Condition Number

Remark. Back to the condition of a problem. We may compute a **condition number** κ of a problem to quantify how relative error in inputs may amplify that in outputs. It roughly satisfies the following

$$\mathcal{E}_{\text{output}} \approx \kappa \mathcal{E}_{\text{input}}.$$

Example. Consider function evaluation, finding $f(x_0)$ at $x = x_0$ (just like in nested polynomials). We perturb the input by δ. Then, we consider the relative error in output,

$$\begin{aligned} \mathcal{E}_{output} &= \frac{f(x_0+\delta) - f(x_0)}{f(x_0)} = \frac{\delta f'(\xi)}{f(x_0)} = \left(\frac{x_0 f'(\xi)}{f(x_0)}\right)\left(\frac{\delta}{x_0}\right) \\ &\approx \left(\frac{x_0 f'(x_0)}{f(x_0)}\right)\left(\frac{\delta}{x_0}\right) =: \kappa_f(x_0)\,\mathcal{E}_{input} \end{aligned}$$

where $\xi \in (x_0, x_0+\delta)$ (motivated by Mean Value Theorem). The second term in the parenthesis, $\frac{\delta}{x_0}$, is the relative error in input ($\frac{x_0+\delta-x_0}{x_0}$). Thus, we call

$$\kappa_f(x_0) = \left|\frac{x_0 f'(x_0)}{f(x_0)}\right|,$$

the condition number of evaluating $f(x)$ at $x = x_0$. As you can tell, this number can grow pretty large if we have a steep slope (consider evaluating $x = 1$ for $f(x) = \frac{x}{1-x}$).

Remark. The "interface" between well-conditioned/ill-conditioned is not that sharp. However, we do find $\kappa < 1$ to represent a well-conditioned problem, and otherwise ill-conditioned.

Example. (Well-conditioned problem vs Unstable Algorithm)

Let's revisit function evaluation. Consider $f(x) = \sqrt{1+x} - 1$ and we wish to evaluate numbers $x = x_0 \approx 0$.

$$\kappa_f(x) = \frac{\sqrt{1+x}+1}{2\sqrt{1+x}}$$

so $\kappa_f(0) = 1$, which means it is at least not ill-conditioned (error propagation stays neutral).

However, to compute $f(x_0)$, we have a simple three-step algorithm:

(1) $s_1 = 1 + x_0$ (stable with $\kappa_{s_1}(0) = \frac{x_0 \times 1}{1+x_0}\,|_{x_0=0} = 0$)
(2) $s_2 = \sqrt{s_1}$ (stable with $\kappa_{s_2}(s_1) = \frac{s_1 \frac{1}{2\sqrt{s_1}}}{\sqrt{s_1}} = \frac{1}{2}$, regardless of evaluation point)
(3) $s_3 = s_2 - 1$ (unstable with $\kappa_{s_3} = \frac{s_2}{\sqrt{s_2 - 1}}$ near $s_2 \approx 1$ – consider catastrophic cancellation).

We would say this algorithm is unstable because one of the steps is unstable. The remedy is against to use rationalization to bypass the catastrophic cancellation. The analysis of the algorithm using rationalization is left as a HW exercise (HW2 Problem 1).

From the last example, we see that the most obvious algorithm to solve a well-conditioned problem may not be stable. Here are a few more takeaways:

(1) Well-/ill-conditioned refers to the problem.
(2) Stable/Unstable refers to an algorithm or a numerical process.
(3) If the problem is well-conditioned, then there is a stable way to solve it (but not vice versa).
(4) If the problem is ill-conditioned, then there is no reliable way to solve it in a stable way.

One must remember that there are always two facets of scientific computing: the problem and the algorithm. You must study both the condition of the problem AND the stability of your algorithm, and then decide whether the problem should be relaxed to improve its condition, or the algorithm should be revamped to improve stability. It is crucial to know both.

Rates of Convergence

Consider two functions $f(x) = \frac{1}{x}$ and $g(x) = e^{-x}$. Send x to ∞. We know both limits are zero. But, which one does it faster? We actually compare their rates via L'Hôpital,

$$\lim_{x\to\infty} \frac{f(x)}{g(x)} = \lim_{x\to\infty} \frac{1/x}{e^{-x}} = \lim_{x\to\infty} \frac{e^x}{x} = \lim_{x\to\infty} \frac{e^x}{1} = \infty.$$

So, the quotient blows up, meaning that $f(x)$ is running away faster than $g(x)$ – or in other words, $g(x) \to 0$ faster than $f(x)$ does.

Can we quantify this "rate" when we no longer have differentiability to help us, i.e. bye-bye L'Hôpital?

How about first just a sequence of numbers? For example, compare the rate at which

$$\alpha_n = \frac{1}{n}, \quad \gamma_n = \frac{1}{n^2}$$

converge to 0? We anticipate that γ_n goes to zero faster, in the sense that we can compute

$$\lim_{n\to\infty} \frac{\alpha_n}{\gamma_n} = \lim_{n\to\infty} \frac{1/n}{1/n^2} = \infty$$

which means α_n is much larger than γ_n when n is large – γ_n is much closer to zero than $1/n$ and thus converges faster! Now, the task is to quantify the rate of convergence of the sequences.

Definition. ("big O notation") Suppose $\{\beta_n\}_{n=1}^{\infty}$ is a sequence known to converge to zero, and $\{\alpha_n\}_{n=1}^{\infty}$ converges to a number α. If a positive constant K exists with

$$|\alpha_n - \alpha| \le K\beta_n, \quad \text{for large } n,$$

then we say that $\{\alpha_n\}_{n=1}^{\infty}$ converges to α with **rate**, or **order**, **of convergence** $O(\beta_n)$ ("big O of β_n"). We also write this as

$$\alpha_n = \alpha + O(\beta_n), \quad n \to \infty.$$

Though $\{\beta_n\}$ is an arbitrary sequence that goes to zero, we almost always use the form

$$\beta_n = \frac{1}{n^p}, \quad p > 0.$$

We are interested in the largest value of p such that $\alpha_n = \alpha + O(1/n^p)$. For example, if $p = 1$, we say it is 1st-order/linear convergence; $p = 2$, 2nd order/quadratic convergence, and so on.