A generalized global Arnoldi method based on tensor format for ill-posed tensor equations

Fatemeh Panjeh Ali Beik
(joint with Mehdi Najafi)

Department of Mathematics
Vali-e-Asr University of Rafsanjan
email: f.beik@vru.ac.ir

The third conference on Numerical Analysis and Scientific Computation with Applications (NASCA 2018)
July 2nd–6th, 2018
Outline of the presentation:

1. Preliminaries
   - Basic concepts
   - Introducing the main problem

2. Conditioning
   - An overview on an existing result in the literature
   - A lower bound for the condition number
   - An upper bound for the condition number

3. Tikhonov regularization

4. Numerical experiments
   - Notes
   - Example 1
   - Example 2
   - Example 3
What is called a tensor?

In the literature, a multidimensional array is called a tensor. An $N$-way or $N$th-order tensor is an element of the tensor product of $N$ vector spaces.

**Notations**

Vectors (tensors of order one) are represented by lowercase letter, matrices (tensors of order two) are denoted by capital letters and Higher-order tensors (order three or higher) are indicated by Euler script letters.

Representation of a third order tensor

A third-order tensor: $\mathcal{X} \in \mathbb{R}^{I \times J \times K}$

Frontal slices: $X::_k$ (or $X_k$)
Inner product and norm of a tensor

\[ \mathbf{X} = [x_{i_1 i_2 \ldots i_N}] \in \mathbb{R}^{l_1 \times l_2 \times \cdots \times l_N} \quad \text{and} \quad \mathbf{Y} = [y_{i_1 i_2 \ldots i_N}] \in \mathbb{R}^{l_1 \times l_2 \times \cdots \times l_N} \]

The inner product of two same-sized tensors \( \mathbf{X}, \mathbf{Y} \in \mathbb{R}^{l_1 \times l_2 \times \cdots \times l_N} \):

\[ \langle \mathbf{X}, \mathbf{Y} \rangle = \sum_{i_1=1}^{l_1} \sum_{i_2=1}^{l_2} \cdots \sum_{i_N=1}^{l_N} x_{i_1 i_2 \ldots i_N} y_{i_1 i_2 \ldots i_N} . \]

The norm of a tensor:

\[ \| \mathbf{X} \|^2 = \sum_{i_1=1}^{l_1} \sum_{i_2=1}^{l_2} \cdots \sum_{i_N=1}^{l_N} x_{i_1 i_2 \ldots i_N}^2 . \]
The $n$-mode product

The $n$-mode product $\mathcal{X} \times_n U$ of a tensor $\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_{n-1} \times I_n \times I_{n+1} \cdots \times I_N}$ and a matrix $U \in \mathbb{R}^{J \times I_n}$

$$(\mathcal{X} \times_n U)_{i_1 i_2 \cdots i_{n-1} j i_{n+1} \cdots i_N} = \sum_{i_n=1}^{I_n} x_{i_1 i_2 \cdots i_{n-1} i_n i_{n+1} \cdots i_N} u_{ji_n}$$

is a tensor of order $I_1 \times I_2 \times \cdots \times I_{n-1} \times J \times I_{n+1} \times \cdots \times I_N$
Consider the following Sylvester tensor equation

$$\mathcal{X} \times_1 A^{(1)} + \mathcal{X} \times_2 A^{(2)} + \cdots + \mathcal{X} \times_N A^{(N)} = \mathcal{D},$$

(1)

where the matrices $A^{(n)} \in \mathbb{R}^{I_n \times I_n}$ ($n = 1, 2, \ldots, N$) and the right-hand-side tensor $\mathcal{D} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}$ are known and $\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}$ is the unknown tensor to be determined.

The global schemes for some of the well-known iterative methods are developed in their tensor forms for solving (1) where the coefficient matrices $A^{(i)}$'s are positive definite matrices. Here, the matrices $A^{(1)}, \ldots, A^{(N)}$ can be extremely ill-conditioned.

Assume that the right-hand-side tensor $\mathcal{D}$ in (1) contains an error $\mathcal{E}$ which is called “noise”.
Assume that the right-hand-side tensor $\mathcal{D}$ in (1) contains an error $\mathcal{E}$ which is called “noise”.

For $\mathcal{D} = \hat{\mathcal{D}} + \mathcal{E}$, then the tensor $\hat{\mathcal{D}}$ denotes the unknown noise-free unavailable right-hand-side.
Assume that the right-hand-side tensor $\mathcal{D}$ in (1) contains an error $\mathcal{E}$ which is called “noise”.

For $\mathcal{D} = \hat{\mathcal{D}} + \mathcal{E}$, then the tensor $\hat{\mathcal{D}}$ denotes the unknown noise-free unavailable right-hand-side.

Let $\hat{\mathcal{X}}$ denote the solution of minimum norm of the tensor equation with the error-free right-hand-side,

$$\hat{\mathcal{X}} \times_1 A^{(1)} + \hat{\mathcal{X}} \times_2 A^{(2)} + \cdots + \hat{\mathcal{X}} \times_N A^{(N)} = \hat{\mathcal{D}}, \quad (2)$$

for simplicity, Eq. (2) is assumed to be consistent.
Toward sensitivity analysis of (1)

Note

We recall that (1) is equivalent to the following linear system of equations $Ax = b$ with $x = \text{vec}(X)$, $b = \text{vec}(D)$ and

$$
\mathcal{A} = \sum_{j=1}^{N} I(I_N) \otimes \ldots \otimes I(I_{j+1}) \otimes A^{(j)} \otimes I(I_{j-1}) \otimes \ldots \otimes I(I_1),
$$

where $\otimes$ denotes the Kronecker product, $I^{(n)}$ denotes the identity matrix of order $n$ and for a given matrix (or a tensor), the operator “vec” stands for the standard vectorization operator which transforms a tensor to vector [Kolda and Bader, 2009].
Assume that the matrices $A^{(i)}$ for $i = 1, 2, \ldots, N$ are diagonalizable, i.e., there exists nonsingular matrices $T_i$ and diagonal matrices $\Lambda_i$ such that $T_i^{-1}A^{(i)}T_i = \Lambda_i$ for $i = 1, 2, \ldots, N$. Then the following inequality holds:

$$\frac{\|\hat{X} - X\|}{\|\hat{X}\|} \leq \sum_{i=1}^{N} \left\|A^{(i)}\right\|_F \frac{\prod_{i=1}^{N} \text{cond}(T_i)}{\min_{\lambda_i \in \sigma(A^{(i)})} \left| \sum_{i=1}^{N} \lambda_i \right|} \left\|\hat{D} - D\right\|$$

An upper bound for $\text{cond}(A)$

[Linear Multilinear Algebra (2018)]

For the case that $A^{(1)} = A^{(2)} = \cdots = A^{(N)} = A$ such that $A$ is positive stable, then some results are given in the following paper in which the derive expression contains $\|A^{-1}\|_2$.

Toward a lower bound for $\text{cond}(A)$

**Proposition**

Let the matrix $A$ be invertible. Then

$$\frac{1}{\|A^{-1}\|_2^2} = \lambda_{\min}(AA^T) \leq \left( \sum_{i=1}^{N} \sigma_{\min}(A^{(i)}) \right)^2,$$

$$\lambda_{\max}(AA^T) \geq \sum_{i=1}^{N} \sigma_{\max}^2(A^{(i)}) + 2 \sum_{i=1}^{N} \sum_{j=i+1}^{N} \left( y_i^T \mathcal{H}(A^{(i)}) y_i \right) \left( y_j^T \mathcal{H}(A^{(j)}) y_j \right),$$

where $\mathcal{H}(A^{(i)}) = \frac{1}{2}(A^{(i)} + (A^{(i)})^T)$, $A^{(i)}(A^{(i)})^T y_i = \sigma_{\max}^2(A^{(i)}) y_i$ and $\|y_i\|_2 = 1$ for $i = 1, 2, \ldots, N$. 
A Lower bound for $\text{cond}(A)$

In some cases Proposition 2.1 can provide a (useful) lower bound for the $\text{cond}(A)$ in terms of the singular values of the matrices $A^{(1)}, A^{(2)}, \ldots, A^{(N)}$. For instance,

- If the matrices are all equal, i.e., $A^{(i)} = A$ for $i = 1, 2, \ldots, n$, then
  \[
  \text{cond}(A) \geq \frac{\sigma_{\text{max}}(A)}{\sqrt{N} \sigma_{\text{min}}(A)};
  \]

- If the matrices are all (nonsymmetric) positive definite then
  \[
  \text{cond}(A) \geq \left( \frac{N}{\sum_{i=1}^{N} \sigma_{\text{max}}^{2}(A^{(i)})} \right)^{1/2} \frac{1}{\sqrt{N} \sum_{i=1}^{N} \sigma_{\text{min}}(A^{(i)})}.
  \]
Toward an upper bound for \( \text{cond}(\mathcal{A}) \)

**Remark (Zak and Toutounian (2013))**

Let \( F \) and \( G \) be two nonsymmetric matrices. Let \( H_F \) and \( S_F \) be the symmetric and skew-symmetric parts of \( F \) and similarly for \( H_G \) and \( S_G \). If

\[
\lambda_{\min}(H_G)\lambda_{\min}(H_F) + \min(-\lambda(S_G)\lambda(S_F)) > 0,
\]

then the symmetric part of \( G^T \otimes F \) is a positive definite matrix.

---

For simplicity, let us assume that $N = 3$ and 

$$A^{(1)} = A, \quad A^{(2)} = B \quad \text{and} \quad A^{(3)} = C.$$ 

**Proposition**

*Let the matrix $A$ be invertible. Then*

$$\lambda_{\max}(AA^T) \leq (\sigma_{\max}(A) + \sigma_{\max}(B) + \sigma_{\max}(C))^2, \quad (7)$$

*and if $B^T \otimes A$, $C^T \otimes A$ and $C^T \otimes B$ are positive definite or if inequality (6) is true for each pair of matrices $A, B$ and $C$, then*

$$\frac{1}{\lambda_{\min}(AA^T)} \leq \frac{1}{\sigma_{\min}^2(A) + \sigma_{\min}^2(B) + \sigma_{\min}^2(C)}. \quad (8)$$
An upper bound for $\text{cond}(A)$

**Remark**

Let all matrices be (nonsymmetric) positive definite and inequality (6) be true for each pair of matrices $A, B$ and $C$, then

$$\text{cond}(A) \leq \frac{\sigma_{\max}(A) + \sigma_{\max}(B) + \sigma_{\max}(C)}{\sqrt{\sigma_{\min}^2(A) + \sigma_{\min}^2(B) + \sigma_{\min}^2(C)}}.$$
A brief overview


- In [Appl. Numer. Math. 62 (2012), 1215–1228], Reichel et al. developed the above algorithm for solving the following problem

\[
\min_{x \in \mathbb{R}^n} \left\{ \|Ax - b\|_2^2 + \mu \|Lx\|_2^2 \right\}.
\]

- In [J. Comput. Appl. Math. 236 (2012) 2078–2089], the work by Reichel et al. has been extended (by Bouhamidi et al.) for solving the following problem

\[
\min_{x \in \mathbb{R}^{n \times n}} \left\{ \left\| \sum_{i=1}^{p} A_i X B_i - G \right\|_F^2 + \mu \left\| \sum_{j=1}^{q} L_j X L_j' \right\|_F^2 \right\}.
\]
Using the results in [Beik et al., 2016] and some computations, we may also use the generalized global Arnoldi process based on tensor format (called GGAT\_BTF method) to solve

\[
\min_{\mathcal{X} \in \mathbb{R}^{l_1 \times l_2 \times \cdots \times l_N}} \left\{ \left\| \sum_{i=1}^{N} \mathcal{X} \times_i A^{(i)} - \mathcal{D} \right\|^2 + \lambda \left\| \sum_{i=1}^{\tilde{N}} \mathcal{X} \times_i L^{(i)} \right\|^2 \right\},
\]

where \( L^{(i)} \) are given regularization matrices \( i = 1, \ldots, \tilde{N} \) and the positive constant \( \lambda \) is a regularization parameter.
Using the results in [Beik et al., 2016] and some computations, we may also use the generalized global Arnoldi process based on tensor format (called GGAT–BTF method) to solve

\[
\min_{X \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}} \left\{ \left\| \sum_{i=1}^{N} X \times_i A^{(i)} - D \right\|_2^2 + \lambda \left\| \sum_{i=1}^{\tilde{N}} X \times_i L^{(i)} \right\|_2^2 \right\},
\]

(9)

where \( L^{(i)} \) are given regularization matrices \( i = 1, \ldots, \tilde{N} \) and the positive constant \( \lambda \) is a regularization parameter.
Again, consider the following problem

\[
\min_{X \in \mathbb{R}^{n \times n}} \left\{ \left\| \sum_{i=1}^{p} A_i X B_i - G \right\|_F^2 + \mu \left\| \sum_{j=1}^{q} L_j X L_j' \right\|_F^2 \right\}.
\]
Again, consider the following problem

$$\min_{x \in \mathbb{R}^{n \times n}} \left\{ \left\| \sum_{i=1}^{p} A_i X B_i - G \right\|_F^2 + \mu \left\| \sum_{j=1}^{q} L_j X L'_j \right\|_F^2 \right\}.$$ 


In 2018, Huang et al. have proved that there exists a matrix $\tilde{L}_k \in \mathbb{R}^{k \times k}$ such that

$$\left\| \sum_{j=1}^{q} L_j X_k L'_j \right\|_F^2 = \left\| \tilde{L}_k^T y_k \right\|_2^2$$

where $X_k = \sum_{i=1}^{k} y_{k}^{(i)} V_i$ where $V_1, V_2, \ldots, V_k$ results after applying $k$ steps of the Arnoldi process for $\mathcal{K}_k(A_1, G)$ where $A_1(X) = \sum_{i=1}^{p} A_i X B_i$. 
To see how the matrix $\tilde{L}_k$ is obtained, first we need to define

$$M_i = \sum_{j=1}^{q} L_j^{(1)} V_i L_j^{(2)},$$

and

$$N_k = [n_{ij}] \quad \text{and} \quad n_{ij} = \text{trace}(M_i^T M_j).$$

- If $N_k$ is nonsingular, then $N_k = \tilde{L}_k \tilde{L}_k^T$.
- If $N_k$ is singular, spectral factorization is used. In fact, in this case we have $\tilde{L}_k^T = D_k^{1/2} Q_k^T$ where $N_k = Q_k D_k Q_k^T$ and $Q_k \in \mathbb{R}^{k \times k}$ is an orthogonal matrix.

**Note**

In the suggested strategy to solve the reduced (regularized) least squares problem, the inverse of $\tilde{L}_k$ is used...
It can be shown that $\tilde{L}_k$ is invertible,

- If the following matrix is invertible

$$
\sum_{j=1}^{q} \left(L_j^{(2)}\right)^T \otimes L_j^{(1)}.
$$

- Or if the following condition holds

$$
\mathcal{N} \left( \sum_{j=1}^{q} \left(L_j^{(2)}\right)^T \otimes L_j^{(1)} \right) \cap \text{span} \{ V_1, ..., V_k \} = \{0\}
$$

Here $\mathcal{N}(L)$ stands for the null space of a given matrix $L$. Note that

$$
\mathcal{K}_k(A_1, G) \equiv \text{span} \{ V_1, ..., V_k \}$$
In this section, we report some numerical results for solving (1) with the right-hand-side tensor $D$ contaminated by an error tensor $E$.

All of the results were performed using MATLAB R2014a with an Intel Core i7-4770K CPU @ 3.50GHz processor and 24GB RAM.

The noise-tensor $E$ has normally distributed random entries with zero mean and with variance chosen so that $\|E\|/\|\hat{D}\| = \ell$.

We comment that the Tensor Toolbox was utilized for solving the mentioned test problems, see the following reference for more details:

Preliminaries
Conditioning
Tikhonov regularization
Numerical experiments

Notes
Example 1
Example 2
Example 3


Stopping criterion

\[
\frac{\| X_{\mu_k,k} - X_{\mu_{k-1},k-1} \|}{\| X_{\mu_{k-1},k-1} \|} \leq \tau,
\]
Stopping criterion

\[
\frac{\| X_{\mu k, k} - X_{\mu k-1, k-1} \|}{\| X_{\mu k-1, k-1} \|} \leq \tau,
\]

Relative error

\[
e_k =: \frac{\| X_{\mu k, k} - \hat{X} \|}{\| \hat{X} \|}
\]
Stopping criterion

\[ \frac{\|X_{\mu_k,k} - X_{\mu_{k-1},k-1}\|}{\|X_{\mu_{k-1},k-1}\|} \leq \tau, \]

\( \tau \) is given separately in each problem.

Relative error

\[ e_k =: \frac{\|X_{\mu_k,k} - \hat{X}\|}{\|\hat{X}\|} \]

\( \hat{X} \) is the exact solution and \( X_{\mu_k,k} \) is the \( k \)th computed approximation.
In the implementations of the Chebyshev collocation spectral method (CSM) for solving 3D transient coupled radiative and conductive heat transfer problems (see the following paper for more details), at each time step, it is required to solve two Sylvester tensor equations of the following form

\[ \mathcal{G} \times_1 A + \mathcal{G} \times_2 A + \mathcal{G} \times_3 A = \mathcal{D}. \]  

In this example, we use the zero tensor as the initial guess and the regularization matrices are chosen to be bidiagonal and equal to the one given in [1, Eq. (35)],

\[
L = \begin{bmatrix}
0 & 0 & 0 & 0 \\
-1 & 2 & -1 & 0 \\
-1 & 2 & -1 & 0 \\
& & & \\
& & & \\
& & & \\
& & & \\
-1 & 2 & -1 & 0 \\
0 & 0 & 0 & 0
\end{bmatrix}
\]

Results for Example 1.

\[ \tau = 5e - 4 \]

<table>
<thead>
<tr>
<th>( \ell )</th>
<th>Grid</th>
<th>Method</th>
<th>Iter ((k))</th>
<th>(e(k))</th>
<th>CPU-times(sec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.01</td>
<td>55 \times 55 \times 55 [\text{cond}(A)=4.24 \times 10^{16}]</td>
<td>GAT_BTF</td>
<td>13</td>
<td>(2.08 \times 10^{-1})</td>
<td>2.57</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GGAT_BTF</td>
<td>21</td>
<td>(2.81 \times 10^{-2})</td>
<td>6.76</td>
</tr>
<tr>
<td>0.001</td>
<td>55 \times 55 \times 55 [\text{cond}(A)=4.24 \times 10^{16}]</td>
<td>GAT_BTF</td>
<td>12</td>
<td>(6.60 \times 10^{-2})</td>
<td>2.15</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GGAT_BTF</td>
<td>15</td>
<td>(2.67 \times 10^{-2})</td>
<td>3.73</td>
</tr>
<tr>
<td>0.001</td>
<td>75 \times 75 \times 75 [\text{cond}(A)=2.34 \times 10^{16}]</td>
<td>GAT_BTF</td>
<td>8</td>
<td>(2.86 \times 10^{-1})</td>
<td>2.40</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GGAT_BTF</td>
<td>15</td>
<td>(3.14 \times 10^{-2})</td>
<td>7.66</td>
</tr>
<tr>
<td>0.001</td>
<td>110 \times 110 \times 110 [\text{cond}(A)=8.01 \times 10^{16}]</td>
<td>GAT_BTF</td>
<td>10</td>
<td>(3.92 \times 10^{-1})</td>
<td>4.14</td>
</tr>
<tr>
<td></td>
<td></td>
<td>GGAT_BTF</td>
<td>18</td>
<td>(2.01 \times 10^{-2})</td>
<td>14.60</td>
</tr>
</tbody>
</table>
In the second example, we set $\tau = 5 e - 4$

- the $512 \times 512 \times 3$ tensor $\hat{X}$ corresponding to “Sailboat on lake” image is used as an exact (noise-free) image which is available at http://sipi.usc.edu/database.

- We consider Eq. (1) such that $A$ and $B$ are respectively Gaussian Toeplitz and uniform Toeplitz matrices of order 512, (in which $r = 7$ and $\sigma = 2$).

- Regularization matrices

$$L_1 = L_2 = \begin{bmatrix} 2 & -1 & \cdot & \cdot & \cdot & -1 \\ -1 & 2 & -1 \\ \cdot & \cdot & \cdot \\ -1 & 2 & -1 \\ -1 & 2 \end{bmatrix}, \quad L_3 = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix},$$
Figure: Example 2.; Noise level $\ell = 0.01$
Results for Example 2.

\[ \text{cond}(A) = 1.77e + 06 \quad \text{and} \quad \text{cond}(B) = 4.92e + 16 \]

<table>
<thead>
<tr>
<th>$\ell$</th>
<th>Method</th>
<th>Iter($k$)</th>
<th>$e(k)$</th>
<th>CPU-times(sec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.01</td>
<td>GAT_BTF</td>
<td>11</td>
<td>0.0741</td>
<td>8.60</td>
</tr>
<tr>
<td></td>
<td>GGAT_BTF</td>
<td>5</td>
<td>0.1086</td>
<td>1.53</td>
</tr>
<tr>
<td>0.001</td>
<td>GAT_BTF</td>
<td>13</td>
<td>0.0722</td>
<td>12.23</td>
</tr>
<tr>
<td></td>
<td>GGAT_BTF</td>
<td>5</td>
<td>0.1085</td>
<td>1.54</td>
</tr>
</tbody>
</table>
In the third example,

- the solution \( \hat{X} \) is a \( 1017 \times 1340 \times 33 \) tensor associated with a hyperspectral image, see the following reference:


- We consider Eq. (1) such that \( A, B \) and \( C \) are respectively the Gaussian Toeplitz matrices of order 1017, 1340 and 33 with \( r = 5 \) and \( \sigma = 45 \) for \( A \) (and \( r = 7 \) and \( \sigma = 2 \)). Here, \( \tau = 8e - 4 \) and

\[
L^{(1)}, L^{(2)}, L^{(3)} := \begin{bmatrix}
2 & -1 & & \\
-1 & 2 & -1 & \\
& \ddots & \ddots & \ddots \\
& & -1 & 2 & -1 \\
& & & -1 & 2
\end{bmatrix},
\]

\( \tau = 8e - 4 \) and
Results for Example 3.

\[
\begin{align*}
\text{cond}(A) &= 7.02e + 12 \\
\text{cond}(B) &= 3.55e + 06 \\
\text{cond}(C) &= 1.74e + 04
\end{align*}
\]

<table>
<thead>
<tr>
<th>( \ell )</th>
<th>Method</th>
<th>Iter ((k))</th>
<th>(e(k))</th>
<th>CPU-times (sec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.01</td>
<td>GAT_BTF</td>
<td>7</td>
<td>0.0231</td>
<td>248.91</td>
</tr>
<tr>
<td></td>
<td>GGAT_BTF</td>
<td>5</td>
<td>0.5860</td>
<td>142.64</td>
</tr>
<tr>
<td>0.001</td>
<td>GAT_BTF</td>
<td>10</td>
<td>0.0477</td>
<td>514.78</td>
</tr>
<tr>
<td></td>
<td>GGAT_BTF</td>
<td>5</td>
<td>0.5870</td>
<td>143.22</td>
</tr>
</tbody>
</table>
Figure: Exact image (left) and noisy image (right).

Figure: Restored image by: GGAT\_BTF (left) and GAT\_BTF (right) for noise level = 0.001.
Thanks for your attention!
Algorithm 1 The global Arnoldi method based on tensor form

1: Let $\mathcal{V}_1 = D/\|D\|_F$
2: for $j = 1, \ldots, k$ do
3: \hspace{1em} $\mathcal{V} = M(\mathcal{V}_j)$;
4: \hspace{1em} for $i = 1, \ldots, j$ do
5: \hspace{2em} $h_{i,j} = \langle \mathcal{V}, \mathcal{V}_i \rangle_F$;
6: \hspace{2em} $\mathcal{V} = \mathcal{V} - h_{i,j} \mathcal{V}_i$;
7: \hspace{1em} end for
8: \hspace{1em} $h_{j+1,j} = \| \mathcal{V} \|_F$;
9: \hspace{1em} if $h_{j+1,j} > 0$ then
10: \hspace{2em} $\mathcal{V}_{j+1,j} = \mathcal{V} / h_{j+1,j}$;
11: \hspace{1em} else
12: \hspace{2em} exit;
13: \hspace{1em} end if
14: end for
Algorithm 2 Global Arnoldi-Tikhonov based on tensor form

**Input:** $A^{(i)}, i = 1, \ldots , p$; $L^{(j)}, j = 1, \ldots , q$; tensor $\mathcal{D}$ and $\varepsilon, \eta > 1$;

**Output:** Approximation solution $\mathcal{X}_k$;

1. for $k = 1, 2, \ldots$ until convergence do
2. Construct $\mathcal{Y}_k$ with the column tensors $\mathcal{Y}_i$’s and $\mathcal{H}_k$ by Algorithm 1;
3. Compute $M_i = \sum_{j=1}^{q} \mathcal{Y}_i \times_j L^{(j)}$ ($i = 1, 2, \ldots , k$);
4. Compute $\mathcal{N}_k = [n_{i,j}] \in \mathbb{R}^{k \times k}$, $n_{i,j} = \langle M_i, M_j \rangle_F$;
5. Compute Choleski factorization $\mathcal{N}_k = \tilde{L}_k \tilde{L}_k^T$;
6. Compute $\hat{H}_k = \mathcal{H}_k \tilde{L}_k^T$;
7. Compute the zero $\nu$ of $\phi(\nu) = \|\mathcal{D}\|_F c_1 - \hat{H}_k z_{1/\nu,k}\|_2^2$ by the discrepancy principle;
8. Define the regularization parameter $\lambda = 1/\nu$;
9. Compute $z_{\lambda,k}$ by

$$
\min_{z \in \mathbb{R}^k} \left\| \begin{bmatrix} \hat{H}_k \\ \lambda^{1/2} I_k \end{bmatrix} z - \begin{bmatrix} \|\mathcal{D}\|_F \\ 0 \end{bmatrix} \right\|_2^2,
$$

and let $y_k = \tilde{L}_k^T z_{\lambda,k}$;
10. Compute $\mathcal{X} = \sum_{i=1}^{k} y_k^{(i)} \mathcal{Y}_i = \mathcal{Y}_k \times_{(N+1)} y_k$, where $y_k = (y_k^{(1)}, \ldots , y_k^{(k)})^T$;
11. end for
Algorithm 3 GGAT\textsubscript{\textit{BTF}}(k) process.

\textbf{Input: } Choose an integer $k > 0$ and the initial tensor $\mathcal{V}_1$ such that $\|\mathcal{V}_1\| = 1$;

\textbf{Output: } approximation solution $\mathcal{X}_k$;

1: Set $N = 1$;
2: for $j = 1, \ldots, k$ do
3: \hspace{1em} if $j > N$ then
4: \hspace{2em} exit;
5: \hspace{1em} end if
6: \hspace{1em} $\mathcal{W}_j := \mathcal{M}(\mathcal{V}_j)$
7: \hspace{1em} for $i = 1, \ldots, N$ do
8: \hspace{2em} $\mathcal{H}_\mathcal{M}(i, j) := \langle \mathcal{W}, \mathcal{V}_i \rangle$
9: \hspace{2em} $\mathcal{W} = \mathcal{W} - \mathcal{H}_\mathcal{M}(i, j) \mathcal{V}_i$
10: \hspace{1em} end for
11: \hspace{1em} $\mathcal{H}_\mathcal{M}(N + 1, j) := \|\mathcal{W}\|
12: \hspace{1em} if $\mathcal{H}_\mathcal{M}(N + 1, j) > 0$ then
13: \hspace{2em} $N = N + 1$;
14: \hspace{2em} $\mathcal{V}_N = \mathcal{W} / \mathcal{H}_\mathcal{M}(N, j)$
15: \hspace{1em} else
16: \hspace{2em} exit;
17: \hspace{1em} end if
18: \hspace{1em} $\mathcal{W}_j := \mathcal{L}(\mathcal{V}_j)$
19: \hspace{1em} for $i = 1, \ldots, N$ do
20: \hspace{2em} $\mathcal{H}_\mathcal{L}(i, j) := \langle \mathcal{W}, \mathcal{V}_i \rangle$
21: \hspace{2em} $\mathcal{W} = \mathcal{W} - \mathcal{H}_\mathcal{L}(i, j) \mathcal{V}_i$
22: \hspace{1em} end for
23: \hspace{1em} $\mathcal{H}_\mathcal{L}(N + 1, j) := \|\mathcal{W}\|
24: \hspace{1em} if $\mathcal{H}_\mathcal{L}(N + 1, j) > 0$ then
25: \hspace{2em} $N = N + 1$;
26: \hspace{2em} $\mathcal{V}_N = \mathcal{W} / \mathcal{H}_\mathcal{L}(N, j)$
27: \hspace{1em} else
28: \hspace{2em} exit;
29: \hspace{1em} end if
30: end for
Uniform Toeplitz

The Uniform Toeplitz matrix $A = [a_{ij}]$ is defined by

$$a_{ij} = \begin{cases} 
\frac{1}{2r-1}, & |i - j| \leq r, \\
0, & otherwise,
\end{cases}$$

Gaussian Toeplitz matrix

The Gaussian Toeplitz matrix $A = [a_{ij}]$ is defined by

$$a_{ij} = \begin{cases} 
\frac{1}{\sigma \sqrt{2\pi}} \exp\left(-\frac{(i-j)^2}{2\sigma^2}\right), & |i - j| \leq r, \\
0, & otherwise,
\end{cases}$$