D5: Databases and Information Systems
Data Mining and Matrices, SS 2017
Homework #1: Prerequisites
Tutorial: **10 May 2017** at 10:15

$$\begin{pmatrix} \mathbf{1} & \mathbf{1} & \mathbf{0} \\ \mathbf{1} & \mathbf{1} & \mathbf{1} \\ \mathbf{0} & \mathbf{1} & \mathbf{1} \end{pmatrix}$$

You can discuss these problems with other students, but everybody must do and present 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). You are of course free to use material from the Internet, but again, you must present the intermediate steps and you must also be able to explain why the steps are valid and why you chose them. You can mark an answer even if it is not complete or correct, as long as you have made significant progress towards solving it. Note, however, that the TA does the final decision on whether your solution is complete (or correct) enough for a mark.

You must answer to every problem in this first homework sheet and be prepared to present your answers. All questions are solvable using the required prerequisites for the course, although you might have to refresh your memory of some topics. Not all concepts mentioned in the problems are necessarily familiar to you. This is intentional; you should still be able to do the problems by using the information provided in the problem statement.

Problem 1 (Matrix multiplication). Consider matrices

$$\boldsymbol{A} = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} \quad \text{and} \quad \boldsymbol{B} = \begin{pmatrix} j & k \\ l & m \\ o & p \end{pmatrix} \ ,$$

and vector

$$\boldsymbol{v} \begin{pmatrix} x \\ y \\ z \end{pmatrix} \ .$$

Now we can write $\boldsymbol{Av}$ as

$$\boldsymbol{Av} = \begin{pmatrix} ax + by + cz \\ dx + ey + fz \\ gx + hy + iz \end{pmatrix} \ .$$

a) Write open $\boldsymbol{v}^T\boldsymbol{A}$.

b) Write open $\boldsymbol{AB}$.

c) Write open $(\boldsymbol{B}^T\boldsymbol{A}^T)^T$.

Problem 2 (Linear programs). Consider the following linear program. It does not have a unique answer. There are many ways to prove that; use an approach of your choice to show that the program doesn't have a unique solution. Explain also why your method is appropriate

$$\begin{aligned}
1x_1 + 2x_2 + 0x_3 + 5x_4 + 0x_5 + 0x_6 + 7x_7 &= 97\tfrac{1}{3} \\
1x_1 + 3x_2 + 1x_3 + 0x_4 + 2x_5 + 0x_6 + 0x_7 &= 42\tfrac{2}{3} \\
0x_1 + 3x_2 + 1x_3 + 0x_4 + 0x_5 + 2x_6 + 2x_7 &= 29 \\
1x_1 + 2x_2 + 7x_3 + 5x_4 + 7x_5 + 0x_6 + 0x_7 &= 190\tfrac{2}{3} \\
1x_1 + 5x_2 + 3x_3 + 1x_4 + 2x_5 + 2x_6 + 7x_7 &= 106\tfrac{1}{3} \\
0x_1 + 3x_2 + 0x_3 + 5x_4 + 0x_5 + 0x_6 + 0x_7 &= 70\tfrac{1}{3} \\
4x_1 + 18x_2 + 12x_3 + 16x_4 + 11x_5 + 4x_6 + 16x_7 &= 4
\end{aligned}$$

max planck institut
informatik

Problem 3 (Orthogonality). The *Euclidean inner product* of two vectors $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^n$ is defined as

$$\langle \boldsymbol{a}, \boldsymbol{b} \rangle = \sum_{i=1}^{n} a_i b_i \; .$$

Two non-zero vectors in Euclidean space are *orthogonal* if the angle between them is $\frac{\pi}{2} + k\pi$ radians for $k \in \{0, 1, 2, \ldots\}$. Show that two non-zero Euclidean vectors are orthogonal if and only if their inner product is 0. You may use the fact that if θ is the angle between $\boldsymbol{a}$ and $\boldsymbol{b}$, then

$$\cos(\theta) = \frac{\langle \boldsymbol{a}, \boldsymbol{b} \rangle}{\|\boldsymbol{a}\| \, \|\boldsymbol{b}\|} \; ,$$

where $\|\boldsymbol{a}\| = \sqrt{\sum_{i=1}^{n} a_i^2}$ is the *Euclidean norm* of $\boldsymbol{a}$.

Problem 4 (Norms). A function $f : \mathbb{R}^n \to \mathbb{R}$ is a *norm* if it satisfies the following properties for all $a \in \mathbb{R}$ and all $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$

$$f(a\boldsymbol{x}) = |a| \, f(\boldsymbol{x}) \qquad (4.1)$$

$$f(\boldsymbol{x} + \boldsymbol{y}) \leq f(\boldsymbol{x}) + f(\boldsymbol{y}) \qquad (4.2)$$

$$\text{If } f(\boldsymbol{x}) = 0 \text{ then } \boldsymbol{x} \text{ is all-zeros vector.} \qquad (4.3)$$

Norms generalize the concept of distance, and the most common norm is the *Euclidean norm* defined in $\mathbb{R}^n$ as

$$\|\boldsymbol{x}\|_2 = \sqrt{\sum_{i=1}^{n} x_i^2} \; . \qquad (4.4)$$

Prove that the Euclidean norm is indeed a norm, that is, prove that it satisfies properties (4.1)–(4.3). You may use the *Cauchy–Schwarz* inequality: for any $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$,

$$\left(\sum_{i=1}^{n} x_i y_i \right)^2 \leq \left(\sum_{i=1}^{n} x_i^2 \right) \left(\sum_{i=1}^{n} y_i^2 \right) \; . \qquad (4.5)$$

Problem 5 (The Frobenius norm). The *Frobenius norm* of an n-by-m real matrix $\boldsymbol{B}$ is defined as

$$\|\boldsymbol{B}\|_F = \left(\sum_{i=1}^{n} \sum_{j=1}^{m} b_{ij}^2 \right)^{1/2} \; . \qquad (5.1)$$

We can compute the Frobenius norm using the *trace* of a matrix: The *trace* function tr computes the sum of the diagonal elements of a square matrix, that is, if $\boldsymbol{A} \in \mathbb{R}^{n \times n}$, then

$$\operatorname{tr}(\boldsymbol{A}) = \sum_{i=1}^{n} a_{ii} \; . \qquad (5.2)$$

Proof that

$$\|\boldsymbol{B}\|_F = \left(\operatorname{tr}(\boldsymbol{B}^T \boldsymbol{B}) \right)^{1/2} \; , \qquad (5.3)$$

where $\boldsymbol{B}^T$ is the transpose of $\boldsymbol{B}$.

Problem 6 (The QR decomposition). Let $\boldsymbol{A} \in \mathbb{R}^{n \times m}$ with $n \geq m$. The *(thin) QR decomposition of* $\boldsymbol{A}$ expresses $\boldsymbol{A}$ as a product of two matrices,

$$\boldsymbol{A} = \boldsymbol{Q}\boldsymbol{R} \,, \tag{6.1}$$

where $\boldsymbol{Q} \in \mathbb{R}^{n \times m}$ is a *column orthogonal* matrix (that is, $\boldsymbol{Q}^T\boldsymbol{Q} = \boldsymbol{I}$, where $\boldsymbol{I}$ is the m-by-m *identity matrix*) and $\boldsymbol{R} \in \mathbb{R}^{m \times m}$ is an upper-triangular matrix.

Let us assume the columns of $\boldsymbol{A}$ are linearly independent. Then the *pseudo-inverse of* $\boldsymbol{A}$, $\boldsymbol{A}^+$, is defined as

$$\boldsymbol{A}^+ = (\boldsymbol{A}^T\boldsymbol{A})^{-1}\boldsymbol{A}^T \,. \tag{6.2}$$

(Then inverse of $\boldsymbol{A}^T\boldsymbol{A}$ exists because $\boldsymbol{A}$ has linearly independent columns.) When the columns of $\boldsymbol{A}$ are linearly independent, matrix $\boldsymbol{R}$ on its QR decomposition is non-singular (i.e. it has inverse $\boldsymbol{R}^{-1}$ such that $\boldsymbol{R}\boldsymbol{R}^{-1} = \boldsymbol{R}^{-1}\boldsymbol{R} = \boldsymbol{I}$).

Write Equation (6.2) using the QR decomposition of $\boldsymbol{A}$ instead of $\boldsymbol{A}$ and simplify the result as much as you can. Remember the following two basic equations:

$$(\boldsymbol{X}\boldsymbol{Y})^T = \boldsymbol{Y}^T\boldsymbol{X}^T \tag{6.3}$$

$$(\boldsymbol{X}\boldsymbol{Y})^{-1} = \boldsymbol{Y}^{-1}\boldsymbol{X}^{-1} \qquad \text{(assuming the inverses exist).} \tag{6.4}$$