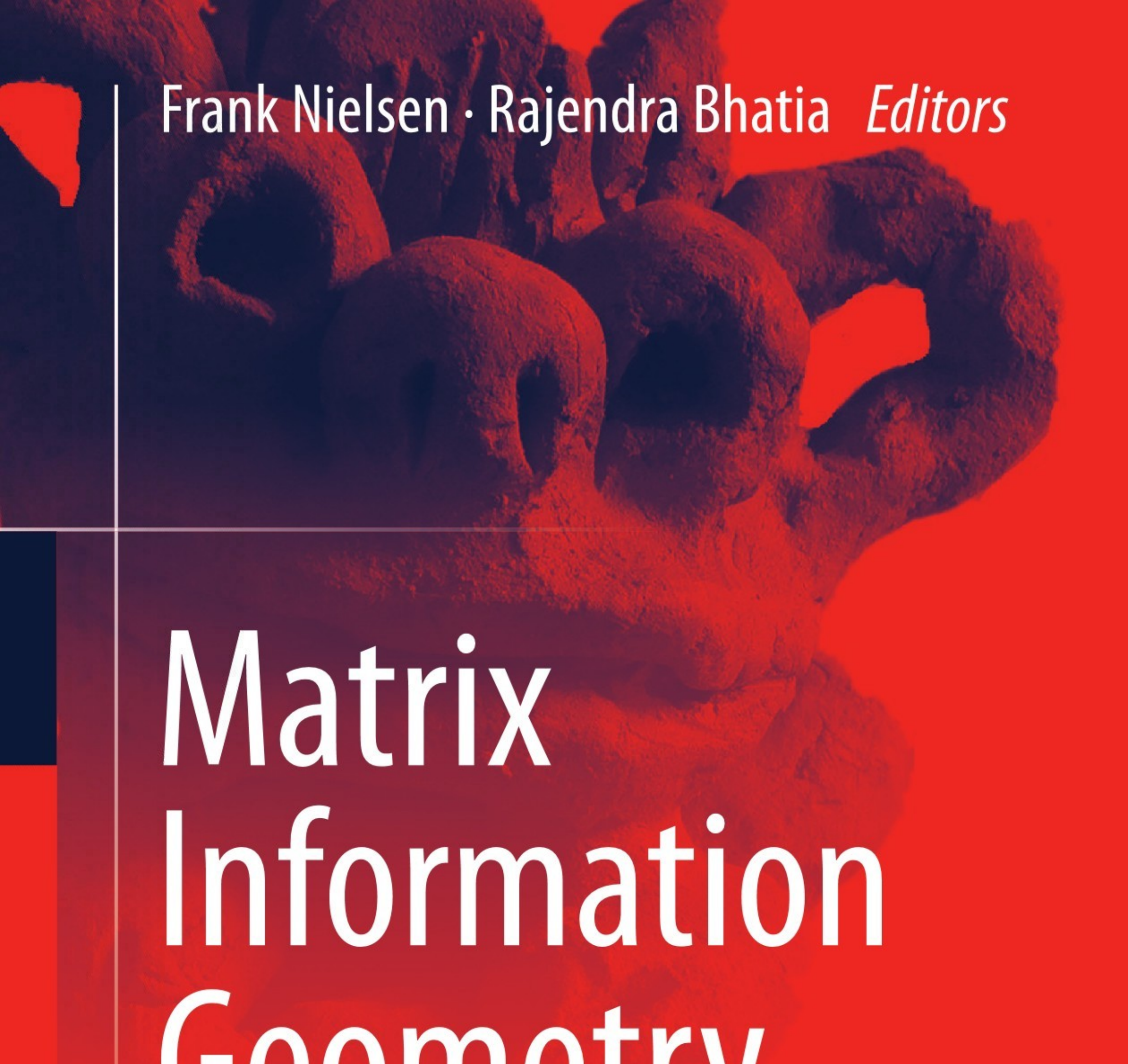

Frank Nielsen · Rajendra Bhatia *Editors*

Matrix Information Geometry

Springer

Matrix Information Geometry

Frank Nielsen · Rajendra Bhatia
Editors

Matrix Information Geometry

Springer

Editors
Frank Nielsen
Higashi Gotanda 3-14-13
Shinagawa-Ku 141-0022
Japan

Rajendra Bhatia
S.J.S. Sansanwal Marg 7
Delhi 110016
India

ISBN 978-3-642-30231-2 ISBN 978-3-642-30232-9 (eBook)
DOI 10.1007/978-3-642-30232-9
Springer Heidelberg New York Dordrecht London

Library of Congress Control Number: 2012941088

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

I would like to dedicate this book to my father

Gudmund Liebach Nielsen
(16.03.1946–25.02.2011)

Preface

1 Welcome to “Matrix Information Geometry”

This book is the outcome of the Indo-French Workshop on “Matrix Information Geometries (MIG): Applications in Sensor and Cognitive Systems Engineering,” which was held at École Polytechnique and Thales Research and Technology Center, Palaiseau, France, in February 23–25, 2011.

The workshop was generously funded mostly by the Indo-French Centre for the Promotion of Advanced Research (IFCPAR). During the event, 22 renowned invited French and Indian speakers gave lectures on their areas of expertise within the field of matrix analysis and processing.

From these speakers, a total of 17 original contributions or state-of-the-art chapters have been prepared in this edited book. All articles were thoroughly peer-reviewed (from 3 to 5 reviewers) and improved according to the suggestions, remarks or comments of the referees.

For the reader’s convenience, the 17 contributions presented in this book are organized into three parts, as follows:

1. State-of-the-art surveys & original matrix theory papers,
2. Advanced matrix theory for radar processing,
3. Matrix-based signal processing applications (computer vision, economics, statistics, etc.)

Further information including the slides of speakers and photos of the event can be found on-line at:

http://www.informationgeometry.org/MIG/

2 Group Photo (24th February 2011)

This photo was taken in the "Cour Ferrié" of École Polytechnique, France

3 Organization

The 17 chapters of the book have been organized into the following three parts:

1. State-of-the-art surveys & original matrix theory work:

- Supremum/infimum and nonlinear averaging of positive definite symmetric matrices (*Jesús Angulo*)
- The Riemannian mean of positive matrices (*Rajendra Bhatia*)
- The geometry of low-rank Kalman filters (*Silvère Bonnabel and Rodolphe Sepulchre*)
- KV cohomology in information geometry (*Michel Nguiffo Boyom and Paul Mirabeau Byande*)
- Derivatives of multilinear functions of matrices (*Priyanka Grover)*
- Jensen divergence-based means of SPD matrices (*Frank Nielsen Meizhu Liu, Baba C. Vemuri*)
- Exponential barycenters of the canonical Cartan connection and invariant means on Lie groups (*Xavier Pennec and Vincent Arsigny*)

2. Advanced matrix theory for radar processing:

- Medians and means in Riemannian geometry: existence, uniqueness and computation (*Marc Arnaudon, Frédéric Barbaresco and Le Yang*)

- Information geometry of covariance matrix: Cartan-Siegel homogeneous bounded domains, Mostow/Berger fibration and Fréchet Median (*Frédéric Barbaresco*)
- On the use of matrix information geometry for polarimetric SAR image classication (*Pierre Formont, Jean-Philippe Ovarlez, and Frédéric Pascal*)
- Doppler information geometry for wake turbulence monitoring (*Zhongxun Liu and Frédéric Barbaresco*)

3. Matrix-based signal processing applications:

- Review of the application of matrix information Theory in Video Surveillance (*M.K. Bhuyan and Malathi.T*)
- Comparative evaluation of symmetric SVD algorithms for real-time face and eye tracking (*Tapan Pradhan, Aurobinda Routray, and Bibek Kabi*)
- Real-time detection of overlapping sound events with non-negative matrix factorization (*Arnaud Dessein, Arshia Cont, Guillaume Lemaitre*)
- Mining matrix data with Bregman matrix divergences for portfolio selection (*Richard Nock, Brice Magdalou, Eric Briys, and Frank Nielsen*)
- Learning mixtures by simplifying kernel density estimators (*Olivier Schwander and Frank Nielsen*)
- Particle filtering on Riemannian manifolds: Application to covariance matrices tracking (*Hichem Snoussi*).

Besides keywords mentioned at the beginning of each chapter, a global index of terms is provided at the end of the book.

4 Sponsors

We gratefully acknowledge the generous financial support of the Indo-French Centre for the Promotion of Advanced Research (IFCPAR/CEFIPRA) and the following sponsor institutions without which we could not have successfully organized this meeting:

- Agence Nationale pour la Recherche (ANR, Contract ANR-07-BLAN-328, GAIA: Computational Information Geometry and its Applications)
- École Polytechnique, and specially the Computer Science Department (LIX) of Ecole Polytechnique
- CEREGMIA, University of Antille-Guyane, Martinique.
- Sony Computer Science Laboratories Inc
- Thales

In particular, we would like to warmly thank Dr. A. Amudeswari, Director of the Indo French Centre for the Promotion of Advanced Research. In addition, we would like to express our deep gratitude to Amit Kumar Mishra (Indian Institute of Technology Guwahati, now a Senior Lecturer at University of Cape Town) who was instrumental in the early stages to kick off the meeting.We gratefully acknowledge the editorial and production staff of Springer-Verlag with special thanks to Dr. Christoph Baumann and Ms. Carmen Wolf.

We would also like to thank Frédéric Barabaresco (Thales), François Le Chevalier (Thales Air Operations), Olivier Schwander (École Polytechnique, LIX), Ms. Corinne Poulain (École Polytechnique, LIX) and Ms. Evelyne Rayssac (École Polytechnique, LIX) for providing us with valuable assistance.

Frank Nielsen (5793b870) expresses his gratitude to Prof. Mario Tokoro and Dr. Hiroaki Kitano, as well as all other members of Sony Computer Science Laboratories, Inc.

CEFIPRA/Ecole Polytechnique/LIX/Sony Computer Science Laboratories/Thales

CEREGMIA, UAG

Agence Nationale de la Recherche: ANR-07-BLAN-328 (GAIA: Computational Information Geometry and its Applications)

It is our hope that this collection of contributed chapters presented in this book will be a valuable resource for researchers working with matrices, and for graduate students. We hope the book will stimulate further research into this fascinating interface of matrices, geometries and applications.

April 2012

Prof. Frank Nielsen
Prof. Rajendra Bhatia

Contents

Part I
State-of-the-Art Surveys and Original Matrix Theory Work

Chapter 1
Supremum/Infimum and Nonlinear Averaging of Positive Definite Symmetric Matrices

Jesús Angulo

1.1 Introduction

Mathematical morphology is a nonlinear image processing methodology originally developed for binary and greyscale images [33]. It is based on the computation of maximum $\bigwedge$ (dilation operator) and minimum $\bigvee$ (erosion operator) in local neighborhoods called structuring elements [36]. That means that the definition of morphological operators needs a partial ordering relationship $\leq$ between the points to be processed. More precisely, for a real valued image $f : E \to \mathbb{R}$, the flat dilation and erosion of image f by structuring element B are defined respectively by

$$\delta_B(f)(\mathbf{x}) = \left\{ f(\mathbf{y}) : f(\mathbf{y}) = \bigwedge_{\mathbf{z}} [f(\mathbf{z})], \mathbf{z} \in B_{\mathbf{x}} \right\}. \tag{1.1}$$

$$\varepsilon_B(f)(\mathbf{x}) = \left\{ f(\mathbf{y}) : f(\mathbf{y}) = \bigvee_{\mathbf{z}} [f(\mathbf{z})], \mathbf{z} \in \check{B}_{\mathbf{x}} \right\}, \tag{1.2}$$

where $B_{\mathbf{x}} \subset E$ is the structuring element centered at point $\mathbf{x} \in E$, and $\check{B}$ is the reflection of structuring element with respect to the origin. Evolved operators are based on dilations and erosions: openings/closings, residues (gradient, top-hats), alternate sequential filters, geodesic operators (opening/closing by reconstruction, levelings). Morphological operators and filters perform noise suppression, contrast image enhancement, structure extraction and multi-scale decomposition, etc. [36].

J. Angulo (✉)
CMM-Centre de Morphologie Mathématique,
Mathématiques et Systèmes, MINES ParisTech,
35, rue Saint Honoré, Cedex, 77305 Fontainebleau, France
e-mail: jesus.angulo@mines-paristech.fr

F. Nielsen and R. Bhatia (eds.), *Matrix Information Geometry*,
DOI: 10.1007/978-3-642-30232-9_1,

Theory of morphological operators has been formulated in the general framework of complete lattices [23]: a complete lattice $(\mathcal{L}, \leq)$ is a partially ordered set $\mathcal{L}$ with order relation $\leq$, a supremum written $\bigvee$, and an infimum written $\bigwedge$, such that every subset of $\mathcal{L}$ has a supremum (smallest upper bound) and an infimum (greatest lower bound). Let $\mathcal{L}$ be a complete lattice. A dilation $\delta : \mathcal{L} \rightarrow \mathcal{L}$ is a mapping commuting with suprema, i.e., $\delta\left(\bigvee_i X_i\right) = \bigvee_i \delta\left(X_i\right)$. An erosion $\varepsilon : \mathcal{L} \rightarrow \mathcal{L}$ commutes with infima, i.e., $\delta\left(\bigwedge_i X_i\right) = \bigwedge_i \delta\left(X_i\right)$. Then the pair (ε, δ) is called an adjunction on $\mathcal{L}$ if for very $X, Y \in \mathcal{L}$, it holds: $\delta(X) \leq Y \Leftrightarrow X \leq \varepsilon(Y)$. Mathematical morphology is also characterized by its domain of invariance in the complete lattice $\mathcal{L}$) of the space of image values. Morphological operators $\psi(\cdot) : \mathcal{L} \rightarrow \mathcal{L}$ commutate with a group of transformations $G(\cdot) : \mathcal{L} \rightarrow \mathcal{L}$ of image values, i.e., for any $f(\mathbf{x}) \in \mathcal{F}(E, \mathcal{L})$ we have $\psi(G(f))(\mathbf{x}) = G(\psi(f))(\mathbf{x})$ or

$$\begin{array}{ccc} f(\mathbf{x}) & \longrightarrow & \psi(f)(\mathbf{x}) \\ \updownarrow & & \updownarrow \\ G(f)(\mathbf{x}) & \longrightarrow & \psi(G(f))(\mathbf{x}) \end{array}$$

Obviously the commutativity of the product $G \circ \psi$ is equivalent to the invariance of the ordering $\leq$ under the transformation $G(\cdot)$. The group of invariant transformations $G(\cdot)$ depends on the physical properties of each particular $\mathcal{L}$, e.g., in gray level images, morphological operators commute with anamorphosis (i.e., $G(\cdot)$ is a strictly increasing mapping).

Dilation and erosion can be also computed using an eikonal PDE [2]:

$$\partial u_t = \pm \parallel \nabla u \parallel, \tag{1.3}$$

with initial conditions $u(x, y, 0) = f(x, y)$. The sign $+$ leads to the dilation and the sign—to an erosion using an isotropic structuring element. Some advantages of the continuous formulation are, on the one hand, the fact that required elements (partial derivatives and Euclidean norm) do not required an ordering and, on the other hand, as other standard methods for numerical solutions of PDEs, the continuous approach allows for sub-pixel accuracy of morphological operators.

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, with the corresponding connection with the Legendre transform [26]. More precisely, the two basic morphological mappings $\mathcal{F}(E, \overline{\mathbb{R}}) \rightarrow \mathcal{F}(E, \overline{\mathbb{R}})$ are given respectively by

$$\delta_b(f)(\mathbf{x}) = (f \oplus b)(\mathbf{x}) = \sup_{\mathbf{h} \in E} \left(f(\mathbf{x} - \mathbf{h}) + b(\mathbf{h})\right), \tag{1.4}$$

and

$$\varepsilon_b(f)(\mathbf{x}) = (f \ominus b)(\mathbf{x}) = \inf_{\mathbf{h} \in E} \left(f(\mathbf{x} + \mathbf{h}) - b(\mathbf{h})\right). \tag{1.5}$$

where the canonical family of structuring functions are the paraboloids $b_a(\mathbf{x}) = -\frac{\|\mathbf{x}\|^2}{2a}$.

Matrix and tensor valued images appear nowadays in various image processing fields and applications [43]:

- Structure tensor images representing the local orientation and edge information [19], which are computed by Gaussian smoothing of the dyadic product $\nabla u \nabla u^T$ of an image $u(x, y)$:

$$G(u)(x, y) = \omega_\sigma * \left(\nabla u(x, y)\nabla u(x, y)^T\right) = \begin{pmatrix} g_{xx}(x, y) & g_{xy}(x, y) \\ g_{xy}(x, y) & g_{yy}(x, y) \end{pmatrix}$$

  where $\nabla u(x, y) = \left(\frac{\partial u(x,y)}{\partial x}, \frac{\partial u(x,y)}{\partial y}\right)^T$ is the 2D spatial intensity gradient and ω_σ stands for a Gaussian smoothing with a standard deviation of σ. Hence, the components of the matrix are $g_{xx}(x, y) = \omega_\sigma * \left(\frac{\partial u(x,y)}{\partial x}\right)^2$, $g_{yy}(x, y) = \omega_\sigma * \left(\frac{\partial u(x,y)}{\partial y}\right)^2$ and $g_{xy}(x, y) = \omega_\sigma * \left(\frac{\partial u(x,y)}{\partial x}\frac{\partial u(x,y)}{\partial y}\right)$.
- Diffusion tensor magnetic resonance imaging (DT-MRI) [10] which describes the diffusive property of water molecules using 3×3 positive semidefinite matrix-field, i.e., image value at each pixel (x, y) is a tensor:

$$D(x, y) = \begin{pmatrix} d_{xx}(x, y) & d_{xy}(x, y) & d_{xz}(x, y) \\ d_{xy}(x, y) & d_{yy}(x, y) & d_{yz}(x, y) \\ d_{xz}(x, y) & d_{yz}(x, y) & d_{zz}(x, y) \end{pmatrix}$$

  where $d_{ii}(x, y)$ describes molecular mobility along each direction i of the space and $d_{ij}(x, y)$ the correlation between directions i and j of the space.
- Covariance matrices in different modalities of radar imaging [8, 9], including matrices of particular structure as the Toeplitz covariance matrices (from reflection coefficients parametrization) [47].

In this chapter we are interested in matrix-valued images considered as a spatial structured matrix field $f(\mathbf{x})$ such that

$$f : E \subset \mathbb{Z}^2,\ \mathbb{Z}^3 \longrightarrow \mathrm{PDS}(n)$$

where E is the support space of pixels and, in particular, we focuss on (real) positive definite symmetric $n \times n$ matrices $\mathrm{PDS}(n)$. The reader interested in positive definite matrices is referred to the excellent monograph [11], which considers issues on functional analysis, harmonic analysis and differential geometry in the manifold of positive definite matrices, and in particular it is explained recent work on the geometric mean of several matrices which will be used in this study.

In order to visualize the $\mathrm{PDS}(n)$ matrices, and operations between them, we consider the classical property which said that a matrix $A \in \mathrm{PDS}(n)$ corresponds to a quadratic form

$$q_A(\mathbf{x}) = \mathbf{x}^{\mathrm{T}} A^{-1}\mathbf{x}, \quad \mathbf{x} \in \mathbb{R}^n.$$

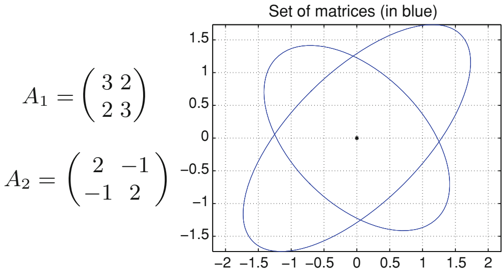

Fig. 1.1 Example of two matrices PDS(2) depicted by their ellipses

Therefore, the matrix A can be represented by the isohypersurface $q_A(\mathbf{x})$, i.e., the ellipsoid $\mathbf{x}^T A^{-1}\mathbf{x} = 1$ centered around 0. Figure 1.1 gives an example for PDS(2). In the context of DTI, the ellipoids have a natural interpretation: if the matrix $A \in \mathrm{PDS}(3)$ represents the diffusivity at a particle, then the ellipsoid encloses the smallest volume within which this particle will be found with some required probability after a short time interval.

The application of classical real-valued morphological operators to vector-valued images such as colour or multispectral images is not straightforward [39, 40]. To consider separately each vector component independently does not generally lead to useful operators [34]. In the framework of matrix-valued spaces, the extension of mathematical morphology to images $f \in \mathcal{F}(E, \mathrm{PDS}(n))$ requires also adapted methods but this extension is neither natural nor unique.

1.1.1 State-of-the-Art

To the best of our knowledge, extension of mathematical morphology to matrix-valued images has been addressed exclusively by Burgeth et al. [15, 16]. They have considered two different approaches. The first one [16] is based on the Löwner partial ordering $\leq^L$: $\forall A, B \in \mathrm{PDS}(n)$, $A \leq^L B \Leftrightarrow B - A \in \mathrm{PDS}(n)$, and where the supremum and infimum of a set of matrices are computed using convex matrix analysis tools (penumbral cones of each matrix, minimal enclosing circle of basis, computation of vertex of associated penumbra matrix). There is a geometrical interpretation viewing the tensors $\mathrm{PDS}(n)$ as ellipsoids: the supremum of a set of tensors is the smallest ellipsoid enclosing the ellipsoids associated to all the tensors; the infimum is the largest ellipsoid which is contained in all the ellipsoids. The second approach [15] corresponds to the generalization of the morphological PDE given in Eq. (1.3) to matrix data: the numerical schema of Osher and Sethian for diffusion equation is generalized to matrices. Both approaches were compared in [15] for various basic morphological operators, mainly for regularization (smoother results for PDE framework than for Löwner ordering) and for edge/details extraction in DT-MRI examples.

Besides the Löwner ordering, there exist a theory on ordering of matrices, which is almost limited to Hermitian nonnegative definite matrices; a recent book [27] on the topic studies in depth this topic. There are three well characterized partial orderings [37, 7, 22]: the Löwner ordering $\leq^L$ (defined above); the minus ordering $A \leq^- B \Leftrightarrow rank(B - A) = rank(B) - rank(A)$; and the star ordering $A \leq^* B \Leftrightarrow A^2 = AB$. They are related between them according to $A \leq^* B \Rightarrow A \leq^- B \Rightarrow A \leq^L B$. It is evident that the minus and the star orderings are two restrictive and consequently without interest for matrix-valued image processing.

As we have just mentioned above, finding the unique smallest enclosing ball of a set of points in a particular space (also known as the minimum enclosing ball or the 1-center problem) is related to the Löwner ordering in the case of PDS(n) matrices. Some recent works in the topic [29, 1] are therefore appropriate for sup/inf computation. In particular, it was introduced in [4] a generic 1-center iterative algorithm for Riemannian geometries, which can be instantiated for example to the case of the manifold of PDS(n) matrices.

From the applications viewpoint, the mean of PDS(n)) matrices is very important in DTI denoising and analysis [41, 25]. However, to our knowledge, the previous theoretical results of mathematical morphology for PDS(n)) matrices [15, 16] have not yet proved their interest for real applications, over and above some illustrative examples from small DTI samples.

1.1.2 Aim of the Study and Chapter Organisation

The goal of this work is to introduce various alternatives ways to extend mathematical morphology to the space PDS(n), which are different from those introduced by Burgeth et al. [15, 16].

More precisely, let $\mathfrak{A} = \{A_i\}_{i=1}^N$ be a finite set of N matrices, where $A_i \in$ PDS(n), we are aiming at computing the supremum $\sup(\mathfrak{A}) = A_\vee$ and the infimum $\inf(\mathfrak{A}) = A_\wedge$ matrices, such that $A_\vee, A_\wedge \in$ PDS(n). As mentioned above, if the operators $\sup(\mathfrak{A})$ and $\inf(\mathfrak{A})$ are defined, dilation and erosion according to Eqs. (1.1) and (1.2) are stated for any image $f \in \mathcal{F}(E, \text{PDS}(n))$ and any structuring element.

Three different families of approaches are explored in the rest of the document.

- Section 1.2 deals with total orderings for sup-inf input-preserving operators. The basic idea 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.

- Section 1.3 discusses partial spectral ordering and inverse eigenvalue problem. By considering as partial ordering the product ordering 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 supremum is not straightforward. We propose two alternatives, the most interesting one based on using as orthogonal basis the one obtained from the geometric mean of the matrices.
- The notion of counter-harmonic mean is introduced in Sect. 1.4 as a nonlinear averaging procedure to calculate pseudo-morphological operators. We have recently shown in [3] how the counter-harmonic mean [14] can be used to introduce nonlinear operators which asymptotically mimic dilation and erosion. It is shown how the extension of counter-harmonic mean to symmetric positive definite matrices is very natural and leads to an efficient operator to robustly estimate the supremum/infimum of a set of matrices.

Application of these supremum/infimum definitions to compute morphological operators on $\mathrm{PDS}(n)$ matrix-valued images is illustrated in Sect. 1.5. The preliminary comparative results are useful to understand the potential interest of nonlinear filtering on matrix-valued images but also to show that there is no universal ordering strategy for all image processing tasks.

Finally, Sect. 1.6 of conclusions and perspectives close the chapter.

1.2 Total Orderings for Sup-Inf Input-Preserving Sets of PDS Matrices

Before introducing total orderings based on lexicographic cascades of spectral invariants as well as on kernelized distances to reference matrices, we start this section by a discussion on the difference between partial and total ordering.

1.2.1 Partial Ordering vs. Total Ordering

We remind that $\leq$ is a partial order (or antisymmetric preorder) over the set of $\mathrm{PDS}(n)$ matrices if for all A, B, and C in $\mathrm{PDS}(n)$, we have that: $A \leq A$ (reflexivity); if $A \leq B$ and $B \leq A$ then $A = B$ (antisymmetry); if $A \leq A$ and $B \leq C$ then $A \leq C$ (transitivity).

For matrices A, B elements of the partially ordered set $\mathrm{PDS}(n)$ according to $\leq$, if $A \leq B$ or $B \leq A$, then A and B are comparable. Otherwise they are incomparable. That involves that using a partial ordering $\leq$ over $\mathrm{PDS}(n)$ the computation of supremum (resp. infimum) of a set of matrices $\mathfrak{A}$ can produce the situation where two matrices which are incomparable are also bigger (smaller) than any of the other matrices of $\mathfrak{A}$.

A typical case of partial ordering for matrices is the one which corresponds to the product order of the matrix components, i.e., the matrix components are taken marginally. For instance for matrices $A, B \in \mathrm{PDS}(2)$ we have

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \leq_{marg} B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} \Leftrightarrow \begin{cases} a_{11} \leq b_{11} \text{ and } a_{12} \leq b_{12} \text{ and} \\ a_{21} \leq b_{21} \text{ and } a_{22} \leq b_{22} \end{cases}$$

The marginal (or componentwise) supremum/infimum associated to the product partial ordering are given respectively by

$$\overset{marg}{\sup}(\mathfrak{A}) = \begin{pmatrix} \bigvee_i a_{11,i} & \bigvee_i a_{12,i} \\ \bigvee_i a_{21,i} & \bigvee_i a_{22,i} \end{pmatrix}$$

and

$$\overset{marg}{\inf}(\mathfrak{A}) = \begin{pmatrix} \bigwedge_i a_{11,i} & \bigwedge_i a_{12,i} \\ \bigwedge_i a_{21,i} & \bigwedge_i a_{22,i} \end{pmatrix}$$

As we can expect, the obtained supremum/infimum can be a new matrix which may not belong to $\mathfrak{A}$: this is known as the "false color" problem in multivariate morphology [35]. Similarly, two different sets of matrices $\mathfrak{A}_1$ and $\mathfrak{A}_2$ can lead to the same supremum/infimum and consequently these subsets will not be comparable between them.

However, the fundamental drawback of the product order of matrices $\leq_{marg}$ applied to $\mathrm{PDS}(n)$ is the fact that it is not guaranteed that $A_\vee$ and $A_\wedge$ belongs to $\mathrm{PDS}(n)$, e.g.,

$$A_1 = \begin{pmatrix} 6 & 3 \\ 3 & 2 \end{pmatrix} \in \mathrm{PDS}(2), \quad A_2 = \begin{pmatrix} 10 & 3 \\ 3 & 1 \end{pmatrix} \in \mathrm{PDS}(2)$$

and the infimum matrix:

$$A_1 \overset{marg}{\bigwedge} A_2 = \begin{pmatrix} 6 & 3 \\ 3 & 1 \end{pmatrix}$$

is symmetric but not positive definite.

A partial order under which every pair of elements is comparable is called a total order (or linear order). A totally ordered set is also called a chain. Hence, we have a total ordering $\leq$ over the set $\mathrm{PDS}(n)$ if for any two different matrices $A \neq B$, we have $A < B$ or $B < A$, i.e., all the $\mathrm{PDS}(n)$ matrices are ordered according to $\leq$.

Besides this practical interest of having comparable elements, the other advantage of total ordering is associated to the following notion.

Definition 1 Given $\mathfrak{A} = \{A_i\}_{i=1}^N$, a finite set of N matrices, where $A_i \in \mathrm{PDS}(n)$, the supremum and infimum are input-preserving iff $A_\vee \in \mathfrak{A}$ and $A_\wedge \in \mathfrak{A}$.

Obviously, any sup/inf input-preserving operator involves necessarily supremum/infimum matrices belonging to the set $\mathrm{PDS}(n)$.

We have now this classical result which can be easily proven.

Proposition 1 *Any total ordering in* $\mathrm{PDS}(n)$ *leads to* sup/inf *input-preserving operators.*

We can now introduce two families of total orderings.

1.2.2 Lexicographic Total Orderings Based on Tensor Invariants

Let us consider that the N matrices of $\mathfrak{A} = \{A_i\}_{i=1}^N$ have been factorized in their spectral form, i.e.,

$$A_i = V_i \Lambda_i V_i^{\mathrm{T}}$$

where Λ_i is the diagonal matrix of ordered eigenvalues

$$\Lambda_i = \mathrm{diag}\left(\lambda_1(A_i), \cdots, \lambda_n(A_i)\right),$$

with $\lambda_1(A_i) \geq \cdots \geq \lambda_n(A_i)$, and $V_i \in SO(n)$ is the orthogonal matrix of eigenvector basis

$$V_i = \left(\overrightarrow{v}_1(A_i), \cdots, \overrightarrow{v}_n(A_i) \right),$$

such that $\| \overrightarrow{v}_1(A_i) \| = 1$ and $\langle \overrightarrow{v}_j(A_i), \overrightarrow{v}_k(A_i) \rangle = 0, \forall j \neq k$. This representation is frequently used in this study.

We introduce the lexicographic spectral partial ordering $\leq_{lex}^0$ as follows.

Definition 2 Let A and B be two $\mathrm{PDS}(n)$ matrices. We define that $A \leq_{lex}^0 B$ if the ordered sequence of the eigenvalues of A is lexicographically smaller or equal to the corresponding sequence of eigenvalues of B, i.e., if there exists an index j, $1 \leq j \leq n$ such that $\lambda_i(A) = \lambda_i(B)$ for all $i < j$, and $\lambda_j(A) < \lambda_j(B)$ if $j \leq n$

To be precise, it is a total ordering for the space of eigenvalues however is only an antisymmetric preorder for $\mathrm{PDS}(n)$. In fact, using their interpretation as ellipsoids, two unequal matrices A and B can have the same shape, given by their eigenvalues but different orientation in the space given by the orthogonal matrix basis. The most natural way to complete the spectral ordering in order to have a total spectral ordering involves fixing a reference orthogonal basis R_0, in such a way that for A and B having the same eigenvalues the biggest is the matrix having an orthogonal basis closer to R_0; this distance should be of course measured in $SO(n)$. An additional question should be taking into account concerning the choice of R_0. If the value of the reference is independent of the image to be morphologically processed involves that a global transformation of the space values will induce a modification of the ordering; for instance a rotation of all the matrix-valued of the image by a change of origin during the acquisition. Consequently, in order to be invariant to the reference R_0, its choice should be intrinsically done from the image values. In particular, we can consider that an useful R_0 corresponds to the mean value of the matrix basis of the image, where the computation of the mean value is done in $SO(n)$.

One of the difficulties of the lexicographic ordering $\leq^0_{lex}$ is the lack of geometric interpretation of the induced supremum and infimum. A more general strategy to define a spectral-based total orderings lies in associating to each PDS(n) matrix a set of (scalar) invariants which have a geometric interpretation. Then, to define a priority between the invariants in order to build a lexicographic ordering according to these invariants. Finally, to complete with additional condition of distance from R_0 to ensure the totality of the ordering.

For instance, given $A \in$ PDS(3), let $(S_1(A), S_2(A), S_3(A))$ be the set of fundamental symmetric polynomials:

- $S_1(A) = \lambda_1(A) + \lambda_2(A) + \lambda_3(A)$ (mean diameter of the ellipsoid),
- $S_2(A) = \lambda_1(A)\lambda_2(A) + \lambda_2(A)\lambda_3(A) + \lambda_1(A)\lambda_3(A)$ (second order relation of diameters),
- $S_3(A) = \lambda_1(A)\lambda_2(A)\lambda_3(A)$ (volume of ellipsoid).

we can define various other orderings, by changing the priorities between these invariants, e.g.,

(i) Priority is given to the mean diameter of the ellipsoid, then to the main eccentricity finally to the volume:

$$A \leq^1_{lex} B \;\Leftrightarrow\; \begin{cases} S_1(A) < S_1(B) \;\text{ or} \\ S_1(A) = S_1(B) \;\text{ and }\; \frac{\lambda_1(A)}{S_1(A)} < \frac{\lambda_1(B)}{S_1(B)} \;\text{ or} \\ S_1(A) = S_1(B) \;\text{ and }\; \frac{\lambda_1(A)}{S_1(A)} = \frac{\lambda_1(B)}{S_1(B)} \;\text{ and }\; S_3(A) \leq S_3(B) \end{cases}$$

(ii) Priority is given to the volume of the ellipsoid, then to the main eccentricity finally to the mean diameter:

$$A \leq^2_{lex} B \;\Leftrightarrow\; \begin{cases} S_3(A) < S_3(B) \;\text{ or} \\ S_3(A) = S_3(B) \;\text{ and }\; \frac{\lambda_1(A)}{S_1(A)} < \frac{\lambda_1(B)}{S_1(B)} \;\text{ or} \\ S_3(A) = S_3(B) \;\text{ and }\; \frac{\lambda_1(A)}{S_1(A)} = \frac{\lambda_1(B)}{S_1(B)} \;\text{ and }\; S_1(A) \leq S_1(B) \end{cases}$$

(iii) Priority is given to the "size" of the ellipsoid, then to the global eccentricity then to the main eccentricity:

$$A \leq^3_{lex} B \Leftrightarrow \begin{cases} \frac{S_3(A)}{S_1(A)} < \frac{S_3(B)}{S_1(B)} \;\text{ or} \\ \frac{S_3(A)}{S_1(A)} = \frac{S_3(B)}{S_1(B)} \;\text{ and }\; \frac{\lambda_1(A)+\lambda_2(A)}{S_1(A)} < \frac{\lambda_1(B)+\lambda_2(B)}{S_1(B)} \;\text{ or} \\ \frac{S_3(A)}{S_1(A)} = \frac{S_3(B)}{S_1(B)} \;\text{ and }\; \frac{\lambda_1(A)+\lambda_2(A)}{S_1(A)} = \frac{\lambda_1(B)+\lambda_2(B)}{S_1(B)} \;\text{ and }\; \frac{\lambda_1(A)}{S_1(A)} \leq \frac{\lambda_1(B)}{S_1(B)} \end{cases}$$

These geometric parameters of the ellipsoids in PDS(n) are often used in DT-MRI [45, 31] (bulk mean diffusivity, isotropy, fractional anisotropy, etc.), therefore the lexicographic orderings yields easy understanding dilation/erosion operators. Other orthogonal tensor invariant as the proposed in [18] can be also considered for the construction of total orderings.

Figure 1.2 provides an example of supremum and infimum computation for a set of 10 PDS(2) matrices using the lexicographic ordering $\leq^0_{lex}$. This result can be compared with the sumpremum and infimum obtained by the product order of matrices $\leq_{marg}$.

From the previous example, we see that, for instance, according to $\leq^0_{lex}$, the matrices are ordered mostly by the first priority in the lexicographic cascade. Generally, it is possible to reduce the "contribution" to the ordering schema of the first considered invariant by a simple quantization of this invariant. Therefore, we can introduce the α-modulus lexicographic ordering $\leq^3_{lex,\alpha}$ as

$$A \leq^3_{lex,\alpha} B \Leftrightarrow \begin{cases} \lfloor \frac{S_3(A)}{\alpha} \rfloor < \lfloor \frac{S_3(B)}{\alpha} \rfloor \text{ or} \\ \lfloor \frac{S_3(A)}{\alpha} \rfloor = \lfloor \frac{S_3(B)}{\alpha} \rfloor \text{ and } \lambda_1(A)/S_1(A) < \lambda_1(B)/S_1(B) \text{ or} \\ \lfloor \frac{S_3(A)}{\alpha} \rfloor = \lfloor \frac{S_3(B)}{\alpha} \rfloor \text{ and } \lambda_1(A)/S_1(A) = \lambda_1(B)/S_1(B) \text{ and } S_1(A) \leq S_1(B) \end{cases}$$

where $\lfloor x \rfloor$ maps to the largest integer not greater than x and where the value of parameter α allows controlling the degree of quantization of the first condition. For this example, the ellipsoids are roughly compared by their volume, and ellipsoid of similar volume are then compared according to their main eccentricities.

We can consider the main properties of the lexicographic-based total orderings.

Proposition 2 *Lexicographic total orderings based on tensor invariants, completed with distance to a reference R_0, have the following properties.*

- *The associated supremum and infimum involve dilation and erosion operators in the sense that the dilation (erosion) commutes with the supremum (infimum) and that the dilation/erosion forms an adjunction.*
- *Since the sumpremum and infimum are input preserving, the dilation and erosion produce symmetric positive definite matrices.*
- *Dilation and erosion are rotationally invariant if the reference R_0 follows the same rotation as the image values.*
- *More generally, dilation and erosion are invariant to any contrast mapping of the matrix image, that is, to any transformation which modifies the eigenvalues values in such a way that ordering is preserved.*

The proofs are relatively straightforward. We can said as conclusion that these orderings yield a totally ordered complete lattice over PDS(n) which is compatible with the general formulation of dilation/erosion and which have good properties of invariance.

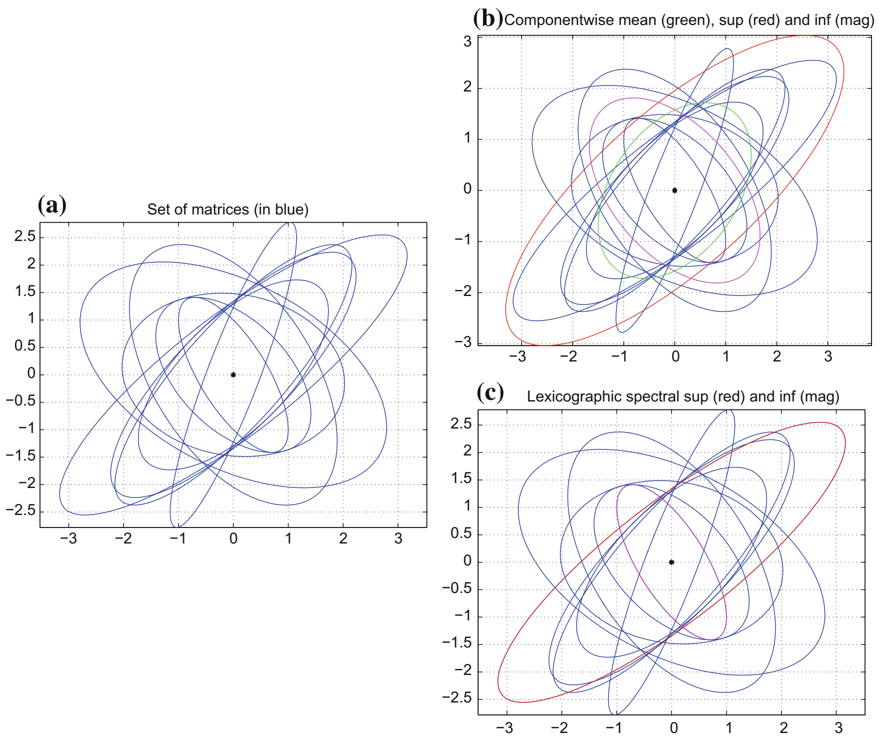

Fig. 1.2 **a** Set $\mathfrak{A}$ of $N = 10$ PDS(2) matrices. **b** Supremum (in *red*) and infimum (in *magenta*) using the product order of matrices $\leq_{marg}$ (componentwise processing); the marginal mean of the matrices is also given in green. **c** Supremum (in *red*) and infimum (in *magenta*) using the lexicographic total ordering $\leq_{lex}^{0}$, which is input-preserving

1.2.3 Lexicographic Total Orderings Based on Prior Sets $(\mathfrak{B}, \mathfrak{F})$

In scalar morphology, the "foreground" is associated to the maximal intensity value $\top$ (i.e., white) and the "background" to the minimal intensity value $\perp$ (i.e., black). The supremum brings towards $\top$ and the infimum towards $\perp$. Using this viewpoint we have recently formulated a general notion of ordering in vector spaces [39] by fixing the references $\top$, $\perp$ and using a supervised learning algorithm. This approach can be naturally extended to non Euclidean spaces such as PDS(n).

Let us consider a training set of I matrices associated to the "foreground" $\mathfrak{F} = \{F_i\}_{i=1}^{I}$ and a training set of J matrices associated to the "background" $\mathfrak{B} = \{B_i\}_{i=1}^{J}$. Let

$$h_{(\mathfrak{B},\mathfrak{F})} : \mathrm{PDS}(n) \rightarrow \mathbb{R}$$

be a surjective mapping, such that for any $A \in \mathrm{PDS}(n)$ we have

$$h_{(\mathfrak{B},\mathfrak{F})}(A) = \left(\sum_{i=1}^{I} K(F_i, A) \right) - \left(\sum_{j=1}^{J} K(B_i, A) \right), \quad (1.6)$$

where the kernel function $K(\cdot, \cdot)$ is a mapping

$$K(R, A) : \mathrm{PDS}(n) \times \mathrm{PDS}(n) \to \mathbb{R}^+ \cup \{0\}$$

Typically, we can consider for instance a radial basis function as kernel

$$K_\alpha(R, A) = e^{-\frac{d(R,A)^2}{\alpha}}, \quad (1.7)$$

where $d(R, A)$ is a distance between the PDS(n) matrices R and A.

Once again, the mapping $h_{(\mathfrak{B},\mathfrak{F})}(\cdot)$ only involves a preorder on the space PDS(n) since two unequal matrices can be mapped on the same real value. The idea to have a complete totally ordered set, i.e., a chain from $\top$ to $\bot$, consists in associating any lexicographic cascade after the computation of the h-mapping.

Definition 3 The lexicographic-completed $(\mathfrak{B}, \mathfrak{F})$-supervised ordering for any pair of matrices A and C is given by

$$A \leq_{sup}^{(\mathfrak{B},\mathfrak{F})} C \Leftrightarrow$$
$$\begin{cases} h_{(\mathfrak{B},\mathfrak{F})}(A) \leq h_{(\mathfrak{B},\mathfrak{F})}(C) & \text{or} \\ h_{(\mathfrak{B},\mathfrak{F})}(A) = h_{(\mathfrak{B},\mathfrak{F})}(C) & \text{and } \{\text{Lexicographic cascade of tensor invariants}\} \end{cases}$$

In practice, we remark that the main ingredient of this kind of approach is the distance between the pair of matrices $d(R, A)$. Many distance and dissimilarity measures have proposed in the literature for the case of DT-MR [31]. We consider that the most useful distances are those which are intrinsically adapted to the geometry of the space PDS(n):

- Riemannian distance. The set of $n \times n$ positive matrices is a differentiable manifold with a natural Riemannian structure (see [11] for a deeper understanding). By integration of its metric over their shortest path on the manifold, given in next section, it is obtained the Riemannian distance for two square positive matrices:

$$d_{Rie}(R, A) = \left| \log \left(R^{-1/2} A R^{-1/2} \right) \right|_F = \left(\sum_{i=1}^{N} \log^2 \lambda_i (R^{-1} A) \right)^{1/2}. \quad (1.8)$$

  This distance is also known as affine-invariant since it is invariant to affine transformation [28].
- Log-Euclidean distance. This notion proposed by [6] coincides with the usual Euclidean (arithmetic) mean in the domain of matrix logarithms:

$$d_{LE}(R, A) = \sqrt{\mathrm{tr}\,(\log(R) - \log(A))^2}. \tag{1.9}$$

We remind that the matrix logarithm $\log(M)$ is defined as the inverse of the matrix exponential $\exp(M) = \sum_{k=0}^{+\infty} M^k/k!$. One should note that for general matrices, neither the uniqueness nor the existence of a logarithm is guaranteed for a given invertible matrix [17]. However, the logarithm of a PDS(n) matrix is well defined and is a symmetric matrix. Distance $d_{LE}(R, A)$ is defined by a Riemannian point of view of a particular vector space structure. Log-Euclidean distance satisfies a number of invariance properties [6]: distance is not changed by inversion (since the inverse of matrices only results in the multiplication by -1 of their logarithms); distance are by construction invariant with respect to any logarithm multiplication (i.e., invariance to any translation in the domain of logarithms); distance is invariant to orthogonal transformation and scaling (but not to any general affine transformation).

Finally, concerning the properties of these total orderings, besides the ones which hold for any total ordering, the invariance properties will depend on the invariance of the chosen distance metric as well as how the training set of matrices $(\mathfrak{B}, \mathfrak{F})$ are selected.

1.3 Partial Spectral Ordering for PDS Matrices and Inverse Eigenvalue Problem

In this section, we continue to use the spectral decomposition of PDS(n) matrices. We start by introducing the spectral product partial ordering $\leq_{sp}$ as follows.

Definition 4 Let A and B be two PDS(n) matrices. We say that $A \leq_{sp} B$ if the ordered sequence of the eigenvalues of A $(\lambda_1(A) \geq \cdots \geq \lambda_n(A) \geq 0)$ is lexicographically smaller or equal to the corresponding sequence of eigenvalues of B, $(\lambda_1(B) \geq \cdots \geq \lambda_n(B) \geq 0)$, i.e., $\lambda_j(A) \leq \lambda_j(B), \forall j = 1, \cdots, n$.

The product ordering $\leq_{sp}$ of eigenvalues does not be confused with their lexicographic ordering $\leq^0_{lex}$. In any case, as we have previously discussed, it is easy to see that $\leq_{sp}$ is only a preorder over PDS(n): the orientation information represented by the eigenvectors is totally ignored (i.e., it does not allow to distinguish between a matrix and rotated version of it).

By using the spectral partial ordering $\leq_{sp}$, the spectral supremum and infimum of a family of matrices $\mathfrak{A} = \{A_i\}_{i=1}^N$ are respectively the matrices

$$A_\vee^{sp} = \sup_{sp}(\mathfrak{A}) = V_\vee \Lambda_\vee V_\vee^{\mathrm{T}}, \tag{1.10}$$

$$A_\wedge^{sp} = \inf_{sp}(\mathfrak{A}) = V_\wedge \Lambda_\wedge V_\wedge^{\mathrm{T}}, \tag{1.11}$$

where the diagonal matrices of the supremum and the infimum are

$$\Lambda_{\vee} = \mathrm{diag}\left(\bigvee_i \lambda_1(A_i), \cdots, \bigvee_i \lambda_n(A_i) \right), \tag{1.12}$$

and

$$\Lambda_{\wedge} = \mathrm{diag}\left(\bigwedge_i \lambda_1(A_i), \cdots, \bigwedge_i \lambda_n(A_i) \right), \tag{1.13}$$

that is, they are obtained as the marginal supremum and infimum of eigenvalues.

Obviously, the question is how to define now supremum/infimum orthogonal basis $V_{\vee}$ and $V_{\wedge}$, which can be interpreted as solving an "inverse eigenvalue problem". In fact, this way of decoupling the shape of the ellipsoids and its orientation have used for instance in [12, 13] for the computation of the distances or geometric means of (fixed) low rank matrices. More precisely, it can be view as a mapping from PDS(n) onto the product space $\mathbb{R}^n \times SO(n)$, where the supremum/infimum on PDS(n) is obtained by an operation on $\mathbb{R}^n$ (the vector space of the eigenvalues) which is simply the marginal vector supremum/infimum and an operation on the space of the eigenvectors $SO(n)$.

We can already remark that in such a case, the supremum/infimum on PDS(n) are not induced by a partial ordering on this space and consequently the operators obtained will not be strictly morphological dilation/erosion.

1.3.1 Spectral Sup/Inf on Geometric Mean Basis

A first alternative is to associate to both $V_{\vee}$ and $V_{\wedge}$ the orthogonal basis of A_{μ}, the matrix mean of $\mathfrak{A}$.

There are different alternatives which have been considered in the literature for computing means of symmetric positive definite matrices [6, 12, 28]. The geometric mean obtained from the Riemannian framework is without any doubt the most interesting. Let us recall the basic elements which can be found in [11]. The Riemannian metric for a matrix A in the manifold PDS(n) is given by the differential

$$ds = \left(\mathrm{tr}(A^{-1} dA \right)^{1/2}. \tag{1.14}$$

The (unique) geodesic between two matrices $A, B \in$ PDS(n) has a parametrization:

$$\gamma(t) = A^{1/2} e^{t \log\left(A^{-1/2} B A^{-1/2}\right)} A^{1/2} = A^{1/2} \left(A^{-1/2} B A^{-1/2} \right)^t A^{1/2}, \tag{1.15}$$

with $t \in [0, 1]$, where $\gamma(0) = A$ and $\gamma(1) = B$. The Riemannian mean between A and B is defined as

$$A \circ B = \gamma(1/2) = A^{1/2} \left(A^{-1/2} B A^{-1/2} \right)^{1/2} A^{1/2}, \tag{1.16}$$

which corresponds to the geometric mean of the matrices, i.e., a symmetrized version of $(AB)^{1/2}$.

The extension of the geometric mean computation of more than two matrices is solved using the notion of Riemannian center, known as Karcher-Frechet barycenter [20, 24]. A fast and efficient algorithm proposed by F. Barbaresco [8, 9] is summarized as follows.

Definition 5 For a set of matrices $\mathfrak{A} = \{A_i\}_{i=1}^N$, the Karcher-Frechet barycenter is computed as $A_\mu(\mathfrak{A}) = X_{k+1}$ such that

$$X_{k+1} = X_k^{1/2} e^{\epsilon \sum_{i=1}^N \log\left(X_k^{-1/2} A_i X_k^{-1/2} \right)} X_k^{1/2}, \tag{1.17}$$

where $\epsilon > 0$ is the step parameter of the gradient descent.

For robustness purposes, it is probably more appropriate to consider the notion of Riemannian median [46, 5].

In summary, the algorithm for supremum matrix $A_\vee^{sp}$:

1. Compute marginal supremum of eigenvalues: $\Lambda_\vee = \mathrm{diag}\left(\bigvee_i \lambda_1(A_i), \cdots, \bigvee_i \lambda_n(A_i) \right)$
2. Compute Karcher-Frechet barycenter: $A_\mu = V_\mu \Lambda_\mu V_\mu^{\mathrm{T}}$
3. Compute inverse spectral matrix: $A_\vee^{sp} = V_\mu \Lambda_\vee V_\mu^{\mathrm{T}}$

Mutatis mutandis $\vee$ by $\wedge$, a similar algorithm is defined for the matrix infimum $A_\wedge^{sp}$.

In Fig. 1.3 is given an example of the supremum/infimum obtained for a set of 10 $\mathrm{PDS}(2)$ matrices: the geometric mean, the supremum and the infimum are ellipsoids with same orientation.

$A_\vee^{sp}$ and $A_\wedge^{sp}$ inherit the properties of the Karcher-Frechet barycenter. This question will be considered in ongoing work. In any case, we insist again that $A_\vee^{sp}$ and $A_\wedge^{sp}$ do not produce dilation/erosion operators since they do not commute with supremum/infimum, i.e., given two sets of $\mathrm{PDS}(n)$ matrices $\mathfrak{A} = \{A_i\}_{i=1}^N$ and $\mathfrak{B} = \{B_j\}_{j=1}^M$ and let $\mathfrak{C} = \mathfrak{A} \cup \mathfrak{B} = \{A_i\}_{i=1}^N \cup \{B_j\}_{j=1}^M$, we have

$$A_\vee^{sp} \bigvee^{sp} B_\vee^{sp} \neq C_\vee^{sp}$$

This is due to the fact that Karcher-Frechet barycenter is not associative, i.e., $A_\mu \circ B_\mu \neq C_\mu$.

1.3.2 Spectral Sup/Inf on Optimized Basis

To complete this section, let us to mention briefly an alternative to tackle the problem of defining the orthogonal basis of the supremum/infimum.

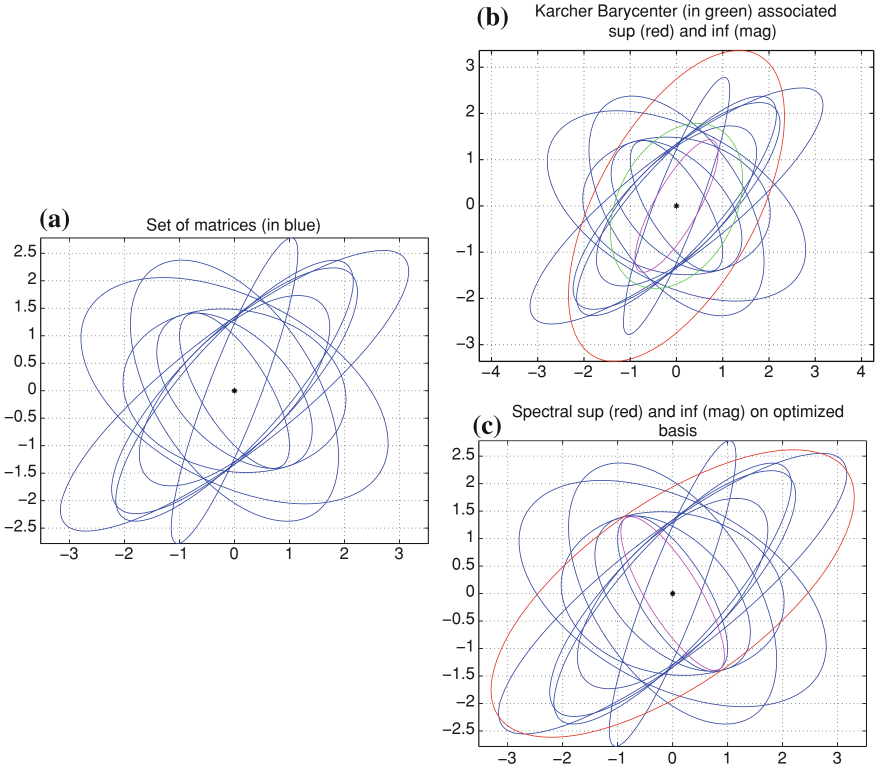

Fig. 1.3 **a** Set $\mathfrak{A}$ of $N = 10$ PDS(2) matrices. **b** spectral sup/inf on geometric mean basis (Karcher-Frechet barycenter computed with parameters $k = 20$ and $\epsilon = 0.1$). **c** spectral sup/inf on optimized basis. The supremum appears in red, the infimum in magenta and the Karcher-Frechet in green

This approach relies on the idea that the largest eigenvalue and corresponding eigenvector should naturally adopted for the matrix $A^{sp}_{\vee}$. Clearly, in the case of PDS(2), the eigenvector basis $V_{\vee}$ is already determined. For general PDS(n), the second eigenvector of $A^{sp}_{\vee}$ can be computed from the given set of matrices by finding the vector lying in the subspace orthogonal to the first eigenvector which is as closer as possible to eigenvector of largest second eigenvalue; and then similarly for the other eigenvectors. Formally, the algorithm to compute the orthogonal basis of supremum: $V_{\vee} = \left(\overrightarrow{v}^{\vee}_{1}, \cdots, \overrightarrow{v}^{\vee}_{n} \right)$ is given as follows.

1. $\overrightarrow{v}^{\vee}_{1} = \overrightarrow{v}_{1}(A_k)$ such that $\lambda_1(A_k) = \bigvee_i \lambda_1(A_i)$;
2. $\overrightarrow{v}^{\vee}_{2} = \overrightarrow{v}$, where $\overrightarrow{v}$ minimizes $\| \overrightarrow{v}_{2}(A_k) - \overrightarrow{v} \|_2$ subject to $\overrightarrow{v}^{\vee}_{1} \perp \overrightarrow{v}$, such that $\lambda_2(A_k) = \bigvee_i \lambda_2(A_i)$;
3. $\overrightarrow{v}^{\vee}_{n-1} = \overrightarrow{v}$, where $\overrightarrow{v}$ minimizes $\| \overrightarrow{v}_{3}(A_k) - \overrightarrow{v} \|_2$ subject to $\overrightarrow{v}^{\vee}_{1} \perp \overrightarrow{v}^{\vee}_{2} \perp \overrightarrow{v}$, such that $\lambda_{n-1}(A_k) = \bigvee_i \lambda_{n-1}(A_i)$.

4. $\overrightarrow{v}_n^{\vee} = \overrightarrow{v}$ such that $\overrightarrow{v}_1^{\vee} \perp \overrightarrow{v}_2^{\vee} \perp \cdots \perp \overrightarrow{v}_{n-1} \perp \overrightarrow{v}$

Mutatis mutandis $\vee$ by $\wedge$, a similar algorithm is defined for the matrix infimum $A_{\wedge}^{sp}$.

An efficient implementation of this algorithm is still an open question, and more important, the properties (existence and uniqueness) of a such orthogonal basis should be also studied in ongoing work.

1.4 Asymptotic Nonlinear Averaging Using Counter-Harmonic Mean for PDS Matrices

We change now our framework and we propose to explore the definition of the supremum/infimum as the asymptotic values of a particular mean which is extended to PDS(n) matrices.

1.4.1 Counter-Harmonic Mean

The counter-harmonic mean (CHM) belongs to the family of the power means [14]. More precisely, the CHM is defined as follows.

Definition 6 Let $\mathbf{a} = (a_1, a_2, \cdots, a_n)$ and $\mathbf{w} = (w_1, w_2, \cdots, w_n)$ be real $n-$tuples, i.e., $\mathbf{a}, \mathbf{w} \in \mathbb{R}^n$. If $P \in \overline{\mathbb{R}}$ then the $P-$th counter-harmonic mean of $\mathbf{a}$ with weight $\mathbf{w}$ is given by [14]

$$\kappa^P(\mathbf{a}; \mathbf{w}) = \begin{cases} \frac{\sum_{i=1}^n w_i a_i^{P+1}}{\sum_{i=1}^n w_i a_i^P} & \text{if } P \in \mathbb{R} \\ \max(a_i) & \text{if } P = +\infty \\ \min(a_i) & \text{if } P = -\infty \end{cases} \tag{1.18}$$

It will be denoted $\kappa^P(\mathbf{a})$ the equal weight case. We notice that $\kappa^0(\mathbf{a}; \mathbf{w})$ is the weighted arithmetic mean and $\kappa^{-1}(\mathbf{a}; \mathbf{w})$ is the weighted harmonic mean.

Used in image processing as a filter, CHM is well suited for reducing the effect of pepper noise for $P > 0$ and of salt noise for $P < 0$ [21]. It is easy to see that for $P \gg 0$ ($P \ll 0$) the pixels with largest (smallest) values in the local neighborhood will dominate the result of the weighted sum. Of course, in practice, the range of P is limited due to the precision in the computation of the floating point operations. In the pioneering paper [38], starting from the natural observation that morphological dilation and erosion are the limit cases of the CHM, it was proposed to use the CHM to calculate robust nonlinear operators which approach the morphological ones but without using max and min operators. In addition, these operators are more robust to outliers (i.e., to noise) and consequently it can be considered as an alternative to rank-based filters in the implementation of pseudo-morphological operators. In our recent study [3] we have also considered empirically how both means converge to the supremum (resp. infimum) when positive P increases (negative P decreases). But let us examine also two properties which are useful to understand the practical interest of the CHM filter.

Proposition 3 *If* $0 \leq P \leq +\infty$ *then* $\kappa^P(\mathbf{a}) \geq \nu^P(\mathbf{a})$*; and if* $-\infty \leq P \leq 0$ *then the following stronger results holds:* $\kappa^P(\mathbf{a}) \leq \nu^{P-1}(\mathbf{a})$*; where* $\nu^P(\mathbf{a}) = \left(\sum_{i=1}^n a_i^P\right)^{1/P}$ *is the* $P-th$ *power-mean filter, or* Minkowski mean of order P, defined for $P \in \mathbb{R}^*$. Inequalities are strict unless $P = 0, +\infty, -\infty$ or if $\mathbf{a}$ is constant.

Proposition 4 *If* $-\infty \leq P \leq Q \leq +\infty$ *then* $\kappa^P(f) \leq \kappa^Q(f)$*, with equality if and only if* $\mathbf{a}$ *is constant.*

Proofs of Propositions 3 and 4 as well as other properties can be found in [14]. Proposition 3 justifies theoretically the suitability of CHM with respect to the alternative approach by high-order Minkowski mean, as considered by Welk [44], in order to propose a nonlinearization of averaging-based filters. We notice that according to Proposition 3, the convergence to the erosion with $P \ll 0$ is faster than to the dilation with equivalent $P \gg 0$, i.e., for $P > 0$

$$|\kappa^P(\mathbf{a}) - \bigvee_i a_i| \geq |\kappa^{-P}(\mathbf{a}) - \bigwedge_i a_i|$$

This asymmetry involves that $\kappa^P(\mathbf{a})$ and $\kappa_B^{-P}(\mathbf{a})$ are not dual operators with respect to the complement, i.e., for $P > 0$

$$\kappa^P(\mathbf{a}) \neq \kappa^{-P}(\complement\mathbf{a})$$

with $\complement\mathbf{a} = -\mathbf{a} = (-a_1, -a_2, \cdots, -a_n)$.

1.4.2 Counter-Harmonic Mean for PDS Matrices

We propose a straightforward generalization of CHM for PDS(n) matrices.

Definition 7 Given $\mathfrak{A} = \{A_i\}_{i=1}^N$, a finite set of N matrices, where $A_i \in \mathrm{PDS}(n)$, the counter-harmonic matrix mean (CHMM) of order P is defined by

$$\kappa^P(\mathfrak{A}) = \left(\sum_{i=1}^N A_i^P\right)^{-1/2} \left(\sum_{i=1}^N A_i^{P+1}\right)\left(\sum_{i=1}^N A_i^P\right)^{-1/2} \tag{1.19}$$

In order to understand the interest of the CHMM, we can study its behavior with respect to P for a numerical example. Let us consider two PDS(2) matrices: