The meaning of Distances in Spectral Analysis

46th IEEE Conference on Decision & Control
Plenary presentation

Tryphon Georgiou

Electrical & Computer Engineering
University of Minnesota
Meaning of distances

- maximal separation ($L_\infty$)
- energy-like content ($L_2$)
- integral of flow-rate ($L_1$)
Power spectra

Periodogram, Blackman-Tukey, Levinson, Durbin, Burg, ...
Spectral analysis

\[ u(k) = \int e^{jk\theta} dX(\theta) \]

\[ E\{u(k)u(k + \ell)\} = \int e^{j\ell\theta} f(\theta) d\theta \]
Signals vs. power densities

\((u_1 - u_2)\) “error signal”

\((f_1 - f_2)\) is not a “signal”
Communications

Speech analysis/coding

New Orleans, December 2007
Medical diagnostics

Noninvasive temperature sensing

Temperature field

with E. Ebbini & A.N. Amini

In IEEE Trans. on Biomedical Engineering, 2005

New Orleans, December 2007
Medical diagnostics

Radar (SAR)

http://www.sandia.gov/radar/images/3dsar.gif

New Orleans, December 2007
Quantitative analysis

How can we compare power spectra?

New Orleans, December 2007
Quantitative analysis

How can we compare power spectra?

New Orleans, December 2007
How can we compare power spectra?

Question:
what is a natural notion of distance
between power spectral densities?
Plan of the talk

*Metrics based on*

- prediction theory
- some parallels with information geometry
- transport geometry

*Case studies & applications*
Setting

\[ \ldots u_{-1}, u_0, u_1, u_2, \ldots \]

\[ \ldots u_{-1}, u_0, u_1, u_2, \ldots \]

\[ f_1(\theta) \]

\[ f_2(\theta) \]
What is it we would like to have?

\[
distance(f_1(\theta), f_2(\theta))
\]

- metric
- meaningful & natural candidates?

Kullback-Leibler, Bregman, Itakura-Saito, Makhoul,..

convex functionals
perceptual qualities

New Orleans, December 2007
Linear prediction

One-step-ahead prediction: \( u_{\text{present}} - \hat{u}_{\text{present} | \text{past}} \)

with \( \hat{u}_{\text{present} | \text{past}} := \sum_{\text{past}} \alpha_k u_k \)

\[
E\{|u_{\text{present}} - \hat{u}_{\text{present} | \text{past}}|^2\} = \text{variance of prediction error}
\]
Szegö’s theorem

One-step-ahead prediction:

\[
\text{least error variance} = \exp \left\{ \frac{1}{2\pi} \int \log f(\theta) d\theta \right\}
\]

it is a geometric mean . . .

\[
\exp \left\{ \frac{1}{3} \left( \log f_1 + \log f_2 + \log f_3 \right) \right\} = \sqrt[3]{f_1 f_2 f_3}
\]
Degradation of prediction error variance

Use $f_2$ to design a predictor (assuming $u_{f_2,\text{time}}$).

Then compare how this performs on $u_{f_1,\text{time}}$ against the optimal based on $f_1$.

\[
\frac{E\{|u_{f_1,\text{present}} - \sum_{\text{past}} a_{f_2,\text{past}} u_{f_1,\text{past}}|^2\} - \text{optimal variance}}{\text{optimal variance}} \geq 0
\]
Degradation of prediction variance

\[
\frac{E\left\{ |u_{f_1,\text{present}} - \sum_{\text{past}} a_{f_2,\text{past}} u_{f_1,\text{past}}|^2 \right\}}{\text{optimal variance}} = \frac{\text{arithmetic mean of } \left( \frac{f_1}{f_2} \right)}{\text{geometric mean of } \left( \frac{f_1}{f_2} \right)}
\]

\[
= \frac{\left( \frac{1}{2\pi} \int \left( \frac{f_1}{f_2} \right) d\theta \right)}{\exp \left( \frac{1}{2\pi} \int \log \left( \frac{f_1}{f_2} \right) d\theta \right)}
\]

*arithmetic over geometric* mean (≥ 1)

New Orleans, December 2007
Riemannian metric

\[ f_1 = f, \]
\[ f_2 = f + \Delta \]

\[
E\left\{ |u_{f_1,\text{present}} - \sum_{\text{past}} a_{f_2,\text{past}} u_{f_1,\text{past}}|^2 \right\} - \text{optimal variance}
\]

\[
\delta(f, f + \Delta) = \frac{1}{2\pi} \int \left( \frac{\Delta}{f} \right)^2 d\theta - \left( \frac{1}{2\pi} \int \left( \frac{\Delta}{f} \right) d\theta \right)^2
\]

\text{variance-like: (mean square) - (arithmetic-mean)}^2
Geodesics

Paths $f_r$ ($r \in [0, 1]$) between $f_0, f_1$ of minimal length $\int_0^1 \sqrt{\delta(f_r, f_{r+dr})}$

each point represents a different power spectral density
The geodesics are exponential families:

\[ f_r = f_0 \left( \frac{f_1}{f_0} \right)^r, \quad r \in [0, 1] \]

\[ = \exp \{ (1 - r) \log(f_0) + r \log(f_1) \} \]
Geodesic distance: metric

The path-length is

\[ d(f_0, f_1) := \sqrt{\frac{1}{2\pi} \int_{-\pi}^{\pi} \left( \log \left( \frac{f_1}{f_0} \right) \right)^2 d\theta - \left( \frac{1}{2\pi} \int_{-\pi}^{\pi} \log \left( \frac{f_1}{f_0} \right) d\theta \right)^2} \]

**variance-like distance on logarithms:** (mean square) - (arithmetic-mean)^2

scale-insensitive, “shape” recognizer

\[ \log \frac{f_1}{f_0} = \log(f_1) - \log(f_0) \]
Prediction-based geometry

Information geometry

New Orleans, December 2007
Information geometry – *parallels*

\[ f \sim p : \text{probability density} \]

\[ I = E_p \{ (\partial_\lambda \log p_\lambda)^2 \} \delta \lambda^2 \]

*Fisher information metric*

\[ I = \sum \frac{\Delta^2}{p} \]

New Orleans, December 2007
Information geometry – *parallels*

Expected “message-length increase”:

\[
H(p_1|p_0) = \left(- \sum p_1 \log(p_0)\right) - \left(- \sum p_1 \log(p_1)\right)
\]

*Fisher information metric*

\[
p_0 = p
\]

\[
p_1 = p + \Delta
\]

\[
I = \sum \frac{\Delta^2}{p}
\]

New Orleans, December 2007
Information geometry – \textit{parallels}

\textbf{Geodesics:} great circles

\[ \mathbf{p} \rightarrow \sqrt{\mathbf{p}} \in \text{Sphere} \]

\[ \begin{pmatrix} p(1) \\ p(2) \\ p(3) \end{pmatrix} \rightarrow \begin{pmatrix} \sqrt{p(1)} \\ \sqrt{p(2)} \\ \sqrt{p(3)} \end{pmatrix} \]

\textbf{Geodesic distance:} Arclength

Battacharyya distance

New Orleans, December 2007
Information vs. prediction-based

\[ \sum \frac{\Delta^2}{p} \quad \text{vs.} \quad \int \left( \frac{\Delta}{f} \right)^2 - \left( \int \frac{\Delta}{f} \right)^2 \]

\[ p \mapsto \sqrt{p} \quad \text{vs.} \quad f \mapsto \log f \]

great circles vs. logarithmic families

New Orleans, December 2007
Information geometry – *parallels*

Chentsov’s theorem:

Stochastic maps are contractive under *Fisher metric*

and

*Fisher metric* is the unique Riemannian metric with this property

New Orleans, December 2007
What is the analog for power spectra?

additive noise
\[ f \mapsto f + f_{\text{noise}} \]

multiplicative noise
\[ f \mapsto f \ast f_{\text{noise}} \]

continuity of moments (second-order statistics)
\[ f \mapsto \text{integrals of } f \]
Transport geometry

Monge-Kantorovich problem

minimize cost of transferring mass

\[ \int \text{cost}(x \rightarrow y) \times \text{mass}(dx, dy) \]

New Orleans, December 2007
Transport for power spectra

Transport-based metric

distances do not increase
under additive noise
and multiplicative noise
with power $\leq 1$

$+$ continuity of statistics

$\text{metric} = \min (\text{cost of transport}(\hat{f}_0, \hat{f}_1) + \text{normalization})$

with Johan Karlsson (KTH) & Mir Shahrouz Takyar
Prediction-based geometry

Transport geometry

applications
Least squares: The theory of motion of heavenly bodies, Gauss, K.F.
Tracking with geodesics

with Xianhua Jiang
Voice & sounds

John Weissmuller’s MGM Tarzan Yell

http://www.complxmind.com
Images & more

Geometric active contours

\[ \frac{\partial}{\partial t} \text{Curve} = \nabla_{\text{Curve \ metric}} (f_{\text{inside}}, f_{\text{outside}}) \]

with Romeil Sandhu and Allen Tannenbaum

New Orleans, December 2007
Images & more

with Romeil Sandhu and Allen Tannenbaum

New Orleans, December 2007
Concluding thoughts

Metrics in spectral analysis

- Operational significance
- Effect of natural transformations
Thank you for your attention

thanks to
Xianhua Jiang  Johan Karlsson  Romeil Sandhu  Mir Shahrouz Takyar

Allen Tannenbaum & Anders Lindquist

National Science Foundation, AFOSR, and Hermes-Luh endowment

New Orleans, December 2007