Identifying Skeleton Curves in Noisy Data

Hanna Jankowski[1] and Larissa Stanberry[2]

[1]York University, [2]University of Washington

Abstract

This paper presents a Bayesian method to reconstruct the centerline in noisy data using B-spline curves. The method is illustrated on simulated two- and three dimensional data and is applied to recover the centerline of the colon in single photon emission computed tomography images.

Keywords: Centerline reconstruction, skeleton reconstruction, B-splines, colon imaging, SPECT data.

1 Introduction

In this paper, we consider the problem of identifying the centerline curve, or skeleton, in noisy data. Skeleton reconstruction frequently arises in practice and examples can be found in medical imaging [Deschamps and Cohen, 2001, Caffo et al., 2008], computer vision [Chaudhuri et al., 2004, Ma et al., 2003], and remote sensing [Banfield and Raftery, 1992]. Mathematically, the interior skeleton of a set $A \subset \mathbb{R}^d$ is defined as a set of points having a nonunique projection on the complement of A. Similarly, the exterior skeleton of A is defined as a set of points having a nonunique projection on A [Delfour and Zolésio, 2001].

A number of existing methods consider a skeleton reconstruction in high-resolution images. For example, the fast marching algorithm proposed in Cohen and Kimmel [1996] was extended to reconstruct a centerline of anatomical structures from high resolution 3D images [Deschamps and Cohen, 2001]. An automated 3D topological thinning was developed to reconstruct the colon skeleton from computed tomographic colonography data [Sadleir and Whelan, 2002]. A medial description of neuroanatomical structures was used to compare ventrical shapes and hippocampi between different subject populations [Styner et al., 2003, Bouix et al., 2005].

Here, we consider estimating the interior skeleton given a random sample of points from the set of interest. That is, given a cloud of points in $\mathbb{R}^d$, we seek to identify a smooth connected curve that passes through the middle of the data. This paper was motivated by the centerline reconstruction from single photon emission computed tomography (SPECT) images of the colon column [Caffo et al., 2008]. The reconstruction method was based on principal curves that were originally introduced as a nonlinear generalization of principal components [Hastie, 1984, Hastie and Stuetzle, 1989]. The principal curve is defined as follows. Let $\mathcal{C}$ be a smooth curve in $\mathbb{R}^d$ parameterized as $C(t)\colon [0,1] \mapsto \mathbb{R}^d$. Let the projection index $t_{\mathcal{C}}(y)$ be the value of the parameter t for which $C(t)$ is closest to y. If the projection is not unique, we take the projection index to be the largest one. The curve $\mathcal{C}$ is called self-consistent or a principal curve of the random variable X if $E[X|t_{\mathcal{C}}(X) = t] = C(t)$ for almost all t. The principal curve is estimated using the iterative rule $C^{(k+1)}(t) = E[X|t_{\mathcal{C}^{(k)}}(X) = t]$, where $\mathcal{C}^{(k)}$ is the kth iterate with parametrization $C^{(k)}(t)$. In practice, the distribution of X is unknown, and the expectation on the right hand side is estimated from the observed data. Although, the convergence of the algorithm was never established, the method was found to perform well on simulated data [Hastie and Stuetzle, 1989].

Since their introduction, principal curves have been adapted to solve a variety of problems. In particular, they were used to estimate the ice floe outlines in satellite images [Banfield and Raftery, 1992]. The curves were estimated using a modified algorithm which improves bias and variance of the iterative procedure for closed curves, but does not apply to open curves. The principal curves were combined with hierarchical clustering to estimate the fault lines in the earthquake zone [Stanford and Raftery, 1997]. Principal curves were also used to model the short term spectrum of the signals in the speech recognition problem [Reinhard and Niranjan, 1998]. In Caffo et al. [2008], the authors employed the original principal curves algorithm to identify the skeleton of the data but propose some modifications to improve the fit and increase the convergence speed. In particular, they introduced a warm-up procedure for the initialization step and resorted to using manually-specified endpoints to improve the fit in the tails. The method was shown to perform well on a variety of idealized (noise-free) images and was applied to reconstruct the centerline of the colon from SPECT data. Although presented as an improvement, the manual identification of end points clearly is a constraint since such points are often hard to identify, especially in high dimensional data.

In this paper, we propose a Bayesian method to reconstruct the smooth connected skeleton curve from noisy data. We model the skeleton as a twice differentiable B-spline curve, without imposing any shape constraints. In what follows, we describe the statistical model in Section 2 and give a brief overview of the B-spline curves. The algorithm for sampling from the posterior is given in Section 3. Further, in Section 4, we demonstrate the performance of the method on simulated data and apply the proposed algorithm to reconstruct the centerline of the colon from SPECT data. Final remarks in Section 4.4 adjourn the paper.

2 The Model

2.1 Likelihood

Let $Y = (Y_1, \ldots, Y_n)$ be the observed data with the unknown skeleton curve $\mathcal{C}$. We assume that $\mathcal{C}$ is twice differentiable and that it is parametrized by a mapping, $C : [0,1] \mapsto \mathbb{R}^d$. In general parametrization, we write $C(t) = (C_1(t), C_2(t), \ldots, C_d(t))$ for the position vector along the curve at time t in $\mathbb{R}^d$. Here, we consider planar and space curves in $\mathbb{R}^2$ and $\mathbb{R}^3$, respectively.

Let $C(s)$ be a natural parametrization of the curve $\mathcal{C}$ with arc length s as a parameter [Kreyszig, 1968]. Denote by $T(s) = C'(s)$ the unit tangent vector of $\mathcal{C}$ at point $C(s)$, where $C'(s)$ is the gradient vector of $C(s)$ with respect to s. If $T'(s) \neq 0$, then the unit vector $N(s) = T'(s)/||T'(s)||$ is the principal normal vector $N(s)$ of curve $\mathcal{C}$ at $C(s)$. The plane spanned by vectors $T(s)$ and $N(s)$ is known as the osculating plane of the curve at point $C(s)$. For curves in $\mathbb{R}^2$, the osculating plane, if exists, remains constant for all values of s, whilst for nonplanar curves, the tilt of the osculating plane changes with s. For curves in $\mathbb{R}^3$, consider the binormal $B(s)$ given by the vector product $B(s) = T(s) \times N(s)$. The triple $(T(s), N(s), B(s))$ constitutes an orthonormal right basis of $\mathcal{C}$ at $C(s)$ and is known as the Frenet frame. More details of the differential geometry of curves can be found in any standard textbook such as [Kreyszig, 1968].

Denote by δ_i the projection index of point Y_i on curve $\mathcal{C}$, as defined in Hastie and Stuetzle [1989]. For every point Y_i, let D_i be its signed distance to the skeleton curve $\mathcal{C}$, so that $|D_i| = |Y_i - C(\delta_i)|$. For planar curves, the sign of D_i is determined with respect to the normal to the curve $N(\delta_i)$ at $C(\delta_i)$ as

$\mathrm{sgn}\{D_i\} = \mathrm{sgn}\{\cos(Y_i - C(\delta_i), N(\delta_i))\}$. For curves in $\mathbb{R}^3$, the sign of the distance is taken with respect to the Frenet basis at the corresponding projection, i.e. $\mathrm{sgn}\{D_i\} = \mathrm{sgn}\{\cos(Y_i - C(\delta_i), B(\delta_i))\}$.

We assume that given the true skeleton curve $\mathcal{C}$, the distances D_i to the curve are independent draws from a centered normal distribution with known variance σ^2. Note that under this model, we only need to know the squared distances, so the sign computations are not required. From the Bayes theorem, we obtain that $\pi(\mathcal{C}|\, Y) \propto \mathcal{L}\,(Y|\mathcal{C})\,\pi(\mathcal{C})$.

2.2 Boundary Prior

In this section, we construct the boundary prior to control the behavior of the curve. We begin with the planar curves. Recall that for curve $\mathcal{C}$ parametrized by $C(s)$, vector $T(s)$ is the unit tangent vector to the curve at $C(s)$. The direction of the tangent vector depends on the orientation of the curve and points in the direction of increasing parameter values s. The curvature vector $K(s)$ at point $C(s)$ is the derivative of the unit tangent vector with respect to s, i.e. $K(s) = T'(s) = C''(s)$. The curvature $\kappa(s)$ of $\mathcal{C}$ at point $C(s)$ is the length of the corresponding curvature vector, i.e. $\kappa(s) = |C''(s)|$.

It follows from the definition that $\kappa(s)$ measures the rate of change of the tangent vector, or the deviation of curve $\mathcal{C}$ from the tangent line, in a neighbourhood of $C(s)$. In other words, for a relatively straight curve, the tangent vector remains largerly unchanged and the curvature is small, but a sharp turn/drop in the direction of the curve gives rise to large changes in the direction of the tangent vector and results in large curvature values. Thus, the behavior of the curve is characterized by its curvature. To control the global behavior of the curve in 2D, we consider the following prior based on the integrated squared curvature

$$\pi(\mathcal{C}) \propto \exp\left\{-\beta_1 \int_0^L |C'(s)|\mathrm{d}s - \beta_2 \int_0^L \kappa^2(s)\mathrm{d}s\right\}, \tag{1}$$

where L is the length of the curve, the deformation constant $\beta_1 > 0$ limits the length of the curve and $\beta_2 > 0$ determines its elasticity. The second term of the prior (1) is analogous to the roughness penalty in nonparametric smoothing and functional regression, where the latter has a form of the integrated squared second derivative with respect to the argument [Green and Silverman, 1994, Ramsay and Silverman, 2005]. The integrated squared curvature is also used as a penalty term in active contour models [Blake and Isard, 1998, Kass et al., 1988, Rueckert and Burger, 1997, Tauber et al., 2004, Terzopoulos et al., 1987]. Note that the prior in (1) is improper because it is invariant with respect to translations and rotations within the domain of interest.

For curves in $\mathbb{R}^3$, the prior (1) requires some modifications since it does not account for deviations of the curve from the plane. Recall that vectors $T(s)$ and $N(s)$ span an osculating plane at point $C(s)$. For curves in $\mathbb{R}^2$, the osculating plane, if exists, remains constant for all values of s. For nonplanar curves, the tilt of the osculating plane changes with s. These changes, however, are not captured by the curvature.

To capture the behavior of the curve in $\mathbb{R}^3$, we consider another characteristic of the curve known as torstion. The torsion $\tau(s)$ of $\mathcal{C}$ at $C(s)$ is a function that satisfies $B'(s) = -\tau(s)N(s)$. The torsion is independent of the curvature and represents an additional parameter that measures the departure of the curve from the plane in the neighborhood of $C(s)$. In fact, the shape of the curve is completely determined by its curvature κ and torsion τ. The torsion of the curve measures the rate of the change of the binormal vector. Clearly, for planar curves, the binormal remains unchanged and the torsion is zero, the reverse is also true. On the other hand, for nonplanar curves, the direction of the binormal vector is no longer constant and abrupt deviations from the plane lead to large absolute values of $\tau(s)$.

Analogously to the 2D case, we consider the integrated squared torsion and define the following prior for curves in $\mathbb{R}^3$

$$\pi(\mathcal{C}) \propto \exp\left\{-\beta_1 \int_0^L |C'(s)|\mathrm{d}s - \beta_2 \int_0^L \kappa^2(s)\mathrm{d}s - \beta_3 \int_0^L \tau^2(s)\mathrm{d}s\right\}, \tag{2}$$

where, as before, L is the length of the curve and β_i, $i = 1, 2, 3$, are non-negative constants. The first two terms are similar to the prior in (1), whilst the last term gives higher probability to planar curves. Note that this prior is also invariant with respect to translations and rotations in $\mathbb{R}^3$ and hence is imporper.

For computational reasons, we prefer an arbitrary parametrization $C(t)$ with parameter t, in which case the squared curvature and the torsion are given by

$$\kappa^2(t) = \{|C'(t)|^2|C''(t)|^2 - \langle C'(t), C''(t)\rangle^2\} \,/\, |C'(t)|^6, \tag{3}$$

$$\tau(t) = |C'(t)\; C''(t)\; C'''(t)|/|C'(t) \times C''(t)|^2, \tag{4}$$

where $0 \leq t \leq 1$, the numerator of $\tau(t)$ is the scalar triple product and $C'(t) \times C''(t)$ in the denominator stands for the cross product.

2.3 B-spline Curves

We use a class of B-spline curves to model the centerline of the data. In this section, we briefly describe B-spline curves and some of their properties and refer for more details to Piegl and Tiller [1997].

Consider a set $\mathbf{V} = \{\mathbf{V}_0, \ldots, \mathbf{V}_M\}$ of $M+1$ points in $\mathbb{R}^d$, where $\mathbf{V}_i = (V_{i1}, \ldots, V_{id})$ is the vector of coordinates of the ith point. A B-spline curve $\mathcal{C}$ of degree p, $1 \leq p \leq M$, is defined as a weighted linear combination of points $\mathbf{V}$ as

$$C(t; \mathbf{V}) = \sum_{i=0}^{M} \mathbf{V}_i \, Q_{i,p}(t), \quad 0 \leq t \leq 1. \tag{5}$$

where the weight for the ith control vertex is given by the value of the ith B-spline basis function, $Q_{i,p}(t)$, of degree p. The points $\mathbf{V}$ are referred to as the control points, or vertices, of the curve. The B-spline basis functions are computed recursively, starting from the unit step functions, as

$$\begin{aligned} Q_{i,0}(t) &= \begin{cases} 1, & \text{if } u_i \leq t < u_{i+1} \\ 0, & \text{otherwise} \end{cases} \\ Q_{i,p}(t) &= \frac{t - u_i}{u_{i+p} - u_i} \, Q_{i,p-1}(t) + \frac{u_{i+p+1} - t}{u_{i+p+1} - u_{i+1}} \, Q_{i+1,p-1}(t). \end{aligned} \tag{6}$$

The nondecreasing sequence of knots $u_0 \leq \ldots \leq u_l$ in equations (6) partitions the interval $[0,1]$. Here, we use a parametrization corresponding to a non-uniform open knot vector with $u_0 = \ldots = u_p = 0$ and $u_{l-p} = \ldots = u_l = 1$ with the remaining knots equally spaced in the interval $(0,1)$. The knot assignment ensures that $C(0; \mathbf{V}) = \mathbf{V}_0$ and $C(1; \mathbf{V}) = \mathbf{V}_M$. For brevity, in what follows, we write $C(t)$ for $C(t; \mathbf{V})$.

Definition (6) implies that the knot vector $(u_0, \ldots, u_l)$ and the degree p of the curve determine the form of the basis functions. In turn, the curve is identified by $p, \mathbf{V}$ and $(u_0, \ldots, u_l)$; note that $l = M + p + 1$. Equations (6) show that B-spline basis functions have local support, since the ith basis function is zero outside the interval $[u_i, u_{i+p+1})$. Consequently, changing the position of the ith vertex affects the curve only locally in $[u_i, u_{i+p+1})$.

As the definition shows, the complexity of the curve of a given degree depends on the number of its polynomial segments, or equivalently, on the number of control points defining the curve. Because the optimal number of control points is unknown, it is often desirable to estimate it as opposed to fitting the curve with a fixed number of control points. To accomplish this, we modify the prior choosing a penalty similar to the one described in Pievatolo and Green [1998]. The modified prior for the planar B-spline curves is now given by

$$\pi(\mathcal{C}) \propto \exp\left\{-\beta_0(M+1)^\gamma - \beta_1\int_0^L |C'(s)|\mathrm{d}s - \beta_2\int_0^L \kappa^2(s)\mathrm{d}s\right\}. \tag{7}$$

Similarly, the prior for the curves in $\mathbb{R}^3$ becomes

$$\pi(\mathcal{C}) \propto \exp\left\{-\beta_0(M+1)^\gamma - \beta_1\int_0^L |C'(s)|\mathrm{d}s - \beta_2\int_0^L \kappa^2(s)\mathrm{d}s - \beta_3\int_0^L \tau^2(s)\mathrm{d}s\right\}. \tag{8}$$

In (7) and (8), β_0 and γ are positive constants, so that the magnitude of the first term increases in M and the priors penalize curves with a large number of control points.

2.4 Posterior Distribution

Combining the likelihood and the prior, we obtain the posterior distribution of the skeleton curve given the data Y, which for curves in 3D, becomes

$$\begin{aligned}\pi(\mathcal{C}|\,Y) &\propto \left(\frac{1}{\sigma^2}\right)^{n/2}\exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^n d_i^2\right\}\times \\ &\times \exp\left\{-\beta_0(M+1)^\gamma - \beta_1\int_0^L |C'(s)|\mathrm{d}s - \beta_2\int_0^L \kappa^2(s)\mathrm{d}s - \beta_3\int_0^L \tau^2(s)\mathrm{d}s\right\}.\end{aligned} \tag{9}$$

For 2D curves, the torsion term vanishes.

3 MCMC Sampler

Because the posterior density of $\mathcal{C}$ is not available in closed-form, we use Monte Carlo methods to generate samples from $\pi(\mathcal{C}|\,\sigma^2, Y)$ in (9). Denote by $(\mathbf{V}^{(k)}, \mathcal{C}^{(k)}),\ k = 1, \ldots, K$, the successive sweeps of the Markov chain Monte Carlo sampler initialized at $(\mathbf{V}^{(0)}, \mathcal{C}^{(0)})$.

To update the curve we construct a sampler similar to the one used for updating polygons in stochastic boundary reconstruction problem [Pievatolo and Green, 1998]. The idea is to modify the curve by updating the position of its vertices. The latter are updated using three different move types: moving, adding or deleting a vertex. The vertex move is a fixed-dimension random walk Metropolis algorithm. In turn, vertex addition and deletion change the dimensionality of the vertex state space, so we resort to a reversible jump MCMC method [Green, 1995].

1. *Move a vertex.* With probability m_k, replace randomly chosen vertex $\mathbf{V}_i^{(k)}$ from the current set of control vertices $\mathbf{V}^{(k)}$ with $\mathbf{V}_i^* = \mathbf{V}_i^{(k)} + Z_i$, where Z_i has a d-variate centered normal distribution with covariance Σ_m. For the proposal $\mathbf{V}^*$, identify the corresponding curve $\mathcal{C}^*$ and compute an approximation of the Metropolis ratio $r_m(\mathcal{C}, \mathcal{C}^*|\sigma^{(k)}, Y) = \pi(\mathcal{C}^*|\sigma^{(k)}, I^{\mathrm{D}})/\pi(\mathcal{C}^{(k)}|\sigma^{(k)}, Y)$. Set $\mathbf{V}^{(k+1)} = \mathbf{V}^*$ and $\mathcal{C}^{(k+1)} = \mathcal{C}^*$ with probability $\min(1, r_m)$ and maintain $\mathbf{V}^{(k+1)} = \mathbf{V}^{(k)}$, $\mathcal{C}^{(k+1)} = \mathcal{C}^{(k)}$, otherwise.

2. *Add a vertex.* With probability a_k, for randomly chosen i, add a new vertex $\mathbf{V}_i^* = 0.5(\mathbf{V}_i^{(k)} + \mathbf{V}_{i-1}^{(k)}) + Z, i = 0, \ldots, M+1$, where $\mathbf{V}_{-1} \equiv \mathbf{V}_0, \mathbf{V}_{M+1} \equiv \mathbf{V}_M$ and Z is a centered d-variate normal variable with covariance Σ_a. For a new set $\mathbf{V}^* = \{\mathbf{V}_0^{(k)}, \ldots, \mathbf{V}_{i-1}^{(k)}, \mathbf{V}_i^*, \mathbf{V}_i^{(k)}, \ldots, \mathbf{V}_M^{(k)}\}$, compute the corresponding curve $\mathcal{C}^*$ and calculate the acceptance ratio

$$r_a(\mathcal{C}, \mathcal{C}^*|\sigma^{(k)}, Y) = \min\left\{1, \frac{\pi(\mathcal{C}^*)(M+1)d_k}{\pi(\mathcal{C}^k)\phi_d(z)(M+2)a_k}\left|\frac{\partial \mathbf{V}^*}{\partial \mathbf{V} \partial Z}\right|\right\}.$$

Set $\mathbf{V}^{(k+1)} = \mathbf{V}^*$, $\mathcal{C}^{(k+1)} = \mathcal{C}^*$ with probability $\min(1, r_a)$ and maintain $\mathbf{V}^{(k+1)} = \mathbf{V}^{(k)}$, $\mathcal{C}^{(k+1)} = \mathcal{C}^{(k)}$, otherwise.

3. *Delete a vertex.* For $n > 3$, with probability d_k, delete a randomly chosen vertex $\mathbf{V}_i^{(k)}$, compute curve $\mathcal{C}^*$, and calculate the acceptance ratio

$$r_d(\mathcal{C}, \mathcal{C}^*|\sigma^{(k)}, Y) = \min\left\{1, \frac{\pi(\mathcal{C}^k)\phi_d(z)(M+2)a_k}{\pi(\mathcal{C}^*)(M+1)d_k}\left|\frac{\partial \mathbf{V}^*}{\partial \mathbf{V} \partial Z}\right|^{-1}\right\}.$$

Set $\mathbf{V}^{(k+1)} = \mathbf{V}^*$, $\mathcal{C}^{(k+1)} = \mathcal{C}^*$ with probability $\min(1, r_d)$ and maintain $\mathbf{V}^{(k+1)} = \mathbf{V}^{(k)}$, $\mathcal{C}^{(k+1)} = \mathcal{C}^{(k)}$, otherwise.

The acceptance ratios r_m, r_a, r_d are computed using a fine partition of $[0, 1]$ in (7) and (8) and a trapezoidal integration in (3) or (4).

4 Results and Discussion

4.1 2D Example

Here we consider a function with oscillations of exponentially varying maginitude and a power trend given by $f(x) = 5e^{-.07t}\cos(.7t) + \sqrt{t},\ x \in [2, 30]$. Data points were generated according to $Y_i = f(x_i) + \varepsilon_i,\ i =$

$1, \ldots, 100$, where x_i are chosen uniformly at random on $[2, 30]$ and ε_i are independent samples from the centered normal distribution with variance $\sigma = 0.05$. Figure 1, left, shows the simulated data with the true centerline overlayed in red.

The skleleton curve was modeled as a cubic B-spline curve, i.e. $p = 3$. To initialize the sampler, six vertices were randomly positioned about the horizontal midline of the data. The prior parameters were set to $\gamma = 3, \beta_0 = 0.1, \beta_1 = 0.1, \beta_2 = 0.001, \Sigma_a = \Sigma_m = 0.5I_2, m_k = a_k = d_k = 1/3$. The first 5K realizations of the Markov chain are discarded as burn-in. The following 25K sweeps of the chain were acquired. The algorithm was implemented in Matlab 7.6.0 (R2008a, The MathWorks) using the C++ NURBS library (ARIA). The computations were performed on a MacBookPro with 2.2.GHz Inter Core 2 Duo processor and 2GB of RAM with 30K sweeps completed in approximately 1.4 minutes. To monitor the convergence, we compute the number of points within distance 3σ from a sample curve for every sweep (Figure 2, right).

Figure 1, right, shows a realization from the posterior with the true curve overlayed in red. Figure 2 (left) shows the number of control points for every sweep. The estimated number of vertices needed to model the curve is 10.

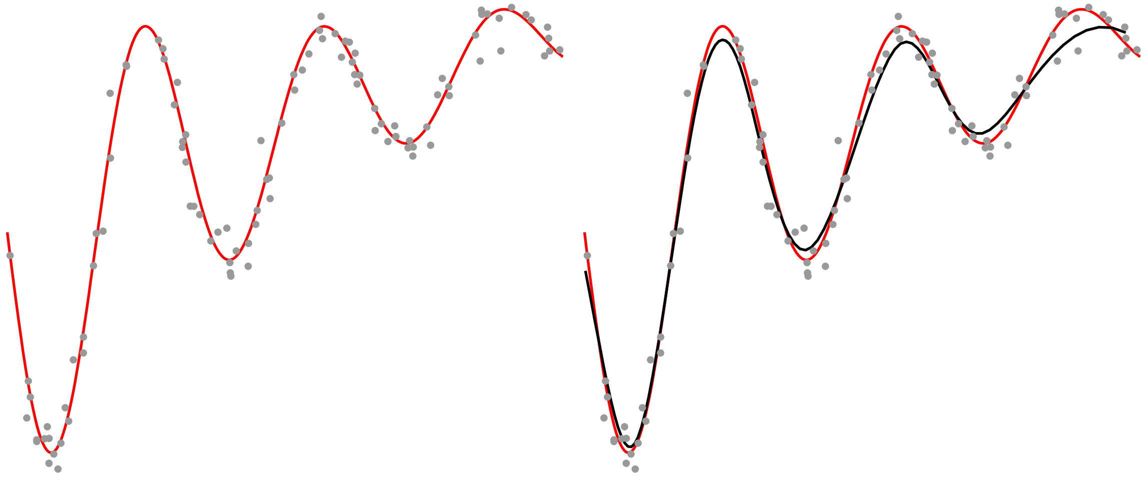

Figure 1: (Left) Simulated data and (right) the fitted curve (black). The red curve is the true curve.

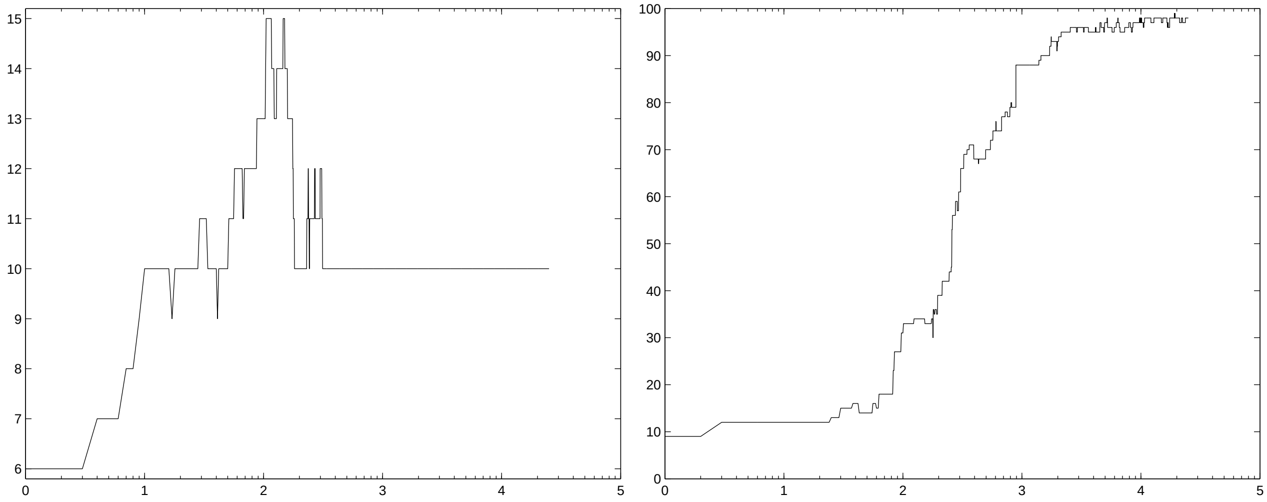

Figure 2: The number of control points (left) and the number of points within 3σ distance from the sample curve (right) for every sweep, for the 2D example. Both statistics are plotted on the $\log_{10}$ scale.

4.2 3D Example

Here, we use the method to estimate the centerline of three dimensional data. We take the true skeleton to be a helix curve $C(t) = (\cos t, \sin t, t)$. We simulate $M = 100$ data points as $Y_i = C(t_i) + N(t_i)\varepsilon_i$, where t_i are chosen uniformly at random on $[0, 1]$, vector $N(t_i) = (-\cos t, -\sin t, 0)$ is the unit normal to the curve at $C(t_i)$, and ε_i are random realizations from the centered normal distribution with variance $\sigma = 0.5$. Figure 3 shows the simulated data in $\mathbb{R}^3$ at three different view angles relative to the true skeleton curve (red).

The cubic curve was initialized by positioning six control vertices to be approximately evenly spaced in the x direction and choosing y and z coordinates from the standard normal distribution. The constants for the 3D prior (8) were set to $\gamma = 2, \beta_0 = 0.05, \beta_1 = \beta_2 = \beta_3 = 0.02, \Sigma_a = \Sigma_m = I_3, m_k = a_k = d_k = 1/3$. The first 5K realizations of the Markov chain were discarded as burn-in and the chain was run for 25K more sweeps. The run time for the helix example was about 1.9 min. The acceptance rate was approximately 4%. The convergence was assessed by monitoring the number of points within 3σ distance from the curve.

Figure 4 shows a good agreement between the true curve (red) and the sample from the posterior (black). Figure 5 (left) shows the number of control points for every sweep. The number of vertices needed to model the curve is approximately ten. The convergence is achieved by 5K iterations as shown in Figure 5 (right).

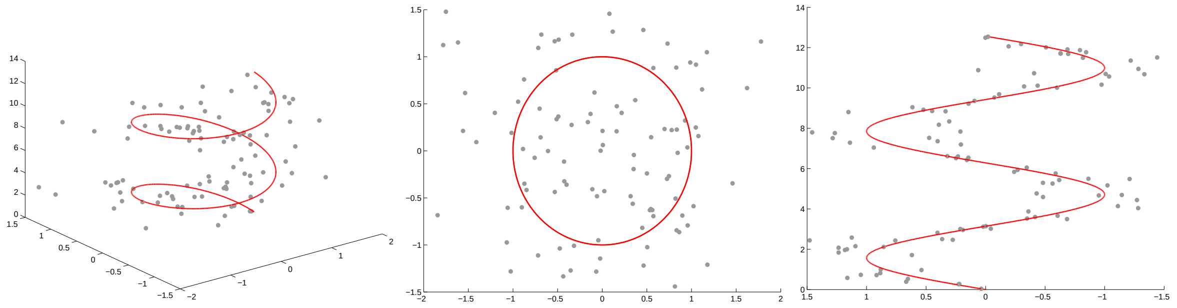

Figure 3: Simulated data for 3D example at three different view angles; the true skeleton curve $C(t) = (\cos t, \sin t, t)$ is shown in red.

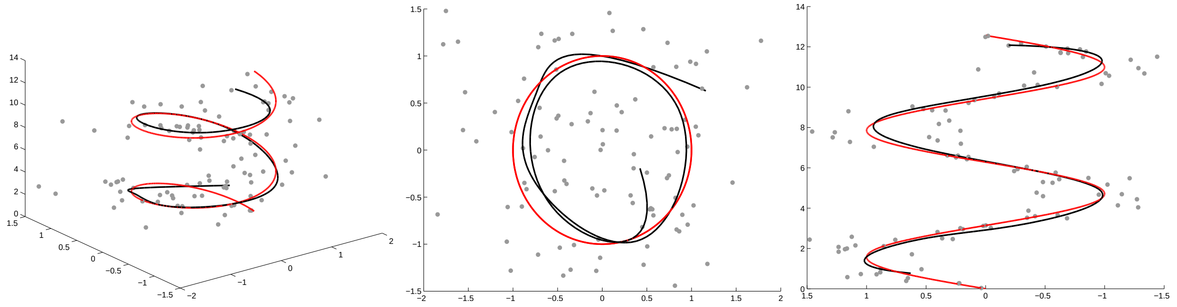

Figure 4: A comparison between a sample from the posterior (black) after the chain convergence and the true helix skeleton curve (red) at three different view angles.

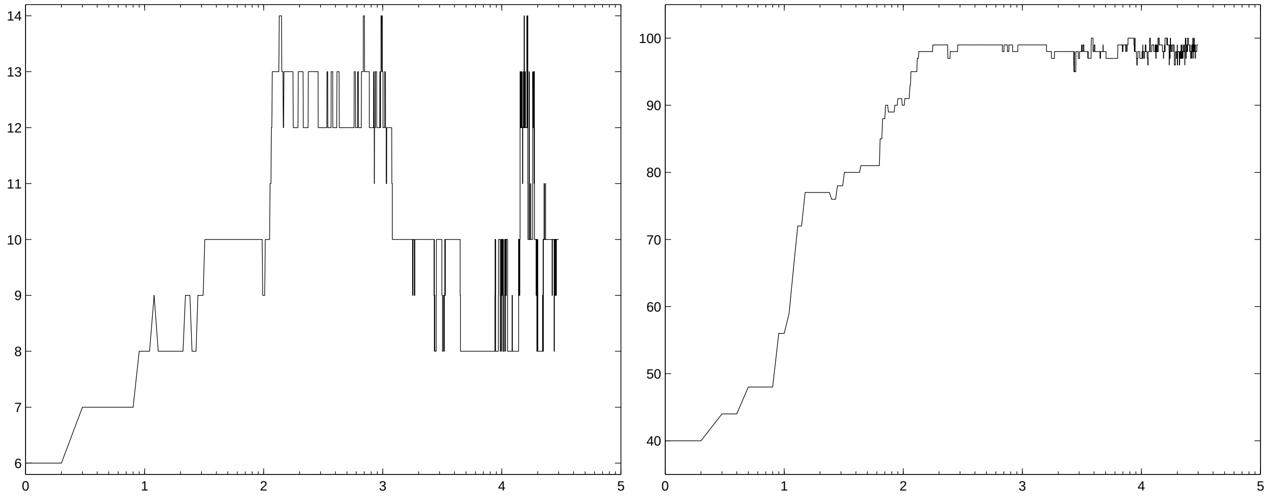

Figure 5: The number of control points (left) and the number of points within 3σ distance from the sample curve (right) for every sweep, for the 3D example. Both statistics are plotted on the $\log_{10}$ scale.

4.3 Application to SPECT Imaging

In this example, we use SPECT data acquired on a single subject who was administered a dose of radiolabeled lubricant through the intracolonic injection. The subject then underwent an experimental procedure designed to mimic the distribution of the lubricant in the colon. SPECT images were acquired using a dual-head VG SPECT-CT imaging system (GE Medical Systems, Waukesha, WI). X-ray computed tomography images were also collected for anatomical reference. The study was approved by the Johns Hopkins Institutional Review Board. The reconstructed SPECT data represent a three-dimensional array of 128^3 voxels. The acquired images were thresholded to remove the background noise. The final data set consists of 1000 voxels sampled uniformly at random after the thresholding. More details on the data acquisition and image preprocessing can be found in Caffo et al. [2008].

To estimate the skeleton of the SPECT data, we initialize a cubic curve with six control vertices spaced approximately evenly in the x direction and with y and z coordinates given by independent random samples from the normal distribution with mean 60 and variance 2. The constants for the prior (8) were set to $\gamma = 2, \beta_0 = 0.1, \beta_1 = \beta_2 = \beta_3 = 0.004, \Sigma_m = \Sigma_a = 20I_2, m_k = a_k = d_k = 1/3, \sigma = 10$. Note that the variance in the likelihood model was determined from the data as a pooled variance estimator based on 100 nearest neighbors of 200 randomly chosen data points. The first 5K realizations of the Markov chain were discarded as burn-in and the chain was run for 25K sweeps. The run time was about 41.9 min.

Figure 6 shows the fitted curve overlayed on the original data at three view angles. The curve reflects basic features of the colon anatomy and is in agreement with the reconstruction results reported in Caffo et al. [2008].

4.4 Concluding Remarks

In this paper, we describe a Bayesian method based on B-spline curves to estimate the skeleton from the noisy data. The recursive definition of the basis functions ensures efficient computation of B-spline curves, while the piecewise polynomial structure provides the desired flexibility to model complex contours. In addition, the affine equivariance property of B-spline curves allows to modify the position of the curve by manipulating only a small number of its control vertices. The method does not require a user-specified end

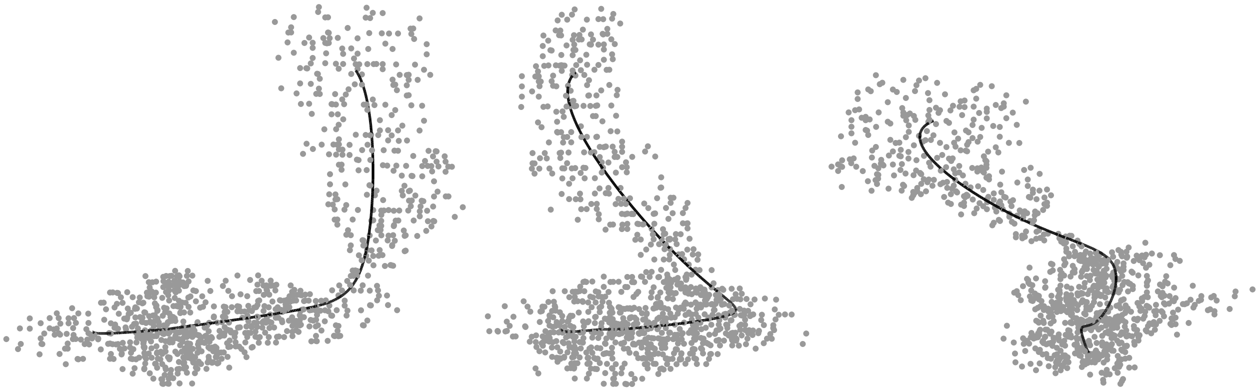

Figure 6: The fitted curve (black) at three view angles.

points, which is advantageous, especially when dealing with highdimensional data.

The resulting estimates naturally depend on the parameters of the prior distribution. In particular, large values of $\beta_i, i = 1, 2, 3$ result in a rigid skeleton curve with a small number of control points which captures the overall trend, but does not reflect important features. In turn, small values of the parameters lead to erratic centerlines. The parameter value also depends on the variance of the proposal densities, i.e. larger variance will require smaller parameters and vice versa. Similarly, choosing values for β_0 and γ leads to stiff curves with small number of vertices. The prior parameters in our examples were determined experimentally by monitoring the behavior of the curve.

With respect to the reversible jump MCMC, from our experiments we found that we gain no advantages when using unequal variances in the proposals. Also, the parameter values $\beta_i, i = 1, 2, 3$ depend on the variance of the proposal densities, i.e. larger variance will require smaller parameters and vice versa. The covariance matrix Σ_a in RJMCMC was determined experimentally by monitoring the behavior of the curve.

Given that the curve is controlled through its vertices, large variance values tend to substantially affect the shape of the curve. This is advantageous at the beginning of the run, but once the shape of the curve is learned, large jumps make it harder to expose local details of the curve. Our simulations indicate that the shape of the curve is mainly learned through repositioning of the vertices. In the 2D example, the vertices are added and deleted mainly at the beginning of the cycle, while in 3D the number of vertices fluctuates throughout the run; see Figures 2 and 5, left panels. This could be explained by larger variability of the 3D data, more complicated helix curve, and an extra dimension that gives more leeway for the vertix repositioning.

In this paper, we considered estimating the number of control points and the skeleton curve simultaneously. In practice, it might be advantageous to use a curve with a fixed number of vertices, for example when dealing with relatively simple contours. The sampler in this case contains only a move step and the updates take less time. A similar approach was used to reconstruct the boundary of an object in a noisy image [Stanberry and Besag, in preparation] .

We should also note that the execution time of 41.9 min is longer for the SPECT data as compared to the simulated examples. This is due to larger number of points in the data set which leads to increased computation time for the distances in the likelihood function. The coding, however, was not optimized and most certainly the adjustments can be made to speed it up.

Another important issue to mention is the acceptance rate which is considerably lower than commonly accepted 20-50%. This, however, is not surprising given the number of coordinates of the control points on which the sampler is based.

In our examples, we focused on modeling open skeleton curves. As an example we consider reconstructing a closed curve by simulating 100 data points uniformly scattered about the unit circle, i.e. $Y = (R\cos\Theta, R\sin\Theta), R \sim U[0.8, 1.2], \Theta \sim U[-\pi, \pi]$. The cubic curve was initialized as a straight line with 6 control points positioned horizontally about the midline of the data. The reconstruction parameters were $\gamma = 2, \beta_0 = 1, \beta_1 = 0.1, \beta_2 = 0.05, \sigma^2 = \text{var}(R), \sigma_m^2 = .05, \sigma_a^2 = .1, m_k = a_k = b_k = 1/3$ for all k. Figure 7 shows the MCMC samples from three independent runs of the Markov chain. The reconstructed curve is not closed, because of the prior that gives higher probability to shorter and "less curvy" curves. Not that in case of the closed curve reconstruction, the original principal curve algorithm exhibits similar behavior Hastie and Stuetzle [1989].

This example shows that in its present form, the method prohibits the transition of the curve from open to closed. Certainly, if the underlying curve is known to be closed, it ought to be modeled as such, e.g. closed principal curves were used to identify the outline of ice floes in Banfield and Raftery [1992]. In particular, the closed cubic B-spline curves are constructed by simply wrapping the first three and the last three control points. However, in practice the exact geometry of the curve is unknown and it would be of interest to develop a method that would allow for changes in the “topology” of the curves.

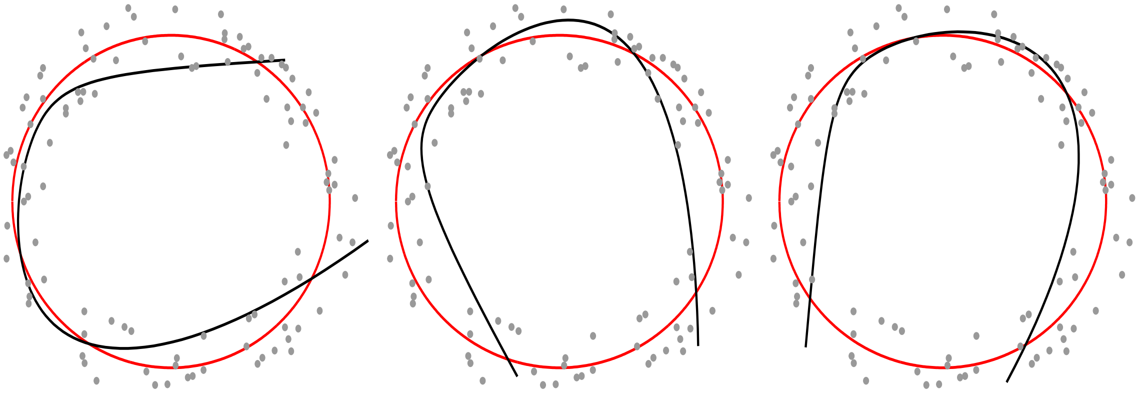

Figure 7: Curve samples from the three independent runs of the Markov Chain.

5 Acknowledgments

The authors thank Brian Caffo, Ciprian Crainiceanu and Craig Hendrix at Johns Hopkins University for sharing the SPECT data used in this paper.

References

J. Banfield and A. Raftery. Ice flow identification in satellite images using mathematical morphology and clustering about principal curves. *Journal of American Statistical Association*, 87:7–16, 1992.

A. Blake and M. Isard. *Active Contours: The Application of Techniques from Graphics, Vision, Control Theory and Statistics to Visual Tracking of Shapes in Motion.* Springer-Verlag New York, Inc., Secaucus, NJ, USA, 1998.

S. Bouix, J. Pruessner, L. Collins, and K. Siddiqi. Hippocampal shape analysis using medial surfaces. *NeuroImage*, 25:1077–1089, 2005.

B. Caffo, C. Crainiceanu, L. Deng, and C. W. Hendrix. A case study of pharmacological colon imaging using principal curves in single photon emission computed tomography. *J Am Stat Assoc*, 103(484):1470–1480, 2008.

P. Chaudhuri, R. Khandekar, D. Sethi, and P. Kalra. An efficient central path algorithm for virtual navigation. *Computer Graphics International Conference*, pages 188–195, 2004.

L. Cohen and R. Kimmel. Fast marching the global minimum of active contours. In *Proc. IEEE International Conference on Image Processing (ICIP'96)*, pages 473–476, 1996.

M. Delfour and J.-P. Zolésio. *Shapes and Geometries*, volume 4 of *Advances in Design and Control*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2001.

T. Deschamps and L. Cohen. Fast extraction of minimal paths in 3D images and applications to virtual endoscopy. *Medical Image Analysis*, 5(4), Dec. 2001.

P. Green. Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. *Biometrika*, 82(4):711–732, 1995.

P. J. Green and B. W. Silverman. *Nonparametric regression and generalized linear models*, volume 58 of *Monographs on Statistics and Applied Probability*. Chapman & Hall, London, 1994. ISBN 0-412-30040-0. A roughness penalty approach.

T. Hastie. *Principal curves and surfaces*. PhD thesis, Stanford University, Department of Statistics, 1984.

T. Hastie and W. Stuetzle. Principal curves. *J Am Stat Assoc*, 84:502–516, 1989.

M. Kass, A. P. Witkin, and D. Terzopoulos. Snakes: Active contour models. *International Journal of Computer Vision*, 1(4):321–331, 1988.

E. Kreyszig. *Introduction to Differential Geometry and Riemannian Geometry*. University of Toronto Press, Toronto, 1968.

W. Ma, F. Wu, and M. Ouhyoung. Skeleton extraction of 3d objects with radial basis functions. In *Proceedings of Shape Modeling International 2003*, pages 207–215, 2003.

L. Piegl and W. Tiller. *The NURBS Book*. Springer-Verlag New York, Inc., New York, NY, USA, 2nd edition, 1997.

A. Pievatolo and P. Green. Boundary detection through dynamic polygons. *Journal of the Royal Statistical Society, Series B, Statistical Methodology*, 60(3):609–626, 1998.

J. O. Ramsay and B. W. Silverman. *Functional data analysis*. Springer Series in Statistics. Springer, New York, second edition, 2005. ISBN 978-0387-40080-8; 0-387-40080-X.

K. Reinhard and M. Niranjan. Subspace models for speech transitions using principal curves. In *Proceedings of Institute of Acoustics*, pages 53–60, 1998.

D. Rueckert and P. Burger. Geometrically deformable templates for shape-based segmentation and tracking in cardiac MR images. In *EMMCVPR '97: Proceedings of the First International Workshop on Energy Minimization Methods in Computer Vision and Pattern Recognition*, pages 83–98, London, UK, 1997. Springer-Verlag.

R. Sadleir and P. Whelan. Colon centreline calculation for CT colonography using optimised 3D topological thinning. *3D Data Processing Visualization and Transmission, International Symposium on*, 0:800, 2002.

L. I. Stanberry and J. Besag. Boundary reconstruction in binary images using splines. in preparation.

D. Stanford and A. E. Raftery. Principal curve clustering with noise. Technical report, University of Washington, 1997.

M. Styner, G. Gerig, J. Lieberman, D. Jones, and D. Weinberg. 2003. *Medical Image Analysis*, 7:207–220, 2003.

C. Tauber, H. Batatia, G. Morin, and A. Ayache. Robust B-spline snakes for ultrasound image segmentation. *Computers in Cardiology*, pages 325–328, 2004.

D. Terzopoulos, A. Witkin, and M. Kass. Snakes: Active contour models. In *IEEE International Conference on Computer Vision*, pages 259–268, 1987.