Machine Learning
Assignment # 1
Universität Bern

Due date: 10/10/2017

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. Suppose that the matrices A and C satisfy the equations $Ax = b$ and $Cx = b$ with the same set of solutions x for every b. Can we conclude that $A = C$? Justify your answer in detail. **[10 points]**
   Solution.
   Yes, $A = C$. Because A and C satisfy $Ax = b$ and $Cx = b$ with the same x and b they must also have the same shape and nullspace (when $b = 0$). If b is the first column of A, then $x = (1, 0, ..., 0)$ solves $Ax = b$ but it also must solve $Cx = b$. Then the first column of C must be equal to b, which is the first column of A. The other columns follow suit.

2. Suppose that the $j - th$ column of a matrix B is a combination of the other columns of B. Show the relationship between the $j - th$ column of the matrix AB and the other columns of AB. **[10 points]**
   Solution.
   By matrix multiplication, each column of AB is A times the corresponding column of B. So if column j of B is a combination of earlier columns, then column j of AB is the same combination of earlier columns of AB.

3. Suppose that matrices A and B are invertible. Show that the inverse of the product AB is **[10 points]**

$$(AB)^{-1} = B^{-1}A^{-1}$$

   Solution. To show that let us multiply AB by $B^{-1}A^{-1}$

$$ABB^{-1}A^{-1} = A(BB^{-1})A^{-1} = AIA^{-1} = AA^{-1} = I$$

   Similarly we can multiply from the left side

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

4. Use the definition of trace to show that $trAB = trBA$, where $A \in \mathbb{R}^{m\times n}, B \in \mathbb{R}^{n\times m}$. **[10 points]**
   Solution.

$$\begin{aligned}
\mathrm{tr}AB &= \sum_{i=1}^{m}(AB)_{ii} = \sum_{i=1}^{m}\Big(\sum_{j=1}^{n} A_{ij}B_{ji}\Big) \\
&= \sum_{i=1}^{m}\sum_{j=1}^{n} A_{ij}B_{ji} = \sum_{j=1}^{n}\sum_{i=1}^{m} B_{ji}A_{ij} \\
&= \sum_{j=1}^{n}\Big(\sum_{i=1}^{m} B_{ji}A_{ij}\Big) = \sum_{j=1}^{n}(BA)_{jj} = \mathrm{tr}BA
\end{aligned}$$

5. Consider the matrix $G = A^\top A$, where $A \in \mathbb{R}^{m\times n}$ and $E = B^\top B$, where $B \in \mathbb{R}^{m\times n}$. **[20 points]**
Show that G, E, and $G+E$ are all positive semi-definite.
Solution.

$$x^\top G x = x^\top A^\top A x = (Ax)^\top (Ax) = y^\top y \geq 0$$

$$x^\top E x = x^\top B^\top B x = (Bx)^\top (Bx) = z^\top z \geq 0$$

$$x^\top (G+E) x = x^\top G x + x^\top E x = y^\top y + z^\top z \geq 0$$

6. Suppose C is positive definite and A has independent columns. Show that $S = A^\top C A$ is positive definite. **[20 points]**
Solution.

$$x^\top S x = x^\top A^\top C A x = (Ax)^\top C A x$$

Since C is assumed positive definite, $x^\top S x = 0$ only when $Ax = 0$. Since A has independent columns, $Ax = 0$ only when $x = 0$. Thus $S = A^\top C A$ is positive definite.

7. Given two sets of vectors $\{x_1, ...x_n\} \subset \mathbb{R}^n$ and $\{y_1, ..., y_n\} \subset \mathbb{R}^n$, show that rank $\left[\sum_{i=1}^m x_i y_i^\top\right] \leq m$.
Hint: First show that the square matrix $x_i y_i^\top$ has rank 1. **[20 points]**

Solution.

$$A^{(i)} = x_i y_i^\top = \begin{bmatrix} -a_1^{(i)}- \\ \vdots \\ -a_n^{(i)}- \end{bmatrix} \tag{1}$$

and $a_1^{(i)} = (x_i)_1 y_i, ..., a_m^{(i)} = (x_i)_n y_i$. So every column of $A^{(i)}$ is a scalar product of y_i. This implies that $A^{(i)}$ is of rank 1. **10 points**

We know that $rank(A+B) \leq rank(A) + rank(B)$. Considering that $rank(x_i y_i^\top) = 1, \forall i$ we have rank $\left[\sum_{i=1}^m x_i y_i^\top\right] \leq m$. **10 points**