THE CONVEX PROGRAMMING

Elena GELOVA[1], Aleksandar KRSTEV[1], Jordan ZIVANOVIK[1], Aleksandra STOJANOVA[1]

[1]*Assistant, UGD-Stip, Faculty of Computer Science, Macedonia, elena.gelova@ugd.edu.mk*

[1]*Assistant Professor, UGD-Stip, Faculty of Computer Science, Macedonia, aleksandar.krstev@ugd.edu.mk*

[1]*Full Professor, UGD-Stip, Faculty of Computer Science, Macedonia, jordan.zivanovik@ugd.edu.mk*

[1]*Assistant, UGD-Stip, Faculty of Computer Science, Macedonia, aleksandra.stojanova@ugd.edu.mk*

ABSTRACT

Convex programming is the simplest and best processed area of nonlinear programming. Many properties of linear programs are transmitted to the convex programs. In this paper properties of convex programs and methods for their solution like gradient method and method of convergence are listed, also an example of solving convex program is given.

KEYWORDS

Convex program, Convex set, Minimum, Target function.

1. Properties of convex programs

The problem in which we need to find the minimum point of a function $f^0(x)$ at a specified set F, is called program (or mathematical program).

When there are convex function and a convex set, then it is a convex program. A set in convex program is usually given implicitly in this way:

$$F = \{x \in R^n : f^1(x) \leq 0, f^2(x) \leq 0, \dots, f^p(x) \leq 0\}$$

where $f^1, f^2, \dots, f^p$ are convex functions. In this case convex program can be formulated as follows:

(K)

$$Min\, f^0(x)$$

$$w.l$$
$$f^1(x) \leq 0$$
$$f^2(x) \leq 0 \qquad (1)$$
$$\dots \dots \dots$$
$$f^p(x) \leq 0.$$

The function f^0 is called target function, while $f^1, f^2, \dots, f^p$ are limiting functions. ("Min" means the required minimum point of the function, while the "w.l" is short for the words "while limiting".) The set of allowed solutions can be written as:

$$F = \{x \in R^n : f^i(x) \leq 0, i = 1, \dots, p\}$$

in which case the program (K) has the form

(K)
$$Min\, f^0(x)$$
$$w.l$$
$$f^i(x) \leq 0, i = 1, \dots, p.$$

We say that a point x is allowed solution of the program (K) if x is part of F.

If $f^i, i = 1, \dots, p$ are convex functions, then the set $F = \{x \in R^n : f^i(x) \leq 0, i = 1, \dots, p\}$ is convex.

2. Methods for solving convex programs

2.1 Gradient method

Gradient method is a modification of the Cauchy's method with most steep decline for programs with constraints. For simplicity, assume that you need to solve convex program with linear constraints.

$$Minf(x)$$
$$w.l$$
$$Ax \leq b$$
$$x \geq 0 \qquad (2)$$

where f is a convex continuous differential function. We denote with F the set of allowed solutions

$$F = \{x : Ax \leq b, x \geq 0\}.$$

The method is iterative, i.e. the optimal solution x^* of the program (2) is the limit point of the sequence with approximation x^k, $\kappa = 0,1,2, \dots$ Any approximation x^k is obtained by solving a linear program and one-dimensional search.

Algorithm

1. Initial allowed approximation is selected $x^0 \in F$. The gradient $\nabla f(x^0)$ is calculated and the stopping rule is specified. Example, "sufficiently small" number $\varepsilon > 0$, with capacity to stop the algorithm when:

$$\|x^{k+1} - x^k\| < \varepsilon. \tag{3}$$

   for $\kappa = 0,1,2, \ldots$:

2. The linear program is solved.

$$\begin{aligned} &Min \nabla f(x^k) \\ &w.l \\ &Ax \leq b \\ &x \geq 0 \end{aligned} \tag{4}$$

The optimal solution is denoted by $\bar{x}^k$.

3. The one-dimensional program is solved.

$$\begin{aligned} &Min f(x^k + \lambda(\bar{x}^k - x^k)) \\ &w.l \\ &0 \leq \lambda \leq 1. \end{aligned}$$

   (here f is searched on linear segment that spans x^k and $\bar{x}^k$).
   The optimal solution is denoted by λ_k.

4. The new approximation is calculated.

$$x^{k+1} = x^k + \lambda_k(\bar{x}^k - x^k).$$

5. If the stopping rule is satisfied, the process is halted, x^{k+1} is permissible approximation of the optimal solution. Otherwise, the process is backed to step 2, $\nabla f(x^{k+1})$ is calculated and the algorithm continues with new x^{k+1}.

The method is graphically described in figure 1.1 Since (4) is a linear program, its solution x^k can be taken as extreme point of the set of permissible solutions F.

2.2 Method of convergence

The proof that some method really converges to its solution is not usually simple. This assertion will be demonstrated on the gradient method. We want to prove the following result:

Let $f(x)$ is convex continuous differential function and if the set of allowed solutions F is limited, than every limit point x^* from the sequence $\{x^k\}$ of calculated Gradient method is a global minimum program (1).

Since $x^*$is a limit point of the sequence, and the set $\{x^k\}$ is limited (and closed), there is a subsequence $\{x^{k_l}\}$ of sequence $\{x^{k_l}\}$ with property:

$$\lim_{l\to\infty} x^{k_l} = x^*.$$

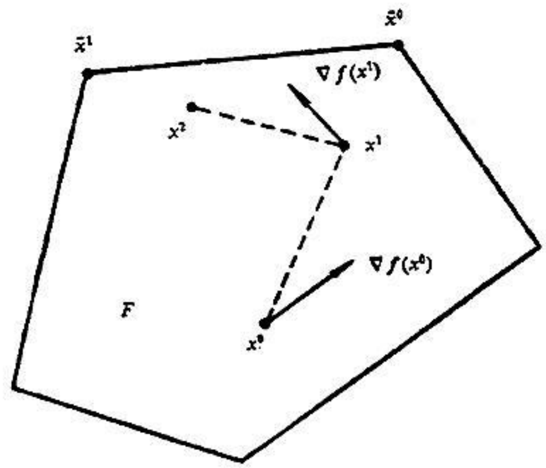

Figure1.1 Gradient method

For each x^{k_l}, the method determines $\bar{x}^{k_l}$ as solution of the linear program (4). Without loss of simplicity we assume that sequences $\{\bar{x}^{k_l}\}$ and $\{\bar{x}^{k_l+1}\}$ are convergent. Their limit point are denoted as $\bar{x}^*$ and x^{**}. Since $\bar{x}^{k_l+1}$ minimizes f on the interval $[x^{k_l},\bar{x}^{k_l}]$, and f is continuous, than x^{**} minimizes f on the interval $[x^*,\bar{x}^*]$. So

$$f(x^*) \le f(x) \text{ for each } x = x^* + \lambda(\bar{x}^* - x^*), 0 \le \lambda \le 1. \quad (5)$$

Specially, $f(x^{**}) \le f(x^*)$. But $f(x^{**}) < f(x^*)$ is impossible because $\{f(x^{k_l})\}$ can only decline, so $f(x^{**}) = f(x^*)$ and (5) becomes

$$f(x^*) \le f(x) \text{ for each } x = x^* + \lambda(\bar{x}^* - x^*), 0 \le \lambda \le 1. \quad (6)$$

Now we write the well-known Taylor's formula for point x:

$$f(x) = f(x^*) + \lambda\nabla f(x^*)(\bar{x}^* - x^*) + O(\lambda) \quad (7)$$

where $O(\lambda)$ is the "residual function" of property

$$\lim_{\lambda\to 0} \frac{O(\lambda)}{\lambda} = 0. \quad (8)$$

This formula, due to (6) pulls

$$\lambda\nabla f(x^*)(\bar{x}^* - x^*) + O(\lambda) \ge 0.$$

After splitting with $\lambda > 0$ and $\lambda \to 0$ have

$$\nabla f(x^*)(\bar{x}^* - x^*) \geq 0. \tag{9}$$

Inequation (9) used only for two specific points $\bar{x}^*$ and x^*. In order to join the inequation allowed arbitrary decisions, first need to apply that

$$\nabla f(x^{k_l})(x - x^{k_l}) \geq \nabla f(x^{k_l})(\bar{x}^{k_l} - x^{k_l}) \text{ for each } x \in F.$$

(This is true because $\bar{x}^{k_l}$ a minimal solution of program (4).) For f a continuous differential function, in border case when $l \to \infty$, the above expression gives for each $x \in F$:

$$\nabla f(x^*)(x - x^*) \geq \nabla f(x^*)(\bar{x}^* - x^*)$$
$$\nabla f(x^*)(x - x^*) \geq 0 \tag{10}$$

optimality of point x^* is now imminent: For all $x \in F$ have:

$$f(x) \geq f(x^*) + \nabla f(x^*)(x - x^*)$$

because the convex function f and

$$f(x) \geq f(x^*)$$

by (10). So x^* is the global minimum.

We will look an example in which will use Kuhn-Tucker conditions and to enumerate them. Primary Kuhn-Tucker condition of optimality.

We denote by $\mathcal{P}(x^*)$ the set of active constraints for point x^*, i.e

$$\mathcal{P}(x^*) = \{i \in \mathcal{P}: f^i(x^*) = 0\}.$$

The primary condition for optimality of programs (K) can be defined as follows: The accepted solution x^* is optimal if and only if, the system of linear inequation

$$\nabla f^0(x^*)d < 0$$
$$\nabla f^i(x^*)d \leq 0, i \in \mathcal{P}(x^*) \tag{11}$$

no solution d.

Dual-Kuhn Tucker condition of optimality.
Dual conditions of optimality are derived from primary conditions of optimality. Without loss of completeness (and to avoid the introduction of double indices) we can assume that the first r constraints are active at the point x^*. The accepted solution x^* is optimal if and only if, there are nonnegative numbers $\lambda_i \geq 0, i \in \mathcal{P}(x^*)$ such that

$$\nabla f^0(x^*) + \sum_{i \in \mathcal{P}(x^*)} \lambda_i \nabla f^i(x^*) = 0.$$

Example 1. You need to solve the program

$$Minf(x) = (x_1 - 2)^2 + (x_2 - \tfrac{1}{2})^2 + x_1^2 x_2^2$$
$$p.o.$$
$$0 \leq x_1 \leq 1.$$
$$0 \leq x_2 \leq 1.$$

First we show that the target function $f(x)$ is convex on set of allowed solutions F. Heseo's matrix of function is

$$\nabla^2 f(x) = 2\begin{bmatrix} 1 + x_1^2 & 2x_1x_2 \\ 2x_1x_2 & 1 + x_2^2 \end{bmatrix}.$$

Minor diagonal of that matrix are nonnegative on F, because it is

$$1 + x_1^2 > 0$$
$$1 + x_2^2 > 0$$
$$(1 + x_1^2)(1 + x_2^2) - 4x_1^2x_2^2 = 1 + x_1^2 + x_2^2 - 3x_1^2x_2^2 \geq 0.$$

Matrix $\nabla^2 f(x)$ is a positive samidefined on F.

This drew f to be a convex function on F.

If you start the iteration for example from $x^0 = (1{,}1)^T$, the optimal solution is found after only three iterations shown in Table 1.2.

Table 1.2

k	x_1^k	x_2^k	λ_k	$f(x^k)$
0	1,0000	1,0000	0,3379	2,2500
1	0,6621	0,6621	1,0000	2,0084
2	1,0000	0,0000	0,2500	1,2500
3	1,0000	0,2500	0,0000	1,1250

Approximation $x^0, x^1, x^2, x^3 = x^*$shown on Figure 1.3

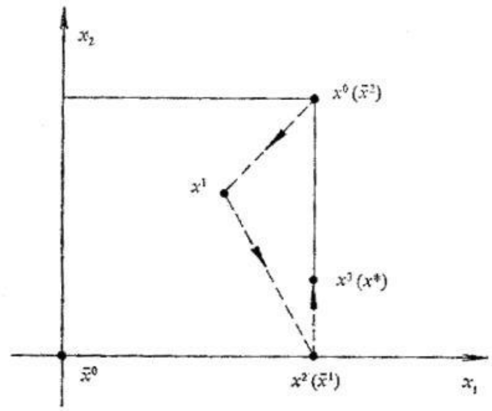

Figure 1.3

Optimality of point $x^* = \begin{bmatrix} 1 \\ \frac{1}{4} \end{bmatrix}$ you can check using the Kuhn - Tucker condition.

First, constraints are written in a manner that meets the program (K):

$$\begin{aligned} f^1(x) &= x_1 - 1 \leq 0 \\ f^2(x) &= -x_1 \quad\ \leq 0 \\ f^3(x) &= x_2 - 1 \leq 0 \\ f^4(x) &= -x_2 \quad\ \leq 0. \end{aligned}$$

In the point $x = x^*$ only the first limit is active, i.e $\mathcal{P}(x^*) = \{1\}$. Therefore Kuhn-Tucker dual system has the form

$$\begin{gathered} \nabla f(x^*) + \lambda_1 \nabla f^1(x^*) = 0 \\ \lambda_1 \geq 0. \end{gathered} \tag{12}$$

$$\nabla f(x^*) = \left[2(x_1 - 2) + 2x_1x_2^2,\ 2\left(x_2 - \tfrac{1}{2}\right) + 2x_1^2x_2\right] |x = x^*$$

$$= \left[-\tfrac{15}{8},\ 0\right]$$

$$\nabla f^1(x^*) = [1,0]$$

Kun- Tucker system (12), after the transposition of the vector will look like this:

$$\begin{gathered} \begin{bmatrix} -\frac{15}{8} \\ 0 \end{bmatrix} + \lambda_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \\ \lambda_1 \geq 0. \end{gathered}$$

Obviously, this system has a solution $\left(\lambda_1 = \frac{15}{8}\right)$. Conclusion: $x_1^* = 1, x_2^* = \frac{1}{4}$ is really optimal solution.

CONCLUSION

When solving convex programs differentiation continuous target functions and arbitrary convex set F of permissible solutions, convex of set F must not be ignored, because otherwise we ran the approximation x^k that is not in the set F. Important properties of Gradient method is to all approximations of the optimal solution is found in the set of allowed solutions. Gradient method is slow but generally reliable method.

REFERENCES

[1] Dantzing G.B., Linear programming and extensions, Princenton University Press,

[2] Princenton, New Jersey, 1963[2] Dr. Jovan J.Petrić , Operaciona istraźivanja, Naučna knjiga, Beograd 1989

[3] Sanjo Zlobec, Jovan Petrić, Nelinearno programiranje, Naučna knjiga, Beograd, 1989

[4] P.Venkataraman, Applied Optimization with Mat lab Programming, John Wiley and Sons,inc. 2002