Effect van de argumentreductie op de foutenanlalyse van $\arctan$

Johan Vervloet

18 maart 2003

Foutenanalyse

relatieve fout. Een functie $f(x)$ wordt benaderd door $\tilde{f}(x)$ zodat

$$\tilde{f}(x) = f(x)(1+\varepsilon) \tag{1}$$

afschatting. Aan de hand van een aantal parameters (werkprecisie, convergent, argument) kan de fout ε afgeschat worden.

$$|\varepsilon| \leq \delta(p_1, \ldots, p_n) \tag{2}$$

begrenzing. De bedoeling is dat $|\varepsilon|$ kleiner blijft dan een bovengrens $\overline{\varepsilon}$ (typisch 1 ULP in de doelprecisie). Hiertoe kiezen we $p_1, \ldots p_n$ zodanig dat

$$|\varepsilon| \leq \delta(p_1, \ldots, p_n) \leq \overline{\varepsilon} \tag{3}$$

Argumentreductie

in theorie : $f(z) = p_z \circ g \circ a(z)$

$a(x)$	reductie van het argument
$g(x)$	berekening zelf
$p_z(x)$	compensatie voor a

in praktijk : $\tilde{f}(z) = \tilde{p}_z \circ \tilde{g} \circ \tilde{a}(z)$

Propagatie van afrondingsfouten

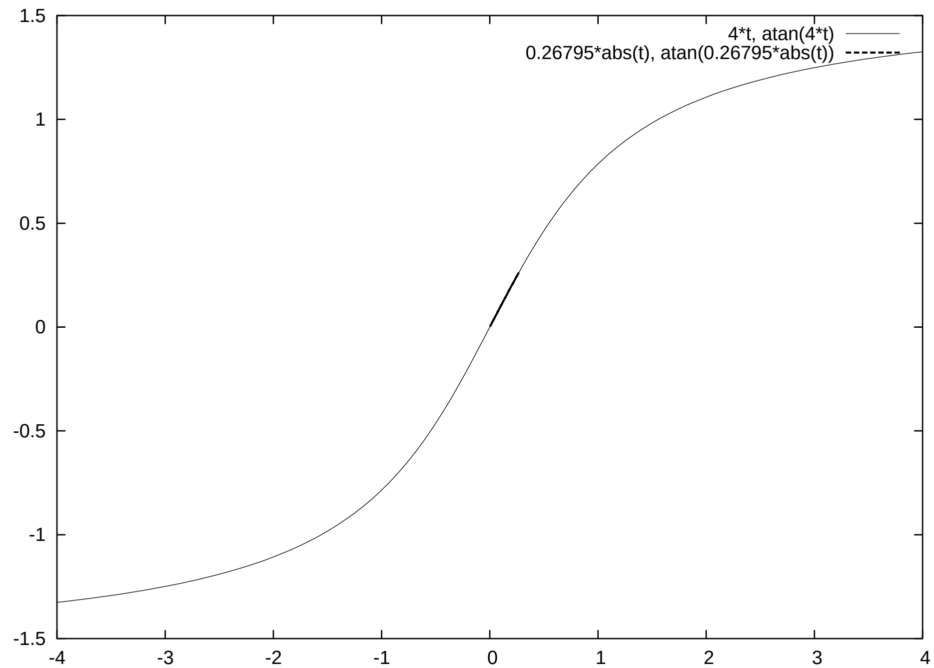
Als de functie $f : \mathbb{R} \to \mathbb{R}$ afleidbaar is op $[x(1 - |\varepsilon_1|), x(1 + |\varepsilon_1|)]$ met $|\varepsilon_1| \leq 1$, dan geldt dat

$$f(x(1 + \varepsilon_1)) = f(x)(1 + \varepsilon_2) \tag{4}$$

waarbij

$$|\varepsilon_2| \leq \delta_f(x, \varepsilon_2) = \max_{|\eta| \leq |\varepsilon_1|} \left| \frac{f'(x(1 + \eta))x\varepsilon_1}{f(x)} \right| \tag{5}$$

De arctan **functie**



Argumentreductie voor arctan

$$\arctan(z) = \begin{cases} \arctan(z) & \text{voor } 0 \le z < 2-\sqrt{3} \\ \frac{\pi}{6} + \arctan\left(\frac{z\sqrt{3}-1}{\sqrt{3}+z}\right) & \text{voor } 2-\sqrt{3} \le z < 1 \\ \frac{\pi}{3} - \arctan\left(\frac{\sqrt{3}-z}{z\sqrt{3}+1}\right) & \text{voor } 1 \le z < 2+\sqrt{3} \\ \frac{\pi}{2} - \arctan\left(\frac{1}{z}\right) & \text{voor } 1 \le z \end{cases} \tag{6}$$

arctan **in** a, g, p_z

$$a(x) = \begin{cases} x & \text{als } 0 \leq x < 2-\sqrt{3} \\ \frac{x\sqrt{3}-1}{\sqrt{3}+x} & \text{als } 2-\sqrt{3} \leq x < 1 \\ \frac{\sqrt{3}-x}{x\sqrt{3}+1} & \text{als } 1 \leq x < 2+\sqrt{3} \\ \frac{1}{x} & \text{als } 2+\sqrt{3} \leq x \end{cases} \tag{7}$$

$$g(x) = \arctan(x) \tag{8}$$

$$p_z(x) = \begin{cases} x & \text{als } 0 \leq z < 2-\sqrt{3} \\ x + \frac{\pi}{6} & \text{als } 2-\sqrt{3} \leq z < 1 \\ \frac{\pi}{3} - x & \text{als } 1 \leq z < 2+\sqrt{3} \\ \frac{\pi}{2} - x & \text{als } 2+\sqrt{3} \leq z \end{cases} \tag{9}$$

Propagatie van de fouten bij arctan

Rekening houdend met dom $g(x) = \left[0, 2-\sqrt{3}\right[$, krijgen we

$$g(x(1+\delta)) = g(x)(1+\varepsilon) \qquad \text{waarbij } |\varepsilon| < 2|\delta| \tag{10}$$

$$p_z(x(1+\delta)) = p_z(x)(1+\varepsilon)$$

$$\text{waarbij} \begin{cases} |\varepsilon| \leq |\delta| & \text{als } 0 \leq z < 2-\sqrt{3} \\ |\varepsilon| \leq \frac{|\delta|}{3} & \text{als } 2-\sqrt{3} \leq z < 1 \\ |\varepsilon| \leq \frac{|\delta|}{5} & \text{als } 1 \leq z < 2+\sqrt{3} \\ |\varepsilon| \leq \frac{|\delta|}{7} & \text{als } 2+\sqrt{3} \leq z \end{cases} \tag{11}$$

Totale fout bij de evaluatie

$$(p \circ g)(a(x)(1+\delta_a)) = (p \circ g \circ a)(x)(1+\varepsilon_2+\delta_p+\varepsilon_2\delta_p) \tag{12}$$

Waarbij δ_a, δ_g en δ_p de relatieve fouten zijn op de functies a, g en p. Verder is

$$|\varepsilon_2| \leq |\varepsilon_1+\delta_g+\varepsilon_1\delta_g| \tag{13}$$

$$|\varepsilon_1| \leq 2|\delta_a| \tag{14}$$

Als $\delta \leq \frac{1}{4}$, dan zal hieruit volgen dat

$$(p \circ g)(a(x)(1+\delta_a)) \leq (p \circ g \circ a)(x)(1+\varepsilon) \tag{15}$$

met $|\varepsilon| \leq 3|\delta|$

Overzicht

Als de maximale relatieve fout $\overline{\varepsilon}$ gegeven is, dan zorgen we ervoor dat

1. $\delta_a \leq \frac{\overline{\varepsilon}}{6}$ (kiezen van goede werkprecisie)
2. $\delta_g \leq \frac{\overline{\varepsilon}}{6}$ (kiezen van goede werkprecisie en goede convergent)
3. $\delta_p \leq \frac{\overline{\varepsilon}}{6}$ (kiezen van goede werkprecisie)

Niet te schatten

Onderstel dat $z \in \left[2-\sqrt{3}, 1\right[$. Er geldt dat

$$a(z) = \frac{z\sqrt{3}-1}{\sqrt{3}+z} \tag{16}$$

Een floating-pointevaluatie van de teller levert

$$z\sqrt{3}(1+\delta_1) - 1 = \left(z\sqrt{3}-1\right)\left(1 + \frac{z\sqrt{3}\delta_1}{z\sqrt{3}-1}\right) \tag{17}$$

waarbij δ_1 de relatieve fout op de berekening van $z\sqrt{3}$ is.

Gered door $p_z(x)$

$$p_z(x) = x + \frac{\pi}{6} \tag{18}$$

$$p_z(x(1+\delta)) = \left(x + \frac{\pi}{6}\right)\left(1 + \frac{\delta x}{x + \frac{\pi}{6}}\right) \tag{19}$$

$$x = \arctan\left(\frac{z\sqrt{3}-1}{\sqrt{3}+z}\right)\left(1 + \frac{z\sqrt{3}\delta_1}{z\sqrt{3}-1}\right)(1+\delta_2) \tag{20}$$

$$\varepsilon = \frac{z\sqrt{3}\delta_1}{z\sqrt{3}-1+\frac{\pi}{6}} \tag{21}$$

$$\lim_{z \to \frac{\sqrt{3}}{3}} \frac{z\sqrt{3}\delta_1}{z\sqrt{3}-1+\frac{\pi}{6}} = \frac{6}{\pi}\delta_1 \tag{22}$$