

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
- 4 Auswertung



Motivation



Motivation

- Spektraldichte einer Matrix in der Physik interessant

Motivation

- Spektraldichte einer Matrix in der Physik interessant
- DOS zeigt Wahrscheinlichkeit für Eigenwerte nahe $p \in \mathbb{R}$



Motivation

- Spektraldichte einer Matrix in der Physik interessant
- DOS zeigt Wahrscheinlichkeit für Eigenwerte nahe $p \in \mathbb{R}$
- Matrizen oft zu groß um Eigenwerte effizient zu berechnen



Motivation

- Spektraldichte einer Matrix in der Physik interessant
- DOS zeigt Wahrscheinlichkeit für Eigenwerte nahe $p \in \mathbb{R}$
- Matrizen oft zu groß um Eigenwerte effizient zu berechnen
- Bedarf an Methoden, die DOS kostengünstig abzuschätzen



Motivation

- Spektraldichte einer Matrix in der Physik interessant
- DOS zeigt Wahrscheinlichkeit für Eigenwerte nahe $p \in \mathbb{R}$
- Matrizen oft zu groß um Eigenwerte effizient zu berechnen
- Bedarf an Methoden, die DOS kostengünstig abzuschätzen
- Beschränkung auf reelle, symmetrische, spärlich besetzte Matrizen



Motivation

- Spektraldichte einer Matrix in der Physik interessant
- DOS zeigt Wahrscheinlichkeit für Eigenwerte nahe $p \in \mathbb{R}$
- Matrizen oft zu groß um Eigenwerte effizient zu berechnen
- Bedarf an Methoden, die DOS kostengünstig abzuschätzen
- Beschränkung auf reelle, symmetrische, spärlich besetzte Matrizen
- Dieser Vortrag basiert auf dem Paper APPROXIMATING SPECTRAL DENSITIES OF LARGE MATRICES von Lin Lin, Yousef Saad, and Chao Yang



Ein paar Definitionen vorab



Ein paar Definitionen vorab

Definition (Funktional)

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



Ein paar Definitionen vorab

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



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 $\delta(f) = \langle \delta, f \rangle = f(0)$



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



Spektraldichte

Definition (Spektraldichte)

Sei $A \in \mathbb{R}^{n \times n}$, $A^T = A$ und A spärlich besetzt.
Dann ist die Spektraldichte definiert als

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

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



Spektraldichte

Definition (Spektraldichte)

Sei $A \in \mathbb{R}^{n \times n}$, $A^T = A$ und A spärlich besetzt.
Dann ist die Spektraldichte definiert als

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

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:

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



Problemstellung



Problemstellung

- Spektraldichte trivial wenn Eigenwerte von A bekannt

Problemstellung

- Spektraldichte trivial wenn Eigenwerte von A bekannt
- Unpraktisch wenn A sehr groß, da Berechnung teuer

Problemstellung

- Spektraldichte trivial wenn Eigenwerte von A bekannt
- Unpraktisch wenn A sehr groß, da Berechnung teuer
- Bedarf effizienter Alternativen um $\phi(t)$ abzuschätzen

Problemstellung

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

Intuitive Idee



Intuitive Idee

- Sei $\sigma(A) \subseteq I \subseteq \mathbb{R}$

Intuitive Idee

- Sei $\sigma(A) \subseteq I \subseteq \mathbb{R}$
- Sei $\{t_i\}_{i=1}^k \subseteq I$ mit $\bigcup_{i=1}^k [t_i, t_{i+1}]$

Intuitive Idee

- Sei $\sigma(A) \subseteq I \subseteq \mathbb{R}$
- Sei $\{t_i\}_{i=1}^k \subseteq I$ mit $\bigcup_{i=1}^k [t_i, t_{i+1}]$
- Zähle Eigenwerte in jedem Teilintervall

Intuitive Idee

- Sei $\sigma(A) \subseteq I \subseteq \mathbb{R}$
- Sei $\{t_i\}_{i=1}^k \subseteq I$ mit $\bigcup_{i=1}^k [t_i, t_{i+1}]$
- Zähle Eigenwerte in jedem Teilintervall
- Berechne Durchschnittswert von $\phi(t)$ in jedem Intervall mit $\nu_{[a,b]}$

Intuitive Idee

- Sei $\sigma(A) \subseteq I \subseteq \mathbb{R}$
- Sei $\{t_i\}_{i=1}^k \subseteq I$ mit $\bigcup_{i=1}^k [t_i, t_{i+1}]$
- Zähle Eigenwerte in jedem Teilintervall
- Berechne Durchschnittswert von $\phi(t)$ in jedem Intervall mit $\nu_{[a,b]}$
- Histogramme nähern sich Spektraldichte für $(t_{i+1} - t_i) \rightarrow 0$



Intuitive Idee

- Sei $\sigma(A) \subseteq I \subseteq \mathbb{R}$
- Sei $\{t_i\}_{i=1}^k \subseteq I$ mit $\bigcup_{i=1}^k [t_i, t_{i+1}]$
- Zähle Eigenwerte in jedem Teilintervall
- Berechne Durchschnittswert von $\phi(t)$ in jedem Intervall mit $\nu_{[a,b]}$
- Histogramme nähern sich Spektraldichte für $(t_{i+1} - t_i) \rightarrow 0$
- Problem: Zerlegung $A - t_i I = LDL^T$ für alle t_i zu zeitaufwendig



Intuitive Idee

- Sei $\sigma(A) \subseteq I \subseteq \mathbb{R}$
- Sei $\{t_i\}_{i=1}^k \subseteq I$ mit $\bigcup_{i=1}^k [t_i, t_{i+1}]$
- Zähle Eigenwerte in jedem Teilintervall
- Berechne Durchschnittswert von $\phi(t)$ in jedem Intervall mit $\nu_{[a,b]}$
- Histogramme nähern sich Spektraldichte für $(t_{i+1} - t_i) \rightarrow 0$
- Problem: Zerlegung $A - t_i I = LDL^T$ für alle t_i zu zeitaufwendig
- Besser: A nur mit Vektoren multiplizieren



Kernel-Polynom-Methode (KPM)



Kernel-Polynom-Methode (KPM)

- Polynomiale Annäherung der Spektraldichte



Kernel-Polynom-Methode (KPM)

- Polynomiale Annäherung der Spektraldichte
- Leitet Koeffizienten aus der Momentenmethode ab



Kernel-Polynom-Methode (KPM)

- Polynomiale Annäherung der Spektraldichte
- Leitet Koeffizienten aus der Momentenmethode ab
- Zwei eng verwandte Varianten



Kernel-Polynom-Methode (KPM)

- Polynomiale Annäherung der Spektraldichte
- Leitet Koeffizienten aus der Momentenmethode ab
- Zwei eng verwandte Varianten
- Benutzt stochastische Sampling-Methode, die auf folgendem Resultat basiert:



Theorem



Theorem

Theorem

Sei $A = A^T \in \mathbb{R}^{n \times n}$ mit Spektralzerlegung $A = U \operatorname{diag}(\lambda_1, \dots, \lambda_n) U^T$ wobei $UU^T = \mathbb{1}_n$.

Sei außerdem $\beta, v \in \mathbb{R}^n$ mit $v = U\beta$

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

$$\mathbb{E}[\beta\beta^T] = \mathbb{1}_n$$



Beweisskizze



Beweis Theorem 1.

Es gilt

$$\mathbb{E}[v] =$$

Beweis Theorem 1.

Es gilt

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

Beweis Theorem 1.

Es gilt

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

Beweis Theorem 1.

Es gilt

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

Beweis Theorem 1.

Es gilt

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

Beweis Theorem 1.

Es gilt

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

Weiterhin gilt, dass

$$\mathbb{E}[vv^T] =$$

Beweis Theorem 1.

Es gilt

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

Weiterhin gilt, dass

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

Beweis Theorem 1.

Es gilt

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

Weiterhin gilt, dass

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

Beweis Theorem 1.

Es gilt

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

Weiterhin gilt, dass

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

Beweis Theorem 1.

Es gilt

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

Weiterhin gilt, dass

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

Beweis Theorem 1.

Es gilt

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

Weiterhin gilt, dass

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

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



Resultat



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:

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



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



Transformation der Eigenwerte

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



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.



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

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



Der Pseudocode

Algorithm 1: The Kernel Polynomial Method.

Input: Real symmetric matrix A with eigenvalues between $[-1, 1]$. A set of points $\{t_i\}$ at which DOS is to be evaluated, the degree M of the expansion polynomial.

Output: Approximate DOS $\{\tilde{\phi}_M(t_i)\}$.

- 1: Set $\zeta_k = 0$ for $k = 0, \dots, M$;
 - 2: **for** $l = 1 : n_{\text{vec}}$ **do**
 - 3: Select a new random vector $v_0^{(l)}$;
 - 4: **for** $k = 0 : M$ **do**
 - 5: Compute $\zeta_k \leftarrow \zeta_k + \left(v_0^{(l)}\right)^T v_k^{(l)}$;
 - 6: Compute $v_{k+1}^{(l)}$ via the three-term recurrence $v_{k+1}^{(l)} = 2Av_k^{(l)} - v_{k-1}^{(l)}$ (for $k = 0$, $v_1^{(l)} = Av_0^{(l)}$);
 - 7: **end for**
 - 8: **end for**
 - 9: Set $\zeta_k \leftarrow \zeta_k / n_{\text{vec}}$, $\mu_k \leftarrow \frac{2 - \delta_{k0}}{n\pi} \zeta_k$ for $k = 0, 1, \dots, M$;
 - 10: Evaluate $\{\tilde{\phi}_M(t_i)\}$ using $\{\mu_k\}$ and Eq. (3.11);
-

Problemstellung

- Sei $\tilde{\phi}(t)$ eine reguläre Funktion die die Spektraldichte schätzt
- Alle Annäherungen stetige Funktionen
- $\phi(t)$ keine Funktion im eigentlichen Sinne
- $\left\| \phi(t) - \tilde{\phi}(t) \right\|_{L^p}$ nicht definiert
- 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 Distribution 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

- "Regularisierung" der δ -Distribution
- Ersetzen durch stetige und glatte Funktionen
- Z.B. Gaussche Normalverteilung mit angemessener Standardabweichung σ
- Diese Funktion $\phi_\sigma(t)$ ist dann wohldefiniert
- Berechne Fehler für $p = 1, 2$ und ∞ :

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

Diese beiden Methoden sind eng verwandt!



Der Begriff der Auflösung

- Selten Bedarf für exakte Annäherung aller Eigenwerte von A
- Oft nur Eigenwerte in $[a, b] \subset \sigma(A)$
- $b - a$ heißt *Auflösung* der Schätzung
- Je kleiner das σ in $\epsilon_2 = \left\| \phi_\sigma(t) - \tilde{\phi}(t) \right\|_p$, desto höher die Auflösung
- Für $\epsilon_1 = \sup_{g \in \mathcal{S}} \left| \langle \phi, g \rangle - \langle \tilde{\phi}, g \rangle \right|$ reguliert man über die Testfunktionen



Noch mehr Probleme mit Dirac

Betrachte

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

Definiere entsprechend

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

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

Noch mehr Probleme mit Dirac (2)

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

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

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

Hierbei sind λ_{us} und λ_{os} wie vorher und der Parameter σ die *Zielauflösung*.

Wir können nun die folgende Metrik zur Qualitätsbewertung nutzen:

$$E \left[\tilde{\phi}; \mathcal{S}(\sigma; [\lambda_{lb}, \lambda_{ub}]) \right] = \sup_{g \in \mathcal{S}(\sigma; [\lambda_{lb}, \lambda_{ub}])} \left| \langle \phi, g \rangle - \langle \tilde{\phi}, g \rangle \right| \quad (1)$$



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

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

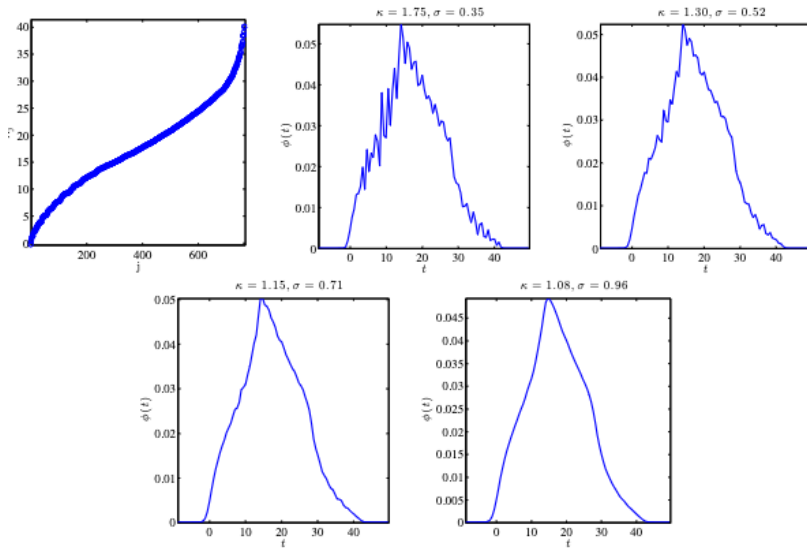
Dann ist

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

der L^{∞} -Fehler zwischen zwei wohldefinierten Funktionen



Veranschaulichung



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

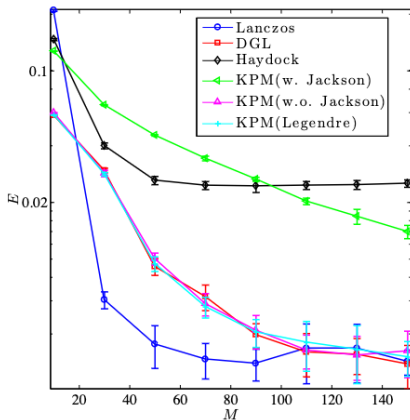
Dies ist z.B. für die Haydock-Methode relevant.

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



Festgelegte Parameter:

- $n = 750$ relativ klein
- $\sigma = 0.35$
- $n_{vec} = 100$

Die Ergebnisse sind der Mittelwert aus 10 Läufen

Festgelegte Parameter:

- $n = 750$ relativ klein
- $\sigma = 0.05$
- $n_{vec} = 100$
- $M = 100$

Die Ergebnisse sind der Mittelwert
aus 10 Läufen

