

Name: \_\_\_\_\_

Math 614 Numerical Linear Algebra (Bueler)

Wednesday 15 November 2023

## Midterm Quiz 2

**In-class or proctored. No book, notes, electronics, calculator, internet access, or communication with other people. 100 points possible.**

**65 minutes maximum!**

- 1. (10 pts)** Suppose  $f : X \rightarrow Y$  is a problem, where  $X$  and  $Y$  are normed vector spaces. For  $x \in X$ , define the *relative condition number* of the problem:

$$\kappa(x) =$$

- 2. (15 pts)** Our textbook TREFETHEN & BAU defines an idealized floating point system  $\mathbf{F}$ , also written  $\mathbb{F}$ . Define/describe it. (*Hints.* A floating point system is scientific notation based on a base  $\beta$  and a precision  $t$ . Both  $\beta$  and  $t$  are integers; what are their ranges? Now precisely describe the allowed fractions and exponents.)

**3. (a)** (5 pts) State axiom (13.5).

**(b)** (5 pts) State axiom (13.7).

**4.** (15 pts) Suppose  $f : X \rightarrow Y$  is a problem and  $\tilde{f} : X \rightarrow Y$  is an algorithm to compute (approximate) that problem on a computer satisfying axioms (13.5) and (13.7). Define what it means for the algorithm  $\tilde{f}$  to be *backward stable* for the input  $x \in X$ .

**5.** (7 pts) Show that  $(1 + O(t))(1 + O(t)) = 1 + O(t)$  as  $t \rightarrow 0$ .

**6.** (7 pts) Consider the problem (function)  $f(x) = x^4$  on real numbers. Compute the absolute condition number  $\hat{\kappa}(x)$  and the relative condition number  $\kappa(x)$ .

**7.** (8 pts) Suppose  $A \in \mathbb{C}^{m \times m}$  is invertible, and that  $b \in \mathbb{C}^m$ . Explain, via major steps, how to use the QR factorization to solve the linear system  $Ax = b$ . How much work,<sup>1</sup> i.e. how many floating point operations, is required for each step?

**8.** (8 pts) Suppose  $A \in \mathbb{C}^{m \times n}$  is full rank, and that  $m \geq n$ . Suppose  $b \in \mathbb{C}^m$ . Explain, via major steps, how to use the reduced SVD factorization to solve the overdetermined system “ $Ax = b$ ” by least squares. How much work is required for each step?

---

<sup>1</sup>For problems **7** and **8**, use big-O notation to communicate the amount of work at leading order in  $m$  and/or  $n$ , as they go to infinity. You do not need to prove your big-O usage.

**9.** Suppose  $x \in \mathbb{R}^2$  and that  $f(x) = x_1^2 + x_2^2$ .

**(a)** (4 pts) Write the obvious floating point algorithm for computing  $f$ , using notation  $\text{fl}(\cdot)$ ,  $\oplus$ ,  $\otimes$ :

$$\tilde{f}(x) =$$

**(b)** (8 pts) Assuming a computer satisfying axioms (13.5) and (13.7), show that the above algorithm is backward stable. You may assume here, without proof, that  $(1 + O(t))(1 + O(t)) = 1 + O(t)$  and  $\sqrt{1 + O(t)} = 1 + O(t)$  as  $t \rightarrow 0$ .

**Extra Credit (use space on blank pages).** (2 pts) Assuming that the result in **9 (b)** above can be extended to  $x \in \mathbb{R}^m$  for any  $m$ , argue that the obvious algorithm for computing the 2-norm of a vector is backward stable. Along the way you will need to describe, and briefly justify, the expected stability properties of a fifth arithmetic operation.

**10.** (8 pts) Suppose I invent a new way of solving linear systems which is even more stable than the Householder reflection QR method. The Bueler algorithm solves  $Ax = b$ , for  $A \in \mathbb{C}^{m \times m}$  invertible and  $b \in \mathbb{C}^m$ , in a backward stable manner, with numerical result  $\tilde{x} \in \mathbb{C}^m$  satisfying  $(A + \delta A)\tilde{x} = b$  where  $\|\delta A\|_2/\|A\|_2 \leq 30 \log_{10}(m)\epsilon_{\text{machine}}$ .<sup>2</sup> On a computer with  $\epsilon_{\text{machine}} = 10^{-16}$ , I apply the Bueler algorithm to solve a linear system for a certain matrix  $A \in \mathbb{C}^{1000 \times 1000}$  for which I know that the 2-norm condition number is  $\kappa_2(A) = 10^9$ . How many digits of accuracy will I have in the answer  $\tilde{x}$ ? (*Hints.* Start by being clear on what is the problem. Apply big ideas precisely, but avoid little algebra.)

---

<sup>2</sup>That is,  $\frac{\|\delta A\|_2}{\|A\|_2} = O(\epsilon_{\text{machine}})$  with constant  $C = 30 \log_{10}(m)$ . This  $C$  is much smaller than for Householder QR.

BLANK SPACE (FULL PAGE)

BLANK SPACE (FULL PAGE)