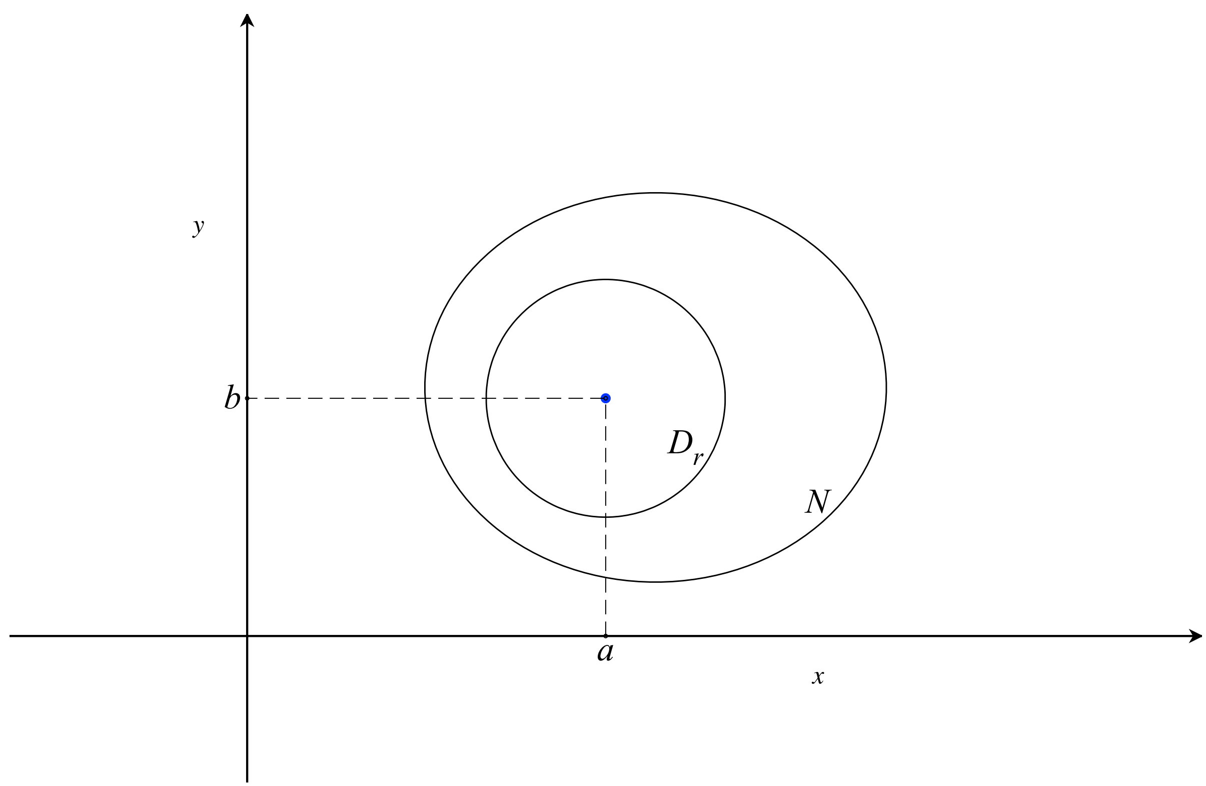
Optimization

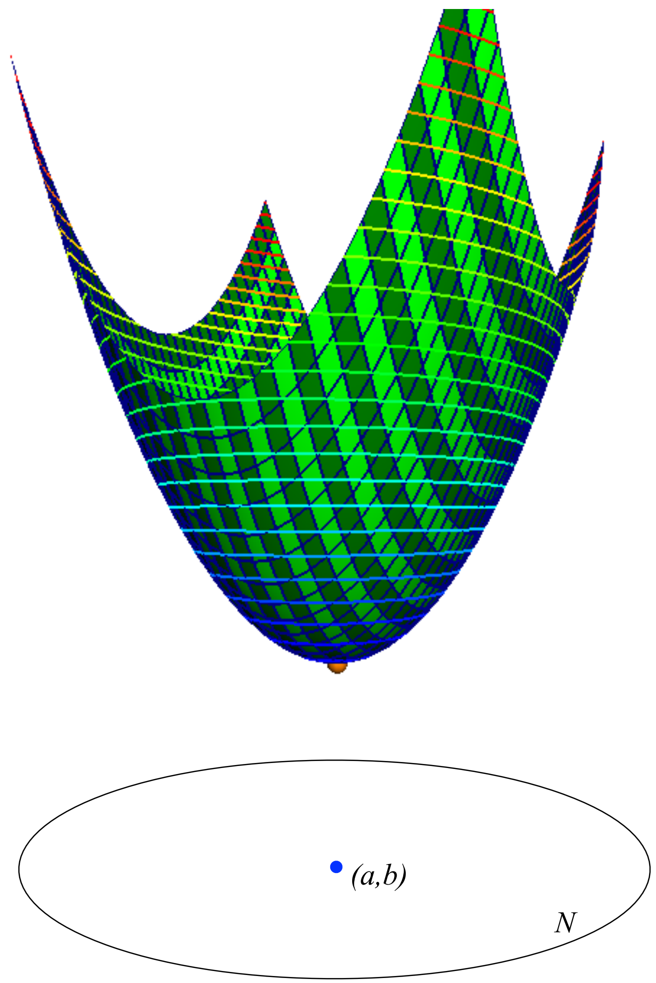
Definition: If (a, b) is a point in the plane, then an open disk $D_r(a, b)$ (of radius $r > 0$) centered at (a, b) is a set of the form

$$D_r(a, b) = \{(x, y) : \sqrt{(x-a)^2 + (y-b)^2} < r\}$$

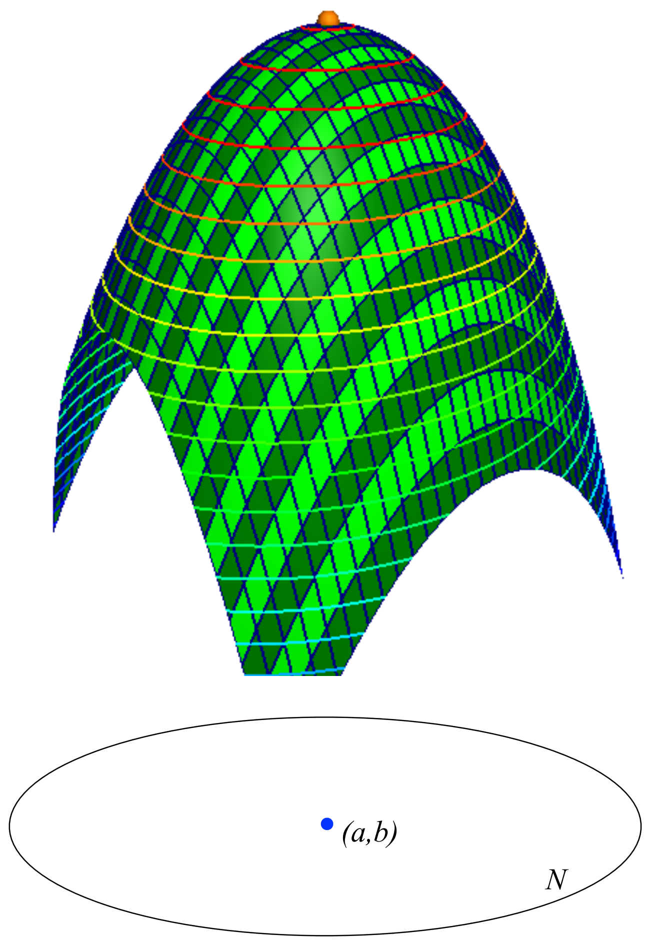
A ***neighborhood*** N of (a, b) is any set that contains an open disk $D_r(a, b)$ centered at (a, b).



Definition: $f(a,b)$ is a ***relative minimum*** value of the function $z = f(x,y)$ if $f(a,b) \le f(x,y)$ for *all* points (x,y) in *some neighborhood* N of (a,b).



Definition: $f(a,b)$ is a ***relative maximum*** value of the function $z = f(x,y)$ if $f(a,b) \geq f(x,y)$ for *all* points (x,y) in *some neighborhood* N of (a,b).



Key Fact: If $f(a,b)$ is a relative minimum or relative maximum value and if $f(x,y)$ is differentiable (in a neighborhood of (a,b)), then

$$f_x(a,b) = 0 \qquad \text{and} \qquad f_y(a,b) = 0.$$

Definition: If $f_x(a,b) = 0$ and $f_y(a,b) = 0$, then (a,b) is a ***critical point*** (or *stationary point*) of $f(x,y)$ and $f(a,b)$ is a ***critical value***.

Restating key fact: *If $f(x,y)$ is differentiable, then its relative extreme values can **only occur at critical points**.*

Explanation: If (x,y) is close to (a,b), then

$$f(x,y) \approx T_1(x,y) = f(a,b) + f_x(a,b)(x-a) + f_y(a,b)(y-b).$$

If $f_x(a,b) \neq 0$, $y = b$ and $x \approx a$, then

$$\begin{aligned} f(x,b) \approx T_1(x,b) &= f(a,b) + f_x(a,b)(x-a) + f_y(a,b)(b-b) \\ &= f(a,b) + f_x(a,b)(x-a), \end{aligned}$$

so $\boxed{f(x,b) - f(a,b) \approx f_x(a,b)(x-a)}$.

Case 1. $f_x(a,b) > 0$. If $x > a$, then $x - a > 0$ so

$$f(x,b) - f(a,b) \approx \overbrace{f_x(a,b)}^{+}\overbrace{(x-a)}^{+} > 0,$$

which means that $f(a,b)$ is ***not*** a maximum value.

If $x < a$, then $x - a < 0$ and

$$f(x,b) - f(a,b) \approx \overbrace{f_x(a,b)}^{+}\overbrace{(x-a)}^{-} < 0,$$

so $f(a,b)$ is ***not*** *a minimum value.*

Case 2. $f_x(a,b) < 0$.

If $x > a$, then $x - a > 0$ and

$$f(x,b) - f(a,b) \approx \overbrace{f_x(a,b)}^{-}\overbrace{(x-a)}^{+} < 0,$$

so $f(a,b)$ is *not a minimum value.*

If $x < a$, then $x - a < 0$ and

$$f(x,b) - f(a,b) \approx \overbrace{f_x(a,b)}^{-}\overbrace{(x-a)}^{-} > 0,$$

so $f(a,b)$ is *not a maximum value.*

If $f_y(a,b) \neq 0$, then the analogous argument with $x = a$ and $y \approx b$ shows that $f(a,b)$ is neither a maximum nor a minimum value.

Conclusion: If $f_x(a,b) \neq 0$, then $f(a,b)$ is ***not*** a relative extreme value. The same argument shows that if $f_y(a,b) \neq 0$, then $f(a,b)$ is ***not*** a relative extreme value.

Therefore, if $f(a,b)$ ***is*** a relative extreme value, then $f_x(a,b)$ and $f_y(a,b)$ ***must both be 0***.

Example: Find the critical point(s) and critical values of the function

$$f(x,y) = x^2 + y^2 - xy + x^3$$

1. First order conditions:

$$f_x = 0 \implies 2x - y + 3x^2 = 0$$

$$f_y = 0 \implies 2y - x = 0$$

2. Critical points: $f_y = 0 \implies x = 2y$ and substituting $2y$ for x in the first equation gives

$$2x - y + 3x^2 = 0 \implies \overbrace{4y}^{2\cdot 2y} - y + \overbrace{12y^2}^{3(2y)^2} = 0 \implies 3y(1+4y) = 0.$$

The critical y-values are $y_1 = 0$ and $y_2 = -1/4$. Remember that at the critical points $x = 2y$, and therefore the critical points are

$$(x_1, y_1) = (0,0) \text{ and } (x_2, y_2) = (-1/2, -1/4).$$

3. Critical values: $f(0,0) = 0$ and $f(-1/2, -1/4) = \dfrac{1}{16}$.

Observation: The definitions of ***relative extreme values*** and of ***critical points*** generalize to functions of any number of variables, as does the connection between relative extreme values and critical points...

Definition: The point $(x_1, \ldots, x_k)$ is ***close to*** the point $(a_1, \ldots, a_k)$ if

$$x_1 \approx a_1,\ x_2 \approx a_2,\ \ldots,\ x_{k-1} \approx a_{k-1} \text{ and } x_k \approx a_k.$$

Definition: $f(a_1, \ldots, a_k)$ is a relative maximum (minimum) value of the function $y = f(x_1, \ldots, x_k)$ if

$$f(a_1, \ldots, a_k) \geq f(x_1, \ldots, x_k) \qquad \big(\, f(a_1, \ldots, a_k) \leq f(x_1, \ldots, x_k)\,\big)$$

for all points $(x_1, \ldots, x_k)$ that are ***sufficiently close*** to $(a_1, \ldots, a_k)$.

Definition: The point $(a_1, \ldots, a_k)$ is a ***critical point*** of the function $f(x_1, \ldots, x_k)$ if

$$f_{x_1}(a_1, \ldots, a_k) = 0,\ f_{x_2}(a_1, \ldots, a_k) = 0,\ \ldots \text{ and } f_{x_k}(a_1, \ldots, a_k) = 0$$

Fact: *If $f(x_1, \ldots, x_k)$ is differentiable and $f(a_1, \ldots, a_k)$ is a relative extreme value, then $(a_1, \ldots, a_k)$ is a critical point of $f(x_1, \ldots, x_k)$.*

Conclusion: To find the relative extreme values of a differentiable function $f(x_1, \ldots, x_k)$, we need to find its critical points. To do this, we need to solve the system of k equations in k variables:

$$\begin{aligned} f_{x_1}(x_1, \ldots, x_k) &= 0 \\ f_{x_2}(x_1, \ldots, x_k) &= 0 \\ \vdots \qquad\qquad & \quad \vdots \\ f_{x_k}(x_1, \ldots, x_k) &= 0 \end{aligned}$$

These equations are called the *first order conditions for relative extrema.*

Example: Find the critical point(s) of the function

$$w = x^2 + 2y^2 - 3z^2 + xy - 2xz + yz + 2x - 3y - 2z + 1$$

First order conditions:

$$\begin{array}{llcr} (1) & w_x = 2x + y - 2z + 2 = 0 & \Longrightarrow & 2x + y - 2z = -2 \\ (2) & w_y = 4y + x + z - 3 = 0 & \Longrightarrow & x + 4y + z = 3 \\ (3) & w_z = -6z - 2x + y - 2 = 0 & \Longrightarrow & -2x + y - 6z = 2 \end{array}$$

If we add equation (1) to equation (3) (eliminating the x s) we get

$$(4) \qquad 2y - 8z = 0.$$

Adding 2× equation (2) to equation (3) (eliminating the x s again) gives

$$(5) \qquad 9y - 4z = 8.$$

From equation (4) it follows that $y = 4z$, and substituting $y = 4z$ into equation (5) gives...

$$36z - 4z = 8 \implies z^* = \frac{8}{32} = \frac{1}{4}$$

which implies that $y^* = 4z^* = 1$.

Finally plugging $y^* = 1$ and $z^* = 1/4$ back into equation (2) we find that

$$x + 4 + \frac{1}{4} = 3 \implies x^* = -\frac{5}{4},$$

so there is only one critical point,

$$(x^*, y^*, z^*) = (-5/4, 1, 1/4)$$

and the critical value is

$$w^* = w(x^*, y^*, z^*) = w(-5/4, 1, 1/4) = 2$$

Example: Find the critical point(s) and critical value(s) of the function

$$F(u, v, w, \lambda) = 5 \ln u + 8 \ln v + 12 \ln w - \lambda(10u + 15v + 25w - 3750).$$

First order conditions:

$$
\begin{array}{lll}
(1) & F_u = 0 \implies & \dfrac{5}{u} - 10\lambda = 0 \\
(2) & F_v = 0 \implies & \dfrac{8}{v} - 15\lambda = 0 \\
(3) & F_w = 0 \implies & \dfrac{12}{w} - 25\lambda = 0 \\
(4) & F_\lambda = 0 \implies & -(10u + 15v + 25w - 3750) = 0
\end{array}
$$

Equation (1) implies that

$$\frac{5}{u} = 10\lambda \implies \boxed{\lambda = \frac{1}{2u}}$$

Likewise, equations (2) and (3) imply that

$$\frac{8}{v} = 15\lambda \implies \boxed{\lambda = \frac{8}{15v}} \quad \text{and} \quad \frac{12}{w} = 25\lambda \implies \boxed{\lambda = \frac{12}{25w}}.$$

Comparing the first and second boxed equations shows that

$$\lambda = \frac{1}{2u} = \frac{8}{15v} \implies 15v = 16u \implies v = \frac{16u}{15}$$

and comparing the first and third boxed equations show that

$$\lambda = \frac{1}{2u} = \frac{12}{25w} \implies 25w = 24u \implies w = \frac{24u}{25}.$$

Equation (4) simplifies

$$-(10u + 15v + 25w - 3750) = 0 \implies 10u + 15v + 25w - 3750 = 0$$

$$\implies 10u + 15v + 25w = 3750$$

and substituting for v and w in this equation gives

$$10u + 15 \cdot \frac{16u}{15} + 25 \cdot \frac{24u}{25} = 3750 \implies 50u = 3750 \implies u^* = 75.$$

It follows that $v^* = \frac{16}{15}u^* = 80$, $w^* = \frac{24}{25}u^* = 72$ and $\lambda^* = \frac{1}{2u^*} = \frac{1}{150}$. I.e., the critical point is $(u^*, v^*, w^*, \lambda^*) = (75, 80, 72, 1/150)$ and the critical value is

$$F(75, 80, 72, 1/150) \approx 107.964.$$