# LINMA2380 Matrix Computation

Homework 2
Group 3

**Okyay Kerem**
**Rochet Robin**
**Tahiri Ghali**
**Quesnel Emeric**

# 1 Exercise A: Krylov subspaces

## 1.1 A1

By Definition 1 and the assumption from A1 we can write :

$$\mathcal{K}_{r+1}(A,b) \subseteq \mathcal{K}_r(A,b)$$

$$span\{b, Ab, ..., A^r b\} \subseteq span\{b, Ab, ..., A^{r-1}b\}$$

As $K_{r+1}$ is included in $\mathcal{K}_r$ we can express $A^r B$ as a linear combination of $b, Ab, ..., A^{r-1}b$. Then we can write that since the dimension of $\mathcal{K}_{r+1}(A,b)$ is equal to the dimension of $\mathcal{K}_r(A,b)$ :

$$\mathcal{K}_{r+1}(A,b) = \mathcal{K}_r(A,b)$$

Knowing this, and by induction the same reasoning can be done with $r+2, r+3, ...$ and for any $s \geq r$ and therefore :

$$\mathcal{K}_s(A,b) = \mathcal{K}_r(A,b)$$

## 1.2 A2

Induction:

- For $r = 1$ we have $\dim(\mathcal{K}_1(A,b)) = 1$ (since $b \neq 0$)
- For $r = 2$ we have $\dim(\mathcal{K}_1(A,b)) = \dim(\{b, Ab\}) = 2$

Since, the number of independent columns of $\mathcal{K}_n(A,b)$ is $s$, we will add a dimension for each $r$ going from 1 to $s$. And from (A1) we saw that after this limit of Krylov subspace order $s$, Krylov subspace is invariant.

# 2 Exercise B: Arnoldi's iteration

## 2.1 B1

Each column $i$ of $K_s(A,b)$ can be considered as a vector $x_i$. We know that $K_s(A,b)$ is a full rank matrix. Then, the set $X = \{x_1, ..., x_s\}$ forms a basis for $Im(K_s(A,b))$.
By the Graham-Schmidt procedure shown in Theorem 2.7 of the lecture notes, we can construct the columns $y_j$ of a matrix $Y$ such that $A = YC$, where :

$$\begin{aligned}
&c_{p,p} = 1 \\
&c_{j,p} = \frac{< x_p, y_j >}{< y_j, y_j >} \quad \forall j < p \quad \text{(where } y_j \text{ is the } j^{\text{th}} \text{ column of } y\text{)} \\
&c_{j,p} = 0 \quad \forall j > p
\end{aligned}$$

$C$ is an upper triangle matrix with only ones on the main diagonal.
Let's $N = diag\{n_1, ..., n_s\},\ \ n_i = ||y_i||$. We get that $A = (YN^{-1})(NC) = Q_s R_s$.
Then, $R = NC$ which is an upper diagonal matrix with a positive diagonal.

## 2.2 B2

$$\mathcal{K}_{r+1}(A,b) = [b, AK_r(A,b)] = Q_{r+1}R_{r+1}$$

$$Q_{r+1}[1:r]R_{r+1}[1:r, 1:r] = \begin{bmatrix} \vdots & \vdots & \cdots & \vdots \\ q_1 & q_2 & \cdots & q_r \\ \vdots & \vdots & \cdots & \vdots \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1r} \\ r_{21} & r_{22} & \cdots & r_{2r} \\ \cdots & \vdots & \ddots & \vdots \\ r_{r1} & r_{r2} & \cdots & r_{rr} \end{bmatrix} = \mathcal{K}_r(A,b)$$

Using the hint:

$$A\mathcal{K}_r(A,b) = [Q_{r+1}R_{r+1}][\text{2:r+1}]$$

The last $r$ columns of $\mathcal{K}_{r+1}(A,b)$ is equal to $A\mathcal{K}_r(A,b)$. We can also write it as:

$$= Q_{r+1}[1:r+1]R_{r+1}[1:r+1, 2:r+1]$$

Thus it holds that:

$$AQ_{r+1}[1:r]R_{r+1}[1:r,1:r] = Q_{r+1}[1:r+1]R_{r+1}[1:r+1,2:r+1]$$

And more generally we can replace $r+1$ with $s$.

## 2.3 B3

From question (B2):

$$AQ_s[1:r]R_s[1:r,1:r] = Q_s[1:r']R_s[1:r',2:r']$$

Can be written as:

$$AQ_s[1:r] = Q_s[1:r']R_s[1:r',2:r']R_s^{-1}[1:r,1:r]$$

Now looking at the structure of $R_s[1:r',2:r']$:

$$R_s[1:r',2:r'] = \begin{bmatrix} r_{12} & r_{13} & \cdots & r_{1r'} \\ r_{22} & r_{23} & \cdots & r_{2r'} \\ 0 & r_{33} & \cdots & r_{3r'} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & r_{r'r'} \end{bmatrix}$$

Which is a upper hessenberg matrix. Since the inverse of the upper triangular is another upper triangular, we can see that the product of $R_s[1:r',2:r']R_s^{-1}[1:r,1:r]$ is going to have zeros for $i > j+1$. If we show the product as follows,

$$\begin{bmatrix} r_{12} & r_{13} & \cdots & r_{1r'} \\ r_{22} & r_{23} & \cdots & r_{2r'} \\ 0 & r_{33} & \cdots & r_{3r'} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & r_{r'r'} \end{bmatrix} \begin{bmatrix} r'_{11} & r'_{12} & \cdots & r'_{1r} \\ 0 & r'_{22} & \cdots & r'_{2r} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & r'_{rr} \end{bmatrix}$$

We can see that the result will be an upper hessenberg matrix. Thus the statement holds.

## 2.4 B4

**Inductive proof:**
For $r=1$,

$$q_1 = \frac{b}{||b||}$$

For $r=2$,

$$q_1 = \frac{b}{||b||}$$

and $q_2$ calculated as follows:

$$w = Aq_1 - \frac{< q_1, Aq_1 >}{||q_1||^2} q_1$$

Assuming $w \neq 0$

$$q_2 = \frac{w}{||w||}$$

So we can observe that:

$$Aq_1 = \frac{Ab}{||b||} \in \mathcal{K}_2(A,b) = \text{span}\{b, Ab\}$$

Take $r+1$ (e.g, r=2), the orthonormal vector $q_{r+1}$ will be generated as follows

$$w_{r+1} = Aq_r - \sum_{i=1}^{r} c_i q_i$$

For scalars $c_i$ coming from the projections. Assuming $w_{r+1} \neq 0$

$$q_{r+1} = \frac{w}{||w||}$$

We can observe again:

$$Aq_r \in \mathcal{K}_{r+1}(A,b)$$

Since $Aq_r$ is a linear combination of $b, Ab, \ldots, A^r b$, and the subtraction terms $c_i q_i$ are in $\text{span}\{q_1, \ldots, q_r\} \subseteq \mathcal{K}_r(A,b)$.
So we can conclude $\text{span}\{q_1, \ldots, q_{r+1}\} \subseteq \mathcal{K}_{r+1}(A,b)$.

## 2.5 B5

In question (B4), when we were looking at $r = 2$ we saw that $q_1 = \frac{b}{||b||}$ and in the calculation for $q_2$ we first calculated $w$.

$$w = Aq_1 - \frac{< q_1, Aq_1 >}{||q_1||^2} q_1$$

We can see that $Aq_1$ is $\frac{Ab}{||b||}$, and the subtracted term is a scalar times the vector $q_1$. So from here we can observe that $\text{span}\{q_1, q_2\} = \text{span}\{b, Ab\} = \mathcal{K}_2(A,b)$. And generalizing in the same way, we can write this for $\text{span}\{q_1, \ldots, q_r\} = \text{span}\{b, Ab, \ldots, A^{r-1}b\} = \mathcal{K}_r(A,b)$

## 2.6 B6

$$v = Aq_s$$

$$v' = \text{Proj}_{\{q_1, \ldots, q_s\}}(v) = \sum_{i=1}^{s} \frac{< Aq_s, q_i >}{||q_i||^2} q_i$$

$$w = v - v'$$

Observing, $v \in \text{span}\{b, Ab, \ldots, A^s b\}$, $q_s \in \text{span}\{b, Ab, \ldots, A^{s-1}b\}$
We can conclude that if $w = 0$, we can represent $v$ as a linear combination of $q_1, \ldots, q_s$. This means that $A^s b \in \text{span}\{b, Ab, \ldots, A^{s-1}b\}$. Therefore, when the algorithm terminates, we get $s = \dim(\mathcal{K}_n(A,b))$.

## 2.7 B7

From (B5), $\text{span}\{q_1, \ldots, q_s\} = \mathcal{K}_s(A,b) = \text{span}\{b, Ab, \ldots, A^{s-1}b\}$. We can see that $\mathcal{K}_s(A,b)$ is a full-rank matrix, so from Theorem 2.8 we can factorize it to $\mathcal{K}_s(A,b) = QR$.
To see the parallel between algorithm 1 and Gram-Schmidt procedure. Consider $\text{span}\{b, Ab, \ldots, A^{s-1}b\}$ as the input to Theorem 2.7.

$$y_1 = b$$

$$y_p = A^{p-1}b - \sum_{i=1}^{p-1} \frac{< A^{p-1}b, y_i >}{< y_i, y_i >} y_i \qquad p = 2, \ldots, s$$

Then, setting

$$q_i = \frac{y_i}{||y_i||}$$

We can observe it is the same procedure as the algorithm 1. So, $\mathcal{K}_s(A,b) = Q_s R_s$ for some $R_s$ upper triangular matrix.

# 3 Exercise C: GMRES for linear system solution approximation

## 3.1 C1

Since $Q$ forms a basis for the Krylov subspace, any vector $x \in \mathcal{K}_r(A, b)$ can be written as a linear combination of columns.

$$x = Qy$$

For some vector $y \in \mathbb{R}^r$.

$$\min_{y \in \mathbb{R}^r} ||AQy - b||$$

Using

$$AQ = QH + \beta q e_{r,r}^\top$$

$$= \min_{y \in \mathbb{R}^r} ||QHy + \beta q e_{r,r}^\top y - b||$$

$$= \min_{y \in \mathbb{R}^r} ||\ [Q, q] \begin{bmatrix} H \\ \beta e_{r,r}^\top \end{bmatrix} y - b\ ||$$

Also, since $q_1 = \frac{b}{||b||}$ we can write

$$b = ||b||\ [Q, q]\ e_{r+1,1}$$

Pluggin in,

$$= \min_{y \in \mathbb{R}^r} ||\ [Q, q] \begin{bmatrix} H \\ \beta e_{r,r}^\top \end{bmatrix} y - ||b||\ [Q, q]\ e_{r+1,1}\ ||$$

$[Q, q]$ is an isometry, which means it preserves norm. So we can remove it from the optimization problem.

$$= \min_{y \in \mathbb{R}^r} ||\ \begin{bmatrix} H \\ \beta e_{r,r}^\top \end{bmatrix} y - ||b||\ e_{r+1,1}\ ||$$

Thus, if $y^*$ minimizes the RHS, since we set $x = Qy$, $Qy^*$ will minimize the LHS.

## 3.2 C2

Solving for (3) is easier due to following reasons:

- Possible that $b \notin \text{Im}(A)$, this makes $Ax = b$ not solvable since it requires an exact solution.
- $A$ needs to be square invertible for $Ax = b$
  - This is not always the case in real life. In overdetermined systems you have more equations than unkowns.
- Robust algorithms for solving (3), least squares is more flexible and handles numerical issues better. Thus, easier in practice.

$$\min_{y \in \mathbb{R}^r} ||\ \tilde{H}y - ||b||\ e_{r+1,1}\ ||$$

Minimized when,

$$\tilde{H}^\top \tilde{H} y = \tilde{H}^\top ||b|| e_{r+1,1}$$

Decomposing $\tilde{H} = QR$

$$(QR)^\top (QR) y = \tilde{H} ||b|| e_{r+1,1}$$

$$R^T R y = \tilde{H} ||b|| e_{r+1,1}$$

Cost of QR decomposition using Householder QR decomposition in section 2.4:

$$2(r + 1) r^2 = 2r^3 + 2r^2$$

Cost of $R^\top R$

$$r * r * (2r-1) = 2r^3 - r^2$$

Cost of $\tilde{H}||b||e_{r+1,1}$

$$r * 1 * (2r+1) = 2r^2 + r$$

Cost of solving $r$ equations with $r$ unknowns:

$$\approx r^3$$

Therefore,

$$\text{flops } = 5r^3 + 3r^2 + r$$

### 3.3 C3

Since $x \in \mathcal{K}_r(A,b)$, we can say that $x = A^n b$ with $n \in [0, r-1]$.Thus,

$$\begin{aligned} b - Ax &= b - AA^n b \\ &= b - A^{n+1} b \\ &= (I - A^{n+1})b \quad \text{(I is the identity matrix of size } n \times n) \end{aligned}$$

Let's now prove that $(I - A^{n+1})$ is a polynomial $p(A)$ with $p \in \mathcal{P}_r^0$. First, we will show that $p(0) = 1$. In our case, we are working with matrices so the equivalent of 1 in scalar is the identity matrix. For $A = 0$, we have :

$$\begin{aligned} p(0) &= I - 0 \\ &= I \end{aligned}$$

We have in fact that $p(0) = I$. Let's see now if $p$ is in fact in the set of polynomials of degree at most $r$. We have $p(A) = I - A^{n+1}$. Previously, we said that $n \in [0, r-1]$. Thus, as we are taking the $(n+1)^{th}$ power of $A$, we can consider that we have a polynomial of degree at most $r$ $(r - 1 + 1 = r)$ considering that the identity matrix will not "change" the degree of the matrix $A$.
We proved that $b - Ax = p(A)b$. Moreover, thanks to the invariant rule of the norm, we have that

$$||Ax - b|| = || - (Ax - b)|| = ||b - Ax||$$

Thus,

$$\min_{x \in K_r(A,b)} ||Ax - b|| = \min_{x \in K_r(A,b)} ||b - Ax|| = \min_{p \in \mathcal{P}_r^0} ||p(A)b||$$

### 3.4 C4

Since $A$ is symmetric it has an eigenvalue decomposition $A = U\Lambda U^T$, where $U$ is an orthonormal matrix (i.e. $UU^T = I$) and $\Lambda$ is a diagonal matrix containing the eigenvalues of A. The spectral norm of $A$ is defined as :

$$\|A\| = \sup_{x=1} \|Ax\|$$

Using the eigenvalue decomposition we can write $Ax = U\Lambda U^T x$ and since $U^T$ is orthogonal it preserves the norm, so :

$$\|Ax\| = \|\Lambda U^T x\|$$

Let $y = U^T x$, then $\|y\| = \|x\| = 1$ becaus $U^T$ is orthogonal therefore,

$$\|Ax\| = \|\Lambda y\| = \sqrt{\sum_{i=1}^{n} \lambda_i^2 y_i^2} \leq \sqrt{\sum_{i=1}^{n} \max_{1 \leq i \leq n} \lambda_i^2 y_i^2} = \sqrt{\max_{1 \leq i \leq n} \lambda_i^2 \sum_{i=1}^{n} y_i^2} = \sqrt{\max_{1 \leq i \leq n} \lambda_i^2 \|y\|^2} = \max_{1 \leq i \leq n} |\lambda_i| \cdot \|y\|$$

Thus we have :

$$\|Ax\| = \|\Lambda y\| \leq \max_{1 \leq i \leq n} |\lambda_i| \cdot \|y\|$$

Because $\Lambda$ is a diagonal matrix. To show that the spectral norm is exactly $\max_{1\leq i\leq n}|\lambda_i|$ consider an eigenvector $\nu_i$ corresponding to the eigenvalue $\lambda_i$ with the largest absolute value. then,

$$A\nu_i = \lambda_i\nu_i \text{ and thus } \|A\nu_i\| = |\lambda_i|\cdot\|\nu_i\| = |\lambda_i| \text{ since } \|\nu_i\| = 1$$

Since we have shown that $\|Ax\| \leq \max_{1\leq i\leq n}|\lambda_i|$, then for all unit vectors $x$, and there exists a specific vector $\nu_i$ for which

$$\|A\nu_i\| = \max_{1\leq i\leq n}|\lambda_i|$$

we can conclude that

$$\|A\| = \max_{1\leq i\leq n}|\lambda_i|$$

### 3.5 C5

As $A$ is symmetric, we can do its eigenvalue decomposition. Thus,

$$A = U\Lambda U^T$$

where $U \in \mathbb{R}^{n\times n}$ is an orthogonal matrix (i.e. $U^TU = I$), and $\Lambda$ is a diagonal matrix which contains the eigenvalues of $A$. We saw in C3 that

$$\min_{x\in K_r(A,b)}||Ax-b|| = \min_{p\in\mathcal{P}_r^0}||p(A)b||$$

Then,

$$\begin{aligned}
\min_{x\in K_r(A,b)}||Ax-b|| &= \min_{p\in\mathcal{P}_r^0}||p(A)b|| \\
&= \min_{p\in\mathcal{P}_r^0}||Up(\Lambda)U^Tb|| \\
&= \min_{p\in\mathcal{P}_r^0}||p(\Lambda)U^Tb|| \quad \text{(as } U \text{ is orthogonal)} \\
&\leq ||U^Tb|| \min_{p\in\mathcal{P}_r^0}\max_{1\leq i\leq n}|p(\lambda_i)| \quad \text{(by Cauchy-Schwartz)} \\
&= ||b|| \min_{p\in\mathcal{P}_r^0}\max_{1\leq i\leq n}|p(\lambda_i)| \quad \text{(as } U^T \text{ is orthogonal)}
\end{aligned}$$

In conclusion, we have

$$\min_{x\in K_r(A,b)}||Ax-b|| \leq ||b|| \min_{p\in\mathcal{P}_r^0}\max_{1\leq i\leq n}|p(\lambda_i)|$$

## 4 Exercise D: Arnoldi's method for eigenvalue approximation

### 4.1 D1

In the expression $||Ax-\lambda x||$ we can substitute $x = Qy$ :

$$||AQy - \lambda Qy|| = |\beta e_r^T y|$$

Then by using the equality $AQ = QH + \beta q e_r^T$ :

$$\begin{aligned}
||QHy + \beta q e_r^T y - \lambda Qy|| &= |\beta e_r^T y| \\
||QHy - \lambda Qy + \beta q e_r^T y|| &= |\beta e_r^T y| \\
||Q(Hy - \lambda y) + \beta q e_r^T y|| &= |\beta e_r^T y|
\end{aligned}$$

As $(\lambda, y)$ is an eigenpair of H, the term $Q(Hy-\lambda y) = 0$ and we end up with :

$$||\beta q e_r^T y|| = |\beta e_r^T y|$$

which is valid since q is a vector from an orthonormal basis (so it does not affect the norm) and $||y|| = 1$.

# 5 Exercise E: Implementation

## 5.1 E1

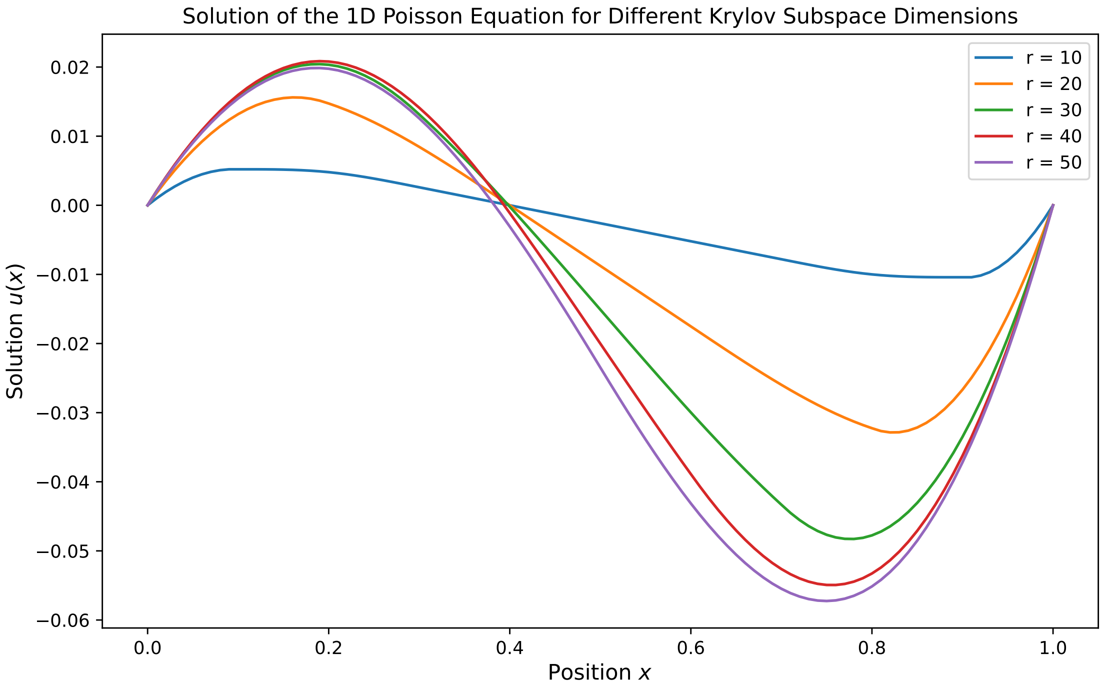


Figure 1: Solution of (3) for r = 10,20,30,40,50

## 5.2 E2

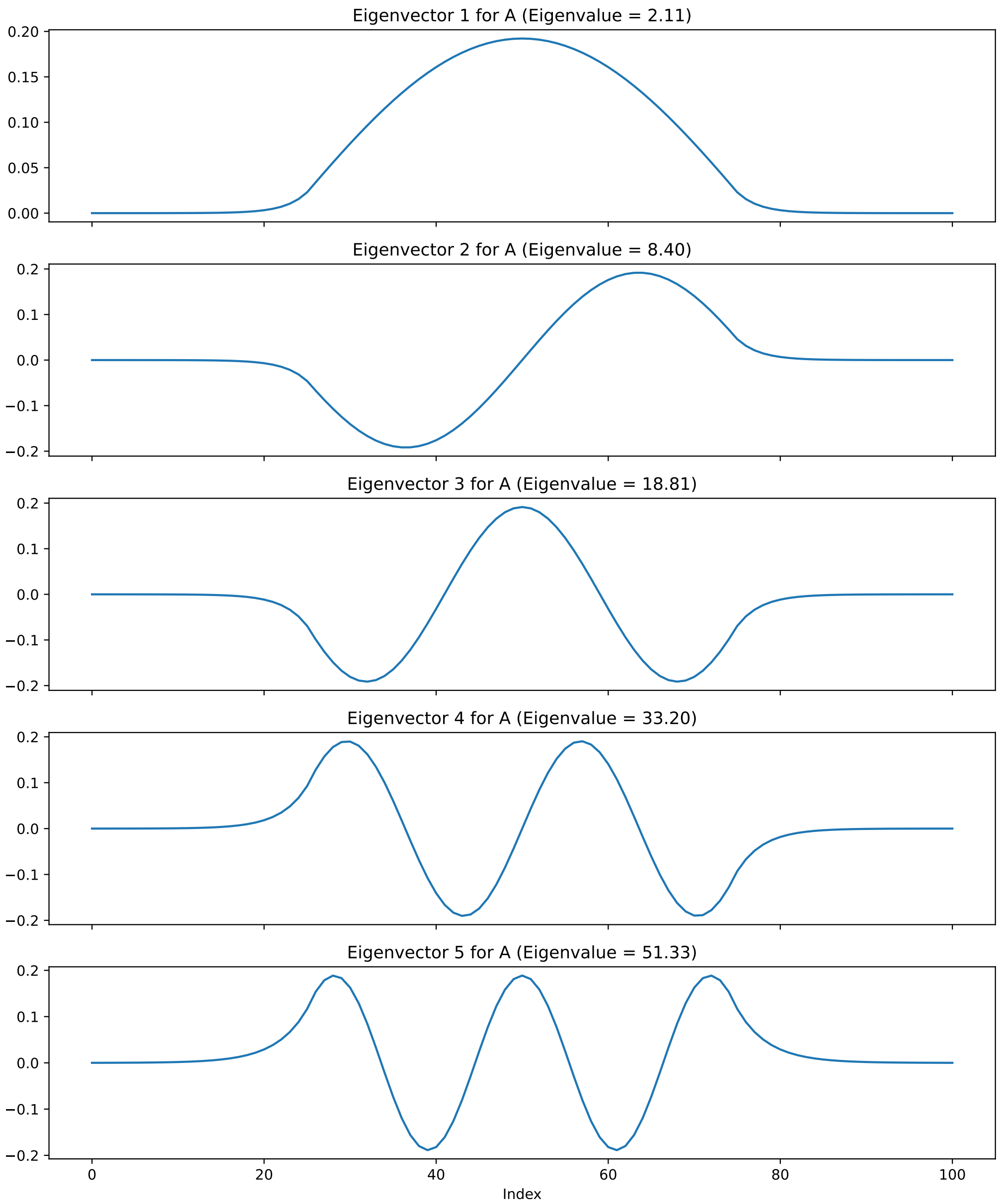


Figure 2: Plot of the actual eigenvectors associated to the smallest 5 eigenvalues of A

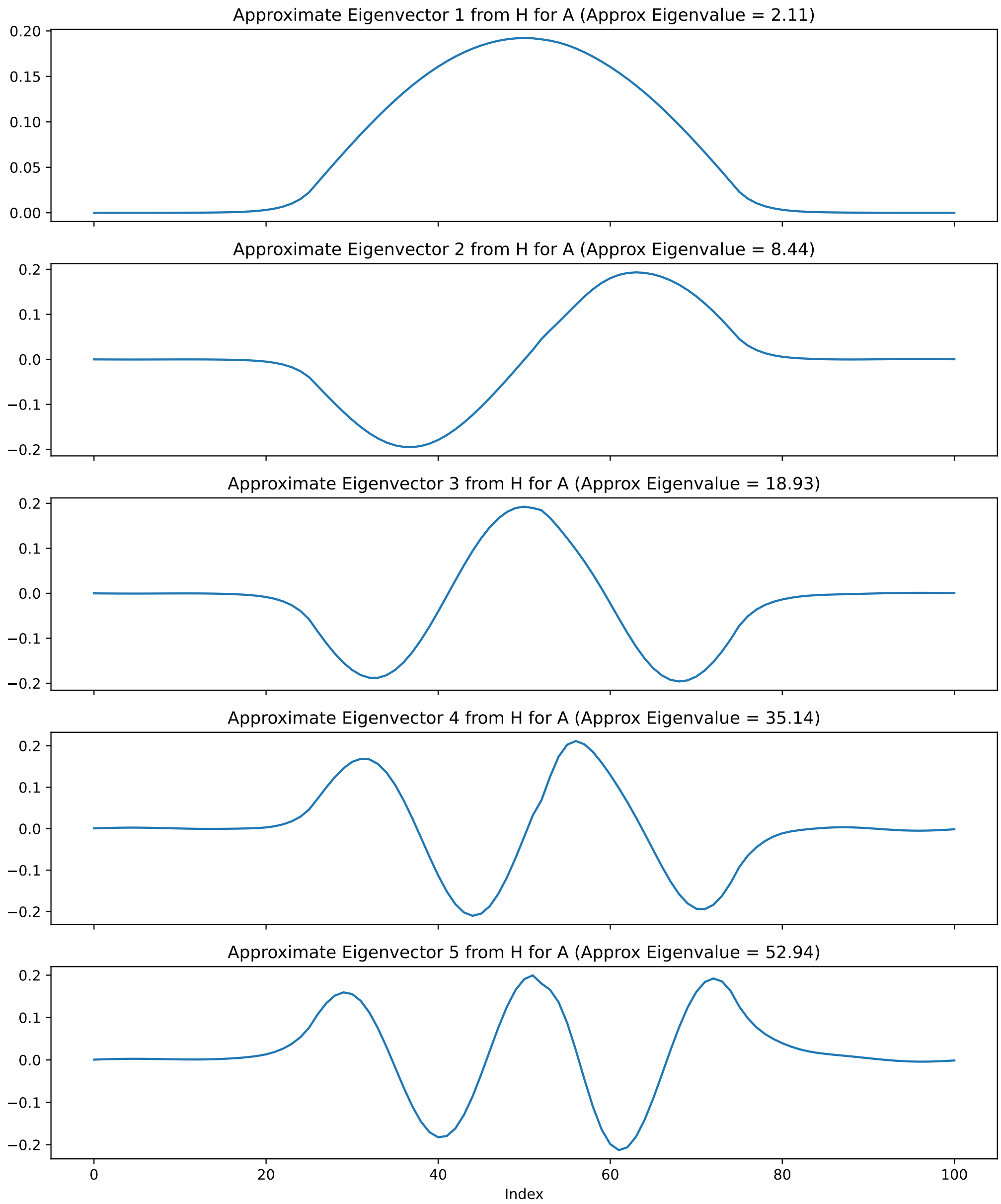


Figure 3: Plot of the approximated eigenvectors for A associated to the smallest 5 eigenvalues of H