LECTURE 1.

Start with Strang's elegant description of *errors and blunders. Errors* are unavoidable aspects of any computation, whether because of a computer's inherent limitations or our own human capacity for mistakes. *Blunders* are avoidable mistakes that come from careless interpretations of data. I expect *you* not to make blunders–it would be too much for me to expect you to make no errors!

Main example: Suppose we have a system we've analyzed–amount of two chemicals, for instance. So suppose we have $\vec{x}_t$ be the state of the system at time t, and suppose that we find

$$\vec{x}_9 = \begin{bmatrix} 1 & 1 \\ 1 & 1.00001 \end{bmatrix} \vec{x}_0.$$

Note that we seem to see that, at time 9, we expect to have approximately equal amounts of each chemical. Write the matrix above as A.

Now suppose we know how much chemical there is at time 9, and want to work backwards–given $\vec{x}_9$, how do we find $\vec{x}_0$? Well, that's easy–we just have

$$\vec{x}_0 = A^{-1}\vec{x}_9.$$

So suppose you find $\vec{x}_9 = \begin{bmatrix} 100 \\ 100 \end{bmatrix}$. Then you can easily compute

$$A^{-1} = \begin{bmatrix} -100000 & 100000 \\ 100001 & -100000 \end{bmatrix}$$

and

$$A^{-1} \begin{bmatrix} 100 \\ 100 \end{bmatrix} = \begin{bmatrix} 0 \\ 100 \end{bmatrix}$$

Great! But maybe not so great. After all, there's naturally some experimental error. But your instruments are really good, and you know that you've measured the chemicals correctly with an error of at most 0.0004 grams. So you say, well, OK, maybe $\vec{x}_9 = \begin{bmatrix} 100 \\ 100.0004 \end{bmatrix}$. That's not a very big difference. But when you work backwards, you find

$$\vec{x}_0 = A^{-1} \begin{bmatrix} 100 \\ 100.0004 \end{bmatrix} = \begin{bmatrix} 40 \\ 60 \end{bmatrix}$$

In other words, a really tiny uncertainty in measurement–less than one one-thousandth of a percent–translates into complete cluelessness about the actual value of $\vec{x}_0$. If you said confidently that $\vec{x}_0$ was $\begin{bmatrix} 0 \\ 100 \end{bmatrix}$, you would have committed a blunder!

So let's analyze this more closely. Suppose you're trying to solve the **Main problem**, $A\vec{x} = \vec{b}$. But suppose $\vec{b}$ you have a certain unknown error $\delta\vec{b}$; that is, maybe what you actually measure is $\vec{b} + \delta\vec{b}$. How far off is your computation of $\vec{x}$? Write $\vec{x}$ for the solution to $A\vec{x} = \vec{b}$ and $\vec{x}'$ for what you actually measure.

Well, you get

$$\vec{x}' = A^{-1}(\vec{b} + \delta\vec{b}) = A^{-1}\vec{b} + A^{-1}\delta\vec{b} = \vec{x} + A^{-1}\delta\vec{b}.$$

So the question is, how big is $\delta\vec{x} = A^{-1}\delta\vec{b}$?

Clearly this depends on the nature of A^{-1}. But note–we're not really so interested in the size of $\delta\vec{x}$ itself; we're more interested in the *percentage* error. An error of 1000 is no big deal if the coordinates of $\vec{x}$ are $10,000,000$.

Relative error in $\vec{b}$ is given by

$$\|\delta\vec{b}\|/\|\vec{b}\|$$

and the relative error in $\vec{x}$ by

$$\|\delta\vec{x}\|/\|\vec{x}\|$$

We say a matrix is *well-conditioned* if $\|\delta\vec{x}\|/\|\vec{x}\|$ is small whenever $\|\delta\vec{b}\|/\|\vec{b}\|$ is small.

Definition. The *condition number* of a matrix is the smallest number c such that

$$\|\delta\vec{x}\|/\|\vec{x}\| \leq c\|\delta\vec{b}\|/\|\vec{b}\|$$

for any $\vec{b}$ and $\delta\vec{b}$.

Example: in the case we did above, we had

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 1.00001 \end{bmatrix}, \vec{b} = \begin{bmatrix} 100 \\ 100 \end{bmatrix}, \delta\vec{b} = \begin{bmatrix} 0 \\ .0004 \end{bmatrix}, \vec{x} = \vec{0}100, \delta\vec{x} = 40{-}40.$$

So we find

$$\|\delta\vec{x}\|/\|\vec{x}\| = 40\sqrt{2}/100 \sim 0.565$$

and

$$\|\vec{\delta b}\|/\|\vec{b}\| = 0.0004/100\sqrt{2} \sim 0.0000282$$

In particular, we find that the condition number is at *least* $200,000$! That's just awful. A very tiny percentage error in the measurement yields a huge percentage error in the conclusion.

Silly example: take A the identity matrix. Then of course $\vec{x} = \vec{b}$ and $\vec{\delta x} = \vec{\delta b}$, so the condition number is 1.

So if the matrix in our problem has a small condition number, we should feel happy–we know that small errors won't propagate into big ones. So how do we compute the condition number? Actually, it takes some doing–but it's easy in the very often encountered case where A (whence also A^{-1}) is a positive definite symmetric matrix.

We want to try to work out a condition number. So we ask: How big can $\vec{\delta x} = A^{-1}\vec{\delta b}$ be, compared to $\vec{b}$ How small can $\vec{x} = A^{-1}\vec{b}$ be, compared to $\vec{b}$?

Groupwork. Let A be a matrix such that

$$A^{-1} = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}.$$

Suppose $\vec{\delta b}$ has length c. How long can $A^{-1}\vec{\delta b}$ be? Suppose $\vec{b}$ has length c. How short can $\vec{x} = A^{-1}\vec{b}$ be?

Maybe let them think about this a while. Try some points.

Hint: To say $\|\vec{b}\| = 1$ is to say

$$\vec{b}^T\vec{b} = c$$

Now $A\vec{x} = \vec{b}$, so we have

$$\vec{x}^T A^T A\vec{x} = \vec{x}^T \begin{bmatrix} 5 & -4 \\ -4 & 5 \end{bmatrix} = c.$$

What do the solutions $\vec{x}$ to the above equation look like? And among those points, which is shortest?

OK, so what we should find at the end of all this is that the way to make $\vec{x}$ long is to have $\vec{b}$ point in the direction of an eigenvector of A with an eigenvalue as close as possible to 0.

Fact. Let λ_n be the largest eigenvalue of A. Then $|\vec{x}\| \geq \lambda_n^{-1}\|\vec{b}\|$. Let λ_n be the smallest eigenvalue of A. Then $\|\vec{\delta x}\| \leq \lambda_1^{-1}\|\vec{\delta b}\|$.

In particular, we have that the condition number is given by

$$c = \lambda_n/\lambda_1.$$

"A symmetric matrix is well-conditioned if the ratio between the largest and smallest eigenvalue is not too great."

This might seem to be the issue in general. After all, the 2×2 matrix A we started with has an eigenvalue of 2 and an eigenvalue which is about 5×10^{-5}. So is the above motto true for *all* matrices? No, not quite. We can write down a matrix like

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

which has the eigenvalues very close together–in fact equal! So is this matrix well-conditioned? No. Because

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

and we find

$$A^{-1} \begin{bmatrix} 1000 \\ 0 \end{bmatrix} = 10000, A^{-1} \begin{bmatrix} 1000 \\ 1 \end{bmatrix} = 01.$$

A small percentage change in the input produced a rather drastic change in the output. So the condition number is big. At least a thousand or so.

We sum up what we have to do in the following question.

Question. Let A be a matrix. Suppose $\|\vec{v}\|$ is known. How large can $\|A\vec{v}\|$ be? How small can $\|A\vec{v}\|$ be?

If there's some time today, we'll investigate this question by drawing some pictures of ellipses.