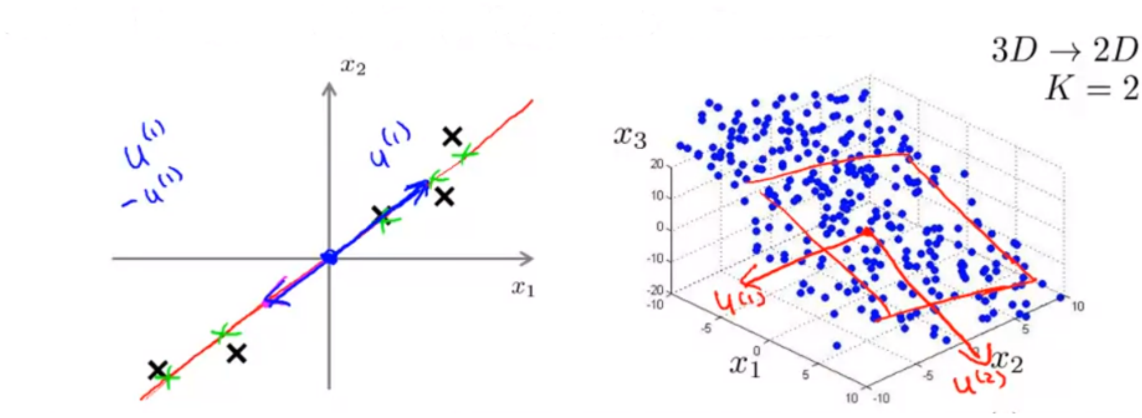
Principal Component Analysis

Insights about PCA:

- PCA is used for the problem of Dimensionality Reduction
- Pre-requisites:
    - Mean Normalization
    - Feature Scaling
- Consolidates higher dimensional features
- PCA is not Linear Regression

Problem Formulation

For a problem where we want to reduce features from 2D to 1D, find a direction (a vector $\mu^{(1)} \in R^n$) onto which to project the data so as to minimize the projection error.



General Case: Reduce from n-dimensions to k-dimensions, find k vectors $\mu^{(1)}, \mu^{(2)}, \ldots, \mu^{(k)}$ onto which to project the data, so as to minimize the projection error.

Data Preprocessing

Training Set: $x^{(1)}, x^{(2)}, \ldots, x^{(m)}$

Preprocessing (Feature Scaling/Mean Normalization):

$$\mu_j = \frac{1}{m} \sum_{i=1}^{m} x_j^{(i)}$$

- Replace each $x_j^{(i)}$ with $x_j - \mu_j$
- If different features on different scales (For example, $x_1 =$ Size of house, $x_2 =$ Number of bedrooms), scale features to have comparable range of values

Algorithm

For reducing data from n-dimensions to k-dimensions,

- Compute **Covariance Matrix**

$$\sum = \frac{1}{m}\sum_{i=1}^{n}(x^{(i)})(x^{(i)})^T$$

- Compute **Eigen Vectors** of matrix $\sum$

  ```
  [U, S, V] = svd(Sigma);
  U_reduce = U(:, 1, :k);
  z = U_reduce'*x;
  ```

  SVD = Single Value Decomposition

We'll get the $n \times n$ U matrix as:

$$U = \begin{bmatrix} \vdots & \vdots & \dots & \vdots \\ u^{(1)} & u^{(2)} & \dots & u^{(m)} \\ \vdots & \vdots & \dots & \vdots \end{bmatrix} \in R^{n \times n}$$