Robust Mean Estimation

- $d = 1$: remove $\alpha n$ points from each side. Determine the mean

$$T = \frac{\sum_{i=\alpha n}^{(1-\alpha)n} x_i}{(1 - 2\alpha)n}$$

- $d = 2$: remove $2\alpha n$ points from outermost layers (one by one).

- Trivial algorithm $O(n^2 \log n)$
- Jarvis march $O(n^2)$
- Overmars and van Leeuven algorithm $O(n(\log n)^2)$
- Optimal algorithm exists $\Theta(n \log n)$

$\text{depth}(p) = 3$
Isotonic regression

An isotone function (not best) approximating a finite point set.

\[ \| f - f^* \| \rightarrow \min \]

A best isotone fit is a step function. The number of steps and the points at which they break must both be determined.
Isotonic Regression

- **Given:** $n$ points $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$ in the plane.
- **Find:** a non-decreasing step-function $f$ minimizing
  \[ \sum_{i=1}^{n} (y_i - f(x_i))^2 \]
- Jumps occur at $x_1, x_2, \ldots, x_n$ only.
- Assume consecutive jumps at $x_i$ and $x_k$, $i < k$. Then
  \[ f(x_p) = (\sum_{i=i}^{k-1} y_i) / (k - i) \]
- The slope of the tangent to the convex hull under $p_i$ is the value of $f(x_i)$.

\[
\begin{align*}
  p_0 &= (0, 0), \\
p_j &= (j, \sum_{i=1}^{j} y_i)
\end{align*}
\]

- $f(x)$ in $O(n \log n)$
  - if sorted,
    - then $O(n)$
Diameter of a Point Set

- **Given**: \( n \) points in the plane.
- **Find**: Pair of points farthest apart.
- Diameter of a point sets is equal to the diameter of its convex hull.
- Distance between extreme points of a convex set is greatest between parallel supporting lines.

We need to solve this problem many times in the minimum diameter \( k \)-clustering problem.

- Yaglom-Boltyanski (1961)
  - The diameter of a convex figure is the greatest distance between parallel lines of support.
Characterization of antipodal pairs. The intersection of $\alpha_i$ and $\alpha_s$ is a pair of planar wedges sharing their vertex.

- a pair of points that does admit parallel supporting lines is called antipodal.
procedure ANTIPODAL PAIRS
begin
  $d$ := $d$.
  while $d < | [d]$ NEXT $< | [b]$ NEXT do
    if $d$ $=$ $b$ then print ("**end of parallel edges**")
    $d$ := $d$.
  end
end

The Adjacent of Parallel Edges
If the point $(x_d, y_d) \neq (x_d, y_d)$
then
  $d$ := $d$.
end

if $(x_d, y_d) \neq (x_d, y_d)$
then print $(b, d)$.

end

while $(b, d) < (d, b)$ do
  $d$ := $d$.
end

end

procedure ANTIPODAL PAIRS
begin
  $d$ := $d$.
  while $d < | [d]$ NEXT $< | [b]$ NEXT do
    if $d$ $=$ $b$ then print ("**end of parallel edges**")
    $d$ := $d$.
  end
end

Code example:

Example:
Diameter of a Point Set - Lower Bound

- **Given:** Two sets $A = \{a_1, ..., a_n\}$ and $B = \{b_1, ..., b_n\}$ of $n$ real numbers.

- Deciding whether $A \cap B = \emptyset$ requires $\Theta(n \log n)$ time.

- Consider a circle $C$ with center in the origo. Map each number $a_i$ to the intersection point of the line $y = a_i x$ with $C$ (in the first quadrant). Map each number $b_i$ to the intersection point of the line $y = b_i x$ with $C$ (in the third quadrant).

- Let $S$ denote the set of $2n$ points.

- $A \cap B = \emptyset$ iff $diam(S) = diam(C)$. 

![Diagram showing the intersection points and the circle C with labeled points.](image)