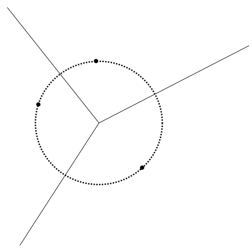
Group 6: Timm Behner, Philipp Bruckschen, Patrick Kaster, Markus Schwalb
MA-INF 4111 - Intelligent Learning and Analysis Systems: Machine Learning
Exercise Sheet 4

1. Voronoi diagrams

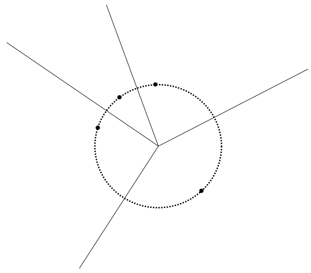
(a) Proof by induction over the number of points on a circle with radius r in $\mathbb{R}^2$:

Induction start: $n = 3$ Three point lying on a circle with radius r are non-collinear and the resulting cells are unbounded (figure 1a).

Induction step: $n \to n+1$ Let the induction assumption be true for n points. Let p, q be two abitrary neighboring points on the circle. We insert a new point o exactly at the intersection of the bisector of p, q and the circles radius r. The resulting cells are unbounded (figure 1b).



(a) induction start



(b) induction step

Figure 1: proof by induction over points on a circle

(b) Let $\mathrm{Vor}(P)$ denote the Voronoi diagram of points P. Euler's formula for planer graphs states that

$$v - e + f = 2,$$

where v, e, f denote the number of vertices, edges and faces of the graph.

The number of faces is $f = |P| = n = \#\text{cells}\,(\mathrm{Vor}(P))$. To make use of Euler's formula, we have to construct a planar Graph from $\mathrm{Vor}(P)$, by adding an additional vertex to the diagram, situated at infinity, and connecting every infinite edge (from the unbounded cells) to this vertex (figure 2).

We note that each edge in $\mathrm{Vor}(P)$ has exactly two incident vertices and each vertex of $\mathrm{Vor}(P) + \infty$ has at least degree (deg) 3. Therefore we conclude

$$\sum_{\text{vertex} \in \mathrm{Vor}(P)+\infty} \deg(\text{vertex}) = 2e \geq 3(v+1) \tag{1}$$

Since we inserted our additional point ∞, Euler's Formula for our constructed graph is:

$$(v+1) - e + f = (v+1) - e + n = 2$$

Inserting this into (1) gives us:

$$2e \geq 3(2 + e - n) = 6 + 3e - 3n \Rightarrow$$
$$e \leq 3n - 6$$

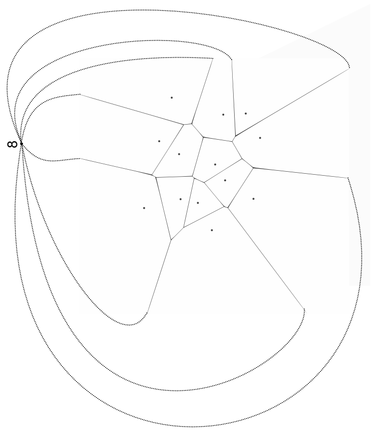


Figure 2: augmenting Voronoi diagram to planar graph

Inserting this result into (1):

$$v \leq 2n - 5$$

■

3. Hypothesis Testing

(a) $n = 100, \#'' + '' = 87$

$$\begin{aligned}
\text{error}_S(h) &= \frac{n - \#'' + ''}{n} = \frac{13}{100} = 0.13 \\
\text{error}_{\text{true}}^{95\%}(h) &= \text{error}_S(h) \pm z_N \sqrt{\frac{\text{error}_S(h)(1 - \text{error}_S(h))}{n}} \\
&= 0.13 \pm 1.96 \sqrt{\frac{0.13(1 - 0.13)}{100}} \approx [0.196, 0.064]
\end{aligned}$$

(b) With a probability of approximately 95% the true error of the hypothesis h lies within the above interval.

■