Journal of Computational Mathematics, Vol.25, No.6, 2007, 661–671.

CONDITIONS FOR OPTIMAL SOLUTIONS OF UNBALANCED PROCRUSTES PROBLEM ON STIEFEL MANIFOLD*

Zhenyue Zhang, Yuyang Qiu and Keqin Du
(*Department of Mathematics, Zhejiang University, Hangzhou 310027, China*
Email: zyzhang@zju.edu.cn, yuyangqiu@gmail.com, du_keqin1978@hotmail.com)

Abstract

We consider the unbalanced Procrustes problem with an orthonormal constraint on solutions: given matrices $A \in \mathcal{R}^{n\times n}$ and $B \in \mathcal{R}^{n\times k}$, $n > k$, minimize the residual $\|AQ - B\|_F$ over the Stiefel manifold of orthonormal matrices. Theoretical analysis on necessary conditions and sufficient conditions for optimal solutions of the unbalanced Procrustes problem is given.

Mathematics subject classification: 65F05, 15A06.
Key words: Procrustes problem, Stiefel manifold, Necessary condition, Sufficient condition.

1. Introduction

Given two matrices $A \in \mathcal{R}^{m\times n}$ and $B \in \mathcal{R}^{m\times k}$ with $n > k$, we consider the orthonormal Procrustes problem: Find an orthonormal matrix $Q \in \mathcal{R}^{n\times k}$ that solves the following constrained optimization problem:

$$\min_{Q^TQ=I} \|AQ - B\|_F^2, \tag{1.1}$$

where $\|\cdot\|_F$ denote the Frobenius norm. The set of orthonormal matrices in $\mathcal{R}^{n\times k}$ defines the orthonormal Stiefel manifold

$$\mathcal{S} = \{Q \in \mathcal{R}^{n\times k} : Q^TQ = I\}. \tag{1.2}$$

In general, $m \geq n$. If $m \gg n$, the size of the problem can be reduced by QR decomposition with no difficulties. Therefore, without loss of generality, we assume that the matrix A is square with order n, i.e., write (1.1) as

$$\min_{Q^TQ=I} \|AQ - B\|_F^2, \quad A \in \mathcal{R}^{n\times n}, \quad B \in \mathcal{R}^{n\times k}, \quad n \geq k. \tag{1.3}$$

We refer to (1.3) as the *balanced* Procrustes problem if $k = n$, and the *unbalanced* Procrustes problem when $k < n$.

The balanced problem is simple and has been solved analytically [4, 8]; solutions are given by the singular value decomposition (SVD) or pole decomposition of A^TB. However, the unbalanced Procrustes problem seems to be quite difficult. First, if A is rank deficient in column, i.e., $r = \text{rank}(A) < n$, then by SVD $A = U_r\Sigma_rV_r^T$ of A,

$$\|AQ - B\|_F^2 = \|\Sigma_rV_r^TQ - U_r^TB\|_F^2 + \|(U_r^\perp)^TB\|_F^2,$$

* Received July 26, 2005; final revised July 7, 2006; accepted August 31, 2006.

where $U_r^{\perp}$ denotes the orthogonal complement of U_r. Hence the problem (1.3) is equivalent to $\min_{Q^TQ=I_k} \|\Sigma_r V_r^T Q - U_r^T B\|_F$. Though $\|\Sigma_r V_r^T Q - U_r^T B\|_F = \|\Sigma_r X - U_r^T B\|_F$ with $X = V_r^T Q$ and $\|X\| \leq 1$, the optimization problem mentioned above is not equivalent to the constrained LS problem

$$\min_{\|X\|\leq 1} \|\Sigma_r X - U_r^T B\|_F.$$

Second, necessary and sufficient conditions for a global minimum of the unbalanced problem are still not clear, though some necessary conditions and sufficient conditions for local and/or global optimal solutions have been reported [3]. Third, except the special case when $k = 1$ for which a quadratic convergent method is proposed and careful analysis for its quadratic convergence is also given in [10], no sufficient analysis is reported for the convergence of existing iterative algorithms for solving the unbalanced problem when $k > 1$, partially due to lack of sufficient theoretical analysis for optimal solutions. Indeed, these iterative algorithms may be divergent or/and not efficient when the problem scale is large or A is ill conditioned.

The purpose of this paper is to show a deep understanding to the unbalanced Procrustes problem. We are interested in conditions of optimal solutions for the unbalanced problem. An analysis for local or global optimal solutions is given, which simplifies the discussions given in [3]. Based upon the analysis presented in this paper, a successive projection (SP) method for solving the unbalanced Procrustes problem will be proposed in a separate paper, together with a careful analysis for the convergence of the successive projection method and reports of numerical experiments.

The rest of this paper is arranged as follows. In Section 2, we review the structures of optimal solutions of the balanced problem. A discussion of necessary conditions for global optimization solutions of the unbalanced Procrustes problem is given in Section 3. In Section 4, we present some sufficient conditions of a local and/or global minimum of the unbalanced problem.

Notations. Given an orthonormal matrix $Q \in \mathcal{R}^{n\times k}$, we call $H \in \mathcal{R}^{n\times(n-k)}$ an orthogonal complement of Q if $[Q, H]$ is a (square) orthogonal matrix. The spectral norm of a vector or a matrix is simply denoted as $\|\cdot\|$, while $\|\cdot\|_F$ denotes the Frobenius norm of matrices. As used in general, I denotes an identity matrix with certain matrix size.

2. Structure of Solutions to Balanced Procrustes Problem

The balanced Procrustes problem, i.e., $k = n$, is relatively simple; it can be solved by the SVD of the matrix A^TB. Here we cite a theorem that illustrates the structure of solutions of the balanced Procrustes problem.

Theorem 2.1. [4, p.695] *Let*

$$A^TB = U \begin{bmatrix} \Sigma_1 & 0 \\ 0 & 0 \end{bmatrix} V^T$$

be the singular value decomposition of A^TB, where $\Sigma_1 = \mathrm{diag}(\sigma_1, \cdots, \sigma_r)$, and $r = \mathrm{rank}(A^TB)$. Then all solutions of the balanced Procrustes problem (1.3) *can be formulated as*

$$Q = U \begin{bmatrix} I_r & 0 \\ 0 & T \end{bmatrix} V^T, \tag{2.1}$$