

1 Matrix-Vector Multiplication

Coordinate Representation. If we write b as $A(A^{-1}b)$, one can see that $A^{-1}b$ is exactly the coefficient of expansion of b in the basis of A .

Geometric representation. We can view Ax as $\text{range}(A)$, i.e. subspace spanned by column vectors a_1, a_2, \dots, a_n . If vector b is perpendicular to this subspace, one has $A^T b = 0$, which corresponds to $\langle a_i, b \rangle = a_i^T b = 0$.

Problem 1

By considering what space is spanned by the first n columns of R and using (1.8), show that if R is a nonsingular $m \times m$ upper-triangular matrix, then R^{-1} is also upper-triangular. (The analogous result also holds for lower-triangular matrices.)

Proof. This is a simple extension of solving upper triangle system

$$Rx = b$$

and when we consider appropriate b 's, saying

$$R[x_1, x_2, \dots, x_n] = [\delta_1, \delta_2, \dots, \delta_n]$$

One can see $x_{ij} = 0$ when $i < j$ since the first i columns of R is spanned by $[\delta_1, \dots, \delta_i]$. \square

One of the most important transformations is projection and let's start with simple example-projection of vector v to unit q . Assume we move both vectors at origin and the projection length is the inner product of two vectors q^*v (minus sign means the angle between them is greater than $\pi/2$). Hence the projection is $(q^*v)q$, which can be also rewritten as $(qq^*)v$ through the lens of transformation qq^* .

Think what remains about v after subtracting the projection part $(q^*v)q$? If we decompose the Euclidean space into $H \oplus q$, then it's clear that $v - (q^*v)q$ falls into the subspace H which can be seen as a hyperplane orthogonal to q . Moreover, one can see that $v - 2(q^*v)q$ is exactly the image of v in the **mirror** H .

An isometric transformation from v to w originates from **reflection** through the hyperplane H which is orthogonal to $v - w$!

Problem 2

Suppose a, b are column vectors in \mathbb{R}^n and they are orthogonal, please investigate $M := ab^T + ba^T$.

Geometrically, it can be interpreted as a reflection within the plane spanned by a and b , combined with a scaling.

Let's consider the action of M on a vector $x \in \mathbb{R}^n$: $Mx = ab^T x + ba^T x$. The term $ab^T x$ projects x onto b (along the direction of b), then becomes a scalar to dilate vector a .

Because a and b are orthogonal, these projections are independent. The transformation effectively exchanges the components of x along a and b . If $\|a\| = \|b\|$, the transformation is a pure reflection in the hyperplane orthogonal to both a and b . If $\|a\| \neq \|b\|$, a scaling is also involved.

Example 1.1. 1. Let $a = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ and $b = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$. Then

$$ab^T + ba^T = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Applying this to $x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ yields $\begin{bmatrix} x_2 \\ x_1 \\ x_3 \end{bmatrix}$, which swaps x_1 and x_2 and leaves x_3 unchanged.

2. Let $a = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}$ and $b = \begin{bmatrix} 1 \\ -1 \\ 1 \\ 1 \end{bmatrix}$. Note that a and b are orthogonal, as their dot product is $a \cdot b = 0$.

We compute the matrix $M = ab^T + ba^T$,

$$M = ab^T + ba^T = \begin{bmatrix} 2 & 0 & 1 & 1 \\ 0 & -2 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{bmatrix}$$

One can show that M is of rank 2.

2 Orthogonal Vectors and Matrices

Definition 2.1 (Isometric matrix). Let $n, k \in \mathbb{N}$, and let $Q \in \mathbb{F}^{n \times k}$. If $Q^*Q = I$ holds, Q is called *isometric*. A square isometric matrix is called **unitary**.

Remark 2.2. Given isometric Q , one cannot have $Q^*Q = QQ^*$ and QQ^* is just a orthogonal projection in \mathbb{R}^n . For example,

$$Q = \begin{bmatrix} 0 & 1 \\ 1 & 0 \\ 0 & 0 \end{bmatrix}$$

We will use this technique in invariant subspace by simultaneous iterations.

Let me give a general example about **orthogonal projection**. Given

$$Q = \begin{bmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & 0 \end{bmatrix}$$

One can check $Q^T Q = I_2$ which justifies the isometric of Q and

$$QQ^T = \begin{bmatrix} 5/6 & -1/6 & 1/3 \\ -1/6 & 5/6 & 1/3 \\ 1/3 & 1/3 & 1/3 \end{bmatrix}$$

which is usually denoted by P .

One can see the action of QQ^T to v by

$$\begin{bmatrix} q_1 & q_2 \end{bmatrix} \begin{bmatrix} q_1^* \\ q_2^* \end{bmatrix} v = \begin{bmatrix} q_1 & q_2 \end{bmatrix} \begin{bmatrix} q_1^* v \\ q_2^* v \end{bmatrix} = C_1 q_1 + C_2 q_2 \quad (1)$$

which means QQ^T sends v to the subspace spanned by q_1 and q_2 . Of course, one can guess that it is similar to

$$\begin{bmatrix} 1 & & \\ & 1 & \\ & & 0 \end{bmatrix}$$

because we have $Pq_i = q_i, i = 1, 2$, i.e. two eigenvectors with eigenvalue 1. It's easy to use eq(1) to check that QQ^T has the diagonal form under basis

$$\begin{bmatrix} q_1 & q_2 & q_3 \end{bmatrix}$$

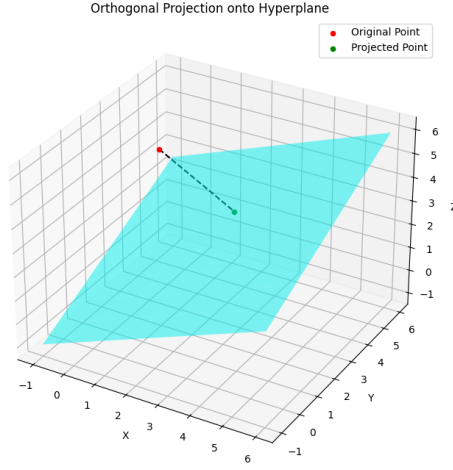


Figure 1: QQ^T projects vector v to the hyperplane generated by the column vector of Q , i.e. $q_1 = [1/\sqrt{3}, 1/\sqrt{3}, 1/\sqrt{3}]$ and $q_2 = [1/\sqrt{2}, -1/\sqrt{2}, 0]$. For instance, point $[1 \ 2 \ 6]^T$ in red is projected to $[1.5 \ 2.5 \ 3]^T$

where q_3 is the extended normalized vector. The illustration is given in the following figure.

We need prepare some lemmas to be at bottom of simultaneous iteration, assume V is matrix consisting of linearly independent columns $[v_1, v_2, \dots, v_k]$ and we have the following proposition

- Lemma 2.3.**
1. V is injective is equivalent to \widehat{V} is injective.
 2. PV is injective is equivalent to $\widehat{P}\widehat{V}$ is injective.
 3. PV is injective implies that V is injective.

Proof. The first can be easily proven by $v_i = QQ^*v_i = Q\widehat{v}_i$.

We can view PV as new matrix consisting of w_i and find that $\widehat{PV} = \widehat{P}\widehat{V}$. □

3 Norms

Since the norm of vector x induces a metric space, one can see that the closed unit ball

$$\{x \in \mathbb{R}^2 : \|x\| = 1\}$$

are different under different norm. Since the norm of matrix is defined by vector norm, it's worthwhile to give an illustration.

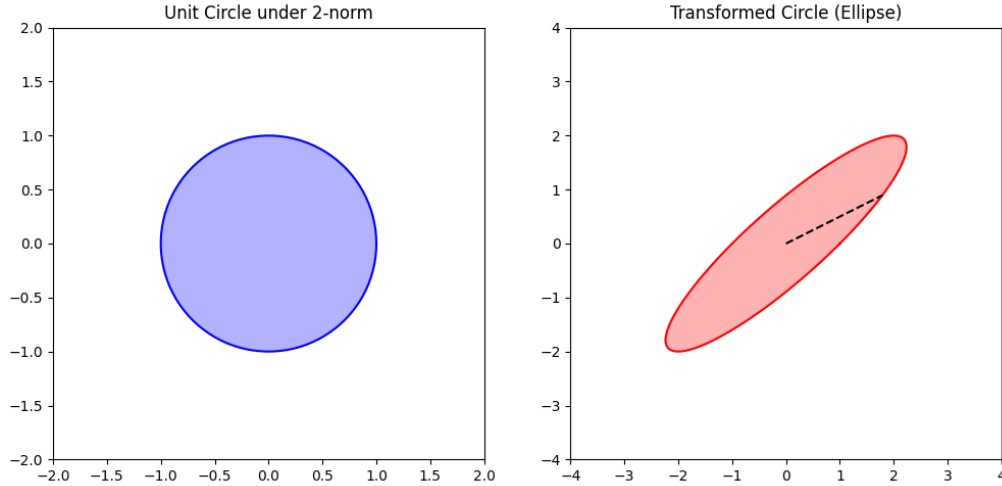


Figure 2: The black dashed line is eigenvector of transformed matrix $\begin{bmatrix} 1 & 2 \\ 0 & 2 \end{bmatrix}$. The spectral of matrix A is strictly less than $\|A\|$ since it's non-symmetric.

Problem 3

Let $\|\cdot\|$ denote any norm on \mathbb{C}^m and also the induced matrix norm on $\mathbb{C}^{m \times m}$. Show that $\rho(A) \leq \|A\|$, where $\rho(A)$ is the spectral radius of A , i.e., the largest absolute value $|\lambda|$ of an eigenvalue λ of A .

Proof. Suppose ξ is one eigenvector corresponding to largest eigenvalue λ , then

$$x = \frac{\xi}{\|\xi\|}$$

is a unit vector under $(\mathbb{C}^m, \|\cdot\|)$.

The conclusion comes from

$$\|Ax\| = \|\lambda x\| = |\lambda| \|x\| = |\lambda| = \rho$$

and $\|Ax\| \leq \|A\| \|x\| = \|A\|$.

□

Remark 3.1. Next, we show that

$$\rho(A) = \{\langle Ax, x \rangle, \|x\| = 1\} = \|A\|$$

when A is symmetric.

By assumption, there is a unitary matrix Q such that Q^*AQ is a diagonal matrix with real number entry. Then

$$\langle Ax, x \rangle = \langle QDQ^*x, x \rangle = \langle D\hat{x}, \hat{x} \rangle$$

where \hat{x} is the coordinates of x under basis of Q . It's easy to see that $\lambda_{\min} \leq \langle Ax, x \rangle \leq \lambda_{\max}$. (Here we used the conclusion about unitary matrix $\|Qx\| = \|x\|$.)

Next, we prove second equality, thanks to

$$\langle Ax, Ax \rangle = \langle A^*Ax, x \rangle = \langle A^2x, x \rangle$$

and the previous, we know that the maximum of $\langle Ax, Ax \rangle$ is λ_{\max}^2 .

Problem 4

Show that $\|A\| \leq \|A\|_F$.

Proof. This question reveals the importance of SYMMETRIC matrix.

Suppose A is symmetric, then by $A = QDQ^T$, one can see that $\|A\| = \lambda_{\max} < \|QDQ^T\|_F$.

For general case, we have $\langle Ax, Ax \rangle = \langle A^T Ax, x \rangle$ where there is a symmetric matrix and one can deduce the conclusion easily once he notes that sum of diagonal elements of $A^T A$ is exactly the Frobenius norm of matrix A . \square

4 Machine error and Roundoff

Double precision.

Example 4.1. Take $x = 9.4$ as an example, one can see it's stored in the following binary string

$$+1.\boxed{00101100110011001100110011001100110011001100110011001101} \times 2^3. \quad (2)$$

In computer arithmetic, the real number x is replaced with the string of bits $\text{fl}(x)$. According to this definition, $\text{fl}(9.4)$ is the number in the binary representation (2). We arrived at the floating point representation by discarding the infinite tail $\overline{.1100} \times 2^{-52} \times 2^3 = \overline{.0110} \times 2^{-51} \times 2^3 = .4 \times 2^{-48}$ from the right end of the number and then adding $2^{-52} \times 2^3 = 2^{-49}$ in the rounding step. Therefore,

$$\begin{aligned} \text{fl}(9.4) &= 9.4 + 2^{-49} - 0.4 \times 2^{-48} \\ &= 9.4 + (1 - 0.8)2^{-49} \\ &= 9.4 + 0.2 \times 2^{-49}. \end{aligned}$$

In other words, a computer using double precision representation and the Rounding to Nearest Rule makes an error of 0.2×2^{-49} when storing 9.4. We call 0.2×2^{-49} the rounding error.

The important message is that the floating point number representing 9.4 is not equal to 9.4, although it is very close. To quantify that closeness, we use the standard definition of error.

Definition 4.2. Let x_c be computed version of the exact quantity x , then we call

$$\frac{|x - x_c|}{x}$$

the **relative error**.

Example 4.3. Continuing the example about 9.4, one can see

$$\frac{|\text{fl}(9.4) - 9.4|}{|9.4|} = \frac{0.2 \times 2^{-49}}{9.4} = \frac{8}{47} \times 2^{-52} < \frac{1}{2}\mathbf{u}.$$

where \mathbf{u} is machine epsilon.

More general, one has

$$\text{fl}(a \circ b) = (a \circ b)(1 + \delta), \quad \text{where } |\delta| < \mathbf{u}.$$

where $a \circ b$ is the operation between two numbers like $+$, $-$, \times , $/$.

Backward Stability. Algorithms in numerical linear algebra are usually classified from the point of view of the numerical stability as forward and backward stable. Say, that for some data x you want to compute the solution y given by some function $f : y = f(x)$. Because our world (and computers as well) is not ideal, we get some \tilde{y} instead of y . We distinguish here between the forward error, that is, some norm of $y - \tilde{y}$ and backward error.

The backward error answers the question of how much we have to perturb the input data x for which we actually solved the problem, that is, we look for \tilde{x} such that $\tilde{y} = f(\tilde{x})$. This problem is not usually uniquely solvable (there might be many \tilde{x} satisfying this condition), so one usually seeks \tilde{x} which is close to the original data x .

5 LU factorization

Problem 5

This is talking about backward stability in solving

$$Ly = b$$

where L is $n \times n$ upper triangle matrix. Our algorithm is used to output y after inputting L and b , hence we want to find ΔL such that

$$(L + \Delta L)\tilde{y} = b \quad (3)$$

Assuming the simplest case

$$l_{11}y_1 = b_1, \quad (l_{11} + \Delta l_{11})\tilde{y}_1 = b_1$$

where \tilde{y} is computed by forward substitution using machine. Subtract the first eq from the second, we get

$$l_{11}(\tilde{y} - y) + \Delta l_{11}\tilde{y} = 0,$$

Hence

$$\frac{|\Delta l_{11}|}{|l_{11}|} = \frac{|\tilde{y} - y|}{|\tilde{y}|} \leq \mathbf{u}.$$

where the last relative error about \tilde{y} can be seen in Example 4.3.

When $i \geq 2$, things becomes perplexing because there are more errors propagating in the process. In exact arithmetic,

$$l_{ii}y_i = b_i - \sum_{j=1}^{i-1} l_{ij}y_j, \quad i = 1, 2, \dots, n.$$

The algorithm in detail is

```

1: for  $i = 1$  to  $n$ , step 1 do
2:    $w^{(1)} \leftarrow b_i$ 
3:   for  $j = 0$  to  $i - 1$ , step 1 do
4:      $w^{(j+1)} \leftarrow w^{(j)} - l_{ij}y_j$ 
5:    $y_i \leftarrow w^{(i)}/l_{ii}$ 

```

There are two sources of error,

$$\tilde{w}^{(j+1)} = (\tilde{w}^{(j)} - l_{ij}\tilde{y}_j(1 + \delta_j))(1 + \delta'_j), \quad |\delta_j| \leq \mathbf{u}, \quad |\delta'_j| \leq \mathbf{u}$$

and

$$l_{ii}\tilde{y}_i = (1 + \delta_i)\tilde{w}^{(i)}, \quad |\delta_i| \leq \mathbf{u}.$$

Putting these together gives

$$\frac{l_{ii}\tilde{y}_i}{1+\delta_i} = b_i(1+\delta'_1)\cdots(1+\delta'_{i-1}) - \sum_{j=1}^{i-1} l_{ij}\tilde{y}_j(1+\delta_j)(1+\delta'_j)\cdots(1+\delta'_{i-1})$$

and dividing by the factor in the term containing b_i leads to

$$\frac{l_{ii}\tilde{y}_i}{(1+\delta_i)(1+\delta'_1)\cdots(1+\delta'_{i-1})} = b_i - \sum_{j=1}^{i-1} l_{ij}\tilde{y}_j \frac{1+\delta_j}{(1+\delta'_1)\cdots(1+\delta'_{j-1})} \quad (4)$$

Comparing eq(4) with our goal eq(3) written as followings

$$\sum_{j \leq i} (l_{ij} + \Delta l_{ij}) \tilde{y}_j = b_i,$$

we deduce

$$\tilde{l}_{ii} = \frac{l_{ii}}{(1+\delta_i)(1+\delta'_1)\cdots(1+\delta'_{i-1})} = (1+\epsilon_{ii})l_{ii}, \quad |\epsilon_{ii}| \leq i\mathbf{u} + O(\mathbf{u}^2)$$

and

$$\tilde{l}_{ij} = \frac{l_{ij}(1+\delta_j)}{(1+\delta'_1)\cdots(1+\delta'_{j-1})} = l_{ij}(1+\epsilon_{ij}), \quad |\epsilon_{ij}| \leq j\mathbf{u} + O(\mathbf{u}^2).$$

6 QR Decomposition

In this section, we investigate QR factorization and Householder reflection. Just like LU factorization, we view QR factorization as **encoding** following process

$$AR^{-1} = Q$$

i.e. manipulating the columns of A to get orthogonal matrix.

First, we investigate the algorithm derived from Grad-Schmit

Require: $m \times n$ matrix A

```

1: for  $i = 1$  to  $n$  do
2:    $q_i \leftarrow A(1 : m; i)$ 
3:   for  $j = 1$  to  $i - 1$  do
4:      $r_{ji} \leftarrow A(1 : m; j)^T q_i$ 
5:      $q_i = q_i - r_{ji}A(1 : m; j)$ 
6:    $A(1 : m; i) \leftarrow q_i / \|q_i\|_2^2$ 
```

Example 6.1. *This is about the stability of QR factorization.*

We have seen the example like when $\delta = 10^{-10}$ and

$$A = \begin{bmatrix} 1 & 1 & 1 \\ \delta & 0 & 0 \\ 0 & \delta & 0 \\ 0 & 0 & \delta \end{bmatrix}.$$

One have to understand that one discard ε in the process of roundoff

$$\text{fl}(1 + \delta^2) = 1,$$

while one cannot view $\sqrt{\delta^2 + \delta^2}$ as 0. That's why we call 10^{-15} as relative representation error and $[10^{-308}, 10^{308}]$ is the range of double precision.

Next, we use Householder reflector to realize QR factorization, where we use orthogonal matrix Q to put zeros in $a_{ij}, i > j$, in other words, we view $A = QR$ as

$$Q^{-1}A = Q^{-1}[\alpha_1 \mid \alpha_2 \mid \cdots \mid \alpha_n] = R.$$

Therefore we need find an orthogonal matrix Q^{-1} such that $Q^{-1}\alpha_1 = c e_1$ and if we can do this, we can use the same pattern to finish the factorization. (induction idea is very common!)

Remark 6.2. *I realized the QR algorithm [2][Algorithm 10.1, 10.3] in Python, which is very worthwhile. I learned some points like*

- $(I - 2qq^*)v$ can be executed as $v - 2q(q^*v)$ which avoids multiplication between matrix qq^* and vector v .
- Count computational cost: first find the predominant part and analyze the flops per capita!
- Implicit calculation of a product Qx .

7 Eigenvalue and Eigenvector

Here I state a result which will be used again and again in the following text.

Problem 6

Suppose $A, E \in \mathbb{C}^{n \times n}$ are self-adjoint with eigenvalues $\lambda_1 \geq \dots \geq \lambda_n, \epsilon_1 \geq \dots \geq \epsilon_n$ respectively, and $B = A + E$ has eigenvalues $\mu_1 \geq \dots \geq \mu_n$. Prove that $\lambda_i + \epsilon_1 \geq \mu_i \geq \lambda_i + \epsilon_n$

Proof. When eigenvalues of self-adjoint matrix A are arranged in descending order (from large to small), we have

$$\lambda_k = \max_{\dim(V)=k} \min_{v \in V} \langle Av, v \rangle.$$

When eigenvalues are arranged in ascending order, we have

$$\mu_k = \min_{\dim(V)=n-k+1} \max_{v \in V} \langle Av, v \rangle.$$

Taking the above statement for granted, we now prove $\mu_i \leq \lambda_i + \epsilon_1$.

$$\begin{aligned} \mu_i &= \min_{\dim(V)=n-i+1} \max_v (v^*(A + E)v) \\ &\leq \min_{\dim(V)=n-i+1} \left(\max_v v^*Av + \max_v v^*Ev \right) \\ &\leq \lambda_i + \epsilon_1. \end{aligned}$$

□

7.1 Jacobi Iteration

Problem 7

Write a program to find the eigenvalues of an $m \times m$ real symmetric matrix by the Jacobi algorithm with the standard row-wise ordering, plotting the sum of the squares of the off-diagonal entries on a log scale as a function of the number of sweeps. Apply your program to random matrices of dimensions 20, 40, and 80.

7.2 Power Method

Let me brush up the main points in [1].

Rayleigh Quotient

Assume we get a vector $v^{(k)}$ after k -step iterations, then we face the problem in solving approximate eigenvalue x , i.e.

$$xv = Av$$

Numerically, Av and v are not parallel and this boils down to solving inconsistent linear system. One can apply the conclusion of Least Square to get **Rayleigh quotient**:

$$x = \Lambda_A(v) := \frac{v^* Av}{v^* v} \quad (5)$$

Also one can get the same result(5) by solving the following optimization problem

$$\min_x \|Av - xv\|_2^2 \quad (6)$$

Geometrically, one find a scalar x such that $Av - xv$ is orthogonal to v . Algebraically, eq(6) (take real case as example) boils down to find the critical point of

$$f(x) = v^T v^2 - v^T (A^T + A)vx + v^T A^T Av$$

If $v^T A^T v$ is real number, then $v^T A^T v = v^T Av$, one can use derivative of $f(x)$ to determine the critical point and the result matches the Least Square. When we consider complex number field, the expansion of $f(x) = \min_x \|Av - xv\|_2^2$ is subtle, saying

$$f(x) \neq v^* vx^2 - v^* (A^* + A)vx + v^* A^* Av$$

especially x is complex.

Residue Analysis Based on the analysis above, there are orthogonal vectors v and $xv - Av$ and this offers a way to create a self adjoint matrix B such that v is exactly the eigenvector of B , which means little difference between A and B .

Review example1.1, one knows that $vr^* + rv^*$ is a self-adjoint matrix sending v to $r = xv - Av$. We can add this matrix by A to eliminate Av , which yields $B = vr^* + rv^* + A$.

When considering simultaneous iteration to get invariant subspace (eigenvector space), the difference between theory and practice appears. For example

Numerical Instability and Orthogonalization

Consider the symmetric matrix:

$$A = \begin{bmatrix} 4 & 1 & 1 \\ 1 & 3 & 1 \\ 1 & 1 & 2 \end{bmatrix}$$

We will start with the matrix V with two columns in the initial step:

$$V^{(0)} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$$

Iteration Without Orthogonalization

Each column of $V^{(m)}$ is updated via:

$$V^{(m+1)} = AV^{(m)}$$

Perform a few iterations to illustrate the convergence:

First Iteration:

$$V^{(1)} = AV^{(0)} = \begin{bmatrix} 4 & 1 & 1 \\ 1 & 3 & 1 \\ 1 & 1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 4 & 1 \\ 1 & 3 \\ 1 & 1 \end{bmatrix}$$

Second Iteration:

$$V^{(2)} = AV^{(1)} = \begin{bmatrix} 4 & 1 & 1 \\ 1 & 3 & 1 \\ 1 & 1 & 2 \end{bmatrix} \begin{bmatrix} 4 & 1 \\ 1 & 3 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 18 & 8 \\ 8 & 11 \\ 7 & 6 \end{bmatrix}$$

... 10th Iteration:

$$V^{(10)} = AV^{(9)} = \begin{bmatrix} 8490718 & 5843148 \\ 5843148 & 4032371 \\ 4458347 & 3073546 \end{bmatrix}$$

Observation

As iterations continue, each column increasingly aligns with the dominant eigenvector

$\mathbf{v}_1 = \begin{bmatrix} 0.7558 \\ 0.5207 \\ 0.3971 \end{bmatrix}$ corresponding to $\lambda_{\max} = 5.2143$. For large m , the columns of $V^{(m)}$ converge to:

$$V^{(m)} \approx \begin{bmatrix} 0.7558 & 0.5207 \\ 0.3971 & 0.7558 \\ 0.5207 & 0.3971 \end{bmatrix}$$

This convergence indicates numerical instability, as the columns lose linear independence and all vectors converge toward the dominant eigenvector's direction.

Apply QR into Simultaneous Iteration

Orthogonalize (using QR Factorization): Perform QR factorization when computing $V^{(2)} = AV^{(1)} = A$ where

$$V^{(1)} = \begin{bmatrix} 4 & 1 \\ 1 & 3 \\ 1 & 1 \end{bmatrix} = Q^{(1)} R^{(1)} = \begin{bmatrix} -0.9428 & 0.2851 \\ -0.2357 & -0.9366 \\ -0.2357 & -0.2036 \end{bmatrix} \begin{bmatrix} -4.2426 & -1.8856 \\ 0 & -2.7285 \end{bmatrix}$$

, then we compute the QR of $AQ^{(1)} = Q^2 \hat{R}^{(2)}$,

$$\begin{aligned} V^{(2)} &= Q^{(2)} \hat{R}^{(2)} R^{(1)} \\ &= \begin{bmatrix} -0.8611 & 0.4685 \\ -0.3827 & -0.8531 \\ -0.3349 & -0.2297 \end{bmatrix} \begin{bmatrix} 4.9272 & 1.3987 \\ 0 & 2.5708 \end{bmatrix} \begin{bmatrix} -4.2426 & -1.8856 \\ 0 & -2.7285 \end{bmatrix} \end{aligned}$$

Update $V^{(m)}$ with orthogonal columns: After orthogonalization:

$$V^{(m)} = Q^{(m)} R^{(m)} = \begin{pmatrix} -0.7559 & 0.6316 \\ -0.5205 & -0.7395 \\ -0.3971 & -0.2330 \end{pmatrix} 1.0e + 07 \begin{pmatrix} -5.8556 & -4.0339 \\ 0 & -0.0024 \end{pmatrix}$$

with $m = 10$ as example. One can verify that the columns of $Q^{(10)}$ closely approximate the eigenvectors of A corresponding to its two largest eigenvalues (in magnitude).

This orthogonalization corrects the numerical errors that lead to loss of linear independence, thus ensuring each iteration step mainly jumps to the invariant subspace.

7.3 QR Iteration

The QR method mentioned to compute eigenvalues has some points which can be improved. For example,

7.4 Improved QR Iteration

8 Exercise

Problem 8

Let $A \in \mathbb{R}^{n \times n}$ and $S \in \mathbb{R}^{r \times r}$, $r < n$, be symmetric matrices such that

$$AQ_1 - Q_1S = E_1 \quad (7)$$

where $Q_1 \in \mathbb{R}^{n \times r}$ and $Q_1^T Q_1 = I$. Show that, there exist $\mu_1, \dots, \mu_r \in \lambda(A)$ such that

$$|\mu_k - \lambda_k(S)| \leq \sqrt{2} \|E_1\|_2, \quad k = 1, \dots, r$$

where $\lambda(B)$ denotes the set of all eigenvalues of B .

A natural way to deal with non-square isometric matrix is to extend it. Suppose $Q = [Q_1 \ Q_2]$ is $n \times n$ isometric and then $Q^T A Q$ is similar to A and has the same eigenvalue. We can write $Q^T A Q$ into block form, i.e.

$$\begin{aligned} \begin{bmatrix} Q_1^T \\ Q_2^T \end{bmatrix} A \begin{bmatrix} Q_1 & Q_2 \end{bmatrix} &= \begin{bmatrix} Q_1^T A \\ Q_2^T A \end{bmatrix} \begin{bmatrix} Q_1 & Q_2 \end{bmatrix} \\ &= \begin{bmatrix} Q_1^T A Q_1 & Q_1^T A Q_2 \\ Q_2^T A Q_1 & Q_2^T A Q_2 \end{bmatrix} \\ &= \begin{bmatrix} S & \\ & Q_2^T A Q_2 \end{bmatrix} + \begin{bmatrix} Q_1^T E_1 & Q_1^T A Q_2 \\ Q_2^T A Q_1 & O \end{bmatrix} \end{aligned}$$

where the first equation comes from the following rule of block matrix: assume we split matrix M into upper and lower block with the same number of columns, then

$$Mx = \begin{bmatrix} M_{1:k,1:n} \\ M_{k+1:n,1+n} \end{bmatrix} x = \begin{bmatrix} M_{1:k,1:n}x \\ M_{k+1:n,1+n}x \end{bmatrix}$$

This can be interpreted as we split a system of n equations into two system of equations (in formal).

By the equation of $Q^T A Q$, we know that the difference of $|\mu - \lambda|$ can be controlled by the norm of matrix

$$\begin{bmatrix} Q_1^T E_1 & Q_1^T A Q_2 \\ Q_2^T A Q_1 & O \end{bmatrix} = \begin{bmatrix} Q_1^T E_1 & E_1^T Q_2 \\ Q_2^T E_1 & O \end{bmatrix},$$

where we use the property $Q_2^T Q_1 = O$ and assumed equation (7)

Let's derive the bound in the following.

$$E = \begin{bmatrix} Q_1^T E_1 & E_1^T Q_2 \\ Q_2^T E_1 & 0 \end{bmatrix}, \quad y = \begin{bmatrix} \hat{y} \\ y_\perp \end{bmatrix}$$

Then:

$$Y = Ey = \begin{bmatrix} Q_1^T E_1 \hat{y} + E_1^T Q_2 y_\perp \\ Q_2^T E_1 \hat{y} \end{bmatrix}$$

The squared norm of Y is:

$$\|Y\|^2 = \|Q_1^T E_1 \hat{y} + E_1^T Q_2 y_\perp\|^2 + \|Q_2^T E_1 \hat{y}\|^2$$

Using the triangle inequality:

$$\|Y\|^2 \leq (\|Q_1^T E_1 \hat{y}\| + \|E_1^T Q_2 y_\perp\|)^2 + \|Q_2^T E_1 \hat{y}\|^2$$

Using $(a + b)^2 \leq 2(a^2 + b^2)$:

$$\|Y\|^2 \leq 2(\|E_1\|^2 \|\hat{y}\|^2 + \|E_1\|^2 \|y_\perp\|^2) + \|E_1\|^2 \|\hat{y}\|^2$$

Since $\|y\|^2 = \|\hat{y}\|^2 + \|y_\perp\|^2$:

$$\|Y\|^2 \leq 2\|E_1\|^2 \|y\|^2 + \|E_1\|^2 \|\hat{y}\|^2 \leq 3\|E_1\|^2 \|y\|^2$$

Therefore:

$$\|Y\| \leq \sqrt{3}\|E_1\|\|y\|$$

This implies:

$$\|E\| \leq \sqrt{3}\|E_1\|$$

Therefore, the best bound we can obtain with this approach is $\sqrt{3}\|E_1\|$. To achieve $\sqrt{2}\|E_1\|$ would require additional assumptions or a different approach. Final Answer: The final answer is $\boxed{\sqrt{3}\|E_1\|}$

References

- [1] Steffen Börm and Christian Mehl. *Numerical methods for eigenvalue problems*. Walter de Gruyter, 2012.
- [2] Lloyd N Trefethen and David Bau. *Numerical linear algebra*. SIAM, 2022.