### **Numerical Linear Algebra - Krylov subspaces**

NGSSC

Uppsala University 2011

# Why Krylov subspaces are so much used?

NGSSC, Sci. Comp., January 2011

Maya Nevtcheva, IT, Uppsala University maya@it.uu.se-p.1/16

The Idea Behind Krylov Methods Ilse C. F. Ipsen and Carl D. Meyer The American Mathematical Monthly, Vol. 105, No. 10, Dec., 1998 Why Krylov methods make sense, and why it is natural to represent a solution to a linear system as a member of a Krylov space?

Want to solve  $A\mathbf{x} = \mathbf{b}, \mathbf{b}, \mathbf{x} \in \mathbb{R}^n, A \in \mathbb{R}^{n,n}$ 

Use the projection framework, i.e., we seek an approximate solution  $\widetilde{\mathbf{x}} = \mathbf{x}^0 + \delta$ , where  $\delta \in K$ ,  $dim(K) = m \ll n$ , such that

 $\mathbf{b} - A\widetilde{\mathbf{x}} \perp L, dim(L) = m$ 

 $\mathbf{x}^0$  is arbitrary.

Major results:

- (A) The matrix  $B = W^T A V$  is nonsingular for any W and V either if A is positive definite and L = K, or if  $A^{-1}$  exists and L = AK.
- (B) Properties

(I) 
$$K = L$$
,  $A$ -spd  $\Rightarrow ||\mathbf{x}^* - \widetilde{\mathbf{x}}|| \le ||\mathbf{x}^* - \mathbf{x}||$  for any  
 $\mathbf{x} = \mathbf{x}^0 + \mathbf{y}, \mathbf{y} \in L$ 

(II) 
$$L = AK$$
,  $\Rightarrow \|\mathbf{b} - A\widetilde{\mathbf{x}}\| \le \|\mathbf{b} - A\mathbf{x}\|$  for any  $\mathbf{x} = \mathbf{x}^0 + \mathbf{y}, \mathbf{y} \in L$ 

The importand question is now how to choose K. We let

$$K \equiv \mathcal{K}^m(A, \mathbf{v}) = span\{\mathbf{v}, A\mathbf{v}, \cdots, A^{m-1}\mathbf{v}\}$$

for some vector v. Usual choices:  $\mathbf{v} = \mathbf{b}$  or  $\mathbf{v} = \mathbf{r}^0 \equiv \mathbf{b} - A\mathbf{x}^0$ .

Relevant questions:

- Why is  $\mathcal{K}(A, \mathbf{b})$  often a good space from which to construct an approximate solution?
- Why are eigenvalues important for Krylov methods
- Why do Krylov methods often do so well for Hermitian matrices?

One can show that the solution of  $A\mathbf{x} = \mathbf{b}$  has a natural representation in  $\mathcal{K}_k(A, \mathbf{b})$  for some k.

If k happens to be small, we have a fast convergence.

The minimal polynomial of A,  $q_d(t)$  of degree d, is the unique monoic polynomial of minimal degree, for which

$$q(A) = 0.$$

It has the form

$$q_d(t) = \prod_{j=1}^d (t - \lambda_j)^{m_j},$$

where

-  $\lambda_1, \cdots, \lambda_d$  are distinkt eigenvalues of A,

-  $m_1, \dots, m_d$  are the corresponding indeces of  $\lambda_j$  (the sizes of the largest Jordan block, associated with  $\lambda_j$ ).

$$q_d(t) = \prod_{j=1}^d (t - \lambda_j)^{m_j} = \sum_{s=0}^m \alpha_s t^s,$$
 (1)

where 
$$m \sum_{j=1}^{d} m_j$$
.  
Example:  $A = \begin{bmatrix} 3 & 1 \\ & 3 \\ & & 4 \\ & & 4 \end{bmatrix}$ .  
Then we have  $\lambda_1 = 3, m_1 = 2, \lambda_2 = 4, m_2 = 1$ .  
Note that, since we have assumed that  $A$  is nonsingular, in  
(1), the coefficient  $\alpha_0 = \prod_{j=1}^{d} (-\lambda_j)^{m_j} \neq 0$ .

$$q(A) = \alpha_0 I_n + \alpha_1 A + \alpha_2 A^2 + \dots + \alpha_m A^m = 0, \ \alpha_0 \neq 0$$
  
Then  $A^{-1}q(A) = 0$ , thus,

$$A^{-1} = \frac{1}{\alpha_0} \sum_{j=0}^{m-1} \alpha_{j+1} A^j$$

However,  $\mathbf{x} = A^{-1}\mathbf{b}$  !

If the minimal polynomial of A ( $A^{-1}\exists$ ) has degree m, then  $\mathbf{x} = A^{-1}\mathbf{b} \in \mathcal{K}^m(A, \mathbf{b})$ .



Remarks:

- If d is small, then the convergence is fast.
- We also see that the eigenvalues of A, not its singular values, are important, because the dimension of the solution space is determined by the degree of the minimal polynomial.

Suppose that A is singular. One can show that even if a solution exists, it may not lie in the Krylov space  $\mathcal{K}^m(A, \mathbf{b})$ .

Example: Consider a consistent linear system  $N\mathbf{x} = \mathbf{c}$ , where N is a nulpotent matrix, i.e., there exists some integer  $\ell$ , such that  $N^{\ell} = 0$  but  $N^{\ell-1} \neq 0$ . Suppose that the solution  $\mathbf{x}$  is a linear combination of Krylov vectors, i.e.,

$$\mathbf{x} = \beta_0 \mathbf{c} + \beta_1 N \mathbf{c} + \beta_2 N^2 \mathbf{c} + \dots + \beta_{\ell-1} N^{\ell-1} \mathbf{c}$$

Then,  $\mathbf{c} = N\mathbf{x} = \beta_0 N\mathbf{c} + \beta_1 N^2 \mathbf{c} + \dots + \beta_{\ell-2} N^{\ell-1} \mathbf{c}$  and  $(I - \beta_0 N - \beta_1 N^2 - \dots - \beta_{\ell-2} N^{\ell-1}) \mathbf{c} = 0.$ 

The matrix  $Q = I - \beta_0 N - \beta_1 N^2 - \cdots - \beta_{\ell-2} N^{\ell-1}$  is nonsingular, because of the following reasons. The eigenvalues of any nilpotent matrix are all equal to zero, thus, the eigenvalues of Q are all equal to 1. Therefore, **c** must be zero.

Moral: the solution of a system with a nilpotent matrix and a nonzero right hand side cannot lie in the Krylov subspace, generated by the matrix and the rhs.

#### What happens if $A^{-1}$ does not exist?

Apply the following trick: Decompose the space  $C^n = \mathcal{R}(A^{\ell}) \oplus \mathcal{N}(A^{\ell})$ , where  $\ell$  is the index of the zero eigenvalue of  $A \in C^{n \times n}$  and  $\mathcal{R}(\cdot)$  and  $\mathcal{N}(\cdot)$  denote range and nullspace. Then

$$A = \begin{bmatrix} R & 0 \\ 0 & N \end{bmatrix},$$

where R is nonsingular and N is nilpotent of index  $\ell$ .

Suppose now that  $A\mathbf{x} = \mathbf{b}$  has a Krylov solution

$$\mathbf{x} = \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{bmatrix} = \sum_{j=1}^d \alpha_j A \mathbf{b} = \sum_{j=0}^d \alpha_j \begin{bmatrix} R^j & 0 \\ 0 & N^j \end{bmatrix} \begin{bmatrix} \mathbf{b}_1 \\ \mathbf{b}_2 \end{bmatrix}$$

thus

$$\mathbf{x}_1 = \sum_{j=0}^d \alpha_j R^j \mathbf{b}_1$$
 and  $\mathbf{x}_2 = \sum_{j=0}^d \alpha_j N^j \mathbf{b}_2$ .

## What happens if $A^{-1}$ does not exist?

From  $A\mathbf{x} = \mathbf{b}$  we have that  $N\mathbf{x}_2 = \mathbf{b}_2$ , so  $\sum_{j=0}^{d-1} \alpha_j N^{j+1} \mathbf{b}_2 = \mathbf{b}_2$  and

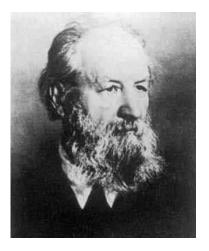
$$(I - \sum_{j=0}^{d-1} \alpha_j N^{j+1})\mathbf{b}_2 = \mathbf{0}$$

and following analogous reasons we obtain that  $b_2 = 0$ .

In other words, The existence of a Krylov solution requires that  $\mathbf{b} \in \mathcal{R}(A^{\ell})$ . The converse statement is also true.

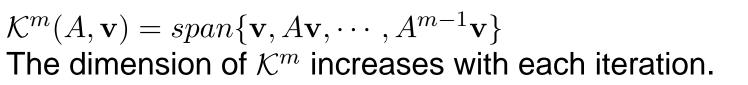
**Theorem 1** A square linear system  $A\mathbf{x} = \mathbf{b}$  has a Krylov solution if and only if  $\mathbf{b} \in \mathcal{R}(A^{\ell})$ , where  $\ell$  is the index of the zero eigenvalue of A.

# **Alexei Nikolaevich Krylov**



1863-1945, Maritime Engineer

- 300 papers and books on: shipbuilding, magnetism,artilery,math, astronomy, geodesy
- 1890: theory of oscillating motions of the ship
- **1904:** he built the first machine in Russia for integrating ODEs
- 1931: Krylov subspace methods



- Theorem [Cayley-Hamilton]:  $d \le n$
- $\mathcal{K}^d$  is invariant under A, thus,  $\mathcal{K}^m = \mathcal{K}^d$  for m > d, thus,

 $dim(\mathcal{K}^m) = \min(m, d)$