Les 2 Taylor reeksen

We hebben in Wiskunde 1 een aantal belangrijke reële functies gezien, bijvoorbeeld de exponentiële functie $\exp(x)$ of de trigonometrische functies $\sin(x)$ en $\cos(x)$. Toen hebben we wel eigenschappen van deze functies aangegeven, bijvoorbeeld dat $\exp(x)$ gekarakteriseerd is door $\exp(x)' = \exp(x)$ en $\exp(0) = 1$. Maar hoe we de waarde van zo'n functie echt kunnen berekenen, of hoe een zakrekenmachine, GRM of computer de waarden van zo'n functie berekent, hebben we toen niet gezien.

De methode die hiervoor (met zekere variaties) wordt gebruikt, is een functie te benaderen door een veelterm. Dit lijkt op het eerste gezicht te eenvoudig om efficiënt te kunnen werken, maar men kan bewijzen dat op een begrensd gebied een continue functie door veeltermen willekeurig goed benaderd kan worden. Naarmate de gewenste nauwkeurigheid toeneemt zijn hiervoor echter veeltermen van hogere graad nodig.

2.1 Interpolatie

Eén mogelijkheid om een benaderende veelterm te vinden, gebruikt de methode van *interpolatie*. We weten dat door 2 punten (x_1, y_1) en (x_2, y_2) met verschillende x-coördinaten $x_1 \neq x_2$ een eenduidige lineaire functie vastgelegd is, met als grafiek de lijn door deze twee punten, te weten

$$l(x) = \frac{y_2 - y_1}{x_2 - x_1}(x - x_1) + y_1.$$

Net zo leggen 3 punten (weer met verschillende x-waarden) in het $x-y$-vlak eenduidig een parabool, dus een kwadratische functie vast. Op een analoge manier laat zich aantonen dat door $n+1$ punten met verschillende x-waarden een eenduidige veelterm van graad n vastgelegd is. De punten x_i waarop de waarden gegeven zijn noemt men ook *basispunten* of *roosterpunten*. De veelterm door de punten laat zich met behulp van de *Lagrange interpolatie* zelfs expliciet aangeven:

Stelling: De eenduidige veelterm $p(x)$ van graad n die door de $n+1$ punten $(x_0, y_0), (x_1, y_1), \ldots, (x_n, y_n)$ met verschillende x-coördinaten x_i loopt, is gegeven door

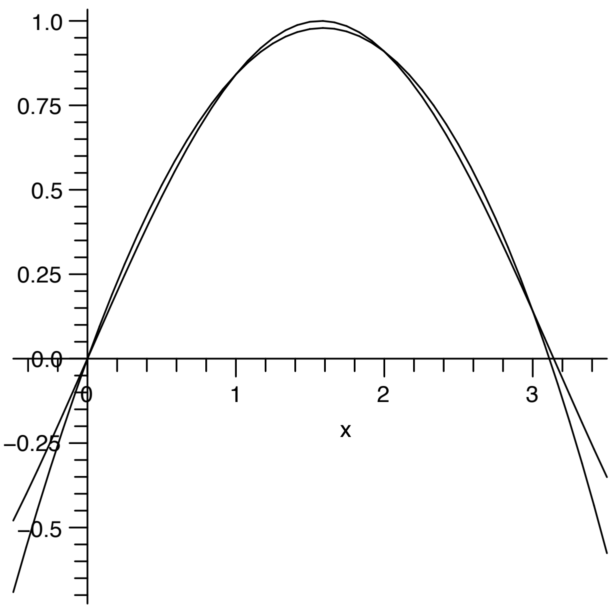
$$p(x) := y_0 \cdot L_0(x) + y_1 \cdot L_1(x) + \ldots + y_n \cdot L_n(x), \text{ waarbij}$$

$$L_k(x) := \frac{x - x_0}{x_k - x_0} \cdot \frac{x - x_1}{x_k - x_1} \cdots \frac{x - x_{k-1}}{x_k - x_{k-1}} \cdot \frac{x - x_{k+1}}{x_k - x_{k+1}} \cdots \frac{x - x_n}{x_k - x_n}.$$

De grap bij deze stelling ligt in het feit dat de hulpfuncties $L_k(x)$ juist zo gemaakt zijn dat $L_k(x_k) = 1$ en $L_k(x_i) = 0$ voor $i \neq k$.

In Figuur I.5 zien we dat de sinus functie goed door een derdegraads interpolatie $p(x)$ benaderd wordt, die door de vier basispunten 0, 1, 2 en 3 loopt en gegeven is door:

$$\begin{aligned}
p(x) &= \frac{\sin(1)}{2}x(x-2)(x-3) - \frac{\sin(2)}{2}x(x-1)(x-3) + \frac{\sin(3)}{6}x(x-1)(x-2) \\
&= (\frac{\sin(1)}{2} - \frac{\sin(2)}{2} + \frac{\sin(3)}{6})\,x^3 + (-\frac{5\sin(1)}{2} + 2\sin(2) - \frac{\sin(3)}{2})\,x^2 \\
&+ (3\sin(1) - \frac{3\sin(2)}{2} + \frac{\sin(3)}{3})\,x \\
&\approx -0.0104\,x^3 - 0.3556\,x^2 + 1.2075\,x.
\end{aligned}$$



Figuur I.5: Benadering van $\sin(x)$ met behulp van de Lagrange interpolatie door de punten met x-waarden 0, 1, 2 en 3.

De interpolatie laat zich op een enigszins voor de hand liggende manier ook op functies van meerdere variabelen veralgemenen. Dit heeft in het bijzonder voor punten in de 3-dimensionale ruimte veel toepassingen. Zo worden bijvoorbeeld voor oppervlakken (zo als gezichten) alleen maar een aantal roosterpunten opgeslagen en de punten ertussen door interpolatie berekend om een glad oppervlak te krijgen. Op deze manier laat zich de beweging van een *computer animated* figuur in de *virtual reality* reconstrueren uit gemeten bewegingen van een echte figuur, waarbij alleen maar op enkele punten sensoren zitten.

Er zijn echter ook een paar problemen als we de interpolatie willen gebruiken om een functie te benaderen:

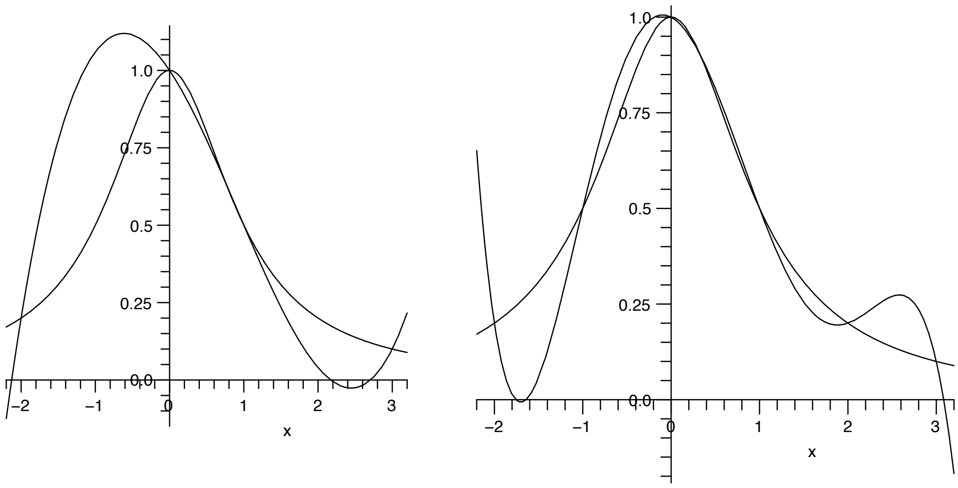
(i) Omdat we een functie $f(x)$ door een veelterm willen benaderen, moeten we de y_k in principe al als functiewaarden $y_k = f(x_k)$ kennen, terwijl we eigenlijk de veelterm juist bepalen om de functiewaarden te kunnen benaderen.

(ii) Afhankelijk van hoe we de x_k kiezen, kan de veelterm $p(x)$ sterk van de functie $f(x)$ afwijkingen, bijvoorbeeld als de x_k te grote afstanden van elkaar hebben.

Het laatste punt is in Figuur I.6 geïllustreerd: We benaderen de functie

$$f(x) := \frac{1}{1+x^2}$$

op het interval $[-2, 3]$ door interpolaties van graad 3 en van graad 5. In het linkerplaatje is de interpolatie met de vier basispunten $-2, 0, 1, 3$ te zien, in het rechterplaatje de interpolatie met de zes basispunten $-2, -1, 0, 1, 2, 3$. In beide gevallen wijkt $f(x)$ sterk van de interpolatie af, alleen maar op het interval $[0, 1.5]$ is de benadering enigszins redelijk.



Figuur I.6: Interpolatie van $\frac{1}{1+x^2}$ met 4 en 6 basispunten.

In principe zouden we op intervallen waar de functie $f(x)$ heel sterk verandert veel basispunten x_k willen hebben, terwijl op stukken waar de functie bijna lineair is niet zo veel basispunten nodig zijn. Om hierover te kunnen beslissen moeten we naar de hogere afgeleiden van $f(x)$ kijken en dit idee geeft aanleiding tot een andere manier om een benaderende veelterm te bepalen. Deze gebruikt inderdaad de hogere afgeleiden van $f(x)$, maar slechts in een enkele punt.

2.2 Taylor veeltermen

Het idee bij de *Taylor veelterm* is, een functie $f(x)$ in de omgeving van één punt x_0 te benaderen. Hiervoor wordt een veelterm geconstrueerd die in het punt x_0

niet alleen maar dezelfde functiewaarde als $f(x)$ heeft, maar ook dezelfde eerste afgeleide, dezelfde tweede afgeleide, enzovoorts.

Definitie: De *Taylor veelterm* $p(x) := p_{f,n,x_0}(x)$ van graad n in het punt x_0 is voor een in x_0 (minstens) n keer differentieerbare functie $f(x)$ vastgelegd door de eigenschappen:

$$p(x_0) = f(x_0), \quad p'(x_0) = f'(x_0), \quad p''(x_0) = f''(x_0), \ \ldots, \ p^{(n)}(x_0) = f^{(n)}(x_0),$$

de Taylor veelterm $p(x)$ heeft in het punt x_0 dus precies dezelfde afleidingen (tot orde n) als $f(x)$.

Men gaat na dat deze eigenschappen inderdaad een eenduidige veelterm van graad n bepalen, namelijk de veelterm

$$p(x) = c_n(x - x_0)^n + c_{n-1}(x - x_0)^{n-1} + \ldots + c_1(x - x_0) + c_0$$

met coëfficiënten

$$c_0 = f(x_0), \quad c_1 = f'(x_0), \quad c_2 = \frac{f''(x_0)}{2!}, \quad \ldots \quad c_n = \frac{f^{(n)}(x_0)}{n!}.$$

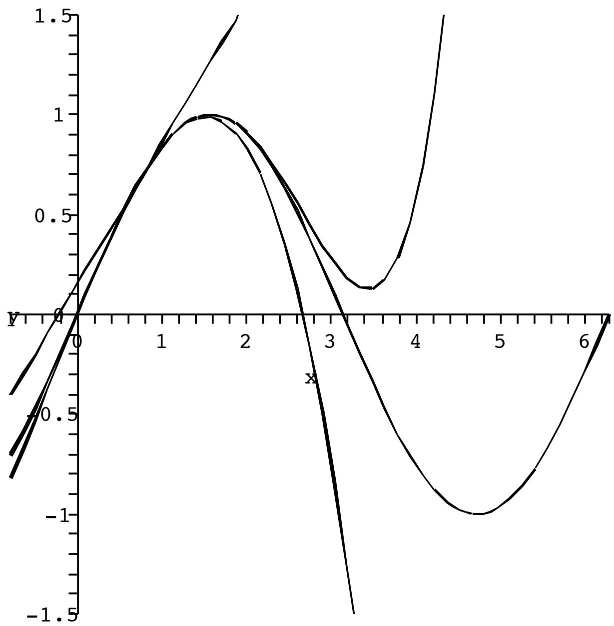
Dat dit inderdaad de coëfficiënten zijn, ziet men door het uitschrijven van de afgeleiden van $p(x)$, invullen van x_0 en vergelijken met $f^{(n)}(x_0)$:

$$\begin{aligned}
p(x) &= c_n(x - x_0)^n + \ldots + c_3(x - x_0)^3 + c_2(x - x_0)^2 + c_1(x - x_0) + c_0 \\
&\Rightarrow p(x_0) = c_0, \text{ uit } f(x_0) = p(x_0) = c_0 \text{ volgt dan } c_0 = f(x_0) \\
p'(x) &= n \cdot c_n(x - x_0)^{n-1} + \ldots + 3 \cdot c_3(x - x_0)^2 + 2 \cdot c_2(x - x_0) + c_1 \\
&\Rightarrow p'(x_0) = c_1, \text{ uit } f'(x_0) = p'(x_0) = c_1 \text{ volgt dan } c_1 = f'(x_0) \\
p''(x) &= n(n-1) \cdot c_n(x - x_0)^{n-2} + \ldots + 3 \cdot 2 \cdot c_3(x - x_0) + 2 \cdot c_2 \\
&\Rightarrow p''(x_0) = 2 \cdot c_2, \text{ uit } f''(x_0) = p''(x_0) = 2 \cdot c_2 \text{ volgt dan } c_2 = \frac{f''(x_0)}{2} \\
p'''(x) &= n(n-1)(n-2) \cdot c_n(x - x_0)^{n-3} + \ldots + 3 \cdot 2 \cdot c_3 \\
&\Rightarrow p'''(x_0) = 3 \cdot 2 \cdot c_3 \Rightarrow c_3 = \frac{f'''(x_0)}{3!} \quad \text{enzovoorts.}
\end{aligned}$$

Het idee achter de geëiste eigenschappen is, dat het overeenstemmen van de hogere afgeleiden ervoor zorgt dat de grafiek van de Taylor veelterm zich rond het punt x_0 steeds beter aan de grafiek van $f(x)$ vlijt.

In Figuur I.7 zien we het effect van Taylor veeltermen van verschillende graad. Terwijl de veelterm van graad 1 de grafiek van $\sin(x)$ alleen maar raakt (omdat $\sin(x)$ geen rechte stukken heeft), is de veelterm van graad 3 al een redelijke benadering tussen ongeveer 0.5 en 1.5, voordat hij naar beneden wegduikt. Met de veelterm van graad 5 gaat het iets langer goed, maar dan loopt deze naar boven weg.

Meestal wordt de Taylor veelterm niet in de boven aangegeven volgorde van de termen geschreven (dus beginnend met de hoogste), maar andersom. Verder wordt het feit dat de Taylor veelterm de functie slechts in een kleine omgeving



Figuur I.7: Benadering van $\sin(x)$ door de Taylor veeltermen van graad 1, 3 en 5 in het punt $x_0 = \frac{\pi}{4}$

van x_0 goed benadert weergegeven door x in de vorm $x = x_0 + h$ te schrijven (waarbij men stiekem verondersteld dat h klein is). Dit geeft de volgende vorm van de Taylor veelterm:

$$p(x_0 + h) = f(x_0) + f'(x_0) \cdot h + \frac{f''(x_0)}{2} \cdot h^2 + \ldots + \frac{f^{(n)}(x_0)}{n!} \cdot h^n.$$

Opdracht 8 *Bepaal voor $f(x) := \sqrt{x}$ de Taylor veelterm van graad 4 in het punt $x_0 = 1$.*

Over het verband tussen de benaderde functie $f(x)$ en de Taylor veelterm $p(x)$ valt echter nog veel meer te zeggen, in het bijzonder laat zich expliciet een afschatting voor de fout $|f(x_0 + h) - p(x_0 + h)|$ aangeven (die natuurlijk van h afhangt). We definiëren hiervoor de n-de *foutterm* $R_n(h)$ door

$$f(x_0 + h) = p(x_0 + h) + R_n(h) \quad \text{of} \quad R_n(h) = f(x_0 + h) - p(x_0 + h).$$

Er zijn verschillende manieren om de foutterm $R_n(h)$ aan te geven, de drie meest belangrijke zijn de volgende:

$$R_n(h) = \frac{1}{(n+1)!} \cdot f^{(n+1)}(t) \cdot h^{n+1} \text{ voor een } t \in [0, h] \qquad \text{(Lagrange vorm)}$$

$$R_n(h) = \frac{1}{n!} \cdot f^{(n+1)}(t) \cdot (h-t)^n h \text{ voor een } t \in [0, h] \qquad \text{(Cauchy vorm)}$$

$$R_n(h) = \frac{1}{n!} \int_0^h f^{(n+1)}(t) \cdot (h-t)^n \, dt \qquad \text{(integraal vorm)}$$

Het bewijs dat de foutterm op deze manieren geschreven kan worden is niet eens zo moeilijk. Bijvoorbeeld berust de Lagrange vorm op de *middelwaardestelling*, die zegt, dat er voor een (differentieerbare) functie $f(x)$ op een interval $[a, b]$ een punt $c \in [a, b]$ in het interval ligt waar de raaklijn aan $f(x)$ dezelfde stijging heeft als de gemiddelde stijging $\frac{f(b)-f(a)}{b-a}$ van $f(x)$ op het interval $[a, b]$, d.w.z. waar geldt dat $f(b) - f(a) = f'(c)(b - a)$. Dit klopt omdat de raaklijn niet overal een grotere of overal een kleinere stijging dan de gemiddelde stijging kan hebben.

Vooral de Lagrange vorm van de foutterm is vaak geschikt om de fout bij de benadering door de Taylor veelterm af te schatten. Dit lukt in het bijzonder als zich makkelijk een grens voor de $(n + 1)$-de afgeleide van $f(x)$ op het interval $[0, h]$ laat aangeven.

Bijvoorbeeld geldt voor $f(x) = \sin(x)$ dat $|f^{(n+1)}(t)| \leq 1$ voor alle t, omdat de afgeleiden $\pm \sin(t)$ of $\pm \cos(t)$ zijn en deze functies alleen maar waarden tussen -1 en 1 hebben.

Net zo goed ziet men in dat voor $f(x) = \exp(x)$ voor $t \in [0, h]$ (met $h > 0$) geldt dat $|f^{(n+1)}(t)| \leq e^h$ omdat de afgeleiden de exponentiële functie reproduceren en de exponentiële functie monotoon stijgend is.

Voorbeelden:

(1) Voor $f(x) := \sin(x)$ en $x_0 = 0$ hebben we $f(x_0) = \sin(0) = 0$, $f'(x_0) = \cos(0) = 1$, $f''(x_0) = -\sin(0) = 0$, $f'''(x_0) = -\cos(0) = -1$ en $f^{(4)}(x_0) = \sin(0) = 0$. De Taylor veelterm van graad 4 van $\sin(x)$ rond $x_0 = 0$ is dus

$$p(x) = 0 + x - 0 - \frac{1}{6}x^3 + 0 = x - \frac{1}{6}x^3.$$

Als we nu bijvoorbeeld de waarde van $\sin(\frac{\pi}{10})$ willen bepalen, krijgen we de benadering $\sin(\frac{\pi}{10}) \approx \frac{\pi}{10} - \frac{1}{6}(\frac{\pi}{10})^3 = \frac{\pi}{10} - \frac{\pi^3}{6000} \approx 0.30899$.

De fout kunnen we afschatten met $|R_4(\frac{\pi}{10})| < \frac{1}{5!} \cdot 1 \cdot (\frac{\pi}{10})^5 < 2.6 \cdot 10^{-5} = 0.000026$, dus hebben we aangetoond dat $0.30896 < \sin(\frac{\pi}{10}) < 0.30902$.

(2) We kunnen het Euler getal e benaderen met behulp van de Taylor ontwikkeling van $f(x) := \exp(x)$ in $x_0 = 0$. Omdat $f^{(n)}(x) = \exp(x)$, geldt $f^{(n)}(0) = 1$ voor alle n. We krijgen daarom de n-de Taylor veelterm

$$p(h) = 1 + h + \frac{h^2}{2} + \frac{h^3}{3!} + \ldots + \frac{h^n}{n!}.$$

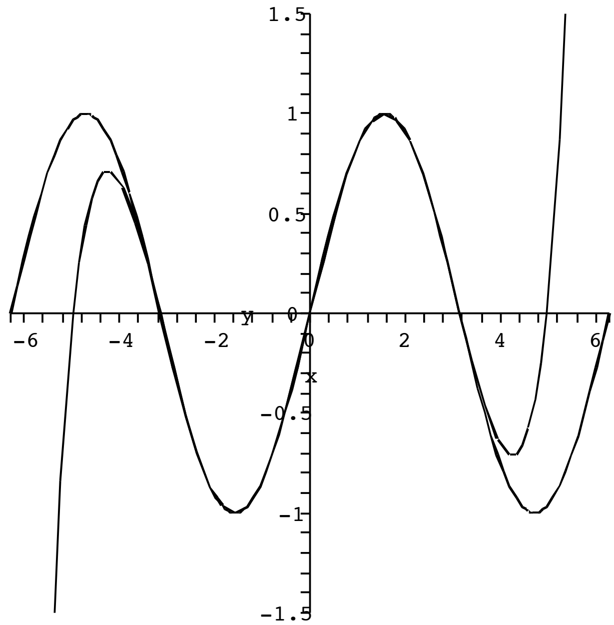
Als benadering van $e = f(1)$ krijgen we met $n = 6$ bijvoorbeeld $p(1) = 1 + 1 + \frac{1}{2} + \frac{1}{6} + \frac{1}{24} + \frac{1}{120} + \frac{1}{720} \approx 2.718$.

De fout kunnen we afschatten met $\frac{1}{(n+1)!}\, e^t \cdot 1^{n+1}$ voor een $t \in [0, 1]$, maar omdat we e juist willen berekenen, moeten we hier een grove afschatting voor e nemen, bijvoorbeeld $e^t \leq e < 3$. Hieruit volgt dat $|R_n(1)| < \frac{3}{7!} = \frac{3}{5040} \approx 0.0006$, dus hebben we aangetoond dat $2.7174 < e < 2.7186$.

OPDRACHT 9 *Bepaal met behulp van een Taylor veelterm de waarde van* $\sin(1)$ *op* 8 *decimalen, d.w.z. met een fout van hoogstens* 10^{-8}.

2.3 Taylor reeksen

Door de graad van de Taylor veelterm te laten groeien, krijgen we (in goedaardige gevallen) een steeds betere benadering van een functie, en meestal ook op een groter interval. In Figuur I.8 zien we bijvoorbeeld, dat de Taylor veelterm van graad 10 (deze keer in het punt $x_0 = 0$ berekend) tussen $-\pi$ en π bijna niet van de functie $\sin(x)$ te onderscheiden is.



Figuur I.8: Benadering van $\sin(x)$ door de Taylor veelterm van graad 10 in het punt $x_0 = 0$

Men krijgt dus het idee dat de benadering steeds beter wordt als we de graad verhogen, en het beste zou zijn, helemaal niet te stoppen maar oneindig door te gaan. Dit doen we dus!

Hier wordt nu duidelijk waarom het nuttig is de Taylor veelterm niet met de hoogste term $\frac{f^{(n)}(x_0)}{n!} h^n$ te beginnen, maar met de laagste term $f(x_0)$, want op deze manier kunnen we op een geschikte plek *enzovoorts* zeggen om aan te duiden, dat de veelterm met hogere termen door gaat.

Machtreeksen

Als men een algemene veeltermen met opstijgende termen schrijft, krijgt men iets als

$$p(x) = a_0 + a_1 x + a_2 x^2 + \ldots + a_n x^n = \sum_{i=0}^{n} a_i x^i.$$

Als we nu niet met een term $a_n x^n$ stoppen, maar ook voor een willekeurig hoge graad i steeds nog een coëfficiënt a_i van x^i definiëren (die weliswaar 0 mag zijn), noemen we dit geen veelterm meer, maar een *reeks*, soms voor alle duidelijkheid zelfs een *oneindige reeks*. Een reeks is dus een uitdrukking van de vorm

$$r(x) = a_0 + a_1 x + a_2 x^2 + \ldots + a_n x^n + \ldots = \sum_{n=0}^{\infty} a_n x^n.$$

Veeltermen zijn nu een speciale vorm van reeksen, een veelterm van graad n is namelijk een reeks waarbij vanaf $i = n + 1$ geldt dat $a_i = 0$ is.

Er zijn twee verschillende manieren hoe men tegen een oneindige reeks aan kan kijken, van een meer algebraïsch standpunt of van een meer analytisch standpunt.

- Algebraïsch kan men een reeks zien als een abstracte uitdrukking die gerepresenteerd is door de oneindige rij $(a_0, a_1, a_2, \ldots)$ van coëfficiënten. Twee reeksen laten zich net zo optellen als men veeltermen optelt, namelijk door de coëfficiënten van iedere term x^n bij elkaar op te tellen.

  Desnoods laten zich twee reeksen ook met elkaar vermenigvuldigen, dit geeft het zogeheten *convolutieproduct* of *Cauchy product* waarbij men in het product van de twee reeksen weer de coëfficiënten van de termen x^n bepaald. Dit gebeurt in principe ook hetzelfde als bij veeltermen, omdat alleen maar de termen tot graad n in de twee reeksen een bijdrage aan de term x^n in het product kunnen leveren. Er geldt

  $$(\sum_{n=0}^{\infty} a_n x^n) \cdot (\sum_{n=0}^{\infty} b_n x^n) = \sum_{n=0}^{\infty} c_n x^n \text{ met } c_n = \sum_{i=0}^{n} a_i b_{n-i},$$

  bijvoorbeeld is $c_3 = a_0 b_3 + a_1 b_2 + a_2 b_1 + a_3 b_0$.
- Analytisch zien we $f(x) = \sum_{n=0}^{\infty} a_n x^n$ als functie, waarbij we voor de functiewaarde in het punt x een limiet moeten bekijken, namelijk de limiet

  $$f(x) = \lim_{n\to\infty} \sum_{i=0}^{n} a_i x^i.$$

  Deze limiet hoeft niet voor iedere x te bestaan, dus hebben we het alleen maar voor degene waarden x echt met een functie te maken, waar deze limiet inderdaad bestaat.

Voorbeeld: Als we een voorbeeld van een reeks willen bekijken die geen veelterm is, schiet misschien als eerste de reeks door het hoofd waarbij alle coëfficiënten $a_n = 1$ zijn, dus de reeks

$$r(x) = 1 + x + x^2 + \ldots = \sum_{n=0}^{\infty} x^n.$$

Deze reeks heet de *meetkundige reeks* en heeft de mooie eigenschap dat voor $|x| < 1$ geldt dat

$$\sum_{n=0}^{\infty} x^n = \frac{1}{1-x}.$$

Dit gaat men na, door het product $(1 + x + \ldots + x^n)(1 - x)$ uit te schrijven als $(1+x+\ldots+x^n)(1-x) = (1+x+\ldots+x^n) - (x+x^2+\ldots+x^{n+1}) = 1 - x^{n+1}$. Voor $|x| < 1$ is $\lim_{n\to\infty} x^{n+1} = 0$, dus gaat $(1 + x + \ldots + x^n)(1 - x) \to 1$ en we moeten alleen nog door $(1 - x)$ delen om de boven aangegeven formule te krijgen.

De Taylor reeks van een functie

We gaan nu van de Taylor veeltermen een reeks maken door de graad van de veeltermen naar oneindig te laten gaan. De reeks die we zo krijgen noemen we natuurlijk Taylor reeks.

Definitie: De oneindige reeks $T(x) := T_{f,x_0}(x)$ gedefinieerd door

$$\begin{aligned}
T(x) &:= \lim_{n\to\infty} p_{f,n,x_0}(x) \\
&= f(x_0) + f'(x_0)(x - x_0) + \ldots + \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n + \ldots \\
&= \sum_{n=0}^{\infty} \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n.
\end{aligned}$$

heet de *Taylor reeks* van $f(x)$ in het punt x_0.

Over de Taylor reeks als limiet van de Taylor veeltermen moeten we een paar belangrijke opmerkingen kwijt:

(1) Natuurlijk geldt $T(x_0) = f(x_0)$, omdat alle termen in de som voor $n > 0$ wegvallen. Maar het kan zijn dat de reeks $T(x)$ voor geen enkele $x \neq x_0$ convergeert, dus een limiet heeft. Voor goedaardige functies (zo als we die in de praktijk kunnen verwachten) geldt dit wel, tenminste in een (voldoende klein) interval rond x_0.

De vraag of en waar de Taylor reeks van een functie convergeert, geeft aanleiding tot belangrijke en diepe stellingen in de wiskunde, maar is in deze cursus een minder belangrijke vraagstelling, omdat we veronderstellen dat we het met voldoende gladde functies te maken hebben.

(2) Zelfs als de Taylor reeks $T(x)$ voor een zekere waarde x convergeert, hoeft de limiet niet de juiste functiewaarde $f(x)$ te zijn. Een afschrikkend voorbeeld hiervoor is de functie

$$f(x) := e^{-\frac{1}{x^2}}$$

die door $f(0) := 0$ continu naar 0 voortgezet wordt. Deze functie is zelfs willekeurig vaak differentieerbaar, en er geldt $f^{(n)}(0) = 0$ voor alle n. Dit betekent, dat de coëfficiënten van de Taylor reeks van $f(x)$ in het punt $x_0 = 0$ alle 0 zijn en dus de Taylor reeks $T(x) = 0$ is. Aan de andere kant is $f(x) \neq 0$ voor $x \neq 0$, dus geeft de Taylor reeks de functie alleen maar in het nulpunt $x_0 = 0$ weer.

(3) De grootste afstand r zo dat $T(x)$ voor alle x met $|x - x_0| < r$ naar de goede waarde $f(x)$ convergeert, noemt men de *convergentiestraal* van $T(x)$.

Als we bij een functie $f(x)$ ervan uit gaan dat de Taylor reeks $T(x)$ in de punten waar hij convergeert ook naar de goede waarde $f(x)$ convergeert, schrijven we gewoon $f(x)$ in plaats van $T(x)$. Net als bij de Taylor veeltermen is het ook hier gebruikelijk, duidelijk te maken dat x dicht bij het punt x_0 ligt door $x = x_0 + h$ te schrijven. Op deze manier krijgt men de volgende vorm van de Taylor reeks:

$$f(x_0 + h) = \sum_{n=0}^{\infty} \frac{f^{(n)}(x_0)}{n!}\, h^n.$$

Voorbeelden:

(1) $\exp(x)$:
We bekijken de Taylor reeks in het punt $x_0 = 0$, daar geldt $\exp^{(n)}(x_0) = \exp(0) = 1$ voor alle $n \in \mathbb{N}$. De Taylor veelterm van graad n voor $\exp(x)$ in het punt 0 is dus

$$1 + \frac{1}{1!}x + \frac{1}{2!}x^2 + \ldots + \frac{1}{n!}x^n = \sum_{i=0}^{n} \frac{x^i}{i!}$$

en de Taylor reeks in het punt 0 is de limiet $n \to \infty$ van deze veeltermen, dus

$$T(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!} = 1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \frac{x^4}{24} + \ldots.$$

Er laat zich aantonen dat deze Taylor reeks convergentiestraal ∞ heeft, d.w.z. dat $T(x)$ voor elke $x \in \mathbb{R}$ naar $\exp(x)$ convergeert. Omdat de noemers met $n!$ heel snel groeien, is de convergentie erg goed en kunnen we de exponentiële functie al met weinig termen goed benaderen.

(2) $\sin(x)$:
Ook voor de sinus functie bepalen we de Taylor reeks in $x_0 = 0$. Merk op dat $\sin''(x) = -\sin(x)$, $\sin'''(x) = -\sin'(x)$ en $\sin^{(4)}(x) = \sin(x)$, hieruit

volgt dat de afgeleiden (beginnend met de 0-de afgeleide $\sin(0) = 0$) van $\sin(x)$ in het punt $x = 0$ periodiek $0, 1, 0, -1, 0, 1, 0, -1, 0$ enz. zijn. Dit betekent voor algemene n dat

$$\sin^{(2n)}(0) = 0 \quad \text{en} \quad \sin^{(2n+1)}(0) = (-1)^n.$$

De Taylor reeks van $\sin(x)$ in het punt 0 is dus

$$T(x) = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!} = x - \frac{x^3}{6} + \frac{x^5}{120} - \frac{x^7}{5040} + \ldots.$$

(3) $\cos(x)$:
Analoog met de sinus geldt voor de cosinus dat $\cos''(x) = -\cos(x)$, $\cos'''(x) = -\cos'(x)$ en $\cos^{(4)}(x) = \cos(x)$, dus zijn hier de afgeleiden in het punt $x_0 = 0$ periodiek $1, 0, -1, 0, 1, 0, -1, 0, 1$ enz. We hebben dus

$$\cos^{(2n)}(0) = (-1)^n \quad \text{en} \quad \cos^{(2n+1)}(0) = 0.$$

De Taylor reeks van $\cos(x)$ in het punt 0 is dus

$$T(x) = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!} = 1 - \frac{x^2}{2} + \frac{x^4}{24} - \frac{x^6}{720} + \ldots.$$

Net als bij de exponentiële functie is de convergentiestraal ook bij de Taylor reeks van de sinus en cosinus functies oneindig. Verder is ook hier de convergentie van de Taylor reeks erg goed, zo dat we snel een goede benadering van $\sin(x)$ of $\cos(x)$ vinden.

(4) $\log(x)$:
Omdat de logaritme voor $x = 0$ niet gedefinieerd is, moeten we hier de Taylor reeks in een andere punt bepalen, en we kiezen hiervoor $x_0 = 1$. Maar omdat het uiteindelijk toch prettiger is om een reeks met termen x^n en niet $(x-1)^n$ te hebben, gebruiken we een klein trucje: In plaats van $\log(x)$ kijken we naar de functie $f(x) := \log(x+1)$ en bepalen hiervoor de Taylor reeks in het punt $x_0 = 0$.

Er geldt $f'(x) = \frac{1}{x+1} = (x+1)^{-1}$, $f''(x) = (-1) \cdot (x+1)^{-2}$, $f'''(x) = (-1)(-2) \cdot (x+1)^{-3}$, $f^{(4)}(x) = (-1)(-2)(-3) \cdot (x+1)^{-4}$ en algemeen

$$f^{(n)}(x) = (-1)(-2) \ldots (-(n-1)) \cdot (x+1)^{-n} = (-1)^{n-1}(n-1)! \frac{1}{(x+1)^n}.$$

In het bijzonder is $f^{(n)}(0) = (-1)^{n-1}(n-1)!$ en dus $\frac{f^{(n)}(0)}{n!} = \frac{(-1)^{n-1}}{n}$. De Taylor reeks van $\log(x+1)$ in het punt 0 is dus

$$T(x) = \sum_{n=1}^{\infty} (-1)^{n-1} \frac{x^n}{n} = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \ldots.$$

Merk op dat deze Taylor reeks hoogstens convergentiestraal 1 kan hebben, omdat de logaritme in 0 niet gedefinieerd is. Dit is echter ook de convergentiestraal, we kunnen met deze reeks dus alleen maar waarden in het interval $(0, 2)$ bepalen. Maar dit stelt geen probleem voor het berekenen van $\log(x)$ voor, want voor $x \geq 2$ is $\frac{1}{x} \leq \frac{1}{2}$ en dit ligt wel in het interval $[0, 2]$ en we berekenen $\log(x)$ met behulp van de relatie $\log(x) = -\log(\frac{1}{x})$.

Er valt wel nog op te merken, dat de convergentie van de reeks voor de logaritme veel slechter is dan die voor $\exp(x)$, $\sin(x)$ of $\cos(x)$, omdat de noemers met n en niet met $n!$ groeien.

OPDRACHT 10 *Bepaal de Taylor reeks van* $f(x) := \sqrt{x+1}$ *in het punt* $x_0 = 0$.

2.4 Taylor reeksen voor functies van meerdere variabelen

Tot nu toe hebben we in deze les erna gekeken hoe we een gewone functie van één variabel door een oneindige reeks kunnen beschrijven of door een veelterm kunnen benaderen die we door afbreken van de Taylor reeks krijgen. De benadering van een functie $f(x)$ door de Taylor veelterm van graad n wordt vaak ook kort de n-de benadering van $f(x)$ genoemd. De meest belangrijke van deze benaderingen zijn de 0de (ook al lijkt het flauw), de 1ste of *lineaire* en de 2de of *kwadratische* benadering.

We zullen nu kijken, hoe we Taylor veeltermen op functies van meerdere veranderlijken kunnen veralgemenen. Het idee dat we hierbij hanteren is hetzelfde: We benaderen een functie $f(\mathbf{x}) : \mathbb{R}^n \to \mathbb{R}$ door een veelterm $p(\mathbf{x})$ in de n variabelen die in een vaste punt $\mathbf{x}_0$ dezelfde functiewaarde heeft als $f(\mathbf{x})$ en ook dezelfde eerste, tweede enz. partiële afgeleiden.

Om de notatie overzichtelijk te houden, zullen we weer vooral naar functies van twee variabelen kijken, de resultaten laten zich dan makkelijk ook voor n variabelen formuleren.

Om te beginnen moeten we afspreken wat een veelterm van meerdere variabelen is, en wat de graad daarvan is.

- Een *veelterm* in de twee variabelen x en y is een (eindige) som van termen van de vorm $c\,x^i y^j$, waarbij we c de *coëfficiënt* van $x^i y^j$ noemen. Een zuiver product $x^i y^j$ van machten van de variabelen noemt men ook en *monoom.*
- De *graad* van een term $c\,x^i y^j$ is de som $i + j$ van de machten van de variabelen. De graad van een veelterm is het maximum van de graden van de termen.
- Een veelterm van graad 0 heet een *constante functie*, een veelterm van graad 1 een *lineaire veelterm* of *lineaire functie* en een veelterm van graad 2 een *kwadratische veelterm* of *kwadratische functie.*

Voorbeeld: De algemene kwadratische veelterm van twee variabelen is van de vorm

$$p(x, y) = a + b_1 x + b_2 y + c_1 x^2 + c_2 xy + c_3 y^2.$$

We zullen nu de algemene definitie van de begrippen *veelterm* en *graad* voor n variabelen geven. Dit vergt enigszins veel indices en ingewikkelde notaties, maar met het geval van twee variabelen voor ogen zal het geen probleem zijn.

Definitie: Een *veelterm* in de n variabelen $x_1, \ldots, x_n$ is een eindige som van termen van de vorm

$$c\, x_1^{i_1} \cdot \ldots \cdot x_n^{i_n}.$$

De *graad* van een term $c\, x_1^{i_1} \cdot \ldots \cdot x_n^{i_n}$ is de som $i_1 + \ldots + i_n$ van de machten van de variabelen.

De *graad van een veelterm* is het maximum van de graden van de termen in de veelterm.

Voorbeelden:

(1) De algemene lineaire veelterm in $\mathbf{x} = (x_1, \ldots, x_n)$ is

$$p(\mathbf{x}) = a + b_1 x_1 + \ldots + b_n x_n = a + \sum_{i=1}^{n} b_i x_i.$$

(2) De algemene kwadratische veelterm is van de vorm

$$p(\mathbf{x}) = a + \sum_{i=1}^{n} b_i x_i + \sum_{i=1}^{n} \sum_{j=i}^{n} c_{ij} x_i x_j.$$

Lineaire benadering

Als we een functie $f(x, y)$ door een lineaire veelterm $p(x, y)$ willen benaderen, kunnen we eisen dat $p(x_0, y_0) = f(x_0, y_0)$ voor een vast gekozen punt (x_0, y_0) en dat de eerste partiële afgeleiden van $p(x, y)$ overeenkomen met de eerste partiële afgeleiden van $f(x, y)$ in het punt (x_0, y_0).

Voor het gemak nemen we nu nog aan, dat het vast gekozen punt (x_0, y_0) het nulpunt $(0, 0)$ is, we komen naar een algemeen punt terug door x door $h_1 := x - x_0$ en y door $h_2 := y - y_0$ te vervangen.

Voor de algemene lineaire veelterm

$$p(x, y) = a + b_1 x + b_2 y$$

geldt

$$\frac{\partial p}{\partial x} = b_1 \quad \text{en} \quad \frac{\partial p}{\partial y} = b_2.$$

Hieruit volgt

$$p(0, 0) = a, \quad \frac{\partial p}{\partial x}(0, 0) = b_1, \quad \frac{\partial p}{\partial y}(0, 0) = b_2$$

en om te bereiken dat

$$p(0, 0) = f(0, 0), \quad \frac{\partial p}{\partial x}(0, 0) = \frac{\partial f}{\partial x}(0, 0), \quad \frac{\partial p}{\partial y}(0, 0) = \frac{\partial f}{\partial y}(0, 0)$$

moet dus gelden dat

$$a = f(0, 0), \quad b_1 = \frac{\partial f}{\partial x}(0, 0), \quad b_2 = \frac{\partial f}{\partial y}(0, 0).$$

Voor de *lineaire benadering* van $f(x,y)$ in het punt $(x,y) = (0,0)$ krijgen we dus de lineaire veelterm

$$p(x,y) = f(0,0) + \frac{\partial f}{\partial x}(0,0)\,x + \frac{\partial f}{\partial y}(0,0)\,y.$$

Omdat $p(x,y)$ zo is gekozen dat de functiewaarde en de eerste partiële afgeleiden met $f(x,y)$ overeenkomen, geeft $p(x,y)$ juist het raakvlak aan de grafiek van $f(x,y)$ in het punt (x_0, y_0) aan.

Met behulp van de gradiënt kunnen we de lineaire benadering nog iets eenvoudiger schrijven: Met de notaties $\mathbf{x}_0 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ en $\mathbf{h} = \begin{pmatrix} x \\ y \end{pmatrix}$ geldt:

$$p(\mathbf{x}_0 + \mathbf{h}) = f(\mathbf{x}_0) + \nabla f(\mathbf{x}_0) \cdot \mathbf{h}.$$

Met deze notatie hebben we inderdaad de algemene vorm van de Taylor veelterm van graad 1 gevonden:

Stelling: Voor een functie $f(\mathbf{x}) : \mathbb{R}^n \to \mathbb{R}$ van n variabelen is de Taylor veelterm van graad 1 in het punt $\mathbf{x}_0$ gegeven door

$$p(\mathbf{x}_0 + \mathbf{h}) = f(\mathbf{x}_0) + \nabla f(\mathbf{x}_0) \cdot \mathbf{h} = f(\mathbf{x}_0) + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\mathbf{x}_0)\,h_i$$

waarbij h_i de i-de component van de vector $\mathbf{h}$ is, dus $\mathbf{h} = \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix}$.

Kwadratische benadering

Voor de algemene kwadratische veelterm

$$p(x,y) = a + b_1 x + b_2 y + c_1 x^2 + c_2 xy + c_3 y^2$$

geldt

$$\frac{\partial p}{\partial x} = b_1 + 2c_1 x + c_2 y \quad \text{en} \quad \frac{\partial p}{\partial y} = b_2 + c_2 x + 2c_3 y.$$

Om naar de Taylor veelterm van graad 2 te komen, moeten we ervoor zorgen dat naast de eerste partiële afgeleiden ook de tweede partiële afgeleiden van de algemene kwadratische veelterm

$$p(x,y) = a + b_1 x + b_2 y + c_1 x^2 + c_2 xy + c_3 y^2$$

in het punt $(0,0)$ met de tweede partiële afgeleiden van $f(x,y)$ in dit punt overeenkomen.

Voor de eerste partiële afgeleiden geldt

$$\frac{\partial p}{\partial x} = b_1 + 2c_1 x + c_2 y \quad \text{en} \quad \frac{\partial p}{\partial y} = b_2 + c_2 x + 2c_3 y.$$

en net als boven volgt uit $\frac{\partial p}{\partial x}(0,0) = \frac{\partial f}{\partial x}(0,0)$ en $\frac{\partial p}{\partial y}(0,0) = \frac{\partial f}{\partial y}(0,0)$ dat

$$a = f(0,0), \quad b_1 = \frac{\partial f}{\partial x}(0,0), \quad b_2 = \frac{\partial f}{\partial y}(0,0).$$

Voor de tweede partiële van $p(x,y)$ krijgen we:

$$\frac{\partial^2 p}{\partial x^2} = 2c_1, \quad \frac{\partial^2 p}{\partial x\,\partial y} = \frac{\partial^2 p}{\partial y\,\partial x} = c_2, \quad \frac{\partial^2 p}{\partial y^2} = 2c_3$$

en gelijk zetten met de tweede partiële afgeleiden van $f(x,y)$ in het punt $(0,0)$ geeft:

$$c_1 = \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(0,0), \quad c_2 = \frac{\partial^2 f}{\partial x\,\partial y}(0,0), \quad c_3 = \frac{1}{2}\frac{\partial^2 f}{\partial y^2}(0,0).$$

Voor de *kwadratische benadering* van $f(x,y)$ in het punt $(x,y) = (0,0)$ krijgen we dus de kwadratische veelterm

$$\begin{aligned}
p(x,y) &= f(0,0) + \frac{\partial f}{\partial x}(0,0)\,x + \frac{\partial f}{\partial y}(0,0)\,y \\
&\quad + \frac{1}{2}\frac{\partial^2 f}{\partial x^2}(0,0)\,x^2 + \frac{\partial^2 f}{\partial x\,\partial y}(0,0)\,xy + \frac{1}{2}\frac{\partial^2 f}{\partial y^2}(0,0)\,y^2 \\
&= f(0,0) + \frac{\partial f}{\partial x}(0,0)\,x + \frac{\partial f}{\partial y}(0,0)\,y \\
&\quad + \frac{1}{2}\left(\frac{\partial^2 f}{\partial x^2}(0,0)\,x^2 + \frac{\partial^2 f}{\partial x\,\partial y}(0,0)\,xy + \frac{\partial^2 f}{\partial y\,\partial x}(0,0)\,yx + \frac{\partial^2 f}{\partial y^2}(0,0)\,y^2\right).
\end{aligned}$$

Het opsplitsen van de term $\frac{\partial^2 f}{\partial x\,\partial y}(0,0)\,xy$ in de som $\frac{1}{2}(\frac{\partial^2 f}{\partial x\,\partial y}(0,0)\,xy + \frac{\partial^2 f}{\partial y\,\partial x}(0,0)\,yx)$ lijkt niet echt een vereenvoudiging, maar we zullen nu zien dat we hier wel iets aan hebben.

Net zo als we de eerste partiële afgeleiden in een *vector*, die we de gradiënt noemen, samengevat hebben, kunnen we de tweede partiële afgeleiden in een *matrix* samenvatten. Voor een functie $f(x,y)$ van twee variabelen geeft dit de 2×2-matrix

$$H := H_f := \begin{pmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial y\,\partial x} \\ \frac{\partial^2 f}{\partial x\,\partial y} & \frac{\partial^2 f}{\partial y^2} \end{pmatrix}$$

Deze matrix heet de *Hesse matrix* van f. Voor een algemene functie $f(\mathbf{x})$ van n variabelen is de Hesse matrix een $n \times n$-matrix met in de i-de rij de partiële afgeleiden naar de i-de variabel x_i en in de j-de kolom de partiële afgeleiden naar x_j, dus op positie (i,j) staat de tweede partiële afgeleide naar x_i en x_j.

Definitie: Voor een functie $f(\mathbf{x}) : \mathbb{R}^n \to \mathbb{R}$ heet de matrix

$$H := H_f := \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_2\,\partial x_1} & \cdots & \frac{\partial^2 f}{\partial x_n\,\partial x_1} \\ \vdots & \vdots & & \vdots \\ \frac{\partial^2 f}{\partial x_1\,\partial x_n} & \frac{\partial^2 f}{\partial x_2\,\partial x_n} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{pmatrix}$$

met $H_{ij} = \frac{\partial^2 f}{\partial x_j \, \partial x_i}$ de *Hesse matrix* van $f(\mathbf{x})$.

Omdat volgens de Stelling van Schwarz geldt dat $\frac{\partial^2 f}{\partial x_j \, \partial x_i} = \frac{\partial^2 f}{\partial x_i \, \partial x_j}$, is de Hesse matrix een symmetrische matrix, d.w.z. er geldt $H = H^{tr}$.

We noteren met $H(\mathbf{x}_0)$ (of $H(x, y)$ in het geval van twee variabelen) de Hesse matrix waarvoor de partiële afgeleiden in het punt $\mathbf{x}_0$ (of het punt (x, y)) geëvalueerd zijn.

Voor een vector $\mathbf{h} = \begin{pmatrix} x \\ y \end{pmatrix}$ geldt dan:

$$\begin{aligned} \mathbf{h}^{tr} \cdot H(\mathbf{x}_0) \cdot \mathbf{h} &= \frac{\partial^2 f}{\partial x^2}(\mathbf{x}_0)\, x^2 + \frac{\partial^2 f}{\partial y \, \partial x}(\mathbf{x}_0)\, xy + \frac{\partial^2 f}{\partial x \, \partial y}(\mathbf{x}_0)\, yx + \frac{\partial^2 f}{\partial y^2}(\mathbf{x}_0)\, y^2 \\ &= \frac{\partial^2 f}{\partial x^2}(\mathbf{x}_0)\, x^2 + 2\frac{\partial^2 f}{\partial y \, \partial x}(\mathbf{x}_0)\, xy + \frac{\partial^2 f}{\partial y^2}(\mathbf{x}_0)\, y^2. \end{aligned}$$

Met behulp van de Hesse matrix kunnen we Taylor veelterm van graad 2 daarom als volgt herschrijven:

$$p(\mathbf{x}_0 + \mathbf{h}) = f(\mathbf{x}_0) + \nabla f(\mathbf{x}_0) \cdot \mathbf{h} + \frac{1}{2}\, \mathbf{h}^{tr} \cdot H(\mathbf{x}_0) \cdot \mathbf{h}.$$

Ook hier hebben we met deze notatie de algemene vorm van de Taylor veelterm van graad 2 gevonden, er geldt:

Stelling: Voor een functie $f(\mathbf{x}) : \mathbb{R}^n \to \mathbb{R}$ van n variabelen is de Taylor veelterm van graad 2 in het punt $\mathbf{x}_0$ gegeven door

$$p(\mathbf{x}_0 + \mathbf{h}) = f(\mathbf{x}_0) + \nabla f(\mathbf{x}_0) \cdot \mathbf{h} + \frac{1}{2}\, \mathbf{h}^{tr} \cdot H(\mathbf{x}_0) \cdot \mathbf{h}.$$

waarbij h_i de i-de component van de vector $\mathbf{h}$ is.

Voorbeeld 1: We bekijken de functie

$$f(x, y) := e^x \cos(y)$$

in het punt $(x_0, y_0) = (0, 0)$. Er geldt

$$\frac{\partial f}{\partial x} = e^x \cos(y), \quad \frac{\partial f}{\partial y} = -e^x \sin(y),$$

en dus geldt voor de gradiënt

$$\nabla f = \begin{pmatrix} e^x \cos(y) \\ -e^x \sin(y) \end{pmatrix} \quad \text{en} \quad \nabla f(0, 0) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

Voor de tweede partiële afgeleiden geldt

$$\frac{\partial^2 f}{\partial x^2} = e^x \cos(y), \quad \frac{\partial^2 f}{\partial y \, \partial x} = -e^x \sin(y), \quad \frac{\partial^2 f}{\partial y^2} = -e^x \cos(y),$$

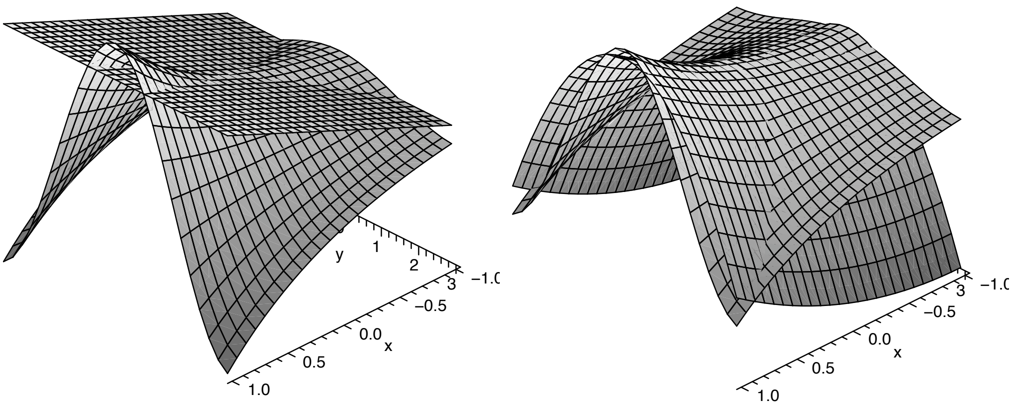
en dus krijgen we de Hesse matrix

$$H = \begin{pmatrix} e^x \cos(y) & -e^x \sin(y) \\ -e^x \sin(y) & -e^x \cos(y) \end{pmatrix} \quad \text{met} \quad H(0, 0) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Omdat $f(0,0) = 1$, zijn de Taylor veeltermen van graad 1 en 2 van $f(x,y)$ gegeven door

$$p_1(x,y) = 1 + x \quad \text{en} \quad p_2(x,y) = 1 + x + \frac{1}{2}x^2 - \frac{1}{2}y^2.$$

In Figuur I.9 zijn de grafiek van de functie en de grafieken van de benadering door de Taylor veeltermen van graad 1 en 2 te zien. Het is duidelijk dat de lineaire benadering het raakvlak aan de grafiek geeft en men ziet goed dat de kwadratische benadering in een omgeving van $(0,0)$ al redelijk goed is.



Figuur I.9: Benadering van $e^x \cos(y)$ door Taylor veeltermen van graad 1 en 2.

Voorbeeld 2: We bekijken de functie

$$f(x,y) := \sin(xy)$$

in het punt $(x_0, y_0) = (1, \frac{\pi}{2})$. Er geldt

$$\frac{\partial f}{\partial x} = y\,\cos(xy), \quad \frac{\partial f}{\partial y} = x\,\cos(xy),$$

en dus geldt voor de gradiënt

$$\nabla f = \begin{pmatrix} y\,\cos(xy) \\ x\,\cos(xy) \end{pmatrix} \quad \text{en} \quad \nabla f(1, \frac{\pi}{2}) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

De lineaire benadering van $f(x,y)$ is dus de constante $f(0,0) = 1$.

Voor de tweede partiële afgeleiden geldt

$$\frac{\partial^2 f}{\partial x^2} = -y^2 \sin(xy), \; \frac{\partial^2 f}{\partial y\,\partial x} = \cos(xy) - xy\,\sin(xy), \; \frac{\partial^2 f}{\partial y^2} = -x^2 \sin(xy),$$

en dus krijgen we de Hesse matrix

$$H = \begin{pmatrix} -y^2 \sin(xy) & \cos(xy) - xy\,\sin(xy) \\ \cos(xy) - xy\,\sin(xy) & -x^2 \sin(xy) \end{pmatrix}$$

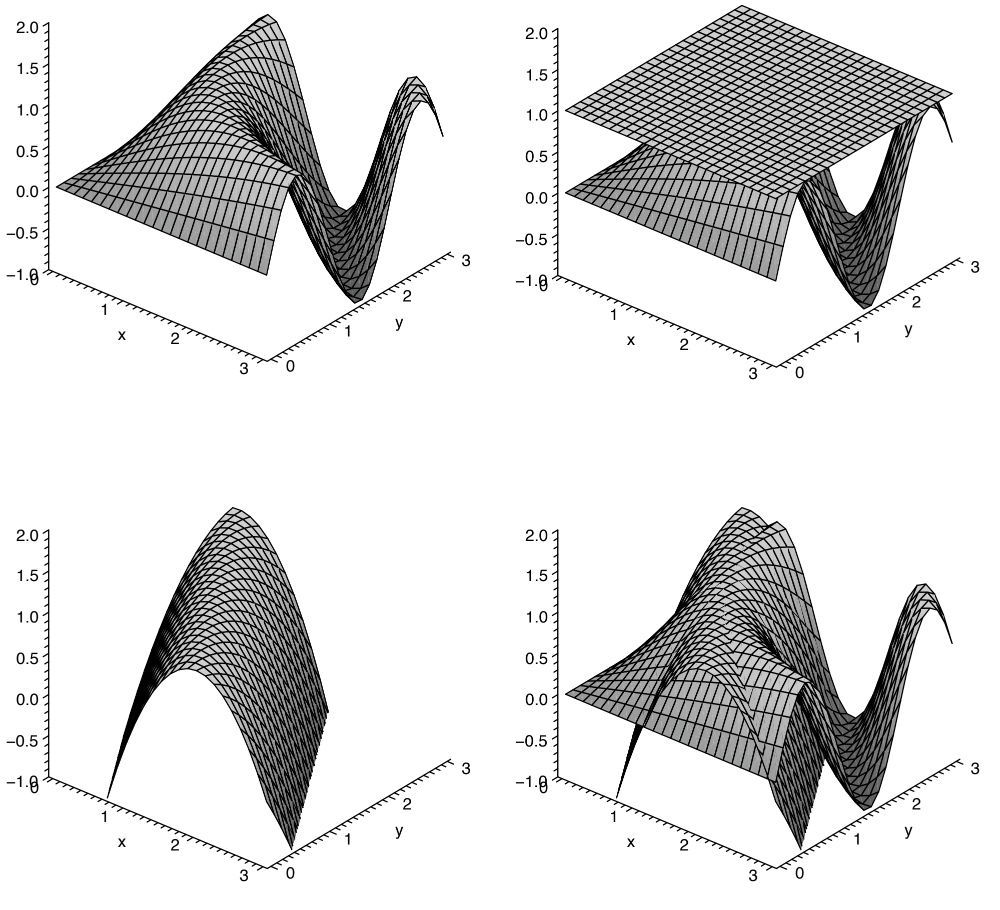
en dus

$$H(1, \frac{\pi}{2}) = \begin{pmatrix} -\frac{\pi^2}{4} & -\frac{\pi}{2} \\ -\frac{\pi}{2} & -1 \end{pmatrix}.$$

De Taylor veeltermen van graad 2 van $f(x, y)$ is dus gegeven door

$$\begin{aligned} p(x, y) &= 1 + \frac{1}{2}\left(-\frac{\pi^2}{4}(x-1)^2 - 2\frac{\pi}{2}(x-1)(y-\frac{\pi}{2}) - (y-\frac{\pi}{2})^2\right) \\ &= 1 - \frac{\pi^2}{8}(x-1)^2 - \frac{\pi}{2}(x-1)(y-\frac{\pi}{2}) - \frac{1}{2}(y-\frac{\pi}{2})^2). \end{aligned}$$

In Figuur I.10 zijn boven de grafiek van de functie en de (constante) lineaire benadering te zien, beneden de grafiek van de kwadratische benadering apart en de kwadratische benadering samen met de grafiek van de functie.



Figuur I.10: Benadering van $\sin(xy)$ door Taylor veeltermen van graad 1 (boven) en 2 (beneden) in het punt $(x_0, y_0) = (1, \frac{\pi}{2})$.

Opdracht 11 *Bepaal voor* $f(x, y) := e^{x^2+y^2}$ *de Taylor veelterm van graad* 2 *in het punt* $(0, 0)$.

Algemene vorm van de Taylor reeks

Als we tot Taylor veeltermen van hogere graden dan 2 en misschien zelfs tot Taylor reeksen voor functies van meerdere variabelen willen komen, wordt het vergelijken van de partiële afgeleiden van een veelterm met algemene coëfficiënten met de partiële afgeleiden van $f(\mathbf{x})$ snel erg onhandig. Maar gelukkig kunnen we de Taylor veeltermen en Taylor reeks voor functies van meerdere variabelen met een klein trucje ook afleiden uit de Taylor reeks voor een functie van één variabel.

Als we een functie $f(\mathbf{x})$ in een kleine omgeving van een vast gekozen punt $\mathbf{x}_0$ bekijken, kunnen we dit schrijven als $f(\mathbf{x}_0 + \mathbf{h})$ voor een (korte) vector $\mathbf{h}$. Hieruit maken we nu (kunstmatig) een nieuwe functie $g(t)$ van een enkele variabel t, namelijk

$$g(t) := f(\mathbf{x}_0 + t \cdot \mathbf{h}).$$

We krijgen de Taylor reeks van $f(\mathbf{x})$ in het punt $\mathbf{x}_0$ door $g(t)$ rond $t = 0$ in een Taylor reeks te ontwikkelen en vinden $f(\mathbf{x}_0 + \mathbf{h})$ door $t = 1$ in deze Taylor reeks in te vullen, want $f(\mathbf{x}_0 + \mathbf{h}) = g(1)$. Op deze manier berekenen we $f(\mathbf{x}_0 + \mathbf{h})$ door:

$$\begin{aligned} f(\mathbf{x}_0 + h) = g(1) &= g(0) + g'(0) \cdot 1 + \frac{1}{2} g''(0) \cdot 1^2 + \ldots + \frac{1}{n!} g^{(n)}(0) \cdot 1^n + \ldots \\ &= \sum_{n=0}^{\infty} \frac{1}{n!} g^{(n)}(0). \end{aligned}$$

Voor de Taylor reeks van $f(\mathbf{x})$ moeten we dus alleen maar de afgeleiden van $g(t)$ in het punt $t = 0$ bepalen. De cruciale waarneming is nu als volgt:

Merk op: De gewone afgeleide $g'(t)$ komt precies overeen met de definitie van de richtingsafgeleide van $f(\mathbf{x})$ in de richting van $\mathbf{h}$.

Voor een vector $\mathbf{h} = \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix}$ met componenten h_i geldt daarom

$$g'(t) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\mathbf{x}_0 + t\mathbf{h})\, h_i \quad \text{en dus} \quad g'(0) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\mathbf{x}_0)\, h_i.$$

Voor de tweede afgeleide $g''(t)$ moeten we weer de richtingsafgeleide in de richting van $\mathbf{h}$ nemen, dit geeft

$$g''(t) = \sum_{j=1}^{n} \sum_{i=1}^{n} \frac{\partial^2 f}{\partial x_j\, \partial x_i}(\mathbf{x}_0 + t\mathbf{h})\, h_i h_j \quad \text{en} \quad g''(0) = \sum_{j=1}^{n} \sum_{i=1}^{n} \frac{\partial^2 f}{\partial x_j\, \partial x_i}(\mathbf{x}_0)\, h_i h_j.$$

Op een soortgelijke manier krijgen we voor de derde afgeleide

$$\begin{aligned} g'''(t) &= \sum_{k=1}^{n} \sum_{j=1}^{n} \sum_{i=1}^{n} \frac{\partial^3 f}{\partial x_k\, \partial x_j\, \partial x_i}(\mathbf{x}_0 + t\mathbf{h})\, h_i h_j h_k, \\ g'''(0) &= \sum_{k=1}^{n} \sum_{j=1}^{n} \sum_{i=1}^{n} \frac{\partial^3 f}{\partial x_k\, \partial x_j\, \partial x_i}(\mathbf{x}_0)\, h_i h_j h_k \end{aligned}$$

en we kunnen in principe zo doorgaan tot hogere afgeleiden, de n-de term is dan een n-voudige som van de n-de partiële afgeleiden van $f(\mathbf{x})$. Maar ook deze sommen zijn nog niet echt prettig om op te schrijven, daarom zullen we nu nog naar een veel gebruikte notatie kijken, waarmee de Taylor reeks voor meerdere variabelen bijna net zo eenvoudig wordt als de Taylor reeks voor functies van één variabel.

Het idee is, de partiële afgeleide als een soort bewerkingsvoorschrift te beschouwen, die we op een functie toepassen en die we een *operator* noemen. In principe is dit niets nieuws, want ook de gewone afgeleide kunnen we zien als een operator $'$ die uit een functie $f(x)$ een andere functie $f'(x)$ maakt. Net zo interpreteren we nu de partiële afgeleide als een operator $\frac{\partial}{\partial x}$ die uit de functie $f(\mathbf{x})$ de nieuwe functie $\frac{\partial}{\partial x} f = \frac{\partial f}{\partial x}$ maakt.

Met deze operatoren kunnen we nu op een voor de hand liggende manier rekenen, voor de som $\frac{\partial}{\partial x} + \frac{\partial}{\partial y}$ van twee operatoren moeten we hiervoor aangeven, welke nieuwe functie de toepassing van $\frac{\partial}{\partial x} + \frac{\partial}{\partial y}$ op een functie $f(\mathbf{x})$ geeft. Als verdere bewerkingen definiëren we ook de vermenigvuldiging van twee operatoren en het vermenigvuldigen van een operator met een constante:

$$\begin{aligned}(c \cdot \frac{\partial}{\partial x_i}) f &:= c \cdot \frac{\partial f}{\partial x_i} \\ (\frac{\partial}{\partial x_i} + \frac{\partial}{\partial x_j}) f &:= \frac{\partial f}{\partial x_i} + \frac{\partial f}{\partial x_j} \\ (\frac{\partial}{\partial x_i} \cdot \frac{\partial}{\partial x_j}) f &:= \frac{\partial^2 f}{\partial x_i \, \partial x_j}.\end{aligned}$$

Hiermee krijgen we bijvoorbeeld

$$(\frac{\partial}{\partial x} + \frac{\partial}{\partial y})^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y \, \partial x} + \frac{\partial^2}{\partial x \, \partial y} + \frac{\partial^2}{\partial y^2} = \frac{\partial^2}{\partial x^2} + 2 \frac{\partial^2}{\partial y \, \partial x} + \frac{\partial^2}{\partial y^2}.$$

Merk op: Dit is in principe dezelfde formule als de binomische formule

$$(a + b)^2 = a^2 + 2ab + b^2$$

en dit betekent dat we met de partiële afgeleiden als operatoren juist zo rekenen als met gewone variabelen.

Met behulp van de operatoren kunnen we het inproduct van een vector $\mathbf{h} = \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix}$ met de gradiënt ∇f herschrijven als

$$\nabla f(\mathbf{x}_0) \cdot \mathbf{h} = (h_1 \cdot \frac{\partial}{\partial x_1} + \ldots + h_n \cdot \frac{\partial}{\partial x_n}) f(\mathbf{x}_0)$$

waarbij het invullen van $\mathbf{x}_0$ aan de rechterkant natuurlijk *na* het toepassen van de operatoren gebeurt.

Net zo krijgen we voor het product met de Hesse matrix

$$\mathbf{h}^{tr} \cdot H(\mathbf{x}_0) \cdot \mathbf{h} = (h_1 \cdot \frac{\partial}{\partial x_1} + \ldots + h_n \cdot \frac{\partial}{\partial x_n})^2 f(\mathbf{x}_0).$$

De Taylor reeks voor een functie van n variabelen laat zich nu schrijven als:

$$f(\mathbf{x}_0 + \mathbf{h}) = \sum_{n=0}^{\infty} \frac{1}{n!}(h_1 \cdot \frac{\partial}{\partial x_1} + \ldots + h_n \cdot \frac{\partial}{\partial x_n})^n f(\mathbf{x}_0).$$

Voorbeeld: We berekenen de derdegraads term van de Taylor reeks voor een functie $f(x, y)$ van twee variabelen. Er geldt $(a+b)^3 = a^3+3a^2b+3ab^2+b^3$, hieruit volgt

$$(h_1\frac{\partial}{\partial x} + h_2\frac{\partial}{\partial y})^3 = h_1^3\frac{\partial^3}{\partial x^3} + 3h_1^2h_2\frac{\partial^3}{\partial y\, \partial x^2} + 3h_1h_2^2\frac{\partial^3}{\partial y^2\, \partial x} + h_2^3\frac{\partial^3}{\partial y^3}$$

dus is de derdegraads term van de Taylor reeks

$$\frac{1}{3!}\left(\frac{\partial^3 f}{\partial x^3}(\mathbf{x}_0)\, h_1^3 + 3\frac{\partial^3 f}{\partial y\, \partial x^2}(\mathbf{x}_0)\, h_1^2h_2 + 3\frac{\partial^3 f}{\partial y^2\, \partial x}(\mathbf{x}_0)\, h_1h_2^2 + \frac{\partial^3 f}{\partial y^3}(\mathbf{x}_0)\, h_2^3\right).$$

Belangrijke begrippen in deze les

- Lagrange interpolatie
- Taylor veelterm
- foutterm
- oneindige reeksen
- Taylor reeks
- lineaire/kwadratische benadering
- Hesse matrix
- partiële afgeleide als operator

Opgaven

12. Bepaal voor $f(x) := (1 + x)^a$ met $a \in \mathbb{R}$ de Taylor veelterm van graad 3 in $x_0 = 0$.
13. Bepaal voor de volgende functies de Taylor veelterm van graad 5 in $x_0 = 0$:
    (i) $f(x) := \sqrt[3]{1 + x} = (1 + x)^{\frac{1}{3}}$;
    (ii) $f(x) := \frac{1}{1+x^2}$;
    (iii) $f(x) := \frac{1}{\sqrt{1-x^2}}$;

(iv) $f(x) := \frac{1}{3+x}$;

(v) $f(x) := \sin(2x^2)$;

(vi) $f(x) := e^{-3x}$.

14. Zij $f(x) := \sqrt[3]{8+x} = (8+x)^{\frac{1}{3}}$.

    (i) Bepaal voor $f(x)$ de Taylor veelterm van graad 4 in $x_0 = 0$.

    (ii) Vind met behulp van de Taylor veelterm in (i) een benadering voor $\sqrt[3]{9}$.

    (iii) Geef met behulp van de Lagrange vorm van de foutterm een afschatting voor de mogelijke fout van de benadering in (ii).

15. Bepaal voor de volgende functies de Taylor reeksen in de aangeven punten:

    (i) $f(x) := \frac{1}{x}$ in $x_0 = 1$;

    (ii) $f(x) := e^x$ in $x_0 = 2$;

    (iii) $f(x) := \sin(x+1)$ in $x_0 = -1$;

    (iv) $f(x) := \log(x)$ in $x_0 = 2$.

16. Vind de Taylor reeks voor $f(x) := \arctan(x)$ in $x_0 = 0$. (Hint: Er geldt $\arctan'(x) = \frac{1}{1+x^2}$ en dit laat zich voor $x^2 < 1$ schrijven als meetkundige reeks $\frac{1}{1+x^2} = 1 - x^2 + x^4 - x^6 + \ldots = \sum_{n=0}^{\infty}(-1)^n x^{2n}$.)

17. Bepaal voor de functie $f(x,y) := \sqrt{x^2+y^3}$ de Taylor veelterm van graad 2 in het punt $(1,2)$. Gebruik de Taylor veelterm om de waarde van $\sqrt{1.02^2+1.97^3}$ te benaderen en vergelijk het resultaat met de *juiste* waarde (volgens een rekenmachine).

18. Bepaal voor de volgende functies de Taylor veelterm van graad 2 in het punt $(x_0, y_0) = (0,0)$:

    (i) $f(x,y) := \frac{1}{1+x^2+y^2}$;

    (ii) $f(x,y) := e^{x+y}$;

    (iii) $f(x,y) := e^{-x^2-y^2}\cos(xy)$;

    (iv) $f(x,y) := \sin(xy) + \cos(xy)$.

19. Bepaal voor de volgende functies de Taylor veelterm van graad 2 in de aangegeven punten (x_0, y_0):

    (i) $f(x,y) := \frac{1}{2+xy^2}$ in $(0,0)$;

    (ii) $f(x,y) := \log(1+x+y+xy)$ in $(0,0)$;

    (iii) $f(x,y) := \arctan(x+xy)$ in $(0,-1)$;

    (iv) $f(x,y) := x^2+xy+y^3$ in $(1,-1)$;

    (v) $f(x,y) := \sin(2x+3y)$ in $(0,0)$;

    (vi) $f(x,y) := \frac{\sin(x)}{y}$ in $(\frac{\pi}{2}, 1)$;

    (vii) $f(x,y) := \frac{1+x}{1+x^2+y^4}$ in $(0,0)$;

    (viii) $f(x,y) := e^{(x-1)^2}\cos(y)$ in $(1,0)$.

20. Bepaal voor de volgende functies de Taylor veelterm van graad 3 in de aangegeven punten (x_0, y_0):

    (i) $f(x,y) := \frac{1}{2+x-2y}$ in $(2,1)$;

    (ii) $f(x,y) := \log(x^2+y^2)$ in $(1,0)$;

    (iii) $f(x,y) := \cos(x+\sin(y))$ in $(0,0)$.

Les 3 Extrema van functies van meerdere variabelen

Bij gewone functies van één variabel hebben we in Wiskunde 1 de vraag behandeld hoe we minima en maxima van een functie kunnen vinden. Het belangrijkste criterium was dat een differentieerbare functie in een lokaal extremum (minimum of maximum) een horizontale raaklijn heeft, de afgeleide van de functie in een extremum is dus noodzakelijk nul. Punten met deze eigenschap hebben we *kritieke punten* genoemd, naast de speciale punten waar de functie niet differentieerbaar is en de randpunten van het interval waarop we de functie bekijken.

3.1 Classificatie van kritieke punten

Het kan zijn dat een functie in een punt een horizontale raaklijn heeft, zonder in dit punt een extremum te hebben. Dit is bijvoorbeeld het geval voor de functie $f(x) = x^3$ in het punt $x = 0$. Zo'n punt noemt men ook een *zadelpunt*. Het verschil tussen een zadelpunt en een echt extremum laat zich aan de hand van de tweede afgeleide $f''(x)$ beschrijven:

Bij een minimum in het punt x_0 is de functie links van x_0 (dus voor $x < x_0$) dalend en rechts van x_0 stijgend, dus is de eerste afgeleide $f'(x)$ links van x_0 negatief en rechts van x_0 positief. Dit betekent dat $f'(x)$ rond x_0 stijgend is en dus geldt $f''(x_0) > 0$. Net zo volgt uit $f'(x_0) = 0$ en $f''(x_0) < 0$, dat de functie $f(x)$ in het punt x_0 een maximum heeft.

Met onze kennis van Taylor reeksen kunnen we dit nu ook van een andere kant bekijken. We ontwikkelen $f(x)$ rond een kritiek punt x_0 met $f'(x_0) = 0$ in een Taylor reeks, dit geeft:

$$f(x_0 + h) = f(x_0) + \frac{1}{2} f''(x_0)\, h^2 + \ldots$$

In een kleine omgeving van x_0 (dus voor kleine waarden van h) kunnen we de hogere termen h^3, h^4, enz. tegenover h^2 verwaarlozen en we kunnen $f(x)$ door de kwadratische functie

$$g(h) := f(x_0) + \frac{1}{2} f''(x_0)\, h^2$$

benaderen. Omdat het bestaan van een lokaal extremum in x_0 alleen maar van een kleine omgeving van x_0 afhangt, heeft $f(x)$ in x_0 een maximum/minimum dan en slechts dan als $g(h)$ in 0 een maximum/minimum heeft. Maar de vergelijking van $g(h)$ beschrijft juist een parabool met toppunt $(0, f(x_0))$ en deze parabool is naar boven geopend als $f''(x_0) > 0$ en naar beneden geopend als $f''(x_0) < 0$. De volgende twee gevallen zijn dus duidelijk:

(1) Voor $f''(x_0) > 0$ heeft $g(h)$ in $h = 0$ een minimum, dus heeft ook $f(x)$ een minimum in x_0.

(2) Voor $f''(x_0) < 0$ heeft $g(h)$ in $h = 0$ een maximum, dus heeft ook $f(x)$ een maximum in x_0.