Analyse d'erreur de récurrences linéaires
à l'aide de séries génératrices

Marc Mezzarobba

CNRS, LIX

Séminaire AriC, 21 janvier 2021

arXiv:2011.00827 [math.NA]

Polynômes de Legendre

$$(n+1)\,P_{n+1}(x) = (2\,n+1)\,x\,P_n(x) - n\,P_{n-1}(x)$$

$x = \diamond(17/18)$

	Intervalles	Flottants
$n = 0$	1.000000000000000	1.00000000000000
10	$[-0.386103889710 \pm 3.61\mathrm{e}-13]$	-0.386103889709935
20	$[\ \ 0.29953942 \pm 5.12\mathrm{e}-9]$	$0.29953941584394\textcolor{red}{1}$
30	$[-0.25184 \pm 7.04\mathrm{e}-6]$	-0.251843502730021
40	$[\ \ 0.2 \pm 0.0280]$	0.214139934045693
50	$[\pm 56.2]$	-0.179433059907698
60	$[\pm 2.31\mathrm{e}+5]$	0.145569066400857
70	$[\pm 9.66\mathrm{e}+8]$	$-0.1120286960441\textcolor{red}{30}$
80	$[\pm 4.09\mathrm{e}+12]$	$0.078978713542448\textcolor{red}{8}$
90	$[\pm 1.74\mathrm{e}+16]$	$-0.04690359400211\textcolor{red}{42}$
100	$[\pm 7.45\mathrm{e}+19]$	$0.01642951460849\textcolor{red}{27}$

Un exemple jouet

[Boldo 2009]

$$c_{n+1} = 2\,c_n - c_{n-1} \qquad (c_0 = \diamond(1/3),\, c_{-1} = 0)$$

$n =$	Intervalles	Flottants
0	$[0.333333333333333 \pm 1.49e-17]$	0.333333333333333
5	$[2.00000000000000 \pm 3.78e-15]$	2.00000000000000
10	$[3.666666666667 \pm 5.74e-13]$	3.66666666666667
15	$[5.3333333333 \pm 5.29e-11]$	5.33333333333334
20	$[7.00000000 \pm 1.60e-9]$	7.00000000000001
25	$[8.666667 \pm 4.65e-7]$	8.66666666666668
30	$[10.3333 \pm 4.41e-5]$	10.3333333333333
35	$[12.000 \pm 8.82e-4]$	12.0000000000000
40	$[1.4e+1 \pm 0.406]$	13.6666666666667
45	$[\pm 21.3]$	15.3333333333334
50	$[\pm 5.04e+2]$	17.0000000000000

Analyse d'erreur naïve

(erreur absolue / "virgule fixe" pour simplifier)

$$
\begin{aligned}
c_{n+1} &= 2\,c_n - c_{n-1} \\
\tilde{c}_{n+1} &= \diamond(2\,\tilde{c}_n - \tilde{c}_{n-1}) \\
&= 2\,\tilde{c}_n - \tilde{c}_{n-1} + \varepsilon_n \qquad |\varepsilon_n| \leqslant \mathbf{u}
\end{aligned}
$$

Analyse d'erreur naïve

(erreur absolue / “virgule fixe” pour simplifier)

$$\begin{aligned} c_{n+1} &= 2\,c_n - c_{n-1} \\ \tilde{c}_{n+1} &= \diamond(2\,\tilde{c}_n - \tilde{c}_{n-1}) \\ &= 2\,\tilde{c}_n - \tilde{c}_{n-1} + \varepsilon_n \qquad |\varepsilon_n| \leqslant \mathbf{u} \end{aligned}$$

$$|\tilde{c}_{n+1} - c_{n+1}| \leqslant 2\,|\tilde{c}_n - c_n| + |\tilde{c}_{n-1} - c_{n-1}| + \mathbf{u}$$

Analyse d'erreur naïve

(erreur absolue / "virgule fixe" pour simplifier)

$$
\begin{aligned}
c_{n+1} &= 2\,c_n - c_{n-1} \\
\tilde{c}_{n+1} &= \diamond(2\,\tilde{c}_n - \tilde{c}_{n-1}) \\
&= 2\,\tilde{c}_n - \tilde{c}_{n-1} + \varepsilon_n \qquad |\varepsilon_n| \leqslant \mathbf{u}
\end{aligned}
$$

$$|\tilde{c}_{n+1} - c_{n+1}| \leqslant 2\,|\tilde{c}_n - c_n| + |\tilde{c}_{n-1} - c_{n-1}| + \mathbf{u}$$

- récurrence : $|\tilde{c}_n - c_n| \leqslant 3^n\,\mathbf{u}$

Analyse d'erreur naïve

(erreur absolue / "virgule fixe" pour simplifier)

$$\begin{aligned} c_{n+1} &= 2\,c_n - c_{n-1} \\ \tilde{c}_{n+1} &= \diamond(2\,\tilde{c}_n - \tilde{c}_{n-1}) \\ &= 2\,\tilde{c}_n - \tilde{c}_{n-1} + \varepsilon_n \qquad |\varepsilon_n| \leqslant \mathbf{u} \end{aligned}$$

$$|\tilde{c}_{n+1} - c_{n+1}| \leqslant 2\,|\tilde{c}_n - c_n| + |\tilde{c}_{n-1} - c_{n-1}| + \mathbf{u}$$

- récurrence : $|\tilde{c}_n - c_n| \leqslant 3^n\,\mathbf{u}$

Analyse d'erreur naïve

(erreur absolue / "virgule fixe" pour simplifier)

$$\begin{aligned} c_{n+1} &= 2\,c_n - c_{n-1} \\ \tilde{c}_{n+1} &= \diamond(2\,\tilde{c}_n - \tilde{c}_{n-1}) \\ &= 2\,\tilde{c}_n - \tilde{c}_{n-1} + \varepsilon_n \qquad |\varepsilon_n| \leqslant \mathbf{u} \end{aligned}$$

$$|\tilde{c}_{n+1} - c_{n+1}| \leqslant 2\,|\tilde{c}_n - c_n| + |\tilde{c}_{n-1} - c_{n-1}| + \mathbf{u}$$

- récurrence : $|\tilde{c}_n - c_n| \leqslant 3^n\,\mathbf{u}$
- un poil plus fin :
  $|\tilde{c}_n - c_n| \leqslant (\lambda_+\,\alpha_+^n + \lambda_-\,\alpha_-^n - 4)\,\mathbf{u} \approx 2.4^n\,\mathbf{u}$

$$\alpha_\pm = 1 \pm \sqrt{2}, \quad \lambda_\pm = 4 \pm 3\sqrt{2}$$

Analyse d’erreur naïve

(erreur absolue / “virgule fixe” pour simplifier)

$$
\begin{aligned}
c_{n+1} &= 2\, c_n - c_{n-1} \\
\tilde{c}_{n+1} &= \diamond(2\, \tilde{c}_n - \tilde{c}_{n-1}) \\
&= 2\, \tilde{c}_n - \tilde{c}_{n-1} + \varepsilon_n \qquad |\varepsilon_n| \leqslant \mathbf{u}
\end{aligned}
$$

$$|\tilde{c}_{n+1} - c_{n+1}| \leqslant 2\, |\tilde{c}_n - c_n| + |\tilde{c}_{n-1} - c_{n-1}| + \mathbf{u}$$

- récurrence : $|\tilde{c}_n - c_n| \leqslant 3^n\, \mathbf{u}$
- un poil plus fin :
  $|\tilde{c}_n - c_n| \leqslant (\lambda_+\, \alpha_+^n + \lambda_-\, \alpha_-^n - 4)\, \mathbf{u} \approx 2.4^n\, \mathbf{u}$
  $$\alpha_\pm = 1 \pm \sqrt{2}, \quad \lambda_\pm = 4 \pm 3\sqrt{2}$$

C’est à peu près ce que fait l’arithmétique d’intervalles.

Analyse d'erreur plus raisonnable

("virgule fixe")

$$c_{n+1} = 2\,c_n - c_{n-1}$$

$$\tilde{c}_{n+1} = \diamond(2\,\tilde{c}_n - \tilde{c}_{n-1}) \qquad \delta_n = \tilde{c}_n - c_n$$

$$= 2\,\tilde{c}_n - \tilde{c}_{n-1} + \varepsilon_n \qquad |\varepsilon_n| \leqslant \mathbf{u}$$

$$\delta_{n+1} = 2\,\delta_n - \delta_{n-1} + \varepsilon_n \qquad (\delta_0 = \delta_1 = 0)$$

erreur globale ⤴ ⤶ *erreur locale*

Analyse d'erreur plus raisonnable

("virgule fixe")

$$\begin{aligned} c_{n+1} &= 2\,c_n - c_{n-1} \\ \tilde{c}_{n+1} &= \diamond(2\,\tilde{c}_n - \tilde{c}_{n-1}) \\ &= 2\,\tilde{c}_n - \tilde{c}_{n-1} + \varepsilon_n \qquad |\varepsilon_n| \leqslant \mathbf{u} \end{aligned}$$

$$\delta_n = \tilde{c}_n - c_n$$

$$\delta_{n+1} = 2\,\delta_n - \delta_{n-1} + \varepsilon_n \qquad (\delta_0 = \delta_1 = 0)$$

erreur globale *erreur locale*

- $\delta_n = \sum_{k=1}^{n-1} k\,\varepsilon_{n-k}$

  $|\delta_n| \leqslant \frac{n\,(n-1)}{2}\,\mathbf{u}$

Analyse d'erreur plus raisonnable

("virgule fixe")

$$\begin{aligned} c_{n+1} &= 2\,c_n - c_{n-1} \\ \tilde{c}_{n+1} &= \diamond(2\,\tilde{c}_n - \tilde{c}_{n-1}) \\ &= 2\,\tilde{c}_n - \tilde{c}_{n-1} + \varepsilon_n \qquad |\varepsilon_n| \leqslant \mathbf{u} \end{aligned} \qquad \boxed{\delta_n = \tilde{c}_n - c_n}$$

$$\delta_{n+1} = 2\,\delta_n - \delta_{n-1} + \varepsilon_n \qquad (\delta_0 = \delta_1 = 0)$$

erreur globale *erreur locale*

- $\delta_n = \sum_{k=1}^{n-1} k\,\varepsilon_{n-k}$

  $|\delta_n| \leqslant \frac{n\,(n-1)}{2}\,\mathbf{u}$

Il faut transporter une expression linéaire en les ε_k.

Les calculs peuvent devenir indigestes.

Un mot d'histoire

Peu de littérature sur les récurrences dans le sens direct

> in the backward direction. There has been less attention devoted to computation which utilizes the difference equation in the forward direction, not because a forward algorithm is more difficult to analyze, but rather for the opposite reason—that its analysis was considered straightforward. Of the above

[Wimp 1972]

von Neumann & Goldstine 1947, Turing 1948 — Systèmes triangulaires

Henrici 1962 — Schémas numériques pour les EDO

Oliver 1965 — Propagation des erreurs relatives dans les réc. linéaires

Liu & Kaneko 1969 — Filtres linéaires en flottant ($\to$ fns de transfert)

Hilaire & Lopez 2013 — Filtres linéaires, pire cas

Séries génératrices

$$(a_n)_{n\in\mathbb{Z}} \quad \longleftrightarrow \quad a(z) = \sum_{n=-\infty}^{\infty} a_n\, z^n$$

$$\delta_{n+1} = 2\,\delta_n - \delta_{n-1} + \varepsilon_n$$

$$\Big\downarrow \; \textstyle\sum_n \square\, z^n$$

$$z^{-1}\,\delta(z) = 2\,\delta(z) - z\,\delta(z) + \varepsilon(z)$$

$$\delta(z) = \frac{z}{(1-z)^2}\,\varepsilon(z)$$

- Formules (produit, composée, ...)
- Méthodes analytiques
  - Extraction de coefficients
  - Asymptotique
- Séries majorantes
- Calcul efficace
- ...

$$z\sum_n a_n\, z^n = \sum_n a_{n-1}\, z^n$$

Séries majorantes

On note $\mathbf{f} \ll \hat{\mathbf{f}}$ lorsque $\forall \mathbf{n}, \quad |\mathbf{f_n}| \leqslant \hat{\mathbf{f}}_\mathbf{n}.$

$$\left(f(z) = \textstyle\sum_n f_n \, z^n, \quad \hat{f}(z) = \textstyle\sum_n \hat{f}_n \, z^n\right)$$

Séries majorantes

On note $\mathbf{f} \ll \hat{\mathbf{f}}$ lorsque $\forall \mathbf{n}, \quad |\mathbf{f_n}| \leqslant \hat{\mathbf{f}}_\mathbf{n}$.

$\left(f(z) = \sum_n f_n z^n, \quad \hat{f}(z) = \sum_n \hat{f}_n z^n\right)$

$$\delta(z) = \frac{z}{(1-z)^2}\,\varepsilon(z)$$

Séries majorantes

On note $\mathbf{f} \ll \hat{\mathbf{f}}$ lorsque $\forall \mathbf{n}, \quad |\mathbf{f_n}| \leqslant \hat{\mathbf{f}}_{\mathbf{n}}.$

$$\left(f(z) = \sum_n f_n\, z^n, \quad \hat{f}(z) = \sum_n \hat{f}_n\, z^n\right)$$

$$\delta(z) = \frac{z}{(1-z)^2}\, \varepsilon(z)$$

- $\varepsilon(z) = \sum_n \varepsilon_n\, z^n \ll \frac{\mathbf{u}}{1-z}$

Séries majorantes

On note $\mathbf{f} \ll \hat{\mathbf{f}}$ lorsque $\forall \mathbf{n}, \quad |\mathbf{f_n}| \leqslant \hat{\mathbf{f}}_\mathbf{n}.$

$\left(f(z) = \sum_n f_n z^n, \quad \hat{f}(z) = \sum_n \hat{f}_n z^n\right)$

$$\delta(z) = \frac{z}{(1-z)^2}\,\varepsilon(z)$$

- $\varepsilon(z) = \sum_n \varepsilon_n z^n \ll \frac{\mathbf{u}}{1-z}$
- $\frac{z}{(1-z)^2} \gg 0$

Séries majorantes

On note $\mathbf{f} \ll \hat{\mathbf{f}}$ lorsque $\forall \mathbf{n},\quad |\mathbf{f_n}| \leqslant \hat{\mathbf{f}}_\mathbf{n}.$

$\left(f(z) = \sum_n f_n z^n, \quad \hat{f}(z) = \sum_n \hat{f}_n z^n\right)$

$$\delta(z) = \frac{z}{(1-z)^2}\,\varepsilon(z)$$

- $\varepsilon(z) = \sum_n \varepsilon_n z^n \ll \frac{\mathbf{u}}{1-z}$
- $\frac{z}{(1-z)^2} \gg 0$
- si $f(z) \ll \hat{f}(z)$ et $g(z) \ll \hat{g}(z)$,

  $f(z)\,g(z) \ll \hat{f}(z)\,\hat{g}(z)$

$$\left|[z^n]\left(f(z)\,g(z)\right)\right| = \left|\sum_{i+j=n} f_i\,g_j\right| \leqslant \sum_{i+j=n} \hat{f}_i\,\hat{g}_j = [z^n]\left(\hat{f}(z)\,\hat{g}(z)\right)$$

Séries majorantes

On note $\mathbf{f} \ll \hat{\mathbf{f}}$ lorsque $\forall \mathbf{n}, \quad |\mathbf{f_n}| \leqslant \hat{\mathbf{f}}_\mathbf{n}$.

$\left(f(z) = \sum_n f_n z^n, \quad \hat{f}(z) = \sum_n \hat{f}_n z^n\right)$

$$\delta(z) = \frac{z}{(1-z)^2}\,\varepsilon(z)$$

$$\ll \frac{\mathbf{u}\,z}{(1-z)^3}$$

- $\varepsilon(z) = \sum_n \varepsilon_n z^n \ll \frac{\mathbf{u}}{1-z}$
- $\frac{z}{(1-z)^2} \gg 0$
- si $f(z) \ll \hat{f}(z)$ et $g(z) \ll \hat{g}(z)$,

  $f(z)\,g(z) \ll \hat{f}(z)\,\hat{g}(z)$

$$\left|[z^n]\,\big(f(z)\,g(z)\big)\right| = \left|\sum_{i+j=n} f_i\,g_j\right| \leqslant \sum_{i+j=n} \hat{f}_i\,\hat{g}_j = [z^n]\,\big(\hat{f}(z)\,\hat{g}(z)\big)$$

Séries majorantes

On note $\mathbf{f} \ll \hat{\mathbf{f}}$ lorsque $\forall \mathbf{n}, \quad |\mathbf{f_n}| \leqslant \hat{\mathbf{f}}_\mathbf{n}.$

$\left(f(z) = \sum_n f_n z^n, \quad \hat{f}(z) = \sum_n \hat{f}_n z^n\right)$

$$\delta(z) = \frac{z}{(1-z)^2}\,\varepsilon(z)$$

$$\ll \frac{\mathbf{u}\,z}{(1-z)^3}$$

$$|\delta_n| \leqslant \frac{n\,(n-1)}{2}\,\mathbf{u}$$

- $\varepsilon(z) = \sum_n \varepsilon_n z^n \ll \frac{\mathbf{u}}{1-z}$
- $\frac{z}{(1-z)^2} \gg 0$
- si $f(z) \ll \hat{f}(z)$ et $g(z) \ll \hat{g}(z)$,

  $f(z)\,g(z) \ll \hat{f}(z)\,\hat{g}(z)$

$$\left|[z^n]\,\big(f(z)\,g(z)\big)\right| = \left|\sum_{i+j=n} f_i\,g_j\right| \leqslant \sum_{i+j=n} \hat{f}_i\,\hat{g}_j = [z^n]\,\big(\hat{f}(z)\,\hat{g}(z)\big)$$

Flottants, erreur relative

$$
\begin{aligned}
c_{n+1} &= 2\,c_n - c_{n-1} \\
\tilde{c}_{n+1} &= \diamond(2\,\tilde{c}_n - \tilde{c}_{n-1}) \\
&= (2\,\tilde{c}_n - \tilde{c}_{n-1})\,(1+\varepsilon_n) \qquad |\varepsilon_n| \leqslant \mathbf{u}
\end{aligned}
$$

Flottants, erreur relative

$$
\begin{aligned}
(1+\varepsilon_n) \qquad c_{n+1} &= 2\,c_n - c_{n-1} \\
(-1) \qquad \tilde{c}_{n+1} &= \diamond(2\,\tilde{c}_n - \tilde{c}_{n-1}) \\
&= (2\,\tilde{c}_n - \tilde{c}_{n-1})\,(1+\boldsymbol{\varepsilon_n}) \qquad |\varepsilon_n| \leqslant \mathbf{u}
\end{aligned}
$$

$$\delta_{n+1} - c_{n+1}\,\varepsilon_n = (2\,\delta_n - \delta_{n-1})\,(1+\varepsilon_n)$$

$$\delta_n = \tilde{c}_n - c_n$$

$$\delta_{n+1} - 2\,\delta_n + \delta_{n-1} = \varepsilon_n\,(c_{n+1} + 2\,\delta_n - \delta_{n-1})$$

$$(z^{-1} - 2 + z)\,\delta(z) = \varepsilon(z) \odot \big(z^{-1}\,c(z) + (2-z)\,\delta(z)\big)$$

$$\delta(z) = \frac{(z\,\varepsilon(z)) \odot \big(c(z) + z\,(2-z)\,\delta(z)\big)}{(1-z)^2}$$

Flottants, erreur relative

$$
\begin{aligned}
(1+\varepsilon_n) \qquad & c_{n+1} = 2\,c_n - c_{n-1} \\
(-1) \qquad & \tilde{c}_{n+1} = \diamond(2\,\tilde{c}_n - \tilde{c}_{n-1}) \\
& \qquad\quad = (2\,\tilde{c}_n - \tilde{c}_{n-1})\,(1+\boldsymbol{\varepsilon_n}) \qquad |\varepsilon_n| \leqslant \mathbf{u}
\end{aligned}
$$

$$\delta_{n+1} - c_{n+1}\,\varepsilon_n = (2\,\delta_n - \delta_{n-1})\,(1+\varepsilon_n)$$

$$\delta_n = \tilde{c}_n - c_n$$

$$\delta_{n+1} - 2\,\delta_n + \delta_{n-1} = \varepsilon_n\,(c_{n+1} + 2\,\delta_n - \delta_{n-1})$$

$$(z^{-1} - 2 + z)\,\delta(z) = \varepsilon(z) \odot \big(z^{-1}\,c(z) + (2-z)\,\delta(z)\big)$$

$$\delta(z) = \frac{(z\,\varepsilon(z)) \odot \big(c(z) + z\,(2-z)\,\delta(z)\big)}{(1-z)^2}$$

Flottants, erreur relative

$$
\begin{aligned}
(1+\varepsilon_n) \qquad & c_{n+1} = 2\,c_n - c_{n-1} \\
(-1) \qquad & \tilde{c}_{n+1} = \diamond(2\,\tilde{c}_n - \tilde{c}_{n-1}) \\
& \qquad = (2\,\tilde{c}_n - \tilde{c}_{n-1})\,{\color{blue}(1+\boldsymbol{\varepsilon}_n)} \qquad |\varepsilon_n| \leqslant \mathbf{u}
\end{aligned}
$$

$$\delta_{n+1} - c_{n+1}\,\varepsilon_n = (2\,\delta_n - \delta_{n-1})\,(1+\varepsilon_n)$$

$$\boxed{\delta_n = \tilde{c}_n - c_n} \qquad \boxed{\delta_{n+1} - 2\,\delta_n + \delta_{n-1} = \varepsilon_n\,(c_{n+1} + 2\,\delta_n - \delta_{n-1})}$$

$$(z^{-1} - 2 + z)\,\delta(z) = \varepsilon(z) \odot \big(z^{-1}\,c(z) + (2-z)\,\delta(z)\big)$$

$$\delta(z) = \frac{(z\,\varepsilon(z)) \odot \big(c(z) + z\,(2-z)\,\delta(z)\big)}{(1-z)^2}$$

$${}^\sharp f(z) = \sum_n |f_n|\,z^n$$

$$c(z) \ll \hat{c}(z)$$

$$\delta(z) \ll \frac{z\,(2+z)\,\mathbf{u}}{(1-z)^2}\,{}^\sharp\delta(z) + \frac{\hat{c}(z)\,\mathbf{u}}{(1-z)^2}$$

Équations majorantes

Lemme **[Cauchy ?]**

Soient $\hat{a}(z), \hat{b}(z) \in \mathbb{R}_+[[z]]$ avec $\hat{a}(0) = 0$. Suppons $y \in \mathbb{R}_+[[z]]$ t.q.

$$y(z) \ll \hat{a}(z)\, y(z) + \hat{b}(z).$$

Alors $y(z)$ est bornée par la solution de $\hat{y}(z) = \hat{a}(z)\, \hat{y}(z) + \hat{b}(z)$, i.e.,

$$y(z) \ll \hat{y}(z) = \frac{\hat{b}(z)}{1 - \hat{a}(z)}.$$

Preuve.

- $y_0 \leqslant \hat{b}_0 = \hat{y}_0$
- $y_n \leqslant \sum_{i=0}^{n} \hat{a}_i\, y_{n-i} + \hat{b}_n$

Équations majorantes

Lemme **[Cauchy ?]**

Soient $\hat{a}(z), \hat{b}(z) \in \mathbb{R}_+[[z]]$ avec $\hat{a}(0) = 0$. Suppons $y \in \mathbb{R}_+[[z]]$ t.q.

$$y(z) \ll \hat{a}(z)\, y(z) + \hat{b}(z).$$

Alors $y(z)$ est bornée par la solution de $\hat{y}(z) = \hat{a}(z)\, \hat{y}(z) + \hat{b}(z)$, i.e.,

$$y(z) \ll \hat{y}(z) = \frac{\hat{b}(z)}{1 - \hat{a}(z)}.$$

Preuve.

- $y_0 \leqslant \hat{b}_0 = \hat{y}_0$
- $y_n \leqslant \sum_{i=1}^{n} \hat{a}_i\, y_{n-i} + \hat{b}_n$

Équations majorantes

Lemme **[Cauchy ?]**

Soient $\hat{a}(z), \hat{b}(z) \in \mathbb{R}_+[[z]]$ avec $\hat{a}(0) = 0$. Suppons $y \in \mathbb{R}_+[[z]]$ t.q.

$$y(z) \ll \hat{a}(z)\, y(z) + \hat{b}(z).$$

Alors $y(z)$ est bornée par la solution de $\hat{y}(z) = \hat{a}(z)\, \hat{y}(z) + \hat{b}(z)$, i.e.,

$$y(z) \ll \hat{y}(z) = \frac{\hat{b}(z)}{1 - \hat{a}(z)}.$$

Preuve.

- $y_0 \leqslant \hat{b}_0 = \hat{y}_0$
- $y_n \leqslant \sum_{i=1}^{n} \hat{a}_i\, y_{n-i} + \hat{b}_n \leqslant \sum_{i=1}^{n} \hat{a}_i\, \hat{y}_{n-i} + \hat{b}_n = \hat{y}_n$

Borne sur l'erreur en flottants

$$^\sharp\delta(z) \ll \underbrace{\frac{z\,(2+z)\,\mathbf{u}}{(1-z)^2}}_{\hat{a}(z)}\,^\sharp\delta(z) + \underbrace{\frac{\hat{c}(z)\,\mathbf{u}}{(1-z)^2}}_{\hat{b}(z)}$$

$$c(z) = \frac{c_0}{(1-z)^2} \ll \frac{|c_0|}{(1-z)^2}$$

$$\begin{aligned}
\Delta(z) \quad &\ll \frac{\hat{b}(z)}{1-\hat{a}(z)} \\
&= \frac{\hat{c}(z)\,\mathbf{u}}{1-2\,(1+\mathbf{u})\,z+(1-\mathbf{u})\,z^2} \\
&= \frac{\hat{c}(z)\,\mathbf{u}}{(1-\boldsymbol{\alpha}\,z)\,(1-\boldsymbol{\beta}\,z)} \\
&\ll \frac{|c_0|\mathbf{u}}{(1-\boldsymbol{\alpha}\,z)^4} \qquad \longrightarrow
\end{aligned}$$

Exponentiel, mais $O(1)$ si $n = O(u^{-1/2})$

$\boldsymbol{\alpha} = 1 + \sqrt{3\,\mathbf{u}} + O(u)$

$\boldsymbol{\alpha} \leqslant 1 + 2\sqrt{\mathbf{u}}$ pour $u \leqslant 0.008$

$$|\delta_n| \leqslant \frac{|c_0|}{6}\,(n+3)^3\,\alpha^n\,\mathbf{u}$$

($\delta_n =$ erreur *absolue* sur c_n)

D'autres exemples

Polynômes de Legendre

- $\mathbf{P}_{n+1}(x) = \dfrac{1}{n+1}\left[(2n+1)\,x\,\mathbf{P}_n(x) - n\,\mathbf{P}_{n-1}(x)\right]$
- En arithmétique virgule fixe :

$$|\tilde{p}_n - P_n(x)| \leqslant \frac{3}{4}\,(n+1)\,(n+2)\,\mathbf{u} \qquad (-1 \leqslant x \leqslant 1)$$

- Coefficients polynomiaux $\longrightarrow$ équa. diff.

Utile pour le calcul de règles de quadrature de Gauss-Legendre [Johansson-M. 2018]

Corde vibrante [Boldo 2009, Boldo *et al.* 2013, 2014]

- $\dfrac{\partial^2 p}{\partial t^2} - c^2\,\dfrac{\partial^2 p}{\partial x^2} = 0 \quad \xrightarrow{\text{discr.}} \quad p_i^{k+1} = 2\,p_i^k - p_i^{k-1} + a\,(p_{i+1}^k - 2\,p_i^k + p_{i-1}^k)$
- En `double` :

$$|\tilde{p}_i^k - p_i^k| \leqslant 39 \cdot 2^{-52} \cdot (k+1)\,(k+2)$$

- Normes sur $\Omega = \mathbb{R}[x]\,/\,\langle x^{2n} - 1\rangle$ et séries dans $\Omega[[t]]$

Nombres de Bernoulli

$$B_n = 1, \frac{-1}{2}, \frac{1}{6}, 0, \frac{-1}{30}, 0, \frac{1}{42}, 0, \frac{-1}{30}, 0, \frac{5}{66}, 0, \frac{-691}{2730}, 0, \frac{7}{6}, 0, \frac{-3617}{510}, \ldots \qquad |B_{2k}| \sim \frac{2\,(2\,k)!}{(2\,\pi)^{2k}}$$

$$b_k = \frac{B_{2k}}{(2\,k)!} \qquad b(z) = \sum_{k=0}^{\infty} b_k\, z^k = \frac{\sqrt{z}/2}{\tanh(\sqrt{z}/2)}$$

Algorithme

[Brent 1980, sugg. de Reinsch]

$$b_k = \frac{1}{(2\,k)!\,4^k} - \sum_{j=0}^{k-1} \frac{b_j}{(2\,k+1-2\,j)!\,4^{k-j}}$$

be used with sufficient guard digits, or a more stable recurrence must be used. If we multiply both sides of (30) by $\sinh(x/2)/x$ and equate coefficients, we get the recurrence

$$C_k + \frac{C_{k-1}}{3!\,4} + \cdots + \frac{C_1}{(2k-1)!\,4^{k-1}} = \frac{2k}{(2k+1)!\,4^k} \tag{36}$$

If (36) is used to evaluate C_k, using precision n arithmetic, the error is only $O(k^2 2^{-n})$. Thus,

[Brent 1980]

Erreur relative dans le calcul de b_k

> **Exercise 4.35** Prove (or give a plausibility argument for) the statements made in §4.7 that: (a) if a recurrence based on (4.59) is used to evaluate the scaled Bernoulli number C_k, using precision n arithmetic, then the relative error is of order $4^k 2^{-n}$; and (b) if a recurrence based on (4.60) is used, then the relative error is $O(k^2 2^{-n})$.

[Brent & Zimmermann 2010]

Conjecture [Brent & Zimmermann ?]

On a $\tilde{b}_k = b_k(1+\eta_k)$ où $\eta_k = \mathbf{O}(\mathbf{k}\cdot\mathbf{u})$.

$\tilde{b}_k =$ valeur calculée en flottants, $\quad u =$ unité d'arrondi

Remarque

À comprendre comme $\eta_k = O(k \cdot u)$ pour $k = O(u^{-1})$,
ou encore $|\eta_k| \leqslant C_k u$ quand $u \to 0$ où $C_k = O(k)$ (resp. $O(k^2)$)

Borne explicite sur l'erreur relative

$$b(z) = \frac{\sqrt{z}/2}{\tanh(\sqrt{z}/2)} \qquad b_k = \frac{1}{(2\,k)!\,4^k} - \sum_{j=0}^{k-1} \frac{b_j}{(2\,k+1-2\,j)!\,4^{k-j}}$$

$$\tilde{b}_k = b_k\,(1+\eta_k)$$

Proposition

L'erreur relative sur le résultat satifait

$$|\eta_k| \leqslant (1+21.2\,u)^k\,(1.1\,k+446)\,u$$

Corollaire : $|\eta_k| \leqslant (3\,k+1213)\,u \quad$ pour $\quad u < 2^{-16}$ et $43\,k\,u \leqslant 1$.

Schéma de la preuve

$\tilde{b}_k =$ valeurs calculées

Analyse d'erreur locale.

$$\tilde{b}_k = \frac{1+s_k}{(2k)!\,4^k} - \sum_{j=0}^{k-1} \frac{\tilde{b}_j\,(1+t_{k,j})}{(2k+1-2j)!\,4^{k-j}}$$

$$|s_k| \leqslant \hat{\theta}_{2k}$$

$$|t_{k,j}| \leqslant \hat{\theta}_{3(k-j)+2}$$

où $\hat{\theta}_n = (1+u)^n - 1$

Schéma de la preuve

$\tilde{b}_k = \text{valeurs calculées}$

Analyse d'erreur locale.

$$\tilde{b}_k = \frac{1 + \mathbf{s_k}}{(2k)!\,4^k} - \sum_{j=0}^{k-1} \frac{\tilde{b}_j\,(1 + \mathbf{t_{k,j}})}{(2k+1-2j)!\,4^{k-j}}$$

$$|\mathbf{s_k}| \leqslant \hat{\theta}_{2k}$$
$$|\mathbf{t_{k,j}}| \leqslant \hat{\theta}_{3(k-j)+2}$$
où $\hat{\theta}_n = (1+u)^n - 1$

Linéarité. $\delta_k := \tilde{b}_k - b_k = \text{erreur globale}$

$$\delta_k = \frac{\mathbf{s_k}}{(2k)!\,4^k} - \sum_{j=0}^{k-1} \frac{\delta_j + \tilde{b}_j\,\mathbf{t_{k,j}}}{(2k+1-2j)!\,4^{k-j}}$$

Schéma de la preuve

$\tilde{b}_k = $ valeurs calculées

Analyse d'erreur locale.

$$\tilde{b}_k = \frac{1+\mathbf{s_k}}{(2\,k)!\,4^k} - \sum_{j=0}^{k-1} \frac{\tilde{b}_j\,(1+\mathbf{t_{k,j}})}{(2\,k+1-2\,j)!\,4^{k-j}}$$

$$|\mathbf{s_k}| \leqslant \hat{\theta}_{2k}$$
$$|\mathbf{t_{k,j}}| \leqslant \hat{\theta}_{3(k-j)+2}$$
où $\hat{\theta}_n = (1+u)^n - 1$

Linéarité. $\delta_k := \tilde{b}_k - b_k = $ erreur globale

$$\delta_k = \frac{\mathbf{s_k}}{(2\,k)!\,4^k} - \sum_{j=0}^{k-1} \frac{\delta_j + \tilde{b}_j\,\mathbf{t_{k,j}}}{(2\,k+1-2\,j)!\,4^{k-j}}$$

Inéquation sur l'erreur globale.

$$\delta(z) \ll \check{S}(z)\,\tilde{C}(z) + \check{S}(z)\,\tilde{S}(z)\,{}^\sharp b(z) + \check{S}(z)\,\tilde{S}(z)\,{}^\sharp\delta(z)$$

$${}^\sharp f(z) = \sum_k |f_k|\,z^k$$

où $C(z) = \cosh(\sqrt{z}/2)$, $S(z) = (\sqrt{z}/2)^{-1}\sinh(\sqrt{z}/2)$, $\check{S}(z) = \dfrac{(\sqrt{z}/2)}{\sin(\sqrt{z}/2)}$,

$\tilde{C}(z) = C(a^2\,z) - C(z)$, $\tilde{S}(z) = S(a^4\,z) - S(z) - (a^2-1)$ avec $a = 1+u$

Schéma de la preuve

$$\delta(z) \ll \check{S}(z)\,\tilde{C}(z) + \check{S}(z)\,\tilde{S}(z)\,{}^{\sharp}\mathfrak{b}(z) + \check{S}(z)\,\tilde{S}(z)\,{}^{\sharp}\delta(z)$$

Majorant « explicite ».

$$\delta(z) \ll \frac{\check{S}(z)\,\tilde{C}(z) + \check{S}(z)\,\tilde{S}(z)\,{}^{\sharp}\mathfrak{b}(z)}{1 - \check{S}(z)\,\tilde{S}(z)} \quad =: \hat{\delta}(z)$$

Schéma de la preuve

$$\delta(z) \ll \check{S}(z)\,\tilde{C}(z) + \check{S}(z)\,\tilde{S}(z)\,{}^\sharp\mathfrak{b}(z) + \check{S}(z)\,\tilde{S}(z)\,{}^\sharp\delta(z)$$

Majorant « explicite ».

$$\delta(z) \ll \frac{\check{S}(z)\,\tilde{C}(z) + \check{S}(z)\,\tilde{S}(z)\,{}^\sharp\mathfrak{b}(z)}{1 - \check{S}(z)\,\tilde{S}(z)} \quad =: \hat{\delta}(z)$$

Asymptotique.

$$\hat{\delta}(z) = \left(\frac{2\,(1-\cosh w)\cos(w)}{w^{-2}\sin(w)^2} + \frac{4\,(\cosh w - 1) + w \sinh w}{w^{-1}\sin w} \right) u + O(u^2)$$

$$w = \sqrt{z}\,/\,2$$

Unique pôle de module minimal en $z = 4\,\pi^2$,
multiplicité (en z) $= 2$

$\Rightarrow \hat{\delta}_k = O(k\,(2\,\pi)^{-2k}) \cdot u + O(u^2)$

$\Rightarrow \eta_k = \text{“}O(k \cdot u)\text{”}$

Schéma de la preuve

$$\delta(z) = \hat{\delta}(z) \ll \frac{\check{S}(z)\,\tilde{C}(z) + \check{S}(z)\,\tilde{S}(z)\,{}^{\sharp}b(z)}{1 - \check{S}(z)\,\tilde{S}(z)}$$

Contrôle du pôle dominant.

Supposons $u \leqslant 2^{-16}$.

Alors $\hat{\delta}(z)$ a un pôle en $\gamma = \left(\frac{2\,\pi}{1+\varphi(u)}\right)^2$ où $0 \leqslant \varphi(u) \leqslant 2\,(\cosh \pi - 1)\,u$.

C'est le seul avec $|z| < 153.7 \approx (3.9\,\pi)^2$.

(Un peu d'analyse, et comparaison avec le cas limité par le théorème de Rouché.)

Schéma de la preuve

$$\delta(z) = \hat{\delta}(z) \ll \frac{\check{S}(z)\,\tilde{C}(z) + \check{S}(z)\,\tilde{S}(z)\,{}^{\sharp}b(z)}{1 - \check{S}(z)\,\tilde{S}(z)}$$

Contrôle du pôle dominant.

Supposons $u \leqslant 2^{-16}$.

Alors $\hat{\delta}(z)$ a un pôle en $\gamma = \left(\frac{2\,\pi}{1+\varphi(u)}\right)^2$ où $0 \leqslant \varphi(u) \leqslant 2\,(\cosh \pi - 1)\,u$.

C'est le seul avec $|z| < 153.7 \approx (3.9\,\pi)^2$.

(Un peu d'analyse, et comparaison avec le cas limité par le théorème de Rouché.)

Symbolique-numérique.

$$\hat{\delta}(z) \ll \frac{2\,|\mathbf{R}(\mathbf{u}) - 1|}{1 - z/\gamma} + \frac{2\,\varphi(u)\,(2+\varphi(u))\,z/(2\,\alpha)^2}{(1-z/\gamma)^2} + \frac{\mathbf{A}_{\boldsymbol{\lambda}}(\mathbf{u})}{1 - z/(\lambda\,\gamma)}$$

Formule de Cauchy + arithmétique d'intervalles.

Sommes partielles de séries D-finies

Contexte : résolution numérique de $\mathbf{L} \cdot \mathbf{u} = 0$ où $L \in \mathbb{C}[z]\langle {}^{d}/_{dz} \rangle$

Calculer un encadrement de $\displaystyle\sum_{n=0}^{N-1} u_n \, \zeta^n$.

L, $u_{0:r-1}$, ζ, N donnés

- **Précision cible** p
- **Hypothèses**

$$\begin{aligned} \text{point ordinaire} &\quad \text{—} \quad a_r(0) \neq 0 \\ \text{cvgce « évidente »} &\quad \text{—} \quad |\zeta| < \min\{|\xi| : a_r(\xi) = 0\} \\ \text{cvgce géométrique} &\quad \text{—} \quad N = \Theta(p) \end{aligned}$$

$$L = a_r(z) \left(\frac{d}{dz}\right)^r + \cdots$$

Naïve approach

Recurrence step: (naïve interval arithmetic, with $u_n \approx \Theta(1)$)

$$\mathbf{u_n} = \frac{-1}{b_s(n)} \left[b_{s-1}(n)\, \mathbf{u_{n-1}} + \cdots + b_1(n)\, \mathbf{u_{n-s+1}} + b_0(n)\, \mathbf{u_{n-s}} \right]$$

$\mathrm{rad} \approx \rho$

$$\mathrm{rad} \gtrsim \left(\sum_i \left| \frac{b_i(s)}{b_s(n)} \right| \right) \rho \gtrapprox s\, \rho$$

$\mathrm{rad}(\mathbf{u_n}) = 2^{\Theta(n)}$ $\qquad$ $\mathrm{rad}(\sum^N u_n \zeta^n) = 2^{\Theta(N)}$ (unless ζ small)

Accuracy target 2^{-p} $\Rightarrow$ Need $\mathbf{\Omega(N)} = \Omega(p)$ guard bits

(This is not a numerical stability issue.)

Running Error Analysis

$$\begin{aligned} \mathbf{u_n} &= \frac{-1}{b_s(n)}\left[b_{s-1}(n)\,\mathbf{u_{n-1}} + \cdots + b_1(n)\,\mathbf{u_{n-s+1}} + b_0(n)\,\mathbf{u_{n-s}}\right] \\ \tilde{\mathbf{u}}_\mathbf{n} &= \frac{-1}{b_s(n)}\left[b_{s-1}(n)\,\tilde{\mathbf{u}}_\mathbf{n-1} + \cdots + b_1(n)\,\tilde{\mathbf{u}}_\mathbf{n-s+1} + b_0(n)\,\tilde{\mathbf{u}}_\mathbf{n-s}\right] + \varepsilon_n \end{aligned}$$

$\tilde{u}_n =$ computed sequence (e.g. floating-point)

local error,
known bound $|\varepsilon_n| \leqslant \hat{\varepsilon}_n$
(can be computed on the fly)

The global error $\delta_n = \tilde{u}_n - u_n$ satisfies

$$b_s(n)\,\delta_n + b_{s-1}(n)\,\delta_{n-1} + \cdots + b_0(n)\,\delta_{n-s} = b_s(n)\,\varepsilon_n$$

Running Error Analysis

$$\mathbf{u_n} = \frac{-1}{b_s(n)} \left[b_{s-1}(n)\, \mathbf{u_{n-1}} + \cdots + b_1(n)\, \mathbf{u_{n-s+1}} + b_0(n)\, \mathbf{u_{n-s}} \right]$$

$$\tilde{\mathbf{u}}_\mathbf{n} = \frac{-1}{b_s(n)} \left[b_{s-1}(n)\, \tilde{\mathbf{u}}_\mathbf{n-1} + \cdots + b_1(n)\, \tilde{\mathbf{u}}_\mathbf{n-s+1} + b_0(n)\, \tilde{\mathbf{u}}_\mathbf{n-s} \right] + \varepsilon_n$$

$\tilde{u}_n =$ computed sequence (e.g. floating-point)

local error,
known bound $|\varepsilon_n| \leqslant \hat{\varepsilon}_n$
(can be computed on the fly)

The global error $\delta_n = \tilde{u}_n - u_n$ satisfies

$$b_s(n)\, \delta_n + b_{s-1}(n)\, \delta_{n-1} + \cdots + b_0(n)\, \delta_{n-s} = b_s(n)\, \varepsilon_n$$

Therefore:

$$a_r(z)\, \delta^{(r)}(z) + \cdots + a_1(z)\, \delta'(z) + a_0(z)\, \delta(z) = Q(z\, {}^{d}\!/\!{}_{dz})\, \varepsilon(z)$$

$$\delta(z) = \sum_n \delta_n z^n, \quad \varepsilon(z) = \sum_n \varepsilon_n z^n$$

Compute a **bound** on δ_n given one on ε_n?

Global Error

$$a_r(z)\,\delta^{(r)}(z)+\cdots+a_0(z)\,\delta(z)=Q(z\,{}^{d}\!/_{dz})\,\varepsilon(z) \quad \xrightarrow{\text{maj.}} \quad \hat{\delta}'(z)=\hat{a}(z)\,\hat{\delta}(z)+\hat{\varepsilon}(z)$$

$$|\delta_0|,\ldots,|\delta_{r-1}|\leqslant\hat{\delta}_0 \quad \Rightarrow \quad \delta(z)\ll\hat{\delta}(z)$$

Global Error

$$a_r(z)\,\delta^{(r)}(z)+\cdots+a_0(z)\,\delta(z)=Q(z\,{}^{d}\!/_{dz})\,\varepsilon(z)\quad\xrightarrow{\text{maj.}}\quad\hat\delta'(z)=\hat a(z)\,\hat\delta(z)+\hat\varepsilon(z)$$

$$|\delta_0|,\ldots,|\delta_{r-1}|\leqslant\hat\delta_0\quad\Rightarrow\quad\delta(z)\ll\hat\delta(z)$$

$$\hat\delta(z)=\hat h(z)\left(\mathrm{cst}+\int_0^z\frac{\hat\varepsilon(w)}{\hat h(w)}\,dw\right),\qquad\hat h(z)=\exp\int_0^z\hat a(z)\,dz$$

Global Error

$$a_r(z)\,\delta^{(r)}(z)+\cdots+a_0(z)\,\delta(z)=Q(z\,{}^{d}\!/_{dz})\,\varepsilon(z)\quad\xrightarrow{\text{maj.}}\quad\hat\delta'(z)=\hat a(z)\,\hat\delta(z)+\hat\varepsilon(z)$$

$$|\delta_0|,\ldots,|\delta_{r-1}|\leqslant\hat\delta_0\quad\Rightarrow\quad\delta(z)\ll\hat\delta(z)$$

$$\hat\delta(z)=\hat h(z)\left(\mathrm{c\cancel{s}t}+\int_0^z\frac{\hat\varepsilon(w)}{\hat h(w)}\,dw\right),\qquad\hat h(z)=\exp\int_0^z\hat a(z)\,dz$$

Global Error

$$a_r(z)\,\delta^{(r)}(z)+\cdots+a_0(z)\,\delta(z)=Q(z\,{}^{d}\!/_{dz})\,\varepsilon(z)\quad\xrightarrow{\text{maj.}}\quad\hat\delta'(z)=\hat a(z)\,\hat\delta(z)+\hat\varepsilon(z)$$

$$|\delta_0|,\ldots,|\delta_{r-1}|\leqslant\hat\delta_0\quad\Rightarrow\quad\delta(z)\ll\hat\delta(z)$$

$$\hat\delta(z)=\hat h(z)\left(\cancel{\mathrm{cst}}+\int_0^z\frac{\hat\varepsilon(w)}{\hat h(w)}\,dw\right),\qquad\hat h(z)=\exp\int_0^z\hat a(z)\,dz$$

take $\hat\varepsilon_n=\bar\varepsilon\,\hat h_n$ $\qquad\bar\varepsilon\lesssim n^r\,\mathbf{u}$ since $|u_n|\lesssim\hat h_n$

Global Error

$$a_r(z)\,\delta^{(r)}(z)+\cdots+a_0(z)\,\delta(z)=Q(z\,{}^{d}\!/_{dz})\,\varepsilon(z)\quad\xrightarrow{\text{maj.}}\quad\hat\delta'(z)=\hat a(z)\,\hat\delta(z)+\hat\varepsilon(z)$$

$$|\delta_0|,\ldots,|\delta_{r-1}|\leqslant\hat\delta_0\quad\Rightarrow\quad\delta(z)\ll\hat\delta(z)$$

$$\hat\delta(z)=\hat h(z)\left(\cancel{\text{cst}}+\int_0^z\frac{\hat\varepsilon(w)}{\hat h(w)}\,dw\right),\qquad \hat h(z)=\exp\int_0^z\hat a(z)\,dz$$

take $\hat\varepsilon_n=\bar\varepsilon\,\hat h_n$ $\qquad$ $\bar\varepsilon\lesssim n^r\,\mathbf{u}$ since $|u_n|\lesssim\hat h_n$

$$|\tilde u(\zeta)-u(\zeta)|\leqslant\hat\delta(|\zeta|)=O(\bar\varepsilon)=O(2^{-p_{\text{work}}+\log N})$$

$\#$guard bits $=O(\log N)$ $\quad$ for fixed L, ini, ζ

Global Error

$$a_r(z)\,\delta^{(r)}(z)+\cdots+a_0(z)\,\delta(z)=Q(z\,{}^{d}\!/_{dz})\,\varepsilon(z)\qquad\xrightarrow{\text{maj.}}\qquad\hat{\delta}'(z)=\hat{a}(z)\,\hat{\delta}(z)+\hat{\varepsilon}(z)$$

$$|\delta_0|,\ldots,|\delta_{r-1}|\leqslant\hat{\delta}_0\quad\Rightarrow\quad\delta(z)\ll\hat{\delta}(z)$$

$$\hat{\delta}(z)=\hat{h}(z)\left(c\cancel{s}t+\int_0^z\frac{\hat{\varepsilon}(w)}{\hat{h}(w)}\,dw\right),\qquad\hat{h}(z)=\exp\int_0^z\hat{a}(z)\,dz$$

take $\hat{\varepsilon}_n=\bar{\varepsilon}\,\hat{h}_n$ $\qquad$ $\bar{\varepsilon}\lesssim n^r\,\mathbf{u}$ since $|u_n|\lesssim\hat{h}_n$

$$|\tilde{u}(\zeta)-u(\zeta)|\leqslant\hat{\delta}(|\zeta|)=O(\bar{\varepsilon})=O(2^{-p_{\text{work}}+\log N})$$

$\#$guard bits $=O(\log N)$ $\quad$ for fixed L, ini, ζ

The same computation yields a bound on the **truncation** error!

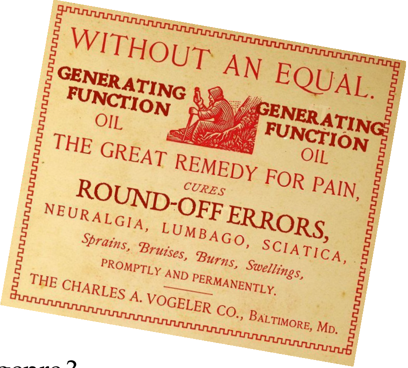
(Replace $\varepsilon(z)$ by a residual accounting for the negelected tail.)

Practical Issues

- The Cauchy majorants are *far* too coarse
  - Use sharper variants [M. 2019]
- Simple majorants cannot be sharp for small n
  - Switch from interval summation to running error analysis
- Need to choose the working precision ($\leftrightarrow$ cutoff point) in advance
  - Heuristics based on asymptotics…
- Good choice of $\hat{\varepsilon}(z)$ (tight & easily computable) not clear

Current status: works well for (some) large equations met in practice, but sometimes slower than naïve interval summation.

Conclusion



Une analyse d'erreur pénible ?

Essayez les séries génératrices !

- Erreur locale $\leftrightarrow$ erreur globale
- Expressions exactes ou équations
- Séries majorantes

Qui a déjà vu des choses de ce genre ?

Algorithme pratique pour les D-finies
(y compris cas singulier)

Récurrences "à rebours" (Miller) ?

Polynômes orthogonaux généraux ?

Schémas d'intégration numérique ?

Polynômes de Legendre

[Johansson & M. 2018]

$$p_{n+1} = \frac{1}{n+1}\left[(2\,n+1)\,x\,p_n - n\,p_{n-1}\right] \qquad x \text{ fixé, } p_n = P_n(x)$$

$$\tilde{p}_{n+1} = \frac{1}{n+1}\left[(2\,n+1)\,x\,\tilde{p}_n - n\,\tilde{p}_{n-1}\right] + \varepsilon_{n+1} \qquad |\varepsilon_n| \leqslant 3\,\mathbf{u} \text{ (FxP)}$$

$$\delta_n = \tilde{p}_n - p_n$$

$$(n+1)\,\delta_{n+1} = (2\,n+1)\,x\,\delta_n - n\,\delta_{n-1} + (n+1)\,\varepsilon_{n+1}$$

$$(1 - 2\,x\,z + z^2)\,\delta'(z) = z\,(x - z)\,\delta(z) + \varepsilon'(z)$$

$$\delta(z) = \frac{1}{\sqrt{1 - 2\,x\,z + z^2}} \int_0^z \frac{\varepsilon'(z)}{\sqrt{1 - 2\,x\,z + z^2}}$$

$$z \sum_n a_n\,z^n = \sum_n a_{n-1}\,z^n \qquad \frac{d}{dz}\left(\sum_n f_n\,z^n\right) = \sum_n (n+1)\,f_{n+1}\,z^n$$

Majoration de la solution

$$\delta(z) = \frac{1}{\sqrt{1 - 2\,x\,z + z^2}} \int_0^z \frac{\varepsilon'(z)}{\sqrt{1 - 2\,x\,z + z^2}}$$

Majoration de la solution

$$\delta(z) = \frac{1}{\sqrt{1 - 2\,x\,z + z^2}} \int_0^z \frac{\varepsilon'(z)}{\sqrt{1 - 2\,x\,z + z^2}} \qquad \varepsilon(z) \ll \frac{3\,\mathbf{u}}{1 - z}$$

$$\frac{3\,\mathbf{u}}{(1 - z)^2}$$

Majoration de la solution

$$\delta(z) = \frac{1}{\sqrt{1 - 2\,x\,z + z^2}} \int_0^z \frac{\varepsilon'(z)}{\sqrt{1 - 2\,x\,z + z^2}} \qquad \varepsilon(z) \ll \frac{3\,\mathbf{u}}{1 - z}$$

$$\frac{3\,\mathbf{u}}{(1 - z)^2}$$

- $\dfrac{1}{\sqrt{1 - 2\,x\,z + z^2}} = \dfrac{1}{\sqrt{(1 - e^{i\theta}\,z)\,(1 - e^{-i\theta}\,z)}}$ $\qquad x = \cos\theta$
- or $\dfrac{1}{\sqrt{1 - e^{\pm i\theta}\,z}} \ll \dfrac{1}{\sqrt{1 - z}}$ $\qquad$ par composition de $\frac{1}{\sqrt{1 - z}} \in \mathbb{R}_+[[z]]$ et $e^{i\theta}\,z$
- donc $\dfrac{1}{\sqrt{1 - 2\,x\,z + z^2}} \ll \dfrac{1}{1 - z}$

(En fait $\dfrac{1}{\sqrt{1 - 2\,x\,z + z^2}} = \displaystyle\sum_n P_n(x)\,z^n$, et il est classique que $|P_n(x)| \leqslant 1$.)

Majoration de la solution

$$\delta(z) = \frac{1}{\sqrt{1-2\,x\,z+z^2}} \int_0^z \frac{\varepsilon'(z)}{\sqrt{1-2\,x\,z+z^2}} \qquad\qquad \varepsilon(z) \ll \frac{3\,\mathbf{u}}{1-z}$$

$$\frac{3\,\mathbf{u}}{(1-z)^2}\,\frac{1}{1-z}$$

- $$\frac{1}{\sqrt{1-2\,x\,z+z^2}} = \frac{1}{\sqrt{(1-e^{i\theta}\,z)\,(1-e^{-i\theta}\,z)}} \qquad x = \cos\theta$$
- or $\dfrac{1}{\sqrt{1-e^{\pm i\theta}\,z}} \ll \dfrac{1}{\sqrt{1-z}}$ par composition de $\frac{1}{\sqrt{1-z}} \in \mathbb{R}_+[[z]]$ et $e^{i\theta}\,z$
- donc $\dfrac{1}{\sqrt{1-2\,x\,z+z^2}} \ll \dfrac{1}{1-z}$

(En fait $\frac{1}{\sqrt{1-2\,x\,z+z^2}} = \sum_n P_n(x)\,z^n$, et il est classique que $|P_n(x)| \leqslant 1$.)

Majoration de la solution

$$\delta(z) = \frac{1}{\sqrt{1-2\,x\,z+z^2}} \int_0^z \frac{\varepsilon'(z)}{\sqrt{1-2\,x\,z+z^2}} \qquad\qquad \varepsilon(z) \ll \frac{3\,\mathbf{u}}{1-z}$$

$$\int \frac{3\,\mathbf{u}}{(1-z)^2}\,\frac{1}{1-z}$$

- $\dfrac{1}{\sqrt{1-2\,x\,z+z^2}} = \dfrac{1}{\sqrt{(1-e^{i\theta}\,z)\,(1-e^{-i\theta}\,z)}}$ $\qquad x = \cos\theta$
- or $\dfrac{1}{\sqrt{1-e^{\pm i\theta}\,z}} \ll \dfrac{1}{\sqrt{1-z}}$ $\qquad$ par composition de $\dfrac{1}{\sqrt{1-z}} \in \mathbb{R}_+[[z]]$ et $e^{i\theta}\,z$
- donc $\dfrac{1}{\sqrt{1-2\,x\,z+z^2}} \ll \dfrac{1}{1-z}$

(En fait $\dfrac{1}{\sqrt{1-2\,x\,z+z^2}} = \sum_n P_n(x)\,z^n$, et il est classique que $|P_n(x)| \leqslant 1$.)

Majoration de la solution

$$\delta(z) = \frac{1}{\sqrt{1-2\,x\,z+z^2}} \int_0^z \frac{\varepsilon'(z)}{\sqrt{1-2\,x\,z+z^2}} \qquad\qquad \varepsilon(z) \ll \frac{3\,\mathbf{u}}{1-z}$$

$$\frac{1}{1-z}\int \frac{3\,\mathbf{u}}{(1-z)^2}\,\frac{1}{1-z}$$

- $\dfrac{1}{\sqrt{1-2\,x\,z+z^2}} = \dfrac{1}{\sqrt{(1-e^{i\theta}\,z)\,(1-e^{-i\theta}\,z)}}$ $\qquad x = \cos\theta$
- or $\dfrac{1}{\sqrt{1-e^{\pm i\theta}\,z}} \ll \dfrac{1}{\sqrt{1-z}}$ $\qquad$ par composition de $\frac{1}{\sqrt{1-z}} \in \mathbb{R}_+[[z]]$ et $e^{i\theta}\,z$
- donc $\dfrac{1}{\sqrt{1-2\,x\,z+z^2}} \ll \dfrac{1}{1-z}$

(En fait $\frac{1}{\sqrt{1-2\,x\,z+z^2}} = \sum_n P_n(x)\,z^n$, et il est classique que $|P_n(x)| \leqslant 1$.)

Majoration de la solution

$$\delta(z) = \frac{1}{\sqrt{1-2\,x\,z+z^2}} \int_0^z \frac{\varepsilon'(z)}{\sqrt{1-2\,x\,z+z^2}} \qquad\qquad \varepsilon(z) \ll \frac{3\,\mathbf{u}}{1-z}$$

$$\delta(z) \ll \frac{1}{1-z} \int \frac{3\,\mathbf{u}}{(1-z)^2} \frac{1}{1-z}$$

- $\dfrac{1}{\sqrt{1-2\,x\,z+z^2}} = \dfrac{1}{\sqrt{(1-e^{i\theta}\,z)\,(1-e^{-i\theta}\,z)}}$ $\qquad x = \cos\theta$
- or $\dfrac{1}{\sqrt{1-e^{\pm i\theta}\,z}} \ll \dfrac{1}{\sqrt{1-z}}$ $\qquad$ par composition de $\frac{1}{\sqrt{1-z}} \in \mathbb{R}_+[[z]]$ et $e^{i\theta}\,z$
- donc $\dfrac{1}{\sqrt{1-2\,x\,z+z^2}} \ll \dfrac{1}{1-z}$

(En fait $\frac{1}{\sqrt{1-2\,x\,z+z^2}} = \sum_n P_n(x)\,z^n$, et il est classique que $|P_n(x)| \leqslant 1$.)

Majoration de la solution

$$\delta(z) = \frac{1}{\sqrt{1-2\,x\,z+z^2}} \int_0^z \frac{\varepsilon'(z)}{\sqrt{1-2\,x\,z+z^2}} \qquad\qquad \varepsilon(z) \ll \frac{3\,\mathbf{u}}{1-z}$$

$$\delta(z) \ll \frac{1}{1-z} \int \frac{3\,\mathbf{u}}{(1-z)^2} \frac{1}{1-z} = \frac{3}{2} \frac{1}{(1-z)^3} \mathbf{u}$$

- $\dfrac{1}{\sqrt{1-2\,x\,z+z^2}} = \dfrac{1}{\sqrt{(1-e^{i\theta}\,z)\,(1-e^{-i\theta}\,z)}}$ $\qquad x = \cos\theta$
- or $\dfrac{1}{\sqrt{1-e^{\pm i\theta}\,z}} \ll \dfrac{1}{\sqrt{1-z}}$ $\qquad$ par composition de $\frac{1}{\sqrt{1-z}} \in \mathbb{R}_+[[z]]$ et $e^{i\theta}\,z$
- donc $\dfrac{1}{\sqrt{1-2\,x\,z+z^2}} \ll \dfrac{1}{1-z}$

(En fait $\frac{1}{\sqrt{1-2\,x\,z+z^2}} = \sum_n P_n(x)\,z^n$, et il est classique que $|P_n(x)| \leqslant 1$.)

Source des images

- Icône « travaux en cours » : `https://commons.wikimedia.org/wiki/File:Work_in_progress_icon.svg`

  Copyright © Sławek Borewicz, diffusé sous les conditions ci-dessous :

- Les autres icônes utilisées dans ce document sont issues du jeu d'icônes Oxygen (`http://www.kde.org/`). Elles peuvent être reproduites sous les conditions de la licence GNU LGPLv3 (`http://www.gnu.org/copyleft/lesser.html`). Voir `https://techbase.kde.org/Projects/Oxygen/`.
- Image « Generating series oil » adaptée de « St. Jacobs Oil », « *Advertising found printed on the end papers of an edition of "Ivanhoe" published by "International Book Company" of New York. The date 1892 is on the cover.* »

  Photo par Tim & Selena Middleton sous licence Creative Commons CC BY 2.0.

  `https://www.flickr.com/photos/tim_and_selena/5052201329/`