The Abdus Salam International Centre for Theoretical Physics

United Nations Educational, Scientific and Cultural Organization

IAEA International Atomic Energy Agency

2044-10

Summer School and Advanced Workshop on Trends and Developments in Linear Algebra

22 June - 10 July, 2009

The differential geometry of the space of symmetric positive-definite matrices and its applications in engineering

Maher Moakher

National Engineering School at Tunis
Tunisia

The differential geometry of the space of symmetric positive-definite matrices and its applications in engineering

Maher Moakher

National Engineering School at Tunis

Advanced Workshop on Trends and Developments in Linear Algebra

ICTP, Trieste, July 2009

Outline

1. **Space of symmetric positive-definite matrices**
   - Geometry
   - Geodesics
   - Differential operators
2. **Means and interpolation**
   - Means
   - Interpolation
   - Anisotropy
3. **PDE regularization**
   - Geometric heat flow
   - Geometric Perona-Malik flow
   - Mean curvature flow

Space of symmetric positive-definite matrices

Let

$$\mathscr{S}(n) := \{\boldsymbol{A} \in \mathcal{M}(n),\ \boldsymbol{A}^T = \boldsymbol{A}\},$$

and

$$\mathscr{P}(n) := \{\boldsymbol{A} \in \mathscr{S}(n),\ \boldsymbol{A} > 0\}.$$

Then $\mathscr{P}(n)$ is the interior of a pointed convex cone.

cone: $\boldsymbol{A} \in \mathscr{P}(n) \Rightarrow t\boldsymbol{A} \in \mathscr{P}(n), t > 0.$

convex: $\boldsymbol{A}, \boldsymbol{B} \in \mathscr{P}(n) \Rightarrow t\boldsymbol{A} + (1-t)\boldsymbol{B} \in \mathscr{P}(n), 0 \leq t \leq 1.$

The boundary of $\mathscr{P}(n)$ is the set of singular positive semidefinite matrices.

Symmetric positive-definite 2-by-2 matrices

$$\mathscr{P}(2) := \left\{ \begin{bmatrix} a & c \\ c & b \end{bmatrix},\ a > 0,\ ab - c^2 > 0 \right\}.$$

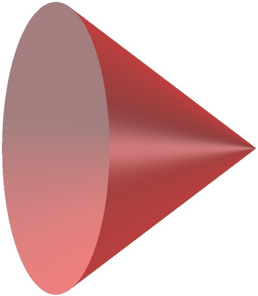
With the change of variables

$$u = \tfrac{1}{2}(a+b), \quad v = \tfrac{1}{2}(a-b),$$

the conditions of positive definiteness become

$$c^2 + v^2 < u^2,\ u > 0,$$

which are clearly the equations of the forward light cone.



Differential geometry of $\mathscr{P}(n)$

The set $\mathscr{P}(n)$ is a differentiable manifold of dimension $\frac{1}{2}n(n+1)$.

The tangent space to $\mathscr{P}(n)$ at any of its points $\boldsymbol{P}$ is

$$T_{\boldsymbol{P}}\mathscr{P}(n) = \{\boldsymbol{P}\} \times \mathscr{S}(n).$$

On $T_{\boldsymbol{P}}\mathscr{P}(n)$, we introduce the inner product

$$\langle \boldsymbol{A}, \boldsymbol{B} \rangle_{\boldsymbol{P}} := \operatorname{tr}(\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}^{-1}\boldsymbol{B})$$

which depends smoothly on the base point $\boldsymbol{P}$.

This inner product yields a natural Riemannian metric on $\mathscr{P}(n)$ given by

$$ds^2 = \operatorname{tr}\left(\boldsymbol{P}^{-1}d\boldsymbol{P}\boldsymbol{P}^{-1}d\boldsymbol{P}\right) = \|P^{-1/2}dPP^{-1/2}\|_F^2,$$

where $d\boldsymbol{P}$ is the symmetric matrix with elements (dP_{ij}).

Some references

This metric was first introduced by Siegel, and studied by Atkinson & Mitchell, Burbea, Maaß, Skovgaard, etc.

- C. L. Siegel, Symplectic Geometry, Academic Press, New York, 1964.
- C. Atkinson, A. F. Mitchell, *Rao's distance measure*, Sankhya, Ser. A **43**, (1981), pp. 345–365.
- J. Burbea, *Informative geometry of probability spaces*, Expo. Math., **4** (1986), pp. 347–378.
- H. Maass, Siegel's Modular Forms and Dirichlet Series, Lecture Notes in Math. 216, Springer-Verlag, Heidelberg, 1971.
- L.T. Skovgaard, *A Riemannian geometry of the multivariate normal model*, Scandinavian J. Statist., **11** (1984), pp. 211–233.

More recently, Amari, Bhatia, Bhatia & Holbrook, Calvo & Oller, Lawson & Lim, M.M., Petz, Zéraï & M.M., etc.

What's special about this metric?

This metric arises naturally from the log barrier function

$$\Psi(\boldsymbol{P}) = -\ln\det\boldsymbol{P}$$

used in cone programming. (Y. Nesterov, ...)

In statistical mechanics the negative of this function is called the Boltzmann entropy. (D. Petz, ...)

In information theory the negative of this function is called the information potential.

The unidimensional picture

When $n = 1$, this metric becomes $ds = \frac{dp}{p}$ which can be integrated to yields

$$\ln p = s + c,$$

or equivalently

$$p = C \exp s.$$

So the Riemannian distance between two positive numbers p and q is

$$|\log p - \log q|.$$

Recall that in the hyperbolic geometry of the Poincaré upper half-plane, the length of the geodesic segment joining the points $P(a, y_1)$ and $Q(a, y_2)$, $y_1, y_2 > 0$, is $|\log y_1 - \log y_2|$.

Invariance properties

Proposition

The Riemannian metric is invariant under

i) *Congruent transformations:* $\boldsymbol{P} \to \boldsymbol{C}\boldsymbol{P}\boldsymbol{C}^T$, *i.e.,*

$$ds^2(\boldsymbol{C}\boldsymbol{P}\boldsymbol{C}^T) = ds^2(\boldsymbol{P}),$$

ii) *Inversion:* $\boldsymbol{P} \to \boldsymbol{P}^{-1}$, *i.e.,*

$$ds^2(\boldsymbol{P}^{-1}) = ds^2(\boldsymbol{P}).$$

$$\boldsymbol{P} \to \boldsymbol{C}\boldsymbol{P}\boldsymbol{C}^T \implies d\boldsymbol{P} \to \boldsymbol{C}d\boldsymbol{P}\boldsymbol{C}^T,$$

$$\boldsymbol{P} \to \boldsymbol{P}^{-1} \implies d\boldsymbol{P} \to -\boldsymbol{P}^{-1}d\boldsymbol{P}\boldsymbol{P}^{-1}.$$

Parametrization of $\mathscr{P}(n)$

$\mathscr{P}(n)$ is a Riemannian manifold of dimension $\frac{1}{2}n(n+1)$.

An elements $\boldsymbol{P}$ of $\mathscr{P}(n)$ can be parametrized in many different ways. For example, it can be parametrized by:

- the $\frac{1}{2}n(n+1)$ entries of an invertible upper triangular matrix $\boldsymbol{L}$ (Cholesky decomposition: $\boldsymbol{P} = \boldsymbol{L}\boldsymbol{L}^T$).
- the $\frac{1}{2}n(n+1)$ entries of a symmetric matrix $\boldsymbol{S}$ (Exponential map: $\boldsymbol{P} = \exp \boldsymbol{S}$).
- the $\frac{1}{2}n(n-1)$ entries of an orthogonal matrix $\boldsymbol{R}$ and the n elements of a diagonal matrix $\boldsymbol{D}$ (Spectral decomposition: $\boldsymbol{S}\ \boldsymbol{P} = \boldsymbol{R}\boldsymbol{D}\boldsymbol{R}^T$).
- the $\frac{1}{2}n(n+1)$ entries of $\boldsymbol{P}$

$$\boldsymbol{P} = \sum_{i=1}^{n} P_{ii}\boldsymbol{E}_{ii} + \sum_{1\le i<j\le n} P_{ij}(\boldsymbol{E}_{ij} + \boldsymbol{E}_{ji}).$$

Matrix vectorization

Let $\boldsymbol{A}$ be an $n \times n$ matrix and let $A_{.j}$ be its j-th column. We denote by $\operatorname{vec} \boldsymbol{A}$ the n^2-column vector

$$\boldsymbol{A} = \begin{bmatrix} A_{.1} & \dots & A_{.n} \end{bmatrix}, \qquad \operatorname{vec} \boldsymbol{A} = \begin{bmatrix} A_{.1} \\ \vdots \\ A_{.n} \end{bmatrix}.$$

We recall that

$$(\operatorname{vec} \boldsymbol{A})^T \operatorname{vec} \boldsymbol{B} = \operatorname{tr}(\boldsymbol{A}^T \boldsymbol{B}),$$

and

$$(\operatorname{vec} \boldsymbol{A})^T (\boldsymbol{B} \otimes \boldsymbol{C}) \operatorname{vec} \boldsymbol{D} = \operatorname{tr}(\boldsymbol{D} \boldsymbol{B}^T \boldsymbol{A}^T \boldsymbol{C}).$$

Vectorization of a symmetric matrix

If $\boldsymbol{A}$ is symmetric

$$\boldsymbol{A} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1,n} \\ a_{21} & a_{22} & \ddots & \vdots \\ \vdots & \ddots & \ddots & a_{n-1,n} \\ a_{n1} & \dots & a_{n,n-1} & a_{nn} \end{bmatrix}$$

then $\frac{1}{2}n(n-1)$ elements of $\text{vec}(\boldsymbol{A})$ are redundant.
We denote by $\upsilon(\boldsymbol{A})$ the $\frac{1}{2}n(n+1)$-vector that is obtained from $\text{vec}(\boldsymbol{A})$ by eliminating the redundant elements

$$\text{vec}\,\boldsymbol{A} = \boldsymbol{D}_n \upsilon(\boldsymbol{A}).$$

$\boldsymbol{D}_n$ is called the duplication matrix. It has full column rank $\frac{1}{2}n(n+1)$, so that

$$\upsilon(\boldsymbol{A}) = \boldsymbol{D}_n^{+}\,\text{vec}\,\boldsymbol{A}.$$

Duplication matrix in $\mathscr{S}(3)$

When $n = 3$, we have

$$\operatorname{vec} \boldsymbol{A} = \begin{bmatrix} a_{11} & a_{21} & a_{31} & \color{red}{a_{12}} & a_{22} & a_{32} & \color{red}{a_{13}} & \color{red}{a_{23}} & a_{33} \end{bmatrix}^T,$$

$$v(\boldsymbol{A}) = \begin{bmatrix} a_{11} & a_{22} & a_{33} & a_{21} & a_{32} & a_{31} \end{bmatrix}^T,$$

and

$$\boldsymbol{D}_3 = \left[\begin{array}{ccc|cc|c} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ \hline 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{array}\right], \quad \boldsymbol{D}_3^{+T} = \left[\begin{array}{ccc|cc|c} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{2} \\ \hline 0 & 0 & 0 & \frac{1}{2} & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{2} & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & \frac{1}{2} \\ 0 & 0 & 0 & 0 & \frac{1}{2} & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{array}\right]$$

Some useful properties

For an invertible matrix $\boldsymbol{A}$ in $\mathscr{S}(n)$, we have

$$[\boldsymbol{D}_n^T(\boldsymbol{A} \otimes \boldsymbol{A})\boldsymbol{D}_n]^{-1} = \boldsymbol{D}_n^+(\boldsymbol{A}^{-1} \otimes \boldsymbol{A}^{-1})\boldsymbol{D}_n^{+T},$$

and

$$\det[\boldsymbol{D}_n^T(\boldsymbol{A} \otimes \boldsymbol{A})\boldsymbol{D}_n] = 2^{\frac{1}{2}n(n-1)}(\det \boldsymbol{A})^{n+1}.$$

J. R. MAGNUS, Linear Structures, Oxford University Press, London, 1988.

J. R. MAGNUS AND H. NEUDECKER, *The elimination matrix: some lemmas and applications*, SIAM J. Alg. Disc. Meth., **1** (1980), pp. 422–449.

Metric tensor

Using the reduced vectorization of $d\boldsymbol{P}$ we have

$$ds^2(\boldsymbol{P}) = \upsilon(d\boldsymbol{P})^T \boldsymbol{G}(\boldsymbol{P})\upsilon(d\boldsymbol{P}),$$

where $\boldsymbol{G}(\boldsymbol{P})$ is a symmetric positive-definite matrix called the metric tensor.

Proposition (Zéraï & M.M.)

The matrix $\mathbf{G}$ *of components of the metric tensor is*

$$[g_{\alpha\beta}(\boldsymbol{P})] = \boldsymbol{G}(\boldsymbol{P}) = \boldsymbol{D}_n^+ \left(\boldsymbol{P}^{-1} \otimes \boldsymbol{P}^{-1}\right) \boldsymbol{D}_n^{+T},$$

the matrix $\mathbf{G}^{-1}$ *of components of the inverse metric tensor is*

$$[g^{\alpha\beta}(\boldsymbol{P})] = \boldsymbol{G}^{-1}(\boldsymbol{P}) = \boldsymbol{D}_n^+ \left(\boldsymbol{P} \otimes \boldsymbol{P}\right) \boldsymbol{D}_n^{+T},$$

and $\det(\boldsymbol{G}(\boldsymbol{P})) = 2^{n(n-1)/2} \left(\det(\boldsymbol{P})\right)^{(n+1)}$.

Christoffel symbols

The *Christoffel symbols* (of the second kind) are:

$$\Gamma^{\delta}_{\alpha\beta} = \frac{1}{2} g^{\gamma\delta} (\partial_\alpha g_{\gamma\beta} + \partial_\beta g_{\gamma\alpha} - \partial_\gamma g_{\alpha\beta}).$$

Proposition (Zéraï & M.M.)

In the coordinate system (p^α)*, the components of the Christoffel symbols are given by*

$$\Gamma^{\gamma}_{\alpha\beta}(\boldsymbol{P}) = -[D_n^T \left(P^{-1} \otimes E^\gamma\right) D_n]_{\alpha\beta}, \quad 1 \leq \alpha, \beta, \gamma \leq \frac{1}{2} n(n+1),$$

where $\{E^\gamma\}_{1 \leq \gamma \leq n(n+1)/2}$ *is the dual basis to* $\{E_\gamma\}_{1 \leq \gamma \leq n(n+1)/2}$.

Christoffel symbols for $\mathscr{P}(3)$

$$\Gamma^1 = \frac{-1}{\rho}\begin{bmatrix} s^1 & 0 & 0 & s^4 & 0 & s^6 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ s^4 & 0 & 0 & s^2 & 0 & s^5 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ s^6 & 0 & 0 & s^5 & 0 & s^3 \end{bmatrix}, \quad \Gamma^2 = \frac{-1}{\rho}\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & s^2 & 0 & s^4 & s^5 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & s^4 & 0 & s^1 & s^6 & 0 \\ 0 & s^5 & 0 & s^6 & s^3 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

$$\Gamma^3 = \frac{-1}{\rho}\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & s^3 & 0 & s^5 & s^6 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & s^5 & 0 & s^2 & s^4 \\ 0 & 0 & s^6 & 0 & s^4 & s^1 \end{bmatrix}, \quad \Gamma^4 = \frac{-1}{2\rho}\begin{bmatrix} 0 & s^4 & 0 & s^1 & s^6 & 0 \\ s^4 & 0 & 0 & s^2 & 0 & s^5 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ s^1 & s^2 & 0 & 2s^4 & s^5 & s^6 \\ s^6 & 0 & 0 & s^5 & 0 & s^3 \\ 0 & s^5 & 0 & s^6 & s^3 & 0 \end{bmatrix},$$

$$\Gamma^5 = \frac{-1}{2\rho}\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & s^5 & 0 & s^2 & s^4 \\ 0 & s^5 & 0 & s^6 & s^3 & 0 \\ 0 & 0 & s^6 & 0 & s^4 & s^1 \\ 0 & s^2 & s^3 & s^4 & 2s^5 & s^6 \\ 0 & s^4 & 0 & s^1 & s^6 & 0 \end{bmatrix}, \quad \Gamma^6 = \frac{-1}{2\rho}\begin{bmatrix} 0 & 0 & s^6 & 0 & s^4 & s^1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ s^6 & 0 & 0 & s^5 & 0 & s^3 \\ 0 & 0 & s^5 & 0 & s^2 & s^4 \\ s^4 & 0 & 0 & s^2 & 0 & s^5 \\ s^1 & 0 & s^3 & s^4 & s^5 & 2s^6 \end{bmatrix}.$$

Geodesics on $\mathscr{P}(n)$

Let $\boldsymbol{P} : [a, b] \to \mathscr{P}(n)$ such that $\boldsymbol{P}(a) = \boldsymbol{P}_1$ and $\boldsymbol{P}(b) = \boldsymbol{P}_2$, be a "sufficiently smooth" curve on $\mathscr{P}(n)$.

The length of the curve $\boldsymbol{P}(t)$ is defined by

$$\mathscr{L}(P) := \int_a^b [g_{\boldsymbol{P}}(\dot{\boldsymbol{P}}(t), \dot{\boldsymbol{P}}(t))]^{1/2}\, dt.$$

Extremal points of $\mathscr{L}$ are called geodesics. They satisfy the second-order differential equation

$$\ddot{\boldsymbol{P}} - \dot{\boldsymbol{P}}\boldsymbol{P}^{-1}\dot{\boldsymbol{P}} = \boldsymbol{0}.$$

The geodesic emanating from $\boldsymbol{P}_0 \in \mathscr{P}(n)$ in the direction of $\boldsymbol{S} \in \mathscr{S}(n)$ is

$$\boldsymbol{P}(t) = \boldsymbol{P}_0^{1/2} \exp(t\boldsymbol{P}_0^{-1/2}\boldsymbol{S}\boldsymbol{P}_0^{-1/2})\boldsymbol{P}_0^{1/2}, \quad t \in \mathbb{R}.$$

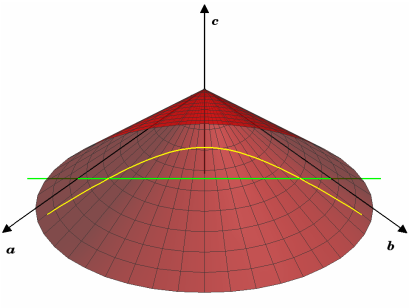
Geodesics on $\mathscr{P}(n)$

Euclidean geodesics:

$$P_E(t) = (1-t)A + tB, \quad t \in [0,1].$$

Riemannian geodesics:

$$P_R(t) = A(A^{-1}B)^t, \quad t \in \mathbb{R}.$$



It follows from Hopf-Rinow theorem that the metric space $(\mathscr{P}(n), d_R)$ is a complete metric space, whereas $(\mathscr{P}(n), d_E)$ is not complete.

Totally geodesic submanifolds of $\mathscr{P}(n)$

- Positive-definite diagonal matrices

$$\mathscr{D}(n) = \{\boldsymbol{D} \in \mathscr{P}(n),\ \boldsymbol{D} \text{ is diagonal}\}.$$

- Geodesic lines

$$\{\boldsymbol{P} \in \mathscr{P}(n),\ \boldsymbol{P} = \boldsymbol{A} \exp(t\boldsymbol{A}^{-1}\boldsymbol{S}\boldsymbol{A}^{-1})\boldsymbol{A},\ \boldsymbol{A} \in GL(n),\ \boldsymbol{S} \in \mathscr{S}(n),\ t \in \mathbb{R}\}.$$

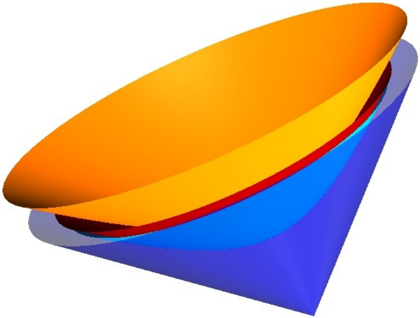
- Positive-definite matrices of constant determinant

$$\mathscr{S}\mathscr{P}_c(n) = \{\boldsymbol{p} \in \mathscr{P}(n),\ \det \boldsymbol{D} = c\}.$$

The hyperbolic geometry of $\mathscr{P}(n)$

The space $\mathscr{P}(n)$ can be seen as a foliated manifold whose codimension-one leaves are isomorphic to the hyperbolic space $\mathbb{H}^p$, where $p = \frac{1}{2}n(n+1) - 1$:

$$\mathscr{P}(n) = \mathscr{S}\mathscr{P}_1(n) \times \mathbb{R}_*^+$$



Riemannian distance

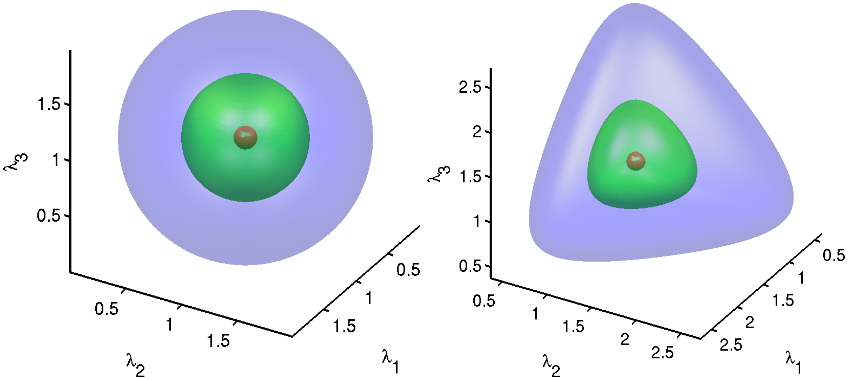
The Riemannian distance between $\boldsymbol{P}_1$ and $\boldsymbol{P}_2$ in $\mathscr{P}(n)$ is the length of the geodesic curve joining them

$$d_{\mathscr{P}(n)}(\boldsymbol{P}_1, \boldsymbol{P}_2) = \| \operatorname{Log}(\boldsymbol{P}_1^{-1/2} \boldsymbol{P}_2 \boldsymbol{P}_1^{-1/2}) \|_F = \left[\sum_{i=1}^{n} \ln^2 \lambda_i \right]^{\frac{1}{2}},$$

where $\lambda_i, i = 1, \ldots, n$, are the (positive) eigenvalues of $\boldsymbol{P}_1^{-1} \boldsymbol{P}_2$.

Note that this distance goes to infinity as one or both of the matrices approaches the boundary of $\mathscr{P}(n)$ consisting of singular positive semi-definite matrices.

Spheres on $\mathscr{P}(3)$ centered at the identity



Differential operators on $\mathscr{P}(n)$

- Gradient

$$\nabla_g f := g^{\alpha\beta} \frac{\partial f}{\partial p^\beta} = \boldsymbol{D}_n^+ (\boldsymbol{P} \otimes \boldsymbol{P}) \boldsymbol{D}_n^{+T} \frac{\partial f}{\partial \boldsymbol{p}},$$

- Divergence

$$\operatorname{div}_g V := \sqrt{g} \frac{\partial^T}{\partial p^\alpha} \left(\frac{1}{\sqrt{g}} V^\alpha \right) = \sqrt{g} \frac{\partial^T}{\partial \boldsymbol{p}} \left(\frac{1}{\sqrt{g}} V \right),$$

- Laplacian

$$\Delta_g f := \operatorname{div}_g \nabla_g f = \alpha \frac{\partial^T}{\partial \boldsymbol{p}} \left(\frac{1}{\alpha} \boldsymbol{D}_n^+ (\boldsymbol{P} \otimes \boldsymbol{P}) \boldsymbol{D}_n^{+T} \frac{\partial f}{\partial \boldsymbol{p}} \right),$$

where

$$\boldsymbol{p} = \upsilon(\boldsymbol{P}) = [p^1, \ldots, p^{n(n+1)/2}]^T, \quad g = \det(g^{\alpha\beta}) = 2^{n(n-1)/2} (\det \boldsymbol{P})^{n+1}.$$

Means and Interpolation

What does a mean mean?

Webster dictionary

mean: occupying a position about midway between extremes.
Etymology: mean is derived from the French root word mien whose origin is the Latin word medius, a term used to refer to a place, time, quantity, value, kind, or quality which occupies a middle position.

average: a single value (as a mean, mode, or median) that summarizes or represents the general significance of a set of unequal values.
Etymology: from earlier average proportionally distributed charge for damage at sea, modification of Middle French avarie damage to ship or cargo, from Old Italian avaria, from Arabic 'awAriyah damaged merchandise.

Generalization of the notion of mean

- Means of two positive numbers have been constructed by the ancient Greeks using geometric proportions. (Two overlapping lists of ten means by Nicomachus and Pappus.)
- The three principal means are generalized to more than two positive numbers, to positive functions, etc.
- In 1975, Anderson and Trapp, and Pusz and Woronowicz introduced the geometric mean for a pair of positive operators on a Hilbert space.
- Since, several researchers have been interested in means of positive definite matrices (Ando, Bhatia, Holbrook, Hiai, Kosaki, Kubo, Lawdon, Li, Lim, Mathias, Petz, Tamesi, etc.)

W. N. Anderson and G. E. Trapp, *Shorted operators*, SIAM J. Appl. Math., **28** (1975), pp. 60–71.

W. Pusz and S. L. Woronowicz, *Functional calculus for sesquilinear forms and the purification map*, Rep. Math. Phys., **8** (1975), pp. 159–170.

Metric-based definition of a mean

For m positive numbers $x_1, \dots, x_m$ we have

- The arithmetic mean is the unique minimizer of $\sum_{k=1}^{m} \mathrm{d}_e(x, x_k)^2$, where $\mathrm{d}_e(x,y) = |x-y|$ is the Euclidean distance in $\mathbb{R}$.
- The geometric mean is the unique minimizer of $\sum_{k=1}^{m} \mathrm{d}_h(x_k, x)^2$, where $\mathrm{d}_h(x,y) = |\log x - \log y|$ is the hyperbolic distance in $\mathbb{R}^{*+}$.

Definition (M.M. '02)

Let $(\mathcal{M}, \mathrm{d})$ be a differentiable manifold. A mean associated with $\mathrm{d}(\cdot,\cdot)$ of m points in $\mathcal{M}$ is defined by

$$\mathsf{M}(x_1, \dots, x_m) := \arg\min_{x \in \mathcal{M}} \sum_{k=1}^{m} \mathrm{d}(x_k, x)^2.$$

M.M., *Means and averaging in the group of rotations*, SIAM J. Matrix Anal. Appl., **24** (2002), pp. 1–16.

Means on $\mathscr{P}(n)$

Definition

A weighted mean associated with a distance $\mathrm{d}(\cdot,\cdot)$ of K symmetric positive-definite matrices $\boldsymbol{A}_1,\ldots,\boldsymbol{A}_K$ with weights $w_1,\ldots w_K$ is

$$\mathcal{M}(\boldsymbol{A}_1,\ldots,\boldsymbol{A}_K;w_1,\ldots,w_K)=\operatorname*{arg\,min}_{\boldsymbol{A}\in\mathscr{P}(n)}\sum_{k=1}^{K}w_k\,\mathrm{d}^2(\boldsymbol{A},\boldsymbol{A}_k).$$

M.M., *A differential-geometric approach to the geometric mean of symmetric positive-definite matrices*, SIAM J. Matrix Anal. Appl., **26** (2005), pp. 735–747.

Means on $\mathscr{P}(n)$

Euclidean mean (Weighted arithmetic mean):

$$\mathcal{M}_E(\boldsymbol{A}_1, \ldots, \boldsymbol{A}_K; w_1, \ldots, w_K) = \frac{1}{K}(w_1\boldsymbol{A}_1 + \cdots + w_k\boldsymbol{A}_K).$$

Riemannian mean (Weighted geometric mean):
$\mathcal{M}_R(\boldsymbol{A}_1, \ldots, \boldsymbol{A}_K; w_1, \ldots, w_K)$ is the unique solution to

$$\sum_{k=1}^{K} w_k \operatorname{Log}(\boldsymbol{X}\boldsymbol{A}_k^{-1}) = \boldsymbol{0}.$$

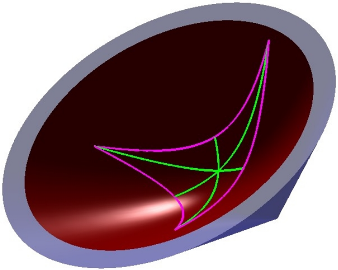
This nonlinear matrix equation can be solved in closed form for $K = 2$ or in special cases.

Otherwise, it can be solved numerically: M.M, Bini, etc.

Proposition (M.M. '05)

Given K positive-definite symmetric matrices $\{\boldsymbol{P}_k\}_{1\leq k\leq K}$ *in* $\mathscr{P}(n)$, *set* $\alpha_k = \sqrt[n]{\det \boldsymbol{P}_k}$ *and* $\tilde{\boldsymbol{P}}_k = \boldsymbol{P}_k/\alpha_k$. *Then the geometric mean of* $\{\boldsymbol{P}_k\}_{1\leq k\leq K}$ *is the geometric mean of* $\{\tilde{\boldsymbol{P}}_k\}_{1\leq k\leq K}$ *multiplied by the geometric mean of* $\{\alpha_k\}_{1\leq k\leq K}$, *i.e.,*

$$G(\boldsymbol{P}_1, \ldots, \boldsymbol{P}_K) = \sqrt[K]{\alpha_1 \cdot \alpha_K}\, G(\tilde{\boldsymbol{P}}_1, \ldots, \tilde{\boldsymbol{P}}_K).$$

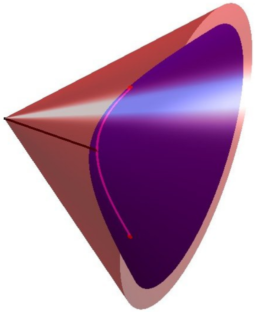


Proposition (M.M. '05)

The geometric mean of two positive-definite symmetric matrices $\boldsymbol{P}_1$ *and* $\boldsymbol{P}_2$ *in* $\mathscr{P}(2)$ *is given by*

$$G(\boldsymbol{P}_1, \boldsymbol{P}_2) = \sqrt{\alpha_1 \alpha_2} \frac{\sqrt{\alpha_2}\boldsymbol{P}_1 + \sqrt{\alpha_1}\boldsymbol{P}_2}{\sqrt{\det(\sqrt{\alpha_2}\boldsymbol{P}_1 + \sqrt{\alpha_1}\boldsymbol{P}_2)}}.$$

where α_1 *and* α_2 *are the determinants of* $\boldsymbol{P}_1$ *and* $\boldsymbol{P}_2$*, respectively.*



Multivariate interpolation of SPD matrices

- Classical multivariate interpolation on an n-simplex is a barycentric coordinate-weighted arithmetic average

$$u(x_1, \dots, x_{n+1}) = \sum_{i=1}^{n+1} \mu_i(x_1, \dots, x_{n+1}) u_i.$$

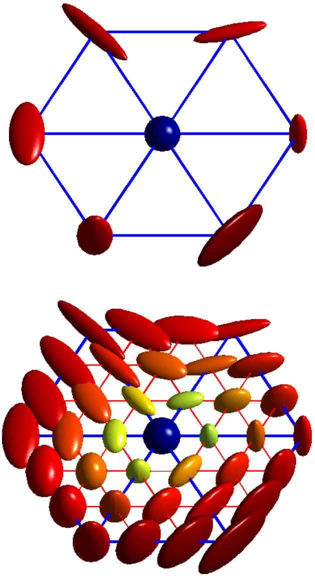
- By analogy, we define the geodesic multivariate interpolation on an n-simplex as a barycentric coordinate-weighted geometric average

$$\boldsymbol{A}(x_1, \dots, x_{n+1}) = \mathcal{M}_R(\boldsymbol{A}_i, \mu_i(x_1, \dots, x_{n+1})),$$

i.e., $\boldsymbol{A}(x_1, \dots, x_{n+1})$ is the solution of

$$\sum_{i=1}^{n+1} \mu_i(x_1, \dots, x_{n+1}) \operatorname{Log}(\boldsymbol{X}^{-1}\boldsymbol{A}_i) = \boldsymbol{0}.$$

Multivariate interpolation of SPD matrices



Closest isotropic matrix

Definition

Given a SPD matrix $\boldsymbol{A}$, the closest isotropic matrix, with respect to a distance $\mathrm{d}(\cdot,\cdot)$, is the matrix $\alpha\boldsymbol{I}$ where α is the minimizer of $\mathrm{d}(\boldsymbol{A},\beta\boldsymbol{I})$ over all $\beta > 0$..

Closest Euclidean: (equal trace)

$$\boldsymbol{A}_E = \frac{1}{3}\,\mathrm{tr}(\boldsymbol{A})\boldsymbol{I}.$$

Closest Riemannian: (equal determinant)

$$\boldsymbol{A}_R = (\det \boldsymbol{A})^{1/3}\boldsymbol{I}.$$

Anisotropy indices

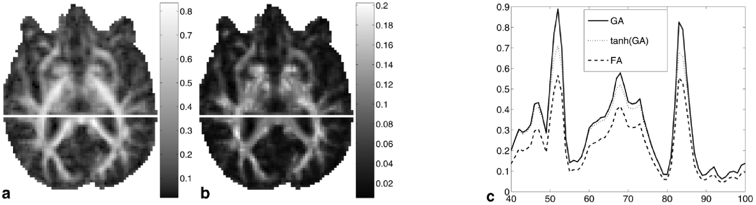
An anisotropy index of a symmetric positive-definite matrix $\boldsymbol{P}$ is a measure of nearness to isotropic matrices:

Euclidean anisotropy index:

$$A_E = \operatorname{tr}(\boldsymbol{P}^2) - \tfrac{1}{3}\operatorname{tr}^2 \boldsymbol{P}.$$

Riemannian anisotropy index:

$$A_R = \operatorname{tr}(\operatorname{Log}^2 \boldsymbol{P}) - \tfrac{1}{3}\operatorname{tr}^2(\operatorname{Log} \boldsymbol{P}).$$



References

- M.M., *Means and averaging in the group of rotations*, SIAM J. Matrix Anal. Appl., **24** (2002), pp. 1–16.
- M.M., *A differential-geometric approach to the geometric mean of symmetric positive-definite matrices*, SIAM J. Matrix Anal. Appl., **26** (2005), pp. 735–747.
- M.M AND P. G. BATCHELOR, *The symmetric space of positive definite tensors: From geometry to applications and visualization*, Ch. 17 in: Visualization and Processing of Tensor Fields, J. Weickert and H. Hagen, eds., Springer, Berlin, pp. 285–298, 2005.
- P. G. BATCHELOR, M.M., D. ATKINSON, F. CALAMANTE, AND A. CONNELLY, *Rigorous Framework for Diffusion Tensor Calculus*, Magn. Reson. Med., **53** (2005), pp 221–225.
- M.M., *On the averaging of symmetric positive-definite tensors*, J. Elasticity, **82** (2006), pp. 273–296.

Regularisation of SPD matrix fields

Diffusion tensor MRI

- Diffusion Tensor Magnetic Resonance Imaging (DT-MRI) is a new modality that provides a non-invasive probe into the microstructure of biological tissues.
- It measures the probability density function for the displacements of particles that undergo Brownian motion due to thermal fluctuations:

$$p(\boldsymbol{x}, \boldsymbol{x}_0; \tau) = \frac{1}{\sqrt{(4\pi\tau)^3 \det \boldsymbol{D}}} \exp\left(-\frac{(\boldsymbol{x} - \boldsymbol{x_0})^T \boldsymbol{D}^{-1} (\boldsymbol{x} - \boldsymbol{x}_0)}{4\tau}\right),$$

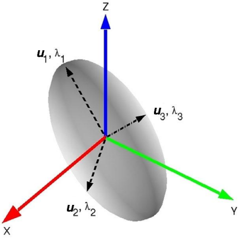
where $\boldsymbol{D}$ is a diffusion tensor.

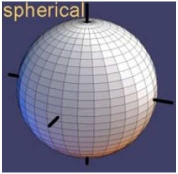
Diffusion ellipsoids

At each voxel, DT-MRI delivers a diffusion tensor (SPD)

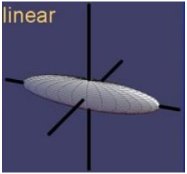
$$\boldsymbol{D} = \begin{pmatrix} D_{xx} & D_{xy} & D_{xz} \\ D_{yx} & D_{yy} & D_{yz} \\ D_{zx} & D_{zy} & D_{zz} \end{pmatrix} = \sum_{i=1}^{3} \lambda_i \boldsymbol{u}_i \otimes \boldsymbol{u}_i,$$

with $0 < \lambda_1 \leq \lambda_2 \leq \lambda_3$.

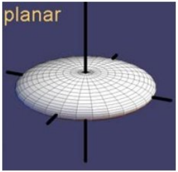




$\lambda_1 \sim \lambda_2 \sim \lambda_3$

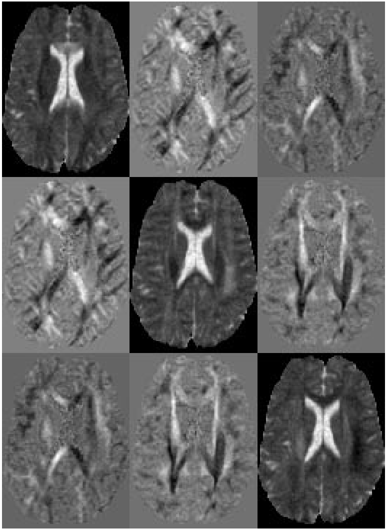


$\lambda_1 \sim \lambda_2 << \lambda_3$



$\lambda_1 << \lambda_2 \sim \lambda_3$

DT-MR image



PDE regularization of a scalar-valued image

Let $f : \Omega \to \mathbb{R}$ ($\Omega \subset \mathbb{R}^d$, $d = 2, 3$), be a given "noisy" image.
To regularize the image, we solve the PDE

$$\begin{cases} \partial_t u = \operatorname{div}\left(\boldsymbol{D}\nabla u\right), & \text{in } \Omega \times [0, T), \\ (\boldsymbol{D}\nabla u_{ij}) \cdot \boldsymbol{\nu} = 0, & \text{on } \partial\Omega \times [0, T), \\ u(\boldsymbol{x}, 0) = f(\boldsymbol{x}), & \text{in } \Omega. \end{cases}$$

- Linear diffusion (linear scale space): $\boldsymbol{D} = c\boldsymbol{I}$
- Nonlinear isotropic diffusion (Perona-Malik): $\boldsymbol{D} = g(|\nabla u|^2)\boldsymbol{I}$, with $g(\cdot)$ is a decreasing nonnegative function, e.g., $g(s) = 1/(1 + s/\lambda^2)$, or $g(s) = \exp(-s/\lambda^2)$.
- Nonlinear anisotropic diffusion: $\boldsymbol{D} = g(\nabla u \otimes \nabla u)$.
  If $\boldsymbol{A} = \boldsymbol{Q}\,\mathrm{diag}(\lambda_1, \ldots, \lambda_n)\boldsymbol{Q}^T$ with $\boldsymbol{Q}\boldsymbol{Q}^T = \boldsymbol{I}$, then
  $g(\boldsymbol{A}) = \boldsymbol{Q}\,\mathrm{diag}(g(\lambda_1), \ldots, g(\lambda_n))\boldsymbol{Q}^T$

Multi-channel images

A multi-channel image is a vector- or matrix-valued image:

$$f : \Omega \to \mathbb{R}^n \text{ (or } \mathbb{R}^{n\times m}\text{)}$$

Examples:

- Color images: 3 channels (RGB)
- DT-MRI images: 6 channels (6 independent diffusivities)

The simplest way of regularization of a multi-channel image is to regularize each channel independently.

But, because of the lack of correlation between channels, edges in each channel move independently by diffusion.

Therefore, the regularization of the different channels better be coupled.

Manifold-valued images as embedding maps

Let

(M, g) be a compact Riemannian manifold of dimension m

(N, h) be a complete Riemannian manifold of dimension n

A manifold-valued image is a map

$$\begin{aligned} \mathbf{\Phi} : M &\longrightarrow N \\ \boldsymbol{x} &\longmapsto \mathbf{\Phi}(\boldsymbol{x}) = (\Phi^1(\boldsymbol{x}), \ldots, \Phi^n(\boldsymbol{x})) \end{aligned}$$

M is called the domain manifold and N is called target manifold.

In the DT-MRI context, we take:

$M = \Omega \subset \mathbb{R}^3$ with the Euclidean metric, i.e., $g_{ij} = \delta_{ij}$.

$N = \mathscr{P}(3)$ with its Riemannian metric.

Harmonic energy functional

To the embedding map

$$\begin{aligned} \boldsymbol{\Phi} : M &\longrightarrow N \\ \boldsymbol{x} &\longmapsto \boldsymbol{\Phi}(\boldsymbol{x}) \end{aligned}$$

we associate the energy

$$E(\boldsymbol{\Phi}) = \frac{1}{2} \int_{\Omega} g^{ij}(\boldsymbol{x}) \frac{\partial \Phi^{\alpha}}{\partial x^{i}} \frac{\partial \Phi^{\beta}}{\partial x^{j}} h_{\alpha\beta}(\boldsymbol{\Phi}) \sqrt{g(\boldsymbol{x})} d\boldsymbol{x}.$$

This is a generalization of the Dirichlet integral.

In string theory, it is known as the Polyakov action.

It was introduced to the field of image processing by Kimmel, Sochen, and Malladi (1997)

Euler-Lagrange equation

Minimization of the energy $E(\Phi)$ yields the Euler-Lagrange equation

$$\frac{\delta E}{\delta \Phi^\mu} = -\left(\partial_i\left(\sqrt{g}g^{ij}\partial_j\Phi^\alpha\right) + \Gamma^\alpha_{\beta\gamma}(\mathbf{\Phi})\partial_i\Phi^\beta\partial_j\Phi^\gamma g^{ij}\right) h_{\alpha\mu}$$

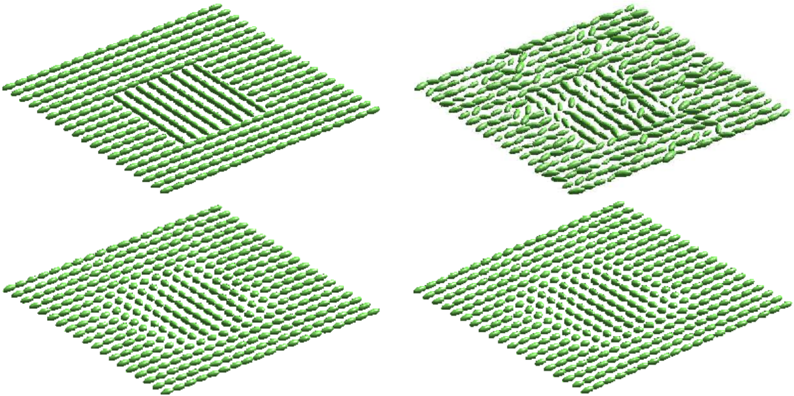
where $\Gamma^\alpha_{\beta\gamma}$ are the Christoffel symbols with respect to the metric h.
The regularization of a multi-channel image can be performed by the gradient descent

$$\Phi^\alpha_t = -\frac{1}{\sqrt{g}} h^{\alpha\mu} \frac{\delta E}{\delta \Phi^\mu}$$

With $g_{ij} = \delta_{ij}$, we have

$$\Phi^\alpha_t = \Delta\Phi^\alpha + \Gamma^\alpha_{\beta\gamma}(\mathbf{\Phi})\partial_i\Phi^\beta\partial_i\Phi^\gamma.$$

Smoothing by the geometric heat flow



Geometric Perona-Malik flow

Minimize the functional

$$E(\mathbf{\Phi}) = \int_{\Omega} \psi(\|\nabla^g \mathbf{\Phi}\|) d\Omega,$$

where

$$\psi(s) = \lambda^2 \log(1 + s^2/\lambda^2)/2.$$

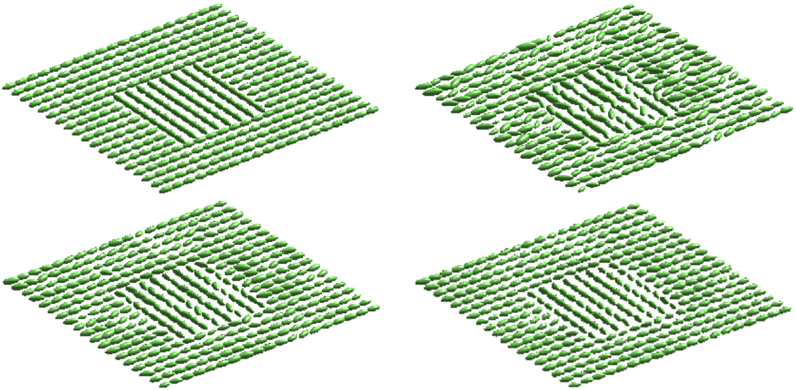
Gradient descent yields

$$\Phi_t^{\alpha} = \partial_i \left(D(\|\nabla^g \mathbf{\Phi}\|)\partial_i \Phi^{\alpha}\right) + D(\|\nabla^g \mathbf{\Phi}\|)\Gamma^{\alpha}_{\beta\gamma}(\mathbf{\Phi})\partial_i \Phi^{\beta} \partial_i \Phi^{\gamma},$$

with

$$D(s) := \frac{\psi'(s)}{2s} = \tfrac{1}{2}\left[1 + \left(\frac{s}{\lambda}\right)^2\right]^{-1}.$$

Smoothing by the geometric Perona-Malik flow



DT-MR images as immersions

Now we view a DT-MR image as a 3-dimensional surface in a 9-dimensional manifold:

$$\begin{aligned} \boldsymbol{\Psi} : M = \Omega &\to N = \Omega \otimes \mathscr{P}(3) \\ \boldsymbol{x} &\mapsto \boldsymbol{\Psi}(\boldsymbol{x}) = (\boldsymbol{x}, \boldsymbol{P}(\boldsymbol{x})) \end{aligned}$$

Here, M is the domain manifold and N is the space-feature manifold. Knowing the metric h on N and the map $\boldsymbol{\Psi}$ we can construct a metric g on M by the process of "pullback"

$$g_{ij}(\boldsymbol{x}) = h_{\alpha\beta}(\boldsymbol{\Phi})\partial_i \Psi^\alpha \partial_j \Psi^\beta$$

The Riemannian metric associated with this surface is then

$$g_{ij}(\boldsymbol{P}) = \delta_{ij} + \mathrm{tr}(\boldsymbol{P}^{-1}\partial_i \boldsymbol{P}\boldsymbol{P}^{-1}\partial_j \boldsymbol{P}).$$

Minimal immersion flow

Minimize the functional

$$E(\boldsymbol{P}) = \int_\Omega \sqrt{\det(g_{ij}(\boldsymbol{P}))}\,d\Omega.$$

Gradient descent yields

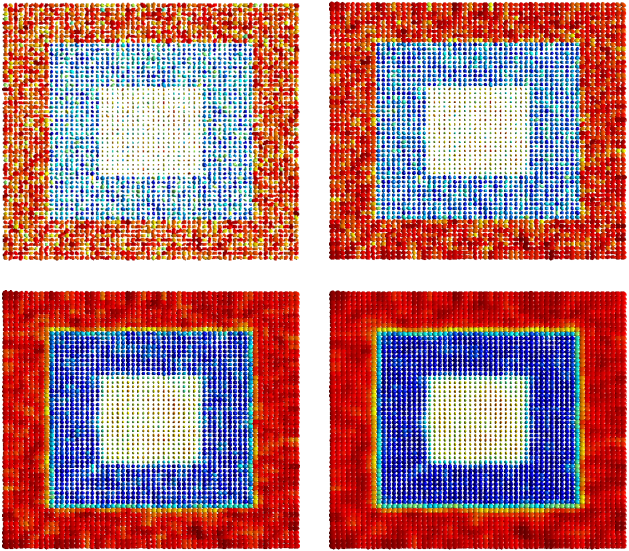
$$\partial_t p^\alpha = \Delta_g p^\alpha + \operatorname{tr}\left\{\Gamma^\alpha(\boldsymbol{P})\nabla\boldsymbol{p}\left[I_3 + (\nabla\boldsymbol{p})^T G\nabla\boldsymbol{p}\right]^{-1}(\nabla\boldsymbol{p})^T\right\},$$

with

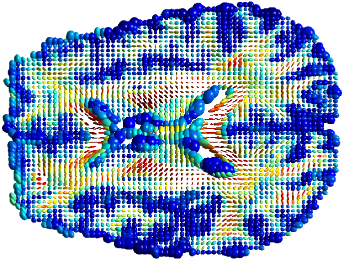
$$\boldsymbol{p} := \begin{pmatrix} P_{11} & P_{22} & P_{33} & P_{21} & P_{32} & P_{31} \end{pmatrix}^T$$

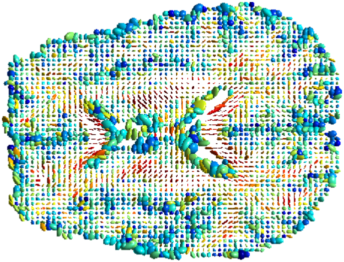
This is the mean curvature flow.

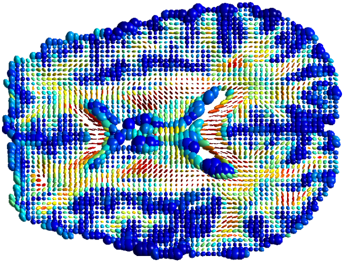
Smoothing by the minimal immersion flow



Smoothing of real DT-MRI data







References

- M. ZÉRAÏ AND M.M., *Riemannian Curvature-Driven Flows for Tensor-Valued Data*, In: Scale Space and Variational Methods, F. Sgallari, A. Murli and N. Paragios, Eds., LNCS 4485, pp. 592–602, 2007.
- M. ZÉRAÏ AND M.M., *The Riemannian Geometry of the Space of Positive-Definite Matrices and its Application to the Regularization of Diffusion Tensor MRI Data*, submitted to: Journal of Mathematical Imaging and Vision.

Acknowledgements

Collaborators:

- **Philip Batchelor**, UCL: Anisotropies
- **Mourad Zéraï**, ENIT: PDE regularization