Mittelmäßigkeit

Henning Thielemann

2008-08-25 und 2008-08-27

Problem

MALDI-TOF-Massenspektrometrie:

- Ziel: Bestandteile einer Stoffprobe nach Massen auftrennen
- Verdampfen der Probe
- Messen der Flugzeit in elektrischem Feld
- Beschleunigung proportional zu Masse-Ladung-Verhältnis

Problem:

- Spektrum „sitzt auf einem Hügel“, genannt Basislinie
- ich weiß nicht warum
- muss jedenfalls weg

Mögliche Lösung

Idee aus morphologischer Bildverarbeitung: Closing-Operation

- gleitendes Minimum über Spektrum (Erosion)
- gleitendes Maximum über Ergebnis (Dilatation)
- gleitendes Minimum effizient berechnen mit Prioritäten-Warteschlange (beispielsweise Heap)

Alternative Lösung I

Irgendwas mit Wavelets

- Idee: Basislinie entspricht geglättetem Spektrum
- schwierige Randbehandlung
- auch sonst irgendwie unbefriedigend

Alternative Lösung II

Glättung mit alternativen gleitenden Mitteln (Nicht Gleitmitteln!)

- Eigenschaft gleitendes arithmetisches Mittel:
  Fläche ober- und unterhalb gleich
- d.h. als Basislinie zu hoch
- abgewandelt trotzdem brauchbar

Gleitendes Mittel: Effiziente Implementierung

$$A(x_0, \ldots, x_{n-1}) = \frac{x_0 + x_1 + \cdots + x_{n-1}}{n}$$
$$A(x_1, \ldots, x_n) = \frac{x_1 + \cdots + x_{n-1} + x_n}{n}$$

Beobachtung:

$$A(x_1, \ldots, x_n) = A(x_0, \ldots, x_{n-1}) + \frac{x_n - x_0}{n}$$

Gleitendes Mittel: Probleme

Problem:
Anhäufung von Rundungsfehlern

Lösungen:

- verlustbehafteter Integrator

$$A(x_1, \ldots, x_n) = (1 - \varepsilon) \cdot \left(A(x_0, \ldots, x_{n-1}) - \frac{x_0}{n} \right) + \frac{x_n}{n}$$

- dyadische Summation
  strukturell kompliziert

Gewichtetes Mittel

Glättung gefälliger,
wenn Werte an Rändern des Filterfensters schwächer gewichtet

Zentraler Grenzwertsatz:
Mehrfache Glättung entspricht ungefähr Faltung mit
GAUSS-Glocke

Potenzmittel (Nicht Viagra!)

Arithmetisches Mittel	$Ax = \frac{x_1 + \cdots + x_n}{n}$
Potenzmittel	$A_p x = \sqrt[p]{\frac{x_1^p + \cdots + x_n^p}{n}}$
Maximum	$\lim_{p \to \infty} A_p x = \max\{x_1, \ldots, x_n\}$
Minimum	$\lim_{p \to -\infty} A_p x = \min\{x_1, \ldots, x_n\}$
Geometrisches Mittel	$\lim_{p \to 0} A_p x = \sqrt[n]{x_1 \cdot \cdots \cdot x_n}$

$$(x = (x_1, \ldots, x_n))$$

Potenzmittel: Eigenschaften

Monotonie:

$$p < q \Rightarrow A_p x \leq A_q x$$

strenge Monotonie:

$$p \neq q \wedge A_p x = A_q x \quad \Rightarrow \quad x_1 = \cdots = x_n$$

Assoziativität:

$$\{x_1, \ldots, x_m\} \subset \mathbb{R}^n \quad \Rightarrow \quad A_p(A_p x_1, \ldots, A_p x_m) = A_p(x_1 +\!\!+ \ldots +\!\!+ x_m)$$

$$\text{mit } (y_1, \ldots, y_n) +\!\!+ (z_1, \ldots, z_n) = (y_1, \ldots, y_n, z_1, \ldots, z_n)$$

Gleitendes Potenzmittel

Erkenntnis:
Potenzmittel ist wie arithmetisches Mittel
bei umbeschrifteter y-Achse.

Implementierung gleitendes Potenzmittel:

- potenziere alle Werte des Spektrums mit p
- glätte
- ziehe p. Wurzel aus geglätteten Werten

LEHMER-Mittel

Mittel nach LEHMER:

$$L_p x = \frac{x_1^p + \cdots + x_n^p}{x_1^{p-1} + \cdots + x_n^{p-1}}$$

Effiziente Implementierung für gleitendes LEHMER-Mittel:

- Potenziere Spektrenwerte mit p, glätte $\rightarrow$ Signal a
- Potenziere Spektrenwerte mit $p-1$, glätte $\rightarrow$ Signal b
- dividiere a werteweise durch b

Zwischenergebnis

- Potenzmittel und LEHMER-Mittel
  bieten Alternativen zu Minimum und Maximum
- relativ einfache effiziente Implementierung
- erlauben feinere Justierung durch Exponent als Parameter
- Qual der Wahl: Welcher Exponent passt am besten?

Fragen über Fragen

- Wie entwirft man ein Mittel?
- …eines, bei dem ZeTeM überdurchschnittlich abschneidet?
- … eines, das auch als gleitendes Mittel
  effizient berechnet werden kann?
- Gibt es eine Form,
  in der alle Mittel dargestellt werden können?
- Und überhaupt: Was ist eigentlich ein Mittel?

Was ist ein Mittel?

Eine Abbildung $M : \mathbb{R}^n_{>0} \to \mathbb{R}_{>0}$ soll (gewichtetes) <u>Mittel</u> heißen, wenn alle der folgenden Bedingungen erfüllt sind:

Fixpunkt		$M(1, 1, \ldots, 1) = 1$
Homogenität	$\forall \lambda > 0\ \forall x$	$M(\lambda \cdot x) = \lambda \cdot Mx$
Monotonie	$\forall x\ \forall y$	$(\forall i\ x_i \leq y_i) \Rightarrow Mx \leq My$

Daraus folgen bereits:

Beschränktheit	$\forall x$	$Mx \in [\min x, \max x]$
Stetigkeit		$\lim_{x \to y} Mx = My$

Was ist ein Mittel?

Eine Abbildung $M : \mathbb{R}^n_{>0} \to \mathbb{R}_{>0}$ soll ungewichtetes Mittel heißen, wenn sie (gewichtetes) Mittel ist, und bei Vertauschung der Eingaben unverändert bleibt.

$$\text{Symmetrie} \quad \forall x \ \forall \text{ Permutation } \pi \quad Mx = M(\pi x)$$

Beispiele für ungewichtete Mittel in diesem Sinne:

- Potenzmittel
- Lehmer-Mittel
- Rangordnungsfilter (Median, Maximum, Minimum)
- logarithmisches Mittel
- arithmetisch-geometrisches Mittel (Gauß-Suppe)

Bald stehen wir am Mittelmeer und haben keine Mittel mehr

Neue Mittel kann man aus bekannten Mitteln $A, M_1, \ldots, M_m$ zusammenbauen:

gemittelte Mittel	$x \mapsto A(M_1 x, \ldots, M_m x)$
verallgemeinertes Potenzmittel	$x \mapsto \sqrt[p]{A(x_1^p, \ldots, x_n^p)}$

Arithmetisch-geometrisches Mittel

Idee: wiederholtes Mitteln von Mitteln
Betrachte Folgen a und b mit

$$a_{i+1} = \frac{a_i + b_i}{2}$$
$$b_{i+1} = \sqrt{a_i \cdot b_i}$$

Es gilt $\lim_{i\to\infty} a_i = \lim_{i\to\infty} b_i$
Grenzwert $\lim_{i\to\infty} a_i$ heißt
arithmetisch-geometrisches Mittel von a_0 und b_0,
kurz $\mathrm{AGM}(a_0, b_0)$.

Arithmetisch-Geometrisches Mittel: Eigenschaften

arithmetisch-geometrisches Mittel erlaubt effiziente Berechnung von

- elliptischen Integralen (Gauss 1818)
- π (Salamin 1976)

$$\pi = \frac{4 \cdot \mathrm{AGM}\left(1, \frac{1}{\sqrt{2}}\right)^2}{1 - \sum_{j=1}^{\infty} 2^{j+1} \cdot c_j^2}$$

mit $c_j^2 = a_j^2 - b_j^2$

äquivalent: $c_j = \frac{1}{2} \cdot (a_{j-1} - b_{j-1})$

rekursiv: $c_j^2 = 4 \cdot a_{j+1} \cdot c_{j+1}$

Modifiziertes Geometrisch-Harmonisches Mittel

nochmal π:
einfacher beschrieben, aber aufwendiger berechnet

$$\begin{aligned} a_0 &= 2 \\ b_0 &= 4 \\ a_{n+1} &= \sqrt{a_n \cdot b_n} \\ b_{n+1} &= \frac{2}{\frac{1}{a_{n+1}} + \frac{1}{b_n}} \end{aligned}$$

Besonderheit: harmonisches Mittel hängt
vom aktuell berechneten geometrischen Mittel ab
Ergebnis:

$$\pi = \lim_{j \to \infty} a_j$$

Gauss-Suppe

Verallgemeinerung von arithmetisch-geometrischem Mittel
zu iterierten Mitteln oder „Gauss-Suppe“ (Ulrich Krause, Bremen)

Mittelwertsatz der Differentialrechnung

Für differenzierbare Funktion f gilt

$$\forall a, b \; a < b \quad \Rightarrow \quad \exists \xi \in [a, b] \quad f'(\xi) = \frac{f(a) - f(b)}{a - b}$$

Umstellen nach ξ erlaubt Mitteln: $\xi = f'^{-1}\left(\frac{f(a)-f(b)}{a-b}\right)$

Logarithmisches Mittel	$f = \ln$	$\xi = \frac{x - y}{\ln x - \ln y}$
Stolarsky-Mittel	$f(x) = x^p$	$\xi = \sqrt[p-1]{\frac{x^p - y^p}{p \cdot (x - y)}}$

Mittelwertsatz der Differentialrechnung

Verallgemeinerung: Höhere Ableitung und dividierte Differenzen

$$M_f(x_0, \ldots, x_n) = {f^{(n)}}^{-1}(n! \cdot f[x_0, \ldots, x_n])$$

Ungewiss, für welche f (außer Logarithmus und Potenzfunktionen) Mittelaxiome für M erfüllt

Integralrechnung

Logarithmisches Mittel:
Fläche unter Exponentialkurve

$$L(x, y) = \int_0^1 x^{1-t} \cdot y^t \, \mathrm{d}\, t = \frac{x - y}{\ln x - \ln y}$$

Verallgemeinerung: Integral über Simplex
Darstellbar mit dividierten Differenzen:

$$L_{\mathrm{I}}(x_0, \dots, x_n) = n! \cdot \exp[\ln x_0, \dots, \ln x_n]$$

Integralverallgemeinerung $\neq$ Differentialverallgemeinerung

Optimierung

arithmetisches Mittel $$Ax = \operatorname*{argmin}_{\xi} \sum_{i=1}^{n} (\xi - x_i)^2$$

Potenzmittel $$A_p x = \sqrt[p]{\operatorname*{argmin}_{\xi} \sum_{i=1}^{n} \left(\xi - x_i^p\right)^2}$$

Median $$Mx = \operatorname*{argmin}_{\xi} \sum_{i=1}^{n} |\xi - x_i|$$

Optimierungsformulierung erlaubt Verallgemeinerung wie:
Median von positiv definiten Matrizen,
der selbst immer positiv definit ist (WELK 2003)

Zusammenfassung

- Vielzahl an Konstruktionsmethoden
- ... erlauben Vielzahl an Mitteln
- Weitgehend offen, ob eine Darstellung andere enthält
- Viele Zusammenhänge von Mitteln
  zu anderen Gebieten der Mathematik