Using the Tractometer to assess performance of a new statistical tractography technique

Lyudmila Sakhanenko*

Department of Statistics and Probability, Michigan State University, East Lansing, MI 48824, USA

We evaluate a recent statistical tractography technique for analyzing the diffusion tensor imaging data using new criteria and the template from the Tractometer. This novel method simultaneously achieves excellent percentages on 3 out of 7 criteria. Compared to deterministic tractography methods, it is superior to all of them. Compared to probabilistic tractography methods, it has similar results for so-called ABC metric and better performance for all the other metrics. On top of this, it is faster and it additionally provides more geometrical information through confidence ellipsoids rather than through a bundle of repeated tracts. Journal of Nature and Science, 1(7):e130, 2015

integral curve | tractography | diffusion tensor imaging

How to cite this article? Here is an example:


Conflict of interest: No conflicts declared.

* Corresponding Author. Department of Statistics and Probability, Michigan State University, East Lansing, MI 48824-1027, USA. Email: luda@stt.msu.edu

This Research was partially supported by NSF grant DMS-1208238

© 2015 by the Journal of Nature and Science (JNSCI).
Using the Tractometer to assess performance of a new statistical tractography technique

Lyudmila Sakhanenko *
Michigan State University
June 10, 2015

Abstract

We evaluate a recent statistical tractography technique for analyzing the diffusion tensor imaging data using new criteria and the template from the Tractometer. This novel method simultaneously achieves excellent percentages on 3 out of 7 criteria. Compared to deterministic tractography methods, it is superior to all of them. Compared to probabilistic tractography methods, it has similar results for so-called ABC metric and better performance for all the other metrics. On top of this, it is faster and it additionally provides more geometrical information through confidence ellipsoids rather than through a bundle of repeated tracts.

Key words: integral curve, tractography, diffusion tensor imaging

1 Introduction

Recently Carmichael and Sakhanenko (2014) proposed a new statistical technique (which we will call CS method below for brevity) for estimating integral curves driven by the dominant eigenvector field of a tensor field, whose components are slopes in a system of regression equations. They use kernel smoothing to estimate tensor components, then take the field of the main eigenvectors of the estimated tensor field, and trace the integral curves along this vector field in small steps. In essence, the integral curve estimator is a plug-in estimator based on kernel smoothing of the tensor field. They showed theoretically

*Department of Statistics and Probability, East Lansing, MI 48824-1027, luda@stt.msu.edu. Research partially supported by NSF grant DMS-1208238.
that these estimators are asymptotically normal under certain regularity conditions on the underlying tensor field. The asymptotical variance of the estimator and its bias can be estimated together with the estimator in small steps sequentially. Geometrically the results are presented by an estimated curve together with confidence ellipsoids along it, see an example in Figure 1 (a). One can easily assess how the uncertainty in image components has propagated to the trajectory, and how it changes as the curve goes through regions of higher uncertainty levels such as near branching points. This tractography method is motivated by and used for estimation of neural fibers in a brain from a dataset of observed noisy diffusion tensor imaging (DTI) components; see the model in the next section.

There are many different tractography methods for DTI, a popular en vivo brain imaging technique; see Assemal et.al. (2011) for a complete review. However, the comparison of them is tricky. First, there is no agreement on a golden standard in simulated studies or in real scans. Secondly, the methods are designed to achieve different goals. For example, some model the tensor field or the vector field while tracing of the fibers is secondary. Recently Côté et.al. (2013) attempted to alleviate this problem and proposed a novel evaluation system for tractography methods. It is based on an artificial template which we reproduce from their website in Figure 1 (b). It mimics the typical behavior of neural fibers. Then they select a set of metrics used for quantitative assessment of the methods. These metrics are average bundle coverage (ABC), valid connections (VC), invalid connections (IC), no connections (NC), valid bundles (VB), and invalid bundles (IB). They are motivated by needs of human brain imaging practitioners. The metrics target the global connectivity. It is the first systematic attempt to go from a qualitative comparison to a quantitative one.

They also collect various approaches into a database, which are then rated based on the metrics. A recent summarized report can be found in Daducci et.al. (2013). The goal of this paper is to compare the method of CS against others using Tractometer and also to show how the selected metrics might not do justice to statistical tractography methods.

2 Model and algorithm

At a fixed location $x$ in a 3D region $G$ we assume that the observations $Y(x) \in \mathbb{R}^N, N \geq 6$, arise from a mixture of two underlying tensors $M^{(1)}(x)$ and $M^{(2)}(x)$, which have the same simple maximal eigenvalue but different leading eigenvectors $v(M^{(1)}(x))$ and $v(M^{(2)}(x))$:

$$Y(x) = B(\pi M^{(1)}(x) + (1 - \pi) M^{(2)}(x)) + \Sigma^{1/2}(x) \Xi_x. \quad (1)$$

Here the symmetric tensors $M^{(i)}(x), i = 1, 2$, are represented by 6D vectors, $B$ is a fixed known $N \times 6$ design matrix. The random noise 6D vector $\Xi$ has mean 0 and identity co-
variance matrix. \( \Sigma(x) \) is a positive definite \( 6 \times 6 \) unknown matrix. The mixing coefficient \( \pi \) determines the relative contributions of \( M^{(1)}(x) \) and \( M^{(2)}(x) \) to \( Y(x) \). A variety of standard clustering techniques can be applied to estimate \( M^{(1)}(x) \) and \( M^{(2)}(x) \) as well as \( \pi \). In our simulation study we use the standard clustering procedures based on the mixture of two Gaussian distributions.

The model (1) is motivated by low angular resolution DTI. It is based on measurements of water diffusion on a grid of points. Locally the relative amount of water diffusion along a spatial direction \( b \in \mathbb{R}^3, \|b\| = 1 \), at a voxel \( x \), \( S(x, b) \), is estimated as follows:

\[
\log \left( \frac{S(x, b)}{S_0(x)} \right) = -cb^*D(x)b + \sigma(x, b)\xi_b,
\]

where \( S_0(x) \) is the amount of water diffusion without gradient application; \( \sigma(x, b) > 0; \xi_b \) describes noise; the constant \( c \) depends only on the proton gyromagnetic ratio, the gradient pulse sequence shape, duration and other timing parameters of the imaging procedure; \( D \) is a \( 3 \times 3 \) symmetric positive definite matrix. On practice a set of \( N \) directions \( b \) is chosen and the corresponding log-losses are measured. They are stacked into a vector-column \( Y \) in the model (1). Meanwhile the upper triangular part of matrix \( D \) is stacked into a 6D vector-column \( M \). We remark that the mixture of tensors is used to model branching phenomenon.

Now we observe a set of \( Y(x_j), j = 1, \ldots, n \), for a collection of sites \( x_j, j = 1, \ldots, n \), in \( G \), which usually forms a 3D grid. CS method first takes the least squares estimators \( \tilde{M}_n(x_{ij}) \) for \( \pi M^{(1)}(x_{ij}) + (1 - \pi)M^{(2)}(x_{ij}) \) at each grid point \( x_{ij}, j = 1, \ldots, n \), and smooths the tensor field using kernel regression estimation to obtain the estimated tensor field \( \tilde{M}_n(\cdot) \) everywhere in the region \( G \). Also at each \( x_{ij} \) the clustering is done if the Gaussian mixture test finds more than 1 cluster. Secondly, taking into account the clusters, the tracing of the estimated integral curve \( \tilde{X}_n(\cdot) \) is done in small steps \( \delta > 0 \) in the direction of the dominant eigenvector field \( v(M^{(i)}(x)), x \in G, i = 1, 2 \). Approximation to the asymptotical covariance matrix \( C(\cdot) \) and the bias \( \mu(\cdot) \) of the estimated curve are also calculated on each step from the previous step values because they do not require the knowledge of the whole trajectory \( x(t), t \in [0, T] \).

For completeness we list the steps of the algorithm for one cluster.

1. Initialize \( q = 0, t_q = 0, \tilde{X}_n(t_0) = x_0, \mu(t_0) = 0, C(t_0) = 0 \).

2. Track the estimated integral curve using Euler’s method. Given a fixed \( \delta > 0 \) and
small steps $t_q = q\delta, q = 0, 1, 2, \ldots$

$$\tilde{M}_n(x_j) = (B^*B)^{-1}B^*Y(x_j), j = 1, \ldots, n, \quad \tilde{M}_n(x) = \frac{1}{nh_n^3} \sum_{j=1}^{n} K \left( \frac{x - x_j}{h_n} \right) \tilde{M}(x_j),$$

$$\tilde{X}_n(t_q) = \tilde{X}_n(t_{q-1}) + \delta v(\tilde{M}_n(\tilde{X}_n(t_{q-1}))).$$

Here $B^*$ stands for the transpose of $B$; $v(A)$ denotes the main eigenvector of matrix $A$; $K$ is a symmetric about 0, probability density with a finite second moment, for example the standard Gaussian density in 3D; the averaging window (bandwidth) size is $h_n = (\beta/n)^{1/6}$ with some finite pre-selected constant $\beta > 0$. This tuning parameter can be chosen as the minimizer of the mean integrated squared error (MISE) between the estimated and the true integral curves as proposed in Carmichael and Sakhanenko (2014). Then $\beta = 0.25(d - 1) \int_0^T \text{tr} C(t, t) dt \left[ \int_0^T \mu_1(t)^\ast \mu_1(t)dt \right]^{-1}$, where one uses $\beta = 1$ in $\mu_1$. One would start with an initial $\beta$, estimate $\mu_1$ and $C$ accordingly, re-evaluate $\beta$ based on these $\mu_1$ and $C$, and repeat.

3. Approximate the gradient of the tensor field by $\nabla \tilde{M}_n(x)$, the tensor of second order derivatives of the tensor field by $\nabla^2 \tilde{M}_n(x)$ and covariance matrix of the noise by $\tilde{\Sigma}_n(x)$ using the following formulas: for $j = 1, \ldots, n$

$$\tilde{\Sigma}_n(x_j) = \frac{1}{m-1} \sum_{i=1}^{m} (Y(L_i) - B\tilde{M}(L_i) - \tilde{Y} + BM)(Y(L_i) - B\tilde{M}(L_i) - \tilde{Y} + BM)^\ast,$$

where $L_i, i = 1, \ldots, m$, are the measurement locations in a local neighborhood of $x_j$ and the averages $\tilde{Y}$ and $\tilde{M}$ are taken over the $m$ values $Y(L_i)$ and $\tilde{M}(L_i), i = 1, \ldots, m$, respectively. For instance, for a regular 3D grid $x_j, j = 1, \ldots, n$, we could set the $L_i$ to be the points in the cube of 26 immediate neighbors surrounding $x_j$ in all 3 spatial directions. Then for $x \in G$

$$\tilde{\Sigma}_n(x) = \frac{1}{nh_n^3} \sum_{j=1}^{n} K \left( \frac{x - x_j}{h_n} \right) \tilde{\Sigma}_n(x_j),$$

$$\nabla \tilde{M}_n(x) = \frac{1}{nh_n^4} \sum_{j=1}^{n} \nabla K \left( \frac{x - x_j}{h_n} \right) \tilde{M}(x_j), \quad \tilde{h}_n = \left( \frac{\log n}{n} \right)^{1/4},$$

$$\nabla^2 \tilde{M}_n(x) = \frac{1}{nh_n^5} \sum_{j=1}^{n} \nabla^2 K \left( \frac{x - x_j}{h_n} \right) \tilde{M}(x_j), \quad \tilde{h}_n = \left( \frac{\log n}{n} \right)^{1/5}.$$

4. And, approximate the gradient of the eigenvector field $\nabla v(M(x))$ by $\nabla v(\tilde{M}_n(x))$ with
components: for $1 \leq p, k, l \leq 3$

$$\frac{\partial v_p(\hat{M}_n(x))}{\partial M_{kl}} = (1 - 0.5\delta_{kl}) \left[ (\lambda(\hat{M}_n(x))I_d - \hat{M}_n(x))_{pk}^+ v_k(\hat{M}_n(x)) + \frac{\partial}{\partial M_{kl}}(\lambda(\hat{M}_n(x))) v_k(\hat{M}_n(x)) \right],$$

where $A^+$ stands for the Moore-Penrose inverse of a matrix $A$ and $\lambda(A)$ denotes the eigenvalue of $A$. Here $I$ is the $3 \times 3$ identity matrix and $\delta_{kl} = 0$ for $k \neq l$ and 1 for $k = l$.

5. Approximate the asymptotical bias $\mu(t_q), q \geq 1$, as

$$\hat{\mu}(t_q) = \hat{\mu}(t_{q-1}) + 0.5\sqrt{\beta}v(\hat{M}_n(\hat{X}_n(t_{q-1})))\nabla \hat{M}_n(\hat{X}_n(t_{q-1}))\hat{\mu}(t_{q-1})$$

$$+ 0.5\delta\nabla \hat{v}(\hat{M}_n(\hat{X}_n(t_{q-1}))) \int_{\mathbb{R}^3} K(z)(\nabla^2 \hat{M}_n(\hat{X}_n(t_{q-1}))z, z)dz.$$

6. Approximate the asymptotical covariance $C(t_q), q \geq 1$, as

$$\hat{C}(t_q) = \hat{C}(t_{q-1}) + \delta\nabla \hat{v}(\hat{M}_n(\hat{X}_n(t_{q-1})))\nabla \hat{M}_n(\hat{X}_n(t_{q-1}))\hat{C}(t_{q-1})$$

$$+ \delta\psi(t_{q-1})\nabla \hat{M}_n(\hat{X}_n(t_{q-1}))^\ast \nabla \hat{v}(\hat{M}_n(\hat{X}_n(t_{q-1})))^\ast$$

$$\times [\hat{M}_n(\hat{X}_n(t_{q-1}))\hat{M}_n^\ast(\hat{X}_n(t_{q-1}))+\hat{\Sigma}_n(\hat{X}_n(t_{q-1}))]$$

$$\times \nabla \hat{v}(\hat{M}_n(\hat{X}_n(t_{q-1})))^\ast,$$

where $\psi(v) = \int_{\mathbb{R}} \int_{\mathbb{R}^3} K(z)K(z + \tau v)dzd\tau$.

7. Approximate the 100$(1 - \alpha)%$ confidence ellipsoid for $x(t_q), q \geq 1$, as

$$P\left\{ \left| \hat{C}(t_q)^{-1/2} \left( \sqrt{n\hat{h}_n^2(X_n(t_q) - x(t_q))} - \hat{\mu}(t_q) \right) \right| \leq R_{\alpha} \right\} \approx 1 - \alpha,$$

where $P(|Z| \leq R_{\alpha}) = 1 - \alpha$ for a standard normal 3D vector $Z$.

8. Repeat steps 2–6 until $t_q$ reaches $T$.

3 Simulation results

We first used the estimation method without any testing for clusters, thus we ignored branching. Then we added clustering test based on Gaussian mixtures. We choose to
use this approach since in DTI model Rician noise at typical noise-to-signal ratios is well approximated by Gaussian noise; see for example discussion in Zhu et al. (2007). Before estimation we performed the standard clustering test at each voxel $x_j$ and identified locations where the test found 2 significant clusters of tensors $M^{(1)}$ and $M^{(2)}$. Then near these locations we clustered the tensor field and performed the whole technique separately in each cluster, that is the kernel smoothing was applied to estimated $\hat{M}_n^{(i)}$, $i = 1, 2$, separately, followed by finding the fields of main eigenvectors and tracing 2 fibers. To our dismay adding clustering did not dramatically change the performance of CS method. We even tried to go to an extreme and forced clustering at each voxel. In this case the global connectivity metrics got worse, the frequency of fibers going off the template and just stopping increased significantly.

For each curve on the template we started its tracing at a point in the regions of interest at both ends. The tracing of curve 1 is good. One direction gives better results than when we start from the other end region. The method fails to trace the complete curve 2. We varied the initial point inside both end regions and the tracing was unstable. When starting from one end it tends to go into curve 3, from the opposite end it goes into curve 5. The tracing of curve 3 is also unstable. When we varied the initial point from one end curve 3 tends to go into curve 2 and then into curves 5-6-7. Starting from the opposite end it goes into curve 7. The tracing of curve 4 is good. One direction gives better results for the confidence ellipsoids than when we start from the other end region. Curves 5-6-7 go together when started from left end region. Curve 5 starts ok but at the end it turns into curve 2. The tracing of curve 6 is good. Starting at any endpoint yields equally good results with tight confidence ellipsoids. Curve 7 either immediately turns into curve 3 or goes ok till it meets with curve 2 where it turns into it. To summarize curves 2, 3, 5, and 7 present the most challenge for CS method. Figure 4 illustrates what happens when the initial point varied in the end regions of these curves. Let us stress that unlike probabilistic tractography methods outputs, these tracks are not realizations of a random procedure. They are different because the initial point is different, leading to different averaging windows moving along the estimated curve. Again we stress that if the method is repeated for the same dataset with the same initial location and the same parameters $(K, h_n, \beta)$ it results in the same estimated curve with the same confidence ellipsoids. Thus, ABC and PC for curves 1, 4, and 6 are 100% and they are 0% for curves 2, 3, and 5. Since curve 7 is recovered ok if the tracing starts from one endpoint but not from the other, ABC and PC are 50%.

The overall behaviour described above remained mostly the same when we analyzed
Table 1: Metrics from the Tractometer evaluated for CS method. Columns "the best probab" and "best determ" contain the best values for each metric by different methods from Côté et al. (2013). The column "the goal" gives the best desired levels of each metric.

<table>
<thead>
<tr>
<th>Metric</th>
<th>b=650</th>
<th>b=1500</th>
<th>b=2000</th>
<th>The goal</th>
<th>best probab</th>
<th>best determ</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABC</td>
<td>.8014</td>
<td>.7845</td>
<td>.7698</td>
<td>1</td>
<td>.918</td>
<td>.809</td>
</tr>
<tr>
<td>VC</td>
<td>.5714</td>
<td>.5714</td>
<td>.5714</td>
<td>1</td>
<td>.146</td>
<td>.247</td>
</tr>
<tr>
<td>IC</td>
<td>.1026</td>
<td>.1538</td>
<td>.1538</td>
<td>0</td>
<td>.008</td>
<td>.122</td>
</tr>
<tr>
<td>NC</td>
<td>.0513</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>.757</td>
<td>.604</td>
</tr>
<tr>
<td>VB</td>
<td>.4286</td>
<td>.4286</td>
<td>.4286</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>IB</td>
<td>.1026</td>
<td>.1538</td>
<td>.1538</td>
<td>0</td>
<td>.192</td>
<td>.122</td>
</tr>
</tbody>
</table>

Data with 3 different levels of noise $b = 650, 1500, 2000 \, \text{mm}^2$ given on the Tractometer website. As expected, the confidence ellipsoids became wider and covered more volume. For example, along curve 4 the norms of the ellipsoids increase by an approximate factor 2 when we switch from $b = 650$ to $b = 1500$, and by approximately 1.5 when we go from $b = 1500$ to $b = 2000$.

Finally, we consider the metrics proposed in the Tractometer. The overall results are summarized in Table 1. For average bundle coverage (ABC) we consider one estimated trajectory together with the 95% confidence ellipsoids along it in place of the bundle of trajectories usually produced by a probabilistic tractography method. Valid connections (VC), valid bundles (VB) and invalid bundles (IB) are calculated as ratios out of 7. Invalid connections (IC) and no connections (NC) are calculated as ratios out of 39. The numbers for the best probabilistic and best deterministic methods are taken from the paper of Daducci et al. (2013).

CS method does not recover 7/7 VB but it simultaneously achieves high ABC of 80.14%, high VC of 57.14% and low IB 6/39. If one compares it against deterministic tractography methods, it is superior to all of them. If one compares it against probabilistic tractography methods, it has comparable results for ABC metric and better for all the other metrics. It is worthwhile to note that the best values for different metrics in Table 1 are achieved by different methods. So for example, the best probabilistic method that gives ABC 91.8% has very low VC 6.7%, very high IB 94% and very high NC 90.1%. Furthermore, we note that CS method is faster and it additionally provides more geometrical information through confidence ellipsoids rather than through a bundle of repeated tracts.
4 Discussion and conclusion

CS method does not recover 7/7 VB but it simultaneously achieves high ABC of 80.14%, high VC of 57.14% and low IB 6/39. If one compares it against deterministic tractography methods, it is superior to all of them. If one compares it against probabilistic tractography methods, it has comparable results for ABC metric and better for all the other metrics. But it is faster and additionally provides more geometrical information through confidence ellipsoids rather than through a bundle of repeated tracts.

Note that the comparison metrics in Tractometer are mainly image-based rather than geometry-based. Meaning that for example ABC counts the proportion of voxels that are inside the fiber bundle on the template through which a prospective method traces the estimated fiber. A geometry-based approach would be to consider fibers as curves in 3D space and study local or global proximity to them of the estimated curves, their geometric features such as sharp changes of curvature, branching, and stopping. For such a geometric-based comparison having estimated curves together with confidence ellipsoids along them is more meaningful than having bundles of 10K simulated estimated streamlines. The later often exhibit different geometric features, and the overall average behavior in a bundle of streamlines is difficult to ascertain.

The image-based metrics are designed for probabilistic tractography methods and put deterministic or statistical methods at disadvantage from the start. But even then the method of CS performs quite well. Even though CS method does not recover 7/7 VB, it simultaneously achieves high ABC, high VC, and low IB. It performs very well compared to both deterministic and probabilistic tractography methods. It is also computationally fast and additionally it provides more information than the probabilistic methods, since it geometrically shows how the uncertainty changes along the estimated fiber as it moves through regions with several fibers located close by or near branching points.

References


(a) Fibers are estimated for an X pattern.  

(b) Template from Côté et al. (2013).

Figure 1: Reference images.
Figure 2: Fibers are estimated without clustering tests.

(a) Estimated fibers when $b$ is 650.
(b) Estimated curves in (a) with 95% confidence ellipsoids.
(c) Estimated fibers when $b$ is 1500 with 95% confidence ellipsoids.
(d) Estimated fibers when $b$ is 2000 with 95% confidence ellipsoids.

Figure 3: Reference images of typical curve estimators.

(a) Fibers are estimated without clustering tests.
(b) Estimators in (a) with 95% confidence ellipsoids.
(c) Fibers are estimated with clustering tests.
(a) Tracings of fiber 2 started at different locations.

(b) Tracings of fiber 3 started at different locations.

(c) Tracings of fiber 5 started at different locations.

(d) Tracings of fiber 7 started at different locations.

Figure 4: Fibers are estimated with and without clustering tests, $b$ is 650.