An Optimal Linear Time Algorithm for Quasi-Monotonic Segmentation

Daniel Lemire (UQAM),

Martin Brooks (NRC),

Yuhong Yan (NRC)

A definition to begin with!

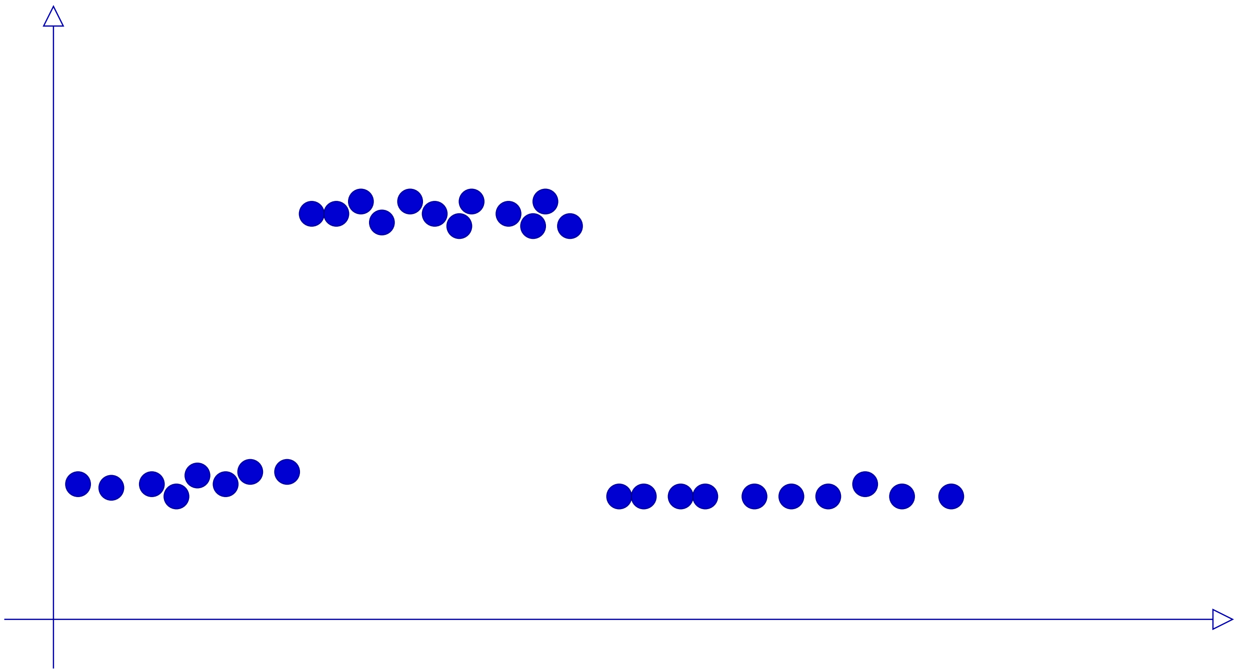
A function F is *monotonic* if

- $F(x) \geq F(z)$ whenever $x > z$ (increasing) **or**
- $F(x) \leq F(z)$ whenever $x > z$ (decreasing).

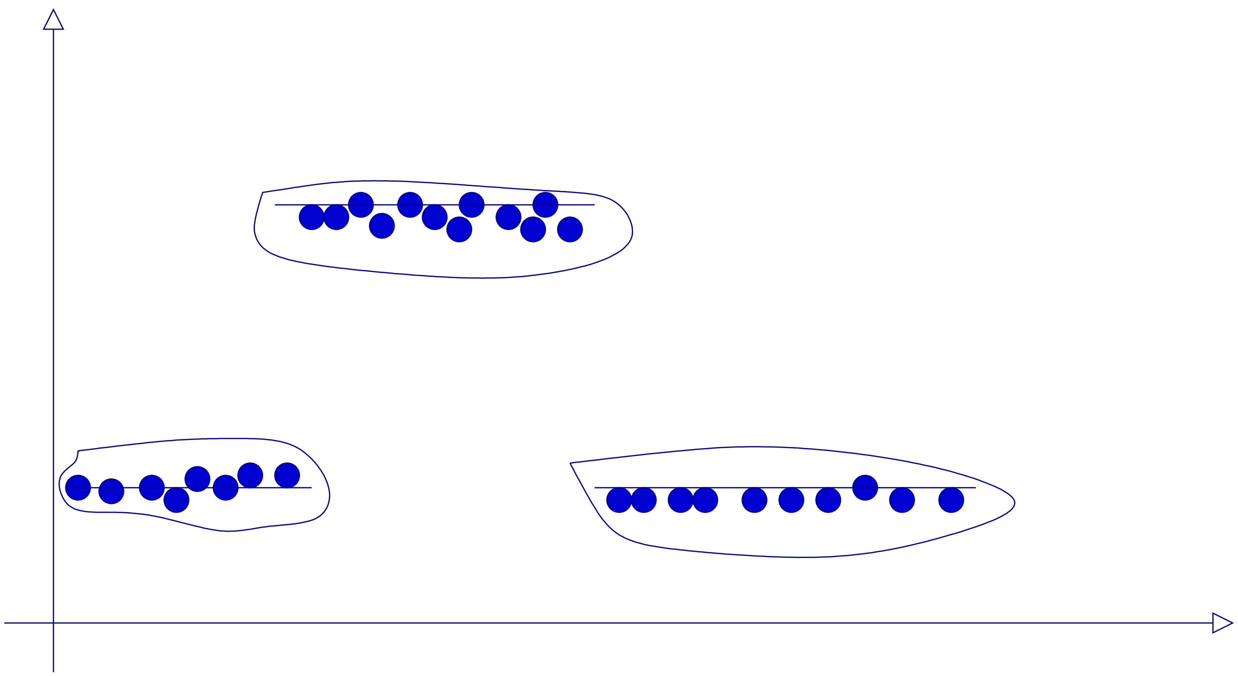
Monotonicity is:

- morphologic: slopes don't matter, you can "time warp"!
- sample-removal oblivious: missing points don't matter.

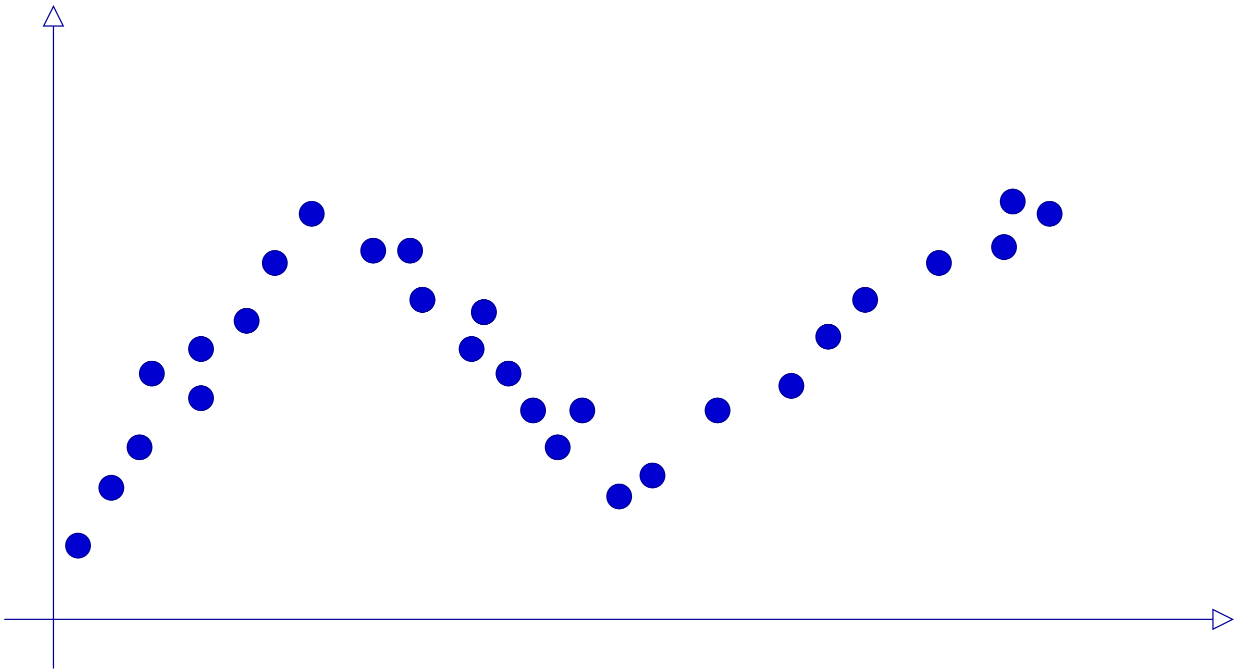
Piecewise quasi-constant scattered data



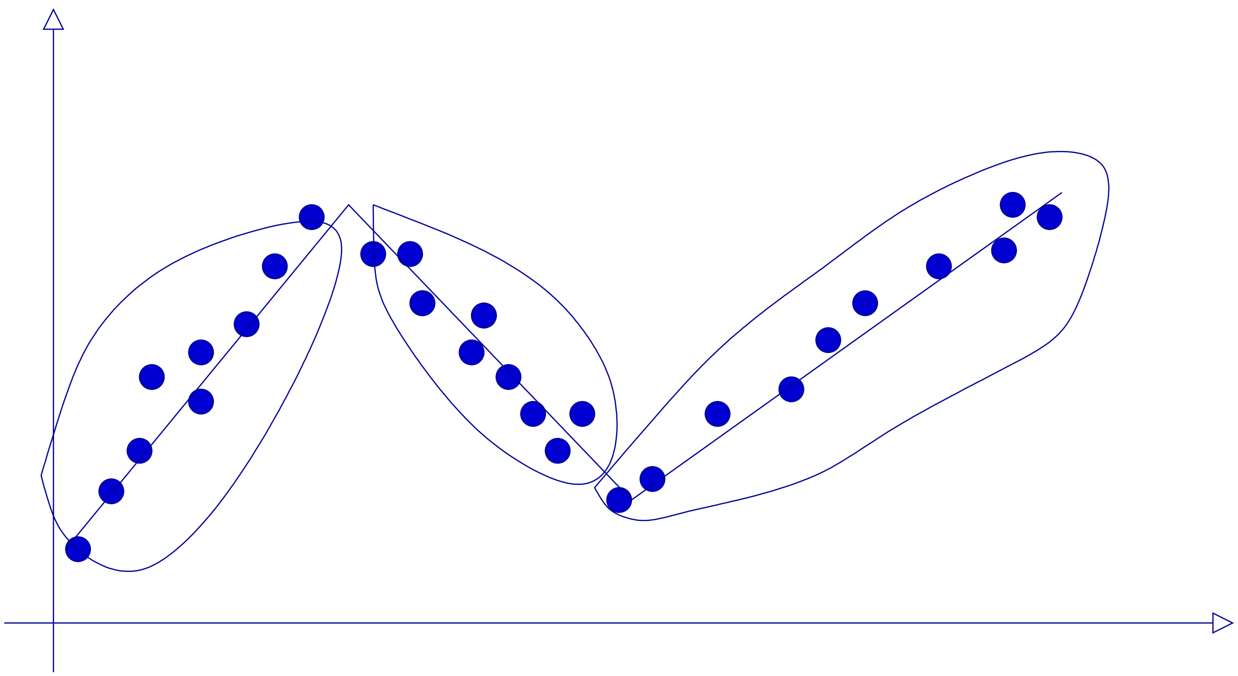
Segmentation of piecewise quasi-constant scattered data



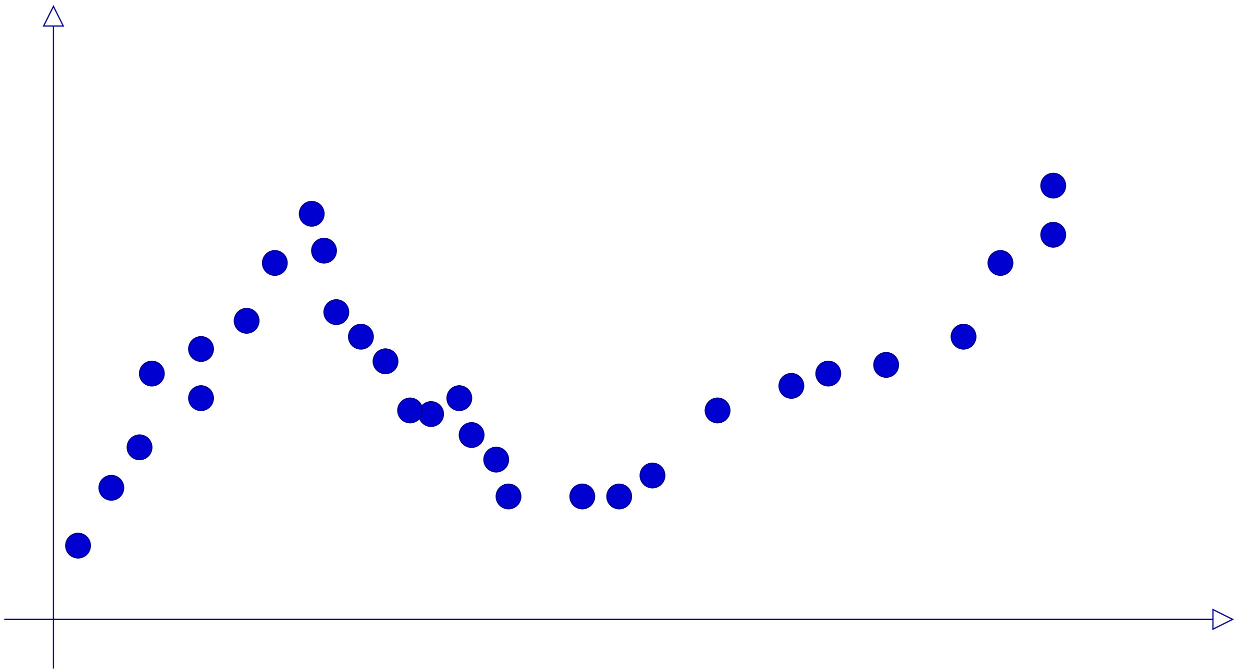
Piecewise quasi-linear scattered data



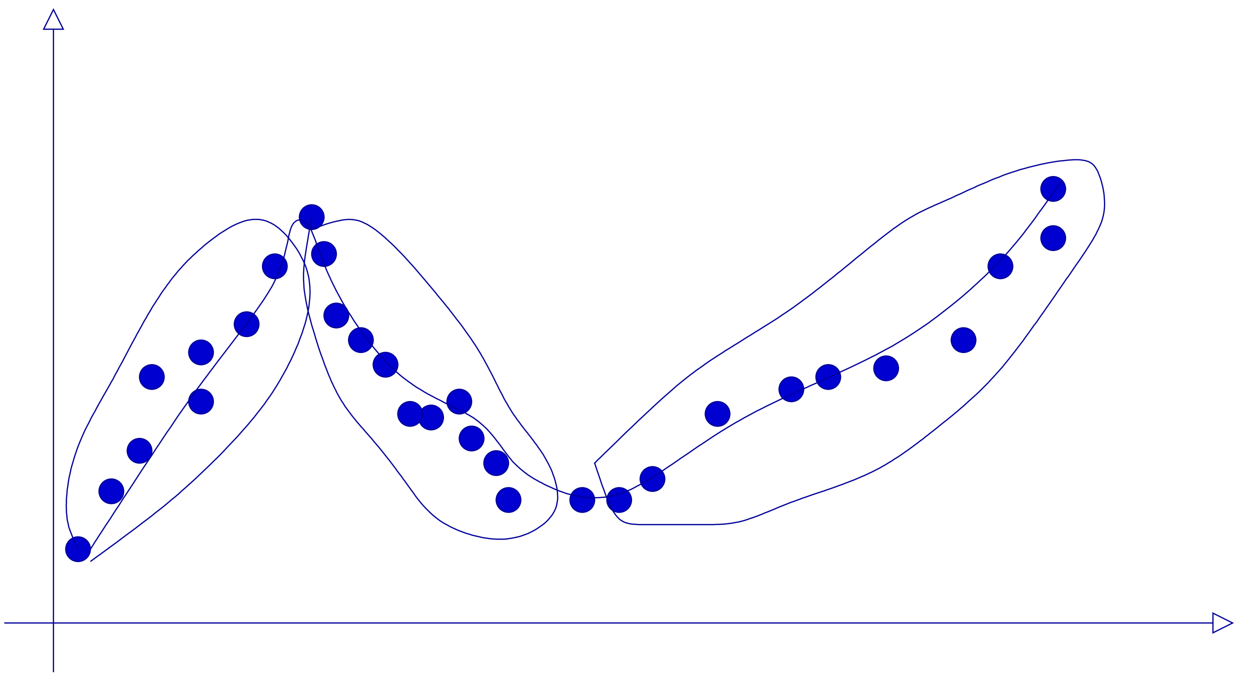
Segmentation of piecewise quasi-linear scattered data



Piecewise quasi-monotonic scattered data



Segmentation of piecewise quasi-monotonic scattered data

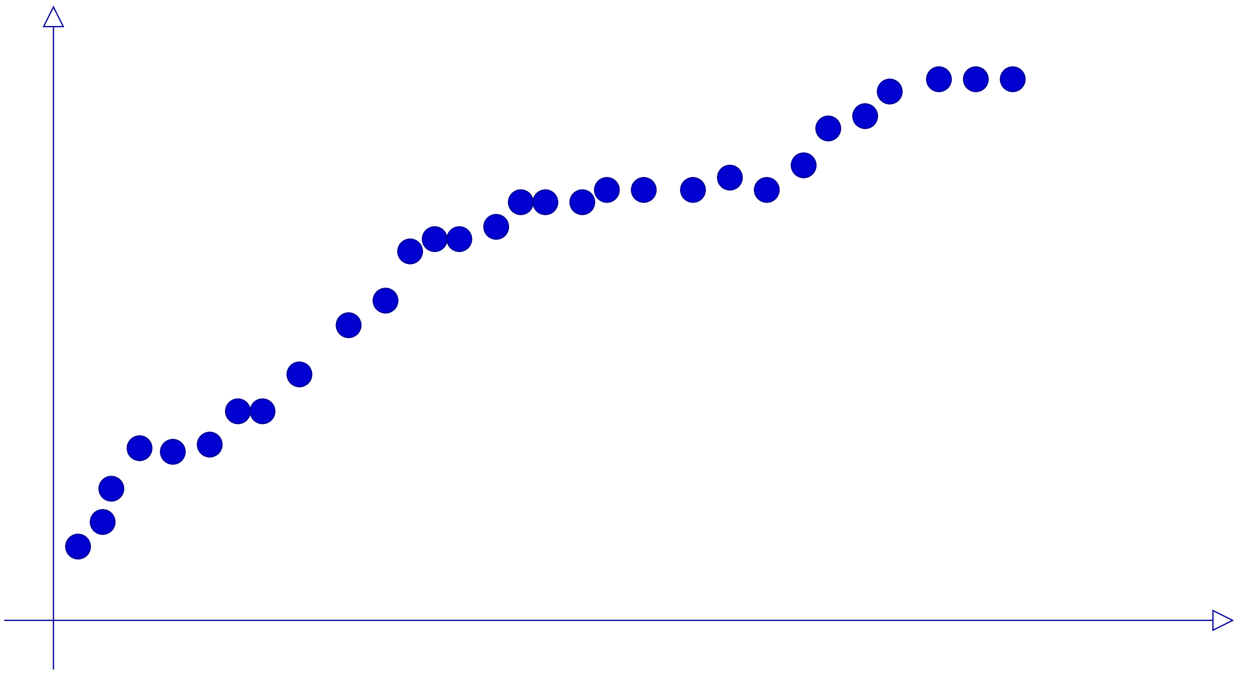


A Generic Segmentation Problem

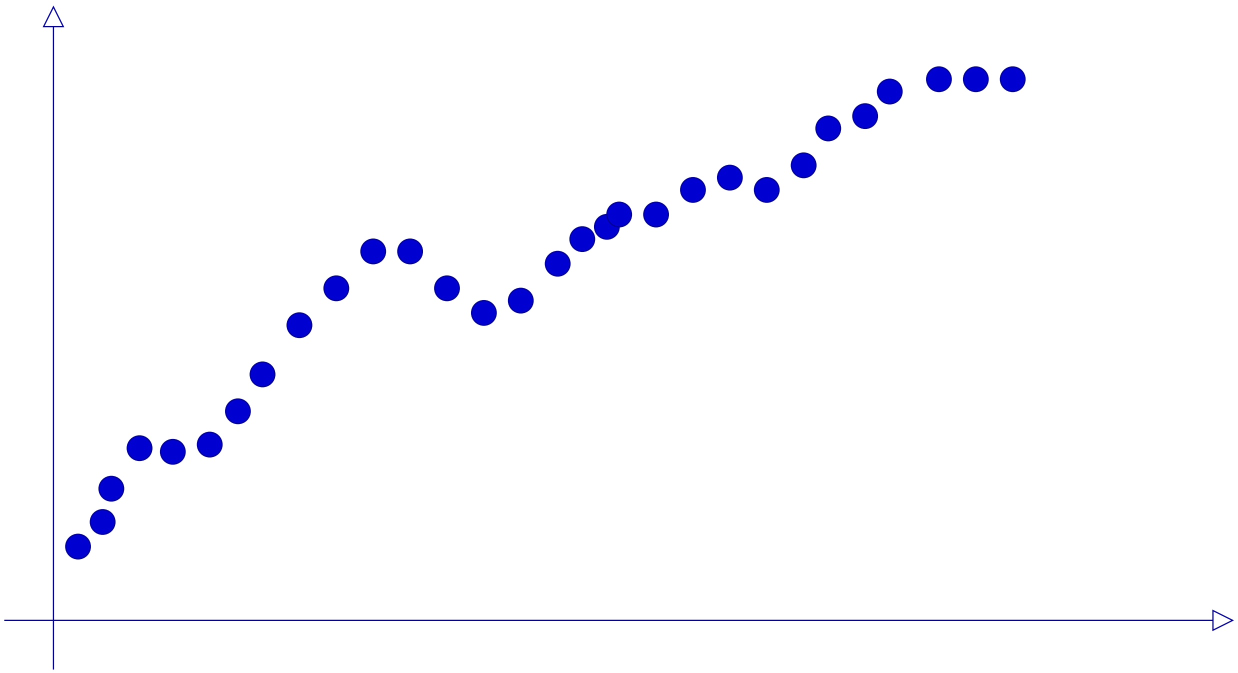
- for segmentation: $S = \{S_1, S_2, ...\}$
- define an **error metric** Q : e.g. linearity, monotonicity, ...
- Given error ε
  - $\varepsilon = \max(Q(S_1), Q(S_2), \ldots)$

  find segmentation with least number of segments.

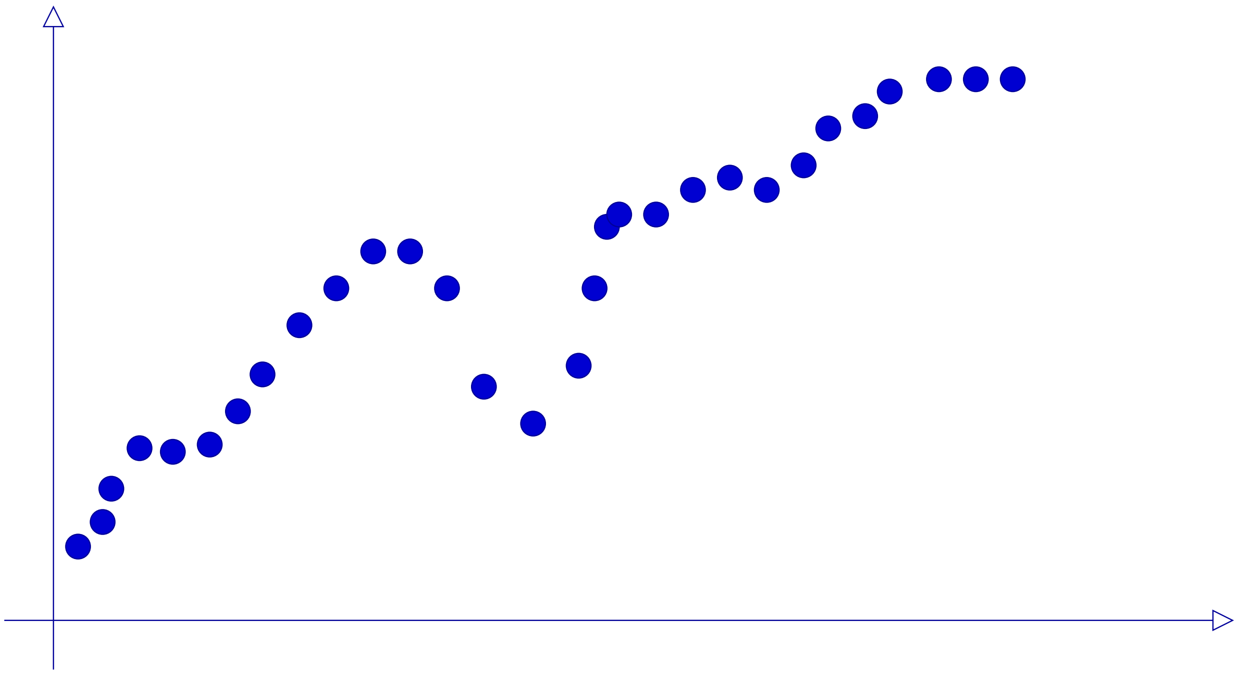
Quantifying monotonicity: pretty good



Quantifying monotonicity: good



Quantifying monotonicity: not good



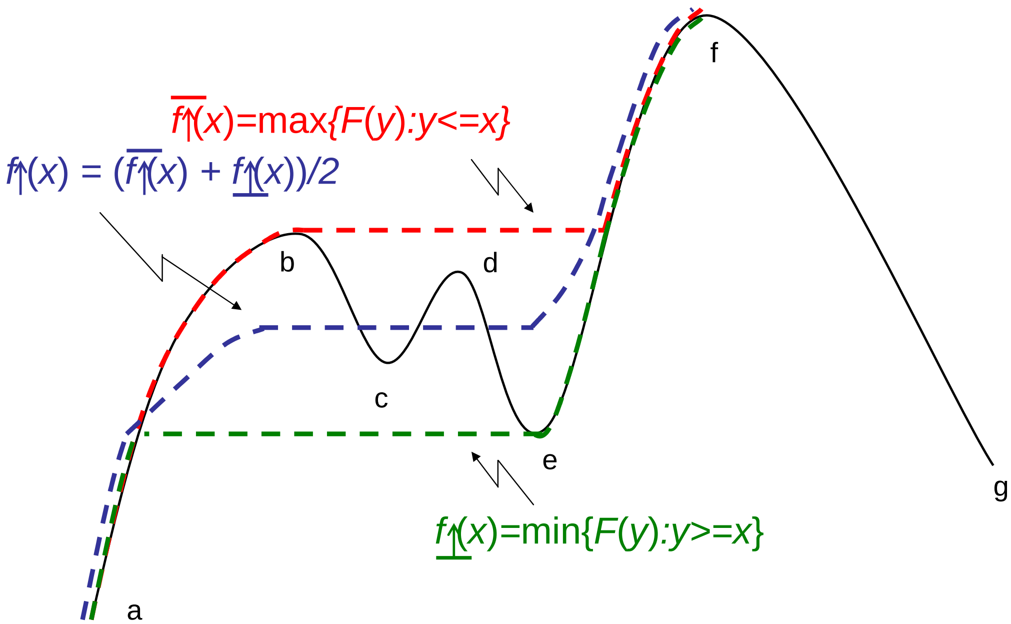
How do we measure how monotone something is?

Given y_i, find monotonic sequence z_i minimizing

▷ $Q = \max_i |y_i - z_i|$ (l_∞ error)

Surprise! It is an easy problem.

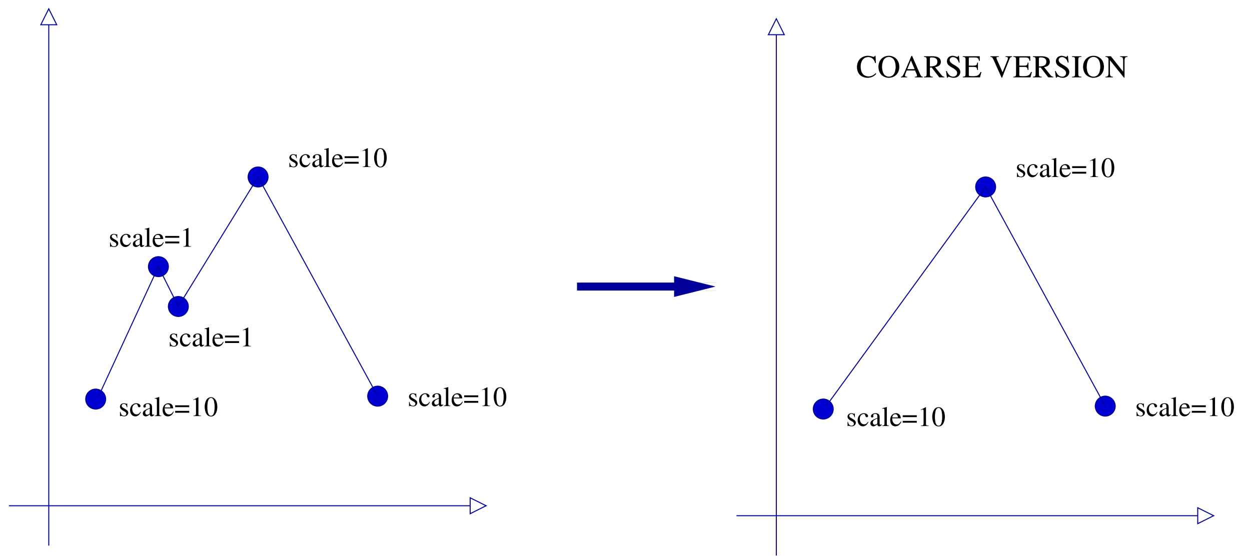
Demonstration: seek the best monotonic increasing approximation function



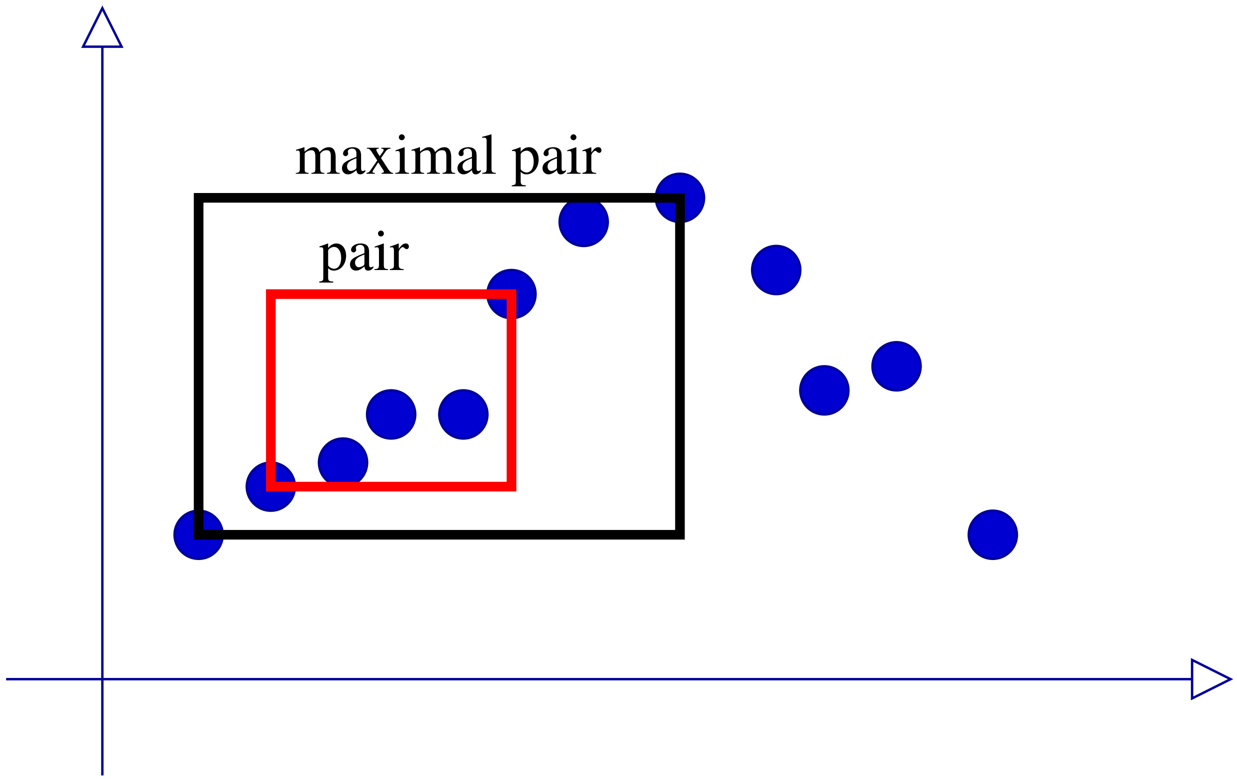
A multiscale approach for monotonicity

- Label each extrema with a scale parameter.
- Intuition is that some extrema are “details” others are important (coarse).
- Scale is defined by the y-axis metric (large variation: large scale).

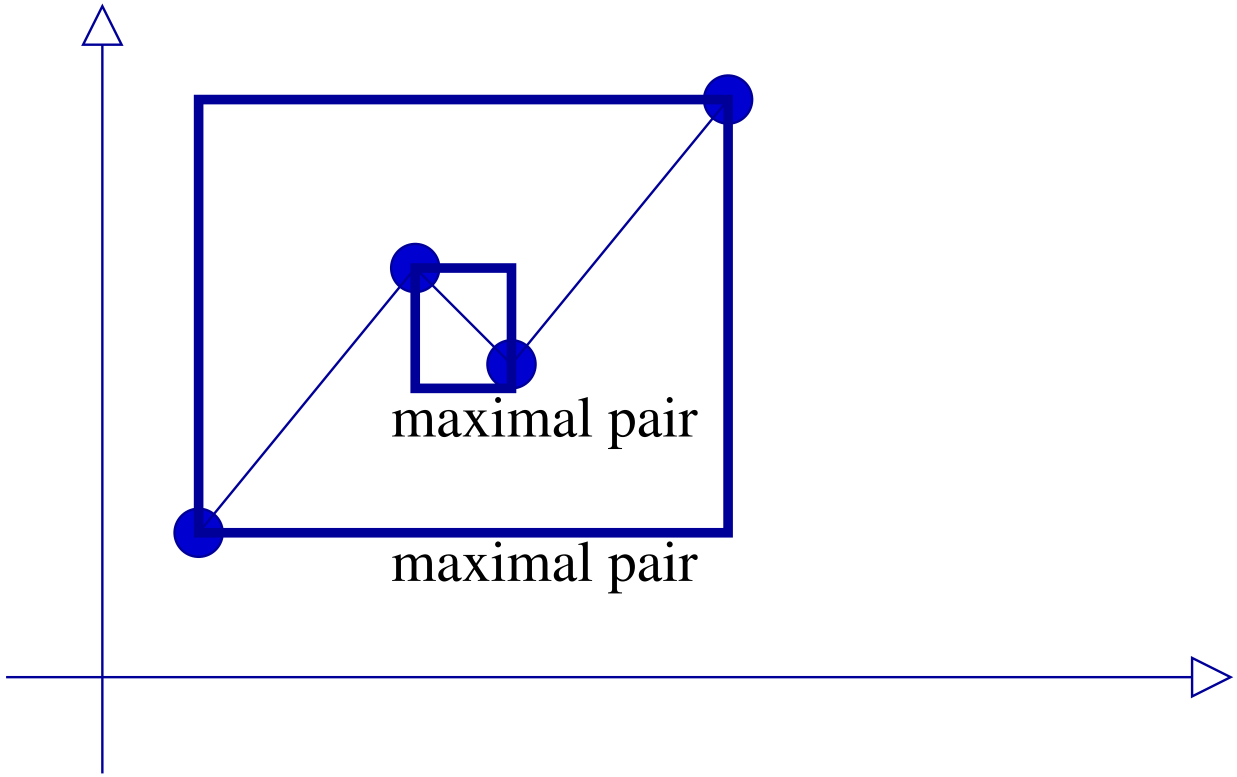
A multiscale approach for monotonicity (2)



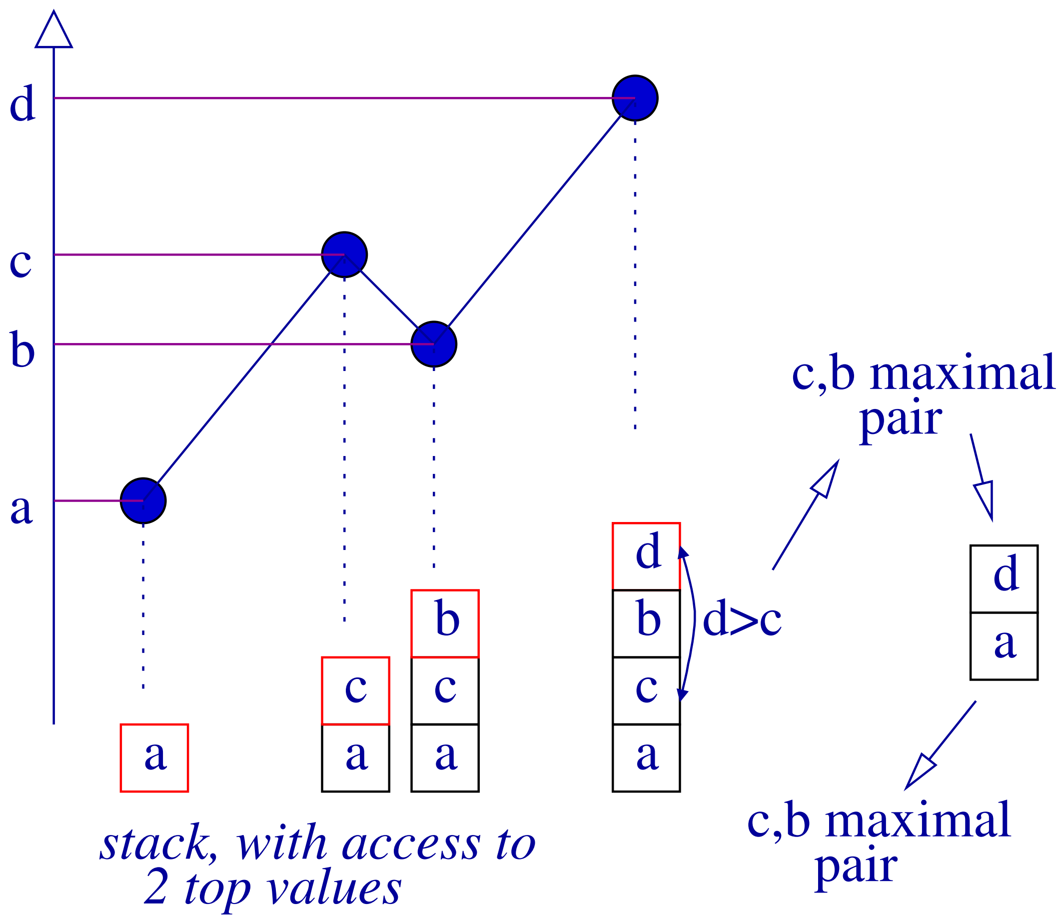
Maximal pairs (2)



Scale of an extrema versus max. pairs



Stack-based Labelling: $O(n)$ (one pass)



THM: top $k+1$ extrema give k-segmentation

- ▷ **Pick all extrema at scale δ or more.** This is an optimal monotonic segmentation!
- ▷ The converse is also true: **pick the top $k+1$ extrema** having higher scale than the rest, and it is an optimal segmentation.

A simple $O(nk)$ optimal segmentation algorithm

- ▷ We can label the extrema in time $O(n)$ (one pass).
- ▷ Filter to find the top k extrema in time $O(nk)$.