# Number Recognition

Mooseok Jang

December 12, 2007

# Table of Contents

# What is it?

# Training and Recognition

- Training
  - Training sample $\rightarrow$ Data ($\in$ $n$-dimensional space)
  - Same number(training sample) $\rightarrow$ specific pattern
  - Building standard!
- Recognition
  - Determine class what test sample is in

# Training sample $\longrightarrow$ Data

The training sample(1, 0) in $4 \times 4$ form

$$\begin{matrix} 0 & x & 0 & 0 \\ 0 & x & 0 & 0 \\ 0 & x & 0 & 0 \\ 0 & x & 0 & 0 \end{matrix} \qquad \begin{matrix} 0 & x & x & x \\ 0 & x & 0 & x \\ 0 & x & 0 & x \\ 0 & x & x & x \end{matrix}$$

convert...

# Training sample $\longrightarrow$ Data

The training sample(1, 0) in $4 \times 4$ form

$$\begin{matrix} 0 & x & 0 & 0 \\ 0 & x & 0 & 0 \\ 0 & x & 0 & 0 \\ 0 & x & 0 & 0 \end{matrix} \qquad \begin{matrix} 0 & x & x & x \\ 0 & x & 0 & x \\ 0 & x & 0 & x \\ 0 & x & x & x \end{matrix}$$

convert...

$$\begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix}$$

$$\begin{bmatrix} 0 & 1 & 1 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 1 & 1 \end{bmatrix}$$

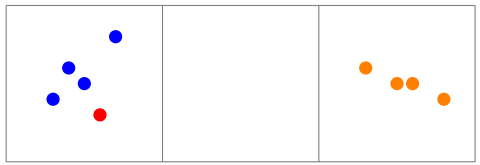
can be mapped to 16-dimensional space

# Scheme

- Suppose training sample is given by

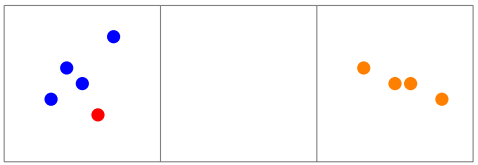
# Scheme

- Suppose training sample is given by



# Scheme

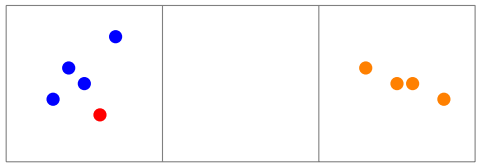
- Suppose training sample is given by



- Data(from training sample)

# Scheme

- Suppose training sample is given by



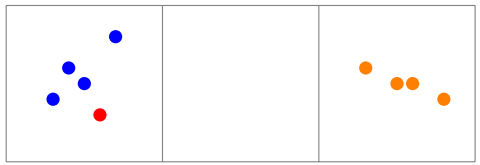
- Data(from training sample)

$$X_{\text{blue}} = [(0.5, 0.5), (0.4, 0.6), (0.3, 0.4), (0.7, 0.8)] \quad (1)$$
$$X_{\text{orange}} = [(2.6, 0.5), (2.3, 0.6), (2.5, 0.5), (2.8, 0.4)] \quad (2)$$

subscript of $X$ = the name of class
red point $(0.6, 0.3)$ = test sample

# Scheme

- Suppose training sample is given by



- Data(from training sample)

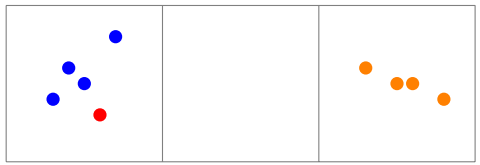
$$X_{\text{blue}} = [(0.5, 0.5), (0.4, 0.6), (0.3, 0.4), (0.7, 0.8)] \quad (1)$$
$$X_{\text{orange}} = [(2.6, 0.5), (2.3, 0.6), (2.5, 0.5), (2.8, 0.4)] \quad (2)$$

subscript of $X$ = the name of class
red point $(0.6, 0.3)$ = test sample

- By training, standard = first value of data
  between 0 and 1 $\rightarrow$ "blue" class
  between 2 and 3 $\rightarrow$ "orange" class

# Scheme

- Suppose training sample is given by



- Data(from training sample)

$$X_{\text{blue}} = [(0.5, 0.5), (0.4, 0.6), (0.3, 0.4), (0.7, 0.8)] \quad (1)$$
$$X_{\text{orange}} = [(2.6, 0.5), (2.3, 0.6), (2.5, 0.5), (2.8, 0.4)] \quad (2)$$

subscript of $X$ = the name of class
red point $(0.6, 0.3)$ = test sample

- By training, standard = first value of data
  between 0 and 1 $\rightarrow$ "blue" class
  between 2 and 3 $\rightarrow$ "orange" class
- By recognition, red point $\in$ "blue" class($\because 0 < 0.6 < 1$)

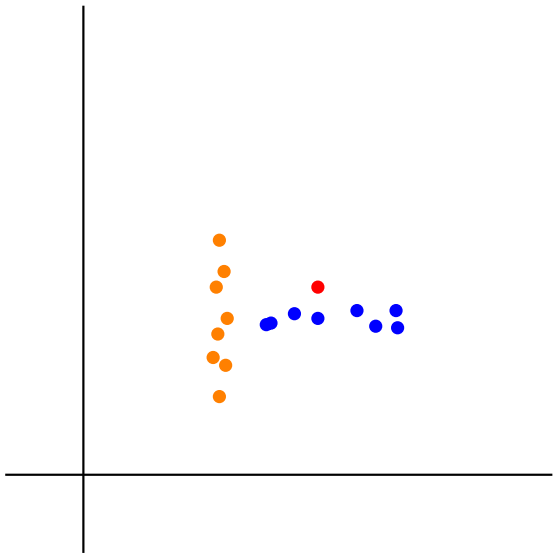
# Most simple method

# Most simple method

Standard : Compare $|X - E(X_i)|$!

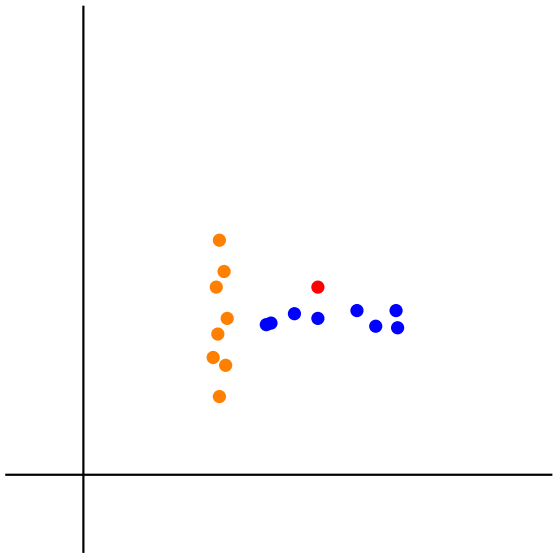
# Most simple method

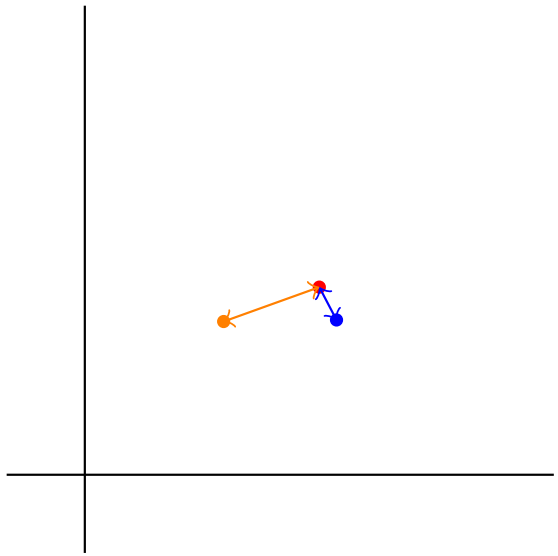
Standard : Compare $|X - E(X_i)|$!



# Most simple method

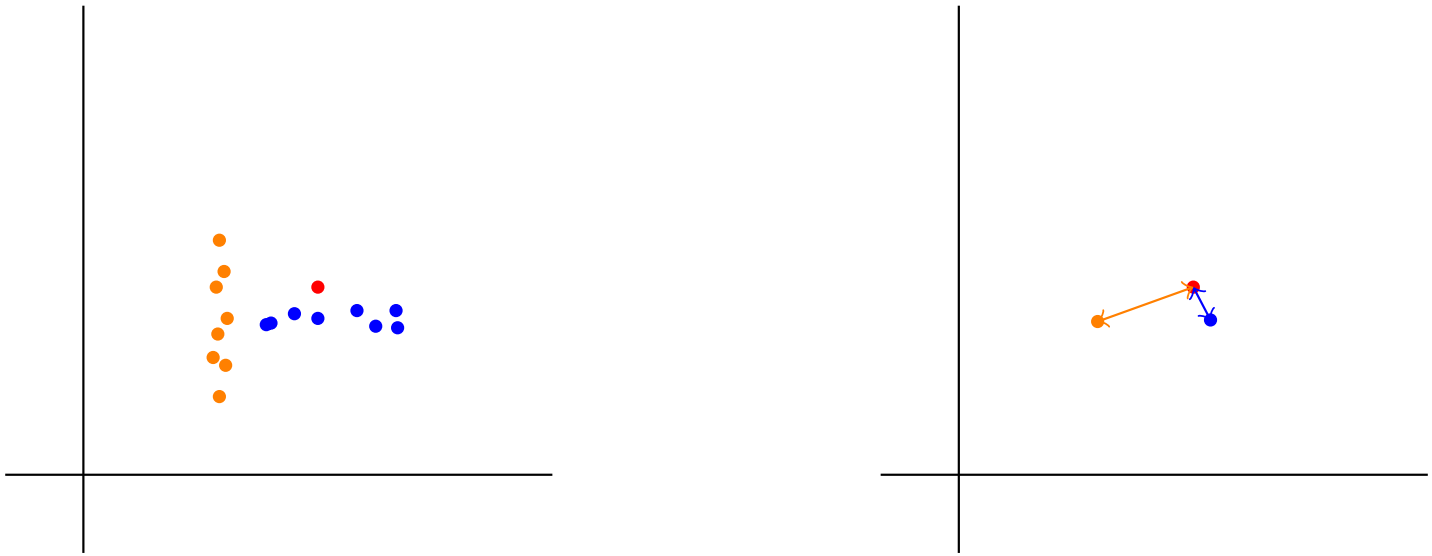
Standard : Compare $|X - E(X_i)|$!





# Most simple method

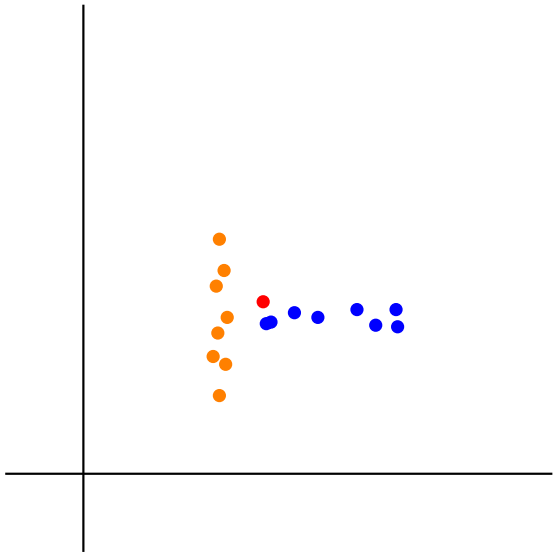
Standard : Compare $|X - E(X_i)|$!



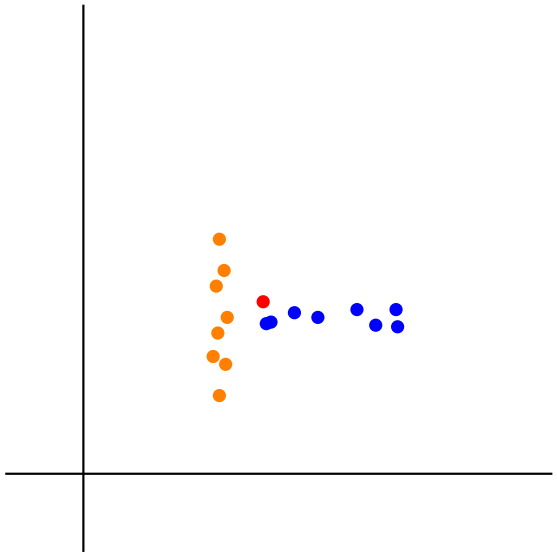
Standard says, and it's real

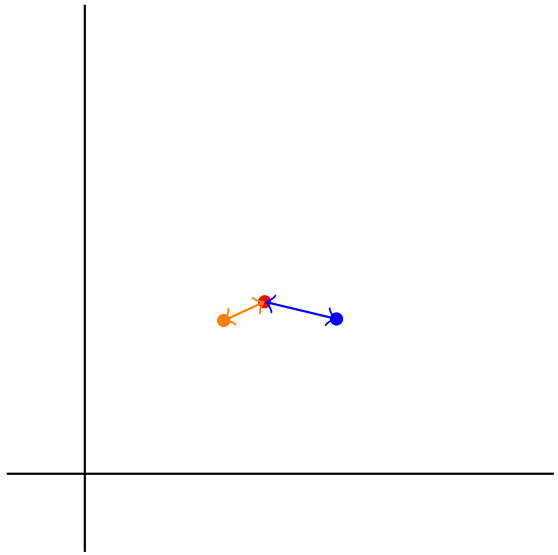
red point $\in$ "blue" class

# Limitation

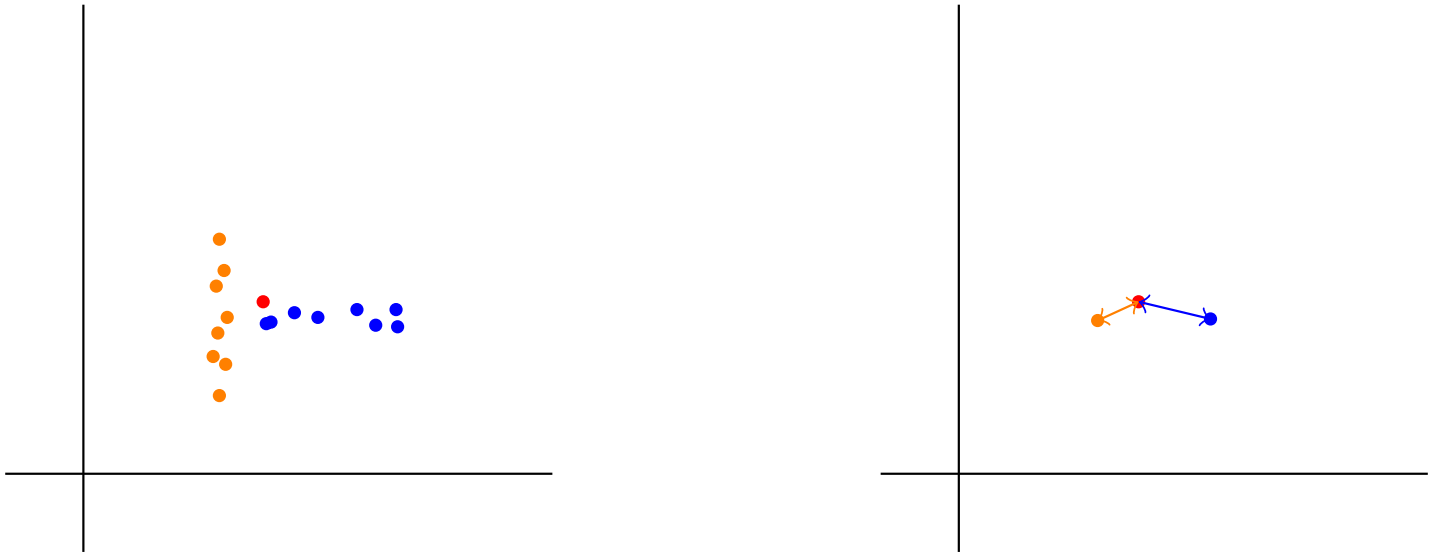


# Limitation





# Limitation



Standard says, but not real

$$\text{red point} \in \text{"blue" class}$$

# Covariance Matrix

- Definition of covariance matrix

$$C_{ij} = E\left[\left(X_i - E\left(X_i\right)\right)\left(X_j - E\left(X_j\right)\right)\right] \tag{3}$$

  Set

$$\lambda_{ik}, \lambda_{jk}, V_{ik}, V_{jk} \tag{4}$$

  as eigenvalue of $X_i$, $X_j$ and eigenvector of $X_i$, $X_j$ respectively.
- Large $\lambda_{ik} \longrightarrow$ sample data is well distributed along eigenvector $V_{ik}$ through long range.
- Dominant eigenvector : its eigenvalue largest one

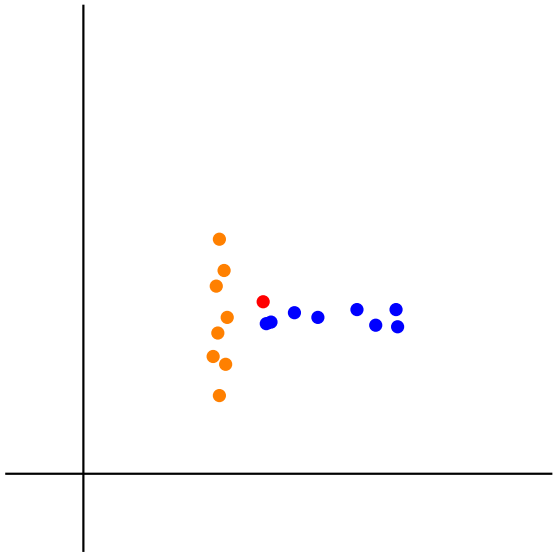
# The method using dominant eigenvector

# The method using dominant eigenvector

Standard : Compare distance to dominant eigenvector!

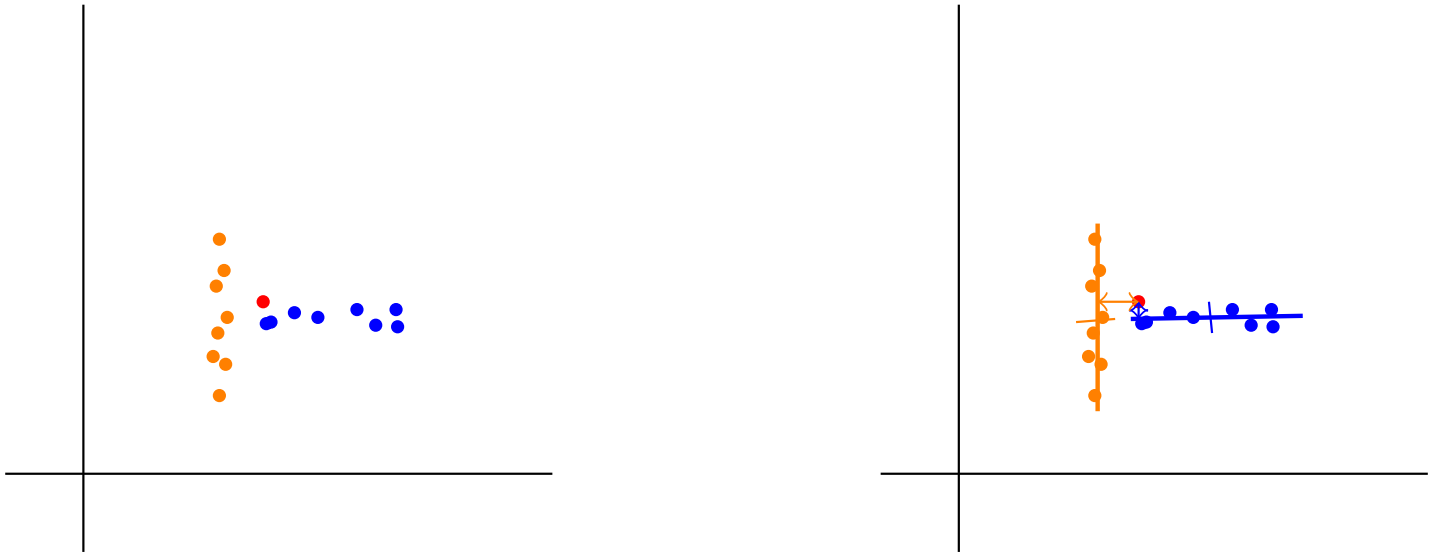
# The method using dominant eigenvector

Standard : Compare distance to dominant eigenvector!



# The method using dominant eigenvector

Standard : Compare distance to dominant eigenvector!



Standard says, and it's real

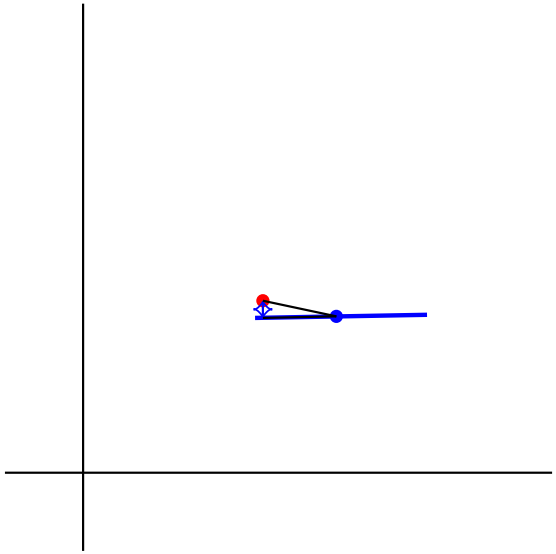
$$\text{red point} \in \text{"blue" class}$$

# The method using dominant eigenvector

Evaluate real distance by equation

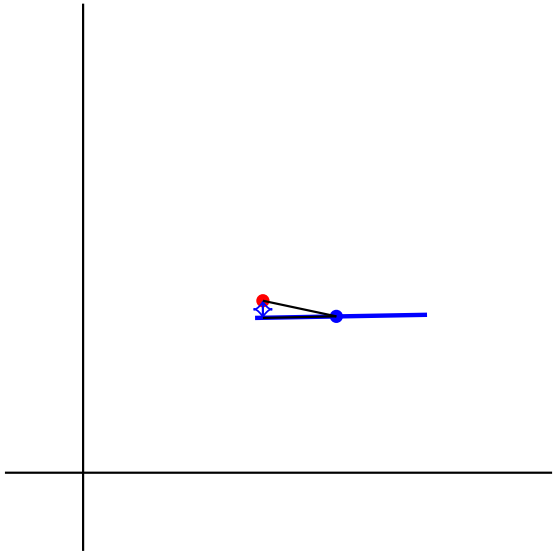
# The method using dominant eigenvector

Evaluate real distance by equation



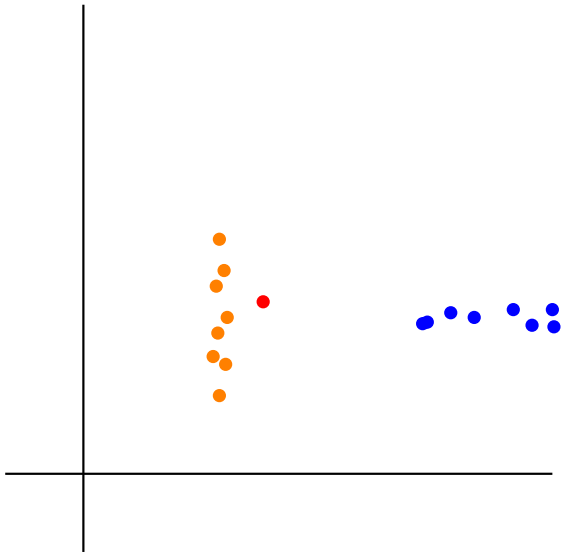
# The method using dominant eigenvector

Evaluate real distance by equation

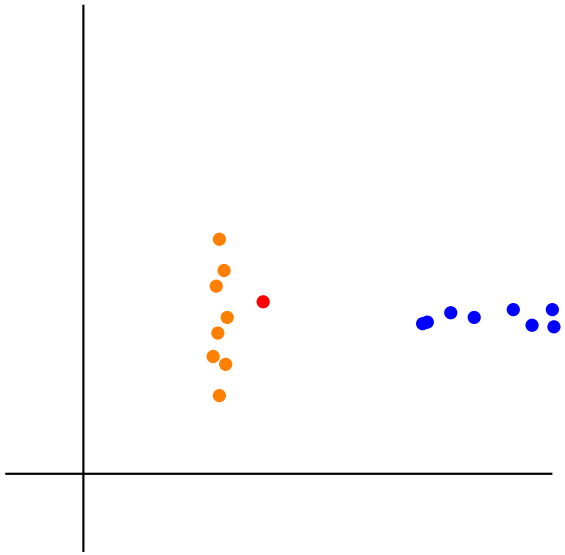


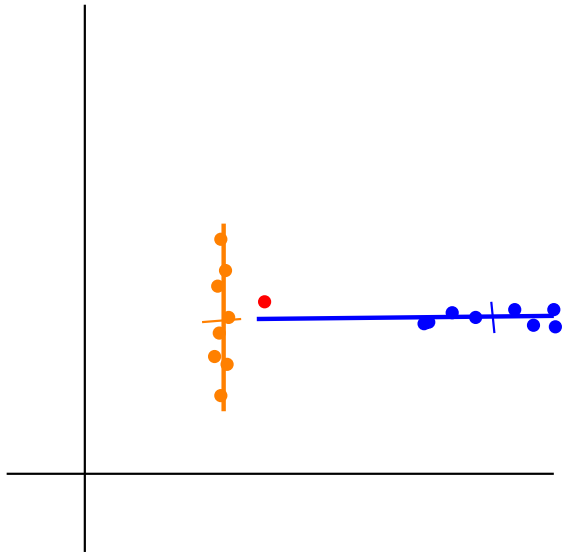
$$d^2 = |X - E(X_i)|^2 - \left(X \cdot V_{i(dominant)}\right)^2 \quad (5)$$

# Limitation

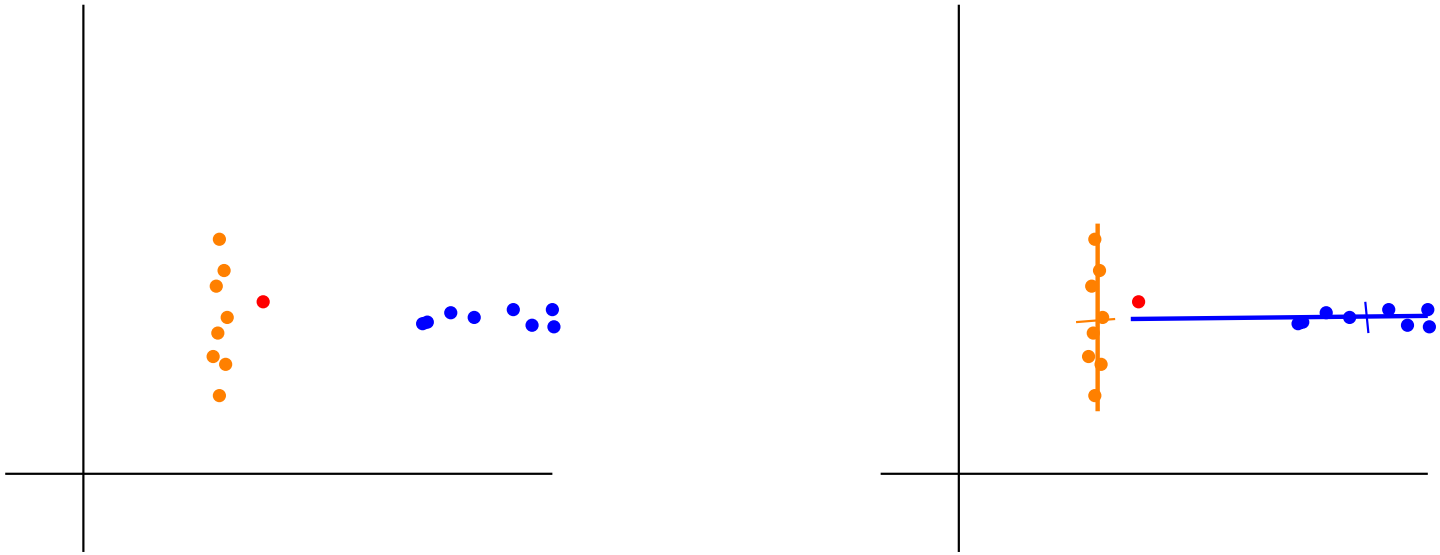


# Limitation





# Limitation



Standard says, but not real

$$\text{red point} \in \text{"blue" class}$$

The method using all eigenvectors will be give the solution of this problem.

# Probable application

Handwriting or any analogue data can be stored in electronic devices as meaningful code.