Math. Operationsforsch. Statist., Ser. Optimization, Vol. 10 (1979) No. 3, 317–323

A Barrier Method with Arbitrary Starting Point[1]

W. GAUL[2] and J. HARTUNG[3]

Summary. To solve a nonlinear optimization problem a barrier method is given that may start in an arbitrary point and the sequential solutions of which can be taken as starting points for solving the subsequent sequential problems, respectively. At each step the violation of the constraints becomes diminished in such a way that after a finite number of steps the interior of the feasible region of the original program is reached.

1. Introduction

Widely used methods for solving a nonlinear constrained extremal problem are the various 'sequential unconstrained' optimization algorithms, cf. for instance [1]–[6].

Exterior penalty methods produce sequential solutions that are infeasible and are feasible only in the limit. Interior penalty methods like the usual barrier methods yield sequential solutions that are admissible but the disadvantage is that a starting point in the interior of the feasible set has to be determined which can take as long as it takes to solve the original problem, cf. [4, p. 213]. In [3, p. 51] a so-called regularized interior point penalty method is given that may start outside the feasible region but it can not be guaranteed that the sequential solutions become feasible within a finite number of steps and it even may also happen that at each step a new starting point has to be computed.

The method proposed here possesses all favoured properties of the usual barrier methods, allowing however to start at an arbitrary point not necessarily in the constrained region. The sequential solutions at each step are admissible points of the sequential programs of the subsequent step, and after a finite number of steps all sequential solutions are interior to the feasible region of the original program. The sequential programs can be solved by algorithms for unconstrained extremal problems. The method can be chosen in such a way that at each step the violation of the restrictions becomes strictly diminished.

---

1 This research has been supported by Sonderforschungsbereich 72, University of Bonn.
2 Universität Bonn, D-5300 Bonn, Wegelerstr. 6, BRD.
3 Universität Münster, D-4400 Münster i. W., Roxeler Str. 64, BRD.

2. The method

Let

$$f: \mathbf{R}^n \to \mathbf{R},$$
$$g_i: \mathbf{R}^n \to \mathbf{R}, \quad i=1, \ldots, m,$$

be continuous real valued functions and

$$X := \{x \mid x \in \mathbf{R}^n, \quad g_i(x) \geqq 0, \quad i=1, \ldots, m\},$$
$$X^0 := \{x \mid x \in \mathbf{R}^n, \quad g_i(x) > 0, \quad i=1, \ldots, m\}.$$

The optimization problem

$$(X, f): \text{minimize} \quad f(x)$$
$$\text{subject to: } x \in X$$

is solved by a sequence of programs

$$(X_\mu, F_\mu): \text{minimize} \quad F_\mu(x)$$
$$\text{subject to: } x \in X_\mu$$

where for $\mu \in \mathbf{N}$

$$X_\mu := \{x \mid x \in \mathbf{R}^n, \quad g_i(x) > -[-g_i(x_{\nu(\mu)})]_+ - b_i(\nu(\mu)), \quad i=1, \ldots, m\} \tag{1}$$

and with

$$G(\mu, x) := \sum_{i=1}^{m} \left(\left[1 - p_i(\mu) \log \left(\frac{g_i(x)}{b_i(\nu(\mu)) + [-g_i(x_{\nu(\mu)})]_+} + 1 \right) \right]_+ \right)^2$$

the sequential functionals F_μ are given by

$$F_\mu(x) := \begin{cases} f(x) + p_0(\mu) G(\mu, x), & \text{if } x \in X_\mu \\ \infty, & \text{otherwise} \end{cases} \tag{2}$$

where $[y]_+ := \max\{y, 0\}$, for $y \in \mathbf{R}$,

$$\nu(1) := 0, \quad \nu(\mu+1) \in \{\nu(\mu), \quad \nu(\mu)+1, \ldots, \mu\}. \tag{3}$$

$$b_i(\nu(\mu)) \in \mathbf{R}_+, \qquad b_i(\nu(\mu)) + [-g_i(x_{\nu(\mu)})]_+ > 0, \tag{4}$$
$$\sum_{\substack{\nu(\mu) \\ \mu \in \mathbf{N}}} b_i(\nu(\mu)) < \infty, \quad i=1, \ldots, m,$$

$$p_i(\mu) \in \mathbf{R}_+ \setminus \{0\}, \quad \lim_{\mu \to \infty} p_i(\mu) = \infty, \quad i=0, 1, \ldots, m. \tag{5}$$

x_μ is a δ-optimal solution of (X_μ, F_μ), $\delta > 0$ (6)
(i.e. $x_\mu \in X_\mu$ and $F_\mu(x_\mu) \leqq \inf\{F_\mu(x) \mid x \in X_\mu\} + \delta$),

x_0 is arbitrary starting point in $\mathbf{R}^n$. (7)

The method is based on the following

Theorem: *Let*

f be bounded below on (8)

$$X_0 := \{x \mid x \in \boldsymbol{R}^n,\ \ g_i(x) \geqq -[-g_i(x_0)]_+ - \sum_{\substack{\nu(\mu)\\ \mu\in N}} b_i(\nu(\mu)),\ \ i=1,\dots,m\},$$

$$\inf_{x\in X} f(x) = \inf_{x\in \mathrm{cl} X^0} f(x), \quad X^0 \neq \varnothing, \tag{9}$$

then

(i)

a) *there exists a δ-optimal solution x_μ of (X_μ, F_μ), $\delta > 0$,*

b) *$F_\mu(x)$ possesses the barrier property on X_μ,* i.e.

$$F_\mu(x) \to \infty \quad as \quad x \to \mathrm{bd}\, X_\mu\,, \quad x \in X_\mu\,,$$

and

c) *if $x \in X^0$ there is $F_\mu(x) = f(x)$ for μ sufficiently large, $\mu \in \boldsymbol{N}$,*

(ii)

a) *there exists a constant $M \in \boldsymbol{N}$ so that $\mu \geqq M$, $\mu \in \boldsymbol{N}$, implies $g_i(x_\mu) > 0$,*

b) $\inf\limits_{x\in X} f(x) \leqq \liminf\limits_{\mu\to\infty} f(x_\mu) \leqq \limsup\limits_{\mu\to\infty} f(x_\mu) \leqq \inf\limits_{x\in X} f(x) + \delta$,

(iii)

if $\{x_\mu\}_{\mu\in N}$ has an accumulation point $\hat{x}$, so $\hat{x}$ is a δ-optimal solution of (X, f),

(iv)

if there is a number $\lambda > 0$ such that

$$U(\lambda) := X \cap \{x \mid x \in \boldsymbol{R}^n,\ f(x) \leqq \inf_{x\in X} f(x) + \delta + \lambda\}$$

is compact, the sequence $\{x_\mu\}_{\mu\in N}$ has accumulation points.

Proof: (i) One has $\varnothing \neq X^0 \subset X \subset X_1 \subset X_0$ because of (4) and (9). So, with (8) $f(x)$ is bounded below on X_1 and there exists $\inf\limits_{x\in X_1} F_1(x)$, resp. a δ-optimal solution ($\delta > 0$) x_1 of (X_1, F_1) with $x_1 \in X_1$. Now, let there exist x_μ up to $\mu = j$ and denote by $k_1, \dots, k_{q_j}$ the different values of $\nu(\mu)$ for $\mu = 1, \dots, j$, $q_j \in \{1, \dots, j\}$ then

$$0 = k_1 < k_2 < \dots < k_{q_j} \leqq j - 1\,.$$

For $x \in X_\mu$, $k_{q_j} < \mu \leqq j$, $i = 1, \dots, m$ one has

$$\begin{aligned} g_i(x) &> -[-g_i(x_{k_{q_j}})]_+ - b_i(k_{q_j}) \\ &> -[-g_i(x_{k_{q_j-1}})]_+ - b_i(k_{q_j-1}) - b_i(k_{q_j}) \\ &> -[-g_i(x_0)]_+ - \sum_{r=1}^{q_j} b_i(k_r) \end{aligned} \tag{10}$$

and so $\varnothing \neq X^0 \subset X \subset X_\mu \subset X_0$.

If $\nu(j+1) = \nu(j)$ (which means $q_{j+1} = q_j$) then $X_{j+1} = X_j$ and nothing remains to be shown. For $\nu(j+1) = k_{q_{j+1}} > k_{q_j}$ (which means $q_{j+1} = q_j + 1$) one gets by (10) because of $x_{k_{q_{j+1}}} \in X_{k_{q_{j+1}}},\ k_{q_{j+1}} \leqq j$

$$-[-g_i(x_{k_{q_{j+1}}})]_+ - b_i(k_{q_{j+1}}) > -[-g_i(x_0)]_+ - \sum_{r=1}^{q_{j+1}} b_i(k_r) \tag{11}$$

$$i = 1, \ldots, m,$$

thus, $\emptyset \neq X^0 \subset X \subset X_{j+1} \subset X_0$, which by (8) and (9) implies the existence of $\inf_{x \in X_{j+1}} F_{j+1}(x)$, resp. a δ-optimal solution x_{j+1} of (X_{j+1}, F_{j+1}). So, part a) of (i) is shown.

Further

$$\emptyset \neq X^0 \subset X \subset X_\mu \subset X_0 \quad \text{for all} \quad \mu \in \boldsymbol{N}, \tag{12}$$

$$\beta_i := [-g_i(x_0)]_+ + \sum_{\substack{\nu(\mu)\\ \mu \in \boldsymbol{N}}} b_i(\nu(\mu)) \geqq [-g_i(x_{\nu(\mu)})]_+ + b_i(\nu(\mu)) > 0 \tag{13}$$

for all $\mu \in \boldsymbol{N}$, $i = 1, \ldots, m$.

Now let $y \in \operatorname{bd} X_\mu$ and $\{y_k\}_{k \in \boldsymbol{N}} \subset X_\mu$ a sequence converging to y. Then there is an $i_0 \in \{1, \ldots, m\}$ such that $g_{i_0}(y_k)$ tends to $-[-g_{i_0}(x_{\nu(\mu)})]_+ - b_{i_0}(\nu(\mu))$ from the right and so $G(\mu, y_k)$ tends to infinity as $k \to \infty$, which gives the barrier property of F_μ on X_μ, $\mu \in \boldsymbol{N}$, establishing (i), b).

Taking into account (13) one gets for $i = 1, \ldots, m$ and $x \in X^0$

$$\frac{g_i(x)}{b_i(\nu(\mu)) + [-g_i(x_{\nu(\mu)})]_+} + 1 \geqq \frac{g_i(x)}{\beta_i} + 1 > 1$$

$$p_i(\mu) \log\left(\frac{g_i(x)}{\beta_i} + 1\right) \to \infty, \quad \text{as} \quad p_i(\mu) \to \infty \quad \text{for} \quad \mu \to \infty$$

and so, for $p_i(\mu)$ sufficiently large

$$\left[1 - p_i(\mu) \log \left(\frac{g_i(x)}{b_i(\nu(\mu)) + [-g_i(x_{\nu(\mu)})]_+} + 1\right)\right]_+ = 0, \tag{14}$$

respectively, there is a $\mu'(x) \in \boldsymbol{N}$ such that

$$G(\mu, x) = 0 \quad \text{for} \quad \mu \geqq \mu'(x), \quad \mu \in \boldsymbol{N},\ x \in X^0, \tag{15}$$

which states (i), c).

(ii): Let be $x' \in X^0$ an α-optimal solution of (X, f), $\alpha > 0$, which exists because of (9), then with (12) and (15) we have with a $\mu'(x')$ for all $\mu \geqq \mu'(x')$, $\mu \in \boldsymbol{N}$,

$$\begin{aligned} -\infty < \inf_{x \in X_0} f(x) &\leqq f(x_\mu) \\ &\leqq f(x_\mu) + p_0(\mu)\, G(\mu, x_\mu) \\ &\leqq f(x') + p_0(\mu)\, G(\mu, x') + \delta \quad \text{(because } x_\mu \text{ is a } \delta\text{-optimal solution of } (X_\mu, F_\mu)) \\ &\leqq \inf_{x \in X} f(x) + \alpha + p_0(\mu)\, G(\mu, x') + \delta \\ &= \inf_{x \in X} f(x) + \alpha + \delta, \end{aligned} \tag{16}$$

thus

$$p_0(\mu)\, G(\mu, x_\mu) \text{ is uniformly bounded,} \tag{17}$$

and since $p_0(\mu)$ is assumed to be unbounded there exists a constant M such that

$$G(\mu, x_\mu) < 1 \quad \text{for} \quad \mu \geqq M\,, \quad \mu \in N\,, \tag{18}$$

implying (for $i=1, \ldots, m$ and $\mu \geqq M$)

$$1 - p_i(\mu) \log \left(\frac{g_i(x_\mu)}{b_i(\nu(\mu)) + [-g_i(x_{\nu(\mu)})]_+} + 1 \right) < 1\,,$$

or

$$\log \left(\frac{g_i(x_\mu)}{b_i(\nu(\mu)) + [-g_i(x_{\nu(\mu)})]_+} + 1 \right) > 0\,,$$

respectively,

$$g_i(x_\mu) > 0\,, \quad \mu \geqq M\,, \quad i=1, \ldots, m\,, \tag{19}$$

which gives (ii), a).

Now let

$$\mu'' := \max \{\mu'(x'), M\}$$

then we get for all $\mu \geqq \mu''$, $\mu \in N$, by (19) $x_\mu \in X$ and from (16)

$$\begin{aligned} \inf_{x \in X} f(x) &\leqq f(x_\mu) \\ &\leqq \inf_{x \in X} f(x) + \alpha + \delta\,, \end{aligned} \tag{20}$$

which yields

$$\begin{aligned} \inf_{x \in X} f(x) &\leqq \liminf_{\mu \to \infty} f(x_\mu) \\ &\leqq \limsup_{\mu \to \infty} f(x_\mu) \\ &\leqq \inf_{x \in X} f(x) + \delta\,, \end{aligned} \tag{21}$$

because α is chosen arbitrarily.

(iii): If $\hat{x}$ is an accumulation point of $\{x_\mu\}_{\mu \in N}$, so by (19) $\hat{x} \in X$ and since

$$\liminf_{\mu \to \infty} f(x_\mu) \leqq f(\hat{x}) \leqq \limsup_{\mu \to \infty} f(x_\mu)$$

we have by (21) that $\hat{x}$ is a δ-optimal solution of (X, f).

(iv): Choosing $\alpha = \lambda$, (20) and (19) give that $x_\mu \in U(\lambda)$ for μ sufficiently large. Thus $\{x_\mu\}_{\mu \in N}$ has accumulation points if $U(\lambda)$ is bounded.

This proves the Theorem. ∎

Corollary: *If X_0 is compact, then (X, f) and (X_μ, F_μ) have optimal solutions and the statements of the Theorem hold with $\delta = 0$.*

Remark: Letting $p_0(\mu) \to \infty$, $\mu \to \infty$, has the consequence that the sequence $\{F_\mu(x)\}_{\mu \in N}$ is unbounded for $x \in \mathrm{bd}X$. If we choose $p_0(\mu) \equiv p_0 > 0$, then $F_\mu(x) \leqq f(x) + p_0 m$ for all $x \in X$ and $\mu \in N$, which is of some numerical advantage. With $p_0 > c := \inf_{x \in X} f(x) - \inf_{x \in X_0} f(x) + \alpha + \delta$ the theorem remains valid, cf. (16), but to find such a p_0 may be as difficult as to solve the original problem.

Taking an arbitrary $p_0 > 0$ the theorem still holds with some minor alterations: (ii) is valid for convergent subsequences of $\{x_\mu\}_{\mu \in \boldsymbol{N}}$, when x_μ is a δ-optimal solution of (F_μ, X_μ) where $0 < \delta < 1$, or some compactness condition implying that $\{x_\mu\}_{\mu \in \boldsymbol{N}}$ is bounded has to be added, and in (iv) the set $U(\lambda)$ has to be exchanged by $\tilde{U}(\lambda) := U(\lambda) \cap \{x \mid g_i(x) \geqq \lambda_i, \lambda_i < 0, \quad i = 1, \ldots, m\}$.

Method

Step 1: Choose an arbitrary starting point $x_0 \in \boldsymbol{R}^n$, an accuracy parameter $\delta > 0$, the penalty parameters $p_0(1)$, $p_i(1)$, the barrier parameters $b_i(0)$, $i = 1, \ldots, m$, and determine a δ-optimal solution x_1 of (X_1, F_1).

Step μ: Let $\nu(\mu) = \nu(\mu-1)$ or make an updating according to $\nu(\mu) \in \{\nu(\mu-1)+1, \ldots, \mu-1\}$, choose $p_0(\mu)$, $p_i(\mu)$ and $b_i(\mu)$, $i = 1, \ldots, m$, and determine a δ-optimal solution x_μ of (X_μ, F_μ).

3. Conclusion

Under the conditions (4), (5), (8) and (9) the convergence properties of the method are given by the theorem.

For finding a δ-optimal solution of the sequential problem (X_μ, F_μ) solution algorithms for unconstrained minimization problems are applicable with the possible starting point $x_{\nu(\mu)}$. Because of the barrier property of $F_\mu(x)$ these algorithms don't leave the feasible region X_μ of the sequential problem and x_μ doesn't lie on the boundary of X_μ, thus they work as on the whole space.

$F_\mu(x)$ is differentiable if $f(x)$ and the $g_i(x)$ are differentiable, $F_\mu(x) = f(x)$ if the $g_i(x)$ are greater than a certain positive value, $i = 1, \ldots, m$, $f(x)$ need not be bounded on the whole space, after M steps the sequential solutions are strictly admissible for (X, f) and at the step μ with $\mu < M$ the method insures that the values $g_i(x_\mu)$, as long as $g_i(x_{\nu(\mu)})$ are not positive, are greater than $g_i(x_{\nu(\mu)}) - b_i(\nu(\mu))$. So from the sequential point of view it would be the best to choose $\nu(\mu) = \mu - 1$ and $b_i(\mu-1) = 0$ if $g_i(x_{\mu-1}) < 0$. Then at each step the values $g_i(x_\mu)$ are strictly increased as long as they are not positive.

Using this method up to step $\mu = M$ instead of that one given in [3, p. 195] for finding a starting point for an interior point penalty method has the additional advantage of producing not an arbitrary feasible point but one 'in the direction of a solution'.

Acknowledgement

The authors would like to thank the referees for their comments to improve an earlier version of the paper. In particular, the introduction of the parameter $p_0(\mu)$ with $p_0(\mu) \to \infty$, $\mu \to \infty$, was suggested by one of the referees. Without this parameter, resp. $p_0(\mu) \equiv 1$, the theorem holds in a form as indicated in the remark.

References

[1] Blum, E., und W. Oettli: Mathematische Optimierung. Springer-Verlag, Berlin-Heidelberg-New York (1975).
[2] Elster, K.-H., und C. Grossmann: Untersuchungen zu Regularitätsbedingungen bei den Barriere- und Zentrenmethoden. Math. Operationsforsch. Statist. 5 (1974), 191–206.
[3] Fiacco, A. V., and G. P. McCormick: Nonlinear Programming: Sequential Unconstrained Minimization Techniques. Wiley, New York (1968).
[4] Fletcher, R. (ed.): Optimization. Academic Press, London (1969).
[5] Lasdon, L. S.: An efficient algorithm for minimizing barrier and penalty functions. Math. Programming 2 (1972), 65–106.
[6] Lootsma, F. A.: A survey of methods for solving constrained minimization problems via unconstrained minimization. In: F. A. Lootsma (ed.): Numerical Methods for Non-Linear Optimization, pp. 313–347, Academic Press, London-New York (1972).

Zusammenfassung

Zur Lösung eines nichtlinearen Optimierungsproblems wird eine Barrierenmethode angegeben, die in einem beliebigen Punkt starten kann und deren sequentielle Lösungen als Startpunkte zur Lösung der jeweils nächstfolgenden sequentiellen Probleme genommen werden können. Die Verletzungen der Nebenbedingungen werden in jedem Schritt so vermindert, daß man nach endlich vielen Schritten das Innere des zulässigen Bereiches des Ausgangsproblems erreicht.

Résumé

On suggère une méthode de solution pour un problème de programmation non-linéaire qui débute dans un point arbitraire. Les solutions sequentielles qui en résultent constituent les points de départ pour la solution des problèmes sequentielles qui s'en suivent. La violation des constraintes peut être diminueé á chaque coup de façon que l'on atteigne dans un nombre fini de coups l'intérieur du domaine admissible du problème initiel.

Eingereicht bei der Redaktion: 4. 7. 1978; revidierte Fassung: 29. 11. 1978

Book Review

D. Plachky, L. Baringhaus, N. Schmitz: **Stochastik I.** Akademische Verlagsgesellschaft, Wiesbaden 1978, 246 S., 11 Abb., kart., DM 22.–

Die Autoren haben zu vorliegendem Band folgenden erklärenden Untertitel gewählt: „Eine elementare Einführung in Grundbegriffe der Wahrscheinlichkeitsrechnung und Statistik. Studienbuch für Studierende der Mathematik, Natur- und Wirtschaftswissenschaften ab 2. Semester". Unter den zahlreichen einführenden Lehrbüchern zur Wahrscheinlichkeitsrechnung kommt dieses Buch hinsichtlich der Stoffauswahl und der Diktion dem Kapitel „Diskrete Wahrscheinlichkeitsräume" des schönen Hochschultextes „Grundbegriffe der Wahrscheinlichkeitstheorie" von K. Hinderer sehr nahe. Allerdings wird diesem Problemkreis hier nun weitaus mehr Platz eingeräumt. Die einem Studienbuch wohl eigentümliche Redundanz betrifft vor allem die detaillierte Beweisführung und die bei fast jedem Beispiel praktizierte peinliche Genauigkeit der formalisierten Darstellung. Dadurch erweist es sich als ein brauchbares, eine anregende Vorlesung ergänzendes Arbeitsmittel für das Selbststudium. Im abschließenden Kapitel über Schätztheorie werden die Begriffe erwartungstreuer Schätzer, Suffizienz und Vollständigkeit für den diskreten Fall mathematisch präzis eingeführt. Als propädeutische Statistik kann dieser Teil noch nicht ausreichen.

H.-J. Girlich