5. Optimización: Funciones de varias variables

En los capítulos 3 y 4 hemos trabajado con funciones de una sola variable, lo cual implicaba relacionar dos variables, x e y, estableciendo que y dependiera del valor que tomara x: $y = f(x)$.

Por ejemplo, si estamos estudiando la tasa de mortalidad materna (TMM) en Uruguay, puede pensarse que uno de los factores que inciden en la misma es el gasto público en salud (GPS) y podríamos establecer una relación del tipo:

$$TMM = f(GPS)$$

Sin embargo, la tasa de mortalidad materna no depende en la realidad de un único factor (el gasto público en salud, por ejemplo), sino que en ella inciden un conjunto mucho más amplio de factores, como pueden ser: el porcentaje de partos atendidos por personal capacitado (PAPC), el número de habitantes por médico (HPM), el número de camas de hospital por habitante (CHPH), el gasto público en educación (GPE).

De esta manera, podría plantearse una relación mucho más amplia:

$$TMM = f(GPS, PAPC, HPM, CHPH, GPE)$$

Así, lo que tenemos es que la tasa de mortalidad materna sería una función de varias variables, con lo cual ganamos en profundidad y realismo en nuestro análisis. La forma concreta en que se relacionen estas variables para explicar la tasa de mortalidad materna dependerá de la teoría que tenga asociada. A continuación se presentan tres ejemplos de cómo podría presentarse en forma explícita esa relación:

$$TMM = -3 \times GPS - 4 \times PAPC + 2 \times HPM - 5 \times CHPH - 6 \times GPE$$
$$TMM = -3 \times GPS - PAPC^2 + 2 \times HPM - HPM^2 - 4 \times CHPH \times GPE$$
$$TMM = e^{-GPS - 3 \times PAPC + 5 \times HPM - \times CHPH - 2 \times GPE}$$

La relevancia de identificar una forma funcional concreta radica en saber cómo un cambio en una de las variables independientes o explicativas afecta a la variable dependiente, en este caso, la tasa de mortalidad materna. Por ejemplo, es interesante conocer cómo un cambio en el gasto público en salud afecta la tasa de mortalidad materna. ¿Un incremento en el gasto público en salud ayuda a disminuir la tasa de mortalidad materna? Si es que influye, ¿cuánto debería incrementar el gobierno el gasto público en salud para conseguir disminuir la tasa de mortalidad materna en 1 unidad?

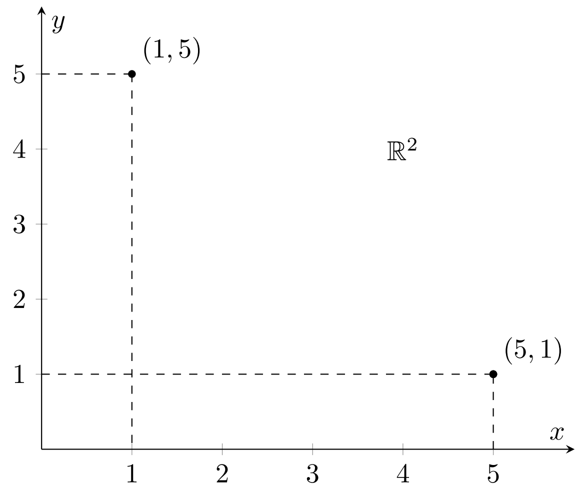
En todas las disciplinas científicas, pero en particular en las Ciencias Sociales es poco frecuente encontrar que cierta variable dependa exclusivamente de una única variable. Es decir, son poco frecuentes las relaciones del tipo: $y = f(x)$. Lo común es que una variable dependa del valor que tomen varias variables, por lo cual se vuelve imperioso considerar funciones que incluyan más de una variable independiente o explicativa a la vez. En este capítulo analizaremos las funciones definidas en dos variables de la forma: $z = f(x, y)$ a fin de dar una idea del análisis en varias variables sin complejizar el análisis. Todo lo que se vea en el presente capítulo puede ser extensible a relaciones del tipo: $z = f(x_1, x_2, ..., x_k)$.

5.1. Análisis de funciones de dos variables

Definición 1. *Llamaremos* $\mathbb{R}^2$ *al conjunto de todos los pares ordenados de números reales.*

Cuando consideramos pares ordenados de numeros, *el orden importa*: $(5,1) \in \mathbb{R}^2$, $(1,5) \in \mathbb{R}^2$ y los consideramos como elementos distintos $(1,5) \neq (5,1)$.

Es posible realizar una representación gráfica de el conjunto $\mathbb{R}^2$ identificandolo con un plano. Para ellos, establecemos un sistema de dos ejes en el plano, esto nos permite asociar cada par ordenado de $\mathbb{R}^2$ con un punto único punto de dicho plano.



Definición 2. *Llamaremos* $\mathbb{R}^3$ *al conjunto de todas las ternas ordenadas de números reales.*

De forma análoga, en $\mathbb{R}^3$ el orden también importa ya que $(1,2,3) \in \mathbb{R}^3$, $(2,3,1) \in \mathbb{R}^3$ pero $(1,2,3) \neq (2,3,1)$. El conjunto $\mathbb{R}^3$ puede asimismo representarse geometricamente al identificarse con el espacio. En este caso, al establecer un sistema de tres ejes en el espacio podemos asociar cada terna ordenada de $\mathbb{R}^3$ con un punto único punto del espacio.

Definición 3. *Decimos que* f *es una* ***función de dos variables*** *si su dominio es un subconjunto de* $\mathbb{R}^2$ *y cuyo codominio son los reales:* $f : D \to \mathbb{R}$ *donde* $D \subset \mathbb{R}^2$

Definición 4. *Decimos que* f *es una* ***función de tres variables*** *si su dominio es un subconjunto de* $\mathbb{R}^3$ *y cuyo codominio son los reales:* $f : D \to \mathbb{R}$ *donde* $D \subset \mathbb{R}^3$

Ejemplo 1: Supongamos que la función en varias variables:

$$\begin{aligned} TMM &= f(GPS, PAPC, HPM, CHPH, GPE) \\ &= -3 \times GPS - 4 \times PAPC + 2 \times HPM - 5 \times CHPH - 6 \times GPE \end{aligned}$$

explica el comportamiento de la tasa de mortalidad materna (en número de muertes maternas cada 100.000 nacidos vivos) a partir de las variables Gasto público en salud (en porcentaje del Producto Interno Bruto - PIB), Porcentaje de partos atendidos por personal capacitado, Número de habitantes por médico, Número de camas de hospital por habitante y Gasto público en educación (en porcentaje del Producto Interno Bruto – PIB).

Supongamos que se tienen datos para dos países:

País	GPS	PAPC	HPM	CHPH	GPE
A	4	90	300	0,02	5
B	6	100	250	0,03	6

Vamos a calcular la tasa de mortalidad materna para uno y otro país, de acuerdo a la función explicitada:

$$TMM_{\text{País A}} = f(4, 90, 300, 0{,}02, 5) = -3(4) - 4(90) + 2(300) - 5(0{,}02) - 6(5) = 197{,}9$$
$$TMM_{\text{País B}} = f(6, 100, 250, 0{,}03, 6) = -3(6) - 4(100) + 2(250) - 5(0{,}03) - 6(6) = 45{,}85$$

Se concluye por lo tanto que el país B tiene una tasa de mortalidad materna mucho menor que el país A.[1]

Ejemplo 2: El Producto Interno Bruto (PIB) de un país se define como el valor de la producción de bienes de uso final realizada en un país o región durante cierto período de tiempo, en general un año. El destino de tal producción puede ser los hogares, las empresas, el gobierno o el exterior. Es así que puede plantearse la siguiente relación:

$$PIB = Consumo_{hogares} + Inversion_{empresas} + Consumo_{gobierno} + Exportaciones - Importaciones$$

En el caso en que la producción sea utilizada por los hogares se denomina Consumo de los hogares (C). Si es comprada por las empresas recibe el nombre de Inversión (I), mientras que si tiene como destino el gobierno se denomina Consumo del gobierno (G). Si se vende al exterior se denomina Exportaciones (X). Por último hay que señalar que los hogares, las empresas y el mismo gobierno compran productos del exterior, denominadas Importaciones (M). Dado que el PIB solo contabiliza la producción realizada en el país, es preciso restar de los demás ítems las Importaciones.

Desde el punto de vista matemático, el PIB es una función de cinco variables:

$$PIB = f(C, I, G, X, M) = C + I + G + X - M$$

Sabiendo el valor de dichas cinco variables se puede calcular el valor del PIB.

Por ejemplo, el valor de dichas variables (en miles de millones de pesos) para los años 2008, 2009 y 2010 se presenta en la siguiente tabla:

[1] A modo informativo, según datos de CEPAL, la tasa de mortalidad materna en Uruguay en 2013 fue de 14 muertes maternas cada 100.000 nacidos vivos, mientras que la de Bolivia fue de 200 muertes maternas cada 100.000 nacidos vivos. Fuente: CEPAL.

Año	C	I	G	X	M
2008	456	148	77	191	219
2009	485	128	92	188	182
2010	554	144	103	209	202

Fuente: INE (Instituto Nacional de Estadística).

Para obtener el PIB de 2010, debemos calcular:

$$\begin{aligned} PIB_{2010} &= f(C_{2010}, I_{2010}, G_{2010}, X_{2010}, M_{2010}) \\ &= f(554, 144, 103, 209, 202) = 554 + 144 + 103 + 209 - 202 = 808 \end{aligned}$$

De la misma manera, para obtener el PIB de 2008 y de 2009, habría que calcular:

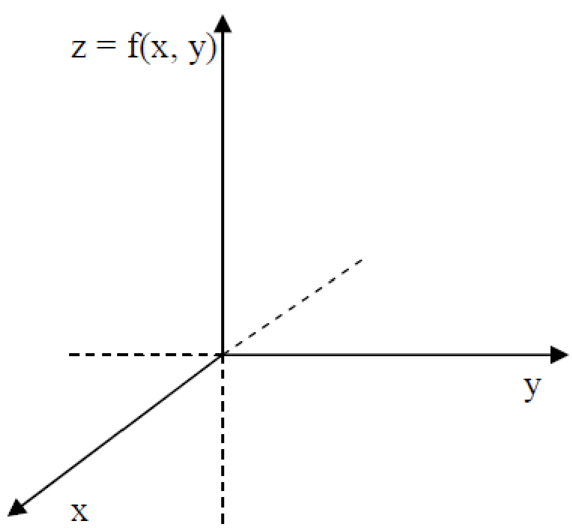
$$PIB_{2008} = f(C_{2008}, I_{2008}, G_{2008}, X_{2008}, M_{2008}) = f(456, 148, 77, 191, 219)$$
$$PIB_{2009} = f(C_{2009}, I_{2009}, G_{2009}, X_{2009}, M_{2009}) = f(485, 128, 92, 188, 182)$$

Realice los cálculos usted mismo y confirme que el PIB de 2008 fue de 653 mil millones de pesos y el de 2009 fue de 711 mil millones de pesos.

Ejemplo 3: Sea la función f tal que $f(x, y) = 2x - 10y - 40$. Calcular el valor que toma la función cuando $x = 1$ y $y = 3$.

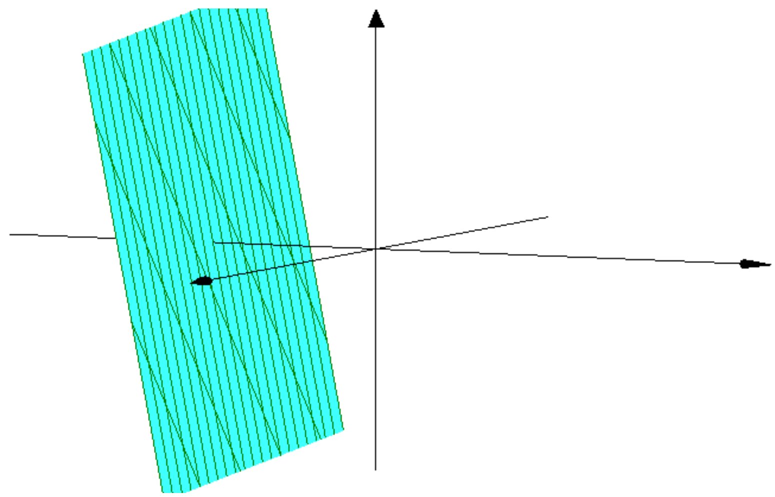
Eso implica calcular: $f(1, 3) = 2(1) - 10(3) - 40 = -68$.

La representación gráfica de funciones de varias variables suele ser mas compleja que en el caso de funciones de una única variable. Nosotros nos concentraremos en la representación gráfica de funciones de dos variables (las representaciones en el caso de 3 o mas variables son significativamente mas complejas y menos intuitivas). Si en el caso de funciones de una variable, considerábamos dos ejes cartesianos perpendiculares para efectuar la representación gráfica, en el caso de funciones de dos variables, deberemos considerar tres ejes, de la siguiente forma:



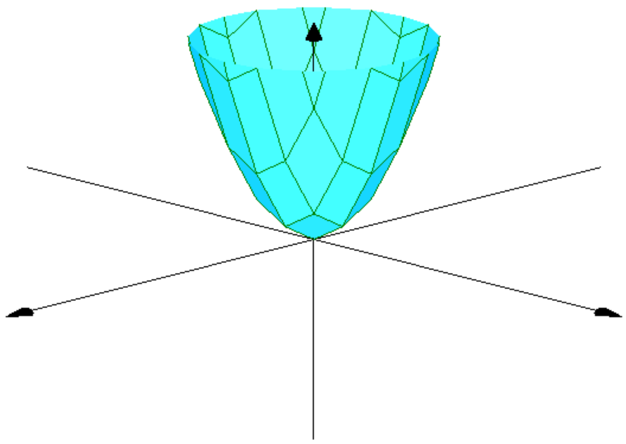
Si bien no discutiremos métodos generales para representar gráficamente cualquier función de dos variables, presentaremos y discutiremos los gráficos de algunos casos sencillos.

Ejemplo 4: Consideremos la función del ejemplo 3: $f : f(x, y) = 2x - 10y - 40$ Su representación gráfica sería:

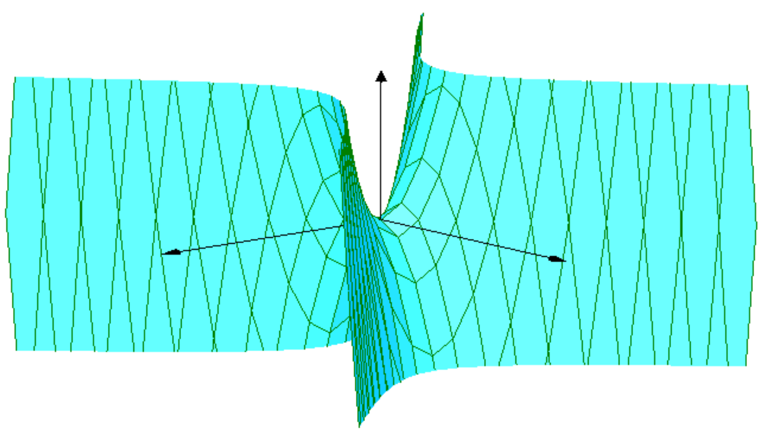


Los gráficos de funciones lineales en x e y, como el caso de la función anterior, tienen la forma de planos.

Ejemplo 5: $f : f(x, y) = x^2 + y^2$



Ejemplo 6: $f : f(x, y) = x^2 - y^2$



Para el caso de más de dos variables independientes, la representación gráfica en el sistema de ejes cartesianos perpendiculares no puede realizarse.

Una manera alternativa de representar el gráfico de una función de dos variables es mediante los denominados ***conjuntos*** o ***curvas de nivel***.

Definición 5. *Sea una función $f : f(x, y)$. Definimos las* ***curvas o conjuntos de nivel*** *$\alpha \in R$ asociadas a la función f, las que denotaremos por $C_{f,\alpha}$ de la siguiente manera:*

$$C_{f,\alpha} = \{(x, y) | f(x, y) = \alpha\}$$

En otras palabras, las curvas de nivel α asociadas a la función f están formadas por todos los puntos del dominio de la función cuya imagen a través de la función es α.

Ejemplo 7: Volvamos a considerar la función del ejemplo 4: $f : f(x,y) = 2x - 10y - 40$.

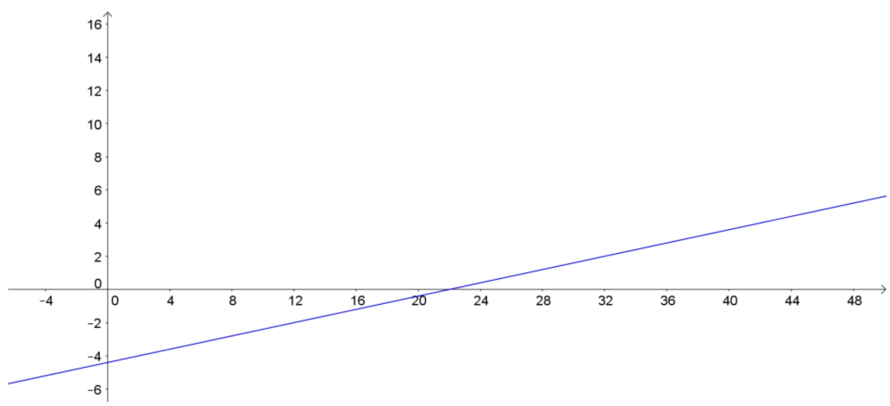
En la página anterior podemos observar su representación gráfica utilizando tres ejes cartesianos para ello. Ahora vamos a servirnos de las curvas de nivel para obtener una representación alternativa de su gráfico, pero utilizando para ello solamente dos ejes.

Vamos a preguntarnos en primer lugar qué puntos (x,y) tienen como valor funcional el número 4, por ejemplo. Esta pregunta la podemos responder determinando la curva de nivel 4 ($\alpha = 4$) para esta función f:

$$C_{f,4} = \{(x,y) | f(x,y) = 4\} = \{(x,y) | 2x - 10y - 40 = 4\}$$

Observando la ecuación que nos queda determinada, apreciamos que esta curva de nivel 4 no es otra cosa que una recta de ecuación $2x - 10y = 44$, o en forma equivalente: $y = 0{,}2x - 4{,}4$.

Además, podemos representar esta recta, considerando dos ejes cartesianos perpendiculares, como se presenta a continuación:

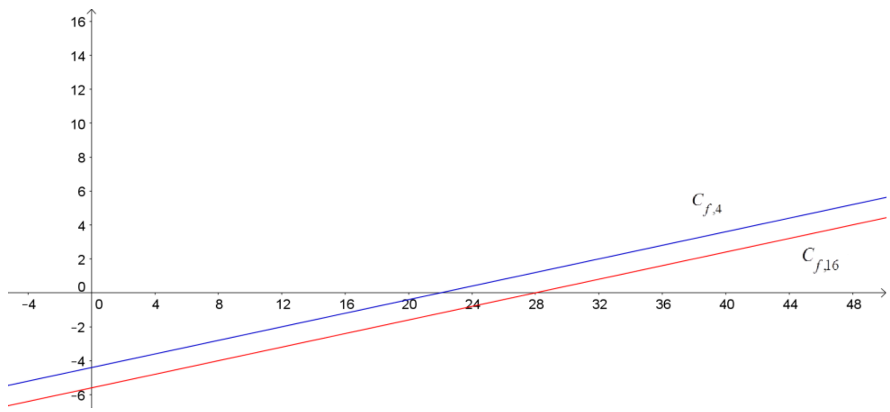


En todos los puntos que forman parte de la recta, $f(x,y)$ toma el valor 4.

¿Y si en lugar del valor 4, nos interesara encontrar aquellos puntos en que la función vale 16? En ese caso, calcularíamos la curva de nivel 16 para la función f:

$$C_{f,10} = \{(x,y) | f(x,y) = 10\} = \{(x,y) | 2x - 10y - 40 = 16\}$$

En este caso, la curva de nivel 10 se trata de una recta de ecuación: $y = 0{,}2x - 5{,}6$; la cual podemos representar junto con la anterior en un nuevo gráfico:



Podemos seguir agregando otras curvas de nivel para reforzar nuestra comprensión del gráfico de la función f. Estas tendrán la siguiente expresión general:

$$C_{f,\alpha} = \{(x, y)|f(x, y) = \alpha\} = \{(x, y)|2x - 10y - 40 = \alpha\}$$

Obtenga las curvas de nivel 32 y de nivel 88 asociadas a la función f, y represéntelas gráficamente. Compruebe que se ubican a la derecha de las ya graficadas.

Se puede deducir, dada la forma general de las curvas de nivel α en este caso, que se tratarán siempre de rectas de pendiente positiva. Además podemos intuir que la función alcanza valores más altos si nos situamos en puntos ubicados en rectas ubicadas más hacia a la derecha. Volveremos más adelante a utilizar las curvas de nivel cuando nos ocupemos de optimizar funciones.

Ejemplo 8. Curvas de indiferencia del consumidor.

Una persona consume dos bienes, lo cual le reporta cierta utilidad o satisfacción. Esta utilidad se expresa a través de la siguiente función (denominada **función de utilidad**):

$$U : U(x, y) = xy$$

donde x es la cantidad consumida del primer bien e y es la cantidad consumida del segundo bien.

Podemos representar gráficamente la función de utilidad a través de las respectivas curvas de nivel:

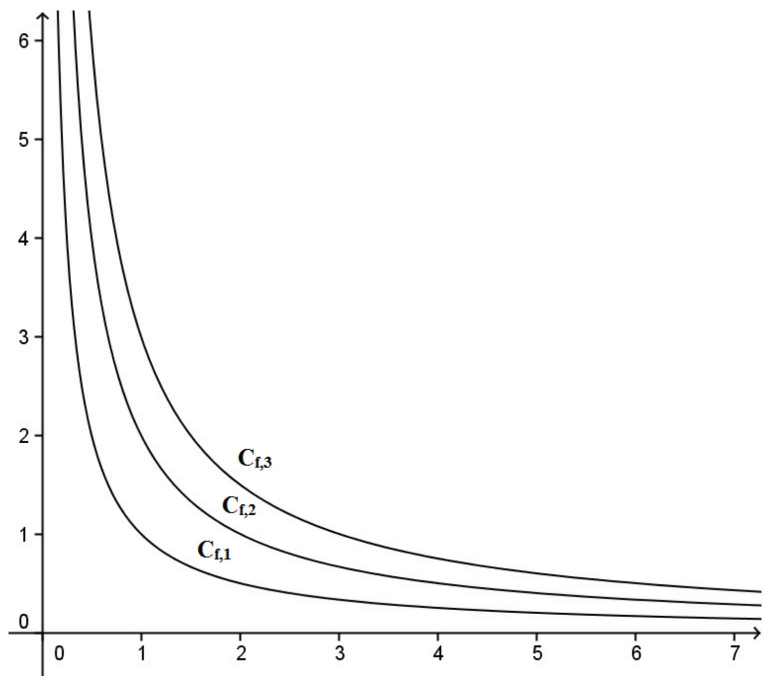
$$C_{U,\alpha} = \{(x, y)|U(x, y) = \alpha\} = \{(x, y)|x \times y = \alpha\}$$

Estas curvas de nivel también reciben el nombre de curvas de indiferencia del consumidor, pues todas las combinaciones de cantidad consumida del primer bien y del segundo bien que conforman la curva de nivel α generan la misma utilidad (α), con lo cual al consumidor le resulta indiferente consumir una combinación u otra.

Consideremos algunas de dichas curvas y representémoslas con el fin de extraer información sobre el crecimiento de la función de utilidad.

$$C_{U,1} = \{(x,y)|U(x,y) = 1\} = \{(x,y)|xy = 1\} = \left\{(x,y)|y = \frac{1}{x}\right\}$$
$$C_{U,2} = \{(x,y)|U(x,y) = 2\} = \{(x,y)|xy = 2\} = \left\{(x,y)|y = \frac{2}{x}\right\}$$
$$C_{U,3} = \{(x,y)|U(x,y) = 3\} = \{(x,y)|xy = 3\} = \left\{(x,y)|y = \frac{3}{x}\right\}$$

Si representamos gráficamente en el sistema de dos ejes cada una de las anteriores curvas de nivel (por ejemplo, para la curva de nivel 1 implicaría representar la ecuación $x \times y$ o su forma equivalente $y = \frac{1}{x}$), obtenemos esto:



Puede apreciarse que valores más elevados de la función se alcanzan en curvas de nivel más alejadas del origen.

El estudio analítico de las funciones de más de una variable puede razonarse en forma análoga a lo realizado para las funciones de una sola variable. Por tal motivo, extenderemos las definiciones más importantes vistas para este caso, para el caso más general de funciones de varias variables.

Definición 6. *Dado un punto $(a,b) \in \mathbb{R}^2$ y un numero $r > 0$, definimos el* ***entorno de centro*** (a,b) ***y radio*** r *como el conjunto de puntos de $\mathbb{R}^2$ cuya distancia al punto (a,b) es menor a r. A este conjunto lo denotaremos como $B_r(a,b)$*

Definición 7. *Dado un conjunto $D \subset \mathbb{R}^2$ y un punto $(a,b) \in D$, decimos que (a,b) es un **punto interior** de D si existe $r > 0$ tal que $B_r(a,b) \subset D$*

Ejemplo:

1. Si $D = \{(x,y) \in \mathbb{R}^2 / x \geq 0, y \geq 0\}$ entonces el punto $(3,2)$ es interior a D pero $(0,2)$ no lo es.

2. Si $D = \mathbb{R}^2$ entonces todos los puntos de D son interiores.

3. Si $D = \{(x,y) \in \mathbb{R}^2 / x = 0\}$ entonces no existen puntos interiores a D.

Definición 8. *Consideremos una función de dos variables $f : D \to \mathbb{R}$ una función de dos variables ($D \subset \mathbb{R}^2$). Se dice que la función f **tiene límite c cuando** (x,y) **tiende al punto** (a,b) si al considerar un entorno $B_r(a,b)$ arbitrariamente chico, la función f toma valores tan cercanos a **c** como se desee dentro de dicho entorno.*
En ese caso diremos que

$$\lim_{(x,y)\to(a,b)} f(x,y) = c$$

Las técnicas para el cálculo de límites de funciones de varias variables son similares a las utilizadas en el caso de funciones de una sola variable.

Definición 9. *Sea una función $f : f(x,y)$.*

- *Se dice que f **es continua en un punto (a,b)** interior a su dominio si:*

$$\lim_{x\to a,\ y\to b} f(x,y) = f(a,b)$$

- *Se dice que f es **continua en un conjunto** D si es continua en cada punto (a,b) de dicho conjunto.*

Nota: Se puede observar que la definición sigue la misma idea que en la continuidad de funciones de una sola variable. De manera informal podemos decir que si una función f es continua en un punto (a,b), al evaluar f en otro punto (x,y) "muy cercano" a (a,b) los valores $f(a,b)$ y $f(x,y)$ son "muy parecidos".

Nota: Durante el curso nosotros nos concentraremos de forma casi exclusiva en funciones continuas. En particular, podemos afirmar que son continuas todas las funciones que se obtienen como sumas y productos de polinomios y exponenciales.

Uno de los ejemplos presentados al inicio del capítulo planteaba una relación entre la Tasa de mortalidad materna y un conjunto de variables [el gasto público en salud (GPS), el porcentaje de partos atendidos por personal capacitado (PAPC), el número de habitantes por médico (HPM), el número de camas de hospital por habitante (CHPH) y el gasto público en educación (GPE)]:

$$TMM = f(GPS, PAPC, HPM, CHPH, GPE)$$

Un análisis completo de esta relación sería interesante e informativo que incluyera cómo el cambio en alguna de las variables independientes afectaría a la variable dependiente. Por ejemplo, cómo un cambio en el gasto público en salud afectaría a la tasa de mortalidad materna. Uno, a priori, esperaría que ese efecto fuese de signo negativo: que un incremento en el gasto público en salud provocase una caída en la tasa de mortalidad materna. Pero, ¿ese efecto sería grande o pequeño? Este tipo de preguntas se vuelve fundamental responderlas para que el gobierno implemente políticas adecuadas, en este caso, en el terreno de la salud maternal.

En el párrafo anterior nos cuestionamos cómo una variación en el gasto público en salud afectaría a la tasa de mortalidad materna. En el contexto de funciones de varias variables no tendría mucho sentido preguntarse cómo una variación en varias variables a la vez afecta a la variable dependiente, sino que es mucho más útil preguntarse por el efecto aislado de una sola de las variables independientes. En otras palabras, no es muy informativo preguntarse cómo una variación a la vez en el gasto público en salud (GPS), en el porcentaje de partos atendidos por personal calificado (PAPC) y en el gasto público en educación (GPE) afectaría a la tasa de mortalidad materna (TMM), pues los efectos de los cambios señalados se mezclarían unos con los otros. Sí tiene sentido intentar aislar el efecto que cada una de las variables independientes tiene sobre la variable dependiente. Esto lo conseguiremos con las denominadas derivadas parciales.

Definición 10. *Sea una función $f : D \to \mathbb{R}$ una función de dos variables ($D \subset \mathbb{R}^2$) y el punto (a, b) interior D. Se denomina **derivada parcial** de la función f respecto a la variable x en el punto (a, b) y se denotará como $\dfrac{\partial f}{\partial x}(a, b)$, al resultado del siguiente límite, siempre y cuando exista y sea finito:*

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

La definición, y por ende su interpretación, es análoga a la de derivada de una función de una sola variable, en la cual se calculó el límite del cociente incremental cuando el incremento en la variable independiente (h) tendía a 0.

Observación: Algo importante a destacar es que la única variable que se hace variar es x (o y) quedando la otra variable constante.

La ***derivada parcial*** de la función f respecto a la variable y en el punto (a, b) es análogo al caso anterior, y se obtiene del resultado del siguiente límite, siempre y cuando exista y sea finito:

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

Nota:

$$\frac{\partial f}{\partial x}(a,b) \text{ también se puede escribir como: } f_x(a,b)$$
$$\frac{\partial f}{\partial y}(a,b) \text{ también se puede escribir como: } f_y(a,b)$$

Como se vio anteriormente al analizar funciones de una variable, el valor de la derivada en un punto puede ser utilizado como "aproximación"del incremento de la función. Lo mismo vale para funciones de varias variables:

- La derivada parcial f_x puede servir para aproximar el cambio de f, cuando la variable x aumenta en una unidad y la variable y se mantiene constante.

$$\frac{\partial f}{\partial x}(x,y) \approx f(x+1,y) - f(x,y)$$

- La derivada parcial f_y puede servir para aproximar el cambio de f, cuando la variable y aumenta en una unidad y la variable x se mantiene constante.

$$\frac{\partial f}{\partial y}(x,y) \approx f(x,y+1) - f(x,y)$$

Ejemplo 12: Anteriormente se señaló que sería interesante estudiar cómo afectaría a la tasa de mortalidad materna un cambio en alguna de las variables de las cuales dependía:

$$\begin{aligned} TMM &= f(GPS, PAPC, HPM, CHPH, GPE) \\ &= -3 \times GPS - 4 \times PAPC + 2 \times HPM - 5 \times CHPH - 6 \times GPE \end{aligned}$$

A su vez en el ejemplo 1 se consideró un cierto país A, para el cual se calculó el valor de la tasa de mortalidad materna, dados los valores de las variables independientes; es decir se calculó el valor de la función f en el punto (4,90,300,0.02,5):

$$TMM_{\text{País A}} = f(4, 90, 300, 0{,}02, 5) = -3(4) - 4(90) + 2(300) - 5(0{,}02) - 6(5) = 197{,}9$$

Intentaremos ahora calcular cómo se vería afectada la tasa de mortalidad materna ante una variación en una unidad en el gasto público social, lo cual es precisamente lo que nos reporta la derivada parcial de la función con respecto a la variable GPS en el punto (4,90,300,0.02,5).

Siguiendo la definición 3, deberemos calcular el límite del cociente incremental correspondiente:

$$\frac{\partial f}{\partial GPS}(4, 90, 300, 0{,}02, 5) = \lim_{h \to 0} \frac{f(4+h, 90, 300, 0{,}02, 5) - f(4, 90, 300, 0{,}02, 5)}{h} =$$

$$= \lim_{h \to 0} \frac{-3(4+h) - 4(90) + 2(300) - 5(0{,}02) - 6(5) - [-3(4) - 4(90) + 2(300) - 5(0{,}02) - 6(5)]}{h} =$$

$$= \lim_{h \to 0} \frac{-3h}{h} = -3$$

El resultado final entonces es $\frac{\partial f}{\partial GPS}(4, 90, 300, 0{,}02, 5) = -3$, el cual se interpreta así: dados los valores iniciales de las variables independientes, por cada unidad que se incremente el gasto público en salud (GPS), dejando las demás variables constantes, la tasa de mortalidad materna (TMM) se reducirá aproximadamente en 3 unidades.

Si queremos saber el efecto que una variación en el porcentaje de partos atendidos por personal calificado (PAPC), segunda variable independiente, tiene sobre la tasa de mortalidad materna, deberemos calcular la derivada parcial de f con respecto a dicha variable en el punto (4,90,300,0.02,5):

$$\frac{\partial f}{\partial PAPC}(4, 90, 300, 0{,}02, 5) = \lim_{h\to 0} \frac{f(4, 90+h, 300, 0{,}02, 5) - f(4, 90, 300, 0{,}02, 5)}{h} =$$

$$= \lim_{h\to 0} \frac{-3(4) - 4(90+h) + 2(300) - 5(0{,}02) - 6(5) - [-3(4) - 4(90) + 2(300) - 5(0{,}02) - 6(5)]}{h} =$$

$$= \lim_{h\to 0} \frac{-4h}{h} = -4$$

En este caso, $\frac{\partial f}{\partial PAPC}(4, 90, 300, 0{,}02, 5) = -4$, lo cual nos quiere decir que, dados los valores iniciales de las variables independientes, por cada unidad que se incremente el porcentaje de partos atendidos por personal calificado (PAPC), dejando las demás variables constantes, la tasa de mortalidad materna (TMM) se reducirá aproximadamente en 4 unidades.

Siguiendo el mismo razonamiento podemos calcular el efecto que una variación en cada una de las restantes variables independientes provoca en la tasa de mortalidad materna.

De manera similar a como definíamos en funciones de una sola variable la función derivada, podemos hacerlo en este caso para funciones de derivadas parciales.

Definición 11. *Las derivadas parciales de una función f son funciones $\dfrac{\partial f}{\partial x}$ y $\dfrac{\partial f}{\partial x}$ definidas de la siguiente forma:*

$$\frac{\partial f}{\partial x}(x, y) = \lim_{h\to 0} \frac{f(x+h, y) - f(x, y)}{h}$$
$$\frac{\partial f}{\partial y}(x, y) = \lim_{h\to 0} \frac{f(x, y+h) - f(x, y)}{h}$$

Para realizar dicho cálculo pueden aplicarse las reglas de derivación ya vistas en el caso de derivadas de funciones de una sola variable.

Ejemplo 13: Sea la función $f : f(x, y) = 5x^2 - 3y^2$. Calcularemos la derivada parcial de la función f respecto a la variable x $\left[\frac{\partial f}{\partial x}(x, y)\right]$ y la derivada parcial de f respecto a la variable y $\left[\frac{\partial f}{\partial y}(x, y)\right]$.

$\frac{\partial f}{\partial x}(x, y) = 10x$ Para obtener este resultado consideramos "y" como una constante (pues la derivada parcial es con respecto a x) y aplicamos las reglas de derivación ya empleadas anteriormente.

$\frac{\partial f}{\partial y}(x,y) = -6y$ En este caso, dado que calculamos la derivada parcial con respecto a y, consideramos como constante "x".

Ejemplo 14: Sea la función $f : f(x,y) = 4x^3 - 2xy^2 + 6y^3$. Calcular sus derivadas parciales.

$$\frac{\partial f}{\partial x}(x,y) = 12x^2 - 2y^2$$
$$\frac{\partial f}{\partial y}(x,y) = -4xy + 18y^2$$

Ejemplo 15: Sea la función $f : f(x,y) = e^{x^2+y^4}$. Calcular sus derivadas parciales.

$$\frac{\partial f}{\partial x}(x,y) = 2xe^{x^2+y^4}$$
$$\frac{\partial f}{\partial y}(x,y) = 4y^3e^{x^2+y^4}$$

Las derivadas parciales que hemos definido anteriormente se denominan también derivadas parciales de primer orden, y son equivalentes a la derivada primera en las funciones de una sola variable. Igual que en el caso de las funciones de una sola variable, se pueden determinar derivadas de segundo orden para las funciones bivariadas. Cada función $f(x,y)$ tendrá a lo sumo cuatro diferentes **derivadas de segundo orden**.

Las cuatro derivadas parciales segundas son: $\dfrac{\partial^2 f}{\partial x \partial x}$, $\dfrac{\partial^2 f}{\partial y \partial y}$, $\dfrac{\partial^2 f}{\partial x \partial y}$ y $\dfrac{\partial^2 f}{\partial y \partial x}$.

La derivada parcial $\dfrac{\partial^2 f}{\partial x \partial x}$ se calcula obteniendo primero $\dfrac{\partial f}{\partial x}$ y luego diferenciando nuevamente respecto de x. De manera análoga, $\dfrac{\partial^2 f}{\partial x \partial y}$ se obtiene determinando la expresión $\dfrac{\partial f}{\partial x}$ y luego diferenciando respecto de y.
Por economía, frecuentemente abreviaremos la notación de la siguiente manera
Notación:

- $\dfrac{\partial^2 f}{\partial x \partial x} = f_{xx}$
- $\dfrac{\partial^2 f}{\partial x \partial y} = f_{xy}$
- $\dfrac{\partial^2 f}{\partial y \partial x} = f_{yx}$
- $\dfrac{\partial^2 f}{\partial y \partial y} = f_{yy}$

Ejemplo 17: Retomamos el ejemplo 14, en que se trabajó con $f : f(x, y) = 4x^3 - 2xy^2 + 6y^3$.

En dicho ejemplo se halló que las derivadas parciales de primer orden la función f eran:

$$\frac{\partial f}{\partial x}(x, y) = 12x^2 - 2y^2$$
$$\frac{\partial f}{\partial y}(x, y) = -4xy + 18y^2$$

Derivando cada una de las derivadas parciales con respecto a la variable x y a la variable y, obtendremos las derivadas parciales de segundo orden. De esta manera, obtendremos en este caso cuatro derivadas parciales de segundo orden.

$$\frac{\partial^2 f}{\partial x \partial x}(x, y) = 24x \qquad \frac{\partial^2 f}{\partial x \partial y}(x, y) = -4y$$
$$\frac{\partial^2 f}{\partial y \partial x}(x, y) = -4y \qquad \frac{\partial^2 f}{\partial y \partial y}(x, y) = -4x + 36y$$

Observación: En este caso $\frac{\partial^2 f}{\partial x \partial y}(x, y) = \frac{\partial^2 f}{\partial y \partial x}(x, y) = -4y$. Este no es un resultado casual. Si las derivadas parciales de segundo orden son funciones continuas, se cumple que $\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x}$. En este curso trabajaremos siempre con funciones que cumplan esta condición.

Definición 12. *Se denomina* ***matriz hessiana*** *de una función f una matriz $H_f(x, y)$ cuyos coeficientes son las derivadas parciales de segundo orden de f ordenados de la siguiente manera:*

$$H_f(x, y) = \begin{pmatrix} \frac{\partial^2 f}{\partial x \partial x}(x, y) & \frac{\partial^2 f}{\partial x \partial y}(x, y) \\ \frac{\partial^2 f}{\partial y \partial x}(x, y) & \frac{\partial^2 f}{\partial y \partial y}(x, y) \end{pmatrix}$$

Dado que en el presente curso trabajaremos siempre con funciones que tengan derivadas de segundo orden continuas, la matriz hessiana siempre será una matriz simétrica, pues $\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x}$.

Ejemplo 18: En el caso de la función del ejemplo 17, su matriz hessiana es:

$$H_f(x, y) = \begin{pmatrix} 24x & -4y \\ -4y & -4x + 36y \end{pmatrix}$$

5.2. Optimización de las funciones de dos variables (sin restricciones)

Uno de los aspectos más interesantes del estudio de funciones de una variable, como vimos en el capítulo 3, es la optimización de las mismas, es decir la búsqueda de sus valores máximo y mínimo. De la misma manera es sumamente útil la optimización de funciones de varias variables, en lo cual nos enfocaremos en la presente sección. Como paso previo al hallazgo de los extremos (máximo y mínimo) absolutos, es preciso buscar los extremos (máximo y mínimo) relativos.

Definición 13. *Sea $f : D \to \mathbb{R}$ una función de dos variables ($D \subset \mathbb{R}^2$) y un punto $(a, b) \in D$.*

- *Decimos que f alcanza un **máximo relativo** en (a, b) si existe $r > 0$ de forma que $B_r \subset D$ y se cumple que $f(x, y) \leq f(a, b)$ para todo $(x, y) \in B_r(a, b)$*
- *Decimos que f alcanza un **mínimo relativo** en (a, b) si existe $r > 0$ de forma que $B_r \subset D$ y se cumple que $f(x, y) \geq f(a, b)$ para todo $(x, y) \in B_r(a, b)$*

¿Cómo pueden identificarse los extremos relativos de una función? A partir de la siguiente propiedad.

Definición 14. *Sea $f : D \to \mathbb{R}$ una función de dos variables ($D \subset \mathbb{R}^2$) y sea (a, b) un punto interior a su dominio. Si todas las derivadas parciales de f en el punto (a, b) valen 0, se dice que (a, b) es un **punto estacionario** de la función f.*

Propiedad 1. Condición necesaria de existencia de extremos relativos.

Sea $f : D \to \mathbb{R}$ una función de dos variables ($D \subset \mathbb{R}^2$) y (a, b) un punto interior de D donde existen ambas derivadas parciales f_x y f_y.
Si f alcanza un extremo relativo en (a, b) entonces (a, b) es un punto estacionario de f.

Definición 15. *Sea $f : D \to \mathbb{R}$ una función de dos variables ($D \subset \mathbb{R}^2$) y (a, b) un punto estacionario de f. Si f no alcanza un extremo relativo en (a, b), decimos que este es un **punto silla** de f.*

Ejemplo 19: Hallar todos los puntos estacionarios de la función $f : f(x, y) = x^2 + y^2$.

Para hallar los puntos estacionarios, debemos resolver el siguiente sistema de ecuaciones:

$$\begin{cases} \frac{\partial f}{\partial x}(x,y) & = 0 \\ \\ \frac{\partial f}{\partial y}(x,y) & = 0 \end{cases} \Leftrightarrow \begin{cases} 2x & = 0 \\ 2y & = 0 \end{cases} \Rightarrow x = y = 0$$

La función f tiene un único punto estacionario, que es (0,0).

Ejemplo 20: Hallar todos los puntos estacionarios y críticos de la función f tal que: $f : f(x, y) = 2x^3 + xy^2 - 24x + y^2$.

Resolvemos entonces el siguiente sistema:

$$\begin{cases} \frac{\partial f}{\partial x}(x,y) &= 0 \\ \\ \frac{\partial f}{\partial y}(x,y) &= 0 \end{cases} \Leftrightarrow \begin{cases} 6x^2+y^2-24 &= 0 \\ 2xy+2y &= 0 \end{cases} \Rightarrow \begin{cases} 6x^2+y^2-24 &= 0 \\ 2y(x+1) &= 0 \end{cases}$$

Para que la última ecuación se verifique, basta que $x=-1$, o que $y=0$, con lo cual debemos considerar dos sistemas:

$$\begin{cases} 6x^2+y^2-24 &= 0 \\ y &= 0 \end{cases} \Rightarrow y=0 \Rightarrow 6x^2+0^2-24=0 \Rightarrow x^2=4 \Rightarrow x=\pm 2$$

De lo cual concluimos que (2,0) y (-2,0) son puntos estacionarios de la función f.

El otro sistema a considerar es:

$$\begin{cases} 6x^2+y^2-24 &= 0 \\ x &= -1 \end{cases} \Rightarrow x=-1 \Rightarrow 6(-1)^2+y^2-24=0 \Rightarrow y^2=18 \Rightarrow y=\pm\sqrt{18}$$

De lo cual concluimos que (-1,$\sqrt{18}$) y (-1,-$\sqrt{18}$) también son puntos estacionarios de la función f.

En total hemos hallado cuatro puntos estacionarios de la función f: (2,0), (-2,0), (-1,$\sqrt{18}$) y (-1,-$\sqrt{18}$).

Nos queda pendiente ahora, identificar si los puntos estacionarios hallados son puntos de extremos relativos o puntos de silla, es decir, clasificar dichos puntos. Para ello aplicaremos la siguiente regla.

Propiedad 2. Clasificación de puntos estacionarios. Sea f una función de dos variables, con derivadas parciales de segundo orden continuas. Sea (a,b) un punto estacionario de f.

Entonces se puede asegurar que:

- Si $\det[H]_f(a,b)>0$ y $\frac{\partial^2 f}{\partial x \partial x}(a,b)>0$, entonces f en (a,b) tiene un mínimo relativo.
- Si $\det[H]_f(a,b)>0$ y $\frac{\partial^2 f}{\partial x \partial x}(a,b)<0$, entonces f en (a,b) tiene un máximo relativo.
- Si $\det[H]_f(a,b)<0$, entonces f en (a,b) tiene un punto de silla.

donde $[H]_f(a,b)$ es la matriz hessiana de la función f evaluada en el punto (a,b).

Notas:

1. Si $\det[H]_f(a,b)=0$, entonces no se puede saber por aplicación de esta regla qué ocurre en (a,b) y se requieren otros procedimientos para seguir analizando, los cuales exceden el ámbito de este curso.

2. La regla solo es válida para funciones de dos variables. Para funciones de más variables, otra regla más general se hace necesaria, la cual no trataremos en este curso.

Ejemplo 21: Clasificar el punto estacionario de la función $f : f(x,y) = x^2 + y^2$ hallado en el ejemplo 19.

En el ejemplo 19 se halló que la función f tenía un único punto estacionario, que era (0,0). Ahora procederemos a clasificarlo, es decir, a analizar si es un punto de extremo relativo o un punto de silla, para lo cual aplicaremos la propiedad 2 (nótese que f es una función de dos variables, por lo cual aplicar tal propiedad está permitido).

Dado que la propiedad referida se basa en el cálculo del determinante de la matriz hessiana, debemos calcular antes las derivadas parciales de segundo orden de f:

$$\begin{aligned}
&\frac{\partial^2 f}{\partial x \partial x}(x,y) = 2 && \Rightarrow \frac{\partial^2 f}{\partial x \partial x}(0,0) = 2 \\
&\frac{\partial^2 f}{\partial x \partial y}(x,y) = \frac{\partial^2 f}{\partial y \partial x}(x,y) = 0 && \Rightarrow \frac{\partial^2 f}{\partial x \partial y}(0,0) = \frac{\partial^2 f}{\partial y \partial x}(0,0) = 0 \\
&\frac{\partial^2 f}{\partial y \partial y}(x,y) = 2 && \Rightarrow \frac{\partial^2 f}{\partial y \partial y}(0,0) = 2
\end{aligned}$$

$$H_f(0,0) = \begin{pmatrix} \frac{\partial^2 f}{\partial x \partial x}(0,0) & \frac{\partial^2 f}{\partial x \partial y}(0,0) \\ \frac{\partial^2 f}{\partial y \partial x}(0,0) & \frac{\partial^2 f}{\partial y \partial y}(0,0) \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$$

Entonces: $\det[H]_f(0,0) = 4 > 0$ y $\frac{\partial^2 f}{\partial x \partial x}(0,0) = 2 > 0$, por lo tanto, la función f tiene en $(0,0)$ un mínimo relativo.

Ejemplo 22: Clasificar los puntos estacionarios de la función $f : f(x,y) = 2x^3 + xy^2 - 24x + y^2$ hallados en el ejemplo 20.

Nótese que nuevamente se trata de una función de dos variables, por lo cual estamos habilitados a aplicar la propiedad 2. Calculamos entonces en primer lugar las derivadas parciales de segundo orden.

$$\begin{aligned}
&\frac{\partial^2 f}{\partial x \partial x}(x,y) = 12x \\
&\frac{\partial^2 f}{\partial x \partial y}(x,y) = \frac{\partial^2 f}{\partial y \partial x}(x,y) = 2y \\
&\frac{\partial^2 f}{\partial y \partial y}(x,y) = 2x + 2
\end{aligned}$$

La matriz hessiana de f queda entonces:

$$H_f(x,y) = \begin{pmatrix} \dfrac{\partial^2 f}{\partial x \partial x}(x,y) & \dfrac{\partial^2 f}{\partial x \partial y}(x,y) \\ \dfrac{\partial^2 f}{\partial y \partial x}(x,y) & \dfrac{\partial^2 f}{\partial y \partial y}(x,y) \end{pmatrix} = \begin{pmatrix} 12x & 2y \\ 2y & 2x+2 \end{pmatrix}$$

Clasificaremos uno por uno los cuatro puntos estacionarios de f hallados: (2,0), (-2,0), (0,2) y (0,-2).

Clasificación del punto estacionario (2,0):

$H_f(2,0) = \begin{pmatrix} 24 & 0 \\ 0 & 6 \end{pmatrix} \Rightarrow \det[H]_f(2,0) = 144 > 0$ y $\frac{\partial^2 f}{\partial x \partial x}(2,0) = 24 > 0$. Entonces (2,0) es un punto de **mínimo relativo** de f.

Clasificación del punto estacionario (-2,0): $H_f(-2,0) = \begin{pmatrix} -24 & 0 \\ 0 & -2 \end{pmatrix} \Rightarrow \det[H]_f(-2,0) = 48 > 0$ y $\frac{\partial^2 f}{\partial x \partial x}(2,0) = -24 < 0$. Entonces (-2,0) es un punto de **máximo relativo** de f.

Clasificación del punto estacionario $(-1,\sqrt{18})$**:**
$H_f(-1,\sqrt{18}) = \begin{pmatrix} -12 & 2\sqrt{18} \\ 2\sqrt{18} & 0 \end{pmatrix} \Rightarrow \det[H]_f(-1,\sqrt{18}) = -72 < 0$. Entonces $(-1,\sqrt{18})$ es un **punto de silla** de f.

Clasificación del punto estacionario $(-1,-\sqrt{18})$**:**
$H_f(-1,-\sqrt{18}) = \begin{pmatrix} -12 & -2\sqrt{18} \\ -2\sqrt{18} & 0 \end{pmatrix} \Rightarrow \det[H]_f(-1,-\sqrt{18}) = -72 < 0$. Entonces $(-1,-\sqrt{18})$ es un **punto de silla** de f.

5.3. Optimización sujeta a restricciones

El análisis de los métodos de optimización basados en el cálculo se centró en la optimización no restringida. En muchas aplicaciones del modelado matemático interviene la optimización de una función objetivo sujeta a ciertas condiciones restrictivas, o simplemente restricciones. Estas restricciones representan limitaciones capaces de influir en el grado que se optimizan las funciones objetivo. Y pueden reflejar limitaciones como escasez de recursos (por ejemplo, mano de obra, materiales o capital), poca demanda de productos, metas de ventas, etc. Típicamente, estos problemas tienen la forma

$$max\ f(x,y) \text{ sujeto a } g(x,y) = c \ (*)$$

o alternativamente

$$min\ f(x,y) \text{ sujeto a } g(x,y) = c \ (**)$$

Los problemas presentan esta estructura se consideran problemas de optimización restringida. En esta sección nos ocuparemos del *método de sustitución* para resolver ciertos problemas de optimización no lineal restringida.

Método de substitución

Presentaremos este método a través de un ejemplo: Una compañía ha recibido una orden de 200 unidades para uno de sus productos. El pedido será surtido con la producción combinada de sus dos plantas. La función conjunta de costo de la fabricación de este producto es:

$$C = f(q_1, q_2) = 2q_1^2 + q_1 q_2 + q_2^2 + 500$$

donde q_1 y q_2 son las cantidades producidas en las plantas 1 y 2, respectivamente. Si el objetivo es minimizar los costos totales, sujeto a la condición de suministrar 200 unidades procedentes de ambas plantas, ¿qué cantidades deberá proporcionar cada una?

En definitiva, el problema puede plantearse como

$$min\ f(q_1, q_2) \text{ sujeto a } q_1 + q_2 = 200\ (***)$$

En este caso, de la restricción de igualdad tenemos que $q_2 = 200 - q_1$. Observemos ademas que existe una restricción implícita en la letra: las magnitudes q_1 y q_2 tienen que ser positivas, lo que implica que $0 \leq q_1 \leq 200$. Sustituyendo en la función de costos obtenemos una nueva función h de una sola variable:

$$h : [0, 200] \longrightarrow \mathbb{R} / h(q_1) = f(q_1, q_2) = 2q_1^2 + q_1(200 - q_1) + (200 - q_1)^2 + 500$$

De esta forma hemos reducido el problema $(***)$ a un problema de optimización en una variable: determinar el mínimo absoluto de h en el intervalo $[0, 200]$

En resumen, esta estrategia de resolución se basa en despejar una de la variables en función de la otra (a partir de la restricción de igualdad) y substituir dicha expresión en la función a optimizar. De esta forma el problema quedará reducido a optimizar una función de una sola variable.