Equivariant Topology and Regression Depth

Marshall Bern
Xerox PARC

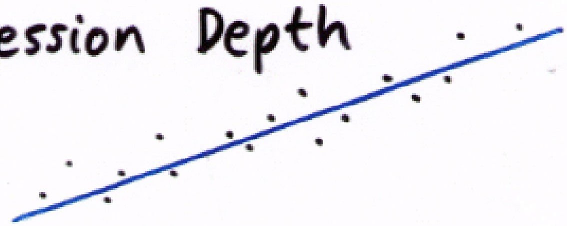
Outline

1. What is Equivariant Topology?

2. Survey of Applications in Computational Geometry

   [For Combinatorics, see Björner or Lovász]

3. Regression Depth



[Živaljević User's Guide to Equivariant Methods in Combinatorics]

[Matoušek Topological Methods in Combinatorics and Geometry]

1. What is Equivariant Topology?

Topological Spaces X, Y

Group G acting on X, Y

$g \in G$ $\quad \phi_g : X \to X$ homeomorphism

ϕ_1 – identity

$\phi_{g \cdot h} = \phi_g \circ \phi_h$

Notation: $\phi_g(x) = g \cdot x$

Does there exist a cont.

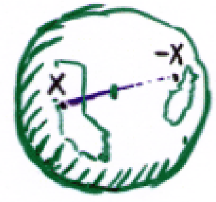
$f : X \to Y$ that is <u>G-equivariant</u>?

$$f(g \cdot x) = g \cdot f(x) \quad \forall g \in G \ \forall x \in X$$

Example: Borsuk-Ulam Theorem

Cont. $f: S^d \to \mathbb{R}^d$

$\exists\ x$ s.t. $f(x) = f(-x)$



Equivariant Version

$G = \mathbb{Z}_2 = \{1, -1\}$ $\quad X = S^d \quad Y = S^{d-1}$

G-equivariant $\equiv$ Antipodal

$\forall x \quad f(x) = -f(-x)$

$\nexists$ cont. G-equivariant

$f: S^d \to S^{d-1}$

$\neg$ Equiv. Version $\Rightarrow$ $\neg$ Borsuk-Ulam

Assume cont., antipodal $f: S^d \to S^{d-1}$

Then $f: S^d \to \mathbb{R}^d$ is cont., antipodal and nowhere 0. $\forall x \;\; f(x) = -f(-x)$

But then $\nexists$ x s.t. $f(x) = f(-x)$

$\neg$ Borsuk-Ulam $\Rightarrow$ $\neg$ Equiv. Version

Assume cont. $f: S^d \to \mathbb{R}^d$ s.t. $\forall x \;\; f(x) \neq f(-x)$

Then let $g(x) = f(x) - f(-x)$

g is cont., antipodal, & nowhere 0.

Let $h(x) = \dfrac{g(x)}{\|g(x)\|}$

$h: S^d \to S^{d-1}$ cont., antipodal

- Changing orientations of planes
- Dir'ns of oriented planes

- ?

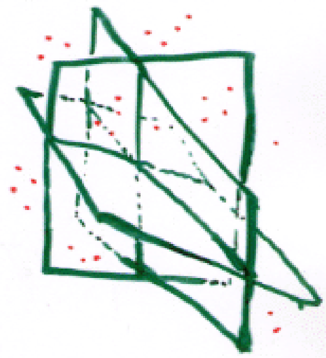
Vector of $\overset{\text{excess}}{\text{measures}}$ $(M_1-\frac{1}{8},\dots,M_8-\frac{1}{8})\in\mathbb{R}^8$

To show $f: X\to\mathbb{R}^8$ hits 0,
show non-existence of $f: X\to S^7$

Example: Equipartition in $\mathbb{R}^3$

[F. Yao - Dobkin - Edelsbrunner]

Any measurable set in $\mathbb{R}^3$ can be split into 8 equal sets by 3 planes.



Equivariant Version

$G = \mathbb{Z}_2 \times \mathbb{Z}_2 \times \mathbb{Z}_2$

$X = S^2 \times S^2 \times S^2$

$Y = S^7$

$\nexists$ cont. G-equiv $f: X \to Y$

Why the Equivariant Versions?

More general results!

[Dold] X, Y simplicial complexes

Group action is free on X, Y

$g \cdot x = x \Rightarrow g = 1$

X is k-connected

and $\dim(Y) \leq k$.

$\Rightarrow \nexists$ G-equiv $f: X \to Y$

[Fadell-Husseini]

[Sarkaria] Index Theory measures

"G-complexity" of X, Y

Caveats

- 90% of Geometric applications seem to use Borsuk-Ulam

- General results may not resolve the question.

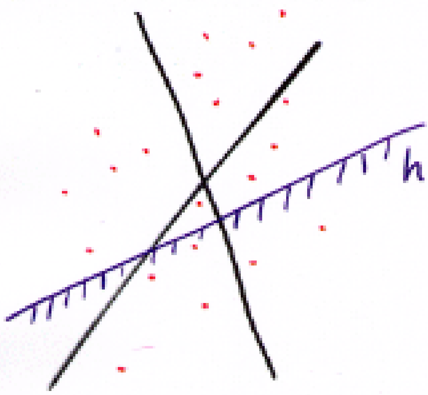
Open: Can any measurable set in $\mathbb{R}^4$ be equipartitioned into 16 by 4 hyperplanes?

$G = (\mathbb{Z}_2)^4 \quad X = (S^2)^4 \quad Y = S^{15}$

$\exists$ G-equiv $f: X \to Y$?

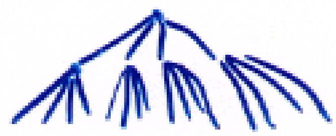
2. Survey of ATMCG

- Data Structures using Equipartitions

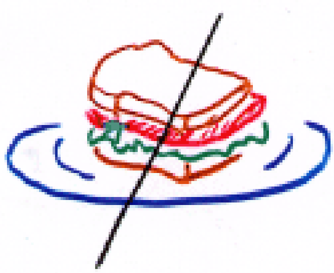


[Willard 82]

Half-space counting query:
How many pts in h?



$$O(3^{\log_4 n}) = O(n^{\log_4 3})$$



Ham Sandwich Thm

Equipartitions themselves

[Ramos96]

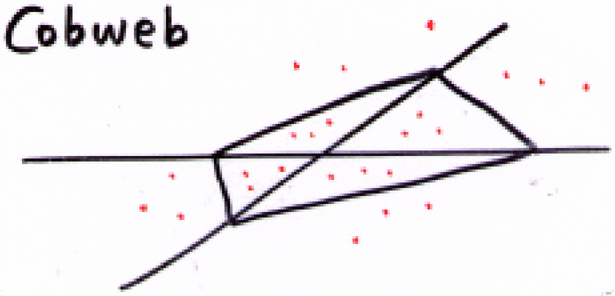
$\mathbb{R}^d$ - Divide j sets by k hyperplanes

$$\begin{array}{ccc} d & j & k \end{array}$$

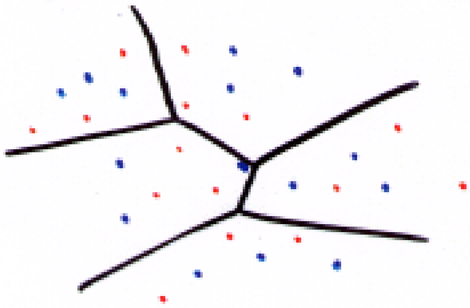
(9, 3, 3) ✓ (9, 5, 2) ✓ (5, 1, 4) ✓

(4, 1, 4) ?

[Schulman93] Cobweb

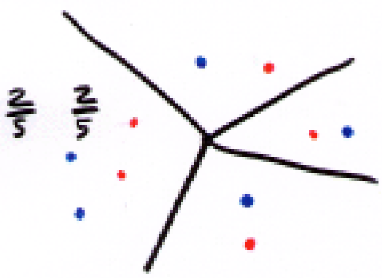


More Equipartitions



[Bespamyatnikh et al 98]
Sakai 98

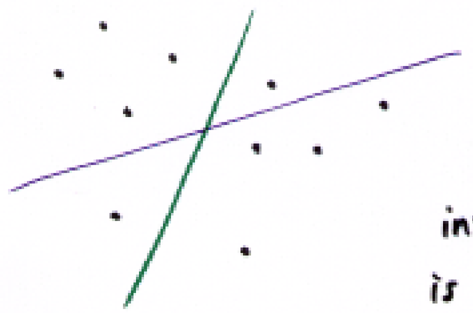
Red-Blue
Equitable Partition



(Bárány-Matoušek 99)

Equipartitions
by K-Fans

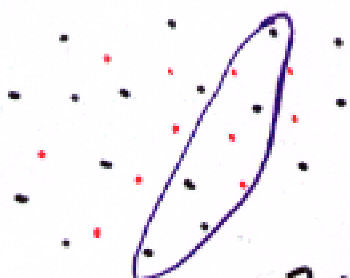
Halving Hyperplanes



[Vrećica · Živaljević]

diff't ways to divide n pts in $\mathbb{R}^d$ into equal halves is $O(n^{d-\delta_d})$.

Weak ε-Nets



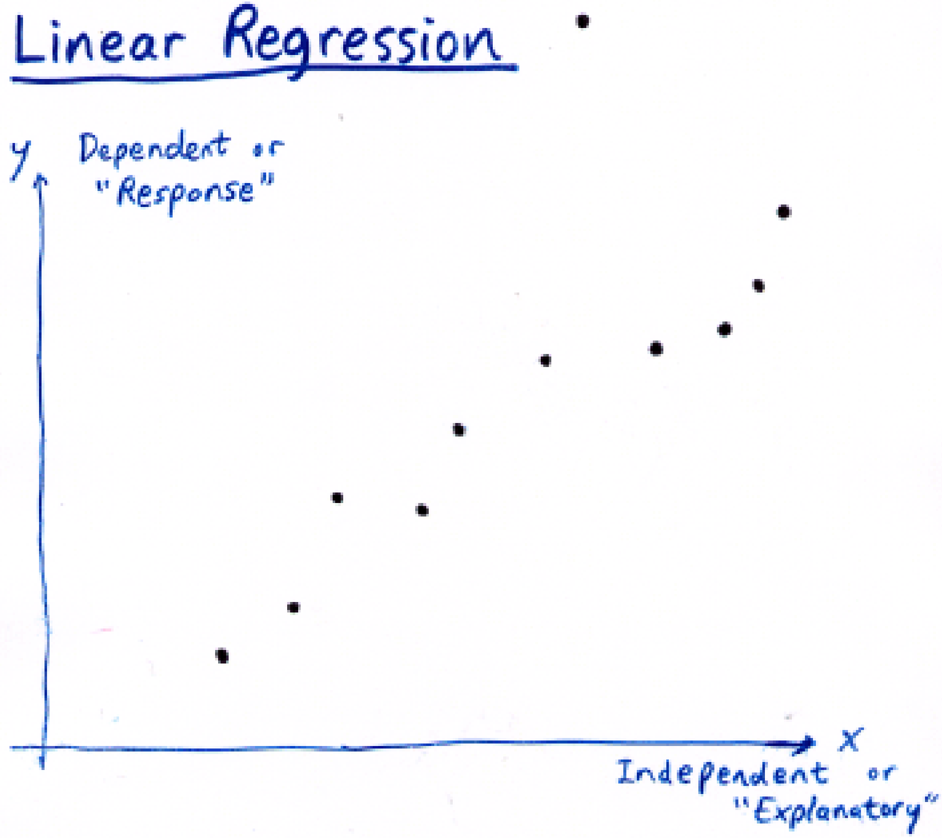
Any convex set containing $\varepsilon \cdot n$ pts of X contains a pt of W

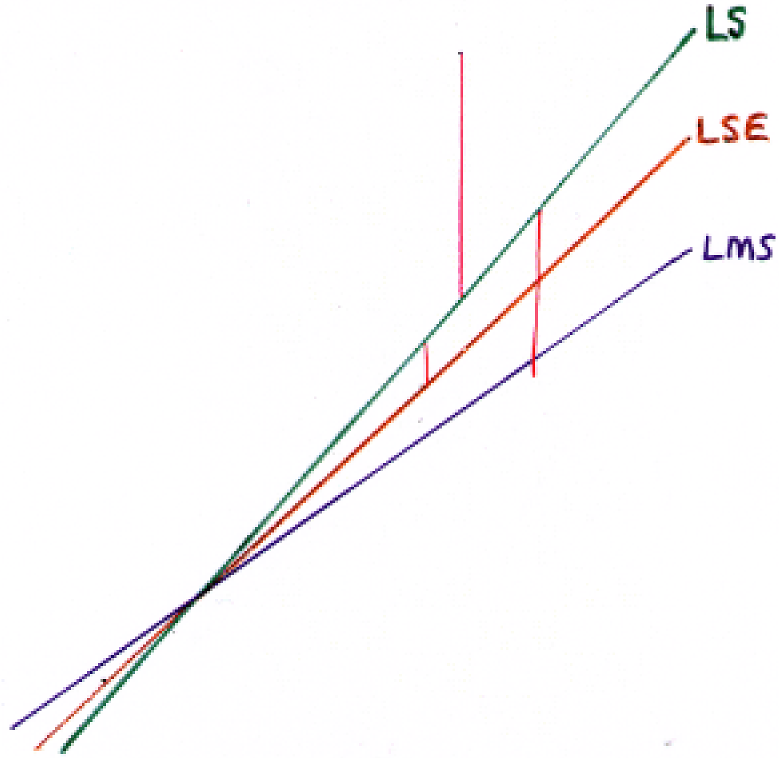
[Alon et al]

$\exists$ W of size $O\left(\left(\frac{1}{\varepsilon}\right)^{d+1-\delta'_d}\right)$

3. Regression Depth

Linear Regression



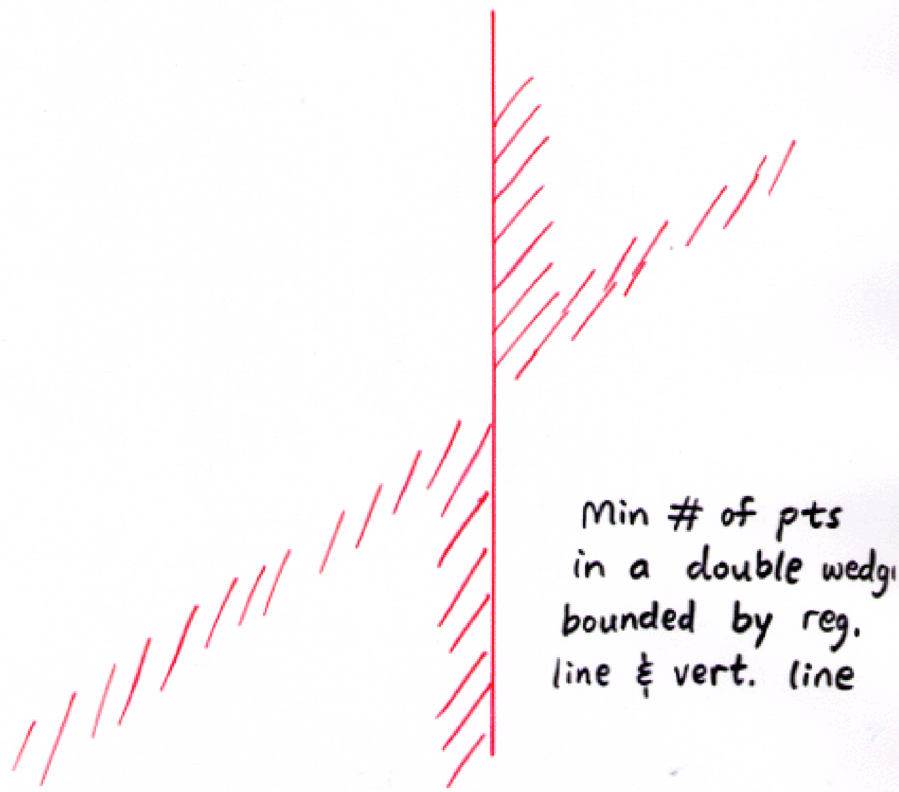


Least Squares $\sum (y_i - \hat{y}_i)^2$

Least Sum of Errors $\sum |y_i - \hat{y}_i|$

Least Median of Squares $\text{med} \{ (y_i - \hat{y}_i)^2 \}$

[Rousseeuw 84]



Deepest Line Regression

- Affine Invariant, Non-metric, No counter-intuitive behavior

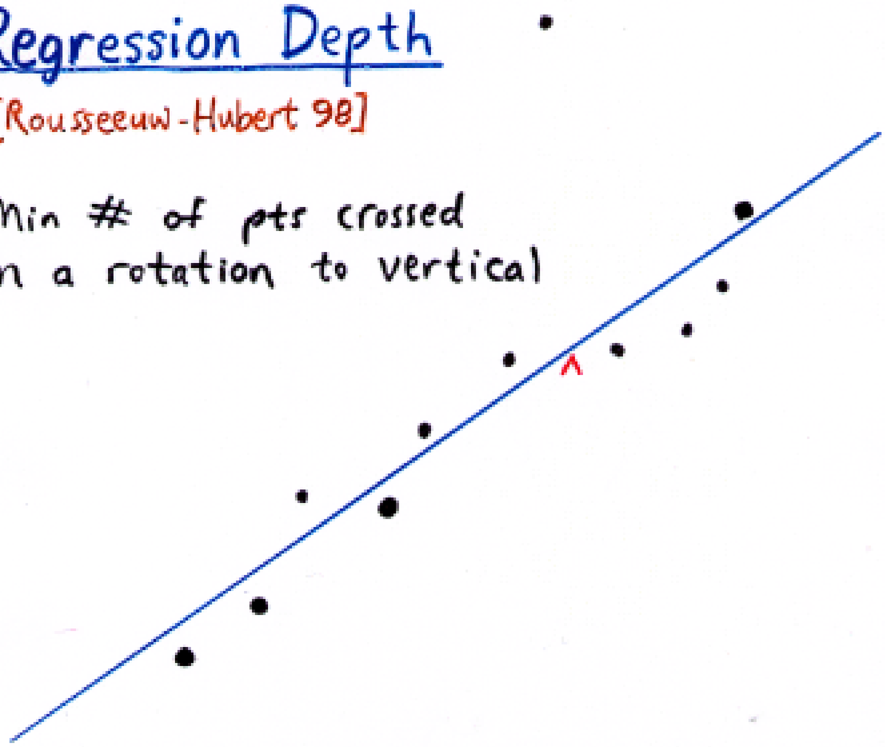
Reg. Depth as a Checker

- "Breakdown Point" of a regression line

Regression Depth

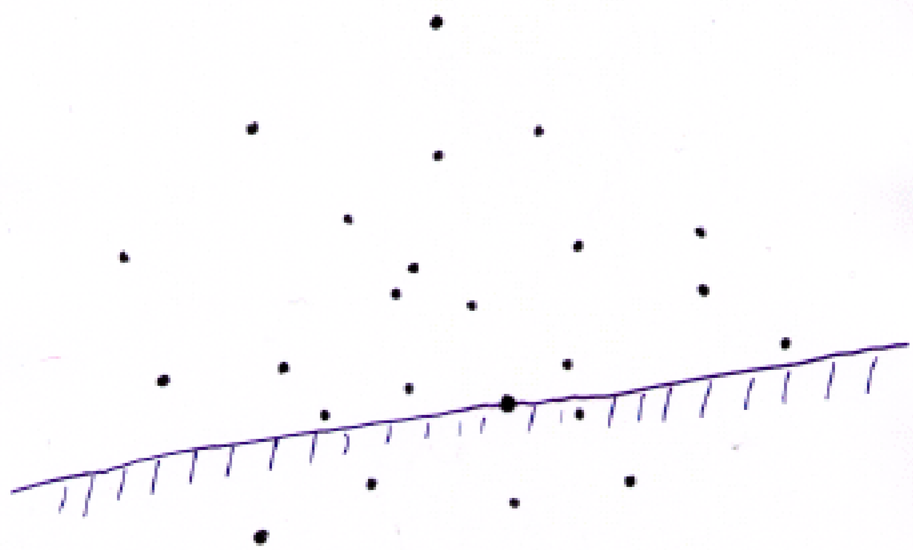
[Rousseeuw-Hubert 98]

Min # of pts crossed
in a rotation to vertical



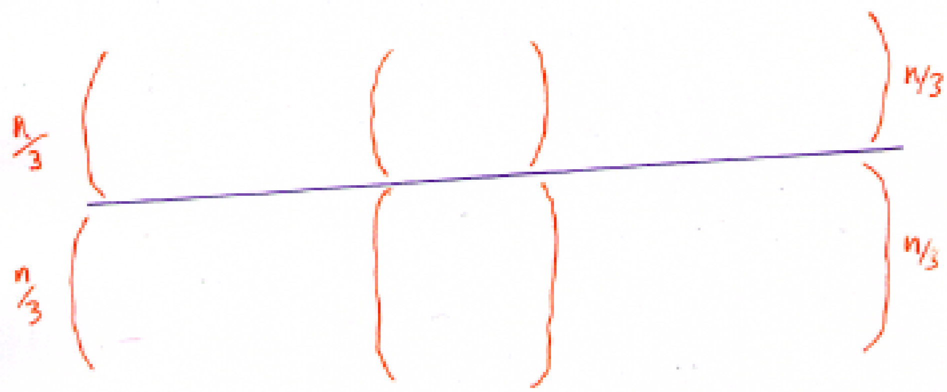
Cf. Data Depth [Tukey, Singh, etc.]

Min # pts in any half-space

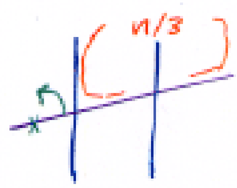


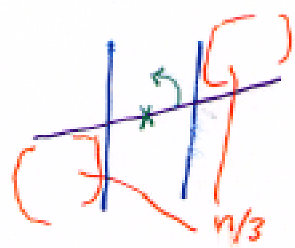
[Rado] $\exists$ pt of data depth $\left\lceil \frac{n}{d+1} \right\rceil$

"Center point"

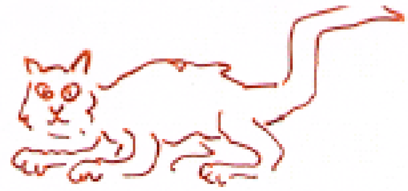


Cat line bisects left $\frac{2n}{3}$ and right $\frac{2n}{3}$ simultaneously. ("Ham sandwich cut")

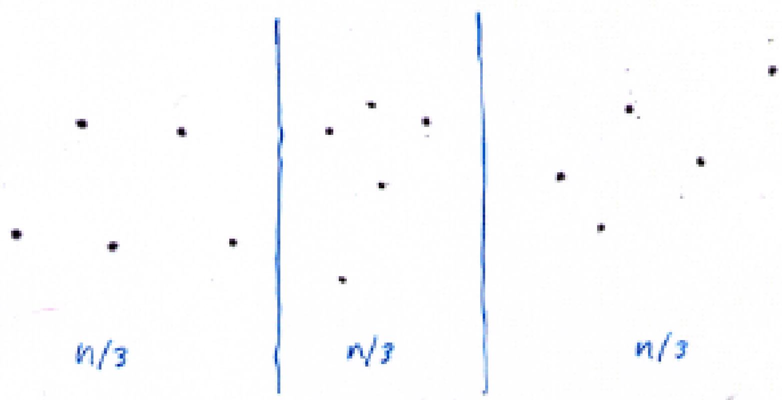




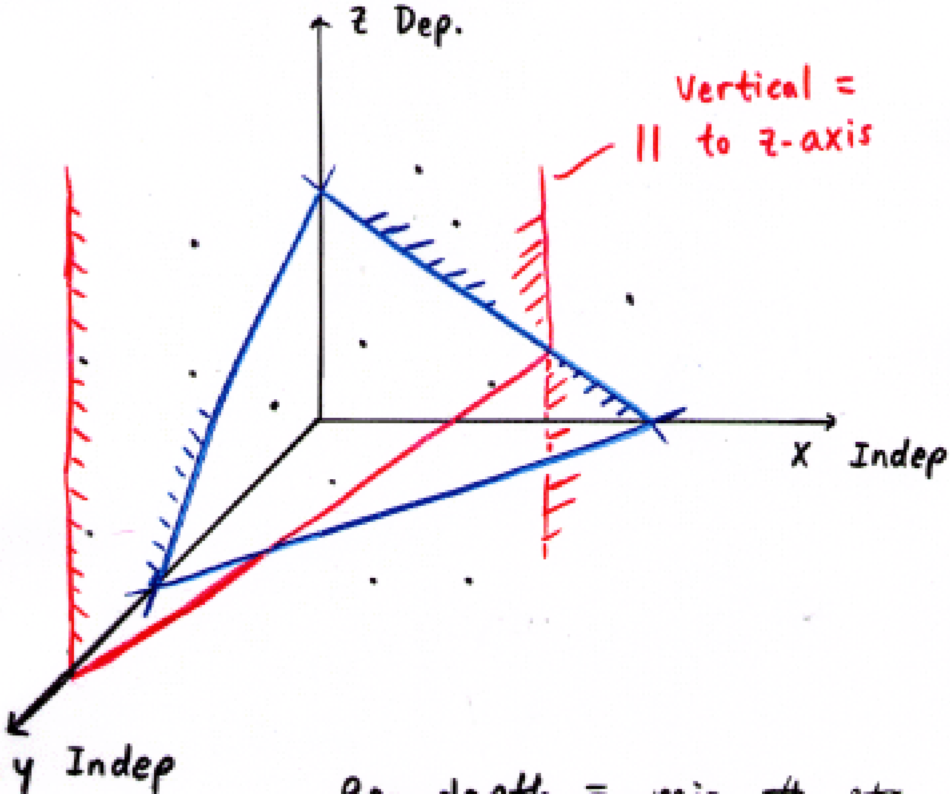
Cat line



For any set of pts in $\mathbb{R}^2$, $\exists$ a line of regression depth $\lceil \frac{n}{3} \rceil$. [Rousseuw-Hubert]



Higher Dimensions



Reg. depth = min # pts
crossed in a rotation
to vertical

Regression Depth Results

[Rousseeuw-Hubert 98] For n pts in $\mathbb{R}^d$
(Amenta-B-Eppstein-Teng 98) $\exists$ hyperplane of depth $\lceil \frac{n}{d+1} \rceil$
[Mizera 98]

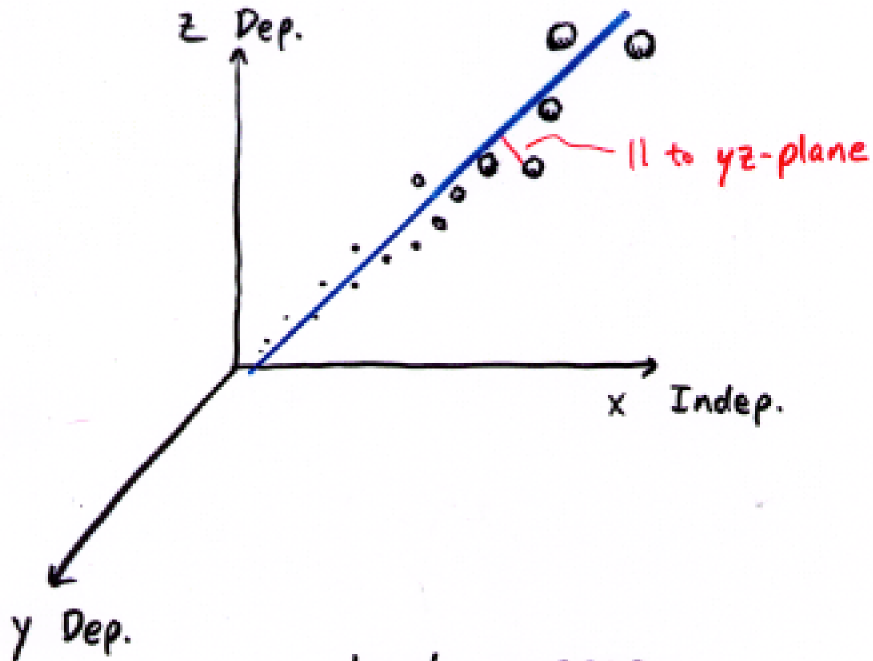
(van Kreveld et al. 99) (Langerman-Steiger 00) Fast alg'ms for $d=2$
$O(n \log n)$

[B-Eppstein 00] [Mizera 00] General'n of reg. depth to fitting a k-flat to pts in $\mathbb{R}^d$

Depth $\left\lceil \frac{n}{(d-k)(k+1)+1} \right\rceil$

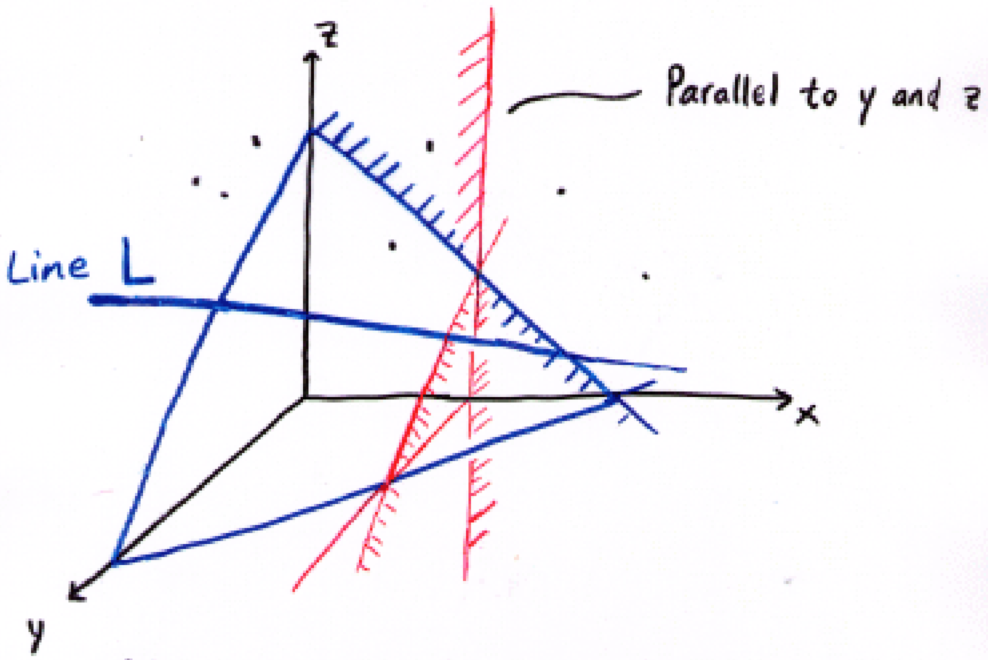
Multivariate Regression

More than one dep't variable.



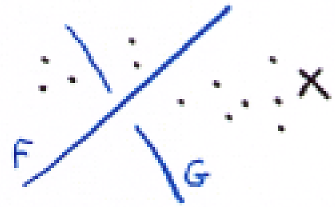
Least squares –
Do y and z separately

Double Wedge Definition



Min # pts in a double wedge s.t.
One bounding hyperplane contains L, &
other " " is vertical
(|| to all dep. axes).

Why?



Crossing Distance

between flats F and G w.r.t points X

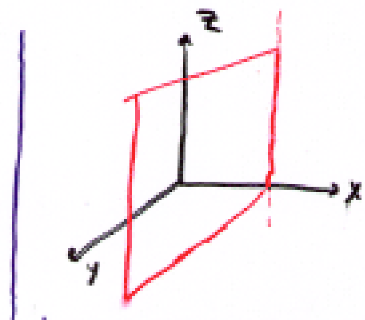
(Primal) Min # pts of X in double wedge w. one boundary containing F, and other " " G.

(Dual) Min # hyperplanes of D(X) crossed by a line segment connecting D(F) and D(G).

Unification

$k=0$ **Data Depth** — Crossing dist. of pt p from hyperplane at ∞

$0 \le k \le d-1$ **Multivariate Reg. Depth** — Crossing dist. of k-flat F from $(d-k-1)$-flat at vertical ∞.



Line contained in all yz-parallel planes

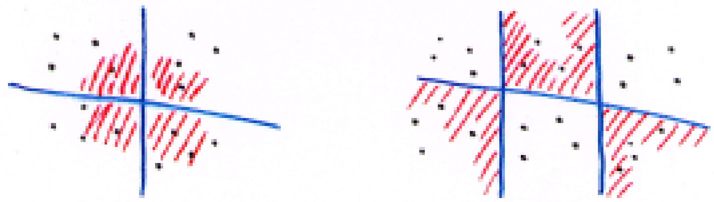
$k=d-1$ **Hyperplane Reg. Depth** — Crossing dist. of hyperplane H from point at vertical ∞.

Pt contained in all z-parallel planes

Proof that Deep k-flats Exist

Nontransversal Coll'n of Sets

cannot all be cut by a hyperplane



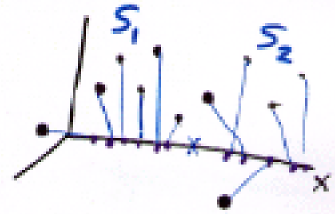
[Cor. to Alon-Kalai] $\exists\ C_k$ s.t. for any n pts in $\mathbb{R}^k$, $\exists$ nontransversal coll'n of $k+1$ subsets, each with $\geq n/C_k$ pts.

[Živaljević-Vrećica, "Center Transversal Thm"]
Given $k+1$ $\leq d$ pt sets in $\mathbb{R}^d$, each w. $\geq m$ pts $\exists$ k-flat F s.t. any half-space contain'g F contains $\geq \lceil \frac{m}{d-k+1} \rceil$ pts of each set.

$k=0$ Center Point $\qquad$ $k=d-1$ Ham Sandwich

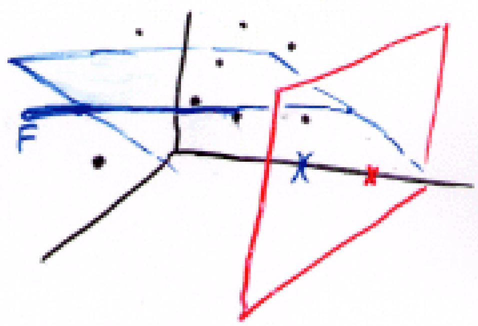
Proof, cont.

1. Find sets $S_1, \ldots, S_{K+1}$ that give a nontransversal coll'n when projected onto span of indep. axes.

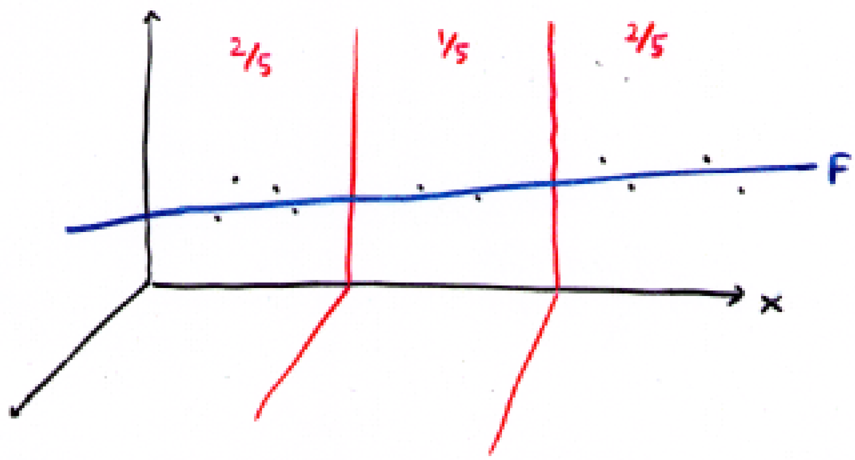


2. Use Center Transversal Thm to find F through $S_1, \ldots, S_{K+1}$.

Vert. Bndry of double wedge misses some S_i. In this S_i, double wedge = half-space through F



Catline Generalizes



Center Transversal Thm [Dol'nikov 90, Živaljević - Vrećica]

Can cut $k+1$ sets in $\mathbb{R}^d$ w. a k-flat F s.t. any half-space contain'g F has $\geq \frac{1}{d-k+1}$ of each set.

$k=1 \quad d=3 \implies$ Line of depth $\lceil \frac{n}{5} \rceil$

Mizera's Improvement

[Mizera 00] For n pts in gen'l pos'n in $\mathbb{R}^d$

$\exists$ k-flat of depth $\left\lceil \frac{n}{p+1} \right\rceil$

$p = (d-k)(k+1) + 1$

k-flats $\equiv$ Pts in p-dim'l space Θ

Pts $\equiv$ Fcns $R_i : \Theta \longrightarrow \mathbb{R}$

Residuals

$$\text{grad } R_i : \Theta \longrightarrow \mathbb{R}^p$$

$$\underset{\text{of } F \in \Theta}{\text{Depth}} \equiv \min_{v \in S^{d-1}} \#\{ i \mid v \cdot \text{grad } R_i(F) > 0 \}$$

Now use Borsuk-Ulam !

Future?

1. Bounds on C_k
2. "Tverberg version" of reg. depth

   n pts in $\mathbb{R}^d$ $\exists$ hyperplane H with nonzero depth in each of $\lceil \frac{n}{d+1} \rceil$ disj. subsets.
3. (4,1,4) Equipartition
4. General Equipart'n Results
5.