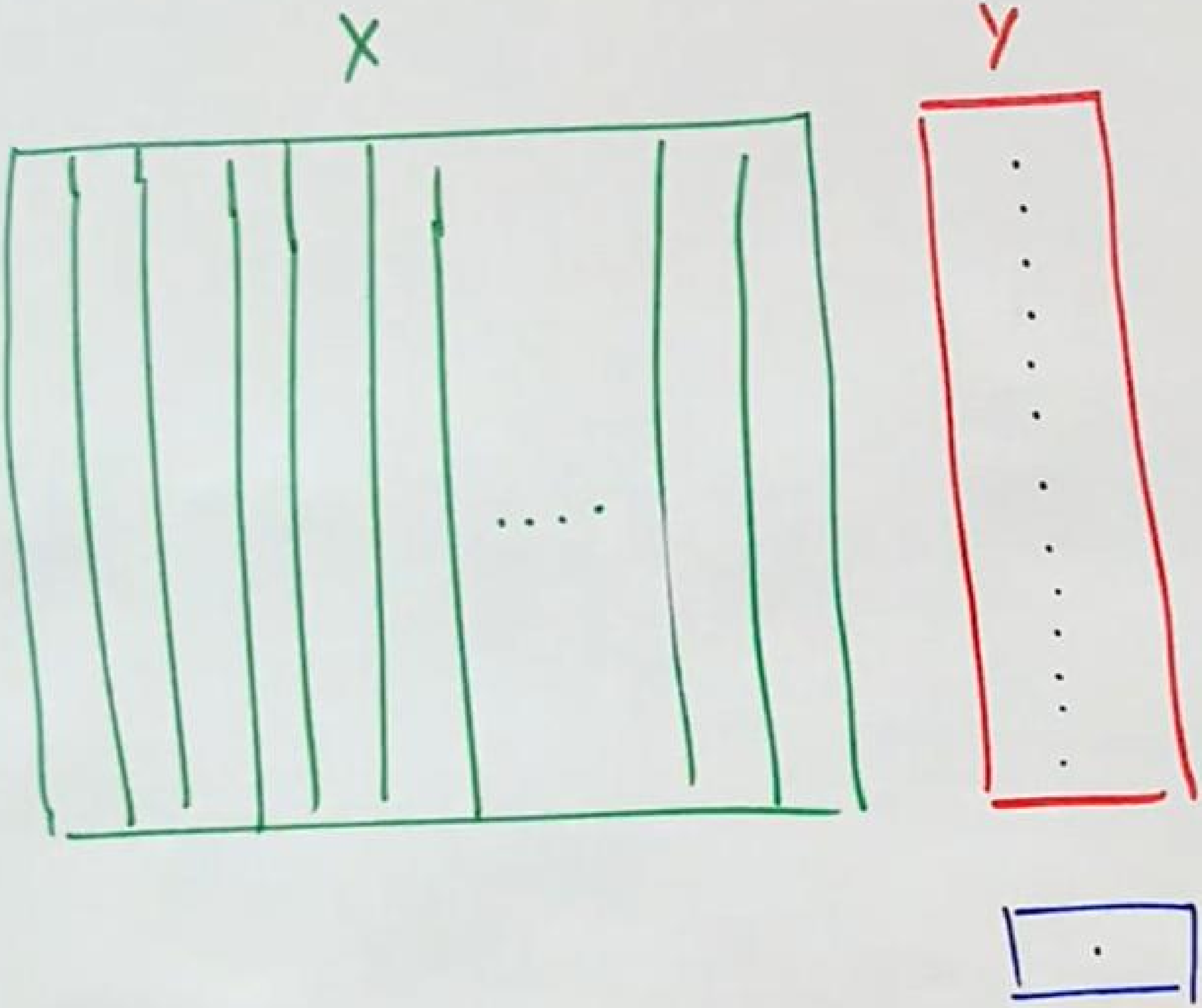
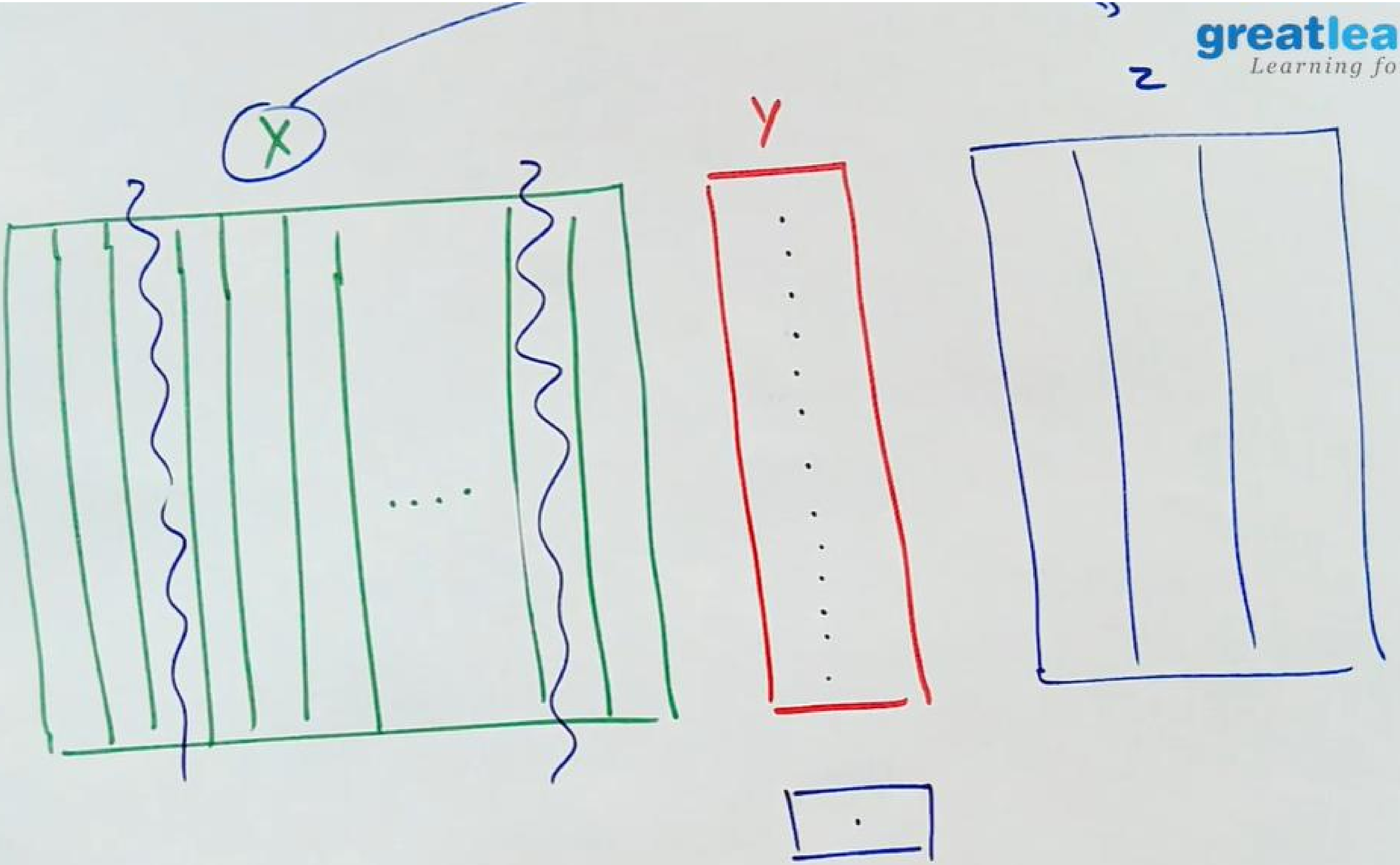
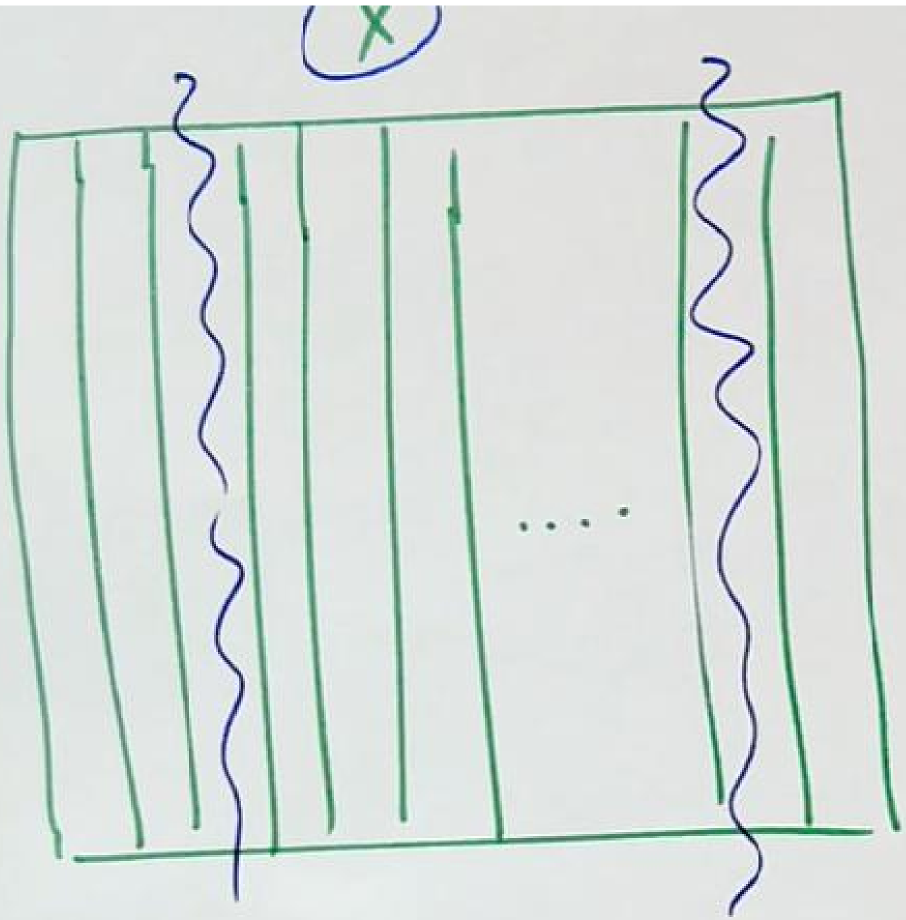

X
Y

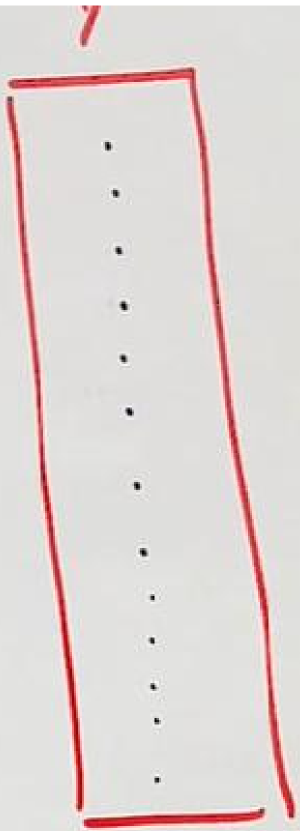
# Dimensionality Reduction

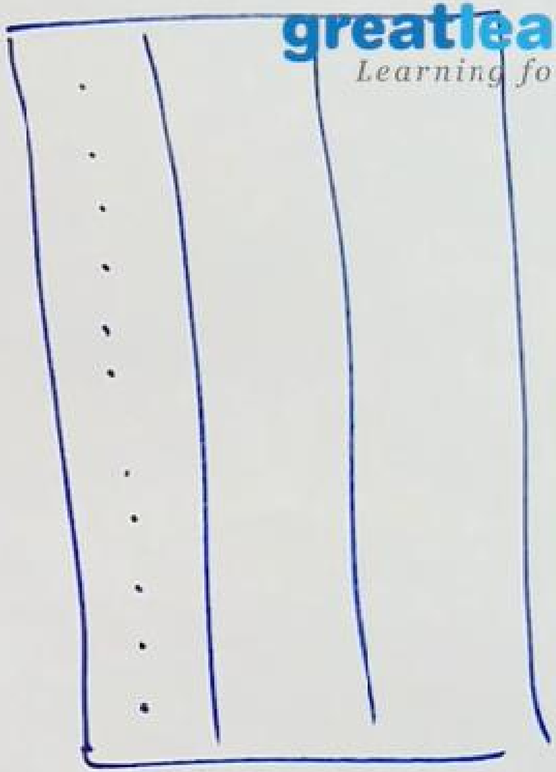
- The process of reducing the number of independent variables
- Reducing dimensionality of independent variables helps in many ways
  - removes multi-collinearity to improve ML model performance
  - decreases computational times for fitting models
  - makes visualization easier
  - decreases storage requirements
  - avoids curse of dimensionality
  - helps reduce over fitting
- Hence dimensionality reduction plays a significant role in analyzing data

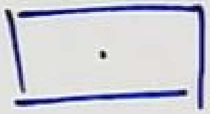
- Feature elimination
  - Simply identify and remove variables (columns) that are not important
  - The disadvantage is that we would gain no insight from those dropped variables and loose any information they contain
- Feature extraction
  - Create a few new variables from the old variables
  - **PCA** Principal Component Analysis: is the most popular feature extraction technique


X
Y
z



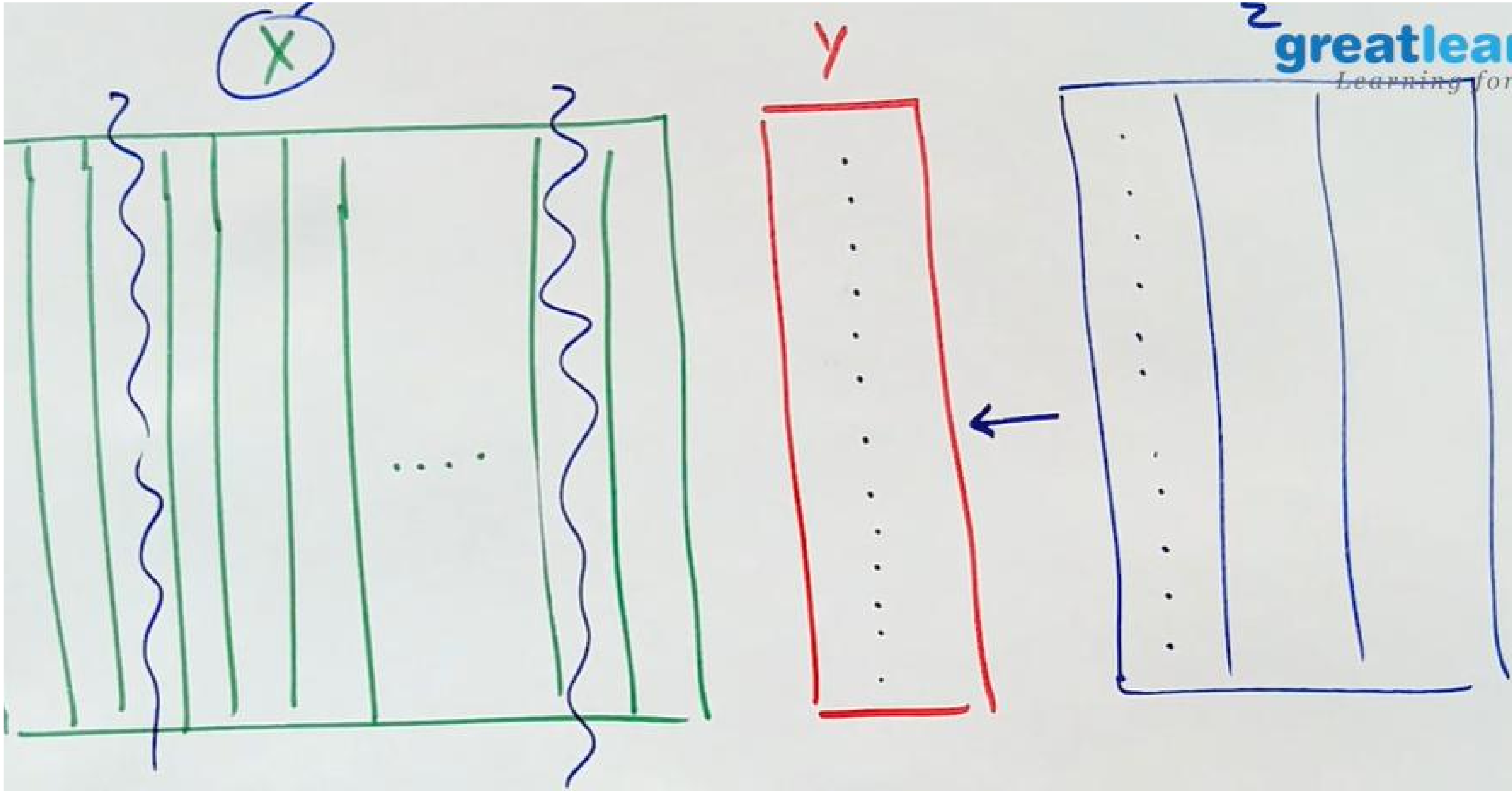


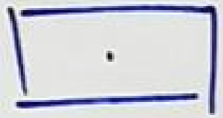
$$z_1 = 0.2\,x_1 + 0.4\,x_2 + 0.6\,x_7$$

$$\cancel{z_2 = x_3^{0.7} \times x}$$

$$z_2 = x_3^{0.2} \times x_4^{0.5} \times \dots \times x_6^{0.3}$$

- Feature elimination
  - Simply identify and remove variables (columns) that are not important
  - The disadvantage is that we would gain no insight from those dropped variables and loose any information they contain
- Feature extraction
  - Create a few new variables from the old variables
  - **PCA** Principal Component Analysis: is the most popular feature extraction technique



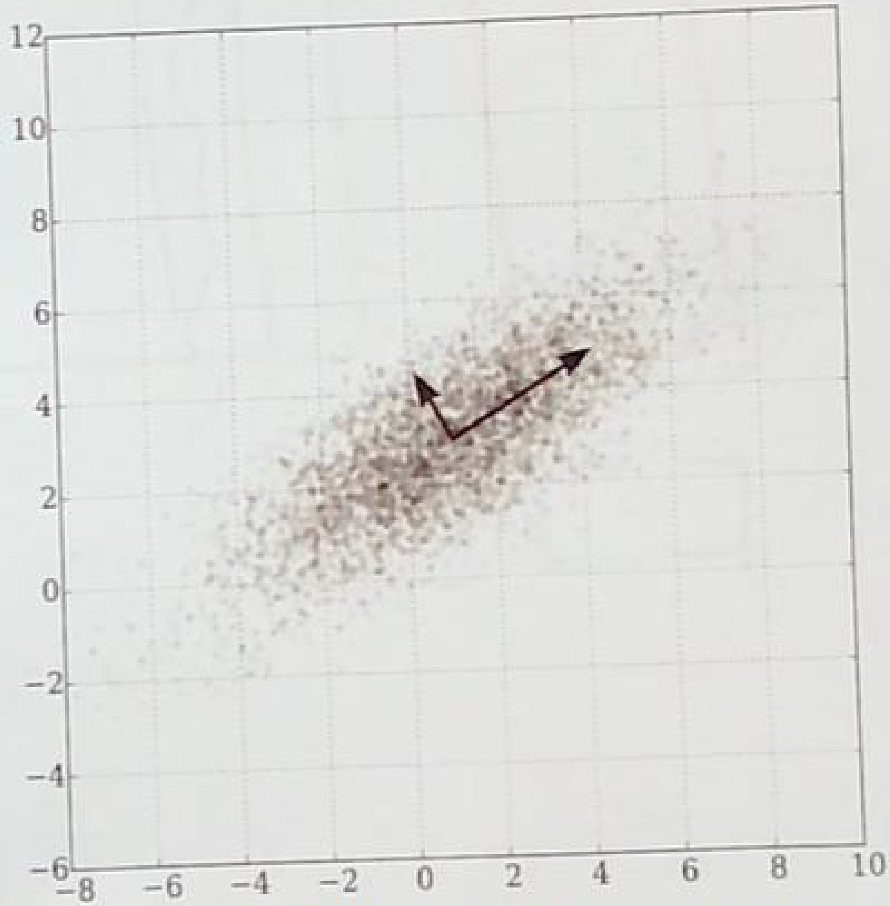



$$Z_1 = 0.2\,x_1 + 0.4\,x_2 + 0.4\,x_7$$

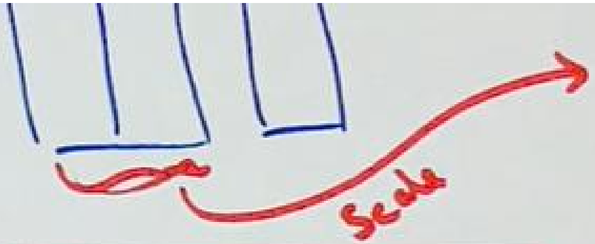
$$\cancel{Z_2 = x_3^{0.5} \times x}$$

$$Z_2 = x_3^{0.2} \times x_4^{0.5} \times \ldots \times x_6^{0.3}$$

- creates new variables using linear combinations of old variables
- is designed to create variables that are independent of one another
- also manages to tell us how important each of these new variables are
- this "importance", helps us to choose how many variables we will use



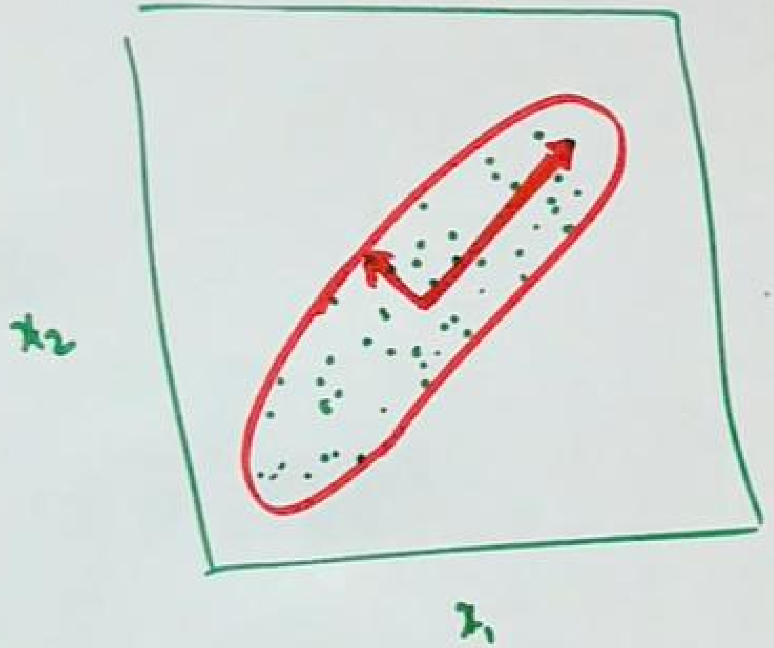

$x_2$
$x_1$




Cov Matrix C

eigen Vect.

eigen Value




$$z_1 = a_{11} x_1 + a_{12} x_2$$

$$z_2 = a_{21} x_1 + a_{22} x_2$$

Scale

eigen Vector

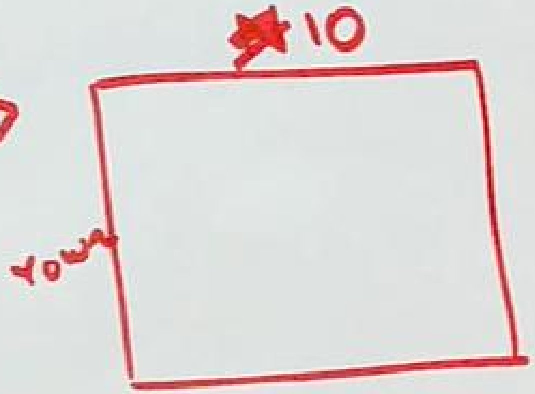

$$z_1 = a_{11}x_1 + a_{12}x_2$$

$$z_2 = a_{21}x_1 + a_{22}x_2$$

$10 \times 3$

$$V = \left( v_1 \mid v_2 \mid v_3 \right)$$

X ⇒

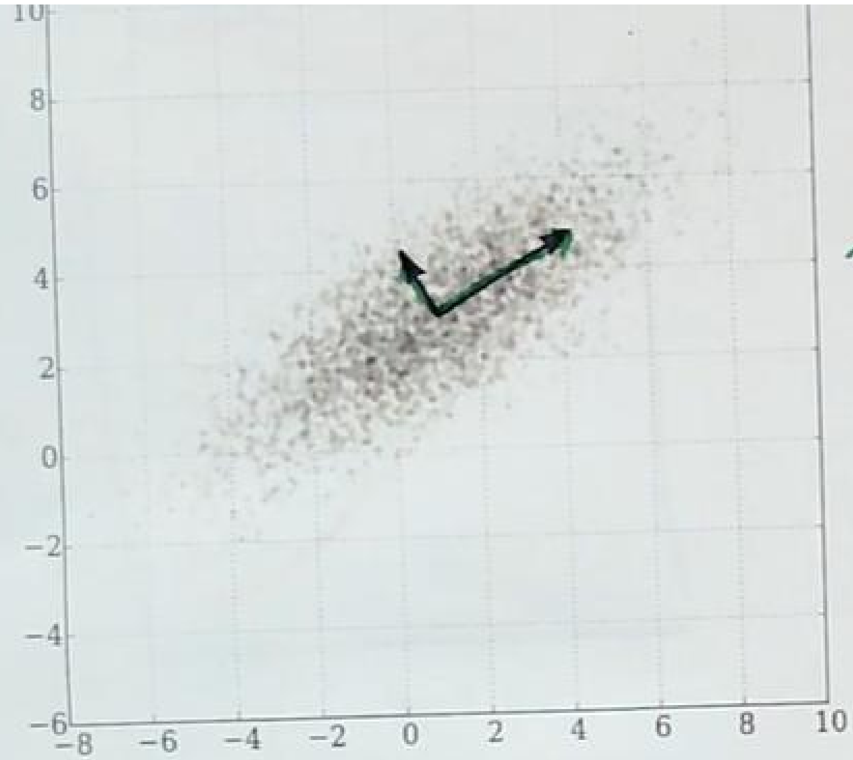



$$Z = V^T \cdot X^T$$

$$\underbrace{3\times 10 \quad 10\times \text{rows}}$$

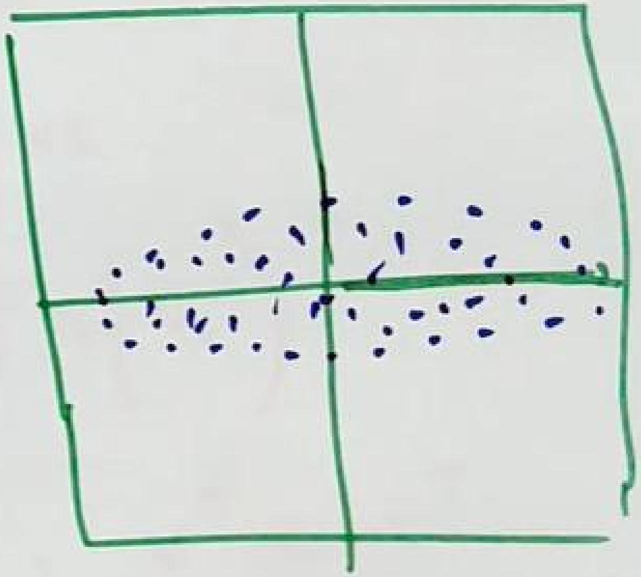
$$\rightarrow Z = \underline{3} \times \underline{\underline{\text{rows}}}$$

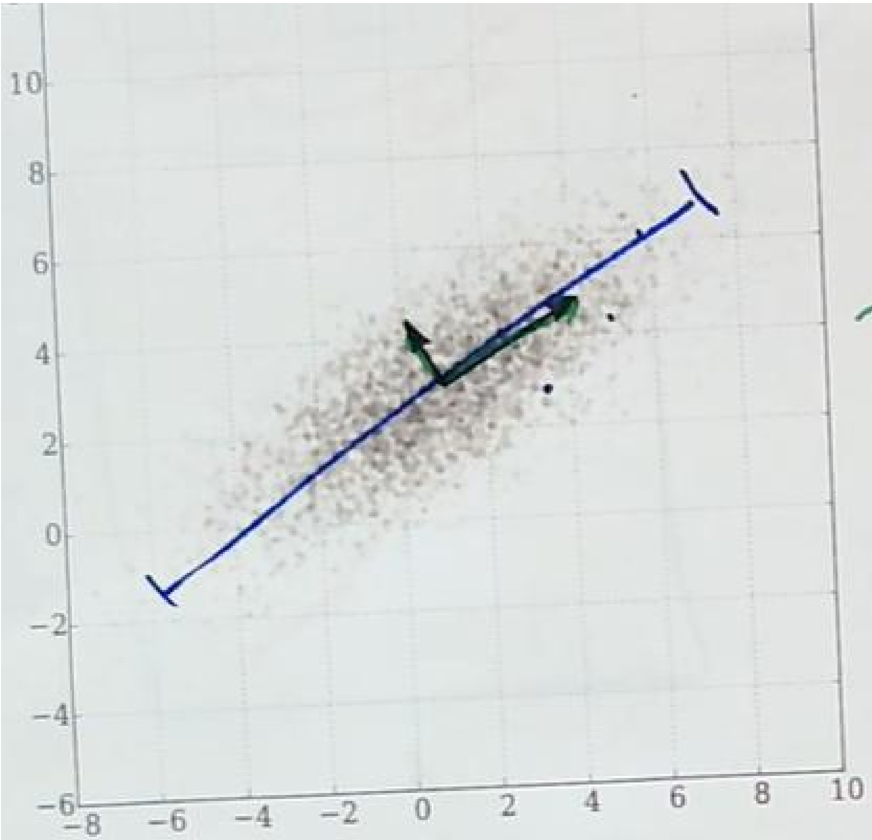
greatlearning
Learning for Life



$$z_1 = a_{11} x_1 + a_{12} x_2$$

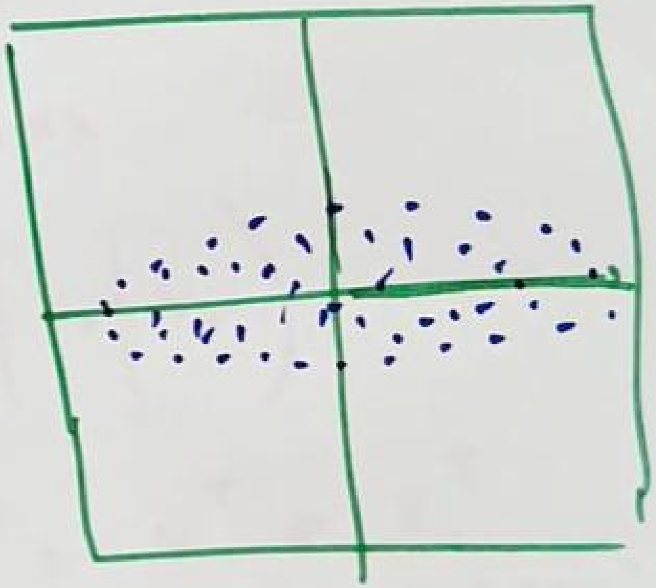
$$z_2 = \cancel{a_{12}}\, a_{21} x_1 + a_{22} x_2$$






$$\left.\begin{aligned} z_1 &= a_{11}x_1 + a_{12}x_2 \\ z_2 &= \cancel{a_{12}}\,a_{21}x_1 + a_{22}x_2 \end{aligned}\right\} 2$$

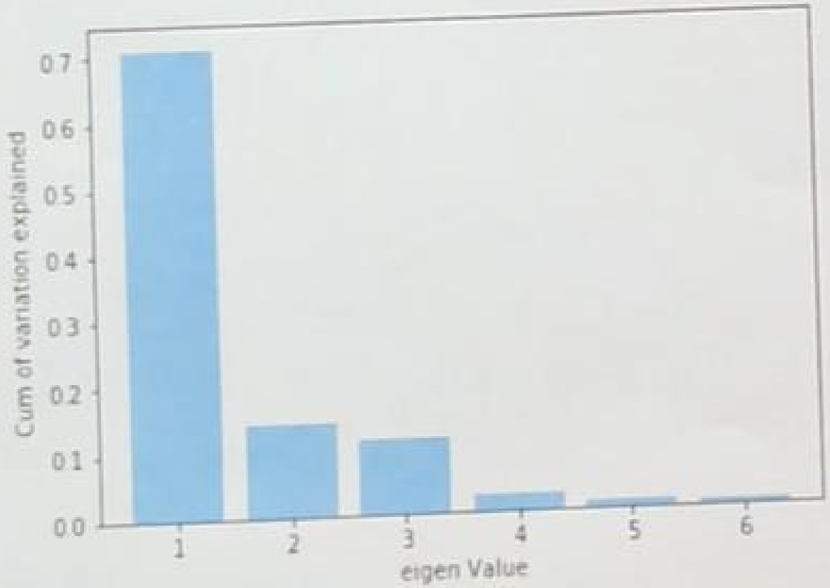
2D

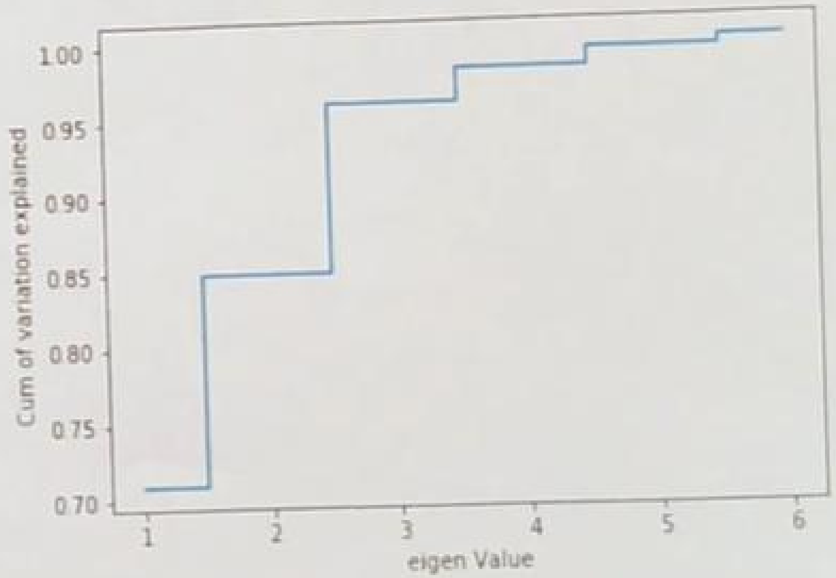


$z_1$

greatlearning
Learning for Life

- Break the covariance matrix into magnitude and direction Eigen Vectors and the Eigen Va covariance matrix can be thought of as the natural axis and magnitudes along those axis
  - Eigen Values of the covariance matrix are the principal components
  - They are all orthogonal to each other - independent
  - The eigen Values also can be used to calculate the percentage of variation explained by each direction
- Sort in the Eigen values in defending order and calculate the cumulative percentage of variation explained
- Pick the number of principal components you will use
- Transform to new variables



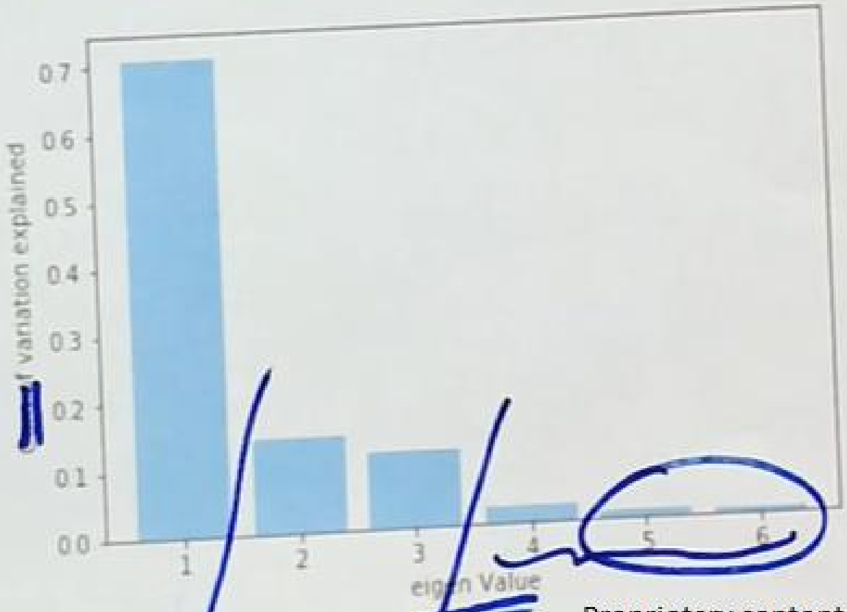
$\lambda_1 \, . \, . \, . \, . \, . \, . \, . \, \lambda_{10}$

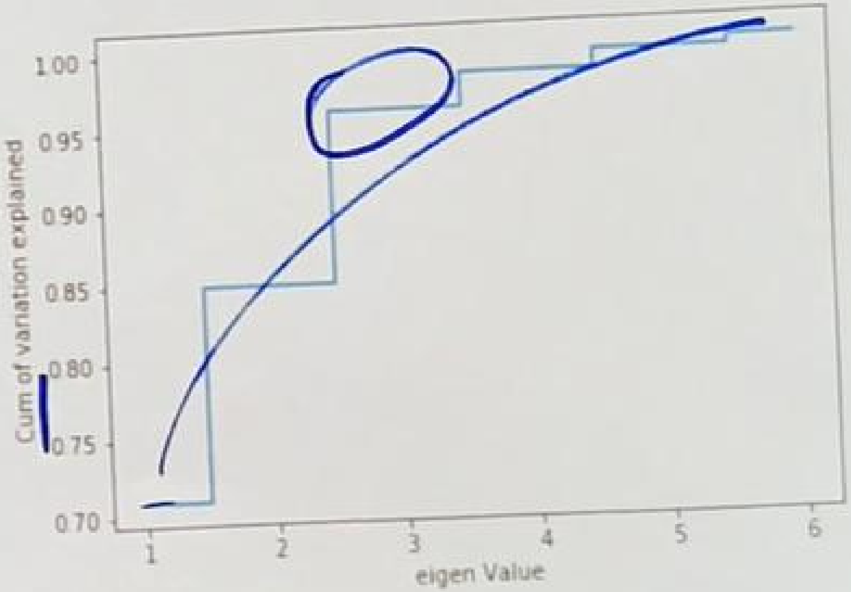
$$\left.\begin{array}{l}\text{\% of var. expl} \\ \text{by the } i^{th} \\ \text{eigen vector}\end{array}\right\} = \frac{\lambda_i}{\sum_{j=1}^{10} \lambda_j}$$

- Break the covariance matrix into magnitude and direction: the eigen vectors of the covariance matrix can be thought of as the natural axis and magnitudes along those axis
  - Eigen Values of the covariance matrix are the principal components
  - They are all orthogonal to each other - independent
  - The eigen Values also can be used to calculate the percentage of variation explained by each direction
- Sort in the Eigen values in defending order and calculate the cumulative percentage of variation explained
- Pick the number of principal components you will use
- Transform to new variables

$\lambda_i \ldots \ldots$

% of var expl. by $i \} = \dfrac{\lambda_i}{\sum_j \lambda_j}$



$y = a + b_1 x_1 + b_2 x_2$

$\cancel{+ b_3 x_3}$

- Feature elimination
  - Simply identify and remove variables (columns) that are not important
  - The disadvantage is that we would gain no insight from those dropped variables and loose any information they contain
- Feature extraction
  - Create a few new variables from the old variables
  - **PCA** Principal Component Analysis: is the most popular feature extraction technique

$$y = a + b_1 x_1 + b_2 x_2 + \cancel{b_3 x_3}$$

- Feature elimination
  - Simply identify and remove variables (columns) that are not important
  - The disadvantage is that we would gain no insight from those dropped variables and loose any information they contain
- Feature extraction
  - Create a few new variables from the old variables
  - **PCA** Principal Component Analysis: is the most popular feature extraction technique

$z_1$ $z_2$ $z_3$

$$y = a + b_1 z_1 + b_2 z_2$$