Zufällige Graphen

Christoph Schumacher

25. Juni 2020

RGCN I, Kap. 7.1, 7.2, 7.3:
Das Konfigurationsmodell

Zusammenfassung

Das Konfigurationsmodell (CM):
- Einführung,
- das Erased CM.

Inhaltsverzeichnis

Fragen

1. —

1 Motivation/Ziel/Fragestellung

Wunschziel: Ein zufälliger Graph, der gleichverteilt aus der Menge aller Graphen mit einer vorgegebenen Gradfolge gezogen wird (mikrokanonisches Ensemble)

Problem: Gibt es überhaupt einen Graphen zu der gegebenen Gradfolge? Insbesondere, wenn man die Gradfolge zufällig generiert, kann das schiefgehen.

Ausweg: Generiere zunächst einen Multigraph (mit mehrfachen Kanten und Schleifen) und „repariere" das Ergebnis! Zwei Möglichkeiten:

I. Mehrfachkanten und Schleifen löschen, oder
II. so lange versuchen, bis der Multigraph ein einfacher Graph ist (das ergibt dasselbe wie auf Einfachheit bedingen).

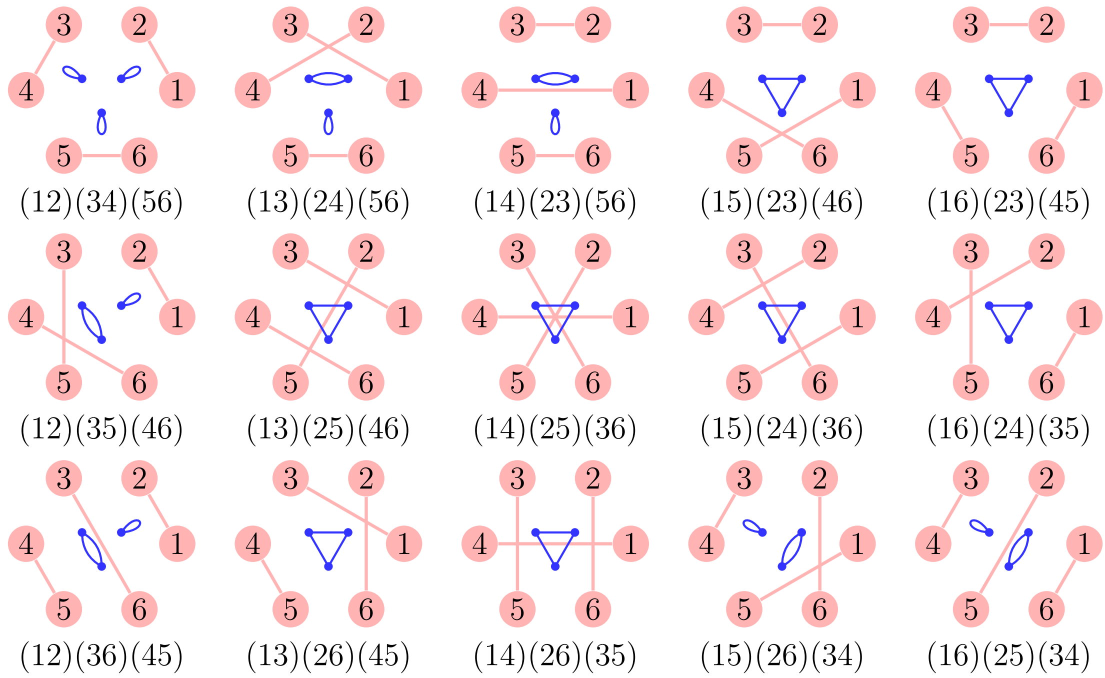
Fragen:

A. Nach dem Löschen hat man ja nicht mehr die richtige Gradfolge. Wie geht man damit um?

B. Hat man nach dem Löschen noch eine Gleichverteilung?
C. Kann man immer bedingen? Es kann doch immer noch passieren, dass keine Graphen mit der geforderten Gradfolge existieren.
D. Was kann man mit dem Modell (den Modellen?) anstellen? (Graphen zählen?)

2 Definition des CM als Multigraph

Hier sind die $5!! = 15$ möglichen Konfigurationen für $\mathrm{Conf}_3(\mathbf{d})$ mit $\mathbf{d} := (2, 2, 2)$. In der Mitte sind in blau die zugehörigen Multigraphen, also die Werte von $\mathrm{CM}_3(\mathbf{d})$. Unter den Konfigurationen steht die jeweilige Abbildung σ geschrieben als Permutation in Zykelschreibweise.



Der einfache Graph K_3 kommt $8 = (2!)^2$ Mal vor. Die drei Graphen mit einer Schleife tauchen je $2! = 2$ Mal auf.

—

Zu Definition 7.6: Ein *pairing scheme* besteht aus ZVn x_i, y_i, $i \in [\ell_n/2]$, mit Werten in $[\ell_n]$, deren gemeinsame Verteilung durch folgenden Algorithmus bestimmt ist:

Initialisierung: Die Menge der zu paarenden halben Kanten ist $L_1 := [\ell_n]$, und wir beginnen mit $j = 1$.

Iterationsschritt:

- Wähle eine halbe Kante $x_j \in L_j$ aus, beispielsweise $x_j := \min L_j$, oder man wählt uniform zufällig. Wenn diese Auswahl nur von $[\ell_n] \setminus L_j$ (und dem Zufall) abhängt, heißt das pairing scheme *adapted*.
- Wähle eine halbe Kante $y_j \in L_j \setminus \{x_j\}$ aus. Wenn diese Auswahl zufällig und gleichverteilt stattfindet, heißt das pairing scheme *uniform*.
- Definiere $L_{j+1} := L_j \setminus \{x_j, y_j\}$ und erhöhe j um 1.

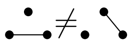
Ende: Wiederhole den Iterationsschritt bis $j > \ell_n/2$ erfüllt ist.

Die zufällige Konfiguration, die man damit generiert hat, ist $\{(x_j, y_j) : j \in [\ell_n/2]\}$ bzw. $\sigma(x_j) := y_j$, $\sigma(y_j) := x_j$, $j \in [\ell_n/2]$.

Wie könnte ein nicht adaptiertes pairing scheme aussehen? — Man darf zum Beispiel nicht den Wert von y_i benutzen, um x_i auszuwählen. Damit könnte man nämlich folgendermaßen Schindluder treiben: Wenn $y_i = 1$ vorliegt, dann wählen wir $x_i := 2$, und bei $y_i = 2$ bestimmen wir $x_i := 1$. In allen anderen Fällen wählen wir x_i unter den verbleibenden Möglichkeiten, aber nicht in $\{1, 2\}$. Auf diese Weise kann man erzwingen, dass die Konfiguration die Kante 12 enthält, und das würde die gewünschte Gleichverteilung auf der Menge der Konfigurationen zerstören.

Man könnte auch auf die Idee kommen, nur Paare zu bilden, die nicht zu multiplen Kanten und nicht zu Schleifen führen. Das wäre dann ein nicht uniformes pairing scheme, bzw. eigentlich wäre es das nicht, denn wenn die Gradfolge nicht graphisch ist, bleiben wir mittendrin stecken.

—

Zu Proposition 7.7: Jetzt sieht man, wozu die Konfigurationen gut sind: Sie bilden den zugrundeliegenden Laplace-Raum. Man erkennt an Formel (7.2.6), dass alle einfachen Graphen dieselbe Wahrscheinlichkeit bekommen, nämlich $\frac{\prod_i d_i!}{(\ell_n-1)!!}$, ganz wie gewünscht. (Dabei können aber isomorphe Graphen als ungleich gelten: $\mathbf{d} = (1,1,0) \neq (0,1,1)$: )

Beispiel.

- $\mathbf{d} = (2,2,2)$, $x_{11} = x_{22} = x_{33} = 1$, $x_{12} = x_{13} = x_{23} = 0$, $\ell_3 = 6$:
  $\mathbb{P}(\mathrm{CM}_3(\mathbf{d}) = G) = \frac{1}{5!!}\frac{2!\,2!\,2!}{2^1 2^1 2^1\; 1!\,1!\,1!\,0!\,0!\,0!} = \frac{1}{15}\checkmark$
- $\mathbf{d} = (2,2,2)$, $x_{11} = x_{22} = 0$, $x_{33} = 1$, $x_{12} = 1$, $x_{13} = x_{23} = 0$, $\ell_3 = 6$:
  $\mathbb{P}(\mathrm{CM}_3(\mathbf{d}) = G) = \frac{1}{5!!}\frac{2!\,2!\,2!}{2^0 2^0 2^1\; 0!\,0!\,1!\,2!\,0!\,0!} = \frac{2}{15}\checkmark$
- $\mathbf{d} = (2,2,2)$, $x_{11} = x_{22} = x_{33} = 0$, $x_{12} = x_{13} = x_{23} = 1$, $\ell_3 = 6$:
  $\mathbb{P}(\mathrm{CM}_3(\mathbf{d}) = G) = \frac{1}{5!!}\frac{2!\,2!\,2!}{2^0 2^0 2^0\; 0!\,0!\,0!\,1!\,1!\,1!} = \frac{8}{15}\checkmark$
- $\mathbf{d} = (2,4)$, $x_{11} = 1$, $x_{22} = 2$, $x_{12} = 0$, $\ell_2 = 6$:
  $\mathbb{P}(\mathrm{CM}_3(\mathbf{d}) = G) = \frac{1}{5!!}\frac{2!\,4!}{2^1 2^2\; 1!\,2!} = \frac{1}{5} = \mathbb{P}(\sigma(1) = 2)\checkmark$
- $\mathbf{d} = (8)$, $x_{11} = 4$, $\ell_1 = 8$:
  $\mathbb{P}(\mathrm{CM}_3(\mathbf{d}) = G) = \frac{1}{7!!}\frac{8!}{2^4 4!} = 1\checkmark$

Die Doppelfakultät der geraden Zahl ℓ_n kann man nämlich auch so ausdrücken:

$$(\ell_n - 1)!! = (\ell_n - 1)(\ell_n - 3)(\ell_n - 5)\cdots 3 \cdot 1 \tag{1}$$

$$= \frac{\ell_n!}{\ell_n(\ell_n - 2)(\ell_n - 4)\cdots 4 \cdot 2} = \frac{\ell_n!}{2^{\ell_n/2}(\ell_n/2)!} = \frac{\ell_n!}{2^{\frac{\ell_n}{2}}(\frac{\ell_n}{2})!} \tag{2}$$

3 Regularitätsbedingungen an die Folge der Gradfolgen

Die Konstruktion (7.2.18):

$$n_k := \lceil nF(k)\rceil - \lceil nF(k-1)\rceil \tag{3}$$

ist neu. Wir prüfen besser, ob die Bedingungen 7.8(a)–(b) erfüllt sind.

$$F_n(x) = \sum_{k=0}^{\lfloor x\rfloor} \frac{n_k}{n} \tag{4}$$

$$= \frac{1}{n}\sum_{k=0}^{\lfloor x\rfloor}\big(\lceil nF(k)\rceil - \lceil nF(k-1)\rceil\big) \tag{5}$$

$$= \frac{\lceil nF(\lfloor x\rfloor)\rceil}{n} \tag{6}$$

$$= \frac{\lceil nF(x) \rceil}{n} \tag{7}$$

$$\begin{cases} \geq F(x), \\ \leq F(x) + \frac{1}{n} \xrightarrow{n\to\infty} F(x). \end{cases} \tag{8}$$

Aus der Rechnung folgt auch

$$F_n(x) = \frac{\lceil nF(x) \rceil}{n} \geq F(x), \tag{9}$$

also $D_n \preceq D$, so dass nach Theorem 2.15 $\mathbb{E}[D_n] \leq \mathbb{E}[D] < \infty$. Bedingung 7.8(b) ist etwas aufwändiger. Zunächst einmal ist bekannt, dass

$$\infty > \mathbb{E}[D] = \sum_{j=1}^{\infty} j\big(F(j) - F(j-1)\big) \tag{10}$$

$$= \sum_{j=1}^{\infty} \sum_{k=0}^{j-1} \underbrace{\big(F(j) - F(j-1)\big)}_{\geq 0} \tag{11}$$

$$= \sum_{k=0}^{\infty} \sum_{j=k+1}^{\infty} \big(F(j) - F(j-1)\big) \tag{12}$$

$$= \sum_{k=0}^{\infty} \lim_{N\to\infty} \sum_{j=k+1}^{N} \big(F(j) - F(j-1)\big) \tag{13}$$

$$= \sum_{k=0}^{\infty} \lim_{N\to\infty} \big(F(N) - F(k)\big) \tag{14}$$

$$= \sum_{k=0}^{\infty} \big(1 - F(k)\big). \tag{15}$$

Analoges gilt für F_n. Das heißt:

$$0 \overset{(9)}{\leq} \mathbb{E}[D] - \mathbb{E}[D_n] = \sum_{k=0}^{\infty} \big(1 - F(k) - (1 - F_n(k))\big) \tag{16}$$

$$\overset{F_n \preceq F}{\leq} \sum_{k=0}^{N} \big(F_n(k) - F(k)\big) + 2 \sum_{k=N+1}^{\infty} \big(1 - F(k)\big) \tag{17}$$

$$\xrightarrow[n\to\infty]{7.8(a)} 2 \sum_{k=N+1}^{\infty} \big(1 - F(k)\big) \xrightarrow[N\to\infty]{\mathbb{E}[D]<\infty} 0. \tag{18}$$

4 Das Erased CM

Beim Erased CM löscht man alle Schleifen und fasst alle multiplen Kanten zu jeweils einer Kante zusammen. Dadurch verändert man natürlich die Grade. Theorem 7.10 besagt, dass die Änderung der Grade asymptotisch nicht ins Gewicht fällt.

Beweisstruktur von Theorem 7.10:

- Verwende $\varepsilon/2$ um p_k gegen $p_k^{(n)}$ auszutauschen.
- Führe die Differenz von $P_k^{(\mathrm{er})}$ und $p_k^{(n)}$ auf die Anzahl aller Schleifen S_n und die Anzahl aller Mehrfachkanten M_n zurück.
- Verwende die first moment method, zeige also $\mathbb{E}[S_n], \mathbb{E}[M_n] \to 0$.

Das Schöne daran ist, dass man beim Erwartungswert die Korrelationen zwischen den verschiedenen Kanten nicht berücksichtigen muss.