

Caleb Logemann
MATH 562 Numerical Analysis II
Homework 1

1. Problem 1.1 Let B be a 4×4 matrix to which the following operations are applied in the given order.

1. double column 1
2. halve row 3
3. add row 3 to row 1
4. interchange columns 1 and 4
5. subtract row 2 from each other rows
6. replace column 4 by column 3
7. delete column 1

The result can be written as a product of 8 matrices one of which is B .

(a) What are the other 7 matrices and what order do they appear in the matrix?

The matrix that doubles column 1 is

$$C = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

when right multiplied. The following matrix halves row 3 when left multiplied.

$$D = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & .5 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

The following matrix adds row 3 to the row 1 when left multiplied.

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

The following matrix interchanges columns 1 and 4 when right multiplied.

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

The following matrix subtracts row 2 from every other row, when left multiplied.

$$G = \begin{bmatrix} 1 & -1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & -1 & 0 & 1 \end{bmatrix}$$

The following matrix replaces column 4 with column 3 when right multiplied.

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

The following matrix deletes column 1 when right multiplied.

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

The resulting matrix product is given by $GEDBCFHI$, where the matrices are given above.

- (b) The result can also be written as a product ABC what are A and C ?
In this case A and C are given by the product of the matrices to the left and the right of B in the part (a). Therefore

$$A = \begin{bmatrix} 1 & -1 & 0.5 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & -1 & 0.5 & 0 \\ 0 & -1 & 0 & 1 \end{bmatrix}$$

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

2. Show that if a matrix $A \in \mathbb{C}^{m \times m}$ is upper triangular and unitary, then it is diagonal.

Proof. Let $A \in \mathbb{C}^{m \times m}$ be unitary and upper triangular. I will denote A be its entries a_{ij} and by its columns as vectors \mathbf{a}_j . The fact that A is upper triangular implies that $a_{ij} = 0$, for $i > j$. Since A is unitary, $AA^* = I$ or $\mathbf{a}_j^* \mathbf{a}_i = 0$ for $i \neq j$ and $\mathbf{a}_j^* \mathbf{a}_i = 1$ for $i = j$. We will procede by induction. Consider $\mathbf{a}_1^* \mathbf{a}_j = 0$ for $j = 2, \dots, m$.

However since A is upper triangular $\mathbf{a}_1 = \mathbf{e}_1$, thus $\mathbf{a}_1^* \mathbf{a}_j = a_{1j}$. Therefore $a_{1j} = 0$ for $j = 2, \dots, m$. Now assume that $a_{ij} = 0$ for $i = 1, \dots, k-1$ and $j = i+1, \dots, m$. Consider $\mathbf{a}_k^* \mathbf{a}_j = 0$ for $j = k+1, \dots, m$. However we know that $a_{ik} = 0$ for $i < k$ and $a_{ik} = 0$ for $i > k$, therefore $\mathbf{a}_k = \mathbf{e}_k$. This implies that $\mathbf{a}_k^* \mathbf{a}_j = a_{kj}$. We can now conclude that $a_{kj} = 0$ for $j = k+1, \dots, m$. Now mathematical induction implies that $a_{ij} = 0$ for $i = 1, \dots, m$ and $j = i+1, \dots, m$. This means that all entries above the diagonal are zero, and thus A is a diagonal matrix. Essential each row can be shown to be zero except on the diagonal by noting that the previous row is zero everywhere except the diagonal. \square

3. Let $A \in \mathbb{C}^{m \times m}$ be Hermitian, that is $A = A^*$. Suppose that $A\mathbf{x} = \lambda\mathbf{x}$, where $\mathbf{x} \in \mathbb{C}^{m \times m}$ and $\lambda \in \mathbb{C}$, so \mathbf{x} is an eigenvector and λ is an eigenvalue.

(a) Prove that λ must be real.

Proof. Consider $\mathbf{x}^* A \mathbf{x}$.

$$\begin{aligned} (\mathbf{x}^* A) \mathbf{x} &= \mathbf{x}^* (A \mathbf{x}) \\ (A^* \mathbf{x})^* \mathbf{x} &= \mathbf{x}^* (A \mathbf{x}) \\ (\lambda \mathbf{x})^* \mathbf{x} &= \mathbf{x}^* (\lambda \mathbf{x}) \\ \bar{\lambda} \mathbf{x}^* \mathbf{x} &= \lambda \mathbf{x}^* \mathbf{x} \\ \bar{\lambda} &= \lambda \end{aligned}$$

Since $\bar{\lambda} = \lambda$, λ must be real. \square

- (b) Prove that if \mathbf{x} and \mathbf{y} are eigenvectors corresponding to different eigenvalues, then \mathbf{x} and \mathbf{y} are orthogonal.

Proof. Let λ_x and λ_y be the eigenvalues corresponding to \mathbf{x} and \mathbf{y} respectively, that is $A\mathbf{x} = \lambda_x \mathbf{x}$, $A\mathbf{y} = \lambda_y \mathbf{y}$, and $\lambda_x \neq \lambda_y$. Since A is Hermitian, we have just shown that $\lambda_x, \lambda_y \in \mathbb{R}$. Also because A is Hermitian

$$\begin{aligned} A &= A^* \\ \mathbf{x}^* A &= \mathbf{x}^* A^* \\ \mathbf{x}^* A &= (A \mathbf{x})^*. \end{aligned}$$

Since \mathbf{x} is an eigenvector

$$\mathbf{x}^* A = (\lambda_x \mathbf{x})^*$$

Since $\lambda_x \in \mathbb{R}$

$$\begin{aligned}\mathbf{x}^* A &= \lambda_x \mathbf{x}^* \\ \mathbf{x}^* A \mathbf{y} &= \lambda_x \mathbf{x}^* \mathbf{y} \\ \mathbf{x}^* \lambda_y \mathbf{y} &= \lambda_x \mathbf{x}^* \mathbf{y} \\ (\lambda_y - \lambda_x) \mathbf{x}^* \mathbf{y} &= 0\end{aligned}$$

Since $\lambda_x \neq \lambda_y$, $\lambda_y - \lambda_x \neq 0$.

$$\mathbf{x}^* \mathbf{y} = 0$$

Therefore \mathbf{x} and \mathbf{y} are orthogonal. □

4. If $\mathbf{u}, \mathbf{v} \in \mathbb{R}^m$, then $A = I + \mathbf{u}\mathbf{v}^T$ is called a rank one perturbation of the identity.
- (a) Show that if A is invertible, then its inverse has the form $A^{-1} = I + \alpha \mathbf{u}\mathbf{v}^T$, and give an expression for the scalar α .
- First let us find an expression for α .

$$\begin{aligned}I &= AA^{-1} \\ &= (I + \mathbf{u}\mathbf{v}^T)(I + \alpha \mathbf{u}\mathbf{v}^T) \\ &= I + \alpha \mathbf{u}\mathbf{v}^T + \mathbf{u}\mathbf{v}^T + \alpha \mathbf{u}\mathbf{v}^T \mathbf{u}\mathbf{v}^T \\ &= I + (\alpha + 1 + \alpha \mathbf{v}^T \mathbf{u}) \mathbf{u}\mathbf{v}^T\end{aligned}$$

In order for this equality to be true

$$\begin{aligned}0 &= \alpha + 1 + \alpha \mathbf{v}^T \mathbf{u} \\ \alpha &= -\frac{1}{1 + \mathbf{v}^T \mathbf{u}}\end{aligned}$$

Proof. If A is invertible then A^{-1} exists and is unique. Consider the matrix $B = I + \alpha \mathbf{u}\mathbf{v}^T$, where $\alpha = -\frac{1}{1 + \mathbf{v}^T \mathbf{u}}$, then the following can be stated.

$$\begin{aligned}AB &= (I + \mathbf{u}\mathbf{v}^T)(I + \alpha \mathbf{u}\mathbf{v}^T) \\ &= I + \alpha \mathbf{u}\mathbf{v}^T + \mathbf{u}\mathbf{v}^T + \alpha \mathbf{u}\mathbf{v}^T \mathbf{u}\mathbf{v}^T \\ &= I + (\alpha + 1 + \alpha \mathbf{v}^T \mathbf{u}) \mathbf{u}\mathbf{v}^T \\ &= I + \left((1 + \mathbf{v}^T \mathbf{u}) \left(-\frac{1}{1 + \mathbf{v}^T \mathbf{u}} \right) + 1 \right) \mathbf{u}\mathbf{v}^T \\ &= I + (-1 + 1) \mathbf{u}\mathbf{v}^T \\ &= I\end{aligned}$$

Similarly it can be shown that $BA = I$. Thus B is an inverse of A , and since the inverse is unique $B = I + \alpha \mathbf{u}\mathbf{v}^T$ is the inverse of A . Therefore if A is invertible it's inverse will be of the form $I + \alpha \mathbf{u}\mathbf{v}^T$. \square

- (b) If A is not invertible than $\mathbf{v}^T \mathbf{u} = -1$, otherwise $I + \alpha \mathbf{u}\mathbf{v}^T$, where $\alpha = -\frac{1}{1+\mathbf{v}^T \mathbf{u}}$ would be an inverse. In this case $(A) = \{\mathbf{u}\}$. Note that

$$\begin{aligned} A(c\mathbf{u}) &= (I + \mathbf{u}\mathbf{v}^T)(c\mathbf{u}) \\ &= c\mathbf{u} + c\mathbf{u}\mathbf{v}^T \mathbf{u} \\ &= c\mathbf{u} - c\mathbf{u} \\ &= \mathbf{0} \end{aligned}$$

for any $c \in \mathbb{R}$. Thus any multiple of \mathbf{u} is in (A) .

5. Let $\|\cdot\|$ denote any norm on \mathbb{C}^m and also the corresponding induced norm on $\mathbb{C}^{m \times m}$, so that $\|A\mathbf{x}\| \leq \|A\|\|\mathbf{x}\|$. Show that $\rho(A) \leq \|A\|$, where $\rho(A) = \max |\lambda|$: where λ is an eigenvalue of A is the spectral radius of A .

Proof. Let \mathbf{x} be an eigenvector of A with corresponding eigenvalue λ , such that $|\lambda| = \rho(A)$, that is \mathbf{x} is the eigenvector corresponding to the eigenvalue whose magnitude is the spectral radius. Consider $\|A\mathbf{x}\|$

$$\begin{aligned} \|A\mathbf{x}\| &= \|\lambda\mathbf{x}\| \\ &= |\lambda|\|\mathbf{x}\| \\ &= \rho(A)\|\mathbf{x}\| \end{aligned}$$

Also note that $\|A\mathbf{x}\| \leq \|A\|\|\mathbf{x}\|$, therefore

$$\begin{aligned} \rho(A)\|\mathbf{x}\| &\leq \|A\|\|\mathbf{x}\| \\ \rho(A) &\leq \|A\| \end{aligned}$$

Since $\|\mathbf{x}\| \neq 0$, because \mathbf{x} is an eigenvector. \square

6. Let $\theta \in (0, 2\pi)$ and define the matrix $Q \in \mathbb{R}^{2 \times 2}$ by

$$Q = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}.$$

Show that $\mathbf{y} = Q\mathbf{x}$ is the vector obtained by rotating vector \mathbf{x} by θ radians.

Proof. Let $\mathbf{x} \in \mathbb{R}^2$, have entries x_1 and x_2 , then the following statements are true $x_1 = r \cos(\phi)$ and $x_2 = r \sin(\phi)$, where $r = \sqrt{x_1^2 + x_2^2}$ is the radius or length of the vector and $\phi = \arctan\left(\frac{x_1}{x_2}\right)$ is the angle between the vector and the x-axis. If the vector is in the third or fourth quadrants the angle must be shifted as necessary. Now the vector \mathbf{y} can be computed as follows.

$$\begin{aligned}\mathbf{y} &= Q\mathbf{x} \\ &= \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} \begin{bmatrix} r \cos(\phi) \\ r \sin(\phi) \end{bmatrix} \\ &= \begin{bmatrix} r(\cos(\theta) \cos(\phi) - \sin(\theta) \sin(\phi)) \\ r(\sin(\theta) \cos(\phi) + \cos(\theta) \sin(\phi)) \end{bmatrix}\end{aligned}$$

Using some trigonometric identities

$$\mathbf{y} = \begin{bmatrix} r \cos(\theta + \phi) \\ r \sin(\theta + \phi) \end{bmatrix}$$

However this is just vector \mathbf{x} except rotated by θ radians. It has the same length or radius, but the angle was changed by θ . \square

7. Below I have shown my function for finding the SVD decomposition of a 2×2 real matrix. The code does use the `circle_image` and `arrow` functions provided in class.

```
A = [1, 2; 0, 2];
figure;
[U, S, V] = SVD2D(A);

A = [1, 1; 2, 2];
figure;
[U, S, V] = SVD2D(A);
```

```
function [U, S, V] = SVD2D(A)
    % create possible unit vectors v1
    theta = 0:.01:2*pi;
    v = [cos(theta); sin(theta)];

    Av = A*v;
    normAv = arrayfun(@(i) norm(Av(:, i), 2), 1:length(theta));
    [s1, i] = max(normAv);

    % v_i maximizes 2-norm of Av
    v1 = v(:, i);
```

```

% s1*u1 = A*v1
Av1 = A*v1;
s1 = norm(Av1,2);
u1 = Av1/s1;

% v2 should be orthogonal to v1
% generate random vector
w = rand(2,1);
v2 = w - (w'*v1)*v1;
v2 = v2/norm(v2);

% s2*u2 is A*v2
Av2 = A*v2;
s2 = norm(Av2,2);
u2 = Av2/s2;

U = [u1, u2];
V = [v1, v2];
S = diag([s1, s2]);

% now plot
circle_image(A);
% plot v vectors
arrow([0, 0], v1, 'b');
arrow([0, 0], v2, 'b');
% plot u vectors
arrow([0, 0], s1*u1, 'r');
arrow([0, 0], s2*u2, 'r');
end

```

This code produces the following two figures for

$$A = \begin{bmatrix} 1 & 2 \\ 0 & 2 \end{bmatrix}$$

and

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

respectively.

