Condition Number, When to Approximate for Ax = b?

Kerem Turgutlu

When a matrix is not solvable in the form of the equation Ax = b, or in detail if it has infinitely many solutions we may introduce a method in order to approximate a solution to this unsolvable system. Which is basically to add noise to b or in other words to change the direction of b just as ϵ so that there will be one unique solution for Ax = b.

This seems simple and straightforward but it may cause very unwanted results if it's not carefully investigated.

Here is a naive example how changing b by ϵ effects the change in x, which we call δ.

$$Ax = b$$
$$A(x + \delta) = b + \epsilon$$
$$A\vec{x} + A\vec{\delta} = \vec{b} + \vec{\epsilon}$$

We want $||\delta||$ to be small while $||\epsilon||$ is big.

Or inequality for good approximation would be $||\delta|| \leq ||\epsilon||$, but we should normalized them for better comparison, or in other words we shoul look at their marginal changes.

$$\frac{||\delta||}{||x||} \leq \frac{||\epsilon||}{||b||}$$

$$\frac{||b||}{|||x||} \frac{||\delta||}{||\epsilon||} \leq 1$$

How do we define these norms:

$$Ax = \lambda x$$
$$b = \lambda x$$

$$||b|| = ||\lambda||||x||$$
$$\frac{||b||}{||x||} = ||\lambda||$$

$$||\lambda_{min}|| \leq \frac{||b||}{||x||} \leq ||\lambda_{max}||$$

Similarly:

$$A\delta = \epsilon$$

$$||\delta|| = \frac{||\epsilon||}{||\lambda||}$$
$$\frac{||\epsilon||}{||\delta||} = ||\lambda||$$

$$||\lambda_{min}|| \leq \frac{||\epsilon||}{||\delta||} \leq ||\lambda_{max}||$$
$$\frac{1}{||\lambda_{min}||} \geq \frac{||\delta||}{||\epsilon||} \geq \frac{1}{||\lambda_{max}||}$$

Then:

$$\frac{||b||}{|||x||}\frac{||\delta||}{||\epsilon||} \leq \frac{||\lambda_{max}||}{||\lambda_{min}||}$$

This upper bound is our condition number. As we stated at the beginning a number less than or close to 1 makes a matrix well conditioned.

Problem

What if matrix A is not diagonalizable or symmetric, then eigenvalues might be all 0s and this boundary won't be defined. Actually we will define this ineqaulity in a different way so that it will be applicable to all matrices out there.

If matrix A is diagonalizable, $A = P\Lambda P^{-1}$ then:

- Λ : n diagonal eigenvalues.
- P: n independent eigenvectors.
- $||Ax|| = ||\lambda||||x||$
- $||A|| = ||\lambda_{max}||$
- $||A^{-1}|| = ||\lambda_{min}||$

Defining Norm of a Matrix:

- $||A|| = max_{||x|| \neq 0} \frac{||Ax||}{||x||}$

By taking the square of the norm we see a relationship with SVD.

$$||A||^2 = \frac{||Ax||^2}{||x||^2}$$
$$||A||^2 = \frac{||x^T A^T Ax||}{||x^T x||}$$
$$||A||^2 = \frac{||x^T A^T Ax||}{||x^T x||}$$

Since $A^T A$ is always symmetric:

$$||A||^2 = ||\lambda_{A^TA}||_{max}$$

$$||A|| = \sqrt{||\lambda_{A^TA}||_{max}}$$

From SVD we know that $A^TA = V\Sigma^2V^T$:

$$||A|| = max\{\sigma_A\}$$

$$||A^{-1}|| = \frac{1}{min\{\sigma_A\}}$$

$$C.N. = \frac{max(\sigma_A)}{min(\sigma_A)}$$

In conclusion, condition number of any given matrix A is the ratio of largest singular value to smallest.