

Geometric Methods for Spectral Analysis

**A DISSERTATION
SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL
OF THE UNIVERSITY OF MINNESOTA
BY**

Xianhua Jiang

**IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
Doctor of Philosophy**

Professor Tryphon T. Georgiou

November, 2011

© Xianhua Jiang 2011
ALL RIGHTS RESERVED

Acknowledgements

First and foremost, I want to thank my advisor Professor Tryphon T. Georgiou. It has been such an honor to work with him, a great mind whose knowledge and intellectual sharpness marvels me all the time. I appreciate the excellent example he provided as one who really enjoys learning and conducts research with a sincere heart. I am grateful not only for his guidance on my research and study, but also his encouragement and confidence in me, which is especially invaluable during my difficult times. This thesis would not have been possible without his guidance and help. This is also an opportunity for me to acknowledge the financial support from the National Science Foundation and the US Air Force Office of Scientific Research.

I want to express my deep gratitude to my oral exam committee members: Professor Murti V. Salapaka for serving as the chair, Professor Zhi-Quan (Tom) Luo for his insight and suggestions on the problem of spectral tracking, Professor Mihailo Jovanovic who kindly let me share the office with his students for all these years, and Professor Stergios I. Roumeliotis to whom I am especially grateful for his help and kindness towards my entire family.

It has been my privilege to get to know Professor Robert R. Snapp at the University of Vermont. Without his encouragement and help during my two-year study there, it would not have been possible for me to transfer to the University of Minnesota.

I would like to acknowledge my colleague, Lipeng Ning, who worked closely with me during my last year of graduate study. Without him, the work on the multivariate power spectra would not be possible. My former colleague Mir Shahrouz Takyar collaborated with me on the problem of speech morphing, and also provided invaluable help and advice on my job searching, for which I am always grateful. Many thanks to Johan Karlsson at the Royal Institute of Technology for assisting me on the project of phoneme separation. I also appreciate the friendship with my fellow officemates: Fu Lin, Rashad Moarref, Binh Lieu, Armin Zare, Neil Dhingra, and Eric Dahlberg. They have been always there when I needed help.

My special thanks go to Uncle Har and Auntie Char with whom I have spent six years at Thanksgiving. Their kindness and friendship makes Minnesota a warm place to stay. I am also thankful to have gotten to know Weiyang Xie, who has been very dear to my family.

I am grateful to my parents in China, who raised me up and took good care of me, to my brother for his selfless support and encouragement. I would also like to thank Douglas and Linda Peacock who have been my second parents in America. I will never forget the love and care they have devoted to me. Dr. and Mrs. Welner have been my second grandparents in the US. Their long trips to visit us and many delicious lobsters we had in their home will always reside in my precious memory.

Finally, my deepest love and gratitude goes to my husband Ke. His support, encouragement, care and love walked me through my PhD years. Sometimes I have felt he has had more patience for me than I have had for myself. I am mostly grateful to have my daughter Grace. She is such a sweetheart to me. Her hugs and kisses bring me so much joy that my life has never been so beautiful.

Dedication

To Ke and Grace Zhou

Abstract

The purpose of this thesis is to study the geometry of power spectra and develop geometric methods for spectral analysis, modeling, and filtering of time series. We first study notions of distance suitable to quantify mismatch between probability distributions, power spectra, and covariance matrices. To this end, we review and relate relevant concepts from statistical inference, information theory, signal processing, and quantum mechanics. We introduce several alternative metrics for scalar as well as multivariate power spectra, and explore concepts from Riemannian geometry, in particular geodesics and means, to model slowly varying time series, and to interpolate and fuse statistics from different sources.

We first study mismatch between power spectra in the context of linear prediction. We introduce two alternative metrics. The first quantifies the degradation of predictive error variance when one power spectrum is used to design an optimal filter which is then applied to a process corresponding to another power spectrum. The second metric reflects the flatness of the innovations process. For this second metric, we provide closed-form expressions for the induced geodesics and geodesic distances. The theory we develop applies to multivariate power spectra. The metrics based on the prediction paradigm lack the important property of weak continuity. To this end we explore an alternative framework based on the transportation problem. In this, energy content has a physical significance and metrics quantify transportation cost between frequencies.

Throughout we explore geometric methods on the Riemannian manifold of power spectra, in a way analogous to the familiar Euclidean geometry. In particular, we use the notion of a geodesic to model the evolution of power spectral densities for slowly varying time series. More specifically, we develop the idea of geodesic path-fitting as an analogue of the least squares line-fitting in the Euclidean space. Further, we investigate means or averages of distributions and of positive semi-definite matrices arising in signal processing, computer vision and pattern recognition applications. The transportation mean, as well as the median, for normalized scalar power spectra are provided analytically. For multivariate Gaussian distributions, we show that the transportation mean requires solving a linear matrix inequality problem, which is computationally tractable. Furthermore, linear structural constraints on the means, based on prior knowledge, can be easily incorporated and solved efficiently using the linear matrix inequality

formulation. We highlight the relevance of the geometric framework with several applications in signal processing, such as spectral tracking, speech morphing, filtering, and spectral averaging.

Contents

Acknowledgements	i
Dedication	iii
Abstract	iv
List of Tables	ix
List of Figures	x
1 Introduction	1
1.1 Time Series and Spectral Analysis	1
1.2 Outline and Contributions	4
2 Metrics on Distributions and Positive Definite Matrices	6
2.1 One-dimensional Distributions	6
2.2 Positive Definite Matrices	10
2.3 Relationship among Metrics	15
3 Metrics on Multivariate Power Spectra	22
3.1 Preliminaries on Multivariate Prediction	23
3.1.1 Geometry of multivariable processes	23
3.1.2 Spectral factors and optimal prediction	26
3.2 Comparison of PSD's	27
3.2.1 Prediction errors and innovations processes	27
3.2.2 The color of innovations and PSD mismatch	28

3.2.3	Suboptimal prediction and PSD mismatch	30
3.2.4	Alternative divergence measures	32
3.3	Geometry on Multivariate Spectra	34
3.3.1	Riemannian structures	34
3.3.2	Geometry on positive matrices	37
3.3.3	Geodesics and geodesic distances	43
3.4	Examples	47
3.5	Extensions	52
4	Geodesics and Curve Fitting	55
4.1	Geodesic Path Fitting	56
4.2	Simulation Results	58
4.3	Extensions	61
5	Metrics and Means	66
5.1	The Welch Method and Arithmetic Mean	67
5.2	The Transportation Mean	69
5.2.1	The one-dimensional case	71
5.2.2	The multivariate-normal-distribution case	77
5.3	The Transportation Median in 1-D	82
5.4	The Kullback-Leibler Mean	86
5.5	Means of Positive Definite Matrices	89
5.6	Extensions	92
6	Applications	93
6.1	Speech Morphing	93
6.2	Phoneme Separation	97
6.3	Spectral Averaging	100
6.4	The Gaussian Particle Filter	102
6.4.1	Linearized transformation and unscented transformation	104
6.4.2	The Gaussian particle transformation	105
6.4.3	Examples	107
6.4.4	Extensions	110

References	111
Appendix A. Review of Differential Geometry and Fibre Bundles	122
Appendix B. An Example of Using Christoffel Symbols to Compute Geodesics	126
Appendix C. Probability and Measure Review	130
C.1 Topology and Topological Space	130
C.2 Measure and Measurable Space	132
C.3 Convergence	135
Appendix D. Another Interpretation of KL Divergence and Fisher Metric	138

List of Tables

6.1 Instantaneous Frequency Estimation	104
--	-----

List of Figures

3.1	Plots of $\log f_0(\theta)$ (upper) and $\log f_1(\theta)$ (lower) for $\theta \in [0, \pi]$	48
3.2	Locus of the roots of $a_\tau(z)$ for $\tau \in [0, 1]$	49
3.3	$\log f_{\tau, \text{AR}}(\theta)$ for $\tau = \frac{1}{3}, \frac{2}{3}, \frac{4}{3}$ (blue), $\tau = 0, 1$ (red).	50
3.4	$\log f_{\tau, \text{linear}}(\theta)$ for $\tau = \frac{1}{3}, \frac{2}{3}, \frac{4}{3}$ (blue), $\tau = 0, 1$ (red).	50
3.5	$\log f_{\tau, \text{geodesic}}(\theta)$ for $\tau = \frac{1}{3}, \frac{2}{3}, \frac{4}{3}$ (blue), $\tau = 0, 1$ (red).	51
3.6	Subplots (1,1), (1,2) and (2,2) show $\log f_0(1, 1), \log f_0(1, 2) $ (same as $\log f_0(2, 1) $) and $\log f_0(2, 2)$. Subplot (2,1) shows $\arg(f_0(2, 1))$	52
3.7	Subplots (1,1), (1,2) and (2,2) show $\log f_1(1, 1), \log f_1(1, 2) $ (same as $\log f_1(2, 1) $) and $\log f_1(2, 2)$. Subplot (2,1) shows $\arg(f_1(2, 1))$	53
3.8	Subplots (1,1), (1,2) and (2,2) show $\log f_\tau(1, 1), \log f_\tau(1, 2) $ (same as $\log f_\tau(2, 1) $) and $\log f_\tau(2, 2)$. Subplot (2,1) shows $\arg(f_\tau(2, 1))$, for $\tau \in [0, 1]$	54
4.1	Power spectra $f_{0, \text{ideal}}$ and $f_{1, \text{ideal}}$ as functions of θ/π	59
4.2	STFT spectrogram and estimated geodesic path.	60
4.3	First row: time-series data; second row: STFT spectra based on the highlighted parts of the time-series; third row: samples of geodesic fit to the STFT spectra.	61
4.4	STFT spectrogram and the prediction geodesic path.	62
4.5	First row: time-series data; second row: STFT spectra based on the highlighted parts of the time-series; third row: samples of prediction geodesic fit to STFT spectra.	63
4.6	STFT spectrogram.	63
4.7	Top row: time-series with two chirp signals and additive noise; second row: STFT spectra corresponding to windows marked with blue; third row: estimated geodesic-fit samples with penalty terms; last row: the estimate geodesic path.	64

4.8	Top row: time-series with two chirp signals and additive noise; second row: STFT spectra corresponding to windows marked with blue; third row: estimated geodesic-fit samples without penalty terms; last row: the estimated geodesic path.	64
4.9	STFT spectrogram.	65
4.10	Top row: time-series with two chirp signals and additive noise; second row: STFT spectra corresponding to windows marked with blue; third row: estimated pairwise geodesic-fit samples; last row: the estimated spectrum.	65
5.1	Arithmetic mean of two Gaussian distributions.	70
5.2	An example of transportation mean. The center of the second row shows the centroid, while others are the given spectra.	75
5.3	Transportation mean of two Gaussian distributions	76
6.1	Linear prediction analysis and synthesis framework	96
6.2	Time signals corresponding to the phoneme “a” for A (female) and B (male)	97
6.3	Power spectra for subjects A (female) and B (male)	98
6.4	Geodesic path between the two spectra following L_1 distance	99
6.5	Geodesic path between the two spectra following prediction metric	100
6.6	Geodesic path between the two spectra following transportation metric	101
6.7	Speech sentence and phoneme segmentation. In the top sub-figure, diamonds designate segmentation by an “expert”. Asterisks, circles, squares and crosses correspond to segmentation based on prediction geodesic distance, L_2 -log distance, L_1 -log distance and the Itakura-Saito distance respectively.	102
6.8	Part of the sentence and phoneme segmentation results. In the top sub-figure, diamonds indicate markings by an “expert”. Asterisks, circles, squares and crosses are marked based on prediction geodesic distance, L_2 -log distance, L_1 -log distance and the Itakura-Saito distance respectively.	103
6.9	Illustration of the area/volumn filter	110
B.1	The red curve is the geodesic path by using the Christoffel symbols. The green curve is based on the closed-form expression (B.3), whose corresponding arc on the unit sphere is plotted in black.	129
C.1	Topological spaces	131
D.1	The ellipsoid uncertainty region on the simplex (in red) and the corresponding disc on the sphere (in blue).	141

Chapter 1

Introduction

A power spectral density (PSD) function represents the energy distribution of a second-order stationary time series over frequencies [1]. In sonar and radar systems, the PSD of a received signal provides information on the locations of sources with respect to the receiver. For speech signal, the spectral content can be used to detect formants, which is essential in speech recognition and synthesis. For stock market data, spectral density may reveal hidden periodic patterns.

Accordingly, much effort has been devoted to the techniques of spectral estimation, which is non-trivial due to the finite record of observations. Below, we first provide a brief account on the spectral analysis methods for stationary or slowly varying time series. We then propose our geometric methods and describe these in detail in the following chapters.

1.1 Time Series and Spectral Analysis

We first briefly review spectral estimation methods for stationary processes. These methods can be divided into two categories, parametric or non-parametric. For more details, we refer to the book by Stoica and Moses [1].

Parametric approaches postulate a model for the data, thus convert the task into estimating model parameters. When the data indeed satisfy the assumed model, parametric methods normally provide a quite accurate estimation. Some well-known parametric estimation algorithms are Yule-Walker, least-squares, Burg, Pisarenko and MUSIC, Min-norm, ESPRIT, and so forth.

Non-parametric methods typically estimate the energy in each frequency by passing the signal through a narrow-banded filter centered at that frequency. The energy of the filtered

signal divided by the bandwidth of the filter approximates the spectral content of the signal at that frequency. The classical methods in this category are Blackman-Tukey, Bartlett, Welch, Daniell, Slepian method, RFB, Capon, and so on.

The above spectral estimation methods are based on the assumption that the underlying process is second-order stationary, while in reality most signals we encounter are time-varying. In speech signal, formants change from one phoneme to another. In seismology, the spectral content of signal changes with ground movement. In sonar and radar systems, spectral density also changes if a target moves. This brings up two questions. First, how can we decide whether a process is stationary? Second, what can we do if the process is not stationary?

To determine whether a process is stationary, Silverman [2] defined a class of random processes to be locally stationary if the covariance can be expressed as a stationary covariance multiplied by a sliding power factor. Martin [3] studied the class of quasi-stationary harmonizable processes which can be approximated as piecewise stationary processes. He then used so-called stability and precision of the pseudo Wigner estimator as a measure of non-stationarity in each frequency band. Xiao et al. [4, 5] generated a set of surrogate data to serve as the null hypothesis of stationarity. These surrogate data have the same marginal spectrum as the non-stationary process but with the uniformly distributed spectrum phase. Then they compared the selected statistic significance of the actual observation with the surrogate data to decide the degree of non-stationarity. Priestley and Rao [6] employed the log transform of the evolutionary spectrum estimator to test whether the process is stationary. The test is equivalent to a two-factor analysis of variance procedure with the known residual variance. By parameterizing a locally wide sense stationary Gaussian random process as a time-varying autoregressive (AR) process with each coefficient in the linear span of some known basis functions, Kay [7] proposed to use a Rao test to check the stationarity of the data.

Regarding the spectral analysis of non-stationary signals, the techniques proposed can still be divided into parametric or non-parametric methods. Along the line of parametric methods, Rao [8] proposed to approximate the time-varying coefficients of an AR model by the first three terms of the Taylor expansion. Unknown parameters are estimated by minimizing a weighted least square problem. Grenier [9] extended well-known Levinson's AR model, Burg's method, Prony's method, and Cadzow's autoregressive-moving-average (ARMA) model to the time-varying case, under the restriction that the unknown coefficients are the weights of the

linear combinations of known basis functions. Dahlhaus [10] proposed to estimate slowly time-varying parameters of a non-stationary process by minimizing a generalized Whittle function over periodograms constructed for successive time-windows. Under the assumption that the coefficients of ARMA model change slowly with time, Kaderli and Kayhan [11] obtained the AR coefficients from the time-varying modified Yule-Walker equations. The time-varying auto-correlations are estimated from the inverse Fourier transform of the evolutionary periodogram. Based on the estimated evolutionary cepstrum, they obtained the MA coefficients by a simple recursion.

Nonparametric methods generally assume that the signal doesn't change too fast. The signal is considered to be stationary in a sufficiently small time interval, and spectral estimation can be carried out in each of such intervals. This is called the time-frequency analysis, of which the short-time Fourier transform (STFT) is the most widely used tool. The STFT can be thought of as the inner product of the signal with basis sinusoidal signals having support over a window which is shifted in time. It is straightforward and easy to implement, but it is limited by the uncertainty principle. A long window may give smaller variance of the estimation, but reduces the time resolution. A short window, on the other hand, increases the time resolution with the sacrifice of the frequency resolution. Along a similar spirit, wavelet transforms take the inner product of the signal with translated and dilated square integrable functions, which are also called wavelets. Since the wavelet will be highly concentrated at the high frequency range, it's believed that wavelet transform has better time-resolution than STFT at high frequency range [12].

Another important method is based on so called time-frequency distributions (TFD), which is a general extension of the classical Fourier analysis. Its concept is closely related to the instantaneous frequency (IF), which corresponds to the frequency of a sinusoid wave. TFD approximates the nonstationary signal as the sum of several monocomponent time varying signals. For each monocomponent signal, TFD tracks the spectral variation based on IF, and provides the energy concentration of the signal around IF at each time. Cohen generalized many well-known TFDs into a unified formulation, which is called Cohen's bilinear class. The desired properties of a TFD are as follows. (1) TFD must be real; (2) Integration of TFD over both time and frequency yields the energy of the signal; (3) Integration of TFD over time yields the spectral density of the signal; (4) Integration of TFD over frequency yields the instantaneous power of the signal; (5) The first moments of TFD yield the instantaneous frequency and the time-delay

of the signal. Four major TFDs have all these properties: the Wigner-Ville distribution, the Rihaczek distribution, the Cohen distribution, and the Choi-Williams distribution. However, these distributions can't guarantee the positivity of TFD. To ensure the positivity of TFD, property (5) has to be sacrificed, which is the case of spectrogram. Because of the possible negative values of TFD, it is more proper to interpret TFD as an energy gradient representation, rather than the distribution of energy over time and frequency [13, 14].

Several generalizations have been made to extend the concept of spectrum to non-stationary processes. One of the early attempts is Page's instantaneous power spectra, which is defined as the differentiation of magnitude squared Fourier transform of the signal from the past to the current with respect to time [15]. Under the assumption that the non-stationary characteristics change slowly over time, Priestley [16] proposed the concept of evolutionary spectrum, which can be interpolated as the local energy distribution over frequency. The evolutionary spectrum can be estimated by using the "double windows" technique. By choosing proper windows, the estimator can be unbiased, and the variance of the estimator was also derived by Priestly.

1.2 Outline and Contributions

In this thesis, we propose geometric methods for spectral analysis. Naturally, a metric geometry allows power spectral density functions to be thought of as points in a suitable space and provides a framework for problems of approximation, smoothing, and averaging, in a way analogous to the familiar Euclidean geometry.

To proceed, we first need suitable metrics to quantify the distance between two power spectral density functions. In Chapter 2, we review various metrics and their geometries for probability distributions, power spectra, as well as positive definite matrices. Such distances arise naturally in statistic inference, information theory, signal processing, or quantum mechanics. We discuss their relationship in the context of Gaussian random processes.

In Chapter 3, we generalize the prediction-based distance for multivariate power spectra. More specifically, two categories of divergences are proposed. The first quantifies the degradation of the predictive error variance: a choice between the two power spectra is used to design an optimal filter, which is then applied to a process corresponding to the second power spectrum. The second quantifies the "flatness" of the innovations process. We further identify corresponding Riemannian metrics, and for the second metric we provide closed-form expressions for the

induced geodesics and geodesic distances.

We then utilize notions of distances and geodesics to model the evolution of power spectral densities in Chapter 4. To be more precise, we propose geodesic path fitting on a Riemannian manifold in a similar way as the least squares line fitting in the Euclidean space. We put our emphasis on the transportation distance and its induced geodesic due to its weak continuity property.

In Chapter 5, we investigate means or averages of distributions or positive definite matrices. The transportation mean, as well as the median, for normalized scalar-valued distributions are provided analytically. For multivariate zero-mean Gaussian distributions, we show that the computation of the transportation mean requires solving a linear matrix inequality problem, which is computationally tractable. As a result, linear structural constraints on the means can be easily incorporated.

Finally, in Chapter 6 we discuss some applications of these geometric methods to signal processing problems, such as speech morphing and spectral averaging.

Chapter 2

Metrics on Distributions and Positive Definite Matrices

A variety of distance measures have been proposed to quantify changes in probability distributions or power spectral densities [17, 18, 19, 20]. These have been linked to probabilistic concepts, statistical inference, quantum mechanics, as well as practical issues motivated by applications such as speech analysis. This chapter provides a general view of this topic. More specifically, Section 2.1 focuses on distances and metrics for one-dimensional distributions. Section 2.2 reviews metrics for positive definite matrices, such as covariance matrices, and density matrices in quantum mechanics. Section 2.3 links various metrics naturally by considering Gaussian random processes.

2.1 One-dimensional Distributions

The *Kullback-Leibler (KL) divergence* is firmly rooted in probability theory for its relevance in data compression and hypothesis testing (see Appendix D). In information theory, the KL divergence quantifies the degradation in source coding efficiency when a code is based on a particular probability distribution f_1 while the pdf of the source f_0 differs. It can be expressed as

$$d_{\text{KL}}(f_0, f_1) := \int_X f_0 \log \frac{f_0}{f_1} dx,$$

where X is the underlying probability space. When f_1 is an infinitesimal perturbation of f_0 , the KL divergence gives rise to the *Fisher information metric*

$$\mathfrak{g}_{f,\text{Fisher}}(\delta) := \int_X \frac{\delta^2}{f} dx \quad (2.1)$$

by ignoring the higher order terms on δ [21]. This turns out to be natural from one additional perspective. It is the unique Riemannian metric for which the stochastic maps are contractive [22] –a property that motivates a family of metrics in the context of matricial counterparts of probability distributions (see Section 2.2). The induced geodesic of the Fisher metric can be expressed as

$$\sqrt{f_\tau} = \frac{\sin(1-\tau)\vartheta}{\sin\vartheta} \sqrt{f_0} + \frac{\sin\tau\vartheta}{\sin\vartheta} \sqrt{f_1} \quad (2.2)$$

for $\tau \in [0, 1]$, and the geodesic distance is the angle ϑ between $\sqrt{f_0}$ and $\sqrt{f_1}$, i.e.,

$$\cos\vartheta = \int_{-\pi}^{\pi} \sqrt{f_0 f_1} \frac{d\theta}{2\pi}.$$

The latter is also known as the *Bhattacharyya coefficient*. Geodesics in the Fisher metric in fact correspond to great circles on the “sphere” $\{\sqrt{f} : \int f = 1\}$ and geodesic distances to the corresponding arclength. A closely related metric is the Hellinger distance [23], also called the Matusita distance [24], which measures the straight line distance between $\sqrt{f_0}$ and $\sqrt{f_1}$.

The KL divergence has been widely used in information theory, however it may not be the best choice in other applications. In particular, in the context of speech analysis and synthesis, the *Itakura-Saito distance* [25, 19] has been commonly used to measure the mismatch between two spectra f_0 and f_1 . This is given by

$$d_{\text{IS}}(f_0, f_1) := \int_{-\pi}^{\pi} \left(\frac{f_0}{f_1} - \log \frac{f_0}{f_1} - 1 \right) \frac{d\theta}{2\pi}.$$

In this the density functions are not required to have the same integrals. It is well-known that infinitesimal perturbations from a given $f(\theta)$ lead the Itakura-Saito distance to a Riemannian metric [17]

$$\mathfrak{g}_{f,\text{IS}}(\delta) := \|\mathfrak{f}^{-1}\delta\|_2^2. \quad (2.3)$$

Moreover, the unique geodesic path from f_0 to f_1 is parameterized as

$$f_\tau = f_0^{1-\tau} f_1^\tau,$$

and the geodesic distance is the *log spectral deviation* [18]

$$d_{\text{Log}}(f_0, f_1) := \|\log f_0 - \log f_1\|_2. \quad (2.4)$$

A gain-optimized expression of the Itakura-Saito distance, which is often called the *Itakura distance* [26, 19], has also been suggested and is given by

$$\begin{aligned} d_I(f_0, f_1) &:= \min_{\alpha > 0} d_{\text{IS}}(f_0, \alpha f_1) \\ &= \log \int_{-\pi}^{\pi} \frac{f_0(\theta) d\theta}{f_1(\theta) 2\pi} - \int_{-\pi}^{\pi} \log \frac{f_0(\theta) d\theta}{f_1(\theta) 2\pi}. \end{aligned}$$

The exact same expression was obtained in [17] by considering the ratio of the “degraded” predictive error variance over the optimal variance by analogy to how the KL divergence quantifies degradation in source coding efficiency. When one spectral density is a perturbation of another, the Itakura distance leads (modulo a scaling factor of 2) to the Riemannian pseudo-metric [17]

$$\mathfrak{g}_{f, \text{Prediction}}(\delta) := \int_{-\pi}^{\pi} \left(\frac{\delta(\theta)}{f(\theta)} \right)^2 \frac{d\theta}{2\pi} - \left(\int_{-\pi}^{\pi} \frac{\delta(\theta) d\theta}{f(\theta) 2\pi} \right)^2 \quad (2.5)$$

on density functions. This is analogous to the Fisher metric and we refer to it as the “*prediction metric*” due to its roots in quadratic prediction theory [17]. It turns out that geodesic paths f_τ ($\tau \in [0, 1]$) connecting spectral densities f_0, f_1 and having minimal length can be explicitly computed as

$$f_\tau(\theta) = f_0^{1-\tau}(\theta) f_1^\tau(\theta), \quad (2.6)$$

for $\tau \in [0, 1]$, see [17]. Furthermore, the length along such “prediction geodesics” can be also *explicitly* computed in terms of the end points

$$d_{\text{pg}}(f_0, f_1) := \sqrt{\int_{-\pi}^{\pi} \left(\log \frac{f_1(\theta)}{f_0(\theta)} \right)^2 \frac{d\theta}{2\pi} - \left(\int_{-\pi}^{\pi} \log \frac{f_1(\theta) d\theta}{f_0(\theta) 2\pi} \right)^2}, \quad (2.7)$$

which is closely related to the L_2 norm of the difference of the log spectra used in speech processing [19].

While these geometric concepts have proven relevant in many applications, they lack a very important and natural property that is often desirable, weak continuity. By definition, a sequence of measures $\{\mu_k\}$ converges to μ weakly if for any continuous bounded function ϕ , $\int \phi d\mu_k \rightarrow \int \phi d\mu$ as $k \rightarrow \infty$. This is desirable for a notion of convergence of power spectra since small changes in the power spectral density ought to reflect small changes in the statistics and

vice versa. For instance, a “proper” distance between two line-spectra ought to decrease as two lines move closer, but it is not the case for distances that fail to be weakly continuous. In order to bring in this additional property, we consider a certain family of distances that have a long history in mathematics and relate to the so-called Monge-Kantorovich transportation problem [27, pg. 212].

In the context of “transportation” problem, density functions represent mass distributions and a cost is associated with transferring one unit of mass from one location to another. The idea is that by doing so, one gives a precise correspondence between two densities while the transportation distance quantifies the optimal cost for establishing this correspondence. The original formulation goes back to Gaspar Monge in 1781, while Leonid Kantorovich developed the modern formulation and theory after the World War II –interestingly, Leonid Kantorovich received the Nobel prize in Economics for related work and application of the theory .

Throughout, we will only consider quadratic transference-cost which is known as the *Wasserstein distance* of order 2. We do so due to the computational simplicity of the metric. Thus, for two density functions f_0 and f_1 supported on the space X (e.g., $X = [-\pi, \pi]$), the transportation distance is defined as follows:

$$d_{W_2}(f_0, f_1) := \left\{ \inf_{\psi(x)} \int_X |x - \psi(x)|^2 f_0(x) dx \right\}^{\frac{1}{2}}$$

where, in general, $\psi(x)$ is a measure-preserving map between f_0 and f_1 . That is,

$$|\det \nabla \psi(x)| f_1(\psi(x)) = f_0(x). \quad (2.8)$$

In the case that f_0 and f_1 are one dimensional power spectra, the optimal transportation plan can be computed *explicitly* [27, pg. 75] through

$$F_0(\theta) = F_1(\psi(\theta)), \quad (2.9)$$

where

$$F_i(\theta) = \int_{-\pi}^{\theta} f_i(\vartheta) d\vartheta$$

is the cumulative function of f_i ($i = 0, 1$). It allows determining $\psi(\theta)$ by integrating f_0, f_1 and comparing the respective values of their integrals [28, section 3.2]. The scaling function $\psi(\theta)$ is monotonically increasing since both f_0 and f_1 are positive. Now, let $t = F_0(\theta)$ and $t \in [0, 1]$, then we can represent θ by an inverse function of F_0 , that is

$$\theta = F_0^{-1}(t), \text{ and similarly } \psi(\theta) = F_1^{-1}(t).$$

Consequently in this one dimensional case, the transportation distance has a closed form

$$d_{W_2}(f_0, f_1) = \left\{ \int_0^1 (F_0^{-1}(t) - F_1^{-1}(t))^2 dt \right\}^{\frac{1}{2}}. \quad (2.10)$$

When f_1 is a small perturbation of f_0 , the transportation distance results in a Riemannian metric; see [27, section 7.6], where the more general high dimensional space is also considered. In the one dimensional case, this as well can be computed explicitly as

$$g_{f, W_2}(\delta) := \int_{-\pi}^{\pi} \frac{\Phi(\theta)^2}{f(\theta)} \frac{d\theta}{2\pi},$$

where

$$\Phi(\theta) = \int_{-\pi}^{\theta} \delta(\vartheta) \frac{d\vartheta}{2\pi}.$$

It is interesting to see how this metric may be seen as related to the Fisher information metric. Perturbations δ of the density function having integral zero, can be thought of as “tangent vectors” in the space of densities. The requirement that the integral is zero follows from the requirement that all densities have the same mass. Thus, in order to compute $d_{W_2}^2(f, f + \delta)$ one needs to integrate this tangent vector δ and compute Φ —this is a linear operation. It is important to recall that f_0 and f_1 can be thought of as defined on the unit circle. Accordingly $\Phi(\theta)$ can be represented as the integral of δ starting from any point on $[-\pi, \pi]$, not necessary $-\pi$, and returning to it modulo 2π .

Geodesics f_τ ($\tau \in [0, 1]$) between two end “points” f_0 and f_1 are determined by a gradient flow [27, page 252], which, in this very special one-dimensional case, specifies the geodesic via

$$F_\tau((1 - \tau)\theta + \tau\psi(\theta)) = F_0(\theta), \quad (2.11)$$

for $\psi(\theta)$ computed from (2.9) with F_τ being the cumulative function of f_τ .

2.2 Positive Definite Matrices

In quantum mechanics, the uncertainty of a quantum system is measured by the von Neumann entropy [29] as a direct analog of the Shannon entropy. We shall consider only the manifold \mathcal{M} of invertible finite density matrices to avoid boundary problems:

$$\mathcal{M} := \{\rho : \rho = \rho^*, \rho > 0, \text{tr}(\rho) = 1\}.$$

For a quantum system described by a density matrix $\rho \in \mathcal{M}$, the von Neumann entropy is defined as:

$$S_{\text{vN}}(\rho) := -\text{tr}(\rho \log \rho)$$

The von Neumann entropy has also been interpreted as the capacity of a quantum communication channel as the quantum analog of the Shannon entropy in the classical information theory [30]. Furthermore, Umegaki [31] considered the quantum relative entropy between two density matrices ρ_0 and ρ_1 :

$$S_{\text{vN}}(\rho_0 || \rho_1) = \text{tr}(\rho_0 \log \rho_0) - \text{tr}(\rho_0 \log \rho_1).$$

It measures the difficulty of distinguishing the state ρ_0 from the state ρ_1 [32, 33]. Just as the Kullback-Leibler divergence gives rise to the Fisher information metric, the quantum relative entropy induces the Kubo-Mori inner product as a Riemannian metric [34]. Let X and Y be tangent vectors at ρ , which are self-adjoint matrices with trace 0, the Kubo-Mori inner product of X and Y is defined to be

$$g_{\rho, \text{KM}}(X, Y) := \int_0^\infty \text{tr} X(\rho + s)^{-1} Y(\rho + s)^{-1} ds.$$

Its derivation is based on the observation that

$$\partial_t \log(\rho + tX)|_{t=0} = \int_0^\infty (\rho + s)^{-1} X(\rho + s)^{-1} ds,$$

thus the Hessian of the relative entropy is¹

$$-\partial_{t_1} \partial_{t_2} S_{\text{vN}}(\rho + t_1 X || \rho + t_2 Y)|_{t_1, t_2=0} = \partial_{t_2} \text{tr}(X \log(\rho + t_2 Y)) = \int_0^\infty \text{tr} X(\rho + s)^{-1} Y(\rho + s)^{-1} ds.$$

As noted earlier, the Fisher information metric is the unique Riemannian metric for which the stochastic maps are contractive. In quantum mechanics, a similar property has been sought

¹ There is an alternative expression for the derivative of $\log(\rho)$ in terms of its Loewner matrix [35, p. 154]. Suppose ρ has the eigen-decomposition $\rho = U \Lambda U^*$. The Loewner matrix $F(\rho)$ of $\log(\rho)$ is

$$[F(\rho)]_{i,j} = \begin{cases} \lambda_i^{-1}, & i = j \\ \frac{\log \lambda_i - \log \lambda_j}{\lambda_i - \lambda_j}, & i \neq j \end{cases}$$

Then the derivative of $\log(\rho)$ is $\partial_t \log(\rho + tX)|_{t=0} = U(F(\rho) \circ (U^* X U))U^*$ where \circ denotes the Schur product. As a result, the Kubo-Mori metric can also be expressed as

$$g_{\rho, \text{KM}}(X, Y) = \langle U^* X U, F(\rho) \circ (U^* Y U) \rangle.$$

for the non-commutative analog of probability vectors, namely, density matrices. These are positive semi-definite and have trace equal to one. In this setting, there are several metrics for which stochastic maps (these are now linear maps between spaces of density matrices, preserving positivity and trace) are contractive. They take the form

$$\text{tr}(XD_\rho(Y))$$

where $D_\rho(Y)$ can be thought as a “non-commutative division” of a matrix Y by the matrix ρ . Thus, if ρ , X and Y are scalars, the above is simply XY/ρ . Examples of such a “non-commutative division” are

$$D_{\rho,1}(Y) := \rho^{-1}Y, \quad (2.12a)$$

$$D_{\rho,2}(Y) := \int_0^\infty (\rho + sI)^{-1}Y(\rho + sI)^{-1}ds, \quad (2.12b)$$

$$D_{\rho,3}(Y) := M, \text{ where } \frac{1}{2}(\rho M + M\rho) = Y, \quad (2.12c)$$

see e.g., [36]. The metric which corresponds to (2.12a) was introduced by Petz [36], the one corresponding to (2.12b) is the Kubo-Mori metric [36], while the one corresponding to (2.12c) is known as the Bures metric [37].

We refer to Appendix A for background information on the Bures metric. As long as $\rho > 0$, the Hermitian matrix M is determined uniquely [38]. The Bures length of a curve $\rho(t)$ is thus defined to

$$L(\rho) = \int \sqrt{\text{tr}(M\dot{\rho})}dt, \text{ s.t. } \dot{\rho} = \frac{1}{2}(M\rho + \rho M).$$

It's easy to see that given two density matrices $\rho_0 = W_0W_0^*$ and $\rho_1 = W_1W_1^*$, where W_0 and W_1 are not unique since W_0U_0 and W_1U_1 are also factors for U_0 and U_1 being unitary matrices. The parallel purification of the geodesic path between ρ_0 and ρ_1 is the arc with the minimum arc length on great circles passing between W_0 's and W_1 's. Thus, the geodesic distance between ρ_0 and ρ_1 is the minimum angle between W_0 's and W_1 's. To be more specific, the angle ϑ between W_0 and W_1 can be computed from

$$\cos \vartheta = \text{Real } \text{tr}(W_0^*W_1).$$

Clearly the angle ϑ is minimized when $|\text{tr}(W_0^*W_1)|$ achieves its maximum. It can be shown that the maximum of $\cos \vartheta$ is achieved if $W_1 = (\rho_1 \sharp \rho_0^{-1})W_0$, where $\rho_1 \sharp \rho_0^{-1}$ is the geometric mean between ρ_0^{-1} and ρ_1 :

$$\rho_1 \sharp \rho_0^{-1} = \rho_1^{\frac{1}{2}}(\rho_1^{\frac{1}{2}}\rho_0\rho_1^{\frac{1}{2}})^{-\frac{1}{2}}\rho_1^{\frac{1}{2}}.$$

As a result, the maximum is $\cos \vartheta = \text{tr}(\rho_0^{\frac{1}{2}} \rho_1 \rho_0^{\frac{1}{2}})^{\frac{1}{2}}$, which is also called the root fidelity in quantum mechanics. The corresponding angle is called the Bures angle. The arc between W_0 and W_1 can be parameterized as

$$W(\tau) = \frac{\sin(1-\tau)\vartheta}{\sin\vartheta} W_0 + \frac{\sin\tau\vartheta}{\sin\vartheta} W_1,$$

and the geodesic path of the density matrices becomes $\rho(\tau) = W(\tau)W(\tau)^*$ for $\tau \in [0, 1]$.

The Bures length can be seen as the arc length between W_0 and W_1 . The straight-line distance between them is the Bures distance [39], which corresponds to the following optimization problem:

$$\begin{aligned} d_B(\rho_0, \rho_1) &:= \min_{U, V} \left\{ \|\rho_0^{\frac{1}{2}} U - \rho_1^{\frac{1}{2}} V\|_F \mid UU' = I, VV' = I \right\} \\ &= \min_U \left\{ \|\rho_0^{\frac{1}{2}} U - \rho_1^{\frac{1}{2}}\|_F \mid UU' = I \right\}, \end{aligned} \quad (2.13)$$

since, clearly, only one unitary transformation U can attain the same minimal value. Note that the Bures distance has been considered as a generalization of the standard Hellinger distance by Ferrante *et al.* [40], and this differs from the quantum Hellinger distance $\|\rho_0^{1/2} - \rho_1^{1/2}\|_F$ proposed in [41].

Interestingly, as shown in [37, 40],

$$d_B(\rho_0, \rho_1) = \left(\text{tr}(\rho_0 + \rho_1 - 2(\rho_0^{\frac{1}{2}} \rho_1 \rho_0^{\frac{1}{2}})^{\frac{1}{2}}) \right)^{\frac{1}{2}}.$$

Also, the optimizing unitary matrix U in (2.13) is

$$U = \rho_0^{-\frac{1}{2}} \rho_1^{-\frac{1}{2}} (\rho_1^{\frac{1}{2}} \rho_0 \rho_1^{\frac{1}{2}})^{\frac{1}{2}}.$$

We note that the Bures distance applies equally well to positive definite matrices without any need for normalization on the traces, and as such, it has been used to compare multivariate power spectral densities [40].

We now depart from density matrices, and focus on the positive definite matrices. We use \mathbb{P}_n to denote the manifold of $n \times n$ positive definite matrices. In recently years, a particular Riemannian metric has received much attention, which coincides with the Rao metric used in statistical inference in the context of Gaussian distributions, see Section 2.3. It is defined as

$$\mathfrak{g}_{q, \text{Rao}}(dq) := \|q^{-\frac{1}{2}} dq q^{-\frac{1}{2}}\|_F^2 = \text{tr}(q^{-1} dq)^2$$

at the point q on the manifold \mathbb{P}_n . The length of a curve $\gamma(t)$ in \mathbb{P}_n can be computed as

$$L(\gamma) = \int \|\gamma^{-\frac{1}{2}}(t)\dot{\gamma}(t)\gamma^{-\frac{1}{2}}(t)\|_F dt.$$

Our exposition on this metric follows mainly the work by Bhatia and Holbrook [42]. First of all, this metric is congruence invariant, i.e., for any invertible $g \in GL_n$, the metric at q with a tangent vector dq is the same as the metric at gqg^* with the tangent vector $gdqg^*$. Consequently

$$L(\gamma) = L(g\gamma g^*).$$

Given two points γ_0 and γ_1 in \mathbb{P}_n , we are interested in the shortest path between them. The geodesic distance is

$$d(\gamma_0, \gamma_1) = \inf\{L(\gamma) : \gamma(0) = \gamma_0, \gamma(1) = \gamma_1\}.$$

To derive the geodesic, we shall first introduce the infinitesimal exponential metric increasing property (IEMI).

Proposition 1 (IEMI property [42, pg. 203]). *For any Hermitian matrices H and K in \mathbb{H}_n , let $De^H(K)$ denote the derivative of the exponential map at the point H in the direction K , i.e.,*

$$De^H(K) = \lim_{t \rightarrow 0} \frac{e^{H+tK} - e^H}{t}.$$

Then for any unitary invariant norm $\|\cdot\|$, the following inequality holds:

$$\|e^{-H/2}De^H(K)e^{-H/2}\| \geq \|K\|.$$

Since $\gamma(t) \in \mathbb{P}_n$, it can be written as $\gamma(t) = e^{H(t)}$, where $H(t) = \log \gamma(t)$. Consequently $\dot{\gamma}(t) = De^{H(t)}(\dot{H}(t))$. Using the IEMI property, we get

$$\begin{aligned} \int_0^1 \|\gamma^{-\frac{1}{2}}\dot{\gamma}\gamma^{-\frac{1}{2}}\|_F dt &\geq \int_0^1 \|\dot{H}(t)\|_F dt \\ &= \|\log \gamma_0 - \log \gamma_1\|_F. \end{aligned}$$

If γ_0 and γ_1 commute, then by choosing the path

$$\gamma(t) = e^{(1-t)\log \gamma_0 + t\log \gamma_1} = \gamma_0^{1-t}\gamma_1^t, \quad (2.14)$$

we have

$$\dot{\gamma}(t) = \gamma(t)(\log \gamma_1 - \log \gamma_0),$$

and

$$L(\gamma) = \|\log \gamma_0 - \log \gamma_1\|_F.$$

In addition, since the shortest path between $\log \gamma_0$ and $\log \gamma_1$ is the unique straight line segment, the path $\gamma(t)$ in (2.14) is the unique shortest path between two commuting matrices. Furthermore, $\gamma(t)$ in (2.14) is the geodesic parameterized by the arc length, i.e.,

$$d(\gamma_0, \gamma(t)) = td(\gamma_0, \gamma_1).$$

If γ_0 and γ_1 do not commute, using the congruence invariant property of the metric, we have

$$d(\gamma_0, \gamma_1) = d(I, \gamma_0^{-1/2} \gamma_1 \gamma_0^{-1/2}).$$

Since the identity matrix I commutes with $\gamma_0^{-1/2} \gamma_1 \gamma_0^{-1/2}$, we readily obtain the geodesic distance

$$d(\gamma_0, \gamma_1) = \|\log I - \log \gamma_0^{-1/2} \gamma_1 \gamma_0^{-1/2}\|_F = \|\log \gamma_0^{-1/2} \gamma_1 \gamma_0^{-1/2}\|_F, \quad (2.15)$$

which can be seen as the log deviation between γ_0 and γ_1 . The unique geodesic is

$$\gamma(t) = \gamma_0^{1/2} (\gamma_0^{-1/2} \gamma_1 \gamma_0^{-1/2})^t \gamma_0^{1/2}.$$

2.3 Relationship among Metrics

For simplicity, we consider a stationary zero-mean Gaussian random process $\{Y_k; k \in \mathbb{Z}\}$. The probability distribution p_n of any random vector $Y_1^n = (Y_1, \dots, Y_n)'$ is fully described by its Toeplitz covariance matrix $P_n = E(Y_1^n Y_1^{n'}) \in \mathcal{T}_n$, where

$$\mathcal{T}_n = \{X \in \mathbb{P}_n : X \text{ Toeplitz}\}$$

and \mathbb{P}_n being the manifold of n -by- n positive definite matrices as noted earlier. To simplify the notation, we shall use $\int dy$ to denote $\int dy_1 \cdots dy_n$. The differential entropy of Y_1^n is

$$h(Y_1^n) = - \int p_n \log p_n dy = \frac{1}{2} (n \log 2\pi + n + \log \det(P_n)).$$

As n goes to ∞ , the average of the differential entropy gives rise to the differential entropy rate $h(Y)$ of the random process $\{Y_k\}$:

$$h_\infty(Y) := \lim_{n \rightarrow \infty} \frac{1}{n} h(Y_1^n)$$

provided the limit exists. Let $f(\theta)$ denote the power spectral density of $\{Y_k\}$. It has been shown [43, pg. 273-274] that for Gaussian random process $\{Y_k\}$,

$$h_\infty(Y) = \frac{1}{2} \left(\log(2\pi) + 1 + \int_{-\pi}^{\pi} \log f(\theta) \frac{d\theta}{2\pi} \right).$$

If owing to a lack of correct information, one considers Y_1^n distributed as $q_n = N(0, Q_n)$. We can define the relative entropy (or the KL divergence) between p_n and q_n as

$$S(p_n || q_n) = \int p_n \log \frac{p_n}{q_n} dy.$$

Plug in the probability density functions for p_n and q_n . After some simple manipulations, we get

$$S(p_n || q_n) = \frac{1}{2} (\log \det(Q_n) - \log \det(P_n) + \text{tr}(Q_n^{-1} P_n) - n). \quad (2.16)$$

As n goes to ∞ , we define the notion of relative entropy rate as

$$S_\infty(p || q) := \lim_{n \rightarrow \infty} \frac{1}{n} S(p_n || q_n)$$

if the limit exists. The following proposition evaluates the relative entropy rate for the Gaussian random process $\{Y_k\}$. We shall restrict ourselves to the set of power spectral densities

$$\mathcal{F} := \{f : f \text{ differentiable on } [-\pi, \pi], f(\theta) > 0, \text{ and } f, f^{-1} \text{ bounded}\}.$$

Proposition 2. *Let $f_1(\theta)$ and $f_2(\theta)$ denote power spectral densities corresponding to the Toeplitz covariance matrices P_n and Q_n respectively, $P_n, Q_n \in \mathcal{T}_n$, ($n \in \mathbb{Z}$), and $f_1, f_2 \in \mathcal{F}$, then the relative entropy rate is half the Itakura-Saito distance $d_{\text{IS}}(f_1, f_2)$ [25]:*

$$S_\infty(p || q) = \frac{1}{2} d_{\text{IS}}(f_1, f_2),$$

and

$$d_{\text{IS}}(f_1, f_2) = \int_{-\pi}^{\pi} \left(\frac{f_1(\theta)}{f_2(\theta)} - \log \frac{f_1(\theta)}{f_2(\theta)} - 1 \right) \frac{d\theta}{2\pi}.$$

Proof. According to the fundamental eigenvalue distribution theorem of Szegö [44, 45], we have

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log \det(P_n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \log f_1(\theta) d\theta \quad (2.17)$$

$$\lim_{n \rightarrow \infty} \frac{1}{n} \log \det(Q_n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \log f_2(\theta) d\theta. \quad (2.18)$$

Also since $f_2(\boldsymbol{\theta})$ is strictly positive, $f_2^{-1}(\boldsymbol{\theta})$ is the power spectral density corresponding to Q_n^{-1} as n goes to ∞ [45, pg. 63]. Likewise, the product of two Toeplitz matrices $Q_n^{-1}P_n$ is asymptotically equivalent to the Toeplitz matrix corresponding to $f_1(\boldsymbol{\theta})/f_2(\boldsymbol{\theta})$. Consequently

$$\lim_{n \rightarrow \infty} \frac{1}{n} \text{tr}(Q_n^{-1}P_n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{f_1(\boldsymbol{\theta})}{f_2(\boldsymbol{\theta})} d\boldsymbol{\theta}. \quad (2.19)$$

Substituting (2.17) and (2.19) into (2.16), we obtain

$$\lim_{n \rightarrow \infty} \frac{1}{n} S(p_n || q_n) = \frac{1}{2} \int_{-\pi}^{\pi} \left(\frac{f_1(\boldsymbol{\theta})}{f_2(\boldsymbol{\theta})} - \log \frac{f_1(\boldsymbol{\theta})}{f_2(\boldsymbol{\theta})} - 1 \right) \frac{d\boldsymbol{\theta}}{2\pi} = \frac{1}{2} d_{\text{IS}}(f_1, f_2).$$

□

As mentioned in Section 2.1, the Riemannian metric induced by the Itakura-Saito distance is

$$\mathfrak{g}_{f, \text{IS}}(\boldsymbol{\delta}) = \|f^{-1} \boldsymbol{\delta}\|_2^2,$$

and the unique geodesic path from f_0 to f_1 is parameterized by

$$f_\tau = f_0^{1-\tau} f_1^\tau, \quad \tau \in [0, 1],$$

with the geodesic distance

$$d_{\text{Log}}(f_0, f_1) = \|\log f_0 - \log f_1\|_2.$$

Remark 3. When $\{Y_k, k \in \mathbb{Z}\}$ is an m -dimensional stationary zero-mean Gaussian random process, a similar result as in Proposition 2 can be established [46, 47, 48], where the Itakura-Saito distance between two matrix-valued power spectral densities is

$$d_{\text{IS}}(f_1, f_2) = \int_{-\pi}^{\pi} (\text{tr}(f_1(\boldsymbol{\theta})f_2^{-1}(\boldsymbol{\theta})) - \log \det(f_1(\boldsymbol{\theta})f_2^{-1}(\boldsymbol{\theta})) - m) \frac{d\boldsymbol{\theta}}{2\pi}.$$

It's not hard to derive that the infinitesimal perturbation from a given $f(\boldsymbol{\theta})$ leads to a Riemannian metric

$$\mathfrak{g}_{f, \text{IS}}(\boldsymbol{\delta}) = \int \|f^{-1/2} \boldsymbol{\delta} f^{-1/2}\|_F^2 \frac{d\boldsymbol{\theta}}{2\pi}. \quad (2.20)$$

For probability distributions $f(x, \boldsymbol{\theta})$ parameterized by a vector $\boldsymbol{\theta}$, the corresponding metric is often referred to as the Fisher-Rao metric [49]

$$\mathfrak{g}_{f, \text{Fisher-Rao}}(\boldsymbol{\delta}_\theta) = \boldsymbol{\delta}_\theta' E \left[\left(\frac{\partial \log f}{\partial \boldsymbol{\theta}} \right) \left(\frac{\partial \log f}{\partial \boldsymbol{\theta}} \right)' \right] \boldsymbol{\delta}_\theta,$$

while for zero-mean Gaussian distributions parameterized by the corresponding covariance matrices, the metric becomes

$$\mathfrak{g}_{P,\text{Rao}}(\Delta) = \left\| P^{-1/2} \Delta P^{-1/2} \right\|_F^2 \quad (2.21)$$

and is often named after C.R. Rao. We summarize this below.

Proposition 4. *Consider a zero-mean, normal distribution p with covariance $P > 0$, and a perturbation p_ε with covariance $P + \varepsilon\Delta$. Provided $\|P^{-1/2} \varepsilon\Delta P^{-1/2}\|_F < 1$,*

$$S(p||p_\varepsilon) = \frac{1}{4} \mathfrak{g}_{P,\text{Rao}}(\varepsilon\Delta) + O(\varepsilon^3). \quad (2.22)$$

Moreover, for $\delta = p_\varepsilon - p$,

$$\mathfrak{g}_{P,\text{Rao}}(\varepsilon\Delta) = 2\mathfrak{g}_{p,\text{Fisher}}(\delta) + O(\varepsilon^4), \quad (2.23)$$

where $\mathfrak{g}_{p,\text{Fisher}}(\delta)$ is the Fisher metric in (2.1).

Proof. The relative entropy between p and p_ε is

$$S(p||p_\varepsilon) = \frac{1}{2} \left(\log \det(P + \varepsilon\Delta) - \log \det(P) + \text{tr} \left((P + \varepsilon\Delta)^{-1} P \right) - n \right)$$

Define $\Delta_P = P^{-1/2} \Delta P^{-1/2}$, then

$$\begin{aligned} S(p||p_\varepsilon) &= \frac{1}{2} \left(\log \det \left(P^{1/2} (I + \varepsilon\Delta_P) P^{1/2} \right) - \log \det(P) + \text{tr} \left(P^{-1/2} (I + \varepsilon\Delta_P)^{-1} P^{-1/2} P \right) - n \right) \\ &= \frac{1}{2} \left(\log \det(I + \varepsilon\Delta_P) + \text{tr} (I + \varepsilon\Delta_P)^{-1} - n \right). \end{aligned} \quad (2.24)$$

We expand $(I + \varepsilon\Delta_P)^{-1}$ into the Taylor series

$$(I + \varepsilon\Delta_P)^{-1} = I - \varepsilon\Delta_P + \varepsilon^2 \Delta_P^2 - \varepsilon^3 \Delta_P^3 + \dots \quad (2.25)$$

Let λ_i , $i = 1, \dots, n$ represent eigenvalues of Δ_P , then

$$\begin{aligned} \log \det(I + \varepsilon\Delta_P) &= \sum_{i=1}^n \log(1 + \varepsilon\lambda_i) \\ &= \sum_{i=1}^n \left(\varepsilon\lambda_i - \frac{1}{2} \varepsilon^2 \lambda_i^2 + \frac{1}{3} \varepsilon^3 \lambda_i^3 + \dots \right) \\ &= \varepsilon \text{tr}(\Delta_P) - \frac{1}{2} \varepsilon^2 \text{tr}(\Delta_P^2) + \frac{1}{3} \varepsilon^3 \text{tr}(\Delta_P^3) + \dots \end{aligned} \quad (2.26)$$

We substitute (2.25) and (2.26) into (2.24) to obtain

$$S(p||p_\varepsilon) = \frac{1}{4}\varepsilon^2 \text{tr}(\Delta_P^2) + O(\varepsilon^3).$$

By following the similar computation, one can easily see that $S(p_\varepsilon||p)$ gives rise to the same metric, though the coefficients of higher order terms are different from those corresponding to $S(p||p + \delta)$.

To draw a connection with the Fisher metric, we substitute $\delta = N(0, P + \Delta) - N(0, P)$ into the Fisher metric:

$$g_{p, \text{Fisher}}(\delta) = \left(\int \frac{\det(P)^{1/2}}{(2\pi)^{n/2} \det(P + \varepsilon\Delta)} e^{-\frac{1}{2}y'(2(P + \varepsilon\Delta)^{-1} - P^{-1})y} dy - 1 \right).$$

Since $\|\varepsilon\Delta_P\|_F < 1$, $\varepsilon^2\Delta_P^2 < I$, and hence

$$-I < \varepsilon\Delta_P < I.$$

Multiplying $P^{1/2}$ from left and right on both sides of the above inequality, we get

$$-P < \varepsilon\Delta < P,$$

or equivalently

$$0 < \frac{1}{2}P + \frac{1}{2}\varepsilon\Delta < P.$$

It follows that

$$-\left(\frac{1}{2}P + \frac{1}{2}\varepsilon\Delta\right)^{-1} < -P^{-1},$$

or equivalently

$$2(P + \varepsilon\Delta)^{-1} - P^{-1} > 0.$$

Consequently

$$\frac{1}{(2\pi)^{n/2} \det(2(P + \varepsilon\Delta)^{-1} - P^{-1})^{-1/2}} e^{-\frac{1}{2}y'(2(P + \varepsilon\Delta)^{-1} - P^{-1})y}$$

is a Gaussian distribution with mean 0 and covariance $(2(P + \varepsilon\Delta)^{-1} - P^{-1})^{-1}$. Since the integration of a Gaussian distribution is 1, we obtain

$$g_{p, \text{Fisher}}(\delta) = \left(\frac{\det(P)^{1/2}}{\det(2(P + \varepsilon\Delta)^{-1} - P^{-1})^{1/2} \det(P + \varepsilon\Delta)} - 1 \right).$$

Since

$$(P + \varepsilon\Delta)^{-1} = P^{-1/2}(I + \varepsilon\Delta_P)^{-1}P^{-1/2},$$

and

$$2(P + \varepsilon\Delta)^{-1} - P^{-1} = P^{-1/2} (2(I + \varepsilon\Delta_P)^{-1} - I) P^{-1/2}.$$

Consequently

$$\begin{aligned} \det(2(P\varepsilon\Delta^{-1} - P^{-1})^{1/2} \det(P + \varepsilon\Delta) &= \det((P + \varepsilon\Delta) (2(P + \varepsilon\Delta)^{-1} - P^{-1}) (P + \varepsilon\Delta))^{1/2} \\ &= \det\left(P^{1/2}(I + \varepsilon\Delta_P) (2(I + \varepsilon\Delta_P)^{-1} - I) (I + \varepsilon\Delta_P)P^{1/2}\right)^{1/2} \\ &= \det(P)^{1/2} \det(I - \varepsilon^2\Delta_P^2)^{1/2}, \end{aligned}$$

and

$$g_{P,\text{Fisher}}(\delta) = \left(\det(I - \varepsilon^2\Delta_P^2)^{-1/2} - 1\right) = \left(\det(I + \varepsilon^2\Delta_P^2 + \varepsilon^4\Delta_P^4 + \dots)^{1/2} - 1\right).$$

Using eigenvalues of Δ_P , we get

$$\begin{aligned} \det(I + \varepsilon^2\Delta_P^2 + \varepsilon^4\Delta_P^4 + \dots)^{1/2} &= \left(\prod_{k=1}^n \left(\sum_{i=0}^{\infty} (\varepsilon\lambda_k)^{2i}\right)\right)^{1/2} \\ &= \left(1 + \sum_{k=1}^n \varepsilon^2\lambda_k^2 + \sum_{k \leq l} \varepsilon^4\lambda_k^2\lambda_l^2 + \dots\right)^{1/2} \\ &= 1 + \frac{1}{2}\varepsilon^2\|\Delta_P\|_F^2 + O(\varepsilon^4), \end{aligned}$$

where in the last equality we used the fact that $\sum_{k=1}^n \lambda_k^2 = \text{tr}(\Delta_P^2) = \|\Delta_P\|_F^2$. Therefore,

$$g_{P,\text{Fisher}}(\delta) = \frac{1}{2}g_{P,\text{Rao}}(\Delta) + O(\varepsilon^4).$$

□

The geometry of the Rao metric has been introduced in Section 2.2. Here we shall only mention a simple observation.

Corollary 5. *The geodesic path induced from the Fisher metric on the manifold of standard normal distributions is*

$$p_\tau = N(0, P_\tau)$$

where

$$P_\tau = P_0^{1/2} \left(P_0^{-1/2} P_1 P_0^{-1/2} \right)^\tau P_0^{1/2},$$

for two end points $p_0 = N(0, P_0)$ and $p_1 = N(0, P_1)$, and the geodesic length is

$$L(p_\tau) = \|\log(P_0^{-1/2} P_1 P_0^{-1/2})\|_F.$$

Chapter 3

Metrics on Multivariate Power Spectra

In Chapter 2, we have reviewed various metrics that have been proposed to quantify the difference between two scalar-valued power spectra. Distances between multivariate power spectra have only recently received any attention. In this direction we mention generalizations of the Hellinger and Itakura-Saito distances by Ferrante *et al.* [46, 40, 50, 51] and the use of the Umegaki-von Neumann relative entropy [52].

In this chapter, we generalize the geometric framework in [17] to the matrix-valued power spectra [53]. We compare two power spectra in the context of linear prediction: a choice between the two is used to design an optimal filter which is then applied to a process corresponding to the second power spectrum. The “flatness” of the innovations process, as well as the degradation of the prediction error variance, when compared to the best possible, are used to quantify the mismatch between the two. This rationale provides us with natural divergence measures. We then identify corresponding Riemannian metrics that dictate the underlying geometry. For a certain case we compute closed-form expressions for the induced geodesics and geodesic distances. These provide a multivariable counterpart to the logarithmic intervals in [17] and the logarithmic spectral deviation [19, page 370]. It is noted that the geodesic distance has certain natural desirable properties; it is inverse-invariant and congruence-invariant. Moreover, the manifold of the multivariate spectral density functions endowed with this geodesic distance is a complete metric space. A discrete counter part of certain of these Riemannian metrics, on the manifold of positive definite matrices (equivalent to power spectra which are constant across frequencies), has been studied extensively in connection to the geometry of positive operators [35] and relates to the Rao-Fisher geometry on probability models restricted to the case

of Gaussian random vectors.

3.1 Preliminaries on Multivariate Prediction

Consider a multivariate discrete-time, zero mean, weakly stationary stochastic process $\{\mathbf{u}(k), k \in \mathbb{Z}\}$ with $\mathbf{u}(k)$ taking values in $\mathbb{C}^{m \times 1}$. Throughout, boldface denotes random variables/vectors, E denotes expectation, $j = \sqrt{-1}$ the imaginary unit, and $*$ the complex conjugate transpose. Let

$$R_k = E \{ \mathbf{u}(\ell) \mathbf{u}^*(\ell - k) \} \text{ for } l, k \in \mathbb{Z}$$

denote the sequence of matrix covariances and $d\mu(\theta)$ be the corresponding matricial power spectral measure for which

$$R_k = \int_{-\pi}^{\pi} e^{-jk\theta} \frac{d\mu(\theta)}{2\pi}.$$

For the most part, we will be concerned with the case of non-deterministic processes of full rank with an absolutely continuous power spectrum. Hence, unless we specifically indicate otherwise, $d\mu(\theta) = f(\theta)d\theta$ with $f(\theta)$ being a matrix-valued power spectral density (PSD) function. Further, for such a non-deterministic process $\log \det f(\theta)$ needs to be integrable, and this will be assumed throughout as well.

Our interest is in comparing PSD's and in studying possible metrics between such. The evident goal is to provide a means to quantify deviations and uncertainty in the spectral domain in a way that is consistent with particular applications. More specifically, we present metrizations of the space of PSD's which are dictated by optimal prediction and reflect dissimilarities that have an impact on the quality of prediction.

3.1.1 Geometry of multivariable processes

We will be considering least-variance linear prediction problems. To this end, we define $L_{2,\mathbf{u}}$ to be the closure of $m \times 1$ -vector-valued finite linear combinations of $\{\mathbf{u}(k)\}$ with respect to convergence in the mean [54, pg. 135]:

$$L_{2,\mathbf{u}} := \overline{\left\{ \sum_{\text{finite}} P_k \mathbf{u}(-k) : P_k \in \mathbb{C}^{m \times m}, k \in \mathbb{Z} \right\}}.$$

Here, “bar” denotes closure. The indices in P_k and $\mathbf{u}(-k)$ run in opposite directions so as to simplify the notation later on where prediction is based on past observations. This space is

endowed with both, a matricial inner product

$$\llbracket \sum_k P_k \mathbf{u}(-k), \sum_k Q_k \mathbf{u}(-k) \rrbracket := E \left\{ \left(\sum_k P_k \mathbf{u}(-k) \right) \left(\sum_k Q_k \mathbf{u}(-k) \right)^* \right\},$$

as well as a scalar inner product

$$\langle \sum_k P_k \mathbf{u}(-k), \sum_k Q_k \mathbf{u}(-k) \rangle := \text{tr} \llbracket \sum_k P_k \mathbf{u}(-k), \sum_k Q_k \mathbf{u}(-k) \rrbracket.$$

Throughout, “tr” denotes the trace of a matrix. It is standard to establish the correspondence between

$$\begin{aligned} \mathbf{p} &:= p(\mathbf{u}) := \sum_k P_k \mathbf{u}(-k) \text{ and} \\ p(z) &:= \sum_k P_k z^k \end{aligned}$$

with $z = e^{j\theta}$ for $\theta \in [-\pi, \pi]$. A time-shift on \mathbf{p} generates the process

$$\mathbf{p}(\ell) := \sum_k P_k \mathbf{u}(\ell - k), \text{ with } \ell \in \mathbb{Z},$$

having

$$p(e^{j\theta}) \frac{d\mu(\theta)}{2\pi} p(e^{j\theta})^*$$

as its power spectrum. The correspondence is the Kolmogorov isomorphism between the “temporal” space $L_2(\mathbf{u})$ and “spectral” space $L_{2,d\mu}$,

$$\varphi : L_2(\mathbf{u}) \rightarrow L_{2,d\mu} : \sum_k P_k \mathbf{u}(-k) \mapsto \sum_k P_k z^k.$$

It is convenient to endow the latter space $L_{2,d\mu}$ with the matricial inner product

$$\llbracket p, q \rrbracket_{d\mu} := \int_{-\pi}^{\pi} \left(p(e^{j\theta}) \frac{d\mu(\theta)}{2\pi} q(e^{j\theta})^* \right)$$

as well as the scalar inner product

$$\langle p, q \rangle_{d\mu} := \text{tr} \llbracket p, q \rrbracket_{d\mu}.$$

The additional structure due to the matricial inner product is often referred to as *Hilbertian* (as opposed to *Hilbert*) [55].

Throughout, $p(e^{j\theta}) = \sum_k P_k e^{jk\theta}$, $q(e^{j\theta}) = \sum_k Q_k e^{jk\theta}$, where we use lower case p, q for matrix functions and upper case P_k, Q_k for their matrix coefficients. For non-deterministic processes of full rank with absolutely continuous spectral measure $d\mu(\theta) = f(\theta)d\theta$, we simplify the notation into

$$\begin{aligned} \llbracket p, q \rrbracket_f &:= \llbracket p, q \rrbracket_{fd\theta}, \text{ and} \\ \langle p, q \rangle_f &:= \langle p, q \rangle_{fd\theta}. \end{aligned}$$

Least-variance linear prediction

$$\min \left\{ \text{tr} E \{ \mathbf{p} \mathbf{p}^* \} : \mathbf{p} = \mathbf{u}(0) - \sum_{k>0} P_k \mathbf{u}(-k), P_k \in \mathbb{C}^{m \times m} \right\} \quad (3.1)$$

can be expressed equivalently in the spectral domain

$$\min \left\{ \llbracket p, p \rrbracket_f : p(z) = I - \sum_{k>0} P_k z^k, P_k \in \mathbb{C}^{m \times m} \right\} \quad (3.2)$$

where the minimum is sought in the positive-definite sense, see [55, pg. 354], [54, pg. 143]. We use “ I ” to denote the identity matrix of suitable size. It holds that, although non-negative definiteness defines only a partial order on the cone of non-negative definite Hermitian matrices, a minimizer for (3.1) always exists. Of course this corresponds to a minimizer for (3.2). The existence of a minimizer is due to the fact that $\text{tr} E \{ \mathbf{p} \mathbf{p}^* \}$ is matrix-convex. Here $d\mu = f d\theta$ is an absolutely continuous measure and the quadratic form is not degenerate; see [56, Proposition 1] for a detailed analysis and a treatment of the singular case where μ is a discrete matrix-valued measure. Further, the minimizer of (3.1) coincides with the minimizer of

$$\min \left\{ \langle p, p \rangle_f : p(z) = I - \sum_{k>0} P_k z^k, P_k \in \mathbb{C}^{m \times m} \right\}. \quad (3.3)$$

From here on, to keep notation simple, $p(z)$ will denote the minimizer of such a problem, with f specified accordingly, and the minimal matrix of (3.1) will be denoted by Ω . That is,

$$\Omega := \llbracket p, p \rrbracket_f$$

while the minimal value of (3.3) is $\text{tr} \Omega$. The minimizer p is precisely the image under the Kolmogorov isomorphism of the optimal *prediction error* \mathbf{p} and Ω the *prediction-error variance*.

3.1.2 Spectral factors and optimal prediction

For a non-deterministic process of full rank, the determinant of the error variance Ω is non-zero. Equivalently, the product of its eigenvalues is non-zero. The well-known Szegö-Kolmogorov formula [55, pg. 369]

$$\det \Omega = \exp \left\{ \int_{-\pi}^{\pi} \log \det f(\theta) \frac{d\theta}{2\pi} \right\} \quad (3.4)$$

relates the product of the eigenvalues of the optimal one-step-ahead prediction error variance with the corresponding PSD. No expression is available in general that would relate f to Ω directly in the matricial case. It is only the product of the eigenvalues that is known to relate to the geometric mean of the determinant of the power spectrum as given above.

We consider only non-deterministic processes of full rank and hence we assume that

$$\log \det f(\theta) \in L_1[-\pi, \pi].$$

In this case, $f(\theta)$ admits a unique factorization

$$f(\theta) = f_+(e^{j\theta}) f_+(e^{j\theta})^*, \quad (3.5)$$

with $f_+(e^{j\theta}) \in \mathcal{H}_2^{m \times m}(\mathbb{D})$,

$$\det(f_+(z)) \neq 0 \text{ in } \mathbb{D} := \{z : |z| < 1\},$$

and normalized so that $f_+(0) = \Omega^{\frac{1}{2}}$. Throughout, $M^{\frac{1}{2}}$ denotes the Hermitian square root of a Hermitian matrix M . The factor f_+ is known as the *canonical (left) spectral factor*. In the case where f is a scalar function ($m = 1$) the canonical spectral factor is explicitly given by

$$f_+(z) = \exp \left\{ \frac{1}{2} \int_{-\pi}^{\pi} \left(\frac{1 + ze^{-j\theta}}{1 - ze^{-j\theta}} \right) \log f(\theta) \frac{d\theta}{2\pi} \right\}, \quad |z| < 1,$$

As usual, $\mathcal{H}_2(\mathbb{D})$ denotes the Hardy space of functions which are analytic in the unit disk \mathbb{D} with square-integrable radial limits. Spectral factorization presents an “explicit” expression of the optimal prediction error in the form

$$p(z) = f_+(0) f_+^{-1}(z). \quad (3.6)$$

Thus, $p(z)^{-1}$ is a “normalized” (left) *outer* factor of f . The terminology “outer” refers to a (matrix-valued) function $g(e^{j\theta})$ for $\theta \in [-\pi, \pi]$ that can be extended into an analytic function

in the open interior of the unit disc \mathbb{D} which is also invertible in \mathbb{D} . It is often standard not to differentiate between such a function in \mathbb{D} and the function on the boundary of radial-limits since these are uniquely defined from one another. In the engineering literature outer functions are also referred to as “minimum phase.”

Remark 6. *Right-outer factors, where $f(\theta) = f_{+,right}(e^{j\theta})^* f_{+,right}(e^{j\theta})$ instead of (3.5) relate to a post-diction optimal estimation problem; in this, the present value of the process is estimated via linear combination of future values (see e.g., [56]). Only left factorizations will be used in the present paper.*

3.2 Comparison of PSD's

We present two complementing viewpoints on how to compare two PSD's, f_1 and f_2 . In both, the optimal one-step-ahead predictor for one of the two stochastic processes, is applied to the other and compared to the corresponding optimal. The first is to consider how “white” the power spectrum of the innovations' process is. The second viewpoint is to compare how the error variance degrades with respect to the optimal predictor. Either principle provides a family of divergence measures and a suitable generalization of the Riemannian geometry of scalar PSD's given in [17]. There is a close relationship between the two.

3.2.1 Prediction errors and innovations processes

Consider two matrix-valued spectral density functions f_1 and f_2 . Since an optimal filter will be designed based on one of the two and then evaluated with respect to the other, some notation is in order.

First, let us use a subscript to distinguish between two processes $\mathbf{u}_i(k)$, $i \in \{1, 2\}$, having the f_i 's as the corresponding PSD's. They are assumed nondeterministic of full rank, vector-valued, and of compatible size. The optimal filters in the spectral domain are

$$p_i := \operatorname{argmin}\{ \llbracket p, p \rrbracket_{f_i} \mid p(0) = I, \\ \text{and } p \in \mathcal{H}_2^{m \times m}(\mathbb{D}) \},$$

and their respective error covariances

$$\Omega_i := \llbracket p_i, p_i \rrbracket_{f_i}.$$

Now define

$$\Omega_{i,j} := \llbracket p_j, p_j \rrbracket_{f_i}.$$

Clearly, $\Omega_{i,j}$ is the variance of the prediction error when the filter p_j is used on a process having power spectrum f_i . Indeed, if we set

$$\mathbf{p}_{i,j} := \mathbf{u}_i(0) - P_{j,1}\mathbf{u}_i(-1) - P_{j,2}\mathbf{u}_i(-2) - \dots \quad (3.7)$$

the prediction-error covariance is

$$\llbracket \mathbf{p}_{i,j}, \mathbf{p}_{i,j} \rrbracket = \llbracket p_j, p_j \rrbracket_{f_i}.$$

The prediction error $\mathbf{p}_{i,j}$ can also be thought of as a time-process, indexed at time-instant $k \in \mathbb{Z}$,

$$\mathbf{p}_{i,j}(k) := \mathbf{u}_i(k) - P_{j,1}\mathbf{u}_i(k-1) - P_{j,2}\mathbf{u}_i(k-2) - \dots$$

for $i, j \in \{1, 2\}$. This is an *innovations process*. Clearly, from stationarity,

$$\llbracket \mathbf{p}_{i,i}, \mathbf{p}_{i,i} \rrbracket = \Omega_i,$$

whereas

$$\llbracket \mathbf{p}_{i,j}, \mathbf{p}_{i,j} \rrbracket \geq \Omega_i,$$

since in this case p_j is suboptimal for \mathbf{u}_i , in general.

3.2.2 The color of innovations and PSD mismatch

We choose to normalize the innovations processes as follows:

$$\mathbf{h}_{i,j}(k) = \Omega_j^{-\frac{1}{2}} \mathbf{p}_{i,j}(k), \text{ for } k \in \mathbb{Z}.$$

The Kolmogorov isomorphism takes

$$\varphi : \mathbf{h}_{i,j}(k) \mapsto f_{j+}^{-1},$$

with the expectation/inner-product being that induced by f_i , and hence, the power spectral density of the process $\mathbf{h}_{i,j}(k)$ is

$$f_{\mathbf{h}_{ij}} = f_{j+}^{-1} f_i f_{j+}^{-*},$$

where $(\cdot)^{-*}$ is a shorthand for $((\cdot)^*)^{-1}$. When $f_i = f_j$, evidently $\{\mathbf{h}_k^{i,i}\}$ is a white noise process with covariance matrix equals to the identity.

Naturally, in an absolute sense, the mismatch between the two power spectra f_i, f_j can be quantified by the distance of $f_{\mathbf{h}_{ij}}$ to the identity. To this end we may consider any symmetrized expression:

$$\int_{-\pi}^{\pi} \mathbf{d}(f_{j+}^{-1} f_i f_{j+}^{-*}, I) \frac{d\theta}{2\pi} + \int_{-\pi}^{\pi} \mathbf{d}(f_{i+}^{-1} f_j f_{i+}^{-*}, I) \frac{d\theta}{2\pi} \quad (3.8)$$

for a suitable distance $\mathbf{d}(\cdot, \cdot)$ between positive definite matrices. In general, it is deemed desirable that distances between power spectra are invariant to scaling (as is the case when distances depend on ratios of spectra, [19]). Researchers and practitioners alike have insisted on such a property, especially for speech and image systems, due to an apparent agreement with subjective qualities of sound and images. It is thus interesting to seek a multivariable analogues inherent in the above comparison.

Due to the non-negative definiteness of power spectra, a convenient option is to take “d” as the trace:

$$\int_{-\pi}^{\pi} \text{tr} \left(f_{j+}^{-1} f_i f_{j+}^{-*} - I \right) + \text{tr} \left(f_{i+}^{-1} f_j f_{i+}^{-*} - I \right) \frac{d\theta}{2\pi}.$$

This indeed defines a distance measure since $(x + x^{-1} - 2)$ is a non-negative function for $0 < x \in \mathbb{R}$ that vanishes only when $x = 1$. Thus, we define

$$\mathbf{D}_1(f_1, f_2) := \int_{-\pi}^{\pi} \text{tr} \left(f_2^{-1} f_1 + f_1^{-1} f_2 - 2I \right) \frac{d\theta}{2\pi}. \quad (3.9a)$$

Interestingly, $\mathbf{D}_1(f_1, f_2)$ can be re-written as follows:

$$\mathbf{D}_1(f_1, f_2) = \int_{-\pi}^{\pi} \|f_1^{-1/2} f_2^{1/2} - f_1^{1/2} f_2^{-1/2}\|_{\text{Fr}}^2 \frac{d\theta}{2\pi} \quad (3.9b)$$

where $\|M\|_{\text{Fr}}^2 := \text{tr}MM^*$ denotes the square of the Frobenius norm¹. It can be readily verified starting from the right hand side of (3.9b) and simplifying this to match (3.9a). It is now easily seen that $\mathbf{D}_1(f_i, f_j)$ has a number of desirable properties listed in the following proposition.

Proposition 7. *Consider f_i, f_j being PSD's of non-deterministic processes of full rank and $g(e^{j\theta})$ an arbitrary outer matrix-valued function in $\mathcal{H}_2^{m \times m}(\mathbb{D})$. The following properties hold:*

- (i) $\mathbf{D}_1(f_i, f_j) \geq 0$.

¹ $\sqrt{\text{tr}MM^*}$ is also referred to as the Hilbert-Schmidt norm.

$$(ii) \quad \mathbf{D}_1(f_i, f_j) = 0 \text{ iff } f_i = f_j \text{ (a.e.)}.$$

$$(iii) \quad \mathbf{D}_1(f_i, f_j) = \mathbf{D}_1(f_j, f_i).$$

$$(iv) \quad \mathbf{D}_1(f_i, f_j) = \mathbf{D}_1(f_i^{-1}, f_j^{-1}).$$

$$(v) \quad \mathbf{D}_1(f_i, f_j) = \mathbf{D}_1(gf_i g^*, gf_j g^*).$$

Proof. Properties (i-iv) follow immediately from (3.9b) while the invariance property (v) is most easily seen by employing (3.9a). \square

3.2.3 Suboptimal prediction and PSD mismatch

We now attempt to quantify how suboptimal the performance of a filter is when this is based on the incorrect choice between the two alternative PSD's. To this end, we consider the error covariance and compare it to that of the optimal predictor. A basic inequality between these error covariances is summarized in the following proposition.

Proposition 8. *Under our earlier standard assumptions, for $i, j \in \{1, 2\}$ and $\Omega_i, \Omega_j > 0$, it holds that*

$$\Omega_{i,j} \geq \Omega_i. \quad (3.10a)$$

Further, the above holds as an equality iff $p_i = p_j$.

Proof. It follows from the optimality of p_i since

$$\llbracket p_j, p_j \rrbracket_{f_i} \geq \llbracket p_i, p_i \rrbracket_{f_i} = \Omega_i.$$

\square

Corollary 9. *The following hold:*

$$\Omega_i^{-\frac{1}{2}} \Omega_{i,j} \Omega_i^{-\frac{1}{2}} \geq I \quad (3.10b)$$

$$\det(\Omega_{i,j}) \geq \det(\Omega_i) \quad (3.10c)$$

$$\text{tr}(\Omega_{i,j}) \geq \text{tr}(\Omega_i) \quad (3.10d)$$

$$\Omega_j^{-\frac{1}{2}} \Omega_{i,j} \Omega_j^{-\frac{1}{2}} \geq \Omega_j^{-\frac{1}{2}} \Omega_i \Omega_j^{-\frac{1}{2}}. \quad (3.10e)$$

Further, each “ \geq ” holds as equality iff $p_i = p_j$.

Thus, a mismatch between the two spectral densities can be quantified by the strength of the above inequalities. To this end, we may consider a number of alternative “divergence measures”. First we consider:

$$D_2(f_i, f_j) := \log \det \left(\Omega_i^{-\frac{1}{2}} \Omega_{i,j} \Omega_i^{-\frac{1}{2}} \right). \quad (3.11)$$

Equivalent options leading to the same Riemannian structure are:

$$\frac{1}{m} \operatorname{tr}(\Omega_i^{-\frac{1}{2}} \Omega_{i,j} \Omega_i^{-\frac{1}{2}}) - 1, \text{ and} \quad (3.12a)$$

$$\det(\Omega_i^{-\frac{1}{2}} \Omega_{i,j} \Omega_i^{-\frac{1}{2}}) - 1. \quad (3.12b)$$

Using the generalized Szegő-Kolmogorov expression (3.4) we readily obtain that

$$\begin{aligned} D_2(f_i, f_j) &= \quad (3.13) \\ &= \log \det \left(\int_{-\pi}^{\pi} f_{j+}^{-1} f_i f_{j+}^{-*} \frac{d\theta}{2\pi} \right) - \int_{-\pi}^{\pi} \log \det \left(f_{j+}^{-1} f_i f_{j+}^{-*} \right) \frac{d\theta}{2\pi} \\ &= \operatorname{tr} \left(\log \int_{-\pi}^{\pi} f_{j+}^{-1} f_i f_{j+}^{-*} \frac{d\theta}{2\pi} - \int_{-\pi}^{\pi} \log f_{j+}^{-1} f_i f_{j+}^{-*} \frac{d\theta}{2\pi} \right). \end{aligned}$$

This expression takes values in $[0, \infty]$, and is zero if and only if the normalized spectral factors $p^{-1} = \Omega^{-1/2} f_+$ are identical for the two spectra. Further, the formula in (3.13) provides a natural generalization of the divergence measures in [17, Eq. (10b)] and of the Itakura distance to the case of multivariable spectra. It satisfies “congruence invariance.” This is stated next.

Proposition 10. *Consider two PSD’s f_i, f_j of non-deterministic processes of full rank and $g(e^{j\theta})$ an outer matrix-valued function in $\mathcal{H}_2^{m \times m}(\mathbb{D})$. The following hold:*

- (i) $D_2(f_i, f_j) \geq 0$.
- (ii) $D_2(f_i, f_j) = 0$ iff $p_i = p_j$.
- (iii) $D_2(f_i, f_j) = D_2(g f_i g^*, g f_j g^*)$.

Proof. Properties (i-ii) follow immediately from (3.11) while the invariance property (iii) is most easily seen by employing (3.13). To this end, first note that $g f_+$ obviously constitutes the spectral factor of $g f g^*$. Substituting the corresponding expressions in (3.13) establishes the invariance. \square

3.2.4 Alternative divergence measures

Obviously, a large family of divergence measures between two matrix-valued power spectra can be obtained based on (3.8). For completeness, we suggest representative possibilities some of which have been independently considered in recent literature.

Frobenius distance

If we use the Frobenius norm in (3.8) we obtain

$$D_{\text{F}}(f_1, f_2) := \frac{1}{2} \sum_{i,j} \int_{-\pi}^{\pi} \|f_{j+}^{-1} f_i f_{j+}^{-*} - I\|_{\text{Fr}}^2 \frac{d\theta}{2\pi} \quad (3.14a)$$

where $\sum_{i,j}$ designates the ‘‘symmetrized sum’’ taking $(i, j) \in \{(1, 2), (2, 1)\}$. It’s straightforward to see that all of

$$f_{j+}^{-1} f_i f_{j+}^{-*}, f_j^{-\frac{1}{2}} f_i f_j^{-\frac{1}{2}} \text{ and } f_j^{-1} f_i$$

share the same eigenvalues for any $\theta \in [-\pi, \pi]$. Thus,

$$\|f_{j+}^{-1} f_i f_{j+}^{-*} - I\|_{\text{Fr}}^2 = \|f_j^{-\frac{1}{2}} f_i f_j^{-\frac{1}{2}} - I\|_{\text{Fr}}^2,$$

and

$$D_{\text{F}}(f_1, f_2) = \frac{1}{2} \sum_{i,j} \int_{-\pi}^{\pi} \|f_j^{-\frac{1}{2}} f_i f_j^{-\frac{1}{2}} - I\|_{\text{Fr}}^2 \frac{d\theta}{2\pi}. \quad (3.14b)$$

Obviously (3.14b) is preferable over (3.14a) since no spectral factorization is involved.

Hellinger distance

A generalization of the Hellinger distance has been proposed in [40] for comparing multivariable spectra. Briefly, given two positive definite matrices f_1 and f_2 one seeks factorizations $f_i = g_i g_i^*$ so that the integral over frequencies of the Frobenius distance $\|g_1 - g_2\|_{\text{Fr}}^2$ between the factors is minimal. The factorization does not need to correspond to analytic factors. When one of the two spectra is the identity, the optimization is trivial and the Hellinger distance becomes

$$\int_{-\pi}^{\pi} \|f^{\frac{1}{2}} - I\|_{\text{Fr}}^2 \frac{d\theta}{2\pi}.$$

Following a variation of this idea we compare the normalized innovation spectra $(f_{j+}^{-1} f_i f_{j+}^{-*})^{\frac{1}{2}}$, for $i, j \in \{1, 2\}$, to the identity in a symmetrized fashion. Thus, we define

$$\begin{aligned} D_H(f_1, f_2) &:= \sum_{i,j} \int_{-\pi}^{\pi} \|(f_{j+}^{-1} f_i f_{j+}^{-*})^{\frac{1}{2}} - I\|_{\text{Fr}}^2 \frac{d\theta}{2\pi} \\ &= \sum_{i,j} \int_{-\pi}^{\pi} \|(f_j^{-\frac{1}{2}} f_i f_j^{-\frac{1}{2}})^{\frac{1}{2}} - I\|_{\text{Fr}}^2 \frac{d\theta}{2\pi}. \end{aligned} \quad (3.15)$$

The second equality follows by the fact that $f_{j+} f_j^{-\frac{1}{2}}$ is a frequency-dependent unitary matrix.

Multivariable Itakura-Saito distance

The classical Itakura-Saito distance can be readily generalized by taking

$$d(f, I) = \text{tr}(f - \log f - I).$$

The values are always positive for $I \neq f > 0$ and equal to zero when $f = I$. Thus, we may define

$$\begin{aligned} D_{\text{IS}}(f_1, f_2) &= \int_{-\pi}^{\pi} d(f_{2+}^{-1} f_1 f_{2+}^{-*}, I) \frac{d\theta}{2\pi} \\ &= \int_{-\pi}^{\pi} (\text{tr}(f_2^{-1} f_1) - \log \det(f_2^{-1} f_1) - m) \frac{d\theta}{2\pi}. \end{aligned} \quad (3.16)$$

The Itakura-Saito distance has its origins in maximum likelihood estimation for speech processing and is related to the Kullback-Leibler divergence between the probability laws of two Gaussian random processes [19, 47]. The matrix-version of the Itakura-Saito distance has been used for multivariate spectral estimation problems in [46, 50, 51].

Log-spectral deviation

It has been argued that a logarithmic measure of spectral deviations is in agreement with perceptive qualities of sound and for this reason it has formed the basis for the oldest distortion measures considered [19]. In particular, the L_2 distance between the logarithms of power spectra is referred to as ‘‘Log-spectral deviation’’ or the ‘‘logarithmic energy.’’ A natural multivariable version is to consider

$$d(f, I) = \|\log(f)\|_{\text{Fr}}^2.$$

This expression is already symmetrized, since $d(f, I) = d(f^{-1}, I)$ by virtue of the fact that the eigenvalues of $\log(f)$ and those of $\log(f^{-1})$ differ only in their sign. Thereby,

$$\|\log(f_{j+}^{-1} f_i f_{j+}^{-*})\|_{\text{Fr}}^2 = \|\log(f_{i+}^{-1} f_j f_{i+}^{-*})\|_{\text{Fr}}^2.$$

Thus we define

$$\begin{aligned} D_{\text{Log}}(f_1, f_2) &:= \left(\int_{-\pi}^{\pi} \|\log(f_{1+}^{-1} f_2 f_{1+}^{-*})\|_{\text{Fr}}^2 \frac{d\theta}{2\pi} \right)^{1/2} \\ &= \left(\int_{-\pi}^{\pi} \|\log(f_1^{-\frac{1}{2}} f_2 f_1^{-\frac{1}{2}})\|_{\text{Fr}}^2 \frac{d\theta}{2\pi} \right)^{1/2}. \end{aligned} \quad (3.17)$$

This represents a multivariable version of the log-spectral deviation (see [19, page 370]). Interestingly, as we will see later on, $D_{\text{Log}}(f_1, f_2)$ possesses several useful properties and, in fact, turns out to be precisely a geodesic distance in a suitable Riemannian geometry.

3.3 Geometry on Multivariate Spectra

Consider a “small” perturbation $f + \Delta$ away from a nominal power spectral density f . All divergence measures that we have seen so far are continuous in their arguments and, in-the-small, can be approximated by a quadratic form in Δ which depends continuously on f . This is what is referred to as a *Riemannian metric*. The availability of a metric gives the space of power spectral densities its properties. It dictates how perturbations in various directions compare to each other. It also provides additional important concepts: geodesics, geodesic distances, and curvature. Geodesics are paths of shortest length connecting the start to the finish; this length is the geodesic distance. Thus, geodesics in the space of power spectral densities represent deformations from a starting power spectral density f_0 to an end “point” f_1 . Curvature on the other hand is intimately connected with approximation and convexity of sets.

In contrast to a general divergence measure, the geodesic distance obeys the triangular inequality and thus, it is a metric (or, a pseudo-metric when by design it is unaffected by scaling or other group of transformations). Geodesics are also natural structures for modeling changes and deformations. In fact, a key motivation behind the present work is to model time-varying spectra via geodesic paths in a suitable metric space. This viewpoint provides a non-parametric model for non-stationary spectra, analogous to a spectrogram, but one which takes into account the inherent geometry of power spectral densities.

3.3.1 Riemannian structures

In the sequel we consider infinitesimal perturbations about a given power spectral density function. We explain how these give rise to nonnegative definite quadratic forms. Throughout, we

assume that all functions are smooth enough so that the indicated integrals exist. This can be ensured if all spectral density functions are bounded with bounded derivatives and inverses. Thus, we will restrict our attention to the following class of PDF's:

$$\mathcal{F} := \{f \mid m \times m \text{ positive definite, differentiable} \\ \text{on } [-\pi, \pi], \text{ with continuous derivative}\}.$$

In the above, we identify the end points of $[-\pi, \pi]$ since f is thought of as a function on the unit circle. Since the functions f are strictly positive definite and bounded, tangent directions of \mathcal{F} consist of admissible perturbations Δ . These need only be restricted to be differentiable with square integrable derivative, hence the tangent space at any $f \in \mathcal{F}$ can be identified with

$$\mathcal{D} := \{\Delta \mid m \times m \text{ Hermitian, differentiable on } [-\pi, \pi] \\ \text{with continuous derivative}\}.$$

Based on the “flatness” of innovations spectra

We first consider the divergence D_1 in (3.9a-3.9b) which quantifies how far the PSD of the normalized innovations process is from being constant and equal to the identity. The induced Riemannian metric takes the form

$$g_{1,f}(\Delta) := \int_{-\pi}^{\pi} \|f^{-1/2}\Delta f^{-1/2}\|_{\text{Fr}}^2 \frac{d\theta}{2\pi}. \quad (3.18a)$$

Proposition 11. *Let $(f, \Delta) \in \mathcal{F} \times \mathcal{D}$ and $\varepsilon > 0$. Then, for ε sufficiently small,*

$$D_1(f, f + \varepsilon\Delta) = g_{1,f}(\varepsilon\Delta) + O(\varepsilon^3).$$

Proof. First note that

$$\begin{aligned} \text{tr}(f(f + \varepsilon\Delta)^{-1}) &= \text{tr}\left(f^{1/2}(I + f^{-1/2}\varepsilon\Delta f^{-1/2})^{-1}f^{-1/2}\right) \\ &= \text{tr}\left(I + f^{-1/2}\varepsilon\Delta f^{-1/2}\right)^{-1} \\ \text{tr}(f(f + \varepsilon\Delta)^{-1}) &= m - \text{tr}(f^{-1/2}\varepsilon\Delta f^{-1/2}) \\ &\quad + \|f^{-1/2}\varepsilon\Delta f^{-1/2}\|_{\text{Fr}}^2 + O(\varepsilon^3). \end{aligned}$$

Likewise,

$$\begin{aligned}\mathrm{tr}(f + \varepsilon\Delta)f^{-1} &= m + \mathrm{tr}(\varepsilon\Delta f^{-1}) \\ &= m + \mathrm{tr}(f^{-1/2}\varepsilon\Delta f^{-1/2}).\end{aligned}$$

Therefore,

$$\begin{aligned}\mathsf{D}_1(f, f + \varepsilon\Delta) &= \mathrm{tr} \int_{-\pi}^{\pi} (f(f + \varepsilon\Delta)^{-1} + (f + \varepsilon\Delta)f^{-1} - 2I) \frac{d\theta}{2\pi} \\ &= \int_{-\pi}^{\pi} \|f^{-1/2}\varepsilon\Delta f^{-1/2}\|_{\mathrm{Fr}}^2 \frac{d\theta}{2\pi} + O(\varepsilon^3).\end{aligned}$$

□

Obviously, an alternative expression for $\mathfrak{g}_{1,f}$ that requires neither spectral factorization nor the computation of the Hermitian square root of f , is the following:

$$\mathfrak{g}_{1,f}(\Delta) := \int_{-\pi}^{\pi} \mathrm{tr}(f^{-1}\Delta f^{-1}\Delta) \frac{d\theta}{2\pi}. \quad (3.18b)$$

It is interesting to also note that any of (3.14), (3.15), (3.16), and (3.17) leads to the same Riemannian metric.

Based on suboptimality of prediction

The paradigm in [17] for a Riemannian structure of scalar power spectral densities was originally built on the degradation of predictive error variance, as this is reflected in the strength of the inequalities of Proposition 8. In this section we explore the direct generalization of that route. Thus, we consider the quadratic form which \mathcal{F} inherits from the relevant divergence D_2 , defined in (3.11). The next proposition shows that this defines the corresponding metric:

$$\begin{aligned}\mathfrak{g}_{2,f}(\Delta) &:= \mathrm{tr} \int_{-\pi}^{\pi} (f_+^{-1}\Delta f_+^{-*})^2 \frac{d\theta}{2\pi} - \mathrm{tr} \left(\int_{-\pi}^{\pi} f_+^{-1}\Delta f_+^{-*} \frac{d\theta}{2\pi} \right)^2 \\ &= \mathfrak{g}_{1,f}(\Delta) - \mathrm{tr} \left(\int_{-\pi}^{\pi} f_+^{-1}\Delta f_+^{-*} \frac{d\theta}{2\pi} \right)^2.\end{aligned} \quad (3.19)$$

Proposition 12. *Let $(f, \Delta) \in \mathcal{F} \times \mathcal{D}$ and $\varepsilon > 0$. Then, for ε sufficiently small,*

$$\mathsf{D}_2(f, f + \varepsilon\Delta) = \frac{1}{2}\mathfrak{g}_{2,f}(\varepsilon\Delta) + O(\varepsilon^3).$$

Proof. In order to simplify the notation let

$$\Delta_\varepsilon := f_+^{-1} \varepsilon \Delta f_+^{-*}.$$

Since Δ, f are both bounded, $|\text{tr}(\Delta_\varepsilon^k)| = O(\varepsilon^k)$ as well as $|\text{tr}(\int_{-\pi}^{\pi} \Delta_\varepsilon \frac{d\theta}{2\pi})^k| = O(\varepsilon^k)$. Using a Taylor series expansion,

$$\begin{aligned} & \text{tr} \log \left(\int_{-\pi}^{\pi} f_+^{-1} (f + \varepsilon \Delta) f_+^{-*} \frac{d\theta}{2\pi} \right) \\ &= \text{tr} \log \left(I + \int_{-\pi}^{\pi} \Delta_\varepsilon \frac{d\theta}{2\pi} \right) \\ &= \text{tr} \left(\int_{-\pi}^{\pi} \Delta_\varepsilon \frac{d\theta}{2\pi} \right) - \frac{1}{2} \text{tr} \left(\int_{-\pi}^{\pi} \Delta_\varepsilon \frac{d\theta}{2\pi} \right)^2 + O(\varepsilon^3), \end{aligned}$$

while

$$\begin{aligned} & \text{tr} \left(\int_{-\pi}^{\pi} \log(f_+^{-1} (f + \varepsilon \Delta) f_+^{-*}) \frac{d\theta}{2\pi} \right) \\ &= \int_{-\pi}^{\pi} \text{tr} \log(I + \Delta_\varepsilon) \frac{d\theta}{2\pi} \\ &= \int_{-\pi}^{\pi} \text{tr} \left(\Delta_\varepsilon - \frac{1}{2} \Delta_\varepsilon^2 \right) \frac{d\theta}{2\pi} + O(\varepsilon^3). \end{aligned}$$

Thus

$$D_2(f, f + \varepsilon \Delta) = \frac{1}{2} \text{tr} \left(\int_{-\pi}^{\pi} \Delta_\varepsilon^2 d\theta - \left(\int_{-\pi}^{\pi} \Delta_\varepsilon \frac{d\theta}{2\pi} \right)^2 \right) + O(\varepsilon^3).$$

□

Interestingly, using of the divergences in (3.12) yields the same metric $\mathfrak{g}_{2,f}$. We also remark that although $\mathfrak{g}_{1,f}$ and $\mathfrak{g}_{2,f}$ are closely related, there is a substantial difference between them. In contrast to $\mathfrak{g}_{2,f}$, evaluation of $\mathfrak{g}_{1,f}$ does not require computing f_+ . However, on the other hand, both $\mathfrak{g}_{1,f}$, and $\mathfrak{g}_{2,f}$ are similarly unaffected by consistent scaling of f and Δ .

Remark 13. *A dual theory based on post-diction error (see Remark 6 and [17]) leads to swapping the left outer factor with the star of the right one. In it, while the analog of $\mathfrak{g}_{1,f}$ is the same as before, the analogous expression for $\mathfrak{g}_{2,f}$ is different.*

3.3.2 Geometry on positive matrices

As indicated earlier, a Riemannian metric on the space of Hermitian $m \times m$ matrices is a family of quadratic forms originating from inner products that depend smoothly on the Hermitian base

point M —the standard Hilbert-Schmidt metric $g_{\text{HS},M}(\Delta) = \langle \Delta, \Delta \rangle := \text{tr}(\Delta^2)$ being one such. Of particular interest are metrics on the space of positive definite matrices that ensure the space is complete and geodesically complete². For our purposes, matrices typically represent covariances. To this end a standard recipe for constructing a Riemannian metric is to begin with an information potential, such as the Boltzmann entropy of a Gaussian distribution with zero mean and covariance M ,

$$S(M) := -\frac{1}{2} \log(\det(M)) + \text{constant},$$

and define an inner product via its Hessian

$$\begin{aligned} \langle X, Y \rangle_M &:= \frac{\partial^2}{\partial x \partial y} S(M + xX + yY)|_{x=0, y=0} \\ &= \text{tr}(M^{-1}XM^{-1}Y). \end{aligned}$$

The Riemannian metric so defined,

$$\begin{aligned} g_M(\Delta) &:= \text{tr}(M^{-1}\Delta M^{-1}\Delta) \\ &= \|M^{-\frac{1}{2}}\Delta M^{-\frac{1}{2}}\|_{\text{Fr}}^2, \end{aligned}$$

is none other than the Fisher-Rao metric on Gaussian distributions expressed in the space of the corresponding covariance matrices.

The relationship of the Fisher-Rao metric on Gaussian distributions with the metric $g_{1,f}$ in (3.18b) is rather evident. Indeed, g_M coincides with $g_{1,f}$ for power spectra which are constant across frequencies, i.e., taking $f = M$ to be a constant Hermitian positive definite matrix.

It is noted that $g_M(\Delta)$ remains invariant under congruence, that is,

$$g_M(\Delta) = g_{TMT^*}(T\Delta T^*)$$

for any square invertible matrix-function T . This is a natural property to demand since it implies that the distance between covariance matrices does not change under coordinate transformations. The same is inherited by $g_{1,f}$ for power spectra. It is for this reason that g_M has in fact been extensively studied in the context of general C^* -algebras and their positive elements; we refer to [35, pg. 201-235] for a nice exposition of relevant material and for further references. Below we highlight certain key facts that are relevant to this chapter. But first, and for future reference, we recall a standard result in differential geometry.

² A space is complete when Cauchy sequences converge to points in the space. It is geodesically complete when the definition domain of geodesics extends to the complete real line \mathbb{R} ; i.e., extrapolating the path beyond the end points remains always in the space.

Proposition 14. Let \mathcal{M} be a Riemannian manifold with $\|\Delta\|_M^2$ denoting the Riemannian metric at $M \in \mathcal{M}$ and Δ a tangent direction at M . For each pair of points $M_0, M_1 \in \mathcal{M}$ consider the path space

$$\Theta_{M_0, M_1} := \{M_\tau : [0, 1] \rightarrow \mathcal{M} : M_\tau \text{ is a piecewise smooth path connecting the two given points}\}.$$

Denote by $\dot{M}_\tau := dM_\tau/d\tau$. The arc-length

$$\int_0^1 \|\dot{M}_\tau\|_M d\tau,$$

as well as the “action/energy” functional

$$\int_0^1 \|\dot{M}_\tau\|_M^2 d\tau$$

attain a minimum at a common path in Θ_{f_0, f_1} . Further, the minimal value of the arclength is the square root of the minimal value of the energy functional, and on a minimizing path the “speed” $\|\dot{M}_\tau\|_M$ remains constant for $\tau \in [0, 1]$.

Proof. See [57, pg. 137]. □

The insight behind the statement of the proposition is as follows. The arclength is evidently unaffected by a re-parametrization of a geodesic connecting the two points. The “energy” functional on the other hand, is minimized for a specific parametrization of geodesic where the velocity stays constant. Thus, the two are intimately related. The proposition will be applied first to paths between matrices, but in the next section it will also be invoked for geodesics between power spectra.

Herein we are interested in geodesic paths M_τ , $\tau \in [0, 1]$, connecting positive definite matrices M_0 to M_1 and in computing the corresponding geodesic distances

$$d_g(M_0, M_1) = \int_0^1 \left\| M_\tau^{-1/2} \frac{dM_\tau}{d\tau} M_\tau^{-1/2} \right\|_{\text{Fr}} d\tau.$$

Recall that a geodesic M_τ is the shortest path on the manifold connecting the beginning to the end.

Theorem 15. *Given Hermitian positive matrices M_0, M_1 , the geodesic between them with respect to g_M is unique (modulo re-parametrization) and given by*

$$M_\tau = M_0^{1/2} (M_0^{-1/2} M_1 M_0^{-1/2})^\tau M_0^{1/2}, \quad (3.20)$$

for $0 \leq \tau \leq 1$. Further, it holds that

$$d_g(M_0, M_\tau) = \tau d_g(M_0, M_1), \text{ for } \tau \in [0, 1],$$

and the geodesic distance is

$$d_g(M_0, M_1) = \|\log(M_0^{-1/2} M_1 M_0^{-1/2})\|_{\text{Fr}}.$$

Proof. A proof is given in [35, Theorem 6.1.6, pg. 205]. However, since this is an important result for our purposes and for completeness, we provide an independent short proof relying on Pontryagin's minimum principle.

We first note that, since g_M is congruence invariant, the path $TM_\tau T^*$ is a geodesic between $TM_0 T^*$ and $TM_1 T^*$, for any invertible matrix T . Further, the geodesic length is independent of T . Thus, we set

$$T = M_0^{-\frac{1}{2}},$$

and seek a geodesic path between

$$X_0 = I \text{ and } X_1 = M_0^{-\frac{1}{2}} M_1 M_0^{-\frac{1}{2}}. \quad (3.21)$$

Appealing to Proposition 14 we seek

$$\min\left\{\int_0^1 \text{tr}(X_\tau^{-1} U_\tau X_\tau^{-1} U_\tau) d\tau, \text{ subject to } \dot{X}_\tau = U_\tau, \text{ and } X_0, X_1 \text{ specified}\right\}. \quad (3.22)$$

Now, (3.22) is a standard optimal control problem. The value of the optimal control must annihilate the variation of the Hamiltonian with respect to the ‘‘control’’ U_τ

$$\text{tr}(X_\tau^{-1} U_\tau X_\tau^{-1} U_\tau) + \text{tr}(\Lambda_\tau U_\tau).$$

Here, Λ_τ represents the co-state (i.e., Lagrange multiplier functions). The variation is

$$\text{tr}(2X_\tau^{-1} U_\tau X_\tau^{-1} \delta U + \Lambda_\tau \delta U)$$

and this being identically zero for all δ_U implies that

$$U_\tau = -\frac{1}{2}X_\tau\Lambda_\tau X_\tau. \quad (3.23)$$

Similarly, the co-state equation is obtained by considering the variation with respect to X . This gives

$$\dot{\Lambda}_\tau = 2X_\tau^{-1}U_\tau X_\tau^{-1}U_\tau X_\tau^{-1}.$$

Substitute the expression for U_τ into the state and the co-state equations to obtain

$$\begin{aligned} \dot{X}_\tau &= -\frac{1}{2}X_\tau\Lambda_\tau X_\tau \\ \dot{\Lambda}_\tau &= \frac{1}{2}\Lambda_\tau X_\tau \Lambda_\tau. \end{aligned}$$

Note that

$$\dot{X}_\tau\Lambda_\tau + X_\tau\dot{\Lambda}_\tau = 0,$$

identically, for all τ . Hence, the product $X_\tau\Lambda_\tau$ is constant. Set

$$X_\tau\Lambda_\tau = -2C. \quad (3.24)$$

The state equation becomes

$$\dot{X}_\tau = CX_\tau.$$

The solution with initial condition $X_0 = I$ is

$$X_\tau = \exp(C\tau).$$

Matching (3.21) requires that $\exp(C) = X_1 = M_0^{-\frac{1}{2}}M_1M_0^{-\frac{1}{2}}$. Thus, $X_\tau = (M_0^{-\frac{1}{2}}M_1M_0^{-\frac{1}{2}})^\tau$ and the geodesic is as claimed. Further,

$$C = \log(M_0^{-\frac{1}{2}}M_1M_0^{-\frac{1}{2}})$$

while $U_\tau = CX_\tau$ from (3.24) and (3.23). So finally, for the minimizing choice of U_τ we get that the cost

$$\begin{aligned} \int_0^\tau \text{tr}(X_\tau^{-1}U_\tau X_\tau^{-1}U_\tau) d\tau &= \int_0^\tau \text{tr}(C^2) d\tau \\ &= \tau \|\log(M_0^{-1/2}M_1M_0^{-1/2})\|_{\text{Fr}}^2 \end{aligned}$$

as claimed. □

Remark 16. *It's important to point out the lower bound*

$$d_g(M_0, M_1) \geq \|\log M_0 - \log M_1\|_{\text{Fr}} \quad (3.25)$$

on the geodesic distance which holds with equality when M_0 and M_1 commute. This is known as the exponential metric increasing property [35, page 203] and will be used later on.

The mid point of the geodesic path in (3.20) is known as the geometric mean of the two matrices M_0 and M_1 . This is commonly denoted by

$$M_{\frac{1}{2}} := M_0 \# M_1.$$

Similar notation, with the addition of a subscript τ , will be used to designate the complete geodesic path

$$M_\tau = M_0 \#_\tau M_1 := M_0^{1/2} (M_0^{-1/2} M_1 M_0^{-1/2})^\tau M_0^{1/2}$$

(see [35]). A number of useful properties can be easily verified:

i) Congruence invariance: for any invertible matrix T ,

$$d_g(M_0, M_1) = d_g(TM_0T^*, TM_1T^*).$$

ii) Inverse invariance:

$$d_g(M_0, M_1) = d_g(M_0^{-1}, M_1^{-1}).$$

iii) The metric satisfies the semiparallelogram law.

iv) The space of positive definite matrices metrized by d_g is complete; that is, any Cauchy sequence of positive definite matrices converges to a positive definite matrix.

v) Given any three “points” M_0, M_1, M_2 ,

$$d_g(M_0 \#_\tau M_1, M_0 \#_\tau M_2) \leq \tau d_g(M_1, M_2),$$

which implies that geodesics diverge at least as fast as “Euclidean geodesics”.

Remark 17. *Property v) implies that the Riemannian manifold of positive definite matrices with metric d_g has nonpositive sectional curvature [58, pg. 39–40]. The nonpositive sectional curvature of a simply connected complete Riemannian manifold has several important geometric consequences. It implies the existence and uniqueness of a geodesic connecting any two points on the manifold [58, pg. 3–4]. Convex sets on such a manifold are defined by the requirement*

that geodesics between any two points in the set lie entirely in the set [58, pg. 67]. Then, “projections” onto the set exist in that there is always a closest point within convex set to any given point. Evidently, such a property should be valuable in applications, such as speaker identification or speech recognition based on a database of speech segments; e.g., models may be taken as the “convex hull” of prior sample spectra and the metric distance of a new sample compared to how far it resides from a given such convex set. Another property of such a manifold is that the center of mass of a set of points is contained in the closure of its convex hull [58, pg. 68]; this property has been used to define the geometric means of symmetric positive matrices in [59].

3.3.3 Geodesics and geodesic distances

Power spectral densities are families of Hermitian matrices parametrized by the frequency θ , and as such, can be thought of as positive operators on a Hilbert space. Geometries for positive operators have been extensively studied for some time now, and power spectral densities may in principle be studied with similar tools. However, what it may be somewhat surprising is that the geometries obtained earlier, based on the innovations flatness and optimal prediction, have points of contact with this literature. This was seen in the correspondence between the metrics that we derived.

In the earlier sections we introduced two metrics, g_1 and g_2 . Although there is a close connection between the two, as suggested by (3.19), it is only for the former that we are able to identify geodesics and compute the geodesic lengths, based on the material in Section 3.3.2. We do this next.

Theorem 18. *There exists a unique geodesic path f_τ with respect to $g_{1,f}$, connecting any two spectra $f_0, f_1 \in \mathcal{F}$. The geodesic path is*

$$f_\tau = f_0^{1/2} (f_0^{-1/2} f_1 f_0^{-1/2})^\tau f_0^{1/2}, \quad (3.26)$$

for $0 \leq \tau \leq 1$. The geodesic distance is

$$d_{g_1}(f_0, f_1) = \sqrt{\int_{-\pi}^{\pi} \|\log f_0^{-1/2} f_1 f_0^{-1/2}\|_{\text{Fr}}^2 \frac{d\theta}{2\pi}}. \quad (3.27)$$

Proof. As before, in view of Proposition 14, instead of the geodesic length we may equivalently consider minimizing the energy/action functional

$$\begin{aligned} E &= \int_0^1 \int_{-\pi}^{\pi} \|f_{\tau}^{-1/2} \dot{f}_{\tau} f_{\tau}^{-1/2}\|_{\text{Fr}}^2 \frac{d\theta}{2\pi} d\tau \\ &= \int_{-\pi}^{\pi} \int_0^1 \|f_{\tau}^{-1/2} \dot{f}_{\tau} f_{\tau}^{-1/2}\|_{\text{Fr}}^2 d\tau \frac{d\theta}{2\pi}. \end{aligned}$$

Clearly, this can be minimized point-wise in θ invoking Theorem 15. Now, inversion as well as the fractional power of symmetric (strictly) positive matrices represent continuous and differentiable maps. Hence, it can be easily seen that, because f_0, f_1 are in \mathcal{F} so is

$$f_{\tau} = f_0^{1/2} (f_0^{-1/2} f_1 f_0^{-1/2})^{\tau} f_0^{1/2}.$$

Therefore, this path is the sought minimizer of

$$\int_0^1 \|f_{\tau}^{-1/2} \dot{f}_{\tau} f_{\tau}^{-1/2}\|_{\text{Fr}}^2 d\tau$$

and the geodesic length is as claimed. \square

Remark 19. *In the case of scalar power spectra, the (unique) geodesic (3.26) for $\mathfrak{g}_{1,f}$ is*

$$f_{\tau} = f_0 \left(\frac{f_1}{f_0} \right)^{\tau}, \text{ with } \tau \in [0, 1].$$

This coincides with a particular choice of a geodesic ([17, Eq. (19)]) for the prediction-based metric. Prediction-based geodesics are nonunique (see [17]).

Corollary 20. *Given any $f_0, f_1, f_2 \in \mathcal{F}$, the function $d_{\mathfrak{g}_1}(f_0 \sharp_{\tau} f_1, f_0 \sharp_{\tau} f_2)$ is convex on τ .*

Proof. The proof is a direct consequence of the convexity of the metric $d_{\mathfrak{g}}(\cdot, \cdot)$. \square

The importance of the statement in the corollary is that the metric space has nonpositive curvature. Other properties are similarly inherited. For instance, $d_{\mathfrak{g}_1}$ satisfies the semi-parallellogram law.

Next we explain that the closure of the space of positive differentiable power spectra, under \mathfrak{g}_1 , is simply power spectra that are squarely log integrable. This is not much of a surprise in view of the metric and the form of the geodesic distance. Thus, the next proposition shows that the completion, denoted by “bar,” is in fact

$$\tilde{\mathcal{F}} := \{f \mid m \times m \text{ positive definite a.e. on } [-\pi, \pi], \log f \in L_2[-\pi, \pi]\}. \quad (3.28)$$

It should be noted that the metric d_{g_1} is not equivalent to an L_2 -based metric $\|\log(f_1) - \log(f_2)\|_2$ for the space. Here,

$$\|h\|_2 := \sqrt{\int_{-\pi}^{\pi} \|h\|_{\text{Fr}}^2 \frac{d\theta}{2\pi}}.$$

In fact, using the latter $\tilde{\mathcal{F}}$ has zero curvature while, using d_{g_1} , $\tilde{\mathcal{F}}$ becomes a space with non-positive (non-trivial) curvature.

Proposition 21. *The completion of \mathcal{F} under d_{g_1} is as indicated in (3.28).*

Proof. Clearly, for $f \in \mathcal{F}$, $\log f \in L_2[-\pi, \pi]$ since f is continuous on the closed interval and positive definite. Further, the logarithm maps positive differentiable matrix-functions to Hermitian differentiable ones, bijectively. Our proof of $\tilde{\mathcal{F}}$ being the completion of \mathcal{F} is carried out in three steps. First we will show that the limit of every Cauchy sequence in \mathcal{F} belongs to $\tilde{\mathcal{F}}$. Next we argue that every point in $\tilde{\mathcal{F}}$ is the limit of a sequence in \mathcal{F} , which together with the first step shows that \mathcal{F} is dense in $\tilde{\mathcal{F}}$. Finally, we need to show that $\tilde{\mathcal{F}}$ is complete with d_{g_1} .

First, consider a Cauchy sequence $\{f_n\}$ in \mathcal{F} which converges to f . Hence, there exists an N , such that for any $k \geq N$, $d_{g_1}(f_k, f) < 1$. Using the triangular inequality for d_{g_1} , we have that

$$d_{g_1}(I, f) \leq d_{g_1}(I, f_N) + d_{g_1}(f_N, f),$$

or, equivalently,

$$\|\log f\|_2 < \|\log f_N\|_2 + 1.$$

Since $\|\log f_N\|_2$ is finite, $f \in \tilde{\mathcal{F}}$.

Next, for any point f in $\tilde{\mathcal{F}}$ which is not continuous, we show that it is the limit of a sequence in \mathcal{F} . Let $h = \log f$, then $h \in L_2[-\pi, \pi]$. Since the set of differentiable functions $C^1[-\pi, \pi]$ is dense in $L_2[-\pi, \pi]$, there exists a sequence $\{h_n \in C^1[-\pi, \pi]\}$ which converges to h in the L_2 norm. Using Theorem 3 in [60, pg. 86], there exists a subsequence $\{h_{n_k}\}$ which converges to h almost everywhere in $[-\pi, \pi]$, i.e.,

$$\|h_{n_k}(\theta) - h(\theta)\|_{\text{Fr}} \rightarrow 0 \text{ a.e., as } n_k \rightarrow \infty.$$

Since the exponential map is continuous [61, pg. 430], $\|e^{h_{n_k}(\theta)} - e^h(\theta)\|_{\text{Fr}}$ converges to 0 almost everywhere as well. Using the sub-multiplicative property of the Frobenius norm, we have that

$$\|I - e^{-h(\theta)} e^{h_{n_k}(\theta)}\|_{\text{Fr}} \leq \|e^{-h(\theta)}\|_{\text{Fr}} \|e^{h_{n_k}(\theta)} - e^h(\theta)\|_{\text{Fr}},$$

where the right side of the above inequality goes to zero. Thus the spectral radius of $(I - e^{-h(\theta)} e^{h_{n_k}(\theta)})$ goes to zero [62, pg. 297]. Hence, all the eigenvalues $\lambda_i(e^{-h(\theta)} e^{h_{n_k}(\theta)})$, $1 \leq i \leq m$, converge to 1 as $k \rightarrow \infty$. Then, $f_{n_k} = e^{h_{n_k}} \in \mathcal{F}$ and

$$\begin{aligned} d_{g_1}(f_{n_k}, f) &= \sqrt{\int_{-\pi}^{\pi} \|\log f^{-1/2} f_{n_k} f^{-1/2}\|_{\text{Fr}}^2 \frac{d\theta}{2\pi}} \\ &= \sqrt{\int_{-\pi}^{\pi} \sum_{i=1}^m \log^2 \lambda_i(f^{-1} f_{n_k}) \frac{d\theta}{2\pi}} \\ &= \sqrt{\int_{-\pi}^{\pi} \sum_{i=1}^m \log^2 \lambda_i(e^{-h} e^{h_{n_k}}) \frac{d\theta}{2\pi}}. \end{aligned}$$

Since $\log \lambda_i(e^{-h} e^{h_{n_k}}) \rightarrow 0$ a.e., for $1 \leq i \leq m$, $d_{g_1}(f_{n_k}, f) \rightarrow 0$ as well. Therefore, f is the limit of $\{f_{n_k}\}$.

Finally we show that $\tilde{\mathcal{F}}$ is complete under d_{g_1} . Let $\{f_n\}$ be a Cauchy sequence in $(\tilde{\mathcal{F}}, d_{g_1})$, and let $h_n = \log f_n$. Using the inequality (3.25), we have

$$d_{g_1}(f_k, f_l) \geq \sqrt{\int_{-\pi}^{\pi} \|h_k - h_l\|_{\text{Fr}}^2 \frac{d\theta}{2\pi}}.$$

Thus $\{h_n\}$ is also a Cauchy sequence in $L_2[-\pi, \pi]$, which is a complete metric space. As a result, $\{h_n\}$ converges to a point h in $L_2[-\pi, \pi]$. Following the similar procedure as in the previous step, there exists a subsequence $\{f_{n_k}\}$ which converges to $f = e^h \in \tilde{\mathcal{F}}$. This completes our proof. \square

Remark 22. *Geodesics of $g_{2,f}$ for scalar power spectra were constructed in [17]. At the present time, a multivariable generalization appears to be a daunting task. The main obstacle is of course non-commutativity of matricial density functions and the absence of an integral representation of analytic spectral factors in terms of matrix-valued power spectral densities. In this direction we point out that some of the needed tools are in place. For instance, a square matrix-valued function which is analytic and non-singular in the unit disc \mathbb{D} , admits a logarithm*

which is also analytic in \mathbb{D} . To see this, consider such a matrix-function, say $f_+(z)$. The matrix logarithm is well defined locally in a neighborhood of any $z_0 \in \mathbb{D}$ via the Cauchy integral

$$g(z) = \frac{1}{2\pi i} \int_{L_{z_0}} \ln(\zeta) (\zeta I - f_+(z))^{-1} d\zeta.$$

Here, L_{z_0} is a closed path in the complex plane that encompasses all of the eigenvalues of $f_+(z_0)$ and does not separate the origin from the point at ∞ . The Cauchy integral gives a matrix-function $g(z)$ which is analytic in a sufficiently small neighborhood of z_0 in the unit disc \mathbb{D} —the size of the neighborhood being dictated by the requirement that the eigenvalues stay within L_{z_0} , and $\exp(g(z)) = f_+(z)$. To define the logarithm consistently over \mathbb{D} we need to ensure that we always take the same principle value. This is indeed the case if we extend $g(z)$ via analytic continuation: since $f_+(z)$ is not singular anywhere in \mathbb{D} and the unit disc is simply connected, the values for $g(z)$ will be consistent, i.e., any path from z_0 to an arbitrary $z \in \mathbb{D}$ will lead to the same value for $g(z)$. Thus, one can set $\log(f_+) = g$ and understand this to be a particular version of the logarithm. Similarly, powers of f_+ can also be defined using Cauchy integrals,

$$\frac{1}{2\pi i} \int_{L_{z_0}} \zeta^\tau (\zeta I - f_+(z))^{-1} d\zeta$$

for $\tau \in [0, 1]$, first in a neighborhood of a given $z_0 \in \mathbb{D}$, and then by analytic continuation to the whole of \mathbb{D} . As with the logarithm, there may be several versions. Geodesics for $\mathfrak{g}_{2,f}$ appear to require paths in the space of canonical spectral factors for the corresponding matricial densities, such as $f_{\tau+} = f_{0+} (f_{0+}^{-1} f_{1+})_+^\tau$. However, the correct expression remains elusive at present.

3.4 Examples

We first demonstrate geodesics connecting two power spectral densities that correspond to all-pole models, i.e., two autoregressive (AR) spectra. The geodesic path between them does not consist of AR-spectra, and it can be considered as a non-parametric model for the transition. The choice of AR-spectra for the end points is only for convenience. As discussed earlier, the aim of the theory is to serve as a tool in non-parametric estimation, path following, morphing, etc., in the spectral domain.

A scalar example:

Consider the two power spectral densities

$$f_i(\theta) = \frac{1}{|a_i(e^{j\theta})|^2}, \quad i \in \{0, 1\},$$

where

$$a_0 = (z^2 - 1.96 \cos(\frac{\pi}{5}) + 0.98^2)(z^2 - 1.7 \cos(\frac{\pi}{3}) + 0.85^2)(z^2 - 1.8 \cos(\frac{2\pi}{3}) + 0.9^2),$$

$$a_1 = (z^2 - 1.96 \cos(\frac{2\pi}{15}) + 0.98^2)(z^2 - 1.5 \cos(\frac{7\pi}{30}) + 0.75^2)(z^2 - 1.8 \cos(\frac{5\pi}{8}) + 0.9^2).$$

Their roots are marked by \times 's and \circ 's respectively, in Figure 3.2, and shown with respect to the unit circle in the complex plane. We consider and compare the following three ways of interpolating power spectra between f_0 and f_1 .

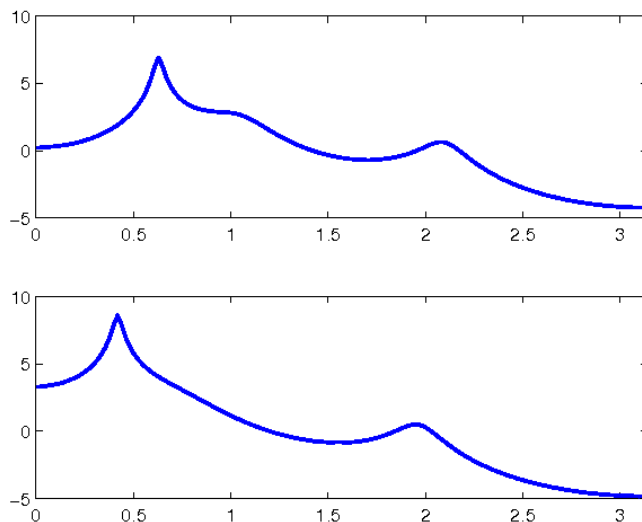


Figure 3.1: Plots of $\log f_0(\theta)$ (upper) and $\log f_1(\theta)$ (lower) for $\theta \in [0, \pi]$.

First, a parametric approach where the AR-coefficients are interpolated:

$$f_{\tau, \text{AR}}(\theta) = \frac{1}{|a_{\tau}(e^{j\theta})|^2}, \quad (3.29a)$$

with $a_{\tau}(z) = (1 - \tau)a_0(z) + \tau a_1(z)$. Clearly, there is a variety of alternative options (e.g., to interpolate partial reflection coefficients, etc.). However, our choice is intended to highlight the

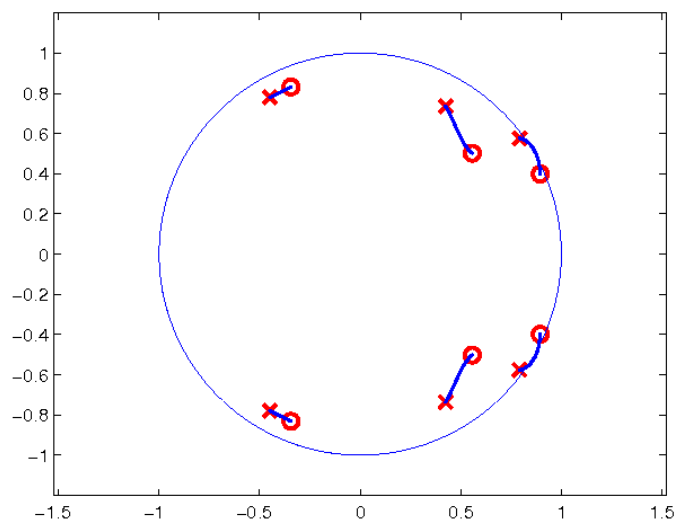


Figure 3.2: Locus of the roots of $a_\tau(z)$ for $\tau \in [0, 1]$.

fact that in a parameter space, admissible models may not always form a convex set. This is evidently the case here as the path includes factors that become “unstable.” The locus of the roots of $a_\tau(z) = 0$ for $\tau \in [0, 1]$ is shown in Figure 3.2.

Then we consider a linear segment connecting the two spectra:

$$f_{\tau,\text{linear}} = (1 - \tau)f_0 + \tau f_1. \quad (3.29b)$$

Again, this is to highlight the fact that the space of power spectra is not linear, and in this case, extrapolation beyond the convex linear combination of the two spectra leads to inadmissible function (as the path leads outside of the cone of positive functions). Finally, we provide the g_1 -geodesic between the two

$$f_{\tau,\text{geodesic}} = f_0 \left(\frac{f_1}{f_0} \right)^\tau. \quad (3.29c)$$

We compare $f_{\tau,\text{AR}}$, $f_{\tau,\text{linear}}$ and $f_{\tau,\text{geodesic}}$ for $\tau \in \{\frac{1}{3}, \frac{2}{3}, \frac{4}{3}\}$. We first note that in plotting $\log f_{\tau,\text{AR}}$ in Figure 3.3, that $f_{\frac{2}{3},\text{AR}}$ is not shown since it is not admissible. Likewise $\log f_{\tau,\text{linear}}$ in Figure 3.4 breaks up for $\tau = \frac{4}{3}$, since $f_{\frac{4}{3},\text{linear}}$ becomes negative for a range of frequencies –dashed curve indicates the absolute value of the logarithm when this takes complex values. The plot of $\log f_{\tau,\text{geodesic}}$ is defined for all the τ and shown in Figure 3.5. It is worth pointing out how two

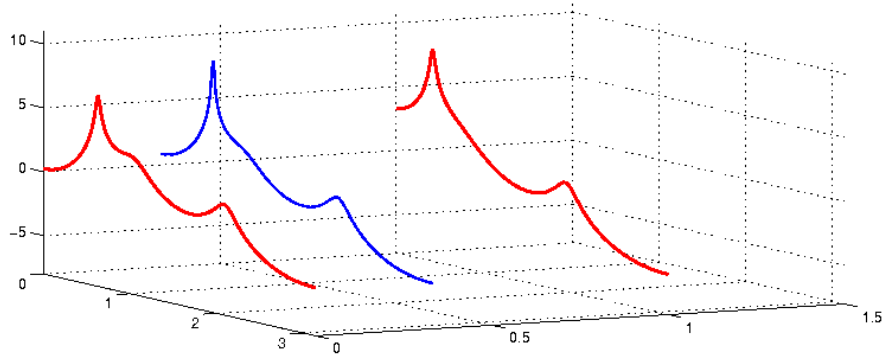


Figure 3.3: $\log f_{\tau,AR}(\theta)$ for $\tau = \frac{1}{3}, \frac{2}{3}, \frac{4}{3}$ (blue), $\tau = 0, 1$ (red).

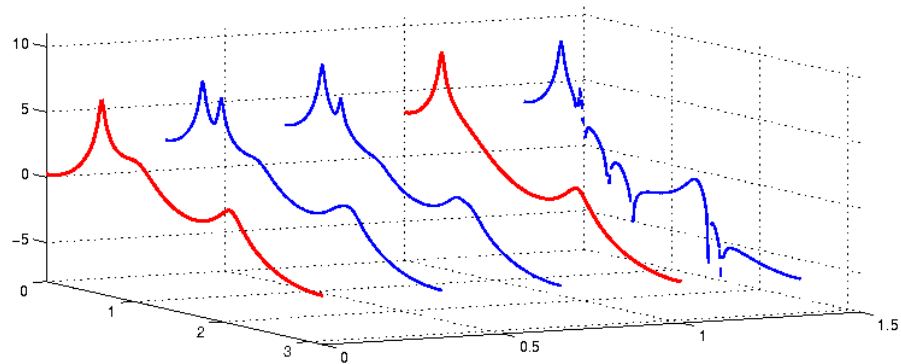


Figure 3.4: $\log f_{\tau,linear}(\theta)$ for $\tau = \frac{1}{3}, \frac{2}{3}, \frac{4}{3}$ (blue), $\tau = 0, 1$ (red).

apparent “modes” in $f_{\tau,linear}$ and $f_{\tau,geodesic}$ are swapping their dominance, which does not occur when following $f_{\tau,AR}$.

A multivariable example:

Consider the two matrix-valued power spectral densities

$$f_0 = \begin{bmatrix} 1 & 0 \\ 0.1e^{j\theta} & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{|a_0(e^{j\theta})|^2} & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0.1e^{-j\theta} \\ 0 & 1 \end{bmatrix}$$

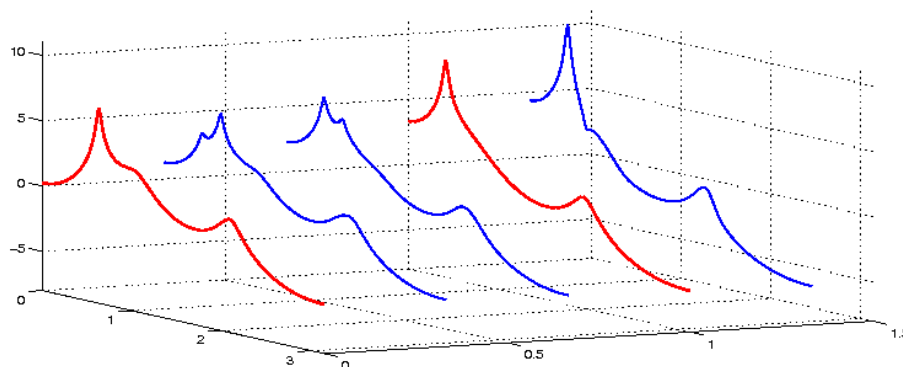


Figure 3.5: $\log f_{\tau, \text{geodesic}}(\theta)$ for $\tau = \frac{1}{3}, \frac{2}{3}, \frac{4}{3}$ (blue), $\tau = 0, 1$ (red).

$$f_1 = \begin{bmatrix} 1 & 0.1e^{j\theta} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{|a_1(e^{j\theta})|^2} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0.1e^{-j\theta} & 1 \end{bmatrix}.$$

Typically, these reflect the dynamic relationship between two time series; in turn these may represent noise input/output of dynamical systems or measurements across independent array of sensors, etc. The particular example reflects the typical effect of an energy source shifting its signature from one of two sensors to the other as, for instance, a possible scatterer moves with respect to the two sensors.

Below f_0 and f_1 are shown in Fig. 3.6 and Fig. 3.7, respectively. Since the value of a power spectral density f , at each point in frequency, is a Hermitian matrix, our convention is to show in the (1,1), (1,2) and (2,2) subplots the log-magnitude of the entries $f(1,1)$, $f(1,2)$ (which is the same as $f(2,1)$) and $f(2,2)$, respectively. Then, since only $f(1,2)$ is complex (and the complex conjugate of $f(2,1)$), we plot its phase in the (2,1) subplot.

Three dimensional surface shows the geodesic connecting f_0 to f_1 in Figure 3.8. Here, $f_{\tau, \text{geodesic}}$ is drawn using

$$f_{\tau, \text{geodesic}} = f_0^{\frac{1}{2}} (f_0^{-\frac{1}{2}} f_1 f_0^{-\frac{1}{2}})^{\tau} f_0^{\frac{1}{2}}.$$

It is interesting to observe the smooth shift of the energy across frequency and directionality.

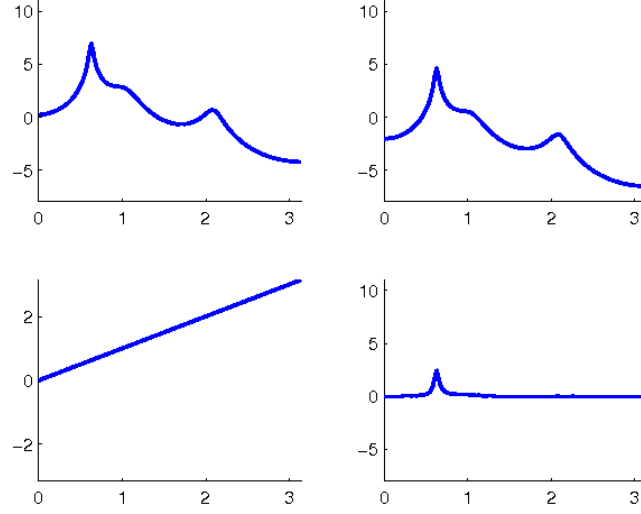


Figure 3.6: Subplots (1,1), (1,2) and (2,2) show $\log f_0(1, 1)$, $\log |f_0(1, 2)|$ (same as $\log |f_0(2, 1)|$) and $\log f_0(2, 2)$. Subplot (2,1) shows $\arg(f_0(2, 1))$.

3.5 Extensions

The aim of this study has been to develop multivariable divergence measures and metrics for matrix-valued power spectral densities. These are expected to be useful in quantifying uncertainty in the spectral domain, detecting events in non-stationary time series, smoothing and spectral estimation in the context of vector valued stochastic processes.

The prediction-based geodesic distance (3.27) has deep roots in statistical inference and information geometry. However, it doesn't metrize the weak convergence. A sequence of probability measures $\{\mu_n\}$ is said to converge weakly to μ , if $\int \phi d\mu_k$ converges to $\int \phi d\mu$ for every bounded continuous function ϕ . This is a desirable notion of convergence for power spectra since small changes in the power spectral density ought to reflect small changes in the statistics and vice versa. Thus it is natural to seek a transportation-like distance between two multivariate power spectral densities which ensures weak convergence.

Let us consider two multivariate power spectra f_0 and f_1 , and let ϕ be a bounded Lipschitz-continuous function, i.e. $\phi \in \Phi$ where

$$\Phi := \{\phi(\theta) : -MI \leq \phi(\theta) \leq MI, \phi(\theta) - \phi(\xi) \leq |\theta - \xi|I, \forall \theta, \xi \in [-\pi, \pi]\}$$

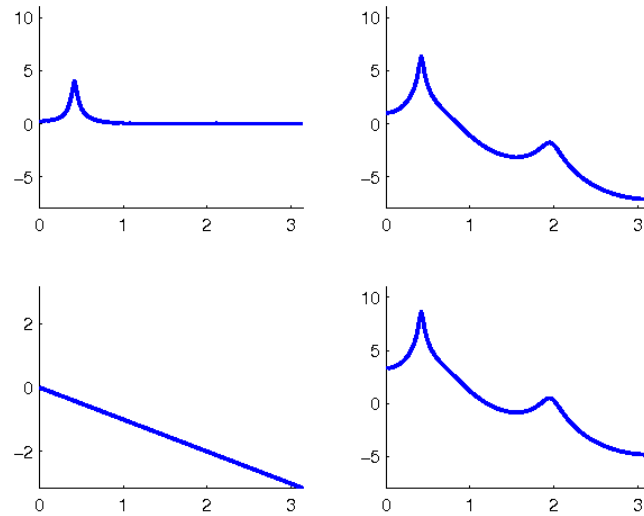


Figure 3.7: Subplots (1,1), (1,2) and (2,2) show $\log f_1(1, 1)$, $\log |f_1(1, 2)|$ (same as $\log |f_1(2, 1)|$) and $\log f_1(2, 2)$. Subplot (2,1) shows $\arg(f_1(2, 1))$.

and M is a constant large number. Then the following distance $d(f_0, f_1)$ metrizes weak convergence:

$$d(f_0, f_1) = \left\{ \sup_{\phi \in \Phi} \text{tr} \left(\int f_1 \phi d\theta - \int f_2 \phi d\theta \right) \right\}. \quad (3.30)$$

The dual of this problem turns out to be

$$\begin{aligned} \min_{\pi} \quad & \text{tr} \int |\theta - \xi| \pi(\theta, \xi) d\theta d\xi \\ \text{s.t.} \quad & \int \pi(\theta, \xi) d\xi - \int \pi(\xi, \theta) d\xi = f_1(\theta) - f_2(\theta) \\ & \pi(\theta, \xi) \geq 0, \forall \theta, \xi \end{aligned}$$

In the scalar case, this dual problem is also called the Kantorovich-Rubinstein transshipment

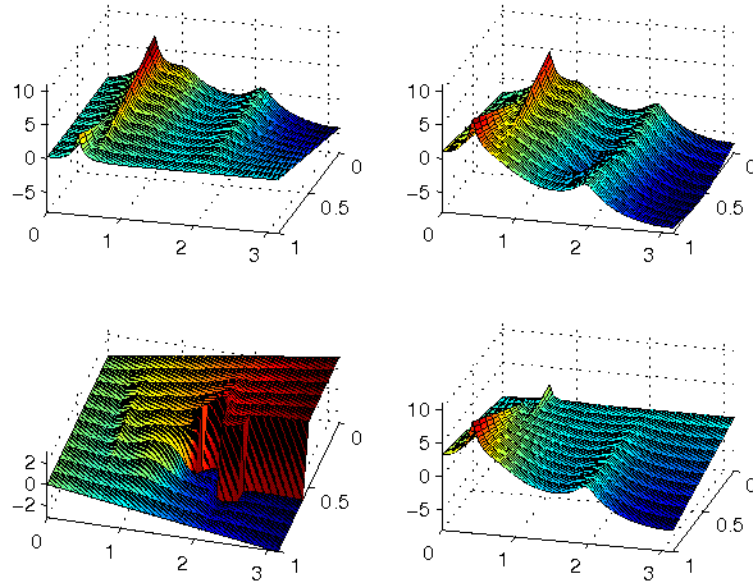


Figure 3.8: Subplots (1,1), (1,2) and (2,2) show $\log f_\tau(1,1)$, $\log |f_\tau(1,2)|$ (same as $\log |f_\tau(2,1)|$) and $\log f_\tau(2,2)$. Subplot (2,1) shows $\arg(f_\tau(2,1))$, for $\tau \in [0, 1]$.

problem [63, pg. 165], and it can be shown to be equivalent to the transportation problem

$$\begin{aligned}
 \min_{\pi} \quad & \int |\theta - \xi| \pi(\theta, \xi) d\theta d\xi \\
 \text{s.t.} \quad & \int \pi(\theta, \xi) d\xi = f_1(\theta) \\
 & \int \pi(\xi, \theta) d\xi = f_2(\theta) \\
 & \pi(\theta, \xi) \geq 0, \forall \theta, \xi.
 \end{aligned}$$

In the multivariate case, it is worth studying whether the distance (3.30) is equivalent to some form of transportation problems, and how it behaves when one distribution is a small rotation or translation of the other.

Chapter 4

Geodesics and Curve Fitting

Curve fitting is a process of constructing a curve that has the best fit to a set of given data. It is typically used to smooth the data, or reduce the dimensionality by fitting a parametric model. In Euclidean geometry, most familiar curves are lines, splines, and ellipses. As an analogy, we propose the geodesic path fitting on a Riemannian manifold. It can be used to model the non-stationary time series in the time-frequency domain [64, 65]. Suppose we are given a sequence of power spectral densities

$$\mathcal{G} := \{g_{\tau_i}(\theta) : \theta \in [-\pi, \pi] \text{ for } i = 0, 1, \dots, n\},$$

where τ_i is an increasing sequence of time-indices, normalized so that $\tau_0 = 0$ and $\tau_n = 1$. These power spectra may typically be obtained from time-series data using STFT, and τ_i ($i = 0, 1, \dots, n$) may represent the mid-points of the corresponding time-windows. We propose to use the concept of the distance and geodesic of spectra to regularize the tracking of spectral density at different time intervals. To be more specific, let d denote a geodesic distance. We seek a geodesic f_τ , $\tau \in [0, 1]$, which minimizes

$$J_{\mathcal{G}}(f_\tau) := \sum_{i=0}^n (d(f_\tau, g_{\tau_i}))^2.$$

When d is the Euclidean distance, we obtain the least-square line fitting.

We will focus on the transportation metric and its geodesic to illustrate the idea [65]. As usual, all spectra are normalized whenever the transportation metric or its geodesic is applied.

4.1 Geodesic Path Fitting

The global geodesic fitting aims to determine a W_2 -geodesic f_τ , $\tau \in [0, 1]$, which minimizes

$$J_{\mathcal{G}}(f_\tau) := \sum_{i=0}^n (d_{W_2}(f_{\tau_i}, g_{\tau_i}))^2.$$

Any geodesic f_τ is completely specified by two ‘‘points,’’ in our case f_0, f_1 . Alternatively, it is also specified by the transference plan ψ according to (2.9) and (2.11). The optimal choice of f_0, f_1, ψ needs to be determined from the data, i.e., the spectra \mathcal{G} and the times τ_i ($i = 0, \dots, n$). The aim of this section is to solve the corresponding optimization problem. This is done next.

In view of (2.10), we have that

$$(d_{W_2}(f_{\tau_i}, g_{\tau_i}))^2 = \int_0^1 (F_{\tau_i}^{-1}(t) - G_{\tau_i}^{-1}(t))^2 dt, \quad (4.1)$$

where $F_{\tau_i}^{-1}$ and $G_{\tau_i}^{-1}$ are the inverse functions of the cumulative distributions of f_{τ_i} and g_{τ_i} respectively. On the other hand, since $\psi(\theta)$ is the optimal transference plan from f_0 to f_1 as in (2.9), the geodesic f_τ has the explicit expression in (2.11). Thus

$$F_0(\theta) = F_1(\psi(\theta)) = F_{\tau_i}((1 - \tau_i)\theta + \tau_i\psi(\theta)).$$

Let $t = F_0(\theta)$ for $t \in [0, 1]$, then equivalently

$$\theta = F_0^{-1}(t), \quad \psi(\theta) = F_1^{-1}(t),$$

and

$$(1 - \tau_i)\theta + \tau_i\psi(\theta) = F_{\tau_i}^{-1}(t).$$

Therefore,

$$F_{\tau_i}^{-1}(t) = (1 - \tau_i)F_0^{-1}(t) + \tau_i F_1^{-1}(t), \quad (4.2)$$

and by plugging (4.2) into (4.1), the objective function can be re-written as

$$J_{\mathcal{G}}(f_0, f_1) = \sum_{i=0}^n \int_0^1 ((1 - \tau_i)F_0^{-1}(t) + \tau_i F_1^{-1}(t) - G_{\tau_i}^{-1}(t))^2 dt.$$

Since the g_{τ_i} 's are given, $G_{\tau_i}^{-1}(t)$ can be computed by taking the inverse function of the cumulative distribution G_{τ_i} , and the only unknowns in the above are $F_0^{-1}(t)$ and $F_1^{-1}(t)$ which are functions of f_0 and f_1 .

To determine f_0 and f_1 numerically, we divide the range $[0, 1]$ into N subintervals of equal length $1/N$ and denote by $\theta_k, k = 0, 1, \dots, N$, the values of θ for which

$$F_0(\theta_k) = \frac{k}{N}. \quad (4.3)$$

Similarly, let $\vartheta_{i,k} (i = 0, \dots, n, k = 0, 1, \dots, N)$ denote the values for which

$$G_{\tau_i}(\vartheta_{i,k}) = \frac{k}{N},$$

and $\hat{\theta}_k$ the values for which

$$F_1(\hat{\theta}_k) = \frac{k}{N}. \quad (4.4)$$

Thereby, $J_{\mathcal{G}}(f_0, f_1)$ is approximated by the following finite sum

$$\mathbb{J} = \frac{1}{N} \sum_{i=0}^n \sum_{k=1}^N ((1 - \tau_i)\theta_k + \tau_i\hat{\theta}_k - \vartheta_{i,k})^2. \quad (4.5)$$

The values of $\vartheta_{i,k} (i = 0, \dots, n, k = 0, 1, \dots, N)$ can be readily computed from the problem data \mathcal{G} , and the only unknowns in this “discretization” of $J_{\mathcal{G}}(f_0, f_1)$ are the vector of θ ’s, namely $\theta_k, k = 0, 1, \dots, N$ (which help determine F_0) and the vector of corresponding $\hat{\theta}$ ’s (which help determine F_1 , and then T). Therefore, the spectral tracking problem can be solved numerically via the following convex quadratic program with linear constraints:

$$\begin{aligned} \min \{ \mathbb{J} : \text{subject to } & -\pi \leq \theta_k \leq \theta_{k+1} \leq \pi \\ & \text{and } -\pi \leq \hat{\theta}_k \leq \hat{\theta}_{k+1} \leq \pi \\ & \text{for } 0 \leq k \leq N - 1 \}. \end{aligned} \quad (4.6)$$

The objective function is convex, so the optimal solution can be found efficiently [66]. The cumulative functions of optimal f_0 and f_1 can be recovered from θ_k and $\hat{\theta}_k$ through (4.3) and (4.4) respectively.

If instead we use the prediction-geodesic distance, the problem can still be formulated into a convex optimization problem. We aim to determine a prediction-geodesic $f_\tau, \tau \in [0, 1]$, which minimizes

$$J_{\mathcal{G}}(f_\tau) := \sum_{i=0}^n (d_{\text{pg}}(f_{\tau_i}, g_{\tau_i}))^2.$$

Recall that the geodesic path f_{τ_i} between f_0 and f_1 belongs to the exponential family

$$f_{\tau_i}(\theta) = f_0^{1-\tau_i}(\theta) f_1^{\tau_i}(\theta).$$

If we discretize θ into N intervals from $-\pi$ to π , and let $\hat{f}_{\tau_i} = \log f_{\tau_i}$, $\hat{g}_{\tau_i} = \log g_{\tau_i}$, for the simplicity of the notation, we can work on the problem numerically by minimizing the following expression

$$\begin{aligned}
\mathbb{J} &= \sum_{i=0}^n \sum_{k=1}^N (\hat{g}_{\tau_i}(k) - (1 - \tau_i)\hat{f}_0(k) - \tau_i\hat{f}_1(k))^2 \\
&\quad - \frac{1}{N} \sum_{i=0}^n \left(\sum_{k=1}^N (\hat{g}_{\tau_i}(k) - (1 - \tau_i)\hat{f}_0(k) - \tau_i\hat{f}_1(k)) \right)^2 \\
&= \sum_{i=0}^n (\hat{g}_{\tau_i} - (1 - \tau_i)\hat{f}_0 - \tau_i\hat{f}_1)^T \left(\mathbf{I} - \frac{1}{N} \mathbf{1}\mathbf{1}^T \right) \\
&\quad (\hat{g}_{\tau_i} - (1 - \tau_i)\hat{f}_0 - \tau_i\hat{f}_1), \tag{4.7}
\end{aligned}$$

where \mathbf{I} is an $N \times N$ identity matrix, $\mathbf{1}$ is an $N \times 1$ vector with each component being 1. Since $(\mathbf{I} - \frac{1}{N} \mathbf{1}\mathbf{1}^T)$ is positive semi-definite, formulation (4.7) is a convex function. Note that because predictive metric is scale-invariant, the solution is not unique. To make the computation stable, one can add constraints on the norms of \hat{f}_0 and \hat{f}_1 .

4.2 Simulation Results

As a numerical example, we generate time-series data by driving a time-varying system with unit-variance white noise and then superimposing white measurement noise with variance equal to 2. The time-varying system consists of a succession of (15th-order) auto-regressive filters chosen to match the spectral character of a W_2 -geodesic between an ideal power spectrum

$$f_{0,\text{ideal}}(\theta) = \left| \frac{1 - 0.5z^{-1} + 0.6z^{-2}}{1 + 0.8z^{-1} + 0.9z^{-2}} \right|_{z=e^{j\theta}}^2$$

and a final

$$f_{1,\text{ideal}}(\theta) = \left| \frac{1 + 0.5z^{-1} + 0.6z^{-2}}{1 - 0.8z^{-1} + 0.9z^{-2}} \right|_{z=e^{j\theta}}^2.$$

These are shown in Figure 4.1. The STFT with a window of 128 points and an overlapping between successive windows by 64 points provides us with a collection of power spectral \mathcal{G} as before. Figure 4.3 shows the time-series data (in the first row) and then, below, it compares STFT power spectra ($g_{\tau_i}(\theta)$) in the second row with corresponding spectra obtained via a geodesic fit ($f_{\tau_i}(\theta)$). Figure 4.2 compares the spectrogram obtained by STFT and the one by

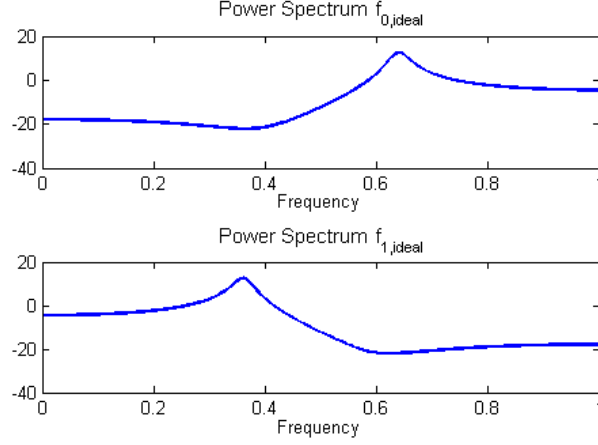


Figure 4.1: Power spectra $f_{0,\text{ideal}}$ and $f_{1,\text{ideal}}$ as functions of θ/π .

the geodesic path fitting. It is clear that the geodesic path captures quite accurately the drift of power in the spectrum over time. Furthermore, the corresponding “frozen time” spectra $f_{\tau_i}(\theta)$ for $i = 0, \dots, n$ appear to reproduce quite accurately the expected power distribution at the particular points in time. On the other hand, due to the small signal to noise ratio (SNR) the STFT seem quite unreliable.

Comparing with the transportation distance, Figure 4.4 and 4.5 show the estimated spectra by using the prediction geodesic. We can see that the prediction geodesic results in a fade-in and fade-out effect on the spectral tracking, thus may not be a proper choice in this situation.

To deal with a long record of data where signal gradually undergoes significant changes, we can simply divide all spectra into several batches and determine a geodesic for each batch. However, this method can lead to discrepancies at the connecting points of successive batches. In order to ensure that no abrupt changes happen at these connecting points, we can add the distances between the ends of geodesics and the beginnings of the successive geodesics to the overall objective function. In addition, to guarantee that the signal changes slowly in each batch, we can also include all the lengths of geodesics as a penalty term. Therefore, our objective function becomes

$$\begin{aligned}
 J_{\mathcal{G}}(f_{\tau}) := & \sum_{j=1}^m \sum_{i=0}^n (d_{W_2}(f_{j,\tau_i}, g_{j,\tau_i}))^2 + \alpha \sum_{j=1}^{m-1} (d_{W_2}(f_{j,\tau_n}, f_{j+1,\tau_1}))^2 \\
 & + \beta \sum_{j=1}^m (d_{W_2}(f_{j,\tau_1}, f_{j,\tau_n}))^2,
 \end{aligned}$$

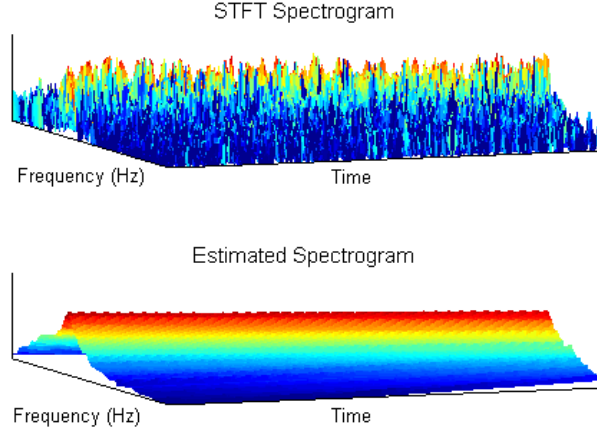


Figure 4.2: STFT spectrogram and estimated geodesic path.

where j is the index of batch, and m is the total number of batches. α and β are weights. In a similar way as before, $J_{\mathcal{G}}(f_{\tau})$ can be approximated as

$$\begin{aligned} \mathbb{J} = & \frac{1}{N} \sum_{j=1}^m \sum_{i=0}^n \sum_{k=1}^N (\tilde{\theta}_{j,i,k} - ((1 - \tau_i)\theta_{j,k} + \tau_i\hat{\theta}_{j,k}))^2 \\ & + \frac{\alpha}{N} \sum_{j=1}^{m-1} \sum_{k=1}^N (\hat{\theta}_{j,k} - \theta_{j+1,k})^2 + \frac{\beta}{N} \sum_{j=1}^m \sum_{k=1}^N (\hat{\theta}_{j,k} - \theta_{j,k})^2, \end{aligned}$$

where the definitions of $\tilde{\theta}_{j,i,k}$, $\theta_{j,k}$, and $\hat{\theta}_{j,k}$ are similar as before with j as the index of batch. The constraints resemble before by replacing θ_k and $\hat{\theta}_k$ with $\theta_{j,k}$ and $\hat{\theta}_{j,k}$. As an example, we generate two chirp signals with additive noise of standard variance 1.5. The spectrogram using STFT is shown in Figure 4.6. Since we allow the chirp to exceed the Nyquist frequency, there is apparent discontinuity in the path of spectral lines. Figure 4.7 shows the result with $\alpha = 50$, $\beta = 1$ and the batch size being 20. Compared with Figure 4.8 where no penalty terms are added, the improvement in the smoothness of the geodesic samples f_{j,τ_i} 's is rather evident.

If computational cost is not an issue, we can fit a geodesic for every pair of successive spectra so that the total geodesic length and the least square deviations from the given spectra are minimized. Therefore, our purpose is to determine f_{τ_i} , which minimizes

$$J_{\mathcal{G}}(f_{\tau}) := \sum_{i=0}^n (d_{W_2}(f_{\tau_i}, g_{\tau_i}))^2 + \alpha \sum_{i=0}^{n-1} (d_{W_2}(f_{\tau_i}, f_{\tau_{i+1}}))^2.$$

Note that in this case f_{τ_i} 's are not confined in a single geodesic. Let $F_{\tau_i}(\theta_{i,k}) = \frac{k}{N}$, where F_{τ_i} is the mass distribution function of f_{τ_i} as before, then the above objective function can be

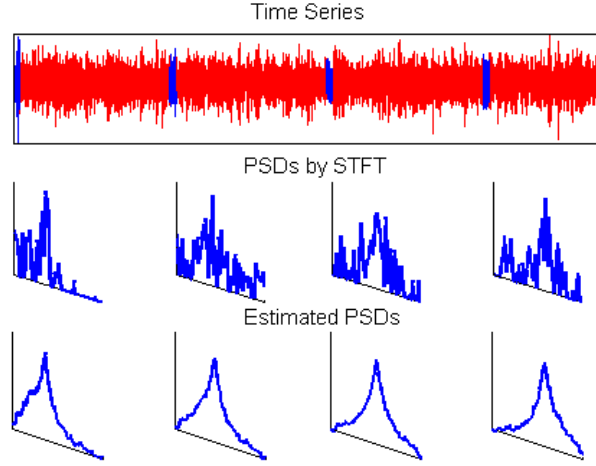


Figure 4.3: First row: time-series data; second row: STFT spectra based on the highlighted parts of the time-series; third row: samples of geodesic fit to the STFT spectra.

approximated numerically as

$$\mathbb{J} = \frac{1}{N} \sum_{i=0}^n \sum_{k=1}^N (\tilde{\theta}_{i,k} - \theta_{i,k})^2 + \frac{\alpha}{N} \sum_{i=0}^{n-1} \sum_{k=1}^N (\theta_{i,k} - \theta_{i+1,k})^2.$$

The unknown variables $\theta_{i,k}$ are subject to the constraints that they are non-decreasing with respect to k and their values are between 0 and 1. Since we have $n * N$ unknowns, it can be computationally expensive when the number of spectra n is large.

To illustrate the idea, we generate a signal with two chirps and additive noise of standard variance 1.5. There are 19 spectra by STFT, which is shown in Figure 4.9. From Figure 4.10, we can see that the estimated spectra by pair-wise geodesic fitting achieve much better accuracy than those by STFT.

4.3 Extensions

So far we have focused on the geodesic path fitting for scalar-valued spectral density functions. In the case of vectorial time series, we may need to consider geodesic fitting for multivariate power spectra. The prediction-based geodesic distance proposed in Chapter 3 leads to a seemly

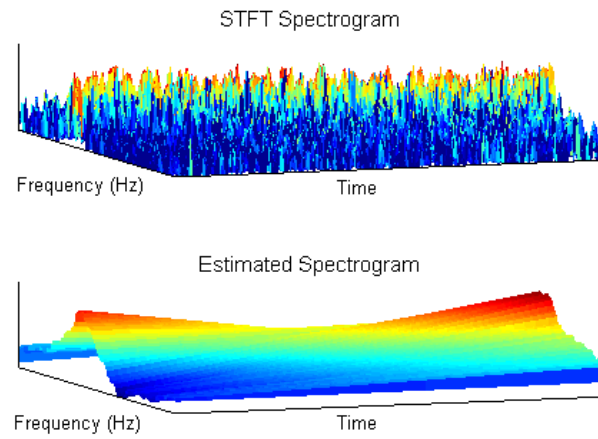


Figure 4.4: STFT spectrogram and the prediction geodesic path.

very difficult optimization problem. To get around, one may consider the log distance

$$d(f_0, f_1) = \left\{ \int \left\| \log f_0(\theta) - \log f_1 \theta \right\|_F^2 \frac{d\theta}{2\pi} \right\}^{\frac{1}{2}},$$

which is the Euclidean distance after the log transformation. The geodesic path f_τ is simply

$$\log f_\tau = (1 - \tau) \log f_0 + \tau \log f_1.$$

The resulting geodesic path fitting problem is convex in $\log f_0$ and $\log f_1$. Further research and study is needed to derive simple algorithms for the path fitting based on the prediction geodesic.

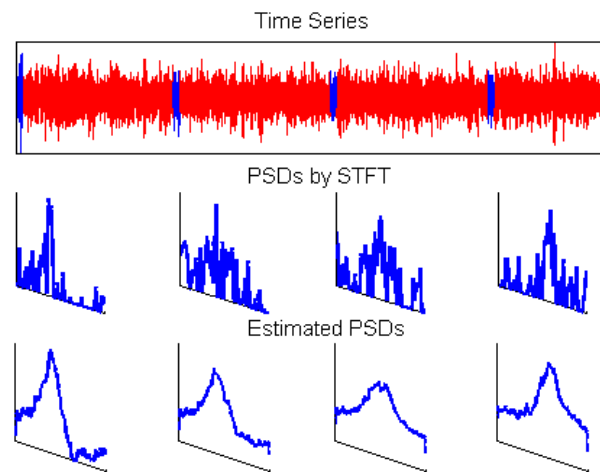


Figure 4.5: First row: time-series data; second row: STFT spectra based on the highlighted parts of the time-series; third row: samples of prediction geodesic fit to STFT spectra.

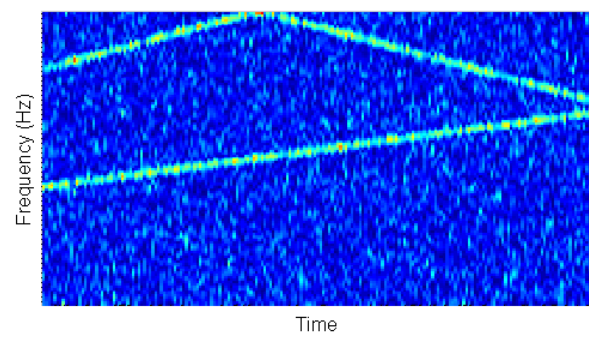


Figure 4.6: STFT spectrogram.

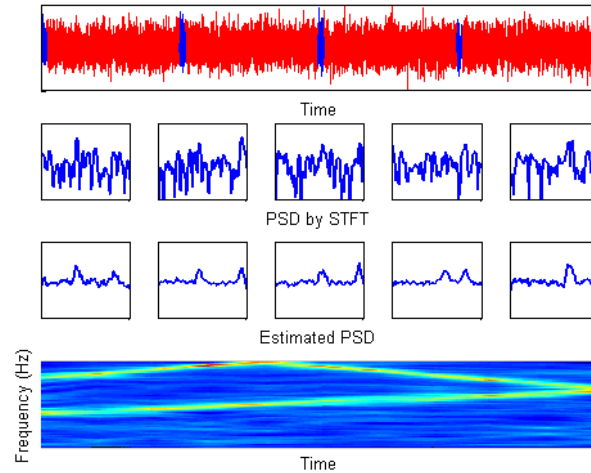


Figure 4.7: Top row: time-series with two chirp signals and additive noise; second row: STFT spectra corresponding to windows marked with blue; third row: estimated geodesic-fit samples with penalty terms; last row: the estimate geodesic path.

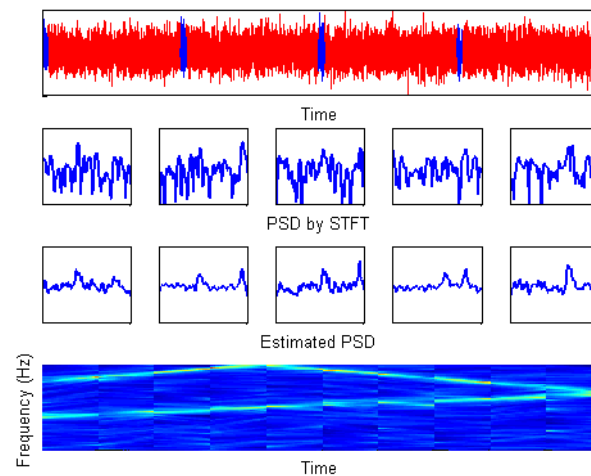


Figure 4.8: Top row: time-series with two chirp signals and additive noise; second row: STFT spectra corresponding to windows marked with blue; third row: estimated geodesic-fit samples without penalty terms; last row: the estimated geodesic path.

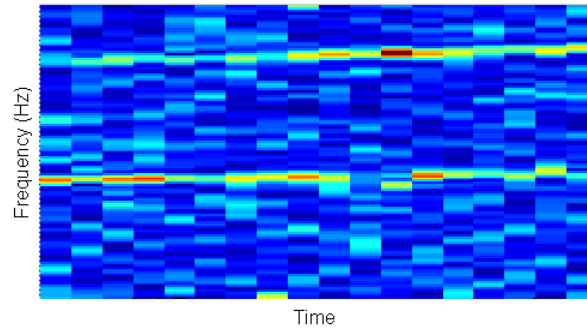


Figure 4.9: STFT spectrogram.

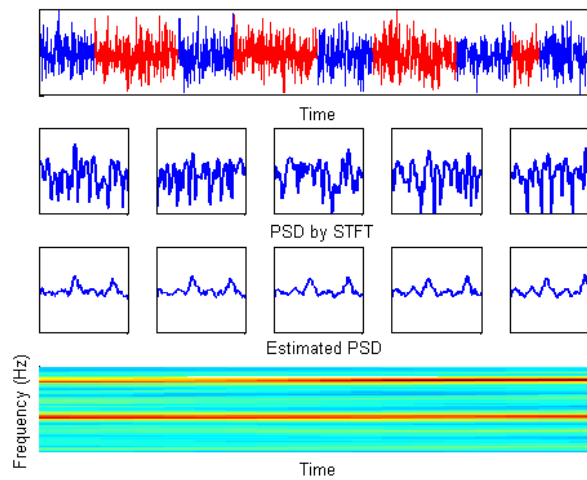


Figure 4.10: Top row: time-series with two chirp signals and additive noise; second row: STFT spectra corresponding to windows marked with blue; third row: estimated pairwise geodesic-fit samples; last row: the estimated spectrum.

Chapter 5

Metrics and Means

Constructing means, or averages, of a given data set has been of great importance in the areas of signal processing, computer vision and pattern recognition, optimal filtering and smoothing [67, 68, 69, 70, 71]. A typical example in spectral analysis is the Welch method which takes the arithmetic mean of the periodograms of overlapped data segments to reduce the variance of the estimation [72, 73]. In [68], the geometric mean is taken over windowed periodograms to reduce both the bias and variance.

From a geometric viewpoint means, such as the arithmetic, geometric, and harmonic mean, represent solutions to certain optimization problems. Examples in different contexts include the Fréchet mean, or Karcher mean. To be more specific, given a set of points $\{f_i, i = 1, \dots, m\}$, a general notion of mean is defined to be the point \bar{f} such that

$$\bar{f} := \arg \min_f \sum_i^m w_i d^2(f, f_i)$$

where d is a suitable distance and w_i 's are normalized weights which sum to 1. It is easy to check that when d is the Euclidean distance, \bar{f} is the arithmetic mean $\sum_i^m w_i f_i$.

In this chapter, we first illustrate the importance of means with the classical Welch method in spectral analysis. We then focus on the transportation mean for the weak continuity property of this distance. The transportation mean, as well as the median, for normalized scalar power spectra is worked out explicitly. In the multivariate case, we review some classical results for Gaussian distributions, and show that it coincides with the Bures distance between covariance matrices. Most importantly, besides the physically appealing interpretation, computation of the metric requires solving a linear matrix inequality (LMI). As a consequence, computations scale

nicely for problems involving large-size matrices, and linear prior constraints on the structure of means are easy to handle. At last, we consider the structured mean estimation problem from a set of positive semi-definite matrices, which has recently become a hot topic in the computer vision and pattern recognition area.

5.1 The Welch Method and Arithmetic Mean

Let $\{\mathbf{x}(t), t \in \mathbb{Z}\}$ be a zero-mean, second-order stationary process with covariance sequence

$$r(k) = E(\mathbf{x}(t)\mathbf{x}(t-k)^*), \quad \forall k, t \in \mathbb{Z}.$$

The spectral density function (PSD) of $\mathbf{x}(t)$ is defined to be the DTFT of the covariance sequence $r(k)$

$$f(\theta) = \sum_{k=-\infty}^{\infty} r(k)e^{-j\theta k}.$$

Given only a finite sample record $\{x(t), t = 1, 2, \dots, N\}$, the spectra density estimation by the periodogram is

$$\hat{f}(\theta) = \frac{1}{N} \left| \sum_{t=1}^N x(t)e^{-j\theta t} \right|^2.$$

There are extensive references on the statistical analysis of the periodogram and its smoothed versions [1, 74, 75, 76]. We shall only mention key results and skip the derivations.

First, the expectation of the periodogram estimator can be shown to be

$$E(\hat{f}(\theta)) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(\theta) W_B(\theta - \phi) d\phi,$$

where $W_B(\theta)$ is the DTFT of the triangular window (also called the Bartlett window)

$$w_b(k) = \begin{cases} 1 - \frac{|k|}{N}, & k = 0, \pm 1, \dots, \pm(N-1) \\ 0, & \text{otherwise.} \end{cases}$$

The main lobe of $W_B(\theta)$ has width about $1/N$ in frequency, which leads to the resolution limit of the periodogram. The side lobes of $W_B(\theta)$ cause the leakage of the energy from energy concentrated bands to the bands with less energy.

As the sample size N goes to infinity, the expectation of $\hat{f}(\theta)$ goes to $f(\theta)$ asymptotically. However, its variance does not converge to zero. Instead, for a general linear signal, which is

generated by passing the Gaussian white noise through a linear filter,

$$\lim_{N \rightarrow \infty} E\{[\hat{f}(\theta_1) - f(\theta_1)][\hat{f}(\theta_2) - f(\theta_2)]\} = \begin{cases} f^2(\theta), & \theta_1 = \theta_2 \\ 0, & \text{otherwise.} \end{cases}$$

Thus for a large sample record, the periodogram behaves like uncorrelated random variables with standard deviations being the true value of PSD. To deal with these limitations, several modified periodogram methods have been proposed, among which we will focus on Welch method due to its interpretation as the arithmetic mean of windowed periodograms.

The Welch method [72, 77] splits the data record into several overlapped segments, and then takes an average of all the periodograms of windowed segments. Consequently, the variance of the estimated PSD is reduced. Windowing each segment prior to the computation of periodogram helps reducing the correlation between two successive segments, and at the same time offers more flexibility in controlling the resolution of the estimated PSD.

To analyze statistic properties of the Welch method, let $\{x_k(t), t = 1, \dots, L\}$ represents the k -th segment of data with the length being L , and let the starting points of two successive segments being D units apart. Then

$$x_k(t) = x(t + (k-1)D), \quad k = 1, \dots, K$$

where $K = \lfloor (N - L + D)/D \rfloor$ with N being the total length of the data record. To each segment, the windowed periodogram is estimated as

$$\hat{f}_k(\theta) = \left(\frac{1}{L} \left| \sum_{t=1}^L w(t) x_k(t) e^{-j\theta t} \right|^2 \right) / \left(\frac{1}{L} \sum_{t=1}^L w^2(t) \right)$$

where the denominator denotes the power of the window $w(t)$. The Welch estimation of PSD is computed as:

$$\hat{f}_W(\theta) = \frac{1}{K} \sum_{k=1}^K \hat{f}_k(\theta).$$

It can be shown that the expectation of the Welch estimator is

$$E(\hat{f}_W(\theta)) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(\theta) h(\theta - \phi) d\phi,$$

where $h(\theta)$ is the normalized periodogram of window $w(t)$:

$$h(\theta) = \left(\frac{1}{L} \left| \sum_{t=1}^L w(t) e^{-j\theta t} \right|^2 \right) / \left(\frac{1}{L} \sum_{t=1}^L w^2(t) \right).$$

If the sample size N goes to infinity, the length L of each segment also goes to infinity. Then the estimated PSD by Welch method is asymptotically unbiased. In the case that $x(t)$ is a sample from a Gaussian process, $f(\theta)$ is flat over the passband of the estimator, and $h(\theta - \phi) = 0$ for $\theta < 0$ and $\theta > \pi$, then

$$\text{Var}\{\hat{f}_w(\theta)\} = \begin{cases} \frac{f^2(\theta)}{K} \left\{ 1 + 2 \sum_{k=1}^K \frac{K-i}{K} \rho(i) \right\}, & \theta \neq 0 \text{ or } \pi \\ \frac{2f^2(\theta)}{K} \left\{ 1 + 2 \sum_{i=1}^K \frac{K-i}{K} \rho(i) \right\}, & \theta = 0 \text{ or } \pi \end{cases}$$

where $\rho(i)$ is the correlation between \hat{f}_k and \hat{f}_{k+i} . Under the above conditions,

$$\rho(i) = \frac{\left[\sum_{k=1}^L w(k)w(k+iD) \right]^2}{\left[\sum_{k=1}^L w^2(k) \right]^2}.$$

If the length of the data record N is sufficiently large, we can split it into K segments without any overlapping. In this case, the correlation $\rho(i)$ will be zero, and the variance of Welch method is about

$$\text{Var}\{\hat{f}_w(\theta)\} = \begin{cases} \frac{f^2(\theta)}{K}, & \theta \neq 0 \text{ or } \pi \\ \frac{2f^2(\theta)}{K}, & \theta = 0 \text{ or } \pi. \end{cases}$$

Therefore, by dividing the data record into several segments and taking the arithmetic mean of their periodograms, the Welch estimation achieves smaller variance than the classical periodogram over the whole data record.

It is worth mentioning that the log-Welch estimation, which is the geometric mean of local periodograms, is believed to have variance-stabilizing property [75, 78, 79].

5.2 The Transportation Mean

The arithmetic mean mentioned in the previous section is often used to track the central tendency of the data, however it can be easily influenced by outliers. In some applications, arithmetic mean doesn't seem to be meaningful. For instance, if we choose \mathbf{x}_i 's to be Gaussian distributions with mean μ_i and same variance σ^2 , we would expect that the mean of \mathbf{x}_i 's can still be a Gaussian-type distribution which is peaked at the arithmetic mean of μ_i . However, it is not the case as shown in the following lemma.

Lemma 23. *Let*

$$f_1(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu_1)^2}{2\sigma^2}}, \quad f_2(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu_2)^2}{2\sigma^2}},$$

then their arithmetic mean $\bar{f}(x) = \frac{1}{2}[f_1(x) + f_2(x)]$ is bimodal if $|\mu_1 - \mu_2| > 2\sigma$, and unimodal otherwise.

The proof is referred to [80]. Figure 5.1 shows examples for both cases.

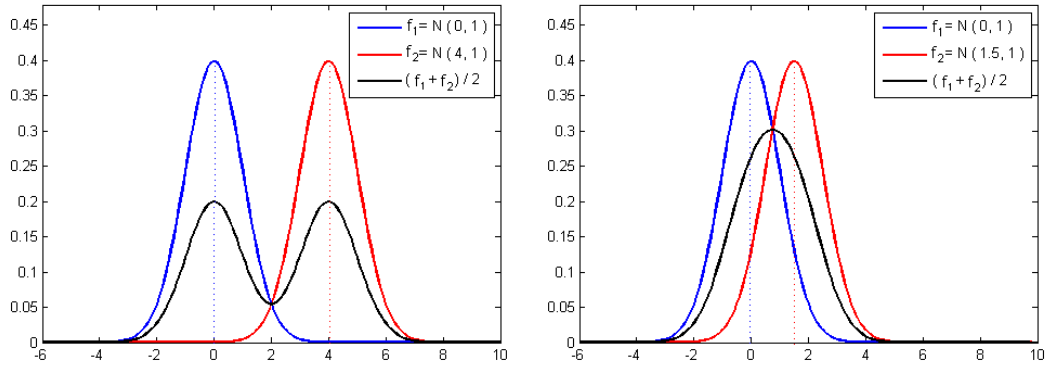


Figure 5.1: Arithmetic mean of two Gaussian distributions.

To this end, we introduce the transportation-based mean which preserves the Gaussianity. In addition, the transportation distance, which relates to the so-called Monge-Kantorovich transportation problem [27, pg. 212], has the weak continuity property. By definition, a sequence of measures $\{\mu_k\}$ converges to μ weakly if for any continuous bounded function ϕ , $\int \phi d\mu_k \rightarrow \int \phi d\mu$ as $k \rightarrow \infty$. This is desirable for a notion of convergence of power spectra since small changes in the power spectral density ought to reflect small changes in the statistics and vice versa. For instance, a “proper” distance between two line-spectra ought to decrease as two lines move closer, but it is not the case for distances that fail to be weakly continuous. Let

$$f_1(\theta) = \delta(\theta - \phi_1), \quad f_2(\theta) = \delta(\theta - \phi_2), \quad \text{and } \phi_1 \neq \phi_2,$$

where $\delta(\theta)$ is a Dirac delta function, which is infinity at $\theta = 0$ and 0 elsewhere such that its total integral is 1.

Let’s first consider the L_2 distance, which is

$$\|f_1 - f_2\|_2^2 = \int (f_1(\theta) - f_2(\theta))^2 d\theta = \int (\delta(\theta - \phi_1) - \delta(\theta - \phi_2))^2 d\theta = \infty.$$

The infinity of the total integral of the squared delta function can be reasoned in the following way. Consider a function which has the constant value $1/\varepsilon$ over an interval ε and 0 elsewhere. As $\varepsilon \rightarrow 0$, the function approaches to be a delta function. The total integral of its square is $1/\varepsilon$, which goes to infinity as $\varepsilon \rightarrow 0$.

If we consider the L_1 distance, then

$$\|f_1 - f_2\|_1 = \int |f_1(\theta) - f_2(\theta)| d\theta = \int |\delta(\theta - \phi_1) - \delta(\theta - \phi_2)| d\theta = 2.$$

Therefore, the L_1 distance between two delta functions is always 2 as far as $\phi_1 \neq \phi_2$.

For the line spectral estimation problems, one would like to quantify errors consistently with the corresponding errors in the frequency between identified lines. Therefore, the transportation distance, which has the property of weak continuity, is much more desired in this case than the L_1 and L_2 distances.

5.2.1 The one-dimensional case

For the completeness of this chapter, we provide the definition of the transportation distance in one dimensional case again. Let f_1 and f_2 be two nonnegative functions on $[a, b]$ with equal total integrals. Throughout, we consider the cost for moving one unit of mass from location θ to $\psi(\theta)$ to be $(\theta - \psi(\theta))^2$. Then their transportation distance is the minimum cost of moving f_1 to f_2 . Mathematically, it is defined to be

$$d_{w_2}(f_1, f_2) := \left\{ \int_a^b |\theta - \psi(\theta)|^2 f_1(\theta) d\theta \right\}^{\frac{1}{2}}$$

where the optimal transference plan $\psi(\theta)$ satisfies the constraint that

$$\int_a^\theta f_1(\sigma) d\sigma = \int_a^{\psi(\theta)} f_2(\sigma) d\sigma. \quad (5.1)$$

In general this transportation distance defines a metric on the space of (normalized) density functions where the geodesic distance between two densities in fact equals the actual distance between the two. In this very special one-dimensional case, the geodesic path f_τ ($\tau \in [0, 1]$) between points f_1 and f_2 is specified via

$$((1 - \tau) + \tau\psi'(\theta)) f_\tau((1 - \tau)\theta + \tau\psi(\theta)) = f_1(\theta), \quad (5.2)$$

where θ and $\psi(\theta)$ satisfy the constraint (5.1).

When both f_1 and f_2 are delta functions with peaks located at ϕ_1 and ϕ_2 respectively, the only way to transport f_1 to f_2 is to move the unit of mass from ϕ_1 to ϕ_2 . Thus, their transportation distance is clear to be

$$d_{W_2}(f_1, f_2) = |\phi_1 - \phi_2| \quad (5.3)$$

whose value decreases as ϕ_1 approaches ϕ_2 .

The transportation mean arises naturally in facility allocation. For example, one has several possible holes to fill each with probability w_i . Now one needs to allocate a pile of sand so that the cost of moving the sand to the possible holes will be minimized. Mathematically, it can be described in the following way:

Given a set of nonnegative functions $f_i(\theta)$, $i = 1, \dots, m$, $\theta \in [a, b]$ with equal total integrals being 1, the transportation mean is obtained through the following optimization problem:

$$\bar{f} = \arg \min_f \sum_{i=1}^m w_i (d_{W_2}(f, f_i))^2, \quad (5.4)$$

where w_i 's are weights with the sum 1. This problem has been studied in [81, 82, 83] for general multidimensional distributions. Below we shall provide the results for the one dimensional case where proofs are relatively straightforward.

Using the cumulative distribution functions, the problem in (5.4) is equivalent to

$$\begin{aligned} \bar{f} = \arg \min_f \quad & \sum_{i=1}^m w_i \int_a^b |\theta - \psi_i(\theta)|^2 dF(\theta) \\ \text{s.t.} \quad & F(\theta) = F_i(\psi_i(\theta)), \quad i = 1, \dots, m \end{aligned} \quad (5.5)$$

Let $t = F(\theta)$ and represent θ by an inverse function of F , that is

$$\theta = F^{-1}(t), \quad \text{and} \quad \psi_i(\theta) = F_i^{-1}(F(\theta)) = F_i^{-1}(t).$$

Since $f(\theta)$ is nonnegative, $F(\theta)$ must increase monotonically with respect to θ . Equivalently, $F^{-1}(t)$ has to increase monotonically with respect to t . In addition, $F^{-1}(t)$ has to satisfy the boundary condition, i.e., $F^{-1}(0) = a$ and $F^{-1}(1) = b$. Instead of solving the problem (5.5) directly, we can first seek the optimal F^{-1} . Defining $g(t) = F^{-1}(t)$, we consider the following

optimization problem:

$$\begin{aligned}
g^*(t) &= \arg \min_{g(t)} \sum_{i=1}^m w_i \int_0^1 |g(t) - F_i^{-1}(t)|^2 dt & (5.6) \\
\text{s.t. } & g'(t) \geq 0, \forall t \in [0, 1] \\
& g(0) = a \\
& g(1) = b.
\end{aligned}$$

After obtaining $g^*(t)$, we have the optimal cumulative distribution function $\bar{F}(g^*(t)) = t$, and the optimal mean can be calculated as

$$\bar{f}(g^*(t)) = 1/(g^*(t))'.$$

Theorem 24. *Let*

$$f_i(\theta) > 0, \forall \theta \in [a, b], \text{ and } \int_a^b f_i(\theta) d\theta = 1,$$

then the transportation mean \bar{f} defined in (5.5) is unique and has an explicit form

$$\bar{F}\left(\sum_{i=1}^m w_i F_i^{-1}(t)\right) = t$$

where \bar{F} is the cumulative distribution function of \bar{f} , and $F_i^{-1}(t)$ is the inverse cumulative distribution function of f_i .

Proof. We will first solve the problem (5.6). Note that $f_i(\theta)$ is given, therefore $F_i^{-1}(t)$ is known or can be calculated numerically. Since $f_i(\theta)$ is nonnegative,

$$(F_i^{-1})'(t) \geq 0, \forall t \in [0, 1] \quad (5.7)$$

and

$$F_i^{-1}(0) = a, \quad F_i^{-1}(1) = b. \quad (5.8)$$

If we ignore the constraints on $g(t)$, the unique minimizer is

$$\hat{g}^*(t) = \sum_{i=1}^m w_i F_i^{-1}(t). \quad (5.9)$$

Because of the conditions on $F_i^{-1}(t)$ in (5.7) and (5.8), $\hat{g}^*(t)$ satisfies

$$(\hat{g}^*)'(t) \geq 0, \forall t \in [0, 1], \quad F_i^{-1}(0) = a, \text{ and } F_i^{-1}(1) = b$$

which are exactly the constraints in problem (5.6). Therefore, (5.9) is the optimal solution for (5.6). Consequently, the optimal transportation mean \bar{f} in (5.5) can be obtained explicitly from

$$\bar{F}\left(\sum_{i=1}^m w_i F_i^{-1}(t)\right) = t.$$

□

In the case where f_i 's are delta functions

$$f_i(\theta) = \delta(\theta - \phi_i),$$

we have

$$F_i^{-1}(t) = \phi_i \text{ for } 0 < t < 1.$$

Let $\bar{\phi}$ denote the arithmetic mean of $\{\phi_i\}$

$$\bar{\phi} = \sum_{i=1}^m w_i \phi_i.$$

Then the cumulative function \bar{F} of the transportation mean satisfies

$$\bar{F}^{-1}(t) = \bar{\phi} \text{ for } 0 < t < 1.$$

Consequently,

$$\bar{f}(\theta) = \delta(\theta - \bar{\phi}).$$

As an example, we choose $\{f_i, i = 1, \dots, 8\}$ as spectral estimations of some sinusoidal signals. Figure 5.2 shows the results where $w_i = 1/8$. We can see that the transportation mean is still like a line spectra, with the location of the peak as the average of those of the given data.

To compare the transportation mean with the arithmetic mean for Gaussian distributions, we have the following corollary:

Corollary 25. *Let $\{f_i(\theta) = N(\mu_i, \sigma_i^2), i = 1, \dots, m\}$, then their transportation mean $\bar{f}(\theta) = N(\mu, \sigma^2)$ where μ and σ are the arithmetic means of $\{\mu_i\}$ and $\{\sigma_i\}$ respectively.*

Proof. Let $f_0(\theta) = N(0, 1)$, then $f_i(\theta)$ is obtained by translating and dilating $f_0(\theta)$, that is

$$\sigma_i f_i(\sigma_i \theta + \mu_i) = f_0(\theta), \quad i = 1, \dots, m.$$

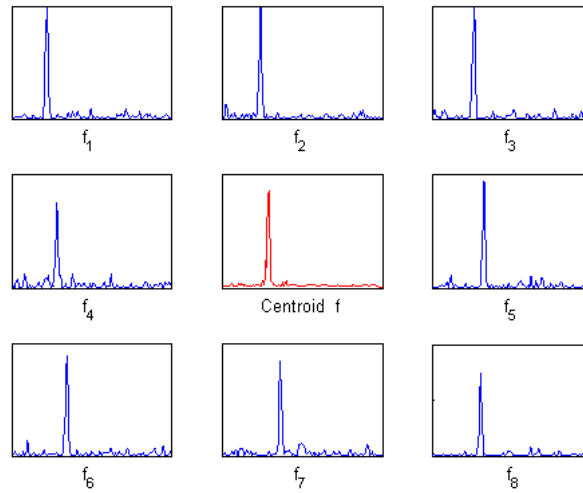


Figure 5.2: An example of transportation mean. The center of the second row shows the centroid, while others are the given spectra.

Consequently,

$$F_i(\sigma_i\theta + \mu_i) = F_0(\theta), \quad i = 1, \dots, m.$$

Using Theorem 24, the transportation mean satisfies

$$\bar{F} \left(\sum_{i=1}^m w_i (\sigma_i\theta + \mu_i) \right) = F_0(\theta),$$

which is the same as

$$\bar{F} \left(\left(\sum_{i=1}^m w_i \sigma_i \right) \theta + \sum_{i=1}^m w_i \mu_i \right) = F_0(\theta).$$

Thus $\bar{f}(\theta)$ is still a Gaussian distribution with mean and standard deviation being

$$\mu = \sum_{i=1}^m w_i \mu_i, \quad \sigma = \sum_{i=1}^m w_i \sigma_i,$$

which are the arithmetic means of $\{\mu_i\}$ and $\{\sigma_i\}$ respectively. \square

As an numerical example, we choose the given data to be two Gaussian distributions with same variance 1 and different means. Figure 5.3 shows the results for $w_i = 1/2$. Comparing with Figure 5.1, we can see clearly that the transportation mean is still a Gaussian distribution which is peaked at the center of two peaks.

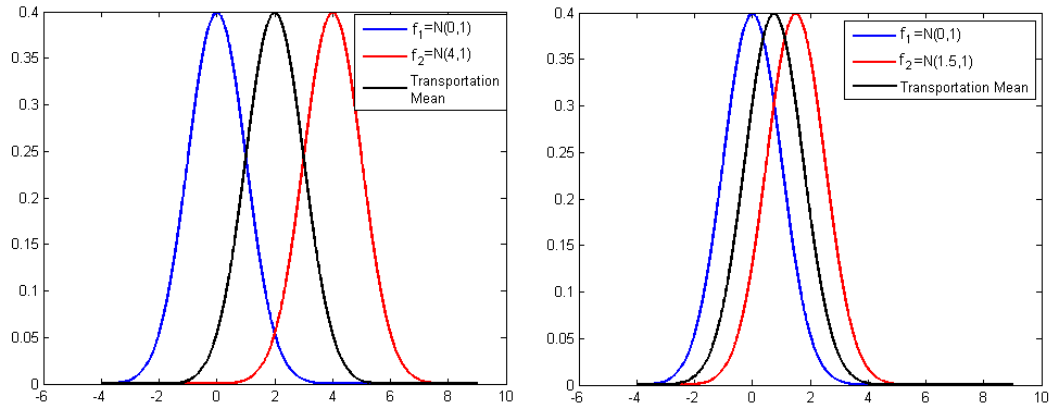


Figure 5.3: Transportation mean of two Gaussian distributions

To see the connection between the transportation mean and geodesic, we have the following lemma:

Lemma 26. *Let*

$$f_i(\theta) \geq 0, \forall \theta \in [a, b], \text{ and } \int_a^b f_i(\theta) d\theta = 1, i = 1, 2,$$

then the transportation mean \bar{f} defined in (5.5) lies on the geodesic path, that is

$$\bar{f}(\theta) = f_\tau(\theta)$$

where $\tau = w_2$, which is the weight for $d_{W_2}^2(\bar{f}, f_2)$.

Proof. Using Theorem 24, we have the cumulative function of the transportation mean

$$\bar{F}(w_1 F_1^{-1}(t) + w_2 F_2^{-1}(t)) = t.$$

Since

$$w_1 = 1 - w_2,$$

and

$$F_1(F_1^{-1}(t)) = F_2(F_2^{-1}(t)) = t,$$

according to the condition on the transportation geodesic in (5.2) which is equivalent to

$$F_\tau((1 - \tau)\theta + \tau\psi(\theta)) = F_1(\theta) = F_2(\psi(\theta)),$$

we have

$$\bar{F} = F_\tau$$

where $\tau = w_2$. Thus

$$\bar{f}(\theta) = f_\tau(\theta).$$

□

5.2.2 The multivariate-normal-distribution case

Consider now the transportation mean problem for the multidimensional distributions. Agueh and Carlier [81] have shown that it is equivalent to the multi-marginal optimal transportation problem which is defined as follows:

$$\begin{aligned} \inf_{\pi \in \Pi(\mu_1, \dots, \mu_m)} \sum_{i=1}^m w_i \int_{X_1 \times \dots \times X_m} |x_i - x|^2 d\pi(x_1, \dots, x_m) \\ \text{s.t. } \mathbf{x} = \sum_{i=1}^m w_i \mathbf{x}_i \end{aligned} \quad (5.10)$$

where \mathbf{x}_i is a random valuable with distributions f_i , \mathbf{x} corresponds to the transportation mean \bar{f} , and Π is the set of all possible joint distributions.

A sketch of argument is as follows. The transportation mean problem considers all the joint distributions π_i between the marginal f_i and the mean \bar{f} . Since the optimal transference plan ψ_i between \bar{f} and f_i determines the relation $\mathbf{x}_i = \psi_i(\mathbf{x})$, the set of optimal π_i 's dictates the unique relation between \mathbf{x}_i and \mathbf{x}_j as

$$\mathbf{x}_i = \psi_i \circ \psi_j^{-1}(\mathbf{x}_j).$$

Then using the inverse limit theorem [84, pg. 70], there exists a unique joint distribution $\pi(x, x_1, \dots, x_m)$ which yields all the marginals π_i 's. Then from the dual problem, it is easy to see that the optimal mean $\mathbf{x} = \sum_{i=1}^m w_i \mathbf{x}_i$. Therefore, the transportation mean problem is equivalent to (5.10).

The existence and uniqueness of the optimal transference plan of this multi-marginal transportation problem has been proved by Gangbo and Świąch [85]. The characterization of the optimal mapping can be found in [82]. To obtain the distribution of the mean \mathbf{x} , one usually needs to solve its dual problem which is an infinite-dimensional linear programming problem. The case that \mathbf{x}_i 's are Gaussian distributions has been of great interest from the view point of

optimal coupling, and the explicit expression for the optimal map has been obtained by various researchers. In particular, Olkin [86], Knott and Smith [83] derived the optimal transportation map between two multi-dimensional Gaussian distributions. This result was later generalized to three Gaussian distributions in [87], [88]. Below we start with the transportation distance between two zero-mean Gaussian distributions.

Let f_1 and f_2 be distributions of two zero-mean Gaussian random variables \mathbf{x}_1 and \mathbf{x}_2 with covariance matrices T_1 and T_2 respectively, and let $S = E(\mathbf{x}_1 \mathbf{x}_2')$ denote their correlation. Further, assuming that their joint distribution is Gaussian we obtain

$$d_{W_2}^2(f_1, f_2) = \min_S \left\{ \text{tr}(T_1 + T_2 - S - S') \mid \begin{bmatrix} T_1 & S \\ S' & T_2 \end{bmatrix} \geq 0 \right\}. \quad (5.11)$$

A closed form solution is easy to obtain [86, 83]:

$$S_0 = (T_1 \# T_2^{-1}) T_2, \quad (5.12)$$

where $(T_1 \# T_2^{-1})$ denotes the geometric mean between T_1 and T_2^{-1}

$$T_1 \# T_2^{-1} = T_2^{-\frac{1}{2}} (T_2^{\frac{1}{2}} T_1 T_2^{\frac{1}{2}})^{\frac{1}{2}} T_2^{-\frac{1}{2}}.$$

The transportation distance is given alternatively by

$$d_{W_2}(f_1, f_2) = \left(\text{tr}(T_1 + T_2 - 2(T_2^{\frac{1}{2}} T_1 T_2^{\frac{1}{2}})^{\frac{1}{2}}) \right)^{\frac{1}{2}},$$

and the optimal transferring map from \mathbf{x}_2 to \mathbf{x}_1 is

$$\mathbf{x}_1 = (T_1 \# T_2^{-1}) \mathbf{x}_2.$$

Since this is central to the rest of this subsection, we provide details as below.

Proof. Consider the Shur complement

$$P := T_1 - S T_2^{-1} S'$$

which is clearly nonnegative definite. Then, $S T_2^{-\frac{1}{2}} = (T_1 - P)^{\frac{1}{2}} U$, where $U U' = I$, and

$$S = (T_1 - P)^{\frac{1}{2}} U T_2^{\frac{1}{2}}. \quad (5.13)$$

Moreover,

$$\mathrm{tr}(S) = \mathrm{tr}((T_1 - P)^{\frac{1}{2}} U T_2^{\frac{1}{2}}) = \mathrm{tr}(T_2^{\frac{1}{2}} (T_1 - P)^{\frac{1}{2}} U). \quad (5.14)$$

Since T_1 and T_2 are given, minimizing $\mathrm{tr}(T_1 + T_2 - S - S')$ is the same as maximizing $\mathrm{tr}(S)$. Let $U_S \Lambda_S V_S'$ be the singular value decomposition of $T_2^{\frac{1}{2}} (T_1 - P)^{\frac{1}{2}}$, and

$$U_0 := \arg \max_U \{ \mathrm{tr}(T_2^{\frac{1}{2}} (T_1 - P)^{\frac{1}{2}} U) \mid U U' = I \}.$$

Then, U_0 must satisfy $V_S' U_0 = U_S'$ and

$$T_2^{\frac{1}{2}} (T_1 - P)^{\frac{1}{2}} U_0 = (T_2^{\frac{1}{2}} (T_1 - P) T_2^{\frac{1}{2}})^{\frac{1}{2}}. \quad (5.15)$$

From (5.14) we have $\mathrm{tr}(S) = \mathrm{tr}((T_2^{\frac{1}{2}} (T_1 - P) T_2^{\frac{1}{2}})^{\frac{1}{2}})$. Since $P \geq 0$, $\mathrm{tr}(S)$ is maximal when $P = 0$. Moreover, if $P = 0$,

$$\mathrm{rank} \left(\begin{bmatrix} T_1 & S \\ S' & T_2 \end{bmatrix} \right) \leq \mathrm{rank}(T_1),$$

and $T_2 = S_0' T_1^{-1} S_0$. Thus, setting $P = 0$ into (5.15), we have

$$U_0 = T_1^{-\frac{1}{2}} T_2^{-\frac{1}{2}} (T_2^{\frac{1}{2}} T_1 T_2^{\frac{1}{2}})^{\frac{1}{2}},$$

and consequently $S_0 = T_2^{-\frac{1}{2}} (T_2^{\frac{1}{2}} T_1 T_2^{\frac{1}{2}})^{\frac{1}{2}} T_2^{\frac{1}{2}}$. \square

Comparing now with the corresponding expression for the Bures distance in (2.13), we readily have the following.

Proposition 27. *For f_1 and f_2 Gaussian zero mean distributions with covariances T_1 and T_2 , respectively,*

$$d_B(T_1, T_2) = d_{W_2}(f_1, f_2).$$

We are now ready to state the transportation mean of several Gaussian distributions.

Theorem 28. *Given a set of Gaussian random variables $\{\mathbf{x}_i \sim N(0, T_i), i = 1, \dots, m\}$, there exists a unique transportation mean \mathbf{x} which is also a Gaussian with distribution $\mathbf{x} \sim N(0, T)$, and T satisfies*

$$T = \sum_{i=1}^m w_i (T^{\frac{1}{2}} T_i T^{\frac{1}{2}})^{\frac{1}{2}}. \quad (5.16)$$

We refer the complete proof of this theorem to [81]. Below, we provide a different derivation for the optimality condition in (5.16).

Proof. Let the transportation mean $\bar{f}(x)$ be a zero-mean Gaussian with covariance T , then

$$\sum_{i=1}^m w_i d_{W_2}^2(\bar{f}, f_i) = \sum_{i=1}^m w_i d_{\mathbb{B}}^2(T, T_i) = \sum_{i=1}^m w_i \operatorname{tr} \left\{ T + T_i - 2(T_i^{\frac{1}{2}} T T_i^{\frac{1}{2}})^{\frac{1}{2}} \right\}.$$

Small perturbations around the optimal T would not change the value of the objective function, i.e.,

$$\sum_{i=1}^m w_i d_{\mathbb{B}}^2(T + \Delta, T_i) - \sum_{i=1}^m w_i d_{\mathbb{B}}^2(T, T_i) = \text{HOT}(\Delta^2)$$

where $\text{HOT}(\Delta^2)$ means the high order terms on Δ^2 . We first consider the derivative of $X^{\frac{1}{2}}$ for a positive definite matrix X in terms of its Loewner matrix [35, p. 154]. Suppose X has the eigen-decomposition $X = U\Lambda U^*$. The Loewner matrix $F(X)$ of $X^{\frac{1}{2}}$ is

$$[F(X)]_{i,j} = \begin{cases} \frac{1}{2} \lambda_i^{-\frac{1}{2}}, & i = j \\ \frac{\lambda_i^{\frac{1}{2}} - \lambda_j^{\frac{1}{2}}}{\lambda_i - \lambda_j}, & i \neq j \end{cases}$$

Then the derivative of $X^{\frac{1}{2}}$ is $\partial_t(X + t\Delta)^{\frac{1}{2}}|_{t=0} = U(F(X) \circ (U^* \Delta U))U^*$ where \circ denotes the Schur product. As a result, for small perturbation Δ ,

$$\begin{aligned} \operatorname{tr}(X + \Delta)^{\frac{1}{2}} - \operatorname{tr}X^{\frac{1}{2}} &= \operatorname{tr}U(F(X) \circ (U^* \Delta U))U^* + \text{HOT}(\Delta^2) \\ &= \operatorname{tr}(F(X) \circ (U^* \Delta U)) + \text{HOT}(\Delta^2) \\ &= \sum_i \frac{1}{2} \lambda_i^{-\frac{1}{2}} (U_i^* \Delta U_i) + \text{HOT}(\Delta^2) \\ &= \sum_i \frac{1}{2} \operatorname{tr}(U_i \lambda_i^{-\frac{1}{2}} U_i^* \Delta) + \text{HOT}(\Delta^2) \\ &= \frac{1}{2} \operatorname{tr}(X^{-\frac{1}{2}} \Delta) + \text{HOT}(\Delta^2). \end{aligned}$$

Substitute X with $T_i^{\frac{1}{2}} T T_i^{\frac{1}{2}}$ and Δ with $T_i^{\frac{1}{2}} \Delta T_i^{\frac{1}{2}}$, we have

$$\operatorname{tr} \left(T_i^{\frac{1}{2}} (T + \Delta) T_i^{\frac{1}{2}} \right)^{\frac{1}{2}} - \operatorname{tr} \left(T_i^{\frac{1}{2}} T T_i^{\frac{1}{2}} \right)^{\frac{1}{2}} = \frac{1}{2} \operatorname{tr} \left((T_i^{\frac{1}{2}} T T_i^{\frac{1}{2}})^{-\frac{1}{2}} (T_i^{\frac{1}{2}} \Delta T_i^{\frac{1}{2}}) \right) + \text{HOT}(\Delta^2).$$

Consequently,

$$\begin{aligned} \sum_{i=1}^m w_i d_{\mathbb{B}}^2(T + \Delta, T_i) - \sum_{i=1}^m w_i d_{\mathbb{B}}^2(T, T_i) &= \text{tr} \Delta - \sum_{i=1}^m w_i \text{tr} \left((T_i^{\frac{1}{2}} T T_i^{\frac{1}{2}})^{-\frac{1}{2}} (T_i^{\frac{1}{2}} \Delta T_i^{\frac{1}{2}}) \right) + \text{HOT}(\Delta^2) \\ &= \text{tr} \left(\left(I - \sum_{i=1}^m w_i T_i^{\frac{1}{2}} (T_i^{\frac{1}{2}} T T_i^{\frac{1}{2}})^{-\frac{1}{2}} T_i^{\frac{1}{2}} \right) \Delta \right) + \text{HOT}(\Delta^2). \end{aligned}$$

The optimality condition is

$$I = \sum_{i=1}^m w_i T_i^{\frac{1}{2}} (T_i^{\frac{1}{2}} T T_i^{\frac{1}{2}})^{-\frac{1}{2}} T_i^{\frac{1}{2}}.$$

Since

$$T_i^{\frac{1}{2}} (T_i^{\frac{1}{2}} T T_i^{\frac{1}{2}})^{-\frac{1}{2}} T_i^{\frac{1}{2}} = T_i \# (T^{-1}) = T^{-\frac{1}{2}} (T^{\frac{1}{2}} T_i T^{\frac{1}{2}})^{\frac{1}{2}} T^{-\frac{1}{2}},$$

the equivalent optimality condition is

$$I = \sum_{i=1}^m w_i T^{-\frac{1}{2}} (T^{\frac{1}{2}} T_i T^{\frac{1}{2}})^{\frac{1}{2}} T^{-\frac{1}{2}}$$

or

$$T = \sum_{i=1}^m w_i (T^{\frac{1}{2}} T_i T^{\frac{1}{2}})^{\frac{1}{2}}.$$

□

Note that in the scalar case, $T = (\sum_{i=1}^m w_i T_i^{\frac{1}{2}})^2$ is the power mean of exponent $1/2$. In the matricial case, T can be seen as the unique power mean which has the minimum total squared Frobenius distance between $T^{\frac{1}{2}}$ and all the possible square roots of the T_i 's.

The existence and uniqueness of solution to the equation (5.16) has been proven in [81] by using the fixed-point theorem. The common way to solve this equation is by the iterative method, i.e.,

$$T_{(k+1)} = \sum_{i=1}^m w_i (T_{(k)}^{\frac{1}{2}} T_i T_{(k)}^{\frac{1}{2}})^{\frac{1}{2}}.$$

Although it works well in many numerical examples, it remains unknown whether this iterative procedure converges.

In the case that there are only two Gaussian distributions $f_0 = N(0, T_0)$ and $f_1 = N(0, T_1)$, the covariance matrix of their transportation mean can be easily obtained from a different perspective. Without loss of generality, let us assume $w_1 = w_2 = 1/2$. From Proposition 27, we have

$$d_{W_2}(f_0, f_1) = d_{\mathbb{B}}(T_0, T_1),$$

and the Bures distance corresponds to the straight line distance between $T_0^{\frac{1}{2}}$ and $T_1^{\frac{1}{2}}U$, where the unitary matrix $U = T_1^{-\frac{1}{2}}T_0^{-\frac{1}{2}}(T_0^{\frac{1}{2}}T_1T_0^{\frac{1}{2}})^{\frac{1}{2}}$. Therefore, the geodesic between them is

$$T_\tau = \left((1-\tau)T_0^{\frac{1}{2}} + \tau T_1^{\frac{1}{2}}U \right) \left((1-\tau)T_0^{\frac{1}{2}} + \tau T_1^{\frac{1}{2}}U \right)^*.$$

The mid point on the geodesic

$$T_{0.5} = \frac{1}{4} (T_0 + T_1 + (T_0 \# T_1^{-1})T_1 + T_1(T_0 \# T_1^{-1}))$$

has the property

$$d_B(T_0, T_{0.5}) = d_B(T_{0.5}, T_1) = \frac{1}{2}d_B(T_0, T_1).$$

Let $f_{0.5} = N(0, T_{0.5})$, equivalently we have

$$d_{W_2}(f_0, f_{0.5}) = d_{W_2}(f_{0.5}, f_1) = \frac{1}{2}d_{W_2}(f_0, f_1).$$

Thus $f_{0.5}$ is the transportation mean between f_0 and f_1 .

5.3 The Transportation Median in 1-D

Now let us consider geometric median which minimizes the sum of distance to each point:

$$\bar{x}_{GM} = \arg \min_x \sum_{i=1}^m w_i d(x, x_i) \quad (5.17)$$

where $d(\cdot, \cdot)$ is the chosen distance function and $\{x_i, i = 1, \dots, m\}$ are given points. Compared with the arithmetic mean, the geometric median is more robust to outliers, but it is in general difficult to compute.

When d is the Euclidean distance, it is well known that the geometric median is unique when the given points are not collinear. In the case of collinear, at least one of given points is a geometric median. For instance, when the number of given points is 2, these points always lie on a straight line. As a result, when $w_i = 1/2, i = 1, 2$, any point on the interval between these two points is their geometric median. When $w_1 \neq w_2$, then the point with the larger weight is the median. If points are not collinear, some numerical algorithms have been proposed to solve this convex optimization problem, such as Weiszfeld's algorithm [89, 90, 91] and semidefinite programming [92].

We shall call the transportation-distance-based geometric median as the transportation median, while reserve the name “geometric median” for the one with Euclidean distance. By following the similar transformation as we did for the transportation mean and using the same notation as in (5.6), we change the problem into the following one:

$$\begin{aligned}
g^*(t) &= \arg \min_{g(t)} \sum_{i=1}^m w_i \sqrt{\int_0^1 |g(t) - F_i^{-1}(t)|^2 dt} \\
\text{s.t. } &g'(t) \geq 0, \forall t \in [0, 1] \\
&g(0) = a \\
&g(1) = b.
\end{aligned} \tag{5.18}$$

If we discretize $g(t)$ and $F_i^{-1}(t)$ along t , the resulting problem becomes obtaining the geometric median of $\{F_i^{-1}(t), i = 1, \dots, m\}$. Thus we can use conventional algorithms for the geometric median to solve the problem. In the case that the given points are Gaussian distributions, we have the following lemma:

Lemma 29. *Let $\{f_i(\theta) = N(\mu_i, \sigma_i^2), i = 1, \dots, m\}$, then their transportation median $\bar{f}_{TM}(\theta) = N(\mu_{GM}, \sigma_{GM}^2)$ where $[\mu_{GM}, \sigma_{GM}]^T$ is the geometric median of $\{[\mu_i, \sigma_i]^T, i = 1, \dots, m\}$.*

Proof. Let

$$f_0(\theta) = N(0, 1),$$

then $f_i(\theta)$ is obtained by translating and dilating $f_0(\theta)$, that is

$$\sigma_i f_i(\sigma_i \theta + \mu_i) = f_0(\theta), \quad i = 1, \dots, m.$$

We use $F(\theta)$ to denote the cumulative distribution function of $f(\theta)$, then

$$F_i(\sigma_i \theta + \mu_i) = F_0(\theta), \quad i = 1, \dots, m.$$

Now let us use $g_0(t)$ to denote $F_0^{-1}(t)$, that is $F_0(g_0(t)) = t$. Then from the above equality, we obtain

$$\sigma_i g_0(t) + \mu_i = F_i^{-1}(t).$$

Let $g(t) = F^{-1}(t)$, then the optimal $g^*(t)$ corresponding to the optimal F^* is obtained through

the following optimization problem:

$$\begin{aligned}
g^*(t) &= \arg \min_{g(t)} \sum_{i=1}^m w_i \sqrt{\int_0^1 |g(t) - \sigma_i g_0(t) - \mu_i|^2 dt} \\
\text{s.t. } &g'(t) \geq 0, \forall t \in [0, 1] \\
&g(0) = -\infty \\
&g(1) = \infty.
\end{aligned} \tag{5.19}$$

Let $u(t) = 1, \forall t \in [0, 1]$, then the above optimization problem is simplified to

$$\begin{aligned}
g^*(t) &= \arg \min_{g(t)} \sum_{i=1}^m w_i \|g - \sigma_i g_0 - \mu_i u\|_2 \\
\text{s.t. } &g'(t) \geq 0, \forall t \in [0, 1] \\
&g(0) = -\infty \\
&g(1) = \infty
\end{aligned} \tag{5.20}$$

where $\| \cdot \|_2$ is the L_2 norm. We can see that $\{\sigma_i g_0 + \mu_i u, i = 1, \dots, m\}$ are points on the plane spanned by $g_0(t)$ and $u(t)$. If we only consider the objective function and do not take the constraints into account, then the optimal $g(t)$ must lie in the same plane, that means

$$g(t) = \sigma g_0(t) + \mu u(t).$$

We can see that if $\sigma > 0$, $g(t)$ above satisfies all the constraints. Since $F(g(t)) = F(\sigma g_0(t) + \mu u(t)) = t$, the optimal $f(\theta)$ must be a Gaussian distribution with standard deviation σ and mean μ . To obtain the optimal σ and μ , we need to solve the following optimization problem:

$$\arg \min_{\alpha > 0, \beta} \sum_{i=1}^n w_i \|(\alpha - \sigma_i) g_0 + (\beta - \mu_i) u\|_2$$

Since f_0 is symmetric, we have

$$F_0(\theta) = 1 - F_0(-\theta),$$

which leads to

$$g_0(t) = -g_0(1-t).$$

Thus

$$\int_0^1 g(t) u(t) dt = 0$$

which means $g(t)$ and $u(t)$ are orthogonal. It is easy to see that $\|u\|_2 = 1$. Now we need to compute the norm of $g_0(t)$. Let us start from the variance of $f_0(t)$:

$$\int_{-\infty}^{\infty} \theta^2 dF_0(\theta) = 1$$

If we change the variable on the left side to $t = F_0(\theta)$, and let $\theta = g_0(t)$, we have

$$\int_0^1 g_0(t)^2 dt = 1$$

which means

$$\|g_0\|_2 = 1.$$

Therefore, $g(t)$ and $u(t)$ are orthonormal. The optimal α and β can be obtained through

$$\arg \min_{\alpha > 0, \beta} \sum_{i=1}^m w_i \sqrt{(\alpha - \sigma_i)^2 + (\beta - \mu_i)^2}$$

which means $[\alpha^* \ \beta^*]^T$ is the geometric median of $\{[\alpha_i \ \beta_i]^T, i = 1, \dots, m\}$. \square

Note that since the points $\{\mu_i, i = 1, \dots, m\}$ are collinear, their geometric median may not be unique. For example, let $\mu_i = i$ and $w_i = 1/4$ for $i = 1, \dots, 4$, then it is easy to verify that both μ_2 and μ_3 are geometric medians. As a consequence, the transportation median may not be unique in contrast to the uniqueness of the transportation mean.

The connection between the transportation median and geodesic is established through the following lemma:

Lemma 30. *Let*

$$f_i(\theta) \geq 0, \forall \theta \in [a, b], \text{ and } \int_a^b f_i(\theta) d\theta = 1, i = 1, 2,$$

then the transportation median \bar{f}_{TM} defined in (5.17) is

$$\bar{f}_{TM}(\theta) = \begin{cases} f_1(\theta), & w_1 > w_2; \\ f_2(\theta), & w_2 > w_1; \\ f_\tau(\theta), \forall \tau \in [0, 1] & w_1 = w_2; \end{cases}$$

where $f_\tau(\theta)$ is any point on the geodesic path between f_1 and f_2 .

Proof. Using (5.6), the optimal $g^*(t)$ is the geometric median of $F_1^{-1}(t)$ and $F_2^{-1}(t)$. Since two points are always collinear, it is easy to see that

$$g^*(t) = \begin{cases} F_1^{-1}(t), & w_1 > w_2; \\ F_2^{-1}(t), & w_2 > w_1; \\ (1 - \tau)F_1^{-1}(t) + \tau F_2^{-1}(t), \forall \tau \in [0, 1] & w_1 = w_2. \end{cases}$$

Since the cumulative function of the transportation median satisfies

$$\bar{F}_{TM}(g^*(t)) = t,$$

using the condition on the transportation geodesic in (5.2) which is equivalent to

$$F_\tau((1 - \tau)\theta + \tau\psi(\theta)) = F_1(\theta) = F_2(\psi(\theta)),$$

we have

$$\bar{f}_{TM}(\theta) = \begin{cases} f_1(\theta), & w_1 > w_2; \\ f_2(\theta), & w_2 > w_1; \\ f_\tau(\theta), \forall \tau \in [0, 1] & w_1 = w_2. \end{cases}$$

□

5.4 The Kullback-Leibler Mean

The transportation mean is not unique in preserving Gaussianity. Below we shall provide an alternative with the same property, e.g. the Kullback-Leibler (KL) mean.

Recall that the KL divergence is defined as

$$d_{\text{KL}}(f_0||f_1) := \int_{\theta} f_0 \log \frac{f_0}{f_1} d\theta,$$

where f_0 and f_1 are two normalized density functions with total integrals 1. We define the KL mean in a way similar to the transportation mean. For a set of given density functions f_i 's, $i = 1, \dots, m$, we seek the density function \bar{f} such that the sum of the Kullback-Leibler distances between \bar{f} and f_i is minimized, that is

$$\bar{f} = \left\{ \arg \min_f \frac{1}{m} \sum_{i=1}^m d_{\text{KL}}(f||f_i), \text{ such that } \int f = 1 \right\}.$$

Theorem 31. *The KL mean of $\{f_i, i = 1, \dots, m\}$ is the scaled geometric mean, i.e.,*

$$\bar{f} = \sqrt[m]{\prod_{i=1}^m f_i} / \int \sqrt[m]{\prod_{i=1}^m f_i}.$$

Proof. Consider the Lagrangian

$$L(f, \lambda) = \frac{1}{m} \sum_{i=1}^m d_{\text{KL}}(f || f_i) + \lambda \left(\int f - 1 \right),$$

which is equivalent to

$$L(f, \lambda) = \int f \log f - \int f \left(\frac{1}{m} \sum_{i=1}^m \log f_i \right) + \lambda \left(\int f - 1 \right).$$

To minimize $L(f, \lambda)$ with respect to f , we use the calculus of variation, that is to consider

$$L(f + \delta, \lambda) = L(f, \lambda) + \int \delta \left(1 + \lambda + \log f - \frac{1}{m} \sum_{i=1}^m \log f_i \right) + HOT(\delta).$$

Then the optimal f must satisfy the condition that

$$1 + \lambda + \log f - \frac{1}{m} \sum_{i=1}^m \log f_i = 0,$$

which means

$$f = \sqrt[m]{\prod_{i=1}^m f_i} \cdot e^{-1-\lambda}.$$

The Lagrangian multiplier λ is chosen such that the integration of f is 1. Therefore, the KL mean is

$$\bar{f} = \sqrt[m]{\prod_{i=1}^m f_i} / \int \sqrt[m]{\prod_{i=1}^m f_i}.$$

□

When f_i 's are Gaussian distributions, the KL mean is again a Gaussian distribution according to the following corollary.

Corollary 32. *Let each given density function $f_i \sim N(\mu_i, T_i)$, then the KL mean*

$$\bar{f} \sim N(\mu, T),$$

where

$$T = \left(\frac{1}{m} \sum_{i=1}^m T_i^{-1} \right)^{-1},$$

and

$$\mu = T \left(\frac{1}{m} \sum_{i=1}^m T_i^{-1} \mu_i \right).$$

Proof. Since each density function is

$$f_i = \frac{1}{(2\pi)^{\frac{n}{2}} |T_i|^{\frac{1}{2}}} e^{-\frac{1}{2}(x-\mu_i)^T T_i^{-1} (x-\mu_i)},$$

we have

$$\begin{aligned} \sqrt[m]{\prod_{i=1}^m f_i} &= \frac{1}{(2\pi)^{\frac{n}{2}} \sqrt[m]{\prod_{i=1}^m |T_i|^{\frac{1}{2}}}} e^{-\frac{1}{2} \left(\frac{1}{m} \sum_{i=1}^m (x-\mu_i)^T T_i^{-1} (x-\mu_i) \right)} \\ &= \frac{1}{(2\pi)^{\frac{n}{2}} \sqrt[m]{\prod_{i=1}^m |T_i|^{\frac{1}{2}}}} e^{-\frac{1}{2}(x-\mu)^T T^{-1} (x-\mu)} e^{-\frac{1}{2} \left(\frac{1}{m} \sum_{i=1}^m \mu_i^T T_i^{-1} \mu_i - \mu^T T^{-1} \mu \right)}, \end{aligned}$$

where

$$T = \left(\frac{1}{m} \sum_{i=1}^m T_i^{-1} \right)^{-1},$$

and

$$\mu = T \left(\frac{1}{m} \sum_{i=1}^m T_i^{-1} \mu_i \right).$$

Therefore, the KL mean

$$\bar{f} = c e^{-\frac{1}{2}(x-\mu)^T T^{-1} (x-\mu)},$$

where c is the constant such that the integration of \bar{f} is 1. Equivalently, we obtain

$$\bar{f} \sim N(\mu, T).$$

□

Note that since the Kullback-Leibler distance is asymmetrical, an alternative of the KL mean could be

$$\bar{f} = \left\{ \arg \min_f \frac{1}{m} \sum_{i=1}^m d_{\text{KL}}(f_i || f), \text{ such that } \int f = 1 \right\}.$$

It is easy to verify that the resulting \bar{f} is nothing but the arithmetic mean of f_i . In this case, the mean of Gaussian distributions is no longer Gaussian.

5.5 Means of Positive Definite Matrices

Thus far we have focused on the means of distributions. In the case of multivariate Gaussian distributions, we equivalently obtain the means of covariance matrices. The problem of constructing means of several positive matrices has received much attention in recent years. In particular, generalizing the geometric mean to more than two positive matrices is mathematically highly non-trivial [35]. In [93], Ando, Li and Mathias proposed to define the geometric mean of more than two matrices as the limit of a sequence. This work motivated Petz and Temesi to generalize other types of means through a similar limit procedure [94]. Although such definitions of means are conceptually simple, they are in general very demanding from a computational standpoint.

The optimization-problem-based definition of means can be readily extended to positive matrices. To be more specific, given a (finite) set of positive semi-definite matrices $\{T_i, i = 1, \dots, m\}$, we can define the mean

$$T^* := \arg \min_{T \geq 0} \frac{1}{m} \sum_{i=1}^m d^2(T, T_i) \quad (5.21)$$

associated with a suitable distance $d(T, T_i)$, provided that such a unique minimizer exists. For example, Moahker [59], Bhatia and Holbrook [42] defined a notion of geometric mean as the unique minimizer with respect to the log-deviation distance. To reduce the computational burden, a third type of geometric mean was proposed in [95] where the distance $d(T, T_i) = \|\log T - \log T_i\|_F$ was used. The latter has the closed-form expression

$$T^* = \exp\left(\frac{1}{m} \sum_{i=1}^m \log T_i\right).$$

Note that unlike the geometric mean defined by Ando, Li and Mathias, the above mean can be extended to the case of a countably infinite set or even a continuum of matrices (e.g., in the form $\exp(\int \log T(\theta) d\theta)$).

In this section, we are particularly interested in the structured mean estimation problems, such as estimating a structured covariance matrices from multiple sample covariance matrices, where the corresponding time series share certain characteristics. This is the case when we collect multiple time series from different sensors and seek to estimate a common component. This problem arises in several areas such as computer vision [96, 70, 71] and financial data analysis [97]. Below, we consider the concept of means as a way to fuse together sampled

statistics and the relevant information from different sources. Accordingly, we modify (5.21) into

$$T^* := \arg \min_{T \in \mathcal{T}} \frac{1}{m} \sum_{i=1}^m d^2(T, T_i), \quad (5.22)$$

for a metric $d(T, T_i)$, or a divergence $d(T||T_i)$ in its place (and a suitable choice of exponent), while \mathcal{T} represents a class based on priors and the structure of the problem at hand. Below, we consider each of the distances/divergences mentioned earlier.

Mean based on KL divergence

The optimization problem based on d_{KL} is

$$\min_{T \in \mathcal{T}} \left\{ \frac{1}{m} \sum_{i=1}^m (\log |T_i| - \log |T| + \text{tr}(T T_i^{-1}) - n) \right\},$$

which is equivalent to

$$\min_{T \in \mathcal{T}} \left\{ -\log |T| + \text{tr}(T \hat{T}^{-1}) \right\}, \quad (5.23)$$

where $\hat{T} = \left(\frac{1}{m} \sum_{i=1}^m T_i^{-1} \right)^{-1}$ is the harmonic mean of the T_i 's. On the other hand, if the likelihood divergence $d_{\text{KL}}(T_i||T)$ is used, the optimization problem

$$\min_{T \in \mathcal{T}} \left\{ \frac{1}{m} \sum_{i=1}^m (-\log |T_i| + \log |T| + \text{tr}(T_i T^{-1}) - n) \right\}$$

is equivalent to

$$\min_{T \in \mathcal{T}} \left\{ \log |T| + \text{tr}(\hat{T} T^{-1}) \right\}, \quad (5.24)$$

where $\hat{T} = \frac{1}{m} \sum_{i=1}^m T_i$ is the arithmetic mean of T_i 's.

Mean based on log-deviation

Using the log-deviation distance

$$d_{\text{Log}}(T, T_i) = \|\log(T_i^{-1/2} T T_i^{-1/2})\|_F, \quad (5.25)$$

which is a geodesic distance on the manifold of positive definite matrices, the optimization problem

$$\min_{T \in \mathcal{T}} \left\{ \frac{1}{m} \sum_{i=1}^m \|\log(T_i^{-1/2} T T_i^{-1/2})\|_F^2 \right\}$$

is not convex in T . If the admissible set \mathcal{T} is relaxed to the set of positive definite matrices, the minimizer is precisely the geometric mean considered by Moakher [59], Bhatia and Holbrook [42], and it is the unique positive definite solution to the following equation

$$\sum_{i=1}^m \log(T_i^{-1} T) = 0. \quad (5.26)$$

When $N = 2$, T is the unique geometric mean between two positive definite matrices

$$T = T_1 \# T_2 = T_1^{\frac{1}{2}} (T_1^{-\frac{1}{2}} T_2 T_1^{-\frac{1}{2}})^{\frac{1}{2}} T_1^{\frac{1}{2}}.$$

This notion of mean has been used in the area of computer vision and pattern recognition to detect e.g., humans in still images and the covariance matrices are used as object descriptors [70]. One may consider instead the linearized optimization problem

$$\min_{T \in \mathcal{T}} \left\{ \sum_{i=1}^m \|T_i^{-1/2} T T_i^{-1/2} - I\|_F^2 \right\}$$

which is clearly convex in T .

Mean based on transportation/Bures distance

The transportation mean problem with only the positive semi-definite constraints has been discussed in great details in Section 5.2.2. Considering the admissible set \mathcal{T} , the optimization problem becomes

$$\min_{T \in \mathcal{T}, S_i} \left\{ \frac{1}{m} \sum_{i=1}^m \text{tr}(T + T_i - S_i - S_i') \mid \begin{bmatrix} T & S_i \\ S_i' & T_i \end{bmatrix} \geq 0, \forall i = 1, \dots, m \right\}. \quad (5.27)$$

Although the result in (5.16) is no longer valid when \mathcal{T} is further constrained, the problem (5.27) is a semi-definite program and thus computationally advantageous when compared with the other options.

5.6 Extensions

So far we have discussed various notions of means for scalar-valued distributions and positive definite matrices. It would be interesting to apply these means to the clustering and classification problems. It is also fascinating to consider classification problems on a Riemannian manifold. Simple concepts in the Euclidean geometry may not have analogies on a manifold. For example, we know that a straight line divides a plane into two. However on a torus, one geodesic can not divide the torus into halves. Computations, such as the shortest distance of a point to a geodesic, may not be trivial tasks either.

Chapter 6

Applications

The geometric properties of spectral densities provide a variety of tools for spectral analysis. By looking at each spectral density as a point on a manifold, we can apply our intuition from the Euclidean geometry to develop a viewpoint for spectral analysis. For example, we may model the deformation between two given density functions by using a geodesic between them. We can automatically separate the phonemes of speech data based on the distance between two estimated spectral densities. These ideas will be developed in more details in the following sections.

6.1 Speech Morphing

Modeling and synthesizing speech is a well studied subject driven by the availability of cheap, high speed processors and by a wide range of applications; these include wireless communications, voice mail, data compression, speech recognition, speaker identification, text-to-speech translation, altering or synthesizing voice with particular characteristics as well as a multitude of applications in the entertainment industry [98]. The purpose of this work is to consider and study geodesic paths between power spectral density functions as a means to morph voice of one individual to the voice of another. Despite great strides in the theory and technology of speech processing, speech morphing is still in an early phase and far from being a standard application [99]. We first provide a rudimentary exposition of certain standard techniques and practices in modeling speech and then we discuss our experimental results.

Speech sounds are produced by acoustic excitation of cavities—the vocal and nasal tracts,

manipulated by the position and stiffness of the tongue, lips, jaw, etc. The excitation is produced by the vocal chords in the larynx, or by turbulence at other parts of the vocal system (as with fricatives **f**, **sh** or affricates **j**, etc., where the tongue and lips create suitable constrictions that generate turbulence). Typically, the shape of the vocal tract results in certain resonances called formants. These dominate the character of voiced sounds. However, coupling of the oral and nasal tracts, as in nasal sounds, causes the power spectra to have deep “valleys” as well. This is also the case when the excitation originates in the interior of the vocal tract, instead of the glottis. Typical models for speech suggest ~ 25 [ms] intervals of “quasi-stationarity”. Each ~ 25 [ms] fragment is represented by a linear filter driven by an excitation signal. Typically, the latter is either a periodic excitation, e.g., a train of pulses, which creates a pitch, or white noise. Voiced sounds typically require the filter to be autoregressive (AR) since the formant-structure is mainly responsible for their character. However, in general, pole-zero models are appropriate. Thus, common practice for speech analysis and synthesis is mainly based on either construction of detailed models or on linear time-frequency techniques which identify/generate the spectral content over sufficiently small time-intervals.

Speech morphing is the process of transforming one person’s speech pattern (e.g., Alice’s) into another’s (Bob’s), gradually, creating a new pattern with a distinct identity, while preserving the speech-like quality and content of the spoken sentence. In practice, there are several technical issues that need to be dealt with depending on the application. A typical experiment requires first that A (i.e., Alice) and B (i.e., Bob) recite the same sentence. These sentences need to be segmented appropriately and an exact correspondence be drawn between the beginning and end of the various quasi-stationary intervals that correspond to similar sounds. Time-frequency analysis can be used to automate marking of such intervals. Then, a suitable algorithm is required to interpolate the spectral qualities of the sounds produced by A and B.

Earlier attempts can be roughly classified in two groups, those who use direct time-frequency models and those who use linear predictive and other nonlinear methods for identifying suitable models. For instance, Abe [100] modified spectra by stitching together the low frequency part of Alice’s spectrum, below a pre-specified frequency, with the part of Bob’s power spectrum above the given frequency. This pre-specified frequency is then used as the control parameter that regulates the mixing. Ye [101] proposed a “perceptually weighted” linear transformation method based on a Gaussian mixture model for the power spectra of the two subjects. Kawahara [102]

manually selected anchor points to define correspondence between two time-frequency representations for Alice and Bob. Based on these anchor points, a piecewise bilinear transformation was used to map the target time-frequency coordinate onto the reference coordinate. Others focus on the modification of the excitation signal and AR spectral parameter; Goncharoff [103] interpolated the AR-spectra by pole shifting. He first applied a perceptually-based pole pairing algorithm, then generated a path between the pole pairs to achieve linear changes in pole’s frequency and bandwidth. Pfitzinger [99] on the other hand applied a dynamic programming technique to align the residuals and power spectra of AR models of the respective speech signals.

The approach taken in our work shares a number of basic steps with those earlier attempts, such as marking and processing separately quasi-stationary segments over 25 [ms] time intervals, etc. The key difference is in the algorithm for interpolating the resulting power spectra—in our work we suggest that this can be most conveniently, and perhaps naturally, effected by following geodesics in suitable metrics (such as those discussed in Chapter 2). Below, we briefly discuss the steps taken for analyzing a voiced phoneme for A and B, generating the respective spectra, and then generating the morphed sound. Examples of geodesics of complete words and sentences are posted at [104] in an audible format.

For analyzing a voiced phoneme we use linear prediction techniques. The basic framework is shown in Figure 6.1. Linear prediction analysis is followed by generation of a power spectrum at $\tau \in [0, 1]$ on the geodesic path, and then followed by synthesis. Linear prediction is used to obtain the coefficients for a modeling filter as well as to estimate the pitch. The voiced sound that is being analyzed (an “a”) is shown as two superimposed time-waveforms in Figure 6.2 (one for speaker A and one for B); it was sampled at 8KHz. For modeling, we choose an AR model of order 14. The frame size is 25 [ms], and the frame interval is 12 [ms]. A standard “pre-emphasis” filter is used to reduce the low-frequency content in the signal. The filtered data is weighted using a Hamming window. For modeling, we choose to use standard methods; we use the autocorrelation method for estimating the covariance lags and then the Levinson-Durbin method for obtaining the coefficients of the AR model. The AR model of the phoneme, for each of the two speakers, provides a corresponding power spectral density. A power spectrum at a point $\tau \in [0, 1]$ on the geodesic path is then determined. We estimate the pitch period for speakers A and B, using either residual-based estimation or Harmonic sinusoidal wave-based estimation. These periods are linearly interpolated along the path. The synthesized power

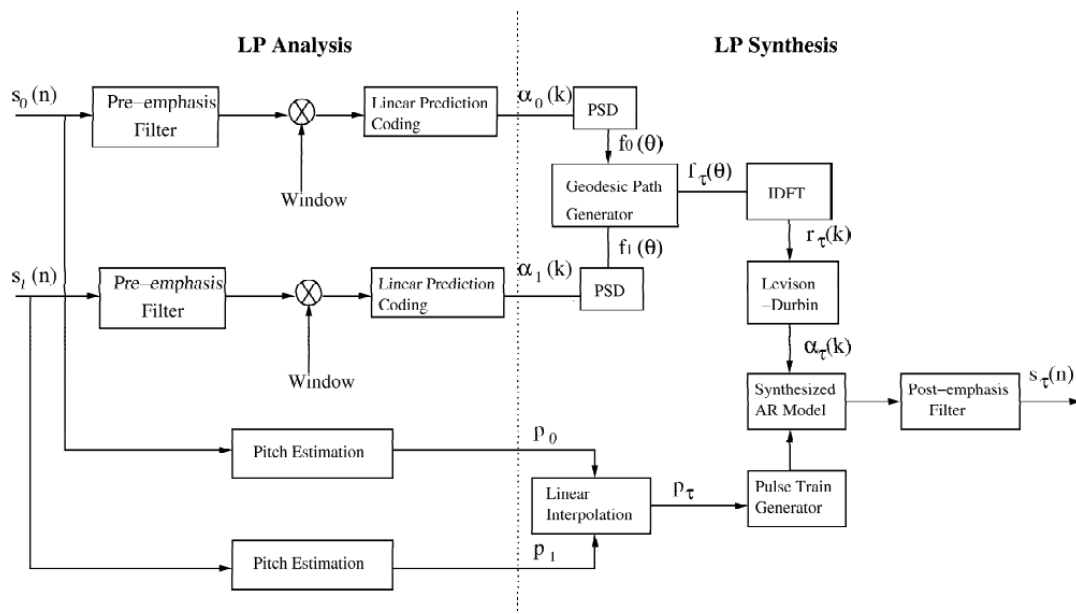


Figure 6.1: Linear prediction analysis and synthesis framework

spectral densities are approximated as AR-spectra and, finally, a pulse train with the interpolated pitch period is used to drive the synthesized AR-filter (at $\tau \in [0, 1]$ on the geodesic) in order to produce the synthesized sound. A post-emphasis filter is used to compensate the effect of the pre-emphasis filter—a standard practice.

In synthesizing complete sentences, a number of issues are critical but relatively standard. For instance, dynamic time warping is needed for an automatic alignment between two persons' speech, which may have different durations for the same phoneme and detecting voiced and unvoiced parts within a sentence. More complicated methods have been suggested to further improve the voice quality and some of these could be incorporated in the above framework. For example, Wong *et al.* [105] suggest that the AR coefficients be estimated when the glottis is closed. This requires further processing and, for high-pitch speakers, it becomes challenging. Others suggest a glottal-flow-like excitation [106], such as in the Liljencrants-Fant (LF) model, instead of simple pulse trains. For the purpose of illustrating the idea of spectral morphing, we just used the simple Linear Prediction model and did not involve ourselves with these advanced techniques.

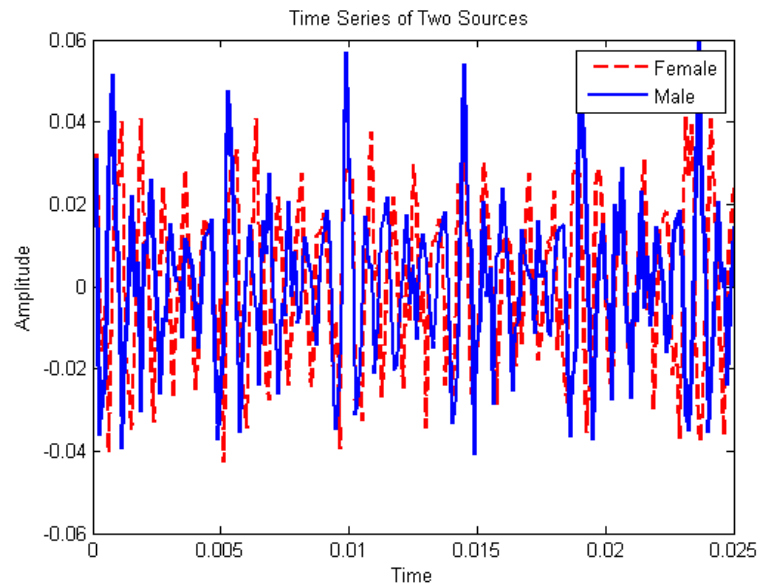


Figure 6.2: Time signals corresponding to the phoneme “a” for A (female) and B (male)

The spectral deformations by the three alternative geodesics (2.2), (2.6), and (2.11) are shown in Figures 6.4, 6.5 and 6.6 respectively. Though the effect on the acoustic quality is very subjective, we find that geodesic (2.6) has surprisingly good acoustic qualities, in spite of the fact that visually, in Figures 6.4 and 6.5, there is an apparent “fade in” and “fade out” of the respective formants in the two power spectra. The voice quality by geodesic (2.8) has some artifacts, which may due to the gradually transport of formant locations and our ears may be very sensitive to it.

6.2 Phoneme Separation

We consider the classic problem of segmenting speech signals into individual phonemes. This is a key step in current-day speech recognition applications, but can be equally well motivated based on speech coding and data compression. The mathematical problem is that of identifying time instances where a non-stationary waveform undergoes significant changes. It is customary to identify and employ a variety of features and cues for knowledge-based recognition and, accordingly, classification of sounds. There is a substantial literature on the subject for which

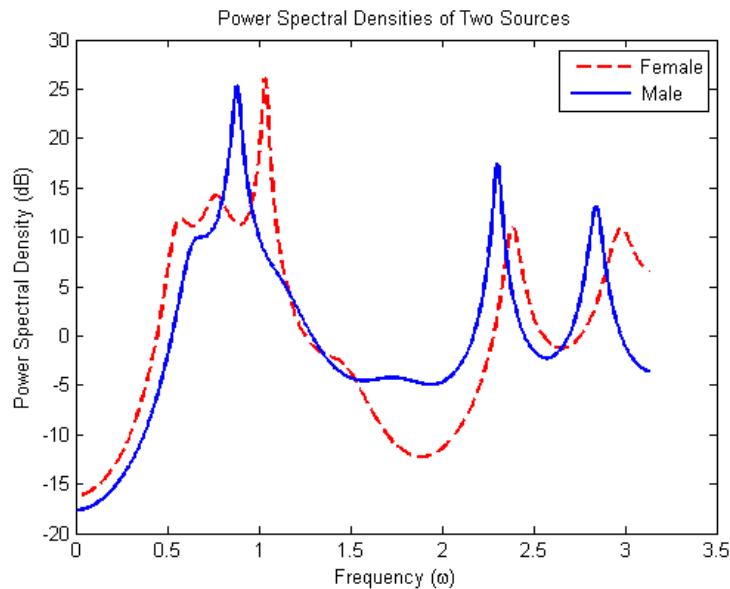


Figure 6.3: Power spectra for subjects A (female) and B (male)

we refer to [107, 108, 109, 110, 111] and the references therein. We want to explore the potential of certain natural metrics for comparing power spectra. Thus, we consider signals as consisting of “stationary segments” where second-order statistics contain most useful information. We estimate the power spectra based on sliding windows, and base our segmentation process on the relative distance between power spectra of neighboring windows. We utilize a (new) metric which quantifies differences in predictive qualities of two power spectra, and compare this with two other natural notions of distance (which can be motivated by analogous rationale). We also compare these with the well-known Itakura-Saito distance which has similarities with the aforementioned metric. We present results on a particular voice signal which has been marked by a human specialist (“expert”). These data come from the American-English DARPA-TIMIT database of [112]. Interestingly, the result of our, rather direct usage of metrics, is very similar to the marking by the “expert.”

For the experimental component of this work, we apply the prediction based geodesic distance to speech signal analysis. We compare neighboring sliding windows and compute the distance between the corresponding power spectra. We use the following four distances:

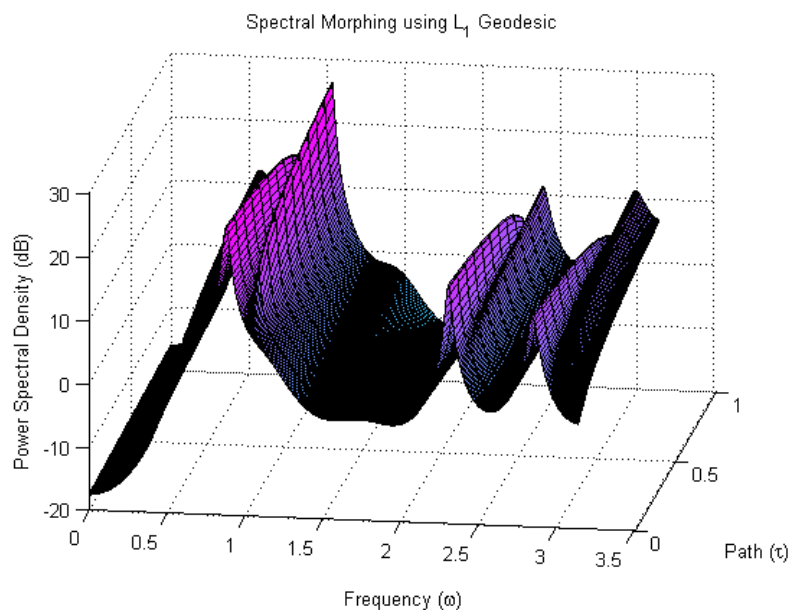


Figure 6.4: Geodesic path between the two spectra following L_1 distance

prediction geodesic dist.	$\sqrt{\int_{-\pi}^{\pi} \left(\log \frac{f_1(\theta)}{f_0(\theta)} \right)^2 \frac{d\theta}{2\pi} - \left(\int_{-\pi}^{\pi} \log \frac{f_1(\theta)}{f_0(\theta)} \frac{d\theta}{2\pi} \right)^2}$
L_2 -log	$\ \log(f_0) - \log(f_1)\ _2$
L_1 -log	$\ \log(f_0) - \log(f_1)\ _1$
Itakura-Saito	$\left\ \frac{f_0}{f_1} - 1 - \log\left(\frac{f_0}{f_1}\right) \right\ _1$

The demarcation corresponding to maximal distance is taken as the separation between nearby phonemes. The segmentation is compared with that drawn by an “expert”, and it was often found to be remarkably accurate. The speech signal and phoneme segmentation results are shown in Figure 6.7.

The speech segment corresponds to the precise sentence “She has your dark suit in greasy wash water all year”. In the top sub-figure, diamonds mark phoneme boundaries as decided by a human “expert”. Asterisks, circles, squares and crosses indicate the boundaries between

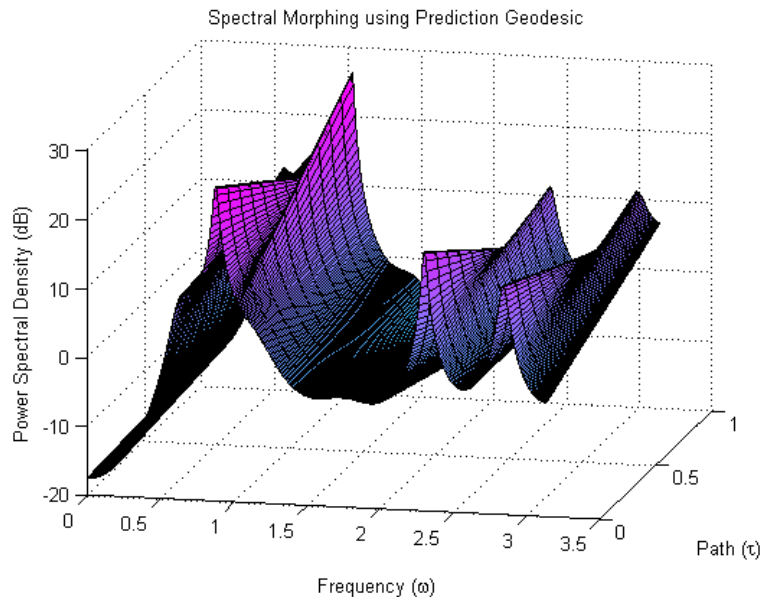


Figure 6.5: Geodesic path between the two spectra following prediction metric

phonemes that were obtained by comparing the prediction-based geodesic distance, L_2 -log distance, L_1 -log distance and the Itakura-Saito distance, respectively, between spectra of neighboring windows. Sub-figures 2, 3, and 4, are plots of the distance between such neighboring windows obtained by the respective metrics, as a function of the border between neighboring windows. Local maxima suggest boundaries between phonemes. Because the Itakura-Saito distance fluctuates considerably, we plot this on a logarithmic scale. Figure 6.8 shows more clearly the detail about a part of the sentence that corresponds to “she has”. We can see that the segmentation markers by all these methods are consistent with those drawn by the “expert”. The segmentation in this last figure separates phonemes corresponding to the sounds S , i , h , a , s , e .

6.3 Spectral Averaging

The idea of the transportation mean developed in Chapter 5 can be applied to the smoothing of spectral densities. For a slowly time-varying signal, the spectral densities are estimated by short time Fourier transformation (STFT). At each “frozen” time, we use the transportation mean of

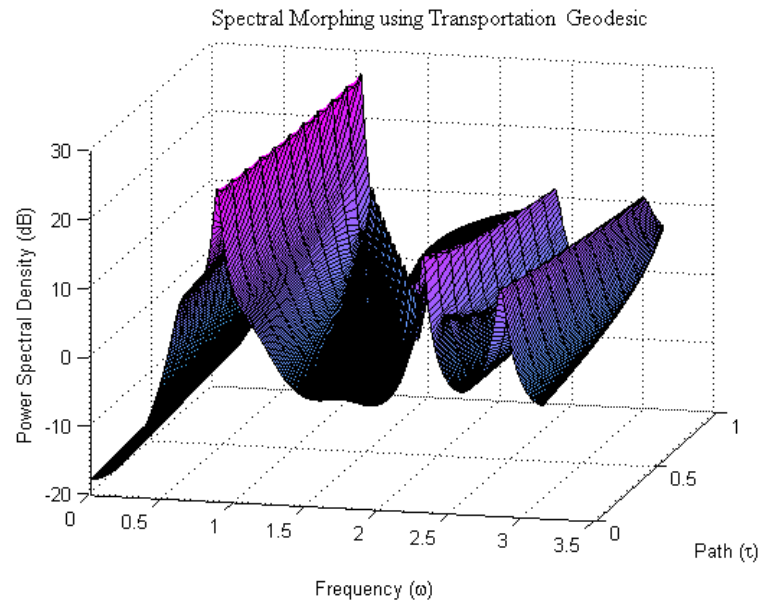


Figure 6.6: Geodesic path between the two spectra following transportation metric

its spectral density and its neighbors to represent the smoothed spectral estimation.

As an example, we generated a chirp signal corrupted by additive noise with variance 1. The instantaneous frequency at time t is given by $(10 + 290t)$ Hz with the sampling rate 1KHz. The STFT is computed over a window of 256 points with an overlap of 128 points between successive windows. At each time, transportation mean is calculated based on two neighbors from the past, the present, and two neighbors from the future with weights

$$[0.1250 \ 0.2188 \ 0.3125 \ 0.2188 \ 0.1250].$$

Table 6.1 compares the estimation performance by the STFT and the spectral averaging over 5000 simulations, where the instantaneous frequency is estimated as the peak location of the corresponding power spectral density. It is rather evident that the transportation-mean-based spectral averaging leads to much smaller variance.

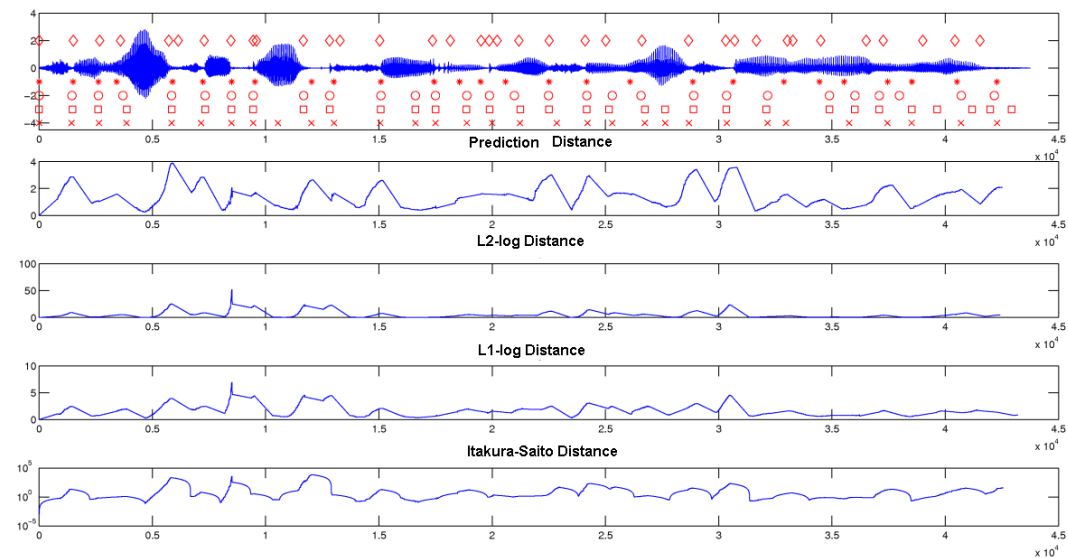


Figure 6.7: Speech sentence and phoneme segmentation. In the top sub-figure, diamonds designate segmentation by an “expert”. Asterisks, circles, squares and crosses correspond to segmentation based on prediction geodesic distance, L_2 -log distance, L_1 -log distance and the Itakura-Saito distance respectively.

6.4 The Gaussian Particle Filter

Let x be an n -dimensional Gaussian random variable with mean \bar{x} and variance P_x , and $y = h(x)$ be a nonlinear transformation of x . We are interested in how the mean and variance of x are propagated through the nonlinear filter h and how to estimate them.

To evaluate the mean and variance of y , we can expand y in a Taylor series around \bar{x} , that is

$$y = h(\bar{x}) + D_{\bar{x}}h + \frac{1}{2!}D_{\bar{x}}^2h + \frac{1}{3!}D_{\bar{x}}^3h + \dots,$$

where $\tilde{x} = x - \bar{x}$, and the operator $D_{\bar{x}}^k h$ is a compact format of

$$D_{\bar{x}}^k h = \left(\sum_{i=1}^n \tilde{x}_i \frac{\partial}{\partial x_i} \right)^k h.$$

As shown in [113, pp 436-439], the mean of y can be expanded as

$$\bar{y} = h(\bar{x}) + \frac{1}{2!}E(D_{\bar{x}}^2h) + \frac{1}{4!}E(D_{\bar{x}}^4h) + \dots, \quad (6.1)$$

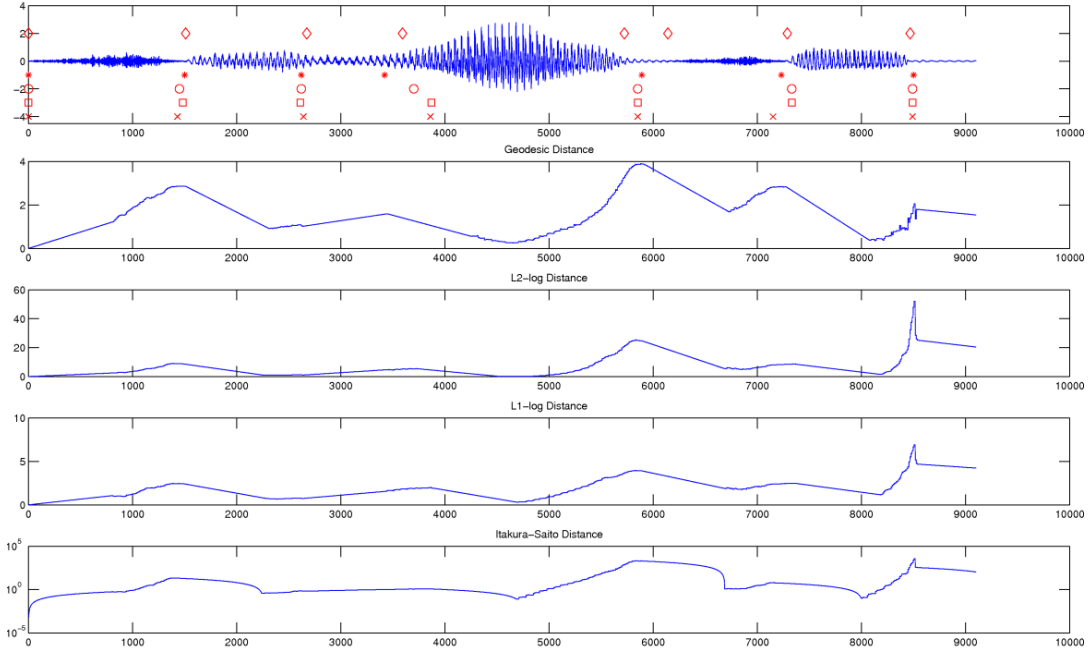


Figure 6.8: Part of the sentence and phoneme segmentation results. In the top sub-figure, diamonds indicate markings by an “expert”. Asterisks, circles, squares and crosses are marked based on prediction geodesic distance, L_2 -log distance, L_1 -log distance and the Itakura-Saito distance respectively.

and the variance of y as

$$P_y = HP_x H^T + E \left[\frac{D_{\bar{x}} h (D_{\bar{x}}^3 h)^T}{3!} + \frac{D_{\bar{x}}^2 h (D_{\bar{x}}^2 h)^T}{2!2!} + \frac{D_{\bar{x}}^3 h (D_{\bar{x}} h)^T}{3!} \right] + E \left(\frac{D_{\bar{x}}^2 h}{2!} \right) E \left(\frac{D_{\bar{x}}^2 h}{2!} \right)^T + \dots, \quad (6.2)$$

where

$$H = \frac{\partial h}{\partial x} \Big|_{x=\bar{x}}.$$

Time Instant (s)		0.384	0.512	0.640	0.7680
Instantaneous Frequency (Hz)		121.36	158.48	195.60	232.72
STFT	Mean	122.08	159.04	195.90	232.58
	Variance	282.24	331.60	359.48	183.87
Spectral Averaging	Mean	120.18	157.29	194.62	231.59
	Variance	27.56	29.05	30.80	29.70

Table 6.1: Instantaneous Frequency Estimation

6.4.1 Linearized transformation and unscented transformation

A simple way to approximate the mean and variance of y is through the linearization of $h(x)$ around \bar{x} , as adopted by the Extended Kalman Filter (EKF). Let \bar{y}_l and P_l denote the approximated mean and variance respectively, then

$$\bar{y}_l = h(\bar{x}) \quad (6.3)$$

$$P_l = HP_x H^T. \quad (6.4)$$

However this type of approximation does not take the higher order terms into consideration. A more sophisticated method is through the unscented transformation. The unscented transformation requires a set of deterministic sigma points $x_u^{(i)}$'s whose ensemble mean and covariance are equal to \bar{x} and P_x respectively, and then passes these points through the nonlinear function. The ensemble mean and covariance of the resulting $y_u^{(i)}$'s give the estimation of \bar{y} and P_y . To be more specific, the $2n + 1$ sigma points $x^{(i)}$ and their weights $w^{(i)}$ are commonly chosen as

$$\begin{aligned} x^{(0)} &= \bar{x}, & w^{(0)} &= \kappa / (n + \kappa), \\ x^{(i)} &= \bar{x} + (\sqrt{(n + \kappa)P_x})_i^T, & w^{(i)} &= 1/2(n + \kappa), \\ x^{(n+i)} &= \bar{x} - (\sqrt{(n + \kappa)P_x})_i^T, & w^{(n+i)} &= 1/2(n + \kappa), \end{aligned}$$

for $i = 1, \dots, n$, and $\kappa \in \mathbb{R}$, and $(\sqrt{P_x})_i^T$ being the transpose of the i -th row of $\sqrt{P_x}$. The transformed sigma points are calculated as

$$y_u^{(i)} = h(x_u^{(i)}),$$

and the mean and covariance of y are estimated as

$$\bar{y}_u = \sum_{i=0}^{2n} w^{(i)} y_u^{(i)} \quad (6.5)$$

$$P_u = \sum_{i=1}^{2n} w^{(i)} (y_u^{(i)} - \bar{y}_u)(y_u^{(i)} - \bar{y}_u)^T. \quad (6.6)$$

It is shown in [114] that \bar{y}_u and P_u match \bar{y} and P_y up to the third order of Taylor series, and the sign of the fourth order of P_u matches that of P_y .

6.4.2 The Gaussian particle transformation

The unscented transformation uses pulses at sigma points to approximate a Gaussian distribution. To take one step further, we propose to approximate it with several Gaussian particles $x_G^{(i)}$. Each of them is a Gaussian distribution with mean $\bar{x}_G^{(i)}$ and variance $P_{xG}^{(i)}$, such that their transportation mean is the original distribution of x . After passing the nonlinear filter $h(x)$, we have

$$y_G^{(i)} = h(x_G^{(i)}).$$

Linearizing $h(x_G^{(i)})$ around $\bar{x}_G^{(i)}$, the resulting $y_G^{(i)}$'s are approximated as Gaussian distributions with means and covariances being

$$\bar{y}_G^{(i)} = h(\bar{x}_G^{(i)}) \quad (6.7)$$

$$P_G^{(i)} = H_G P_{xG}^{(i)} H_G^T, \quad (6.8)$$

where

$$H_G = \left. \frac{\partial h}{\partial x} \right|_{x=\bar{x}_G^{(i)}}.$$

We shall call the transportation mean of $N(\bar{y}_G^{(i)}, P_G^{(i)})$'s as the Gaussian particle (GP) estimation of y . The detailed procedure is as follows. First, choose Gaussian particles as $x_G^{(i)} \sim N(\bar{x}_G^{(i)}, P_{xG}^{(i)})$, such that

$$\bar{x} = \frac{1}{2n} \sum_{i=1}^{2n} \bar{x}_G^{(i)},$$

and

$$P_x = \frac{1}{2n} \sum_{i=1}^{2n} (P_x^{\frac{1}{2}} P_{xG}^{(i)} P_x^{\frac{1}{2}})^{\frac{1}{2}}.$$

A practical choice could be

$$\begin{aligned}\bar{x}_G^{(i)} &= \bar{x} + \varepsilon(\sqrt{nP_x})_i^T, \\ \bar{x}_G^{(n+i)} &= \bar{x} - \varepsilon(\sqrt{nP_x})_i^T, \quad i = 1, \dots, n,\end{aligned}$$

and

$$P_{xG}^{(i)} = P_x.$$

Then compute the estimated mean and covariance of $y_G^{(i)}$ as in (6.7). The GP estimation of y is their transportation mean, which is calculated as

$$\begin{aligned}\bar{y}_G &= \frac{1}{2n} \sum_{i=1}^{2n} y_G^{(i)}, \\ P_G &= \frac{1}{2n} \sum_{i=1}^{2n} (P_G^{\frac{1}{2}} P_G^{(i)} P_G^{\frac{1}{2}})^{\frac{1}{2}}.\end{aligned}$$

Proposition 33. *For the linear system $y = Ax$ where $x, y \in \mathbb{R}^n$ and $x \sim N(0, P)$, the LT, UT, and GPT give the same estimation on the mean and covariance of y .*

Proof. Since the system is linear, the estimation of the LT method is

$$\begin{aligned}\bar{y}_l &= A\bar{x} = 0, \\ P_l &= APA^T.\end{aligned}$$

For the unscented transform, the sigma points are selected as:

$$\begin{aligned}x_u^{(i)} &= (\sqrt{nP_x})_i^T, \\ x_u^{(n+i)} &= -(\sqrt{nP_x})_i^T, \quad i = 1, \dots, n,\end{aligned}$$

where $(\sqrt{nP_x})_i^T$ is again the transpose of the i -th row of $\sqrt{nP_x}$. The transformed sigma points are

$$y_u^{(i)} = Ax_u^{(i)}.$$

Thus the estimated mean and covariance of y are calculated as

$$\begin{aligned}\bar{y}_u &= \frac{1}{2n} \sum_{i=1}^{2n} y_u^{(i)} = 0 \\ P_u &= \frac{1}{2n} \sum_{i=1}^{2n} (Ax_u^{(i)})(Ax_u^{(i)})^T = APA^T.\end{aligned}$$

For the GPT, we choose the mean and covariance of Gaussian particles $x_G^{(i)}$ as

$$\begin{aligned}\bar{x}_G^{(i)} &= \varepsilon(\sqrt{nP_x})_i^T, \\ \bar{x}_G^{(n+i)} &= -\varepsilon(\sqrt{nP_x})_i^T, \quad i = 1, \dots, n,\end{aligned}$$

and

$$P_{xG}^{(i)} = P.$$

Then compute the estimated mean and covariance of $y_G^{(i)}$ as

$$\bar{y}_G^{(i)} = A\bar{x}_G^{(i)} \quad (6.9)$$

$$P_G^{(i)} = APA^T, \quad (6.10)$$

The GP estimation of y is their transportation mean, which is calculated as

$$\begin{aligned}\bar{y}_G &= \frac{1}{2n} \sum_{i=1}^{2n} \bar{y}_G^{(i)} = 0, \\ P_G &= \left(\frac{1}{2n} \sum_{i=1}^{2n} \sqrt{P_G^{(i)}} \right)^2 = APA^T.\end{aligned}$$

Thus for the linear system, the LT, UT, and GPT yield the same estimation. \square

6.4.3 Examples

We shall use a simple nonlinear function $y = x^q$ to compare the estimations of linearized transformation (LT), unscented transformation (UT), and Gaussian particle transformation (GPT). Let $x \sim N(0, \sigma^2)$ and q be a nonzero natural number. First, we shall provide the theoretic mean and variance of y :

$$\bar{y} = E(x^q) = \begin{cases} 0, & q \text{ odd} \\ c_1 \sigma^q, & q \text{ even} \end{cases} \quad (6.11)$$

$$P_y = E(x^{2q}) - \bar{y}^2 = \begin{cases} c_2 \sigma^{2q}, & q \text{ odd} \\ (c_2 - c_1^2) \sigma^{2q}, & q \text{ even} \end{cases}, \quad (6.12)$$

where

$$c_1 = (q-1)!!,$$

and

$$c_2 = (2q-1)!!,$$

with $n!!$ denoting the double factorial.

Using the LT, we obtain

$$y \simeq \bar{x}^q + q\bar{x}^{q-1}(x - \bar{x}),$$

therefore

$$\bar{y}_l = 0 \tag{6.13}$$

$$P_l = \begin{cases} \sigma^2, & q = 1 \\ 0, & q > 1. \end{cases} \tag{6.14}$$

On the other hand, for the UT, we choose sigma points and their corresponding weights as

$$\begin{aligned} x_u^{(0)} &= 0, & w_u^{(0)} &= \kappa/(1 + \kappa) \\ x_u^{(1)} &= \sqrt{(1 + \kappa)}\sigma, & w_u^{(1)} &= 1/2(1 + \kappa) \\ x_u^{(2)} &= -\sqrt{(1 + \kappa)}\sigma, & w_u^{(2)} &= 1/2(1 + \kappa) \end{aligned}$$

The resulting

$$\begin{aligned} y_u^{(0)} &= 0, \\ y_u^{(1)} &= (1 + \kappa)^{q/2} \sigma^q, \\ y_u^{(2)} &= (1 + \kappa)^{q/2} (-\sigma)^q. \end{aligned}$$

Therefore,

$$\bar{y}_u = \sum_{i=0}^2 w_u^{(i)} y_u^{(i)} = \begin{cases} 0, & q \text{ odd} \\ (1 + \kappa)^{(q-2)/2} \sigma^q, & q \text{ even} \end{cases} \tag{6.15}$$

$$P_u = \sum_{i=0}^2 w_u^{(i)} (y_u^{(i)} - \bar{y}_u)^2 = \begin{cases} (1 + \kappa)^{q-1} \sigma^{2q}, & q \text{ odd} \\ \kappa(1 + \kappa)^{q-2} \sigma^{2q}, & q \text{ even} \end{cases} \tag{6.16}$$

When q is odd, we can choose κ to satisfy

$$(1 + \kappa)^{q-1} = (c_2 - c_1^2).$$

In the case that q is even, κ can be chosen as

$$(1 + \kappa)^{(q-2)/2} = c_1$$

so that the estimation is unbiased. The corresponding variance becomes $c_1^2(c_1^{2/(q-2)} - 1)\sigma^{2q}$.

For the GPT, the Gaussian particles are chosen to be $\bar{x}_G^{(1)} = \varepsilon\sigma$, $\bar{x}_G^{(2)} = -\varepsilon\sigma$, and $P_{x_G}^{(1)} = P_{x_G}^{(2)} = \sigma^2$. The linearized filters for these Gaussian particles are:

$$y_G^{(i)} \simeq (\bar{x}_G^{(i)})^q + q(\bar{x}_G^{(i)})^{q-1}(x_G^{(i)} - \bar{x}_G^{(i)}), \quad i = 1, 2.$$

Consequently,

$$\begin{aligned} \bar{y}_G^{(1)} &= (\varepsilon\sigma)^q, \\ \bar{y}_G^{(2)} &= (-\varepsilon\sigma)^q, \end{aligned}$$

and

$$P_G^{(1)} = P_G^{(2)} = q^2 \varepsilon^{2(q-1)} \sigma^{2q}.$$

Thus we have

$$\bar{y}_G = \frac{1}{2}(\bar{y}_G^{(1)} + \bar{y}_G^{(2)}) = \begin{cases} 0, & q \text{ odd} \\ \varepsilon^q \sigma^q, & q \text{ even} \end{cases} \quad (6.17)$$

$$P_G = \left(\frac{1}{2}(\sqrt{P_G^{(1)}} + \sqrt{P_G^{(2)}}) \right)^2 = q^2 \varepsilon^{2(q-1)} \sigma^{2q}. \quad (6.18)$$

When q is odd, we can choose $q^2 \varepsilon^{2(q-1)} = c_2$ so that both mean and variance match the theoretic values. When q is even, we can choose $\varepsilon^q = c_1$ so that the estimation is unbiased. In this case, the estimated variance becomes $q^2 (\sqrt[q]{c_1})^{2q-2}$.

The following table shows the comparison among the above three methods.

		theoretic	LT	UT	GPT
x^2	mean	σ^2	0	σ^2	σ^2
	variance	$2\sigma^4$	0	$2\sigma^4$	$4\sigma^4$
x^3	mean	0	0	0	0
	variance	$15\sigma^6$	0	$15\sigma^6$	$15\sigma^6$
x^4	mean	$3\sigma^4$	0	$3\sigma^4$	$3\sigma^4$
	variance	$96\sigma^8$	0	$18\sigma^8$	$83.1384\sigma^8$
x^5	mean	0	0	0	0
	variance	$945\sigma^{10}$	0	$945\sigma^{10}$	$945\sigma^{10}$
x^6	mean	$15\sigma^6$	0	$15\sigma^6$	$15\sigma^6$
	variance	$10170\sigma^{12}$	0	$646.4213\sigma^{12}$	$3284.4\sigma^{12}$

Follow the same procedure, we can consider the nonlinear system to be

$$y = x + x^q.$$

The following table shows the comparison among the above three methods.

		theoretic	LT	UT	GPT
$x + x^2$	mean	σ^2	0	σ^2	σ^2
	variance	$\sigma^2 + 2\sigma^4$	σ^2	$\sigma^2 + 2\sigma^4$	$\begin{cases} \sigma^2, \sigma < 0.5; \\ 4\sigma^4, \sigma > 0.5. \end{cases}$
$x + x^3$	mean	0	0	0	0
	variance	$\sigma^2 + 6\sigma^4 + 15\sigma^6$	σ^2	$\sigma^2 + 6\sigma^4 + 9\sigma^6$	$\sigma^2 + 6\sigma^4 + 9\sigma^6$
$x + x^4$	mean	$3\sigma^4$	0	$3\sigma^4$	$3\sigma^4$
	variance	$\sigma^2 + 96\sigma^8$	0	$\sigma^2 + 18\sigma^8$	$\begin{cases} \sigma^2, \sigma < 0.4787; \\ 83.1384\sigma^8, \sigma > 0.4787. \end{cases}$
$x + x^5$	mean	0	0	0	0
	variance	$\sigma^2 + 30\sigma^6 + 945\sigma^{10}$	σ^2	$\sigma^2 + 30\sigma^6 + 225\sigma^{10}$	$\sigma^2 + 30\sigma^6 + 225\sigma^{10}$
$x + x^6$	mean	$15\sigma^6$	0	$15\sigma^6$	$15\sigma^6$
	variance	$\sigma^2 + 10170\sigma^{12}$	0	$\sigma^2 + 646.4213\sigma^{12}$	$\begin{cases} \sigma^2, \sigma < 0.445; \\ 3284.4\sigma^{12}, \sigma > 0.445. \end{cases}$

6.4.4 Extensions

Compared with EKF and UKF, the Gaussian particle filter is computationally more demanding. Further work is needed before this becomes a liable alternative by substantially reducing the computational cost. Another interesting alternative is to consider an area/volumn filter. More specifically, let $y = h(x)$ be a nonlinear system where x is Gaussian distributed. To estimate the distribution of the output y , we select several points $x^{(i)}$'s on the σ -curve of the input Gaussian distribution. We postulate that the output of those points $y^{(i)}$'s are still on the σ -curve of the output distribution, i.e., the mass is conserved. We then estimate the mean and variance of the output distribution through minimizing the volume or area of the ellipse so that those points stay close to the ellipsoidal.

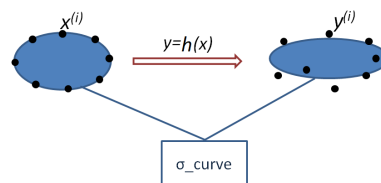


Figure 6.9: Illustration of the area/volumn filter

References

- [1] P. Stoica and R. Moses. *Spectral Analysis of Signals*. Prentice Hall, 2005.
- [2] R. A. Silverman. Locally stationary random processes. *Information Theory, IRE Transactions on*, 3:182–187, 1957.
- [3] W. Martin. Measuring the degree of non-stationarity by using the Wigner-Ville spectrum. In *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal (ICASSP'84)*, volume 9, pages 262–265, 1984.
- [4] J. Xiao, P. Borgnat, and P. Flandrin. Testing stationarity with time-frequency surrogates. In *Proceedings of the European Signal Processing Conference (EUSIPCO'07)*, volume 7, pages 2020–2024, 2007.
- [5] J. Xiao, P. Borgnat, P. Flandrin, and C. Richard. Testing stationarity with surrogates—a one-class SVM approach. In *Proceedings of the IEEE 14th Workshop on Statistical Signal Processing (SSP'07)*, pages 720–724, 2007.
- [6] M. B. Priestley and T. Subba Rao. A test for non-stationarity of time-series. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 31(1):140–149, 1969.
- [7] S. Kay. A new nonstationarity detector. *Signal Processing, IEEE Transactions on*, 56(4):1440–1451, 2008.
- [8] T. S. Rao. The fitting of non-stationary time-series models with time-dependent parameters. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 32(2):312–322, 1970.

- [9] Y. Grenier. Time-dependent ARMA modeling of nonstationary signals. *Audio, Speech and Language Processing, IEEE Transactions on*, 31(4):899–911, 1983.
- [10] R. Dahlhaus. Fitting time series models to nonstationary processes. *The Annals of Statistics*, 25(1):1–37, 1997.
- [11] A. Kaderli and A. S. Kayhan. Spectral estimation of nonstationary ARMA processes using the evolutionary cepstrum. *IEEE Signal Processing Letters*, 9(4):130–132, 2002.
- [12] I. Daubechies. The wavelet transform, time-frequency localization and signal analysis. *Information Theory, IEEE Transactions on*, 36(5):961–1005, 1990.
- [13] S. Haykin. *Advances in spectrum analysis and array processing (vol. I)*. Prentice Hall, 1991.
- [14] L. Cohen. *Time-Frequency Analysis*. Prentice Hall, 1995.
- [15] C. H. Page. Instantaneous power spectra. *Journal of Applied Physics*, 23(1):103–106, 1952.
- [16] M. B. Priestley. Evolutionary spectra and non-stationary processes. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 27(2):204–237, 1965.
- [17] T. T. Georgiou. Distances and Riemannian metrics for spectral density functions. *Signal Processing, IEEE Transactions on*, 55(8):3995–4003, 2007.
- [18] M. Basseville. Distance measures for signal processing and pattern recognition. *Signal processing*, 18(4):349–369, 1989.
- [19] R. Gray, A. Buzo, A. Gray Jr, and Y. Matsuyama. Distortion measures for speech processing. *Acoustics, Speech and Signal Processing, IEEE Transactions on*, 28(4):367–376, 1980.
- [20] T. T. Georgiou, J. Karlsson, and M. S. Takyar. Metrics for power spectra: an axiomatic approach. *Signal Processing, IEEE Transactions on*, 57(3):859–867, 2009.
- [21] L. L. Campbell. The relation between information theory and the differential geometry approach to statistics. *Information Sciences*, 35(3):199–210, 1985.

- [22] N. N. Cencov. *Statistical decision rules and optimal inference*. Number 53. American Mathematical Society, 1982.
- [23] H. Cramér. *Mathematical methods of statistics*. Princeton University, 1946.
- [24] K. Matusita. Decision rule, based on the distance, for the classification problem. *Annals of the Institute of Statistical Mathematics*, 8(1):67–77, 1956.
- [25] F. Itakura and S. Saito. Analysis synthesis telephony based on the maximum likelihood method. *Proceedings of the 6th International Congress on Acoustics*, pages C17–C20, 1968.
- [26] F. Itakura. Minimum prediction residual principle applied to speech recognition. *Acoustics, Speech and Signal Processing, IEEE Transactions on*, 23(1):67–72, 1975.
- [27] C. Villani. *Topics in optimal transportation*. American Mathematical Society, 2003.
- [28] S. Haker, L. Zhu, A. Tannenbaum, and S. Angenent. Optimal mass transport for registration and warping. *International Journal of Computer Vision*, 60(3):225–240, 2004.
- [29] J. von Neumann. *Mathematische Grundlagen der Quantenmechanik*. Verlag von Julius Springer, 1932.
- [30] D. Petz. Entropy, von Neumann and the von Neumann entropy. *John von Neumann and the Foundations of Quantum Physics*, pages 83–92, 2001.
- [31] H. Umegaki. Conditional expectation in an operator algebra. IV. entropy and information. *Kodai Mathematical Journal*, 14(2):59–85, 1962.
- [32] F. Hiai and D. Petz. The proper formula for relative entropy and its asymptotics in quantum probability. *Communications in mathematical physics*, 143(1):99–114, 1991.
- [33] V. Vedral. The role of relative entropy in quantum information theory. *Reviews of Modern Physics*, 74(1):197–234, 2002.
- [34] R. S. Ingarden, H. Janyszek, A. Kossakowski, and T. Kawaguchi. Information geometry of quantum statistical systems. *Tensor, NS*, 37:105–111, 1982.
- [35] R. Bhatia. *Positive definite matrices*. Princeton University Press, 2007.

- [36] D. Petz. Geometry of canonical correlation on the state space of a quantum system. *Journal of Mathematical Physics*, 35:780–795, 1994.
- [37] A. Uhlmann. The metric of Bures and the geometric phase. *Quantum Groups and Related Topics*, page 267, 1992.
- [38] R. Bhatia and P. Rosenthal. How and why to solve the operator equation $AX-XB=Y$. *Bulletin of the London Mathematical Society*, 29(1):1–21, 1997.
- [39] M. Hübner. Explicit computation of the Bures distance for density matrices. *Physics Letters A*, 163(4):239 – 242, 1992.
- [40] A. Ferrante, M. Pavon, and F. Ramponi. Hellinger versus Kullback–Leibler multivariable spectrum approximation. *Automatic Control, IEEE Transactions on*, 53(4):954–967, 2008.
- [41] S. Luo and Q. Zhang. Informational distance on quantum-state space. *Physical Review A*, 69:032106, 2004.
- [42] R. Bhatia and J. A. R. Holbrook. Riemannian geometry and matrix geometric means. *Linear Algebra and its Applications*, 413:594–618, 2006.
- [43] T. M. Cover and J. A. Thomas. *Elements of information theory*. John Wiley, 1991.
- [44] U. Grenander and G. Szegö. *Toeplitz forms and their applications*. University of California Press, 1958.
- [45] R. M. Gray. *Toeplitz and circulant matrices: A review*. Now Publishers Inc., 2006.
- [46] A. Ferrante, C. Masiero, and M. Pavon. Time and spectral domain relative entropy: A new approach to multivariate spectral estimation. *Arxiv preprint arXiv:1103.5602*, 2011.
- [47] M. S. Pinsker. Information and information stability of random variables and processes. 1960. English translation: San Francisco, CA: Holden-Day, 1964.
- [48] A. Stoorvogel and J. van Schuppen. System identification with information theoretic criteria. *Identification, Adaptation, Learning*, pages 289–338, 1996.

- [49] S-I. Amari. Differential-geometrical methods in statistics. *Lecture Notes in Statistics, Berlin, Springer Verlag*, 28, 1985.
- [50] B. Musicus and R. Johnson. Multichannel relative-entropy spectrum analysis. *Acoustics, Speech and Signal Processing, IEEE Transactions on*, 34(3):554–564, 1986.
- [51] P. Enqvist and J. Karlsson. Minimal Itakura-Saito distance and covariance interpolation. In *47th IEEE Conference on Decision and Control*, pages 137–142, 2008.
- [52] T. T. Georgiou. Relative entropy and the multivariable multidimensional moment problem. *Information Theory, IEEE Transactions on*, 52(3):1052–1066, 2006.
- [53] X. Jiang, L. Ning, and T. T. Georgiou. Distances and Riemannian metrics for multivariate spectral densities. accepted, to appear in *Automatic Control, IEEE Transactions on*.
- [54] N. Wiener and P. Masani. The prediction theory of multivariate stochastic processes, I. *Acta Mathematica*, 98:111–150, 1957.
- [55] P. Masani. *Recent trends in multivariable prediction theory. (Krishnaiah, P.R., Editor), Multivariate Analysis*, pp. 351-382. Academic Press, 1966.
- [56] T. T. Georgiou. The Carathéodory–Fejér–Pisarenko decomposition and its multivariable counterpart. *Automatic Control, IEEE Transactions on*, 52(2):212–228, 2007.
- [57] P. Petersen. *Riemannian geometry*. Springer Verlag, 2006.
- [58] J. Jost. *Nonpositive curvature: geometric and analytic aspects*. Birkhäuser, 1997.
- [59] M. Moakher. A differential geometric approach to the geometric mean of symmetric positive-definite matrices. *SIAM Journal on Matrix Analysis and Applications*, 26(3):735–747, 2005.
- [60] A. N. Kolmogorov and S. V. Fomin. *Elements of the theory of functions and functional analysis, Volume 2*. Graylock Press, 1961.
- [61] R. A. Horn and C. R. Johnson. *Topics in matrix analysis*. Cambridge University press, 1994.
- [62] R. A. Horn and C. R. Johnson. *Matrix analysis*. Cambridge University press, 2005.

- [63] S. Rachev and L. Rüschendorf. *Mass transportation problems. Vol. I: theory*. Springer-Verlag, New York, 1998.
- [64] X. Jiang, Z.-Q. Luo, and T. T. Georgiou. Geometric methods for spectral analysis. accepted, to appear in *Signal Processing*, *IEEE Transactions on*.
- [65] X. Jiang, Z.-Q. Luo, and T. T. Georgiou. Power spectral geodesics and tracking. In *Proceedings of the IEEE 47th Conference on Decision and Control (CDC'08)*, pages 1315–1319, 2008.
- [66] S. Boyd and L. Vandenberghe. *Convex optimization*. Cambridge University Press, 2004.
- [67] R. O. Duda, P. E. Hart, and D. G. Stork. *Pattern classification*. Wiley, New York, 2001.
- [68] F. Attivissimo, M. Savino, and A. Trotta. Power spectral density estimation via overlapping nonlinear averaging. *Instrumentation and Measurement, IEEE Transactions on*, 50(5):1418–1424, 2001.
- [69] I. Pitas and A. Venetsanopoulos. Nonlinear mean filters in image processing. *Acoustics, Speech and Signal Processing, IEEE Transactions on*, 34(3):573 – 584, 1986.
- [70] O. Tuzel, F. Porikli, and P. Meer. Human detection via classification on Riemannian manifolds. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 1–8, 2007.
- [71] A. Cherian, S. Sra, A. Banerjee, and N. Papanikolopoulos. Efficient similarity search for covariance matrices via the Jensen-Bregman logdet divergence. In *International Conference on Computer Vision (ICCV)*, 2011.
- [72] P. D. Welch. The use of fast Fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms. *Audio and Electroacoustics, IEEE Transactions on*, 15:70–73, 1967.
- [73] P. D. Welch. On the relationship between batch means, overlapping means and spectral estimation. In *Proceedings of the 19th Conference on Winter simulation (WSC'87)*, pages 320–323, New York, NY, USA, 1987. ACM.

- [74] M. S. Bartlett. Periodogram analysis and continuous spectra. *Biometrika*, 37(1/2):1–16, 1950.
- [75] P. J. Brockwell and R. A. Davis. *Time Series: Theory and Methods*. Springer-Verlag, New York, 2 edition, 1991.
- [76] U. Grenander and M. Rosenblatt. *Statistical Analysis of Stationary Time Series*. Chelsea Publishing Company, New York, 2 edition, 1984.
- [77] D. R. Brillinger. *Time Series: Data Analysis and Theory*. SIAM, 2001.
- [78] G. E. P. Box and D. R. Cox. An analysis of transformations. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 26(2):211–252, 1964.
- [79] M. B. Priestley and T. Subba Rao. A test for non-stationarity of time-series. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 31(1):140–149, 1969.
- [80] A. C. Jr. Cohen and P. J. Burke. Compound normal distribution. *American Mathematical Monthly*, 63(2):129, 1956.
- [81] M. Agueh and G. Carlier. Barycenters in the Wasserstein space. *SIAM Journal on Mathematical Analysis*, 43:904–924, 2011.
- [82] L. Rüschendorf and L. Uckelmann. On the n-coupling problem. *Journal of Multivariate Analysis*, 81(2):242–258, 2002.
- [83] M. Knott and C. S. Smith. On the optimal mapping of distributions. *Journal of Optimization Theory and Applications*, 43(1):39–49, 1984.
- [84] C. Dellacherie and P. A. Meyer. *Probabilities and potential*, volume 29. Elsevier, 1978.
- [85] W. Gangbo and A. Święch. Optimal maps for the multidimensional Monge-Kantorovich problem. *Communications on Pure and Applied Mathematics*, 51(1):23–45, 1998.
- [86] I. Olkin and F. Pukelsheim. The distance between two random vectors with given dispersion matrices. *Linear Algebra and its Applications*, 48:257–263, 1982.
- [87] I. Olkin and S. T. Rachev. Maximum submatrix traces for positive definite matrices. *SIAM Journal on Matrix Analysis and Applications*, 14:390–397, 1993.

- [88] M. Knott and C. S. Smith. On a generalization of cyclic monotonicity and distances among random vectors. *Linear Algebra and its Applications*, 199:363–371, 1994.
- [89] R. Chandrasekaran and A. Tamir. Open questions concerning Weiszfeld’s algorithm for the Fermat-Weber location problem. *Mathematical Programming*, 44(1):293–295, 1989.
- [90] I. Norman Katz. Local convergence in Fermat’s problem. *Mathematical Programming*, 6(1):89–104, 1974.
- [91] H. W. Kuhn. A note on Fermat’s problem. *Mathematical Programming*, 4(1):98–107, 1973.
- [92] J. Nie, P. A. Parrilo, and B. Sturmfels. Semidefinite representation of the k-ellipse. *IMA Journal of Applied Mathematics*, 146:117–132, 2008.
- [93] T. Ando, C. K. Li, and R. Mathias. Geometric means. *Linear Algebra and its Applications*, 385:305–334, 2004.
- [94] D. Petz and R. Temesi. Means of positive numbers and matrices. *SIAM Journal on Matrix Analysis and Applications*, 27(3):712–720, 2006.
- [95] V. Arsigny, P. Fillard, X. Pennec, and N. Ayache. Geometric means in a novel vector space structure on symmetric positive-definite matrices. *SIAM Journal on Matrix Analysis and Applications*, 29(1):328–347, 2007.
- [96] C. Qiu and N. Vaswani. ReProCS: A missing link between recursive robust PCA and recursive sparse recovery in large but correlated noise. *Arxiv preprint arXiv:1106.3286*, 2011.
- [97] H. Wang, A. Banerjee, and D. Boley. Common component analysis for multiple covariance matrices. Technical report, TR-10-017, University of Minnesota, Twin Cities, 2010.
- [98] D. G. Childers. *Speech processing and synthesis toolboxes*. Wiley, 1999.
- [99] H. R. Pfitzinger. Unsupervised speech morphing between utterances of any speakers. In *Proceedings of the 10th Australian International Conference on Speech Science and Technology (SST’04)*, pages 545–550, 2004.

- [100] M. Abe. Speech morphing by gradually changing spectrum parameter and fundamental frequency. In *Proceedings of the 4th International Conference on Spoken Language Processing (ICSLP'96)*, pages 2235–2238, 1996.
- [101] H. Ye and S. Young. Perceptually weighted linear transformations for voice conversion. In *Proceedings of the 8th European Conference on Speech Communication and Technology (EUROSPEECH'03)*, 2003.
- [102] H. Kawahara and H. Matsui. Auditory morphing based on an elastic perceptual distance metric in an interference-free time-frequency representation. In *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP'03)*, volume 1, pages 256–259, 2003.
- [103] V. Goncharoff and M. Kaine-Krolak. Interpolation of LPC spectra via pole shifting. In *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP'95)*, volume 1, pages 780–783, 1995.
- [104] T. T. Georgiou. The meaning of distances in spectral analysis. Plenary lecture, IEEE Conference on Decision and Control, 2007. <http://www.ieeecss-011.org/video/meaning-distances-spectral-analysis>.
- [105] D. Y. Wong, J. D. Markel, and A. H. Gray. Least squares glottal inverse filtering from the acoustic speech waveform. *Audio, Speech, and Language Processing, IEEE Transactions on*, 27(4):350–355, 1979.
- [106] M. R. Sambur, A. E. Rosenberg, L. R. Rabiner, and C. A. McGonegal. On reducing the buzz in LPC synthesis. *Journal of the Acoustical Society of America*, 63(3):918–924, 1978.
- [107] A. M. A. Ali, J. Van der Spiegel, P. Mueller, G. Haentjens, and J. Berman. An acoustic-phonetic feature-based system for automatic phoneme recognition in continuous speech. In *Proceedings of the IEEE International Symposium on Circuits and Systems (ISCAS'99)*, volume 3, pages 118–121, 1999.
- [108] D. B. Grayden and M. S. Scordilis. Phonemic segmentation of fluent speech. In *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP'94)*, volume 1, pages 73–76, 1994.

- [109] B. Ziółko, S. Manandhar, and R. C. Wilson. Phoneme segmentation of speech. In *Proceedings of the 18th International Conference on Pattern Recognition (ICPR'06)*, volume 4, pages 282–285, 2006.
- [110] C. J. Weinstein, S. S. McCandless, L. F. Mondschein, and V. W. Zue. A system for acoustic-phonetic analysis of continuous speech. *Audio, Speech, and Language Processing, IEEE Transactions on*, 23(1):54–67, 1975.
- [111] K. Demuynck and T. Laureys. A comparison of different approaches to automatic speech segmentation. In *Proceedings of the 11th International Conference on Text, Speech and Dialogue (TSD'08)*, pages 385–406. Springer, 2006.
- [112] J. S. Garofolo, L. F. Lamel, W. M. Fisher, J. G. Fiscus, D. S. Pallett, and N. L. Dahlgren. DARPA TIMIT acoustic-phonetic continuous speech corpus CD-ROM. *NTIS order number PB91-100354*, 1993.
- [113] D. Simon. *Optimal State Estimation*. John Wiley & Sons, 2006.
- [114] S. J. Julier and J. K. Uhlmann. A new extension of the Kalman filter to nonlinear systems. In *International Symposium on Aerospace /Defense Sensing, Simulation, and Controls*, volume 3, page 26, 1997.
- [115] S. Kobayashi and K. Nomizu. *Foundations of differential geometry I, II*. Interscience, 1963.
- [116] G. Corach and A. L. Maestripieri. Differential and metrical structure of positive operators. *Positivity*, 3(4):297–315, 1999.
- [117] I. Bengtsson and K. Życzkowski. *Geometry of quantum states: an introduction to quantum entanglement*. Cambridge University Press, 2006.
- [118] M. P. Do Carmo. *Riemannian geometry*. Birkhauser, 1992.
- [119] Mathematica program to compute Christoffel symbols and geodesic equations. <http://www.physics.ucsb.edu/~gravitybook/math/christoffel.pdf>.
- [120] B. Mendelson. *Introduction to topology*. Allyn and Bacon, 1962.

- [121] R. L. Wheeden and A. Zygmund. *Measure and integral: an introduction to real analysis*. Dekker, New York, 1977.
- [122] R. M. Dudley. *Real analysis and probability*. Cambridge University Press, 2002.
- [123] S. Kullback. *Information theory and statistics*. Dover Publishers, 1997.
- [124] B. V. Gnedenko. *Theory of probability*. CRC, 1997.
- [125] W. K. Wootters. Statistical distance and Hilbert space. *Physical Review D*, 23(2):357–362, 1981.

Appendix A

Review of Differential Geometry and Fibre Bundles

This appendix provides some background on fibre bundles and connections in the differential geometry in order to appreciate the Bures metric. We refer to the book by Kobayashi and Nomizu [115] for more details on this subject.

Let \mathbb{M}_n denote the set of $n \times n$ matrices, and let \mathbb{H}_n be the set of Hermitian matrices. The set of skew Hermitian matrices \mathfrak{u}_n is identified with the Lie algebra of the unitary group \mathbb{U}_n . Let GL_n be the general linear group, i.e.,

$$GL_n = \{X \in \mathbb{M}_n : \det(X) \neq 0\},$$

and \mathbb{P}_n be the set of positive definite matrices

$$\mathbb{P}_n = \{X \in GL_n \cap \mathbb{H}_n : X > 0\}.$$

There is an action of GL_n acting on \mathbb{P}_n

$$GL_n \times \mathbb{P}_n \rightarrow \mathbb{P}_n : (g, g_+) \rightarrow gg_+g^*.$$

This action is transitive since for any $x_1, x_2 \in \mathbb{P}_n$, there exists $g = x_2^{1/2}x_1^{-1/2}$ which maps x_1 to x_2 . Thus \mathbb{P}_n is a homogeneous space of GL_n .

In addition, every $g \in GL_n$ has a polar decomposition $g = g_+u$ where $g_+ \in \mathbb{P}_n$ and $u \in \mathbb{U}_n$. The projection π

$$\pi : GL_n \rightarrow \mathbb{P}_n, \pi(g) = g_+ (= \sqrt{gg^*}).$$

maps GL_n into \mathbb{P}_n , where we have used \sqrt{X} to denote the positive square root of matrix X . Thus we have a trivial fibre bundle with \mathbb{P}_n being the base manifold, GL_n the total space and \mathbb{U}_n the structure group. The fibre through $g_+ \in \mathbb{P}_n$ is the set $\{g_+u : u \in \mathbb{U}_n\}$. The tangent space $(TGL_n)_g$ of GL_n at g is identified with \mathbb{M}_n . Likewise, the tangent space $(T\mathbb{P}_n)_{g_+}$ is \mathbb{H}_n , and the tangent space $(T\mathbb{U}_n)_u$ is identified with $u\mathfrak{u}_n$. For any tangent vector \dot{g} at $g \in GL_n$, we can express it as a mapping from $(T\mathbb{P}_n)_{g_+}$ and $(T\mathbb{U}_n)_u$

$$\dot{g} = (\dot{g_+}u) = \dot{g_+}u + g_+\dot{u}.$$

The following lemma establishes the existence and the uniqueness of this map.

Lemma 34. *Given any $\dot{g} \in \mathbb{M}_n$, $g_+ \in \mathbb{P}_n$, and $u \in \mathbb{U}_n$, there exists unique $\dot{g_+} \in \mathbb{H}_n$ and \dot{u} satisfying $u^*\dot{u} \in \mathfrak{u}_n$, such that*

$$\dot{g} = \dot{g_+}u + g_+\dot{u}. \quad (\text{A.1})$$

Proof. We first rewrite (A.1) as

$$\dot{g} = \dot{g_+}u + g_+uu^*\dot{u} = \dot{g_+}u + gu^*\dot{u}.$$

We want to prove that there exist unique $a_h = \dot{g_+} \in \mathbb{H}_n$ and $a_{ah} = u^*\dot{u} \in \mathfrak{u}_n$ such that

$$\dot{g} = a_hu + ga_{ah}.$$

To do so, we first multiply g^* on the right

$$\dot{g}g^* = a_hug^* + ga_{ah}g^* = a_hg_+ + ga_{ah}g^*,$$

where we have used the fact that $g = g_+u$ to reach the second equality. Therefore, $\dot{g}g^* - a_hg_+ \in \mathfrak{u}_n$, i.e.,

$$\dot{g}g^* + g\dot{g}^* = a_hg_+ + g_+a_h.$$

Since $g_+ > 0$, the above equation has a unique solution [38]

$$a_h = \int_0^\infty \exp^{-\tau g_+} (\dot{g}g^* + g\dot{g}^*) \exp^{-\tau g_+} d\tau.$$

Consequently,

$$a_{ah} = g^{-1}(\dot{g} - a_hu).$$

Thus we have obtained the unique solution for (A.1)

$$\dot{g}_+ = \int_0^\infty \exp^{-\tau g_+} (\dot{g} g^* + g \dot{g}^*) \exp^{-\tau g_+} d\tau,$$

and

$$\dot{u} = u g^{-1} (\dot{g} - a_h u) = g_+^{-1} (\dot{g} - \dot{g}_+ u).$$

□

A connection in GL_n splits the tangent space $(TGL_n)_g$ into a vertical space V_g and a horizontal space H_g , such that $(TGL_n)_g$ is a direct sum of V_g and H_g . In addition, H_g has to be differentiable on g , and H_g is invariant by the group \mathbb{U}_n , i.e., $H_{gu} = H_g u$. The vertical space V_g consists of vectors tangent to the fibre at g . More specifically, let g moves along the fibre over g_+ , $\dot{g}_+ = 0$ which implies $\dot{g} \in g u_n$. Thus the vertical subspace is

$$V_g = \{X \in \mathbb{M}_n : X = g a_{ah} \text{ for } a_{ah} \in \mathfrak{u}_n\} = \{X \in \mathbb{M}_n : X g^* + g X^* = 0\}.$$

The construction of a horizontal space H_g is not unique. One way is to take $H_g = g \mathbb{H}_n$ as in [116], i.e.,

$$H_g = g \mathbb{H}_n = \{X \in \mathbb{M}_n : X g^* = g X^*\}, \quad (\text{A.2})$$

which corresponds to the range space of the projection map

$$\phi_g(X) = \frac{1}{2}(X + g X^* g^{*-1}),$$

while V_g is identified with its null space.

Another way of constructing a horizontal space is to define the inner product in $(TGL_n)_g$. That is for $X_1, X_2 \in (TGL_n)_g$,

$$\langle X_1, X_2 \rangle = \frac{1}{2} \text{tr}(X_1^* X_2 + X_2^* X_1).$$

The horizontal space consists of vectors which are orthogonal to the vertical space V_g with respect to the metric defined above. Some simple calculations yield

$$H_g = \mathbb{H}_n g = \{Y \in \mathbb{M}_n : Y^* g = g^* Y\}. \quad (\text{A.3})$$

To distinguish them, we shall call (A.2) the horizontal condition and (A.3) the parallel condition. For a curve $\gamma(t) \in \mathbb{P}_n$, a smooth curve $\Gamma(t) \in GL_n$ is called a lift of $\gamma(t)$ if $\gamma(t) = \Gamma(t)\Gamma(t)^*$. If

$\gamma(t)$ is a curve of density matrices, then $\Gamma(t)$ is also called a standard purification which lies on the sphere \mathcal{S}_n defined by the Frobenius norm

$$\mathcal{S}_n = \{g \in GL_n : \text{tr}(g^*g) = 1\}.$$

A purification satisfying the parallel condition (A.3) is also called a parallel purification [37][117, p. 239-241]. The Bures metric defines the length of the tangent vector $d\rho$ as the Frobenius norm of the corresponding tangent vector in the horizontal space:

$$\mathfrak{g}_{\rho, \text{Bures}}(d\rho) := \{\langle dW, dW \rangle, \text{ s.t. } d\rho = dW W^* + W dW^*, dW^* W = W^* dW\},$$

or equivalently using $dW = HW$ for some Hermitian matrix H ,

$$\mathfrak{g}_{\rho, \text{Bures}}(d\rho) = \left\{ \frac{1}{2} \text{tr} H d\rho, \text{ s.t. } d\rho = H\rho + \rho H \right\}.$$

Appendix B

An Example of Using Christoffel Symbols to Compute Geodesics

The purpose of this example is to go through the standard procedure of using Christoffel symbols to compute geodesics. Though it is conceptually simple, the computation can be rather complicated and tedious. We start with the Fisher information metric, which defines an inner product for two tangent vectors δ_1 and δ_2 at a point p :

$$\langle \delta_1, \delta_2 \rangle_p = \int \frac{\delta_1 \delta_2}{p}.$$

We shall consider discrete probability distributions on the simplex

$$S = \{(x_1, x_2, x_3) \in \mathbb{R}^3 : x_1, x_2, x_3 \geq 0, x_1 + x_2 + x_3 = 1\}.$$

Each probability distribution can be parameterized by the mapping

$$\mathbf{x} : U \subset \mathbb{R}^2 \rightarrow S : \mathbf{x}(u_1, u_2) = (u_1, u_2, 1 - u_1 - u_2)$$

where

$$U = \{(u_1, u_2) \in \mathbb{R}^2 : u_1 \geq 0, u_2 \geq 0, u_1 + u_2 \leq 1\}.$$

This parametrization determines a basis $\{X_1, X_2\}$ of the tangent plane $T_p S$ to S at p :

$$X_1 = \partial \mathbf{x} / \partial u_1 = (1, 0, -1)^T,$$

and

$$X_2 = \partial \mathbf{x} / \partial u_2 = (0, 1, -1)^T.$$

The local representation (g_{ij}) of the Fisher metric in the above coordinate system is

$$\begin{aligned} g_{11} &= \langle X_1, X_1 \rangle_p = \frac{1}{u_1} + \frac{1}{1 - u_1 - u_2} \\ g_{12} = g_{21} &= \langle X_1, X_2 \rangle_p = \frac{1}{1 - u_1 - u_2} \\ g_{22} &= \langle X_2, X_2 \rangle_p = \frac{1}{u_2} + \frac{1}{1 - u_1 - u_2}. \end{aligned}$$

Its inverse (g^{km}) is

$$(g^{km}) = \begin{pmatrix} (1 - u_1)u_1 & -u_1u_2 \\ -u_1u_2 & (1 - u_2)u_2 \end{pmatrix}.$$

According to the Levi-Civita theorem [118], there exists a unique affine connection on S which is symmetric and compatible with the Fisher metric. The compatibility property of this connection ensures that the inner product of two parallel vector fields along a curve being constant. The Christoffel symbols of this Levi-Civita connection can be computed as follows:

$$\Gamma_{jk}^i = \frac{1}{2} \sum_m \left\{ \frac{\partial g_{mj}}{\partial u_k} + \frac{\partial g_{mk}}{\partial u_j} - \frac{\partial g_{jk}}{\partial u_m} \right\} g^{im}.$$

Below are the full expressions of the Christoffel symbols in our setting:

$$\begin{aligned} \Gamma_{11}^1 &= \frac{1 + u_1(-2 + u_2) - u_2}{2u_1(-1 + u_1 + u_2)} \\ \Gamma_{12}^1 &= \Gamma_{21}^1 = \frac{-u_1}{2(-1 + u_1 + u_2)} \\ \Gamma_{22}^1 &= \frac{(-1 + u_1)u_1}{2u_2(-1 + u_1 + u_2)} \\ \Gamma_{11}^2 &= \frac{(-1 + u_2)u_2}{2u_1(-1 + u_1 + u_2)} \\ \Gamma_{12}^2 &= \Gamma_{21}^2 = \frac{-u_2}{2(-1 + u_1 + u_2)} \\ \Gamma_{22}^2 &= \frac{1 + u_1(-1 + u_2) - 2u_2}{2u_2(-1 + u_1 + u_2)}. \end{aligned}$$

Let $V = v_1X_1 + v_2X_2$ be a vector field along a curve $c(t) = (x_1(t), x_2(t), x_3(t))$ on S , then the covariant derivative of V along $c(t)$ is

$$\frac{DV}{dt} = \sum_i \left\{ \frac{dv_i}{dt} + \sum_{j,k} \Gamma_{jk}^i v_j \frac{dx_k}{dt} \right\} X_i.$$

For a curve $\gamma(t)$ to be a geodesic, its field of tangent vectors $\dot{\gamma}(t)$ has covariant derivative being 0, i.e. $D\dot{\gamma}/dt = 0$. For our parameterized curve $\gamma(t) = \mathbf{x}(u_1(t), u_2(t))$, its tangent vectors are $\dot{\gamma} = \dot{u}_1 X_1 + \dot{u}_2 X_2$. The geodesic equations are computed as

$$\ddot{u}_1 = -\frac{(-1+u_1)u_2^2\dot{u}_1^2 + (-1+u_1)u_1^2\dot{u}_2^2 + u_2\dot{u}_1(\dot{u}_1 - 2u_1\dot{u}_1 - 2u_1^2\dot{u}_2)}{2u_1u_2(-1+u_1+u_2)}, \quad (\text{B.1})$$

$$\ddot{u}_2 = \frac{-u_2^3\dot{u}_1^2 + (-1+u_1)u_1\dot{u}_2^2 - (-2+u_1)u_1u_2\dot{u}_2^2 + u_2^2\dot{u}_1(\dot{u}_1 + 2u_1\dot{u}_2)}{2u_1u_2(-1+u_1+u_2)}. \quad (\text{B.2})$$

We have followed the routine in [119] for computing the Christoffel symbols and the geodesic equation using Mathematica.

The easier way to construct the geodesic in this example is to map the probability distributions on the simplex to the unit sphere by taking the square root. Then the geodesic between p_1 and p_2 on the simplex corresponds to the arc on the unit sphere. The geodesic has a closed form

$$\sqrt{\gamma(t)} = \frac{\sin(1-t)\vartheta}{\sin\vartheta} \sqrt{p_1} + \frac{\sin t\vartheta}{\sin\vartheta} \sqrt{p_2} \quad (\text{B.3})$$

for $\tau \in [0, 1]$, and ϑ denotes the angle between $\sqrt{p_1}$ and $\sqrt{p_2}$, i.e.,

$$\cos\vartheta = p_1 \cdot p_2,$$

where \cdot denotes the dot product.

We verify the geodesic equation (B.1) obtained through covariant derivative by comparing with the closed form (B.3). We first set the starting point to be $p(1) = (0.1, 0.2, 0.7)$ and the initial velocity to be $\dot{u}_1(0) = \dot{u}_2(0) = 0.01$. A geodesic path is generated according to (B.1), where the Matlab toolbox ode45 is used to solve the differential equation. Figure B.1 shows geodesics by two methods.

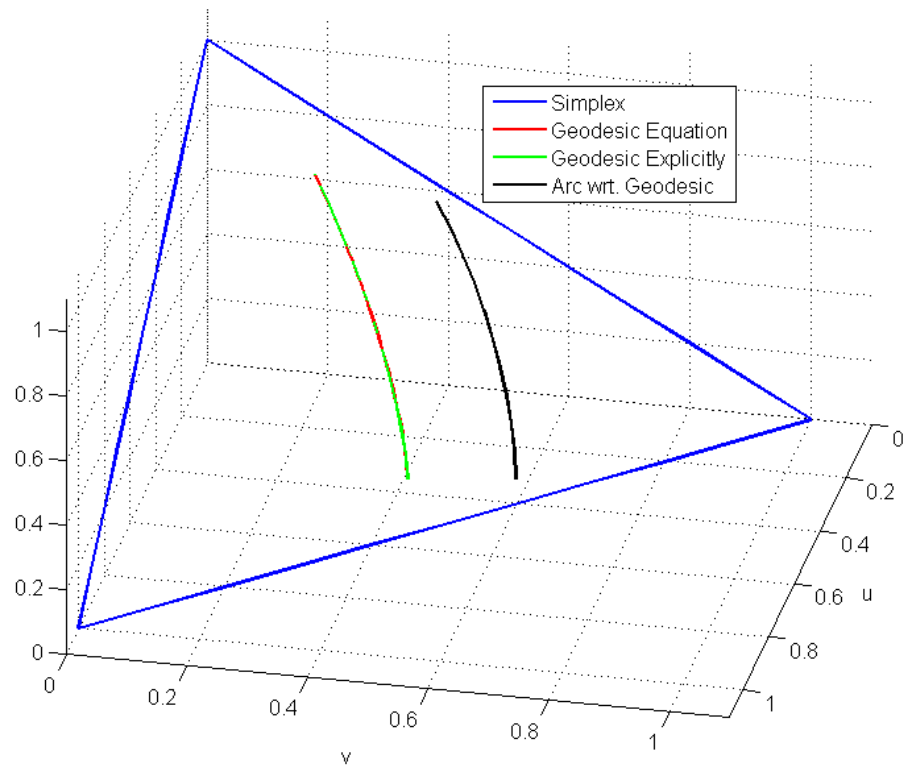


Figure B.1: The red curve is the geodesic path by using the Christoffel symbols. The green curve is based on the closed-form expression (B.3), whose corresponding arc on the unit sphere is plotted in black.

Appendix C

Probability and Measure Review

This appendix provides a brief review of some key concepts and theorems in the probability and measure theory. For more details and derivations of theorems, we refer to classic books in this area, such as [120, 121, 122].

C.1 Topology and Topological Space

Given a set X , a **topology** in X is a collection \mathfrak{T} of subsets of X such that the following properties hold:

1. The set X and empty set \emptyset are in \mathfrak{T} .
2. The intersection of any finite collection of sets in \mathfrak{T} is still in \mathfrak{T} .
3. The union of any collection of sets (finite, countable, or uncountable) in \mathfrak{T} is also in \mathfrak{T} .

The set X together with its topology \mathfrak{T} is called a **topological space**. The most familiar topological spaces arise from **metric spaces**. A metric space is a set X with a distance function (**metric**) d defined in it, such that:

1. $0 \leq d(x, y) \leq \infty$ for all x, y in X .
2. $d(x, y) = 0$ if and only if $x = y$.
3. $d(x, y) = d(y, x)$ for all $x, y \in X$.

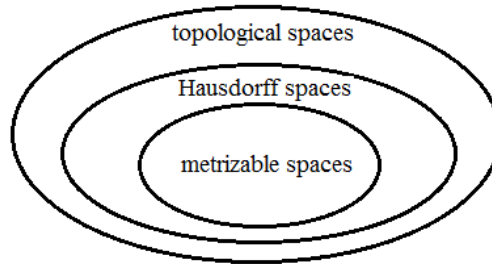


Figure C.1: Topological spaces

$$4. d(x, y) \leq d(x, z) + d(z, y) \text{ for all } x, y, \text{ and } z \in X.$$

A subset in a metric space is said to be **open** if it is a union of open balls. Let (X, d) be a metric space, and \mathfrak{T} be a collection of open sets in X , then (X, \mathfrak{T}) is called the topological space associated with the metric space (X, d) .

Given a topological space (X, \mathfrak{T}) , members of \mathfrak{T} are called open sets in X . When this topological space arises from a metric space, the open sets associated with the topological space are the open sets in the associated metric space. However there are topological spaces which do not associate with metric spaces. For example, let Y be a set of all positive integers, and \mathfrak{T} be a collection of subsets $\{n, n + 1, \dots\}$, then (Y, \mathfrak{T}) is a topological space. In this case, the open sets in (Y, \mathfrak{T}) are not the familiar open sets in the metric spaces. The sub-collection of topological spaces associated with metric spaces is called the **metrizable topological spaces**.

If any pair of distinct points x and y in a topological space (X, \mathfrak{T}) has open sets V_x and V_y containing x and y respectively, and $V_x \cap V_y = \emptyset$, then (X, \mathfrak{T}) is called a **Hausdorff space**. A metric space is obviously a Hausdorff space, but a topological space is not necessary a Hausdorff space as indicated in Figure C.1. For example, let X be $\{1, 2\}$, and \mathfrak{T} be $\{\emptyset, \{1\}, \{1, 2\}\}$, then \mathfrak{T} is a topology in X . But the open sets containing 1 are $\{1\}$ and $\{1, 2\}$, and the only open set containing 2 is $\{1, 2\}$, thus the intersection of open sets containing 1 and 2 respectively is not empty.

Let X, Y be two topological spaces, a function f mapping X into Y is called **continuous** if for every open set V in Y , $f^{-1}(V)$ is an open set in X .

Let f be a mapping from a topological space X to the real line, then f is said to be **lower semicontinuous** if the set $\{x | f(x) > a\}$ is open for any real a . The function f is **upper semicontinuous** if the set $\{x | f(x) < a\}$ is open for any real a . The real function f is continuous if it

is both lower and upper semicontinuous.

Let $\{a_n\}$ be a sequence in $[-\infty, \infty]$, and $b_k = \sup\{a_k, a_{k+1}, \dots\}$, for $k = 1, 2, \dots$, then $\beta = \lim_{k \rightarrow \infty} b_k$ is called the **upper limit** of $\{a_n\}$, and it is denoted as

$$\beta = \limsup_{n \rightarrow \infty} a_n.$$

In a similar way, the **lower limit** of $\{a_n\}$ is defined to be

$$\liminf_{n \rightarrow \infty} a_n = - \limsup_{n \rightarrow \infty} -a_n.$$

The set $\{a_n\}$ converges if

$$\liminf_{n \rightarrow \infty} a_n = \limsup_{n \rightarrow \infty} a_n = \lim_{n \rightarrow \infty} a_n.$$

Let X be a topological space, an **open covering** of a subset B in X is a covering of B consisting of only open sets of X . The space X is said to be **compact** if for every open covering of X , there is a finite sub-covering. A subset of \mathbb{R}^n is compact if and only if it is closed and bounded. A metric space is said to be compact if its associated topological space is compact. A metric space X is compact if and only if every infinite subset A of X has at least an “accumulation point”, which is a point in X such that its neighborhood contains infinitely many distinct points of A .

In a metric space (X, d) , a sequence a_n is said to be a **Cauchy sequence** if $\forall \varepsilon > 0, \exists M$, such that for any $i, j > M, d(a_i, a_j) < \varepsilon$. If any Cauchy sequence in (X, d) converges to a point in X , then X is said to be a **complete metric space**. A compact metric space is complete.

A subset A of a topological space X is **dense** in X if $\bar{A} = X$, where \bar{A} is the closure of A such that every neighborhood of points in \bar{A} contains points in A . If for any nonempty open set V in $X, A \cap V \neq \emptyset$, then A is dense in X . As an example, the set of all rational numbers in the real line is dense. A topological space is said to be **separable** if it contains a countable dense subset. The real line is a separable space.

C.2 Measure and Measurable Space

As an analogy to topology, a collection \mathfrak{F} of subsets of X is said to be a **field** if the following properties hold:

1. $X \in \mathfrak{F}$.

2. For any $A \in \mathfrak{F}$, its complement $A^c \in \mathfrak{F}$.

3. The union of finite subsets in \mathfrak{F} is also in \mathfrak{F} .

A field \mathfrak{M} is said to be a **σ -field** or **σ -algebra** if it is also closed under the countable unions:

4. The union of countable subsets in \mathfrak{M} is also in \mathfrak{M} .

Each member of \mathfrak{M} is called a **measurable set** in X , and such a set X is called a **measurable space**. In a topological space X , a set is called a **Borel set** if it can be formed by countable union, countable intersection, or complement of open sets in X . The collection \mathfrak{B} of all Borel sets on X is called a **Borel σ -algebra**. A function f from a measurable space X to a topological space Y is said to be **measurable** if for any open set V in Y , $f^{-1}(V)$ is measurable. It is easy to see that any continuous function mapping a topological space into another topological space is also a measurable function.

Given a non-empty set X and its σ -field \mathfrak{M} , a **measure** μ on \mathfrak{M} is a non-negative extended real-valued set function, such that for countable unions of pair-wise disjoint sets A_n ,

$$\mu\left(\bigcup_n A_n\right) = \sum_n \mu(A_n).$$

A measure is called a **probability measure** if $\mu(X) = 1$. The triple (X, \mathfrak{M}, μ) is said to be a **measure space** if \mathfrak{M} is a σ -field of the set X , and μ is a measure on \mathfrak{M} . If μ is a probability measure, then (X, \mathfrak{M}, μ) is called a **probability space**.

In particular, let us consider the space \mathbb{R}^n . The **volume** of the closed n -dimensional intervals $I = \{\mathbf{x} : a_i \leq x_i \leq b_i\}$ is

$$v(I) = \prod_{i=1}^n (b_i - a_i).$$

For an arbitrary set E in \mathbb{R}^n , we can find a countable collection S of intervals $\{I_k\}$ to cover E . Considering all such coverings S , the **Lebesgue outer measure** of E is defined to be

$$\mu^*(E) = \inf_S \sum_{I_k \in S} v(I_k).$$

If $E = \bigcup_k E_k$, then the Lebesgue outer measure $\mu^*(E) \leq \sum_k \mu^*(E_k)$. If $E_1 \subset E_2$, then $\mu^*(E_1) \leq \mu^*(E_2)$. Any countable subset of \mathbb{R}^n has outer measure 0. A set $E \subset \mathbb{R}^n$ is said to be Lebesgue measurable if for any $\varepsilon > 0$, there exists an open set G , such that $E \subset G$ and $\mu^*(G - E) < \varepsilon$.

The outer measure of a measurable set is called its **Lebesgue measure**. The union of countable number of measurable sets is measurable. The complement of a measurable set is measurable.

There exist sets that are **nonmeasurable**. The nonmeasurable Vitali set constructed as follows is based on the **axiom of choice**, which states that for a family of arbitrary nonempty sets indexed by a set A , $\{E_a, a \in A\}$, there exists one set consisting of exactly one element from each set E_a . For $x, y \in (0, 1]$, we use $x \oplus y$ to denote $x + y$ or $x + y - 1$ so that $x \oplus y$ lies in $(0, 1]$. x and y are called equivalent if $x \oplus r = y$ for some rational number r in $(0, 1]$. Let H be a subset in $(0, 1]$ formed by taking exactly one represent from each equivalent class according to the axiom of choice. Now let us consider the countably many sets $G_k = H \oplus r_k$ for all rational r_k 's in $(0, 1]$. These sets G_k 's are disjoint. (Otherwise if $h_i + r_i = h_j + r_j$, then h_i and h_j are equivalent. But each member of H is not equivalent, which means $h_i = h_j$. Consequently we have $r_i = r_j$.) Since each point in $(0, 1]$ is included in $\{G_k, k = 1, 2, \dots\}$, we have $(0, 1] = \bigcup_k G_k$. If the set H has measure 0, then $(0, 1]$ has measure 0, which is impossible. If H has a positive measure, then $(0, 1]$ has measure infinite. Thus the Vitali set constructed above is nonmeasurable.

A real-valued function f defined on a set $E \subset \mathbb{R}^n$ is said to be **Lebesgue measurable** if for any finite a , the set $\{\mathbf{x} \in E : f(\mathbf{x}) > a\}$ is measurable.

Now we shall introduce the Lebesgue integral. Let f be a nonnegative real-valued function defined on a set $E \subset \mathbb{R}^n$. The region under f over E is

$$R(f, E) = \{(\mathbf{x}, y) \in \mathbb{R}^{(n+1)} : \mathbf{x} \in E, 0 \leq y \leq f(\mathbf{x}) \text{ if } f(x) < +\infty \\ \text{and } 0 \leq y < +\infty \text{ if } f(x) = +\infty\}.$$

Then the Lebesgue measure of the region $R(f, E)$ is called the **Lebesgue integral** of f over the set E :

$$\int_E f(\mathbf{x}) d\mathbf{x} = \mu^*(R(f, E)).$$

For an arbitrary measurable function f over a set E , it can be written as $f = f^+ - f^-$, where both f^+ and f^- are two measurable nonnegative functions over E . The Lebesgue integral of f over E is

$$\int_E f(\mathbf{x}) d\mathbf{x} = \int_E f^+(\mathbf{x}) d\mathbf{x} - \int_E f^-(\mathbf{x}) d\mathbf{x}.$$

A function f is called **Lebesgue integrable** if its Lebesgue integral exists and finite. The set of all Lebesgue integrable functions is denoted as $L(E) = \{f : |\int_E f| < +\infty\}$.

The difference between the Riemann and Lebesgue integrals is that the Riemann integral starts from the partition of the domain of f , while the Lebesgue integral partitions the range of the function f .

Below we shall state some important theorems regarding the Lebesgue integral.

Theorem 35 (Monotone Convergence Theorem). *Let $\{f_k\}$ be a sequence of Lebesgue integrable functions on the set E ,*

1. *If f_k increases to f a.e., and there exists a Lebesgue integrable function g , such that $f_k > g$ a.e., then $\int_E f_k \rightarrow \int_E f$.*
2. *If f_k decreases to f a.e., and there exists a Lebesgue integrable function g , such that $f_k < g$ a.e., then $\int_E f_k \rightarrow \int_E f$.*

Theorem 36 (Uniform Convergence Theorem). *If a sequence of Lebesgue integrable functions $\{f_k\}$ converges uniformly to f on a set E and E has a finite Lebesgue measure, then f is Lebesgue integrable and $\int_E f_k \rightarrow \int_E f$.*

Theorem 37 (Fatou's Lemma). *Let $\{f_k\}$ be a sequence of Lebesgue integrable functions on the set E , and there exists a Lebesgue integrable function g , such that $f_k > g$ a.e. on E for all k , then*

$$\int_E (\liminf_{k \rightarrow \infty} f_k) \leq \liminf_{k \rightarrow \infty} \int_E f_k$$

Theorem 38 (Lebesgue's Dominated Convergence Theorem). *Let $\{f_k\}$ be a sequence of Lebesgue integrable functions on the set E and $f_k \rightarrow f$ a.e. in E . If there exists a Lebesgue integrable function g , such that $|f_k| \leq g$ a.e. in E for all k , then $\int_E f_k \rightarrow \int_E f$.*

C.3 Convergence

There are several notions of convergence in functional analysis and probability measures.

- **Convergence in measure**

Let (X, \mathfrak{M}, μ) be a measure space. If a sequence of functions $\{f_k\}$ and f are measurable and finite a.e. in a set $E \in \mathfrak{M}$, then f_k converges to f in μ -measure if

$$\forall \varepsilon > 0, \mu\{\mathbf{x} \in E : |f_k(\mathbf{x}) - f(\mathbf{x})| > \varepsilon\} \rightarrow 0.$$

- **Convergence in probability**

Let $\{\mu_k\}$ be a sequence of probability measures and $\{Y_k\}$ be their associated random variables, Y_k is said to converge in probability to Y if μ_k converges to μ in measure.

- **Strong convergence**

In a metric space (X, d) , a sequence $\{x_n\}$ in X is said to converge strongly to x if

$$d(x_n, x) \rightarrow 0.$$

- **Weak convergence in functional analysis**

From the functional analysis viewpoint, the sequence $\{x_n\}$ converges weakly to x in a normed linear space X if

$$\forall f \in X^*, f(x_n) \rightarrow f(x),$$

where X^* is the dual space of X consisting of continuous linear functionals on X . Strong convergence implies weak convergence. They are equivalent if the metric space X is finite dimensional.

- **Weak convergence of probability measures**

Let X be a metric space, and \mathfrak{B} its Borel σ -algebra. A sequence of probability measures $\{\mu_k\}$ on \mathfrak{B} is said to converge weakly in measure to μ if

$$\forall A \in \mathfrak{B}, \mu_k(A) \rightarrow \mu(A),$$

which is equivalent to

$$\forall \text{ bounded continuous } f : X \rightarrow \mathbb{R}, \int f d\mu_k \rightarrow \int f d\mu.$$

If $\{\mu_k\}$ are finite measures on $\mathfrak{B}(\mathbb{R})$, and let $\{F_k\}$ be their corresponding distribution functions, then μ_k converges weakly to μ is equivalent to $F_k(a, b] \rightarrow F(a, b]$ for all $a, b \in \mathbb{R}$.

- **Convergence in distribution**

Let $\{\mu_k\}$ be a sequence of probability measures and $\{Y_k\}$ be their associated random variables, Y_k is said to converge in distribution to Y if μ_k converges weakly to μ . Convergence in probability implies convergence in distribution.

- **Almost sure convergence (or convergence with probability 1)**

A sequence of random variables $\{Y_n\}$ is said to converge almost surely to Y if

$$\Pr(\lim_{n \rightarrow \infty} Y_n = Y) = 1.$$

Almost sure convergence implies convergence in probability.

Appendix D

Another Interpretation of KL Divergence and Fisher Metric

The Kullback-Leibler divergence between two probability distributions p and q of discrete random variables is defined as

$$d_{\text{KL}}(p||q) = \sum_i p_i \log \frac{p_i}{q_i}.$$

From the information theory point of view, the KL divergence measures the average extra bits needed to code a random signal with the incorrect knowledge of distribution q , while the true distribution is p . The KL divergence also arises from the hypothesis testing [43, 123], which we explain as follows.

Assume that you are given a coin, without knowledge of whether the coin is fair. If it is biased, you know that $\text{Pro}(\text{head}) = q_1$ and $\text{Pro}(\text{tail}) = q_2$ with $q_1 + q_2 = 1$. If the coin is fair, then $\text{Pro}(\text{head}) = p_1 = 1/2$ and $\text{Pro}(\text{tail}) = p_2 = 1/2$. Now you need to toss the coin several times to decide whether the coin is fair or not. This is a hypothesis testing problem. Let two hypotheses be H_0 : the coin is distributed as p , and H_1 : the coin is distributed as q . We start with

$$\text{Pro}(H_0) = \text{Pro}(H_1) = \frac{1}{2}.$$

After tossing the coin n times, we find that the outcome $E_n = \{m \text{ heads, } (n - m) \text{ tails}\}$. Then

we can compute the posterior as:

$$\begin{aligned}\text{Pro}(H_0|E_n) &= \frac{\text{Pro}(E_n|H_0) \text{Pro}(H_0)}{\text{Pro}(E_n)} \\ &= \frac{\binom{n}{m} p_1^m p_2^{n-m} \frac{1}{2}}{\text{Pro}(E_n)}\end{aligned}$$

and

$$\text{Pro}(H_1|E_n) = \frac{\binom{n}{m} q_1^m q_2^{n-m} \frac{1}{2}}{\text{Pro}(E_n)}.$$

Their likelihood ratio is

$$\frac{\text{Pro}(H_0|E_n)}{\text{Pro}(H_1|E_n)} = \left(\frac{p_1}{q_1}\right)^m \left(\frac{p_2}{q_2}\right)^{n-m}.$$

The average log likelihood ratio is

$$\frac{1}{n} \log \left(\frac{\text{Pro}(H_0|E_n)}{\text{Pro}(H_1|E_n)} \right) = \frac{m}{n} (\log p_1 - \log q_1) + \frac{n-m}{n} (\log p_2 - \log q_2).$$

If the coin is distributed as p , then $\frac{m}{n} \rightarrow p_1$, and $\frac{n-m}{n} \rightarrow p_2$ as n becomes large. Consequently,

$$\frac{1}{n} \log \left(\frac{\text{Pro}(H_0|E_n)}{\text{Pro}(H_1|E_n)} \right) \rightarrow p_1 (\log p_1 - \log q_1) + p_2 (\log p_2 - \log q_2),$$

where the right hand side is $d_{\text{KL}}(p||q)$. Therefore, the likelihood ratio

$$\frac{\text{Pro}(H_0|E_n)}{\text{Pro}(H_1|E_n)} \rightarrow e^{nd_{\text{KL}}(p||q)}.$$

On the other hand, if the coin is distributed as q , then

$$\frac{\text{Pro}(H_0|E_n)}{\text{Pro}(H_1|E_n)} \rightarrow e^{-nd_{\text{KL}}(q||p)}.$$

If the KL divergence between p and q is large, then from the likelihood ratio of posteriors, one can easily tell the true distribution of the coin.

The small perturbation of the KL divergence gives rise to the Fisher metric. The same metric arises as the uncertainty region of the distribution of histograms.

Let us consider Bernoulli trials, where the outcome of each trial is independent. Assume there are possibly k mutually exclusive outcomes in each experiment, and each with probability

p_1, \dots, p_k . After n experiments, we can count the number of each possible outcome to form a histogram, say $E_n = \{m_1, \dots, m_k\}$ where $n = \sum_i^k m_i$. Since each experiment is independent, the histogram E_n has a multinomial distribution

$$\text{Pro}(E_n) = \frac{n!}{m_1! m_2! \dots m_k!} p_1^{m_1} p_2^{m_2} \dots p_k^{m_k}.$$

According to the Local DeMoivre-Laplace Theorem for multinomial distributions [124, pg. 70], for large n ,

$$\text{Pro}(E_n) \rightarrow \frac{1}{(2\pi n)^{\frac{k-1}{2}} (p_1 p_2 \dots p_k)^{\frac{1}{2}}} e^{-\frac{1}{2} \sum_i^k \frac{(m_i/n - p_i)^2}{p_i/n}},$$

or equivalently by introducing the frequency of occurrence $\{\xi_i = m_i/n, i = 1, \dots, k\}$,

$$\text{Pro}(\{\xi_i\}) \rightarrow \sqrt{2\pi n^{-(2k+1)}} \prod_{i=1}^k N(p_i, p_i/n).$$

The proof is essentially based on the Stirling's formula

$$n! = \sqrt{2\pi n} n^n e^{-n} e^{c_n}$$

where

$$|c_n| \leq \frac{1}{12n}.$$

Therefore, as n increases, $\{\xi_i\}$ is more and more concentrated around p . We can define the uncertainty region of $\{\xi_i\}$ as

$$\sum_i^k \frac{(\xi_i - p_i)^2}{p_i/n} \leq 1$$

or

$$\sum_i^k \frac{(\xi_i - p_i)^2}{p_i} \leq \frac{1}{n}.$$

Let $\xi_i = p_i + \delta_{p_i}$, the left hand side of the above expression becomes

$$\sum_i^k \frac{\delta_{p_i}^2}{p_i},$$

which is exactly the Fisher metric. Furthermore, the transformation $p \rightarrow \sqrt{p}$ maps the simplex to the unit sphere, and for large n , the ellipsoid uncertainty region is mapped to a disc on the unit sphere

$$\sum_i^k (\sqrt{\xi_i} - \sqrt{p_i})^2 \leq \frac{1}{4n},$$

which can be easily verified by expanding $\sqrt{\xi_i}$ in Taylor series around p_i . Figure D.1 (original from [104]) shows the ellipsoid uncertainty region on the simplex (in red) and the corresponding disc on the sphere (in blue). When the uncertainty regions of two distributions overlap

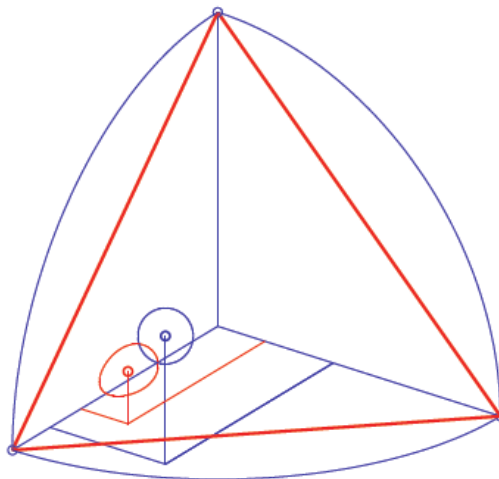


Figure D.1: The ellipsoid uncertainty region on the simplex (in red) and the corresponding disc on the sphere (in blue).

with each other, it is hard to distinguish them. For two distributions p and q far apart, one can find paths connecting them such that locally it is hard to distinguish the point on the path with its neighbors. The one with the minimum length defines a geodesic:

$$\min \int_t \left(\sum_i^k \frac{\dot{p}_i(t)^2}{p_i(t)} \right)^{\frac{1}{2}} dt,$$

which is exactly the geodesic of the Fisher metric. The same idea has been used to define the distance between two preparations of the same quantum system [125].