Algorithmique numérique - ANUM

Academic year 2022-23

Additional Notes

Instructor : Prof. Stef Graillat

written by Hayato Ishida

*This note only contains additional information to lecture slides. This note may be useful only if use with lecture slides.

Contents

I. Possibly useful study materials

1. Scientific Computing, An Introductory Survey, Michael T. Heath, McGraw-Hill, 2002
   Available online (I also have an pdf copy)

2. A First Course in Numerical Methods, Uri M. Ascher, Chen Greif, SIAM, 2011: should be available at the library on campus

3. Introduction to Scientific Computing and Data Analysis, Mark H. Holmes, Springer, 2016: should be available at the library on campus

II. Lecture 1: Introduction to MATLAB and to floating-point arithmetic

<u>Rounding</u>

- *Machine numbers*: Real numbers that are exactly representable in a given floating-number system.

- In other cases, it is approximated by a "nearby" floating-point number. This approximation is called *rounding* and the error caused by this approximation is called *rounding error*.

- The most accurate rounding method is *rounding to nearest* but it is more expensive to implement. It is also a default set by IEEE.

<u>Cancellation*</u>

- An source of rounding error that occurs with subtraction.

- It occurs when subtracting two nearby numbers. the most significant digits in operand matches and cancel each other

- Suppose a and b already contain some rounding error, perform subtraction on these number leads to large error. Example : consider three floating-number $a = 1.22, b = 3.34, c = 2.27$ and perform $b^2 - 4ac$. The exact result of this is 0.0292, however, with rounding: $b^2 = 11.1556$ round to 11.2 and $4ac = 11.0776$ round to 11.1, then result of subtraction is 0.1. *note that this is just a example and result might differ in different settings.

<u>Absorption</u>

- with floating-point we lose some algebraic property of real number. For instance, we might face the situation where:

$$(a+b)+c \neq a+(b+c)$$

III. Lecture 2: Matrix computation

- Transpose: $a_{ij} \to a_{ji}^T$

- SVD can solve most of problems but it is quite expensive. It referred as Swiss knife

- Between dense and sparse, there is no specified boundry.

Algebraic properties of SVD:

- SVD can be rewritten as follow by writing $V = [v_1, v_2, ..., v_n]^T$, and: $U = [u_1, u_2, ..., u_n]^T$:

$$A = U\Sigma V^T = \sum_{i=1}^{r} \sigma_i u_i v_i^T$$

Perron-Frobenius theorem:

(1) If $A \in \mathbb{R}^{n\times n}$ is nonnegative then

- $\rho(A)$ is an eigenvalue of A.

- there is a nonnegative eigenvector x such that $Ax = \rho(A)x$.

(2) If $A \in \mathbb{R}^{n\times n}$ is nonnegative and irreducible then

- $\rho(A)$ is an eigenvalue of A.

- $\rho(A) > 0$

- there is a nonnegative eigenvector x such that $Ax = \rho(A)x$.

- $\rho(A)$ is a simple eigenvalue.

(3) If $A \in \mathbb{R}^{n\times n}$ is positive then (2) holds and, in addition, $|\lambda| < \rho(A)$ for any eigenvalue λ with $\lambda \neq \rho(A)$.

Stochastic matrix:

A matrix $P \in \mathbb{R}^{n\times m}$ said to be stochastic if:

$P \geq 0$ and,

$$\sum_{j=1}^{n} P_{ij} = 1 \quad \text{all} \quad i = 1, 2, ..., n$$