Die Probabilistische Methode

Daniel Werner, Michael Schnürmacher

26. Mai 2005

1 Einführung und Motivation

Die Probabilistische Methode ist ein Verfahren für den Beweis der Existenz von kombinatorischen Objekten oder Algorithmen, die nur schwer oder gar nicht explizit konstruiert werden können. Sie bezieht ihre Mächtigkeit vor allem aus den Gesetzen der Wahrscheinlichkeitstheorie und ihren Folgerungen (Linearität des Erwartungswertes, Tschebytscheff-Ungleichung, Markov-Ungleichung etc.). Sie stützt sich im Wesentlichen auf die beiden folgenden, zunächst trivial erscheinenden Tatsachen:

PM1 Eine Zufallsvariable nimmt wenigstens einen Wert an, der $\geq$ seinem Erwartungswert ist.

PM2 Wenn wir aus einer Menge von zulässigen Objekten ein Element mit bestimmten Eigenschaften mit positiver Wahrscheinlichkeit ziehen, so existiert mindestens ein solches Objekt mit diesen Eigenschaften.

Begründer dieser Methode ist Paul Erdös (1913-1996), der ab 1947 eine beachtliche Anzahl von Resultaten mit dieser Methode erziehlte.
Anwendungsgebiete sind z.B. die Graphentheorie ($\rightarrow$ 2.2, 2.4, 2.5), Kombinatorik ($\rightarrow$ 2.5), Zahlentheorie, Statistik ($\rightarrow$ 2.1, 4.1), Geometrie, Analysis, Numerik und die Informatik, speziell randomisierte Algorithmen ($\rightarrow$ 2.1, 2.3, 2.4).

2 Anwendungen

2.1 Das Set-Balancing-Problem

Definition: Gegeben ist eine $n \times n$ Matrix $\mathbf{A}$ mit Einträgen aus $\{0, 1\}$. Es soll nun ein Vektor $\mathbf{b} \in \{1, -1\}^n$ gefunden werden, der $\|\mathbf{Ab}\|_\infty$ minimiert.

Ein „Algorithmus“, der jedes b_i mit Wahrscheinlichkeit $1/2$ mit 1 und mit Wahrscheinlichkeit $1/2$ mit -1 belegt, findet mit Wahrscheinlichkeit $1 - 2/n$ einen Wert für $\mathbf{b}$, sodass $\|\mathbf{Ab}\|_\infty$ nicht größer als $4\sqrt{n \ln n}$ ist. Daraus können wir wg. **PM2** folgern, dass es mindestens einen solchen Vektor gibt.

2.2 Das MAX-CUT-Problem

Gegeben ein ungerichteter Graph $G = (V, E)$ mit n Knoten und m Kanten. Wir möchten eine Partition der Knoten in zwei disjunkte Mengen A und B finden, sodass

$$|\{(u, v) \in E | u \in A, v \in B\}|$$

maximal wird. Dieses Problem ist **NP**-hart, mit Hilfe der Probabilistischen Methode können wir jedoch folgende (überraschende?) Tatsache zeigen:

Satz: Es existiert immer eine Zerlegung der Knoten in disjunkte Teilmengen, sodass

$$|\{(u,v) \in E | u \in A, v \in B\}| \geq m/2$$

Beweisidee: Wir machen folgendes Gedankenexperiment: Jeder Knoten wird zufällig und gleichverteilt einer der beiden Mengen zugeordnet. Die Wahrscheinlichkeit dafür, dass eine Kante Endpunkte in verschiedenen Mengen hat, ist also 1/2. Wegen der Linearität des Erwartungswertes gilt für die erwartete Anzahl k der Kanten mit Ecken in verschiedenen Mengen $k \geq m/2$. Also existiert wegen **PM2** ein solcher Schnitt. □

2.3 k-MAX-SAT-Problem

Gegeben ist eine Klauselmenge Φ mit n Variablen, m Klauseln und minimaler Klauselläng k. Finde eine Belegung der Variablen von Φ, die die Anzahl der erfüllten Klauseln maximiert.

Satz: Es existiert immer eine Belegung der Variablen, die mindestens $(1-2^{-k}) \cdot m$ Klauseln erfüllt.

Beweisidee: Wir belegen jede Variable mit Wahrscheinlichkeit $\frac{1}{2}$ mit $true$ bzw. $false$. Eine Klausel C_j mit $C_j \geq k$ ist nicht erfüllt $\Leftrightarrow$ alle Literale in C_j sind $false$ $\Rightarrow$ $Pr[C_j = false] \leq 1 - 2^{-k}$ und daraus folgt mit

$$\hat{C}_i = \begin{cases} 1 & \text{falls } C_i = true \\ 0 & \text{sonst} \end{cases}$$

dass die erwartete Anzahl der erfüllten Klauseln

$$E[\sum_{i=1}^{m} \hat{C}_i] = \sum_{i=1}^{m} E[\hat{C}_i] \geq m \left(1 - \frac{1}{2^k}\right)$$

ist. Aus **PM1** folgt dann, dass eine solche Belegung existiert. □

2.4 Expandergraphen

2.4.1 Motivation

Expandergraphen haben zahlreiche Anwendungen in der Informatik und bei Telefonvermittlungsnetzwerken. Es handelt sich hierbei intuitiv um Multigraphen, die stark zusammenhängend aber wenig dicht sind. Wir beschäftigen uns hier mit dem Spezialfall des OR-concentrators.

Definition: Ein (n, d, α, c)-OR-concentrator ist ein bipartiter Multigraph $G(L, R, E)$ mit unabhängigen Mengen von Knoten L und R, die jeweils Größe n haben und für den gilt:

- Jeder Knoten hat höchstens Grad d.
- Für jede Teilmenge $S \subset V$ mit $|S| \leq \alpha n$ gibt es mindestens $c|S|$ Nachbarn in R.

In den meisten Anwendungen möchte man d möglichst klein und c möglichst groß machen. Es ist jedoch nicht klar, ob solche Graphen i.A. überhaupt existieren, da die Bedingungen sehr streng sind.

2.4.2 Beispiel

Theorem: Für alle $n > 0$ existiert ein $(n, 18, \frac{1}{3}, 2)$-OR-concentrator.

Beweisidee: Sei $\varepsilon_s := \{$Es existiert eine s-Untermenge $S \subset L$ mit weniger als cs Nachbarn in $R\}$. Wir zeigen

$$\sum_{s \geq 1}^{\alpha n} Pr[\varepsilon_s] < 1.$$

Die Existenz folgt dann wieder aus **PM2**. □

2.5 Ramsey-Zahlen und ein Satz von Erdös

2.5.1 Allgemeine Problemformulierung

Sei $G = (V, E)$ ein Graph. Wie groß muss $N = |V|$ sein, damit G immer entweder eine k-Clique oder eine l-Anticlique besitzt? Die kleinste solche Zahl wird Ramseyzahl $R(k, l)$ genannt.

2.5.2 Obere Schranke

Satz (Ramsey): Es gilt

$$R(k, l) \leq \binom{k + l - 2}{k - 1}$$

Beweisidee: Doppelinduktion über k, l. Wir zeigen

$$R(k, l) \leq R(k - 1, l) + R(k, l - 1).$$

Aus der Rekursionsgleichung des Binomialkoeffizienten folgt dann direkt der Satz von Ramsey. □

Korollar: Es ist $R(k, k) \leq 2^{2(k-1)}$.

2.5.3 Untere Schranke

Man kann eine untere Schranke für $R(k, l)$ herleiten, wir beschränken uns hier aber auf $R(k, k)$.

Satz (Erdös): Für alle $k \geq 2$ gilt

$$R(k, k) \geq 2^{k/2}.$$

Beweisidee: Wir konstruieren einen Zufallsgraphen der Größe $N < 2^{k/2}$ und zeigen, dass mit positiver Wahrscheinlichkeit weder eine k-Clique noch eine k-Anticlique existiert. Daraus folgt wg. **PM2**, dass ein solcher Graph existiert. □

3 Das Lovász Local Lemma

3.1 Motivation

Angenommen, wir haben paarweise unabhängige Ereignisse $\varepsilon_1, \ldots, \varepsilon_n$ mit $Pr[\varepsilon_i] \leq \frac{1}{2}$ für alle i. Dann ist die Wahrscheinlichkeit dafür, dass keines dieser Ereignisse eintritt $Pr[\overline{\varepsilon_1} \cap \cdots \cap \overline{\varepsilon_n}] \geq 2^{-n}$. Problem: Bei abhängigen Ereignissen können wir solche Aussagen nicht so einfach treffen. Das LLL hilft uns jedoch dann, wenn alle Ereignisse nur von maximal einer konstanten Anzahl d anderer Ereignisse abhängen.

3.2 Das symmetrische LLL

Satz: Sei $G = (V, E)$ ein Abhängigkeitsgraph für die Ereignisse $\varepsilon_1, \ldots, \varepsilon_n$ in einem Wahrscheinlichkeitsraum Ω. Wenn

- $\forall i \in \{1, \ldots, n\} : Pr[\varepsilon_i] \leq p$ mit $p \in [0, 1]$
- jedes Ereignis ist von höchstens d Ereignissen abhängig
- $ep(d + 1) \leq 1$

gilt, so ist $Pr[\overline{\varepsilon_1} \cap \cdots \cap \overline{\varepsilon_n}] > 0$.

Beweisidee: Wir zeigen

$$\forall i \in \{1, \ldots, n\} \quad \forall S \subset \{1, \ldots, n\} - i : \qquad Pr[\varepsilon_i | \bigcap_{j \in S} \overline{\varepsilon_j}] \leq \frac{1}{1 + d}$$

denn dann ist

$$Pr[\bigcap_{i=1}^{n} \overline{\varepsilon_i}] > 0 \qquad \square$$

3.3 Das allgemeine LLL

Satz: Sei $G = (V, E)$ ein Abhängigkeitsgraph für die Ereignisse $\varepsilon_1, \ldots \varepsilon_n$ in einem Wahrscheinlichkeitsraum Ω. Wenn $\forall i \in \{1, \ldots, n\} \quad \exists x_i \in [0, 1] :$

$$Pr[\varepsilon_i] \leq x_i \prod_{(i,j) \in E} (1 - x_j)$$

dann gilt

$$Pr[\bigcap_{i=1}^{n} \overline{\varepsilon_i}] \geq \prod_{i=1}^{n} (1 - x_i) > 0$$

4 Die Methode der bedingten Wahrscheinlichkeiten

Wir illustrieren diese Methode anhand des bereits erwähnten Set-Balancing Problems.

4.1 Derandomisierung des Set-Balancing Algorithmus'

Wir haben einen randomisierten Algorithmus entwickelt, dessen „erwartete" Lösung vernünftig[1] ist. Wir möchten jedoch diesen Algorithmus derandomisieren, um eine garantierte Güte zu bekommen. Dazu betrachten wir den Entscheidungsbaum und bestimmen für jedes der Kinder eines Knotens die[2] Wahrscheinlichkeiten dafür, dass unser Wert in einer Zeile größer als $4\sqrt{n \ln n}$ wird. In unserem speziellen Beispiel setzen wir

$$\hat{P}(a) := \sum_{i=1}^{n} Pr[\varepsilon_i | a] \quad (\geq Pr[\varepsilon | a])$$

wobei $\varepsilon_1, \ldots, \varepsilon_n$ die Wahrscheinlichkeiten dafür sind, dass das Produkt mit der i-ten Zeile $4\sqrt{n \ln n}$ übersteigt. Diesen Wert können wir in Polynomialzeit berechnen und wir können zeigen, dass jeder Knoten ein Kind hat, dessen bedingte Wahrscheinlichkeit eines Versagens kleiner als seine eigene ist $(*)$. Da die Wahrscheinlichkeit der Wurzel $\hat{P}(r) \leq 2/n < 1$ ist, jedes Blatt ($\hat{=}$ kompletter Vektor $\mathbf{b}$) nur die Wahrscheinlichkeit eines Versagens von 0 oder 1 hat (klar?) und die Wahrscheinlichkeit nie vergrößert wird, erreichen wir auf diese Weise einen Vektor $\mathbf{b}$ mit $\|\mathbf{Ab}\|_\infty \leq 4\sqrt{n \ln n}$. Im Allgemeinen lassen sich jedoch die entsprechenden Wahrscheinlichkeiten nicht so einfach berechnen, insbesondere gilt $(*)$ i.A. nicht.

[1] $\leq 4\sqrt{n \ln n}$

[2] durch die Wahl der bisherigen Werte von $b' \in \{1, -1\}^i$ bedingten