Condition Numbers

A problem is *well conditioned* if a small change in the input always creates a small change in the output (solution).

A problem is *ill-conditioned* if a small change in the input can create a large change in the solution (output).

There is actually a continuum here, ranging from the extremely well conditioned (e.g. "evaluate the fcn $f(x) = 1$") to the extremely ill-conditioned, (e.g. "evaluate a fcn at a discontinuity", infinitely ill-conditioned problems are often called ill-posed). In order to quantify the notion, we will define a *condition number*.

A condition number is simply a number which describes how well or ill-conditioned a problem is; the bigger the number the more ill-conditioned the problem. Ideally, an absolute condition number, ν, will behave as follows:

$$\|\text{ change in solution }\| \quad = \quad \nu \ \cdot \ \|\text{ change in input }\|,$$

or a relative condition number, κ, would satisfy

$$\frac{\|\text{ change in solution }\|}{\|\text{ solution }\|} \quad = \quad \kappa \ \cdot \ \frac{\|\text{ change in input }\|}{\|\text{ input }\|}.$$

In most cases, quantities like ν or κ above are impossible to compute; impossible even to define, except in the abstract sense above, for these true condition numbers are functions like $\nu = \nu($ input, change in input, solution, change in solution $)$, and we just don't know these quantities.

But that's ok. The *purpose* of a condition number is to *estimate* things like the difficulty of a problem, or the size of error in a computation. For this we want a condition number that we can easily compute, and which will satisfy

$$\|\text{ change in solution }\| \quad \approx \quad \bar{\nu} \ \cdot \ \|\text{ change in input }\|,$$

in the case of an absolute condition estimator. To get this kind of result we usually restrict input perturbations to be very small and use some notion of derivative.

There have been many condition numbers proposed for the problem "solve $Ax = b$", but the most common is the relative condition number $\kappa = \|A\|\|A^{-1}\|$. Here is one (of many) justifications for its use. Let $x(\epsilon)$ be the solution to $(A + \epsilon E)x(\epsilon) = b + \epsilon e$. Then

$$\dot{x} \equiv \frac{dx}{d\epsilon} = (A + \epsilon E)^{-1}(e - Ex),$$

and (Taylor's theorem)

$$\begin{aligned} x(\epsilon) - x(0) &= \epsilon \dot{x}(0) + O(\epsilon^2) \\ &= \epsilon A^{-1}(e - Ex) + O(\epsilon^2). \end{aligned}$$

Thus

$$\frac{\|x(\epsilon) - x(0)\|}{\|x(0)\|} \leq \|A^{-1}\|\|A\| \left(\frac{\|\epsilon e\|}{\|b\|} + \frac{\|\epsilon E\|}{\|A\|} \right).$$