The condition number and perturbations

Suppose we want to look at the linear system

$$Ax = b$$

and suppose that we want to know how the solution changes as the RHS, b changes. Consider two specific cases

$$\begin{aligned} Ax &= b_1 \\ Ax &= b_2 \end{aligned}$$

then (assuming A is nonsingular)

$$x_1 - x_2 = A^{-1}(b_1 - b_2),$$

so then the relative error in x_2 is an approximation to x_1 iven by

$$\frac{||x_1 - x_2||}{||x_1||} \leq ||A^{-1}|| \frac{||b_1 - b_2||}{||x_1||}.$$

It is a good idea to bound the error by something that doesn't depend on the solution. So, replace x_1 in the RHS denominator. Noting $||A||||x_1|| \geq ||b_1||$, write

$$\frac{1}{||x_1||} \leq \frac{||A||}{||b_1||}$$

and therefore

$$\frac{||x_1 - x_2||}{||x_1||} \leq ||A||||A^{-1}|| \frac{||b_1 - b_2||}{||b_1||}.$$

The quantity $||A||||A^{-1}||$ is the **condition number**. It depends only on the matrix in the problem and not on the RHS but it appears on the RHS of the equation above to amplify the relative change.

Definition: For a given matrix $A \in \mathbf{R}^{n \times n}$ and a given matrix norm $|| \cdot ||$, the **condition number** with respect to the given norm is

$$\kappa(A) = ||A||||A^{-1}||.$$

If A is singular then $\kappa(A) = \infty$, justified as follows.

Theorem: Let $A \in \mathbf{A}^{n \times n}$ be given and singular. Then for any singular matrix $B \in \mathbf{B}^{n \times n}$

$$\frac{1}{\kappa(A)} \leq \frac{||A - B||}{||A||}.$$

Proof: We have that

$$\begin{aligned} \frac{1}{\kappa(A)} &= \frac{1}{||A||||A^{-1}||} \\ &= \frac{1}{||A||} \left(\frac{1}{\underset{x \neq 0}{max} \frac{||A^{-1}x||}{||x||}} \right) \\ &\leq \frac{1}{||A||} \left(\frac{1}{\frac{||A^{-1}x||}{||x||}} \right) \\ &\leq \frac{1}{||A||} \left(\frac{||Ay||}{||y||} \right), \end{aligned}$$

where $y \in \mathbf{R}^n$ is arbitrary. Let y be a nonzero vector such that $By = 0$ since B is singular we know such things exist. Then

$$\frac{1}{\kappa(A)} \leq \frac{||(A - B)y||}{||A||||y||} \leq \frac{||(A - B)||||y||}{||A||||y||} = \frac{||(A - B)||}{||A||}$$

completing the proof.

This tells us that if A is close to a singular matrix then the reciprocal of the condition number is near zero and then $\kappa(A)$ must be large. So, the condition number can be used as a measure of how close a matrix is to being singular. We have seen how solving a system that is nearly singular can lead to large errors.

You know that solutions of the system $Ax = b$ are affected by rounding error and while each occurrance is small the cumulative effect (if n is large) can be catastrophic and produce meaningless results.

Theorem (the effects of perturbation in n): Let $A \in \mathbf{R}^{n \times n}$ and $b \in \mathbf{R}^n$ given and define $x \in \mathbf{R}^n$ the solution of the linear system $Ax = b$. Let $\delta b \in \mathbf{R}^n$ a small perturbation of b and define $x + \delta x \in \mathbf{R}^n$ as the solution of the systme $A(x+\delta x) = (b+\delta b)$. Then

$$\frac{||\delta x||}{||x||} \leq \kappa(A)\frac{||\delta b||}{||b||}.$$

Proof: Have $Ax = b$ and $A(x+\delta x) = (b+\delta b)$ so $A\delta x = \delta b$. Thus

$$\delta x = A^{-1}\delta b \Rightarrow ||\delta x|| \leq ||A^{-1}||||\delta b||$$

so that

$$\frac{||\delta x||}{||x||} \leq \frac{||A||||A^{-1}||||\delta b||}{||A||||x||}$$

and using $||A||||x|| \geq ||b||$ and the definition of the condition number gives

$$\frac{||A||||A^{-1}||||\delta b||}{||A||||x||} \leq \kappa(A)\frac{||\delta b||}{||b||}$$

completing proof.

The theorem demonstrates that the effects of perturbations of the problem on the end result are amplified by the condition number. With an ill-conditioned matrix (large κ) a small change in the data can mean a large change in the solution. This can also be expressed using the **residual** of the solution. If x_c is the computed solution then the residual r is the amount by which x_c fails to solve the problem, i.e. $r = b - Ax_c$. Although $r = 0$ means x_c the exact solution it isn't the case that r small means x_c a good solution.

Theorem: Let $A \in \mathbf{R}^{n\times n}$ nonsingular and $b \in \mathbf{R}^n$ known and $x_c \in \mathbf{R}^n$ the computed solution of $Ax = b$ then

$$\frac{||x - x_c||}{||x||} \leq \kappa(A)\frac{||r||}{||b||}$$

Proof: as exercise

Theorem: Let $A \in \mathbf{R}^{n\times n}$ given and nonsingular and $E in \mathbf{R}^{n\times n}$ a perturbation of A. Let $x \in \mathbf{R}^n$ be the unique solution of $Ax = b$. If $\kappa(A)||E|| < ||A||$

then the perturbed system $(A+E)x_c = b$ has a unique solution and

$$\frac{||x - x_c||}{||x||} \leq \frac{\theta}{1-\theta},$$

where

$$\theta = \kappa(A)\frac{||E||}{||A||}.$$

Proof: as exercise

These perturbation results mean an computed (approximate) solution to a linear algebra system is the *exact* solution to a nearby problem.

Theorem: Let $A \in \mathbf{R}^{n\times n}$ given and nonsingular, $b \in \mathbf{R}^n$ and let $x \in \mathbf{R}^n$ be the exact solution of $Ax = b$. Let x_c be the computed solution then there exists δb such that x_c is the exact solution of the system

$$Ax_c = b + \delta b$$

Proof: Write

$$Ax_c = A(x + (x_c - x)) = b + A(x_c - x) = b + \delta b$$

and also

$$\delta b = A(x_c - x) = Ax_c - b = r$$

completing the proof.

Estimating the condition number

Although singular matrices occur infrequently in computations it can be shown with mathematical rigour that all singular matrices are arbitrarily close to a nonsingular matrix. In numerical computation the odds are high that the rounding error can perturb the matrix away from being singular but the resulting nonsingular matrix will be illconditioned. We know that algorithms fail if a pivot is zero as happens with singular matrices but even small pivots can mean the matrix is nearly singular.

As a rule of thumb: if the solution of a linear system changes by a lot when the problem changes by a little then the matrix is probably ill-conditioned.

If would be useful to know this in advance and the condition number can tell us. So can we (easily) determine it? *No!*. But we can estimate $\kappa_\infty(A)$ with approx the same numerical effort it takes to solve $Ax = b$ once.

What's the problem with computing $\kappa(A)$? While we can easily compute $||A||_\infty$, it's the $||A^{-1}||_\infty$ piece that is difficult. And of course if A *is* ill-conditioned then a calculation of A^{-1} will be unreliable.

So, go back to the norm definition

$$||A^{-1}||_\infty = \overset{max}{x \neq 0} \frac{||A^{-1}x||_\infty}{||x||_\infty}$$

and introduce $\omega \in \mathbf{R}^n, \omega \neq 0$ such that

$$||A^{-1}||_\infty \geq \frac{||A^{-1}\omega||_\infty}{||x||_\infty}$$

Now set $y = A^{-1}\omega$ and substitute in the condition number definition

$$\kappa_\infty(A) \geq ||A||_\infty \frac{||y||_\infty}{||\omega||_\infty}, \tag{8.5}$$

true for any $y \in \mathbf{R}^n, y \neq 0, \omega = Ay$. The trick then is to find a y or ω that maximes the RHS of eqn **??**. Since the inequality holds for any $y, \omega = Ay$ then the larger the RHS the better the estimate of κ.

A singular matrix has at least one zero eigenvalue and a matrix close

to singular therefore has at least one eigenvalue close to zero. If y is the eigenvector of a matrix A with the smallest eigenvalue then it follows that

$$\frac{||y||_\infty}{||\omega||_\infty} = \frac{||y||_\infty}{||Ay||_\infty} = \frac{||y||_\infty}{||\lambda y||_\infty} = |\lambda|^{-1}$$

and we can use this to maximise the value of the condition number.

Use a result from more advanced linear algebra that

$$y^{(i+1)} = A^{-1}/||y^{(i)}||_\infty \; i = 1, 2, \ldots$$

a recursion relation that produces a series of vectors tending to the eigenvector with smallest eigenvalue as required. Since computing A^{-1} is so difficult what one actually calculates are the solutions of

$$Ay^{(i+1)} = y^{(i)}/||y^{(i)}||_\infty$$

given some initial $y^{(i)}$. E.g. LAPACK, a public-domain linear algebra (see for example `http://www.netlib.org/lapack/`) uses 5 iterations of such a relation. Imagine you (or lapack) have done this, you can estimate the condition number then using

$$\kappa_\infty(A) \geq ||A||_\infty \frac{||y(5)||_\infty}{||Ay^{(5)}||_\infty} = \frac{\alpha v}{\omega}$$

with $\alpha = ||A||_\infty, v = ||y^{(5)}||_\infty, \omega = ||Ay^{(5)}||_\infty$. In fact the way this algorithm is set up you can write

$$\omega = ||Ay^{(5)}||_\infty = \left|\left|\frac{y^{(4)}}{||y^{(4)}||_\infty}\right|\right|_\infty = 1$$

and so

$$\kappa(A) = \alpha v$$

.