Partial derivatives

**Notation and Terminology.** Thoughout this section $X, Y, Z, \ldots$ will denote finite-dimen-
sional normed spaces.

**Definition 1.** Let $E \subset X \times Y$, $f : E \to Z$, and $(a, b) \in E^o$. Put $g = f(\cdot, b)$. The partial
derivative of $f$ with respect to the 1st coordinate at $(a, b)$ is

$$D_1f(a, b) := g'(a).$$

Similarly, the partial derivative of $f$ with respect to the 2nd coordinate at $(a, b)$ is

$$D_2f(a, b) := h'(b),$$

where $h = f(a, \cdot)$.

Similarly for functions of $n$ variables.

**Observation 2.** With the above notation, letting

$$C = \{x \in X : (x, b) \in E\},$$

we have $a \in C^0$, so $g'(a)$ makes sense.

Thus, the partial derivative with respect to one coordinate is the derivative of the function
obtained by holding all other coordinates constant. With the above notation,

$$D_1f(a, b) \in B(X, Z) \text{ and } D_2f(a, b) \in B(Y, Z).$$

However, if $X = \mathbb{R}$ we identify $D_1f(a, b)$ with an element of $Z$, so in particular if $E \subset \mathbb{R}^n$
and $Z = \mathbb{R}$ then the partial derivatives $D_i f(a_1, \ldots, a_n)$ are numbers.

**Examples 3.**

1. Let $E = \{(x, y) \in \mathbb{R}^2 : x \geq 0, y > 0\}$, and define $f : E \to \mathbb{R}$ by

$$f(x, y) = \sqrt{x} \log y.$$  

Then

$$D_1f(x, y) = \frac{\log y}{2\sqrt{x}},$$  

$$D_2f(x, y) = \frac{\sqrt{x}}{y}.$$

2. Define $f : \mathbb{R}^3 \to \mathbb{R}$ by

$$f(x, y, z) = xy^2z^3.$$  

Then

$$D_2f(x, y, z) = 2xyz^3,$$

so

$$D_2f(-1, 3, 2) = -48.$$
Notation and Terminology. If confusion is unlikely we sometimes write
\[
\frac{\partial f}{\partial x}(x, y) = D_1 f(x, y)
\]
or
\[
\frac{\partial}{\partial x_i} f(x_1, \ldots, x_n) = D_i f(x_1, \ldots, x_n).
\]

Examples 4.

(1) If \( f(x, y) = x^3 y^2 \) then \( \frac{\partial f}{\partial y}(x, y) = 2x^3 y \).

(2) \( \frac{\partial}{\partial z} x^2 y^3 z^4 = 4x^2 y^3 z^3 \).

Proposition 5. Let \( E \subset X \times Y, f : E \to Z \), and \( (a, b) \in E \). Assume that \( f \) is differentiable at \( (a, b) \). Then:

(1) Both partial derivatives \( D_1 f \) and \( D_2 f \) exist at \( (a, b) \), and
\[
f'(a, b) = (D_1 f(a, b) \ D_2 f(a, b))
\]

(2) \( f' \) is continuous at \( (a, b) \) if and only if both \( D_1 f \) and \( D_2 f \) are.

Proof. (1) Let \( C = \{ x \in X : (x, b) \in E \} \), and define \( h : C \to E \) by \( h(x) = (x, b) \). Then
\[
h(x) = (x, 0) + (0, b),
\]
so \( h \) is (the restriction to \( C \) of) a linear function plus a constant function. Thus \( h \) is differentiable at \( a \), with
\[
h'(a)x = (x, 0).
\]
We have \( f(\cdot, b) = f \circ h \), so by the Chain rule \( D_1 f(a, b) \) exists and
\[
f'(a, b)(x, 0) = f'(h(a))h'(a)x = (f \circ h)'(a)x = D_1 f(a, b)x,
\]
and similarly
\[
f'(a, b)(0, y) = D_2 f(a, b)y.
\]
Since
\[
f'(a, b)(x, y) = f'(a, b)((x, 0) + (0, y)) = f'(a, b)(x, 0) + f'(a, b)(0, y),
\]
we see that the matrix of the linear map \( f'(a, b) \) is \( (D_1 f(a, b) \ D_2 f(a, b)) \).

(2) follows from
\[
B(X \times Y, Z) \cong B(X, Z) \times B(Y, Z).
\]

Corollary 6. Let \( E \subset \mathbb{R}^n \), \( f : E \to \mathbb{R}^m \), and \( a \in E \). If \( f \) is differentiable at \( a \), then all the partial derivatives \( D_j f_i(a) \) exist for \( i = 1, \ldots, n, j = 1, \ldots, m \), and the matrix \( (D_j f_i(a)) \) represents \( f'(a) \) relative to the standard bases. Moreover, \( f' \) is continuous at \( a \) if and only if every \( D_j f_i \) is.
Proof. We have

\[ f'(a) = \begin{pmatrix} f'_1(a) \\ \vdots \\ f'_m(a) \end{pmatrix}, \]

and by Proposition 5 each \( f'_i(a) \) is represented by the row matrix

\[ (D_1f_i(a) \quad \cdots \quad D_nf_i(a)). \]

The result follows. \( \square \)

Example 7. Let

\[ A = \{(x, y) \in \mathbb{R}^2 : x, y > 0\}, \]

and put \( f = \chi_A \). Then \( D_1f \) and \( D_2f \) both exist at \((0, 0)\), but \( f \) is not differentiable there. The nondifferentiability can be shown directly, or by noticing that \( D_{(1,1)}f(0,0) \) does not exist.

Example 8. Let

\[ A = \{(x, y) \in \mathbb{R}^2 : 0 < y < x^2\}, \]

and let \( f = \chi_A \). Then \( D_uf(0,0) = 0 \) for all \( u \in \mathbb{R}^2 \), but \( f \) is nondifferentiable at \((0,0)\).

Observation 9. Let \( E \subset \mathbb{R}^n \), \( f : E \to \mathbb{R} \), and \( a \in E \).

1. If all the partial derivatives of \( f \) exist at \( a \), then the gradient vector is given by

\[ \nabla f(a) = (D_1f(a), \ldots, D_nf(a)). \]

2. Now also let \( I \subset \mathbb{R} \), \( g : I \to E \), and \( t \in I \). Suppose that \( g \) is differentiable at \( t \) and \( f \) is differentiable at \( g(t) \). Then

\[ (f \circ g)'(t) = \sum_{i=1}^{n} D_i f(g(t))g'_i(t). \]

Example 10. For all \( r \in \mathbb{R} \) we have

\[ \nabla \|x\|^r = r\|x\|^{r-2}x \]

because

\[ D_i\|x\|^r = \frac{\partial}{\partial x_i} \left( \sum_{j=1}^{n} x_j^2 \right)^{r/2} = r \left( \sum_{j=1}^{n} x_j^2 \right)^{r/2-1} \cdot 2x_i = r\|x\|^{r-2}x_i. \]