Chapter 7

7.1 Accuracy

Problem: Compute $f(x)$ given x. Result in floating point arithmetic $\tilde{f}(x)$.

Definition: An algorithm is accurate if

$$\frac{\left\|\tilde{f}(x) - f(x)\right\|}{\|f(x)\|} = O(\epsilon_{\text{machine}}) \tag{7.1}$$

Usually too much to ask if f is ill conditioned.

7.2 Stability

An algorithm $\tilde{f}$ for a problem f is stable if for all x

$$\frac{\left\|\tilde{f}(x) - f(\tilde{x})\right\|}{\|f(\tilde{x})\|} = O(\epsilon_{\text{machine}}) \tag{7.2}$$

For some $\tilde{x}$ such that:

$$\frac{\|\tilde{x} - x)\|}{\|x\|} = O(\epsilon_{\text{machine}}) \tag{7.3}$$

Stable algorithm gives almost the right answer to nearly the right question.

7.3 Backward Stability

Satisfied by most algorithms in this course.

Definition: f is backward stable if

$$\tilde{f}(x) = f(\tilde{x}), \quad \text{with} \quad \frac{\|\tilde{x} - x)\|}{\|x\|} = O(\epsilon_{\text{machine}}) \tag{7.4}$$

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

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

Converges to 0 linearly with $\epsilon_{\text{machine}}$ as $\epsilon_{\text{machine}} \to 0$.

In practice $O(\epsilon_{\text{machine}})$ means $\leq C \cdot \epsilon_{\text{machine}}$, where C is a low order polynomial in the size of the problem. If solving $Ax = b$, size$(A) = [m, n]$.

Usually $O(\epsilon_{\text{machine}})$ means 100ϵ, 1000ϵ.

7.5 Stability of Floating Point

Consider the problem of computing $x^T y$, $x, y \in \mathbb{R}^n$

$$\alpha = x^T y = \sum_{i=1}^{n} x_i y_i \tag{7.5}$$

Start with $n = 1$:

$$\begin{aligned} \alpha &= x_1 y_1 && (7.6) \\ \tilde{\alpha} &= x_1 y_1 (1 + \delta_1) \quad |\delta_1| \leq \epsilon_{\text{machine}} \\ &= x_1 (1 + \delta_1) y_1 && (7.7) \end{aligned}$$

$\tilde{\alpha}$ is the exact answer for slightly perturbed data $x_1(1 + \delta_1)$ and y_1.

$n = 2$:

$$\begin{aligned} fl\left(\sum x_i y_j\right) &= (x_1 y_1 (1 + \delta_1) + x_2 y_2 (1 + \delta_2))(1 + \mu_1) \quad |\delta, \mu| \leq \epsilon_{\text{machine}} \\ &= x_1 y_1 (1 + \delta_1)(1 + \mu_1) + x_2 y_2 (1 + \delta_2)(1 + \mu_1) \\ &= \underbrace{x_1 (1 + \delta_1)(1 + \mu_1)}_{\tilde{x}_1} y_1 + \underbrace{x_2 (1 + \delta_2)(1 + \mu_1)}_{\tilde{x}_2} y_2 \\ &= \tilde{x}_1 y_1 + \tilde{x}_2 y_2 \\ &= \tilde{x}^T y && (7.8) \end{aligned}$$

where,

$$\begin{aligned} \tilde{x} &= (\tilde{x}_1, \tilde{x}_2) \\ &= (x_1 (1 + \delta_1)(1 + \mu_2), x_2 (1 + \delta_2)(1 + \mu_2)) && (7.9) \\ \frac{\|\tilde{x} - x\|}{\|x\|} &\leq 2\epsilon + O(\epsilon^2) && (7.10) \end{aligned}$$

General case: Define

$$\begin{aligned}
\tilde{x}_i &= (1+\delta_i)(1+\mu_{i-1})(1+\mu_i)\cdots(1+\mu_{n-1}) && (7.11)\\
\tilde{y}_i &= y && (7.12)
\end{aligned}$$

$$\begin{aligned}
fl(x^T y) &= fl\left(\sum_{i=1}^{n} x_i y_y\right)\\
&= ((x_1y_1(1+\delta_1)+x_2y_2(1+\delta_2))(1+\mu_1)+x_3y_3(1+\delta_3))(1+\mu_2)+\cdots\\
&= x_1y_1(1+\delta_1)(1+\mu_1)(1+\mu_2)\cdots(1+\mu_{n-1})+\\
&\quad x_2y_2(1+\delta_2)(1+\mu_1)(1+\mu_2)\cdots(1+\mu_{n-1})+\\
&\quad x_3y_3(1+\delta_3)(1+\mu_2)(1+\mu_3)\cdots(1+\mu_{n-1})+\\
&\quad \cdots+\\
&\quad x_ny_n(1+\delta_n)(1+\mu_{n-1}) && (7.13)\\
&= \tilde{x}^T y && (7.14)
\end{aligned}$$

where,

$$\frac{\|\tilde{x}-x\|}{\|x\|} \le n\epsilon \tag{7.15}$$

Example: Not backward stable, the outer product

$$fl(xy^T) = [x_iy_j(1+\delta_{ij})] \tag{7.16}$$

n^2 errors, $2n$ parameters. $fl(xy^T)$ is unlikely to be rank 1.