Advanced Calculus (II)

WEN-CHING LIEN

Department of Mathematics
National Cheng Kung University

2009

11.5: Mean Value Theorem and Taylor's Formula

Theorem (11.31 Mean Value Theorem on $\mathbf{R}^n$)

Let V be open in $\mathbf{R}^n$ and suppose that $f : V \to \mathbf{R}^m$ is differentiable on V. If $\mathbf{x}, \mathbf{a} \in V$ and $L(\mathbf{x}; \mathbf{a}) \subseteq V$, then for each $\mathbf{u} \in \mathbf{R}^m$, there is a $\mathbf{c} \in L(\mathbf{x}; \mathbf{a})$ such that

$$\mathbf{u} \cdot (f(\mathbf{x}) - f(\mathbf{a})) = \mathbf{u} \cdot (Df(\mathbf{c})(\mathbf{x} - \mathbf{a})).$$

Proof.

Let

$$g(t) = \mathbf{a} + t(\mathbf{x} - \mathbf{a}), \quad t \in \mathbf{R},$$

and notice by Exercise 8, p.338, that $g : \mathbf{R} \to \mathbf{R}^n$ is differentiable with $Dg(t) = \mathbf{x} - \mathbf{a}$ for all $t \in \mathbf{R}$. Since $L(\mathbf{x}; \mathbf{a}) \subseteq V$ and V is open, choose $\delta > 0$ such that $g(t) \in V$ for all $t \in I_\delta := (-\delta, 1 + \delta)$. By the Chain Rule,

$$D(f \circ g)(t) = Df(g(t))(\mathbf{x} - \mathbf{a}), \quad t \in I_\delta. \tag{24}$$

Fix $\mathbf{u} \in \mathbf{R}^m$, and consider the function

$$F(t) = \mathbf{u} \cdot (f \circ g)(t), \quad t \in I_\delta.$$

Proof.

Let

$$g(t) = \mathbf{a} + t(\mathbf{x} - \mathbf{a}), \quad t \in \mathbf{R},$$

and notice by Exercise 8, p.338, that $g : \mathbf{R} \to \mathbf{R}^n$ is differentiable with $Dg(t) = \mathbf{x} - \mathbf{a}$ for all $t \in \mathbf{R}$. Since $L(\mathbf{x}; \mathbf{a}) \subseteq V$ and V is open, choose $\delta > 0$ such that $g(t) \in V$ for all $t \in I_\delta := (-\delta, 1 + \delta)$. By the Chain Rule,

(24) $$D(f \circ g)(t) = Df(g(t))(\mathbf{x} - \mathbf{a}), \quad t \in I_\delta.$$

Fix $\mathbf{u} \in \mathbf{R}^m$, and consider the function

$$F(t) = \mathbf{u} \cdot (f \circ g)(t), \quad t \in I_\delta.$$

Proof.

Let

$$g(t) = \mathbf{a} + t(\mathbf{x} - \mathbf{a}), \quad t \in \mathbf{R},$$

and notice by Exercise 8, p.338, that $g : \mathbf{R} \to \mathbf{R}^n$ is differentiable with $Dg(t) = \mathbf{x} - \mathbf{a}$ for all $t \in \mathbf{R}$. Since $L(\mathbf{x}; \mathbf{a}) \subseteq V$ and V is open, choose $\delta > 0$ such that $g(t) \in V$ for all $t \in I_\delta := (-\delta, 1 + \delta)$. By the Chain Rule,

$$D(f \circ g)(t) = Df(g(t))(\mathbf{x} - \mathbf{a}), \quad t \in I_\delta. \tag{24}$$

Fix $\mathbf{u} \in \mathbf{R}^m$, and consider the function

$$F(t) = \mathbf{u} \cdot (f \circ g)(t), \quad t \in I_\delta.$$

Proof.

Let

$$g(t) = \mathbf{a} + t(\mathbf{x} - \mathbf{a}), \quad t \in \mathbf{R},$$

and notice by Exercise 8, p.338, that $g : \mathbf{R} \to \mathbf{R}^n$ is differentiable with $Dg(t) = \mathbf{x} - \mathbf{a}$ for all $t \in \mathbf{R}$. Since $L(\mathbf{x}; \mathbf{a}) \subseteq V$ and V is open, choose $\delta > 0$ such that $g(t) \in V$ for all $t \in I_\delta := (-\delta, 1 + \delta)$. By the Chain Rule,

$$D(f \circ g)(t) = Df(g(t))(\mathbf{x} - \mathbf{a}), \quad t \in I_\delta. \tag{24}$$

Fix $\mathbf{u} \in \mathbf{R}^m$, and consider the function

$$F(t) = \mathbf{u} \cdot (f \circ g)(t), \quad t \in I_\delta.$$

Proof.

Let

$$g(t) = \mathbf{a} + t(\mathbf{x} - \mathbf{a}), \quad t \in \mathbf{R},$$

and notice by Exercise 8, p.338, that $g : \mathbf{R} \to \mathbf{R}^n$ is differentiable with $Dg(t) = \mathbf{x} - \mathbf{a}$ for all $t \in \mathbf{R}$. Since $L(\mathbf{x}; \mathbf{a}) \subseteq V$ and V is open, choose $\delta > 0$ such that $g(t) \in V$ for all $t \in I_\delta := (-\delta, 1 + \delta)$. By the Chain Rule,

$$D(f \circ g)(t) = Df(g(t))(\mathbf{x} - \mathbf{a}), \quad t \in I_\delta. \tag{24}$$

Fix $\mathbf{u} \in \mathbf{R}^m$, and consider the function

$$F(t) = \mathbf{u} \cdot (f \circ g)(t), \quad t \in I_\delta.$$

□

Proof.

The function F is a real-valued function on I_δ. By the Dot Product Rule: $D(f \cdot g)(\mathbf{a}) = g(\mathbf{a})Df(\mathbf{a}) + f(\mathbf{a})Dg(\mathbf{a})$ and (24), F is differentiable on I_δ with

$$F'(t) = \mathbf{u} \cdot D(f \circ g)(t) = \mathbf{u} \cdot (Df(g(t))(\mathbf{x} - \mathbf{a})).$$

Hence, by the one-dimensional Mean Value Theorem, there is a $t_0 \in (0, 1)$ such that

$$\mathbf{u} \cdot (f(\mathbf{x}) - f(\mathbf{a})) = F(1) - F(0) = F'(t_0) = \mathbf{u} \cdot (Df(g(t_0))(\mathbf{x} - \mathbf{a})).$$

Thus set $\mathbf{c} = g(t_0)$. □

Proof.

The function F is a real-valued function on I_δ. By the Dot Product Rule: $D(f \cdot g)(\mathbf{a}) = g(\mathbf{a})Df(\mathbf{a}) + f(\mathbf{a})Dg(\mathbf{a})$ and (24), F is differentiable on I_δ with

$$F'(t) = \mathbf{u} \cdot D(f \circ g)(t) = \mathbf{u} \cdot (Df(g(t))(\mathbf{x} - \mathbf{a})).$$

Hence, by the one-dimensional Mean Value Theorem, there is a $t_0 \in (0, 1)$ such that

$$\mathbf{u} \cdot (f(\mathbf{x}) - f(\mathbf{a})) = F(1) - F(0) = F'(t_0) = \mathbf{u} \cdot (Df(g(t_0))(\mathbf{x} - \mathbf{a})).$$

Thus set $\mathbf{c} = g(t_0)$.

Proof.

The function F is a real-valued function on I_δ. By the Dot Product Rule: $D(f \cdot g)(\mathbf{a}) = g(\mathbf{a})Df(\mathbf{a}) + f(\mathbf{a})Dg(\mathbf{a})$ and (24), F is differentiable on I_δ with

$$F'(t) = \mathbf{u} \cdot D(f \circ g)(t) = \mathbf{u} \cdot (Df(g(t))(\mathbf{x} - \mathbf{a})).$$

Hence, by the one-dimensional Mean Value Theorem, there is a $t_0 \in (0, 1)$ such that

$$\mathbf{u} \cdot (f(\mathbf{x}) - f(\mathbf{a})) = F(1) - F(0) = F'(t_0) = \mathbf{u} \cdot (Df(g(t_0))(\mathbf{x} - \mathbf{a})).$$

Thus set $\mathbf{c} = g(t_0)$.

Proof.

The function F is a real-valued function on I_δ. By the Dot Product Rule: $D(f \cdot g)(\mathbf{a}) = g(\mathbf{a})Df(\mathbf{a}) + f(\mathbf{a})Dg(\mathbf{a})$ and (24), F is differentiable on I_δ with

$$F'(t) = \mathbf{u} \cdot D(f \circ g)(t) = \mathbf{u} \cdot (Df(g(t))(\mathbf{x} - \mathbf{a})).$$

Hence, by the one-dimensional Mean Value Theorem, there is a $t_0 \in (0, 1)$ such that

$$\mathbf{u}\cdot(f(\mathbf{x})-f(\mathbf{a})) = F(1)-F(0) = F'(t_0) = \mathbf{u}\cdot(Df(g(t_0))(\mathbf{x} - \mathbf{a})).$$

Thus set $\mathbf{c} = g(t_0)$.

Proof.

The function F is a real-valued function on I_δ. By the Dot Product Rule: $D(f \cdot g)(\mathbf{a}) = g(\mathbf{a})Df(\mathbf{a}) + f(\mathbf{a})Dg(\mathbf{a})$ and (24), F is differentiable on I_δ with

$$F'(t) = \mathbf{u} \cdot D(f \circ g)(t) = \mathbf{u} \cdot (Df(g(t))(\mathbf{x} - \mathbf{a})).$$

Hence, by the one-dimensional Mean Value Theorem, there is a $t_0 \in (0, 1)$ such that

$$\mathbf{u}\cdot(f(\mathbf{x})-f(\mathbf{a})) = F(1)-F(0) = F'(t_0) = \mathbf{u}\cdot(Df(g(t_0))(\mathbf{x} - \mathbf{a})).$$

Thus set $\mathbf{c} = g(t_0)$.

Proof.

The function F is a real-valued function on I_δ. By the Dot Product Rule: $D(f\cdot g)(\mathbf{a}) = g(\mathbf{a})Df(\mathbf{a}) + f(\mathbf{a})Dg(\mathbf{a})$ and (24), F is differentiable on I_δ with

$$F'(t) = \mathbf{u}\cdot D(f\circ g)(t) = \mathbf{u}\cdot(Df(g(t))(\mathbf{x}-\mathbf{a})).$$

Hence, by the one-dimensional Mean Value Theorem, there is a $t_0 \in (0,1)$ such that

$$\mathbf{u}\cdot(f(\mathbf{x})-f(\mathbf{a})) = F(1)-F(0) = F'(t_0) = \mathbf{u}\cdot(Df(g(t_0))(\mathbf{x}-\mathbf{a})).$$

Thus set $\mathbf{c} = g(t_0)$.

Proof.

The function F is a real-valued function on I_δ. By the Dot Product Rule: $D(f\cdot g)(\mathbf{a}) = g(\mathbf{a})Df(\mathbf{a}) + f(\mathbf{a})Dg(\mathbf{a})$ and (24), F is differentiable on I_δ with

$$F'(t) = \mathbf{u}\cdot D(f\circ g)(t) = \mathbf{u}\cdot(Df(g(t))(\mathbf{x}-\mathbf{a})).$$

Hence, by the one-dimensional Mean Value Theorem, there is a $t_0 \in (0,1)$ such that

$$\mathbf{u}\cdot(f(\mathbf{x})-f(\mathbf{a})) = F(1)-F(0) = F'(t_0) = \mathbf{u}\cdot(Df(g(t_0))(\mathbf{x}-\mathbf{a})).$$

Thus set $\mathbf{c} = g(t_0)$.

Proof.

The function F is a real-valued function on I_δ. By the Dot Product Rule: $D(f \cdot g)(\mathbf{a}) = g(\mathbf{a})Df(\mathbf{a}) + f(\mathbf{a})Dg(\mathbf{a})$ and (24), F is differentiable on I_δ with

$$F'(t) = \mathbf{u} \cdot D(f \circ g)(t) = \mathbf{u} \cdot (Df(g(t))(\mathbf{x} - \mathbf{a})).$$

Hence, by the one-dimensional Mean Value Theorem, there is a $t_0 \in (0, 1)$ such that

$$\mathbf{u} \cdot (f(\mathbf{x}) - f(\mathbf{a})) = F(1) - F(0) = F'(t_0) = \mathbf{u} \cdot (Df(g(t_0))(\mathbf{x} - \mathbf{a})).$$

Thus set $\mathbf{c} = g(t_0)$.

Proof.

The function F is a real-valued function on I_δ. By the Dot Product Rule: $D(f\cdot g)(\mathbf{a}) = g(\mathbf{a})Df(\mathbf{a}) + f(\mathbf{a})Dg(\mathbf{a})$ and (24), F is differentiable on I_δ with

$$F'(t) = \mathbf{u}\cdot D(f\circ g)(t) = \mathbf{u}\cdot(Df(g(t))(\mathbf{x}-\mathbf{a})).$$

Hence, by the one-dimensional Mean Value Theorem, there is a $t_0 \in (0,1)$ such that

$$\mathbf{u}\cdot(f(\mathbf{x})-f(\mathbf{a})) = F(1)-F(0) = F'(t_0) = \mathbf{u}\cdot(Df(g(t_0))(\mathbf{x}-\mathbf{a})).$$

Thus set $\mathbf{c} = g(t_0)$.

Proof.

The function F is a real-valued function on I_δ. By the Dot Product Rule: $D(f \cdot g)(\mathbf{a}) = g(\mathbf{a})Df(\mathbf{a}) + f(\mathbf{a})Dg(\mathbf{a})$ and (24), F is differentiable on I_δ with

$$F'(t) = \mathbf{u} \cdot D(f \circ g)(t) = \mathbf{u} \cdot (Df(g(t))(\mathbf{x} - \mathbf{a})).$$

Hence, by the one-dimensional Mean Value Theorem, there is a $t_0 \in (0, 1)$ such that

$$\mathbf{u} \cdot (f(\mathbf{x}) - f(\mathbf{a})) = F(1) - F(0) = F'(t_0) = \mathbf{u} \cdot (Df(g(t_0))(\mathbf{x} - \mathbf{a})).$$

Thus set $\mathbf{c} = g(t_0)$.

□

Definition (11.32)

A subset E of $\mathbf{R}^n$ is said to be *convex* if and only if $L(\mathbf{x}; \mathbf{a}) \subseteq E$ for all $\mathbf{x}, \mathbf{a} \in E$.

Corollary (11.33)

Let V be convex and open in $\mathbf{R}^n$ and suppose that $f : V \to \mathbf{R}$. If f is differentiable on V and $\mathbf{a}+\mathbf{h}$, $\mathbf{a}$ both belong to V, then there is a $0 < t < 1$ such that

$$(25) \quad f(\mathbf{a}+\mathbf{h})-f(\mathbf{a}) = Df(\mathbf{a}+t\mathbf{h})\cdot\mathbf{h} := \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(\mathbf{a}+t\mathbf{h})h_j.$$

Proof.

Let u be a nonzero scalar, and suppose that $\mathbf{a} + \mathbf{h}$, $\mathbf{a}$ both belong to V. Since V is convex, $L(\mathbf{a} + \mathbf{h}; \mathbf{a}) \subseteq V$. Hence, by Theorem 11.31,

$$u(f(\mathbf{a} + \mathbf{h}) - f(\mathbf{a})) = u(\nabla f(\mathbf{c}) \cdot \mathbf{h})$$

for some $\mathbf{c} \in L(\mathbf{a} + \mathbf{h}; \mathbf{a})$. Dividing this inequality by u and choosing $t \in (0, 1)$ such that $\mathbf{c} = \mathbf{a} + t\mathbf{h}$, we conclude that (25) holds. □

Proof.

Let u be a nonzero scalar, and suppose that $\mathbf{a}+\mathbf{h}$, $\mathbf{a}$ both belong to V. Since V is convex, $L(\mathbf{a}+\mathbf{h};\mathbf{a}) \subseteq V$. Hence, by Theorem 11.31,

$$u(f(\mathbf{a}+\mathbf{h})-f(\mathbf{a})) = u(\nabla f(\mathbf{c})\cdot\mathbf{h})$$

for some $\mathbf{c}\in L(\mathbf{a}+\mathbf{h};\mathbf{a})$. Dividing this inequality by u and choosing $t\in(0,1)$ such that $\mathbf{c}=\mathbf{a}+t\mathbf{h}$, we conclude that (25) holds. □

Proof.

Let u be a nonzero scalar, and suppose that $\mathbf{a}+\mathbf{h}$, $\mathbf{a}$ both belong to V. Since V is convex, $L(\mathbf{a}+\mathbf{h};\mathbf{a}) \subseteq V$. Hence, by Theorem 11.31,

$$u(f(\mathbf{a}+\mathbf{h}) - f(\mathbf{a})) = u(\nabla f(\mathbf{c}) \cdot \mathbf{h})$$

for some $\mathbf{c} \in L(\mathbf{a}+\mathbf{h};\mathbf{a})$. Dividing this inequality by u and choosing $t \in (0,1)$ such that $\mathbf{c} = \mathbf{a} + t\mathbf{h}$, we conclude that (25) holds. □

Proof.

Let u be a nonzero scalar, and suppose that $\mathbf{a}+\mathbf{h}$, $\mathbf{a}$ both belong to V. Since V is convex, $L(\mathbf{a}+\mathbf{h};\mathbf{a}) \subseteq V$. Hence, by Theorem 11.31,

$$u(f(\mathbf{a}+\mathbf{h}) - f(\mathbf{a})) = u(\nabla f(\mathbf{c}) \cdot \mathbf{h})$$

for some $\mathbf{c} \in L(\mathbf{a}+\mathbf{h};\mathbf{a})$. Dividing this inequality by u and choosing $t \in (0,1)$ such that $\mathbf{c} = \mathbf{a} + t\mathbf{h}$, we conclude that (25) holds.

□

Proof.

Let u be a nonzero scalar, and suppose that $\mathbf{a}+\mathbf{h}$, $\mathbf{a}$ both belong to V. Since V is convex, $L(\mathbf{a}+\mathbf{h};\mathbf{a}) \subseteq V$. Hence, by Theorem 11.31,

$$u(f(\mathbf{a}+\mathbf{h}) - f(\mathbf{a})) = u(\nabla f(\mathbf{c}) \cdot \mathbf{h})$$

for some $\mathbf{c} \in L(\mathbf{a}+\mathbf{h};\mathbf{a})$. Dividing this inequality by u and choosing $t \in (0,1)$ such that $\mathbf{c} = \mathbf{a} + t\mathbf{h}$, we conclude that (25) holds.

□

Corollary (11.34)

Let V be an open set in $\mathbf{R}^n$, let H be a compact subset of V, and suppose that $f : V \to \mathbf{R}^m$ is $\mathcal{C}^1$ on V. If E is a convex subset of H, then there is a constant M (which depends on H and f but not on E) such that

$$\|f(\mathbf{x}) - f(\mathbf{a})\| \leq M\|\mathbf{x} - \mathbf{a}\|$$

for all $\mathbf{x}, \mathbf{a} \in E$.

Proof.

Since H is compact and the entries of Df are continuous on H, we have by the Extreme Value Theorem (Theorem 9.32 or 10.63) and the proof of Theorem 8.17 that the operator norm of Df is bounded on H, i.e., that

$$M := \sup_{c \in H} \|Df(\mathbf{c})\|$$

is finite. Notice that M depends only on H and f.

Let $\mathbf{x}, \mathbf{a} \in E$ and $\mathbf{u} = f(\mathbf{x}) - f(\mathbf{a})$. Since E is convex, $L(\mathbf{x}; \mathbf{a}) \subseteq E$. Hence, by Theorem 11.31, there is a $\mathbf{c} \in L(\mathbf{x}; \mathbf{a})$ such that

$$\|f(\mathbf{x}) - f(\mathbf{a})\|^2 = \mathbf{u} \cdot (f(\mathbf{x}) - f(\mathbf{a})) = \mathbf{u} \cdot (Df(\mathbf{c})(\mathbf{x} - \mathbf{a}))$$

$$= (f(\mathbf{x}) - f(\mathbf{a})) \cdot (Df(\mathbf{c})(\mathbf{x} - \mathbf{a})).$$

Proof.

Since H is compact and the entries of Df are continuous on H, we have by the Extreme Value Theorem (Theorem 9.32 or 10.63) and the proof of Theorem 8.17 that the operator norm of Df is bounded on H, i.e., that

$$M := \sup_{c \in H} \|Df(\mathbf{c})\|$$

is finite. Notice that M depends only on H and f.

Let $\mathbf{x}, \mathbf{a} \in E$ and $\mathbf{u} = f(\mathbf{x}) - f(\mathbf{a})$. Since E is convex, $L(\mathbf{x}; \mathbf{a}) \subseteq E$. Hence, by Theorem 11.31, there is a $\mathbf{c} \in L(\mathbf{x}; \mathbf{a})$ such that

$$\|f(\mathbf{x}) - f(\mathbf{a})\|^2 = \mathbf{u} \cdot (f(\mathbf{x}) - f(\mathbf{a})) = \mathbf{u} \cdot (Df(\mathbf{c})(\mathbf{x} - \mathbf{a}))$$

$$= (f(\mathbf{x}) - f(\mathbf{a})) \cdot (Df(\mathbf{c})(\mathbf{x} - \mathbf{a})).$$

Proof.

Since H is compact and the entries of Df are continuous on H, we have by the Extreme Value Theorem (Theorem 9.32 or 10.63) and the proof of Theorem 8.17 that the operator norm of Df is bounded on H, i.e., that

$$M := \sup_{c \in H} \|Df(\mathbf{c})\|$$

is finite. Notice that M depends only on H and f. Let $\mathbf{x}, \mathbf{a} \in E$ and $\mathbf{u} = f(\mathbf{x}) - f(\mathbf{a})$. Since E is convex, $L(\mathbf{x}; \mathbf{a}) \subseteq E$. Hence, by Theorem 11.31, there is a $\mathbf{c} \in L(\mathbf{x}; \mathbf{a})$ such that

$$\|f(\mathbf{x}) - f(\mathbf{a})\|^2 = \mathbf{u} \cdot (f(\mathbf{x}) - f(\mathbf{a})) = \mathbf{u} \cdot (Df(\mathbf{c})(\mathbf{x} - \mathbf{a}))$$

$$= (f(\mathbf{x}) - f(\mathbf{a})) \cdot (Df(\mathbf{c})(\mathbf{x} - \mathbf{a})).$$

Proof.

Since H is compact and the entries of Df are continuous on H, we have by the Extreme Value Theorem (Theorem 9.32 or 10.63) and the proof of Theorem 8.17 that the operator norm of Df is bounded on H, i.e., that

$$M := \sup_{c \in H} \|Df(\mathbf{c})\|$$

is finite. Notice that M depends only on H and f.

Let $\mathbf{x}, \mathbf{a} \in E$ and $\mathbf{u} = f(\mathbf{x}) - f(\mathbf{a})$. Since E is convex, $L(\mathbf{x}; \mathbf{a}) \subseteq E$. Hence, by Theorem 11.31, there is a $\mathbf{c} \in L(\mathbf{x}; \mathbf{a})$ such that

$$\|f(\mathbf{x}) - f(\mathbf{a})\|^2 = \mathbf{u} \cdot (f(\mathbf{x}) - f(\mathbf{a})) = \mathbf{u} \cdot (Df(\mathbf{c})(\mathbf{x} - \mathbf{a}))$$

$$= (f(\mathbf{x}) - f(\mathbf{a})) \cdot (Df(\mathbf{c})(\mathbf{x} - \mathbf{a})).$$

Proof.

Since H is compact and the entries of Df are continuous on H, we have by the Extreme Value Theorem (Theorem 9.32 or 10.63) and the proof of Theorem 8.17 that the operator norm of Df is bounded on H, i.e., that

$$M := \sup_{c \in H} \|Df(\mathbf{c})\|$$

is finite. Notice that M depends only on H and f.
Let $\mathbf{x}, \mathbf{a} \in E$ and $\mathbf{u} = f(\mathbf{x}) - f(\mathbf{a})$. Since E is convex, $L(\mathbf{x}; \mathbf{a}) \subseteq E$. Hence, by Theorem 11.31, there is a $\mathbf{c} \in L(\mathbf{x}; \mathbf{a})$ such that

$$\begin{aligned}\|f(\mathbf{x}) - f(\mathbf{a})\|^2 = \mathbf{u} \cdot (f(\mathbf{x}) - f(\mathbf{a})) &= \mathbf{u} \cdot (Df(\mathbf{c})(\mathbf{x} - \mathbf{a})) \\ &= (f(\mathbf{x}) - f(\mathbf{a})) \cdot (Df(\mathbf{c})(\mathbf{x} - \mathbf{a})).\end{aligned}$$

Proof.

Since H is compact and the entries of Df are continuous on H, we have by the Extreme Value Theorem (Theorem 9.32 or 10.63) and the proof of Theorem 8.17 that the operator norm of Df is bounded on H, i.e., that

$$M := \sup_{c \in H} \|Df(\mathbf{c})\|$$

is finite. Notice that M depends only on H and f.

Let $\mathbf{x}, \mathbf{a} \in E$ and $\mathbf{u} = f(\mathbf{x}) - f(\mathbf{a})$. Since E is convex, $L(\mathbf{x}; \mathbf{a}) \subseteq E$. Hence, by Theorem 11.31, there is a $\mathbf{c} \in L(\mathbf{x}; \mathbf{a})$ such that

$$\|f(\mathbf{x}) - f(\mathbf{a})\|^2 = \mathbf{u} \cdot (f(\mathbf{x}) - f(\mathbf{a})) = \mathbf{u} \cdot (Df(\mathbf{c})(\mathbf{x} - \mathbf{a}))$$

$$= (f(\mathbf{x}) - f(\mathbf{a})) \cdot (Df(\mathbf{c})(\mathbf{x} - \mathbf{a})).$$

Proof.

Since H is compact and the entries of Df are continuous on H, we have by the Extreme Value Theorem (Theorem 9.32 or 10.63) and the proof of Theorem 8.17 that the operator norm of Df is bounded on H, i.e., that

$$M := \sup_{c \in H} \|Df(\mathbf{c})\|$$

is finite. Notice that M depends only on H and f.

Let $\mathbf{x}, \mathbf{a} \in E$ and $\mathbf{u} = f(\mathbf{x}) - f(\mathbf{a})$. Since E is convex, $L(\mathbf{x}; \mathbf{a}) \subseteq E$. Hence, by Theorem 11.31, there is a $\mathbf{c} \in L(\mathbf{x}; \mathbf{a})$ such that

$$\|f(\mathbf{x}) - f(\mathbf{a})\|^2 = \mathbf{u} \cdot (f(\mathbf{x}) - f(\mathbf{a})) = \mathbf{u} \cdot (Df(\mathbf{c})(\mathbf{x} - \mathbf{a}))$$

$$= (f(\mathbf{x}) - f(\mathbf{a})) \cdot (Df(\mathbf{c})(\mathbf{x} - \mathbf{a})).$$

Proof.

Since H is compact and the entries of Df are continuous on H, we have by the Extreme Value Theorem (Theorem 9.32 or 10.63) and the proof of Theorem 8.17 that the operator norm of Df is bounded on H, i.e., that

$$M := \sup_{c \in H} \|Df(\mathbf{c})\|$$

is finite. Notice that M depends only on H and f.

Let $\mathbf{x}, \mathbf{a} \in E$ and $\mathbf{u} = f(\mathbf{x}) - f(\mathbf{a})$. Since E is convex, $L(\mathbf{x}; \mathbf{a}) \subseteq E$. Hence, by Theorem 11.31, there is a $\mathbf{c} \in L(\mathbf{x}; \mathbf{a})$ such that

$$\|f(\mathbf{x}) - f(\mathbf{a})\|^2 = \mathbf{u} \cdot (f(\mathbf{x}) - f(\mathbf{a})) = \mathbf{u} \cdot (Df(\mathbf{c})(\mathbf{x} - \mathbf{a}))$$

$$= (f(\mathbf{x}) - f(\mathbf{a})) \cdot (Df(\mathbf{c})(\mathbf{x} - \mathbf{a})).$$

Proof.

Since H is compact and the entries of Df are continuous on H, we have by the Extreme Value Theorem (Theorem 9.32 or 10.63) and the proof of Theorem 8.17 that the operator norm of Df is bounded on H, i.e., that

$$M := \sup_{c \in H} \|Df(\mathbf{c})\|$$

is finite. Notice that M depends only on H and f.

Let $\mathbf{x}, \mathbf{a} \in E$ and $\mathbf{u} = f(\mathbf{x}) - f(\mathbf{a})$. Since E is convex, $L(\mathbf{x}; \mathbf{a}) \subseteq E$. Hence, by Theorem 11.31, there is a $\mathbf{c} \in L(\mathbf{x}; \mathbf{a})$ such that

$$\|f(\mathbf{x}) - f(\mathbf{a})\|^2 = \mathbf{u} \cdot (f(\mathbf{x}) - f(\mathbf{a})) = \mathbf{u} \cdot (Df(\mathbf{c})(\mathbf{x} - \mathbf{a}))$$

$$= (f(\mathbf{x}) - f(\mathbf{a})) \cdot (Df(\mathbf{c})(\mathbf{x} - \mathbf{a})).$$

Proof.

Since H is compact and the entries of Df are continuous on H, we have by the Extreme Value Theorem (Theorem 9.32 or 10.63) and the proof of Theorem 8.17 that the operator norm of Df is bounded on H, i.e., that

$$M := \sup_{c \in H} \|Df(\mathbf{c})\|$$

is finite. Notice that M depends only on H and f.

Let $\mathbf{x}, \mathbf{a} \in E$ and $\mathbf{u} = f(\mathbf{x}) - f(\mathbf{a})$. Since E is convex, $L(\mathbf{x}; \mathbf{a}) \subseteq E$. Hence, by Theorem 11.31, there is a $\mathbf{c} \in L(\mathbf{x}; \mathbf{a})$ such that

$$\|f(\mathbf{x}) - f(\mathbf{a})\|^2 = \mathbf{u} \cdot (f(\mathbf{x}) - f(\mathbf{a})) = \mathbf{u} \cdot (Df(\mathbf{c})(\mathbf{x} - \mathbf{a}))$$

$$= (f(\mathbf{x}) - f(\mathbf{a})) \cdot (Df(\mathbf{c})(\mathbf{x} - \mathbf{a})).$$

Proof.

It follows from the Cauchy-Schwarz Inequality and the definition of the operator norm that

$$\|f(\mathbf{x}) - f(\mathbf{a})\|^2 \leq \|f(\mathbf{x}) - f(\mathbf{a})\| \|Df(\mathbf{c})\| \|\mathbf{x} - \mathbf{a}\|.$$

If $\|f(\mathbf{x}) - f(\mathbf{a})\| = 0$, there is nothing to prove. Otherwise, we can divide the inequality above by $\|f(\mathbf{x}) - f(\mathbf{a})\|$ to obtain

$$\|f(\mathbf{x}) - f(\mathbf{a})\| \leq \|Df(\mathbf{c})\| \|\mathbf{x} - \mathbf{a}\| \leq M\|\mathbf{x} - \mathbf{a}\|.$$

Proof.

It follows from the Cauchy-Schwarz Inequality and the definition of the operator norm that

$$\|f(\mathbf{x}) - f(\mathbf{a})\|^2 \le \|f(\mathbf{x}) - f(\mathbf{a})\| \|Df(\mathbf{c})\| \|\mathbf{x} - \mathbf{a}\|.$$

If $\|f(\mathbf{x}) - f(\mathbf{a})\| = 0$, there is nothing to prove. Otherwise, we can divide the inequality above by $\|f(\mathbf{x}) - f(\mathbf{a})\|$ to obtain

$$\|f(\mathbf{x}) - f(\mathbf{a})\| \le \|Df(\mathbf{c})\| \|\mathbf{x} - \mathbf{a}\| \le M\|\mathbf{x} - \mathbf{a}\|.$$

Proof.

It follows from the Cauchy-Schwarz Inequality and the definition of the operator norm that

$$\|f(\mathbf{x}) - f(\mathbf{a})\|^2 \leq \|f(\mathbf{x}) - f(\mathbf{a})\| \|Df(\mathbf{c})\| \|\mathbf{x} - \mathbf{a}\|.$$

If $\|f(\mathbf{x}) - f(\mathbf{a})\| = 0$, there is nothing to prove. Otherwise, we can divide the inequality above by $\|f(\mathbf{x}) - f(\mathbf{a})\|$ to obtain

$$\|f(\mathbf{x}) - f(\mathbf{a})\| \leq \|Df(\mathbf{c})\| \|\mathbf{x} - \mathbf{a}\| \leq M \|\mathbf{x} - \mathbf{a}\|.$$

Proof.

It follows from the Cauchy-Schwarz Inequality and the definition of the operator norm that

$$\|f(\mathbf{x}) - f(\mathbf{a})\|^2 \leq \|f(\mathbf{x}) - f(\mathbf{a})\| \|Df(\mathbf{c})\| \|\mathbf{x} - \mathbf{a}\|.$$

If $\|f(\mathbf{x}) - f(\mathbf{a})\| = 0$, there is nothing to prove. Otherwise, we can divide the inequality above by $\|f(\mathbf{x}) - f(\mathbf{a})\|$ to obtain

$$\|f(\mathbf{x}) - f(\mathbf{a})\| \leq \|Df(\mathbf{c})\| \|\mathbf{x} - \mathbf{a}\| \leq M\|\mathbf{x} - \mathbf{a}\|.$$

Proof.

It follows from the Cauchy-Schwarz Inequality and the definition of the operator norm that

$$\|f(\mathbf{x}) - f(\mathbf{a})\|^2 \leq \|f(\mathbf{x}) - f(\mathbf{a})\|\|Df(\mathbf{c})\|\|\mathbf{x} - \mathbf{a}\|.$$

If $\|f(\mathbf{x}) - f(\mathbf{a})\| = 0$, there is nothing to prove. Otherwise, we can divide the inequality above by $\|f(\mathbf{x}) - f(\mathbf{a})\|$ to obtain

$$\|f(\mathbf{x}) - f(\mathbf{a})\| \leq \|Df(\mathbf{c})\|\|\mathbf{x} - \mathbf{a}\| \leq M\|\mathbf{x} - \mathbf{a}\|.$$

Proof.

It follows from the Cauchy-Schwarz Inequality and the definition of the operator norm that

$$\|f(\mathbf{x}) - f(\mathbf{a})\|^2 \leq \|f(\mathbf{x}) - f(\mathbf{a})\| \|Df(\mathbf{c})\| \|\mathbf{x} - \mathbf{a}\|.$$

If $\|f(\mathbf{x}) - f(\mathbf{a})\| = 0$, there is nothing to prove. Otherwise, we can divide the inequality above by $\|f(\mathbf{x}) - f(\mathbf{a})\|$ to obtain

$$\|f(\mathbf{x}) - f(\mathbf{a})\| \leq \|Df(\mathbf{c})\| \|\mathbf{x} - \mathbf{a}\| \leq M \|\mathbf{x} - \mathbf{a}\|.$$

□

Corollary (11.35)

Suppose that V is open and connected in $\mathbf{R}^n$ and that $f : V \to \mathbf{R}^m$ is differentiable on V. If $Df(\mathbf{c}) = O$ for all $\mathbf{c} \in V$, then f is constant on V.

Proof.

Fix $\mathbf{a} \in V$, and let $\mathbf{x} \in V$. Since V is open and connected, V is polygonally connected (see Exercise 10, p.277). Thus, there exist points $\mathbf{x}_0 = \mathbf{a}, \mathbf{x}_1, \ldots, \mathbf{x}_k = \mathbf{x}$ such that $L(\mathbf{x}_{j-1}; \mathbf{x}_j)$ for $j = 1, 2, \ldots, k$ (see Figure 11.4).

Let $\mathbf{u} = f(\mathbf{x}) - f(\mathbf{a})$ and choose by Theorem 11.31 points $\mathbf{c}_j \in L(\mathbf{x}_{j-1}; \mathbf{x}_j)$ such that

$$\mathbf{u} \cdot (f(\mathbf{x}_j) - f(\mathbf{x}_{j-1})) = \mathbf{u} \cdot (Df(\mathbf{c}_j)(\mathbf{x}_j - \mathbf{x}_{j-1})) = 0$$

for $j = 1, 2, \ldots, k$. Summing over j and telescoping, we see by the choice of $\mathbf{u}$ that

$$0 = \sum_{j=1}^{k} \mathbf{u} \cdot (f(\mathbf{x}_j) - f(\mathbf{x}_{j-1})) = \mathbf{u} \cdot (f(\mathbf{x}) - f(\mathbf{a})) = \|f(\mathbf{x}) - f(\mathbf{a})\|^2.$$

Therefore, $f(\mathbf{x}) = f(\mathbf{a})$.

Proof.

Fix $\mathbf{a} \in V$, and let $\mathbf{x} \in V$. Since V is open and connected, V is polygonally connected (see Exercise 10, p.277). Thus, there exist points $\mathbf{x}_0 = \mathbf{a}, \mathbf{x}_1, \ldots, \mathbf{x}_k = \mathbf{x}$ such that $L(\mathbf{x}_{j-1}; \mathbf{x}_j)$ for $j = 1, 2, \ldots, k$ (see Figure 11.4).

Let $\mathbf{u} = f(\mathbf{x}) - f(\mathbf{a})$ and choose by Theorem 11.31 points $\mathbf{c}_j \in L(\mathbf{x}_{j-1}; \mathbf{x}_j)$ such that

$$\mathbf{u} \cdot (f(\mathbf{x}_j) - f(\mathbf{x}_{j-1})) = \mathbf{u} \cdot (Df(\mathbf{c}_j)(\mathbf{x}_j - \mathbf{x}_{j-1})) = 0$$

for $j = 1, 2, \ldots, k$. Summing over j and telescoping, we see by the choice of $\mathbf{u}$ that

$$0 = \sum_{j=1}^{k} \mathbf{u} \cdot (f(\mathbf{x}_j) - f(\mathbf{x}_{j-1})) = \mathbf{u} \cdot (f(\mathbf{x}) - f(\mathbf{a})) = \|f(\mathbf{x}) - f(\mathbf{a})\|^2.$$

Therefore, $f(\mathbf{x}) = f(\mathbf{a})$.

Proof.

Fix $\mathbf{a} \in V$, and let $\mathbf{x} \in V$. Since V is open and connected, V is polygonally connected (see Exercise 10, p.277).
Thus, there exist points $\mathbf{x}_0 = \mathbf{a}, \mathbf{x}_1, \ldots, \mathbf{x}_k = \mathbf{x}$ such that $L(\mathbf{x}_{j-1}; \mathbf{x}_j)$ for $j = 1, 2, \ldots, k$ (see Figure 11.4).
Let $\mathbf{u} = f(\mathbf{x}) - f(\mathbf{a})$ and choose by Theorem 11.31 points $\mathbf{c}_j \in L(\mathbf{x}_{j-1}; \mathbf{x}_j)$ such that

$$\mathbf{u} \cdot (f(\mathbf{x}_j) - f(\mathbf{x}_{j-1})) = \mathbf{u} \cdot (Df(\mathbf{c}_j)(\mathbf{x}_j - \mathbf{x}_{j-1})) = 0$$

for $j = 1, 2, \ldots, k$. Summing over j and telescoping, we see by the choice of $\mathbf{u}$ that

$$0 = \sum_{j=1}^{k} \mathbf{u} \cdot (f(\mathbf{x}_j) - f(\mathbf{x}_{j-1})) = \mathbf{u} \cdot (f(\mathbf{x}) - f(\mathbf{a})) = \|f(\mathbf{x}) - f(\mathbf{a})\|^2.$$

Therefore, $f(\mathbf{x}) = f(\mathbf{a})$.

Proof.

Fix $\mathbf{a} \in V$, and let $\mathbf{x} \in V$. Since V is open and connected, V is polygonally connected (see Exercise 10, p.277). Thus, there exist points $\mathbf{x}_0 = \mathbf{a}, \mathbf{x}_1, \dots, \mathbf{x}_k = \mathbf{x}$ such that $L(\mathbf{x}_{j-1}; \mathbf{x}_j)$ for $j = 1, 2, \dots, k$ (see Figure 11.4).

Let $\mathbf{u} = f(\mathbf{x}) - f(\mathbf{a})$ and choose by Theorem 11.31 points $\mathbf{c}_j \in L(\mathbf{x}_{j-1}; \mathbf{x}_j)$ such that

$$\mathbf{u} \cdot (f(\mathbf{x}_j) - f(\mathbf{x}_{j-1})) = \mathbf{u} \cdot (Df(\mathbf{c}_j)(\mathbf{x}_j - \mathbf{x}_{j-1})) = 0$$

for $j = 1, 2, \dots, k$. Summing over j and telescoping, we see by the choice of $\mathbf{u}$ that

$$0 = \sum_{j=1}^{k} \mathbf{u} \cdot (f(\mathbf{x}_j) - f(\mathbf{x}_{j-1})) = \mathbf{u} \cdot (f(\mathbf{x}) - f(\mathbf{a})) = \|f(\mathbf{x}) - f(\mathbf{a})\|^2.$$

Therefore, $f(\mathbf{x}) = f(\mathbf{a})$.

Proof.

Fix $\mathbf{a} \in V$, and let $\mathbf{x} \in V$. Since V is open and connected, V is polygonally connected (see Exercise 10, p.277). Thus, there exist points $\mathbf{x}_0 = \mathbf{a}, \mathbf{x}_1, \ldots, \mathbf{x}_k = \mathbf{x}$ such that $L(\mathbf{x}_{j-1}; \mathbf{x}_j)$ for $j = 1, 2, \ldots, k$ (see Figure 11.4).

Let $\mathbf{u} = f(\mathbf{x}) - f(\mathbf{a})$ and choose by Theorem 11.31 points $\mathbf{c}_j \in L(\mathbf{x}_{j-1}; \mathbf{x}_j)$ such that

$$\mathbf{u} \cdot (f(\mathbf{x}_j) - f(\mathbf{x}_{j-1})) = \mathbf{u} \cdot (Df(\mathbf{c}_j)(\mathbf{x}_j - \mathbf{x}_{j-1})) = 0$$

for $j = 1, 2, \ldots, k$. Summing over j and telescoping, we see by the choice of $\mathbf{u}$ that

$$0 = \sum_{j=1}^{k} \mathbf{u} \cdot (f(\mathbf{x}_j) - f(\mathbf{x}_{j-1})) = \mathbf{u} \cdot (f(\mathbf{x}) - f(\mathbf{a})) = \|f(\mathbf{x}) - f(\mathbf{a})\|^2.$$

Therefore, $f(\mathbf{x}) = f(\mathbf{a})$.

Proof.

Fix $\mathbf{a} \in V$, and let $\mathbf{x} \in V$. Since V is open and connected, V is polygonally connected (see Exercise 10, p.277). Thus, there exist points $\mathbf{x}_0 = \mathbf{a}, \mathbf{x}_1, \ldots, \mathbf{x}_k = \mathbf{x}$ such that $L(\mathbf{x}_{j-1}; \mathbf{x}_j)$ for $j = 1, 2, \ldots, k$ (see Figure 11.4).

Let $\mathbf{u} = f(\mathbf{x}) - f(\mathbf{a})$ and choose by Theorem 11.31 points $\mathbf{c}_j \in L(\mathbf{x}_{j-1}; \mathbf{x}_j)$ such that

$$\mathbf{u} \cdot (f(\mathbf{x}_j) - f(\mathbf{x}_{j-1})) = \mathbf{u} \cdot (Df(\mathbf{c}_j)(\mathbf{x}_j - \mathbf{x}_{j-1})) = 0$$

for $j = 1, 2, \ldots, k$. Summing over j and telescoping, we see by the choice of $\mathbf{u}$ that

$$0 = \sum_{j=1}^{k} \mathbf{u} \cdot (f(\mathbf{x}_j) - f(\mathbf{x}_{j-1})) = \mathbf{u} \cdot (f(\mathbf{x}) - f(\mathbf{a})) = \|f(\mathbf{x}) - f(\mathbf{a})\|^2.$$

Therefore, $f(\mathbf{x}) = f(\mathbf{a})$.

Proof.

Fix $\mathbf{a} \in V$, and let $\mathbf{x} \in V$. Since V is open and connected, V is polygonally connected (see Exercise 10, p.277). Thus, there exist points $\mathbf{x}_0 = \mathbf{a}, \mathbf{x}_1, \dots, \mathbf{x}_k = \mathbf{x}$ such that $L(\mathbf{x}_{j-1}; \mathbf{x}_j)$ for $j = 1, 2, \dots, k$ (see Figure 11.4).

Let $\mathbf{u} = f(\mathbf{x}) - f(\mathbf{a})$ and choose by Theorem 11.31 points $\mathbf{c}_j \in L(\mathbf{x}_{j-1}; \mathbf{x}_j)$ such that

$$\mathbf{u} \cdot (f(\mathbf{x}_j) - f(\mathbf{x}_{j-1})) = \mathbf{u} \cdot (Df(\mathbf{c}_j)(\mathbf{x}_j - \mathbf{x}_{j-1})) = 0$$

for $j = 1, 2, \dots, k$. Summing over j and telescoping, we see by the choice of $\mathbf{u}$ that

$$0 = \sum_{j=1}^{k} \mathbf{u} \cdot (f(\mathbf{x}_j) - f(\mathbf{x}_{j-1})) = \mathbf{u} \cdot (f(\mathbf{x}) - f(\mathbf{a})) = \|f(\mathbf{x}) - f(\mathbf{a})\|^2.$$

Therefore, $f(\mathbf{x}) = f(\mathbf{a})$.

Proof.

Fix $\mathbf{a} \in V$, and let $\mathbf{x} \in V$. Since V is open and connected, V is polygonally connected (see Exercise 10, p.277). Thus, there exist points $\mathbf{x}_0 = \mathbf{a}, \mathbf{x}_1, \ldots, \mathbf{x}_k = \mathbf{x}$ such that $L(\mathbf{x}_{j-1}; \mathbf{x}_j)$ for $j = 1, 2, \ldots, k$ (see Figure 11.4).

Let $\mathbf{u} = f(\mathbf{x}) - f(\mathbf{a})$ and choose by Theorem 11.31 points $\mathbf{c}_j \in L(\mathbf{x}_{j-1}; \mathbf{x}_j)$ such that

$$\mathbf{u} \cdot (f(\mathbf{x}_j) - f(\mathbf{x}_{j-1})) = \mathbf{u} \cdot (Df(\mathbf{c}_j)(\mathbf{x}_j - \mathbf{x}_{j-1})) = 0$$

for $j = 1, 2, \ldots, k$. Summing over j and telescoping, we see by the choice of $\mathbf{u}$ that

$$0 = \sum_{j=1}^{k} \mathbf{u} \cdot (f(\mathbf{x}_j) - f(\mathbf{x}_{j-1})) = \mathbf{u} \cdot (f(\mathbf{x}) - f(\mathbf{a})) = \|f(\mathbf{x}) - f(\mathbf{a})\|^2.$$

Therefore, $f(\mathbf{x}) = f(\mathbf{a})$.

Proof.

Fix $\mathbf{a} \in V$, and let $\mathbf{x} \in V$. Since V is open and connected, V is polygonally connected (see Exercise 10, p.277). Thus, there exist points $\mathbf{x}_0 = \mathbf{a}, \mathbf{x}_1, \ldots, \mathbf{x}_k = \mathbf{x}$ such that $L(\mathbf{x}_{j-1}; \mathbf{x}_j)$ for $j = 1, 2, \ldots, k$ (see Figure 11.4).

Let $\mathbf{u} = f(\mathbf{x}) - f(\mathbf{a})$ and choose by Theorem 11.31 points $\mathbf{c}_j \in L(\mathbf{x}_{j-1}; \mathbf{x}_j)$ such that

$$\mathbf{u} \cdot (f(\mathbf{x}_j) - f(\mathbf{x}_{j-1})) = \mathbf{u} \cdot (Df(\mathbf{c}_j)(\mathbf{x}_j - \mathbf{x}_{j-1})) = 0$$

for $j = 1, 2, \ldots, k$. Summing over j and telescoping, we see by the choice of $\mathbf{u}$ that

$$0 = \sum_{j=1}^{k} \mathbf{u} \cdot (f(\mathbf{x}_j) - f(\mathbf{x}_{j-1})) = \mathbf{u} \cdot (f(\mathbf{x}) - f(\mathbf{a})) = \|f(\mathbf{x}) - f(\mathbf{a})\|^2.$$

Therefore, $f(\mathbf{x}) = f(\mathbf{a})$.

Proof.

Fix $\mathbf{a} \in V$, and let $\mathbf{x} \in V$. Since V is open and connected, V is polygonally connected (see Exercise 10, p.277). Thus, there exist points $\mathbf{x}_0 = \mathbf{a}, \mathbf{x}_1, \ldots, \mathbf{x}_k = \mathbf{x}$ such that $L(\mathbf{x}_{j-1}; \mathbf{x}_j)$ for $j = 1, 2, \ldots, k$ (see Figure 11.4).

Let $\mathbf{u} = f(\mathbf{x}) - f(\mathbf{a})$ and choose by Theorem 11.31 points $\mathbf{c}_j \in L(\mathbf{x}_{j-1}; \mathbf{x}_j)$ such that

$$\mathbf{u} \cdot (f(\mathbf{x}_j) - f(\mathbf{x}_{j-1})) = \mathbf{u} \cdot (Df(\mathbf{c}_j)(\mathbf{x}_j - \mathbf{x}_{j-1})) = 0$$

for $j = 1, 2, \ldots, k$. Summing over j and telescoping, we see by the choice of $\mathbf{u}$ that

$$0 = \sum_{j=1}^{k} \mathbf{u} \cdot (f(\mathbf{x}_j) - f(\mathbf{x}_{j-1})) = \mathbf{u} \cdot (f(\mathbf{x}) - f(\mathbf{a})) = \|f(\mathbf{x}) - f(\mathbf{a})\|^2.$$

Therefore, $f(\mathbf{x}) = f(\mathbf{a})$.

Proof.

Fix $\mathbf{a} \in V$, and let $\mathbf{x} \in V$. Since V is open and connected, V is polygonally connected (see Exercise 10, p.277). Thus, there exist points $\mathbf{x}_0 = \mathbf{a}, \mathbf{x}_1, \ldots, \mathbf{x}_k = \mathbf{x}$ such that $L(\mathbf{x}_{j-1}; \mathbf{x}_j)$ for $j = 1, 2, \ldots, k$ (see Figure 11.4).

Let $\mathbf{u} = f(\mathbf{x}) - f(\mathbf{a})$ and choose by Theorem 11.31 points $\mathbf{c}_j \in L(\mathbf{x}_{j-1}; \mathbf{x}_j)$ such that

$$\mathbf{u} \cdot (f(\mathbf{x}_j) - f(\mathbf{x}_{j-1})) = \mathbf{u} \cdot (Df(\mathbf{c}_j)(\mathbf{x}_j - \mathbf{x}_{j-1})) = 0$$

for $j = 1, 2, \ldots, k$. Summing over j and telescoping, we see by the choice of $\mathbf{u}$ that

$$0 = \sum_{j=1}^{k} \mathbf{u} \cdot (f(\mathbf{x}_j) - f(\mathbf{x}_{j-1})) = \mathbf{u} \cdot (f(\mathbf{x}) - f(\mathbf{a})) = \|f(\mathbf{x}) - f(\mathbf{a})\|^2.$$

Therefore, $f(\mathbf{x}) = f(\mathbf{a})$.

Theorem (11.37 Taylor's Formula on $\mathbf{R}^n$)

Let $p \in \mathbf{N}$, let V be open in $\mathbf{R}^n$, let $\mathbf{x}, \mathbf{a} \in V$, and suppose that $f : V \to \mathbf{R}$. If the pth total total differential of f exists on V and $L(\mathbf{x};\mathbf{a}) \subseteq V$, then there is a point $\mathbf{c} \in L(\mathbf{x};\mathbf{a})$ such that

$$f(\mathbf{x}) = f(\mathbf{a}) + \sum_{k=1}^{p-1} \frac{1}{k!} D^{(k)} f(\mathbf{a};\mathbf{h}) + \frac{1}{p!} D^{(p)} f(\mathbf{c};\mathbf{h})$$

for $\mathbf{h} := \mathbf{x} - \mathbf{a}$.

Note: These hypotheses are met if V is convex and f is $\mathcal{C}^p$ on V.

Proof.

Let $\mathbf{h} = \mathbf{x} - \mathbf{a}$. As in the proof Theorem 11.31, choose $\delta > 0$ so small that $\mathbf{a} + t\mathbf{h} \subset V$ for $t \in I_\delta := (-\delta, 1 + \delta)$. The function $F(t) = f(\mathbf{a} + t\mathbf{h})$ is differentiable on I_δ and, by the Chain Rule,

$$F'(t) = Df(\mathbf{a} + t\mathbf{h})(\mathbf{h}) = \sum_{k=1}^{n} \frac{\partial f}{\partial x_k}(\mathbf{a} + t\mathbf{h})h_k.$$

In fact, a simple induction argument can be used to verify

$$F^{(j)}(t) = \sum_{i_1=1}^{n} \cdots \sum_{i_j=1}^{n} \frac{\partial^j f}{\partial x_{i_1} \ldots \partial x_{i_j}}(\mathbf{a} + t\mathbf{h})h_{i_1} \ldots h_{i_j}$$

for $j = 1, 2, \ldots, p$. Thus

Proof.

Let $\mathbf{h} = \mathbf{x} - \mathbf{a}$. As in the proof Theorem 11.31, choose $\delta > 0$ so small that $\mathbf{a} + t\mathbf{h} \subset V$ for $t \in I_\delta := (-\delta, 1 + \delta)$. The function $F(t) = f(\mathbf{a} + t\mathbf{h})$ is differentiable on I_δ and, by the Chain Rule,

$$F'(t) = Df(\mathbf{a} + t\mathbf{h})(\mathbf{h}) = \sum_{k=1}^{n} \frac{\partial f}{\partial x_k}(\mathbf{a} + t\mathbf{h})h_k.$$

In fact, a simple induction argument can be used to verify

$$F^{(j)}(t) = \sum_{i_1=1}^{n} \cdots \sum_{i_j=1}^{n} \frac{\partial^j f}{\partial x_{i_1} \ldots \partial x_{i_j}}(\mathbf{a} + t\mathbf{h})h_{i_1} \ldots h_{i_j}$$

for $j = 1, 2, \ldots, p$. Thus

Proof.

Let $\mathbf{h} = \mathbf{x} - \mathbf{a}$. As in the proof Theorem 11.31, choose $\delta > 0$ so small that $\mathbf{a} + t\mathbf{h} \subset V$ for $t \in I_\delta := (-\delta, 1+\delta)$. The function $F(t) = f(\mathbf{a} + t\mathbf{h})$ is differentiable on I_δ and, by the Chain Rule,

$$F'(t) = Df(\mathbf{a} + t\mathbf{h})(\mathbf{h}) = \sum_{k=1}^{n} \frac{\partial f}{\partial x_k}(\mathbf{a} + t\mathbf{h})h_k.$$

In fact, a simple induction argument can be used to verify

$$F^{(j)}(t) = \sum_{i_1=1}^{n} \cdots \sum_{i_j=1}^{n} \frac{\partial^j f}{\partial x_{i_1} \ldots \partial x_{i_j}}(\mathbf{a} + t\mathbf{h})h_{i_1} \ldots h_{i_j}$$

for $j = 1, 2, \ldots, p$. Thus

Proof.

Let $\mathbf{h} = \mathbf{x} - \mathbf{a}$. As in the proof Theorem 11.31, choose $\delta > 0$ so small that $\mathbf{a} + t\mathbf{h} \subset V$ for $t \in I_\delta := (-\delta, 1 + \delta)$. The function $F(t) = f(\mathbf{a} + t\mathbf{h})$ is differentiable on I_δ and, by the Chain Rule,

$$F'(t) = Df(\mathbf{a} + t\mathbf{h})(\mathbf{h}) = \sum_{k=1}^{n} \frac{\partial f}{\partial x_k}(\mathbf{a} + t\mathbf{h})h_k.$$

In fact, a simple induction argument can be used to verify

$$F^{(j)}(t) = \sum_{i_1=1}^{n} \cdots \sum_{i_j=1}^{n} \frac{\partial^j f}{\partial x_{i_1} \ldots \partial x_{i_j}}(\mathbf{a} + t\mathbf{h})h_{i_1} \ldots h_{i_j}$$

for $j = 1, 2, \ldots, p$. Thus

Proof.

Let $\mathbf{h} = \mathbf{x} - \mathbf{a}$. As in the proof Theorem 11.31, choose $\delta > 0$ so small that $\mathbf{a} + t\mathbf{h} \subset V$ for $t \in I_\delta := (-\delta, 1 + \delta)$. The function $F(t) = f(\mathbf{a} + t\mathbf{h})$ is differentiable on I_δ and, by the Chain Rule,

$$F'(t) = Df(\mathbf{a} + t\mathbf{h})(\mathbf{h}) = \sum_{k=1}^{n} \frac{\partial f}{\partial x_k}(\mathbf{a} + t\mathbf{h}) h_k.$$

In fact, a simple induction argument can be used to verify

$$F^{(j)}(t) = \sum_{i_1=1}^{n} \cdots \sum_{i_j=1}^{n} \frac{\partial^j f}{\partial x_{i_1} \ldots \partial x_{i_j}}(\mathbf{a} + t\mathbf{h}) h_{i_1} \ldots h_{i_j}$$

for $j = 1, 2, \ldots, p$. Thus

Proof.

Let $\mathbf{h} = \mathbf{x} - \mathbf{a}$. As in the proof Theorem 11.31, choose $\delta > 0$ so small that $\mathbf{a} + t\mathbf{h} \subset V$ for $t \in I_\delta := (-\delta, 1 + \delta)$. The function $F(t) = f(\mathbf{a} + t\mathbf{h})$ is differentiable on I_δ and, by the Chain Rule,

$$F'(t) = Df(\mathbf{a} + t\mathbf{h})(\mathbf{h}) = \sum_{k=1}^{n} \frac{\partial f}{\partial x_k}(\mathbf{a} + t\mathbf{h})h_k.$$

In fact, a simple induction argument can be used to verify

$$F^{(j)}(t) = \sum_{i_1=1}^{n} \cdots \sum_{i_j=1}^{n} \frac{\partial^j f}{\partial x_{i_1} \ldots \partial x_{i_j}}(\mathbf{a} + t\mathbf{h})h_{i_1} \ldots h_{i_j}$$

for $j = 1, 2, \ldots, p$. Thus

Proof.

Let $\mathbf{h} = \mathbf{x} - \mathbf{a}$. As in the proof Theorem 11.31, choose $\delta > 0$ so small that $\mathbf{a} + t\mathbf{h} \subset V$ for $t \in I_\delta := (-\delta, 1 + \delta)$. The function $F(t) = f(\mathbf{a} + t\mathbf{h})$ is differentiable on I_δ and, by the Chain Rule,

$$F'(t) = Df(\mathbf{a} + t\mathbf{h})(\mathbf{h}) = \sum_{k=1}^{n} \frac{\partial f}{\partial x_k}(\mathbf{a} + t\mathbf{h})h_k.$$

In fact, a simple induction argument can be used to verify

$$F^{(j)}(t) = \sum_{i_1=1}^{n} \cdots \sum_{i_j=1}^{n} \frac{\partial^j f}{\partial x_{i_1} \ldots \partial x_{i_j}}(\mathbf{a} + t\mathbf{h})h_{i_1} \ldots h_{i_j}$$

for $j = 1, 2, \ldots, p$. Thus

Proof.

Let $\mathbf{h} = \mathbf{x} - \mathbf{a}$. As in the proof Theorem 11.31, choose $\delta > 0$ so small that $\mathbf{a} + t\mathbf{h} \subset V$ for $t \in I_\delta := (-\delta, 1+\delta)$. The function $F(t) = f(\mathbf{a} + t\mathbf{h})$ is differentiable on I_δ and, by the Chain Rule,

$$F'(t) = Df(\mathbf{a} + t\mathbf{h})(\mathbf{h}) = \sum_{k=1}^{n} \frac{\partial f}{\partial x_k}(\mathbf{a} + t\mathbf{h})h_k.$$

In fact, a simple induction argument can be used to verify

$$F^{(j)}(t) = \sum_{i_1=1}^{n} \cdots \sum_{i_j=1}^{n} \frac{\partial^j f}{\partial x_{i_1} \ldots \partial x_{i_j}}(\mathbf{a} + t\mathbf{h})h_{i_1} \ldots h_{i_j}$$

for $j = 1, 2, \ldots .p$. Thus

□

Proof.

$$(26) \quad F^{(j)}(0) = D^{(j)}f(\mathbf{a};\mathbf{h}) \text{ and } F^{(p)}(t) = D^{(p)}f(\mathbf{a}+t\mathbf{h};\mathbf{h})$$

for $j = 1, \dots, p-1$, and $t \in I_\delta$.

We have proved that $F : I_\delta \to \mathbf{R}$ has a derivative of order p everywhere on $I_\delta \supset [0,1]$. Therefore, by the one-dimensional Taylor Formula and (26),

$$f(\mathbf{x}) - f(\mathbf{a}) = F(1) - F(0) = \sum_{j=1}^{p-1} \frac{1}{j!} F^{(j)}(0) + \frac{1}{p!} F^{(p)}(t)$$

$$= \sum_{j=1}^{p-1} \frac{1}{j!} D^{(j)} f(\mathbf{a};\mathbf{h}) + \frac{1}{p!} D^{(p)} f(\mathbf{a}+t\mathbf{h};\mathbf{h})$$

for some $t \in (0,1)$. Thus set $\mathbf{c} = \mathbf{a} + t\mathbf{h}$.

Proof.

$$F^{(j)}(0) = D^{(j)}f(\mathbf{a};\mathbf{h}) \textit{ and } F^{(p)}(t) = D^{(p)}f(\mathbf{a}+t\mathbf{h};\mathbf{h}) \tag{26}$$

for $j = 1, \dots, p-1$, and $t \in I_\delta$.
We have proved that $F : I_\delta \to \mathbf{R}$ has a derivative of order p everywhere on $I_\delta \supset [0, 1]$. Therefore, by the one-dimensional Taylor Formula and (26),

$$f(\mathbf{x}) - f(\mathbf{a}) = F(1) - F(0) = \sum_{j=1}^{p-1} \frac{1}{j!} F^{(j)}(0) + \frac{1}{p!} F^{(p)}(t)$$

$$= \sum_{j=1}^{p-1} \frac{1}{j!} D^{(j)} f(\mathbf{a};\mathbf{h}) + \frac{1}{p!} D^{(p)} f(\mathbf{a} + t\mathbf{h};\mathbf{h})$$

for some $t \in (0, 1)$. Thus set $\mathbf{c} = \mathbf{a} + t\mathbf{h}$.

Proof.

$$(26) \quad F^{(j)}(0) = D^{(j)}f(\mathbf{a};\mathbf{h}) \text{ and } F^{(p)}(t) = D^{(p)}f(\mathbf{a}+t\mathbf{h};\mathbf{h})$$

for $j = 1, \dots, p-1$, and $t \in I_\delta$.
We have proved that $F : I_\delta \to \mathbf{R}$ has a derivative of order p everywhere on $I_\delta \supset [0, 1]$. Therefore, by the one-dimensional Taylor Formula and (26),

$$f(\mathbf{x}) - f(\mathbf{a}) = F(1) - F(0) = \sum_{j=1}^{p-1} \frac{1}{j!} F^{(j)}(0) + \frac{1}{p!} F^{(p)}(t)$$

$$= \sum_{j=1}^{p-1} \frac{1}{j!} D^{(j)} f(\mathbf{a};\mathbf{h}) + \frac{1}{p!} D^{(p)} f(\mathbf{a} + t\mathbf{h};\mathbf{h})$$

for some $t \in (0, 1)$. Thus set $\mathbf{c} = \mathbf{a} + t\mathbf{h}$.

Proof.

$$F^{(j)}(0) = D^{(j)}f(\mathbf{a};\mathbf{h}) \text{ and } F^{(p)}(t) = D^{(p)}f(\mathbf{a}+t\mathbf{h};\mathbf{h}) \tag{26}$$

for $j = 1, \ldots, p-1$, and $t \in I_\delta$.
We have proved that $F : I_\delta \to \mathbf{R}$ has a derivative of order p everywhere on $I_\delta \supset [0,1]$. Therefore, by the one-dimensional Taylor Formula and (26),

$$f(\mathbf{x}) - f(\mathbf{a}) = F(1) - F(0) = \sum_{j=1}^{p-1} \frac{1}{j!}F^{(j)}(0) + \frac{1}{p!}F^{(p)}(t)$$

$$= \sum_{j=1}^{p-1} \frac{1}{j!}D^{(j)}f(\mathbf{a};\mathbf{h}) + \frac{1}{p!}D^{(p)}f(\mathbf{a}+t\mathbf{h};\mathbf{h})$$

for some $t \in (0,1)$. Thus set $\mathbf{c} = \mathbf{a} + t\mathbf{h}$.

Proof.

$$(26) \quad F^{(j)}(0) = D^{(j)}f(\mathbf{a};\mathbf{h}) \textit{ and } F^{(p)}(t) = D^{(p)}f(\mathbf{a}+t\mathbf{h};\mathbf{h})$$

for $j = 1, \ldots, p-1$, and $t \in I_\delta$.

We have proved that $F : I_\delta \to \mathbf{R}$ has a derivative of order p everywhere on $I_\delta \supset [0,1]$. Therefore, by the one-dimensional Taylor Formula and (26),

$$f(\mathbf{x}) - f(\mathbf{a}) = F(1) - F(0) = \sum_{j=1}^{p-1} \frac{1}{j!} F^{(j)}(0) + \frac{1}{p!} F^{(p)}(t)$$

$$= \sum_{j=1}^{p-1} \frac{1}{j!} D^{(j)} f(\mathbf{a};\mathbf{h}) + \frac{1}{p!} D^{(p)} f(\mathbf{a}+t\mathbf{h};\mathbf{h})$$

for some $t \in (0,1)$. Thus set $\mathbf{c} = \mathbf{a} + t\mathbf{h}$.

Proof.

$$F^{(j)}(0) = D^{(j)}f(\mathbf{a};\mathbf{h}) \text{ and } F^{(p)}(t) = D^{(p)}f(\mathbf{a}+t\mathbf{h};\mathbf{h}) \tag{26}$$

for $j = 1, \ldots, p-1$, and $t \in I_\delta$.
We have proved that $F : I_\delta \to \mathbf{R}$ has a derivative of order p everywhere on $I_\delta \supset [0,1]$. Therefore, by the one-dimensional Taylor Formula and (26),

$$f(\mathbf{x}) - f(\mathbf{a}) = F(1) - F(0) = \sum_{j=1}^{p-1} \frac{1}{j!}F^{(j)}(0) + \frac{1}{p!}F^{(p)}(t)$$

$$= \sum_{j=1}^{p-1} \frac{1}{j!}D^{(j)}f(\mathbf{a};\mathbf{h}) + \frac{1}{p!}D^{(p)}f(\mathbf{a}+t\mathbf{h};\mathbf{h})$$

for some $t \in (0,1)$. Thus set $\mathbf{c} = \mathbf{a} + t\mathbf{h}$.

Proof.

$$(26) \quad F^{(j)}(0) = D^{(j)}f(\mathbf{a};\mathbf{h}) \textit{ and } F^{(p)}(t) = D^{(p)}f(\mathbf{a}+t\mathbf{h};\mathbf{h})$$

for $j = 1, \ldots, p-1$, and $t \in I_\delta$.
We have proved that $F : I_\delta \to \mathbf{R}$ has a derivative of order p everywhere on $I_\delta \supset [0,1]$. Therefore, by the one-dimensional Taylor Formula and (26),

$$f(\mathbf{x}) - f(\mathbf{a}) = F(1) - F(0) = \sum_{j=1}^{p-1} \frac{1}{j!} F^{(j)}(0) + \frac{1}{p!} F^{(p)}(t)$$

$$= \sum_{j=1}^{p-1} \frac{1}{j!} D^{(j)} f(\mathbf{a};\mathbf{h}) + \frac{1}{p!} D^{(p)} f(\mathbf{a}+t\mathbf{h};\mathbf{h})$$

for some $t \in (0,1)$. Thus set $\mathbf{c} = \mathbf{a} + t\mathbf{h}$.

Proof.

$$(26) \quad F^{(j)}(0) = D^{(j)} f(\mathbf{a}; \mathbf{h}) \textit{ and } F^{(p)}(t) = D^{(p)} f(\mathbf{a}+t\mathbf{h}; \mathbf{h})$$

for $j = 1, \ldots, p-1$, and $t \in I_\delta$.

We have proved that $F : I_\delta \to \mathbf{R}$ has a derivative of order p everywhere on $I_\delta \supset [0, 1]$. Therefore, by the one-dimensional Taylor Formula and (26),

$$f(\mathbf{x}) - f(\mathbf{a}) = F(1) - F(0) = \sum_{j=1}^{p-1} \frac{1}{j!} F^{(j)}(0) + \frac{1}{p!} F^{(p)}(t)$$

$$= \sum_{j=1}^{p-1} \frac{1}{j!} D^{(j)} f(\mathbf{a}; \mathbf{h}) + \frac{1}{p!} D^{(p)} f(\mathbf{a} + t\mathbf{h}; \mathbf{h})$$

for some $t \in (0, 1)$. Thus set $\mathbf{c} = \mathbf{a} + t\mathbf{h}$.

Thank you.