An Overview of Krylov Methods

Benjamin Vines

(Dated: May 13th, 2025)

I. KRYLOV SUBSPACES

In this paper I will attempt to give an explanation of Krylov iterative methods. We should start with subspaces. Simply put, a subspace is a subset of a larger vector space, which is itself a vector space. We want to look at Krylov subspaces. A Krylov subspace is a subspace of the form:

$$K_m(A, v) \equiv \text{span}\{v, Av, A^2v, \ldots, A^{m-1}v\} \tag{1}$$

K_m is the subspace of all vectors R^n, which can be written as $x = p(A)v$, with p being a polynomial of a degree less than m-1. The Krylov subspace can be used for a variety of iterative approximation methods, known as Krylov Methods.

II. KRYLOV METHODS

Krylov methods are a set of iterative methods which employ krylov subspaces, and can be used to solve large systems of equations and large sparse matrices, matrices which only store nonzero elements as a memory saving tactic. Krylov methods work by manipulating vectors of the matrices, thus avoiding matrix operations. This, generally, saves computing power at the cost of memory. We will cover three of the major ones today.

III. CONJUGATE GRADIENT METHOD

One of the most widely known methods for solving sparse SPD systems is the conjugate gradient method. This consists of Orthogonal projection onto the Krylov subspace, $K_m(r_0, A)$, and minimizes the error over the Krylov subspace. A major advantage of the conjugate gradient method is that each iteration is based solely off the last direction/residual. This makes it much less memory-intensive then other methods, as there is no need to store the entire iteration history. The drawback is that it's much less stable. Similar/derived methods include the conjugate residual method and the bi-conjugate residual method.

An example algorithm:

$$
\begin{array}{ll}
1. & r_0 := b - Ax_0, \quad \beta := \|r_0\|_2, \quad v_1 := \dfrac{r_0}{\beta} \\
2. & \text{For } j = 0, 1, \ldots, \text{ until convergence:} \\
3. & \alpha_j := \dfrac{(r_j, r_j)}{(Ap_j, p_j)} \\
4. & x_{j+1} := x_j + \alpha_j p_j \\
5. & r_{j+1} := r_j - \alpha_j A p_j \\
6. & \beta_j := \dfrac{(r_{j+1}, r_{j+1})}{(r_j, r_j)} \\
7. & p_{j+1} := r_{j+1} + \beta_j p_j
\end{array}
\tag{2}
$$

IV. GMRES

The next method we will discuss is the Generalized Minimum Residual Method. This was developed by Yousef Saad in 1986. Like others, it solves for Ax=b, but for a general nonsymmetric A, using Arnoldi iteration (which we will discuss shortly). Takes $K = K_m$, and $L = Ak_m$, and at each step, solves for least-squares to minimize the residual. This makes it a very stable, and gives it very good convergence for A. However, this comes with the drawback that it is very memory intensive. below is it's main algorithm:

$$
\begin{array}{rl}
1. & \text{Compute } r_0 = b - Ax_0, \quad \beta := \|r_0\|_2, \quad v_1 := \dfrac{r_0}{\beta} \\
2. & \text{For } j = 1, 2, \ldots, m: \\
3. & \text{Compute } w_j := Av_j \\
4. & \text{For } i = 1, \ldots, j \text{ do:} \\
5. & h_{ij} := (w_j, v_i) \\
6. & w_j := w_j - h_{ij} v_i \\
7. & \text{EndDo} \\
8. & h_{j+1,j} = \|w_j\|_2. \quad \text{If } h_{j+1,j} = 0, \text{ set } m := j \text{ and go to 11} \\
9. & v_{j+1} = \dfrac{w_j}{h_{j+1,j}} \\
10. & \text{EndDo} \\
11. & \text{Define the } (m+1) \times m \text{ Hessenberg matrix } \bar{H}_m = \{h_{ij}\}_{1 \leq i \leq m+1, 1 \leq j \leq m} \\
12. & \text{Compute } y_m, \text{ the minimizer of } \|\beta e_1 - \bar{H}_m y\|_2, \quad x_m = x_0 + V_m y_m
\end{array}
\tag{3}
$$

GMRES is a generalization of the Minimal Residual method (MINRES), which also forms the basis for the Quasi-minimal residual method (QMR) and Transpose-free Quasi-minimal residual method (TFQMR).

V. ARNOLDI'S METHOD

Finally we have Arnoldi's method. It was developed by W.E. Arnoldi in 1951. Arnoldi was an american engineer who, in 1951 while working for the Hamilton Standard Division of United Aircraft, published "The Principle of Minimized Iterations In the Solution of the Matrix Eigenvalue Problem." which lays out what is now known as Arnoldi's method, and is one of the most cited works concerning linear algebra.

Arnoldi's method is an algorithm which can be used for orthogonal projection onto K_m for general non-hermitian matrices. It works by multiplying the previous vector, V_j, by A at each step, and normalizing the resulting vector W_j, against all previous V_j's, repeating until W_j goes to zero. It is memory intensive, as it relies on the entirety of previous iterations to continue. It's algorithm is below:

$$
\begin{array}{ll}
1. & \text{Choose a vector } v_1 \text{ such that } \|v_1\|_2 = 1 \\
2. & \text{For } j = 1, 2, \ldots, m : \\
3. & \text{Compute } h_{ij} = (Av_j, v_i) \quad \text{for } i = 1, 2, \ldots, j \\
4. & \text{Compute } w_j := Av_j - \sum_{i=1}^{j} h_{ij} v_i \\
5. & h_{j+1,j} = \|w_j\|_2 \\
6. & \text{If } h_{j+1,j} = 0, \text{ then Stop.} \\
7. & v_{j+1} = \dfrac{w_j}{h_{j+1,j}}
\end{array}
\tag{4}
$$

As we saw earlier, Arnoldi's method of iteration can also be used as the basis for other methods, such as the Conjugate Gradient method.

VI. CONCLUSION

These are just some of the more well known Krylov Methods. There are several more, however, I believe these cover the main points of how they work. It's a fascinating subject, and I hope to spend more time learning them.

[1] W. Arnoldi, "the principle of minimized iterations in the solution of the matrix eigenvalue problem", (1951).

[2] Y. Saad, "iterative methods for sparse linear systems", (2003).