Is Lanczos tridiagonalization algorithm essential for solving large eigenvalue problems?

By

Achiya Dax

Hydrological Service
Jerusalem, Israel

e-mail: dax20@water.gov.il
In this talk we propose an alternative approach:

**A new type of restated Krylov method**

The new method

Avoids the Lanczos algorithm

Avoids polynomial filtering

It is neither “explicit restart” nor “implicit restart”
Plan:

Part 1: The new method

Part 2: Applications

Low-rank approximations of large sparse matrices.
Computing small eigenvalues or small singular values via inexact inversions.
Notations

$G$ a large sparse symmetric $n \times n$ matrix, $G^T = G$,

with eigenvalues $\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_n$

and eigenvectors $v_1, v_2, ..., v_n$

\[ Gv_j = \lambda_j v_j, \quad j = 1, ..., n. \quad GV = VD \]

\[ V = [v_1, v_2, ..., v_n], \quad V^TV = V V^T = I \]

\[ D = \text{diag} \{ \lambda_1, \lambda_2, ..., \lambda_n \} \]

\[ G = VDV^T = \Sigma \lambda_j v_j v_j^T \]
**Aim:** Computing a small number, $k$, of exterior eigenpairs.

For example, $k$ eigenvalues with the largest moduli:

$$|\lambda_1| \geq |\lambda_2| \geq ... \geq |\lambda_k| \geq ... \geq |\lambda_{k+\ell}| \geq ... \geq |\lambda_n|$$

$v_1, v_2, \ldots, v_k, \ldots, v_{k+\ell}, \ldots, v_n$

$\ell$ is the length of the restarted Krylov sequences

In our experiments $\ell = k + 40$
The $q$th iteration, $q = 1, 2, \ldots$,

Starts with an $n \times (k + \ell)$ matrix, $X_{q-1} = [V_{q-1}, Y_{q-1}]$, that has orthonormal columns.

$V_{q-1}$ is an $n \times k$ matrix that contains the current Ritz vectors.

$Y_{q-1}$ is an $n \times \ell$ matrix that contains “new” information which is obtained from a Krylov matrix.
The qth iteration, \( q = 1, 2, \ldots \),

**Step 1**: Compute the new Ritz matrix \( V_q \).

**Step 2**: Compute a new Krylov information matrix \( B_q \).

**Step 3**: Obtain \( Z_q \) by orthogonalizing \( B_q \) against \( V_q \).

**Step 4**: Compute, \( Y_q \), an orthonormal basis of \( Z_q \).

**Step 5**: Define \( X_q = [V_q, Y_q] \)

The matrices \( V_q \), \( Y_q \) and \( X_q \) has orthonormal columns.
Step 1: Compute the Rayleigh quotient matrix

\[ S_q = X_{q-1}^T G X_{q-1} \]

and the \( k \) largest eigenvalues of \( S_q \)

\[ |\lambda_1^{(q)}| \geq |\lambda_2^{(q)}| \geq ... \geq |\lambda_k^{(q)}| \geq ... \geq |\lambda_\ell^{(q)}| \]

The related \( k \) eigenvectors are assembled into the \( \ell \times k \) matrix

\[ W_q = [w_1, w_2, ..., w_k] \]

and used to construct the new matrix of Ritz vectors

\[ V_q = X_{q-1} W_q \]

Since \( X_{q-1} \) and \( W_q \) have orthonormal columns, \( V_q \) inherits this property.
The monotonicity property

The eigenvalues of the matrix $V_q^T G V_q$

$$|\lambda_1^{(q)}| \geq |\lambda_2^{(q)}| \geq \ldots \geq |\lambda_k^{(q)}|$$

interlace those of the matrix $[V_q, Y_q]^T G [V_q, Y_q]$

$$|\lambda_1^{(q+1)}| \geq |\lambda_2^{(q+1)}| \geq \ldots \geq |\lambda_k^{(q+1)}|$$

and therefore

$$|\lambda_j| \geq |\lambda_j^{(q+1)}| \geq |\lambda_k^{(q)}|$$

for $j = 1, \ldots, k$ and $q = 1, 2, \ldots$.

This holds for any choice of $B_q$!
The basic Krylov matrix

$$B_q = [b_1, b_2, \ldots, b_\ell]$$

The columns span a Krylov subspace of $G$ that is generated by the following three term recurrence relation

For $j = 2, \ldots, k$

$$b_j = G \ b_{j-1}$$

Orthogonalize $b_j$ against $b_{j-1}$

Orthogonalize $b_j$ against $b_{j-2}$

Normalize $b_j$. 
The starting vector \( b_0 \)

is a unit vector in the direction

\[ V_q e = [v_1, v_2, ..., v_k] e = v_1 + v_2 + ... + v_k \]

where \( v_1 + v_2 + ... + v_k \) are the current Ritz vectors.

**Computing \( b_0 \):**

Set \( b_0 = V_q e \),

normalize \( b_0 \).

**Computing \( b_1 \):**

Set \( b_1 = G b_0 \)

orthogonalize \( b_1 \) against \( b_0 \),

normalize \( b_1 \).
Comparison with “Explicit Restart”

Step 1: Compute the new Ritz matrix $V_q$.
Step 2: Compute a new Krylov information matrix $B_q$.
Step 3: Obtain $Z_q$ by orthogonalizing $B_q$ against $V_q$.
Step 4: Compute, $Y_q$, an orthonormal basis of $Z_q$.
Step 5: Define $X_q = [V_q, Y_q]$.

Step 1: Compute a Ritz matrix $V_q$ from $T_{q-1}$.
Step 2: Compute a new tridiagonal matrix $T_q$.
(Starting from the direction of $V_q e$.)
Comparison with “Implicit Restart”

Step 1: Compute the new Ritz matrix $V_q$.

Step 2: Compute a new Krylov information matrix $B_q$.

Step 3: Obtain $Z_q$ by orthogonalizing $B_q$ against $V_q$.

Step 4: Compute, $Y_q$, an orthonormal basis of $Z_q$.

Step 5: Define $X_q = [V_q, Y_q]$.

Step 1: Compute the new Ritz matrix $V_q$ from $T_{q-1}$.

Step 2: Compute a new tridiagonal matrix $T_q$, (Using “polynomial filtering”.)
Applications

* Low-rank approximations of large sparse matrices.

* Exact/Inexact inversions:
  Inner-outer iterative methods for calculating small eigenvalues and small singular values.
Low-rank approximation of a large sparse $m \times n$ matrix $A$

is carried out by replacing $G$ with $A^T A$

In practice $G$ is never computed. Instead, matrix-vector products of the form $z = Gx$ are computed in two steps: First compute $y = Ax$ then $z = A^T y$.

**Advantage:** The left singular vectors (and left Ritz vectors) are not needed. This leads to significant saving of computer storage with respect to Lanczos bidiagonalization. (The final approximation can be expressed as $A V V^T = U V^T$.)
Computing small eigenvalues

Assume for simplicity that the eigenvalues of $G$ satisfy

$$\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n > 0$$

and we want to compute the $k$ smallest eigenvalues.

In this case there is a minor change in Step 1.
Step 1*:  Compute the Rayleigh quotient matrix

\[ S_q = X_{q-1}^T G X_{q-1} \]

and the \( k \) smallest eigenvalues of \( S_q \)

\[ \lambda_1^{(q)} \geq \lambda_2^{(q)} \geq ... \geq \lambda_{n+1-k}^{(q)} \geq ... \geq \lambda_n^{(q)} \]

The related \( k \) eigenvectors are assembled into the \( \ell \times k \) matrix

\[ W_q = [ w_1, w_2, \ldots, w_k ] \]

and used to construct the new matrix of Ritz vectors

\[ V_q = X_{q-1} W_q \]

Since \( X_{q-1} \) and \( W_q \) have orthonormal columns, \( V_q \) inherits this property.
Exact inversions

This improvement is carried out by replacing $G$ with $G^{-1}$ in the construction of the Krylov matrix $B_q = [b_1, b_2, \ldots, b_\ell]$. Here the columns of $B_q$ span a Krylov subspace of $G^{-1}$ that is generated by the following three term recurrence relation

For $j=2, \ldots, k$

Set $b_j = G^{-1}b_{j-1}$

Orthogonalize $b_j$ against $b_{j-1}$

Orthogonalize $b_j$ against $b_{j-2}$

Normalize $b_j$. 
**Exact inversions**

In practice, the columns of $B_q$ span a Krylov subspace of $G^{-1}$ that is obtained by using a **direct method** to solve the related linear systems.

For $j=2, \ldots, k$

- Compute $b_j$ by solving the linear system $Gx = b_{j-1}$
- Orthogonalize $b_j$ against $b_{j-1}$
- Orthogonalize $b_j$ against $b_{j-2}$
- Normalize $b_j$. 

Comparison with “Implicit Restart”

Step 1: Compute the smallest Ritz eigenpairs $V_q$ of $G$.

Step 2: Compute a new Krylov information matrix $B_q$ of $G^{-1}$.

Step 3: Obtain $Z_q$ by orthogonalizing $B_q$ against $V_q$.

Step 4: Compute, $Y_q$, an orthonormal basis of $Z_q$.

Step 5: Define $X_q = [V_q, Y_q]$.

Step 1: Use $T_{q-1}$ to compute the largest Ritz eigenpairs of $G^{-1}$.

Step 2: Compute a new Lanczos matrix $T_q$ of $G^{-1}$.
Inexact inversions

The columns of $B_q$ approximate a Krylov subspace of $G^{-1}$ that is obtained by using an iterative method to solve the related linear systems. (Such as CG, MINRES or GMRES, equipped with a suitable preconditioner.)

For $j = 2, \ldots, k$

Set $b_j$ to be an approximate solution of the linear system $Gx = b_{j-1}$.

Orthogonalize $b_j$ against $b_{j-1}$.

Orthogonalize $b_j$ against $b_{j-2}$.

Normalize $b_j$. 
Comparison with “Implicit Restart”

Step 1: Compute the smallest Ritz eigenpairs \( V_q \) of \( G \).

Step 2: Compute an approximate Krylov matrix \( B_q \) of \( G^{-1} \).

Step 3: Obtain \( Z_q \) by orthogonalizing \( B_q \) against \( V_q \).

Step 4: Compute, \( Y_q \), an orthonormal basis of \( Z_q \).

Step 5: Define \( X_q = [ V_q, Y_q ] \)

Step 1: Use \( T_{q-1} \) to compute the largest Ritz eigenpairs of \( G^{-1} \).

Step 2: Compute an approximate Lanczos matrix \( T_q \) of \( G^{-1} \).
Computing small singular values

$A$ is a large sparse $m \times n$ matrix with singular values

$$\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_n \geq 0,$$

and we want to compute the $k$ smallest singular triplets.

This task is carried out by introducing the following modifications in Steps 1 and 2 of the basic iteration.
Computing small singular values:

The qth iteration, q=1, 2, ... .

Step 1**: Compute the k smallest singular values of $AX_{q-1}$ and the related matrix of right Ritz vectors $V_k$.

Step 2**: Compute an approximate Krylov matrix $B_q$ of $(AT^T A^{-1})$.

Step 3: Obtain $Z_q$ by orthogonalizing $B_q$ against $V_q$.

Step 4: Compute, $Y_q$, an orthonormal basis of $Z_q$.

Step 5: Define $X_q = [V_q, Y_q]$
Step 1**: Compute the Rayleigh matrix

\[ H_q = AX_{q-1} \]

and the k smallest singular values of \( H_q \)

\[ \sigma_1^{(q)} \geq \sigma_2^{(q)} \geq \ldots \geq \sigma_{n+1-k}^{(q)} \geq \ldots \geq \sigma_n^{(q)} \]

The related k right singular vectors are assembled into the \( \ell \times k \) matrix

\[ W_q = [w_1, w_2, \ldots, w_k] \]

and used to construct the new matrix of Ritz vectors

\[ V_q = X_{q-1} W_q \]

Since \( X_{q-1} \) and \( W_q \) have orthonormal columns, \( V_q \) inherits this property.
Step 2**: Inexact inversion of $A^T A$

The columns of $B_q$ approximate a Krylov subspace of $(A^T A)^{-1}$ using an **iterative method** to solve the related linear systems.

For $j = 2, \ldots, k$

Set $b_j$ to be an **approximate solution**

of the linear system $A^T A x = b_{j-1}$.

Orthogonalize $b_j$ against $b_{j-1}$.

Orthogonalize $b_j$ against $b_{j-2}$.

Normalize $b_j$. 
Computing small singular triplets

The new method

Step 1: Compute $V_q$ the smallest right Ritz vectors of $A$.

Step 2: Compute an approximate Krylov matrix $B_q$ of $(A^TA)^{-1}$.

Step 3: Obtain $Z_q$ by orthogonalizing $B_q$ against $V_q$.

Step 4: Compute, $Y_q$, an orthonormal basis of $Z_q$.

Step 5: Define $X_q = [V_q, Y_q]$.

Lanczos bidiagonalization

Step 1: Use $B_{q-1}$ to compute the largest Ritz eigenpairs of $A^{-1}$.

Step 2: Compute a lower bidiagonal matrix $B_q$ from $A^{-1}$.
Concluding remarks

The new method is useful for the following purposes:

* A reduced storage approach for calculating low-rank approximations of large matrices.

* The use of exact/inexact inversions for calculating small eigenvalues and small singular triplets.
  (In this case most of the computation time is spent on the linear solver.)

* The method is easily adapted to use “shift and invert” techniques.
The END

Thank You