Dimensionality Reduction - Introduction

- In many learning problems, the datasets have large number of variables.
- Sometimes, the number of variables is more than the number of observations.
- For example, in many scientific fields such as
  - image processing,
  - time series analysis,
  - Internet search engines, and
  - automatic text analysis.

Dimensionality Reduction - Introduction

- Statistical and machine learning methods have some difficulty when dealing with such high-dimensional data.
- Normally the number of input variables is reduced before the machine learning algorithms can be successfully applied.
- In statistical and machine learning, dimensionality reduction or dimension reduction is the process of reducing the number of variables under consideration by obtaining a smaller set of principal variables.

Dimensionality Re

10 ⇒ 7

Feature selection

- In feature selection, we are interested in finding k of the total of n features that give us the most information and we discard the other (n−k) dimensions.

Dimensionality Reduction - Types

Feature extraction

- In feature extraction, we are interested in finding a new set of k features that are the combination of the original n features.
- These methods may be supervised or unsupervised depending on whether or not they use the output information.
- The best known and most widely used feature extraction methods are Principal Components Analysis (PCA) and Linear Discriminant Analysis (LDA).

Dimensionality Reduction - Measures of error

- In both methods we require a measure of the error in the model.
- In regression problems, we may use the
  - **Mean Squared Error (MSE)** or the
  - **Root Mean Squared Error (RMSE)**

Dimensionality Reduction - Measures of error

- MSE is the sum, over all the data points, of the square of the difference between the predicted and actual target variables, divided by the number of data points.
- If $y_1, y_2 \dots . y_n$ are the observed values and $\hat{y}_1, \hat{y}_2 \dots . \hat{y}_n$ are the predicted values, then

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$$

Dimensionality Reduction - Measures of error

- In classification problems, we may use the misclassification rate as a measure of the error.

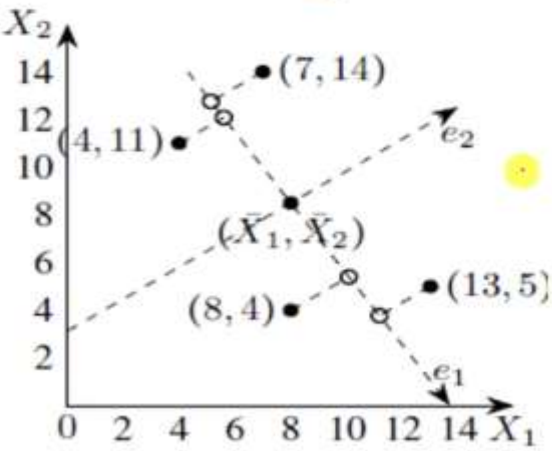
10 → 6 4

- This is defined as follows:

$\frac{6}{10}$ = .06

$$\text{misclassification rate} = \frac{\text{no. of misclassified examples}}{\text{total no. of examples}}$$

Principle Component Analysis



Solved Example

Principle Component Analysis – Solved Example

- Given the data in Table, reduce the dimension from 2 to 1 using the Principal Component Analysis (PCA) algorithm.

Feature	Example 1	Example 2	Example 3	Example 4
X_1	4	8	13	7
X_2	11	4	5	14

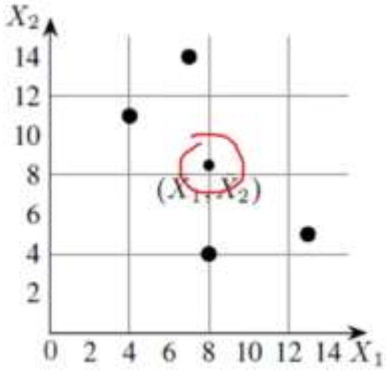
Principle Component Analysis – Solved Example

Step 1: Calculate Mean

$$\bar{X}_1 = \frac{1}{4}(4 + 8 + 13 + 7) = 8,$$

$$\bar{X}_2 = \frac{1}{4}(11 + 4 + 5 + 14) = 8.5.$$

F	Ex 1	Ex 2	Ex 3	Ex 4
X_1	4	8	13	7
X_2	11	4	5	14



Principle Component Analysis – Solved Example

Step 2: Calculation of the covariance matrix.

F	Ex 1	Ex 2	Ex 3	Ex 4
X_1	4	8	13	7
X_2	11	4	5	14

$$S = \begin{bmatrix} \mathrm{Cov}(X_1, X_1) & \mathrm{Cov}(X_1, X_2) \\ \mathrm{Cov}(X_2, X_1) & \mathrm{Cov}(X_2, X_2) \end{bmatrix}$$

$$\overline{X_1} = 8$$

$$\overline{X_2} = 8.5$$

$$\begin{aligned} \mathrm{Cov}(X_1, X_1) &= \frac{1}{N-1}\sum_{k=1}^{N}(X_{1k} - \bar{X}_1)(X_{1k} - \bar{X}_1) \\ &= \tfrac{1}{3}\left((4-8)^2 + (8-8)^2 + (13-8)^2 + (7-8)^2\right) \\ &= 14 \end{aligned}$$

Principle Component Analysis – Solved Example

Step 2: Calculation of the covariance matrix.

F	Ex 1	Ex 2	Ex 3	Ex 4
X_1	4	8	13	7
X_2	11	4	5	14

$$S = \begin{bmatrix} \mathrm{Cov}(X_1, X_1) & \mathrm{Cov}(X_1, X_2) \\ \mathrm{Cov}(X_2, X_1) & \mathrm{Cov}(X_2, X_2) \end{bmatrix}$$

$$\overline{X_1} = 8$$

$$\overline{X_2} = 8.5$$

$$\begin{aligned} \mathrm{Cov}(X_1, X_2) &= \frac{1}{N-1}\sum_{k=1}^{N}(X_{1k} - \bar{X}_1)(X_{2k} - \bar{X}_2) \\ &= \tfrac{1}{3}((4-8)(11-8.5) + (8-8)(4-8.5) \\ &\quad + (13-8)(5-8.5) + (7-8)(14-8.5) \\ &= -11 \end{aligned}$$

Principle Component Analysis – Solved Example

Step 2: Calculation of the covariance matrix.

F	Ex 1	Ex 2	Ex 3	Ex 4
X_1	4	8	13	7
X_2	11	4	5	14

$$\overline{X_1} = \mathbf{8}$$

$$\overline{X_2} = \mathbf{8.5}$$

$$S = \begin{bmatrix} \mathrm{Cov}(X_1, X_1) & \mathrm{Cov}\,(X_1, X_2) \\ \mathrm{Cov}\,(X_2, X_1) & \mathrm{Cov}\,(X_2, X_2) \end{bmatrix}$$

$$\begin{aligned} \mathrm{Cov}\,(X_2, X_1) &= \mathrm{Cov}\,(X_1, X_2) \\ &= -11 \end{aligned}$$

$$\begin{aligned} \mathrm{Cov}\,(X_2, X_2) &= \frac{1}{N-1}\sum_{k=1}^{N}(X_{2k} - \bar{X}_2)(X_{2k} - \bar{X}_2) \\ &= \tfrac{1}{3}\left((11-8.5)^2 + (4-8.5)^2 + (5-8.5)^2 + (14-8.5)^2\right) \\ &= 23 \end{aligned}$$

Principle Component Analysis – Solved Example

F	Ex 1	Ex 2	Ex 3	Ex 4
X_1	4	8	13	7
X_2	11	4	5	14

$\overline{X_1} = 8$

$\overline{X_2} = 8.5$

$$S = \begin{bmatrix} 14 & -11 \\ -11 & 23 \end{bmatrix}$$

Step 3: Eigenvalues of the covariance matrix

The characteristic equation of the covariance matrix is,

$$\begin{aligned} 0 &= \det(S - \lambda I) \\ &= \begin{vmatrix} 14-\lambda & -11 \\ -11 & 23-\lambda \end{vmatrix} \\ &= (14-\lambda)(23-\lambda) - (-11)\times(-11) \\ &= \lambda^2 - 37\lambda + 201 \end{aligned}$$

$\begin{matrix} 1 & 0 \\ 0 & 1 \end{matrix}$ $\begin{matrix} \lambda & 0 \\ 0 & \lambda \end{matrix}$

Principle Component Anal

Step 2: Calculation of the covariance matrix.

F	Ex 1	Ex 2	Ex 3	Ex 4
X_1	4	8	13	7
X_2	11	4	5	14

$\overline{X_1} = 8$

$\overline{X_2} = 8.5$

$$S = \begin{bmatrix} \mathrm{Cov}(X_1, X_1) & \mathrm{Cov}(X_1, X_2) \\ \mathrm{Cov}(X_2, X_1) & \mathrm{Cov}(X_2, X_2) \end{bmatrix} = \begin{bmatrix} 14 & -11 \\ -11 & 23 \end{bmatrix}$$

Step 3: Eigenvalues of the covariance matrix

F	Ex 1	Ex 2	Ex 3	Ex 4
X_1	4	8	13	7
X_2	11	4	5	14

The characteristic equation of the covariance matrix is,

$$\begin{aligned} 0 &= \det(S - \lambda I) \\ &= \begin{vmatrix} 14-\lambda & -11 \\ -11 & 23-\lambda \end{vmatrix} \\ &= (14-\lambda)(23-\lambda) - (-11)\times(-11) \\ &= \lambda^2 - 37\lambda + 201 \end{aligned}$$

$$\begin{aligned} \lambda &= \tfrac{1}{2}(37 \pm \sqrt{565}) \\ &= 30.3849,\ 6.6151 \\ &= \lambda_1,\ \lambda_2 \quad \text{(say)} \end{aligned}$$

$\overline{X_1} = 8$

$\overline{X_2} = 8.5$

$$S = \begin{bmatrix} 14 & -11 \\ -11 & 23 \end{bmatrix}$$

Principle Component Analysis – Solved Example

Step 4: Computation of the eigenvectors

F	Ex 1	Ex 2	Ex 3	Ex 4
X_1	4	8	13	7
X_2	11	4	5	14

$$U = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$

$$\begin{bmatrix} 0 \\ 0 \end{bmatrix} = (S - \lambda I)\, U$$

$$= \begin{bmatrix} 14-\lambda & -11 \\ -11 & 23-\lambda \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$

$$= \begin{bmatrix} (14-\lambda)u_1 - 11u_2 \\ -11u_1 + (23-\lambda)u_2 \end{bmatrix}$$

$(14-\lambda)u_1 = 11u_2$

$$(14-\lambda)u_1 - 11u_2 = 0$$

$$-11u_1 + (23-\lambda)u_2 = 0$$

$$\frac{u_1}{11} = \frac{u_2}{14-\lambda} = t$$

$\overline{X_1} = 8$

$\overline{X_2} = 8.5$

$$S = \begin{bmatrix} 14 & -11 \\ -11 & 23 \end{bmatrix}$$

$\lambda_1 = 30.3849$

$\lambda_2 = 6.6151$

Principle Component Analysis – Solved Example

Step 4: Computation of the eigenvectors

F	Ex 1	Ex 2	Ex 3	Ex 4
X_1	4	8	13	7
X_2	11	4	5	14

$$U = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$

$$\frac{u_1}{11} = \frac{u_2}{14-\lambda} = t$$

$u_1 = 11t, \quad u_2 = (14-\lambda)t$

$$U = \begin{bmatrix} 11 \\ 14-\lambda \end{bmatrix}.$$

$\overline{X_1} = 8$

$\overline{X_2} = 8.5$

$$S = \begin{bmatrix} 14 & -11 \\ -11 & 23 \end{bmatrix}$$

$\lambda_1 = 30.3849$

$\lambda_2 = 6.6151$

Principle Component Analysis – Solved Example

Step 4: Computation of the eigenvectors

F	Ex 1	Ex 2	Ex 3	Ex 4
X_1	4	8	13	7
X_2	11	4	5	14

$$U_1 = \begin{bmatrix} 11 \\ 14 - \lambda_1 \end{bmatrix}.$$

$\overline{X_1} = 8$

$\overline{X_2} = 8.5$

- To find a unit eigenvector, we compute the length of U_1 which is given by,

$$S = \begin{bmatrix} 14 & -11 \\ -11 & 23 \end{bmatrix}$$

$$\begin{aligned} \|U_1\| &= \sqrt{11^2 + (14 - \lambda_1)^2} \\ &= \sqrt{11^2 + (14 - 30.3849)^2} \\ &= 19.7348 \end{aligned}$$

$\lambda_1 = 30.3849$

$\lambda_2 = 6.6151$

Principle Component Analysis – Solved Example

Step 4: Computation of the eigenvectors

F	Ex 1	Ex 2	Ex 3	Ex 4
X_1	4	8	13	7
X_2	11	4	5	14

$$U_1 = \begin{bmatrix} 11 \\ 14 - \lambda_1 \end{bmatrix}.$$

$\overline{X_1} = 8$

$\overline{X_2} = 8.5$

- To find a unit eigenvector, we compute the length of U_1 which is given by,

$$e_1 = \begin{bmatrix} 11/\|U_1\| \\ (14 - \lambda_1)/\|U_1\| \end{bmatrix} = \begin{bmatrix} 11/19.7348 \\ (14 - 30.3849)/19.7348 \end{bmatrix} = \begin{bmatrix} 0.5574 \\ -0.8303 \end{bmatrix}$$

e_1

$$S = \begin{bmatrix} 14 & -11 \\ -11 & 23 \end{bmatrix}$$

$$\begin{aligned} \|U_1\| &= \sqrt{11^2 + (14 - \lambda_1)^2} \\ &= \sqrt{11^2 + (14 - 30.3849)^2} \\ &= 19.7348 \end{aligned}$$

$\lambda_1 = 30.3849$

$\lambda_2 = 6.6151$

Principle Component Analysis – Solved Example

Step 4: Computation of the eigenvectors

F	Ex 1	Ex 2	Ex 3	Ex 4
X_1	4	8	13	7
X_2	11	4	5	14

$$U_1 = \begin{bmatrix} 11 \\ 14 - \lambda_1 \end{bmatrix}$$

$\overline{X_1} = 8$

$\overline{X_2} = 8.5$

- To find a unit eigenvector, we compute the length of U_1 which is given by,

$$\|U_1\| = \sqrt{11^2 + (14 - \lambda_1)^2}$$
$$= \sqrt{11^2 + (14 - 30.3849)^2}$$
$$= 19.7348$$

$$e_1 = \begin{bmatrix} 11/\|U_1\| \\ (14 - \lambda_1)/\|U_1\| \end{bmatrix} = \begin{bmatrix} 11/19.7348 \\ (14 - 30.3849)/19.7348 \end{bmatrix}$$

$$e_1 = \begin{bmatrix} 0.5574 \\ -0.8303 \end{bmatrix}$$

$$e_2 = \begin{bmatrix} 0.8303 \\ 0.5574 \end{bmatrix}$$

$$S = \begin{bmatrix} 14 & -11 \\ -11 & 23 \end{bmatrix}$$

$\lambda_1 = 30.3849$

$\lambda_2 = 6.6151$

Principle Component Analysis – Solved Example

Step 5: Computation of first principal components

F	Ex 1	Ex 2	Ex 3	Ex 4
X_1	4	8	13	7
X_2	11	4	5	14

$$e_1^T \begin{bmatrix} X_{1k} - \bar{X}_1 \\ X_{2k} - \bar{X}_2 \end{bmatrix}$$

$$e_1 = \begin{bmatrix} 0.5574 \\ -0.8303 \end{bmatrix}$$

$\overline{X_1} = 8$

$\overline{X_2} = 8.5$

$$e_1^T \begin{bmatrix} X_{1k} - \bar{X}_1 \\ X_{2k} - \bar{X}_2 \end{bmatrix} = \begin{bmatrix} 0.5574 & -0.8303 \end{bmatrix} \begin{bmatrix} X_{11} - \bar{X}_1 \\ X_{21} - \bar{X}_2 \end{bmatrix}$$
$$= 0.5574(X_{11} - \bar{X}_1) - 0.8303(X_{21} - \bar{X}_2)$$
$$= 0.5574(4 - 8) - 0.8303(11 - 8,5)$$
$$= -4.30535$$

$$e_2 = \begin{bmatrix} 0.8303 \\ 0.5574 \end{bmatrix} \quad S = \begin{bmatrix} 14 & -11 \\ -11 & 23 \end{bmatrix}$$

$\lambda_1 = 30.3849$

$\lambda_2 = 6.6151$

Principle Component Analysis – Solved Example

Step 5: Computation of first principal components

F	Ex 1	Ex 2	Ex 3	Ex 4
X_1	4	8	13	7
X_2	11	4	5	14

Feature	Ex 1	Ex 2	Ex 3	Ex 4
X_1	4	8	13	7
X_2	11	4	5	14
First Principle Components	-4.3052	3.7361	5.6928	-5.1238

$$\overline{X_1} = 8$$

$$\overline{X_2} = 8.5$$

$$S = \begin{bmatrix} 14 & -11 \\ -11 & 23 \end{bmatrix}$$

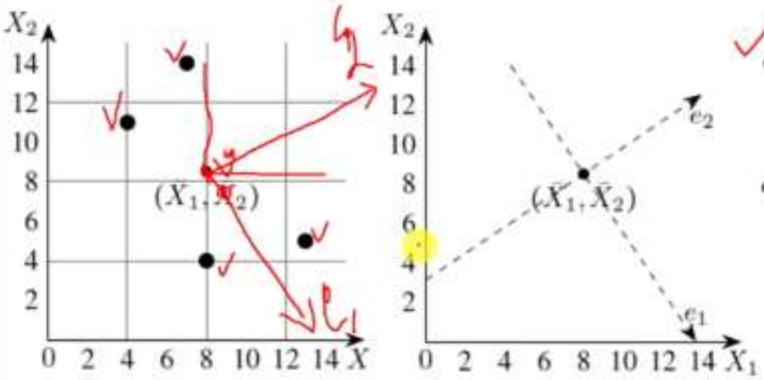
$$\lambda_1 = 30.3849$$

$$\lambda_2 = 6.6151$$

Principle Component Analysis – Solved Example

Step 6: Geometrical meaning of first principal components

F	Ex 1	Ex 2	Ex 3	Ex 4
X_1	4	8	13	7
X_2	11	4	5	14



$$e_1 = \begin{bmatrix} 0.5574 \\ -0.8303 \end{bmatrix}$$

$$e_2 = \begin{bmatrix} 0.8303 \\ 0.5574 \end{bmatrix}$$

$$\overline{X_1} = 8$$

$$\overline{X_2} = 8.5$$

$$S = \begin{bmatrix} 14 & -11 \\ -11 & 23 \end{bmatrix}$$

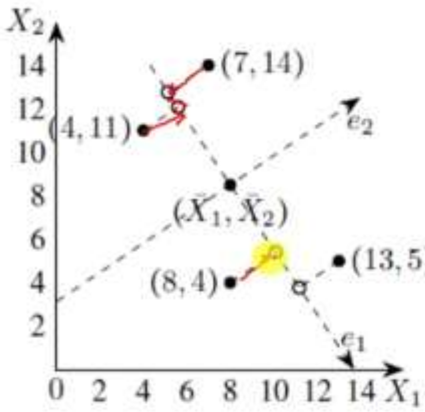
$$\lambda_1 = 30.3849$$

$$\lambda_2 = 6.6151$$

Principle Component Analysis – Solved Example

Step 6: Geometrical meaning of first principal components

F	Ex 1	Ex 2	Ex 3	Ex 4
X_1	4	8	13	7
X_2	11	4	5	14



$e_1 = \begin{bmatrix} 0.5574 \\ -0.8303 \end{bmatrix}$

$e_2 = \begin{bmatrix} 0.8303 \\ 0.5574 \end{bmatrix}$

$\overline{X_1} = 8$

$\overline{X_2} = 8.5$

$S = \begin{bmatrix} 14 & -11 \\ -11 & 23 \end{bmatrix}$

$\lambda_1 = 30.3849$

$\lambda_2 = 6.6151$

Principle Component Analysis – Solved Example

Step 5: Computation of first principal components

F	Ex 1	Ex 2	Ex 3	Ex 4
X_1	4	8	13	7
X_2	11	4	5	14

Feature	Ex 1	Ex 2	Ex 3	Ex 4
X_1	4	8	13	7
X_2	11	4	5	14
First Principle Components	-4.3052	3.7361	5.6928	-5.1238

$\overline{X_1} = 8$

$\overline{X_2} = 8.5$

$S = \begin{bmatrix} 14 & -11 \\ -11 & 23 \end{bmatrix}$

$\lambda_1 = 30.3849$

$\lambda_2 = 6.6151$