Krylov Methods

Krylov space methods are considered as one of the most important methods for solving eigenvalues/eigenvectors. Consider the following linear transformation,

\[ x \rightarrow \text{black box} \rightarrow Ax \]

We want to build an iterative method such that \( x^{(k)} \rightarrow x \) with \( Ax = b \) as \( k \rightarrow \infty \).

We decide to consider \( x^{(k)} \in K_k(A, b) = \text{span}\{b, Ab, \cdots, A^{k-1}b\} \)

This has the following benefits:

1. We can construct \( K_k(A, b) \) with just the black box,
2. \( K_k(A, b) \subseteq K_{k+1}(A, b) \),
3. Denote \( P_k(A) \) as a polynomial of \( A \). Since any linear combination of \( b, Ab, \cdots \) is a polynomial of \( A \) times \( b \), we have \( x^{(k)} = P_k(A)b \).

If we consider the following so-called Krylov matrix

\[ K_n = [b, Ab, \cdots, A^{n-1}b] \]

This matrix is very ill conditioned. The reason is that by convergence of power method, \( A^n b \) approaches a multiple of the dominant eigenvector of \( A \) as \( n \) gets large, making \( K_n \) almost singular. Therefore we need to work with an orthonormal basis.

Given \( n \times n \) matrix \( A \), we can compute an orthogonal matrix \( Q \) and an upper Hessenberg matrix \( H \) (upper triangular + one sub-diagonal) s.t.

\[ A = QHQ^* \]

Let \( Q_k = [q_1, \cdots, q_k] \) be the first \( k \) columns of matrix \( Q \). And let

\[ \tilde{H}_k = \begin{pmatrix} h_{11} & h_{12} & \cdots & h_{1k} \\ h_{21} & h_{22} & \cdots & h_{2k} \\ 0 & h_{32} & \cdots & h_{3k} \\ \vdots & \vdots & & \vdots \\ 0 & \cdots & h_{k,k-1} & h_{kk} \\ 0 & 0 & \cdots & h_{k+1,k} \end{pmatrix} \]

be the upper \((k+1) \times k\) block of \( H \). Then

\[ AQ_k = Q_k \tilde{H}_k \]

\[ Aq_k = h_{1k}q_1 + \cdots + h_{kk}q_k + h_{k+1,k}q_{k+1} \]

This gives us a \((k+1) \) term recurrence for \( q_{k+1} \). We can show that the \( h \)'s correspond to modified Gram-Schmidt orthogonalization.
Arnoldi Algorithm

Algorithm 1 Arnoldi Algorithm

1: procedure
2: Given $A$, $b$, let $v_1 = b/\|b\|_2$
3: for $k = 1, 2, \ldots$ do
4: $q = Av_k$
5: for $j = 1, \ldots, k$ do
6: $h_{jk} = v_j^*q$
7: $q = q - h_{jk}v_j$
8: $h_{k+1,k} = \|q\|_2$
9: $v_{k+1} = q/h_{k+1,k}$

Then $V = [v_1, v_2, \ldots, v_k]$ is an orthonormal basis for $K_k(A, b)$. And we have

$$AV_k = V_{k+1} \tilde{H}_k$$

$$\Rightarrow V_k^*AV_k = \begin{bmatrix} I & 0 \\ \tilde{H}_k \end{bmatrix} = H_k$$

This yields an alternative interpretation of the Arnoldi iteration as a (partial) orthogonal reduction of $A$ to Hessenberg form. The matrix $H_k$ can be viewed as the representation in the basis formed by the Arnoldi vectors of the orthogonal projection of $A$ onto the Krylov subspace $K_k$.

If $A$ is Hermitian, then so is $V_k^*AV_k$. Thus $H_k$ is tridiagonal and we get a three term recurrence which is a further reduction of $A$. This leads us to the Lanczos algorithm.

Lanczos Algorithm

Algorithm 2 Lanczos Algorithm

1: procedure
2: Given $A = A^*$, $b$, $\beta_0$, $v_0 = 0$, let $v_1 = b/\|b\|_2$
3: for $k = 1, 2, \ldots$ do
4: $q = Av_k$
5: $\alpha_k = v_k^*q$
6: $q = q - \beta_{k+1}v_{k+1} - \alpha_kv_k$
7: $\beta_k = \|q\|_2$
8: $v_{k+1} = q/\beta_k$

Define

$$T_k = \begin{bmatrix} \alpha_1 & \beta_1 & 0 \\ \beta_1 & \ddots & \ddots \\ \vdots & \ddots & \beta_{k-1} \\ 0 & \beta_{k-1} & \alpha_k \end{bmatrix}$$

(1)

Then

$$AV_k = V_k\tilde{T}_k = V_k\begin{bmatrix} T_k \\ \beta_k e_k^T \end{bmatrix}$$
While formally the three-term recurrence is exact and no orthogonality is lost, numerically its use results in $V_k$ with columns that are not particularly orthogonal.

We want to solve $Ax = b$ with $x^{(k)} \in K_k(A,b)$ with 0 as our initial guess. $x^{(k)}$ should be the "best" vector in $K_k(A,b)$. We define "best" in terms of:

$$x^{(k)} = \arg\min_{x \in K_k(A,b)} \|Ax - b\|_2^2$$  

Can do this if $A = A^* \Rightarrow$ MINRES. Generally $A \Rightarrow$ GMRES.

What if we define

$$x^{(k)} = \arg\min_{x \in K_k(A,b)} \|x - A^{-1}b\|_2^2$$

$$= x^T x + b^T (A^{-1})^T A^{-1} B - x^T A^{-1} b - b^T (A^{-1})^T x$$

cannot eliminate $(A^{-1})^T$.

**Conjugate gradient (CG)**

$$x^{(k)} = \arg\min_{x \in K_k(A,b)} \|x - A^{-1}b\|_A^2$$

if $A \succ 0$.

We want to solve $\min_{x \in K_k} \|x - A^{-1}b\|_A^2$. Assume that $A$ is symmetric and that $A \succ 0$. We can write this as

$$\min_{y \in \mathbb{R}^k} \|V_k y - A^{-1}b\|_A^2 = \min_{y \in \mathbb{R}^k} (V_k y - A^{-1}b)^T A (V_k y - A^{-1}b)$$

$$= \min_{y \in \mathbb{R}^k} y^T V_k^T A V_k y + b^T A^{-1} b - 2y^T V_k b$$

$$= \min_{y \in \mathbb{R}^k} y^T T_k y - 2y^T e_1 \|b\|$$

Take derivative with respect to $y$ set it to zero, we get

$$T_k y - \|b\| e_1 = 0$$

$$\Rightarrow T_k y = \|b\| e_1$$

CG conceptually:

- Run Lanczos with $A, b$ at each step $k$ to obtain $V_k, T_k$
- Solve $T_k y^{(k)} = \|b\| e_1$
- Then $x^{(k)}$ solves $\min_{x \in K_k} \|x - A^{-1}b\|_A^2$.


**References**
