May 9, 2023

4. OPTIMIZATION WITH INEQUALITY CONSTRAINTS: THE METHOD OF KUHN-TUCKER

All throughout this chapter, D denotes an **open subset** of $\mathbb{R}^n$.

1. Introduction

We study next optimization problems with inequality constraints.

$$\max \; f(x) \quad \text{s.t.: } x \in S, \tag{1.1}$$

where $S = \{x \in D : \; g_1(x) \leq b_1, \ldots, g_m(x) \leq b_m\}$ and where $f : D \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$, $g = (g_1, \ldots, g_m) : D \longrightarrow \mathbb{R}^m$. Contrary to the Lagrange case, there is no limitation in the (finite) number of constraints.

Remark 1.1. A constraint of the type $g_i(x) \geq b_i$ is equivalent to $-g_i(x) \leq -b_i$. The problem

$$\min \; f(x) \quad \text{s.t.: } x \in S,$$

has the same solution(s) that

$$\max \; -f(x) \quad \text{s.t.: } x \in S.$$

2. Khun-Tucker necessary conditions

Definition 2.1. Given a point $x_0 \in D$, we say that the restriction $i = 1, 2, \ldots, m$ is **binding** at the point x_0 for problem 1.1 if $g_i(x_0) = b_i$. If $g_i(x_0) < b_i$, then we say that the restriction i is not binding at the point x_0.

Definition 2.2. Let g be the of class C^1 in D. The point $x_0 \in D$ is **regular** if either no restriction binds at x_0 or the gradient vectors of the constraints than binds at x_0 form a matrix of maximal rank (that is, if it is $k \leq m$ the number of constraints that bind at x_0, then the gradient vectors for these constraints form a matrix of rank k).

Definition 2.3. The Lagrangian function associated to (1.1) is

$$\mathcal{L}(x, \lambda) = f(x) + \lambda \cdot (b - g(x)), \tag{2.1}$$

where $\lambda = (\lambda_1, \ldots, \lambda_m)$.

Theorem 2.4 (Kuhn-Tucker's Method)**.** Suppose that the functions $f : D \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$, $g = (g_1, \ldots, g_m) : D \longrightarrow \mathbb{R}^m$ are of class C^1 in D and x_0 is a regular point of (1.1). If x_0 is a solution of problem 1.1, then there is a vector $\lambda_0 = (\lambda_1^0, \ldots, \lambda_m^0) \in \mathbb{R}^m$ such that

$$\forall i \in \{1, \ldots, n\} \quad \frac{\partial \mathcal{L}}{\partial x_i}(x_0, \lambda_0) = 0,$$

$$\forall i \in \{1, \ldots, m\} \quad \left\{ \begin{array}{rcl} \lambda_i^0 (b_i - g_i(x_0)) & = & 0, \\ \lambda_i^0 & \geq & 0, \\ g_i(x_0) & \leq & b_i. \end{array} \right.$$

Remark 2.5. The equations

(1) $\nabla_x L(x, \lambda) = 0,$

(2) $\lambda_1(b_1 - g_1(x)) = 0, \cdots, \lambda_m(b_m - g_m(x)) = 0$,
(3) $\lambda_1 \geq 0, \ldots, \lambda_m \geq 0$,
(4) $g_1(x) \leq b_1, \ldots, g_m(x) \leq b_m$.

are the Kuhn-Tucker equations of problem 1.1.

Remark 2.6. A way to apply the necessary conditions of Theorem 2.4 is to find the solutions of the system of $n+m$ equations and $n+m$ unknowns

$$\left\{\begin{array}{ccc} \frac{\partial \mathcal{L}}{\partial x_1}(x,\lambda) & = & 0 \\ \vdots & \vdots & \vdots \\ \frac{\partial \mathcal{L}}{\partial x_n}(x,\lambda) & = & 0 \\ \lambda_1^0(b_1 - g_1(x_0)) & = & 0 \\ \vdots & \vdots & \vdots \\ \lambda_m^0(b_m - g_m(x_0)) & = & 0. \end{array}\right.$$

From the solutions (x,λ) found, only are retained those that satisfy the rest of conditions: $x \in S$ and $\lambda_i \geq 0$.

Example 2.7 (perfect substitutes). Suppose that an agent has income 5 and that his utility function over consumption bundles is $u(x,y) = 2x + y$. If the prices of the goods are $p_1 = 3$, $p_2 = 1$ what are the demand functions of the agent?
The maximization problem of the agent is

$$\begin{array}{rl} \max & 2x + y \\ \text{s.a.} & 3x + y \leq 5 \\ & x \geq 0 \\ & y \geq 0 \end{array}$$

We write first this problem in the form 1.1

$$\begin{array}{rl} \max & 2x + y \\ \text{s.a.} & 5 - 3x - y \geq 0 \\ & x \geq 0 \\ & y \geq 0 \end{array}$$

The associated Lagrangian is

$$\mathcal{L}(x,y,\lambda_1,\lambda_2,\lambda_3) = 2x + y + \lambda_1(5 - 3x - y) + \lambda_2 x + \lambda_3 y$$

and the Kuhn-Tucker equations are

$$\frac{\partial L}{\partial x} = 2 - 3\lambda_1 + \lambda_2 = 0 \tag{2.2}$$

$$\frac{\partial L}{\partial y} = 1 - \lambda_1 + \lambda_3 = 0 \tag{2.3}$$

$$\lambda_1(5 - 3x - y) = 0 \tag{2.4}$$

$$\lambda_2 x = 0 \tag{2.5}$$

$$\lambda_3 y = 0 \tag{2.6}$$

$$3x + y \leq 5$$

$$x \geq 0$$

$$y \geq 0$$

$$\lambda_1, \lambda_2, \lambda_3 \geq 0$$

We try to solve first equations (2.2)-(2.6). Note that if $\lambda_1 = 0$ then the first equation implies that $\lambda_2 = -2 < 0$, which contradicts equation $\lambda_1, \lambda_2, \lambda_3 \geq 0$. Therefore, $\lambda_1 > 0$. From equation $\lambda_1(5 - 3x - y) = 0$ we conclude that $5 - 3x - y = 0$ so that

$$y = 5 - 3x$$

Suppose that $x > 0$. In this case, the equation $\lambda_2 x = 0$ implies that $\lambda_2 = 0$. From the first equation we see that $\lambda_1 = 2/3$ and substituting in the equation we obtain that $\lambda_3 = \lambda_1 - 1 = -1/3 < 0$ which contradicts the equation $\lambda_1, \lambda_2, \lambda_3 \geq 0$.
We conclude that $x = 0$ and $y = 5$. Then,

$$x = 0, \quad y = 5, \quad \lambda_1 = \lambda_2 = 1, \quad \lambda_3 = 0$$

is the unique solution of the system.

Remark 2.8. If S is compact, then the Theorem of Weierstrass assures the existence of global solutions of (1.1). Assuming that the condition of regularity holds, these global solutions are critical points of f relative to S. We will locate the global solutions by evaluating those critical points with the objective function f. The extremal values will determine the global maximum and minimum of f on S.

3. Sufficient conditions. Convex programs

Let $f : D \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$, $g = (g_1, \ldots, g_m) : D \longrightarrow \mathbb{R}^m$ be of class C^1 in D.

Theorem 3.1. If in problem (1.1) the function f is concave and the functions $g_1, \ldots, g_m$ are convex, then the Khun-Tucker necessary conditions of Theorem 2.4 are also sufficients and if $x_0 \in S$ fulfills them, then x_0 is a global maximum if f in S.

Remark 3.2. In the conditions of the above theorem, if f is strictly concave, then the global maximum, if exists, is unique. Hence, once a solution $x_0, \lambda_0)$ of the KT conditions has been found, we can stop searching, since it is the only solution of the KT conditions.

4. Non-negativity constraints

Most often, inequality constrained problems include non-negativity of the variables. Although this is a particular case of the theory developed so far, you will find in textbooks an apparently different set of Khun-Tucker conditions. Actually, they are equivalent to the original ones, but applied to this specific problem of non-negativity of the variables, adopt the following structure. Let the optimization problem with inequality constraints and non-negativity constraints

$$\max \ f(x) \quad \text{s.t.: } x \in S, \tag{4.1}$$

where $S = \{x \in D : \ g_1(x) \leq b_1, \ldots, g_m(x) \leq b_m, x_1 \geq 0, \ldots, x_n \geq 0\}$ and where $f : D \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$, $g = (g_1, \ldots, g_m) : D \longrightarrow \mathbb{R}^m$.

Definition 4.1. The Lagrangian function associated to (4.1) is

$$L(x, \lambda) = f(x) + \lambda \cdot (b - g(x)), \tag{4.2}$$

where $\lambda = (\lambda_1, \ldots, \lambda_m)$.

Note that this Lagrangian does not incorporate the non-negativity constraints. It only attaches multipliers to the other kind of constraints, $g_i(x) \leq b_i$.

Theorem 4.2 (Kuhn-Tucker's Method for problems with non-negativity cosntraints)**.** Suppose that the functions $f : D \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$, $g = (g_1, \ldots, g_m) : D \longrightarrow \mathbb{R}^m$ are of class C^1 in D and x_0 is a regular point of[1] (4.1). If x_0 is a solution of problem 1.1, then there is a vector $\lambda_0 = (\lambda_1^0, \ldots, \lambda_m^0) \in \mathbb{R}^m$ such that

$$\forall i \in \{1, \ldots, n\} \quad \left\{ \begin{array}{rcl} x_i \frac{\partial L}{\partial x_i}(x_0, \lambda_0) & = & 0, \\ \frac{\partial L}{\partial x_i}(x_0, \lambda_0) & \leq & 0 \end{array} \right.$$

$$\forall i \in \{1, \ldots, m\} \quad \left\{ \begin{array}{rcl} \lambda_i^0 (b_i - g_i(x_0)) & = & 0, \\ \lambda_i^0 & \geq & 0, \\ g_i(x_0) & \leq & b_i. \end{array} \right.$$

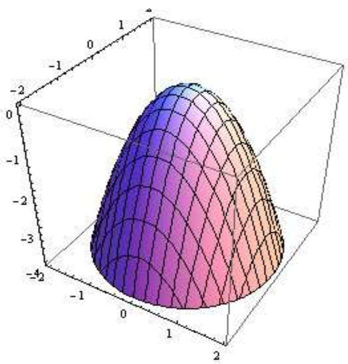
5. Optimization of convex (concave) functions

Let C be a convex subset of $\mathbb{R}^n$. Now we consider either of the following problems:

(1) The function $f : C \to \mathbb{R}$ is concave in C and we study the problem

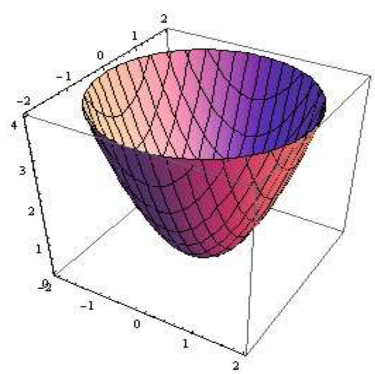
$$\max_{x \in C} f(x)$$

[1] For this, it suffices that the matrix formed whose columns are the gradient vectors $\nabla g_1(x_0)^t, \ldots, \nabla g_m(x_0)^t$, has rank m; there is no need to consider the non-negativity constraints to check this condition.



(2) The function $f : C \to \mathbb{R}$ is convex in C and we study the problem

$$\min_{x \in C} f(x)$$



Proposition 5.1. Let C be a convex subset of $\mathbb{R}^n$. Let $f : C \to \mathbb{R}$.

(1) if f is concave and x_0 is a local maximum of f on C, then x_0 is a global maximum of f on C.
(2) if f is convex, and $x_0 \in C$ is a local minimum of f on C, then x_0 is a global minimum of f on C.

Proposition 5.2. Let $C \subset \mathbb{R}^n$ be convex and open, $x_0 \in C$ and f of class C^1 on C.

(1) If f is concave on C then, x_0 is a global maximum of f on C if and only if $\nabla f(x_0) = 0$.
(2) If f is convex on C, then x_0 is a global minimum of f on C if and only if $\nabla f(x_0) = 0$.

Proof In either case, if f has a maximum at x_0 then, $\nabla f(x_0) = 0$. If, for example, f is concave then, for each $x \in C$ we have that

$$f(x) \leq f(x_0) + \nabla f(x_0)(x - x_0) = f(x_0)$$

Hence, if $\nabla f(x_0) = 0$, we have that$f(x) \leq f(x_0)$ for other $x \in C$.

Remark 5.3. If a function is strictly concave (resp. convex)then,

$$f(x) < f(x_0) + \nabla f(x_0)(x - x_0) = f(x_0)$$

and we see that if it has a maximum (resp. minimum) point, then it is unique. This can be proved directly from the definition, without using the first order conditions.

Proposition 5.4. Let $C \subseteq \mathbb{R}^n$ be nonempty and convex and let $f : C \to \mathbb{R}$. Then

(1) The minimum set of a convex function is a convex set.
(2) The maximum set of a concave function is a convex set.

Proof Let us prove the first assertion. The second one follows from this by considering $-f$. Let M denote the set of minimum points of f. If $M = \emptyset$, then there is nothing to prove. Otherwise, let $x_1, x_2 \in M$ and let $\lambda \in [0,1]$. We have $m = f(x_i) \leq f(x)$ for all $x \in M$, for $i = 1, 2$. Since f is convex, $f(\lambda x_1 + (1-\lambda)x_2) \leq \lambda f(x_1) + (1-\lambda)f(x_2) = m$. Hence $\lambda x_1 + (1-\lambda)x_2$ is also a minimum point of f and thus $\lambda x_1 + (1-\lambda)x_2 \in M$ and M is convex.

Theorem 5.5. Let $C \subseteq \mathbb{R}^n$ be nonempty and convex, let $f : C \to \mathbb{R}$ and let x_0 be an interior point of f. Then

(1) If f is convex and x_0 is a global maximum of f in C, then f is constant.
(2) If f is concave and x_0 is a global minimum of f in C, then f is constant.

Definition 5.6. Let $C \subseteq \mathbb{R}^n$ be a convex set. A point $x_0 \in C$ is a **vertex** or **extreme point** of C if $x_0 = \lambda x_1 + (1-\lambda)x_2$ for $x_1, x_2 \in C$ and $\lambda \in [0,1]$, implies $\lambda = 0$ or $\lambda = 1$.

Hence, a point of C is a vertex if it cannot be written as a non-trivial convex combination of points of C. Note that a vertex is a boundary point of C, but not every boundary point that belongs to C is a vertex. For instance, consider the sets

$$\begin{aligned} C_1 &= \{(x,y) \in \mathbb{R}^2 : x + y < 1\}, \\ C_2 &= \{(x,y) \in \mathbb{R}^2 : x + y \leq 1\}, \\ C_3 &= \{(x,y) \in \mathbb{R}^2 : x + y \leq 1, x \geq 0, y \geq 0\}, \\ C_4 &= \{(x,y) \in \mathbb{R}^2 : x^2 + y^2 \leq 1\}. \end{aligned}$$

All are convex sets. Neither C_1 nor C_2 has vertices. C_3 has three vertices, the points $(0,0)$, $(0,1)$ and $(1,0)$. C_4 has infinitely many vertices (every point in the circumference $x^2 + y^2 = 1$ is a vertex).

Theorem 5.7. Let $C \subseteq \mathbb{R}^n$ be nonempty, convex and compact and let $f : C \to \mathbb{R}$ be continuous. Then

(1) If f is convex, then f attains its global maximum in some vertex of C.
(2) If f is concave, then f attains its global minimum in some vertex of C.

Example 5.8. The function $f(x,y) = -x^2 - y^2 + \ln(1 + x + y)$ is strictly concave in the compact and convex set C_3 defined above. The derivatives of f are

$$\begin{aligned} \nabla f(x,y) &= \left(-2x + \frac{1}{1+x+y}, -2y + \frac{1}{1+x+y}\right) \\ Hf(x,y) &= \begin{pmatrix} -2 - \dfrac{1}{(1+x+y)^2} & -\dfrac{1}{(1+x+y)^2} \\ -\dfrac{1}{(1+x+y)^2} & -2 - \dfrac{1}{(1+x+y)^2} \end{pmatrix}. \end{aligned}$$

The principal minors of the Hessian matrix of f are: $\Delta_1 = -2 - \frac{1}{(1+x+y)^2} < 0$, and $\Delta_2 = 4 + \frac{4}{(1+x+y)^2} > 0$.

The vertices of C_3 are $(0,0)$, $(0,1)$ y $(1,0)$. From the above theorem, the global minima of f in C_3 are among these points. Since $f(0,0)=0$, $f(0,1)=f(1,0)=-1+\ln 2<-1+\ln e=-1+1=0$, the points $(0,1)$ y $(1,0)$ are the global minima of f in C_3.
The theorem gives no information about maxima (other than existence). They could be located in the interior or in the boundary of C_3. To find them, we could set the corresponding KT problem. However, in this case the FOCs

$$\frac{\partial f}{\partial x}(x,y)=-2x+\frac{1}{1+x+y}=0,$$

$$\frac{\partial f}{\partial y}(x,y)=-2y+\frac{1}{1+x+y}=0,$$

have a unique solution, $(\frac{1}{4}(\sqrt{5}-1),\frac{1}{4}(\sqrt{5}-1))$, which belongs to C_3, since its components are non-negative and add up $\frac{1}{2}(\sqrt{5}-1)\leq 1$. Thus, this point is the global maximum of f in C_3 (and in $\mathbb{R}^2_+$ indeed).