

Samenvatting Numerieke Wiskunde

Tom Sydney Kerckhove

Started: 20 februari 2014

Compiled: 1 maart 2014

1 Foutenanalyse

Definitie 1. *Absolute fout Δx .*

$$\Delta x = \bar{x} - x$$

Definitie 2. *Relatieve fout δx .*

$$\delta x = \frac{\bar{x} - x}{x}$$

Definitie 3. *Klassieke voorstelling.*

$$x = \sum_{i=m}^n c_i r^i$$

Grondtal (radix) r , Getallen voor de komma n , Getallen na de komma m .

Definitie 4. *Bewegende kommavoorstelling.*

$$x = yb^e$$

Mantisse m , Basis b , Exponent e . (b^e is de schaalfactor)

Definitie 5. *Exacte waarde van x : x .*

$$x = \sum_{i=m}^n c_i b^i$$

Mantisse m , Basis b , Exponent e .

Definitie 6. *Benadering voor x : \bar{x} .*

$$x = \sum_{i=m}^n \bar{c}_i b^i$$

Mantisse m , Basis b , Exponent e .

Definitie 7. *Juist cijfer c_i :*

$$|\bar{x} - x| \leq \frac{1}{2}b^i$$

Definitie 8. *Verband tussen absolute fout en aantal juiste cijfers na de komma.*

$$\frac{1}{2}b^{-p-1} < |\bar{x} - x| \leq \frac{1}{2}b^{-p}$$

Aantal juiste cijfers na de komma p .

Definitie 9. Verband tussen relatieve fout en aantal juiste cijfers na de komma.

$$\frac{1}{2}b^{j-k-1} < \frac{|\bar{x} - x|}{\bar{x}} \leq \frac{1}{2}b^{j-k+1} \quad \text{of} \quad \frac{1}{2}b^{-q-1} < \frac{|\bar{x} - x|}{\bar{x}} \leq \frac{1}{2}b^{-q+1}$$

Positie van het eerste beduidende cijfer k , Positie van het laatste beduidende cijfer $j + 1$, Aantal juiste beduidende cijfers $q = k - j$.

Definitie 10. Machineprecisie ϵ_{mach} .

$$\epsilon_{mach} = \frac{1}{2}b^{1-p}$$

Definitie 11. Elementaire bewerking τ .

$$fl(x\tau y) = (x\tau y)(1 + \eta)$$

($|\eta| \leq \epsilon_{mach}$ en $x, y \in O_{real}$)

Definitie 12. Absolute fout op een som.

$$\Delta \left(\sum_{i=1}^n x_i \right) = \sum_{i=1}^n \Delta x_i$$

Bovengrens:

$$\left| \Delta \left(\sum_{i=1}^n x_i \right) \right|_{max} = n\epsilon_+$$

Relatieve fout op een som.

$$\delta_s = \frac{\Delta x + \Delta y}{x + y}$$

Definitie 13. Zij $\bar{x} = x(1 + \delta x)$ en $\bar{y} = y(1 + \delta y)$.
Absolute fout op een vermenigvuldiging.

$$\Delta p = y\Delta x + x\Delta y$$

Relatieve fout op een vermenigvuldiging.

$$\Delta xy = xy(\delta x + \delta y + \delta x \delta y) \approx \delta x + \delta y$$

Bovengrens:

$$|\delta(xy)| \leq 2\epsilon.$$

Definitie 14. Absolute fout op differentieerbare functies.

$$\Delta f(x) = f(\bar{x}) - f(x) = \Delta x f'(x') \approx f'(\bar{x}) \Delta x \text{ met } x' \text{ tussen } x \text{ en } \bar{x}$$

Bovengrens:

$$|\Delta f(x)|_{\max} \approx |\Delta x|_{\max} \max_t |f'(t)|$$

$$|\Delta f(x_1, \dots, x_n)|_{\max} \approx \sum_{i=1}^n |\Delta x_i|_{\max} \max_{t_1, \dots, t_n} |f'_i(t_1, \dots, t_n)|$$

Definitie 15. Norm $\|\cdot\|$.

De norm van een mathematisch object geeft de ‘grootte’ ervan aan.

Definitie 16. Gegevens en resultaten.

Exact gegeven g .

Gewijzigd gegeven \bar{g} .

Absoluut verschil in gegevens Δg .

$$\Delta g = \bar{g} - g$$

Relatief verschil in gegevens δg .

$$\delta g = \frac{\bar{g} - g}{\|g\|} = \frac{\Delta g}{\|g\|}$$

Exact resultaat r .

$$r = F(g)$$

Berekend resultaat \bar{r}

$$\bar{r} = F(\bar{g})$$

Absoluut verschil in resultaat Δr

$$\Delta r = F(\bar{g}) - r = \bar{r} - r$$

Relatief verschil in resultaat δr

$$\delta r = \frac{F(\bar{g}) - r}{\|r\|} = \frac{\Delta r}{\|r\|}$$

Definitie 17. Conditie van een probleem.

Absoluut Conditiegetal k_A .

$$k_A = \lim_{\epsilon \rightarrow 0} \sup_{\|\Delta g\| \leq \epsilon} \frac{\|\Delta r\|}{\|\Delta g\|}$$

Relatief Conditiegetal k_R

$$k_R = \lim_{\epsilon \rightarrow 0} \sup_{\|\Delta g\| \leq \epsilon} \frac{\|r\|}{\|g\|}$$

Het conditiegetal van een probleem geeft aan hoeveel een fout op de gegevens wordt opgeblazen in de resultaten.

Voorbeeld: Als het relatieve conditiegetal van een probleem F a is, dan worden fouten op de gegevens met een factor a opgeblazen.

Definitie 18. *Benaderingen.*

De exacte berekening van een resultaat F .

$$r = F(g)$$

De exacte berekening van het resultaat gebaseerd op een gewijzigd gegeven \bar{g} .

$$\bar{r} = F(\bar{g})$$

Eindige discretisatie van F , exact berekend: \tilde{F} . Discretisatiefout $err_{dis} = \tilde{F} - F$.

$$\tilde{F} = F + err_{dis}$$

$$\tilde{r} = \tilde{F}(\bar{g})$$

Implementatie van \tilde{F} in eindige precisie.

$$\tilde{\tilde{r}} = \bar{F}(\bar{g})$$

De waarde van de gegevens als $\tilde{\tilde{r}}$ een exacte berekening zou zijn: $\bar{\bar{g}}$.

$$\bar{F}(\bar{\bar{g}}) = \tilde{\tilde{r}} = \tilde{F}(\bar{\bar{g}})$$

Definitie 19. *Stabiliteit van een methode*

- *Voorwaartse/Sterke Stabiliteit.*

Absolute voorwaartse stabiliteit: De absolute grootte van het verschil tussen het berekende resultaat van de methode toegepast op een gewijzigd resultaat \bar{g} in eindige precisie: $\bar{F}(\bar{g})$ en het exacte resultaat van de methode toegepast op een gewijzigd resultaat \bar{g} : $\tilde{F}(\bar{g})$.

$$\|\bar{F}(\bar{g}) - \tilde{F}(\bar{g})\|$$

Relatieve voorwaartse stabiliteit: De relatieve grootte van het verschil tussen het berekende resultaat van de methode toegepast op een gewijzigd resultaat \bar{g} in eindige precisie: $\bar{F}(\bar{g})$ en het exacte resultaat van de methode toegepast op een gewijzigd resultaat \tilde{g} : $\tilde{F}(\tilde{g})$.

$$\frac{\|\bar{F}(\bar{g}) - \tilde{F}(\tilde{g})\|}{\|\tilde{F}(\tilde{g})\|}$$

Kleine waarde(n) \Rightarrow Voorwaarts stabiele methode \tilde{F} .

- *Achterwaartse Stabiliteit.*

Absolute achterwaartse stabiliteit: De grootte van het verschil tussen de gegevens zoals ze zouden zijn als \tilde{r} een exact resultaat was $\bar{\bar{g}}$ en de echte (gewijzigde) gegevens \bar{g} .

$$\|\bar{\bar{g}} - \bar{g}\|$$

Relatieve achterwaartse stabiliteit: De relatieve grootte van het verschil tussen de gegevens zoals ze zouden zijn als \tilde{r} een exact resultaat was $\bar{\bar{g}}$ en de echte (gewijzigde) gegevens \bar{g} .

$$\frac{\|\bar{\bar{g}} - \bar{g}\|}{\|\bar{g}\|}$$

Relatieve waarde is $O(\epsilon_{mach}) \Rightarrow$ Achterwaarts stabiele methode.

- *Zwakke Stabiliteit.*

Het de grootte van het verschil tussen $\bar{\bar{r}}$ en \tilde{r} . ten opzichte van de grootte van het verschil tussen $\tilde{\bar{r}}$ en \tilde{r} : S .

$$S = \frac{\|\bar{\bar{r}} - \tilde{r}\|}{\|\tilde{\bar{r}} - \tilde{r}\|}$$

$S \approx 1 \Rightarrow$ zwak stabiele methode.