Sponsorisé par le CEntre Franco-Indien pour la Promotion de la Recherche Avancée CEFIPRA | IFCPAR Sponsored by Indo-French Centre for the Promotion of Advanced Research

Indo-French Seminar

Matrix Information Geometry

MIG 2011
Palaiseau, France, February 2011
Extended Abstracts

Frédéric Barbaresco
Frank Nielsen
Olivier Schwander

Matrix Information Geometry

INDO-FRENCH WORKSHOP
Applications in Sensor and Cognitive Systems Engineering-
23rd February 2011 (Thales TRT)
24th-25th February 2011 (Ecole Polytechnique)
Extended Abstracts

Volume Editors

Frédéric Barbaresco
Thales Air Systems S.A.
Surface Radar Domain, Advanced Developments Department
E-mail : frederic.barbaresco@thalesgroup.com

Frank Nielsen
École Polytechnique
Sony Computer Science Laboratories,Inc.
E-mail: frank.nielsen@acm.org

Olivier Schwander
École Polytechnique
E-mail: olivier.schwander@polytechnique.edu

Introduction

Matrix and Tensor Data Processing is a breakthrough in the domain of signal, image and information processing with many potential applications in sensor and cognitive systems engineering. The participants are experts in the areas of theoretical mathematics or engineering sciences.

Topics

- Information Geometry
- Differential Geometry of structured Matrix
- Positive Definite Matrix
- Covariance Matrix
- Application for Sensors (EM, EO, Acoustic,. . .)
- Applications for Cognitive systems (Data Mining, Data analysis, . . .)

Organization

Co-chairs

- F. Barbaresco (Thales, France)
- A. K. Mishra (Indian Institute of Technology Guwahati, India)
- F. Nielsen (Ecole Polytechnique, France/Sony Computer Science Laboratories, Inc, Japan)

Organizing committee

- F. Barbaresco (Thales, France)
- A. K. Mishra (Indian Institute of Technology Guwahati, India)
- F. Nielsen (École Polytechnique, France/Sony Computer Science Laboratories, Inc, Japan)
- O. Schwander (École Polytechnique, France)

Sponsoring Institutions

- CEFIPRA/IFCPAR: Indo-French Centre for the Promotion of Advanced Research
- École Polytechnique
- Sony CSL
- Thales

Table of Contents

Mathematical Morphology for Matrix-Valued Images

Jesús Angulo

CMM-Centre de Morphologie Mathématique, Mathématiques et Systèmes, MINES ParisTech; 35, rue Saint Honoré, 77305 Fontainebleau Cedex, France
jesus.angulo@mines-paristech.fr

Context. Mathematical morphology is a nonlinear image processing methodology originally developed for binary and greyscale images [9]. It is based on the computation of maximum (dilation operator) and minimum (erosion operator) in local neighbourhoods called structuring elements [10]. That means that the definition of morphological operators needs an ordering relationship between the points to be processed. Theory of morphological operators was then formulated in the framework of complete lattices [7], where the basic structure is a partial ordering and the existence of supremum and infimum for any set of points should be guaranteed. Dilation and erosion can be also computed using an eikonal PDE [1]. In addition, dilation and erosion can be also studied in the framework of convex analysis, as the supremum/infimum convolution in the $(\max, +)/(\min, +)$ algebras, which the corresponding connection with the Legendre transform [8].

Matrix and tensor valued images appear nowadays in various image processing fields and applications [11], e.g., structure tensor images representing the local orientation and edge information, diffusion tensor magnetic resonance imaging and covariance matrices in radar imaging.

State-of-the-art. Morphological operators and filters perform noise suppression, contrast image enhancement, structure extraction and decomposition, etc. [10]. The extension of mathematical morphology to matrix-valued images has been addressed exclusively by Burgeth et al. [5] [6]. They have considered two different approaches: the first one is based on the Löwner partial ordering [12], where the sup/inf of matrices are computed using convex matrix analysis tools; the other one corresponds to the generalization of the morphological PDE to matrix data.

Aim of the study. The goal of this work is to introduce various alternatives ways to extend mathematical morphology, which are different from those introduced by Burgeth et al. In particular, focussing on positive definite symmetric matrices, three different families of approaches are explored.

- *Partial spectral ordering and inverse eigenvalue problem.* By considering the partial ordering based on a lexicographic cascade of eigenvalues, it is possible to define the sup/inf of a set of matrices as the matrix having as eigenvalues the sup/inf of eigenvalues. However, the definition of the orthogonal basis of corresponding sup is not trivial. A few possible algorithms for this inverse eigenvalue problem will be briefly mentioned.

– *Total orderings for input-preserving* sup-inf *operators.* An alternative methodology to use the spectral ordering consists in defining as supremum of a set of matrices, the matrix which is bigger according to the lexicographic priority of eigenvalues or according to a given priority between some matrix invariants associated to the eigenvalues. This kind of approaches is valid when a total ordering is defined. Consequently, the spectral information should be completed with additional conditions in the lexicographic cascade.
  In cases where a pair of reference matrix sets is defined (typically, a training set of matrices associated to the foreground and a training set of matrices associated to the background), it is also possible to define a total ordering according to the distances of each matrix to both reference sets. In such a technique, the distance between matrices is the key element for the ordering.
– *Asymptotic pseudo-morphological operators using counter-harmonic mean.* We have recently shown in [2] how the counter-harmonic mean [4] can be used to introduce nonlinear operators which asymptotically mimic dilation and erosion. The generalization of the counter-harmonic mean filter to matrix images leads to a simple methodology for nonlinearization of matrix PDE diffusion filtering.

References

1. L. Alvarez, F. Guichard, P-L. Lions, J-M. Morel. Axioms and fundamental equations of image processing. *Arch. for Rational Mechanics,* 123(3): 199–257, 1993.
2. J. Angulo. Pseudo-Morphological Image Diffusion using the Counter-Harmonic Paradigm. In *Proc.of Acivs'2010 (2010 Advanced Concepts for Intelligent Vision Systems),* LNCS Vol. 6474, Part I, p. 426–437, Springer, 2010.
3. J.K. Baksalary, F. Pukelsheim. On the Löwner, Minus, and Start Partial Orderings of Nonnegative Definite Matrices and Their Squares. *Linear Algebra and its Applications,* 151: 135–141, 1991.
4. P.S. Bullen. *Handbook of Means and Their Inequalities,* 2nd edition, Springer, 1987.
5. B. Burgeth, A. Bruhn, S. Didas, J. Weickert, M. Welk. Morphology for tensor data: Ordering versus PDE-based approach. *Image and Vision Computing,* 25(4): 496–511, 2007.
6. B. Burgeth, N. Papenberg, A. Bruhn, M. Welk, J. Weickert. Mathematical morphology for matrix fields induced by the Loewner ordering in higher dimensions. *Signal Processing,* 87(2): 277–290, 2007.
7. H.J.A.M. Heijmans. *Morphological Image Operators,* Academic Press, Boston, 1994.
8. P Maragos. Slope Transforms: Theory and Application to Nonlinear Signal Processing. *IEEE Trans. on Signal Processing,* 43(4): 864–877, 1995.
9. J. Serra. *Image Analysis and Mathematical Morphology,* London: Academic Press, 1982.
10. P. Soille. *Morphological Image Analysis,* Springer-Verlag, Berlin, 1999.
11. J. Weickert, H. Hagen (Eds.). *Visualization and Processing of Tensor Fields,* Springer-Verlag, Berlin, 2006.
12. J.K. Baksalary, F. Pukelsheim. On the Löwner, Minus, and Start Partial Orderings of Nonnegative Definite Matrices and Their Squares. *Linear Algebra and its Applications,* 151: 135–141, 1991.

Stochastic Algorithms for Computing p-Means of Probability Measures

Marc Arnaudon and Le Yang

Laboratoire de Mathématiques et Applications, CNRS : UMR 6086,
Université de Poitiers, 86962 Futuroscope Chasseneuil Cedex, France
Marc.Arnaudon@math.univ-poitiers.fr
Le.Yang@math.univ-poitiers.fr

Abstract. We give a stochastic algorithm to compute the p-mean of a probability measure μ whose support is contained in a regular geodesic ball of a Riemannian manifold. We show the almost sure convergence of this algorithm. Under a further regularity condition, an associated central limite theorem is also proved. Moreover, it is shown that a deterministic subgradient algorithm can also be used to compute the Riemmanian median of μ.

1 Stochastic algorithms for computing p-means of probability measures

Let M be a Riemannian manifold with Riemannian distance ρ. Assume that the sectional curvatures K satisfy $-\beta^2 \le K \le \alpha^2$. Let μ be a probability measure with support included in a geodesic ball $B(a,r)$ in M. Fix $p \in [1,\infty)$. We will always make the following assumptions on (r,p,μ):

Assumption 1 *The support of μ is not reduced to one point. Either $p > 1$ or the support of μ is not contained in a line, and the radius r satisfies*

$$r < r_{\alpha,p} \quad \text{with} \begin{cases} r_{\alpha,p} = \frac{1}{2}\min\left\{\mathrm{inj}(M), \frac{\pi}{2\alpha}\right\} & \text{if } p \in [1,2) \\ r_{\alpha,p} = \frac{1}{2}\min\left\{\mathrm{inj}(M), \frac{\pi}{\alpha}\right\} & \text{if } p \in [2,\infty) \end{cases} \tag{1}$$

Under assumption 1,the function $H_p : B(a,r) \longrightarrow \mathbb{R}_+$, $x \longmapsto \int_M \rho^p(x,y)\mu(dy)$ has a unique minimizer e_p in M, the p-mean of μ, and moreover $e_p \in B(a,r)$. If $p=1$, e_1 is the median of μ.

Theorem 1. *Let $(P_k)_{k\ge1}$ be a sequence of independent $B(a,r)$-valued random variables, with law μ. Let $(t_k)_{k\ge1} \subset (0, C_{p,\mu}]$ be a sequence of positive numbers satisfying*

$$\sum_{k=1}^{\infty} t_k = +\infty \quad \text{and} \quad \sum_{k=1}^{\infty} t_k^2 < \infty. \tag{2}$$

Letting $x_0 \in B(a,r)$, define inductively the random walk $(X_k)_{k\ge0}$ by

$$X_0 = x_0 \quad \text{and for } k \ge 0 \quad X_{k+1} = \exp_{X_k}\left(-t_{k+1}\,\mathrm{grad}_{X_k} F_p(\cdot, P_{k+1})\right) \tag{3}$$

where $F_p(x,y) = \rho^p(x,y)$, *with the convention* $\operatorname{grad}_x F_p(\cdot,x) = 0$.

The random walk $(X_k)_{k\geq 1}$ *converges in* L^2 *and almost surely to* e_p.

Theorem 2. *Let* $(X_k)_{k\geq 0}$ *be the time inhomogeneous* M*-valued Markov chain defined in Theorem 1 with* $t_k = \min\left(\frac{\delta}{k}, C_{p,\mu}\right)$, $k \geq 1$ *for some* $\delta > 0$. *We define for* $n \geq 1$ *the rescaled* $T_{e_p}M$*-valued Markov chain* $(Y_k^n)_{k\geq 0}$ *by*

$$Y_k^n = \frac{k}{\sqrt{n}} \exp_{e_p}^{-1} X_k. \tag{4}$$

Assume that H_p *is* C^2 *in a neighborhood of* e_p, *and that* $\delta > C_{p,\mu,K}^{-1}$. *The sequence of processes* $\left(Y_{[nt]}^n\right)_{t\geq 0}$ *weakly converges in* $\mathbb{D}((0,\infty), T_{e_p}M)$ *to a diffusion process* y_δ *given by*

$$y_\delta(t) = \sum_{i=1}^d t^{1-\delta\lambda_i} \int_0^t s^{\delta\lambda_i - 1} \langle \delta\sigma\, dB_s, e_i\rangle e_i, \quad t \geq 0, \tag{5}$$

where B_t *is a standard Brownian motion on* $T_{e_p}M$, $\sigma \in \operatorname{End}(T_{e_p}M)$ *satisfies*

$$\sigma\sigma^* = \mathbb{E}\left[\operatorname{grad}_{e_p} F_p(\cdot, P_1) \otimes \operatorname{grad}_{e_p} F_p(\cdot, P_1)\right],$$

$(e_i)_{1\leq i\leq d}$ *is an orthonormal basis diagonalizing the symmetric bilinear form* $\nabla dH_p(e_p)$ *and* $(\lambda_i)_{1\leq i\leq d}$ *are the associated eigenvalues.*

1.1 A subgradient algorithm for computing median

Initializing $x_1 \in B(a,r)$,

Do $x_{k+1} = \exp_{x_k}(-t_k \frac{H(x_k)}{|H(x_k)|})$,

While $H(x_k) \neq 0$,

where

$$H(x_k) = \int_{M\setminus\{x\}} \frac{-\exp_x^{-1} p}{\rho(x,p)} \mu(dp),$$

$$t_k \in (0,\varepsilon),\ t_k \to 0,\ \sum_{k=1}^{\infty} t_k = +\infty.$$

References

1. M. Arnaudon, C. Dombry, A. Phan and L. Yang, "*Stochastic algorithms for computing means of probability measures*", preprint hal-00540623, version 1, (2010).
2. L. Yang, "R*iemannian median and its estimation*", LMS J. Comput. Math. Vol. 13, 461-479, (2010).

Fréchet Metric Space, Homogeneous Siegel Domains & Radar Matrix Signal Processing

Frédéric Barbaresco

Thales Air Systems S.A.
Surface Radar Domain, Advanced Developments Department
frederic.barbaresco@thalesgroup.com

Information Geometry has been introduced by Indian scientist Calyampudi Radhakrishna Rao [15] (http://www.crraoaimscs.org/) PhD student of R.A. Fisher, and axiomatized by N. N. Chentsov [5], to define a distance between statistical distributions that is invariant to nonsingular parameterization transformations. C.R. Rao introduced this geometry in his 1945 seminal paper on Cramer-Rao Bound (this bound was discovered in parallel by Maurice René Fréchet [6] in 1939 and extended later to multivariate case by Georges Darmois). We will see that this geometry could be considered in the framework of Positive Definite Matrices Geometry for Complex circular Multivariate Laplace-Gauss Law (see R. Bhatia [2]). For Doppler / Array / STAP Radar Processing, Information Geometry Approach will give key role to Symmetric spaces and Homogeneous bounded domains geometry. For Radar, we will propose Information Geometry metric as Kähler metric, given by Hessian of Khler potential (Entropy of Radar Signal given by $-\log\det(R)$). To take into account Toeplitz structure of Time/Space Covariance Matrix or Toeplitz-Block-Toeplitz structure of Space-Time Covariance matrix, Parameterization known as Partial Iwasawa Decomposition [9] could be applied through Complex Autoregressive Model or Multi-channel Autoregressive Model. Then, Hyperbolic Geometry of Poincaré Unit Disk [14] or Symplectic Geometry of Siegel Unit Disk [17, 18] will be used as natural space to compute p-mean (p=2 for mean, p=1 for median) of covariance matrices via Fréchet/Karcher Flow [7, 10] for Weiszfeld algorithm extension on Manifold and on Frchet Metric spaces. We have tested also stochastic flow proposed by M. Arnaudon. This new mathematical framework will allow developing concept of OS (Ordered Statistic) for Hermitian Positive Definite Covariance Space/Time Toeplitz matrices or for Space-Time Toeplitz-Block-Toeplitz matrices. We will then define OS-HDRCFAR (Ordered Statistic High Doppler Resolution CFAR) and OS-STAP (Ordered Statistic Space-Time Adaptive Processing). This approach is based on the existence of a center of mass in the large for manifolds with non-positive curvature that was proven and used by Elie Cartan back in the 1920s [3]. The general case was employed by Calabi in an unpublished note. In 1977, Hermann Karcher [10] has proposed intrinsic flow to compute this barycenter, that we adapt for covariance matrices. This geometric foundation of Radar Signal Processing is based on general concept of Siegel domains [17, 18]. We will then give a brief history of Siegel domains studies in Europe, Russia and China. In 1935, Elie Cartan [4] proved that irreductible homogeneous bounded symmetric domains could be reduced to six types, included two exceptional ones. Four non-exceptionnal Cartans domains are now called classical models, and considered as the higher dimensional analogues of the Poincars unit disk [14] in the complex plane. After in the framework of Symplectic Geometry [17, 18], Carl Ludwig Siegel has introduced first explicit descriptions of symmetric domains, where the realization of bounded domains as unbounded domains played fundamental role (for an important class of them, these unbounded domains are Siegel domains of the first kind, with important particular case of Siegel Upper Half Plane). In 1953, Loo-Keng Hua [8] obtained the orthonormal system and the Bergman/Cauchy/Poisson kernel functions for each of the four classical domains using group representation theory. Elie Cartan proved that all bounded homogeneous complex domains in dimension 2 and 3 are symmetric and conjectured that is true for dimension greater than 3. Ilya Piatetski-Shapiro [13], after Hua works, has extended Siegel description to other symmetric domains and has disproved the Elie Cartan conjecture that all transitive domains are symmetric. A. Borel showed that if in a bounded homogeneous region a semisimple Lie group operates transitively, then that region is symmetric. These results were strengthened by Hano and obtained in parallel by Jean-Louis Koszul [11, 12] who also studied affinely homogeneous regions that are fundamental for Information Geometry and real Hessian or complex Kählerian geometries [16]. Piatetski-Shapiro introduced finally general definition of a Siegel domain of the second kind (all symmetric domains allow a generalization of Siegel tube domains), and has proved in 1963 with S.G. Gindikin and E. Vinberg that any bounded homogeneous domain has a realization as a Siegel domain of the second kind with transitive action of linear transformation. In parallel, E. Vinberg [19] worked on the theory of homogeneous convex cones, as fundamental construction of Siegels domains (he introduced a special class of generalized matrix T-algebras), and S.G. Gindikin worked on analytic aspects of Siegels domains. More recently, classical complex Symmetric spaces have been studied by F. Berezin [1]. With Karpelevitch, Piatetski-Shapiro explored underlying geometry of these complex homogeneous domains manifolds, and more especially, the fibering of domains over components of the boundary. Let a bounded domain, he constructed a fibering by looking at all the geodesic that end in each boundary component and associating the end point to every point on the geodesic. For our Radar STAP and Toeplitz-Block-Toeplitz covariances matrices, we have used Berger fibering in Unit siegel Disk based on the theorem that all symmetric spaces are fibered on a compact symmetric space (Mostow decomposition).

References

1. Feliks Aleksandrovich Berezin. Quantization in complex symmetric spaces. *Math. USSR*, 9:341–379, 1975.
2. Rajendra Bhatia. *Positive Definite Matrices (Princeton Series in Applied Mathematics).* Princeton University Press, illustrated edition edition, December 2006.
3. Elie Cartan. Groupes simples clos et ouverts et géométrie riemannienne. *Journal de mathématiques pures et appliquées*, (8):1–34, 1929.

4. Elie Cartan. Sur les domaines bornés de l'espace de n variables complexes. *Abh. Math. Seminar Hamburg*, 1935.
5. Nikolai Nikolaevich Chentsov. *Statistical Decision Rules and Optimal Inferences.* Transactions of Mathematics Monograph, numero 53, 1982. Published in russian in 1972.
6. Maurice Fréchet. Sur l'extension de certaines évaluations statistiques au cas de petits échantillons. *Review of the International Statistical Institute*, 11(3/4):182–205, 1943. published in IHP Lecture in 1939.
7. Maurice Fréchet. Les éléments aléatoires de nature quelconque dans un espace distancié. *Annales de l'institut Henri Poincaré*, 10(4):215–310, 1948.
8. Loo-Keng Hua. *Harmonic analysis of functions of several complex variables in the classical domains.* American Mathematical Society, 1958.
9. Kenkichi Iwasawa. On some types of topological groups. *Ann. of Math. (2)*, 50:507–558, 1949.
10. Hermann Karcher. Riemannian center of mass and mollifier smoothing. *Communications on Pure and Applied Mathematics*, 30(5):509–541, 1977.
11. Jean-Louis Koszul. Sur la forme hermitienne canonique des espaces homogénes complexes. *Can. J. of. Math.*, (7):562–576, 1955.
12. Jean-Louis Koszul. Exposés sur les espaces homogènes symétriques. *Publicacao da Sociedade de Matematica de Sao Paulo*, 1959.
13. Ilya Piatetskii-Shapiro. *Automorphic Functions and the Geometry of Classical Domains.* Gordon & Breach Science Publishers, 1969. 340 pages.
14. Henri Poincaré. Les fonctions analytiques de deux variables et la reprsentation conforme. *Rendiconti del Circolo Matematico di Palermo (1884 - 1940)*, 23:185–220, 1907.
15. Calyampudi Radhakrishna Rao. Information and the accuracy attainable in the estimation of statistical parameters. *Bulletin of the Calcutta Mathematical Society*, 37:81–89, 1945.
16. Hirohiko Shima. *The geometry of Hessian structures.* World Scientific, 2007.
17. C. L. Siegel. Sympletic geometry. *Amer. J. Math.*, (65):1–86, 1943.
18. C. L. Siegel. *Analytic functions of several complex variables.* Princeton Univ. Press, 1949.
19. Ernest Borisovich Vinberg. The theory of convex homogeneous cones. *Trans. Amer. Math. Soc.*, (12), 1963. Trudy Moskov. Mat. Obshch.

Matrix Inequalities

Rajendra Bhatia

Theoretical Statistics & Mathematics Unit,
Indian Statistical Institute,
7, S. J. S. Sansanwal Marg,
New Delhi-110 016, India
rbh@isid.ac.in

Inequalities, like the arithmetic-geometric mean inequality and the Cauchy-Schwarz inequality, have long been used in almost all areas of mathematics. Their matrix versions have been discovered in the last few years, and are equally important in several subjects. In this talk we will illustrate some general principles that have been found to be very useful in deriving such inequalities.

We will focus on two themes. The first is the arithmetic-geometric mean inequality $\|A^{1/2}B^{1/2}\| \leq \frac{1}{2}\|A + B\|$, where A and B are positive semidefinite matrices. The second is the problem of estimating $\|f(A) - f(B)\|$ in terms of $\|A-B\|$ for functions of matrices. Both problems have common ingredients, and both have generalisations that make them more attractive as well as useful.

Keywords: Matrix inequalities, positive definite matrix, arithmdetic-geometric mean, perturbation bound

Matrix Means

Rajendra Bhatia

Theoretical Statistics & Mathematics Unit,
Indian Statistical Institute,
7, S. J. S. Sansanwal Marg,
New Delhi-110 016, India
rbh@isid.ac.in

Averaging operations on positive definite matrices have been of interest in operator theory, engineering, physics and statistics. Because of noncommutativity of matrix multiplication and because of subtleties of the partial order $A \geq B$ on positive definite matrices, the geometric mean presents special difficulties.

The problem for a pair of matrices was solved in the 1970's, and the theory continues to find new uses in matrix analysis and applications. When more than two matrices are involved, a satisfactory definition has been found only recently. In 2005 the Riemannian barycentre of m matrices was proposed as a candidate for their geometric mean. One of its important properties—monotonicity in the m variables—has been established in 2010. Meanwhile the concept has been used in diverse applications such as elasticity, imaging, radar, machine learning etc. We will discuss the problem from the perspective of matrix analysis and operator theory.

Keywords: Geometric mean of positive definite matrices, Riemannian manifold, barycentre, matrix monotonicity.

Use of Matrix Information Theory in Video Surveillance

Manas Kamal Bhuyan

Department of Electronics and Electrical Engineering,
Indian Institute of Technology Guwahati, India-781039.
mkb@iitg.ernet.in
http://www.iitg.ernet.in

The problem of visual inspection of outdoor environments (e.g., airports, railway stations, roads, etc.) has received growing attention in recent times. The work presented in this talk is a component of an extensive research work/project to apply intelligent Closed-Circuit Television (CCTV) to enhance counter-terrorism capability for the protection of mass transport systems. The purpose of intelligent surveillance systems is to automatically perform surveillance tasks by applying cameras in the place of human eyes. Recently, with the development of video hardware such as digital cameras, the video surveillance system is becoming more widely applied and is attracting more researchers to develop fast and robust algorithms. In this talk, the main focus will be given on the basic concept of matrix recovery from corrupted sampled entries i.e., fundamental principle to develop probably correct and efficient algorithms for recovery of low-dimensional linear structure from non-ideal observations. The separation problem (i.e, how to recover original data from the incomplete data) is one of the fundamental problems in Computer Vision. In this talk, some practical applications of Computer Vision like Face Recognition, Handling Occlusions in Target Tracking and Background modelling for video surveillance by using Matrix Information Theory will be highlighted. Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA) are probably the most widely used subspace projection techniques for face recognition. PCA and LDA are useful statistical techniques that has found application in fields such as face/object recognition, and are a common techniques for finding patterns in data of high dimension. One approach to coping with the problem of excessive dimensionality is to reduce the dimensionality by combining features. Linear combinations are particularly attractive because they are simple to compute and analytically tractable. In effect, linear methods project the high-dimensional data onto a lower dimensional space. Principal Component Analysis seeks a projection that best represents the data in a least-squares sense whereas, Linear Discriminant Analysis seeks a projection that best separates the data in a least-squares sense. At one level, PCA and LDA are very different: LDA is a supervised learning technique that relies on class labels, whereas PCA is an unsupervised technique. Nonetheless, in circumstances where class labels are available either technique can be used. One characteristic of both PCA and LDA is that they produce spatially global feature vectors. In other words, the basis vectors produced by PCA and LDA are non-zero for almost all dimensions, implying that a change to a single input pixel will alter every dimension of its subspace projection. The talk will introduce the basics of pattern recognition, with specific emphasis on statistical parametric methods for face recognition and background modelling in a surveillance video. The audience will be introduced to mathematical perspective (matrix information theory) of the problem of pattern recognition for feature discriminatory and classification techniques.

Keywords: Face Recognition, Principal Component Analysis, Linear Discriminant Analysis, Background Modelling.

References

1. Bhuyan, M.K., Kong Suyu., Sanderson, C., Lovell, Brian C., Tracking of Persons for Video Surveillance of Unattended Environments, IEEE international Conference on Pattern Recognition (ICPR), (2008), 1-4.
2. Zhao, W. *et.al.* Face Recognition: A Literature Survey, ACM Computing Surveys, (2003), 399-458.
3. Duda, Richard O., Hart, Peter E. Pattern Classification, 2nd Edition, John Wiley and Sons, Inc. (2006).

Geometric means of fixed rank positive semi-definite matrices.

Silvère Bonnabel, Anne Collard, and Rodolphe Sepulchre

Mines-ParisTech
60 bd Saint-Michel 75272 Paris Cedex 06
University of Liège (Ulg)
B-4000 Liège, Belgium
{silvere.bonnabel}@mines-paristech.fr,{anne.collard,r.sepulchre}@ulg.ac.be

Much research has been devoted to the matrix generalization of the geometric mean $\sqrt{ab}$ of two positive numbers a and b to positive definite matrices (see for instance Chapter 4 in [1] for an expository and insightful treatment of the subject). The further extension of a geometric mean from two to an arbitrary number of matrices is an active current research area [2, 3]. It has been increasingly recognized that from a theoretical point of view as well as in numerous applications, matrix geometric means are to be preferred to their arithmetic counterparts for developing a calculus in the cone of positive definite matrices.

The fundamental and axiomatic approach of Ando [2] reserves the adjective "geometric" to a definition of mean that enjoys the following properties:

(P1) Consistency with scalars: if A, B commute $M(A,B) = (AB)^{1/2}$.
(P2) Joint homogeneity $M(\alpha A, \beta B) = (\alpha\beta)^{1/2} M(A,B)$.
(P3) Permutation invariance $M(A,B) = M(B,A)$.
(P4) Monotonicity. If $A \leq A_0$ (i.e. $(A_0 - A)$ is a positive matrix) and $B \leq B_0$, the means are comparable and verify $M(A,B) \leq M(A_0, B_0)$.
(P5) Continuity from above. If $\{A_n\}$ and $\{B_n\}$ are monotonic decreasing sequence (in the Lowner matrix ordering) converging to A, B then $\lim(A_n \circ B_n) = M(A,B)$.
(P6) Congruence invariance. For any $G \in \mathrm{Gl(n)}$ we have $M(GAG^T, GBG^T) = GM(A,B)G^T$.
(P7) Self-duality $M(A,B)^{-1} = M(A^{-1}, B^{-1})$.

The most famous geometric mean is probably the "Ando geometric mean": $M(A,B) = A\#B = A^{1/2}(A^{-1/2}BA^{-1/2})^{1/2}A^{1/2}$. There are many equivalent definitions of the Ando geometric mean. Notably, it coincides with the Riemannian mean when the set of positive definite matrices is equipped with the Fisher metric provided by information geometry (positive definite matrices are then viewed as covariances of zero mean Gaussian random vectors).

The present work seeks to extend any geometric mean defined on the open cone of positive definite matrices $\mathrm{P_n}$ to the the set of positive semi-definite matrices of fixed rank p, denoted by $S_+(p,n)$. Our motivation is primarily computational: with the growing use of low-rank approximations of matrices as a way to retain tractability in large-scale applications, there is a need to extend the calculus of positive definite matrices to their low-rank counterparts.

The classical approach in the literature is to extend the definition of a mean from the interior of the cone to the boundary of the cone by a continuity argument. As a consequence, this topic has not received much attention. This approach has however serious limitations from a computational viewpoint because it is not rank-preserving. For instance the Ando's geometric mean of two semi-definite positive matrices of rank $p < n/2$ is almost surely **null**. Thus in applications involving low-rank matrices, the use of such an extension is moot.

In this talk we will show how to turn any geometric mean of an arbitrary number of positive definite matrices into a rank-preserving geometric mean well-defined on the set of positive semi-definite matrices. The proposed rank-preserving geometric mean is rooted in a Riemannian geometry recently studied in [4], and which can be viewed as an extension of the Fisher metric to positive semi-definite matrices. We will also discuss some filtering applications.

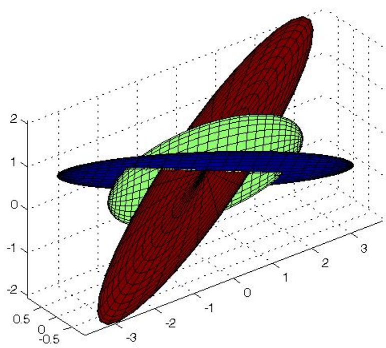

Fig. 1. Proposed mean in $S^{+}(2,3)$. Matrices are viewed as flat ellipsoids. The computation of the mean ellipsoid (in green) decouples into the computation of a mean subspace of the ranges, and a geometric mean in $\mathrm{P_p}$.

References

1. R. Bhatia, *Positive definite matrices*, Princeton University Press, Princeton, NJ, 2006.
2. T. Ando, C.K. Li, and R. Mathias. Geometric means. *Linear Algebra Appl.*, 385:305–334, 2004.
3. R. Bhatia and J. Holbrook. Riemannian geometry and matrix geometric means. *Linear Algebra Appl.*, 413:594–618, 2006
4. S. Bonnabel and R. Sepulchre. Riemannian metric and geometric mean for positive semidefinite matrices of fixed rank. *SIAM J. Matrix Anal. Appl.*, 31(3):1055–1070, 2009.

Koszul-Vinberg cohomology in Information geometry.

Paul Mirabeau Byande and Michel Nguiffo Boyom

Universit Montpellier2, Place Eugne bataillon,
34095 Montpellier Cedex 5, cc 051
paulbyande@yahoo.fr
boyom@math.univ-montp2.fr

Abstract. The main concern of this work is to relate some homological algebra formalism to some topics in statistical theory. The goal is the study of relationships between geometric invariants and statistical estimator in statistical models.

Keywords: Koszul-Vinberg algebras, Maurer-Cartan polynomial, affine structures, statistical model, Fisher metric, α-connections.

1 Introduction

A statistical model for a measurable set $\mathcal{X}$ is a family $S = \{p_\theta = p(., \theta), \quad \theta \in \Theta\}$ of probability distributions on $\mathcal{X}$ where Θ is an open submanifold of an euclidian space $\mathbb{R}^n$ and the map $\theta \in \Theta \to p_\theta \in \mathbb{R}^{\mathcal{X}}$ is injective and smooth. We are interested in the group G of transformation which preserve the family S under sufficient statistics.
In 1945 C.Rao [5] defined the so called Fisher (riemannian) metric, namely

$$g_{ij}(\theta) = \int \frac{\partial}{\partial\theta_i} ln(p_\theta(x)) \frac{\partial}{\partial\theta_j} ln(p_\theta(x)) p_\theta(x) dx.$$

He showed that the group G is a subgroup of isometries of (Θ, g_{ij}). The pair (Θ, g_{ij}) is called the information geometry associated to S. He had written on the importance of analysing statistical model from the perspective of Riemannian geometry.
After this, there were a variety of efforts to pursue this line of research, but unfortunately very few of these results which related directly back to statistical problems [1]. For more than twenty years, the Fisher metric has been the only known geometric invariant of a statistical model.
In 1972 Chentsov [2] pointed out that G-invariant linear connections are closely related to Fisher metric. These connexions are called α-connections, $\alpha \in \mathbb{R}$. Let $f : X \to \mathbb{R}, \quad x \in X$ and $\theta \in \Theta$ we denote E_θ the expectation with respect to the distribution p_θ,

$$E_\theta(f) = \int f(x) p_\theta(x) dx, \quad l_\theta(x) = ln(p_\theta(x)).$$

Consider the functions $\Gamma_{ij,k}^{(\alpha)}$ which maps point each θ to the following value:

$$(\Gamma_{ij,k}^{(\alpha)})_\theta = E_\theta[(\partial_i\partial_j l_\theta + \frac{1-\alpha}{2}\partial_i l_\theta \partial_j l_\theta)(\partial_k l_\theta)]$$

where α is some arbitrary real number. The α-connexion is a connexion $\nabla^{(\alpha)}$ with

$$g(\nabla_{\partial_i}^{(\alpha)}\partial_j, \partial_k) = \Gamma_{ij,k}^{(\alpha)}.$$

In general, G-invariants structure in a statistical model S play a crucial role when analysing the relationship between statistics/probability theory and the structure formed by introducing a metric and a connection on S. The Fisher metric and the α-connections are uniquely characterized by this invariance property. It is well known that locally flat connections correspond to affine structures[3].
We plan to investigate G-invariant affine structures in an information geometry. The main statement to be proved here works in statistical model S admitting coordinate functions yelding an efficient estimator.

Theorem 11 *Let S be a model statistical admitting coordinates functions yelding an efficient estimator.*

– *If the Fisher metric is flat then all of the α-connections are locally flat.*
– *If the Fisher metric is not flat then there are exactly two flat α-connections with $\alpha \in \{-1, 1\}$.*

The proof of the theorem makes case of the Koszul-Vinberg cohomology [4]. More generally it is to be noticed that the use of Koszul-Vinberg cohomology is relevant in studying the developpement geometric aspect of the Jauge theory.

Acknowledgments. Paul Mirabeau Byande is supported by AUF/Grant: PC-420/2460/247/BAC 2007.

References

1. Amari S. and Nagaoka H., Methods of Information Geometry . Translation of Mathetmatical Monogragphs. AMS- OXFORD, vol 191.
2. N.N. Chentsov, statistical decision rules and optimal inference , *Trans. of Math. Mono. vol. 53, AMS, Rhodes Island, NY, 1972.*
3. S. Kobayashi et K. Nomizu, *Foundations of differential Geometry,* Vols 1 et 2 Wiley-Interscience Publishers, New york et London, 1969.
4. Nguiffo Boyom M. The cohomology of Koszul-Vinberg algebras, *Pacific J. Math. Vol 225,N1 (2006) 119-153.*
5. C.R. Rao, Information and accuracy attainable in the estimation of statistical parameters, *Bull. of the Cal. Math. Soc., 37:81-91, 1945*

Applications of Information Geometry to Audio Signal Processing

Arnaud Dessein⋆ and Arshia Cont

IRCAM, CNRS UMR 9912
1 place Stravinsky, 75004 Paris, France
{arnaud.dessein,arshia.cont}@ircam.fr

In this talk, we present some applications of information geometry to audio signal processing. We seek a comprehensive framework that allows to quantify, process and represent the information contained in audio signals. In digital audio, a sound signal is generally encoded as a waveform, and a common problematic is to extract relevant information about the signal by computing sound features from this waveform. A key issue in this context is then to bridge the gap between the raw signal or low-level features (e.g. attack time, frequency content), and the symbolic properties or high-level features (e.g. speaker, instrument, music genre).

We address this issue by employing the theoretical framework of information geometry. In general terms, information geometry is a field of mathematics that studies the notions of probability and of information by the way of differential geometry [1]. The main idea is to analyze the geometrical structure of differential manifold owned by certain families of probability distributions which form a statistical manifold. We aim to investigate the intrinsic geometry of families of probability distributions that represent audio signals, and to manipulate informative entities of sounds within this geometry.

We focus on the statistical manifolds related to exponential families. Exponential families are parametric families of probability distributions that encompass most of the distributions commonly used in statistical learning. Moreover, exponential families equipped with the dual exponential and mixture affine connections possess two dual affine coordinate systems, respectively the natural and the expectation parameters. The underlying dually flat geometry exhibits a strong Hessian dualistic structure, induced by a twice differentiable convex function, called potential, together with its Legendre-Fenchel conjugate. This geometry generalizes the standard self-dual Euclidean geometry, with two dual Bregman divergences instead of the self-dual Euclidean distance, as well as dual geodesics, a generalized Pythagorean theorem and dual projections.

However, the Bregman divergences are generalized distances that are not symmetric and do not verify the triangular inequality in general. From a computational viewpoint, several machine learning algorithms that rely on strong metric properties possessed by the Euclidean distance are therefore not suitable anymore. Yet, recent works have proposed to generalize some of these algorithms to the case of exponential families and of their associated Bregman

⋆ Part of this work was completed while the author was visiting the Japanese-French Laboratory for Informatics, Tokyo, Japan.

divergences [2–6]. It is thus possible, with a single generic implementation, to consider numerous and widely used statistical models or divergences, in algorithms such as centroid computation and hard clustering (k-means), parameter estimation and soft clustering (expectation-maximization), proximity queries in ball trees (nearest-neighbors search, range search).

We discuss the use of this powerful computational framework for applications in audio. The general paradigm is the following. The audio signal is first represented with sound features. We then model these features with probability distributions and apply the tools of information geometry onto these distributions. In particular, it allows to redefine the notion of similarity between two signals in an information setting by employing the canonical divergence of the underlying statistical manifold. This paradigm has been recently investigated for audio data mining in [7]. We show in particular how to segment audio streams into quasi-stationary chunks that form consistent informative entities. These entities can then be treated as symbols for applications such as music similarity analysis, musical structure discovery, query by similarity, audio recombination by concatenative synthesis, and computer-assisted improvisation.

Acknowledgments. This work was supported by a doctoral fellowship from the UPMC (EDITE) and by a grant from the JST-CNRS ICT (Improving the VR Experience).

References

1. Amari, S.-i. and Nagaoka, H.: Methods of Information Geometry. American Mathematical Society, Translations of Mathematical Monographs, Vol. 191 (2000)
2. Banerjee, A., Merugu, S., Dhillon, I. S., Ghosh, J.: Clustering with Bregman Divergences. Journal of Machine Learning Research. 6, 1705–1749 (2005)
3. Cayton, L.: Fast Nearest Neighbor Retrieval for Bregman Divergences. In: 25th International Conference on Machine Learning. Helsinki, Finland (2008)
4. Cayton, L.: Efficient Bregman Range Search. Advances in Neural Information Processing Systems. Curran Associates, Inc. 22, 243–251 (2009)
5. Garcia, V., Nielsen, F., Nock, R.: Levels of Details for Gaussian Mixture Models. In: 9th Asian Conference on Computer Vision, pp. 514–525. Xi'an, China (2009)
6. Nielsen, F., Nock, R.: Sided and Symmetrized Bregman Centroids. IEEE Transactions on Information Theory. 55(6), 2882–2904 (2009)
7. Cont A., Dubnov S., and Assayag G.: On the Information Geometry of Audio Streams with Applications to Similarity Computing. IEEE Transactions on Audio, Speech and Language Processing. 19(1), to appear (2011)

H/α Unsupervised Classification for Highly Textured Polinsar Images Using Information Geometry of Covariance Matrices

P. Formont[1,2⋆], J.P. Ovarlez[1,2], F. Pascal[1], G. Vasile[3], and L. Ferro-Famil[4]

[1] SONDRA Research Alliance
Plateau du Moulon, 3 rue Joliot-Curie, 91192 Gif-sur-Yvette, France
Tel: +33 169 851 803 - Fax: +33 169 851 809
pierre.formont@supelec.fr, frederic.pascal@supelec.fr
[2] French Aerospace Lab
ONERA DEMR/TSI - Chemin de la Hunière, 91761 Palaiseau, France
Tel: +33 180 386 313 - Fax: +33 +33 180 386 269
jean-philippe.ovarlez@onera.fr
[3] Grenoble-Image-sPeech-Signal-Automatics Lab, CNRS
GIPSA-lab DIS/SIGMAPHY, Grenoble INP - BP 46, 38402 Grenoble, France Tel: +33 476 826 334
gabriel.vasile@gipsa-lab.grenoble-inp.fr
[4] IETR, Image and Remote Sensing Group
SAPHIR Team, University of Rennes 1, Rennes - France
laurent.ferro-famil@univ-rennes1.fr

The recently launched POLSAR systems are now capable of producing high quality polarimetric SAR images of the Earth surface under meter resolution. The additional polarimetric information allows the discrimination of different scattering mechanisms. In [1] was introduced the entropy-alpha-anisotropy (H/α/A) classification based on the eigenvalues of the polarimetric covariance matrix (CM). This CM is usually estimated, under homogeneous and Gaussian assumptions, with the well known Sample Covariance Matrix (SCM) which is Wishart distributed. Based on this decomposition, the unsupervised classification of the SAR images can be performed by an iterative algorithm based on complex Wishart density function. It uses the H/α decomposition results to get an initial segmentation into eight clusters, then the K-means clustering is implemented by considering the polarimetric CM as the feature vectors. This technique needs however to derive by a classical Euclidian mean operation the averaged CM of each center of class and to compute by Wishart distance the miminal distance between each pixel CM and with all the class centers.

The decrease of the resolution cell offers the opportunity to observe much thinner spatial features than the decametric resolution of the up-to-now available SAR images but also lead to more complicated effects like spatial heterogeneity. Hence, some areas usually considered as random backscattering mechanisms can become punctual deterministic backscattering mechanisms. The usual techniques of classification, detection, speckle filtering, used for decametric resolution, have to be adapted to these new challenging problems. For high resolution SAR im-

⋆ The author would like to thank the DGA for funding this research.

ages, recent studies have shown that the spatial heterogeneity of the observed scene leads to non-Gaussian clutter modelling. One commonly used fully polarimetric non-Gaussian clutter model is the compound Gaussian model: the spatial heterogeneity of the SAR image intensity is taken into account by modelling the polarimetric clutter information m-vector $\mathbf{k}$ (in the monostatic case, $m = 3$) as a SIRV (Spherically Invariant Random Vector), i.e. the product between the square root of a scalar random variable τ (texture) and an independent, zero mean, complex circular Gaussian random vector $\mathbf{z}$ (speckle) and characterized by an unknown Covariance Matrix $\mathbf{M}$:

$$\mathbf{k} = \sqrt{\tau}\,\mathbf{z}, \tag{1}$$

in this model, the variable τ can represent the spatial variation of the intensity of the wave vector $\mathbf{k}$ from pixel to pixel. All the polarimetric information (phase relationships within the wave vector) is so contained only in the covariance matrix $\mathbf{M}$. Relatively to a given pixel (equivalently to a given τ), the wave vector is then Gaussian.

In homogeneous and Gaussian clutter assumption, the texture τ is assumed to be constant and to be the same for all the pixels. In that case, the statistic of the secondary data is Gaussian and the covariance matrix (called the Sample Covariance Matrix) can be estimated by the Maximum Likelihood (ML) Theory with a set of N secondary data $\mathbf{k}_i$, $i \in [1, N]$ as:

$$\widehat{M}_{SCM} = \frac{1}{N} \sum_{i=1}^{N} \mathbf{k}_i\,\mathbf{k}_i^H \tag{2}$$

In the SIRV model, the covariance matrix is generally an unknown parameter which can be estimated from Maximum Likelihood (ML) Theory. In [2], Gini et al. derived the ML estimate $\widehat{\mathbf{M}}_{FP}$ of the covariance matrix $\mathbf{M}$ for deterministic texture, which is the solution of the following equation:

$$\widehat{\mathbf{M}}_{FP} = f(\mathbf{M}) = \frac{m}{N} \sum_{i=1}^{N} \frac{\mathbf{k}_i\,\mathbf{k}_i^H}{\mathbf{k}_i^H\,\widehat{\mathbf{M}}_{FP}^{-1}\,\mathbf{k}_i}, \tag{3}$$

This approach has been used in [3] by Conte et al. to derive a recursive algorithm for estimating the solution matrix $\mathbf{M}_{FP}$ called the Fixed Point Covariance matrix. This algorithm consists in computing the Fixed Point of f using the sequence defined by $\mathbf{M}_{i+1} = f(\mathbf{M}_i)$ and $\mathbf{M}_0 = \mathbf{I}$. It has been shown in [2] and [3] that the estimation scheme from (2), developed under the deterministic texture case, yields also an Approximate ML (AML) estimator under stochastic texture hypothesis. This study has been completed by the work of Pascal et al. [4], which recently established the existence and the uniqueness of the Fixed Point estimator of the normalized covariance matrix, as well as the convergence of the recursive algorithm whatever the initialization.

The aim of this proposed paper is twofold. Firstly, we propose in this paper to briefly recall original results obtained recently in [5] for the joint Maximum

Likelihood estimation of the texture and the polarimetric CM. These results, based on the above Fixed Point CM estimate, allowed to derive a new distance for SIRV CM and to propose a new technique of speckle filtering (PWF) in heterogeneous environment. Secondly, we introduce a metric-based mean for the space of positive-definite Hermitian covariance matrices. An emerging theory [6–8] allows to take into account the fact that Euclidian space can not describe the space of positive-definite Hermitian CM. Rigorously, the averaged covariance matrix $\mathbf{M}_{\omega_l}$ (SCM or Fixed Point) of a H/α/A cluster l can not be computed with the Euclidean metric, i.e. usual arithmetic mean as:

$$\mathbf{M}_{\omega_l} = \frac{1}{K}\sum_{k=1}^{K}\mathbf{M}_k^l \tag{4}$$

where $\mathbf{M}_k^l$, $k \in [1, K]$ are the K covariances matrices of all pixels belonging to the class ω_l in the H/α plane. It is well known that after few iterations of the unsupervised classification, all the centers of class move significantly within the H/α plane leading a more difficult physical interpretation to the final classification. The mean associated with the Riemannian metric corresponds to the geometric mean:

$$\mathbf{M}_{\omega_l} = \underset{\mathbf{M}_\omega \in \mathcal{P}(m)}{\arg\min} \sum_{k=1}^{K} \left|\left|\log\left(\mathbf{M}_\omega\, {\mathbf{M}_k^l}^{-1}\right)\right|\right|_F^2 \tag{5}$$

where $||.||_F$ stands for the Frobenius norm and $\mathcal{P}(m)$ is the set of the Hermitian definite-positive covariance matrices of size m. The solution can easily be found using a simple gradient algorithm.

We discuss further in the paper the use of the Riemannian mean and we use differential geometric tools to give a characterization of this mean. We can show that the centers of class will remain more stable during the iteration process, leading to a different interpretation of the H/α/A classification. This technique can be applied both on classical SCM and on Fixed Point CM. Used jointly with the Fixed Point CM estimate, this technique can give nice results when dealing with high resolution and highly textured polarimetric SAR images classification.

References

1. S. R. Cloude and E. Pottier, "An entropy based classification scheme for land applications of polarimetric SAR", IEEE Trans. on GRS, Vol. 35, No. 1, pp.68-78, Jan. 1997.
2. F. Gini and M. V. Greco, Covariance Matrix Estimation for CFAR Detection in Correlated Heavy Tailed Clutter," Signal Processing, Vol.82, No.12, pp.1847-1859, 2002.
3. E. Conte, A. DeMaio, and G. Ricci, "Recursive Estimation of the Covariance Matrix of a Compound-Gaussian Process and its Application to Adaptive CFAR Detection," IEEE Trans. on Image Processing, Vol.50, No.8, pp.1908-1915, 2002.
4. F. Pascal, Y. Chitour, J.P. Ovarlez, P. Forster and P. Larzabal, "Covariance Structure Maximum Likelihood Estimates in Compound Gaussian Noise: Existence and Algorithm Analysis", *IEEE Trans. on SP*, Vol. 56, No. 1, pp. 34-48, Jan. 2008.

5. G. Vasile, J.P. Ovarlez, F. Pascal and C. Tison, "Coherency Matrix Estimation of Heterogeneous Clutter in High Resolution Polarimetric SAR Images", IEEE Trans. on GRS, Vol.48, No.4, pp.1809-1826, April 2010.
6. V. Devlaminck and P. Terrier, "Geodesic distance on non-singular coherency matrix space in polarization optics", J. Opt. Soc. Am. A, Vol.27, No.8, August 2010
7. M. Moakher, "Differential Geometric Approach to the Geometric Mean of Symmetric Positive-Definite Matrices", SIAM J. Matrix Anal. Appl., Vol.26, No.3, pp.735-747, 2005
8. F. Barbaresco, "Innovative Tools for Radar Signal Processing Based on Cartan's Geometry of SPD Matrices and Information Geometry", IEEE International Radar Conference, Rome, May 2008

Derivatives of Matrix Functions and their Norms

Priyanka Grover

Theoretical Statistics and Mathematics Unit,
Indian Statistical Institute, Delhi Centre
7, S.J.S. Sansanwal Marg
New Delhi-110016, India
pgrover8r@isid.ac.in

Perturbation or error bounds of various functions have been of great interest for a long time. Mean value theorem and Taylor's theorem come handy for this purpose. While the former is useful in estimating $\|f(A+X) - f(A)\|$ in terms of $\|X\|$ and requires the norms of the first derivative of the function, the latter is useful in computing higher order perturbation bounds and needs norms of the higher order derivatives of the function.

In this talk, we shall discuss derivatives of all orders of some well known functions of matrices (determinant, permanent, tensor powers, symmetric and antisymmetric tensor powers etc.). We shall also give norms of these derivatives and derive their perturbation bounds using the above tools.

Keywords: Derivatives, norms, perturbation bound, determinant, permanent, tensor power, antisymmetric tensor power, symmetric tensor power

On the Computational Instability of the Matrizant Analysis of the Coupled Waveguides and a Solution to Overcome

Trinath Karr

The BuG Design, Plot No. 17, Jagamohan Nagar, Bhubaneswar 751030, India
kar@mecheng.iisc.ernet.in

Abstract. Wave coupling is a phenomenon between multiple interconnecting domains that have one or more physical parameters getting exchanged at their interface. Aero-acoustic wave coupling between interacting domains is due to the vibrating fluid columns at the interface. It is due to the independently sustainable pressure fields of the individual acoustic domains across the finite valued interface impedance. An alternate model of solution to such systems begins with the representative matrix of the set of differential equations, where the matrix has the differential coefficients as its constituent elements.[1, 2] A disordered matrix, representing a set of coupled waveguides, i.e., acoustic streamlines, is a common prospect of the analysis. It yields numerical instability during application of boundary conditions leading to inaccurate prediction of performances and, mathematically, functional discontinuities. From the application point of view, such systems are always evaluated by the relative strength of its physical parameters across the streamline with the help of a set of boundary conditions. The philosophy of virtual (pseudo) boundary conditions at the intermediate nodes of the vector to contain the matrix elements and thus the consequent numerical instabilities have been discussed herewith. Physical applications with its historical development and numerous applied insights have been provided for the visualization of its theoretical contexts. [3, 4] So also, it verifies the effectiveness of simplified algorithms/numerical schemes that can effectively replace its complex counterparts in present day practices.[5, 6]

1 Introduction

Many of the computational and numerical analysis have working limitations. Some of those limitations are inherent to their mathematical model and its subsequent convergence. On the other hand, for some, though they are theoretically stable and convergence is not an issue, there are practical limitations owing to machine limitations. One such case arises from the mere sizes of numbers being generated during matrix analysis. Mathematically, astronomical sizes of the matrix elements are not a subject of concern as long as the normalised/reduced matrix has not been the subject of interest.

In matrix applications, the purpose of the matrix is purely a linear operator that transforms the system variables/vectors across a definite space. If the practicality lies only with few elements of the vector and their relative values, the corresponding operator, in the process, gets reduced upon the application of the suitable boundary conditions. This is one way of normalising the system matrix. Among many practical applications, the wave attributes of phase, impedance or pressure perturbation across the boundary of the traversing domains are of subjective importance. In this work, the case of multiple interacting waveguides those interact through a weak separator has been considered. A weak separator is technically symbolised by a low value of boundary impedance.

2 Concept

In the presented case, the reduction of the matrix by applying the boundary conditions leads to the numerical instabilities. This error is due to the normalisation of the large numerical valued matrix components. The present method works out a simplified scheme that consists of three sub-steps.

1. Splitting the system matrix to several sub-matrices through segmentation.
2. Generating the pseudo boundary conditions (as there exist no boundary conditions physically) at the intermediate nodes of those sub-matrices.
3. Reducing the smaller sub-matrices with those pseudo boundary conditions independently before multiplying the reduced matrices to generate the final transfer matrix.

References

1. E. Dokumaci, “Matrizant approach to acoustic analysis of perforated multiple pipe mufflers carrying mean flow”. Journal of Sound and Vibration, **191**(4): 505-518 (1996).
2. R. A. Frazer, W. J. Duncan and A. R. Collar, *Elementary matrices and Some Applications to Dynamics and Differential Equations*, Cambridge University Press, Cambridge (1952).
3. L. L. Beranek and H. P. Sleeper, “The design and construction of anechoic sound chambers”. Journal of the Acoustical Society of America **18**(1): 140-150 (1946).
4. W. Koidan, G. R. Hruska and M. A. Pickett, “Wedge design for National Bureau of Standards anechoic chamber”. Journal of the Acoustical Society of America, **52**(4): 1071-1076 (1972).
5. V. Easwaran and M. L. Munjal, “Finite element analysis of wedges used in anechoic chambers”. Journal of Sound and Vibration, 160(2): 333-350 (1993).
6. C. N. Wang and M. K. Tang, “Boundary element evaluation on the performance of sound absorbing wedges for anechoic chambers”. Engg. Analysis with Boundary Elements, **18**: 103-110 (1996).

Doppler Information Geometry for Wake Turbulence Monitoring

Zhongxun Liu and Frédéric Barbaresco

Thales Air systems
zhongxun.liu@gmail.com
frederic.barbaresco@thalesgroup.com

Radar Doppler processing is closely related to the robust estimation of covariance matrices. With the methods of information geometry, we consider all the Toeplitz Herimitian Positive Definite covariance matrices of order n as a manifold Ω_n. Here, geometries of covariance matrices based on two kinds of radar data models are presented. For the radar time series modeled by complex circular multivariate Gaussian distribution, the robust distance between two Radar Hermitian SPD (Symmetric Positive Definite) matrices is derived with the theory of information geometry. For the radar time series modeled by a complex autoregressive process, Kahler geometry is introduced and the coordinate of Ω_n is parameterized by the reflection coefficients $G = [P_0\, \mu_1\, \cdots\, \mu_n]$, which are derived from maximum entropy method of Doppler spectral analysis and uniquely determined by the covariance matrix of radar time series. Based on Affine Information Geometry theory, Kahler metric on G is defined by the Hessian of Kahler potential function given by entropy of autoregressive process and the distance between any two autoregressive models is derived. Hence, the Doppler entropy for a radar cell is defined by the distance between regularized maximum order autoregressive model (maximum Doppler lines) and the autoregressive model of order 1 (minimum Doppler lines). Finally, a radar detector based on Doppler entropy assessment is proposed. This advanced Doppler processing chain will be implemented by GPU processing for wake vortex real time monitoring in the airport.

Keywords: Information Geometry, Kahler Geometry, Wake Turbulence, Radar Monitoring

Computational Information Geometry on Matrix Manifolds

Frank Nielsen

École Polytechnique
Sony Computer Science Laboratories, Inc.
Frank.Nielsen@acm.org
www.informationgeometry.org

In this talk, I will illustrate information-theoretically flavored algorithmic methods on the space of symmetric positive definite (SPD) matrices. SPD matrices occur as tensors in many science and engineering fields such as materials science, image analysis, statistics, finance, machine learning, radar, and robotics, just to name a few. The space of SPD matrices can be endowed with a geometric structure in several ways: vector space, Riemannian space (Lie group), and more broadly under the framework of differential information geometry (Finsler space, etc).
We shall focus on the matrix information manifolds implied by a divergence function, representing the dissimilarity measure of matrices. Three classes of generic divergences built on top of a convex contrast function, termed Csiszár, Burbea-Rao, and Bregman divergences, and their interactions will be discussed. We then present basic algorithmic tools for processing and characterizing efficiently finite sets of SPD matrices: center points, clustering, and Voronoi diagrams.

References

1. Marc Arnaudon and Frank Nielsen. Medians and means in Finsler geometry. *CoRR*, abs/1011.6076, 2010.
2. Marc Arnaudon and Frank Nielsen. On Approximating the Riemannian 1-Center. HAL-00560187:1, 2011.
3. Jean-Daniel Boissonnat, Frank Nielsen, and Richard Nock. Bregman Voronoi diagrams. *Discrete and Computational Geometry*, April 2010.
4. Vincent Garcia and Frank Nielsen. Simplification and hierarchical representations of mixtures of exponential families. *Signal Processing (Elsevier)*, 90(12):3197–3212, 2010.
5. Meizhu Liu and Baba C. Vemuri and Shun-ichi Amari and Frank Nielsen. Total Bregman divergence and its applications to shape retrieval. In *International Conference on Computer Vision (CVPR)*, pages 3463–3468, 2010.
6. Frank Nielsen. A family of statistical symmetric divergences based on Jensen's inequality. *CoRR*, abs/1009.4004, 2010.
7. Frank Nielsen and Sylvain Boltz. The Burbea-Rao and Bhattacharyya centroids. *Computing Research Repository (CoRR)*, http://arxiv.org/, April 2010.
8. Frank Nielsen and Richard Nock. On the smallest enclosing information disk. *Inf. Process. Lett.*, 105(3):93–97, 2008.

9. Frank Nielsen and Richard Nock. Quantum Voronoi diagrams and Holevo channel capacity for 1-qubit quantum states. In E.-H. Yang F. Kschischang, editor, *IEEE International Symposium on Information Theory (ISIT)*. IEEE CS Press, 2008.
10. Frank Nielsen and Richard Nock. The dual Voronoi diagrams with respect to representational Bregman divergences. In *International Symposium on Voronoi Diagrams (ISVD)*, pages 71–78, 2009.
11. Frank Nielsen and Richard Nock. Sided and symmetrized Bregman centroids. *IEEE Transactions on Information Theory*, 55(6):2048–2059, June 2009.
12. Frank Nielsen and Richard Nock. Entropies and cross-entropies of exponential families. In *Proceedings of the International Conference on Image Processing (ICIP)*, pages 3621–3624, 2010.
13. Frank Nielsen and Richard Nock. Hyperbolic Voronoi diagrams made easy. In *International Conference on Computational Science and its Applications (ICCSA)*, volume 1, pages 74–80, Los Alamitos, CA, USA, 2010. IEEE Computer Society.
14. Frank Nielsen and Richard Nock. Jensen-Bregman Voronoi diagrams and centroidal tessellations. In *International Symposium on Voronoi Diagrams (ISVD)*, pages 56–65, 2010.
15. Frank Nielsen, Paolo Piro, and Michel Barlaud. Bregman vantage point trees for efficient nearest neighbor queries. In *Proceedings of the 2009 IEEE International Conference on Multimedia and Expo (ICME)*, pages 878–881, 2009.
16. Richard Nock and Frank Nielsen. Fitting the smallest enclosing bregman ball. In *European Conference on Machine Learning (ECML)*, pages 649–656, 2005.
17. Baba Vemuri, Meizhu Liu, Shun ichi Amari, and Frank Nielsen. Total Bregman divergence and its applications to DTI analysis. *IEEE Transactions on Medical Imaging*, 2011. 10.1109/TMI.2010.2086464.

Mining Matrix Data with Bregman Matrix Divergences

Richard Nock

Ceregmia — Université Antilles-Guyane, Schoelcher, France
rnock@martinique.univ-ag.fr

Bregman divergences are a key component of information geometry ([1],[2]), for which it has been formally shown that they shape supervised and unsupervised classification to such a precision that they bring remarkable design capabilities and analyses for algorithms ([11, 9, 10]). Pioneered by earlier works on quantum entropy [12], more recent works have begun to extend Bregman divergences to handle linear operators ([5, 7, 13, 14]). In a first part of this talk, I will review and compare some possible matrix-based extensions of Bregman divergences, two of the main being ([4, 7, 14]):

$$D_\psi(\mathbf{L}||\mathbf{N}) = \mathrm{Tr}\left(\mathbf{D}_\psi\right) + \boldsymbol{\lambda}^\top(\mathbf{I}-\mathbf{Q})\tilde{\boldsymbol{\nu}} \ , \quad (1)$$

$$D_\psi(\mathbf{L}||\mathbf{N}) = \mathrm{Tr}\left(\mathbf{D}_\psi\mathbf{Q}\right) \ , \quad (2)$$

where $\mathbf{Q}$ is a particular doubly stochastic matrix, $\mathbf{D}_\psi$ is a matrix collecting Bregman divergences between eigenspectra, and $\boldsymbol{\lambda}, \tilde{\boldsymbol{\nu}}$ are vectors related to eigenvalues of $\mathbf{L}, \mathbf{N}$ (respectively). Divergences like von Neumann divergences, Umegaki's relative entropy, logdet divergence or Mahalanobis divergences can be brought in such forms. In the meantime, algorithms known as on-line learning algorithms, one of the best examples of a sophisticated mix between (Bregman divergences-based) geometric and algorithmic features, have gradually been tailored ([6, 15]) to crafting portfolios in a theory pioneered by Harry Markowitz more than fifty years ago ([8]). Markowitz narrowed down the traditional expected utility model and assumed that investors only care for mean and variance. The mean-variance portfolio theory was born. Recently, authors have begun to show that the mean-variance theory is in fact a particular case of a mean-divergence theory ([3]), which calls to exponential families of distributions to model a market's behavior (and not simply Gaussians, that fit to Markowitz model), and, hence, to Bregman divergences to model distortions between the market's "natural" and observed behaviors.

In a second part of this talk, I will present the basics of the model, in particular when observed behaviors are matrices. Connection with on-line learning algorithms for matrix data shall also be presented, calling to results and algorithms of ([14, 7]).

References

1. S.-I. Amari and H. Nagaoka. *Methods of Information Geometry.* Oxford University Press, 2000.

2. L. M. Bregman. The relaxation method of finding the common point of convex sets and its application to the solution of problems in convex programming. *USSR Comp. Math. and Math. Phys.*, 7:200–217, 1967.
3. E. Briys, B. Magdalou, and R. Nock. Portfolios, information and geometry: *simplex orbis non sufficit*. In T. Lagoarde-Segot, editor, *Financial Crisis: Thinking Otherwise*. NOVA Eds, New York, 2010.
4. I. Dhillon and J.-A. Tropp. Matrix nearness problems with Bregman divergences. *SIAM Journal on Matrix Analysis and Applications*, 29:1120–1146, 2007.
5. J.-C. Duchi, S. Shalev-Shwartz, Y. Singer, and A. Tewari. Composite objective mirror descent. In *Proc. of the 23 rd COLT*, pages 14–26, 2010.
6. E. Even-Dar, M. Kearns, and J. Wortman. Risk-sensitive online learning. In *17th ALT*, pages 199–213, 2006.
7. B. Kulis, M.-A. Sustik, and I.-S. Dhillon. Low-rank kernel learning with Bregman matrix divergences. *J. of Mach. Learn. Res.*, 10:341–376, 2009.
8. H. Markowitz. Portfolio selection. *Journal of Finance*, 6:77–91, 1952.
9. R. Nock and F. Nielsen. Bregman divergences and surrogates for learning. *IEEE Trans. PAMI*, 31(11):2048–2059, 2009a.
10. R. Nock and F. Nielsen. Intrinsic geometries in learning. In F. Nielsen, editor, *Emerging Trends in Visual Computing*, pages 175–215. Springer Verlag, 2009b.
11. R. Nock and F. Nielsen. On the efficient minimization of classification-calibrated surrogates. In *NIPS*21*, 2008.
12. M. Ohya and D. Petz. *Quantum entropy and its use*. Springer-Verlag, 1993.
13. D. Petz. Bregman divergence as relative operator entropy. *Acta Math. Hungarica*, 116:127–131, 2007.
14. K. Tsuda, G. Rätsch, and M. Warmuth. Matrix exponentiated gradient updates for on-line learning and Bregman projection. *J. of Mach. Learn. Res.*, 6:995–1018, 2005.
15. M. Warmuth and D. Kuzmin. Online variance minimization. In *19 th COLT*, pages 514–528, 2006.

Laplacian matrix of a graph

Sukanta Pati

Department of Mathematics
IIT Guwahati-781039, India.
pati@iitg.ernet.in

Let G be a simple graph on vertices $1, \ldots, n$. The *adjacency matrix* $A(G)$ (or A) of G is a $n \times n$ matrix with (i, j)-th entry $a_{ij} = 1$ if i is adjacent to j ($i \sim j$) and $a_{ij} = 0$ otherwise. The *Laplacian matrix* $L(G)$ (or L) is defined as $D - A(G)$, where D is the diagonal degree matrix of G.

• It is known that $L = QQ^t$, where Q is the vertex-edge incidence matrix. For any vector x we have $x^t L x = \sum\limits_{i \sim j} (x_i - x_j)^2$. The Laplacian matrix is a positive semidefinite matrix. The smallest eigenvalue $\lambda_1(L) = 0$ and $\mathbb{1}$, the vector of all ones, is a corresponding eigenvector.

• Fiedler [6] observed that the second smallest eigenvalue $\lambda_2(L)$ is positive if and only if G is connected. He termed this eigenvalue as the *algebraic connectivity* of G viewing it as an algebraic measure of the connectivity of the graph. The corresponding eigenvectors of L are popularly known as the *Fiedler vectors* of G.

• Many researcher have studied different properties and applications of $L(G)$. We refer the reader to the survey article by Merris [12] and references therein for a general background. It is said [9] that "$\lambda_2(L(G))$ is a measure of the stability and the robustness of the network dynamic system".
We shall note here a few interesting results relating Laplacian matrix with the graph structure.

• [G. Kirchhoff, Anal. Phys. Chem, 1847] Let $L(i|j)$ denote the submatrix obtained from L by deleting row i and column j. Then $(-1)^{i+j} \det L(i|j)$ is the number of spanning trees in G.

• [Fiedler, 1975] Let G be a connected graph. Let Y be a Fiedler vector. Then the subgraph induced by the vertices v in G for which $Y(v) \geq 0$ is connected and the subgraph induced by the vertices v in G for which $Y(v) \leq 0$ is connected.
Let G be a connected graph and Y a Fiedler vector. We call v a *characteristic vertex* if $Y(v) = 0$ and $Y(w) \neq 0$ for some vertex w. We call an edge uv a *characteristic edge* if $Y(u)Y(v) < 0$. The *characteristic set* $C(G, Y)$ is the collection of all characteristic elements.

• [10, 3] Let G be a connected graph and Y a Fiedler vector. Then any two characteristic elements lie on a simple cycle which contains no more characteristic elements. Either $C(G, Y)$ is a single vertex or it is contained in a block B of G. In the first case $C(G, X) = \{v\}$, for any Fiedler vector X and in the second case $C(G, X)$ is contained in B for any Fiedler vector X.

• [7, 11, 3] Let T be a tree and Y a Fiedler vector. Suppose that $C(T, Y) = \{k\}$. Let P be a path that starts from k. Then (i) either $Y(v_i) > 0$, increase and

concave down along P, (ii) or $Y(v_i) < 0$, decrease and concave up along P, (iii) or $Y(v_i) = 0$, along P. A similar statement can be made when $C(T,Y)$ is an edge.
Extremizing algebraic connectivity among graphs with some graph theoretic constraints have drawn a good amount of attention, see for example [5, 8, 13, 14].

• Let T be a tree, k it's characteristic vertex and B a branch at a vertex u which does not contain k (a component of $T - u$ not containing k). Let w be the vertex of B adjacent to u; v be a vertex in T such that $v \sum u$ and $dist(k,v) = dist(k,u) + 1$. Put $T' = T - uw + vw$. Then $\lambda_2(L(T')) \leq \lambda_2(L(T))$. That is, the algebraic connectivity decreases if move the branch B away from the characteristic set.

More general results can be found in [2].
The concept of 'resistance distance' has been studied by many researchers, see for example [1] and the references therein. Let $M = L^+$, the Moore-Penrose inverse of L. The *resistance distance* between i and j is defined as $r(i,j) = m_{ii} + m_{jj} - 2m_{ij}$. Let R denote the matrix with entries $r(i,j)$. We note a few known results below.

• Let G be a connected graph. Then $r(i,j) \leq dist(i,j)$, equality holds if and only if there is a unique i-j-path. Furthermore, $r(i,j) + r(j,k) \geq r(i,k)$ (triangle inequality).

• $R^{-1} = -\frac{1}{2}L + \frac{1}{\tau^t R \tau}\tau\tau^t$, where $\tau_i = 2 - \sum\limits_{j \sim i} r(i,j)$. Generalizes inverse of distance matrix for a tree, as $dist(i,j) = r(i,j)$ holds for trees.

• Let G be connected and $i \sim j$. Denote $k(G)$ the number of spanning trees in G and by $k'(G)$ the number of spanning trees in G containing ij. Then $r(i,j) = \frac{k'(G)}{k(G)}$.

• Let G be connected and λ_i denote the ith smallest Laplacian eigenvalue of G. Then $\sum\limits_i \sum\limits_j r(i,j) = 2\sum\limits_{i=2}^{n} \frac{1}{\lambda_i}$.

References

1. R. B. Bapat. Graphs and Matrices. Universitext, Springer (2010); Also Hindustan Book Agency, TRIM (2010).
2. R. B. Bapat, A. K. Lal, and S. pati. On algebraic connectivity of graphs with at most two points of articulation. Communicated.
3. R. B. Bapat, S. Kirkland, and S. Pati. The Perturbed Laplacian Matrix of a Graph. *Linear and Multilinear Alg.*, 49, 219–242, (2001).
4. T. Biyikoglu, J. Leydold, and P. F. Stadler. *Laplacian eigenvectors of graphs. Perron-Frobenius and Faber-Krahn type theorems.* Lecture Notes in Mathematics, 1915, Springer, Berlin, 2007.
5. S. Fallat and S. Kirkland. Extremizing algebraic connectivity subject to graph theoretic constraints. *Electron. J. of Linear Algebra*, 3, 48-74, (1998).
6. M. Fiedler. Algebraic connectivity of graphs. *Czechoslovak Math. J.*, 23(98), 298-305, (1973).

7. M. Fiedler. A property of eigenvectors of nonnegative symmetric matrices and its application to graph theory. *Czechoslovak Math. J.*, 25, 607-618, (1975).
8. J. M. Guo. A conjecture on the algebraic connectivity of connected graphs with fixed girth. *Discrete Math.*, 308 (23), 5702-5711, (2008).
9. Y. Kim and M. Mesbahi. On maximizing the second smallest eigenvalue of a state dependent graph Laplacian. 2005 American control conference, Portland, USA.
10. S. Kirkland and S. Fallat. Perron components and algebraic connectivity for weighted graphs. *Linear and Multilinear Alg.*, 44, 131-138, (1998).
11. S. Kirkland, M. Neumann and B. L. Shader. Characteristic vertices of weighted trees via Perron values. *Linear and Multilinear Alg.*, 40(4), 311-325, (1996).
12. R. Merris. Laplacian matrices of graphs: a survey. *Linear Alg. Appl.*, 197/198, 143-176, (1994).
13. K. L. Patra, and A. K. Lal. The effect on the algebraic connectivity of a tree by grafting or collapsing of edges. *Linear Algebra Appl.*, 428(4), 855-864, (2008).
14. J. Y. Shao, J. M. Guo, and H. Y. Shan. The ordering of trees and connected graphs by algebraic connectivity. *Linear Alg. Appl.*, 428(7), 1421-1438, (2008).

Geometry and Statistics for Computational Anatomy

Xavier Pennec

INRIA Sophia Antipolis - Asclepios Team, 2004 Rte des Lucioles BP 93
06902 Sophia Antipolis Cedex, France
xavier.pennec@inria.fr

Over the last 30 years, there was an explosion of imaging modalities allows observing both the anatomy *in vivo* and *in situ* at multiple spatial scales (from cells to the whole body), multiple time scales (beating heart, growth, aging, evolution of species), and on multiple subjects. The combination of these new observation means and of the computerized methods is at the heart of *computational anatomy*, an emerging discipline at the interface of geometry, statistics and image analysis which aims at developing algorithms to model and analyze the biological shape of tissues and organs. The goal is to estimate representative organ anatomies across diseases, populations, species or ages, to model the organ development across time (growth or aging), to establish their variability, and to correlate this variability information with other functional, genetic or structural information (e.g. fiber bundles extracted from diffusion tensor images). From an applicative point of view, a first objective is to understand and to model how life is functioning at the population level, for instance by classifying pathologies from structural deviations (taxonomy). A second application objective is to better drive the adaptation of generic models of the anatomy (atlas) into patient-specific data (personalization) in order to help therapy planning (before), control (during) and follow-up (after).
Understanding and modeling the shape of organs is made difficult by the absence of physical models for comparing different subjects, the complexity of shapes, and the high number of degrees of freedom implied. The general method is to identify anatomically representative geometric features (points, tensors, curves, surfaces, volume transformations), and to describe and compare their statistical distribution in different populations. As these geometric features most often belong to manifolds that have no canonical Euclidean structure, we have to rely on more elaborated algorithmic basis. The Riemannian structure proves to be a powerful and consistent framework for computing simple statistics on finite dimensional manifolds [7, 8] and can be extend to a complete computing framework on manifold-valued images [9]. For instance, the choice of a convenient Riemannian metric on the space of positive define symmetric matrices (tensors) allows to generalize consistently to tensor fields many important geometric data processing algorithms such as interpolation, filtering, diffusion and restoration of missing data. This framework is particularly well suited to the statistical estimation of Diffusion Tensor Images [5], and can also be used for modeling the brain variability from sulcal lines drawn at the surface of the cerebral cortex [4].

To move from simple point-wise features to curves, surfaces and deformations, we believe that the embedding vector space of currents provide an interesting computational environment. It was introduced in the field by Glaunes and extended by Durrleman [2, 1] to a generative shape model which combines a random diffeomorphic deformation model a la Grenander & Miller, that encodes the geometric variability of the anatomical template, with a random residual shape variability model (a la Kendall) on the deformed template. We applied the efficient algorithmic toolbox developed for handling statistics on currents to the analysis of the shape of the right ventricle of the heart in a population of Tetralogy of Fallot patients The resulting statistical model of the remodeling of the ventricle during growth turns out to have an anatomically meaningful interpretation [6]. The extensions of this type of methodology to longitudinal evolution estimations in populations is currently one of the most active topic in computational anatomy. We present here a simple model where we combine a static inter-subject change of coordinate system with a time-warp to transform the generic scenario of deformation at the population level to the subject specific longitudinal observations. When applied to different species (here bonobos vs chimpanzees) or to diseases (autism vs control), this model suggests that the change in the speed of evolution might be more important than the shape differences [3].

References

1. S. Durrleman, P. Fillard, X.r Pennec, A. Trouvé, and N. Ayache. Registration, atlas estimation and variability analysis of white matter fiber bundles modeled as currents. *NeuroImage*, 2011. In press.
2. S. Durrleman, X. Pennec, A. Trouvé, and N. Ayache. Statistical models on sets of curves and surfaces based on currents. *Med. Image Analysis*, 13(5):793–808, 2009.
3. S. Durrleman, X. Pennec, A. Trouvé, N. Ayache and J. Braga. Comparison of the endocast growth of chimpanzees and bonobos via temporal regression and spatiotemporal registration. Proc. of *Miccai STIA Workshop*, Beijing, Sep. 2010.
4. P. Fillard, V. Arsigny, X. Pennec, K.M. Hayashi, P.M. Thompson, and N. Ayache. Measuring brain variability by extrapolating sparse tensor fields measured on sulcal lines. *Neuroimage*, 34(2):639–650, January 2007.
5. P. Fillard, X. Pennec, V. Arsigny, and N. Ayache. Clinical DT-MRI estimation, smoothing and fiber tracking with log-euclidean metrics. *IEEE Transactions on Medical Imaging*, 26(11):1472–1482, November 2007.
6. T. Mansi, S. Durrleman, B. Bernhardt, M. Sermesant, H. Delingette, I. Voigt, P. Lurz, A.M Taylor, J. Blanc, Y. Boudjemline, X. Pennec, and N. Ayache. A statistical model of right ventricle in tetralogy of Fallot for prediction of remodelling and therapy planning. In *Proc. of MICCAI'09*, LNCS 5761, pages 214–221, London, UK, Sep. 2009. Springer.
7. X. Pennec. Probabilities and statistics on Riemannian manifolds: Basic tools for geometric measurements. In *Proc. of Nonlinear Signal and Image Processing (NSIP'99)*, vol. 1, p.194–198, June 20-23, Antalya, Turkey, 1999. IEEE-EURASIP.
8. X. Pennec. Intrinsic statistics on Riemannian manifolds: Basic tools for geometric measurements. *J. of Mathematical Imaging and Vision*, 25(1):127–154, July 2006.

9. X. Pennec, P. Fillard, and N. Ayache. A Riemannian framework for tensor computing. *International Journal of Computer Vision*, 66(1):41–66, January 2006.

Fast and Fixed-Point SVD Algorithm for Face and Eye Tracking

Tapan Pradhan Aurobinda Routray and Supratim Gupta

Department of Electrical Engineering,
IIT Kharagpur, Kharagpur, India
tapanpradhan78@hotmail.com,aroutray@iitkgp.ac.in,
sgupta.iitkgp@gmail.com

This article proposes an application of fast fixed-point Singular Value Decomposition (SVD) algorithm for face and eye tracking in real-time. The application is meant for on-board detection and classification of drowsiness and loss of attention of human drivers from the ocular features such as PERCLOS (PERcentage CLOSure of eyes) and saccadic (quick) movement of iris.There is an increasing trend in accidents on road due to drowsiness and loss of attention. Non intrusive drowsiness detection system will be effective for alerting drowsy vehicle drivers. PERCLOS has been a standard parameter to indicate drowsiness. However, recently saccadic movements have been reported to carry vital information during early stages of drowsiness. For measurement and quantification of these parameters we propose eigen space based classification of the eye images so as to determine the state of eyelids (fully open, partially open or closed) and state of iris (left, right, centre) with respect to incoming frames with time.The following issues need to be considered while implementing the system on a real-time embedded platform.
1. Response time and power consumption of the embedded system
2. Vibration and noise problems in the vehicle
3. Variation of illumination level during operation The system is expected to

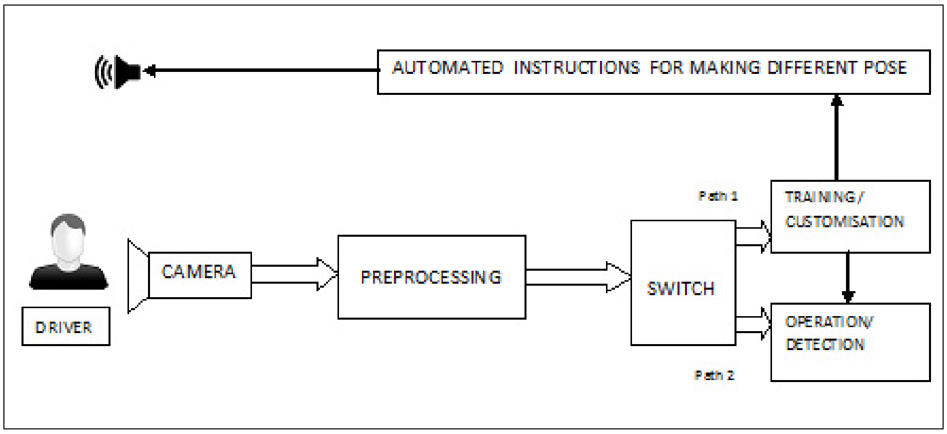

Fig. 1: Block diagram of the proposed system

operate in two phases i.e. *training or customization* and *operation or detection*. While in training phase (Path 1, fig. 1) the subject is expected to position himself in front of the camera as per the automated instructions such that a set of different face and eye positions are acquired. Eigen eye space is created online using fast fixed-point SVD algorithm from the training set. In detection phase (Path 2, fig. 1) test image is projected on to the eigen space and weight vectors obtained are used for detecting face. Similarly we can detect eyes also. The weight vectors of eyes obtained are compared for eye classification. Similar method is applied for iris detection and classification.The following variants of SVD algorithm with floating-point format have been tested for the purpose.
1. Jacobi Method 2. QR Like Algorithm 3. Golub-Kahan-Reinsch Algorithm 4. Tridiagonalization and Symmetric QR Iteration [alternatively Divide-and-Conquer Method or Bisection and Inverse Iteration may be used]Eye detection algorithms have been implemented in single board computer (SBC) and in smart camera NI 1742. Fixed-point implementation of SVD algorithm was carried out in DSP (TMS320C55XX). Eye detection with spectacles needs to be explored further.

Particle Filtering on Riemannian Manifolds. Application to Visual Tracking

Hichem Snoussi and Cédric Richard

Charles Delaunay Institute,
UMR STMR 6279 CNRS , University of Technology of Troyes,
12, rue Marie Curie, 10010, France.
`Hichem.Snoussi@utt.fr`

Recently, a general scheme of particle filtering on Riemannian manifolds has been proposed in [1]. In addition to the nonlinear dynamics, the system state is constrained to lie on a Riemannian manifold $\mathcal{M}$, which dimension is much lower than the whole embedding space dimension. The Riemannian manifold formulation of the state space model avoids the curse of dimensionality from which suffer most of the particle filter methods. Furthermore, this formulation is the only natural tool when the embedding Euclidean space cannot be defined (the state space is defined in an abstract geometric way) or when the constraints are not easily handled (space of positive definite matrices). In order to illustrate the effectiveness of the proposed differential-geometric framework, we consider the problem of visual tracking. The specificity of our modeling is the extension of the hidden state (velocity and position) by jointly estimating the state covariance. As the state covariance is a positive definite matrix, the Euclidean space is not suitable when tracking this covariance. Instead, one should exploit the differential geometric properties of the space of positive definite matrices, by constraining the estimated matrix to move along the geodesics of this Riemannian manifold. The proposed sequential Bayesian updating consists thus in drawing state samples while moving on the manifold geodesics.

1 Particle filtering on Riemannian manifolds

The aim of this section is to propose a general scheme for the extension of the particle filtering method on a Riemannian manifold. The hidden state $\boldsymbol{x}$ is constrained to lie in a Riemannian manifold $(\mathcal{M}, \boldsymbol{g}, \nabla)$ endowed with a Riemannian metric $\boldsymbol{g}$ and an affine connection ∇. The system evolves according to the following nonlinear dynamics:

$$\begin{cases} \boldsymbol{x}_t \sim p_x(\boldsymbol{x}_t \mid \boldsymbol{x}_{t-1}, u_t) \ , \ \boldsymbol{x} \in \mathcal{M} \\ \boldsymbol{y}_t \sim \qquad p_y(\boldsymbol{y}_t \mid \boldsymbol{x}_t, u_t), \end{cases} \tag{1}$$

where the Markov chain (random walk) $p_x(\boldsymbol{x}_t \mid \boldsymbol{x}_{t-1}, u_t)$ on the manifold $\mathcal{M}$ is defined according to the following generating mechanism:

1. Draw a sample $\boldsymbol{v}_t$ on the tangent space $\mathcal{T}_{\boldsymbol{x}_{t-1}}\mathcal{M}$ according to a pdf $p_{\boldsymbol{v}}(.)$.
2. $\boldsymbol{x}$ is obtained by the exponential mapping of $\boldsymbol{v}_t$ according to the affine connection ∇.

In other words, a random vector $\boldsymbol{v}_t$ is drawn on the tangent space $\mathcal{T}_{\boldsymbol{x}_{t-1}}\mathcal{M}$ by the usual Euclidean random technics. Then, the exponential mapping allows the transformation of this vector to a point $\boldsymbol{x}_t$ on the Riemannian manifold. The point $\boldsymbol{x}_t$ is the endpoint of the geodesic starting from $\boldsymbol{x}_{t-1}$ with a random initial velocity vector $\boldsymbol{v}_t$.

As a generating stochastic mechanism is properly defined on the manifold, the particle filtering is naturally extended. It simply consists in propagating the trajectories on the manifold by the random walk process, weighting the particles by the likelihood function and sampling with replacement.

Based on particle trajectories $\{\hat{\boldsymbol{x}}_{0:T}^{(i)}\}$, classical particle filtering algorithm provides a simple way to approximate point estimates. In fact, any quantity of interest $h(\boldsymbol{x})$ can be estimated by its *a posteriori* expectation, minimizing the expected mean square error. The empirical mean of the transformed particles $h(\boldsymbol{x}_t^{(i)})$ represents an unbiased Mont-Carlo estimation of the *a posteriori* expectation. Averaging in the manifold context is no more a valid operation: The empirical mean could be located outside the manifold or the averaging itself does not have a meaning in the absence of a summation operator on the manifold. In order to obtain a valid point estimate, one should rather minimize the mean square error, where the error is evaluated by the geodesic distance $\mathcal{D}$ on the manifold (related to the connection ∇). Following the work of Fréchet [2], the point estimate can be defined by the intrinsic mean (also called Riemannian barycenter). The intrinsic mean has the following expression:

$$\hat{\boldsymbol{x}}_t = \arg\min_{\boldsymbol{x}_t \in \mathcal{M}} \mathrm{E}\big[(\mathcal{D}(\boldsymbol{x}_t, \boldsymbol{s}_t))^2\big] = \arg\min_{\boldsymbol{x}_t \in \mathcal{M}} \int (\mathcal{D}(\boldsymbol{x}_t, \boldsymbol{s}_t))^2 p(\boldsymbol{s}_t \mid y_{1..T}) d_{\boldsymbol{\mu}} \boldsymbol{s}_t$$

where the expectation operator is computed with respect to the *a posteriori* probability density $p(\boldsymbol{s}_t \mid y_{1..T})$ and a dominating measure $d_{\boldsymbol{\mu}}$.

This seminar will illustrate the effectiveness of the geometric filtering framework in visual tracking, where the state to be estimated is the motion affine transformation [4].

References

1. Hichem Snoussi and Ali Mohammad-Djafari, "Particle Filtering on Riemannian Manifolds," in *Bayesian Inference and Maximum Entropy Methods*, Ali Mohammad-Djafari, Ed. MaxEnt Workshops, July 2006, pp. 219–226, Amer. Inst. Physics.
2. M. Fréchet, "Les éléments aléatoires de nature quelconque dans un espace distancié," *Ann. Inst. H. Poincaré*, 1948.
3. H. Karcher, "Riemannian centre of mass and mollifier smoothing," *Comm. Pure Appl. Math*, vol. 30, pp. 509–541, 1977.
4. Yunpeng Liu, Guangwei Li and Zelin Shi. , "Covariance Tracking via Geometric Particle Filtering," *EURASIP Journal on Advances in Signal Processing. Vol. 2010, Article ID 583918, 9 pages, 2010.*

Covariance Matrix Based Spectrum Sensing for Cognitive Radios

Ganapathy Viswanath

Advanced Technology Labs, Honeywell Technology Solutions,
Bangalore 560 037, India.
gviswa@gmail.com

Abstract. Spectrum sensing is one of the most important components in any cognitive radio system that allows usage of the underutilized portions of the radio spectrum. First worldwide standard for cognitive radios to operate in the recently available TV bands is being formulated by IEEE 802.22 Working Group. For this standard to succeed, it is necessary that the presence of TV signals is detected using a reliable sensing mechanism. Spectrum sensing algorithm using statistical covariance and its variants have attracted lot of attention recently. Spectrum sensing algorithms based on statistical covariances ([1]) have advantages over the conventional approach based on energy detection. Covariance based signal detection does not require apriori knowledge of noise power. Spectral covariance of the received signal for signal detection was proposed in [2]. The proposed algorithm exploits statistical correlation of the signal, in particular the pilot signal in frequency domain. Detection performance of this technique shows improved sensitivity compared to other pilot detection algorithms. The properties of the eigenvalue of the covaraince matrix have also been used to detect the presence of radio signal. Random matrix theory has been employed to derive the probability of false alarm and probability of missed detection for eigenvalue based signal sensing. In this paper we propose to discuss different spectrum sensing approaches based on statistical covariance matrix and compare the performance metrics of each approach. Further, we will also suggest techniques to improve the performance of covariance based spectrum sensing approach. Techniques to apply these alogrithms to a wider class of signals will also be discussed in this paper.

1 Introduction

Explosive growth in the number of wireless devices operating in the unlicensed as well licensed bands has resulted in severe shortage of radio spectrum. The multitude of wireless networks and protocols (e.g., Wi-Fi, Bluetooth, WiMax etc.) operating in the unlicensed bands and vying for their share of the spectrum has lead to interference and performance degradation for all the users. However, recent studies by the Federal Communication Commission (FCC) [3] in US and OFCOM have shown that at any given time and in any given geographic locality, less than 10% of the available spectrum in the licensed band is utilized. To exploit these under utilized parts of the spectrum (also referred to as white spaces or spectrum holes), the FCC has advocated development of a new generation of programmable, smart radios that can dynamically access various parts of the spectrum, including the licensed bands. Such radios would operate as secondary users in the licensed bands. These radios are required to possess the capabilities of spectrum usage sensing, environment learning and interference avoidance with the primary users of the licensed spectrum bands while simultaneously ensuring the quality of service (QoS) requirements of both the primary and secondary users. Radios with such capabilities are referred to as cognitive radios (CRs).

2 Signal Model

In a single radio based sensing approach, even the weak signals must be detected to avoid causing interference to primary receivers within its transmission zone. The basic hypothesis problem for transmitter detection is usually formulated as:

$$x(t) = \begin{cases} n(t) & H_0, \\ h \cdot s(t) + n(t) & H_1 \end{cases} \tag{1}$$

where $x(t)$ is the signal received by the cognitive radio, $s(t)$ is the transmitted signal of the primary user, $n(t)$ is the *Additive White Gaussian Noise* (AWGN) and h is the amplitude gain of the channel. For the theoretical analysis we assume h is equal to one for simplicity. H_0 is the *null hypothesis* for the scenario that there is no primary user on the channel. H_1 is the *alternative hypothesis* that there exists a primary user currently transmitting on the channel. In general, by increasing the duration of time (up to a certain extent) for which the test statistics is averaged, the hypothesis can be tested arbitrarily well. The probabilities of interest for spectrum sensing are the probability of detection (P_d), which defines, at hypothesis H_1, the probability of sensing algorithm having detected the presence of the primary signal and the probability of false alarm (P_{fa}), which defines, at hypothesis H_0, the probability of sensing algorithm detecting the presence of the primary signal.
The statistical covariance matrix is computed based on the received signal $x(t)$. The test statistics are derived based on the difference in the correleation properties of the signal and noise.

References

1. Y. Zeng and Y.-C. Liang, "Spectrum-sensing algorithms for cognitive radio based on statistical covariances", *IEEE Trans. Vehicular Technology*, Vol. 58, No. 4, pp. 1804-1815,May 2009.
2. Jaewon Kim and Jeffry G. Andrews, "Sensitive White Space Detection with Spectral Covariance Sensing ", *IEEE Trans. Wireless Communication*, 2010.
3. Federal Communications Commission, Spectrum Policy Task Force, *Report ET docket No.* 02-135, Nov. 2002.

Author Index