(2.6)

The Gradient

Consider $f: \mathcal{U} \subset \mathbb{R}^n \to \mathbb{R}$. The Jacobian matrix of f at $a \in \mathcal{U}$, $Jf(a)$, is the $1 \times n$ matrix $\left(\frac{\partial f}{\partial x_1}(a) \; \frac{\partial f}{\partial x_2}(a) \; \dots \; \frac{\partial f}{\partial x_n}(a) \right)$. This matrix is very similar to to the n-component vector $\left(\frac{\partial f}{\partial x_1}(a), \frac{\partial f}{\partial x_2}(a), \dots, \frac{\partial f}{\partial x_n}(a) \right)$ called the gradient of f at a. The gradient is denoted compactly by $\nabla f(a)$ and, on first glance, can be distinguished from $Jf(a)$ by the presence of commas that separate the entries (components) of the gradient. As we shall see, the gradient has many important geometric properties.

Def: The gradient of $f: \mathcal{U} \subset \mathbb{R}^n \to \mathbb{R}$ is the vector-valued function given by

$$\nabla f(x) = \left(\frac{\partial f}{\partial x_1}(x), \frac{\partial f}{\partial x_2}(x), \dots, \frac{\partial f}{\partial x_n}(x) \right).$$

Ex. Let $f(x, y, z) = yz - xz + xy$. Calculate $\nabla f(x, y, z)$.

Solution:

$$\nabla f(x,y,z) = \left(\frac{\partial f}{\partial x}(x,y,z), \frac{\partial f}{\partial y}(x,y,z), \frac{\partial f}{\partial z}(x,y,z) \right) =$$

$$= (y - z, \; z + x, \; y - x)$$

Notice that, for $f: \mathcal{U} \subset \mathbb{R}^n \to \mathbb{R}$, $Df(a)(x) = \nabla f(a) \cdot x$.

Directional derivatives

By thinking of $f: \mathcal{U} \subset \mathbb{R}^2 \to \mathbb{R}$ as a function describing the altitude of each point $(x, y, f(x,y))$ on the landscape $G_f = \{(x, y, f(x,y)) : (x,y) \in \mathcal{U}\}$, the partial derivatives $\frac{\partial f}{\partial x}(x,y)$ and $\frac{\partial f}{\partial y}(x,y)$ may be understood geometrically as measurements of the steepness of climb in the x and y-axis directions respectively. That is, if you find yourself standing on a point with coordinates $(a, b, f(a,b))$, $\frac{\partial f}{\partial x}(a,b)$ desribes whether, by moving in the positive x-axis direction, you'll be walking up or downhill. The magnitude $|\frac{\partial f}{\partial x}(a,b)|$ describes the steepness of this walk. Similarly $\frac{\partial f}{\partial y}(a,b)$ measures the steepness and direction of the incline along the positive y-axis.

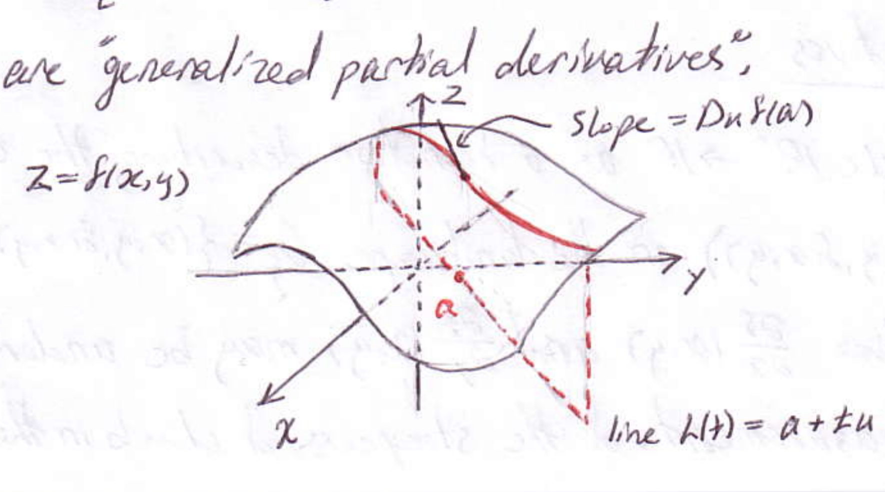
Generally, we might be interested to measure the slope of incline in the direction u, where u is a unit vector (u_1, u_2). To define this notion precisely, notice that $\frac{\partial f}{\partial x}(a) = \lim_{h \to 0} \frac{f(a+he_1) - f(a)}{h}$ and $\frac{\partial f}{\partial y}(a) = \lim_{h \to 0} \frac{f(a+he_2) - f(a)}{h}$.

Def: Let $f: U \subset \mathbb{R}^n \to \mathbb{R}$, let $a \in U$, and let u be a unit vector in $\mathbb{R}^n$. The directional derivative of f at a in the direction u is

$$D_u f(a) = \lim_{h \to 0} \frac{f(a+hu) - f(a)}{h}$$

provided that the limit exists.

Notice that $\frac{\partial f}{\partial x_i}(a) = D_{e_i} f(a)$ (Why?). In other words, directional derivatives are "generalized partial derivatives".



The drawing above illustrates the geometric significance of the directional derivative. If we think of $z = f(x,y)$ as the altitude of the point $(x, y, f(x,y))$ on the surface, $z = f(L(t))$ is a curve in the L-z axis. where L is the line generated by $L(t) = a + tu$.

If we let $g: \mathbb{R} \to \mathbb{R}$ be defined by $g(t) = f(L(t))$, we see that

$$D_u f(a) = \lim_{h \to 0} \frac{f(a+hu) - f(a)}{h} = \lim_{h \to 0} \frac{g(0+h) - g(0)}{h} = g'(0).$$

This motivates the following theorem, which will be useful in computing partial derivatives.

Thm: If $f: \mathcal{U} \subset \mathbb{R}^n \to \mathbb{R}$ is differentiable at a and u is a unit vector, then

$D_u f(a) = \nabla f(a) \cdot u$

Proof: Let $L(t) = a + tu$ and $g(t) = f(L(t))$. Then, as mentioned above, $D_u f(a) = g'(0)$. In terms of Calc III, $g'(0) \equiv Jg(0)$. In other words, since $g: \mathbb{R} \to \mathbb{R}$, $Dg(0)(t)$ is the total derivative of g at 0, which is the tangent line with slope $\frac{\partial g}{\partial t}(0) = \frac{dg}{dt}(0) \equiv \left(\frac{dg}{dt}(0)\right) = Jg(0)$.

By the chain rule, $Jg(0) = J(f \circ L)(0) = Jf(L(t))\Big|_{t=0} JL(0) =$

$= Jf(a)\,(L'(0))^T = Jf(a)\, u^T \equiv \nabla f(a) \cdot u$

Ex. Find the directional derivative of

$$f(x_1, x_2, x_3) = 2x_1^3 x_2^2 x_3$$

at the point $(-1, 1, -1)$ in the direction $(6, 11, 7)$

Solution: A unit vector in the prescribed direction is $u = \frac{(6,11,7)}{\sqrt{36+121+49}} =$

$= \frac{(6,11,7)}{\sqrt{206}}$

Thus $D_u f(-1,1,-1) = \left(6x_1 x_2^2 x_3,\ 4x_1^3 x_2 x_3,\ 2x_1^3 x_2^2\right)\Big|_{(-1,1,-1)} \cdot \frac{(6,11,7)}{\sqrt{206}} =$

$= \frac{(-6,4,-2)\cdot(6,11,7)}{\sqrt{206}} = \frac{-6}{\sqrt{206}}$

For a fixed value a, we can vary the unit vector u. Since u has constant magnitude, this amounts to viewing $D_u f(a)$ as a function of the angle between a and u: $D_u f(a) = \nabla f(a)\cdot u = \|\nabla f(a)\|\,\|u\|\cos\theta =$
$= \|\nabla f(a)\|\cos\theta$ since $\|u\| = 1$ (why?).

Observe that $\|\nabla f(a)\|\cos\theta \leq \|\nabla f(a)\|$, with equality when $\theta = 0$.
Thus, when u has the same direction as the vector $\nabla f(a)$, the directional derivative is largest. In other words, when $\theta = 0$, $\nabla f(a)$ and u are linearly dependent. Since u is a unit vector and since u is a scaler multiple of $\nabla f(a)$ in the direction of $\nabla f(a)$, $u = \frac{\nabla f(a)}{\|\nabla f(a)\|}$,
The inequality $\|\nabla f(a)\|\cos\theta \leq \|\nabla f(a)\|$ implies that $D_u f(a) \leq \|\nabla f(a)\|$ for all unit vectors u with equality if and only if $u = \frac{\nabla f(a)}{\|\nabla f(a)\|}$.
It follows that $D_{\frac{\nabla f(a)}{\|\nabla f(a)\|}} f(a) = \|\nabla f(a)\|$.

This means that $\frac{\nabla f(a)}{\|\nabla f(a)\|}$ is the direction of steepest climb from position $(a, f(a))$ on the graph. Similar reasoning indicates that $-\frac{\nabla f(a)}{\|\nabla f(a)\|}$ is the direction of steepest descent from $(a, f(a))$ (why?)

Ex. Olga Olaafsen is the world's toughest mountain climber; she always takes the steepest route up the mountainside. Suppose that she finds herself at the point $(1,3,67.38)$ on the surface of Mount Gauss, whose elevation in feet at each point is given by $10,000e^{-x^2-(y-1)^2}$. In what direction should she head to maintain her reputation?

Solution: If $f(x,y) = 10,000e^{-x^2-(y-1)^2}$ she should head in the direction

$$\frac{\nabla f(1,3)}{\|\nabla f(1,3)\|} = \left.\frac{10,000\left(-2xe^{-x^2-(y-1)^2}, -2(y-1)e^{-x^2-(y-1)^2}\right)}{\left\|10,000\left(-2xe^{-x^2-(y-1)^2}, -2(y-1)e^{-x^2-(y-1)^2}\right)\right\|}\right|_{(x,y)=(1,3)} =$$

$$= \left.\frac{-(x, y-1)}{\|(x,y-1)\|}\right|_{(1,3)} = \frac{-(1,2)}{\sqrt{5}} = \left(\frac{-1}{\sqrt{5}}, -\frac{2}{\sqrt{5}}\right)$$

There is a close relationship between the gradient of a function of two variables and the level curves of the function. Let f be differentiable at $a \in \mathbb{R}^2$ and let $f(x,y) = c$ be the equation for the level curve of f that passes through a. We can think of the level curve $L_c = \{x \in \mathbb{R}^2; f(x) = c\}$. as parametrized by $x = r(t)$, so that $r(0) = a$.

The equation $f(x,y) = c$ can now be written as $f(r(t)) = c$.

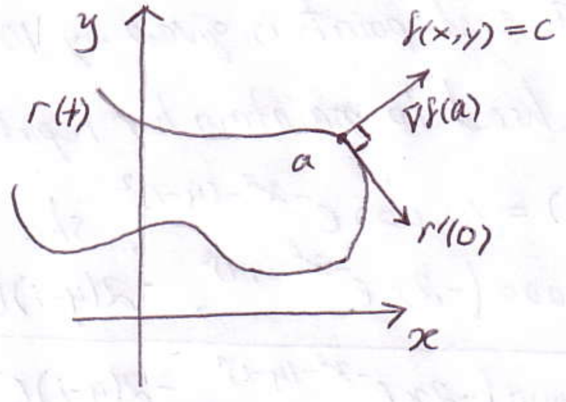
Differentiating both sides and using the chain rule on the left, we have

$$\frac{d}{dt}\left(f(r(t))\right) = \frac{d}{dt}c \;;$$

$$\nabla f(r(t))\cdot r'(t) = 0 \quad \text{or, when } t=0, \quad \nabla f(a)\cdot r'(0) = 0$$

If $r'(0) \neq 0$, this means that $\nabla f(a)$ is perpendicular to a vector that is tangent to the level curve through a. Since this reasoning works for any point on the level curve, we have proven the following theorem.

Thm: If $\nabla f(a) \neq 0$ and the level curve of f through a has a tangent vector T at a, then $\nabla f(a)$ is perpendicular to T.



Ex. Find the equation of the line tangent to the curve $3x^2y^4 - x^2 = 8$ at the point $(-2,1)$.

Solution: let $f(x,y) = 3x^2y^4 - x^2$ and let $r(t)$ be a parametrization of the level curve $L_8 = \{(x,y) \in \mathbb{R}^2;\ f(x,y) = 8\}$ with $r(0) = (-2,1)$.

The tangent line to the curve $\{(x,y) \in \mathbb{R}^2; f(x,y) = 8\} = \{(x,y);\ 3x^2y^4 - x^2 = 8\}$ at $(-2,1)$ is the line parametrized by $L(s) = (-2,1) + s\,r'(0)$.

We can obtain $r'(0)$ from the equation $\nabla f(-2,1) \cdot r'(0) = 0$:

$$\nabla f(-2,1) = \left(6xy^4 - 2x,\ 12x^2y^3\right)\Big|_{(-2,1)} = (-8, 48)$$

Since $r'(0) = (r_1'(0), r_2'(0)) \in \mathbb{R}^2$, all vectors perpendicular to $\nabla f(-2,1)$ are scaler multiples of each other. If (α, β) is perpendicular to $\nabla f(-2,1)$,

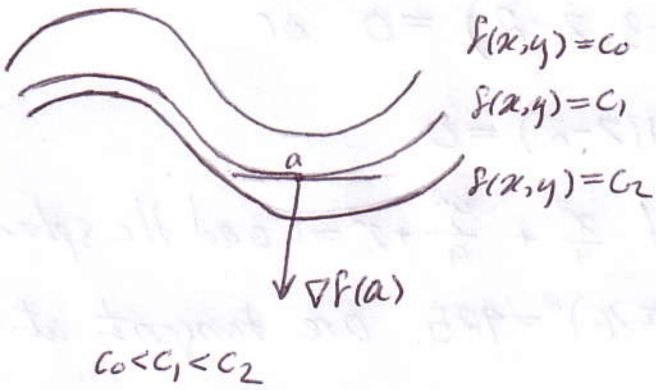
$-8\alpha + 48\beta = 0 \Rightarrow \beta = \frac{-8}{-48}\alpha = \frac{1}{6}\alpha$. Hence $r'(0)$ is a scalar multiple of $(1, \frac{1}{6})$

Thus the tangent line may be parametrized by $L(s) = (-2,1) + s(1, \frac{1}{6})$ or by $p(s) = (-2,1) + s(6,1)$ etc.

Alternatively, recall that a line in $\mathbb{R}^2$ is just "like a plane" in $\mathbb{R}^3$; we can think of $\nabla f(a)$ as normal to the line. Thus, the line through $(-2,1)$ with normal vector $\nabla f(a)$ is the set $\{(x,y) : \nabla f(a)\cdot((x,y)-(-2,1))=0\}$

Hence the equation of this line is given by $(-8,48)\cdot((x,y)-(-2,1))=0$ which simplifies to $-x+6y-8=0$.

Note too that since $\nabla f(a)$ points in the direction of greatest increase in f, if $\bar{c}>c$ and the two constants are close, then $\nabla f(a)$ points towards the side of the level curve $f(x,y)=c$ on which the level curve $f(x,y)=\bar{c}$ lies.



The same ideas carry over to higher-dimensional spaces. For example if $f: U\subset\mathbb{R}^3\to\mathbb{R}$, $f(x,y,z)=c$ describes a level surface. If a is a point on this level surface and $r(t)$ is any path s.t. $r(0)=a$ and $r(t)$ is a point on the level surface for any $t\in\mathbb{R}$, then $f(r(t))=c$, implying that $\nabla f(a)\cdot r'(0)=0$. Thus $\nabla f(a)$ is perpendicular to the curve at a.

Since the choice of curve was arbitrary, $\nabla f(a)$ is perpendicular to each vector in the "bundle" of possible tangents at a. These vectors span a plane through a which we call the plane tangent to $f(x,y,z)=c$ at a.

Its equation is $\nabla f(a)\cdot(x-a)=0$ or in terms of components,

$$\frac{\partial f}{\partial x}(a)(x-a_1)+\frac{\partial f}{\partial y}(a)(y-a_2)+\frac{\partial f}{\partial z}(a)(z-a_3)=0$$

Note that $\frac{\nabla f(a)}{\|\nabla f(a)\|}$ is the direction of most rapid climb from position $(a, f(a))$ for $f: U \subset \mathbb{R}^n \to \mathbb{R}$. The rate of climb in the direction $\frac{\nabla f(a)}{\|\nabla f(a)\|}$ is $\|\nabla f(a)\|$.

Ex. Find an equation for the plane tangent to the surface $x^2 - y^2 + z^2 = 0$ at $(\sqrt{5}, 3, 2)$

Solution: $\nabla f(\sqrt{5}, 3, 2) = (2x, -2y, 2z)\Big|_{(\sqrt{5}, 3, 2)} = (2\sqrt{5}, -6, 4)$

Consequently, the plane is given by the equation

$(2\sqrt{5}, -6, 4) \cdot (x - \sqrt{5}, y - 3, z - 2) = 0$ or

$2\sqrt{5}(x - \sqrt{5}) - 6(y - 3) + 4(z - 2) = 0.$

Ex. Show that the ellipsoid $\frac{x^2}{4} + \frac{y^2}{9} + z^2 = 1$ and the sphere $(x-10)^2 + (y-5)^2 + (z - 37\sqrt{23}/6)^2 = 925$ are tangent at their point of intersection $(1, 1, \sqrt{23}/6)$.

Solution: let $f(x,y,z) = \frac{x^2}{4} + \frac{y^2}{9} + z^2$ and $g(x,y,z) = (x-10)^2 + (y-5)^2 + (z - 37\sqrt{23}/6)^2$

Then the level surfaces $f(x,y,z) = 1$ and $g(x,y,z) = 925$ contain the point $(1, 1, \sqrt{23}/6)$.

Observe that $\nabla f(1, 1, \sqrt{23}/6) = \left(\frac{x}{2}, \frac{2y}{9}, 2z\right)\Big|_{(1,1,\sqrt{23}/6)} = \left(\frac{1}{2}, \frac{2}{9}, \frac{\sqrt{23}}{3}\right)$

and $\nabla g(1, 1, \sqrt{23}/6) = \left(2(x-10), 2(y-5), 2z - \frac{37}{3}\sqrt{23}\right)\Big|_{(1,1,\sqrt{23}/6)} = (-18, -8, -12\sqrt{23})$

$= -36\left(\frac{1}{2}, \frac{2}{9}, \frac{\sqrt{23}}{3}\right)$. Thus ∇f and ∇g are parallel at $(1, 1, \sqrt{23}/6)$.

This means that $\nabla f(1, 1, \sqrt{23}/6) \cdot ((x,y,z) - (1, 1, \sqrt{23}/6)) = 0$ and $\nabla g(1, 1, \sqrt{23}/6) \cdot ((x,y,z) - (1, 1, \sqrt{23}/6)) = 0$ are equations describing the same plane.

Application to single-variable Calculus

Ex. Find the derivative of $f(x) = \int_1^x \frac{\cos xt}{t}\,dt$.

Solution: If $f(x)$ were equal to $\int_1^x \frac{\cos t}{t}\,dt$, we could have applied the fundamental theorem of calculus and $\frac{df}{dx}$ would have been $\frac{\cos x}{x}$. However, f is not of the form $f(x) = \int_a^x g(t)\,dt$ and the fundamental theorem of calculus cannot be directly applied. Notice that if we define $P(x,y) = \int_1^x \frac{\cos yt}{t}\,dt$, then

$$\frac{d}{dx} f(a) = \lim_{h\to 0} \frac{f(a+h) - f(a)}{h} = \lim_{h\to 0} \frac{P(a+h, a+h) - P(a,a)}{h} =$$

$$= \lim_{h\to 0} \frac{P[(a,a)+h(1,1)] - P(a,a)}{h}.$$

Let $L(t) = (a,a) + t(1,1)$. Then $\frac{d}{dx} f(a) = \frac{d}{dt} P(L(t))\Big|_{t=0} =$

$$= \nabla P \cdot L'(t)\Big|_{t=0} = \left(\frac{\partial P}{\partial x}(a,a), \frac{\partial P}{\partial y}(a,a)\right)\cdot(1,1) =$$

$$= \frac{\partial P}{\partial x}(a,a) + \frac{\partial P}{\partial y}(a,a) = \frac{\partial}{\partial x}\int_1^x \frac{\cos yt}{t}\,dt\Big|_{(a,a)} + \frac{\partial}{\partial y}\int_1^x \frac{\cos yt}{t}\,dt\Big|_{(a,a)} =$$

$$= \frac{\cos yx}{x}\Big|_{(a,a)} + \int_1^x \frac{\partial}{\partial y}\frac{\cos yt}{t}\,dt\Big|_{(a,a)} = \frac{\cos a^2}{a} + \int_1^x -\sin yt\,dt\Big|_{(a,a)} =$$

$$= \frac{\cos a^2}{a} + \frac{\cos yt}{y}\Big|_1^x\Big|_{(a,a)} = \frac{\cos a^2}{a} + \left(\frac{\cos xy}{y} - \frac{\cos y}{y}\right)\Big|_{(a,a)} =$$

$$= \frac{\cos a^2}{a} + \frac{\cos a^2}{a} - \frac{\cos a}{a} = \frac{2\cos a^2 - \cos a}{a}.$$

Hence $f'(x) = \frac{2\cos x^2 - \cos x}{x}$.