Einführung in die Grundlagen der Numerik

Institut für Numerische Simulation
Rheinische Friedrich-Wilhelms-Universität Bonn

Wintersemester 2014/2015

Was ist das Ziel einer Approximation?

Die schnelle Bestimmung einer Funktion u_N für die der Fehler

$$e_N = \|u_N - u\|_V$$

in einer bestimmten Norm $\|\cdot\|_V$ klein ist. Eine häufig verwendete Norm ist die L^2-Norm

$$\|e_N\|^2_{L^2(\Omega)} := \int_\Omega |e_N(x)|^2 \, dx.$$

Integration & Quadratur

Berechnung des bestimmten Integrals

$$\int_a^b f(x)w(x)\;dx = I(f) \qquad \int_\Omega f(x)w(x)\;dx = I(f)$$

- (a, b) bzw. Ω Integrationsgebiet; $f : \Omega \subset \mathbb{R}^d \to \mathbb{R}$ Integrand
- besondere Herausforderung wenn d groß (z.B. $d = 100$)
- Gewicht $w \in (0, \infty)^\Omega$ mit existierenden Momenten

Approximation des bestimmten Integrals

$$\tilde{I}(f) := \sum_{i=0}^{N} \omega_i f(x_i)$$

- x_i Quadraturpunkte / Stützstellen; ω_i Quadraturgewichte
- Quadraturfehler

$$|I(f) - \tilde{I}(f)| \leq \epsilon(N, f, d)$$

Konstruktion von Quadraturregeln

Approximation von f

Für ein beliebiges Approximationsverfahren mit

$$\tilde{f}(x) = \sum_{i=0}^{N} c_i \phi_i(x)$$

kann man einfach ein Quadraturverfahren über

$$\tilde{I}(f) = I(\tilde{f}) = \sum_{i=0}^{N} c_i I(\phi_i) = \sum_{i=0}^{N} c_i \Phi_i\Big|_a^b$$

formal definieren.

- Stammfunktionen Φ_i zur Basis ϕ_i
- Koeffizienten der Approximation c_i erst berechnen
- Stabilität?

Exaktheitsgrad

Falls für alle Polynome p vom Grad $\leq m$ gilt $I(p) = \tilde{I}(p)$, dann heißt die Quadraturformel exakt vom ***Grad*** m. (oder von der Ordnung $m+1$.)

Verfahren aus ALMA 2

- Newton-Cotes-Formeln (äquidistante Stützstellen)
- Extrapolation (Interpolation der numerischen Integrationsresultate)
- Romberg-Quadratur (Extrapolation mit Halbierung der zusammengesetzten Sehnentrapezregel)
- Adaptive Integration

Gauß-Quadratur

Ziel

Was ist der höchste Exaktheitsgrad einer Quadraturformel

$$\tilde{I}(f) := \sum_{0=1}^{N} \omega_i f(x_i)$$

mit $N+1$ Stützstellen x_i und $N+1$ Gewichten ω_i zu einem vorgegebenen nichtnegativen Gewicht w?

Minimaler Exaktheitsgrad

Es seien ϕ_i, $i = 0, \ldots, N$ eine Basis von Π_N. Dann hat $\tilde{I}$ Exaktheitsgrad mindestens N.

Maximaler Exaktheitsgrad

$\tilde{I}$ sei interpolatorische Quadraturformel mit mindestens Exaktheitsgrad N. Dann hat $\tilde{I}$ höchstens Exaktheitsgrad $2N+1$.

Orthogonale Polynome

Es sei $w \in (0, \infty)^{(a,b)}$ mit $-\infty \leq a \leq b \leq \infty$ sodass alle

$$c_k := \int_a^b w(x) x^k$$

für $k \in \mathbb{N}$ existieren und $c_0 > 0$. Dann

Nullstellen

haben die orthogonalen Polynome p_n zu w nur einfache Nullstellen im Inneren von (a, b).

Unter diesen Voraussetzungen:

Hauptsatz zur Gauß-Quadratur

Es existiert genau eine Quadraturformel

$$\tilde{I}(f) = \sum_{0=1}^{N} \omega_i f(x_i)$$

mit Exaktheitsgrad $2N+1$. Die Stützstellen x_i sind die $N+1$ Nullstellen des $N+1$-ten orthogonalen Polynoms p_{N+1} zum Skalarprodukt

$$\langle f, g \rangle = \int_a^b f(x)g(x)w(x)\mathrm{d}x \,.$$

Gewichte

$$\omega_i = \int_a^b w(x) \prod_{j \neq i} \left(\frac{x - x_j}{x_i - x_j} \right)^2 \mathrm{d}x > 0$$

Restglied

Für $f \in C^{2n}(a, b)$ gilt

$$(I - \tilde{I})(f) = \frac{f^{(2n)}(\eta)}{(2n)!} \int_a^b w(x) \prod_{i=0}^{N} (x - x_i)^2 \mathrm{d}x \quad \text{mit } \eta \in (a, b)$$

Konvergenz

Falls (a, b) beschränkt, konvergiert die Folge der $\tilde{I}$ mit $N \to \infty$ für jede stetige Funktion.

Beispiele

Name	(a, b)	w
Gauß-Legendre	$(-1, 1)$	1
Gauß-Chebyshev	$(-1, 1)$	$\frac{1}{\sqrt{1-x^2}}$
Gauß-Hermite	$(-\infty, \infty)$	$\exp(-x^2)$
Gauß-Laguerre	$(0, \infty)$	$\exp(-x)$

Effiziente Berechnung

Stützstellen

Lanzcos-Prozess für die orthogonalen Polynome, Dreitermrekursion

$$xp_k = \beta_{k-1}p_{k-1} + \alpha_k p_k + \beta_k p_{k+1} \quad \beta_{k-1} = \|p_k\| \quad \alpha_k = \langle xp_k, p_k \rangle$$

bezüglich w-Skalarprodukt, und damit

$$x \begin{pmatrix} p_0(x) \\ \vdots \\ \vdots \\ p_n(x) \end{pmatrix} = \underbrace{\begin{pmatrix} \alpha_0 & \beta_0 & & \\ \beta_0 & \ddots & \ddots & \\ & \ddots & \ddots & \beta_{n-1} \\ & & \beta_{n-1} & \alpha_n \end{pmatrix}}_{=:T} \underbrace{\begin{pmatrix} p_0(x) \\ \vdots \\ \vdots \\ p_n(x) \end{pmatrix}}_{=:p(x)} + \beta_n \begin{pmatrix} 0 \\ \ddots \\ 0 \\ p_{n+1}(x) \end{pmatrix}$$

sind die Nullstellen x_i von p_{n+1} die Eigenwerte von T zu den Eigenvektoren $p(x_i)$ und mit QR-Verfahren für Tridiagonalmatrizen effizient berechenbar.

Gewichte

LGS für Gewichte ω_i

$$\underbrace{\begin{pmatrix} p_0(x_0) & p_0(x_1) & \dots & p_0(x_n) \\ p_1(x_0) & p_1(x_1) & \dots & p_1(x_n) \\ \vdots & \vdots & \ddots & \vdots \\ p_n(x_0) & p_n(x_1) & \dots & p_n(x_n) \end{pmatrix}}_{=:P} \underbrace{\begin{pmatrix} \omega_0 \\ \vdots \\ \vdots \\ \omega_n \end{pmatrix}}_{=:g} = \begin{pmatrix} \sqrt{b-a} \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \sqrt{b-a} e_1$$

QR-Verfahren: $P = QD$, dann folgt mit

$$(Q_{1,i})_i = ((PD^{-1})_{1,i})_i = \frac{1}{\sqrt{b-a}}(1, \dots, 1)$$

dass

$$g = \sqrt{b-a} D^{-1} Q^T e_1 = \dots = (b-a)(Q_{1,1}^2, \dots, Q_{1,N+1}^2)$$

Gauss-Christoffel-Quadratur

einzelne Teile, Gautschi 1968, Golub, Welsh, 1969

Algorithmus

1. Lanczos-Prozess: T, nur α_i und β_i zu speichern.
2. Q, $(x_i)_i$ aus QR-Verfahren zu T
3. $(\omega_i)_i = (b - a)(Q_{1,i}^2)_i$

Hierarchische Quadratur auf $[0, 1]$

- Setze $\tilde{I} := \frac{1}{2}(f(0) + f(1))$
- Setze $(x, h) := (\frac{1}{2}, \frac{1}{2})$.
- Berechne den hierarchischen Überschuss

$$H(x, h) = f(x) - \frac{1}{2}\left(f(x-h) + f(x+h)\right)$$

- Falls $|H(x, h)| > \epsilon$:

$$\tilde{I} = \tilde{I} + hH(x, h),$$

  Und betrachte

$$(x, h) = (x - \frac{h}{2}, \frac{h}{2}) \qquad (x, h) = (x + \frac{h}{2}, \frac{h}{2})$$

Tensorprodukt-Quadratur

- Grundlage: mehrdimensionale Interpolation, z.B. durch Interpolationspolynom p

$$\int_{a_1}^{b_1} \cdots \int_{a_d}^{b_d} f = I(f) \approx \tilde{I}(f) = \int_{a_1}^{b_1} \cdots \int_{a_d}^{b_d} p \quad \text{mit}$$

$$p \in \Pi_n^d = \{p(x_1, \ldots, x_d) = \sum_{0 \leq \sum_{i=1}^d k_i \leq n} a_{k_1,\ldots,k_d} \prod_{i=1}^d x_i^{k_i}\}$$

- Nicht für beliebige Stützstellen wohldefiniert. ok, z.B. für kartesische Produkte von 1D Stützstellen $x_{i,j} \in (a_i, b_i), j = 0, \ldots, n_i, \forall j_1 \neq j_2 : x_{i,j_1} \neq x_{i,j_2}$ mit Gewichten $\omega_{i,j}$

$$\tilde{I}(f) = \left(\prod_{i=1}^d \tilde{I}_i \right)(f) = \sum_{j_1=0}^{n_1} \omega_{1,j_1} \cdots \sum_{j_d=0}^{n_d} \omega_{d,j} f(x_{1,j_1}, \ldots, x_{d,j_d}) \ .$$

- Für verschiedene Richtungen andere Quadraturformeln verwendbar
- Mit hinreichender Beschränkheit der iterierten Integrale/Quadraturformeln I_i mit Stützstellen $x_{i,j}, \omega_{i,j}$

$$
\Big| \int_{a_i}^{b_i} f(y_1, \ldots; x; \ldots, y_d) \mathrm{d}x - \underbrace{\sum_{j=0}^{n_i} \omega_{i,j} f(y_1, \ldots; x_{i,j}; \ldots, y_d)}_{=\tilde{I}_i(x \mapsto f(y_0, \ldots; x; \ldots, y_d)} \Big|
\leq (b_i - a_i) C_i(f) h_i^{k_i}
$$

für alle $(y_0, \ldots, y_d) \in \prod_{i=1}^d [a_i, b_i]$ gilt am Ende

$$
\left| I(f) - \tilde{I}(f) \right| \leq \left(\prod_{i=1}^{d} (b_i - a_i) \right) \sum_{i=1}^{d} C_i(f) h_i^{k_i}
$$

- Kann wie in $1D$ auch Stützstellen und Gewichte zur exakten Integration von Polynomen aus Gleichungssystem konstruieren.

Standardgebiete

- Für andere Gebiete (nicht Rechtecke, Quader...) zweckmäßiger, exakte Integration von Monomen

$$\prod_{i=1}^{d} x_i^{k_i}, 0 \le \sum_{i=1}^{d} k_i \le l$$

  zu fordern
- Satz von Tschakalov, 1957: Es gibt für beliebige Gebiete solche Integrationsformeln mit positiven Gewichten und allen Stützstellen im Gebiet.

Beispiele

Integration exakt für Monome $1, x, y, x^2, xy, y^2$.

Dreieck mit Ecken $(0,0), (0,1), (1,0)$

$$\tilde{I}(f) = \frac{1}{6}\left(f(\frac{1}{2},0) + f(0,\frac{1}{2}) + f(\frac{1}{2},\frac{1}{2})\right)$$

Tetraeder mit Ecken $(0,0,0), (1,0,0), (0,1,0), (0,0,1)$

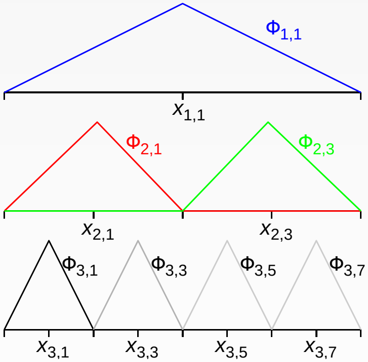
$$r = \frac{5-\sqrt{5}}{20} \qquad s = \frac{5+3\sqrt{5}}{20}$$

$$\tilde{I}(f) = \frac{1}{24}\left(f(r,r,r) + f(s,r,r) + f(r,s,r) + f(r,r,s)\right)$$

Monte-Carlo

- Für große Dimensionen Integration kostspielig, da n^d Stützstellen auf vollem Gitter
- Integrale sind im Wesentlichen Mittelwerte, also aus Stochastik mit Gesetz der großen Zahlen: Es sei w eine Wahrscheinlichkeitsdichte ($\int_a^b w = 1$) und x_i nach w verteilt dann
$$P\left(\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} f(x_i) = \int_a^b f(x)\mathrm{d}w(x)\right) = 1$$
- Schwierigkeit typischerweise Ziehen der Zufallszahlen und langsame Konvergenz
- Analog im Mehrdimensionalen

Hierarchische Basen

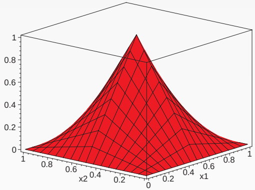


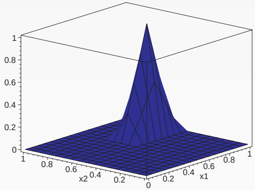
hierarchische Basis in 1D

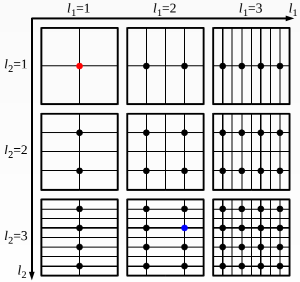
Notation

- Gebiet $\Omega = [0,1]^d$
- Gitter beschrieben durch Multiindex $\underline{l} = (l_1, \ldots, l_d) \in \mathbb{N}^d$
- Gitterweite $\underline{h}_{\underline{l}} = (h_1, \ldots, h_d) = (2^{-l_1}, \ldots, 2^{-l_d})$
- $|\underline{l}|_1 = l_1 + \ldots + l_d$
- $\Phi_{\underline{l},\underline{i}}(\underline{x}) = \prod_{j=1}^{d} \Phi_{l_j,i_j}(x_j)$

2D Beispiel, $\Phi_{(1,1),(1,1)}$ and $\Phi_{(2,3),(3,5)}$

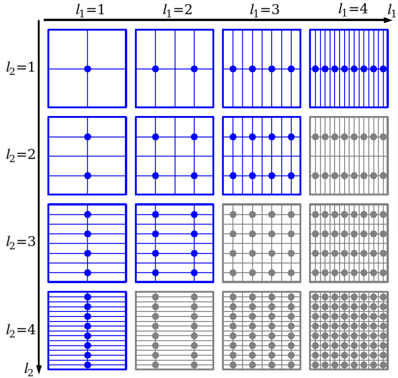




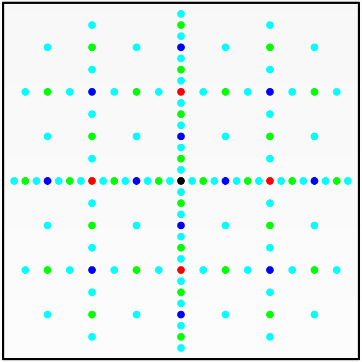


Dünne Gitter

Kosten/Nutzen also Konstant auf “Diagonalen” (i.e. konstante $|\underline{l}|_1$).
Also “optimales” Gitter.



Punkte/Stützstellen gleicher "Wichtigkeit" in gleicher Farbe



Die Bilder in Formeln:

$$V_n^1 := \bigoplus_{|\underline{l}|_1 \leq n+d-1} W_{\underline{l}}$$

wobei

$$W_{\underline{l}} := \langle \Phi_{\underline{l},\underline{i}} \rangle_{\underline{i} \in \{\underline{1} \leq \underline{i} \leq 2^{\underline{l}},\ \text{all } i_j \text{ are uneven}\}}$$

Jedes $v \in V_n^1$ hat eindeutige Darstellung

$$v = \sum_{|\underline{l}|_1 \leq n+d-1} w_{\underline{l}}$$

with $w_{\underline{l}} \in W_{\underline{l}}$.

Fehlerabschätzungen

Interpolation eines $u : [0,1]^d \to \mathbb{R}$ with $v \in V_n^1$ liefert für sogenannte ***hierarchische Überschüsse*** (Quadraturfehler wie in $1D$ aus Interpolationsfehler ableitbar)

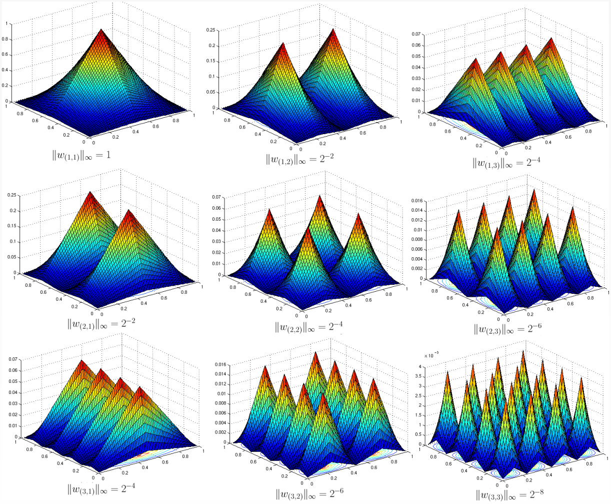
$$\|w_{\underline{l}}\|_\infty \leq 2^{-2|\underline{l}|_1 - d} \, \| \frac{\partial^{2d} u}{\partial x_1^2 \cdot \ldots \cdot \partial x_d^2} \|_\infty$$

$$\|w_{\underline{l}}\|_2 \leq 2^{-2|\underline{l}|_1} \; 3^{-d} \, \| \frac{\partial^{2d} u}{\partial x_1^2 \cdot \ldots \cdot \partial x_d^2} \|_2$$

Im allgemeinen nur Beweise nur für einfache Probleme vorhanden!

Beispiel

Interpolation von $16x \cdot (x-1) \cdot y \cdot (y-1)$ auf $[0,1]^2$:



Kombitechnik

- Rechnen mit hierarchischer Basis von V_n^1 problematisch, multi-rekursiv
- Können aber $v_n \in V_n^1, n > 1$ auch zusammensetzen aus

$$v_n \approx \sum_{q=0}^{d-1} (-1)^q \binom{d-1}{q} \sum_{\underline{l}=n+d-1-q} v_{\underline{l}}$$

wobei

$$v_{\underline{l}} \in V_{\underline{l}} = \sum_{\underline{m} \leq \underline{l}} W_{\underline{m}}$$

die Approximanden in den vollen, aber anisotropen (unterschiedliche h in verschiedene Dimensionen) Räumen $V_{\underline{l}}$ and die gesuchte Funktion u ist.

- in 2D

$$v_n \approx \sum_{\underline{l}=n+1} u_{\underline{l}} - \sum_{\underline{l}=n} u_{\underline{l}}$$

- Konvergenzbeweis nur für einfachste Probleme vorhanden!

Volle Gitter

- Kosten: $(2^n - 1)^d \in O(2^{nd}) = O(h^{-d})$
- Genauigkeit bei Interpolationsproblemen: $||u - u_h||_2 \in O(2^{-2n}) = O(h^2)$

Dünne Gitter

- Kosten: $dim V_n^1 = \cdots = \sum_{k=0}^{n-1} \binom{k+d-1}{d-1} 2^k = \cdots = (-1)^d + 2^n \sum_{k=0}^{d-1} \binom{n+d-1}{k} (-2)^{d-1-k} = \cdots \in O(2^n n^{d-1}) = O(h^{-1} \log(h^{-1})^{d-1})$
- Genauigkeit: $||u - u_h||_2 \in O(2^{-2n}) = O(h^2)$

Konkreter

$d = 2$

n	2	4	8	10	20	30
volles Gitter	9	225	$\sim 10^5$	$\sim 10^6$	$\sim 10^{12}$	$\sim 10^{18}$
dünnes Gitter	5	49	1793	$\sim 10^5$	$\sim 10^7$	$\sim 10^{10}$

$d = 3$

n	2	4	8	10	20	30
volles Gitter	27	3375	$\sim 10^8$	$\sim 10^9$	$\sim 10^{18}$	$\sim 10^{27}$
dünnes Gitter	7	111	7423	47103	$\sim 10^8$	$\sim 10^{12}$