Diskrete Optimierung: Fallstudien aus der Praxis

Barbara Wilhelm | Michael Ritter

Randomisierte Methoden für k-Matching

Im Folgenden finden Sie eine formale Beschreibung des gewählten Approximationsalgorithmus. Arbeiten Sie die Beschreibung durch, machen Sie sich Notizen. Versuchen Sie, die Schritte des Algorithmus und die Beweise mit Hilfe von kleinen Skizzen nachzuvollziehen. Überlegen Sie sich, wie Sie die Erklärung dieses Algorithmus gestalten (Struktur, wichtige Ideen, Resultate). Wenn Sie Fragen haben, wenden Sie sich an Ihre Betreuer.

Problem: k-Matching
Input: Ein Graph $G = (V, E)$, eine Zahl $k \in \mathbb{N}$.

Aufgabe: Finde ein k-Matching in G, das möglichst viele Kanten enthält. Ein k-Matching ist eine Kantenmenge $M \subseteq E$ mit der Eigenschaft, dass jeder Knoten höchstens mit k Kanten aus M inzident ist.

Der Approximationsalgorithmus für k-Matching, den wir hier behandeln werden, hebt sich etwas von den anderen Approximationsalgorithmen ab, die Sie in der Galerie gesehen haben. Er produziert nämlich nicht unbedingt eine zulässige Lösung und garantiert auch nicht unter allen Umständen eine gute Approximation. (Damit handelt es sich eigentlich gar nicht um einen Approximationsalgorithmus, sondern nur um eine Heuristik.) Wir werden aber beweisen, dass zumindest mit einer gewissen Wahrscheinlichkeit eine zulässige Lösung und eine gute Approximation gefunden werden. Benötigt man eine zulässige Lösung oder will man die Güte der Approximation erhöhen, so kann man das einfach durch mehrmaliges Anwenden des Algorithmus erreichen.

Im Folgenden sei der Graph $G = (V, E)$ gegeben durch seine (Knoten-Kanten-) Inzidenzmatrix $A \in \{0, 1\}^{n \times m}$. Das k-Matching-Problem lässt sich dann als ganzzahliges lineares Programm formulieren:

$$\begin{aligned} \max \sum_{i=1}^{m} x_i & \\ Ax &\leq k\mathbb{1} \\ x &\in \{0, 1\}^m \end{aligned}$$

Die Variable x_i einer ganzzahligen optimalen Lösung ist genau dann 1, wenn die zugehörige Kante im k-Matching enthalten ist, sonst 0.

Die Methode des randomisierten Rundens beruht nun darauf, zunächst die LP-Relaxation des obigen ILP zu lösen, man erhält dann eine evtl. nicht-ganzzahlige Lösung $x' \in [0, 1]^m$. Anschließend rundet man das Ergebnis zufällig. Eine ganzzahlige „Lösung“ x^* entsteht, indem für jedes $i \in \{1, \dots, m\}$ eine „unfaire Münze“ geworfen wird: Mit Wahrscheinlichkeit x'_i wird x^*_i auf 1 gesetzt, mit Wahrscheinlichkeit $1 - x'_i$ auf 0. Die Komponenten der relaxierten Optimallösung, die ohnehin schon nahe an 1 sind, werden also mit großer Wahrscheinlichkeit auf 1 gesetzt (analog für 0) und man kann hoffen, dass das Ergebnis nicht zu weit von der ganzzahligen Optimallösung entfernt ist (immerhin ist es ja „nahe“ am relaxierten Optimum). Dabei kann es natürlich vorkommen, dass ein Vektor x^* entsteht, der keine zulässige Lösung für das k-Matching-Problem darstellt. Um diesem Problem zu begegnen, rundet man etwas vorsichtiger (beachte, dass $x = 0$ immer eine zulässige Lösung des k-Matching-Problems ist). Der Algorithmus im Detail:

1. Eingabe: Inzidenzmatrix A eines Graphen, $k \in \mathbb{N}$ und $\varepsilon \in (0, 1)$.
2. Löse die LP-Relaxation des k-Matching-ILP. Die relaxierte Optimallösung sei $x' \in [0, 1]^m$, der Zielfunktionswert dieser Lösung sei ζ'.
3. Erzeuge einen Vektor $x^* \in \{0, 1\}^m$ durch randomisiertes Runden aus x' mit

$$\begin{aligned} &\mathbb{P}[x_i^* = 1] = (1 - \varepsilon/2)x_i' \\ \text{und}\quad &\mathbb{P}[x_i^* = 0] = 1 - \mathbb{P}[x_i^* = 1] = 1 - (1 - \varepsilon/2)x_i'. \end{aligned}$$

Satz 1
Sei $G = (V, E)$ ein Graph mit Inzidenzmatrix $A \in \{0, 1\}^{n \times m}$, $k \in \mathbb{N}$ und OPT sei die Kardinalität eines maximalen k-Matchings in G. Weiter sei $\varepsilon \in (0, 1)$ so gewählt, dass $k \geq \frac{6\ln(4n)(2-\varepsilon)}{\varepsilon^2}$. Dann produziert obiger Algorithmus einen Vektor $x^* \in \{0, 1\}^m$, für den mit einer Wahrscheinlichkeit von mindestens $1/2$ gilt:

1. *$Ax^* \leq k$ (x^* ist also ein zulässiges k-Matching)*
2. *$\sum_{i=1}^m x_i^* \geq (1 - \varepsilon)OPT$*

Für den Beweis benötigen wir eine Ungleichung, die Sie vielleicht aus der Stochastik kennen, die *Angluin-Valiant-Ungleichung*.

Lemma 2 (Angluin-Valiant)
Seien $Z_1, \dots, Z_r$ unabhängige Zufallsvariablen mit Werten in $\{0, 1\}$, $w_1, \dots, w_r \in [0, 1]$, und sei $p_i := \mathbb{P}[Z_i = 1]$, $q_i := \mathbb{P}[Z_i = 0] = 1 - p_i$ für alle $i \in \{1, \dots, r\}$. Dann gelten für $Z := \sum_{i=1}^r w_i Z_i$ und für jedes $\beta \in (0, 1]$ die Ungleichungen

$$\begin{aligned} &\mathbb{P}\Big[Z \geq (1 + \beta)\mathbb{E}[X]\Big] \leq \exp\left(\frac{-\beta^2\mathbb{E}[X]}{3}\right) \\ \textit{und}\ &\mathbb{P}\Big[Z \geq (1 - \beta)\mathbb{E}[X]\Big] \leq \exp\left(\frac{-\beta^2\mathbb{E}[X]}{2}\right) \end{aligned}$$

Beweis (von Satz 1). Wir definieren zunächst Zufallsvariablen

$$X_i := (Ax^*)_i \text{ für alle } i \in \{1, \dots, m\}.$$

Sei x' die LP-Optimallösung aus dem Algorithmus, dann gilt für jedes i:

$$\mathbb{E}[X_i] = \mathbb{E}\left[\sum_{j=1}^m a_{ij}x_j^*\right] = \sum_{j=1}^m a_{ij}\mathbb{E}\left[x_j^*\right] = \sum_{j=1}^m a_{ij}(1 - \frac{\varepsilon}{2})x_j' = (1 - \frac{\varepsilon}{2})\sum_{j=1}^m a_{ij}x_j' \leq (1 - \frac{\varepsilon}{2})k,$$

da x' ein zulässiges k-Matching ist.

Wir setzen $\delta := (1 - \varepsilon/2)k - \mathbb{E}[X_i] \geq 0$ und definieren Zufallsvariablen $Z_i := X_i + \delta$. Damit gilt $\mathbb{E}[Z_i] = (1 - \varepsilon/2)k$. Mit $\beta := \frac{\varepsilon}{2-\varepsilon}$ ist $0 < \beta \leq 1$ und es gilt

$$\mathbb{P}[X_i > k] \leq \mathbb{P}[X_i + \delta > k] = \mathbb{P}[Z_i > k] = \mathbb{P}\left[Z_i > (1+\beta)(1-\varepsilon/2)k\right] = \mathbb{P}\Big[Z_i > (1+\beta)\mathbb{E}[Z_i]\Big]$$

Wir definieren nun konstante Zufallsvariablen $X_{m+1}, \dots, X_{\lceil\delta\rceil}$ mit $\mathbb{P}[X_i = 1] = 1$ für $j \in \{m+1, \dots, \lceil\delta\rceil\}$ und setzen $w_{m+1} = \cdots = w_{\lceil\delta\rceil-1} = 1$, $w_{\lceil\delta\rceil} = \delta - \lfloor\delta\rfloor$. Mit $Z_i = \sum_{i=1}^m a_{ij}x_j^* + \sum_{i=m+1}^{\lceil\delta\rceil} w_i X_i$ und der Ungleichung von Angluin-Valiant folgt dann aus dem vorherigen

$$\mathbb{P}[X_i > k] \leq \exp\left(\frac{-\beta^2\mathbb{E}[Z_i]}{3}\right) =: B.$$

Da $\mathbb{E}[Z_i] = (1 - \varepsilon/2)k$, gilt weiter

$$\frac{\beta^2\mathbb{E}[Z_i]}{3} = \frac{1}{3}\frac{\varepsilon^2}{(2-\varepsilon)^2}(1-\varepsilon/2)k = \frac{k}{6}\frac{\varepsilon^2}{2-\varepsilon}.$$

Es gilt weiterhin

$$\begin{aligned}
& B \leq \frac{1}{4n} \\
\Leftrightarrow & \exp\left(\frac{-\beta^2\mathbb{E}[Z_i]}{3}\right) \leq \frac{1}{4n} \\
\Leftrightarrow & \frac{-\beta^2\mathbb{E}[Z_i]}{3} \leq \ln\left(\frac{1}{4n}\right) \\
\Leftrightarrow & \frac{\beta^2\mathbb{E}[Z_i]}{3} \geq \ln(4n) \\
\Leftrightarrow & \frac{k}{6}\frac{\varepsilon^2}{2-\varepsilon} \geq \ln(4n) \\
\Leftrightarrow & k \geq \frac{6(2-\varepsilon)\ln(4n)}{\varepsilon^2}
\end{aligned}$$

Nach Voraussetzung folgt also $B \leq \frac{1}{4n}$ und damit

$$\mathbb{P}\left[(Ax^*)_i > k \text{ für mindestens ein } i \in \{1, \dots, n\}\right] \leq \sum_{i=1}^{n} \mathbb{P}[(Ax^*)_i > k] \leq \frac{1}{4}.$$

Ähnlich (etwas einfacher) rechnet man auch nach, dass $\mathbb{P}\left[\sum_{i=1}^m x_i^* < (1-\varepsilon)OPT\right] \leq \frac{1}{4}$ ist. Zusammen ergibt sich dann die Behauptung (warum?). $\square$