Homework 4

ZHANG Dehong

55199998.

Exercise 1

[5 points]. Prove that the eigenvalues of a symmetric matrix $A \in \mathbb{R}^{n\times n}$, where $A = A^T$, are real-valued (*i.e.*, not complex-valued). Prove also that the eigenvectors of A corresponding to different eigenvalues are orthogonal to each other.

Assume a Complex number λ is the eigenvalue of the symmetric matrix $A \in \mathbb{R}^{n\times n}$; $\vec{x}$ is the corresponding eigenvector:

$$\begin{cases} A\vec{x} = \lambda\vec{x}, \quad \vec{x} \neq 0. \\ A = A^T \end{cases}$$

① Assume $\vec{y} = A\vec{x}$

$$\|y\|_2^2 = \vec{y}\cdot\vec{y} = (Ax)^T(Ax) = x^TA^TAx = x^TAAx$$
$$= x^TA(\lambda x) = \lambda x^TAx = \lambda^2 x^Tx = \lambda^2\|x\|_2^2$$

$\lambda = \sqrt{\frac{\|y\|_2^2}{\|x\|_2^2}}$ $(\|y\|_2^2 > 0,\ \|x\|_2^2 > 0)$. therefore λ is real number

② Assume.

$$\begin{cases} A\vec{x}_1 = \lambda_1\vec{x}_1 \\ A\vec{x}_2 = \lambda_2\vec{x}_2 \\ \lambda_1 \neq \lambda_2 \\ A = A^T. \end{cases}$$

$$(Ax_2)^T(x_1) = x_2^TA^Tx_1 = x_2^TAx_1 = \lambda_1 x_2^Tx_1$$
$$(x_1)^T(Ax_2) = x_1^TAx_2 = \lambda_2 x_1^Tx_2$$

According to the property of dot product.

$$a\cdot b = |a||b|\cos\theta = |b||a|\cos\theta = b\cdot a.$$

$$\begin{cases} (Ax_2)^Tx_1 = x_1^TAx_2 \\ \vec{x_1}^T\vec{x_2} = \vec{x_2}^T\vec{x_1}. \end{cases}$$

$$\lambda_1 x_2^Tx_1 = \lambda_2 x_1^Tx_2.$$
$$\lambda_1 x_1^Tx_2 = \lambda_2 x_1^Tx_2$$
$$(x_1^Tx_2)(\lambda_1 - \lambda_2) = 0.$$

Since. $\lambda_1 \neq \lambda_2$, $\lambda_1, \lambda_2 \in \mathbb{R}$.

$$x_1^Tx_2 = 0.$$

Exercise 2

[5 points]. Derive the lth largest direction of variance in principal component analysis (PCA).

To find the l-th largest direction of variance it's equivalent is.

$$\max_v v^T\Sigma v.$$

subject to $\|v\|_2^2 = 1$.

$$v^Tv_1 = v^Tv_2 = \cdots = v^Tv_{l-1} = v_1^Tv_2 = \cdots = v_{l2}^Tv_l = 0.$$

$$L(v, \lambda, \nu) = v^T\Sigma v + \lambda(1 - \|v\|_2^2) + \sum_{i=1}^{l-1}\nu_i v^Tv_i$$

$$\frac{\partial L}{\partial v} = 2\Sigma v - 2\lambda v + \sum_{i=1}^{l-1}\nu_i v_i = 0.$$

Left multiply v_i^T on both sides for each $i \in [0, l-1]$

$$2v_i^T\Sigma v - 2\lambda v_i^Tv + \nu_i v_i^Tv_i = 0.$$
$$2(\Sigma v_i)^Tv - 0 + \nu_i = 0$$
$$\nu_i = 0 \quad (\text{for each } i \in [0, l-1]).$$

we have $\Sigma v = \lambda v$.

Therefore v^* is the eigenvector corresponding to the l-th largest eigenvalue.

Exercise 3

[5 points]. 1) Suppose that a discrete-time linear system has outputs $y[n]$ for the given inputs $x[n]$, as shown in Fig. 1. Determine the response $y_4[n]$ when the input is as shown in Fig. 2.

a) [1 point]. Express $x_4[n]$ as a linear combination of $x_1[n]$, $x_2[n]$, and $x_3[n]$.

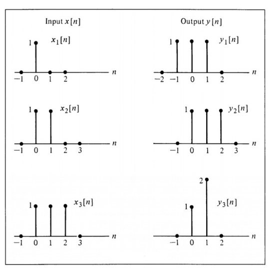


Figure 1:

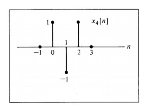


Figure 2:

$$X = [x_1, x_2, x_3] = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \qquad x_4 = \begin{array}{c} \begin{bmatrix} 0 \\ 1 \\ -1 \\ 1 \\ 0 \end{bmatrix} \end{array}\begin{array}{c} -1 \\ 0 \\ 1 \\ 2 \\ 3 \end{array}$$

$$Xw = x_4$$

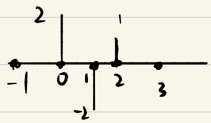
$$A = \begin{array}{c} a \\ b \\ c \end{array}\left[\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & -1 \\ 0 & 0 & 1 & 1 \end{array}\right] \qquad w = \begin{bmatrix} 2 \\ -2 \\ 1 \end{bmatrix}$$

$c = 1$, $b = -2$, $a = 2$

therefore, $X w = x_4$

where $X = [x_1, x_2, x_3]$

$$w = \begin{array}{c|c} \begin{bmatrix} 0 \\ 2 \\ -2 \\ 1 \\ 0 \end{bmatrix} & \begin{matrix} -1 \\ 0 \\ 1 \\ 2 \\ 3 \end{matrix} \end{array} \qquad W:$$



b) [1 point]. Using the fact that the system is linear, determine $y_4[n]$, the response to $x_4[n]$.

The system is linear, therefore

$X^T A = Y^T$ where $X = [x_1, x_2, x_3]$ $Y = [y_1, y_2, y_3]$

$x_4^T A = y_4^T$

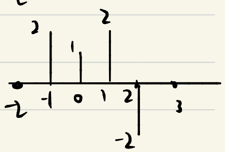
$w^T X^T A = w^T Y^T = y_4^T$

$y_4 = Y w$

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

$$= \begin{bmatrix} 0 \\ 2 \\ 2 \\ 2 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ -2 \\ -2 \\ -2 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 1 \\ 2 \\ 0 \\ 0 \end{bmatrix}$$

$$= \begin{bmatrix} 0 \\ 2 \\ 1 \\ 2 \\ -2 \\ 0 \end{bmatrix} \begin{matrix} -2 \\ -1 \\ 0 \\ 1 \\ 2 \\ 3 \end{matrix}$$



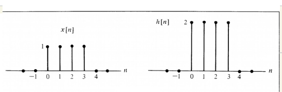
c) [1 point]. From the input-output pairs in Fig. 1, determine whether the system is time-invariant.

According to the last pair (compared to the first two) the system is not time-invariant.

2) Determine the discrete-time convolution of $x[n]$ and $h[n]$ for the following two cases.

a) [1 point]. As shown in Fig. 3.

b) [1 point]. As shown in Fig. 4.



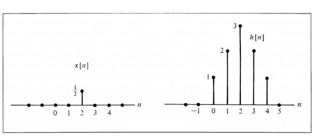
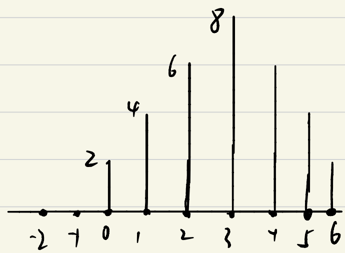


Figure 4:

$$f(n) = (x*h)(n) = \sum_{k=-\infty}^{+\infty} x(k)\, h(n-k)$$

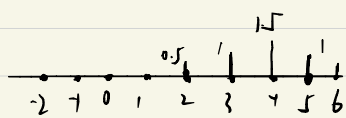
a)

$f(-2) = 0$, $f(-1) = 0$, $f(0) = 2$, $f(1) = 4$, $f(2) = 6$ $f(3) = 8$.
$f(4) = 6$, $f(5) = 4$, $f(6) = 2$



(b)

$f(-2) = 0$, $f(-1) = 0$, $f(0) = 0$, $f(1) = 0$, $f(2) = 0.5$
$f(3) = 1$ $f(4) = 1.5$ $f(5) = 1$, $f(6) = 0.5$



Code in C++.

```cpp
#include <iostream>
#include <string.h>
using namespace std;
#define MAX 100
#define NUM 8
double x[MAX], h[MAX];
int t;
double getxval(int k){
    return k < NUM-2 && k>=-2 ? x[k+2] : 0.0;
}
double gethval(int k){
    return k < NUM-2 && k>=-2 ? h[k+2] : 0.0;
}
int main(){
    cin >> t;
    while(t--){
        for(int i=0;i<NUM;i++) cin >> x[i];
        for(int i=0;i<NUM;i++) cin >> h[i];
        for(int n=-2; n<NUM-2 ; n++){
            double ans = 0;
            for(int k=-2;k<NUM-2;k++){
                ans += getxval(k)*gethval(n-k);
                // cout << k << " "<<getxval(k) << " " <<n-k<<" "<< gethval(n-k) << endl;
            }
            cout << "n: " << n << " f(n): " << ans << endl;
        }
        cout << endl;
    }

    return 0;
}
```

Exercise 4 - Optional

In lecture notes, we have proven the Eckart-Young-Mirsky Theorem under the Frobenius norm. Here prove that the same theorem holds under the spectral norm[1] as well. Specifically, given an $m \times n$ matrix X of rank $r \leq \min\{m, n\}$ and its singular value decomposition $X = U\Sigma V^T$, with singular values $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_r$, among all $m \times n$ matrices of lower rank $l < r$, the best approximation is $Y^\star = U\Sigma_l V^T$, where Σ_l is the diagonal matrix with singular values $\sigma_1, \sigma_2, \ldots, \sigma_l$ in the sense that

$$\|X - Y^\star\|_2 = \min\{\|X - Y\|_2; Y \in \mathbb{R}^{m\times n}, \text{rank}Y \leq l\},$$

where $\|A\|_2 = \max_z \frac{\|Az\|_2}{\|z\|_2}$ is the spectral norm of $A \in \mathbb{R}^{m\times n}$ and is equal to the largest singular value of A.

For orthogonal matrices U, V, we can prove UV is also orthogonal matrix.

$$\begin{cases} U^TU = I \\ V^TV = I \end{cases} \quad (UV)^T(UV) = V^TU^TUV = I.$$

Therefore: according to SVD

$\|A\|_2 = \|U\Sigma V^T\|_2$ $\quad$ W is an orthogonal matrix $W^TW = I$

$\|WA\|_2 = \|WU\Sigma V^T\|_2$ define $U' = WU$

$= \|U'\Sigma V^T\|_2.$

$= \|U\Sigma V^T\|_2 = \|A\|_2 \leftarrow$ (since Σ is not changed)

$\|AW\|_2 = \|U\Sigma V^TW\|_2$

$= \|U\Sigma(W^TV)^T\|_2$ (W^T is orthogonal)

$= \|U\Sigma V'^T\|_2 = \|U\Sigma V^T\| = \|A\|_2.$

$\min\|X - Y\|_2 = \min\|U\Sigma V^T - Y\|_2$ $\quad$ ($\text{rank}(Y) \leq l$)

$= \min\|\Sigma - U^TYV\|_2.$

Denote $Z = U^TYV$, an $m\times n$ matrix of rank l

$m\times m$ $m\times n$ $n\times n$

$\|\Sigma - Z\|_2 = \|$

CS4487: Home Assignment №4

Department of Computer Science
City University of Hong Kong

Due on November 20, 2019, 7pm

Exercise 1

[5 points]. Prove that the eigenvalues of a symmetric matrix $A \in \mathbb{R}^{n \times n}$, where $A = A^T$, are real-valued (*i.e.*, not complex-valued). Prove also that the eigenvectors of A corresponding to different eigenvalues are orthogonal to each other.

Exercise 2

[5 points]. Derive the lth largest direction of variance in principal component analysis (PCA).

Exercise 3

[5 points]. 1) Suppose that a discrete-time linear system has outputs $y[n]$ for the given inputs $x[n]$, as shown in Fig. 1. Determine the response $y_4[n]$ when the input is as shown in Fig. 2.

a) [1 point]. Express $x_4[n]$ as a linear combination of $x_1[n]$, $x_2[n]$, and $x_3[n]$.

b) [1 point]. Using the fact that the system is linear, determine $y_4[n]$, the response to $x_4[n]$.

c) [1 point]. From the input-output pairs in Fig. 1, determine whether the system is time-invariant.

2) Determine the discrete-time convolution of $x[n]$ and $h[n]$ for the following two cases.

a) [1 point]. As shown in Fig. 3.

b) [1 point]. As shown in Fig. 4.

Exercise 4 - Optional

In lecture notes, we have proven the Eckart-Young-Mirsky Theorem under the Frobenius norm. Here prove that the same theorem holds under the spectral norm[1] as well. Specifically, given an $m \times n$ matrix X of rank $r \leq \min\{m, n\}$ and its singular value decomposition $X = U\Sigma V^T$, with singular values $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_r$, among all $m \times n$ matrices of lower rank $l < r$, the best approximation is $Y^\star = U\Sigma_l V^T$, where Σ_l is the diagonal matrix with singular values $\sigma_1, \sigma_2, \ldots, \sigma_l$ in the sense that

$$\|X - Y^\star\|_2 = \min\{\|X - Y\|_2; Y \in \mathbb{R}^{m\times n}, \mathrm{rank}Y \leq l\},$$

where $\|A\|_2 = \max_z \frac{\|Az\|_2}{\|z\|_2}$ is the spectral norm of $A \in \mathbb{R}^{m\times n}$ and is equal to the largest singular value of A.

[1] http://mathworld.wolfram.com/SpectralNorm.html

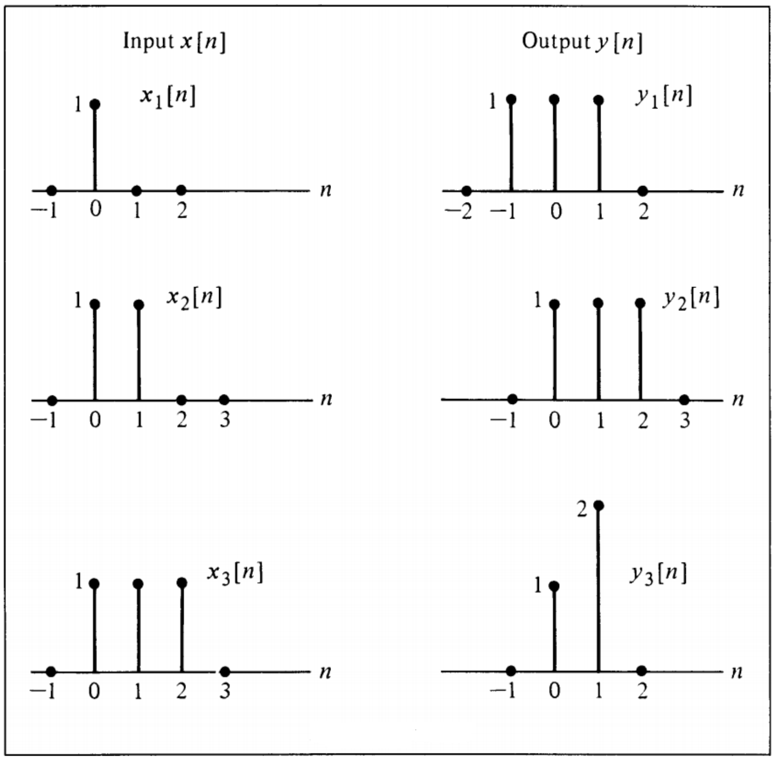


Figure 1:

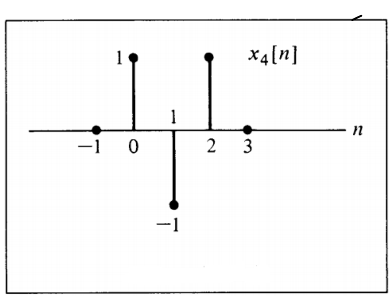


Figure 2:

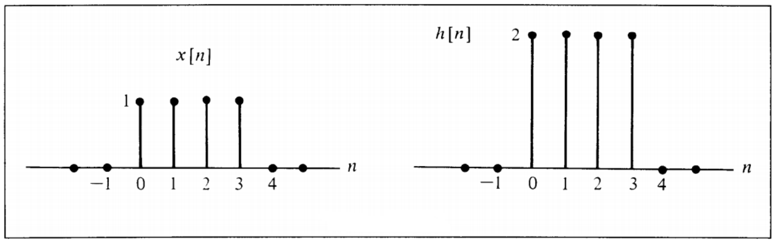


Figure 3:

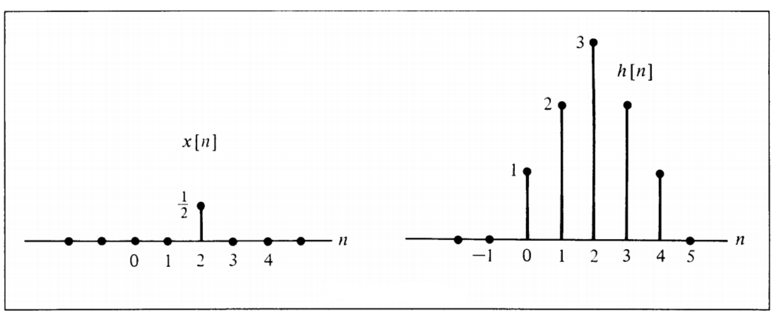


Figure 4: