feasible solution of LSAP, i.e., $x_{ij}^* = 1$ for exactly n components of X^*. Let us apply a parametric linear programming approach to increase the optimal value z. The parameters here are quadratic. Since there are $2n-1$ linearly independent constraints, the solution X^* is a degenerate solution. Let $\mathrm{dr} = (\mathrm{dr}_i)$ and $\mathrm{dc} = (\mathrm{dc}_j)$ denote the row and column dual variables, respectively, and let $\mathcal{B} := \{(i,j): x_{ij}^* \text{ is a basic variable}\}$ denote the optimal basis set. The reduced costs are then

$$\tilde{c}_{ij} = c_{ij} - \mathrm{dr}_i - \mathrm{dc}_j.$$

If we perturb LSAP according to (3.23) in the direction (e, g, r, s) with step size $\alpha \geq 0$, then the new costs as a function of α are

$$\bar{c}_{ij} = c_{ij} + \alpha k_{ij} + \alpha^2 l_{ij}, \tag{3.27}$$

where

$$k_{ij} = 2\left[g_j \sum_k a_{ik} + e_i \sum_k b_{kj} \right] + a_{ii}s_j + r_i b_{jj}$$

and

$$l_{ij} = -2\left\{ g_j \sum_k e_k + 2ne_i g_j - (2e_i + r_i)s_j - r_i(2g_j + s_j) + r_i s_j \right\}. \tag{3.28}$$

With the above definitions we conclude the following about the stepsize α and the value $z(\alpha)$.

Proposition 3.2. *Let* dr^*, dc^* (*resp.* dr^{**}, dc^{**}) *denote the row and column dual variables found by replacing the costs* c_{ij} *by* k_{ij} (*resp.* l_{ij}), *but with the same basis set* $\mathcal{B}$. *Let the reduced costs for the perturbations be denoted*

$$\tilde{c}_{ij}^* = k_{ij} - \mathrm{dr}_i^* - \mathrm{dc}_j^*$$

and

$$\tilde{c}_{ij}^{**} = l_{ij} - \mathrm{dr}_i^{**} - \mathrm{dc}_j^{**}. \tag{3.29}$$

Let

$$\bar{\alpha} = \min_{i,j}\{\textit{nonnegative real roots of the quadratic } \tilde{c}_{ij} + \alpha \tilde{c}_{ij}^{*} + \alpha^2 \tilde{c}_{ij}^{**}\}. \quad (3.30)$$

Then the optimal basis $\mathcal{B}$ is unchanged if a stepsize $0 \leq \alpha \leq \bar{\alpha}$ is chosen. Moreover, the new optimal value of LSAP is

$$z(\alpha) = z(0) + \sum_{x_{ij}^{*}>0} (\alpha k_{ij} + \alpha^2 l_{ij}). \quad (3.31)$$

Proof. The result follows as in the usual LP sensitivity analysis, i.e., for $0 \leq \alpha \leq \bar{\alpha}$ with basis set $\mathcal{B}$, the optimality criteria is unchanged, since the reduced costs for the perturbed problem remain nonnegative. □

Thus we do not need to resolve LSAP if we restrict the stepsize $\alpha \leq \bar{\alpha}$. If we choose the stepsize $\alpha = \bar{\alpha}$, e.g., when $\bar{\alpha} = 0$, then we need to change the optimal basis. We choose (i, j) where $\bar{\alpha}$ is attained in (3.30), and let the corresponding x_{ij} enter the basis, and possibly increase to 1 from 0. Since the corresponding reduced cost $\tilde{c}_{ij} = 0$, this can be done.

In the case that the current X^*, obtained after a stepsize $\alpha = \bar{\alpha}$, is an optimal nonunique solution of LSAP, it can happen that z is non-differentiable. In this case we can use subdifferential calculus to find the direction of steepest ascent of z. (See [2, 16].)

Lemma 3.1. *The direction of steepest ascent of z is*

$$d = \arg \min_{\phi \in \partial z(0)} \|\phi\|,$$

where the subdifferential of z is

$$\partial z(0) = \left\{ \phi : \phi = \sum_{(i,j)} x_{ij}^{*} \nabla c_{ij},\ X^{*} \textit{ is an optimal solution of LSAP} \right\}.$$

Proof. Let $z(\alpha d)$ denote the solution of LSAP if the costs are changed using αd, where $\alpha \geq 0$, $d = (e, g, r, s)$, and the cost matrix C is changed using (3.23). Then the direction of steepest descent d of $-z$ is the solution of

$$\min_{\|d\| \leq 1} \lim_{\alpha \to 0^{+}} ((-z)(\alpha d) - (-z)(0))/\alpha, \quad (3.32)$$

or equivalently

$$\min_{\|d\| \leq 1} (-z)'(0; d), \quad (3.33)$$

where $(-z)'(0; d)$ denotes the directional derivatives of $-z$ in the direction d at 0.

Since the subdifferential is a convex compact set, we get

$$\min_{\|d\|\leq 1} (-z)'(0; d) = \min_{\|d\|\leq 1} \max_{\phi\in\partial(-z)(0)} \phi^t d$$

$$= \max_{\phi\in\partial(-z)(0)} \min_{\|d\|\leq 1} \phi^t d = \max_{\phi\in\partial(-z)(0)} \phi^t(-\phi/\|\phi\|)$$

$$= -\min_{\phi\in\partial z(0)} \|\phi\|. \tag{3.34}$$

Thus the direction of steepest ascent for z is found to be the element of minimal norm of the subdifferential

$$d = \arg\min_{\phi\in\partial z(0)} \|\phi\|. \tag{3.35}$$

Note that we found the direction of steepest descent for $-z$, since the theory of subgradients works for minimization rather than maximization (see [2]).

If the optimum X^* is unique, then the optimal value z is differentiable. By applying the chain rule we see that

$$\partial z/\partial e_l = \sum_{x_{ij}^*>0} \partial c_{ij}/\partial e_l. \tag{3.36}$$

The other partial derivations are found similarly. By (3.36) and since (see [2])

$$\partial z(0) = \operatorname{conv}\left\{\phi\colon \phi = \lim_n \nabla z(u_n),\ u_n \to 0\right\},$$

we conclude that

$$\partial z(0) = \left\{\phi\colon \phi = \sum_{(i,j)} x_{ij}^* \nabla c_{ij},\ X^* \text{ is an optimum solution of LSAP}\right\}, \tag{3.37}$$

where ∇c_{ij} denotes the gradient with respect to the change (e, g, r, s). Thus the direction of steepest ascent d in (3.35) is found by solving the quadratic program

$$\min\{\|\nabla C X^*\|^2\colon X^* \text{ solves LSAP}\}, \tag{3.38}$$

where ∇C is the Jacobian matrix with columns ∇c_{ij}. □

Since $\|\nabla C X^*\| \leq \|\nabla C\|\|X^*\|$, we can get an approximate solution to (3.38) by finding the minimal norm solution of LSAP,

$$\min\{\|X^*\|\colon X^* \text{ solves LSAP}\},$$

i.e., if there are k basic optimal solutions, X_i^*, then the minimal norm approximate solution to (3.38) is the mean

$$X^* = (1/k)\sum_{i=1}^{k} X_i^*. \tag{3.39}$$

To summarize, we now present an algorithm which increases the lower bound for QAP.

Input: Matrices A, B and C; w^* (estimate of the maximum of the objective function value); λ (fixed stepsize factor); i_{max} (maximum number of iterations); ε (bound on the norm of the subgradient);

Initialization: $i = 0$ (iteration number); $d_i = 0$ (perturbation vector);

Iteration: while $i < i_{max}$ do Steps 1 to 5:

Step 1. Compute the current objective function value with perturbation d_i, i.e., $w_i = m(d_i) + z(d_i)$;

Step 2. Compute the gradient ∇m at d_i and a subgradient ϕ of z at d_i;

Step 3. Test if the norm of the subgradient is small, i.e., if $\|\nabla m + \phi\| < \varepsilon$ then stop;

Step 4. Compute the increment

$$\delta d = \lambda \frac{(w^* - w_i)}{\|\nabla m + \phi\|^2} (\nabla m + \phi);$$

Step 5. Update

$$d_{i+1} \leftarrow d_i + \delta d, \qquad i \leftarrow i + 1.$$

The composition of the increment δd is done analogous to the subgradient optimization technique analysed in [9]. The sequence $(w_i)_{i \geq 1}$ is in general not monotonically increasing. It is not clear whether the function $m(d) + z(d)$ is indeed concave. Therefore, the convergence results of [9] cannot be applied. The practical experience, however, shows that the chosen approach is well suited for the problem.

We choose w^* to be a certain percentage p below the best known feasible solution of the QAP, where $0\% \leq p \leq 10\%$. The stepsize factor λ was chosen from the interval $[0.05, 0.5]$. The procedure stops when either the norm of the gradient gets too small or when a certain maximum number of iterations is reached.

The computational effort for one iteration can be estimated as follows

$m(d_i)$ requires the diagonalization of A and B, so is $O(n^3)$.

$z(d_i)$ is computed in $O(n^3)$ steps by solving LSAP. Given the eigenvectors and eigenvalues, ∇m and ϕ can be determined in $O(n^3)$, yielding a total of $O(n^3)$ per iteration.

All computations were programmed on an IBM PC/AT in FORTRAN. The running time to obtain all the improved eigenvalue bounds contained in Table 1 was less than one hour.

4. Numerical results and conclusion

We have taken the eight examples from Nugent et al. [14], which were also used in [4]. First we applied the above mentioned steepest ascent algorithm to increase the lower bound. The results are summarized in Table 1. Column GLB corresponds

Table 1
Lower bounds for QAPs

Size n	Best known value	GLB	EVB	MEVB
5	50	50	47	50
6	86	82	70	70
7	148	137	123	130
8	214	186	160	174
12	578	493	446	495
15	1150	963	927	989
20	2570	2057	2075	2229
30	6124	4539	4982	5349

to the Gilmore-Lawler bound applied to the unreduced problems. These are the bounds commonly used in branch and bound algorithms (see, e.g., [1, 17]). Column EVB contains the eigenvalue bounds obtained in [4] using the reduction scheme (3.14). MEVB corresponds to the new bound using the naïve ascent algorithm described above. 200 iterations were allowed ($i_{\max} = 200$), and $\varepsilon = 0.1$ was chosen. It turns out that for $n \geq 12$ the new bound outperforms all the existing bounds.

Ultimately, the new bound should be used in a branch and bound scheme, so the following two issues become relevant:

(i) efficient computation of the new bound;

(ii) computation of the new bound in each node of the branch and bound decision tree.

In answer to (i), Table 1 shows that the steepest ascent method of Section 3.3 works well, but takes many (200) iterations to reach the given bounds. Using more refined techniques for nonsmooth optimization allows us to speed up the computations. We used the BT-algorithm for minimizing a non-smooth function (see [19] and [21]) and obtained the results summarized in Table 2. We only focused on the problems for $n \geq 12$. The number of iterations was (arbitrarily) limited to 70. The first of the pair of numbers in each column $i \in \{1, 5, 10, 20, 70\}$ contains the current lower bound, while the second number gives the norm of a subgradient after iteration i. Note that after only five iterations, the lower bounds obtained by the BT-algorithm outperform the Gilmore-Lawler bound GLB for $n \geq 15$.

Table 2
The new bound using the BT-algorithm [19]

Size n	1	5	10	20	70
12	−909, 405.2	470, 20.3	474, 6.3	483, 3.7	498, 0.7
15	−1745, 630.2	967, 42.7	976, 8.5	982, 3.5	1002, 1.5
20	−3198, 1007.0	2166, 52.7	2197, 27.3	2214, 9.7	2286, 2.9
30	−7836, 1954.5	5239, 59.5	5263, 23.6	5326, 11.1	5443, 3.0

The actual CPU-time for the problem of size $n = 30$ was roughly 10 seconds per iteration on a SUN 3/60 workstation. This seems an acceptable price for the considerable improvement of the lower bound.

For question (ii), it should be pointed out that the improved eigenvalue bound can be used in a branch and bound scheme, because fixing an assignment leads to a QAP of lower dimension. So in principle, the improvement strategies can be used at each node of the decision tree. In practice, it seems advisable to spend a big effort at the root node of the branching tree to have a good quality bound to start with. Then the original data A, B, C should be replaced by the reduced data $\bar{A}$, $\bar{B}$, $\bar{C}$ corresponding to the best reduction at the root node. This implies that only a few iterations are likely to be necessary to obtain good bounds in subsequent nodes of the branching tree. As pointed out above, this can be achieved in a reasonable time frame. In summary, the proposed improvement technique has to be considered a new powerful tool to solve QAPs.

Acknowledgement

We thank Helga Schramm and Jochem Zowe for providing us with the BT-algorithm, the software for non-differentiable optimization.

References

[1] R.E. Burkard and U. Derigs, "Assignment and matching problems: solution methods for Fortran-programs," in: *Lecture Notes in Economics and Mathematical Systems No. 184* (Springer, Berlin, 1980).

[2] F.H. Clarke, *Optimization and Nonsmooth Analysis* (Wiley, New York, 1983).

[3] M.E. Dyer, A.M. Frieze and C.J.H. McDiarmid, "On linear programs with random costs," *Mathematical Programming* 35 (1986) 3–16.

[4] G. Finke, R.E. Burkard and F. Rendl, "Quadratic assignment problems," *Annals of Discrete Mathematics* 31 (1987) 61–82.

[5] G. Finke and E. Medova-Dempster, "Approximation approach to combinatorial optimization problems," Research Report, Technical University of Nova Scotia (Halifax, 1987).

[6] P.E. Gill, W. Murray and M.H. Wright, *Practical Optimization* (Academic Press, San Diego, 1981).

[7] B. Gollan, "Eigenvalue perturbations and non-linear parametric optimization," *Mathematical Programming Study* 30 (1987) 67–81.

[8] S. Hadley and H. Wolkowicz, "The Hessian of a function of the eigenvalues," Research Report, The University of Waterloo (Waterloo, Ontario, 1988).

[9] M. Held, P. Wolfe and P. Crowder, "Validation of subgradient optimization," *Mathematical Programming* 6 (1974) 62–88.

[10] A.J. Hoffman and H.W. Wielandt, "The variation of the spectrum of a normal matrix," *Duke Mathematical Journal* 20 (1953) 37–39.

[11] R. Horn and C. Johnson, *Matrix Analysis* (Cambridge University Press, Cambridge, 1985).

[12] L. Mirsky, "The spread of a matrix," *Mathematika* 3 (1956) 127–130.

[13] K.G. Murty *Linear Programming* (Wiley, New York, 1983).

[14] C.E. Nugent, T.E. Vollman and J. Ruml, "An experimental comparison of techniques for the assignment of facilities to locations," *Operations Research* 16 (1968) 150–173.

[15] M.L. Overton, "On minimizing the maximum eigenvalue of a symmetric matrix," *SIAM Journal of Matrix Analysis and Applications* 9 (1988) 256-268.
[16] R.T. Rockafellar, *Convex Analysis* (Princeton University Press, Princeton, NJ, 1970).
[17] C. Roucairol, "A parallel branch and bound algorithm for the quadratic assignment problem," *Discrete Applied Mathematics* 18 (1987) 211-225.
[18] S. Sahni and T. Gonzales, "P-complete approximation problems," *Journal of the Association for Computing Machinery* 23 (1976) 555-565.
[19] H. Schramm and J. Zowe, "A combination of the bundle approach and the trust region concept," in: J. Guddat et al., eds., *Advances in Mathematical Optimization* (Akademie Verlag, Berlin, 1988).
[20] H. Wolkowicz and G.P.H. Styan, "Bounds for eigenvalues using traces," *Linear Algebra and its Applications* 29 (1980) 471-506.
[21] J. Zowe, "The BT-algorithm for minimizing a nonsmooth functional subject to linear constraints," Working paper 1088, University of Bergen (Bergen, 1988).