Homework 1: Numerics, Intro to LU

CS 205A: Mathematical Methods for Robotics, Vision, and Graphics (Fall 2013), Stanford University

Due Monday, October 7, at midnight

This homework is is probably one of the most involved "analytical" homeworks of CS 205A, to give you a taste for the science of computing error bounds and so on. Again, feel free to make ample use of office hours, Piazza, and other resources, and get started early. We do not necessarily expect students to finish every problem.

Problem 1 (15 points). *Suppose $f : \mathbb{R} \to \mathbb{R}$ is infinitely differentiable, and we wish to write algorithms for finding x^* minimizing $f(x^*)$. Our algorithm outputs x_{est}, an approximation of x^*. Assuming that in our context this problem is equivalent to finding roots of $f'(x)$, write expressions for:*

(a) Forward error of the approximation.

(b) Backward error of the approximation.

(c) Conditioning of this minimization problem near x^*.

Problem 2 (50 points; adapted from CS 205A 2012). *As we saw in the first lecture, thanks to floating-point arithmetic we cannot expect that computations involving decimal points can be carried out with 100% precision. Instead, with rounding and truncation, each time we do a numerical operation we induce the potential for error.*

Suppose we care about an operation $\diamond$ between two scalars x and y; here $\diamond$ might stand for $+, -, \times$, and so on. As a model for the error that occurs when computing $x \diamond y$, we will say that evaluating $x \diamond y$ on the computer yields a number $(1+\varepsilon)(x \diamond y)$ for some number ε satisfying $0 \leq |\varepsilon| < \varepsilon_{max} \ll 1$; we will assume ε can depend on $\diamond$, x, and y.

(a) Why is this a reasonable model for numerical mistakes in floating point arithmetic? For example, why does this make more sense than assuming that the output of evaluating $x \diamond y$ is $(x \diamond y) + \varepsilon$?

(b) As a convenience for future problems prove the following lemma: Suppose $\varepsilon_1, \ldots, \varepsilon_k$ satisfy $0 \leq |\varepsilon_i| < \varepsilon_{max} \ll 1$ for all i. Then, there exists some ε satisfying $0 \leq |\varepsilon| < \varepsilon_{max}$ such that $(1+\varepsilon_1)(1+\varepsilon_2)\cdots(1+\varepsilon_k) = (1+\varepsilon)^k$.

(c) In reality, we also only know the values of x and y relative to numerical precision; so define $\bar{x} \equiv (1+\varepsilon_x)x$ and $\bar{y} \equiv (1+\varepsilon_y)y$. Accounting for this error as well as that incurred by evaluating $x \diamond y$ and assuming $0 \leq |\varepsilon_x|, |\varepsilon_y| < \varepsilon_{max}$, write bounds for the error of computing the following expressions relative to $x \diamond y$ in terms of ε_{max}:

(i) $x \times y$

(ii) $x \div y$

Note: *It is acceptable to write answers in the form* $c\varepsilon_{max}^k + O(\varepsilon_{max}^{k+1})$*; these higher-order terms generally are negligible since* ε_{max} *is small.*

Hint: *You will most likely need to apply the lemma you just proved as well as the triangle inequality.*

(d) *Argue that the relative error of computing* $x - y$ *can be unbounded. This phenomenon is known as "catastrophic cancellation" and can cause serious numerical issues.*

(e) *Suppose we compute* nx *using the recurrence*

$$\begin{aligned} s_1 &= x \\ s_k &= s_{k-1} + x. \end{aligned}$$

Prove a bound for the relative error of such a computation that is approximately $n\varepsilon_{max}/2$*. Here and for the next part, you can assume* n *is large but still satisfies* $n \ll 1/\varepsilon_{max}$*.*

(f) *Now, suppose* $n = 2^k$*, and suppose we compute* nx *using the recurrence*

$$\begin{aligned} q_0 &= x \\ q_k &= q_{k-1} + q_{k-1} \end{aligned}$$

Prove a similar error bound.

Extra credit. *Show that the Kahan strategy for evaluating this sum exhibits error that does not scale with* n*.*
Warning: *Your instructor was unable to do this in an elegant way!*

Problem 3 (35 points)**.** *This week's linear algebra gymnastics (these are well-known theorems; it goes without saying that copying your proof out of a book/web page comprises a violation of the Honor Code):*

(a) *Show that the product of upper triangular matrices is upper triangular. Use this result to show (in a line or two) that the product of lower triangular matrices is lower triangular.*

(b) *Show that the eigenvalues of upper triangular matrices* $U \in \mathbb{R}^{n\times n}$ *are exactly their diagonal elements.*
Extra credit: *If we assume that the eigenvectors of* U *are* $\vec{v}_k$ *satisfying* $U\vec{v}_k = u_{kk}\vec{v}_k$*, characterize* $\operatorname{span}\{\vec{v}_1, \ldots, \vec{v}_k\}$ *for* $1 \leq k \leq n$ *when the diagonal values* u_{kk} *of* U *are distinct (tricky!).*

(c) *Show that the inverse of an (invertible) lower triangular matrix is lower triangular.*