NOTE ON THE METRIC AND AFFINE CONNECTIONS ON THE SPACE OF FINITE MEASURES

HUAIYU ZHU

ABSTRACT. Explicit formulas for the metric and affine connections on the space of finite measures are derived in different coordinate systems. They are transformed and compared with previous formulas of Čencov, Amari and others. These formulas are given for finite sample spaces but in a form that is easily generalizable to infinite sample spaces. An error in Čencov's book is corrected.

1. INTRODUCTION

1.1 Let X be a sample space with a σ-field of measurable sets. Let $\mathcal{P}(X)$ and $\mathcal{M}_+(X)$ be the spaces of all probability measures and finite measures on X, respectively. That is,

$$(1.1) \qquad \mathcal{P}(X) = \left\{ p \in \mathcal{M}_+(X) : \int p = 1 \right\}.$$

1.2 Čencov (1982) proved that for $X = \mathbb{N}_m := \{1, 2, \dots\}$ being a finite set, there is only one (modulo a constant multiplication factor) invariant metric g on $\mathcal{P}(X)$ in the category of congruent (sufficient) Markov morphisms, and there is only one family of equivariant affine connections $\overset{\gamma}{\nabla}$ indexed by a real parameter γ. The metric is defined by the Fisher information matrix.

1.3 Amari (1985) developed the dual affine geometry which associates these concepts beautifully with the information divergence D_γ commonly used in information theory and statistics. He also extended these concepts to $\mathcal{M}_+(X)$ and made clear that the γ-affine structure is simply the affine

structure induced by the γ-coordinate $l_\gamma(p) = p^\gamma/\gamma$. The parameter γ here corresponds to α of Amari with $\alpha = 1 - 2\gamma$. Conversely, Eguchi (1983) showed that the metric g and affine connections $\overset{\gamma}{\nabla}$ are uniquely determined by the second and third order derivatives of the information deviation D_γ.

1.4 Zhu and Rohwer (1997, 1995) applied these results to Bayesian decision theory and revealed the importance of the dual affine structure on $\mathcal{M}_+(X)$. Additional formulas for

$$(1.2) \qquad g_{ij} := g(e_i, e_j), \qquad \overset{\gamma}{\Gamma}_{ijk} = g(e_i \overset{\gamma}{\nabla} e_j, e_k),$$

where e_i is the ith unit coordinate vector, are also given in different τ-coordinates [1].

1.5 Unfortunately, the representations used in these works are somewhat different and details are often omitted, making it difficult to associate these results. In this note we explicitly spell out all these formulas concerning g, $\overset{\gamma}{\Gamma}$ and D_γ in any τ-coordinates for both $\mathcal{M}_+(X)$ and $\mathcal{P}(X)$, where $X = \mathbb{N}_m$.

1.6 We do not derive formulas for arbitrary (ie. infinite) X, but in the final section we discuss what the corresponding formulas would look like. The algebraic forms are similar, but there will be additional topological issues involved. Some additional related topics are also discussed there.

2. METRIC AND AFFINE CONNECTIONS ON $\mathcal{M}_+(X)$

2.1 Notations Except when explicitly noted, we shall assume $X = \mathbb{N}_m$ and omit explicit reference to it. For differential geometry equations, we shall use the index notation but without the summation convention. Generally, $\gamma, \tau \in [0, 1]$, but the boundary cases $\gamma, \tau \in \{0, 1\}$ often require special considerations.

---

[1] The covariant differentiation of v with respect to u is usually written as $\nabla_u v$, but the notation $u\nabla v$ makes explicit the multiplicative role of u and is typographically nicer.

2.2 Coordinates The γ-coordinate ξ of $p \in \mathcal{M}_+$ is an element of a space (Amari, 1985, p. 66)

$$(2.1) \qquad l_\gamma(p) := p^\gamma/\gamma \in \mathcal{M}_+^\gamma \cong \mathbb{R}_+^m,$$

$$(2.2) \qquad l_0(p) := \log p \in \mathcal{M}_+^0 \cong \mathbb{R}_+^m.$$

The tangent space is a linear space,

$$(2.3) \qquad T_p\mathcal{M}_+^\gamma = \mathcal{M}^\gamma \cong \mathbb{R}^m,$$

with norm (except for $\gamma = 0$)

$$(2.4) \qquad \|u\|_{1/\gamma} = \left(\sum_i u_i^{1/\gamma}\right)^\gamma = \left(\sum_i p_i\right)^\gamma/\gamma.$$

2.3 Tangent spaces and bundles We shall from now on fix the association

$$(2.5) \qquad l_\gamma: \quad \mathcal{M}_+ \to \mathcal{M}_+^\gamma,$$

$$(2.6) \qquad T_p l_\gamma: \quad T_p\mathcal{M}_+ \to T_p\mathcal{M}_+^\gamma = \mathcal{M}^\gamma.$$

Although for different γ the space $\mathcal{M}^\gamma \cong \mathbb{R}^m$ is the same linear space, their norms are different. Incidentally, the topologies induced by these norms are the same, even for infinite X, but that is beyond the scope of this note.

At each fixed point p, the tangent mappings $T_p l_\gamma$ are linear, so the pull-back of the affine structures by $T_p l_\gamma$ from $T_p\mathcal{M}^\gamma$ to $T_p\mathcal{M}_+$ are the same for all γ. That is, all the $T_p\mathcal{M}^\gamma$ are isomorphic to the same tangent space $T_p\mathcal{M}_+$. However, because l_γ itself is nonlinear, the isomorphism between $T_p\mathcal{M}^\gamma$ and $T_p\mathcal{M}_+$ is different at different p. This gives rise to different affine connections depending on γ as shown below. For most γ the whole tangent bundle $T\mathcal{M}_+$ is not isomorphic to a single linear space.

2.4 Change of coordinates A change of coordinate from $\xi = p^\gamma/\gamma$ to $u = p^\tau/\tau$ gives

$$(2.7) \qquad \frac{\partial \xi_i}{\partial u_j} = \delta_{ij}p_i^{\gamma-\tau},$$

$$(2.8) \qquad \frac{\partial \xi_i}{\partial u_j \partial u_k} = \delta_{ijk}p_i^{\gamma-2\tau}(\gamma - \tau).$$

These relations will be used repeatedly in the sequel.

2.5 Metric   The metric g on $T\mathcal{M}_+$ is induced by the inner product of $T_p\mathcal{M}_+^{1/2} = \mathcal{M}^{1/2}$, which is independent of p. That is

$$(2.9) \qquad \forall u, v \in \mathcal{M}^{1/2}: \quad g(u,v) = \sum_i u_i v_i.$$

If $u = l_{1/2}(p), v = l_{1/2}(q)$, then

$$(2.10) \qquad g(u,v) = 4 \sum_i \sqrt{p_i q_i}.$$

Let e_i denote the ith standard unit vector in $\mathcal{M}^{1/2}$, then

$$(2.11) \qquad g_{ij} := g(e_i, e_j) = \delta_{ij}.$$

Changing to τ-coordinates transforms the tangent vectors

$$(2.12) \qquad u_i \in \mathcal{M}^{1/2} \to p_i^{\tau - 1/2} u_i \in \mathcal{M}^\tau,$$

and the metric is represented as (Zhu and Rohwer, 1995, p. 30)

$$(2.13) \qquad g_{ij} = p_i^{1-2\tau} \delta_{ij}.$$

2.6 Affine connection   The γ-affine connection on $T\mathcal{M}_+$ is induced by the affine structure of $T_p\mathcal{M}_+^\gamma = \mathcal{M}^\gamma$. That is, for constant vectors,

$$(2.14) \qquad u, v \in \mathcal{M}^\gamma: \quad u \overset{\gamma}{\nabla} v = 0.$$

$$(2.15) \qquad \overset{\gamma}{\Gamma}_{ijk} := g(e_i \overset{\gamma}{\nabla} e_j, e_k) = 0.$$

Changing to τ-coordinates, using both (2.8) and (2.13), gives (Zhu and Rohwer, 1995, p. 30)

$$(2.16) \qquad \overset{\gamma}{\Gamma}_{ijk} = (\gamma - \tau) p_i^{1-3\tau} \delta_{ijk}.$$

Obviously, it vanishes exactly when $\gamma = \tau$. That is, only in $\mathcal{M}^\gamma$ the constant vector fields are γ-parallel transforms of themselves.

2.7 Dual affine   The inner product of γ and $1 - \gamma$ transforms of two vectors remain constant on any curve. That is (Amari, 1985, p. 68)

(2.17)

$$g(u \overset{\gamma}{\nabla} v, w) + g(v, u \overset{1-\gamma}{\nabla} w) = (1 - 2\tau) \sum_i p^{1-3\tau} u_i v_i w_i = (u \nabla g)(v, w)$$

This means $\overset{\gamma}{\nabla}$ and $\overset{1-\gamma}{\nabla}$ are dual-affine with respect to g (Amari, 1985).

2.8 Potential and deviation The γ-potential is defined as (Amari, 1985, p. 83)

$$(2.18) \qquad \psi_\gamma(p) := \sum_i p_i/(1-\gamma), \qquad \psi_1(p) := \sum p_i \log p_i.$$

The γ-deviation is defined as (Amari, 1985, p. 85)

$$(2.19) \qquad D_\gamma(p,q) = \psi_\gamma(p) + \psi_{1-\gamma}(q) - l_\gamma(p) \cdot l_{1-\gamma}(q).$$

This expands to (Zhu and Rohwer, 1995, p. 16)

$$(2.20) \qquad D_\gamma(p,q) = \sum_i \frac{\gamma p_i + (1-\gamma)q_i - p_i^\gamma q_i^{1-\gamma}}{\gamma(1-\gamma)}.$$

This expression was also studied in (Vajda, 1989, p. 228) without using the dual affine geometry.

2.9 On $\mathcal{P}$, because $\sum p = \sum q = 1$, it simplifies to (Amari, 1985, p. 87)

$$(2.21) \qquad D_\gamma(p,q) = \frac{1 - \sum_i p_i^\gamma q_i^{1-\gamma}}{\gamma(1-\gamma)}.$$

Familiar examples include the Hellinger distance, the Kullback-Leibler deviation and the χ^2-deviation:

$$(2.22) \qquad D_{1/2}(p,q) = 2\sum_i (\sqrt{p_i} - \sqrt{q_i})^2,$$

$$(2.23) \qquad D_1(p,q) = \sum_i p_i \log \frac{p_i}{q_i},$$

$$(2.24) \qquad D_2(p,q) = \sum_i \frac{(p_i - q_i)^2}{2q_i}.$$

2.10 Differentiation Let $u_i = p_i^\tau/\tau$, $v_i = q_i^\tau/\tau$, then

$$(2.25) \qquad \partial_{u_i} \partial_{v_j} D_\gamma(p,q) = -p_i^{\gamma-\tau} q_i^{1-\gamma-\tau} \delta_{ij}.$$

$$(2.26) \qquad \partial_{u_i} \partial_{u_j} \partial_{v_k} D_\gamma(p,q) = -(\gamma-\tau)p_i^{\gamma-2\tau} q_i^{1-\gamma-\tau} \delta_{ijk}.$$

Obviously this is indeed consistent with (Eguchi, 1983),

$$(2.27) \qquad g_{ij} = -\partial_{u_i}\partial_{v_j} D_\gamma(p,q)\big|_{q=p},$$

$$(2.28) \qquad \overset{\gamma}{\Gamma}_{ijk} = -\partial_{u_i}\partial_{u_j}\partial_{v_k} D_\gamma(p,q)\big|_{q=p}.$$

2.11 Conjugate If the two vectors are represented in mutually dual coordinates, then

$$(2.29) \qquad u \in \mathcal{M}^\tau,\ v \in \mathcal{M}^{1-\tau} \implies g(u,v) = \int uv.$$

2.12 Dual coordinates We have (Zhu and Rohwer, 1995, p. 30)

$$(2.30) \qquad g^{ij} = p^{2\tau-1}\delta^{ij}.$$

$$(2.31) \qquad \overset{\gamma}{\Gamma}_{ij}{}^k = (\gamma-\tau)p_i^{-\tau}\delta_{ij}{}^k.$$

In particular, for $\tau = 0$, ie. in exponential coordinates,

$$(2.32) \qquad \overset{\gamma}{\Gamma}_{ij}{}^k = \gamma\delta_{ij}{}^k.$$

This will be shown (§3) to be consistent with the equations (12.12) and (12.13) in (Čencov, 1982, p. 175).

3. THE METRIC AND AFFINE CONNECTIONS ON $\mathcal{P}$

3.1 Equivalent projections The space of probability measures $\mathcal{P}$ is a submanifold of $\mathcal{M}_+$ characterized by $\sum_i p_i = 1$. Therefore not all p_i are independent. There are at least three ways to "project" the metric and affine connections from $\mathcal{M}_+$ to $\mathcal{P}$:

- Projecting the tangent vectors to $T\mathcal{P}$ after an infinitesimal parallel transport (Amari, 1985, p. 38–40).
- Using the same formula with all the redundant components as if they are independent (Lauritzen, 1987, p. 187 (3.10)).
- Restricting D_γ to $\mathcal{P}$, and using (2.27) and (2.28).

Their equivalence is due to the fact that vectors in the tangent space $T\mathcal{P}$ are constrained by a differential relation implied by $\sum_i p_i = 1$, which in the τ

coordinate is

$$(3.1) \qquad \forall u \in T_p \mathcal{P}^\tau \subset \mathcal{M}^\tau : \qquad \sum_i p_i^{1-\tau} u_i = 0.$$

The expressions for g and $\overset{\gamma}{\Gamma}$ in $T_p \mathcal{P}^\tau$ are therefore

$$(3.2) \qquad g_{ij} = p_i^{1-2\tau} \delta_{ij}. \qquad \overset{\gamma}{\Gamma}_{ijk} = (\gamma - \tau) p_i^{1-3\tau} \delta_{ijk}$$

$$(3.3) \qquad g^{ij} = p^{2\tau-1} \delta^{ij}. \qquad \overset{\gamma}{\Gamma}_{ij}{}^k = (\gamma - \tau) p_i^{-\tau} \delta_{ij}{}^k.$$

3.2 Čencov's affine connections Čencov (1982, p. 175) showed that the following family of affine connections is unique in the category of Markov morphism on $T_p \mathcal{P}$,

$$(3.4) \qquad X_i \nabla X_j = -\gamma(p_j X_i + p_i X_j), \qquad i \neq j$$

$$(3.5) \qquad X_i \nabla X_i = \gamma(1 - 2p_i) X_i,$$

where $X_i \in T_p \mathcal{P}^0$. This corresponds to

$$(3.6) \qquad \overset{\gamma}{\Gamma}_{ij}{}^k = \gamma(\delta_{ij}{}^k - p_i \delta_j^k - p_j \delta_i^k).$$

Since $\forall u, v \in T_p \mathcal{P}^0$,

$$(3.7) \qquad \sum_i p_i u_i = \sum_i p_i v_i = 0,$$

this is also equivalent to

$$(3.8) \qquad \sum_{ij} \overset{\gamma}{\Gamma}_{ij}{}^k u_i v_j = \gamma u_k v_k, \qquad \overset{\gamma}{\Gamma}_{ij}{}^k = \gamma \delta_{ij}{}^k,$$

which is identical to (2.32).

The difference between (2.32) and (3.6) for $u \overset{\gamma}{\nabla} v$ only shows up when $u, v \in T_p \mathcal{M}_+^0 \setminus T_p \mathcal{P}^0$. In that case (3.6) gives the covariant differentiation of their projections into $T_p \mathcal{P}^0$ while (2.32) gives differentiations in $T_p \mathcal{M}_+^0$.

3.3 Special cases Here we correct a small but confusing error on (Čencov, 1982, p. 177), where he gave the mixture, metric, and exponential connections as corresponding to $\gamma = -1, 3/2, 0$, while in fact they should be

$\gamma = 1, 1/2, 0$. Here we follow the convention of regarding covariant vectors as differentials. Let

(3.9) $V := \partial_{p_1}, \qquad X_i := fV, \qquad f := p_1(1 - p_1).$

(3.10) $u_i = 2\sqrt{p_i}, \qquad U = \partial_{u_1} = \sqrt{f}V = f^{-1/2}X_1.$

Then the mixture connection is given by $VV = 0$,

(3.11) $X_1X_1 = fV(fV) = fVfV = (1 - 2p_1)X_1,$

corresponding to $\gamma = 1$. The metric connection is given by $UU = 0$,

(3.12) $X_1X_1 = \sqrt{f}U(\sqrt{f}U) = fV\sqrt{f}U = \frac{1}{2}(1 - 2p_1)X_1,$

corresponding to $\gamma = 1/2$. The exponential connection is of course given by $X_1X_1 = 0$, corresponding to $\gamma = 0$.

The proof for all γ goes almost identically. Let $w := l_\gamma(p)$, $W := \partial_{w_1} = f^{-\gamma}X_1 = f^{1-\gamma}V$. Then the γ-connection as defined here is given by $WW = 0$,

(3.13)

$X_1X_1 = f^\gamma W(f^\gamma W) = f^\gamma W f^\gamma W = \gamma X_1 f f^{\gamma-1}W = \gamma(1 - 2p_1)X_1,$

which shows the γ is the same parameter as used by Čencov.

4. ELIMINATING REDUNDANT COORDINATES

4.1 Constraint on tangent space   The formulas in (Amari, 1985) eliminates redundant coordinate using $\sum_i p_i = 1$. Here we use a slightly different approach which is perhaps somewhat simpler. Denote $\sum' := \sum_1^{n-1}$, and generally use a primed symbol to denote corresponding object on $\mathcal{P}$ with index running from 1 through $n - 1$. For $u \in T_p\mathcal{P}^\tau$,

(4.1) $\sum_i' p_i^{1-\tau}u_i + p_n^{1-\tau}u_n = 0.$

(4.2) $u_n = -\sum_i' \left(\frac{p_i}{p_n}\right)^{1-\tau} u_i.$

This can be used to simply substitute out the last element in each sum to arrive at the desired formula, as the following paragraphs show.

4.2 Metric   We have

$$(4.3) \quad g(u,v) = \sum_{ij} g_{ij} u_i v_j = \sum_i {}' p_i^{1-2\tau} u_i v_i + p_n^{1-2\tau} u_n v_n = \sum_i {}' g'_{ij} u_i v_i$$

$$(4.4) \qquad\qquad g'_{ij} := p_i^{1-2\tau} \delta_{ij} + p_n^{-1} (p_i p_j)^{1-\tau}.$$

In particular,

$$(4.5) \qquad\qquad \tau = 0 \implies g'_{ij} = p_i \delta_{ij} + p_n^{-1} p_i p_j.$$

$$(4.6) \qquad\qquad \tau = 1 \implies g'_{ij} = p_i^{-1} \delta_{ij} + p_n^{-1},$$

the $\tau = 1$ formula being identical to that on (Amari, 1985, p. 31).

4.3 γ-affine connection   We have

$$g(u\overset{\gamma}{\nabla}v, w) = \sum_{ijk} \overset{\gamma}{\Gamma}_{ijk} u_i v_j w_k$$

$$(4.7) \qquad = (\gamma - \tau) \left(\sum_i {}' p_i^{1-3\tau} u_i v_i w_i + p_n^{1-3\tau} u_n v_n w_n \right)$$

$$= \sum_{ijk} {}' \overset{\gamma}{\Gamma}'_{ijk} u_i v_j w_k$$

$$(4.8) \qquad \overset{\gamma}{\Gamma}'_{ijk} := (\gamma - \tau) \left(p_i^{1-3\tau} \delta_{ijk} - p_n^{-2} (p_i p_j p_k)^{1-\tau} \right)$$

In particular,

$$(4.9) \qquad\qquad \tau = 0 \implies \overset{\gamma}{\Gamma}'_{ijk} = \gamma \left(p_i \delta_{ijk} - p_n^{-2} p_i p_j p_k \right).$$

$$(4.10) \qquad\qquad \tau = 1 \implies \overset{\gamma}{\Gamma}'_{ijk} = (\gamma - 1) \left(p_i^{-2} \delta_{ijk} - p_n^{-2} \right),$$

the $\tau = 1$ formula being identical to that on (Amari, 1985, p. 43).

4.4 Dual coordinate   The corresponding formulas for the metric are

$$(4.11) \qquad\qquad g'^{ij} = p_i^{2\tau - 1} \delta^{ij} - (p_i p_j)^{\tau}.$$

$$(4.12) \qquad\qquad \tau = 0 \implies g'^{ij} = p_i^{-1} \delta^{ij} - 1.$$

$$(4.13) \qquad\qquad \tau = 1 \implies g'^{ij} = p_i \delta_{ij} - p_i p_j.$$

and those for the affine connection are

(4.14) $\overset{\gamma}{\Gamma'}_{ij}{}^{k} = (\gamma - \tau)\left(p_i^{-\tau}\delta_{ij}{}^{k} - \delta_{ij}p_i^{1-2\tau}p_k^{\tau} - p_n^{-1}(p_ip_j)^{1-\tau}p_k^{\tau}\right).$

(4.15) $\tau = 0 \implies \overset{\gamma}{\Gamma'}_{ij}{}^{k} = \gamma\left(\delta_{ij}{}^{k} - \delta_{ij}p_i - p_n^{-1}p_ip_j\right).$

(4.16) $\tau = 1 \implies \overset{\gamma}{\Gamma'}_{ij}{}^{k} = (\gamma - 1)\left(p_i^{-1}\delta_{ij}{}^{k} - \delta_{ij}p_i^{-1}p_k - p_n^{-1}p_k\right).$

5. DISCUSSION OF RELATED ISSUES

5.1 It is obvious that the generalization of these equations to arbitrary sample space X requires calculus on the fractional powers of measures. It can be shown (Zhu, 1998), using ideas in (Neveu, 1965) that the linear spaces $\mathcal{M}^{\gamma}(X)$ spanned by these elements are Banach spaces with the desired duality, if we restrict attention to $\gamma \in (0,1)$. These spaces generalizes the corresponding classical Lebesgue function spaces and act as enveloping tangent spaces. The cases of $\{0,1\}$ are very important but need special treatment because of nonequivalence of topologies. In this paper the issue of topology does not arise because X is assumed to be finite.

5.2 The formulas for the metric and affine connections on arbitrary X look very similar to those given here. In nonparametric fashion, we have, for $p \in \mathcal{M}_+, u, v, w \in T_p\mathcal{M}_+^{\tau} = \mathcal{M}^{\tau}$,

(5.1) $g(u, v) = \int p^{1-2\tau}uv,$

(5.2) $g(u\overset{\gamma}{\nabla}v, w) = (\gamma - \tau)\int p^{1-3\tau}uvw.$

(5.3) $u\overset{\gamma}{\nabla}v = (\gamma - \tau)p^{1-3\tau}uv \in \mathcal{M}^{1-\tau}.$

These are calculations in the generalized Lebesgue spaces.

5.3 Campbell (1986) showed that the extension of g from $P(X)$ to $\mathcal{M}_+(X)$ is not unique under classical Markov morphisms. It is likely that the same is true for $\overset{\gamma}{\nabla}$. The metric and affine connections defined here may be called the canonical versions. We conjecture that if the Markov morphisms are relaxed to allow unnormalized mappings, the canonical versions are again unique equivariant extensions. On the other hand, it is shown elsewhere that

the extensions of the canonical versions from $X = \mathbb{N}_m$ to arbitrary X are unique.

5.4 Amari (1985) uses functions of random variables and their differentials where we uses functions of measures and their differentials as models of tangent vectors. It is trivial to verify that they are equivalent, as long as all variables with the same distributions are identified.

REFERENCES

Amari, S. (1985). *Differential-Geometrical Methods in Statistics*, volume 28 of *Springer Lecture Notes in Statistics*. Springer-Verlag, New York.

Campbell, L. L. (1986). An extended Čencov characterization of a riemannian metric. *Trans. Amer. Math. Soc.*, **98**(1), 135–141.

Čencov, N. N. (1982). *Optimal Decision Rules and Optimal Inference*. Amer. Math. Soc., Rhode Island. Translation from Russian, 1972, Nauka, Moscow.

Eguchi, S. (1983). Second order efficiency of minimum contrast estimators in a curved exponential family. *Ann. Statist.*, **11**, 793–803.

Lauritzen, S. L. (1987). Statistical manifolds. In S. Amari, O. E. Barndorff-Nielsen, R. E. Kass, S. L. Lauritzen, and C. R. Rao, editors, *Differential Geometry in Statistical Inference*, volume 10 of *IMS Lecture Notes Monograph*, chapter 4, pages 163–216. Inst. Math. Statist., Hayward, CA.

Neveu, J. (1965). *Mathematical Foundations of the Calculus of Probability*. Holden-Day, San Francisco. Translated from French, 1964, Masson.

Vajda, I. (1989). *Theory of Statistical Inference and Information*. Kluwer Academic, Dordrecht. Translation of: Teória informácie a štatistického rozhodovania.

Zhu, H. (1998). Generalized Lebesgue spaces and applications to statistics. Working Paper SFI-98-06-44, Santa Fe Institute. Submitted.

Zhu, H. and Rohwer, R. (1995). Information geometric measurements of generalisation. Technical Report NCRG/4350, Aston University. `ftp:/` `/cs.aston.ac.uk/neural/zhuh/generalisation.ps.Z.`

Zhu, H. and Rohwer, R. (1997). Measurements of generalisation based on information geometry. In S. W. Ellacott, J. C. Mason, and I. J. Anderson, editors, *Mathematics of Neural Networks: Models Algorithms and Applications*, pages 394–398. Kluwer, Boston. Proc. Math. of Neural Networks and Appl. Conf. (MANNA), Oxford, 3–7 July 1995.

SANTA FE INSTITUTE, 1399 HYDE PARK ROAD, SANTA FE, NM 87501
E-mail address: `zhuh@santafe.edu`