Formula for a derivative

We used the tangent plane to the graph of a function f of 2 variables to argue that the linear map

$$L(h,k) = f_x(c,d)h + f_y(c,d)k$$

approximates $f(c+h,d+k) - f(c,d)$ very well when h and k are small. It can be shown, with the help of the Mean Value Theorem, that if the partial derivatives $f_x(x,y)$ and $f_y(x,y)$ are continuous at (c,d) then indeed.

$$\lim_{||(h,k)|| \to 0} \frac{|f(c+h,d+k) - f(c,d) - L(h,k)|}{||(h,k)||} = 0$$

Therefore if f has continuous partial derivatives then its derivative at (c,d) has formula

$$Df(c,d)(h,k) = hf_x(c,d) + kf_y(c,d) \tag{1}$$

In matrix form, it is

$$Df(c,d)(h,k) = \begin{pmatrix} f_x(c,d) & f_y(c,d) \end{pmatrix} \begin{pmatrix} h \\ k \end{pmatrix}$$

The 1×2 matrix $\begin{pmatrix} f_x(c,d) & f_y(c,d) \end{pmatrix}$ is called the Jacobian matrix for f at (c,d).

Exercise 1

1. *Use formula (1) to determine $Df(c,d)(h,k)$ given f and point (c,d):*

   (a) $f(x,y) = 4 - 3xy^2$, *(i)* $(2,-1)$, *(ii)* (x,y)

   (b) $f(x,y) = x^2y^2 + 3x^3 - 4y^3$, *(i)* $(1,-2)$, *(ii)* (x,y)

   (c) $f(x,y) = 1 + \frac{x}{y} - \frac{y^2}{x}$, *(i)* $(-1,3)$, *(ii)* (x,y)

   (d) $2x\sin y - y^2\cos 2x$ *at (i)* $\left(\frac{\pi}{6}, \frac{\pi}{4}\right)$, *(ii)* (x,y)

   (e) $f(x,y) = \frac{3}{xy} - \frac{y^2}{x^2}$, *(i)* $(1,1)$, *(ii)* (x,y)

   (f) $f(x,y) = (2x+3y)\ln(e+xy)$, *(i)* $(0,1)$, *(ii)* (x,y)

The Derivative of a Function from a Subset of $\mathbb{R}^2$ into $\mathbb{R}^2$

A function from a subset of $\mathbb{R}^2$ into $\mathbb{R}^2$ assigns a pair of real numbers to every pair of real numbers in its domain. Examples:

(1) $f(x,y) = \left(x^2y, x+y^3\right)$ (2) $f(x,y) = \left(x+3y^2, x^2-4y^3\right)$ (3) $f(x,y) = \left(x\sin y, y^2e^{xy}\right)$

We cannot draw the graph of such a function $f : A \subseteq \mathbb{R}^2 \to \mathbb{R}^2$ in our 3-dimensional space, therefore we skip the part about a tangent plane and define it to be differentiable at a point (c,d) in its domain if $f(c+h,d+k) - f(c,d)$ can be approximated by a linear map $L : \mathbb{R}^2 \to \mathbb{R}^2$ such that the norm $||f(c+h,d+k) - f(c,d) - L(h,k)||$ of the error term is small compared to $||(h,k)||$. More precisely,

Definition 2 *A function $f : A \subseteq \mathbb{R}^2 \to \mathbb{R}^2$ is differentiable at a point (c,d) in its domain if there is a linear map, denoted by $Df(c,d)$ such that*

$$\lim_{(h,k) \to (0,0)} \frac{||f(c+h,d+k) - f(c,d) - Df(c,d)(h,k)||}{||(h,k)||} = 0$$

Example 3 *Consider the function $f(x,y) = \left(x^2y, x+y^3\right)$. Let (c,d) be a point in $\mathbb{R}^2$. Then for any $(h,k) \in \mathbb{R}^2$*

$$f(c+h,d+k) - f(c,d) = \left(c^2k + 2cdh + 2chk + h^2d + h^2k, h + 3d^2k + 3dk^2 + k^3\right)$$

The terms which are linear in h or k are $c^2k + 2cdh$ in the first component and $h + 3d^2k$ in the second component. When we separate them from the non-linear ones we get

$$f(c+h,d+k) - f(c,d) = \left(c^2k + 2cdh, h + 3d^2k\right) + \left(2chk + h^2d + h^2k, 3dh^2 + k^3\right)$$

We have to show that the norm of the error in approximating $f(c{+}h, d{+}k) - f(c,d)$ *with* $\left(c^2k + 2cdh, h + 3d^2k\right)$ *is small compared to* $||(h,k)||$. *The error is* $\left(2chk + h^2d + h^2k, 3dh^2 + k^3\right)$. *Recall the inequalities* $|h| \leq ||(h,k)||$, $|k| \leq ||(h,k)||$. *Combined with the triangle inequality, they imply that*

$$\begin{aligned}
\left|\left|\left(2chk + h^2d + h^2k, 3dk^2 + k^3\right)\right|\right| &\leq (|2c|)\,|hk| + (|d|)\,|h|^2 + |h|^2\,|k| + (|3d|)\,|k|^2 + |k|^3 \\
&\leq (|2c|)\,||(h,k)||^2 + (|d|)\,||(h,k)||^2 + ||(h,k)||^3 + (|3d|)\,||(h,k)||^2 + ||(h,k)||^3 \\
&= (|2c| + 4\,|d| + 2\,||(h,k)||)\,||(h,k)||^2
\end{aligned}$$

It follows that $\displaystyle\lim_{(h,k)\to(0,0)} \frac{\left|\left|\left(2chk + h^2d + h^2k, 3dk^2 + k^3\right)\right|\right|}{||(h,k)||} \leq \lim_{(h,k)\to(0,0)} (|2c| + 4\,|d| + 2\,||(h,k)||)\,||(h,k)|| = 0$. *Therefore* f *is differentiable at* (c,d) *and its derivative* $Df(c,d)$ *has formula*

$$Df(c,d)\,(h,k) = \left(2cdh + c^2k, h + 3d^2k\right)$$

The Mean Value Theorem may be used to show that if $f(x,y) = (f_1(x,y), f_2(x,y))$ and the first order partial derivatives of f_1 and f_2 are continuous then the derivative of f at (c,d) has formula

$$Df(c,d)\,(h,k) = \left(h\frac{\partial f_1}{\partial x} + k\frac{\partial f_1}{\partial y}, h\frac{\partial f_2}{\partial x} + k\frac{\partial f_2}{\partial y}\right)$$

All the partial derivatives are evaluated at (c,d). If we agree to write $\left(h\frac{\partial f_1}{\partial x} + k\frac{\partial f_1}{\partial y}, h\frac{\partial f_2}{\partial x} + k\frac{\partial f_2}{\partial y}\right)$ not as a row but as the 2×1 column matrix

$$\begin{pmatrix} h\dfrac{\partial f_1}{\partial x} + k\dfrac{\partial f_1}{\partial y} \\ \\ h\dfrac{\partial f_2}{\partial x} + k\dfrac{\partial f_2}{\partial y} \end{pmatrix}$$

then the expression for $Df(c,d)\,(h,k)$ may be written as a matrix product:

$$Df(c,d)\,(h,k) = \begin{pmatrix} \dfrac{\partial f_1}{\partial x} & \dfrac{\partial f_1}{\partial y} \\ \\ \dfrac{\partial f_2}{\partial x} & \dfrac{\partial f_2}{\partial y} \end{pmatrix} \begin{pmatrix} h \\ \\ k \end{pmatrix}$$

Following what we did for a function from $\mathbb{R}^2$ into $\mathbb{R}$, we call $\begin{pmatrix} \dfrac{\partial f_1}{\partial x} & \dfrac{\partial f_1}{\partial y} \\ \\ \dfrac{\partial f_2}{\partial x} & \dfrac{\partial f_2}{\partial y} \end{pmatrix}$ the Jacobian matrix of f at (c,d).

Example 4 *In Example 3,* $f_1(x,y) = x^2y$ *and* $f_2(x,y) = x + y^3$. *Therefore:*

$$\frac{\partial f_1(c,d)}{\partial x} = 2cd,\ \frac{\partial f_1(c,d)}{\partial x} = c^2,\ \frac{\partial f_2(c,d)}{\partial x} = 1 \text{ and } \frac{\partial f_2(c,d)}{\partial y} = 3d^2$$

It follows that

$$Df(c,d)\,(h,k) = \begin{pmatrix} 2cd & c^2 \\ \\ 1 & 3d^2 \end{pmatrix} \begin{pmatrix} h \\ \\ k \end{pmatrix}$$

The Jacobian matrix of f *at* (c,d) *is*

$$\begin{pmatrix} 2cd & c^2 \\ \\ 1 & 3d^2 \end{pmatrix}$$

It follows that

$$Df(c,d)\,(h,k) = \begin{pmatrix} 2cd & c^2 \\ 1 & 3d^2 \end{pmatrix} \begin{pmatrix} h \\ k \end{pmatrix}$$

The Jacobian matrix of f at (c,d) is

$$\begin{pmatrix} 2cd & c^2 \\ 1 & 3d^2 \end{pmatrix}$$

Example 5 *Let $f(x,y) = (f_1(x,y), f_2(x,y)) = \left(x \sin y, y^2 e^{xy}\right)$. The partial derivatives of f_1 and f_2 are $\dfrac{\partial f_1}{\partial x} = \sin y$, $\dfrac{\partial f_1}{\partial y} = x\cos y$, $\dfrac{\partial f_2}{\partial x} = y^3 e^{xy}$ and $\dfrac{\partial f_2}{\partial y} = \left(2y + y^2 x\right) e^{xy}$. They are continuous on $\mathbb{R}^2$. Let (c,d) be a point in $\mathbb{R}^2$. Then $\dfrac{\partial f_1(c,d)}{\partial x} = \sin d$, $\dfrac{\partial f_1(c,d)}{\partial y} = c\cos d$, $\dfrac{\partial f_2(c,d)}{\partial x} = d^3 e^{cd}$ and $\dfrac{\partial f_2(c,d)}{\partial y} = \left(2d + d^2 c\right) e^{cd}$. Therefore the derivative of f at (c,d) is given by*

$$\begin{aligned} Df(c,d)\,(h,k) &= \begin{pmatrix} \sin d & c\cos d \\ d^3 e^{cd} & \left(2d + d^2 c\right) e^{cd} \end{pmatrix} \begin{pmatrix} h \\ k \end{pmatrix} \\ &= \begin{pmatrix} h \sin d + ck \cos d \\ hd^3 e^{cd} + k\left(2d + d^2 c\right) e^{cd} \end{pmatrix} \end{aligned}$$

In general, if (x,y) is any point in $\mathbb{R}^2$, then the derivative of f at (x,y) is given by

$$\begin{aligned} Df(x,y)\,(h,k) &= \begin{pmatrix} \sin y & x\cos y \\ y^3 e^{xy} & \left(2y + y^2 x\right) e^{xy} \end{pmatrix} \begin{pmatrix} h \\ k \end{pmatrix} \\ &= \begin{pmatrix} h \sin y + xk \cos y \\ hy^3 e^{xy} + k\left(2y + y^2 x\right) e^{xy} \end{pmatrix} \end{aligned}$$

We may also write this as

$$Df(x,y)\,(h,k) = \left(h \sin y + xk\cos y, hy^3 e^{xy} + k\left(2y + y^2 x\right) e^{xy}\right)$$

The Derivative of a Function from a Subset of $\mathbb{R}^p$ into $\mathbb{R}^q$

This is a straight-forward generalization of what we have observed so far. To simplify notation, we denote a point $(c_1, \ldots, c_p)$ in $\mathbb{R}^p$ by c. By the same token, $(h_1, \ldots h_p)$ may be written as h and $(c_1 + h_1, \ldots, c_p + h_p)$ as $c + h$.

Let $f : A \subseteq \mathbb{R}^p \to \mathbb{R}^q$ be a given function and c be a point in its domain. We say that f is differentiable at c if $f(c+h)-f(c)$ can be approximated by a linear function $L(h)$ such that the error term $f(c+h)-f(c)-L(h)$ is small compared to $||h||$. Note that $f(x_1, \ldots x_p) = (f_1(x_1, \ldots x_p), \ldots, f_q(x_1, \ldots x_p))$. Thus f has q components and each component is a function of the p variables $x_1, \ldots x_p$. It can be shown that if the components $f_1, \ldots, f_q$ have continuous partial derivatives $\dfrac{\partial f_i}{\partial x_j}$, $i = 1, \ldots, q$ and $j = 1, \ldots, p$ at c then f is differentiable at c and its derivative is given by

$$Df(c)\,(h) = \begin{pmatrix} \dfrac{\partial f_1(c)}{\partial x_1} & \cdots & \dfrac{\partial f_1(c)}{\partial x_p} \\ \vdots & \vdots & \vdots \\ \dfrac{\partial f_q(c,d)}{\partial x_1} & \cdots & \dfrac{\partial f_q(c,d)}{\partial x_p} \end{pmatrix} \begin{pmatrix} h_1 \\ \vdots \\ k \end{pmatrix}$$

The matrix M below is called the Jacobian matrix for f at c.

$$M = \begin{pmatrix} \dfrac{\partial f_1(c)}{\partial x_1} & \cdots & \dfrac{\partial f_1(c)}{\partial x_p} \\ \vdots & \vdots & \vdots \\ \dfrac{\partial f_q(c,d)}{\partial x_1} & \cdots & \dfrac{\partial f_q(c,d)}{\partial x_p} \end{pmatrix}$$

Two properties of this matrix you should not miss are:

1. It is a $q \times p$ matrix whereas f is a function from $\mathbb{R}^p$ into $\mathbb{R}^q$, (in other words, do not miss the reversal).
2. The first row consists of the partial derivatives of f_1, (the first component of f), with respect to the p variables $x_1, \ldots, x_p$, the second row consists of the partial derivatives of f_2, (the second component of f), with respect to the p variables $x_1, \ldots, x_p$, and so on.

Exercise 6 *Write down $Df(x)h$ as a matrix product for each given function f. In each case, x is an arbitrary point in the domain of f.*

1. $f(x_1, x_2, x_3) = (x_1x_2, x_2x_3)$, $h = (h_1, h_2, h_3)$.
2. $f(x_1, x_2) = (x_1 + x_2, x_1x_2, x_1^2 + x_2^3)$, $h = (h_1, h_2)$.
3. $f(x) = (x^2, 3e^x, x)$, $h = h$
4. $f(x_1, x_2, x_3) = x_2 \sin x_1 - x_2x_3e^{2x_1}$, $h = (h_1, h_2, h_3)$.