Arch. Math. 114 (2020), 215–225

0003-889X/20/020215-11
published online October 4, 2019
https://doi.org/10.1007/s00013-019-01382-x

Archiv der Mathematik

Characterization of solution sets of convex optimization problems in Riemannian manifolds

A. Barani and S. Hosseini

Abstract. In this paper, a characterization of the solution sets of convex smooth optimization programmings on Riemannian manifolds, in terms of the Riemannian gradients of the cost functions, is obtained.

Mathematics Subject Classification. 58C99, 58C05.

Keywords. Riemannian optimization, Convex functions, Solution sets.

1. Introduction. This paper is concerned with a characterization of the solution sets of convex optimization programmings on Riemannian manifolds. Characterizations of the solution sets of the nonlinear programming problems with multiple solutions, provided that one minimizer is known, play an important role in many fields, such as optimization problems, variational inequalities, and equilibrium problems. Mangasarian [16] presented several characterizations of the solution sets for differentiable convex programs on linear spaces and applied them to study monotone linear complementarity problems. Further investigation has been done by Burke and Ferris [3]. In the last decade, Mangasarian type characterizations were derived for several smooth and nonsmooth convex or generalized convex problems on linear spaces; see [12,13,23] and references therein.

Extensions of concepts and techniques from Euclidean spaces to Riemannian manifolds are natural and lead to successful tools in optimization, therefore such topics with practical and theoretical purposes have been the subject of several research papers. Udriste [22] and Rapscák [17] introduced the theory of convex functions on Riemannian manifolds motivated by the fact that some constrained optimization problems can be seen as unconstrained ones from the Riemannian geometry point of view. In addition, another advantage is that optimization problems with nonconvex objective functions can be written as convex optimization problems by endowing the space with an appropriate Riemannian metric (see Example 2.1). Recently, a number of important results have been obtained on various aspects of optimization theory and applications on Riemannian manifolds, which introduced several important techniques and methods for existence of solutions of optimization problems on Riemannian manifolds; see [1,2,8,20].

The purpose of this paper is to present a simple characterization of solution sets of convex optimization problems on Riemannian manifolds in terms of the Riemannian gradient of the cost function. To the best of our knowledge, it has not been given before and to formulate and prove this result on Riemannian manifolds, we need to use several tools and techniques from Riemannian geometry. One of the applications of this characterization is for the problem of minimizing convex quadratic functions defined on a convex subset of a sphere, which arises in solving fixed point theorems, surjectivity theorems, existence theorems for complementarity problems, and variational inequalities by calculating the scalar derivatives; for more details about these theorems and their applications see [6] and references therein. In particular, some existence theorems could be reduced to optimizing a quadratic function on a convex subset of the sphere. Moreover, the minimization problems of quadratic functions defined on the sphere occur as subproblems in methods of nonlinear programming; see [18].

2. Preliminaries. In this section, we introduce some fundamental properties and notations of Riemannian manifolds. These basic facts can be found in any introductory book on Riemannian geometry; see for example [19]. Throughout this paper, M is an n-dimensional Riemannian manifold with a Riemannian metric $\langle \cdot, \cdot \rangle_x$ on the tangent space $T_xM \cong \mathbb{R}^n$ for every $x \in M$. The corresponding norm is denoted by $\| \cdot \|$. Let us recall that the length of a piecewise C^1 curve $\gamma : [a, b] \to M$ is defined by

$$L(\gamma) = \int_a^b \|\gamma'(t)\| dt.$$

By minimizing the length functional over the set of all piecewise C^1 curves with $\gamma(0) = x$ and $\gamma(1) = y$ for $x, y \in M$, we obtain a Riemannian distance on M denoted by $d(x, y)$. The space of vector fields on M is denoted by $\mathcal{X}(M)$ and ∇ is the Levi-Civita connection associated to M. A geodesic is a smooth curve γ satisfying the equation $\nabla_{\gamma'(t)}\gamma'(t) = 0$. The exponential mapping $\exp : \tilde{T}M \to M$ is defined as $\exp(v) = \gamma(1)$, where γ is the geodesic defined by its starting point x and the velocity $\gamma'(0) = v$ at x and $\tilde{T}M$ is an open neighborhood in TM. The restriction of exp to T_xM in $\tilde{T}M$ is denoted by $\exp_x$ for every $x \in M$. For a minimizing geodesic $\gamma : [0, l] \to M$ connecting x to y in M, and for a vector $v \in T_xM$, there is a unique parallel vector field P along γ such that $P(0) = v$, this is called the parallel translation of v along γ. The mapping $T_xM \ni v \mapsto P(1) \in T_yM$ is a linear isometry from T_xM onto T_yM. This map is denoted by P_x^y.

The Riemannian metric induces a map $f \mapsto \operatorname{grad} f \in \mathcal{X}(M)$ which associates to each differentiable function f at $x \in M$, its gradient via the rule

$$\langle \operatorname{grad} f(x), v\rangle_x = df(v) = \frac{d}{dt} f(\exp_x(tv))|_{t=0}, \quad v \in T_xM.$$

The Riemannian Hessian of f at a point $x \in M$ is the linear mapping

$$\operatorname{Hess} f(x) : T_xM \to T_xM$$

defined by

$$\operatorname{Hess} f(x)[v] = \nabla_v \operatorname{grad} f$$

for every $v \in T_xM$. Note that $\operatorname{Hess} f(x)$ satisfies

$$\operatorname{Hess} f(x)(v, v) = \frac{d^2}{dt^2} f(\exp_x(tv))|_{t=0}, \quad v \in T_xM,$$

and this formula fully defines $\operatorname{Hess} f(x)$.

A subset S of a Riemannian manifold is called convex if any two points x, $y \in S$ can be joined by a unique minimizing geodesic (denoted by γ_{xy}) which lies entirely in S. Note that there is little consistency in the meanings attached to the terms "convex set" and "strongly convex set" (see page 105 in [9] and page 2488 in [15] and references therein).

It is known that $\exp_x^{-1}$ is well-defined on every convex set S, $d(x, y) = ||\exp_x^{-1}(y)||$ for every $x, y \in S$, and

$$\gamma_{xy}(t) = \exp_x(t \exp_x^{-1} y) \quad \text{for all } t \in [0, 1];$$

see [11]. Let S be a nonempty convex subset of M, a function $f : S \to \mathbb{R}$ is said to be convex if for every x, $y \in S$ and every $t \in [0, 1]$,

$$f(\gamma_{xy}(t)) \le (1 - t)f(x) + tf(y).$$

The following example illustrates a nonconvex function which can be written as a convex function on a Riemannian manifold with an appropriate metric (see [4]).

Example 2.1. The function $f : \mathbb{R}^2 \to \mathbb{R}$ defined by

$$f(x) = e^{x_1}(\cosh(x_2) - 1), \quad x = (x_1, x_2),$$

is not convex. Endowing $\mathbb{R}^2$ with the metric $g(x) := \operatorname{diag}(1, e^{2x_1})$, we obtain a Riemannian manifold M_g. The Hessian matrix

$$\operatorname{Hess} f(x) = \operatorname{diag}(e^{x_1}(\cosh(x_2) - 1), e^{x_1}\cosh(x_2) + e^{3x_1}(\cosh(x_2) - 1)),$$

is positive semidefinite, therefore f is convex on M_g.

3. Characterization of the solution sets. Our aim is to characterize the solution set of the following optimization problem

$$\min_{x \in S} f(x), \tag{1}$$

where $S \subseteq M$ is a convex subset of M and f is a twice continuously differentiable convex function on some open convex set containing S. We denote the solution set of the optimization problem (1) by

$$\bar{S} = \operatorname*{argmin}_{x \in S} f(x),$$

and assume that $\bar{S} \neq \emptyset$. If $\bar{x} \in \bar{S}$, then

$$\bar{S} = \{x \in S : f(x) = f(\bar{x})\},$$

and $\bar{S}$ is a convex subset of S. The following theorem is a generalization of [7, Proposition 15] from the sphere S^n to a general setting which describes the relation between the solution sets of (1) and a variational inequality.

Theorem 3.1. *Let $S \subseteq M$ be a convex subset of M and f be a twice continuously differentiable convex function on some open convex set containing S. Then $\bar{x} \in \bar{S}$ if and only if*

$$\langle \operatorname{grad} f(\bar{x}), \exp_{\bar{x}}^{-1} y \rangle_{\bar{x}} \geq 0 \quad \textit{for all } y \in S. \tag{2}$$

Proof. Let $\bar{x} \in \bar{S}$, $y \in S$, and $\gamma_{\bar{x}y}$ be the minimal geodesic connecting $\bar{x}$ and y. By convexity of S, $f(\gamma_{\bar{x}y}(t)) \geq f(\bar{x})$ for all $t \in [0,1]$. Therefore

$$\langle \operatorname{grad} f(\bar{x}), \exp_{\bar{x}}^{-1} y \rangle_{\bar{x}} = \lim_{t \to 0} \frac{f(\gamma_{\bar{x}y}(t)) - f(\bar{x})}{t} \geq 0.$$

Now suppose that (2) holds. By convexity of f, we have

$$f(y) - f(\bar{x}) \geq \langle \operatorname{grad} f(\bar{x}), \gamma'_{\bar{x}y}(0) \rangle_{\bar{x}} = \langle \operatorname{grad} f(\bar{x}), \exp_{\bar{x}}^{-1} y \rangle_{\bar{x}} \geq 0 \quad \text{for all } y \in S.$$

□

Now we present a characterization for the solution set of a convex optimization problem on a convex subset of a Riemannian manifold which is our main result.

Theorem 3.2. *Let $S \subseteq M$ be a convex subset of M, f be a twice continuously differentiable convex function on some open convex set containing S, and $\bar{x} \in \bar{S}$. Then*

$$\bar{S} = \{x \in S : \langle \operatorname{grad} f(\bar{x}), \exp_{\bar{x}}^{-1} x \rangle_{\bar{x}} = 0,\ P_x^{\bar{x}}[\operatorname{grad} f(x)] = \operatorname{grad} f(\bar{x})\}. \tag{3}$$

Proof. We denote the right hand side of (3) by S^*. On the contrary, we assume that $x \in S^* \setminus \bar{S}$. By convexity of f,

$$\langle \operatorname{grad} f(x), \exp_x^{-1}(\bar{x}) \rangle_{\bar{x}} \leq f(\bar{x}) - f(x) < 0.$$

Now, by properties of the parallel translation, we get

$$\begin{aligned} \langle \operatorname{grad} f(x), \exp_x^{-1}(\bar{x}) \rangle_x &= \langle P_x^{\bar{x}}[\operatorname{grad} f(x)], P_x^{\bar{x}}[\exp_x^{-1} \bar{x}] \rangle_{\bar{x}} \\ &= \langle \operatorname{grad} f(\bar{x}), -\exp_{\bar{x}}^{-1} x \rangle_{\bar{x}} = 0, \end{aligned}$$

which is a contradiction.

For the converse, let $x \in \bar{S}$ and $\gamma_{\bar{x}x}(t)$ be the minimal geodesic connecting $\bar{x}$ and x. Since $\bar{S}$ is convex, this geodesic lies entirely in $\bar{S}$ so $f(\gamma_{\bar{x}x}(t)) = f(\bar{x})$ for all $t \in [0,1]$. Therefore

$$\langle \operatorname{grad} f(\bar{x}), \exp_{\bar{x}}^{-1} x \rangle_{\bar{x}} = \lim_{t \to 0} \frac{f(\gamma_{\bar{x}x}(t)) - f(\bar{x})}{t} = 0.$$

Similarly, we have $\langle \operatorname{grad} f(x), \exp_x^{-1} \bar{x} \rangle_x = 0$. These two equations and properties of the parallel transport imply

$$\langle \operatorname{grad} f(\bar{x}) - P_x^{\bar{x}}[\operatorname{grad} f(x)], \exp_{\bar{x}}^{-1} x \rangle_{\bar{x}} = 0. \tag{4}$$

Now, we define $F : [0,1] \to T_{\bar{x}}M$ as follows

$$F(t) = P_{\gamma_{x\bar{x}}(t)}^{\bar{x}}[\operatorname{grad} f(\gamma_{x\bar{x}}(t))].$$

Note that $\gamma'_{x\bar{x}}(t) = P_x^{\gamma_{x\bar{x}}(t)}[\exp_x^{-1} \bar{x}]$ and

$$F(t-s) = P_{\gamma_{x\bar{x}}(t)}^{\bar{x}} P_{\gamma_{x\bar{x}}(t-s)}^{\gamma_{x\bar{x}}(t)}[\operatorname{grad} f(\gamma_{x\bar{x}}(t-s))] \quad \text{for all } t \in [0,1],$$

hence

$$\begin{aligned} F'(t) &= -\frac{d}{ds} F(t-s)|_{s=0} \\ &= -P_{\gamma_{x\bar{x}}(t)}^{\bar{x}} \frac{d}{ds} P_{\gamma_{x\bar{x}}(t-s)}^{\gamma_{x\bar{x}}(t)}[\operatorname{grad} f(\gamma_{x\bar{x}}(t-s))]|_{s=0} \\ &= P_{\gamma_{x\bar{x}}(t)}^{\bar{x}}[\operatorname{Hess} f(\gamma_{x\bar{x}}(t))(\gamma'_{x\bar{x}}(t))]. \end{aligned}$$

Since F is C^1 and by the previous equality, we get

$$\begin{aligned} \operatorname{grad} f(\bar{x}) - P_x^{\bar{x}}[\operatorname{grad} f(x)] &= F(1) - F(0) = \int_0^1 F'(t)dt \\ &= \int_0^1 P_{\gamma_{x\bar{x}}(t)}^{\bar{x}} \operatorname{Hess} f(\gamma_{x\bar{x}}(t))(\gamma'_{x\bar{x}}(t))dt \\ &= \int_0^1 P_{\gamma_{x\bar{x}}(t)}^{\bar{x}} \operatorname{Hess} f(\gamma_{x\bar{x}}(t))(P_{\bar{x}}^{\gamma_{x\bar{x}}(t)}(\exp_{\bar{x}}^{-1} x))dt \\ &= \left(\int_0^1 P_{\gamma_{x\bar{x}}(t)}^{\bar{x}} \operatorname{Hess} f(\gamma_{x\bar{x}}(t) \right) P_{\bar{x}}^{\gamma_{x\bar{x}}(t)}(\cdot)dt) \exp_{\bar{x}}^{-1} x \\ &= A \exp_{\bar{x}}^{-1} x, \end{aligned} \tag{5}$$

where

$$\int_0^1 P_{\gamma_{x\bar{x}}(t)}^{\bar{x}} \operatorname{Hess} f(\gamma_{x\bar{x}}(t)) P_{\bar{x}}^{\gamma_{x\bar{x}}(t)}(\cdot)dt : T_{\bar{x}}M \to T_{\bar{x}}M$$

is a linear map. We claim that this linear map is positive semidefinite and therefore corresponds to a positive semidefinite matrix A. To prove the claim, note that $\operatorname{Hess} f(\gamma_{x\bar{x}}(t))$ is a positive semidefinite linear map, hence

$$\langle w, P^{\bar{x}}_{\gamma_{x\bar{x}}(t)}[\mathrm{Hess}\, f(\gamma_{x\bar{x}}(t))(P^{\gamma_{x\bar{x}}(t)}_{\bar{x}}(w))]\rangle_{\bar{x}}$$
$$= \langle P^{\gamma_{x\bar{x}}(t)}_{\bar{x}}(w), \mathrm{Hess}\, f(\gamma_{x\bar{x}}(t))[P^{\gamma_{x\bar{x}}(t)}_{\bar{x}}(w)]\rangle_{\gamma_{x\bar{x}}(t)} \geq 0$$

for every $w \in T_{\bar{x}}M$, therefore by integration with respect to t, it implies that

$$\langle w, Aw\rangle_{\bar{x}} = \left\langle w, \int_0^1 P^{\bar{x}}_{\gamma_{x\bar{x}}(t)}[\mathrm{Hess}\, f(\gamma_{x\bar{x}}(t))(P^{\gamma_{x\bar{x}}(t)}_{\bar{x}}(w))]dt \right\rangle_{\bar{x}} \geq 0,$$

which proves our claim. Combining (4) and (5), we deduce that

$$\langle \mathrm{grad}\, f(\bar{x}) - P^{\bar{x}}_{x}[\mathrm{grad} f(x)], \exp^{-1}_{\bar{x}} x\rangle_{\bar{x}} = \langle A \exp^{-1}_{\bar{x}} x, \exp^{-1}_{\bar{x}} x\rangle_{\bar{x}} = 0. \tag{6}$$

Since A is symmetric positive semidefinite, it follows from (6) and [10, p. 431] that

$$\mathrm{grad}\, f(\bar{x}) - P^{\bar{x}}_{x}[\mathrm{grad}\, f(x)] = A \exp^{-1}_{\bar{x}} x = 0,$$

which completes the proof. □

The following corollary is an immediate consequence of Theorem 3.2.

Corollary 3.1. *Let $S \subseteq M$ be a convex subset of M, f be a twice continuously differentiable convex function on some open convex set containing S, and $\bar{x} \in \bar{S}$. Then*

(i) *the function $x \mapsto \|\mathrm{grad}\, f(x)\|$ is constant on $\bar{S}$.*
(ii) $\bar{S} = \tilde{S} = \{x \in S : \langle \mathrm{grad}\, f(\bar{x}), \exp^{-1}_{\bar{x}} x)\rangle_{\bar{x}} \leq 0, P^{\bar{x}}_{x}[\mathrm{grad}\, f(x)] = \mathrm{grad} f(\bar{x})\}$.

Proof. The first part is obtained by Theorem 3.2 and the isometric property of the parallel translation. For proving the second part, we have that the inclusion $\bar{S} \subseteq \tilde{S}$ holds by Theorem 3.2. For the converse, assume that $x \in \tilde{S}$. Since the parallel translation is an isometry and f is convex, we deduce that

$$f(\bar{x}) - f(x) \geq \langle \mathrm{grad}\, f(x), \exp^{-1}_{x} \bar{x}\rangle_{x} = \langle P^{\bar{x}}_{x}[\mathrm{grad}\, f(x)], P^{\bar{x}}_{x}[\exp^{-1}_{x} \bar{x}]\rangle_{\bar{x}}$$
$$= -\langle \mathrm{grad}\, f(\bar{x}), \exp^{-1}_{\bar{x}} x\rangle_{\bar{x}} \geq 0.$$

This shows that $f(\bar{x}) \geq f(x)$. Since $\bar{x}$ is the minimizer of f on S, we have $f(\bar{x}) = f(x)$, which proves that $\tilde{S} \subseteq \bar{S}$. □

Now, we present some examples to illustrate how our characterization of solution sets of optimization problems works in particular nontrivial settings of Riemannian manifolds.

Example 3.1. Let M be a Hadamard manifold and U be a nonempty open convex subset of M. Pick $x_0, y_0 \in U$ with $x_0 \neq y_0$. Choose $\varepsilon > 0$ such that $S := \bar{B}(x_0, \varepsilon) \subset U$, $A := \bar{B}(y_0, \varepsilon) \subset U$ are convex and $S \cap A = \emptyset$. Define the function $f : S \to \mathbb{R}$ by

$$f(x) := \frac{1}{2} d^2_A(x).$$

Note that the function $d^2_A(\cdot)$ is convex and twice continuously differentiable on U. Let $\bar{x} \in \bar{S}$ and $p := \pi_A(\bar{x})$, where $\pi_A(\bar{x})$ is the metric projection of $\bar{x}$ on A. We have that

$$\mathrm{grad} f(\bar{x}) = -\exp^{-1}_{\bar{x}} p, \quad \|\mathrm{grad} f(\bar{x})\| = d(\bar{x}, p) = d(\bar{x}, A).$$

Thus, by using Theorem 3.2, we deduce

$$\bar{S} = \{x \in S : \langle \exp_{\bar{x}}^{-1} p, \exp_{\bar{x}}^{-1} x\rangle_{\bar{x}} = 0, \operatorname{grad} f(x) = -P_{\bar{x}}^{x}[\exp_{\bar{x}}^{-1} p]\}.$$

Example 3.2. The general matrix rank minimization problem (RMP) expressed as

$$\min \operatorname{rank} X, \qquad X \in S \subset P_n,$$

where $X \in \mathbb{R}^{n\times n}$, P_n is the set of positive semidefinite $n \times n$ matrices, and S is a convex set, is computationally hard to solve. This problem arises in many areas such as control, system identification, statistics, signal processing, and computational geometry; see [5] and references therein. Rather than solving the RMP, one can use the function

$$\log\det(X + \delta I),$$

as a smooth surrogate for rank$X$ and instead solve the following problem

$$\min \log\det(X + \delta I), \quad X \in S,$$

where $\delta > 0$ can be interpreted as a small regularization constant; see [5]. Note that this surrogate is not convex on the linear space $\mathbb{R}^{n\times n}$. This application motivated us to consider the problem

$$\min \log\det(X), \qquad X \in S = \{X \in P_n^+ :\ 0 < A \le X\}, \tag{7}$$

which is not convex with respect to the Euclidean metric on $\mathbb{R}^{n\times n}$. Here P_n^+ is the set of positive definite $n \times n$ matrices and $A \in P_n^+$.

The set of symmetric positive definite matrices, as a Riemannian manifold, is the most studied example of manifolds of nonpositive curvature. The tangent space to P_n^+ at any of its points P is the space $T_P P_n^+ = \{P\} \times S_n$, where S_n is the space of symmetric $n \times n$ matrices. On each tangent space $T_P P_n^+$, the inner product is defined by

$$\langle A, B\rangle_P = \operatorname{tr}(P^{-1}AP^{-1}B).$$

The Riemannian distance between $P, Q \in P_n^+$ is given by

$$\operatorname{dist}(P, Q) = \left(\sum_{i=1}^{n} \ln^2(\lambda_i)\right)^{1/2},$$

where $\lambda_i,\ i = 1, ..., n$, are eigenvalues of $P^{-1}Q$. The exponential map

$$\exp_P : S_n \to P_n^+$$

is defined by

$$\exp_P(v) = P^{1/2}\exp(P^{-1/2}vP^{-1/2})P^{1/2}.$$

Moreover, if $P \in P_n^+$, then

$$\exp_P^{-1} : P_n^+ \to S_n$$

is defined by

$$\exp_P^{-1}(Q) = P\log(P^{-1}Q),$$

where log and exp denote the logarithm and exponential functions on the matrix space; for more details see [21]. The parallel transport along the unique geodesic connecting X and Y, is defined by

$$P_X^Y(Z) = (YX^{-1})^{1/2} Z (X^{-1}Y)^{1/2}.$$

Moreover, the Riemannian gradient of a function f defined on P_n^+ is given by using the Euclidean gradient, denoted by ∇f, using the following formula,

$$\operatorname{grad} f(X) = X \operatorname{symm}(\nabla f(X)) X,$$

where $\operatorname{symm}(\nabla f(X)) = 1/2(\nabla f(X) + \nabla f(X)^T)$. For $f(X) = \log\det(X)$, $\nabla f(X) = X^{-1}$, therefore $\operatorname{grad} f(X) = X$. First we claim that S is a convex subset of P_n^+. Assume that $X, Y \in S$, then $X \geq A$ and $Y \geq A$. The unique geodesic connecting these two points is defined by

$$\gamma(t) := X^{1/2}(X^{-1/2} Y X^{-1/2})^t X^{1/2},$$

by using the Löwner–Heinz inequality (see [14, Lemma 2.1]), we have $\gamma(t) \geq A$ and therefore S is convex in P_n^+. We claim that A is a solution for the problem (7). By using Theorem 3.1, we need to prove that $\langle A, \exp_A^{-1}(Y)\rangle_A \geq 0$ for all $Y \in S$. Note that

$$\begin{aligned}\langle A, \exp_A^{-1}(Y)\rangle_A &= \operatorname{tr}(A^{-1}AA^{-1}A\log(A^{-1}Y)) \\ &= \operatorname{tr}\log(A^{-1}Y) = \log\det(A^{-1}Y) \geq 0.\end{aligned}$$

Therefore,

$$\bar{S} = \{X \in S : \log\det A = \log\det X\}.$$

To illustrate Theorem 3.2, we will see that

$$\bar{S} = \{X \in S : \langle \operatorname{grad} f(A), \exp_A^{-1} X\rangle_A = 0,\ P_X^A[\operatorname{grad} f(X)] = \operatorname{grad} f(A)\}.$$

Note that $\operatorname{grad} f(X) = X, \operatorname{grad} f(A) = A$, and

$$P_X^A(X) = (AX^{-1})^{1/2} X (X^{-1}A)^{1/2} = A \text{ for all } \ X \in S.$$

Moreover,

$$\begin{aligned}\langle \operatorname{grad} f(A), \exp_A^{-1} X\rangle_A &= \langle A, A\log(A^{-1}X)\rangle_A \\ &= \operatorname{tr}(A^{-1}AA^{-1}A\log(A^{-1}X)) \\ &= \operatorname{tr}\log(A^{-1}X) = \log\det(A^{-1}X),\end{aligned}$$

which shows the required equation.

Recall that the unit sphere $S^2 := \{x \in \mathbb{R}^3 :\ ||x|| = 1\}$ is a 2-dimensional manifold with the usual Riemannian distance function defined as

$$d(x, y) = \arccos\langle x, y\rangle \quad \text{for all } x, y \in S^2.$$

For every $\bar{x} \in S^2$, it follows from the definition of the Riemannian metric on S^2 that

$$\langle u, v\rangle_{\bar{x}} = \langle u, v\rangle \quad \text{for all } u, v \in T_{\bar{x}} S^2,$$

where $\langle \cdot, \cdot \rangle$ denotes the standard inner product in $\mathbb{R}^3$. For $x \in S^2$, the exponential map $\exp_x : T_xS^2 \to S^2$ is defined by

$$\exp_x(v) = \cos(||v||)x + \sin(||v||)\frac{v}{||v||}, \quad v \in T_xS^2. \tag{8}$$

Moreover, $\exp_x^{-1} : S^2 \to T_xS^2$ is

$$\exp_x^{-1}(y) = \frac{\theta}{\sin\theta}(y - x\cos\theta), \quad y \in S^2, \tag{9}$$

where $\theta = \arccos\langle x, y\rangle$. Let $t \to \gamma(t)$ be the unique minimal geodesic in S^2 joining $\gamma(0) = x$ to $\gamma(0) = y$, and let $u := \frac{\gamma'(0)}{||\gamma'(0)||}$. The parallel translation of a vector $v \in T_xS^2$ along the geodesic γ is given by

$$\begin{aligned} P_x^{\gamma(t)}(v) = \big(&- \sin(||\gamma'(0)||t)u'v\big)x \\ &- \big(\cos(||\gamma'(0)||t)u'v\big)u + (I - uu')v; \end{aligned} \tag{10}$$

see [1]. In the following example, which is an improvement of [16, Corollary 1, p. 1988], we consider the optimization problem on convex subsets of the unit sphere involving quadratic cost functions.

Example 3.3. Let S be a convex subset of S^2 and f be the quadratic convex function on an open convex subset of S^2 containing S defined by

$$f(x) := \langle Ax, x\rangle = x^TAx, \tag{11}$$

where $A \in \mathbb{R}^{n\times n}$ is an $n \times n$ matrix. Suppose that $\bar{x} \in \bar{S}$. By using formulas (8)–(10) and [1, p. 74], for every $x \in \bar{S}$, we get

$$\operatorname{grad} f(x) = 2(Ax - (xx^T)Ax),$$

and

$$\exp_{\bar{x}}^{-1}(x) = \frac{\theta}{\sin\theta}(x - \bar{x}\cos\theta),$$

where $\theta = \arccos\langle \bar{x}, x\rangle$. Therefore, by Theorem 3.2, we obtain

$$\bar{S} = \{x \in S : \langle c, x - \bar{x}\cos\theta\rangle_{\bar{x}} = 0, (I - xx^T)Ax = d\},$$

where $c = A\bar{x} - (\bar{x}\bar{x}^T)A\bar{x}$ and $d = P_{\bar{x}}^{x}(c)$.

References

[1] Absil, P.A., Mahony, R., Sepulchre, R.: Optimization Algorithms on Matrix Manifolds. Princeton University Press, Princeton (2009)

[2] Barani, A.: Subdifferentials of perturbed distance function in Riemannian manifolds. Optimization **67**(11), 1849–1868 (2018)

[3] Burke, J.V., Ferris, M.C.: Characterization of solution sets of convex programs. Oper. Res. Lett. **10**, 57–60 (1991)

[4] Da Cruz Neto, J.X., Oliveira, O.P., Lucambio Pérez, L.R., Nemeth, S.Z.: Convex- and monotone-transformable mathematical programming problems and a proximal-like point method. J. Glob. Optim. **35**(1), 53–69 (2006)

[5] Fazel, M., Hindi, H., Boyd, S.P.: Log-det heuristic for matrix rank minimization with applications to hankel and euclidean distance matrices. In: Proceedings of the 2003 American Control Conference, 2003, vol. 3, pp. 2156–2162 (2003)

[6] Ferreira, O.P., Nemeth, S.Z.: On the spherical convexity of quadratic functions. J. Glob. Optim. **73**(3), 537–545 (2019)

[7] Ferreira, O.P., Iusem, A.N., Nemeth, S.Z.: Projections onto convex sets on the sphere. J. Glob. Optim. **57**, 663–676 (2013)

[8] Grohs, P., Hosseini, S.: ε-Subgradient algorithms for locally Lipschitz functions on Riemannian manifolds. Adv. Comput. Math. **42**(2), 333–360 (2016)

[9] Grossir, D.: Newton's method, zeroes of vector fields, and the Riemannian center of mass. Adv. Appl. Math. **33**, 95–135 (2004)

[10] Horn, R.A., Johnson, C.A.: Matrix Analysis. Cambridge University Press, Cambridge (2013)

[11] Hosseini, S., Pouryayevali, M.R.: On the metric projection onto prox-regular subsets of Riemannian manifolds. Proc. Am. Math. Soc. **141**, 233–244 (2013)

[12] Jeyakumar, V., Yang, X.Q.: Convex composite multiobjective nonsmooth programming. Math. Program. **59**, 325–343 (1993)

[13] Jeyakumar, V., Lee, G.M., Dinh, N.: Lagrange multiplier conditions characterizing the optimal solution set of cone-constrained convex programs. J. Optim. Theory Appl. **123**, 83–103 (2004)

[14] Lim, Y.: Factorizations and geometric means of positive definite matrices. Linear Algebra Appl. **437**(9), 2159–2172 (2012)

[15] Li, C., Yao, J.C.: Variational inequalities for set-valued vector fields: convexity of the solution set and proximal point algorithm. SIAM J. Control Optim. **50**(4), 2486–2514 (2012)

[16] Mangasarian, O.L.: A simple characterization of solution sets of convex programs. Oper. Res. Lett. **7**, 21–26 (1988)

[17] Rapscák, T.: Smooth Nonlinear Optimization in $\mathbb{R}^n$. Kluwer Academic, Dordrecht (1997)

[18] Rendl, F., Wolkowicz, H.: A semidefinite framework for trust region subproblems with applications to large scale minimization. Math. Program. **77**(2 Ser. B), 237–299 (1997)

[19] Sakai, T.: Riemannian Geometry. American Mathematic Society, Providence (1992)

[20] Schneider, R., Uschmajew, A.: Convergence results for projected line-search methods on varieties of low-rank matrices via Lojasiewicz inequality. SIAM J. Optim. **25**(1), 622–646 (2015)

[21] Sra, S., Hosseini, R.: Conic geometric optimization on the manifold of positive definite matrices. SIAM J. Optim. **25**(1), 713–739 (2015)

[22] Udriste, C.: Convex Functions and Optimization Methods on Riemannian Manifolds. Mathematics and its Applications, vol. 297. Kluwer Academic, Dordrecht (1994)

[23] Wu, Z.L., Wu, S.Y.: Characterizations of the solution sets of convex programs and variational inequality problems. J. Optim. Theory Appl. **130**, 339–358 (2006)

A. Barani
Department of Mathematics
Lorestan University
P.O. Box 465
Khorramabad
Iran
e-mail: barani.a@lu.ac.ir

S. Hosseini
Hausdorff Center for Mathematics
Institute for Numerical Simulation
University of Bonn
53115 Bonn
Germany
e-mail: hosseini@ins.uni-bonn.de

Received: 2 March 2019