

Technische Universiteit Delft
Faculteit Elektrotechniek, Wiskunde en Informatica
Delft Institute of Applied Mathematics

**Bewijs van Wigners halve cirkel stelling
(Engelse titel: Proof of Wigner's semicircle
theorem)**

Verslag ten behoeve van het
Delft Institute of Applied Mathematics
als onderdeel ter verkrijging

van de graad van

**BACHELOR OF SCIENCE
in
TECHNISCHE WISKUNDE**

door

MARK LOOIJJE

**Delft, Nederland
Oktober 2018**



BSc verslag TECHNISCHE WISKUNDE

“Bewijs van Wigners halve cirkel stelling”
(Engelse titel: “Proof of Wigner’s semicircle theorem”)

MARK LOOIJE

Technische Universiteit Delft

Begeleider

Dr.ir. W.G.M. Groenevelt

Overige commissieleden

Prof. dr. F.H.J. Redig

Drs. E.M. van Elderen

...

...

Oktober, 2018

Delft

Samenvatting

Voor grote random matrices bestaat er een verband tussen de eigenwaarden van deze random matrices en halve cirkels. Om meer precies te zijn bestaan er voor random matrices die aan een aantal voorwaarden voldoen een vorm van convergentie tussen de verdeling van eigenwaarden van de matrices en een halve cirkel verdeling.

Een stelling die deze convergentie beschrijft is Wigners halve cirkel stelling. Het doel van dit verslag is het bewijzen van deze stelling.

De vorm van convergentie die we hier tegen komen is wat in de kansrekening bekend staat als zwakke convergentie. In ons geval betekent dit dat voor μ_n de empirische spectraal maat, voor μ_{sc} de halve cirkel verdeling en voor f een willekeurige continue begrensde functie moet gelden dat $\int f d\mu_n \rightarrow \int f d\mu_{sc}$ bijna zeker voor $n \rightarrow \infty$. De maat μ_n hangt hier af van de eigenwaarden van een $n \times n$ Wigner matrix.

Al vroeg wordt aangetoond dat het voldoende is om te laten zien dat alle momenten convergeren. In het vervolg van het verslag wordt er dan ook veel gekeken naar $\int x^k d\mu_{sc}$ en $\mathbb{E}[\int x^k d\mu_n]$. We bekijken $\mathbb{E}[\int x^k d\mu_n]$ in plaats van $\int x^k d\mu_n$, omdat over de verwachting makkelijker uitspraken te doen zijn.

Het bekijken van deze verwachting levert een uitdrukking op in de vorm van een gigantische sommatie over verwachtingen van matrixproducten (spoorformule). Door de elementen waarover gesommeerd wordt als graaf of als wandeling over een graaf te beschouwen kunnen we aantonen dat in veel gevallen de sommand gelijk aan nul of willekeurig klein moet zijn.

Vervolgens blijft er een telprobleem over van het aantal grafen waarvoor de sommand geen nul is. Deze lijken moeilijk te tellen, maar het lukt om een bijjectie te vinden tussen deze grafen en een speciaal type wandelingen op $\mathbb{Z}_{\geq 0}$ (Dyck-wandelingen) welke wel goed te tellen zijn.

Resultaat hiervan is dat $\int x^k d\mu_{sc}$ en $\mathbb{E}[\int x^k d\mu_n]$ nu willekeurig dicht bij elkaar liggen voor grote n .

In de latere hoofdstukken wordt onder andere een bewijs gegeven van een interessant stukje maattheorie, namelijk het Borel Cantelli lemma.

Door dit lemma toe te passen op ons momentenprobleem en een aantal ongelijkheden (waaronder Chebyshev) te gebruiken lukt het om te bewijzen dat $\mathbb{P}(|\int x^k d\mu_n - \int x^k d\mu_{sc}| \geq \epsilon) = 0$ voor $n \rightarrow \infty$, voor $\epsilon > 0$ en voor alle waarden van k .

Inhoudsopgave

Samenvatting	5
Inleiding	9
1 Wigner halve cirkel stelling	11
2 Halve cirkel verdeling	15
3 Empirische spectraal maat	17
4 Beperking tot momenten	23
5 Momenten bij de halve cirkel verdeling	29
6 Spoorformule	35
7 Grafenformule	41
8 Grafenformule voor grote n	47
9 Equivalentie met wandelingen op $\mathbb{Z}_{\geq 0}$	53
10 Borel Cantelli	57
11 Gevolgen van Chebyshev ongelijkheid	61
12 Bewijs van Wigners halve cirkel stelling	67
Conclusie	69
Bijlage A - Uitbreidingen van de stelling	71
Bijlage B - MatLab code EigenvaluesRandomMatrix.m	73
Lijst terugkerende symbolen	75
Bibliografie	77

Inleiding

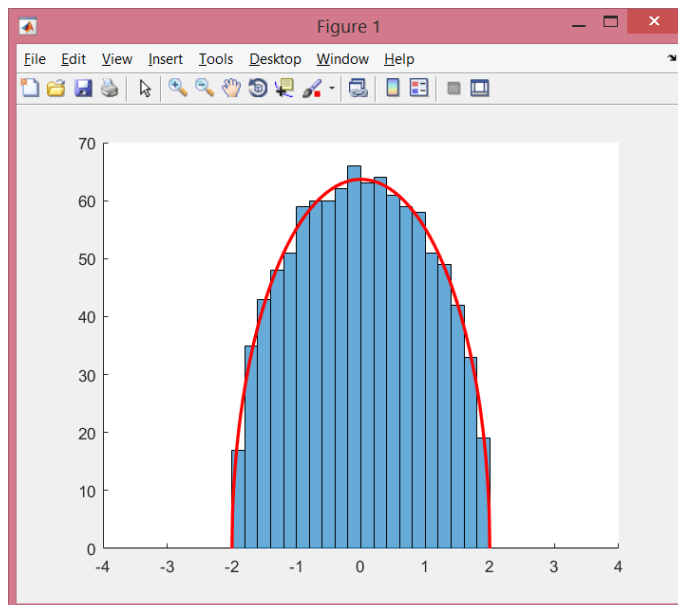
De theorie voor random matrices is rond 1950 ontstaan met als doel het bestuderen van energieniveaus van zware atoomkernen. Dit kan namelijk erg goed met random matrices gemodelleerd worden.

Een interessant onderwerp voor deze random matrices is dat voor bepaalde voorwaarden op de matrix je uitspraken kan doen over de eigenwaarden van deze matrix.

Zo is het het geval dat als de afmetingen van de matrix groot genoeg zijn, de matrix symmetrisch is en aan nog een aantal kleinere voorwaarden voldoet, dat er dan een sterk verband bestaat tussen de eigenwaarden van de random matrix en een halve cirkel.

Ter illustratie kunnen we in bijvoorbeeld MatLab een $n \times n$ matrix W genereren, waarbij voor elk element op en boven de diagonaal een willekeurig getal wordt getrokken uit een normale verdeling met verwachting 0 en variantie 1. De elementen onder de diagonaal verkrijgen we door de matrix te spiegelen door de diagonaal.

We rekenen de eigenwaarden van W/\sqrt{n} uit en geven deze weer door middel van een histogram. Ter vergelijking staat in dezelfde figuur een halve cirkel geplott met dezelfde oppervlakte onder de grafiek.



Figuur 1: De eigenwaarden van een willekeurig gegenereerde 1000x1000 matrix

In figuur 1 hebben we dit gedaan voor matrixgrootte $n = 1000$. De gelijkenis met een halve cirkel is duidelijk zichtbaar.

De stelling die dit fenomeen beschrijft is Wigner's halve cirkel stelling. Voor bepaalde voorwaarden op random matrices geldt dat voor matrixgrootte $n \rightarrow \infty$ er een vorm van convergentie bestaat tussen de eigenwaarden van de matrix en de halve cirkel verdeling.

Wat deze voorwaarden precies zijn en wat voor type convergentie dit is wordt duidelijk gemaakt in het eerste hoofdstuk. In het vervolg van het verslag wordt deze stelling bewezen aan de hand van de momentenmethode. Daar waar extra kennis nodig is om bepaalde stappen van het bewijs te kunnen volgen wordt ook deze informatie uitbundig belicht.

Op deze manier is het de bedoeling dat de stelling en het bewijs daarvan voor een geïnteresseerde tweede- of derdejaars wiskundestudent goed te volgen is.

Hoofdstuk 1

Wigner halve cirkel stelling

Benodigde definities

Zoals gezegd in de inleiding bestaat er een verband tussen de eigenwaarden van bepaalde typen random matrices en een halve cirkel. Om dit verband te kunnen beschrijven hebben we een aantal nieuwe termen nodig.

Definitie 1.1 (Wigner matrices).

Zij W een $n \times n$ matrix waarvoor

- $W_{i,j}$, $1 \leq i \leq j$ zijn onafhankelijke random variabelen op de reële getallen
- Voor elke $j < i$ laat $W_{i,j} = W_{j,i}$
- Voor elke i, j is $W_{i,j}$ begrensd
- Voor elke i, j geldt $\mathbb{E}[W_{i,j}] = 0$
- Voor elke $i \neq j$ geldt $\text{Var}(W_{i,j}) = \mathbb{E}[W_{i,j}^2] = 1$

Matrices W die aan deze voorwaarden voldoen noemen we in dit verslag **Wigner matrices**.

Kijkend naar deze eigenschappen is belangrijk op te merken dat W een reële symmetrische matrix is. Ander belangrijk punt is dat elk kental in de matrix onafhankelijk van alle andere kentallen is, behalve dan het kental gespiegeld in de diagonaal.

Voor bijvoorbeeld een 3×3 matrix ziet dit er dan als volgt uit.

$$\begin{bmatrix} W_{1,1} = RV(\text{beg}, \mathbb{E} = 0) & W_{1,2} = RV(\text{beg}, \mathbb{E} = 0, \text{Var} = 1) & W_{1,3} = RV(\text{beg}, \mathbb{E} = 0, \text{Var} = 1) \\ W_{1,2} = W_{2,1} & W_{2,2} = RV(\text{beg}, \mathbb{E} = 0) & W_{2,3} = RV(\text{beg}, \mathbb{E} = 0, \text{Var} = 1) \\ W_{3,1} = W_{1,3} & W_{3,2} = W_{2,3} & W_{3,3} = RV(\text{beg}, \mathbb{E} = 0) \end{bmatrix}$$

Waarbij bijvoorbeeld $RV(\text{beg}, \mathbb{E} = 0, \text{Var} = 1)$ betekent dat dit een random variabele is die begrensd is, verwachting 0 heeft en variantie 1.

Notatie 1.2.

In dit verslag is de W gereserveerd voor een matrix als in definitie 1.1. Dit betekent dat als je een W tegenkomt dan dit voor een matrix van dit type staat.

Verder kunnen extra indices aan deze matrix worden gegeven, waarbij op verschillende punten de nadruk gelegd kan worden:

- W_n wordt gebruikt als de nadruk erop gelegd wordt dat het een matrix is met afmetingen $n \times n$.
- $W_{i,j}$ is het kental in W op positie i, j .
- $W_{\vec{e}}$ is eigenlijk gelijk aan $W_{i,j}$, in latere hoofdstukken wordt een tweetal getallen (i, j) weleens geïnterpreteerd als zijde \vec{e} van een graaf

Definitie 1.3 (Empirische Spectraal Maat (ESM)).

Voor een symmetrische $n \times n$ matrix W_n , zij $\lambda_1(\frac{1}{\sqrt{n}}W_n) \leq \dots \leq \lambda_n(\frac{1}{\sqrt{n}}W_n)$ de geordende eigenwaarden van $\frac{1}{\sqrt{n}}W_n$.

Dan is de **empirische spectraal maat (ESM)** gedefinieerd als de kansmaat

$$\mu_n := \frac{1}{n} \sum_{i=1}^n \delta_{\lambda_i(\frac{1}{\sqrt{n}}W_n)}$$

Waarbij $\delta_{\lambda_i(\frac{1}{\sqrt{n}}W_n)}$ de Dirac-maat is ten opzichte van het punt $\lambda_i(\frac{1}{\sqrt{n}}W_n)$.

In het verloop van het verslag zullen we vaak enkel λ_i schrijven in plaats van $\lambda_i(\frac{1}{\sqrt{n}}W_n)$. Dit met het oog op een betere overzichtelijkheid.

Aan de ESM wordt dan ook vaak gerefereerd als $\mu_n := \frac{1}{n} \sum_{i=1}^n \delta_{\lambda_i}$.

In hoofdstuk 3 wordt er meer aandacht gegeven aan de ESM en hoe hiermee te rekenen.

In de kansrekening bestaan er meerdere typen convergentie voor rijen random variabelen. Een daarvan is zwakke convergentie, welke vele equivalente definities heeft. We gebruiken de volgende definitie

Definitie 1.4 (Zwakke convergentie).

Zij $\{\mu_n\}$ een rij van kansmaten. We zeggen dat μ_n **zwak convergeert** naar een kansmaat μ als

$$\int f d\mu_n \rightarrow \int f d\mu$$

voor alle continue en begrensde functies $f : D \rightarrow \mathbb{R}$ waarbij D een deelverzameling van \mathbb{R} .

Stelling

Stelling 1.5 (Wigner halve cirkel stelling).

De rij $\{\mu_n\}$ van empirische spectraal maten van een rij Wigner matrices $\{W_n\}$ convergeert zwak bijna zeker naar de halve cirkel verdeling op $[-2, 2]$.

Dit betekent dat voor elke begrensde continue functie f geldt dat

$$\int_{\mathbb{R}} f d\mu_n \rightarrow \int_{\mathbb{R}} f d\mu_{sc} = \int_{-2}^2 f(x) \frac{1}{2\pi} \sqrt{4-x^2} dx \text{ bijna zeker}$$

De toevoeging bijna zeker houdt in dat de verzameling met uitzonderingen eventueel niet-leeg zou kunnen zijn, maar dat deze kans nul heeft. Dit is een kleine nuance die je zou kunnen opvatten als

$$\mathbb{P}\left(\left|\int_{\mathbb{R}} f d\mu_n - \int_{\mathbb{R}} f d\mu_{sc}\right| \geq \epsilon\right) = 0$$

voor $\epsilon > 0$ willekeurig en n groot genoeg.

Er bestaan meer uitgebreidere versies van deze stelling. Dit is echter de versie die we in het hoofddeel van het verslag zo duidelijk mogelijk proberen te bewijzen.

Hoofdstuk 2

Halve cirkel verdeling

Inleiding

Het doel in dit verslag is het aantonen dat de halve cirkel verdeling en de ESM voor grote n sterk op elkaar lijken. In dit hoofdstuk daarom aandacht voor de halve cirkel verdeling. In het volgende hoofdstuk ligt de focus op de ESM.

Voor de kansmaat μ_{sc} (sc staat voor semicircle, oftewel halve cirkel) bestaat er een mooie continue verdelingsfunctie. Dit is mooi, dan kunnen we gewoon Riemann-integreren. In hoofdstuk 5 zullen we zien dat dat ons het leven een stuk makkelijker maakt als we daadwerkelijk gaan moeten integreren over de kansmaat μ_{sc} .

In dit hoofdstuk wordt nog niet echt gerekend, maar worden kort wat algemene eigenschappen van de halve cirkel verdeling bekeken.

Uitwerking

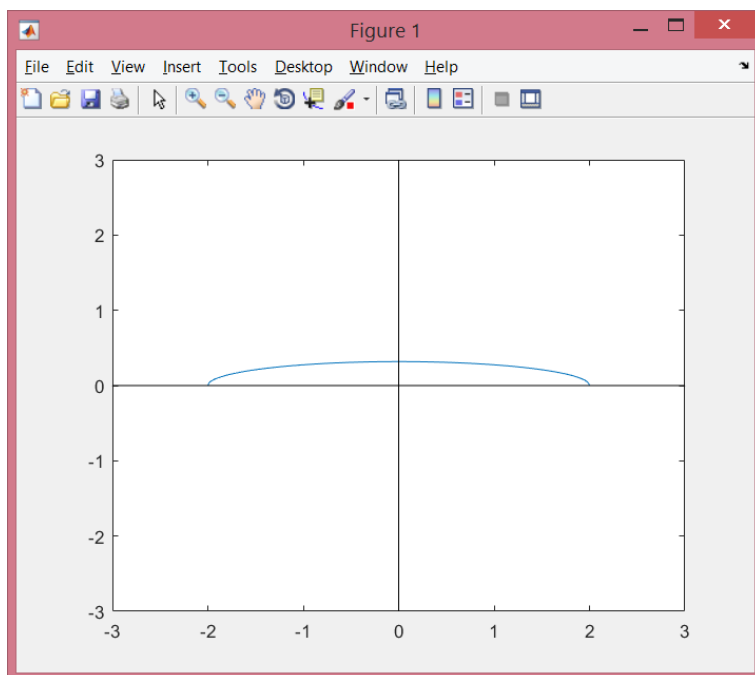
De algemene halve cirkel verdeling op het interval $[-R, R]$ wordt beschreven door de verdelingsfunctie

$$g(x) = \begin{cases} \frac{2}{\pi R^2} \sqrt{R^2 - x^2} & |x| \leq R \\ 0 & |x| > R \end{cases}$$

In ons geval zijn we geïnteresseerd in de verdeling op $[-2, 2]$, hier geldt dan $\frac{1}{2\pi} \sqrt{4 - x^2}$ op $[-2, 2]$ en 0 daarbuiten.

Enigszins opmerkelijk is dat de formule $\frac{1}{2\pi} \sqrt{4 - x^2}$ geen halve cirkel beschrijft. Hoe de grafiek hiervan er wel uitziet is te zien in figuur 2.1.

Voor de keuze $R = \sqrt{2/\pi}$ is de grafiek wel een halve cirkel, voor andere R wordt de grafiek wat opgerekt. $\sqrt{R^2 - x^2}$ beschrijft immers een halve cirkel, de factor $\frac{2}{\pi R^2}$ zorgt er vervolgens voor dat de oppervlakte onder de grafiek 1 blijft. De resulterende vorm is een halve ellips.



Figuur 2.1: De grafiek van $\frac{1}{2\pi}\sqrt{4-x^2}$

Een verdere eigenschap van de halve cirkel verdeling, die we weliswaar in het bewijs niet gebruiken, maar die zeker wel opmerkelijk is, is dat de halve cirkel verdeling een bijzonder geval is van de β -verdeling.

Herinner de verdelingsfunctie van de β -verdeling:

$$g(x) = \frac{1}{B(\alpha, \beta)} x^{\alpha-1} (1-x)^\beta, \quad x \in [0, 1]$$

Neem als parameters $\alpha = \frac{3}{2}$ en $\beta = \frac{3}{2}$. In dat geval geldt $B(\alpha, \beta) = \frac{\pi}{8}$.

$$g(x) = \frac{8}{\pi} x^{\frac{1}{2}} (1-x)^{\frac{1}{2}} = \frac{8}{\pi} \sqrt{x-x^2}, \quad x \in [0, 1]$$

Omschrijven van de wortel geeft

$$g(x) = \frac{8}{\pi} \sqrt{\frac{1}{4} - \left(x - \frac{1}{2}\right)^2}, \quad x \in [0, 1]$$

Dit is precies de verdelingsfunctie van de halve cirkel verdeling voor $R = \frac{1}{2}$, alleen dan met centrum $\frac{1}{2}$ in plaats van centrum 0.

De halve cirkel verdeling is dus om precies te zijn een geschaalde en verschoven β -verdeling.

Hoofdstuk 3

Empirische spectraal maat

Inleiding

In tegenstelling tot bij μ_{sc} is $\int_{\mathbb{R}} f d\mu_n$ niet Riemann-integreerbaar. Het is wel mogelijk om net te doen alsof dit kan en je kan ook beredeneren hoe het dan ongeveer zou moeten werken. Uiteindelijk ben je dan echter dingen aan het doen die niet wiskundig gedefinieerd zijn, het is in ieder geval niet netjes.

Om deze integraal toch netjes te kunnen beschrijven moeten we ons verdiepen in een stukje Reële analyse, of om meer precies te zijn maattheorie en Lebesgue-integreerbaarheid.

We beperken ons hierbij tot die informatie die noodzakelijk is om onze integraal te kunnen beschrijven. Niet elke term hoeft hiervoor uitbundig behandeld te worden als het voor ons specifieke geval niet uitmaakt.

Waar we wel een stukje extra aandacht aan geven is de Dirac-maat en integreren over deze Dirac-maat. Dit omdat de ESM en de Dirac-maat sterk aan elkaar verwant zijn en begrijpen hoe het voor de Dirac-maat werkt extra intuïtie creëert voor de ESM.

Uitwerking

Definitie 3.1 (Maat,kansmaat,maatruimte).

Zij \mathcal{S} een halfring van deelverzamelingen van een verzameling X . Een verzamelingen functie $\mu : \mathcal{S} \rightarrow [0, \infty]$ heet een **maat** op \mathcal{S} als het voldoet aan de eigenschappen

(1.) $\mu(\emptyset) = 0$ en

(2.) Als $\{A_n\}$ een rij disjuncte deelverzamelingen in \mathcal{S} is waarvoor $\bigcup_{n=1}^{\infty} A_n \in \mathcal{S}$, dan

$$\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mu(A_n)$$

Als μ ook nog voldoet aan (3.) noemen we μ een **kansmaat**

(3.) $\mu(X) = 1$

Een triple (X, \mathcal{S}, μ) , met X een niet-lege verzameling, \mathcal{S} een halfring van deelverzamelingen van X en μ een maat op \mathcal{S} , noemen we een **maatruimte**.

Wat hierbij precies een halfring is, is niet per se belangrijk. Het is een structuur waarop je

maattheorie kan uitvoeren. Wat wel belangrijk is, is dat bijvoorbeeld $\mathcal{P}(X)$, de machtsverzameling met alle deelverzamelingen van X een halfring is. Voor $X = \mathbb{R}$ is de verzameling met daarin alle deelverzamelingen van \mathbb{R} een halfring.

We bekijken als voorbeeld de Dirac-maat en laten zien dat deze aan de definitie van een kansmaat voldoet.

Voorbeeld 3.2 (Dirac-maat).

Zij X een niet-lege verzameling en zij $\mathcal{S} = \mathcal{P}(X)$. Voor een element $a \in X$, definieer

$$\mu : \mathcal{S} \rightarrow [0, \infty)$$

$$\mu(A) = \begin{cases} 1 & \text{als } a \in A \\ 0 & \text{als } a \notin A \end{cases}$$

Het is duidelijk dat μ hier een kansmaat is, controleer immers de punten (1.) tot en met (3.) uit definitie 3.1

(1.) $a \notin \emptyset$, dus $\mu(\emptyset) = 0$

(2.) Zij $\{A_n\}$ een rij disjuncte deelverzamelingen van X . Er zijn twee gevallen te onderscheiden. De eerste is dat a in geen enkele van de A_n voorkomt, dan komt a uiteraard ook niet voor in de vereniging van de A_n en dus $\sum_{n=1}^{\infty} \mu(A_n) = 0$ en $\mu(\bigcup_{n=1}^{\infty} A_n) = 0$.

Het tweede geval is dat a in precies een van A_n voorkomt en daarmee dus ook in de vereniging zit. Dan $\sum_{n=1}^{\infty} \mu(A_n) = 1$ en $\mu(\bigcup_{n=1}^{\infty} A_n) = 1$.

In beide gevallen zijn $\mu(\bigcup_{n=1}^{\infty} A_n)$ en $\sum_{n=1}^{\infty} \mu(A_n)$ gelijk.

(3.) $a \in X$, dus $\mu(X) = 1$.

Voorbeeld 3.3 (Toepassing op de ESM).

Zij nu $X = \mathbb{R}$ en $\mathcal{S} = \mathcal{P}(\mathbb{R})$. Schrijf de Dirac-maat ten opzichte van het punt $\lambda_i \in \mathbb{R}$ als δ_{λ_i} . Bekijk dan nu μ_n die we als volgt gedefinieerd hadden:

$$\mu_n = \frac{1}{n} \sum_{i=1}^n \delta_{\lambda_i}$$

Dan is μ_n ook een kansmaat. Dit bewijzen gaat op grotendeels dezelfde manier als bewijzen voor de Dirac-maat even hiervoor.

(1.) $\lambda_1, \dots, \lambda_n \notin \emptyset$, dus $\mu_n(\emptyset) = 0$.

(2.) Zij $\{A_k\}$ een rij disjuncte deelverzamelingen van X . Als index schrijven we hier k omdat de n al gebruikt is in de notatie van de maat. Dan zijn er $l \leq n$ eigenwaarden bevat in de vereniging $\bigcup_{k=1}^{\infty} A_k$, dus $\mu_n(\bigcup_{k=1}^{\infty} A_k) = \frac{l}{n}$.

Elke eigenwaarde in $\bigcup_{k=1}^{\infty} A_k$ zit in exact één deelverzameling A_k , dus ook $\sum_{k=1}^{\infty} \mu_n(A_k) = \frac{l}{n}$.

(3.) Alle n punten $\lambda_1, \dots, \lambda_n \in \mathbb{R}$, dus $\mu_n(\mathbb{R}) = \frac{1}{n} \sum_{i=1}^n \delta_{\lambda_i} = \frac{1}{n} \sum_{i=1}^n 1 = 1$.

Voordat we integreerbaarheid breder kunnen definiëren zijn er nog een paar termen nodig om op te halen. Als eerste het concept bijna overal.

Voor (X, \mathcal{S}, μ) een maatruimte, zeggen dat een relatie **bijna overal** geldt (dit zullen we afkorten met b.o.), betekend dat deze relatie bijna overal geldt ten opzichte van de buitenmaat μ^* gegenereerd uit μ .

Zij nu (X, \mathcal{S}, μ) een maatruimte en f, g reëelwaardige functies gedefinieerd op X , dan:

- (1.) $f = g$ b.o. als $\mu^*(\{x \in X : f(x) \neq g(x)\}) = 0$
- (2.) $f \leq g$ b.o. als $\mu^*(\{x \in X : f(x) > g(x)\}) = 0$
- (3.) $f_n \rightarrow f$ b.o. als $\mu^*(\{x \in X : f_n(x) \not\rightarrow f(x)\}) = 0$
- (4.) $f_n \uparrow f$ b.o. als $f_n \leq f$ b.o. voor alle n en $f_n \rightarrow f$ b.o.

In ons geval waarbij we μ gedefinieerd hebben voor alle deelverzamelingen van \mathbb{R} hoeven we ons niet druk te maken over μ^* , die is voor ons gelijk aan μ .

Wat verder nog nodig is, dat zijn de karakteristieke functies en stapfuncties.

Zij A een deelverzameling van X , dan is de **karakteristieke functie** χ_A gedefinieerd als

$$\chi_A(x) = \begin{cases} 1 & \text{als } x \in A \\ 0 & \text{als } x \notin A \end{cases}$$

Een functie ϕ noemen we een **stapfunctie** als deze te schrijven is in de vorm $\phi = \sum_{i=1}^n a_i \chi_{A_i}$, waarbij $a_i \in \mathbb{R}$ en A_i een meetbare verzameling met eindige maat.

Handige van deze stapfuncties is dat er eenvoudig integralen over genomen kunnen worden.

Definitie 3.4 (Lebesgue integralen voor stapfuncties).

Zij ϕ een stapfunctie met standaardrepresentatie $\phi = \sum_{i=1}^n a_i \chi_{A_i}$, dat wil zeggen dat a_1, \dots, a_n de verschillende niet-nul waarden van ϕ zijn en dat $A_i = \{x \in X : \phi(x) = a_i\}$. De Lebesgue integraal van ϕ is dan gedefinieerd door

$$\int_X \phi \, d\mu = \sum_{i=1}^n a_i \mu^*(A_i)$$

Maar we willen natuurlijk kunnen integreren over meer dan enkel stapfuncties. Daarvoor is het volgende begrip, bovenfuncties, belangrijk.

Een functie $f : X \rightarrow \mathbb{R}$ heet een **bovenfunctie** als er een rij $\{\phi_n\}$ van stapfuncties bestaat zo dat

- (1.) $\phi_n \uparrow f$ b.o. en
- (2.) $\lim \int \phi_n \, d\mu < \infty$

Definitie 3.5 (Lebesgue integralen voor bovenfuncties).

Zij f een bovenfunctie en zij $\{\phi_n\}$ een rij van stapfuncties zo dat $\phi_n \uparrow f$ b.o. geldt. Dan is de

Lebesgue integraal van f gedefinieerd door

$$\int f \, d\mu = \lim_{n \rightarrow \infty} \int \phi_n \, d\mu$$

De waarde van de Lebesgue integraal van een bovenfunctie hangt niet af van welke genererende rij van stapfuncties je hebt gebruikt. Hierdoor is het zo dat als f een bovenfunctie is en g een functie is waarvoor $g = f$ b.o., dat g dan ook een bovenfunctie is en dat $\int g \, d\mu = \int f \, d\mu$. Een uitbreiding van deze redenering gebruiken we in voorbeeld 3.7 en voorbeeld 3.8.

Lang niet alle functies zijn bovenfuncties, dus we moeten nog een stap verder. Deze stap verder noemen we Lebesgue integreerbare functies.

Definitie 3.6 (Lebesgue integreerbaarheid, Lebesgue integraal).

Een functie $f : X \rightarrow \mathbb{R}$ heet **Lebesgue integreerbaar** als er twee bovenfuncties u en v bestaan zo dat $f = u - v$ b.o. geldt. De **Lebesgue integraal** van f is gedefinieerd als

$$\int f \, d\mu = \int u \, d\mu - \int v \, d\mu$$

Twee opmerkingen hierbij: de waarde van de integraal hangt er niet van af welke functies je als u en v neemt, zolang maar $f = u - v$ geldt.

Tweede punt is dat als f Lebesgue integreerbaar is en g is een andere functie met $f = g$ b.o., dan is g ook Lebesgue integreerbaar en $\int g \, d\mu = \int f \, d\mu$. De reden hiervoor is dat als $f = g$ b.o. en $f = u - v$ b.o., dat dan ook $g = u - v$ b.o., dus je kan dezelfde bovenfuncties gebruiken.

Voorbeeld 3.7 (Integreren t.o.v. de Dirac-maat).

Zij X een niet-lege verzameling en zij μ de Dirac maat met respect tot een punt $a \in X$. Dan is iedere functie $f : X \rightarrow \mathbb{R}$ integreerbaar en $\int f \, d\mu = f(a)$.

Immers, $\phi(x) = f(a)\chi_a(x)$ is een stapfunctie die hoort bij $f(x)$ met $f(a)\chi_a(x) = f(x)$ b.o.

Uitrekenen van de integraal gaat met behulp van definitie 3.4 en de opmerking bij definitie 3.6 dan snel:

$$\begin{aligned} \int f \, d\mu &= \int \phi \, d\mu \\ &= \int f(a)\chi_a(x) \, d\mu \\ &= f(a)\mu(a) \\ &= f(a) \end{aligned}$$

Als afsluiting van dit hoofdstuk bekijken we integratie over de ESM. Eerst voor n eindig en daarna voor n oneindig. Dan weten we voor volgende hoofdstukken hoe we hier mee moeten rekenen.

Voorbeeld 3.8 (Integreren over de ESM met eindige n).

Zij nu $X = \mathbb{R}$ en $\mathcal{S} = \mathcal{P}(\mathbb{R})$. Schrijf de Dirac-maat ten opzichte van het punt $\lambda_i \in \mathbb{R}$ als δ_{λ_i} . Bekijk dan nu μ_n die we als volgt gedefinieerd hadden:

$$\mu_n = \frac{1}{n} \sum_{i=1}^n \delta_{\lambda_i}$$

Zij f een begrensde functie $\mathbb{R} \rightarrow \mathbb{R}$.

$\phi(x) = \sum_{i=1}^n f(\lambda_i) \chi_{\lambda_i}(x)$ is een bijbehorende stapfunctie. Immers $\phi(x) = f(x)$ b.o. De bijbehorende integraal is

$$\begin{aligned} \int f \, d\mu_n &= \int \phi \, d\mu_n \\ &= \int \sum_{i=1}^n f(\lambda_i) \chi_{\lambda_i}(x) \, d\mu_n \\ &= \sum_{i=1}^n f(\lambda_i) \mu_n(\lambda_i) \\ &= \frac{1}{n} \sum_{i=1}^n f(\lambda_i) \end{aligned}$$

Voorbeeld 3.9 (Integreren over de ESM voor $n \rightarrow \infty$).

Eerste stuk van voorbeeld 3.8 kunnen we overnemen:

Zij $X = \mathbb{R}$ en $\mathcal{S} = \mathcal{P}(\mathbb{R})$. Schrijf de Dirac-maat ten opzichte van het punt $\lambda_i \in \mathbb{R}$ als δ_{λ_i} . Bekijk dan μ_n die we als volgt gedefinieerd hadden:

$$\mu_n = \frac{1}{n} \sum_{i=1}^n \delta_{\lambda_i}$$

Zij f een begrensde functie $\mathbb{R} \rightarrow \mathbb{R}$.

Introduceer de functies f^+ en f^- als volgt:

$$f^+ = \max\{f, 0\}, \quad f^- = -\min\{f, 0\}$$

Dan geldt $f = f^+ - f^-$. Voor de functies f^+ en f^- kunnen we aantonen dat dit bovenfuncties zijn.

Definieer de rijen stapfuncties $\{\phi_n\}$ en $\{\psi_n\}$ als

$$\phi_n(x) = \sum_{i=1}^n f^+(\lambda_i) \chi_{\lambda_i}(x), \quad \psi_n(x) = \sum_{i=1}^n f^-(\lambda_i) \chi_{\lambda_i}(x)$$

Dan $\phi_n \leq \phi_{n+1}$ b.o. voor alle n en $\phi_n \rightarrow f^+$ b.o. .

Ook $\psi_n \leq \psi_{n+1}$ b.o. voor alle n en $\psi_n \rightarrow f^-$ b.o. .

Met andere woorden $\phi_n \uparrow f^+$ b.o. en $\psi_n \uparrow f^-$ b.o. .

Verder geldt:

$$\lim_{n \rightarrow \infty} \int \phi_n \, d\mu = \lim_{n \rightarrow \infty} \sum_{i=1}^n f^+(\lambda_i) \mu^*(\lambda_i) = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n f^+(\lambda_i) \text{ en}$$

$$\lim_{n \rightarrow \infty} \int \psi_n \, d\mu = \lim_{n \rightarrow \infty} \sum_{i=1}^n f^-(\lambda_i) \mu^*(\lambda_i) = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n f^-(\lambda_i)$$

Waarbij deze limieten beide bestaan en eindig zijn, want f en daarmee f^+ en f^- zijn begrensd.

Nu is aangetoond dat f^+ en f^- bovenfuncties zijn met

$$\int f^+ \, d\mu = \lim_{n \rightarrow \infty} \int \phi_n \, d\mu = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n f^+(\lambda_i) \text{ en}$$

$$\int f^- \, d\mu = \lim_{n \rightarrow \infty} \int \psi_n \, d\mu = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n f^-(\lambda_i)$$

Is het eenvoudig te concluderen dat

$$\int f \, d\mu = \int f^+ \, d\mu - \int f^- \, d\mu$$

$$= \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n f^+(\lambda_i) - f^-(\lambda_i)$$

$$= \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n f(\lambda_i)$$

Hoofdstuk 4

Beperking tot momenten

Inleiding

De vorige hoofdstukken dienden voornamelijk om termen te introduceren, de stelling die we willen bewijzen te benoemen en om wat intuïtie te creëren bij het onderwerp. Zo hebben we wat plaatjes bekeken in MatLab en we hebben bekeken wat het betekend om te integreren over de ESM.

In dit hoofdstuk gaan we daadwerkelijk beginnen aan het bewijzen van stelling 1.5. Herinner:

Stelling 1.5 (Wigner halve cirkel stelling).

De rij $\{\mu_n\}$ van empirische spectraal maten van een rij Wigner matrices $\{W_n\}$ convergeert zwak bijna zeker naar de halve cirkel verdeling op $[-2, 2]$.

Dit betekent dat voor elke begrensde continue functie f geldt dat

$$\int_{\mathbb{R}} f d\mu_n \rightarrow \int_{\mathbb{R}} f d\mu_{sc} = \int_{-2}^2 f(x) \frac{1}{2\pi} \sqrt{4-x^2} dx \text{ bijna zeker}$$

Bewijzen dat integralen over willekeurige functies f convergeren is natuurlijk niet eenvoudig. Doel van dit hoofdstuk is om het probleem een stuk te vereenvoudigen. We gaan namelijk laten zien dat het voldoende is om te laten zien dat de integralen over de momenten convergeren.

Het is vervolgens in volgende hoofdstukken een stuk behapbaarder om onderzoek te doen naar $\int x^k d\mu_{sc}$ en $\int x^k d\mu_n$ voor willekeurige $k \in \mathbb{Z}_{\geq 0}$, dan voor dezelfde integralen over willekeurige functies f .

Uitwerking

Om dit te bewijzen zullen we onder andere meerdere keren Weierstraß en de driehoeksongelijkheid toe moeten passen. Uiteraard komen daar slimme vermenigvuldigingen en het gebruiken van eigenschappen van integralen en de functie f bij, maar de hoofdrolspeler is Weierstraß. Om deze reden geven we hier de benaderingsstelling van Weierstraß en een gevolg hiervan. Kunnen we hier makkelijk naar refereren.

Stelling 4.1 (Benaderingsstelling van Weierstraß).

Zij f een continue reële functie op $[a, b]$. Dan voor iedere $\epsilon > 0$, bestaat er een polynoom p op $[a, b]$ zo dat, $|f(x) - p(x)| < \epsilon$ voor alle $x \in [a, b]$.

Gevolg 4.2 (Weierstraß voor integralen).

Zij f een continue reële functie op $[a, b]$. Zij μ een kansmaat. Dan voor iedere $\epsilon > 0$, bestaat er een polynoom p op $[a, b]$ zo dat $|\int_{x=a}^b f(x) - p(x) d\mu| < \epsilon$ voor alle $x \in [a, b]$.

Bewijs.

Zij $\epsilon > 0$. Dan bestaat er volgens de benaderingsstelling van Weierstraß een polynoom p op $[a, b]$ zo dat $|f(x) - p(x)| < \epsilon$ voor alle $x \in [a, b]$. Dan

$$\left| \int_{x=a}^b f(x) - p(x) d\mu \right| \leq \int_{x=a}^b |f(x) - p(x)| d\mu \leq \int_{x=a}^b \epsilon d\mu < \epsilon \int_{x=a}^b d\mu \leq \epsilon \quad \square$$

Het is trouwens ook zo dat het polynoom $p(x)$ die werkt voor stelling 4.1 ook voldoet voor gevolg 4.2.

Stelling 4.3 (Beperking tot momenten).

Zij f een continue en begrensde functie op \mathbb{R} , zij $[-R, R]$ een interval dat $[-2, 2]$ bevat en zij $\epsilon > 0$.

Zij $p(x)$ een polynoom als in stelling 4.1, dus $|f(x) - p(x)| < \epsilon$ voor alle $x \in [-R, R]$.

Als nu voor alle $k \in \mathbb{Z}_{\geq 0}$ geldt

$$\left| \int x^k d\mu_n - \int x^k d\mu_{sc} \right| \rightarrow 0$$

Dan

$$\left| \int f(x) d\mu_n - \int f(x) d\mu_{sc} \right| \rightarrow 0$$

Bewijs.

Zij f een continue, begrensde functie op een begrensd interval $[-R, R]$, waarbij $R > 2$. Zij $\epsilon > 0$.

Dan bestaat er een polynoom p waarvoor $|f(x) - p(x)| < \epsilon$ voor alle $x \in [-R, R]$

Voor dit polynoom p bestaan er c_j zo dat $p(x) = \sum_{j=1}^k c_j x^j$, waarbij k de graad is van p .

Uit de driehoeksongelijkheid volgt:

$$\begin{aligned} & \left| \int_{x \in \mathbb{R}} f d\mu_n - \int_{x \in \mathbb{R}} f d\mu_{sc} \right| \\ &= \left| \int_{x \in \mathbb{R}} f d\mu_n - \int_{x \in \mathbb{R}} p d\mu_n + \int_{x \in \mathbb{R}} p d\mu_n - \int_{x \in \mathbb{R}} p d\mu_{sc} + \int_{x \in \mathbb{R}} p d\mu_{sc} - \int_{x \in \mathbb{R}} f d\mu_{sc} \right| \\ &\leq \left| \int_{x \in \mathbb{R}} (f - p) d\mu_n \right| + \left| \int_{x \in \mathbb{R}} p d\mu_n - \int_{x \in \mathbb{R}} p d\mu_{sc} \right| + \left| \int_{x \in \mathbb{R}} (f - p) d\mu_{sc} \right| \\ &\leq \int_{x \in \mathbb{R}} |f - p| d\mu_n + \left| \int_{x \in \mathbb{R}} p d\mu_n - \int_{x \in \mathbb{R}} p d\mu_{sc} \right| + \int_{x \in \mathbb{R}} |f - p| d\mu_{sc} \end{aligned}$$

$$= \star_1 + \star_2 + \star_3$$

Waarbij

$$\star_1 = \int_{x \in \mathbb{R}} |f - p| d\mu_n$$

$$\star_2 = \left| \int_{x \in \mathbb{R}} p d\mu_n - \int_{x \in \mathbb{R}} p d\mu_{sc} \right|$$

$$\star_3 = \int_{x \in \mathbb{R}} |f - p| d\mu_{sc}$$

Nu is het zaak om te laten zien dat zowel \star_1 , \star_2 als \star_3 naar nul convergeren voor grote n . We zullen dit laten zien in volgorde van makkelijk, naar moeilijk. Dit komt neer op eerst \star_3 , dan \star_2 en als laatste \star_1 .

Afschatten van \star_3 :

$$\begin{aligned} & \int_{x \in \mathbb{R}} |f - p| d\mu_{sc} \\ &= \int_{x=-2}^{x=2} |f - p| d\mu_{sc} \end{aligned}$$

We integreren over een eindig interval voor $x \in [-2, 2]$, dus hier mogen we Weierstraß toepassen.

$$\begin{aligned} &< \int_{x=-2}^{x=2} \epsilon d\mu_{sc} \\ &= \epsilon \int_{x=-2}^{x=2} d\mu_{sc} \\ &= \epsilon \end{aligned}$$

Afschatten van \star_2 :

$$\begin{aligned} \star_2 &= \left| \int_{x \in \mathbb{R}} p d\mu_n - \int_{x \in \mathbb{R}} p d\mu_{sc} \right| \\ &= \left| \int_{x \in \mathbb{R}} \sum_{j=1}^k c_j x^j d\mu_n - \int_{x \in \mathbb{R}} \sum_{j=1}^k c_j x^j d\mu_{sc} \right| \end{aligned}$$

Het gaat om een eindige sommatie, dus we mogen sommatie en integratie verwisselen

$$\begin{aligned} &= \left| \sum_{j=1}^k \int_{x \in \mathbb{R}} c_j x^j d\mu_n - \sum_{j=1}^k \int_{x \in \mathbb{R}} c_j x^j d\mu_{sc} \right| \\ &= \left| \sum_{j=1}^k \left(\int_{x \in \mathbb{R}} c_j x^j d\mu_n - \int_{x \in \mathbb{R}} c_j x^j d\mu_{sc} \right) \right| \end{aligned}$$

$$\begin{aligned}
&= \left| \sum_{j=1}^k c_j \left(\int_{x \in \mathbb{R}} x^j d\mu_n - \int_{x \in \mathbb{R}} x^j d\mu_{sc} \right) \right| \\
&\leq \sum_{j=1}^k |c_j| \left| \int_{x \in \mathbb{R}} x^j d\mu_n - \int_{x \in \mathbb{R}} x^j d\mu_{sc} \right|
\end{aligned}$$

Uit de aanname volgt dat alle elementen in de sommatie naar nul convergeren. Aangezien de sommatie een eindige sommatie is geldt dat $\star_2 \rightarrow 0$.

Afschatten van \star_1 :

Deze is met afstand de lastigste. Ookal lijkt deze op het eerste gezicht erg sterk op \star_3 , hier mogen we niet zo eenvoudig overheen gaan. Dit omdat we niet weten of de λ_i 's begrensd zijn.

Spits de integraal over x over \mathbb{R} in een integraal over het begrensde interval $[-R, R]$ en een integraal over het complement van $[-R, R]$.

$$\begin{aligned}
&\int_{x \in \mathbb{R}} |f - p| d\mu_n \\
&= \int_{|x| < R} |f - p| d\mu_n + \int_{|x| > R} |f - p| d\mu_n
\end{aligned}$$

De linkerintegraal convergeert natuurlijk wel naar nul, op dit stukje kan je namelijk wel Weierstrass toepassen. Blijft over om naar te kijken de rechterintegraal.

Merk op dat f begrensd is, er bestaat dus een $A \in \mathbb{R}$ zodat $f(x) < A$ voor alle x .

Verder geldt dat $|p(x)| = \left| \sum_{j=1}^k c_j x^j \right| \leq \sum_{j=1}^k |c_j| |x|^j$.

Als we enkel die waarden van $|x| > R > 1$ bekijken, wat we zo dadelijk ook gaan doen, dan geldt verder dat $|p(x)| \leq \sum_{j=1}^k |c_j| |x|^k \leq \left(\sum_{j=1}^k |c_j| \right) |x|^k = B|x|^k$ voor $B = \sum_{j=1}^k |c_j|$.

Dit is belangrijk voor het afschatten van de rechterintegraal.

$$\begin{aligned}
&\int_{|x| > R} |f - p| d\mu_n \\
&\leq \int_{|x| > R} |f| + |p| d\mu_n \\
&\leq \int_{|x| > R} A + B|x|^k d\mu_n \\
&\leq \int_{|x| > R} (A + B)|x|^k d\mu_n
\end{aligned}$$

De laatste stap kan gemaakt worden omdat $|x|^k > 1$, dus $A < A|x|^k$

Vermenigvuldig de integrand met $\frac{|x|^k}{R^k} \frac{|x|^{2l}}{R^{2l}}$, waarbij k de graad van het polynoom p en $l \in \mathbb{Z}_{\geq 1}$ willekeurig. Voor $x > R$ zijn beide breuken groter dan 1. Dus

$$\int_{|x| > R} |f - p| d\mu_n$$

$$\begin{aligned}
&\leq \int_{|x|>R} (A+B)|x|^k d\mu_n \\
&\leq \int_{|x|>R} (A+B) \frac{|x|^{2k+2l}}{R^{k+2l}} d\mu_n \\
&= \frac{A+B}{R^{k+2l}} \int_{|x|>R} x^{2k+2l} d\mu_n \\
&\leq \frac{A+B}{R^{k+2l}} \int_{x \in \mathbb{R}} x^{2k+2l} d\mu_n
\end{aligned}$$

Korte uitleg bij de laatste twee stappen. De absolute waarde strepen rond x kunnen weggehaald worden aangezien we een even macht bekijken, die is altijd positief. Wegens dat de integrand altijd positief is, is de integraal over $|x| > R$ ook altijd kleiner dan de integraal over heel \mathbb{R} .

De integraal geeft nu eigenlijk het $2k + 2l$ -de moment weer. Voor de momenten weten we uit de aanname dat voor grote n de momenten bij kansmaat μ_n convergeren naar de momenten bij kansmaat μ_{sc} . Dus

$$\rightarrow \frac{A+B}{R^{k+2l}} \int_{x \in \mathbb{R}} x^{2k+2l} d\mu_{sc} \quad (n \rightarrow \infty)$$

Gebruik makend van de eigenschappen van μ_{sc} volgt nu dat

$$\begin{aligned}
&= \frac{A+B}{R^{k+2l}} \int_{x=-2}^{x=2} x^{2k+2l} d\mu_{sc} \\
&\leq \frac{A+B}{R^{k+2l}} \int_{x=-2}^{x=2} 2^{2k+2l} d\mu_{sc} \\
&= \frac{A+B}{R^{k+2l}} 2^{2k+2l} \int_{x=-2}^{x=2} d\mu_{sc} \\
&= \frac{A+B}{R^{k+2l}} 2^{2k+2l} \\
&= 2(A+B) \frac{2^{k+2l}}{R^{k+2l}} \\
&= 2(A+B) \left(\frac{2}{R}\right)^{k+2l}
\end{aligned}$$

We hebben aangenomen dat $R > 2$, dus de breuk $\frac{2}{R} < 1$.

Laat nu $l \rightarrow \infty$ (dit is toegestaan, want l was willekeurig), dan $\left(\frac{2}{R}\right)^{k+2l} \rightarrow 0$ en $2(A+B)\left(\frac{2}{R}\right)^{k+2l} \rightarrow 0$. Dus

$$\star_1 = \int_{|x|<R} |f-p| d\mu_n + \int_{|x|>R} |f-p| d\mu_n \rightarrow 0 + 0 = 0 \quad (n \rightarrow \infty)$$

En

$$\left| \int_{x \in \mathbb{R}} f d\mu_n - \int_{x \in \mathbb{R}} f d\mu_{sc} \right| \leq \star_1 + \star_2 + \star_3 \rightarrow 0 \quad (n \rightarrow \infty)$$

□

Hoofdstuk 5

Momenten bij de halve cirkel verdeling

Inleiding

In hoofdstuk 4 hebben we bepaald dat het voldoende is te laten zien dat de momenten convergeren. In dit hoofdstuk rekenen we uit wat deze momenten dan zijn, of tenminste dit doen we voor de semicircle verdeling μ_{sc} .

Bij deze berekening moeten we onderscheid maken tussen de even en de oneven momenten. Dit vanwege eigenschappen voor het integreren over even danwel oneven functies. We zullen dan vervolgens vrij snel zien dat voor de oneven momenten we nul als antwoord vinden. Voor de even momenten wordt het een flinke rekenpartij en eindigen we met aardig wat geschuif met faculteiten.

Uitwerking

We beginnen met de makkelijkere waarden van k , oftewel de oneven kant.

Stelling 5.1.

Zij k oneven. Dan is het k -de moment gelijk aan

$$\int x^k d\mu_{sc} = 0$$

Bewijs.

$$\begin{aligned} \int x^k d\mu_{sc} &= \int_{x=-2}^{x=2} x^k \frac{1}{2\pi} \sqrt{4-x^2} dx \\ &= \frac{1}{2\pi} \int_{-2}^2 x^k \sqrt{4-x^2} dx \end{aligned}$$

Bij oneven machten k is x^k een oneven functie. $\sqrt{4-x^2}$ is juist even. Samen is $x^k \sqrt{4-x^2}$ dan een oneven functie. Integralen van een oneven functie over een begrend interval $[-R, R]$, in dit geval $R = 2$, zijn gelijk aan nul. \square

Voor het bepalen van de even momenten moet er meer gerekend worden. We gebruiken de substitutie $x = 2 \sin(\theta)$ en er wordt ook partieel geïntegreerd. Daarom in plaats van direct de uitkomst te geven eerst wat voorbereidingen.

Lemma 5.2.

Voor $k \in \mathbb{Z}_{\geq 2}$ geldt een reductieformule voor $\int_a^b \sin^k(\theta) d\theta$, namelijk

$$\int_a^b \sin^k(\theta) d\theta = \frac{[-\sin^{k-1}(x) \cos(x)]_a^b}{k} + \frac{k-1}{k} \int_a^b \sin^{k-2}(\theta) d\theta$$

Bewijs.

$$\int_a^b \sin^k(\theta) d\theta = \int_a^b \sin^{k-1}(\theta) \sin(\theta) d\theta$$

Nu partieel integreren met

$$f(\theta) = \sin^{k-1}(\theta), \quad f'(\theta) = (k-1) \sin^{k-2}(\theta) \cos(\theta)$$

$$g(\theta) = -\cos(\theta), \quad g'(\theta) = \sin(\theta)$$

Geeft de volgende uitdrukking voor $\int_a^b \sin^k(\theta) d\theta$:

$$= \left[-\sin^{k-1}(\theta) \cos(\theta)\right]_a^b - \int_a^b (k-1) \sin^{k-2}(\theta) \cos(\theta) \cdot -\cos(\theta) d\theta$$

$$= \left[-\sin^{k-1}(\theta) \cos(\theta)\right]_a^b + (k-1) \int_a^b \sin^{k-2}(\theta) \cos^2(\theta) d\theta$$

Gebruiken van de trigonometrische eigenschap $\sin^2 + \cos^2 = 1$ leidt tot

$$= \left[-\sin^{k-1}(\theta) \cos(\theta)\right]_a^b + (k-1) \int_a^b \sin^{k-2}(\theta) (1 - \sin^2(\theta)) d\theta$$

$$= \left[-\sin^{k-1}(\theta) \cos(\theta)\right]_a^b + (k-1) \int_a^b \sin^{k-2}(\theta) - \sin^k(\theta) d\theta$$

$$= \left[-\sin^{k-1}(\theta) \cos(\theta)\right]_a^b + (k-1) \int_a^b \sin^{k-2}(\theta) d\theta - (k-1) \int_a^b \sin^k(\theta) d\theta$$

Dus een samenvatting van wat we tot nu toe hebben bepaald

$$\int_a^b \sin^k(\theta) d\theta = \left[-\sin^{k-1}(\theta) \cos(\theta)\right]_a^b + (k-1) \int_a^b \sin^{k-2}(\theta) d\theta - (k-1) \int_a^b \sin^k(\theta) d\theta$$

Tel aan allebei de kanten $(k-1) \int_a^b \sin^k(\theta) d\theta$ erbij op:

$$k \int_a^b \sin^k(\theta) d\theta = \left[-\sin^{k-1}(\theta) \cos(\theta)\right]_a^b + (k-1) \int_a^b \sin^{k-2}(\theta) d\theta$$

Nu beide kanten door k delen geeft het resultaat

$$\int_a^b \sin^k(\theta) d\theta = \frac{[-\sin^{k-1}(\theta) \cos(\theta)]_a^b}{k} + \frac{k-1}{k} \int_a^b \sin^{k-2}(\theta) d\theta$$

□

Lemma 5.3.

Voor de grenzen $a = 0$ en $b = \frac{1}{2}\pi$ geldt voor de reductieformule 5.2 dat

$$\int_0^{\frac{1}{2}\pi} \sin^k(\theta) d\theta = \frac{k-1}{k} \int_0^{\frac{1}{2}\pi} \sin^{k-2}(\theta) d\theta$$

Bewijs.

Er geldt dat $[-\sin^{k-1}(\theta) \cos(\theta)]_0^{\frac{1}{2}\pi} = 0$, immers

$$\begin{aligned} & [-\sin^{k-1}(\theta) \cos(\theta)]_0^{\frac{1}{2}\pi} \\ &= -\sin^{k-1}(0) \cos(0) + \sin^{k-1}\left(\frac{1}{2}\pi\right) \cos\left(\frac{1}{2}\pi\right) \\ &= 0^{k-1} \cdot 1 + 1^{k-1} \cdot 0 = 0 \end{aligned}$$

Lemma 5.2 met $a = 0$ en $b = \frac{1}{2}\pi$:

$$\begin{aligned} \int_0^{\frac{1}{2}\pi} \sin^k(\theta) d\theta &= \frac{[-\sin^{k-1}(\theta) \cos(\theta)]_0^{\frac{1}{2}\pi}}{k} + \frac{k-1}{k} \int_0^{\frac{1}{2}\pi} \sin^{k-2}(\theta) d\theta \\ &= \frac{k-1}{k} \int_0^{\frac{1}{2}\pi} \sin^{k-2}(\theta) d\theta \end{aligned}$$

□

We willen nu een nieuwe notatie invoeren voor $\int_0^{\frac{1}{2}\pi} \sin^k(\theta) d\theta$. Het resultaat van deze integraal is enkel afhankelijk van de waarde van k , we korten deze integraal dan ook af met a_k .

Notatie 5.4.

$$a_k = \int_0^{\frac{1}{2}\pi} \sin^k(\theta) d\theta$$

Met deze nieuwe notatie kunnen we lemma 5.3 een stuk korter en eenvoudiger opschrijven:

$$a_k = \frac{k-1}{k} a_{k-2}.$$

Verder willen we voor de even getallen k a_k uitdrukken in a_0 . Dit omdat a_0 eenvoudig uit te rekenen is en het bovendien handig is een meer algemene reductieformule te hebben.

Lemma 5.5.

Zij k even, zij $a_k = \int_0^{\frac{1}{2}\pi} \sin^k(\theta) d\theta$. Dan

$$a_k = \frac{k!}{2^k (\frac{1}{2}k)! (\frac{1}{2}k)!} a_0$$

Bewijs.

Beginnen met het aantonen van een eerdere reductieformule

$$a_k = \frac{k-1}{k} \frac{k-3}{k-2} \cdots \frac{1}{2} a_0 \quad (*)$$

Dit is eenvoudig aan te tonen, hij klopt voor $k = 2$, immers $\frac{2-1}{2}a_{2-2} = \frac{1}{2}a_0$.

Nu aangenomen dat het klopt voor zekere waarde k , dan geldt het ook voor $k + 2$, immers $a_{k+2} = \frac{k+1}{k+2}a_k = \frac{k+1}{k+2} \frac{k-1}{k} \frac{k-3}{k-2} \dots \frac{1}{2}a_0$. Uit inductie volgt nu het gewenste resultaat.

Nu willen we formule (*) omschrijven naar een nieuwe reductieformule waarin we faculteiten tegenkomen. Vermenigvuldig daarop de breuk $\frac{k-1}{k}$ met $\frac{k}{k}$, de breuk $\frac{k-3}{k-2}$ met $\frac{k-2}{k-2}$, enzovoorts tot en met $\frac{2}{2}$ vermenigvuldigt met $\frac{1}{2}$.

$$\begin{aligned} a_k &= \frac{k-1}{k} \frac{k-3}{k-2} \dots \frac{1}{2} a_0 \\ &= \frac{k(k-1)}{k^2} \frac{(k-2)(k-3)}{(k-2)^2} \dots \frac{2 \cdot 1}{2^2} a_0 \end{aligned}$$

In de teller is $k!$ al te herkennen, de noemer moet eerst verder worden omgeschreven.

$$= \frac{k!}{(2(\frac{1}{2}k))^2 (2(\frac{1}{2}k-1))^2 \dots (2(1))^2} a_0$$

Verander in de noemer de volgorde van de vermenigvuldiging

$$= \frac{k!}{2^k (\frac{1}{2}k)^2 (\frac{1}{2}k-1)^2 \dots (1)^2} a_0$$

Nu is duidelijk te zien dat

$$= \frac{k!}{2^k (\frac{1}{2}k)! (\frac{1}{2}k)!} a_0$$

□

Lemma 5.6.

$$a_0 = \frac{1}{2}\pi$$

Bewijs.

$$a_0 = \int_0^{\frac{1}{2}\pi} \sin^0(\theta) d\theta = \int_0^{\frac{1}{2}\pi} 1 d\theta = \frac{1}{2}\pi$$

□

Nu hebben we genoeg kennis om de even momenten uit te gaan rekenen.

Stelling 5.7.

Zij k even. Dan is het k -de moment gelijk aan

$$\int x^k d\mu_{sc} = \frac{k!}{(\frac{1}{2}k+1)! (\frac{1}{2}k)!}$$

Bewijs.

$$\begin{aligned} \int x^k d\mu_{sc} &= \int_{-2}^2 x^k \frac{1}{2\pi} \sqrt{4-x^2} dx \\ &= \frac{1}{2\pi} \int_{-2}^2 x^k \sqrt{4-x^2} dx \end{aligned}$$

Integreren over een even functie, dus

$$= \frac{1}{\pi} \int_0^2 x^k \sqrt{4-x^2} dx$$

Nu integreren met de substitutiemethode

$$x = 2 \sin(\theta), \quad d\theta = 2 \cos(\theta) dx$$

De keuze voor deze substitutie zorgt ervoor dat vooral $\sqrt{4-x^2}$ erg mooi uitkomt.

$$x^k = (2 \sin(\theta))^k = 2^k \sin^k(\theta)$$

$$\sqrt{4-x^2} = \sqrt{4-4\sin^2(\theta)} = \sqrt{4} \sqrt{1-\sin^2(\theta)} = 2\sqrt{\cos^2(\theta)} = 2 \cos(\theta)$$

De reden dat bij het wegwerken van de wortel er geen absolute waarde hoeft te worden genomen is omdat $\cos(\theta)$ voor de interessante waarden van θ positief is. De grenzen van de integraal veranderen dan van $x=0$ en $x=2$ naar $\theta=0$ en $\theta=\frac{1}{2}\pi$. Dit geeft

$$\begin{aligned} &= \frac{1}{\pi} \int_0^{\frac{1}{2}\pi} 2^k \sin^k(\theta) 2 \cos(\theta) 2 \cos(\theta) d\theta \\ &= \frac{2^{k+2}}{\pi} \int_0^{\frac{1}{2}\pi} \sin^k(\theta) \cos^2(\theta) d\theta \\ &= \frac{2^{k+2}}{\pi} \int_0^{\frac{1}{2}\pi} \sin^k(\theta) (1-\sin^2(\theta)) d\theta \\ &= \frac{2^{k+2}}{\pi} \int_0^{\frac{1}{2}\pi} \sin^k(\theta) - \sin^{k+2}(\theta) d\theta \end{aligned}$$

Gebruik nu de notatie als in 5.4

$$= \frac{2^{k+2}}{\pi} (a_k - a_{k+2})$$

En de recursieve formule 5.5

$$= \frac{2^{k+2}}{\pi} \left(\frac{k!}{2^k (\frac{1}{2}k)! (\frac{1}{2}k)!} a_0 - \frac{(k+2)!}{2^{k+2} (\frac{1}{2}k+1)! (\frac{1}{2}k+1)!} a_0 \right)$$

Haal de factor $a_0 = \frac{1}{2}\pi$ buiten de haakjes en schrijf de tweede breuk zo op dat deze meer op de eerste lijkt.

$$= 2^{k+1} \left(\frac{k!}{2^k (\frac{1}{2}k)! (\frac{1}{2}k)!} - \frac{(k+2)(k+1)}{2^2 (\frac{1}{2}k+1)^2} \frac{k!}{2^k (\frac{1}{2}k)! (\frac{1}{2}k)!} \right)$$

Haal nu de factor 2^k in de noemer buiten de haakjes

$$\begin{aligned}
 &= 2 \left(\frac{k!}{(\frac{1}{2}k)!(\frac{1}{2}k)!} - \frac{(k+2)(k+1)}{(k+2)^2} \frac{k!}{(\frac{1}{2}k)!(\frac{1}{2}k)!} \right) \\
 &= 2 \left(\frac{k!}{(\frac{1}{2}k)!(\frac{1}{2}k)!} - \frac{k+1}{k+2} \frac{k!}{(\frac{1}{2}k)!(\frac{1}{2}k)!} \right)
 \end{aligned}$$

Vermenigvuldig de breuk $\frac{k!}{(\frac{1}{2}k)!(\frac{1}{2}k)!}$ met $\frac{\frac{1}{2}k+1}{\frac{1}{2}k+1}$ en de breuk $\frac{k+1}{k+2}$ met $\frac{1/2}{1/2}$

$$\begin{aligned}
 &= 2 \left(\frac{(\frac{1}{2}k+1)k!}{(\frac{1}{2}k+1)!(\frac{1}{2}k)!} - \frac{\frac{1}{2}k+\frac{1}{2}}{\frac{1}{2}k+1} \frac{k!}{(\frac{1}{2}k)!(\frac{1}{2}k)!} \right) \\
 &= 2 \left(\frac{(\frac{1}{2}k+1)k!}{(\frac{1}{2}k+1)!(\frac{1}{2}k)!} - \frac{(\frac{1}{2}k+\frac{1}{2})k!}{(\frac{1}{2}k+1)!(\frac{1}{2}k)!} \right)
 \end{aligned}$$

Nu zijn de noemers gelijk, kan je de breuken van elkaar aftrekken

$$\begin{aligned}
 &= 2 \left(\frac{\frac{1}{2}k!}{(\frac{1}{2}k+1)!(\frac{1}{2}k)!} \right) \\
 &= \frac{k!}{(\frac{1}{2}k+1)!(\frac{1}{2}k)!}
 \end{aligned}$$

□

Voor het overzicht sluiten we dit hoofdstuk af met een korte samenvatting, namelijk

$$\int x^k d\mu_{sc} = \begin{cases} 0 & k \text{ oneven} \\ \frac{k!}{(\frac{1}{2}k+1)!(\frac{1}{2}k)!} & k \text{ even} \end{cases}$$

Hoofdstuk 6

Spoorformule

Inleiding

In het vorige hoofdstuk hebben we $\int x^k d\mu_{sc}$ uitgerekend, nu is het de beurt aan $\int x^k d\mu_n$. In feite weten we daar al wat over, immers in hoofdstuk 3 hebben we laten zien hoe we integreren over de empirische spectraal maat:

$$\int x^k d\mu_n = \frac{1}{n} \sum_{i=1}^n \lambda_i^k$$

In deze uitdrukking zien we nog de λ_i -s, die dus random zijn. Ze hangen af van alle stochasten $W_{i,j}$.

We willen iets kunnen zeggen over $\int x^k d\mu_{sc}$, zonder dat we daar nog random elementen in tegen komen. Dat is ook nodig om te kunnen vergelijken bij de momenten die we vonden bij de halve cirkel verdeling. In plaats van naar $\int x^k d\mu_{sc}$, kijken we daarom $\mathbb{E}[\int x^k d\mu_{sc}]$.

$\mathbb{E}[\int x^k d\mu_{sc}]$ ziet er misschien dubbelop uit, dit is het niet en dat komt omdat de x in de integraal niet de random variabele is. In het begin van de uitwerking wordt er aandacht aan gegeven wat $\mathbb{E}[\int x^k d\mu_{sc}]$ nu precies betekend.

Het echte doel van dit hoofdstuk is een begin te maken aan de analyse van $\mathbb{E}[\int x^k d\mu_{sc}]$. Het uitwerken van dit probleem is te groot om in een enkel hoofdstuk te passen, dus eerst wordt er toegewerkt richting een uitdrukking die we de spoorformule noemen.

$$\mathbb{E}[\int x^k d\mu_n] = \frac{1}{n^{k/2+1}} \sum_{1 \leq i_1, \dots, i_k \leq n} \mathbb{E}[W_{i_1, i_2} W_{i_2, i_3} \dots W_{i_k, i_1}]$$

Hiervoor gebruiken we begrippen en trucs uit de lineaire algebra.

Uitwerking

$\mathbb{E}[\int x^k d\mu_{sc}]$ lijkt er zoals in de inleiding al gezegd dubbelop uit te zien en dat is begrijpelijk. Voor bijvoorbeeld een enkele stochast Y en $y \in Y$ met verdelingsfunctie $g(y)$ geldt inderdaad $\int y^k g(y) dy = \mathbb{E}[y^k]$ en $\mathbb{E}[\int y^k g(y) dy] = \mathbb{E}[\mathbb{E}[y^k]] = \mathbb{E}[y^k]$.

Wil je daarenegen de verwachting uitrekenen voor een functie over meerdere stochasten, zeg bijvoorbeeld $\mathbb{E}[f(y_1, y_2, y_3)]$ met onafhankelijke stochasten Y_1, Y_2, Y_3 en verdelingsfuncties

$g_1(y_1), g_2(y_2), g_3(y_3)$, dan $\mathbb{E}[f(y_1, y_2, y_3)] = \int \int \int f(y_1, y_2, y_3) g_1(y_1) g_2(y_2) g_3(y_3) dy_1 dy_2 dy_3$

De integraal $\int x^k d\mu_{sc}$ kan je daarom ook beter opvatten als functie die afhangt van heel veel stochasten $W_{i,j}$ (zoals we gezien hebben in hoofdstuk 3), soortgelijk aan het voorbeeld met Y_1, Y_2, Y_3 hierboven, alleen dan dus veel groter.

Zoals in de inleiding gezegd, we gebruiken begrippen en trucs uit de lineaire algebra. Verder komt de naam spoorformule natuurlijk ook niet nergens vandaan. Daarom introduceren we het spoor en kijken we wat voor verband er is tussen het spoor en de eigenwaarden van een matrix.

Definitie 6.1 (Spoor).

Zij A een $n \times n$ -matrix. Dan is het **spoor** van A gedefinieerd als de som van de elementen op de diagonaal:

$$\text{Tr}(A) = \sum_{i=1}^n a_{ii}$$

De Engelse benaming van het spoor is de trace, vandaar de afkorting Tr.

Iets dat nu nodig is om eigenschappen voor het spoor aan te tonen is enig begrip en notatie bij herhaalde matrixvermenigvuldiging. Zo gebruiken we bijvoorbeeld de afkorting

$$\sum_{1 \leq k_1, \dots, k_{m-1} \leq n} \text{iets dat je wil optellen} = \sum_{k_1=1}^n \sum_{k_2=1}^n \dots \sum_{k_{m-1}=1}^n \text{iets dat je wil optellen}$$

De aankomende twee rekenregels (lemma 6.2 en gevolg 6.3) zijn in feite gewoon herhaalde matrixvermenigvuldiging, maar worden wel gegeven opdat het duidelijk is waar de lange sommaties in verdere lemma's en stellingen vandaan komen.

Lemma 6.2.

Zij A_1, A_2, \dots, A_m $n \times n$ -matrices. Dan

$$(A_1 A_2 \dots A_m)_{i,j} = \sum_{1 \leq k_1, \dots, k_{m-1} \leq n} (A_1)_{i,k_1} (A_2)_{k_1,k_2} \dots (A_m)_{k_{m-1},j}$$

Bewijs.

Voor het geval $m = 2$ is het duidelijk, dit is gewoon de definitie:

$$(A_1 A_2)_{i,j} = \sum_{k=1}^n (A_1)_{i,k} (A_2)_{k,j}$$

Neem nu aan dat het geldt voor zekere waarde van m , dan

$$\begin{aligned} (A_1 A_2 \dots A_m A_{m+1})_{j,k} &= \sum_{k=1}^n (A_1 A_2 \dots A_m)_{i,k} (A_{m+1})_{k,j} \\ &= \sum_{k=1}^n \left[\sum_{1 \leq k_1, \dots, k_{m-1} \leq n} (A_1)_{i,k_1} (A_2)_{k_1,k_2} \dots (A_m)_{k_{m-1},k} \right] (A_{m+1})_{k,j} \end{aligned}$$

Nu mogen we de haakjes weghalen en de volgorde van de sommaties veranderen.

$$= \sum_{1 \leq k_1, \dots, k_{m-1} \leq n} \sum_{k=1}^n (A_1)_{i, k_1} (A_2)_{k_1, k_2} \dots (A_m)_{k_{m-1}, k} (A_{m+1})_{k, j}$$

Als je in plaats van index k hier index k_m leest, staat er exact het lemma voor $m + 1$ matrices.

$$= \sum_{1 \leq k_1, \dots, k_m \leq n} (A_1)_{i, k_1} (A_2)_{k_1, k_2} \dots (A_m)_{k_{m-1}, k_m} (A_{m+1})_{k_m, j}$$

Uit het principe van volledige inductie weten we nu dat de uitspraak werkt voor alle waarden van $m \in \mathbb{N}$. \square

Gevolg 6.3.

$$(A^m)_{i, j} = \sum_{1 \leq k_1, \dots, k_{m-1} \leq n} A_{i, k_1} A_{k_1, k_2} \dots A_{k_{m-1}, j}$$

Bewijs.

Toepassen van lemma 6.2 met $A_1 = A_2 = \dots = A_m = A$. \square

Nu dat bekend is, kunnen we wat eigenschappen van het spoor aantonen, zoals het aparte gedrag dat je ziet als je het spoor neemt over een product van matrices.

Lemma 6.4 (cyclische eigenschap).

Het spoor is invariant onder cyclische permutaties, of met andere woorden: voor A_1, A_2, \dots, A_m $n \times n$ -matrices, geldt

$$\text{Tr}(A_1 A_2 \dots A_m) = \text{Tr}(A_2 A_3 \dots A_m A_1)$$

Bewijs.

We beginnen met de definitie van de trace

$$\text{Tr}(A_1 A_2 \dots A_m) = \sum_{i=1}^n (A_1 A_2 \dots A_m)_{i, i}$$

Nu volgens lemma 6.2 is dit gelijk aan

$$= \sum_{i=1}^n \sum_{1 \leq k_1, \dots, k_{m-1} \leq n} (A_1)_{i, k_1} (A_2)_{k_1, k_2} \dots (A_m)_{k_{m-1}, i}$$

Kunnen we volgordes verwisselen

$$= \sum_{1 \leq k_1, \dots, k_{m-1}, i \leq n} (A_2)_{k_1, k_2} \dots (A_m)_{k_{m-1}, i} (A_1)_{i, k_1}$$

$$= \sum_{k_1=1}^n (A_2 \dots A_m A_1)_{k_1, k_1}$$

$$= \text{Tr}(A_2 \dots A_m A_1)$$

\square

Lemma 6.5 (som van eigenwaarden).

Zij A een symmetrische $n \times n$ -matrix. Zij $\lambda_1(A), \dots, \lambda_n(A)$ de n eigenwaarden van A . Dan geldt

$$\text{Tr}(A) = \sum_{i=1}^n \lambda_i(A)$$

Bewijs.

A is een symmetrische matrix, dus A is diagonaliseerbaar: $A = PDP^{-1}$, waarbij P inverteerbaar en D een diagonaalmatrix.

Als we verder nog gebruik maken van lemma 6.4 vinden we dat

$$\text{Tr}(A) = \text{Tr}(PDP^{-1}) = \text{Tr}(DP^{-1}P) = \text{Tr}(DI) = \text{Tr}(D)$$

In de diagonaalmatrix D staan op de diagonaal de eigenwaarden van A , dus

$$\text{Tr}(A) = \text{Tr}(D) = \sum_{i=1}^n \lambda_i(A)$$

□

Stelling 6.6 (Spoorformule).

Voor onze Wigner matrix W geldt het volgende

$$\mathbb{E} \left[\int x^k d\mu_n \right] = \frac{1}{n^{k/2+1}} \sum_{1 \leq i_1, \dots, i_k \leq n} \mathbb{E} [W_{i_1, i_2} W_{i_2, i_3} \dots W_{i_k, i_1}]$$

Bewijs.

In het bewijs gebruiken we de definitie voor integreren over kansmaat μ_n uit hoofdstuk 3 en de zojuist bewezen lemma's.

Voor de eerste stap gebruiken we de kennis van hoofdstuk 3, integreren μ_n .

$$\mathbb{E} \left[\int x^k d\mu_n \right] = \mathbb{E} \left[\frac{1}{n} \sum_{i=1}^n \lambda_i \left(\frac{1}{\sqrt{n}} W \right)^k \right]$$

Gebruiken van lemma 6.5 is toegestaan, W is immers een symmetrische matrix. Dit geeft

$$= \mathbb{E} \left[\frac{1}{n} \text{Tr} \left(\left(\frac{1}{\sqrt{n}} W \right)^k \right) \right]$$

Omdat voor c een scalar en A een matrix geldt dat $\text{Tr}(cA) = c\text{Tr}(A)$ en we ook bij de verwachting de scalars naar buiten mogen halen kunnen we dit schrijven als

$$= \frac{1}{n^{k/2+1}} \mathbb{E} \left[\text{Tr}(W^k) \right]$$

Iets dat je wel op moet merken is dat je het spoor pakt van de matrix W^k en dat dit heel iets anders is dan de k -de macht nemen van het spoor.

$$= \frac{1}{n^{k/2+1}} \mathbb{E} \left[\sum_{i=1}^n (W^k)_{i,i} \right]$$

Nu gebruiken van gevolg 6.3

$$= \frac{1}{n^{k/2+1}} \mathbb{E} \left[\sum_{i=1}^n \sum_{1 \leq i_1, \dots, i_{k-1} \leq n} W_{i, i_1} W_{i_1, i_2} \dots W_{i_{k-1}, i} \right]$$

En dan je indexjes anders nummeren: $i \mapsto i_1, i_1 \mapsto i_2, \dots, i_{k-1} \mapsto i_k$.

$$= \frac{1}{n^{k/2+1}} \mathbb{E} \left[\sum_{1 \leq i_1, \dots, i_k \leq n} W_{i_1, i_2} W_{i_2, i_3} \dots W_{i_k, i_1} \right]$$

$$= \frac{1}{n^{k/2+1}} \sum_{1 \leq i_1, \dots, i_k \leq n} \mathbb{E} [W_{i_1, i_2} W_{i_2, i_3} \dots W_{i_k, i_1}]$$

□

Hoofdstuk 7

Grafenformule

Inleiding

Uit vorige hoofdstukken weten we dat

$$\mathbb{E} \left[\int x^k d\mu_n \right] = \frac{1}{n^{k/2+1}} \sum_{1 \leq i_1, \dots, i_k \leq n} \mathbb{E} [W_{i_1, i_2} W_{i_2, i_3} \dots W_{i_k, i_1}]$$

Doel is nu om die sommatie handiger te schrijven. Dit kan door de combinaties $\mathbf{i} = (i_1, \dots, i_k) \in \{1, \dots, n\}^k$ beter te analyseren en ook als graaf of als wandeling over een graaf te beschouwen.

Verder in het hoofdstuk bekijken we ook equivalentieclasses voor deze grafen en wandelingen. Deze equivalentieclasses bepalen voor een groot deel hoe we onze sommatie opschrijven en hoe we er aan gaan rekenen.

We sluiten het hoofdstuk af met de grafenformule, een tussenresultaat in het berekenen van de bewuste integraal.

$$\mathbb{E} \left[\int x^k d\mu_n \right] = \frac{1}{n^{k/2+1}} \sum_{t=1}^{\lfloor k/2 \rfloor + 1} \sum_{\mathbf{i} \in R_{k,t}} \sum_{\mathbf{j} \sim \mathbf{i}} \mathbb{E} \left[\prod_{\vec{e} \in E(\mathbf{j})} W_{\vec{e}} \right]$$

Dit ziet er misschien nog niet uit als een verbetering, maar omdat meer elementen die qua structuur sterk op elkaar lijken nu beter in de sommatie gesorteerd staan gaat het rekenen nu een stuk makkelijker zijn.

Dat rekenen volgt dan in het volgende hoofdstuk.

Uitwerking

Voor de combinaties $\mathbf{i} = (i_1, \dots, i_k) \in \{1, \dots, n\}^k$ introduceren we een aantal notaties. Een met betrekking tot verwachtingen, $C(\mathbf{i})$, en een aantal met betrekking tot grafen: $V(\mathbf{i})$, $E(\mathbf{i})$, $\tilde{E}(\mathbf{i})$, $G(\mathbf{i})$ en $P(\mathbf{i})$.

Notatie 7.1.

Met betrekking tot verwachtingen:

$$C(\mathbf{i}) = \mathbb{E} [W_{i_1, i_2} W_{i_2, i_3} \dots W_{i_k, i_1}]$$

Met betrekking tot grafen:

$V(\mathbf{i}) = \{i_1, \dots, i_k\}$; de knopen, zonder multipliciteiten

$E(\mathbf{i}) = \{(i_1, i_2), \dots, (i_k, i_1)\}$; de gerichtte zijden, met multipliciteiten

$\tilde{E}(\mathbf{i}) = \{(i_1, i_2), \dots, (i_k, i_1)\}$; de zijden, zonder multipliciteiten of richting

$G(\mathbf{i}) = (V(\mathbf{i}), \tilde{E}(\mathbf{i}))$; het skelet van \mathbf{i} , een ongerichte graaf zonder dubbele zijden.

$P(\mathbf{i}) = \{i_1, (i_1, i_2), i_2, \dots, i_k, (i_k, i_1), i_1\}$; een wandeling over het skelet $G(\mathbf{i})$

Puntje om op te merken is dat $C(\mathbf{i})$ een getal is terwijl de overige verzamelingen zijn. De V gebruiken we vanwege het Engelse woord vertices, de E komt van edges, \tilde{E} komt ook van edges, waarbij de tilde is om aan te geven dat de verzameling weliswaar op E lijkt, maar wel degelijk anders is, de G komt van het Engelse graph, de P komt tenslotte van het Engelse path.

Voorbeeld 7.2.

We bekijken in dit voorbeeld $\mathbf{i} = (1, 4, 3, 6, 4, 3, 4)$. Het is duidelijk dat $\mathbf{i} \in \{1, \dots, 6\}^7$. Dit voorbeeld is klein genoeg dat het overzichtelijk blijft maar tegelijkertijd zitten voor de bepaling van $V(\mathbf{i})$, $E(\mathbf{i})$ en $\tilde{E}(\mathbf{i})$ alle interessante scenarios erin. Dan:

$$C(\mathbf{i}) = \mathbb{E} [W_{1,4}W_{4,3}W_{3,6}W_{6,4}W_{4,3}W_{3,4}W_{4,1}]$$

$$V(\mathbf{i}) = \{1, 3, 4, 6\}$$

$$E(\mathbf{i}) = \{(1, 4), (4, 3), (3, 6), (6, 4), (4, 3), (3, 4), (4, 1)\}$$

$$\tilde{E}(\mathbf{i}) = \{(1, 4), (3, 4), (3, 6), (4, 6)\}$$

$$G(\mathbf{i}) = (\{1, 3, 4, 6\}, \{(1, 4), (3, 4), (3, 6), (4, 6)\})$$

$$P(\mathbf{i}) = \{1, (1, 4), 4, (4, 3), 3, (3, 6), 6, (6, 4), 4, (4, 3), 3, (3, 4), 4, (4, 1), 1\}$$

Toelichting: De verzameling $V(\mathbf{i})$ volgt uit het weglaten van de dubbele elementen van \mathbf{i} en vervolgens de overgebleven nummers op volgorde te zetten.

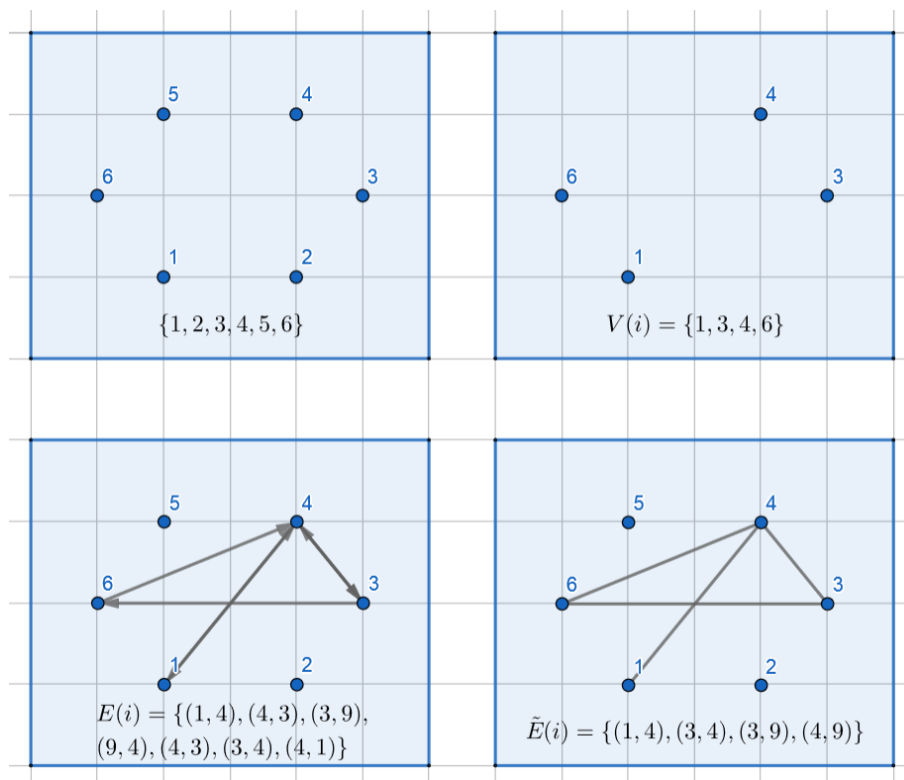
De verzameling $E(\mathbf{i})$ bestaat uit de koppels (i_1, i_2) tot en met (i_k, i_1) . Deze laatste is wel belangrijk om mee te nemen. \mathbf{i} eindigt dan wel op $\dots, 3, 4$, voor $E(\mathbf{i})$ moet ook nog $(4, 1)$ meege-
nomen worden.

Voor $\tilde{E}(\mathbf{i})$ moeten de multipliciteiten uit $E(\mathbf{i})$ verwijderd worden. Het koppel $(4, 3)$ komt dubbel voor, dus daarvan moet er een weg. Ook bekijken we de koppels nu als ware ze richtingloos, met andere woorden we zien $(1, 4)$ en $(4, 1)$ als zijnde hetzelfde koppel, waardoor een van deze twee uit de verzameling gehaald moet worden. Hetzelfde geldt voor $(4, 3)$ en $(3, 4)$. Dit geeft uiteindelijk voor $\tilde{E}(\mathbf{i})$ de verzameling als dat die hierboven is weergegeven.

$G(\mathbf{i})$ is wat je krijgt als je tegelijkertijd naar $V(\mathbf{i})$ en $\tilde{E}(\mathbf{i})$ kijkt.

$P(\mathbf{i})$ verkrijgt je tenslotte door te beginnen bij het eerste punt i_1 , in dit geval 1, gevolgd door de zijde naar het tweede punt, in dit geval $(1,4)$, gevolgd door dit tweede punt, oftewel 4, gevolgd door de zijde van het tweede punt naar het derde punt, enzovoorts. Ook hier is het belangrijk dat je uiteindelijk eindigt terug in het beginpunt.

Ook kunnen we deze verzamelingen visueel weergeven, zoals in afbeelding 7.1



Figuur 7.1: Visuele weergave voorbeeldgraaf

Lemma 7.3.

Zij $\vec{e}_1, \dots, \vec{e}_k$ de elementen van $E(\mathbf{i})$, oftewel de zijden van de graaf behorend bij de combinatie \mathbf{i} . Zij \vec{e}_j een zijde waarvoor geldt dat \vec{e}_j en zijn omgekeerde $\overleftarrow{\vec{e}_j}$ samen slechts een keer voorkomen in $E(\mathbf{i})$.

Dan $C(\mathbf{i}) = 0$.

Bewijs.

Als \vec{e}_j en zijn omgekeerde $\overleftarrow{\vec{e}_j}$ samen slechts een keer voorkomen in $E(\mathbf{i})$, dan weten we dat de random variabele $W_{\vec{e}_j}$ onafhankelijk is van de andere random variabelen $\{W_{\vec{e}} : \vec{e} \neq \vec{e}_j, \vec{e} \in E(\mathbf{i})\}$.

Dit weten we vanuit de voorwaarden op de matrix W , alle random variabelen op en boven de diagonaal zijn onafhankelijk van elkaar.

Herinner dat voor onafhankelijke random variabelen X, Y geldt dat $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$. Deze rekenregel kunnen we nu ook gebruiken voor het uitrekenen van $C(\mathbf{i})$.

$$\begin{aligned}
 C(\mathbf{i}) &= \mathbb{E} [W_{i_1, i_2} W_{i_2, i_3} \dots W_{i_k, i_1}] \\
 &= \mathbb{E} \left[\prod_{\vec{e}} W_{\vec{e}} \right] \\
 &= \mathbb{E} \left[W_{\vec{e}_j} \prod_{\vec{e} \neq \vec{e}_j} W_{\vec{e}} \right] \\
 &= \mathbb{E} [W_{\vec{e}_j}] \mathbb{E} \left[\prod_{\vec{e} \neq \vec{e}_j} W_{\vec{e}} \right]
 \end{aligned}$$

De verwachting van $W_{\vec{e}_j}$ is bekend wegens de voorwaarden op de matrix W , deze verwachting is namelijk gelijk aan nul. Er geldt dus ook

$$C(\mathbf{i}) = 0$$

□

We bekijken de invloed van lemma 7.3 op voorbeeld 7.2. Dan $\vec{e}_1 = (1, 4)$, $\vec{e}_2 = (4, 3)$, enzovoort. Merk op dat bijvoorbeeld $\vec{e}_1 = (1, 4)$ en $\vec{e}_7 = (4, 1)$ elkaars omgekeerde zijn, je zou dus kunnen zeggen dat $\vec{e}_7 = \check{e}_1$. \vec{e}_1 en zijn omgekeerde komen samen dus twee keer voor in $E(\mathbf{i})$. \vec{e}_3 komt echter maar een keer voor in $E(\mathbf{i})$ en zijn omgekeerde is helemaal afwezig. In voorbeeld 7.2 geldt dus $C(\mathbf{i}) = 0$.

Een handig gevolg van lemma 7.3 is het aankomende lemma 7.5. Om dat te bewijzen moet wel eerst het volgende worden opgemerkt.

Opmerking 7.4.

Met $V(\mathbf{i})$, $\tilde{E}(\mathbf{i})$ en $G(\mathbf{i})$ gedefinieerd als in 7.1 is $G(\mathbf{i})$ een verbonden graaf.

Omdat $G(\mathbf{i})$ een verbonden graaf is geldt dat $|V(\mathbf{i})| \leq |\tilde{E}(\mathbf{i})| + 1$ met gelijkheid enkel als $G(\mathbf{i})$ een boom is.

Deze ongelijkheid wordt ook in het volgende hoofdstuk nog een keer gebruikt.

Lemma 7.5.

Zij \mathbf{i} een combinatie van lengte k waarvoor $C(\mathbf{i}) \neq 0$

Dan geldt $|\tilde{E}(\mathbf{i})| \leq \lfloor \frac{k}{2} \rfloor$ en $|V(\mathbf{i})| \leq \lfloor \frac{k}{2} \rfloor + 1$.

Bewijs.

We hebben in lemma 7.3 geleerd dat als er een zijde \vec{e}_j bestaat in $E(\mathbf{i})$ waarvoor \vec{e}_j en zijn omgekeerde \check{e}_j samen slechts een keer voorkomen, dat we dan vinden dat $C(\mathbf{i}) = 0$.

Willen we dat $C(\mathbf{i}) \neq 0$, dan moet dus gelden dat voor alle $\vec{e} \in E(\mathbf{i})$ steeds \vec{e} en zijn omgekeerde samen ten minste twee keer voorkomen.

Als dit het geval is moet $\tilde{E}(\mathbf{i})$ tenminste de helft kleiner zijn dan $E(\mathbf{i})$. Dus

$$|\tilde{E}(\mathbf{i})| \leq \frac{|E(\mathbf{i})|}{2} = \frac{k}{2}$$

Omdat $|\tilde{E}(\mathbf{i})|$ een geheel getal moet zijn is dit equivalent met

$$|\tilde{E}(\mathbf{i})| \leq \left\lfloor \frac{k}{2} \right\rfloor$$

Bedenk dat voor een ongerichtte verbonden graaf het aantal punten nooit meer kan zijn het aantal zijdes + 1. Omdat $G(\mathbf{i})$ zo 'n ongerichtte verbonden graaf is volgt nu dat

$$|V(\mathbf{i})| \leq |\tilde{E}(\mathbf{i})| + 1 \leq \left\lfloor \frac{k}{2} \right\rfloor + 1$$

□

Notatie 7.6.

Definieer een equivalentierelatie \sim over de combinaties op $\{1, \dots, n\}^k$.
 Noem twee combinaties \mathbf{i} en \mathbf{j} equivalent, $\mathbf{j} \sim \mathbf{i}$, als er een bijectie $\pi : \{1, \dots, n\} \rightarrow \{1, \dots, n\}$ bestaat zo dat $(j_1, \dots, j_k) = (\pi(i_1), \dots, \pi(i_k))$.

Het is duidelijk dat twee combinaties enkel equivalent kunnen zijn als ze hetzelfde aantal stappen k nemen en hetzelfde aantal verschillende punten t bezoeken.

In voorbeeld 7.2 keken we naar de combinatie $\mathbf{i} = \{1, 4, 3, 6, 4, 3, 4\}$. Een voorbeeld van een equivalente combinatie $\mathbf{j} \sim \mathbf{i}$ is $\mathbf{j} = \{2, 3, 4, 1, 3, 4, 3\}$.

Een voorbeeld van een bijectie is in cykelnotatie: $\pi = (1\ 2\ 6)(3\ 4)$

We verdelen $\{1, \dots, n\}^k$ in equivalentieklassen, in verzamelingen van combinaties die equivalent zijn met een bepaalde combinatie.

Voor elke equivalentieklasse nemen we een vertegenwoordiger, een voor de hand liggende keuze voor deze vertegenwoordiger is die combinatie die begint bij 1 en voor iedere keer dat je een nieuw punt aan doet noemen we deze een hoger dan de hoogste die je eerder hebt aangedaan.

De logische vertegenwoordiger van de equivalentieklasse waarin ook

$\mathbf{i} = \{1, 4, 3, 6, 4, 3, 4\}$ zit is dan bijvoorbeeld $\{1, 2, 3, 4, 2, 3, 2\}$.

Laat $R_{k,t}$ de verzameling zijn van alle vertegenwoordigers van equivalentieklassen met k stappen op t punten.

Met al deze nieuwe notatie kunnen we

$$\mathbb{E} \left[\int x^k d\mu_n \right] = \frac{1}{n^{k/2+1}} \sum_{1 \leq i_1, \dots, i_k \leq n} \mathbb{E} [W_{i_1, i_2} W_{i_2, i_3} \dots W_{i_k, i_1}]$$

anders opschrijven. Door lemma 7.5 te gebruiken kunnen we bovendien concluderen dat we over minder combinaties hoeven te sommeren.

Stelling 7.7 (Grafenformule).

Voor onze Wigner matrix W geldt

$$\begin{aligned} \mathbb{E} \left[\int x^k d\mu_n \right] &= \frac{1}{n^{k/2+1}} \sum_{1 \leq i_1, \dots, i_k \leq n} \mathbb{E} [W_{i_1, i_2} W_{i_2, i_3} \dots W_{i_k, i_1}] \\ &= \frac{1}{n^{k/2+1}} \sum_{t=1}^{\lfloor k/2 \rfloor + 1} \sum_{\mathbf{i} \in R_{k,t}} \sum_{\mathbf{j} \sim \mathbf{i}} \mathbb{E} \left[\prod_{\vec{e} \in E(\mathbf{j})} W_{\vec{e}} \right] \end{aligned}$$

Bewijs.

Voor dit bewijs is het het belangrijkste om te realiseren dat de twee sommaties hieronder eigenlijk

hetzelfde zijn, op voorwaarde dan dat wat links en rechts op de puntjes staat hetzelfde is.

$$\sum_{1 \leq i_1, \dots, i_k \leq n} \dots = \sum_{t=1}^{k+1} \sum_{\mathbf{i} \in R_{k,t}} \sum_{\mathbf{j} \sim \mathbf{i}} \dots$$

Het is duidelijk dat de linkersommatie sommeert over de combinaties $\mathbf{i} = \{i_1, \dots, i_k\}^k$, waarbij voor elke mogelijke van deze combinaties er exact een keer iets wordt opgeteld.

In de rechtersommatie is dit ook het geval, bekijk zo bijvoorbeeld een willekeurige combinatie $\mathbf{j} = \{j_1, \dots, j_k\}^k$, we nemen hier een j in plaats van een i omdat dat met de notatie wat beter uitkomt.

Dan is \mathbf{j} equivalent met exact een vertegenwoordiger \mathbf{i} , welke vervolgens exact een keer voorkomt in de verzameling van vertegenwoordigers $R_{k,t}$. Ook is met de keuze van \mathbf{j} bepaald hoe veel verschillende punten t de combinatie bezoekt, waarbij $1 \leq t \leq k+1$ gezien de eigenschappen van grafen.

Dus

$$\begin{aligned} \mathbb{E} \left[\int x^k d\mu_n \right] &= \frac{1}{n^{k/2+1}} \sum_{1 \leq i_1, \dots, i_k \leq n} \mathbb{E} [W_{i_1, i_2} W_{i_2, i_3} \dots W_{i_k, i_1}] \\ &= \frac{1}{n^{k/2+1}} \sum_{t=1}^{k+1} \sum_{\mathbf{i} \in R_{k,t}} \sum_{\mathbf{j} \sim \mathbf{i}} \mathbb{E} [W_{i_1, i_2} W_{i_2, i_3} \dots W_{i_k, i_1}] \end{aligned}$$

Verder weten we uit lemma 7.5 dat als $t > \lfloor \frac{k}{2} \rfloor + 1$, dat dan $C(\mathbf{i}) = 0$. We hoeven dus over minder verschillende waarden voor t te sommeren.

$$= \frac{1}{n^{k/2+1}} \sum_{t=1}^{\lfloor k/2 \rfloor + 1} \sum_{\mathbf{i} \in R_{k,t}} \sum_{\mathbf{j} \sim \mathbf{i}} \mathbb{E} [W_{i_1, i_2} W_{i_2, i_3} \dots W_{i_k, i_1}]$$

Een andere schrijfwijze van het matrixproduct in de verwachting geeft dan

$$= \frac{1}{n^{k/2+1}} \sum_{t=1}^{\lfloor k/2 \rfloor + 1} \sum_{\mathbf{i} \in R_{k,t}} \sum_{\mathbf{j} \sim \mathbf{i}} \mathbb{E} \left[\prod_{\vec{e} \in E(\mathbf{j})} W_{\vec{e}} \right]$$

□

Hoofdstuk 8

Grafenformule voor grote n

Inleiding

In het vorige hoofdstuk hebben we een uitdrukking gevonden voor de verwachting van het k -de moment, dit in de vorm van een gigantische sommatie over verwachtingen van matrixproducten. Ook hebben we gezien dat voor een flink deel van deze matrixproducten we verwachting nul vinden. Die kunnen we dus uit de sommatie gooien. We waren geeindigd met

$$\mathbb{E} \left[\int x^k d\mu_n \right] = \frac{1}{n^{k/2+1}} \sum_{t=1}^{\lfloor k/2 \rfloor + 1} \sum_{\mathbf{i} \in R_{k,t}} \sum_{\mathbf{j} \sim \mathbf{i}} \mathbb{E} \left[\prod_{\vec{e} \in E(\mathbf{j})} W_{\vec{e}} \right]$$

In dit hoofdstuk zullen we zien dat voor grote matrixgrootte n we nog veel meer combinaties kunnen elimineren. Voor de ene groep gevallen zal het zijn omdat de verwachting willekeurig klein en daarmee niet meer interessant wordt, voor een andere groep gaan de verwachtingen daadwerkelijk gelijk aan nul zijn.

Er worden afschattingen gemaakt en er wordt ook weer gekeken naar de eigenschappen van de graaf die hoort bij een bepaalde combinatie, als ook de eigenschappen op de matrix W .

Op een gegeven moment wordt ook het onderscheid gemaakt tussen oneven waarden van k en even waarden van k .

Voor oneven waarden van k zal de verwachting van het k -de moment willekeurig dicht bij nul komen te liggen. Voor even waarden van k eindigen we het hoofdstuk met een telprobleem. Dit telprobleem schuiven we dan door naar hoofdstuk 9.

Uitwerking

Lemma 8.1.

Twee handige afschattingen zijn

$$\left| \mathbb{E} \left[\prod_{\vec{e} \in E(\mathbf{j})} W_{\vec{e}} \right] \right| \leq T^k \text{ en}$$

$$|\{\mathbf{j} : \mathbf{j} \sim \mathbf{i}\}| \leq n^t$$

Bewijs.

In de voorwaarden van de matrix W staat dat iedere $W_{\vec{e}}$ in absolute waarde begrensd is door $T \in \mathbb{R}$. Dan is het product van k van zulke elementen in absolute waarde dan ook begrensd door T^k .

Als $\left| \prod_{\vec{e} \in E(\mathbf{j})} W_{\vec{e}} \right| \leq T^k$, dan ook zeker $\left| \mathbb{E} \left[\prod_{\vec{e} \in E(\mathbf{j})} W_{\vec{e}} \right] \right| \leq T^k$

Het bepalen van het aantal combinaties dat binnen een equivalentieklasse (specifiek één met k stappen en t verschillende punten) zit, komt neer op het tellen op hoeveel manieren je t punten uit een mogelijke n punten kunt tellen.

Herhaling is niet toegestaan. Dit geeft $|\{\mathbf{j} : \mathbf{j} \sim \mathbf{i}\}| = \frac{n!}{(n-t)!}$.

De rest volgt uit uitschrijven van de faculteiten en het feit dat $n-1, \dots, n-t+1$ allemaal kleiner zijn dan n .

$$|\{\mathbf{j} : \mathbf{j} \sim \mathbf{i}\}| = \frac{n!}{(n-t)!} = \frac{n(n-1)\dots(1)}{(n-t)(n-t-1)\dots(1)} = n(n-1)\dots(n-t+1) \leq n^t$$

□

Met hulp van deze afschattingen kunnen we de t -de sommand in de grafenformule begrenzen.

Gevolg 8.2 (Afschatting t -de sommand grafenformule).

De t -de sommand van de grafenformule kunnen we begrenzen door

$$\left| \frac{1}{n^{k/2+1}} \sum_{\mathbf{i} \in R_{k,t}} \sum_{\mathbf{j} \sim \mathbf{i}} \mathbb{E} \left[\prod_{\vec{e} \in E(\mathbf{j})} W_{\vec{e}} \right] \right| \leq |R_{k,t}| T^k \frac{n^t}{n^{k/2+1}}$$

Bewijs.

$$\left| \frac{1}{n^{k/2+1}} \sum_{\mathbf{i} \in R_{k,t}} \sum_{\mathbf{j} \sim \mathbf{i}} \mathbb{E} \left[\prod_{\vec{e} \in E(\mathbf{j})} W_{\vec{e}} \right] \right| \leq \frac{1}{n^{k/2+1}} \sum_{\mathbf{i} \in R_{k,t}} \sum_{\mathbf{j} \sim \mathbf{i}} \left| \mathbb{E} \left[\prod_{\vec{e} \in E(\mathbf{j})} W_{\vec{e}} \right] \right|$$

Afschatting 1 van lemma 8.1 geeft

$$\leq \frac{1}{n^{k/2+1}} \sum_{\mathbf{i} \in R_{k,t}} \sum_{\mathbf{j} \sim \mathbf{i}} T^k$$

In de sommatie over $\mathbf{j} \sim \mathbf{i}$ sommeren we nu steeds over hetzelfde, dus

$\sum_{\mathbf{j} \sim \mathbf{i}} T^k = |\{\mathbf{j} : \mathbf{j} \sim \mathbf{i}\}| T^k$. In afschatting 2 van lemma 8.1 hebben we een bovengrens voor $|\{\mathbf{j} : \mathbf{j} \sim \mathbf{i}\}|$, dus

$$\leq \frac{1}{n^{k/2+1}} \sum_{\mathbf{i} \in R_{k,t}} n^t T^k$$

De breuk die nu nog voor de sommatie staat kunnen we natuurlijk ook in de sommatie zetten

$$\begin{aligned} &= \sum_{\mathbf{i} \in R_{k,t}} T^k \frac{n^t}{n^{k/2+1}} \\ &= |R_{k,t}| T^k \frac{n^t}{n^{k/2+1}} \end{aligned}$$

□

De breuk $\frac{n^t}{n^{k/2+1}}$ maakt nu dat voor grote matrixgrootte n de t -de sommand vaak naar nul convergeert, immers

$$\frac{n^t}{n^{k/2+1}} \rightarrow 0 \text{ voor } n \rightarrow \infty \text{ als } t < \frac{k}{2} + 1$$

Lemma 8.3 (Convergentie t -de sommand grafenformule).

De t -de sommand van de grafenformule is $o(1)$ voor alle $t \leq \lfloor \frac{k}{2} \rfloor$ voor even getallen k en voor alle $t \leq \lfloor \frac{k}{2} \rfloor + 1$ voor oneven getallen k .

Hierbij betekent $o(1)$ dat voor iedere $\epsilon > 0$ een getal $N \in \mathbb{R}$ bestaat zo dat je functie in absolute waarde kleiner is dan ϵ voor alle $n \geq N$.

Deze notatie komt overeen met de kleine o notatie $f(x) = o(g(x))$ voor $x \rightarrow \infty$ toegepast op f de t -de sommand, g de functie 1 en het bekijken van gehele getallen n .

Bewijs.

Zij $t < \frac{k}{2} + 1$.

Zij $\epsilon > 0$ willekeurig. Zij $N = \left(\frac{\epsilon}{|R_{k,t}|T^k} \right)^{k/2+1-t}$. Dan voor $n \geq N$

$$\begin{aligned} \left| |R_{k,t}|T^k \frac{n^t}{n^{k/2+1}} \right| &\leq |R_{k,t}|T^k \frac{N^t}{N^{k/2+1}} \\ &= |R_{k,t}|T^k N^{t-k/2-1} \end{aligned}$$

Substitueer $N = \epsilon^{k/2+1-t}$

$$\begin{aligned} &= |R_{k,t}|T^k \left(\left(\frac{\epsilon}{|R_{k,t}|T^k} \right)^{k/2+1-t} \right)^{t-k/2-1} \\ &= |R_{k,t}|T^k \frac{\epsilon}{|R_{k,t}|T^k} \\ &= \epsilon \end{aligned}$$

Dus we hebben bewezen dat de t -de sommand $o(1)$ is voor $t < \frac{k}{2} + 1$

Zij nu k een even getal en t geheel, dan is $t < \frac{k}{2} + 1$ equivalent met $t \leq \frac{k}{2}$.

Voor k een even getal is $\frac{k}{2}$ geheel, dus het naar beneden afronden verandert verder niets: $t < \frac{k}{2} + 1$ is equivalent met $t \leq \lfloor \frac{k}{2} \rfloor$.

Dus de t -de sommand is bij even getallen k $o(1)$ voor $t \leq \lfloor \frac{k}{2} \rfloor$.

Zij nu k een oneven getal en t geheel, dan is $t < \frac{k}{2} + 1$ equivalent met $t \leq \lfloor \frac{k}{2} \rfloor + 1$.

Dus de t -de sommand is bij oneven getallen k $o(1)$ voor $t \leq \lfloor \frac{k}{2} \rfloor + 1$. □

Omdat we in de grafenformule t sommeren van 1 tot $\lfloor \frac{k}{2} \rfloor + 1$ krijgen we de volgende interessante resultaten. Stelling 8.4 voor k oneven en stelling 8.5 voor k even.

Stelling 8.4 (Verwachting oneven momenten).

Voor k oneven geldt

$$\mathbb{E} \left[\int x^k d\mu_n \right] \rightarrow 0 \text{ voor } n \rightarrow \infty$$

Bewijs.

Beginnen met de grafenformule als aan het einde van hoofdstuk 7:

$$\mathbb{E} \left[\int x^k d\mu_n \right] = \frac{1}{n^{k/2+1}} \sum_{t=1}^{\lfloor k/2 \rfloor + 1} \sum_{\mathbf{i} \in R_{k,t}} \sum_{\mathbf{j} \sim \mathbf{i}} \mathbb{E} \left[\prod_{\vec{e} \in E(\mathbf{j})} W_{\vec{e}} \right]$$

Haal de breuk $\frac{1}{n^{k/2+1}}$ binnen de sommatie

$$= \sum_{t=1}^{\lfloor k/2 \rfloor + 1} \frac{1}{n^{k/2+1}} \sum_{\mathbf{i} \in R_{k,t}} \sum_{\mathbf{j} \sim \mathbf{i}} \mathbb{E} \left[\prod_{\vec{e} \in E(\mathbf{j})} W_{\vec{e}} \right]$$

Uit lemma 8.3 weten we nu dat voor oneven machten k voor $t = 1$ tot $\lfloor \frac{k}{2} \rfloor + 1$ de t -de sommand $o(1)$ is. Dus

$$\begin{aligned} &= \sum_{t=1}^{\lfloor k/2 \rfloor + 1} o(1) \\ &= o(1) \end{aligned}$$

Oftewel $\mathbb{E} \left[\int x^k d\mu_n \right] \rightarrow 0$ voor $n \rightarrow \infty$. □

De even momenten geven een stuk meer werk. We moeten nog een keer wat beter naar de grafen kijken en omdat niet alles nul of $o(1)$ wordt moeten we ook meer rekenen.

Stelling 8.5 (Verwachting even momenten).

Voor k even geldt

$$\mathbb{E} \left[\int x^k d\mu_n \right] \rightarrow |S_{k,k/2+1}| \text{ voor } n \rightarrow \infty$$

Waarbij $S_{k,k/2+1}$ de verzameling is die bestaat uit combinaties \mathbf{i} waarvoor:

- \mathbf{i} heeft lengte k en doet $\frac{k}{2} + 1$ verschillende punten aan
- \mathbf{i} is een vertegenwoordiger van een equivalentieklasse
- $G(\mathbf{i})$ is een boom
- $P(\mathbf{i})$ is een wandeling die elke zijde in $G(\mathbf{i})$ exact 2 maal aandoet (eenmaal in beide richtingen)

Bewijs.

Als in het bewijs van stelling 8.4 gebruiken we

$$\mathbb{E} \left[\int x^k d\mu_n \right] = \sum_{t=1}^{\lfloor k/2 \rfloor + 1} \frac{1}{n^{k/2+1}} \sum_{\mathbf{i} \in R_{k,t}} \sum_{\mathbf{j} \sim \mathbf{i}} \mathbb{E} \left[\prod_{\vec{e} \in E(\mathbf{j})} W_{\vec{e}} \right]$$

Uit lemma 8.3 weten we dat voor even machten k voor $t = 1$ tot $\lfloor \frac{k}{2} \rfloor$ de t -de sommand $o(1)$ is. Dus

$$= \frac{1}{n^{k/2+1}} \sum_{\mathbf{i} \in R_{k,k/2+1}} \sum_{\mathbf{j} \sim \mathbf{i}} \mathbb{E} \left[\prod_{\vec{e} \in E(\mathbf{j})} W_{\vec{e}} \right] + o(1) \quad (*)$$

De combinaties \mathbf{i} en \mathbf{j} waar nu over gesommeerd wordt zijn precies die combinaties met k stappen en die daarbij $\frac{k}{2} + 1$ verschillende punten aandoet. Ook in deze sommatie zitten nog veel combinaties \mathbf{i} en \mathbf{j} waarvoor de verwachting over die combinatie nul is.

Voor de combinaties in (*) geldt $|E(\mathbf{i})| = k$ en $|V(\mathbf{i})| = \frac{k}{2} + 1$.

We willen iets zeggen over die combinaties hierin die geen verwachting nul opleveren.

Herinner lemma 7.5: Voor \mathbf{i} een combinatie die geen verwachting nul geeft geldt $|\tilde{E}(\mathbf{i})| \leq \lfloor \frac{k}{2} \rfloor$, waarbij we wegens het even zijn van k mogen zeggen dat $|\tilde{E}(\mathbf{i})| \leq \frac{k}{2}$.

Dus voor \mathbf{i} een combinatie met verwachting ongelijk aan nul geldt $|\tilde{E}(\mathbf{i})| + 1 \leq \frac{k}{2} + 1 = V(\mathbf{i})$.

Maar uit opmerking ... weten we dat $|V(\mathbf{i})| \leq |\tilde{E}(\mathbf{i})| + 1$ met gelijkheid precies wanneer het skelet van \mathbf{i} , $G(\mathbf{i}) = (V(\mathbf{i}), \tilde{E}(\mathbf{i}))$, een boom is.

Combineren van deze uitspraken geeft dat de enige interessante combinaties \mathbf{i} voor onze sommatie die combinaties zijn waarvoor het skelet een boom is.

Het is duidelijk dat als \mathbf{i} een combinatie is die voldoet aan alle voorwaarden uit de vorige alinea, dat dan ook alle combinaties \mathbf{j} in zijn equivalentieklasse daaraan voldoen. Andersom als \mathbf{i} hier niet aan voldoet, dan doen de combinaties in zijn equivalentieklasse dit ook niet.

Definieer nu $S_{k,k/2+1}$ als de deelverzameling van $R_{k,k/2+1}$ die bestaat uit combinaties \mathbf{i} waarvoor:

- \mathbf{i} heeft lengte k en doet $\frac{k}{2} + 1$ verschillende punten aan
- \mathbf{i} is een vertegenwoordiger van een equivalentieklasse
- $G(\mathbf{i})$ is een boom
- $P(\mathbf{i})$ is een wandeling die elke zijde in $G(\mathbf{i})$ exact 2 maal aandoet (eenmaal in beide richtingen)

De eerste twee eigenschappen worden overgenomen van $R_{k,k/2+1}$, de derde eigenschap zorgt ervoor dat je de combinaties die verwachting nul geven kwijtraakt en eigenschap vier is een gevolg van de eerste en de derde.

We kunnen nu sommatie over $R_{k,k/2+1}$ vervangen door sommatie over $S_{k,k/2+1}$.

$$\begin{aligned} \mathbb{E} \left[\int x^k d\mu_n \right] &= \frac{1}{n^{k/2+1}} \sum_{\mathbf{i} \in R_{k,k/2+1}} \sum_{\mathbf{j} \sim \mathbf{i}} \mathbb{E} \left[\prod_{\vec{e} \in E(\mathbf{j})} W_{\vec{e}} \right] + o(1) \\ &= \frac{1}{n^{k/2+1}} \sum_{\mathbf{i} \in S_{k,k/2+1}} \sum_{\mathbf{j} \sim \mathbf{i}} \mathbb{E} \left[\prod_{\vec{e} \in E(\mathbf{j})} W_{\vec{e}} \right] + o(1) \end{aligned}$$

Nu we alleen nog kijken naar combinaties in $S_{k,k/2+1}$ of equivalent met combinaties in $S_{k,k/2+1}$ kunnen we uitspraken doen over de sommand $\mathbb{E} \left[\prod_{\vec{e} \in E(\mathbf{j})} W_{\vec{e}} \right]$. Omdat je elke zijde exact twee keer doorloopt, een keer in beide richtingen, geldt het volgende.

$$\mathbb{E} \left[\prod_{\vec{e} \in E(\mathbf{j})} W_{\vec{e}} \right] = \mathbb{E} \left[\prod_{\vec{e} \in \tilde{E}(\mathbf{j})} W_{\vec{e}} W_{\overleftarrow{\vec{e}}} \right]$$

Merk op: links zonder tilde, rechts met tilde.
 Gebruik de symmetrie van de matrix W .

$$= \prod_{\vec{e} \in \tilde{E}(\mathbf{j})} \mathbb{E} [W_{\vec{e}}^2] = \prod_{\vec{e} \in \tilde{E}(\mathbf{j})} \text{Var}(W_{\vec{e}})$$

De variantie van alle matrixelementen $W_{\vec{e}}$ zijn bekend, namelijk $\text{Var}(W_{\vec{e}}) = 1$.

$$= \prod_{\vec{e} \in \tilde{E}(\mathbf{j})} 1 = 1$$

Dit geeft ons de volgende uitdrukking voor het k -de moment

$$\begin{aligned} \mathbb{E} \left[\int x^k d\mu_n \right] &= \frac{1}{n^{k/2+1}} \sum_{\mathbf{i} \in S_{k, k/2+1}} \sum_{\mathbf{j} \sim \mathbf{i}} \mathbb{E} \left[\prod_{\vec{e} \in E(\mathbf{j})} W_{\vec{e}} \right] + o(1) \\ &= \frac{1}{n^{k/2+1}} \sum_{\mathbf{i} \in S_{k, k/2+1}} \sum_{\mathbf{j} \sim \mathbf{i}} 1 + o(1) \\ &= \frac{1}{n^{k/2+1}} \sum_{\mathbf{i} \in S_{k, k/2+1}} |\{\mathbf{j} : \mathbf{j} \sim \mathbf{i}\}| + o(1) \end{aligned}$$

Het aantal combinaties in een equivalentieklasse is bekend: $\frac{n!}{(n-t)!}$ of in ons geval $\frac{n!}{(n-(k/2+1))!}$

$$= \frac{1}{n^{k/2+1}} \sum_{\mathbf{i} \in S_{k, k/2+1}} \frac{n!}{(n - (k/2 + 1))!} + o(1)$$

Aangezien de breuk binnen de sommatie niet van de keuze van \mathbf{i} afhangt kunnen we schrijven

$$= \frac{1}{n^{k/2+1}} |S_{k, k/2+1}| \frac{n!}{(n - (k/2 + 1))!} + o(1)$$

Breuken bij elkaar brengen

$$= \frac{n!}{n^{k/2+1} (n - (k/2 + 1))!} |S_{k, k/2+1}| + o(1)$$

Voor $n \rightarrow \infty$ convergeert de breuk naar een dus

$$= |S_{k, k/2+1}| + o(1)$$

□

Hoofdstuk 9

Equivalentie met wandelingen op $\mathbb{Z}_{\geq 0}$

Inleiding

Dit hoofdstuk bestaat eigenlijk uit twee delen. Enerzijds het vinden van een bijectie tussen de vertegenwoordigers van equivalentieklassen in $S_{k,k/2+1}$ en random wandelingen op $\mathbb{Z}_{\geq 0}$ met k stappen en start en einde in 0.

De reden dat we dit doen is omdat het moeilijk is het aantal vertegenwoordigers van equivalentieklassen te tellen. Tellen van wandelingen op \mathbb{Z} of op een deelverzameling van \mathbb{Z} is veel eenvoudiger.

Tweede deel is dus het uitrekenen hoeveel van dit type wandelingen er nu eigenlijk zijn. Niet geheel toevallig blijken dit er $\frac{k!}{(\frac{1}{2}k+1)!(\frac{1}{2}k)!}$ te zijn.

Uitwerking

Deze wandelingen op $\mathbb{Z}_{\geq 0}$, welke wandelingen zijn het en aan welke voorwaarden moeten deze voldoen? De wandelingen waarin we geïnteresseerd zijn, zijn de zogenaamde Dyck-wandelingen van lengte k .

Definitie 9.1 (Dyck-wandelingen).

Dyck-wandelingen van lengte k zijn wandelingen met zowel start als einde in 0, die per stap ofwel één toeneemt ofwel één afneemt en die enkel de getallen op $\mathbb{Z}_{\geq 0}$ aandoet.

Herinner verder de eigenschappen van de elementen in $S_{k,k/2+1}$. Het zijn die combinaties \mathbf{i} waarvoor:

- \mathbf{i} heeft lengte k en doet $\frac{k}{2} + 1$ verschillende punten aan
- \mathbf{i} is een vertegenwoordiger van een equivalentieklasse
- $G(\mathbf{i})$ is een boom
- $P(\mathbf{i})$ is een wandeling die elke zijde in $G(\mathbf{i})$ exact 2 maal aandoet (eenmaal in beide richtingen)

Stelling 9.2 (Equivalentie met Dyck-wandelingen).

Er bestaat een bijectie tussen de elementen van $S_{k,k/2+1}$ en Dyck-wandelingen van lengte k .

Bewijs.

Geef iedere zijde in de graaf lengte 1. Definieer voor een vertegenwoordiger \mathbf{i} de wandeling dan als volgt:

X_n is de afstand (langs de graaf) van het n -de punt tot de wortel (het beginpunt i_1).

Nummering is hier een beetje onhandig, maar zie i_1 als het nulde punt van \mathbf{i} , i_2 als de eerste, i_k als de $k - 1$ -de en tenslotte i_1 weer als de k -de.

Het is duidelijk dat voor deze afbeelding geldt $X_0 = 0$ en $X_k = 0$, dat je per stap ofwel een toeneemt ofwel een afneemt en dat je enkel getallen op $\mathbb{Z}_{\geq 0}$ aandoet.

Als je op deze manier voor een combinatie \mathbf{i} een wandeling maakt, krijg je dus een unieke Dyck-wandeling.

Andersom, als je begint vanuit een Dyck-wandeling, dan komt toenemende X_n overeen met het bezoeken van een niet eerder bezocht punt. Welk nieuwe punt dit is is al uniek bepaald aangezien we naar een vertegenwoordiger kijken.

Afnemende X_n komt overeen met het terugvolgen van de unieke zijde die loopt in de richting van de wortel. Gezien de eigenschappen van de graaf is ook hiervoor maar 1 mogelijkheid.

Bij een Dyck-wandeling hoort dus ook een unieke combinatie \mathbf{i} .

We hebben dus een bijectie tussen de combinaties \mathbf{i} in $S_{k,k/2+1}$ en Dyck-wandelingen van lengte k . \square

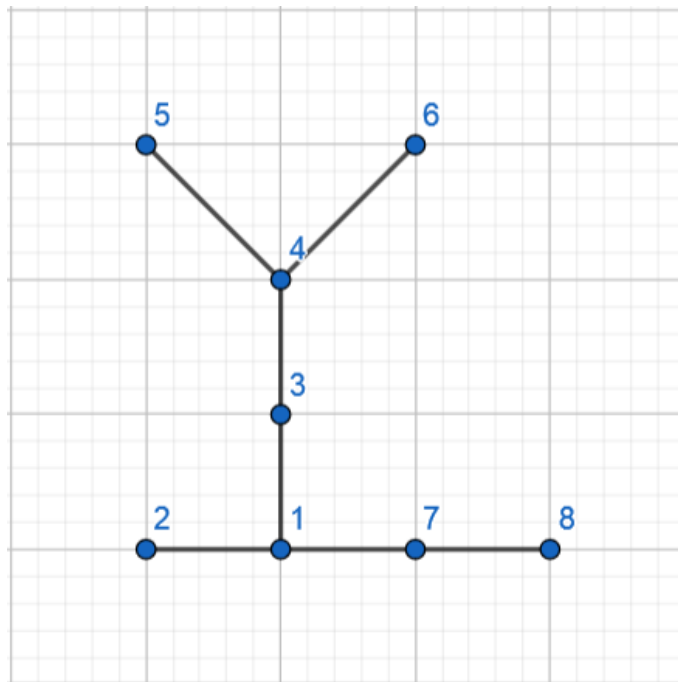
Voorbeeld 9.3.

Bekijk de combinatie $\mathbf{i} = \{1, 2, 1, 3, 4, 5, 4, 6, 4, 3, 1, 7, 8, 7\}$

Dan $E(\mathbf{i}) =$

$\{(1, 2), (2, 1), (1, 3), (3, 4), (4, 5), (5, 4), (4, 6), (6, 4), (3, 1), (1, 7), (7, 8), (8, 7), (7, 1)\}$.

De graaf van \mathbf{i} ziet er als volgt uit:



Figuur 9.1: De combinatie \mathbf{i} geïnterpreteerd als graaf

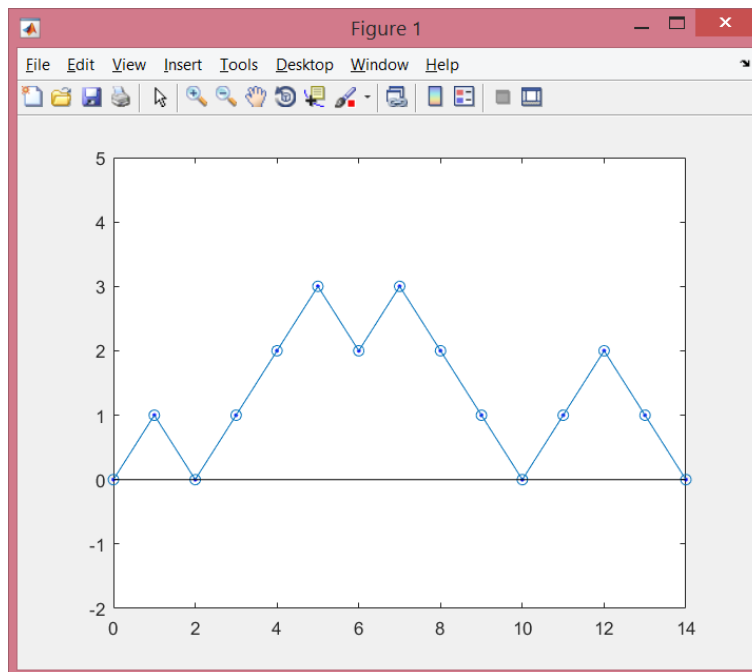
We nemen de lengte van elke zijde gelijk aan 1, dan is de afstand tussen de verschillende punten eenvoudig af te lezen.

$$d(1, 1) = 0, d(1, 2) = 1, d(1, 3) = 0, d(1, 4) = 1, d(1, 5) = 2, \\ d(1, 6) = 3, d(1, 7) = 2, d(1, 8) = 1, d(1, 9) = 0, d(1, 10) = 1, d(1, 11) = 2, d(1, 12) = 3, \\ d(1, 13) = 2, d(1, 14) = 1, d(1, 15) = 0.$$

De bijbehorende wandeling in $\mathbb{Z}_{\geq 0}$ is dus

$$0, 1, 0, 1, 2, 3, 2, 3, 2, 1, 0, 1, 2, 1, 0$$

Of de wandeling in verloop van de tijd



Figuur 9.2: De combinatie i geïnterpreteerd als wandeling

We hebben dus laten zien dat er een bijectie is tussen de wandelingen op de grafen in $S_{k,k/2+1}$ en de Dyck-wandelingen.

Dat is mooi, want het aantal Dyck-wandelingen kunnen we tellen. Daarmee weten we dan vervolgens ook $|S_{k,k/2+1}|$

Stelling 9.4 (Dyck-wandelingen tellen).

$$|S_{k,k/2+1}| = |\{\text{Dyck-wandelingen van lengte } k\}| = \frac{k!}{(\frac{1}{2}k + 1)!(\frac{1}{2}k)!}$$

Bewijs.

De eerste gelijkheid $|S_{k,k/2+1}| = |\{\text{Dyck-wandelingen van lengte } k\}|$ volgt uit de bijectie van

stelling 9.2. Vanaf hier schrijven we uit wat die Dyck-wandelingen eigenlijk zijn

$$\begin{aligned}
& |\{\text{Dyck-wandelingen van lengte } k\}| \\
&= |\{\text{wandelingen met } X_0 = X_k = 0 \text{ en } X_j \text{ voor alle } 0 \leq j \leq k\}| \\
&= |\{\text{wandelingen met } X_0 = X_k = 0\}| - \\
&\quad |\{\text{wandelingen met } X_0 = X_k = 0 \text{ en } X_j = -1 \text{ voor een } j \in \{1, \dots, k-1\}\}|
\end{aligned}$$

Hier is de eerste term gelijk aan $\binom{k}{\frac{k}{2}}$, immers $\frac{k}{2}$ sprongen omhoog en $\frac{k}{2}$ sprongen omlaag. Voor de tweede term kun je beweren dat die wandelingen in bijjectie zijn met wandelingen die beginnen in 0 en eindigen in -2. Wat je namelijk kunt doen, voor j de eerste keer dat de wandeling in -1 komt, is om vanaf punt j de wandeling te spiegelen door de lijn $y = -1$. Dit geeft een uniek pad met start in 0, einde in -2 en evenveel bezoeken aan -1. Het aantal wandelingen met $X_0 = 0$ en $X_k = -2$ is $\binom{k}{\frac{k}{2}-1}$, want $\frac{k}{2} - 1$ sprongen omhoog en $\frac{k}{2} + 1$ sprongen omlaag.

Dit geeft met wat rekenen met faculteiten

$$\begin{aligned}
& |\{\text{Dyck-wandelingen van lengte } k\}| \\
&= \binom{k}{\frac{k}{2}} - \binom{k}{\frac{k}{2}-1} \\
&= \frac{k!}{(\frac{1}{2}k)!(\frac{1}{2}k)!} - \frac{k!}{(\frac{1}{2}k+1)!(\frac{1}{2}k-1)!} \\
&= \frac{(\frac{1}{2}k+1)k!}{(\frac{1}{2}k+1)!(\frac{1}{2}k)!} - \frac{(\frac{1}{2}k)k!}{(\frac{1}{2}k+1)!(\frac{1}{2}k)!} \\
&= \frac{k!}{(\frac{1}{2}k+1)!(\frac{1}{2}k)!}
\end{aligned}$$

□

Zoals in de inleiding al gezegd is deze uitkomst niet geheel toevallig. Voor oneven waarden van k wisten we immers al dat de momenten bij de halve cirkel verdeling en de verwachting van de momenten van de ESM erg dicht bij elkaar liggen.

$$\int x^k d\mu_{sc} = 0 \tag{hoofdstuk 5}$$

$$\mathbb{E} \left[\int x^k d\mu_n \right] = o(1) \tag{hoofdstuk 8}$$

Voor de even waarden van k is dit nu ook het geval.

$$\int x^k d\mu_{sc} = \frac{k!}{(\frac{1}{2}k+1)!(\frac{1}{2}k)!} \tag{hoofdstuk 5}$$

$$\mathbb{E} \left[\int x^k d\mu_n \right] = |S_{k,k/2+1}| + o(1) = \frac{k!}{(\frac{1}{2}k+1)!(\frac{1}{2}k)!} + o(1) \tag{hoofdstuk 9}$$

Hoofdstuk 10

Borel Cantelli

Inleiding

Doel van dit hoofdstuk is het bewijzen van een interessant resultaat uit de maattheorie, namelijk het Borel Cantelli Lemma. Omdat een kansruimte een speciaal geval is van een maatruimte koppelt zich dit uiteindelijk aan het probleem waar me mee bezig zijn. Naar het einde van het hoofdstuk komt dan ook aan bod hoe we dit resultaat gaan toepassen.

Maar het bewijs gaat over algemene maatruimtes. We bekijken hiervoor twee equivalente definities voor de lim sup voor verzamelingen. Deze equivalentie wordt ook aangetoond en vervolgens gebruikt voor het bewijzen van Borel Cantelli:

Voor (X, \mathcal{S}, μ) een maatruimte en $A_n \subseteq \mathcal{S}$ een aftelbare collectie van meetbare verzamelingen geldt: Als $\sum_{n=1}^{\infty} \mu(A_n) < \infty$, dan $\mu(\limsup_{n \rightarrow \infty} A_n) = 0$

In dit hoofdstuk laten we het rekenen met momenten dus kort even los en gaan we even een totaal andere kant op.

Uitwerking

Zoals gezegd in de inleiding, om Borel Cantelli te bewijzen gebruiken we lim sup's voor verzamelingen. De eerste van twee equivalente definities is als volgt:

Definitie 10.1.

Zij $\{A_n : n \in \mathbb{N}\}$ een rij verzamelingen.

Dan definiëren we het lim sup als

$$\begin{aligned} \limsup_{n \rightarrow \infty} A_n &:= \bigcap_{i=0}^{\infty} \bigcup_{n=i}^{\infty} A_n \\ &= (A_0 \cup A_1 \cup A_2 \cup \dots) \cap (A_1 \cup A_2 \cup \dots) \cap \dots \end{aligned}$$

We gebruiken ook een tweede equivalente definitie voor de limsup. Aantonen dat deze twee definities equivalent zijn is niet heel triviaal, dus we introduceren deze equivalente formulering als stelling.

Stelling 10.2.

Zij $\{A_n : n \in \mathbb{N}\}$ een rij verzamelingen.

Dan wordt een equivalente definitie voor $\limsup A_n$ gegeven door

$$\limsup_{n \rightarrow \infty} A_n := \{x : x \in A_i \text{ voor } \infty \text{ veel } i\}$$

Bewijs.

Bekijk de volgende twee verzamelingen:

$$(1) = \bigcap_{i=0}^{\infty} \bigcup_{n=i}^{\infty} A_n$$

$$(2) = \{x : x \in A_i \text{ voor } \infty \text{ veel } i\}$$

In verzameling (1) herkennen we de definitie 10.1, in verzameling (2) de definitie gesuggereerd in stelling 10.2. Doel is te laten zien dat (1) = (2).

Bewijs dat (2) \subseteq (1):

Zij $x \in A_i$ voor oneindig veel $i \in \mathbb{N}$.

Laat nu ϕ het rijtje zijn met de nummers van deze i , in toenemende volgorde. Dat wil zeggen stel x komt voor in $A_1, A_4, A_5, A_8, A_9, A_{11}, \dots$, dan $\phi = 1, 4, 5, 8, 9, 11, \dots$

Dan voor elk willekeurig getal k , bestaat er een getal l waarvoor $\phi_l \geq k$. Dan $x \in A_{\phi_l} \subseteq \bigcap_{n=k}^{\infty} A_n$. Maar aangezien deze k willekeurig gekozen was, geldt $x \in A_{\phi_l} \subseteq \bigcap_{n=k}^{\infty} A_n$ voor iedere k . Dus ook $x \in \bigcap_{i=0}^{\infty} \bigcup_{n=i}^{\infty} A_n$.

Bewijs dat (1) \subseteq (2):

Zij $x \in \bigcap_{i=0}^{\infty} \bigcup_{n=i}^{\infty} A_n$.

Als nu x in slechts eindig veel A_i voorkomt, dan bestaat er dus een grootste i waarvoor $x \in A_i$. Noem deze grootste i i_{max} .

Dan $x \notin (A_{i_{max}+1} \cup A_{i_{max}+2} \cup \dots) = \bigcap_{n=i_{max}}^{\infty} A_n$ en daarmee $x \notin \bigcap_{i=0}^{\infty} \bigcup_{n=i}^{\infty} A_n$.

Hier komen we tegenspraak tegen. Dus x moet juist in oneindig veel A_i voorkomen. \square

Met deze equivalentie aangetoond kunnen we Borel Cantelli gaan bewijzen. In het bewijs worden verder wat eigenschappen van maten gebruikt, zoals het feit dat maten monotoon zijn, dat ze subadditief zijn, en dat ze uiteraard altijd groter gelijk nul zijn.

Stelling 10.3 (Borel Cantelli).

Zij (X, \mathcal{S}, μ) een maatruimte, $A_n \subseteq \mathcal{S}$ een aftelbare collectie van meetbare verzamelingen.

Als

$$\sum_{n=1}^{\infty} \mu(A_n) < \infty$$

Dan

$$\mu\left(\limsup_{n \rightarrow \infty} A_n\right) = 0$$

Bewijs.

Beginnen met de definitie van de lim sup.

$$\mu \left(\limsup_{n \rightarrow \infty} A_n \right) = \mu \left(\bigcap_{i=0}^{\infty} \bigcup_{j=i}^{\infty} A_j \right)$$

Er geldt dat $\bigcap_{i=0}^{\infty} \bigcup_{j=i}^{\infty} A_j$ een deelverzameling is van $\bigcup_{j=i}^{\infty} A_j$. Omdat maten monotoon zijn (d.w.z. als $B \subseteq C$, dan $\mu(B) \leq \mu(C)$), volgt nu dat

$$\leq \mu \left(\bigcup_{j=i}^{\infty} A_j \right)$$

Subadditiviteit van maten (d.w.z. voor verzamelingen B en C geldt $\mu(B \cup C) \leq \mu(B) + \mu(C)$) geeft vervolgens

$$\leq \sum_{j=i}^{\infty} \mu(A_j)$$

Dit geldt voor iedere willekeurige waarde voor i . Verder hebben we de aanname dat $\sum_{n=1}^{\infty} \mu(A_n)$ convergeert, dus de staart van de reeks convergeert naar nul: $\lim_{i \rightarrow \infty} \sum_{n=i}^{\infty} \mu(A_n) = 0$. Dit werkt omdat μ een maat is en de termen $\mu(A_j)$ daardoor nooit negatief kunnen zijn.

Korte samenvatting van wat we tot nu toe hebben:

$$(1) \mu \left(\limsup_{n \rightarrow \infty} A_n \right) \leq \sum_{j=i}^{\infty} \mu(A_j) \text{ voor iedere } i$$

$$(2) \lim_{i \rightarrow \infty} \sum_{j=i}^{\infty} \mu(A_j) = 0$$

Combineren van deze twee stukken informatie geeft dat

$$\mu \left(\limsup_{n \rightarrow \infty} A_n \right) \leq 0$$

Maten zijn natuurlijk nooit negatief, dit geeft het gewenste resultaat. □

Kansruimtes zijn speciale gevallen van maatruimtes. Dit kunnen we gebruiken om Borel Cantelli te koppelen aan ons probleem:

Gevolg 10.4.

Zij $\epsilon > 0$, als

$$\sum_{n=1}^{\infty} \mathbb{P} \left(\left| \int x^k d\mu_n - \int x^k d\mu_{sc} \right| \geq \epsilon \right) < \infty$$

Dan

$$\mathbb{P} \left(\left| \int x^k d\mu_n - \int x^k d\mu_{sc} \right| \geq \epsilon, \text{ voor oneindig veel } n \right) = 0$$

Bewijs.

Iedere kansruimte is een maatruimte, dus voor iedere aftelbare rij gebeurtenissen A_n geldt er:

Als $\sum_{n=1}^{\infty} \mathbb{P}(A_n) < \infty$, dan $\mathbb{P}(\limsup_{n \rightarrow \infty} A_n) = 0$.

Vul dan voor A_n de volgende gebeurtenis in:

$$A_n = \left| \int x^k d\mu_n - \int x^k d\mu_{sc} \right| \geq \epsilon$$

En gebruik de tweede definitie voor de \limsup 10.2. □

Wat voor impact heeft dit resultaat op het grotere geheel, op het bewijzen van stelling 1.5? Wel, zie het belangrijke gevolg 10.5.

Gevolg 10.5.

Zij $\epsilon > 0$, als

$$\sum_{n=1}^{\infty} \mathbb{P} \left(\left| \int x^k d\mu_n - \int x^k d\mu_{sc} \right| \geq \epsilon \right) < \infty$$

Dan geldt voor $n \rightarrow \infty$ dat

$$\left| \int x^k d\mu_n - \int x^k d\mu_{sc} \right| < \epsilon \text{ bijna zeker}$$

Bewijs.

Als $\sum_{n=1}^{\infty} \mathbb{P} \left(\left| \int x^k d\mu_n - \int x^k d\mu_{sc} \right| \geq \epsilon \right) < \infty$.

En daardoor $\mathbb{P} \left(\left| \int x^k d\mu_n - \int x^k d\mu_{sc} \right| \geq \epsilon, \text{ voor oneindig veel } n \right) = 0$.

Dan weten we dat er slechts eindig veel n zijn waarvoor het verschil tussen de twee integralen groter dan ϵ is: $\left| \int x^k d\mu_n - \int x^k d\mu_{sc} \right| \geq \epsilon$. Er is dus een grootste n , n_{max} , waarvoor dit geldt.

Voor $n > n_{max}$ weten we dan dat de integralen wel dicht bij elkaar liggen, oftewel dat

$\left| \int x^k d\mu_n - \int x^k d\mu_{sc} \right| < \epsilon$ bijna zeker □

In het volgende hoofdstuk tonen we aan dat de sommatie

$$\sum_{n=1}^{\infty} \mathbb{P} \left(\left| \int x^k d\mu_n - \int x^k d\mu_{sc} \right| \geq \epsilon \right)$$

inderdaad eindig is.

Hoofdstuk 11

Gevolgen van Chebyshev ongelijkheid

Inleiding

In de hoofdstukken 5 tot en met 9 zijn we er mee bezig geweest dat $\int x^k d\mu_{sc}$ en $\mathbb{E}[\int x^k d\mu_n]$ willekeurig dicht bij elkaar liggen voor grote n .

De vraag is wat we hier mee kunnen, dit is immers niet de convergentie die we uiteindelijk zoeken. We willen dat de momenten van de halve cirkel verdeling en die van de ESM convergeren, ook zonder dat er een verwachting omheen staat.

Als we wat kunnen zeggen over het verschil tussen $\int x^k d\mu_n$ en $\mathbb{E}[\int x^k d\mu_n]$, dan hebben we wel genoeg informatie om uitspraken te doen over het verschil tussen de momenten m.b.t. de halve cirkel verdeling en die m.b.t. de ESM.

Zo'n verband bestaat, de Chebyshev ongelijkheid geeft een begrenzing voor de kans dat een random variabele en zijn verwachting meer dan een ϵ uit elkaar liggen.

Doel van dit hoofdstuk is om deze Chebyshev ongelijkheid op te halen en te kijken wat de gevolgen hiervan zijn. Hierna moeten we wat trucs uithalen om een aantal varianties klein te praten. Aan het einde van dit hoofdstuk hebben we dan vervolgens genoeg informatie verzameld om Wigners halve cirkel stelling te bewijzen.

Uitwerking

Herinner Chebyshev's ongelijkheid

Stelling 11.1 (Chebyshev's ongelijkheid).

Zij Y een random variabele en $\mathbb{E}[Y^2] < \infty$. Dan

$$\mathbb{P}(|Y| \geq t) \leq \frac{1}{t^2} \mathbb{E}[Y^2]$$

voor iedere $t > 0$.

Een voor ons interessante versie van deze ongelijkheid is

Gevolg 11.2 (Chebyshev's ongelijkheid uitgedrukt in varianties).

Zij $\epsilon > 0$ willekeurig, dan geldt

$$\mathbb{P} \left(\left| \int x^k d\mu_n - \mathbb{E} \left[\int x^k d\mu_n \right] \right| \geq \epsilon \right) \leq \frac{1}{\epsilon^2} \text{Var} \left(\int x^k d\mu_n \right)$$

Bewijs.

Zeg $Y = \int x^k d\mu_n - \mathbb{E} \left[\int x^k d\mu_n \right]$, zeg $t = \epsilon$. Dit uitschrijven geeft

$$\mathbb{P} \left(\left| \int x^k d\mu_n - \mathbb{E} \left[\int x^k d\mu_n \right] \right| \geq \epsilon \right) \leq \frac{1}{\epsilon^2} \mathbb{E} \left[\left(\int x^k d\mu_n - \mathbb{E} \left[\int x^k d\mu_n \right] \right)^2 \right]$$

Nu is aan de rechterkant de definitie van de variantie van $\int x^k d\mu_n$ te herkennen. Dus

$$= \frac{1}{\epsilon^2} \text{Var} \left(\int x^k d\mu_n \right)$$

□

We gaan terug naar de sommatie die in hoofdstuk 10 zo belangrijk was:

$$\sum_{n=1}^{\infty} \mathbb{P} \left(\left| \int x^k d\mu_n - \int x^k d\mu_{sc} \right| \geq \epsilon \right)$$

Om deze eindig te praten bekijken we de n -de sommand en begrenzings daarop voor grote waarden van n .

Het is ook voldoende om naar de staart van de sommatie te kijken, als de staart van de sommatie immers naar nul convergeert, dan is de hele sommatie eindig. Immers

$$\sum_{n=1}^{\infty} \mathbb{P} \left(\left| \int x^k d\mu_n - \int x^k d\mu_{sc} \right| \geq \epsilon \right) \leq N + \sum_{n=N}^{\infty} \mathbb{P} \left(\left| \int x^k d\mu_n - \int x^k d\mu_{sc} \right| \geq \epsilon \right)$$

voor iedere waarde van $N \in \mathbb{N}$.

Lemma 11.3.

Zij $\epsilon > 0$ willekeurig, dan geldt voor grote n dat

$$\mathbb{P} \left(\left| \int x^k d\mu_n - \int x^k d\mu_{sc} \right| \geq \epsilon \right) \leq \frac{4}{\epsilon^2} \text{Var} \left(\int x^k d\mu_n \right)$$

Bewijs.

We beginnen met de driehoeksongelijkheid

$$\mathbb{P} \left(\left| \int x^k d\mu_n - \int x^k d\mu_{sc} \right| \geq \epsilon \right)$$

$$\begin{aligned}
&= \mathbb{P} \left(\left| \int x^k d\mu_n - \mathbb{E} \left[\int x^k d\mu_n \right] + \mathbb{E} \left[\int x^k d\mu_n \right] - \int x^k d\mu_{sc} \right| \geq \epsilon \right) \\
&\leq \mathbb{P} \left(\left| \int x^k d\mu_n - \mathbb{E} \left[\int x^k d\mu_n \right] \right| + \left| \mathbb{E} \left[\int x^k d\mu_n \right] - \int x^k d\mu_{sc} \right| \geq \epsilon \right)
\end{aligned}$$

Dit is ook toegestaan binnen de kans operator, immers de kans dat bijvoorbeeld $X \geq \epsilon$ is kleiner dan de kans dat $X + \delta \geq \epsilon$ voor $\delta > 0$.

Wat we nu op moeten merken is dat $|\mathbb{E} [\int x^k d\mu_n] - \int x^k d\mu_{sc}|$ klein is voor grote n , daar zijn we in hoofdstukken 6 tot en met 9 mee bezig geweest.

Vanaf zekere waarde van n geldt dat $|\mathbb{E} [\int x^k d\mu_n] - \int x^k d\mu_{sc}| \leq \frac{1}{2}\epsilon$.

Als je wilt dat $|\int x^k d\mu_n - \mathbb{E} [\int x^k d\mu_n]| + |\mathbb{E} [\int x^k d\mu_n] - \int x^k d\mu_{sc}| \geq \epsilon$ moet dus op zijn minst gelden dat $|\int x^k d\mu_n - \mathbb{E} [\int x^k d\mu_n]| \geq \frac{1}{2}\epsilon$.

Oftewel

$$\leq \mathbb{P} \left(\left| \int x^k d\mu_n - \mathbb{E} \left[\int x^k d\mu_n \right] \right| \geq \frac{1}{2}\epsilon \right)$$

Chebyshev ongelijkheid 11.2 geeft nu

$$\leq \frac{4}{\epsilon^2} \text{Var} \left(\int x^k d\mu_n \right)$$

waarbij voor de variantie de breuk $\frac{4}{\epsilon^2}$ staat in plaats van de $\frac{1}{\epsilon^2}$ die we in gevolg 11.2 hebben gezien. Dit omdat we in de kans $\frac{1}{2}\epsilon$ hebben staan in plaats van enkel ϵ . \square

Volgende opdracht is nu de varianties $\text{Var} (\int x^k d\mu_n)$ te begrenzen.

Lemma 11.4 (Spoorformule toegepast op variantie).

Voor de variantie van $\int x^k d\mu_n$ geldt

$$\text{Var} \left(\int x^k d\mu_n \right) = \frac{1}{n^{k+2}} \sum_{\mathbf{i}, \mathbf{j} \in \{1, \dots, n\}^k} C(\mathbf{i}, \mathbf{j}) - C(\mathbf{i})C(\mathbf{j})$$

waarbij

$$C(\mathbf{i}) = \mathbb{E} [W_{i_1, i_2} \dots W_{i_k, i_1}]$$

$$C(\mathbf{j}) = \mathbb{E} [W_{j_1, j_2} \dots W_{j_k, j_1}]$$

$$C(\mathbf{i}, \mathbf{j}) = \mathbb{E} [W_{i_1, i_2} \dots W_{i_k, i_1} W_{j_1, j_2} \dots W_{j_k, j_1}]$$

Bewijs.

Begin met uitrekenen van $\mathbb{E} \left[\left(\int x^k d\mu_n \right)^2 \right]$ en $\mathbb{E} \left[\int x^k d\mu_n \right]^2$.

Voor de eerste (de verwachting van het kwadraat) volgen we grotendeels het bewijs van stelling 6.6 (de spoorformule), voor de tweede (het kwadraat van de verwachting) gebruiken we eigenlijk alleen het resultaat van deze stelling.

Met behulp van de spoorformule vinden we dus eerst

$$\begin{aligned}
\mathbb{E} \left[\left(\int x^k d\mu_n \right)^2 \right] &= \mathbb{E} \left[\left(\frac{1}{n} \sum_{i=1}^n \lambda_i^k \right)^2 \right] \\
&= \mathbb{E} \left[\left(\frac{1}{n} \text{Tr} \left(\frac{1}{\sqrt{n}} W \right)^k \right)^2 \right] \\
&= \mathbb{E} \left[\frac{1}{n^2} \frac{1}{n^{k/2}} \text{Tr}(W^k) \frac{1}{n^{k/2}} \text{Tr}(W^k) \right] \\
&= \frac{1}{n^{k+2}} \mathbb{E} \left[\text{Tr}(W^k) \text{Tr}(W^k) \right] \\
&= \frac{1}{n^{k+2}} \mathbb{E} \left[\sum_{1 \leq i_1, \dots, i_k \leq n} W_{i_1, i_2} \dots W_{i_k, i_1} \sum_{1 \leq j_1, \dots, j_k \leq n} W_{j_1, j_2} \dots W_{j_k, j_1} \right] \\
&= \frac{1}{n^{k+2}} \mathbb{E} \left[\sum_{1 \leq i_1, \dots, i_k \leq n} \sum_{1 \leq j_1, \dots, j_k \leq n} W_{i_1, i_2} \dots W_{i_k, i_1} W_{j_1, j_2} \dots W_{j_k, j_1} \right] \\
&= \frac{1}{n^{k+2}} \sum_{1 \leq i_1, \dots, i_k \leq n} \sum_{1 \leq j_1, \dots, j_k \leq n} \mathbb{E} [W_{i_1, i_2} \dots W_{i_k, i_1} W_{j_1, j_2} \dots W_{j_k, j_1}] \\
&= \frac{1}{n^{k+2}} \sum_{\mathbf{i}, \mathbf{j} \in \{1, \dots, n\}^k} C(\mathbf{i}, \mathbf{j})
\end{aligned}$$

Voor meer uitleg in tekst bij de verschillende rekenstappen zie het bewijs van stelling 6.6. Daar volgen we precies dezelfde stappen.

En vervolgens vinden we ook dat

$$\begin{aligned}
\mathbb{E} \left[\int x^k d\mu_n \right]^2 &= \left(\frac{1}{n^{k/2+1}} \sum_{\mathbf{i} \in \{1, \dots, n\}^k} C(\mathbf{i}) \right)^2 \\
&= \frac{1}{n^{k+2}} \sum_{\mathbf{i}, \mathbf{j} \in \{1, \dots, n\}^k} C(\mathbf{i}) C(\mathbf{j})
\end{aligned}$$

Gebruiken van de rekenregel $\text{Var}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2$ geeft nu het gewenste resultaat.

$$\begin{aligned}
\text{Var} \left(\int x^k d\mu_n \right) &= \mathbb{E} \left[\left(\int x^k d\mu_n \right)^2 \right] - \mathbb{E} \left[\int x^k d\mu_n \right]^2 \\
&= \frac{1}{n^{k+2}} \sum_{\mathbf{i}, \mathbf{j} \in \{1, \dots, n\}^k} C(\mathbf{i}, \mathbf{j}) - \frac{1}{n^{k+2}} \sum_{\mathbf{i}, \mathbf{j} \in \{1, \dots, n\}^k} C(\mathbf{i}) C(\mathbf{j}) \\
&= \frac{1}{n^{k+2}} \sum_{\mathbf{i}, \mathbf{j} \in \{1, \dots, n\}^k} C(\mathbf{i}, \mathbf{j}) - C(\mathbf{i}) C(\mathbf{j})
\end{aligned}$$

□

Lemma 11.5.

$$\text{Var} \left(\int x^k d\mu_n \right) = O(n^{-2})$$

Bewijs.

Bekijk het uitgangspunt van lemma 11.4.

$$\text{Var} \left(\int x^k d\mu_n \right) = \frac{1}{n^{k+2}} \sum_{\mathbf{i}, \mathbf{j} \in \{1, \dots, n\}^k} C(\mathbf{i}, \mathbf{j}) - C(\mathbf{i})C(\mathbf{j})$$

waarbij

$$C(\mathbf{i}) = \mathbb{E} [W_{i_1, i_2} \dots W_{i_k, i_1}]$$

$$C(\mathbf{j}) = \mathbb{E} [W_{j_1, j_2} \dots W_{j_k, j_1}]$$

$$C(\mathbf{i}, \mathbf{j}) = \mathbb{E} [W_{i_1, i_2} \dots W_{i_k, i_1} W_{j_1, j_2} \dots W_{j_k, j_1}]$$

Uit lemma 7.5 weten we dat $C(\mathbf{i})$ en $C(\mathbf{j})$ enkel ongelijk nul kunnen zijn als in zowel de combinatie \mathbf{i} als de combinatie \mathbf{j} hooguit $\frac{k}{2} + 1$ verschillende punten bezocht kunnen worden.

Voor $C(\mathbf{i})C(\mathbf{j}) \neq 0$ moet dus gelden dat \mathbf{i} en \mathbf{j} samen hooguit $k + 2$ verschillende punten kunnen aandoen.

Interpreteer vervolgens \mathbf{i}, \mathbf{j} als combinatie van lengte $2k$ en zij $C(\mathbf{i}, \mathbf{j}) \neq 0$. Dan geldt volgens lemma 7.3 dat $\tilde{E}(\mathbf{i}, \mathbf{j}) \leq \frac{2k}{2} = k$. Omdat de graaf bij \mathbf{i}, \mathbf{j} niet per se een samenhangende graaf is, maar wel uit hooguit twee samenhangende delen bestaat geldt $V(\mathbf{i}, \mathbf{j}) \leq k + 2$.

Dus voor $C(\mathbf{i})C(\mathbf{j}) \neq 0$ moet net als voor $C(\mathbf{i})C(\mathbf{j}) \neq 0$ gelden dat \mathbf{i} en \mathbf{j} samen hooguit $k + 2$ verschillende punten kunnen aandoen.

Voor $C(\mathbf{i})C(\mathbf{j}) - C(\mathbf{i}, \mathbf{j}) \neq 0$ geldt het dan uiteraard ook.

Stel dat \mathbf{i} en \mathbf{j} geen zijdes delen, dan geldt wegens onafhankelijkheid $C(\mathbf{i}, \mathbf{j}) = C(\mathbf{i})C(\mathbf{j})$, oftewel de sommand verdwijnt. Met andere woorden de enige paren van combinaties die bijdragen aan de som delen ten minste een zijde.

Maar we hebben ook gezegd dat alle zijdes in de grafen van \mathbf{i} en \mathbf{j} tenminste twee keer doorlopen moeten worden, anders valt immers alles weg wegens de onafhankelijkheid van de $W_{\vec{e}}$'s.

De enige paren van combinaties die bijdragen aan de som delen dus ten minste 2 zijdes.

Voor $C(\mathbf{i})C(\mathbf{j}) - C(\mathbf{i}, \mathbf{j}) \neq 0$ moet nu gelden dat \mathbf{i} en \mathbf{j} samen hooguit k verschillende punten bezoeken. Deze k verschillende punten mogen dan de waarden 1 tot en met n aannemen.

Het aantal mogelijkheden dat je de 1, 2, 3, ..., k punten uit $2k$ punten kiest is eindig en hangt niet af van n . Het aantal combinaties waarvoor de sommand geen nul geeft is dus $O_k(n^k)$, waarbij de O staat voor het Landau grote O symbool en de k voor het feit dat de constante in de definitie van deze notatie af mag hangen van k .

De sommand is tevens altijd eindig, dit gezien de voorwaarden op de matrix W , elk kental

is immers begrensd. Dit maakt dat de variantie $O(n^{-2})$ is, of beter gezegd $O_k(n^{-2})$. Reden hiervoor is de breuk $\frac{1}{n^{k+2}}$.

$$\begin{aligned}\text{Var}\left(\int x^k d\mu_n\right) &= \frac{1}{n^{k+2}} \sum_{\mathbf{i}, \mathbf{j} \in \{1, \dots, n\}^k} C(\mathbf{i}, \mathbf{j}) - C(\mathbf{i})C(\mathbf{j}) \\ &= \frac{1}{n^{k+2}} O_k(n^k) \\ &= O_k(n^{-2})\end{aligned}$$

□

Met deze uitspraak over de orde van de variantie $\text{Var}\left(\int x^k d\mu_n\right) = O_k(n^{-2})$, kunnen we via eerdere ongelijkheden ook wat zeggen over de kans dat $\int x^k d\mu_n$ en $\int x^k d\mu_{sc}$ meer dan ϵ uit elkaar liggen. Ook deze blijkt voldoende klein te zijn.

Gevolg 11.6.

$$\mathbb{P}\left(\left|\int x^k d\mu_n - \int x^k d\mu_{sc}\right| \geq \epsilon\right) = O_{k,\epsilon}(n^{-2})$$

Bewijs.

Combineren van lemma 11.3 en lemma 11.5 geeft dat voor grote n

$$\begin{aligned}\mathbb{P}\left(\left|\int x^k d\mu_n - \int x^k d\mu_{sc}\right| \geq \epsilon\right) &\leq \frac{4}{\epsilon^2} \text{Var}\left(\int x^k d\mu_n\right) \\ &= \frac{4}{\epsilon^2} O_k(n^{-2}) \\ &= O_{k,\epsilon}(n^{-2})\end{aligned}$$

□

Deze is sommeerbaar voor alle k, ϵ , er geldt immers $\sum_{i=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}$ (Euler). Omdat $O_{k,\epsilon}(n^{-2})$ dus sommeerbaar is voor alle k en ϵ hebben we aangetoond dat de sommatie uit gevolg 10.5 begrensd is.

Hoofdstuk 12

Bewijs van Wigners halve cirkel stelling

We zijn nu aangekomen op het punt dat alle informatie die we nodig hebben om de stelling te bewijzen bekend is. Rest alleen om deze informatie met elkaar te combineren tot het gewenste bewijs.

Voor de laatste keer in dit verslag herinner stelling 1.5.

Stelling 1.5 (Wigner halve cirkel stelling).

De rij $\{\mu_n\}$ van empirische spectraal maten van een rij Wigner matrices $\{W_n\}$ convergeert zwak bijna zeker naar de halve cirkel verdeling op $[-2, 2]$.

Dit betekent dat voor elke begrensde continue functie f geldt dat

$$\int_{\mathbb{R}} f d\mu_n \rightarrow \int_{\mathbb{R}} f d\mu_{sc} = \int_{-2}^2 f(x) \frac{1}{2\pi} \sqrt{4-x^2} dx \text{ bijna zeker}$$

Bewijs.

Het doel is dus te bewijzen dat voor continue en begrensde functies f geldt

$$\left| \int f d\mu_n - \int f d\mu_{sc} \right| \rightarrow 0 \text{ voor } n \rightarrow \infty$$

In hoofdstuk 4 hebben we geleerd dat het voldoet te laten zien dat

$$\left| \int x^k d\mu_n - \int x^k d\mu_{sc} \right| \rightarrow 0 \text{ voor } n \rightarrow \infty$$

Voor de halve cirkel verdeling hebben we in hoofdstuk 5 de k -de momenten uitgerekend.

$$\int x^k d\mu_{sc} = \begin{cases} 0 & k \text{ oneven} \\ \frac{k!}{(\frac{1}{2}k+1)!(\frac{1}{2}k)!} & k \text{ even} \end{cases}$$

Tevens geldt wegens hoofdstuk 6 tot en met 9 dat

$$\mathbb{E} \left[\int x^k d\mu_n \right] = \begin{cases} o(n) & k \text{ oneven} \\ \frac{k!}{(\frac{1}{2}k+1)!(\frac{1}{2}k)!} + o(n) & k \text{ even} \end{cases} \quad (n \rightarrow \infty)$$

Volgens Borel Cantelli in hoofdstuk 10 voldoet het te bewijzen dat voor willekeurige $\epsilon > 0$ en voor alle k de volgende som eindig is

$$\sum_{n=1}^{\infty} \mathbb{P} \left(\left| \int x^k d\mu_n - \int x^k d\mu_{sc} \right| \geq \epsilon \right) < \infty$$

Maar omdat $\int x^k d\mu_{sc}$ en $\mathbb{E} [\int x^k d\mu_n]$ zo dicht bij elkaar liggen hebben we in hoofdstuk 11 aangetoond dat

$$\mathbb{P} \left(\left| \int x^k d\mu_n - \int x^k d\mu_{sc} \right| \geq \epsilon \right) = O_{k,\epsilon}(n^{-2}) \quad (n \rightarrow \infty)$$

En deze is sommerbaar voor alle k en ϵ . Dus $\sum_{n=1}^{\infty} \mathbb{P} (|\int x^k d\mu_n - \int x^k d\mu_{sc}| \geq \epsilon)$ is eindig. \square

Conclusie

We kiezen ervoor om Wigner's halve cirkel stelling te bewijzen aan de hand van de momentenmethode. Vrij vroeg wordt er dan ook al bewezen dat het voldoende is te laten zien dat de momenten van de twee verdelingen naar elkaar convergeren.

De reden dat dit bewijs werkt is omdat het lukt om het k -de moment van de halve cirkel verdeling en de verwachting van het k -de moment bij de ESM zo dicht bij elkaar te praten. Op het moment dat je dat voor elkaar hebt, dan is de rest van het bewijs wel in te vullen.

Cruciaal hiervoor is om in de spoorformule op verschillende manieren naar het matrixproduct te kijken. Juist door een groep matrixelementen aan een graaf te koppelen kan je voor grote delen van de sommatie laten zien dat het gelijk aan nul is of dat het naar nul convergeert.

Vervolgens koppelen van de overgebleven grafen met Dyck-wandelingen zorgt ervoor dat je kan tellen hoeveel zinvolle grafen en daarmee matrixproducten daadwerkelijk moeten sommeren.

In het slotdeel van het bewijs gebruiken we het Borel Cantelli lemma om te laten zien dat het nu voldoet om een andere sommatie eindig te praten.

Met de Chebyshev ongelijkheid en andere afschattingen krijgen we dit uiteindelijk voor elkaar.

Bijlage A - Uitbreidingen van de stelling

Hermitische matrices

Een mogelijke veralgemenisering van de stelling is de toevoeging dat de elementen in de matrix niet op de diagonaal in principe ook complex mogen zijn. Dan moet de matrix wel Hermitisch zijn in plaats van symmetrisch (een matrix A noemen we Hermitisch als $\overline{A}^T = A$, waarbij het streepje boven de A de complex geconjugeerde betekent).

Dit geeft dus de volgende voorwaarden voor een $n \times n$ Wigner matrix W .

- $W_{i,j}$, $1 \leq i < j$ zijn onafhankelijke random variabelen op \mathbb{C}
- $W_{i,i}$, $1 \leq i \leq n$ zijn onafhankelijke random variabelen op \mathbb{R}
- Voor elke $j < i$ laat $W_{i,j} = \overline{W_{j,i}}$
- Voor elke i, j is $W_{i,j}$ begrensd
- Voor elke i, j geldt $\mathbb{E}[W_{i,j}] = 0$
- Voor elke $i \neq j$ geldt $\text{Var}(W_{i,j}) = \mathbb{E}[|W_{i,j}|^2] = 1$

Om deze verandering te kunnen verantwoorden is het nodig te bedenken waar in het bewijs we de symmetrie-eigenschap van W hebben gebruikt.

Een van deze momenten is wanneer we gebruiken dat alle $W_{i,j}$ onafhankelijk van elkaar zijn, behalve dan de random variabele aan de andere kant van de diagonaal $W_{j,i}$. Dit is duidelijk nog steeds het geval.

Het andere moment is wanneer er gebruikt wordt dat de eigenwaarden van een reële symmetrische matrix reëel zijn. Voor Hermitische matrices is dit ook het geval, immers

Zij λ een willekeurige eigenwaarde van een Hermitische matrix A en zij \mathbf{x} een eigenvector bij deze eigenwaarde λ . Dan geldt

$$A\mathbf{x} = \lambda\mathbf{x}$$

Vanaf de linkerkant vermenigvuldigen met $\overline{\mathbf{x}}^T$ geeft

$$\overline{\mathbf{x}}^T A\mathbf{x} = \overline{\mathbf{x}}^T \lambda\mathbf{x}$$

$$\overline{\mathbf{x}}^T A\mathbf{x} = \lambda\overline{\mathbf{x}}^T \mathbf{x}$$

$$\bar{\mathbf{x}}^T A \mathbf{x} = \lambda \|\mathbf{x}\|^2$$

Nu beide kanten conjugereren en transponeren

$$\overline{(\bar{\mathbf{x}}^T A \mathbf{x})}^T = \overline{(\lambda \|\mathbf{x}\|)^T}$$

$$(\mathbf{x}^T \bar{A} \bar{\mathbf{x}})^T = \bar{\lambda} \|\mathbf{x}\|^2$$

$$\bar{\mathbf{x}}^T \bar{A}^T \mathbf{x} = \bar{\lambda} \|\mathbf{x}\|^2 \tag{*}$$

Aangezien A Hermitisch is geldt $\bar{A}^T = A$.

$$\bar{\mathbf{x}}^T \bar{A}^T \mathbf{x} = \bar{\mathbf{x}}^T A \mathbf{x}$$

$$\bar{\mathbf{x}}^T \bar{A}^T \mathbf{x} = \bar{\mathbf{x}}^T \lambda \mathbf{x}$$

$$\bar{\mathbf{x}}^T \bar{A}^T \mathbf{x} = \lambda \|\mathbf{x}\|^2 \tag{**}$$

Uit vergelijking (*) en (**) volgt nu $\bar{\lambda} \|\mathbf{x}\|^2 = \lambda \|\mathbf{x}\|^2$.

\mathbf{x} is hierin een eigenvector, dus $\mathbf{x} \neq \mathbf{0}$ en $\|\mathbf{x}\| \neq 0$.

Dus $\bar{\lambda} = \lambda$, oftewel λ is reëel. Dus Hermitische matrices hebben enkel reële eigenwaarden.

Bijlage B - MatLab code

EigenvaluesRandomMatrix.m

```
clear all;
close all;

% Vul in de stapgrootte van de matrix.
N = 1000;

% Genereer een willekeurige NxN-matrix A.
for i=1:N;
    % Kentallen op de diagonaal zijn normaal verdeeld met verwachting 0 en
    % variantie 1.
    A(i,i)=normrnd(0,1);
    for j=i+1:N;
        % Ook alles boven de diagonaal is normaal verdeeld met verwachting
        % 0 en variantie 1.
        A(i,j)=normrnd(0,1);
        % Maak de matrix symmetrisch.
        A(j,i)=A(i,j);
    end
end

% We zijn geïnteresseerd niet in de eigenwaarden van de matrix A, maar in
% de eigenwaarden van de matrix A/sqrt(N).
% Herdefinieer dus de matrix A.
A=A/sqrt(N)
% Reken de eigenwaarden uit en sla deze op
e=eig(A)

figure;
hold on
% Maak een histogram voor de gevonden eigenwaarden.
histwidth=0.2;
x=[-4:0.2:4];
histogram(e,x)

% En plot in dezelfde afbeelding een halve cirkel.
% De extra scalar hier wordt gebruikt opdat het oppervlakte onder de halve
% cirkel gelijk is aan de oppervlakte van het histogram.
```

```
scalar = histwidth*N;  
x=[-2:0.01:2];  
plot(x,scalar*1/(2*pi)*sqrt(4-x.^2),'r','Linewidth',2)  
hold off
```

Lijst terugkerende symbolen

In chronologische volgorde:

W	Wigner matrix
W_n	Wigner matrix met dimensies $n \times n$
$W_{i,j}$	het element op positie i, j in een Wigner matrix
$W_{\vec{e}}$	\vec{e} stelt een zijde i, j voor, $W_{\vec{e}}$ is dan hetzelfde als $W_{i,j}$
T	begrenzing op de random variabelen $W_{i,j}$
μ	een willekeurige maat
μ_n	empirische spectraal maat (ESM)
μ_{sc}	halve cirkel verdeling
δ	Dirac-maat
δ_{λ_i}	Dirac-maat t.o.v. λ_i
\mathcal{S}	een willekeurige halfring
a_k	afkorting voor $\int_0^{\frac{1}{2}\pi} \sin^k(\theta) d\theta$
A	een willekeurige matrix, verzameling of gebeurtenis
$\text{Tr}(A)$	spoor van een matrix A
$C(\mathbf{i})$	de verwachting van het matrixproduct behorend bij \mathbf{i} :
$V(\mathbf{i})$	de punten in de graaf behorend bij \mathbf{i}
$E(\mathbf{i})$	de gerichte zijden in de graaf behorend bij \mathbf{i}
$\tilde{E}(\mathbf{i})$	de ongerichte zijden in de graaf behorend bij \mathbf{i}
$G(\mathbf{i})$	$(V(\mathbf{i}), \tilde{E}(\mathbf{i}))$ het skelet behorend bij \mathbf{i}
$P(\mathbf{i})$	de wandeling over de graaf $G(\mathbf{i})$
\vec{e}	elementen van $E(\mathbf{i})$
\overleftarrow{e}	omgekeerde van \vec{e}
$R_{k,t}$	verzameling van vertegenwoordigers van equivalentieclasses
$S_{k,t}$	deelverzameling van $R_{k,t}$, voldoet aan wat extra voorwaarden
X_n	locatie op tijdstip n in een wandeling op $\mathbb{Z}_{\geq 0}$

Bibliografie

- [1] Nicholas Cook, *Two Proofs of Wigner's semicircular law*. Lecture notes, 2012
- [2] Slava Kargin, Elena Yudovina, *Random Matrix Theory, Lecture 1*. Lecture notes, 2013
- [3] Charalambos D. Aliprantis, Owen Burkinshaw, *Principles of Real Analysis*. Academic Press, 1998
- [4] Avner Friedmann *Foundations of Modern Analysis*. Holt, Rinehart, and Winston, 1982
- [5] René L. Schilling *Measures, Integrals and Martingales*. Cambridge University Press, 2005