Least Squares Matching Problems

R. W. Brockett

Department of Applied Sciences
Harvard University
Cambridge, Massachusetts 02138

Submitted by Jan C. Willems

ABSTRACT

A large class of matching problems can be thought of in the following way. One is given a group $G$ which acts on a Riemannian manifold and one is given two points $x_1$ and $x_2$ in the manifold. The problem is to find $g \in G$ such that $d(g(x_1), x_2)$ is as small as possible. If the group has finite cardinality, this is a combinatorial optimization problem whereas if $G$ is a Lie group, one can use the calculus to obtain necessary conditions. Our main results establish circumstances under which one can relax the combinatorial problem to an appropriate Lie group problem and then solve it by the method of steepest descent. An interesting class of applications involves embedding the permutation group on $n$ letters in the orthogonal group $O(n)$ with the latter acting on $R^{n(n+1)/2}$ via a symmetricized tensor product. The convergence properties of the descent equation which arises in this way are analyzed in some detail.

INTRODUCTION

In this paper we show that, in a number of cases, combinatorial optimization problems can be recast as function minimization problems with the functions to be minimized taking the form

$$\eta(\Theta) = \sum_{i=1}^{n} \text{tr} \Theta Q_i \Theta N_i - 2 \text{tr} M \Theta^T$$

*This work was supported in part by the U.S. Department of the Air Force under grant AFOSR-96-00197, in part by the U.S. Army Research Office under grant DAAL03-86-K-0171, and in part by the National Science Foundation under grant CDR-85-00108.
with $\Theta$ restricted to belong to a matrix Lie group. Not only is this of interest in itself, but because we are able to show that the method of steepest descent can be an effective tool for minimizing such functions, it immediately yields new computational procedures. Matching problems of the type studied here play an important role in computer vision, pattern analysis, and some types of learning theories. Often they are thought of as having both a continuous (geometrical) aspect involving the selection of a best-fitting transformation and a combinatorial aspect involving a search over all possible permutations of some finite set. Karmarkar's work [1] already suggested that the difference between continuous and combinatorial optimization may not be all that profound; our work reinforces this point of view while considerably extending its scope.

The mathematics at the heart of our approach may be explained as follows. When the $n(n-1)/2$ parameter orthogonal group acts on $\mathbb{E}^n$ according to $(\Theta, x) \mapsto \Theta x$, it includes the group of permutations as a subgroup in the obvious way. Linear functions on $\Theta x$ take the form $\langle \eta, \Theta x \rangle$ and, for $x$ fixed, they are extremized by a value of $\Theta$ which aligns $\Theta x$ and $\eta$. The best $\Theta$ is a permutation only when the components of $\eta$ are a scaled version of some reordering of the components of $x$. On the other hand, the $n(n-1)/2$ parameter orthogonal group acts on a $n(n+1)/2$ dimensional space of symmetric matrices via $(\Theta, Q) \mapsto \Theta^T Q \Theta$. This is just a concrete form of the restriction of the tensor product $\Theta \otimes \Theta$ to symmetric forms.

This representation also contains the permutation group on $n$ letters in that for $Q$, a diagonal matrix, $Q = \text{diag}(\alpha_1, \alpha_2, \ldots, \alpha_n)$, $\Theta^T Q \Theta$ is $\text{diag}(\alpha_{\pi(1)}, \alpha_{\pi(2)}, \ldots, \alpha_{\pi(n)})$ if $\Theta$ is a permutation matrix. The advantage of this representation is that linear functionals on $\Theta^T Q \Theta$, which can always be expressed as $\text{tr} \Theta^T Q \Theta N$, are extremized when $\Theta$ is a permutation matrix provided that $Q$ and $N$ are diagonal matrices and at least one of them has distinct eigenvalues. This provides a rich class of "calculus problems" having a strong combinatorial flavor, and hence it provides the means to formulate combinatorial optimization problems in a differentiable setting.

To introduce least squares matching, we recall a result of Nadás [2]. He showed that if $x_1, x_2, \ldots, x_p$ and $y_1, y_2, \ldots, y_p$ are given vectors in Euclidean $n$-space, then the translation $b$ and orthogonal matrix $\Theta$ which minimizes

$$\eta = \sum_{i=1}^{p} ||(\Theta x_i + b - y_i)||^2$$

can be expressed explicitly. One sees rather easily that the given data enter $\eta$ only through the matrices

$$N = \sum x_i x_i^T,$$

$$M = \sum x_i y_i^T.$$
and the centroids $\bar{x} = (x_1 + x_2 + \cdots + x_p)/p$, $\bar{y} = (y_1 + y_2 + \cdots + y_p)/p$. In fact the best values of $\Theta$ and $b$ do not even depend on $N$ and are simply

$$\Theta = (M^TM)^{-1/2}M$$

and

$$b = \bar{y} - \Theta \bar{x},$$

where the square root is the symmetric, positive definite square root. For a given $\Theta$, the choice of $b$ is unique. The choice of $\Theta$ is unique if $M^TM$ is nonsingular and has no repeated eigenvalues.

This can be viewed as solving a certain matching problem in which two patterns, taking the form of ordered sets of ordered $n$-tuples, are to be matched in a least squares sense, through the choice of an element of a given transformation group. It is a special case of the following more general situation.

**Matching Problem A.** Given a representation $\Phi$ of a Lie group acting on an inner product space $(V, \langle \cdot, \cdot \rangle)$ and given two ordered sets of points $\{x_1, x_2, \ldots, x_p\}$ and $\{y_1, y_2, \ldots, y_p\}$ in $V$, find $\phi \in \Phi$ such that

$$\eta = \sum_{i=1}^{p} \langle \phi(x_i) - y_i, \phi(x_i) - y_i \rangle$$

is as small as possible.

In this generality, it is not possible to express the solution of this problem explicitly, but algorithmic solutions based on the steepest descent are possible. Special cases lead to the consideration of the relationship between the standard polar representation of linear maps of an inner product space into itself, i.e. the process of expressing a map $M$ as $M = R\Theta = (M^TM)^{1/2}M(M^TM)^{-1/2}$, and the steady state solution of the (not necessarily symmetric) Riccati-like matrix differential equation $\dot{\Theta} = \Theta M\Theta - M^T$.

Problem A has no obvious combinatorial aspect in that the $x$’s and the $y$’s are ordered in advance and there is no question as to which of the $x$’s is to be paired with which of the $y$’s. In terms of the language commonly used in computer vision, it presumes that there is an *a priori* solution of the correspondence problem; otherwise it makes no sense to ask that $||\phi(x_i) - y_i||^2$ should be small.
The object of the present paper is to provide a unified point of view for understanding a range of matching problems, of which this is but one special case. The following question extends the scope further.

**Matching Problem B.** Given a representation \( \Phi \) of a Lie group acting on an inner product space \((V, \langle \cdot, \cdot \rangle)\) and given two sets of point \( \{ x_1, x_2, \ldots, x_p \} \) and \( \{ y_1, y_2, \ldots, y_p \} \) in \( V \), find \( \phi \in \Phi \) and a permutation \( \pi \in \Pi_p \), the set of all permutations of the integers \( 1, 2, \ldots, p \), such that

\[
\eta = \sum_{i=1}^{p} \langle \phi(x_{\pi(i)}) - u_i, \phi(x_{\pi(i)}) - u_i \rangle
\]

is as small as possible.

If, in such a situation, there exists a perfect match (corresponding to \( \eta = 0 \)), then of course the value of any symmetric function of the \( x \)'s will map under \( \phi \) to the value of the same symmetric function evaluated on the \( y \)'s. For example, the centroid of the \( x \)'s will map to the centroid of the \( y \)'s. If the \( x \)'s and \( y \)'s are points in \( \mathbb{E}^n \), then the second moments, \( \sum x_i x_i^T \), define \( n(n+1)/2 \) linearly independent, degree two polynomials which will map to the corresponding \( y \)-moments \( \sum y_i y_i^T \). If \( \phi(x) = Tx \) acts linearly, then \( T \) acts on the degree two forms via the restriction of the tensor product \( T \otimes T \) to symmetric forms. If \( n = 2 \), as would be the case for the most obvious computer vision application, then the three moments of inertia are symmetric functions of degree two and we can determine certain relations on \( T \) without finding \( \pi \in \Pi_p \). In real situations perfect matches will fail to exist. It is natural to pick \( T \) in such a way as to minimize the differences between the corresponding symmetric functions, which leads to a problem of the type considered under A.

Finally we consider a third matching problem which contains aspects of each of the previous two. Instead of postulating complete \textit{a priori} knowledge of the correspondence, or complete \textit{a priori} ignorance about it, we take an intermediate position. Suppose one is given a disjoint partition of the set of \( x \)'s and \( y \)'s, say \( \{ x_1, x_2, \ldots, x_r \}, \{ x_{r+1}, x_{r+2}, \ldots, x_p \}, \ldots, \{ x_t+1, x_t+2, \ldots, x_p \} \) and \( \{ y_1, y_2, \ldots, y_r \}, \{ y_{r+1}, y_{r+2}, \ldots, y_{r+s} \}, \ldots, \{ y_{t+1}, y_{t+2}, \ldots, y_p \} \). Suppose that the first subset of \( x \)'s is to be matched with the first subset of \( y \)'s, the second subset of \( x \)'s with the second subset of \( y \)'s, etc., but that within the subsets themselves, nothing is known about which of the \( x \)'s matches with which of the \( y \)'s. This happens, for example, when there are isolated clusters of points with an \textit{a priori} match between clusters but no information on individual points.
MATCHING PROBLEM C. Given a representation $\Phi$ of a Lie group acting on an inner product space $(V, \langle \cdot, \cdot \rangle)$ and given two collections of sets of points in $V$, $\{(x_{11}, x_{12}, \ldots, x_{1r}), (x_{21}, x_{22}, \ldots, x_{2r}), \ldots, (x_{p1}, x_{p2}, \ldots, x_{pr})\}$ and $\{(y_{11}, y_{12}, \ldots, y_{1r}), (y_{21}, y_{22}, \ldots, y_{2r}), \ldots, (y_{p1}, y_{p2}, \ldots, y_{pr})\}$, find $\phi$ in $\Phi$ and $\pi_1, \pi_2, \ldots, \pi_r$, permutations of $1, 2, \ldots, r$, such that

$$\eta = \sum_{i=1}^{p} \sum_{j=1}^{r} \left( \phi(x_{i\pi(j)}) - y_{ij} \right)$$

is as small as possible.

EXAMPLE. Suppose we have $p$ line segments in $E^2$ whose endpoints define $p$ pairs of points in $E^2$. We denote these points by $\{((a_i, b_i), (c_i, d_i))\}_{i=1}^{p}$. We do not, however, have an orientation for the line segments. We can say that the configuration admits a $(Z_2)^p \times \Pi_p$ symmetry, where $Z_2$ is the two element group which acts to reverse orientation and $\Pi_p$ acts to renumber the line segments. Clearly the average

$$[\bar{x}, \bar{y}] = \frac{1}{2p} \sum_{i=1}^{p} \begin{bmatrix} a_i \\ b_i \\ c_i \\ d_i \end{bmatrix}$$

is invariant with respect to renumbering of the pairs and/or changing the orientation of any given line segment. This is the only linear combination of $a_i, b_i, c_i, d_i$ which enjoys this property, since any linear combination which contains any $a_i$ must contain all $a_i$'s and all $c_i$'s with equal coefficients, etc. On the other hand, suppose that $\bar{x} = 0$. Equivalently, redefine $((a_i, b_i), (c_i, d_i))$ to be $((a_i - \bar{x}, b_i - \bar{y}), (c_i - \bar{x}, d_i - \bar{y}))$. Then the second moment associated with the point set consisting of all $2p$ points, ignoring pairings, is

$$\begin{bmatrix} m_{xx} \\ m_{xy} \\ m_{yy} \end{bmatrix} = \sum_{i=1}^{p} \begin{bmatrix} a_i^2 + c_i^2 \\ 2a_i b_i + 2c_i d_i \\ b_i^2 + d_i^2 \end{bmatrix}.$$

These three moments are also invariant. Given that the points appear in
pairs, we can form three more invariant quadratic forms

\[
\begin{bmatrix}
a_{xx} \\
a_{xy} \\
a_{yy}
\end{bmatrix} = \sum_{i=1}^{p} \begin{bmatrix}
(a_i - c_i)^2 \\
(a_i - c_i)(b_i - d_i) \\
(b_i - d_i)^2
\end{bmatrix}.
\]

These are the sum of the squares of the x-axis differences, the sum of the (signed) areas of the rectangles whose sides are aligned with the x and y axes and whose diagonal is defined by \(((a_i, b_i), (c_i, d_i))\), and the sum of the squares of the lengths of the y-axis differences, respectively.

One strategy to use in solving Problem C is to form a suitable set of symmetric functions of degree 1, 2, ..., \(p\) and then choose \(\phi\) in such a way as to obtain a least squares fit for the values of these functions. Having chosen \(\phi\), we can treat the determination of the \(\tau_i\) as a special case of Problem B.

**OPTIMALITY CONDITIONS FOR PROBLEM A**

We begin by recasting Problem A as a problem about minimizing \(\text{tr}(\Theta^T Q \Theta N) - 2 \text{tr} M \Theta^T\). Consider a least squares matching criterion defined through a symmetric positive definite \(n\) by \(n\) matrix \(Q\):

\[
\eta = \sum_{i=1}^{m} (\Theta x_i - b - y_i)^T Q(\Theta x_i - b - y_i),
\]

with the requirement that \(\Theta\) be an element of a closed matrix Lie group. We restrict our attention to closed matrix groups so as to avoid problems associated with subgroups of matrices which are not submanifolds, nonexistence of minima, etc.

It simplifies the notation somewhat to eliminate \(b\) and this can be done without loss of generality. The idea is to replace \(x\), \(y\), and \(\Theta\) by "homogeneous" versions

\[
x_i \mapsto \begin{bmatrix} x_i \\ 1 \end{bmatrix}, \quad y_i \mapsto \begin{bmatrix} y_i \\ 1 \end{bmatrix}
\]

and

\[
\Theta \mapsto \begin{bmatrix} \Theta \\ b \\ 1 \end{bmatrix}, \quad Q \mapsto \begin{bmatrix} Q & 0 \\ 0 & 0 \end{bmatrix}.
\]
This lets us write

$$\eta = \sum_{i=1}^{p} (\Theta x_i - y_i)^T Q(\Theta x_i - y_i).$$

It will be convenient to write this as

$$\eta = \text{tr} \left( \sum_{i=1}^{p} \Theta^T Q \Theta x_i x_i^T - Q(\Theta x_i y_i^T - y_i x_i^T \Theta) + y_i y_i^T \right),$$

where $\text{tr}$ denotes trace, and to introduce (here there is a slight shift in notation from the previous section)

$$N = \sum_{i=1}^{p} x_i x_i^T,$$

$$M = Q \sum_{i=1}^{p} y_i x_i^T.$$

Since $\text{tr} AB = \text{tr} BA$, we have

$$\eta = \text{tr}(\Theta^T Q \Theta N - 2M \Theta^T) + \sum_{i=1}^{p} y_i^T Q y_i.$$

In order to minimize $\eta$ we introduce the Lie algebra $L$ associated with the group of possible $\Theta$'s. An admissible perturbation in a neighborhood of $\Theta$ is a perturbation of the form $\Theta \rightarrow \Theta(I + \epsilon L)$ for $L$ in the Lie algebra and $\epsilon$ real. Expanding $\eta(\Theta(I + \epsilon L))$ to first order in $\epsilon$ shows that if $\Theta_0$ is optimal, then for all $L$ in the Lie algebra

$$\text{tr}(\Theta_0^T Q \Theta_0 L N + \Theta_0^T Q \Theta_0 N L^T - 2ML^T \Theta_0^T) = 0.$$

Recasting this slightly, we obtain the following theorem. [Here and below $\text{Gl}(n)$ stands for the set of all $n$ by $n$ nonsingular matrices.]

**Theorem 1.** Let $G$ be a Lie subgroup of $\text{Gl}(n)$. A necessary condition for $\Theta_0$ to be a stationary point for $\text{tr}(\Theta^T Q \Theta N - 2M \Theta^T)$ with respect to all
\( \Theta \) belonging to \( G \) is that

\[
\text{tr}(\Theta_0^TQ\Theta_0N - M^T\Theta_0)L = 0
\]

for all \( L \) in the Lie algebra of \( G \).

If there were no restrictions on \( L \), we would obtain

\[
N\Theta_0^TQ\Theta_0 = M^T\Theta_0,
\]

or

\[
\Theta_0 = Q^{-1}MN^{-1}.
\]

On the other hand, if, say, \( G \) is the set of all \( n \) by \( n \) matrices of determinant one, then \( L \) is the set of all matrices having zero trace and \( \Theta \) is the solution of

\[
N\Theta_0^TQ\Theta_0 - M^T\Theta_0 = \alpha I
\]

with \( \alpha \) being chosen so that \( \text{det} \Theta = 1 \). In general, if the Lie algebra has dimension \( k < n^2 \), then the variational equation represents \( k \) equations in \( k \) unknowns, and setting \( \Theta \) equal to \( Q^{-1}MN^{-1} \) will not meet the constraints. The remaining issue is, then, to obtain an effective algorithm for finding the best \( \Theta \) which does satisfy the constraints.

**THE DESCENT EQUATION FOR PROBLEM A**

In order to generate an effective algorithm for minimizing \( \text{tr}(\Theta^TQ\ThetaN - 2M\Theta^T) \) we give \( \text{Gl}(n) \) the structure of a Riemannian manifold and investigate the gradient vector field associated with the function to be minimized. As noted above, in \( \text{Gl}(n) \) we can parametrize a neighborhood of a nonsingular matrix \( G_0 \) by \( G = G_0e^\Omega \) with \( \Omega \) an arbitrary matrix. Moreover, if \( G \) is a closed Lie subgroup of \( \text{Gl}(n) \) with Lie algebra \( L \), then we can parametrize a neighborhood of \( G_0 \) in \( G \) by \( G_0e^\Omega \) with \( \Omega \) now in \( L \). We can consider the space of \( n \) by \( n \) matrices (or any subspace of it) as an inner product space with inner product

\[
\langle \Omega_1, \Omega_2 \rangle = \text{tr} \Omega_1^T\Omega_2.
\]

This inner product defines a positive definite quadratic form on the tangent space of the set of nonsingular matrices and hence lets us regard the tangent space as a Euclidean space. It also defines a Riemannian metric on the
LEAST SQUARES MATCHING PROBLEMS

manifold of nonsingular matrices; using it, we can pass from the linear term of the Taylor series expansion of a smooth function defined on $\text{Gl}(n)$ to a vector field on $\text{Gl}(n)$. The restriction of this inner product to the Lie algebra $L$ of a closed subgroup of $\text{Gl}(n)$ gives a Riemannian structure to that subgroup. Examples include the Euclidean group acting on $\mathbb{E}^n$, considered, as above, as a subgroup of $\text{Gl}(n+1)$, and the orthogonal group $O(n)$ considered as a subgroup of $\text{Gl}(n)$.

**Theorem 2.** Consider $\text{Gl}(n)$ as a Riemannian manifold with $(ds)^2 = -\text{tr}(d\Omega)^T(d\Omega)$, where $G_0e^\theta$ parametrizes a neighborhood of $G_0$. Let $L$ be a Lie algebra of matrices corresponding to a closed subgroup $G$ of $\text{Gl}(n)$. The gradient flow on $G$ corresponding to the function $\frac{1}{2}\text{tr}(\Theta^TQ\Theta N - 2M\Theta^T)$ is expressible as

$$\dot{\Theta} = \Theta \sum_{i=1}^{\dim L} L_i \text{tr} \left( L_i^T (N\Theta^T Q\Theta - M\Theta^T) \right),$$

where $\{L_i\}$ is any orthonormal basis for $L$.

**Proof.** As we have seen above, the gradient of $\eta$ evaluated at $L$, $\nabla \eta(L)$, is $\text{tr}[(Q\Theta N\Theta^T - M^T\Theta)L] = \left< \Theta N\Theta^T Q - \Theta^T M, L \right>$. This means that

$$\left< \Theta^{-1}\dot{\Theta} - \sum Q\Theta N\Theta^T + M^T\Theta, L \right> = 0,$$

and the formula in the theorem is simply a restatement of this. \hfill \blacksquare

It is possible to express this same idea without the use of the trace, and it is sometimes convenient to do so. Under the same hypothesis as appears in Theorem 2, there exists $P_i$ and $Q_i$ such that the gradient flow on $G$ corresponding to the function $\frac{1}{2}\text{tr}(\Theta^TQ\Theta N - 2M\Theta^T)$, $\Theta \in G$, is expressible in the form

$$\dot{\Theta} = \Theta \left( \sum P_i (Q\Theta N\Theta^T - M\Theta^T)Q_i \right).$$

Moreover, if $G$ is a subgroup of the orthogonal group, the gradient flow can be expressed as

$$\dot{\Theta} = \Theta \sum P_i (Q\Theta N\Theta^T - \Theta N\Theta^T Q + \Theta M^T - M\Theta^T)P_i^T.$$

To see this, note that for any given linear map $\phi$ of the space of $n$ by $n$ matrices into itself, there exists $n$ by $n$ matrices $P_i$ and $Q_i$, $i = 1, 2, \ldots, \alpha,$
such that the mapping is given by

\[ \phi(M) = \sum_{i=1}^{\alpha} P_i(M)Q_i. \]

In fact, if \( E_{ij} \) denotes the \( n \) by \( n \) matrix whose \( ij \)th element is one with all other elements being zero, then \( E_{ij}ME_{kl} = (M_{jk}) \cdot E_{il} \) is an operator which maps the \( jk \)th element of the domain onto the \( il \)th element of the range. Clearly a suitable linear combination of the form

\[ A(\cdot) = \sum \alpha_{ijkl} E_{ij}(\cdot)E_{kl} \]

will generate any linear map, and \( n^4 \) is an upper bound on the number of terms required. As for the second part, notice that for \( M = -M^T \) we have

\[ (E_{ij} - E_{ji})M(E_{ij} - E_{ji})^T = (E_{ij} - E_{ji})M(E_{ji} - E_{ij}) = (M_{ij})(E_{ji} - E_{ij}). \]

The function \( f(\Theta) = \text{tr} M\Theta^T \) with \( \Theta \) restricted to the orthogonal group is especially interesting. Since \( \text{tr}[M^T\Theta(I + \Delta)] = \text{tr} M^T\Theta + \text{tr} M^T\Theta\Delta = \text{tr} M^T\Theta + \langle I, M^T\Theta\Delta \rangle \), we see that the linear functional \( \langle \Theta^T M, \cdot \rangle \) represents the gradient. Projecting the linear functional \( \langle \Theta^T M, \cdot \rangle \) onto the tangent space at \( \Theta \) gives \( \langle \Theta^T M - M^T\Theta, \cdot \rangle \), and setting \( \Theta^T \dot{\Theta} \) equal to this gives the descent equation for \( \text{tr} M^T\Theta \):

\[ \Theta^T \dot{\Theta} = (M^T\Theta - \Theta^T M), \]

or

\[ \dot{\Theta} = (\Theta M^T\Theta - M). \]

This differential equation on the orthogonal group can be solved explicitly in terms of the pseudo-Hamiltonian pair

\[
\begin{bmatrix}
\dot{X} \\
\dot{P}
\end{bmatrix} =
\begin{bmatrix}
0 & -M^T \\
M & 0
\end{bmatrix}
\begin{bmatrix}
X \\
P
\end{bmatrix}
\]

by setting \( \Theta = XP^{-1} \). If we introduce a change of variables \( H = \Theta^T M \), we see that

\[ \dot{H} = H^2 - M^TM. \]

Clearly the singular values of \( H \) do not change as \( H \) evolves in time.

Since \( O(n) \) is a compact group, the solution of the descent equation will approach a stationary point. In view of stability considerations the equilib-
rium point will, except for trajectories initialized on thin (codimension one) subsets, not only be a stationary point but also be a local minimum of $\text{tr} \ M \Theta^T$. There is, however, no assurance that it will be a global minimum. More interesting claims can be made about the effectiveness of a descent algorithm if one can completely categorize the stationary points and enumerate the local minima. As it turns out, it is possible to be quite specific about stationary values of $\text{tr} \ M \Theta^T$. The following theorem completely resolves the matter in a "generic" situation. The main idea is to be found in Shayman's work on Morse functions for the classical groups. (See Theorem 2 of [3].)

**Theorem 3.** Let $M^T M$ be nonsingular with unrepeated eigenvalues $\lambda_1 > \lambda_2 > \cdots > \lambda_n > 0$. As $\Theta$ ranges over $O(n)$ there are $2^n$ values of $\Theta$ such that the gradient of $\text{tr} \ M \Theta^T$ vanishes. At these values $\text{tr} \ M \Theta^T = (\pm \lambda_1 \pm \lambda_2 \pm \cdots \pm \lambda_n)$. On each of the two connected components of $O(n)$ there is one local maximum, and at that point $\text{tr} \ M \Theta^T$ takes on the value of $\text{tr} \sqrt{M^T M}$ or $\sqrt{M^T M} - 2\lambda_n$. On each of the two connected components of $O(n)$ there is just one local minimum, at which point $\text{tr} \ M \Theta^T$ takes on the value $-\text{tr} \sqrt{M^T M}$ or $-\text{tr} \sqrt{M^T M} + 2\lambda_n$. The descent equation, initialized with $\det \Theta = +1$,

$$\dot{\Theta} = \Theta M \Theta - M^T,$$

has $2^{n-1}$ stationary points, exactly one of which is asymptotically stable. The $n(n-1)/2$ by $n(n-1)/2$ dimensional Hessian corresponding to the expansion of $\text{tr} \ M \Theta^T$ at $M \Theta^T = \text{diag}(\alpha_1 \lambda_1, \alpha_2 \lambda_2, \ldots, \alpha_n \lambda_n)$, with $\alpha_i = \pm 1$, has eigenvalues $(\alpha_i \bar{\lambda}_i + \alpha_j \bar{\lambda}_j)$ for $1 \leq i < j \leq n$.

**Proof.** This is just a calculation. If $M^T M$ has distinct positive eigenvalues, then it has $2^n$ distinct real symmetric square roots, and therefore there are $2^n$ different values of $\Theta$ which make $\text{tr} \ M \Theta^T$ stationary. If $M^T M$ has repeated eigenvalues, the situation is different and there will be manifolds of minima. They play a central role in the paper of Frankel [4].

**Optimality Conditions for Problem B**

We will now show that Problem B can be recast as one which depends on minimizing functions of the form

$$\eta = \sum_{i=1}^{n} \Theta^T Q_i \Theta N_i$$
with \( \Theta \) restricted to the orthogonal group. Before doing so, however, we
investigate the properties of this function. Already in 1937 von Neumann [5]
carried out an investigation of this situation in the special case \( n = 1 \).
(Actually von Neumann investigated a complex unitary version but it makes
little difference.) In order to appreciate this result it is helpful to recall a
result of Hardy, Littlewood, and Polya [6] which asserts that if \( x_1, x_2, \ldots, x_n \)
and \( y_1, y_2, \ldots, y_n \) are given sequences of real numbers and if \( \pi \in \Pi \), then in
order to maximize

\[
\eta = \sum x_{\pi(i)} y_i,
\]

\( \pi \) should reorder the \( x \)'s so as to match the largest of the \( x \)'s against the
largest of the \( y \)'s the next largest of the \( x \)'s against the next largest of the
\( y \)'s, etc.

**Theorem 4** (cf. von Neumann [5]). If \( Q \) and \( N \) are \( n \) by \( n \) symmetric
matrices with eigenvalues \( \lambda_1 > \lambda_2 > \cdots > \lambda_n \) and \( \mu_1 > \mu_2 > \cdots > \mu_n \), then
as \( \Theta \) varies over the orthogonal group, \( \text{tr} \Theta^T Q \Theta N \) has \( 2^n \cdot n! \) stationary
points, exactly \( 2^n \) of which are local maxima, at which points \( \text{tr} \Theta \Theta N \Theta^T \)
takes on the value \( \lambda_1 \mu_1 + \lambda_2 \mu_2 + \cdots + \lambda_n \mu_n \) and exactly \( 2^n \) of which are
local minima, at which points \( \text{tr} \Theta \Theta N \Theta^T \) takes on the value of \( \lambda_1 \mu_n + \lambda_2 \mu_{n-1} + \cdots + \lambda_n \mu_1 \). If \( \Psi \) and \( \Phi \) are orthogonal matrices such that
\( \Psi \Phi \Psi^T \) and \( \Phi N \Phi^T \) are diagonal, then the values of \( \Psi \) which make
\( \text{tr} \Theta \Theta N \Theta^T \) stationary are of the form \( \Theta = \Psi D \Pi \Phi \), with \( \Pi \) a permutation
matrix and \( D \) a diagonal square root of \( I \). The eigenvalues of the Hessian
of \( \text{tr} \Theta \Theta N \Theta^T \) at the value \( \Theta = \Psi D \Pi \Phi \) are the \( n(n-1)/2 \) products
\( D = \text{diag}(\alpha_1, \alpha_2, \ldots, \alpha_n) \); \( \alpha_i = \pm 1 \)

\[
\tau_{ij} = (\lambda_i - \lambda_j)(\mu_{\pi(i)} - \mu_{\pi(j)}), \quad 1 \leq i < j \leq n.
\]

In particular, the eigenvalues \( \tau_{ij} \) are all positive for just one choice of \( \pi \), and
they are all negative for just one choice of \( \pi \).

**Proof.** Suppose that \( \Psi \) and \( \Phi \) are orthogonal such that \( \Psi Q \Psi^T \) and
\( \Phi N \Phi^T \) are diagonal. Then

\[
\text{tr} \Theta \Theta N \Theta^T = \text{tr}(\Psi Q \Psi^T \Theta \Phi \Phi^T \Phi N \Phi^T \Theta \Theta^T \Psi^T) = \text{tr}(D \Theta \Theta^T D^T \Phi \Theta \Theta^T \Psi^T),
\]

where \( D \) is diagonal.
where $D_Q$ and $D_N$ are the diagonal forms of $Q$ and $N$. Redefine the arbitrary orthogonal matrix $\Theta$ as $\Psi \Theta \tilde{\Theta}^T = \tilde{\Theta}$. Now $\text{tr} D_Q \tilde{\Theta} D_N \tilde{\Theta}^T$ is clearly stationary if $\tilde{\Theta}$ is any permutation matrix and

$$\text{tr} D_Q \tilde{\Theta} D_N \tilde{\Theta}^T = \sum \lambda_i \mu_{\pi(i)}.$$ 

On the other hand, $\text{tr} D_Q \tilde{\Theta} D_N \tilde{\Theta}^T$ is not stationary unless $D_Q \tilde{\Theta} D_N \tilde{\Theta}^T$ is symmetric, which, because the eigenvalues of $Q$ and $N$ are unrepeated, happens only when $\tilde{\Theta} N \tilde{\Theta}^T$ is diagonal, i.e., when $\tilde{\Theta}$ is a permutation or when $\tilde{\Theta}$ is a diagonal square root of $I$ or when $\tilde{\Theta}$ is the product of a permutation and a diagonal square root of $I$.

We can interpret the results of Theorem 4 this way. Suppose that we let $N$ be the diagonal matrix $\text{diag}(1, 2, \ldots, n)$. If we chose $Q$ to be any symmetric matrix with eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$, then as time evolves, the descent equation

$$\dot{\Theta} = \Theta Q \Theta^T N \Theta - N \Theta Q$$

will evolve so as to make $\Theta Q \Theta^T$ approach a diagonal matrix whose entries are sorted by size with the largest appearing in the $11$ entry and the smallest in the $nn$ entry.

There are many ways to realize the group of permutations of a list of $m$ numbers as a subgroup of a representation of the orthogonal group. Here we are interested in

$$\text{diag}\left\{ \Theta \text{diag}(\alpha_1, \alpha_2, \ldots, \alpha_m) \Theta^T \right\} = \text{diag}(\beta_1, \beta_2, \ldots, \beta_m).$$

If $\Theta$ is a permutation matrix, this action defines a permutation of $(\alpha_1, \alpha_2, \ldots, \alpha_m)$. We can think of this as the restriction to the permutation matrices of the symmetrized tensor product of the orthogonal group with itself.

Suppose that we have two sets of $m$ points, $x_1, x_2, \ldots, x_m$ and $y_1, y_2, \ldots, y_m$, in $\mathbb{R}^n$, and that we want to find a permutation $\pi$ such that

$$\eta = \sum_{i=1}^{m} ||x_{\pi(i)} - y_i||^2$$

is as small as possible. In view of the fact that we can permute the elements
of a diagonal matrix by conjugation with a permutation matrix, i.e. in view of

$$\Pi^T \text{diag}(\alpha_1, \alpha_2, \ldots, \alpha_m) \Pi = \text{diag}(\alpha_{\pi(1)}, \alpha_{\pi(2)}, \ldots, \alpha_{\pi(n)})$$

we see that in the case \(n = 1\) this problem is solved by finding the permutation matrix \(\Pi\) which maximizes

$$\text{tr}[\Pi^T \text{diag}(x_1, x_2, \ldots, x_n) \Pi \text{diag}(y_1, y_2, \ldots, y_n)]$$

This comes about because

$$\sum_{i=1}^{m} (x_{\pi(i)} - y_i)^2 = \sum_{i=1}^{m} (x_{\pi(i)}^2 + y_i^2) - 2 \sum_{i=1}^{m} x_{\pi(i)} y_i$$

and the first two terms on the right are independent \(\pi\). If instead of restricting \(\Pi\) to be a permutation matrix we could allow it to be any orthogonal matrix, then of course

$$\max_{\Pi} \prod_{\pi(i)} y_i \leq \max_{\Theta} \text{tr}[\Theta \text{diag}(x_1, x_2, \ldots, x_m) \Theta^T \text{diag}(y_1, y_2, \ldots, y_m)]$$

However, in view of Theorem 4 this inequality is actually an equality. Embedding the permutation group in the orthogonal group and expanding the domain to the entire orthogonal group does not alter the value of the minimum or change the points at which it is achieved.

We now turn to a more general situation. In order to proceed we adopt a notation in which the \(j\)th element of the \(i\)th vector \(x_i\) is written as \(x_{ij}\) and similarly for \(y_i\). Thus

$$x_i = \begin{bmatrix} x_{i1} \\ x_{i2} \\ \vdots \\ x_{in} \end{bmatrix} \quad \text{and} \quad y_i = \begin{bmatrix} y_{i1} \\ y_{i2} \\ \vdots \\ y_{in} \end{bmatrix}$$

We may then write

$$\eta = \sum_{i=1}^{m} \|x_{\pi(i)} - y_i\|^2 = \sum_{i=1}^{m} \sum_{j=1}^{n} (x_{\pi(i)j} - y_{ij})^2$$

$$= \sum_{i=1}^{m} \sum_{j=1}^{n} x_{\pi(i)j}^2 + y_{ij}^2 - 2 \sum_{i=1}^{m} \sum_{j=1}^{n} x_{\pi(i)j} y_{ij}.$$
Thus to minimize $\eta$ we should maximize the trace of a sum of terms of the type dealt with above, i.e.

$$\sum_{j=1}^{n} \Pi^T \text{diag}(x_{1j}, x_{2j}, \ldots, x_{mj}) \Pi \text{diag}(y_{1j}, y_{2j}, \ldots, y_{mj}).$$

Again, by replacing $\Pi$ with $\Theta$ we can only increase the maximum value. If we define $Q_j$ and $N_j$ as

$$Q_j = \text{diag}(x_{1j}, x_{2j}, \ldots, x_{mj}),$$

$$N_j = \text{diag}(y_{1j}, y_{2j}, \ldots, y_{mj}),$$

then we have

$$\min_{\Pi} \|x_{\pi(i)} - y_i\|^2 \leq \sum_{i} \|x_i\|^2 + \|y_i\|^2 - 2 \max_{\Theta} \sum_{i=1}^{n} \text{tr} \Theta^T Q_j \Theta N_j,$$

Again, it turns out that for a number of problems of interest, this inequality is, in fact, an equality.

To explain this, notice that $m(m-1)/2$ by $m(m-1)/2$ Hessian matrix which represents the leading term in the expansion of

$$\phi(\Theta) = \sum_{i=1}^{n} \text{tr} \Theta^T Q_j \Theta N_j$$

valid near $\Theta = \Pi$ is just the sum of $n$ Hessians of the type which appear in Theorem 4. Since these Hessians, which may be thought of as $m(m-1)/2$ by $m(m-1)/2$ matrices, are all diagonal, their sums evaluated at a permutation $\Pi$ have eigenvalues

$$\lambda_{ik} = \sum_{j=1}^{n} (q_{ij} - q_{kj})(n_{ij} - n_{kj}).$$

If we identify $q_{ij}$ with $x_{ij}$ and $n_{ij}$ with $y_{ij}$, then it is natural to introduce an error term

$$x_{ij} = y_{ij} + e_{ij}.$$
That is, $e_{ij}$ is the error in the $j$th component of the $i$th vector $x_i$. This yields

$$\lambda_{ik} = \sum_{j=1}^{n} (n_{ij} - n_{kj})^2 + (e_{ij} - e_{kj})(n_{ij} - n_{kj}).$$

Identifying the first sum as the square of the distance $\|y_i - y_j\|^2$, we see that if for some $\pi$ the magnitude of error between $x_{\pi(i)}$ and $y_i$ is smaller than one-half the distance between $y_i$ and its nearest neighbor, then the permutation solution is a local minimum. The same argument applies with the roles of $x$ and $y$ reversed.

We summarize with a theorem.

**Theorem 5.** Let $x_1, x_2, \ldots, x_m$ and $y_1, y_2, \ldots, y_m$ be vectors in $\mathbb{E}^n$. For $i = 1, 2, \ldots, n$ let $Q_i$ and $N_i$ be $m$ by $m$ matrices defined as

$$Q_i = \text{diag}(x_{1i}, x_{2i}, \ldots, x_{mi}),$$

$$N_i = \text{diag}(y_{1i}, y_{2i}, \ldots, y_{mi}).$$

If the permutation $\pi$ which minimizes

$$\eta = \sum_{i=1}^{m} \|x_{\pi(i)} - y_i\|^2$$

produces an approximate match which is sufficiently accurate so that for all $i$

$$\|x_{\pi(i)} - y_i\|^2 \leq \frac{1}{2} \min_j \|x_i - x_j\|^2,$$

then equating $\Theta$ to the corresponding permutation matrix $\Pi$ yields a local minimum for

$$\eta(\Theta) = \sum_{i=1}^{n} \text{tr} \Theta^T Q_i \Theta N_i.$$

Thus the best match is an attractor for the descent equation on the group of $m$ by $m$ orthogonal matrices

$$\dot{\Theta} = \sum_{i=1}^{n} \Theta Q_i \Theta^T N_i \Theta - N_i \Theta Q_i.$$
REFERENCES


*Received September 1988; final manuscript accepted 18 January 1989*