MANCHESTER
1824

The University of Manchester

Information geometry for control of some stochastic processes

Dodson, CTJ

MIMS EPrint: **2022.4**

Manchester Institute for Mathematical Sciences
School of Mathematics

The University of Manchester

ISSN 1749-9097

Journal of Mathematical Sciences, Vol. xxx, No. y, Month, 20zz

INFORMATION GEOMETRY FOR CONTROL OF SOME STOCHASTIC PROCESSES

C.T.J. Dodson UDC

Abstract. A basic requirement in control systems is a metric that measures discrepancies between actual and desired states. For statistically influenced systems information geometric methods provide natural Riemannian metrics on smooth spaces of states; such manifolds arise in minimum-phase linear systems and multi-input systems with known stochastic noise. Commonly recurring practical situations are 'nearly' Poisson or 'nearly' Uniform with a complementarity in the geometry of these two; another involves multivariate Gaussians and their mixtures. Similarly we encounter 'nearly' independent Poisson, and 'nearly' independent Gaussian processes. For such cases we have information geometric results and examples. Some of these methods are applicable to control systems for statistically influenced processes, such as monitoring essential features in continuous production of threads, films, foils and fibre networks, and batch processing of stochastic textures.

***Key words and phrases*:** information metric, statistical state space, geometry of near-random, near-uniform, multivariate Gaussians, mixture distributions.

***AMS Subject Classification*:** 60D05, 53B20

1. Introduction

The geometrization of models of real phenomena have long been known to give valuable insights because of the established value of analytic geometric features, such as a natural metric, parallelism, perpendicularity, curvature and geodesic curves that are invariant under the choice of coordinate representation. For *Information* geometry this corresponds to the invariance of measure functions of probability density functions under changes of parameters. For example, the family of gamma distributions and the family of log-gamma distributions yield Riemannian 2-manifolds that are isometric isomorphs. The natural information metric provides distances between states and along state trajectories. Using information theory with its underpinning information geometry brings important concepts mirroring physical theory of statistical mechanics: eg. entropy (ie the 'mean log probability density') and its relation to maximum likelihood methods for model optimization. Information geometric methods provide natural neighbourhoods that respect the intrinsic geometry of the space of states, for visualization and monitoring of algorithms, and dynamics of stochastic behaviour trajectories.

Amari [1] Chapter 10, discusses linear systems and complex random variables where the phase is uniformly distributed and represented via a family of probability density functions that are of exponential type. Such families admit information geometric representation. Choi and Mullhaupt [5] showed that the information geometry of a signal filter corresponds to a Kähler manifold and such a manifold arises as a minimum-phase linear system. Zhenning Zhang et al [22] showed that a multi-input system with known stochastic noise and a single output has conditional output probability density shape determined by the control input vector. This control input vector and the output feedback provide a coordinate system for a statistical manifold. They developed a steepest descent algorithm for the control input vector to make the conditional output probability density function as close as possible to the given probability density function, and illustrated with a 2-input 1-output worked example.

Translated from Itogi Nauki i Tekhniki, Seriya Sovremennaya Matematika i Ee Prilozheniya. Tematicheskie Obzory, Vol. 137, Mathematical Physics, 2017.

1. Near-random and near-uniform processes [4, 8]. Gamma probability density functions for random variable $x \in [0, \infty)$ are given by:

$$\mathcal{G} = \{f(x; \nu, \kappa) = \nu^\kappa \frac{x^{\kappa-1}}{\Gamma(\kappa)} e^{-x\nu} \mid \nu, \kappa \in \mathbb{R}^+\} \equiv \mathbb{R}^+ \times \mathbb{R}^+ \tag{1}$$

with mean $\mu = \frac{\kappa}{\nu}$, standard deviation $\sigma = \frac{\sqrt{\kappa}}{\nu}$ so $\frac{\sigma}{\mu} = \frac{1}{\sqrt{\kappa}}$. Gamma distributions are of exponential type [4], since $\log f = \kappa \log \nu - \log \Gamma(\kappa) + \log (x^{\kappa-1} e^{-x\nu})$ and $\int f = 1$, so the gamma potential function is

$$\varphi(\nu, \kappa) = \log \Gamma(\kappa) - \kappa \log \nu. \tag{2}$$

with (ν, κ) as natural parameters. This potential function provides an embedding of the gamma manifold $\mathcal{G}$ in $\mathbb{R}^3$ [10]:

$$e : \mathcal{G} \to \mathbb{R}^3 : (\nu, \kappa) \mapsto (\nu, \kappa, \varphi(\nu, \kappa)). \tag{3}$$

The case $\kappa = 1$ corresponds to the 1-dimensional space of exponential distributions, which characterises Poisson random processes and equation(2) reduces to $\varphi(\nu, 1) = -\log \nu$. Figure 2 shows part of the embedding equation (3) and a tubular neighbourhood of the curve of all exponential distributions.

The smooth Riemannian manifold $\mathcal{G}$ (here a curved surface) of probability gamma density functions §1 has the Fisher-Rao metric tensor equation [2, 4]

$$[g_{ij}](\nu, \kappa) = \begin{bmatrix} \frac{\kappa}{\nu^2} & -\frac{1}{\nu} \\ -\frac{1}{\nu} & \frac{d^2}{d\kappa^2} \log(\Gamma(\kappa)) \end{bmatrix} \tag{4}$$

$$\text{and arc length function} \quad ds^2 = \frac{\kappa}{\nu^2} d\nu^2 - \frac{2}{\nu} d\nu d\kappa + \left(\frac{d^2}{d\kappa^2} \log(\Gamma(\kappa)) \right) d\kappa^2. \tag{5}$$

so a curve in the space $\mathcal{G}$ of gamma distributions

$$c : [0, 1] \to \mathcal{G} : t \mapsto (c_1(t), c_2(t)) \tag{6}$$

has tangent vector $(\dot{c}_1(t), \dot{c}_2(t))$ where the dot signifies differentiation by t and its information length up to $t = T$ is

$$\ell(T) = \int_0^T \sqrt{\sum_{i,j} \dot{c}_i(t) \dot{c}_j(t)\, g_{ij}(c_1(t), c_2(t))}\, dt \tag{7}$$

Via the change of random variables

$$\mathbb{R}^+ \to [0, 1] : x \mapsto a = e^{-x}, \tag{8}$$

we obtain the family of log-gamma probability density functions:

$$\mathcal{L} = \{P(a, \nu, \kappa) = \frac{a^{\nu-1} \nu^\kappa (\log \frac{1}{a})^{\kappa-1}}{\Gamma(\kappa)} \mid \nu, \kappa \in \mathbb{R}^+\} \tag{9}$$

for random variable $a \in [0, 1]$ and parameters $\nu, \kappa > 0$. The limiting case $\kappa = \nu \to 1$ is the uniform distribution, so they give neighbourhoods of uniformity. The limiting case $\kappa = \nu \to \infty$ is the delta function. Moreover, they contain good approximations to truncated univariate Gaussians, which can yield realistic models for real processes with bounded random variable.

The information geometry of these two families of distributions is well-understood; so in each case we have a 2-dimensional Riemannian manifold of states for processes, with the Fisher-Rao information metric equation (4). Indeed, the two manifolds $\mathcal{G}, \mathcal{L}$ are mutual isometric isomorphs, since the metric is invariant under reparametrization via the logarithmic transformation equation (8) [4].

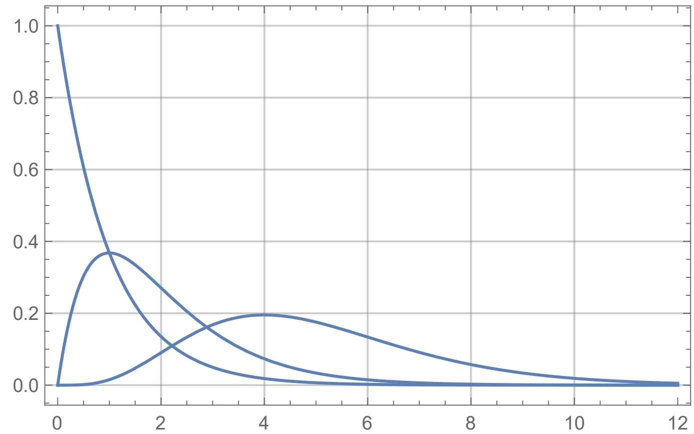


Fig. 1. *Plots of the gamma probability density function equation (1) for* $\kappa = 1, 2, 5,$ and $\nu = 1$. *The case* $\kappa = 1$ *is the exponential function, which corresponds to a Poisson random process, in the depicted case with unit mean.*

2. Relevant results [4, 8].

- Every neighbourhood of a uniform process contains a neighbourhood of processes subordinate to log-gamma probability density functions.
- Every neighbourhood of an independent pair of identical Poisson processes contains a neighbourhood of bivariate processes subordinate to Freund bivariate exponential probability density functions.
- The 5-manifold of bivariate Gaussians admits a 2D subspace in which is a neighbourhood of independence for bivariate Gaussian processes.
- Via the Central Limit Theorem, by continuity, the tubular neighbourhoods of the curve of zero covariance for bivariate Gaussian processes contains all limiting bivariate processes sufficiently close to the independence case for all processes with marginals that converge in probability density function to Gaussians.

Theorem 1 [14, 8]
For independent positive random variables with a common probability density function f*, having independence of the sample mean and the sample coefficient of variation is equivalent to* f *being the gamma distribution.*
This characterization, which implies that the sample standard deviation is proportional to the mean, $\sigma \propto \mu$, is one of the main reasons for the large number of applications of gamma distributions, because many near-random natural processes have observed standard deviation approximately proportional to the mean, as we illustrate in [4]. Examples include the distribution of spacings between consecutive occurrencies of each of the 20 amino acids in protein chains of genomes, which exhibit a 1-dimensional stochastic texture Figure 2 and [4, 8]. Continuously produced 1-dimensional materials include threads [7] and wires, controllable for diameter and linear density variability via radiography, approximately following log-gamma distributions. Paper and board made from water suspensions of cellulose fibres, and air-laid glass and carbon fibre networks are continuously manufactured 2-dimensional stochastic textures [7, 17]. Such fibrous networks tend to have log-gamma distributed local areal density [4] and gamma distributed pore structures [11, 17].

Fig. 2. *Human DNA amino acid sequence: a 20-colour-coded 1-dimensional stochastic texture; and a corresponding sequence using 20 grey levels. All 20 amino acids in genome protein chains have near-random gamma-distributed spacings but all have* $\kappa < 1$*, so they all tend to cluster to differing degrees.*

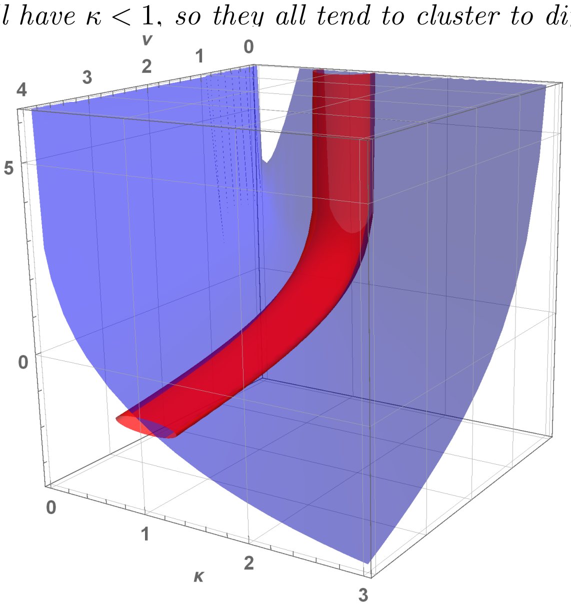


Fig. 3. *A tubular neighbourhood of the curve of all Poisson random processes,*$\kappa = 1$*, in the 2-manifold* $\mathcal{G}$ *of gamma probability densities embedded in* $\mathbb{R}^3$ *via potential function* φ*, equation (2),* [4].

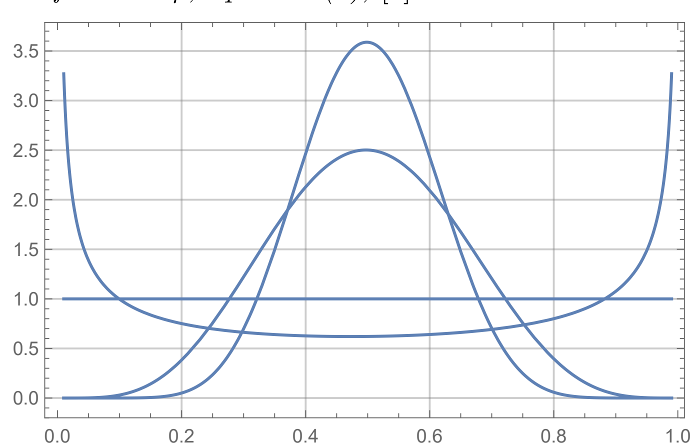
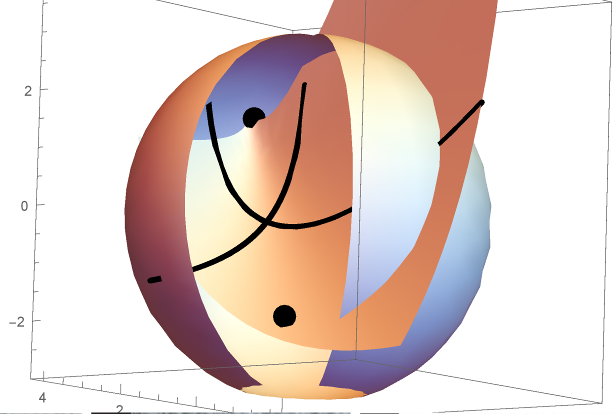


Fig. 4. *Plots of the log-gamma probability density function equation (9) on* $[0, 1]$ *with central mean for* $\kappa = 0.5, 1, 5, 10$*. The case* $\kappa = 1$ *is the uniform distribution.*



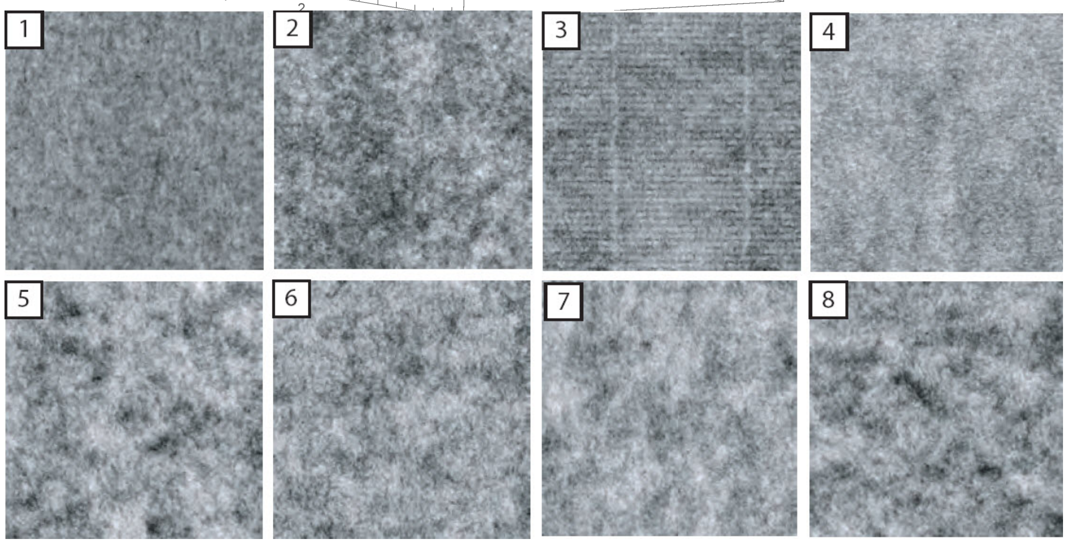


Fig. 6. *Examples of continuously manufactured 2-dimensional stochastic structures [12] which typically have distributions of local areal density that are approximated by log-gamma probability densities with $\kappa > 1$.*

2. Freund 4-Manifold $\mathcal{F}$ and Neighbourhoods of Independence

The 4-manifold $\mathcal{F}$ of Freund bivariate mixture exponential density functions has positive parameters $\alpha_i, \beta_i, i = 1, 2$

$$\mathcal{F} \equiv \{f | f(x, y; \alpha_1, \beta_1, \alpha_2, \beta_2) = \left\{ \begin{array}{ll} \alpha_1 \beta_2 e^{-\beta_2 y - (\alpha_1 + \alpha_2 - \beta_2)x} & \text{for } 0 \leq x < y \\ \alpha_2 \beta_1 e^{-\beta_1 x - (\alpha_1 + \alpha_2 - \beta_1)y} & \text{for } 0 \leq y \leq x \end{array} \right\} | \alpha_i, \beta_i \in \mathbb{R}^+\} \qquad (10)$$

Theorem 2 [4]
Every neighbourhood of an independent pair of identical Poisson processes contains a neighbourhood of bivariate processes subordinate to Freund bivariate exponential probability density functions. A Freund submanifold F_2 is the representation of a bivariate process for which the marginals are identical exponentials.

This provides neighbourhoods of that subspace W in F_2 of the bivariate processes with zero covariance: equivalently, neighbourhoods of independence for Poisson random (ie exponentially distributed) processes. The submanifold $F_2 \subset \mathcal{F} : \alpha_1 = \alpha_2,\ \beta_1 = \beta_2$ forms an exponential family, with parameters (α_1, β_1) and potential function $\varphi = -\log(\alpha_1 \, \beta_1)$.

In F_2, $Cov(X, Y) = 0$ if and only if $\alpha_1 = \beta_1$ and then

$$f(x, y; \alpha_1, \alpha_1) = \alpha_1^2 \; e^{\alpha_1 |y - x|} = f_X(x) f_Y(y)$$

so that here we do have independence of these exponentials if and only if the covariance is zero; this is illustrated in Figure 3.

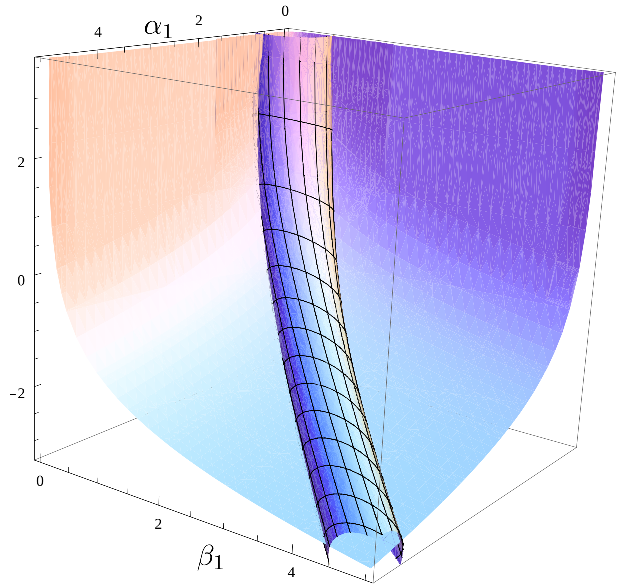


Fig. 7. *Neighbourhood of independent Poisson random processes embedded in a Freund manifold [4].*

3. Spaces of k-variate Gaussians

A k-variate Gaussian probability density function f is defined by its mean k-vector μ, and its $k \times k$ symmetric covariance matrix Σ

$$f(x,\mu,\Sigma) = \frac{e^{-\frac{1}{2}(x-\mu)^T \Sigma^{-1}(x-\mu)}}{\sqrt{(2\pi)^k |\Sigma|}}, \quad x \in \mathbb{R}^k \tag{11}$$

We have used multivariate Gaussians for face recognition using the neighbourhoods of colour pixel features at landmark points in face images [20], where we found that the spatial covariances among pixel colours was important. Craciunesco and Murari et al [6, 16] used geodesic distance on Gaussian manifolds to interpret time series in very large databases from Tokomak measurements in fusion research. Verdoolaege, Shabbir et al [18, 21] used multivariate generalized Gaussians for colour texture discrimination in the wavelet domain. In these studies the discrimination used approximations to the information distance between pairs of multivariate Gaussian probability density functions.

The information distance between two k-variate Gaussians f^A, f^B is known analytically in two particular cases:

1. **Fixed covariance Σ:** $\quad f^A(\mu^A,\Sigma), f^B(\mu^B,\Sigma)$
   The positive definite symmetric quadratic form Σ gives the distance between mean vectors:

$$D_\mu(f^A,f^B) = \sqrt{(\mu^A-\mu^B)^T \cdot \Sigma^{-1} \cdot (\mu^A-\mu^B)}. \tag{12}$$

2. **Fixed mean μ:** $\quad f^A(\mu,\Sigma^A), f^B(\mu,\Sigma^B)$
   Information distance between covariances:

$$D_\Sigma(f^A,f^B) \;=\; \sqrt{\frac{1}{2}\sum_{j=1}^{k} \log^2(\lambda_j^{AB})} \tag{13}$$

$$\{\lambda_j^{AB}\} = \mathrm{Eig}(\Sigma^{A^{-1/2}} \cdot \Sigma^B \cdot \Sigma^{A^{-1/2}})$$

There is no general analytic solution for the geodesic distance between two k-variate Gaussians, but for many practical applications the absolute information distance may be inessential and comparative values suffice. Then, for example,

$$D = \sqrt{D_\mu^2 + D_\Sigma^2} \tag{14}$$

from equations (12) and (13) may be a sufficient approximation. Indeed, equation (13) gives the geodesic distance between f^A with $\Sigma^A = I$ and f^B with $\mu^A = \mu^B = 0$ and the information metric is invariant under affine transformations of the mean. Note also that whereas equation (13) is independent of the mean vectors, equation (12) depends on both covariance and mean vectors and defines a norm on mean vectors for each $f^A(\mu^A, \Sigma)$:

$$||\mu^A|| = \sqrt{(\mu^A)^T \cdot (\Sigma^A)^{-1} \cdot (\mu^A)} \tag{15}$$

which is evidently sensitive to the covariance.

In principle, equation (13) yields all of the geodesic distances since the information metric is invariant under affine transformations of the mean [3] Appendix 1; see also the article of P.S. Eriksen [13]. The equations for the geodesics were shown by Skovgaard [19] to be

$$\begin{aligned} \ddot{\mu} &= \dot{\Sigma}\Sigma^{-1}\dot{\mu} \\ \ddot{\Sigma} &= \dot{\Sigma}\Sigma^{-1}\dot{\Sigma} - \dot{\mu}\dot{\mu}^T. \end{aligned} \tag{16}$$

Eriksen [13] observed that the family $\mathcal{N}^k$ of k-variate Gaussians is isometric to the space $GA^+(k)/SO(k)$ where $GA^+(k)$ consists of positive affine transformations. Hence by a translation it is sufficient to restrict the geodesic to one through $\Sigma = I$ the identity, in the direction $(-B, v)$. Then, through the change of coordinates, $\Delta = \Sigma^{-1}$, $\delta = \Sigma^{-1}\mu$, equation (16) becomes

$$\begin{aligned} \dot{\Delta} &= -B\Delta + v\delta^T \ \text{ with } \ \Delta(0) = I, \delta(0) = 0, \\ \dot{\delta} &= -B\delta + (1 + \delta^T\Delta^{-1}\delta)v. \end{aligned} \tag{17}$$

Then using

$$A = \begin{pmatrix} -B & v & 0 \\ v^T & 0 & -v^T \\ 0 & -v & B \end{pmatrix} \tag{18}$$

Eriksen proved that the geodesic solution curve is given by

$$\Lambda : \mathbb{R} :\to \mathcal{N}^k : t \mapsto e^{At} \ = \ \begin{pmatrix} \Delta & \delta & \Phi \\ \delta^T & 1 + \delta^T\Delta^{-1}\delta & \gamma^T \\ \Phi^T & \gamma & \Gamma \end{pmatrix} \tag{19}$$

$$\text{where } \gamma \ = \ \Delta^{-1}\delta + \Phi^T\Delta^{-1}\delta, \text{ and } \delta^T\Delta^{-1}\delta = \gamma^T\Gamma^{-1}\gamma. \tag{20}$$

$$\text{So } (\Delta(-t), \delta(-t)) \ = \ (\Gamma(t), \delta(t)). \tag{21}$$

Of course, the analytic difficulty is the requirement to find the length of the geodesic between two points in $\mathcal{N}^k$ to obtain a distance function, that being the infimum of arc length over all curves joining the points; so numerical computation would be required for more precise information distances.

4. Spaces of arbitrary mixtures of multivariate Gaussians

We consider a mixture distribution consisting of a linear combination of k-variate Gaussians with an increasing sequence of $k = 2, 3, \ldots, N$ variables:

$$f_2 = (2, \mu_2, \Sigma_2), f_3 = (3, \mu_3, \Sigma_3)..., f_N = (N, \mu_N, \Sigma_N), \ \text{ with } \forall k \int_{\mathbb{R}^k} f_k = 1. \tag{22}$$

where $\mu_k \in \mathbb{R}^k$ is the k-vector of means and $\Sigma_k \in \mathbb{R}^{(k^2+k)/2}$ is the positive definite symmetric $(k \times k)$ covariance matrix with components $(\sigma_{ij}), i \leq j = 1, 2, ..., k$. The standard basis for the space of k-variate covariance matrices Σ, is $E_{ij} = 1_{ii}$ for $i = j$, $E_{ij} = 1_{ij} + 1_{ji}$ for $i \neq j$ so

$$\Sigma = \sum_{i \leq j=1}^{k} \sigma_{ij} E_{ij}.$$

We presume that the parameters and relative weights w_k of these component probability density functions equation (22) have been obtained empirically, giving a mixture density:

$$f = \sum_{k=2}^{N} w_k f_k, \quad \text{with } w_k \geq 0 \text{ and } \sum_{k=2}^{N} w_k = 1. \tag{23}$$

We wish to estimate the information distance $D(f^A, f^B)$ between two such mixture distributions $f^A = (\mu^A, \Sigma^A, w^A)$ and $f^B = (\mu^B, \Sigma^B, w^B)$.

We do not have analytically the distance between two *mixtures* of k-variate Gaussians: $f^A(\mu^A, \Sigma^A, w^A)$ and $f^B(\mu^B, \Sigma^B, w^B)$ so we must resort to approximations for incorporating the weightings of components.

1. Mixtures projected onto the complex plane [9]. The idea here is simple: for each mixture distribution f^A given by a weighted sum (23) we obtain two numbers $||\mu^A||$ and $||\Sigma^A||$ being the weighted sums of norms of means and covariances. The norm on mean vectors is given by D_μ and for the covariance matrices we need a matrix norm, which here we choose as the Frobenius norm given for an $n \times n$ matrix $M_{\alpha\beta}$ by the square root of the sum of squares of its elements $m_{\alpha\beta}$,

$$||M_{\alpha\beta}||^2 = \sum_{\alpha=1}^{n} \sum_{\beta=1}^{n} (m_{\alpha\beta})^2$$

Note that if $M_{\alpha\beta}$ has eigenvalues $\{\lambda_\alpha\}$ and is represented on a basis of eigenvectors then

$$||M_{\alpha\beta}||^2 = \sum_{\alpha=1}^{n} (\lambda_\alpha)^2.$$

Given a mixture distribution f^A consisting of M different multivariate Gaussians: $G^A = \{G_i^A(\mu_i^A, \Sigma_i^A)\}_{i=1,M}$ with weights $w^A = \{w_i^A\}_{i=1,M}$ we have

$$f^A = \sum_{m=1}^{M} w_m^A G_m^A$$

$$||\mu^A|| = \sqrt{\sum_{m=1}^{M} w_m^A \left((\mu_m^A)^T.(\Sigma_m^A)^{-1}.\mu_m^A\right)} \tag{24}$$

$$||\Sigma^A|| = \sqrt{\sum_{m=1}^{M} w_m^A ||\Sigma_m^A||^2}. \tag{25}$$

Now we can represent f^A by the complex number $\phi^A = ||\mu^A|| + i||\Sigma^A||$, and this 2-dimensional representation may be useful in some contexts, for example when following or controlling the trajectory of a process in (μ, Σ) space. Correspondingly, its difference from another such complex number ϕ^B for f^B gives us a distance measure in our reduced space of mixtures:

$$\Delta(f^A, f^B) = |\phi^B - \phi^A|. \tag{26}$$

In [9] we illustrated the effects of differing weighting sequences on given mixtures of k-variate Gaussians.

5. Concluding Comments

1. Information geometry provides a Riemannian manifold structure on the smooth space of parameters of probability density functions.

2. We have illustrated aspects of several families of distributions, in particular the cases of the gamma and multivariate Gaussian density functions.

3. The gamma family contains as a special case the exponential distribution arising from a Poisson process, the log-gamma family contains as a special case the uniform distribution and both cases arise commonly.

4. Some of these methods are applicable to control systems for statistically influenced processes, such as monitoring essential features in continuous production of threads, films, foils and fibre networks and batch processing of stochastic textures.

REFERENCES

1. S-I. Amari. **Methods of Information Geometry and its Applications.** Oxford, Springer Applied Mathematics Series, 2016. https://www.springer.com/gb/book/9784431559771
2. S-I. Amari and H. Nagaoka. **Methods of Information Geometry.** Oxford, American Mathematical Society, Oxford University Press, 2000.
3. C. Atkinson and A. F. S. Mitchell, *Rao's distance measure*, Sankhya: Indian Journal of Statistics, 43 (1981), 345-365. https://www.jstor.org/stable/25050283
4. Khadiga Arwini and C.T.J. Dodson. **Information Geometry Near Randomness and Near Independence**. Lecture Notes in Mathematics, Springer-Verlag, New York, Berlin, 2008. http://www.springer.com/mathematics/geometry/book/978-3-540-69391-8
5. J. Choi and A.P. Mullhaupt. Kählerian information geometry for signal processing. *Entropy* 17 (2015) 1581-1605. DOI: 10.3390/e17041581
6. T. Craciunescu and A. Murari, Geodesic distance on Gaussian manifolds for the robust identification of chaotic systems, *Nonlinear Dynamics*, 86 (2016), 677-693. https://link.springer.com/article/10.1007/s11071-016-2915-x
7. M. Deng and C.T.J. Dodson. **Paper: An Engineered Stochastic Structure**, Tappi Press, Atlanta, 1994. https://imisrise.tappi.org/TAPPI/Products/01/R/0101R238
8. C.T.J. Dodson. Some illustrations of information geometry in biology and physics. In **Handbook of Research on Computational Science and Engineering: Theory and Practice** Eds. J. Leng, W. Sharrock, IGI-Global, Hershey, PA, 2012, pp 287-315. DOI: 10.4018/978-1-61350-116-0
9. C.T.J. Dodson. Information distance estimation between mixtures of multivariate Gaussians. *AIMS Mathematics* 3, 4 (2018) 439-447. DOI: 10.3934/Math.2018.4.448
10. C.T.J. Dodson, and H. Matsuzoe. *An affine embedding of the gamma manifold.* Applied Sciences 5 1 (2003) 1-6. https://eudml.org/doc/229929
11. C.T.J. Dodson and W.W. Sampson. The effect of paper formation and grammage on its pore size distribution. *J. Pulp Pap. Sci.* 22 5 (1996) J165-J169. DOI: 10.1.1.51.405
12. C.T.J. Dodson and W.W. Sampson. Information geometry and dimensionality reduction for statistical structural features of paper. In **Proc. 15th Fundamental Research Symposium**, Cambridge, 8-13 September 2013, ISBN .
13. P. S. Eriksen, Geodesics connected with the Fisher metric on the multivariate normal manifold. In C.T.J. Dodson, Editor, **Proceedings of the Geometrization of Statistical Theory Workshop**, Lancaster (1987), 225-229. https://trove.nla.gov.au/work/18676980

14. R.G. Laha. On a characterization of the gamma distribution. *Annals of Mathematical Statistics* 25 (1954) 784-787. https://www-jstor-org.manchester.idm.oclc.org/stable/2236664
15. M. Leok and J Zhang. Connecting information geometry and geometric mechanics. *Entropy* 19 (2017) 518-550. DOI: 10.3390/e19100569
16. A. Murari, T. Craciunescu, E. Peluso, et al. Detection of Causal Relations in Time Series Affected by Noise in Tokamaks Using Geodesic Distance on Gaussian Manifolds, *Entropy* 19 (2017) 569-580. DOI: 10.3390/e19100569
17. W.W. Sampson. **Modelling Stochastic Fibrous Materials with Mathematica**. Springer-Verlag, London, 2009. https://www.springer.com/gb/book/9781848009905
18. A. Shabbir, G. Verdoolaege and G. Van Oost. Multivariate texture discrimination based on geodesics to class centroids on a generalized Gaussian manifold. In F. Nielsen and F. Barbaresco (Eds) *Geometric Science of Information*, LNCS 8085, Springer, Berlin-Heidelberg, 2013, 853-860.
19. L. T. Skovgaard. A Riemannian geometry of the multivariate normal model, *Scand. J. Stat.* 11 (1984), 211-223. https://www.jstor.org/stable/4615960
20. J. Soldera, C.T.J. Dodson and J. Scharcanski. Face recognition based on texture information and geodesic distance approximations between multivariate normal distributions. *Measurement Science and Technology* 29, 11, 2018. DOI: 10.1088/1361-6501/aade18
21. G. Verdoolaege and A. Shabbir. *Color Texture Discrimination Using the Principal Geodesic Distance on a Multivariate Generalized Gaussian Manifold*, International Conference on Networked Geometric Science of Information, Springer, Cham Berlin-Heidelberg, 2015, 379-386.
22. Zhenning Zhang, Huafei Sun and Linyu Peng. Natural gradient algorithm for stochastic distribution systems with output feedback. *Differential Geometry and its Applications* 31, (2013) 682-690. DOI: 10.1016/j.difgeo.2013.07.004

C.T.J. Dodson
School of Mathematics, University of Manchester, Manchester, M13 9PL, UK.
E-mail: `ctdodson@manchester.ac.uk`