PCA

Problem 3.

(a). $\tilde{Y}_1$ preserves 70% variance of Y_1. $Cov(X) = \frac{1}{N} X^T X$. $Cov(Y_1) = \lambda^2 \cdot \frac{1}{N} X^T X$

for Y_1, $\sum_{i=1}^{k} \lambda_i^{Y_1} = \lambda^2 \sum_{i=1}^{k} \lambda_i^{X} \geq 0.7 \cdot \lambda^2 \sum_{i=1}^{D} \lambda_i^{X} = 0.7 \sum_{i=1}^{D} \lambda_i^{Y_1}$

(b). $\tilde{Y}_2$ preserves 70% variance of Y_2

R is an orthonogal matrix, it only rotates the data matrix X

doesn't chang the variance (eigenvalues) of the covariance matrix.

(c). $\tilde{Y}_3$ preserves 70% variance of Y_3

$Cov(Y_3) = \frac{1}{N} \cdot (XP)^T \cdot XP = \frac{1}{N} \cdot P^T X^T X P = P^T \Sigma P$

for Y_3, $\sum_{i=1}^{k} \lambda_i^{Y_3} = 25 \sum_{i=1}^{k} \lambda_i^{X} \geq 0.7 \times 25 \cdot \sum_{i=1}^{D} \lambda_i^{X} = 0.7 \cdot \sum_{i=1}^{D} \lambda_i^{Y_3}$

so $\sum_{i=1}^{k} \lambda_i^{Y_3} \geq 0.7 \cdot \sum_{i=1}^{D} \lambda_i^{Y_3}$

(d). can not tell without additional information. it scales $\overbrace{\text{each}}^{\text{data in}}$ dimension with different number, which has changed the covariance distribution.

(e). $\tilde{Y}_5$ preserves 70% variance of Y_5

PCA will center the data, which will remove the shifting amount μ.

(f). $\tilde{Y}_6$ preserves 100% variance of Y_6

$rank(Y_6) \leq 5$. when $k = 5$, the projected data has preserved all the dimensions.

Problem 4.

(a). 1. Center the data: $\bar{x} = [\bar{x}_1, \bar{x}_2, \bar{x}_3] = \frac{1}{4} \begin{bmatrix} 4+2+4-2 \\ 3+1-1+1 \\ 2-2+2+2 \end{bmatrix}^T = \frac{1}{4} \begin{bmatrix} 8 \\ 4 \\ 4 \end{bmatrix}^T = [2, 1, 1]$

$$\tilde{X} = X - \bar{X} = \begin{bmatrix} 2 & 2 & 1 \\ 0 & 0 & -3 \\ 2 & -2 & 1 \\ -4 & 0 & 1 \end{bmatrix}$$

2. Compute the covariance matrix:

$$\Sigma_{\tilde{X}} = \frac{1}{N} \cdot \tilde{X}^T \tilde{X} = \frac{1}{4} \cdot \begin{bmatrix} 2 & 0 & 2 & -4 \\ 2 & 0 & -2 & 0 \\ 1 & -3 & 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} 2 & 2 & 1 \\ 0 & 0 & -3 \\ 2 & -2 & 1 \\ -4 & 0 & 1 \end{bmatrix}$$

$$= \frac{1}{4} \begin{bmatrix} 24 & 0 & 0 \\ 0 & 8 & 0 \\ 0 & 0 & 12 \end{bmatrix} = \begin{bmatrix} 6 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

3. Eigenvector Decomposition of $\Sigma_{\hat{x}}$

$$\Sigma_{\hat{x}} = \Gamma^T \Lambda \Gamma = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}^T \begin{bmatrix} 6 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$PC_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \quad PC_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \quad PC_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

$$Var(PC_1) = 6 \quad Var(PC_2) = 2 \quad Var(PC_3) = 3$$

(b). the top 2 PC is $PC_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ and $PC_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$

$$\Gamma_{truncated} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix}$$

$$Y = \tilde{X} \cdot \Gamma_{truncated} = \begin{bmatrix} 2 & 2 & 1 \\ 0 & 0 & -3 \\ 2 & -2 & 1 \\ -4 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 0 & -3 \\ 2 & 1 \\ -4 & 1 \end{bmatrix}$$

$$\sum_{i=1}^{2} \lambda_i = 6+3 = 9$$

$$\sum_{i=1}^{3} \lambda_i = 6+3+2 = 11$$

$$\text{fraction} = \frac{9}{11} = 81.8\%$$

(c). x_5 should not change the mean $\bar{x} = [2, 1, 1]$, or the covariance matrix will change.

so set $x_5 = [2, 1, 1]$

Problem 5. $P = MV \Rightarrow P_{Leslie} = [0, 3, 0, 0, 4] \cdot \begin{bmatrix} 0.58 & 0 \\ 0.58 & 0 \\ 0.58 & 0 \\ 0 & 0.71 \\ 0 & 0.71 \end{bmatrix} = [\underset{\text{Sci-Fi-Concept}}{1.74}, \underset{\text{Romance-Concept}}{2.84}]$

according to the projected data, Leslie likes the Romance-concept movies more than the SciFi-concept movies, so she will rate higher for Casablanca than the Matrix and Star Wars.

Problem 6. the optimal solution $w^* = (X^T X^{-1}) X^T y$

$X = U \Sigma V^T$

$U \in \mathbb{R}^{N \times D}$

$\Sigma \in \mathbb{R}^{D \times D}$

$V^T \in \mathbb{R}^{D \times D}$

$$\begin{aligned}
&= (U \Sigma V^T)^T (U \Sigma V^T)^{-1} (U \Sigma V^T)^T y \\
&= (V \Sigma U^T \cdot U \Sigma V^T)^{-1} \cdot (V \Sigma U^T) y \\
&= (V \Sigma^2 V^T)^{-1} \cdot V \Sigma U^T \cdot y \\
&= (V^T)^{-1} \Sigma^{-2} \cdot V^{-1} \cdot V \Sigma U^T \cdot y \\
&= V \Sigma^{-2} \cdot \Sigma U^T y \\
&= V \Sigma^{-1} U^T y
\end{aligned}$$

$U^T y$: $O(D*N \times N*1) = O(D \cdot N)$

$\Sigma^{-1} U^T y$: $O(D*D \cdot D*1)$

Σ is diagnal $= O(D)$

$V \Sigma^{-1} U^T y$: $O(D*D \cdot D*1)$. $= O(D^2)$

$N > D$, so the dominant complexity is $O(D \cdot N)$

Programming task 10: Dimensionality Reduction

In [2]:
```
import numpy as np
import matplotlib.pyplot as plt

%matplotlib inline
```

Exporting the results to PDF

Once you complete the assignments, export the entire notebook as PDF and attach it to your homework solutions. The best way of doing that is

1. Run all the cells of the notebook.
2. Export/download the notebook as PDF (File -> Download as -> PDF via LaTeX (.pdf)).
3. Concatenate your solutions for other tasks with the output of Step 2. On a Linux machine you can simply use `pdfunite`, there are similar tools for other platforms too. You can only upload a single PDF file to Moodle.

Make sure you are using `nbconvert` Version 5.5 or later by running `jupyter nbconvert --version`. Older versions clip lines that exceed page width, which makes your code harder to grade.

PCA

Given the data in the matrix X your tasks is to:

- Calculate the covariance matrix Σ.
- Calculate eigenvalues and eigenvectors of Σ.
- Plot the original data X and the eigenvectors to a single diagram. What do you observe? Which eigenvector corresponds to the smallest eigenvalue?
- Determine the smallest eigenvalue and remove its corresponding eigenvector. The remaining eigenvector is the basis of a new subspace.
- Transform all vectors in X in this new subspace by expressing all vectors in X in this new basis.

The given data X

In [3]:
```
X = np.array([(-3,-2),(-2,-1),(-1,0),(0,1),
              (1,2),(2,3),(-2,-2),(-1,-1),
              (0,0),(1,1),(2,2), (-2,-3),
              (-1,-2),(0,-1),(1,0), (2,1),(3,2)])
```

Task 1: Calculate the covariance matrix Σ

In [6]:
```
def get_covariance(X):
    """Calculates the covariance matrix of the input data.

    Parameters
    ----------
    X : array, shape [N, D]
        Data matrix.

    Returns
    -------
    Sigma : array, shape [D, D]
        Covariance matrix

    """
    # TODO
    mean = np.mean(X, axis=0)
    X_centered = X - mean
    N = X.shape[0]
    Sigma = (X_centered.T @ X_centered) / N
    return Sigma
```

Task 2: Calculate eigenvalues and eigenvectors of Σ.

In [7]:

```python
def get_eigen(S):
    """Calculates the eigenvalues and eigenvectors of the input matrix.

    Parameters
    ----------
    S : array, shape [D, D]
        Square symmetric positive definite matrix.

    Returns
    -------
    L : array, shape [D]
        Eigenvalues of S
    U : array, shape [D, D]
        Eigenvectors of S

    """
    L, U = np.linalg.eigh(S)

    idx = L.argsort()[::-1]
    L = L[idx]
    U = U[:, idx]

    return L, U
```

Task 3: Plot the original data X and the eigenvectors to a single diagram.

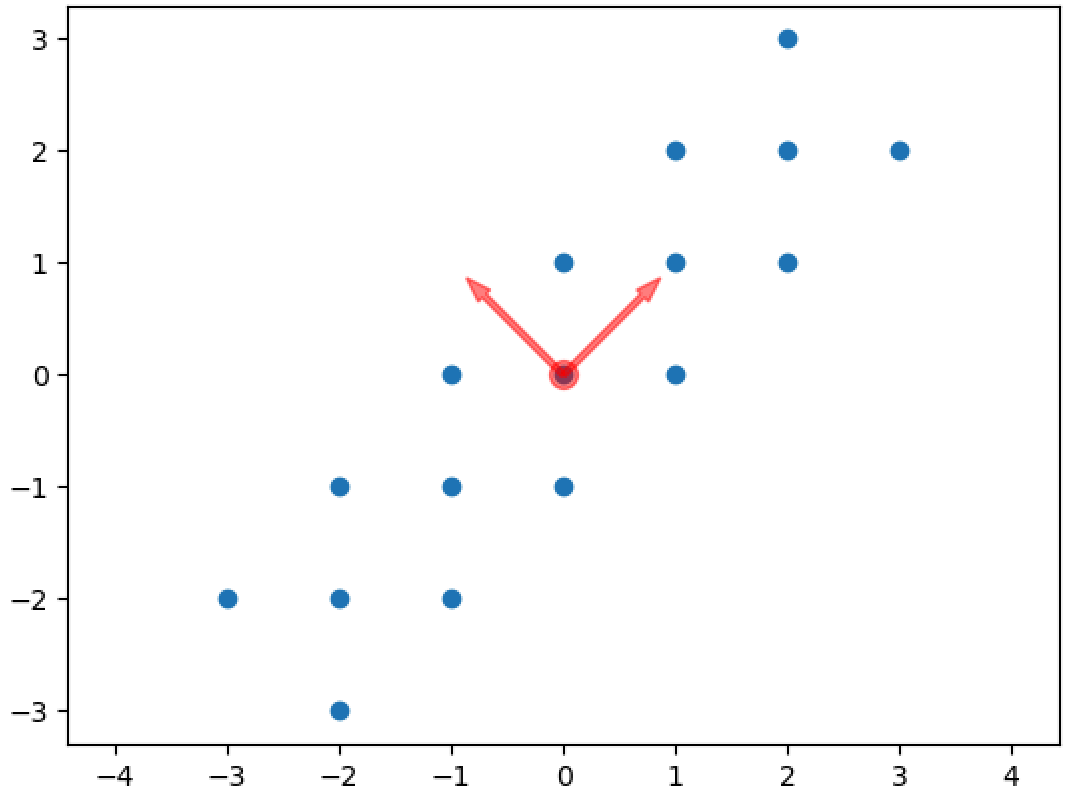
In [8]:

```python
# plot the original data
plt.scatter(X[:, 0], X[:, 1])
plt.axis('equal')

# plot the mean of the data
mean_d1, mean_d2 = X.mean(0)
plt.plot(mean_d1, mean_d2, 'o', markersize=10, color='red', alpha=0.5)

# calculate the covariance matrix
Sigma = get_covariance(X)
# calculate the eigenvector and eigenvalues of Sigma
L, U = get_eigen(Sigma)

plt.arrow(mean_d1, mean_d2, U[0, 0], U[1, 0], width=0.05, color='red', alpha=0.5)
plt.arrow(mean_d1, mean_d2, U[0, 1], U[1, 1], width=0.05, color='red', alpha=0.5);
```



What do you observe in the above plot? Which eigenvector corresponds to the smallest eigenvalue?

Write your answer here:

```
we can see the data can be represented in 2 directions(axis).
The smallest eigenvalue coresponds to the smallest variance in the
dataset, its direction is towards upper left, we need to reduce
this dimension if we want to do PCA.
```

Task 4: Transform the data

Determine the smallest eigenvalue and remove its corresponding eigenvector. The remaining eigenvector is the basis of a new subspace. Transform all vectors in X in this new subspace by expressing all vectors in X in this new basis.

In [11]:
```
def transform(X, U, L):
    """Transforms the data in the new subspace spanned by the eigenvector correspon

    Parameters
    ----------
    X : array, shape [N, D]
        Data matrix.
    L : array, shape [D]
        Eigenvalues of Sigma_X
    U : array, shape [D, D]
        Eigenvectors of Sigma_X
```

```
    Returns
    -------
    X_t : array, shape [N, 1]
        Transformed data

    """
    mean = np.mean(X, axis=0)
    X_centered = X - mean
    U_reduced = U[:, 0:1]
    X_t = X_centered @ U_reduced
    return X_t
```

In [12]:
```
X_t = transform(X, U, L)
```

SVD

Task 5: Given the matrix M find its SVD decomposition $M = U \cdot \Sigma \cdot V$ and reduce it to one dimension using the approach described in the lecture.

In [13]:
```
M = np.array([[1, 2], [6, 3],[0, 2]])
```

In [14]:
```
def reduce_to_one_dimension(M):
    """Reduces the input matrix to one dimension using its SVD decomposition.

    Parameters
    ----------
    M : array, shape [N, D]
        Input matrix.

    Returns
    -------
    M_t: array, shape [N, 1]
        Reduce matrix.

    """
    U, S, Vt = np.linalg.svd(M, full_matrices=False)

    U_reduced = U[:, 0:1]   #[N,1]
    s_reduced = S[0]    #scalar

    # 计算降维后的矩阵
    M_t = U_reduced * s_reduced

    return M_t
```

In [10]:
```
M_t = reduce_to_one_dimension(M)
```