UML2014

Exercise set 1

Solutions to be presented in the 18.3.2014 session

Exercise 1:

Understand and be able to explain the formula

$$\frac{\partial}{\partial A_{ij}} A_{nm} = \delta_{in}\delta_{jm}, \tag{1}$$

where A is a matrix.

Advice: Suppose a function $f : \mathbb{R}^2 \to \mathbb{R}$ was defined by a formula $f(x, y) = x^2y^3 + 2xy^4$. If you were asked to compute the partial derivatives $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$, you would (at least implicitly) use the formulas

$$\frac{\partial}{\partial x} x = 1, \qquad \frac{\partial}{\partial x} y = 0, \qquad \frac{\partial}{\partial y} x = 0, \qquad \frac{\partial}{\partial y} y = 1. \tag{2}$$

How would you generalize these identities with the Kronecker delta for a function $f : \mathbb{R}^N \to \mathbb{R}$, whose input parameter is $(x_1, x_2, \ldots, x_N)$? Finally, consider some function $f : \mathbb{R}^{N\times M} \to \mathbb{R}$, $A \mapsto f(A)$, whose input parameter is a matrix.

Exercise 2:

Suppose $t \mapsto A(t)$ and $t \mapsto B(t)$ are some differentiable mappings of form $\mathbb{R} \to \mathbb{R}^{N\times M}$ and $\mathbb{R} \to \mathbb{R}^{M\times K}$. So for each fixed $t \in \mathbb{R}$, $A(t)$ and $B(t)$ are some matrices. Let's define the derivatives of the matrices componentwisely, so that $\frac{dA(t)}{dt}$ and $\frac{dB(t)}{dt}$ are also matrices, defined by formulas

$$\Big(\frac{dA(t)}{dt}\Big)_{nm} := \frac{d}{dt}\Big(A(t)_{nm}\Big), \qquad \Big(\frac{dB(t)}{dt}\Big)_{mk} := \frac{d}{dt}\Big(B(t)_{mk}\Big). \tag{3}$$

Prove the formula

$$\frac{d}{dt}\Big(A(t)B(t)\Big) = \frac{dA(t)}{dt}B(t) + A(t)\frac{dB(t)}{dt}. \tag{4}$$

Advice: You can assume that the formula

$$\frac{d}{dt}\Big(f(t)g(t)\Big) = f'(t)g(t) + f(t)g'(t) \tag{5}$$

is already known for real valued functions $f : \mathbb{R} \to \mathbb{R}$ and $g : \mathbb{R} \to \mathbb{R}$. Notice that this is not the same thing as Equation (4), since (4) involves matrix multiplications.

Exercise 3:

Solve some useful formula for the partial derivative

$$\frac{\partial}{\partial A_{ij}} f(A), \tag{6}$$

when the function is defined by a formula

$$f(A) = \mathbf{v}^T A^T C A \mathbf{v} \tag{7}$$

where $\mathbf{v} \in \mathbb{R}^{N\times 1}$ is some constant vertical vector, $C \in \mathbb{R}^{N\times N}$ is some constant symmetric matrix.

Advice: One of the most primitive and secure ways to obtain some result is to first expand all matrix multiplications as

$$f(A) = \sum_{n,n',n'',n'''=1}^{N} v_n A_{n'n} C_{n'n''} A_{n''n'''} v_{n'''}, \tag{8}$$

and use the result from the first exercise, but this is not the most efficient way. We recommend you to contemplate on ways to avoid some unnecessary index work.

Exercise 4:

Solve simple formulas for partial derivatives

$$\frac{\partial}{\partial \mathrm{Re}(v_i)} f(\mathbf{v}) \quad \text{and} \quad \frac{\partial}{\partial \mathrm{Im}(v_i)} f(\mathbf{v}) \tag{9}$$

when a function $f : \mathbb{C}^N \to \mathbb{R}$ has been defined by a formula $f(\mathbf{v}) = \mathbf{v}^\dagger \mathbf{v}$. Here $\mathbf{v}$ is a $N \times 1$ vertical vector with complex elements, and $\mathbf{v}^\dagger$ is a $1 \times N$ conjugate transpose, whose elements are the complex conjugates of the elements of $\mathbf{v}$.

Advice: For a complex variable $z \in \mathbb{C}$ we denote its real and imaginary parts as $\mathrm{Re}(z)$ and $\mathrm{Im}(z)$, so that always $z = \mathrm{Re}(z) + i\mathrm{Im}(z)$. Computing partial derivatives with respect to $\mathrm{Re}(z)$ and $\mathrm{Im}(z)$ works the same way as computing partial derivatives with respect to x_1 and x_2 with a two dimensional vector $\mathbf{x} \in \mathbb{R}^2$. For example, it would not be difficult to solve $\frac{\partial f}{\partial x_1}$ and $\frac{\partial f}{\partial x_2}$ for $f : \mathbb{R}^2 \to \mathbb{R}$, $f(\mathbf{x}) = \|\mathbf{x}\|^2$. Then, what would $\frac{\partial f}{\partial \mathrm{Re}(z)}$ and $\frac{\partial f}{\partial \mathrm{Im}(z)}$ be for $f : \mathbb{C} \to \mathbb{R}$, $f(z) = |z|^2$?