

集中講義: 高次元エクスパンダーとその応用

清水 伸高 (東工大)

2024 年 5 月

Contents

1	グラフ上のランダムウォーク	1
1.1	グラフ	1
1.2	グラフ上の単純ランダムウォーク	2
1.3	単純ランダムウォークの収束性と定常分布	3
1.4	辺上のランダムウォーク	4
1.5	混交時間とエクスパンダーグラフ	5
2	高次元エクスパンダー	5
2.1	定義	5
2.2	単体複体上のランダムウォーク	5
2.3	局所スペクトルエクスパンダー	5
2.4	Oppenheim のトリクルダウン定理	5
3	マトロイド	5
3.1	定義	5
3.2	例	5
3.2.1	グラフ的マトロイド	5
3.2.2	線形マトロイド	5
3.3	モチベーション	5
3.3.1	組合せ最適化	5
3.3.2	組合せ論	5
3.4	基の数え上げ	5
4	Anari, Liu, Gharan, Vinzant の定理	5

1 グラフ上のランダムウォーク

1.1 グラフ

まず基礎的な概念の定義を与える. 本講義では有限単純無向グラフを単に**グラフ (graph)**と呼ぶ. すなわち, グラフとは有限集合 V とその二元部分集合 $E \subseteq \binom{V}{2}$ の組 $G = (V, E)$ である. V の元を**頂点 (vertex)**, E の元を**辺 (edge)**と呼ぶ. 特に混乱が生じない限りは辺 $\{u, v\}$ を省略して uv と記す. 二頂点 $u, v \in V$ が $uv \in E$ を満たすとき, u は v に隣接しているという (v もまた u に隣接している). 頂点 $u \in V$ に対し $\deg(u) = |\{v \in V : uv \in E\}|$ を**次数 (degree)**と呼ぶ. 全ての頂点の次数が d に

等しいとき, G は d -正則 (d -regular) であるという. 以下で定義される行列 $A \in \mathbb{R}^{V \times V}$ を隣接行列 (adjacency matrix) という:

$$A(u, v) = \begin{cases} 1 & \text{if } uv \in E, \\ 0 & \text{otherwise.} \end{cases}$$

考えているグラフ G が明らかな場合は次数や隣接行列などを $\deg(u)$, A などと表す. また, 考えているグラフ G が曖昧であったり特別に指定したい場合は $\deg_G(u)$, A_G などと表す.

グラフ $G = (V, E)$ の連結性を定義する. 二頂点 $a, b \in V$ に対し, ある頂点列 $v_0 = a, v_1, \dots, v_{\ell-1}, v_\ell = b$ が存在して $\{v_0, v_1\}, \dots, \{v_{\ell-1}, v_\ell\}$ が全て G の辺になっているとき, $a \sim b$ と表す. この関係は同値関係になっており, V を \sim で割った商集合 V/\sim の各同値類 $[v]$ を G の連結成分 (connected component) という. 商集合 V/\sim が単一の連結成分からなるとき, G は連結 (connected) であるという.

グラフ $G = (V, E)$ の二部性を定義する. ある頂点分割 $V = L \sqcup R$ が存在して $E \cap \binom{L}{2} = \emptyset$ かつ $E \cap \binom{R}{2} = \emptyset$ が成り立つとき, G は二部 (bipartite) であるといい, 頂点部分集合 L, R を G の部集合 (partite set) と呼ぶ. 直感的には, G が二部グラフであるというのは, ある頂点分割 $V = L \sqcup R$ に対して G の全ての辺が L と R の間を跨いでいることを意味する. なお, 部集合への分割 $V = L \sqcup R$ は必ずしも一意であるとは限らない.

1.2 グラフ上の単純ランダムウォーク

ランダムウォークは高次元エクスペンダーの定義やその解析に不可欠な概念である. そこで本節ではまず最も基本的なグラフ上の単純ランダムウォークを定義し, その基本的な性質を紹介する. 後に単純ランダムウォークを拡張した一般的なランダムウォークを導入し, その重要なクラスである可逆なランダムウォークについて説明する.

定義 1.1 (単純ランダムウォーク)

グラフ $G = (V, E)$ を考える. 頂点集合 V 上に値をとる確率変数の列 $(X_t)_{t=0,1,\dots}$ であって, 任意の $t \geq 0$, 頂点列 $(v_0, \dots, v_{t-1}) \in V^t$, および $v \in V$ に対して

$$\Pr[X_t = v \mid X_0 = v_0, \dots, X_{t-1} = v_{t-1}] = \Pr[X_t = v_t \mid X_{t-1} = v_{t-1}] = \frac{1}{\deg(v_{t-1})}$$

を満たすものを G 上の単純ランダムウォーク (simple random walk) という. さらに,

$$P(u, v) = \begin{cases} \frac{1}{\deg(u)} & \text{if } uv \in E, \\ 0 & \text{otherwise} \end{cases}$$

で定義される行列 $P \in [0, 1]^{V \times V}$ を遷移確率行列 (transition matrix) と呼ぶ.

要するに, 初期地点 X_0 を選び, 現在いる位置から一様ランダムな隣接点を選びそこに遷移するという確率的な操作を繰り返して得られる系列が単純ランダムウォークである. 初期地点の選び方は何でもよく, 例えば X_0 は決定的な頂点 $X_0 = u$ であったり一様ランダムな頂点であってもよい. 初期頂点 X_0 の分布が決まれば各時刻 t における X_t の分布は一意に定まる. 実際, $t \geq 0$ に対し $x_t \in [0, 1]^V$

を X_t の分布とする (すなわち, $x_t(u) = \Pr[X_t = u]$). 任意の $t \geq 1$ に対し

$$\begin{aligned} x_t(v) &= \Pr[X_t = v] \\ &= \sum_{u \in V} \Pr[X_t = v \text{ and } X_{t-1} = u] \\ &= \sum_{u \in V} \Pr[X_t = v \mid X_{t-1} = u] \Pr[X_{t-1} = u] \\ &= \sum_{u \in V} P(u, v) x_{t-1}(u) \end{aligned}$$

という漸化式を得る. これは $x_t = x_{t-1}P$ と表せる (ここで x_t は行ベクトルとして扱う) ので

$$x_t = x_0 P^t \quad (1)$$

を得る.

グラフ $G = (V, E)$ の各頂点を対角に並べた $V \times V$ 行列を**次数行列 (degree matrix)** という. すなわち, 次数行列 $D \in \mathbb{R}^{V \times V}$ は

$$D(u, v) = \begin{cases} \deg(u) & \text{if } u = v, \\ 0 & \text{otherwise.} \end{cases}$$

グラフ G 上の単純ランダムウォークの遷移確率行列 P は, 次数行列 D と隣接行列 A を用いて $P = D^{-1}A$ と表せる.

1.3 単純ランダムウォークの収束性と定常分布

グラフ G 上の単純ランダムウォーク $(X_t)_{t \geq 0}$ を考え, 時刻 t における分布 $x_t \in [0, 1]^V$ の $t \rightarrow \infty$ における収束性について述べる.

まず, 分布間の距離として全変動距離を定義する.

定義 1.2 (全変動距離)

有限集合 V 上の二つの分布 $\mu, \nu \in [0, 1]^V$ に対し, **全変動距離 (total variation distance)** を

$$d_{\text{TV}}(\mu, \nu) := \frac{1}{2} \sum_{u \in V} |\mu(u) - \nu(u)| = \frac{1}{2} \|\mu - \nu\|_1$$

で定める.

全変動距離は単に ℓ^1 ノルムを 2 で割った値だが, 次の性質を持つがゆえに統計学, 情報理論, 機械学習, 計算機科学を含む様々な分野で非常に重要な役割を果たしている.

命題 1.3

有限集合 V を考え, 分布 $\pi \in [0, 1]^V$ と部分集合 $U \subseteq V$ に対し $\pi(U) := \sum_{u \in U} \pi(u)$ とする. 任意の二つの分布 $\mu, \nu \in [0, 1]^V$ と任意の部分集合 $U \subseteq V$ に対して

$$|\mu(U) - \nu(U)| \leq d_{\text{TV}}(\mu, \nu).$$

すなわち、全変動距離が小さいということは任意の事象の発生確率の差が小さいことを意味する。なお、この不等式はタイトである。実際、 $U = \{u \in V : \mu(u) > \nu(u)\}$ とすれば等号が成り立つ。

グラフ G 上の単純ランダムウォークの分布を x_t とする ($t \geq 0$)。ある分布 $\pi \in [0, 1]^V$ が存在して任意の初期分布 x_0 に対し $d_{TV}(x_t, \pi) \rightarrow 0$ ($t \rightarrow \infty$) となるための条件を与える。まず、これが満たされないような G の条件を考えてみよう。

- 単純ランダムウォークは同じ連結成分内でしか遷移しない。従ってグラフ G が非連結ならば初期頂点 X_0 の場所によって x_t は異なってしまう。
- グラフ G が二部グラフであり、部集合への分割の一つを $V = L \sqcup R$ とする。仮に $X_0 \in L$ であったとすると t が偶数のとき $X_t \in L$ 、 t が奇数のとき $X_t \in R$ となるので、 x_t は収束しない。

実は、収束性を阻害するのは上の二つのケースのみであることが知られている (証明は割愛)。

定理 1.4 (単純ランダムウォークの収束性)

連結グラフ $G = (V, E)$ が二部グラフでないならば、ある分布 $\pi \in [0, 1]^V$ が存在して G 上の任意の単純ランダムウォークの分布 x_t は初期頂点の分布 x_0 に依らず $d_{TV}(x_t, \pi) \rightarrow 0$ を満たす。さらに、収束先の分布 π は次で与えられる：

$$\pi(u) = \frac{\deg(u)}{2|E|}. \quad (2)$$

定義 1.5 (定常分布)

グラフ $G = (V, E)$ に対し、式 (2) で定義される V 上の分布 $\pi \in [0, 1]^V$ を単純ランダムウォークの**定常分布 (stationary distribution)** と呼ぶ。

グラフ G が非連結や二部グラフであったとしても定常分布は定義できることに注意せよ。単純ランダムウォークでは定常分布は常に存在するが、定理 1.4 の条件を満たすときかつその時に限り定常分布への一意収束性が保証される。

1.4 辺上のランダムウォーク

単純ランダムウォークの定常分布 π を導出してみよう。

1.5 混交時間とエクспанダーグラフ

2 高次元エクспанダー

2.1 定義

2.2 単体複体上のランダムウォーク

2.3 局所スペクトルエクспанダー

2.4 Oppenheim のトリクルダウン定理

3 マトロイド

3.1 定義

3.2 例

3.2.1 グラフ的マトロイド

3.2.2 線形マトロイド

3.3 モチベーション

3.3.1 組合せ最適化

3.3.2 組合せ論

3.4 基の数え上げ

4 Anari, Liu, Gharan, Vinzant の定理

定理 4.1 (Oppenheim のトリクルダウン定理)

hoge

定理 4.1より, 以下を得る.

系 4.2

jimei na kei

系 4.2 定理 4.1