Chapter 12

ON MINIMIZATION OF SUMS OF HETEROGENEOUS QUADRATIC FUNCTIONS ON STIEFEL MANIFOLDS.

T. Rapcsák[1]

Senior Researcher, Computer and Automation Institute, Hungarian Academy of Sciences, Budapest, Hungary.

[1]This research was supported in part by the Hungarian National Research Foundation, Grant Nos. OTKA-T016413 and OTKA-T017241

Abstract The minimization of functions $\sum_{i=1}^{k} \frac{1}{2}\mathbf{x}_i^T \mathbf{A}_i \mathbf{x}_i$ is studied under the constraint that vectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k \in R^n$ form an orthonormal system and $\mathbf{A}_1, \ldots, \mathbf{A}_k$, $(k \leq n)$ are given symmetric $n \times n$ matrices. The set of feasible points determines a differentiable manifold introduced by Stiefel in 1935. The optimality conditions are obtained by the global Lagrange multiplier rule, and variable metric methods along geodesics are suggested as solving methods for which a global convergence theorem is proved. Such problems arise in various situations in multivariate statistical analysis.

Keywords: Quadratic optimization, quadratic equality constraints, Stiefel manifolds.

1. Introduction

In 1935, Stiefel introduced a differentiable manifold consisting of all the orthonormal vector systems $\mathbf{x_1}, \mathbf{x_2}, \ldots, \mathbf{x_k} \in R^n$, where R^n is the n-dimensional Euclidean space and $k \leq n$. James (1976) dealt with

A. Migdalas et al. (eds.), From Local to Global Optimization, 277–290.

the topology of Stiefel manifolds. Bolla, Michaletzky, Tusnády and Ziermann (1998) analysed the maximization of sums of heterogeneous quadratic functions on Stiefel manifolds assuming the positive definiteness of the quadratic forms. They formulated the optimality conditions, studied some structural properties of the problem, proposed an iterative algorithm and proved its convergence to a critical point. Statistical applications of these global optimization models (e.g., dynamical factor analysis) are quoted. Their results are mainly based on the matrix theory and the given representation of Stiefel manifolds.

Here, the aim is to reformulate the above optimization problem as a smooth nonlinear optimization one to obtain the first-order and second-order necessary and sufficient optimality conditions based on classical and new results in optimization theory and to show how it is possible to use a generalization of the classical nonlinear optimization methods for solving them. A new theoretical approach elaborated for smooth nonlinear optimization is suggested based on the Riemannian geometry and the global Lagrange multiplier rule [7,9].

2. Optimization problem

Consider the following optimization problem:

$$\begin{aligned} &\min \sum_{i=1}^{k} \frac{1}{2} \mathbf{x}_i^T \mathbf{A}_i \mathbf{x}_i \\ &\mathbf{x}_i^T \mathbf{x}_j = \delta_{ij}, \ 1 \le i,j \le k \le n, \\ &\mathbf{x}_i \in R^n, \ i = 1,\dots,k, \ n \ge 2, \end{aligned} \tag{2.1}$$

where $\mathbf{A}_i$, $i = 1,\dots,k$, are given symmetric $n \times n$ matrices and δ_{ij} is the Kronecker's delta. Since the constraint set is compact and the objective function is continuous, optimization problem (2.1) has a global minimum point. Obviously, the lower bound of the global minimum is $\sum_{i=1}^{k} \lambda_{i,min}$, where $\lambda_{i,min}$ denote the minimal eigenvalues of the matrices A_i, $i = 1,\dots,k$, but in general, this value is not attained because the minimal eigenvectors of the matrices A_i, $i = 1,\dots,k$, are pairwise orthogonal in some cases only. The constraints of problem (2.1) can be written as

$$\mathbf{x}_i^T \mathbf{x}_i = 1, \qquad i = 1,\dots,k, \tag{2.2a}$$

$$\mathbf{x}_i^T \mathbf{x}_j = 0, \qquad i,j = 1,\dots,k, \quad i \ne j, \tag{2.2b}$$

$$\mathbf{x}_i \in R^n, \qquad i = 1,\dots,k, \quad n \ge 2.$$

It follows from orthogonality that $n \geq k \geq 2$. Equalities (2.2a) and (2.2b), and equalities (2.2b) determine a compact and a noncompact differentiable manifold, respectively. So, optimization problem (2.1) consists of the minimization of the sums of heterogeneous quadratic functions over a special differentiable manifold, a Stiefel manifold.

First, optimization problem (2.1) is reformulated as a smooth nonlinear optimization one in order to use its machinery. Let us introduce the following notations:

$$\mathbf{x} = (\mathbf{x}_1, \ldots, \mathbf{x}_k) \in R^{kn}, \qquad J = \{(i,j) |\ i,j = 1, \ldots, k,\ i \neq j\},$$

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \mathbf{A}_k \end{pmatrix}, \quad \mathbf{C}_l = \begin{pmatrix} 0 & \cdots & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & \cdots & I_n & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \cdots & 0 \end{pmatrix}, \quad l = 1, \ldots, k,$$

$$\mathbf{C}_{ij} = \begin{pmatrix} 0 & \cdots & 0 & \cdots & 0 & \cdots \\ \vdots & \ddots & \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 0 & \cdots & I_n & \cdots \\ \vdots & \ddots & \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & I_n & \cdots & 0 & \cdots \\ 0 & \cdots & 0 & \cdots & 0 & \cdots \end{pmatrix}, \quad (i,j) \in J,$$

where $\mathbf{A}$, $\mathbf{C}_l$, $\mathbf{C}_{ij}$ are $kn \times kn$ blockdiagonal matrices, I_n is the identity matrix in R^n, $\mathbf{C}_l$ and $\mathbf{C}_{ij}$ contain I_n in the l-th diagonal block and in the (i,j) as well as (j,i) blocks, respectively. The $kn \times kn$ symmetric matrices $\mathbf{C}_{ij}$ are defined for all the pairs of different indices belonging to J, given by the $k(k-1)/2$ combinations of the indices $1, \ldots, k$.

In the case of a compact Stiefel manifold, problem (2.1) is equivalent to

$$\begin{aligned} &\min \frac{1}{2}\mathbf{x}^T\mathbf{A}\mathbf{x} \\ h_l(\mathbf{x}) = \frac{1}{2}\mathbf{x}^T\mathbf{C}_l\mathbf{x} - \frac{1}{2} = 0, \qquad &l = 1, \ldots, k, \\ h_{ij}(\mathbf{x}) = \frac{1}{2}\mathbf{x}^T\mathbf{C}_{ij}\mathbf{x} = 0, \qquad &(i,j) \in J, \\ \mathbf{x} \in R^{kn}, \qquad &n \geq 2. \end{aligned} \tag{2.3}$$

Problem (2.3) is of the basic equality constrained problems in smooth optimization which type of constraints is studied in most of the classical literature (e.g., Luenberger, 1973). The difficulty in the solution of problems (2.3) originated mostly from the intersections of the quadratic equality constraints, which results in the fact that the feasible region is a nonconvex and possibly disconnected subset of the hypersphere $\sum_{i=1}^{k}\sum_{j=1}^{n} x_{ij}^2 = k$ in R^{kn}. Adams (1962) gave necessary and sufficient conditions depending on n and k for ensuring that every point of the whole $(n-1)$-dimensional hypersphere should admit an orthonormal k-frame.

The next step is the investigation of the structure of the constraint set. Let

$$
\begin{aligned}
f(\mathbf{x}) &= \frac{1}{2}\mathbf{x}^T\mathbf{A}\mathbf{x}, \qquad \mathbf{x} \in R^{kn},\\
M_{n,k} &= \{\mathbf{x} \in R^{kn} \mid h_l(\mathbf{x}) = 0,\ l = 1,\dots,k; \quad h_{ij}(\mathbf{x}) = 0, \quad (i,j) \in J\}, \qquad (2.4)\\
TM_{n,k}(\mathbf{x}) &= \{\mathbf{v} \in R^{kn} \mid \nabla h_l(\mathbf{x})\mathbf{v} = 0,\ l = 1,\dots,k;\\
&\qquad \nabla h_{ij}(\mathbf{x})\mathbf{v} = 0, \quad (i,j) \in J\}, \quad \mathbf{x} \in M_{n,k},
\end{aligned}
$$

where the symbol ∇ denotes the gradient vector of a function. Though the next statement is known in the literature, the optimization theoretical reformulation of the original problem makes a simple proof possible.

Theorem 2.1 *The set $M_{n,k}$ is a C^∞ differentiable manifold with the dimension $kn - \frac{k(k+1)}{2}$.*

Proof It is sufficient to show that the gradients of the equality constraints given by the row vectors

$$
\begin{aligned}
\mathbf{x}^T\mathbf{C}_l &= (\dots,\mathbf{x}_i,\dots), \qquad l = 1,\dots,k,\\
\mathbf{x}^T\mathbf{C}_{ij} &= (\dots,\mathbf{x}_i,\dots,\mathbf{x}_j,\dots), \qquad (i,j) \in J,
\end{aligned} \qquad (2.5)
$$

are linearly independent on $M_{n,k}$. In problem (2.3), the number of equalities is equal to $k + k(k-1)/2 = \frac{k(k+1)}{2} < kn$. As

$$
\begin{aligned}
\mathbf{x}^T\mathbf{C}_l\mathbf{C}_l\mathbf{x} &= 1, && l = 1,\dots,k, && \mathbf{x} \in M_{n,k},\\
\mathbf{x}^T\mathbf{C}_{ij}\mathbf{C}_{ij}\mathbf{x} &= 2, && (i,j) \in J, && \mathbf{x} \in M_{n,k},\\
\mathbf{x}^T\mathbf{C}_i\mathbf{C}_j\mathbf{x} &= 0, && (i,j) \in J, && \mathbf{x} \in M_{n,k}, \qquad (2.6)\\
\mathbf{x}^T\mathbf{C}_l\mathbf{C}_{ij}\mathbf{x} &= 0, && l = 1,\dots,k; \ (i,j) \in J, && \mathbf{x} \in M_{n,k},\\
\mathbf{x}^T\mathbf{C}_{ij}\mathbf{C}_{kl}\mathbf{x} &= 0, && (i,j),(k,l) \in J,\ (i,j) \neq (k,l), && \mathbf{x} \in M_{n,k},
\end{aligned}
$$

these relations mean that the gradients are different from zero, and any two of them are orthogonal, from which the statement follows. ∎

3. Optimality conditions

In this part, the first-order and second-order necessary and sufficient optimality conditions of problem (2.3) are stated. The gradients of the equality constraints are linearly independent, the feasible set is a differentiable manifold, moreover, the feasible set $M_{n,k}$ is a Riemannian smooth $(kn - \frac{k(k+1)}{2})$-manifold in R^{kn} where the Riemannian metric on $M_{n,k}$ is induced by the Euclidean one of R^{kn}. Now, the global Lagrange multiplier rule developed for the case of equality constraints (Rapcsák, 1991, 1997) is used to obtain the first-order and second-order optimality conditions of problem (2.3). The geometric background of the rule can be found in these references as well.

Let the Lagrangian function be

$$L(\mathbf{x}, \mu(\mathbf{x}), \eta(\mathbf{x})) =$$

$$f(\mathbf{x}) - \sum_{l=1}^{k} \mu_l(\mathbf{x}) h_l(\mathbf{x}) - \sum_{(i,j)\in J} \eta_{ij}(\mathbf{x}) h_{ij}(\mathbf{x}), \qquad \mathbf{x} \in M_{n,k}, \tag{3.1}$$

the geodesic gradient vector of the Lagrangian function,

$$\nabla^g L(\mathbf{x}, \mu(\mathbf{x}), \eta(\mathbf{x})) =$$

$$\nabla f(\mathbf{x}) - \sum_{l=1}^{k} \mu_l(\mathbf{x}) \nabla h_l(\mathbf{x}) - \sum_{(i,j)\in J} \eta_{ij}(\mathbf{x}) \nabla h_{ij}(\mathbf{x}), \qquad \mathbf{x} \in M_{n,k}, \tag{3.2}$$

and the geodesic Hessian matrix of the Lagrangian function

$$H^g L(\mathbf{x}, \mu(\mathbf{x}), \eta(\mathbf{x})) =$$

$$Hf(\mathbf{x}) - \sum_{l=1}^{k} \mu_l(\mathbf{x}) Hh_l(\mathbf{x}) - \sum_{(i,j)\in J} \eta_{ij}(\mathbf{x}) Hh_{ij}(\mathbf{x}), \qquad \mathbf{x} \in M_{n,k}, \tag{3.3}$$

where the symbol H denotes the Hessian matrix of a function, and

$$\mu(\mathbf{x})^T = \nabla f(\mathbf{x}) J\mathbf{h}^T(\mathbf{x})[J\mathbf{h}(\mathbf{x}) J\mathbf{h}(\mathbf{x})^T]^{-1}, \qquad \mathbf{x} \in M_{n,k}, \tag{3.4}$$

$$J\mathbf{h}(\mathbf{x}) = \begin{pmatrix} \nabla h_1(\mathbf{x}) \\ \vdots \\ \nabla h_l(\mathbf{x}) \end{pmatrix}, \qquad \mathbf{x} \in M_{n,k},$$

$$\eta(\mathbf{x})^T = \nabla f(\mathbf{x}) J\hat{\mathbf{h}}^T(\mathbf{x})[J\hat{\mathbf{h}}(\mathbf{x}) J\hat{\mathbf{h}}(\mathbf{x})^T]^{-1}, \qquad \mathbf{x} \in M_{n,k}, \tag{3.5}$$

$$J\hat{\mathbf{h}}(\mathbf{x}) = \begin{pmatrix} \nabla h_{12}(\mathbf{x}) \\ \vdots \\ \nabla h_{ij}(\mathbf{x}) \\ \vdots \end{pmatrix}, \qquad \mathbf{x} \in M_{n,k}.$$

The global Lagrange multiplier rule based on Theorem 15.2.1 and Corollary 15.2.1 in [9, p.276] is formulated for the case of problem (2.3) in the following statement. First, a definition of geodesic convex sets is recalled where the geodesic is used in the classical meaning. If M is a Riemannian C^2 manifold, then a set $\mathcal{C} \subseteq \mathcal{M}$ is geodesic convex if any two points of $\mathcal{C}$ are joined by a geodesic belonging to $\mathcal{C}$, moreover, a singleton is geodesic convex.

Theorem 3.1 *If the point $\mathbf{x}_0 \in M_{n,k}$ is a (strict) local minimum of problem (2.3), then*

$$\begin{aligned} \nabla^g L(\mathbf{x}_0, \mu(\mathbf{x}_0), \eta(\mathbf{x}_0)) &= 0, \\ \mathbf{v}^T H^g L(\mathbf{x}_0, \mu(\mathbf{x}_0), \eta(\mathbf{x}_0))\mathbf{v} \geq (>)0, &\qquad \mathbf{v} \in TM_{n,k}(\mathbf{x}_0). \end{aligned} \tag{3.6}$$

If $\mathcal{C} \subseteq M_{n,k}$ is an open geodesic convex set and there exists a point $\mathbf{x}_0 \in \mathcal{C}$ such that

$$\begin{aligned} &\nabla^g L(\mathbf{x}_0, \mu(\mathbf{x}_0), \eta(\mathbf{x}_0)) = 0, \\ \mathbf{v}^T H^g L(\mathbf{x}, \mu(\mathbf{x}), \eta(\mathbf{x}))\mathbf{v} \geq (>)0, \qquad &\mathbf{v} \in TM_{n,k}(\mathbf{x}), \qquad \mathbf{x} \in \mathcal{C}, \end{aligned} \tag{3.7}$$

then the point $\mathbf{x}_0$ is a (strict) global minimum of the function f on $\mathcal{C}$.

Let

$$S(\mathbf{x}) = \tag{3.8}$$

$$\begin{pmatrix} \mathbf{x}_1^T \mathbf{A}_1 \mathbf{x}_1 I_n & \frac{1}{2}(\mathbf{x}_2^T \mathbf{A}_2 \mathbf{x}_1 + \mathbf{x}_1^T \mathbf{A}_1 \mathbf{x}_2) I_n & \cdots & \frac{1}{2}(\mathbf{x}_k^T \mathbf{A}_k \mathbf{x}_1 + \mathbf{x}_1^T \mathbf{A}_1 \mathbf{x}_k) I_n \\ \frac{1}{2}(\mathbf{x}_1^T \mathbf{A}_1 \mathbf{x}_2 + \mathbf{x}_2^T \mathbf{A}_2 \mathbf{x}_1) I_n & \mathbf{x}_2^T \mathbf{A}_2 \mathbf{x}_2 I_n & \cdots & \frac{1}{2}(\mathbf{x}_k^T \mathbf{A}_k \mathbf{x}_2 + \mathbf{x}_2^T \mathbf{A}_2 \mathbf{x}_k) I_n \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1}{2}(\mathbf{x}_1^T \mathbf{A}_1 \mathbf{x}_k + \mathbf{x}_k^T \mathbf{A}_k \mathbf{x}_1) I_n & \frac{1}{2}(\mathbf{x}_2^T \mathbf{A}_2 \mathbf{x}_k + \mathbf{x}_k^T \mathbf{A}_k \mathbf{x}_2) I_n & \cdots & \mathbf{x}_k^T \mathbf{A}_k \mathbf{x}_k I_n \end{pmatrix}$$

where $\mathbf{x} = (\mathbf{x}_1, \ldots, \mathbf{x}_k) \in M_{n,k}$.

By Theorem 3.1, the first-order optimality condition of problem (2.3) at a point $\mathbf{x}_0$ is

$$\mathbf{A}\mathbf{x}_0 - \sum_{l=1}^{k} (\mathbf{x}_0^T \mathbf{A} \mathbf{C}_l \mathbf{x}_0) \mathbf{C}_l \mathbf{x}_0 - \sum_{(i,j) \in J} \frac{1}{2} (\mathbf{x}_0^T \mathbf{A} \mathbf{C}_{ij} \mathbf{x}_0) C_{ij} \mathbf{x}_0 = 0, \qquad \text{i.e.,}$$

$$\mathbf{A}\mathbf{x}_0 = S(\mathbf{x}_0)\mathbf{x}_0, \tag{3.9}$$

and the second-order necessary condition at a point $\mathbf{x}_0$ is equivalent to the positive semidefiniteness of the matrix

$$\left(\mathbf{A} - \sum_{l=1}^{k}(\mathbf{x}_0^T\mathbf{A}\mathbf{C}_l\mathbf{x}_0)\mathbf{C}_l - \sum_{(i,j)\in J}\frac{1}{2}(\mathbf{x}_0^T\mathbf{A}\mathbf{C}_{ij}\mathbf{x}_0)\mathbf{C}_{ij}\right)_{|TM_{\mathbf{x}_0}} = (\mathbf{A} - S(\mathbf{x}_0))_{|TM_{\mathbf{x}_0}}, \tag{3.10}$$

where the symbol $|TM_{\mathbf{x}}$ denotes the restriction to the tangent space $TM_{n,k}(\mathbf{x})$ given in the form of

$$\begin{aligned} TM_{n,k}(\mathbf{x}) = \{\mathbf{v} \in R^{kn} |\ & \mathbf{x}_l^T\mathbf{v}_l = 0, \quad l = 1,\ldots,k; \\ & \mathbf{x}_i^T\mathbf{v}_j + \mathbf{x}_j^T\mathbf{v}_i = 0, \quad (i,j) \in J\}, \mathbf{x} \in M_{n,k}. \end{aligned} \tag{3.11}$$

It is obvious that the tangent spaces $TM_{n,k}(\mathbf{x})$, $\mathbf{x} \in M_{n,k}$, take part of the corresponding tangent spaces of the hypersphere $\sum_{i=1}^{k}\sum_{j=1}^{n} x_{ij}^2 = k$ in R^{kn}. As problem (2.3) has at least one stationary point and $\det(\mathbf{A} - S(\mathbf{x}_0)) = 0$, thus several stationary points can be expected.

Corollary 3.1*If we consider the maximization problem*

$$\begin{aligned} & \max \sum_{i=1}^{k} \frac{1}{2}\mathbf{x}_i^T\mathbf{A}_i\mathbf{x}_i \\ & \mathbf{x}_i^T\mathbf{x}_j = \delta_{ij}, \qquad 1 \le i,j \le k \le n, \\ & \mathbf{x}_i \in R^n, \quad i = 1,\ldots,k, \qquad n \ge 2, \end{aligned} \tag{3.12}$$

then the first-order and second-order necessary optimality conditions at a point $\mathbf{x}_0$ are as follows:

$$\begin{aligned} & \mathbf{A}\mathbf{x}_0 = S(\mathbf{x}_0)\mathbf{x}_0, \\ & (\mathbf{A} - S(\mathbf{x}_0))_{|TM_{\mathbf{x}_0}} \quad \textit{is negative semidefinite}, \end{aligned} \tag{3.13}$$

Bolla, Michaletzky, Tusnády and Ziermann (1997) studied problem (3.12) assuming the positive definiteness of the matrices $\mathbf{A}_1,\ldots,\mathbf{A}_k$ and proved that the first-order and second-order optimality conditions at a point $\mathbf{x}_0$ are

$$\begin{aligned} & \mathbf{A}\mathbf{x}_0 = \hat{S}(\mathbf{x}_0)\mathbf{x}_0, \\ & \hat{S}(\mathbf{x}_0) \quad \text{is symmetric}, \\ & (\mathbf{A} - \hat{S}(\mathbf{x}_0))_{|TM_{\mathbf{x}_0}} \quad \text{is negative semidefinite}, \end{aligned} \tag{3.14}$$

$$\hat{S}(\mathbf{x}) = \begin{pmatrix} \mathbf{x}_1^T \mathbf{A}_1 \mathbf{x}_1 I_n & \mathbf{x}_1^T \mathbf{A}_1 \mathbf{x}_2 I_n & \cdots & \mathbf{x}_1^T \mathbf{A}_1 \mathbf{x}_k I_n \\ \mathbf{x}_2^T \mathbf{A}_2 \mathbf{x}_1 I_n & \mathbf{x}_2^T \mathbf{A}_2 \mathbf{x}_2 I_n & \cdots & \mathbf{x}_2^T \mathbf{A}_2 \mathbf{x}_k I_n \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{x}_k^T \mathbf{A}_k \mathbf{x}_1 I_n & \mathbf{x}_k^T \mathbf{A}_k \mathbf{x}_2 I_n & \cdots & \mathbf{x}_k^T A_k \mathbf{x}_k I_n \end{pmatrix}, \qquad (3.15)$$

where $\mathbf{x} = (\mathbf{x}_1, \dots, \mathbf{x}_k) \in M_{n,k}$.

The main difference between the first-order and second-order optimality conditions (3.13) and (3.14) is that $\hat{S}(\mathbf{x})$, $\mathbf{x} \in M_{n,k}$, are, in general, nonsymmetric matrices, which is unusual in optimization theory, while $S(\mathbf{x})$, $\mathbf{x} \in M_{n,k}$, are symmetric.

Lemma 3.1 *Optimality conditions (3.13) and (3.14) are equivalent.*

Proof If all the k blocks consisting of one n-vector each are multiplied by $\mathbf{x}_j$, $j = 1, \dots, k$, in the first-order optimality condition (3.13), then equalities

$$\mathbf{x}_j A_i \mathbf{x}_i = \mathbf{x}_j A_j \mathbf{x}_i, \qquad (i, j) \in J, \qquad (3.16)$$

are obtained at the point $\mathbf{x}_0$, from which the statement follows. ∎

Example 3.1 Let $n = 2$, $k = 2$ and

$$\mathbf{A}_1 = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}, \qquad \mathbf{A}_2 = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix},$$

$$\mathbf{x} = (1, 0, 0, 1), \qquad \mathbf{y} = (0, 1, 1, 0),$$

then $f(\mathbf{x}) = 5$, $f(\mathbf{y}) = \mathbf{3}$,

$$S(\mathbf{x}) = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}, \qquad S(\mathbf{y}) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix},$$

$$\mathbf{A} - S(\mathbf{x}) = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \qquad \mathbf{A} - S(\mathbf{y}) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

$$TM_{n,k}(\mathbf{x}) \subseteq \{\mathbf{v} \in R^4 \mid v_1 = 0,\ v_4 = 0\},$$

$$TM_{n,k}(\mathbf{y}) \subseteq \{\mathbf{v} \in \mathbf{R}^4 \mid \mathbf{v_2} = \mathbf{0},\ \mathbf{v_3} = \mathbf{0}\}.$$

It follows that $(\mathbf{A} - S(\mathbf{x}))\mathbf{x} = 0$, $\mathbf{A} - S(\mathbf{x})$ is negative definite on $TM_{n,k}(\mathbf{x})$, i.e., $\mathbf{x}$ is a strict local maximum point, $(\mathbf{A} - S(\mathbf{y}))\mathbf{y} = 0$, $\mathbf{A} - S(\mathbf{y})$ is positive definite on $TM_{n,k}(\mathbf{y})$, i.e., $\mathbf{y}$ is a strict local minimum point.

4. Variable metric methods along geodesics

A general framework for variable metric methods along geodesics was developed to find a stationary point or a local optimum of smooth nonlinear optimization problems defined on a Riemannian manifold (Gabay, 1982; Rapcsák and Thang, 1995; Rapcsák, 1997, 1998). These algorithms represent a generalization of unconstrained optimization methods and constrained optimization methods under equalities and proceed in the case of problem (2.3) as follows.

Let us consider problem (2.3). Starting from an initial feasible solution $\mathbf{x_0} \in M_{n,k} \subseteq R^{kn}$, let $\mathbf{x_i} \in M_{n,k} \subseteq R^{kn}$ be the feasible solution of the ith iteration step, D_i a $kn \times kn$ symmetric matrix defining a linear map $D_i : TM_{n,k}(\mathbf{x_i}) \to TM_{n,k}(\mathbf{x_i})$, and a positive definite quadratic form on $TM_{n,k}(\mathbf{x_i})$ with a uniform lower bound at every iteration point (i.e., $\mathbf{v}_i D_i \mathbf{v}_i \geq \varepsilon |\mathbf{v}_i|^2$, $\mathbf{v}_i \in TM_{n,k}(\mathbf{x_i})$, $\varepsilon > 0$ for all i) and G the induced Riemannian metric of $M_{n,k}$.

1. Compute the direction $\mathbf{p_i} \in R^{kn}$ as follows:

$$\mathbf{p_i} = -D_i(\mathbf{A} - S(\mathbf{x}_i))\mathbf{x_i}. \qquad (4.1)$$

2. Let

$$\mathbf{x_{i+1}} = \gamma_{\mathbf{x_i}}(t_i, \mathbf{p}_i), \qquad (4.2)$$

where $\gamma_{\mathbf{x_i}}(t_i, \mathbf{p}_i)$ means the arc of the geodesic starting from $\mathbf{x_i}$ with tangent $\mathbf{p_i}$ and the stepsize t_i is determined by an exact geodesic search to find the first local minimum along the above geodesic, i.e.,

$$t_i = \arg\min\{f(\gamma_{\mathbf{x_i}}(t, \mathbf{p}_i))|\ t \in R_+\} \qquad (4.3)$$

or chosen according to the Armijo principle (e.g., Ortega and Rheinboldt, 1970), i.e., given $\alpha \in (0, \frac{1}{2})$, let $t_i = 2^{-l_i}$, with l_i the smallest integer such that

$$f(\gamma_{\mathbf{x}_i}(2^{-l_i}, \mathbf{p}_i)) \leq f(\mathbf{x}_i) - \alpha 2^{-l_i} |\mathbf{x}_i^T(\mathbf{A} - S(\mathbf{x}_i)) D_i (\mathbf{A} - S(\mathbf{x}_i))\mathbf{x}_i|. \qquad (4.4)$$

Note that $(\mathbf{A} - S(\mathbf{x}))\mathbf{x}$, $\mathbf{x} \in M_{n,k}$, are the orthogonal projections of the gradient vectors ∇f with respect to the Euclidean metric of R^{kn} to the tangent spaces of the Stiefel manifold. So, this general framework contains several well-known nonlinear optimization algorithms: gradient projection method and reduced gradient method by choosing $D_i = I$ for all i; Newton-type methods and Newton-type methods along geodesics by choosing D_i based on the Hessian matrix HL of the Lagrangian function and the geodesic Hessian matrix H^gL at the ith iteration point, respectively. Thus, numerical techniques of smooth nonlinear optimization can be applied in these cases as well.

An algorithm is said to be globally convergent if the algorithm is guaranteed to generate a sequence of points converging to a critical point for arbitrary starting point. The global convergence theorems proved for variable metric methods along geodesics (Rapcsák and Thang, 1995; Rapcsák, 1997, 1998) can be adapted to this case. An advantage of variable metric methods along geodesics is that global convergence can be proved for a class of methods without assuming the positive or negative definiteness of the matrices $\mathbf{A}_1, \ldots, \mathbf{A}_k$.

Theorem 4.1 *If the sequence $\{\mathbf{x}_i\}$ is generated by (4.1), (4.2) and (4.3) or (4.1), (4.2) and (4.4), and if a mapping is defined by $D_i(\mathbf{A} - S(\mathbf{x}_i))\mathbf{x}_i \in TM_{n,k}(\mathbf{x}_i)$, for all i at the iteration points and D_i are positive definite on $TM_{n,k}(\mathbf{x}_i)$ with a uniform lower bound for all i, then the sequence $\{\mathbf{x}_i\}$ is either finite terminating at a critical point, or infinite, and every accumulation point is critical. If the critical values of f are distinct, the whole sequence $\{\mathbf{x}_i\}$ converges to a critical point.*

If $D_i = I$ for all i at the iteration points, then the statement is as follows.

Corollary 4.1 *If the sequence $\{\mathbf{x}_i\}$ is generated by (4.1), (4.2) and (4.3) or (4.1), (4.2) and (4.4), then the sequence $\{\mathbf{x}_i\}$ is either finite terminating at a critical point, or infinite, and every accumulation point is critical. If the critical values of f are distinct, the whole sequence $\{\mathbf{x}_i\}$ converges to a critical point.*

5. Proof of Theorem 4.1

Let W_i denote the connected component containing the $\mathbf{x}_i$ of the level set

$$\{\mathbf{x} \in M \subseteq \mathcal{M} |\ f(\mathbf{x}) \leq f(\mathbf{x}_i)\},$$

where $\mathcal{M}$ is equal to R^n endowed with a Riemannian metric $\mathcal{G}(\mathbf{x})$, $\mathbf{x} \in \mathcal{M}$, and M is a Riemannian submanifold with the induced metric G,

$$\mathbf{p}_i = -D_i^2 \nabla^G f(\mathbf{x}_i)^T, \qquad \nabla^G f(\mathbf{x}_i)^T \in R^n, \tag{5.1}$$

where $\nabla^G f^T$ is the Riemannian gradient vector (e.g., Rapcsák, 1997) and

$$f(\gamma_{\mathbf{x}_i}(2^{-l_i}, \mathbf{p}_i)) \leq f(\mathbf{x}_i) - \alpha 2^{-l_i} |\nabla^G f(\mathbf{x}_i) \mathcal{G}(\mathbf{x}_i) D_i^2 \nabla^G f(\mathbf{x}_i)^T|. \tag{5.2}$$

The following statement is a combination of two global convergence theorems found in Rapcsák and Thang, 1995 and Rapcsák, 1997, 1998, respectively.

Theorem 5.1 *If f is continuously differentiable, W_0 is a compact set in a complete Riemannian submanifold $M \subseteq \mathcal{M}$, the sequence $\{\mathbf{x}_i\}$ is generated by (5.1), (4.2) and (4.3) or (5.1), (4.2) and (5.2), and if a mapping is defined by $D_i \nabla^G f(\mathbf{x}_i)^T \in TM(\mathbf{x}_i)$, for all i at the iteration points and satisfies there the Lipschitz condition, D_i are positive definite on $TM(\mathbf{x}_i)$ with a uniform lower bound for all i, and D_i and $\mathcal{G}$ commute on $TM(\mathbf{x}_i)$ at every iteration point, then the sequence $\{\mathbf{x}_i\}$ is either finite terminating at a critical point, or infinite, and every accumulation point is critical. If the critical values of f are distinct, the whole sequence $\{\mathbf{x}_i\}$ converges to a critical point.*

By Theorem 11.1.1 in [9], if R^{kn} is endowed with a Riemannian metric $\mathcal{G}$ in problem (2.3), then the Riemannian gradient vector is equal to

$$\nabla^G f^T = (I - \mathcal{G}^{-1}[J\mathbf{h}^T, J\hat{\mathbf{h}}^T]([J\mathbf{h}, J\hat{\mathbf{h}}]\mathcal{G}^{-1}[J\mathbf{h}^T, J\hat{\mathbf{h}}^T])^{-1}[J\mathbf{h}^T, J\hat{\mathbf{h}}])\mathcal{G}^{-1}\nabla f^T. \tag{5.3}$$

In the case of problem (2.3), the space R^{kn} is endowed with the Euclidean metric, i.e., $\mathcal{G}(\mathbf{x}) = I_{kn}$, $\mathbf{x} \in R^{kn}$, the Riemannian metrics of the Stiefel manifolds are induced by the Euclidean metric, the objective function is, obviously, continuously differentiable, the set W_0 is compact, the Stiefel manifolds are complete, the matrices D_i^2 can be replaced by D_i for all i, because $\mathbf{p}_i$ given by (4.1) is a descent direction for every i, thus the statement of Theorem 4.1 is a consequence of Theorem 5.1. ∎

Acknowledgement. This is to express my acknowledgement to Gy. Michaletzky whose precious ideas greatly contributed to this paper.

References

[1] Adams, J. F., Vector fields on spheres, Annales of Mathematics 75 (1962) 603-632.

[2] Bolla, M., Michaletzky, Gy., Tusnády, G. and Ziermann, M., Extrema of sums of heterogeneous quadratic forms, Linear Algebra and Applications 269 (1998) 331-365.

[3] Gabay, D., Minimizing a differentiable function over a differentiable manifold, Journal of Optimization Theory and Applications 37 (1982) 177-219.

[4] James, J. M., The topology of Stiefel manifolds, London Mathematical Society, Lecture Notes Series 24, Cambridge University Press, Cambridge, 1976.

[5] Luenberger, D. G., Introduction to linear and nonlinear programming, Addison-Wesley Publishing Company, Reading, 1973.

[6] Ortega, J.M. and Rheinboldt, N.C., Iterative solution of nonlinear equations in several variables, Academic Press, New York, New York, 1970.

[7] Rapcsák, T., Geodesic convexity in nonlinear optimization, Journal of Optimization Theory and Applications 69 (1991) 169-183.

[8] Rapcsák, T. and Thang, T. T., Nonlinear coordinate representations of smooth optimization problems, Journal of Optimization Theory and Applications 86 (1995) 459-489.

[9] Rapcsák, T., Smooth nonlinear optimization in R^n, Kluwer Academic Publishers, Boston, London, Dordrecht, 1997.

[10] Rapcsák, T., Variable metric methods along geodesics, in: New trends in mathematical programming, F. Giannessi, S. Komlósi and T. Rapcsák (eds.), Kluwer Academic Publishers (1998) 257-275.

[11] Stiefel, E., Richtungsfelder und Fernparallelismus in n-dimensionalen Mannigfaltigkeiten, Commentarii Mathematici Helvetici 8 (1935/6) 305-353.