1. Calculate $f_x$ and $f_y$ for the homogeneously defined functions $f(x, y) = e^{xy + siny}$ and $f(x, y) = x^2y$. Finally, for $f(x, y) = x^3 + cos(xy)$, find $f_{xx}$, $f_{xy}$, $f_{yx}$ and $f_{yy}$. The application of partial derivatives is important later for Back-Propagation and Gradient Descent Algorithms used in Energy Minimization for training more complicated Neural Network.

2. Check whether the inhomogeneously defined function $f$ defined over $\Omega \subset R^2$

$$f(x, y) = \begin{cases} \frac{xy}{x^2+y^2}, & \text{if } (x, y) \neq (0, 0); \\ 0, & \text{if } (x, y) = (0, 0) \end{cases}$$

has its partial derivatives over $R^2$.

3. Prove that the function $f_\alpha(x) = \frac{1}{1+\exp(-\alpha x)}$ (with arbitrary $\alpha > 0$) satisfies

$$f'_\alpha(x) = \alpha f_\alpha(x)(1 - f_\alpha(x)) \quad \text{and} \quad 0 < f_\alpha(x) < 1.$$ 

Infer the graph of $f_\alpha$. Show that for a fixed $x$ the limit $\lim_{\alpha \to -\infty} f_\alpha(x) = f_\infty(x)$, where

$$f_\infty(x) = \begin{cases} 1, & \text{if } x > 0; \\ 1/2, & \text{if } x = 0; \\ 0, & \text{if } x < 0 \end{cases}$$

which means that sigmoidal-like function $f_\alpha$ approximates asymptotically a standard perceptron discontinuous activation function. Note that the derivative $f'_\alpha$ ($g'_\beta$ defined below) is expressed in terms of the function $f_\alpha$ (or $g_\beta$) itself. This reduces computation time given $f_\alpha$ (or $g_\beta$) is known.

4. Prove that the function $g_\beta(x) = \frac{1-\exp(-\beta x)}{1+\exp(-\beta x)}$ (with $\beta > 0$) satisfies

$$g'_\beta(x) = (\beta/2)(1 - g_\beta^2(x)) \quad \text{and} \quad -1 < g_\beta(x) < 1.$$ 

Infer the graph of $g_\beta$. Show that for a fixed $x$ the limit $\lim_{\beta \to -\infty} g_\beta(x) = g_\infty(x)$, where

$$g_\infty(x) = \text{sgn}(x) = \begin{cases} 1, & \text{if } x > 0; \\ 0, & \text{if } x = 0; \\ -1, & \text{if } x < 0 \end{cases}$$

which means that tangensoidal function $g_\beta$ approximates asymptotically one of the standard perceptron discontinuous activation function.
5. Define a single McCulloch-Pitts neurons which implement the Boolean Switching logic functions AND, OR and logical NOT.

6. Construct a two-layer network containing the 3 perceptrons (2 in the hidden layer and one in the output layer and of course three signals in the input layer) which implements the exclusive OR logical function: i.e. \( \text{XOR} = (x \text{ AND NOT(y)}) \text{ OR (NOT(x) AND y)} \). Hint: rearrange the expression into conjunction of two other linearly separable Boolean functions. Each of them can be modelled by the perceptron. Finally, note that the AND is also linearly separable Boolean function.

Assoc. Prof. Ryszard Kozera