

MEAN AND PRINCIPAL CURVATURE ESTIMATION FROM NOISY POINT CLOUD DATA OF MANIFOLDS EMBEDDED IN $\mathbb{R}^n$

YOUSUF M. SOLIMAN

1. INTRODUCTION

Data is often provided as a finite set of points embedded in Euclidean space. In many applications, we have reason to believe that these data points are generated from a distribution with support over a manifold, rather than from a distribution with support over all of $\mathbb{R}^n$. In many applications, the shape of interest is a low dimensional submanifold embedded in a higher dimensional space; we refer to this assumption as the *manifold assumption*. A natural question to ask under what conditions can we reconstruct the manifold.

Unfortunately, fully reconstructing a submanifold is not directly within the realm of what is possible now. Although it is true that if we have a manifold that is sampled densely enough, that we can construct a triangulation of the data points that will triangulate the manifold, we have no reason to believe that the triangulation will preserve any of the geometric properties of the underlying manifold in the presence of any noise. Topological statistics and inference has been studied extensively over the past decade; unfortunately, however, geometric inference has not been studied nearly as much. I believe that the main reason that topological inference has gained so much momentum, while geometric inference has not, is due to the fact that when we only care about recovering the topological properties we can strongly reason about the noise and show that the topological properties of the manifold are robust to noise. On the other hand, small perturbations of the data have drastic effects on the underlying manifold. There has not been much work in a statistical sense that will guarantee convergence of the manifold as we collect more data.

Rather than asking to reconstruct the submanifold, one can instead ask if we can infer the geometric and topological properties of these surfaces from a finite sample of points that lie *near* the true manifold.

A classical example of this problem is the surface reconstruction problem in $\mathbb{R}^3$. Over the past several years the problem of surface reconstruction has been relatively well understood and effective and efficient algorithms have been developed [HDD⁺92, ABK98, KBH06, KSO04, CL96]. The main interest in the surface reconstruction problem come from the tangible representations of data in computer graphics and computer vision. Many algorithms can take a noisy data set that is contaminated with outliers in $\mathbb{R}^3$ and produce a tessellation that approximates the underlying surface.

The manifold reconstruction problem isn't as easy, nor has it been as well studied, in higher dimensions. It is important to note that although a lot of the topological properties of manifolds can be inferred from a high dimensional *noisy* point cloud, the same cannot be said about the geometry. There is relatively little work on rigorous inference of *geometric*, not topological, quantities of manifolds based on point cloud samples, even in the noiseless case.

In this paper, I will consider the problem of recovering the local curvature of the underlying manifold based on noisy point samples. Curvature is a widely used invariant feature in pattern classification and computer vision algorithms. Furthermore, understanding the curvature of the data gives one a better understanding of the local manifold geometry, which can then be used to better construct a sufficiently fine triangulation of the underlying manifold. Since data is only given as a sample of the underlying manifold, one is required to estimate curvature, or principal curvatures, from discrete values.

Clearly, classical differential geometry cannot be used directly on discrete objects such as triangle meshes or point cloud data. However, the study of discrete differential geometry provides an elegant extension of the classical theory over discrete manifolds, such as triangle meshes. Most of the use and application of discrete differential geometry has been contained in the field of computer graphics. Furthermore, I am not familiar with many ideas of discrete differential geometry that have been applied to high dimensional simplicial complexes. There has been a bit of work done on learning geometric operators, such as the Laplacian

E-mail address: ysoliman@cmu.edu.

Date: May 10, 2016.

[BSW09], from an arbitrary point cloud embedded in $\mathbb{R}^d$. Using the Laplacian, one can estimate the mean curvature. Furthermore, it is possible to extract the principal curvatures using concepts of discrete differential geometry, and the ideas of traditional differential geometry provide the intuition and justification for many of these different techniques. Furthermore, these techniques, in some cases, *can* handle noisy data, but I would argue that this is neither the main intent or strength of discrete differential geometry. This position comes from the observation that discrete differential geometry is not built around any sort of a statistical framework, and as such it would be difficult to strongly reason about the presence of noise.

The structure of this paper is as follows: first, I will precisely define curvature in the smooth setting. Next, I will provide a survey of the different techniques of inferring the curvature in the discrete case of arbitrary point cloud data. In this survey, I will describe the limitations of these algorithms as the dimensionality rises, as well as in the presence of noise.

I will then present an extension of integral invariants to submanifolds with arbitrary codimension as a methodology for inferring the curvature of point cloud data at different scales. I will then discuss the limitations of this framework and some ideas that can be used to solve these problems. Note that the extensions that I present here are not necessarily novel or computationally difficult, but I do think that I present a series of interesting ideas that I have not seen in the literature that I think may have some considerable importance in terms of high dimensional geometric inference. I will end by providing some of my own ideas of how understanding the curvature of very high dimensional submanifolds can be used in practical applications.

2. CURVATURE OF SURFACES AND DIFFERENTIABLE MANIFOLDS

I will begin by introducing curvature in the classical language of differential geometry. I consider the notion of principal curvatures, which is the extrinsic notion of curvature which intuitively looks at how quickly normal vectors change.

2.1. Principal Curvatures. I first informally discuss the notion of curvature of surfaces in $\mathbb{R}^3$. The principal curvatures of a surface in space can be described by two numbers at each point, called the principal curvatures. Informally, we can compute them in the following way: Let $\mathcal{M}$ be a surface in $\mathbb{R}^3$, p a point in $\mathcal{M}$, and η is a unit normal vector to $\mathcal{M}$ at p .

At each point $p \in \mathcal{M}$ we define the tangent plane as the flat space that is spanned by all of the tangent vectors of $\mathcal{M}$ at p . We denote this space as $T_p\mathcal{M}$. When we consider an m -dimensional manifold we have that the tangent plane $T_p\mathcal{M} \cong \mathbb{R}^m$ is the m -dimensional tangent plane. When we consider a submanifold of Euclidean space, we can define the normal space as follows:

$$(T_p\mathcal{M})^\perp = T_p\mathbb{R}^n \setminus T_p\mathcal{M}.$$

- (1) Choose a tangent vector $v \in T_p\mathcal{M}$ and consider the plane Π through p that contains both η and v . The intersection of Π with $\mathcal{M}$ is then a plane curve $\gamma \subset \Pi$ passing through p .
- (2) Compute the signed curvature $\kappa_p(v)$ of γ at p .
- (3) Repeat this for all tangent vectors in $T_p\mathcal{M}$.
- (4) Define the principal curvatures of $\mathcal{M}$ at p , denoted as κ_1 and κ_2 , as the minimum and maximum signed curvatures of planar curves γ in every tangent direction, respectively.

Note that we say that the curvature of a curve is positive if the normal vector is chosen such that it lies within the osculating circle.

One should note that the entire curvature of the manifold is *entirely* determined by the principal curvatures and directions. Furthermore, the principal curvature is not an intrinsic notion of the manifold, however, the Gaussian curvature ($\kappa = \kappa_1\kappa_2$) is. We can similarly define the mean curvature as $\mathcal{H} = \frac{1}{2}(\kappa_1 + \kappa_2)$.

2.2. Curvature of m -dimensional Submanifold Embedded in $\mathbb{R}^n$. Now we begin with a *rigorous* treatment the curvature of an m -dimensional submanifold $\mathcal{M}$ of $\mathbb{R}^n$. We will assume that $\mathcal{M}$ is at least $\mathcal{C}^2$ to ensure that the curvature is well defined everywhere.

We will begin by considering the first fundamental form. Let (U, φ) , where $\varphi : U \rightarrow \mathcal{M}$ be a local chart of $\mathcal{M}$. We will consider everything locally, so $\varphi : U \rightarrow \varphi(U) \subset \mathcal{M}$ is indeed a diffeomorphism. Let $u \in U$

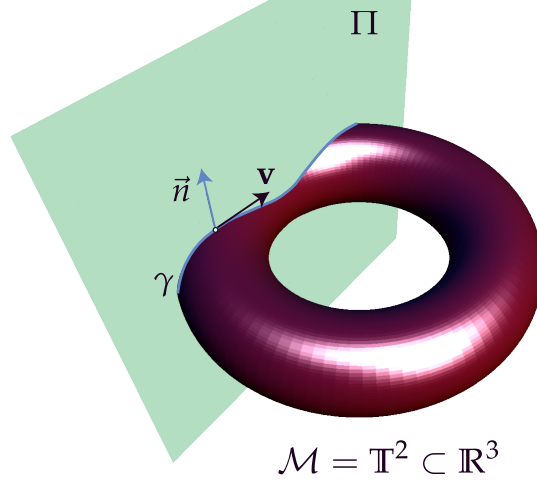


FIGURE 1. Principal Curvatures of Surfaces embedded in $\mathbb{R}^3$. We have that $\eta \in (T_p\mathcal{M})^\perp$, $v \in T_p\mathcal{M}$, we then define Π to be the plane defined by the normal vector $\eta \times v$. Now let γ be the planar curve found by intersecting Π with $\mathcal{M}$. $\kappa_p(v)$ is the curvature of the planar curve γ .

and let $p := \varphi(u) \in \mathcal{M}$. We have that the natural Riemannian metric (the standard inner product) on $\mathbb{R}^n$ induces a Riemannian metric on $T_p\mathcal{M}$ as follows:

$$I_p(\mathbf{v}) = \|\mathbf{v}\|^2, \quad \mathbf{v} \in T_p\mathcal{M}.$$

Now we will consider the second fundamental form. Again, consider $u \in U$ and let $p = \varphi(u) \in \mathcal{M}$. We now define the second fundamental tensor of the manifold $\mathcal{M}$ at p along the normal η as the mapping $H_p^{(\eta)} : T_p\mathcal{M} \times T_p\mathcal{M} \rightarrow (T_p\mathcal{M})^\perp$.

Without going into too much of the machinery of Riemannian geometry, we can define the affine connection, ∇ , as a way of connecting nearby tangent planes in a way that allows differentiability. In the case of $\mathbb{R}^n$, we have that the covariant derivative is just component-wise differentiation. Now if we consider a submanifold $\mathcal{N}$ that is embedded in $\mathcal{M}$, we say that $\bar{\nabla}$ is the connection on $\mathcal{M}$ and ∇ is the connection on $\mathcal{N}$ using the induced metric from $\mathcal{M}$. Now, we define the second fundamental tensor as the symmetric bilinear quadratic form

$$H_p^{(\eta)}(\mathbf{v}, \mathbf{w}) = \langle \bar{\nabla}_{\mathbf{v}} \bar{\mathbf{w}} - \nabla_{\mathbf{v}} \mathbf{w}, \eta \rangle = \langle D_{\mathbf{v}} \eta, \mathbf{w} \rangle, \quad \mathbf{v}, \mathbf{w} \in T_p\mathcal{M}.$$

Note that $\bar{\nabla}_{\mathbf{v}} \bar{\mathbf{w}} - \nabla_{\mathbf{v}} \mathbf{w}$ is just the normal component of the component-wise differentiation of the vector field in $\mathcal{M}$ with respect to $\mathcal{N}$. We now can write the second fundamental form as

$$\Pi_p^{(\eta)}(\mathbf{v}) = H_p^{(\eta)}(\mathbf{v}, \mathbf{v}), \quad \mathbf{v} \in T_p\mathcal{M}.$$

Note that in local coordinates, we have that

$$\left(\Pi_p^{(\eta)} \right)_{ij} = \frac{\partial^2 \varphi}{\partial x_i \partial x_j}(0) \cdot \eta$$

From this it is clear that the second fundamental form does indeed coincide with the Hessian, this justifies the Taylor series approximation of $\mathcal{M}$ *locally*, since all of the second partial derivatives coincide.

We can now define the curvature vector in the direction $\mathbf{v}$ of $\mathcal{M}$ at p as

$$\kappa(\mathbf{v}) = \frac{\Pi_p^{(\eta)}(\mathbf{v})}{I_p(\mathbf{v})}.$$

We call the eigenvectors of Π the principal directions of curvature and the eigenvalues the principal curvatures with respect to some normal η_i .

Note that if the manifold is of codimension 1, we can write locally that the manifold is a surface with the following Taylor Series expansion around some point x_0 as

$$f(x_1, \dots, x_m) = \frac{1}{2} \left(\sum_{i=1}^m \kappa_i x_i^2 \right) + \mathcal{O}(\|x\|^3).$$

Note that we can write the second fundamental form in this way since it is a quadratic form, so we can diagonalize it using the eigenvectors and since by definition κ_i are the eigenvalues, this becomes our diagonalization. We will discuss this framework of principal curvature on manifolds with arbitrary codimension in more depth later and show how we can use these Taylor expansions to work with manifold locally in a nice way.

Note that we could also consider a second fundamental form with respect to the entire normal space, which we could use to write the following Taylor Series expansion:

$$(\varphi(u+h) - \varphi(u))^\perp = \frac{1}{2} \Pi_p(h) + \mathcal{O}(|h|^3).$$

3. A NOTE ON CURVATURE IN THE DISCRETE SETTING

When attempting to extract any geometric properties in the discrete setting, we have to understand that we are aiming to extract the geometric properties of the underlying *smooth* manifold structure. It is important to think of the discrete structure, whether it be a discrete manifold or an arbitrary point cloud, as a proxy for the true manifold.

Many algorithms for estimating curvature are restricted to the case of a triangle mesh; however, most of these techniques can be generalized to the case of an arbitrary point cloud by estimating normal vectors, small geodesic neighborhoods, or a boundary over which to integrate. Similarly, we could generalize most of the techniques that work on triangle meshes to arbitrary discrete manifolds represented as simplicial complexes.

4. CURVATURE ESTIMATION

This section is concerned with discussing the various techniques that have been previously studied for estimation of principal, or mean, curvatures.

When thinking about the following curvature estimation algorithms, it is important to note that one of the main challenges that is unique to the point cloud data case, that is not an issue for a discrete manifold, is that the underlying neighborhood structure is *a priori* unknown. This can introduce some complications when, for example, estimating the curvature in all of the directions that are neighboring some point. However, this shortcoming is not always a huge issue since we can estimate small neighborhoods using ε -balls or k -Nearest Neighbors with a small k .

4.1. Local Chart Estimation. Local chart estimation concerns itself with fitting an analytic surface to points in a small neighborhood, then computing the curvatures of this surface analytically. These methods will produce exact results *only if* the points lie exactly on the surface being fit, and furthermore, if the model used to fit the surface indeed coincides with the underlying manifold.

One of the primary weaknesses of local chart estimation is that they become extremely unstable near discontinuities of the smooth structure. Furthermore, these algorithms are not guaranteed to converge once noise is introduced into the dataset.

When attempting to do local chart estimation, there are two primary methods of approaching this problem. The first technique, which provides a much more accurate approximation is interpolation, but interpolation (i.e. Lagrange interpolation) is extremely sensitive to noise, especially at when many data points that are close together are perturbed by even small amounts. The second technique is to do some sort of nonparametric regression, which is to find a surface that minimizes the least squares error.

Without knowing what the underlying neighborhood structure is, we are unable to consider the curvatures in the directions of the *neighboring* points around some point $p \in \mathcal{M}$. There has been extensive research done on local chart estimation in low dimensions as seen in [CP05, MW00, DB02]. It is important to note that, in practice, we are only able to do polynomial approximation due to the computational intractability of nonparametric regression as well as its need for a much larger sample size.

There have been also numerous reports across the years that shown that curvature estimation using local chart estimation is *extremely* sensitive to noise [FJ89].

4.2. Vertex Normal Curvature-Based Methods. (This is the technique that was used in the class earlier on.)

These techniques first estimate the normal curvature in the direction of each edge leaving a vertex, then use the κ_n estimates to find the second fundamental form. Very commonly, they estimate the normal curvature at a point p_i in the direction of a neighboring point p_j as follows:

$$\kappa_{ij} = \frac{2n_i \cdot (p_i - p_j)}{\|p_i - p_j\|^2}.$$

The principal curvatures may then be found from a function of the eigenvalues of

$$\sum_j \kappa_{ij} (p_i - p_j)(p_i - p_j)^\top.$$

One of the main issues with this algorithm is that the principal curvatures are only accurate when the distribution of the neighboring points is uniform. That is to say that methods start to break down in the case where the triangulation consists of many thin long triangles. Some of this issues can be mitigated by using a weighting scheme based on face areas.

We can show that these methods are special cases of local chart estimation, when we use circles to fit the data points instead of polynomials or other complex analytic surfaces. Since this is the case, these techniques also suffer the same issues that the local chart estimation techniques have.

This sort of a method is nice when we have a simplicial surface in $\mathbb{R}^3$, since we can compute vertex normals as a weighted average of the neighboring face normals.

When we compute vertex normals in higher dimensions, a difficulty that I have been running into is making sure that the choice of the normal at each point varies consistently, that is to say that the vertex normals do not switch signs since there are two natural choices, or signed values, of a normal in each codimension. This becomes much more difficult when the submanifold has codimension greater than 1. I want to work on a pipeline that takes the initial estimates of the normal vectors and produces a smoothly varying normal field over all of the data points. Although choosing a consistent normal orientation will not be possible on a nonorientable manifold, we will still be able to consider the curvature since it is only a local notion. As an aside, this might be a lot easier than I am making it out to be, if we choose a normal orientation that is pointing inwards in the local chart.

4.3. Tensor Voting Methods. Tensor voting, or averaging, methods compute the average of the curvature tensor over a small area of the polygonal mesh. These methods aim to compute the curvature tensor, which is related to both the first and second fundamental forms, as described above in the smooth setting. We note that if we are only computing the curvature in unit tangent directions, which is what we will want to do, that the curvature tensor reduces to the second fundamental form. The second fundamental form is indeed a quadratic form, and in the case of surfaces and curves it is easy to compute the eigenvalues of a small matrix.

This type of algorithm first computes the vertex normals in the direction of the neighboring points as described above. we can then compute the second fundamental form according to some orthonormal basis of the tangent plane. We then compute the eigenvalues of this matrix, which gives us the principal directions and the curvature.

This technique was originally presented by Taubin in [Tau95]. Recently, tensor voting has been proposed as a way to estimate the tensor of arbitrary point clouds [LCC08, GM12, KSNS]. They also also variable size neighborhoods as a method of smoothing to allow the algorithm to be robust against noise. However, I have not come across applications of this method to high dimensional manifold estimation.

4.4. Integral Invariants. Multiresolution analysis refers to the idea that we can estimate the curvature at different scales. The reason that multiresolution analysis is desired is that when noise is introduced into the dataset, we can estimate the curvature at a larger scale which acts as a way to smooth our estimates of the curvature or other differential properties of the manifold.

An interesting method of estimating the curvature is the integral invariant. The idea of the integral invariant is that if we integrate some function on a local patch of the manifold that the surface area, signed distance, or volume, will be a function of the principal curvatures at these points. This comes about by

considering the covariance matrix of the data points in some local patch, which results in several integrals which can then be approximated using the curvature.

Integration is inherently robust to noise, especially when considered to differentiation, since we are not taking differences, but rather sums. If we perturb the data points slightly, the value of the integral will not change significantly. That is that if we are computing integrals at some scale S_0 , features that are at a scale significantly smaller than S_0 , will not effect the value of the integration.

Integral invariants are a relatively new technique for estimating differential quantities, as seen in [PWY⁺07, MCH⁺06]. We note that integral invariants are can be used for estimating many differential quantities, such as the principal directions, principal moments of inertia, and any other quantities that can be related to the value of the integral of a surface patch of the manifold.

The intuition behind integral invariants is that we can locally approximate $\mathcal{M}$ as a surface that has coefficients that are exactly the principal curvatures, using the second fundamental form. We then define a functional which integrates some function over a region, that we approximate using our surface estimation that has the principal curvatures. We can then numerically compute this integral, and using our approximations solve for the curvature or a function of it. Consider Figure 2 below as an example.

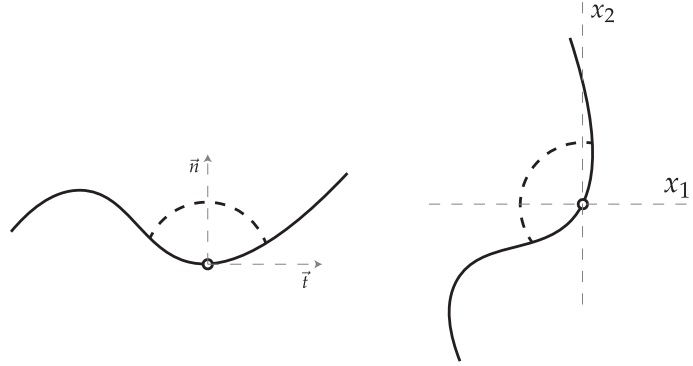


FIGURE 2. On the left, we see the manifold as it is presented to us in the original coordinate frame. We compute the length of the dotted region. On the right, we can consider the same manifold locally rotated such that it is aligned with the principal curvatures and directions. In this frame we can show how the length of this dotted region is related to the curvature.

The length of the curve or the value of any integral over this dashed region is *invariant* under this rotation into the proper frame. In the proper frame, we can relate the integration over this region to the curvature of the manifold, locally.

My extensions to higher dimensional submanifolds will be based on the ideas of integral invariants.

4.5. Discrete Varifolds. While learning geometric measure theory (GMT), I started to wonder about the notion of discrete varifolds since I have seen other concepts from GMT, such as normal cycles, manifold reach, and distance measures, show up in manifold approximation. In this search, I came across a small body of work that presents a notion of discrete varifolds and has shown extremely promising results for their use in extracting geometric information, specifically mean curvature. I just wanted to briefly note that this seems like a promising line of research that should be considered and may have deep properties when attempting submanifold reconstruction.

All of the ideas that that I am referencing have only been introduced over the past several years in [Bue14, BLM15].

Many of the ideas of discrete varifolds for estimating discrete curvature originates from the connection between mean curvature and the first variation of area. Varifolds were initially developed to study critical points of the area functional. It has been shown that *both* smooth and discrete objects can be endowed with a varifold structure, which provides a unified framework for analyzing geometric properties.

The main idea to take away from the definition of a varifold is that it encodes both spatial and tangential information. In a sense, a varifold structure is a measure whose support is the tangent bundle $T\mathcal{M}$. The

authors consider volumetric varifolds, that is that they are associated with the volumetric representation of an object. They then provide a consistent notion of ε -mean curvature at some scale in both the discrete and smooth case. They end by providing extremely promising results of estimating the mean curvature from point cloud samples that have extremely high accuracy.

5. THE FRAMEWORK FOR HIGHER DIMENSIONAL INTEGRAL INVARIANTS

In all of the literature on integral invariants that I have come across, I have not found anything on generalizing these techniques for higher dimensions. My belief for why this is the case is that these techniques have primarily only been used in computational geometry and low-dimensional shape matching, and the manifold learning community is still relatively small and therefore there hasn't been much work done on estimating the curvature of higher dimensional manifolds. Furthermore, the applications that follow understanding the curvature in high dimensions have also not been discussed, so there has been relatively little motivation for this generalization. However, one should note that extensions of these notions to higher dimensions is not necessarily very difficult, besides the conceptual complexity and many of the technical nuances that arise when considering spaces of arbitrary dimension. Note that my work in the following sections closely follows the development of integral invariants in [PWY⁺07], and differs only in the fact that these notions are generalized to arbitrary dimensions. In that regard, the work is not novel and I am primarily interested in the experimental results of these generalizations in arbitrary dimension.

Throughout my analysis, I assume that the dimension of $\mathcal{M}$ is known. Furthermore, I assume that the tangent and normal spaces are known. Note that in practice these are not trivial problems to solve. However, there is no lack of research done on dimensionality estimation of arbitrary point cloud data. I also discuss some ideas on how to accurately estimate the basis of the tangent and normal spaces.

We build on the background discussed at the beginning of the paper on principal curvatures. Let $\mathcal{M}$ be an m -dimensional manifold embedded in $\mathbb{R}^n$. Let $d = n - m$ denote the codimension of $\mathcal{M}$. Recall that if $d = 1$ then we can approximate the changes of $\mathcal{M}$ in the normal direction as we move along a tangent vector locally using a function of the tangent vector. Let $p \in \mathcal{M}$ and consider a local patch around p , we can then describe the coordinates of $\mathcal{M}$ in $\mathbb{R}^n$ of the form $(x_1, \dots, x_n, y)$ where we have that

$$y(x_1, \dots, x_n) = \frac{1}{2} \left(\sum_{i=1}^n \kappa_i x_i^2 \right) + \mathcal{O}(\|\mathbf{x}\|^3).$$

Note that for this to be true, we need to align the coordinate vectors with the principal directions at p .

Now, we will consider the case when $\mathcal{M}$ is of arbitrary codimension. Consider $p \in \mathcal{M}$. Recall that $\dim(T_p \mathcal{M}) = \dim \mathcal{M} = m$, and furthermore that $\dim(T_p \mathbb{R}^n) = n$. So we have that $\dim((T_p \mathcal{M})^\perp) = d$. At p , let $\{\eta_1, \dots, \eta_d\}$ be an orthonormal basis of the normal space $(T_p \mathcal{M})^\perp$. Similar to the case of $d = 1$, we can locally write *each* remaining $(n - m)$ coordinates of $\mathcal{M}$ in $\mathbb{R}^n$ as a function of the principal curvatures corresponding to the $n - m$ normals at $p \in \mathcal{M}$. So we have $(n - m)$ functions of the form

$$y_i = f_i(x_1, \dots, x_n) = \frac{1}{2} \left(\sum_{j=1}^n \kappa_j^{(i)} x_j^2 \right) + \mathcal{O}(\|\mathbf{x}\|^3),$$

where $\kappa_j^{(i)}$ is the j th principal curvature corresponding to the normal η_i . That is, we can write the coordinates of $\mathcal{M}$ that is embedded in $\mathbb{R}^n$ as having the following local coordinates:

$$(x_1, \dots, x_m, y_1, y_2, \dots, y_{n-m}).$$

Due to standard linear operator theory, we can indeed choose a coordinate basis that diagonalizes this quadratic form. Note that this basis is that which is aligned to the principal directions. Hence, we have coordinates, using the proper basis,

$$(x_1, \dots, x_n, y_1, \dots, y_d)$$

that locally approximate the manifold quadratically.

Fix $r \in \mathbb{R}$. Let Π_S be the projection of the manifold $\mathcal{M}$ onto the subspace S . Now let

$$\mathcal{N}_i := \Pi_{T_p \mathcal{M} \times \text{span}\{\eta_i\}} \cong \mathbb{R}^{m+1}.$$

Let $d\mathbb{R}^m$ be the standard (induced) volume form on $\mathbb{R}^m$ and let $d\mathbb{R}^{m+1}$ be the standard volume form on $\mathbb{R}^{m+1} \cong \mathcal{N}_i$.

Consider the region

$$\Omega_r = \{\mathbf{x} \in \mathcal{N}_i : \mathbf{x}^\top \mathbf{x} \leq r^2, x_{\eta_i} \geq f_i(x_1, \dots, x_m)\},$$

where x_{η_i} is the component of $\mathbf{x}$ that corresponds to the normal vector η_i , as visualized in Figure ?? . This region is the portion of the ball in $\mathbb{R}^{m+1}$ that lies above the curve. Define the integral invariant I as the functional

$$I(h) = \int_{\Omega_r} h(\mathbf{x}) d\mathbb{R}^{m+1}.$$

Using this framework, I am now in a position to develop approximations for these integral invariants and show how they can be used for estimating the mean curvature with respect to each normal η_i .

6. APPROXIMATING THE MEAN CURVATURE USING INTEGRAL INVARIANTS

To approximate the mean and principal curvatures, we need to first approximate $I(h)$ in a manner that is related to the principal curvatures of the manifold. Once we understand how $I(h)$ is related to the principal curvatures, we will show that we can estimate the mean curvature by $I(1)$ and furthermore, we can estimate all of the principal curvatures by consider the covariance matrix which consists of terms of the form $I(x_i x_j)$.

There are two lemmas presented that fully characterize this estimation. The first lemma is of primary importance when estimating all of the principal curvatures. Note that throughout this section, I somewhat abuse the big-O notation by not always treating it as a set (i.e. $f(x) = \mathcal{O}(x)$).

Lemma 6.1. *This Lemma provides an accurate method to estimate the functional I with provable error bounds.*

Assume that $h \in \mathcal{O}(\rho^k \cdot (x_{\eta_i})^l)$. Consider

$$\hat{I}(h) = \int_{\mathbb{B}_r^+} h(\mathbf{x}) d\mathbb{R}^{m+1} - \int_{\mathbb{B}_r^m} \left(\int_{x_{\eta_i}=0}^{x_{\eta_i}=\frac{1}{2} \left(\sum_{j=1}^m \kappa_j^{(i)} x_j^2 \right)} h(\mathbf{x}) dx_{\eta_i} \right) d\mathbb{R}^m,$$

where

$$\mathbb{B}_r^+ = \{\mathbf{x} \in \mathcal{N}_i : \mathbf{x}^\top \mathbf{x} \leq r^2, x_{\eta_i} \geq 0\} \quad \text{and} \quad \mathbb{B}_r^m = \{\mathbf{z} \in \mathbb{R}^m : \mathbf{z}^\top \mathbf{z} \leq r^2\}.$$

We claim that

$$I(h) - \hat{I}(h) \in \mathcal{O}(r^{k+2l+m+3}).$$

Proof. Consider

$$\tilde{I}(h) = \int_{\mathbb{B}_r^+} h(\mathbf{x}) d\mathbb{R}^{m+1} - \int_{\mathbb{B}_r^m} \left(\int_{x_{\eta_i}=0}^{f_i} h(\mathbf{x}) dx_{\eta_i} \right) d\mathbb{R}^m.$$

Notice that the error between $\tilde{I}(h)$ and $I(h)$ is the region that is outside of $\mathbb{B}_r^+$, but that is still bounded by the surface in a cylindrical radius r . Let E^* denote this error region. The first thing to estimate is how wide is the region, E^* , that is contained in the cylinder, but that is outside of the sphere. Consider δ , a constant, such that for all $\|\mathbf{x}\| \leq r$ we have that $|f_i(x_1, \dots, x_m)| \leq \delta \|\mathbf{x}\|^2$. Notice that our error region goes at most as far as r , and begins at least at $r - \frac{1}{2}\delta r^3$. This is found by looking at where a ball of radius r intersects $\delta \|\mathbf{x}\|^2$. Since we are integrating over $\mathbb{B}_r^m$, we wrap this width around the surface area of this sphere many times, that is

$$\frac{2\pi^{m/2}}{\Gamma(m/2)} r^{m-1},$$

which is clearly on the order r^{m-1} . Lastly, we can bound the height of $\mathcal{M}$ locally using our Taylor series quadratic expansion on the order of r^3 . We are now in a position to bound

$$\int_{E^*} h(\mathbf{x}) d\mathbb{R}^{m+1}.$$

Note that

$$\begin{aligned}
 \int_{E^*} h(\mathbf{x}) \, d\mathbb{R}^{m+1} &\leq \max_{\mathbf{z} \in E^*} |h(\mathbf{z})| \, \text{vol}(E^*) \\
 &\approx \mathcal{O}(r^{m-1} \cdot r^3 \cdot r^2) \max_{\mathbf{z} \in E^*} |h(\mathbf{z})| \\
 &\approx \mathcal{O}(r^{m+4}) \cdot \mathcal{O}(r^k r^{2l}) \\
 &\approx \mathcal{O}(r^{k+2l+m+4}).
 \end{aligned}$$

Now we cut off the surface f_i to it's quadratic term. That is, we are aiming to estimate the error when using $\hat{I}(h)$. Notice that the region E^* now is the region of the cylinder bounded by $f_i(x_1, \dots, x_m)$ and is outside of the sphere. Clearly, the region still has width on the order of r^3 since we can choose δ as before. Now we have that $\text{vol}(E^*)$ now is of the order

$$r^{m-1} \cdot \mathcal{O}(r^{k+2l}) \int_0^r \mathcal{O}(\rho^3) \rho \, d\rho \approx \mathcal{O}(r^{k+2l+m+3}).$$

So we have shown that

$$I(h) = \hat{I}(h) + \mathcal{O}(r^{k+2l+m+3}).$$

□

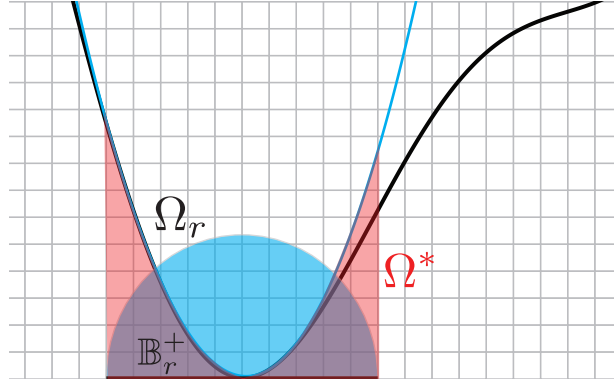


FIGURE 3. The regions denoted in the proof are visualized in this simple case of $m = 1$ and $n = 2$. Ω_r is the blue region contained above the function f . Ω^* , the error region, is only the bright red. $\mathbb{B}_r^+$ is the entire hemisphere. Lastly, the total red area that is shaded under the polynomial approximation of the curve is the term that is subtracted in $\hat{I}(h)$.

Lemma 6.2. *We can estimate the mean curvature $\mathcal{H}_i$ at some scale r as*

$$\hat{\mathcal{H}}_i^{(r)} \approx \frac{\Gamma\left(\frac{m+1}{2}\right) \cdot (m+2)}{\pi^{\frac{m+1}{2}} r^{m+2}} \left(\text{vol}(\mathbb{B}_r^+) - \text{vol}(\Omega_r) + \mathcal{O}(r^{m+3}) \right).$$

Proof. We can estimate the mean curvature with respect to the normal η_i , denoted as $\mathcal{H}_i$, by computing $I(1)$.

So consider

$$I(1) \approx \int_{\mathbb{B}_r^+} d\mathbb{R}^{m+1} - \int_{\mathbb{B}_r^m} f_i(x_1, \dots, x_m) d\mathbb{R}^m$$

Notice that for all $i, j \in \{1, \dots, m\}$ and $i \neq j$ that

$$\int_{\mathbb{B}_r^m} x_i x_j \, d\mathbf{x} = 0,$$

due to symmetry. On the other hand, we have that

$$\begin{aligned} \int_{\mathbb{B}_r^m} x_i^2 \, d\mathbb{R}^m &= \frac{1}{m} \int_{\mathbb{B}_r^m} \langle \mathbf{x}, \mathbf{x} \rangle \, d\mathbb{R}^m = \frac{1}{m} \int_{\Theta^*} \int_0^r \rho^2 \rho^{m-1} d\rho d\Theta = \frac{r^{m+2}}{m(m+2)} (\text{surface area of } \mathbb{B}_1^m) \\ &= \frac{2\pi^{(m+1)/2}}{\Gamma((m+1)/2) \cdot m(m+2)} r^{m+2}, \end{aligned}$$

where the ρ^{m-1} term comes from the Jacobian of the differential. We can now estimate

$$\int_{\mathbb{B}_r^m} f_i(\mathbf{x}) \, d\mathbb{R}^m \approx \frac{\pi^{(m+1)/2} \mathcal{H}_i}{\Gamma((m+1)/2) \cdot (m+2)} r^{m+2} + \mathcal{O}(r^{m+3}),$$

where $\mathcal{H}_i$ is the mean curvature. We get this by recalling that $\mathcal{H}_i = \frac{1}{m} \text{trace}(S_{\eta_i}) = \frac{1}{m} \text{trace}(\Pi_p^{(\eta_i)})$ and that we can locally express $\mathcal{M}$ using a Taylor expansion based on $\Pi_p^{(\eta_i)}$.

Now consider $I(1)$ again. We have that

$$\begin{aligned} \hat{I}(1) &\approx \int_{\mathbb{B}_r^+} d\mathbb{R}^{m+1} - \left(\int_{\mathbb{B}_r^m} f_i(x_1, \dots, x_m) \, d\mathbb{R}^m \right) \\ &\approx \int_{\mathbb{B}_r^+} d\mathbb{R}^{m+1} - \left(\frac{\pi^{(m+1)/2} \mathcal{H}_i}{\Gamma((m+1)/2) \cdot (m+2)} r^{m+2} + \mathcal{O}(r^{m+3}) \right). \end{aligned}$$

So we have that

$$\hat{\mathcal{H}}_i^{(r)} \approx \frac{\Gamma\left(\frac{m+1}{2}\right) \cdot (m+2)}{\pi^{\frac{m+1}{2}} r^{m+2}} \left(\text{vol}(\mathbb{B}_r^+) - \text{vol}(\Omega_r) + \mathcal{O}(r^{m+3}) \right).$$

□

One can easily check that the case of a surface embedded in $\mathbb{R}^3$ (when $m = 2$), that this estimation of the mean curvature at some scale r does indeed coincide with what has been found in [PWY⁺07].

Note that the mean curvature vector in this direction is just $H_i \eta_i$. Furthermore, we can derivate the total mean curvature vector of $\mathcal{M}$ by summing all of the mean curvatures in each codimension. Above, I did all of my analysis with respect to the specific codimensions due to the simplicity in extending this analysis, and to furthermore, be able to estimate principal curvatures.

I was originally working on just using the mean curvature vector in arbitrary codimension, but I was running into some initial mathematical and conceptual difficulty. So I reverted to this simpler case that we can use when we have estimates for the tangent and normal spaces. However, after going through this preliminary analysis I am not necessarily sure that working with the mean curvature vector in the full ambient space is directly possible. The difficulty of using integral invariants when we have very high codimension is that we are generally computing the integral over some region and then removing an estimate of the error, so when we have a high codimension the bounds that I have been able to show on the error of different estimates have not been very tight.

Furthermore, notice that we can compute these integrals, even without knowing the underlying manifold using Monte Carlo methods or other numerical integration techniques. Another approach that has been used for curvature estimation is based on convolutions [PWY⁺07] This integration is extremely easy to do when we have a triangulated manifold exactly, and we can approximate the boundary in the case of a point set. However, I think it would be extremely interesting to look into integration of manifolds based on a finite sample of the manifold from a less statistical standpoint and a more traditional differential geometric and information theoretic perspective.

7. ESTIMATING THE PRINCIPAL DIRECTIONS WITH RESPECT TO THE NORMAL η_i

To estimate the principal curvatures, we will consider the covariance matrix of the data, which we can approximate using our integral invariants. We then compute the eigenvalues and eigenvectors and show their relationship to the principal curvatures and directions. We can strongly reason that if there is a small amount of noise that the estimates of the principal curvatures and directions will still be relatively low using random matrix theory and eigenvalue perturbation theory.

Let $\Sigma(\Omega_r)$ be the covariance matrix, whose entries are the covariance of the coordinates over the region Ω_r . Explicitly, we will be computing

$$\Sigma(\Omega_r) = \int_{\Omega_r} (\mathbf{x} - \mathbf{m})(\mathbf{x} - \mathbf{m})^\top d\mathbb{R}^{m+1} = \int_{\Omega_r} \mathbf{x}\mathbf{x}^\top d\mathbb{R}^{m+1} - \text{vol}(\Omega_r)\mathbf{m}\mathbf{m}^\top,$$

where $\mathbf{m}$ is the barycenter of the data. Note that the ℓ^{th} coordinate of the matrix $\int_{\Omega_r} \mathbf{x}\mathbf{x}^\top d\mathbb{R}^{m+1}$ is $\hat{I}(x_\ell x_j)$. Notice that for all ℓ, j such that $\ell \neq j$, we have that $\hat{I}(x_\ell x_j) = 0$, due to the symmetry. So we only need to compute $\hat{I}(x_\ell^2)$ for all $\ell = 1, \dots, m$. Furthermore, to compute the barycenter $\mathbf{m}$, we only need to compute $\hat{I}(x_\ell)$ for all $\ell = 1, \dots, m$ and $\hat{I}(x_{\eta_i})$. The computations of the barycenter are as follows:

$$\begin{aligned} \hat{I}(x_\ell) &= 0 \quad (\text{since } x_\ell \text{ is odd}), \\ \hat{I}(x_{\eta_i}) &= \int_{\mathbb{B}_r^+} x_{\eta_i} d\mathbb{R}^{m+1} - \int_{\mathbb{B}_r^m} \left(\frac{1}{2} \left(\frac{1}{2} \sum_{j=1}^m \kappa_j^{(i)} x_j^2 \right)^2 \right) d\mathbb{R}^m \\ &= \int_{\mathbb{B}_r^+} x_{\eta_i} d\mathbb{R}^{m+1} + \mathcal{O}(r^{4+m}). \end{aligned}$$

We then consider a standard spherical parameterization of the n -sphere and integrate to get

$$\begin{aligned} \hat{I}(x_{\eta_i}) &= \frac{r^{m+2}}{m+2} \int_{\theta_m=0}^{\theta_m=\pi} \cdots \int_{\theta_1=0}^{\theta_1=\pi} \sin^m(\theta_1) \sin^{m-1}(\theta_2) \cdots \sin(\theta_m) d\theta_1 d\theta_2 \cdots d\theta_m + \mathcal{O}(r^{4+m}) \\ &= \frac{r^{m+2}}{m+2} \prod_{k=1}^m \frac{\sqrt{\pi} \Gamma\left(\frac{k+1}{2}\right)}{\Gamma\left(\frac{k}{2} + 1\right)} + \mathcal{O}(r^{m+4}). \end{aligned}$$

Notice that the integration $\hat{I}(x_\ell^2)$ will weight $\kappa_\ell^{(i)}$ by a factor of 3 more than any of the other principal curvatures. I need to do the compute all of these in arbitrary dimension and compute their Taylor series expansion before I can actually write any code to do this.

Notice that these computations are indeed simple, but start become very complex and convoluted when considering the computations of the general case of dimension m . I haven't computed a closed form expression yet for all of the eigenvalues of this covariance matrix yet.

Using the expression for $\Sigma(\Omega_r)$, we can compute the eigenvalues, which are just the diagonal entries since all of the other entries are zero. Using these eigenvalues, which are expressed in terms of the principal curvatures, we will be able to extract the individual principal curvatures, as well as the principal directions, by careful algebraic manipulation. We will then have that the unit normal vector η_i will be parallel to one of the eigenvectors. We then have that the rest of the eigenvectors correspond to the principal curvatures in this codimension.

8. DIFFICULTY, ISSUES, AND NOTES RELATING TO THIS FRAMEWORK

In this section, I will discuss the main practical issues with this framework and this analysis. I will then informally discuss some ideas on how to fix these problems so that we can proceed with the analysis. I will conclude by providing some general remarks and notes about the framework and assumptions that I am making.

8.1. Difficulty of Estimating the Tangent and Normal Spaces. It is important to note that this framework implicitly requires an estimate of the normal space at a point $p \in \mathcal{M}$. The *only* reason that they are needed is that we are doing many of our computations in the projected space $\mathcal{N}_i$.

The problem is now to estimate all of the normal vectors of $\mathcal{M}$ as it lies embedded in $\mathbb{R}^n$. If the points lie exactly on the manifold and are sampled densely enough, we can consider the covariance matrix of the data in a sufficiently small neighborhood, compute the eigenvectors and the $n - m$ eigenvectors corresponding to the $n - m$ smallest eigenvalues will provide a basis for the normal space at some point p .

However, once we introduce outliers into our dataset, we run into issues where our estimation of the normal vectors can become extremely unstable. Especially if we are considering a neighborhood size that does indeed contain curvature elements. Note that for estimation of the normal basis vectors we want to

consider a radius that is much smaller than the radius that we want to consider for estimating the mean curvature. I have just started looking into this, and don't have much done yet as I'm still trying to formalize all of the details and work on the actual analysis. The idea is very similar to statistical bootstrapping at two levels. The first is given a neighborhood of radius r around a point p , we compute the original estimate of the tangent plane using a Robust PCA. We can then do another bootstrapping type procedure to determine the optimal scale at which we use to determine the basis of our tangent and normal spaces. That is we choose different neighborhood sizes, and many times choose a random sample of the points in this space. Apply the Robust PCA algorithm described to these data points. Repeat. We then choose the radius size r that minimizes the average variance of all of the random bootstrap samples in that neighborhood size with the original estimate that we derived for some fixed neighborhood radius r_0 .

Note that if there is some Gaussian noise in the neighborhood of radius r , we should be able to use eigenvalue perturbation theory to show that the error in the estimates around this point are relatively low. Something that I would like to look into is bounding the error of the mean curvature estimates in each codimension as a function of the error of the normal estimate.

8.2. The Effects of Noise on Curvature Estimation. It is clear that noise may corrupt the inference of any geometric property, such as curvature. What I hope to do after I complete this analysis is to study the first variations of the volume Ω_r with respect to some vector field Y . The idea is then to try and bound the variation with respect to the maximum value of Y . Using this first variation, we can determine the maximum scale of the variance that will bound the error of our mean curvature estimate.

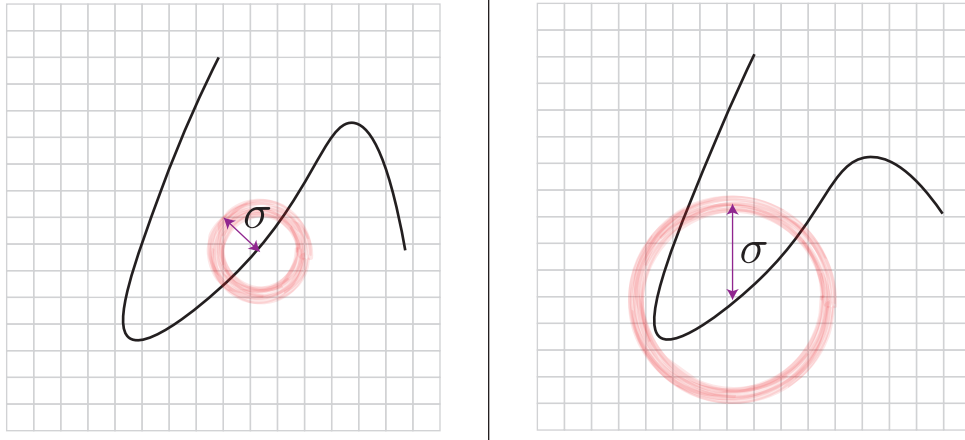


FIGURE 4. If the variance of the data is relatively low, the noise does not corrupt the curvature estimates; however, if the variance becomes too large, as is the case of the illustration on the right, the underlying structure of the manifold cannot be recovered.

Consider Figure 4 for an illustration of the issues of introducing noise when attempting to estimate the curvature of the underlying manifold. In light of this illustration, we need to note that the scale of the variance of the data needs to be considerably smaller than the scale of the curvature. Besides understanding this connection, there is relatively little analysis that I have done in regards to this noise-curvature interaction so far.

9. APPLICATIONS OF KNOWING THE CURVATURE OF HIGH DIMENSIONAL SUBMANIFOLDS

A central problem when considering discrete manifolds is determining whether two objects in a space are equivalent in some manner. When working with such complex structures, it is often beneficial to abstract the individual data points to higher level features. Understanding the mean and principal curvatures are indeed very strong features that are indicative of the underlying structure of the data. One idea is that you can use the curvature of submanifolds as a signature that can be used to distinguish different discrete submanifolds from each other.

One of the reasons that understanding if two submanifolds is, in general, not a simple problem is that if the codimension of the space is large and two submanifolds are the same up to rigid transformation,

without a complex realignment procedure we cannot determine whether they are the same or not. Furthermore, we cannot talk about how much they differ. The idea is to use differential invariants to compare these shapes. The study of invariants is extremely prevalent in differential geometry since it abstracts the complicated notions of an object to a few much simpler descriptors. This idea is also seen a lot in computational topology, where we are considering the persistent homology as a descriptor of the data.

Curvature has been used in low-dimensions as a shape descriptor and I would like to explore the effectiveness of curvature as an invariant when considering high dimensional submanifolds. To do this, I want to start looking at real datasets, understand the curvature of these spaces, and try and look if there are any meaningful connections that we can derive from these signatures.

One can refer to [Epp15, Joh01] to see how curvature has been used in the computer vision community in the past as a shape descriptor.

Another application of understanding the curvature, is to have a build towards a better understanding of submanifold reconstruction via triangulations. If we can estimate the curvature at some scale, in a manner that is robust to error, it should be possible to incorporate this information at the very least in understanding the proper neighborhood structure. The information could be taken a step further, to compute more accurate estimations of the tangent plane and to estimate geodesic distances without necessarily traversing a neighborhood graph.

10. NONPARAMETRIC BOOTSTRAP FOR OPTIMAL SCALE SELECTION

Note that the linear functional I performs the integration of some function h over some region that is part of the boundary of an open ball with some radius r . Since we can choose this scale parameter r , it is natural to ask what is the optimal scale that captures the underlying curvature with the least error?

To answer this question we can turn to the extensive literature of nonparametric statistics. We can think of our mean (or principal) curvature estimators, $\hat{\kappa}_r$, as nonparametric estimators of the underlying curvature at some scale r . Just as any nonparametric estimators have smoothing parameters, we have that our scale parameter is indeed a smoothing parameter. Consider Figure 5 to see the effects of changing the scale r .

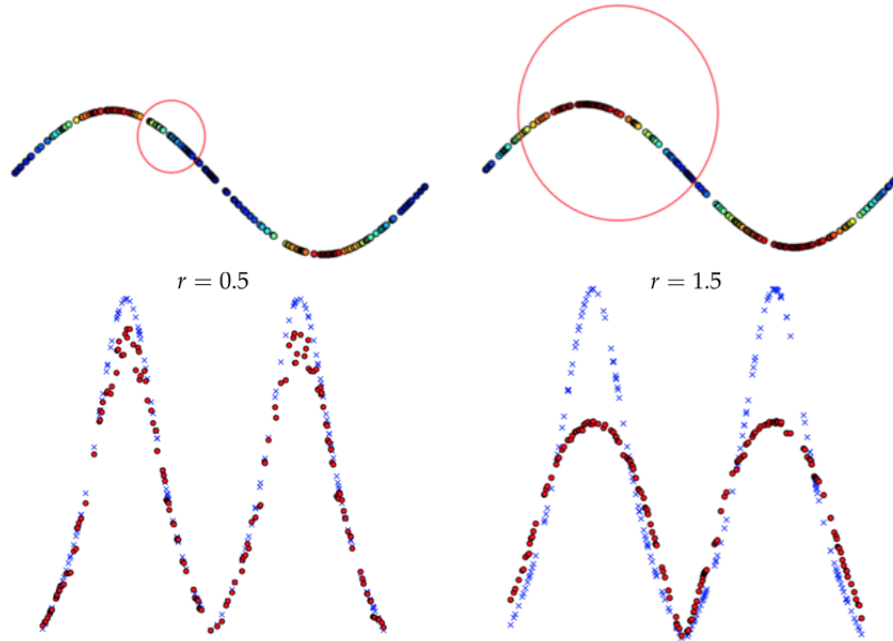


FIGURE 5. We consider an example where we compute the estimates of the curvature of a sine wave at two different scales. The top graph has the curve embedded in $\mathbb{R}^2$. The bottom graph plots the estimated curvature (in red) and the true curvature (in blue) as a function of time, which completely parameterizes the curve.

Since we can choose different smoothing parameters we intuitively want to choose the largest scale that we can that does not capture any features besides the curvature. Note that when we choose a larger scale, any noise that occurs at a scale that is much less than the scale of integration has a very small effect on the curvature estimate.

With this idea in mind, we can aim to minimize the mean squared error, which is defined as

$$\mathbb{E}(\hat{\kappa}_r - \kappa)^2.$$

The most important thing to note is that the mean squared error can be decomposed into the variance of the estimator and the squared bias. Notice that at a small scale, we have that the bias is relatively low, however the variance is relatively high. On the other hand, if we choose a larger smoothing parameter, we have that the variance is relatively low, but the bias is relatively high. Of course, we cannot consider the variance and bias if we only have one dataset, so we think of the data as random variables that are sampled from a uniform distribution over the submanifold.

Of course, we do not have the underlying variance or bias of our estimator, so we instead estimate it using a nonparametric bootstrap. The nonparametric bootstrap procedure is as follows:

- (1) Compute an initial estimate of κ by choosing some scale according to some prior, denote this as $\hat{\kappa}_0$
- (2) Estimate the empirical cumulative distribution function $\hat{F}_n(x)$
- (3) Draw n observations from the empirical CDF
- (4) Compute the bootstrap estimate $(\hat{\kappa}_r^*)_i$
- (5) Repeat B times
- (6) Compute the estimated mean squared error by replacing the true κ with $\hat{\kappa}_0$.

We can then continue drawing scale parameters r according to some prior and we choose the scale r that minimizes the mean squared error. We can then update our distribution of r to some posterior such that our search becomes more and more refined as we traverse the submanifold.

In the next section we provide some results and show how this optimal scale selection improves performance drastically.

11. RESULTS AND CONCLUSIONS

I ran the curvature estimator on several different curves. I primarily tested the effects of noise and sparsity on the standard sine wave:

$$x_1, \dots, x_n \sim \text{Uniform}(0, 2\pi), \quad y_i = \sin(x_i) + \varepsilon_i, \quad \varepsilon_i \sim \mathcal{N}(0, \sigma^2).$$

We can see the results on the circle and parabola below in Figure 6.

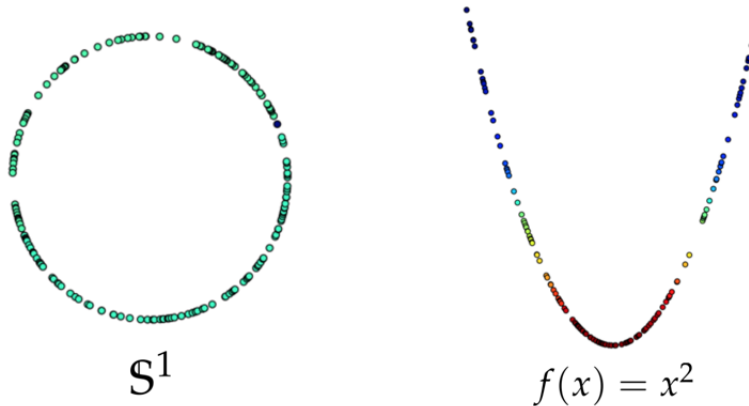


FIGURE 6. Visually, the results of the curvature estimator on the circle S^2 and the parabola are correct.

Consider Figure 7 below to see the effects of sparsity and noise.

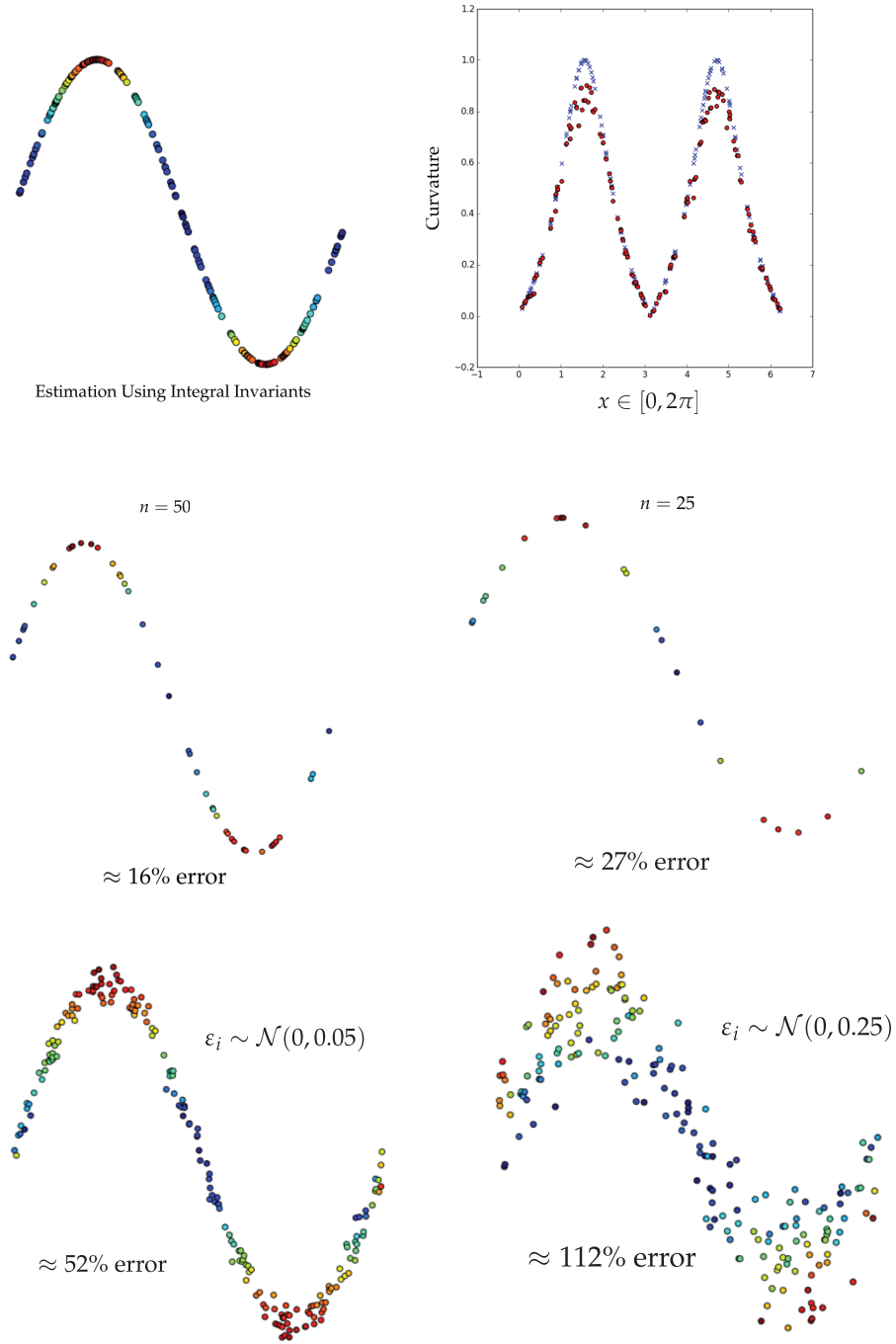


FIGURE 7. Here we consider a sine wave and we can see the effects of sparsity and noise on the overall curvature estimation. We note the average percent error of the estimate for each of these curves.

Consider Figure 8, on the next page, to see how this optimal scale selection improves the estimation accuracy.

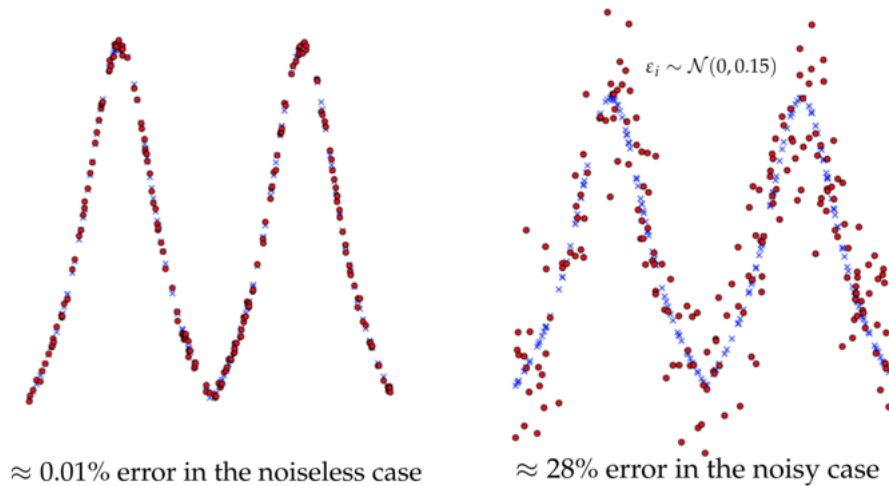


FIGURE 8. By applying the optimal radius selection, we see that the accuracy drops significantly in the noiseless case. However, in the noisy case, the optimal scale selection cannot attempt to do any fine tuning so it chooses a scale that will oversmooth the submanifold.

12. SHORTCOMINGS AND FUTURE WORK

I was planning on completing the eigenanalysis of the covariance matrix in arbitrary dimensions, as well as properly analyze the bootstrap methodology for estimating the tangent and normal spaces. I aimed to analyze, in a rigorous manner, the effects of noise on the data through the first variation of the volume. *Unfortunately*, I did not have time to do all of the theoretical work on showing tight error bounds. This is something that I plan on completing soon.

The last issue that I ran into has to do with the numerical integration itself of high dimensional submanifolds where the boundary is only determined by some noisy points.

Overall, I think that the problem of geometric inference of differential properties has to be thought of in an invariant way due to the fact that small perturbations of the data *will* grossly impact the true differential properties of the manifold that fits these points exactly. However, if we are looking at invariants, we know how to bound the first variations of area and we will be able to provide theoretical bounds on the error *before* observing our data.

REFERENCES

- [ABK98] Nina Amenta, Marshall Bern, and Manolis Kamvyselis. A new voronoi-based surface reconstruction algorithm. In *Proceedings of the 25th annual conference on Computer graphics and interactive techniques*, pages 415–421. ACM, 1998.
- [BLM15] Blanche Buet, Gian Paolo Leonardi, and Simon Masnou. Surface approximation, discrete varifolds, and regularized first variation, 2015.
- [BSW09] Mikhail Belkin, Jian Sun, and Yusu Wang. Constructing laplace operator from point clouds in $\mathbb{R}^d$. In *Proceedings of the twentieth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 1031–1040. Society for Industrial and Applied Mathematics, 2009.
- [Bue14] Blanche Buet. Quantitative conditions of rectifiability for varifolds. *arXiv preprint arXiv:1409.4749*, 2014.
- [CL96] Brian Curless and Marc Levoy. A volumetric method for building complex models from range images. In *Proceedings of the 23rd annual conference on Computer graphics and interactive techniques*, pages 303–312. ACM, 1996.
- [CP05] Frédéric Cazals and Marc Pouget. Estimating differential quantities using polynomial fitting of osculating jets. *Computer Aided Geometric Design*, 22(2):121–146, 2005.
- [DB02] Ioannis Douros and Bernard F Buxton. Three-dimensional surface curvature estimation using quadric surface patches. *Proceedings of the Scanning*, 2002.
- [Epp15] Sagi Eppel. Using curvature to distinguish between surface reflections and vessel contents in computer vision based recognition of materials in transparent vessels. *arXiv preprint arXiv:1506.00168*, 2015.
- [FJ89] Patrick J Flynn and Anil K Jain. On reliable curvature estimation. In *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, volume 89, pages 110–116. Citeseer, 1989.
- [GM12] Dian Gong and Gérard Medioni. Probabilistic tensor voting for robust perceptual grouping. In *Computer Vision and Pattern Recognition Workshops (CVPRW), 2012 IEEE Computer Society Conference on*, pages 1–8. IEEE, 2012.
- [GPPVW10] Christopher Genovese, Marco Perone-Pacífico, Isabella Verdinelli, and Larry Wasserman. Minimax manifold estimation. *arXiv preprint arXiv:1007.0549*, 2010.

- [GVL12] Gene H Golub and Charles F Van Loan. *Matrix computations*, volume 3. JHU Press, 2012.
- [HDD⁺92] Hugues Hoppe, Tony DeRose, Tom Duchamp, John McDonald, and Werner Stuetzle. *Surface reconstruction from unorganized points*, volume 26. ACM, 1992.
- [Joh01] Björn Johansson. Multiscale curvature detection in computer vision. 2001.
- [KBH06] Michael Kazhdan, Matthew Bolitho, and Hugues Hoppe. Poisson surface reconstruction. In *Proceedings of the fourth Eurographics symposium on Geometry processing*, volume 7, 2006.
- [KSNS] Evangelos Kalogerakis, Patricio Simari, Derek Nowrouzezahrai, and Karan Singh. Robust statistical estimation of curvature on discretized surfaces. Citeseer.
- [KSO04] Ravikrishna Kolluri, Jonathan Richard Shewchuk, and James F O'Brien. Spectral surface reconstruction from noisy point clouds. In *Proceedings of the 2004 Eurographics/ACM SIGGRAPH symposium on Geometry processing*, pages 11–21. ACM, 2004.
- [Kul70] Ravindra Shripad Kulkarni. Curvature and metric. *Annals of Mathematics*, 91(2):311–331, 1970.
- [LCC08] Gabriele Lombardi, Elena Casiraghi, and Paola Campadelli. *Advanced Concepts for Intelligent Vision Systems: 10th International Conference, ACIVS 2008, Juan-les-Pins, France, October 20-24, 2008. Proceedings*, chapter Curvature Estimation and Curve Inference with Tensor Voting: A New Approach, pages 613–624. Springer Berlin Heidelberg, Berlin, Heidelberg, 2008.
- [MCH⁺06] Siddharth Manay, Daniel Cremers, Byung-Woo Hong, Anthony J Yezzi, and Stefano Soatto. Integral invariants for shape matching. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 28(10):1602–1618, 2006.
- [MW00] Dereck S Meek and Desmond J Walton. On surface normal and gaussian curvature approximations given data sampled from a smooth surface. *Computer Aided Geometric Design*, 17(6):521–543, 2000.
- [PWY⁺07] Helmut Pottmann, Johannes Wallner, Yong-Liang Yang, Yu-Kun Lai, and Shi-Min Hu. Principal curvatures from the integral invariant viewpoint. *Computer Aided Geometric Design*, 24(8):428–442, 2007.
- [SN07] Márta Szilvási-Nagy. About curvatures on triangle meshes. *KoG*, 10(10.):13–18, 2007.
- [Tau95] Gabriel Taubin. Estimating the tensor of curvature of a surface from a polyhedral approximation. In *Computer Vision, 1995. Proceedings., Fifth International Conference on*, pages 902–907. IEEE, 1995.
- [TT05] Wai-Shun Tong and Chi-Keung Tang. Robust estimation of adaptive tensors of curvature by tensor voting. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 27(3):434–449, 2005.