Stability

The last topic we'll look at before moving on to multivariable calculus is <u>stability</u>. Floating-point operations are inherently imprecise, but some matrices will magnify this imprecision when we solve them. We'll cover two techniques here: one, how to refine initial results when they are imprecise, and two, how to determine when a matrix is dangerous.

Error reduction

(from Lengyel's book on 3D graphics math)

Suppose we solve $Mx = r$ for x and get x_0, which is a little off: $Mx_0 = r_0$ and $r_0 \neq r$. We'll call the differences $x_0 - x = \Delta x$ and $r_0 - r = \Delta r$. Then $M(x + \Delta x) = r + \Delta r$, and subtracting $Mx = r$ from both sides gives $M\Delta x = Mx_0 - r$.

The RHS is known, so we can solve for our original error Δx, which is probably smaller in magnitude than x and therefore smaller in absolute error. We can repeat this process still more, for more accuracy.

Condition number

Rule of thumb: you lose a digit of precision for every digit in the condition #

A matrix's condition number tells us how much it will amplify any error in the input into error in the solution. Nearly singular matrices have large condition numbers.

Let the system we're solving be $Ax=b$, and let ϵ be some error in b. The condition number is a ratio of two ratios:

$$\frac{\;\dfrac{|A^{-1}\epsilon|}{|A^{-1}b|}\;}{\dfrac{|\epsilon|}{|b|}}$$

The ratio of the size of $A^{-1}\epsilon$ to the size of the true solution $(x=A^{-1}b)$

The ratio of the size of the error to the size of the original vector

So we're measuring how much the error gets amplified relative to the vector.

Some algebra shows this is $\dfrac{|A^{-1}\epsilon|}{|\epsilon|}\cdot\dfrac{|b|}{|A^{-1}b|}=\dfrac{|A^{-1}\epsilon|}{|\epsilon|}\cdot\dfrac{|Ax|}{|x|}$.

If we're worried about the <u>max</u> value that this could have, there's a handy concept we haven't explored: the <u>norm</u> of a matrix, written like $\|A\|$, which is the maximum amount the matrix can increase the size of a vector. Looking at the above, if we want

$$\max_{\epsilon}\frac{|A^{-1}\epsilon|}{|\epsilon|}\cdot\max_{x}\frac{|Ax|}{|x|},$$

this is $\underbrace{\|A^{-1}\|\cdot\|A\|}_{\text{condition number}}$. So all we need is to figure out how much either of these two matrices could possibly scale vectors.

That sounds tricky, but we're allowed to use a definition of length for our vectors that isn't the standard "square root of squared components." Different ways of measuring vector length are also called "norms."

and there are three to be aware of:

Vectors

l_2 norm: The familiar vector length, $\sqrt{x_1^2 + \cdots + x_n^2}$

l_1 norm: Just sum absolute values of components: $|x_1| + \cdots + |x_n|$

l_∞ norm: Just return the component with largest absolute value, $\max_i |x_i|$.

Those definitely don't give the same results. But using any of these norms produces a reasonable condition number, and the corresponding norms for the matrices are rather easier to compute for the l_1 and l_∞ norms.

l_2 norm for matrices: $\|A\|_2 = \max_i \sqrt{\lambda_i}$ where λ_i is an eigenvalue of A^TA

l_1 norm for matrices: Is just the sum of absolute values of a column, $\max_j \sum_{i=1}^{n} |a_{ij}|$

l_∞ norm for matrices: Just the max sum of absolute values of a row, $\|A\|_\infty = \max_i \sum_{j=1}^{n} |a_{ij}|$

So if $A = \begin{bmatrix} 1 & 4 & 7 \\ -2 & 5 & 8 \\ 3 & 6 & 9 \end{bmatrix}$ then $\|A\|_1 = 7+8+9 = 24$
and $\|A\|_\infty = 3+6+9 = 18$

but to find $\|A\|_2$ we'd need to calculate eigenvalues — and we'd get a number that's not too far off from those.

So if we calculate $\|A^{-1}\| \cdot \|A\|$ for some norm, we get an idea as to how much the matrix will amplify error (just # of digits gives some idea). There are ways to estimate condition number without calculating inverse, but we'll skip them.

Major midterm topics (in chron order)

Gauss-Jordan elimination for finding
- solutions to linear systems
- rank
- inverses

Framing dynamical systems as matrices w/ exponentiation
Dot products & finding angles & lengths
Composing matrices (in the right order)
Deciding whether vectors are linearly independent
Change of basis
Finding dimension of a span
Understanding how perceptrons classify - what they can & can't
Projections onto lines & planes, including to fit lines to data
Purpose of determinant
Eigenvectors & eigenvalues: what they mean, how to find them
Kernel, nullity, range space, and relationship to rank
Differences between singular and nonsingular matrices

Equations provided

$$a \cdot b = a_1 b_1 + \dots + a_n b_n = |A||B|\cos\theta$$

$$\text{proj}_{\vec{s}} \vec{v} = \frac{\vec{v} \cdot \vec{s}}{\vec{s} \cdot \vec{s}} \vec{s}$$

$$\text{proj}_A \vec{v} = A\vec{c} = A(A^T A)^{-1} A^T \vec{v}$$

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc \qquad \begin{vmatrix} a & b & c \\ d & e & f \\ h & i & j \end{vmatrix} = \dots$$

$$\det(A - \lambda I) = 0$$

R②

$$\begin{bmatrix} 1 & 0 & 1 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$ Rank? Inverse?

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & 0 & 0 \\ -1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{array}\right] \to \left[\begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{array}\right]$$

$$\downarrow$$

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 & -1 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{array}\right]$$

$$\downarrow$$

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 0 & -1 \\ 0 & 1 & 0 & 1 & 1 & -1 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{array}\right]$$

Eigenvalues?

$$\begin{vmatrix} 1-\lambda & 0 & 1 \\ -1 & 1-\lambda & 0 \\ 0 & 0 & 1-\lambda \end{vmatrix} = (1-\lambda)^3 = 0 \to \text{all } 1$$

Eigenvectors?

$$\left[\begin{array}{ccc|c} 0 & 0 & 1 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

What does this mean?

$z = 0$

$x = 0$

(y is anything)

$$c\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

Check:

$$\begin{bmatrix} 1 & 0 & 1 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & 0 & -1 \\ 1 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \checkmark$$

Check:

$$\begin{bmatrix} 1 & 0 & 1 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \checkmark$$

Are $\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$ a valid basis?

Are they orthogonal? (the last two are)

How could I replace $\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$ with a vector orthogonal to both of the others?

$$\frac{\begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}}{\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}} \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ -1 \\ 0 \end{bmatrix} \quad \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} - \begin{bmatrix} 0 \\ -1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

$$\frac{\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}}{\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}} = \frac{1}{2}\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} - \frac{1}{2}\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} \\ 0 \\ -\frac{1}{2} \end{bmatrix}$$

How would we express $\begin{bmatrix}1\\2\\3\end{bmatrix}$ in our other basis? $, \begin{bmatrix}1\\-1\\0\end{bmatrix} \begin{bmatrix}0\\1\\0\end{bmatrix} \begin{bmatrix}1\\0\\1\end{bmatrix}$

$$c_1\begin{bmatrix}1\\-1\\0\end{bmatrix} + c_2\begin{bmatrix}0\\1\\0\end{bmatrix} + c_3\begin{bmatrix}1\\0\\1\end{bmatrix} = \begin{bmatrix}1\\2\\3\end{bmatrix}$$

$$A\vec{c} = \begin{bmatrix}1\\2\\3\end{bmatrix}$$

Do I need to redo my G-J work, or is there a shortcut?

$$\vec{c} = A^{-1}\begin{bmatrix}1\\2\\3\end{bmatrix} = \begin{bmatrix}1 & 0 & -1\\1 & 1 & -1\\0 & 0 & 1\end{bmatrix}\begin{bmatrix}1\\2\\3\end{bmatrix} = \begin{bmatrix}-2\\0\\3\end{bmatrix}$$

Check: $-2\begin{bmatrix}1\\-1\\0\end{bmatrix} + 0\begin{bmatrix}\dots\end{bmatrix} + 3\begin{bmatrix}1\\0\\1\end{bmatrix} = \begin{bmatrix}1\\2\\3\end{bmatrix}$

What's a matrix for swapping the x and z coordinates?

$$\begin{bmatrix}0 & 0 & 1\\0 & 1 & 0\\1 & 0 & 0\end{bmatrix}$$

What's the rank?

What's the inverse?

~~What must its eigenvalues be?~~ Name an eigenvalue and eigenvector.

(actual computation $\lambda^2(1-\lambda)-(1-\lambda)=-\lambda^3+\lambda^2-(1)=0 \Rightarrow 1 \text{ or } -1$)

$$\begin{bmatrix}1 & 2 & 3\\1 & 2 & 4\\1 & 2 & 5\end{bmatrix}$$

Can we tell the rank just by looking at this?

Does it have an inverse?

What's a general solution to the homogeneous matrix? (use x y z)

R (4)

I want to fit a line of form $y = mx + b$ to datapoints $(3, 9), (4, 13), (5, 14)$

- What is my vector I am projecting, and what matrix defines the plane? $\vec{v} = \begin{bmatrix} 9 \\ 13 \\ 14 \end{bmatrix}$ $A = \begin{bmatrix} 3 & 1 \\ 4 & 1 \\ 5 & 1 \end{bmatrix}$

What's an equation for my coefficents m, b if the whole thing is $proj_m \vec{v} = A(A^TA)^{-1}A^T\vec{v}$?

$= A\vec{c}$

Perceptrons: What can they learn and what can't they? Exampls of each category?