Tikhonov Regularization of Sphere-Valued Signals

Laurent Condat

July 25, 2022

Abstract—It is common to have to process signals, whose values are points on the 3-D sphere. We consider a Tikhonov-type regularization model to smoothen or interpolate sphere-valued signals defined on arbitrary graphs. We propose a convex relaxation of this nonconvex problem as a semidefinite program, which is easy to solve numerically and is efficient in practice.

Index Terms—spherical data, Tikhonov regularization, smoothing, convex relaxation, directional statistics

I. INTRODUCTION

In a wide range of applications, one has to deal with signals or data records, like time series, whose values are points on the 3-D unit sphere, like the orientation of a camera, direction of a moving object, axis of rotation of a robot’s articulation. We consider the general setting, where the (possibly multidimensional) signal is defined on a graph, with values located at the nodes; every value is a point on the sphere. Two values are adjacent if there is an edge between their nodes. A 2-D image is a particular case with edges between every pair of neighboring pixels horizontally and vertically, forming a square grid. Then, to regularize signals on graphs, it is natural to promote the property that adjacent values are close to each other, in some sense. For scalar values, Tikhonov regularization consists in penalizing the squared differences of adjacent values. In this work, we focus on an equivalent model for sphere-valued signals. The sphere is nonconvex, so that the corresponding Tikhonov regularization problem is nonconvex too, and challenging to solve. We propose a new convex relaxation and show by some numerical examples that it is tight enough to yield exactly the global solution of the Tikhonov nonconvex problem in practical cases.

II. TIKHONOV SMOOTHING FOR SPHERE-VALUED SIGNALS

A. Sphere-Valued Signals on Graphs

We denote by $\mathbb{S} = \{ x = (x^1, x^2, x^3) \in \mathbb{R}^3 : \| x \|_2 = 1 \}$ the 3-D unit sphere, where $\| \cdot \|_2$ is the Euclidean norm. We want to estimate a signal $x = (x_n)_{n \in V}$, with values $x_n \in \mathbb{S}$, defined on a connected undirected graph $(V, E)$, where $V$ is the set of nodes and $E$ is the set of edges, which are sets of two distinct nodes. Typically, we are given a noisy signal $y = (y_n)_{n \in V}$ defined on the same graph and the sought signal $x$ is a smoothed, or denoised, version of $y$, which achieves a tradeoff between closeness to $y$ and smoothness, in some sense. Another typical setting is interpolation, or inpainting:

$y$ is defined on a subset $U \subset V$ of nodes and we want to estimate its missing samples; that is, $x$ is the smoothest signal defined on $V$ such that $x_n = y_n$, for every $n \in U$.

B. Classical Tikhonov Regularization

For vector-valued signals with values in $\mathbb{R}^d$, for any $d \geq 1$, Tikhonov-regularized smoothing consists in solving the following convex optimization problem. Given $y = (y_n)_{n \in V}$ and nonnegative weights $(w_n)_{n \in V}$ and $(\lambda_{n,n'})_{\{n,n'\} \in E}$, $x = (x_n)_{n \in V}$ is the solution to

$$\min_{x_n \in \mathbb{R}^d : n \in V} \sum_{n \in V} \frac{w_n}{2} \| x_n - y_n \|_2^2 + \sum_{\{n,n'\} \in E} \frac{\lambda_{n,n'}}{2} \| x_n - x_{n'} \|_2^2.$$  

(1)

For the interpolating task with $y$ defined only on $U \subset V$, we want to solve, instead:

$$\min_{x_n \in \mathbb{R}^d : n \in V} \sum_{\{n,n'\} \in E} \frac{\lambda_{n,n'}}{2} \| x_n - x_{n'} \|_2^2 \quad \text{s.t.} \quad x_n = y_n, \forall n \in U.$$  

(2)

Formally, (3) can be viewed as a particular case of (1) with $w_n = (+ \infty$ if $n \in U$, 0 otherwise), so that we can focus on the form (1), with the weights $w_n$ allowed to be $+\infty$.

We can notice that (1) corresponds to the maximum-a-posteriori (MAP) estimate of an unknown signal $x^0$ given $y$, which is $x^0$ plus white Gaussian noise, assuming a Gaussian Markov Random Field prior for $x^0$, with nonzero dependencies between its Gaussian variables along the edges of $V$. That is, $y^0 = x^0 + \mathcal{N}(1/w_n)$ and $x_{n,k}^0 - x_{n,k'}^0 \sim \mathcal{N}(1/\lambda_{n,n'})$, where $\mathcal{N}(\sigma^2)$ denotes the normal distribution with zero mean and variance $\sigma^2$, independently on each coordinate $k = 1, \ldots, d$.

C. Proposed Model

We want to formulate an equivalent problem to (1) for signals $x$ and $y$ with values in $\mathbb{S}$. The natural counterpart of (1) is to keep the same cost function to minimize, with values in $\mathbb{S}$ instead of $\mathbb{R}^d$:

$$\min_{x_n \in \mathbb{S} : n \in V} \sum_{n \in V} \frac{w_n}{2} \| x_n - y_n \|_2^2 + \sum_{\{n,n'\} \in E} \frac{\lambda_{n,n'}}{2} \| x_n - x_{n'} \|_2^2.$$  

(3)

We can develop the squared norm $\| x_1 - x_2 \|_2^2$, for any $(x_1, x_2) \in \mathbb{S}^2$, as $\| x_1 \|_2^2 + \| x_2 \|_2^2 - 2 x_1 \cdot x_2$, where $x_1 \cdot x_2 = x_1^1 x_2^1 + x_1^2 x_2^2 + x_1^3 x_2^3$ denotes the Euclidean inner product.
product of \( x_1 \) and \( x_2 \). Since \( \|x_1\| = \|x_2\| = 1 \), we have \( \|x_1 - x_2\|^2 = 2 - 2x_1 \cdot x_2 \). Hence, we can rewrite (3) as
\[
\text{minimize } \Psi_{\text{orig}}(x) = \sum_{n \in \mathbb{V}} w_n (1 - x_n \cdot y_n) + \sum_{\{n, n'\} \in E} \lambda_{n, n'} (1 - x_n \cdot x_n').
\] (4)

We can note that \( x_n \cdot x_n' \in [-1, 1] \) is the cosine of the angle at the origin between \( x \) and \( x' \), so that \( 1 - x_n \cdot x_n' \) is indeed a kind of distance between \( x_n \) and \( x_n' \). Like in the classical setting, there is a Bayesian interpretation of (4) as a MAP estimate of an unknown signal \( x^2 \) given \( y \), which is \( x^2 \) plus noise, this time following a von Mises–Fisher distribution, widely used in directional statistics \( 1 \). That is, \( y_n \in \mathbb{S} \) is the outcome of a random variable with probability density function (p.d.f.) proportional to \( e^{w_n y_n \cdot x_n^2} \). Similarly, \( x^2 \) is supposed to be a Markov Random Field with von Mises–Fisher dependency along every edge, with p.d.f. proportional to \( e^{\lambda_{n, n'} x_n^2 \cdot x_n'^2} \).

D. A Basic Convex Relaxation

Because \( \mathbb{S} \) is nonconvex, the problem (3) is nonconvex and difficult to solve. A straightforward way to relax it into a convex problem is to minimize the objective function, which is convex, with every variable \( x_n \) in the convex Euclidean ball \( \mathbb{B} := \{ x \in \mathbb{R}^3 : \| x \|_2 \leq 1 \} \) instead of the sphere \( \mathbb{S} \). After this problem has been solved, every \( x_n \in \mathbb{B} \) is resized as \( x_n/\|x_n\| \) to project it back on \( \mathbb{S} \). We call the combination of this convex problem and the rescaling postprocessing step the baseline method.

III. A New Convex Relaxation

To formulate a better convex relaxation than simply replacing \( \mathbb{S} \) by \( \mathbb{B} \) in (3), we start with the equivalent problem (4) and introduce an auxiliary variable \( d_{n, n'} \in \mathbb{R} \) at every edge \( \{n, n'\} \in E \): we can rewrite (4) as
\[
\text{minimize } \sum_{n \in \mathbb{V}} w_n (1 - x_n \cdot y_n) + \sum_{\{n, n'\} \in E} \lambda_{n, n'} (1 - d_{n, n'})
\text{ s.t. } d_{n, n'} = x_n \cdot x_n', \forall \{n, n'\} \in E.
\] (5)

In this last problem, the objective function to minimize is convex, and even linear, which is beneficial. Indeed, a general property in optimization is that minimizing a linear functional over a set yields a solution on the boundary of the set. It seems intractable to express the convex hull of the \( x_n \) and \( d_{n, n'} \) such that \( x_n \in \mathbb{S} \) and \( d_{n, n'} = x_n \cdot x_n' \), for every \( n, n' \). So, we “marginalize” the relaxation and design instead a convex set \( \Omega_{n, n'} \) at every edge \( \{n, n'\} \in E \), so that its boundary consists of points satisfying \( x_n \in \mathbb{S} \), \( x_n' \in \mathbb{S} \) and \( d_{n, n'} = x_n \cdot x_n' \).

In a previous work \( 2 \), the author proposed a convex relaxation of (5) for Tikhonov regularization on the 2-D unit circle \( \{ x = (x_1, x_2) \in \mathbb{R}^2 : \| x \|_2 = 1 \} \), where the \( \Omega_{n, n'} \) are complex ellipsoids, which are complex Hermitian matrices with particular structure. Indeed, it is convenient to represent points on the circle as complex numbers with modulus one, since multiplication of two such numbers corresponds to a rotation, or geodesic path between two points on the circle. Extending these ideas from the 2-D circle to the 3-D sphere is not straightforward, since there is no easy representation of 3-D rotations using complex numbers. However, it is known that quaternions are well suited to perform geometric operations on the sphere \( 3 \), and that a vector quaternion of the form \( q = ai + bj + ck \), for some \( (a, b, c) \in \mathbb{R}^3 \), can be represented as a \( 2 \times 2 \) complex matrix, which is a scaled Pauli matrix:
\[
M_{a, b, c} := \begin{pmatrix} -ci & -b - ai \\ b - ai & ci \end{pmatrix}
\] (6)

(where the complex \( i = \sqrt{-1} \) is different from the quaternionic \( i \) mentioned above).

For every \((a, b, c) \in \mathbb{R}^3\), \( M_{a, b, c}^H = -M_{a, b, c} \), where \( H \) denotes the Hermitian transpose, and the product \( M_{a, b, c}^H M_{a, b, c} \) is
\[
\begin{pmatrix} ci & b + ai \\ -b + ai & -ci \end{pmatrix} \begin{pmatrix} -ci & -b - ai \\ b - ai & ci \end{pmatrix} = \begin{pmatrix} a^2 + b^2 + c^2 & 0 \\ 0 & a^2 + b^2 + c^2 \end{pmatrix} = (a^2 + b^2 + c^2) \text{Id},
\]

where \( \text{Id} \) denotes the \( 2 \times 2 \) identity. Hence, \( M_{a, b, c}^H M_{a, b, c} = \text{Id} \) if and only if \( a^2 + b^2 + c^2 = 1 \). More generally, for any \( x = (a, b, c) \in \mathbb{S} \) and \( x' = (a', b', c') \in \mathbb{S} \), the product \( M_{a, b, c}^H M_{a', b', c'} \) encodes information about the geodesic path between \( x \) and \( x' \); for every \( x = (a, b, c) \in \mathbb{R}^3 \) and \( x' = (a', b', c') \in \mathbb{R}^3 \), the product \( M_{a, b, c}^H M_{a', b', c'} \) is
\[
\begin{pmatrix} ci & b + ai \\ -b + ai & -ci \end{pmatrix} \begin{pmatrix} c'i & -b - a'i \\ b - a'i & c'i \end{pmatrix} = \begin{pmatrix} a'c' + bb' + cc' + (ab' - a'b)i & a'c' - ac' + (bc' - b'c)i \\ ac' - a'c + (bc' - b'c)i & a' + bb' + cc' + (a'b - ab')i \end{pmatrix}.
\]

The trace of this matrix equals twice the inner product \( x \cdot x' = aa' + bb' + cc' \), which we are interested in for our formulation (5).

Therefore, for every \( x = (a, b, c) \in \mathbb{R}^3 \) and \( x' = (a', b', c') \in \mathbb{R}^3 \), the Hermitian matrix
\[
\begin{pmatrix} 1 & 0 & -ci & -b - ai \\ 0 & 1 & c'i & -b' - a'i \\ -ci & c'i & 1 & 0 \\ -b - ai & b + ai & 0 & 1 \end{pmatrix}
\] (9)

is of rank 2 and positive semidefinite, in which case it can be written as the product
\[
\begin{pmatrix} 1 & 0 \\ 0 & 1 \\ -b + ai & -ci \\ b + ai & c'i \end{pmatrix} \times \begin{pmatrix} ci & b + ai \\ -b + ai & -ci \\ -b' + a'i & -c'i \\ b' + a'i & c'i \end{pmatrix},
\] (10)

if and only if \( a^2 + b^2 + c^2 = 1 \), \( a^2 + b'^2 + c'^2 = 1 \), \( d = aa' + bb' + cc' \), \( e = b'c - bc' \), \( f = ac' - a'c \), \( g = a'b - ab' \). Since the
boundary of the convex set of positive semidefinite matrices of the form (9) is actually the subset of rank-2 matrices, such matrices are well suited to promote the properties that \( x \in S, x' \in S, d = x \cdot x' \).

Hence, the proposed convex relaxation is:

\[
\begin{align*}
\minimize_{x_n \in \mathbb{R}^3, n \in V} & \quad \Psi_{\text{conv}}(x, d, e, f, g) \\
\text{s.t.} & \quad P_{n,n'} \succeq 0, \quad \forall \{n, n'\} \in E,
\end{align*}
\]

where \( \succeq 0 \) denotes semidefiniteness and, for every \( \{n, n'\} \in E \), \( P_{n,n'} \) is a Hermitian matrix of the form (9), with \( a, a', b, b', c, c', d, e, f, g \) replaced by \( x^1_n, x^3_n, x^2_n, x^3_n, x^1_n, x^3_n, x^2_n, x^3_n, d_{n,n'}, e_{n,n'}, f_{n,n'}, g_{n,n'} \), respectively. There is no guarantee that when solving the problem (11), every matrix \( P_{n,n'} \) will be of rank 2, but this is what we hope for.

Thus, we solve the convex problem (11), and if the obtained solution \( (x^*, d^*, e^*, f^*, g^*) \) is such that \( x_n \in S \) for every \( n \in V \) and \( d_{n,n'} = x_n \cdot x_{n'} \), for every \( \{n, n'\} \in E \), then \( x^* \) is the exact global solution to the original nonconvex problem (4).

The problem (11) consists of minimizing a linear function under a composite constraint, like in [2]. A well suited algorithm is the Proximal Method of Multipliers [4], [5]; its complexity is dominated by projecting, at every iteration and every edge, a \( 6 \times 6 \) matrix onto the cone of positive semidefinite matrices.

IV. Experiments

In progress.

REFERENCES

[1] P. E. Jupp and K. V. Mardia, Directional statistics. John Wiley & Sons, 2009.
[2] L. Condat, “Tikhonov regularization of circle-valued signals,” IEEE Trans. Signal Process., 2022, in press.
[3] V. I. Arnol’d, “The geometry of spherical curves and the algebra of quaternions,” Russian Mathematical Surveys, vol. 50, no. 1, 1995.
[4] R. T. Rockafellar, “Augmented Lagrangians and applications of the proximal point algorithm in convex programming,” Math. Oper. Res., vol. 1, pp. 97–116, 1976.
[5] L. Condat, D. Kitahara, A. Contreras, and A. Hirabayashi, “Proximal splitting algorithms for convex optimization: A tour of recent advances, with new twists,” SIAM Review, 2022, to appear.
[6] U. R. Alim, T. Möller, and L. Condat, “Gradient estimation revitalized,” IEEE Trans. Vis. Comput. Graphics, vol. 16, no. 6, pp. 1494–1503, Nov.-Dec. 2010.