МИНОБРНАУКИ РОССИИ
САНКТ-ПЕТЕРБУРГСКИЙ ГОСУДАРСТВЕННЫЙ
ЭЛЕКТРОТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ
«ЛЭТИ» ИМ. В. И. УЛЬЯНОВА (ЛЕНИНА)
Кафедра математического обеспечения и применения ЭВМ

ОТЧЕТ

по практической работе №2
по дисциплине «Машинное обучение»

Студент гр. 1310	__________________	Комаров Д. Е.
Преподаватель	__________________	Жангиров Т.Р.

Санкт-Петербург
2025

Задание 1

Постановка задачи

Рассмотрим данные, представленные в таблице 1.

Таблица 1 – Данные для первого задания

i	a1	a1
x1	4	2.9
x2	2.5	1
x3	3.5	4
x4	2	2.1

Строки соответствуют наблюдениям, столбцы признакам. $x_{i,1}$ – Означает 1-й признак i-го наблюдения.

Есть ядро (функция сходства)

$$K(x_i, x_j) = \|x_i - x_j\|^2 = (x_{i1} - x_{j1})^2 + (x_{i2} - x_{j2})^2.$$

Рассчитайте ядерную матрицу. Ядерная матрица должна получиться размером 4x4, на главной диагонали которой стоят 0.

Код программы

```
Xa1=[4, 2.5,3.5,4]
Xa2=[2.9,1,4,2.1]

Core=[]
for i in range(4):
    Core.append([])
    for j in range(4):
         Core[i].append((Xa1[i]-Xa1[j])**2+(Xa2[i]-Xa2[j])**2)

for a in Core:
    for b in a:
        print("%.2f" % b,end=' ')
    print()
```

Результат выполнения

В результате выполнения программы получилась матрица

$$Core = \begin{pmatrix} 0 & 5.86 & 1.46 & 0.64 \\ 5.86 & 0 & 10 & 3.46 \\ 1.46 & 10 & 0 & 3.86 \\ 0.64 & 3.46 & 3.86 & 0 \end{pmatrix}.$$

Задание 2

Постановка задачи

Рассмотрим данные в виде матрицы D, представленной в таблице 2.

Таблица 2 – Матрица D

	a1	a2
X1	8	-20
X2	0	-1
X3	10	-19
X4	10	-20
X5	2	0

Строки соответствуют наблюдениям, столбцы признакам. Выполните задания:

1. рассчитайте среднее $\boldsymbol{\mu}$ и ковариационную матрицу $\boldsymbol{\Sigma}$ для матрицы D;
2. рассчитайте собственные числа для матрицы $\boldsymbol{\Sigma}$;
3. рассчитайте матрицу корреляций.

Код программы

```
import numpy as np
D=np.array([[8,-20],[0,-1],[10,-19],[10,-20],[2,0]])
m=np.mean(D,axis=0)
print('D mean: ',m)
covMat=np.cov(D.transpose())
print('Covariation matrix:\n',covMat)
eig=np.linalg.eig(covMat)
lam1,lam2=max(eig.eigenvalues),min(eig.eigenvalues)
print('Eigenvalues: %.2f, %.2f'% (lam1,lam2))
print('Corelation matrix:\n',np.corrcoef(D.transpose()))
```

Результат выполнения

1. Математическое ожидание представлено вектором

$$\boldsymbol{\mu} = \begin{pmatrix} 6 \\ -12 \end{pmatrix},$$

ковариационная матрица

$$\boldsymbol{\Sigma} = \begin{pmatrix} 22 & -47.5 \\ -47.5 & 110.5 \end{pmatrix}.$$

2. Собственные числа матрицы $\boldsymbol{\Sigma}$ равны 131.17 и 1.33

3. Матрица корреляций

$$\mathbf{r} = \begin{pmatrix} 1 & -0.96 \\ -0.96 & 1 \end{pmatrix}.$$

Задание 3

Постановка задачи

Для данных и полученных результатов из второго задания:

1. рассчитайте первые две главные компоненты;
2. на основании собственных чисел рассчитайте остаточную дисперсию для каждой компоненты;
3. в декартовых координатах исходного пространства, отобразите вектора, соответствующие главным компонентам. (таким образом будет видно, на какие вектора проводилась проекция). В качестве точки, откуда начинаются вектора, используйте **μ**.

Код программы

```
import numpy as np
import matplotlib.pyplot as plt

D=np.array([[8,-20],
            [0,-1],
            [10,-19],
            [10,-20],
            [2,0]])

covMat=np.cov(D.transpose())
eig=np.linalg.eigh(covMat)
lam1,lam2=max(eig.eigenvalues),min(eig.eigenvalues)
pc1=eig.eigenvectors[0]    if    eig.eigenvalues[0]>eig.eigenvalues[1]    else
eig.eigenvectors[1]
pc2=eig.eigenvectors[1]    if    eig.eigenvalues[0]>eig.eigenvalues[1]    else
eig.eigenvectors[0]
print('pc1:',pc1)
print('pc2:',pc2)
print("Residual var for pc1: %.2f"%(sum(np.var(D,axis=0,ddof=1))-lam1))
print("Residual var for pc2: %.2f"%(sum(np.var(D,axis=0,ddof=1))-lam1-lam2))
vlen=10
plt.arrow(m[0],m[1],vlen*pc1[0],vlen*pc1[1],head_width=0.5,
head_length=0.7,label='pc1',fc='red')
plt.arrow(m[0],m[1],vlen*pc2[0],vlen*pc2[1],head_width=0.5,
head_length=0.7,label='pc2',fc='yellow')
plt.scatter(*D.transpose())
plt.grid()
plt.legend()
plt.xlabel('a1')
plt.ylabel('a2')
plt.show()
```

Результат выполнения

1. Первая главная компонента

$$pc1 = \binom{-0.4}{0.92},$$

вторая

$$pc2 = \begin{pmatrix} -0.92 \\ -0.4 \end{pmatrix}.$$

2. Остаточная дисперсия для первой компоненты равна 1.33, для второй равна 0.

3. Вектора, соответствующие главным компонентам представлены на рисунке 1.

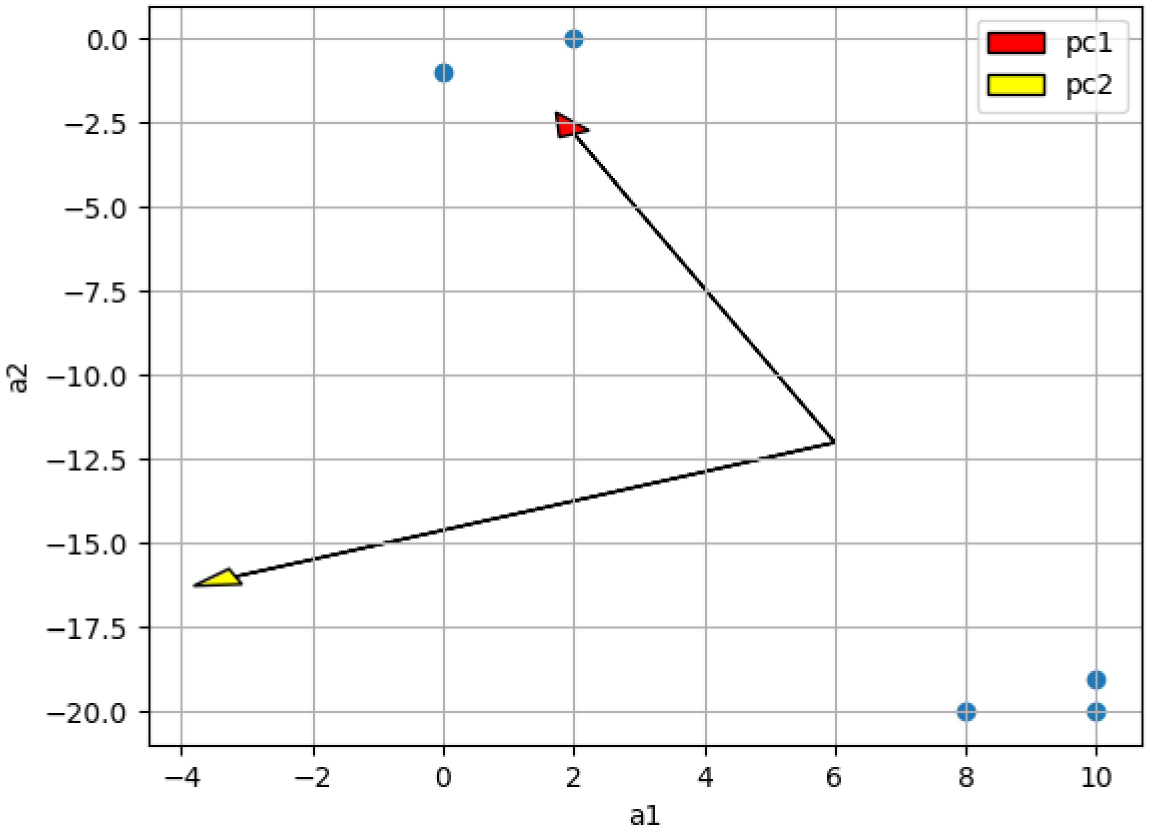

Рисунок 1 – Вектора, соответствующие главным компонентам