(*) **Optimization - introduction**

Our focus for the rest of the quarter is *optimization in several variables*, which means finding the maximum and/or minimum values of functions of several variables. We will study two flavors of this problem:

Unconstrained optimization: '*Find the maximum/minimum value(s) of the function* $z = f(x, y)$'. This is an *un*constrained problem because there are no additional constraints (restrictions) on the variables x and y (besides providing an optimal value).

E.g., if $P(p_1, p_2)$ is a monopolistic firm's profit function, where p_1 and p_2 are the prices of two competing products that the firm sells, find the prices that the firm should set to maximize its profit.

Constrained optimization: '*Find the maximum/minimum value(s) of the function* $w = f(x, y)$ ***subject to the constraint*** $g(x, y) = c$'. This is a constrained problem because the variables x and y are restricted (constrained) by the condition $g(x, y) = c$. I.e., we can't choose the variables x and y freely in the optimization problem.

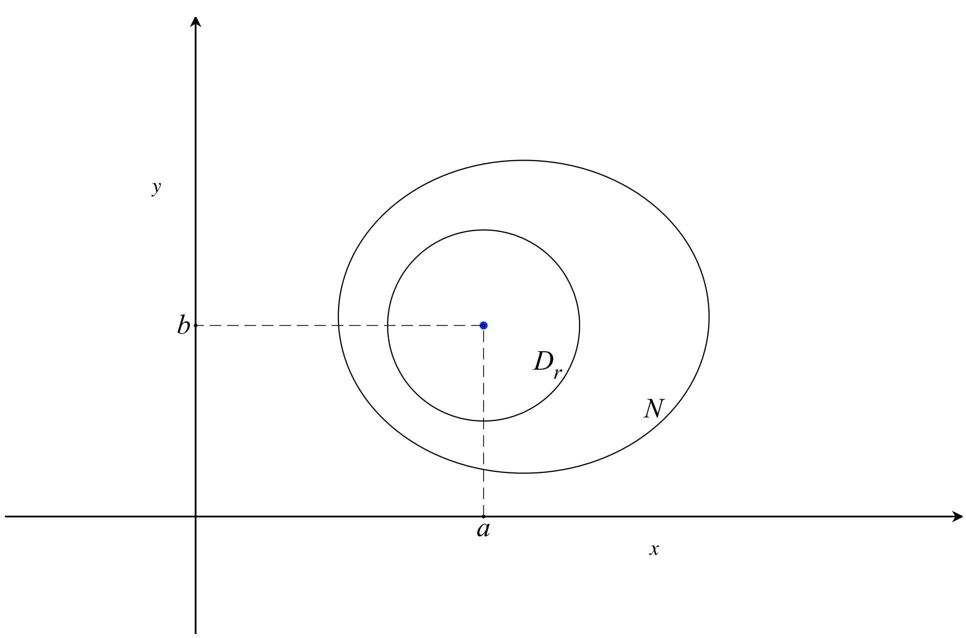
For example, if $C(k, l)$ is the cost to a firm of using k units of capital input and l units of labor input in their production process, and $Q(k, l)$ is the output generated by using k units of capital and l units of labor, then the *cost minimization problem* is that of minimizing the function $C(k, l)$ subject the the constraint $Q(k, l) = q_0$. That is to say, this is the problem of *minimizing the cost of producing* q_0 *units of output.*

(*) **Basic terminology and definitions** (with pretty pictures)

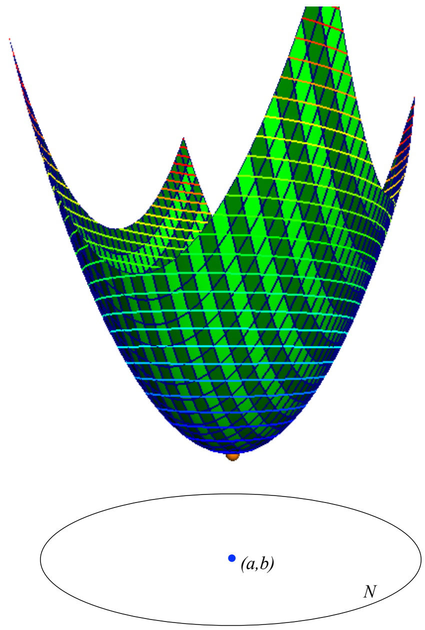
Definition: If (a, b) is a point in the plane, then an open disk $D_r(a, b)$ (of radius $r > 0$) centered at (a, b) is a set of the form

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

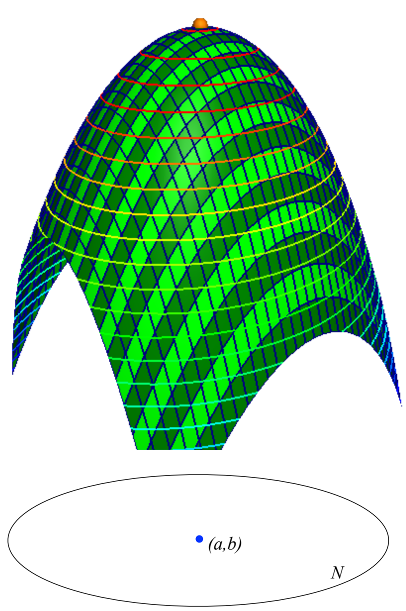
A ***neighborhood*** N of (a, b) is any set that contains an open disk $D_r(a, b)$ centered at (a, b).



Definition: $f(a,b)$ is a ***relative minimum*** value of the function $z = f(x,y)$ if $f(a,b) \leq f(x,y)$ 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).



(*) **Critical points**

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 called a ***critical point*** (or *stationary point*) of $f(x,y)$ and $f(a,b)$ is called a ***critical value***.

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

Note that this implication only goes one way — if $f(a,b)$ is a relative extreme value then (a,b) is a critical point, but not every critical value is necessarily a relative extreme value.

(*) **Explanation (of the key fact):** 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 1. $f_x(a,b) < 0$.
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 $f_x(a,b) \neq 0$ or $f_y(a,b) \neq 0$, 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**.*

Terminology: The (system of) equations

$$\begin{aligned} f_x(x,y) &= 0 \\ f_y(x,y) &= 0 \end{aligned}$$

(whose solutions are the critical points of $f(x,y)$) are sometimes referred to as the *first order conditions* for relative maximum/minimum value.

(*) **Example**

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

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

1. First order conditions:

$$\begin{aligned} f_x = 0 &\implies 2x - y + 3x^2 = 0 \\ f_y = 0 &\implies 2y - x = 0 \end{aligned}$$

2. Critical points: $\quad f_y = 0 \implies x = 2y$ 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: $\quad f(0,0) = 0$ and $f(-1/2,-1/4) = \dfrac{1}{16}$.

(*) **Generalization**

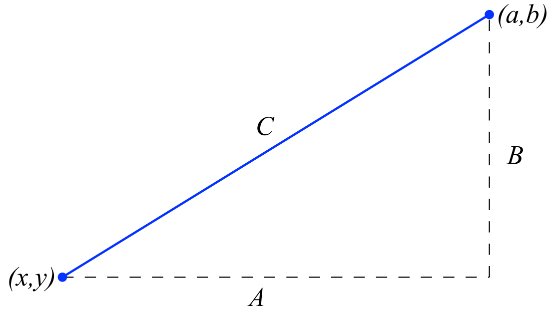
The definitions of neighborhoods, relative extreme values, critical points and critical values generalize in a straightforward way to functions of any number of variables, as does the relation between critical points and relative extreme values. I will list these generalizations below for a generic function of n variables.

Definitions.

- The ***distance*** between two points, $(x_1, x_2, \ldots, x_n)$ and $(a_1, a_2, \ldots, a_n)$ in n-dimensional space is given by

$$\operatorname{dist}\big((x_1, x_2, \ldots, x_n), (a_1, a_2, \ldots, a_n)\big) = \sqrt{(x_1 - a_1)^2 + (x_2 - a_2)^2 + \cdots + (x_n - a_n)^2}.$$

This formula is a direct generalization of the usual distance formula in two dimensions based on the Pythagorean theorem, as illustrated below.



The distance between the points (a,b) and (x,y) is C, and by the Pythagorean theorem, $C^2 = A^2 + B^2 = (a-x)^2 + (b-y)^2$, so

$$\operatorname{dist}\big((x,y),(a,b)\big) = C = \sqrt{(a-x)^2 + (b-y)^2}.$$

- The n-dimensional ball of radius r centered at the point $(a_1, a_2, \ldots, a_n)$, denoted $B_r(a_1, a_2, \ldots, a_n)$, is the collection of all points in n-dimensional space whose distance to this point is less than r. I.e.,

  $$B_r(a_1, a_2, \ldots, a_n) = \Big\{(x_1, x_2, \ldots, x_n) : \operatorname{dist}\big((x_1, x_2, \ldots, x_n), (a_1, a_2, \ldots, a_n)\big) < r\Big\}$$

  This is also a direct generalization of the two-dimensional case: in two dimensions, B_r is the *disk*, D_r defined above.

- A ***neighborhood*** $\mathcal{N}$ of the point $(a_1, a_2, \ldots, a_n)$ is any set that contains some n-dimensional ball $B_r(a_1, a_2, \ldots, a_n)$ centered at the point.

- Given a function of n variables, $y = f(x_1, x_2, \ldots, x_n)$, $f(a_1, a_2, \ldots, a_n)$ is a ***relative maximum value*** if

  $$f(a_1, a_2, \ldots, a_n) \geq f(x_1, x_2, \ldots, x_n)$$

  for all points $(x_1, x_2, \ldots, x_n)$ in some neighborhood $\mathcal{N}$ of the point.

- Likewise, $f(a_1, a_2, \ldots, a_n)$ is a ***relative minimum value*** if

  $$f(a_1, a_2, \ldots, a_n) \leq f(x_1, x_2, \ldots, x_n)$$

  for all points $(x_1, x_2, \ldots, x_n)$ in some neighborhood $\mathcal{N}$ of the point.

- The point $(a_1, a_2, \ldots, a_n)$ is a ***critical point*** of the *differentiable* function $f(x_1, x_2, \ldots, x_n)$ if all the first order partial derivatives of the function are equal to 0 at this point:

  $$f_{x_1}(a_1, a_2, \ldots, a_n) = f_{x_2}(a_1, a_2, \ldots, a_n) = \cdots = f_{x_n}(a_1, a_2, \ldots, a_n) = 0.$$

Fact: *If $f(x_1, x_2, \ldots, x_n)$ is a differentiable function and $f(a_1, a_2, \ldots, a_n)$ is a relative minimum/maximum value, then $(a_1, a_2, \ldots, a_n)$ is a critical point of the function.*

Conclusion: *To find the relative extreme value(s) of a function of several variables, the first step is to find the critical point(s) of the function.*

(*) **Example**

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

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

First order conditions:

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

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

$$2y - 8z = 0. \tag{4}$$

Adding $2\times$ equation (2) to equation (3) (eliminating the x s again) gives

$$9y - 4z = 8. \tag{5}$$

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

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

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:

$$\begin{aligned}
F_u = 0 &\implies & \frac{5}{u} - 10\lambda &= 0 && (6)\\
F_v = 0 &\implies & \frac{8}{v} - 15\lambda &= 0 && (7)\\
F_w = 0 &\implies & \frac{12}{w} - 25\lambda &= 0 && (8)\\
F_\lambda = 0 &\implies & -(10u + 15v + 25w - 3750) &= 0 && (9)
\end{aligned}$$

Equation (6) implies that

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

Likewise, equations (7) and (8) imply that

$$\frac{8}{v} = 15\lambda \implies \lambda = \frac{8}{15v} \tag{11}$$

and

$$\frac{12}{w} = 25\lambda \implies \lambda = \frac{12}{25w}. \tag{12}$$

Comparing equations (10) and (11) shows that

$$\lambda = \frac{1}{2u} = \frac{8}{15v} \implies 15v = 16u \implies v = \frac{16u}{15}. \tag{13}$$

Likewise, comparing equations (10) and (12) shows that

$$\lambda = \frac{1}{2u} = \frac{12}{25w} \implies 25w = 24u \implies w = \frac{24u}{25}. \tag{14}$$

Now, equation (9) simplifies as follows

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

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

and substituting for v and w (from equations (13) and (14)) 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

$$v^* = \frac{16}{15}u^* = 80, \; w^* = \frac{24}{25}u^* = 72 \; \text{ and } \; \lambda^* = \frac{1}{2u^*} = \frac{1}{150}.$$

I.e., the critical point is $(u^*, v^*, w^*, \lambda^*) = (75, 80, 72, 1/150)$ and the critical value is

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