HANDOUT #5
BEWEIS DES WIGNER'SCHEN HALBKREISSATZES

FOLKMAR BORNEMANN

Aus Terry Taos Blog vom 2. Februar 2010 habe ich gelernt,[1] wie sich der Wigner'sche Halbkreissatz mit Hilfe der Stieltjes-Transformation besonders einfach dann beweisen lässt, wenn man das Phänomen der Maßkonzentration benutzt. Verwendet man hierzu das stochastische Landau-Symbol aus dem Handout #4, so wird der Beweis noch einmal sehr viel durchsichtiger.

1. Das Setting

Voraussetzungen. Es sei M_n eine Folge n-dimensionaler hermitescher Zufallsmatrizen, deren Diagonalen Null sind. Für gegebenes n bestehe das obere Dreieck der Matrix M_n aus unabhängigen komplexen Zufallsvariablen, für die gelte:[2]

$$|(M_n)_{ij}| \leqslant c, \qquad \mathbb{E}(M_n)_{jk} = 0, \qquad \mathbb{E}|(M_n)_{jk}|^2 = 1 \qquad (j < k).$$

Die Folge der zugehörigen skalierten *Wigner-Matrizen* sei $X_n = n^{-1/2} M_n$.

Stieltjes-Transformation des empirischen Spektralmaßes. Die Stieltjes-Transformation des empirischen Spektralmaßes

$$L_n = \frac{1}{n} \sum_{j=1}^{n} \delta_{\lambda_j(X_n)}$$

von X_n ist

$$s_n(z) = \int_{-\infty}^{\infty} \frac{dL_n(x)}{x - z} = \frac{1}{n} \sum_{j=1}^{n} \frac{1}{\lambda_j(X_n) - z} = \frac{1}{n} \operatorname{tr}(X_n - zI)^{-1} \qquad (z \in \mathbb{H}).$$

Der schwache Grenzwert der empirischen Spektralmaße L_n lässt sich nun anhand des punktweisen Limes von $s_n(z)$ bestimmen (siehe Satz 3 auf Handout #3).

Partitionierung. Wir partitionieren entlang der letzten Zeile und Spalte gemäß

$$X_n = \begin{pmatrix} \check{X}_n & \frac{1}{\sqrt{n}} v_n \\ \frac{1}{\sqrt{n}} v_n^H & 0 \end{pmatrix}$$

und definieren

$$\check{s}_n(z) = \frac{1}{n} \sum_{j=1}^{n-1} \frac{1}{\lambda_j(\check{X}_n) - z} = \frac{1}{n} \operatorname{tr}(\check{X}_n - zI)^{-1}.$$

Date: 17. Juni 2010.

[1] `http://terrytao.wordpress.com/2010/02/02/254a-notes-4-the-semi-circular-law/`.

[2] Beachte, dass hier weder Gleichverteilung vorausgesetzt wird, noch eine Beziehung zwischen den Matrizen M_n für verschiedene n.

Konstanten. In den folgenden Abschnitten hängen die Konstanten der O-Terme und der Tailabschätzungen in den $\mathbb{O}$-Termen nur von $\operatorname{Im} z$ und der gemeinsamen Schranke c aller Elemente der M_n ab.

2. Trennung des Spektrums

Wir verwenden die Partitionierung der Matrix X_n auf zweifache Weise. Zum ersten benutzen wir, dass sich die Eigenwerte von X_n und $\check{X}_n$ gegenseitig trennen:

$$\lambda_1(X_n) \leqslant \lambda_1(\check{X}_n) \leqslant \lambda_2(X_n) \leqslant \lambda_2(\check{X}_n) \leqslant \cdots \leqslant \lambda_{n-1}(X_n) \leqslant \lambda_{n-1}(\check{X}_n) \leqslant \lambda_n(X_n).$$

(Das folgt aus der variationellen Charakterisierung von Eigenwerten hermitescher Matrizen.) Wenn wir nun Lemma 2 auf die Funktion $f_z(x) = 1/(x-z)$ anwenden, so erhalten wir daraus (mit Hilfe von Lemma 3)

$$\check{s}_n(z) - s_n(z) = \frac{1}{n}\sum_{j=1}^{n-1} \frac{1}{\lambda_j(\check{X}_n) - z} - \frac{1}{n}\sum_{j=1}^{n} \frac{1}{\lambda_j(X_n) - z} = O(n^{-1}).$$

Da $\check{s}_n(z)$ gar nicht von v_n abhängt, variiert $s_n(z)$ bezüglich beliebiger[3] Änderungen von v_n demnach maximal um den Betrag $O(n^{-1})$. Aus Symmetriegründen gilt das nämliche für jede Spalte; also erst recht für den Teil jener Spalte im oberen Dreieck $(X_n)_{ij}$ $(i<j)$. Alle solche Teilspalten sind aber voneinander stochastisch unabhängig, so dass nach McDiarmid gilt:

$$s_n(z) = \mathbb{E}s_n(z) + \mathbb{O}(\sqrt{n}\cdot n^{-1}) = \mathbb{E}s_n(z) + \mathbb{O}(n^{-1/2}).$$

Borel–Cantelli liefert

$$s_n(z) - \mathbb{E}s_n(z) \overset{\text{a.s.}}{\to} 0.$$

Es bleibt, den Limes der Folge $\mathbb{E}s_n(z)$ zu ermitteln; dabei ist

$$\mathbb{E}s_n(z) = \frac{1}{n}\mathbb{E}\operatorname{tr}(X_n - zI)^{-1} = \frac{1}{n}\sum_{j=1}^{n} \mathbb{E}((X_n - zI)^{-1})_{jj}.$$

3. Resolventen

Zum zweiten entsteht aus der Partitionierung von X_n, dass

$$X_n - zI = \begin{pmatrix} \check{X}_n - zI & \frac{1}{\sqrt{n}}v_n \\ \frac{1}{\sqrt{n}}v_n^H & -z \end{pmatrix}.$$

Da die Spektren von X_n und $\check{X}_n$ reell sind und $\operatorname{Im} z > 0$ gilt, sind nach dem Spektralsatz $X_n - zI$ und $\check{X}_n - zI$ invertierbar, so dass Lemma 1 anwendbar ist:

$$((X_n - zI)^{-1})_{nn} = -\frac{1}{z + \frac{1}{n}v_n^H(\check{X}_n - zI)^{-1}v_n}.$$

Nach dem Spektralsatz und Lemma 3 ist

$$\|(\check{X}_n - zI)^{-1}\| = \max_{j=1,\dots,n} \frac{1}{|\lambda_j(\check{X}_n) - z|} = O(1).$$

[3]Wir verwenden hier also *nicht* die gleichmäßige Beschränktheit $|(v_n)_j| \leqslant c$. Ebenso hätte X_n eine beliebige Diagonale haben dürfen. Das Resultat dieses §2 ist daher für beliebige hermitesche Zufallsmatrizen X_n richtig, deren Elementen auf und über der Diagonale *unabhängig* sind.

Beispiel 2 in Handout #4 liefert daher zusammen mit den Ergebnissen aus §2, dass

$$\frac{1}{n} v_n^H (\check{X}_n - zI)^{-1} v_n = \check{s}_n(z) + \mathbb{O}(n^{-1/2}) = \mathbb{E}s_n(z) + \mathbb{O}(n^{-1/2}).$$

Nach Lemma 3 ist die Funktion $x \mapsto 1/(z+x)$ Lipschitz-stetig, so dass

$$((X_n - zI)^{-1})_{nn} = -\frac{1}{z + \frac{1}{n} v_n^H (\check{X}_n - zI)^{-1} v_n} = -\frac{1}{z + \mathbb{E}s_n(z)} + \mathbb{O}(n^{-1/2}).$$

Da dies aus Symmetriegründen für jedes Diagonalelement $((X_n - zI)^{-1})_{jj}$ richtig sein muss, gilt für ihre gemittelten Erwartungswerte die asymptotische Gleichung

$$\mathbb{E}s_n(z) = -\frac{1}{z + \mathbb{E}s_n(z)} + O(n^{-1/2}).$$

4. Grenzwertbestimmung

Wegen $s_n(z) = O(1)$ gilt $\mathbb{E}s_n(z) = O(1)$; aus $\operatorname{Im} s_n(z) > 0$ folgt $\operatorname{Im} \mathbb{E}s_n(z) > 0$. Für beliebiges, aber festes $z \in \mathbb{H}$ gibt es daher eine Teilfolge n', so dass

$$\mathbb{E}s_{n'}(z) \to s(z), \qquad \operatorname{Im} s(z) \geqslant 0.$$

Dieser Häufungspunkt genügt nach §3 der Gleichung

$$s(z) = -\frac{1}{z + s(z)};$$

daher gilt $\operatorname{Im} s(z) > 0$ (anderenfalls wäre z reell). Auflösung ergibt

$$s(z) = \frac{1}{2}(-z \pm \sqrt{z^2 - 4});$$

wobei wir anhand von $\operatorname{Im} z > 0$ feststellen, dass nur eine einzige Wahl des Vorzeichens auf $\operatorname{Im} s(z) > 0$ führen kann. Der Häufungspunkt $s(z)$ ist demnach *eindeutig* bestimmt und wir erhalten ohne jede Teilfolgenauswahl

$$\mathbb{E}s_n(z) \to s(z) = \frac{1}{2}(-z \pm \sqrt{z^2 - 4}) \qquad (z \in \mathbb{H}).$$

Die auf diese Weise eindeutig definierte Funktion $s(z)$ lässt sich auf $\mathbb{C} \setminus [-2, 2]$ analytisch forsetzen, wobei $s(z) = -z^{-1} + O(z^{-3})$ für $z \to \infty$ gilt. Nach dem Herglotz'schen Darstellungssatz ist $s(z)$ daher die Stieltjes-Transformierte eines W-Maßes L. Da nun punktweise in $x \in \mathbb{R}$

$$\lim_{\epsilon \downarrow 0} \frac{1}{\pi} \operatorname{Im} s(x + i\epsilon) = \frac{1}{2\pi} \sqrt{(4 - x^2)_+}$$

gilt (dominiert durch $1/\pi$), liefert der Satz von der dominierten Konvergenz

$$\frac{1}{\pi} \operatorname{Im} s(x + i\epsilon)\, dx \xrightarrow{v} L = \frac{1}{2\pi} \sqrt{(4 - x^2)_+}\, dx,$$

so dass L tatsächlich die *Halbkreisverteilung* ist. Zusammenfassend erhalten wir den *Wigner'schen Halbkreissatz*:

$$\boxed{L_n \xrightarrow{w} L = \frac{1}{2\pi} \sqrt{(4 - x^2)_+}\, dx \quad \text{fast sicher.}}$$

Bemerkung 1. Der Beweis funktioniert ohne große Änderung auch mit einer von Null verschiedenen Diagonalen: Wenn wir die Elemente auf und über der Diagonalen von M_n als unabhängig voraussetzen und verlangen, dass $(M_n)_{jj} = \mathbb{O}(1)$ ist, so bleibt die Fehlerabschätzung $O(n^{-1/2})$ der asymptotischen Gleichung für $\mathbb{E}s_n(z)$ richtig. Jetzt überträgt sich der Beweis auch ohne weitere Änderungen auf Gauß'sche Wigner-Matrizen, da Beispiel 2 aus Handout #4 hier richtig bleibt.

5. Einige Hilfsresultate

Lemma 1. *Die Matrix $A \in \mathbb{C}^{n\times n}$ sei entlang der letzten Zeile und Spalte partitioniert:*

$$A = \begin{pmatrix} \breve{A} & v \\ u^H & \alpha \end{pmatrix}.$$

Sind A und $\breve{A}$ invertierbar, so ist $\alpha \neq u^H \breve{A}^{-1} v$ und es gilt

$$(A^{-1})_{nn} = \frac{1}{\alpha - u^H \breve{A}^{-1} v}.$$

Beweis. Setze $e_n = (0, \dots, 0, 1)^H$. Die letzte Komponente der Lösung x des Gleichungssystems $Ax = e_n$ liefert dann die gesuchte Größe $x_n = (A^{-1})_{nn}$. Block-Gauß-Elimination ergibt das System

$$\begin{pmatrix} \breve{A}^{-1} & 0 \\ -u^H \breve{A}^{-1} & 1 \end{pmatrix} e_n = \begin{pmatrix} \breve{A}^{-1} & 0 \\ -u^H \breve{A}^{-1} & 1 \end{pmatrix} \begin{pmatrix} \breve{A} & v \\ u^H & \alpha \end{pmatrix} x = \begin{pmatrix} I & \breve{A}^{-1} v \\ 0 & \alpha - u^H \breve{A}^{-1} v \end{pmatrix} x,$$

dessen letzte Zeile $1 = (\alpha - u^H \breve{A}^{-1} v) x_n$ lautet.[4] □

Lemma 2. *Es sei $x_1 \leqslant x_1' \leqslant x_2 \leqslant x_2' \leqslant \cdots \leqslant x_{n-1} \leqslant x_{n-1}' \leqslant x_n$; $f : \mathbb{R} \to \mathbb{C}$ sei stetig differenzierbar und von beschränkter Variation. Dann gilt:*

$$\left| \sum_{k=1}^{n-1} f(x_k') - \sum_{k=1}^{n} f(x_k) \right| \leqslant \|f\|_\infty + \int_{-\infty}^{\infty} |f'(x)|\, dx.$$

Beweis. Wegen der Disjunktheit der Integrationsintervalle gilt

$$\left| \sum_{k=1}^{n-1} (f(x_k') - f(x_k)) \right| = \left| \sum_{k=1}^{n-1} \int_{x_k}^{x_k'} f'(x)\, dx \right| \leqslant \sum_{k=1}^{n-1} \int_{x_k}^{x_k'} |f'(x)|\, dx \leqslant \int_{-\infty}^{\infty} |f'(x)|\, dx.$$

Ferner ist $|f(x_n)| \leqslant \|f\|_\infty$, womit auch schon alles bewiesen ist. □

Lemma 3. *Es seien $z \in \mathbb{H}$ und $x, x' \in \mathbb{R}$. Dann gilt für $f_z(x) = 1/(x - z)$:*

$$\text{(i) } \|f_z\|_\infty \leqslant \frac{1}{\operatorname{Im} z}; \quad \text{(ii) } \int_{-\infty}^{\infty} |f_z'(x)|\, dx = \frac{\pi}{\operatorname{Im} z}; \quad \text{(iii) } \operatorname{Lip} f_z \leqslant \frac{1}{(\operatorname{Im} z)^2}.$$

Beweis. Setze $z = a + ib$ mit $a \in \mathbb{R}$ und $b > 0$. Eine direkte Rechnung zeigt

$$|f_z(x)| = \frac{1}{\sqrt{b^2 + (x-a)^2}} \leqslant \frac{1}{b}, \qquad f_z'(x) = -\frac{1}{(x-z)^2},$$

woraus unmittelbar (i), (ii) und (iii) folgt. □

[4] Der Ausdruck $\alpha - u^H \breve{A}^{-1} v$ ist das *Schur-Komplement* von $\breve{A}$ in A.