Foundations of Math: Neural Networks

Winter 2023/24

Version 1 (January 30, 2024)

As discussed in class, some of the tasks regarding differentiation and Jacobians in the exam may be inspired by neural networks. Here is a collection of exercises for practicing this. Note that none of this material is specific to neural networks – rather, it reflects applications of partial derivatives, the chain rule, and the Jacobian.

There are a few hints in footnotes to help you in case you get stuck [these are bits of knowledge that we covered and that you should know for the exam].

I will provide solutions at the beginning of next week.

These exercises are all very similar, and build from a simple calculat[illegible] complex calculations.

因此，雅可比矩阵是一个 行向量：

$$Jac(f) = [2v_1^2x_1, 2v_2^2x_2, \ldots, 2v_n^2x_n]$$

或者写成矩阵形式：

$$Jac(f) = \begin{bmatrix} 2v_1^2x_1 & 2v_2^2x_2 & \ldots & 2v_n^2x_n \end{bmatrix}$$

这个结果是一个 $1 \times n$ 矩阵，因为 f 是一个标量函数，对 n 个变量求导后会形成一个行向量。

1. Let $f(x_1, x_2, x_3) = x_1 + 7x_2 - 5x_3$. What is $Jac(f)$?

2. Let $a, b, c \in \mathbb{R}$. Let $f(x_1, x_2, x_3) = ax_1 + bx_2 + cx_3$. What is $Jac(f)$?

3. Let $v \in \mathbb{R}^n$ a vector. Let $f(x_1, \ldots, x_n) = \sum_{i=1}^n v_i x_i$. What is $Jac(f)$?

4. Let $v \in \mathbb{R}^n$ a vector. Let $f(x_1, \ldots, x_n) = \sum_{i=1}^n v_i^2 x_i^2$. What is $Jac(f)$, written as a matrix? Hint: Use the multidimensional chain rule, in one of its several forms that we covered in class.

5. Let $v \in \mathbb{R}^n$ a vector. Let $f(x_1, \ldots, x_n) = \log(1 + \sum_{i=1}^n v_i^2 x_i^2)$. What is $Jac(f)$?

6. Let $f(x_1, x_2, x_3) = g(x_1)x_1 + g(x_2)x_2 + g(x_3)x_3$, where g is some function. What is $Jac(f)$, in terms of g'?

7. Let $g : \mathbb{R} \to \mathbb{R}$ be some function. Let $f(x_1, \ldots, x_n) = \sum_{i=1}^n g(x_i)x_i$. What is $Jac(f)$, in terms of g'?

8. Let $g : \mathbb{R} \to \mathbb{R}$ be some function. Let $f(x_1, \ldots, x_n) = \sum_{i=1}^n g(x_1)x_i$ Note the difference to the preceding question: $g(x_1)$ vs $g(x_i)$. What is $Jac(f)$, in terms of g'?

9. Let $g : \mathbb{R} \to \mathbb{R}$ be some function. Let $f(x_1, \ldots, x_n) = \log(1 + \sum_{i=1}^n g(x_i))$. What is $Jac(f)$, in terms of g'?

10. What is $\frac{\partial}{\partial x_2}(x_1 + 7x_2 - 5x_3)$?

11. What is $\frac{\partial}{\partial x_2}(x_1^2 + 7x_2^2 - 5x_3^2)$?

12. What is $\frac{\partial}{\partial x_2} \log(1 + x_1^2 + 7x_2^2 - 5x_3^2)$? [assuming the argument to the logarithm is positive]

13. Let $x \in \mathbb{R}^n$ be a vector.

    What is $\frac{\partial}{\partial x_1} x_1$?

    What is $\frac{\partial}{\partial x_1} x_2$?

    What is $\frac{\partial}{\partial x_2} x_1$?

    Let i be any element of $\{1, \ldots, n\}$. What is $\frac{\partial}{\partial x_i} x_i$?

    Let i, j be any element of $\{1, \ldots, n\}$. What is $\frac{\partial}{\partial x_i} x_j$? Give the result separately in the case where $i = j$ and where $i \neq j$.

    Let $a \in \mathbb{R}^n$ be a vector. Let i, j be any element of $\{1, \ldots, n\}$. What is $\frac{\partial}{\partial x_i}[a_j x_j]$? Give the result separately in the case where $i = j$ and where $i \neq j$.

    Let $a \in \mathbb{R}^n$ be a vector. Let i be any element of $\{1, \ldots, n\}$. What is $\frac{\partial}{\partial x_i} \sum_{j=1}^n a_j x_j$?

(2) $(Ax)_1$ 用 A 和 x 的元素表示

$$(Ax)_1 = \sum_{j=1}^{n} A_{1j} x_j$$

其中，A_{1j} 是矩阵 A 的第一行第 j 列的元素，x_j 是向量 x 的第 j 个分量。

一般化：

$$(Ax)_i = \sum_{j=1}^{n} A_{ij} x_j$$

14. Let $A \in \mathbb{R}^{n \times n}$ be a matrix. Let $x \in \mathbb{R}^n$ be a vector.

    Note that $Ax \in \mathbb{R}^n$? [why?]

    What is $(Ax)_1$, written in terms of the entries of A and x?[1]

    What is $\frac{\partial}{\partial x_i}((Ax)_1)$?[2]

    What is $\frac{\partial}{\partial x_i}((Ax)_2)$?[3]

    If $f(x) = Ax$, then what is $Jac(f)$? A

    Let $B \in \mathbb{R}^{n \times n}$ be another matrix. If $f(x) = BAx$, then what is $Jac(f)$? That is, give its entries by providing a formula for $\frac{\partial}{\partial x_i} f(x)$.

    Let $g : \mathbb{R}^n \to \mathbb{R}^n$ be any function. If $f(x) = g(Ax)$, then what is $Jac(f)$, in terms of $Jac(g)$? That is, give its entries by providing a formula for $\frac{\partial}{\partial x_i} f(x)$.

    Let $v \in \mathbb{R}$ be a vector, and g as in the last question. If $f(x) = v \cdot g(Ax)$, then what is $Jac(f)$? That is, give its entries by providing a formula for $\frac{\partial}{\partial x_i} f(x)$.

    What is $\frac{\partial}{\partial (A_{12})}((Ax)_1)$?[4]

    If $i, j \in \{1, \ldots, n\}$, then what is $\frac{\partial}{\partial (A_{ij})}((Ax)_1)$?

    If $i, j, k \in \{1, \ldots, n\}$, then what is $\frac{\partial}{\partial (A_{ij})}((Ax)_k)$?

因此，$f(x)$ 可以写为：

$$f(x) = BAx$$

2. 计算雅可比矩阵 雅可比矩阵的定义是：

$$\text{Jac}(f) = \left[\frac{\partial f_i}{\partial x_j}\right]_{i,j}$$

由于 $f(x) = BAx$ 是一个线性变换，其对 x 的导数就是矩阵 BA 本身。因此：

$$\text{Jac}(f) = BA$$

3. 结论 所以，函数 $f(x) = BAx$ 的雅可比矩阵 **就是** BA 矩阵本身，即：

$$\frac{\partial f}{\partial x} = BA$$

这表明 **线性变换的雅可比矩阵就是变换本身的矩阵表达形式。**

[1]Hint: it is $\sum_{l=1}^{n} A_{1l} x_l$ – why?
[2]Hint: $\frac{\partial}{\partial x_i} \sum_{l=1}^{n} A_{1l} x_l = A_{1i}$ – why?
[3]Hint: $\frac{\partial}{\partial x_i} \sum_{l=1}^{n} A_{2l} x_l = A_{2i}$ – why?
[4]Hint: $\frac{\partial}{\partial (A_{12})} \sum_{l=1}^{n} A_{1l} x_l = x_2$ – why?