

Abschlussarbeit Fallstudien der math. Modellbildung

Manuela Lambacher, Dominik Otto, Andreas Wiedemann

26. März 2014

1 Das Marchenko-Pastur-Gesetz

Das Marchenko-Pastur-Gesetz

Sei Y_N eine $N \times M(N)$ -Matrix mit unabhängigen zentrierten Einträgen mit Varianz 1,

$$\sup_{j,k,N} \mathbb{E}[|Y_N(j,k)|^q] = C_q < \infty \quad \forall q \in \mathbb{N}$$

und $M(N) \in \mathbb{N}$ so, dass

$$\lim_{N \rightarrow \infty} \frac{M(N)}{N} = \alpha \in [1, \infty).$$

Sei weiterhin die Wishart-Matrix gegeben als

$$W_N = \frac{1}{N} Y_N Y_N^T,$$

und habe die empirische Eigenwertverteilung

$$L_N = \frac{1}{N} \sum_{j=1}^N \delta_{\lambda_j}$$

und das Zustandsdichtemaß $\overline{L}_N = \mathbb{E}[L_N]$. Dann gilt die Konvergenz

$$\overline{L}_N \xrightarrow{w} f_\alpha(x) dx \quad (N \rightarrow \infty)$$

im Raum der Wahrscheinlichkeitsmaße auf \mathbb{R} , wobei

$$f_\alpha(x) = \frac{1}{2\pi x} \sqrt{(x - (1 - \sqrt{\alpha})_+^2)((1 + \sqrt{\alpha})_+^2 - x)}$$

Beweis des Marcenko-Pastur Gesetzes

Die Konvergenz soll im Folgenden mittels der Momentenmethode gezeigt werden. Als ersten Schritt bringt man $N^{l+1}\langle \overline{L_N}, x^l \rangle$ in eine Form, die eine weitergehende Untersuchung ermöglicht:

$$\begin{aligned}
 N^{l+1}\langle \overline{L_N}, x^l \rangle &= N^{l+1} \cdot \int x^l \overline{L_N}(dx) = N^{l+1} \cdot \frac{1}{N} \cdot \mathbb{E}[\text{tr}(W_N^l)] = N^l \sum_{j_1, \dots, j_l=1}^N \mathbb{E} \left[\prod_{p=1}^l W_{j_p, j_{p+1}} \right] \\
 &= N^l \sum_{j_1, \dots, j_l=1}^N \mathbb{E} \left[\prod_{p=1}^l \frac{1}{N} \sum_{k=1}^{M(N)} Y_N(j_p, k) \cdot Y_N(j_{p+1}, k) \right] \\
 &= \sum_{j_1, \dots, j_l=1}^N \mathbb{E} \left[\left(\sum_{k=1}^{M(N)} Y_N(j_1, k) \cdot Y_N(j_2, k) \right) \cdot \left(\prod_{p=2}^l \sum_{k=1}^{M(N)} Y_N(j_p, k) \cdot Y_N(j_{p+1}, k) \right) \right] \\
 &= \sum_{j_1, \dots, j_l=1}^N \mathbb{E} \left[\prod_{p=2}^l \sum_{k_1, k_2=1}^{M(N)} Y_N(j_1, k_1) \cdot Y_N(j_2, k_1) \cdot Y_N(j_p, k_2) \cdot Y_N(j_{p+1}, k_2) \right] \\
 &= (\dots) \text{ (es wird analog zum vorherigen Schritt für alle } p \text{ verfahren)} \\
 &= \sum_{j_1, \dots, j_l=1}^N \sum_{k_1, \dots, k_l=1}^{M(N)} \mathbb{E}[Y_N(j_1, k_1) Y_N(j_2, k_1) Y_N(j_2, k_2) Y_N(j_3, k_2) \dots Y_N(j_l, k_l) Y_N(j_1, k_l)] \\
 &= \sum_{s_1, s_2=1}^l \sum_{\substack{J: v(J)=s_1 \\ K: v(K)=s_2}} \mathbb{E}[Y_N(J, K)]
 \end{aligned} \tag{1.1}$$

wobei

$$\begin{aligned}
 J &= (j_1, j_2, j_2, \dots, j_l, j_l, j_1), K = (k_1, k_1, \dots, k_l, k_l), \\
 v : \mathbb{N}^{2l} &\rightarrow \mathbb{N}, \quad v(X) := \text{Anzahl der verschiedenen Indizes in } X
 \end{aligned}$$

Die einzelnen Summanden können also als Eulergraphen auf $s_1 + s_2$ verschiedenen Knoten und $2l$ Kanten zwischen der Menge der k -Knoten und der j -Knoten für alle Paare $(j_i, k_i), (j_{i+1}, k_i)$ interpretiert werden.

Setze $s = s_1 + s_2$, dann ergeben sich folgende drei Fälle:

- $s < l + 1$

$$\begin{aligned}
 \mathbb{E}[Y_N(J, k)] &\leq \prod_{n=1}^l \left(\sup_{j, k, N} \mathbb{E} \left[|Y_N(j, k)|^l \right] \right)^{\frac{1}{l}} \\
 &= \prod_{n=1}^l C_l^{\frac{1}{l}} = C_l
 \end{aligned} \tag{1.2}$$

Außerdem gilt:

$$\#\{J : v(J) = s_1\} \leq \binom{N}{s_1} s_1^l \leq N^{s_1} s_1^l \quad (1.3)$$

$$\#\{K : v(K) = s_2\} \leq \binom{M(N)}{s_2} s_2^l \leq M(N)^{s_2} s_2^l \quad (1.4)$$

Somit ergibt sich aus (1.2) – (1.4):

$$\frac{1}{N^{l+1}} \sum_{\substack{J:v(J)=s_1 \\ K:v(K)=s_2}} \mathbb{E}[Y_N(J, K)] < C_l(l+1)^l \frac{N^{s_1} M(N)^{s_2}}{N^{l+1}} \xrightarrow{N \rightarrow \infty} 0 \quad (1.5)$$

- $s > l + 1$

Nach Lemma aus der Vorlesung existiert in diesem Fall eine einfache, echte Kante. Da die Matrixeinträge unabhängig sind, kann der Erwartungswert für diesen Matrixeintrag ausgeklammert werden. Alle Einträge sind jedoch zentriert, und damit folgt direkt $\mathbb{E}[Y_N(J, K)] = 0$.

- $s = l + 1$

Die verbliebenen Graphen auf $l + 1$ verschiedenen Knoten haben die Struktur von Doppelbäumen und sollen im Weiteren genauer untersucht werden.

Zusammenfassung: Zu $\lim_{N \rightarrow \infty} \langle \overline{L_N}, x^l \rangle$ tragen nur die Graphen auf $l + 1$ verschiedenen Knoten bei:

$$\begin{aligned} \beta_l &:= \lim_{N \rightarrow \infty} \langle \overline{L_N}, x^l \rangle = \lim_{N \rightarrow \infty} \frac{1}{N^{l+1}} \sum_{s_1, s_2=1}^l \sum_{\substack{J:v(J)=s_1 \\ K:v(K)=s_2}} \mathbb{E}[Y_N(J, K)] \\ &= \lim_{N \rightarrow \infty} \frac{1}{N^{l+1}} \sum_{J, K: v(J)+v(K)=l+1} \mathbb{E}[Y_N(J, K)] \end{aligned} \quad (1.6)$$

Berechnung der β_l mittels Catalanpfade

Die Doppelbaumstruktur dieser Graphen lässt sich wie folgt nutzen:

Wähle für einen Doppelbaum r Knoten aus den k -Knoten und $l + 1 - r$ Knoten aus den j -Knoten. Dann gilt:

$$\begin{aligned} \sum_{J, K: v(J)+v(K)=l+1} \mathbb{E}[Y_N(J, K)] &= \sum_{r=1}^l \binom{N}{l+1-r} (l+1-r)! \binom{M(N)}{r} r! \\ &\quad \cdot \#\{\text{Doppelbäume mit } l+1-r \text{ } j\text{-Knoten und } r \text{ } k\text{-Knoten}\} \end{aligned} \quad (1.7)$$

Ein Doppelbaum mit r k -Knoten und $(l + 1 - r)$ j -Knoten kann nun in folgender Weise als Catalan-Pfad der Länge l interpretiert werden:

Wähle als Wurzel des Baumes einen j -Knoten und gliedere den Baum in Ebenen, wobei die Wurzel in der 0.Ebene liegt. Die k -Knoten liegen also in ungeraden Ebenen, die j -Knoten in geraden Ebenen. Verweise jede Kante mit einer Richtung, sodass bei jeder Doppelkante eine Kante von dem Knoten wegführt und eine zu ihm hinführt. Durchlaufe den Baum mit Hilfe der Tiefensuche und nummeriere die Kanten in der Reihenfolge, wie sie hierbei durchlaufen werden. Konstruiere den Catalan-Pfad wie folgt:

- Wird ein Knoten zum ersten Mal erreicht: $(+1)$
- Kehrt man zu einem bereits besuchten Knoten zurück (-1)

Da der Baum durch Tiefensuche durchlaufen wird, gibt die Entfernung eines Knotens zur Wurzel die Anzahl der nötigen Schritte wieder, die benötigt werden, um ihn zum ersten Mal zu erreichen. Da die j -Knoten in den geraden Ebenen und die k -Knoten in den ungeraden Ebenen liegen, ergibt sich, dass sich die j -Knoten stets auf den geraden Niveaus und die k -Knoten stets auf den ungeraden Niveaus des Catalan-Pfades befinden.

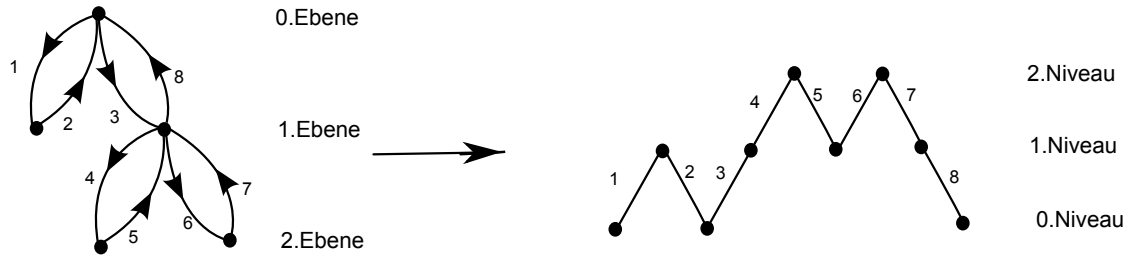


Abbildung 1.1: Doppelbaum auf 5 Knoten und dem zugehörigen Catalan-Pfad

Die Abbildung von den Doppelbäumen auf die Catalanpfade ist eine Bijektion:

- Wohldefiniertheit: Da der Graph eulersch ist, ist $\sum_{i=1}^{2l} s_i = 0$; Der Graph ist immer auf oder überhalb des 0.Niveaus, ansonsten würde es ein m geben, sodass $\sum_{i=1}^{2m-1} s_i = -1$, $\sum_{i=1}^{2m} s_i = 0$, $s_{2m-1} = -1$. Also könnte ein Doppelbaum mit Knoten $1, 2, \dots, 2m$ konstruiert werden, und da $s_{2m-1} = -1$ gilt, würde eine Kante von diesem Knoten zurück zu einem der ersten $2m$ Knoten gehen, was dem Aufbau eines Doppelbaums widersprechen würde. Insgesamt wurde also tatsächlich ein Catalanpfad konstruiert.
- Surjektivität: Starte bei der Wurzel. Füge für jede Aufstiegs-kante einen Knoten in die nächst höhere Ebene und eine Kante als Verbindung hinzu. Füge für jede Abstiegs-kante eine Kante in die nächst niedrigere Ebene hinzu, bis die Doppelbaumstruktur vollständig ist. Damit ist jeder Catalanpfad Urbild eines Doppelbaums.
- Injektivität: kann analog zur Übung durch rekursives Entfernen von identischen Kanten gezeigt werden.

1 Das Marchenko-Pastur-Gesetz

Damit ist die Anzahl der Doppelbäume in (1.7) gleich der Anzahl der dazugehörigen Catalanpfade.

Die betrachteten Doppelbäume haben nach (1.7) ein kombinatorisches Gewicht von

$$\binom{N}{l+1-r} (l+1-r)! \binom{M(N)}{r} r! \quad (1.8)$$

Für N hinreichend groß ist dies genähert $N^{l+1-r} M(N)^r$. Damit gilt:

$$\frac{1}{N^{l+1}} N^{l+1-r} M(N)^r = \left(\frac{M(N)}{N} \right)^r \rightarrow \alpha^r \quad (1.9)$$

und schließlich folgt mit (1.6):

$$\begin{aligned} \beta_l &:= \lim_{N \rightarrow \infty} \langle \overline{L_N}, x^l \rangle \\ &= \lim_{N \rightarrow \infty} \sum_{J, K: v(J)+v(K)=l+1} \frac{1}{N^{l+1}} \mathbb{E}[Y_N(J, K)] \\ &= \lim_{N \rightarrow \infty} \sum_{r=1}^l \frac{1}{N^{l+1}} \binom{N}{l+1-r} (l+1-r)! \binom{M(N)}{r} r! \cdot \#\{p_r\} \\ &= \sum_{r=1}^l \sum_{p_r \in C_l} \alpha^r = \sum_{p \in C_l} \alpha^r \end{aligned}$$

wobei $r = \#\{\text{Abstiege von ungeraden auf gerade Niveaus in } C_l\}$, und p_r Catalan-Pfad mit r entsprechenden Abstiegen.

!!! Explizite Formel für die Anzahl der p_r gefunden:

$$\binom{l-1}{r} \binom{l}{r} \frac{1}{r+1}$$

Wer Spaß dran hat, kann sie beweisen. Ansonsten lassen wir sie raus, wir brauchen sie ja nicht.

Relation für β_l

Zu zeigen ist: Mit $\beta_0 := 1, \gamma_0 := 1$ und

$$\gamma_l := \sum_{p \in C_l} \alpha^{l-r} \quad (l \geq 1)$$

gelten die Relationen

$$\beta_l = \alpha \gamma_l = \alpha \sum_{r=0}^{l-1} \beta_r \gamma_{l-1-r} \quad \forall l \geq 1 \quad (1.10)$$

($l-r$ ist dabei die Anzahl der Abstiegsanten von geraden auf ungerade Niveaus im Catalan-Pfad.)

- Betrachte die Position im Pfad, an der zum ersten Mal die 0 erreicht wird. Diese ist immer gerade, da die 0 genau dann erreicht wird, wenn man im zugehörigen Doppelbaum zur Wurzel zurückkehrt. Sei also $2j$, $j \in \{1, \dots, l\}$, diese Position. Teile nun den Pfad in einen vorderen Teil P_1 der Länge $2j$ und einen hinteren Teil P_2 der Länge $2l - 2j$. In dem erzeugenden Baum entspricht P_1 dem äußersten Teilbaum, P_2 dem restlichen Baum. P_2 ist also ein beliebiger Catalan-Pfad. P_1 hat die besondere Struktur, dass die 0 erst an letzter Position erreicht wird, die erste und letzte Kante sind somit als Aufstiegs- bzw. Abstiegs-kante festgelegt und können für die kombinatorische Analyse vernachlässigt werden. Löscht man diese beiden Kanten und subtrahiert 1 von allen Elementen von P_1 , erhält man einen neuen Pfad \tilde{P}_1 , der ein beliebiger Catalan-Pfad der Länge $2j - 2$ ist. Dies folgt direkt aus der Struktur des Doppelbaumes: \tilde{P}_1 entspricht dem äußersten Teilbaum ohne die ursprüngliche Wurzel, wodurch man einen neuen, beliebigen Doppelbaum erhält. Die geraden/ungeraden Niveaus im ursprünglichen Pfad sind in \tilde{P}_1 ungerade/gerade (Durch die Verschiebung um 1 nach unten). r und $l - r$ haben sich also genau vertauscht. Damit ergibt sich aus den Definition für β_l und γ_l die Formel:

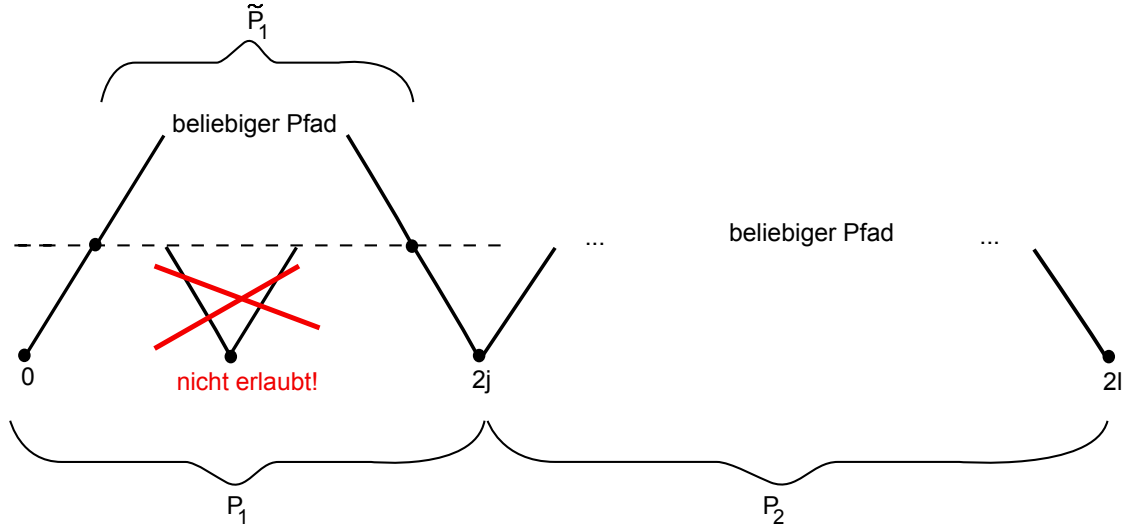


Abbildung 1.2: Schematische Darstellung der Rekursion

$$\begin{aligned}
 \beta_l &= \alpha \sum_{j=1}^{l-1} \gamma_{j-1} \beta_{l-j} = \alpha \sum_{j=0}^l \gamma_j \beta_{l-j-1} \\
 &= \alpha \sum_{k=0}^l \beta_k \gamma_{l-1-k}
 \end{aligned} \tag{1.11}$$

Das zusätzliche α wird in der Formel durch das Löschen der gewichteten Abwärts-kante in P_1 hervorgerufen.

- γ wird analog dargestellt: Im Gegensatz zu vorher tragen nun die Abwärtskanten von gerade zu ungerade zum Gewicht von α bei. Sei $2j$, $j \in \{1, \dots, l\}$, die Position, an der zum ersten Mal die 0 erreicht wird.

Teile wiederum den Pfad in einen vorderen Teil P_1 der Länge $2j$ und einen hinteren beliebigen Catalanpfad P_2 der Länge $2l - 2j$. In P_1 sind wieder die erste und letzte Kante festgelegt. Beim Löschen dieser Kante werden die Ebenen vertauscht.

Anders als bei den β_l wurde durch das Löschen der Abwärtskante allerdings kein α verloren, da nun lediglich die Abwärtskanten von gerade zu ungerade das Gewicht α haben.

Damit gilt:

$$\begin{aligned}\gamma_l &= \sum_{k=0}^l \beta_k \gamma_{l-1-k} \\ &= \frac{\beta_l}{\alpha}\end{aligned}\tag{1.12}$$

Verallgemeinerten Momente von f_α

Man betrachtet

$$xf_\alpha(x) = \frac{1}{2\pi} \sqrt{(x - (1 - \sqrt{\alpha})_+^2)((1 + \sqrt{\alpha})_+^2 - x)}$$

und vergleicht es mit dem aus der Vorlesung bekanntem Halbkreisgesetz

$$\sigma(x) = \frac{1}{2\pi} \sqrt{(4 - x^2)_+}$$

Dabei wird ersichtlich, dass xf_α als ein modifizierter Halbkreis beschrieben werden kann: Es ist null außerhalb seiner Schnittpunkte mit der x-Achse, $xf_\alpha(1 + \alpha + x) = xf_\alpha(1 + \alpha - x)$, d.h. symmetrisch um seinen "Mittelpunkt" $1 + \alpha$. Man kann den Graph der Funktion somit als einen um $\alpha + 1$ verschobenen und um einen konstanten Faktor nach oben verzerrten Halbkreis beschreiben. Auch in der nun folgenden Rechnung wird durch Substitution wieder das Halbkreisgesetz erreicht.

Anmerkung: Berechnung liefert $\int_{\mathbb{R}} f_\alpha(x) dx = 1$, f_α ist ein Wahrscheinlichkeitsmaß.

Beweis der Formel für Q_n

$$\begin{aligned}Q_n &:= \alpha^{-1-n/2} \int_{\mathbb{R}} f_\alpha(x) x(x - \alpha - 1)^n dx \\ &= \frac{1}{2\pi} \alpha^{-1-n/2} \int_{(1-\sqrt{\alpha})^2}^{(1+\sqrt{\alpha})^2} \sqrt{(x - (1 - \sqrt{\alpha})^2)((1 + \sqrt{\alpha})^2 - x)} (x - \alpha - 1)^n dx \\ &= \frac{1}{2\pi} \alpha^{-1-n/2} \int_{(1-\sqrt{\alpha})^2}^{(1+\sqrt{\alpha})^2} \sqrt{-\alpha^2 + 2\alpha x + 2\alpha - x^2 + 2x - 1} (x - \alpha - 1)^n dx \\ &\stackrel{x=y+\alpha+1}{=} \frac{1}{2\pi} \alpha^{-1-n/2} \int_{-2\sqrt{\alpha}}^{2\sqrt{\alpha}} \sqrt{4\alpha - y^2} y^n dy\end{aligned}$$

$$y=2\sqrt{\alpha}z \quad \frac{1}{2\pi}\alpha^{-1-n/2} \cdot 2^n \alpha^{n/2} \cdot 4\sqrt{\alpha} \int_{-1}^1 \sqrt{1-z^2} z^n dz = \frac{2}{\pi} \cdot 2^n \int_{-1}^1 \sqrt{1-z^2} z^n dz$$

$$\stackrel{\text{Übung 1}}{=} \sigma(z^n) = \begin{cases} 0, & n \text{ ungerade} \\ C_{\frac{n}{2}}, & n \text{ gerade} \end{cases}$$

Also

$$Q_n = \begin{cases} 0, & n \text{ ungerade} \\ C_{\frac{n}{2}}, & n \text{ gerade} \end{cases} \quad (1.13)$$

Nachdem $\{\alpha^{-1-n/2}x(x-\alpha-1)^n | n \in \mathbb{N}_0\} \cup \{1\}$ ebenso eine Basis aller Polynome darstellt wie $\{x^l | l \in \mathbb{N}_0\}$, ist f_α durch die Q_n und $\int_{\mathbb{R}} f_\alpha(x) dx$ genauso bestimmt wie durch die normalen Momente.

Theorem 1.0.1. Sei $\mu \in \mathcal{M}(\mathbb{R})$ mit $m_l(\mu) = \int x^l \mu(dx) < \infty \forall l \in \mathbb{N}_0$. Dann ist μ das einzige Wahrscheinlichkeitsmaß mit diesen Momenten, wenn die Potenzreihe

$$\sum_{l=0}^{\infty} m_l(f_\alpha) \frac{z^l}{l!}$$

einen positiven Konvergenzradius besitzt.

Da die Momente m_l endliche Linearkombinationen der verallgemeinerten Momente sind, wird das Theorem jetzt auf die Q_n angewandt.

Also ist zu zeigen: $Q_n < \infty$ (folgt direkt aus (1.13)) und $\sum_{n=0}^{\infty} Q_n \frac{z^n}{n!}$ besitzt positiven Konvergenzradius.

$$\sum_{n=0}^{\infty} Q_n \frac{z^n}{n!} = \sum_{n=0}^{\infty} a_n z^n$$

wobei

$$a_n = \begin{cases} 0, & n \text{ ungerade} \\ \frac{C_{\frac{n}{2}}}{n!} = ((\frac{n}{2} + 1)! \frac{n!}{2!})^{-1}, & n \text{ gerade} \end{cases}$$

nach der Definition der Catalanzahlen. Wurzelkriterium für Konvergenzradius ergibt:

$$(\frac{n}{2} + 1)! \frac{n!}{2!} \geq 1 \Rightarrow |a_n| \leq 1 \Rightarrow$$

$$r = (\limsup_{n \rightarrow \infty} \sqrt[n]{|a_n|})^{-1} > 1$$

Damit ist das Wahrscheinlichkeitsmaß f_α eindeutig durch seine verallgemeinerten Momente bestimmt.

Verallgemeinerte Momente von $\overline{L_N}$ für $N \rightarrow \infty$

Nachdem bereits f_α durch seine Momente bestimmt wurde, wird im folgenden Abschnitt der Grenzwert der verallgemeinerten Momente von $\overline{L_N}$ untersucht. Hierfür werden die Ergebnisse für β_l benötigt.

$$R_n := \lim_{N \rightarrow \infty} \alpha^{-1-n/2} \int_{\mathbb{R}} \overline{L_N}(dx) x(x - \alpha - 1)^n \quad (1.14)$$

Anwendung des binomischen Lehrsatzes ergibt:

$$\begin{aligned} R_n &= \alpha^{-1-n/2} \lim_{N \rightarrow \infty} \int_{\mathbb{R}} \overline{L_N}(dx) x \sum_{k=0}^n \binom{n}{k} x^{n-k} (-\alpha - 1)^k \\ &= \alpha^{-1-n/2} \lim_{N \rightarrow \infty} \langle \overline{L_N}, \sum_{k=0}^n \binom{n}{k} x^{n+1-k} (-\alpha - 1)^k \rangle \\ &= \alpha^{-1-n/2} \sum_{k=0}^n \binom{n}{k} (-\alpha - 1)^k \beta_{n+1-k} \end{aligned} \quad (1.15)$$

wegen $\beta_l = \lim_{N \rightarrow \infty} \langle \overline{L_N}, x^l \rangle$ und der Linearität des Integrals.

Um die Rekursionsformel zu zeigen benötigt man folgende Formel:

$$\sum_{n=t}^{t+u} \binom{m-n}{m-u-t} \binom{n}{t} = \binom{m+1}{u} \quad \forall t: 0 \leq t \leq m-u \quad (1.16)$$

Beweis von (1.16) Man führt den Beweis anschaulich anhand des Pascalschen Dreiecks durch. Jeder Binomialkoeffizient wird mit einem Knoten im Dreieck identifiziert, sein Wert ist gleich der Anzahl der Pfade, die zu ihm führen.

Betrachte den Punkt $\binom{m+1}{m+1-u}$. Für ein festes $t+1$ sodass $m-u+1 \geq t+1 \geq 1$ definiert man eine Diagonale

$$T = \left\{ \binom{n+1}{t+1} \mid n+1 \geq t+1 \right\}$$

Jeder Pfad von $\binom{0}{0}$ zu $\binom{m+1}{m+1-u}$ muss diese Diagonale überqueren.

Sei $\binom{n+1}{t+1}$ der Punkt, an dem der Pfad T zum ersten Mal berührt. Der vorhergehende Schritt ist damit eindeutig festgelegt, er muss von $\binom{n}{t}$ gekommen sein, sonst wäre er bereits auf T gewesen.

Man summiert nun über alle möglichen ersten Punkte auf T . Die dazugehörigen Pfade im Pascalschen Dreieck werden in zwei Teilpfade unterteilt. Der erste, um $\binom{n+1}{t+1}$ zu erreichen, wobei der letzte Schritt vorgegeben ist. Es bleiben $\binom{n}{t}$ Möglichkeiten für diesen Abschnitt. Der zweite Teilpfad zwischen dem Punkt auf T und dem Ziel hat das kombinatorische Gewicht von $\binom{m-n}{m-u-t}$.

Nun muss man noch betrachten, welches das größtmögliche $n+1$ ist, das noch einen Beitrag zu der Summe liefert. Dies ist der Fall, wenn

$$\binom{m+1-n-1}{m+1-u-t-1} = 1,$$

d.h. es gibt nur noch eine Möglichkeit für den zweiten Teilpfad. Dies ist gleichbedeutend mit $m - n = m - u - t$.

Insgesamt wurde gezeigt:

$$\sum_{n=t}^{t+u} \binom{m-n}{m-u-t} \binom{n}{t} = \binom{m+1}{m+1-u} = \binom{m+1}{u}$$

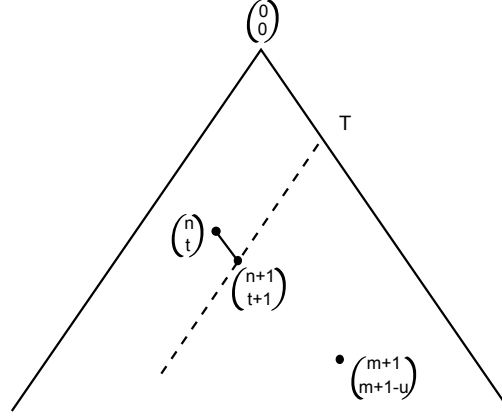


Abbildung 1.3: Darstellung von 1.16

Beweis der Rekursionsformel

Nun betrachtet man:

$$\begin{aligned} \sum_{n=0}^m R_{m-n} R_n &= \sum_{n=0}^m \sum_{k=0}^{m-n} \binom{m-n}{k} (-\alpha - 1)^k \beta_{m-n+1-k} \alpha^{-1-\frac{m-n}{2}} \sum_{l=0}^n \binom{n}{l} (-\alpha - 1)^l \beta_{n+1-l} \alpha^{-1-\frac{n}{2}} \\ &= \alpha^{-1-\frac{m}{2}} \sum_{n=0}^m \sum_{l=0}^n \sum_{k=0}^{m-n} \binom{m-n}{k} \binom{n}{l} (-\alpha - 1)^{k+l} \beta_{m-n+1-k} \gamma_{n+1-l} \end{aligned}$$

Der Index k wird anschließend durch $u := k + l$ ersetzt.

$$\sum_{n=0}^m R_{m-n} R_n = \alpha^{-1-\frac{m}{2}} \sum_{n=0}^m \sum_{l=0}^n \sum_{u=l}^{m+l-n} \binom{m-n}{u-l} \binom{n}{l} (-\alpha - 1)^u \beta_{m-n+1+u-l} \gamma_{n+1-l}$$

Mit $t := n - l$ kann l umgeschrieben werden. Die weiteren Umformungen gelten aufgrund der Eigenschaften des Binomialkoeffizienten und $\gamma_i = \alpha^{-1} \beta_i \forall i \geq 1$.

$$\begin{aligned}
\sum_{n=0}^m R_{m-n} R_n &= \alpha^{-1-\frac{m}{2}} \sum_{n=0}^m \sum_{t=0}^n \sum_{u=n-t}^{m-t} \binom{m-n}{u-n+t} \binom{n}{t} (-\alpha-1)^u \beta_{m+1-u-t} \gamma_{1+t} \\
&= \alpha^{-1-\frac{m}{2}} \sum_{t=0}^m \sum_{u=0}^{m-t} \sum_{n=t}^{m-t} \binom{m-n}{u-n+t} \binom{n}{t} (-\alpha-1)^u \beta_{m+1-u-t} \gamma_{1+t} \\
&= \alpha^{-1-\frac{m}{2}} \sum_{u=0}^{m+1} \sum_{t=0}^{m-u} (-\alpha-1)^u \beta_{m+1-u-t} \gamma_{1+t} \sum_{n=t}^{u+t} \binom{m-n}{m-u-t} \binom{n}{t}
\end{aligned}$$

Mit Formel (1.16) gilt schließlich:

$$\begin{aligned}
\sum_{n=0}^m R_{m-n} R_n &= \alpha^{-1-\frac{m}{2}} \sum_{u=0}^{m+1} \sum_{t=0}^{m-u} (-\alpha-1)^u \beta_{m+1-u-t} \gamma_{1+t} \binom{m+1}{u} \\
&= \alpha^{-1-\frac{m}{2}} \sum_{u=0}^{m+1} \binom{m+1}{u} (-\alpha-1)^u \sum_{t=0}^{m-u} \beta_{m+1-u-t} \gamma_{1+t}
\end{aligned}$$

Durch die Substitution $r := m+1-u-t$ und die Anwendung von Formel (1.10) und (1.15) ergibt sich Folgendes:

$$\begin{aligned}
\sum_{n=0}^m R_{m-n} R_n &= \alpha^{-1-\frac{m}{2}} \sum_{u=0}^{m+1} \binom{m+1}{u} (-\alpha-1)^u \left(\sum_{r=0}^{m+1-u} \beta_r \gamma_{m+2-u-r} - \beta_0 \gamma_{m+2-u} \right) \\
&= \alpha^{-1-\frac{m}{2}} \sum_{u=0}^{m+1} \binom{m+1}{u} (-\alpha-1)^u \left(-\beta_{m+2-u} + \sum_{r=0}^{m+2-u} \beta_r \gamma_{m+2-u-r} - \alpha^{-1} \beta_{m+2-u} \right) \\
&= \alpha^{-1-\frac{m}{2}} \sum_{u=0}^{m+1} \binom{m+1}{u} (-\alpha-1)^u \left(\frac{1}{\alpha} \beta_{m+3-u} - \frac{\alpha+1}{\alpha} \beta_{m+2-u} \right) \\
&= \alpha^{-1-\frac{m}{2}} \sum_{u=0}^{m+2} \binom{m+1}{u} (-\alpha-1)^u \frac{1}{\alpha} \beta_{m+3-u} \\
&\quad - \alpha^{-1-\frac{m}{2}} \sum_{u=0}^{m+1} \binom{m+1}{u} (-\alpha-1)^u \frac{\alpha+1}{\alpha} \beta_{m+2-u} \\
&= \alpha^{-1-\frac{m}{2}} \sum_{u=0}^{m+2} \left(\binom{m+2}{u} - \binom{m+1}{u-1} \right) (-\alpha-1)^u \frac{1}{\alpha} \beta_{m+3-u} \\
&\quad - \alpha^{-1-\frac{m}{2}} \sum_{u=0}^{m+1} \binom{m+1}{u} (-\alpha-1)^u \frac{\alpha+1}{\alpha} \beta_{m+2-u} \\
&= R_{m+2} - \alpha^{-2-m/2} \sum_{u=0}^{m+2} \binom{m+1}{u-1} (-\alpha-1)^u \beta_{m+3-u} \\
&\quad - \alpha^{-2-m/2} (1+\alpha) \sum_{u=0}^{m+1} \binom{m+1}{u} (-\alpha-1)^u \beta_{m+2-u}
\end{aligned}$$

$$\begin{aligned}
 &= R_{m+2} + \alpha^{-2-m/2}(1+\alpha) \sum_{i=0}^{m+1} \binom{m+1}{i} (-\alpha-1)^i \beta_{m+2-i} \\
 &\quad - \alpha^{-2-m/2}(1+\alpha) \sum_{u=0}^{m+1} \binom{m+1}{u} (-\alpha-1)^u \beta_{m+2-u}
 \end{aligned}$$

Also gilt:

$$\sum_{n=0}^m R_{m-n} R_n = R_{m+2} \quad (1.17)$$

Beweis der schwachen Konvergenz

Nun soll gezeigt werden, dass $R_m = Q_m$:

Es wird überprüft, dass $R_0 = Q_0$ und $R_1 = Q_1$, sowie dass die weiteren Folgenglieder von Q_m durch die gleiche Rekursionsformel (1.17) gebildet werden können. Daraus folgt unmittelbar $R_m = Q_m \forall m \in \mathbb{N}_0$

- $R_0 = Q_0, R_1 = Q_1$:

$$\begin{aligned}
 \beta_1 &= \alpha \beta_0 \gamma_0 = \alpha \\
 \beta_2 &= \alpha \beta_0 \gamma_1 + \alpha \beta_1 \gamma_0 = \alpha \gamma_1 + \alpha \beta_1 = \beta_1 + \alpha \beta_1 \\
 R_0 &= \lim_{N \rightarrow \infty} \alpha^{-1} \int_{\mathbb{R}} \overline{L_N}(dx) x = \frac{\beta_1}{\alpha} = 1 \\
 R_1 &= \alpha^{\frac{3}{2}} \lim_{N \rightarrow \infty} \int_{\mathbb{R}} \overline{L_N}(dx) (x^2 - (\alpha+1)x) = \alpha^{\frac{3}{2}} (\beta_2 - (\alpha+1)\beta_1) = 0 \\
 Q_0 &= C_0 = 1, \quad Q_1 = 0
 \end{aligned}$$

- Rekursion für Q_m :

$$\begin{aligned}
 \text{m ungerade: } \sum_{n=0}^m Q_{m-n} Q_n &= Q_0 \underbrace{Q_m}_{=0} + \underbrace{Q_1}_{=0} Q_{n-1} + \dots = 0 = Q_{m+2} \\
 \text{m gerade: } Q_{m+2} &= C_{\frac{m}{2}+1} = \sum_{k=0}^{\frac{m}{2}} C_k C_{\frac{m}{2}-k} = \sum_{k=0}^{\frac{m}{2}} Q_{2k} Q_{m-2k} = \sum_{n=0}^m Q_n Q_{m-n}
 \end{aligned}$$

wobei der letzte Schritt aus $Q_n Q_{m-n} = 0$ für $n = 2k+1$ folgt.

Mit folgendem Theorem aus der Vorlesung kann das Marchenko-Pastur-Gesetz nun bewiesen werden.

1 Das Marchenko-Pastur-Gesetz

Theorem 1.0.2. Sei $(\mu_n)_{n \in \mathbb{N}}$ eine Folge in $\mathcal{M}(\mathbb{R})$ mit endlichen Momenten und $\mu \in \mathcal{M}(\mathbb{R})$ eindeutig durch seine Momente bestimmt. Dann gilt:

$$\forall l \in \mathbb{N}_0 : m_l(\mu_n) \rightarrow m_l(\mu) \Rightarrow \mu_n \xrightarrow{W} \mu$$

Die Voraussetzungen dieses Theorems werden nun überprüft.

- $\forall l \in \mathbb{N}_0 : m_l(\mu_n) \rightarrow m_l(\mu)$
Für $l \geq 1$ existieren $\lambda_0, \dots, \lambda_{l-1} \in \mathbb{R}$, sodass $x^l = \sum_{n=0}^{l-1} \lambda_n x(x - \alpha - 1)^n$. Daraus folgt:

$$\begin{aligned} \forall l \in \mathbb{N} : \lim_{N \rightarrow \infty} m_l(\overline{L_N}) &= \lim_{N \rightarrow \infty} \int_{\mathbb{R}} x^l \overline{L_N}(dx) = \lim_{N \rightarrow \infty} \int_{\mathbb{R}} \sum_{n=0}^{l-1} \lambda_n x(x - \alpha - 1)^n \overline{L_N}(dx) \\ &= \sum_{n=0}^{l-1} \lambda_n \lim_{N \rightarrow \infty} \int_{\mathbb{R}} x(x - \alpha - 1)^n \overline{L_N}(dx) \\ &= \sum_{n=0}^{l-1} \tilde{\lambda}_n R_n = \sum_{n=0}^{l-1} \tilde{\lambda}_n Q_n = m_l(f_\alpha) \end{aligned}$$

Wobei $\tilde{\lambda}_n = \lambda \alpha^{1+n/2}$

Für den Fall $l = 0$ erhält man:

$$\lim_{N \rightarrow \infty} m_0(\overline{L_N}) = \lim_{N \rightarrow \infty} \int_{\mathbb{R}} \overline{L_N}(dx) = 1 = \int_{\mathbb{R}} f_\alpha(x) dx$$

da $\overline{L_N}$ nach Vorlesung ebenfalls ein Wahrscheinlichkeitsmaß ist.

- Es wurde bereits gezeigt, dass f_α eindeutig bestimmt ist.
- Es bleibt zu zeigen, dass R_n endlich ist $\forall n \in \mathbb{N}_0$. Daraus folgt wieder, da $m_l(\overline{L_N})$ endliche Linearkombination von R_n ist, dass die Momente endlich sind.

$$\begin{aligned} \alpha^{-1-n/2} \int_{\mathbb{R}} \overline{L_N}(dx) x(x - \alpha - 1)^n &= \alpha^{-1-n/2} \sum_{k=0}^n \binom{n}{k} (-\alpha - 1)^k \langle \overline{L_N}, x^{n+1-k} \rangle \\ &= \alpha^{-1-n/2} \sum_{k=0}^n \binom{n}{k} (-\alpha - 1)^k \mathbb{E} \left[\frac{1}{N} \text{tr}(W_N^{n+1-k}) \right] < \infty \quad \forall N, n \in \mathbb{N}_0 \end{aligned}$$

wegen $\sup_{j,k,N} \mathbb{E}[|Y_N(j,k)|^q] < \infty$

Damit sind alle Voraussetzungen des Theorems erfüllt \Rightarrow

$$\overline{L_N} \xrightarrow{w} f_\alpha(x) dx \quad (N \rightarrow \infty)$$

Schluss

Eine Anwendung dieses berühmten Marchenko-Pastur-Gesetzes findet man zum Beispiel in der Neurobiologie. Um das Nervensystem zu erforschen, werden Impulsspitzen der Nervenzellen statistisch ausgewertet. Dies geschieht mithilfe des sogenannten “spiked population model“. Es handelt sich um ein multivariates statistisches Modell, bei dem die Populationsgröße sehr groß ist und die Stichprobenanzahl vergleichbar anwächst, so dass beide im Idealfall gegen Unendlich gehen. Mathematisch wird so vorgegangen, dass eine $N \times M(N)$ Zufallsmatrix betrachtet wird, wobei N die Größe der Stichprobe und M die Größe der Population ist, sodass $\frac{M}{N} \rightarrow \alpha$ für $N \rightarrow \infty$. Bestimmte Eigenwerte dieser Matrix werden als bekannt für diese Population vorausgesetzt, und das Ziel ist nun, die restlichen Eigenwerte bzw. deren Erwartungswert herauszufinden. Wie man sieht, ist dafür Marchenko-Pastur von großem Nutzen.

Verweise: J.Baik, J.W.Silverstein, Eigenvalues of Large Sample Covariance Matrices of Spiked Population Models, 2004
und S.Johnston, Comparative Investigation into Classical and Spiking Neurin Implementations on FPGAs, 2005