An Analysis of Numerical Practice Problems towards Multi-Linear Algebra

Alok Tripathi*

Associate Professor, Department of Mathematics, Galgotias University, Greater Noida, Uttar Pradesh, India

Abstract – Numerical multi-linear algebra, in which instead of matrices and vectors the higher-order tensors are considered in numerical viewpoint, is a new branch of computational mathematics. Although the linear numerical algebra is an extension of it, it has many essential differences and more difficulties than the linear numerical algebra. This paper presents an incomplete survey of state-of-the-art knowledge on this issue and shows new trends in further research. Our survey also includes an important part of a detailed bibliography. A new branch of computer math is numerical multilinear algebra. It deals with the numerical handling by replacing matrix of higher-order tensors. Various computational issues related to the higher order tensors are included, such as decomposition of the tensor, tensor range calculation, own-value computation of the tensor, lower rank tensor approximations, numerical stability, and tensor calculation disturbance analysis, etc. This new business branch has a strong practical background and broad applications in the fields of digital image restore psychometrics, chemometrics, econometrics and multi-way data analysis.

Key Words – Numerical Multi-Linear Algebra, Higher Order Tensor, Tensor Decomposition, Lower Rank Approximation of Tensor, Multi-Way Data Analysis

- X - - - - - - - - - - - - -- - - - - - - - - - - - - - - - - - -

INTRODUCTION

The traditional Multi-linear algebra is branch and shows how the operations of tensor R work (over a commutative ring). They discuss algebra, algebra, symmetrical algebra, co-algebra and algebra of hop, etc. In modern multi-way data analysis and signal processing, however, more features of a higher-order tensor are needed. How the tensor is broken down into a whole of vector devices, how the tensor is approximated with a lower range tensor, how the tensor itself and its specific values are calculated and how higher tensors range are applied when blind source is separated (BSS). These are all a modern branch of machine math—a multi-linear algebra. Nearly all changing systems can describe differential equations. Science and engineering, business, social science, biology, enterprise, health and more are omnipresent. Several mathematicians study the nature of these equations over hundreds of years and many advanced solutions are available. Differential equations often contain systems that are so complex or systems that prevent traceability of a strictly analytical solution. In these complex systems, computer simulations and numerical procedures are helpful. Techniques for solving differential equations based on numerical approximations were developed before programmable computers. In rooms where differential equation systems are numerically determined for military calculations, people (usually female) workings on mechanical computers have been commonly found. In the development of analog computers for analysis of mechanical, thermal or chemical systems before programmable computers, the analogies to electrical systems have also been popular. The increased speed of programmable computers and the lower cost means that even more complex differential equation systems can be solved by simply writing programs on a

common PC. Your desktop computer can currently solve problems that the fastest supercomputers were unavailable for only 5 to 10 years before.

Although it is a very young field, the multilinear numerical algebra recently received great attention and dramatic developments because of a strong motivation for practical background and applications. In this field various experts and experts in linear numerical algebra and engineering put their energy. Several international workshops and conferences on this new branch have been held in the USA, France and Switzerland. For example, the Tensor Decomposition Workshop was held at the United States Institute in Palo Alto, California between 19 and 23 July 2004 by Golub, Kolda, Nagy and Van Loan. The workshop was attended by about 35 participants, from 11 countries - informaticians, mathematicians and a broad range of researchers using tensor decline. See the SIAM News article and the website of this workshop. De Lathauwer and Comon also organized the CIRM workshop on tensor decompositions and applications in Lumini and Marseille France, from 29 August to 2 2005. Around 43 people from 13 countries attended the workshop. In the workshop the main problems will be discussed: topological properties of tensor area, tensor decompositions accurately or roughly, mathematical characteristics of tensor decomposition, analysis of separate components, the application of telecommunications, pattern analysis, and statistical modeling, analysis of data and sensor arrays. The Modern Massive Dataset Workshop in Golub, Mahony, Drineas, Lim, hosted a Modern Massive Dataset Workshopping Set at Stanford University June 21-24, 2006. The workshop was attended. The last day of the workshop was the subject of tensor-based data applications. Golub stated that at the conclusion of the workshop a new branch of applied mathematics was developed. In Zurich, Switzerland, from 16-20 July 2007, Golub, Comon, De Lathauwer and Lim held the ICIAM Mini Symposium on the 'Numerical multilinear algebra: a New Beginning.' Golub has written the following: "For the 'numeric multilinear algebra' no common name yet exists. This we usually define as the testing and use of tensor / multilinear algebra, tensor symmetry / symmetric algebra, alternating tensor / external algebra, spinning / Clifford algebra in computational mathematics. There's been an error. This minisymposium is a major step towards the definition and development of this new computer mathematics discipline." It is our hope.

1. FIRST ORDER SYSTEMS

A simple first order differential equation has general form

$$\frac{dy}{dt} = f(y, t) \tag{1.1}$$

When dy/dt means the shift in y as far as time and f(y, t) is concerned, y and time are both functions. Note that the y variable derivatives depend on themselves. For d/dt there are several different scores, popular are y and y'.

One of the simplest differential equations is

$$\frac{dy}{dt} = -y. \tag{1.2}$$

We will focus on this equation in order to introduce the many definitions. The equation is convenient since the simple analysis solution allows us to verify the accuracy of our numerical method. It also controls the efficiency of a heating and a cooling procedure, radioactive decay of chemicals, drug absorption in the body, the charge of a condenser, and population growth only to name a few. This equation is also important.

To analytically solve the equation, we first rearrange the equation as

$$\frac{dy}{y} = -dt \qquad (1.3)$$

and integrate once with respect to time to obtain

$$ln(y) = -t + C \qquad (1.4)$$

Where C is an integration constant. By taking the exponential of the entire equation we remove the normal log term

$$e^{ln(y)} = e^{C-t} \qquad (1.5)$$

which finally can be written as

$$y = Ce^{-t}. \qquad (1.6)$$

You can verify that this answer complies with the equation by replacing the solution in the original equation. Since we have obtained the solution via integration, an integration constant would still be established. This constant is defined according to the initial conditions or the initial state of the system (C in our above solution). We will continue with the initial condition for simplicity of this analysis that $y(t = 0) = 1$, yielding C=1.

2.   NOTATIONS

The N-way tensor, i.e. its entries, are accessed via the N index, say i1,i2, is an N-way array. ,iN range from 1 to Ik for every ik. For instance, a vector has an order 1 tensor and a matrix is an order 2 tensor. In this paper, variables are given their values in the real world, unless stated otherwise. However, in this complex field all statements remain valid. Bold lowercase letters (e.g. u) are denoted to vectors, while matrices are represented with uppercase letters (e.g. A). Calligraphic, uppercase letters denote higher-order tensors (e.g. A). The array entries are scalar in size and labelled with single letters, such as ui or Ti,j,...,l, i,j,., l., l., T. After a change of coordinate system, an N-order tensor has the multilinearity property. To classify the ideas, consider T with the entruments Ti1i2i3, the inversible matrices A, B and C of three square squares with the elements Ai1j1, Bi2j2 and Ci3j3, respectively, and change of coordinates. To describe the ideas, consider T. The tensor T can be written as

$$\mathcal{T}_{i_1 i_2 i_3} = \sum_{j_1=1}^{J_1} \sum_{j_2=1}^{J_2} \sum_{j_3=1}^{J_3} \mathcal{S}_{j_1 j_2 j_3} A_{i_1 j_1} B_{i_2 j_2} C_{i_3 j_3}, \qquad (2.1)$$

and is denoted by

$$\mathcal{T} = A \overset{\mathcal{S}}{\star} B \overset{\mathcal{S}}{\star} C, \qquad (2.2)$$

where $\mathcal{S}$ is a $(J_1 \times J_2 \times J_3)$-tensor, A, B and C are matrices of order $I_1 \times J_1$, $I_2 \times J_2$ and $I_3 \times J_3$, respectively. The Tucker model (or the Tucker product), is commonly used in factor analyses, multi-way data analysis and psychometrics. (2.1)-(2.2) is a tensor representation. Two vectors' external product is defined as

$$C = \mathbf{u} \circ \mathbf{v} = \mathbf{u}\mathbf{v}^{\mathrm{T}} \tag{2.3}$$

with element $C_{ij} = u_i v_j$, which is a matrix. The outer product of N vectors, $\mathbf{u}^{(i)}$ $(i = 1, \ldots, N)$, is defined as

$$\mathscr{T} = \mathbf{u}^{(1)} \circ \mathbf{u}^{(2)} \circ \cdots \circ \mathbf{u}^{(N)} \tag{2.4}$$

with $(i_1, i_2, \ldots, i_N)$-element

$$\mathscr{T}_{i_1 i_2 \cdots i_N} = u_{i_1}^{(1)} u_{i_2}^{(2)} \cdots u_{i_N}^{(N)},$$

which is said to be a rank-1 tensor of order N. In general, for two tensors $\mathscr{A}$ and $\mathscr{B}$ of order M and N respectively, one can define outer product $\mathscr{C} = \mathscr{A} \circ \mathscr{B}$ as a tensor of order $M + N$:

$$\mathscr{C}_{i_1 \cdots i_M j_1 \cdots j_N} = \mathscr{A}_{i_1 \cdots i_M} \mathscr{B}_{j_1 \cdots j_N}. \tag{2.5}$$

Given two tensors, $\mathscr{A} = \{\mathscr{A}_{i_1 \cdots i_M}\}$ and $\mathscr{B} = \{\mathscr{B}_{j_1 \cdots j_N}\}$ having the same first dimension $(I_1 = J_1)$, one can define the mode-1 contraction product (or inner product)

$$(\mathscr{A} \bullet \mathscr{B})_{i_2 \cdots i_M j_2 \cdots j_N} = \sum_{i_1=1}^{I_1} \mathscr{A}_{i_1 \cdots i_M} \mathscr{B}_{j_1 \cdots j_N}. \tag{2.6}$$

This product is caused by multiplication of the regular matrix. Indeed, if two matrices A and B are similar, then the regular matrix product if column number A and row number B are identic (= p), is

$$(AB)_{ij} = \sum_{k=1}^{p} A_{ik} B_{kj}, \tag{2.7}$$

which can be written in mode-1 contraction product

$$AB = A^{\mathrm{T}} \bullet B \tag{2.8}$$

with element

$$(AB)_{ij} = (A^{\mathrm{T}} \bullet B)_{ij} = \sum_{k=1}^{p} (A^{\mathrm{T}})_{ki} B_{kj}. \tag{2.9}$$

Similarly, as long as the tensors A and B have the same p-th dimension, we can define a mode-p contraction product. The Tucker product (2.1) is therefore often known to be a contraction product and often is referred to as by

$$\mathcal{T} = \mathcal{S} \bullet A \bullet B \bullet C \qquad (2.10)$$

or

$$\mathcal{T} = \mathcal{S} \times_1 A \times_2 B \times_3 C, \qquad (2.11)$$

where $\times k$ denotes summing on i_k. This representation also induces several higher-order rank-factorization and higher-order singular value decomposition.

Given two tensors $\mathcal{A}$ and $\mathcal{B}$ of order N with the same dimensions. One defines their Hadamard product $\mathcal{C} = \mathcal{A} \odot \mathcal{B}$ with element as

$$\mathcal{C}_{i_1 \cdots i_N} = \mathcal{A}_{i_1 \cdots i_N} \mathcal{B}_{i_1 \cdots i_N}.$$

It says that the Hadamard product of two tensors gives a tensor with the same order and same dimension with A and B. As usual, the Kronecker product of two vectors u and v of $m \times 1$ and $n \times 1$, respectively, is defined as the vector of $mn \times 1$ with all the possible cross-product as follows:

$$\mathbf{a} \otimes \mathbf{b} := \begin{pmatrix} a_1 \mathbf{b} \\ a_2 \mathbf{b} \\ \vdots \\ a_m \mathbf{b} \end{pmatrix} = \begin{pmatrix} a_1 b_1 \\ \vdots \\ a_1 b_n \\ \vdots \\ a_m b_n \end{pmatrix}.$$

We can give some definitions of super symmetric tensor, tensor scalar product and tensor Frobenius norm, which are the generalization of matrix case in a straightforward way.

3. NUMERICAL ANALYSIS PRACTICE PROBLEMS

The following issues are descriptive of the classroom approaches. They are representative of the problems that are going to be checked.

1. Solving Equations

Problem 1. Suppose that $f : R \rightarrow R$ is continuous and suppose that for $a < b \in R$, $f(a) \cdot f(b) < 0$. Show that there is a c with $a < c < b$ such that $f(c) = 0$.

Problem 2. Solve the equation $x^5 - 3x^4 + 2x^3 - x^2 + x = 3$. Solve using the Bisection method. Solve using the Newton-Raphson method. How many solutions are there?

Problem 3. The Bisection method and Newton Raphson method solve the equation $x = \cos x$. What are the numbers of solutions? The $\sin(x) = \cos x$ equation can be solved by means of bisection and Newton-Raphson. How many solutions are there?

Problem 4. Let h be a continuous function $h : Rn \rightarrow Rn$. Let $x0 \in Rn$. Suppose that $h^n(x0) \rightarrow z$ as $n \rightarrow \infty$. Show that $h(z) = z$.

Problem 5. Solve the equation $x\,4 = 2$ by the Newton-Raphson method. How many real solutions are there? For which starting values x0 will the method converge?

Problem 6. Suppose that $f : \mathbb{R} \rightarrow \mathbb{R}$ is continuous and that $f(z) = 0$. Suppose that $f'(z) \neq 0$. Let $g(x) = x - \frac{f(x)}{f'(x)}$. Show that there is an $\varepsilon > 0$ such that for any $z_0 \in [z - \varepsilon, z + \varepsilon]$, $g^n(z_0) \rightarrow z$ as $n \rightarrow \infty$

Problem 7. Show that the Newton-Raphson method converges quadratically. That is, suppose that the fixed point is z and that the error of the nth iteration is $|x_n - z| = h$, then $|x_{n+1} - z| \approx h^2$ for h small enough.

Solution. Let let $f(x)$ be a function on $[0, \pi]$. Then the estimate will be $\int_0^\pi p(x)dx$ where $p(x)$ is the Lagrange polynomial which is $f\left(\frac{\pi}{2}\right)$ at $\frac{\pi}{2}$ and $f\left(\frac{3\pi}{4}\right)$ at $\frac{3\pi}{4}$. Now $p(x) = f\left(\frac{\pi}{2}\right) \cdot p_0(x) + f\left(\frac{3\pi}{4}\right) \cdot p_1(x)$ where $p_0(x) = \frac{\left(x - \frac{3\pi}{4}\right)}{\left(\frac{\pi}{2} - \frac{3\pi}{4}\right)}$ and $p_1(x) = \frac{\left(x - \frac{\pi}{2}\right)}{\left(\frac{3\pi}{4} - \frac{\pi}{2}\right)}$. Now $\int_0^\pi p(x)dx = \int_0^\pi \left(f\left(\frac{\pi}{2}\right) \cdot p_0(x) + f\left(\frac{3\pi}{4}\right) \cdot p_1(x)\right) dx$. This shows that $A_0 = \int_0^\pi p_0(x)dx$ and $A_1 = \int_0^\pi p_1(x)dx$. Thus, $A_0 = \pi$ and $A_1 = 0$.

Differential Equations

Problem 1. Solve the differential equation for $\frac{dx}{dt} = f(t, x) = t \cdot x^2$ with $x(0) = 1$. Solve using Picard iteration for five iterations. Solve using the Taylor method of order 3,4, and 5. Solve using the Euler method, modified Euler, Heun, and Runge-Kutta methods using $h = \frac{1}{20}$ and $n = 20$. Compare the answers and the errors for each of these methods.

Problem 2. How would you go about solving the differential equation $\frac{d^2x}{dt^2} = x$ with $x(0) = 1$ and $x'(0) = 1$ with each of the methods listed in the previous problem. What changes would need to be made in the programs? Solve this problem as a linear differential equation using the linear ode program. Solve on the interval [0, 1] with $h = \frac{1}{10}$.

Problem 3. Consider the following differential equation.

$$\frac{dx}{dt} = t \cdot x$$
$$x(0) = 1$$

Solve on the interval [0, 1] using h = .1.

Problem 4. Compare Euler, Heun, and Runge-Kutta on [0, 1] using h = .1.

$$\frac{dx}{dt} = t \cdot x$$
$$x(0) = 1$$

Problem 5. Use the Euler method to solve the following differential equation

$$\frac{dx}{dt} = x$$
$$x(0) = 1$$

Solve on [0, 1] using $h = \frac{1}{n}$. Do this by hand to show that $x_n = \left(1 + \frac{1}{n}\right)^n$. What does this say about the following limit

$$\lim_{n \to \infty} \left(1 + \frac{1}{n}\right)^n$$

CONCLUSION

In this study, we discuss the motivation with the context and development of the key areas of the multi-linear numerical algebra. We have examined decompositions of tensors (for example,

decomposition in high-order, decomposition of high-order singular value, canonical and pseudo-canonical decomposition, decomposition of tensors, etc.), best approximation for rank 1 and rank rank r, related high-order value-adding problems, polynomial optimization in multiple variations, and typical applications of tensors; (for example, BSS, BD, SDP, and other optimization problems). The growth of this field is currently greatly stimulated by various applications of digital image restore, signal processing, wireless communications, psychometry, and multi-way data analysis and high-order statistics. So far, the first move is only beginning with several topics in the field. There are also minimal numerical analysis and numerical performance.

REFERENCES

1.   Alon N, de la Vega W F, Kannan R, et. al. (2003). Random sampling and approximation of max-csps. J Comput System Sci., 67: pp. 212–243

2.   Bergman G. M. (1969). Ranks of tensors and change of base field. J Algebra, 11: pp. 613–621

3.   Bienvenu G, Kopp L. (1983). Optimality of high-resolution array processing using the eigensystem approach. IEEE Trans ASSP, 31: pp. 1235–1248

4.   Cao X R, Liu R W (1996). General approach to blind source separation. IEEE Trans Signal Processing, 44: pp. 562–570

5.   Cardoso J F (1991). Super-symmetric decomposition of the forth-order cumulant tensor. Blind identification of more sources than sensors. In: Proceedings of the IEEE International Conference on Acoust, Speech, and Signal Processing (ICASSP'91), Toronto, Canada.

6.   Cardoso J F (1999). High-order contrasts for independent component analysis. Neural Computation, 11: pp. 157–192

7.   Caroll J D, Chang J J (1970). Analysis of individual differences in multidimensional scaling via n-way generalization of Eckart-Young decomposition. Psychometrika, 35: pp. 283–319

8.   Comon P (1994). Tensor diagonalization, a useful tool in signal processing. In: Blanke M, Soderstrom T, eds. IFAC-SYSID, 10th IFAC Symposium on System Identification (Copenhagen, Denmark, July 4-6, 1994. Invited Session). Vol. 1, pp. 77–82

9.   Comon P (1994). Independent component analysis, a new concept? Signal Processing, Special Issue on Higher-Order Statistics, 36: pp. 287–314

10.   Comon P, Mourrain B. (1996). Decomposition of quantics in sums of powers of linear forms. Signal Processing, Special Issue on Higher-Order Statistics, 53: pp. 96–107

11.   Comon P. (2000). Block methods for channel identification and source separation. In: IEEE Symposium on Adaptive Systems for Signal Process, Commun Control (Lake Louise, Alberta, Canada, Oct 1-4, 2000. Invited Plenary). Pp. 87–92

12.   Comon P, Chevalier P (2000). Blind source separation: Models, concepts, algorithms, and performance. In: Haykin S, ed. Unsupervised Adaptive Filtering, Vol 1. New York: John Wiley.

13.   Comon P (2002). Tensor decompositions: State of the art and applications. In: McWhirter J G, Proundler I K, eds. Mathematics in Signal Processing V. Oxford: Oxford University Press.

14. Comon P (2004). Canonical tensor decompositions. In: ARCC Workshop on Tensor Decompositions, Americal Institute of Mathematics (AIM), Palo Alto, California, USA, July 19–23.

15. Comon P, Golub G, Lim L H, et. al. (2007). Symmetric tensors and symmetric tensor rank. SIAM J Matrix and Applications, 2007 (in press)

16. Coppi R, Bolasco S, eds. (1989). Multiday Data Analysis. Amsterdam: Elsevier.

17. de la Vega W F, Kannan R, Karpinski M, et. al. (2005). Tensor Decomposition and Approximation Algorithms for Constraint Satisfaction Problems. New York: ACM Press, pp. 747–754

18. Defant A, Floret K (1993). Tensor Norms and Operator Ideals. North-Holland Mathematics Studies, No 176. Amsterdam: North-Holland.

19. Ferrier C (1998). Hilbert's 17th problem and best dual bounds in quadatic minimization. Cybernet Systems Anal, 34: pp. 696–709

20. Golub G, Van Loan C F (1996). Matrix Computations. 3rd ed. Baltimore: Johns Hopkins University Press.

21. Golub G, Kolda T G, Nagy J, et. al. (2004). Workshop on Tensor Decompositions, American Institute of Mathematics, Palo Alto, California, 19–23 July, 2004. http://www.aimath.org/WWN/tensordecomp/

22. Golub G, Mahoney M, Drineas P, et. al. (2006). Workshop for Modern Massive Data Sets, 21–24 June, 2006. http://www.stanford.edu/group/mmds/

23. Golub G, Mahoney M, Drineas P, et. al. (2006). Bridge the gap between numerical linear algebra, theoretical computer science, and data applications. SIAM News, 39(8). http://www.siam.org/pdf/news/1019.pdf

24. Greub W H (1967). Multilinear Algebra. Berlin: Springer-Verlag.

25. Harshman R A (1972). Determination and proof of minimum uniqueness conditions of PARAFAC. UCLA Working Papers in Phonetics, 22: pp. 111–117.

26. He X, Sun W (1990). Introduction to Generalized Inverses of Matrices. Nanjing: Jiangsu Sci & Tech Publishing House, 1990 (in Chinese)

Corresponding Author

Alok Tripathi*

Associate Professor, Department of Mathematics, Galgotias University, Greater Noida, Uttar Pradesh, India