Curvatures of statistical structures

Barbara Opozda

Paris, October 2015

Statistical structures - statistical setting

M - open subset of $\mathbb{R}^n$

Λ - probability space with a fixed σ-algebra

$p : M \times \Lambda \ni (x, \lambda) \to p(x, \lambda) \in \mathbb{R}$ - smooth relative to x such that $p_x(\lambda) := p(x, \lambda)$ is a probability measure on Λ — probability distribution

$\ell(x, \lambda) := log(p(x, \lambda))$

$g_{ij}(x) := \mathbb{E}_x[(\partial_i \ell)(\partial_j \ell)]$, where $\mathbb{E}_x$ is the expectation relative to the probability p_x $\forall x \in M$, $\partial_1, ..., \partial_n$ - the canonical frame on M

g – **Fisher information metric tensor field on** M

$C_{ijk}(x) = \mathbb{E}_x[(\partial_i \ell)(\partial_j \ell)(\partial_k \ell)]$ - **cubic form**

(g, C) – **statistical structure on** M

Statistical structures (Codazzi structures)– geometric setting; three equivalent definitions

M – manifold, $\dim M = n$

I) (g, C), C - totally symmetric $(0,3)$-tensor field on M, that is,

$$C(X,Y,Z) = C(Y,X,Z) = C(Y,Z,X) \quad \forall X,Y,Z \in T_xM, \ x \in M$$

C – **cubic form**

II) (g, K), K – symmetric $(1,2)$-tensor field (i.e., $K(X,Y) = K(Y,X)$) and symmetric relative to g, that is,

$$g(X, K(Y,Z)) = g(Y, K(X,Z))$$

is symmetric for all arguments.

$$C(X,Y,Z) = g(X, K(Y,Z))$$

III) (g, ∇), ∇ - torsion-free connection such that

$$(\nabla_X g)(Y, Z) = (\nabla_Y g)(X, Z) \tag{1}$$

∇ — statistical connection

T – any tensor field of type (p, q) on M, ∇T – of type $(p, q+1)$

$$\nabla T(X, Y_1, ..., Y_q) = (\nabla_X T)(Y_1, ..., Y_q)$$

In particular, $\nabla g(X, Y, Z) = (\nabla_X g)(Y, Z)$
(1) $\Leftrightarrow$ ∇g is a symmetric cubic form
$\hat{\nabla}$ - Levi-Civita connection for g

$$K(X, Y) := \nabla_X Y - \hat{\nabla}_X Y$$

K – difference tensor

$$\nabla g(X, Y, Z) = -2g(X, K(Y, Z)) = -2C(X, Y, Z)$$

A statistical structure is **trivial** if and only if $K = 0$ or equivalently $C = 0$ or equivalently $\nabla = \hat{\nabla}$.

$K_X Y := K(X, Y)$

$E := \mathrm{tr}_g K = K(e_1, e_1) + ... + K(e_n, e_n) = (\mathrm{tr}\, K_{e_1})e_1 + ... + (\mathrm{tr}\, K_{e_n})e_n$

E – **mean difference vector field**

$$E = 0 \quad \Leftrightarrow \quad \mathrm{tr}\, K_X = 0 \quad \forall X \in TM \quad \Leftrightarrow \quad \mathrm{tr}_g C(X, \cdot, \cdot) = 0 \quad \forall X \in TM$$

$E = 0 \Rightarrow$ **trace-free statistical structure**

Fact. (g, ∇) – trace-free if and only if $\nabla \nu_g = 0$, where ν_g – volume form determined by g

Examples

Riemannian geometry of the second fundamental form

M – locally strongly hypersurface in $\mathbb{R}^{n+1}$
– the second fundamental form h satisfies the Codazzi equation

$$\nabla h(X, Y, Z) = \nabla h(Y, X, Z),$$

where ∇ is the induced connection (the Levi-Civita connection of the first fundamental form)
(h, ∇) - statistical structure

Similarly one gets statistical structures on hypersurfaces in space forms.

Equiaffine geometry of hypersurfaces in the standard affine space $\mathbb{R}^{n+1}$

M – locally strongly convex hypersurface in $\mathbb{R}^{n+1}$
ξ – a transversal vector field
D – standard flat connection on $\mathbb{R}^{n+1}$, $X, Y \in \mathcal{X}(M)$, ξ - transversal vector field

$$D_X Y = \nabla_X Y + h(X,Y)\xi \qquad - \textit{Gauss formula}$$

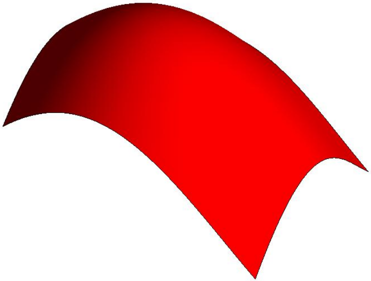
∇ – induced connection, h – second fundamental form (metric tensor field)

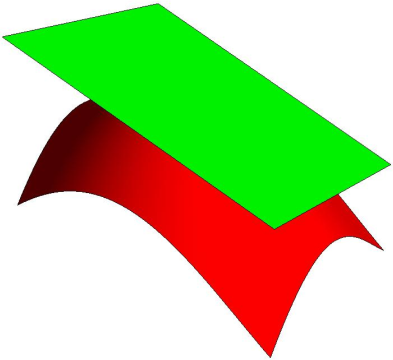
$$D_X \xi = -SX + \tau(X)\xi \qquad - \textit{Weingarten formula}$$

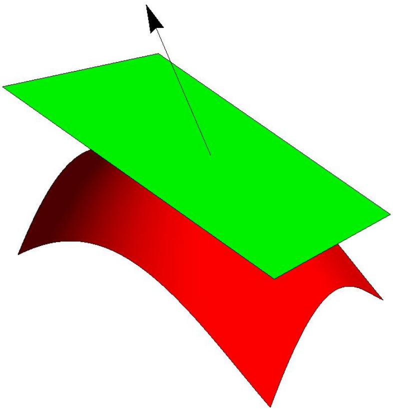
If $\tau = 0$, ξ is called equiaffine. In this case the Codazzi equation is satisfied

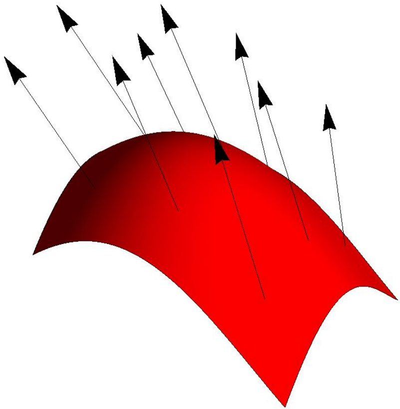
$$\nabla h(X,Y,Z) = \nabla h(Y,X,Z)$$

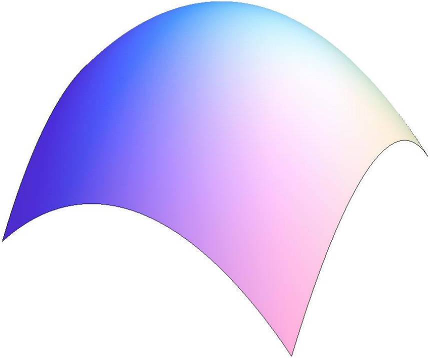
(h, ∇) – statistical structure

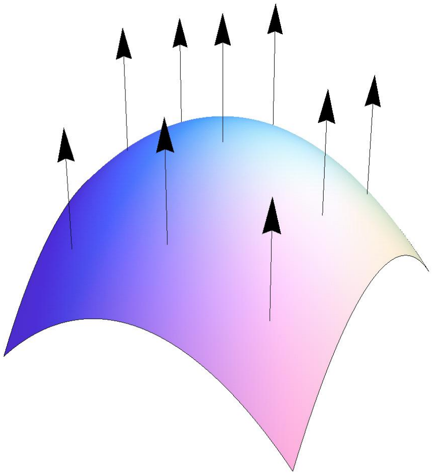


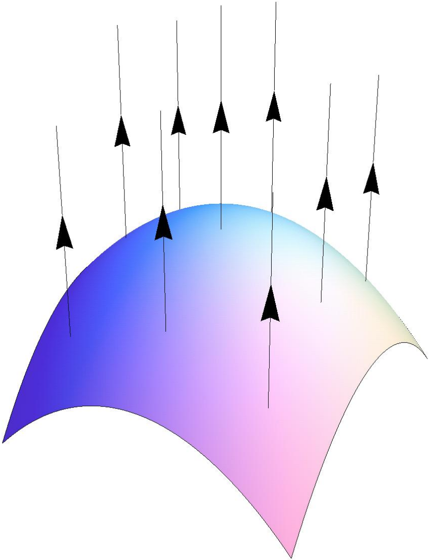


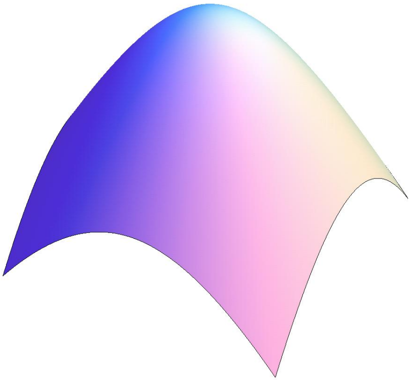


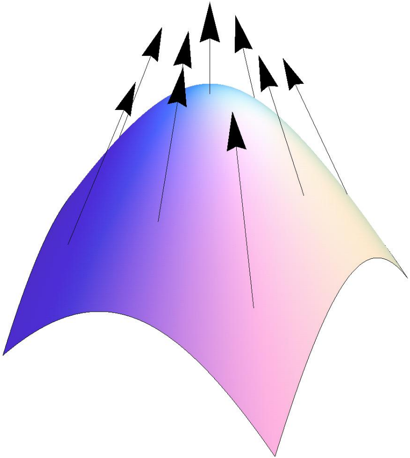


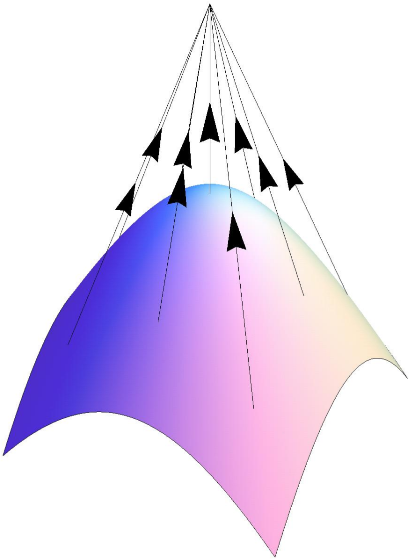












Geometry of Lagrangian submanifolds in Kaehler manifolds

N – Kaehler manifold of real dimension $2n$ and with complex structure J
M – Lagrangian submanifold of N - n-dimensional submanifold such that JTM orthogonal to TM, i.e. JTM is the normal bundle (in the metric sense) for $M \subset N$
D – the Kaehler connection on N

$$D_X Y = \nabla_X Y + JK(X, Y)$$

g – induced metric tensor field on M
(g, K) – statistical structure
It is trace-free $\Leftrightarrow$ M is minimal in N.

Most of statistical structures are outside the three classes of examples. For instance, in order that a statistical structure is locally realizable on an equiaffine hypersurface it is necessary that $\overline{\nabla}$ is projectively flat.

Dual connections, curvature tensors

g – metric tensor field on M, ∇ – any connection

$$Xg(Y,Z) = g(\nabla_X Y, Z) + g(Y, \overline{\nabla}_X Z) \tag{2}$$

$\overline{\nabla}$ – **dual connection**

(g, ∇) – statistical structure if and only if $(g, \overline{\nabla})$ – statistical structure

$R(X,Y)Z$ – $(1,3)$ - curvature tensor for ∇

If $R = 0$ the structure is called **Hessian**

$\overline{R}(X,Y)Z$ – curvature tensor for $\overline{\nabla}$

$$g(R(X,Y)Z, W) = -g(\overline{R}(X,Y)W, Z) \tag{3}$$

In particular, $R = 0 \Leftrightarrow \overline{R} = 0$.

$\hat{\nabla}$ – Levi-Civita connection for g,
$\nabla = \hat{\nabla} + K$, $\overline{\nabla} = \hat{\nabla} - K$
$\hat{R}$ – curvature tensor for $\hat{\nabla}$

$$R(X,Y) = \hat{R}(X,Y) + \quad (\hat{\nabla}_X K)_Y - (\hat{\nabla}_Y K)_X \quad + [K_X, K_Y] \tag{4}$$

,

where

$$[K_X, K_Y] = K_X K_Y - K_Y K_X$$

$$\overline{R}(X,Y) = \hat{R}(X,Y) - \quad (\hat{\nabla}_X K)_Y + (\hat{\nabla}_Y K)_X \quad + [K_X, K_Y] \tag{5}$$

$$\overline{R}(X,Y) + R(X,Y) = 2\hat{R}(X,Y) + 2[K_X, K_Y] \tag{6}$$

Sectional curvatures

R does not have to be skew-symmetric relative to g, i.e. $g(R(X,Y)Z,W) \neq -g(R(X,Y)W,Z)$, in general.

Lemma *

The following conditions are equivalent:

1) $g(R(X,Y)Z,W) = -g(R(X,Y)W,Z) \quad \forall X,Y,Z,W$
2) $R = \overline{R}$
3) $\hat{\nabla} K$ is symmetric, that is,

$$(\hat{\nabla} K)(X,Y,Z) = (\hat{\nabla}_X K)(Y,Z) = (\hat{\nabla}_Y K)(X,Z) = (\hat{\nabla} K)(Y,X,Z)$$

$\forall X,Y,Z$.

For hypersurfaces in $\mathbb{R}^{n+1}$ each of the above conditions describes an affine sphere

$\mathcal{R} := \frac{R+\overline{R}}{2}$

$[K,K](X,Y)Z := [K_X, K_Y]Z$

$\mathcal{R}(X,Y)Z$ and $[K,K](X,Y)Z$ are Riemann-curvature-like tensors – they are skew-symmetric in X, Y, satisfy the first Bianchi identity,
$R(X,Y)$, $[K,K](X,Y)$ are skew-symmetric relative to g $\forall X, Y$
π – vector plane in T_xM, X, Y – orthonormal basis of π

sectional curvature for g – $\hat{k}(\pi) := g(\hat{R}(X,Y)Y,X)$

sectional K-curvature – $k(\pi) := g([K,K](X,Y)Y,X)$

sectional ∇-curvature – $k^{\nabla}(\pi) := g(\mathcal{R}(X,Y)Y,X)$

In general, Schur's lemma does not hold for k^{∇} and k. We have, however,

Lemma

*Assume that M is connected, $\dim M > 2$ and the sectional ∇- curvature (the sectional K-curvature) is point-wise constant. If one of the equivalent conditions in Lemma * holds then the sectional ∇-curvature (the sectional K-curvature) is constant on M.*

sectional K-curvature

The easiest situation which should be taken into account is when the sectional K-curvature is constant for all vector planes in $T_x M$. In this respect we have

Theorem

If the sectional K-curvature is constant and equal to A for all vector planes in T_xM then there is an orthonormal basis $e_1, ..., e_n$ of T_xM and numbers $\lambda_1, ..., \lambda_n$, $\mu_1, ..., \mu_{n-1}$ such that

$$K_{e_1} = \begin{bmatrix} \lambda_1 & & & \\ & \mu_1 & & \\ & & \ddots & \\ & & & \mu_1 \end{bmatrix} \quad K_{e_i} = \begin{bmatrix} & & & \mu_1 & & & \\ & & & \vdots & & & \\ & & & \mu_{i-1} & & & \\ \mu_1 & \cdots & \mu_{i-1} & \lambda_i & & & \\ & & & & \mu_i & & \\ & & & & & \ddots & \\ & & & & & & \mu_i \end{bmatrix}$$

$$K_{e_n} = \begin{bmatrix} & & & \mu_1 \\ & & & \vdots \\ & & & \mu_{n-1} \\ \mu_1 & \cdots & \mu_{n-1} & \lambda_n \end{bmatrix}$$

continuation of the theorem

Moreover

$$\mu_i = \frac{\lambda_i - \sqrt{\lambda_i^2 - 4A_{i-1}}}{2},$$

$$A_i = A_{i-1} - \mu_i^2,$$

for $i = 1, ..., n-1$ where $A_0 = A$. The above representation of K is not unique, in general. If additionally $\mathrm{tr}_g K = 0$ then $A \leqslant 0$, $\lambda_n = 0$ and λ_i, μ_i for $i = 1, ..., n-1$ are expressed as follows

$$\lambda_i = (n-i)\sqrt{\frac{-A_{i-1}}{n-i+1}}, \quad \mu_i = -\sqrt{\frac{-A_{i-1}}{n-i+1}}.$$

In particular, in the last case the numbers λ_i, μ_i depend only on A and the dimension of M.

Example 1.

$$K_{e_1} = \begin{bmatrix} \lambda & & & \\ & \lambda/2 & & \\ & & \ddots & \\ & & & \lambda/2 \end{bmatrix} \quad K_{e_i} = \begin{bmatrix} & & & \lambda/2 & & & \\ & & & \vdots & & & \\ & & & 0 & & & \\ \lambda/2 & \cdots & 0 & 0 & & & \\ & & & & 0 & & \\ & & & & & \ddots & \\ & & & & & & 0 \end{bmatrix}$$

$$K_{e_n} = \begin{bmatrix} & & & \lambda/2 \\ & & & \vdots \\ & & & 0 \\ \lambda/2 & \cdots & 0 & 0 \end{bmatrix}$$

The sectional K-curvature is constant $= \lambda^2/4$

Example 2.

K-curvature vanishes, i.e. $[K, K] = 0$. There is an orthonormal frame $e_1, ..., e_1$ such that

$$K_{e_1} = \begin{bmatrix} \lambda_1 & & & \\ & 0 & & \\ & & \ddots & \\ & & & 0 \end{bmatrix} \quad K_{e_i} = \begin{bmatrix} & & & 0 & & & \\ & & & \vdots & & & \\ & & & 0 & & & \\ 0 & \cdots & 0 & \lambda_i & & & \\ & & & & 0 & & \\ & & & & & \ddots & \\ & & & & & & 0 \end{bmatrix}$$

$$K_{e_n} = \begin{bmatrix} & & & 0 \\ & & & \vdots \\ & & & 0 \\ 0 & \cdots & 0 & \lambda_n \end{bmatrix}$$

Some theorems on the sectional K-curvature

(g, K) – trace-free if $E = \mathrm{tr}_g K = 0$

Theorem

Let (g, K) be a trace-free statistical structure on M with symmetric $\hat{\nabla} K$. If the sectional K-curvature is constant then either $K = 0$ (the statistical structure is trivial) or $\hat{R} = 0$ and $\hat{\nabla} K = 0$.

Theorem

Let $\hat{\nabla} K = 0$. Each of the following conditions implies that $\hat{R} = 0$:
1) the sectional K-curvature is negative,
2) [K,K]=0 and K is non-degenerate, i.e. $X \to K_X$ is a monomorphism.

Theorem

K is as in Example 1. at each point of M, $\hat{\nabla}K$ is symmetric, $\operatorname{div} E$ is constant on M ($E = \operatorname{tr}_g K$). Then the sectional curvature for g by any plane containing E is non-positive. Moreover, if M is connected it is constant. If $\hat{\nabla}E = 0$ then $\hat{\nabla}K = 0$ and the sectional curvature (of g) by any plane containing E vanishes.

Theorem

If the sectional K-curvature is non-positive on M and $[K, K] \cdot K = 0$ then the sectional K-curvature vanishes on M.

Corollary

If (g, K) is a Hessian structure on M with non-negative sectional curvature of g and such that $\hat{R} \cdot K = 0$ then $\hat{R} = 0$.

Theorem

The sectional K-curvature is negative on M, $\hat{R} \cdot K = 0$. Then $\hat{R} = 0$.

Theorem

Let M be a Lagrangian submanifold of N, where N is a Kaehler manifold of constant holomorphic curvature $4c$, the sectional curvature of the first fundamental form g on M is smaller than c on M and $\hat{R} \cdot K = 0$, where K is the second fundamental tensor of $M \subset N$. Then $\hat{R} = 0$.

∇-sectional curvature

All affine spheres are statistical manifolds of constant sectional ∇-curvature

A Riemann curvature-like tensor defines the curvature operator. For instance, for the curvature tensor $\mathcal{R} = (R + \overline{R})/2$ we have the curvature operator $\mathfrak{R} : \Lambda^2 TM \to \Lambda^2 TM$ given by

$$g(\mathfrak{R}(X \wedge Y), Z \wedge W) = g(\mathcal{R}(Z, W)Y, X)$$

A curvature operator is symmetric relative to the canonical extension of g to the bundle $\Lambda^2 TM$. Hence it is diagonalizable. In particular, it can be positive definite, negative definite etc.
The assumption that $\mathfrak{R}$ is positive definite is stronger than the assumption that the sectional ∇-curvature is positive.

Theorem

Let M be a connected compact oriented manifold and (g, ∇) be a trace-free statistical structure on M. If $R = \overline{R}$ and the curvature operator determined by the curvature tensor $\hat{R}$ is positive definite on M then the sectional ∇-curvature is constant.

Theorem

Let M be a connected compact oriented manifold and (g, ∇) be a trace-free statistical structure on M. If the curvature operator for $\mathcal{R} = \frac{R+\overline{R}}{2}$ is positive on M then the Betti numbers $b_1(M) = \ldots = b_{n-1}(M) = 0$.

sectional curvature for g

$\hat{k}(\pi) = g(\hat{R}(X,Y)Y,X)$, X, Y – an orthonormal basis for π

Theorem

Let M be a compact manifold equipped with a trace-free statistical structure (g,∇) such that $R = \overline{R}$. If the sectional curvature $\hat{k}$ for g is positive on M then the structure is trivial, that is $\nabla = \hat{\nabla}$.

In the 2-dimensional case we have

Theorem

Let M be a compact surface equipped with a trace-free statistical structure (g,∇). If M is of genus 0 and $R = \overline{R}$ then the structure is trivial.

B. Opozda, *Bochner's technique for statistical manifolds*, Annals of Global Analysis and Geometry, DOI 10.1007/s10455-015-9475-z

B. Opozda, *A sectional curvature for statistical structures*, arXiv:1504.01279[math.DG]

Hessian structures

(g, ∇) – Hessian if $R = 0$. Then $\overline{R} = 0$ and $\hat{R} = -[K, K]$.
(g, ∇) is Hessian if and only if $\hat{\nabla} K$ is symmetric and $\hat{R} = -[K, K]$.
All Hessian structure are locally realizable on affine hypersurfaces in $\mathbb{R}^{n+1}$ equipped with Calabi's structure. If they are trace-free they are locally realizable on improper affine spheres.
If the difference tensor is as in Example 1. and the structure is Hessian then $K = 0$.