Geodesics in the Riemannian Geometry for EEG Artifact detection

Author

Giovanni Grego - s202287

July 01, 2022

Contents

Abstract **1**

1 Introduction **2**
1.1 EEG and Artifact sources 2

2 Methods - Geodetics in the Riemannian geometry **4**
2.1 Introduction 4
2.2 Curves 5
2.2.1 Tangent vector to a curve 5
2.3 Surfaces 6
2.3.1 Tangent plane to a surface 7
2.4 Metric 10
2.4.1 Metric tensor 10
2.4.2 Line element 11
2.5 Geodesics 13
2.5.1 Covariant derivative 13
2.5.2 Geodesics 15

3 Application for artifact detection **16**
3.1 Distance and mean 17
3.2 Properties 17

4 Conclusions **19**

Resources **I**

Abstract

Electroencephalography (EEG) is nowadays one of the most popular methods to acquire cortical electrical brain activity from the scalp, in a non-invasive manner. One of the aspects that make the recording difficult, is the presence of important sources of noise, that overlap with the source brain electrical activity.

Many researchers are now focusing on the removal of these noisy sources in an automatic, online, and computational light way. For these reasons, in the last years one theoretical knowledge has emerged, showing its power in this context: the Riemannian geometry.

This report wants to give a non-formal introduction to the differential geometry, in an intuitive way to be followed easily by a non-specialist reader. The goal of this part is to introduce geodesics, that in the differential geometry are the curves of minimal length between two points, i.e. the distance.

Finally, those concepts will be applied to the practical case of the EEG artifact detection, showing the potentialities and giving some hints of why this technique is always more widely used.

Chapter 1

Introduction

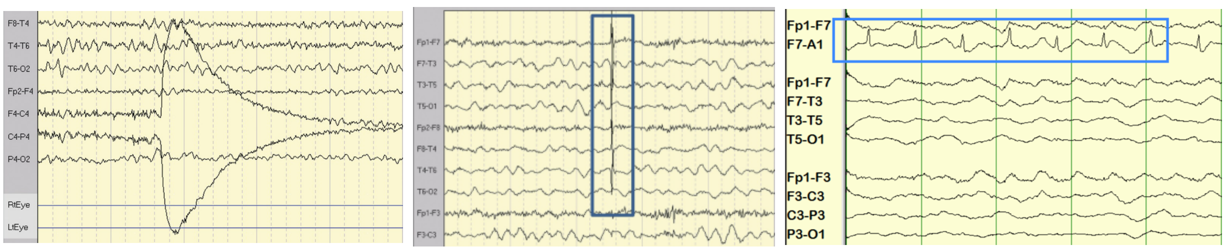
1.1 EEG and Artifact sources

Electroencephalography (EEG) is a non-invasive method for recording brain electrical activity by measuring the voltage between pairs of electrodes located uniformly on the scalp. Nowadays, EEG is one of the most important techniques for recording cortical brain activities, given its balance between accuracy and practicality, cost and easiness of use.

One of the drawbacks of this method is the sensitivity to external sources of information other than the brain activity, that contaminate the signal. These external sources are called artifacts, and they can interfere with the brain signal in the time-frequency-space domain, making the tasks of detection and the removal of the noise complicated.

The sources of artifacts can be classified in physiological activities of the user (e.g. movements, muscle tensions, ECG..) or external sources such as environmental interferences, electrode pop-ups or cable movement. They can be periodic events (e.g. ECG) or they can appear irregularly. In addition, they can influence a group of channels (global) or only a single channel (local) [25].

Some examples are illustrated in the figures below:



(a) Double-channel not periodic (b) Multi-channel not periodic (c) Single-channel periodic

Figure 1.1: Example of three different types of artifacts.

While the problem of artifacts is well known in the literature [25] [5] [17], a single universal and complete method for artifact detection/removal does still not exist.

Different methods have been elaborated over the last years. The most popular techniques are based on blind source separation (e.g ICA and CCA) [12], time-frequency representation (e.g. wavelet transform), adaptive filtering and machine learning applications [5] [17] [20] [19].

Many of these methods make use, at a certain point of the algorithm, the concept of distance among two points, for computing errors or for deciding the best choice to take. If in the past the most intuitive and straight-forward metric was the Euclidean distance, nowadays the literature tends to put more focus on the Riemannian geometry for computing distances and thus means.

Some examples of popular algorithms that use Riemannian geometry are the Riemannian Potato Field algorithm [4] [2] and the Riemannian Artifact Subspace Reconstruction [6].

It is believed that Riemannian geometry will improve the accuracy and the possibilities of the already known algorithms, as well as suggesting new solutions and methods [9]. For this reason, it is important to acquire the fundamentals of the Riemannian geometry, in order to be able to apply it in different scenario.

In this report, we will trace an overview of the geometry until the concepts of geodesic first, giving then explanations of why it suits well the EEG artifact detection/removal task.

Chapter 2

Methods - Geodetics in the Riemannian geometry

In order to follow this part, we make as assumption that the reader already has some basic knowledge about differential calculus (e.g. derivatives, integrals, differential operators...) [1], linear algebra (e.g. vectors, norm, inner and vector product, matrices...) [13] and the properties of R^3 (Banach space, Hilbert space, metric space and topological space, in particular) [26]. The purpose of this report is not to give an extended and complete treatment of the Riemannian geometry, but to give an easy to understand introduction to allow a general user to understand the concepts behind. In addition, we will discuss only the necessary parts for computing the differences among two points, i.e. the length of the geodesic. For these reasons, in this report we will focus only on the general ideas, limiting all the explanations to the R^3 domain, for making the report easier and to be able to give graphical examples. All the details and extensions can be found on the references [21][7][18][23].

2.1 Introduction

Riemannian geometry is a particular case of differential geometry, thus it is based on concepts like derivative, integral, and tangent vectors. The idea is that any surface in this geometry can be approximated, locally, with a tangent plane in which all the properties owned by the Euclidean space are valid. In this way, we can trace back every problem in a manifold to a Euclidean space. In particular, and for simplicity, we are interested in surfaces (subsets of R^3 with two dimensions) that are smooth, i.e. infinitely differentiable at any point [21].

2.2 Curves

A curve α in R^3 can be described with a parametrisation, thus: $\alpha : (t_1, t_2) \subseteq R \to R^3$, with (t_1, t_2) open interval. The components in R^3 of the curve are:

$$\alpha(t) = \begin{pmatrix} \xi_1(t) \\ \xi_2(t) \\ \xi_3(t) \end{pmatrix}, t \in R \tag{2.1}$$

2.2.1 Tangent vector to a curve

The tangent vector to a curve is defined as: $\dot{\alpha} = \lim_{\epsilon \to 0} \frac{\alpha(t+\epsilon)-\alpha(t)}{\epsilon}$ with $t \in R$ the parameter. The tangent vector approximates the curve at that point. A graphical example is shown below:

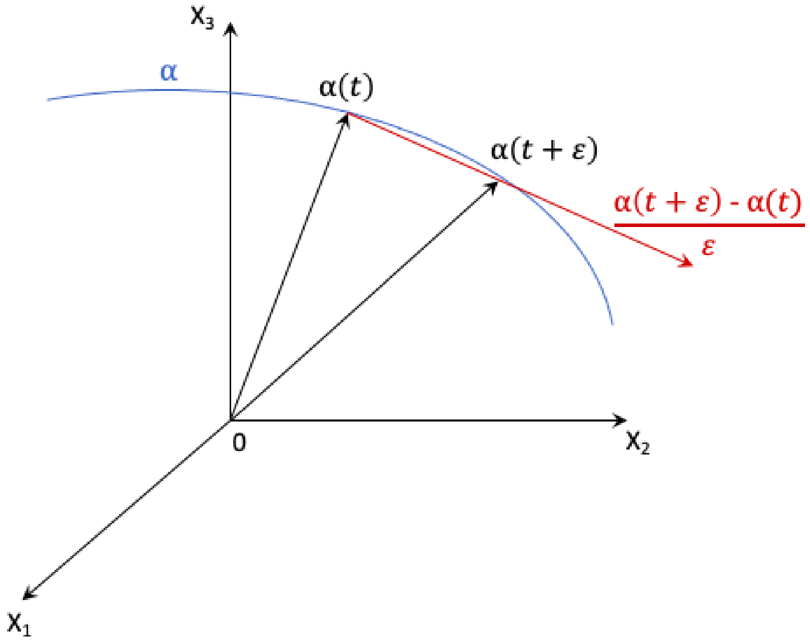


Figure 2.1: Graphical example of the construction of the tangent vector

The components are:

$$\dot{\alpha}(t) = \begin{pmatrix} \dot{\xi}_1(t) \\ \dot{\xi}_2(t) \\ \dot{\xi}_3(t) \end{pmatrix}, t \in R \tag{2.2}$$

2.3 Surfaces

A surface S is a bidimensional subset of R^3. It can be described by a vector function, parametrisation: $x : A \subseteq R^2 \to R^3$, with A open subset of R^2, with all regularities properties. A graphical example is shown below:

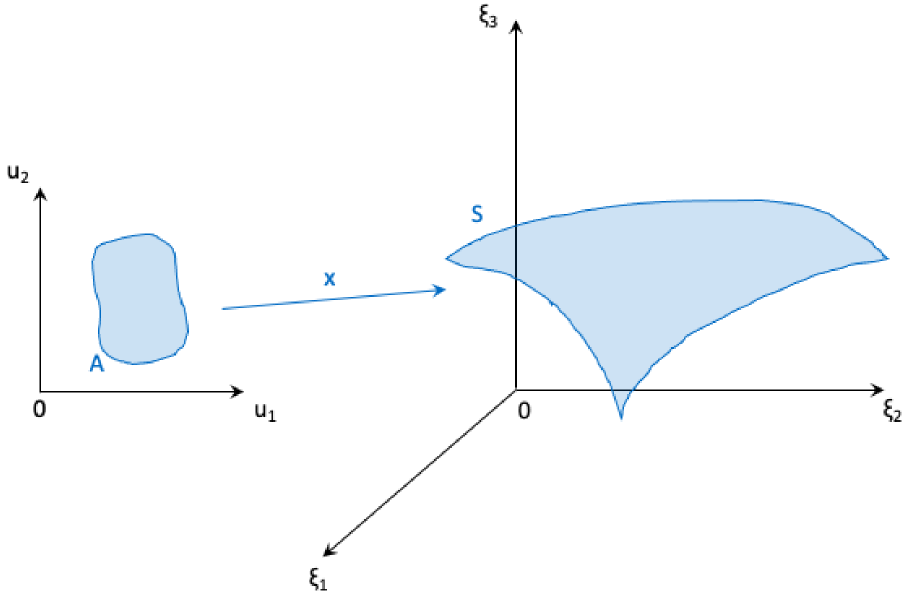


Figure 2.2: Graphical example of a surface parametrisation

The components of the parametrisation are:

$$x(u_1, u_2) = \begin{pmatrix} \xi_1(u_1, u_2) \\ \xi_2(u_1, u_2) \\ \xi_3(u_1, u_2) \end{pmatrix}, (u_1, u_2) \in A \tag{2.3}$$

Thus, a parametrisation determines in the surface a system of local coordinates: $(u_1, u_2) \to (\xi_1, \xi_2, \xi_3)$.

2.3.1 Tangent plane to a surface

The idea is the same as before, but now we need one step more for going from the line parametrized by t to the set A and then from A to R^3. So, we take a curve $\alpha(t)$ (N.B. the points in $\alpha(t)$ are (u_1, u_2)) and through the parametrisation x we obtain a new curve in R^3: $\beta(t) = x(\alpha(t))$, which lies on S. Below, the image to give a geometric idea:

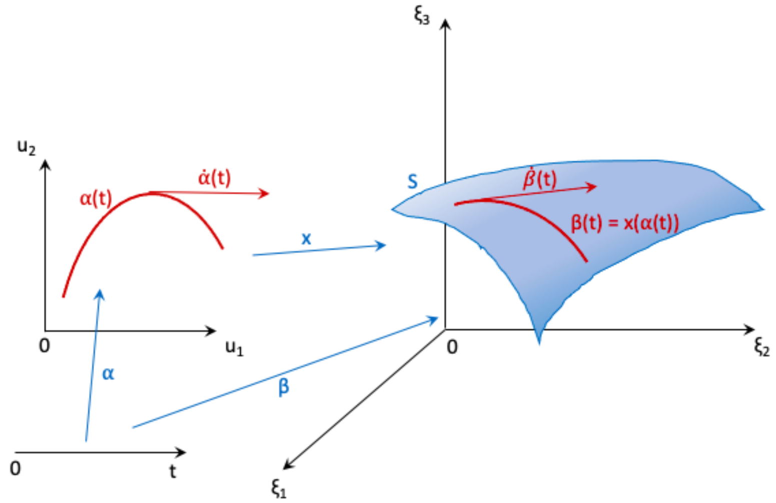


Figure 2.3: Graphical example of a curve parametrisation into a surface

The components are:

$$\beta(t) = \begin{pmatrix} \xi_1(u_1(t), u_2(t)) \\ \xi_2(u_1(t), u_2(t)) \\ \xi_3(u_1(t), u_2(t)) \end{pmatrix} \tag{2.4}$$

Of particular interest are the coordinate lines, thus the lines in S derived from lines parallel to the axes in A. These lines in A are $\alpha_1(t) = \begin{pmatrix} t \\ \bar{u}_2 \end{pmatrix}$ and $\alpha_2(t) = \begin{pmatrix} \bar{u}_1 \\ t \end{pmatrix}$ with $\bar{u}_1$ and $\bar{u}_2$ constants.
Thus, in S:

$$\beta_1(t) = \begin{pmatrix} \xi_1(t, \bar{u}_2) \\ \xi_2(t, \bar{u}_2) \\ \xi_3(t, \bar{u}_2) \end{pmatrix}, \beta_2(t) = \begin{pmatrix} \xi_1(\bar{u}_1, t) \\ \xi_2(\bar{u}_1, t) \\ \xi_3(\bar{u}_1, t) \end{pmatrix} \tag{2.5}$$

Now, the tangent vector is thus: (remember formula (2.4))

$$\dot{\beta}(t) = \begin{bmatrix} \frac{\partial \xi_1}{\partial u_1} & \frac{\partial \xi_1}{\partial u_2} \\ \frac{\partial \xi_2}{\partial u_1} & \frac{\partial \xi_2}{\partial u_2} \\ \frac{\partial \xi_3}{\partial u_1} & \frac{\partial \xi_3}{\partial u_2} \end{bmatrix} \begin{pmatrix} \dot{u_1} \\ \dot{u_2} \end{pmatrix} \tag{2.6}$$

Where $\begin{bmatrix} \frac{\partial \xi_1}{\partial u_1} & \frac{\partial \xi_1}{\partial u_2} \\ \frac{\partial \xi_2}{\partial u_1} & \frac{\partial \xi_2}{\partial u_2} \\ \frac{\partial \xi_3}{\partial u_1} & \frac{\partial \xi_3}{\partial u_2} \end{bmatrix}$ is the differential of x.

So we can rewrite the formula above as:

$$\dot{\beta}(t) = dx \begin{pmatrix} \dot{u_1} \\ \dot{u_2} \end{pmatrix} \tag{2.7}$$

This formula tells us how to transform the tangent vector $\dot{\alpha}$ to the curve α to the tangent vector $\dot{\beta}$ to the surface!

The formula (2.6) can also be written as:

$$\dot{\beta}(t) = \dot{u_1} x_{u_1} + \dot{u_2} x_{u_2} \tag{2.8}$$

where $x_{u_1} = \begin{bmatrix} \frac{\partial \xi_1}{\partial u_1} \\ \frac{\partial \xi_2}{\partial u_1} \\ \frac{\partial \xi_3}{\partial u_1} \end{bmatrix}, x_{u_2} = \begin{bmatrix} \frac{\partial \xi_1}{\partial u_2} \\ \frac{\partial \xi_2}{\partial u_2} \\ \frac{\partial \xi_3}{\partial u_2} \end{bmatrix}$.

Formula (2.8) is another fundamental piece. It tells us that all the derivatives of a point p in S are linear combinations of two vectors! All the vectors $v = a_1 x_{u_1} + a_2 x_{u_2}, a_1, a_2 \in R$ are the tangent plane of S in p! The tangent plane is therefore a bidimensional vector space in which $\{x_{u_1}, x_{u_2}\}$ is a base. We will call $T_p(S)$ this plane.

It is easy to show that x_{u_1} and x_{u_2} are the tangent vectors to the coordinate lines! These vectors have a fundamental role in the theory. They determine the tangent plane to a surface in a point, through which we can linearise the surface locally and explore it. It is also easy to show that the tangent plane is independent of the parametrisation, as expected intuitively.

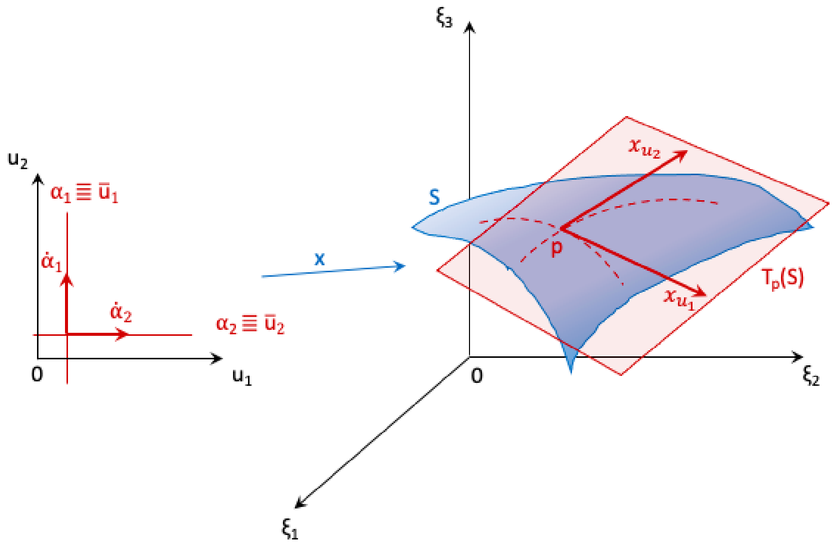


Figure 2.4: Graphical example of the tangent plane at a point on a surface.

Therefore, finally, we can parametrise the tangent plane. A generic point $q \in T_p(S)$ is defined as: $q = p + a_1 x_{u_1} + a_2 x_{u_2}, a_1, a_2 \in R, p \in A$. Thus:

$$T_p(S) = \begin{cases} \eta_1 = p_1 + a_1 \dfrac{\partial \xi_1}{\partial u_1}(p) + a_2 \dfrac{\partial \xi_1}{\partial u_2}(p) \\ \eta_2 = p_2 + a_1 \dfrac{\partial \xi_2}{\partial u_1}(p) + a_2 \dfrac{\partial \xi_3}{\partial u_2}(p) \\ \eta_3 = p_3 + a_1 \dfrac{\partial \xi_3}{\partial u_1}(p) + a_2 \dfrac{\partial \xi_3}{\partial u_2}(p) \end{cases} \tag{2.9}$$

2.4 Metric

One of the most important concepts to understand in the Riemannian geometry is that, in the differential geometry, all the metric properties of a surface depend only on the local coordinates. In other words, we can measure the properties and explore the surface without needing a container space. We will see that the metric depends only on the components x_{u_1}, x_{u_2}, thus only on the local coordinates u_1, u_2.

2.4.1 Metric tensor

There are many ways to introduce this object. I decided to first define it and then show some examples of why it is so powerful.

We define the metric tensor as:

$$g_{ij} = \begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{bmatrix} \tag{2.10}$$

Where:

$$g_{11} = \langle x_{u_1}, x_{u_2} \rangle = (\frac{\partial \xi_1}{\partial u_1})^2 + (\frac{\partial \xi_2}{\partial u_1})^2 + (\frac{\partial \xi_3}{\partial u_1})^2$$

$$\begin{aligned} g_{12} = g_{21} = \langle x_{u_1}, x_{u_2} \rangle = \langle x_{u_2}, x_{u_1} \rangle = \\ = \frac{\partial \xi_1}{\partial u_1}\frac{\partial \xi_1}{\partial u_2} + \frac{\partial \xi_2}{\partial u_1}\frac{\partial \xi_2}{\partial u_2} + \frac{\partial \xi_3}{\partial u_1}\frac{\partial \xi_3}{\partial u_2} \end{aligned}$$

$$g_{11} = \langle x_{u_2}, x_{u_2} \rangle = (\frac{\partial \xi_1}{\partial u_2})^2 + (\frac{\partial \xi_2}{\partial u_2})^2 + (\frac{\partial \xi_3}{\partial u_2})^2$$

We can see that, in the formulas above, everything depends only on u_1 and u_2, thus the local coordinates. As before, we need to question what happens if we change the arbitrary parametrisation of the surface, so if we change the local coordinates. It is easy to show that, with g_{ij} the metric tensor in the old parametrisation, and g'_{ij} the one in the new:

$$g'_{ij} = \sum_{k,l=1}^{2} \frac{\partial u_k}{\partial u'_i}\frac{\partial u_l}{\partial u'_j} g_{kl} \tag{2.11}$$

This is the base of the tensorial calculus, and it shows that it does not depend on anything else than, again, local coordinates.

One of the easiest examples to show the metric tensor in action is the inner product in a Riemannian surface. We take two vectors a, b in $T_p(S)$. Their inner product is:

$$\begin{aligned}\langle a, b\rangle &= \langle a_1 x_{u_1} + a_2 x_{u_2}, b_1 x_{u_1} + b_2 x_{u_2}\rangle = \\ &= a_1 b_1 \langle x_{u_1}, x_{u_1}\rangle + a_1 b_2 \langle x_{u_1}, x_{u_2}\rangle + a_2 b_1 \langle x_{u_2}, x_{u_1}\rangle + a_2 b_2 \langle x_{u_2}, x_{u_2}\rangle\end{aligned}$$

But this, from the (2.10) is equal to:

$$\begin{aligned}\langle a, b\rangle &= a_1 b_1 g_{11} + a_1 b_2 g_{12} + a_2 b_1 g_{21} + a_2 b_2 g_{22} = \\ &= \sum_{i,j=1}^{2} a_i b_j g_{ij}\end{aligned} \tag{2.12}$$

Another example is the formula for the norm of a vector in the tangent plane, also known as first fundamental form:

$$\begin{aligned}||a||^2 = \langle a, a\rangle &= (a_1)^2 g_{11} + 2 a_1 a_2 g_{12} + (a_2)^2 g_{22} = \\ &= \sum_{i,j=1}^{2} a_i a_j g_{ij}\end{aligned} \tag{2.13}$$

2.4.2 Line element

The line element is fundamental to be able to measure the length of lines on a surface. The length of an arch l of a curve β in S, where the parameter is $t \in (t_1, t_2)$ is:

$$\begin{aligned}l &= \int_{t_1}^{t_2} ds = \int_{t_1}^{t_2} \sqrt{(\frac{\mathrm{d}\xi_1}{\mathrm{d}t})^2 + (\frac{\mathrm{d}\xi_2}{\mathrm{d}t})^2 + (\frac{\mathrm{d}\xi_3}{\mathrm{d}t})^2 dt} \\ &= \int_{t_1}^{t_2} \sqrt{g_{11} \dot{u_1}^2 + 2 g_{12} \dot{u}_1 \dot{u}_2 + g_{22} \dot{u_2}^2 dt} = \int_{t_1}^{t_2} \sqrt{\sum_{i,j=1}^{2} g_{ij} \dot{u}_i \dot{u}_j dt}\end{aligned} \tag{2.14}$$

So, the line element ds is:

$$\begin{aligned}ds &= \sqrt{g_{11} \dot{u_1}^2 + 2 g_{12} \dot{u}_1 \dot{u}_2 + g_{22} \dot{u_2}^2 dt} = \\ &= \sqrt{g_{11} du_1^2 + 2 g_{12} du_1 du_2 + g_{22} du_2^2} = \sqrt{\sum_{i,j=1}^{2} g_{ij} du_i du_j}\end{aligned} \tag{2.15}$$

Thus, ds^2:

$$ds^2 = \sum_{i,j=1}^{2} g_{ij} du_i du_j \tag{2.16}$$

Formula (2.16) is of incredible importance in the theory of general relativity (extended in four dimensions). So, now, we can compute lengths of lines just knowing the metric sensor that, as we already know, depends only on the local coordinates. It is easy to show that the line element is independent of the parametrisation chosen.

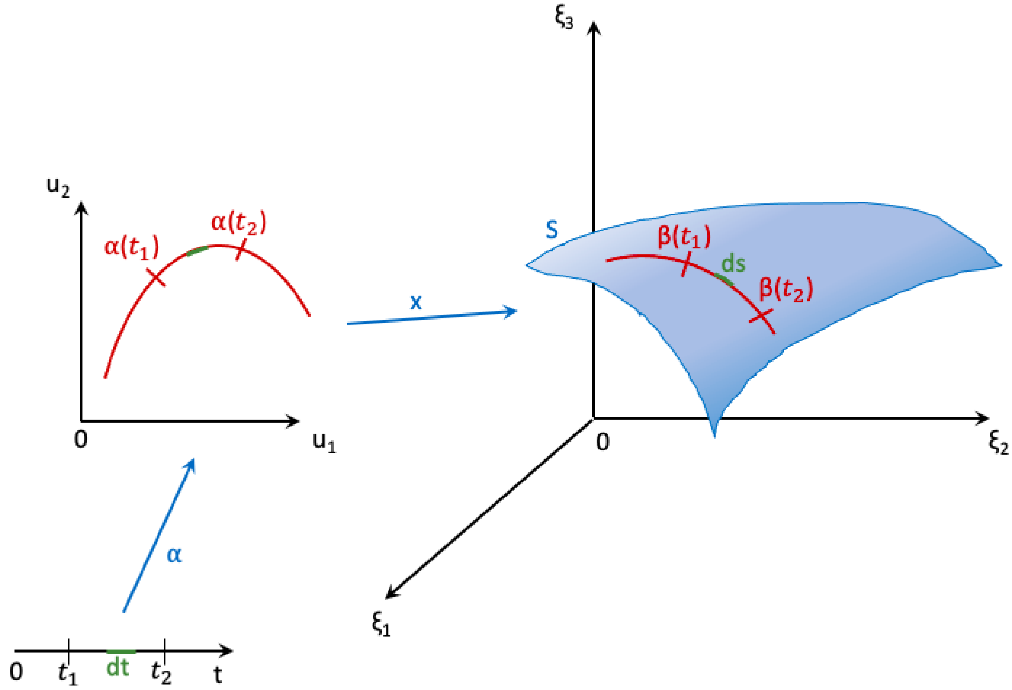


Figure 2.5: Graphical example of the line element on a surface

2.5 Geodesics

A geodesic can be defined in two equivalent ways: it is a line of minimum length between two points on a surface or it is a line with constant acceleration. The geodesic in a manifold is analogous to the distance in R^3. Or better, the geodesic between two points in the Euclidean manifold R^3 is the straight line that links the two points. The geodesics are, with the tangent planes, the basic objects of the manifold theory.

We know from the Euclidean space that a geodesic is a line with zero acceleration at any point, thus $\ddot{\beta}(t) = 0$. We want to define something similar for the manifolds. The problem is that, in general, the acceleration of a curve in a surface is a vector that has components not included in the tangent plane, thus not explainable with local coordinates. We need to define a new concept of derivative, such that the resulting vector is always contained in the tangent plane. We, therefore, define the covariant derivative.

2.5.1 Covariant derivative

The idea is to compute the derivative in R^3 of a vector field over a curve of a surface, and then project the resulting vector in the tangent plane. The vector found is so the covariant derivative of the vector field with respect to a curve.

Let's introduce on the manifold S, parametrised by $x(u_1, u_2)$, the vector field: $A = a_1(u_1, u_2)x_{u_1} + a_2(u_1, u_2)x_{u_2} \quad , a_1, a_2 : R^2 \rightarrow R$; and the curve $\beta = x(\alpha)$. The derivative of the vector field in a point is a vector that lies on R^3, in which $\{x_{u_1}, x_{u_2}, N\}$ are the basis vectors. Two components are the local coordinates, the third can be re-written as: $N = \frac{x_{u_1} \times x_{u_2}}{\sqrt{det(g_{ij})}}$. The idea is to compute the derivative of A in the $\{x_{u_1}, x_{u_2}, N\}$ basis and then to take only the first two components (equivalent to the geometric projection). For simplicity, I skip the procedure step by step for computing the derivative of A and I give the final result.

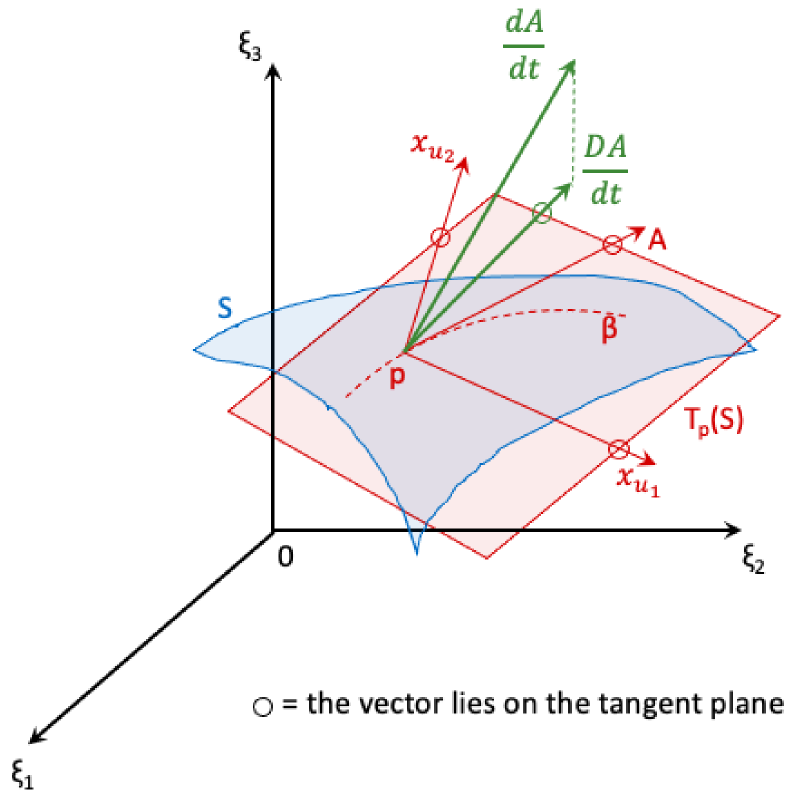


Figure 2.6: Graphical example of the covariant derivative

The covariant derivative $\frac{DA}{dt}$ is described as:

$$\frac{DA}{dt} = (\dot{a}_1 + \sum_{i,j=1}^{2} a_i \dot{u}_j \Gamma^1_{ij})x_{u_1} + (\dot{a}_2 + \sum_{i,j=1}^{2} a_i \dot{u}_j \Gamma^2_{ij})x_{u_2} \tag{2.17}$$

Where Γ^i_{jk} are the Christoffel symbols and they are derived from the combinations of vector and inner products of the permutations of the basis vectors. As before, I only give the final formula:

$$\Gamma^i_{jk} = \frac{1}{2}\sum_{l=1}^{2}[(\frac{\partial g_{kl}}{\partial u_j} + \frac{\partial g_{lj}}{\partial u_k} - \frac{\partial g_{jk}}{\partial u_l})g^{il}] \tag{2.18}$$

Note again that everything depends only on intrinsic properties: metric tensor and local coordinates. The definition of the covariant derivative is one of the most important concepts of the differential geometry, as the possibility of computing the derivatives in a surface given only the local coordinates, give the power of extending all the knowledge we have about the differential calculus without exit from the manifold itself.

2.5.2 Geodesics

Finally, we can now describe a geodesic from a vector approach. We define the geodesic on a surface (analogously of the line in R^3) as the curve with zero acceleration. So we impose that the second derivative of the curve $\beta = x(\alpha(t))$ on the surface S is equal to zero:

$$\frac{D\dot{\beta}}{dt} = 0 \tag{2.19}$$

But we know that $\dot{\beta} = \dot{u}_1 x_{u_1} + \dot{u}_2 x_{u_2}$, so we need to solve the system of differential equations:

$$\ddot{u}_i + \sum_{j,k=1}^{2} \Gamma^i_{jk} \dot{u}_j \dot{u}_k = 0 \quad i = 1, 2 \tag{2.20}$$

With initial conditions $u_1(t_0), u_2(t_0), \dot{u}_1(t_0), \dot{u}_2(t_0)$.
The geodesic is, so, the curve along which the field of the vector tangent to it is parallel. Therefore, the tangent vector in R^3 to a geodesic has a constant norm!

At this point, we would introduce the most interesting and astonishing part of the differential geometry: the curvature. This topic, though, is beyond the scope of this report and it is therefore left to the reader [7][21].

Chapter 3

Application for artifact detection

We saw the theory behind geodesics in a manifold, and we understood that they can be used for computing distances and means. But why they are emerging in the EEG signal analysis? Which are the advantages of this method, and how to apply it to EEG artifact detection?

The idea is that the variance of a signal (single time series, single-channel), when an artifact is produced, is higher than the variance of a signal produced by normal brain activities [9][4]. The variance in the single-channel case is a real number, and it can be computed in a sliding window matter, having for each data point the variance of the signal in the last n data points. When we use multiple channels, we have multiple signals and so multiple variances. In this case, we don't speak anymore about the variance of the signal but we construct the covariance matrix, which is the extension of the variance in multiple dimensions. The covariance matrix is a symmetric positive definite matrix. In the space of the positive matrices, all the assumptions used in the theory part are valid, and it can be shown that the space of the covariance matrices is a differentiable manifold [9]. So, we can also use the concept of geodesic as a metric of the distance between two points, i.e. two covariance matrices.

Now the idea is obvious: we compute a baseline covariance matrix and we compute in real-time the covariance matrix of the signal at that moment; we then compute the distance between these two matrices and, if the distance is higher than a threshold, an artifact is detected. The reason why the Riemannian approach gives good results is that, as we will see, this distance has some invariance properties that make it robust against noise and outliers and it has incredible generalization capabilities across sessions and subjects. It is important to highlight that, the generalization capability is more important for the classifiers than not for the easier task of covariance distance. The Riemannian distance, as well as many other useful methods based on the Riemannian geometry for positive-definite matrices are implemented in the python library pyRiemann [24].

3.1 Distance and mean

It can be shown that the length of the geodesic between two points (i.e. matrices) in the manifold of the positive matrices has a closed-form solution:

$$d(M_1, M_2) = || \log(M_1^{-\frac{1}{2}} M_2 M_1^{-\frac{1}{2}}) || = \sqrt{\sum_{n=1}^{N} \log^2 \lambda_n} \tag{3.1}$$

where λ_n are the N eigenvalues of $M_1^{-\frac{1}{2}} M_2 M_1^{-\frac{1}{2}}$.

A better look at this formula could suggest a more efficient way of computing the eigenvalues. It is well known that matrix operations (especially multiplication and power elevation) are heavy tasks for a computer. Thus, we want to keep the number of operations on matrices low. For positive definite matrices, it can be shown that the set of eigenvalues of the matrix $M_1^{-\frac{1}{2}} M_2 M_1^{-\frac{1}{2}}$ is the same as the matrix $M_1^{-1} M_2$. This form gives the same results above, but it implies only one multiplication and one power elevation. Thus, it will be preferable in a real implementation. On the other side, the pyRiemann implementation uses the first and less efficient form. Further investigations will be executed in the future to address if improvements can be done on this side.

Now that we have defined the concept of the distance between two matrices, it is easy to extend it to define the concept of Riemannian mean (i.e. Frechet's variational approach): if there exists a unique point X for which the dispersion $\frac{1}{k}\sum_{k=1}^{k} d^2(M_k, X)$ is minimal, then X is called the mean of the points $\{M_1, ..., M_k\}$. Thus the problem of finding the mean corresponds to: $\arg\min_G \frac{1}{k}\sum_{k=1}^{k} d_G^2(M_k, G)$. With d as defined in formula (3.1).

3.2 Properties

Let's see some useful and important properties of the Riemannian distance and the Riemannian mean.

The distance is invariant under any congruence:

$$d_G(XM_1X^T, XC_2X^T) = d_G(M_1, M_2) \tag{3.2}$$

and under inversion:

$$d_G(M_1^{-1}, M_2^{-1}) = d_G(M_1, M_2) \tag{3.3}$$

These two formulas lead to a generalization for the mean. These are the congruence invariance for the mean G:

$$G(XM_1X^T, ..., XC_kX^T) = XG(M_1, ..., M_k)X^T \tag{3.4}$$

for any invertible X; and the self-duality property:

$$G(M_1^{-1}, ..., M_k^{-1}) = (G(M_1, ..., M_k))^{-1} \tag{3.5}$$

The congruence invariance is an incredibly important property in the EEG signals analysis that, for nature, is very sensitive to shifts in the sources of data, especially true for different subjects and different recording sessions. The invariance tells us that the manipulations we do in the sensor space manifold are equivalent to those done in a manifold of the same dimensions, in our case the source space manifold. In other words, it means that, even though we are observing the noisy samples taken from the electrodes (i.e. the sensor space), the relative distances, means, and general transformations we can do in that space, would give the same results of those done in the "real" underneath source space, the original brain activity signal space.

In addition, the Riemannian mean, being an extension of the geometric mean, is more robust to outliers than, for example, the arithmetic mean [10]. The difference is greater for not Gaussian distributions, of particular importance in EEG signals. Many studies provide shreds of evidence of the effectiveness of using the Riemannian distance, In the following, just some well known references of studies using Riemannian geometry: [11][4][3][27] [28][16][15][14][22][8].

Chapter 4

Conclusions

In this elaborate, we started with an introduction to the artifact detection and classification tasks and we focused on a particular concept: the Riemannian distance.

The choice for this method was due to its contemporary power and simplicity, promising good results, and for being implemented in embedded systems in a real-time setup. We had a really fast and introductory idea to the Riemannian geometry, focused on giving the basis for understanding the geodesics in a Riemannian manifold.

Therefore we treated the important topics of curves, surfaces, metric tensor, and covariant derivative for, eventually, introducing the geodesics as a measure of distance between two points in the manifold.

Finally, we gave the intuition of how to apply this theoretical knowledge in the EEG artifact detection field, showing how to compute distances between two covariance matrices. Lastly, we showed some properties of the Riemannian distance and, at a glance, some hints of why it is so suitable for EEG signals.

Resources

[1] André Avez. *Differential calculus*. Courier Dover Publications, 2020.

[2] Alexandre Barachant, Anton Andreev, and Marco Congedo. "The Riemannian Potato: an automatic and adaptive artifact detection method for online experiments using Riemannian geometry". In: *TOBI Workshop lV*. 2013, pp. 19–20.

[3] Alexandre Barachant et al. "Extraction of motor patterns from joint EEG/EMG recording: A Riemannian Geometry approach". In: *6th International Brain-Computer Interface Meeting*. 2016.

[4] Quentin Barthélemy et al. "The Riemannian potato field: a tool for online signal quality index of EEG". In: *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 27.2 (2019), pp. 244–255.

[5] Shaibal Barua and Shahina Begum. "A review on machine learning algorithms in handling EEG artifacts". In: *The Swedish AI Society (SAIS) Workshop SAIS, 14, 22-23 May 2014, Stockholm, Sweden*. 2014.

[6] Sarah Blum et al. "A Riemannian Modification of Artifact Subspace Reconstruction for EEG Artifact Handling". In: *Frontiers in Human Neuroscience* 13 (2019).

[7] Isaac Chavel. *Riemannian geometry: a modern introduction*. Vol. 98. Cambridge university press, 2006.

[8] Marco Congedo and Alexandre Barachant. "A special form of SPD covariance matrix for interpretation and visualization of data manipulated with Riemannian geometry". In: *AIP Conference Proceedings*. Vol. 1641. 1. American Institute of Physics. 2015, pp. 495–503.

[9] Marco Congedo, Alexandre Barachant, and Rajendra Bhatia. "Riemannian geometry for EEG-based brain-computer interfaces; a primer and a review". In: *Brain-Computer Interfaces* 4.3 (2017), pp. 155–174.

[10] Marco Congedo, Alexandre Barachant, and Ehsan Kharati Koopaei. "Fixed point algorithms for estimating power means of positive definite matrices". In: *IEEE Transactions on Signal Processing* 65.9 (2017), pp. 2211–2220.

[11] Marco Congedo et al. "Approximate joint diagonalization and geometric mean of symmetric positive definite matrices". In: *PloS one* 10.4 (2015), e0121423.

[12] Laura Frølich, Tobias S Andersen, and Morten Mørup. "Classification of independent components of EEG into multiple artifact classes". In: *Psychophysiology* 52.1 (2015), pp. 32–45.

[13] Werner H Greub. *Linear algebra*. Vol. 23. Springer Science & Business Media, 2012.

[14] Mehrtash Harandi, Mathieu Salzmann, and Richard Hartley. "Dimensionality reduction on SPD manifolds: The emergence of geometry-aware methods". In: *IEEE transactions on pattern analysis and machine intelligence* 40.1 (2017), pp. 48–62.

[15] Mehrtash T Harandi, Mathieu Salzmann, and Richard Hartley. "From manifold to manifold: Geometry-aware dimensionality reduction for SPD matrices". In: *European conference on computer vision*. Springer. 2014, pp. 17–32.

[16] Inbal Horev, Florian Yger, and Masashi Sugiyama. "Geometry-aware principal component analysis for symmetric positive definite matrices". In: *Asian Conference on Machine Learning*. PMLR. 2016, pp. 1–16.

[17] Md Kafiul Islam, Amir Rastegarnia, and Zhi Yang. "Methods for artifact detection and removal from scalp EEG: A review". In: *Neurophysiologie Clinique/Clinical Neurophysiology* 46.4-5 (2016), pp. 287–305.

[18] Jürgen Jost and Jeurgen Jost. *Riemannian geometry and geometric analysis*. Vol. 42005. Springer, 2008.

[19] Demetres Kostas, Stephane Aroca-Ouellette, and Frank Rudzicz. "BENDR: using transformers and a contrastive self-supervised learning task to learn from massive amounts of EEG data". In: *arXiv preprint arXiv:2101.12037* (2021).

[20] Vernon Lawhern, W David Hairston, and Kay Robbins. "DETECT: A MATLAB toolbox for event detection and identification in time series, with applications to artifact detection in EEG signals". In: *PloS one* 8.4 (2013), e62944.

[21] John Lee. *Introduction to topological manifolds*. Vol. 202. Springer Science & Business Media, 2010.

[22] Xavier Navarro-Sune et al. "Riemannian geometry applied to detection of respiratory states from eeg signals: The basis for a brain–ventilator interface". In: *IEEE Transactions on Biomedical Engineering* 64.5 (2016), pp. 1138–1148.

[23] Peter Petersen. *Riemannian geometry*. Vol. 171. Springer, 2006.

[24] *pyRiemann: Biosignals classification with Riemannian Geometry*. `https://pyriemann.readthedocs.io/en/latest/index.html`. Accessed: 2022-06-30.

[25] Mona Sazgar and Michael G Young. "EEG artifacts". In: *Absolute epilepsy and EEG rotation review*. Springer, 2019, pp. 149–162.

[26] Wilson A Sutherland. *Introduction to metric and topological spaces*. Oxford University Press, 2009.

[27] Takashi Uehara, Toshihisa Tanaka, and Simone Fiori. "Robust averaging of covariance matrices by Riemannian geometry for motor-imagery brain–computer interfacing". In: *Advances in Cognitive Neurodynamics (V)*. Springer, 2016, pp. 347–353.

[28] Florian Yger and Masashi Sugiyama. "Supervised logeuclidean metric learning for symmetric positive definite matrices". In: *arXiv preprint arXiv:1502.03505* (2015).