J Glob Optim (2008) 40:375–388
DOI 10.1007/s10898-007-9212-7

# Sectional curvatures in nonlinear optimization

**Tamás Rapcsák**

Received: 31 May 2007 / Accepted: 11 July 2007 / Published online: 3 August 2007


**Abstract** The aim of the paper is to show how to explicitly express the function of sectional curvature with the first and second derivatives of the problem's functions in the case of submanifolds determined by equality constraints in the $n$-dimensional Euclidean space endowed with the induced Riemannian metric, which is followed by the formulation of the minimization problem of sectional curvature at an arbitrary point of the given submanifold as a global minimization one on a Stiefel manifold. Based on the results, the sectional curvatures of Stiefel manifolds are analysed and the maximal and minimal sectional curvatures on an ellipsoid are determined.

**Keywords** Nonlinear programming · Special Riemannian manifolds · Critical points

**AMS Subject Classifications 2000:** 90C30 · 53C25 · 57R70

## 1 Introduction

In differential geometry, one of the main tools for studying the structure of the Riemannian manifolds seems to be the sectional curvature [10]. A famous result is that the fundamental Euclidean, Riemannian elliptic and Bolyai-Lobachevsky hyperbolic manifolds are characterized by zero, positive and negative constant sectional curvature, respectively. The definition of the sectional curvature is based on the fourth-order Riemannian curvature tensor field and the Riemannian metric (e.g., [12]), so concrete calculations should need extremely difficult procedures. Besides the sectional curvature having interesting geometric interpretations, its importance comes from the fact that the knowledge of sectional curvatures determines

This research was supported in part by the Hungarian Scientific Research Fund, Grant No. OTKA-T043276 and OTKA-K60480.

T. Rapcsák (✉)
Computer and Automation Research Institute,
Hungarian Academy of Sciences, Kende u. 13-17, 1111 Budapest, Hungary
e-mail: rapcsak@sztaki.hu

the Riemannian curvature tensor field completely. For classifying the constraints of smooth nonlinear optimization problems, a possibility is to use the sectional curvatures.

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 $k \leq n$. In the paper, it will be shown that the computation of the minimal sectional curvature in the case of submanifolds determined by equality constraints in the Euclidean space $R^n$ endowed with the induced Riemannian metric leads to solving minimization problems on Stiefel manifolds.

Let

$$M[\mathbf{h}] = \{\mathbf{x} \in R^n \mid h_j(\mathbf{x}) = 0, \ j = 1, \ldots, n-k\}, \tag{1}$$

where $k > 0$, $h_j \in C^2$, $j = 1, \ldots, n-k$, and $\mathbf{0}$ is a regular value of the map $\mathbf{h}$, i.e., the $(n-k) \times n$ Jacobian matrix $J\mathbf{h}(\mathbf{x})$ of $\mathbf{h}$ at $\mathbf{x}$ is of full rank $(n-k)$ for all $\mathbf{x} \in M[\mathbf{h}]$. Under these assumptions, the set $M[\mathbf{h}]$ is a $k$-dimensional submanifold of $C^2$ in $R^n$ (e.g., [4]) which can be endowed with a Riemannian metric $G$. In optimization theory, the Riemannian metric is often induced by the Euclidean metric of $R^n$ [7].

The aim of the paper is to show how to explicitly express the function of sectional curvature with the first and second derivatives of the problem's functions on the submanifold $M[\mathbf{h}]$ endowed with the induced Riemannian metric, and how to formulate the minimization problems of sectional curvatures related to $M[\mathbf{h}]$. At an arbitrary point $\mathbf{x}_0 \in M[\mathbf{h}]$, this minimization problem leads to a global minimization one on Stiefel manifolds [11], which seems to be an interesting new branch of nonlinear optimization [3,9]. After obtaining the optimality conditions for the minimization problems of sectional curvatures, the sectional curvatures of Stiefel manifolds are analysed and the maximal and minimal sectional curvatures on an ellipsoid are determined. These curvatures are proportional with the condition number of the given matrix. The paper is ended with some open problems.

## 2 Sectional curvatures on Riemannian manifolds

Let $M$ be an $n$-dimensional differentiable manifold, and let the tangent space of $M$ at an arbitrary point $m \in M$ be denoted by $TM(m)$. The tangent space $TM(m)$ is a linear space and has the same dimension as $M$. Because we restrict ourselves to real manifolds, $TM(m)$ is isomorphic to $R^n$. If $M$ is endowed with a Riemannian metric $G$, then $M$ is a Riemannian manifold denoted by $(M, G)$. The inner product of two tangent vectors $\mathbf{v}_1, \mathbf{v}_2 \in TM(m)$ is equal to

$$\langle \mathbf{v}_1, \mathbf{v}_2 \rangle_m = G(m, \mathbf{v}_1, \mathbf{v}_2), \tag{2}$$

where $G(m)$ is the Riemannian metric at the point $m$. The norm of a tangent vector $\mathbf{v} \in TM(m)$ is defined by

$$\| \mathbf{v} \|_m = \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle_m}\,. \tag{3}$$

**Definition 1** [12, p.8] Let $(M, G)$ be a Riemannian manifold with the Riemannian curvature tensor field $R$. Let $m$ be a point in $M$ and $V$ a two-dimensional vector subspace of the tangent space $TM(m)$. Suppose that $\{\mathbf{v}_1, \mathbf{v}_2\}$ is a basis of $V$. The real number

$$K(m, V) = \frac{R(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_1, \mathbf{v}_2)}{G(m, \mathbf{v}_1, \mathbf{v}_1)G(m, \mathbf{v}_2, \mathbf{v}_2) - G(m, \mathbf{v}_1, \mathbf{v}_2)^2} \tag{4}$$

is said to be the sectional curvature of $M$ at $m$ along the section $V$.

It is an important property that the sectional curvature $K(m, V)$ does not depend on the particular choice of the basis of $V$. Oprea [6] studied the minimization of sectional curvatures (4) on Riemannian manifolds, and gave a first-order characterization of critical sections.

A beautiful theorem of global differential geometry is the sphere theorem which states as follows:

**Theorem 1** [2] *Let $(M, G)$ be a compact simply connected Riemannian manifold whose sectional curvatures $K(m, V)$, $m \in M$, for all two-dimensional vector subspaces of the tangent space $TM(m)$, satisfy*

$$0 < \alpha\, K_{\max} < K(m, V) \leq K_{\max}, \qquad m \in M, \qquad \alpha = 1/4. \tag{5}$$

*Then, $(M, G)$ is homeomorphic to a sphere.*

A Riemannian manifold $(M, G)$ has a constant sectional curvature $K_0$ if for all $m \in M$ and all two-dimensional vector subspaces $V$ of $TM(m)$, we have $K(m, V) = K_0$. If $M$ is two-dimensional, this implies that $K(m) = K_0$ for all $m \in M$. It is not difficult to verify that if a Riemannian metric is multiplied by a positive constant $c$, then the sectional curvatures are multiplied by 1/c. Two examples for Riemannian manifolds with constant sectional curvature $K_0$: the Euclidean space $R^n$ with $K_0 = 0$ and the unit sphere $S^n \subset R^{n+1}$ with $K_0 = 1$.

A point $m$ in an $n$-dimensional Riemannian manifold $(M, G)$ is called isotropic if the sectional curvatures $K(m, V)$ for all two-dimensional vector subspaces $V$ of $TM(m)$ have the same value $K_0$.

**Schur theorem (1886)** *If $M$ is a connected Riemannian manifold of dimension $n \geq 3$ and all points are isotropic, then $M$ has a constant sectional curvature.*

It is known that the Gaussian curvature of a surface in the three-dimensional Euclidean space is the sectional curvature. For $n = 2$, the statement does not apply: there are surfaces with nonconstant Gaussian curvature. The sphere theorem can be considered a generalization of Schur theorem in which the sectional curvatures belong to the same interval at every point.

**Hilbert theorem (1901)** *A complete surface $M$ with constant curvature $K_0 = -1$ cannot be immersed in $R^3$.*

## 3 Sectional curvatures on Euclidean submanifolds with the induced Riemannian metric

The problem to be solved is to determine the minimal sectional curvature of the submanifold $M[\mathbf{h}]$ endowed with the induced Riemannian metric at an arbitrary point $\mathbf{x}_0 \in M[\mathbf{h}]$. First, the function of sectional curvatures is explicitly expressed with the first and second derivatives of the problem's functions.

**Theorem 2** *Let $\mathbf{x}_0 \in M[\mathbf{h}]$ be an arbitrary point, $\mathbf{n}_1, \ldots, \mathbf{n}_{n-k}$ an orthonormal basis with respect to the Euclidean metric of the orthogonal subspace of the tangent space $TM[\mathbf{h}](\mathbf{x}_0)$, and $\mathbf{w}_1, \mathbf{w}_2 \in R^n$ an orthonormal pair of vectors with respect to the Euclidean metric in $TM[\mathbf{h}](\mathbf{x}_0)$. Then, the sectional curvature of $M[\mathbf{h}]$ at the point $\mathbf{x}_0$ along the section $\{\mathbf{w}_1, \mathbf{w}_2\}$ is equal to*

$$K_{M[\mathbf{h}]}(\mathbf{x}_0, \mathbf{w}_1, \mathbf{w}_2)$$
$$= \sum_{i=1}^{n-k} \left[ \left( \sum_{j=1}^{n-k} \mu_{ij}(\mathbf{x_0}) \mathbf{w}_1^T H_{\mathbf{x}} h_j(\mathbf{x}_0) \mathbf{w}_1 \right) \left( \sum_{j=1}^{n-k} \mu_{ij}(\mathbf{x_0}) \mathbf{w}_2^T H_{\mathbf{x}} h_j(\mathbf{x_0}) \mathbf{w}_2 \right) - \left( \sum_{j=1}^{n-k} \mu_{ij}(\mathbf{x_0}) \mathbf{w}_1^T H_{\mathbf{x}} h_j(\mathbf{x_0}) \mathbf{w}_2 \right)^2 \right], \tag{6}$$

*where*

$$\boldsymbol{\mu}_i^T(\mathbf{x}_0) = \mathbf{n}_i^T J\mathbf{h}(\mathbf{x_0})^T \left[ J\mathbf{h}(\mathbf{x_0}) J\mathbf{h}(\mathbf{x_0})^T \right]^{-1}, \qquad i = 1, \ldots, n-k.$$

*Proof* Let $M_1$ be an $n$-dimensional Riemannian manifold with metric $G_1$ and $M$ a $k$-dimensional submanifold with $n - k > 0$. An immersion of $M$ into $M_1$ is a $C^2$ map $F : M \to M_1$ such that the derivative $dF$ ($JF$ in a coordinate representation) of $F$ is one-to-one on every tangent space $TM(m)$, $m \in M$. The induced Riemannian metric or the first fundamental form of the immersion is given by $\langle dF, dF \rangle$ ($= JF^T G_1 JF$ in a coordinate representation) which makes $M$ a Riemannian manifold.

Let $K_M(m, \mathbf{v}_1, \mathbf{v}_2)$ and $K_{M_1}(m, \mathbf{v}_1, \mathbf{v}_2)$ denote the sectional curvatures of the manifolds $M$ and $M_1$ at any point $m \in M$ and for every pair of tangents $\mathbf{v}_1, \mathbf{v}_2 \in TM(m)$, respectively. Let $\mathbf{n}_1, \ldots, \mathbf{n}_{n-k}$ be an orthonormal basis of the orthogonal subspace of the tangent space $TM(m)$ denoted by $TM^{\perp}(m)$ and $\mathbf{v}_1, \mathbf{v}_2$ an orthonormal pair of vectors with respect to the induced Riemannian metric in $TM(m)$. Then, the sectional curvature of $M$ at the point $m$ along the section $\{\mathbf{v}_1, \mathbf{v}_2\}$ is equal to

$$K_M(m, \mathbf{v}_1, \mathbf{v}_2) = K_{M_1}(m, \mathbf{v}_1, \mathbf{v}_2) + \sum_{i=1}^{n-k} \left( B_{\mathbf{n}_i}(\mathbf{v}_1, \mathbf{v}_1) B_{\mathbf{n}_i}(\mathbf{v}_2, \mathbf{v}_2) - B_{\mathbf{n}_i}(\mathbf{v}_1, \mathbf{v}_2)^2 \right), \tag{7}$$

where $B_{\mathbf{n}_i}$, $i = 1, \ldots, n-k$, is the second fundamental form in the direction $\mathbf{n}_i$, $i = 1, \ldots, n-k$, respectively. Thus, the induced sectional curvature on $M$ is equal to the sectional curvature on $M_1$ plus the sum of the "squares of areas" with respect to the second fundamental forms (see, e.g., [1, p. 193]).

Let us consider the manifold $M[\mathbf{h}]$ and assume that the Riemannian metric is induced by the Euclidean metric of $R^n$. Consider an arbitrary coordinate representation of the manifold $M[\mathbf{h}]$ in any neighbourhood given by the smooth vector function $\mathbf{x}(\mathbf{u})$, $\mathbf{u} \in U \subseteq R^k$, $\mathbf{x} \in R^n$, where $U$ is an open set. The second fundamental form of the manifold $M$ immersed in $R^n$ at $\mathbf{u}_0 \in U$ in the direction of a normal unit vector $\mathbf{n}$ is the quadratic form $\mathbf{v}^T B_{\mathbf{n}}(\mathbf{u}_0)\mathbf{v}$, $\mathbf{v} \in R^k$, where the elements of the $k \times k$ matrix $B_{\mathbf{n}}(\mathbf{u}_0)$, $\mathbf{u}_0 \in U$, are

$$b_{ij}(\mathbf{u}_0) = \left( \frac{\partial^2 \mathbf{x}(\mathbf{u}_0)}{\partial u_i \, \partial u_j} \right)^T \mathbf{n}, \qquad i, j = 1, \ldots, k, \qquad \mathbf{u}_0 \in U \subseteq R^k,$$

(e.g., [7, p. 31]). The formula of the second fundamental form implies an operation between vectors and three-dimensional hypermatrices. Let

$$H\mathbf{x}(\mathbf{u}_0) = \begin{pmatrix} Hx_1(\mathbf{u}_0) \\ \vdots \\ Hx_n(\mathbf{u}_0) \end{pmatrix}, \qquad \mathbf{u}_0 \in U \subseteq R^k,$$

then, $\mathbf{n}^T H\mathbf{x}(\mathbf{u}_0) = H\mathbf{x}(\mathbf{u}_0)\mathbf{n} = \sum_{l=1}^{n} n_l H x_l(\mathbf{u}_0)$.

By formula (9.3.6) on p. 150 in [7], the second covariant derivatives

$$D^2\left(\mathbf{n}_i^T\mathbf{x}(\mathbf{u})\right) = J\mathbf{x}(\mathbf{u})^T H_{\mathbf{x}}\left(\mathbf{n}_i^T\mathbf{x}(\mathbf{u})\right) J\mathbf{x}(\mathbf{u}) + \nabla_{\mathbf{x}}\left(\mathbf{n}_i^T\mathbf{x}(\mathbf{u})\right) H\mathbf{x}(\mathbf{u})$$
$$-\nabla_{\mathbf{x}}\left(\mathbf{n}_i^T\mathbf{x}(\mathbf{u})\right) J\mathbf{x}(\mathbf{u})\Gamma(\mathbf{u}), \qquad \mathbf{u}\in U\subseteq R^k, \qquad i=1,\ldots,n-k, \tag{8}$$

where $\mathbf{n}_i,\ i=1,\ldots,n-k$, are constant orthonormal vectors of $TM[\mathbf{h}]^{\perp}(\mathbf{x}_0)$, $J\mathbf{x}(\mathbf{u})$ is the Jacobian matrix, the matrix multiplication $J\mathbf{x}(\mathbf{u})\Gamma(\mathbf{u}), \mathbf{u}\in U\subseteq R^k$, is defined by the rule related to the multiplication of a row vector and a three-dimensional matrix, applied consecutively to every row vector of $J\mathbf{x}(\mathbf{u})$.

From (8), we obtain that

$$D^2\left(\mathbf{n}_i^T\mathbf{x}(\mathbf{u}_0)\right) = B_{\mathbf{n}_i}(\mathbf{u}_0), \quad i=1,\ldots,n-k, \qquad \mathbf{u}_0\in U\subseteq R^k. \tag{9}$$

By Theorem 9.5.3 on p. 159 in [7], if $M[\mathbf{h}]$ is given by (1), then the second covariant derivatives of the functions $\mathbf{n}_i^T\mathbf{x},\ \mathbf{x}\in M[\mathbf{h}], i=1,\ldots,n-k$, can be formulated in an equivalent form of

$$D^2\left(\mathbf{n}_i^T\mathbf{x}\right) = \left(\sum_{j=1}^{n-k}\mu_{ij}(\mathbf{x})H_{\mathbf{x}}h_j(\mathbf{x})\right)_{|TM}, \qquad \mathbf{x}\in M[\mathbf{h}], \qquad i=1,\ldots,n-k, \tag{10}$$

where

$$\boldsymbol{\mu}_i^T(\mathbf{x}) = \mathbf{n}_i^T J\mathbf{h}(\mathbf{x})^T\left[J\mathbf{h}(\mathbf{x})J\mathbf{h}(\mathbf{x})^T\right]^{-1}, \qquad \mathbf{x}\in M[\mathbf{h}] \qquad i=1,\ldots,n-k, \tag{11}$$

and the tangent vectors $\mathbf{w}\in R^n$ of $TM[\mathbf{h}](\mathbf{x}_0)$ are $\mathbf{w}=J\mathbf{x}(\mathbf{u}_0)\mathbf{v},\ \mathbf{v}\in R^k$. Thus, in the case of an orthonormal pair of tangent vectors with respect to the induced Riemannian metric $\mathbf{v}_1,\mathbf{v}_2\in R^k$, we have that $\mathbf{w}_1=J\mathbf{x}(\mathbf{u}_0)\mathbf{v}_1$ and $\mathbf{w}_2=J\mathbf{x}(\mathbf{u}_0)\mathbf{v}_2$ are orthonormal with respect to the Euclidean metric.

By substituting relations (10) for (7) and taking the equality $K_{R^n}(\mathbf{x}_0,\mathbf{v}_1,\mathbf{v}_2)=0$ into account, the statement is proved. □

*Example 1* Consider the sphere in $R^3$ given by

$$M_S = M[h] = \left\{\mathbf{x}\in R^3 \mid h(\mathbf{x}) = \frac{1}{2}\left(x_1^2+x_2^2+x_3^2\right) - \frac{1}{2} = 0\right\}. \tag{12}$$

By differentiating the function $h$ in (12), we have that

$$\nabla h(\mathbf{x}) = (x_1,x_2,x_3), \qquad \|\nabla h(\mathbf{x})\| = \nabla h(\mathbf{x})^T\nabla h(\mathbf{x}) = 1, \qquad \mathbf{x}\in M_S,$$

$$Hh(\mathbf{x}) = I_3 = \begin{pmatrix}1&&\\&1&\\&&1\end{pmatrix}, \qquad \mathbf{x}\in M_S.$$

Thus, the unique unity normal vectors are $\nabla h(\mathbf{x}),\ \mathbf{x}\in M_S$, and $\mu(\mathbf{x})=1,\ \mathbf{x}\in M_S$.

Let $\mathbf{w}_1,\ \mathbf{w}_2$ be an orthonormal pair of tangent vectors with respect to the Euclidean metric at an arbitrary point of $M_S$, then by Theorem 2,

$$K_{M_S}(\mathbf{x},\mathbf{w}_1,\mathbf{w}_2) = 1, \qquad \mathbf{x}\in M_S, \tag{13}$$

which proves that $M_S$ has a constant sectional curvature.

In order to determine the minimal sectional curvature of the manifold $M[\mathbf{h}]$ with the induced Riemannian metric at an arbitrary given point $\mathbf{x}_0 \in M[\mathbf{h}]$, the following nonlinear global optimization problem has to be solved:

$$\begin{gathered} \min K_{M[\mathbf{h}]}(\mathbf{x}_0, \mathbf{w}_1, \mathbf{w}_2) \\ \mathbf{w}_1^T \mathbf{w}_1 = 1, \qquad \mathbf{w}_2^T \mathbf{w}_2 = 1, \qquad \mathbf{w}_1^T \mathbf{w}_2 = 0, \qquad \mathbf{w}_1, \mathbf{w}_2 \in R^n. \end{gathered} \tag{14}$$

We remark that the section $\{\mathbf{w}_1, \mathbf{w}_2\}$ corresponds to the section

$$\begin{aligned} \{\mathbf{v}_1 &= J\mathbf{h}(\mathbf{x}_0)^T [J\mathbf{h}(\mathbf{x}_0) J\mathbf{h}(\mathbf{x}_0)^T]^{-1} \mathbf{w}_1 \in R^k, \\ \mathbf{v}_2 &= J\mathbf{h}(\mathbf{x}_0)^T [J\mathbf{h}(\mathbf{x}_0) J\mathbf{h}(\mathbf{x}_0)^T]^{-1} \mathbf{w}_2 \in R^k\} \end{aligned}$$

at the tangent space $TM[\mathbf{h}](\mathbf{x}_0)$.

## 4 Minimization on Stiefel manifolds

Consider the following optimization problem:

$$\begin{gathered} \min f(\mathbf{w}_1, \mathbf{w}_2, \ldots, \ldots, \mathbf{w}_k) \\ \mathbf{w}_i^T \mathbf{w}_j = \delta_{ij}, \qquad 1 \le i, j \le k \le n, \\ \mathbf{w}_i \in R^n, \quad i = 1, \ldots, k, \qquad n \ge 2, \end{gathered} \tag{15}$$

where $f$: $R^{kn} \to R$ is a twice continuously differentiable function and $\delta_{ij}$ is the Kronecker's delta. Since the feasible set is compact and the objective function is continuous, optimization problem (15) has, at least, a global minimum point and a global maximum point, thus several stationary points. The feasible set of problem (15), denoted by $M_{n,k}$, consists of all the orthonormal vector systems $\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_k \in R^n$, $k \le n$, and can be written as

$$\mathbf{w}_i^T \mathbf{w}_i = 1, \qquad i = 1, \ldots, k, \tag{16}$$

$$\mathbf{w}_i^T \mathbf{w}_j = 0, \qquad i, j = 1, \ldots, k, \quad i \ne j, \tag{17}$$

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

Equalities (16), (17), and equalities (17) determine a compact and a noncompact set, respectively. These constraints seem to be useful to describe independent groups of variables in the phase of modelling real-life problems.

If $k = 1$ and $f(\mathbf{w}) = \frac{1}{2}\mathbf{w}^T A\mathbf{x}$, then problem (15) is equal to the classical eigenvalue problem

$$\begin{gathered} \min \frac{1}{2}\mathbf{w}^T A\mathbf{w} \\ ||\mathbf{w}||^2 = 1, \qquad \mathbf{w} \in R^n. \end{gathered} \tag{18}$$

In order to study the geometric structure of the feasible set, a new representation of the same was suggested in [8,9] providing a decomposition of the feasible set as well. Let us introduce the following notations:

$$\mathbf{w} = \left(\mathbf{w}_1^T, \ldots, \mathbf{w}_k^T\right)^T \in R^{kn}, \qquad J = \{(i, j) \mid i, j = 1, \ldots, k, \; i < j\},$$

$$C_l = \begin{pmatrix} 0 & \ldots & 0 & \ldots & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & \ldots & I_n & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \ldots & 0 & \ldots & 0 \end{pmatrix}, \qquad l = 1, \ldots, k,$$

$$C_{ij} = \begin{pmatrix} 0 & \ldots & 0 & \ldots & 0 & \ldots \\ \vdots & \ddots & \vdots & \ddots & \vdots & \vdots \\ 0 & \ldots & 0 & \ldots & I_n & \ldots \\ \vdots & \ddots & \vdots & \ddots & \vdots & \vdots \\ 0 & \ldots & I_n & \ldots & 0 & \ldots \\ 0 & \ldots & 0 & \ldots & 0 & \ldots \end{pmatrix}, \qquad (i, j) \in J,$$

where $C_l$, $l = 1, \ldots, k$, are $kn \times kn$ blockdiagonal matrices, $C_{ij}$ $kn \times kn$ block matrices, $I_n$ is the identity matrix in $R^n$, $C_l$ and $C_{ij}$ 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 $C_{ij}$ are defined for all the pairs of different indices belonging to $J$, given by the $k\,(k-1)\,/2$ combinations of the indices $1, \ldots, k$.

It follows that in the case of a compact Stiefel manifold, the feasible set $M_{n,k}$ given by (16) and (17) is equivalent to

$$\begin{aligned} h_l\,(\mathbf{w}) &= \frac{1}{2}\mathbf{w}^T C_l \mathbf{w} - \frac{1}{2} = 0, \qquad l = 1, \ldots, k, \\ h_{ij}\,(\mathbf{w}) &= \frac{1}{2}\mathbf{w}^T C_{ij} \mathbf{w} = 0, \qquad (i, j) \in J, \\ \mathbf{w} &\in R^{kn}, \qquad n \geq 2. \end{aligned} \tag{19}$$

In the definition of the index set $J$, the restriction $i < j$ ensures that only one of the identical equalities $h_{ij}\,(\mathbf{w}) = 0$ and $h_{ji}\,(\mathbf{w}) = 0$, $i, j = 1, \ldots, k,\; i \neq j$ appears in (19).

Thus, the feasible set $M_{n,k}$ and its tangent space at the point $\mathbf{w} \in M_{n,k}$ can be described by

$$M_{n,k} = \left\{\mathbf{w} \in R^{kn} \mid h_l\,(\mathbf{w}) = 0,\; l = 1, \ldots, k; \quad h_{ij}\,(\mathbf{w}) = 0, \quad (i, j) \in J\right\}, \tag{20}$$

$$\begin{aligned} TM_{n,k}\,(\mathbf{w}) &= \left\{\mathbf{v} \in R^{kn} \mid \nabla h_l\,(\mathbf{w})\,\mathbf{v} = 0, l = 1, \ldots, k; \quad \nabla h_{ij}\,(\mathbf{w})\,\mathbf{v} = 0, \quad (i, j) \in J\right\} \\ &= \left\{\mathbf{v} \in R^{kn} \mid \mathbf{w}_l^T \mathbf{v}_l = 0, \quad l = 1, \ldots, k; \quad \mathbf{w}_i^T \mathbf{v}_j + \mathbf{w}_j^T \mathbf{v}_i = 0, \quad (i, j) \in J\right\}, \\ &\quad \mathbf{w} \in M_{n,k}, \end{aligned} \tag{21}$$

where the symbol $\nabla$ denotes the gradient vector of a function which is a row vector.

The following statements characterizing the structure of the feasible set can be found with a simple proof in [9].

**Theorem 3** *The set $M_{n,k}$ is a compact $C^\infty$ differentiable manifold (Stiefel manifold) with dimension $kn - \frac{k(k+1)}{2}$ for every pair of positive integers $(k, n)$ satisfying $k \leq n$. The Stie-*

*fel manifolds are connected if $k < n$. In the cases $k = n$, the Stiefel manifolds are of two components.*

Now, following [9], the first-order and second-order, necessary and sufficient, local and global optimality conditions of problem (15) are stated.

By using the equality representations of the compact Stiefel manifolds $M_{n,k}$, problem (15) is equivalent to

$$\begin{aligned} &\min \ f(\mathbf{w}) \\ &h_l(\mathbf{w}) = \frac{1}{2}\mathbf{w}^T C_l \mathbf{w} - \frac{1}{2} = 0, \qquad l = 1, \ldots, k, \\ &h_{ij}(\mathbf{w}) = \frac{1}{2}\mathbf{w}^T C_{ij}\mathbf{w} = 0, \qquad (i, j) \in J, \\ &\mathbf{w} \in R^{kn}, \qquad n \geq 2. \end{aligned} \tag{22}$$

Problem (22) is one of the basic equality constrained problems in smooth optimization studied in most of the classical literature (e.g., [5]). The difficulty in the solution of problems (22) 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 $\mathbf{w}^T\mathbf{w} = k$ in $R^{kn}$, and from the nonconvexity of the objective function. It is emphasized that by Theorem 3, the feasible set of problem (15) is connected if $k < n$, and it has two connected components if $n = k$.

Before stating the optimality conditions, the 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 $\mathcal{C} \subseteq M$ 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. Let us introduce the symmetric matrix function

$$\begin{aligned} &S(\mathbf{w}) \\ &= \begin{pmatrix} (\nabla f(\mathbf{w})C_1\mathbf{w})I_n & \frac{1}{2}(\nabla f(\mathbf{w})C_{12}\mathbf{w})I_n & \ldots & \frac{1}{2}(\nabla f(\mathbf{w})C_{1k}\mathbf{w})I_n \\ \frac{1}{2}(\nabla f(\mathbf{w})C_{12}\mathbf{w})I_n & (\nabla f(\mathbf{x})C_2\mathbf{w})I_n & \ldots & \frac{1}{2}(\nabla f(\mathbf{w})C_{2k}\mathbf{w})I_n \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1}{2}(\nabla f(\mathbf{w})C_{1k}\mathbf{w})I_n & \frac{1}{2}(\nabla f(\mathbf{w})C_{2k}\mathbf{w})I_n & \ldots & (\nabla f(\mathbf{w})C_k\mathbf{w})I_n \end{pmatrix}, \\ &\mathbf{w} = \left(\mathbf{w}_1^T, \ldots, \mathbf{w}_k^T\right)^T \in M_{n,k}. \end{aligned} \tag{23}$$

**Theorem 4** [9] *If the point $\mathbf{w}_0 \in M_{n,k}$ is a (strict) local minimum of problem (22), then*

$$\nabla f(\mathbf{w}_0) = \mathbf{w}_0^T S(\mathbf{w}_0), \qquad \textit{and} \tag{24}$$

$$\Big(Hf(\mathbf{w}_0) - S(\mathbf{w}_0)\Big)_{|TM_{n,k}(\mathbf{w}_0)} \tag{25}$$

*is a positive semidefinite (definite) matrix where the symbol $|TM_{n,k}(\mathbf{w}_0)$ denotes the restriction to the tangent space at the point $\mathbf{w}_0$ and $H$ the Hessian matrix of a function.*

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

$$\nabla f(\mathbf{w}_0) = \mathbf{w}_0^T S(\mathbf{w}_0), \qquad and$$
$$\left(Hf(\mathbf{w}) - S(\mathbf{w})\right)_{|TM_{n,k}(\mathbf{w})}, \qquad \mathbf{w} \in \mathcal{C}, \tag{26}$$

*are positive semidefinite (definite) matrices, then the point $\mathbf{w}_0$ is a (strict) global minimum of the function $f$ on the set $C$.*

## 5 Sectional curvatures on Stiefel manifolds with the induced Riemannian metric

The aim of this part is to determine the sectional curvatures of Stiefel manifolds endowed with the induced Riemannian metric.

**Theorem 5** *Let $\tilde{\mathbf{x}} \in M_{n,k}$ be an arbitrary point,*

$$C_l\tilde{\mathbf{x}}, \qquad l = 1, \ldots, k,$$
$$\frac{1}{\sqrt{2}} C_{ij}\tilde{\mathbf{x}}, \qquad (i, j) \in J, \tag{27}$$

*an orthonormal basis with respect to the Euclidean metric of the orthogonal subspace of the tangent space $TM_{n,k}(\tilde{\mathbf{x}})$, and $\tilde{\mathbf{w}}_1, \tilde{\mathbf{w}}_2$ an orthonormal pair of vectors with respect to the Euclidean metric in $TM_{n,k}(\tilde{\mathbf{x}})$. Then, the sectional curvature of $M_{n,k}$ at the point $\tilde{\mathbf{x}}$ along the section $\{\tilde{\mathbf{w}}_1, \tilde{\mathbf{w}}_2\}$ is equal to*

$$K_{M_{n,k}}(\tilde{\mathbf{x}}, \tilde{\mathbf{w}}_1, \tilde{\mathbf{w}}_2) = \sum_{l=1}^{k}\left(\mathbf{w}_{1l}^2\mathbf{w}_{2l}^2 - (\mathbf{w}_{1l}^T\mathbf{w}_{2l})^2\right)$$
$$+\frac{1}{\sqrt{2}}\sum_{(i,j)\in J}\left(\left(2\mathbf{w}_{1i}^T\mathbf{w}_{1j}\right)\left(2\mathbf{w}_{2i}^T\mathbf{w}_{2j}\right)\right.$$
$$\left.-\left(\mathbf{w}_{1i}^T\mathbf{w}_{2j} + \mathbf{w}_{1j}^T\mathbf{w}_{2i}\right)^2\right). \tag{28}$$

*Proof* In the case of Stiefel manifolds $M_{n,k}$, due to formulae (19), (21), let the orthonormal basis of $TM_{n,k}^{\perp}(\mathbf{x})$, $\mathbf{x} \in M_{n,k}$, be given by the gradients of the equality constraints

$$C_l\mathbf{x}, \qquad l = 1, \ldots, k,$$
$$\frac{1}{\sqrt{2}} C_{ij}\mathbf{x}, \qquad (i, j) \in J.$$

By (11) and the orthonormality of the basis, the vector functions

$$\boldsymbol{\mu}_l(\mathbf{x}), \qquad l = 1, \ldots, k, \qquad \mathbf{x} \in M_{n,k},$$
$$\boldsymbol{\mu}_{(i,j)}(\mathbf{x}), \qquad (i, j) \in J, \qquad \mathbf{x} \in M_{n,k},$$

are of the constant unity vectors.

Consider an orthonormal pair of tangent vectors $\widetilde{\mathbf{w}}_1, \widetilde{\mathbf{w}}_2 \in TM_{n,k}(\tilde{\mathbf{x}})$ at an arbitrary point $\tilde{\mathbf{x}} \in M_{n,k}$. By Theorem 2,

$$
\begin{aligned}
&K_{M_{n,k}}(\tilde{\mathbf{x}}, \widetilde{\mathbf{w}}_1, \widetilde{\mathbf{w}}_2)\\
&\quad= \sum_{i=1}^{k}\left[\left(\sum_{l=1}^{k}\mu_{il}(\tilde{\mathbf{x}})\widetilde{\mathbf{w}}_1^T C_l \widetilde{\mathbf{w}}_1\right)\left(\sum_{l=1}^{k}\mu_{il}(\tilde{\mathbf{x}})\widetilde{\mathbf{w}}_2^T C_l \widetilde{\mathbf{w}}_2\right) - \left(\sum_{l=1}^{k}\mu_{il}(\tilde{\mathbf{x}})\widetilde{\mathbf{w}}_1^T C_l \widetilde{\mathbf{w}}_2\right)^2\right]\\
&\quad+\sum_{l=1}^{k}\left[\left(\sum_{(i,j)\in J}\mu_{lij}(\tilde{\mathbf{x}})\widetilde{\mathbf{w}}_1^T C_{ij}\widetilde{\mathbf{w}}_1\right)\left(\sum_{(i,j)\in J}\mu_{lij}(\tilde{\mathbf{x}})\widetilde{\mathbf{w}}_2^T C_{ij}\widetilde{\mathbf{w}}_2\right)\right.\\
&\quad\left.-\left(\sum_{(i,j)\in J}\mu_{lij}(\tilde{\mathbf{x}})\widetilde{\mathbf{w}}_1^T C_{ij}\widetilde{\mathbf{w}}_2\right)^2\right]\\
&\quad=\sum_{l=1}^{k}\left((\widetilde{\mathbf{w}}_1^T C_l \widetilde{\mathbf{w}}_{\mathbf{1}})(\widetilde{\mathbf{w}}_2^T C_l \widetilde{\mathbf{w}}_{\mathbf{2}}) - (\widetilde{\mathbf{w}}_1^T C_l \widetilde{\mathbf{w}}_{\mathbf{2}})^2\right)\\
&\quad+\frac{1}{\sqrt{2}}\sum_{(i,j)\in J}\left((\widetilde{\mathbf{w}}_1^T C_{ij}\widetilde{\mathbf{w}}_1)(\widetilde{\mathbf{w}}_2^T C_{ij}\widetilde{\mathbf{w}}_{\mathbf{2}}) - (\widetilde{\mathbf{w}}_1^T C_{ij}\widetilde{\mathbf{w}}_{\mathbf{2}})^2\right) = \sum_{l=1}^{k}\left(\mathbf{w}_{1l}^2\mathbf{w}_{2l}^2 - (\mathbf{w}_{1l}^T\mathbf{w}_{2l})^2\right)\\
&\quad+\frac{1}{\sqrt{2}}\sum_{(i,j)\in J}\left((2\mathbf{w}_{1i}^T\mathbf{w}_{1j})(2\mathbf{w}_{2i}^T\mathbf{w}_{2j}) - (\mathbf{w}_{1i}^T\mathbf{w}_{2j} + \mathbf{w}_{1j}^T\mathbf{w}_{2i})^2\right),
\end{aligned}
$$

which is the statement. □

*Example 2* Let $n = 3,\ k = 2$, then the dimension of the Stiefel manifold $M_{3,2}$ is equal to

$$kn - \frac{k(k+1)}{2} = 6 - \frac{2\cdot 3}{2} = 3.$$

Let $S^2$ be the hypersphere in $R^3$, then the restrictions for the tangent vectors $\widetilde{\mathbf{w}} = (\mathbf{w}_1^T, \mathbf{w}_2^T)^T \in TM_{3,2}(\tilde{\mathbf{x}})$ of the Stiefel manifold $M_{3,2}$ at the point

$$\tilde{\mathbf{x}} = (\mathbf{x}_1^T, \mathbf{x}_2^T)^T, \qquad \mathbf{x}_1, \mathbf{x}_2 \in S^2,$$

are as follows:

$$
\begin{aligned}
\mathbf{x}_l^T\mathbf{w}_l = 0, \qquad l = 1, 2,\\
\mathbf{x}_1^T\mathbf{w}_2 + \mathbf{x}_2^T\mathbf{w}_1 = 0.
\end{aligned}
$$

In order to determine the sectional curvatures of the Stiefel manifolds $M_{3,2}$, the corresponding orthonormal pairs of vectors, representing some two-dimensional subspaces of the given tangent space, have to be given.

At the point $\tilde{\mathbf{x}}$, the tangent plane $TM_{3,2}(\tilde{\mathbf{x}})$ can be spanned by the orthonormal vectors

$$
\begin{aligned}
\widetilde{\mathbf{w}}_1 &= \left(-\frac{1}{\sqrt{2}}\mathbf{x}_2^T, \frac{1}{\sqrt{2}}\mathbf{x}_1^T\right)^T = \frac{1}{\sqrt{2}}\left(-\mathbf{x}_2^T, \mathbf{x}_1^T\right)^T,\\
\widetilde{\mathbf{w}}_2 &= \left(\frac{1}{\sqrt{2}}\mathbf{x}_3^T, \frac{1}{\sqrt{2}}\mathbf{x}_3^T\right)^T = \frac{1}{\sqrt{2}}\left(\mathbf{x}_3^T, \mathbf{x}_3^T\right)^T,\\
\widetilde{\mathbf{w}}_3 &= \left(\frac{1}{\sqrt{2}}\mathbf{x}_3^T, -\frac{1}{\sqrt{2}}\mathbf{x}_3^T\right)^T = \frac{1}{\sqrt{2}}\left(\mathbf{x}_3^T, -\mathbf{x}_3^T\right)^T,
\end{aligned}
$$

where $\mathbf{x}_1^T\mathbf{x}_3 = 0,\ \mathbf{x}_2^T\mathbf{x}_3 = 0$ and $\mathbf{x}_3^T\mathbf{x}_3 = 1$.

By using formula (28) and the fact that the sectional curvature does not depend on the choice of the vectors belonging to a given two-dimensional subspace of the tangent space, we have that

$$\begin{aligned}
K(\tilde{\mathbf{x}}, \widetilde{\mathbf{w}}_1, \widetilde{\mathbf{w}}_2) &= \mathbf{w}_{11}^2 \mathbf{w}_{21}^2 - \left(\mathbf{w}_{11}^T \mathbf{w}_{21}\right)^2 + \mathbf{w}_{12}^2 \mathbf{w}_{22}^2 - \left(\mathbf{w}_{12}^T \mathbf{w}_{22}\right)^2 \\
&\quad + \frac{1}{\sqrt{2}}\Big((2\mathbf{w}_{11}^T \mathbf{w}_{12})(2\mathbf{w}_{21}^T \mathbf{w}_{22}) \\
&\quad - (\mathbf{w}_{11}^T \mathbf{w}_{22} + \mathbf{w}_{12}^T \mathbf{w}_{21})^2\Big) = \frac{1}{2} > 0, \\
K(\tilde{\mathbf{x}}, \widetilde{\mathbf{w}}_2, \widetilde{\mathbf{w}}_3) &= \mathbf{w}_{21}^2 \mathbf{w}_{31}^2 - \left(\mathbf{w}_{21}^T \mathbf{w}_{31}\right)^2 + \mathbf{w}_{22}^2 \mathbf{w}_{32}^2 - \left(\mathbf{w}_{22}^T \mathbf{w}_{32}\right)^2 \\
&\quad + \frac{1}{\sqrt{2}}\Big((2\mathbf{w}_{21}^T \mathbf{w}_{22})(2\mathbf{w}_{31}^T \mathbf{w}_{32}) \\
&\quad - (\mathbf{w}_{21}^T \mathbf{w}_{32} + \mathbf{w}_{22}^T \mathbf{w}_{31})^2\Big) \\
&= \frac{1}{\sqrt{2}}\left(2\left(\frac{1}{2}\right) \cdot 2\left(-\frac{1}{2}\right) - \left(-\frac{1}{2} + \frac{1}{2}\right)\right) = -\frac{1}{\sqrt{2}} < 0. \\
K(\tilde{\mathbf{x}}, \widetilde{\mathbf{w}}_1, \widetilde{\mathbf{w}}_3) &= \mathbf{w}_{11}^2 \mathbf{w}_{31}^2 - \left(\mathbf{w}_{11}^T \mathbf{w}_{31}\right)^2 + \mathbf{w}_{12}^2 \mathbf{w}_{32}^2 - \left(\mathbf{w}_{12}^T \mathbf{w}_{32}\right)^2 \\
&\quad + \frac{1}{\sqrt{2}}\Big((2\mathbf{w}_{11}^T \mathbf{w}_{12})(2\mathbf{w}_{31}^T \mathbf{w}_{32}) \\
&\quad - (\mathbf{w}_{11}^T \mathbf{w}_{32} + \mathbf{w}_{12}^T \mathbf{w}_{31})^2\Big) = \frac{1}{2} > 0.
\end{aligned}$$

## 6 Minimization of sectional curvatures on Euclidean submanifolds with the induced Riemannian metric

In order to determine the minimal sectional curvature of the manifold $M[\mathbf{h}]$ with the induced Riemannian metric at an arbitrary given point $\mathbf{x}_0 \in M[\mathbf{h}]$, the following nonlinear global optimization problem has to be solved:

$$\begin{gathered}
\min K_{M[\mathbf{h}]}(\mathbf{x}_0, \mathbf{w}_1, \mathbf{w}_2) \\
\mathbf{w}_1^T \mathbf{w}_1 = 1, \qquad \mathbf{w}_2^T \mathbf{w}_2 = 1, \qquad \mathbf{w}_1^T \mathbf{w}_2 = 0, \qquad \mathbf{w}_1, \mathbf{w}_2 \in R^n,
\end{gathered} \tag{29}$$

which is equivalent to

$$\begin{gathered}
\min K_{M[\mathbf{h}]}(\mathbf{x}_0, \mathbf{w}) \\
\mathbf{w} = (\mathbf{w}_1^T, \mathbf{w}_2^T)^T \in M_{k,2},
\end{gathered} \tag{30}$$

By (23),

$$\begin{gathered}
S(\mathbf{w}) = \begin{pmatrix} \nabla_{\mathbf{w}} K_{M[\mathbf{h}]}(\mathbf{x}_0, \mathbf{w}) \begin{pmatrix} \mathbf{w}_1 \\ \mathbf{0} \end{pmatrix} I_n & \frac{1}{2} \nabla_{\mathbf{w}} K_{M[\mathbf{h}]}(\mathbf{x}_0, \mathbf{w}) \begin{pmatrix} \mathbf{w}_2 \\ \mathbf{w}_1 \end{pmatrix} I_n \\ \frac{1}{2} \nabla_{\mathbf{w}} K_{M[\mathbf{h}]}(\mathbf{x}_0, \mathbf{w}) \begin{pmatrix} \mathbf{w}_2 \\ \mathbf{w}_1 \end{pmatrix} I_n & \nabla_{\mathbf{w}} K_{M[\mathbf{h}]}(\mathbf{x}_0, \mathbf{w}) \begin{pmatrix} \mathbf{0} \\ \mathbf{w}_2 \end{pmatrix} I_n \end{pmatrix}, \\
\mathbf{w} \in \left(\mathbf{w}_1^T, \mathbf{w}_2^T\right)^T \in M_{k,2}.
\end{gathered} \tag{31}$$

and by Theorem 4, we have the first-order and second-order optimality conditions of problem (31).

**Theorem 6** *If the point* $\tilde{\mathbf{w}} \in M_{k,2}$ *is a (strict) local minimum of problem (30), then*

$$\nabla_{\mathbf{w}} K_{M[\mathbf{h}]}(\mathbf{x}_0, \tilde{\mathbf{w}}) = \tilde{\mathbf{w}}^T S(\tilde{\mathbf{w}}), \tag{32}$$

*and*

$$\Big(H_{\mathbf{w}} K_{M[\mathbf{h}]}(\mathbf{x}_0, \tilde{\mathbf{w}}) - S(\tilde{\mathbf{w}})\Big)_{|TM_{k,2}(\tilde{\mathbf{w}})}, \tag{33}$$

*is a positive semidefinite (definite) matrix.*

*If* $\mathcal{C} \subseteq M_{k,2}$ *is an open geodesic convex set, and there exists a point* $\tilde{\mathbf{w}} \in \mathcal{C}$ *such that (32) holds, and*

$$\big(H_{\mathbf{w}} K_{M[\mathbf{h}]}(\mathbf{x}_0, \mathbf{w}) - S(\mathbf{w})\big)_{|TM_{k,2}(\mathbf{w})}, \qquad \mathbf{w} \in \mathcal{C}, \tag{34}$$

*are positive semidefinite (definite) matrices, then the point* $\tilde{\mathbf{w}}$ *is a (strict) global minimum.*

Based on Theorems 2 and 6, let us determine the maximal and minimal sectional curvatures on an ellipsoid.

**Theorem 7** *Consider the feasible set*

$$M[h] = \left\{ \mathbf{x} \in R^n \mid h(\mathbf{x}) = \frac{1}{2}\mathbf{x}^T A\mathbf{x} - \frac{1}{2}c = 0 \right\},$$

*where A is a symmetric and positive definite* $n \times n$ *matrix and* $c \neq 0$*. Then,*

$$\max K_{M[h]}(\mathbf{w}_1, \mathbf{w}_2) = \frac{\lambda_1 \lambda_2}{\lambda_n c}$$

*and*

$$\min K_{M[h]}(\mathbf{w}_1, \mathbf{w}_2) = \frac{\lambda_n \lambda_{n-1}}{\lambda_1 c},$$

*where* $\lambda_1, \lambda_2$ *and* $\lambda_{n-1}, \lambda_n$ *are the two largest and smallest eigenvalues of the matrix A, respectively.*

*Proof* By differentiating the function $h$, we have that

$$\nabla h(\mathbf{x}) = \mathbf{x}^T A, \quad \| \nabla h(\mathbf{x}) \| = (\mathbf{x}^T AA\mathbf{x})^{1/2}, \quad Hh(\mathbf{x}) = A, \qquad \mathbf{x} \in M[h],$$

and since the unique normal vectors are $\dfrac{\nabla h(\mathbf{x})}{\| \nabla h(\mathbf{x}) \|}$, $\mathbf{x} \in M[h]$,

$$\mu(\mathbf{x}) = \frac{1}{\| \nabla h(\mathbf{x}) \|}, \qquad \mathbf{x} \in M[h].$$

Let $\mathbf{w}_1$ and $\mathbf{w}_2$ be an orthonormal pair of tangent vectors with respect to the Euclidean metric at an arbitrary point $\mathbf{x}_0$ of $M[h]$, then by Theorem 2,

$$K_{M[h]}(\mathbf{x}_0, \mathbf{w}_1, \mathbf{w}_2) = \frac{1}{\| \nabla h(\mathbf{x}_0) \|^2}\Big((\mathbf{w}_1^T A\mathbf{w}_1)(\mathbf{w}_2^T A\mathbf{w}_2) - (\mathbf{w}_1^T A\mathbf{w}_2)^2\Big). \tag{35}$$

Thus,

$$\nabla_{\mathbf{w}} K_{M[h]}(\mathbf{x}_0, \mathbf{w}_1, \mathbf{w}_2) = \frac{2}{\| \nabla h(\mathbf{x}_0) \|^2} \Big(\mathbf{w}_1^T A(\mathbf{w}_2^T A\mathbf{w}_2) - \mathbf{w}_2^T A(\mathbf{w}_1^T A\mathbf{w}_2), \mathbf{w}_2^T A(\mathbf{w}_1^T A\mathbf{w}_1) - \mathbf{w}_1^T A(\mathbf{w}_1^T A\mathbf{w}_2)\Big), \quad \mathbf{w} \in M_{n-1,2}, \tag{36}$$

$$S(\mathbf{w}) = \begin{pmatrix} \nabla_{\mathbf{w}} K_{M[h]}(\mathbf{x}_0, \mathbf{w}) \begin{pmatrix} \mathbf{w_1} \\ \mathbf{0} \end{pmatrix} I_n & \frac{1}{2}\nabla_{\mathbf{w}} K_{M[h]}(\mathbf{x}_0, \mathbf{w}) \begin{pmatrix} \mathbf{w_2} \\ \mathbf{w_1} \end{pmatrix} I_n \\ \frac{1}{2}\nabla_{\mathbf{w}} K_{M[h]}(\mathbf{x}_0, \mathbf{w}) \begin{pmatrix} \mathbf{w_2} \\ \mathbf{w_1} \end{pmatrix} I_n & \nabla_{\mathbf{w}} K_{M[h]}(\mathbf{x}_0, \mathbf{w}) \begin{pmatrix} \mathbf{0} \\ \mathbf{w}_2 \end{pmatrix} I_n \end{pmatrix}$$

$$= \frac{2}{\| \nabla h(\mathbf{x}_0) \|^2} \begin{pmatrix} \Big((\mathbf{w}_1^T A\mathbf{w}_1)(\mathbf{w}_2^T A\mathbf{w}_2) - (\mathbf{w}_1^T A\mathbf{w}_2)^2\Big) I_n & 0 \\ 0 & \Big((\mathbf{w}_1^T A\mathbf{w}_1)(\mathbf{w}_2^T A\mathbf{w}_2) - (\mathbf{w}_1^T A\mathbf{w}_2)^2\Big) I_n \end{pmatrix}, \quad \mathbf{w} \in M_{n-1,2}. \tag{37}$$

The first-order optimality conditions of the minimization or maximization of sectional curvatures on $M[h]$ with the induced Riemannian metric are as follows:

$$\mathbf{w}_1^T A(\mathbf{w}_2^T A\mathbf{w}_2) - \mathbf{w}_2^T A(\mathbf{w}_1^T A\mathbf{w}_2) = \Big((\mathbf{w}_1^T A\mathbf{w}_1)(\mathbf{w}_2^T A\mathbf{w}_2) - (\mathbf{w}_1^T A\mathbf{w}_2)^2\Big)\mathbf{w}_1^T, \tag{38}$$

$$\mathbf{w}_2^T A(\mathbf{w}_1^T A\mathbf{w}_1) - \mathbf{w}_1^T A(\mathbf{w}_1^T A\mathbf{w}_2) = \Big((\mathbf{w}_1^T A\mathbf{w}_1)(\mathbf{w}_2^T A\mathbf{w}_2) - (\mathbf{w}_1^T A\mathbf{w}_2)^2\Big)\mathbf{w}_2^T, \quad \mathbf{w} \in M_{n-1,2}. \tag{39}$$

The pairs of eigenvectors of $A$ fulfil equations (38) and (39), thus, they are of stationary points of the sectional curvature optimization problem. A consequence is that if $\mathbf{w}_1$ and $\mathbf{w}_2$ are the orthonormal eigenvectors related to the two largest eigenvalues $\lambda_1$ and $\lambda_2$ or the two smallest eigenvalues $\lambda_n$ and $\lambda_{n-1}$, then we obtain the maximum value $\frac{1}{\| \nabla h(\mathbf{x}_0) \|^2} \lambda_1\lambda_2$ or the minimum value $\frac{1}{\| \nabla h(\mathbf{x}_0) \|^2} \lambda_n\lambda_{n-1}$ of the sectional curvatures at the point $\mathbf{x}_0$.

Since $\max_{\mathbf{x}\in M[h]} \| \nabla h(\mathbf{x}) \|^2$ and $\min_{\mathbf{x}\in M[h]} \| \nabla h(\mathbf{x}) \|^2$ are $\lambda_1 c$ and $\lambda_n c$, respectively, we have proved the statement. □

We remark that the maximal and minimal sectional curvatures on $M[h]$ with the induced Riemannian metric are proportional with the condition number of the matrix $A$.

Let $\lambda_n = 0.4$, $\lambda_{n-1} = 0.7, \ldots, \lambda_2 = 1.7$, $\lambda_1 = 2$ be the eigenvalues of a matrix $A$. Then, the inequality of the sphere theorem fulfils, thus, by the sphere theorem, the ellipsoid $M[h]$ is homeomorphic to a sphere.

## 7 Concluding remarks

In the paper, the function of sectional curvature is explicitly expressed with the first and second derivatives of the problem's functions in the case of submanifolds determined by equality constraints in the $n$-dimensional Euclidean space endowed with the induced Riemannian metric and the minimization problem of sectional curvature is formulated as a global minimization one on a Stiefel manifolds.

The sectional curvature of an arbitrary point of the manifold is not a function on the manifold, but it is a continuous function on the two-dimensional vector subspaces of the given tangent space. It follows that the sectional curvatures on a compact subset of the manifold are bounded.

Some open questions are as follows:

- characterization of the Stiefel manifolds based on sectional curvatures;
- analysis of special equality constraints based on sectional curvatures (e.g., Theorem 7);
- solution of the minimization problem of sectional curvatures on special submanifolds;
- development of methods for solving the minimization problem of sectional curvatures on submanifolds;
- generalization of Schur theorem.

## References

1. Bishop, R.L., Crittenden, R.J.: Geometry of Manifolds. Academic Press, New York (1964)
2. do Carmo, M.P.: Riemannian Geometry. Birkhäuser, Boston (1992)
3. Edelman, A., Arias, T.A., Smith, S.T.: The geometry of algorithms with orthogonality constraints. SIAM J. Matrix Anal. Appl. **20**, 303–353 (1998)
4. Hicks, N.J.: Notes on differential geometry. Van Nostrand Publishing Company, Princeton, New Jersey (1965)
5. Luenberger, D.G.: Introduction to linear and nonlinear programming. Addison-Wesley Publishing Company, Reading (1973)
6. Oprea, T.: Optimization methods on Riemannian submanifolds. Anal. Univ. Bucureşti Mat. Anul LIV, 127–136 (2005)
7. Rapcsák T.: Smooth Nonlinear Optimization in $R^n$. Kluwer Academic Publishers (1997)
8. Rapcsák, T.: On minimization of sums of heterogeneous quadratic function on Stiefel manifolds. In: Migdalas, A., Pardalos, P., Varbrand, P. (eds.) From Local to Global Optimization, pp. 277–290. Kluwer Academic Publishers (2001)
9. Rapcsák, T.: On minimization on Stiefel manifolds. Eur. J. Operat. Res. **143**, 365–376 (2002)
10. Spivak, M.: A Comprehensive Introduction to Differential Geometry I-V. Publish or Perish, Berkeley (1979)
11. Stiefel, E.: Richtungsfelder und Fernparallelismus in $n$-dimensional Mannigfaltigkeiten. Comm. Math. Helve. **8**, 305–353 (1935/6)
12. Udriste, C.: Convex Functions and Optimization Methods on Riemannian Manifolds. Kluwer Academic Publishers, Dordrecht (1994)