STA 141C - Big Data & High Performance Statistical Computing **Spring 2022**

Lecture 8: Iterative method for solving linear equations

Lecturer: Bo Y.-C. Ning *February 1, 2022*

Disclaimer: *My notes may contain errors, distribution outside this class is allowed only with the permission of the Instructor.*

Announcement

- HW 2 due this week
- Final project proposal announced

Last time

- QR method
- Sweep operator
- Summary of numerical linear algebra

Today

- Condition number

1 Condition number

A condition number for a matrix and computational task measures how sensitive the answer is to perturbations in the input data and to roundoff errors made during the solution process.

For a nonsingular matrix $A \in \mathbb{R}^{n\times n}$, consider estimating the linear equation $Ax = b \Rightarrow x = A^{-1}b$. We want to know how the solution changes with a small perturbation of the input b (or A)

Let $\tilde{b} = b + \Delta b$, then

$$\tilde{x} = A^{-1}(b + \Delta b) = x + \Delta x,$$

thus,

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

Since $b = Ax$, we have $\frac{1}{\|x\|} \leq \|A\|\frac{1}{\|b\|}$. This leads to

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

We thus call $\kappa(A) = \|A\| \|A^{-1}\|$ the condition number for matrix inversion.

The condition number depends on matrix norm being used. In general, κ_p means condition number defined using p-norm. Large condition number is "bad".

Below are a list of some useful facts for the condition number:

$$\begin{aligned}
\kappa(A) &= \kappa(A^{-1}) \\
\kappa(cA) &= c\kappa(A) \\
\kappa(A) &\geq 1 \\
\kappa_1(A) &= \kappa_\infty(A') \\
\kappa_2(A) &= \kappa_2(A') = \frac{\sigma_1(A)}{\sigma_n(A)} \\
\kappa_2(A'A) &= \frac{\lambda_1(A'A)}{\lambda_n(A'A)} = \kappa_2^2(A) \geq \kappa_2(A)
\end{aligned}$$

The last point says that the condition number of $A'A$ can be much larger than that of A.

Normalization can help improve the condition number. There are several different normalization techniques, including log-transformation, z-score, and scaling ($\tilde{a} = \frac{a - \text{mean}(a)}{\text{sd}(a)}$.).

More examples can be found [here]