261: 1st November

So far, we have looked at the equation $Ax = b$ and how to calculate x, when we know both A and b. However, we have not yet considered whether any uncertainty in the values of A or b would affect our answer.

Uncertainty

Uncertainty, usually comes from one of two places. Firstly, there's experimental uncertainty. For example, if we were studying the aerodynamics of a wing and the vector b represented the velocity of the air at some point near the wing, then there would be a limit to how accurately we could measure b. If the sensor measured the velocity to three decimal places, then this would leave an error margin of $\pm 0.0005 m/s$.

The second most common source of uncertainty comes from the computer. When a program is run, each variable is given a fixed number of bytes, in which to represent a value. This means that for each number, there's a limited number of digits available to represent it. If we try to give the variable a value with more significant figure than it can handle, the computer automatically rounds it. In doing so, it introduces a rounding error.

So we must ask the question, "Will these errors change the value of x and if so, by how much?"

The answer depends on the problem that you're looking at. In some cases, the errors will have a dramatic effect and in others, it will be small. This tutorial will focus on ways of quantifying the susceptability of a problem to these errors.

Just to convince yourself, try the following. In (a) and (b) below, you are given A, x, b and δb. In both cases, x is the solution to $Ax = b$. What you should do is try to work out what x would be if you replace b with $b + \delta b$. By adding δb to b, you are effectively changing b by a small amount. It is easiest, if you use Cramer's rule.

a)

$$A = \begin{pmatrix} 1000 & 1 \\ 1 & 2 \end{pmatrix} \quad x = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad b = \begin{pmatrix} 1001 \\ 3 \end{pmatrix} \quad \delta b = \begin{pmatrix} 0.1 \\ 0.1 \end{pmatrix}$$

b)

$$A = \begin{pmatrix} 100 & 0.5 \\ 1 & 10 \end{pmatrix} \quad x = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad b = \begin{pmatrix} 100.5 \\ 11 \end{pmatrix} \quad \delta b = \begin{pmatrix} -0.5 \\ -0.5 \end{pmatrix}$$

When it comes to quantifying uncertainty, the best you can do is put a limit on it. With this in mind, we will have a look at a very common equation that is often used; the Cauchy-Schwartz inequality.

Cauchy-Schwartz Inequality

Imagine a vector in two dimensions. It's sitting on the $x - y$ plane. We will call this vector a. Now imagine a second two-dimensional vector also sitting in the $x - y$ plane. This is vector b. The length of a is $|a|$ and of b is $|b|$. The Cauchy-Schwartz inequality says that the dot product of the two vectors can never be greater than the product of their lengths. This makes some sense as the dot product is defined to be $|a||b|cos\theta$ where θ is the angle between the vectors. Obviously, $cos\theta$ has values in the range $[-1, 1]$ so $|a||b|cos\theta$ can never be greater than $|a||b|$.

The application of this inequality does not stop with vectors. We can apply the same idea to matrices, if we replace the dot product operation with matrix vector multiplication. This would give us $|Ax| \leq \|A\||x|$. Of course, the definition of $\|A\|$ is not quite as clear as $|x|$, but by using the definition of $|x|$, we can work out how to define $\|A\|$.

If $|Ax| \leq \|A\||x|$ and $Ax = b$, then $\|A\| \geq \frac{|b|}{|x|}$. Calculate $\frac{|b|}{|x|}$ for each of the following.

c)

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \quad x = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

d)

$$A = \begin{pmatrix} 0 & 2 \\ 2 & 4 \end{pmatrix} \quad x = \begin{pmatrix} 1 \\ 3 \end{pmatrix}$$

Now, the length of a vector is defined as $\sqrt{(\sum_{i=1} x_i^2)}$, but it is not the only definition that used for $|x|$. A more general definition is $(\sum_{i=1} x_i^p)^{\frac{1}{p}}$. This is known as the l_p-norm and the value of p is a matter of choice. The definition of the length of a vector is the same as the l_2-norm.

A commonly used norm is the l_∞-norm, which is known as just the infinity norm. Although it is difficult to see from the expression for the general p-norm, this is equal to $max\{|x_i|\}$.

To go with the new definitions of $|x|$, there are also new definitions of $\|A\|$. In the case of the infinity norm, this is equal to the value of the row with the greatest sum, ie $max\{\sum_{j=1}^{n} |a_{ij}|\}$.

Measures of Instabiltiy

Now lets have a proper look at the effects of small variations in b and A. Let us say that b is changed by a small amount, δb, and that δb changes the solution, x, by a small amount, δx. This would mean that

$$A(x + \delta x) = b + \delta b$$

Now, if $Ax = b$, then we can subtract Ax from the left hand side and b from the right hand side to get

$$A\delta x = \delta b$$

and so

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

Although this is a handy relation to know, it would be more useful to know the size of δx relative to x., ie to know if x changes by a little or a lot.

So, if $\delta x = A^{-1}\delta b$, then using the Cauchy-Schwartz inequality, we have that

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

We also know that

$$|b| = |Ax| \leq \|A\||x|$$

so

$$\frac{|b|}{\|A\|} \leq |x|$$

If we combine this with the equation for $|\delta x|$, we get

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

which neatly relates $\frac{|\delta x|}{|x|}$ to $\frac{|\delta b|}{|b|}$.

$\|A^{-1}\|\|A\|$ is known as tha condition number of A and is denoted by $C(A)$. The larger it is the more susceptable the problem is, to errors.

Now lets see what happens when there is an error in A,

$$(A + \delta A)(x + \delta x) = b$$

Now we subtract out $Ax = b$ to give

$$\delta A(x + \delta x) + A\delta x = 0$$

Rearranging this we get

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

If we use the Cauchy-Schwartz inequality again, then we obtain

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

so

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

If we introduce a factor of $\frac{\|A\|}{\|A\|}$ then we have

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

Rearranging, we get

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

And so, we end up with

$$\frac{|\delta x|}{|x|} \leq C(A)\frac{\|\delta A\|}{\|A\|}$$

where C(A) is the condition number again.

Using the infinity norms, calculate the condition number of the following matrices.

e)

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$$

f)

$$\begin{pmatrix} 1 & 0 \\ 2 & 5 \end{pmatrix}$$

g)

$$\begin{pmatrix} 1000 & 1 \\ 1 & 1 \end{pmatrix}$$