Utrecht, 9 december 2014

Numerieke Wiskunde

Gerard Sleijpen

Universiteit Utrecht
Department of Mathematics

http://www.staff.science.uu.nl/∼sleij101/

Evaluatiefouten

Voorbeeld. Beschouw $f(x) \equiv -x(x-1)(x-2)\cdot\ldots\cdot(x-8)$.
Hoe groot is de evaluatie fout in $f(7)$?
Schrijf uit als machten van x:

$$f(x) = a_1x^9 + a_2x^8 + a_3x^7 + \ldots + a_9x + a_{10},$$

waarbij $a = [-1, 36, -546, 4536, -22449, 67284, -118124, 109584, -40320, 0]$

We zoomen in rond $x = 7$

Evalueren van f in x levert $f(x) + \xi g(x)$
met $g(x) = 18|a_1||x|^9 + 17|a_2||x|^8 + \ldots + 3|a_9||x| + |a_{10}|$.
In $x = 7$ is $g(x) = 2.4\,10^{10}$.

Approximatiefouten

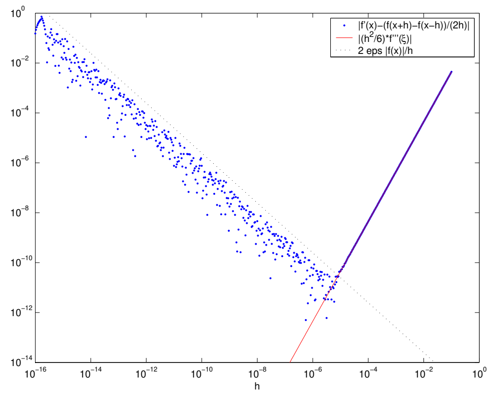
Approximatiefouten (ook wel benaderingsfouten genoemd) hebben een structuur. Die kan je uitbuiten.

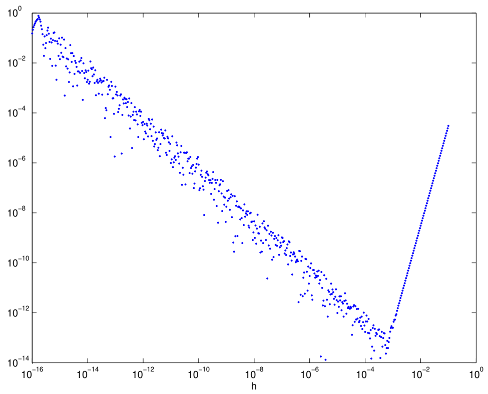
Voorbeeld. Als f voldoende glad is, dan zijn er c_i zodat

$$f'(x) - D_h(f)(x) = c_1h + c_2h^2 + c_3h^3 + \ldots \quad \text{voor} \quad h \to 0.$$

Hierbij is

$$D_h(f)(x) \equiv \frac{f(x+h) - f(x)}{h} \quad \text{voor} \quad h > 0$$





Als f voldoende glad

$$f'(x) - D_h(f)(x) = c_1h + c_2h^2 + c_3h^3 + \ldots$$

Schrijf $I \equiv f'(x)$ en $D(h) = D_h(f)(x)$.

Als h 1) voldoende klein (zodat $c_2h^2 \ll c_1h$) en $\qquad(*)$
2) niet te klein (rekenfouten spelen geen rol)

dan $I - D(h) \approx c_1h$ en $\boxed{I - D(\tfrac{1}{2}h) \approx D(\tfrac{1}{2}h) - D(h)}$

Dus $I \approx D(\tfrac{1}{2}h) + (D(\tfrac{1}{2}h) - D(h))$.

Is $\boxed{T(h) \equiv 2D(\tfrac{1}{2}h) - D(h)}$ beter dan $D(\tfrac{1}{2}h)$?

$$\begin{aligned} I - T(h) &= 2(I - D(\tfrac{1}{2}h)) - (I - D(h)) \\ &= 2(c_1\tfrac{1}{2}h + c_2\tfrac{1}{4}h^2 + \ldots) - (c_1h + c_2h^2 + \ldots) \\ &= -\tfrac{1}{2}c_2h^2 - \tfrac{3}{4}c_3h^3 - \ldots = c_2'h^2 + c_3'h^3 + \ldots \end{aligned}$$

Als f voldoende glad

$$f'(x) - D_h(f)(x) = c_1h + c_2h^2 + c_3h^3 + \ldots$$

Schrijf $I \equiv f'(x)$ en $D(h) = D_h(f)(x)$.

Als h 1) voldoende klein (zodat $c_2h^2 \ll c_1h$) en $\qquad(*)$
2) niet te klein (rekenfouten spelen geen rol)

dan $I - D(h) \approx c_1h$ en $\boxed{I - D(\tfrac{1}{2}h) \approx D(\tfrac{1}{2}h) - D(h)}$

$$\boxed{V(h) \equiv \frac{D(\tfrac{1}{2}h) - D(h)}{D(\tfrac{1}{4}h) - D(\tfrac{1}{2}h)} \approx 2}$$

$$V(h) = \frac{(I - D(h)) - (I - D(\tfrac{1}{2}h))}{(I - D(\tfrac{1}{2}h)) - (I - D(\tfrac{1}{4}h))} \approx \frac{c_1h - c_1\tfrac{1}{2}h}{c_1\tfrac{1}{2}h - c_1\tfrac{1}{4}h} = 2\frac{1 - \tfrac{1}{2}}{1 - \tfrac{1}{2}} = 2.$$

Als $V(h) \approx 2$ (en $V(\tfrac{1}{2}h) \approx 2, \ldots$)
dan hebben we er **vertrouwen** in dat $(*)$ geldt.

Stel voor $p, q \in \mathbb{N}, \quad p < q,$

$$I - D(h) = ch^p + dh^q + \ldots$$

h voldoende klein: $I - D(h) \approx ch^p$, $I - D(\tfrac{1}{2}h) \approx (\tfrac{1}{2})^p ch^p$

$$\boxed{I - D(\tfrac{1}{2}h) \approx \frac{1}{2^p - 1}[D(\tfrac{1}{2}h) - D(h)]}$$

Dus $I \approx D(\tfrac{1}{2}h) + \frac{1}{2^p-1}[D(\tfrac{1}{2}h) - D(h)] = \frac{1}{2^p-1}[2^pD(\tfrac{1}{2}h) - D(h)]$.

Is $\boxed{T(h) \equiv \frac{1}{2^p - 1}[2^pD(\tfrac{1}{2}h) - D(h)]}$ beter dan $D(\tfrac{1}{2}h)$?

$$\begin{aligned} I - T(h) &= \tfrac{1}{2^p-1}[2^p(I - D(\tfrac{1}{2}h)) - (I - D(h))] \\ &= \tfrac{1}{2^p-1}\left[\left(\tfrac{2^p}{2^q} - 1\right)h^q + \ldots\right] \\ &= d'h^q + \ldots \end{aligned}$$

Romberg schema

$$\begin{array}{lllllll}
h-D(h) & & & & & & \\
 & \diagdown & T_1(h) & & & & \\
\tfrac{1}{2}h-D(\tfrac{1}{2}h) & \diagup\,\diagdown & & \diagdown & T_2(h) & \diagdown & \\
 & & T_1(\tfrac{1}{2}h) & \diagup\,\diagdown & & & T_3(h) \\
\tfrac{1}{4}h-D(\tfrac{1}{4}h) & \diagup\,\diagdown & & \diagup & T_2(\tfrac{1}{2}h) & \diagup & \\
 & & T_1(\tfrac{1}{4}h) & & & & \\
\tfrac{1}{8}h-D(\tfrac{1}{8}h) & \diagup & & & & &
\end{array}$$

Als in j-de kolom fout evenredig h^p (bereken de $V_j(h)$) dan

- schat fout $I - T_j(\frac{1}{2}h)$ volgens $\boxed{\dfrac{1}{2^p-1}\left[T_j(\tfrac{1}{2}h) - T_j(h)\right]}$
- corrigeer volgens $\boxed{T_{j+1}(h) = \dfrac{1}{2^p-1}\left[2^p T_j(\tfrac{1}{2}h) - T_j(h)\right]}$

Romberg schema

Romberg schema is een interpolatie schema.

$$h_0 > h_1 > h_2 > h_3 > 0$$

Ter herinnering. Als p_0 $D(h)$ interpoleert op (h_0, h_1, h_2)

p_1 $D(h)$ interpoleert op (h_1, h_2, h_3)

dan interpoleert p gedefinieerd door

$$p(x) \equiv \frac{x-h_3}{h_0-h_3}p_0(x) + \frac{x-h_0}{h_3-h_0}p_1(x)$$

D op (h_0, h_1, h_2, h_3) en

$$\begin{aligned} p(0) &= \frac{-h_3}{h_0-h_3}p_0(0) + \frac{-h_0}{h_3-h_0}p_1(0) \\ &= \frac{1}{h_0-h_3}[h_0 p_1(0) - h_3 p_0(0)] \end{aligned}$$

Fout: $I - p(0) = (-h_0)(-h_1)(-h_2)(-h_3)\frac{1}{4!}D^{(4)}(\xi)$

Romberg schema

Als $\quad I - D(h) = c_1 h \ + c_2 h^2 + c_3 h^3 + \ldots$

$$\begin{array}{lllllll}
h_0-D(h_0) & & & & & & \\
 & \diagdown & T_{1,0} & & & & \\
h_1-D(h_1) & \diagup\,\diagdown & & \diagdown & T_{2,0} & \diagdown & \\
 & & T_{1,1} & \diagup\,\diagdown & & & T_{3,0} \\
h_2-D(h_2) & \diagup\,\diagdown & & \diagup & T_{2,1} & \diagup & \\
 & & T_{1,2} & & & & \\
h_3-D(h_3) & \diagup & & & & &
\end{array}$$

- schat fout $I - T_{j,1}$ volgens $\boxed{\dfrac{h_{j+1}}{h_0-h_{j+1}}\left[T_{j,1} - T_{j,0}\right]}$
- corrigeer volgens $\boxed{T_{j+1,0} = \dfrac{1}{h_0-h_{j+1}}\left[h_0 T_{j,1} - h_{j+1} T_{j,0}\right]}$

$$I - T_{j,0} = (-h_0)(-h_1)\ldots(-h_j)c_{j+1} + \mathcal{O}(h^{j+2})$$

Romberg schema

Als $\quad I - D(h) = c_2 h^2 + c_4 h^4 + c_6 h^6 + \ldots$

$$\begin{array}{lllllll}
h_0-D(h_0) & & & & & & \\
 & \diagdown & T_{1,0} & & & & \\
h_1-D(h_1) & \diagup\,\diagdown & & \diagdown & T_{2,0} & \diagdown & \\
 & & T_{1,1} & \diagup\,\diagdown & & & T_{3,0} \\
h_2-D(h_2) & \diagup\,\diagdown & & \diagup & T_{2,1} & \diagup & \\
 & & T_{1,2} & & & & \\
h_3-D(h_3) & \diagup & & & & &
\end{array}$$

- schat fout $I - T_{j,1}$ volgens $\boxed{\dfrac{h_{j+1}^2}{h_0^2-h_{j+1}^2}\left[T_{j,1} - T_{j,0}\right]}$
- corrigeer volgens $\boxed{T_{j+1,0} = \dfrac{1}{h_0^2-h_{j+1}^2}\left[h_0^2 T_{j,1} - h_{j+1}^2 T_{j,0}\right]}$

$$I - T_{j,0} = (-h_0^2)(-h_1^2)\ldots(-h_j^2)c_{2j+2} + \mathcal{O}(h^{2j+2})$$

Romberg schema

Als $\quad I - D(h) = c_1 h^2 + c_2 h^4 + c_3 h^6 + \ldots$

dan interpoleer

de functie $\widetilde{D}(h) \equiv c_1 h + c_2 h^2 + c_3 h^3 + \ldots$

op $h_0^2, h_1^2, \ldots$

Dan $\widetilde{D}(h^2) = D(h)$ en op $\widetilde{D}(h)$ is het schema gebaseerd op $c_1 h + c_2 h^2 + \ldots$ van toepassing.

Romberg schema

Als $\quad I - D(h) = c_1 h \ + c_2 h^2 + c_3 h^3 + \ldots$

$$\begin{array}{lllllll}
h_0-D(h_0) & & & & & & \\
 & \diagdown & T_{1,0} & & & & \\
h_1-D(h_1) & \diagup\,\diagdown & & \diagdown & T_{2,0} & \diagdown & \\
 & & T_{1,1} & \diagup\,\diagdown & & & T_{3,0} \\
h_2-D(h_2) & \diagup\,\diagdown & & \diagup & T_{2,1} & \diagup & \\
 & & T_{1,2} & & & & \\
h_3-D(h_3) & \diagup & & & & &
\end{array}$$

- Hoe fout $I - T_{3,0}$ te schatten?

– Gebruik $I - T_{j,0} = (-h_0)(-h_1)\ldots(-h_j)c_{j+1} + \mathcal{O}(h^{j+2})$?
c_j is (gewoonlijk) onbekend. Grootte $\mathcal{O}(h^{j+2})$ onbekend.

– Voeg een rij vanuit een h_4 toe? Kan, maar is extra werk (en wellicht 'zit $T_{3,0}$ al in de afrondfouten')

Gewoonlijk $|I - T_{3,0}| \ll |I - T_{2,0}|$:

schat $|I - T_{3,0}|$ door $|I - T_{3,0}| \lesssim |T_{2,0} - T_{3,0}|$.

Romberg schema

Waarom andere rij h dan $h, \frac{1}{2}h, \frac{1}{4}h, \ldots$?

Voorbeeld. $h, \alpha h, \alpha^2 h, \alpha^3 h, \ldots$ met $\alpha \in (\frac{1}{2}, 1)$.

Voordeel: Voor dezelfde minimale $\alpha^q h$ ($> \frac{1}{2}^p h$ voor $p < q$) verder in het Romberg schema $T_{q,0}$. Gewoonlijk is $T_{q,0}$ (met $h_j = \alpha^j h$) nauwkeuriger dan $T_{p,0}$ (met $h_j = \frac{1}{2}^p h$). Het even nauwkeurig resultaat $T_{p,\ell}$ vereist een kleinere $h_{j+\ell} = (\frac{1}{2})^{j+\ell}h$ en kan aanleiding geven tot grotere effecten van afrondfouten.

Voorbeeld. Bulirch rij $h, \frac{1}{2}h, \frac{1}{3}h, \frac{1}{4}h, \frac{1}{6}h, \ldots$

Voordeel: grotere efficiëntie bij numeriek integreren

Samenvatting

- Evaluatiefouten hebben **geen** duidelijke structuur.
  Een majorant kan relatief eenvoudig afgeleid worden.

- Approximatiefouten hebben een mooie structuur en die kan uitgebuit worden:
  - fouten kunnen 'automatisch' geschat worden
  - er kan 'automatisch' vastgesteld worden of deze schatting betrouwbaar is
  - benaderingen kunnen m.b.v. de geschatte fout gecorrigeerd worden
  - deze procedure kan recursief worden toegepast.