Solving Multi-Linear Systems with $\mathcal{M}$-Tensors

Yimin WEI

School of Mathematical Sciences
Fudan University, Shanghai, China

Joint work with Weiyang DING

Definition ($\mathcal{Z}$-tensor)

We call a tensor $\mathcal{A}$ as a *$\mathcal{Z}$-tensor*, if all of its off-diagonal entries are non-positive.

Spectral radius of a tensor is defined by

$$\rho(\mathcal{A}) = \left\{ |\lambda| : \mathcal{A}\mathbf{x}^{m-1} = \lambda \mathbf{x}^{[m-1]}, \lambda \in \mathbb{C}, \mathbf{x} \in \mathbb{C}^n \setminus \{\mathbf{0}\} \right\},$$

where

$$(\mathcal{A}\mathbf{x}^{m-1})_i = \sum_{i_2,\ldots,i_m=1}^{n} a_{ii_2\ldots i_m} x_{i_2} \ldots x_{i_m}$$

and

$$\mathbf{x}^{[m-1]} = [x_1^{m-1}, x_2^{m-1}, \ldots, x_n^{m-1}]^\top.$$

$\mathcal{M}$-Tensor II

Definition ($\mathcal{M}$-tensor)

We call a $\mathcal{Z}$-tensor $\mathcal{A} = s\mathcal{I} - \mathcal{B}$ ($\mathcal{B} \geq 0$) as an *$\mathcal{M}$-tensor* if $s \geq \rho(\mathcal{B})$; We call it as a *nonsingular $\mathcal{M}$-tensor* if $s > \rho(\mathcal{B})$.

K. Chang, L. Qi, and T. Zhang, A survey on the spectral theory of nonnegative tensors. Numer. Linear Algebra Appl., 20(6), 891–912, 2013.

W. Ding, L. Qi, and Y. Wei. $\mathcal{M}$-tensors and nonsingular $\mathcal{M}$-tensors. Linear Algebra Appl., 439(10):3264–3278, 2013.

L. Zhang, L. Qi, and G. Zhou. $\mathcal{M}$-tensors and some applications. SIAM J. Matrix Anal. Appl., 35(2):437–452, 2014.

If $\mathcal{A}$ is a $\mathcal{Z}$-tensor, then the following conditions are equivalent:

(D1) $\mathcal{A}$ is a nonsingular $\mathcal{M}$-tensor;

(D2) Every real eigenvalue of $\mathcal{A}$ is positive;

(D3) The real part of each eigenvalue of $\mathcal{A}$ is positive;

(D4) $\mathcal{A}$ is semi-positive; that is, there exists $\mathbf{x} > \mathbf{0}$ with $\mathcal{A}\mathbf{x}^{m-1} > \mathbf{0}$;

(D5) There exists $\mathbf{x} \geq \mathbf{0}$ with $\mathcal{A}\mathbf{x}^{m-1} > \mathbf{0}$;

(D6) $\mathcal{A}$ has all positive diagonal entries, and there exists a positive diagonal matrix D such that $\mathcal{A}D^{m-1}$ is strictly diagonally dominant;

(D7) $\mathcal{A}$ has all positive diagonal entries, and there exist two positive diagonal matrices D_1 and D_2 such that $D_1\mathcal{A}D_2^{m-1}$ is strictly diagonally dominant;

(D8) There exists a positive diagonal tensor $\mathcal{D}$ and a nonsingular $\mathcal{M}$-tensor $\mathcal{C}$ with $\mathcal{A} = \mathcal{D}\mathcal{C}$;

(D9) There exists a positive diagonal tensor $\mathcal{D}$ and a nonnegative tensor $\mathcal{E}$ such that $\mathcal{A} = \mathcal{D} - \mathcal{E}$ and there exists $\mathbf{x} > \mathbf{0}$ with

Multi-Linear Equations I

Let $\mathcal{A}$ be an m^{th}-order tensor in $\mathbb{C}^{n\times n\times\cdots\times n}$ and $\mathbf{b}$ be a vector in $\mathbb{C}^{n}$.

Then a *multi-linear equation* can be expressed as

$$\mathcal{A}\mathbf{x}^{m-1} = \mathbf{b}.$$

E.g., if $\mathcal{A} \in \mathbb{R}^{2\times 2\times 2}$, then $\mathcal{A}\mathbf{x}^2 = \mathbf{b}$ is a condense form of

$$\begin{cases} a_{111}x_1^2 + (a_{112} + a_{121})x_1x_2 + a_{122}x_2^2 = b_1, \\ a_{211}x_1^2 + (a_{212} + a_{221})x_1x_2 + a_{222}x_2^2 = b_2. \end{cases}$$

Denote the *solution set* of the multi-linear equation $\mathcal{A}\mathbf{x}^{m-1} = \mathbf{b}$ as

$$\mathcal{A}^{-1}\mathbf{b} := \left\{\mathbf{x} \in \mathbb{C}^n : \mathcal{A}\mathbf{x}^{m-1} = \mathbf{b}\right\}.$$

Furthermore, when $\mathcal{A}$ and $\mathbf{b}$ are both real, we define the real solution set, the nonnegative solution set, and the positive solution set as

$$\begin{aligned}(\mathcal{A}^{-1}\mathbf{b})_{\mathbb{R}} &:= \left\{\mathbf{x} \in \mathbb{R}^n : \mathcal{A}\mathbf{x}^{m-1} = \mathbf{b}\right\},\\ (\mathcal{A}^{-1}\mathbf{b})_{+} &:= \left\{\mathbf{x} \in \mathbb{R}^n_{+} : \mathcal{A}\mathbf{x}^{m-1} = \mathbf{b}\right\},\\ (\mathcal{A}^{-1}\mathbf{b})_{++} &:= \left\{\mathbf{x} \in \mathbb{R}^n_{++} : \mathcal{A}\mathbf{x}^{m-1} = \mathbf{b}\right\}.\end{aligned}$$

An $\mathcal{M}$-equation is referred to

$$\mathcal{A}\mathbf{x}^{m-1} = \mathbf{b}, \tag{1}$$

where $\mathcal{A} = s\mathcal{I} - \mathcal{B}$ is an $\mathcal{M}$-tensor. Particularly, when the right-hand side $\mathbf{b}$ is nonnegative or positive, we also require a nonnegative or positive solution, respectively.

- Find one of the sparsest solutions to a tensor complementarity problem

$$\min \|\mathbf{x}\|_0, \quad \text{s.t. } \mathcal{A}\mathbf{x}^{m-1} - \mathbf{b} \geq \mathbf{0}, \ \mathbf{x} \geq \mathbf{0}, \ \mathbf{x}^\top(\mathcal{A}\mathbf{x}^{m-1} - \mathbf{b}) = 0; \tag{2}$$

This is generally an NP-hard problem. Luo, Qi and Xiu (Optimization Letters, 2016, online) suggest that if the tensor $\mathcal{A}$ is a $\mathcal{Z}$-tensor, then a sparsest solution of the above tensor complementarity problem can be achieved by solving the following polynomial programming problem

$$\min \|\mathbf{x}\|_1, \quad \text{s.t. } \mathcal{A}\mathbf{x}^{m-1} = \mathbf{b}, \ \mathbf{x} \geq \mathbf{0}. \tag{3}$$

They prove that if $\mathcal{A}$ is a nonsingular $\mathcal{M}$-tensor, then the problem (3) is uniquely solvable and the unique solution is also an optimal solution to the problem (2).

Applications II

- Discretize the partial differential equation (nonlinear Klein-Gordon equation) with Dirichlet's boundary condition like

$$\begin{cases} u(\mathbf{x})^{m-2} \cdot \Delta u(\mathbf{x}) = -f(\mathbf{x}) \text{ in } \Omega, \\ u(\mathbf{x}) = g(\mathbf{x}) \text{ on } \partial\Omega, \end{cases} \qquad (m = 3, 4, \dots)$$

- Analyze a necessary and sufficient condition for existence of a positive Perron vector. (S. Hu and L. Qi, arXiv:1511.07759, SIMAX, 2016)

Diagonal Systems I

The diagonal of a tensor $\mathcal{A}$ contains the entries $a_{ii\ldots i}$ with $i = 1, 2, \ldots, n$, and other entries are called off-diagonal.

A tensor is called diagonal if all its off-diagonal entries are zeros.

A diagonal equation is referred to

$$\mathcal{D}\mathbf{x}^{m-1} = \mathbf{b},$$

where the coefficient m^{th}-order tensor $\mathcal{D}$ is diagonal.

Diagonal Systems II

When m is even, the real solution set $(\mathcal{D}^{-1}\mathbf{b})_{\mathbb{R}}$ has a unique element $\mathbf{x}$ with $x_i = (b_i/d_{ii\ldots i})^{1/(m-1)}$.

When the diagonal of $\mathcal{D}$ and the vector $\mathbf{b}$ are positive, the solution above is also the unique element in $(\mathcal{D}^{-1}\mathbf{b})_{++}$.

When m is odd, the real solution set $(\mathcal{D}^{-1}\mathbf{b})_{\mathbb{R}}$ has at most 2^n elements $\mathbf{x}$ with $x_i = \pm(b_i/d_{ii\ldots i})^{1/(m-1)}$ if $d_{ii\ldots i}b_i \geq 0$ for all i, or else the real solution set is empty.

Further when the diagonal of $\mathcal{D}$ and the vector $\mathbf{b}$ are positive, the positive solution set $(\mathcal{D}^{-1}\mathbf{b})_{++}$ has a unique element $\mathbf{x}$ with $x_i = (b_i/d_{ii\ldots i})^{1/(m-1)}$.

Triangular Systems

The lower triangular part of a tensor $\mathcal{A}$ contains the entries $a_{i_1 i_2 \ldots i_m}$ with $i_1 = 1, 2, \ldots, n$ and $i_2, \ldots, i_m \leq i_1$, and other entries are said to be in the off-lower triangular part. The strictly lower part consists of the entries $a_{i_1 i_2 \ldots i_m}$ with $i_1 = 1, 2, \ldots, n$ and $i_2, \ldots, i_m < i_1$.

A tensor is called lower triangular if all its entries in the off-lower triangular part are zeros.

A lower triangular equation is referred to

$$\mathcal{L}\mathbf{x}^{m-1} = \mathbf{b},$$

where the coefficient tensor $\mathcal{L}$ is lower triangular.

The (strictly) upper triangular part of a tensor, upper triangular tensors and upper triangular equations can be defined analogously.

Forward Substitution

Algorithm (Forward Substitution)

If $\mathcal{L} \in \mathbb{C}^{n\times n\times\cdots\times n}$ is lower triangular and $\mathbf{b} \in \mathbb{C}^n$, then this algorithm overwrites $\mathbf{b}$ with one of the solutions to $\mathcal{L}\mathbf{x}^{m-1} = \mathbf{b}$.

$b_1 =$ one of the $(m-1)$-st roots of $b_1/l_{11\dots1}$
for $i = 2 : n$
  for $k = 1 : m$

$$p_k = \sum \Big\{ l_{ii_2\dots i_m} \cdot \prod_{\substack{l=2,\dots,m \\ l\neq p_1,\dots,p_{k-1}}} b_{i_l} \;:\; i_2,\dots,i_m \le i, \; i_{p_1},\dots,i_{p_{k-1}} \text{ are the only } k-1 \text{ indices equal } i \Big\}$$

  end
  $b_i =$ one of the roots of $p_1 + p_2 t + \cdots + p_m t^{m-1} = b_i$
end

Triangular $\mathcal{M}$-Equations

Now we consider the triangular equations satisfying that the coefficient tensor is a nonsingular $\mathcal{M}$-tensor.

For a triangular tensor, it is a nonsingular $\mathcal{M}$-tensor if and only if its diagonal entries are positive and its off-diagonal entries are nonpositive.

Proposition

Let $\mathcal{L}$ be an m^{th}-order n-dimensional lower triangular $\mathcal{M}$-tensor. If $\mathbf{b}$ is a nonnegative vector, then $\mathcal{L}\mathbf{x}^{m-1} = \mathbf{b}$ has at least one nonnegative solution. Furthermore, if $\mathbf{b}$ is a positive vector, then $\mathcal{L}\mathbf{x}^{m-1} = \mathbf{b}$ has a unique positive solution.

Triangular $\mathcal{M}$-Equations

When we solve the equation, the coefficients $p_1, p_2, \ldots, p_{m-1}$ are nonpositive and p_m is positive in each step. Thus the companion matrix of $p_1 + p_2 t + \cdots + p_m t^{m-1} = b_i$, i.e.,

$$C_i = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ \frac{b_i - p_1}{p_m} & \frac{-p_2}{p_m} & \frac{-p_3}{p_m} & \cdots & \frac{-p_{m-1}}{p_m} \end{bmatrix},$$

is a nonnegative matrix when $b_i \geq 0$, and it is an irreducible nonnegative matrix, which is not similar via a permutation to a block upper triangular matrix, when $b_i > 0$.

Thus the polynomial equation in each step has at least one nonnegative solution $x_i = \rho(C_i)$ if $\mathbf{b}$ is nonnegative ($\rho(C_i)$ is the spectral radius of matrix C_i), and it has a unique positive solution if $\mathbf{b}$ is positive.

A Fixed-Point Iteration

An $\mathcal{M}$-equation is referred to

$$\mathcal{A}\mathbf{x}^{m-1} = \mathbf{b},$$

where $\mathcal{A} = s\mathcal{I} - \mathcal{B}$ is an $\mathcal{M}$-tensor. When the right-hand side $\mathbf{b}$ is nonnegative, we require a nonnegative solution.

Consider the fixed-point iteration

$$\mathbf{x}_{k+1} = T_{s,\mathcal{B},\mathbf{b}}(\mathbf{x}_k) := (s^{-1}\mathcal{B}\mathbf{x}_k^{m-1} + s^{-1}\mathbf{b})^{[1/(m-1)]}, \ k = 0, 1, 2, \dots,$$

where $\mathbf{v}^{[\alpha]} = [v_1^\alpha, v_2^\alpha, \dots, v_n^\alpha]^\top$ denotes the componentwise power of a vector.

It is easy to understand that each fixed point of this iteration is a solution of the above nonsingular $\mathcal{M}$-equation, and vice versa.

Cones

Let $\mathbb{E}$ be a real Banach space. If $\mathbb{P}$ is a nonempty closed and convex set in $\mathbb{E}$ and satisfies that

- $\mathbf{x} \in \mathbb{P}$ and $\lambda \geq 0$ imply $\lambda\mathbf{x} \in \mathbb{P}$, and
- $\mathbf{x} \in \mathbb{P}$ and $-\mathbf{x} \in \mathbb{P}$ imply $\mathbf{x} = \mathbf{0}$,

then $\mathbb{P}$ is called a cone in $\mathbb{E}$.

Furthermore, a cone $\mathbb{P}$ induces a semi-order in $\mathbb{E}$, i.e., $\mathbf{x} \leq \mathbf{y}$ if $\mathbf{y} - \mathbf{x} \in \mathbb{P}$. Let $\{\mathbf{x}_n\}$ be an arbitrary increasing series in $\mathbb{E}$ with an upper bound, i.e., there exists $\mathbf{y} \in \mathbb{E}$ such that

$$\mathbf{x}_1 \leq \mathbf{x}_2 \leq \cdots \leq \mathbf{x}_n \leq \cdots \leq \mathbf{y}.$$

If there must be $\mathbf{x}_* \in \mathbb{E}$ such that $\|\mathbf{x}_n - \mathbf{x}_*\| \to 0$ $(n \to \infty)$, then we call $\mathbb{P}$ a regular cone.

Let $T : \mathbb{D} \to \mathbb{E}$ be a map, where $\mathbb{D}$ is a subset in $\mathbb{E}$. If $\mathbf{x} \leq \mathbf{y}$ $(\mathbf{x}, \mathbf{y} \in \mathbb{D})$ implies $T(\mathbf{x}) \leq T(\mathbf{y})$, then we call T an increasing map on $\mathbb{D}$.

A Fixed-Point Theorem

Theorem (Amann 1976)

Let $\mathbb{P}$ be a regular cone in an ordered Banach space $\mathbb{E}$ and $[\mathbf{u}, \mathbf{v}] \subset \mathbb{E}$ be a bounded order interval. Suppose that $T : [\mathbf{u}, \mathbf{v}] \to \mathbb{E}$ is an increasing continuous map which satisfies

$$\mathbf{u} \leq T(\mathbf{u}) \quad \text{and} \quad \mathbf{v} \geq T(\mathbf{v}).$$

Then T has at least one fixed point in $[\mathbf{u}, \mathbf{v}]$. Moreover, there exists a minimal fixed point $\mathbf{x}_*$ and a maximal fixed point $\mathbf{x}^*$ in the sense that every fixed point $\bar{\mathbf{x}}$ satisfies $\mathbf{x}_* \leq \bar{\mathbf{x}} \leq \mathbf{x}^*$. Finally, the iteration method

$$\mathbf{x}_{k+1} = T(\mathbf{x}_k), \quad k = 0, 1, 2, \ldots$$

converges to $\mathbf{x}_*$ from below if $\mathbf{x}_0 = \mathbf{u}$, i.e., $\mathbf{u} = \mathbf{x}_0 \leq \mathbf{x}_1 \leq \cdots \leq \mathbf{x}_*$, and converges to $\mathbf{x}^*$ from above if $\mathbf{x}_0 = \mathbf{v}$, i.e., $\mathbf{v} = \mathbf{x}_0 \geq \mathbf{x}_1 \geq \cdots \geq \mathbf{x}^*$.

By the above fixed-point theorem, we can study the existence of the positive solutions of the $\mathcal{M}$-equations.

Theorem

If $\mathcal{A}$ is a nonsingular $\mathcal{M}$-tensor, then for every positive vector $\mathbf{b}$ the multi-linear system of equations $\mathcal{A}\mathbf{x}^{m-1} = \mathbf{b}$ has a unique positive solution.

We can rewrite the above theorem into an equivalent condition for nonsingular $\mathcal{M}$-tensors, which generalizes the 'nonnegative inverse' property of M-matrix to the tensor case.

Existence and Uniqueness of Positive Solutions II

Theorem

Let $\mathcal{A}$ be a $\mathcal{Z}$-tensor. Then it is a nonsingular $\mathcal{M}$-tensor if and only if $(\mathcal{A}^{-1}\mathbf{b})_{++}$ has a unique element for every positive vector $\mathbf{b}$.

We introduce the notation $\mathcal{A}_{++}^{-1}\mathbf{b}$ to denote the unique positive solution to $\mathcal{A}\mathbf{x}^{m-1} = \mathbf{b}$ for a nonsingular $\mathcal{M}$-tensor $\mathcal{A}$ and a positive vector $\mathbf{b}$. Then $\mathcal{A}_{++}^{-1}\widehat{\mathbf{b}} \geq \mathcal{A}_{++}^{-1}\widetilde{\mathbf{b}} > \mathbf{0}$ if $\widehat{\mathbf{b}} \geq \widetilde{\mathbf{b}} > \mathbf{0}$.

Proof of existence

Recall that a nonnegative solution of $\mathcal{A}\mathbf{x}^{m-1} = \mathbf{b}$ is a fixed point of

$$T_{s,\mathcal{B},\mathbf{b}} : \mathbb{R}^n_+ \to \mathbb{R}^n_+, \ \mathbf{x} \mapsto \left(s^{-1}\mathcal{B}\mathbf{x}^{m-1} + s^{-1}\mathbf{b}\right)^{[1/(m-1)]}.$$

Note that $\mathbb{R}^n_+$ is a regular cone and $T_{s,\mathcal{B},\mathbf{b}}$ is an increasing continuous map. When $\mathcal{A}$ is a nonsingular $\mathcal{M}$-tensor, i.e., $s > \rho(\mathcal{B})$, there exists a positive vector $\mathbf{z} \in \mathbb{R}^n_{++}$ such that $\mathcal{A}\mathbf{z}^{m-1} > \mathbf{0}$. Denote

$$\underline{\gamma} = \min_{i=1,2,\ldots,n} \frac{b_i}{(\mathcal{A}\mathbf{z}^{m-1})_i} \quad \text{and} \quad \overline{\gamma} = \max_{i=1,2,\ldots,n} \frac{b_i}{(\mathcal{A}\mathbf{z}^{m-1})_i}.$$

Then $\underline{\gamma}\mathcal{A}\mathbf{z}^{m-1} \leq \mathbf{b} \leq \overline{\gamma}\mathcal{A}\mathbf{z}^{m-1}$, which indicates that

$$\underline{\gamma}^{1/(m-1)}\mathbf{z} \leq T_{s,\mathcal{B},\mathbf{b}}\left(\underline{\gamma}^{1/(m-1)}\mathbf{z}\right) \quad \text{and} \quad \overline{\gamma}^{1/(m-1)}\mathbf{z} \geq T_{s,\mathcal{B},\mathbf{b}}\left(\overline{\gamma}^{1/(m-1)}\mathbf{z}\right).$$

By the above fixed-point theorem, there exists at least one fixed point $\bar{\mathbf{x}}$ of $T_{s,\mathcal{B},\mathbf{b}}$ with

$$\underline{\gamma}^{1/(m-1)}\mathbf{z} \leq \bar{\mathbf{x}} \leq \overline{\gamma}^{1/(m-1)}\mathbf{z},$$

which is obviously a positive vector when $\mathbf{b}$ is also positive.

Proof of uniqueness

Furthermore, we can prove that the positive fixed point is unique when $\mathbf{b}$ is positive. Assume that there are two positive fixed points $\mathbf{x}$ and $\mathbf{y}$, i.e.,

$$T_{s,\mathcal{B},\mathbf{b}}(\mathbf{x}) = \mathbf{x} > \mathbf{0} \quad \text{and} \quad T_{s,\mathcal{B},\mathbf{b}}(\mathbf{y}) = \mathbf{y} > \mathbf{0}.$$

Denote $\eta = \min_{i=1,2,\ldots,n} \frac{x_i}{y_i}$, thus $\mathbf{x} \geq \eta\mathbf{y}$ and $x_j = \eta y_j$ for some j. If $\eta < 1$, then $\mathcal{A}(\eta\mathbf{y})^{m-1} = \eta^{m-1}\mathbf{b} < \mathbf{b}$, which indicates

$$T_{s,\mathcal{B},\mathbf{b}}(\eta\mathbf{y}) = \left[s^{-1}\mathcal{B}(\eta\mathbf{y})^{m-1} + s^{-1}\mathbf{b}\right]^{[1/(m-1)]} > \eta\mathbf{y}.$$

However, since $T_{s,\mathcal{B},\mathbf{b}}$ is nonnegative and increasing, we have

$$T_{s,\mathcal{B},\mathbf{b}}(\eta\mathbf{y})_j \leq T_{s,\mathcal{B},\mathbf{b}}(\mathbf{x})_j = x_j = \eta y_j.$$

This is a contradiction. Thus $\eta \geq 1$, which implies $\mathbf{x} \geq \mathbf{y}$. Similarly, we can also show that $\mathbf{y} \geq \mathbf{x}$, so $\mathbf{x} = \mathbf{y}$. Therefore, the positive fixed point of $T_{s,\mathcal{B},\mathbf{b}}$ is unique, and equivalently the positive solution to $\mathcal{A}\mathbf{x}^{m-1} = \mathbf{b}$ is unique.

Theorem

Let $\mathcal{A}$ be an $\mathcal{M}$-tensor. If there exists a nonnegative vector $\mathbf{v}$ such that $\mathcal{A}\mathbf{v}^{m-1} \geq \mathbf{b}$, then $(\mathcal{A}^{-1}\mathbf{b})_+$ is nonempty.

Generally speaking, the nonnegative solution set $(\mathcal{A}^{-1}\mathbf{b})_+$ in the above theorem has more than one element, and these nonnegative solutions lay on a hypersurface in $\mathbb{R}^n$.

E.g., we construct a $3 \times 3 \times 3$ singular $\mathcal{M}$-tensor $\mathcal{A} = \mathcal{I} - \mathcal{B}$, where $\mathcal{B}$ is a nonnegative tensor with $\mathcal{B}\mathbf{1}^2 = \mathbf{1}$, so that $\rho(\mathcal{B}) = 1$, and $b_{ijk} = 0$ if $i \in \{2, 3\}$ and either $j = 1$ or $k = 1$. Apparently, for each right-hand side $\mathbf{b} = (b_1, 0, 0)^\top$, the nonnegative solutions to $\mathcal{A}\mathbf{x}^2 = \mathbf{b}$ have the form $\mathbf{x} = (\alpha, \beta, \beta)^\top$. We display the procedures of the iteration $\mathbf{x}_{k+1} = (\mathcal{B}\mathbf{x}_k^2 + \mathbf{b})^{[1/2]}$ with two kinds of initial points $(\beta, \beta, \beta)^\top$ and $(2 - \beta, \beta, \beta)^\top$ $(0 \leq \beta \leq 1)$.

Systems with General $\mathcal{M}$-Tensors

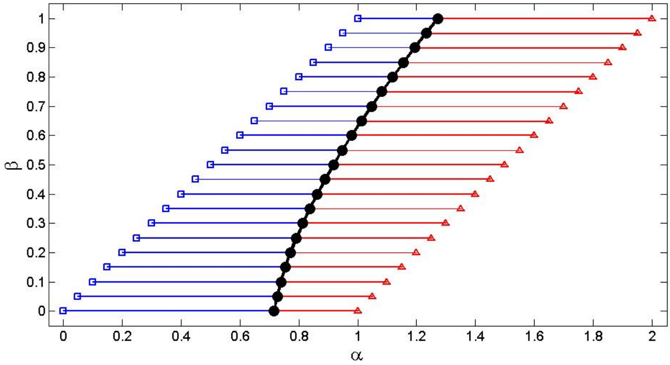


Figure: The iterations for a singular $\mathcal{M}$-equation.

Non-Positive Right-Hand Side I

We discuss the nonsingular $\mathcal{M}$-equations with non-positive right-hand sides.

When the coefficient tensor is of even order, the situation is simple.

Assume that $\mathcal{A}$ is an even-order nonsingular $\mathcal{M}$-tensor and $\mathbf{b}$ is a non-positive vector. Then the equation $\mathcal{A}\mathbf{x}^{m-1} = \mathbf{b}$ is equivalent to $\mathcal{A}(-\mathbf{x})^{m-1} = -\mathbf{b}$, which is a nonsingular $\mathcal{M}$-equation with nonnegative right-hand side. However, the case is totally different when the coefficient tensor is of odd order.

Non-Positive Right-Hand Side II

Theorem

Let $\mathcal{A}$ be a $\mathcal{Z}$-tensor. Then $\mathcal{A}$ is a nonsingular $\mathcal{M}$-tensor if and only if $\mathcal{A}$ does not reverse the sign of any vector; that is, if $\mathbf{x} \neq \mathbf{0}$ and $\mathbf{b} = \mathcal{A}\mathbf{x}^{m-1}$, then for some subscript i,

$$x_i^{m-1} b_i > 0.$$

From the above theorem, we can easily understand that there is no real vector $\mathbf{x}$ such that $\mathbf{b} = \mathcal{A}\mathbf{x}^{m-1}$ is non-positive when m is odd, since $\mathbf{x}^{[m-1]}$ is always nonnegative.

Non-Homogeneous Left-Hand Side I

All the multi-linear equations that we discuss above have homogeneous left-hand sides.

However, similar results can be established for some special equations with non-homogeneous left-hand sides. Consider the following equation

$$\mathcal{A}\mathbf{x}^{m-1} - \mathcal{B}_{m-1}\mathbf{x}^{m-2} - \cdots - \mathcal{B}_2\mathbf{x} = \mathbf{b} > \mathbf{0},$$

where $\mathcal{A} = s\mathcal{I} - \mathcal{B}_m$ is an m^{th}-order nonsingular $\mathcal{M}$-tensor and $\mathcal{B}_p$ is a p^{th}-order nonnegative tensor for $p = 2, 3, \ldots, m$.

Non-Homogeneous Left-Hand Side II

Theorem

Let $\mathcal{A}$ be an m^{th}-order $\mathcal{Z}$-tensor and $\mathcal{B}_p$ be a p^{th}-order nonnegative tensor for $p = 2, 3, \ldots, m$. Then the equation

$$\mathcal{A}\mathbf{x}^{m-1} - \mathcal{B}_{m-1}\mathbf{x}^{m-2} - \cdots - \mathcal{B}_2\mathbf{x} = \mathbf{b} \tag{4}$$

has a unique positive solution for every positive vector $\mathbf{b}$ *if and only if $\mathcal{A}$ is a nonsingular $\mathcal{M}$-tensor.*

Absolutely $\mathcal{M}$-Equations

A tensor is called an absolutely $\mathcal{M}$-tensor if it can be written as $\mathcal{A} = s\mathcal{I} - \mathcal{B}$ with $s > \rho(|\mathcal{B}|)$, i.e., $s\mathcal{I} - |\mathcal{B}|$ is a nonsingular $\mathcal{M}$-tensor.
It can be verified directly that an absolutely $\mathcal{M}$-tensor must be a nonsingular $\mathcal{H}$-tensor, since $s - |b_{ii\ldots i}| \leq |s - b_{ii\ldots i}|$. An *absolutely $\mathcal{M}$-equation* is referred to

$$\mathcal{A}\mathbf{x}^{m-1} = \mathbf{b},$$

where $\mathcal{A}$ is an absolutely $\mathcal{M}$-tensor.
An *absolutely $\mathcal{M}$-equation* is referred to

$$\mathcal{A}\mathbf{x}^{m-1} = \mathbf{b},$$

where $\mathcal{A}$ is an absolutely $\mathcal{M}$-tensor.

Employ the Brouwer fixed-point theorem

Theorem (Brouwer)

Let Ω be a bounded closed convex set in $\mathbb{R}^n$. If map $F : \Omega \to \Omega$ is continuous, then there exists $\mathbf{x}_* \in \Omega$ such that $F(\mathbf{x}_*) = \mathbf{x}_*$.

When m is even, we also consider the fixed-point iteration

$$\mathbf{x}_k = F(\mathbf{x}_{k-1}) := (s^{-1}\mathcal{B}\mathbf{x}_{k-1}^{m-1} + s^{-1}\mathbf{b})^{[1/(m-1)]}, \ k = 1, 2, \ldots .$$

Applying Brouwer's fixed point theorem, we can prove that $F(\mathbf{x})$ must have a fixed point in $\mathbb{R}^n$.

Theorem

An even order absolutely $\mathcal{M}$-equation has at least one real solution for every real right-hand side.

The inverse matrix of a banded nonsingular M-matrix has a so-called decay property that is, its entries decay exponentially from the diagonal to the corners. It is interesting that a banded nonsingular $\mathcal{M}$-tensor has a similar property.

Although there is no "inverse tensor" for a general $\mathcal{M}$-tensor, we can express this property as that the entries of the minimal nonnegative solution to $\mathcal{A}\mathbf{x}^{m-1} = \mathbf{e}_p$ decay exponentially from the p-th entries to both the ends, where $\mathbf{e}_p$ is the p-th orthonormal column vector.

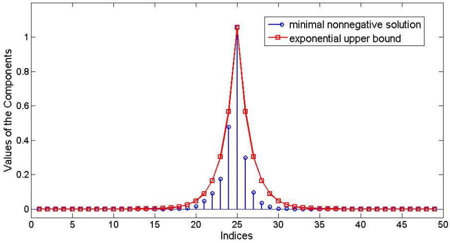
Banded $\mathcal{M}$-Equation

Let $\mathcal{A} = \mathcal{I} - \mathcal{B}$ be an m^{th}-order nonsingular $\mathcal{M}$-tensor satisfying that $\mathcal{A}$ is strictly diagonally dominant, i.e., $\mathcal{A}\mathbf{1}^2 > \mathbf{0}$. Furthermore, we also assume that $\mathcal{B}$ is d-banded, i.e., $b_{i_1 i_2 \dots i_m} = 0$ if $(i_1, i_2, \dots, i_m)$ lays out of

$$\Omega_{i_1} := \left\{(i_1, i_2, \dots, i_m) : |i_s - i_t| \leq d \text{ for all } s, t = 1, 2, \dots, m\right\}.$$

The minimal nonnegative solution solution $\mathbf{x}_*$ of $\mathcal{A}\mathbf{x}^{m-1} = \mathbf{e}_p$ satisfies

$$(\mathbf{x}_*)_i \leq (\mathbf{x}_*)_p \cdot \widehat{\xi}^{|i-p|}.$$



Split the coefficient tensor $\mathcal{A}$ into $\mathcal{A} = \mathcal{M} - \mathcal{N}$ satisfying that the equations with coefficient tensor $\mathcal{M}$ are easy to solve and $\mathcal{N}$ is nonnegative. Then the iteration

$$\mathbf{x}_k = \mathcal{M}_{++}^{-1}(\mathcal{N}\mathbf{x}_{k-1}^{m-1} + \mathbf{b}), \; k = 1, 2, \ldots$$

offers a nonnegative solution to the equation above if it converges.

Different iterative methods for nonsingular $\mathcal{M}$-equations.

Iterative Methods	Alternatives of $\mathcal{M}$
Jacobi	diagonal part
forward G-S	lower triangular part
simplified forward G-S	strictly lower triangular part & diagonal part
backward G-S	upper triangular part
simplified backward G-S	strictly upper triangular part & diagonal part

Consider an operator $\phi : \mathbb{R}^n \to \mathbb{R}^n$. Let $\mathbf{x}_*$ be a fixed point of $\phi(\mathbf{x})$. Then we call $\mathbf{x}_*$ an attracting fixed point, if there exists $\delta > 0$ such that the sequence $\{\mathbf{x}_k\}$ defined by $\mathbf{x}_{k+1} = \phi(\mathbf{x}_k)$ converges to $\mathbf{x}_*$ for any $\mathbf{x}_0$ such that $\|\mathbf{x}_0 - \mathbf{x}_*\| \leq \delta$.

Theorem (Rheinboldt 1998)

Let $\mathbf{x}_*$ be a fixed point of $\phi : \mathbb{R}^n \to \mathbb{R}^n$, and let $\nabla\phi : \mathbb{R}^n \to \mathbb{R}^{n\times n}$ be the Jacobian of ϕ. Then $\mathbf{x}_*$ is an attracting fixed point if $\sigma := \rho\big(\nabla\phi(\mathbf{x}_*)\big) < 1$; further, if $\sigma > 0$, then the convergence of $\mathbf{x}_{k+1} = \phi(\mathbf{x}_k)$ to $\mathbf{x}_*$ is linear with rate σ.

Convergence of the Classical Iterations I

Derive the Jacobian of the operator

$$\phi(\mathbf{x}) = \mathcal{M}_{++}^{-1}(\mathcal{N}\mathbf{x}^{m-1} + \mathbf{b}).$$

take gradients on both sides of

$$\mathcal{M}\phi(\mathbf{x})^{m-1} = \mathcal{N}\mathbf{x}^{m-1} + \mathbf{b},$$

and we obtain

$$\mathcal{M}\phi(\mathbf{x})^{m-2} \cdot \nabla\phi(\mathbf{x}) = \mathcal{N}\mathbf{x}^{m-2}.$$

When we take $\mathbf{x}$ as the positive fixed point $\mathbf{x}_*$, then the matrix $\mathcal{M}\phi(\mathbf{x}_*)^{m-2} = \mathcal{M}\mathbf{x}_*^{m-2}$ is a nonsingular M-matrix, since $\mathbf{x}_* > \mathbf{0}$ and

$$\mathcal{M}\mathbf{x}_*^{m-2} \cdot \mathbf{x}_* = \mathcal{N}\mathbf{x}_*^{m-1} + \mathbf{b} \geq \mathbf{b} > \mathbf{0}.$$

The Jacobian of $\phi(\mathbf{x})$ at $\mathbf{x}_*$ is

$$\nabla\phi(\mathbf{x}_*) = \left(\mathcal{M}\mathbf{x}_*^{m-2}\right)^{-1}\mathcal{N}\mathbf{x}_*^{m-2},$$

which is a nonnegative matrix. Since

$$\mathcal{N}\mathbf{x}_*^{m-2}\cdot\mathbf{x}_* = \mathcal{M}\mathbf{x}_*^{m-1} - \mathbf{b} \le \theta\mathcal{M}\mathbf{x}_*^{m-1}$$

with $0 \le \theta < 1$, then $\nabla\phi(\mathbf{x}_*)\cdot\mathbf{x}_* \le \theta^{1/(m-1)}\mathbf{x}_*$.

Therefore the spectral radius $\rho\big(\nabla\phi(\mathbf{x}_*)\big) \le \theta^{1/(m-1)} < 1$, which indicates that $\mathbf{x}_*$ is an attracting fixed point of ϕ.

Convergence of the Classical Iterations III

Since $\mathcal{A}$ is a nonsingular $\mathcal{M}$-tensor, we can take an initial vector $\mathbf{x}_0$ satisfying that

$$\mathbf{0} < \mathcal{A}\mathbf{x}_0^{m-1} \leq \mathbf{b},$$

then we shall prove that the iteration

$$\mathbf{x}_k = \mathcal{M}_{++}^{-1}(\mathcal{N}\mathbf{x}_{k-1}^{m-1} + \mathbf{b}), \;\; k = 1, 2, \ldots$$

converges to the solution to the positive nonsingular $\mathcal{M}$-equation with a positive right-hand side.

An SOR-Like Acceleration

An SOR-like acceleration can also be applied to this iterative method. For instance, we can choose a proper $\omega > 0$ so that the iterative scheme

$$\mathbf{x}_k = (\mathcal{M} - \omega\mathcal{I})_{++}^{-1}\left[(\mathcal{N} - \omega\mathcal{I})\mathbf{x}_{k-1}^{m-1} + \mathbf{b}\right], \; k = 1, 2, \ldots$$

converges faster than the original scheme. The acceleration effect is due to a smaller α in the above discussion. There are some restrictions when choosing the parameter ω, such as

1. $\omega > 0$,
2. $\mathcal{M} - \omega\mathcal{I}$ is still a nonsingular $\mathcal{M}$-tensor,
3. $(\mathcal{N} - \omega\mathcal{I})\mathbf{x}_{k-1}^{m-1} + \mathbf{b} > \mathbf{0}$ for all $k = 1, 2, \ldots$.

However, whether there is an optimal parameter ω and how to choose it still remain as open questions.

A Numerical Example

We construct a 3rd-order nonsingular $\mathcal{M}$-tensor $\mathcal{A} = s\mathcal{I} - \mathcal{B}$ as follows. First, we generate a nonnegative tensor $\mathcal{B} \in \mathbb{R}_+^{n \times n \times n}$ containing random values drawn from the standard uniform distribution on $(0, 1)$. Next, set the scalar

$$s = (1 + \varepsilon) \cdot \max_{i=1,2,\ldots,n} (\mathcal{B}\mathbf{1}^2)_i, \ \varepsilon > 0,$$

where $\mathbf{1} = (1, 1, \ldots, 1)^\top$. Obviously, $\mathcal{A}$ is a diagonally dominant $\mathcal{Z}$-tensor, i.e., $\mathcal{A}\mathbf{1}^2 > \mathbf{0}$. Thus $\mathcal{A}$ is a nonsingular $\mathcal{M}$-tensor. In this example, we take $n = 10$ and $\varepsilon = 0.01$.

The way we select the acceleration parameter is

$$\omega = \tau \cdot \min_{i=1,2,\ldots,n} a_{ii\ldots i}, \ 0 < \tau < 1.$$

This ensures the first two conditions discussed in the above subsection and the last condition when $\mathbf{x}_{k-1}$ is close to the solution. In this experiment, we take $\tau = 0.35$, which is chosen by experience.

A Numerical Example I

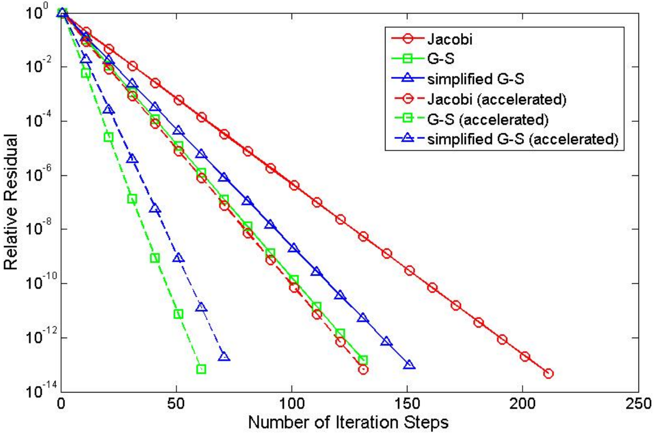


Figure: The comparison of different classical iterative methods for $\mathcal{M}$-equations.

The Newton Method I

When the coefficient tensor $\mathcal{A}$ is a symmetric nonsingular $\mathcal{M}$-tensor, computing the positive solution is equivalent to solving the optimization problem

$$\min_{\mathbf{x}>\mathbf{0}} \varphi(\mathbf{x}) := \frac{1}{m}\mathcal{A}\mathbf{x}^m - \mathbf{x}^\top \mathbf{b},$$

where

$$\nabla\varphi(\mathbf{x}) = \mathcal{A}\mathbf{x}^{m-1} - \mathbf{b} =: -\mathbf{r} \text{ and } \nabla^2\varphi(\mathbf{x}) = (m-1)\mathcal{A}\mathbf{x}^{m-2}.$$

Note that when $\mathcal{A}\mathbf{x}^{m-1} > \mathbf{0}$, matrix $\mathcal{A}\mathbf{x}^{m-2}$ is obviously a symmetric Z-matrix and

$$\mathcal{A}\mathbf{x}^{m-2} \cdot \mathbf{x} = \mathcal{A}\mathbf{x}^{m-1} > \mathbf{0}.$$

Therefore, $\mathcal{A}\mathbf{x}^{m-2}$ is a symmetric nonsingular M-matrix, and thus a positive definite matrix.

The Newton Method II

Then the Newton step

$$\mathbf{p}_k = -\left[\nabla^2 \varphi(\mathbf{x}_k)\right]^{-1} \nabla \varphi(\mathbf{x}_k) = \frac{1}{m-1}(\mathcal{A}\mathbf{x}_k^{m-2})^{-1}\mathbf{r}_k$$

is ensured to be a descending direction.

Then the iterations are as follows

$$\begin{cases} M_k = \mathcal{A}\mathbf{x}_k^{m-2}, \\ \mathbf{r}_k = \mathbf{b} - M_k \mathbf{x}_k, \\ \mathbf{p}_k = \frac{1}{m-1} M_k^{-1} \mathbf{r}_k, \\ \mathbf{x}_{k+1} = \mathbf{x}_k + \mathbf{p}_k, \end{cases} \quad k = 0, 1, 2, \ldots$$

Numerical Example II

We construct a symmetric $\mathcal{M}$-tensor of size $10 \times 10 \times 10$ by the following way. Let $\mathcal{B} \in \mathbb{R}^{10\times 10\times 10}$ be a nonnegative tensor with

$$b_{i_1 i_2 i_3} = |\tan(i_1 + i_2 + i_3)|.$$

It can be computed that $\rho(\mathcal{B}) \approx 1450.3$. Thus $\mathcal{A} = 1500\mathcal{I} - \mathcal{B}$ is a symmetric nonsingular $\mathcal{M}$-tensor. We apply the Newton method, the accelerated Jacobi method, the accelerated Gauss-Seidel method, and the accelerated simplified Gauss-Seidel method to this equation, respectively. And the acceleration parameter for the Jacobi and Gauss-Seidel method is taken as $\tau = 400$, which is experimentally optimal. The figure shows the comparison on the number of iteration steps, from which we can see that the Newton method converges much faster than other algorithms in this situation.

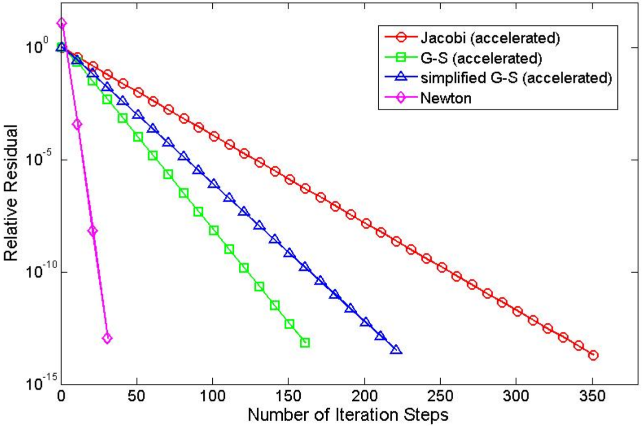


Figure: Comparison of the Newton method and other iterative methods.

A Toy Example

Consider the ordinary differential equation

$$\frac{\mathrm{d}^2 x(t)}{\mathrm{d}t^2} = -\frac{f(t)}{x(t)^2} \text{ in } (0,1),$$

with Dirichlet's boundary conditions

$$x(0) = c_0, \quad x(1) = c_1.$$

Assume that $f(t) > 0$ on $[0,1]$, $c_0 > 0$, $c_1 > 0$, and we require a positive solution $x(t)$ on $[0,1]$. This equation can describe a particle's movement under the gravitation

$$m\frac{\mathrm{d}^2 x}{\mathrm{d}t^2} = -G\frac{Mm}{x^2} \ \rightarrow \ x^2 \cdot \frac{\mathrm{d}^2 x}{\mathrm{d}t^2} = -GM,$$

where $G \approx 6.67 \times 10^{-11} \mathrm{N} \cdot \mathrm{m}^2/\mathrm{kg}^2$ is the gravitational constant and $M \approx 5.98 \times 10^{24} \mathrm{kg}$ is the mass of the earth.

A Toy Example

After the discretization, we get a system of polynomial equations

$$\begin{cases} x_1^3 = c_0^3, \\ 2x_i^3 - x_i^2 x_{i-1} - x_i^2 x_{i+1} = \frac{GM}{(n-1)^2}, \ \ i = 2, 3, \dots, n-1, \\ x_n^3 = c_1^3. \end{cases}$$

Then this can be rewritten into a multi-linear equation $\mathcal{A}\mathbf{x}^3 = \mathbf{b}$, where the coefficient tensor $\mathcal{A}$ is a tensor with seven diagonals, i.e., $\mathcal{A}$ is 1-banded, and

$$\begin{cases} a_{1,1,1,1} = a_{n,n,n,n} = 1, \\ a_{i,i,i,i} = 2, \ \ i = 2, 3, \dots, n-1, \\ a_{i,i-1,i,i} = a_{i,i,i-1,i} = a_{i,i,i,i-1} = -1/3, \ \ i = 2, 3, \dots, n-1, \\ a_{i,i+1,i,i} = a_{i,i,i+1,i} = a_{i,i,i,i+1} = -1/3, \ \ i = 2, 3, \dots, n-1, \end{cases}$$

and the right-hand side is a positive vector with

$$\begin{cases} b_1 = c_0^2, \\ b_i = \frac{GM}{(n-1)^2}, \ \ i = 2, 3, \dots, n-1, \\ b_n = c_1^2. \end{cases}$$

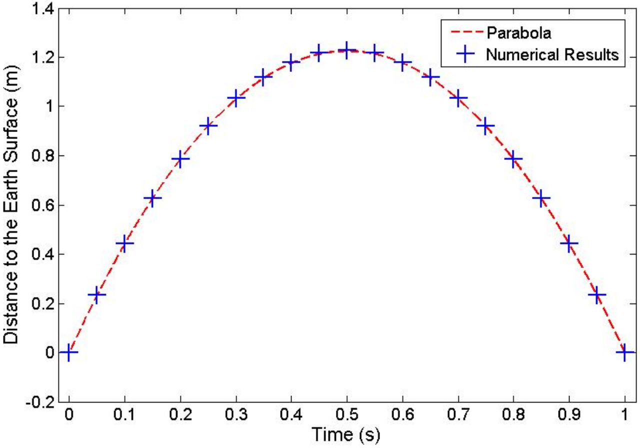


Figure: A particle's movement under the earth's gravitation.

The power method for matrix eigenproblems is extended to nonnegative tensor eigenproblems, often called the NQZ method, and its convergence is widely studied.

L. Elsner. Inverse iteration for calculating the spectral radius of a non-negative irreducible matrix. Linear Algebra and Appl., 15(3):235–242, 1976.

Z. Jia, W.-W. Lin, and C.-S. Liu. A positivity preserving inexact Noda iteration for computing the smallest eigenpair of a large irreducible M-matrix. Numer. Math., 130:645–679, 2015.

M. Ng, L. Qi, and G. Zhou. Finding the largest eigenvalue of a nonnegative tensor. SIAM J. Matrix Anal. Appl., 31(3):1090–1099, 2009.

K.-C. Chang, K. J. Pearson, and T. Zhang. Primitivity, the convergence of the NQZ method, and the largest eigenvalue for nonnegative tensors. SIAM J. Matrix Anal. Appl., 32(3):806–819, 2011.

Inverse Iteration II

Algorithm (Inverse Iteration)

If $\mathcal{B}$ is an m^{th}-order n-dimensional nonnegative tensor, then this algorithm gives the spectral radius of $\mathcal{B}$ and the corresponding eigenvector when converges.

$\mathbf{x}_0 = $ an approximated eigenvector
$s_0 = (1+\varepsilon) \cdot \max_i (\mathcal{B}\mathbf{x}_0^{m-1})_i / (\mathbf{x}_0^{[m-1]})_i$
for $k = 1, 2, \ldots$
    $\mathbf{y}_k = (s_{k-1}\mathcal{I} - \mathcal{B})_{++}^{-1} (\mathbf{x}_{k-1}^{[m-1]})$
    if $\angle(\mathbf{y}_k^{[m-1]}, s_{k-1}\mathbf{y}_k^{[m-1]} - \mathbf{x}_{k-1}^{[m-1]})$ is small enough
        break
    end
    $s_k = (1+\varepsilon) \cdot \left[s_{k-1} - \left(\min_i (\mathbf{x}_{k-1})_i / (\mathbf{y}_k)_i \right)^{m-1} \right]$
    $\mathbf{x}_k = \mathbf{y}_k / \|\mathbf{y}_k\|$
end

Numerical Examples

I. We generate a nonnegative tensor $\mathcal{B} \in \mathbb{R}_+^{10\times10\times10}$ containing random values drawn from the standard uniform distribution on $(0,1)$. This is an irreducible example.

II. We also generate a nonnegative tensor $\mathcal{B} \in \mathbb{R}_+^{10\times10\times10}$ containing random values drawn from the standard uniform distribution on $(0,1)$ first. Then set $b_{i_1 i_2 i_3} = 0$ if $i_1 \geq 7$ and $i_2, i_3 < 7$. Thus this tensor is reducible.

III. The third test tensor is a nonnegative tensor $\mathcal{B} \in \mathbb{R}_+^{10\times10\times10}$ with $b_{i_1 i_2 i_3} = |\tan(i_1 + i_2 + i_3)|$, which is symmetric.

IV. Let $\mathcal{B} \in \mathbb{R}_+^{3\times3\times3}$ with $b_{133} = b_{233} = b_{311} = b_{322} = 1$ and other entries being zeros. Since the power method does not work for this example, we compare the inverse iteration with the shifted power method. And we select an experimental optimal shift. Moreover, the multi-linear equation $(s\mathcal{I} - \mathcal{B})\mathbf{x}^2 = \mathbf{b}$ is equivalent to a linear equation

$$\begin{bmatrix} s & 0 & -1 \\ 0 & s & -1 \\ -1 & -1 & s \end{bmatrix} \cdot \begin{bmatrix} x_1^2 \\ x_2^2 \\ x_3^2 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix},$$

Numerical Example III

Compare of the inverse iteration and the (shifted) power method.

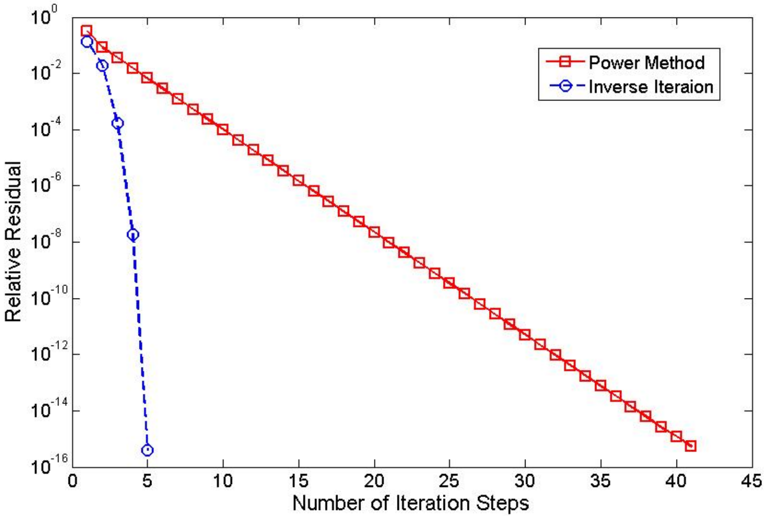
	I		II		III		IV	
	Steps	Time	Steps	Time	Steps	Time	Steps	Time
Inverse	4	7.38	6	11.00	5	9.89	5	0.58
Power	11	2.65	33	7.93	41	9.90	NC	NC
Shifted	–	–	–	–	–	–	25	1.41

Some Remarks

Each algorithm's number of iteration steps and the running time of 1000 times of experiments are listed in Table. Since the shifted power method has no apparent advance to the original power method for the first three examples, we just compare the inverse iteration with the power method. From Table, we can conclude that the inverse iteration always converges faster than the power method or the shifted power method.

The following Figure displays the convergence of the inverse iteration and the power method for the third example. From this Figure, we can guess that the inverse iteration converges quadratically, and this is true for the matrix case. Nevertheless, we cannot prove this conjecture and remain it as an open question.

The comparison of inverse iteration and power method



Consider the following the ordinary differential equation with Dirichlets boundary condition

$$\begin{cases} u(x)^{m-2} \cdot u''(x) = -f(x), \ x \in (0,1), \\ u(0) = g_0, \ u(1) = g_1, \end{cases}$$

where $f(x) > 0$ in $(0,1)$ and $g_0, g_1 > 0$.
Partition the interval $[0,1]$ into $n-1$ small intervals with the same length $h = 1/(n-1)$, and denote

$$\mathbf{u}_h = \Big[u(0), u(h), \ldots, u\big((n-1)h\big)\Big]^\top,$$
$$\mathbf{v}_h = \Big[-u^{(4)}(0), -u^{(4)}(h), \ldots, -u^{(4)}\big((n-1)h\big)\Big]^\top,$$
$$\mathbf{f}_h = \Big[g_0/h^2, f(h), \ldots, f\big((n-2)h\big), g_1/h^2\Big]^\top,$$

where $u(x)$ is the exact solution of the above boundary-value problem.

The discretization tensor $\mathcal{L}_h$ of the operator $u \mapsto u^{m-2} \cdot u''$ is introduced in the first section, and our numerical solution $\widehat{\mathbf{u}}_h$ is obtained by solving the unique positive solution of an $\mathcal{M}$-equation $\mathcal{L}_h \widehat{\mathbf{u}}_h^{m-1} = \mathbf{f}_h$. It is well-known that the truncated error of the discretization

$$\frac{u(x-h)-2u(x)+u(x+h)}{h^2} - u''(x) = \frac{h^2}{12} u^{(4)}(x) + O(h^4).$$

Thus we have

$$\mathcal{L}_h \mathbf{u}_h^{m-1} - \mathcal{L}_h \widehat{\mathbf{u}}_h^{m-1} = \mathcal{L}_h \mathbf{u}_h^{m-1} - \mathbf{f}_h = \frac{h^2}{12} \cdot \mathbf{u}_h^{[m-2]} \circ \mathbf{v}_h + O(h^4),$$

which further implies that

$$d_h(\mathbf{u}_h, \widehat{\mathbf{u}}_h) := \left\| \mathcal{L}_h \mathbf{u}_h^{m-1} - \mathcal{L}_h \widehat{\mathbf{u}}_h^{m-1} \right\|_\infty \leq \frac{h^2}{12} \cdot \|u\|_{L^\infty}^{m-2} \cdot \|u^{(4)}\|_{L^\infty} + O(h^4).$$

It can be verified that $d_h(\cdot,\cdot)$ is a metric in the cone $\{\mathbf{x} > \mathbf{0} : \mathcal{L}_h \mathbf{x}^{m-1} > \mathbf{0}\}$. Then we can say that the numerical solution $\widehat{\mathbf{u}}_h$ is very close to the exact solution $\mathbf{u}_h$ when the parameter h is small enough. Next, we shall estimate the convergence of the discretization scheme.

Note that $\mathcal{L}_h \mathbf{u}_h^{m-1}$ is also a positive vector when h is small enough, then the matrix $\mathcal{L}_h \mathbf{u}_h^{m-2}$ is a nonsingular M-matrix as discussed in Section 4. Hence we have the first order approximation

$$\mathbf{u}_h - \widehat{\mathbf{u}}_h \approx (\mathcal{L}_h \mathbf{u}_h^{m-2})^{-1}(\mathcal{L}_h \mathbf{u}_h^{m-1} - \mathcal{L}_h \widehat{\mathbf{u}}_h^{m-1})$$

when h is small enough, and thus

$$\|\mathbf{u}_h - \widehat{\mathbf{u}}_h\|_\infty \lesssim \left\|(\mathcal{L}_h \mathbf{u}_h^{m-2})^{-1}\right\|_\infty \cdot \left\|\mathcal{L}_h \mathbf{u}_h^{m-1} - \mathcal{L}_h \widehat{\mathbf{u}}_h^{m-1}\right\|_\infty.$$

We thus need to bound the ∞-norm of the inverse of the M-matrix $\mathcal{L}_h \mathbf{u}_h^{m-2}$. First denote $\mathcal{L}_h = s_h \mathcal{I} - \mathcal{A}_h$, where $s_h = 2/h^2$ and $\mathcal{A}_h$ is nonnegative. Then we can write

$$\begin{aligned}
(\mathcal{L}_h \mathbf{u}_h^{m-2})^{-1} &= \left((s_h \mathcal{I} - \mathcal{A}_h)\mathbf{u}_h^{m-2}\right)^{-1} \\
&= \left(s_h U_h^{m-2} - \mathcal{A}_h \mathbf{u}_h^{m-2}\right)^{-1} \\
&= s_h^{-1} U_h \left[I - s_h^{-1} U_h^{-(m-1)}(\mathcal{A}_h \mathbf{u}_h^{m-2}) U_h\right]^{-1} U_h^{-(m-1)},
\end{aligned}$$

where $U_h = \mathrm{diag}\big((\mathbf{u}_h)_1, (\mathbf{u}_h)_2, \ldots, (\mathbf{u}_h)_n\big)$.

Denote $W_h = U_h^{-(m-1)}(\mathcal{A}_h \mathbf{u}_h^{m-2})U_h$, which is a nonnegative matrix. Note that $(\mathcal{A}_h \mathbf{u}_h^{m-2})U_h \mathbf{1} = \mathcal{A}_h \mathbf{u}_h^{m-1}$, thus the summations of all the rows of W_h are

$$\begin{aligned} W_h \mathbf{1} = U_h^{-(m-1)} \mathcal{A}_h \mathbf{u}_h^{m-1} &\leq \mathbf{1} \cdot \max_{i=1:n} \frac{(\mathcal{A}_h \mathbf{u}_h^{m-1})_i}{(\mathbf{u}_h)_i^{m-1}} \\ &= \mathbf{1} \cdot \max_{i=1:n} \frac{s_h (\mathbf{u}_h)_i^{m-1} - (\mathcal{L}_h \mathbf{u}_h^{m-1})_i}{(\mathbf{u}_h)_i^{m-1}} \\ &= \mathbf{1} \cdot \left[s_h - \min_{i=1:n} \frac{(\mathcal{L}_h \mathbf{u}_h^{m-1})_i}{(\mathbf{u}_h)_i^{m-1}} \right] \\ &=: \mathbf{1} \cdot (s_h - \gamma_h). \end{aligned}$$

Similarly, we have $W_h^k \mathbf{1} \leq W_h^{k-1} \mathbf{1} \cdot (s_h - \gamma_h) \leq \cdots \leq \mathbf{1} \cdot (s_h - \gamma_h)^k$. Also, because $W_h \mathbf{1} \leq \mathbf{1} \cdot (s_h - \gamma_h) < \mathbf{1} \cdot s_h$ and W_h is an irreducible nonnegative matrix, we have $\rho(s_h^{-1} W_h) < 1$.

Employ the Taylor expansion of the matrix $(I-X)^{-1} = I + X + X^2 + \dots$ for $\rho(X) < 1$, and we can obtain that

$$(I - s_h^{-1} W_h)^{-1}\mathbf{1} = \sum_{k=0}^{\infty} (s_h^{-1} W_h)^k \mathbf{1} \leq \sum_{k=0}^{\infty} \mathbf{1} \cdot (1 - \gamma_h / s_h)^k = \mathbf{1} \cdot (s_h / \gamma_h).$$

Finally, we get a upper bound of the ∞-norm of the nonnegative matrix $(\mathcal{L}_h \mathbf{u}_h^{m-2})^{-1}$ that

$$\begin{aligned}
\|(\mathcal{L}_h \mathbf{u}_h^{m-2})^{-1}\|_\infty &= \max_{i=1:n} \left((\mathcal{L}_h \mathbf{u}_h^{m-2})^{-1} \mathbf{1} \right)_i \\
&= \max_{i=1:n} \left(s_h^{-1} U_h \left(I - s_h^{-1} W_h \right)^{-1} U_h^{-(m-1)} \mathbf{1} \right)_i \\
&\leq \left(\min_{i=1:n} \mathbf{u}_h \right)^{-(m-1)} \cdot \max_{i=1:n} \frac{(\mathbf{u}_h)_i^{m-1}}{(\mathcal{L}_h \mathbf{u}_h^{m-1})_i} \cdot \max_{i=1:n} \mathbf{u}_h \\
&\approx \left(\min_{i=1:n} \mathbf{u}_h \right)^{-(m-1)} \cdot \max_{i=1:n} \frac{(\mathbf{u}_h)_i^{m-1}}{(\mathbf{f}_h)_i} \cdot \max_{i=1:n} \mathbf{u}_h \\
&\leq \frac{\max_x u(x)^m}{\min_x u(x)^{m-1} \cdot \min_x f(x)}.
\end{aligned}$$

Note that $u(x)$ can also be regarded as the solution of the elliptic problem

$$\begin{cases} u''(x) = f_1(x) := -f(x)/u(x)^{m-2}, \ x \in (0,1), \\ u(0) = g_0, \ u(1) = g_1. \end{cases}$$

So we know that $u(x) \geq \min\{g_0, g_1\}$ since $f(x) > 0$ and $g_0, g_1 > 0$. Then

$$\|\mathbf{u}_h - \widehat{\mathbf{u}}_h\|_\infty \lesssim \frac{\|u\|_{L^\infty}^m}{\min\{g_0,g_1\}^{m-1} \cdot \min_x f(x)} \cdot \frac{h^2}{12} \cdot \|u\|_{L^\infty}^{m-2} \cdot \|u^{(4)}\|_{L^\infty} =: Kh^2,$$

where the constant K is independent with the parameter h. Therefore, the sequence $\{\widehat{\mathbf{u}}_h\}$ converges to the exact solution when $h \to 0$.

Conclusions

Let $\mathcal{A}$ be a $\mathcal{Z}$-tensor. Then the following three conditions are equivalent:

1. $\mathcal{A}$ is a nonsingular $\mathcal{M}$-tensor;
2. $(\mathcal{A}^{-1}\mathbf{b})_{++}$ has a unique element for every positive vector $\mathbf{b}$;
3. $\mathcal{A}$ does not reverse the sign of any vector; that is, if $\mathbf{x} \neq \mathbf{0}$ and $\mathbf{b} = \mathcal{A}\mathbf{x}^{m-1}$, then $x_i^{m-1} b_i > 0$ for some subscript i.

We also extend some algorithms and concepts in matrix computations, such as the Jacobi method, the Gauss-Seidel method, the Newton method, the inverse iteration, etc., to the higher-order tensor cases.

Furthermore, many difficulties appear when extending the results from the linear case to the multi-linear case, many of which are left as open questions throughout this paper.

Thank You for Your Attention!