1 Slide 125

Which of the following two ways to compute a polynomial is stable? Which one is the more efficient algorithm?

1. $p(x) = a_0 + a_1 x + \cdots + a_n x^n, x \in \mathbb{R}$
2. $p(x) = a_0 + x(a_1 + x(a_2 + x(a_3 + \cdots + x(a_{n-1} + xa_n) \ldots)))$

Solution:

2 Slide 135

Compute the condition number for the following two matrices.

$$A = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix} \qquad B = \begin{bmatrix} 168 & 113 \\ 113 & 76 \end{bmatrix}$$

Solution:

3 Slide 135

In a genetic study, m study subjects have been genotyped at n bi-alleleic loci (i.e. marker scores $X = 0, 1$ or 2). The goal is to compute the genetic variance/co-variance matrix. Describe how you would compute it and why. Discuss also the complexity and memory storage requirements. You might discuss a standard way first and then something cool.

Solution:

Reading:
`http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3397261/`

4

Please write matrix (1) in triplet format and csr format

5

Given 2 sparse $(n \times n)$-times matrices,A and B. How is the expected sparsity of $A + B$ and $A \times B$

6

For genotype matrices, define a new csr format that takes the special data structure of genotype data into account.

- Using simulation studies (by drawing from the provided MAF data),how much storage space to you save on average?
- At least how much storage space do you save in 95% of thecases/simulation studies?