The condition number for differential equations

Brian Sutton

25 Jan 2020

Numerical Solution of Ordinary Differential Equations by L. Fox and D. F. Mayers takes an unusual approach. Before attempting a single numerical computation, the authors introduce the concept of *inherent instability*. They write the following:

> In most scientific problems not all the data are known exactly. They will contain 'physical constants' ...whose values are generally obtained by experiment.... They may also contain the results of measurement, say with a ruler or protractor.... [For purely mathematical problems,] numbers like e or π cannot be stored exactly....
>
> The data therefore have uncertainties, hopefully reasonably small, and as a result the required answers will also have uncertainties which may or may not be small.... If the relevant uncertainty [in the answer] is large we say that the problem is *inherently unstable*, or *ill-conditioned....*

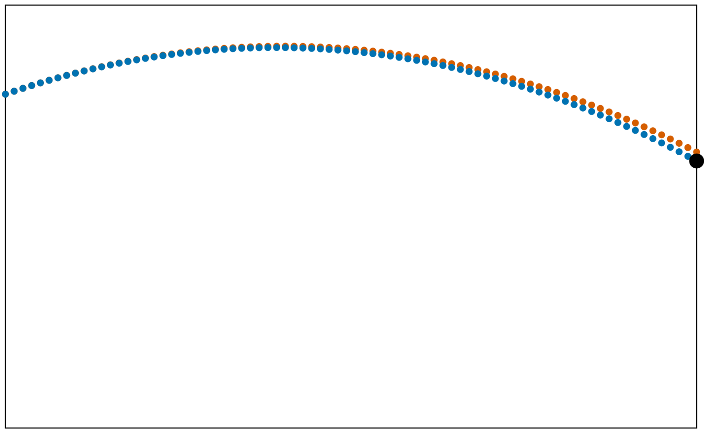
In the preface of the book, which was published in 1987, Fox and Mayers discuss their early and frequent consideration of inherent instability: "We believe that this feature is very important but largely neglected, and that it is desirable to know (and to practice that knowledge) how to determine the effects of small changes in the relevant data on the answers, and how, if necessary, some reformulation of the original problem can reduce the sensitivity."

I discovered the Fox–Mayers book not long before I finished my own book, *Numerical Analysis: Theory and Experiments,* which also considers the conditioning of differential equations. In this post, I explore the similarities and differences between the books' approaches. Briefly, the goal is the same: whereas many discussions on the numerical solution of differential equations mention instability as a qualitative phenomenon to recognize, we want to *measure* sensitivity quantitatively and predict instability in advance. The two books work toward this goal in somewhat different ways.

1 Illustrations

To illustrate inherent stability/instability, Fox and Mayers begin with a game of darts, measuring the effects of a small change in initial velocity or angle of elevation. They show that the dart throw is well conditioned, in the sense that small adjustments produce small changes in outcomes.

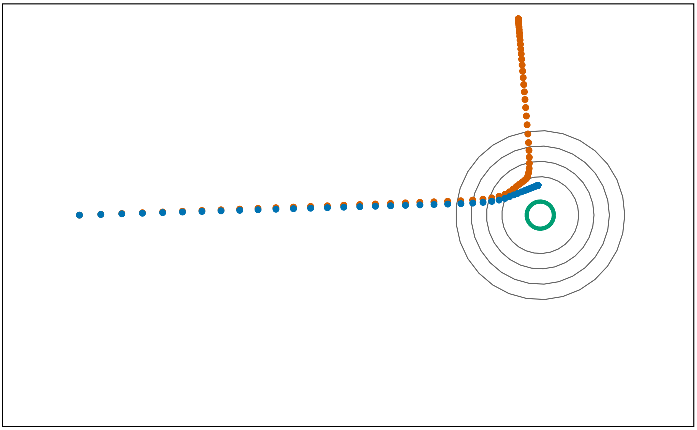
Side view of a dart throw



In the figure above, the blue curve is the trajectory of a dart hitting the bullseye. The orange curve comes from a slightly greater initial velocity; the dart lands a little above the bullseye. A small change in the initial condition produces a small change in the result.

A quite different example is the “mole hole” from my favorite miniature golf course when I was a kid. The target is a cup at the top of a small circular plateau. While a perfect strike sends the ball up the hill and into the hole, a small error can lead to catastrophic results: the ball rolling back down the hill, sometimes leaving the player in a worse position than before.

Top–down view of a miniature golf putt



Above, the blue trajectory leaves the ball at the top of the plateau for an easy tap-in, while the orange trajectory ends far from the base of the plateau. A tiny difference in the initial angle of attack leads to a great difference in the final result. This is sensitivity to initial conditions. We also say that the problem of computing the ball's path is ill-conditioned (for certain starting data).

2 Measuring sensitivity

Fox and Mayers focus on the implications of uncertainty in a single parameter at a time. Their model for the dart throw is as follows:

- Variables: t (time), x (horizontal position), y (vertical position)
- Parameters: g (acceleration by gravity), D (horizontal distance between thrower and dartboard)
- Differential equations:

$$\frac{d^2x}{dt^2} = 0, \quad \frac{d^2y}{dt^2} = -g$$

- Initial data: V (initial velocity), θ (initial angle of elevation), H (height of release)
- Initial conditions:

$$x(0) = 0, \quad \left.\frac{dx}{dt}\right|_{t=0} = V\cos\theta, \quad y(0) = H, \quad \left.\frac{dy}{dt}\right|_{t=0} = V\sin\theta$$

- Terminal event: $x(T) = D$

Note that the above equations imply $dx/dt = V\cos\theta$ and

$$\frac{dy}{dx} = \frac{dy/dt}{dx/dt} = \frac{1}{V\cos\theta}\frac{dy}{dt}$$

and

$$\frac{d^2y}{dx^2} = \frac{d}{dt}\left[\frac{dy}{dx}\right]\frac{1}{dx/dt} = \frac{1}{V^2\cos^2\theta}\frac{d^2y}{dt^2} = -\frac{g}{V^2\cos^2\theta}.$$

Therefore, the dart's trajectory in the xy–plane satisfies

$$\frac{d^2y}{dx^2} = -\frac{g}{V^2\cos^2\theta}, \quad y|_{x=0} = H, \quad \left.\frac{dy}{dx}\right|_{x=0} = \tan\theta.$$

Fox and Mayers consider the effect of a perturbation of the initial velocity V. Let $y_V = \partial y/\partial V$, the change in height with respect to initial velocity. Differentiating all of the previous quantities with respect to V gives the following initial-value problem for y_V:

$$\frac{d^2y_V}{dx^2} = \frac{2g}{V^3\cos^2\theta}, \quad y_V|_{x=0} = 0, \quad \left.\frac{dy_V}{dx}\right|_{x=0} = 0.$$

The blue dart trajectory was computed using $g = 32$ ft/s^2, D = 8 ft, $V = 20$ ft/s, and $\theta = \pi/12$. Solving for y_V using these values gives

$$y_V|_{x=D} = 0.274.$$

(Note that this could be found either by hand or using a numerical solver. For more difficult examples, a numerical solver is necessary.) What is the implication? Suppose, for example, that the initial velocity imparted by the dart thrower were increased by 0.3 ft/s. The sensitivity analysis predicts that the point of impact on the dart board would rise by approximately

$$0.274 \times 0.3 = 0.082.$$

In fact, the orange dart trajectory was found by increasing V from 20 ft/s to 20.3 ft/s. A closer inspection finds that the point of impact rose by about 0.080 ft, in line with the sensitivity analysis.

Fox and Mayers compute similar perturbations for $y_g = \partial y/\partial g$ and $y_\theta = \partial y/\partial\theta$. With the data above, their analysis produces $y_g = -0.09$ and $y_\theta = 7.10$ at $x = D$. Because none of the partial derivatives is especially large, the system is not particularly sensitive to changes in the initial velocity, force of gravity, or angle of elevation.

My book introduces a related approach using a *condition number*. The condition number quantifies the ratio of uncertainty in the solution to uncertainty in the data, focusing on the worst-case scenario. For the dart example, the absolute condition number is about 30; for the mini golf example, it is about 50 000. Although a direct comparison of these numbers is not fair because they are based on different quantities measured in different units of measurement, the greater condition number is cause for a closer look.

In my presentation, the condition number for an initial-value problem or boundary-value problem is exactly the same condition number that is introduced earlier for systems of linear equations. From a tool we already possess, we get warnings of instability for free (not to mention rigorous error bounds).

3 Historical context

Historically, there has been much discussion of ill-conditioning in the numerical solution of differential equations, but relatively little mention of the condition *number* for differential equations. That is, the discussion is often qualitative rather than quantitative.

I suspect one reason is the historical prevalence of explicit methods. An explicit method marches forward in time by small steps, evaluating a sequence defined by an explicit recurrence relation. (For example, Euler's method is defined by the recurrence $y_{k+1} = y_k + hf'(y_k)$.) A small error in one step propagates forward into subsequent steps leading to a cascade of errors. An ill-conditioned problem can amplify the errors so that the global error is large even though the local error at each step is small. Analyzing this cascade of errors is complicated.

By comparison, the sensitivity analysis of the linear algebraic system $Ax = b$ is simple. The residual $b - A\hat{x}$ and the error $x - \hat{x}$ are related by $\|x - \hat{x}\| \leq \|A^{-1}\| \|b - A\hat{x}\|$, regardless of how $\hat{x}$ is computed.

For a one-semester course on numerical analysis, I am a proponent of implicit methods for differential equations. With these methods, sample values $y_1, y_2, y_3, \ldots$ are determined by a system of algebraic equations, which must be solved by a method such as Gaussian elimination or Newton's method. In either case, a matrix equation $Ax = b$ arises—directly in the linear case or, in the nonlinear case, as the inner loop of a nonlinear method. A large condition number for the matrix A can indicate instability in the original differential equation. What we learn about conditioning of matrices translates directly to problems involving differential equations.

One technical point should be mentioned. Many implicit methods are implemented in a stepwise fashion. Rather than computing a separate condition number for each step, which would ignore the effects of error propagation across steps, we want a single condition number for the entire problem. The resolution is this: It is possible to express the entire computation in terms of a single block matrix, with each block index corresponding to a time step. Fortunately, the condition number of this huge matrix can be estimated efficiently without explicitly constructing the matrix in computer memory.

4 Conclusion

A large condition number warns that danger lurks. A sensitivity analysis can reveal the source of the danger. My book *Numerical Analysis: Theory and Experiments* discusses the condition number and provides code for computing it, allowing the detection of ill-conditioned problems with no extra human effort. The more nuanced sensitivity analysis of Fox and Mayers helps us to identify the nature of the sensitivity.