Lecture 14: Stability

Big O

$f = O(\phi(\varepsilon))$ as $\varepsilon \to \varepsilon_0$ means there exists constants k_0 and ε_1 so that

$$|\varepsilon - \varepsilon_0| < \varepsilon_1 \Rightarrow |f(\varepsilon)| \leq k_0 \cdot |\phi(\varepsilon)|$$

examples:

1. $\sin(\varepsilon)$

$\to \sin(\varepsilon) = O(1)$
$\sin(\varepsilon) = O(\varepsilon)$

$>$ gives information about how fast something goes to a limit.

$\sin(\varepsilon) \neq O(\varepsilon^2)$

~~proof~~

$$\sin(\varepsilon) = k_0 \varepsilon^2$$

$$\Rightarrow \frac{\sin(\varepsilon)}{\varepsilon} = k_0 \cdot \varepsilon$$

$$\Rightarrow \lim_{\varepsilon \to 0} \frac{\sin(\varepsilon)}{\varepsilon} = 1 = 0 \,!$$

Algorithms:

problem: $f: X \mapsto Y$
algorithm: $\tilde{f}: X \mapsto Y$

example:

$f: \mathbb{R} \to \mathbb{R}$ defined by $f(x) = x^2$

$\tilde{f}: \mathbb{R} \to \mathbb{R}$ defined by $\tilde{f}(x) = fl[fl(x) \cdot fl(x)]$

Accuracy:

Absolute error $AE = \|\tilde{f}(x) - f(x)\|$

Relative error $RE = \dfrac{\|\tilde{f}(x) - f(x)\|}{\|f(x)\|}$

An algorithm is accurate if

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

Stability:

An algorithm $\tilde{f}$ is stable if for each $x \in X$

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

for some $\tilde{x}$ with

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

☆ An algorism is stable if it gives nearly the right answer to nearly the right question.

Backward Stability:

An algorithm $\tilde{f}$ is backward stable if for each x

$\tilde{f}(x) = f(\tilde{x})$ for some $\tilde{x}$ with $\dfrac{\|x - \tilde{x}\|}{\|x\|} = \mathcal{O}(\varepsilon_{machine})$

☆ An algorithm is backwards stable if it gives exactly the right answer to nearly the right question

examples:

$f : \mathbb{R}^2 \to \mathbb{R}$ $\quad f(x_1, x_2) = x_2 - x_1$

$\tilde{f} : \mathbb{R}^2 \to \mathbb{R}$, $\quad \tilde{f}(x_1, x_2) = fl(x_2) \ominus fl(x_1)$

$$fl(x_1) = x_1(1+\varepsilon_1)$$
$$fl(x_2) = x_2(1+\varepsilon_2)$$

$$\rightarrow \tilde{f}(x) = fl(x_2) \ominus fl(x_1) = (fl(x_2) - fl(x_1))(1+\varepsilon_3)$$
$$= [x_2(1+\varepsilon_2) - x_1(1+\varepsilon_1)](1+\varepsilon_3)$$
$$= x_2(1 + \underbrace{\varepsilon_2 + \varepsilon_3 + \varepsilon_2\varepsilon_3}_{\varepsilon_4}) - x_1(1 + \underbrace{\varepsilon_1 + \varepsilon_3 + \varepsilon_1\varepsilon_3}_{\varepsilon_5})$$
$$= x_2(1+\varepsilon_4) - x_1(1+\varepsilon_5)$$

$|\varepsilon_4| \le 2\,\varepsilon_{machine} + O(\varepsilon^2_{machine})$

$|\varepsilon_5| \le 2\,\varepsilon_{machine} + O(\varepsilon^2_{machine})$

$$\Rightarrow \tilde{f}(x_1, x_2) - f(\underbrace{x_1(1+\varepsilon_4),\ x_2(1+\varepsilon_5)}_{\tilde{x}}) = 0$$

$$\frac{\|x-\tilde{x}\|}{\|x\|} = \frac{\|(\varepsilon_4 x_1,\ \varepsilon_5 x_2)\|}{\|x\|} \le \max\{\varepsilon_4, \varepsilon_5\} \cdot \frac{\|x\|}{\|x\|}$$

$$\le \varepsilon_{machine}.$$

Accuracy of Backward Stable:

Theorem - If a backward stable algorithm is applied to solve a problem $f: X \to Y$ with condition number K then

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

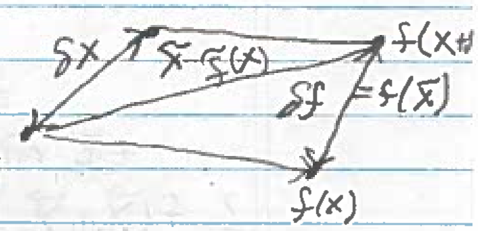
~~proof~~:

By backward stability.

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

for some $\tilde{x}$ satisfying

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



From the definition of condition number

$$K(x) \ge \frac{\|f(x+\delta x) - f(x)\|}{\|f(x)\|} \Big/ \frac{\|x\|}{\|\delta x\|}$$

$$= \frac{\|\tilde{f}(x) - f(x)\|}{\|f(x)\|} \Big/ \cdot \frac{\|x\|}{\|\tilde{x}-x\|}$$

$$\Rightarrow \frac{\|\tilde{x}-x\|}{\|x\|} \ge \frac{\|\tilde{f}(x) - f(x)\|}{\|f(x)\|}$$

$$\Rightarrow \frac{\|\tilde{f}(x) - f(x)\|}{\|f(x)\|} \le \varepsilon_{machine} \cdot K(x) \Rightarrow \frac{\|\tilde{f}(x) - f(x)\|}{\|f(x)\|} = O(K(x)\varepsilon$$

$\boxed{\text{Backward stability} + \text{conditioning number} \Longrightarrow \text{accuracy!!!}}$

examples:

1. $f(x_1, x_2) = x_1 \cdot x_2$

$$\tilde{f}(x_1, x_2) = fl(x_1) \otimes fl(x_2)$$
$$= x_1(1+\varepsilon_1) \otimes x_2(1+\varepsilon_2)$$
$$= [x_1(1+\varepsilon_1) \cdot x_2(1+\varepsilon_2)](1+\varepsilon_3)$$
$$= x_1(1+\varepsilon_1+\varepsilon_3+\varepsilon_1\varepsilon_3) \cdot x_2(1+\varepsilon_2)$$
$$= f(\underbrace{x_1(1+\varepsilon_1+\varepsilon_3+\varepsilon_1\varepsilon_3), x_2(1+\varepsilon_2)}_{\tilde{x}})$$

$$\frac{\|\tilde{x}-x\|}{\|x\|} = \frac{\|(x_1(\varepsilon_1+\varepsilon_3+\varepsilon_1\varepsilon_3), x_2(\varepsilon_2))\|}{\|(x_1, x_2)\|}$$

$$\le \max\{(\varepsilon_1+\varepsilon_3+\varepsilon_1\varepsilon_3, \varepsilon_2\} \frac{\cancel{\|(x_1, x_2)\|}}{\cancel{\|(x_1, x_2)\|}}$$

$$= O(\varepsilon_{machine}).$$

2. $f(\underset{\in\mathbb{R}^2}{\vec{x}}, \underset{\in\mathbb{R}^2}{\vec{y}}) = \vec{x}^T\vec{y} = x_1y_1 + x_2y_2$

$$\tilde{f}(\vec{x}, \vec{y}) = [fl(x_1) \otimes fl(y_1)] \oplus [fl(x_2) \otimes fl(y_2)]$$

$$= (\tilde{x}_1 \cdot \tilde{y}_1) \oplus (\tilde{x}_2 \cdot \tilde{y}_2)$$

(from previous result) $\frac{\|(x_1-\tilde{x}_1, y_1-\tilde{y}_1)\|_\infty}{\|(x_1, y_1)\|_\infty} \le O(\varepsilon_{machine}), \frac{\|(x_2-\tilde{x}_2, y_2-\tilde{y}_2)\|_\infty}{\|(x_2, y_2)\|_\infty} = O(\varepsilon_{machine})$

$$\Rightarrow \tilde{f}(\vec{x}, \vec{y}) = (\tilde{x}_1\tilde{y}_1 + \tilde{x}_2\tilde{y}_2)(1+\varepsilon_m)$$

$$= f(\vec{\tilde{x}}, \vec{\tilde{y}})$$

$$\vec{\tilde{x}} = \left(\tilde{x}_1(1+\varepsilon_m)^{1/2},\ \tilde{x}_2(1+\varepsilon_m)^{1/2}\right)$$
$$\vec{\tilde{y}} = \left(\tilde{y}_1(1+\varepsilon_m)^{1/2},\ \tilde{y}_2(1+\varepsilon_m)^{1/2}\right)$$

$$(1+\varepsilon_m)^{1/2} = 1 + \tfrac{1}{2}\varepsilon_m + O(\varepsilon_m)^2$$

$$\frac{\|(\vec{\tilde{x}},\vec{\tilde{y}}) - (\vec{x},\vec{y})\|}{\|(\vec{x},\vec{y})\|} = \frac{\|\tilde{x}_1(1+\frac{1}{2}\varepsilon_m)-x_1,\ \tilde{x}_2(1+\frac{1}{2}\varepsilon_m)-x_2,\ y_1(1+\frac{1}{2}\varepsilon_m)-y_1,\ y_2(1+\frac{1}{2}\varepsilon_m)\|}{\|(x_1,x_2,y_1,y_2)\|}$$

$$\leq O(\varepsilon_m)$$

example:

$x+1 = f(x)$

$\tilde{f}(x) = fl(x) \oplus 1.$

$$= x(1+\varepsilon_1) \oplus 1$$
$$= (x(1+\varepsilon_1)+1)(1+\varepsilon_2)$$
$$= \underbrace{x(1+\varepsilon_1+\varepsilon_2+\varepsilon_1\varepsilon_2)}_{\tilde{x}} + \varepsilon_2 + 1$$
$$= f(\tilde{x})$$

$$\frac{\|\tilde{x}-x\|}{\|x\|} = \frac{\|x(\varepsilon_1+\varepsilon_2+\varepsilon_1\varepsilon_2)+\varepsilon_2\|}{\|x\|}$$

$$\leq (\varepsilon_1+\varepsilon_2+\varepsilon_1\varepsilon_2) + \frac{\varepsilon_2}{\|x\|}$$

$\Rightarrow$ diverges if $x \approx 0$.

example:

Outer product

$A = xy^*$

$\downarrow$

Not possible to be backwards stable.

$xy^* \Rightarrow$ rank 1

$$\tilde{f}(x,y) = \begin{bmatrix} fl(x_1) \otimes fl(y_1) & \cdots & fl(x_1) \otimes fl(y_n) \\ \vdots & & \vdots \\ fl(x_n) \otimes fl(y_1) & \cdots & fl(x_n) \otimes fl(y_n) \end{bmatrix} \rightarrow \text{Not rank } 1$$