

Spektraldichte großer Matrizen

Eine numerische Annäherung

Carina Seidel

Universität Potsdam

3. Juli 2024



Inhaltsverzeichnis

- 1 Einleitung
- 2 Die Kernel-Polynom-Methode
- 3 Qualitätsanalyse der Annäherungen



Motivation

- Eigenwerte einer Matrix sind in vielen Bereichen der Mathematik und Physik interessant
- Oftmals sind Matrizen zu groß um diese effizient zu berechnen
- Stattdessen berechnen wir die Spektraldichte (Density of States)
- Der Einfachheit halber beschränken wir uns auf symmetrische, reelle Matrizen, die spärlich besetzt sind



Distribution

Definition (Funktional)

Sei V ein \mathbb{R} -Vektorraum. Ein Funktional T ist eine Abbildung $T : V \rightarrow \mathbb{R}$

Definition (Distribution)

Sei $\emptyset \neq \Omega \subset \mathbb{R}^n$ offen. Sei \mathcal{E} der Raum der *Testfunktionen* über Ω . Eine Distribution T ist eine Abbildung $T : \mathcal{E} \rightarrow \mathbb{R}$ wobei für alle $g, g_1, g_2, \{g_n\}_{n \in \mathbb{N}} \in \mathcal{E}$ mit $\lim_{n \rightarrow \infty} g_n \rightarrow g$ gilt:

$$T(g_1 + \lambda g_2) = T(g_1) + \lambda T(g_2) \quad \text{und} \quad \lim_{n \rightarrow \infty} T(g_n) \rightarrow T(g)$$

Kurz: Eine Distribution T ist ein stetiges und lineares Funktional auf \mathcal{E}



Delta-Distribution

Definition (Delta-Distribution)

Sei $\mathcal{E} = \mathcal{C}^\infty(\Omega)$ mit $0 \in \Omega \subset \mathbb{R}^n$

Dann ist $\delta : \mathcal{E} \rightarrow \mathbb{R}, f \mapsto f(0)$ mit $\delta(f) = \langle \delta, f \rangle = f(0)$

Wichtige Eigenschaft:

$$\int_{-\infty}^{\infty} f(x) \delta(x - a) \, dx = \int_{-\infty}^{\infty} f(x) \delta(a - x) \, dx = f(a) \implies \int_{-\infty}^{\infty} \delta(x - a) \, dx = 1$$



Spektraldichte

Sei $A \in \mathbb{R}^{n \times n}$, $A^T = A$ und A spärlich besetzt.

Dann ist die Spektraldichte (engl. Density of States (DOS)) definiert als

$$\phi(t) = \frac{1}{n} \sum_{j=1}^n \delta(t - \lambda_j) \quad (1)$$

wobei δ die Delta-Distribution und λ_j die Eigenwerte von A in nicht-absteigender Reihenfolge sind.

Die Anzahl der Eigenwerte in einem Intervall $[a, b]$ kann dann wie folgt ausgedrückt werden:

$$\nu_{[a,b]} = \int_a^b \sum_j \delta(t - \lambda_j) dt \equiv \int_a^b n\phi(t) dt \quad (2)$$



Problemstellung

- Spektraldichte trivial wenn Eigenwerte von A bekannt
- Unpraktisch wenn A sehr groß, da Berechnung teuer
- Wir brauchen effiziente Alternativen um $\phi(t)$ abzuschätzen
- Allerdings: $\phi(t)$ keine "Funktion" im eigentlichen Sinne



- Sei $I \subseteq \mathbb{R}$ das Intervall, dass das Spektrum von A beinhaltet.
- Teile nun I in kleinere Teilintervalle $[t_i, t_{i+1}]$
- Benutze den Silvestreschen Trägheitssatz um die Eigenwerte in jedem Teilintervall zu zählen.
- Berechne den Durchschnittswert von $\phi(t)$ in jedem dieser Intervalle mithilfe von
- Für $(t_{i+1} - t_i) \rightarrow 0$ nähern sich die Histogramme der Spektraldichte.
- Problem: Berechnung der Zerlegung $A - t_i I = LDL^T$ für alle t_i ist zu zeitaufwendig.
- Besser: A nur mit Vektoren multiplizieren.

Annahmen

- Wir betrachten zwei Methoden zur Annäherung der Spektraldichte
- Der Einfachheit halber sei im Folgenden immer $A \in \mathbb{R}^{n \times n}$, $A^T = A$
- Die Verallgemeinerung auf hermitesche Matrizen ist im Nachhinein unkompliziert
- Zunächst die Kernel-Polynom-Methode (KPM)
- Danach das klassische Lanczos-Verfahren zur teilweisen Diagonalisierung von A
- Schwierig zu beurteilen welche Methode die beste ist



Kernel-Polynom-Methode (KPM)

- Formelle polynomiale Erweiterung der Spektraldichte.
- Macht von der Moment Matching Methode gebrauch.
- Wir zeigen, wie das Lanczos-Spektroskopieverfahren mit der KPM zusammenhängt
- Eine weitere Variante ist die Delta-Gauss-Legendre Methode



- Zwei Klassen, KPM und Lanczos-Spektroskopieverfahren
- Zwei Methoden sind äquivalent zur KPM
- Die andere Klasse basiert auf der partiellen Tridiagonalisierung
- Zwei Methoden, Gaußscher und Lorentzscher Weichzeichnung (Regularisierung)
- Alle Methoden benutzen eine stochastische Sampling-Methode, die auf folgendem Resultat basiert:

Theorem

Sei $A \in \mathbb{R}^{n \times n}$ mit Spektralzerlegung $A = \sum_{j=1}^n \lambda_j u_j u_j^T$ wobei $u_i u_j^T = \delta_{ij}$.

Sei außerdem $v \in \mathbb{R}^n$ mit $v = \sum_{j=1}^n \beta_j u_j$

Gilt $v_i \sim \mathcal{N}(0, 1)$ i.i.d. für die Komponenten $\{v_i\}_{i=1}^n$ von v , also $\mathbb{E}[v] = 0$ und $\mathbb{E}[vv^T] = \mathbb{1}_n$, dann

$$\mathbb{E}[\beta_i \beta_j] = \delta_{ij}, \quad i, j \in \mathbb{N}_1^n$$



Beweis Theorem 1.

Für $U = [u_1, u_2, \dots, u_n]$ und $\beta = (\beta_1, \beta_2, \dots, \beta_n)^T$ gilt $v = U\beta$
Daher folgt, dass

$$\mathbb{E}[v] = \mathbb{E}[U\beta] = U\mathbb{E}[\beta] = 0 \implies \mathbb{E}[\beta] = 0$$

Weiterhin gilt, dass

$$\mathbb{E}[vv^T] = \mathbb{E}[(U\beta)(U\beta)^T] = \mathbb{E}[U\beta\beta^T U^T] = U\mathbb{E}[\beta\beta^T]U^T = \mathbb{1}_n$$

woraus folgt, dass $\mathbb{E}[\beta\beta^T] = \mathbb{1}_n$



Resultat

Sei $f(A)$ eine Matrixfunktion. Dann haben wir

$$\begin{aligned}\mathbb{E} [v^T f(A) v] &= \mathbb{E} [(U\beta)^T f(U\Lambda U^T)(U\beta)] \\ &= \mathbb{E} [\beta^T U^T U f(\Lambda) U^T U \beta] \\ &= \mathbb{E} [\beta^T f(\Lambda) \beta] \\ &= \mathbb{E} \left[\sum_{j=1}^n \beta_j^2 f(\lambda_j) \right] \\ &= \sum_{j=1}^n f(\lambda_j) \mathbb{E} [\beta_j^2] \\ &= \sum_{j=1}^n f(\lambda_j)\end{aligned}$$



Tschebyschev-Polynome

Mit Hilfe der trigonometrischen Funktionen können die Tschebyschev wie folgt ausgedrückt werden:

$$T_k(t) = \begin{cases} \cos(k \arccos(t)) & \text{für } k \in [-1, 1] \\ \cosh(k \operatorname{arcosh}(t)) & \text{für } k > 1 \\ (-1)^k \cosh(k \operatorname{arcosh}(-t)) & \text{für } k < -1 \end{cases}$$

Es gilt außerdem die Rekursionsformel

$$T_{k+1}(t) = 2tT_k(t) - T_{k-1}(t)$$



Transformation der Eigenwerte

Wir benutzen im Folgenden $T_k(t) = \cos(k \arccos(t))$ um die Dirac-Dichte zu erweitern.

Wir müssen uns daher auf Matrizen beschränken, deren Eigenwerte im Intervall $[-1, 1]$ liegen.

Seien daher λ_{us} und λ_{os} jeweils die untere bzw. obere Schranke für die Eigenwerte von A . Definiere

$$c := \frac{\lambda_{us} + \lambda_{os}}{2} \quad \text{und} \quad d := \frac{\lambda_{os} - \lambda_{us}}{2}$$

Dann ist $B = \frac{A - c * \mathbb{1}_n}{d}$ eine Matrix mit Eigenwerten im Intervall $[-1, 1]$



Erweiterung der Delta-Distribution

Berechne zunächst

$$\hat{\phi}(t) = \sqrt{1-t^2} \phi(t) = \sqrt{1-t^2} \times \frac{1}{n} \sum_{j=1}^n \delta(t - \lambda_j)$$

Nun erweitern wir

$$\hat{\phi}(t) = \sum_{k=0}^{\infty} \mu_k T_k(t)$$

im Sinne von

$$\int_{-1}^1 \hat{\phi}(t) g(t) \, dt = \int_{-1}^1 \sum_{k=0}^{\infty} \mu_k T_k(t) g(t) \, dt$$

für $g \in \mathcal{S}$



$$\begin{aligned}\mu_k &= \frac{2 - \delta_{k0}}{\pi} \int_{-1}^1 \frac{1}{\sqrt{1-t^2}} T_k(t) \hat{\phi}(t) \, dt \\&= \frac{2 - \delta_{k0}}{\pi} \int_{-1}^1 \frac{1}{\sqrt{1-t^2}} T_k(t) \sqrt{1-t^2} \phi(t) \, dt \\&= \frac{2 - \delta_{k0}}{n\pi} \sum_{j=1}^n T_k(\lambda_j) \\&= \frac{2 - \delta_{k0}}{n\pi} \operatorname{Spur}(T_k(A))\end{aligned}$$

Weitere Parameter

Es folgt, dass

$$\zeta_k = \frac{1}{n_{vec}} \sum_{l=1}^{n_{vec}} \left(v_0^{(l)} \right)^T T_k(A) v_0^{(l)}$$

ein guter Schätzer für $\text{Spur}(T_k(A))$ ist und damit

$$\mu_k \approx \frac{2 - \delta_{k0}}{n\pi} \zeta_k$$



Berechnung der ζ_k

Sei im Folgenden $v_0 \equiv v_0^{(l)}$ Berechne nun

$$T_{k+1}(A)v_0 = 2AT_k(A)v_0 - T_{k-1}(A)v_0$$

Für $v_k \equiv T_k(A)v_0$ gilt also, dass

$$v_{k+1} = 2Av_k - v_{k-1}$$



Definition $\tilde{\phi}_M(t)$

Sobald die $\{\mu_k\}$ bestimmt sind, wäre die Erweiterung theoretisch durch

$$\phi(t) = \frac{1}{\sqrt{1-t^2}} \hat{\phi}(t)$$

gegeben. Allerdings gilt

$$\lim_{k \rightarrow \infty} \mu_k \rightarrow 0$$

und wir interessieren uns nur für $T_k(t)$ mit $k \leq M$

Daher schätzen wir ϕ durch

$$\tilde{\phi}_M(t) = \frac{1}{\sqrt{1-t^2}} \sum_{k=0}^M \mu_k T_k(t)$$



Problemstellung

- Sei im Folgenden $\tilde{\phi}(t)$ eine reguläre Funktion die die Spektraldichte schätzt
- Alle Annäherungen sind stetige Funktionen
- $\phi(t)$ ist keine Funktion im eigentlichen Sinne
- Wir können nicht die L^p -Norm benutzen, um $\phi(t) - \tilde{\phi}(t)$ abzuschätzen
- Zwei Möglichkeiten, dies zu umgehen



Schwartz-Raum über \mathbb{R}

$$\mathcal{S}(\mathbb{R}) := \left\{ f \in \mathcal{C}^\infty(\mathbb{R}) \mid \forall p, k \in \mathbb{N}_0 : \sup_{x \in \mathbb{R}} \left| x^p f^{(k)}(x) \right| < \infty \right\}$$



Erste Methode

Wir benutzen die Tatsache, dass $\delta(t)$ eine Verteilung ist:

Sei $g \in \mathcal{C}^\infty(\mathbb{R})$ eine Testfunktion aus dem Schwartz-Raum \mathcal{S} , dann

$$\langle \delta(\cdot - \lambda), g \rangle \equiv \int_{-\infty}^{\infty} \delta(t - \lambda) g(t) \, dt \equiv g(\lambda)$$

und für alle $p, k \in \mathbb{N}_0$

$$\sup_{t \in \mathbb{R}} |t^p g^{(k)}(t)| < \infty$$

Dann wird der Fehler wie folgt gemessen:

$$\epsilon_1 = \sup_{g \in \mathcal{S}} \left| \langle \phi, g \rangle - \langle \tilde{\phi}, g \rangle \right|$$



Zweite Methode

Wir "regularisieren" die $\delta(t)$ -Funktionen

Dazu ersetzen wir sie durch stetige und glatte Funktionen

Zum Beispiel die Gaussche Normalverteilung mit einer angemessene Standardabweichung σ

Die daraus entstandene Funktion $\phi_\sigma(t)$ ist wohldefiniert

Für $p = 1, 2$ und ∞ können wir folgenden Fehler berechnen:

$$\epsilon_2 = \left\| \phi_\sigma(t) - \tilde{\phi}(t) \right\|_p \quad (3)$$

Diese beiden Methoden sind eng verwandt!



Der Begriff der Auflösung

Selten ist eine exakte Annäherung aller Eigenwerte von A gewünscht. Oftmals genügt es, die Anzahl der Eigenwerte in einem beliebigen Teilintervall $[a, b]$ des Spektrums zu wissen.

Die Größe $b - a$ dieses Teilintervalls bezeichnet man als Auflösung der Schätzung:

Je kleiner das Teilintervall, desto höher die Auflösung.

Die Genauigkeit der Annäherung ist nur bis zur gewünschten Auflösung aussagekräftig.

Für $\epsilon_2 = \left\| \phi_\sigma(t) - \tilde{\phi}(t) \right\|_p$ aus (3) gilt:

Je kleiner das σ , desto höher die Auflösung.



Noch mehr Probleme mit Dirac

Betrachte

$$\nu_{[a,b]} = \int_a^b n\phi(t) \, dt$$

aus (2). Definiere entsprechend

$$\tilde{\nu}_{[a,b]} = \int_a^b n\tilde{\phi}(t) \, dt$$

mit $\tilde{\phi}(t) \in \mathcal{C}^\infty$

Noch mehr Probleme mit Dirac (2)

Angenommen, $n = 1$ und $\phi(t) = \delta(t)$.

Unendliche Auflösung bedeutet $|\nu_{[a,b]} - \tilde{\nu}_{[a,b]}|$ soll für $[a,b]$ beliebig klein ebenfalls klein sein.

Sei also $a = -\varepsilon, b = \varepsilon$. Aus der Definition der δ -Funktion folgt dann, dass

$$\lim_{\varepsilon \rightarrow 0+} \nu_{[-\varepsilon, \varepsilon]} = 1$$

während für glatte Funktionen $\tilde{\phi}$ selbstverständlich gilt, dass

$$\lim_{\varepsilon \rightarrow 0+} \tilde{\nu}_{[-\varepsilon, \varepsilon]} = 0$$

Fazit: Keine glatte Funktion konvergiert zur Spektraldichte unter stetiger Erhöhung der Auflösung



Einschränkung des Schwartz-Raums

Eine endliche Auflösung ist oftmals genug.

Wir können den Schwartz-Raum \mathcal{S} also einschränken.

Beispiel: Betrachte nur Gaussche Verteilungsfunktionen der Form

$$g_\sigma(t) = \frac{1}{(2\pi\sigma^2)^{\frac{1}{2}}} e^{-\frac{t^2}{2\sigma^2}}$$

und schränke \mathcal{S} auf den Unterraum

$$\mathcal{S}(\sigma; [\lambda_{lb}, \lambda_{ub}]) = \{g \mid g(t) \equiv g_\sigma(t - \lambda), \lambda \in [\lambda_{lb}, \lambda_{ub}]\}$$

Hierbei sind λ_{lb} und λ_{ub} jeweils das Infimum und Supremum der Eigenwerte von A und der Parameter σ die *Zielauflösung*.

Wir können nun die folgende Metrik zur Qualitätsbewertung nutzen:

$$E \left[\tilde{\phi}; \mathcal{S}(\sigma; [\lambda_{lb}, \lambda_{ub}]) \right] = \sup_{g \in \mathcal{S}(\sigma; [\lambda_{lb}, \lambda_{ub}])} \left| \langle \phi, g \rangle - \langle \tilde{\phi}, g \rangle \right|$$



Regularisierung der Spektraldichte

- Konstruiere zunächst eine glatte Darstellung der δ -Funktion.
- Dies muss verhältnismäßig zur gewünschten Auflösung sein.
- Der Fehler kann dann direkt berechnet werden
- Wahl des σ : so groß wie möglich für leichte Annäherung, so klein wie möglich für Genauigkeit



Regularisierung der Spektraldichte mit Gauss

Sei

$$\phi_\sigma(t) = \langle \phi(\cdot), g_\sigma(\cdot - t) \rangle = \sum_{j=1}^n g_\sigma(t - \lambda_j)$$

Dies ist dann nicht anderes als die "Weichzeichnung" der Spektraldichte durch Gauß-Funktionen der Breite σ

Genauso sei

$$\tilde{\phi}_\sigma(t) = \langle \tilde{\phi}(\cdot), g_\sigma(\cdot - t) \rangle$$

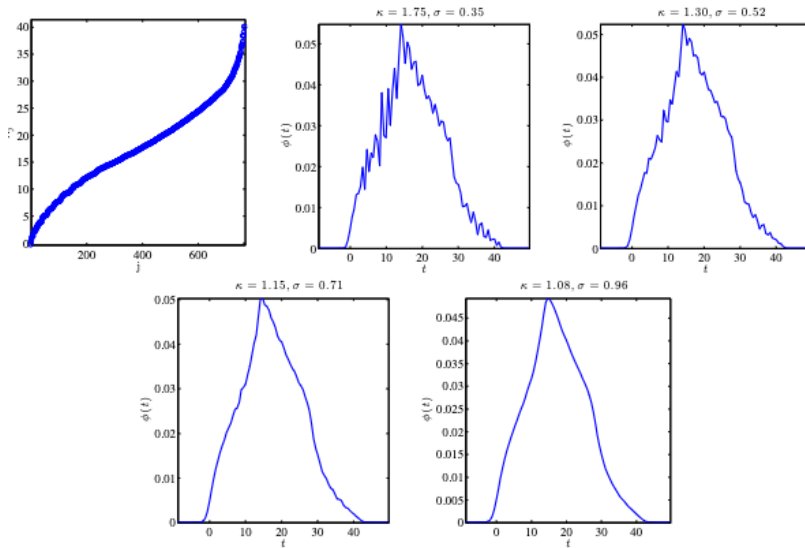
Dann ist

$$E \left[\tilde{\phi}; \mathcal{S}(\sigma; [\lambda_{lb}, \lambda_{ub}]) \right] = \sup_{g \in \mathcal{S}(\sigma; [\lambda_{lb}, \lambda_{ub}])} \left| \langle \phi(\cdot), g_\sigma(\cdot - t) \rangle - \langle \tilde{\phi}(\cdot), g_\sigma(\cdot - t) \rangle \right|$$

der L^∞ -Fehler zwischen zwei wohldefinierten Funktionen



Schöne Bilder



Regularisierung der Spektraldichte mit Lorentz

Die Lorentz-Funktion ist definiert durch

$$\frac{\eta}{(t - \lambda)^2 + \eta^2} = -\operatorname{Im} \left(\frac{1}{t - \lambda + i\eta} \right),$$

wobei η eine kleine Regularisierungskonstante ist.

Für $\eta \rightarrow 0$ nähert sich die Lorentz-Funktion der Dirac-Funktion um den Eigenwert λ an

Dies ist später für die Haydock-Methode relevant.

Die Bedingung der Nicht-Negativität

- Die Spektraldichte ist als Wahrscheinlichkeitsverteilung nicht-negativ, also

$$\forall g \in \mathcal{S}, g \geq 0 : \langle \phi, g \rangle \geq 0$$

- Einige numerische Annäherungen brechen mit dieser Eigenschaft
- Das führt zu großen Fehlern