Optimization of a concave function using a modified gradient projection method. Application to the determination of optimal controls in production systems

Optimisation d'une fonction concave par une méthode modifiée de gradient projeté. Application à la détermination de contrôles optimaux dans des systèmes de production

Diagne S. G[1], Gningue Y.[2], Ndiaye S[3],

Abstract :
We propose here a new descent method for the constrained minimization of concave functions in a domain determined by the dynamic equations for a given system. We use the concavity properties, the characteristics of the optimal solution and some modified form of gradient projection in order to find the direction of descent, and to do largest steps of moves towards the optimum, without getting out of the admissible domain. The results are then used for the determination of optimal control problems in production systems.

Key words :
Optimal control, projection gradient, concave, dynamic systems, production systems.

Résumé :
Nous proposons une nouvelle méthode de descente accélérée pour la minimisation sous contraintes de fonctions concaves dans un domaine défini par les équations d'évolution d'un système donné. Nous utilisons les inéquations caractérisant la concavité des fonctions différentiables, les propriétés de la solution optimale et la configuration du domaine admissible pour agir sur la direction indiquée par le gradient projeté et les pas de déplacement afin d'assurer l'accélération. Contrairement aux pas habituels infiniment petits que l'on fait dans les méthodes de gradient, ici nous faisons des pas relativement grands en garantissant notre maintien dans le domaine admissible. Nous donnons ensuite des exemples d'application au contrôle optimal de systèmes dynamiques.

Mots clés :
Contrôle optimal, gradient projeté, concave, système dynamique, système de production.

[1]* Diagne Salimata Guèye, Université Cheikh Anta Diop de Dakar Sénégal
Correspondant : Tél…………….. Email : **gueyesa@hotmail.com** .
[2] Université Laurentienne Sudbury Canada.
[3] Université Cheikh Anta Diop de Dakar Sénégal

1. Problem statement and notations

We consider a system governed by the set of dynamic equations of the following type:

$$x_t^j = x_{t-1}^j + u_{t-1}^j - q_t^j \;; \quad j = 1, 2, \ldots, m \;; t = 1, 2, \ldots, n \qquad \textbf{(1)}$$

with $u_{t-1}^j \geq 0$; $x_{t-1}^j \geq 0$; $x_n^j = 0$; x_o^j and the q_t^j a set of given constants.

$(x_t^j)_{j=1,..,m}$ stands for the state at stage t,

$(u_t^j)_{j=1,\ldots,m}$ stands for the control at stage t.

$$\sigma_{i,k}^j = \begin{cases} \sum_{t=i}^{k} q_t^j & \text{if } i \leq k \\ 0 & \text{else} \end{cases}$$

and

$i_j = n(j-1) + 1$, j = 1, 2, …, m

the vectors $v = (v_1, \ldots, v_{nm})$,

$y = (y_1, \ldots, y_{nm})$ and $d = (d_1, \ldots, d_{nm})$

are defined by

$$v_1 = u_o^1 \;;\ldots v_n = u_{n-1}^j \;;\ldots v_{nm-n+1} = u_o^m \;; v_{nm} = u_{n-1}^m$$
$$y_1 = x_o^1 \;;\ldots v_n = x_{n-1}^j \;;\ldots y_{nm-n+1} = x_o^m \;; y_{nm} = x_{n-1}^m \qquad \textbf{(2)}$$
$$d_1 = q_1^1 \;;\ldots d_n = q_n^1 \;;\ldots d_{nm-n+1} = q_1^m \;;\ldots d_{nm} \; q_n^m$$

Thus for i = 1, …, nm we use the variables v_i and d_i instead of u_t^j and q_t^j in order to avoid any possible confusion with their iterate when we describe our algorithm.

With these notations, the dynamic constraints become, for $i_j \leq i \leq nj$

$$g_i(v) = -\sum_{k=i_j}^{i} v_k + \sum_{k=i_j}^{i} d_k - x_o^j \leq 0$$
$$g_{nj}(v) = -\sum_{k=i_j}^{nj} v_k + \sum_{k=i_j}^{nj} d_k - x_o^j = 0 \qquad \textbf{(3)}$$

Together with the nonegativity of control variables,

$$g_t(v) = - v_{i-nm} \leq 0 \;; I = nm + 1, \ldots, 2nm \qquad \textbf{(4)}$$

these mn contraints define the domain of realisable solutions, or admissible domain :

$$D = \{v \in \mathbb{R}_+^{mn} \;/\; g_i(v) \leq 0 \;; i = 1, \ldots, 2nm \text{ with } g_{jn}(v) = 0 \;; j = 1, \ldots, m\} \qquad \textbf{(5)}$$

Let

$$D_\sigma = \{v = (v_i) \in D \;/\; v_i = \sigma_{i+1,k}^j ,\; i < k \leq nj\} \cup \{0\} \qquad \textbf{(6)}$$

In the sequel, we assume

$$x_o^j = 0 \;; \forall j = 1, \ldots, m.$$

$$(P) \qquad \min_{v \in D} J(v) \qquad \textbf{(7)}$$

We are interested in the following problem where $J : \mathbb{R}_+^{mn} \longrightarrow \mathbb{R}$ is defined by

$$J(v) = \sum_{t=1}^{n} \sum_{j=1}^{m} \left[c_{t-1}^j (u_{t-1}^j) + f_t^j (y_t^j) \right] \qquad \textbf{(8)}$$

with :

c_{t-1}^j concave, non negative, non decreasing, with continuous first, and second order partial derivatives; f_t^j affine, non decreasing, non negative functions.

2. Existence of solutions

Theorem 1: There exists an optimal solution to problem (P).

Proof: In $\mathbb{R}^{mn}$ endowed with the Euclidean scalar product, the domain D is non empty, since it contains the control v = d. The non negativity constraints, and the conditions $g_{nj}(v) = 0$ imply

$$0 \leq u_i^j \leq \sum_{t=1}^{n} q_t^j \;; j = 1, .., m \qquad \textbf{(9)}$$

from which we conclude that D is bounded.

By continuity of the functions g_{ij}, the half spaces determined by the hyperplanes defined by these functions are closed. Since D is an intersection of closed half-spaces, it is closed and bounded in a finite dimensional space, hence it is compact. On the other hand, the functions f_t^j are continuous, and so are the functions c_t^j (since - c_t^j being a proper convex function on a finite dimensional space, is continuous on the interior of its effective domain). Hence J is continuous, as a sum of continuous functions. Thus, by Weirstrass theorem, J being continuous on a compact set, reaches its minimum on D, i.e. (P) has an optimal solution.

3. Solution method

The idea is to use:

- the concavity properties of the functions,
- the characteristics of the optimal solution given by Wagner-Within theorem,
- a modified gradient projection,

to design a descent method in order to build a sequence $v^o, v^1, \ldots, v^k, v^{k+1}, \ldots$ of iterates, in such a way as to reduce the value of J while remaining inside D. So that the objective fonction decreases until an optimal solution is reached.

We will take the two following theorems into account :

Theorem 2 : For $v^k \in$ D, let A^k be the matrix of active constraints on this point.

$$-r^k = -\nabla J(v^k) + [(A^k)^t (A^k(A^k)^t)^{-1} A^k]\, \nabla J(v^k) \quad \textbf{(10)}$$

which points towards the opposite direction to the orthogonal projection of the $\nabla J(v^k)$ on the tangent supspace at v^k is a direction of descent.

Theorem 3 (Wagner Within)} :
If there exists an optimal solution to problem (P) , then there exists an optimal solution ($\hat{u}_t^j$) such as

$$\hat{x}_t^j \; \hat{u}_t^j = 0 \; ; \forall\, t = 1, \ldots, n \; ; \forall j = 1,\ldots, m \quad \textbf{(11)}$$

4. Algorithm

We exhibit our algorithm with the index i such that $n(j-1) + 1 \le i \le jn$ and $1 \le j \le m$.

Algorithm

Step 0. Initialization.
Set the initial control $v^o \equiv d$
Determine the set I^o of the active constraints at v^o.

Step 1. Projected gradient.
Determine the matrix A whose lines correspond to the contraints $i \in I^k$
Determine projection matrix
$P = I - A^t [A A^t]^{-1} A$
Set the direction
$r^k = -P\, grad(J(v^k))$.
If $r^k = 0$ then go to step 4
Else continue

Step 2. New v vector
For $i = n(j-1) + 1$, to jn do

a. If $r_i^k = 0$ then $v_i^{k+1} = v_i^k$
b. If $r_i^k > 0$ let $v_i^{k+1} = v_i^k$ and $l = i + 1$
while $r_l^k \le 0$ do
if $r_l^k < 0$ do $v_i^{k+1} = v_i^{k+1} + v_l^k$
let $v_\ell^{k+1} = 0$
else $r_l^k = 0$ do $v_l^{k+1} = v_l^k$
let $l = l + 1$
end (while)
c. let $i = l$ if $i < jn$ return to **b**
else ($i = jn$) **continue**

Step 3. Constraints reset
Let $k = k + 1$
Determine the set I^k of active constraints at v^k and return to Step 1.

Step 4. Coefficients reset
Compute the values of the coefficients
$\mu = -[AA^t]^{-1} A\, grad(J)(v^k)$
If $\mu^i \ge 0$; $\forall$ i
END
Else let μ^i be the coordinate of μ with the smallest negative value then do $I^k = I^k - \{i\}$ and return to step 1.

Remark : ∗ By construction of the current vector, we have

$$\sum_{i=1}^{n} v_i^{k+1} = \sum_{i=1}^{n} v_i^k \quad \text{et } v_i^{k+1} \ge 0 \; ; i = 1, \ldots, nm$$

since every amount added to one of the components is subtracted from another one with larger index; conversely, every amount subtracted from some component is added to another component with smaller index.

* Since the constraints are linear, the matrices of active constraints at iterations k and k + 1, differ just by one row. So a projection matrix is always computed from the precedent.
* As in projected gradient method rg A = card I(v), so AA^t has an inverse.
* In projected gradient method, since $y_{n(j-1)+1} = 0$ and $g_n(v) \le 0$, we always have $r_{n(j-1)+1} \ge 0$ and $\forall\ r_j < 0$, there exists an associated $r_i > 0$ with $i \le j$.

5. Results

Theorem 4: There exists an optimal solution in D_σ.

Proof: From Wagner Within theorem there exists an optimal solution such as:

$$\hat{y}_i\ \hat{v}_i = 0\ \forall\ i \in [i_j, nj]$$

Let m = 1, it will be the same proof for every fixed m. Let us proof that if $v = (v_1, \ldots, v_n)$ is an admissible solution such that $v_i\, y_i = 0$, $\forall\ i = \{1,\ldots, n\}$; then $v_i = 0$ or $v_i = \sigma^1_{i+1,k}$ with $k \in \{i+1,\ldots, n\}$. From constrants **(3)** and **(4)** $v_i = 0$ or $v_i \ge \sigma^1_{i+1,i+1}$ else v is not admissible.

By contradiction assume that

$v_i\, y_i = 0$; $\forall$ i = 1, …, n with $v_i \ne 0$ and $v_i \ne \sigma^1_{i+1,k}$ $\forall$ k $\in$ {i + 1, …, n} that is, assume $v_i \ne 0$ and $\exists\ \ell \in \{i + 1,\ldots,n\}$ such that $\sigma^1_{i+1,\ell} < v_i < \sigma^1_{i+1,\ell+1}$.

$v_i \ne 0 \quad \Rightarrow y_i = 0$ since $v_i\, y_i = 0$; $\forall$ i

$v_i > \sigma^1_{i+1,\ell} \quad \Rightarrow y_{i+1} = y_i + v_i - d_{i+1} = v_i - d_{i+1} > \sigma^1_{i+1,\ell} - d_{i+1} > 0$

$\Rightarrow y_{i+1} \ne 0 \Rightarrow v_{i+1} = 0$

$v_i > \sigma^1_{i+1,\ell} \quad \Rightarrow y_{i+1} \ne 0 \Rightarrow v_{i+1} = 0$

$\Rightarrow y_{i+2} \ne 0 \Rightarrow v_{i+2} = 0$

$\vdots$

$\Rightarrow y_\ell \ne 0 \Rightarrow v_\ell = 0$

$$y_{\ell+1} = v_i + v_{i+1} + \ldots + v_\ell - d_{i+1} - \ldots - d_{\ell+1} = v_i + 0 - \sigma^{\ell+1}_{i+1}$$

$v_i < \sigma^1_{i+1,\ell+1}$ implies $y_{\ell+1} < 0$ which contradicts the fact that $(y_1, \ldots, y_n) \in \mathbb{R}^n_+$ so $v_i = 0$ or $\exists$ k $\in$ {i + 1, …, n} such that $v_i = \sigma^1_{i+1,k}$.

According to Wagner Within theorem there exists an optimal solution v^* such that $v^*_i = 0$ or $v^*_i = \sigma^1_{i+1,k}$, and for every fixed m $\forall$j = i, …, m $\forall$i $\in$ [n(j – 1) + 1, jn] , $v^*_i = 0$ or $v_i = \sigma^j_{i+1,k}$, k $\in$ [i, n j] $\Rightarrow v^*_i \in D_\sigma$.

Remark 4: This proposition justifies the steps 0 and 2 of the algorithm. The moves are such that the components of the control variables take the form $\sigma^j_{i,k}$ while remaining in D_σ and so in D. In other methods and situations just very small moves take us out of the admissible domain. In addition it is more interesting to look for the optimal solution in D_σ than in D.

Lemma 1. If at iterate k, v^k is an admissible control vector, i.e $v^k \in D$, then, for the next vector v^{k+1} we have

$$\sum_{i=i_j}^{h} v_i^{k+1} \ge \sum_{i=i_j}^{h} v_i^{k}\ ;\ h = i_j, i_j+1,\ldots,j_n\ ;\ j = 1,\ldots,m \quad \textbf{(12)}$$

Proof: In this sum, for every fixed h, consider the set of all i ≤ h, and let L stand for those for which $v_i^{k+1} \ge v_i^k$ and L^c for the complement of L.

If $l \in L$, then $v_l^{k+1} \ge v_l^k$.

If $l \in L^c$, i.e. if $v_l^{k+1} < v_l^k$ then, by construction, $r_l^k < 0$; $v_l^{k+1} = 0$; $\exists\ l_o$ such that $l_o < l \le jn$ and $v_{l_o}^{k+1} \ge v_{l_o}^k + v_l^k$.

Thus $v_{l_o}^{k+1} + v_l^{k+1} \ge v_{l_o}^k + v_l^k$, with $l_o \le h$, and $l \le h$ hence

$$\sum_{i=i_j}^{h} v_i^{k+1} \ge \sum_{i=i_j}^{h} v_i^{k}.$$

Proposition 1. The sequence of iterates $v^o, v^1, .., v^k, v^{k+1}$, is a sequence of admissible control vectors.

Proof : We prove the statement by induction. For the initial control, we have $v^o \in D$. Suppose $v^k \in D$ i.e. $g_i(v^k) \leq 0$; i = 1, …, 2nm with $g_{jn}(v^k) = 0$; j = 1, …, m i.e.

$$\sum_{i=i_j}^{h} v_i^k \geq \sum_{i=i_j}^{h} d_i \quad h = 1, 2, ..., n \qquad \sum_{i=i_j}^{nj} v_i^k = \sum_{i=i_j}^{nj} d_i$$

$$\sum_{i=i_j}^{h} v_i^k \geq \sum_{i=i_j}^{h} d_i \quad \Rightarrow \quad \sum_{i=i_j}^{h} v_i^{k+1} \geq \sum_{i=i_j}^{h} d_i \text{ from}$$

Lemma 1 and

$$\sum_{i=i_j}^{nj} v_i^k = \sum_{i=i_j}^{nj} d_i \Rightarrow \sum_{i=i_j}^{nj} v_i^{k+1} = \sum_{i=i_j}^{nj} d_i$$

according to the remark. In addition, $v_i^{k+1} \geq 0$ for i = 1, 2, …, nm. Thus v^{k+1} satisfies all the constraints $g_i(v^{k+1}) \leq 0$; i = 1, …, 2nm with $g_{jn}(v^{k+1}) = 0$; j = 1, …, m, i.e. $v^{k+1} \in D$.

Proposition 2. Let r be the direction determined at step 1 of the algorithm. If, for the value i of the index, such that $r_i^k > 0$ there corresponds the value l > i of the index such that $r_\ell^k < 0$, then:

$$\frac{\partial J}{\partial v_i}(v^k) \leq \frac{\partial J}{\partial v_l}(v^k) \qquad \textbf{(13)}$$

Proof : By contradiction, assume this is not the case, i.e.

$$\frac{\partial J}{\partial v_i}(v^k) > \frac{\partial J}{\partial v_l}(v^k)$$

Since $r_i^k > 0$, then

$$r_i^k \frac{\partial J}{\partial v_i}(v^k) > r_i^k \frac{\partial J}{\partial v_l}(v^k)$$

Let δ be the direction such that

$$\delta_i = 0 \; ; \; \delta_l = r_i^k + r_\ell^k \; ; \; \delta_j = r_j^k \; ; \; \forall\, j \neq i \text{ et } j \neq l$$

then we have :

$$\delta_i \frac{\partial J}{\partial v_i}(v^k) + \delta_\ell \frac{\partial J}{\partial v_\ell}(v^k) = r_j^k \frac{\partial J}{\partial v_\ell}(v^k) + r_\ell^k \frac{\partial J}{\partial v_\ell}(v^k) < r_j^k \frac{\partial J}{\partial v_i}(v^k) + r_j^k \frac{\partial J}{\partial v_i}(v^k)$$

Since $\delta_j = r_j^k$; $\forall\, j \neq i$, $j \neq l$

$$\delta_j \frac{\partial J}{\partial v_j}(v^k) = r_j^k \frac{\partial J}{\partial v_j}(v^k)$$

This implies

$$\delta^t \nabla J(v^k) < r^k \nabla J(v^k)$$

which contradicts the optimality of the direction r^k given by the projected gradient method.

Proposition 3. At each iterate, we have:

$$(v^{k+1} - v^k)^t \nabla J(v^k) \leq 0 \quad \textbf{(14)}$$

Proof : By construction, to every l such that $r_l < 0$ there corresponds an i such that $r_i > 0$, and, to every i such that $r_i > 0$ there corresponds a set C(i) of indexes l such that $r_l < 0$, and

$$v_i^{k+1} = v_i^k + \sum_{l \in C(i)} v_l^k \text{ and } \frac{\partial J}{\partial v_i}(v^k) \leq \frac{\partial J}{\partial v_l}(v^k)$$

Multiplying the last inequality by v_l^k ; $l \in C(i)$, and taking the sum, we get

$$(v_i^{k+1} - v_i^k)\frac{\partial J}{\partial v_i}(v^k) = \sum_{l C(i)} v_l^k \frac{\partial J}{\partial v_i}(v^k) \leq \sum_{l C(i)} v_l^k \frac{\partial J}{\partial v_i}(v^k)$$

Which implies that

$$\sum_{i/r_i>0} (v_i^{k+1} - v_i^k)\frac{\partial J}{\partial v_i}(v^k) \leq \sum_{i/r_i>0}\sum_{l \in C(i)} v_l^k \frac{\partial J}{\partial v_l}(v^k)$$

i.e.

$$\sum_{i/r_i>0} (v_i^{k+1} - v_i^k)\frac{\partial J}{\partial v_i}(v^k) \leq \sum_{i/r_i \leq 0} v_l^k \frac{\partial J}{\partial v_l}(v^k) \qquad \textbf{(15)}$$

On the other hand, since $v_l^{k+1} = 0$ if $r_l < 0$ and $v_l^{k+1} = v_l^k$ if $r_l = 0$ then

$$\sum_{l/r_l<0} (v_l^{k+1} - v_l^k)\frac{\partial J}{\partial v_\ell}(v^k) = - \sum_{\ell/r_l \leq 0} v_\ell^k \frac{\partial J}{\partial v_\ell}(v^k) \qquad \textbf{(16)}$$

Taking the sum on both sides of (15) and (16) yields

$$\sum_{l/r_l>0}(v_l^{k+1}-v_l^{k})\ \frac{\partial J}{\partial v_i}(v^k) + \sum_{l/r_l\leq 0}(v_l^{k+1}-v_l^{k})\ \frac{\partial J}{\partial v_i}(v^k) \leq 0$$

i.e.

$$(v^{k+1} - v^k)^t\ \nabla J(v^k) \leq 0.$$

Proposition 4. The algorithm converges.
Proof : From the characterization of concave functions, the hypersurface lies below its tangent plane at each of its points, i.e. :

$$J(v^{k+1}) \leq J(v^k) + (v^{k+1} - v^k)^t\ \nabla J(v^k).$$

From Proposition 1 and 3, v^k, $v^{k+1} \in$ D and

$$(v^{k+1} - v^k)^t\ \nabla J(v_k) \leq 0$$

Which implies that $J(v^{k+1}) \leq J(v^k)$.

Thus, at each iterate k, we have $J(v^{k+1}) \leq J(v^k)$ moreover, by the existence theorem, the optimal solution v^* exists. Therefore, for every vk, we have : $J(v^k) \geq J(v^*)$.

D_σ is closed hence it is a part of R^n reduced to a finite number of points.

It is bounded; hence $D_\sigma \subset D$ and we have seen that D is bounded.

This implies D_σ is a compact set and the continuous function J has a minimum in D_σ.
Let $w^* \in D_\sigma$ such a $J(w^*) = \min_{v\in D_\sigma} J(v)$.

From theorem 4, there exists an optimal solution $v^* \in D_\sigma$.

$v^* \in D_\sigma \Rightarrow J(v^*) \geq J(w^*)$.

v^* optimal solution of (P) $\Rightarrow J(v^*) \leq J(w^*)$

$\Rightarrow J(v^*) = J(w^*)$

$v^* \in D_\sigma$ and $J(v^{k+1}) < J(v^k)$.

The sequence $\{(J(v^k))_k\}$ is decreasing in D_σ and bounded from below by $J(w^*)$ and therefore converges. D_σ is reduced to a finite number of points.

This implies that $J(w^*)$ will be reached in a finite number of iterations

$$J(v^k) \rightarrow J(w^*) = J(v^*).$$

and so the algorithm converges to the optimal solution.

6. Examples of application

6.1. Domains

1) These results may be applied to optimal control, in storage management :

- of a dam, where the power station is generally combined with a reservoir upstream and a river downstream, for the control of water releases, the optimal policy of fillings and the follow-up of water volumes in the reservoirs,
- of the use of reservoirs in irrigation systems, or by Water Boards,
- of continental fresh water, for the control and availability follow-up both in quantity and quality, in a context of scarcity or of non uniform space and time repartition,
- of oil or mineral ressources; the control of exploitation and refinery,
- of material ressources, ammunition, raw material, or in-process in a manufacturing factory or a business enterprise,
- of vehicule fleet, in transportation or distribution companies,
- of a set of turbines or electricity generators to be started according to the needs, and intensity of demand,
- of radio-active waste, or other toxic wastes from petrochemical plant, for the control of polluting emissions, and the control of toxic emissions in the atmosphere, rivers or lands.

2) The choice of concave functions is justified by the fact that in some cases, costs increase less than proportionally when quantities increase and in other cases, to discourage the use of available products in context of scarcity, or the production of toxic or polluting products, the costs increase less than proportionally.

In the first cases, we assume functions are concave, in the second cases we assume that the opposite functions are concave.

6.2. Application

Let us consider m products noted j = 1,…, m on a time horizon of n time periods t = 1, …, n.

We are interested in the determination of the optimal controls which will manage the storage in order to satisfy the demand $q_1^j, q_2^j, \dots, q_n^j$ for each product j.

The main variables of the model are :

q_t^j demand, volume of product j which may be dealt with during period t

u_t^j control (decision, policy, order, action, volume sent in the turbines, etc... of product j during period t

x_t^j storage state of product j during period t.

Assume initial storage level to be zero, i.e. $x_o^j = 0$; j = 1, 2, …, m, the dynamic equations for the evolution of storage are written as:

$$x_t^j = x_{t-1}^j + u_t^j - q_t^j \; ; \; t = 1, 2, \dots, n \quad \textbf{(17)}$$

$u_t^j \geq 0$ and $x_t^j \geq 0$.

For simplicity, we take n = 2, m = 2 and the functions $f_t^j = 0$.

$$q_1 = 5 \; ; q_2 = 3 \; ; q_3 = 2 \; ; q_4 = 7$$

$c_1^1 = \log(v_1+10)$; $c_2^1 = 4v_2$; $c_1^2 = 7$; $c_2^2 = 3v_4$

The objective function becomes

$$J(v) = \log(v_1 + 10) + 4v_2 + 7 + 3v_4 \quad \textbf{(18)}$$

and the gradient is $\nabla J(v) = \left(\frac{1}{v_1+10}, 4, \; 0, \; 3\right)$

The dynamic constraints are:

$g_1(v) = - v_1 + 5 \leq 0$; $g_2(v) = - v_1 - v_2 + 8 = 0$;

$g_3(v) = -v_3 + 2 \leq 0$; $g_4(v) = -v_3 - v_4 + 9 = 0$.

While the nonegativity constraints state :

$g_j(v) = - v_{j-4} \leq 0$; j = 5, 6, 7, 8.

Itération 0 : Let us choose the initial vector

$$V^o = (5, 3, ; 2, 7) \;\; ; \nabla (J(v^o)) = \left(\frac{1}{15} \;\; 4, 0, 3\right)$$

The active constraints being

$I(v^o) = (1, 2, 3, 4)$ we get

$$A = \begin{bmatrix} -1 & 0 & 0 & 0 \\ -1 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & -1 & -1 \end{bmatrix}$$

His yields the projection matrix

$$P^o = I - A^t[AA^t]^{-1} A = 0$$

$\mu = [[A\,A^t]^{-1} a](-\nabla J(v^o))$

$$= \begin{bmatrix} -1 & 1 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} -\frac{1}{15} \\ -4 \\ 0 \\ -3 \end{bmatrix} = \begin{bmatrix} -\frac{59}{15} \\ 4 \\ -3 \\ 3 \end{bmatrix}$$

$$\mu = (\mu^i) = \left(\frac{-59}{15}, 4, -3, 3\right)$$

μ^1 is the smallest negative value, so

$$A = \begin{bmatrix} -1 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & -1 & -1 \end{bmatrix}$$

$$P = \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} & 0 & 0 \\ -\frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\mu = (\mu^i) = \left(\frac{59}{30}, -\frac{59}{30}, 0, 0\right).$$

Itération 1 :

$r_1 > 0$; $r_2 < 0$; $r_3 = r_4 = 0 \Rightarrow v^1 = (8, 0, 2, 7)$

$$\Rightarrow I(v^1) = (2, 3, 4, 6)$$

$$\nabla (J(v^1)) = \left(\frac{1}{18}, \; 4, \; 0, \; 3\right)$$

$$A = \begin{bmatrix} -1 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & -1 & -1 \\ 0 & -1 & 0 & 0 \end{bmatrix}$$

$$P = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{2} & -\frac{1}{2} \\ 0 & 0 & -\frac{1}{2} & \frac{1}{2} \end{bmatrix}$$

$$\mu = (\mu^i) = \left(0, 0, \frac{3}{2}, -\frac{3}{2}\right).$$

Itération 2 :

$r_1 = r_2 = 0$; $r_3 = \frac{3}{2}$; $r_4 = -\frac{3}{2} \Rightarrow$

$v^2 = (8, 0, 9, 0) \Rightarrow I(v^2) = (2, 4, 6, 8)$

$$A = \begin{bmatrix} -1 & -1 & 0 & 0 \\ 0 & 0 & -1 & -1 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}$$

$$P = 0 \; ; \; \mu = (\mu^i) = \left(\frac{1}{18}, 0, \frac{75}{18}, 3\right).$$

P = 0 and all the μ^i are positives, the kuhn et Tucker conditions are satisfied,
$V^2 = (8, 0, 9, 0)$ is optimal.

7. Références

[1] H. Brezis: Analysis functional, Theory and Application, Dunod (1999).

[2] Céa J., Optimisation, théorie et algorithmes, Dunod, Paris, (1971).

[3] Ekeland I. , Temam et R. : Convex Analysis and Variational Problems, North-Holland (American Elseveier), (1976).

[4] Ekeland I. : Mathematical models in economics The Mathematical Sciences after the Year 2000. Beirut (1999) 37, 44, Word sci. Publisching, River Edge.

[5] Hager,W.W and Park,S. :The gradient projection method with exact line search. Journal of global optimization, 30(1): 103-118, (2004).

[6] Hiriart-Urruty, J.B. and Lemaréchal, C. : Convex Analysis and Minimization Algorithms, Volume I and II. Springer-Verlag, Berlin, (1996).

[7] Luenberger D.G. : Introduction to linear and non linear programming, Addison-Wesey, (1973).

[8] Minoux, M. : Programmation Mathématique, Théorie et Algorithmes, Tome 1 et 2, Dunod, (1983).

[9] Rosen, J.B. : The gradient projection method for non-linear programming, part 1: linear constraints, S.I.A.M. Journal, 8, (1960), p.181-217.

[10] Rockafellar R.T. : Convex Analysis, Princeton University Press, Princeton, New Jersey , (1970).

[11] Sakarovitch M. : Optimisation combinatoire, Hermann, (1984).

[12] Wagner, H. M., and T. Within : Dynamic Problems in the Theory of the Firm, Theory of Inventory Management, second ed. Princeton, University Press, Princeton, N.J, (1957).