Institut für Numerische Simulation
Rheinische Friedrich-Wilhelms-Universität Bonn

universitätbonn

Übungen zu Einführung in die Numerische Mathematik (V2E2)

Sommersemester 2016

Prof. Dr. Martin Rumpf — Pascal Huber — Sascha Tölkes

Übungsblatt 5 — Abgabe: 31.05.2016

Aufgabe 15 (Fehlerabschätzung für Ableitungen) — 4 Punkte

Sei $T \subset \mathbb{R}^d$ ein nicht-degeneriertes Simplex mit Durchmesser $h > 0$ und sei $f \in C^{k+1}(T)$. Verfahren Sie ähnlich zum Beweis von Satz 2.13 aus der Vorlesung und zeigen Sie, dass für die Ableitungen von f die Fehlerabschätzungen

$$\|\partial_x^\beta (f - I_h(f))\|_{\infty,T} \leq C\, h^{k+1-|\beta|} \|D_x^{k+1} f\|_{\infty,T}$$

gelten, wobei $I_h(f) \in \mathcal{P}_k^d$ das Interpolationpolynom von f auf dem Simplex T und $\beta = (\beta_1, \ldots, \beta_d)$ ein Multiindex mit $|\beta| \leq k$ ist.

Hinweise:

- *Falls Ihnen Schritt (ii) des Beweises Schwierigkeiten macht, können Sie stattdessen versuchen, die Aussage für $f \in C^{k+|\beta|+1}(T)$ zu zeigen.*
- *Zeigen Sie für Schritt (iv) des Beweises, dass eine Ungleichung der Form*

  $$\|\partial_x^\beta (f - I_h(f))\|_{\infty,T} \leq C_\beta(A_T^{-1}) \max_{|\gamma|=|\beta|} \|\partial_{\hat{x}}^\gamma (\hat{f} - \widehat{I}(\hat{f}))\|_{\infty,\widehat{T}}$$

  gilt und finden Sie eine passende Abschätzung für $C_\beta(A_T^{-1})$.

Lösung

Wir gehen wie im Beweis von Satz 2.13 vor.

(i) *Stabilität*: Es bezeichne $\widehat{T}$ den d-dimensionalen Einheitssimplex und es sei $\gamma = (\gamma_1, \ldots, \gamma_d)$ ein beliebiger Multiindex mit $|\gamma| \leq k$. Dann gilt für das Interpolationspolynom $\widehat{I}(\hat{f})$ einer Funktion $\hat{f} \in C^{k+1}(\widehat{T})$

$$\|\partial_{\hat{x}}^\gamma \widehat{I}(\hat{f})\|_{\infty,\widehat{T}} \leq \binom{d+k}{k} \max_\alpha \|\partial_{\hat{x}}^\gamma L_\alpha\|_{\infty,\widehat{T}}\, \|\hat{f}\|_{\infty,\widehat{T}} \leq C(k)\, \|\hat{f}\|_{\infty,\widehat{T}}.$$

Hierbei sind L_α die Lagrange-Basispolynome.

(ii) *Stabilitätsargument*: Für alle Polynome $\hat{q} \in \mathcal{P}_k^d$ gilt die Abschätzung

$$\|\partial_{\hat{x}}^{\gamma}(\hat{f} - \widehat{I}(\hat{f}))\|_{\infty,\widehat{T}} \leq \|\partial_{\hat{x}}^{\gamma}(\hat{f} - \hat{q})\|_{\infty,\widehat{T}} + \|\partial_{\hat{x}}^{\gamma}(\widehat{I}(\hat{f}) - \widehat{I}(\hat{q}))\|_{\infty,\widehat{T}}. \tag{1}$$

Leider kann hier nicht wie in der Vorlesung für $\hat{q}$ direkt das Tensorpolynom von $\hat{f}$ verwendet werden. Zwar könnten wir nach Anwendung der Stabilität aus Schritt (i) auf den zweiten Term, diesen problemlos gegen $\|D_{\hat{x}}^{k+1}\hat{f}$ abschätzen, über den ersten Term hätten wir dann allerdings wegen der Ableitung $\partial_{\hat{x}}^{\gamma}$ keine Kontrolle.
Wir lösen dieses Problem, in dem wir einen weiteren Term hinzufügen. Sei $\hat{p} \in \mathcal{P}_k^d$ beliebig. Dann folgt wieder mit der Dreiecksungleichung

$$\|\partial_{\hat{x}}^{\gamma}(\hat{f} - \widehat{I}(\hat{f}))\|_{\infty,\widehat{T}} \leq \|\partial_{\hat{x}}^{\gamma}\hat{f} - \hat{p}\|_{\infty,\widehat{T}} + \|\hat{p} - \partial_{\hat{x}}^{\gamma}\hat{q}\|_{\infty,\widehat{T}} + \|\partial_{\hat{x}}^{\gamma}\widehat{I}(\hat{f} - \hat{q})\|_{\infty,\widehat{T}}.$$

Nach Anwendung der Stabilität ergibt dies

$$\|\partial_{\hat{x}}^{\gamma}(\hat{f} - \widehat{I}(\hat{f}))\|_{\infty,\widehat{T}} \leq \|\partial_{\hat{x}}^{\gamma}\hat{f} - \hat{p}\|_{\infty,\widehat{T}} + \|\hat{p} - \partial_{\hat{x}}^{\gamma}\hat{q}\|_{\infty,\widehat{T}} + C(k)\|\hat{f} - \hat{q}\|_{\infty,\widehat{T}}.$$

Nun wählen wir $\hat{p}$ als das Taylorpolynom von $\partial_{\hat{x}}^{\gamma}\hat{f}$ vom Grad $(k - |\gamma|)$ und $\hat{q}$ als das Taylorpolynom von $\hat{f}$ vom Grad k. Damit können der erste und letzte Term auf der rechten Seite gegen $C\|D_{\hat{x}}^{k+1}\hat{f}\|_{\infty,\widehat{T}}$ abgeschätzt werden. Der verbliebene Term verschwindet, da nach Wahl der Polynome $\hat{p} = \partial_{\hat{x}}^{\gamma}\hat{q}$. Also erhalten wir insgesamt

$$\|\partial_{\hat{x}}^{\gamma}(\hat{f} - \widehat{I}(\hat{f}))\|_{\infty,\widehat{T}} \leq C\|D_{\hat{x}}^{k+1}\hat{f}\|_{\infty,\widehat{T}}$$

(iii) *Transformation auf T*: Analog zum Beweis aus der Vorlesung erhalten wir

$$|||A_T||| \leq \frac{h}{\rho(\widehat{T})} \qquad \text{und} \qquad \left|\left|\left|A_T^{-1}\right|\right|\right| \leq \frac{h(\widehat{T})}{\rho},$$

wobei $\rho, \rho\widehat{T}$ die jeweiligen Innenkugeldurchmesser bezeichnen. Da wir hier nur von einem Simplex ausgehen, können wir ρ durch h abschätzen, sodass gilt:

$$|||A_T||| \leq c_1 h \qquad \text{und} \qquad \left|\left|\left|A_T^{-1}\right|\right|\right| \leq \frac{c_2}{h}.$$

(Dies trifft insbesondere auf reguläre Familien von Triangulierungen zu.)

(iv) *Transformation der Abschätzung*: Wir definieren $\hat{f}(\hat{x}) := f(A_T\hat{x} + b_T)$, d.h. es gilt $f(x) = \hat{f}(A_T^{-1}(x - b_T))$. Die Kettenregel liefert also $D_x f(x) = D_{\hat{x}}\hat{f}(\hat{x})A_T^{-1}$, wobei $\hat{x} = A_T^{-1}(x - b_T)$. Eine iterierte Anwendung dieser Beziehung liefert demnach

$$\|\partial_x^{\beta}(f - I_h(f))\|_{\infty,T} \leq C\left|\left|\left|A_T^{-1}\right|\right|\right|^{|\beta|} \max_{|\gamma|=|\beta|}\|\partial_{\hat{x}}^{\gamma}(\hat{f} - \widehat{I}(\hat{f}))\|_{\infty,\widehat{T}}.$$

Nun wenden wir Teil (ii) an und erhalten

$$\|\partial_x^{\beta}(f - I_h(f))\|_{\infty,T} \leq C\left|\left|\left|A_T^{-1}\right|\right|\right|^{|\beta|}\|D_{\hat{x}}^{k+1}\hat{f}\|_{\infty,\widehat{T}}.$$

Wie in der Vorlesung kann das nun wie folgt abgeschätzt werden

$$\|\partial_x^\beta (f - I_h(f))\|_{\infty,T} \le C \left|\left|\left| A_T^{-1} \right|\right|\right|^{|\beta|} |||A_T|||^{k+1} \|D_x^{k+1} f\|_{\infty,T}.$$

Schließlich wenden wir die Abschätzungen der Operatornormen aus Schritt (iii) an und erhalten die gewünschte Aussage:

$$\|\partial_x^\beta (f - I_h(f))\|_{\infty,T} \le C h^{-|\beta|} h^{k+1} \|D_x^{k+1} f\|_{\infty,T}.$$

Bemerkung: Falls wir in (1) annehmen, dass $\hat{f} \in C^{k+|\gamma|+1}(\widehat{T})$, können wir direkt $\hat{q}$ als Taylorpolynom von $\hat{f}$ vom Grad k wählen. Der zweite Term wird dann wie bereits beschrieben abgeschätzt. Für den ersten Term nutzen wir die Taylorformel und können den Restterm wegen der erhöhten Regularität abschätzen, sodass wir die folgende Ungleichung erhalten:

$$\|\partial_{\hat{x}}^\gamma (\hat{f} - \widehat{I}(\hat{f}))\|_{\infty,\widehat{T}} \le C \|D_{\hat{x}}^{k+|\gamma|+1} \hat{f}\|_{\infty,\widehat{T}} + C(k) \|D_{\hat{x}}^{k+1} \hat{f}\|_{\infty,\widehat{T}}$$

Wie oben erhalten wir durch Hin- und Rücktransformation die gewünschten h-Potenzen:

$$\|\partial_x^\beta (f - I_h(f))\|_{\infty,T} \le C h^{-|\beta|} \left(h^{k+1} \|D_x^{k+1} f\|_{\infty,T} + h^{k+|\beta|+1} \|D_x^{k+|\beta|+1} f\|_{\infty,T} \right).$$

Da wir üblicherweise von kleinen $h > 0$ ausgehen, können wir den zweiten Summanden in der Klammer auf kosten einer etwas größeren Konstanten C vernachlässigen und erhalten auch in diesem Fall die Aussage.

Aufgabe 16 (Diskrete Energieminimierung) 4 Punkte

Gesucht seien Minimierer der Energie

$$E[u] = \int_0^1 |u'|^2 - u \, dx$$

mit $u \in C^1([0,1])$ und Randwerten $u(0) = u(1) = 0$. Dazu betrachten wir folgende Approximation: Zu $h = \frac{1}{N}$ definieren wir einen endlich dimensionalen Raum von Funktionen

$$\mathcal{V}_h = \Big\{ u_h : [0,1] \to \mathbb{R} \Big| \; u_h|_{[ih,(i+1)h]} \text{ ist affin für } i = 0, \cdots, N-1, \\ u_h \in C^0([0,1]), u_h(0) = u_h(1) = 0 \Big\}$$

und eine Approximation der Energie

$$E_h[u_h] = \sum_{i=0}^{N-1} \int_{ih}^{(i+1)h} |u_h'|^2 - u_h \, dx.$$

(i) Geben Sie eine Basis $\{\varphi_h^i\}_{i=1,\cdots,N-1}$ von $\mathcal{V}_h$ an.

(ii) Für $u_h(x) = \sum_{i=0}^{N-1} u_h^i \varphi_h^i(x)$, $u_h^i \in \mathbb{R}$ betrachten Sie die Abbildung

$$\bar{E}_h : \mathbb{R}^{N-1} \to \mathbb{R};\ (u_h^i)_{i=1,\cdots,N-1} \to E_h[u_h(x)].$$

Berechnen Sie $\frac{\partial \bar{E}_h}{\partial u_h^i}$.

(iii) Stellen Sie die notwendige Bedingung für ein Minimum von E_h auf $\mathcal{V}_h$ als Gleichungssystem dar.

Bemerkung: Dieses Gleichungssystem sollte ähnlich eines Ihnen aus der Vorlesung bekannten sein.

Lösung

(i) $\varphi_h^i(jh) = \delta_{ij}$ legt die Werte an den Intervallgrenzen fest und definiert damit für jedes $i \in \{1, N-1\}$ genau eine Funktion in $\mathcal{V}_h$. Betrachtet man $u_h \in \mathcal{V}_h$, so ist $u_h(x) = \sum_{i=1}^{N-1} u_h(ih)\varphi_h^i(x)$ die eindeutige Darstellung von u_h, also bilden die φ_h^i eine Basis.

(ii) Nun kann man $\bar{E}_h$ explizit angeben. Dazu beachte man, dass der Träger der Basisfunktionen sich jeweils auf zwei Intervalle beschränkt.

$$\begin{aligned}
\bar{E}_h[u_h] &= \sum_{i=0}^{N-1} \int_{ih}^{(i+1)h} |u_h'|^2 - u_h \, dx \\
&= \sum_{i=0}^{N-1} \int_{ih}^{(i+1)h} |\sum_{j=1}^{N-1} u_h^j (\varphi_h^j)'(x)|^2 - \sum_{j=1}^{N-1} u_h^j \varphi_h^j(x) \, dx \\
&= \sum_{i=0}^{N-1} \int_{ih}^{(i+1)h} (\sum_{j=1}^{N-1} u_h^j (\varphi_h^j)'(x))(\sum_{k=1}^{N-1} u_h^k (\varphi_h^k)'(x)) - \sum_{j=1}^{N-1} u_h^j \varphi_h^j(x) \, dx \\
&= \sum_{j=1}^{N-1} \sum_{k=j-1}^{j+1} \int_{(j-1)h}^{(j+1)h} u_h^j (\varphi_h^j)'(x) u_h^k (\varphi_h^k)'(x) \, dx - \sum_{j=1}^{N-1} u_h^j \int_{(j-1)h}^{(j+1)h} \varphi_h^j(x) \, dx
\end{aligned}$$

Nun berechnet man die Integrale der Basisfunktion und der Produkte von Ableitungen:

$$\begin{aligned}
\int_{(j-1)h}^{(j+1)h} \varphi_h^j(x) \, dx &= h \\
\int_{(j-1)h}^{(j+1)h} (\varphi_h^j)'(x)^2 \, dx &= \frac{2}{h} \\
\int_{jh}^{(j+1)h} (\varphi_h^j)'(x)(\varphi_h^{j+1})'(x) \, dx &= -\frac{1}{h}
\end{aligned}$$

Damit ergibt sich

$$\bar{E}_h[u_h] = \sum_{i,j=1}^{N-1} L_{ij} u_h^i u_h^j - \sum_{i,j=1}^{N-1} h u_h^i \,,$$

$$L_{ij} = \begin{cases} 2/h & : i = j\,, \\ -1/h & : i = j \pm 1\,, \\ 0 & : \text{ sonst.} \end{cases}$$

$$\frac{\partial \bar{E}_h}{\partial u_h^k}[u_h] = \sum_{i=1}^{N-1} L_{ik} u_h^i + \sum_{j=1}^{N-1} L_{kj} u_h^j - h$$

$$= 2 \sum_{i=1}^{N-1} L_{ik} u_h^i - h$$

(iii) Wir erhalten das Gleichungssystem

$$\sum_{i=1}^{N-1} L_{ik} u_h^i = \frac{1}{h}(-u_h^{i-1} + 2u_h^i - u_h^{i+1}) = \frac{h}{2}$$

$$\text{bzw.} \quad \frac{2}{h^2}(-u_h^{i-1} + 2u_h^i - u_h^{i+1}) = 1\,,$$

analog dem entsprechenden Gleichungssystem, wenn man die notwendige Bedingung zur Minimierung von E mit Differenzenquotienten diskretisiert.

Aufgabe 17 (Tensorproduktinterpolation im $\mathbb{R}^m$) 4 Punkte

Sei $\Omega = (a,b)^m$ und $t_0, \dots t_n \in [a,b]$ seien paarweise verschieden. Weiterhin sei

$$\mathcal{N} := \{x^\alpha | \alpha = (\alpha_1, \dots \alpha_m), \alpha_i \in 0, \dots n, x^\alpha = (t_{\alpha_1}, \dots t_{\alpha_m})\}\,.$$

Sei $\mathcal{P} := \underbrace{\mathcal{P}_n \otimes \cdots \otimes \mathcal{P}_n}_{m \text{ Faktoren}}$, d. h.

$$\mathcal{P} = \{p : \mathbb{R}^m \to \mathbb{R} \,|\, t \mapsto p(x_1, \dots x_{i-1}, t, x_{i+1}, \dots x_m) \in \mathcal{P}_n$$
$$\text{für } 1 \le i \le m \text{ und } x_1, \dots x_{i-1}, x_{i+1}, \dots x_m \in \mathbb{R}\ \}\,.$$

(i) Geben Sie eine Lagrange-Basis von $\mathcal{P}$ an.

(ii) Zeigen Sie, dass zu $f \in C^0(\bar{\Omega})$ genau ein $p \in \mathcal{P}$ existiert, das $f(x) = p(x)$ für alle $x \in \mathcal{N}$ erfüllt.

(iii) Geben Sie die Lagrange-Basis für $m = 3$, $a = t_0 = 0$, $b = t_1 = 1$ und $n = 1$ explizit an.

Lösung

(i)

$$L_\alpha(x) = \prod_{i=1}^{m} \prod_{j \neq \alpha_i} \frac{x_i - t_j^i}{t_{\alpha_i}^i - t_j}$$

(ii) Analog zum 1D Fall.

(iii)

$$\{(1-x)(1-y)(1-z), (1-x)(1-y)z, (1-x)y(1-z),$$
$$x(1-y)(1-z), x(1-y)z, xy(1-z), (1-x)yz, xyz\}$$

Aufgabe 18 (Basiskonstruktion auf Prisma) **4 Punkte**

Gegeben sei ein Prisma, konstruieren Sie über einen Tensorproduktansatz (bzgl. eines Intervalls und eines Dreiecks) eine Basis eines bilinearen und eines biquadratischen Polynomraums. Verwenden Sie als Koordinaten $(\mu_0, \mu_1, \lambda_0, \lambda_1, \lambda_2)$, wobei μ_i bzw. λ_i die baryzentrischen Koordinaten auf dem Intervall als 1-Simplex bzw. dem Dreieck als 2-Simplex bezeichnen.

Lösung

- **Lineare Basen:**

  Intervall: $B_0(\mu) = \mu_0$, $B_1(\mu) = \mu_1$
  Dreieck: $B_0(\lambda) = \lambda_0$, $B_1(\lambda) = \lambda_1$, $B_2(\lambda) = \lambda_2$

- **Bilineare Basis, Prisma:**

  $B_{ij}(\mu, \lambda) = \mu_i \lambda_j$ für $i = 0, 1$, $j = 0, 1, 2$

- **Quadratische Basen:**

  Intervall:

$$B_0(\mu) = 4\mu_0\mu_1$$
$$B_1(\mu) = 2\mu_0^2 - \mu_0$$
$$B_2(\mu) = 2\mu_1^2 - \mu_1$$

  Dreieck:

$$B_0(\lambda) = 4\lambda_0\lambda_1$$
$$B_1(\lambda) = 4\lambda_1\lambda_2$$
$$B_2(\lambda) = 4\lambda_0\lambda_2 B_0(\lambda) = 4\lambda_0\lambda_1$$
$$B_3(\lambda) = 2\lambda_0^2 - \lambda_0$$
$$B_4(\lambda) = 2\lambda_1^2 - \lambda_1$$
$$B_5(\lambda) = 2\lambda_2^2 - \lambda_2$$

- **Biquadratische Basis: Prisma**

$$B_{ij}(\mu,\lambda) = (2\mu_i^2 - \mu_i)(2\lambda_j^2 - \lambda_j) \text{ für } i = 0,1 \text{ und } j = 0,1,2$$
$$\tilde{B}_{ij}(\mu,\lambda) = (2\mu_i^2 - \mu_i)(4\lambda_j - \lambda_{(j+1) mod\, 3}) \text{ für } i = 0,1 \text{ und } j = 0,1,2$$
$$B_j(\mu,\lambda) = (4\mu_0\mu_1)(2\lambda_j^2 - \lambda_j) \text{ für } j = 0,1,2$$
$$\tilde{B}_j(\mu,\lambda) = (4\mu_0\mu_1)4\lambda_j - \lambda_{(j+1) mod\, 3}) \text{ für } j = 0,1,2$$