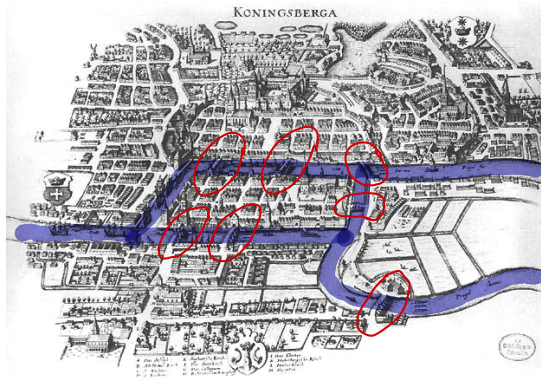


ランダムグラフの基礎および連結グラフの総数

1. 研究背景

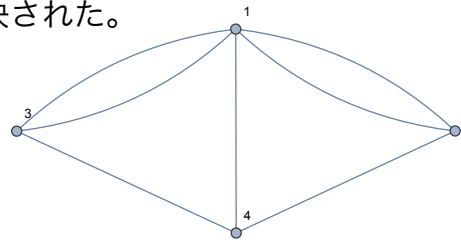
右は18世紀初頭の東プロイセン首都ケーニヒスベルクの地図である。

地図中の青線は首都を東西に両断するプレーゲル川、赤囲みはそこにかかる橋である。この7つの橋において、どの橋も2回以上通らずに、すべて1回渡って、元のところに帰還できるだろうか？
ただし、どこからでも出発してよい。



これは「**ケーニヒスベルクの橋渡し**」と呼ばれる、数学史上の有名問題である。問題提起からおおよそ30年後の1736年、**レオンハルト・オイラー**の論文「Solutio problematis ad geometriam situs pertinentis. (位置の幾何学に関する一問題の解)」によって、地図上の一筆書き問題に帰着され、「**帰還できない**」と否定的に解決された。

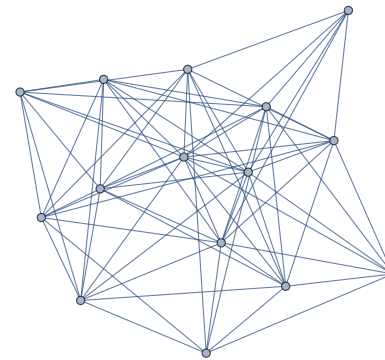
同論文で、**オイラー**は地図を右のように単純化した。右が一筆書き可能であれば、題意を満たす7つの橋の渡りかたが存在する。



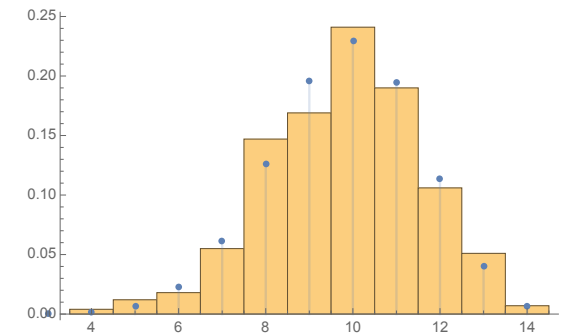
右のような図形は数学的には「**グラフ**」と呼ばれる。オイラーはこのグラフ（正確には**連結グラフ**）において、

- ・すべての頂点の**次数**（頂点から伸びる辺の本数）が偶数である。
- ・2つの頂点の**次数**が奇数で、その他のすべての頂点の**次数**は偶数であるのどちらかが成り立つときに限り一筆書きできることを示した。

異分野・異業種研究交流会2021(日本数学会、日本応用数理学会、統計関連学会連合)
九州大学大学院数理学府D3 蓮井太郎 (hasui.taro.206@s.kyushu-u.ac.jp)



エルデシュ・レンニグラフの例
($n = 15, p = 0.7$)



左のようなグラフを1000個集めた場合の頂点次数の分布。青丸は二項分布
これにより、 $n \rightarrow \infty$ で頂点次数はポアソン分布に従うことが分かる

ケーニヒスベルクの橋渡し問題のように、とある性質を満たす**グラフ**が存在するか？ という動機に対して数学的な解答を与えることを目的とする**グラフ理論**は、1736年のこのときから始まったと言える。その後、1959年、**ポール・エルデシュ**と**アルフレッド・レンニ**によって、**グラフ**にランダム性を導入した**ランダムグラフ理論**が研究され始める。

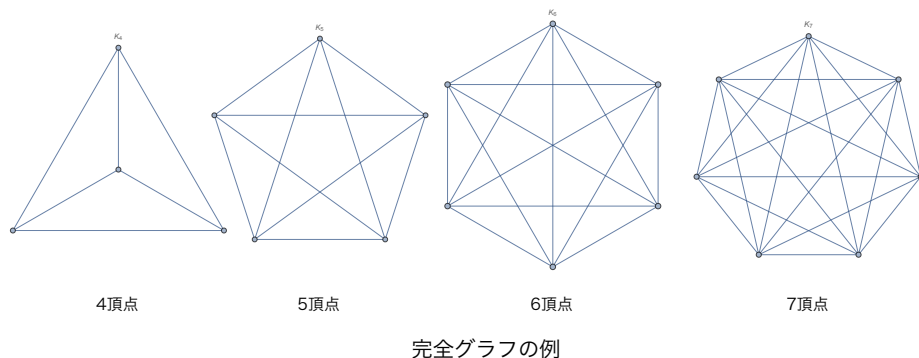
本稿では複雑な数学の知識、もしくはその応用に関するあらゆる知識を仮定せずに、**グラフ理論**の基礎を、そしてエルデシュらによる初期の**ランダムグラフ理論**の概要を述べる。なお、**ランダムグラフ**は応用上**ネットワーク理論**と関係が深い。例えば古典的な**ランダムグラフ**のひとつである「**エルデシュ・レンニグラフ**」は、そのまま現実世界での様々な**ネットワーク理論**に応用できる。中でも集中的に研究されているのは、グラフの**連結性**に関しであり、本稿の後半では辺と頂点を指定したときに連結なグラフがいくつできるか？ という**連結グラフの総数の問題**について論じる。なお本講演は**白井朋之氏** (九大IMI)と**藪奥哲史氏**(北九州高専)との共同研究による。

2. ランダムグラフの基本的な知識

まず（ランダム）グラフ理論の定義や用語を導入する。

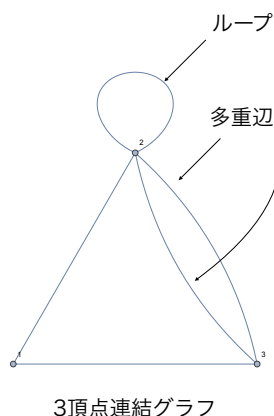
◆基本的な定義

- ・有限集合 V に対して、 $E \subset V \times V$ から決まる構造 $G = (V, E)$ を **グラフ** と呼ぶ。（ V : **頂点集合**、 E : **辺集合**）
- ・つながった辺の列を **道**、始点と終点と同じ道を **閉道**、同じ頂点对を持つ複数本の辺を **多重辺** と呼ぶ。
- ・ある辺の両端点が等しいとき、**ループ** と呼ぶ。
- ・2頂点間に道が存在することを **連結** とよび、グラフ上の任意の2 頂点が連結なグラフを **連結グラフ** と呼ぶ。
- ・任意の2頂点間に道があるグラフを **完全グラフ** と呼ぶ。



頂点集合に対し、確率で辺を入れて完成するグラフを **ランダムグラフ** という。辺や確率の導入方法によって様々なランダムグラフが存在する。なお、ランダムグラフ理論においては、（分野にもよるが）基本的に右のような **ループ** や **多重辺** を持つグラフは考慮しないことが多い。このようなループや多重辺を持たないグラフを **単純グラフ** と定義する。

たとえば、もっとも古典的で有名なランダムグラフが次の **エルデシュ・レンニグラフ** である。

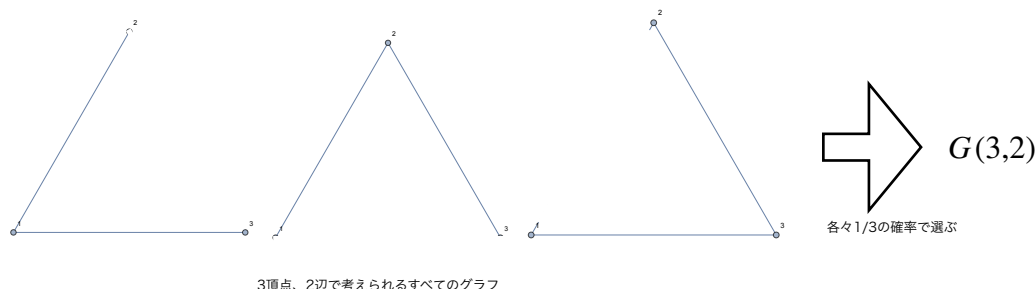


まず **エルデシュ・レンニグラフ** には、以下の2パターンが存在する。論分や書籍ではこれらは区別されていないこともあるので、文脈からどちらか判断しなければならない。なお、歴史的には $G(n, p)$ 型で研究され始めているが、近年この **エルデシュ・レンニモデル** を一般化した **ハイパーグラフ理論** においては、 $G(n, M)$ 型も盛んに研究されている。

◆ $G(n, M)$ 型

n 頂点、 M 辺のグラフから **一様ランダムにひとつのグラフを選択するタイプ**

例. $n = 3, M = 2$ の場合、



◆ $G(n, p)$ 型

辺をひとつも持たない n 点から始め、その **孤立点** に対し、**辺を独立に確率 p で加える** タイプ。加えられる辺の数は $\binom{n}{2}$ 本

時間発展させる場合、 n 頂点 **完全グラフ** K_n ができたら終了。頂点 v の

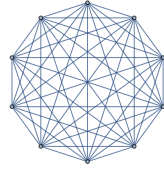
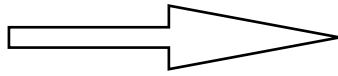
次数 を $\deg(v)$ とすると、 $P[\deg(v) = k] = \binom{n-1}{k} p^k (1-p)^{n-1-k}$

また、簡単な計算で $E[\text{次数}] = (n-1)p$ 、 $E[\text{辺の本数}] = \binom{n}{2} p$ も分かる。

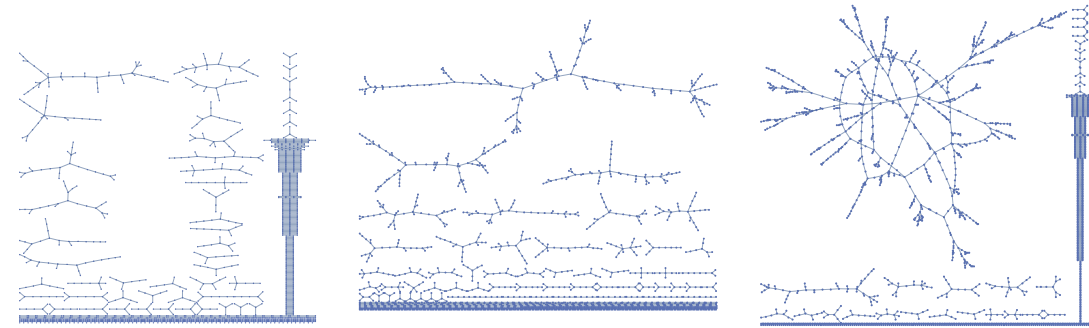
例. $G(10, p)$ すなわち $n = 10$ の場合



$p = 0$ (孤立点)



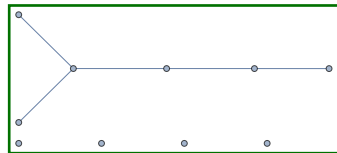
$p = 1$ (完全グラフ)



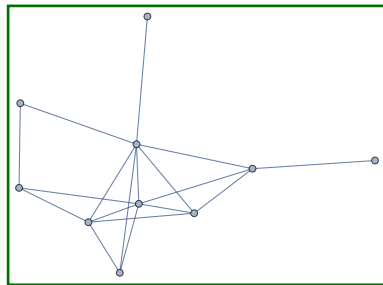
$n = 2000, c = 0.8$ すなわち $p = \frac{0.8}{2000}$

$n = 2000$ すなわち $p = \frac{1}{2000}$

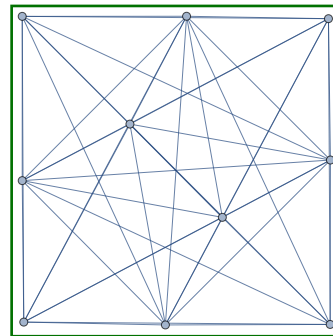
$n = 2000, c = 1.2$ すなわち $p = \frac{1.2}{2000}$



$p = 0.1$



$p = 0.5$



$p = 0.9$

さて、そもそも連結グラフはいくつあるのだろうか。より正確には、辺と頂点の数を指定した場合、連結グラフは全部で何個できるのか。連結性を目を向ければ当然こういう疑問が出てくる。

グラフの種類によっては、連結グラフの総数は知られている。以下、 $G(V, E)$ を単純グラフする。このとき $|E| - |V|$ の値によってグラフを分類する。

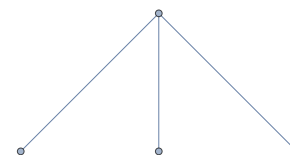
◆グラフの分類

- $|E| - |V| = -1$ となるグラフをツリーと呼ぶ。
- グラフ (V, E) の任意の辺集合 $T \subset E$ においてその部分グラフ (V, T) が木であるなら、この (V, T) を (V, E) のスパニングツリーと呼ぶ。
- $|E| - |V| = 0$ となるグラフをサイクルグラフと呼ぶ。
- $|E| - |V| = k \geq 1$ となるグラフを k -コンプレックスグラフと呼ぶ。

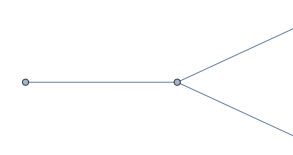
3. グラフの連結性および先行研究

ランダムグラフの研究で重要な概念はグラフの連結性である。連結性は読んで字の如くグラフがどの程度結びついているかを表している。

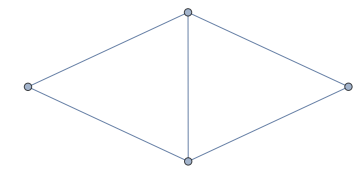
例えばエルデシュ・レンニグラフでは、 $p(n) = c/n$ のように辺の導入確率 p を頂点数 n に依存させて変化させると、グラフの各成分の連結性が激変する。このような変化を相転移と呼ぶ。なお、エルデシュ・レンニグラフは漸近的な閾値が $\log n/n$ になることも知られている。



頂点数4のツリー



頂点数4のサイクルグラフ



頂点数4の1-コンプレックスグラフ

ここで、頂点と辺の数を強調するため、 n 個の頂点と q 本の辺を持つグラフを (n, q) -グラフと書き、連結な (n, q) -グラフの数を $f(n, q)$ とする。ただし、 $k < -1$ のとき $f(n, n+k) = 0$ である。

さてこのとき次が成り立つことが知られている。

◆Cayleyの定理 (Cayley, A. 1889)

$$f(n, n-1) = n^{n-2}$$

つまり単純グラフにおいて、その全域木の数はずべて計算可能である。またサイクルグラフの総数に関しては以下が知られている。

◆サイクルグラフの総数 (E. M. Wright. 1977)

$$f(n, n) = \frac{1}{2} \left(\frac{h(n)}{n} - n^{n-2}(n-1) \right)$$

ただし、 $h(n) = \sum_{s=1}^{n-1} \binom{n}{s} s^s (n-s)^{n-s}$ とする。

このサイクルグラフの総数の漸近挙動に関しては以下が知られる。ちなみにコンプレックスグラフの総数は具体的には知られていない。

◆サイクルグラフの総数の漸近挙動 (E. M. Wright. 1977)

$$f(n, n) \sim \sqrt{\frac{\pi}{8}} n^{n-1/2} \quad (n \rightarrow \infty)$$

次に2部グラフで同じ議論を行う。まずグラフ $G(V, E)$ の頂点集合 V の部分集合 V_1, V_2 について、 $V = V_1 \cup V_2$ とし $V_1 \cap V_2 = \emptyset$ とする。このようなグラフを2部グラフと呼び、 $BG(V_1, V_2, E)$ と書く。

また $|V_1| = r, |V_2| = s, |E| = q$ となる2部グラフ $BG(V_1, V_2, E)$ を (r, s, q) -グラフと書く。さらに $f(r, s, q)$ を連結な (r, s, q) -グラフの総数とする。ただし $f(1, 0, 0) = f(0, 1, 0) = 1$ $f(0, 0, -1) = 0$ である。このとき2部グラフのツリーの総数に関して以下が知られる。

◆2部グラフの木の総数 (H. I. Scoins. 1962)

$$f(r, s, r+s-1) = r^{s-1} s^{r-1}$$

4.主結果

以上先行研究を述べた。ここでは本講演の蓮井-白井-藪奥による主結果である連結な2部サイクルグラフの総数とその漸近挙動について述べる。まず以下のような q に関する漸化式がある。

◆ (r, s) -グラフ上の q -コンプレックスグラフの総数の間の漸化式

$$(q+1)f(r, s, q+1) = (rs-q)f(r, s, q) + Q(r, s, q)$$

ただし $Q(r, s, q)$ は以下、

$$Q(r, s, q) = \frac{1}{2} \sum_{r_1=0}^r \sum_{s_1=0}^s \sum_{t=0}^q \binom{r}{r_1} \binom{s}{s_1} \left\{ (r-r_1)s_1 + r_1(s-s_1) \right\} f(r_1, s_1, t) f(r-r_1, s-s_1, q-t)$$

$Q(r, s, q)$ は2つの連結成分を $q+1$ 本目の辺で結んで完成する $(r, s, q+1)$ -グラフの総数である。さて、 $q = r+s-1+k$ の場合の連結グラフの総数 $f(r, s, r+s-1+k)$ において、以下のように指数型母関数を定義する。

$$F_k(x, y) := \sum_{r, s=0}^{\infty} \frac{f(r, s, r+s-1+k)}{r!s!} x^r y^s$$

上記の漸化式の指数型母関数表示は次のようになる。

◆漸化式の母関数表示

$D_x = x\partial_x$ とし $D_y = y\partial_y$ とする。 $k = -1, 0, 1, 2, \dots$ に対し、

$$\begin{aligned} & (D_x + D_y + k) F_{k+1} \\ &= (D_x D_y - D_x - D_y + 1 - k) F_k + \sum_{l=0}^{k+1} D_x F_l \cdot D_y F_{k+1-l} \end{aligned}$$

今、 $T(x, y) := F_0(x, y) = \sum_{r,s=0}^{\infty} \frac{r^{s-1} s^{r-1}}{r! s!} x^r y^s$ とする。上記母関数表示において

$k = -1$ とすると

$$(D_x + D_y - 1) F_0 = D_x F_0 \cdot D_y F_0$$

と書ける。 $T_x := D_x F_0$ $T_y := D_y F_0$ とすると、さらに $T_x + T_y - T = T_x T_y$ と書きかせる。

このとき母関数表示において $k = 1$ すなわち2部サイクルグラフの総数 $f(r, s, r+s)$ に関する母関数 $U := F_1$ について、以下が成り立つ。

◆ $f(r, s, r+s)$ の母関数

$$U(x, y) = -\frac{1}{2} \left(\log(1 - T_x T_y) + T_x T_y \right)$$

なおこの母関数の漸近挙動も判明している。以下、指数型母関数

$A(x) = \sum_{n=0}^{\infty} a_n \frac{x^n}{n!}$ について、記号 $\langle x^n \rangle A(x)$ で a_n を表す。このとき

$u_n := \langle x^n \rangle U(x, x)$ とおくと、

$$u_n = \sum_{r+s=n} \binom{n}{r} f(r, s, r+s)$$

となる。これは頂点数 n の完全2部グラフ上のサイクルグラフの個数をあらわし、その漸近挙動は次のようになる。

◆2部サイクルグラフの総数の漸近挙動

$n = 4, 5, \dots$ に対し、

$$u_n = n^{n-1} \sum_{2 \leq k \leq n/2} \frac{n!}{(n-2k)! n^{2k}}$$

$n \rightarrow \infty$ のとき、

$$u_n \sim \sqrt{\frac{\pi}{8}} n^{n-1/2}$$

ここで、サイクルグラフの総数の漸近挙動は

$$f(n, n) \sim \sqrt{\frac{\pi}{8}} n^{n-1/2} \quad (n \rightarrow \infty)$$

であった。つまりサイクルグラフの総数と2部サイクルグラフの総数の漸近挙動の主要項は一致することがわかる。

以下は $f(n, n)$ と u_n の具体的な値の比較である。この表を見るに有限の n では両者異なる。

n	3	4	5	6	7	8	9	10	11
$f(n, n)$	1	15	222	3660	68295	1436568	33779340	880107840	25201854045
u_n	0	6	120	2280	46200	1026840	25102224	673706880	19745850960

一方 u_n と $f(n, n)$ の比率のグラフは以下。1に収束する、すなわち確かに漸近的に等しい。

