PART III - Conditioning and Stability

Lecture 12 - *Conditioning and Condition Numbers*

OBJECTIVE:
We now turn to a systematic discussion of two fundamental issues in numerical analysis: conditioning and stability.
Conditioning pertains to the sensitivity of a *mathematical problem*.
Stability pertains to the sensitivity of an *algorithm* used to solve a mathematical problem on a computer.

$\diamond$ CONDITIONING OF A PROBLEM

In a very abstract sense, solving a problem is like evaluating a function

$$y = f(x).$$

Here, x represents the input to the problem (the data), f represents the "problem" itself, and y represents its solution.

We are interested in studying the effect on y when a given x is perturbed slightly.

If small changes in x lead to small changes in y, we say the problem is *well-conditioned*.

If small changes in x lead to large changes in y, we say the problem is *ill-conditioned*.

Of course what constitutes “large” or “small” may depend on the problem
$\longrightarrow$ *it only makes sense to solve well-conditioned problems*.

Because floating-point arithmetic used by computers introduces relative errors (see Lecture 13) not absolute errors, we define conditioning in terms of a *relative* condition number.

$\diamond$ RELATIVE CONDITION NUMBER

Let δx denote a small perturbation of x, and

$$\delta f = f(x + \delta x) - f(x)$$

be the corresponding perturbation in f.
Then, the *relative condition number* $\kappa = \kappa(x)$ is defined to be

$$\kappa(x) = \lim_{\delta \to 0} \max_{\|\delta x\| \le \delta} \left(\frac{\|\delta f\|}{\|f(x)\|} \Big/ \frac{\|\delta x\|}{\|x\|} \right)$$

Or, if you just assume δx and δf are infinitesimal

$$\kappa(x) = \max_{\delta x} \left(\frac{\|\delta f\|}{\|f(x)\|} \Big/ \frac{\|\delta x\|}{\|x\|} \right)$$

$\rightarrow$ maximum value of the ratio "relative change in f" to "relative change in x".

If f has a derivative , we can write

$$\frac{\delta f}{\delta x} = \mathbf{J}(x)$$

where $\mathbf{J}$ is known as the *Jacobian* of f at x. It is the matrix of first partial derivatives of f.

e.g., suppose

$$f(x_1, x_2, x_3) = \begin{pmatrix} x_1 x_2 + \sin(x_3) + x_1^2 \\ 7 + e^{x_2} \end{pmatrix}$$

then,

$$\begin{aligned} \mathbf{J} &= \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \frac{\partial f_1}{\partial x_3} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \frac{\partial f_2}{\partial x_3} \end{bmatrix} \\ &= \begin{bmatrix} x_2 + 2x_1 & x_1 & \cos(x_3) \\ 0 & e^{x_2} & 0 \end{bmatrix} \end{aligned}$$

i.e., the (i,j) entry of $\mathbf{J}$ is $\frac{\partial f_i}{\partial x_j}$.

Note 1. *$\delta f \approx \mathbf{J}(x)\delta x$ with $\delta f = \mathbf{J}(x)\delta x$ in the limit $\|\delta x\| \to 0$.*

In terms of $\mathbf{J}$,

$$\kappa = \frac{\|\mathbf{J}(x)\|}{\|f(x)\|/\|x\|}$$

Note 2. *There is also a concept of* absolute condition number, *but this is usually less useful then the relative condition number because roundoff errors on computers are relative errors (not absolute errors); see Lecture 13.*

We say a problem is *well-conditioned* if κ is small (e.g., $\approx 1, 10, 10^2$), and *ill-conditioned* if it is large (e.g., $\approx 10^6, 10^{14}$).

Note 3. *What constitutes "large" depends on the precision you are working in!*

A general rule of thumb is that if $\kappa = 10^p$, then you cannot really trust the last p digits of your answer.

So, in single precision, where $\epsilon_{\text{machine}} \approx 10^{-8}$, $\kappa = 10^6$ is pretty ill-conditioned because you will only be able to trust the first 2 digits of your answer (this may be sufficient for some applications!).

But, in double precision, where $\epsilon_{\text{machine}} \approx 10^{-16}$, $\kappa = 10^6$ is not such a big deal.

Example 12.1 DIVISION BY 2

Consider the (trivial) problem of dividing a number by 2. This can be described by the function

$$f : x \to \frac{x}{2}$$

So,

$$\mathbf{J} = \left[\frac{\partial f}{\partial x}\right] = \frac{1}{2}$$

and

$$\kappa = \frac{\|\mathbf{J}\|}{\|f(x)\|/\|x\|} = \frac{\frac{1}{2}}{\frac{1}{2}|x|/|x|} = 1$$

$\to$ a well-conditioned problem!

Example 12.2 SUBTRACTION

Consider the problem of subtracting two numbers. This can be described by the function

$$f(x) : (x_1, x_2) \to x_1 - x_2$$

For simplicity, let $\| \cdot \| = \| \cdot \|_\infty$. Then,

$$\mathbf{J} = \left[\begin{array}{cc} \frac{\partial f}{\partial x_1} & \frac{\partial f}{\partial x_2} \end{array} \right] = \left[\begin{array}{cc} 1 & -1 \end{array} \right]$$

So,

$$\|\mathbf{J}\|_\infty = 2$$

and

$$\kappa = \frac{\|\mathbf{J}\|}{\|f(x)\|/\|x\|} = \frac{2}{|x_1 - x_2|/\max\{|x_1|, |x_2|\}}$$

So we see κ is large if $|x_1 - x_2|$ is small, i.e., $x_1 \approx x_2$. This leads us to the well-known result that subtraction of nearly equal quantities leads to large (cancellation) errors in the result.

Example 12.3 FINDING EIGENVALUES OF A NONSYMMETRIC MATRIX

This problem is often ill-conditioned

e.g.,

$$\mathbf{A} = \begin{bmatrix} 1 & 1000 \\ 0 & 1 \end{bmatrix}$$

and

$$\tilde{\mathbf{A}} = \begin{bmatrix} 1 & 1000 \\ 0.001 & 1 \end{bmatrix}$$

The eigenvalues of $\mathbf{A}$ are $\{1, 1\}$, whereas those of $\tilde{\mathbf{A}}$ are $\{0, 2\}$. *(verify!)*

$\rightarrow$ a large change in the output (eigenvalues) for a small change ($\sim 10^{-3}$) of the input ($\mathbf{A} \rightarrow \tilde{\mathbf{A}}$).

Note 4. *On the other hand if* $\mathbf{A}$ *is symmetric (or more generally, if it is* normal[1]*) then finding its eigenvalues is a well-conditioned problem.*

It can be shown that if λ and $\lambda + \delta\lambda$ are the eigenvalues of $\mathbf{A}$ and $\mathbf{A} + \delta\mathbf{A}$ respectively, then

$$|\delta\lambda| \leq \|\delta\mathbf{A}\|_2$$

[1] $\mathbf{A}$ is normal if $\mathbf{A}\mathbf{A}^T - \mathbf{A}^T\mathbf{A} = 0$.

$\longrightarrow$ using the 2-norm, we can take

$$\|\mathbf{J}\| = \max \left\| \frac{\delta f}{\delta x} \right\| = \max \left\| \frac{\delta \lambda}{\delta \mathbf{A}} \right\| = 1$$

thus

$$\kappa = \frac{1}{\|\lambda\| / \|\mathbf{A}\|_2} = \|\mathbf{A}\|_2 / |\lambda|$$

We'll come back to this in Lecture 26.

$\diamond$ CONDITION OF MATRIX-VECTOR MULTIPLICATION

Consider the problem of computing $\mathbf{Ax}$ for fixed $\mathbf{A}$ and input $\mathbf{x}$
i.e., we wish to determine the conditioning of matrix-vector multiplication for perturbations in $\mathbf{x}$ but not $\mathbf{A}$.

In this case, $\mathbf{J} = \mathbf{A}$ *(verify!)*

So, by definition,

$$\kappa = \frac{\|\mathbf{A}\|}{\|\mathbf{Ax}\| / \|\mathbf{x}\|}$$

If $\mathbf{A}$ is square and invertible, we can use

$$\mathbf{x} = \mathbf{A}^{-1}\mathbf{A}\mathbf{x}$$

to write

$$\begin{aligned} \|\mathbf{x}\| &= \|\mathbf{A}^{-1}\mathbf{A}\mathbf{x}\| \\ &\leq \|\mathbf{A}^{-1}\|\|\mathbf{A}\mathbf{x}\| \end{aligned}$$

or

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

to write

$$\kappa \leq \|\mathbf{A}\|\|\mathbf{A}^{-1}\|$$

It can be shown that we can take

$$\kappa = \|\mathbf{A}\|\|\mathbf{A}^{-1}\|$$

and also more generally for non-square or non-invertible matrices,

$$\kappa = \|\mathbf{A}\| \|\mathbf{A}^{\dagger}\|$$

The problem we have just analyzed is given $\mathbf{A}$ and $\mathbf{x}$, what is the condition number of forming $\mathbf{b} = \mathbf{A}\mathbf{x}$?

Note 5. *There is a corresponding inverse problem: given $\mathbf{A}$ and $\mathbf{b}$, what is the condition number of solving for $\mathbf{x}$? i.e., $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$*

$\longrightarrow$ we can see that formally this is the same problem as before, except with $\mathbf{A}$ replaced by $\mathbf{A}^{-1}$.

This leads to the following result:

Theorem 1. *Let $\mathbf{A} \in \mathbb{R}^{m \times m}$ be nonsingular and consider $\mathbf{A}\mathbf{x} = \mathbf{b}$. The problem of computing $\mathbf{b}$ given $\mathbf{x}$ has condition number*

$$\kappa = \|\mathbf{A}\| \frac{\|\mathbf{x}\|}{\|\mathbf{b}\|} \leq \|\mathbf{A}\| \|\mathbf{A}^{-1}\|$$

with respect to perturbations in $\mathbf{x}$.

The problem of computing $\mathbf{x}$ given $\mathbf{b}$ has the condition number

$$\kappa = \|\mathbf{A}^{-1}\| \frac{\|\mathbf{b}\|}{\|\mathbf{x}\|} \leq \|\mathbf{A}\| \|\mathbf{A}^{-1}\|$$

with respect to perturbations in $\mathbf{b}$.

◇ CONDITION NUMBER OF A MATRIX

In both cases, the inequalities in Theorem 12.1 can be made into equalities. So we define the *condition number of $\mathbf{A}$* (relative to the norm $\|\cdot\|$) as

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

$\rightarrow$ In this case the condition number applies to the matrix, *not the problem*.

As usual, if $\kappa(\mathbf{A})$ is small, $\mathbf{A}$ is said to be well-conditioned; if $\kappa(\mathbf{A})$ is large, $\mathbf{A}$ is said to be ill-conditioned.

Note 6. *In the 2-norm,* $\|\mathbf{A}\| = \sigma_1$ *and* $\|\mathbf{A}^{-1}\| = \frac{1}{\sigma_m}$ *Thus,*

$$\kappa(\mathbf{A}) = \frac{\sigma_1}{\sigma_m}$$

$\rightarrow$ This is how the 2-norm condition numbers of matrices are computed in practice.

The ratio $\frac{\sigma_1}{\sigma_m}$ can be interpreted as the eccentricity of the hyperellipse that is the image of the unit hypersphere in $\mathbb{R}^m$ (recall Lecture 4).

For rectangular $\mathbf{A} \in \mathbb{R}^{m \times n}, \quad m \geq n$, with full rank,

$$\kappa(\mathbf{A}) = \|\mathbf{A}\| \|\mathbf{A}^{\dagger}\|$$

Since $\mathbf{A}^{\dagger}$ was motivated by least-squares problem, the 2-norm condition number often makes sense in which case

$$\kappa(\mathbf{A}) = \frac{\sigma_1}{\sigma_n}$$

$\diamond$ CONDITION NUMBER OF A SYSTEM OF EQUATIONS

In Theorem 12.1, we fixed $\mathbf{A}$ and perturbed $\mathbf{x}$ or $\mathbf{b}$ $\rightarrow$ what about if we perturb $\mathbf{A}$?

Fix $\mathbf{b}$ and consider the problem

$$f : \mathbf{A} \rightarrow \mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$$

where $\mathbf{A}$ is perturbed by $\delta\mathbf{A}$. Then, $\mathbf{x}$ is perturbed by $\delta\mathbf{x}$, where

$$(\mathbf{A} + \delta\mathbf{A})(\mathbf{x} + \delta\mathbf{x}) = \mathbf{b}$$

Using $\mathbf{A}\mathbf{x} = \mathbf{b}$, we have to leading order

$$(\delta\mathbf{A})\mathbf{x} + \mathbf{A}(\delta\mathbf{x}) = 0$$

So,

$$\delta\mathbf{x} = -\mathbf{A}^{-1}(\delta\mathbf{A})\mathbf{x}$$

and

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

So,

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

Again, it can be shown that equality in the above expression can be attained. This leads us to the following result:

Theorem 2. *Fix* $\mathbf{b}$ *and consider the problem of computing* $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$ *for square, nonsingular* $\mathbf{A}$. *Then the condition number of this problem with respect to perturbations in* $\mathbf{A}$ *is*

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

Theorems 12.1 and 12.2 are of fundamental importance in numerical linear algebra
$\rightarrow$ they determine how many digits of accuracy you can expect in $\mathbf{x}$ when solving $\mathbf{A}\mathbf{x} = \mathbf{b}$.