Lecture 10: Conditioning and Stability

Amos Ron

University of Wisconsin - Madison

February 22, 2021

Outline

1. Conditioning
   - Introduction to the notion of conditioning
   - The condition number of a matrix

2. Stability
   - Stability of factorization algorithms
   - Stability of QR-factorization

Outline

1. Conditioning
   - Introduction to the notion of conditioning
   - The condition number of a matrix

2. Stability
   - Stability of factorization algorithms
   - Stability of QR-factorization

Blank page

Solving a linear system in a finite precision environment

We are given a square invertible $m \times m$ A, and two equations:

- An equation
$$Ax = b_0, \quad x_0 := A^{-1}b_0.$$
- An equation
$$Ax = b_1, \quad x_1 := A^{-1}b_1,$$

with $||x_0||$ and $||x_1||$ of about the same order.

Solving a linear system in a finite precision environment

We are given a square invertible $m \times m$ A, and two equations:

- An equation

$$Ax = b_0, \quad x_0 := A^{-1}b_0.$$

- An equation

$$Ax = b_1, \quad x_1 := A^{-1}b_1,$$

with $||x_0||$ and $||x_1||$ of about the same order.

Question:

How large/small can the ratio

$$\frac{||b_0||}{||b_1||}$$

be?

Solving a linear system in a finite precision environment

Why should we care about such question?
Assume, for example, that

$$\frac{||b_0||}{||b_1||} = 10^{12}$$

Choose $b_2 = 10^{-4} b_1$. Then

- Denote $x_2 := A^{-1} b_2$, then $x_2 = 10^{-4} x_1$, therefore, more or less,

$$||x_2|| \approx 10^{-4} ||x_0||.$$

- But:

$$||b_2|| = 10^{-4} ||b_1|| = 10^{-16} ||b_0||.$$

Solving a linear system in a finite precision environment

Summary of the current discussion

- We assumed that, for the given A, there exist x_0, x_1, such that
$$||x_0|| \approx ||x_1||, \quad \text{and} \quad ||Ax_0|| \approx 10^{12}||Ax_1||.$$
- We concluded that there exist b_0, b_2, such that
$$||b_0|| \approx 10^{16}||b_2|| \quad \text{and} \quad ||A^{-1}b_0|| \approx 10^{4}||A^{-1}b_2||.$$

Solving a linear system in a finite precision environment

Summary of the current discussion

- We assumed that, for the given A, there exist x_0, x_1, such that
$$||x_0|| \approx ||x_1||, \quad \text{and} \quad ||Ax_0|| \approx 10^{12}||Ax_1||.$$
- We concluded that there exist b_0, b_2, such that
$$||b_0|| \approx 10^{16}||b_2|| \quad \text{and} \quad ||A^{-1}b_0|| \approx 10^{4}||A^{-1}b_2||.$$

This is a disaster!

Consider the two equations $Ax = b_0$, and $Ax = b_0 + b_2$. In Matlab, the two systems are the same. But the solutions of the two systems are very different!

The definition of condition number

Definition: Condition Number

A is any matrix (not neccesarily square). $||\cdot||$ is some norm.

$$\mathrm{cond}(A) := \frac{\max\{||Av|| : ||v|| = 1\}}{\min\{||Av|| : ||v|| = 1\}}.$$

The definition of condition number

Definition: Condition Number

A is any matrix (not neccesarily square). $||\cdot||$ is some norm.

$$\text{cond}(A) := \frac{\max\{||Av|| : ||v|| = 1\}}{\min\{||Av|| : ||v|| = 1\}}.$$

Note: If A is square and invertible, then

$$\text{cond}(A) = ||A||||A^{-1}||.$$

The 2-condition number

Example: A is 2X2

$$\mathrm{cond}(A) = s(1)/s(2).$$

The 2-condition number

Example: Q is orthogonal

$$\mathrm{cond}(Q) = 1.$$

Conditioning and linear systems

Theorem

A is square invertible. $b_1, b_2 \in R^m$, and $x_i = A^{-1}b_i$, $i = 1, 2$. $||\cdot||$ is some norm. Then:

$$\frac{||x_1 - x_2||}{||x_1||} \leq \text{cond}(A)\frac{||b_1 - b_2||}{||b_1||}.$$

Conditioning and linear systems

Theorem

A is square invertible. $b_1, b_2 \in R^m$, and $x_i = A^{-1}b_i$, $i = 1, 2$. $||\cdot||$ is some norm. Then:

$$\frac{||x_1 - x_2||}{||x_1||} \leq \text{cond}(A)\frac{||b_1 - b_2||}{||b_1||}.$$

Proof: We just need to find two vectors v_1, v_2 such that

$$\frac{||x_1 - x_2||}{||x_1||}\frac{||b_1||}{||b_1 - b_2||} = \frac{||Av_1||}{||v_1||}\frac{||v_2||}{||Av_2||}.$$

However, the LHS above is

$$\frac{||x_1 - x_2||}{||b_1 - b_2||}\frac{||b_1||}{||x_1||} = \frac{||Ax_1||}{||x_1||}\frac{||x_1 - x_2||}{||A(x_1 - x_2)||}.$$

Outline

1. Conditioning
   - Introduction to the notion of conditioning
   - The condition number of a matrix

2. Stability
   - Stability of factorization algorithms
   - Stability of QR-factorization

Condition numbers in the context of factorization algorithms

When solving $Ax = b$, we factored $A = BC$. We need now to solve

- $By = b$.
- $Cx = y$.

So, the condition numbers that matter are of B and of C!

Condition numbers in the context of factorization algorithms

When solving $Ax = b$, we factored $A = BC$. We need now to solve

- $By = b$.
- $Cx = y$.

So, the condition numbers that matter are of B and of C!
Not hard to show:

$$\mathrm{cond}(A) \leq \mathrm{cond}(B)\mathrm{cond}(C).$$

The problem is that it is possible that

$$\mathrm{cond}(A) << \mathrm{cond}(B)\mathrm{cond}(C).$$

In this case the factorization algorithm is unstable.

Stability of QR-factorization

Theorem: QR-factorization is as stable as it gets!

Assume $A = QR$, then:

- $\mathrm{cond}(A) = \mathrm{cond}(R)$.
- $\mathrm{cond}(Q) = 1$.