Methods for Solving Tensor Eigenvalue Problems

Rhea Shroff

Advisor : Dr. Sara Pollock, University of Florida

September 23, 2022

Outline

Tensors

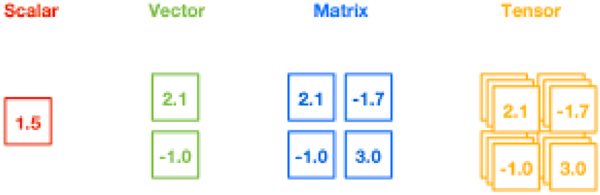
Definition

A real m^{th} order n^{th} dimensional tensor $\mathcal{A}$ consists of n^m real entries $\mathcal{A}_{i_1,\ldots,i_m} \in \mathbb{R}$ where $i_j = 1, \ldots, n$ for $j = 1, \ldots, m$.

Here, $\mathcal{A} \in \mathbb{R}^{n \times n \ldots \ldots \times n}$ or $\mathbb{R}^{[m,n]}$, where the n in the superscript is multiplied m times, i.e., it has m modes.

Since each mode of the tensor has the same dimension, we say that the tensor is **square**.

A tensor is an algebraic object that describes a multi-linear relationship between sets of algebraic objects related to a vector space.



But a tensor is more than a supermatrix.

Inspiration from Matrix Problems

Since tensors can be represented as supermatrices, one hopes that we can extend the known results for matrices to them. Some examples :

- Hyperdeterminants, Cayley (1845)
- Positive definite tensors
- Characteristic Equations

Tensor Eigenvalues

Definition

A tensor $\mathcal{A}$ is called supersymmetric if its entries are invariant under any permutation of the indices.

Definition

Assume that tensor $\mathcal{A}$ is a symmetric m^{th} order n^{th} dimensional real valued tensor. For any n dimensional vector x.

$$(Ax^{m-1})_{i_1} = \sum_{i_2=1}^{n} \dots \sum_{i_m=1}^{n} A_{i_1 i_2 \dots i_m} x_{i_2} \dots x_{i_m} \text{ ; for } i_1 = 1, \dots, n$$

General Definition

Let $1 < p < +\infty$. A real number $\lambda \in \mathbb{R}$ is said to be the ℓ^p eigenvalues of $\mathcal{A}$, a symmetric m^{th} order n^{th} dimensional real valued tensor, if the following holds

$$\mathcal{A}x^{m-1} = \lambda\phi_p(x) \ , \ ||x||_p = 1,$$

where $x \in \mathbb{R}^n$ and $||x||_p$ denotes the usual ℓ^p norm and $\phi_p : \mathbb{R}^n \to \mathbb{R}^n$ is the map entry wise defined by $\phi_p(x)_i = |x_i|^{p-2}x_i = sign(x_i)|x_i|^{p-1}$ for $i = 1, ..., n$. Here, x is the eigenvector corresponding to λ. (Cipolla et al., 2020)

Definition

Assume that tensor $\mathcal{A}$ is a symmetric m^{th} order n^{th} dimensional real valued tensor. Then $\lambda \in \mathbb{R}$ is an ℓ^2 eigenvalue (or E - eigenvalue) of $\mathcal{A}$ if there exists $x \in \mathbb{R}^n$ such that

$$Ax^{m-1} = \lambda x \text{ and } x^T x = 1.$$

Here x is the E - eigenvector associated with *lambda* and (λ, x) is the E - eigenpair.

If however, the E eigenvalue has a real E eigenvector then it is called a Z eigenvalue. This is what the methods we will be discussing here primarily focus on.

- This is similar to the eigenpair definition for matrices.
- One important difference - Normalization Requirement

Another important class of eigenvalues for tensors is the ℓ^m-eigenvalues, where m is the order of the tensor.

- When the ℓ^m-eigenvalue has a real eigenvector, it is called an H - eigenvalue.
- Classical properties of eigenvalues of matrices can be extrapolated for these
  - Number of eigenvalues
  - Diagonal tensors
  - Gerschgorin type theorem

These properties do not however hold true for the E-eigenvalues.

- There are no fixed number of E - eigenvalues.
- They are invariant for the tensor.

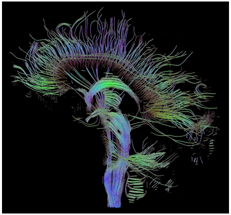
But why is this a reasonable problem to tackle?

Theorem

A supersymmetric tensor/ supermatrix always has Z - eigenvalues. (Qi, 2007)

Areas of Applications

- Spectral hypergraph theory
- Diffusion magnetic resonance imaging
- Blind source separation
- Geometric measure of entanglement



Tensors are not just supermatrices but for our purpose we will use them interchangeably.

Theorem

The E eigenvalues of a tensor are the same as the E eigenvalues of its representation supermatrix under an orthonormal basis. (Qi, 2007)

Additionally, we will refer to E - eigenvalues as eigenvalues for the rest of the presentation.

Problem Statement

Using an approach similar to that for eigenvalues of a symmetric matrix A, where we want critical points of the Rayleigh Quotient, $x^T Ax/||x||_2^2$.

Our problem for finding eigenpairs can be reduced to finding the KKT point (Karush - Kuhn - Tucker point or constrained stationary point) of the non linear optimization problem

$$\max_{x \in \mathbb{R}^n} \mathcal{A}x^m \text{ subject to } x^T x = 1, \text{ where } \mathcal{A}x^m = \sum_{i_1=1}^{n} \dots \sum_{i_m=1}^{n} a_{i_1 \dots i_m} x_{i_1} \dots . x_{i_m}.$$

(Lim, 2005)

Symmetric Higher Order Power Method (S-HOPM)

For symmetric tensors, we use the following algorithm :

Given a symmetric tensor $\mathcal{A} \in \mathbb{R}^{[m,n]}$.
Given $x_0 \in \mathbb{R}^n$ with $||x_0|| = 1$. Let $\lambda_0 = \mathcal{A}x_0^m$.
for k $= 0, 1, ...$ do
  $\hat{x}_{k+1} \leftarrow \mathcal{A}x_k^{m-1}$
  $x_{k+1} \leftarrow \hat{x}_{k+1}/||\hat{x}_{k+1}||$
  $\lambda_{k+1} \leftarrow \mathcal{A}x_{k+1}^m$
end for

Higher Order Power Method (HOPM)

This is a generalized form of SHOPM, i.e., we remove the symmetric requirement. The input tensors for this need not even be square tensors.

In some special cases we get linear convergence for this method, such as, when there is an orthogonal decomposition of the tensor. (Kofidis and Regalia, 2002)

But this generalized system requires more depth. So we will stick to the simpler case for now.

Convergence for SHOPM

Theorem

Let $\mathcal{A}$ be any supersymmetric m^{th} order n^{th} dimensional tensor, where the associated function $f(x) = \mathcal{A}x^m$ is convex (or concave) on $\mathbb{R}^n$. The SHOPM method converges to a local maximum (or minimum) of the restriction of f to the unit sphere Σ for any initialization, except for saddle points and crest lines leading to such saddle points.

(Kofidis and Regalia, 2002), (Kolda and Mayo, 2011)

Numerical Results - Odd Order Tensor

Example 1.

Let us consider the following odd order tensor $\mathcal{A} \in \mathbb{R}^{[3,3]}$ defined:

$$a_{111} = -0.1281,\ a_{112} = 0.0516,\ a_{113} = -0.0954,\ a_{122} = -0.1958,$$

$$a_{123} = -0.1790,\ a_{133} = -0.2676,\ a_{222} = 0.3251,\ a_{223} = 0.2513,$$

$$a_{233} = 0.1773,\ a_{333} = 0.0338.$$

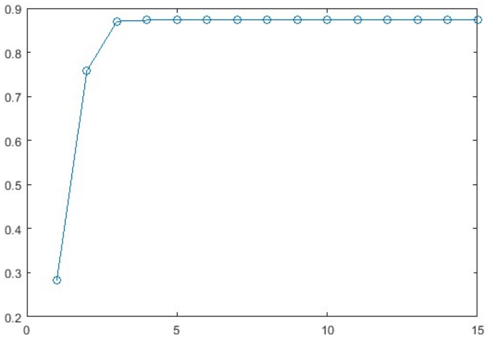


Figure 1: Using S-HOPM for Example 1 : Lambda Converging

Numerical Results - Even Order Tensor

Example 2.

Let us consider the following even order tensor $\mathcal{A} \in \mathbb{R}^{[4,3]}$ defined:

$$a_{1111} = 0.2883,\ a_{1112} = -0.0031,\ a_{1113} = 0.1973,\ a_{1122} = -0.2485,$$

$$a_{1123} = -0.2939,\ a_{1133} = 0.3847,\ a_{1222} = 0.2972,\ a_{1223} = 0.1862,$$

$$a_{1233} = 0.0919,\ a_{1333} = -0.3619,\ a_{2222} = 0.1241,\ a_{2223} = 0.3420,$$

$$a_{2233} = 0.2127,\ a_{2333} = 0.2727,\ a_{3333} = -0.3054.$$

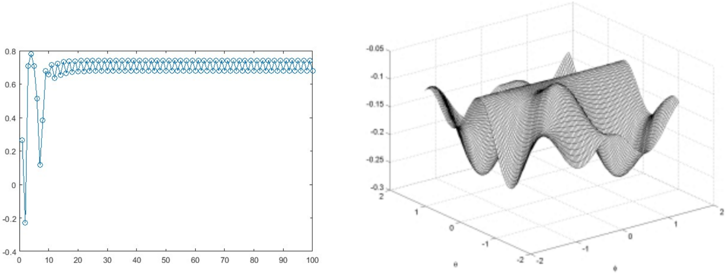


Figure 2: Using S-HOPM for Example 2 : Lambda iterates (left) and underlying function $f(x)$ (right)

Shifted Symmetric HOPM (SS-HOPM)

To overcome the lack convexity (or concavity) in the function, we can modify the function f to enforce it. We can do so by adding a shift. (Kolda and Mayo, 2011)

Given a symmetric tensor $\mathcal{A} \in \mathbb{R}^{[m,n]}$.
Given $x_0 \in \mathbb{R}^n$ with $||x_0|| = 1$. Let $\lambda_0 = \mathcal{A}x_0^m$.
Given $\alpha \in \mathbb{R}$
for k $= 0, 1, ...$ do
  if $\alpha \geq 0$ then   (Assume Convex)
    $\hat{x}_{k+1} \leftarrow \mathcal{A}x_k^{m-1} + \alpha x_k$
  else   (Assume Concave)
    $\hat{x}_{k+1} \leftarrow -(\mathcal{A}x_k^{m-1} + \alpha x_k)$
  end if
  $x_{k+1} \leftarrow \hat{x}_{k+1}/||\hat{x}_{k+1}||$
  $\lambda_{k+1} \leftarrow \mathcal{A}x_{k+1}^m$
end for

Convergence for SSHOPM

For an appropriate choice of α we can guarantee convergence of the eigenvalues. In the convex case :

$$\alpha > \beta(\mathcal{A}) \equiv (m-1) \max_{x \in \Sigma} \rho(\mathcal{A}x^{m-2}).$$

In the concave case : $\alpha < -\beta(\mathcal{A})$.

The underlying function here becomes $\hat{f}(x) \equiv f(x) + \alpha(x^T x)^{m/2}$

So can we take any really large value for α? Yes, but it is not advisable.

Convergence for SSHOPM

Theorem

Let $\mathcal{A} \in \mathbb{R}^{[m,n]}$ be symmetric. For $\alpha > \beta(\mathcal{A})$, the iterates $\{\lambda_k, x_k\}$ produced by the SS-HOPM algorithm satisfy the following properties.

(a) The sequence $\{\lambda_k\}$ is non decreasing and there exists a λ_* such that $\lambda_k \to \lambda_*$.

(b) The sequence $\{x_k\}$ has an accumulation point. For every such accumulation point x_*, the pair (λ_*, x_*) is an eigenpair of $\mathcal{A}$.

(c) If $\mathcal{A}$ has finitely many real eigenvectors, then there exists x_* such that $x_k \to x_*$.

(Kolda and Mayo, 2011)

Numerical Results - Odd Order Tensor

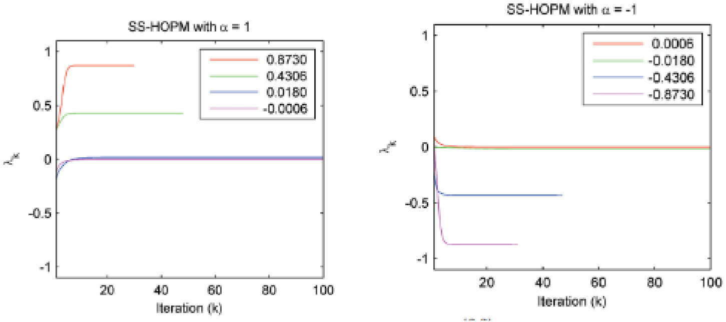


Figure 3: Using SSHOPM for Example 1 : $\alpha = 1$ (left) and $\alpha = -1$ (right)

Comparison with the SHOPM :

# Occurrences	λ	$\mathbf{x}$	Median its.
62	0.8730	$[\,-0.3922 \quad 0.7249 \quad 0.5664\,]$	19
38	0.4306	$[\,-0.7187 \; -0.1245 \; -0.6840\,]$	184

# Occurrences	λ	$\mathbf{x}$	Median its.
40	0.8730	$[\,-0.3922 \quad 0.7249 \quad 0.5664\,]$	32
29	0.4306	$[\,-0.7187 \; -0.1245 \; -0.6840\,]$	48
18	0.0180	$[\,0.7132 \quad 0.5093 \; -0.4817\,]$	116
13	-0.0006	$[\,-0.2907 \; -0.7359 \quad 0.6115\,]$	145

# Occurrences	λ	$\mathbf{x}$	Median its.
19	0.0006	$[\,0.2907 \quad 0.7359 \; -0.6115\,]$	146
18	-0.0180	$[\,-0.7132 \; -0.5093 \quad 0.4817\,]$	117
29	-0.4306	$[\,0.7187 \quad 0.1245 \quad 0.6840\,]$	49
34	-0.8730	$[\,0.3922 \; -0.7249 \; -0.5664\,]$	33

Figure 4: Eigenvalue occurrences for Example 1: SHOPM (top), SSHOPM with $\alpha = 1$ (middle), SSHOPM with $\alpha = -1$ (bottom)

Numerical Results - Even Order Tensor

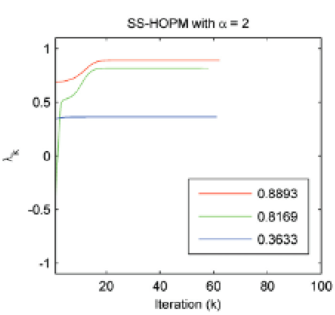


Figure 5: Using SSHOPM for Example 2 : $\alpha = 2$ (left) and $\alpha = -2$ (right)

# Occurrences	λ	x	Median its.
46	0.8893	[0.6672 0.2471 −0.7027]	63
24	0.8169	[0.8412 −0.2635 0.4722]	52
30	0.3633	[0.2676 0.6447 0.7160]	65

# Occurrences	λ	x	Median its.
15	−0.0451	[−0.7797 −0.6135 −0.1250]	35
40	−0.5629	[−0.1762 0.1796 −0.9678]	23
45	−1.0954	[−0.5915 0.7467 0.3043]	23

Figure 6: Eigenvalue occurrences for Example 2: SSHOPM with $\alpha = 2$ (top), SSHOPM with $\alpha = -2$ (bottom)

Simple Extrapolation Method for Matrices

- This is a variant of the power method.
- The method accelerates convergence to the dominant eigenpair.
- Particularly useful for problems with a small spectral gap.
- Basic Idea - Using a linear combination of the last two iterations to approximate the next step. The method uses a ratio of the previous two residuals for this.

(Nigam and Pollock, 2021)

Simple Extrapolation for Matrices

The algorithm for the matrix case:

Choose u_0 and set $h_0 = ||u_0||$
for $k = 0, ..., m-1 \geq 1$ **do**
  Set $x_k = h_k^{-1} u_k, u_{k+1} = v_{k+1} = Ax_k, h_{k+1} = ||u_{k+1}||$
  Set $\lambda_k = (u_{k+1}, x_k)$ and $d_{k+1} = u_{k+1} - \lambda_k x_k$
end for
for k = m, m+1, ... **do**
  Set $x_k = h_k^{-1} u_k,\ v_{k+1} = Ax_k$
  Compute $\gamma_k = -||d_k||/||d_{k-1}||$
  Set $u_{k+1} = (1-\gamma_k)v_{k+1} + \gamma_k v_k,\ h_{k+1} = ||u_{k+1}||$
  $x_k^\gamma = (1-\gamma_k)x_k + \gamma_k x_{k-1}$
  $\lambda_k = (u_{k+1}, x_k^\gamma)/(x_k^\gamma, x_k^\gamma)$
  $d_{k+1} = u_{k+1} - \lambda_k x_k^\gamma$
  STOP if $||d_{k+1}|| <$ tolerance
end for

S-HOPM + Simple Extrapolation - Even Order

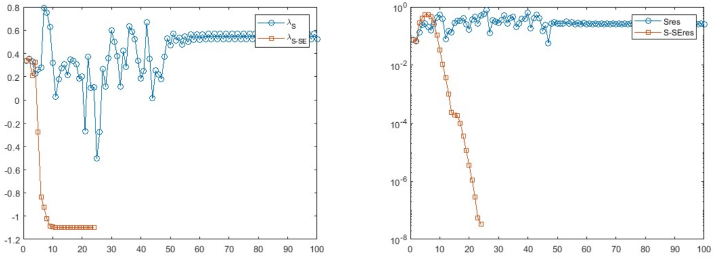


Figure 7: Using the Simple Extrapolation with SHOPM, for even tensors : lambda values (left) and residual (right)

Using the extrapolation gives us convergence.

S-HOPM + Simple Extrapolation - Odd Order

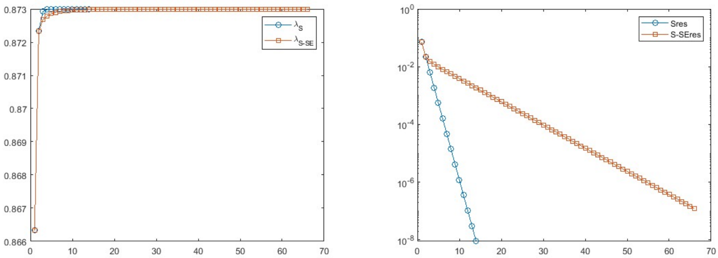


Figure 8: Using the Simple Extrapolation with SHOPM, for odd tensors : lambda values (left) and residual (right)

The acceleration does not work here!

SS-HOPM + Simple Extrapolation - Even Order

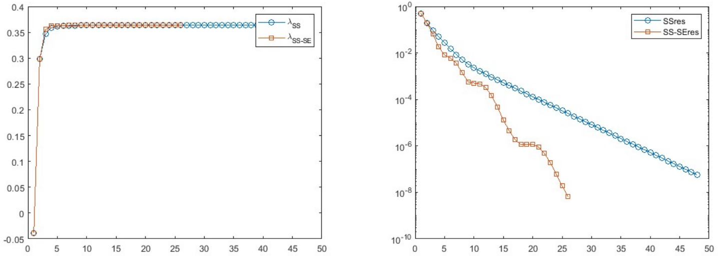
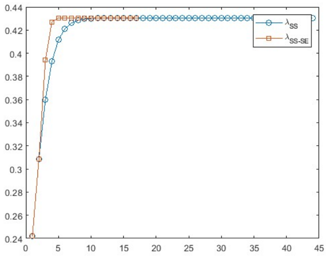


Figure 9: Using the Simple Extrapolation with SS-HOPM, for even tensors : lambda values (left) and residual (right)

SS-HOPM + Simple Extrapolation - Odd Order



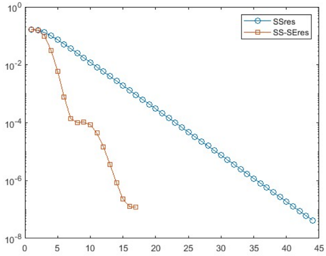


Figure 10: Using the Simple Extrapolation with SS-HOPM, for odd tensors : lambda values (left) and residual (right)

The proposed acceleration works well with the SS-HOPM.

Current Work

Numerically, we have seen that the method works well. But we need convergence results. More importantly, we need acceleration.

- Key - The spectral radius of the Jacobian.

References

[1] B. W. Bader, T. G. Kolda, et. al., *MATLAB Tensor Toolbox version* 3.4, www.tensortoolbox.org, September 21, 2022.

[2] S. Cipolla, M. Redivo-Zaglia, and F. Tudisco, *Shifted and extrapolated power methods for tensor ℓ^p-Eigenpairs*. ETNA - Electronic Transactions on Numerical Analysis, 53 (2020), 1–27. https://doi.org/10.1553/etna_vol53s1

[3] E. Kofidis and P. A. Regalia, *On the best rank 1 approximation of higher order supersymmetric tensors*, SIAM Journal on Matrix Analysis and Applications, 23 (2002), pp. 863-884.

[4] T. G. Kolda, and J. R. Mayo, *An adaptive shifted power method for computing generalized tensor eigenpairs*. SIAM Journal on Matrix Analysis and Applications, 35(4) (2011), pp. 1563-1581.

[5] T. G. Kolda, and J. R. Mayo, *Shifted power method for computing tensor eigenpairs*. SIAM Journal on Matrix Analysis and Applications, 32(4) (2011), pp. 1095–1124.

References (contd.)

[6] L.-H. Lim, *Singular values and eigenvalues of tensors: A variational approach*, in CAMSAP'05: Proceeding of the IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing, 2005, pp. 129-132.

[7] N. Nigam, and S. Pollock, *A simple extrapolation method for clustered eigenvalues*. Numerical Algorithms, 89(1) (2021), pp. 115–143. https://doi.org/10.1007/s11075-021-01108-7

[8] L. Qi, *Eigenvalues of a real supersymmetric tensor*, J. Symbolic Comput., 40 (2005), pp. 1302-1324.

[9] L. Qi, *Eigenvalues and invariants of tensors*, J. Math. Anal. Appl., 325 (2007), pp. 1363-1377.

The End

Thank you!!