[MATH-20] The Kuhn-Tucker method

Miguel-Angel Canela
Associate Professor, IESE Business School

The Lagrange function

In the Lagrange method, when X^* is a local extremum point of f, there exist $\lambda_1, \ldots, \lambda_m$ such that

$$\nabla f(X^*) = \lambda_1 \nabla g_1(X^*) + \cdots + \lambda_m \nabla g_m(X^*).$$

Then, $\lambda_1, \ldots, \lambda_m$ are said to be the **Lagrange multipliers** for X^*. In the classical presentations of the Lagrange method (in most textbooks), the multipliers are introduced through the so called **Lagrange function**,

$$L(X, \lambda) = f(X) - \lambda_1\, g_1(X) - \cdots - \lambda_m\, g_m(X).$$

It is easy to see that the definition of the multipliers is equivalent to the system of equations

$$\frac{\partial L}{\partial x_1} = \cdots = \frac{\partial L}{\partial x_n} = 0,$$

or, in vector notation, $\nabla_X L(X, \lambda) = \mathbf{0}$. The subscript in ∇_X means that only derivatives with respect to the coordinates of X are included. λ stands for $(\lambda_1, \ldots, \lambda_m)$.

Example 1. In the Example 5 of the preceding lecture,

$$L(x_1, x_2, \lambda) = x_1 x_2 - \lambda(x_1 + x_2 - 1),$$

hence

$$\nabla_X L(x_1, x_2) = \begin{bmatrix} x_2 - \lambda \\ x_1 - \lambda \end{bmatrix} = \mathbf{0} \Longrightarrow x_1 = x_2 = \lambda.$$

So, in this example, $\lambda = 1/2$. $\square$

Example 2. The Lagrange method can be used in combination with the search for local extremum to solve optimization problems in closed bounded domains. Let me consider the problem of maximizing the objective $f(x_1, x_2, x_3) = x_1 + x_2 + x_3$, subject to $x_1^2 + 2x_2^2 + 3x_3^2 \le 1$.

First, there must be a maximum and minimum value of f in this domain (which you can imagine like a watermelon), since it is closed and bounded. If one of these extremum values is attained at an interior point, it will be a local extremum point, with null gradient. But

$$\nabla f(x_1, x_2, x_3) = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \ne \mathbf{0},$$

so the solution must be found among the boundary points ($x_1^2 + 2x_2^2 + 3x_3^2 = 1$). We can apply now the Lagrange method. The Lagrange function is

$$L(x_1, x_2, x_3, \lambda) = x_1 + x_2 + x_3 - \lambda\big(x_1^2 + 2x_2^2 + 3x_3^2 - 1\big),$$

hence

$$\nabla_X L(x_1, x_2, x_3, \lambda) = \begin{bmatrix} 1 - 2\lambda x_1 \\ 1 - 4\lambda x_2 \\ 1 - 6\lambda x_3 \end{bmatrix} = \mathbf{0}.$$

Since none of the x_i can be null, we can write

$$\lambda = \frac{1}{2x_1} = \frac{1}{4x_2} = \frac{1}{6x_3}\,,$$

leading to $x_1 = 2x_2 = 3x_3$. This is leaves us with two potential solutions,

$$\pm\big(\sqrt{6/11}, \sqrt{3/11}, \sqrt{2/11}\big).$$

One must be the maximum and the other the minimum, and it is obvious which is which. □

The Kuhn-Tucker method

The **Kuhn-Tucker method** is a general method for solving optimization problems when the constraints are inequalities, which, in practice, consists of packing together two methods as I did in Example 2. It is based on the following theorem.

Theorem. Suppose that the restriction of f to the domain defined by

$$g_1(X) \le 0,\ \ldots\ ,\ g_m(X) \le 0$$

has a local maximum at X^*. Then, there exists $\lambda = (\lambda_1, \ldots, \lambda_m)$ such that:

(1) $\nabla_X L(X^*, \lambda) = \mathbf{0}$.

(2) $\lambda_j g_j(X^*) = 0$, for $j = 1, \ldots, m$.

(3) $\lambda_j \ge 0$, $g_j(X^*) \le 0$, for $j = 1, \ldots, m$.

If f has a local minimum, the same conditions hold, except for (3), where we have $\lambda_j \le 0$.

In most cases, the domain is closed and bounded, so that one knows in advance that there is a maximum and a minimum value of f. Then, if the Kuhn-Tucker method gives more than one potential maximum (or minimum) point, we identify it by comparing the values of f at those points.

Example 3. Let us consider the objective $f(x_1, x_2) = x_1 x_2$ and the constraint

$$g(x_1, x_2) = x_1^2 + x_2^2 - 1 \le 0.$$

Here,

$$L(x_1, x_2, \lambda) = x_1 x_2 - \lambda\big(x_1^2 + x_2^2 - 1\big),$$

and condition (1) is

$$\nabla_X L(x_1, x_2, \lambda) = \begin{bmatrix} x_2 - 2\lambda x_1 \\ x_1 - 2\lambda x_2 \end{bmatrix} = \mathbf{0}.$$

Condition (2) gives two cases:

- Case 1: $\lambda = 0$. Then $x_1 = x_2 = 0$. We already know that $(0, 0)$ is a saddle point of f (Example 1, lecture 15).
- Case 2: $g(x_1, x_2) = 0$. Adding this equation to condition (1), we have a system of three equations with three unknowns, whose solutions are:

  (a) $x_1 = x_2 = \pm 1/\sqrt{2}$, $\lambda = 1/2$. This gives the maximum value $1/2$.

  (b) $x_1 = -x_2 = \pm 1/\sqrt{2}$, $\lambda = -1/2$. This gives the minimum value $-1/2$. □

Example 4. Consider the objective $f(x_1, x_2) = x_2 - x_1$ and the constraints

$$g_1(x_1, x_2) = x_1^2 + x_2^2 - 1 \leq 0, \qquad g_2(x_1, x_2) = x_1^2 - x_2 - 1 \leq 0.$$

Here,

$$L(x_1, x_2, \lambda) = x_2 - x_1 - \lambda_1\bigl(x_1^2 + x_2^2 - 1\bigr) - \lambda_2\bigl(x_1^2 - x_2 - 1\bigr).$$

Condition (1) is

$$\nabla_X L(x_1, x_2, \lambda) = \begin{bmatrix} -1 - 2\lambda_1 x_1 - 2\lambda_2 x_1 \\ 1 - 2\lambda_1 x_2 + \lambda_2 \end{bmatrix} = \mathbf{0}.$$

In condition (2), we consider the following cases:

- Case 1: $\lambda_1 = \lambda_2 = 0$. There is no solution here, due to $\nabla_X L(x_1, x_2, \lambda) \neq \mathbf{0}$.
- Case 2: $\lambda_1 = 0$, $\lambda_2 \neq 0$, $g_2(x_1, x_2) = 0$. This gives

  $$1 + 2\lambda_2 x_1 = 1 + \lambda_2 = x_1^2 - x_2 - 1 = 0,$$

  whose solution is $x_1 = 1/2$, $x_2 = -3/4$, $\lambda_2 = -1$.
- Case 3: $\lambda_2 = 0$, $\lambda_1 \neq 0$, $g_1(x_1, x_2) = 0$. This gives

  $$1 + 2\lambda_1 x_1 = 1 - 2\lambda_1 x_2 = x_1^2 + x_2^2 - 1 = 0.$$

  The solution is $x_1 = 1/\sqrt{2}$, $x_2 = -1/\sqrt{2}$, $\lambda_1 = 1/\sqrt{2}$.
- Case 4: $\lambda_1\lambda_2 \neq 0$, $g_1(x_1, x_2) = g_2(x_1, x_2) = 0$. This gives us the three intersection points
  (a) $(1, 0)$, with $\lambda_1 = -3/2$, $\lambda_2 = 1$.
  (b) $(-1, 0)$, with $\lambda_1 = -1/2$, $\lambda_2 = 1$.
  (c) $(0, -1)$, with $\nabla_X L(x_1, x_2, \lambda) \neq \mathbf{0}$.

Therefore, cases 2 and 3 gives us the minimum and maximum points, respectively. □

Homework

A. Find the maximum and the minimum value of $f(x_1, x_2) = x_1^2 + 2x_2^2 - x_1$ subject to $x_1^2 + x_2^2 \leq 1$, $x_1 \geq 0$ and $x_2 \geq 0$.