Betrouwbare functie-evaluatie met behulp van kettingbreukontwikkelingen

Johan Vervloet*

2003–2004

1 Doelstelling van het project

Het voorgestelde onderzoeksproject wil bijdragen tot de ontwikkeling van een softwarebibliotheek die speciale functies betrouwbaar evalueert. Deze betrouwbare evaluatie (elk afgeleverd cijfer moet beduidend zijn) beogen we voor een zo groot mogelijk domein van argumenten en parameters.

Aangezien het aanleggen van een dergelijke functiebibliotheek een omvangrijke onderneming is, beperkt onze bijdrage zich tot de speciale functies die zich goed laten benaderen door kettingbreukontwikkelingen.

2 Van functie naar kettingbreuk

Veronderstel dat we een functie f willen evalueren in z, en we kunnen $f(z)$ herschrijven in functie van een kettingbreukontwikkeling $F(z)$:

$$F(z) = \frac{a_1(z)}{b_1(z)} \,+\, \frac{a_2(z)}{b_2(z)} \,+\, \frac{a_3(z)}{b_3(z)} + \cdots \tag{1}$$

De $a_n(z)$ zijn de partiële tellers van de kettingbreukontwikkeling, en de $b_n(z)$ de partiële noemers.

In het algemeen convergeert $F(z)$ slechts snel als z gelegen is in een beperkt deel van het domein. Vaak wordt de uitdrukking $f(z)$ herschreven in functie van $F(z')$, waarbij z' binnen dat interessante deel van het domein ligt. Men spreekt van argumentreductie.

Het bepalen van een geschikte z', en het berekenen van $f(z)$ in functie van $F(z')$ gebeurt in floating-pointaritmetiek. Onvermijdelijk duiken hierbij afrondingsfouten op. Aangezien de argumentreductie en de relatie tussen $f(z)$ en $F(z')$ voor elke functie verschilt, gaat de implementatie van elke functie $f(z)$ gepaard met een grondige analyse van de afrondingsfouten in kwestie.

3 Een nauwkeurige kettingbreukevaluatie

Als we kunnen terugvallen op een bovenvermelde foutenanalyse, is de berekening van de functiewaarde $f(z)$ nu herleid tot de evaluatie van de kettingbreuk

*Aspirant F.W.O. Vlaanderen

$F(z')$. In praktijk zal de exacte waarde van $F(z')$, zoals gegeven in (1) benaderd worden door een gemodificeerde convergent $F_N(z';w)$, waarbij $F_N(z;w)$ als volgt gedefinieerd is:

$$F_N(z;w) = \frac{a_1(z)}{b_1(z)} + \frac{a_2(z)}{b_2(z)} + \cdots \frac{a_N(z)}{b_N(z)+w} \tag{2}$$

Uitdrukking (2) ontstaat door in uitdrukking (1) de N'de 'staart' te vervangen door een 'modificatie' w. Het spreekt voor zich dat de kwaliteit van de gemodificeerde convergent (2) als benadering voor uitdrukking (1) toeneemt naarmate w deze staart beter benadert.

Als we de kettingbreuk (1) benaderen door de convergent (2), en we gebruiken hiervoor floating-pointaritmetiek, dan hangt de nauwkeurigheid van het resultaat af van twee factoren:

1. De fout die ontstaat door de afrondingen op de floating-pointbewerkingen. (afrondingsfout)

2. De fout die optreedt omdat puur wiskundig $F_N(z;w) \neq F(z)$. (truncatiefout)

De afrondingsfout is enerzijds afhankelijk van de manier waarop $F_N(z;w)$ in floating-pointaritmetiek geëvalueerd wordt, en anderzijds van de precisie van de evaluatie. Als er gekozen wordt voor achterwaartse recursie (waarbij de uitdrukking van rechts naar links wordt geëvalueerd), kan gemakkelijk een werkprecisie bepaald worden zodat de afrondingsfout onder controle blijft[2, 4].

De truncatiefout hangt af van de keuze van N en w. Voor een nauwkeurig resultaat, zal een grote N of een nauwkeurige w nodig zijn. In praktijk is een kleine waarde voor N interessanter dan een grote. Op die manier zijn voor de evaluatie immers minder berekeningen nodig, wat een aanzienlijke tijdswinst oplevert.

Aangezien we opteren voor achterwaartse recursie, is het belangrijk dat de keuze van N en w vaststaat alvorens de eigenlijke evaluatie begint. Vandaar dat we gebruik maken van een zogenaamde 'a priori' foutenbegrenzing.

4 A priori begrenzing van de truncatiefout

De truncatiefout die optreedt bij kettingbreukevaluatie kan op verschillende manieren a priori begrensd worden[1]. In dit project werd gekozen voor het Interval Sequence Theorem (IST), omdat het de truncatiefout scherp begrenst, en omdat het overweg kan met modificaties $w \neq 0$.

Theorema (IST). *Stel dat*

$$F(z) = \frac{a_1(z)}{1} + \frac{a_2(z)}{1} + \frac{a_3(z)}{1} + \cdots \tag{3}$$

Als er een rij intervallen

$$(V_n)_n = ([\ell_n, r_n])_n \tag{4}$$

bestaat zodat voor alle $n > 0$

1. $0 < \ell_n < r_n < \infty$

2. $(1+r_n)\ell_{n-1} \leq a_n(z) \leq (1+\ell_n)r_{n-1}$

dan geldt

$$\frac{|F(z) - F_N(z;w)|}{F(z)} \leq \frac{r_N - \ell_N}{1+\ell_N} \prod_{k=1}^{N-1} \frac{r_k}{1+r_k} \tag{5}$$

voor $w \in V_N$*.*

Dit theorema, afgeleid uit het ‘Oval Sequence Theorem (OST)’[3], geeft een bovengrens voor de relatieve truncatiefout van de kettingbreukontwikkeling. Om het te kunnen gebruiken, moet aan twee voorwaarden voldaan zijn:

1. Voor de te benaderen functie is een kettingbreukontwikkeling nodig met alle partiële noemers gelijk aan 1.
2. We hebben intervallen V_n nodig die aan de criteria van het IST voldoen.

De eerste voorwaarde vormt niet zo'n groot probleem. Een willekeurige kettingbreukontwikkeling $F(z)$, kan door toepassing van een equivalentietransformatie tot de gewenste vorm herleid worden. Moeilijker is het vinden van geschikte V_n'etjes.

In [1] is beschreven hoe geschikte intervallen V_n bepaald kunnen worden voor kettingbreuken $F(z)$ van de vorm (3) die tot een van deze klasses behoren:

1. $(a_{2n}(z))_n \nearrow a_\ell(z), (a_{2n+1}(z))_n \searrow a_r(z)$ en $0 < a_\ell(z) \leq a_r(z) < \infty$. De functies arcsinh, arctanh en log kunnen geformuleerd worden met behulp van een kettingbreuk uit deze klasse.
2. $(a_n(z))_n \searrow a(z) > 0$. Functies die in functie van dergelijke ontwikkelingen geformuleerd kunnen worden, zijn o.m. arccos, arctan en arccot.
3. $(a_n(z))_n \searrow 0$. In deze klasse vinden we tanh en exp terug.
4. $(a_n(z))_n \nearrow 0$. De functie tan kan in functie van een dergelijke kettingbreuk geschreven worden. **Moet ik toch eens nakijken, want het IST vraagt volgens mij dat de a_n'etjes positief zijn.**
5. $(a_n(z))_n \nearrow +\infty$. In deze klasse ontmoeten we de error functie en de exponentiele integraal.

De keuze van de V_n'etjes heeft een grote invloed op de bovengrens bepaald door het IST. Het interval V_N bepaalt immers de keuze van de modificatie w, en zoals gezegd beïnvloedt deze de grootte van de truncatiefout. Zo beschikken we ook over een alternatieve keuze voor de intervallen $(V_n)_n$, die aanleiding geeft tot nauwkeurigere w_n'etjes, en bijgevolg ook tot een kleinere N. Wanneer deze V_n'etjes op een naïeve manier berekend en gebruikt worden, vraagt dit echter veel rekentijd. Een efficiëntere manier om ze te gebruiken wordt onderzocht.

5 Omsluiting van functiewaarden: intervalaritmetiek

Aangezien de bovenstaande techniek toelaat de fout op een berekende functiewaarden betrouwbaar te begrenzen, is de stap naar intervalaritmetiek snel gezet. Een implementatie van multiprecisie-intervalaritmetiek die gebruik maakt van de technieken uit de vorige subsectie is momenteel in de maak.

Referenties

[1] Annie Cuyt, Brigitte Verdonk, Haakon Waadeland, and Johan Vervloet. Efficient and reliable multiprecision implementation of some elementary and special functions. in voorbereiding.

[2] W.B. Jones and W.J. Thron. *Continued fractions: analytic theory and applications*. Encycl. Math.11, Addison Wesley, London, 1981.

[3] Lisa Lorentzen and Haakon Waadeland. *Continued Fractions with Applications*. North-Holland, 1992.

[4] Brigitte Verdonk. On the fast and reliable multiprecision evaluation of some elementary functions. Lecture Notes 'International conference on numerical algorithms', Marrakesh, 2001.

[5] Brigitte Verdonk, Johan Vervloet, and Annie Cuyt. Blending interval and set arithmetic for maximal reliability. ingezonden ter publicatie.