

Prof. Dr. A. Klawonn  
J. Knepper, M. Sc.

26. Juni 2019

## 12. Übung zu Wissenschaftliches Rechnen II

**Aufgabe 1:** Beachten Sie den Anhang.

**Aufgabe 2:** (3 Punkte)

Zeigen Sie, dass die Massenmatrix spektraläquivalent zum Schurkomplement  $BA^{-1}B^T$  des Sattelpunktproblems ohne Strafterm ist. Es gelten die üblichen Voraussetzungen an  $A$  und  $B$ .

**Aufgabe 3:** (5 Punkte)

Für ein polygonales Gebiet  $\Omega \subset \mathbb{R}^d$ ,  $d = 1, 2, 3$ , sei ein reguläres Finite-Elemente-Gitter  $(\tau_h)_h$  mit stückweisen polynomialen nodalen Basisfunktionen  $\{\varphi_i\}_{i=1}^n$  gegeben. Zeigen Sie, dass die Massenmatrix  $M$  spektral äquivalent zu ihrer Diagonalen  $M_D$  ist. Zeigen Sie also, dass Konstanten  $c_L$  und  $c_R$  existieren, die unabhängig von  $h_T := \text{diam}(T)$  für  $T \in \tau_h$  sind, sodass

$$c_L \leq \frac{x^T M x}{x^T M_D x} \leq c_R. \quad (1)$$

Sie dürfen ohne Beweis nutzen, dass Konstanten  $d_1$  und  $d_2$  unabhängig von  $h_T$  existieren, sodass für  $T \in \tau_h$

$$d_1 h_T^d \leq \|\varphi_i\|_{L^2(T)} \leq d_2 h_T^d \quad (2)$$

erfüllt ist; cf. [1]. Sie können ebenfalls die Aussage ohne Beweis nutzen, dass Konstanten  $c_1$  und  $c_2$  unabhängig von  $h_T$  existieren, sodass

$$c_1 \|v\|_{L^2(\Omega)}^2 \leq \underbrace{\sum_{T \in \tau_h} h_T^d \sum_{a_i \in T} |v(a_i)|^2}_{=: g(v)} \leq c_2 \|v\|_{L^2(\Omega)}^2 \quad \forall v \in V^h, \quad (3)$$

wobei  $a_i \in T$  die Knoten der nodalen Basis sind und  $V^h$  der Finite-Elemente-Raum; cf. [2].

*Bemerkung:*  $M_D$  nennt man auch *konzentrierte Massenmatrix* bzw. *lumped mass matrix*.

**Aufgabe 4:** (2 Punkte)

Seien  $A^h, B^h, C^h$  symmetrisch positiv-definite FEM-Matrizen. Zeigen Sie: Sind  $C^h$  und  $B^h$  spektral äquivalent und  $B^h$  und  $A^h$ , dann auch  $C^h$  und  $A^h$ .

---

<sup>1</sup>Lemma B.5, S. 373, Andrea Toselli, Olof Widlund, 2005, *Domain Decomposition Methods – Algorithms and Theory*

<sup>2</sup>Aufgabe 2, Übungsblatt 11, *Numerik partieller Differentialgleichung I*, SS2018

**Aufgabe 5:** (3 + 5 + 1 + 2 = 11 Punkte)

Es sei die Sattelpunktmatrix

$$K := \begin{pmatrix} A & B^T \\ B & 0 \end{pmatrix}$$

gegeben, wobei die zu  $A$  gehörige Bilinearform  $a: X \times X \rightarrow \mathbb{R}$  symmetrisch,  $X$ -elliptisch und stetig sei. Die zu  $B$  gehörige Bilinearform  $b: X \times M \rightarrow \mathbb{R}$  sei stetig und erfülle die inf-sup-Bedingung. Wir betrachten den Rechtsvorkonditionierer

$$B_U^{-1} = \begin{pmatrix} A & B^T \\ 0 & -C \end{pmatrix}^{-1},$$

wobei  $C$  symmetrisch und positiv-definit sei.

- a) Was ist aus theoretischer Sicht eine gute Wahl für  $C$ ? D.h. wie muss  $C$  gewählt werden, damit  $KB_U^{-1}$  im GMRES-Verfahren zu einer niedrigen Anzahl Iterationen führt?

Überlegen Sie sich zunächst intuitiv, welches  $C$  zu einer „einfachen“ Matrix  $KB_U^{-1}$  führt. Bedenken Sie, dass  $C$  die Voraussetzungen erfüllen muss.

- b) Zeigen Sie anschließend, dass bei obiger Wahl von  $C$  das GMRES-Verfahren in 2 Schritten konvergiert.
- c) Ist das oben gewählte  $C$  auch in der Praxis eine gute Wahl? Geben Sie die Vor- und Nachteile an.
- d) Geben Sie nun zwei Alternativen an, die nicht die Nachteile vom oben gewählten  $C$  haben aber weiterhin aus Sicht der Theorie gut geeignet scheinen. Begründen Sie Ihre Wahl.

*Hinweis:* Im GMRES-Verfahren wird zur Lösung von  $Mx = b$  im  $k$ -ten Schritt das Residuum über den affinen Raum  $x_0 + \mathcal{K}_k(M, r_0)$  minimiert, wobei

$$\mathcal{K}_k(M, r_0) = \text{span}(\{r_0, Mr_0, \dots, M^{k-1}r_0\})$$

der  $k$ -te Krylowunterraum ist. Das Residuum  $r_k = b - Mx_k$  erfüllt

$$r_k = b - M(x_0 + y), \quad y \in \mathcal{K}_k(M, r_0).$$

Damit folgt

$$\begin{aligned} r_k &= r_0 - My \\ &= r_0 - Mp_{k-1}(M)r_0, \quad p_{k-1} \in \mathcal{P}_{k-1} \\ &= (I - Mp_{k-1}(M))r_0 \\ &= \tilde{p}_k(M)r_0, \quad \tilde{p}_k \in \tilde{\mathcal{P}}_k := \{p_k \in \mathcal{P}_k : p_k(0) = I\}. \end{aligned}$$

Es folgt

$$\frac{\|r_k\|_2}{\|r_0\|_2} = \min_{y \in \mathcal{K}_k(M, r_0)} \frac{\|b - M(x_0 + y)\|_2}{\|r_0\|_2} = \min_{\tilde{p}_k \in \tilde{\mathcal{P}}_k} \frac{\|\tilde{p}_k(M)r_0\|_2}{\|r_0\|_2} \leq \min_{\tilde{p}_k \in \tilde{\mathcal{P}}_k} \|\tilde{p}_k(M)\|_2.$$

**Programmieraufgabe 1:** (4 Punkte)

Modifizieren Sie Ihr Programm vom letzten Übungsblatt und ersetzen Sie die Lösung von  $A_1x = y$  bzw.  $Cx = y$  mit dem CG-Verfahren durch folgende (inexakte) Löser: Sei  $H := A_1$  oder  $H := C$  in der Notation des letzten Übungsblatts. Approximieren Sie die Inverse über

$$H^{-1} \approx \hat{H}^{-1} := \sum_{i=1}^{N^2} R_i^T H_{T_i, T_i}^{-1} R_i.$$

Sie müssen also lediglich einmal den bereits implementierten 1-Level-Overlapping-Schwarz-Vorkonditionierer anstelle vom CG-Verfahren aufrufen.

- Vergleichen Sie die benötigte Anzahl an Iterationen,
  1. falls global kein Vorkonditionierer verwendet wird (also nur GMRES),
  2. falls der Rechtsvorkonditionierer vom letzten Übungsblatt mit den inneren CG-Iterationen verwendet wird,
  3. falls der Rechtsvorkonditionierer vom letzten Übungsblatt mit der oben beschriebenen Modifikation verwendet wird.

**Programmieraufgabe 2:** (10 Punkte) (**wird als Theorieaufgabe gezählt**)

Modifizieren Sie Ihr Programm aus Programmieraufgabe 1 wie folgt: Ersetzen Sie das Modellproblem, indem Sie für die rechte Seite der partiellen DGL  $f = 0$  wählen und am unteren Rand des Quadrats eine homogene Neumannrandbedingung (*do-nothing*). Am oberen Rand setzen Sie ein parabolisches Einflussprofil

`uInflow = Expression(('0.0', '-(1-4*(x[0]-0.5)*(x[0]-0.5))'), degree=2)`

und sonst überall eine no-slip-Randbedingung. Nutzen Sie  $\mathcal{P}_2$ - $\mathcal{P}_1$ -Elemente mit  $2 \cdot 40^2$  vielen Dreiecken.

Wählen Sie für die Matrix  $C$  des Vorkonditionierers die zwei Alternativen aus Aufgabe 5d) sowie  $C = h^2 I$  mit<sup>3</sup>  $h = 1/40$ . Nutzen Sie den Ansatz aus Programmieraufgabe 1, um ein System  $Cx = y$  zu lösen; jedoch nur dann, wenn  $C$  keine Diagonalmatrix ist (sonst invertieren Sie die Diagonalmatrix direkt).

- Vergleichen Sie die benötigte Anzahl an Iterationen und testen Sie auch GMRES ohne Vorkonditionierer.

*Hinweise:*

- Vergessen Sie nicht Ihre Exportroutine in Python und Ihre Einleseroutine in Matlab an die  $\mathcal{P}_2$ -Elemente anzupassen.
- Für die Gitterpartitionierung benötigen Sie nun unterschiedliche Knotenindexlisten für  $C$  und  $A_1$ .
- Bei entsprechender Wahl von  $C$  kann dieses wieder mit FEniCS assembliert werden.

<sup>3</sup>A. Klawonn, G. Starke, 1999, *Block triangular preconditioners for nonsymmetric saddle point problems: field-of-values analysis*, Numerische Mathematik, Vol. 81, pp. 577-594.

**Abgabe:**

- **Theorie:** Bis Mittwoch, 03. Juli 2019, 14:00 Uhr,
- **Programme:** Bis Mittwoch, 10. Juli 2019, 14:00 Uhr,

im entsprechenden Kasten in Raum 3.01 des Mathematischen Instituts.

# Spektraläquivalenz

Elman et al.<sup>4</sup> definieren zwei symmetrisch positiv definite Matrizen  $A$  und  $M$  als *spektral äquivalent*, falls zwei Konstanten  $c_1$  und  $c_2$  existierten, sodass

$$c_1 \leq R_{A,M}(x) \leq c_2 \quad \forall x \neq 0,$$

wobei

$$R_{A,M}(x) := \frac{\langle Ax, x \rangle}{\langle Mx, x \rangle}$$

der verallgemeinerte Rayleigh-Quotient ist. Wie wir im Folgenden sehen werden folgt daraus die Beschränktheit der Konditionszahl von  $M^{-1}A$  durch  $\frac{c_2}{c_1}$ . Dies alleine ist nicht sonderlich hilfreich, da wir immer  $c_1$  und  $c_2$  klein bzw. groß genug wählen können. Die Definition der Spektraläquivalenz bezieht sich auf die Unabhängigkeit der Konstanten von der Gitterauflösung  $h$ , falls  $A$  und  $M$  auf, z.B., Finite-Elemente-Matrizen basieren.

Da  $A$  und  $M$  s.p.d. sind, ist auch  $B := M^{-1/2}AM^{-1/2}$  s.p.d.; daher erfüllt der Rayleigh-Quotient für  $B$

$$\lambda_{\min}(B) \leq R_B(x) \leq \lambda_{\max}(B) \quad \forall x \neq 0, \quad (4)$$

wobei

$$R_B(x) := \frac{\langle Bx, x \rangle}{\langle x, x \rangle}.$$

Die Eigenwerte werden von  $R_B(x)$  für die entsprechenden Eigenvektoren angenommen. Da  $M^{-1/2}$  symmetrisch und invertierbar ist, ist  $M^{-1/2}x \neq 0$  für  $x \neq 0$  und (4) ist äquivalent zu

$$\lambda_{\min}(B) \leq \frac{\langle A(M^{-1/2}x), M^{-1/2}x \rangle}{\langle M(M^{-1/2}x), M^{-1/2}x \rangle} \leq \lambda_{\max}(B) \quad \forall x \neq 0,$$

und daher

$$\lambda_{\min}(B) \leq \frac{\langle Ay, y \rangle}{\langle My, y \rangle} \leq \lambda_{\max}(B) \quad \forall y \neq 0.$$

Wie bekannt, falls  $M^{-1}$  als Vorkonditionierer von  $A$  im konjugierten Gradientenverfahren dienen soll, entspricht dies theoretisch dem CG-Verfahren angewandt auf  $B$ . Daher hängt die Konvergenz von der Konditionszahl von  $B$  ab, also

$$\kappa_2(B) = \frac{\lambda_{\max}(B)}{\lambda_{\min}(B)}.$$

Daher gilt nach Definition, falls  $A$  und  $M$  spektral äquivalent sind, dass

$$\kappa_2(B) \leq \frac{c_2}{c_1}.$$

*Bemerkung:*  $M^{-1}A$  ist im Allgemeinen nicht s.p.d. und daher gilt i.A.  $\kappa_2(M^{-1}A) \neq \kappa_2(B)$ . Allerdings gilt  $\kappa_2(B) = \frac{\lambda_{\max}(M^{-1}A)}{\lambda_{\min}(M^{-1}A)}$ , da  $M^{-1}A$  und  $B$  dasselbe Spektrum haben.

---

<sup>4</sup>Remark 2.5, p. 85, H. C. Elman, D. J. Silvester, A. J. Wathen (2014) *Finite Elements and Fast Iterative Solvers: With Applications in Incompressible Fluid Dynamics*, Second Edition, Oxford University Press