University of St Andrews

Scotland's First University

School of Mathematics and Statistics

A review of geometric statistics

Peter Jupp

University of St. Andrews

http://www.mcs.st-andrews.ac.uk/~pej/

Plan of talk

1. Generalities
2. Geometric objects (Technicalities)
3. Statistical applications
4. The future?

1. GENERALITIES

Geometric Statistics

(a) Geometry of sample space:

Data analysis in $\mathbb{R}^d$ (data inspection, PCA, . . . ,

persistent homology, . . .)

Spatial statistics (predictors in $\mathbb{R}^d$)

Image analysis

Directional statistics (directions, axes, rotations, . . .)

Shape statistics

(b) Geometry of parameter space:

'Information geometry'

‘Information geometry’

$\sim 50,000$ hits on Google

Differential geometry in parametric statistics (higher-order asymptotics)

Computational geometry

Machine learning; Neural manifolds

Thermodynamics

Encouraging quotations

Differential geometry [in statistics] *is dead*

This reminds me of the time when we believed that differential geometry was the answer to everything

Why differential geometry?

(a) ‘Causes’:

Invariance under re-parameterisation

Parameter spaces not vector spaces but manifolds

Appropriate language for Taylor expansion
(typically, estimator ‘close’ to true parameter value)

(b) Purposes:

conceptual

computational

Differential geometry seems hard!

3 approaches:

(a) intuitive 'tangible' geometry

(b) coordinate-based dialect $g_{ij}d\theta^i d\theta^j$

(c) coordinate-free dialect g

2. GEOMETRIC OBJECTS

Geometric objects

manifolds

tensors

Riemannian metrics

(affine) connections

curvature

yokes

volumes

symplectic forms

manifolds

Manifold Θ locally like $\mathbb{R}^d$

Local coordinates $\theta^1, \ldots \theta^d$

Tangent space $T_\theta\Theta$ to Θ at θ spanned by 'arrows'

$$\frac{\partial}{\partial\theta^1}, \ldots, \frac{\partial}{\partial\theta^d}$$

Cotangent space $T_\theta^*\Theta = (T_\theta\Theta)^*$ spanned by $d\theta^1, \ldots, d\theta^d$

tensors

Tensor = multi-linear form on tangent and cotangent spaces

$$T_\theta\Theta \times \ldots \times T_\theta\Theta \times T_\theta^*\Theta \times \ldots \times T_\theta^*\Theta \to \mathbb{R}$$

E.g.

tangent $a^i \frac{\partial}{\partial \theta^i}$

cotangent $a_i d\theta^i$

score statistic $\frac{\partial l}{\partial \theta^i} d\theta^i$

Riemannian metric

Riemannian metrics

inner product on $T_\theta \Theta$ $g_{ij} d\theta^i d\theta^j$

measures length of infinitessimal vectors

E.g. Fisher information

$$E_\theta \left[\frac{\partial l}{\partial \theta^i} \frac{\partial l}{\partial \theta^j} \right] d\theta^i d\theta^j$$

(affine) connections

connection ($\leftrightarrow$ covariant differentiation
$\leftrightarrow$ parallel translation)

[3 definitions!]

generalisation of directional derivative:

X, Y vector fields on Θ

$\nabla_X Y$ is derivative of Y in direction of X

$$\nabla_{\frac{\partial}{\partial\theta^i}} \frac{\partial}{\partial\theta^j} = \Gamma_{ij}^{\;k} \frac{\partial}{\partial\theta^k}$$

Curvature

(a) Riemann curvature of connection ∇

measures asymmetry of 2nd derivative ∇^2

(b) Embedding curvature (= second fundamental form = Euler–Schouten curvature) of inclusion $\Omega \subset \Theta$ using connection ∇ on Θ

measures how Ω curves inside Θ

Yokes

Yoke ($\simeq$ contrast function $\simeq$ divergence)

$$g : \Theta \times \Theta \to \mathbb{R}$$

Coordinates $(\theta^1, \ldots, \theta^d; \theta'^1, \ldots, \theta'^d)$ on $\Theta \times \Theta$.

g is yoke on Θ if

(i) $\not{g}_{i;}(\theta) = \left. \frac{\partial g(\theta,\theta')}{\partial \theta^i} \right|_{\theta'=\theta} = 0$,

(ii) the matrix $[\not{g}_{i;j}(\theta)] = \left[\left. \frac{\partial^2 g(\theta,\theta')}{\partial \theta^i \partial \theta'^j} \right|_{\theta'=\theta} \right]$ is non-singular

Expected likelihood yoke

Parametric statistical model with parameter space Θ.

Expected likelihood yoke on Θ

$$g(\theta, \theta') = E_{\theta'}[l(\theta; x) - l(\theta'; x)] = -K(\theta', \theta)$$

Yokes yield natural coordinates

g yoke on Θ, $\theta \in \Theta$, $\alpha \in \mathbb{R}$

$$\overset{\alpha}{\Gamma}_\theta \colon \Theta \longrightarrow T_\theta\theta$$

$$\overset{\alpha}{\Gamma}_\theta\,(\theta') = \not{g}^{r;u}(\theta)\left\{\frac{1+\alpha}{2}\frac{\partial g(\theta';\theta)}{\partial\theta^u} + \frac{1-\alpha}{2}\frac{\partial g(\theta;\theta')}{\partial\theta^u}\right\}\frac{\partial}{\partial\theta^r}$$

$$g \circ \left(\overset{\alpha}{\Gamma}_\theta \times \overset{\alpha}{\Gamma}_\theta\right)^{-1} : T_\theta\theta \times T_\theta\theta \longrightarrow \Theta \times \Theta \longrightarrow \mathbb{R}$$

Yokes yield tensors

$$g \circ \left(\overset{\alpha}{\Gamma}_\theta \times \overset{\alpha}{\Gamma}_\theta\right)^{-1} : \underbrace{T_\theta\Theta \times T_\theta\Theta}_{\text{vector space}} \longrightarrow \Theta \times \Theta \longrightarrow \mathbb{R}$$

Derivatives at $(0, 0)$ are tensors

$$\not{T}_{r;s} = \not{g}_{r;s} \qquad \text{Riemannian metric}$$

$$\not{T}_{r;st} = \not{g}_{r;st} - \not{g}_{st;r} \qquad \textit{skewness tensor}$$

$$\overset{\alpha}{\Gamma}{}^r_{st} = \not{g}^{r;u}\left(\frac{1+\alpha}{2}\not{g}_{ru;t} + \frac{1-\alpha}{2}\not{g}_{t;ru}\right) \qquad \text{connection}$$

Volumes

Volume on Θ
$=$ measure on Θ with smooth, everywhere-positive density

$$a(\theta)d\theta^1 \wedge \ldots \wedge d\theta^d = a(\theta)d\theta^1 \ldots d\theta^d$$

Riemannian metrics yield volumes
Riemannian metric g on Θ gives Riemannian volume

$$|g_{i;j}(\theta)|^{1/2} d\theta^1 \wedge \ldots \wedge d\theta^d$$

infinitesimal unit cuboids have unit infinitesimal volume

Fisher information $\mapsto$ Jeffreys' prior

Symplectic forms

Riemannian metric:
non-singular symmetric $(0, 2)$-tensor g $\qquad g_{ij} d\theta^i d\theta^j$

Symplectic form:
non-singular *skew-symmetric* $(0, 2)$-tensor

$$\eta \qquad \eta_{ij} d\theta^i \wedge d\theta^j$$

Symplectic form η gives symplectic volume

$$\eta \wedge \ldots \wedge \eta \qquad |\eta_{ij}|\, d\theta^1 \wedge \ldots \wedge d\theta^d$$

Yokes yield symplectic forms

Coordinates $(\theta^1, \ldots, \theta^d)$ on Θ give coordinates $(\theta^1, \ldots, \theta^d, \theta'^1, \ldots, \theta'^d)$ on $\Theta \times \Theta$

Yoke g on Θ gives symplectic form

$$\eta(\theta, \theta') = g_{i;j}(\theta, \theta') d\theta^i \wedge d\theta'^j$$

on neighbourhood of the diagonal $\{(\theta, \theta) : \theta \in \Theta\}$ of $\Theta \times \Theta$

Yokes yield volumes

Yoke g on Θ gives

(a) Riemannian metric $\not{g}_{i;j}(\theta)d\theta^i d\theta^j$ on Θ

(b) symplectic form $g_{i;j}(\theta,\theta')d\theta^i \wedge d\theta'^j$ on (part of) $\Theta \times \Theta$

and so

(a) Riemannian volume

$$\left|\not{g}_{r;s}(\theta)\right|^{1/2} \left|\not{g}_{i;j}(\theta')\right|^{1/2} d\theta^1 \wedge \ldots \wedge d\theta^d \wedge d\theta'^1 \wedge \ldots \wedge d\theta'^d$$

(b) symplectic volume

$$|g_{i;j}(\theta,\theta')|\, d\theta^1 \wedge \ldots \wedge d\theta^d \wedge d\theta'^1 \wedge \ldots \wedge d\theta'^d$$

on (part of) $\Theta \times \Theta$

Ratios of volumes

Yoke g on Θ gives ratio

$$\frac{|g_{i;j}(\theta, \theta')|}{\left|\not{g}_{i;j}(\theta)\right|^{1/2}\left|\not{g}_{i;j}(\theta')\right|^{1/2}}$$

of symplectic to Riemannian volumes

3. STATISTICAL APPLICATIONS

Applications

Concise formulae for

(a) information loss $\qquad t : \mathcal{X} \longrightarrow \mathcal{Y}$

$$i^X(\theta) - i^{t(X)}(\theta) = \text{functions of curvature}$$

(b) p^* formula for m.l.e.

$$\frac{p^*(\hat{\theta};\theta|a)}{|\not{g}_{i;j}(\hat{\theta})|^{1/2}} = c(\theta,a)\exp\{g(\theta,\hat{\theta})\}$$

g observed likelihood yoke

$$g(\theta,\theta') = l(\theta;\theta',a) - l(\theta';\theta',a)$$

Applications

(c) Bartlett correction of likelihood ratio test

$$w^* = \left(1 + \frac{B}{n}\right) w$$

B quadratic in $\not{T}_{rstu;}$, $\not{T}_{rs;tu}$ and $\not{T}_{r;st}$

(d) prediction limits

(e) modified profile likelihood, adjusted profile likelihood (ratios of volumes)

Applications

Sequential estimation

Geometric Wald tests

(natural coordinates $\overset{\alpha}{\Gamma}_\theta$)

Further developments

Parameters of interest ($\Theta \to \Omega$)

Infinite-dimensional/non-parametric (Pistone *et al.*)

Estimating functions (covectors)

Preferred point geometry
($\Theta \longrightarrow$ Riemannian metrics on Θ)

Decision theory (Dawid, Lauritzen, Parry)

Sensitivity analysis (Anaya, Critchley, Marriott, Vos)

Intrinsic van Trees theorem (information on the prior)

Quantum geometries

Quantum geometries

state

complex $d \times d$ matrix ρ with $\rho = \rho^*, \rho \geq 0, \mathrm{tr}\ \rho = 1$

$$\mathcal{S}(\mathbb{C}^d) = \{\rho : \rho = \rho^*, \rho \geq 0, \mathrm{tr}\ \rho = 1\}$$

quantum model $\quad \Theta \longrightarrow \mathcal{S}(\mathbb{C}^d)$

measurement : quantum model $\mapsto$ statistical model

Quantum score, quantum information,
quantum Cramér–Rao Theorem

4. THE FUTURE

What next?

That is up to us!