MAU23203: Analysis in Several Real Variables Michaelmas Term 2022 Disquisition IX: An Alternative Proof of the Multidimensional Chain Rule

David R. Wilkins

Lemma A *Let X be an open set in $\mathbb{R}^m$, let $\varphi\colon X \to \mathbb{R}^n$ be a function mapping X into $\mathbb{R}^n$, let $T\colon \mathbb{R}^m \to \mathbb{R}^n$ be a linear transformation from $\mathbb{R}^m \to \mathbb{R}^n$ and let $\mathbf{p}$ be a point belonging to the domain X of the function φ. Also let $\sigma\colon X \to \mathbb{R}^n$ be the function defined throughout the domain X of the function φ that is uniquely characterized by the properties that $\sigma(\mathbf{p}) = \mathbf{0}$ and*

$$\varphi(\mathbf{x}) = \varphi(\mathbf{p}) + T(\mathbf{x} - \mathbf{p}) + |\mathbf{x} - \mathbf{p}|\, \sigma(\mathbf{x})$$

for all points $\mathbf{x}$ of the domain X of the function φ. Then the function $\varphi\colon X \to \mathbb{R}^n$ is differentiable at the point $\mathbf{p}$, with derivative $T\colon \mathbb{R}^m \to \mathbb{R}^n$, if and only if the associated function σ is continuous at the point $\mathbf{p}$.

Proof Note that

$$\sigma(\mathbf{x}) = \begin{cases} \dfrac{1}{|\mathbf{x} - \mathbf{p}|} \left(\varphi(\mathbf{x}) - \varphi(\mathbf{p}) - T(\mathbf{x} - \mathbf{p})\right) & \text{if } \mathbf{x} \neq \mathbf{p}; \\ \mathbf{0} & \text{if } \mathbf{x} = \mathbf{p}. \end{cases}$$

The very definition of differentiability therefore ensures that the function φ is differentiable at the point $\mathbf{p}$, with derivative T, if and only if

$$\lim_{\mathbf{x} \to \mathbf{p}} \sigma(\mathbf{x}) = \mathbf{0} = \sigma(\mathbf{p}).$$

Moreover $\lim\limits_{\mathbf{x} \to \mathbf{p}} \sigma(\mathbf{x}) = \sigma(\mathbf{p})$ if and only if the function σ is continuous at the point $\mathbf{p}$ (see Proposition 6.5). The result follows. ∎

We recall the statement of the version of the Chain Rule that is applicable to compositions of vector-valued functions of several real variables (see Proposition 8.20).

The Chain Rule. *Let X and Y be open sets in $\mathbb{R}^m$ and $\mathbb{R}^n$ respectively, let $\varphi: X \to \mathbb{R}^n$ and $\psi: Y \to \mathbb{R}^k$ be functions mapping X and Y into $\mathbb{R}^n$ and $\mathbb{R}^k$ respectively, where $\varphi(X) \subset Y$, and let $\mathbf{p}$ be a point of X. Suppose that φ is differentiable at $\mathbf{p}$ and that ψ is differentiable at $\varphi(\mathbf{p})$. Then the composition $\psi \circ \varphi: X \to \mathbb{R}^k$ is differentiable at $\mathbf{p}$, and*

$$D(\psi \circ \varphi)_{\mathbf{p}} = (D\psi)_{\varphi(\mathbf{p})} \circ (D\varphi)_{\mathbf{p}}.$$

Thus the derivative of the composition $\psi \circ \varphi$ of the functions at the point $\mathbf{p}$ is the composition of the derivatives of the functions φ and ψ at $\mathbf{p}$ and $\varphi(\mathbf{p})$ respectively.

Proof Let $\mathbf{q} = \varphi(\mathbf{p})$, and let $\sigma: X \to \mathbb{R}^n$ and $\tau: Y \to \mathbb{R}^k$ be the uniquely-determined functions defined throughout the domains X and Y of the functions φ and ψ respectively so that $\sigma(\mathbf{p}) = \mathbf{0}$, $\tau(\mathbf{q}) = \mathbf{0}$,

$$\varphi(\mathbf{x}) = \varphi(\mathbf{p}) + (D\varphi)_{\mathbf{p}}(\mathbf{x} - \mathbf{p}) + |\mathbf{x} - \mathbf{p}|\, \sigma(\mathbf{x})$$

for all points $\mathbf{x}$ of the domain X of the function φ, and

$$\psi(\mathbf{y}) = \psi(\mathbf{q}) + (D\psi)_{\mathbf{q}}(\mathbf{y} - \mathbf{q}) + |\mathbf{y} - \mathbf{q}|\, \tau(\mathbf{y})$$

for all points $\mathbf{y}$ of the domain Y of the function ψ. The differentiability of the functions φ and ψ at the points $\mathbf{p}$ and $\mathbf{q}$ then ensures that the functions σ and τ are continuous at the points $\mathbf{p}$ and $\mathbf{q}$ respectively, where $\mathbf{q} = \varphi(\mathbf{p})$ (see Lemma A). Moreover the composition function $\tau \circ \varphi$ is continuous at the point $\mathbf{p}$, because the functions φ and τ are continuous at the points $\mathbf{p}$ and $\varphi(\mathbf{p})$ respectively (see Proposition 5.1).

The linearity of $(D\psi)_{\mathbf{q}}: \mathbb{R}^n \to \mathbb{R}^k$ then ensures that

$$\begin{aligned}
\psi(\varphi(\mathbf{x})) &= \psi(\mathbf{q}) + (D\psi)_{\mathbf{q}}(\varphi(\mathbf{x}) - \mathbf{q})) + |\varphi(\mathbf{x}) - \mathbf{q}|\, \tau(\varphi(\mathbf{x})) \\
&= \psi(\varphi(\mathbf{p})) + (D\psi)_{\mathbf{q}}(\varphi(\mathbf{x}) - \varphi(\mathbf{p}))) + |\varphi(\mathbf{x}) - \varphi(\mathbf{p})|\, \tau(\varphi(\mathbf{x})) \\
&= \psi(\varphi(\mathbf{p})) + (D\psi)_{\mathbf{q}}(D\varphi)_{\mathbf{p}}(\mathbf{x} - \mathbf{p}) + |\mathbf{x} - \mathbf{p}|(D\psi)_{\mathbf{q}}(\sigma(\mathbf{x})) \\
&\qquad + |\varphi(\mathbf{x}) - \varphi(\mathbf{p})|\, \tau(\varphi(\mathbf{x})) \\
&= \psi(\varphi(\mathbf{p})) + (D\psi)_{\mathbf{q}}(D\varphi)_{\mathbf{p}}(\mathbf{x} - \mathbf{p}) + |\mathbf{x} - \mathbf{p}|\chi(\mathbf{x})
\end{aligned}$$

for all $\mathbf{x} \in X$, where $\chi: X \to \mathbb{R}^k$ is the uniquely-determined function on the domain X of the function φ defined so that $\chi(\mathbf{p}) = 0$ and

$$\chi(\mathbf{x}) = (D\psi)_{\mathbf{q}}(\sigma(\mathbf{x})) + \frac{|\varphi(\mathbf{x}) - \varphi(\mathbf{p})|}{|\mathbf{x} - \mathbf{p}|}\, \tau(\varphi(\mathbf{x}))$$

for all points $\mathbf{x}$ of the set X that are distinct from the point $\mathbf{p}$. Thus, in order to complete the proof of the differentiability of the composition function $\psi \circ \varphi$ at the point $\mathbf{p}$, it suffices to show that that the function χ is continuous at the point $\mathbf{p}$ (see Lemma A), and moreover the continuity of the function χ at the point $\mathbf{p}$ can be established by verifying that $\lim_{\mathbf{x}\to\mathbf{p}} \chi(\mathbf{p}) = \mathbf{0}$.

Now $\lim_{\mathbf{x}\to\mathbf{p}} \sigma(\mathbf{x}) = \mathbf{0}$. The continuity of the linear transformation $(D\psi)_{\mathbf{q}}$ therefore ensures that

$$\lim_{\mathbf{x}\to\mathbf{p}} (D\psi)_{\mathbf{q}}(\sigma(\mathbf{x})) = (D\psi)_{\mathbf{q}} \left(\lim_{\mathbf{x}\to\mathbf{p}} \sigma(\mathbf{x}) \right) = (D\psi)_{\mathbf{q}}(\mathbf{0}) = \mathbf{0}.$$

Also there exist positive real numbers M and δ_0 such that $|\varphi(\mathbf{x}) - \varphi(\mathbf{p})| \leq M|\mathbf{x} - \mathbf{p}|$ whenever $|\mathbf{x} - \mathbf{p}| < \delta_0$ (see Proposition 8.18). Then, given any positive real number ε, there exists some real number δ satisfying $0 < \delta < \delta_0$ which is small enough to ensure that $|\tau(\varphi(\mathbf{x}))| < \varepsilon/M$ whenever $|\mathbf{x} - \mathbf{p}| < \delta$, because $\tau(\varphi(\mathbf{p})) = \tau(\mathbf{q}) = \mathbf{0}$ and the composition function $\tau \circ \varphi$ is continuous at the point $\mathbf{p}$. It follows that

$$\frac{|\varphi(\mathbf{x}) - \varphi(\mathbf{p})|}{|\mathbf{x} - \mathbf{p}|} |\tau(\varphi(\mathbf{x}))| < \varepsilon$$

whenever $|\mathbf{x} - \mathbf{p}| < \delta$. Consequently

$$\lim_{\mathbf{x}\to\mathbf{p}} \left(\frac{|\varphi(\mathbf{x}) - \varphi(\mathbf{p})|}{|\mathbf{x} - \mathbf{p}|} |\tau(\varphi(\mathbf{x}))| \right) = \mathbf{0}.$$

We can now conclude that

$$\lim_{\mathbf{x}\to\mathbf{p}} \chi(\mathbf{x}) = \lim_{\mathbf{x}\to\mathbf{p}} (D\psi)_{\mathbf{q}}(\sigma(\mathbf{x})) + \lim_{\mathbf{x}\to\mathbf{p}} \left(\frac{|\varphi(\mathbf{x}) - \varphi(\mathbf{p})|}{|\mathbf{x} - \mathbf{p}|} \tau(\varphi(\mathbf{x})) \right) = \mathbf{0} = \chi(\mathbf{p}),$$

and consequently the composition function $\psi \circ \varphi$ is differentiable at the point $\mathbf{p}$, with derivative $(D\psi)_{\varphi(\mathbf{p})} \circ (D\varphi)_{\mathbf{p}}$, as required. ∎