Distances between spectral densities

"The shortest distance between two points is always under construction". (R. McClanahan)

R. Sepulchre -- University of Cambridge

Celebrating Anders Lindquist 75th Birthday

Stockholm, November 2017

Genesis of this talk

Best control scientists in the world gather at KTH

Publicerad 2009-09-17

They have come to celebrate two milestones in the careers of Chris Byrnes and Anders Lindquist. But this international symposium in the field of systems and control is also a tribute to the best control scientists in the world who have gathered in Stockholm for this occasion.

2009: *Spectrum approximation problem* (Byrnes, Georgiou, Lindquist)

Problem 1 (Spectral estimation): Given $\Psi \in \mathcal{S}_+^{m\times m}(\mathbb{T})$ and $\Sigma \in \mathcal{S}_+^{n\times n}$, find $\hat{\Phi}$ that solves

$$\min_{\Phi \in \mathcal{S}_+^{m\times m}(\mathbb{T})} d(\Psi, \Phi)$$

$$\text{s.t.} \int_{-\pi}^{\pi} G(e^{j\vartheta})\Phi(e^{j\vartheta})G^*(e^{j\vartheta})\frac{d\vartheta}{2\pi} = \Sigma$$

where $d: \mathcal{S}_+^{m\times m}(\mathbb{T}) \times \mathcal{S}_+^{m\times m}(\mathbb{T}) \to [0,\infty)$ is a suitable (pseudo-)distance function in the cone $\mathcal{S}_+^{m\times m}(\mathbb{T})$.

Sounds like distance between two systems,

but with a new twist …

Today's talk : based on

arXiv.org > math > arXiv:1708.02818

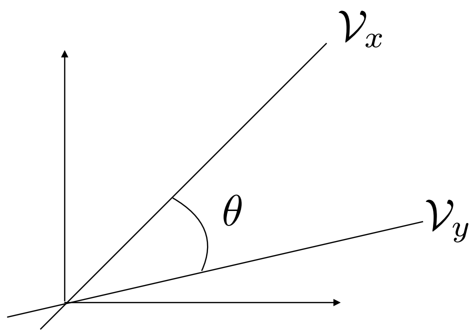
Mathematics > Optimization and Control

Finslerian Metrics in the Cone of Spectral Densities

Giacomo Baggio, Augusto Ferrante, Rodolphe Sepulchre

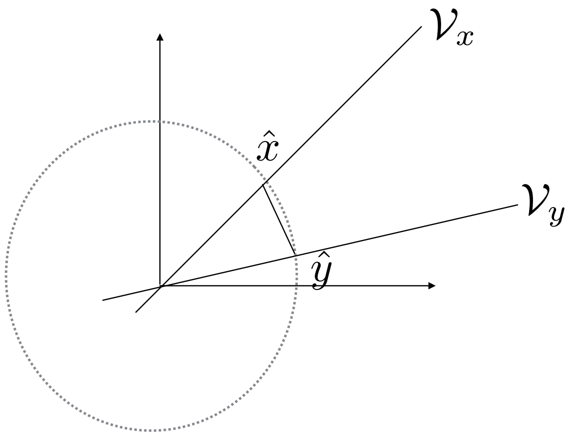
(Submitted on 9 Aug 2017)

The key point : the value of *chordal* distances



Example: how to compute the distance between two lines in the plane ?

The classical answer: chordal distance



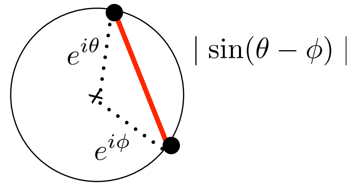
$$\| \hat{x} \|_2 = \| \hat{y} \|_2 = 1$$

identify the lines with unit vectors and compute

$$\| \hat{x} - \hat{y} \|_2$$

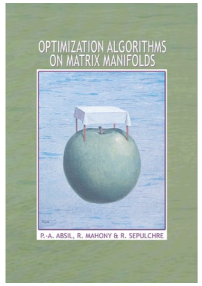
Unitary chordal distances are popular in approximation problems

e.g. planar rotations:



Intrinsic distances between subspaces can be expressed in terms of *principal angles*. *Unitary chordal* distances replace the angles with their sinus. They retain the invariance by *rotation*.

A property at the core of matrix approximation problems in engineering



The gap metric is a chordal distance between LTI systems

The **gap metric** is a distance between subspaces (graphs of pairs (u,y)=Gw).

To be computable, the distance should be a **chordal** distance.

A chordal distance will be invariant by rotation (unitary transformations) provided that the subspaces are images of **unitary** operators.

$$\mathrm{gap}(\mathcal{V}_1, \mathcal{V}_2) = \| P_{\mathcal{V}_1} - P_{\mathcal{V}_2} \|_2$$

Gap metrics are computed via H-infty norms of transfer functions

Proceedings of the 27th Conference on Decision and Control Austin, Texas • December 1988

ON THE COMPUTATION OF THE GAP METRIC*

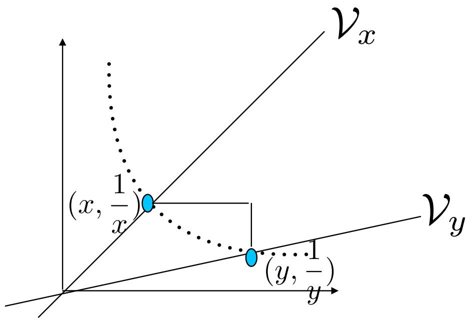
Tryphon T. Georgiou

PROPOSITION 1 The following hold

$$\delta(P_1,P_2) = \max\left\{ \|\mathbf{P}_{(G_1\mathbf{H}_2)^\perp} G_2\|, \|\mathbf{P}_{(G_2\mathbf{H}_2)^\perp} G_1\| \right\} \quad \textbf{(3a)}$$

$$= \max\left\{ \inf_{Q \in \mathbf{H}_\infty} \|\hat{G}_1 - \hat{G}_2 Q\|_\infty, \inf_{Q \in \mathbf{H}_\infty} \|\hat{G}_2 - \hat{G}_1 Q\|_\infty \right\} \quad \textbf{(3b)}$$

An alternative answer: log chordal distance



log chordal $$d(\mathcal{V}_x, \mathcal{V}_y) = \max(\log \frac{x}{y}, \log \frac{y}{x})$$

scale-invariant as opposed to rotation-invariant

What is a good distance between rational spectral densities ?

$$\mathcal{S}_+^{n\times n} := \{\, \Phi \in \mathbb{R}(z)^{n\times n} \;:\; \Phi(e^{j\theta}) \geq 0,\ \theta \in [-\pi, \pi]\}$$

Our hint:

1. Stochastic LTI systems are represented by ***positive*** rather than ***unitary*** operators

2. Computable distances are chordal.

Hence : what is the log-chordal distance between spectral densities ?

Outline

1. Log chordal distances in cones

2. Application to the cone of spectral densities

3. Desirable properties of a distance

4. Comparison with other distances

Log chordal distances in cones

Let $\mathcal{K}$ be a closed, solid, pointed, convex cone defined in a real Banach space $\mathcal{B}$ with norm $\|\cdot\|_{\mathcal{B}}$, that is, a closed subset $\mathcal{K}$ with the properties that: (i) the interior of $\mathcal{K}$, denoted by $\mathring{\mathcal{K}}$, is non-empty, (ii) $\mathcal{K} + \mathcal{K} \subseteq \mathcal{K}$, (iii) $\mathcal{K} \cap -\mathcal{K} = \{0\}$, (iv) $\lambda\mathcal{K} \subseteq \mathcal{K}$ for all $\lambda \geq 0$. The cone $\mathcal{K}$ induces a partial ordering $\leq_{\mathcal{K}}$ on $\mathcal{B}$ by

$$x \leq_{\mathcal{K}} y \iff y - x \in \mathcal{K}.$$

$$M(x,y) := \inf\{\lambda \,:\, x \leq_{\mathcal{K}} \lambda y\}$$

$$d_T(x,y) := \log \max\{M(x,y), M(y,x)\}$$

Thompson (or part) metric (1962)

A close cousin of Hilbert metric

$$m(x,y) := \sup\{\mu \,:\, \mu y \leq_{\mathcal{K}} x\} \qquad d_H(x,y) := \log \frac{M(x,y)}{m(x,y)}$$

Application to the cone of spectral densities

Theorem 1: Consider two full normal rank spectral densities $\Phi_1, \Phi_2 \in \mathcal{S}_{+,\mathrm{rat}}^{n\times n}(\mathbb{T})$ and let $W_1, W_2 \in \mathbb{R}^{n\times n}(z)$ denote the corresponding minimum-phase spectral factors. If $W_2^{-1}W_1$ has no zero/pole on $\mathbb{T}$, then the Hilbert and Thompson metrics between Φ_1 and Φ_2 are given, respectively, by

$$d_H(\Phi_1, \Phi_2) = \log \|W_2^{-1}W_1\|_{\mathcal{H}_\infty}^2 \|W_1^{-1}W_2\|_{\mathcal{H}_\infty}^2,$$

$$d_T(\Phi_1, \Phi_2) = \log \max\left\{\|W_2^{-1}W_1\|_{\mathcal{H}_\infty}^2, \|W_1^{-1}W_2\|_{\mathcal{H}_\infty}^2\right\}.$$

Otherwise, it holds $d_H(\Phi_1, \Phi_2) = d_T(\Phi_1, \Phi_2) = \infty$.

Proof:

$$M(\Phi_1, \Phi_2) = \inf\{\lambda \in \mathbb{R} : \Phi_1(e^{j\vartheta}) \le \lambda \Phi_2(e^{j\vartheta}),\ \vartheta \in [-\pi, \pi]\}$$

$$= \inf\{\lambda \in \mathbb{R} : \Phi_2^{-\frac{1}{2}}(e^{j\vartheta})\Phi_1(e^{j\vartheta})\Phi_2^{-\frac{1}{2}}(e^{j\vartheta}) \le \lambda I_n, \vartheta \in [-\pi, \pi]\}$$

$$= \left\|\Phi_2^{-\frac{1}{2}}\Phi_1\Phi_2^{-\frac{1}{2}}\right\|_{\mathcal{L}_\infty}$$
$$= \|W_2^{-1}\Phi_1 W_2^{-*}\|_{\mathcal{L}_\infty}$$
$$= \|W_2^{-1}W_1W_1^*W_2^{-*}\|_{\mathcal{L}_\infty}$$
$$= \|W_2^{-1}W_1\|_{\mathcal{L}_\infty}^2,$$
$$= \|W_2^{-1}W_1\|_{\mathcal{H}_\infty}^2$$

Desirable properties of a distance

- Computable
- Invariant
- optimisable

“Scale” invariance of the distance

Congruence (or filtering) invariance:

$$d(\Phi_1, \Phi_2) = d(T\Phi_1 T^*, T\Phi_2 T^*)$$

$$\text{for any} \quad T \in \mathbb{R}_*^{n\times n}[z]$$

Using an invariant distance in approximation problems makes the solution unaffected by filtering the data

A source of robustness in modeling !

Invariant properties are the main source of non-euclidean geometries

Differential geometry of log chordal distances

Thompson metric endows the cone with a Finsler manifold structure (similar with Riemannian structure but the norm in the tangent space does not derive from an inner product).

norm $\|v\|_x^T := \inf\{\alpha > 0 : -\alpha x \leq_K v \leq_K \alpha x\}$

length $\ell(\gamma) := \int_a^b \|\gamma'(t)\|_{\gamma(t)}^T \mathrm{d}t.$

'log-chordal' geodesic

$$\varphi(t) = \begin{cases} \left(\frac{\beta^t - \alpha^t}{\beta - \alpha}\right) y + \left(\frac{\beta\alpha^t - \alpha\beta^t}{\beta - \alpha}\right) x, & \text{if } \beta \neq \alpha, \\ \alpha^t x, & \text{if } \beta = \alpha, \end{cases}$$

Outline

1. Log chordal distances in cones
2. Application to the cone of spectral densities
3. Desirable properties of a distance
4. Comparison with other distances

The monovariate case

$$d(\phi_1, \phi_2) = \log(\max(\| \frac{\phi_1}{\phi_2} \|_\infty, \| \frac{\phi_2}{\phi_1} \|_\infty)$$
$$= \max(\| \log \frac{\phi_1}{\phi_2} \|_\infty, \| \log \frac{\phi_2}{\phi_1} \|_\infty)$$

to be compared with

$$d(\phi_1, \phi_2) = \sqrt{\frac{1}{2\pi} \int_{-\pi}^{\pi} (\log \frac{\phi_1}{\phi_2})^2 d\theta - (\frac{1}{2\pi} \int_{-\pi}^{\pi} \log(\frac{\phi_1}{\phi_2}) d\theta)^2}$$ (Georgiou, 2006)

$$d(\phi_1, \phi_2) = \sqrt{\frac{1}{2\pi} \int_{-\pi}^{\pi} (\mathcal{D}^{\frac{1}{2}} \log \frac{\phi_1}{\phi_2})^2 d\theta} \qquad (= \| \mathcal{D}^{\frac{1}{2}} \log \frac{\phi_1}{\phi_2} \|_2)$$

(Martin, 2005)

Scale invariance implies a measure of *distortion*.
The main difference lies in the choice of the two versus infinite norm.

The static case (distance on the SDP cone)

$$d_T(\Phi, I) = \log \max(\lambda_M, \lambda_m^{-1})$$
$$d_H(\Phi, I) = \log \frac{\lambda_M}{\lambda_m}$$

to be compared with

$$d(\Phi, I) = \| \log \Phi \|_F = \sqrt{\sum \log \lambda_i^2}$$

(This is the Fisher-Rao metric, see e.g. Smith 2005)

Invariance implies a logarithmic measure of *spectral* quantities.
The main difference lies in the choice of the two versus infinite norm.

The general case

$$d_T(\Phi_1, \Phi_2) = \log \max \left\{ \|W_2^{-1} W_1\|^2_{\mathcal{H}_\infty}, \|W_1^{-1} W_2\|^2_{\mathcal{H}_\infty} \right\}$$

to be compared with

$$\left(\int_{-\pi}^{\pi} \|\log W_1^{-1} \Phi_2 W_1^{-*}\|_F^2 \frac{d\vartheta}{2\pi} \right)^{1/2}$$ ("two norm" Riemannian analog)

or

$$\|W_2^{-1} W_1\|^2_{\mathcal{H}_2} + \|W_1^{-1} W_2\|^2_{\mathcal{H}_2} - 2n,$$ (divergence measure)

(Jian, Ning, Georgiou 2012)

Thompson metric combines the geometrical properties of the two norm with the computational properties of the divergence measures.

Outline

1. Log chordal distances in cones
2. Application to the cone of spectral densities
3. Desirable properties of a distance
4. Comparison with other distances

Conclusions

1. Thompson metric in the cone of spectral densities enjoys a number of desirable properties
2. The underlying geometry of cones is Finslerian rather than Riemannian
3. A new avenue for distances between systems with a conic representation, e.g. gaussian processes and passive systems

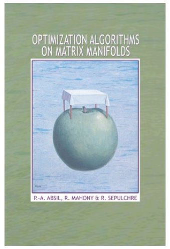
Distances in cones

An overwhelming topic in

Information geometry
Convex analysis
Optimization
Optimal transport
Theory of monotone operators
Differential geometry
…

An overwhelming number of applications in system theory

Covariance matrices
Gaussian distributions
probability vectors
Monotone systems
Consensus theory
Kalman filtering
Spectral estimation
Quantum estimation and control
…



The Magritte picture is perhaps not entirely right here …