Derivatives with respect to metrics and applications: Subgradient Marching Algorithm

Fethallah Benmansour, Guillaume Carlier, Gabriel Peyré, Filippo Santambrogio

Derivatives with Respect to Metrics and Applications: Subgradient Marching Algorithm

F. Benmansour, G. Carlier, G. Peyré, F. Santambrogio *

January 8, 2009

Abstract

This paper describes the Subgradient Marching algorithm to compute the derivative of the geodesic distance with respect to the metric. The geodesic distance being a concave function of the metric, this algorithm computes an element of the subgradient in $O(N^2 \log(N))$ operations on a discrete grid of N points. It performs a front propagation that computes the subgradient of a discrete geodesic distance. Equipped with this Subgradient Marching, a Riemannian metric can be designed through an optimization process. We show applications to landscape modeling and to traffic congestion. Both applications require the maximization of geodesic distances under convex constraints, and are solved by subgradient descent computed with our Subgradient Marching. We also show application to the inversion of travel time tomography, where the recovered metric is the local minimum of a non-convex variational problem involving geodesic distances.

Keywords: Geodesics, Eikonal equation, subgradient descent, Fast Marching Method, traffic congestion, travel time tomography.

1 Introduction

1.1 Riemannian Metric Design

The shortest path between a pair of points for a given Riemannian metric defines a curve that tends to follow areas where the metric is low. It is an object of primary interest in both pure mathematics and applied fields. For instance, as far as applications are concerned, such minimal paths are used intensively in computer vision and medical image analysis to perform segmentation of objects and extraction of tubular vessels [8]. The metric is designed to be low around the boundary of organs and vessels so that shortest paths follow these salient features.

In some applications, the Riemannian metric is the object of interest, and should be computed from a set of constraints or criteria. Some of these constraints might involve the length of geodesic curves between sets of key points, and these geodesic distances should be maximized or minimized. As shown in this paper, the maximization of geodesic lengths leads to convex problems, whereas the minimization of the distance leads to a non-convex problem. A global (for maximization) or local (for minimization) solution to these metric design problems can be found using a subgradient descent.

*CEREMADE, UMR CNRS 7534, Université Paris-Dauphine, Pl. de Lattre de Tassigny, 75775 Paris Cedex 16, FRANCE {benmansour,carlier,peyre,filippoo}@ceremade.dauphine.fr.

This paper proposes the first algorithm to compute such a subgradient of the geodesic distance with respect to the metric. It can thus be used as a building block for an optimization procedure that computes an optimal metric according to criteria on the length of the geodesic curves.

1.2 Geodesic Distances.

Riemannian metric. An isotropic Riemannian metric ξ on a domain $\Omega \subset \mathbb{R}^d$ defines a weight $\xi(x)$ that penalizes a curve $\gamma(t)$ passing through a point $x = \gamma(t) \in \Omega$. The length of the curve according to ξ is

$$L_\xi(\gamma) = \int_0^T |\gamma'(t)|\xi(\gamma(t))dt. \tag{1.1}$$

This metric ξ defines a geodesic distance $d_\xi(x_0, x)$ that is the minimal length of rectifiable curves joining two points $x_0, x \in \Omega$

$$d_\xi(x_0, x) = \min_{\gamma(0)=x_0, \gamma(1)=x} L_\xi(\gamma). \tag{1.2}$$

The distance map

$$\mathcal{U}^\xi(x) = d_\xi(x_0, x) \tag{1.3}$$

to the starting point x_0 is a function of the metric ξ, where we have drop the dependence with respect to x_0 for simplicity. The mapping $\xi \mapsto \mathcal{U}^\xi(x)$ is the one we wish to maximize or minimize in this paper, where x_0 and x are fixed points.

The geodesic curve γ joining x_1 to x_0 is the solution of an ordinary differential equation that corresponds to a gradient descent of $\mathcal{U}^\xi$

$$\frac{\mathrm{d}\gamma(s)}{\mathrm{d}s} = -\mathrm{grad}_{\gamma(s)}\,\mathcal{U}^\xi \quad \text{and} \quad \gamma(0) = x_1, \tag{1.4}$$

where $\mathrm{grad}_x\,\mathcal{U}^\xi \in \mathbb{R}^d$ is the usual gradient of the function $x \mapsto \mathcal{U}^\xi(x)$. This should not be confused with the subgradient with respect to the metric defined in the following paragraph.

Geodesic subgradient. The design of a metric through the maximization or minimization of $d_\xi(x_0, x)$ requires to compute the gradient $g = \nabla_\xi \mathcal{U}^\xi(x)$ of the mapping $\xi \mapsto \mathcal{U}^\xi(x)$. For any location $y \in \Omega$, $g(y)$ tells how much the geodesic distance between x_0 and x is sensitive to variations on $\xi(y)$.

In the continuous framework of (1.1) and (1.2), a small perturbation $\xi_\varepsilon = \xi + \varepsilon h$ defines a geodesic distance map $\mathcal{U}^{\xi_\varepsilon}(x)$ between x and x_0, that can be differentiated with respect to ε at $\varepsilon = 0$

$$\frac{d}{d\varepsilon}\mathcal{U}^{\xi_\varepsilon}(x)\Big|_{\varepsilon=0} = \int_\gamma h\, d\mathcal{H}^1 = \int_0^1 h(\gamma(t))|\gamma'(t)|dt, \tag{1.5}$$

where the curve γ is the geodesic from x_0 to x according to the metric ξ. If γ is unique, this shows that $\xi \mapsto \mathcal{U}^\xi(x)$ is differentiable at ξ, and that the gradient g is a measure supported along the curve γ. In the case where this geodesic is not unique, this quantity may fail to be differentiable. Yet, the map $\xi \mapsto \mathcal{U}^\xi(x)$ is anyway concave (as an infimum of linear quantities in ξ) and for each geodesic we get an element of the super-differential through Equation (1.5). In the sequel we will often refer to subgradients and subdifferentials for the concave function $\xi \mapsto \mathcal{U}^\xi(x)$ instead of superdifferentials and supergradients, this slight abuse of terminology should not create confusion however.

The extraction of geodesics is quite unstable, especially for metrics such that x and x_1 are connected by many curves of length close to the minimum distance $d_\xi(x_0, x)$. It is thus unclear how to discretize in a robust way the gradient of the geodesic distance directly from the continuous definition (1.5). This paper proposes an alternative method, where g is defined unambiguously as a subgradient of a discretized geodesic distance. Furthermore, this discrete subgradient is computed with a fast Subgradient Marching algorithm.

Figure 1 shows two examples of subgradient computations. Near a degenerate configuration, we can see that the subgradient g might be located around several minimal curves.

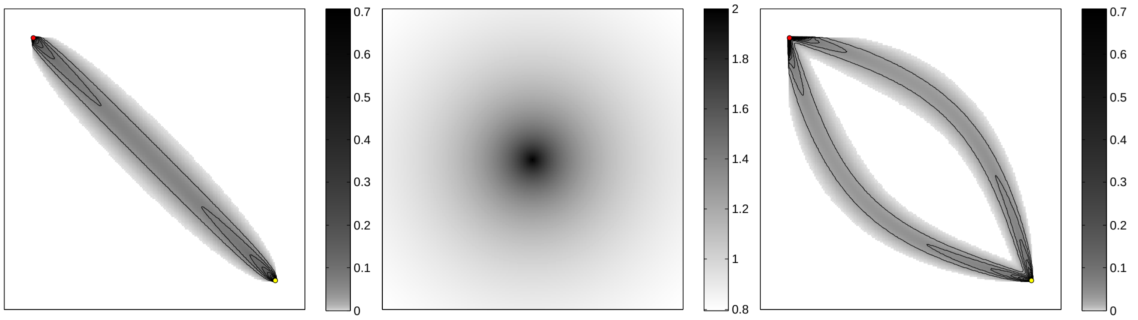

Figure 1: On the left, $\nabla_\xi \mathcal{U}(x_1)$ and some of its iso-levels for $\xi = 1$. In the middle, a non constant metric $\xi(x) = 1/(1.5 - \exp(-||c - x||))$, where c is the center of the image. On the right, an element of the superdifferential of the geodesic with respect to the metric shown in the middle figure.

Anisotropic metrics. The geodesic distance and its subgradient can be defined for more complicated Riemannian metric ξ that depends both on the location $\gamma(t)$ of the curve and on its local direction $\gamma'(t)/|\gamma'(t)|$. The algorithm presented in this paper extends to this more general setting, thus allowing to design arbitrary anisotropic Riemannian metric. We decided however to restrict our attention to the isotropic case, that has many practical applications.

1.3 Previous Works and Contributions

Geodesic distance computation. The estimation of distance maps $\mathcal{U}^\xi$ has been intensively studied in numerical analysis and can be approximated on discrete grid of N with the Fast Marching Method of Sethian [13], and Tsitsiklis [14] in $O(N \log(N))$ operations. This algorithm has opened the door to many application in computer vision where the minimal geodesic curves extracts image features, see for instance [13, 8]. Section 2 recalls the basics of the discretization of geodesic distance and Section 2.3 details the front propagation procedure underlying the Fast Marching method.

Geodesic distance optimization. The optimization of $\mathcal{U}^\xi$ with respect to ξ is much less studied. It is however an important problem in some specific fields, such as for landscape design, traffic congestion and seismic imaging. In these applications, the metric ξ is optimized to meet certain criteria, or is recovered by optimization from a few geodesic distance measures.

This paper tackles directly the problem of optimizing quantities involving the distance function $\mathcal{U}^\xi$ by computing a subgradient $\nabla_\xi \mathcal{U}^\xi(x)$ of the mapping $\xi \mapsto \mathcal{U}^\xi(x)$ for a given point x. The Subgradient Marching algorithm is described in Section 3. It follows the optimal ordering used by the Fast Marching, making the overall process only $O(N^2 \log(N))$ to compute a subgradient of the maps $\xi \mapsto \mathcal{U}^\xi(x)$ for all the grid points x.

This Subgradient Marching computes an exact subgradient of the discrete geodesic distance, so that it can be used to minimize variational problems involving geodesic distances. We believe it is important to first discretize the problem of interest and perform an exact minimization of the discrete problem. As far as geodesic quantities are involved, discretizing optimality condition of a continuous functional is indeed highly unstable.

Landscape design. Shape design requires the modification of the Riemannian metric defined by the first fundamental form of the surface. Minimization of geodesic length distortion is a well studied criterion to perform surface flattening and shape comparison, see for instance [3].

This paper tackles directly the problem of optimizing a Riemannian metric ξ. The example of landscape design using a fixed amount of resources is studied in Section 4.1. The length of geodesics is maximized under local and global constraint on the metric. This problem has a unique solution that can be found using a subgradient computed with our Subgradient Marching algorithm.

An application to travel time tomography is shown in Section 4.3. A subgradient descent allows one to find a local minimum of a variational energy involving geodesic distances.

Traffic congestion. A continuous generalization of the *Wardrop equilibria* [15], originally proposed in [5], involves the maximization of a concave functional depending on the geodesic distances between landmarks. A subgradient descent approximates this continuous solution and [1] describes an algorithm that makes use of our Subgradient Marching. Section 4.2 recalls basic facts of this congestion approximation and shows some numerical examples.

Seismic imaging. Seismic imaging computes an approximation of the underground from few surfaces measurements [6, 11]. This corresponds to an ill posed inverse problem that is regularized using smoothness prior information about the ground and simplifying assumption about wave propagation. Discarding multiple reflexions, the first arrival time of a pressure wave corresponds to the geodesic distance to the source, for a Riemannian metric ξ that reflects the properties of the underground.

The recovery of ξ from a few measurements $d_\xi(x_i, x_j)$ between sources x_i and sensors x_j corresponds to travel time tomography. A least square recovery of ξ involve the optimization of the geodesic distance. It has been carried over using for instance adjoint state methods [6, 11] that involve many computations of the geodesic map $\mathcal{U}^\xi$ for a varying metric ξ. Our Subgradient Marching algorithm allows to find a local minimum of the regularized least square energy using a descent method.

2 Discrete Geodesic Distances

2.1 Discretization

Eikonal Equation Our approach to minimize geodesic distances first defines a discrete geodesic distance $\mathcal{U}^\xi$, solution of a discretized partial differential equation. A discrete subgradient $\nabla_\xi \mathcal{U}^\xi(x)$ of the map $\xi \mapsto \mathcal{U}^\xi(x)$ is then defined to solve exactly discrete

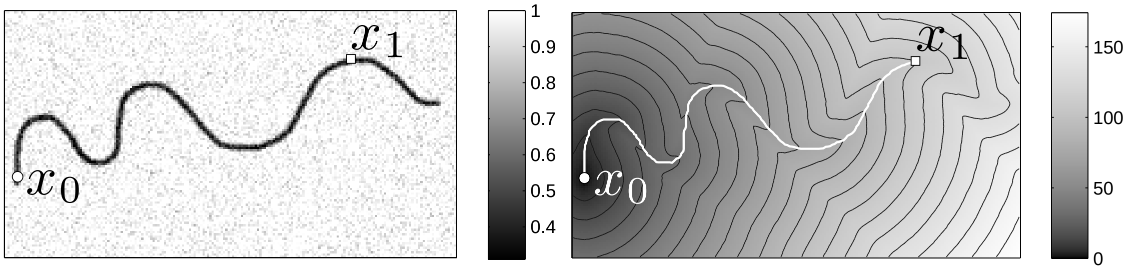

Figure 2: Example of the minimal path computation using the Fast Marching algorithm. On the left: the metric ξ. On the right: The minimal action map $\mathcal{U}$ and the minimal path linking x_1 to x_0.

variational problems involving geodesic distances. This is a general framework that could be extended to a larger class of non-linear partial differential equations.

The geodesic map $\mathcal{U}^\xi(x)$ defined in (1.3) is the unique viscosity solution of the Eikonal non-linear PDE (see [10])

$$\begin{cases} \|\nabla \mathcal{U}^\xi(x)\| = \xi, \\ \mathcal{U}^\xi(x_0) = 0. \end{cases} \tag{2.1}$$

The computation of $\mathcal{U}^\xi(x)$ thus requires the discretization of (2.1) so that a numerical scheme captures the viscosity solution of the equation.

Upwind Discretization In the following, we describe the computation in 2D and assume that the domain is $\Omega = [0,1]^2$, although the scheme carries over for an arbitrary domain in any dimension.

We will also drop the dependence on ξ and x_0 of the distance map $\mathcal{U}^\xi = \mathcal{U}$ to ease the notations. The geodesic distance map $\mathcal{U}^\xi$ is discretized on a grid of $N = n \times n$ points, so that $\mathcal{U}_{i,j}$ for $0 \leq i, j < n$ is an approximation of $\mathcal{U}^\xi(ih, jh)$ where the grid step is $h = 1/n$. The metric ξ is also discretized so that $\xi_{i,j} = \xi(ih, jh)$.

Classical finite difference schemes do not capture the viscosity solution of (2.1). Upwind derivative should be used instead

$$\begin{aligned} D_1\mathcal{U}_{i,j} &:= \max\{(\mathcal{U}_{i,j} - \mathcal{U}_{i-1,j}), (\mathcal{U}_{i,j} - \mathcal{U}_{i+1,j}), 0\}/h, \\ D_2\mathcal{U}_{i,j} &:= \max\{(\mathcal{U}_{i,j} - \mathcal{U}_{i,j-1}), (\mathcal{U}_{i,j} - \mathcal{U}_{i,j+1}), 0\}/h. \end{aligned}$$

As proposed by Rouy and Tourin [12], the discrete geodesic distance map $\mathcal{U} = (\mathcal{U}_{i,j})_{i,j}$ is found as the solution of the following discrete non-linear equation that discretizes (2.1)

$$D\mathcal{U} = \xi \quad \text{where} \quad D\mathcal{U}_{i,j} = \sqrt{D_1\mathcal{U}_{i,j}^2 + D_2\mathcal{U}_{i,j}^2}. \tag{2.2}$$

Rouy and Tourin [12] showed that this discrete geodesic distance $\mathcal{U}$ converges to $\mathcal{U}^\xi$ when h tends to 0.

Figure 2 shows an example of a discrete geodesic distance map $\mathcal{U}$. The metric ξ has a low value along a black curve, so that the geodesic curves tends to follow this feature. An example of geodesic curve is shown on the right, that is obtained by a numerical integration of the ordinary differential equation (1.4).

2.2 Concavity of the Geodesic Distance

To solve variational problems involving the geodesic distance $d_\xi(x_0, x)$, for $x = (ih, jh)$, one would like to differentiate with respect to ξ the discrete distance map $\mathcal{U}^\xi_{i,j}$, obtained by solving (2.2). Actually, this is not always possible, since the mapping $\xi \mapsto \mathcal{U}^\xi_{i,j}$ is not necessary smooth. The following proposition proves that $\mathcal{U}^\xi_{i,j}$ is a concave function of ξ and this allows for superdifferentiation (the correspondent of subdifferential for concave functions instead of convex).

Proposition 2.1. *For a given point (i, j), the functional $\xi \mapsto \mathcal{U}^\xi_{i,j}$ is concave.*

Proof. In the following we drop the dependence on (i, j) and note $\mathcal{U}^\xi = \mathcal{U}^\xi_{i,j}$. Thanks to the homogeneity, it is sufficient to prove super-additivity. We want to prove the inequality

$$\mathcal{U}^{\xi_1+\xi_2} \geq \mathcal{U}^{\xi_1} + \mathcal{U}^{\xi_2}.$$

Thanks to the comparison principle of Lemma 2.2 below, it is sufficient to prove that $\xi_1 + \xi_2 \geq D(\mathcal{U}^{\xi_1} + \mathcal{U}^{\xi_2})$, where the operator D is defined in (2.2). This is easily done if we notice that the operator D is convex (as it is a composition of the function $(s, t) \mapsto \sqrt{s^2 + t^2}$, which is convex and increasing in both s and t, and the operator D_1 and D_2, which are convex since they are produced as a maximum of linear operators) and $1-$homogeneous, and hence it is subadditive, i.e. it satisfies $D(u + v) \leq Du + Dv$. □

Lemma 2.2. *If $\xi \leq \eta$, then $\mathcal{U}^\xi \leq \mathcal{U}^\eta$.*

Proof. Let us suppose at first a strict inequality $\xi < \eta$. Take a minimum point for $\mathcal{U}^\eta - \mathcal{U}^\xi$ and suppose it is not the fixed point x_0. Computing D and using sub-additivity we have

$$\eta = D\mathcal{U}^\eta \leq D(\mathcal{U}^\eta - \mathcal{U}^\xi) + D\mathcal{U}^\xi = D(\mathcal{U}^\eta - \mathcal{U}^\xi) + \xi,$$

which gives $D(\mathcal{U}^\eta - \mathcal{U}^\xi) \geq \eta - \xi > 0$. Yet, at minimum points we should have $D(\mathcal{U}^\eta - \mathcal{U}^\xi) = 0$ and this proves that the minimum is realized at x_0, which implies $\mathcal{U}^\eta - \mathcal{U}^\xi \geq 0$.

To handle the case $\xi \leq \eta$ without a strict inequality, juste replace η by $(1 + \varepsilon)\eta$ and notice that the application $\eta \mapsto \mathcal{U}^\eta$ is continuous. □

2.3 Fast Marching Propagation

The Fast Marching algorithm, introduced by Sethian in [13] and Tsitsiklis in [14], allows to solve (2.2) in $O(N \log(N))$ operations using an optimal ordering of the grid points. This greatly reduces the numerical complexity with respect to iterative methods, because grid points are only visited once.

We recall the basic ideas underlying this algorithm, because our Subgradient Marching computation of $\nabla_\xi \mathcal{U}^\xi(x)$ makes use of the same ordering process.

The values of $\mathcal{U}$ may be regarded as the arrival times of wavefronts propagating from the source point x_0 with velocity $1/\xi$. The central idea behind the Fast Marching method is to visit grid points in an order consistent with the way wavefronts propagates.

In the course of the algorithm, the state of a grid point (i, j) passes successively from *Far* (no estimate of $\mathcal{U}_{i,j}$ is available) to *Trial* (an estimate of $\mathcal{U}_{i,j}$ is available, but it might not be the solution of (2.1)) to *Known* (the value of $\mathcal{U}_{i,j}$ is fixed and solves (2.1)). The set of *Trial* points forms an interface between *Known* points (initially the point x_0 alone) and the *Far* points. The Fast Marching algorithm progressively propagates this front of *Trial* points so that all grid points are visited, see Fig. 3.