Machine Learning Assignment # 2
Universität Bern

Due date: 10/10/2018

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. You are not allowed to work with others.

Calculus review **[Total 100 points]**

Recall that the Jacobian of a function $f : \mathbb{R}^n \to \mathbb{R}^m$ is an $m \times n$ matrix of partial derivatives

$$Df(x) = \begin{bmatrix} \frac{\partial f_1(x)}{\partial x_1} & \frac{\partial f_1(x)}{\partial x_2} & \dots & \frac{\partial f_1(x)}{\partial x_n} \\ \frac{\partial f_2(x)}{\partial x_1} & \frac{\partial f_2(x)}{\partial x_2} & \dots & \frac{\partial f_2(x)}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_m(x)}{\partial x_1} & \frac{\partial f_m(x)}{\partial x_2} & \dots & \frac{\partial f_m(x)}{\partial x_n} \end{bmatrix}$$

where $x = [x_1\ x_2\ \dots\ x_n]^\top$, $f(x) = [f_1(x)\ f_2(x)\ \dots\ f_m(x)]^\top$ and $\frac{\partial f_i(x)}{\partial x_j}$ is the partial derivative of the i-th output with respect to the j-th input.

Answer the following questions (show all the steps of your working)

1. Consider the function $f : \mathbb{R}^m \to \mathbb{R}^n$ and $f(x) = x^\top A$ where $A \in \mathbb{R}^{m \times n}$. Show that $Df(x) = A^\top$. [10 points]
2. Consider the function $g : \mathbb{R}^n \to \mathbb{R}$ and $g(x) = (x^\top x)^{m+1}$, with m an integer larger than 0. Calculate $\nabla g(x)$. [15 points]
3. Given a square matrix $A \in \mathbb{R}^{m \times m}$, when is the equality $tr(A^2) = tr(A^T A)$ true? [10 points]
   Hint: $A^2 = AA$.
4. Assume $A \in \mathbb{R}^{m \times n}$, $X \in \mathbb{R}^{m \times n}$, and $B \in \mathbb{R}^{m \times m}$. Show that
   - $\nabla_X tr(AX^T) = A$ [10 points]
   - $\nabla_X tr(X^T X) = 2X$. [10 points]
5. Suppose $A \in R^{m \times n}$ and is full rank, and $b \in R^m$ is a vector such that $b \notin \mathcal{R}(A)$. [20 points]
   In this case we will not be able to find a vector $x \in R^n$, such that $Ax = b$.
   Find a vector x such that Ax is as close as possible to b, as measured by the square of the Euclidean norm:
   $$\|Ax - b\|_2^2$$
   Hint: $\|x\|_2^2 = x^\top x$. *Review the least squares method.*
6. Solve the following equality constrained optimization problem: [25 points]
   $$\max_{x \in R^n} x^\top Ax \qquad \text{subject to } \|x\|_2^2 = 1$$
   for a symmetric matrix $A \in S^n$.
   Hint: Use the Lagrangian and the Eigen decomposition.