Zufällige Graphen

Christoph Schumacher

18. Juni 2020

RGCN I, Kap. 6.6:
Die Verteilung des Generalized Random Graph

Zusammenfassung

Die Verteilung des GRG wird auf eine alternative Weise charakterisiert. Dabei wird die Bedeutung des GRG klarer.

Inhaltsverzeichnis

Fragen

1. In diesem Kapitel werden die Gewichte $(w_i)_{i\in[n]}$ ersetzt durch Parameter $(u_i)_{i\in[n]}$, und zwar mittels (6.6.7): $u_i := w_i/\sqrt{\ell_n}$. Was hat es damit auf sich? — Die Gewichte $(w_i)_i$ und die zufehörigen Parameter $(u_i)_i$ führen zum selben Wahrscheinlichkeitsmaß auf der Menge aller Graphen mit Knotenmenge $[n]$. Die Umparametrisierung ist aus theoretischer Sicht praktisch, um den Zusammenhang zu anderen Disziplinen zu erkennen, ändert aber nichts am Modell. Die Gewichte w_i sind wegen ihrer Interpretation als asymptotischer Knotengrad parktisch.

2. Zusammenhang mit (6.9.1)? — Die Kantenwahrscheinlichkeiten

$$p_{ij} = \frac{\widetilde{w}_i\widetilde{w}_j}{\mathbb{E}[W]n + \widetilde{w}_i\widetilde{w}_j} \tag{1}$$

führen ebenfalls auf einen GRG, und zwar den mit Parametern $u_i := \widetilde{w}_i/\sqrt{\mathbb{E}[W]n}$. Die Parameter w_i und $\widetilde{w}_i$ kann man ineinander umrechnen:

$$\widetilde{w}_i = u_i\sqrt{\mathbb{E}[W]n} = w_i\sqrt{\frac{\mathbb{E}[W]n}{\ell_n}} = w_i\sqrt{\frac{\mathbb{E}[W]}{\mathbb{E}[W_n]}}. \tag{2}$$

Unter der Bedingung 6.4(b) stimmen w_i und $\widetilde{w}_i$ asymptotisch überein.

3. Was ist so spannend an Theorem 6.15? — Theorem 6.15 stellt eine Brücke zum Konfigurationsmodell her. Dort kann man nämlich auch eine Gleichverteilung auf einer Menge von Graphen erzeugen, und über diese Verbindung können Ergebnisse vom Konfigurationsmodell auf den GRG übertragen werden.

1 Die Verteilung des GRG

Aufgabe 6.23:

$$\begin{aligned}
\prod_{1\le i<j\le n}(u_iu_j)^{x_{ij}} &= \sqrt{\prod_{1\le i<j\le n}(u_iu_j)^{x_{ij}}\prod_{1\le j<i\le n}(u_iu_j)^{x_{ij}}} && (3)\\
&= \sqrt{\prod_{1\le i,j\le n}\left(u_i^{x_{ij}}u_j^{x_{ij}}\right)} && (4)\\
&= \sqrt{\prod_{1\le i,j\le n}u_i^{x_{ij}}\prod_{1\le i,j\le n}u_j^{x_{ij}}} && (5)\\
&= \sqrt{\prod_{i\in[n]}u_i^{\sum_{j\in[n]}x_{ij}}\prod_{j\in[n]}u_j^{\sum_{i\in[n]}x_{ij}}} && (6)\\
&= \prod_{i\in[n]}u_i^{d_i(x)} && (7)
\end{aligned}$$

—

Zur Bedeutung von Formel (6.6.10):

- Bedeutung in der **Statistik**: Man kann den GRG auf n Knoten als parametrisiertes statistisches Modell $(\mathcal{X}_n,\mathcal{A}_n,\mathcal{P}_n)$ auffassen, wobei $\mathcal{X}_n$ die Menge aller Graphen mit n Knoten ist, $\mathcal{A}_n$ die Potenzmenge davon, und

$$\mathcal{P}_n := \{\mathbb{P}_u \colon u=(u_i)_{i\in[n]}\in(0,\infty)^n\} \tag{8}$$

  die Menge aller Wahrscheinlichkeitsmaße der Form (6.6.10):

$$\mathbb{P}_u(X=x)=\frac{1}{G(u)}\prod\nolimits_{i\in[n]}u_i^{d_i(x)}=\exp\Big(\sum\nolimits_{i\in[n]}d_i(x)\log u_i-\log G(u)\Big). \tag{9}$$

  So sieht man: $\mathcal{P}_n$ ist eine exponentielle Familie. Theorem 6.15 hat eine Verallgemeinerung in der Statistik.

- In der statistischen Mechanik schreibt man Formel (6.6.10) mit Hilfe von

$$\theta_i := -\log u_i\in\mathbb{R},\quad Z(\theta):=G((\mathrm{e}^{-\theta_i})_{i\in[n]}) \tag{10}$$

  und $d(x):=(d_i(x))_{i\in[n]}$ als

$$\mathbb{P}(X=x)=\frac{1}{Z(\theta)}\exp\Big(-\sum\nolimits_{i\in[n]}\theta_id_i(x)\Big)=\frac{1}{Z(\theta)}\exp\big(-\langle\theta,d(x)\rangle\big). \tag{11}$$

  Jetzt erkennt man, dass die Verteilung des GRG einer Boltzmann-Statistik folgt und kanonisches Ensemble im Sinne der statistischen Mechanik ist. Die Graphen sind Mikrozustände, Gradfolgen sind Makrozustände, und im thermodynamischen Limes $n\to\infty$ überlebt „nur“ die empirische Verteilung der Grade. Theorem 6.15 beschreibt in der statistischen Physik das mikrokanonische Ensemble. Die Zustandssumme Z (englisch partition function) übernimmt in diesem Kontext die Aufgabe der momentenerzeugenden Funktion aus Proposition 6.16.

Die Herleitung des kanonischen Ensembles in der statistischen Mechanik ermöglicht ein tieferes Verständnis des GRG: Die Verteilungen in $\mathcal{P}_n$ maximieren die Entropie unter der Nebenbedingung, dass die Erwartungwerte der Knotengrade vorgegeben wurden. Das wird so interpretiert, dass die Verteilung zu einem gegebenen Parameter durch die Information über den erwarteten Knotengrad vollständig beschrieben ist, es steckt keine weitere Information in dem Maß. In der Statistik formuliert man den Sachverhalt so: Die Grade $(d_i(x))_{i\in[n]}$ sind eine suffiziente Statistik für den GRG, das heißt, die Gradfolge enthält die ganze Information der Stichprobe über den Parameter.

Für die Maximierung der Entropie kürzen wir $p_x := \mathbb{P}(X = x)$ ab. Die Entropie einer Verteilung $p = (p_x)_{x\in\mathcal{X}_n}$ auf $\mathcal{X}_n$ ist

$$S(p) := -\sum_{x\in\mathcal{X}_n} p_x \log p_x. \tag{12}$$

Die Nebenbedingungen sind

$$\sum_{x\in\mathcal{X}_n} p_x = 1 \quad \text{und} \quad \sum_{x\in\mathcal{X}_n} d_i(x)p_x = \delta_i \tag{13}$$

für $i \in [n]$, denn dem Knotengrad $d_i\colon \mathcal{X}_n \to \mathbb{N}_0$ des Knoten $i \in [n]$ wird der Erwartungswert $\delta_i > 0$ vorgeschrieben. Wir verwenden die Lagrange-Multiplikatoren $\mu, \theta_i \in \mathbb{R}$:

$$\frac{\partial}{\partial p_x}\Bigg(S(p) - \mu\Big(\sum\nolimits_{y\in\mathcal{X}_n} p_y - 1\Big) - \sum_{i\in[n]} \theta_i\Big(\sum\nolimits_{y\in\mathcal{X}_n} d_i(y)p_y - \delta_i\Big)\Bigg) \tag{14}$$

$$= -\log p_x - 1 - \mu - \sum_{i\in[n]} \theta_i d_i(x) \overset{!}{=} 0 \tag{15}$$

$$\iff p_x = \exp\big(-\langle\theta, d(x)\rangle - (\mu + 1)\big). \tag{16}$$

Mit der ersten Nebenbedingung können wir μ bestimmen:

$$1 = \sum_{x\in\mathcal{X}_n} p_x = \mathrm{e}^{-(\mu+1)} \sum_{x\in\mathcal{X}_n} \exp\big(-\langle\theta, d(x)\rangle\big) = \mathrm{e}^{-(\mu+1)} Z(\theta), \tag{17}$$

woraus die Form (11) folgt.

Um die Äquivalenz der Beschreibung mit Hilfe von (6.6.10) und der bisherigen Beschreibung zu sehen, leiten wir aus (6.6.10) die Beziehung (6.6.8) her:

$$p_{ij} = \mathbb{P}(X_{ij} = 1) \tag{18}$$

$$= \sum_{x\in\mathcal{X}_n\colon x_{ij}=1} \frac{1}{G(u)} \prod_{k=1}^{n} u_k^{d_k(x)} \tag{19}$$

$$= \sum_{x\in\mathcal{X}_n\colon x_{ij}=0} \frac{u_i u_j}{G(u)} \prod_{k=1}^{n} u_k^{d_k(x)} \tag{20}$$

$$= u_i u_j\, \mathbb{P}(X_{ij} = 0) \tag{21}$$

$$= u_i u_j (1 - p_{ij}). \tag{22}$$

Das können wir nach p_{ij} auflösen und erhalten (6.6.8).

Die Funktion

$$\hat{d}_r\colon \mathcal{X}^m \to \mathbb{R}, \quad x \mapsto \frac{1}{m}\sum_{k=1}^{m} d_r(x) \tag{23}$$

ist ein erwartungstreuer Schätzer für den Knotengrad des Knoten r:

$$\mathbb{E}[\hat{d}_r(X^{(1)}, \ldots, X^{(m)})] = \frac{1}{m}\sum_{k=1}^{m} \mathbb{E}[d_r(X^{(k)})] \tag{24}$$

$$
\begin{aligned}
&= \sum_{x\in\mathcal{X}} d_r(x) \frac{1}{G(u)} \prod_{i\in[n]} u_i^{d_i(x)} && (25)\\
&= \sum_{x\in\mathcal{X}} \frac{u_r}{G(u)} \frac{\partial}{\partial u_r} \prod_{i\in[n]} u_i^{d_i(x)} && (26)\\
&\overset{(6.6.12)}{=} \frac{u_r}{G(u)} \frac{\partial}{\partial u_r} G(u) && (27)\\
&= u_r \frac{\partial}{\partial u_r} \log G(u) && (28)\\
&= u_r \frac{\partial}{\partial u_r} \sum_{1\le i<j\le n} \log(1+u_i u_j) && (29)\\
&= \sum_{i=1}^{r-1} \frac{u_i u_r}{1+u_i u_r} + \sum_{j=r+1}^{n} \frac{u_r u_j}{1+u_r u_j} && (30)\\
&\overset{(6.6.8)}{=} \sum_{i\in[n]\setminus\{r\}} p_{ri} && (31)\\
&= \sum_{i\in[n]\setminus\{r\}} \mathbb{P}(X_{ri}=1) && (32)\\
&= \sum_{i\in[n]\setminus\{r\}} \mathbb{E}\Big[\textstyle\sum_{i\in[n]\setminus\{r\}} X_{ri}\Big]. && (33)
\end{aligned}
$$

Wir suchen als nächstes den Maximum-Likelihood-Schätzer für den Parameter u. Die Likelihood eines Parameterwert ist die Wahrscheinlichkeit, die tatsächlich beobachtete Stichprobe zu realisieren, und die soll maximal sein. Die Log-Likelihood-Funktion $\ell\colon (0,\infty)^n \to \mathbb{R}$ lautet

$$
\begin{aligned}
\ell(u) &:= \log \prod_{k=1}^{m} \mathbb{P}_u(X = x^{(k)}) && (34)\\
&= \sum_{k=1}^{m} \Big(\log \prod_{i\in[n]} u_i^{d_i(x^{(k)})} - \log G(u)\Big) && (35)\\
&= \sum_{k=1}^{m} \sum_{i\in[n]} d_i(x^{(k)}) \log u_i - m \sum_{1\le i<j\le n} \log(1+u_i u_j). && (36)
\end{aligned}
$$

Zur Bestimmung des Maximums berechnen wir die Ableitungen nach u_r:

$$
\frac{\partial}{\partial u_r} \ell(u) = \sum_{k=1}^{m} \frac{d_r(x^{(k)})}{u_r} - m \sum_{i=1}^{r-1} \frac{u_i}{1+u_i u_r} - m \sum_{j=r+1}^{n} \frac{u_j}{1+u_r u_j}. \tag{37}
$$

Der Maximum-Likelihood-Schätzer $\hat{u} = (\hat{u}_i)_{i\in[n]}$ erfüllt also für alle $r \in [n]$:

$$
\frac{1}{m} \sum_{k=1}^{m} d_r(x^{(k)}) = \sum_{i=1}^{r-1} \frac{\hat{u}_i \hat{u}_r}{1+\hat{u}_i \hat{u}_r} + \sum_{j=r+1}^{n} \frac{\hat{u}_r \hat{u}_j}{1+\hat{u}_r \hat{u}_j} \tag{38}
$$

Das bedeutet: Der Maximum-Likelihood-Schätzer fittet den Parameter u an die Schätzwerte $\hat{d}_r$ für die Knotengrade.