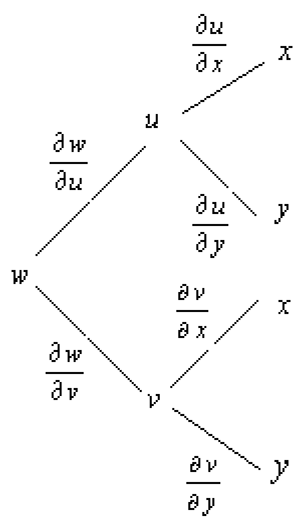
Regla de la cadena

Una de las reglas que en el cálculo de una variable resulta muy útil es la **regla de la cadena**. Dicho grosso modo, esta regla sirve para derivar una composición de funciones, esto es, una función que depende de una variable que a su vez depende de otra variable. Aunque en el caso de varias variables, el aspecto que puede tomar dicho resultado puede parecer muy variado y mucho más complicado, es posible presentarlo con el "mismo aspecto". Así, nuestra intención en la siguiente discusión no será demostrar dicho resultado (lo que cae fuera del nivel de este curso), sino mostrar cómo se le puede plantear bajo un artificio de tipo nemotécnico. Mostramos el resultado en un caso particular que corresponde al de una función de dos variables mismas que dependen a su vez de otras dos variables, insistimos no obstante que el resultado puede y de hecho tiene variantes a esta presentación.

Regla de la cadena
Si $w = F(u,v)$, $u = f(x,y)$ y $v = g(x,y)$ donde F, f y g son diferenciables, entonces:

$$\frac{\partial w}{\partial x} = \left(\frac{\partial w}{\partial u}\right)\left(\frac{\partial u}{\partial x}\right) + \left(\frac{\partial w}{\partial v}\right)\left(\frac{\partial v}{\partial x}\right) \quad \& \quad \frac{\partial w}{\partial y} = \left(\frac{\partial w}{\partial u}\right)\left(\frac{\partial u}{\partial y}\right) + \left(\frac{\partial w}{\partial v}\right)\left(\frac{\partial v}{\partial y}\right)$$

El resultado anterior y otros, pueden ser recordados nemotécnicamente por medio de un símil del diagrama de árbol utilizado, por ejemplo, en probabilidad. La idea es simple: se inicia con la función w y de ahí se van marcando (con líneas) hacia la derecha las variables de las que depende. Posteriormente, cada variable de la que depende w se pone bajo la misma idea en dependencia de sus propias variables. Ahora, cada segmento uniendo variables se marca con una derivada parcial, ésta es la derivada parcial de la variable a la izquierda con respecto a la variable a la derecha. De manera más concreta, considera el siguiente diagrama:



Esquema nemotécnico para la regla de la cadena

Derivada direccional

Hemos hablado ya de la derivada parcial, y con ello tenemos la herramienta matemática que nos permite determinar la razón de cambio de una función cuando variamos una variable a la vez dando “un paso” en cualquiera de las direcciones determinadas por alguno de los vectores de la base canónica de $\mathbb{R}^n$. Claro, debes leer aquí “un paso” relacionándolo con la magnitud de los vectores de la base canónica de $\mathbb{R}^n$, todos los cuales sin distinción son vectores unitarios. Sin embargo, es frecuente que requiramos determinar la razón de cambio de una función en cualquier dirección. Es de esperar que lo que ganaremos en la generalización de este concepto, lo perderemos en la facilidad de cálculo de una derivada parcial. Nuestro trabajo en esta sección será, no obstante, mostrar que podemos ganar la citada generalización sin incurrir en cálculos complicados. Para lograr nuestro cometido nos apoyaremos en la regla de la cadena.

En primer lugar echemos un vistazo a la definición de derivada parcial, ésta es:

$$\frac{\partial f}{\partial x_j}\left(\overrightarrow{P_0}\right)=\lim_{h\to 0}\frac{f\left(\overrightarrow{P_0}+h\overrightarrow{e_j}\right)-f\left(\overrightarrow{P_0}\right)}{h}$$

Con base en lo anterior, nuestra definición ahora es muy simple, se basa únicamente en modificar el **vector unitario** particular de la base canónica por cualquier **vector unitario** $\vec{u}$ De esta manera tenemos:

Definición. Derivación direccional
Sean U un conjunto abierto de $\mathbb{R}^n$, $\overrightarrow{P_0}=\left(x_1,x_2,...,x_n\right)\in U$ y $f:U\to\mathbb{R}$ una función de varias variables. Si $\vec{u}$ es cualquier vector unitario de $\mathbb{R}^n$, **la derivada direccional** de la función f en $\overrightarrow{P_0}$ y en la dirección de $\vec{u}$ se define por:

$$D_u\, f\left(\overrightarrow{P_0}\right)=\lim_{h\to 0}\frac{f\left(\overrightarrow{P_0}+h\vec{u}\right)-f\left(\overrightarrow{P_0}\right)}{h}$$

Dada la importancia del anterior concepto, vale la pena insistir en lo siguiente: *la derivada direccional mide la razón de cambio de la función f cuando al estar en* $\overrightarrow{P_0}$ *se da “un paso” en la dirección del vector unitario* $\vec{u}$.

Ahora tenemos que resolver un asunto: ¿cómo podemos calcular una derivada direccional? Es claro que quisiéramos mantener de alguna manera la comodidad que teníamos en el cálculo de derivadas parciales. La buena noticia es que esto puede hacerse con ayuda de la regla de la cadena.

Para lograr nuestro cometido, introduzcamos la función $g:\mathbb{R}\to\mathbb{R}$ definida por:

$$g\left(t\right)=f\left(\overrightarrow{P_0}+t\vec{u}\right)$$

Observa que $g(0)=f\left(\overrightarrow{P_0}\right)$ y por consiguiente:

$$\lim_{h\to 0}\frac{f\left(\overrightarrow{P_0}+h\vec{u}\right)-f\left(\overrightarrow{P_0}\right)}{h}=\lim_{h\to 0}\frac{g(h)-g(0)}{h}=g'(0)$$

Éste, es un gran adelanto que vale la pena considerar porque hemos hallado una forma para calcular derivadas direccionales sin necesidad de recurrir a su definición. Así, antes de avanzar a la siguiente idea veamos cómo se le puede aplicar.

Ejemplo.

Sean $\vec{u}$ un vector unitario cualquiera y $\overrightarrow{P_0}$ un vector fijo. Calcula $D_u f\left(\overrightarrow{P_0}\right)$ si $f\left(\vec{x}\right)=\left\|\vec{x}\right\|^2$ para todo $\vec{x}\ \varepsilon\ \mathbb{R}^n$.

Solución:

En primer lugar definimos nuestra función auxiliar $g(t)=f\left(\overrightarrow{P_0}+t\vec{u}\right)$, ésta aplicada a nuestro caso particular queda:

$$g(t)=f\left(\overrightarrow{P_0}+t\vec{u}\right)=\|\overrightarrow{P_0}+t\vec{u}\,\|^2$$

Entonces:

$$\begin{aligned}g(t)&=f\left(\overrightarrow{P_0}+t\vec{u}\right)=\left\|\overrightarrow{P_0}+t\vec{u}\right\|^2\\&=\left(\overrightarrow{P_0}+t\vec{u}\right)\cdot\left(\overrightarrow{P_0}+t\vec{u}\right)\\&=\overrightarrow{P_0}\cdot\overrightarrow{P_0}+2t\left(\overrightarrow{P_0}\cdot\vec{u}\right)+t^2\left(\vec{u}\cdot\vec{u}\right)\\&=\left\|\overrightarrow{P_0}\right\|^2+2t\left(\overrightarrow{P_0}\cdot\vec{u}\right)+t^2\left\|\vec{u}\right\|^2\\&=\left\|\overrightarrow{P_0}\right\|^2+2t\left(\overrightarrow{P_0}\cdot\vec{u}\right)+t^2,\text{ puesto que }\left\|\vec{u}\right\|=1.\end{aligned}$$

NOTA: El símbolo $\cdot$ indica **producto punto** (o **escalar**) entre vectores.

El resto tampoco es complicado, sólo nos resta derivar con respecto a t y evaluar la derivada en $t=0$. Tenemos:

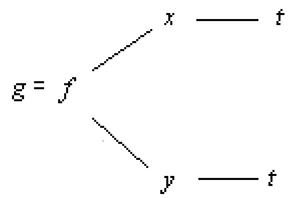
$$g'(t)=2\left(\overrightarrow{P_0}\cdot\vec{u}\right)+2t\text{, pues }\left\|\overrightarrow{P_0}\right\|^2\text{ y }\left(\overrightarrow{P_0}\cdot\vec{u}\right)\text{ son constantes.}$$

Finalmente $D_u f\left(\overrightarrow{P_0}\right)=g'(0)=2\left(\overrightarrow{P_0}\cdot\vec{u}\right)$.

Aunque nuestro procedimiento funcionó muy bien, buscaremos ahora una idea alterna que nos permita ver en breve la conexión existente entre derivada parcial y direccional (que con el enfoque anterior no logra apreciarse). Utilizamos ahora la misma idea pero añadiremos a nuestra estrategia la regla de la cadena. No pierdas de vista que deseamos hallar $g'(0)$.

Primero nuestro esquema (que limitaremos a una función de dos variables).

Suponemos en lo que sigue que $\vec{u}=(a,b)$ es un vector direccional fijo y que $\overrightarrow{P_0}=(x_0, y_0)$ es cualquier punto, entonces $g(t)=f\left(\overrightarrow{P_0}+t\vec{u}\right)=f\left(\underbrace{x_0+at}_{x},\underbrace{y_0+bt}_{y}\right)$. Nota que las coordenadas en la primera y segunda entradas de f son variables a las hemos denotado en la forma habitual x, y, y que éstas dependen a su vez de la variable t. Entonces:



De este esquema deducimos que:

$$g'(t)=\frac{\partial f\left(\overrightarrow{P_0}+t\vec{u}\right)}{\partial x}\frac{\partial x}{\partial t}+\frac{\partial f\left(\overrightarrow{P_0}+t\vec{u}\right)}{\partial y}\frac{\partial y}{\partial t}$$

Como $x=x_0+at$, $y=y_0+bt$ resulta que $\frac{\partial x}{\partial t}=a$, y $\frac{\partial y}{\partial t}=b$. En otras palabras:

$$g'(t)=\frac{\partial f\left(\overrightarrow{P_0}+t\vec{u}\right)}{\partial x}a+\frac{\partial f\left(\overrightarrow{P_0}+t\vec{u}\right)}{\partial y}b\,,$$

De donde: $D_u\, f\left(\overrightarrow{P_0}\right)=g'(0)=\frac{\partial f\left(\overrightarrow{P_0}\right)}{\partial x}a+\frac{\partial f\left(\overrightarrow{P_0}\right)}{\partial y}b$

$$=\left(\frac{\partial f\left(\overrightarrow{P_0}\right)}{\partial x},\frac{\partial f\left(\overrightarrow{P_0}\right)}{\partial y}\right)\cdot(a,b)$$

Nada mal, porque resulta que el primero de los vectores se compone de derivadas parciales. Dada su importancia (no sólo en este tema) se le da un nombre especial, a este vector lo llamaremos el **vector gradiente**. El segundo vector es la dirección dada por el vector $\vec{u}$. Por lo tanto hemos obtenido un importante resultado:

Teorema. Sobre el cálculo de la derivada direccional
Si f es una función diferenciable en una región abierta que contiene a $\overrightarrow{P_0}$, entonces:

$$D_u\, f\left(\overrightarrow{P_0}\right)=\nabla f\left(\overrightarrow{P_0}\right)\cdot\vec{u}$$

Donde

$$\nabla f\left(\overrightarrow{P_0}\right) = \left(\frac{\partial f\left(\overrightarrow{P_0}\right)}{\partial x}, \frac{\partial f\left(\overrightarrow{P_0}\right)}{\partial y}\right)$$

Es el vector gradiente de la función f en $\overrightarrow{P_0}$.

El resultado anterior tiene cuando menos las siguientes dos consecuencias importantes, a saber:

a) Como $\vec{A} \cdot \vec{B} = \left\| \vec{A} \right\| \left\| \vec{B} \right\| \cos(\theta)$, se deduce que:

$$\begin{aligned} D_u\, f\left(\overrightarrow{P_0}\right) &= \nabla f\left(\overrightarrow{P_0}\right) \cdot \vec{u} \\ &= \left\| \nabla f\left(\overrightarrow{P_0}\right) \right\| \left\| \vec{u} \right\| \cos(\theta) = \left\| \nabla f\left(\overrightarrow{P_0}\right) \right\| \cos(\theta), \end{aligned}$$

Pues $\left\| \vec{u} \right\| = 1$ ya que $\vec{u}$ es un vector unitario. Ahora, como $-1 \le \cos(\theta) \le 1$, deducimos que la **máxima razón de cambio** de f (igual a $\left\| \nabla f\left(\overrightarrow{P_0}\right) \right\|$) se obtiene cuando $\cos(\theta) = 1$, es decir, cuando $\theta = 0$, esto es, cuando la dirección de $\vec{u}$ coincide con la del vector gradiente, en otras palabras, cuando $\vec{u} = \dfrac{\nabla f\left(\overrightarrow{P_0}\right)}{\left\| \nabla f\left(\overrightarrow{P_0}\right) \right\|}$. De manera similar, la **mínima razón de cambio** (igual a $-\left\| \nabla f\left(\overrightarrow{P_0}\right) \right\|$) se obtiene cuando $\cos(\theta) = -1$, es decir, cuando $\theta = \pi$, es decir, cuando la dirección de $\vec{u}$ coincide con la del vector menos gradiente, en otras palabras, cuando $\vec{u} = -\dfrac{\nabla f\left(\overrightarrow{P_0}\right)}{\left\| \nabla f\left(\overrightarrow{P_0}\right) \right\|}$.

b) La siguiente consecuencia es junto con la del inciso a) la base sobre la cual se apoya el importante teorema de Lagrange sobre **extremos de funciones con restricciones** (**multiplicadores de Lagrange**) que veremos más adelante.

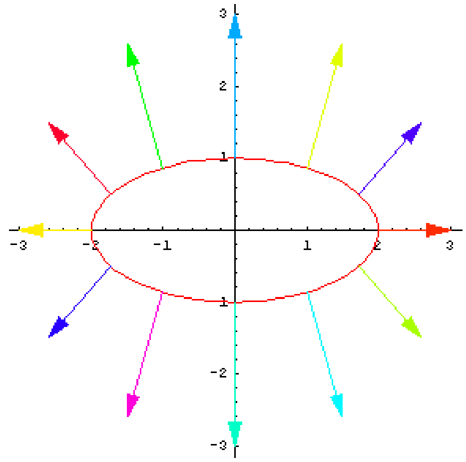
Para fijar ideas, supón que $C_N = \left\{ (x, y) \in D_f : f(x, y) = N \right\}$ es una curva de nivel N para f e imagina que $\vec{r}(t) = \left(x(t), y(t)\right);\ t \in I$, donde I es cierta parte del dominio de la función vectorial $\vec{r}$ es una representación vectorial de tal curva de nivel, es decir que: $f\left(x(t), y(t)\right) = N$. Entonces, al usar la regla de la cadena y derivar con respecto a t encontramos:

$$\frac{\partial f}{\partial x} * \frac{dx}{dt} + \frac{\partial f}{\partial y} * \frac{dy}{dt} = \frac{dN}{dt} = 0$$

Es decir:

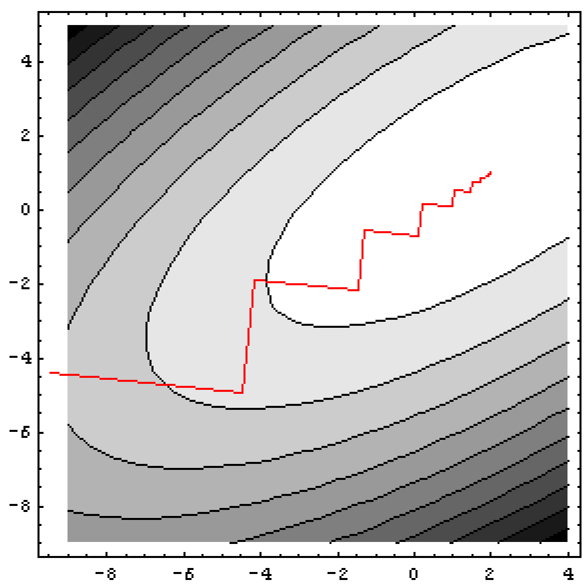
$$\nabla f.\overrightarrow{r'(t)} = 0$$

Como $\overrightarrow{r'(t)}$ es un vector tangente a la trayectoria, esto significa que el vector gradiente es perpendicular a esta trayectoria. Además, ya que $f(x,y) = N$ es una curva de nivel, podemos decir que el vector gradiente se "pega" a las curvas de nivel en un mapa de contorno de forma perpendicular.



El vector gradiente se "pega" a cualquier curva de nivel de forma perpendicular a ésta.

Ésta es la conclusión importante, si consideramos los últimos dos incisos:
"El vector gradiente (menos gradiente) apunta siempre en la dirección del máximo (mínimo) de una función y al "pegarse" a una curva de nivel lo hace de forma ortogonal a ésta".



El vector gradiente (menos gradiente) "localiza" el máximo (mínimo) de una función con restricciones paso a paso, pegándose en todo momento de forma perpendicular a las curvas del mapa de contorno.