Machine Learning
Assignment # 1
Universität Bern

Due date: 9/10/2019

Late submissions will incur a penalty. Submit your answers in ILIAS (as a pdf or as a picture of your written notes if the handwriting is very clear). Submission instructions will be provided via email. For any clarification about the problem set ask the teaching assistant. You are not allowed to work with others.

Linear algebra review **[Total 100 points]**

Solve each of the following problems and show all the steps of your working.

1. $S = \{v_1, \dots, v_n\}$ be an orthogonal set of non-zero vectors in R^n. **[10 points]**
   Prove that the vectors in S are linearly independent.
   Solution.
   Let $c_1, \dots c_n \in \mathbb{R}$ such that $c_1\ v_1 + \cdots c_n v_n = 0$. Take the dot product with v_1 and we get $0 = v_1 \cdot \sum_{j=1}^{n} (c_j v_j) = c_1 \|v_1\|^2$. Therefore, $c_1 = 0$ **[5 points]**. Similarly we find that $c_2, c_3, \dots c_n = 0$. Therefore S is linearly independent **[5 points]**.

2. Given a square matrix $A \in \mathbb{R}^{n\times n}$ and a vector $x \in \mathbb{R}^n$ show that $x^T A x = x^T(\frac{1}{2}A + \frac{1}{2}A^T)x$ **[15 points]**
   Solution.

$$x^T(\frac{1}{2}A + \frac{1}{2}A^T)x = (\frac{1}{2}x^T A + \frac{1}{2}x^T A^T)x = \frac{1}{2}x^T A x + \frac{1}{2}x^T A^T x = \frac{1}{2}x^T A x + \frac{1}{2}(x^T A x)^T$$

[5 points]

   since $x^T A x$ is a scalar $x^T A x = (x^T A x)^T$ and therefore, **[5 points]**

$$\frac{1}{2}x^T A x + \frac{1}{2}(x^T A x)^T = \frac{1}{2}x^T A x + \frac{1}{2}x^T A x = x^T A x$$

[5 points]

3. Show that if $(A + B)^{-1} = A^{-1} + B^{-1}$ then $AB^{-1}A = BA^{-1}B$ **[15 points]**
   Solution.

$$(A + B)(A^{-1} + B^{-1}) = I$$

[5 points]

$$AA^{-1} + AB^{-1} + BA^{-1} + BB^{-1} = I$$
$$AB^{-1} + BA^{-1} = -I$$

[5 points]

$$AB^{-1}A + B = -A \to AB^{-1}A = -A - B$$
$$A + BA^{-1}B = -B \to BA^{-1}B = -A - B$$
$$AB^{-1}A = BA^{-1}B$$

[5 points]

4. Use the definition of trace to show that $tr(A+B) = trA + trB$, where $A, B \in \mathbb{R}^{n\times n}$. **[15 points]**
Solution.

$$\text{tr}(A) = \sum_{i=1}^{n} A_{ii}$$
$$\text{tr}(B) = \sum_{i=1}^{n} B_{ii}$$

[5 points]

$$\text{tr}(A) + \text{tr}(B) = \sum_{i=1}^{n} A_{ii} + \sum_{i=1}^{n} B_{ii} = \sum_{i=1}^{n} (A_{ii} + B_{ii})$$

[5 points]

$$\text{tr}(A+B) = \sum_{i=1}^{n} (A_{ii} + B_{ii})$$

[5 points]

5. Show that if (λ_i, x_i) are the i-th eigenvalue and i-th eigenvector of a non-singular and symmetric matrix $A \in \mathbb{R}^{n\times n}$, then $(\frac{1}{\lambda_i}, x_i)$ are the i-th eigenvalue and i-th eigenvector of A^{-1}. **[15 points]**
Hint: use the eigendecomposition of A.
Solution.

$$Ax = \lambda_i x_i \;\; \textbf{[5 points]} \to x_i = \lambda_i A^{-1} x_i \;\; \textbf{[5 points]} \to A^{-1} x_i = \frac{1}{\lambda_i} x \;\; \textbf{[5 points]}$$

6. Show that rank$(A) \leq \min\{m, n\}$, where $A \in \mathbb{R}^{m\times n}$. **[10 points]**

Solution. We know that column rank and row rank of any matrix is equal. also we know that the column rank is at most equal to the number of columns and the row rank is at most equal to the number of rows **[5 points]**. These two consideration implies that rank$(A) \leq \min\{m, n\}$ **[5 points]**.

7. In each of the following cases, state whether the real matrix A is guaranteed to be singular or not. Justify your answer in each case.

   (a) $A \in \mathbb{R}^{(n+1)\times n}$ is a full rank matrix. **[4 points]**
   Solution. No. A non-singular matrices should be square.

   (b) $|A| = 0$. **[4 points]**
   Solution. No. An square matrix A is non-singular if and only $|A| \neq 0$

   (c) A is an orthogonal matrix. **[4 points]**
   Solution. Yes, For an orthogonal matrix Q we have $Q^\top Q = QQ^\top = I$ so Q is non singular and $Q^{-1} = Q^\top$

   (d) A has no eigenvalue equal to zero. **[4 points]**
   Solution. Yes, we know that $|A| = \prod \lambda_i$. So if A has no zero eigenvalue, then $|A| \neq 0$ so A is non-singular.

   (e) A is a symmetric matrix with non-negative eigenvalues. **[4 points]**
   Solution. No. Eigenvalues could be zero and then it would not be invertible (see (d)).