LINEAR ALGEBRA: ANSWER TO HOMEWORK 3

VISHVAS VASUKI

1. 5.3

Consider the matrix A

```
  -2    11
 -10     5
```

Remark 1.0.1. I have verified my calculations with Matlab, and used it to perform some tedious arithmatic. However, I followed the strategy I would on paper.

1.1. **a.** The real SVD of A is:

```
U =

  -0.7071   -0.7071
  -0.7071    0.7071


S =

  14.1421         0
        0    7.0711


V =

   0.6000   -0.8000
  -0.8000   -0.6000
```

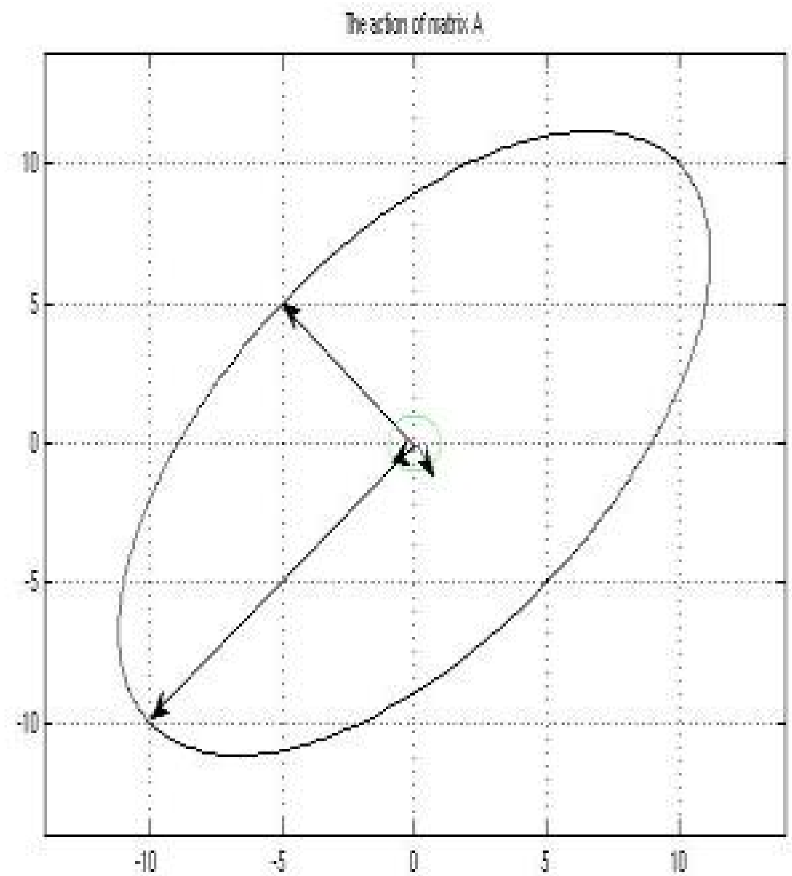
This was calculated by considering the fact that $AA^* = U\Sigma^2U^*$, and $A^*A = V\Sigma^2V^*$. So, U, V and Σ can be found by calculating the eigenvalue decomposition of AA^* and A^*A.

Method used for finding eigenvalues and eigenvectors: Eigenvalues of a square matrix B may be found by solving the polynomial corresponding to $det(B-lI) = 0$. Eigenvectors can then be found by solving for x in the equation $(B-lI)x = 0$.

For verification, the value for Σ^2 found in this manner was:

```
  50     0
   0   200
```

1.2. **b.** The singular values are: 14.1421 and 7.0711. The left singular vectors are the columns of V shown above. The right singular vectors are the columns of U shown above.



Shown in the figure are the vectors corresponding to the columns of V and the columns of matrix

```
US =

  -10.0000   -5.0000
  -10.0000    5.0000
```

1.3. **c.** $\|A\|_1 = 16$. This is obtained from the largest column sum of A.

$\|A\|_2 = 14.1421$.

This is obtained in the following manner: Take vector in the unit ball x*=[a $\sqrt{1-a^2}$]. Then, we get (Ax)* = [-2a + 11$\sqrt{1-a^2}$ -10a + 5$\sqrt{1-a^2}$]. We then maximize the norm of this vector.

$\|A\|_\infty = 15$: $\|Ax\|_\infty$ is maximum when x* = [-1, 1].

$\|A\|_{Frob} = \sqrt{\sum_{i=1}^n \sum_{j=1}^n A_{i,j}} = 15.8114$.

1.4. **d.** $A^{-1} = (U\Sigma V^*)^{-1} = V\Sigma^{-1}U*$. Upon performing this calculation, we find:

```
0.0500   -0.1100
0.1000   -0.0200
```

1.5. **e.** Any single eigenvalue and eigenvector satisfy the equation $Ax = lx$. So, $(A - lI)x = 0$. So, $det(A - lI)$. This is the determinant of the matrix:

```
 -2-l    11
-10     5-l
```

So, we solve for l in the characteristic equation: $(-2-l)(5-l)+110 = 0$ to get the eigenvalues l_1, l_2: $1.5000 \pm 9.8869i$.

1.6. **f.** Indeed, upon calculation, we verify that $l_1l_2 = det(A) = 100$.
Indeed, upon calculation, we verify that $\sigma_1\sigma_2 = |det(A)| = 100$.

1.7. **g.** The area of the ellipsoid: $\pi\sigma_1\sigma_2 = 314.16$

2. 5.4

$A \in C^{m\times m}$ has an SVD A=USV*. Find the eigenvalue decomposition of the $2m \times 2m$ hermetian matrix T =

$$\begin{bmatrix} 0 & A^* \\ A & 0 \end{bmatrix}$$

2.1. **Answer.** The answer as written earlier was incorrect.

3. 6.2

Let E be the $m \times m$ matrix that extracts the 'even part' of an m-vector: Ex = (x+ Fx)/2 where F is the $m \times m$ matrix which flips $(x_1, \dots x_m)^*$ to $(x_m, \dots x_1)^*$. Is E an orthogonal projector or an oblique projector, or not a projector at all? What are its entries?

3.1. **Answer.** F is a matrix whose secondary diagonal is filled with 1 and whose non-secondary diagonal elements are filled with 0.

E = 0.5(F+I). All the entries on its primary and secondary diagonals are non zero. All other elements are 0.

In case m is even, all primary and secondary diagonal entries are 0.5.

In case m is odd, all primary and secondary diagonal entries are 0.5, except for $E_{\lceil m/2\rceil,\lceil m/2\rceil}$, which is 1.

In either case, $E = E^2$; so E is a projector. Also, in either case, $E = E^*$; so E is an orthogonal projector.

4. 7.5

Let A be an $m \times n$ matrix $(m \geq n)$, and let $A = \hat{Q}\hat{R}$ be a reduced QR factorization.

Notation. ith column of A is represented by a_i. r_i^* represents the ith row of R.

4.1. **1.**

Theorem 4.1.1. *A has rank n if and only if all the diagonal entries of $\hat{R}$ are nonzero.*

Proof. $A = \hat{Q}\hat{R}$.

If A has rank n, the vectors $\{a_i\}$ span an n dimensional space. Suppose for the sake of contradiction that $r_{i,i} = 0$ for some i. Then it would imply that the space spanned by i linearly independent vectors $a_1 \dots a_i$ is spanned by only i-1 vectors $q_1 \dots q_{i-1}$, which is an absurdity. Hence, if A has rank n, $\forall i : r_{i,i} \neq 0$.

Suppose that $\forall i : r_{i,i} \neq 0$. Now we consider the space spanned by the matrix $A = \hat{Q}\hat{R}$. All vectors in $\{q_i\}$ are mutually $\perp$, and therefore independent. So, the rank of $\hat{Q}$ is n, and $\{q_i\}$ span an n dimensional space. For every q_i, there is atleast one vector a_i (especially $a_i = Qr_i$) which depends on q_i (that is, has a component in the direction of q_i). Hence, the columns of A span a space whose dimensions are at least n. But, as A is an $m \times n$ matrix ($m \geq n$), its rank is at most n. So, A has rank n. □

4.2. **2.**

Theorem 4.2.1. *$\hat{R}$ has k nonzero diagonal entries for some k with $0 \leq k < n$. Then, $rank(A) \geq k$.*

Proof. $a_i = Qr_i$. More specifically, every a_i can be expressed as a linear combination of vectors $q_1 \dots q_{i-1}$. Suppose that $r_i = 0$ for some i. Then a_i can be expressed as a linear combination of vectors $q_1 \dots q_{i-2}$.

Proof to be completed. □

5. Question

(Gram-Schmidt Process) Let

$$v_1 = \begin{bmatrix} 1 \\ \varepsilon \\ 0 \\ 0 \end{bmatrix}, \quad v_2 = \begin{bmatrix} 1 \\ 0 \\ \varepsilon \\ 0 \end{bmatrix}, \quad v_3 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \varepsilon \end{bmatrix},$$

and ε be such that $fl(1 + \varepsilon^2) = 1$.

- a Apply Classical Gram-Schmidt and show that the computed vectors are not numerically orthogonal, i.e., computed vectors have dot products much larger than ε.
- b Apply Modified Gram-Schmidt and show that the computed vectors are numerically orthogonal, i.e., computed vectors have dot products $= O(\varepsilon)$.

5.1. **Classical gram Schmidt:** We simulate the following code:

```
for j=[1:n]
    v = A(:,j);
    for i=[1:j-1]
        R(i,j) = Q(:,i)'*A(:,j);
        v=v-R(i,j)*Q(:,i);
    end
    R(j,j)=norm(v,2);
    Q(:,j)=v/R(j,j);
```

```
end
```

$v_1^* =$ [1 e 0 0].

$v_2^* =$ [1 0 e 0].

$v_3^* =$ [1 0 0 e].

When j=1, i=1: $v_1^* =$ [1 e 0 0]. $R(1,1) = (fl(1+e^2))^{0.5} = 1$. $q_1^* =$ [1 e 0 0].

When j=2: $v_2^* =$ [1 0 e 0]. When i=1: R(1,2) = 1. $v_2^* =$ [0 -e e 0]. $R(2,2) = e\sqrt{2}$. $q_2^* = [0\ \ -2^{-0.5}\ 2^{-0.5}\ 0]$.

When j=3: $v_3^* =$ [1 0 0 e]. When i=1: R(1,3) = 1. $v_3^* =$ [0 -e 0 e]. When i=2: R(2,3) = 0. $v_3^* =$ [0 -e 0 e]. $R(3,3) = e\sqrt{2}$. $q_3^* = [0\ \ -2^{-0.5}\ 0\ 2^{-0.5}]$.

Now, we check orthogonality: $q_1^* q_2 = -2^{-0.5}e$. $q_1^* q_3 = -2^{-0.5}e$. $q_3^* q_2 = 2^{-1}$.

The computed vectors <u>are not</u> numerically orthogonal, i.e., computed vectors have dot products much larger than ε.

5.2. **Modified gram Schmidt:** We simulate the following code:

```
for j=[1:n]
    R(j,j)=norm(V(:,j),2);
    Q(:,j)=V(:,j)/R(j,j);
    for i=[j+1:n]
        R(j,i) = Q(:,j)'*V(:,i);
        V(:,i)=V(:,i)-R(j,i)*Q(:,j);
    end
end
```

$v_1^* =$ [1 e 0 0]. $v_2^* =$ [1 0 e 0]. $v_3^* =$ [1 0 0 e].

When j=1: $v_1^* =$ [1 e 0 0]. $R(1,1) = fl((1+e^2)^{1/2}) = 1$. $q_1^* =$ [1 e 0 0]. When i=2: R(1,2) = 1. $v_2^* =$ [0 -e e 0]. When i=3: R(1,3) = 1. $v_3^* =$ [0 -e 0 e].

When j=2: $v_2^* =$ [0 -e e 0]. $R(2,2) = e2^{0.5}$. $q_2^* = [0\ \ -2^{-0.5}\ 2^{-0.5}\ 0]$. When i=3: $R(2,3) = e2^{0.5}$. $R(2,3)q_2^* = [0\ e\ e\ 0]$. $v_3^* = [0\ \ -2e\ \ -e\ e]$.

When j=3: $v_3^* = [0\ -2e\ -e\ e]$. $R(3,3) = e6^{0.5}$. $q_3^* = [0\ -2(6^{-0.5})\ -6^{-0.5}\ 6^{-0.5}]$.

Now, we check orthogonality: $q_1^* q_2 = -2^{-0.5}e$. $q_1^* q_3 = -2(6^{-0.5})e$. $q_3^* q_2 = 0$.

The computed vectors <u>are</u> numerically orthogonal, i.e., computed vectors have dot products $= O(\varepsilon)$.