Numerical Linear Algebra
Homework 7

Alexander Hurnisch
hurnisc6@msu.edu

14.1. True or False?

(a) $\sin x = O(1)$ as $x \to \infty$.

(b) $\sin x = O(1)$ as $x \to 0$.

According to

The Meaning of $O(\epsilon_{\text{machine}})$

We now explain the precise meaning of "$O(\epsilon_{\text{machine}})$" in (14.2)–(14.5).

The notation

$$\varphi(t) = O(\psi(t)) \tag{14.6}$$

is a standard one in mathematics, with a precise definition. This equation asserts that there exists some positive constant C such that, for all t sufficiently close to an understood limit (e.g., $t \to 0$ or $t \to \infty$),

$$|\varphi(t)| \le C\psi(t). \tag{14.7}$$

it is very obvious that both (a) and (b) are true, since:

$$|\sin(x)| \le 1 \quad \forall x$$

15.1. Each of the following problems describes an algorithm implemented on a computer satisfying the axioms (13.5) and (13.7). For each one, state whether the algorithm is *backward stable*, *stable but not backward stable*, or *unstable*, and prove it or at least give a reasonably convincing argument. Be sure to follow the definitions as given in the text.

(b) Data: $x \in \mathbb{C}$. Solution: x^2, computed as $x \otimes x$.

Stability means:

is appropriate to aim for in general is *stability*. We say that an algorithm $\tilde{f}$ for a problem f is *stable* if for each $x \in X$,

$$\frac{\|\tilde{f}(x) - f(\tilde{x})\|}{\|f(\tilde{x})\|} = O(\epsilon_{\text{machine}}) \tag{14.3}$$

for some $\tilde{x}$ with

$$\frac{\|\tilde{x} - x\|}{\|x\|} = O(\epsilon_{\text{machine}}). \tag{14.4}$$

In words,

A stable algorithm gives nearly the right answer to nearly the right question.

Backward Stability means:

stronger and simpler than stability. We say that an algorithm $\tilde{f}$ for a problem f is *backward stable* if for each $x \in X$,

$$\tilde{f}(x) = f(\tilde{x}) \quad \text{for some } \tilde{x} \text{ with } \frac{\|\tilde{x} - x\|}{\|x\|} = O(\epsilon_{\text{machine}}). \tag{14.5}$$

This is a tightening of the definition of stability in that the $O(\epsilon_{\text{machine}})$ in (14.3) has been replaced by zero. In words,

A backward stable algorithm gives exactly the right answer to nearly the right question.

Examples are given in the next lecture.

The algorithm here is $\tilde{f}(x) := \text{float}(x) \otimes \text{float}(x)$
For the problem $f(x) = x^2$ $\qquad = \text{float}(\text{float}(x)^2)$

So for some $x \in X$ we have

$$\tilde{f}(x) = (x(1+\varepsilon_1))^2(1+\varepsilon_2) = x^2(1+\varepsilon_1)^2(1+\varepsilon_2)$$

$$= x^2(1+\varepsilon_3)$$

For some ε_3 with $|\varepsilon_3| \leq 3\,\varepsilon_{\text{machine}} + O(\varepsilon_{\text{machine}}^3)$

In other words: The computed result $\tilde{f}(x)$ is exactly equal to $f(\tilde{x})$ where $\tilde{x}$ satisfies

$$\frac{|\tilde{x}-x|}{|x|} = O(\varepsilon_{machine})$$

And any $C > 3$ will suffice for the constant implicit in the "O" symbol.

Which means that the algorithm is backward stable. $\square$

17.3. Let $L \in \mathbb{C}^{m\times m}$ be a unit lower-triangular matrix (i.e., with diagonal entries equal to 1). For convenience, write L in the form

$$L = \begin{bmatrix} 1 & & & & \\ -\ell_{2,1} & 1 & & & \\ -\ell_{3,1} & -\ell_{3,2} & 1 & & \\ \vdots & \vdots & & \ddots & \\ -\ell_{m,1} & -\ell_{m,2} & -\ell_{m,3} & \cdots & 1 \end{bmatrix},$$

and define $M = L^{-1}$.

(a) Derive a formula for m_{ij} (which may involve other entries of M). Which entries of L does m_{ij} depend on?

We can solve this by forward substitution.
We can also already assume that M is also lower triangular, as prooven already in Homework 1 Problem 1 (Textbook Problem 1.3).

So we have:

$$I = LM \ , \quad \underbrace{I_k = L M_k}_{\text{we can solve each of these by forward sub}} \ , \quad m_{ij} = 0 \text{ for } j > i, \ m_{ii} = 1$$

$$I_k = \underbrace{\vec{e}_m}_{\text{unit vector}} = \sum_{i=1}^{m} (L^T)_i M_k I_i = \sum_{i=1}^{m} \Big[\sum_{j=1}^{i} \underbrace{-l_{ij}}_{l_{ij}=0,\ j>i} m_{jk} \Big] I_i$$

So we know only the $i = k$ term of the first sum will be non-trivial.

$$1 = \sum_{j=1}^{k} -l_{kj} m_{jk} \ , \quad \forall k : 1 \le k \le m$$

$$1 = \sum_{j=1}^{k} -l_{kj}\, m_{jk} \quad , \quad \forall k : 1 \leq k \leq m$$

We know already that $m_{jk} = 0$, $k > j$

that leaves: $1 = -\underbrace{l_{kk}}_{=1} m_{kk} \Rightarrow m_{kk} = 1$, as expected

This was not actually necessary, but well. Using forward sub, we get:

$$m_{jk} = \delta_{jk} + \sum_{i=1}^{j-1} m_{ik}\, l_{ji}$$

The diagonal entries of L are 1, so no division needed!

Kronecker delta

$$\delta_{ij} = \begin{cases} 1, & i=j \\ 0, & i \neq j \end{cases}$$

l_{ij} defined negative

(a) Derive a formula for m_{ij} (which may involve other entries of M). Which entries of L does m_{ij} depend on?

As we've seen, the entries m_{ii} don't depend on L, they are allways equal to one.
So do the entries m_{ij} for $j > i$, they allways vanish.

The entries m_{ij} with $j < i$ depend on the $(i-1)\times(i-1)$ upper left sub-matrix of L, which is itself lower-triangular.