THE ORTHOGONALLY CONSTRAINED REGRESSION REVISITED

MOODY T. CHU* AND NICKOLAY T. TRENDAFILOV†

Abstract. Two data analysis problems, the orthonormal Procrustes problem and the Penrose regression problem, are reconsidered in this paper. These problems are known in the literature for their importance as well as difficulty. This work presents a way to calculate the projected gradient and the projected Hessian explicitly. One immediate result of this calculation is the complete characterization of the first order and the second order optimality conditions for both problems. Another application is the natural formulation of a continuous steepest descent flow that can serve as a globally convergent numerical method. Applications are demonstrated by numerical examples.

Key words. Constrained regression, Procrustes rotation, Penrose regression, Projected gradient, Optimality condition, Continuous-time method.

1. Introduction. The problem of matching data matrices to maximal agreement by orthogonal rotations arises in many areas of disciplines. An concise and instructive discussion of its application to factor analysis and multidimensional scaling can be found in [7]. For general consideration, Ten Berge proposes a taxonomy of matching procedures based on gauging criteria, orthogonality, simultaneity, generality and symmetry [13]. In this paper we derive the first order and the second order optimality conditions for two of the most important cases in the family of problems. Our result includes as a special case what is already known in the literature, and appears to be the strongest possible provision for assessing a local optimizer.

The first problem of great interest is the matrix regression subject to orthonormality constraints, also known as the *orthonormal Procrustes problem* (OPP). For fixed matrices $A \in R^{n\times p}$ and $B \in R^{n\times q}$, the OPP is concerned with the following equality constrained optimization problem:

$$\text{Minimize} \quad \|AQ - B\| \tag{1}$$

$$\text{Subject to} \quad Q \in R^{p\times q},\ Q^TQ = I_q \tag{2}$$

where I_q stands for the $q \times q$ identity matrix.

The second problem considered is the more general *Penrose regression problem* (PRP). For fixed $A \in R^{n\times p}, C \in R^{q\times m}$ and $B \in R^{n\times m}$, the PRP concerns the optimization:

$$\text{Minimize} \quad \|AQC - B\| \tag{3}$$

$$\text{Subject to} \quad Q \in R^{p\times q},\ Q^TQ = I_q. \tag{4}$$

When $p = q$, the variable Q becomes an orthogonal matrix. This is the so called symmetric problem according to the taxonomy by Ten Berge, also known as the orthogonal Procrustes problem. In this case, the optimal solution for the OPP is well known.

* Department of Mathematics, North Carolina State University, Raleigh, North Carolina, 27695-8205. This research was supported in part by National Science Foundation under grant DMS-9422280.

† Computer Stochastic Laboratory, Institute of Mathematics, Bulgarian Academy of Sciences, P. O. Box 373, Sofia 1090, Bulgaria. This research was supported partially by contract MM-421 with Bulgarian Ministry of Education, Science and Technologies.

In fact, the solution is given by $Q = VU^T$ where V and U come from the singular value decomposition of $A^TB = V\Sigma U^T$ [8, 6]. The more interesting yet challenging case is the so called asymmetric problem where $p \neq q$ (We shall assume henceforth $p > q$.) We are not aware of any direct solution for the asymmetric problem. An iterative solution suggested by Green and Gower has been discussed in [13]. Iterative methods for the PRP can be found in [9] and [10]. Thus far, the best known optimality condition for the OPP is that at a minimizer Q the corresponding matrix B^TAQ must be symmetric and positive semi-definite. We shall see in the sequel that this necessary condition can be further strengthened.

Our approach in this paper is based on the recent development in the so called continuous realization methods. The idea has been to seek after the connection between the dynamics of a certain differential equation and a certain discrete numerical algorithm. See [2] for a recent review on this subject. The continuous formulation often has the advantage of furnishing better understanding about the properties of the solution for the underlying problem. Also, we have seen problems that are seemingly impossible to tackle by conventional discrete methods but can be solved by properly formulated continuous realization processes. In this paper we shall utilize the continuous realization approach to reconsider the above two problems. Specifically, we shall first study in the topology of the so call Stiefel manifold

$$\mathcal{O}(p,q) := \{Q \in R^{p\times q} | Q^TQ = I_q\}, \tag{5}$$

and then show that gradient flows for either problem can be reformulated without much difficulty. Our approach here is similar in spirit to that in [1], but with much more manipulation. An important feature of our approach is that the problems can be treated in a completely unified manner. We shall derive matrix differential equations describing gradient flows for both problems. Following the solution of these equations lead us to the solution to the corresponding regression problems. More importantly, we can specify the necessary and sufficient optimality conditions characterizing the optimizers.

This paper is organized as follows: Some important topological properties, particularly the tangent space, of the manifold $\mathcal{O}(p,q)$ are briefly discussed in §2. These properties are then applied to the OPP in §3. We derive the projected gradient and the project Hessian of a certain objective function in great details that enable us to specify the first order and the second order sufficient and necessary conditions for a local minimizer of the OPP. The same ideas can be used for the PRP. We state the results in §4 without the tedious technicality. Finally, we present some numerical experiment in §5.

2. Stiefel Manifold. The set $\mathcal{O}(p,q)$ of all $p \times q$ real matrices with orthonormal columns forms a smooth manifold of which the topology has been considered by Stiefel in [12]. For a quick grasp of its main properties, we recommend the paper by Edelman et al [3]. We outline in this section some main points that will be useful for the discussion.

We shall regard $\mathcal{O}(p,q)$ as embedded in the pq dimensional Euclidean space $R^{p\times q}$ equipped with the Frobenius inner product:

$$\langle X, Y\rangle := \text{trace}(XY^T)$$

for any $X, Y \in R^{p\times q}$. The tangent space $\mathcal{T}_Q\mathcal{O}(p,q)$ of $\mathcal{O}(p,q)$ at any $Q \in \mathcal{O}(p,q)$ is given by

$$\mathcal{T}_Q\mathcal{O}(p,q) = \{H \in R^{p\times q} | Q^T H \text{ is skew-symmetric}\}.$$

To further characterize a tangent vector, we recall that a least squares solution X to the equation

$$MX = N$$

is given by

$$X = M^\dagger K + (I - M^\dagger M)W$$

where $M^\dagger$ is the Moore-Penrose inverse of M, I is an identity matrix, and W is an arbitrary matrix of proper dimension. Applied to our case with $M = Q^T$ where $Q \in \mathcal{O}(p,q)$ and $N = K \in R^{q\times q}$ where K is skew-symmetric, we note that $(Q^T)^\dagger = Q$. The following theorem therefore follows.

THEOREM 2.1. *Any tangent vector $H \in \mathcal{T}_Q\mathcal{O}(p,q)$ has the form*

$$H = QK + (I_p - QQ^T)W \tag{6}$$

where $K \in R^{q\times q}$ and $W \in R^{p\times q}$ are arbitrary, and K is skew-symmetric.

For convenience, we shall abbreviate I_p as I. Denote

$$\mathcal{S}(q) := \{\text{all symmetric matrices in } R^{q\times q}\}.$$

It is not difficult to check by the dimension counting arguments that the normal spaces of $\mathcal{O}(p,q)$ at any orthonormal matrix Q is given by:

$$\mathcal{N}_Q\mathcal{O}(p,q) = Q\mathcal{S}(q). \tag{7}$$

It is worthy to mention the following decomposition of the space $R^{p\times q}$.

THEOREM 2.2. *The space $R^{p\times q}$ can be written as the direct sum of three mutually perpendicular subspaces*

$$R^{p\times q} = Q\mathcal{S}(q) \oplus Q\mathcal{S}(q)^\perp \oplus \mathcal{N}(Q^T), \tag{8}$$

where $\mathcal{S}(q)^\perp$ is simply the orthogonal complement of $\mathcal{S}(q)$ with respect to the Frobenius inner product and $\mathcal{N}(Q^T) := \{X \in R^{p\times q} | Q^T X = 0\}$.

We therefore are able to define the following projections.

COROLLARY 2.3. *Let $Z \in R^{p\times q}$. Then*

$$\pi_{\mathcal{T}}(Z) := Q\frac{Q^T Z - Z^T Q}{2} + (I - QQ^T)Z \tag{9}$$

defines the projection of Z onto the the tangent space $\mathcal{T}_Q\mathcal{O}(p,q)$. Similarly,

$$\pi_{\mathcal{N}}(Z) := Q\frac{Q^T Z + Z^T Q}{2} \tag{10}$$

defines the projection of Z onto the normal space $\mathcal{N}_Q\mathcal{O}(p,q)$.

3. Orthogonal Procrustes Problem. Let $A \in R^{n\times p}$ and $B \in R^{n\times q}$ be given and fixed. Consider the function $F : R^{p\times q} \longrightarrow R$ defined by

$$F(Q) := \frac{1}{2}\langle AQ - B, AQ - B\rangle . \tag{11}$$

Apparently, the OPP is equivalent to the minimization of the function $F(Q)$ over the feasible set $\mathcal{O}(p,q)$.

With respect to the Frobenius inner product, the gradient $\nabla F(Q)$ of the objective function $F(Q)$ should be interpreted as the matrix

$$\nabla F(Q) = A^T(AQ - B). \tag{12}$$

Suppose the projection $g(Q)$ of the gradient $\nabla F(Q)$ onto the tangent space $\mathcal{T}_Q\mathcal{O}(p,q)$ can be computed explicitly. Then the differential equation

$$\frac{dQ}{dt} = -g(Q) \tag{13}$$

naturally defines the steepest descent flow for the function F on the feasible set $\mathcal{O}(p,q)$. By applying Corollary 2.3, this projected gradient $g(Q)$ indeed is given by

$$\begin{aligned} g(Q) &= \frac{Q}{2}\left(Q^T\nabla F(Q) - (\nabla F(Q))^TQ\right) + (I - QQ^T)\nabla F(Q) \\ &= \frac{Q}{2}(B^TAQ - Q^TA^TB) + (I - QQ^T)A^T(AQ - B). \end{aligned} \tag{14}$$

Thus the differential equation

$$\frac{dQ}{dt} = \frac{Q}{2}(Q^TA^TB - B^TAQ) - (I - QQ^T)A^T(AQ - B) \tag{15}$$

defines a steepest descent flow on the manifold $\mathcal{O}(p,q)$ for the objective function F in (11). Starting with an initial point, say,

$$Q(0) = \begin{pmatrix} I_q \\ 0 \end{pmatrix}, \tag{16}$$

we may use this flow to approximate a solution to the OPP.

Another advantage of knowing the projected gradient $g(Q)$ explicitly is that we can describe the first order optimality condition for a stationary point.

THEOREM 3.1. *For $Q \in \mathcal{O}(p,q)$ to be a stationary point of the OPP, the following two conditions must hold simultaneously:*

(a). B^TAQ is symmetric, and

(b). $(I - QQ^T)A^T(AQ - B) = 0$.

Proof. For Q to be a stationary point, it is necessarily that $g(Q) = 0$. Since the two factors in (14) are mutually perpendicular, each individual factor must zero by itself. The condition (a) follows from pre-multiplying the first factor in (14) by Q^T. ∎

We remark here that the condition (a) in Theorem 3.1 is a known fact in the literature [13], but the condition (b) somehow has been overlooked. Our result apparently is new and rectify that defect.

We can also derive an explicit projected Hessian formula to further identify the stationary points. The development is based on an extension idea discussed in [1] that significantly cuts short the work of what would be if going through the classical notion of Lagrangian multipliers. The task is to first extend the function $g(Q)$ defined on the manifold $\mathcal{O}(p,q)$ to the function $G(Z)$ defined on the entire space $R^{p\times q}$ by:

$$G(Z) := \frac{Z}{2}(B^TAZ - Z^TA^TB) + (I - ZZ^T)A^T(AZ - B). \tag{17}$$

The Fréchet derivative of G at Z acting on a general H is:

$$\begin{aligned} G'(Z)H &= \frac{1}{2}\Big(H(B^TAZ - Z^TA^TB) + Z(B^TAH - H^TA^TB)\Big) \\ &\quad -(HZ^T + ZH^T)A^T(AZ - B) + (I - ZZ^T)A^TAH. \end{aligned} \tag{18}$$

We then consider the action of G' at a stationary point $Z = Q \in \mathcal{O}(p,q)$ on any tangent vector $H = QK + (I - QQ^T)W$ for arbitrary $K \in \mathcal{S}(q)^\perp$ and $W \in R^{p\times q}$. This action, according to the arguments in [1], produces exactly the same action of the projected Hessian for the OPP. For convenience, we divide the action into four parts:

First of all, we observe

$$\begin{aligned} \langle G'(Q)QK, QK\rangle &= \Big\langle \frac{1}{2}\Big(QK(B^TAQ - Q^TA^TB) + Q(B^TAQK - K^TQ^TA^TB)\Big) \\ &\quad -(QKQ^T + QK^TQ^T)A^T(AQ - B) + (I - QQ^T)A^TAQK, QK\Big\rangle \\ &= \Big\langle \frac{1}{2}\Big(Q(B^TAQK - K^TQ^TA^TB)\Big) + (I - QQ^T)A^TAQK, QK\Big\rangle \\ &= \Big\langle \frac{1}{2}\Big(B^TAQK - K^TQ^TA^TB\Big), K\Big\rangle \\ &= \Big\langle B^TAQK, K\Big\rangle. \end{aligned} \tag{19}$$

In the above, the second equality follows from condition (a) in Theorem 3.1 and the fact that K is skew-symmetric. The third equality utilizes the adjoint property

$$\langle M, NP\rangle = \Big\langle N^TM, P\Big\rangle$$

for any three conformal matrices M, N and P, and the fact that

$$Q^T(I - QQ^T) = 0. \tag{20}$$

Secondly, we have

$$\begin{aligned} \Big\langle G'(Q)QK, (I - QQ^T)W\Big\rangle &= \Big\langle \frac{Q}{2}\Big((B^TAQK - K^TQ^TA^TB)\Big) \\ &\quad +(I - QQ^T)A^TAQK, (I - QQ^T)W\Big\rangle \\ &= \Big\langle (I - QQ^T)A^TAQK, (I - QQ^T)W\Big\rangle \\ &= \Big\langle A^TAQK, (I - QQ^T)W\Big\rangle. \end{aligned} \tag{21}$$

In the above, the third equality follows from (20) and the fact that

$$(I - QQ^T)^2 = I - QQ^T. \tag{22}$$

Thirdly, we have

$$\begin{aligned}
&\left\langle G'(Q)(I-QQ^T)W, (I-QQ^T)W\right\rangle \\
= &\left\langle \frac{Q}{2}\left((B^TA(I-QQ^T)W - W^T(I-QQ^T)A^TB)\right)\right. \\
&-\left((I-QQ^T)WQ^T + QW^T(I-QQ^T)\right)A^T(AQ-B) \\
&\left.+(I-QQ^T)A^TA(I-QQ^T)W, (I-QQ^T)W\right\rangle \\
= &\left\langle -WQ^TA^T(AQ-B) + A^TA(I-QQ^T)W, (I-QQ^T)W\right\rangle. \qquad (23)
\end{aligned}$$

Again, we have used (20) and (22) in the above to induce the last equality.

Finally, we observe that

$$\begin{aligned}
&\left\langle G'(Q)(I-QQ^T)W, QK\right\rangle \\
= &\left\langle \frac{Q}{2}\left((B^TA(I-QQ^T)W - W^T(I-QQ^T)A^TB)\right)\right. \\
&\left.-\left((I-QQ^T)WQ^T + QW^T(I-QQ^T)\right)A^T(AQ-B), QK\right\rangle \\
= &\left\langle B^TA(I-QQ^T)W - W^T(I-QQ^T)A^T(AQ-B), K\right\rangle \\
= &\left\langle -W^T(I-QQ^T)A^TAQ, K\right\rangle = \left\langle A^TAQK, (I-QQ^T)W\right\rangle. \qquad (24)
\end{aligned}$$

Putting all pieces together, the quantity

$$\left\langle G'(Q)(QK + (I-QQ^T)W), QK + (I-QQ^T)W\right\rangle \tag{25}$$

represents the action of the projected Hessian of F at tangent vectors, and hence we obtain the following theorem from the constrained optimization theory [4, Section 3.4]:

THEOREM 3.2. *At a stationary point $Q \in \mathcal{O}(p,q)$ satisfying Theorem 3.1, a second-order necessary condition for Q to be a minimizer of the OPP is that the inequality*

$$\begin{aligned}
&\left\langle B^TAQK, K\right\rangle + 2\left\langle A^TAQK, (I-QQ^T)W\right\rangle + \left\langle A^TA(I-QQ^T)W, (I-QQ^T)W\right\rangle \\
&-\left\langle (I-QQ^T)WQ^TA^T(AQ-B), (I-QQ^T)W\right\rangle \geq 0 \qquad (26)
\end{aligned}$$

holds for all skew-symmetric matrices $K \in R^{q\times q}$ and arbitrary matrices $W \in R^{p\times q}$. If (26) holds with the strict inequality, then it is sufficient that Q is a local minimizer of the OPP.

The condition in Theorem 3.2 is the best one can hope for from the general theory of nonlinear equality constrained optimization. The complete and explicit characterization in (26) appears to be known the first time. Note that we have purposefully rewritten the last term on the left-hand side of (26) in the quadratic form of $(I - QQ^T)W$.

Two special cases of Theorem 3.2 are of great interest. With $W = 0$, the necessary condition in (26) becomes

$$\left\langle B^T AQK, K \right\rangle \geq 0 \tag{27}$$

for all skew-symmetric matrices K in $R^{q\times q}$. With $K = 0$, the necessary condition becomes

$$\left\langle A(I-QQ^T)W, A(I-QQ^T)W \right\rangle \geq \left\langle (I-QQ^T)WQ^TA^T(AQ-B), (I-QQ^T)W \right\rangle \tag{28}$$

for arbitrary matrix $W \in R^{p\times q}$. We shall show below that (27) is equivalent to the condition known in the literature [13]. But (28) apparently is new.

Observe that $K \in R^{q\times q}$ is skew-symmetric if and only if its singular value decomposition is of the form

$$K = U\Sigma J^T U^T, \tag{29}$$

where U is an $q \times q$ orthogonal matrix, $\Sigma = \operatorname{diag}\{\sigma_1, \ldots, \sigma_q\}$ contains singular values with $\sigma_{2i-1} = \sigma_{2i}$, $i = 1, \ldots, \lfloor \frac{q}{2} \rfloor$, and $\sigma_q = 0$ if q is odd, and

$$J = \begin{cases} I_{\lfloor \frac{q}{2} \rfloor} \otimes \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, & \text{if } q \text{ is even;} \\ \begin{bmatrix} I_{\lfloor \frac{q}{2} \rfloor} \otimes \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} & 0 \\ 0 & 1 \end{bmatrix}, & \text{if } q \text{ is odd.} \end{cases}$$

It follows that

$$K^2 = -U\Sigma^2 U^T \tag{30}$$

which in fact is the spectral decomposition of K^2. We know from Theorem 3.1 that B^TAQ is necessarily symmetric at any stationary point Q. Let

$$B^TAQ = V\Lambda V^T$$

denote the corresponding spectral decomposition. Note that

$$\begin{aligned} \left\langle B^TAQK, K \right\rangle &= \left\langle V\Lambda V^T, U\Sigma^2 U^T \right\rangle \\ &= \sum_{i=1}^{q} \lambda_i \left(\sum_{t=1}^{q} p_{it}^2 \sigma_t^2 \right) \end{aligned}$$

where $P = (p_{ij}) = V^T U$. Since the orthogonal matrix $P \in R^{q\times q}$ can be arbitrary, in order to maintain the inequality in (27) it must be that all entries $\lambda_1, \ldots, \lambda_q$ are nonnegative. We have proved the following result:

Corollary 3.3. *A necessary condition for the stationary point $Q \in \mathcal{O}(p, q)$ to be a solution of the OPP is that the matrix B^TAQ be positive semi-definite and that the inequality (28) be held for arbitrary $W \in R^{p\times q}$.*

4. Penrose Regression Problem. The more general Penrose regression problem can be considered parallel to the orthonormal Procrustes problem. For given matrices $A \in R^{n\times p}, C \in R^{q\times m}$ and $B \in R^{n\times m}$, we solve the Penrose regression problem by minimizing the function

$$E(Q) = \frac{1}{2}\langle AQC - B, AQC - B\rangle \tag{31}$$

over the feasible set $Q \in \mathcal{O}(p,q)$. Without repeating the details, we directly present results as follows.

The gradient $\nabla E(Q)$ of $E(Q)$ with respect to the Frobenius inner product should be interpreted as the matrix

$$\nabla E(Q) = A^T(AQC - B)C^T. \tag{32}$$

The projected gradient $g(Q)$ of $\nabla E(Q)$ onto the tangent space $\mathcal{T}_Q\mathcal{O}(p,q)$ is given by:

$$\begin{aligned} g(Q) &= \frac{Q}{2}(Q^TA^T(AQC-B)C^T - C(AQC-B)^TAQ) \\ &\quad +(I-QQ^T)A^T(AQC-B)C^T. \end{aligned} \tag{33}$$

The steepest descent flow for $E(Q)$ in (33) is characterized by the differential equation:

$$\begin{aligned} \frac{dQ}{dt} &= \frac{Q}{2}(C(AQC-B)^TAQ - Q^TA^T(AQC-B)C^T) \\ &\quad -(I-QQ^T)A^T(AQC-B)C^T. \end{aligned} \tag{34}$$

Again, we may approximate a solution of the PRP by solving the differential equation (34) with an initial value, say,

$$Q(0) = \begin{pmatrix} I_q \\ 0 \end{pmatrix}.$$

Similar to Theorem 3.1, we also can specify the first order optimality condition.

THEOREM 4.1. *For $Q \in \mathcal{O}(p,q)$ to be a stationary point of the PRP, the following two conditions must hold simultaneously:*
(a). $C(AQC-B)^TAQ$ is symmetric, and
(b). $(I-QQ^T)A^T(AQC-B)C^T = 0$.

Upon extending the projected gradient $g(Q)$ in (33) to the function $G(Z)$ defined over the entire space $R^{p\times q}$:

$$\begin{aligned} G(Z) &:= \frac{Z}{2}(Z^TA^TAZCC^T - Z^TA^TBC^T - CC^TZ^TA^TAZ + CB^TAZ) \\ &\quad +(I-ZZ^T)A^T(AZC-B)C^T, \end{aligned} \tag{35}$$

we calculate the action of the Fréchet derivative of G at Z on a general H:

$$
\begin{aligned}
G'(Z)H \;:=\;& \frac{H}{2}(Z^TA^TAZCC^T - Z^TA^TBC^T - CC^TZ^TA^TAZ + CB^TAZ)\\
&+\frac{Z}{2}(H^TA^TAZCC^T + Z^TA^TAHCC^T - H^TA^TBC^T\\
&-CC^TH^TA^TAZ - CC^TZ^TA^TAH + CB^TAH)\\
&-(HZ^T + ZH^T)A^T(AZC - B)C^T + (I - ZZ^T)A^TAHCC^T. \qquad (36)
\end{aligned}
$$

We then consider the action of G' at a stationary point $Z = Q \in \mathcal{O}(p,q)$ on any tangent vector $H = QK + (I - QQ^T)W$ for arbitrary $K \in \mathcal{S}(q)^{\perp}$ and $W \in R^{p\times q}$. This action produces exactly the same action of the projected Hessian for the PRP. For the record, we present the final results of our calculation in the following. The details can be filled in by the same arguments as for the OPP.

We claim that

$$
\begin{aligned}
\langle G'(Q)QK, QK\rangle &= \left\langle Q^TA^TAQKCC^T - Q^TA^T(AQC-B)C^TK, K\right\rangle, \qquad (37)\\
\left\langle G'(Q)QK, (I-QQ^T)W\right\rangle &= \left\langle A^TAQKCC^T, (I-QQ^T)W\right\rangle. \qquad (38)
\end{aligned}
$$

We also have

$$
\begin{aligned}
&\left\langle G'(Q)(I-QQ^T)W, (I-QQ^T)W\right\rangle\\
=\;& \left\langle -WQ^TA^T(AQC-B)C^T + A^TA(I-QQ^T)WCC^T, (I-QQ^T)W\right\rangle\\
=\;& \left\langle A^TA(I-QQ^T)WC, (I-QQ^T)WC\right\rangle\\
&-\left\langle (I-QQ^T)WQ^TA^T(AQC-B)C^T, (I-QQ^T)W\right\rangle, \qquad (39)
\end{aligned}
$$

and finally

$$
\left\langle G'(Q)(I-QQ^T)W, QK\right\rangle = \left\langle A^TAQKCC^T, (I-QQ^T)W\right\rangle. \qquad (40)
$$

Similar to Theorem 3.2, we therefore have the following second-order optimality condition

THEOREM 4.2. *At a stationary point $Q \in \mathcal{O}(p,q)$ satisfying Theorem 4.1, a second-order necessary condition for Q to be a minimizer of the PRP is that the inequality*

$$
\begin{aligned}
&\left\langle Q^TA^TAQKCC^T - Q^TA^T(AQC-B)C^TK, K\right\rangle\\
&+2\left\langle A^TAQKCC^T, (I-QQ^T)W\right\rangle\\
&+\left\langle A^TA(I-QQ^T)WC, (I-QQ^T)WC\right\rangle\\
&-\left\langle (I-QQ^T)WQ^TA^T(AQC-B)C^T, (I-QQ^T)W\right\rangle \geq 0 \qquad (41)
\end{aligned}
$$

holds for all skew-symmetric matrices $K \in R^{q\times q}$ and arbitrary matrices $W \in R^{p\times q}$. If (41) holds with the strict inequality, then it is sufficient that Q is a local minimizer for the PRP.

We note that when $m = q$ and $C = I_q$, then the PRP is reduced to the OPP. The above theory generalizes the results in §3 and shows how the presence of C complicates the conditions.

With $W = 0$, the necessary condition in (41) becomes

$$\left\langle Q^T A^T A Q K C C^T, K \right\rangle + \left\langle C(AQC - B)^T AQ, K^2 \right\rangle \geq 0 \tag{42}$$

for all skew-symmetric matrices K in $R^{q \times q}$. Another necessary condition, when $K = 0$, is the inequality

$$\left\langle A(I - QQ^T)WC, A(I - QQ^T)WC \right\rangle \geq \left\langle (I - QQ^T)WQ^T A^T (AQC - B)C^T, (I - QQ^T)R \right\rangle. \tag{43}$$

for arbitrary matrix $W \in R^{p \times q}$.

Denote the spectral decompositions of the following matrices:

$$\begin{aligned} C(AQC - B)^T AQ &= V \Lambda V^T, \\ A^T A &= T \Phi T^T, \\ CC^T &= S \Psi S^T, \end{aligned}$$

where each of $V, S \in R^{q \times q}$ and $T \in R^{p \times p}$ is orthogonal. Note that all entries in $\Phi = \operatorname{diag}\{\phi_1, \ldots, \phi_p\}$ and $\Psi = \operatorname{diag}\{\psi_1, \ldots, \psi_q\}$ are nonnegative. Using (30), we obtain

$$\begin{aligned} \langle G'(Q)QK, QK \rangle &= \langle \Phi R \Psi, R \rangle - \left\langle V \Lambda V^T, U \Sigma^2 U^T \right\rangle \\ &= \sum_{j=1}^{p} \phi_j \left(\sum_{s=1}^{q} r_{js}^2 \psi_s \right) - \sum_{i=1}^{q} \lambda_i \left(\sum_{t=1}^{q} p_{it}^2 \sigma_t^2 \right) \end{aligned} \tag{44}$$

where $P = (p_{it}) := V^T U$ and $R = (r_{js}) := T^T QKS$. Thus the following second order optimality condition is proved.

Corollary 4.3. *A necessary condition for the stationary point $Q \in \mathcal{O}(p, q)$ to be a solution of the PRP is that the matrix $C(AQC - B)^T AQ$ be negative semi-definite and that the inequality (43) be held for arbitrary $W \in R^{p \times q}$.*

5. Numerical Experiment. In this section, we report some experiences of our experiment with the differential equation (15) applied to the OPP. We have observed similar experience of (34) applied to the PRP. To avoid repetition, results for PRP will not be presented here.

The computation is carried out by MATLAB 4.2a on an ALPHA 3000/300LX workstation. The solvers used for the initial value problem are **ode113** from the MATLAB ODE SUITE [11]. The code **ode113** is a PECE implementation of Adams-Bashforth-Moulton methods for non-stiff systems. More details of this code can be found in the document [11]. The reason for using this code is simply for convenience and illustration. Any other ODE solvers can certainly be used instead.

In our experiments, the tolerance for both absolute error and relative error is set at 10^{-12}. This criterion is used to control the accuracy in following the solution path.

The high accuracy we required here has little to do with the dynamics of the underlying vector field, and perhaps is not needed in practical data analysis application. We examine the output values at time interval of 10. The integration terminates automatically when the relative improvement of $F(Q)$ between two consecutive output points is less than 10^{-10} indicating local minimizer has been found. So as to fit the data comfortably in the running text, we report only the case $n = 5$ and display all numbers with five digits. All codes used in this experiment are available upon request.

One important feature of (15) (and similarly of (34)) is that the resulting $Q(t)$ should automatically stay on the manifold $\mathcal{O}(p,q)$. In numerical calculation, however, round-off errors and truncations errors easily throw the computed $Q(t)$ off the manifold of constraint. To remedy this problem, we adopt an additional non-linear projection scheme suggested by Gear [5]: Suppose Q is an approximate solution to (15) satisfying

$$Q^T Q = I + O(h^r)$$

where r represents the order of the numerical method. Let $Q = \tilde{Q}R$ be the unique QR decomposition of Q with $\operatorname{diag}(R) > 0$. Then

$$\tilde{Q} = Q + O(h^r) \tag{45}$$

and $\tilde{Q} \in \mathcal{O}(p,q)$. The condition $\operatorname{diag}(R) > 0$ is important to ensure the transition of $Q(t)$ is smooth in t. In our implementation, the matrix Q on the right-hand side of (15) is replaced by the corresponding $\tilde{Q}$.

Example 1. In practice, the data in A and B often represent two different ordinations of the same samples or populations. Based on this idea, we produce the test data for this experiment by first using the random number generator **rand** in MATLAB to create the matrix

$$A = \begin{bmatrix} 0.2190 & 0.3835 & 0.5297 & 0.4175 \\ 0.0470 & 0.5194 & 0.6711 & 0.6868 \\ 0.6789 & 0.8310 & 0.0077 & 0.5890 \\ 0.6793 & 0.0346 & 0.3834 & 0.9304 \\ 0.9347 & 0.0535 & 0.0668 & 0.8462 \end{bmatrix}.$$

We then use the command **randperm** in MATLAB to generate a random permutation

$$Q_0 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

Finally, we define $B := AQ_0$ so that the underlying OPP, though may have many local solutions due to its non-linearity, has exactly one global solution Q_0 at which $F(Q_0) = 0$. We emphasize here that the data obtained this way are not realistic because in practice B can rarely be a simple permutation of columns of A. In fact, it is precisely for the reason that it is often difficult to determine the relationship between B and A that an

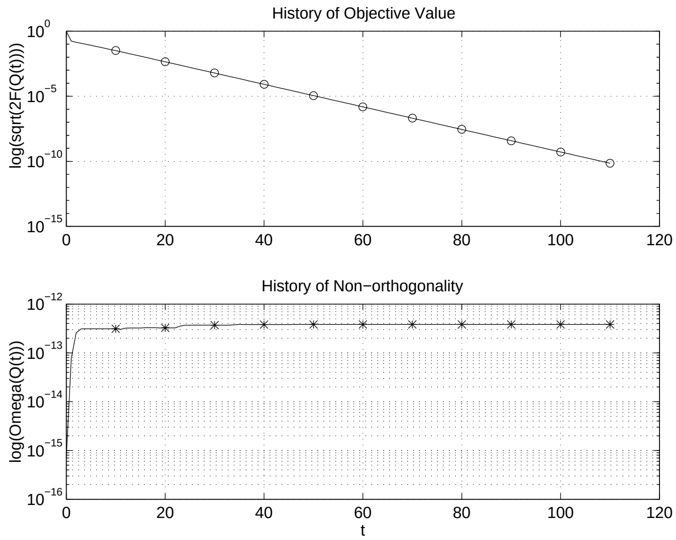

FIG. 1. *A semi-log plot of $F(Q(t))$ and $\Omega(Q(t))$ for Example 1(a).*

orthonormal Procrustes analysis is needed [7]. We present the following examples just to illustrate how the descent flow behaviors.

(a). Obviously, the initial value determines where our descent flow will converges to. For instance, suppose we start with the matrix

$$Q(0) = \begin{bmatrix} 0.2618 & 0.9198 & -0.2333 \\ 0.8912 & -0.3467 & -0.2333 \\ 0.2618 & 0.1301 & 0.9399 \\ 0.2618 & 0.1301 & 0.0876 \end{bmatrix} \tag{46}$$

that represent a non-trivial perturbation of Q_0. Figure 1 records the history of the changes of the objective value $2F(Q(t)) = \|AQ(t) - B\|$ where $Q(t)$ is determined by integrating the differential equation (15). Clearly, the global solution is obtained in this case.

Also recorded in Figure 1 is the history of the function

$$\Omega(Q(t)) := \|I_q - Q(t)^T Q(t)\| \tag{47}$$

that measures the deviation of $Q(t)$ from the manifold of constraint $\mathcal{O}(p, q)$. It is seen that $Q(t)$ is well kept within the local tolerance.

(b). Suppose the initial value $Q(0)$ is taken to be that suggested in (16). Then we can only reach a local minimizer

$$Q_* = \begin{bmatrix} 0.0094 & 0.3811 & -0.5768 \\ 0.9999 & 0.0094 & 0.0088 \\ 0.0088 & -0.5768 & 0.4625 \\ -0.0110 & 0.7225 & 0.6733 \end{bmatrix}$$

with objective value $\|AQ_* - B\| \approx 0.2234$. The test results are presented in Figure 2. It is interesting to note that both conditions in Theorem 3.2 are satisfied.

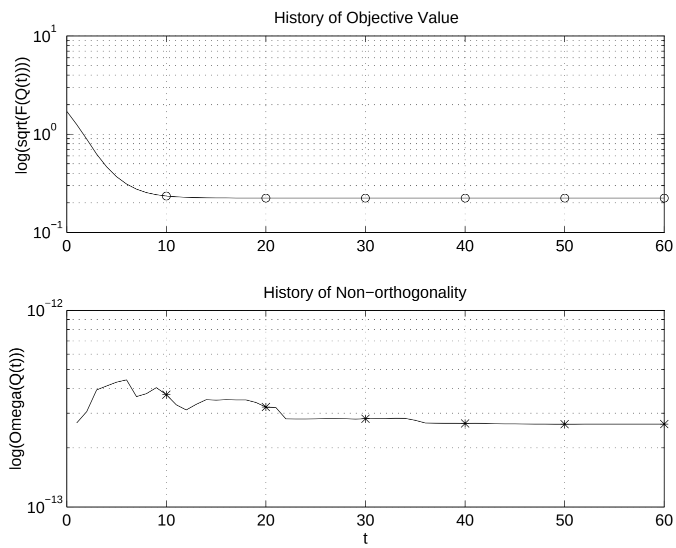

FIG. 2. *A semi-log plot of $F(Q(t))$ and $\Omega(Q(t))$ for Example 1(b).*

Example 2. Suppose $B = AQ_0 + \frac{1}{2}\Delta$ where

$$\Delta = \begin{bmatrix} 0.5383 & 0.9503 & 0.6004 \\ -0.6168 & 0.3468 & 1.0047 \\ -1.2161 & -0.9547 & -0.3608 \\ -0.8900 & -0.7598 & -0.6719 \\ -1.9832 & 0.3192 & -0.6037 \end{bmatrix}$$

represents a random perturbation from normal distribution $N(0,1)$. With $Q = Q_0$, this noise has the magnitude $\|AQ_0 - B\| = 1.7152$. But by following (15) with $Q(0)$ given by (46), we obtain

$$Q_\# = \begin{bmatrix} -0.3582 & 0.8064 & 0.4669 \\ 0.8186 & 0.2608 & 0.2344 \\ 0.0800 & -0.4655 & 0.8275 \\ 0.4418 & 0.2549 & -0.2056 \end{bmatrix}$$

with $\|AQ_\# - B\|$ being reduced to 1.5132. The test results are recorded in Figure 3. Again, we can only report that a local minimizer is found.

6. Conclusion. We have reconsidered the orthonormal Procrustes problem and the Penrose regression problem. By using the projected gradient idea, we are able to completely characterize the first order and the second order optimality conditions for a local minimizer. The results extend what is already known in the literature, and provides a new numerical method for solving these problems.

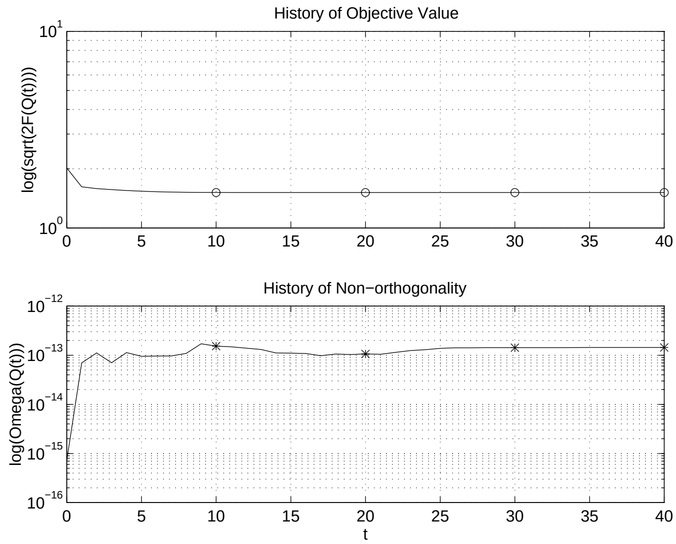

FIG. 3. *A semi-log plot of $F(Q(t))$ and $\Omega(Q(t))$ for Example 2.*

REFERENCES

[1] M. T. Chu and K. R. Driessel, The projected gradient method for least squares matrix approximations with spectral constraints, SIAM J. Numer. Anal., 27(1990), 1050-1060.

[2] M. T. Chu, A list of matrix flows with applications, Fields Institute Communications, 3(1994), 87-97.

[3] A. Edelman, T. Arias and S. T. Smith, Conjugate gradient on the Stiefel and Grassmann manifolds, preprint, Dept. of Math., MIT, 1995.

[4] P. E. Gill, W. Murray, and M. H. Wright, *Practical Optimization*, Academic Press, Florida, 1981.

[5] C. W. Gear, Maintaining solution invariants in the numerical solution of ODEs, SIAM J. Sci. Stat. Comput., 7(1986), 734-743.

[6] G. H. Golub and C. F. Van Loan, *Matrix Computation*, 2nd ed., The Johns Hopkins University Press, Baltmore, 1991.

[7] J. C. Gower, Multivariate analysis: Ordination, multidimensional scaling and allied topics. In *Handbook of Applicable Mathematics*, Vol. VI: Statistics, Part B. ed. Emlyn Lloyd, New York: John Wiley & Sons, 1984.

[8] B. Green, The orthogonal approximation of an oblique structure in factor analysis, Psychometrika, 17(1952), 429-440.

[9] M. A. Koschat and D. F. Swayne, A weighted Procrustes criterion, Psychometrika, 56(1991), 229-239.

[10] A. Mooijaart and J.J.F Commandeur, A general solution of the weighted orthonormal Procrustes problem, Psychometrika, 55(1990), 657-663.

[11] L. F. Shampine and M. W. Reichelt, The MATLAB ODE suite, preprint, 1995.

[12] E. Stiefel, Richtungsfelder und fernparallelismus in n-dimensionalel manning faltigkeiten, Commentarii Mathematici Helvetici, 8(1935-1936), 305-353.

[13] J. M. F. ten Berge and D. L. Knol, Orthogonal rotations to maximal agreement for two or more matrices of different column orders, Psychometrika, 49(1984), 49-55.