Universitat de Barcelona

Numerical Linear Algebra

master in fundamental principles of data science

Project 2

SVD Applications

DATA SCIENCE @ UNIVERSITAT DE BARCELONA

Author

Vladislav Nikolov Vasilev

Faculty of Mathematics and Computer Science

Academic year 2021-2022

Contents

1 Introduction

The goal of this project is to discuss three common applications of the Singular Value Decomposition (SVD). First, let's briefly review what the SVD is.

Given a rectangular matrix $A \in \mathbb{R}^{m \times n}$ with $m \geq n$, we can express it as

$$A = U \Sigma V^T$$

where $U \in \mathbb{R}^{m \times m}$ and $V \in \mathbb{R}^{n \times n}$ are two orthogonal basis and $\Sigma \in \mathbb{R}^{m \times n}$ is a matrix that can be divided in the diagonal block $\Sigma\left[1:n, 1:n\right]$ with the singular values the singular values σ_i in the diagonal and the zero block $\Sigma\left[(m-n):m, 1:n\right]$. The singular values are ordered such that $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n \geq 0$. Since U and V are orthogonal, we have that $U^{-1} = U^T$ and $V^{-1} = V^T$.

There are some cases in which we can also compute a reduced version of the SVD, which is faster and reduces the amount of memory needed to store the matrices. This can be particularly useful in scenarios where the matrix A is rank deficient.

Let $A \in \mathbb{R}^{m \times n}$ be a rectangular matrix with $rank(A) = r$, where $r < n$. For this case, the reduced SVD can be computed as

$$A = U_r \Sigma_r V_r^T$$

where $U_r \in \mathbb{R}^{m \times r}$ and $V_r^T \in \mathbb{R}^{r \times r}$ are the orthogonal basis and $\Sigma \in \mathbb{R}^{r \times r}$ is the diagonal matrix containing the nonzero singular values.

There are many applications of the SVD, but in this project we are going to focus on three of them: solving the Least Squares Problem, graphic compression and Principal Component Analysis.

2 Least Squares Problem

The first application that we are going to address is the Least Squares Problem (LSP).

2.1 Polynomial fitting

2.2 The rank deficient LSP

3 Graphics compression

4 Principal Component Analysis

4.1 Example problem

4.2 Genes problem

References

[1] Texto referencia
`https://url.referencia.com`