

# Hilfsmittel „Mathematik“

Approximation von Funktionen

Igor Geiger, Ralf Bergmann

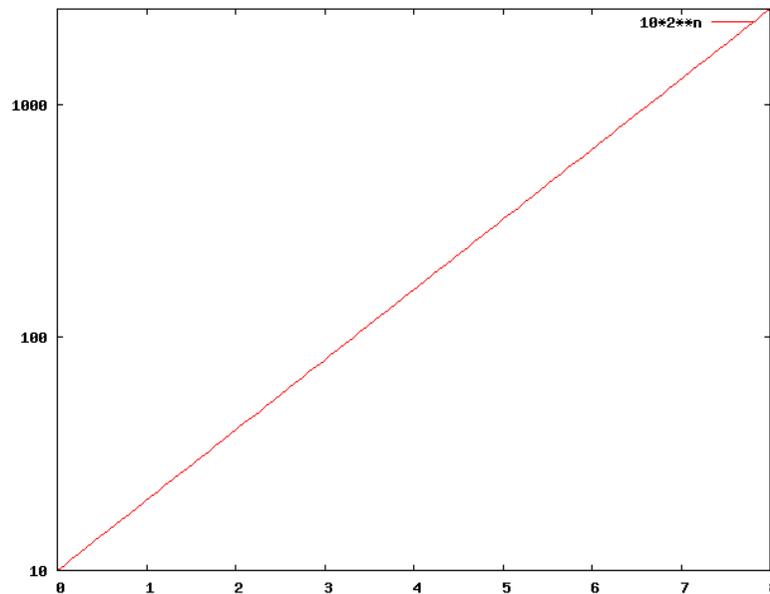
Ib07

# Exponentieller Aufwand

- Bei der Analyse von Algorithmen zur Bestimmung des Aufwandes, für eine bestimmte Anzahl an Eingabegrößen, die von der Anwendung des Algorithmus abhängt, tritt eine Punktmenge im kartesischen Koordinatensystem auf, die exponentiell beschrieben werden kann.
- In realen Messungen ergibt sich aber meistens eine verworfene Punktmenge die einer Exponentialfunktion angenähert werden kann. Jedoch lässt sich anhand dieser Kurve der Aufwand des Algorithmus nicht exakt bestimmen da man dazu einen exakten Anstieg braucht bzw. eine exakte Funktion.
- Dazu würd die entstandene Punktmenge in ein kartesisches Koordinatensystem mit einer logarithmischen y-Achse überführt.

# Exponentieller Aufwand

- Darstellung von  $T(n) = 10 \cdot 2^n$  im neuen System nach Gnuplot.  
Es entsteht folgende Gerade, Schnittpunkt bei  $y=10$  mit der  $y$ -Achse



# Exponentieller Aufwand

- Dies entspricht, vom Aussehen, der Geraden im normalen kartesischen Koordinatensystem:

$$T(n) = \log(T(n)) = \log(2) \cdot n + \log(10)$$

# Exponentieller Aufwand

- Desweiteren hat sich gezeigt das es viel einfacher ist eine Gerade im dargestellten Logarithmischen Koordinatensystem anhand der eingetragenen Punkte zu verlegen als eine Exponentialfunktion im normalen Koordinatensystem zu bestimmen.
- Die mathematische Methode zur Bestimmung einer Geraden im logarithmischen Koordinatensystem, anhand der verstreuten Messpunkte, nennt man **Lineare Regression**. Die Gerade  $f(n) = a \cdot n + b$  soll so bestimmt werden das der maximale Abstand zwischen der Geraden und den Messpunkten möglichst klein wird.

# Lineare Regression

- Mathematische Vorgehensweise:  $\max |T(n) - (a \cdot n + b)| \rightarrow \min$

- Da es für jeden Messpunkt gelten muss, würd die gesuchte Funktion als Summe der Abstände der einzelnen Messpunkte dargestellt:

$$\sum_n |T(n) - (a \cdot n + b)| \rightarrow \min$$

- Von der entstandenen Funktion wird aber zur Bestimmung des Minimums die erste Ableitung benötigt jedoch ist die Betragsfunktion an der Stelle  $x = 0$

nicht stetig differenzierbar. Dazu Quadriert man sie nach der Methode der kleinsten Fehlerquadrate von Gauss, dadurch wird sie Differenzierbar aber die Suche nach dem Minimum wird dadurch nicht beeinflusst.

$$\sum_n (T(n) - (a \cdot n + b))^2 \rightarrow \min$$

# Lineare Regression

- Für das Minimum muss die erste Ableitung der obigen Funktion Null gesetzt werden.

$$f'(x) = 0$$

- Da man aber das Minimum nach a und b haben will muss man auch jeweils nach a und b ableiten.

Ableitung nach a:

$$2 \cdot \sum_{i=1}^r (T(i) - a \cdot i - b)(-i) \rightarrow 0$$

- Ableitung nach b:

$$2 \cdot \sum_{i=1}^r (T(i) - a \cdot i - b)(-1) \rightarrow 0$$

# Lineare Regression

- Durch Vereinfachung erhält man die sogenannte Normalgleichung:

$$a \cdot \sum_{i=1}^r i^2 + b \cdot \sum_{i=1}^r i = \sum_{i=1}^r i + T(i)$$

$$a \cdot \sum_{i=1}^r i + r \cdot b = \sum_{i=1}^r i = \sum_{i=1}^r i + T(i)$$

- Durch weiteres Umstellen nach a bzw. b erhält man die gewünschten Parameter für die Regressionsgerade.

$$a = \frac{r \cdot (\sum_{i=1}^r i \cdot T(i)) - (\sum_{i=1}^r i) \cdot (\sum_{i=1}^r T(i))}{r \cdot \sum_{i=1}^r i^2 - (\sum_{i=1}^r i)^2}$$

$$b = \frac{(\sum_{i=1}^r i^2) \cdot (\sum_{i=1}^r T(i)) - (\sum_{i=1}^r i \cdot T(i)) \cdot \sum_{i=1}^r i}{r \cdot \sum_{i=1}^r i^2 - (\sum_{i=1}^r i)^2}$$

- Somit erhält man die Parameter a und b für die Regressionsgerade

# Lineare Regression

- **Beispielrechnung:**
- folgende Meßwerte:
- $T(1)=3$
- $T(2)=6$
- $T(3)=7$
- $T(4)=9$
- $T(5)=11$
- In die Gleichung für a und b einsetzen nach Regeln

# Lineare Regression

$$\sum_{i=1}^r i \cdot T(i) = 127$$

$$\sum_{i=1}^r i = 15$$

$$\sum_{i=1}^r T(i) = 36$$

$$\sum_{i=1}^r i^2 = 55$$

$$\left(\sum_{i=1}^r i\right)^2 = 225$$

$$\frac{(5 \cdot 127) - (15 \cdot 36)}{(5 \cdot 55) - 15^2} = 1,9$$

$$\frac{(36 \cdot 55) - (127 \cdot 15)}{(5 \cdot 55) - 15^2} = 1,5$$

# Lineare Regression

- Parameter für die Regressionsgerade im kartesischen Koordinatensystem

$$a = 1,9$$

$$b = 1,5$$

$$f(n) = 1,9n + 1,5$$

# Polynomfunktion

- Der Graph wird zur Geraden wenn beide Achsen logarithmisch eingestellt sind

$$T(n) = c \cdot n^k$$

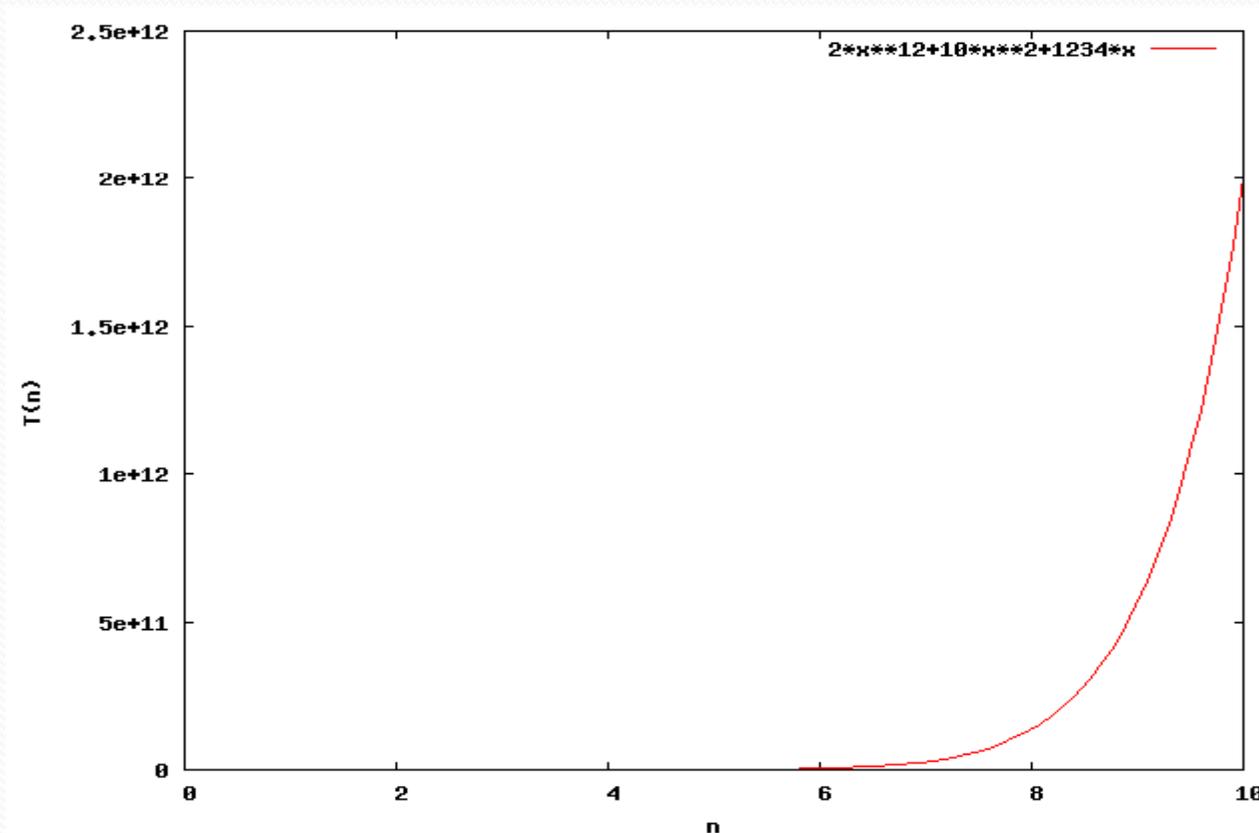
$C \rightarrow$  konst. Faktor  
 $k \rightarrow$  Anstieg

- Eine Polynomfunktion kann beschrieben werden als:

$$T(n) = c_1 \cdot n_1^{k_1} + c_2 \cdot n_2^{k_2} + \dots + c_i \cdot n_i^{k_i}$$

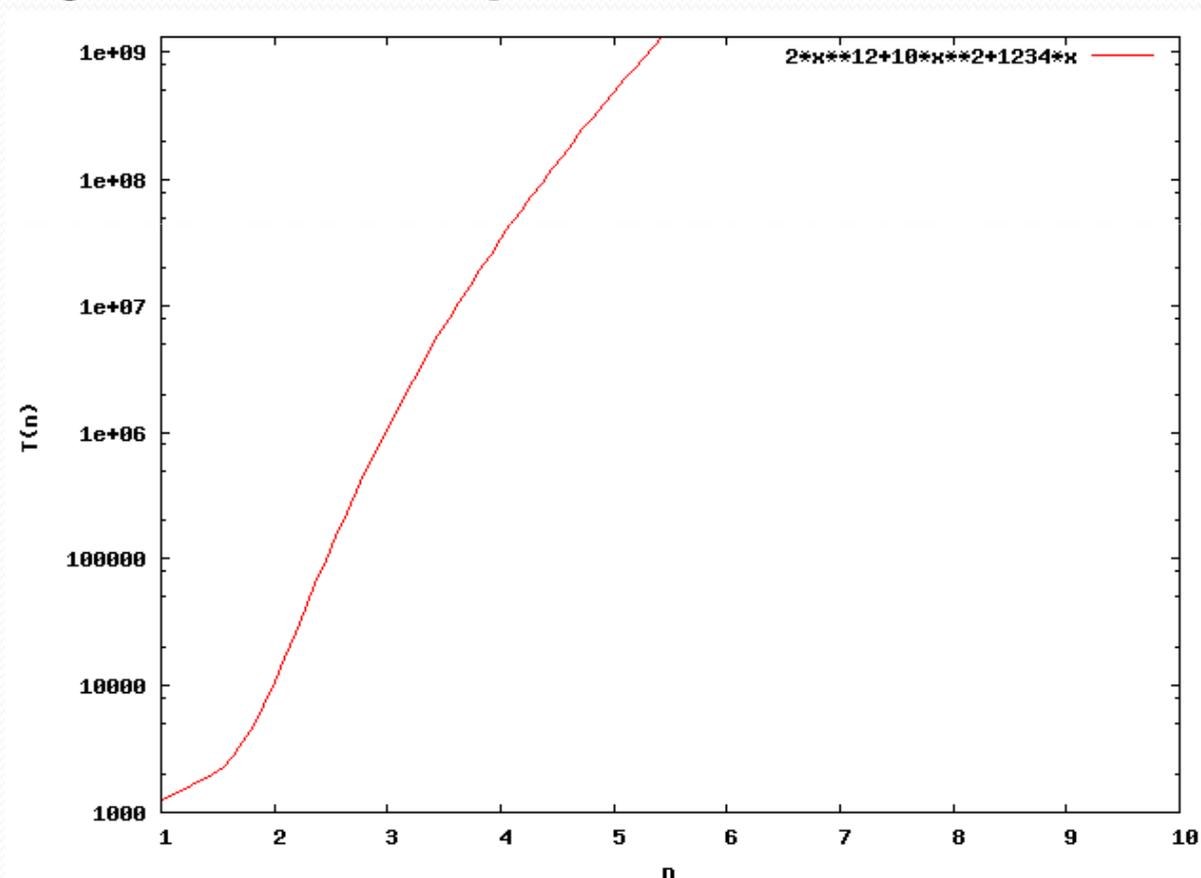
# Polynomfunktion

- Beispiel:  $T(n) = 2 \cdot n^{12} + 10 \cdot n^2 + 1234 \cdot n$



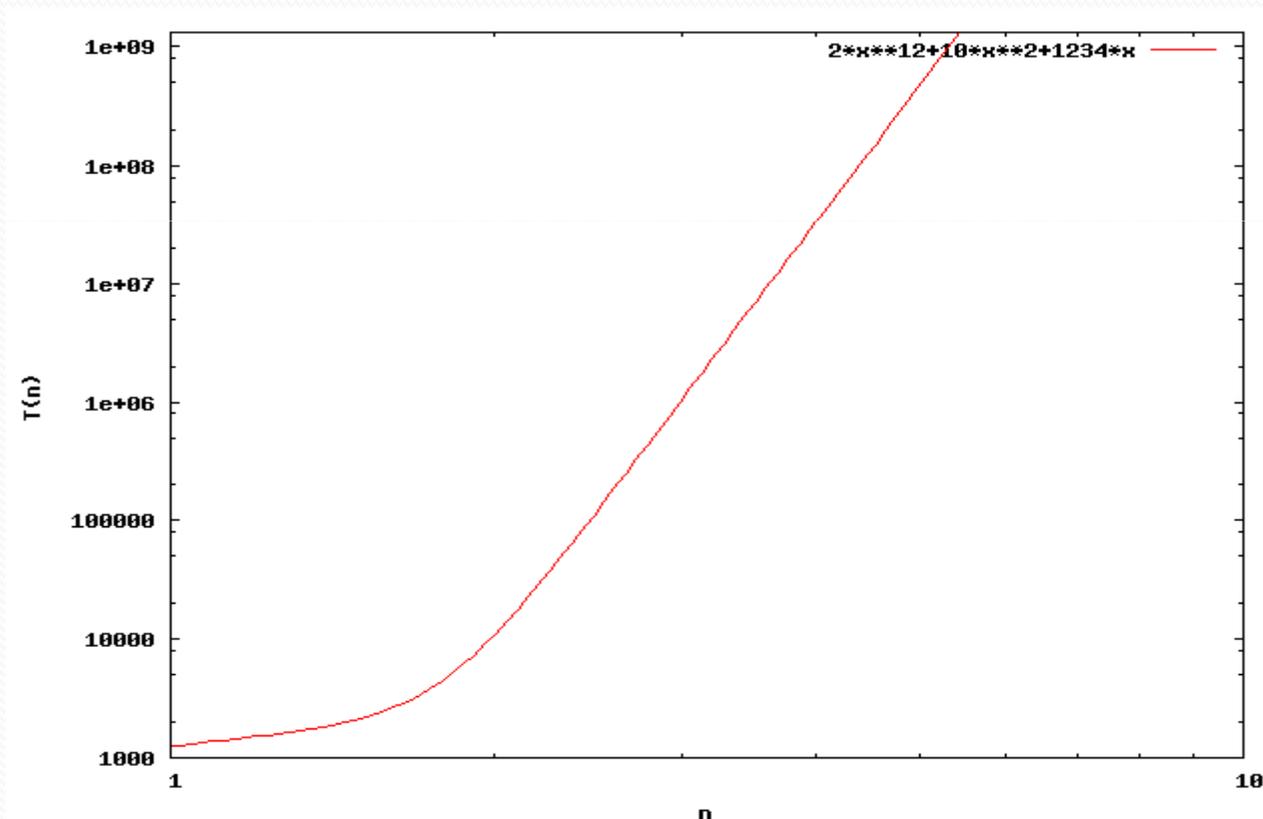
# Polynomfunktion

- mit logarithmischer y-Achse



# Polynomfunktion

- Mit logarithmischer x- und y-Achse



# Polynomfunktion

- Im Allgemeinen kann nicht davon ausgegangen werden das eine Gerade entsteht
- Grund: beim rechnerischen Versuch scheitert es aufgrund der Summe
- Lösung: Methode der kleinsten Fehlerquadrate auf Polynomfunktionen anwenden

$$f(c_0, c_1, \dots, c_m, x) = \sum_{i=0}^m c_i g_i(x)$$

Mit Basisfunktionen  $g_i : x \mapsto x^i$

# Polynomfunktion

$$F(c_1, c_2, \dots, c_m) = \sum_{k=1}^n \left( \sum_{i=0}^m c_i g_i(x_k) - y_k \right)^2 \rightarrow \min$$

- Laut dieser Definition soll dafür gesorgt werden das die Summe minimal wird
- Dazu soll die Ableitung gebildet und 0 gesetzt werden

$$\frac{\partial F}{\partial c_i} = 2 \sum_{k=1}^n \left( \sum_{i=0}^m c_i g_i(x_k) - y_k \right) g_i(x_k) = 0$$

# Polynomfunktion

- Nun multipliziert man die Ableitung aus und bringt sie auf die rechte Seite

$$\sum_{i=1}^m \left( \sum_{k=0}^n g_i(x_k) \cdot g_i(x_k) \right) = \sum_{k=1}^n y_k \cdot g_j(x_k)$$

- Dabei steht  $g_j$  für eine Basisfunktion zwischen 0 und m

# Polynomfunktion

- Nun trägt man die Formel in eine lineares Gleichungssystem ein und erhält  $c$ , wenn man es löst und für  $g_j$  alle Basisfunktionen annimmt
- Dies kann man in eine Matrix darstellen
- Wobei  $g_j$  die Transponierte von  $A$  ist

$$A = \begin{pmatrix} g_0(x_1) & g_1(x_1) & \dots & g_m(x_1) \\ g_0(x_2) & g_1(x_2) & \dots & g_m(x_2) \\ \vdots & \vdots & \dots & \vdots \\ g_0(x_n) & g_1(x_n) & \dots & g_m(x_n) \end{pmatrix}$$

# Polynomfunktion

- Daraus ergibt sich die Form

$$A^T \cdot A \cdot \vec{c} = A^T \cdot \vec{y}$$

- Jetzt löst man das Gleichungssystem nach  $\vec{c}$  und erhält die Faktoren der einzelnen Polynome

$$\vec{c} = (c_0, c_1, \dots, c_m)^T$$

# Polynomfunktion

- Zusammenfassung
  1. Gewünschte Basisfunktion angeben
  2. Messwerte in Funktion einsetzen
  3. Matrix bilden
  4. Gleichungssystem mit Matrizen und Messwerten lösen
  5. Man erhält die Koeffizienten der Polynomfunktion