Department of Applied Mathematics
Faculty of Electrical Engineering, Mathematics and Computer Sciences
Dr. K. Smetana, Dr. M. Walter

July 3, 2019

Resit of the exam for the Course

Scientific Computing (code: 191551200)

2018/2019

The total number of points is 54 points. In order to get the grade 5.5 you will need at most 27 points. No notes or calculators are allowed. **Explain each step in your solution.** This exam has 9 tasks.

Task 1 (LU decomposition) (4 + 3 + 1 points)

We have given the following matrix A:

$$A := \begin{pmatrix} \alpha & -1 & 0 \\ -1 & \alpha & -1 \\ 0 & -1 & \alpha \end{pmatrix} \qquad \text{with } \alpha \in \mathbb{R}. \tag{1}$$

a) Compute the LU decomposition of A defined in (1) without pivoting, where L has to be a unit lower triagonal matrix.

b) Give conditions for α that are sufficient for the existence of a unit lower triagonal matrix L and an upper triagonal matrix U such that $A = LU$ for A defined in (1).

c) Use the just determined LU decomposition of the matrix A defined in (1) to compute $\det(A)$.

Task 2 (QR decomposition via Householder reflections) (4 points)

Let $v \in \mathbb{R}^n$ and $I \in \mathbb{R}^{n \times n}$. Explain why the Householder matrix

$$H(v) := I - 2\frac{vv^T}{v^Tv}$$

is a reflection and say why it is thus suited for the QR decomposition.
Draw a picture to illustrate your explanation.

Task 3 (Conjugate gradient method) (6 points)

Reminder: Conjugate gradient method: Choose $x^{(0)}$, set $r^{(0)} = b - Ax^{(0)}$, $p^{(0)} = r^{(0)}$. Until convergence compute

$$\alpha_k = \frac{p^{(k)*}r^{(k)}}{p^{(k)*}Ap^{(k)}}, \qquad x^{(k+1)} = x^{(k)} + \alpha_k p^{(k)}, \qquad r^{(k+1)} = r^{(k)} - \alpha_k Ap^{(k)},$$
$$\beta_k = \frac{(Ap^{(k)})^* r^{(k+1)}}{(Ap^{(k)})^* p^{(k)}}, \qquad p^{(k+1)} = r^{(k+1)} - \beta_k p^{(k)}.$$

Recall also the definition of the Krylov subspace $K_k(A; r^{(0)}) = \text{span}\{r^{(0)}, Ar^{(0)}, \ldots, A^{k-1}r^{(0)}\}$.
Actual task: Suppose $A \in \mathbb{R}^{n \times n}$ is symmetric positive definite and $b \in \mathbb{R}^n$. We denote by $\bar{x} \in \mathbb{R}^n$ the solution of $A\bar{x} = b$. Show that $x^{(k)}$ minimizes $\|x - \bar{x}\|_A$ in the subspace $x^{(0)} \oplus K_k(A; r^{(0)})$. Hint: Rewrite $x \in x^{(0)} \oplus K_k(A; r^{(0)})$ as $x^{(k)} + w$ with $w \in K_k(A; r^{(0)})$ and include the argumentation why $x \in x^{(0)} \oplus K_k(A; r^{(0)})$ can be written in that way. Moreover, exploit that $K_k(A; r^{(0)})$ can also be spanned by other vectors.

Task 4 (Weak formulation, coercivity) (7 points)
Let $I = (a, b)$ and consider the following strong formulation of a boundary value problem

$$-(p(x)y'(x))' + q(x)y'(x) + s(x)y(x) = f(x) \;\; \forall x \in I \quad \text{and} \quad y(a) = y(b) = 0. \tag{2}$$

Let $f \in L^2(I)$, $p \in C^0(\bar{I})$, $p(x) \geq p_0 > 0$, $q \in C^1(\bar{I})$, and $s \in C^0(\bar{I})$.

(a) Derive the weak formulation corresponding to (2).

(b) Define an appropriate bilinear form $b : ?\times? \to \mathbb{R}$ based on that weak formulation, noting that the terms that you do not include into the bilinear form must not depend on the weak solution. Under which conditions on q and s is this bilinear form coercive with respect to the norm $\|v\|_X := \left(\int_I (v'(x))^2 \, dx\right)^{1/2}$?

Task 5 (A priori error estimate of Finite Element approximation) (4 points)
Let $I = (a, b)$, $f \in L^2(I)$, $p \in C^1(\bar{I})$, $p(x) \geq p_0 > 0$ and define the vector space $X := \{v \in H^1(I) : v(a) = 0\}$ equipped with the H^1-norm. Furthermore, let $y \in X$ be the weak solution satisfying

$$\int_I p(x)y'(x)v'(x) \, dx = \int_I f(x)v(x) \, dx \quad \forall v \in X.$$

Moreover, let us assume additionally that $y \in H^2(a, b)$. Finally, denote by $y_h \in X_h$ the piecewise linear Finite Element approximation.

State the corresponding a priori error estimate for the Finite Element method, i.e. give the respective estimate for $\|y - y_h\|_X$ including the dependency on h, where the latter denotes the grid size. Give the main ideas for the proof. Three key points are expected; one sentence per key point is sufficient.

Task 6 (Gershgorin circles) (3 + 1 + 2 points)
Consider the matrix

$$A := \begin{pmatrix} 8 & 1 & 0 \\ 1 & 5 & 2 \\ 0 & 2 & 10 \end{pmatrix}.$$

a) Use the Gershgorin circles to approximately locate the eigenvalues of A. You do <u>not</u> have to use a similarity transform.

b) Illustrate your findings from a) in a picture.

c) Use the Gershgorin circle theorem to determine an <u>upper bound</u> for the condition number of the matrix A. Recall to that end that the condition number of A is given by $\kappa(A) = \|A\|_2 \|A^{-1}\|_2$.

Task 7 (Comparison of algorithms that approximate singular values and singular vectors) (10 points)
Assume that you want to compute approximations of the singular values $\sigma_1, \ldots, \sigma_{50}$ and the corresponding left singular vectors $u_1, \ldots, u_{50}$ of a large sparse matrix $A \in \mathbb{R}^{m \times n}$, $m, n \gg 50$. Which algorithms that we discussed during the course can be recommended to compute such an approximation and why? Comment also briefly why other methods we discussed are not suited for the current task. Compare the key ideas of those (two) algorithms. Which algorithm would you prefer in this setting; consider the convergence behavior/rate and the computational costs in this context.

Task 8 (Training deep neural networks) (3 points)
Explain why the steepest descent method is in general not suitable to train deep neural networks. Explain the main idea of an algorithm that can be used instead.

Task 9 (Strength of two models for the Steiner tree problem) (6 points)
Reminder: Definition and notation for Steiner trees and Steiner arborescences:
Let $G = (V, E)$ be an undirected graph, let $c \in \mathbb{R}_+^E$ be a nonnegative objective vector and let $T \subseteq V$ be a set of *terminal nodes*. A *Steiner tree* is a subset $F \subseteq E$ of edges such that F connects all terminals, i.e., for each pair $s, t \in T$ of terminals there is an s-t-path that only uses edges of F. The *Steiner tree problem* is to find a Steiner tree with mininum costs $c(F) := \sum_{e \in F} c_e$. Note that a Steiner tree may in fact contain cycles, but due to nonnegativity of c there always exists an optimal solution that is a tree. Similar to the TSP we consider the bidirected version of G, denoted by $D = (V, A)$, i.e.,

$$A := \{(u, v), (v, u) : \{u, v\} \in E\}$$

contains the two arcs corresponding to each edge of G. Additionally, we fix a node $r \in T$. A *Steiner arborescence* is a subset $B \subseteq A$ of arcs such that for every other terminal $k \in T \setminus \{r\}$ there is a k-r-path only using arcs of B.
Remark: One can think of Steiner arborescences as Steiner trees that are "directed towards r".

Actual task:
Show that the two following integer programming models for the Steiner arborescence problem are equivalent with respect to the y-variables!
The first model has only arc variables $y \in \{0, 1\}^A$ with $y_a = 1 \iff a \in B$. It reads

$$\begin{aligned} \min \quad & \sum_{(u,v) \in A} c_{\{u,v\}} y_{(u,v)} & & \\ \text{s.t.} \quad & y(\delta^{\text{out}}(S)) \geq 1 & \forall S \subseteq V : S \cap T \neq \emptyset,\ r \notin S & \quad (3a) \\ & y_a \in \{0, 1\} & \forall a \in A. & \quad (3b) \end{aligned}$$

The second model uses, in addition to arc variables, flow variables $f^{(k)} \in \mathbb{R}^A$ for all $k \in T \setminus \{r\}$ ensuring existence of a k-r-flow of value 1. It reads

$$\begin{aligned} \min \quad & \sum_{(u,v) \in A} c_{\{u,v\}} y_{(u,v)} & & \\ \text{s.t.} \quad & 0 \leq f_a^{(k)} \leq y_a & \forall a \in A,\ \forall k \in T \setminus \{r\} & \quad (4a) \\ & f^{(k)}(\delta^{\text{out}}(v)) - f^{(k)}(\delta^{\text{in}}(v)) = \begin{cases} +1 & \text{if } v = k \\ -1 & \text{if } v = r \\ 0 & \text{otherwise} \end{cases} & \forall v \in V\ \forall k \in T \setminus \{r\} & \quad (4b) \\ & y_a \in \{0, 1\} & \forall a \in A & \quad (4c) \end{aligned}$$

overview points

1	2	3	4	5	6	7	8	9
8	4	6	7	4	6	10	3	6

Table 1: Total: 54 points