Author Proof

Opening the Parallelogram: Considerations on Non-Euclidean Analogies

Pierre-Alexandre Murena[1(✉)], Antoine Cornuéjols[2], and Jean-Louis Dessalles[1]

[1] Télécom ParisTech - Université Paris Saclay, 46 rue Barrault, 75013 Paris, France
last@telecom-paristech.fr

[2] UMR MIA-Paris, AgroParisTech, INRA, Université Paris Saclay, 5 rue Claude Bernard, 75005 Paris, France
antoine.cornuejols@agroparistech.fr

Abstract. Analogical reasoning is a cognitively fundamental way of reasoning by comparing two pairs of elements. Several computational approaches are proposed to efficiently solve analogies: among them, a large number of practical methods rely on either a parallelogram representation of the analogy or, equivalently, a model of proportional analogy. In this paper, we propose to broaden this view by extending the parallelogram representation to differential manifolds, hence spaces where the notion of vectors does not exist. We show that, in this context, some classical properties of analogies do not hold any longer. We illustrate our considerations with two examples: analogies on a sphere and analogies on probability distribution manifold.

AQ1 AQ2

Keywords: Analogy · Non-Euclidean geometry

1 Introduction

Making analogies is considered by psychologists as a basic cognitive ability of human beings [8], yet it remains a challenging task for artificial intelligence. An analogy designates a situation where a parallel can be drawn between two distinct and *a priori* unrelated domains. Computational models of analogical reasoning have been developed either to map semantic domains [6], to solve analogical problems on character strings, either structured [10] or unstructured [12], or to characterize the quality of an analogy [5]. Apart from its major cognitive interpretation, analogy plays an important role in case-based reasoning (CBR) [1]: In order to solve a new case, CBR focuses on previously encountered cases and aims to adapt solutions to the new problem. This adaptation process can be interpreted as *one-domain* analogical reasoning (which means that the source and target domains are identical).

A classical representation of analogies between vectors is the *parallelogram model*, which states that the four elements of the analogy obey a regularity rule

M. T. Cox et al. (Eds.): ICCBR 2018, LNAI 11156, pp. 1–15, 2018.
https://doi.org/10.1007/978-3-030-01081-2_39

close to a parallelogram in the representation space. For instance, the analogy "Paris is to France what Stockholm is to Sweden" may be interpreted in the form of the equality "Paris - France + Sweden = Stockholm". The first occurrences of this representation date back to the earliest researches [17] and have been resurrected in the recent years throughout the neural networks representation skills, in particular in the `Word2Vec` paradigm [15], [14] or even in visual object categorization [11]. The idea can be summed up as follows: Considering the analogical equation $A : B :: C : x$ with variable x, we assume that each element can be represented as a point in a Euclidean space and that the solution x is defined as the vector $x = C + B - A$. This representation is consistent with the axioms of analogical proportion, as shown in Sect. 2.4 of [13].

In this paper, we address the question of what happens when this representation is not true. Our main point is to loosen the structure of the representation space and to consider analogies on *Riemannian manifolds* instead of analogies in Euclidean spaces. A manifold can be understood intuitively as a space which is almost a Euclidean space, in the sense that it is locally Euclidean. Because of their curvature, the notion of vector does not exist in differential manifolds, hence the parallelogram representation is not valid in them. A way to get around that is to consider the notions of *geodesic curve* and *parallel transport* in Riemannian manifolds which allow one to build parallelogram-like shapes. These notions will be explained with more details in Sect. 2.4. We will show that the parallelogram construction is a particular case of the proposed procedure for Euclidean manifolds, but that non-Euclidean structures do not verify the classical axioms of analogy with this setting.

The remainder of this article is organized as follows. In Sect. 2, we present the general problem of analogies in non-Euclidean spaces. The problem is introduced with the help of a trivial example (analogies on spheres), but a more general explanation follows. In particular, we discuss the link between a found analogical dissimilarity and manifold curvature. In Sect. 3, we propose an application of the proposed theory in the case of a very particular space: the space of normal distributions. We will illustrate the developed ideas through a couple of simulations which show the impact of curvature. Lastly, we propose a discussion on proportional analogy in differential manifolds.

2 Non-Euclidean Spaces and Non-commutative Analogies

2.1 Intuition: Analogies on the Sphere $\mathbb{S}^2$

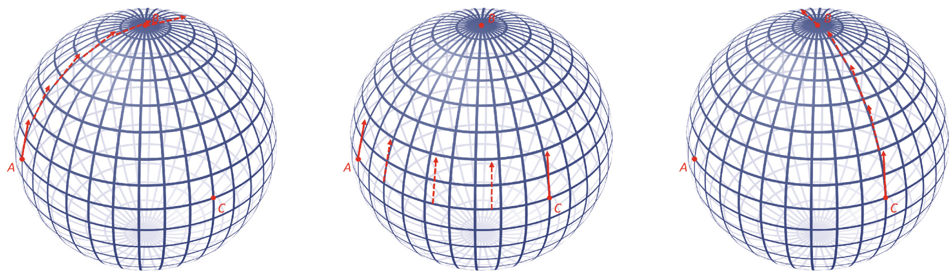
In order to understand the ideas at play, we propose to consider the example of analogies on a sphere. We denote by $\mathbb{S}^2$ the sphere defined as the subset of $\mathbb{R}^3$ defined as $\mathbb{S}^2 = \{x | x_1^2 + x_2^2 + x_3^2 = 1\}$. The sphere can be shown to be a differential manifold, and is obviously not Euclidean.

We consider three points A, B and C on the sphere and we try to solve the analogical equation $A : B :: C : x$. In the context of this example, we will consider three specific points, but the conclusions we will draw would be the same for any 3 points which are "not aligned" (in the sense that the third point is not on the shortest path between the two others).

In order to solve this analogy, an intuitive idea would be to apply the same procedure as described by the parallelogram rule. On Earth, it is possible to use the parallelogram rule directly on a small scale: Since Earth is locally flat, we can consider the floor as a vector space and apply a parallelogram rule by walking from A to C by keeping in mind the direction to go to B from A.

The same procedure can be used when the three points are very distant. In mathematical terms, we can formulate this procedure as a three steps method:

1. **Direction finding:** Estimation of the direction d to reach B from A following a geodesic (i.e. a path of minimal length).
2. **Parallel transport:** The direction vector is transported along the geodesic from A to C.
3. **Geodesic shooting:** Point D is reached by following the transported direction d' from point C.



(a) Step 1: Direction finding (b) Step 2: Parallel transport (c) Step 3: Geodesic shooting

Fig. 1. Step by step resolution of the analogical equation $A : B :: C : x$ on the sphere $\mathbb{S}^2$. The solution found is $x = B$.

We consider for instance the case where B corresponds to the North pole and A and C are located on the equator. For simplicity purpose, we also suppose that the angle between the locations of A and C in the equator plane is $\pi/2$. The solution to this analogy is shown in Fig. 1.

The steps can be intuitively explained as follows. The first step consists in finding the shortest path from A to B: this path is characterized by the initial direction, which is mathematically encoded by a vector in the tangent space. The second step is of a different nature: The idea is to go along the shortest path from A to C while maintaining the initial direction vector "in the same direction" (the exact mathematical terminology will be precised in the next section). As an illustration of this, the second step can be seen as walking from A to C while maintaining one's nose parallel from one position to the other. The shortest path from A to C in our example is the equator and the initial direction is the vector pointing toward the North pole: Hence, step 2 is similar to walking from A to

C along the equator with the nose pointing toward the North pole at any time. The third step consists in following the transported initial direction the same time as done to join B from A in step 1.

Using this procedure, the solution of the analogical equation $A : B :: C : x$ is $x = B$. With the same procedure applied to the analogical equation $A : C :: B : x$, we obtain the solution $x = C$, which is in contradiction with the *exchange of the means* property of analogical proportion. However, we can easily verify that the other properties are verified:

- Symmetry of the 'as' relation: $C : B :: A : B$ and $B : C :: A : C$
- Determinism: the solution of $A : A :: B : x$ is $x = B$

In the following, we will call a **Non-commutative Analogy** an analogy which satisfies the symmetry of the 'as' relation and the determinism property, but not necessarily the exchange of the means. An analogical proportion is a more constrained case of a non-commutative analogy.

2.2 Reminder: Riemannian Geometry

In order to understand our method, we have to introduce some standard definitions of Riemannian geometry. The proposed definitions are not entirely detailed: we refer interested readers to standard references [3] for more details. For each notion, we propose an intuitive and less rigorous explanation.

A *topological manifold* of dimension d is a connected paracompact Hausdorff space for which every point has an open neighborhood U that is homeomorphic to an open subset of $\mathbb{R}^d$ (such a homeomorphism is called a *chart*). A manifold is called *differentiable* when the chart transitions are differentiable, which means that the mapping from one chart representation to another is smooth. Intuitively, a manifold can be seen as a space that is locally close a vector space.

A tangent vector ξ_x to a manifold $\mathcal{M}$ at point x can be defined as the equivalence class of differentiable curves γ such that $\gamma(0) = x$ modulo a first-order contact condition between curves. It can be interpreted as a "direction" from the point x (which only makes sense when $\mathcal{M}$ is a subset of a vector space). The set of all tangent vectors to $\mathcal{M}$ at x is denoted $T_x\mathcal{M}$ and called tangent space to $\mathcal{M}$ at x. The tangent space can be shown to have a vector space structure. When the tangent spaces $T_x\mathcal{M}$ are equipped with an inner-product g_x which varies smoothly from point to point, $\mathcal{M}$ is called a *Riemannian manifold*.

We define a connection ∇ as a mapping $C^\infty(T\mathcal{M}) \times C^\infty(T\mathcal{M}) \to C^\infty(T\mathcal{M})$ satisfying three properties that are not detailed here: A connection can be seen as a directional derivative of vector fields over the tangent space. It measures the way a tangent vector is modified when moving from one point to another in a given direction. A special connection, called the *Levi-Civita connection*, is defined as an intrinsic property of the Riemannian manifold which depends on its metric g only. This connection follows the "shape" of the manifold (here, the word shape is understood in its intuitive meaning).

These tools are used to define two notions that are fundamental in our interpretation of non-commutative analogies: parallel transport and geodesics. Let

$(\mathcal{M}, g)$ be a Riemannian manifold and let $\gamma : [0,1] \to \mathcal{M}$ be a smooth curve on $\mathcal{M}$. The curve γ is called a geodesic if $\nabla_{\dot\gamma}\dot\gamma = 0$ (which means that γ is auto-parallel, or keeps its tangent vector pointing "in the same direction" at any point). This definition of a geodesic can be shown to correspond to a minimum length curve between two points.

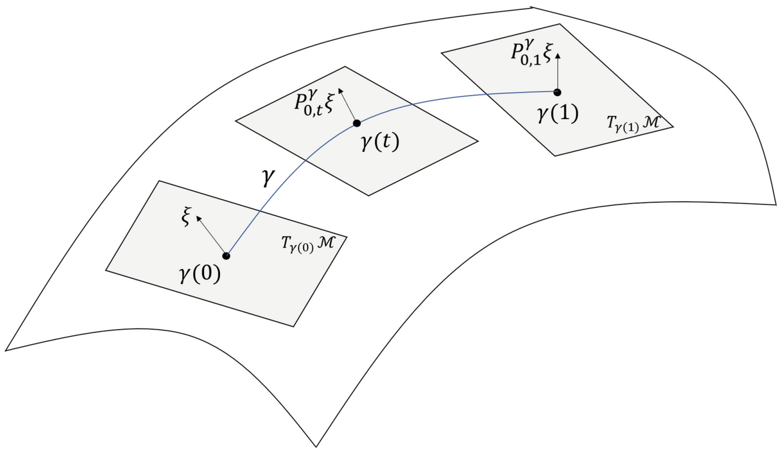


Fig. 2. Illustration of parallel transport on a differential manifolds. Vector ξ is transported along a curve γ. At any position t, we have $P^{\gamma}_{0,t}\xi \in T_{\gamma(t)}\mathcal{M}$.

A vector field X along γ is said to be parallel if $\nabla_{\dot\gamma}X = 0$. One can define the *parallel transport* as the application $P^{\gamma}_{t_0,t} : T_{\gamma(t_0)}\mathcal{M} \to T_{\gamma(t)}\mathcal{M}$ which maps any vector of the tangent space ξ at point $\gamma(t_0)$ to the corresponding value at $\gamma(t)$ for the parallel vector field X such that $X(\gamma(t_0)) = \xi$ (Fig. 2). Intuitively, the parallel transport along a curve keeps a tangent vector "pointing in the same direction".

2.3 Non-commutative Analogies

Following the ideas developed in Sect. 2.1, we propose the following definition for a non-commutative analogical proportion:

Definition 1. *A non-commutative analogy on a set X is a relation on X^4 such that, for every 4-uple $(A, B, C, D) \in X^4$, the following properties are observed:*

- *Symmetry of the 'as' relation: $R(A, B, C, D) \Leftrightarrow R(C, D, A, B)$*
- *Determinism: $R(A, B, A, x) \Rightarrow x = B$*

The second axiom (determinism) is slightly different from the original analogical proportion. For analogical proportion, two possible implications could be used to characterize determinism (the second characterization being the implication $R(A, A, B, x) \Rightarrow x = B$): One being true, the other is a consequence of the first. In non-commutative analogy, these two implications are not equivalent anymore.

Removing the exchange of the means from the definition of an analogy actually makes sense. The symmetry of the means operates in the cross-domain dimension of the analogy: Keeping this observation in mind, the symmetry of the means seems to be a natural property. In practice, it can be observed that the property is perceived as less natural in many examples. Consider for instance the well-known analogy "The sun is to the planets as the nucleus is to the electrons". The symmetrized version of this analogy is "The sun is to the nucleus as the planets are to the electrons", which is less understandable than the original analogy.

Moreover, many examples of common analogies can be found that do not satisfy this property. For instance, the analogy "Cuba is to the USA as North Korea is to China", which is based on a comparison of politics and geographic proximity, while the symmetrized analogy "Cuba is to North Korea as the USA are to China" does not make sense. In this example, the status of the terms is different: In one direction, the analogy is based on a political comparison, while in the other direction it is based on a large-scale geographical comparison. The nature of these two domains is not the same and does not have the same weight in the analogy. This intuition of a directional weighting is coherent with the model of non-euclidean manifolds.

2.4 Non-commutative Analogies on Riemannian Manifolds

Let $\mathcal{M}$ be a Riemannian manifold and $A, B, C, D \in \mathcal{M}$. We propose to find a geometric condition on the four points such that $A : B :: C : D$ defines a non-commutative analogy.

Definition 2. *The parallelogramoid algorithm $\mathcal{A}_p : \mathcal{M}^3 \mapsto \mathcal{M}$ is defined as follows. Consider $(A, B, C) \in \mathcal{M}^3$. Let $\gamma_1 : [0, 1] \to \mathcal{M}$ be a geodesic curve such that $\gamma_1(0) = A$ and $\gamma_1(1) = B$. Let $\xi \in T_A\mathcal{M}$ such that $\xi = \dot{\gamma_1}(0)$. Consider a geodesic curve $\gamma_2 : [0, 1] \to \mathcal{M}$ such that $\gamma_2(0) = A$ and $\gamma_2(1) = C$. Let γ_3 be the geodesic defined by $\gamma_3(0) = C$ and $\dot{\gamma_3}(0) = P_{0,1}^{\gamma_2}\xi$. Then $\mathcal{A}_p(A, B, C) = \gamma_3(1)$.*

Algorithm $\mathcal{A}_p$ corresponds to the procedure used in the case of a sphere. In general, the described procedure is not unique: The unicity of tangent vector ξ is not guaranteed. For instance, in the case of the sphere, if A and B correspond to the North and South poles, there exists an infinite number of such vectors ξ.

Theorem 1. *The relation $R(A, B, C, D) \vDash (\mathcal{A}_p(A, B, C) = D)$ defines a non-commutative analogy on $\mathcal{M}$.*

Proof. We would like to show that $C : D :: A : B$ (symmetry axiom) is correct with our construction. We use the tilde notation to describe the curves for this analogy. For instance, $\tilde{\gamma_1}$ is the geodesic from C to D, hence $\tilde{\gamma_1} = \gamma_3$. Similarly, $\tilde{\gamma_2} = -\gamma_2$, where $-\gamma$ designates the "opposite curve" (i.e. $\tilde{\gamma_2}(s) = \gamma(1-s)$). Since parallel transport is invertible, $\xi = P_{0,1}^{\tilde{\gamma_2}} P_{0,1}^{\gamma_2}\xi$. Thus, $\tilde{\gamma_3}$ is the geodesic curve such that $\tilde{\gamma_3}(0) = A$ and $\dot{\tilde{\gamma_3}}(0) = \xi$ and consequently $\tilde{\gamma_3}(1) = B$ (Fig. 3).

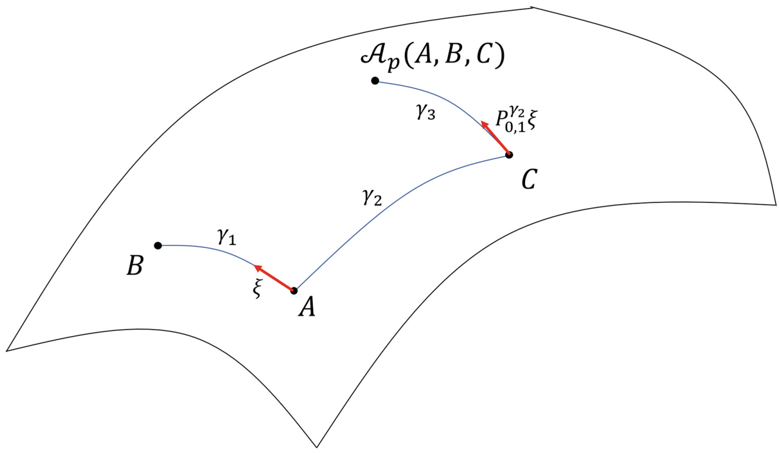


Fig. 3. parallelogramoid procedure on a Riemannian manifold.

AQ3

In general, the relation does not define a proportional analogy since symmetry of the means does not hold: $\mathcal{A}_p(A, B, C) \neq \mathcal{A}_p(A, C, B)$. We will show that we have equality only for a specific metric, called *flat metric*.

When $\mathcal{M} = \mathbb{R}^n$ endowed with the canonical inner-product, the proposed construction can be shown to be equivalent to the usual parallelogram rule, since a geodesic is defined as a straight line and parallel transport over a straight line is a simple translation of the original tangent vector. It can be shown that the converse is almost true: The manifolds for which $\mathcal{A}_p$ designs an analogical proportion have their *Ricci curvature* vanishing at any point.

Theorem 2. *The only Riemannian metrics g such that the relation defined by $R(A, B, C, D) \vDash (\mathcal{A}_p(A, B, C) = D)$ is an analogical proportion for any A, B and C are Ricci-flat.*

Proof. In this demonstration, we will consider the equivalent problem where we are given $A \in \mathcal{M}$ and $\xi_1, \xi_2 \in T_A\mathcal{M}$. With these notations, $B = \gamma_1(1)$ and $C = \gamma_2(1)$ where γ_1 is the geodesic drawn from A with initial vector ξ_1 and γ_2 is the geodesic drawn from A with initial vector ξ_2. Considering an infinitesimal parallelogramoid as defined in Definition 1.1 of [16], where δ is the distance between A and B, and ϵ the distance between A and C. Then the distance between C and $D = \mathcal{A}_p(A, B, C)$ is equal to

$$d = \delta\left(1 - \frac{\epsilon^2}{2}K(v, w) + \mathcal{O}(\epsilon^3 + \epsilon^2\delta)\right)$$

where $K(v, w)$ is the sectional curvature in directions (v, w). In the case of analogical proportion, it can be verified that distance d must be equal to δ. Thus, we have necessarily $K(v, w) = 0$ and, by construction of Ricci curvature $Ric(v)$ as the average value of $K(v, w)$ when w runs over the unit sphere, we have the result.

Obviously, Euclidean spaces endowed with the canonical vector space are Ricci-flat, but there exists other Ricci-flat spaces. A direct consequence of Theorem 2 is that analogical proportions can be defined on some differential manifolds. We will discuss in Sect. 4 the general existence of such relations on general manifolds.

3 Application: Analogies on Fisher Manifold

3.1 Fisher Manifold

By definition, a parametric family of probability distributions $(p_\theta)_\theta$ has a natural structure of a differential manifold and, in this context, is called *statistical manifold*. Unless in general a manifold is not associated to a notion of distance or metric, *information geometry* states that there exists only one natural metric for statistical manifolds [4]. This metric, called *Fisher metric*, is defined as follows [7]:

$$g_{ab}(\theta) = \int p(x|\theta) \frac{\partial \log p(x|\theta)}{\partial \theta^a} \frac{\partial \log p(x|\theta)}{\partial \theta^b} dx \tag{1}$$

It can be related to the variance of the relative difference between one distribution $p(x|\theta)$ and a neighbour $p(x|\theta + d\theta)$. For a more complete introduction to Fisher manifolds and more precise explanations on the nature of Fisher metric, we refer the reader to [2].

Among all possible statistical manifolds, we focus on the set of normal distributions, denoted by $\mathcal{N}(n)$. A complete description of the geometric nature of $\mathcal{N}(n)$ is given in [18]. As mentioned in this paper, a geodesic curve $(\mu(t), \Sigma(t))$ on $\mathcal{N}(n)$ is described by the following geodesic equation:

$$\begin{cases} \ddot{\Sigma} + \dot{\mu}\dot{\mu}^T - \dot{\Sigma}\Sigma^{-1}\dot{\Sigma} = 0 \\ \ddot{\mu} - \dot{\Sigma}\Sigma^{-1}\dot{\mu} = 0 \end{cases} \tag{2}$$

In order to apply the parallelogrammoid algorithm and find non-commutative analogies on $\mathcal{N}(n)$, a fundamental issue has to be overcome. As explained in the reminder on Riemannian geometry, there exists two equivalent definitions of geodesic curves:

1. A geodesic can be interpreted as a curve of shortest length between two points. It is described by two points (A, B).
2. A geodesic can be interpreted as an auto-parallel curve, hence a curve generated by the parallel transport of its celerity. It is described by the initial state: the initial position $A \in \mathcal{M}$ and the initial celerity $\xi \in \mathcal{T}_A\mathcal{M}$.

These two definitions are equivalent but switching from the one to the other is a complex task in general. The second definition offers a simple computational model for geodesic shooting, since it corresponds to integrating a differential equation (Eq. 2 in our case), but using it to find a geodesic between two points requires to find initial celerity ξ.

In the scope of this paper, we consider the algorithm for minimal geodesic on $\mathcal{N}(n)$ proposed by [9]. The proposed algorithm is based on the simple idea to shoot a geodesic using initial celerity ξ using Eq. 2 and to update ξ based on the euclidean difference between the endpoint of the integrated curve and the actual expected endpoint. The algorithm is empirically shown to converge for lower dimensions ($n = 2$ or $n = 3$).

3.2 Experimental Results

We present the results of the parallelogrammoid procedures $\mathcal{A}_p(A, B, C)$ an $\mathcal{A}_p(A, C, B)$ obtained for various bidimensional multinormal distributions. We use the classical representation of the multivariate normal distributions by the isocontour of its covariance matrix, centered at the mean of the distribution. The results we display are presented as follows:

- In black: Intermediate points in the trajectories γ_1, γ_2 and γ_3.
- In blue: Normal distribution A.
- In green: Normal distribution B.
- In cyan: Normal distribution C.
- In red: Normal distribution $D_1 = \mathcal{A}_p(A, B, C)$.
- In magenta: Normal distribution $D_2 = \mathcal{A}_p(A, C, B)$.

Case 1: Fixed Covariance Matrix

For the first case, we fix $\mu_A = (0, 0)$, $\mu_B = (1, 1)$, $\mu_C = (0, 1)$ and $\Sigma_1 = \Sigma_B = \Sigma_C = \begin{pmatrix} 1 & 0 \\ 0 & .1 \end{pmatrix}$.

The space of normal distributions with fixed covariance matrix is euclidean, which implies that algorithm $\mathcal{A}_p$ is equivalent to the parallelogram rule under these conditions and that the defined relation is an analogical proportion. We observe on Fig. 4 that the trajectories of means in the space correspond to a parallelogram and that the two solutions are identical.

Case 2: Fixed Mean in Source Domain, Fixed Covariance from Source to Target

For the second case, we fix $\mu_A = \mu_B = (0, 0)$, $\mu_C = (0, 2)$ and, for covariance matrices, $\Sigma_A = \Sigma_C = \begin{pmatrix} 1 & 0 \\ 0 & .1 \end{pmatrix}$ and $\Sigma_B = \begin{pmatrix} .1 & 0 \\ 0 & 1 \end{pmatrix}$.

With these parameters, we observe that the two results are different (Fig. 5). The result of $\mathcal{A}_p(A, B, C)$ corresponds to the intuition that D will have the same mean as C and the same covariance change as B compared to A. However, for $\mathcal{A}_p(A, C, B)$, the results are non-intuitive: the mean of distribution D is different from the mean of C. It can be explained by the fact that the trajectory varies both in μ and Σ. The geometric properties of information require that these two dimensions are related together and that the change in μ depends on the change in Σ.

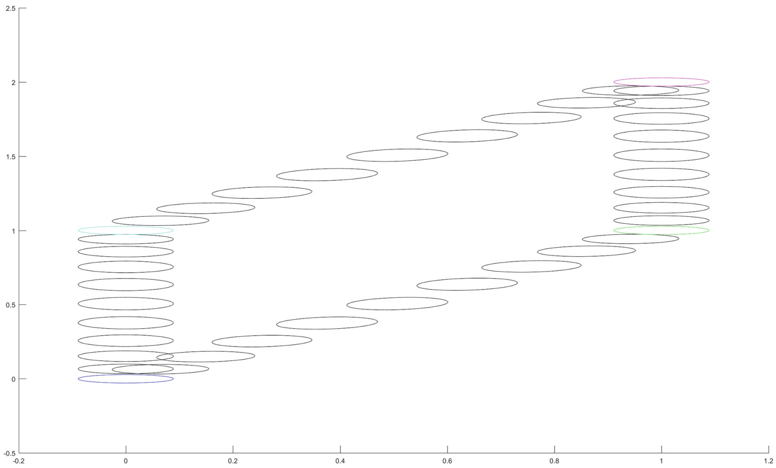


Fig. 4. Results for case 1 (fixed covariance matrix setting)

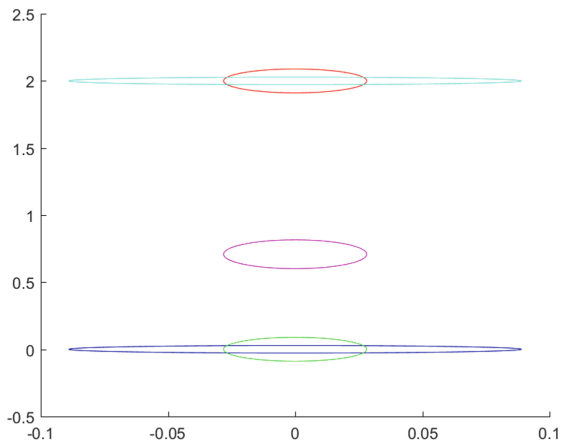


Fig. 5. Results for case 2 (fixed mean in source, fixed covariance from source to target)

Case 3: Symmetric Distributions

For the third case, we fix $\mu_A = (0, 0), \mu_B = (1, 0)$ and $\mu_C = (0, 1)$, and, for covariance matrices, $\Sigma_B = \Sigma_C = \begin{pmatrix} 1 & -.5 \\ -.5 & .5 \end{pmatrix}$ and $\Sigma_A = \begin{pmatrix} 1 & .5 \\ .5 & .5 \end{pmatrix}$. We notice on Fig. 6 that the trajectory leads to a distributions with "flat" covariance matrix (with one large and one very small eigenvalue). No real intuitive interpretation can be given of the observed trajectory (which shows that information geometry cannot explain shape deformations, here ellipse deformations, as expected by human beings).

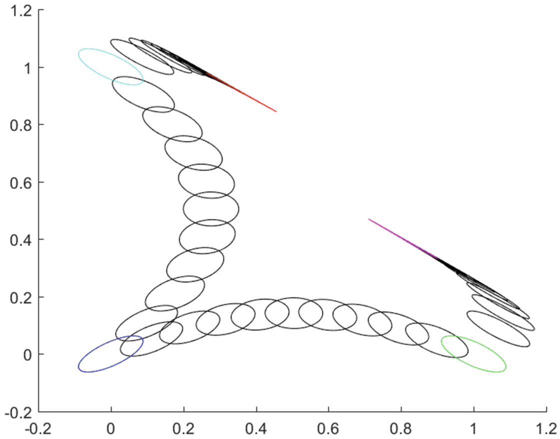


Fig. 6. Results for case 3 (symmetric)

Case 4: Slight Perturbation

For the third case, we fix $\mu_A = (0,0), \mu_B = (1,0)$ and $\mu_C = (0,1)$, and, for covariance matrices, $\Sigma_A = \begin{pmatrix} 1 & .5 \\ .5 & .5 \end{pmatrix}$, $\Sigma_B = \begin{pmatrix} 1 & -.5 \\ -.5 & .5 \end{pmatrix}$ and $\Sigma_C = \begin{pmatrix} 1 & .6 \\ .6 & .6 \end{pmatrix}$. Covariance matrix Σ_C is slightly different from Σ_1. If they were equal, the parallelogramoid would be closed. However, the slight modification introduces a perturbation large enough to make $\mathcal{A}_p(A, B, C) \neq \mathcal{A}_p(A, C, B)$ (Fig. 7. Such artifacts could introduce larger errors in case the distributions are not know with good precision (for instance if they were estimated from data). AQ4

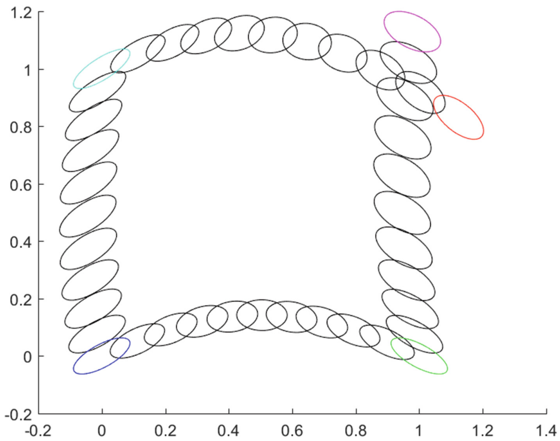


Fig. 7. Results for case 4 (slight perturbation)

4 Proportional Analogies on Manifolds

In previous sections, we have shown that the intuition of what an analogy can be in a differential manifold leads to a less constrained definition of analogies than the definition of proportional analogy. However, at this point of the paper, the existence of proportional analogies on a manifold $\mathcal{M}$ remains an open question. The purpose of this section is to discuss the construction of analogical proportions on a manifold.

Let $\mathcal{M}$ be a differential manifold. Our purpose is to design an algorithm to build analogical proportions. We define an algorithm as a function $\mathcal{A} : \mathcal{M}^3 \mapsto \mathcal{M}$.

Definition 3. *An algorithm $\mathcal{A} : \mathcal{M}^3 \mapsto \mathcal{M}$ is said to design an analogical proportion on $\mathcal{M}$ if, for all $(A, B, C) \in \mathcal{M}^3$, the relation $R(A, B, C, D) \vDash (D = \mathcal{A}(A, B, C))$ satisfies the axioms of analogical proportion.*

Definition 3 can be seen as a reverse way to define solutions of analogical equations. If a relation R is an analogical proportion over $\mathcal{M}$ designed by algorithm $\mathcal{A}$, then $x = \mathcal{A}(a, b, c)$ is the unique solution of equation $R(a, b, c, x)$ where x is the variable.

The following proposition offers an alternative characterization of proportion-designing algorithms based on global characteristics.

Proposition 1. *Algorithm $\mathcal{A}$ designs an analogical proportion if and only if the following three conditions hold true for any $(A, B, C) \in \mathcal{M}^3$:*

1. $\mathcal{A}(A, B, A) = B$ *or* $\mathcal{A}(A, A, B) = B$
2. $\mathcal{A}(A, B, C) = \mathcal{A}(A, C, B)$
3. $B = \mathcal{A}(C, \mathcal{A}(A, B, C), A)$.

Proof. The proof is a direct consequence of the axioms of analogical proportion.

In the case where $\mathcal{M}$ is a vector space, it can be verified easily that the parallelogram rule algorithm $\mathcal{A}(A, B, C) = C + B - A$ designs an analogical proportion. However, it is not the only algorithm to satisfy this property. In Proposition 2, we exhibit a parametered class of analogical proportion designing algorithms.

Proposition 2. *If $\mathcal{M}$ is a vector space and $f : \mathcal{M} \mapsto \mathcal{M}$ is a bijective mapping, then algorithm $\mathcal{A}_f$ defined by $\mathcal{A}_f(A, B, C) = f^{-1}(f(C) + f(B) - f(A))$ designs analogical proportion.*

It can be noticed that, when f is linear, algorithm $\mathcal{A}_f$ corresponds to the parallelogram rule. For other values of f, algorithm $\mathcal{A}_f$ can define proportions of another nature. An interesting perspective would be to study if these non-trivial proportions on a vector space can be related analogical proportions on a manifold.

The result of Proposition 2 can be generalized to any spaces:

Proposition 3. *Consider E and F two isomorphic sets with $f : E \mapsto F$ a corresponding isomorphism. If $\mathcal{A}_F : F^3 \mapsto F$ designs an analogical proportion on F then algorithm $\mathcal{A}_E : E^3 \mapsto E$ defined by $\mathcal{A}_E(a, b, c) = f^{-1}\left(\mathcal{A}_F(f(a), f(b), f(c))\right)$ designs an analogical proportion on E.*

Based on this result, it can be shown that analogical proportion defining algorithms exist on any manifold.

Theorem 3. *For any manifold $\mathcal{M}$, there exists an algorithm that defines analogical proportion on $\mathcal{M}$.*

Proof. Consider a finite atlas $\mathcal{A} = \{(U_\alpha, \psi_\alpha) | \alpha \in \{1, \ldots, m\}\}$. Such an atlas exists for m large enough. In this definition, U_α corresponds to a domain on $\mathcal{M}$ and $\psi_\alpha : U_\alpha \mapsto \mathcal{B}_n(0,1)$ is an homeomorphism from U_α onto the unitary ball $\mathcal{B}_n(0,1)$ on $\mathbb{R}^n$ (where n is the dimension of $\mathcal{M}$). If we denote $E_k = \{x \in \mathbb{R}^n | 2(k-1) < x_1 < 2(k+1)\}$, one can equivalently extend the mapping ψ_k to be homeomorphisms between U_k and E_k (Fig. 8).

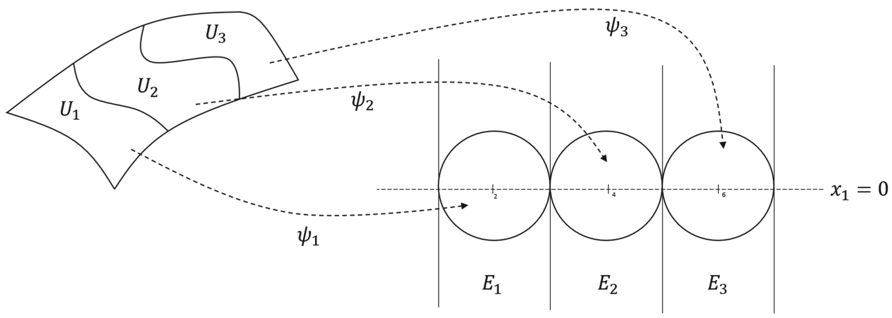


Fig. 8. Construction of a bijective mapping between n-dimensional manifold $\mathcal{M}$ and an open subset of $\mathbb{R}^n$. For simplicity purpose, the subsets U_k are presented as disjoint, which they are not.

We build a function $\psi : \mathcal{M} \mapsto \bigcap_{k=1}^{m} E_k$ as follows: If $x \in U_k \setminus \bigcap_{i>k} U_i$, then $\psi(x) = \psi_k(x) + e_k$ where e_k is the vector with first component equal to $2k$ and all other components equal to 0. This function defines a bijective mapping. Since $\bigcap_{k=1}^{m} E_k$ is an open subset of $\mathbb{R}^n$, there exists a bijection $\bigcap_{k=1}^{m} E_k \mapsto \mathbb{R}^n$. The theorem follows from Proposition 3 and the fact that $\mathbb{A}_{\mathbb{R}^n} \neq \emptyset$.

Theorem 3 is fundamental since it states the existence of analogical proportions on manifolds, which seems to invalidate the intuitions exposed with the parallelogrammoid method. However, the intuitive "validity" of the existing analogies (and in particular of the analogies produced by the proof) is not clear since they appear to be highly irregular since they are not continuous.

These observations point out a deficiency in the definition of analogical proportion, which comes from its main applicative domains. The definitions of analogical proportion were first designed for applications in character-string domains [12] and were discussed for applications in other non-continuous domains [13] such as analogies between finite sets. Among real continuous applications (hence applications which do not involve a discretization of the continuous space), most are based on parallelogram rule on a vector space. When defining analogical proportions on continuous spaces, a continuity property is also desirable, which is not induced by the definition of analogical proportion. Intuitively, this property makes sense: If two analogical problems are close, it is expected that their solutions will be close as well.

The question of the existence of analogical proportion defining algorithms that are also continuous (in the sense of a function $\mathcal{M}^3 \mapsto \mathcal{M}$) remains open at this step. It is impossible to adapt the proof of Theorem 3 in order to make the mapping continuous. More generally, the result cannot be directly adapted from Proposition 3.

5 Conclusion

In this paper, we proposed an extension of the well-known naive parallelogram representation of proportional analogies. We have shown that, when the space is curved (or more precisely when it is a differential manifold), the equality $D = C + B - A$ does not make any sense and more subtle descriptions have to be chosen. The solution we proposed is based on geodesic shooting and parallel transport, and corresponds to the parallelogram representation when the manifold is euclidean. However, the introduction of the curvature is inconsistent with one of the axioms of proportional analogy. However, this change of perspective is necessary since it is required by specific situations and the lost properties did not make sense from a cognitive point of view. We illustrated our proposition on two simple manifolds: the sphere and Fisher manifold for normal distributions. In the future, tests on more complex manifolds would be of interest, especially for analogies between objects which belong naturally to non-euclidean spaces. A study of feature relatedness in concept spaces and how such correlations induce a curve of the space is also directly connected to potential applications.

In addition, a work has to be done in the direction of finding relations of analogical proportions in manifolds. Until now, researches have focused mainly on more simple sets (boolean analogies, analogies between sets, character strings, or vectors) but some structures cannot be represented by simple objects and will require defining proportional analogies on manifolds, for instance shape spaces. We have shown the existence of analogical proportions in manifolds, but could not show the existence of *continuous* analogical proportions, which would be a fundamental property of a good intuitive proportion. The existence of continuous analogies remains an open question that will have to be solved in future researches.

Acknowledgments. This research is supported by the program Futur & Ruptures (Institut Mines Télécom).

References

1. Aamodt, A., Plaza, E.: Case-based reasoning: foundational issues, methodological variations, and system approaches. AI Commun. **7**(1), 39–59 (1994)
2. Amari, S.I.: Differential-Geometrical Methods in Statistics, vol. 28. Springer, New York (2012). https://doi.org/10.1007/978-1-4612-5056-2
3. Boothby, W.M.: An Introduction to Differentiable Manifolds and Riemannian Geometry, vol. 120. Academic Press, New York (1986)
4. Cencov, N.N.: Statistical Decision Rules and Optimal Inference, vol. 53. American Mathematical Soc., Providence (2000)
5. Cornuéjols, A., Ales-Bianchetti, J.: Analogy and induction: which (missing) link? In: Workshop "Advances in Analogy Research: Integration of Theory and Data from Cognitive, Computational and Neural Sciences". Sofia, Bulgaria (1998)
6. Falkenhainer, B., Forbus, K.D., Gentner, D.: The structure-mapping engine: algorithm and examples. Artif. Intell. **41**(1), 1–63 (1989)
7. Fisher, R.A.: Theory of statistical estimation. In: Mathematical Proceedings of the Cambridge Philosophical Society, vol. 22, pp. 700–725. Cambridge University Press, Cambridge (1925)
8. Goswami, U.: Analogical Reasoning in Children. Psychology Press, Hove (2013)
9. Han, M., Park, F.C.: DTI segmentation and fiber tracking using metrics on multivariate normal distributions. J. Math. Imaging Vis. **49**(2), 317–334 (2014)
10. Hofstadter, D., Mitchell, M.: The copycat project: a model of mental fluidity and analogy-making. In: Fluid Concepts and Creative Analogies, pp. 205–267. Basic Books Inc., New York (1995)
11. Hwang, S.J., Grauman, K., Sha, F.: Analogy-preserving semantic embedding for visual object categorization. In: Dasgupta, S., Mcallester, D. (eds.) Proceedings of the 30th International Conference on Machine Learning (ICML 2013), vol. 28, pp. 639–647. JMLR Workshop and Conference Proceedings, May 2013. http://jmlr.org/proceedings/papers/v28/juhwang13.pdf
12. Lepage, Y.: Solving analogies on words: an algorithm. In: Proceedings of the 17th international conference on Computational linguistics, vol. 1, pp. 728–734. Association for Computational Linguistics (1998)
13. Miclet, L., Bayoudh, S., Delhay, A.: Analogical dissimilarity: definition, algorithms and two experiments in machine learning. J. Artif. Intell. Res. **32**, 793–824 (2008). http://dblp.uni-trier.de/db/journals/jair/jair32.html#MicletBD08
14. Mikolov, T., Chen, K., Corrado, G., Dean, J.: Efficient estimation of word representations in vector space. CoRR abs/1301.3781 (2013). http://dblp.uni-trier.de/db/journals/corr/corr1301.html#abs-1301-3781
15. Mikolov, T., Yih, W.T., Zweig, G.: Linguistic regularities in continuous space word representations. In: HLT-NAACL, pp. 746–751 (2013)
16. Ollivier, Y.: A visual introduction to Riemannian curvatures and some discrete generalizations. In: Analysis and Geometry of Metric Measure Spaces: Lecture Notes of the 50th Séminaire de Mathématiques Supérieures (SMS), Montréal, pp. 197–219 (2011)
17. Rumelhart, D.E., Abrahamson, A.A.: A model for analogical reasoning. Cognit. Psychol. **5**(1), 1–28 (1973). http://www.sciencedirect.com/science/article/pii/0010028573900236
18. Skovgaard, L.T.: A Riemannian geometry of the multivariate normal model. Scand. J. Stat. **11**, 211–223 (1984)