COMPUTATIONAL MATHEMATICS
BANACH CENTER PUBLICATIONS, VOLUME 13
PWN – POLISH SCIENTIFIC PUBLISHERS
WARSAW 1984

CONSTRAINED OPTIMIZATION VIA UNCONSTRAINED MINIMIZATION

CH. GROSSMANN

Technische Universität Dresden, Sektion Mathematik,
DDR-8027-Dresden, Mommsenstr. 13, DDR

The aim of this paper is to derive some common aspects of different sequential unconstrained minimization techniques by the use of the close relation between the generated unconstrained subproblems and the perturbed primal optimization problems. Our approach leads also to certain new optimization methods.

1. Introduction; general theory

Let W, F denote reflexive Banach spaces and $X \subset W$ a closed subset. In the space F a partial ordering "$\geqslant$" is induced by a closed convex cone C according to

$$a \leqslant b \Leftrightarrow b-a \in C.$$

In this paper we investigate the following nonlinear programming problem

$$(1.1) \qquad f_0(x) \to \min! \quad \text{subject to} \quad x \in X,\ \boldsymbol{f}(x) \leqslant 0.$$

Here $f_0\colon X \to \boldsymbol{R}^1$ denotes a given continuous functional and $\boldsymbol{f}\colon X \to F$ a continuous operator. To shorten our notation we define the Banach space $H = \boldsymbol{R}^1 \times F$ with the related norm $\|\cdot\|_H = \sqrt{|\cdot|_{\boldsymbol{R}^1}^2 + \|\cdot\|_F^2}$ and the partial ordering induced by the cone $\boldsymbol{R}^1 \times C$. Any $h \in H$ is assumed to be represented by $h = (h_0, \boldsymbol{h})$ with $h_0 \in \boldsymbol{R}^1$ and $\boldsymbol{h} \in F$. We write

$$f(x) = \big(f_0(x), \boldsymbol{f}(x)\big) \quad \text{for any } x \in X.$$

The essential idea of the sequential unconstrained minimization technique consists in transforming the primal problem (1.1) by means of a functional $T\colon X \times Y \to \tilde{\boldsymbol{R}} := \boldsymbol{R}^1 \cup \{+\infty\}$ into auxiliary problem

$$(1.2) \qquad T(x, y) \to \min! \quad \text{s.t.} \quad x \in X.$$

Here y denotes a fixed parameter ranging over a given parameter set Y. The set Y can be of finite or infinite dimension. In the sequel we restrict attention to functionals T defined with use of a functional E: $Y \times H \to \tilde{\boldsymbol{R}}$ by the relation

$$T(x, y) = \inf\{E(y, v) \mid v \geqslant f(x)\} \tag{1.3}$$

for any $x \in X$, $y \in Y$. Then E is called the generating functional. It has to be remarked that most of the sequential unconstrained minimization techniques used in practice actually operate with auxiliary functionals T which are representable in the form (1.3).

As an example of a general penalty technique let us consider the method proposed by Babuška [2] for solving Poisson's equation

$$\begin{cases} -\Delta x = q \quad \text{in } \Omega \\ \text{with homogeneous boundary condition} \\ x = 0 \quad \text{on } \Gamma. \end{cases} \tag{1.4}$$

Ω denotes a bounded domain in $\boldsymbol{R}^n$ with a boundary $\Gamma \in C^\infty$. A weak solution $x \in W_2^1(\Omega)$ of the boundary problem (1.4) can be determined by solving the optimization problem

$$f_0(x) \equiv \int_\Omega \left[\sum_{i=1}^n \left(\frac{\partial x}{\partial \xi_i}\right)^2\right] d\xi - 2\int_\Omega x\cdot q\, d\xi \to \min! \tag{1.5}$$

$$\text{s.t.} \quad x = 0 \text{ on } \Gamma.$$

If we choose $F = L_2(\Gamma)$, then by imbedding theorems (see [14] e.g.) the restriction operator $f(x) := x|_\Gamma$ is known to be continuous. Now, let $C = \{0\}$ and $X = W = W_2^1(\Omega)$. Thus (1.5) forms a general optimization problem of the type (1.1). Taking $Y = \operatorname{int} R_+^1$ and $E(y, v) = v_0 + y\|v\|^2_{L_2(\Gamma)}$, we get the auxiliary problem

$$\int_\Omega \left[\sum_{i=1}^n \left(\frac{\partial x}{\partial \xi_i}\right)^2\right] d\xi - 2\int_\Omega xq\, d\xi + y\oint_\Gamma x^2\, d\sigma \to \min!$$

$$\text{subject to} \quad x \in W_2^1(\Omega)$$

which coincides with the technique described in [2].

In the same way we can apply the penalty technique to problems with mixed boundary conditions as those considered in [10] by defining an appropriate cone C.

Let

$$Q = \{v \in H \mid \exists\, x \in X \text{ such that } f(x) \leqslant v\} \tag{1.6}$$

denote the characteristic set related to the optimization problem (1.1).

If χ denotes the primal functional of (1.1) defined by

$$\chi(u) = \inf\{f_0(x) \mid x \in X,\ f(x) \leqslant u\} \quad \text{for any } u \in F$$

then the relations

$$Q \subset \operatorname{epi}\chi \quad \text{and} \quad \operatorname{cl}Q = \operatorname{cl}\operatorname{epi}\chi$$

with $\operatorname{epi}\chi = \{v \in H \mid v_0 \geqslant \chi(v)\}$ hold.

Now, for fixed parameters $y \in Y$ we consider the following comparison problem

$$E(y, v) \to \min! \quad \text{s.t.} \quad v \in Q. \tag{1.7}$$

The following lemma shows the close relation between the two optimization problems (1.2) and (1.7).

LEMMA 1. *Let $\varepsilon \geqslant 0$ be a real number and $y \in Y$ a fixed parameter. Suppose the auxiliary problem* (1.2) *posseses an approximate solution $x^\varepsilon(y) \in X$ such that*

$$T\big(x^\varepsilon(y), y\big) \leqslant T(x, y) + \varepsilon \quad \textit{for any } x \in X. \tag{1.8}$$

If a solution $v^\varepsilon(y)$ of the related problem

$$E(y, v) \to \min! \ \textit{s.t.}\ v \geqslant f\big(x^\varepsilon(y)\big), \tag{1.9}$$

exists then $v^\varepsilon(y)$ also forms an ε-solution of the comparison problem (1.7), *i.e. $v^\varepsilon(y) \in Q$ and*

$$E\big(y, v^\varepsilon(y)\big) \leqslant E(y, v) + \varepsilon \quad \textit{for any } v \in Q. \tag{1.10}$$

Proof. The constraints in the problem (1.9) and the relation $x^\varepsilon(y) \in X$ guarantee that $v^\varepsilon(y)$ belongs to Q. Now, we suppose that there exists $\tilde{v} \in Q$ such that

$$E(y, \tilde{v}) + \varepsilon < E\big(y, v^\varepsilon(y)\big). \tag{1.11}$$

Because $v^\varepsilon(y)$ solves (1.9), according to (1.3), the equality

$$E\big(y, v^\varepsilon(y)\big) = T\big(x^\varepsilon(y), y\big) \tag{1.12}$$

holds. From $\tilde{v} \in Q$ and the definition (1.6) of the characteristic set we get the existence of an $\tilde{x} \in X$ such that $f(\tilde{x}) \leqslant \tilde{v}$. This together with (1.3) leads to

$$T(\tilde{x}, y) \leqslant E(y, \tilde{v}). \tag{1.13}$$

Combining (1.10)–(1.13) we get a contradiction to the inequality (1.8). ∎

LEMMA 2. *Let $\varepsilon' \geqslant 0$, $\hat{x} \in X$ and $\hat{v} \in H$ be such that $f(\hat{x}) \leqslant \hat{v}$ and*

$$\hat{v} - \sigma e^0 \notin Q \quad \textit{for any } \sigma > \varepsilon' \tag{1.14}$$

where $e^0 = (1, 0) \in H$. Then $\hat{x}$ forms an ε'-solution of the perturbed problem

$$f_0(x) \to \min! \quad \textit{s.t.} \quad x \in X,\ f(x) \leqslant \hat{v}.$$

Proof. If the claim does not hold then a point $\tilde{x} \in X$ exists such that $f(\tilde{x}) \leqslant \hat{v}$ and $f_0(\tilde{x})+\varepsilon' < f_0(\hat{x})$. Set $\tilde{v} = (f_0(\tilde{x}), \hat{v})$. Then $\tilde{v} \in Q$ and

$$\tilde{v} = \hat{v} - \sigma e^0 \quad \text{with} \quad \sigma > f_0(\hat{x}) - f_0(\tilde{x}) > \varepsilon'$$

in opposition to (1.14). ∎

Remarks. 1. Investigating the behaviour of specific generating functionals E one can obtain estimates for the relation between the two accuracies ε and ε'. In particular, if E is represented in the form

$$E(y, v) = v_0 + e(y, v) \quad \text{for any } y \in Y,\ v \in H \tag{1.15}$$

with a given functional e: $Y \times F \to \tilde{\boldsymbol{R}}$ then $\varepsilon' \leqslant \varepsilon$.

2. Lemmas 1 and 2 generalize the well-known theorem of Everett [4]. This theorem follows by setting $\varepsilon = \varepsilon' = 0$ with

$$Y = C^* = \{y \in F^* \mid [y, u] \geqslant 0 \ \forall\, u \in C\}$$

and

$$E(y, v) = v_0 + [y, v].$$

Here $[y, u]$ denotes the value of a continuous linear functional y at the point u.

3. If the parameter $y \in Y$ can be chosen so that the point $v^\varepsilon(y)$ defined by Lemma 1 satisfies the conditions

$$v^\varepsilon(y) = 0, \tag{1.16}$$

$$v^\varepsilon(y) - \sigma e^0 \notin Q \quad \text{for any } \sigma > 0, \tag{1.17}$$

then according to Lemma 2 the point $x^\varepsilon(y)$ solves the given optimization problem (1.1). To guarantee (1.1), however, the auxiliary problem (1.2) has in general to be solved exactly, i.e. with $\varepsilon = 0$.

Now we describe certain aspects of the general duality theory (compare [8]). Let us introduce the following property:

(V) If $\tilde{v} \in Q$ and $\tilde{v} - \sigma e^0 \in Q$ for some $\sigma > 0$ then

$$\inf_{v \in Q} E(y, v) < E(y, \tilde{v}).$$

Condition (V) concerns the family of functionals $E(y, \cdot)$, $y \in Y$ as well as the given nonlinear programming problem (1.1). It should be remarked that (V) automatically holds if E can be represented in form (1.15).

Let us define

$$\tau(y) = \sup\{t \in \boldsymbol{R}^1 \mid E(y, te^0) \leqslant \inf_{v \in Q} E(y, v)\} \tag{1.18}$$

for any $y \in Y$. Then τ is a functional τ: $Y \to \bar{\boldsymbol{R}}$: $= \tilde{\boldsymbol{R}} \cup \{-\infty\}$. The following theorem characterizes weak duality and can be proved analogously to [6].

THEOREM 1. *Let property* (V) *hold. Then for any $x \in X$ with $f(x) \leqslant 0$ and for any $y \in Y$ the inequality*

$$f_0(x) \geqslant \tau(y)$$

holds.

According to Theorem 1 the optimization problem

$$\tau(y) \to \sup! \quad \text{s.t.} \quad y \in Y \tag{1.19}$$

can be interpreted as a dual problem with respect to the primal problem 1.1). Similarly to the well-known saddle point theorem for the Lagrangian

$$L(x, y) = f_0(x) + [y, f(x)], \qquad x \in X, \ y \in C^*,$$

we get the following theorem.

THEOREM 2. *Let the functional E be of type* (1.15) *with $e(y, 0) = 0$, for any $y \in Y$. Suppose that*

$$\sup_{y \in Y} T(x, y) = \begin{cases} f_0(x) & \text{if } x \text{ is feasible for (1.1)}, \\ +\infty & \text{otherwise}. \end{cases} \tag{1.20}$$

(If $(\hat{x}, \hat{y}) \in X \times Y$ is a saddle point of the functional T, i.e.

$$T(\hat{x}, y) \leqslant T(\hat{x}, \hat{y}) \leqslant T(x, \hat{y}) \qquad \text{for any } x \in X, \ y \in Y,$$

then $\hat{x}$ solves the primal problem (1.1) *and $\hat{y}$ solves the dual problem* (1.19), *and we have $f_0(\hat{x}) = \tau(\hat{y})$.*

The proof given in [8] can be used just as well for the general spaces considered here. Furthermore, in [8] it is shown that T has a saddle point $(\hat{x}, \hat{y})$ if and only if the primal functional χ of the problem (1.1) is Φ-subdifferentiable at 0 in a certain sense. In the finite dimensional case this can be ensured by the well-known second order optimality condition involving strict complementarity and linear independence of the active gradients.

Now we describe certain specific sequential unconstrained minimization techniques.

2. Shifting methods

Let $Y = F$ and

$$E(y, v) = v_0 + \varphi(v + y) \qquad \text{for any } v \in H. \tag{2.1}$$

Here φ denotes a certain convex Fréchet-differentiable and coercive functional $\varphi\colon F \to \tilde{\boldsymbol{R}}$. Furthermore, suppose there exists a forcing function δ (i.e. $\delta(t) \geqslant 0$ for any $t \geqslant 0$ and $\delta(t) \to 0$ implies $t \to 0$) such that

$$\varphi(y) \geqslant \varphi(y + u) - \varphi'(y + u)u + \delta(\|u\|) \tag{2.2}$$

for any $y,\ u \in F$. According to (1.3) the auxiliary functional T is given by

$$T(x, y) = f_0(x) + \inf_{v \geqslant f(x)} \varphi(v+y). \tag{2.3}$$

We assume that the optimization problem (1.1) is convex, i.e. the set X is convex and f_0, f are convex. We remark that the convexity of f is closely related to the partial ordering in F induced by the cone C.

ALGORITHM 1

step 1: Select $y^1 \in Y$ and set $k := 1$.

step 2: Define x^k as a solution of the auxiliary problem

$$T(x, y^k) \to \min! \quad \text{s.t.} \quad x \in X. \tag{2.4}$$

step 3: Find a solution $v^k \in F$ of the problem

$$\varphi(v+y^k) \to \min! \quad \text{s.t.} \quad v \geqslant f(x^k). \tag{2.5}$$

step 4: Set $y^{k+1} := y^k + v^k$ and $k := k+1$. Go to step 2.

THEOREM 3. *Let the primal functional χ of* (1.1) *be continuous at the point* 0. *Then any accumulation point of the sequence $\{x^k\}$ generated in algorithm* 1 *forms a solution of the optimization problem* (1.1).

Proof. Since the partial ordering in F was induced by a closed convex cone, the sets $\{v \in F \mid v \geqslant f(x^k)\}$ are nonempty closed and convex. Now, by the properties of the functional φ and because F is a reflexive Banach space we know that solutions v^k of the subproblems (2.5) do exist.

By Lemma 1 the point $\left(f_0(x^k),\ v^k\right) \in H$ solves the problem

$$v_0 + \varphi(y^k + v) \to \min! \quad \text{s.t.} \quad v \in Q. \tag{2.6}$$

The special structure of the generating functional E was taken into account here. The optimality condition (2.6) leads to

$$v_0 - f_0(x^k) + \varphi'(y^k + v^k)(v - v^k) \geqslant 0 \quad \text{for any } v \in Q. \tag{2.7}$$

Let us define $w^k \in H$ by $w^k = 0$ and

$$w_0^k = f_0(x^k) + \varphi'(y^k + v^k)v^k. \tag{2.8}$$

Now, step 4 of algorithm 1 guarantees that the Kuhn–Tucker conditions for the problem

$$\begin{aligned} & v_0 + \varphi(y^{k+1} + v) \to \min! \\ & \text{subject to} \\ & v_0 - f_0(x^k) + \varphi'(y^k + v^k)(v - v^k) \geqslant 0 \end{aligned} \tag{2.9}$$

are satisfied at the point w^k.

Because (2.9) is a convex optimization problem, the point w^k forms an optimal solution of the problem (2.9).

With (2.7) this results in the inequality

$$E(y^{k+1}, w^k) \leqslant \inf_{v\in Q} E(y^{k+1}, v)$$

and according to (1.18) we get

$$\tau(y^{k+1}) \geqslant E(y^{k+1}, w^k) - \varphi(y^k) = w_0^k + \varphi(y^{k+1}) - \varphi(y^k).$$

Now, using (2.2) we have

$$\begin{aligned}\tau(y^k) &\leqslant \tau(y^k) + \delta(\|v^k\|)\\ &= f_0(x^k) + \varphi(y^k+v^k) - \varphi(y^k) + \delta(\|v^k\|)\\ &\leqslant f_0(x^k) + \varphi'(y^k+v^k)v^k = w_0^k \leqslant \tau(y^{k+1}) \leqslant \chi(0)\end{aligned}$$

for any $k = 1, 2, \ldots$

Thus $\{\tau(y^k)\}$ forms a bounded and monotonically increasing sequence. Therefore $\{\tau(y^k)\}$ converges and

$$\lim_{k\to\infty} \delta(\|v^k\|) = 0.$$

Using the forcing property $(\delta(t)\to 0 \Rightarrow t\to 0)$, we get

$$\lim_{k\to\infty} v^k = 0.$$

By Lemmas 1 and 2, the points x^k solve the perturbed problems

$$f_0(x)\to\min! \quad \text{s.t.} \quad x \in X,\ f(x) \leqslant v^k.$$

This results in $f_0(x^k) = \chi(v^k)$, $k = 1, 2, \ldots$ From the continuity of f and the closedness of C and X we get $f(x^*) \leqslant 0$ and $x^* \in X$ for any accumulation point x^* of the sequence $\{x^k\}$.

Furthermore, $f_0(x^*) = \chi(0)$ holds because f_0 is continuous on X and χ is continuous at 0. This completes the proof. ∎

We remark that the statement of Theorem 3 also holds for any weak accumulation point of $\{x^k\}$. This follows from the weak closedness of closed convex sets and the weak lower semicontinuity of continuous convex functionals, similarly to the proof of Theorem 3.

It should be noted that algorithm 1 generalizes the well-known augmented Lagrangian technique (see [5], [13] e.g.) to a wider class of auxiliary functionals.

3. A superlinearly convergent method

In [13] it is shown that algorithm 1 with $\varphi(u) = \langle u, u\rangle$ in Hilbert spaces W, F generates sequences $\{v^k\}$ only linearly converging to 0. More general classes of sequential unconstrained minimization methods for solving finite dimensional problems are investigated in [7].

Now we sketch the principles of a second order technique in shifting methods.

We preserve the assumptions from Section 2. Furthermore, we assume the functionals χ and φ to be twice continuously Fréchet-differentiable on a neighbourhood $U_\varepsilon(0)$ and on F, respectively. Additionally let some $\gamma > 0$ exist such that

$$(3.1) \qquad \big(\varphi''(u)w, w\big) \geqslant \gamma \|w\|^2 \quad \text{for any } u, w \in F.$$

ALGORITHM 2

step 1:
step 2: the same as steps 1–3 of algorithm 1
step 3:
step 4: Find $y^{k+1} \in Y$ such that

$$(3.2) \qquad \varphi'(y^{k+1}) = \varphi'(y^k + v^k) + \chi''(v^k)v^k.$$

Set $k := k+1$. Go to step 2.

We remark that condition (3.1) is sufficient for (2.2) with $\delta(t) = \frac{1}{2}\gamma t^2$. Furthermore, from $y^{k+1} = y^k + v^k$ in step 4 of algorithm 1 we get $\varphi'(y^{k+1}) = \varphi'(y^k + v^k)$. Therefore (3.2) can be considered as a natural extension of step 4 in algorithm 1, taking into account second order information on χ.

THEOREM 4. *Let the auxiliary problems be solvable for any parameter $y \in Y$. Then there exists $\varrho > 0$ such that algorithm 2 can be performed for any $y^1 \in Y$ satisfying*

$$(3.3) \qquad \|\chi'(0) + \varphi'(y^1)\|_{F^*} < \varrho.$$

Furthermore, the sequence $\{v^k\}$ converges to 0 superlinearly.

Proof. Write $N(y) = \{u \in F \mid \chi(u) + \varphi(y+u) \leqslant \chi(0) + \varphi(y)\}$. By the convexity of χ (ensured by the convexity of the problem (1.1)) we get by Taylor's formula

$$\chi(u) + \varphi(y+u) \geqslant \chi(0) + \varphi(y) + \big(\chi'(0) + \varphi'(y)\big)u + \tfrac{1}{2}\int_0^1 \big(\varphi''(y+tu)u, u\big)dt.$$

Using (3.1) we hence obtain

$$(3.4) \qquad \|\chi'(0) + \varphi'(y)\|_{F^*} := \sup_{\|w\|=1} \frac{\|\big(\chi'(0) + \varphi'(y)\big)w\|}{\|w\|} \geqslant \frac{1}{2}\gamma\|u\|$$

for any $u \in N(y)$.

Thus

$$(3.5) \qquad \|\chi'(0) + \varphi'(y)\|_{F^*} < \gamma\varepsilon/2 \Rightarrow N(y) \subset \operatorname{int} U_\varepsilon(0).$$

The definition of the set $N(y)$ implies

(3.6) $$\operatorname*{argmin}_{u\in F}\{\chi(u)+\varphi(y+u)\}\subset N(y).$$

From the differentiability of χ on $U_\varepsilon(0)$ we get

$$\chi'(0)+\varphi'(y+u)+\chi''(u)u = \chi'(u)+\varphi'(y+u)+\int_0^1\big(\chi''(u)-\chi''(tu)\big)u\,dt$$
for any $u\in U_\varepsilon(0)$.

This leads to

(3.7) $$\|\chi'(0)+\varphi'(y+u)+\chi''(u)u\|_{F^*}\leqslant\|\chi'(u)+\varphi'(y+u)\|_{F^*}+o(\|u\|)$$
for any $u\in U_\varepsilon(0)$.

Because of $\lim\limits_{a\to+0}\dfrac{o(a)}{a}=0$ and $\gamma>0$, there exists $\hat{a}>0$ such that

(3.8) $$o(a)\leqslant\tfrac{1}{4}\gamma a\quad\text{for any } a\in[0,\hat{a}].$$

Let $\varrho=\frac{1}{2}\gamma\min\{\hat{a},\varepsilon\}$. Then, by induction

(3.9) $$\|v^k\|<\min\{\hat{a},\varepsilon\},\qquad k=1,2,\ldots$$

Since E is of type $E(y,v)=v_0+\varphi(y+v)$, from the relation $\mathrm{cl}Q=\mathrm{cl\,epi}\,\chi$ and Lemma 1, we get

(3.10) $$v^k\in\operatorname*{argmin}_{u\in F}\{\chi(u)+\varphi(y^k+u)\},\qquad k=1,2,\ldots$$

for the points v^k generated in step 3 of the algorithm. Now, assumption (3.3) yields

$$\|\chi'(0)+\varphi'(y^1)\|_{F^*}<\varrho\leqslant\gamma\varepsilon/2.$$

With (3.5), (3.10) and (3.4) this leads to

$$\|v^1\|<\frac{2}{\gamma}\varrho=\min\{\hat{a},\varepsilon\}.$$

Thus inequality (3.9) holds for $k=1$.

Now suppose that (3.9) holds for some k. From the differentiability of χ and φ on $U_\varepsilon(0)$ and on F, respectively, and from the relation (3.10) we get

$$\chi'(v^k)+\varphi'(y^k+v^k)=0.$$

Using (3.2), (3.6) and (3.8) we hence obtain

$$\|\varphi'(0)+\varphi'(y^{k+1})\|_{F^*}\leqslant\tfrac{1}{4}\gamma\|v^k\|<\tfrac{1}{4}\gamma\varepsilon,$$

and by (3.4), (3.6), (3.7) we get $\|v^{k+1}\|\leqslant\frac{1}{2}\|v^k\|$. Therefore the inequality (3.9) holds also for $k+1$. On the way we have obtained $\lim\limits_{k\to\infty}v^k=0$.

Because the auxiliary problems are supposed to be solvable and each of the problems (2.5) in step 3 of the algorithm has a unique solution, the method can be performed without further restrictions.

From (3.2), (3.4), (3.6) and (3.7) we get $\|v^{k+1}\| = o(\|v^k\|)$, $k = 1, 2, \ldots$ Therefore the sequence v^k converges superlinearly to 0.

The optimality for (1.1) of any accumulation point of the generated sequence $\{x^k\}$ can be shown similarly to the proof of Theorem 3. ∎

We remark that algorithm 2 converges only locally. A result concerning global convergence can be obtained by combining it with algorithm 1 (see [7]).

Another way for deriving superlinearly convergent methods is to construct Newton-like methods for solving the generated dual problem (1.19). The resulting procedure is closely related to the method proposed here.

In computer codes the second order derivatives of the primal functional χ are to be replaced by appropriate approximations.

4. Estimations for the dual value

In this section we will briefly consider a simple but powerful application of the theoretical approach introduced in Section 1.

In the same way as in [7] we can show the following

LEMMA 3. *Let (x^*, a^*) be a saddle point of the Lagrangian L, $L(x, a) = f_0(x) + [a, f(x)]$, related to the problem* (1.1). *Further, let*

$$Q^* = \{v \in H \mid v_0 + [a^*, v] \geqslant \chi(0)\}. \tag{4.1}$$

Then the inclusion $Q \subset Q^*$ holds.

If we define a functional τ^* similar to τ by

$$\tau^*(y) = \sup\{t \in \mathbf{R}^1 \mid E(y, te^0) \leqslant \inf_{v \in Q^*} E(y, v)\}$$

then Lemma 3 leads to the estimation

$$\tau^*(y) \leqslant \tau(y) \quad \text{for any } y \in Y \tag{4.2}$$

for the dual functional τ.

In some specific methods the problems

$$E(y, v) \to \min! \quad \text{s.t.} \quad v \in Q^* \tag{4.3}$$

can be solved explicitly. Then by (4.2) we get a practicable estimation of $\tau(y)$. However, this bound is qualitative only because the component a^* of a saddle point (x^*, a^*) of the Lagrangian is generally unknown, as is the solution itself.

Let us consider a simple example showing the advantage of the proposed estimation. Let F be a Hilbert space and let

$$E(y, v) = v_0 + y\langle v, v\rangle, \qquad Y = \operatorname{int} \boldsymbol{R}^1_+.$$

Then the problem (4.3) has the form

$$v_0 + y\langle v, v\rangle \to \min!$$

subject to

$$v_0 + \langle a, v\rangle \geqslant \chi(0), \qquad v \in H.$$

Minimizing over v_0 we get the equivalent problem

$$\chi(0) + y\langle v, v\rangle - \langle a, v\rangle \to \min! \quad \text{s.t.} \quad v \in F.$$

This problem is solved by $v^* = \frac{1}{2y} a^*$. Thus we get

$$\inf_{v \in Q^*} E(y, v) = \chi(0) - \frac{1}{4y} \|a^*\|^2,$$

and because of $e(y, 0) = 0$, we have

$$\tau^*(y) = \chi(0) - \frac{1}{4y} \|a^*\|^2.$$

For further details the interested reader is referred to [7].

References

[1] А. С. Антипин, *Методы нелинейного программирования основанные на прямой и двойственной модификации функции Лагранжа*, Препринт ВНИИ Сист. Иссл., Москва 1979.

[2] I. Babuška, *The finite element method with penalty*, Math. Computation 27 (1973), 221–228.

[3] I. Ekeland, R. Temam, *Convex analysis and variational problems*, North-Holland, Amsterdam 1976.

[4] H. D. Everett III, *Generalized Lagrange multiplier method for solving problem of optimum allocation of resources*, Operations Res. 11 (1963), 399–417.

[5] Е. Г. Гольштейн, Н. В. Третьяков, *Модифицированные функции Лагранжа*, Экон. Мат. Методы 10.3 (1974), 568–591.

[6] Ch. Grossmann, *Common properties of nonlinear programming algorithms basing on sequential unconstrained minimizations*; in: *Proceedings of the 7-th summer school "Nonlinear Analysis", Berlin 1979*, Akademie-Verlag, pp. 107–117.

[7] —, *Sensitivitätsanalyse als einheitlicher Zugang zu den Verfahren der sukzessiven unrestringierten Minimierung*, Dissertation B, TU Dresden, 1979.

[8] —, *A unified approach to nonlinear programming algorithms basing on sequential unconstrained minimizations*; in: *Proceedings of the 9-th IFIP conference on optimization techniques*, Lect. Notes, Springer, Berlin 1980, pp. 62–68.

[9] Ch. Grossmann, A. A. Kaplan, *Strafmethoden und modifizierte Lagrangefunktionen in der nichtlinearen Optimierung*, Teubner-Text, Leipzig 1979.

[10] I. Hlavacek, *Dual finite element analysis for some unilateral boundary value problems*; in: *Equadiff IV*, Lect. Notes, Springer, Berlin 1979, pp. 152–159.

[11] А. Д. Иоффе, В. М. Тихомиров, *Теория экстремальных задач*, „Наука", Москва 1974.

[12] O. K. Neckermann, *Verallgemeinerte kunjugierte Dualität und Penaltyverfahren*, Dissertation, Würzburg 1976.

[13] R. T. Rockafellar, *Augmented Lagrange multiplier functions and duality in nonconvex programming*, SIAM J. Control 12 (1974), 268–285.

[14] E. Zeidler, *Vorlesungen über nichtlineare Funktionalanalysis*, Teubner-Text, Leipzig 1977.

Presented to the Semester
Computational Mathematics
February 20 — May 30, 1980