Exercise 7

January 18, 2021

1. Compute the derivative $f'(x)$ for:

   $f(x) = log(x^4)sin(x^3)$

   Solution:

   Firstly, let's rewrite $f(x)$ as $f(x) = 4\log(x)\sin(x^3)$.

   Then, $f'(x) = \frac{4}{x}\sin(x^3) + 12x^2\log(x)\cos(x^3)$.

2. Compute the derivative $f'(x)$ of logistic sigmoid:

   $f(x) = \frac{1}{1+exp(-x)}$

   Solution:

   If we rewrite our function as $f(x) = (1 + \exp(-x))^{-1}$, then we have $f'(x) = \frac{\exp(-x)}{(1+\exp(-x))^2}$.

3. Compute the derivative $f'(x)$ of function:

   $f(x) = exp(-\frac{1}{2\sigma^2}(x - \varphi)^2)$
   Where σ and $\varphi \in R$ are constant

   Solution:

   We have $f'(x) = \frac{\mu - x}{\sigma^2}f(x)$.

4. Compute the Taylor polynomial $T_n, n = 0, ..., 5$ of $f(x) = sin(x) + cos(x)$ at $x_0 = 0$

   Solution:

   We compute the first five derivatives of our function at 0. We have $f(0) = f'(0) = 1$, $f^{(2)}(0) = f^{(3)}(0) = -1$, and $f^{(4)}(0) = f^{(5)}(0) = 1$.

   The Taylor polynomial $T_5(x) = 1 + x - \frac{1}{2}x^2 - \frac{1}{6}x^3 + \frac{1}{24}x^4 + \frac{1}{120}x^5$. The lower-order Taylor polynomials can be found by truncating this expression appropriately.

5. Compute the derivatives df/dx of the following functions. Provide the dimensions of every single partial derivative. Describe your steps in detail.

   (a) Use the chain rule. Provide the dimensions of every single partial derivative.
   $f(z) = exp(-\frac{1}{2z})$
   $z = g(y) = y^T S^-1y$
   $y = h(x) = x - \mu$

   (b) $f(x) = tr(xx^T + \sigma^2 y), x \in R^D$ Here $tr(A)$ is the trace of A, i.e., the sum of the diagonal elements A_{ii}.
   Hint: Explicitly write out the outer product

   (c) Use the chain rule. Provide the dimensions of every single partial derivative. You do not need to compute the product of the partial derivatives explicitly.
   $f = tanh(z) \in R^M$
   $z = Ax + b, x \in R^N, A \in R^M, b \in R^M$
   Here, tanh is applied to every component of z.

   Solution:

   (a) We have $\frac{df}{dz}$ has dimension 1×1, and is simply $-\frac{1}{2}\exp(-\frac{1}{2}z)$.

   Now, $\frac{dz}{dy}$ has dimension $1 \times D$, and is given by $y^\mathsf{T}(S^{-1} + (S^{-1})^\mathsf{T})$.

   Finally, $\frac{dy}{dx}$ has dimension $D \times D$, and is just the identity matrix.

Again, we multiply these all together to get our final derivative.

(b) If we explicitly write out $xx^\top + \sigma^2 I$, and compute its trace, we find that $f(x) = x_1^2 + \cdots + x_n^2 + n\sigma^2$.

Hence, $\frac{df}{dx} = 2x^\top$.

(c) Here, $\frac{df}{dz} = \begin{bmatrix} \frac{1}{\cosh^2 z_1} & 0 & \cdots & 0 \\ 0 & \frac{1}{\cosh^2 z_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{1}{\cosh^2 z_M} \end{bmatrix}$, while $\frac{dz}{dx} = A$, as in Question 7b.

Finally, $\frac{df}{dx}$ is given by the product of these two matrices.