Krylov subspace methods and GMRES

(Generalized Minimal Residuals)

21/02/2024

Recap:

- Partial reduction of A to its Heisenberg form:

- $AQ_k = Q_{k+1}\tilde{H}_k$, where

$$\tilde{H}_k = \begin{bmatrix} h_{11} & h_{12} & \cdots & h_{1k} \\ h_{21} & h_{22} & \cdots & h_{2k} \\ & \ddots & \ddots & \vdots \\ & & h_{k,k-1} & h_{kk} \\ & & & h_{k+1,k} \end{bmatrix} \in \mathbb{R}^{(k+1)\times k}$$

$$\tilde{H}_k = \begin{bmatrix} h_{11} & h_{12} & \cdots & h_{1k} \\ h_{21} & h_{22} & \cdots & h_{2k} \\ & \ddots & \ddots & \vdots \\ & & h_{k,k-1} & h_{kk} \\ & & & 0 \end{bmatrix} \in \mathbb{R}^{(k+1)\times k}$$

- $Aq_k = h_{1k}q_1 + h_{2k}q_2 + \cdots + h_{kk}q_k + h_{k+1,k}q_{k+1}$
- $AQ_k = Q_k H_k$
- $\mathcal{K}_k = \mathcal{K}_{k+1} = \mathcal{K}_{k+2} = \cdots$
- if (λ, y) is an eigenpair of H_k , then $(\lambda, Q_k y)$ is an eigenpair of A .

If there is breakdown of the Arnoldi iteration,

i.e., $h_{k+1,k} = 0, AQ_k = Q_k H_k$

- True solution
  $x_* = A^{-1}b \in \mathcal{K}_k$
- We can stop the iteration.
- This is a good breakdown.

From now on, we assume Arnoldi iteration does not break down.

- $Q_k^* Q_{k+1} = I_{k\times(k+1)}$
- What effect does $I_{k\times(k+1)}$ has on a given matrix $M \in \mathbb{R}_{(k+1,m)}$?
- $I_{k\times(k+1)}$ eliminates the last row of M.

GMRES method

- Original problem: Find $x_* \in \mathbb{R}^n$ s.t. $Ax_* = b$
- Minimizing problem: Find $x \in \mathscr{V}$ of dimension $k \ll n$ s.t. $\|Ax - b\|$ is minimum.
- $\rightarrow$Find $y \in \mathbb{R}^k$ s.t. $\|AQ_k y - b\|$ is minimum.
- $\rightarrow$Find $y \in \mathbb{R}^k$ s.t. $\|\tilde{H}_k y - \beta e_1\|$ is minimum.
- $\beta = \|b\|, e_1 = [1,0,\cdots\ 0] \in \mathbb{R}^{k+1}$.
- Once y is found, $x_k = Q_k y$ is the best estimate of x_*.

Assignment 2: Write a Matlab code for GMRES algorithm.

- Hint:
- $q_1 = b/\|b\|$.
- For $i = 1,2,\cdots$
  - The i^{th} step of Arnoldi iteration
  - Find y that minimizes $\|\tilde{H}_k y - \beta e_1\|$
  - $x_i = Q_i y$

- What method could you apply to minimize $\|\tilde{H}_k y - \beta e_1\|$?
- Think about a stopping criterion.

What method could you apply to find y that minimizes $\|\tilde{H}_k y - \beta e_1\|$?

i.e. a solution to

$$\arg\min_{y \in \mathbb{R}^k} \|\tilde{H}_k y - \beta e_1\|$$

- Full QR factorization of $\tilde{Q}_k \tilde{R}_k = \tilde{H}_k \in \mathbb{R}^{(k+1)\times k}$
- $\tilde{Q}_k \in \mathbb{R}^{(k+1)\times(k+1)}$ is orthonormal, i.e., $\tilde{Q}_k^* \tilde{Q}_k = I_{k+1}$
- $\tilde{R}_k \in \mathbb{R}^{(k+1)\times k}$ is upper trianglar matrix
- $\|\tilde{H}_k y - \beta e_1\| = \|\tilde{Q}_k \tilde{R}_k y - \beta e_1\|$
- $= \|\tilde{Q}_k^* (\tilde{Q}_k \tilde{R}_k y - \beta e_1)\|$
- $= \|\tilde{R}_k y - \beta \tilde{Q}_k^* e_1\|$

- What does the last row of $\tilde{R}_k$ look like?

- $$[q_{(1)} \,|\, q_{(2)} \,|\, \cdots \,|\, q_{(k)} \,|\, q_{(k+1)}] \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1,k} \\ & h_{22} & \cdots & r_{2,k} \\ & & \ddots & \vdots \\ & & & r_{kk} \\ & & & r_{k+1,k} \end{bmatrix}$$

- $$\|\tilde{R}_k y - \beta \tilde{Q}_k^* e_1\| = \| \begin{bmatrix} R_k \\ 0_{1\times k} \end{bmatrix} y - \beta \tilde{Q}_k^* e_1 \|$$

- $$= \| \begin{bmatrix} -r_1- \\ -r_2- \\ \vdots \\ -r_k- \\ -0- \end{bmatrix} y - \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \\ \beta_{k+1} \end{bmatrix} \|$$

- $$\left\| \begin{bmatrix} -r_1- \\ -r_2- \\ \vdots \\ -r_k- \\ -0- \end{bmatrix} y - \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \\ \beta_{k+1} \end{bmatrix} \right\|^2$$
- $$= \left\| \begin{bmatrix} r_1 y \\ r_2 y \\ \vdots \\ r_k y \\ 0 \end{bmatrix} - \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_k \\ \beta_{k+1} \end{bmatrix} \right\|^2 = \left\| \begin{bmatrix} r_1 y - \beta_1 \\ r_2 y - \beta_2 \\ \vdots \\ r_k y - \beta_k \\ -\beta_{k+1} \end{bmatrix} \right\|^2$$
- $$= (r_1 y - \beta_1)^2 + (r_2 y - \beta_2)^2 + \cdots + (r_k y - \beta_k)^2 + \beta_{k+1}^2$$

- $\arg\min\limits_{y \in \mathbb{R}^k} [(r_1 y - \beta_1)^2 + (r_2 y - \beta_2)^2 + \cdots + (r_k y - \beta_k)^2 + \beta_{k+1}^2]$
- $? = \arg\min\limits_{y \in \mathbb{R}^k} [(r_1 y - \beta_1)^2 + (r_2 y - \beta_2)^2 + \cdots + (r_k y - \beta_k)^2]$ ✅
- $\arg\min\limits_{y \in \mathbb{R}^k} \| \begin{bmatrix} R_k \\ 0_{1 \times k} \end{bmatrix} y - \beta \tilde{Q}_k^* e_1 \| = \arg\min\limits_{y \in \mathbb{R}^k} \| R_k y - \tilde{\beta}_k \|$
- $\tilde{\beta}_k = [\beta_1, \beta_2, \cdots, \beta_k]^T$

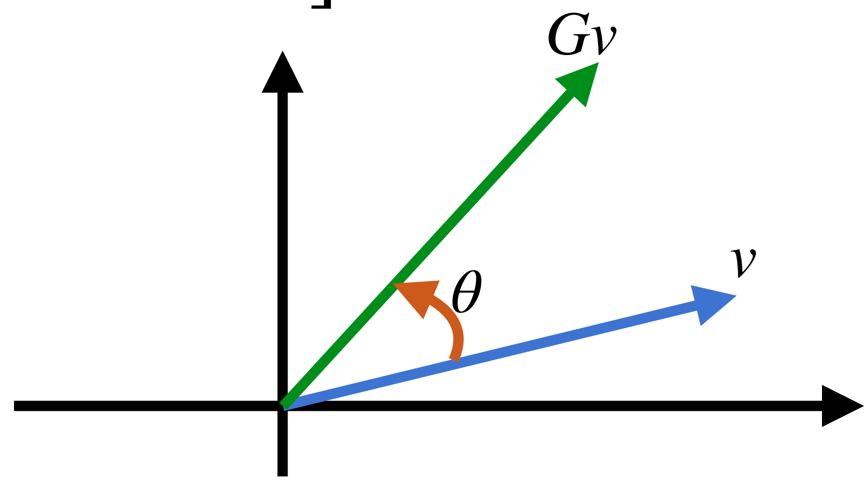
GMRES algorithm:

- Set tolerance t, e.g. $t = 10^{-4}$
- $q_1 = b/\|b\|$
- For $k = 1,2,\cdots$
    - The k^{th} step of Arnoldi iteration to find $\tilde{H}_k \in \mathbb{R}^{(k+1)\times k}$
    - $\tilde{H}_k = \tilde{Q}_k \tilde{R}_k$
      (Matlab command: [Q,R]=qr(H).)

      QR requires $O(k^2)$ flops.
    - Find y that minimizes $\|\tilde{R}_k y - \beta \tilde{Q}_k^* e_1\|$
      (Matlab command: y=R\B for solving Ry=B.)
    - $x_k = Q_k y$
    - If relative residual norm $\dfrac{\|Ax_k - b\|}{\|b\|} < t$
    - End loop.

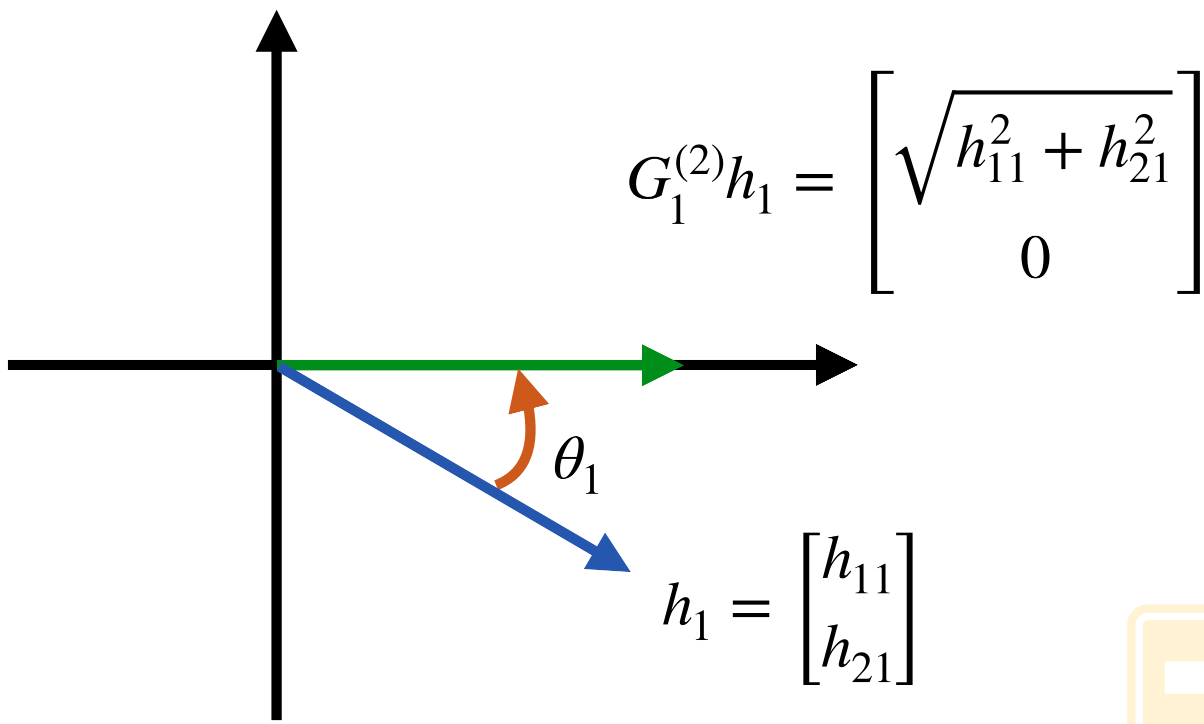
More efficient way for QR factorization of $\tilde{H}_k = \tilde{Q}_k \tilde{R}_k$

- Given's rotation matrix $G = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}$
- $Gv = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$
- $\quad = \begin{bmatrix} v_1\cos\theta + v_2\sin\theta \\ -v_1\sin\theta + v_2\cos\theta \end{bmatrix}$



- $G^TG = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}$
- $\quad = \begin{bmatrix} \cos^2\theta + \sin^2\theta & \cos\theta\sin\theta - \sin\theta\cos\theta \\ \sin\theta\cos\theta - \cos\theta\sin\theta & \sin^2 + \cos^2\theta + \theta \end{bmatrix}$
- $\quad = I_2$

- If $k = 1$, $\tilde{H}_1 = \begin{bmatrix} h_{11} \\ h_{21} \end{bmatrix}$
- $G_1^{(2)} = \begin{bmatrix} c_1 & s_1 \\ -s_1 & c_1 \end{bmatrix}$, where $c_1 = \cos\theta_1, s_1 = \sin\theta_1$



- If $k = 1$, $\tilde{H}_1 = \begin{bmatrix} h_{11} \\ h_{21} \end{bmatrix}$
- $G_1^{(2)} = \begin{bmatrix} c_1 & s_1 \\ -s_1 & c_1 \end{bmatrix}$, where $c_1 = \cos\theta_1, s_1 = \sin\theta_1$
- $G_1^{(2)}\tilde{H}_1 = \begin{bmatrix} \sqrt{h_{11}^2 + h_{21}^2} \\ 0 \end{bmatrix}$
- $= \begin{bmatrix} c_1 & s_1 \\ -s_1 & c_1 \end{bmatrix} \begin{bmatrix} h_{11} \\ h_{21} \end{bmatrix} = \begin{bmatrix} c_1 h_{11} + s_1 h_{21} \\ -s_1 h_{11} + c_1 h_{21} \end{bmatrix}$
- If $k = 2$, $\tilde{H}_2 = \begin{bmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \\ & h_{32} \end{bmatrix}$

- $G_1^{(3)} = \begin{bmatrix} c_1 & s_1 & \\ -s_1 & c_1 & \\ & & 1 \end{bmatrix}$,
- $G_2^{(3)} = \begin{bmatrix} 1 & & \\ & c_2 & s_2 \\ & -s_2 & c_2 \end{bmatrix}$, where $c_2 = \cos\theta_2, s_2 = \sin\theta_2$.
- $G_2^{(3)} G_1^{(3)} \tilde{H}_2 = \begin{bmatrix} 1 & & \\ & c_2 & s_2 \\ & -s_2 & c_2 \end{bmatrix} \begin{bmatrix} c_1 & s_1 & \\ -s_1 & c_1 & \\ & & 1 \end{bmatrix} \begin{bmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \\ & h_{32} \end{bmatrix}$
- $$= \begin{bmatrix} 1 & & \\ & c_2 & s_2 \\ & -s_2 & c_2 \end{bmatrix} \begin{bmatrix} \sqrt{h_{11}^2 + h_{21}^2} & c_1 h_{21} + s_1 h_{22} \\ 0 & s_1 h_{11} + c_1 h_{21} \\ 0 & h_{32} \end{bmatrix} \rightarrow \begin{bmatrix} r_{11} & \alpha \\ 0 & \beta \\ 0 & \boxed{h_{32}} \end{bmatrix}$$

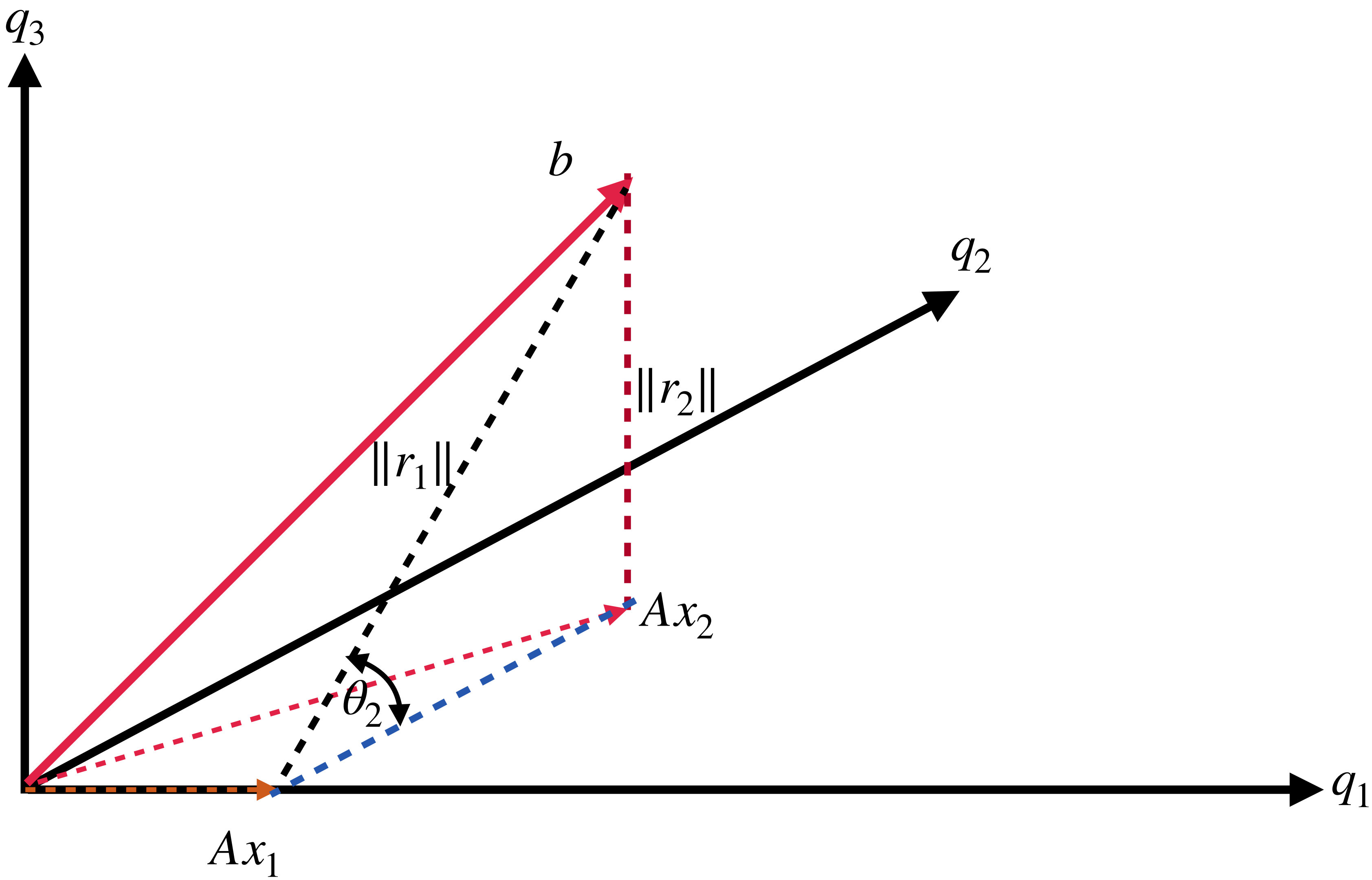
- $= \begin{bmatrix} 1 & & \\ & c_2 & s_2 \\ & -s_2 & c_2 \end{bmatrix} \begin{bmatrix} r_{11} & \alpha \\ 0 & \beta \\ 0 & h_{32} \end{bmatrix}$
- $= \begin{bmatrix} r_{11} & \alpha \\ 0 & c_2\beta + s_2 h_{32} \\ 0 & -s_2\beta + c_2 h_{32} \end{bmatrix}$ ✅ ✅

$$\begin{bmatrix} \beta \\ h_{32} \end{bmatrix} \to \begin{bmatrix} \sqrt{\beta^2 + h_{32}^2} \\ 0 \end{bmatrix}$$

$\boxed{G_1^{(2)}\tilde{H}_1}$

$k = 2,3,\cdots,n$

Updating requires only $O(k)$ flops.



$$\sin\theta_2 = \frac{\|r_2\|}{\|r_1\|}$$

Proposition**: The GMRES residual satisfies the following equation:

$$\frac{\|r_k\|}{\|r_{k-1}\|} = |\sin \theta_k|$$

**Y. Saad and M.H. Schultz, GMRES: a generalized minimal residual algorithm for solving nonsymmetric linear systems, SIAM J. Sci. Statist. Comput. 7 (1986) 856-869.

Residual norm of GRMES method is at least non-increasing.

Example 1: Space-fractional diffusion problem

Consider a dense Toeplitz matrix that arises from the discretization of space-fractional diffusion problems.

$$\frac{\mathrm{d}^\alpha u(x)}{\mathrm{d}_+ x^\alpha} = \frac{1}{\Gamma(n-\alpha)} \frac{\mathrm{d}^n}{\mathrm{d}x^n} \int_L^x \frac{u(\xi)}{(x-\xi)^{\alpha+1-n}} \mathrm{d}\xi$$

$$\frac{\mathrm{d}^\alpha u(x)}{\mathrm{d}_- x^\alpha} = \frac{(-1)^n}{\Gamma(n-\alpha)} \frac{\mathrm{d}^n}{\mathrm{d}x^n} \int_x^R \frac{u(\xi)}{(\xi-x)^{\alpha+1-n}} \mathrm{d}\xi$$

$$A \in \mathbb{R}^{300\times 300}, b \in \mathbb{R}^{300}$$

Reorthogonalizing in Arnoldi iterations:

Example

1. $q_3 = q_3'/h_{32}$

2. $q_3 = q_3 - (q_3^* q_1)q_1 - (q_3^* q_2)q_2$

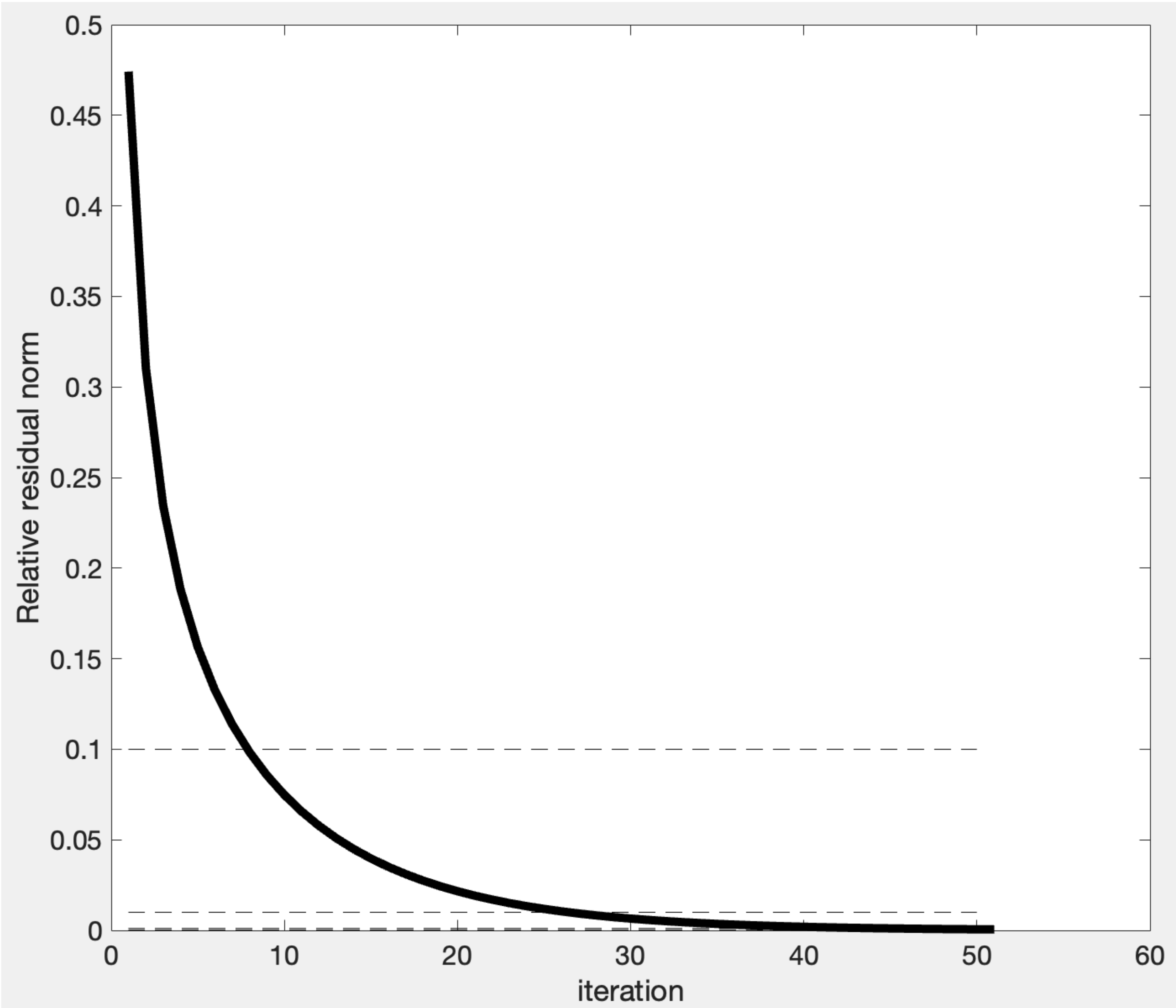
3. $q_3 = q_3/\|q_3\|$

Regular:

	1	2	3
1	1.0000	-6.2450e-17	-7.1774e-17
2	-6.2450e-17	1.0000	1.4408e-15
3	-7.1774e-17	1.4408e-15	1.0000
4	-8.3185e-17	1.5911e-15	-2.7569e-15
5	-9.4759e-17	1.7896e-15	-3.2541e-15
6	-9.3160e-17	2.0143e-15	-3.6949e-15
7	-9.4841e-17	2.2730e-15	-4.1687e-15
8	-1.0519e-16	2.5607e-15	-4.7023e-15
9	-6.5377e-17	2.8840e-15	-5.2955e-15
10	-1.3046e-16	3.2436e-15	-5.9714e-15
11	-2.2253e-16	3.6599e-15	-6.7269e-15
12	-2.8636e-16	4.1199e-15	-7.5823e-15
13	-3.2217e-16	4.6436e-15	-8.5498e-15
14	-4.2769e-16	5.2356e-15	-9.6371e-15
15	-5.9843e-16	5.8983e-15	-1.0858e-14
16	-6.8635e-16	6.6470e-15	-1.2236e-14
17	-8.2193e-16	7.4897e-15	-1.3787e-14
18	-9.8771e-16	8.4393e-15	-1.5539e-14
19	-1.0072e-15	9.5135e-15	-1.7510e-14
20	-1.1611e-15	1.0721e-14	-1.9736e-14

Reorthogonalied:

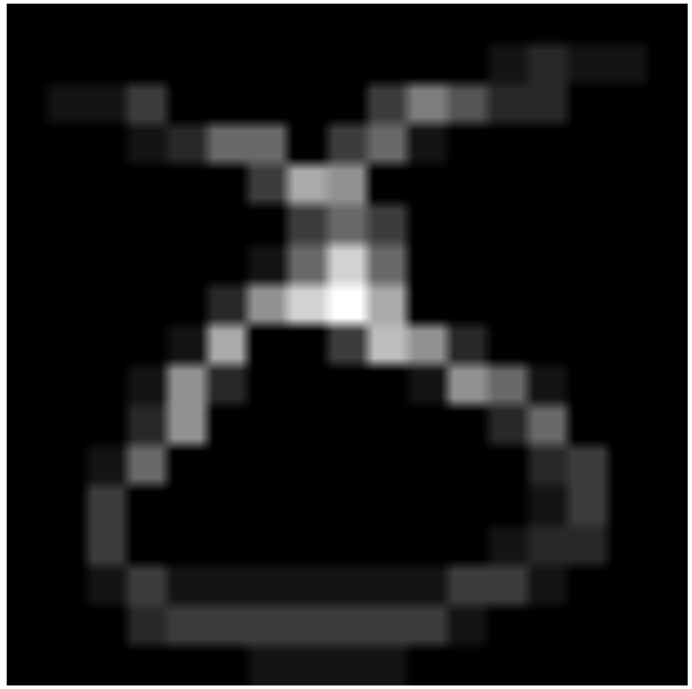
	1	2	3
1	1.0000	-6.2450e-17	1.1791e-17
2	-6.2450e-17	1.0000	1.4403e-15
3	1.1791e-17	1.4403e-15	1.0000
4	4.1200e-18	-1.6263e-18	-2.7485e-15
5	1.8052e-17	1.0842e-19	3.5745e-19
6	-7.3726e-18	-5.4210e-19	-1.9389e-18
7	-6.3426e-18	-1.1926e-18	1.8626e-18
8	5.7463e-18	-1.8431e-18	-2.0896e-18
9	-2.3039e-17	0	6.5560e-19
10	-1.1249e-17	4.3368e-19	3.7778e-19
11	-3.8299e-17	-9.7578e-19	-1.0308e-18
12	2.2579e-17	-1.1926e-18	-3.4610e-18
13	1.9326e-17	1.6263e-18	-1.3722e-18
14	4.0983e-17	-6.5052e-19	6.9626e-19
15	1.0923e-17	2.1684e-19	1.7102e-18
16	5.5023e-18	0	5.9123e-19
17	-1.0625e-17	0	9.4359e-19
18	-6.2613e-18	-2.1684e-19	-1.6094e-20
19	-1.1411e-17	7.5894e-19	-1.6610e-18
20	1.1899e-17	7.5894e-19	-1.8381e-19



1	0.4741
2	0.3105
3	0.2344
4	0.1883
5	0.1563
6	0.1324
7	0.1135
8	0.0982
9	0.0855
10	0.0748
11	0.0656
12	0.0577
13	0.0509
14	0.0449
15	0.0397
16	0.0351
17	0.0311
18	0.0275
19	0.0244
20	0.0216

Example 2: Cameraman

(a)
True Image
494×494 pixels



(b)
non-symmetric
PSF: 17×17 pixels

(c)
Blurred Image
494×494 pixels

$x_* \in \mathbb{R}^{244036}$

Columnized,

Matlab code x(:)

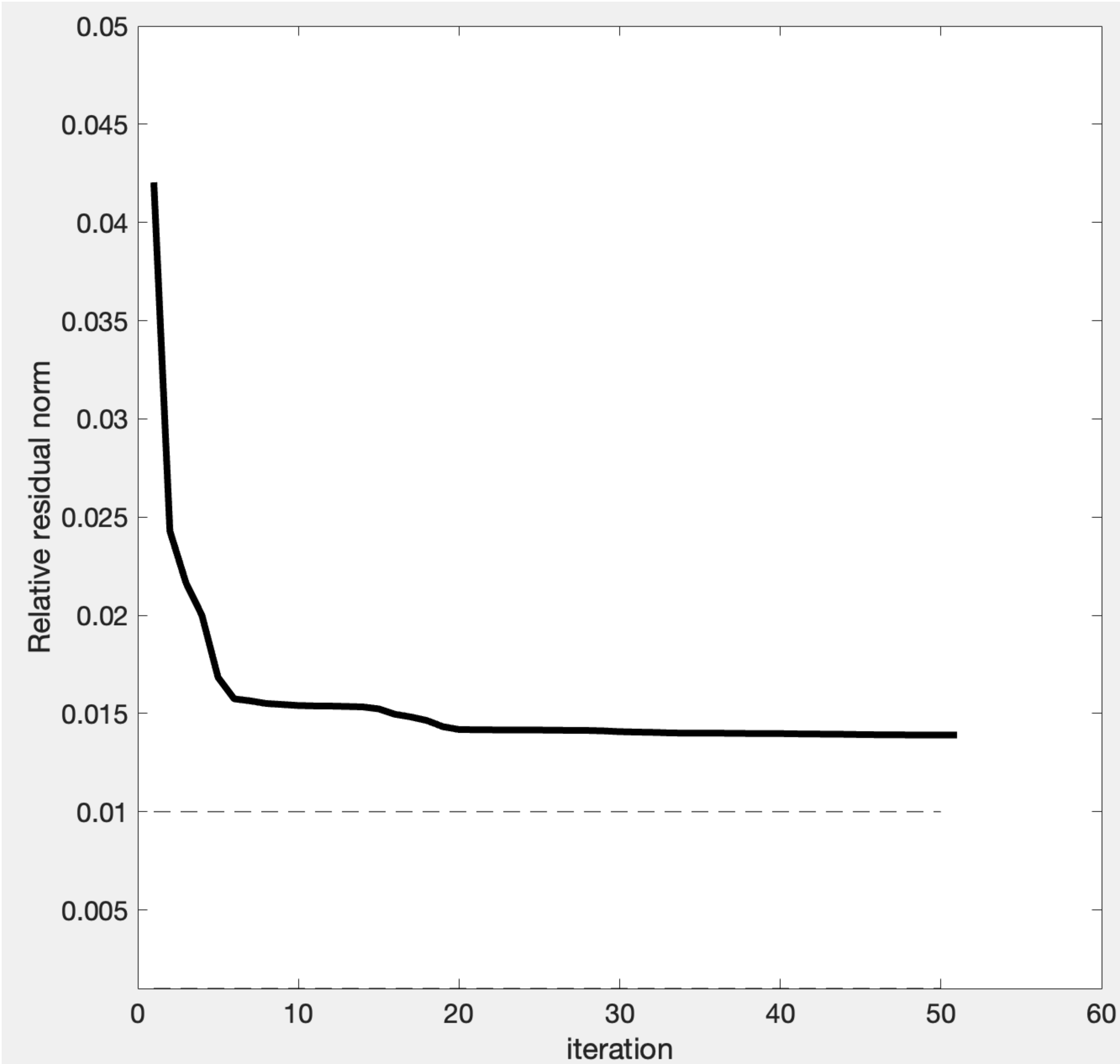
$A \in \mathbb{R}^{244036 \times 244036}$

Sparse,

NEVER explicitly formed

$b \in \mathbb{R}^{244036}$

Columnized



20	0.0142
21	0.0142
22	0.0142
23	0.0142
24	0.0142
25	0.0142
26	0.0141
27	0.0141
28	0.0141
29	0.0141
30	0.0141
31	0.0141
32	0.0140
33	0.0140
34	0.0140

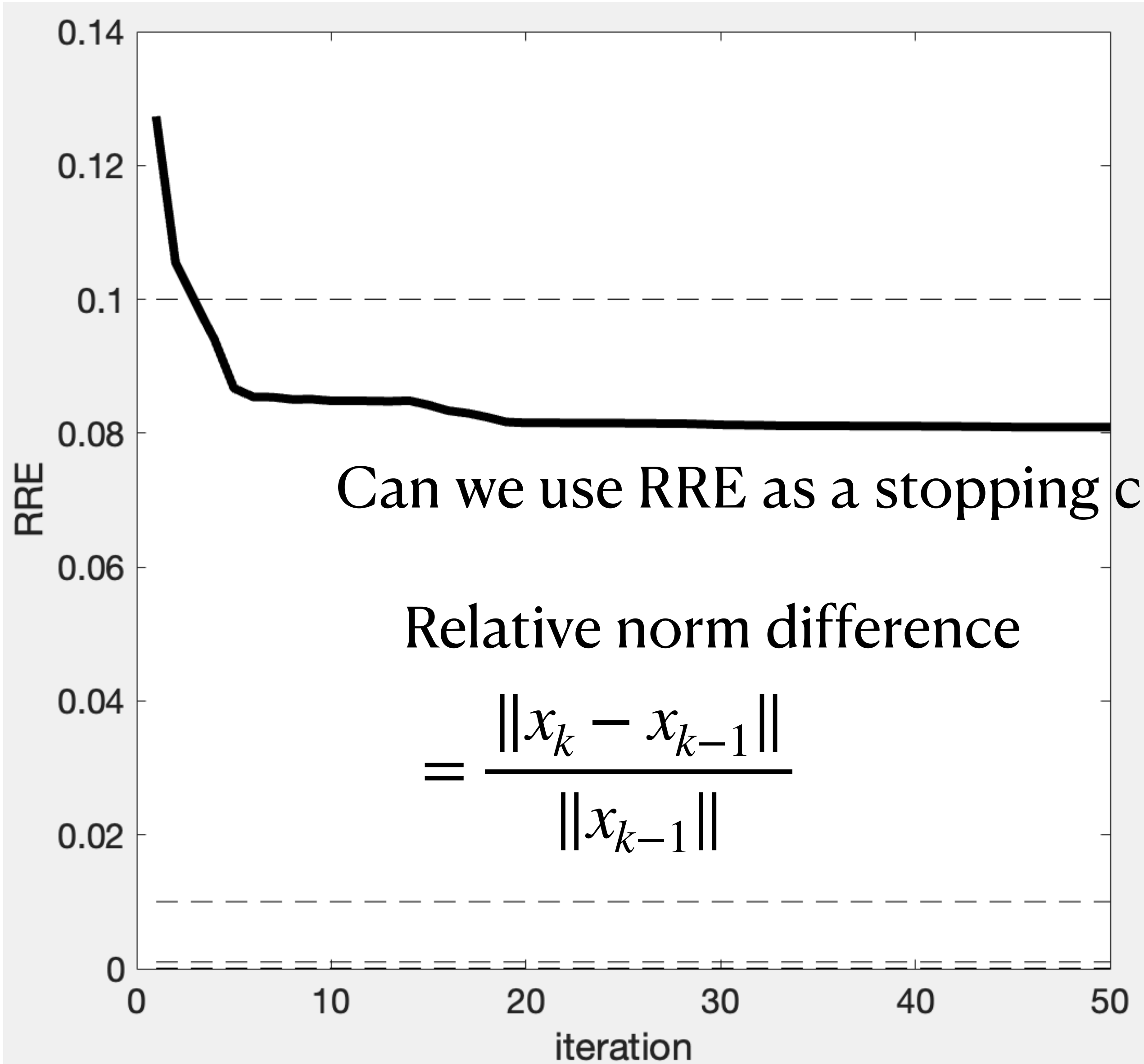

k=1
k=2
k=3


k=6
k=14
k=20

RRE: Relative restoration error

$$\mathrm{RRE}(x_k) = \frac{\|x_k - x_*\|}{\|x_*\|}$$

Unknown



Can we use RRE as a stopping criterion?

Relative norm difference

$$= \frac{\|x_k - x_{k-1}\|}{\|x_{k-1}\|}$$

The convergence of GMRES depends on the property of the matrix A, such as if it is

- Symmetric/Hermitian ($A^* = A$);
- Ill-conditioned (sensitive to perturbation);
- Normal ($AA^* = A^*A$), etc.

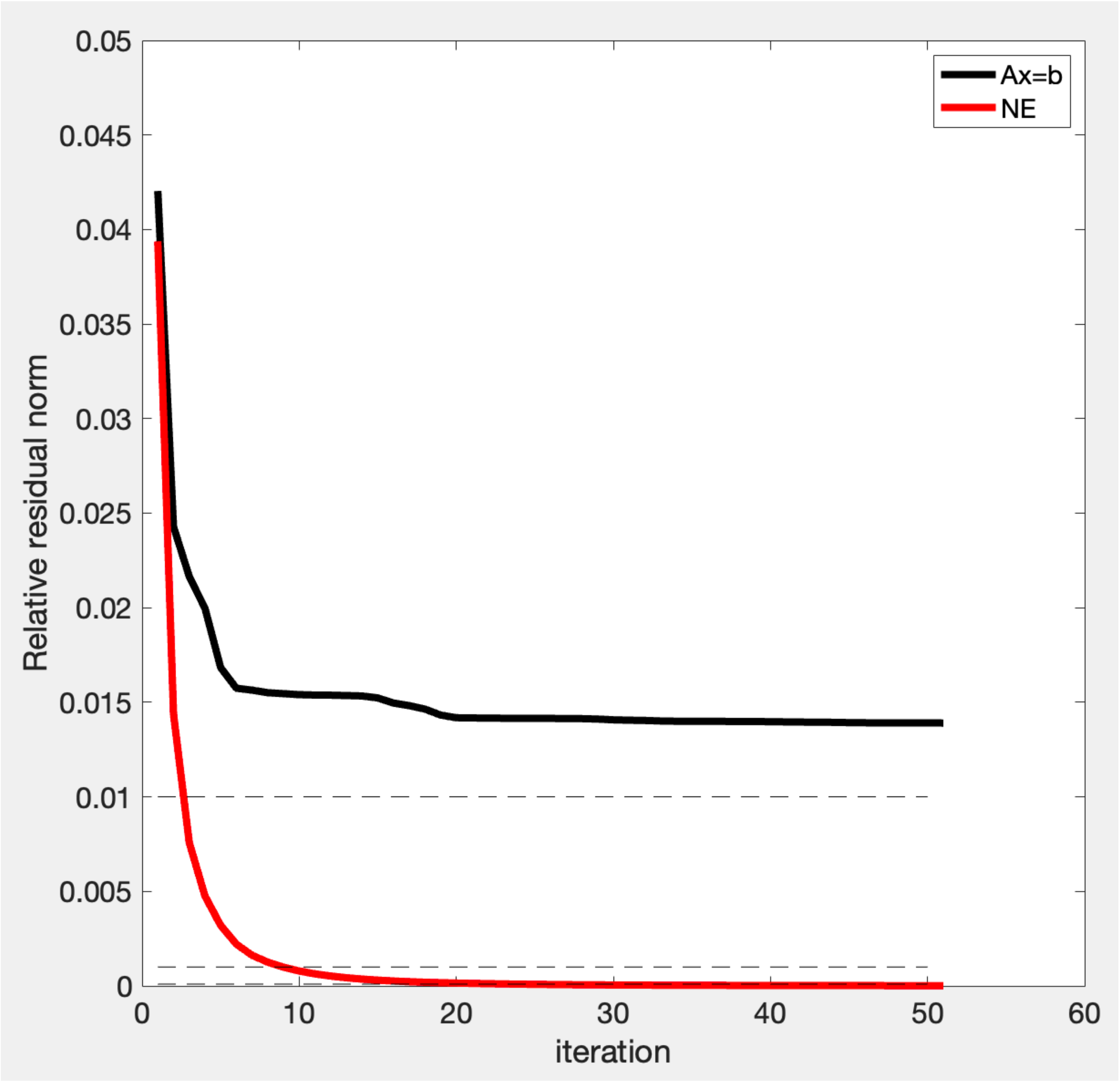
Other Krylov subspace

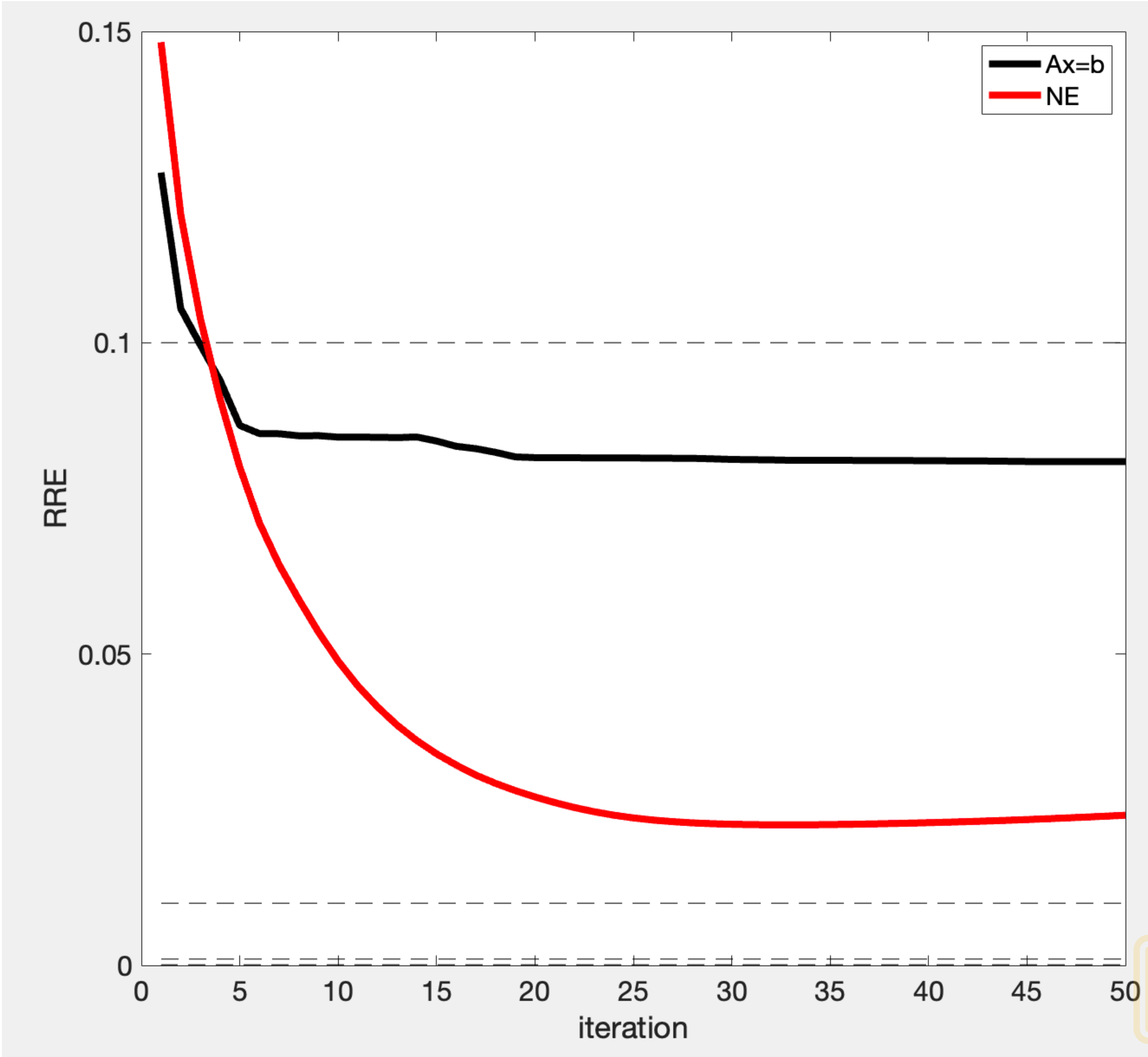
- Normal equation of $Ax = b$
- $\rightarrow A^T A x = A^T b$
- What is the related Krylov subspaces of the normal equation?
- $\mathcal{K}_k(A^T A, A^T b)$.
- $A^T A$ is usually not explicitly formed. How do we modify our code?

Arnoldi iterations for creating $\mathcal{K}_k(A^T A, A^T b)$,

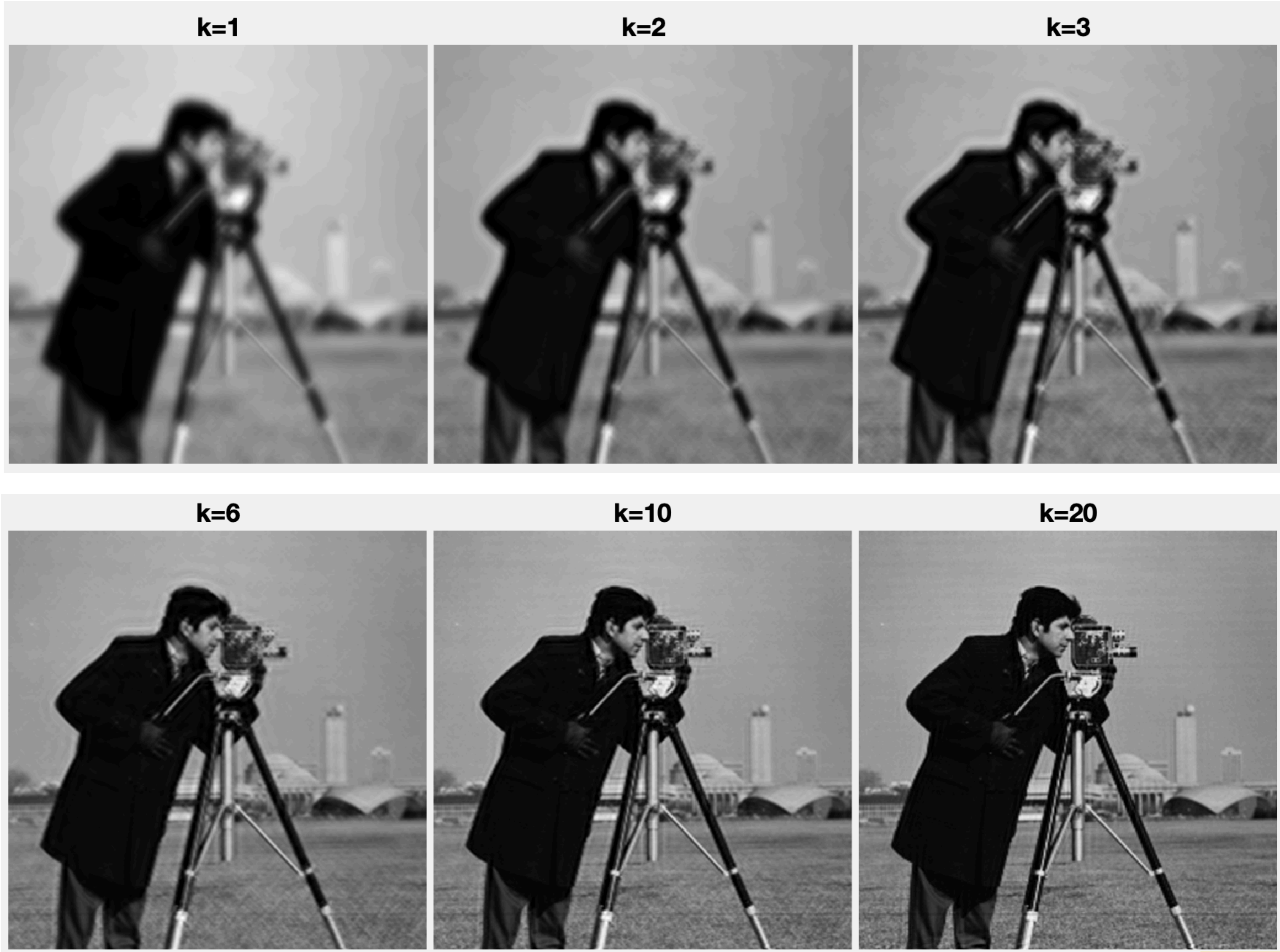
- $b = A^T b$
- $v = A^T (A q_2)$
- ...

Example 3: Cameraman Normal equation

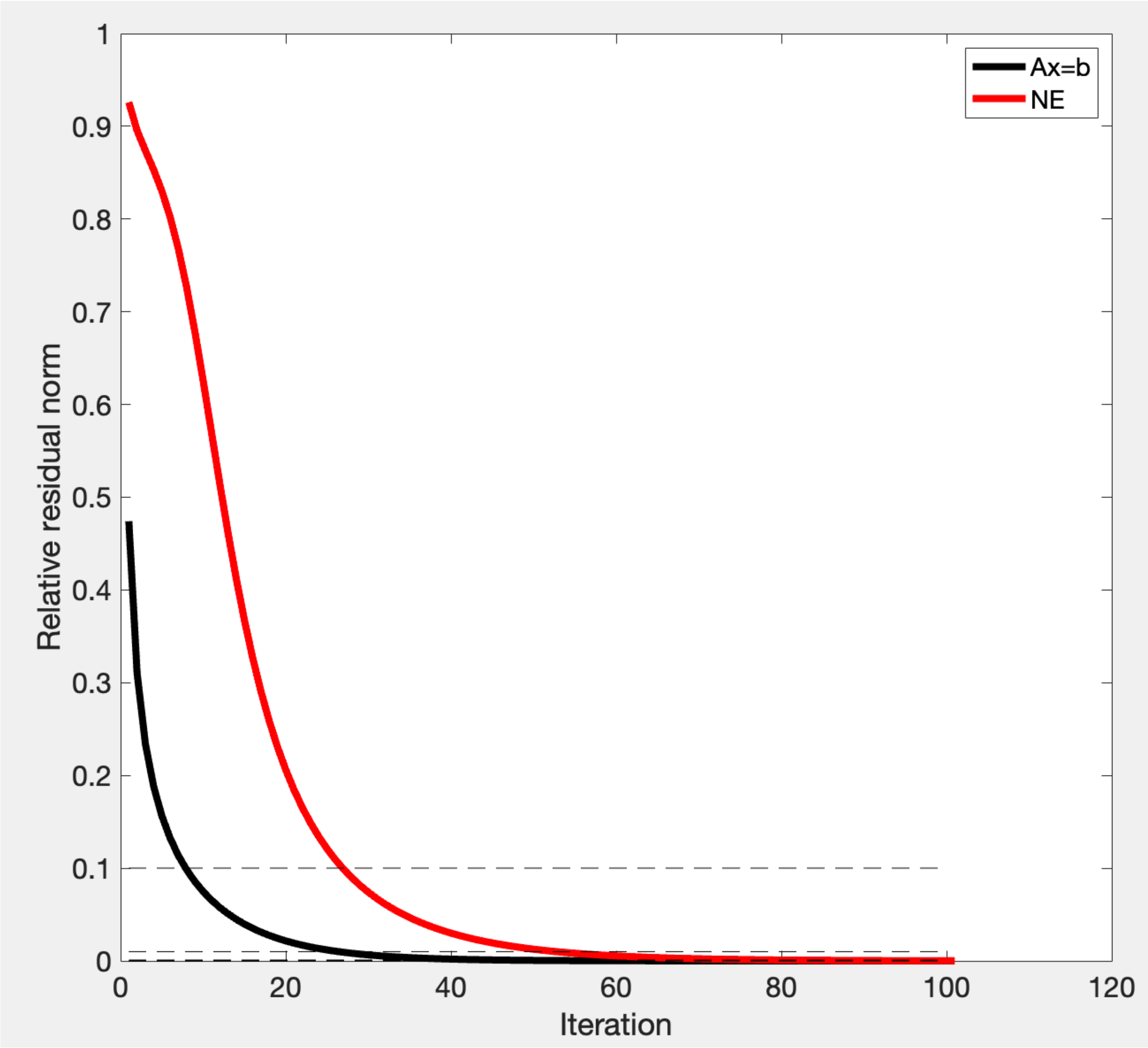




0.15
0.1
0.05
0
RRE
0
5
10
15
20
25
30
35
40
45
50
iteration
Ax=b
NE


k=1
k=2
k=3
k=6
k=10
k=20

Example 4: Space-fractional diffusion problem (Normal equation)



- Since A^TA is symmetric, its reduced Heisenberg form is even simpler:

$$AQ_k = Q_{k+1}\tilde{T}_k$$

$$\tilde{T}_k = \begin{bmatrix} \alpha_1 & \beta_1 & & \\ \beta_1 & \alpha_2 & \ddots & \\ & \ddots & \ddots & \beta_{k-1} \\ & & \beta_{k-1} & \alpha_k \\ & & & \beta_k \end{bmatrix} \in \mathbb{R}^{k\times k}$$

- $\rightarrow$Lanczos iteration

Assignment 3

Write in Matlab an algorithm for GMRES method, with QR factorization as shown on Slide 12.

Save it as GMRES_qr.m

Compress it as a zip file.

Upload it on Blackboard by

11:59 pm, Tuesday, 12/03/2024

To check if your algorithm works:

1. Download assignment.zip on Blackboard.
2. Unzip it under the same folder where you save your codes for GMRES methods in Matlab.
3. Run imageblurring.m
4. Run your own code for solving $\arg \min_{x \in \mathcal{K}_k(A,b)} \|Ax - b\|$ with the following settings:
   1) Maximum number of iteration 50
   2) Tolerance for relative residual norm 0.03
5. Plot a figure for relative residual norm similar to the one on Page 16 of this file. (Code is provided in the zip file.)
6. Plot the restored image when your algorithm is terminated. (Code is provided in the zip file.)