M 384: Assignment 8

Nathan Stouffer

Exercises 10.1.5 — Problem 3

Problem. If f is differentiable at y, show that $d_u f(y)$ is linear in u, meaning $d_{au+bv} f(y) = ad_u f(y) + bd_v f(y)$ for any scalars a and b.

Proof. Before beginning, note that Theorem 10.1.1 says that if $f : D \to \mathbb{R}^m$ for $D \subset \mathbb{R}^n$ is differentiable at y with differential $df(y)$, then $d_u f(y)$ exists at y for any $u \in \mathbb{R}^n$ and $d_u f(y) = df(y) \cdot u$. Now let $w = au + bv$ for $u, v \in \mathbb{R}^n$ and scalars a, b.

Since $\mathbb{R}^n$ is a vector space, we have that $w \in \mathbb{R}^n$. Thus we have

$$d_{au+bvd} f(y) = d_w f(y) = df(y) \cdot w = df(y) \cdot (au + bv) = adf(y) \cdot u + bdf(y) \cdot v = ad_u f(y) + bd_v f(y)$$

where the crucial steps are justified by Theorem 10.1.1 and because scalars commute with matrices.

$\square$

Exercises 10.1.5 — Problem 10

Problem. Let $g : [a, b] \to \mathbb{R}^n$ be differentiable. If $f : \mathbb{R}^n \to \mathbb{R}^1$ is differentiable, what is the derivative $(d/dt)f(g(t))$.

Proof. Here we have $g(t) = (g_1(t), \ldots, g_n(t))$ for $t \in [a, b]$ and $f(z) = f(z_1, \ldots, z_n)$ where $z = (z_1, \ldots, z_n) \in \mathbb{R}^n$. Then $f \circ g = f(g(t)) = f(g_1(t), \ldots, g_n(t)) \in \mathbb{R}$. Consider $f \circ g$ and let $z_k = g_k(t)$. The textbook provides this formula for partial derivatives:

$$\frac{\partial f}{\partial x_j} = \sum_{k=1}^{n} \frac{\partial f}{\partial z_k} \left(\frac{\partial z_k}{\partial x_j} \right) = \begin{bmatrix} \frac{\partial f}{\partial z_1} & \cdots & \frac{\partial f}{\partial z_n} \end{bmatrix} \begin{bmatrix} \frac{\partial z_1}{\partial x_j} \\ \vdots \\ \frac{\partial z_n}{\partial x_j} \end{bmatrix} = \nabla f \cdot \begin{bmatrix} \frac{dg_1}{dt} \\ \vdots \\ \frac{dz_n}{dt} \end{bmatrix} = \nabla f \cdot \frac{dg}{dt}$$

where $\frac{dg}{dt}$ is a vector valued function. In this case, the domain is $[0, 1]$ so there is no need for x_j, which is why it is replaced with t.

$\square$

Exercises 10.1.5 — Problem 13

Problem. Compute df for the following functions.

a. $f : \mathbb{R}^2 \to \mathbb{R}^1$ $\quad f(x_1, x_2) = x_1 e^{x_2}$
b. $f : \mathbb{R}^3 \to \mathbb{R}^2$ $\quad f(x_1, x_2, x_3) = (x_3, x_2)$
c. $f : \mathbb{R}^2 \to \mathbb{R}^3$ $\quad f(x_1, x_2) = (x_1, x_2, x_1 x_2)$

Proof.
a. Since $f : \mathbb{R}^2 \to \mathbb{R}^1$ we know that df is a 1×2 matrix.

$$df = \begin{bmatrix} \frac{\partial f}{\partial x_1} & \frac{\partial f}{\partial x_2} \end{bmatrix} = \begin{bmatrix} e^{x_2} & x_1 e^{x_2} \end{bmatrix}$$

b. Since $f : \mathbb{R}^3 \to \mathbb{R}^2$ we know that df is a 2×3 matrix.

$$df = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \frac{\partial f_1}{\partial x_3} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \frac{\partial f_2}{\partial x_3} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

c. Since $f : \mathbb{R}^2 \to \mathbb{R}^3$ we know that df is a 3×2 matrix.

$$df = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} \\ \frac{\partial f_3}{\partial x_1} & \frac{\partial f_3}{\partial x_2} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ x_2 & x_1 \end{bmatrix}$$

$\square$

Exercises 10.1.5 — Problem 15

Problem. If $f : D \to \mathbb{R}$ is C^1 with $D \subset \mathbb{R}^n$ and D contains the line segment joining x and y, show that $f(y) = f(x) + \nabla f(z) \cdot (y - x)$ for some point z on the line segment. Explain why this is an n-dimensional analog of the mean value theorem.

Proof. For fixed vectors x and y, let $g : [0, 1] \to \mathbb{R}^n$ be defined by taking $t \mapsto x + t(y - x)$ and define $h(t) = (f \circ g)(t) = f(g(t))$ a map from $[0, 1] \to \mathbb{R}$. Since x, y are connected by a line segment contained in D we have that the image of $[0, 1]$ under g is a subset of D, which is where f is differentiable. Then since f is differentiable, we can apply the formula from two problems ago and say that

$$\frac{d}{dt} f(g(t)) = \nabla f \cdot \begin{bmatrix} \frac{d}{dt}(x_1 + t(y_1 - x_1)) \\ \vdots \\ \frac{d}{dt}(x_n + t(y_n - x_n)) \end{bmatrix} = \nabla f \cdot \begin{bmatrix} y_1 - x_1 \\ \vdots \\ y_n - x_n \end{bmatrix} = \nabla f \cdot (y - x)$$

Then since $h : [0, 1] \to \mathbb{R}$ is C^1 we can apply the MVT and say that there exists some $\lambda \in (0, 1)$ such that $h(\lambda) = h(1) - h(0) = f(g(1)) - f(g(0)) = f(y) - f(x)$. Taking $z = g(\lambda)$, we have that

$$f(y) - f(x) = \nabla f(z) \cdot (y - x) \iff f(y) = f(x) + \nabla f(z) \cdot (y - x)$$

This is an n-dimensional analogue of the MVT since it provides with only the existence of some point "in-between" two selected points for which the MVT condition holds.

$\square$