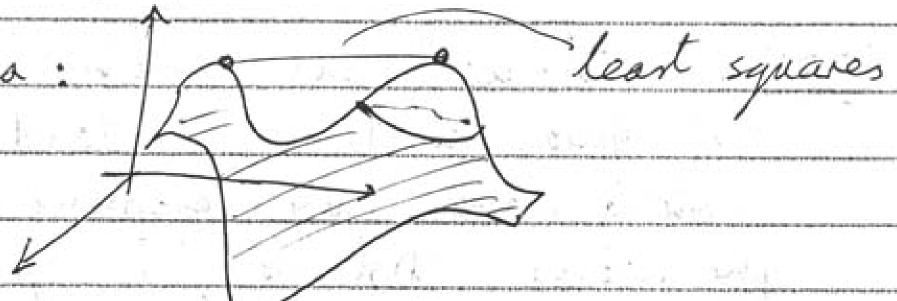
Lecture Notes

BRIDSON – AIM, JUNE 1, 2003, PALO ALTO

– DISTANCES, SPACES + CURVATURE

1) – Why look at different spaces / distances rather than just doing distance estimates in Euclidean space $\mathbb{R}^n$?

Rough idea :



Data involves some measurement, giving points in $\mathbb{R}^n$. You want to talk about "CLOSE" data points. Typically, you might do a least squares approximation, which amounts to taking STRAIGHT-LINE DISTANCE in $\mathbb{R}^n$.

BUT a more natural measure of SEPARATION (DISTANCE) between data points is the LENGTH of path IN THE ~~possib~~ SPACE OF POSSIBLE DATA (drawn as a surface).

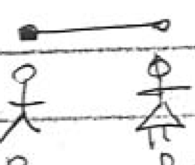
Thus we eschew EXTRINSIC distance (the straight line) for the INTRINSIC distance in the space of outcomes.

- So, we want to work in SPACES with non-Euclidean notions of DISTANCE (separation).

ASIDE : "distance" and intuition / vocabulary of space should not mislead you into thinking we're only describing spatial phenomena. "Distance" is just a metaphor for (quantified) difference

EXAMPLE : Hand-shaking graph.

VERTICES = PEOPLE ON EARTH

EDGES = {  | p_1 has shaken hands with p_2 }

P_2 P_2

- Distance — # edges to be traversed in going from one vertex to another in shortest path.

Points to Observe :

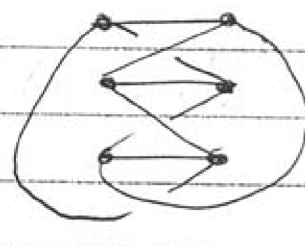
1) We had a notion of <u>length</u> and turned it into a notion of distance (we'll go the other way later).

2) One can embed this graph in $\mathbb{R}^3$ in a length-preserving way, but the straight-line distance of the embedding will have little or nothing to do with the natural notion of distance defined above.
— And why bother! Unnatural

---

3) Although one can embed any graph in $\mathbb{R}^3$, one cannot embed in $\mathbb{R}^2$, necessarily

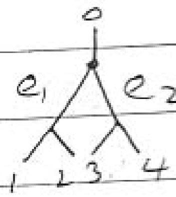
<u>e.g.</u> utilities graph $K_{3,3}$



<u>MORAL</u> : As we look at higher dimensional spaces, we can't hope for low (≤ 4) - dimensional embeddings

— Another example (seen in previous talk)

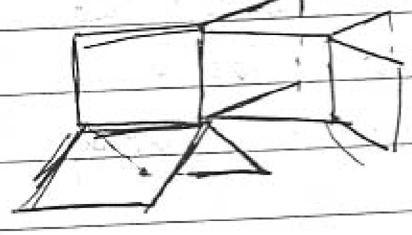
TREE SPACE $n=4$



two edge lengths $|e_1|+|e_2|=1$

— as internal edges degenerate we get change of topology; as edges vary we move around in a square

— We get a space built out of squares:



can't expect a Euclidean embedding, misleading to try a least-squares distance, but there <u>IS</u> a natural distance:

$d(x,y)$ = <u>length</u> of shortest path from x to y

measured <u>LOCALLY</u> in squares

<u>GLOBAL</u> measure of separation.

<u>KEY IDEA</u> : <u>Geometry</u> (of spaces with notion of distance) allows one to use <u>LOCAL</u> information to make <u>GLOBAL</u> conclusions (a process hopeless with discrete spaces — thus we want to arrange our data into spaces)

<u>SLIDE:1</u> — METRIC SPACES

- Deliberately pared down to essential features

- Restricting to <u>GEODESIC</u> Metrics makes geometry possible in a real sense (in particular opens the possibility of a local-to-global analysis of the space).

<u>IMPORTANT REMARKS</u>: (I) Can go from notion of distance <u>TO</u> length. (and can go back in geodesic spaces)

(II) SEPARATION ⇝ distance

In many circumstances one may have a notion of <u>separation</u> of data,

$$s : X \times X \longrightarrow \mathbb{R} \quad (\text{or } \mathbb{N})$$

that does not satisfy the triangle inequality. In such circumstances, one can manufacture a metric by

$d(x,y)$ = length of shortest sequence of hops $x \rightsquigarrow x_1 \rightsquigarrow \cdots \rightsquigarrow x_n = y$

such that $s(x_i, x_{i+1}) \leq$ constant.

— or some similar process.

<u>SLIDES 2+3</u>: Hyperbolic graphs, with explanation. Explanation of <u>GROWTH</u> point — the difficulty of embedding things in Euclidean space is <u>not</u> just a question of <u>dimension</u>

SLIDE 4 — Geodesics
— Core issues

Re. comparison : coordinates ~> Riemannian geometry

comparison of finite ~~cong~~ configurations....

SLIDE 5 : CURVATURE

SLIDE 6 : CAT(0) spaces

Emphasis — the local-to-global POWER

Where is this power coming from - CONVEXITY

SLIDE 7 : CONVEXITY.
SLIDE 8 : CENTRES
- Explanation of Euclidean origins. Discussion.
- Emphasis on the applications to tree space.

- There follows a detailed description of (a variation on) the link of the star vertex in tree space for $n=5$