

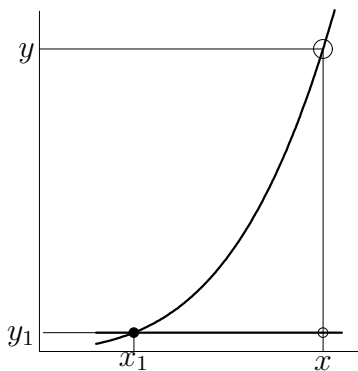
3.18 Inleiding

Het komt vaak voor dat we slechts gedeeltelijke informatie hebben over het vloeiende verloop van een functie f en toch de waarde van de functie $y = f(x)$ in een bepaald punt x willen weten. We gebruiken dan een tabel met tabelpunten van de functie $y = f(x)$. Naarmate de tabel meer punten bevat in de buurt van het gevraagde punt, kunnen we meer punten gebruiken waarmee we kunnen interpoleren naar het gevraagde punt. We behandelen de systematische 'Newton-Gauss' interpolatiemethode, die is genoemd naar de twee beroemde wis-en natuurkundigen Newton en Gauss.

3.18.1 Één tabelpunt: nuldegraads interpolatie

Stel dat alleen het punt $P_1 = (x_1, y_1)$ bekend is, en dat we het punt $P = (x, y)$ voor een bepaalde x willen benaderen met het punt $P_{b1} = (x, y_{b1})$. We kunnen niets anders doen dan de functiewaarde te benaderen met de waarde y_1 van de bekende positie:

$$y_{b1} = y_1$$



De grafiek van de functie wordt zo een rechte lijn evenwijdig aan de x-as door het gegeven punt P_1 . De functie van die rechte lijn is een constante, namelijk $y = y_1$, een 'nuldegraads' veelterm. We noemen dit *nuldegraads interpolatie*.

Bij het vervangen van het punt door de nuldegraads benadering wordt een *fout* ('error') e_{b1} in de functiewaarde gemaakt:

$$y = y_{b1} + e_{b1}$$

$$e_{b1} = y - y_{b1}$$

Deze fout is 0 als $x = x_1$, of $x - x_1 = 0$. De fout kan, volgens de fundamenteelstelling van de algebra over nulpunten, worden geschreven als product:

$$e_{b1} = (x - x_1)D_{b1}$$

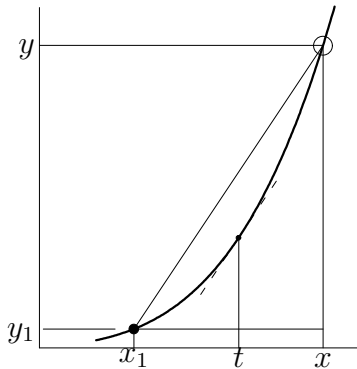
De factor D_{b1} is op te vatten als functie $D(P_1, P)$, afhankelijk van P_1 en P , dus van x, x_1 en y_1 . Na invullen van de definities blijkt:

$$D(P, P_1) = \frac{y - y_1}{x - x_1}$$

$$D_{b1} = \frac{Dy}{Dx} = D(P, P_1) = D(P_1, P)$$

Feitelijk is het differentiequotient $D(P_1, P)$ de *numeriek eerste afgeleide* van de kromme tussen de punten P_1 en P . Belangrijk is dat de volgorde der punten er kennelijk niet toe doet.

De numeriek eerste afgeleide kan worden afgeschat met een afgeleide (mits de kromme differentieerbaar is) in een 'tussenvpunt' t van de kromme: dat punt waar de raaklijn evenwijdig is aan de koorde van P_1 naar P .



$$D_{b1} = y'(t) = y'_t \quad t \approx x_m = (x_1 + x)/2$$

Het tussenvpunt ligt ongeveer midden tussen x en x_1 . Om t te bepalen ontwikkelen we $y - y_1$ rond 'de' nader te bepalen t . Noem $x - t = u$, dan is:

$$\begin{aligned} y - y_1 &= 0 + (u - u_1)y'_t + (u^2 - u_1^2)y''_t/2 + (u - u_1)O(y''') \\ &= u_1y'_t + u_1(u + u_1)y''_t/2 + (u - u_1)O(y''') \\ e_{b1} &= y'(t) + (x + x_1 - 2t)y''_t/2 + O(y''') \end{aligned}$$

Met de keuze in het midden $t = x_m = (x + x_1)/2$ verdwijnt de term met de tweede afgeleide, en verkrijgen we de 'best mogelijke benadering' van het tussenvpunt t : de middenpositie!

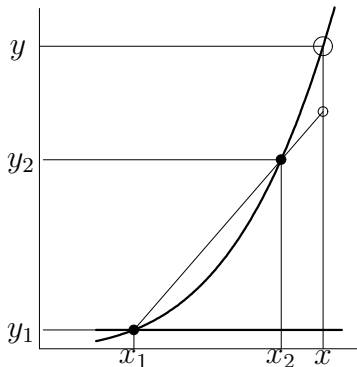
Samenvatting

Voor nuldegraads Gauss interpolatie met één punt P_1 geldt:

$$\begin{aligned} y &= y_{b1} + e_{b1} \\ (1) \quad y_{b1} &= y_1 \quad e_{b1} = (x - x_1)D(P_1, P) \\ D(P_1, P) &= \frac{y - y_1}{x - x_1} = y'(x_m) + O(y''') \end{aligned}$$

3.19 Twee tabelpunten: eerstegraads lineaire interpolatie

Voor een betere benadering betrekken we een tweede tabelpunt bij de benadering. Met het tweede punt $P_2 = (x_2, y_2)$ wordt de nuldegraads fout e_{b1} geschat door in de functie $D(P_1, P)$ het variabele punt $P = (x, y)$ te vervangen door het vaste punt $P_2 = (x_2, y_2)$. Zodoende verkrijgen we een *eerstegraads benadering* y_{b2} (lineair in x):



$$\begin{aligned} y_{b2} &= y_{b1} + e_{b1}(P_2) \\ &= y_1 + (x - x_1)D(P_1, P_2) \end{aligned}$$

De benadering van de kromme is een rechte lijn door de twee posities van de kromme P_1 en P_2 ; bijvoorbeeld met steunpunt P_1 en richtingsvector $P_2 - P_1$. Voor de fout e_{b2} , die gemaakt wordt als deze eersteorde benadering met twee punten wordt gebruikt, geldt uiteraard

$$\begin{aligned} y &= y_{b2} + e_{b2} = y_{b1} + e_{b1} \\ e_{b2} &= e_{b1}(P) - e_{b1}(P_2) \end{aligned}$$

Deze fout is niet alleen 0 bij het eerste punt, als $x = x_1$ is $e_{b1} = 0$, maar ook in het tweede punt, als $x = x_2$. Dus kan de fout worden geschreven als product

$$e_{b2} = (x - x_1)(x - x_2)D_{b2}$$

waarbij D_{b2} een functie $D(P_1, P_2, P)$ is, afhankelijk van twee tabelpunten $P_1 = (x_1, y_1)$, $P_2 = (x_2, y_2)$ en het punt van de kromme $P = (x, y)$. Door invullen van e_{b1} en e_{b2} blijkt:

$$D(P_1, P_2, P) = \frac{D(P_1, P) - D(P_1, P_2)}{x - x_2}$$

$$D_{b2} = D(P_1, P_2, P) = D(P_1, P, P_2) = D(P_2, P_1, P)$$

Feitelijk is de tweedeorde differentiequotient $D(P_1, P_2, P)$ de numeriek tweedeaafgeleide van de kromme van deze drie punten (de volgorde doet er niet toe). De numeriek tweedeaafgeleide $D(P_1, P_2, P)$ is alleen dan nul als òf de drie punten op een lijn liggen, òf twee punten samenvallen. De numeriek tweedeaafgeleide kan worden afgeschat met een tweedeaafgeleide in een 'tussenvpunt' t van de kromme: dit keer het punt waar de raaklijn van de numeriek eersteafgeleide evenwijdig is aan de koorde van P_2 naar P .

$$D(P_1, P_2, P) = y_t''/2 \quad t \approx (x_1 + x_2 + x)/3$$

Ook nu rechtvaardigt de reeksontwikkeling rond t de keuze voor het 'middenpunt' ('zwaartepunt'):

$$D(P_1, P) = y_t' + (u + u_0)y_t''/2 + (u^2 + uu_1 + u_1^2)y_t'''/6 + \dots$$

Na aftrekken van $D(P_1, P_2)$ en delen door $t - x_2 = u - u_2$ volgt

$$D(P_1, P_2, P) = y_t''/2 + (u + u_1 + u_0)y_t'''/6 + \dots$$

Door de keuze $u_1 + u_2 + u = 0$ of $t = (x_1 + x_2 + x)/3$ verdwijnt de afhankelijkheid van de derdeafgeleide.

Samenvatting Voor eerstegraads interpolatie met twee tabelpunten P_1 en P_2 en het punt van de kromme P geldt:

$$y = y_{b2} + e_{b2}$$

$$y_{b2} = y_1 + (x - x_1)D(P_1, P_2)$$

$$e_{b2} = (x - x_1)(x - x_2)D(P_1, P_2, P)$$

$$(2) \quad D(P_1, P_2, P) = \frac{D(P_1, P) - D(P_1, P_2)}{x - x_2}$$

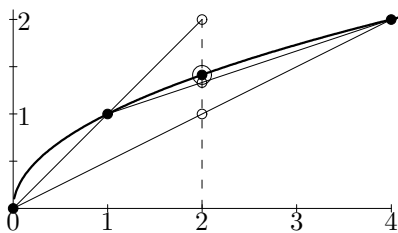
$$= y''(t)/2 + O(y^{(iv)})$$

Voorbeeld Schat de waarde van $\sqrt{2}$ door lineaire interpolatie, met behulp van punten met heeltallige wortel. De tabelpunten zijn:

x	0	1	4
y	0	1	2

Oplossing Voor het schatten van $\sqrt{2}$ gebruiken we de twee dichtstbijzijnde punten. Het dichtst bij $x = 2$ is $x = 1$, op afstand 1. Dan kunnen we kiezen uit twee punten op afstand 2: $x = 0$ en $x = 4$. Om de beste te kunnen kiezen schatten we de fout af met $e_{b2} = (x - x_1)(x - x_2)D_{b2}$. Daartoe maken we eerst een differentiequotienten tabel tot de tweedeafgeleide :

x	y	$D_{b1} \approx y'$	$D_{b2} \approx y''/2$
0	0	$\frac{1-0}{1-0} = 1$	$\frac{1/3-1}{4-0} = -1/6$
1	1	$\frac{2-1}{4-1} = 1/3$	
4	2		



Lineaire interpolatie $y = y_1 + (x - x_1)D(P_1, P_2)$ met $P_2 = (0, 0)$ geeft een te hoge waarde: $y_{b2} = 1 + (2 - 1)1 = 2$, en met $P_2 = (4, 2)$ een te lage waarde: $y_{b2} = 1 + (2 - 4)(1/3) = 1\frac{1}{3}$. De fout-schatting met $x_2 = 0$ is: $e_{b2} = (2 - 1)(2 - 0)(-1/6) = -1/3$. De foutschatting met $x_2 = 4$ is: $e_{b2} = (2 - 1)(2 - 4)(-1/6) = 1/3$. Toevallig (?) is de absolute fout gelijk, alleen het teken verschilt. Als we de interpolatiewaarde verhogen met de foutschatting krijgen we de beste schatting, die in beide gevallen de waarde $\sqrt{2} \approx 1\frac{2}{3} = 1,7$ heeft. Dat beide benaderingen gelijk zijn geworden is geen toeval, omdat precies dezelfde punten zijn gebruikt voor de benadering maar in verschillende volgorde.

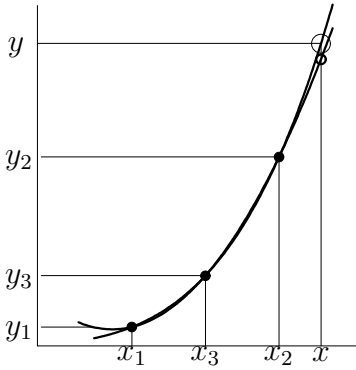
Vergeleken met de bekende wortelwaarde $\sqrt{2} = 1,41$ is de 'inter'polatie waarde $y_{b2} = 1\frac{1}{3} = 1,33$, gevonden met $x = 1$ en $x = 4$, slechts $e_{b2} = 1,41 - 1,33 = 0,08$ fout ('inter' betekent dat 2 binnen het interval $[1,4]$ ligt). De 'extra'polatie waarde $y_{b2} = 2$, gevonden met $x = 1$ en $x = 0$, geeft een veel slechter resultaat: $e_{b2} = 1,41 - 2 = -0,59$ ('extra' betekent dat 2 buiten het interval $[0,1]$ ligt). Algemeen is interpolatie beter dan extrapolatie.

Omdat we de afgeleiden van de wortelfunctie kennen, kunnen we precies aangeven waar het tussenpunt moet zitten. Immers $y = x^{1/2}$, $y' = x^{-1/2}/2$ en $y'' = -x^{-3/2}/4$, zodat de fout $y''(t)/2 = -t^{-3/2}/8$. Bij de extrapolatie met $x_2 = 0$ bleek $e_{b2} = -0,59 = (2 - 1)(2 - 0)(-t^{-3/2}/8) = -t^{-3/2}/4$ waaruit opgelost $t = 0,57$; vergelijk dat met het zwaartepunt $(1 + 0 + 2)/3 = 1$: een behoorlijk verschil. Bij de interpolatie met $x_2 = 4$ bleek $e_{b2} = 0,08 = (2 - 1)(2 - 4)(-t^{-3/2}/8) = t^{-3/2}/4$, zodat $t = 2,12$; vergeleken met het zwaartepunt $(1 + 4 + 2)/3 = 2,33$ is dat niet slecht!

3.20 Drie tabelpunten: tweedegraads parabool interpolatie

Met een derde punt van de kromme $P_3 = (x_3, y_3)$ wordt op zijn beurt de functie $D(P_1, P_2, P)$ afgeschat door $P = P_3$ tot de tweedegraads (kwadratisch in x) interpolatie y_{b3} :

$$\begin{aligned} y_{b3} &= y_{b2} + e_{b2}(P_3) \\ &= y_{b2}(P_1, P_2) + (x - x_1)(x - x_2)D(P_1, P_2, P_3) \end{aligned}$$



De benadering van de kromme is, mits de drie posities niet op een lijn liggen, een parabool door de drie posities P_1, P_2, P_3 .

Voor de fout e_{b3} , die gemaakt wordt als deze tweedegraads benadering wordt gebruikt, geldt uiteraard

$$\begin{aligned} y &= y_{b3} + e_{b3} = y_{b2} + e_{b2} \\ e_{b3} &= e_{b2}(P) - e_{b2}(P_3) \end{aligned}$$

De fout $e_{b3}(x)$ is nul in P_1, P_2 (waar $e_{b2} = 0$) en P_3 (als $P = P_3$). Dus kan de fout worden geschreven als een product veelterm:

$$e_{b3} = (x - x_1)(x - x_2)(x - x_3)D_{b3},$$

met D_{b3} een functie $D(P_1, P_2, P_3, P)$ van de drie tabelpunten P_1, P_2, P_3 en het punt van de kromme P . Door invullen van e_{b2} blijkt:

$$D(P_1, P_2, P_3, P) = \frac{D(P_1, P_2, P) - D(P_1, P_2, P_3)}{x - x_3}$$

$$D_{b3} = D(P_1, P_2, P_3, P) = D(P_1, P_2, P, P_3) = D(P_1, P, P_3, P_2) + \dots$$

De derdeorde differentiequotient $D(P_1, P_2, P_3, P)$ is feitelijk een numeriek derde afgeleide.

Intermezzo

Bij tweedegraads krommen is het eersteorde differentiequotient lineair, het tweedeorde differentiequotient constant en de tweedegraads interpolatie foutloos.

$$\begin{aligned} y &:= y_1 + (x - x_1)y_1' + (x - x_1)^2 y_1''/2 \\ D(P_1, P) &= (y - y_1)/(x - x_1) = y_1' + (x - x_1)y_1''/2 \\ y_2 &= y_1 + (x_2 - x_1)y_1' + (x_2 - x_1)^2 y_1''/2 \\ y - y_2 &= (x - x_2)y_1' + ((x - x_1)^2 - (x_2 - x_1)^2)y_1''/2 \\ &= (x - x_2)(y_1' + (x - x_1 + x_2 - x_1)y_1''/2) \\ D(P_2, P) &= (y - y_2)/(x - x_2) = y_1' + (x - x_1 + x_2 - x_1)y_1''/2 \\ D(P, P_1, P_2) &= (D(P, P_2) - D(P, P_1))/(x_2 - x_1) = y_1''/2 \\ \Rightarrow D(P, P_1, P_2) &= D(P_1, P_2, P) = D(P_1, P_2, P_3) = y_1''/2 \\ D(P_1, P_2, P_3, P) &= (D(P_1, P_2, P) - D(P_1, P_2, P_3))/(x - x_3) = 0 \\ &\Rightarrow e_{b3} = 0 \end{aligned}$$

In het algemeen kan het derdeorde differentiequotient $D(P_1, P_2, P_3, P)$ worden geschat als een derde afgeleide in een tussenpunt t dat ongeveer het 'zwaartepunt' is van de betrokken punten.

$$D(P_1, P_2, P_3, P) = y_t'''/6 \quad t \approx (x_1 + x_2 + x_3 + x)/4$$

Ook nu rechtvaardigt de reeksontwikkeling rond t met $u = x - t$ de keuze voor het zwaartepunt:

$$D(P_1, P_2, P, P) = y_t''/2 + (u + u_1 + u_2)y_t'''/6 + \\ + (u^2 + uu_1 + u_2^2 + u_1^2 + u_2u_1)y_t''''/24 + \dots$$

Na aftrekken van $D(P_1, P_2, P_3)$ en delen door $x - x_3 = u - u_3$ volgt

$$D(P_1, P_2, P_3, P) = y_t''''/6 + (u_1 + u_2 + u_3 + u)y_t''''/24 + \dots$$

Door de keuze $t = (x_1 + x_2 + x_3 + x)/4$ wordt $u_1 + u_2 + u_3 + u = 0$ en verdwijnt de afhankelijkheid van de vierde afgeleide.

Samenvatting

Voor tweedegraads interpolatie met drie tabelpunten P_1, P_2, P_3 en punt van de kromme P geldt:

$$y = y_{b3} + e_{b3} \\ y_{b3} = y_{b2}(P_1, P_2) + (x - x_1)(x - x_2)D(P_1, P_2, P_3) \\ = y_1 + (x - x_1)(D(P_1, P_2) + (x - x_2)D(P_1, P_2, P_3)) \\ (3) \quad e_{b3} = (x - x_1)(x - x_2)(x - x_3)D(P_1, P_2, P_3, P) \\ D(P_1, P_2, P_3, P) = \frac{D(P_1, P_2, P) - D(P_1, P_2, P_3)}{x - x_3} \\ \approx y'''(t)/6 + O(y^{(v)})$$

3.21 n tabelpunten: $n - 1$ -stegraads interpolatie en algoritme

Meer in het algemeen baseren we een $n - 1$ -stegraads interpolatie y_{bn} met n tabelpunten $P_1 = (x_1, y_1), P_2 = (x_2, y_2), \dots, P_n = (x_n, y_n)$ op de interpolatie met de eerste n tabelpunten. Kies nu in de foutformule van de interpolatie met $n - 1$ punten het punt van de kromme in het laatste tabelpunt $P = P_n$.

$$y = y_{bn} + e_{bn} = y_{bn-1} + e_{bn-1} \\ e_{bn} = e_{bn-1}(P) - e_{bn-1}(P_n)$$

Dan verdwijnt de fout e_{bn} in elk van de $n - 1$ tabelpunten waarvoor $e_{bn-1} = 0$ en in het punt P_n . Voor de fout geldt dus dat er een factor D_{bn} is:

$$e_{bn} = (x - x_1) \cdots (x - x_n)D_{bn}$$

De factor D_{bn} is op te vatten als een functie $D(P_1, \dots, P_n, P)$ van alle tabelpunten en het punt van de kromme P . Door

invullen van e_{bn} en e_{bn-1} blijkt een iteratie vergelijking voor het $n - 1$ ste orde differentiequotient te gelden:

$$D_{bn} = D(P_1, \dots, P_n, P) \\ = \frac{D(P_1, \dots, P_{n-1}, P) - D(P_1, \dots, P_{n-1}, P_n)}{x - x_n}$$

De beginwaarde $D(b1) = y_1$ volgt uit de eenpuntsinterpolatie.

Merk op, dat het differentiequotient D symmetrisch is in het variabele punt P en laatste punt P_n . Maar dat geldt ook voor de voorgaande punten. Ze is dus symmetrisch in alle punten, inclusief het variabele punt P .

Zonder bewijs merken we op, dat het n -de orde differentiequotient van de punten $P_1 \dots P_n$ gelijk is aan de n -degraads afgeleide in een tussenpunt t , dat kan worden benaderd met het zwaartepunt van de deelnemende punten (inclusief P)

$$D(P_1, \dots, P_n, P) = y^{(n)}(t)/n! + O(y^{(n+2)}) \\ t \approx (x_1 + \dots + x_n + x)/(n + 1)$$

Voorbeeld Van een functie is een tabel gegeven:

x	0,0	0,1	0,2	0,3	0,4	0,5	0,6
y	1,0000	1,1052	1,2214	1,3499	1,4918	1,6482	1,8221

Benader met behulp van deze tabel de functiewaarde in $x = 0,33$ zo nauwkeurig mogelijk.

Oplossing De punten die het dichtst bij 0,33 liggen zijn om en om links en rechts van het punt. We zullen een differentiequotienttabel maken uitgaande van de oorspronkelijke volgorde:

P	x	y	D	D^2	D^3	D^4
P_7	0,0	1,0000				
			1,052			
P_5	0,1	1,1052		0,550		
			1,162		0,217	
P_3	0,2	1,2214		0,615		-0,09
			1,285		0,183	
P_1	0,3	1,3499		0,670		+0,21
			1,419		0,267	
P_2	0,4	1,4918		0,750		-0,04
			1,569		0,250	
P_4	0,5	1,6487		0,825		
			1,734			
P_6	0,6	1,8221				

De tabel kan kolom voor kolom worden ingevuld, afhankelijk van de gevraagde nauwkeurigheid. De eerste interpolatie met P_1 is $y_1 = 1,3499$.

Voor het berekenen van de tweepuntsinterpolatie berekenen we eerst de kolom van de numeriek eersteafgeleiden D .

Zo is bijvoorbeeld $D(P_1, P_2) = D(P_2, P_1) = (y_2 - y_1)/(x_2 - x_1) = (1, 4918 - 1, 3499)/(0, 4 - 0, 3) = 1, 419$ en $D(P_1, P_3) = (1, 3499 - 1, 2214)/(0, 3 - 0, 2) = 1, 285$. De tweepuntsinterpolatie geeft $y_{b2} = 1, 3499 + (0, 33 - 0, 30)1, 419 = 1, 3925$.

Voor de driepuntsinterpolatie berekenen we de kolom van de numeriek tweedeaafgeleiden D^2 , bijvoorbeeld $D(P_1, P_2, P_3) = (D(P_1, P_2) - D(P_1, P_3))/(x_2 - x_3) = (1, 419 - 1, 285)/(0, 4 - 0, 2) = 0, 670$, enzovoort.

Achtereenvolgens krijgen we waarden die onder en boven de eindwaarde zitten:

$$\begin{aligned} y_{b1} &= 1, 3499 \\ y_{b2} &= 1, 3499 + (0, 03)1, 419 = 1, 3925 \\ y_{b3} &= 1, 3925 + (0, 03)(-0, 07)0, 670 = 1, 3911 \\ y_{b4} &= 1, 3911 + (0, 03)(-0, 07)(1, 03)0, 267 = 1, 3905 \end{aligned}$$

Merk op, dat de getallen in de kolom van de (numeriek) vierdeafgeleide klein zijn met wisselend teken. Dat duidt er op, dat we in de afrondruis komen. De fout is immers van de orde $(0, 03)(-0, 07)(1, 03)(-1, 07)D(P_1, P_2, P_3, P_4, P) \approx 0, 002D^4$. We nemen deze afgeleide dus niet meer mee, om te voorkomen dat er onbetrouwbare bijdragen in de interpolatie komen.

De *beste schatting* voor $x = 0, 33$ is, vanwege de manier waarop we naar de eindwaarde gaan, het gemiddelde van de laatste twee waarden: $y = (1, 3911 + 1, 3905)/2 = 1, 3908$, met een absolute fout die niet groter is dan de helft van het verschil: $|E| \leq (1, 3911 - 1, 3905)/2 = 0, 0003$.

Het volgende algoritme (als Pascal functie) gebruikt het geheugen veel efficiënter:

```
function interpolatieGauss (n: integer; x,y: array[];
                           xb,E: real): real;
var
  yb,                               { interpolatiewaarde }
  Eb,                               { fout }
  px: real;                          { px= (x-x1)..(x-xn) }
  i: integer;
begin
  read(xb);                          { xb -> x }
  read(E);                            { maximale fout }
  n:= 1; read(x[1],y[1]);
  yb:= y[1]; px:= 1;                  { eerste interpolatie }
  repeat
    n:= n+1; read(x[n],y[n]);
    for i:= 1 to n-1 do
      y[n]:= (y[n]-y[i])/(x[n]-x[i]);  { y -> D }
      px:= px*(xb-x[n-1]); Eb:= px*y[n]; yb:= yb+Eb;
    until abs(Eb) < dy
  interpolatieGauss:= yb
end;
```


C.M.Fortuin
van Deldenpad 8
6862DC Oosterbeek NL

fortuinc@xs4all.nl