Max van Horssen
eerstejaars student, Radboud Universiteit Nijmegen
max.vanhorssen@student.ru.nl

Onderwijs Profielwerkstuk

Interpolatiepolynomen en coëfficiënten

Polynomiale interpolatie is het vinden van het polynoom met de laagst mogelijke graad die door een gegeven aantal punten gaat. Over dit vraagstuk schreef Max van Horssen, destijds 6 vwo-leerling van de Stichtse Vrije School in Zeist, zijn profielwerkstuk. Hij stelde zich ten doel een formule te vinden voor de coëfficiënten van het polynoom in standaardvorm (dat wil zeggen als machtreeks). Met elementaire algebra leidde hij drie verschillende uitdrukkingen af, waarvan hij de algemene geldigheid bewees. Deze uitdrukkingen kon hij echter niet direct terugvinden in de literatuur. Wel zijn de polynomen van Newton en Lagrange bekend maar deze geven niet rechtstreeks de coëfficiënten. Max heeft de verschillende polynoomvormen geïmplementeerd in een Javaprogramma en hun algoritmische complexiteit en numerieke stabiliteit vergeleken. Na afronding van het werkstuk kwam hij in contact met wiskundigen die hem op het spoor van de lineaire algebra zetten. Door zo naar het vraagstuk te kijken werd het verband tussen de verschillende polynoomvormen duidelijk.

In het voorjaar van 2017 (in 4 vwo) was ik bezig om te leren programmeren in Java. Eén programma stuurde raketten aan die de maan moesten bereiken. Het vluchtplan werd opgesteld door een genetisch algoritme dat in de loop van generaties steeds beter moest presteren. Om het prestatieverloop te kunnen volgen tekende ik de resultaten in een grafiek, waarbij opeenvolgende punten werden verbonden met rechte lijnen. Ik vroeg me opeens af of alle punten in de grafiek ook met één vloeiende lijn konden worden verbonden. Zo werd mijn interesse in het interpolatieprobleem gewekt.

In de wiskundeles hadden we geleerd de functie $\alpha x+\beta$ te bepalen als de waarden y_0, y_1 in respectievelijk x_0, x_1 gegeven zijn. We hebben dan een stelsel van twee vergelijkingen met twee onbekenden, de coëfficiënten van de functie. De eerste coëfficiënt α vinden we door oplossen van het stelsel:

$$\alpha=\frac{y_0-y_1}{x_0-x_1}.$$

De tweede coëfficiënt β verkregen we door de eerste coëfficiënt te substitueren in een van de vergelijkingen.

Ik begon met het geval van twee punten uit te breiden naar drie punten x_0, x_1, x_2 met respectievelijk waarden y_0, y_1, y_2. Dit geeft een kwadratische functie met drie coëfficiënten: $\alpha x^2+\beta x+\gamma$. Oplossen van dit stelsel geeft voor α:

$$\alpha=\frac{y_0}{(x_0-x_1)(x_0-x_2)}+\frac{y_1}{(x_1-x_0)(x_1-x_2)}+\frac{y_2}{(x_2-x_0)(x_2-x_1)}.$$

Het viel me op dat α voor twee punten in verwante vorm geschreven kan worden:

$$\alpha=\frac{y_0}{x_0-x_1}+\frac{y_1}{x_1-x_0}$$

en dat in het algemene geval van $n+1$ punten de hoogste coëfficiënt er zo uit zou kunnen zien:

$$\alpha=\sum_{j=0}^{n}\frac{y_j}{\prod_{k=0,k\neq j}^{n}(x_j-x_k)}.$$

Vanaf dit moment nummerde ik de coëfficiënten als volgt: C_n voor α, C_{n-1} voor β, et cetera. Ik heb alle coëfficiënten van het interpolatiepolynoom $\sum_{i=0}^{n}C_i x^i$ voor drie en vier punten uitgerekend. Het bleek dat ze alle van de volgende vorm zijn:

$$C_i=\pm\sum_{j=0}^{n}y_j\frac{Q_{i,j}}{\prod_{k=0,k\neq j}^{n}(x_j-x_k)}$$

met een gegeven teken (plus of min) en een nog te bepalen uitdrukking $Q_{i,j}$. Voor $n=3$ en $0 \leq j \leq n$ vond ik voor de vier gevallen $Q_{0,j}, \ldots, Q_{3,j}$ de volgende uitdrukkingen:

$$Q_{0,j} = 1, \quad Q_{1,j} = \sum_{\substack{0 \leq k \leq n \\ k \neq j}} x_k, \quad Q_{2,j} = \sum_{\substack{0 \leq k_1 < k_2 \leq n \\ k_1, k_2 \neq j}} x_{k_1} x_{k_2}, \quad Q_{3,j} = \prod_{\substack{0 \leq k \leq n \\ k \neq j}} x_k.$$

De vraag die ik moest beantwoorden was: welke uitdrukking $Q_{i,j}$ beschrijft de bovenstaande reeks en is deze van toepassing voor elk aantal punten? Met deze vraag in mijn achterhoofd vertrok ik met mijn klas voor een taalreis naar Dresden. Tijdens de lange busreis viel het kwartje, namelijk: $Q_{i,j}$ is de som van het product van alle combinaties x-en, exclusief die met een gegeven index j, met een gegeven lengte $n-i$ en zonder herhaling,

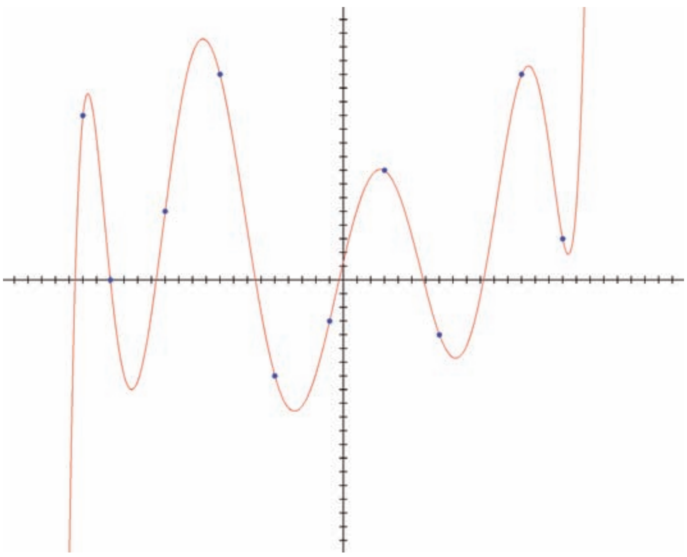
$$Q_{i,j} := \sum_{\substack{V \in \mathcal{P}(X \setminus x_j) \\ |V| = n-i}} \prod_{x' \in V} x'.$$

Hier is X de verzameling van alle x-en en $\mathcal{P}(X)$ de machtsverzameling van X. Zo vond ik als algemene uitdrukking voor het interpolatiepolynoom:

$$f(x) = \sum_{i=0}^{n} (-1)^{n-i} \sum_{j=0}^{n} y_j \frac{Q_{i,j}}{\prod_{k=0, k \neq j}^{n} (x_j - x_k)} x^i.$$

Het is natuurlijk nog de vraag of deze formule inderdaad geldig is voor elk aantal punten. Ik implementeerde de formule in mijn Java-programma zodat ik dit kon testen. Een spannend moment was toen ik de formule voor het eerst uitprobeerde met tien punten. Samen met mijn vader zat ik op het puntje van mijn stoel om te kijken of het zou werken. En inderdaad, de functie ging precies door alle punten! Zie Figuur 1.

Maar dit is slechts een empirisch bewijs. “Nu nog het echte bewijs,” zei mijn vader (informaticus), half serieus, “dat zou met volledige inductie kunnen.” Dat vond ik een mooie uitdaging waarmee ik in de zomervakantie aan de slag ging. Tijdens een bergwandeling in Oostenrijk kreeg ik het centrale idee voor het bewijs. Namelijk: door het polynoom uit te drukken in een polynoom van een lagere graad plus een restterm kon ik de inductiehypothese gebruiken. Met een paar zelfgeformuleerde hulpstellingen lukte het om het bewijs rond te krijgen. Terug van vakantie heb ik samen met mijn vader het bewijs doorgenomen en gestructureerd opgeschreven. Door te zoeken op internet kwam ik erachter dat mijn uitdrukking voor $Q_{i,j}$ bekend is als het elementaire symmetrische polynoom. Na het maken van het bewijs was ik overtuigd van mijn studiekeuze wiskunde en inmiddels was ik zo enthousiast dat ik besloot mijn profielwerkstuk over dit onderwerp te schrijven.



Figuur 1 Output van het programma voor interpolatie door tien punten.

Definities en notaties

Bij het interpoleren komen we de volgende wiskundige begrippen tegen:

1. *Gedeelde differenties*:
$$f[x_0] := y_0,$$
$$f[x_0, \ldots, x_n] := \frac{f[x_0, \ldots, x_{n-1}] - f[x_1, \ldots, x_n]}{x_0 - x_n}.$$
2. *Elementaire symmetrische polynomen*: de som van het product van alle combinaties elementen van een verzameling met een gegeven lengte k en zonder herhaling:
$$e_k(x_0, \ldots, x_n) := \sum_{0 \leq i_0 < \cdots < i_k \leq n} x_{i_0} \cdots x_{i_k}.$$
3. *Compleet homogene symmetrische polynomen*: de som van het product van alle combinaties elementen van een verzameling met een gegeven lengte k en met herhaling:
$$h_k(x_0, \ldots, x_n) := \sum_{0 \leq i_0 \leq \cdots \leq i_k \leq n} x_{i_0} \cdots x_{i_k}.$$
4. *Vandermonde-matrix*:
$$\boldsymbol{V} := \begin{bmatrix} 1 & x_0 & \cdots & x_0^n \\ 1 & x_1 & \cdots & x_1^n \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & \cdots & x_n^n \end{bmatrix}.$$
5. *Vandermonde-determinant* (stelling):
$$|\boldsymbol{V}| = \prod_{0 \leq i < j \leq n} (x_j - x_i).$$

Verschillende vormen

Voor het werkstuk wilde ik weten waar de formule in de literatuur terug te vinden is. Ik vond de interpolatiepolynomen van Newton en Lagrange:

$$N(x) = \sum_{i=0}^{n} f[x_0, \ldots, x_i] \prod_{j=0}^{i-1} (x - x_j),$$
$$L(x) = \sum_{i=0}^{n} y_i \prod_{j=0, j \neq i}^{n} \frac{x - x_j}{x_i - x_j},$$

en een verwijzing naar de Vandermonde-matrix $\boldsymbol{V}$. De lineaire algebra methode die gebruik maakt van de inverse van $\boldsymbol{V}$ en de Vandermonde-determinant leidt tot

$$V(x) = \sum_{i=0}^{n} (-1)^{n-i} \sum_{j=0}^{n} y_j \frac{e_{n-i}(x_0, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n)}{\prod_{k=0, k \neq j}^{n} (x_j - x_k)} x^i.$$

Deze vorm van het interpolatiepolynoom komt overeen met de

Foto: Ed van Rijswijk

Max van Horssen

eerder gevonden vorm $f(x)$ als we daarin de uitdrukking $Q_{i,j}$ vervangen door het elementaire symmetrische polynoom e_{n-i}.

Het Newton-polynoom $N(x)$ maakt gebruik van de gedeelde differenties (GD) en $V(x)$ maakt gebruik van het elementaire symmetrische polynoom (ESP). Ik plaatste de gevonden polynomen in een taxonomie, gebaseerd op de vorm van de uitdrukkingen. Zie Tabel 1.

Wat opviel was de open plaats in de taxonomie. Zou er een vierde vorm bestaan die zowel GD als ESP gebruikt? Inderdaad bleek dat $V(x)$ omgeschreven kan worden naar een vorm die aan de gewenste eigenschappen voldoet. Zo vond ik de volgende vorm:

$$H(x) = \sum_{i=0}^{n} \sum_{j=i}^{n} f[x_0, \ldots, x_j](-1)^{j-i} e_{j-i}(x_0, \ldots, x_{j-1}) x^i.$$

Omdat ik deze vorm niet kon vinden in de (online) literatuur en er toch een naam voor nodig had heb ik deze vorm voor het moment $H(x)$ genoemd, naar de eerste letter van mijn achternaam. (Later bleek de letter H goed te passen bij deze vorm, omdat in de lineaire algebra deze vorm tot stand komt met behulp van de matrix $\boldsymbol{H}$ die de compleet homogene symmetrische polynomen bevat, zie kader 'Lineaire algebra' op de volgende pagina.)

Taxonomie	**Geen ESP**	**Wel ESP**
Geen GD	$L(x)$	$V(x)$
Wel GD	$N(x)$	$\ldots$

Ten slotte kunnen we de coëfficiënten van het polynoom ook op een recursieve manier vinden. We bepalen eerst de hoogste coëfficiënt en na substitutie ontstaat een nieuw stelsel waarvan we weer de hoogste coëfficiënt bepalen, et cetera. Dit geeft de vorm $R(x)$ met coëfficiënten $\boldsymbol{F}_i$, zie kader 'Lineaire algebra'.

Complexiteit en stabiliteit

We kennen nu vijf vormen van het interpolatiepolynoom: N, L, V, H en R, waarvan de laatste drie de coëfficiënten geven van het polynoom in standaardvorm. Al deze vormen heb ik geïmplementeerd in mijn Javaprogramma. Hoewel alle vormen in principe hetzelfde resultaat geven is de manier van uitrekenen steeds anders. Dit heeft gevolgen voor de *orde* (algoritmische complexiteit) en de nauwkeurigheid (numerieke stabiliteit) van de berekeningen.

In mijn implementatie hebben alle vormen orde $O(n^2)$ behalve V, deze heeft orde $O(n^3)$. De nauwkeurigheid bepaalde ik met een experiment. (Schets van het experiment: het interpoleren door punten van een functie die slechts te benaderen is door een polynoom. Dit werd voor elke vorm gedaan met 2 tot en met 30 punten. De gevonden afwijkingen ten opzichte van de functie geven in een grafiek op logaritmische basis vrijwel een rechte lijn. De score van een vorm is de richtingscoëfficiënt van de lijn omgerekend naar een getal tussen 0 (extreem onnauwkeurig) en 100 (perfect).) Uit dit experiment blijkt dat R snel numeriek instabiel wordt en daarom is deze vorm verder uitgesloten. Het blijkt dat L het meest nauwkeurig is (score 96 uit 100), gevolgd door N (88), H (72) en V (58). L is dus de beste vorm als nauwkeurigheid belangrijk is en er veel interpolatiepunten zijn. N scoort redelijk op deze punten en heeft een ander voordeel: er kunnen gemakkelijk punten worden toegevoegd zonder dat de hele berekening opnieuw gedaan hoeft te worden. H, ten slotte, is de beste vorm die de coëfficiënten van het polynoom in standaardvorm geeft. H is nauwkeuriger en efficiënter dan V, en R is te onnauwkeurig.

Tot slot

Na afronding van mijn profielwerkstuk kwam ik in contact met wiskundigen die mij op het spoor van de lineaire algebra zetten. Door zo naar het vraagstuk te kijken werd het verband tussen de gevonden formules duidelijk. De uitwerking daarvan wordt gegeven in het kader 'Lineaire algebra' op de volgende pagina. De betere nauwkeurigheid van de interpolatiepolynomen L en N is waarschijnlijk de reden dat deze vormen wel algemeen bekend zijn en de coëfficiëntgerichte vormen V, H en R niet. V en H kunnen worden beschouwd als respectievelijk L en N in standaardvorm gebracht, met behulp van een bekende eigenschap van de elementaire symmetrische polynomen. Maar ze kunnen dus ook direct worden afgeleid met elementaire algebra of lineaire algebra.

Dankwoord

Ik wil de volgende mensen bedanken die mij verder hebben geholpen in mijn proces: Gerrit de Bruijn ('t Atrium Amersfoort) voor het aangeven hoe de eerste paar coëfficiënten van V uit L kunnen worden afgeleid; Hennie ter Morsche en Wil Schilders (TU Eindhoven) voor de verwijzing naar het boek van Norlund [1] en de aanwijzing dat L en V uit de daar gegeven representatie kunnen worden afgeleid door ontwikkeling naar de laatste kolom of rij; Rob Tijdeman (Universiteit Leiden) voor zijn bemiddeling en het proeflezen van dit artikel; en ten slotte mijn vader Jan-Jaap van Horssen voor zijn rol als sparringpartner.

Lineaire algebra

We gebruiken lineaire algebra om de relaties tussen de verschillende vormen van het interpolatiepolynoom te beschrijven. We schrijven $P(x) = \sum_{i=0}^{n} \lambda_i x^i$, $\boldsymbol{\Lambda} = (\lambda_0, \ldots, \lambda_n)^T$ en $\boldsymbol{Y} = (y_0, \ldots, y_n)^T$ en krijgen zo de vergelijking

$$\boldsymbol{V\Lambda} = \boldsymbol{Y}.$$

We definiëren de matrix $\boldsymbol{A}$:

$$\boldsymbol{A} := \begin{bmatrix} \boldsymbol{V} & \boldsymbol{Y} \\ \boldsymbol{X} & P(x) \end{bmatrix}$$

met $\boldsymbol{X} = (1, x, \ldots, x^n)$. Vanwege de lineaire afhankelijkheid van de kolommen volgt dat $|\boldsymbol{A}| = 0$. Hieruit volgt [1]:

$$P(x) = -\frac{|\boldsymbol{A}_0|}{|\boldsymbol{V}|} \tag{1}$$

met

$$\boldsymbol{A}_0 := \begin{bmatrix} \boldsymbol{V} & \boldsymbol{Y} \\ \boldsymbol{X} & 0 \end{bmatrix}.$$

Als we de teller in (1) naar de laatste kolom ontwikkelen, vinden we het bekende Lagrange-polynoom

$$L(x) = \sum_{i=0}^{n} y_i \prod_{j=0, j\neq i}^{n} \frac{x - x_j}{x_i - x_j}.$$

Ontwikkelen we in plaats daarvan naar de laatste rij dan verschijnen de elementaire symmetrische polynomen en vinden we

$$V(x) = \sum_{i=0}^{n} (-1)^{n-i} \sum_{j=0}^{n} y_j \frac{e_{n-i}(x_0, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n)}{\prod_{k=0, k\neq j}^{n} (x_j - x_k)} x^i.$$

We kunnen het stelsel $\boldsymbol{VY}$ in $\boldsymbol{A}$ ook met Gauss-eliminatie oplossen, dit geeft het stelsel $\boldsymbol{HD}$. Hierin is $\boldsymbol{H}$ de bovendriehoeksvorm van $\boldsymbol{V}$. Dit geeft een matrix

$$\widehat{\boldsymbol{A}} := \begin{bmatrix} \boldsymbol{H} & \boldsymbol{D} \\ \boldsymbol{X} & P(x) \end{bmatrix}$$

waar $\boldsymbol{D} = (f[x_0], \ldots, f[x_0, \ldots, x_n])^T$ en de matrix $\boldsymbol{H}$ bestaat uit de compleet homogene symmetrische polynomen:

$$\boldsymbol{H}_{i,j} := h_{j-i}(x_0, \ldots, x_i).$$

Analoog aan het voorgaande volgt:

$$P(x) = -\frac{|\widehat{\boldsymbol{A}}_0|}{|\boldsymbol{H}|} \tag{2}$$

met

$$\widehat{\boldsymbol{A}}_0 := \begin{bmatrix} \boldsymbol{H} & \boldsymbol{D} \\ \boldsymbol{X} & 0 \end{bmatrix}.$$

Ontwikkelen we de teller in (2) naar de laatste kolom (er geldt $|\boldsymbol{H}| = 1$) dan vinden we het bekende Newton-polynoom

$$N(x) = \sum_{i=0}^{n} f[x_0, \ldots, x_i] \prod_{j=0}^{i-1} (x - x_j).$$

Ontwikkelen we in plaats daarvan naar de laatste rij dan vinden we

$$H(x) = \sum_{i=0}^{n} \sum_{j=i}^{n} f[x_0, \ldots, x_j] (-1)^{j-i} e_{j-i}(x_0, \ldots, x_{j-1}) x^i.$$

Ten slotte geven we de representatie die ontstaat door in de matrix $\boldsymbol{A}$ het stelsel $\boldsymbol{VY}$ met Gauss–Jordan-eliminatie op te lossen, dit geeft het stelsel $\boldsymbol{IF}$. Zo krijgen we de matrix

$$\widetilde{\boldsymbol{A}} := \begin{bmatrix} \boldsymbol{I} & \boldsymbol{F} \\ \boldsymbol{X} & P(x) \end{bmatrix},$$

waarbij de vector $\boldsymbol{F}$ recursief (van n naar 0) gedefinieerd wordt door:

$$\boldsymbol{F}_n := f[x_0, \ldots, x_n],$$
$$\boldsymbol{F}_i := f[x_0, \ldots, x_i] - \sum_{k=1}^{n-i} h_k(x_0, \ldots, x_i) \boldsymbol{F}_{i+k}.$$

Analoog aan het voorgaande vinden we:

$$P(x) = -\frac{|\widetilde{\boldsymbol{A}}_0|}{|\boldsymbol{I}|} \tag{3}$$

met

$$\widetilde{\boldsymbol{A}}_0 := \begin{bmatrix} \boldsymbol{I} & \boldsymbol{F} \\ \boldsymbol{X} & 0 \end{bmatrix}.$$

Ontwikkelen we de teller in (3) naar de laatste kolom óf naar de laatste rij (dat maakt in dit geval geen verschil) dan vinden we:

$$R(x) = \sum_{i=0}^{n} \boldsymbol{F}_i x^i.$$

We hebben nu vijf representaties van het interpolatiepolynoom afgeleid. Hiervan zijn zoals gezegd het Lagrange-polynoom en het Newton-polynoom algemeen bekend. Deze geven niet direct de coëfficiënten $\boldsymbol{\Lambda}$ van het polynoom in standaardvorm. De andere drie representaties (V, H en R) geven elk op een andere wijze de coëfficiënten van de standaardvorm van het interpolatiepolynoom weer.

$$\begin{tikzcd}
\boxed{L = \mathcal{L}_{\boldsymbol{Y}}(\boldsymbol{A}_0)} \arrow[rr, leftrightarrow] \arrow[d, leftrightarrow] & & \boxed{V = \mathcal{L}_{\boldsymbol{X}}(\boldsymbol{A}_0)} \arrow[d, leftrightarrow] \\
\boxed{N = \mathcal{L}_{\boldsymbol{D}}(\widehat{\boldsymbol{A}}_0)} \arrow[rr, leftrightarrow] \arrow[dr, leftrightarrow] & & \boxed{H = \mathcal{L}_{\boldsymbol{X}}(\widehat{\boldsymbol{A}}_0)} \arrow[dl, leftrightarrow] \\
 & \boxed{R = \mathcal{L}_{\boldsymbol{F}/\boldsymbol{X}}(\widetilde{\boldsymbol{A}}_0)} &
\end{tikzcd}$$

Figuur 2 Taxonomie. $\mathcal{L}_{\boldsymbol{W}}(\boldsymbol{M})$ betekent de Laplace-expansie van de matrix $\boldsymbol{M}$ naar de kolom of de rij $\boldsymbol{W}$. De pijlen geven relatief eenvoudige algebraïsche transformaties aan.

Referentie

1 N.E. Nörlund, Vorlesungen über Differenzenrechnung, *Grundlehren der mathematischen Wissenschaften* 13 (1924).