J Glob Optim (2008) 42:609–617
DOI 10.1007/s10898-008-9338-2

Closed-form solution of a maximization problem

Richard W. Cottle · Ingram Olkin

Received: 26 March 2008 / Accepted: 12 August 2008 / Published online: 4 September 2008

Abstract According to the characterization of eigenvalues of a real symmetric matrix A, the largest eigenvalue is given by the maximum of the quadratic form $\langle xA, x\rangle$ over the unit sphere; the second largest eigenvalue of A is given by the maximum of this same quadratic form over the subset of the unit sphere consisting of vectors orthogonal to an eigenvector associated with the largest eigenvalue, etc. In this study, we weaken the conditions of orthogonality by permitting the vectors to have a common inner product r where $0 \leq r < 1$. This leads to the formulation of what appears—from the mathematical programming standpoint—to be a challenging problem: the maximization of a convex objective function subject to nonlinear equality constraints. A key feature of this paper is that we obtain a closed-form solution of the problem, which may prove useful in testing global optimization software. Computational experiments were carried out with a number of solvers.

Keywords Constrained optimization · Test problem · Quadratic forms · Intraclass correlation matrices · Eigenvalues · Eigenvectors

Mathematics Subject Classification (2000) 15A18 · 15A23 · 15A63 · 90C30 · 90C90

1 Introduction

According to the characterization of eigenvalues of a real symmetric matrix A—as developed, for example, by Bellman [2, Chapter 7]—the largest eigenvalue is given by the maximum of

We dedicate this paper to the memory of our great friend and colleague, Gene H. Golub.

R. W. Cottle (✉)
Department of Management Science & Engineering, Stanford University, Stanford, CA 94305-4026, USA
e-mail: rwc@stanford.edu

I. Olkin
Department of Statistics and School of Education, Stanford University, Stanford, CA 94305-4065, USA
e-mail: olkin@stat.stanford.edu

the quadratic form $\langle xA, x\rangle$ over the unit sphere; the second largest eigenvalue of A is given by the maximum of this same quadratic form over the subset of the unit sphere consisting of vectors orthogonal to an eigenvector associated with the largest eigenvalue, etc. Results like this and the Courant-Fischer Min-Max Theorem are, in a sense, related to (or shed some light on) the problem we study here.

Let A denote a positive semidefinite diagonal matrix of order $n \geq 2$. For a given scalar r such that $0 \leq r < 1$, we seek a matrix $X \in \mathbb{R}^{n\times n}$ with rows $x^1, \ldots, x^n$ that (globally) maximizes the function

$$F_n(X) = \sum_{i=1}^{n} \langle x^i A,\ x^i\rangle \tag{1}$$

subject to the constraints

$$g_i(X) = \langle x^i, x^i\rangle = 1, \quad i = 1, \ldots, n, \tag{2}$$

and

$$h_{ij}(X) = \langle x^i, x^j\rangle = r, \quad \text{for all } i \neq j. \tag{3}$$

Denote by $\mathrm{P}_n(r)$ the constrained maximization problem (1)–(3).

It is helpful—but not restrictive—to relabel the variables in such a way that the diagonal elements of A are arranged in descending order. Therefore, assume from the start that

$$A = \mathrm{Diag}(\alpha_1,\ \alpha_2,\ \ldots,\ \alpha_n) \quad \text{and} \quad \alpha_1 \geq \alpha_2 \geq \ldots \geq \alpha_n \geq 0. \tag{4}$$

(The α_i are, of course, eigenvalues.) For the sake of nontriviality, however, assume that $\alpha_1 > \alpha_n$; otherwise, the objective function $F_n(X) = n\alpha_1 = \mathrm{tr}(A)$ for all solutions of (2). As it happens, the maximization problem $\mathrm{P}_n(0)$ is also trivial, for then $R = I$, the feasible matrices X are orthogonal, and $\mathrm{tr}(XAX') = \mathrm{tr}(A)$, that is, the objective function is constant on the feasible region of $\mathrm{P}_n(0)$. For this reason, assume hereafter that $r > 0$. Yet, even though we disallow the cases $\alpha_1 = \alpha_n$ and $r = 0$, our conclusions are valid for these possibilities as well.

There are infinitely many matrices that satisfy the constraints of $\mathrm{P}_n(r)$. For our purpose, it is convenient to single out a particular matrix which we now describe.

Let $I \in \mathbb{R}^{n\times n}$ be the identity matrix and let $E \in \mathbb{R}^{n\times n}$ denote the matrix of all ones. Define the parameters

$$\mu = \sqrt{1-r} \quad \text{and} \quad \nu = \frac{\sqrt{1+(n-1)r} - \sqrt{1-r}}{n}. \tag{5}$$

Then put

$$C_n = \mu I + \nu E \in \mathbb{R}^{n\times n}. \tag{6}$$

The rows of this matrix satisfy the constraints (2) and (3). Accordingly, the feasible region of $\mathrm{P}_n(r)$—the set $\mathcal{X}$ of vectors satisfying its constraints—is nonempty and compact. Because the objective function F_n is continuous, its global extrema are attained on the feasible region, but not uniquely so. For instance, if X satisfies (2) and (3), then so does $-X$. Indeed, if X satisfies (2) and (3), then for any orthogonal $G \in \mathbb{R}^{n\times n}$, XG also satisfies (2) and (3). The nonuniqueness can arise in another way as well. Because $r < 1$, it follows that the vectors x^i and x^j cannot be equal when $i \neq j$. Problem $\mathrm{P}_n(r)$ can be stated more succinctly. Indeed, for any $n \times n$ matrix X, the objective function can be expressed conveniently as the trace of XAX':

$$F_n(X) = \sum_{i=1}^{n} \langle x^i A,\ x^i \rangle = \mathrm{tr}(XAX'). \tag{7}$$

The constraints can be stated (albeit with some redundancy) in the form

$$XX' = R = (1-r)I + rE. \tag{8}$$

In the statistical literature a matrix of this form is called an *intraclass correlation matrix.* Thus the problem $\mathrm{P}_n(r)$ is just

$$\text{maximize } \ \mathrm{tr}(XAX') \quad \text{subject to } XX' = R. \tag{9}$$

The eigenvalues of R are $1 + (n-1)r$ and $1 - r$, the latter with multiplicity $n - 1$, and consequently, for $0 < r < 1$, the matrix is positive definite.

We observe that the feasible region $\mathcal{X}$ of problem $\mathrm{P}_n(r)$ is nonconvex, which rules out the use of convex programming results. To make matters worse, the maximand is convex. It is conceivable, however, that the following observation might prove useful for computational purposes.

The diagonal matrix $D = \alpha_1 I - A$ is positive semidefinite; its diagonal entries $d_1, d_2, \ldots, d_n$ satisfy $0 = d_1 \le d_2 \le \ldots \le d_n$. Because of the n constraints (2), the *restriction* of F_n to the feasible region $\mathcal{X}$ is given by

$$F_n(X) = \sum_{i=1}^{n} \langle x^i A,\ x^i \rangle = \sum_{i=1}^{n} \langle x^i (\alpha_1 I - D),\ x^i \rangle = n\alpha_1 - \sum_{i=1}^{n} \langle x^i D,\ x^i \rangle.$$

This means that the *maximization* of the convex function F_n over $\mathcal{X}$ is equivalent to the *minimization* of the convex function $-F_n(X)$. Whether this alternate formulation would have computational merit is unclear, but, as our analysis will show, the question is moot because we obtain a closed-form solution of $\mathrm{P}_r(n)$ by matrix-theoretic methods. Accordingly, we proceed with the original problem statement.

Remark 1 A solution to (9) can be obtained by letting $R^{-1/2}X = Y$, where $R^{1/2}$ is the unique symmetric positive definite square root of R, in which case the constraint becomes $YY' = I$. Then

$$\mathrm{tr}(YAY'R) \le \sum_{i=1}^{n} \alpha_i \rho_i,$$

where $\alpha_1 \ge \ldots \ge \alpha_n,\ \ \rho_1 \ge \ldots \ge \rho_n$ are the ordered eigenvalues of A and R, respectively. Unless A or R has some special structure, a closed-form solution to (9) is not obtainable. In the present case, the structure of R does permit a closed-form solution of (9). This is significant because it enables verification of the effectiveness of global optimization algorithms and thereby represents a contribution to the literature of global optimization test problems, see [3]. In (12) we state the global maximum value of the objective function. This should prove useful in testing algorithms because of the nonuniqueness of the global optima in (9).

2 First-order optimality conditions

Problems such as $\mathrm{P}_n(r)$ can be treated by the familiar Lagrange multiplier rule. In this case, the theorem states that if $\tilde{X}$ is a local extremum of F_n subject to the constraints (2) and (3), and if the gradients of the constraint functions are linearly independent at $\tilde{X}$, then there exist $m = n(n+1)/2$ scalars $\lambda_1, \ldots, \lambda_n,\ \mu_{12}, \ldots, \mu_{n-1,n}$ such that $\tilde{X}$ satisfies

$$\frac{\partial F_n(X)}{\partial x^k} - \sum_{i=1}^{n} \lambda_i \frac{\partial g_i(X)}{\partial x^k} - \sum_{i<j} \mu_{ij} \frac{\partial h_{ij}(X)}{\partial x^k} = 0 \quad \text{for each } k = 1, \ldots, n, \tag{10}$$

where the operator $\partial/\partial x^k$ signifies taking the (partial) gradient with respect to the vector x^k.

It is customary to use the *stationarity conditions* (10) as a way of finding the matrix $\tilde{X}$ even though it is not clear that the *regularity condition* (linear independence of the gradients) holds at every feasible X. In any event, (10) is a set of necessary first-order optimality conditions. These first-order conditions do not distinguish between local maxima and minima. For that, one usually needs to use second-order optimality conditions.

When these first-order conditions hold, the special structure of problem $\mathrm{P}_n(r)$ permits us to make a statement about the objective function value at any feasible X at which (10) holds. Indeed for each $k = 1, \ldots, n$,

$$\frac{\partial F_n(X)}{\partial x^k} = 2x^k A, \qquad \text{and} \qquad \frac{\partial g_i(X)}{\partial x^k} = 2x^k \text{ (if } k = i),$$

whereas

$$\frac{\partial h_{ij}(X)}{\partial x^k} = x^j \quad (\text{if } k = i) \qquad \text{and} \qquad \frac{\partial h_{ij}(X)}{\partial x^k} = x^i \quad (\text{if } k = j).$$

Now suppose $\tilde{X}$ is a solution of (10). After multiplying each of these n equations by the corresponding $\tilde{x}^k$, we obtain

$$2\langle \tilde{x}^k A, \tilde{x}^k\rangle - 2\lambda_k \langle \tilde{x}^k, \tilde{x}^k\rangle - \sum_{i<k} \mu_{ik} \langle x^i, x^k\rangle - \sum_{k<j} \mu_{kj} \langle x^k, x^j\rangle = 0.$$

If $\tilde{X}$ is also a solution of (2) and (3), substitution into—and then adding—these equations leads to

$$2\sum_{k=1}^{n} \langle \tilde{x}^k A, \tilde{x}^k\rangle - 2\sum_{k=1}^{n} \lambda_k - 2r \sum_{i<j} \mu_{ij} = 0,$$

thereby yielding the interesting relationship

$$F_n(\tilde{X}) = \sum_{k=1}^{n} \lambda_k + r \sum_{i<j} \mu_{ij}. \tag{11}$$

There is, however, another expression for the optimal value of F_n. This formula—which will be developed later—is

$$F_n(X) = \mathrm{tr}(A) + r\,\mathrm{tr}(D). \tag{12}$$

A comparison of the formulas (11) and (12) raises the question: How are the individual Lagrange multipliers related to the diagonal elements of A and D? The answer for $n = 2$ is given below. For larger values of n, the Lagrange multiplier methodology becomes cumbersome. However, the form of the solution for $n = 2$ suggests a potential solution for $n > 2$. Indeed, we manage to obtain a closed form solution to $\mathrm{P}_n(r)$ by other means, and in so doing, verify (12).

2.1 The case of $n = 2$

The Lagrange multiplier rule is especially effective in the case of problem $\mathrm{P}_2(r)$ which calls for maximizing the convex function $F_2(x^1, x^2)$ subject to three nonlinear equality constraints:

$$g_1(x^1, x^2) = \langle x^1, x^1\rangle - 1 = 0, \tag{13}$$

$$g_2(x^1, x^2) = \langle x^2, x^2\rangle - 1 = 0, \tag{14}$$

$$h_{12}(x^1, x^2) = \langle x^1, x^2\rangle - r = 0. \tag{15}$$

In this case, one writes out the first-order optimality conditions as already given. This approach and a bit of elementary algebra lead to the discovery of the values of the Lagrange multipliers and two corresponding stationary points:

$$\widehat{X} = \begin{bmatrix} \sqrt{\frac{1+r}{2}} & \sqrt{\frac{1-r}{2}} \\ \sqrt{\frac{1+r}{2}} & -\sqrt{\frac{1-r}{2}} \end{bmatrix}, \qquad \check{X} = \begin{bmatrix} \sqrt{\frac{1-r}{2}} & \sqrt{\frac{1+r}{2}} \\ -\sqrt{\frac{1-r}{2}} & \sqrt{\frac{1+r}{2}} \end{bmatrix}.$$

Because

$$F_2(\widehat{X}) = a_1 + a_2 + r(a_1 - a_2), \qquad \text{and} \qquad F_2(\check{X}) = a_1 + a_2 - r(a_1 - a_2),$$

assumption (4) implies $F_2(\widehat{X}) > F_2(\check{X})$ meaning that $\widehat{X}$ is a maximum and $\check{X}$ is a minimum for $\mathrm{P}_2(r)$.

As a transition to the next section, we remark that

$$\widehat{X} = \begin{bmatrix} \frac{\sqrt{1+r}+\sqrt{1-r}}{2} & \frac{\sqrt{1+r}-\sqrt{1-r}}{2} \\ \frac{\sqrt{1+r}-\sqrt{1-r}}{2} & \frac{\sqrt{1+r}+\sqrt{1-r}}{2} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix}.$$

In accordance with (5) and (6), the left-hand factor in this matrix equation is C_2. This formula is the gist of what we aim to prove in general.

3 The general case

In this section we generalize the findings of Sect. 2, but not by means of Lagrange multipliers; that approach proved to be unwieldy and less promising than originally expected. Instead we apply some matrix theory and exhibit a closed-form solution of $\mathrm{P}_n(r)$. To this end, we recall the definitions of μ and ν given in (5) from which it will be seen that μ is a function of r, whereas ν is a function of both r and n.

From the form of the matrix R, and the simple fact that $E^2 = nE$, it can be verified by multiplication that

$$R^{1/2} = \sqrt{1-r}\,I + \left(\frac{\sqrt{1+(n-1)r} - \sqrt{1-r}}{n}\right)E = \mu I + \nu E;$$

which was denoted, in (6) by C_n; while feasible, this matrix is neither a maximum nor a minimum for $\mathrm{P}_n(r)$, yet it is well suited for obtaining the desired optimal solution. Indeed for any feasible matrix X, we have

$$XX' = C_n C_n'.$$

A result of Parker [6, Theorem 5] then implies that $X = C_nU$ for some orthogonal matrix U.

For this reason, the problem becomes one of finding an orthogonal matrix U for which

$$X = C_nU \tag{16}$$

maximizes $\operatorname{tr}(XAX')$.

Our choice for U is the transpose of the *Helmert matrix* H_n of order n. The first row of H_n is $(1/\sqrt{n})(1, \ldots, 1)$. For $i = 2, \ldots, n$, the ith row of H_n is $(1/\sqrt{(i-1)i})(1, \ldots, 1, -1, 0, \ldots, 0)$; the negative component is in the i-th position. For example,

$$H_5 = \begin{bmatrix} \frac{1}{\sqrt{5}} & \frac{1}{\sqrt{5}} & \frac{1}{\sqrt{5}} & \frac{1}{\sqrt{5}} & \frac{1}{\sqrt{5}} \\ \frac{1}{\sqrt{1\cdot 2}} & -\frac{1}{\sqrt{1\cdot 2}} & 0 & 0 & 0 \\ \frac{1}{\sqrt{2\cdot 3}} & \frac{1}{\sqrt{2\cdot 3}} & -\frac{2}{\sqrt{2\cdot 3}} & 0 & 0 \\ \frac{1}{\sqrt{3\cdot 4}} & \frac{1}{\sqrt{3\cdot 4}} & \frac{1}{\sqrt{3\cdot 4}} & -\frac{3}{\sqrt{3\cdot 4}} & 0 \\ \frac{1}{\sqrt{4\cdot 5}} & \frac{1}{\sqrt{4\cdot 5}} & \frac{1}{\sqrt{4\cdot 5}} & \frac{1}{\sqrt{4\cdot 5}} & -\frac{4}{\sqrt{4\cdot 5}} \end{bmatrix}.$$

It is well known and easy to verify that for every positive integer n, the Helmert matrix H_n belongs to the orthogonal group, $O(n)$. Hence when C_n is the feasible matrix defined in (6), the matrix

$$\widehat{X} = C_nH_n'$$

is also feasible. It is our contention that just as in the case of $n = 2$, the matrix $\widehat{X} = C_nH_n'$ solves $\mathrm{P}_n(r)$ for *every* integer $n \geq 2$. Hereafter, we write $\widehat{X} = CH'$ with the understanding that the integer $n \geq 2$ is fixed at the order under discussion.

Now, with a bit of straightforward matrix algebra, the elements of $\widehat{X}$ are given by

$$x_{ij} = \begin{cases} \sqrt{\frac{1+(n-1)r}{n}}, & \text{if } j = 1, \\ \sqrt{\frac{1-r}{(j-1)j}}, & \text{if } j > i \geq 1, \\ -(i-1)\sqrt{\frac{1-r}{(i-1)i}}, & \text{if } j = i > 1, \\ 0, & \text{if } i > j > 1. \end{cases} \tag{17}$$

Moreover, it can be verified that the value of the objective function corresponding to this matrix is given by the formula (12), that is

$$F_n(CH') = \operatorname{tr}(A) + r\operatorname{tr}(D). \tag{18}$$

By itself, this does not prove that $\widehat{X}$ is optimal for $\mathrm{P}_n(r)$. For that we turn to some well known matrix-theoretic results. The idea will be to show that the function $\operatorname{tr}(XAX')$ is bounded above by a quantity which it attains when $X = CH'$.

Our objective function is the trace of XAX', where A is a diagonal matrix with elements $a_1 \geq a_2 \geq \ldots \geq a_n \geq 0$, and $a_1 > a_n$. For the reason discussed above, we are interested in letting $X = CU$ where U is orthogonal. At this point we use the well known fact that $\mathrm{tr}(AB) = \mathrm{tr}(BA)$ when A and B are *arbitrary* conformable matrices. We have

$$\mathrm{tr}(XAX') = \mathrm{tr}(CUAU'C) = \mathrm{tr}(UAU'C^2) = \mathrm{tr}(UAU'R) = \mathrm{tr}(AU'RU).$$

Recall that $0 < r < 1$. The matrix R can be decomposed as $V'BV$ where B is diagonal and positive definite with $b_1 \geq b_2 \geq \cdots \geq b_n > 0$ and V is orthogonal. After replacing R by this factorization and then defining the orthogonal matrix $W = VU$, we see that

$$\mathrm{tr}(AU'RU) = \mathrm{tr}(AW'BW) = \mathrm{tr}(WAW'B).$$

In summary, we have

$$\mathrm{tr}(XAX') = \mathrm{tr}\,(WAW'B).$$

Accordingly, we can assert that

$$\max_{X \in \mathcal{X}} \mathrm{tr}(XAX') = \max_{W \in O(n)} \mathrm{tr}(WAW'B) \leq \max_{W,Z \in O(n)} \mathrm{tr}(WAZB) = \sum_{i=1}^{n} \sigma_i(A)\sigma_i(B),$$

where $\sigma_1(A), \ldots, \sigma_n(A)$ and $\sigma_1(B), \ldots, \sigma_n(B)$ denote the singular values of A and B, respectively.

The last equality follows from a theorem of von Neumann [7] (which is discussed by Marshall and Olkin [5, Chapter 20]) stating that for *arbitrary* real square matrices A and B of order n and for any $U,\ V \in O(n)$,

$$\mathrm{tr}(UAVB) \leq \max_{U,V \in O(n)} \mathrm{tr}(UAVB) = \sum_{i=1}^{n} \sigma_i(A)\sigma_i(B). \tag{19}$$

In our case, the matrices A and B are diagonal; A is positive semidefinite and B is positive definite. Hence the singular values $\sigma_i(A)$ and $\sigma_i(B)$ are just the diagonal elements a_i and b_i, respectively, of these matrices. Under these conditions, equality holds in (19) when the orthogonal matrices U and V are permutation matrices that align the diagonal elements so that $\sigma_1(A) \geq \ldots \geq \sigma_n(A)$ and $\sigma_1(B) \geq \ldots \geq \sigma_n(B)$. (See, e.g., Hardy, Littlewood and Pólya [4, Section 10.2] for the extremal values of such expressions.) When this is done, the right-hand side of (19) is

$$\begin{aligned}\sum_{i=1}^{n} \sigma_i(A)\sigma_i(B) &= a_1[1 + (n-1)r] + \sum_{i=2}^{n} a_i(1-r) \\ &= \sum_{i=1}^{n} a_i + \sum_{i=1}^{n} r(a_1 - a_i) = \mathrm{tr}(A) + r\,\mathrm{tr}(D).\end{aligned}$$

Altogether, this amounts to saying that

$$\max_{X \in \mathcal{X}} \mathrm{tr}(XAX') \leq \mathrm{tr}(A) + r\,\mathrm{tr}(D).$$

But, as we have already noted in (18), this upper bound on $F_n(X)$ is *attained* when $X = CH'$.

Remark 2 An alternative argument to the achievement of the bound $\mathrm{tr}(A) + r\,\mathrm{tr}(D)$ is that in (19), $\mathrm{tr}(WAW'B) = \sum_{i=1}^{n} \sigma_i(A)\sigma_i(B)$ for $W = I$ because the diagonals of A and B are both arranged in descending order. (The role of $Z \in O(n)$ in (19) is to assure the ordering we already have.)

The upshot of this argument is

Theorem 1 *For all r such that $0 < r < 1$ and all integers $n \geq 2$, the matrix $\widehat{X} = C_n H_n'$ solves problem $P_n(r)$, and the optimal value is given by $F_n(\widehat{X}) = tr(A) + r\, tr(D)$.*

Note that when $r = 0$ or $r = 1$, this expression for the optimal value of F_n is still valid.

4 Variants of the main problem

Problem $\mathrm{P}_n(r)$ has some interesting variants that are worthy of attention. For instance, a whole class of problems comes about by replacing the objective function $F_n(X) = \mathrm{tr}\,(XAX')$ by one of the elementary symmetric functions of the eigenvalues of XAX'. The k-th elementary symmetric function, denoted by E_k, of the n variables $\lambda_1, \ldots, \lambda_n$ is the sum of all monomials of the form $\lambda_{i_1}\lambda_{i_2}\ldots\lambda_{i_j}$. For $k = 1, 2, \ldots, n$, consider a family of maximization problems $\mathrm{P}_{n,k}(r)$:

$$\mathrm{P}_{n,k}(r): \quad \text{maximize } \; E_k(\lambda(XAX')) \quad \text{subject to } XX' = R, \tag{20}$$

where $\lambda(XAX') = (\lambda_1(XAX'), \ldots, \lambda_n(XAX'))$.

To bring this problem statement into line with that of Sect. 1, note that

$$E_k(\lambda(XAX')) = \mathrm{tr}((XAX')^{(k)}),$$

where $(XAX')^{(k)}$ denotes the k-th compound of XAX'. By the Binet-Cauchy Theorem for compound matrices $(XAX')^{(k)} = X^{(k)}A^{(k)}(X')^{(k)}$, which permits us to state the problem as

$$\mathrm{P}_{n,k}(r): \quad \text{maximize } \; \mathrm{tr}(X^{(k)}A^{(k)}(X')^{(k)}) \quad \text{subject to } XX' = R.$$

(For further details on compounds, see Aitken [1] or Marshall and Olkin [5].) The matrix $A^{(k)}$ is guaranteed to be diagonal, but its diagonal elements are not guaranteed to be in descending order. This property can be produced by a principal rearrangement. *To avoid extra notation, we assume that this has been done from the start.*

Let $Y = R^{-1/2}X$ so that $YY' = I$, and (20) becomes

$$\mathrm{P}_{n,k}(r): \quad \text{maximize } \; \mathrm{tr}(Y^{(k)}A^{(k)}Y^{(k)\prime}R^{(k)}) \quad \text{subject to } Y \in O(n).$$

Replace R by $H\Delta H'$, where $H \in O(n)$ is the Helmert matrix H_n of order n, and $\Delta = \mathrm{diag}(\lambda_1(R), \ldots, \lambda_n(R))$; more specifically, $\Delta = \mathrm{Diag}(1 + (n-1)r, 1-r, \ldots, 1-r)$. With $Z = UY \in O(n)$, the problem becomes

$$\mathrm{P}_{n,k}(r): \quad \text{maximize } \; \mathrm{tr}(Z^{(k)}A^{(k)}(Z')^{(k)}\Delta^{(k)}) \quad \text{subject to } Z \in O(n). \tag{21}$$

From (19) and the fact that $Z \in O(n)$ implies $Z^{(k)} \in O(C(n,k))$,

$$\begin{aligned}\max_{Z \in O(n)} \mathrm{tr}(Z^{(k)}A^{(k)}(Z')^{(k)}\Delta^{(k)}) &\leq \max_{\substack{Z \,\in\, O(n) \\ V \,\in\, O(C(n,k))}} \mathrm{tr}(Z^{(k)}A^{(k)}V\Delta^{(k)}) \\ &= \sum_{i=1}^{C(n,k)} \sigma_i(A^{(k)})\sigma_i(\Delta^{(k)}), \end{aligned} \tag{22}$$

which has the same form as (19). The eigenvalues (positive diagonal elements) of $\Delta^{(k)}$ are arranged in descending order. We have assumed that those of $A^{(k)}$ are as well. If this assumption is not adopted from the start, a suitable principal rearrangement will produce the desired ordering although the permutation matrix that accomplishes this task cannot be stated a priori. To complete the proof, we need to specify a matrix $X \in \mathcal{X}$ that leads (via the string of substitutions $X \to Y \to Z$) to equality in (21). We assert

Theorem 2 *For* $0 < r < 1$ *and* $1 \le k \le n$ *(*$n \neq 1$*), the matrix* $X = C_n H_n'$ *solves problem* $P_{n,k}(r)$ *provided the eigenvalues of* $A^{(k)}$ *are already arranged in descending order.*

5 Some computational experience

As suggested previously, one of the potential uses for an optimization problem with a closed-form solution is testing whether particular algorithms are capable of finding its certified optimal solution. With this in mind, a group of small test cases were run using 8 solvers available on the NEOS server http://neos.mcs.anl.gov/neos/solvers/index.html. Specifically, the solvers used in these experiments (and the categories under which they are listed) were

GLOBAL OPTIMIZATION:	NONLINEAR CONSTRAINED OPTIMIZATION:
ASA, PGAPack, and PSWarm	Ipopt, Minos, Mosek, Pennon, and SNOPT

The problems were all submitted to NEOS using the AMPL format. We are most grateful to Dongdong Ge and Yinghui Wu for their assistance in coding and running these trials. A total of 13 randomly generated problems for $n = 2, 3, 5$ and 10 were created and submitted to the 8 solvers. All problems were run with the corresponding "neutral" starting point $C_n = \mu I + \nu E$ (see (5) and (6)) and again from the (presumed) minimum. The latter were chosen as the matrices obtained from $X = C_n U$ (see (16)) where U was taken as (one of) the permutation matrices that produced the *smallest* objective function value. The outcome of the these trials can be summarized as follows. None of the three global optimization solvers identified above was able to solve any of the submitted problems. It appears that they were not designed to do so. As for the other five solvers, we found that only Pennon found a global maximum in each of the 13 cases. Ipopt and SNOPT were nearly as good, managing to obtain a global maximum in 10 and 9 instances, respectively. Minos solved 4 of the problems whereas Mosek solved none.

References

1. Aitken, A.C.: Determinants and Matrices. Oliver and Boyd, Edinburgh (1964)
2. Bellman, R.: Introduction to Matrix Analysis. McGraw-Hill, New York (1960)
3. Floudas, C.A., Adjiman, C.S., Esposito, W.R., Gümüs, Z.H., Harding, S.T., Klepeis, J.L., Meyer, C.A., Schweiger, C.A. (eds.): Handbook of Test Problems in Local and Global Optimization. Kluwer Academic Publishers, Boston (1999)
4. Hardy, G.H., Littlewood, J.E., Pólya, G.: Inequalities, 2nd edn. Cambridge University Press, Cambridge (1959)
5. Marshall, A.W., Olkin, I.: Theory of Majorization and its Applications. Academic Press, London (1979)
6. Parker, W.V.: The characteristic roots of matrices. Duke Math. J. **12**, 519–526 (1945)
7. von Neumann, J.: Some matrix inequalities and metrization of matrix-space. Tomsk Univ. Rev. **1**, 286–300 (1937)