PCP定理とその証明

清水 伸高 (東京科学大学)

Contents

Chapter 1

誤り訂正符号

誤り訂正符号とは文字列に冗長性を付加してノイズに対する頑健性を達成する手法である. 厳密には, **アルファベット**と呼ばれる有限集合 Σ 上の文字列 $x \in \Sigma^n$ を受け取り, より長い文字列

Chapter 2

エクスパンダーグラフ

2.1 定義

グラフ $G = (V, E)$ は, 単純ランダムウォークの非自明な第二固有値 $\lambda(P)$ が小さいときにエクスパンダーであるという. 多くの文脈では通常, 正則グラフに対してのみエクスパンダー性が定義されるが本講義では一般のグラフに対してエクスパンダー性を定義する.

定義 2.1.1 （エクスパンダー）

グラフ $G = (V, E)$ 上の単純ランダムウォークの遷移確率行列 P が $\lambda(P) \leq \lambda$ を満たすときグラフ G は **λ-エクスパンダー (λ-expander)** という. また, P の第二固有値が $\lambda_2 \leq \lambda$ を満たすとき, グラフ G は**片側 λ-エクスパンダー (one-sided λ-expander)** という.

本講義ではエクスパンダー性を持つ単体複体も取り扱うため, エクスパンダー性を持つグラフのことを**エクスパンダーグラフ**と呼んで区別する.

要するにグラフのエクスパンダー性とは単純ランダムウォークの混交時間が小さいという性質を意味する. 二部グラフは周期的であり特に最小固有値が $\lambda_{|V|} = -1$ となるためこの意味ではエクスパンダーグラフになりえないが, 片側エクスパンダーであるならば遅延単純ランダムウォークの混交時間は小さくなる.

ランダムウォークの混交時間が小さいとはランダムウォークが「すぐに混ざり合う」ことを意味する. この「すぐに混ざり合う」性質から, ランダムウォーク $(X_t)_{t\geq 0}$ が時刻 t までに訪れた頂点の集合を $U_t = \{X_0, \ldots, X_t\}$ とすると, $|U_t|$ はすぐに拡大 (expand) していく.

例 1 完全グラフ. グラフ $\left(V, \binom{V}{2}\right)$ を**完全グラフ (complete graph)** という. n 頂点完全グラフ上の単純ランダムウォークの遷移確率行列 P は, 単位行列 I と全成分が 1 の行列 J を用いて $P = \frac{1}{n-1}(J - I)$ で表せる. 完全グラフは正則グラフなので定常分布 π は V 上の一様分布である. 第一固有値は $\mathbf{1}$ であり, その他の固有ベクトル $x_i (i \geq 2)$ は全て $\mathbf{1}$ に直交し, 特に $Jx_i = 0$ となる. 従って $Px_i = -\frac{1}{n-1}x_i$ なので, $\lambda_1 = 1$, $\lambda_2 = \cdots = \lambda_n = -\frac{1}{n}$ である. よって, n 頂点完全グラフは $(1/n)$-expander であると同時に片側 $(-1/n)$-エクスパンダーグラフである ($\lambda(P)$ の定義では絶対値をつけているが片側エクスパンダー性の定義では絶対値をつけていないことに注意).

例 2 閉路グラフ. 頂点数 n の閉路グラフとは, 頂点集合 $V = \{v_1, \dots, v_n\}$ に対して辺集合 E が $E = \{\{v_1, v_2\}, \dots, \{v_{n-1}, v_n\}, \{v_n, v_1\}\}$ で与えられるグラフ (V, E) である. 頂点数 n が偶数のとき, 閉路グラフは二部グラフとなる.

ここでは $\omega = \exp\left(\frac{2\pi i}{n}\right)$ を 1 の冪根とし, 頂点集合を $V = \{\omega^i : i = 0, \dots, n-1\}$ とし, 各辺を $\{\omega^i, \omega^{i+1}\}$ で表す. 任意の関数 $f : V \to \mathbb{R}$ に P を作用させると

$$Pf(\omega^i) = \frac{f(\omega^{i-1}) + f(\omega^{i+1})}{2}$$

を得る. 関数 $x_k : \omega^j \mapsto \omega^{kj}$ を考えると, $Px_k(\omega_j) = \frac{\omega^{k(j-1)} + \omega^{k(j+1)}}{2} = \frac{\omega^{-k} + \omega^{-k}}{2} \cdot x_k(\omega_j)$ を得る. 従って各 $k = 0, \dots, n-1$ に対し x_k はそれぞれ固有値 $\frac{\omega^k + \omega^{-k}}{2} = \cos\frac{2\pi k}{n}$ に対応する固有ベクトルである.

これらの固有値を降順に並べて $1 = \lambda_1 \geq \cdots \geq \lambda_n \geq -1$ とすると, $\lambda_2 = \cos\frac{2\pi}{n} = 1 - \frac{4\pi^2}{2n^2} + O(n^{-4})$ である. 頂点数 n が偶数のときは $\lambda_n = -1$, 頂点数 n が奇数のときは $\lambda_n = \cos\pi\left(1 - \frac{1}{n}\right) = -\cos\frac{\pi}{n} = -1 + \frac{\pi^2}{2n^2} - O(n^{-4})$ である. 従って頂点数 n が奇数のときの閉路グラフは $\cos\frac{\pi}{n}$-エクスパンダーであり, $n \to \infty$ の漸近を考えると n のスペクトルギャップは $\Theta(1/n^2)$ となる.

2.2 存在性と陽な構成

エクスパンダー性はランダムウォークがすぐに混ざり合うということを意味し, これを満たすグラフは多くの辺を持つべきである. 例えば完全グラフは非常に強いエクスパンダー性を持つ一方で閉路グラフのエクスパンダー性は乏しい. では, 疎でありかつエクスパンダー性をもつグラフは存在するだろうか? また, 陽に構成できるだろうか?

2.2.1 ケイリーグラフ

エクスパンダーグラフを陽に構成する最も重要なアプローチの一つとして幾何学的群論におけるケイリーグラフと呼ばれる概念が知られている.

定義 2.2.1 (ケイリーグラフ)

G を有限群, $A \subseteq G$ を G の生成系[a]であって単位元を含まず, 逆元で閉じている (i.e., $A^{-1} = \{a^{-1} : a \in A\} = A$) ものとする. 頂点集合 $V = G$, 辺集合 $E = \{\{g, ag\} : g \in G, a \in A\}$ に対し (V, E) で与えられるグラフを**ケイリーグラフ (Cayley graph)** といい, $\mathrm{Cay}(A, G)$ で表す. 頂点 $g \in G$ に対し, $E_g = \{\{g, ag\} : a \in A\}$ を g を含む辺の集合とする.

[a] 任意の $g \in G$ に対してある有限個の $a_1, \dots, a_m \in A$ を用いて $g = \prod_{i=1}^{m} a_i$ と表せるとき, $A \subseteq G$ は生成系であるという.

ケイリーグラフは生成系 A の幾何学的な性質を調べる幾何学的群論における重要な研究対象の一つである. A は単位元を含まないため自己ループは存在しない. また, $A^{-1} = A$ より $\{ag, a^{-1}(ag)\} \in E$ となるため $\mathrm{Cay}(A, G)$ は無向グラフとなっている. 同様にケイリーグラフ $\mathrm{Cay}(G, A)$ も考えることができるが, 写像 $x \mapsto x^{-1}$ を考えると $\mathrm{Cay}(A, G)$ と $\mathrm{Cay}(G, A)$

が同型になっているので本質的には同じである. ケイリーグラフ $\mathrm{Cay}(A, G)$ は $|A|$-正則グラフである.

2.2.2 エクスパンダー性の限界とラマヌジャングラフ (*)

グラフの辺数を固定したとき, エクスパンダー性のパラメータ λ はどこまで小さくできるだろうか? ここでは厳密な証明は与えずに直感的な議論によって正則グラフに絞ってエクスパンダー性の限界を説明する.

あとの節 (セクション 2.4) で詳しく述べるが, 応用上は正則なエクスパンダーグラフが重要である. 正則グラフ上のランダムウォークの遷移確率行列は単に隣接固有値を次数で割ったものであり定常分布も一様分布なので単に隣接行列の固有値を考えれば良いことがわかる.[1]

固定した自然数 $d \geq 3$ に対して最もエクスパンダー性の強い (つまり λ が最小となる)d-正則グラフはどのようなグラフだろうか? 問題を言い換えればランダムウォークがより多くの頂点を訪れやすくするにはグラフをどのように構成すれば良いだろうか?

理想的なグラフ: d-正則無限木. 直感的な議論だが, 短い閉路があるとそれに沿って同じ頂点を訪れてしまうので, そのような閉路はない方が良いと思われる. 従ってそのグラフを虫眼鏡でズームすると局所的には木構造になっているべきであろう. そこで「理想的な」グラフとして d-正則で頂点数が無限の木 T を考える. 頂点集合 V は加算無限であるため, 有限グラフに対する隣接行列や固有値の概念を無限グラフに拡張したものが必要である. 集合 $\ell^2(V) \subseteq \mathbb{R}^V$ を $\ell^2(V) = \{f\colon V \to \mathbb{R}\colon \sum_{u\in V} f(u)^2 < \infty\}$ とする. 隣接作用素 $A\colon \ell^2(V) \to \ell^2(V)$ を

$$Af(u) = \sum_{v \in N_T(u)} f(v)$$

で定める. ここで $N_T(u) \subseteq V$ は T において u と隣接している頂点の集合であり, T の d-正則性から有限集合である. 作用素 A のスペクトルを

$$\mathrm{spec}(A) = \{\lambda \in \mathbb{R}\colon A - \lambda I \text{ は可逆でない}\}$$

とする.

定理 2.2.2

d-正則無限木 T の隣接作用素 A のスペクトルは以下を満たす:

$$\mathrm{spec}(A) \subseteq [-2\sqrt{d-1}, 2\sqrt{d-1}].$$

従って, 理想的なグラフを考えるとその隣接行列の非自明な固有値はその絶対値が高々 $2\sqrt{d-1}$ である (第一固有ベクトル $\mathbf{1}$ に対応する関数は $\ell^2(V)$ に属さない). よって任意の d-正則グラフは $\lambda(P) \geq \frac{2\sqrt{d-1}}{d}$ を満たすであろうことが予想される.

Alon-Boppana の定理. 定数次数の正則グラフの直径は $\Omega(\log n)$ を満たす.

[1] これらの理由からエクスパンダーグラフの理論は多くの教科書では正則グラフ上でのみ展開されている.

補題 2.2.3

$d \geq 3$ のとき, n 頂点 d-正則連結グラフ $G = (V, E)$ の直径は $\operatorname{diam}(G) \geq \log_{d-1} \frac{(d-2)(n-1)}{d}$ を満たす.

証明. 頂点 $u \in V$ を固定すると, u から ℓ 本以下の辺を辿って辿り着ける頂点は高々

$$1 + d + d(d-1) + \cdots + d(d-1)^{\ell-1} \leq 1 + d(d-1)^{\ell-1} \sum_{i=0}^{\ell-1} \left(\frac{1}{d-1} \right)^i \leq 1 + \frac{d(d-1)^\ell}{d-2}$$

この値が n より真に大きいとき, u から ℓ 本以下の辺を辿って辿り着けない頂点が存在し, これは $\operatorname{diam}(G) \geq \ell$ を意味する. これを解くと $\ell > \log_{d-1} \frac{(d-2)(n-1)}{d}$ を得る. □

定理 2.2.4 （Alon–Boppana の定理）

ある定数 $c > 0$ が存在し, 任意の n 頂点 d 正則グラフ G 上の単純ランダムウォークの遷移確率行列 P の第二固有値 λ_2 は

$$\lambda_2 \geq \frac{2\sqrt{d-1}}{d} \left(1 - \frac{c}{\operatorname{diam}(G)^2} \right)$$

を満たす.

特に, 補題 2.2.3 より, 次数 $d \geq 3$ を固定して頂点数 $n \to \infty$ の漸近において $\ell = \Omega(\log n)$ であるため, $\lambda_2 \geq \frac{2\sqrt{d-1}}{d}(1 - O(1/\log^2 n))$ を満たす. この結果は次数 d を固定したときの正則グラフのエクスパンダー性のパラメータの限界を表している.

ここでは少し弱い下界として

$$\lambda_2 \geq \frac{2\sqrt{d-1}}{d} \left(1 - O\left(\frac{\log \operatorname{diam}(G)}{\operatorname{diam}(G)} \right) \right) \tag{2.1}$$

を証明する. この下界でも $\lambda_2 \geq \frac{2\sqrt{d-1}}{d}(1 - o(1))$ を示すには十分である.

式 (2.1) の証明. 遷移確率行列を P とし, 隣接行列を A とする. 二頂点 u, v を uv 間の最短路が $\operatorname{diam}(G)$ に等しくなるように固定し関数 $f \colon V \to \mathbb{R}$ を $f = \delta_s - \delta_t$ とすると, 任意の $k \geq 1$

に対して

$$\begin{aligned}
\lambda(P^{2k}) &= \lambda(P^k)^2 \\
&\geq \frac{\mathbf{Var}_\pi[P^k f]}{\mathbf{Var}_\pi[f]} && \because ?? \\
&= \frac{\left\|P^k f\right\|_\pi^2}{\|f\|_\pi^2} && \because \mathbf{E}_\pi f = \mathbf{E}_\pi[P^k f] = 0 \\
&= \frac{\left\langle f, P^{2k} f \right\rangle_\pi}{\|f\|_\pi^2} && \because ?? \\
&= \frac{P^{2k}(u,u) + P^{2k}(v,v) - 2P^{2k}(u,v)}{2} \\
&= \frac{A^{2k}(u,u) + A^{2k}(v,v) - 2A^{2k}(u,v)}{2d^{2k}}. && \because P = \frac{1}{d}A
\end{aligned}$$

??より, $k = \left\lfloor \frac{\mathrm{diam}(G)-1}{2} \right\rfloor$ とすると, u, v の選び方より $A^{2k}(u,v) = 0$ である. ??より, $A^{2k}(u,u)$ は頂点 u を含む長さ $2k$ の閉路の個数に等しい. さらに, この値は d-正則無限木 T において固定した頂点を含む長さ $2k$ の閉路の個数で下から抑えることができる.

補題 2.2.5 （正則グラフの閉路数の下界）

$d \geq 3$ に対し, T を d-正則無限木とし, 頂点を一つ固定する. この頂点を含む長さ $2k$ の閉路の個数を t_{2k} とすると

$$t_{2k} = d(d-1)^{k-1} \cdot \frac{1}{k+1}\binom{2k}{k}$$

である. さらに, 任意の d-正則グラフ $G = (V, E)$ の任意の頂点 u に対し, u を含む長さ $2k$ の閉路の個数は少なくとも t_{2k} である.

まずは補題 2.2.5 を認めて定理 2.2.4 の証明を完成させる (補題 2.2.5 は後で証明する). 二項係数 $\binom{2k}{k}$ は Stirling の近似により $\binom{2k}{k} \geq \frac{4^k}{\sqrt{\pi k}}\left(1 - \frac{1}{8k}\right)$ を満たすことが示せる. 従って, 補題 2.2.5 より,

$$\begin{aligned}
\lambda(P)^{2k} &\geq d^{-2k} t_{2k} \\
&\geq d^{-2k} \cdot (d-1)^k \cdot \frac{2^{2k}}{(k+1)\sqrt{\pi k}}\left(1 - \frac{1}{8k}\right).
\end{aligned}$$

特に, ある定数 $c > 0$ が存在して

$$\begin{aligned}
\lambda(P) &\geq \frac{2\sqrt{d-1}}{d} \cdot k^{-c/k} \\
&\geq \frac{2\sqrt{d-1}}{d} \cdot \left(1 - O\left(\frac{\log k}{k}\right)\right).
\end{aligned}$$

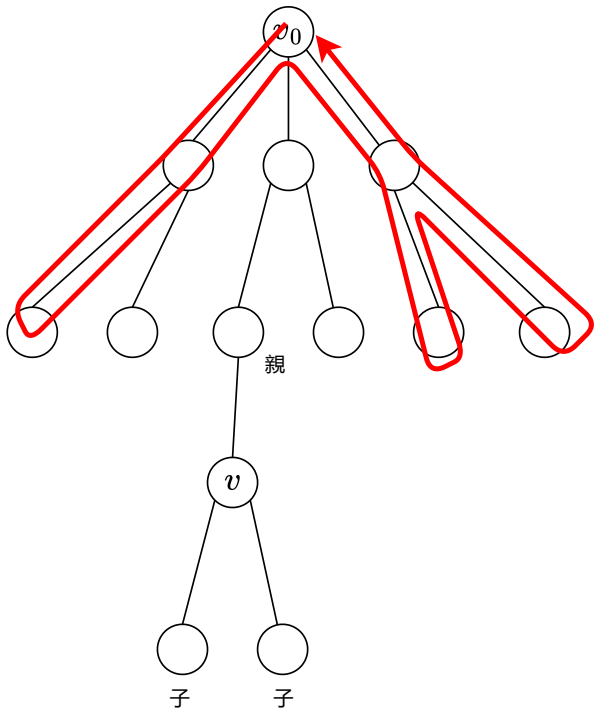

図 2.1: 3 正則木 T 上の長さ 10 の閉路. 親から子への遷移と子から親への遷移を 5 回ずつ行う.

最後の不等号は $k^{-k} = \mathrm{e}^{-\frac{\log k}{k}} = 1 - O\big(\frac{\log k}{k}\big)$ を用いた. $k = \left\lfloor \frac{\mathrm{diam}(G)-1}{2} \right\rfloor$ を代入すれば式 (2.1) を得る. □

最後に補題 2.2.5 を証明する.

補題 *2.2.5* の証明. d-正則無限木 T の特別な頂点 v_0 を一つ固定し, (グラフ理論においてスタンダードな) 幾つかの用語を定義する. 固定した特別な頂点 v_0 を**根 (root)** と呼び, T の頂点 v に対し, $\mathrm{dist}(v_0, v)$ を**深さ (depth)** と呼び $\mathrm{depth}(v)$ で表す (特に $\mathrm{depth}(v_0) = 0$ である). 頂点 v に T 上で隣接している d 個の頂点からなる集合を $N_T(v)$ と表す ($N_T(v)$ に v は含めない). これらの隣接頂点のうち, 深さが $\mathrm{depth}(v) - 1$ となるただ一つの頂点を v の**親 (parent)** と呼び, 残りの深さ $\mathrm{depth}(v) + 1$ の頂点を v の**子 (child)** と呼ぶ.

木 T において根を始点とする長さ $2k$ の閉路 $(v_0, \dots, v_{2k})$ を考える (ここで $v_0 = v_{2k}$). 各 i に対して $d_i = \mathrm{depth}(v_i)$ とし, 列 $(d_0, \dots, d_{2k})$ を考える. まず $d_0 = 0$ であり, その後は $d_{i+1} \in \{d_i \pm 1\}$ であり, 常に非負性を保ちながら最後に $d_{2k} = 0$ となる. このような列 $(d_0, \dots, d_{2k})$ の総数はカタラン数と等しく, $\frac{1}{k+1}\binom{2k}{k}$ に等しい. 特に, 各 $d_i - d_{i-1}$ の符号を見ると $d_0 = d_{2k} = 0$ より正と負がそれぞれ k 個ずつ含まれている.

次に, 各 $(d_1, \dots, d_{2k})$ に対して深さの履歴がこれと等しくなるような閉路の個数を考える. まず, $i = 1$ において, 頂点 v_0 からは子に遷移するため v_1 の取り方は d 通りある. 各 $i \geq 2$ において, $d_i = d_{i-1} + 1$ (すなわち v_i が v_{i-1} の子である) とき, v_i の選び方は $d - 1$ 通りある. 一方で親に遷移する場合はその遷移先は一意である. 子への遷移はちょうど k 回発生するため, 深さの履歴が与えられた $(d_1, \dots, d_{2k})$ に等しくなるような閉路の個数は $d(d-1)^{k-1}$ に等しい. 従って $t_{2k} = \frac{1}{k+1}\binom{2k}{k} \cdot d(d-1)^{k-1}$ を得る.

後半の主張を証明する. d-正則グラフ G の頂点 u_0 を一つ固定する. 木 T の頂点集合を U, グラフ G の頂点集合を V とし, N_T の定義と同様にグラフ G の頂点 $v \in V$ に対しその d 個の隣接頂点の集合を $N_G(v)$ で表す. G から T への準同型写像 $\phi\colon U \to V$ を以下のように構成する[2]: 深さに関して帰納的に定義する.

- まず, $\phi(v_0) = u_0$ とする. 根 v_0 の子 $N_T(v_0)$ から $N_G(u_0)$ への全単射 ϕ_0 を任意に一つ選び, $\phi\colon N_T(v_0) \ni v' \mapsto \phi_1(v') \in N_T(u_0)$ によって $N_T(v_0)$ における ϕ を定義する.
- T における深さ ℓ 以下の全ての頂点に対し $\phi(v)$ が定義されているとする. 深さ ℓ の各頂点 v に対し, その親を p, 子を $c_1, \dots, c_{d-1}$ とする. v とその親 p に対しては $u := \phi(v)$, $u_p := \phi(p) \in U$ が定義されている. このとき, 全単射 $\psi\colon N_T(v) \setminus \{p\} \to N_G(u) \setminus \{u_p\}$ を任意に一つ固定し, 各 $\phi(c_j)$ を $\psi(c_j) \in V$ とする.

このようにして定義された写像 $\phi\colon U \to V$ は確かに準同型なので T の閉路 $(v_0, \dots, v_{2k})$ に対して $(\phi(v_0), \dots, \phi(v_{2k}))$ は G の閉路になっている. さらに, $(\phi(v_0), \dots, \phi(v_{2k}))$ の形になっている G の閉路が与えられたとき, v_0 は一意に定まり, 以降の v_i は ϕ の帰納的な定義で用いた全単射の逆写像を用いて順番に復元することができるため, $(v_0, \dots, v_{2k}) \mapsto (\phi(v_0), \dots, \phi(v_{2k}))$ は単射である. 従ってグラフ G に含まれるある頂点を始点とした長さ $2k$ の閉路の個数は少なくとも t_{2k} である. □

ラマヌジャングラフ. 漸近的に定理 2.2.4 を達成するグラフを**ラマヌジャングラフ (Ramanujan graph)** という.

定義 2.2.6 （ラマヌジャングラフ）

d-正則グラフ $G = (V, E)$ は, その単純ランダムウォークの遷移確率行列 P の第二固有値 λ_2 が $\lambda_2 \leq 2\sqrt{d-1}$ を満たすとき, **ラマヌジャングラフ (Ramanujan graph)** という.

定理 2.2.4 を達成するグラフ列, すなわち, 次数 d を固定したときに頂点数が増大していくグラフ列 $(G_n)_{n\in\mathbb{N}}$ であって各 G_n が d-正則ラマヌジャングラフとなるものは存在するだろうか? この漸近的に最適な正則エクスパンダーグラフの構成は Lubotzky, Phillips, and Sarnak [LPS88] and Margulis [Mar88] によって独立同時期に初めてその構成が与えられた. 彼らは $d-1$ が 4 で割った余りが 1 となる素数であるときに d-正則ラマヌジャングラフの列を構成した. なお, 「ラマヌジャングラフ」という名称は [LPS88] の証明がラマヌジャン予想と呼ばれる予想に依拠しているからである (「予想」と書いたが当時は既に解決している). その後, Morgenstern [Mor94] によって次数が素数べき $+1$ の形であってもラマヌジャングラフが構成できることが示された.

定理 2.2.7 （ラマヌジャングラフの陽な構成）

任意の素数 q と任意の $k \in \mathbb{N}$ に対して, 頂点数が発散するある (q^k+1)-正則ラマヌジャングラフの列が存在し, 陽に構成できる.

[2] グラフ $G = (V, E)$ から $H = (W, F)$ への**準同型写像 (homomorphism)** とは, 写像 $\phi\colon V \to W$ であって $\{u, v\} \in E \Rightarrow \{\phi(u), \phi(v)\} \in F$ を満たすものである. ここでは自然に無限グラフに対してこの概念を拡張している.

ランダム正則グラフ. Lubotzky, Phillips, and Sarnak [LPS88] やその後続研究によりラマヌジャングラフ列については様々な構成方法が知られている. では, そもそもラマヌジャングラフは何個あるのだろうか? n 頂点 d-正則グラフ全体の集合を $\mathcal{G}_{n,d}$ とし, $\mathcal{G}_{n,d}$ から一様ランダムに選ばれたグラフ $G \sim \mathcal{G}_{n,d}$ を考える (nd は常に偶数とする). この確率変数をランダム正則グラフという. ランダム正則グラフは「ほぼ」ラマヌジャングラフであることが知られている [Fri08].

定理 2.2.8 (Friedman の定理)

任意の $d \geq 3$ と任意の $\varepsilon > 0$ に対し,

$$\lim_{n\to\infty} \Pr\left[\lambda(P) \geq \frac{2\sqrt{d-1}}{d} + \varepsilon\right] = 0.$$

すなわち, ほとんど全ての定数次数正則グラフはラマヌジャングラフと同等のスペクトルを持つ.

2.3 性質

エクスパンダーグラフは固有値によって定義されるが, 様々な興味深い性質を持つことが知られている. ここではグラフ理論的な性質, ランダムウォークに関する性質, そして擬似ランダム性について紹介する.

2.3.1 グラフ理論的な性質

グラフ理論的に非常に興味深い多くの性質を有する.

Cheeger, chromatic number, diameter

2.3.2 ランダムウォークの性質

hitting time, cover time

2.3.3 擬似ランダム性 (*)

この節は高次元エクスパンダーの本筋から少し外れるが, エクスパンダーグラフの重要であることの理由の一つとしてその擬似ランダム性について概説する.

加法的組合せ論や計算量理論では**擬似ランダム性 (pseudorandomness)** と呼ばれる概念が非常に重要な役割を果たしている.

定義 2.3.1 (分布の擬似ランダム性)

有限集合 Ω 上のある分布 μ と関数族 $\mathcal{F} = \{f : \Omega \to \{0,1\}\}$ を考える. 分布 μ は, 任意

の $f \in \mathcal{F}$ に対して

$$\left| \mathop{\mathbb{E}}_{x\sim\mu}[f(x)] - \mathop{\mathbb{E}}_{y\sim U_\Omega}[f(y)] \right| \leq \varepsilon$$

を満たすとき, $\mathcal{F}$ **に対して** ε**-擬似ランダムである**という (ここで, $y \sim U_\Omega$ とは Ω 上一様ランダムに y が選ばれたことを意味する).

直感的には, 分布が擬似ランダムであるとは, その分布が任意の $f \in \mathcal{F}$ を使っても一様分布と**識別できない (indistinguishable)** ことを意味する. 例えば全変動距離に関する**??**では, $\mathcal{F}$ を V 上の二値関数全体 (すなわち任意の V の部分集合) の族としたときの識別不可能性のパラメータ ε が全変動距離で与えられることを意味する. すなわち μ は常に $d_{\mathrm{TV}}(\mu, U_\Omega)$)-擬似ランダムである. 関数クラス $\mathcal{F}$ をより制限したときにパラメータ ε がどこまで小さくなるかに興味がある.

組合せ論では $\mathcal{F}$ としてある特殊な関数クラスを仮定することによって**組合せ論的擬似ランダム性**を定義する. 例えばグラフ理論や加法的組合せ論のコーナーストーンの一つと呼ばれる Szemerédi の正則化補題と呼ばれる結果は, 非常に大雑把に言えば任意の密なグラフが定数個の擬似ランダムな二部グラフと疎な部分に分解できることを主張する定理である. 組合せ論的擬似ランダムネスの概念は特に加法的組合せ論において非常な協力な道具となっており, Green–Tao の定理の証明においても重要な役割を果たしている (驚くべきことに, 識別不可能性の枠組みで Green–Tao の定理の証明を理解してそれを学習理論におけるブースティングの証明に応用するという研究もなされている!).

計算量理論では $\mathcal{F}$ を「効率的なアルゴリズムの全体」や「素子数の少ない論理回路の全体」とすることで**計算量的擬似ランダム性**を定義できる. 任意の効率的なアルゴリズムに対して一様ランダムな文字列と識別できないということは, その分布に従って生成されたメッセージを盗み見てもそこから得られる情報が何もない (ランダムな文字列を見てるのと同じ) であることから, 計算量的擬似ランダム性は暗号の計算量的安全性の定義の根幹をなすことがわかる.

エクスパンダーグラフの組合せ論的擬似ランダム性を説明する. 正則 λ-エクスパンダー $G = (V, E)$ を考える. 集合 $\Omega = V \times V$ 上の分布 $\mu = \mu_G$ として一様ランダムな辺 $\{u, v\} \in E$ を選び, (u, v) もしくは (v, u) どちらかを等確率で選んだ時の頂点対の分布とする. すなわち,

$$\Pr_{(u,v)\sim\mu}[(u, v) = (s, t)] = \frac{\mathbf{1}_{\{s,t\}\in E}}{2|E|} = \frac{\mathbf{1}_{\{s,t\}\in E}}{nd} \tag{2.2}$$

とする. 関数族 $\mathcal{F}$ を

$$\mathcal{F} = \{f_{S,T} \colon (s, t) \mapsto \mathbf{1}_{s\in S, t\in T} \colon S, T \subseteq V\} \tag{2.3}$$

で定める.

補題 2.3.2 (エクスパンダー混交補題)

グラフ G が n 頂点 d-正則 λ-エクスパンダーであるとき, 式 (2.2) で定義された分布 μ は式 (2.3) で定義された関数族 $\mathcal{F}$ に関して ε-擬似ランダムである. すなわち, 任意の頂点部分集合 $S, T \subseteq V$ に対して, $e(S, T) = \sum_{s\in S, t\in T} \mathbf{1}_{\{s,t\}\in E}$ を S, T 間の辺の本数 ($S \cap T$

内の辺は 2 回数える) とすると,

$$\left|e(S,T) - \frac{d}{n}|S||T|\right| \leq d\lambda\sqrt{|S||T|\left(1-\frac{|S|}{n}\right)\left(1-\frac{|T|}{n}\right)}.$$

グラフ G の隣接行列 A を考えるとイメージしやすい. この行列は全部で nd 個の 1 を持っているため, 全成分の中で 1 の密度は $\frac{d}{n}$ である. ここで, 部分集合 $S,T \subseteq V$ に対して A の $S \times T$ で定まる部分行列 $A_{S,T}$ を考える. この行列に含まれる 1 の個数 ($e(S,T)$ に等しい) は, $\frac{d}{n} \cdot |S||T|$ に近い値となっている (図 2.2).

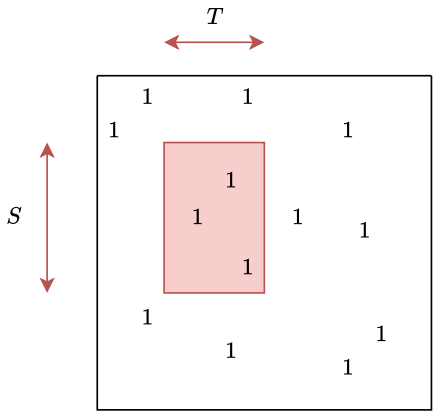

図 2.2: 正則グラフの隣接行列 A を考える. このグラフがエクスパンダーならば, 頂点部分集合 $S,T \subseteq V$ で指定される長方形内に含まれる 1 の密度は行列全体の 1 の密度に近い値をとる.

補題 *2.3.2* の証明. グラフ G 上の単純ランダムウォーク P を考える. 部分集合 $S,T \subseteq V$ に対し関数 $f = \delta_S, g = \delta_T$ として**??**を適用すると

$$\begin{aligned}
\langle f, Pg\rangle_\pi &= \frac{e(S,T)}{nd},\\
\mathbf{E}_\pi f &= \frac{|S|}{n},\\
\mathbf{E}_\pi[Pg] = \mathbf{E}_\pi g &= \frac{|T|}{n},\\
\mathbf{Var}_\pi f &= \frac{|S|}{n}\left(1-\frac{|S|}{n}\right),\\
\mathbf{Var}_\pi g &= \frac{|T|}{n}\left(1-\frac{|T|}{n}\right)
\end{aligned}$$

より整理すると主張を得る. □

演習問題 1

補題 2.3.2 の証明で表した五つの等式を実際に確認せよ.

2.4 エクスパンダーグラフの応用

グラフのエクスパンダー性は組合せ論的な興味だけでなく, 理論計算機科学において多くの定理の証明の道具として非常に重要な役割を果たしている. ここではその一端を軽く紹介する. より詳細の議論は [HLW06] を参照されたい.

2.4.1 脱乱択化

Albert Einstein は「量子は確率的に振る舞う」とする量子力学の枠組みに対して懐疑的であり, 1926 年に Max Born に宛てた手紙において

> Der Alte würfelt nicht. (神はサイコロを振らない)

と述べている. では, アルゴリズムの神はサイコロを振るだろうか? より具体的には, 乱択は計算能力を真に向上させるだろうか? この哲学的な問いは 90 年代から今もなお計算量理論において深く研究されており, その中心的なリーダーの一人である Avi Wigderson は 2021 年に Abel 賞, 2023 年に Turing 賞を受賞している.

ここではエクスパンダーグラフを使って「少ないサイコロで多くのサイコロの出目を hitting 性の意味で模倣できる」という結果を紹介する.

2.4.2 誤り訂正符号

誤り訂正符号 (error-correcting code) または単に**符号 (code)** とは文字列に冗長性を持たせることでノイズに対する頑健性を与える手法である. 数学的には符号はビット列の集合 $\mathcal{C} \subseteq \mathbb{F}_2^n$ であり, その元 $f \in \mathcal{C}$ を**符号語 (codeword)** と呼ぶ[3]. ここで n を符号 $\mathcal{C}$ の**符号長 (code length)** と呼ぶ. 符号には, 任意の相異なる二つの符号語が互いにハミング距離の意味で離れていることが望まれる. 形式的には, 正規化されたハミング距離 $\mathrm{dist}(f,g) = n^{-1}\sum_{i\in[n]} \mathbf{1}_{f(i)\neq g(i)} = \Pr_{i\sim[n]}[f(i) \neq g(i)]$ を考え, $\min_{f\neq g\in\mathcal{C}} \mathrm{dist}(f,g)$ を符号 $\mathcal{C}$ の**距離 (distance)** という. 文字列 $f \in \mathbb{F}_2^n$ と $\mathcal{C} \subseteq \mathbb{F}_2^n$ に対して $\mathrm{dist}(f,\mathcal{C}) = \min_{w\in\mathcal{C}} \mathrm{dist}(f,w)$ を f の $\mathcal{C}$ への距離とする.

符号 $\mathcal{C} \subseteq \mathbb{F}_2^n$ が線形部分空間となるとき, 符号 $\mathcal{C}$ を**線形符号 (linear code)** と呼ぶ. 文脈によっては線形符号のことを単に符号と呼ぶこともあり, 本講義も以降は特に断りのない限りこの慣習に従う. すなわち符号と言えばそれは線形部分空間を意味する. 符号長 n, ランク k の線形符号に対し, k/n をその符号の**レート (rate)** と呼ぶ. 直感的には符号のレートはその符号が空間 $\mathbb{F}_2^n$ 内でどれほど密に充填しているかを表すため, 符号のレートと距離にはトレードオフがある. 線形符号 $\mathcal{C}$ の距離は最小ハミング重みを持つ非ゼロの符号語によって与えられることに注意されたい.

ここでは, ケイリーグラフ (定義 2.2.1) を用いて構成される符号を紹介する.

[3] 文脈によってはビット列の代わりに有限集合 Σ に対して $\mathcal{C} \subseteq \Sigma^n$ を符号と定義することもある. 実際には計算機上では Σ の元を $\lceil \log_2 |\Sigma| \rceil$ ビットで表すため $\Sigma = \mathbb{F}_2$ とすることが多い.

定義 2.4.1 （ケイリーエクスパンダー符号）

ケイリーグラフ $\mathrm{Cay}(A,G)=(V,E)$ と符号 $\mathcal{C}_A\subseteq\mathbb{F}_2^A$ を考える．頂点 $g\in V$ と関数 $f\colon E\to\mathbb{F}_2$ に対し，$f_g=(f(\{g,ag\}))_{a\in A}\in\mathbb{F}_2^A$ と定める．符号 $\mathcal{C}_A\subseteq\mathbb{F}_2^A$ に対して

$$\mathcal{C}(A,G,\mathcal{C}_A)=\{f\in\mathbb{F}_2^E\colon\forall g\in V, f_g\in\mathcal{C}_A\}$$

を**ケイリーエクスパンダー符号**と呼ぶ．

一般の (ケイリーグラフとは限らない) 正則エクスパンダーグラフを用いて定義される**エクスパンダー符号 (expander code)** が有名だが，定義の簡潔さを優先してあえてケイリーグラフに限定したエクスパンダー符号を紹介した．もし仮に A を生成系とするケイリーグラフの列 $(\mathrm{Cay}(A,G_n))_{n\in\mathbb{N}}$ とレートと距離の良い性質をもつ一つの符号 $\mathcal{C}_A$ があったとしよう．すると，この一つの符号から符号の列 $(\mathcal{C}_n)_{n\in\mathbb{N}}:=(\mathcal{C}(A,G_n,\mathcal{C}_A))_{n\in\mathbb{N}}$ を構成できる．

さらに興味深いことに，構成に用いたケイリーグラフがエクスパンダー性を持つならば元の符号 $\mathcal{C}_A$ のレートと距離の性質を符号列 $(\mathcal{C}_n)$ も受け継ぐ．

補題 2.4.2

$\mathcal{C}_A$ のレートが r_A ならば $\mathcal{C}(A,G,\mathcal{C}_A)$ のレートは少なくとも $2r_A-1$ である．

証明．符号 $\mathcal{C}_A\subseteq\mathbb{F}_2^A$ のレートが r_A なので，その任意の符号語 $f_0\in\mathcal{C}_A$ は $|A|(1-r_A)$ 個の線形制約を満たす．ケイリーエクスパンダー符号 $\mathcal{C}(A,G,\mathcal{C}_A)$ の符号語 f は，全ての頂点 $g\in G$ に対して f_g が $|A|(1-r_A)$ 個の線形制約を満たしているので，f は高々 $|G||A|(1-r_A)$ 個の線形制約を満たしている．つまり f の自由度は少なくとも $|E|-|G||A|(1-r_A)=|E|(1-2(1-r_A))$ なので，$\mathcal{C}(A,G,\mathcal{C}_A)$ のレートは少なくとも $1-2(1-r_A)=2r_A-1$ となる． □

補題 2.4.3

符号 $\mathcal{C}_A$ の距離が δ_A，ケイリーグラフ $\mathrm{Cay}(A,G)$ が λ-エクスパンダーならば，ケイリーエクスパンダー符号 $\mathcal{C}(A,G,\mathcal{C}_A)$ の距離は少なくとも $\delta_A(\delta_A-\lambda)$ である．

証明にはエクスパンダー混交補題からすぐに従う以下の系を用いる:

系 2.4.4

d-正則な λ-エクスパンダー $G=(V,E)$ の頂点部分集合 $S\subseteq V$ が $e(S,S)\geq cd|S|$ を満たす (言い換えると，誘導部分グラフ $G[S]$ の平均次数が cd 以上) ならば，$|S|\geq(c-\lambda)n$ を満たす．

系 *2.4.4* の証明．$S=T$ として補題 2.3.2 を適用すると

$$cd|S|\leq e(S,S)\leq\frac{d}{|V|}|S|^2+\lambda d|S|.$$

これを解くと $|S|\geq(c-\lambda)n$ を得る． □

補題 2.4.3 の証明. 任意の非ゼロの符号語 $f \in \mathcal{C}(A, G, \mathcal{C}_A)$ が少なくとも $\delta_A(\delta_A - \lambda)|E|$ 個の 1 を持つことを言えばよい. 符号語 $f \neq 0$ に対し, $F = \{e \in E \colon f(e) = 1\}$ とし, $S = \bigcup_{e \in F} e$ を F の辺と接続している頂点の全体とする (辺 e を要素数 2 の頂点部分集合として見ている). $|F| \geq \delta_A(\delta_A - \lambda)|E|$ を示せばよい.

$\mathcal{C}_A$ の距離の条件より, 各 $g \in S$ に対して $f_g \in \mathbb{F}_2^A$ は少なくとも $\delta_A|A|$ 本の辺が接続している. すなわち, $\mathrm{Cay}(A, G)$ の部分グラフ (S, F) の最小次数は $\delta_A|A|$ を満たすので $|F| \geq \frac{\delta_A|A|}{2}|S|$. 誘導部分グラフ $G[S]$ は (S, F) を部分グラフとして含むので $e(S, S) \geq 2|F|$. さらに (S, F) に対する握手補題より

$$e(S, S) \geq 2|F| \geq \delta_A|A||L|.$$

系 2.4.4 より $|S| \geq (\delta_A - \lambda)|G|$ なので, $|F| \geq \frac{\delta_A|A|}{2}|S| \geq \delta_A(\delta_A - \lambda)\frac{|G||A|}{2} = \delta_A(\delta_A - \lambda)|E|$ を得る. □

下論文の二部エクスパンダーに基づく定義との比較

2.4.3 PCP 定理

2.4.4 Goldreich の擬似乱数生成器

Chapter 3

高次元エクスパンダー概論

高次元エクスパンダーとはグラフのエクスパンダー性を単体複体に拡張した概念である. 単体複体上では, 大域的なランダムウォークと局所的なランダムウォークの二つのタイプのランダムウォークを自然に考えることができ, これらに基づいてそれぞれ大域的なエクスパンダー性と局所的なエクスパンダー性の概念が定義できる. 端的に述べると高次元エクスパンダーの理論はこれら二つの概念がほぼ等価であることを明らかにしており, これは単体複体における「局所大域原理」[1]を体現しているといえる.

本チャプターではまず単体複体とその上でのランダムウォークを定義し, これに基づいて高次元エクスパンダーの定義と重要な性質を紹介する.

3.1 定義

まずは単体複体に関する基礎的な用語を定義していく. 文脈によっては単体複体は多面体などを貼り合わせた幾何的な概念を指すこともあるが, 本講義では組合せ的ないわゆる set system (ハイパーグラフ) としての単体複体を扱う.

定義 3.1.1 （単体複体）

有限集合 V と V の部分集合族 $\mathbb{F} \subseteq 2^V$ であって部分集合で閉じているもの (すなわち, $\sigma \subseteq \tau \in \mathbb{F} \Rightarrow \sigma \in \mathbb{F}$) の組 $X = (V, \mathbb{F})$ を**単体複体 (simplicial complex)** という. 集合族 $\mathbb{F}$ の元を**面 (face)** と呼び, 面 $\sigma \in \mathbb{F}$ の**次元 (dimension)** を $\dim \sigma = |\sigma| - 1$ とする[a]. 単体複体 X の次元を $\dim X = \max\{\dim \sigma \colon \sigma \in \mathbb{F}\}$ とする.
次元 d の単体複体 X は (包含関係に関して) 極大な面の次元が全て d に等しいとき, **純粋 (pure)** であるという. 整数 $-1 \leq k \leq \dim X$ に対し $X(k) = \{\sigma \in \mathbb{F} \colon \dim \sigma = k\}$ とする. 特に断りのない限り, $X(0) = V$ を仮定する (そうでなければ V として $V = X(0)$ とした単体複体を考える).

[a] 特に, 空集合 $\emptyset \in \mathbb{F}$ の次元は -1 である.

面の次元の概念は単体複体の幾何的な表現に由来する. このイメージになぞらえて, 次元 0 の面を**頂点 (vertex)**, 次元 1 の面を**辺 (edge)** と呼ぶことがある. 次元 2 以上の任意の単

[1] 局所大域原理 (local-global principle) とは整数論などで知られる不定方程式の解の存在性に関して, 局所的な情報が大域的な情報を導くという現象の総称である.

体複体 X に対して $(X(0), X(1))$ はグラフである.

例 1. グラフ $G = (V, E)$ に対し, 空集合, V, E からなる部分集合族 $\mathbb{F} = \{\emptyset\} \cup \{\{v\} : v \in V\} \cup E$ 考えると, $(V, \mathbb{F})$ は単体複体である.

例 2. 有限集合 V に対し, $\mathbb{F} = \binom{V}{\leq k} := \{\sigma \subseteq V : |\sigma| \leq k\}$ としたとき, $(V, \mathcal{F})$ は純粋な $(k+1)$-次元の単体複体である.

例 3. 閉路を含まないグラフを**森 (forest)** といい, 連結な森を**木 (tree)** という. 連結グラフ G の部分グラフであって木であるものを**全域木 (spanning tree)** という (cf. **??**). グラフ $G = (V, E)$ に対し, 森であるような部分グラフの辺集合からなる集合族 $\mathbb{F} \subseteq 2^E$ は単体複体である. すなわち,

$$\mathbb{F} = \left\{F \subseteq E : \text{部分グラフ } (V, F) \text{ は森}\right\}$$

に対して $(E, \mathbb{F})$ は単体複体である. 簡単のため G を連結グラフであるとすると, $(E, \mathbb{F})$ の極大面は G の全域木に対応し, その次元は $n-2$ に等しい. すなわち $(E, \mathbb{F})$ は純粋な $(n-2)$-次元単体複体である.

なお, グラフ G が連結でない場合, 異なる連結成分に属する二頂点 u, v を縮約し一つの頂点として扱うことによって $(V, \mathbb{F})$ の構造を変えないまま連結成分数を減らすことができるので, 連結性を仮定しても一般性を失わない.

例 4. 実行列 $A \in \mathbb{R}^{n \times m}$ (ただし $m \geq n$) の行ベクトルを $\mathbf{a}_1, \ldots, \mathbf{a}_n$ とする. 集合 $V = \{1, \ldots, n\}$ の部分集合族であって, 線形独立な行ベクトル集合のインデックスとなるものの全体を $\mathbb{F}$ とする. すなわち

$$\mathbb{F} = \left\{I \subseteq V : (\mathbf{a}_i)_{i \in I} \text{は線形独立}\right\}$$

とすると, $(V, \mathbb{F})$ は純粋な単体複体であり, その次元は A のランク $\mathrm{rank}(A)$ に対し $\mathrm{rank}(A) - 1$ となる.

例 5. 部集合 L, R を持つ二部グラフ $G = (V, E)$ を考える. 辺部分集合 $M \subseteq E$ は, 部分グラフ (V, M) の全ての頂点の次数が高々 1 であるとき**マッチング (matching)** という. マッチング M の部分集合 $M' \subseteq M$ もまたマッチングであるため, グラフ G のマッチング全体からなる辺部分集合族 $\mathbb{F} \subseteq 2^E$ に対し, $(E, \mathbb{F})$ は単体複体である. 一般に極大マッチングのサイズは異なる場合があるのでこの単体複体は純粋ではない (図 3.1).

例 6. グラフ $G = (V, E)$ の頂点部分集合 $U \subseteq V$ は, U に属する任意の二頂点間に辺がある (すなわち $\binom{U}{2} \subseteq E$) とき, **クリーク (clique)**[2] という. 特に, 単一頂点からなる集合 $\{u\}$ や $\emptyset$ もクリークである. クリークの部分集合もまたクリークなので, グラフ G の全てのクリークからなる頂点集合族を $\mathbb{F}$ とすると, $(V, \mathbb{F})$ は単体複体である.

[2] clique とは派閥を意味する英単語である.

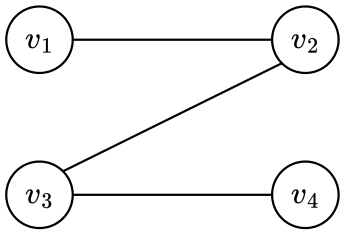

図 3.1: マッチング $M_1 = \{v_1, v_2\}, \{v_3, v_4\}\}$ と $M_2 = \{\{v_2, v_3\}\}$ はどちらも極大である.

定義 3.1.2 （リンクとスケルトン）

単体複体 $X = (V, \mathbb{F})$ を考える. 面 $\sigma \in \mathbb{F}$ の**リンク (link)** とは単体複体 $(V \setminus \sigma, \mathbb{F}_\sigma)$ であって集合族 $\mathbb{F}_\sigma$ が

$$\mathbb{F}_\sigma = \{\tau \setminus \sigma \colon \sigma \subseteq \tau \in \mathbb{F}\}$$

で与えられるものである. 次元 k 以下の面の集合

$$\mathbb{F}_k = \{\sigma \in \mathbb{F} \colon \dim \sigma \leq k\}$$

に対し $(V, \mathbb{F}_k)$ を k**-スケルトン** (k**-skelton**) という.

面 σ のリンクとは, σ をある意味で「縮約」して得られる単体複体であり, 面 σ の周りの局所的な構造を表している. 例えば連結グラフ $G = (V, E)$ の森の全体からなる単体複体 $X = (E, \mathbb{F})$ を考えよう. $\sigma \in \mathbb{F}$ を一つ固定したとき, リンク X_σ はどのような単体複体になっているだろうか? X_σ の全ての面は辺集合 F を含むので, F に含まれる全ての頂点を縮約して得られるより小さなグラフを考え, そこから F の辺を除去して得られるグラフ G' の森全体からなる単体複体とみなせる.

3.2 大域エクスパンダー性

グラフ上のランダムウォークは頂点集合上で遷移するものを考えていたが, 単体グラフ上のランダムウォークは異なる次元の面の間で遷移するものを考える. 具体的には, **??**で考えたようにある次元 i の面から次元 $i+1$ の面に遷移する上昇ウォークと逆に次元 $i+1$ の面から次元 i の面に遷移する下降ウォークである. 上昇ウォークと下降ウォークが互いに随伴の関係になるようにするため, 各 $X(i)$ 上での定常分布を定義し, $X(i)$ と $X(i+1)$ の間で詳細釣り合い条件が満たされるように定義される.

3.2.1 下降ウォークと定常分布

まず, **??**で考えた下降ウォークを単体複体に拡張し, $X(d-1)$ 上では一様分布を定常分布とすることによって帰納的に各 $X(i)$ 上での定常分布を定める.

定義 3.2.1 （下降ウォークと面上の定常分布）

純粋な d 次単体複体 $X = (V, \mathbb{F})$ を考える．各 $i = 0, \dots, d-1$ に対し確率行列 $P_i^{\downarrow} \in [0,1]^{X(i) \times X(i-1)}$ を

$$P_i^{\downarrow}(\tau, \sigma) = \begin{cases} \frac{1}{i+1} & \text{if } \sigma \subseteq \tau, \\ 0 & \text{otherwise} \end{cases}$$

とする．各 $i = 0, \dots, d-1$ に対して，$X(i)$ 上の分布 $\pi_i \in [0,1]^{X(i)}$ を

- $i = d-1$ のとき，π_{d-1} は $X(d-1)$ 上の一様分布．すなわち $\pi_{d-1}(\sigma) = \frac{1}{|X(d-1)|}$
- π_{i+1} が定義されているとき，$\pi_i = \pi_{i+1} P_{i+1}^{\downarrow}$

で定める．分布 π_i を (i 次の) 定常分布と呼ぶ．

面 $\tau \in X(i+1)$ に対して $P_{i+1}^{\downarrow}(\tau, \cdot)$ で定まる $X(i)$ 上の分布は，面 τ に含まれる頂点 u を一様ランダムに選んだときの $\sigma = \tau \setminus \{u\}$ の分布と等しい．この分布は，まず一様ランダムに $X(d)$ から面を選び，その中から一様ランダムに選ばれた $i+1$ 個の頂点からなる $X(i)$ の面のなす分布である．特に全ての $\sigma \in X(i)$ に対し $\pi_i(\sigma) = 0$ である (そうでなければ，π_d が一様分布であることに反する)．ある面 $\tau \in X(i+1)$ から分布 $P_{i+1}^{\downarrow}(\tau, \cdot)$ に従ってランダムに選ばれた面 σ に遷移する過程を**下降ウォーク (down walk)** と呼ぶ．

3.2.2 上昇ウォーク

定義 3.2.1 では次元 i の面から次元 $i-1$ に遷移する下降ウォークを与えた．同様に，次元 i の面から次元 $i+1$ の面に遷移する上昇ウォークを，$X(i)$ と $X(i+1)$ の間の詳細釣り合い条件が成り立つように定義する．

定義 3.2.2 （上昇ウォーク）

純粋な d 次単体複体 $X = (V, \mathbb{F})$ を考える．各 $i = -1, \dots, d-2$ に対し確率行列 $P_i^{\uparrow} \in [0,1]^{X(i) \times X(i+1)}$ を

$$\begin{aligned} P_i^{\uparrow}(\sigma, \tau) &= \frac{\pi_{i+1}(\tau)}{\pi_i(\sigma)} P_{i+1}^{\downarrow}(\tau, \sigma) \\ &= \begin{cases} \frac{\pi_{i+1}(\tau)}{(i+1)\pi_i(\sigma)} & \text{if } \sigma \subseteq \tau, \\ 0 & \text{otherwise} \end{cases} \end{aligned}$$

で定める．

二つの面集合 $X(i)$ と $X(i+1)$ を部集合とする二部グラフを考えればわかりやすい．それぞれの部集合には π_i, π_{i+1} が定常分布として付随しており，上昇ウォーク $P_i^{\uparrow}$ と下降ウォーク $P_{i+1}^{\downarrow}$ は詳細釣り合い条件

$$\forall \sigma \in X(i), \tau \in X(i+1),\ \pi_i(\sigma) P_i^{\uparrow}(\sigma, \tau) = \pi_{i+1}(\tau) P_{i+1}^{\downarrow}(\tau, \sigma)$$

を満たしている.

なお, 下降ウォーク $P_i^{\downarrow}$ と上昇ウォーク $P_i^{\uparrow}$ の添字 i は遷移の開始地点の面の次元としている. 各 $X(i)$ に対して空間 $\ell^2_{\pi_i}(X(i))$, すなわち, **??**と同様にして $\mathbb{R}^{X(i)}$ に内積 $\langle\cdot,\cdot\rangle_{\pi_i}$ が導入された空間を定義できる. このとき,

$$P_i^{\uparrow}\colon \ell^2_{\pi_{i+1}}(X(i+1)) \to \ell^2_{\pi_i}(X(i))$$

と

$$P_{i+1}^{\downarrow}\colon \ell^2_{\pi_i}(X(i)) \to \ell^2_{\pi_{i+1}}(X(i+1))$$

は互いに随伴の関係にある. すなわち, 任意の $f \in \ell^2_{\pi_i}(X(i)), g \in \ell^2_{\pi_{i+1}}(X(i+1))$ に対して

$$\langle P_i^{\uparrow} f, g\rangle_{\pi_{i+1}} = \langle f, P_{i+1}^{\downarrow} g\rangle_{\pi_i} \tag{3.1}$$

が成り立つ.

演習問題 2

式 (3.1) を確認せよ. すなわち, 定義 3.2.1 で定義された定常分布 π_i, π_{i+1} および任意の二つの関数 $f\colon X(i) \to \mathbb{R}, g\colon X(i+1) \to \mathbb{R}$ に対して

$$\sum_{\sigma\in X(i)} \pi_i(\sigma) f(u) = \sum_{\tau\in X(i+1)} \pi_{i+1}(\tau) g(\tau)$$

を確認せよ.

最後に, 上昇ウォークと下降ウォークを組み合わせることによって面 $X(i)$ 上の 2 種類のランダムウォークが定義できる:

定義 3.2.3 （上昇下降と下降上昇ウォーク）

定義 3.2.1, 3.2.2 と同じ設定を考える.

$$P_i^{\wedge} := P_i^{\uparrow} P_{i+1}^{\downarrow},$$
$$P_i^{\vee} := P_i^{\downarrow} P_{i-1}^{\uparrow}$$

を遷移確率行列として持つ $X(i)$ 上のランダムウォークをそれぞれ**上昇下降ウォーク (up-down walk), 下降上昇ウォーク (down-up walk)** と呼ぶ. ここで, $X(-1)$ 上での下降上昇ウォークと $X(d-1)$ 上での上昇下降ウォークは定義されない.

上昇下降ウォークはグラフ上の遅延単純ランダムウォークの自然な一般化になっている (**??**).

補題 3.2.4 （定常分布）

面 $X(i)$ 上の上昇下降ウォークと下降上昇ウォークはどちらも π_i を定常分布としてもつ.

証明. 計算によって簡単に確認できる. 実際,

$$\pi_i P_i^{\wedge} = \pi_i P_i^{\uparrow} P_{i+1}^{\downarrow} = \pi_{i+1} P_{i+1}^{\downarrow} = \pi_i,$$
$$\pi_i P_i^{\vee} = \pi_i P_i^{\downarrow} P_{i-1}^{\uparrow} = \pi_{i-1} P_{i-1}^{\uparrow} = \pi_i$$

より主張を得る. □

注釈 3.2.5 （「上昇」「下降」の名称）

非常にややこしいのだが, 上昇ウォークと下降ウォークの遷移確率行列と左右どちらから作用させるかによって「上昇」「下降」の意味合いが反転してしまう. 確率行列としての上昇ウォークは $P_i^{\uparrow} \in [0,1]^{X(i)\times X(i+1)}$ で表せる. 一般によくある左から作用させる作用素の感覚で考えると $P_i^{\uparrow} \colon \mathbb{R}^{X(i+1)} \to \mathbb{R}^{X(i)}$ であるので, 次元を一つ落とすように見えてしまうのである. 下降ウォークについても同様である. 特に上昇下降ウォーク $P_i^{\wedge} = P_i^{\uparrow} P_{i+1}^{\downarrow}$ を左から作用させると「次元を下げてから上げる」ものになるので, 下降上昇ウォークと混同しやすい.
本講義はランダムウォークを主眼におき, 右から作用させるときの $P_i^{\uparrow}, P_i^{\downarrow}$ に興味があるので, 左固有値以外の文脈ではとにかくランダムウォークの遷移確率行列といえば右から作用させるものであると考えていき, 名称についても定義 3.2.1, 3.2.2 の呼称を採用している.
そもそも, 遷移確率行列 $P(u,v)$ を「u から v に遷移する確率」として定義してしまったのが根本的な原因であり, 転置したものを改めて遷移確率行列と定義しなおせば解決できるのだが, ランダムウォークの文化ではもはや**??**の定義が完全に主流となってしまっておりそれに反するとほとんどの参考文献が読みづらくなってしまうので従った.
なお, 可逆なランダムウォーク (**??**) は内積 $\langle \cdot, \cdot \rangle_{\pi}$ の意味で左右どちらから作用させても本質的に同じであるので左右どちらから作用させるかについてこのような煩雑な話は考えなくて良い.

3.3 局所エクスパンダー性

単体複体の各リンクに対して局所的なランダムウォークを次のように定義する.

3.3.1 局所的なランダムウォーク

全てのリンクの 2-スケルトン上での局所的な重み付きランダムウォークを考える. 重み付きランダムウォークについては**??**を参照されたい.

定義 3.3.1 （局所ランダムウォーク）

純粋な d 次元単体複体 $X = (V, \mathbb{F})$ を考える. 次元 $i \leq d-2$ の面 $\sigma \in \mathbb{F}$ に対し, リンク X_σ の 2-スケルトンを $G_\sigma = (V_\sigma, E_\sigma)$ とする. このグラフの辺重み $w_\sigma \colon E_\sigma \to [0,1]$ を

$$w_\sigma(e) = \pi_i(\sigma \cup e)$$

で定め, これによって定まる V_σ 上の重み付きランダムウォークを面 σ に関する**局所ランダムウォーク (local random walk)** と呼び[a], 遷移確率行列を $P_\sigma \in [0,1]^{V_\sigma \times V_\sigma}$ と

する.

[a] このランダムウォークの概念は高次元エクスパンダーの文脈ではほぼ必ず登場するが, 特に標準的な用語が与えられてはいないので、「局所ランダムウォーク」という用語は本講義だけの局所的なものとする.

必ずしも局所ランダムウォークが既約性や非周期性を持つとは限らない (すなわち, グラフ G_σ が非連結だったり二部グラフになりうる) が, 可逆性は必ず満たすことに注意せよ.

グラフ (1 次元単体複体) だと面 $\emptyset$ に対する局所ランダムウォークのみ存在するが, これは上昇下降ウォーク (すなわち遅延単純ランダムウォーク) と同じである. 従って局所ランダムウォークの概念はより高次元の単体複体を考える際に意味を持つ.

補題 3.3.2

遷移確率行列 P_σ をもつ局所ランダムウォークの定常分布を π_σ とすると,

3.3.2 局所エクスパンダー

純粋な単体複体の局所的なエクスパンダー性を定義する. 任意の面 σ に対し G_σ 上での局所ランダムウォークの第二固有値が小さいとき, その単体複体は局所的エクスパンダー性をもつという.

定義 3.3.3 (局所エクスパンダー性)

純粋な d 次元単体複体 $X = (V, \mathbb{F})$ は, 任意の面 $\sigma \in \mathbb{F}$ に対し $\lambda_2(P) \leq \lambda$ を満たすとき, **局所 λ-エクスパンダー (local λ-expander)** であるという. より一般に, 任意の $i = -1, \ldots, d-2$ と任意の面 $\sigma \in X(i)$ に対して $\lambda_2(P_\sigma) \leq \lambda_i$ を満たすとき, 単体複体 X は局所 $(\lambda_{-1}, \ldots, \lambda_{d-2})$-エクスパンダーであるという.

エクスパンダーグラフ (定義 2.1.1) と比較すると, 片側 (λ_2) だけの上界だけでエクスパンダー性を定義している. これは, 単体複体上の上昇下降ウォークがグラフ上の遅延単純ランダムウォークに対応していることに起因する (遅延単純ランダムウォークの遷移確率行列は半正定値だから λ_2 の上界だけあれば混交性が保証される).

3.4 Oppenheim のトリクルダウン定理

3.5 高次元エクスパンダーの応用

理論計算機科学における高次元エクスパンダーの応用を簡単にまとめる. 特にマトロイドに関するものは次チャプターにて解説する. 本質的には, 局所的な情報 (局所ランダムウォークの混交性) が大域的な情報 (上昇下降ランダムウォークの混交性) を導出するという性質が極めて重要であり, これに基づいて誤り訂正符号などの構成がなされている.

Chapter 4

マトロイド

マトロイド (matroid) は「行列 (matrix) のようなもの (-oid)」という名を冠するが, 線形代数における線型独立性をグラフの全域木などに拡張した概念である.

4.1 定義

定義 4.1.1 （マトロイド）

次の性質を持つ単体複体 $(V, \mathbb{F})$ を**マトロイド (matroid)** という: 任意の $\sigma, \tau \in \mathbb{F}$ に対し, $|\sigma| < |\tau|$ ならば, ある $u \in \tau \setminus \sigma$ が存在して $\sigma \cup \{u\} \in \mathbb{F}$.

4.2 例

4.2.1 グラフ的マトロイド

4.2.2 線形マトロイド

4.3 モチベーション

4.3.1 組合せ最適化

4.3.2 組合せ論

4.4 基の数え上げ

4.5 Anari, Liu, Gharan, Vinzant の定理

Bibliography

[Fri08] J. Friedman. "A Proof of Alon's Second Eigenvalue Conjecture and Related Problems". In: **Memoirs of the American Mathematical Society** 195 (2008) (cit. on p. 14).

[HLW06] S. Hoory, N. Linial, and A. Wigderson. "Expander graphs and their applications". en. In: **Bull. Am. Math. Soc.** 43.4 (2006), pp. 439–561. DOI: 10.1090/S0273-0979-06-01126-8 (cit. on p. 17).

[LPS88] A. Lubotzky, R. Phillips, and P. Sarnak. "Ramanujan graphs". In: **Combinatorica** 8.3 (Sept. 1988), pp. 261–277. DOI: 10.1007/BF02126799 (cit. on pp. 13, 14).

[Mar88] G. A. Margulis. "Explicit group-theoretic constructions of combinatorial schemes and their applications in the construction of expanders and concentrators". In: **Problems Inform. Transmission** (1988) (cit. on p. 13).

[Mor94] M. Morgenstern. "Existence and Explicit Constructions of q + 1 Regular Ramanujan Graphs for Every Prime Power q". In: **Journal of Combinatorial Theory Series B** 62.1 (Sept. 1994), pp. 44–62. DOI: 10.1006/jctb.1994.1054 (cit. on p. 13).