Foundations of Data Science & Machine Learning

Summary — Week 11
Devansh Singh Rathore
111701011

B.Tech. in Computer Science & Engineering
Indian Institute of Technology Palakkad

May 11, 2021

Abstract

In this lecture series we will discuss about Singular Values, Singular Vectors and Singular Value Decomposition.

1 Singular Values and Singular Vectors

Definition: Consider n points $x_1, x_2, ..., x_n$ in $\mathbb{R}^d$ $(d >> 1)$. Can we find a lower dimensional (k) subspace S_k of $\mathbb{R}^d$ s.t. $x_1, x_2, ..., x_n$ can be "approximated well" by points $y_1, y_2, ..., y_n$ in S_k? This is called **Dimensionality Reduction**.

→ "Approximated well" can have various interpretations:

SVD - interpretation: Choose the k-dimensional subspace which minimises the sum of squared Euclidean distances.

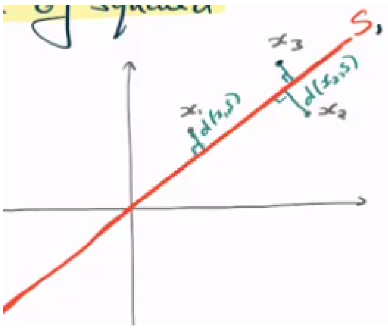

Fig 1.0 S_1 of $\mathbb{R}^2$

That is,

$$S_k = argmin_{S \in s} \sum_{i=1}^{n} dist^2(x_i, S)$$

where the minimisation is over all k-dimensional subspaces of $\mathbb{R}^d$

Notes:

1. It is different from the least square regression line.

   * vertical distance
   * not necessarily through origin.

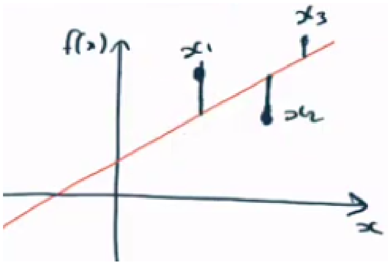

Fig 1.1 Regression line

2. Once you find S_k, $\forall i \in [n]$ let $y_i = Proj_{S_k}(x_i)$ be orthognal projection of x_i on S_k. Then the matrix

$$B_k = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}_{n \times d}$$

is the best k-rank approximation for the data matrix

$$A_k = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}_{n \times d}$$

Note: rank$(B_k) \leq k$
That is

$$B_k = argmin_{B \in b} ||A - B||_F$$

where the minimisation is over all $n \times d$ real matrices with rank $\leq k$.
$(||A||_F^2 = \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij}^2)$

Proof Idea:
$||A - B_k||_F^2 = \sum_{i=1}^{n} \sum_{j=1}^{d} (x_{ij} - y_{ij})^2$
$= \sum_{i=1}^{n} ||x_i - y_i||^2$
$= \sum_{i=1}^{n} dist^2(x, S_k)$

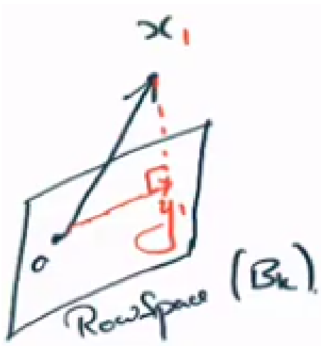

Fig 1.2 proof idea

But how do we find S?

- Equivalently we find an orthogonal basis $\{v_1, v_2, ..., v_k\}$ for S_k.
- In fact we do more. We find an orthogonal basis $\{v_1, v_2, ..., v_k\}$ of $\mathbb{R}^d$