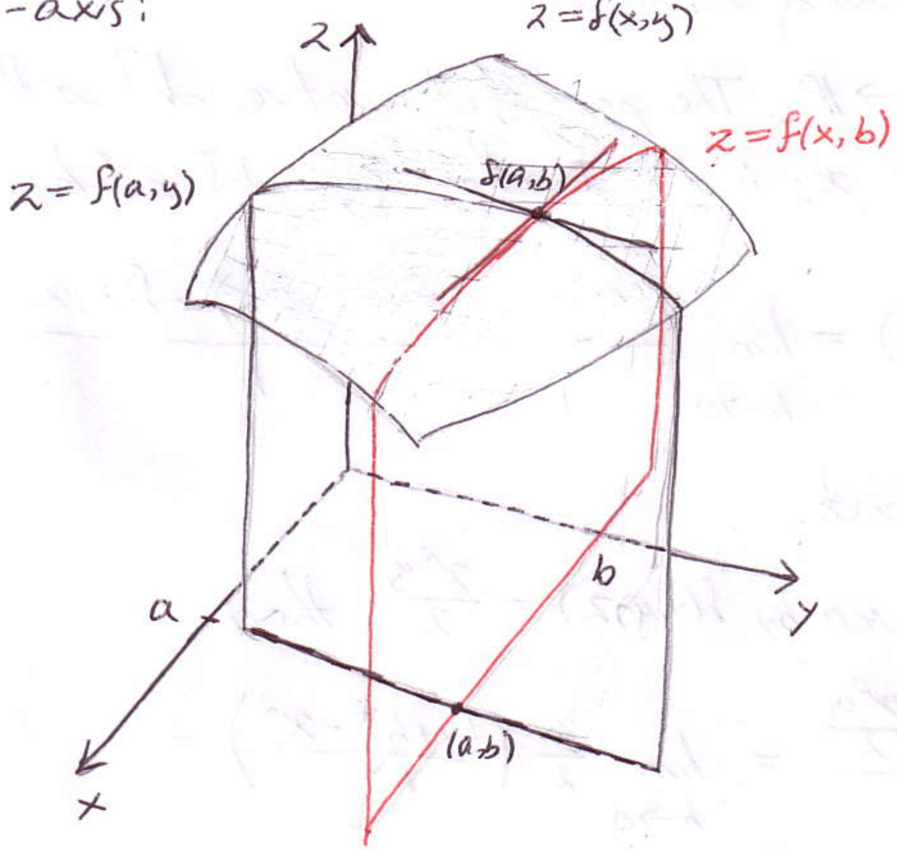
(2.3)

Partial derivatives

Let $z = f(x,y)$ be a function of two variables. Then its graph may be interpreted geometrically as a landscape. If you stand in position $(a, b, f(a,b))$, how steeply would you have to climb if you move parallel to the x-axis? How about the y-axis?



In the figure above, the red curve outlines a function of one variable $z = g(x) = f(x, b)$. Similarly, the black curve outlines another function of one variable $z = p(y) = f(a, y)$. The steepness of climb in the direction of the x-axis is just the slope of the tangent to the black curve, $p'(y)$, for $y=b$. Similarly, the steepness of climb in the direction of the x-axis is just the slope of the tangent line to the red curve, $g'(x)$, for $x=a$.

Observe that $g'(a) = \lim_{h \to 0} \frac{g(a+h)-g(a)}{h} = \lim_{h \to 0} \frac{f(a+h,b)-f(a,b)}{h}$

and $p'(b) = \lim_{h \to 0} \frac{p(b+h)-p(b)}{h} = \lim_{h \to 0} \frac{f(a,b+h)-f(a,b)}{h}$

This motivates the following definition:

Def: Let $f: U \subset \mathbb{R}^n \to \mathbb{R}$. The partial derivative of f with respect to the variable x_i is the function $\frac{\partial f}{\partial x_i}$ defined by

$$\frac{\partial f}{\partial x_i}(x_1, x_2, \ldots, x_n) = \lim_{h \to 0} \frac{f(x_1, x_2, \ldots, x_i+h, \ldots, x_n) - f(x_1, x_2, \ldots, x_i, \ldots, x_n)}{h}$$

whenever this limit exists.

Ex. Let $f: \mathbb{R}^3 \to \mathbb{R}$ given by $f(x,y,z) = \frac{x^2 y}{z}$, then

$$\frac{\partial f}{\partial x} = \lim_{h \to 0} \frac{\frac{(x+h)^2 y}{z} - \frac{x^2 y}{z}}{h} = \lim_{h \to 0} \frac{y}{z}\left(\frac{(x+h)^2 - x^2}{h}\right) =$$

$$= \frac{y}{z} \lim_{h \to 0} \frac{x^2+2xh+h^2-x^2}{h} = \frac{y}{z} \lim_{h \to 0} (2xh+h^2)\frac{1}{h} = \frac{y}{z} \lim_{h \to 0} (2x+h)$$

$$= \frac{y}{z} \cdot 2x = \frac{2xy}{z}$$

Thus $\frac{\partial f}{\partial x}(x,y,z) = \frac{2xy}{z}$

$$\frac{\partial f}{\partial y} = \lim_{h\to 0}\left(\frac{x^2(y+h)}{z} - \frac{x^2y}{z}\right)\frac{1}{h} = \lim_{h\to 0} \frac{x^2}{z}\left(\frac{y+h-y}{h}\right) =$$

$$= \frac{x^2}{z}$$

Similarly

$\frac{\partial f}{\partial z} = -\frac{x^2y}{z^2}$ as you should verify.

Observe that in computing the partial derivative of f with respect to x_i, we treat all other variables as if they were constant. Computing partial derivatives is therefore very similar to the procedures studied in Calc I.

Ex. Calculate the partial derivatives of $f(x,y) = x\sin xy + ye^{-x^2}$

Solution: $\frac{\partial f}{\partial x}(x,y) = \sin xy + x(\cos xy)\cdot y + y(-2xe^{-x^2}) =$

$= \sin(xy) + xy\cos(xy) - 2xye^{-x^2}$.

$$\frac{\partial f}{\partial y}(x,y) = x(\cos(xy)\,x) + (e^{-x^2}) = x^2\cos(xy) + e^{-x^2}.$$

It is frequently cumbersome to write the partial derivative of $f: U \subset \mathbb{R}^n \to \mathbb{R}$ with respect to x_i as $\frac{\partial f}{\partial x_i}(x_1, x_2, \ldots, x_n)$. To save time (and space on paper), this derivative will often be written simply as $\frac{\partial f}{\partial x_i}$ with the variables suppressed. $\frac{\partial f}{\partial x_i}(x_1, x_2, \ldots, x_n)$ is also written as f_{x_i}, although we'll use this notation far less frequently.

Higher-order partial derivatives are analogous to higher-order ordinary derivatives. For instance, the second-order partial derivatives of f are defined by

$$\frac{\partial^2 f}{\partial x_j \partial x_i} = \frac{\partial}{\partial x_j}\left(\frac{\partial f}{\partial x_i}\right), \quad 1 \le i, j \le n.$$

When $i \neq j$, we call these mixed partials. With the subscript notation, we write $f_{x_i x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i}$

Notice that with the subscript notation, the partials are taken in the order, left to right, in which the subscripts appear; with the "∂" notation, the partials are taken in right-to-left order. Also, if $i = j$ we write $\frac{\partial^2 f}{\partial x_i^2}$ for the second partial.

Ex. Let $f(x,y) = 3x^2 + 4xy - 7y^2$

Then $\frac{\partial f}{\partial x} = 6x + 4y$ and $\frac{\partial f}{\partial y} = 4x - 14y$.

$$\frac{\partial^2 f}{\partial x^2} = 6 \qquad \frac{\partial^2 f}{\partial y \partial x} = 4$$

$$\frac{\partial^2 f}{\partial x \partial y} = 4 \qquad \frac{\partial^2 f}{\partial^2 y} = -14.$$

Partial derivatives of order higher than two are defined as you would expect. For instance, the third-order partials are

$$\frac{\partial}{\partial x_k}\left(\frac{\partial^2 f}{\partial x_j \partial x_i}\right), \quad 1 \le i, j, k \le n$$

How many different third-order partials are there?

In the previous example, you may have noticed that $\frac{\partial^2 f}{\partial y \partial x} = \frac{\partial^2 f}{\partial x \partial y}$ This is no coincidence. For almost all functions you encounter throughout this course, the mixed partials are equal at almost every point. The general principle at work here is expressed in the following theorem, which plays an important role in the study of vector calculus.

Thm: (Clairaut's theorem). Suppose that the mixed partials of $f: U \subset \mathbb{R}^2 \to \mathbb{R}$ are defined in an open disk B about a point $a \in U$. If $\frac{\partial^2 f}{\partial y \partial x}$ and $\frac{\partial^2 f}{\partial x \partial y}$ are continuous at a, then

$$\frac{\partial^2 f}{\partial y \partial x}(a) = \frac{\partial^2 f}{\partial x \partial y}(a)$$

Proof: (Optional) The proof presented below relies on the stronger hypothesis that the second-order partials are continuous in some open disk $B_r(a) \subset B$. Let $\vec{a} = (a, b)$.

Observe that $\frac{\partial^2 f}{\partial x \partial y}(a,b) = \lim_{h \to 0} \frac{\frac{\partial f}{\partial y}(a+h, b) - \frac{\partial f}{\partial y}(a,b)}{h} =$

$$= \lim_{h\to 0} \frac{\lim\limits_{k\to 0} \frac{1}{k}\left(f(a+h,b+k) - f(a+h,b)\right) - \frac{1}{k}\left(f(a,b+k) - f(a,b)\right)}{h} =$$

$$= \lim_{h\to 0} \lim_{k\to 0} \frac{f(a+h,b+k) - f(a,b+k) - f(a+h,b) + f(a,b)}{kh} \qquad (1)$$

by the mean-value theorem, there exist $a^*, \bar{a} \in (a, a+h)$

s.t. $\dfrac{f(a+h,b+k) - f(a,b+k)}{h} = \dfrac{\partial f}{\partial x}(a^*, b+k)$ and

$$\frac{f(a+h,b) - f(a,b)}{h} = \frac{\partial f}{\partial x}(\bar{a}, b)$$

With this, (1) becomes

$$\lim_{h\to 0} \lim_{k\to 0} \frac{\frac{\partial f}{\partial x}(a^*, b+k) - \frac{\partial f}{\partial x}(\bar{a}, b)}{k} =$$

$$\lim_{h\to 0} \lim_{k\to 0} \frac{\frac{\partial f}{\partial x}(a^*, b+k) - \frac{\partial f}{\partial x}(a^*, b) + \frac{\partial f}{\partial x}(a^*, b) - \frac{\partial f}{\partial x}(\bar{a}, b)}{k} =$$

$$= \lim_{(h,k)\to(0,0)} \frac{\frac{\partial f}{\partial x}(a^*, b+k) - \frac{\partial f}{\partial x}(a^*, b)}{k} + \frac{\frac{\partial f}{\partial x}(a^*, b) - \frac{\partial f}{\partial x}(\bar{a}, b)}{k} =$$

$$= \lim_{(h,k)\to(0,0)} \frac{\partial^2 f}{\partial y \partial x}(a^*, b^*) + \frac{\frac{\partial f}{\partial x}(a^*, b) - \frac{\partial f}{\partial x}(\bar{a}, b)}{k} \qquad (2)$$

where $b^* \in (b, b+k)$ is a number guaranteed to exist by the mean-value theorem.

We know that $\lim\limits_{(h,k)\to(0,0)} \frac{\partial^2 f}{\partial y \partial x}(a^*, b^*) = \frac{\partial^2 f}{\partial y \partial x}(a,b)$ because,

as h and k go to 0, $a^* \in (a, a+h)$ and $b^* \in (b, b+k)$ are pushed to a and b respectively and because the second-order partials are continuous at (a,b).

We also know that (2) exists and is equal to $\frac{\partial^2 f}{\partial x \partial y}(a,b)$ (that's what we started from!) Hence, it follows that

$$\lim_{(h,k)\to(0,0)} \frac{\frac{\partial f}{\partial x}(a^*, b) - \frac{\partial f}{\partial x}(\bar{a}, b)}{k}$$ exists as well

(it is equal to a difference of (2) and $\lim\limits_{(h,k)\to(0,0)} \frac{\partial^2 f}{\partial y \partial x}(a^*, b^*)$.

Since $(2) - \lim\limits_{(h,k)\to(0,0)} \frac{\partial^2 f}{\partial y \partial x}(a^*, b^*) = \lim\limits_{(h,k)\to(0,0)} \frac{\partial^2 f}{\partial y \partial x}(a^*, b^*) + \frac{\frac{\partial f}{\partial x}(a^*, b) - \frac{\partial f}{\partial x}(\bar{a}, b)}{k}$

$- \frac{\partial^2 f}{\partial y \partial x}(a^*, b^*) = \lim\limits_{(h,k)\to(0,0)} \frac{\frac{\partial f}{\partial x}(a^*, b) - \frac{\partial f}{\partial x}(\bar{a}, b)}{k}$, we know that this limit exists)

The last limit can be written as $\lim\limits_{k\to 0} \lim\limits_{h\to 0} \frac{\frac{\partial f}{\partial x}(a^*, b) - \frac{\partial f}{\partial x}(\bar{a}, b)}{k} =$

$$= \lim_{k\to 0} \frac{0}{k} = 0$$

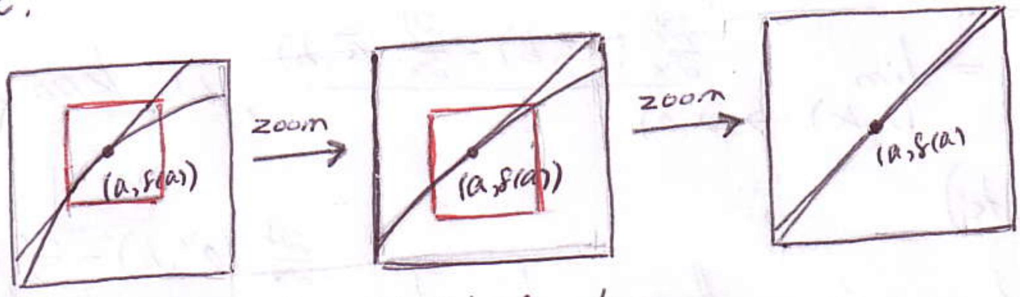
Hence $(2) = \frac{\partial^2 f}{\partial y \partial x}(a,b)$ (Why?)

Differentiability and the total derivative

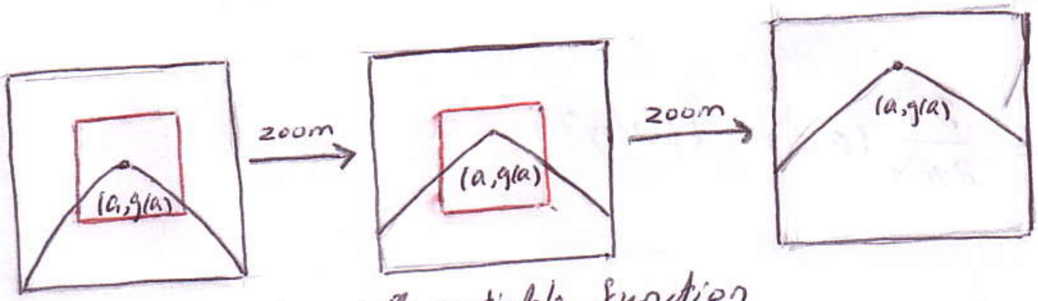
Recall from Calc I that the function $f: \mathcal{U} \subset \mathbb{R} \to \mathbb{R}$ is said to be differentiable at $x = a \in \mathcal{U}$ if $\lim_{x \to a} \frac{f(x) - f(a)}{x - a}$ exists. If we wish to generalize the derivative to functions of several variables, we'll need to modify the definition of the derivative, since, if $f: \mathcal{U} \subset \mathbb{R}^2 \to \mathbb{R}$ for instance,

$$\lim_{(x,y)\to(a,b)} \frac{f(x,y) - f(a,b)}{(x,y) - (a,b)}$$ does not make any sense (Why?).

In Calc I, it might have been mentioned in passing that $f: \mathcal{U} \subset \mathbb{R} \to \mathbb{R}$ is differentiable if the tangent line to the curve f at a point of differentiability $x = a$ is a good approximation to the curve.

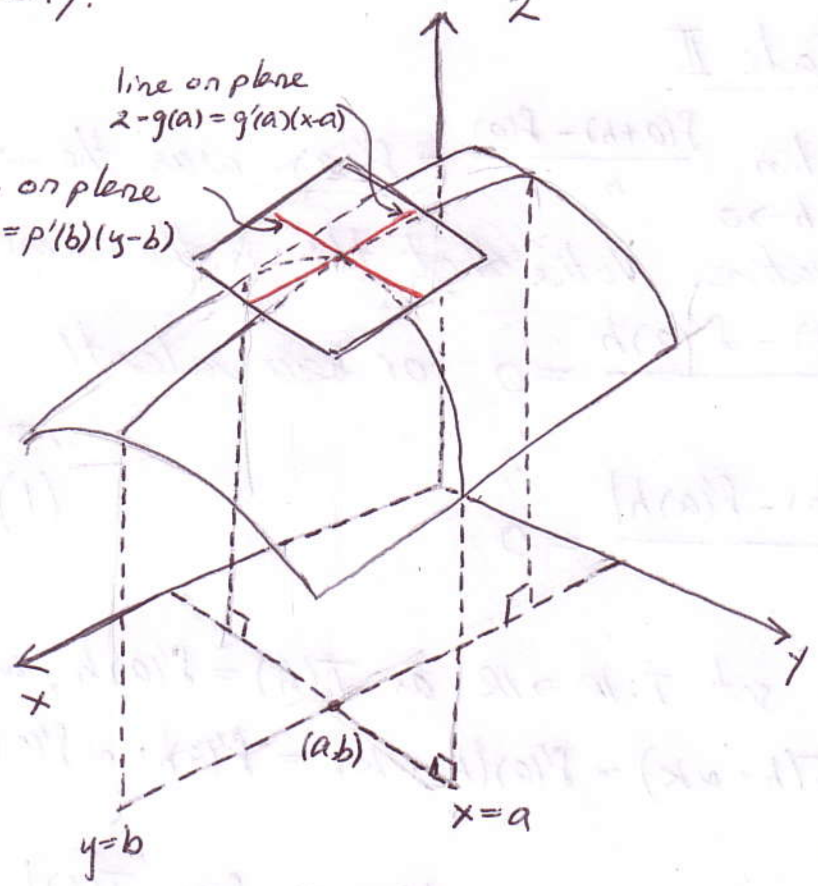


Differentiable function "looks like" its tangent-line under a large zoom.



Non-differentiable function does not "look like" a line under a large zoom.

The equation of the tangent line to $(a, f(a))$ is $y - f(a) = f'(a)(x-a)$. Since the analogue of a line in $\mathbb{R}^2$ is a plane in $\mathbb{R}^3$, we might suspect that if $f: \mathbb{R}^2 \to \mathbb{R}$ is differentiable at $(x,y) = (a,b)$, the derivative would be a plane of the form $z - f(a,b) = A(x-a) + B(y-b)$ where the coefficients A and B are to be determined.
Notice that this tangent plane contains the tangents to the curves $z = f(x,b)$ and $z = f(a,y)$ at the points $x = a$ and $y = b$ respectively.



Let $g(x) = f(x,b)$ and $p(y) = f(a,y)$. Observe that $g'(x) = \frac{\partial f}{\partial x}(x,b)$ and $p'(y) = \frac{\partial f}{\partial y}(a,y)$. Hence $g'(a) = \frac{\partial f}{\partial x}(a,b)$ and $p'(b) = \frac{\partial f}{\partial y}(a,b)$. Notice now that $g(a) = p(b) = f(a,b)$

It follows that $z - g(a) = g'(a)(x-a) \Rightarrow$

$\Rightarrow z - f(a,b) = \frac{\partial f}{\partial x}(a,b)(x-a) = \frac{\partial f}{\partial x}(a,b)(x-a) + B(b-b)$

Also $z - p(b) = p'(b)(y-b) \Rightarrow z - f(a,b) = A(a-a) + \frac{\partial f}{\partial y}(y-b)$.

Putting this together, we get $z - f(a,b) = \frac{\partial f}{\partial x}(a,b)(x-a) + \frac{\partial f}{\partial y}(a,b)(y-b)$

In other words, if $f: U \subset \mathbb{R}^2 \to \mathbb{R}$ "looks like" a plane (is differentiable) near the point $(a,b,f(a,b))$ then its approximating plane has the equation $z - f(a,b) = \frac{\partial f}{\partial x}(a,b)(x-a) + \frac{\partial f}{\partial y}(a,b)(y-b)$.

Derivatives in Calc. III

Observe that $\lim_{h \to 0} \frac{f(a+h) - f(a)}{h} = f'(a)$ was the way we defined the derivative. Notice that this is the same as

$$\lim_{h \to 0} \frac{f(a+h) - f(a) - f'(a)h}{h} = 0 \quad \text{or equivalently}$$

$$\lim_{h \to 0} \frac{|f(a+h) - f(a) - f'(a)h|}{|h|} = 0 \qquad (1)$$

Notice that if we set $T: \mathbb{R} \to \mathbb{R}$ as $T(h) = f'(a)h$, we get a linear map. That is $T(h + \alpha k) = f'(a)(h + \alpha k) = f'(a)h + \alpha f'(a)k =$
$= T(h) + \alpha T(k)$

Thus we can rewrite (1) as $\lim_{h \to 0} \frac{|f(a+h) - f(a) - T(h)|}{|h|} = 0 \qquad (2)$

This formulation states that $f: U \subset \mathbb{R} \to \mathbb{R}$ is differentiable at a if there exists a linear transformation that is a good approximation to the difference $f(a+h) - f(a)$ in the sense of (2).

This motivates the following definition of the derivative of a function $f: U \subset \mathbb{R}^n \to \mathbb{R}^m$.

Def: A function $f: U \subset \mathbb{R}^n \to \mathbb{R}^m$ is differentiable at $a \in U$ if there exists a linear transformation $T: \mathbb{R}^n \to \mathbb{R}^m$ such that

$$\lim_{h \to 0} \frac{\|f(a+h) - f(a) - T(h)\|}{\|h\|} = 0 \text{ where } h \in \mathbb{R}^n. \qquad (3)$$

This limit can also be written in the form

$$\lim_{x \to a} \frac{\|f(x) - f(a) - T(x-a)\|}{\|x-a\|} = 0 \qquad (4)$$

We call T the derivative (or total derivative) of f at a.

It is interesting to observe that at most one linear map can satisfy (3). In other words, if f is differentiable, its derivative is unique.

Proposition: Suppose $f: U \subset \mathbb{R}^n \to \mathbb{R}^m$ satisfies

$$\lim_{h \to 0} \frac{\|f(a+h) - f(a) - T(h)\|}{\|h\|} = \lim_{h \to 0} \frac{\|f(a+h) - f(a) - S(h)\|}{\|h\|} = 0$$

For some linear functions $T, S: \mathbb{R}^n \to \mathbb{R}^m$ and some $a \in \mathbb{R}^n$.

Then $T = S$.

Proof: (Optional)

We would like to show that for any $x \in \mathbb{R}^n$, $T(x) = S(x)$.

If $x = 0$ $\quad T(0) = T(0+0) = T(0) + T(0)$. Hence $T(0) = 0$

Similarly $S(0) = 0$.

Suppose now that $x \neq 0$. To show that $T(x) = S(x)$, it suffices to prove that $\|T(x) - S(x)\| = 0$ or, equivalently, that $\frac{\|T(x)-S(x)\|}{\|x\|} = 0$

Observe that for $t \in \mathbb{R}$ $\frac{\|T(tx) - S(tx)\|}{\|tx\|} = \frac{\|tT(x) - tS(x)\|}{\|tx\|} =$

$$= \frac{|t| \, \|T(x) - S(x)\|}{|t| \, \|x\|} = \frac{\|T(x) - S(x)\|}{\|x\|}$$

Thus $\frac{\|T(x)-S(x)\|}{\|x\|} = \lim_{t \to 0} \frac{\|T(tx) - S(tx)\|}{\|tx\|} = \lim_{h \to 0} \frac{\|T(h) - S(h)\|}{\|h\|}$

where $h = tx \in \mathbb{R}^n$

But $\lim_{h \to 0} \frac{\|T(h) - S(h)\|}{\|h\|} = \lim_{h \to 0} \frac{\|f(a+h) - f(a) - S(h) - f(a+h) + f(a) + T(h)\|}{\|h\|}$

$$\leq \lim_{h \to 0} \left(\frac{\|f(a+h) - f(a) - S(h)\|}{\|h\|} + \frac{\|f(a+h) - f(a) - T(h)\|}{\|h\|} \right) = 0$$

Thus $0 \leq \frac{\|T(x) - S(x)\|}{\|x\|} \leq 0$ which implies that

$\|T(x) - S(x)\| = 0$ so that $T(x) = S(x)$ as desired.

Since the derivative is unique, the often used notation for it is $Df(a)$.
In other words, the derivative of f at a is the linear map $Df(a): \mathbb{R}^n \to \mathbb{R}^m$

Remark: $Df(a)$ is <u>not</u> the function Df evaluated at a.
$Df(a)$ stands for T in expression (3) above.

The total derivative of a Scalar-Valued Function

Suppose $f: U \subset \mathbb{R}^n \to \mathbb{R}$. We would like to know how to compute $Df(a): \mathbb{R}^n \to \mathbb{R}$. By what we know about linear maps $Df(a)(x) = Ax^T$ for some $n \times 1$ matrix A. How might we find this matrix?

Thm: If $f: U \subset \mathbb{R}^n \to \mathbb{R}$ has partial derivatives that are defined on an open ball about a point $a \in U$ and if $\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \ldots, \frac{\partial f}{\partial x_n}$ are continuous at a, then f is differentiable at a and its derivative is given by

$$Df(a)(x) = \left(\frac{\partial f}{\partial x_1}(a) \; \frac{\partial f}{\partial x_2}(a) \; \ldots \; \frac{\partial f}{\partial x_n}(a) \right) x^T.$$

Proof: (By Thomas K. Barr) For the sake of clarity, we prove the theorem for $n=2$. Let $\vec{a} = (a,b)$ and $\vec{x} = (x,y)$ be the variables for f.

Observe that $f(\vec{x}) - f(\vec{a}) = f(x,y) - f(a,b) = \big(f(x,y) - f(a,y)\big) + \big(f(a,y) - f(a,b)\big)$.

For the moment, we regard x and y as fixed and consider the function g of one variable given by

$$g(t) = f(t,y)$$

where t varies over the interval between a and x. Since $\frac{\partial f}{\partial x}$ is defined on the interval, g is continuous and differentiable, so by the mean value theorem there exists a number $\bar{x}$ between x and a (which depends on the choice of y) such that

$$g(x)-g(a)=g'(\bar{x})(x-a).$$

In terms of f, this says that $f(x,y)-f(a,y)=\frac{\partial f}{\partial x}(\bar{x},y)(x-a)$.
Applying the same reasoning to the function $h(t)=f(a,t)$ for t varying over the interval between b and y, leads to a similar conclusion: There exists $\bar{y}$ between b and y such that

$$f(a,y)-f(a,b)=\frac{\partial f}{\partial y}(a,\bar{y})(y-b).$$

Let $Jf(a)=\left(\frac{\partial f}{\partial x}(a,b) \quad \frac{\partial f}{\partial y}(a,b)\right)$. Observe that

$$f(\vec{x})-f(\vec{a})-Jf(\vec{a})(\vec{x}-\vec{a})^T=f(x,y)-f(a,b)-Jf(a,b)(x-a,y-b)^T=$$

$$=\big(f(x,y)-f(a,y)\big)+\big(f(a,y)-f(a,b)\big)-Jf(a,b)(x-a,y-b)^T=$$

$$=\frac{\partial f}{\partial x}(\bar{x},y)(x-a)+\frac{\partial f}{\partial y}(a,\bar{y})(y-b)-Jf(a,b)(x-a,y-b)^T=$$

$$=\frac{\partial f}{\partial x}(\bar{x},y)(x-a)+\frac{\partial f}{\partial y}(a,\bar{y})(y-b)-\frac{\partial f}{\partial x}(a,b)(x-a)-\frac{\partial f}{\partial y}(a,b)(y-b)=$$

$$=\left(\frac{\partial f}{\partial x}(\bar{x},y)-\frac{\partial f}{\partial x}(a,b)\right)(x-a)+\left(\frac{\partial f}{\partial y}(a,\bar{y})-\frac{\partial f}{\partial y}(a,b)\right)(y-b)=$$

$$=\left(\frac{\partial f}{\partial x}(\bar{x},y)-\frac{\partial f}{\partial x}(a,b),\ \frac{\partial f}{\partial y}(a,\bar{y})-\frac{\partial f}{\partial y}(a,b)\right)\cdot(x-a,y-b)$$

By the Cauchy-Schwarz inequality, the absolute value of this dot product is no more than the products of the magnitudes of the vectors. So we see that

$$\frac{|f(\vec{x})-f(\vec{a})-Jf(\vec{a})(\vec{x}-\vec{a})^T|}{\|\vec{x}-\vec{a}\|}\leq\frac{\left\|\left(\frac{\partial f}{\partial x}(\bar{x},y)-\frac{\partial f}{\partial x}(a,b),\ \frac{\partial f}{\partial y}(a,\bar{y})-\frac{\partial f}{\partial y}(a,b)\right)\right\|\|\vec{x}-}{\|\vec{x}-\vec{a}\|}$$

$$= \left\| \left(\frac{\partial f}{\partial x}(\bar{x},y) - \frac{\partial f}{\partial x}(a,b),\ \frac{\partial f}{\partial y}(a,\bar{y}) - \frac{\partial f}{\partial y}(a,b) \right) \right\| \qquad (1)$$

As $\vec{x} \to \vec{a}$, we must have $(\bar{x}, y) \to \vec{a}$ and $(a, \bar{y}) \to \vec{a}$. Since $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$ are continuous at $\vec{a}$, the right-hand side of (1) goes to zero as $\vec{x} \to \vec{a}$.

Thus

$$\lim_{\vec{x} \to \vec{a}} \frac{\| f(\vec{x}) - f(\vec{a}) - Jf(\vec{a})(x-a)^T \|}{\| \vec{x} - \vec{a} \|} = 0$$

Hence $Df(\vec{a})(\vec{x}) = Jf(\vec{a})\, x^T$

The matrix $Jf(\vec{a})$ is called the Jacobian matrix of f at $\vec{a}$.

Remark: In Calc. I you were taught that the derivative of $f: U \subset \mathbb{R} \to \mathbb{R}$ at a is $\frac{df}{dx}(a)$ or the slope of the tangent line. However, from the perspective of Calc III, $\frac{df}{dx}(a) = \frac{\partial f}{\partial x}(a)$ is the Jacobian matrix $Jf(a) = \left(\frac{\partial f}{\partial x}(a) \right)$. The derivative $Df(a)(x) = \frac{\partial f}{\partial x}(a)\, x$ is the line through the origin with slope $\frac{\partial f}{\partial x}(a)$.

In general, if $f: U \subset \mathbb{R}^n \to \mathbb{R}$, $Jf(\vec{a})$ is analogous to slope of a line, while the derivative is analogous to the line through the origin with "slope" $Jf(\vec{a})$.

Ex. Calculate the Jacobian matrix of $f(x,y) = 2x^3 + 5xy^2 + y^3$ at the point $\vec{a} = (-3, 1)$

Solution: $\frac{\partial f}{\partial x}(-3,1) = (6x^2 + 5y^2)\Big|_{(-3,1)} = 59$

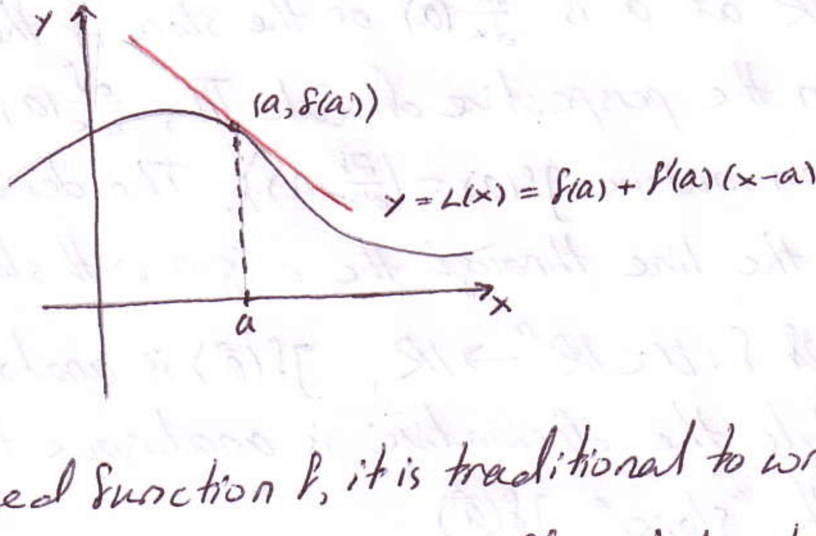
$$\frac{\partial f}{\partial y}(-3,1) = (10xy + 3y^2)\Big|_{(-3,1)} = -27$$

Thus $Jf(-3,1) = (59 \;\; -27)$ and $Df(-3,1)(x,y) =$
$= (59 \;\; -27)\begin{pmatrix} x \\ y \end{pmatrix} = 59x - 27y.$

The linear approximation of a scalar-valued function f at a point $\vec{a}$ is the scalar-valued function

$$L(\vec{x}) = f(\vec{a}) + Jf(\vec{a})(\vec{x}-\vec{a})^T.$$

When f has one variable, the graph of L is the line tangent to the graph of f at the point $(a, f(a))$. When f has two variables, the graph of L is the plane tangent to the graph of f at the point $(a_1, a_2, f(a_1,a_2))$, where $\vec{a} = (a_1, a_2)$



For a real-valued function f, it is traditional to write the total derivative in an alternative way using differential notation:

$$df = \frac{\partial f}{\partial x_1}dx_1 + \frac{\partial f}{\partial x_2}dx_2 + \ldots + \frac{\partial f}{\partial x_n}dx_n$$

In this notation, df is called the total differential of f. It is understood that if the partials are all evaluated at $\vec{a} = (a_1, a_2, \ldots, a_n)$, then the variables $dx_1, dx_2, \dots, dx_n$ stand, respectively, for x_1-a_1, $x_2-a_2, \dots,$ x_n-a_n.

Ex. The total differential of $f(x,y) = xe^{-y}+ye^{2x}$ is

$$df = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy = (e^{-y}+2ye^{2x})dx + (-xe^{-y}+e^{2x})dy$$

<u>The total derivative of a vector-valued function.</u>

Let $f: U \subset \mathbb{R}^n \to \mathbb{R}^m$ be a vector-valued function. Then f is differentiable at $\vec{a} \in U$ if there is a linear map $Df(\vec{a}): \mathbb{R}^n \to \mathbb{R}^m$ that satisfies

$$\lim_{\vec{x}\to\vec{a}} \frac{\|f(\vec{x}) - f(\vec{a}) - Df(\vec{a})(\vec{x}-a)\|}{\|\vec{x}-\vec{a}\|} = 0.$$

By what we know about linear maps, $Df(\vec{a})(\vec{x}) = M(Df(\vec{a}))\vec{x}^T$ $= Jf(\vec{a})\vec{x}^T$ where $Jf(\vec{a})$ is an $m\times n$ matrix.

Thm: let $f: U \subset \mathbb{R}^n \to \mathbb{R}^m$ be given by the formula

$$f(\vec{x}) = (f_1(\vec{x}), f_2(\vec{x}), \dots, f_m(\vec{x}))$$

If $\frac{\partial f_i}{\partial x_j}$ is continuous at $\vec{a}$ for $i=1,\dots,m$ and $j=1,\dots,n$, then f is differentiable at $\vec{a}$ and the derivative of f at $\vec{a}$ is the linear transformation $Df(\vec{a}): \mathbb{R}^n \to \mathbb{R}^m$ given by

$$Df(\vec{a})(\vec{x}) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(\vec{a}) & \frac{\partial f_1}{\partial x_2}(\vec{a}) & \dots & \frac{\partial f_1}{\partial x_n}(\vec{a}) \\ \frac{\partial f_2}{\partial x_1}(\vec{a}) & \frac{\partial f_2}{\partial x_2}(\vec{a}) & \dots & \frac{\partial f_2}{\partial x_n}(\vec{a}) \\ \vdots & & & \\ \frac{\partial f_m}{\partial x_1}(\vec{a}) & \frac{\partial f_m}{\partial x_2}(\vec{a}) & \dots & \frac{\partial f_m}{\partial x_n}(\vec{a}) \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

Proof: It suffices to show that $Df(a) = (Df_1(a), Df_2(a), \ldots, Df_m(a))$ where $Df_i(a)$ is the total derivative of the i^{th} component function f_i of f.

Observe that, because each $\frac{\partial f_i}{\partial x_i}(a)$ exists and is continuous in an open ball about a, it follows that $f_i : U \subset \mathbb{R}^n \to \mathbb{R}$ is differentiable with derivative $Df_i(a)(x) = \left(\frac{\partial f_i}{\partial x_1}(a), \frac{\partial f_i}{\partial x_2}(a), \ldots, \frac{\partial f_i}{\partial x_n}(a)\right) x^T =$

$$= \frac{\partial f_i}{\partial x_1}(a)x_1 + \frac{\partial f_i}{\partial x_2}(a)x_2 + \ldots + \frac{\partial f_i}{\partial x_n}(a)x_n.$$

Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be defined by $T(x) = \left(\sum_{j=1}^{n} \frac{\partial f_1}{\partial x_j}(a)x_j, \ldots, \sum_{j=1}^{n} \frac{\partial f_m}{\partial x_j}(a)x_j\right)$

Then $M(T) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(a) & \ldots & \frac{\partial f_1}{\partial x_n}(a) \\ \vdots & & \\ \frac{\partial f_m}{\partial x_1}(a) & \ldots & \frac{\partial f_m}{\partial x_n}(a) \end{pmatrix}$ and

we may therefore think of $T(x)$ as $\begin{pmatrix} \frac{\partial f_1}{\partial x_1}(a) & \cdots & \frac{\partial f_1}{\partial x_n}(a) \\ \vdots & & \\ \frac{\partial f_m}{\partial x_1}(a) & \cdots & \frac{\partial f_m}{\partial x_m}(a) \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$

To see that $T = Df(a)$ (i.e. that T is the derivative) observe that

$$\lim_{x \to a} \frac{\|f(x) - f(a) - T(x-a)\|}{\|x-a\|} = \lim_{x \to a} \frac{\|(f_1(x) - f_1(a) - Df_1(a)(x-a), \ldots, f_m(x) - f_m(a) - Df_m(a)(x-a))\|}{\|x-a\|}$$

$$\leq \lim_{x \to a} \sum_{i=1}^{m} \frac{|f_i(x) - f_i(a) - Df_i(a)(x-a)|}{\|x-a\|}$$ by triangle inequality.

Since the limit on the right is 0, the desired result follows.

Ex. Find the derivative of $f(x_1, x_2) = (x_1x_2 + x_1^3x_2^2, x_1^2+x_2^2, \frac{x_1}{x_2})$ at the point $(1,-1)$

Solution: The component functions are

$$f_1(x_1, x_2) = x_1x_2 + x_1^3x_2^2$$

$$f_2(x_1, x_2) = x_1^2 + x_2^2$$

$$f_3(x_1, x_2) = \frac{x_1}{x_2}$$

The Jacobian matrix is

$$Jf(1,-1) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(1,-1) & \frac{\partial f_1}{\partial x_2}(1,-1) \\ \frac{\partial f_2}{\partial x_1}(1,-1) & \frac{\partial f_2}{\partial x_2}(1,-1) \\ \frac{\partial f_3}{\partial x_1}(1,-1) & \frac{\partial f_3}{\partial x_2}(1,-1) \end{pmatrix} =$$

$$= \left.\begin{pmatrix} x_2 + 3x_1^2x_2^2 & x_1 + 2x_1^3x_2 \\ 2x_1 & 2x_2 \\ \frac{1}{x_2} & \frac{-x_1}{x_2^2} \end{pmatrix}\right|_{(1,-1)} = \begin{pmatrix} 2 & -1 \\ 2 & -2 \\ -1 & -1 \end{pmatrix}$$

Hence the derivative is

$$Df(1,-1)(\vec{x}) = \begin{pmatrix} 2 & -1 \\ 2 & -2 \\ -1 & -1 \end{pmatrix} \vec{x}^T \quad \text{or}$$

$$Df(1,-1)(x_1,x_2) = \begin{pmatrix} 2 & -1 \\ 2 & -2 \\ -1 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

<u>Differential Approximation (optional)</u>

The definition of differentiability says that the difference

$$f(\vec{x}) - f(\vec{a}) - Df(\vec{a})(\vec{x}-\vec{a})$$

is less than a constant multiple of $\|\vec{x}-\vec{a}\|$ for choices $\vec{x}$ sufficiently close to $\vec{a}$. So for $\vec{x}$ near $\vec{a}$, this difference is approximately 0. Rearranging the terms, we can write this as

$$f(\vec{x}) \approx f(\vec{a}) + Df(\vec{a})(\vec{x}-\vec{a}) \qquad (1)$$

for $\vec{x}$ near $\vec{a}$. We call the right-hand side of (1) the differential approximation of f near $\vec{a}$. When the function is real valued, (1) is often written in differential notation:

$$f(\vec{x}) \approx f(\vec{a}) + df.$$

With $\Delta f = f(\vec{x}) - f(\vec{a})$, this is written succinctly as

$$\Delta f \approx df.$$

Ex. Suppose that f is differentiable at the point $(1,2)$, satisfies $f(1,2) = (2,6,-4)$ and has Jacobian matrix

$$Jf(1,2) = \begin{pmatrix} 1 & -1 \\ 0 & 6 \\ -4 & 3 \end{pmatrix}$$

Find an approximation for $f(1.2, 1.9)$

Solution: $f(1.2, 1.9) \approx f(1,2) + Df(1,2)\big((1.2, 1.9) - (1,2)\big) =$

$$= f(1,2) + Jf(1,2)\begin{pmatrix} 0.2 \\ -0.1 \end{pmatrix} =$$

$$= \begin{pmatrix} 2 \\ 6 \\ -4 \end{pmatrix} + \begin{pmatrix} 1 & -1 \\ 0 & 6 \\ -4 & 3 \end{pmatrix}\begin{pmatrix} 0.2 \\ -0.1 \end{pmatrix} = \begin{pmatrix} 2.3 \\ 5.4 \\ -5.1 \end{pmatrix}$$

Ex. The earth's surface is not actually spherical; it's "flattened" at the poles so that cross sections through the north and south poles are ellipses with semimajor radius 6370 km and semiminor radius 6350 km. Approximately what volume of water is needed to provide 1mm of rainfall over the entire earth's surface?

Solution: The earth is an ellipsoid with volume $V(6370, 6370, 6350)$ where $V(a,b,c) = \frac{4\pi}{3} abc$.

If we cover the entire planet by a film of water $1mm = 0.000001 km$ thick, the volume will be $V(6370+1\times10^{-6}, 6370+1\times10^{-6}, 6350+1\times10^{-6})$. We are interested in the volume of the 1mm thick film which is $V(6370+1\times10^{-6}, 6370+1\times10^{-6}, 6350+1\times10^{-6}) - V(6370, 6370, 6350) = \Delta V \approx dV = \frac{\partial V}{\partial a} da + \frac{\partial V}{\partial b} db + \frac{\partial V}{\partial c} dc$ where

$da = db = dc = 10^{-6}$ and

$$\left.\frac{\partial V}{\partial a}\right|_{(6370, 6370, 6350)} = 1.6943\times10^{8}$$

$$\left.\frac{\partial V}{\partial b}\right|_{(6370, 6370, 6350)} = 1.6943\times10^{8}$$

$$\left.\frac{\partial V}{\partial c}\right|_{(6370, 6370, 6350)} = 1.6997\times10^{8}$$

Hence $dV = 1.6943\times10^{8}\times10^{-6} + 1.6943\times10^{8}\times10^{-6} + 1.6997\times10^{8}\times10^{-6}$

$= 509\ km^3$

Ex. A function may have partials at a point but fail to be differentiable at that point.

Consider, for instance,

$$f(x,y) = \begin{cases} \dfrac{x^3-y^3}{x^2+y^2}, & x^2+y^2 \neq 0 \\ 0. & x^2+y^2 = 0 \end{cases}$$

We can easily calculate $\frac{\partial f}{\partial x}(0,0)$ and $\frac{\partial f}{\partial y}(0,0)$:

$$\frac{\partial f}{\partial x}(0,0) = \lim_{h\to 0} \frac{f(h,0)-f(0,0)}{h} = \lim_{h\to 0} \frac{h^3/h^2 - 0}{h} = 1$$

$$\frac{\partial f}{\partial y}(0,0) = \lim_{h\to 0} \frac{f(0,h)-f(0,0)}{h} = \lim_{h\to 0} \frac{-h^3/h^2 - 0}{h} = -1$$

If f is differentiable at $(0,0)$, then the total derivative is the linear transformation $T(x,y) = 1\cdot x - 1\cdot y = x-y$. But

$$\lim_{(x,y)\to(0,0)} \frac{|f(x,y)-f(0,0)-(x-y)|}{\sqrt{x^2+y^2}} = \lim_{(x,y)\to(0,0)} \frac{x^2y-y^2x}{(x^2+y^2)^{3/2}}.$$

does not exist (why?)