

De Approximatiestelling van Weierstraß

Korteweg-de Vries Instituut voor Wiskunde

Universiteit van Amsterdam

Mastercourse 15 november 2005

Peter Spreij

spreij@science.uva.nl

1 Introductie

In deze mastercourse behandelen we de *Approximatiestelling van Weierstraß*. Deze stelling zegt dat we een gegeven continue functie op $[0, 1]$ willekeurig goed benaderen met behulp van polynomen en dat deze benaderingen uniform zijn. Sjiaker gezegd: de klasse van polynomen ligt dicht in de ruimte van continue functies met de sup-norm. We zullen dit nader gaan preciseren.

In veel analyseboeken is het bewijs van deze stelling niet constructief. We zullen verderop zien dat de Kansrekening uitkomst biedt en precies polynomen beschrijft die een gegeven continue functie benaderen. De ingrediënten die we gaan gebruiken zijn de Binomiale verdeling, de ongelijkheid van Chebychev en de Zwakke Wet van de Grote Aantallen, evenals wat eigenschappen van continue functies.

2 Over continue functies

Eerst halen we op wat een continue functie f is. In gewone taal betekent dit dat de waarden $f(x)$ weinig kunnen veranderen, als we x maar weinig variëren. Dit is geen wiskundige uitspraak en bovendien een bron van verwarring, die er in de historie toe heeft geleid dat pas na verloop van tijd onderscheid werd gemaakt tussen continue functies en *uniform* continue functies.

We herinneren aan de definitie van een continue functie $f : I \rightarrow \mathbb{R}$. De verzameling I is hier een interval, dat open, gesloten, half-open, etc. kan zijn, en ook nog \mathbb{R} zelf. We noemen $f : I \rightarrow \mathbb{R}$ continu, als voor alle $x \in I$ en alle $\varepsilon > 0$ een getal $\delta > 0$ te vinden is zodanig dat voor alle $y \in I$ met $|y - x| < \delta$ geldt dat $|f(y) - f(x)| < \varepsilon$.

Duidelijk zal zijn dat deze definitie apelleert aan onze intuïtieve voorstelling van continuïteit. Belangrijk om op te merken is dat de te vinden δ in het algemeen afhangt van de voorgeschreven ε en van het punt x dat we beschouwen. Zie opgave 7.1.

Er zijn echter (heel wat) continue functies waarbij de δ niet van x afhangt (afhankelijkheid van ε zal i.h.a. natuurlijk gehandhaafd blijven). In dat geval geldt de volgende verscherping van continuïteit: voor alle $\varepsilon > 0$ is er een $\delta > 0$ te vinden zodanig dat voor alle x, y met $|y - x| < \delta$ geldt dat $|f(y) - f(x)| < \varepsilon$. Dit heet *uniforme continuïteit*.

De volgende stelling uit de Analyse is van fundamenteel belang. Deze stelling zullen we voor kennisgeving aannemen. We hebben de stelling nodig voor het bewijs in paragraaf 6.

Stelling 2.1 *Zij I een gesloten en begrensde interval en $f : I \rightarrow \mathbb{R}$ continu. Dan is f begrensd en zelfs uniform continu.*

3 De Binomiale verdeling

We herhalen n keer een experiment dat alleen de uitkomsten ‘succes’ en ‘mislukking’ kent, aangegeven met een 1 of een 0. Zij S_n het totaal aantal successen in deze rij. Denk hierbij aan het herhaald opgooien van een munt. Het bovenkomen van ‘munt’ beschouwen we als een succes en S_n vertelt ons dan hoe vaak dit gebeurt.

We veronderstellen dat elke keer de kans dat de uitkomst ‘succes’ optreedt, gelijk is aan $p \in [0, 1]$ (en de kans op mislukking $1 - p$). We veronderstellen bovendien dat de experimenten onafhankelijk zijn. Dan heeft de stochastisch variabele S_n een binomiale verdeling, aangeduid met $\text{Bin}(n, p)$. Dat wil zeggen dat

$$\mathbb{P}(S_n = k) = \binom{n}{k} p^k (1 - p)^{n-k}, \quad k \in \{0, \dots, n\}.$$

Schrijven we $S_n = X_1 + \dots + X_n$ met X_i de uitkomst van het i -de experiment, dan volgen een aantal karakteristieken eenvoudig. We merken eerst op dat de verwachtingen $\mathbb{E}X_i$ alle gelijk aan p zijn. Omdat $X_i^2 = X_i$ is dan bovendien $\mathbb{E}(X_i^2) = p$. Maar dan is $\text{Var } X_i = \mathbb{E}(X_i^2) - (\mathbb{E}X_i)^2 = p - p^2$. Onmiddellijk volgt dan dat $\mathbb{E}S_n = np$. Voorts hebben we de cruciale observatie dat de X_i stochastisch onafhankelijk zijn. Een bekende regel zegt dat daarom

$$\text{Var } S_n = \text{Var } X_1 + \dots + \text{Var } X_n = n(p - p^2).$$

Omdat $p - p^2$ maximaal is voor $p = 1/2$, en dan gelijk aan $1/4$, hebben we bovendien dat

$$\text{Var } S_n \leq n/4$$

voor alle $p \in [0, 1]$.

4 Enkele hulpproblematies

In deze paragraaf vermelden we een lijstje van bekende ongelijkheden en andere feitjes die we in de paragrafen 5 en 6 zullen gebruiken. Tevens introduceren we enige extra notatie.

driehoeksongelijkheid Voor alle $x, y \in \mathbb{R}$ geldt $|x + y| \leq |x| + |y|$.

driehoeksongelijkheid voor verwachting Als X een stochastische variabele is met eindige verwachting, dan geldt $|\mathbb{E}X| \leq \mathbb{E}|X|$.

indicator functie Voor een verzameling A definiëren we de functie 1_A door $1_A(x) = 1$ als $x \in A$ en $1_A(x) = 0$ als $x \notin A$.

verwachting van een indicator functie Als A een gebeurtenis is, dan is $X = 1_A$ een stochastische variabele die alleen de waarden 0 en 1 aanneemt. De laatste met kans $\mathbb{P}(A)$. Maar dan is $\mathbb{E}X = \mathbb{E}1_A = \mathbb{P}(A)$.

5 De zwakke wet van de grote aantallen

Gevoelsmatig denken we, dat in het experiment met de munt uit paragraaf 3 de fractie $\frac{S_n}{n}$ van het aantal keren dat ‘munt’ valt, in de buurt van p ligt, als n heel groot is. In een wiskundige vertaling hiervan komt het erop neer dat de kans dat $\frac{S_n}{n}$ een klein beetje van p afwijkt heel klein zal zijn. Dit staat bekend als de ‘zwakke wet van de grote aantallen’. Dit wordt precies gemaakt in de volgende propositie.

Propositie 5.1 *Zij $\delta > 0$ willekeurig. Dan geldt dat*

$$\mathbb{P}\left(\left|\frac{S_n}{n} - p\right| \geq \delta\right) \rightarrow 0, \text{ als } n \rightarrow \infty.$$

Het bewijs van deze propositie berust op de ongelijkheid van Chebychev. Deze vermelden we in algemene vorm in het volgende lemma.

Lemma 5.2 *Zij Y een stochast met eindige variantie σ^2 . Zij μ de verwachting van Y . Dan geldt*

$$\mathbb{P}(|Y - \mu| \geq \eta) \leq \frac{\sigma^2}{\eta^2}.$$

Bewijs Het bewijs van dit lemma berust op het achter elkaar zetten van wat trivialiteiten. We schrijven E voor de gebeurtenis $\{|Y - \mu| < \eta\}$. Bekijk

$$\begin{aligned} (Y - \mu)^2 &= 1_E(Y - \mu)^2 + 1_{E^c}(Y - \mu)^2 \\ &\geq 1_{E^c}(Y - \mu)^2 \\ &\geq 1_{E^c}\eta^2. \end{aligned}$$

Nemen we nu verwachtingen, dan volgt $\mathbb{E}(Y - \mu)^2 \geq \eta^2\mathbb{P}(E^c)$. Omdat $\sigma^2 = \text{Var } Y$ gedefinieerd is als $\mathbb{E}(Y - \mu)^2$, volgt de bewering met wat heen en weer schuiven. \square

Bewijs van propositie 5.1 We passen eenvoudig het lemma toe met $Y = S_n$. Dan is $\mu = np$ en $\sigma^2 = n(p-p^2)$. Voorts is $\{|\frac{S_n}{n} - p| \geq \delta\} = \{|S_n - \mu| \geq n\delta\}$. De kans hierop is volgens het lemma (neem $\eta = n\delta$) ten hoogste gelijk aan $\frac{n(p-p^2)}{n^2\delta^2} = \frac{(p-p^2)}{n\delta^2}$. Voor $n \rightarrow \infty$ gaat dit naar nul. \square

Voor het doel van deze mastercourse hebben we de expliciete formulering als Zwakke Wet niet nodig. Wel lichten we uit het bewijs de belangrijke ongelijkheid $\mathbb{P}(|\frac{S_n}{n} - p| > \delta) \leq \frac{(p-p^2)}{n\delta^2}$, die we via $p - p^2 \leq 1/4$ nog kunnen omvormen tot

$$\mathbb{P}(|\frac{S_n}{n} - p| > \delta) \leq \frac{1}{4n\delta^2}.$$

6 De approximatiestelling van Weierstraß

In het bewijs van de stelling van deze paragraaf maken we gebruik van de hulpresultaten uit paragraaf 4. Zoek zelf uit welke waar gebruikt worden!

Stelling 6.1 (Weierstraß) *Zij $f : [0, 1] \rightarrow \mathbb{R}$ continu en $\varepsilon > 0$. Dan bestaat er een polynoom p zodanig dat*

$$\forall x \in [0, 1] : |f(x) - p(x)| < \varepsilon.$$

Bewijs Omdat we wat Kansrekening gaan gebruiken, vervangen we de variabele x door $p \in [0, 1]$. Zij voor iedere $n \in \mathbb{N}$ de stochast S_n gegeven met een $\text{Bin}(n, p)$ -verdeling. Zie paragraaf 3. Definieer nu

$$B_n(p) = \mathbb{E}f(\frac{S_n}{n}) = \sum_{k=0}^n f(\frac{k}{n}) \binom{n}{k} p^k (1-p)^{n-k}.$$

Zij $\varepsilon > 0$. We zullen laten zien dat we n zo groot kunnen kiezen dat voor alle $p \in [0, 1]$ geldt: $|f(p) - B_n(p)| < \varepsilon$. Aan het werk!

Eerst wat voorbereiding. De functie f is begrensd op $[0, 1]$. Laten we zeggen $|f(x)| \leq K$ voor een $K > 0$ en alle $x \in [0, 1]$. Bovendien is f uniform continu. Zij dan $\varepsilon > 0$ gegeven en kies hierbij een δ zodanig $|f(y) - f(x)| < \varepsilon/2$, zo gauw $|y - x| < \delta$.

We schrijven E_n voor de gebeurtenis $\{|\frac{S_n}{n} - p| < \delta\}$ en merk op dat op deze gebeurtenis geldt dat $|f(p) - f(\frac{S_n}{n})| < \varepsilon/2$. Op het complement van deze gebeurtenis (eigenlijk overal) geldt dat $|f(p) - f(\frac{S_n}{n})| \leq 2K$. In de volgende

ketting van berekeningen gaan we deze feiten gebruiken.

$$\begin{aligned}
|f(p) - B_n(p)| &= |f(p) - \mathbb{E}f\left(\frac{S_n}{n}\right)| \\
&= |\mathbb{E}(f(p) - f\left(\frac{S_n}{n}\right))| \\
&\leq \mathbb{E}|f(p) - f\left(\frac{S_n}{n}\right)| \\
&= \mathbb{E}(|f(p) - f\left(\frac{S_n}{n}\right)|(\mathbb{1}_{E_n} + \mathbb{1}_{E_n^c})) \\
&\leq \mathbb{E}(|f(p) - f\left(\frac{S_n}{n}\right)|\mathbb{1}_{E_n}) + \mathbb{E}(|f(p) - f\left(\frac{S_n}{n}\right)|\mathbb{1}_{E_n^c}) \\
&\leq \mathbb{E}\left(\frac{1}{2}\varepsilon \cdot \mathbb{1}_{E_n}\right) + \mathbb{E}(2K \cdot \mathbb{1}_{E_n^c}) \\
&\leq \frac{1}{2}\varepsilon + 2K \mathbb{E}\mathbb{1}_{E_n^c} \\
&= \frac{1}{2}\varepsilon + 2K \mathbb{P}(E_n^c).
\end{aligned}$$

We concentreren ons nu op de laatste kans. We herinneren ons ineens de ongelijkheid aan het einde van paragraaf 5:

$$\mathbb{P}(E_n^c) = \mathbb{P}\left(\left|\frac{S_n}{n} - p\right| \geq \delta\right) \leq \frac{1}{4n\delta^2}.$$

Kies nu $n > \frac{K}{\delta^2}$. Dan is $2K\mathbb{P}(E_n^c) < \frac{1}{2}\varepsilon$. Maar dan wordt de laatste regel uit het kettinglijstje kleiner dan ε . Samenvattend: $|f(p) - B_n(p)| < \varepsilon$, en het bewijs is geleverd. \square

Tot slot een opmerking. Het polynoom B_n in bovenstaand bewijs wordt ook wel Bernstein polynoom genoemd. Bovendien is dit polynoom nu min of meer expliciet bekend, omdat we voor n (zie het bewijs) de ondergrens $\frac{K}{\delta^2}$ hebben, en K en δ in principe te bepalen zijn uit het functievoorschrift van f . In principe, want vooral het preciseren van δ is niet altijd eenvoudig. Echter, voor een gegeven f en ε is door ‘trial and error’ experimenteel–met de computer–wel een voldoende grote waarde van n te vinden.

7 Opgaven

1. Beschouw $f : (0, 1) \rightarrow \mathbb{R}$, $f(x) = \frac{1}{x}$. Laat zien dat f voldoet aan de definitie van een continue functie met $\delta = \frac{x^2\varepsilon}{1+x\varepsilon}$. Merk op dat δ kleiner wordt met x . Laat aan de hand van een schets zien, dat f niet uniform continu op $(0, 1)$ is.