離散最適化基礎論　第 6 回
幾何ハイパーグラフ (2)：ε ネット

岡本 吉央
okamotoy@uec.ac.jp

電気通信大学

2017 年 12 月 1 日

最終更新：2017 年 12 月 1 日　11:04

主題

離散最適化のトピックの１つとして幾何的被覆問題を取り上げ，
その数理的側面と計算的側面の双方を意識して講義する

なぜ講義で取り扱う？

- 「離散最適化」と「計算幾何学」の接点として重要な役割を果たしているから
- 様々なアルゴリズム設計技法・解析技法を紹介できるから
- 応用が多いから

スケジュール 前半 (予定)

1. 幾何的被覆問題とは？ (10/6)
- ⋆ 国内出張のため休み (10/13)
2. 最小包囲円問題 (1)：基本的な性質 (10/20)
3. 最小包囲円問題 (2)：乱択アルゴリズム (10/27)
- ⋆ 文化の日のため休み (11/3)
4. クラスタリング (1)：k-センター (11/10)
5. 幾何ハイパーグラフ (1)：VC 次元 (11/17)
- ⋆ 調布祭 のため 休み (11/24)
6. 幾何ハイパーグラフ (2)：ε ネット (12/1)

注意：予定の変更もありうる

7. 幾何的被覆問題 (1)：線形計画法の利用 (12/8)
8. 幾何的被覆問題 (2)：シフト法 (12/15)
9. 幾何的被覆問題 (3)：局所探索法 (12/22)
10. 幾何的被覆問題 (4)：局所探索法の解析 (1/5)

⋆ センター試験準備 のため 休み (1/12)

11. 幾何ハイパーグラフ (3)：ε ネット定理の証明 (1/19)
12. 幾何アレンジメント (1)：合併複雑度と ε ネット (1/26)
13. 幾何アレンジメント (2)：合併複雑度の例 (2/2)
14. 最近のトピック (2/9)
15. 期末試験 (2/16？)

注意：予定の変更もありうる

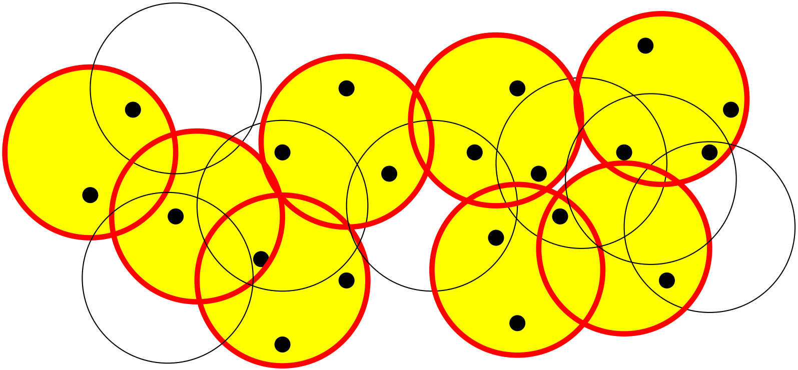
今日の内容

VC 次元と ε ネット

- ε ネットとは？
- ε ネット定理
- ε ネットの例
- ε ネットの要素数の上界

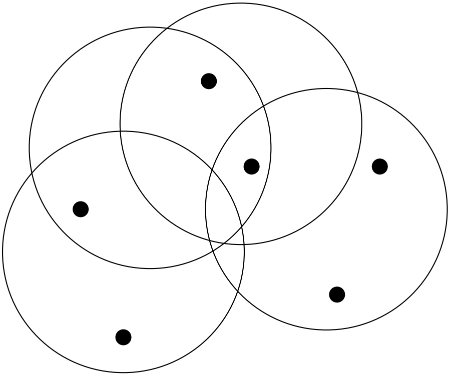
幾何的被覆問題の例 (1)

平面上にいくつかの点といくつかの単位円が与えられたとき
単位円を選んで，点をすべて覆いたい
選ばれる単位円の数を最も少なくするにはどうすればよいか？



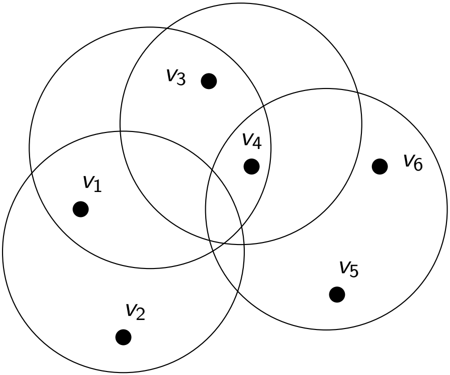
被覆問題としての定式化

- $V = \{v_1, v_2, v_3, v_4, v_5, v_6\}$
- $E = \{e_1, e_2, e_3, e_4\}$
- $e_1 = \{v_1, v_2\}$, $e_2 = \{v_1, v_3, v_4\}$, $e_3 = \{v_3, v_4\}$, $e_4 = \{v_4, v_5, v_6\}$



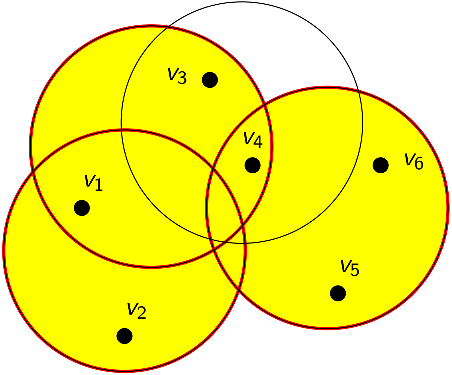
被覆問題としての定式化

- $V = \{v_1, v_2, v_3, v_4, v_5, v_6\}$
- $E = \{e_1, e_2, e_3, e_4\}$
- $e_1 = \{v_1, v_2\}$, $e_2 = \{v_1, v_3, v_4\}$, $e_3 = \{v_3, v_4\}$, $e_4 = \{v_4, v_5, v_6\}$



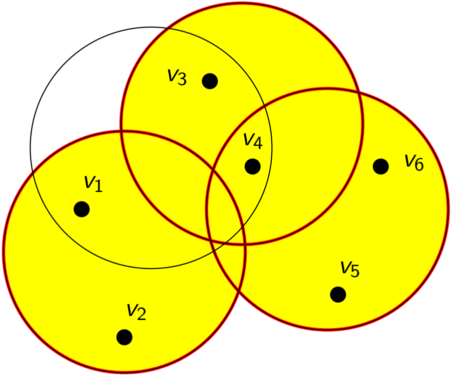
被覆問題としての定式化：最適解と最適値

- $E = \{e_1, e_2, e_3, e_4\}$
- $e_1 = \{v_1, v_2\}$, $e_2 = \{v_1, v_3, v_4\}$, $e_3 = \{v_3, v_4\}$, $e_4 = \{v_4, v_5, v_6\}$
- $E' = \{e_1, e_2, e_4\}$ は最適解で，3 が最適値



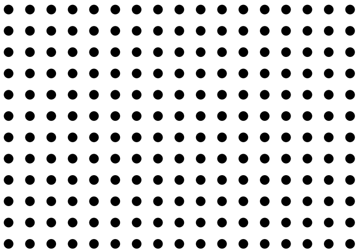
被覆問題としての定式化：最適解と最適値

- $E = \{e_1, e_2, e_3, e_4\}$
- $e_1 = \{v_1, v_2\}$, $e_2 = \{v_1, v_3, v_4\}$, $e_3 = \{v_3, v_4\}$, $e_4 = \{v_4, v_5, v_6\}$
- $E' = \{e_1, e_3, e_4\}$ も最適解で，3 が最適値



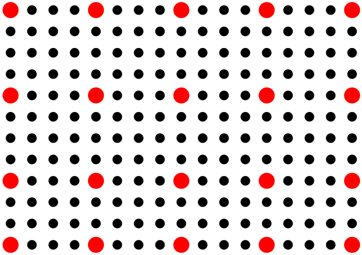
目次

平面上の点集合の粗視化



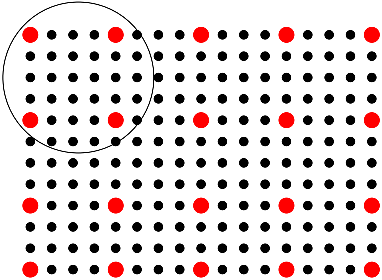
観点：あまり多くの点を含まない辺は無視する

平面上の点集合の粗視化



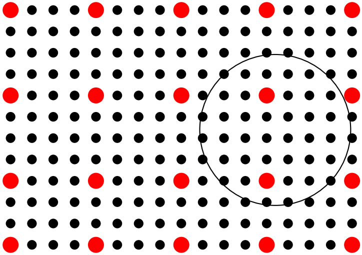
観点：あまり多くの点を含まない辺は無視する

平面上の点集合の粗視化



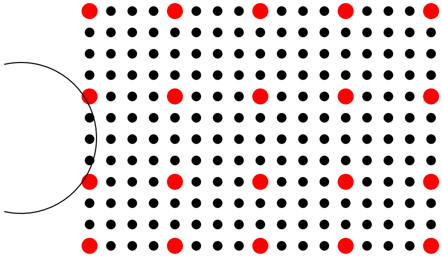
観点：あまり多くの点を含まない辺は無視する

平面上の点集合の粗視化



観点：あまり多くの点を含まない辺は無視する

平面上の点集合の粗視化



観点：あまり多くの点を含まない辺は無視する

ハイパーグラフに対する ε ネット

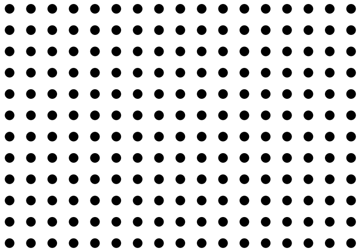
ハイパーグラフ $H=(V,E)$，実数 $\varepsilon \in [0,1]$

定義：ハイパーグラフに対する ε ネット (ε-net)

H に対する ε ネットとは，次を満たす集合 $N \subseteq V$ のこと

$$|e| \geq \varepsilon \cdot |V| \text{ を満たす任意の } e \in E \text{ に対して，} N \cap e \neq \emptyset$$

$|V| = 204$, $\varepsilon = 1/8$ とすると, $\varepsilon \cdot |V| = 25.5$ $\rightsquigarrow$ 1/8 ネット



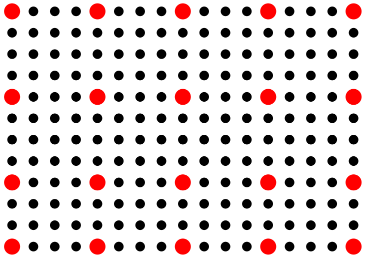
ハイパーグラフ $H=(V,E)$，実数 $\varepsilon \in [0,1]$

定義：ハイパーグラフに対する ε ネット (ε-net)

H に対する ε ネットとは，次を満たす集合 $N \subseteq V$ のこと

$$|e| \geq \varepsilon \cdot |V| \text{ を満たす任意の } e \in E \text{ に対して，} N \cap e \neq \emptyset$$

$|V|=204$, $\varepsilon=1/8$ とすると, $\varepsilon \cdot |V| = 25.5$ $\rightsquigarrow$ 1/8 ネット



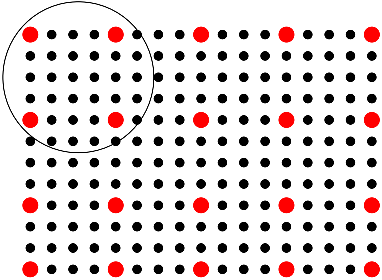
ハイパーグラフ $H=(V,E)$，実数 $\varepsilon \in [0,1]$

定義：ハイパーグラフに対する ε ネット (ε-net)

H に対する ε ネットとは，次を満たす集合 $N \subseteq V$ のこと

$$|e| \geq \varepsilon \cdot |V| \text{ を満たす任意の } e \in E \text{ に対して，} N \cap e \neq \emptyset$$

$|V| = 204$, $\varepsilon = 1/8$ とすると, $\varepsilon \cdot |V| = 25.5$ $\rightsquigarrow$ 1/8 ネット



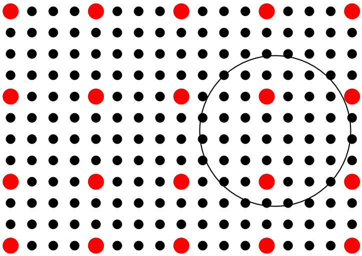
ハイパーグラフ $H = (V, E)$，実数 $\varepsilon \in [0, 1]$

定義：ハイパーグラフに対する ε ネット (ε-net)

H に対する ε ネットとは，次を満たす集合 $N \subseteq V$ のこと

$$|e| \geq \varepsilon \cdot |V| \text{ を満たす任意の } e \in E \text{ に対して，} N \cap e \neq \emptyset$$

$|V| = 204$, $\varepsilon = 1/8$ とすると, $\varepsilon \cdot |V| = 25.5$ $\rightsquigarrow$ 1/8 ネット



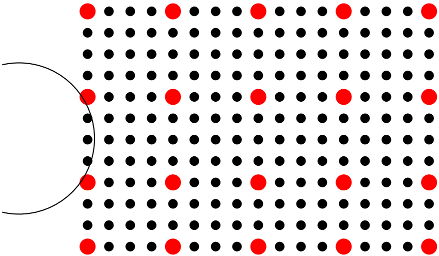
ハイパーグラフ $H = (V, E)$，実数 $\varepsilon \in [0, 1]$

定義：ハイパーグラフに対する ε ネット (ε-net)

H に対する ε ネットとは，次を満たす集合 $N \subseteq V$ のこと

$|e| \geq \varepsilon \cdot |V|$ を満たす任意の $e \in E$ に対して，$N \cap e \neq \emptyset$

$|V| = 204$, $\varepsilon = 1/8$ とすると, $\varepsilon \cdot |V| = 25.5$ $\leadsto$ 1/8 ネット



ハイパーグラフ $H = (V, E)$, 実数 $\varepsilon \in [0, 1]$

問題

H の ε ネットとして，どれくらい小さいものが作れるか？

- 小さければ小さいほどよい
- 小ささは ε に依存する？

問題に対する解答

ハイパーグラフ $H = (V, E)$, 実数 $\varepsilon \in (0, 1]$

定理：小さな ε ネットの存在性

要素数 $O\left(\frac{1}{\varepsilon} \log |E|\right)$ の ε ネットが存在する

定理：ε ネット定理 (Haussler, Welzl '87)

要素数 $O\left(\frac{d}{\varepsilon} \log \frac{d}{\varepsilon}\right)$ の ε ネットが存在する　　ただし，$d = \text{vc-dim}(H)$

$\therefore$ VC 次元が定数 $\Rightarrow$ ε ネットの最小要素数は $|V|$ や $|E|$ に依存しない！

注意

これらは多項式時間で構成できる ($O(|V||E|)$ 時間)

ε ネットの概念と ε ネット定理は次の論文による

- David Haussler, Emo Welzl: ε-Nets and Simplex Range Queries. Discrete & Computational Geometry 2: 127-151 (1987)

ハイパーグラフ $H = (V, E)$, 実数 $\varepsilon \in (0, 1]$

定理：ε ネット定理 (Haussler, Welzl '87)

要素数 $O\left(\frac{d}{\varepsilon}\log\frac{d}{\varepsilon}\right)$ の ε ネットが存在する　　ただし，$d = \mathrm{vc\text{-}dim}(H)$

H が幾何的に得られる場合，要素数を更に小さくできることもある

- 半平面から得られる場合：要素数 $= O\left(\frac{1}{\varepsilon}\right)$ (Komlós, Pach, Woeginger '92)
- 円から得られる場合：要素数 $= O\left(\frac{1}{\varepsilon}\right)$ (Matoušek, Seidel, Welzl '90)
- 軸平行長方形から得られる場合：要素数 $= O\left(\frac{1}{\varepsilon}\log\log\frac{1}{\varepsilon}\right)$ (Aronov, Ezra, Sharir '10)

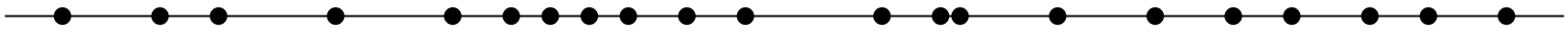
目次

ε ネットの例：区間

ハイパーグラフ $H = (V, E)$ として，次を考える

- $V \subseteq \mathbb{R}$，有限集合
- $E = \{V \cap [a, b] \mid a, b \in \mathbb{R}, a \leq b\}$

前回の講義の帰結：$\mathrm{vc\text{-}dim}(H) \leq 2$

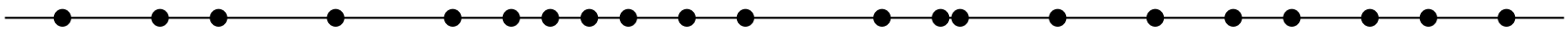


どのように ε ネットを構成できるか？

ε ネットの例：区間 — 構成法

構成法

左から順に V の要素を見て行き，$\lceil \varepsilon \cdot |V| \rceil$ 番目にあるものを次々と N に加えていく

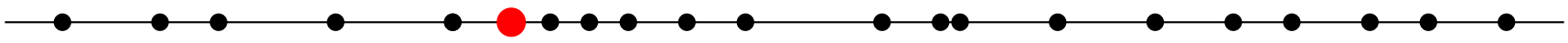


$|V| = 21$, $\varepsilon = 1/4$ のとき，$\varepsilon \cdot |V| = 5.25$

- すると，$\varepsilon \cdot |V|$ 個の頂点を含む区間は必ず N の要素を含む
- $|N| = \left\lfloor \dfrac{|V|}{\lceil \varepsilon \cdot |V| \rceil} \right\rfloor \leq 1/\varepsilon$

構成法

左から順に V の要素を見て行き，$\lceil \varepsilon \cdot |V| \rceil$ 番目にあるものを次々と N に加えていく

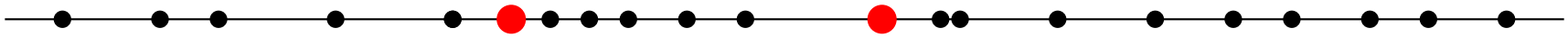


$|V| = 21$, $\varepsilon = 1/4$ のとき，$\varepsilon \cdot |V| = 5.25$

- すると，$\varepsilon \cdot |V|$ 個の頂点を含む区間は必ず N の要素を含む
- $|N| = \left\lfloor \dfrac{|V|}{\lceil \varepsilon \cdot |V| \rceil} \right\rfloor \leq 1/\varepsilon$

構成法

左から順に V の要素を見て行き，$\lceil \varepsilon \cdot |V| \rceil$ 番目にあるものを次々と N に加えていく

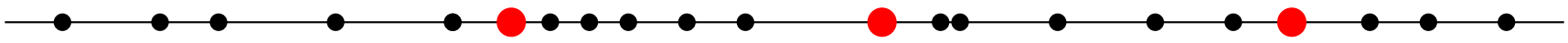


$|V| = 21$, $\varepsilon = 1/4$ のとき，$\varepsilon \cdot |V| = 5.25$

- すると，$\varepsilon \cdot |V|$ 個の頂点を含む区間は必ず N の要素を含む
- $|N| = \left\lfloor \dfrac{|V|}{\lceil \varepsilon \cdot |V| \rceil} \right\rfloor \leq 1/\varepsilon$

構成法

左から順に V の要素を見て行き，$\lceil \varepsilon \cdot |V| \rceil$ 番目にあるものを次々と N に加えていく

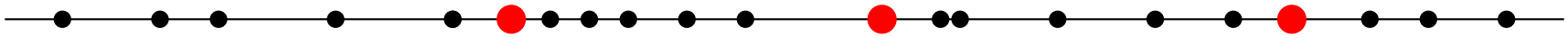


$|V| = 21$, $\varepsilon = 1/4$ のとき，$\varepsilon \cdot |V| = 5.25$

- すると，$\varepsilon \cdot |V|$ 個の頂点を含む区間は必ず N の要素を含む
- $|N| = \left\lfloor \dfrac{|V|}{\lceil \varepsilon \cdot |V| \rceil} \right\rfloor \leq 1/\varepsilon$

構成法

左から順に V の要素を見て行き，$\lceil \varepsilon \cdot |V| \rceil$ 番目にあるものを次々と N に加えていく



$|V| = 21$, $\varepsilon = 1/4$ のとき，$\varepsilon \cdot |V| = 5.25$

- すると，$\varepsilon \cdot |V|$ 個の頂点を含む区間は必ず N の要素を含む
- $|N| = \left\lfloor \dfrac{|V|}{\lceil \varepsilon \cdot |V| \rceil} \right\rfloor \leq 1/\varepsilon$

結論

数直線上の区間から得られるハイパーグラフに対して，要素数 $\dfrac{1}{\varepsilon}$ の ε ネットが存在する

目次

ハイパーグラフ $H = (V, E)$, 実数 $\varepsilon \in (0, 1]$

定理：小さな ε ネットの存在性

要素数 $O\left(\frac{1}{\varepsilon}\log|E|\right)$ の ε ネットが存在する

構成法：次の乱択アルゴリズムを考える

(1) $|e| < \varepsilon \cdot |V|$ を満たす辺 e を E から除去する
(残った辺の集合を E' とする)

(2) $p = \dfrac{c \ln|E'|}{\varepsilon \cdot |V|}$ とする (c は大きな定数)

(3) V の各要素を確率 p で独立に N へ入れる

(4) N を出力

合併上界

事象 A, B に対して

$$\Pr(A \cup B) \leq \Pr(A) + \Pr(B)$$

証明：

$$\Pr(A \cup B) = \Pr(A) + \Pr(B) - \Pr(A \cap B) \leq \Pr(A) + \Pr(B) \qquad \square$$

単純な不等式であるが，きわめて有用

マルコフの不等式

自然数値確率変数 $X \geq 0$ と正実数 $t > 0$ に対して，$\mathrm{E}[X]$ が存在するとき

$$\Pr(X \geq t) \leq \frac{\mathrm{E}[X]}{t}$$

<u>証明</u>：t が自然数である場合のみ示す (t が実数の場合も同様 → 演習問題)

$$\mathrm{E}[X] = \sum_{i=0}^{\infty} i \cdot \Pr(X = i)$$

マルコフの不等式

自然数値確率変数 $X \geq 0$ と正実数 $t > 0$ に対して，$\mathrm{E}[X]$ が存在するとき

$$\Pr(X \geq t) \leq \frac{\mathrm{E}[X]}{t}$$

証明：t が自然数である場合のみ示す (t が実数の場合も同様 → 演習問題)

$$\begin{aligned}
\mathrm{E}[X] &= \sum_{i=0}^{\infty} i \cdot \Pr(X = i) \\
&= \sum_{i=0}^{t-1} \underbrace{i}_{\geq 0} \cdot \Pr(X = i) + \sum_{i=t}^{\infty} \underbrace{i}_{\geq t} \cdot \Pr(X = i)
\end{aligned}$$

マルコフの不等式

自然数値確率変数 $X \geq 0$ と正実数 $t > 0$ に対して，$\mathrm{E}[X]$ が存在するとき

$$\Pr(X \geq t) \leq \frac{\mathrm{E}[X]}{t}$$

証明：t が自然数である場合のみ示す (t が実数の場合も同様 → 演習問題)

$$\begin{aligned}
\mathrm{E}[X] &= \sum_{i=0}^{\infty} i \cdot \Pr(X = i) \\
&= \sum_{i=0}^{t-1} \underbrace{i}_{\geq 0} \cdot \Pr(X = i) + \sum_{i=t}^{\infty} \underbrace{i}_{\geq t} \cdot \Pr(X = i) \\
&\geq \sum_{i=t}^{\infty} t \cdot \Pr(X = i)
\end{aligned}$$

マルコフの不等式

自然数値確率変数 $X \geq 0$ と正実数 $t > 0$ に対して，$\mathrm{E}[X]$ が存在するとき

$$\Pr(X \geq t) \leq \frac{\mathrm{E}[X]}{t}$$

証明：t が自然数である場合のみ示す (t が実数の場合も同様 → 演習問題)

$$\begin{aligned}
\mathrm{E}[X] &= \sum_{i=0}^{\infty} i \cdot \Pr(X = i) \\
&= \sum_{i=0}^{t-1} \underbrace{i}_{\geq 0} \cdot \Pr(X = i) + \sum_{i=t}^{\infty} \underbrace{i}_{\geq t} \cdot \Pr(X = i) \\
&\geq \sum_{i=t}^{\infty} t \cdot \Pr(X = i) = t \sum_{i=t}^{\infty} \Pr(X = i)
\end{aligned}$$

マルコフの不等式

自然数値確率変数 $X \geq 0$ と正実数 $t > 0$ に対して，$\mathrm{E}[X]$ が存在するとき

$$\Pr(X \geq t) \leq \frac{\mathrm{E}[X]}{t}$$

証明：t が自然数である場合のみ示す (t が実数の場合も同様 → 演習問題)

$$\begin{aligned}
\mathrm{E}[X] &= \sum_{i=0}^{\infty} i \cdot \Pr(X = i) \\
&= \sum_{i=0}^{t-1} \underbrace{i}_{\geq 0} \cdot \Pr(X = i) + \sum_{i=t}^{\infty} \underbrace{i}_{\geq t} \cdot \Pr(X = i) \\
&\geq \sum_{i=t}^{\infty} t \cdot \Pr(X = i) = t \sum_{i=t}^{\infty} \Pr(X = i) = t \cdot \Pr(X \geq t)
\end{aligned}$$

$\square$

構成法：次の乱択アルゴリズムを考える

(1) $|e| < \varepsilon \cdot |V|$ を満たす辺 e を E から除去する
(残った辺の集合を E' とする)

(2) $p = \dfrac{c \ln |E'|}{\varepsilon \cdot |V|}$ とする (c は大きな定数)

(3) V の各要素を確率 p で独立に N へ入れる

(4) N を出力

今から行うこと

高い確率で，次の 2 つの事象が同時に生起することの証明

1. 出力 N が H に対する ε ネットであること
2. $|N| = O(\frac{1}{\varepsilon} \log |E|)$ であること

辺 $e \in E'$ を固定すると，

$$\Pr(e \cap N = \emptyset) = (1-p)^{|e|} \le (1-p)^{\varepsilon \cdot |V|}$$

辺 $e \in E'$ を固定すると，

$$\Pr(e \cap N = \emptyset) = (1-p)^{|e|} \leq (1-p)^{\varepsilon \cdot |V|}$$

つまり，

$$\Pr(N \text{ が } \varepsilon \text{ ネットではない}) \quad = \quad \Pr(\exists\ e \in E' \colon e \cap N = \emptyset)$$

辺 $e \in E'$ を固定すると，

$$\Pr(e \cap N = \emptyset) = (1-p)^{|e|} \le (1-p)^{\varepsilon \cdot |V|}$$

つまり，

$$\begin{aligned} \Pr(N \text{ が } \varepsilon \text{ ネットではない}) &= \Pr(\exists\, e \in E' : e \cap N = \emptyset) \\ &\le |E'|(1-p)^{\varepsilon \cdot |V|} \qquad (\text{合併上界}) \end{aligned}$$

辺 $e \in E'$ を固定すると，

$$\Pr(e \cap N = \emptyset) = (1-p)^{|e|} \leq (1-p)^{\varepsilon \cdot |V|}$$

つまり，

$$\begin{aligned}
\Pr(N \text{ が } \varepsilon \text{ ネットではない}) &= \Pr(\exists\, e \in E' \colon e \cap N = \emptyset) \\
&\leq |E'|(1-p)^{\varepsilon \cdot |V|} \qquad (\text{合併上界}) \\
&\leq |E'| \exp(-p\varepsilon \cdot |V|)
\end{aligned}$$

補足：任意の実数 x に対して，$1 + x \leq \exp(x)$

辺 $e \in E'$ を固定すると，

$$\Pr(e \cap N = \emptyset) = (1-p)^{|e|} \le (1-p)^{\varepsilon \cdot |V|}$$

つまり，

$$\begin{aligned}
\Pr(N \text{ が } \varepsilon \text{ ネットではない}) &= \Pr(\exists\, e \in E' \colon e \cap N = \emptyset) \\
&\le |E'|(1-p)^{\varepsilon \cdot |V|} \qquad (\text{合併上界}) \\
&\le |E'| \exp(-p\varepsilon \cdot |V|) \\
&= |E'| \exp\left(-\frac{c \ln |E'|}{\varepsilon \cdot |V|} \varepsilon \cdot |V|\right)
\end{aligned}$$

補足：任意の実数 x に対して，$1 + x \le \exp(x)$

辺 $e \in E'$ を固定すると，

$$\Pr(e \cap N = \emptyset) = (1-p)^{|e|} \leq (1-p)^{\varepsilon \cdot |V|}$$

つまり，

$$\begin{aligned}
\Pr(N \text{ が } \varepsilon \text{ ネットではない}) &= \Pr(\exists\, e \in E' \colon e \cap N = \emptyset) \\
&\leq |E'|(1-p)^{\varepsilon \cdot |V|} \qquad (\text{合併上界}) \\
&\leq |E'| \exp(-p\varepsilon \cdot |V|) \\
&= |E'| \exp\left(-\frac{c \ln |E'|}{\varepsilon \cdot |V|} \varepsilon \cdot |V|\right) \\
&= \frac{1}{|E'|^{c-1}}
\end{aligned}$$

<u>補足</u>：任意の実数 x に対して，$1 + x \leq \exp(x)$

一方で，$\mathrm{E}[|N|] = p|V| = \dfrac{c \ln |E'|}{\varepsilon} = O\left(\dfrac{1}{\varepsilon} \log |E|\right)$ であり，
マルコフの不等式より，

$$\Pr(|N| \geq c\mathrm{E}[|N|]) \leq \frac{\mathrm{E}[|N|]}{c\mathrm{E}[|N|]} = \frac{1}{c}$$

$\square$

一方で，$\mathrm{E}[|N|] = p|V| = \dfrac{c \ln |E'|}{\varepsilon} = O\left(\dfrac{1}{\varepsilon} \log |E|\right)$ であり，
マルコフの不等式より，

$$\Pr(|N| \geq c\mathrm{E}[|N|]) \leq \frac{\mathrm{E}[|N|]}{c\mathrm{E}[|N|]} = \frac{1}{c}$$

ゆえに，N が ε ネットであり，かつ $|N| = O(\frac{1}{\varepsilon} \log |E|)$ を満たす確率は

$$1 - \frac{1}{|E'|^{c-1}} - \frac{1}{c}$$

以上 □

目次

今日の内容

VC 次元と ε ネット

- ε ネットとは？
- ε ネット定理
- ε ネットの例
- ε ネットの要素数の上界

ε ネット定理の証明は 1 月に行う予定

未解決問題

次のギャップを埋められるか？

円から作られるハイパーグラフに対する ε ネット

- 要素数 $\dfrac{13.4}{\varepsilon}$ の ε ネットが必ず存在 (上界)
  (Bus, Gupta, Mustafa, Ray '16)
- 要素数 $\dfrac{2}{\varepsilon} - 1$ を ε ネットが必要とする例が存在 (下界)
  (Komlós, Pach, Woeginger '92)

直線から作られるハイパーグラフに対する ε ネット

- 要素数 $O\left(\dfrac{1}{\varepsilon}\log\dfrac{1}{\varepsilon}\right)$ の ε ネットが必ず存在 (上界)
  (ε ネット定理)
- 要素数 $\dfrac{1}{2\varepsilon}\dfrac{\log^{1/3}\frac{1}{\varepsilon}}{\log\log\frac{1}{\varepsilon}}$ を ε ネットが必要とする例が存在 (下界)
  (Balogh, Solymosi '17)

残った時間の使い方

- 演習問題をやる
  - 相談推奨 (ひとりでやらない)
- 質問をする
  - 教員は巡回
- 退室時，小さな紙に感想など書いて提出する ← 重要
  - 内容は何でも OK
  - 匿名で OK

目次