SIAM J. MATRIX ANAL. APPL.
Vol. 31, No. 3, pp. 1491–1506

CONVEXITY PROPERTIES OF THE CONDITION NUMBER*

CARLOS BELTRÁN†, JEAN-PIERRE DEDIEU‡, GREGORIO MALAJOVICH§, AND MIKE SHUB¶

Abstract. We define in the space of $n \times m$ matrices of rank n, $n \leq m$, the condition Riemannian structure as follows: For a given matrix A the tangent space at A is equipped with the Hermitian inner product obtained by multiplying the usual Frobenius inner product by the inverse of the square of the smallest singular value of A denoted $\sigma_n(A)$. When this smallest singular value has multiplicity 1, the function $A \to \log(\sigma_n(A)^{-2})$ is a convex function with respect to the condition Riemannian structure that is $t \to \log(\sigma_n(A(t))^{-2})$ is convex, in the usual sense for any geodesic $A(t)$. In a more abstract setting, a function α defined on a Riemannian manifold $(\mathcal{M}, \langle , \rangle)$ is said to be self-convex when $\log \alpha(\gamma(t))$ is convex for any geodesic in $(\mathcal{M}, \alpha \langle , \rangle)$. Necessary and sufficient conditions for self-convexity are given when α is C^2. When $\alpha(x) = d(x, \mathcal{N})^{-2}$, where $d(x, \mathcal{N})$ is the distance from x to a C^2 submanifold $\mathcal{N} \subset \mathbb{R}^j$, we prove that α is self-convex when restricted to the largest open set of points x where there is a unique closest point in $\mathcal{N}$ to x. We also show, using this more general notion, that the square of the condition number $\|A\|_F / \sigma_n(A)$ is self-convex in projective space and the solution variety.

Key words. condition number, geodesic, log-convexity, Riemannian geometry, linear group

AMS subject classifications. Primary, 65F35; Secondary, 15A12

DOI. 10.1137/080718681

1. Introduction.

Let two integers $1 \leq n \leq m$ be given, and let us consider the space of matrices $\mathbb{K}^{n \times m}$, $\mathbb{K} = \mathbb{R}$, or $\mathbb{C}$, equipped with the Frobenius Hermitian product

$$\langle M, N \rangle_F = \text{trace } (N^* M) = \sum_{i,j} m_{ij} \overline{n_{ij}}.$$

Given an absolutely continuous path $A(t)$, $a \leq t \leq b$, its length is given by the integral

$$L = \int_a^b \left\| \frac{dA(t)}{dt} \right\|_F dt,$$

and the shortest path connecting $A(a)$ to $A(b)$ is the segment connecting them. Consider now the problem of connecting these two matrices with the shortest possible path in staying, as much as possible, away from the set of "singular matrices," that is, the matrices with nonmaximal rank.

*Received by the editors March 17, 2008; accepted for publication (in revised form) by A. S. Lewis October 2, 2009; published electronically January 6, 2010.

http://www.siam.org/journals/simax/31-3/71868.html

†Departmento de Matemáticas, Estadísticas y Computacíon Universidad de Cantábria, 39005 Santander, Spain (beltranc@gmail.com). This author was supported by MTM2007-62799, a Spanish postdoctoral grant, and an NSERC Discovery grant.

‡Institut de Mathématiques, Université Paul Sabatier, 31062 Toulouse cedex 09, France (jean-pierre.dedieu@math.univ-toulouse.fr). This author was supported by the ANR Gecko.

§Departamento de Matemática Aplicada, Universidade Federal de Rio de Janeiro, Caixa Postal 68530, CEP 21945-970, Rio de Janeiro, RJ, Brazil (gregorio@ufrj.br). This author was partially supported by CNPq grants 303565/2007-1 and 470031/2007-7, by FAPERJ (Fundação Carlos Chagas de Amparo à Pesquisa do Estado do Rio de Janeiro), and by the Brazil-France agreement of cooperation in Mathematics.

¶Department of Mathematics, University of Toronto, Toronto, ON, M5S 2E4, Canada (shub.michael@gmail.com). This author was supported by an NSERC Discovery grant.

The singular values of a matrix $A \in \mathbb{K}^{n\times m}$ are denoted in nonincreasing order:

$$\sigma_1(A) \geq \cdots \geq \sigma_{n-1}(A) \geq \sigma_n(A) \geq 0.$$

We denote by $\mathbb{GL}_{n,m}$ the space of matrices $A \in \mathbb{K}^{n\times m}$ with maximal rank: rank $A = n$; that is, $\sigma_n(A) > 0$ so that the set of singular matrices is

$$\mathcal{N} = \mathbb{K}^{n\times m} \setminus \mathbb{GL}_{n,m} = \left\{A \in \mathbb{K}^{n\times m} \ : \ \sigma_n(A) = 0\right\}.$$

Since the smallest singular value of a matrix is equal to the distance from the set of singular matrices:

$$\sigma_n(A) = d_F(A, \mathcal{N}) = \min_{S\in\mathcal{N}} \|A - S\|_F,$$

given an absolutely continuous path $A(t)$, $a \leq t \leq b$, we define its "condition length" by the integral

$$L_\kappa = \int_a^b \left\|\frac{dA(t)}{dt}\right\|_F \sigma_n(A(t))^{-1} dt.$$

A good compromise between length and distance to $\mathcal{N}$ is obtained in minimizing L_κ. We call "minimizing condition geodesic" an absolutely continuous path, parametrized by arc length, which minimizes L_κ in the set of absolutely continuous paths with given endpoints and condition distance $d_\kappa(A, B)$ between two matrices the length L_κ of a minimizing condition geodesic with endpoints A and B, if any.

In this paper our objective is to investigate the properties of the smallest singular value $\sigma_n(A(t))$ along a condition geodesic. Our main result says that the map $\log\left(\sigma_n(A(t))^{-1}\right)$ is convex. Thus $\sigma_n(A(t))$ is concave, and its minimum value along the path is reached at one of the endpoints.

Note that a similar property holds in the case of hyperbolic geometry where instead of $\mathbb{K}^{n\times m}$ we take $\mathbb{R}^{n-1} \times [0, \infty[$, instead of $\mathcal{N}$ where we have $\mathbb{R}^{n-1} \times \{0\}$, and where the length of a path $a(t) = (a_1(t), \ldots, a_n(t))$ is defined by the integral

$$\int \left\|\frac{da(t)}{dt}\right\| a_n(t)^{-1} dt.$$

Geodesics in that case are arcs of circles centered at $\mathbb{R}^{n-1} \times \{0\}$ or segments of vertical lines, and $\log\left(a_n(t)^{-1}\right)$ is convex along such paths.

The approach used here to prove our theorems is heavily based on Riemannian geometry. We define on $\mathbb{GL}_{n,m}$ the following Riemannian structure:

$$\langle M, N\rangle_{\kappa,A} = \sigma_n(A)^{-2} \mathrm{Re}\, \langle M, N\rangle_F,$$

where $M, N \in \mathbb{K}^{n\times m}$ and $A \in \mathbb{GL}_{n,m}$. The minimizing condition geodesics defined previously are clearly geodesic in $\mathbb{GL}_{n,m}$ for this Riemannian structure so that we may use the toolbox of Riemannian geometry. In fact things are not so simple: the smallest singular value $\sigma_n(A)$ is a locally Lipschitz map in $\mathbb{GL}_{n,m}$, and it is smooth on the open subset

$$\mathbb{GL}^{>}_{n,m} = \{A \in \mathbb{GL}_{n,m} \ : \ \sigma_{n-1}(A) > \sigma_n(A)\},$$

that is, when the smallest singular value of A is simple. On the open subset $\mathbb{GL}^{>}_{n,m}$ the metric $\langle\cdot,\cdot\rangle_\kappa$ defines a smooth Riemannian structure, and we call "condition geodesics"

the geodesics related to this structure. Such a path is not necessarily a minimizing geodesic. Our first main theorem establishes a remarkable property of the condition Riemannian structure.

THEOREM 1. *σ_n^{-2} is logarithmically convex on $\mathbb{GL}_{n,m}^{>}$, i.e., for any geodesic curve $\gamma(t)$ in $\mathbb{GL}_{n,m}^{>}$ for the condition metric the map $\log\left(\sigma_n^{-2}(\gamma(t))\right)$ is convex.*

PROBLEM 1. *The condition Riemannian structure $\langle .,. \rangle_\kappa$ is defined in $\mathbb{GL}_{n,m}$ where it is only locally Lipschitz. Let us define condition geodesics in $\mathbb{GL}_{n,m}$ as the extremals of the condition length L_κ (see, for example,* [3, *Thm.* 4.4.3, *Chap.* 4] *for the definition of such extremals in the Lipschitz case). Is Theorem* 1 *still true for $\mathbb{GL}_{n,m}$? All the examples we have studied confirm that convexity holds, even if $\sigma_n^{-1}(\gamma(t))$ fails to be C^1. See Boito-Dedieu* [2]*. We intend to address this issue in a future paper.*

In a second step we extend these results to other spaces of matrices: the sphere $\mathbb{S}_r(\mathbb{GL}_{n,m}^{>})$ of radius r in $\mathbb{GL}_{n,m}^{>}$ in Corollary 6, the projective space $\mathbb{P}\left(\mathbb{GL}_{n,m}^{>}\right)$ in Corollary 7. We also consider the case of the solution variety of the homogeneous equation $M\zeta = 0$, that is, the set of pairs

$$\left\{(M,\zeta) \in \mathbb{K}^{n\times(n+1)} \times \mathbb{K}^{n+1} \ : \ M\zeta = 0\right\}.$$

Now our function α is the square of the condition number studied by Demmel in [4]. This is done in the affine context in Theorem 3 and in the projective context in Corollary 8.

Since $\sigma_n(A)$ is equal to the distance from A to the set of singular matrices a natural question is to ask whether our main result remains valid for the inverse of the distance from certain sets or for more general functions.

DEFINITION 1. *Let $(\mathcal{M}, \langle\cdot,\cdot\rangle)$ be Riemannian, and let $\alpha : \mathcal{M} \to \mathbb{R}$ be a function of class C^2 with positive values. Let $\mathcal{M}_\kappa$ be the manifold $\mathcal{M}$ with the new metric*

$$\langle\cdot,\cdot\rangle_{\kappa,x} = \alpha(x)\langle\cdot,\cdot\rangle_x$$

called condition Riemann structure. We say that α is self-convex when $\log\alpha(\gamma(t))$ is convex for any geodesic γ in $\mathcal{M}_\kappa$.

For example, with $\mathcal{M} = \{x = (x_1,\ldots,x_n) \in \mathbb{R}^n : x_n > 0\}$ equipped with the usual metric, $\alpha(x) = x_n^{-2}$ is self-convex. The space $\mathcal{M}_\kappa$ is the Poincaré model of hyperbolic space.

In the following theorem we prove self-convexity for the distance function to a C^2 submanifold without boundary $\mathcal{N} \subset \mathbb{R}^j$. Let us denote by

$$\rho(x) = d(x,\mathcal{N}) = \min_{y\in\mathcal{N}} \|x-y\| \quad \text{and } \alpha(x) = \frac{1}{\rho(x)^2}.$$

Let $\mathcal{U}$ be the largest open set in $\mathbb{R}^j$ such that, for any $x \in \mathcal{U}$, there is a unique closest point in $\mathcal{N}$ to x. When $\mathcal{U}$ is equipped with the new metric $\alpha(x)\,\langle .,.\rangle$ we have the following theorem.

THEOREM 2. *The function $\alpha : \mathcal{U} \setminus \mathcal{N} \to \mathbb{R}$ is self-convex.*

Theorem 2 is then extended to the projective case. Let $\mathcal{N}$ be a $\mathcal{C}^2$ submanifold without boundary $\mathbb{P}(\mathbb{R}^j)$. Let us denote by d_R the Riemannian distance in projective space (points in the projective space are lines through the origin and the distance d_R between two lines is the angle they make). Let us denote $d_{\mathbb{P}} = \sin d_R$ (this is also a distance), define $\alpha(x) = d_{\mathbb{P}}(x,\mathcal{N})^{-2}$, and let $\mathcal{U}$ be the largest open subset of $\mathbb{P}(\mathbb{R}^j)$ such that for $x \in \mathcal{U}$ there is a unique closest point from $\mathcal{N}$ to x for the distance $d_{\mathbb{P}}$. Then we have the following corollary.

COROLLARY 1. *The map $\alpha : \mathcal{U} \setminus \mathcal{N} \to \mathbb{R}$ is self-convex.*

The extension of Theorems 1 and 2 to other types of sets or functions is not obvious. In Example 1 we prove that $\alpha(A) = \sigma_1(A)^{-2} + \cdots + \sigma_n(A)^{-2}$ is not self-convex in $\mathbb{GL}_{n,m}$.

In Example 2 we take $\mathcal{N} = \mathbb{R}^2$, and $\mathcal{U}$ the unit disk so that $\mathcal{U}$ contains a point (the center) which has many closest points from $\mathcal{N}$. In that case the corresponding function $\alpha : \mathcal{U} \setminus \mathcal{N} \to \mathbb{R}$ is self-convex, but it fails to be smooth at the center of the disk.

In Example 3 we provide an example of a submanifold $\mathcal{N} \subset \mathbb{R}^2$ such that the function $\alpha(x) = d(x, \mathcal{N})^{-2}$ defined on $\mathbb{R}^2 \setminus \mathcal{N}$ is not self-convex.

Our interest in considering the condition metric in the space of matrices comes from recent papers by Shub [8] and Beltrán and Shub [1] where these authors use condition length along a path in certain solution varieties to estimate step size for continuation methods to follow these paths. They give bounds on the number of steps required in terms of the condition length of the path. If geodesics in the condition metric are followed, the known bounds on polynomial system solving are vastly improved. To understand the properties of these geodesics, we have begun this paper with linear systems where we can investigate their properties more deeply. We find self-convexity in the context of this paper remarkable. We do not know if similar issues may naturally arise in linear algebra even for solving systems of linear equations. Similar issues do clearly arise when studying continuation methods for the eigenvalue problem.

2. Self-convexity. Let us first recall some basic definitions about convexity on Riemannian manifolds. A good reference on this subject is Udrişte [9].

DEFINITION 2. *We say that a function $f : \mathcal{M} \to \mathbb{R}$ is convex whenever*

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

for every $x, y \in \mathcal{M}$, for every geodesic arc γ_{xy} joining x and y and $0 \leq t \leq 1$.

The convexity of f in $\mathcal{M}$ is equivalent to the convexity in the usual sense of $f \circ \gamma_{xy}$ on $[0,1]$ for every $x, y \in U$ and the geodesic γ_{xy} joining x and y or also to the convexity of $g \circ \gamma$ for every geodesic γ [9, Thm. 2.2, Chap. 3]. Thus, we see the following lemma.

LEMMA 1. *Self-convexity of a function $\alpha : \mathcal{M} \to \mathbb{R}$ is equivalent to the convexity of $\log \circ \alpha$ in the condition Riemannian manifold $\mathcal{M}_\kappa$.*

When f is a function of class C^2 in the Riemannian manifold $\mathcal{M}$, we define its second derivative $D^2 f(x)$ as the second covariant derivative. It is a symmetric bilinear form on $T_x\mathcal{M}$. Note [9, Chap. 1] that if $x \in \mathcal{M}$ and $\dot{x} \in T_x\mathcal{M}$, and if $\gamma(t)$ is a geodesic in $\mathcal{M}$, $\gamma(0) = x$, $\frac{d}{dt}\gamma(0) = \dot{x}$, then

$$D^2 f(x)(\dot{x}, \dot{x}) = \frac{d^2}{dt^2}(f \circ \gamma)(0).$$

This second derivative depends on the Riemannian connection on $\mathcal{M}$. Since $\mathcal{M}$ is equipped with two different metrics, $\langle ., . \rangle$ and $\langle ., . \rangle_\kappa$, we have to distinguish between the corresponding second derivatives; they are denoted by $D^2 f(x)$ and $D^2_\kappa f(x)$, respectively. No such distinction is necessary for the first derivative $Df(x)$.

Convexity on Riemannian manifold is characterized by (see [9, Thm. 6.2, Chap. 3]) the following proposition.

PROPOSITION 1. *A function $f : \mathcal{M} \to \mathbb{R}$ of class C^2 is convex if and only if $D^2 f(x)$ is positive semidefinite for every $x \in \mathcal{M}$.*

We use this proposition to obtain a characterization of self-convexity: α is self-convex if and only if the second derivative $D^2_\kappa(\log\circ\alpha)(x)$ is positive semidefinite for any $x \in \mathcal{M}_\kappa$. We get the following proposition.

PROPOSITION 2. *For a function $\alpha : \mathcal{M} \to \mathbb{R}$ of class C^2 with positive values self-convexity is equivalent to*

$$2\alpha(x)D^2\alpha(x)(\dot{x},\dot{x}) + \|D\alpha(x)\|_x^2\|\dot{x}\|_x^2 - 4(D\alpha(x)\dot{x})^2 \geq 0$$

for any $x \in \mathcal{M}$ and for any vector $\dot{x} \in T_x\mathcal{M}$, the tangent space at x.

Proof. Let $x \in \mathcal{M}$ be given. Let $\varphi : \mathbb{R}^m \to \mathcal{M}$ be a coordinate system such that $\varphi(0) = x$ and with first fundamental form $g_{ij}(0) = \delta_{ij}$ (Kronecker's delta) and Christoffel's symbols $\Gamma^i_{jk}(0) = 0$, and let

$$A = \alpha \circ \varphi$$

so that $\alpha(x) = A(0)$. Those coordinates are called "normal" or "geodesic." Note that this implies

$$\frac{\partial g_{ij}}{\partial z_k}(0) = 0$$

for all i, j, k. We denote by $g_{\kappa,ij}$ and $\Gamma^i_{\kappa,jk}$, respectively, the first fundamental form and the Christoffel symbols for φ in $\mathcal{M}_\kappa$. Let us compute them. Note that

$$\begin{aligned} g_{\kappa,ij}(z) &= g_{ij}(z)A(z), \\ \frac{\partial g_{\kappa,ij}}{\partial z_k}(0) &= Dg_{\kappa,ij}(0)(e_k) = D(g_{ij}A)(0)(e_k) \\ &= g_{ij}(0)DA(0)(e_k) + A(0)Dg_{ij}(0)(e_k) = \delta_{ij}\frac{\partial A}{\partial z_k}(0). \end{aligned}$$

Moreover,

$$\begin{aligned} \Gamma^i_{\kappa,jk} = \frac{1}{A(0)}\Gamma^i_{jk} &= \frac{1}{2A(0)}\left(\frac{\partial g_{\kappa,ij}}{\partial z_k}(0) + \frac{\partial g_{\kappa,ik}}{\partial z_j}(0) - \frac{\partial g_{\kappa,jk}}{\partial z_i}(0)\right) \\ &= \frac{1}{2A(0)}\left(\delta_{ij}\frac{\partial A}{\partial z_k}(0) + \delta_{ik}\frac{\partial A}{\partial z_j}(0) - \delta_{jk}\frac{\partial A}{\partial z_i}(0)\right). \end{aligned}$$

That is,

$$\begin{cases} \Gamma^i_{\kappa,ik} = \Gamma^i_{\kappa,ki} = \frac{1}{2A(0)}\frac{\partial A}{\partial z_k}(0) & \text{for all } i,k, \\ \Gamma^i_{\kappa,jj} = \frac{-1}{2A(0)}\frac{\partial A}{\partial z_i}(0), & j \neq i, \\ \Gamma^i_{\kappa,jk} = 0 & \text{otherwise.} \end{cases}$$

The second derivative of the composition of two maps

$$\mathcal{M} \xrightarrow{f} \mathbb{R} \xrightarrow{\psi} \mathbb{R}$$

is given by the identity (see [9, Hessian, Chap. 1.3])

$$D^2(\psi \circ f)(x) = D\psi(f(x))D^2f(x) + \psi''(f(x))Df(x) \otimes Df(x),$$

where $Df(x) \otimes Df(x)$ is the bilinear form on $T_x\mathcal{M}$ by

$$(Df(x) \otimes Df(x))(u,v) = Df(x)(u)Df(x)(v).$$

This gives, in our context (that is, when $f = \alpha$ and $\psi = \log$),

$$D^2_\kappa(\log\circ\alpha)(x) = \frac{1}{\alpha(x)} D^2_\kappa\alpha(x) - \frac{1}{\alpha(x)^2} D\alpha(x)\otimes D\alpha(x).$$

According to Proposition 1 our objective is now to give a necessary and sufficient condition for $D^2_\kappa(\log\circ\alpha)(x)$ to be positive semidefinite for each $x \in \mathcal{M}$. In our system of local coordinates the components of $D^2\alpha(x)$ are (see [9, Chap. 1.3])

$$A_{jk} = \frac{\partial^2 A}{\partial z_j \partial z_k} - \sum_i \Gamma^i_{jk} \frac{\partial A}{\partial z_i} = \frac{\partial^2 A}{\partial z_j \partial z_k}$$

while the components of $D^2_\kappa\alpha(x)$ are

$$A_{\kappa,jk} = \frac{\partial^2 A}{\partial z_j \partial z_k} - \sum_i \Gamma^i_{\kappa,jk} \frac{\partial A}{\partial z_i}.$$

If we replace the Christoffel symbols in this last sum by the values previously computed, then we obtain, when $j = k$,

$$\begin{aligned}\sum_i \Gamma^i_{\kappa,jj} \frac{\partial A}{\partial z_i} &= \Gamma^j_{\kappa,jj} \frac{\partial A}{\partial z_j} + \sum_{i\neq j} \Gamma^i_{\kappa,jj} \frac{\partial A}{\partial z_i} \\ &= \frac{1}{2A}\left(\frac{\partial A}{\partial z_j}\right)^2 - \frac{1}{2A}\sum_{i\neq j}\left(\frac{\partial A}{\partial z_i}\right)^2 = \frac{1}{A}\left(\frac{\partial A}{\partial z_j}\right)^2 - \frac{1}{2A}\sum_i \left(\frac{\partial A}{\partial z_i}\right)^2\end{aligned}$$

while, when $j \neq k$,

$$\sum_i \Gamma^i_{\kappa,jk} \frac{\partial A}{\partial z_i} = \Gamma^j_{\kappa,jk} \frac{\partial A}{\partial z_j} + \Gamma^k_{\kappa,jk} \frac{\partial A}{\partial z_k} = \frac{1}{2A}\frac{\partial A}{\partial z_k}\frac{\partial A}{\partial z_j} + \frac{1}{2A}\frac{\partial A}{\partial z_j}\frac{\partial A}{\partial z_k} = \frac{1}{A}\frac{\partial A}{\partial z_j}\frac{\partial A}{\partial z_k}.$$

Both cases are subsumed in the identity

$$\sum_i \Gamma^i_{\kappa,jk} \frac{\partial A}{\partial z_i} = \frac{1}{A}\frac{\partial A}{\partial z_j}\frac{\partial A}{\partial z_k} - \frac{\delta_{jk}}{2A}\sum_i \left(\frac{\partial A}{\partial z_i}\right)^2.$$

Putting together all these identities gives the following expression for the components of $D^2_\kappa(\log\circ\alpha)(x)$:

$$\begin{aligned}D_k a^2 (\log\circ\alpha)(x)_{jk} &= \frac{1}{A}\left(\frac{\partial^2 A}{\partial z_j \partial z_k} - \frac{1}{A}\frac{\partial A}{\partial z_j}\frac{\partial A}{\partial z_k} + \frac{\delta_{jk}}{2A}\sum_i \left(\frac{\partial A}{\partial z_i}\right)^2\right) - \frac{1}{A^2}\frac{\partial A}{\partial z_j}\frac{\partial A}{\partial z_k} \\ &= \frac{1}{2A^2}\left(2A\frac{\partial^2 A}{\partial z_j \partial z_k} + \delta_{jk}\sum_i \left(\frac{\partial A}{\partial z_i}\right)^2 - 4\frac{\partial A}{\partial z_j}\frac{\partial A}{\partial z_k}\right).\end{aligned}$$

Thus, $D^2_\kappa(\log\circ\alpha)(x) \geq 0$ if and only if

$$2\alpha(x)D^2\alpha(x) + \|D\alpha(x)\|^2_x \langle .,.\rangle_x - 4D\alpha(x)\otimes D\alpha(x)$$

is positive semidefinite, that is, when

$$2\alpha(x)D^2\alpha(x)(\dot{x},\dot{x}) + \|D\alpha(x)\|^2_x \|\dot{x}\|^2_x - 4(D\alpha(x)\dot{x})^2 \geq 0$$

for any $x \in \mathcal{M}$ and for any vector $\dot{x} \in T_x\mathcal{M}$. This finishes the proof. ∎

An easy consequence of Proposition 2 is the following. See also Example 3.

COROLLARY 2. *When a function $\alpha : \mathcal{M} \to \mathbb{R}$ of class C^2 is self-convex, then any critical point of α has a positive semidefinite second derivative $D^2\alpha(x)$. Such a function cannot have a strict local maximum or a nondegenerate saddle.*

PROPOSITION 3. *The following condition is equivalent for a C^2 function $\alpha = 1/\rho^2 : \mathcal{M} \longrightarrow \mathbb{R}$ to be self-convex on $\mathcal{M}$: For every $x \in \mathcal{M}$ and $\dot{x} \in T_x\mathcal{M}$,*

$$\|\dot{x}\|^2\|D\rho(x)\|^2 - (D\rho(x)\dot{x})^2 - \rho(x)D^2\rho(x)(\dot{x},\dot{x}) \geq 0,$$

or, what is the same,

$$2\|\dot{x}\|^2\|D\rho(x)\|^2 \geq D^2\rho^2(x)(\dot{x},\dot{x}).$$

Proof. Note that

$$D\alpha(x)\dot{x} = \frac{-2}{\rho(x)^3}D\rho(x)\dot{x},$$

$$D^2\alpha(x)(\dot{x},\dot{x}) = \frac{6}{\rho(x)^4}(D\rho(x)\dot{x})^2 - \frac{2}{\rho(x)^3}D^2\rho(x)(\dot{x},\dot{x}).$$

Hence, the necessary and sufficient condition of Proposition 2 reads

$$\frac{4\|\dot{x}\|^2\|D\rho(x)\|^2}{\rho(x)^6} - \frac{16}{\rho(x)^6}(D\rho(x)\dot{x})^2 + \frac{12}{\rho(x)^6}(D\rho(x)\dot{x})^2 - \frac{4}{\rho(x)^5}D^2\rho(x)(\dot{x},\dot{x}) \geq 0,$$

and the proposition follows. ∎

COROLLARY 3. *Each of the following conditions is sufficient for a function $\alpha = 1/\rho^2 : \mathcal{M} \longrightarrow \mathbb{R}$ to be self-convex at $x \in \mathcal{M}$: For every $\dot{x} \in T_x\mathcal{M}$,*

$$D^2\rho(x)(\dot{x},\dot{x}) \leq 0,$$

or

$$\|D^2\rho^2(x)\| \leq 2\|D\rho(x)\|^2.$$

In the following proposition we obtain a weaker condition on α to obtain convexity in $\mathcal{M}_\kappa$ instead of self-convexity.

PROPOSITION 4. *$\alpha(x)$ is convex in $\mathcal{M}_\kappa$ if and only if*

$$2\alpha(x)D^2\alpha(x)(\dot{x},\dot{x}) + \|D\alpha(x)\|_x^2\|\dot{x}\|_x^2 - 2(D\alpha(x)\dot{x})^2 \geq 0$$

for any $x \in \mathcal{M}$ and any vector $\dot{x} \in T_x\mathcal{M}$.

Proof. We follow the lines of the proof of Proposition 2 with ψ equal to the identity map instead of $\psi = \log$. ∎

3. Some general formulas for matrices.

PROPOSITION 5. *Let $A = (\Sigma, 0) \in \mathbb{GL}^{>}_{n,m}$, where $\Sigma = diag\ (\sigma_1 \geq \cdots \geq \sigma_{n-1} > \sigma_n) \in \mathbb{K}^{n\times n}$. The map $\sigma_n : \mathbb{GL}^{>}_{n,m} \to \mathbb{R}$ is a smooth map and, for every $U \in \mathbb{K}^{n\times m}$,*

$$\begin{cases} D\sigma_n(A)U = \mathrm{Re}(u_{nn}), \\ D^2\sigma_n^2(A)(U,U) = 2\sum_{j=1}^m |u_{nj}|^2 - 2\sum_{k=1}^{n-1} \frac{|u_{kn}\sigma_n + \overline{u_{nk}}\sigma_k|^2}{\sigma_k^2 - \sigma_n^2}. \end{cases}$$

Proof. Since σ_n^2 is an eigenvalue of AA^* with multiplicity 1, the implicit function theorem proves the existence of smooth functions $\sigma_n^2(B) \in \mathbb{R}$ and $u(B) \in \mathbb{K}^n$, defined in an open neighborhood of A and satisfying

$$\begin{cases} BB^*u(B) = \sigma_n^2(B)u(B), \\ \|u(B)\|^2 = 1, \\ u(A) = e_n = (0,\dots,0,1)^T \in \mathbb{K}^n, \\ \sigma_n^2(A) = \sigma_n^2. \end{cases}$$

Differentiating these equations at B gives, for any $U \in \mathbb{K}^{n\times m}$,

$$\begin{cases} (UB^* + BU^*)u(B) + BB^*\dot{u}(B) = \left(\sigma_n^2\right)' u(B) + \sigma_n^2(B)\dot{u}(B), \\ u(B)^*\dot{u}(B) = 0 \end{cases}$$

with $\dot{u}(B) = Du(B)U$ and $\left(\sigma_n^2\right)' = D\sigma_n^2(B)U$. Premultiplying the first equation by $u(B)^*$ gives

$$u(B)^*(UB^* + BU^*)u(B) + u(B)^*BB^*\dot{u}(B) = \left(\sigma_n^2\right)' u(B)^*u(B) + \sigma_n^2(B)u(B)^*\dot{u}(B)$$

so that

$$D\sigma_n^2(B)U = \left(\sigma_n^2\right)' = 2\mathrm{Re}(u(B)^*UB^*u(B))$$

and

$$D\sigma_n(B)U = \frac{\mathrm{Re}(u(B)^*UB^*u(B))}{\sigma_n(B)}.$$

The derivative of the eigenvector is now easy to compute:

$$Du(B)U = \dot{u}(B) = (\sigma_n^2(B)I_n - BB^*)^\dagger(UB^* + BU^* - \left(\sigma_n^2\right)' I_n)u(B),$$

where $(\sigma_n^2(B)I_n - BB^*)^\dagger$ denotes the generalized inverse (or Moore–Penrose inverse) of $\sigma_n^2(B)I_n - BB^*$.

The second derivative of σ_n^2 at B is given by

$$\begin{aligned} D^2\sigma_n^2(B)(U,U) &= 2\mathrm{Re}(\dot{u}(B)^*UB^*u(B) + u(B)^*UU^*u(B) + u(B)^*UB^*\dot{u}(B)) \\ &= 2\mathrm{Re}(u(B)^*UU^*u(B) + u(B)^*(UB^* + BU^*)\dot{u}(B)) = 2\mathrm{Re}(u(B)^*UU^*u(B) \\ &\quad + u(B)^*(UB^* + BU^*)(\sigma_n^2(B)I_n - BB^*)^\dagger(UB^* + BU^* - \left(\sigma_n^2\right)' I_n)u(B)). \end{aligned}$$

Using $u(A) = e_n$ and $\sigma_n(A) = \sigma_n$ we get

$$\begin{cases} D\sigma_n^2(A)U = 2\mathrm{Re}(UA^*)_{nn} = 2\sigma_n\mathrm{Re}(u_{nn}), \\ D\sigma_n(A)U = \mathrm{Re}(u_{nn}), \end{cases}$$

and the second derivative is given by

$$\begin{aligned} &D^2\sigma_n^2(A)(U,U) \\ &= 2\mathrm{Re}\left((UU^*)_{nn} + \sum_{k=1}^{n-1}(UA^* + AU^*)_{nk}(\sigma_n^2 - \sigma_k^2)^{-1}(UA^* + AU^* - \left(\sigma_n^2\right)' I_n)_{kn}\right) \\ &= 2\mathrm{Re}\left((UU^*)_{nn} + \sum_{k=1}^{n-1}\frac{|(UA^* + AU^*)_{kn}|^2}{\sigma_n^2 - \sigma_k^2}\right) \\ &= 2\sum_{j=1}^{m}|u_{nj}|^2 - 2\sum_{k=1}^{n-1}\frac{|u_{kn}\sigma_n + \overline{u_{nk}}\sigma_k|^2}{\sigma_k^2 - \sigma_n^2}. \end{aligned}$$

∎

COROLLARY 4. *Let $A = (\Sigma, 0) \in \mathbb{GL}^{>}_{n,m}$, where $\Sigma = \operatorname{diag}(\sigma_1 \geq \cdots \geq \sigma_{n-1} > \sigma_n > 0) \in \mathbb{K}^{n\times n}$. Let us define $\rho(A) = \sigma_n(A)/\|A\|_F$. Then, for any $U \in \mathbb{K}^{n\times m}$ such that $\operatorname{Re}\langle A, U\rangle_F = 0$, we have*

$$\begin{cases} D\rho(A)U = \frac{\operatorname{Re}(u_{nn})}{\|A\|_F}, \\ D^2\rho^2(A)(U,U) = \frac{2}{\|A\|_F^2}\left(\sum_{j=1}^m |u_{nj}|^2 - \sum_{k=1}^{n-1} \frac{|u_{kn}\sigma_n + \overline{u_{nk}}\sigma_k|^2}{\sigma_k^2 - \sigma_n^2} - \frac{\|U\|_F^2}{\|A\|_F^2}\sigma_n^2\right). \end{cases}$$

Proof. Note that

$$D\rho(A)U = \frac{D\sigma_n(A)U\|A\|_F - \sigma_n(A)\frac{2\operatorname{Re}\langle A,U\rangle_F}{2\|A\|_F}}{\|A\|_F^2} = \frac{D\sigma_n(A)U}{\|A\|_F},$$

and the first assertion of the corollary follows from Proposition 5. For the second one, note that $h = h_1/h_2$ (for real valued $\mathcal{C}^2$ functions h, h_1, h_2 with $h_2(0) \neq 0$) implies

$$D^2h = \frac{h_2^2 D^2h_1 - h_1h_2D^2h_2 - 2h_2Dh_1Dh_2 + 2h_1(Dh_2)^2}{h_2^3}.$$

Now, $\rho^2(A) = \sigma_n^2(A)/\|A\|_F^2$, $D(\|A\|_F^2)U = 2\operatorname{Re}\langle A, U\rangle_F = 0$, $D^2(\|A\|_F^2)(U,U) = 2\|U\|_F^2$, and $D^2\sigma_n^2(A)(U,U)$ is known from Proposition 5. The formula for $D^2\rho^2(A)$ follows after some elementary calculations. ∎

4. The affine linear case. We consider here the Riemannian manifold $\mathcal{M} = \mathbb{GL}^{>}_{n,m}$ equipped with the usual Frobenius Hermitian product. Let $\alpha : \mathbb{GL}^{>}_{n,m} \to \mathbb{R}$ be defined as $\alpha(A) = 1/\sigma_n^2(A)$.

COROLLARY 5. *The function α is self-convex in $\mathbb{GL}^{>}_{n,m}$.*

Proof. From Proposition 3, it suffices to see that

$$2\|U\|_F^2\|D\sigma_n(A)\|_F^2 \geq D^2\sigma_n^2(A)(U,U).$$

Since unitary transformations are isometries in $\mathbb{GL}^{>}_{n,m}$ with respect to the condition metric we may suppose, via a singular value decomposition, that $A = (\Sigma, 0) \in \mathbb{GL}^{>}_{n,m}$, where $\Sigma = \operatorname{diag}(\sigma_1 \geq \cdots \geq \sigma_{n-1} > \sigma_n) \in \mathbb{K}^{n\times n}$. Now, the inequality to verify is obvious from Proposition 5, as $\|D\sigma_n(A)\|_F = 1$ and

$$D^2\sigma_n^2(A)(U,U) = 2\sum_{j=1}^m |u_{nj}|^2 - 2\sum_{k=1}^{n-1} \frac{|u_{kn}\sigma_n + \overline{u_{nk}}\sigma_k|^2}{\sigma_k^2 - \sigma_n^2} \leq 2\sum_{j=1}^m |u_{nj}|^2 \leq 2\|U\|_F^2. \qquad ∎$$

COROLLARY 6. *Let $r > 0$. The function α is self-convex in the sphere $\mathbb{S}_r(\mathbb{GL}^{>}_{n,m})$ of radius r in $\mathbb{GL}^{>}_{n,m}$.*

Proof. It is enough to prove that any geodesic in $(\mathbb{S}_r(\mathbb{GL}^{>}_{n,m}), \alpha)$ is also a geodesic in $(\mathbb{GL}^{>}_{n,m}, \alpha)$. Indeed, suppose that A and B are matrices in $\mathbb{S}_r(\mathbb{GL}^{>}_{n,m})$ and the minimal geodesic in $(\mathbb{GL}^{>}_{n,m}, \alpha)$ between A and B is $X(t)$, $a \leq t \leq b$. Then we claim that $L_\kappa\left(\frac{rX(t)}{\|X(t)\|_F}\right) \leq L_\kappa(X(t))$. Indeed, for any t,

$$\frac{d}{dt}\left(\frac{rX(t)}{\|X(t)\|_F}\right) = \frac{r\frac{dX(t)}{dt}}{\|X(t)\|_F} - r\frac{X(t)Re(\langle X(t), \frac{dX(t)}{dt}\rangle_F)}{\|X(t)\|_F^3}$$

so that

$$\begin{aligned}
&\left\| \frac{d}{dt}\left(\frac{rX(t)}{\|X(t)\|}\right)\right\|_F \\
&= \left(\frac{r^2 \left\|\frac{dX(t)}{dt}\right\|_F^2}{\|X(t)\|_F^2} + \frac{r^2 Re(\langle X(t), \frac{dX(t)}{dt}\rangle_F)^2}{\|X(t)\|_F^4} - \frac{2r^2 Re(\langle X(t), \frac{dX(t)}{dt}\rangle_F)^2}{\|X(t)\|_F^4} \right)^{1/2} \\
&= \left(\frac{r^2 \left\|\frac{dX(t)}{dt}\right\|_F^2}{\|X(t)\|_F^2} - \frac{r^2 Re(\langle X(t), \frac{dX(t)}{dt}\rangle_F)^2}{\|X(t)\|_F^4} \right)^{1/2} \leq \frac{r\left\|\frac{dX(t)}{dt}\right\|_F}{\|X(t)\|_F}.
\end{aligned}$$

Hence,

$$\left\| \frac{d}{dt}\left(\frac{rX(t)}{\|X(t)\|_F}\right)\right\|_\kappa = \sigma_n^{-1}\left(\frac{rX(t)}{\|X(t)\|_F}\right)\left\| \frac{d}{dt}\left(\frac{rX(t)}{\|X(t)\|}\right)\right\|_F$$

$$= \frac{\|X(t)\|_F \sigma_n^{-1}(X(t))}{r}\left\| \frac{d}{dt}\left(\frac{rX(t)}{\|X(t)\|}\right)\right\|_F \leq \sigma_n^{-1}(X(t))\left\|\frac{dX(t)}{dt}\right\|_F = \left\|\frac{dX(t)}{dt}\right\|_\kappa.$$

Therefore $X(t)$ can only be a minimizing geodesic if it belongs to $\mathbb{S}_r(\mathbb{GL}^{>}_{n,m})$. Since all geodesics are locally minimizing geodesics, Corollary 6 follows. ∎

The following gives an example of a smooth and nonself-convex function in $\mathbb{GL}_{n,m}$.

EXAMPLE 1. *For $n \geq 3$, the function $\alpha(A) = \sigma_1(A)^{-2} + \cdots + \sigma_n(A)^{-2}$ is not self-convex in $\mathbb{GL}_{n,m}$.*

Proof. For simplicity we consider the case of real square matrices. We have $\alpha(A) = \|A^{-1}\|_F^2$,

$$D\alpha(A)\dot{A} = -2\langle A^{-1}, A^{-1}\dot{A}A^{-1}\rangle_F = -2\langle A^{-T}A^{-1}A^{-T}, \dot{A}\rangle_F,$$

$$\|D\alpha(A)\|_F^2 = 4\|A^{-T}A^{-1}A^{-T}\|_F^2,$$

$$D^2\alpha(A)(\dot{A},\dot{A}) = 2\|A^{-1}\dot{A}A^{-1}\|_F^2 + 4\langle A^{-1}, A^{-1}\dot{A}A^{-1}\dot{A}A^{-1}\rangle_F.$$

According to Proposition 4, the self-convexity of $\alpha(A)$ in $\mathbb{GL}_n$ is equivalent to

$$2\|A^{-1}\|_F^2\left(2\|A^{-1}\dot{A}A^{-1}\|_F^2 + 4\langle A^{-1}, A^{-1}\dot{A}A^{-1}\dot{A}A^{-1}\rangle_F\right)$$

$$+ 4\|\dot{A}\|_F^2\|A^{-T}A^{-1}A^{-T}\|_F^2 - 8\langle A^{-1}, A^{-1}\dot{A}A^{-1}\rangle_F^2 \geq 0.$$

This inequality is not satisfied when

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix} \text{ and } \dot{A} = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \qquad \square$$

5. The homogeneous linear case.

5.1. The complex projective space. The matter of this subsection is mainly taken from Gallot–Hulin–Lafontaine [6, sec. 2.A.5].

Let V be a Hermitian space of complex dimension $\dim_{\mathbb{C}} V = d+1$. We denote by $\mathbb{P}(V)$ the corresponding projective space, that is, the quotient of $V \setminus \{0\}$ by the group $\mathbb{C}^*$ of dilations of V; $\mathbb{P}(V)$ is equipped with its usual smooth manifold structure with complex dimension $\dim \mathbb{P}(V) = d$. We denote by p the canonical surjection.

Let V be considered as a real vector space of dimension $\dim_{\mathbb{R}} V = 2d+2$ equipped with the scalar product $\operatorname{Re}\langle .,. \rangle_V$. The sphere $\mathbb{S}(V)$ is a submanifold in V of real dimension $2d+1$. This sphere being equipped with the induced metric becomes a Riemannian manifold and, as usual, we identify the tangent space at $z \in \mathbb{S}(V)$ with

$$T_z\mathbb{S}(V) = \{u \in V \ : \ \operatorname{Re}\langle u, z\rangle_V = 0\}.$$

The projective space $\mathbb{P}(V)$ can also be seen as the quotient $\mathbb{S}(V)/S^1$ of the unit sphere in V by the unit circle in $\mathbb{C}$ for the action given by $(\lambda, z) \in S^1 \times \mathbb{S}(V) \to \lambda z \in \mathbb{S}(V)$. The canonical map is denoted by

$$p_V : \mathbb{S}(V) \to \mathbb{P}(V).$$

p_V is the restriction of p to $\mathbb{S}(V)$.

The horizontal space at $z \in \mathbb{S}(V)$ related to p_V is defined as the (real) orthogonal complement of $\ker Dp_V(z)$ in $T_z\mathbb{S}(V)$. This horizontal space is denoted by H_z. Since V is decomposed in the (real) orthogonal sum

$$V = \mathbb{R}z \oplus \mathbb{R}iz \oplus z^{\perp},$$

and since $\ker Dp_V(z) = \mathbb{R}iz$ (the tangent space at z to the circle $S^1 z$) we get

$$H_z = z^{\perp} = \{u \in V \ : \ \langle u, z\rangle = 0\}.$$

There exists on $\mathbb{P}(V)$ a unique Riemannian metric such that p_V is a Riemannian submersion; that is, p_V is a smooth submersion and, for any $z \in \mathbb{S}(V)$, $Dp_V(z)$ is an isometry between H_z and $T_{p(z)}\mathbb{P}(V)$. Thus, for this Riemannian structure, one has

$$\langle Dp_V(z)u, Dp_V(z)v\rangle_{T_{p(z)}\mathbb{P}(V)} = \operatorname{Re}\langle u, v\rangle_V$$

for any $z \in \mathbb{S}(V)$ and $u, v \in H_z$.

PROPOSITION 6. *Let $z \in \mathbb{S}(V)$ be given.*

1. *A chart at $p(z) \in \mathbb{P}(V)$ is defined by*

$$\varphi_z : H_z \to \mathbb{P}(V), \quad \varphi_z(u) = p(z+u).$$

2. *Its derivative at 0 is the restriction of $Dp(z)$ at H_z:*

$$D\varphi_z(0) = Dp(z) : H_z \to T_{p(z)}\mathbb{P}(V),$$

which is an isometry.

3. *For any smooth mapping $\psi : \mathbb{P}(V) \to \mathbb{R}$, and for any $v \in H_z$ we have*

$$D\psi(p(z))\,(Dp(z)v) = D(\psi \circ \varphi_z)(0)v$$

and

$$D^2\psi(p(z))(Dp(z)v, Dp(z)v) = D^2(\psi \circ \varphi_z)(0)(v, v).$$

Proof. 1 and 2 are easy. We have $D(\psi \circ \varphi_z)(0) = D\psi(p(z))D(\varphi_z)(0)$, which gives 3 since $D(\varphi_z)(0)v = Dp(z)v$ for any $v \in H_z$. For the second derivative, recall that $D^2\psi(p(z))(Dp(z)v, Dp(z)v) = (\psi \circ \tilde{\gamma})''(0)$, where $\tilde{\gamma}$ is a geodesic curve in $\mathbb{P}(V)$ such that $\tilde{\gamma}(0) = p(z)$, $\tilde{\gamma}'(0) = Dp(z)v$. Now, consider the horizontal p_V-lift γ of $\tilde{\gamma}$ to $\mathbb{S}(V)$ with base point z. Note that $\gamma(0) = z$, $\gamma'(0) = v$. Hence,

$$(\psi \circ \tilde{\gamma})''(0) = (\psi \circ p \circ \gamma)''(0) = D^2(\psi \circ p)(z)(v, v) + D\psi(p(z))Dp(z)\gamma''(0).$$

As $\gamma''(0)$ is orthogonal to $T_z\mathbb{S}(V)$, we have $Dp(z)\gamma''(0) = 0$. Finally,

$$D^2(\psi \circ p)(z)(v, v) = (\psi \circ p(z + tv))''(0) = (\psi \circ \varphi_z(tv))''(0) = D^2(\psi \circ \varphi_z)(0)(v, v),$$

and the assertion on the second derivative follows. ∎

The following result will be helpful.

PROPOSITION 7. *Let $\mathcal{M}_1, \mathcal{M}_2$ be Riemannian manifolds, and let $\alpha_2 : \mathcal{M}_2 \to]0, \infty[$ be of class C^2. Let $\pi : \mathcal{M}_1 \to \mathcal{M}_2$ be a Riemannian submersion. Let $\mathcal{U}_2 \subseteq \mathcal{M}_2$ be an open set, and let us assume that $\alpha_1 = \alpha_2 \circ \pi$ is self-convex in $\mathcal{U}_1 = \pi^{-1}(\mathcal{U}_2)$. Then, α_2 is self-convex in $\mathcal{U}_2$.*

Proof. Let $\mathcal{M}_{\kappa,1}$ be $\mathcal{M}_1$, but endowed with the condition metric given by α_1, and let $\mathcal{M}_{\kappa,2}$ be $\mathcal{M}_2$, but endowed with the condition metric given by α_2. Then, $\pi : \mathcal{M}_{\kappa,1} \to \mathcal{M}_{\kappa,2}$ is also a Riemannian submersion.

Now, let $\gamma_2 : [a, b] \to \mathcal{U}_2 \subseteq \mathcal{M}_{\kappa,2}$ be a geodesic, and let $\gamma_1 \subseteq \mathcal{M}_{\kappa,1}$ be its horizontal lift by π. Then, γ_1 is a geodesic in $\mathcal{U}_1 \subseteq \mathcal{M}_1$ (see [6, Cor. 2.109]), and hence $\log \alpha_1(\gamma_1(t))$ is a convex function of t. Now,

$$\log(\alpha_2(\gamma_2(t))) = \log(\alpha_2 \circ \pi(\gamma_1(t))) = \log(\alpha_1(\gamma(t)))$$

is convex as wanted. ∎

COROLLARY 7. *The function $\alpha_2 : \mathbb{P}(\mathbb{GL}^{>}_{n,m}) \to \mathbb{R}$, $\alpha_2(A) = \|A\|_F^2 \sigma_n^{-2}(A)$ is self-convex in $\mathbb{P}(\mathbb{GL}^{>}_{n,m})$.*

Proof. Note that $p : \mathbb{S}(\mathbb{GL}^{>}_{n,m}) \to \mathbb{P}(\mathbb{GL}^{>}_{n,m})$ is a Riemannian submersion, and $\alpha_2 = \alpha \circ p$, where α is as in Corollary 6. The corollary follows from Proposition 7. ∎

5.2. The solution variety. Let us denote by p_1 and p_2 the canonical maps

$$\mathbb{S}_1 \stackrel{p_1}{\to} \mathbb{P}\left(\mathbb{K}^{n\times(n+1)}\right) \quad \text{and} \quad \mathbb{S}_2 \stackrel{p_2}{\to} \mathbb{P}\left(\mathbb{K}^{n+1}\right) = \mathbb{P}_n(\mathbb{K}),$$

where $\mathbb{S}_1$ is the unit sphere in $\mathbb{K}^{n\times(n+1)}$ and $\mathbb{S}_2$ is the unit sphere in $\mathbb{K}^{n+1}$. Consider the affine solution variety,

$$\hat{\mathcal{W}}^{>} = \left\{(M, \zeta) \in \mathbb{S}_1 \times \mathbb{S}_2 \ : \ M \in \mathbb{GL}^{>}_{n,n+1} \text{ and } M\zeta = 0\right\}.$$

It is a Riemannian manifold equipped with the metric induced by the product metric on $\mathbb{K}^{n\times(n+1)} \times \mathbb{K}^{n+1}$. The tangent space to $\hat{\mathcal{W}}^{>}$ is given by

$$T_{(M,\zeta)}\hat{\mathcal{W}}^{>} = \left\{(\dot{M}, \dot{\zeta}) \in T_M\mathbb{S}_1 \times T_\zeta\mathbb{S}_2 \ : \ \dot{M}\zeta + M\dot{\zeta} = 0\right\}.$$

The projective solution variety considered here is

$$\mathcal{W}^{>} = \left\{(p_1(M), p_2(\zeta)) \in \mathbb{P}\left(\mathbb{K}^{n\times(n+1)}\right) \times \mathbb{P}_n(\mathbb{K}) \ : \ M \in \mathbb{GL}^{>}_{n,n+1} \text{ and } M\zeta = 0\right\};$$

that is, also a Riemannian manifold equipped with the metric induced by the product metric on $\mathbb{P}\left(\mathbb{K}^{n\times(n+1)}\right) \times \mathbb{P}_n(\mathbb{K})$.

Let us denote by π_1 the restriction to $\hat{\mathcal{W}}^>$ of the first projection $\mathbb{S}_1 \times \mathbb{S}_2 \to \mathbb{S}_1$, and by $R : \hat{\mathcal{W}}^> \to \mathbb{R}$, $R = \sigma_n \circ \pi_1$. We have the following lemma.

LEMMA 2. *Let $w = (M, \zeta) \in \hat{\mathcal{W}}^>$, and let γ be a geodesic in $\hat{\mathcal{W}}^>$, $\gamma(0) = w$. Then,*

$$D\sigma_n(\pi_1(w))(\pi_1 \circ \gamma)''(0) < 0.$$

Proof. Our problem is invariant by unitary change of coordinates. Hence, using a singular value decomposition, we can assume that $M = (\Sigma, 0) \in \mathbb{GL}^>_{n,n+1}$, where $\Sigma = \operatorname{diag}(\sigma_1 \geq \cdots \geq \sigma_{n-1} > \sigma_n) \in \mathbb{K}^{n\times n}$ and $\zeta = e_{n+1} = (0, \ldots, 0, 1)^T \in \mathbb{S}_2$. As $\gamma = (M(t), \zeta(t))$ is a geodesic of $\hat{\mathcal{W}}^> \subseteq \mathbb{K}^{n\times(n+1)} \times \mathbb{K}^n$, $\gamma''(0)$ is orthogonal to $T_w\hat{\mathcal{W}}$, which contains all the pairs of the form $((A, 0), 0)$ where A is a $n \times n$ matrix, $\operatorname{Re}\langle \Sigma, A\rangle = 0$. Hence, $M''(0)$ has the form

$$M''(0) = (a\Sigma, *)$$

for some real number $a \in \mathbb{R}$. Finally, $M(t)$ is contained in the sphere so $\|M(t)\|_F = 1$ and

$$0 = (||M(t)||_F^2)''(0) = 2||M'(0)||_F^2 + 2\operatorname{Re}\langle M(0), M''(0)\rangle = 2||M'(0)||_F^2 + 2a$$

so that $a = -\|M'(0)\|_F^2$ and $(M''(0))_{nn} = -\|M'(0)\|_F^2\sigma_n$. From Proposition 5,

$$D\sigma_n(\pi_1(w))(\pi_1 \circ \gamma)''(0) = \operatorname{Re}((\pi_1 \circ \gamma)''(0)_{nn}) = \operatorname{Re}(M''(0))_{nn} < 0. \qquad \square$$

THEOREM 3. *The map $\alpha : \hat{\mathcal{W}}^> \to \mathbb{R}$ given by $\alpha(M, \zeta) = \sigma_n(M)^{-2}$ is self-convex.*

Proof. Using unitary invariance we can take $M = (\Sigma, 0) \in \mathbb{GL}^>_{n,n+1}$, where $\Sigma = \operatorname{diag}(\sigma_1 \geq \cdots \geq \sigma_{n-1} > \sigma_n) \in \mathbb{K}^{n\times n}$ and $\zeta = e_{n+1} = (0, \ldots, 0, 1)^T \in \mathbb{S}_2$. According to Proposition 3 we have to prove that

$$2\,\|\dot{w}\|_w^2\,\|DR(w)\|^2 \geq D^2R^2(w)(\dot{w}, \dot{w})$$

for every $w \in \hat{\mathcal{W}}^>$ and $\dot{w} \in T_w\hat{\mathcal{W}}^>$. From Proposition 5 we have

$$DR(w)\dot{w} = D\sigma_n(\pi_1(w))(D\pi_1(w)\dot{w}) = \operatorname{Re}(D\pi_1(w)\dot{w})_{nn},$$

so that $\|DR(w)\| = 1$. On the other hand, assume that $\dot{w} \neq 0$, and let γ be a geodesic in $\hat{\mathcal{W}}^>$, $\gamma(0) = w$, $\dot{\gamma}(0) = \dot{w}$. From Lemma 2,

$$\begin{aligned} D^2R^2(w)(\dot{w}, \dot{w}) &= (\sigma_n^2 \circ \pi_1 \circ \gamma)''(0) \\ &= D^2\sigma_n^2(\pi_1(w))(D\pi_1(w)\dot{w}, D\pi_1(w)\dot{w}) + 2\sigma_n D\sigma_n(\pi_1(w))(\pi_1 \circ \gamma)''(0) \\ &< D^2\sigma_n^2(\pi_1(w))(D\pi_1(w)(\dot{w}), D\pi_1(w)(\dot{w})). \end{aligned}$$

Thus, we have to prove that for $\dot{y} \in \mathbb{K}^{n\times(n+1)}$,

$$2\,\|\dot{y}\|^2 \geq D^2\sigma_n^2(\pi_1(w))(\dot{y}, \dot{y}),$$

which is a consequence of our Proposition 5. $\square$

COROLLARY 8. *The map $\alpha_2 : \mathcal{W}^> \to \mathbb{R}$ given by $\alpha_2(M, \zeta) = \|M\|_F^2/\sigma_n^2(M)$ is self-convex.*

Proof. Consider the Riemannian submersion

$$p_1 \times p_2 : \mathbb{S}_1 \times \mathbb{S}_2 \longrightarrow \mathbb{P}\left(\mathbb{K}^{n\times(n+1)}\right) \times \mathbb{P}_n(\mathbb{K}),\ p_1 \times p_2(M, \zeta) = (p_1(M), p_2(\zeta)).$$

Note that $T_{(M,\zeta)}\hat{\mathcal{W}}^>$ contains the kernel of the derivative $D(p_1 \times p_2)(M, \zeta)$. Thus, the restriction $p_1 \times p_2 : \hat{\mathcal{W}}^> \to \mathcal{W}^>$ is also a Riemannian submersion. The corollary follows combining Proposition 7 and Theorem 2. $\square$

6. Self-convexity of the distance from a submanifold of $\mathbb{R}^j$. Let $\mathcal{N}$ be a C^k submanifold without boundary $\mathcal{N} \subset \mathbb{R}^j$, $k \geq 2$. Let us denote by

$$\rho(x) = d(x, \mathcal{N}) = \inf_{y \in \mathcal{N}} \|x - y\|$$

the distance from $\mathcal{N}$ to $x \in \mathbb{R}^j$ (here $d(x, y) = \|x - y\|$ denotes the Euclidean distance). Let $\mathcal{U}$ be the largest open set in $\mathbb{R}^j$ such that, for any $x \in \mathcal{U}$, there is a unique closest point from $\mathcal{N}$ to x. This point is denoted by $K(x)$ so that we have a map defined by

$$K : \mathcal{U} \to \mathcal{N}, \ \rho(x) = d(x, K(x)).$$

Classical properties of ρ and K are given in the following proposition (see also Foote [5] and Li and Nirenberg [7]).

PROPOSITION 8.

1. *ρ is defined and $1-$Lipschitz on $\mathbb{R}^j$,*
2. *for any $x \in \mathcal{U}$, $x - K(x)$ is a vector normal to $\mathcal{N}$ at $K(x)$, i.e., $x - K(x) \in \left(T_{K(x)}\mathcal{N}\right)^{\perp}$,*
3. *K is C^{k-1} on $\mathcal{U}$,*
4. *ρ^2 is C^k on $\mathcal{U}$, $D\rho^2(x)\dot{x} = 2\langle x - K(x), \dot{x}\rangle$, and $D^2\rho^2(x)(\dot{x}, \dot{x}) = 2\|\dot{x}\|^2 - 2\langle DK(x)\dot{x}, \dot{x}\rangle$,*
5. *ρ is C^k on $\mathcal{U} \setminus \mathcal{N}$,*
6. *$\langle DK(x)\dot{x}, \dot{x}\rangle \geq 0$ for every $x \in \mathcal{U}$ and $\dot{x} \in \mathbb{R}^j$.*

Proof.

1. For any x and y one has $\rho(x) = d(x, K(x)) \leq d(x, K(y)) \leq d(x, y) + d(y, K(y)) = d(x, y) + \rho(y)$. Since x and y play a symmetric role we get $|\rho(x) - \rho(y)| \leq d(x, y)$.
2. This is the classical first order optimality condition in optimization.
3. This classical result may be derived from the inverse function theorem applied to the canonical map defined on the normal bundle to $\mathcal{N}$

$$\text{can} : \text{N}\mathcal{N} \to \mathbb{R}^j, \ \text{can}(y, n) = y + n$$

for every $y \in \mathcal{N}$ and $n \in N_y\mathcal{N} = (T_y\mathcal{N})^{\perp}$. The normal bundle is a C^{k-1} manifold, the canonical map is a C^{k-1} diffeomorphism when restricted to the set $\{(y, n) : y + tn \in \mathcal{U} \text{ for all } 0 \leq t \leq 1\}$, and $K(x)$ is easily given from can^{-1}.
4. The derivative of ρ^2 is equal to $D\rho^2(x)\dot{x} = 2\langle x - K(x), \dot{x} - DK(x)\dot{x}\rangle = 2\langle x - K(x), \dot{x}\rangle$ because $DK(x)\dot{x} \in T_{K(x)}\mathcal{N}$ and $x - K(x) \in \left(T_{K(x)}\mathcal{N}\right)^{\perp}$. Thus $\nabla\rho^2(x) = 2(x - K(x))$ is C^{k-1} on $\mathcal{U}$ so that ρ^2 is C^k. The formula for $D^2\rho^2$ follows.
5. This step is obvious.
6. Let $x(t)$ be a curve in $\mathcal{U}$ with $x(0) = x$. Let us denote $\frac{dx(t)}{dt} = \dot{x}(t)$, $\frac{d^2x(t)}{dt^2} = \ddot{x}(t)$, $y(t) = K(x(t))$, $\frac{dy(t)}{dt} = \dot{y}(t)$, and $\frac{d^2y(t)}{dt^2} = \ddot{y}(t)$. From the first order optimality condition we get

$$\langle x(t) - y(t), \dot{y}(t)\rangle = 0$$

whose derivative at $t = 0$ is

$$\langle \dot{x} - \dot{y}, \dot{y}\rangle + \langle x - y, \ddot{y}\rangle = 0.$$

Thus

$$\langle DK(x)\dot{x},\dot{x}\rangle = \langle \dot{y},\dot{x}\rangle = \langle \dot{y},\dot{y}\rangle - \langle x-y,\ddot{y}\rangle .$$

This last quantity is equal to $\frac{1}{2}\frac{d^2}{dt^2}\|x-y(t)\|^2\big|_{t=0}$. It is nonnegative by the second order optimality condition. □

Proofs of Theorem 2 and Corollary 1. We are now able to prove our second main theorem. Let us denote $\alpha(x) = 1/\rho(x)^2$. We shall prove that α is self-convex on $\mathcal{U}$. From Proposition 3 it suffices to prove that, for every $\dot{x} \in \mathbb{R}^j$,

$$2\|\dot{x}\|^2\|D\rho(x)\|^2 \geq D^2\rho^2(x)(\dot{x},\dot{x})$$

or, according to assertion 4 of Proposition 8 and $\|D\rho\| = 1$, that

$$2\|\dot{x}\|^2 \geq 2\|\dot{x}\|^2 - 2\langle DK(x)\dot{x},\dot{x}\rangle .$$

This is obvious from assertion 4 of Proposition 8.

Now we prove Corollary 1. Let $\mathbb{S}_1(\mathbb{R}^j)$ be the sphere of radius 1 in $\mathbb{R}^j$, and let $p_{\mathbb{R}^j}$ denote the canonical projection $p_{\mathbb{R}^j} : \mathbb{R}^j \to \mathbb{P}(\mathbb{R}^j)$. Note that the preimage of $\mathcal{N}$ by $p_{\mathbb{R}^j}$ satisfies

$$d(y, p_{\mathbb{R}^j}^{-1}(\mathcal{N})) = d_{\mathbb{P}}(p_{\mathbb{R}^j}(y),\mathcal{N})\|y\|.$$

As in the proof of Corollary 6, the mapping $1/\rho(x)^2$ is self-convex in the set $\mathbb{S}_1(\mathbb{R}^j) \cap p_{\mathbb{R}^j}^{-1}(\mathcal{U})$. Now, apply Proposition 7 to the Riemannian submersion $p_{\mathbb{R}^j}$ to conclude the corollary. □

Two examples.

EXAMPLE 2. *Take $\mathcal{U}$ the unit disk in $\mathbb{R}^2$ and $\mathcal{N}$ the unit circle. The corresponding function is given by*

$$\alpha(x) = d(x,\mathcal{N})^{-2} = 1/\left(1-\|x\|\right)^2 .$$

According to Theorem 2, the map $\log\alpha(x)$ is convex along the condition geodesics in

$$\mathcal{U}\setminus\{(0,0)\} = \left\{x\in\mathbb{R}^2 \ : \ 0<\|x\|<1\right\}.$$

This property also holds in $\mathcal{U}$: a geodesic through the origin is a ray $x(t) = (-1+e^t)(\cos\theta,\sin\theta)$ when $-\infty < t \leq 0$, and $x(t) = (1-e^{-t})(\cos\theta,\sin\theta)$ when $0 \leq t < \infty$ for some θ. In that case

$$\log\alpha(x(t)) = 2\,|t|\,,$$

which is convex.

EXAMPLE 3. *Take $\mathcal{N} \subset \mathbb{R}^2$ equal to the union of the two points $(-1,0)$ and $(1,0)$. In that case*

$$\alpha(x)^{-1} = d(x,\mathcal{N})^2 = \min\left((1+x_1)^2+x_2^2,(1-x_1)^2+x_2^2\right).$$

It may be shown that for any $0 < a \leq 1/10$, the straight line segment is the only minimizing geodesic joining the points $(0,-a)$ and $(0,a)$. Since $\log\alpha(0,t) = -\log(1+t^2)$ has a maximum at $t=0$, $g(t)$, $-a \leq t \leq a$ cannot be log-convex. Here $\{0\}\times\mathbb{R}$ is equal to the locus in $\mathbb{R}^2$ of points equally distant from the two nodes, which is the set we avoid in Theorem 2.

REFERENCES

[1] C. Beltrán and M. Shub, *Complexity of Bézout's theorem* VII: *Distance estimates in the condition metric*, Found. Comput. Math., 9 (2009), pp. 179–195.

[2] P. Boito and J.-P. Dedieu, *The condition metric in the space of full rank rectangular Matrices*, available online at http://www.math.univ-toulouse.fr/~dedieu/Boito-Dedieu-future.pdf, SIAM J. Matrix Anal. Appl., to appear.

[3] F. H. Clarke, *Optimization and Nonsmooth Analysis*, 2nd ed., Les Publications CRM, Montreal, 1989.

[4] J. W. Demmel, *The probability that a numerical analysis problem is difficult*, Math. Comput., 50 (1988), pp. 449–480.

[5] R. Foote, *Regularity of the distance function*, Proc. Amer. Math. Soc., 92 (1984), pp. 153–155.

[6] S. Gallot, D. Hulin, and J. Lafontaine, *Riemannian Geometry*, 3rd ed., Springer-Verlag, Berlin, 2004.

[7] Y. Li and L. Nirenberg, *Regularity of the distance function to the boundary*, Rend. Accad. Naz. Sci. XL Mem. Mat. Appl. (5), 29 (2005) pp. 257–264.

[8] M. Shub, *Complexity of Bézout's theorem* VI: *Geodesics in the condition metric*, Found. Comput. Math., 9 (2009), pp. 171–178.

[9] C. Udrişte, *Convex Functions and Optimization Methods on Riemannian Manifolds*, Kluwer Academic Publishers, Dordrecht, The Netherlands, 1994.