Kapitel 10

Zufallsgraphen

Man könnte vermuten, dass ein Graph mit einer großen chromatischen Zahl lokal einen großen vollständigen Graphen enthalten muss, der die Färbungszahl hochtreibt. Dies ist jedoch nicht richtig. Falsch ist auch, dass ein solcher Graph zumindets kleine Kreise enthalten muss. Man kann sogar zeigen, dass Graphen mit beliebig hoher chromatischer Zahl existieren, die eine beliebig große Taillenweite haben, d.h. deren kleinster Kreis beliebig groß ist.

Diese Aussage konstruktiv zu beweisen, hat sich als schwierig erwiesen. Wir wollen in diesem Kapitel die Methode vorstellen, mit der es Erdös 1959 gelungen ist, den Beweis zu führen.

10.1 Informelle Einführung von Zufallsgraphen

Sei $V = \{1, \ldots, n\}$ eine Knotenmenge. Wir erzeugen zufällig einen Graphen aus V, indem wir für jedes Paar i, j würfeln, ob (i, j) eine Kante in E sein soll. Sei $p \in [0, 1]$ und $q = 1 - p$, so sei für alle (i, j) p die Wahrscheinlichkeit, dass die Kante auftritt, und q die Wahrscheinlichkeit, dass sie nicht auftritt.

Dieses Modell ist nicht das einzig denkbare. Es erlaubt u.a. nicht die Anzahl der generierten Kanten zu kontrollieren, was aus Sicht etwa der Analyse von Algorithmen, deren Laufzeit üblicherweise in der Knoten und Kantenzahl gemessen wird, nachteilig ist. Ein alternativer Ansatz könnte darin bestehen, alle Graphen mit m Kanten gleichwahrscheinlich zu erzeugen.

Wir bleiben aber bei dem ersten Modell und würfen also alle Kanten gleich wahrscheinlich und unabhängig voneinander aus. Die Wahrscheinlichkeit, dass wir einen bestimmten Graphen auf n Knoten und m Kanten auswürfeln, beträgt dann

$$p^m q^{\binom{n}{2} - m}.$$

Wir bezeichnen mit $\mathcal{G}(n, p)$ den Wahrscheinlichkeitsraum der Zufallsgraphen auf n Knoten und mit Kantenwahrscheinlichkeit p.

Zur Veranschaulichung: die Wahrscheinlichkeit, dass ein Graph G einen anderen Graphen H mit l Kanten auf den ersten k Knoten als Teilgraphen enthält, beträgt

$$P(H \subseteq G) = p^l.$$

Verlangen wir, dass der Graph induziert ist, so müssen wir die anderen Kanten ausschließen, d.h.

$$P(H \text{ induzierter Teilgraph von } G) = p^l q^{\binom{k}{2}-l}.$$

Die Wahrscheinlichkeit, dass G einen zu H isomorphen Teilgraphen enthält ist schwieriger zu berechnen, da sich die möglichen Kopien von H überlappen können. Daher sind die Eriegnisse, das sie als Untergraphen auftreten, nicht unabhängig voneinander. Sicherlich bildet die Summe der Einzelwahrscheinlichkeiten eine obere Schranke. Als Beispiel:

Lemma 10.1. *Die Wahrscheinlichkeit, dass $G \in \mathcal{G}(n,p)$ eine stabile Menge der Größe k mit $k \geq 2$ enthält, ist*

$$P(\alpha(G) \geq k) \leq \binom{n}{k} q^{\binom{k}{2}}.$$

Beweis: Die Wahrscheinlichkeit, dass eine Teilmenge $U \subseteq V$ mit $|U| = k$ keine Kanten enthält, ist $q^{\binom{k}{2}}$. Da $\binom{n}{k}$ solcher Teilmengen existieren, folgt die Behauptung. □

Ebenso gilt:

$$P(\omega(G) \geq k) \leq \binom{n}{k} p^{\binom{k}{2}}.$$

10.1.1 Zusammenhang zu den Ramseyzahlen

Ramsey hat 1930 in einem Satz speziell die folgende Aussage bewiesen:

Satz 10.2. *Zu jedem $k \geq 1$ existiert eine natürliche Zahl $R(k)$ so dass jeder Graph mit $R(k)$ Knoten entweder eine Clique der Größe k oder eine stabile Menge der Größe k enthält.* □

Der Beweis dieses überraschenden Ergebnisses (genauer einer allgemeineren Aussage) ist konstruktiv. Allerdings ist das genaue Verhalten der Funktion $R(k)$ nicht bekannt. Außer für kleine k, gibt es lediglich Schranken, u.a. über Zufallsgraphen.

Satz 10.3. *Für $k \geq 4$ ist $R(k) > 2^{k/2}$.*

Beweis: Sei $p = q = \frac{1}{2}$. Für $k \geq 4$ ist $k! > 2^k$ (Beweis per Induktion). Daraus und aus den obigen Aussagen über stabile Mengen und Cliquen folgt für $n \leq 2^{k/2}$:

$$\begin{aligned}
P(\alpha(G) \geq k) = P(\omega(G) \geq k) &\leq \binom{n}{k} \frac{1}{2}^{\binom{k}{2}} \\
&< (n^k/2^k) 2^{-\frac{1}{2}k(k-1)} \\
&\leq (2^{k^2/2}/2^k) 2^{-\frac{1}{2}k(k-1)} \\
&= 2^{-k/2} \\
&< \frac{1}{2}
\end{aligned}$$

Damit ist für $n \leq 2^{k/2}$ die Wahrscheinlichkeit, dass eines der Ereignisse echt kleiner als 1, d.h. es existiert mindestens ein Graph auf n Knoten, der weder eine Clique noch eine stabile Menge der Größe jeweils k enthält. □

10.2 Erwartungswerte

Jede Funktion, die einem Graphen einen nichtnegativen Wert zuordnet (chromatische Zahl, maximale Cliquengröße, Stabilitätszahl usw.) wird auf $\mathcal{G}(n,p)$ zu einer Zufallsvariablen:

$$x : \mathcal{G}(n,p) \to \mathbb{R}_+.$$

Der **Erwartungswert**

$$E(x) = \sum_{G \in \mathcal{G}(n,p)} x(G)P(G)$$

ist das mit der Wahrscheinlichkeit des Auftretens gewichtete Mittel der Zufallsvariablen. Insbesondere ist der Erwartungswert linear in seinen Argumenten, d.h. $E(ax+by) = aE(x) + bE(y)$. Der Erwartungswert ermöglicht es, Aussagen über das Verhalten von x zu erhalten.

Lemma 10.4 (Markov-Ungleichung). *Sei x eine Zufallsvariable auf $\mathcal{G}(n,p)$ und $t > 0$. Dann gilt:*

$$P(x \geq t) \leq E(x)/t.$$

Beweis: Es ist

$$\begin{aligned} E(x) &= \sum_{G \in \mathcal{G}(n,p)} x(G)P(G) \\ &\geq \sum_{\substack{G \in \mathcal{G}(n,p) \\ x(G) \geq t}} x(G) \cdot P(G) \\ &\geq \sum_{\substack{G \in \mathcal{G}(n,p) \\ x(G) \geq t}} P(G) \cdot t \\ &= P(x \geq t) \cdot t. \end{aligned}$$

□

Zur Veranschaulichung berechnen wir die erwartete Anzahl von Kreisen der Länge k in einem Graphen.

Satz 10.5. *Die mittlere Anzahl von Kreisen der Länge k in einem Graphen $G \in \mathcal{G}(n,p)$ ist*

$$E(x) = \frac{n!}{(n-k)! \cdot 2k} \cdot p^k.$$

Beweis: Für einen fetsgewählten Kreis C der Länge k auf der Knotenmenge V sei x_C die Zufallsvariable auf $\mathcal{G}(n,p)$, die den Wert 1 animmt, wenn $C \subseteq G$, und sonst 0 ist. Dann ist $E(x_C)$ die Wahrscheinlichkeit, dass $C \subseteq G$, also

$$E(x_C) = P(C \subseteq G) = p^k.$$

Jeder Kreis wird durch eine Folge $u_1, \ldots, u_k$ dargestellt. Davon gibt es $n \cdot (n-1) \cdot \ldots \cdot (n-k+1) = \frac{n!}{(n-k)!}$ viele, wovon jeweils $2k$ den gleichen Kreis darstellen. Da die Zufallsvariable x die Summe der Zufallsvariablen x_c ist, folgt aus der Linearität:

$$E(x) = E(\sum_C x_C) = \sum_C E(x_C) = \frac{n!}{(n-k)! \cdot 2k} \cdot p^k.$$

□

Im Folgenden wollen wir das Resultat von Erdös über Graphen mit großer Taillenweite und großer chromatischer Zahl diskutieren. Es scheint intuitiv klar, dass, wenn wir in unserem wahrscheinlichkeitstheoretischen Modell p klein wählen, also vor allem Graphen mit wenigen Kanten erzeugen, ein zufällig erzeugter Graph mit hoher Wahrscheinlichkeit keine kleine Kreis haben wird. Wählen wir umgekehrt p groß, so werden nur selten große stabile Mengen erzeugt werden. Die Frage ist, ob wir einen Kompromiss für p finden, der beide Zahlen groß werden lässt.

Unser Modell bleibt auch dann gültig, wenn wir p nicht konstant wählen, sondern z.B. als Funktion von n. Wir zeigen zuerst, dass, wenn p etwas langsamer fällt als $\frac{1}{n}$, wir fast sicher keine großen stabilen Mengen finden werden.

Lemma 10.6. *Sei $k > 0$ und $p \geq \frac{6k \ln n}{n}$, so gilt*

$$\lim_{n \to \infty} P(\alpha \ \geq \ \frac{1}{2} \cdot n/k) = 0.$$

Beweis: folgt aus dem Satz 10.1 und geeigneten Abschätzungen. □

Satz 10.7 (Erdös). *Zu jedem $k \in \mathbb{N}$ gibt es einen Graphen G mit Taillenweite $g(G) > k$ und chromatischer Zahl $\chi(G) > k$.*

Beweis: Für das vorgegebene k wähle ε so, dass $0 < \varepsilon < \frac{1}{k}$. Sei $p = n^{\varepsilon - 1}$. Sei $x(G)$ die Zufallsvariable, die jedem $G \in \mathcal{G}(n,p)$ die Anzahl seiner Kreise der Länge höchstens k zuordnet. Nach Satz 10.5 gilt:

$$E(x) = \sum_{i=3}^{k} \frac{n!}{(n-i)! \cdot 2i} \cdot p^i \leq \frac{1}{2} \sum_{i=3}^{k} n^i p^i \leq \frac{1}{2}(k-2) n^k p^k,$$

wobei die letzte Abschätzung aus $(np)^i \leq (np)^k$ folgt, da $np = n^\varepsilon \geq 1$. Aus der Markov-Ungleichung 10.4 folgt weiter:

$$\begin{aligned} P(x \geq \frac{n}{2}) &\leq E(x)/(\frac{n}{2}) \\ &\leq (k-2) n^{k-1} p^k \\ &= (k-2) n^{k-1} n^{(\varepsilon-1)k} \\ &= (k-2) n^{k\varepsilon - 1} \end{aligned}$$

Da $k\varepsilon - 1 < 0$ folgt nach Wahl von ε:

$$\lim_{n \to \infty} P(x \geq \frac{n}{2}) = 0.$$

Damit und nach Lemma 10.1 können wir n so groß wählen, dass $P(x \geq \frac{n}{2}) < \frac{1}{2}$ und $P(\alpha \ \geq \frac{1}{2} \cdot n/k) < \frac{1}{2}$. Dann existiert eine Graph $G \in \mathcal{G}(n,p)$ mit $\alpha(G) < \frac{1}{2} n/k$ und mit weniger als $\frac{n}{2}$ kurzen Kreisen. Sie H der Graph, der aus G entsteht, wenn wir aus jedem kurzen Kreis einen Knoten entfernen. Dann hat H immer noch mindestens $\frac{n}{2}$ Knoten, enthält aber keine kurzen Kreise mehr, d.h. $g(H) > k$. Weiter gilt nach Wahl von G:

$$\chi(H) \geq \frac{|H|}{\alpha(H)} \geq \frac{n/2}{\alpha(G)} > k.$$

□

10.3 Eigenschaften fast aller Graphen

Im folgenden wollen wir unter einer Eigenschaft eines Graphen eine Teilmenge aller Graphen verstehen, die unter Isomorphie abgeschlossen ist, d.h. hat G die Eigenschaft, so auch jeder zu G isomorphe Graph. Wir wollen untersuchen, wie sich für eine gegebene Eigenschaft und eine gegebene Auswahlfunktion $p(n)$ die Wahrscheinlichkeit verhält, dass ein Graph die Eigenschaft hat, wenn n gegen unendlich geht. Wir sagen, dass **fast alle** Graphen diese Eigenschaft haben, falls die Wahrscheinlichkeit gegen 1 strebt, und dass **fast kein** Graph diese Eigenschaft hat, falls der Grenzwert 0 ist. Zur Illustration wollen wir zeigen, dass jeder gegebene Graph als isomorphe Kopie eines induzierten Teilgraphen fast aller Graphen auftaucht.

Lemma 10.8. *Sei H ein Graph und $0 < p < 1$. Dann enthalten fast alle Graphen einen induzierten Teilgraphen, der isomorph zu H ist.*

Beweis: Sei $k = |V(H)|$. Für eine k-elementige Teilmenge U der n Knoten eines Zufallsgraphen $G \in \mathcal{G}(n,p)$ sei $r > 0$ die Wahrscheinlichkeit, dass $G(U)$ isomorph zu H ist. Wegen der Unabhängigkeit des Würfelns der Kanten, hängt diese Wahrscheinlichkeit nur von p, nicht aber von n. G enthält $\lfloor n/k \rfloor$ disjunkte k-elementige Teilmengen. Die Wahrscheinlichkeit, dass keine dieser Teilmengen eine isomorphe Kopie von H enthält, ist daher $(1-r)^{\lfloor n/k \rfloor}$. Damit folgt:

$$P(H \text{ ist nicht induzierter Teilgraph von } G) \leq (1-r)^{\lfloor n/k \rfloor} \xrightarrow[n \to \infty]{} 0.$$

$\square$

Für $i, j \in \mathbb{N}$ bezeichne $T_{i,j}$ die Eigenschaft, dass der Graph mindestens $i + j$ Knoten enthält und zu je zwei disjunkten Teilmengen U_1 und U_2 mit $|U_1| \leq i$ und $|U_2| \leq j$ stets einen Knoten $v \notin (U_1 \cup U_2)$ enthält, der zu allen Knoten aus U_1 und zu keinem Knoten aus U_2 benachbart ist.

Lemma 10.9. *Sei $0 < p < 1$ und $i, j \in \mathbb{N}$. Dann hat fast jeder Graph die Eigenschaft $T_{i,j}$.*

Beweis: Die Wahrscheinlichkeit, dass in einem Graph G zu gegebenen U_1, U_2 ein gegebener Knoten $v \in V \setminus (U_1 \cup U_2)$ zu allen Knoten aus U_1 und zu keinem Knoten aus U_2 benachbart ist, beträgt

$$p^{|U_1|} q^{|U_2|}.$$

Wegen der Unabhängigkeit ist für $n \geq i + j$ somit die Wahrscheinlichkeit, dass im Restgraphen für gegebene Mengen U_1, U_2 kein solcher Knoten existiert

$$(1 - p^{|U_1|} q^{|U_2|})^{n - |U_1| - |U_2|} \leq (1 - p^i q^j)^{n-i-j}.$$

Es existieren höchstens $\binom{n}{i+j} \leq n^{i+j}$ disjunkter Teilmengen U_1, U_2 in V. Die Wahrscheinlichkeit, dass darunter ein Paar ohne ein entsprechendes v ist, beträgt somit höchstens

$$n^{i+j} (1 - p^i q^j)^{n-i-j}.$$

Da $1 - p^i q^j < 1$, geht dieser Wert für $n \to \infty$ gegen 0. $\square$

Korollar 10.10. *Für $0 < p < 1$ und $k \in \mathbb{N}$ ist fast jeder Graph k-zusammenhängend.*

Beweis: Nach Lemma 10.9 reicht es zu zeigen, dass jeder Graph mit Eigenschaft $T_{2,k-1}$ k-zusammenhängend ist. Dies folgt aber aus dem Satz von Menger, da je zwei nichtbenachbarte Knoten von keiner Teilmenge U_2 der Kardinalität $k-1$ getrennt werden können, denn sie haben immer noch einen gemeinsamen Nachbarn außerhalb von U_2. □

Wir haben sogar eine schärfere Aussage gezeigt: fast alle Graphen haben maximale Distanz zwei. Ebenso lässt sich zeigen, dass fast alle Graphen eine hohe chromatische Zahl in der Größenordnung von $n/\log n$ haben.

10.4 Schwellenfunktionen

Wir beschränken uns in diesem Abschnitt auf **monotone Eigenschaften**, d.h. Eigenschaften, die erhalten bleiben, wenn Kanten hinzugefügt werden. Wir wollen versuchen, für monotone Eigenschaften Schwellenfunktionen zu bestimmen. Dabei nennen wir eine Funktion $t(n)$ **Schwellenfunktion** für eine Eigenschaft T, wenn aus $p(n)/t(n) \to 0$ folgt, dass fast kein Graph die Eigenschaft T hat, und aus $p(n)/t(n) \to \infty$ folgt, dass fast alle Graphen die Eigenschaft T haben.

Wir werden dabei für eine Zufallsvariable x mit den bisher bekannten Methoden das Ergebnis $E(x) \to \infty$ ableiten können. Wir würden daraus gerne die Aussage $P(x=0) \to 0$ ableiten. Obwohl dieser Schluss naheliegt, ist er i.A. nicht richtig: betrachte etwa $P(x=0) = 1/2$ und $P(x=n) = 1/2$. Der gewünschte Schluss ist jedoch möglich, wenn wir sicherstellen, dass die Werte sich nicht sehr weit vom Erwartungswert entfernen. Dazu betrachten werden typischerweise **zweite Momente**, d.h. die Erwartung von x^2, und die **Varianz** $E[(x-E(x))^2]$ betrachtet.

Lemma 10.11 (Tschebyschev-Ungleichung). *Für jedes reelle $t > 0$ gilt:*

$$P(|x-E(x)| \geq t) \leq E[(x-E(x))^2]/t^2.$$

Beweis: Aus der Markov-Ungleichung 10.4 folgt:

$$P(|x-E(x)| \geq t) = P((x-E(x))^2 \geq t^2) \leq E[(x-E(x))^2]/t^2.$$

□

Als Folgerung ergibt sich:

Lemma 10.12. *Gilt $E(x^2)/E(x)^2 \to 1$, so folgt $P(x=0) \to 0$.*

Beweis: Für Graphen mit $x(G) = 0$ gilt $|x(G) - E(x)| = E(x)$. Somit folgt aus Lemma 10.11

$$P(x=0) \leq P(|x-E(x)| \geq E(x)) \leq E[(x-E(x))^2]/E(x)^2 = (E(x^2)-E(x)^2)/E(x)^2 \to 0.$$

□

Zur Veranschaulichung des Konzeptes der Schwellenfunktion leiten wir eine Schranke dafür her, dass jeder Knoten mindestens einen Nachbarn hat. Dies ist etwas schwächer als den Zusammenhang zu fordern. Dafür trennen wir das Verhalten schärfer als bei den Schwellenfunktionen gefordert. Dazu einige Vorüberlegungen:

1) Ist $p = c \ln n/n$ für eine beliebige Konstante c, so folgt $np^2(1/2 + p/3 + p^2/4 + \ldots) \to 0$,
2) Für $-1 \leq x < 1$ ist $\ln(1-x) = -x - \frac{x^2}{2} - \frac{x^3}{3} - \ldots$,
3) $ne^{-np} = n^{1-c}$.

Lemma 10.13. *Sei $p = c \ln n/n$ für eine Konstante $c > 1$. Dann haben fast alle Graphen in $\mathcal{G}(n, p)$ keine isolierten Knoten.*

Beweis: Sei x_i die Indikatorfunktion, die anzeigt, ob der Knoten v_i isoliert ist, und x die Zufallsvariable der Anzahl der isolierten Knoten. Dann ist $E(x) = \sum_i E(x_i) = n(1-p)^{(n-1)}$. Mit der obigen Bemerkung 2 folgt:

$$(1-p)^n = e^{n \ln(1-p)} = e^{-np} e^{-np^2(1/2+p/3+\ldots)}.$$

Mit Bemerkung 1 folgt dann $(1-p)^n \sim e^{-np}$, und da $(1-p)^{-1} \sim 1$, ergibt sich $E(x) \sim ne^{-np}$. Nach Bemerkung 3 ist dann $E(x) \sim n^{1-c}$. Da $c > 1$, folgt $E(x) \to 0$. □

Lemma 10.14. *Sei $p = c \ln n/n$ für eine Konstante $c < 1$. Dann haben fast alle Graphen in $\mathcal{G}(n, p)$ isolierte Knoten.*

Beweis: Für $c < 1$ folgt aus den obigen Überlegungen $E(x) \to \infty$. Wegen der Linearität des Erwartungswertes und der Eigenschaft $x_i^2 = x_i$ der Indikatorfunktionen, folgt $E(x^2) = \sum_{i=1}^n E(x_i^2) + 2\sum_{i<j} E(x_i x_j) = E(x) + n(n-1)E(x_i x_j)$. Das Produkt $x_i x_j$ ist wiederum eine Indikatorvariable, die genau dann den Wert 1 annimmt, wenn beide Knoten isoliert sind. Damit ist $E(x_i x_j) = (1-p)^{2n-3}$. Da wie vorher $(1-p)^n \sim e^{-np}$, folgt $E(x_i x_j) \sim e^{-2np}$. Somit ergibt sich $E(x^2) \sim E(x) + E(x)^2$ und somit $E(x^2)/E(x)^2 \to 1$. Aus dem Lemma 10.12 folgt dann $P(x = 0) \to 0$. □

Korollar 10.15. *Die Funktion $\ln n/n$ ist eine Schwellenfunktion für das Verschwinden isolierter Knoten.* □

Mit ähnlichen Methoden lässt sich das folgende allgemeine Resultat zeigen. Dazu heiße ein Graph H **ausgewogen**, wenn der durchschnittliche Knotengrad in einem Teilgraphen nie größer ist als der von H.

Satz 10.16. *Ist H ein ausgewogener Graph mit k Knoten und l Kanten, so ist $t(n) = n^{-k/l}$ eine Schwellenfunktion für das Auftreten von H als Teilgraph.* □

Als Folgerungen ergeben sich aus diesem Satz:

Korollar 10.17. *Für $k \geq 3$ ist $t(n) = 1/n$ eine Schwellenfunktion für das Auftreten eines Kreises der Länge k.* □

Insbesondere ist die Schwellenfunktion unabhängig von k. Damit treten ab der Kantenauswahlwahrscheinlichkeit $1/n$ bereits Kreise jeder konstanten Länge auf.

Korollar 10.18. *Ist T ein Baum auf k Knoten, so ist $t(n) = n^{-k/(k-1)}$ eine Schwellenfunktion für das Auftreten von T als Teilgraph.* □

Lässt man die Kantenwahrscheinlichkeit langsam anwachsen, so zeigt sich folgendes Verhalten:

- unterhalb von $1/n^2$ hat fast jeder Graph nur isolierte Knoten,
- ab $1/n^2$ tauchen erste Kanten auf, die noch isoliert bleiben,
- ab $1/n^{\frac{3}{2}}$ hat fast jeder Graph Komponenten mit mehr als zwei Knoten,
- bei $n^{-(1+1/k)}$ haben die Graphen Bäume mit $k+1$ Knoten, aber sind weiterhin kreisfrei. Für jedes k sind dies Schwellenfunktionen, die das Auftreten von Bäumen mit $k+1$ Knoten von denen mit $k+2$ Knoten trennen,
- ab $1/n$ tauchen Kreise auf,
- mit wachsendem p erhalten die Kreise Sehnen, d.h. die Graphen sind nicht mehr planar. Eine der Komponenten schwillt an,
- bis bei $\ln n/n$ diese Komponente die anderen verschluckt und die Graphen zusammenhängend werden,
- danach hat für alle k jeder Knoten mindestens k Nachbarn.