Pre-class activities
Artifical Neural Networks

Emmanuel Rachelson

1 Derivation: the chain rule and total derivatives

1. Consider a function $f : \mathbb{R}^p \to \mathbb{R}^q$. Recall the expression of this function's Jacobian matrix.
2. The total derivative $Df_{\hat{x}}(h)$ of f in $\hat{x}$ is a linear operator. Recall its definition (for the same $f : \mathbb{R}^p \to \mathbb{R}^q$ function as above).
3. Consider two functions, $g : \mathbb{R} \to \mathbb{R}^p$ and $f : \mathbb{R}^p \to \mathbb{R}$. Let $F = f \circ g$ be the composite function such that $F(x) = f(g(x))$. Write the derivative of F with respect to x as an expression of the partial derivatives of f and g.
4. Now suppose that in the example above, all the g_k functions are identity functions, that is $g(x) = [x, \dots, x]^T$. How does the total derivative of F simplify?
5. Finally, consider two functions $g : \mathbb{R} \to \mathbb{R}^p$ and $f : \mathbb{R}^p \to \mathbb{R}^q$. As previously, lets $F = f \circ g$ be the composite function. Write the total derivative of F as an expression of the partial derivatives of f and g.