Ezra 1

Ezra Miller : stats with samples on stratified spaces

stat methods : mean, variance, PCA, CLN, CLT, ...

Def : manifold = 2^{nd} countable, Hausdorff topological space s.t. every pt has a neighborhood homeomorphic to an open ball in $\mathbb{R}^n$.

(2^{nd} countable = M is covered by a finite # balls)

chart : $\pi : U \in \mathbb{R}^n \xrightarrow[\text{injection}]{} M$.

$M_{\alpha\beta}$ transition for chart α & β

Category : topological manifold
smooth / C^∞ "
analytic "
algebraic variety

Def : A topologically stratified space in a Hausdorff topol space X that is a disjoint union of manifolds $M_1 \cup M_2 \cup \ldots M_\ell$ s.t.

1/ $M_1 \cup \ldots \cup M_k$ is a closed subset $\forall k \le \ell$

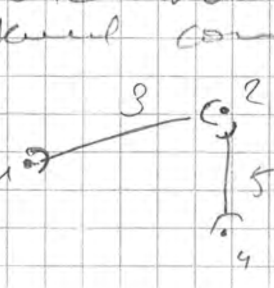
2/ $\forall x, y \in M_i$ (stratum), there exist a homeomorphism $\varphi : X \to X$ with $\varphi(M_k) = M_k \ \forall k$ and $\varphi(x) = y$
stratum preserving (transitive action

3/ Each stratum has finitely many connected components.

Examples

dim 0 : finite discrete set of pts (each pt is closed).
stratification : as one disconnected manifold or as several connected.

dim 1 : graphs. 2 pts + one edge,



or trees -

Trees = connected graphs with no cycles. 2

Spider or stars :

Dim 2+ :

open book = spider × $\mathbb{R}^d$

spine ∪ open pages - no

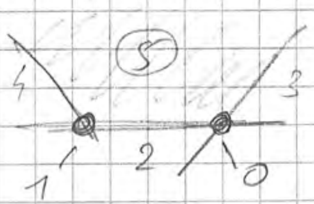
GM example

dim 0: 1

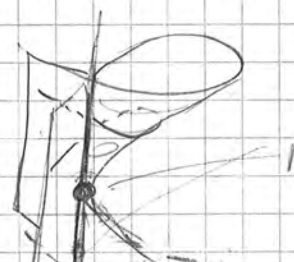
dim 1: open circle.

dim 2: torus + disk

Polyhedron : intersection of finitely many closed half spaces in $\mathbb{R}^d$.



Whitney cusp



dim 0: 1 $\quad$ Sing(X) = |

not equisingular

dim 1: 2 (or 1 disconnected)

dim 2: 4 (or less if disconnected)

Shape spaces $\quad$ A = d×n matrix (n labeled landmarks in $\mathbb{R}^d$).

$A \sim gA \quad g \in$ Transf group G.

$\Rightarrow$ G = SE(d), similarities, projective

$X = G \backslash \mathbb{R}^d$

mod out the left = algebraic variety

Ezra (3)

⚠ Remove unstable orbits to avoid having non-Hausdorff quotients. (~~free~~ action is needed).

Ex Phylogenetic tree is a tree with labeled leaves (vertices of deg 1) an edge length >0

2 positive edges $\mathbb{R}^2_{>0}$

eg

removing an edge creates 2 disconnected comp: an edge can be identified to the set of leaves on both sides.

internal edges: $\mathbb{R}^2_{>0}$

→ a tree is a union of positive orthants.

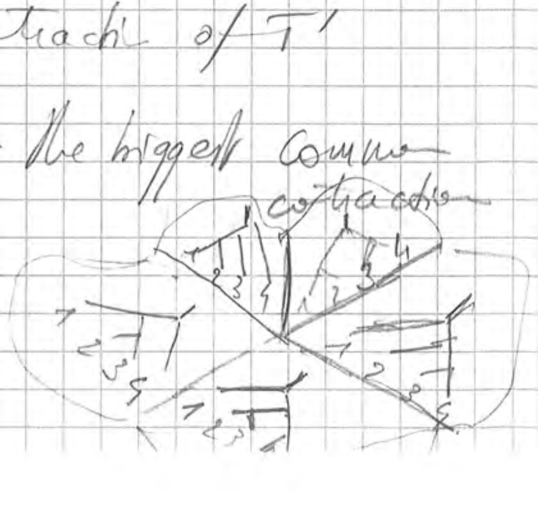
[Billera - Holmes - Vogtmann 1998].

- facets of T_n are orthants $\mathbb{R}^{n-2}_{>0}$ (binary tree topologies).
- faces are orthants $\mathbb{R}^{d}_{>0}$, $d \le n-2$

tree $T \longleftrightarrow x_T \in \mathbb{R}^{edge(T)}_{>0}$ ← splits $=: O_T$

$O_T \subseteq \overline{O_{T'}} \Leftrightarrow T$ is a contraction of T'

$\overline{O_T} = \overline{O_{T'}} \cap \overline{O_{T''}} \Leftrightarrow T$ is the biggest common contraction

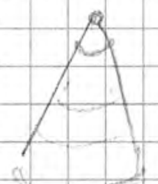


Epa 4.

PSD matrices.

stratification by subsets of equal eigenvalues.

Cone:



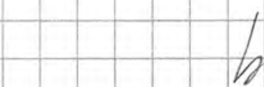
homeomorphic to $\mathbb{R}^2$

and locally isometric to except at the apex.

Volume of the sphere at the vertex is $2\pi r$ in the plane

but only $\alpha r < 2\pi r$ at the vertex of the cone

$\alpha < 2\pi$: positive curvature.



circle length $5\pi/2 \cdot r$

$\alpha > 2\pi$ negatively curved (kale).

$\text{Kale}_{\alpha = 5\pi/2} \subset T_4$

Thm: Tubular neighborhood thm.

X stratified $\Longrightarrow$ a tubular neighborhood N of each stratum S is a fiber bundle over S whose fiber N_S is homeomorphic to a cone over a stratified space L with $\dim L = \dim N - 1$.

$N_p X$

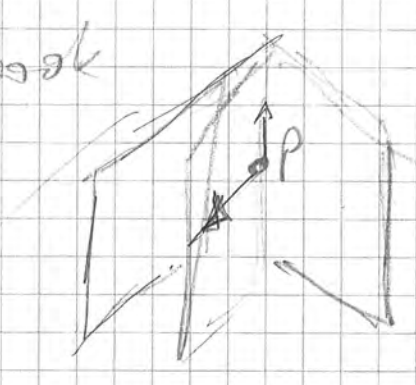
Tg cone $T_p X = T_p S \times N_p X$

$T_p S$: tg space of S

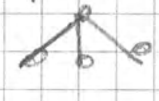
$N_p X$: Normal slice to X, i.e. going to another stratum.

Ora 5

open book



$T_pX = \mathbb{R} \times$ spider



<u>Def</u>: $\dim(X) = \max_\alpha (\dim \Pi_\alpha)$.

for $p \in S$ stratum of dim $d-1$, where $d = \dim X$.

$T_pX = \underbrace{T_pS}_{\mathbb{R}^{d-1}} \times \underbrace{N_p\cancel{X}}_{\text{spider, dim}=1}$ with N_pX = cone over the finite set = "spider"

<u>Thm</u>: if Π_i and Π_j are strata then $\Pi_i \cap \overline{\Pi_j} \neq \emptyset \Rightarrow \Pi_i \subseteq \overline{\Pi_j}$

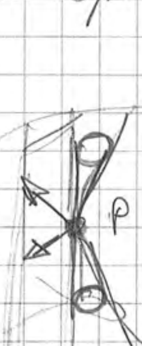
<u>Correlation</u>: the strata are partially ordered.

$$\Pi_i \leq \Pi_j \Leftrightarrow \Pi_i \subseteq \overline{\Pi_j}$$

<u>Rem</u>: the open book is universal in the sense that every co-dim ~~one~~ singularity of a stratified space is an open book.

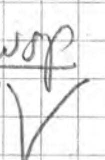
$\dim(S) = 2 - 1 = 1$.

eg for the Whitney cusp

If discontinuous  $p \Rightarrow T_pS = 0$

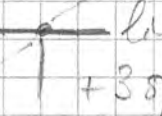
$N_pX = T_pX =$ 

<u>If cusp</u>

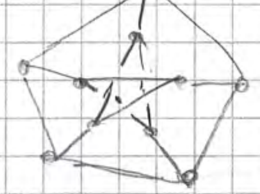


 $\longrightarrow N_pX =$ 

tangent space for the tree space T_4

$T_p X$ at a line = 

$T_0 \overline{T_4} = N_0 T_4$ = cone over petersen graph



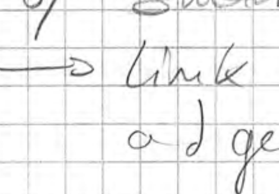
Metrics, smoothness, etc

a Riemannian stratified space should require at least:
a topologically stratified metric space such that
each stratum is Riemannian.

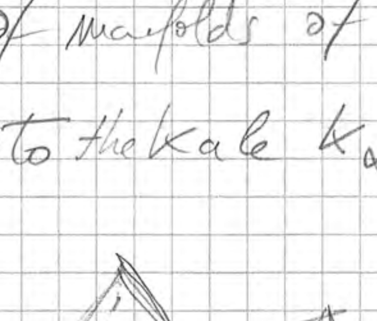
geodesic metric

Sheaf of smooth functions: [book by Pflaum].
→ link with the manifold with corners
and gems of manifolds of Peter Michor.

Exercise: embed T_3 into the kale K_α for $\alpha > 3\pi$



discrete hyperbolic saddle



Thm: Whitney stratified spaces (embedded in an Eucl sp)
a topologically stratified space is a WSS if
limits of secant lines joining strata R_i and R_j
with $R_i \subseteq \overline{R_j}$ is included in the limit of tg planes
to R_j as points cv to R_i.

Examples: real of complex algebraic
(or semi-analytic, or sub analytic)

Thm: Whitney stratified ⟹ triangulable

X is a metric stratified space.

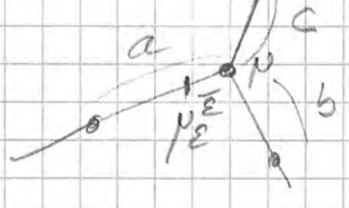
Def: A probability distrib p on X has a frechet function $F(y) = \int_X d^2(y,x)\, p(dx)$

The frechet mean of p is $\arg\min_{y \in X} F(y)$.

What can go wrong:

wiggle point ⟹ ~~wiggle mean~~
the mean can stick

Example: 3 spider



$a = b = c = 1 \Rightarrow$ mean is 0.

$$F(\mu) = a^2 + b^2 + c^2$$

$$F(\mu_\varepsilon) = (a-\varepsilon)^2 + (c+\varepsilon)^2 + (b+\varepsilon)^2$$

$$= a^2 + b^2 + c^2 + 2(c+b-a)\varepsilon + 3\varepsilon^2$$

$$F(\mu_\varepsilon) - F(\mu) = 2\varepsilon(c+b-a) + 3\varepsilon^2.$$

positive for $a < b+c$ for small ε.

Thus the means sticks at 0 for small ε.

On a general stratum, the mean wiggles along the Tg space but may be sticky along the Npt

Epra ⑧

LLN! $x_1, x_2, \dots$ indep random variables with value in X distrib according to p

then $\overline{x_n} \longrightarrow \mu(p)$. as $n \to +\infty$,

fix X = open-book = $S \times$ spider, where $S = \mathbb{R}^d$

given a proba p, there are 3 possibilities:

1/ $\mu(p) \notin S$ (it is in a page).

2/ $\mu(p) \in S$ and the same is true for $\mu(p')$ for ~~all~~ p' near p.

3/ $\mu(p) \in S$ but $\exists p'$ arbitrarily close to p such that $\mu(p') \notin S$

Rem the open-book X is CAT(0) so the Frechet mean is unique.

Def for these three cases:

1/ mean is non-sticky

2/ mean is ~~non~~ sticky

3/ mean is partly sticky.

(the definition implies that there exist directions in which it stays sticky).

Thm [BATTS WG].

If x_1, x_2 are iid in X distrib according to p with supp(p) on at least 3 pages.

then $\bar{x}_n \underset{n\to+\infty}{\longrightarrow} \mu(p)$ (std LLN)

Moreover: (•) If $\mu(p)$ is non-sticky or partly sticky (case 1 or 3), then there is a page L and a random integer N s.t. $\bar{x}_n \in L$ $\forall n \geq N$ almost surely

(•) $\mu(p)$ sticky $\Rightarrow$ $\exists$ N random such that $\bar{x}_n \in S$ $\forall n \geq N$ a.s.

CLT: $x_1 \dots x_n$ iid $\Rightarrow$ $\bar{x}_n$ is a random variable

$\bar{x}_n \to \delta_{\mu(p)}$ and $\sqrt{n}(\bar{x}_n - \mu(p)) \to \tilde{p}$,

where $\tilde{p}$ is a limiting distribution (classically Gaussian).

Thm [BATTS WG] A CLT for square integrable p holds with limiting distrib:

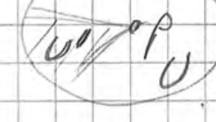
1/ non-sticky mean: Gaussian on $\mathbb{R}^{d+1} \simeq T_{\mu(p)}L = T_{\mu(p)}X$

2/ sticky case: Gaussian on $\mathbb{R}^d = S = T_p S$

3/ partly sticky: Gaussian on L + its reflective projection to S

Ezra (10)

<u>Def</u> Fix a metric space X and a topologized set of $\mathcal{P}$ of proba distrib on X. The mean $\mu(p)$ of $p \in \mathcal{P}$ <u>sticks</u> to a subset $K \subseteq X$ if every neighborhood U of $p \in \mathcal{P}$ contains a non-empty open set $U' \subseteq U$ with $\mu(p') \subseteq K$ for all $p' \in U'$



[partly sticky + sticky case].

<u>Example</u> [Huckemann, Mattingly, Nolan, Miller]

the kale (isolated planar hyperbolic singularity)

Kale with $\alpha > 2\pi$

projection on the sectors.

the shadows of the kale