1 Session 4: Constrained optimization

In the previous section we have considered functions of several variables that are independent from each other. What happens if the variables are, however, not independent from each other? (see excellent material in Richard Weber's web site)

What is the maximum height we can attain constrained to the trajectory of the road?

Let us consider optimizing a function $f(x, y)$ (the height of the hill in the previous exemple) subject to the constrain $g(x, y) = c$ (the equation describing the path, the road).

One can simply substitute $g(x, y) = c$ in $f(x, y)$, but this may not involve trivial algebra.

Alternatively, let us consider the method of the Lagrange multipliers. To optimize $f(x, y)$ we require

$$df = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy = 0$$

we saw that if dx and dy were independent, $\frac{\partial f}{\partial x} = \frac{\partial f}{\partial y} = 0$. But they are not here: they are constrained by $g(x, y)$, which is a constant:

$$dg = \frac{\partial g}{\partial x} dx + \frac{\partial g}{\partial y} dy = 0$$
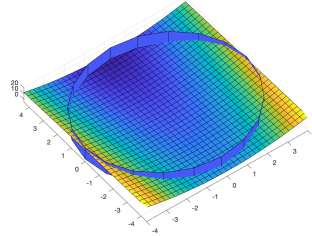
Let us consider a parameter λ (the *Lagrange multiplier*) and obtain an expression for which dx and dy are independent:

$$d(f + \lambda g) = (\frac{\partial f}{\partial x} + \lambda \frac{\partial g}{\partial x}) dx + (\frac{\partial f}{\partial y} + \lambda \frac{\partial g}{\partial y}) dy = 0$$

Now:

$$\begin{aligned}
\frac{\partial f}{\partial x} + \lambda \frac{\partial g}{\partial x} &= 0 \\
\frac{\partial f}{\partial y} + \lambda \frac{\partial g}{\partial y} &= 0 \\
g(x, y) &= c
\end{aligned}$$

are sufficient to find λ, x^* and y^*.

Exercise 1

Find the optimal values of the function $f(x, y) = x^2 - y$ constrained within the circumference $x^2 + y^2 = 4$. (note that, interestingly, in the optimal value x^* we will get a gradient for the constraint that is parallel to the gradient of the function: $\boldsymbol{\nabla} f(x^*) = \lambda^* \boldsymbol{\nabla} g(x^*)$).

The general problem we study takes the form

$$\begin{aligned} \text{maximize} \quad & f(x) \\ \text{subject to} \quad & \begin{cases} x \in X \\ g(x) = b \end{cases} \end{aligned}$$

where

$$\begin{aligned} x &\in \mathbf{R}^n && (n \text{ decision variables}) \\ f &: \mathbf{R}^n \to \mathbf{R} && (\text{objective function}) \\ X &\subseteq \mathbf{R}^n && (\text{regional constraints}) \\ \left. \begin{array}{l} g : \mathbf{R}^n \to \mathbf{R}^m \\ b \subseteq \mathbf{R}^m \end{array} \right\} && (m \text{ functional constraints}) \end{aligned}$$

If the constraint is $g(x) \leq b$ we introduce a **slack variable** z and write:

$$g(x) + z = b, \; z \geq 0$$

Regional and functional constraints define the **feasible set** for x

In the Lagrange approach, the constrained maximization (minimization) problem is rewritten as a Lagrange function

$$L(x, \lambda) = f(x) - \lambda g(x)$$

whose optimal point is a saddle point, i.e. a global maximum (minimum) over the domain of the choice variables and a global minimum (maximum) over the multipliers.

Let us take this example:

$$\begin{aligned} \text{maximize} \quad & f(x) = x_1 - x_2 - 2x_3 \\ \text{subject to} \quad & \begin{cases} x_1 + x_2 + x_3 = 5 \\ x_1^2 + x_2^2 = 4 \\ x = (x_1, x_2, x_3) \in \mathbf{R}^3 \end{cases} \end{aligned}$$

1. We write the Lagrangian as (note we have two constraints):

$$\begin{aligned}
L(x,\lambda) &= x_1 - x_2 - 2x_3 - \lambda_1(x_1 + x_2 + x_3 - 5) - \lambda_2(x_1^2 + x_2^2 - 4) \\
&= [x_1(1-\lambda_1) - \lambda_2 x_1^2] + [x_2(-1-\lambda_1) - \lambda_2 x_2^2] + [x_3(-2-\lambda_1)] \\
&\quad + 5\lambda_1 + 4\lambda_2
\end{aligned}$$

2. We want to minimize the Lagrangian subject to the regional constraints $x = (x_1, x_2, x_3) \in \mathbf{R}^3$.

3. We find the values λ such that will define a feasible set

$$Y = \left\{ \lambda : \underbrace{\min_{x \in \mathbf{R}^3} L(x,\lambda)}> -\infty \right\}$$

So, we only consider the values of λ for which we obtain a finite minimum. In the example

$$Y = \{\lambda : \lambda_1 = -2, \lambda_2 < 0\}$$

4. For $\lambda \in Y$, the minimum will be obtained at some $x(\lambda)$ (that depends on λ in general). Here, $x(\lambda) = (3/2\lambda_2, 1/2\lambda_2, x_3)^T$.

5. Finally, we find a λ value that provides a feasible value $x(\lambda)$. Here

$$x_1^2 + x_2^2 = 4 \Rightarrow \lambda_2 = -\sqrt{5/8}$$

So, x^* is optimal:

$$x^* = (-3\sqrt{2/5}, -\sqrt{2/5}, 5 + 4\sqrt{2/5})^T; \ \lambda^* = (-2, -\sqrt{5/8})^T$$

Exercise 2

Use the Lagrange multipliers to solve this problem:

$$\begin{aligned}
\text{minimize} \quad & f(x,y) = x^2 y \\
\text{subject to} \quad & x^2 + y^2 = 1
\end{aligned}$$

[1] [2]

Let us consider a general optimization problem with constraints: let $f : \mathbf{R}^n \to \mathbf{R}$ be the objective function and $\mathbf{g} : \mathbf{R}^n \to \mathbf{R}^m$ a set of constraint functions ($\mathbf{g}(\mathbf{x}) = (g_1(\mathbf{x}), \ldots, g_m(\mathbf{x}))$) with $\mathbf{x} = (x_1, \ldots, x_n)$. We want to obtain

$$\text{Min } f(\mathbf{x}) \text{ such that } \mathbf{g}(\mathbf{x}) \geq 0$$

Theorem 1. *If $\boldsymbol{x}^*$ is a local minimum for the optimisation problem and the constraints $\boldsymbol{g}(\boldsymbol{x}) \geq 0$ are satisfied at $\boldsymbol{x}^*$, then $\boldsymbol{x}^*$ must satisfy:*

$$\begin{aligned}
\boldsymbol{\nabla} f(\boldsymbol{x}) &= \lambda_1 \boldsymbol{\nabla} g_1(\boldsymbol{x}) + \cdots + \lambda_m \boldsymbol{\nabla} g_m(\boldsymbol{x}) \\
0 &= \lambda_1 g_1(\boldsymbol{x}) = \cdots = \lambda_m g_m(\boldsymbol{x}) \\
0 &\leq \boldsymbol{g}(\boldsymbol{x})
\end{aligned}$$

[1] https://people.maths.bris.ac.uk/~maxmr/opt/kt1.pdf
[2] https://www.cmi.ac.in/~madhavan/courses/dmml2018/literature/Lagrangian_Methods_for_Constrained_Optimization.pdf

Exercise 3

Consider this NLO problem[1]:

$$\text{maximize} \quad f(x, y) = x^2 y + 2y^2$$

$$\text{subject to} \quad \begin{cases} x + 3y \leq 9 \\ x + 2y \leq 8 \\ 3x + 2y \leq 18 \\ 0 \leq x \leq 5 \\ 0 \leq y \leq 2 \end{cases}$$

Solve the program graphically- Begin at point $(0,0)$ and check the gradient. Check the KKT conditions. If they are not satisfied, find an improving direction in the feasible region.

To better understand these exercises, you can make use of geogebra.

Exercise 4

Is it possible to maximize $f(x, y) = xy$ subject to $100 \geq x + y$ and $x \leq 40$?

Exercise 5

What about maximizing $f(x, y) = xy$ subject to $x + y^2 \leq 2$ and $x, y \geq 0$?

References

[1] Michael Carter, Camille C Price, and Ghaith Rabadi. *Operations Research. A Practical Introduction. Second Edition.* CRC Press, 2019.