# Smoothed Analysis

Die bisherige Analyse von Algorithmen war in einen hohen Grad unvollkommen. Es existiert doch eine große Klasse von Algorithmen wie z.B. Quicksort oder Simplex, die verbreitet sind, obwohl ihre Komplexität (O(n^k) bzw. O(2^n)) sehr pessimistisch scheint zu sein. Diese Divergenz zwischen Theorie und Praxis folgt daraus, dass die bisherigen Methoden der Analyse sich auf die Untersuchung von zwei Fällen (worst case and average case) stützten. Die beiden Fälle können uns aber täuschende Aussagen über einen Algorithmus liefern. Z.B. worst case legt uns nahe, dass ein gewisser Algorithmus langsam ist, indem Eingabedaten analysiert werden die in Praxis sehr selten oder sogar nie vorkommen können. Deshalb hat Spielnam und Teng im 2001 Jaher eine neue Methode – „smoothed analysis“ vorgeschlagen, die als eine Interpolation zwischen worst-case und average case gedacht ist. Inzwischen hat sich aber herausgestellt, dass die ausgeglichene Analyse auch in der Numerik eine große Rolle spielt. (Konditionsabschätzung)
Die Idee, die da hinter steckt, ist die Stellen oder ganze Umgebungen zu lokalisieren, wo der Algorithmus eine schlechte Laufzeit hat.

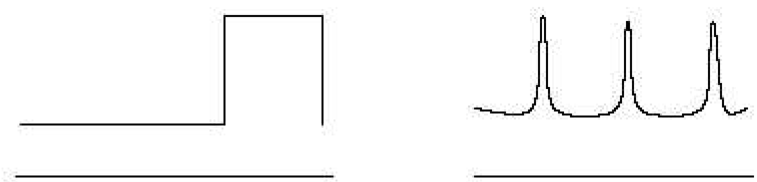
**Def 1.**

Die ausgeglichene Komplexität von einem Algorithmus definiert man als [1] :

$$\max_{x} \mathrm{E}_{y\in U_\epsilon(x)} \quad C(y)$$

wobei x - die Eingabedaten sind
C(y) – die Laufzeit der Algorithmus zur Eingabe y

**Bespiel 1**



Im linken Fall ist die smoothed Komplexität gleich dem worst case Fall. Im rechten Fall die smoothed Komplexität jedoch fällt als eine Funktion von Epsylon.

**Beispiel 1**

Betrachte das lineare Gleichungsystem:

Ax=b

wobei A regulär ist.
Dann ist das o.g. Gleichungssystem eindeutig lösbar
mit der Lösung: $x=A^{-1}*b$

Der Einfluss der  Störung der Koeffizientenmatrix A auf die Lösung ist gegeben durch:

$$\mathrm{Kond}(A) := \left\| A \right\| \cdot \left\| A^{-1} \right\|$$

Nun besteht die Frage, ob ein gut konditioniertes Problem durch Störungen der Koeffizientenmatrix zu einem schlecht konditionierten umgewandelt werden kann.

Zuerst eine Schlechte Nachricht: **Ja !!!**
aber
es ist kaum wahrscheinlich.

**Satz 1. (Smallest singular value)**

Sei $\bar{A}$ eine reguläre nxn Matrix und A die Matrix der Zufallsvariablen wo die Matrix $\bar{A}$ das Mitte bildet.
Dann:

$$\Pr\left( \left\| A^{-1} \right\| \geq x \right) \leq 2.35 \frac{\sqrt{n}}{x \cdot \sigma}$$

Beweis:
Es gilt:

$$\Pr\left( \left\| A^{-1} \cdot v \right\| \geq x \right) \leq \sqrt{\frac{2}{\pi}} \cdot \frac{1}{x \cdot \sigma}$$

Sei u so gewählt, dass

$$\left\| A^{-1} \cdot u \right\| = u \quad \left\| A^{-1} \cdot v \right\| \geq |u \cdot v|$$

Sei c bel. grösser 0.

$$\Pr\left( \left\| A^{-1} \right\| \geq x \cdot \sqrt{\frac{c}{n}} \right) \leq \Pr\left( \left\| A^{-1} \right\| \geq x \quad \text{and} \quad |u \cdot v| \geq \sqrt{\frac{c}{n}} \right) = \Pr\left( \left\| A^{-1} \right\| \geq x \right) \cdot \Pr\left( |u \cdot v| \geq \sqrt{\frac{c}{n}} \right)$$

Also

$$\Pr\left( \left\| A^{-1} \right\| \geq x \right) \leq \frac{\Pr\left( \left\| A^{-1} \cdot v \right\| \geq x \sqrt{\frac{c}{n}} \right)}{\Pr\left( |u \cdot v| \geq \sqrt{\frac{c}{n}} \right)} \leq \sqrt{\frac{2}{\pi}} \cdot \frac{\sqrt{n}}{x \cdot \sigma \cdot \sqrt{c} \cdot \Pr\left( |u \cdot v| \geq \sqrt{\frac{c}{n}} \right)}$$

[redacted]

Wähle nun c = 0.57 so ergibt sich

$$\Pr\left( \left|\left|A^{-1}\right|\right| \geq x \right) \leq 2.35 \frac{\sqrt{n}}{x \cdot \sigma}$$

**Satz 2. (Condition number[2] )**

Sei $\bar{A}$ bzw. $A$ wie oben definiert. Ist dann $\sigma^2 \leq 1$ und $||\bar{A}|| \leq \sqrt{n}$ so gilt:

$$\Pr(K(A) \geq x) \leq \frac{9.4n(1 + \sqrt{\log(x) \cdot 2n}}{x \cdot \sigma}$$

Nach Lemma von Davidson and Szarek [3] für alle k grösser gleich 0 gilt:

$$\Pr\left( ||A - \bar{A}|| \geq \sqrt{n} + k \right) \leq e^{\frac{-k^2}{2}}$$

$$\Rightarrow \Pr\left[ ||A - \bar{A}|| \geq \sqrt{n} + \sqrt{2 \cdot \left( \ln(\varepsilon)^{-1} \right)} \right] \leq \varepsilon$$

$$\Rightarrow \Pr\left[ ||A|| \geq 2\sqrt{n} + \sqrt{2 \cdot \left( \ln(\varepsilon)^{-1} \right)} \right] \leq \varepsilon$$

Nach Satz 1. ergibt sich:

$$\Pr\left[ ||A|| \cdot \left|\left|A^{-1}\right|\right| \geq \frac{4.7 \cdot n + 2.35\sqrt{2 \cdot n \cdot \left( \log(\varepsilon)^{-1} \right)}}{\varepsilon \sigma} \right] \leq 2\varepsilon$$

Durch geeignete Substitution erhält man:

$$\Pr(K(A) \geq x) \leq \frac{9.4n\left( 1 + \sqrt{\frac{\log(x)}{2 \cdot n}} \right)}{x \cdot \sigma}$$

**Beispiel 4 (Smoothed Zeitkomplexität von Quicksort)**

**Def 5. Partielle Zufallspermutation**

Eine partielle Zufallspermutation beschreibt das folgende Experiment:

„Gegeben ist eine List $x_1,\ldots, x_n$ von paarweise verschiedenen Zahlen. Wähle nun eine Teilliste aus dieser Liste aus, wobei wir jeden Element – unabhängig von den anderen Elementen – mit einer Wahrscheinlichkeit von p wählen können. Sei $j_1,\ldots,j_i$ mit $j_1<j_2<\ldots<j_i$ die Positionsfolge der ausgewählten Elemente und
D(i) – eine Zufällige Permutation. Dann ist das Ergebnis dieses Zufallsexperiments die Liste $x_1,\ldots, x_n$ wobei wir die Elemente $x_{j1},\ldots,x_{ji}$ gemäß der Permutation D(i) vertauschen.

**Satz 3.**

Die erwartete Laufzeit von Quicksort auf einer partiellen Zufallspermutation von n Elementen bei dem Wahrscheinlichkeitsparameter p ist O(n/p*log(n))

Beweis: [1]

[1] MFCS 2003, C. Banderier, K. Mehlhorn, R. Beier „Smoothed Analysis of Three Combinatorial Problems“

[2] Boston university 2003, A. Sankar, Daniel A. Spielman „Smoothed Analysis of the Condition Numbers and Growth Factors of Matrices“

[3] Elsevier Science 2001, K.R. Davidson, S.J. Szarek „Handbook on the Geometry of Banach spaces“