D5: Databases and Information Systems
Data Mining and Matrices, SS 2015
Homework #2: Linear algebra & SVD
Due: **19 May 2015** at 12:15 CEST

max planck institut informatik

You can discuss these problems with other students, but everybody must hand in their own answers. You can use computers etc. to perform the algebraic operations, but you must show the intermediate steps (and "computer said so" is never a valid answer). Remember to write your name and matriculation number to every answer sheet!

Problem 1 (Matrix product with orthogonal matrices). Let $\boldsymbol{A}$ and $\boldsymbol{B}$ be n-by-n orthogonal matrices. Show that

$$\left|(\boldsymbol{A}\boldsymbol{B}^T)_{ij}\right| \leq 1$$

for all $1 \leq i, j \leq n$.

Problem 2 (Frobenius norm and SVD). A function $f\colon \mathbb{R}^{n\times m} \to \mathbb{R}$ is *unitarily invariant* if

$$f(\boldsymbol{A}) = f(\boldsymbol{U}\boldsymbol{A}\boldsymbol{V}) \tag{2.1}$$

for any $\boldsymbol{A} \in \mathbb{R}^{n\times m}$ and for any orthogonal $\boldsymbol{U} \in \mathbb{R}^{n\times n}$ and $\boldsymbol{V} \in \mathbb{R}^{m\times m}$. Use the fact that the Frobenius norm is unitarily invariant to show that if $(\sigma_i)_{i=1}^n$ are the singular values of $\boldsymbol{A}$, then

$$\|\boldsymbol{A}\|_F^2 = \sum_{i=1}^n \sigma_i^2 \ . \tag{2.2}$$

Problem 3 (Eckart–Young Theorem). In the lectures it was claimed that the truncated SVD $\boldsymbol{A}_k = \boldsymbol{U}_k\boldsymbol{\Sigma}_k\boldsymbol{V}_k^T$ provides the optimal rank-k approximation of $\boldsymbol{A}$. Prove this claim, that is, show that

$$\|\boldsymbol{A} - \boldsymbol{A}_k\|_F = \min_{\substack{\boldsymbol{B} \\ \operatorname{rank}(\boldsymbol{B})=k}} \|\boldsymbol{A} - \boldsymbol{B}\|_F \ .$$

Problem 4 (Unitarily invariant Frobenius). In Problem 2 we claimed that the Frobenius norm is unitarily invariant. Prove this claim.

Problem 5 (Singular values and rank). Let $\boldsymbol{H}_n = (h_{ij})$ be an n-by-n matrix such that $h_{ij} = 1/(i+j-1)$ (where $i, j \in \{1, 2, \dots, n\}$). Use your favorite linear algebra software (e.g. R or Matlab) to compute the SVD of the 12-by-12 matrix $\boldsymbol{H}_{12}$. Does the matrix have full rank? Why?

Problem 6 (Singular values and condition numbers). Consider again the family of matrices $\boldsymbol{H}_n$ of the previous question. Compute (and report) the condition numbers for $\boldsymbol{H}_4$, $\boldsymbol{H}_5$, and $\boldsymbol{H}_6$. Do you think (some of) these condition numbers are high? Why?