AlgGeo 16.1

Wdh.: ε-Netz-Theorem: $(X, \mathcal{F})$ Mengenverband, $\dim_{VC}(X, \mathcal{F}) = d$; Maß μ

Dann gibt es für $r > 1$ ein $\underbrace{\frac{1}{r}\text{-Netz } N \subseteq X}_{}$ mit $Cdr \ln r$ Elementen.

Dabei: C Konstante.

$$\forall S \in \mathcal{F} : \left(\mu(S) \geq \frac{1}{r} \Rightarrow N \cap S \neq \emptyset\right)$$

Beweis: $s := Cdr \ln r$, $k := \frac{s}{2r}$

entferne aus $\mathcal{F}$ alle S mit $\mu(S) < \frac{1}{r}$.

Ziehe aus X zwei s-Sample N, M μ-verteilt, mit Zurücklegen

(einzeln oder aus $2s$-Sample A)

E_0: $\exists S \in \mathcal{F}$: $N \cap S = \emptyset$ (N ist kein $\frac{1}{r}$-Netz)

E_1: $\exists S \in \mathcal{F}$: $N \cap S = \emptyset$ und $|M \cap S| \geq k$

E_S: $N \cap S = \emptyset$ und $|M \cap S| \geq k \mid A$ für A fest, $S \in \mathcal{F}$

schon gezeigt:

$$P(E_0) < 2 \cdot P(E_1)$$

brauchen noch $P(E_1) \overset{!}{<} \frac{1}{2}$ $\Rightarrow P(E_0) < 1$ $\Rightarrow$ Existenz von $\frac{1}{r}$-Netzen

dafür schon bewiesen: $P(E_S) \leq r^{-\frac{Cd}{4}}$

jetzt: A immer noch fest; $S \in \mathcal{F}$ variabel.

AlgGeo 16.2

<u>Beobachtung</u>: Ereignis E_S hängt nur von $A \cap S$ ab (nicht von ganz S)

$$\Rightarrow P(E_1 | A) = P(\exists S \in \mathcal{F} : E_S) = P\Big(\bigcup_{S \in \mathcal{F}} E_S\Big) = P\Big(\bigcup_{\substack{S \in \mathcal{F} \\ S \cap A \text{ paarw. verschieden}}} E_S\Big)$$

$$\underset{\text{Union bound}}{\leq} \underbrace{\sum_{\substack{S \in \mathcal{F} \\ S \cap A \text{ paarw. versch.}}}}_{\#\text{ Summanden } = |\mathcal{F}|_A| \underset{\text{Shatter Function Lemma}}{\leq} \left(\frac{e \cdot 2s}{d}\right)^d} \overbrace{P(E_S)}^{\leq r^{-\frac{cd}{4}}}$$

$$\underset{\text{einsetzen}}{\leq} \left(\frac{2ecdr\ln r}{d}\right)^d r^{-\frac{cd}{4}}$$

$$= \Bigg(\underbrace{2ec\frac{\ln r}{r^{\frac{c}{4}-1}}}_{<\frac{1}{2}}\Bigg)^d \overset{!}{<} \frac{1}{2} \quad \text{für } d, r \geq 2, \ c \text{ hinreichend groß}$$

$c \approx 13{,}1$

($\rightarrow$ Geogebra)

AlgGeo 16.3

$$\Rightarrow P(E_1) \le \sum_A \underbrace{P(E_1 | A)}_{< \frac{1}{2}} \cdot P(A) < \frac{1}{2}$$ $\boxed{\varepsilon\text{-Thm.}}$ (Haussler, Welzl)

Das bedeutet für Kunstgalerien:

Theorem Es gibt Konstante D: $\forall r > 1$:

Ist P ein einfaches Polygon mit:

$$\forall p \in P: \quad \text{Fläche}(\text{vis}(p)) \ge \frac{1}{r} \text{Fläche}(P),$$

so läßt sich P mit $Dr \ln r$ Wächtern bewachen.

unabhängig von # Ecken von P

Bew: Betrachte $X = P$ und $\mathcal{F} = \{\text{vis}(p) ; p \in P\}$ und wende an: $\dim_{VC}(X, \mathcal{F}) < \infty$

ε-Netz-Theorem. $\square$

(hier: $D \sim 4300$)

Bemerkungen: (i) Schranke $C d r \ln r$ kann für allgemeine $\frac{1}{r}$-Netze nicht verbessert werden, wohl für Spezialfälle

(ii) Durch Vergrößerung von s erhält man untere Schranken für die W'keit, daß N ein $\frac{1}{r}$-Netz ist.

AlgGeo 16.4

→ Literatur: Sariel Har-Peled, Geometric Approximation Algorithms, 2011.

(iii) Statt "nicht-leerer Schnitt" kann man "angemessene Repräsentation" verlangen:

<u>Theorem</u> $(X, \mathcal{F})$, μ wie oben.

Dann hat ein Sample A der Größe $C d r^2 \ln \frac{dr}{\delta}$ mit Wahrscheinlichkeit $\geq 1-\delta$ folgende Eigenschaft:

$$\forall F \in \mathcal{F} : \left| \frac{|A \cap F|}{|A|} - \mu(F) \right| < \frac{1}{r}$$

"$\frac{1}{r}$-Approximation" ; "Diskrepanz"

<u>Algorithmische Anwendung</u>

Gegeben: einfaches Polygon P mit n Ecken, $\mu(P) = 1$

Gesucht: $p \in P$: $\mu(\mathrm{vis}(p))$ maximal

(1) Exakte Lösung ?

(Cheong, Efrat, Har-Peled '04: "p_ε"

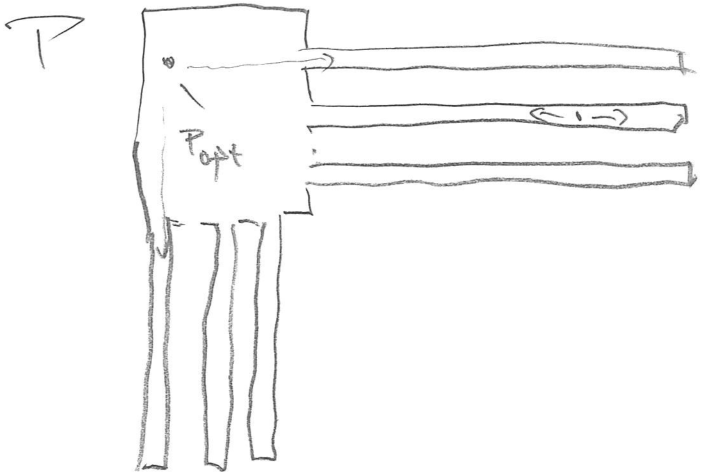
"On the guard who sees most and the shop that sells most".

Alg Geo 16.5

② Random Sampling: * wähle zufällig genügend viele Punkte $p_i \in P$ aus

* berechne $\max\limits_{1 \le i \le m} \mu(\mathrm{vis}(p_i))$

Funktioniert nicht gut:



n Korridore, $\sim 4n$ Ecken

jeder Korridor: Fläche $\frac{1}{n} - \frac{1}{n^{10}}$

$\Rightarrow$ Rechteck hat Fläche

$$1 - n\left(\frac{1}{n} - \frac{1}{n^{10}}\right) = \frac{1}{n^{9}}$$

$\Rightarrow$ W'keit, Rechteck zu treffen, ist klein

$\Rightarrow$ W'keit ist groß, daß jeder Sample-Punkt in einem Korridor liegt und nur diesen sieht, ist groß.

$\Rightarrow$ bestenfalls 2-Approximation

AlgGeo 16.6

(3) Benutze Approximationssatz für $\mathcal{F} = \{vis(p), p \in X\}$, $X = P$

Gegeben: Fehlerschranke δ

Wähle gleichverteilt eine Menge A von

$$M := \underset{\substack{\uparrow \\ 14}}{d} \, C' \frac{n^2}{\delta^2} \log \frac{n}{\delta} \text{ Punkte aus } P$$

Wende Approx-Theorem an mit $d = 14$, $r := \frac{2n}{\delta}$

$\Rightarrow$ mit W'keit $\geq 1 - \delta$ ist A eine $\frac{\delta}{2n}$-Approximation, d.h.

$$\forall p \in P: \quad \mu(vis(p)) \sim \frac{\# a \in A : p \text{ sieht } a}{|A|} \quad \text{bis auf } \frac{\delta}{2n}\text{-Fehler}$$

genauer: $\left| \cdot - \cdot \right| < \frac{\delta}{2n}$

Lemma: Sei p_a der Punkt in P, der die meisten Elemente von A sieht.

Dann: $\mu(vis(p_a)) \geq (1-\delta)\, \mu(vis(p)) \quad \forall p \in P$

Bew: Sei $p_{opt} \in P$ ein Punkt mit maximalem Sichtbarkeitspolygon

Klar (grobe Schranke): $\mu(vis(p_{opt})) \geq \frac{1}{n}$ (denn: in einer Triang. muß es Dreieck mit Fläche $\geq \frac{1}{n-2} > \frac{1}{n}$ geben)

(*)

Alg Geo 16.7.

$$\mu(vis(P_a)) \underset{\text{Approx}}{\geq} \frac{\# a \in A : P_a \text{ sieht } a}{|A|} - \frac{\delta}{2n} \geq \frac{\# a \in A : P_{opt} \text{ sieht } a}{|A|} - \frac{\delta}{2n}$$

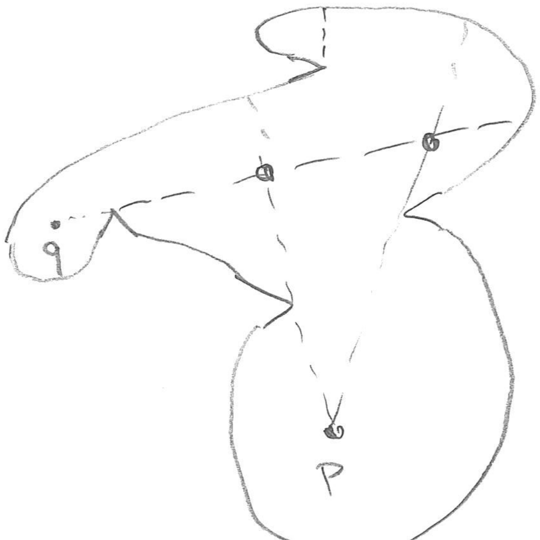
$$\underset{\text{Approx}}{\geq} \left(\mu(vis(P_{opt})) - \frac{\delta}{2n} \right) - \frac{\delta}{2n} = \mu(vis(P_{opt})) - \delta \frac{1}{n}$$

$$\geq \mu(vis(P_{opt})) - \delta \mu(vis(P_{opt})) = (1-\delta)\mu(vis(P_{opt}))$$

$\boxed{\text{Lemma}}$

<u>Also:</u> Um vis(p) zu maximieren, wollen wir P_a finden, das maximal viele Punkte aus A sieht

<u>Algorithmus</u>



* berechne vis(a) für jedes $a \in A$: $O(M \cdot n)$, $M = |A|$
* berechne das Arrangement dieser vis(a)

Komplexität: <u>$O(M^2 n)$</u> (nicht $M^2 n^2$)

<u>denn:</u> jedes vis(a) hat $\leq n$ Kanten

$\Rightarrow$ insgesamt gibt es $\leq M \cdot n$ Kanten

zum Glück: eine Kante von vis(a) kann ein vis(a') höchstens 2x schneiden

$\Rightarrow$ es gibt nur $O(M^2 n)$ Schnittpunkte (Knoten)

$\underset{\text{Euler}}{\Longrightarrow}$ nur $O(M^2 n)$ Zellen

* Berechnung des Arrangements mit Sweep

$$\mathcal{O}(M^2 n \log(M^2 n)) = \mathcal{O}\left(\frac{n^5}{\delta^4} \log^3 \frac{n}{\delta}\right)$$

$$M = c' \frac{n^2}{\delta^2} \log \frac{n}{\delta}$$

* Jetzt: jede Zelle einzeln anschauen,
  für jede Zelle: Anzahl der $vis(a)$, $a \in A$, die z enthalten

  $\to$ max ausgeben! p_a gefunden

<u>Theorem</u> Sei P einfaches Polygon mit n Ecken, $\delta > 0$. Dann kann man in Zeit $\mathcal{O}\left(\frac{n^5}{\delta^4} \log^3 \frac{n}{\delta}\right)$ einen Punkt $p_a \in P$ finden mit

$$P\left(\mu(vis(p_a)) \geq (1-\delta)\,\mu(vis(p_0))\right) \geq 1-\delta$$

„Monte-Carlo", (nicht Las Vegas).

AlgGeo 16.9

VC-Dimension stammt aus der statistischen Lerntheorie (Valiant '84)

PAC-Learning:

Gegeben: Beispielraum: X } Mengensystem
Menge von Konzepten: $\mathcal{F}$

Aufgabe: "Lerne" Konzept $T \in \mathcal{F}$

Algorithmus erhält Folge von markierten Beispielen
$(p_0, \delta_0), (p_1, \delta_1), \ldots$ mit

$$p_i \in X, \quad \delta_i = \begin{cases} 1, & p_i \in T \\ 0, & p_i \notin T \end{cases}$$

Die p_i werden i.i.d. nach unbekannter W'verteilung μ gezogen

Parameter: ε: Fehlerparameter
δ: Konfidenzparameter

Algorithmus liest Beispielfolge, wählt dann $S \in \mathcal{F}$ aus.

S approximately correct $:\Longleftrightarrow \mu(S \Delta T) \leq \varepsilon$

Alg ist probably approximately correct $:\Longleftrightarrow$

$$P(S \text{ approx. correct}) \geq 1-\delta.$$

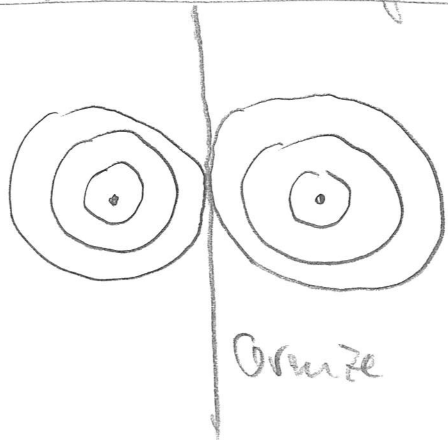
AlgGeo 16.10

Alg. heißt PAC-Alg., falls Obiges für alle μ und alle $T \in \mathcal{F}$ gilt.

Falls es PAC-Alg. gibt, heißt $(X, \mathcal{F})$ PAC-lernbar.

<u>Theorem</u>: $(X, \mathcal{F})$ PAC-lernbar $\iff$ $\dim_{VC}(X, \mathcal{F}) < \infty$.

<u>Jetzt: Voronoi-Diagramme</u>



R. Descartes:
Principia Philosophiae
(1633)