Snelle priemgetaltests

Lenstra, H.W.

Snelle priemgetaltests

Hendrik W. LENSTRA, Jr.
Universiteit van Amsterdam

Deze gehele voordracht zal n een geheel getal groter dan 1 aangeven.
We noemen n priem als er geen gehele getallen a, b bestaan met

$$n = a \cdot b,\ a > 1,\ b > 1.$$

Anders heet n samengesteld. Twee fundamentele problemen uit de elementaire getaltheorie zijn :
A. *Hoe kan men snel zien of een gegeven getal n priem is ?*
B. *Als n niet priem is, hoe vindt men gehele getallen $a > 1$, $b > 1$ zodat $n = a \cdot b$?*

Hoewel deze problemen erg op elkaar lijken zijn de wiskundige technieken die men voor de beantwoording gebruikt totaal verschillend.

Ik zal mij in deze voordracht bijna uitsluitend met probleem A bezighouden. Recente vorderingen ten aanzien van dit probleem zijn gemaakt door de Amerikaanse wiskundigen L.M. Adleman en R.S. Rumely. Een vereenvoudigde versie van hun methode, ontwikkeld in samenwerking met H. Cohen (Bordeaux), is geprogrammeerd voor het CDC Cyber 170-750 computersysteem van het SARA-rekencentrum te Amsterdam ; een en ander met medewerking van D.T. Winter en A.K. Lenstra.

Voordat ik inga op de theorie die aan het programma ten grondslag ligt bespreek ik enkele getallenvoorbeelden. Deze zijn op zich niet meer dan curiositeiten, maar geven toch een idee van wat op het ogenblik wel en niet kan worden gedaan.

Het getal

$$10^{100} + 267 = 100\ldots00267$$

($97\times$)

is het kleinste priemgetal van 101 cijfers. Ons programma wist binnen 42,8 seconden te

bewijzen dat het priem was. Dit is een "typische" rekentijd voor getallen van dezelfde orde van grootte. Oudere programma's waren niet in staat te beslissen of $10^{100} + 267$ priem is.

Een typisch priemgetal van 200 cijfers zal, als ons programma daarvoor geschikt gemaakt wordt, ongeveer 6 minuten in beslag nemen.

Het getal

$$\frac{10^{1031} - 1}{9} = \underbrace{111 \ldots 111}_{1031 \text{ x}}$$

is hoogstwaarschijnlijk priem ; wat dat precies betekent zullen we verderop merken. Een ruwe berekening leert dat onze methode op dezelfde machine voor dit getal ongeveer één week nodig zou hebben.

Voor getallen die een speciale vorm hebben kan men veel verder gaan. Zo werd door D. Slowinski bewezen dat het getal

$$2^{44497} - 1 = 854 \ldots 671 \quad (13395 \text{ cijfers})$$

priem is, en hij had hier minder dan een uur rekentijd (op een CRAY-1) voor nodig. Dit is op het ogenblik het grootst bekende priemgetal.

Al deze voorbeelden betreffen probleem A. Probleem B is veel lastiger. Getallen van 40 à 50 cijfers kan men op het ogenblik met de beste bekende methoden meestal wel binnen een paar uur in factoren ontbinden, maar van een getal als

$$2^{227} - 1 = 215 \ldots 727,$$

dat 69 cijfers heeft, is geen enkele priemfactor bekend ; wel weet men dat het getal zelf niet priem is.

Het is misschien verbazend dat men van een getal zeker kan weten dat het niet priem is zonder er een factor van te kennen. Dit is doorgaans te danken aan de volgende stelling of één van zijn varianten :

Stelling van Fermat (Pierre de Fermat, 1601-1665).

$$n \text{ priem} \Rightarrow \quad \forall\, a \in \mathbb{Z} \quad : a^n \equiv a \bmod n.$$

Bij deze stelling valt op te merken dat het ook voor zeer grote n gemakkelijk is, althans op een computer, om na te gaan of de congruentie $a^n \equiv a \bmod n$ inderdaad geldt. Dit doet men door niet a^n zelf uit te rekenen (dat zou zelfs voor een snelle computer ondoenlijk zijn als $n \approx 10^{100}$ en $a = 2$), maar alleen de rest van a^n bij deling door n ; en deze rest valt te berekenen door middel van een reeks herhaalde kwadrateringen en vermenigvuldigingen modulo n.

Vindt men ook maar één $a \in \mathbb{Z}$ waarvoor de congruentie $a^n \equiv a \bmod n$ niet geldt, dan weet men zeker dat n niet priem is, zonder nochtans een factor van n te kennen.

Willen we bewijzen dat een getal n wél priem is, dan hebben we een omkering van de stelling van Fermat nodig. Hier doen zich twee moeilijkheden voor.

I. *Ten eerste is de directe omkering, waarin "⇒" vervangen is door "⇐", foutief : het getal* $n = 1729 = 7 \cdot 13 \cdot 19$ is niet priem, maar toch geldt

$$a^{1729} \equiv a \bmod 1729$$

voor alle $a \in \mathbb{Z}$.

II. Ten tweede, zelfs als de omkering waar was dan zou dat ons niet veel helpen, want het is volledig ondoenlijk alle gehele getallen a af te proberen.

Hoe kunnen we deze moeilijkheden overwinnen ?

De eerste moeilijkheid wordt overwonnen door scherpere versies van de stelling van Fermat te beschouwen, waarvan de omkering wél correct is.

We beginnen met een algebraïsche verscherping :

Stelling. Als n priem is, dan geldt voor elke commutatieve ring R en alle $a, b \in R$ de congruentie

$$(a + b)^n \equiv a^n + b^n \bmod nR.$$

Hier geeft nR het ideaal $\{ x + x + \ldots + x \text{ (n termen)} \mid x \in R \}$ van R aan.

Het bewijs berust op het binomium van Newton en op de opmerking dat de binomiaalcoëfficiënten $\binom{n}{i}$, voor $0 < i < n$, door n deelbaar zijn als n priem is.

Nemen we $R = \mathbb{Z}$ en $b = 1$ dan vinden we, met inductie naar a, de stelling van Fermat terug.

Men kan bewijzen dat van bovenstaande stelling de omkering ook geldt : geldt de congruentie in de stelling voor alle R en alle $a, b \in R$, dan is n ook priem. Het is in feite voldoende voor R de polynoomring $\mathbb{Z}[X]$ te nemen, en $a = X$, $b = 1$.

Voordat we ingaan op een getaltheoretische verscherping van de stelling van Fermat behandelen we enkele eigenschappen van het Jacobisymbool $\left(\frac{a}{n}\right)$; voor meer bijzonderheden zie men de leerboeken, zoals [2].

Vanaf nu nemen we aan dat n oneven is. Als n priem is, dan volgt uit de stelling van Fermat dat

$$a^{n-1} \equiv 1 \bmod n \quad \text{voor} \quad a \in \mathbb{Z} \ , \text{ggd}(a, n) = 1$$

dus

$$a^{(n-1)/2} \equiv 1 \quad \text{of} \quad -1 \bmod n.$$

Het Jacobisymbool $(\frac{a}{n}) \in \{ 1, -1 \}$ is voor $a \in \mathbb{Z}$, ggd $(a, n) = 1$ nu gedefinieerd door

$(\frac{a}{n}) \equiv a^{(n-1)/2} \bmod n$ als n priem is,

$(\frac{a}{n}) = (\frac{a}{p_1})(\frac{a}{p_2}) \ldots (\frac{a}{p_t})$ als $n = p_1 p_2 \ldots p_t$, p_j priem.

Voor het gemak zetten we $(\frac{a}{n}) = 0$ als ggd $(a, n) > 1$.

Het Jacobisymbool is onderwerp van een door Gauss ontwikkelde theorie, waarvan het middelpunt wordt gevormd door de kwadratische reciprociteitswet, zie [2]. Onder gebruikmaking van deze wet kan men - en dat is voor het volgende van belang om te weten - het Jacobisymbool $(\frac{a}{n})$ snel berekenen, ook zonder dat men de ontbinding van n in priemfactoren kent.

Uit de definitie van $(\frac{a}{n})$ volgt direct de volgende verscherping van de stelling van Fermat :

Stelling. Er geldt :

$$n \text{ priem} \Rightarrow (\quad \forall\, a \in \mathbb{Z} \ : \text{ggd}(a, n) = 1 \Rightarrow a^{(n-1)/2} \equiv (\frac{a}{n}) \bmod n$$

Opnieuw is het, ook voor zeer grote n, op een computer gemakkelijk om na te gaan of voor een gegeven a de congruentie $a^{(n-1)/2} \equiv (\frac{a}{n}) \bmod n$ inderdaad geldt.

Door D.H. Lehmer werd in 1976 bewezen dat van bovenstaande stelling ook de omkering geldt. Preciezer :

Stelling. Als n een oneven samengesteld getal is, dan geldt

$$a^{(n-1)/2} \not\equiv (\frac{a}{n}) \bmod n$$

voor ten minste de helft van alle $a \in \{ 1, 2, \ldots, n-1 \}$ met ggd $(a, n) = 1$.

Met deze resultaten is moeilijkheid I op bevredigende wijze opgelost. In de daadwerkelijke priemgetaltest werken we in feite met een combinatie van beide benaderingen : we maken gebruik van congruenties in uitbreidingsringen van $\mathbb{Z}$, en deze congruenties bevatten een symbool dat het Jacobisymbool veralgemeent.

Met moeilijkheid II blijven we zitten : het is nog steeds niet doenlijk alle a (mod n) te proberen, en helemaal niet om alle ringen R na te gaan.

De eerste methode die we bespreken om moeilijkheid II op te lossen is van probabilistische aard. Hij werkt als volgt. Trek onafhankelijk honderd willekeurige getallen a uit $\{1, 2, \ldots, n-1 \}$, en test voor elk van de getrokken waarden of

 $a^{(n-1)/2} \equiv (\frac{a}{n}) = \pm 1 \bmod n$

geldt. Als er ook maar één a is waarvoor het fout is, dan weet men zeker dat n samengesteld is. Is het evenwel goed voor alle honderd waarden van a, dan is n hoogstwaarschijnlijk priem.

Stel immers dat n niet priem is ; wegens de stelling van Lehmer heeft elke a dan een kans

$\leqslant \frac{1}{2}$ om aan bovenstaande congruentie te voldoen, dus de kans dat alle honderd eraan voldoen is

$$\leqslant (\frac{1}{2})^{100} < 0{,}00000000000000000000000000001.$$

Het is dus erg moeilijk te betwijfelen dat n priem is.

Deze methode, voorgesteld door R. Solovay en V. Strassen, is principieel niet in staat wiskundig sluitende bewijzen te leveren dat een getal priem is. Aan de andere kant is het wel een zeer snelle methode, en men kan hem altijd toepassen vóórdat men de tijdrovende methode die ik straks bespreek toepast.

Door G.L. Miller en M.O. Rabin is een verbetering van bovenstaande test voorgesteld die met minder dan de helft van het rekenwerk dezelfde foutenkans bereikt.

De tweede methode om moeilijkheid II op te lossen is van futuristische aard, aangezien hij berust op een nog onbewezen vermoeden, de gegeneraliseerde Riemannhypothese. Dit is een uitspraak over de ligging van de nulpunten van bepaalde in de analytische getaltheorie optredende complexe functies, waarvan de correctheid vrij algemeen aangenomen wordt, maar die tot op heden nog niet is bewezen.

Onder aanname van dit vermoeden werd door G.L. Miller in 1976 bewezen dat het in de bovenbeschreven test voldoende is alleen de getallen a met

$$a < 70\ (\log n)^2, \quad a \text{ priem}$$

te testen. Voldoen al deze waarden van a aan de congruentie $a^{(n-1)/2} \equiv (\frac{a}{n}) \bmod n$, dan kan men met behulp van de gegeneraliseerde Riemannhypothese aantonen dat n inderdaad priem is.

Aan deze methode kleven twee bezwaren. Ten eerste leert een ruwe schatting, dat hij voor getallen van omstreeks 100 cijfers ongeveer 500 keer zoveel tijd kost als de hierna te bespreken methode, hoewel hij voor veel grotere n sneller is. Ten tweede is de gegeneraliseerde Riemannhypothese nog niet bewezen.

We wenden ons nu tot de methoden die in staat zijn met wiskundige zekerheid te bewijzen dat een getal priem is. Er bestaan er meer dan ik in deze voordracht de tijd heb om op te noemen, maar ze berusten alle op een of andere manier op de volgende stelling, die zo triviaal is dat ik hem naar mijzelf genoemd heb.

Stelling van Lenstra.

n is priem $\Leftrightarrow$ elke deler r van n is een macht van n.

Om $\Rightarrow$ te bewijzen is het voldoende op te merken dat $1 = n^0$ en $n = n^1$. Het bewijs van $\Leftarrow$ laat ik als opgave aan het gehoor over.

Welke rol speelt deze stelling bij priemgetaltests ? Onnauwkeurig geformuleerd komt het op het volgende neer. Stel dat n aan vele voorwaarden van het soort

$$a^{(n-1)/2} \equiv (\frac{a}{n}) \bmod n$$

voldoet. Dan kan men aantonen dat iedere deler r van n zich in een bepaald opzicht gedraagt alsof hij een macht van n is. In zekere omstandigheden kan men daaruit afleiden dat 1 en n de enige delers van n zijn, zodat n inderdaad priem is.

De volgende stelling kan dienen om deze gang van zaken te illustreren. We nemen vanaf nu ook aan dat n niet deelbaar door 3 is.

Stelling. Als

$$a^{(n-1)/2} \equiv (\frac{a}{n}) \bmod n \quad \text{voor } a = -1, 2 \text{ en } 3,$$

en als er bovendien een $a \in \mathbb{Z}$ is waarvoor geldt

$$a^{(n-1)/2} \equiv -1 \bmod n,$$

dan is er voor elke deler r van n een geheel getal $i \geqslant 0$ met

$$r \equiv n^i \bmod 24.$$

De voorwaarde $a^{(n-1)/2} \equiv -1 \bmod n$, voor zekere $a \in \mathbb{Z}$, kan niet worden weggelaten, zoals uit het eerder gegeven voorbeeld n = 1729 blijkt. Als n daadwerkelijk priem is, dan is het meestal gemakkelijk een geheel getal a te vinden dat inderdaad aan deze voorwaarde voldoet.

De stelling zegt, dat iedere deler r van n zich "in een bepaald opzicht" (namelijk : modulo 24) als een macht n^i van n gedraagt. In feite kunnen we i = 0 of i = 1 nemen, want $n^2 \equiv 1 \bmod 24$.

Het bewijs van de stelling geef ik hier niet. Het berust op de volgende bewering, geldig voor positieve gehele getallen n_1 en n_2 die niet deelbaar zijn door 2 of 3 :

$$n_1 \equiv n_2 \bmod 24 \Leftrightarrow (\frac{a}{n_1}) = (\frac{a}{n_2}) \text{ voor } a = -1, 2 \text{ en } 3.$$

Deze bewering is een gevolg van de eerder genoemde kwadratische reciprociteitswet.

De bovengeformuleerde stelling is voor priemgetaltests niet bijzonder nuttig, omdat we met

de conclusie verder niets kunnen aanvangen
Het zou veel nuttiger zijn een dergelijke stelling te hebben waarin het getal 24 door een aanzienlijk groter getal vervangen is Analyseert men het hier niet gegeven bewijs van de stelling dan ontdekt men dat de eigenschap van het getal 24 die de stelling mogelijk maakt de volgende is

$$m^2 \equiv 1 \bmod 24 \text{ voor iedere } m \in \mathbb{Z} \quad \text{met ggd } (m, 24) = 1$$

Men kan aantonen dat 24 het grootste getal met deze eigenschap is Wil men 24 door een groter getal vervangen dan zal men in plaats van naar kwadraten naar hogere machten moeten kijken, en dat voert tot het beschouwen van symbolen die het Jacobisym eralgemenen

Vervangen we kwadraten door twaalfde machten, dan vinden we al dat 24 nzienlijk ver groot kan worden

$$m^{12} \equiv 1 \bmod 65520 \text{ voor iedere } m \in \mathbb{Z} \quad \text{met ggd } (m, 65520) = 1$$

Hier is $65520 = 2^4 \quad 3^2 \quad 5 \quad 7 \quad 13$

In het aan het begin van deze voordracht genoemde computerprogramma worden nog veel grotere getallen gebruikt

$$m^{5040} - 1 \bmod s \text{ voor iedere } m \in \mathbb{Z} \quad \text{met ggd } (m, s) = 1$$

Hier is $5040 = 2^4 \quad 3^2 \quad 5 \quad 7$, en s is een getal van 53 cijfers

$$s = 153219867888544432846626127356636113800104312257712 00$$

$$= 2^6 \quad 3^3 \quad 5^2 \quad 7^2 \quad 11 \quad 13 \quad 17 \quad 19 \quad 29 \quad 31 \quad 37 \quad 41 \quad 43 \quad 61 \quad 71 \quad 73 \quad 113 \quad 127$$

$$181 \quad 211 \quad 241 \quad 281 \quad 337 \quad 421 \quad 631 \quad 1009 \quad 2521$$

Het is voor ons van belang dat $s > \sqrt{n}$ als n niet meer dan 100 cijfers heeft

Ik geef nu een schetsmatige en niet al te nauwkeurige beschrijving van de door ons gebruikte priemgetaltest Voor meer bijzonderheden en verwijzingen naar de literatuur raadplege men [1]

Priemgetaltest voor $n < 10^{100}$
Stap 1 Test of de grootste gemene deler van n en s gelijk is aan 1, met s als hierboven (deze grootste gemene deler kan men met de algoritme van Euclides bepalen) Is dit niet het geval dan is n deelbaar door een van de priemfactoren van s, en we rekenen niet verder

Stap 2 Test 67 congruenties die elk analoog zijn aan

$$a^{(n-1)/2} \equiv \left(\frac{a}{n}\right) \bmod n$$

maar met a vervangen door geschikt gekozen elementen van de ringen

$$\mathbb{Z}\,[e^{2\pi i/p^k}],\quad \text{met p priem, } k \geqslant 1 \text{ en } p^k \text{ een deler van } 5040 = 2^4 \,.\, 3^2 \,.\, 5 \,.\, 7,$$

en met $(\frac{a}{n})$ vervangen door een gegeneraliseerd symbool, waarvan de waarden machten van $e^{2\pi i/p^k}$ zijn (voor $p^k = 2$ geldt $e^{2\pi i/p^k} = -1$, en dan krijgen we het Jacobisymbool zelf).

Indien ten minste één van de 67 congruenties niet vervuld is, dan is n samengesteld, en men stopt.

Stel nu dat alle 67 congruenties inderdaad gelden. De congruenties zijn zó geselecteerd, dat men dan kan bewijzen dat er voor elke deler r van n een $i \in \mathbb{Z}$ is met

$$r \equiv n^i \bmod s, \quad 0 \leqslant i < 5040.$$

Stap 3. Uitgaande van deze informatie gaan we alle delers r van n met $r \leqslant \sqrt{n}$ opsporen. Dit is kennelijk voldoende om n in factoren te ontbinden en dus om te zien of n priem is.

Uit $r \leqslant \sqrt{n}$ volgt $r < s$, dus r is volledig bepaald als we r modulo s weten. Wegens de informatie uit stap 2 kunnen we daarom zó te werk gaan : bepaal, voor elke i = 0, 1, ..., 5039, het getal r_i dat voldoet aan

$$r_i \equiv n^i \bmod s, \quad 0 \leqslant r_i < s.$$

Alle delers $\leqslant \sqrt{n}$ van n komen dan onder de r_i voor, dus we kunnen de algoritme besluiten met 5040 testdelingen.

Uit bovenstaande beschrijving van stap 3 make men niet op dat de algoritme ook kan helpen bij het in factoren ontbinden van getallen. In de praktijk zullen namelijk alle samengestelde getallen n al in stap 2 (of stap 1) ontdekt worden.

Ik besluit met een speciaal voor de VVWL-studiedag vervaardigd getal van honderd cijfers waarvan ons programma in 33,573 seconden bewees dat het priem is :

221201011319050022051805140907091407002309191121 14

040512051801011819001100040503051302051800198200 29.

BIBLIOGRAFIE

1. H. COHEN and H.W. LENSTRA, Jr., *Primality testing and Jacobi sums,* Report 82-18, Mathematisch Instituut, Universiteit van Amsterdam, 1982.
2. A. SCHOLZ und B. SCHOENEBERG, *Einführung in die Zahlentheorie,* Sammlung Göschen Bd. 1131, Walter de Gruyter & Co, Berlin, 1955.