Ivan Hlaváček

Some variational methods for nonlinear mechanics

In: Svatopluk Fučík and Alois Kufner (eds.): Nonlinear Analysis, Function Spaces and Applications, Proceedings of a Spring School held in Horní Bradlo, 1978, [Vol 1]. BSB B. G. Teubner Verlagsgesellschaft, Leipzig, 1979. Teubner Texte zur Mathematik. pp. 128--148.

Persistent URL: http://dml.cz/dmlcz/702408

SOME VARIATIONAL METHODS FOR NONLINEAR MECHANICS

Ivan Hlaváček

1. Introduction

If one has to solve a physical problem as a whole, i.e., from the physical reality to the numbers, one meets usually three major stages:

(i) the mathematical formulation of the problem (creating a mathematical model);

(ii) the approximation of the mathematical model, i.e., a transformation into one or a sequence of problems of a simpler nature, which are solvable in finite-dimensional spaces;

(iii) the algorithm, realizing the numerical solutions of the approximate problems.

Each of the stages gives rise to important theoretical questions, for instance:(i) which is the most suitable mathematical formulation (this point includes proofs of existence and uniqueness of the solution), (ii) the choice of approximations, error estimates or at least a convergence proof for the approximations, (iii) the choice of the algorithm, its convergence and other properties.

In the present lecture we restrict ourselves to nonlinear problems, which admit a natural variational formulation:to find a function minimizing a convex functional over a closed convex set of admissible functions. We shall discuss the stages (ii) and (iii) of the scheme mentioned above. Our exposition is by no means exhaustive - there exists extremely rich literature in this field (see the references [1] - [10]). The aim of this lecture is to give a survey of some methods, which appear to be efficient in practice.

First we present several methods of linearization: method of Kachanov, steepest descent, contraction. (Section 2.)

Then a basic idea of discretization by finite elements will be shown, which transforms the initial problem formulated in a functional space with infinite by many dimensions into approximate problems in finite-dimensional spaces. The latter problems, however, may still be nonlinear. We present a theorem on the convergence of the finite element approximations. The theorem follows from a more general one for the Ritz-Galerkin method. The question of error estimates will also be discussed.

If the approximate problems are nonlinear, their solution may not be immediate. To this end, we present several efficient algorithms of convex programming.

In the conclusion of this section, let us sketch a general scheme to clasify the position of individual steps in the course of the total solution - see Fig. 1.

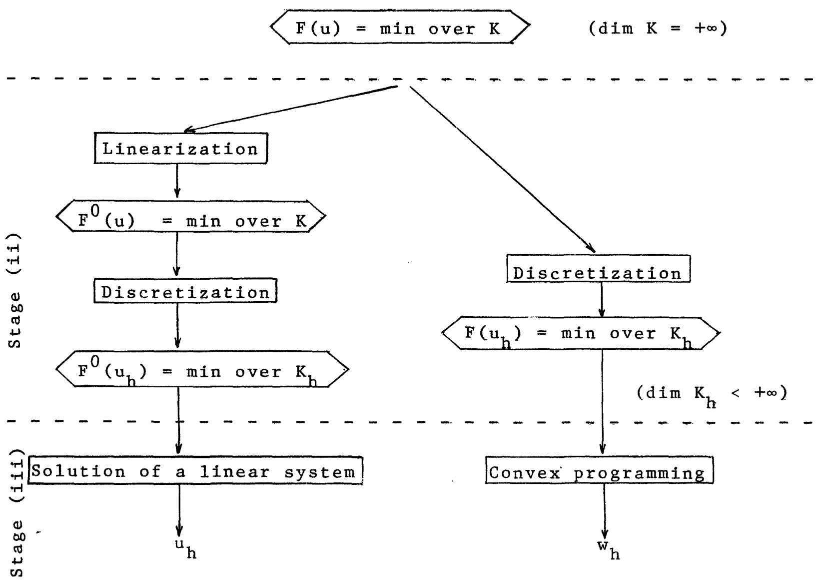

Fig. 1.

2. Methods of linearization

For linear problems many theoretical and practical results have been obtained. It can be therefore advantageous to solve the original problem by generating a sequence of suitable approximate linear problems (i.e. minimizations of quadratic functionals). There are several ways of linearization:

(i) those based on the physical intuition, e.g.

- superposition of small deformations on finite deformations in solid mechanics,
- incremental methods in plasticity,
- Kachanov's method in elasto-plasticity;

(ii) those of abstract mathematical character, e.g.

- method of the steepest descent,
- method of contraction.

Let us present some theoretical results on the Kachanov's method, the steepest descent and the method of contraction.

2.1. Method of Kachanov (secant modulus method)

We shall give an abstract version of the method in a Hilbert space H (see [1], [8]).

Theorem 2.1. Let a functional $\mathcal{F}: H \to R^1$ *be given, which has a continuous and strongly monotone Gâteaux differential, i.e.*

$$D\mathcal{F}(u+v,v) - D\mathcal{F}(u,v) \geqq C_1\|v\|^2 . \tag{2.1}$$

Let a form $B(u;x,y)$ *be given, symmetric and bilinear with respect to* x *and* y *, depending on another element* $u \in H$ *and such that*

$$B(u;y,y) \geqq C_2\|y\|^2 , \tag{2.2}$$

$$B(u;x,y) \leqq C_3\|x\| \, \|y\| , \tag{2.3}$$

$$B(u;u,v) = D\mathcal{F}(u,v), \tag{2.4}$$

(2.5) $\frac{1}{2}B(x;y,y) - \frac{1}{2}B(x;x,x) + \mathcal{F}(x) - \mathcal{F}(y) \geqq 0$.

Let $f \in H$ *and let* u *be the (unique) element minimizing the functional* $F(v) \equiv \{\mathcal{F}(v) - (f, v)\}$ *over* H .

Let $u_0 \in H$ *be arbitrary. We define the iterative solutions* $u_1, u_2, \ldots$ *by the following linear problems:*

(2.6) $B(u_n;\ u_{n+1},\ v) = (f,\ v) \quad \forall v \in H, \quad n = 0, 1, 2, \ldots$.

Then there exists a unique sequence $\{u_n\}$ *and* u_n *converge to* u *in* H .

EXAMPLE 2.1 In the theory of elasto-plastic bodies the method can be easily applied. The problem (2.6) is obtained by inserting u_n into the function of Lamé modulus, which is a natural idea.

2.2. Method of the steepest descent

The following iterative procedure for minimizing a functional is based on a simple idea: to change the approximation in the direction opposite to the gradient of the functional, i.e. in the direction of the steepest descent. Let us present some sufficient conditions for the convergence of this method.

Theorem 2.2. Let the functional $F : H \to R^1$ *have the first and second Gâteaux differentials. Denoting*

$$DF(u,v) = (G(u),v) ,$$

the element $G(u)$ *is called the gradient of* F *at the point* u . *Let the second differential* $D^2F(u;v,w)$ *be continuous with respect to* u , *if* v *and* w *are fixed.*

Assume that there exist positive numbers M *and* m *such that*

(2.7) $$|D^2F(u;v,w)| \leqq M||v||\ ||w|| ,$$

(2.8) $$D^2F(u;v,v) \geqq m||v||^2 .$$

Let u_n *be known. We set* $w_n = G(u_n)$. *If* $w_n = 0$, u_n *is equal to the (unique) element* u *minimizing* F *over* H . *If* $w_n \neq 0$, *we set*

$$u_{n+1} = u_n - \frac{1}{M} w_n \tag{2.9}$$

or

$$u_{n+1} = u_n - \rho_n w_n \ , \tag{2.10}$$

where $\rho_n \in [0,\infty)$ *is such that* u_{n+1} *minimizes* F *on the ray* $u_n - \rho w_n$.

Then u_n *converge to* u *in* H . For the proof we refer e. g. to [4] or [7].

REMARK 2.2 (i) The search of $G(u_n)$ coincides with a linearization of the original problem.

(ii) There are many related procedures - see e.g. the books [3] or [4].

2.3. Method of contraction

The well-known principle of contractive mappings can be employed to the solution of the minimum problem.

Theorem 2.3 *Let operator* $T : H \rightarrow H$ *be strongly monotone, i.e.*

$$(Tu - Tv,\ u - v) \geqq m||u - v||^2$$

and satisfying the Lipschitz condition:

$$||Tu - Tv|| \leqq M||u - v||.$$

Let $y \in H$ *be given. Then the (unique) solution of the equation*

$$Tx = y \tag{2.11}$$

can be found as a fixed point of the operator A , *where*

$$Ax = x - \varepsilon(Tx - y), \quad 0 < \varepsilon < 2m/M^2 .$$

Defining the iterations

$$x_{n+1} = Ax_n \ , \qquad n = 0,1,... \tag{2.12}$$

with $x_0 \in H$ *arbitrary, we have:*

x_n *converge to* x *in* H ,

$$||x_n - x|| \leqq \frac{\alpha^n}{1 - \alpha} ||x_0 - Ax_0|| \ ,$$

where

$$\alpha = (1 - 2\varepsilon m + \varepsilon^2 M^2)^{\frac{1}{2}} < 1 .$$

For the proof see e. g. [7] or [3].

REMARK 2.3 The above method can be applied to the solution of the minimum problem for the functional F over H, if

$$Tx - y = G(x) ,$$

where $G(x)$ denotes the gradient of F at x.
In practice, the problem to evaluate x_{n+1} from (2.12) represents a linear problem, due to the definition of the operator T (see [7]).

3. Discretization by finite elements

Even after the linearization we still have a problem in an infinite-dimensional functional space. Thus we are not able to start a numerical procedure unless the problem is reduced to finite-dimensional spaces by a discretization. To this end, two general ways can be employed - finite differences or finite elements. Having problems on domains with general boundaries in mind, we prefer the finite element method.

3.1. Finite element method

Let us recall briefly the main features of the finite element discretization.

(i) The domain $\Omega \subset R^n$ ($n = 1,2,3$ in most cases) is decomposed into a finite number of subdomains Ω_j (e.g. simplexes if Ω is a polytope) such that

$$\bar{\Omega} = \bigcup_{j=1}^{N} \bar{\Omega}_j , \quad \Omega_j \cap \Omega_k = \emptyset \quad \text{for } j \neq k .$$

(ii) In every subdomain Ω_j we choose
- a polynomial function $p_j(x)$, $x \in \bar{\Omega}_j$,
- a set of nodes $Q_{jk} \in \Omega_j$, $k = 1,2,\ldots,\bar{k}$ (e.g. vertices, midpoints of the sides e.t.c.)

- a set of nodal parameters (e.g. the values of derivatives $D^{|\alpha|}p_j(Q_{jk})$, $|\alpha| = 0,1,\ldots$) such that p_j is uniquely determined by the set of nodal parameters.

The subdomain $\bar{\Omega}_j$ together with the polynomial p_j, the set of nodes and nodal parameters is called a finite element.

(iii) We define a piecewise polynomial function $p(x)$ on the whole domain Ω such that the restrictions are

$$p|_{\Omega_j} = p_j \ , \quad j = 1,2,\ldots,N \ ;$$

by equating nodal parameters at the coinciding nodes common for any two adjacent subdomains, we can guarantee the continuity of p and of some derivatives $D^\alpha p$ if necessary.

EXAMPLE A classical finite element technique consists of a triangulation of a polygonal domain $\Omega \subset R^2$ and of using linear polynomials p_j ; the nodes are identified with the vertices and the nodal parameters with the values of $p_j(Q_{jk})$ $(|\alpha| = 0)$.

Any function p constructed above is determined uniquely by a finite number of nodal parameters. Setting all the nodal parameters equal to zero except one, which is equal to 1 , we obtain a basis function w_m . Then any p can be written in the form

$$p(x) = \sum_{m=1}^{\bar{m}} a_m w_m(x) \ , \tag{3.1}$$

where the coefficients a_m are uniquely determined. Thus we are led to spaces S_h of piecewise polynomial functions of the norm (3.1). For instance, the space of linear finite elements (see the above EXAMPLE) is a subspace of $W^{1,2}(\Omega) \cap C(\bar{\Omega})$.

The two following properties of finite element spaces are of great importance. (For simplicity, we present the first of them for the plane domains $\Omega \subset R^2$ only.)

(i) Approximability: Denote $h = \max_{j=1,\ldots,N} (\operatorname{diam} \bar{\Omega}_j)$. Let the family $\{\mathcal{T}_h\}$, $0 < h \leq h_0$ of triangulations of Ω be regular, i.e. let a

$\gamma > 0$ exist, independent of h and such that all interior angles in $\mathcal{T}_h$ are bounded from below by γ. Then to every function $u \in W^{k,2}(\Omega)$ there exists a function $v_h \in S_h$ such that

$$(3.2) \qquad \|u - v_h\|_{W^{m,2}(\Omega)} \leqq C\, h^{k-m} |u|_{W^{k,2}(\Omega)} \ ;$$

here $k_0 \leqq k \leqq k_1$, $0 \leqq m < k$ and the integers k_0, k_1 are determined by the particular type of finite elements; $|u|_{W^{k,2}}$ denotes the seminorm in $W^{k,2}(\Omega)$, generated by the derivatives of the k-th order.

(E.g. for linear finite elements one has $k_0 = 1$, $k_1 = 2$.) See [5], [6] for the details.

(ii) Small support of the basis functions. In contradiction to the classical Ritz-Galerkin method, the support of any basis function w_m in (3.1) is much smaller than Ω . Therefore, if we employ the Ritz-Galerkin method (see (3.3). (3.4), (3.5) below), in case of a quadratic functional F we obtain a problem with a band matrix, which is of great advantage from the computational point of view.

Let us consider again the problem

$$(3.3) \qquad F(u) = \min \quad \text{over} \quad K \subset H \ ,$$

where F is convex and K is a convex closed subset. Let us apply the general Ritz-Galerkin technique by reducing the set K of admis sible functions to the approximate set

$$K_h = K \cap S_h \ ,$$

where S_h is the space of finite elements.

(In case that the original space H is generated by a Cartesian product

$$H = W^{k_1,2}(\Omega) \times W^{k_2,2}(\Omega) \times \ldots \times W^{k_r,2}(\Omega) \ , \qquad r > 1 \ ,$$

we have to approximate each of the components by a suitable finite element space.)

Thus we obtain an approximate problem

$$(3.4) \qquad F(u_h) = \min \quad \text{over} \quad K_h \ .$$

Substituting for u_h the sum $\sum_m a_m^h w_m$, we are led to an equivalent problem

$$\mathcal{F}(a^h) = \min \quad \text{over } \mathcal{K}_h , \tag{3.5}$$

where $\mathcal{K}_h$ is a closed convex subset of $R^{\bar{m}}$ and $\mathcal{F}$ is a convex function. **The latter** problem can be solved by various algorithms. A few procedures of convex programming will be recommended in Section 4.

3.2. Convergence of the finite element method

Next let us discuss the properties of the solution u_h of the problem (3.4) especially the behaviour of the error $||u - u_h||$ for $h \to 0$.

<u>*Theorem 3.1*</u> *Let* $F : H \to R$ *be a functional with the first and second differentials and assume that*

$$C_1 ||v||^2 \leqq D^2F(u;v,v) \leqq C_2 ||v||^2 \qquad \forall u,v \in H . \tag{3.6}$$

Let K *be a closed convex subset of* H, $K_h \subset H$ *closed convex subsets for any* $0 < h \leqq h_0$. *Let* u *and* u_h *be the (unique) solutions of the minimum problems* (3.3) *and* (3.4), *respectively.*

Assume that

(i) to every $h \in (0,h_0]$ *there exists an element* $v_h \in K_h$ *such that*

$$||u - v_h|| \to 0 \quad \text{for} \quad h \to 0;$$

(ii) $\left.\begin{array}{l} v_h \in K_h,\ u^* \in H , \\ v_h \rightharpoonup u^* \ \ (\textit{weakly})\ \textit{for} \ \ h \to 0 \end{array}\right\} \Rightarrow u^* \in K .$

Then

$$||u - u_h|| \to 0 \quad \text{for} \quad h \to 0. \tag{3.7}$$

For the proof see [3] - Chpt. 4, Th. 06.

<u>REMARK 3.1</u> 1^o Note that if $K_h \subset K$ (i.e., the so called "internal" approximations of K), the assumption (ii) is satisfied due to the fact that $v_h \in K$ for any h and K is weakly closed.

2^{o} The assumption (i) is more difficult to verify even for $K_h \subset K$. Suppose we are able to prove that the intersection

$$K \cap C^{\infty}(\bar{\Omega})$$

is dense in K. (If e.g. $K = W^{k,2}(\Omega)$, this density is a well-known result of Gagliardo.) Then applying estimates of the form (3.2) to a function $u_{\varepsilon} \in K \cap C^{\infty}(\bar{\Omega})$, which is close enough to u, we deduce the convergence $v_h \to u$.

REMARK 3.2 Sometimes it is not suitable or even possible to construct $K_h \subset K$. Then we have to prove (ii) explicitly. Such a case is called the case of "external" approximations.

Sometimes the left inequality in (3.6) is not true and we have only

$$C_1 |v|^2 \leqq D^2F(u;v,v) \leqq C_2 ||v||^2 \quad \forall u \in K, \quad \forall v \in H , \tag{3.8}$$

where $|\cdot|$ denotes a seminorm in H. Then the following result can be useful.

Theorem 3.2 *Let functional* F *be coercive on* K, *i.e.*

$$\lim F(v) = +\infty \text{ for } \quad v \in K , \quad ||v|| \to +\infty$$

and let it satisfy (3.8).

Assume that both the minimum problems (3.3) over K *and (3.4) over* $K_h \subset K$ *have unique solutions* u *and* u_h, *respectively.*

Let to every $h \in (0,h_0]$ *a* $v_h \in K_h$ *exist such that*

$$||u - v_h|| \to 0 \quad \text{for} \quad h \to 0 .$$

Then for $h \to 0$ *it holds*

$$u_h \rightharpoonup u \qquad \text{(weakly) in } H , \tag{3.9}$$

$$|u_h - u| \to 0 . \tag{3.10}$$

The proof is parallel to that of Theorem 3.1.

REMARK 3.3 We often can prove on the basis of (3.9) and (3.10) that $u_h \to u$ in H (strongly). In fact, consider e.g. the case $H = W^{1,2}(\Omega)$,

$$|v|^2 = \int_\Omega |\text{grad } v|^2 \, dx \; .$$

Then (3.9) implies that $u_h \to u$ in $L_2(\Omega)$. Using also (3.10) we can readily see that $u_h \to u$ in $W^{1,2}(\Omega)$.

3.3. Error estimates

The problem to find the rate of convergen e in terms of the parameter h appears to be delicate even in relatively simple cases. It is well-known that if a functional F satisfying (3.6) has to be minimized over a whole Hilbert space H , the error $||u - u_h||$ of the Ritz-Galerkin approximations $u_h \in H_h = H \cap S_h$ is of the same order as the distance of u from the subspace H_h . Therefore one can use any element $v_h \in H_h$ in $C||u - v_h||$ to obtain an upper bound for $||u - u_h||$, and an interpolate of u in H_h is inserted for v_h , as a rule.

For problems, where the set K of admissible functions is not the whole space H (i.e. for variational inequalities), we have only

$$(3.11) \qquad ||u - u_h|| \leq C[\text{dist}(u,K_h)]^{1/2} \; ,$$

in general. In particular cases, however, the estimate (3.11) is not optimal. To the author's knowledge, only two methods for quasi-optimal a priori error estimates have been proposed:

(i) method of one-sided approximations (MOSCO and STRANG [11]),

(ii) method of FALK ([12]) .

Method of one-sided approximations is based on the following lemma.

Lemma 3.1 Let F *be a functional satisfying the conditions (3.6).*

Let K *be a convex closed subset of* H *and* $K_h \subset K$ *a closed convex subset. Let* u *and* u_h *be the (unique) solutions of the problems* (3.3) *and* (3.4), *respectively.*

Let an element $w_h \in K_h$ *exist such that* $2u - w_h \in K$. *Then*

(3.12) $$||u - u_h|| \leq (C_2/C_1)^{\frac{1}{2}}||u - w_h|| \quad .$$

For the proof see e.g. [11] or [13].

Hence the error is bounded from above by the distance from u to w_h. If we succeed in finding a suitable element w_h, for which the estimate is possible, we obtain the same rate of convergence for $u_h - u$. Thus the whole problem is reduced to the construction of $w_h \in K_h$, $2u - w_h \in K$, w_h close to u. (See [11], [13], [14] for such an approach.)

REMARK 3.4 Let a functional F satisfy only (3.8) instead of (3.6). Assuming that $K_h \subset K$ and both (3.3) and (3.4) have unique solution u and u_h, respectively, we obtain

(3.13) $$|u - u_h| \leq (C_2/C_1)^{\frac{1}{2}}||u - w_h||$$

for any $w_h \in K_h$ such that $2u - w_h \in K$.

(iii) Method of Falk consists in the following lemma.

Lemma 3.2 Let

(3.14) $$F(v) = \tfrac{1}{2} A(v,v) - f(v) \ ,$$

where $A(u,v)$ *is a symmetric, positive bilinear form continuous on* $H \times H$, $f \in H'$ *a given linear continuous functional.*

Let K *and* K_h *be closed convex subsets of* H. *Assume that the problems of minimizing* F *over* K *and over* K_h *have solutions* u *and* u_h, *respectively.*

Then it holds

(3.15) $$A(u-u_h,u-u_h) \leq [f(u-v_h) + f(u_h-v) + A(u_h-u,v_h-u) + + A(u,v_h-u) + A(u,v-u_h)] \quad \forall v \in K, \ \forall v_h \in K_h \ .$$

The proof follows from the conditions

$$A(u,v-u) - f(v-u) \geq 0 \qquad v \in K \ ,$$

$$A(u_h,v_h-u) - f(v_h-u) \geq 0 \qquad v_h \in K_h \ .$$

The estimate (3.15) can be utilized provided the solution u is sufficiently regular. In some examples, the two terms

$$[A(u,v_h-u) + f(u-v_h)] \qquad \text{and} \qquad [A(u,v-u_h) + f(u_h-v)]$$

can be transformed by means of Green's theorem into surface integrals and the latter estimated by a suitable choice of $v_h \in K_h$ and $v \in K$ (see e.g. [15], [16], [17] for the detailed proofs in case of unilateral boundary value problems).

4. Some algorithms of convex programming

We are going to discuss several algorithms, which are suitable for the solution of the approximate problem (3.5), i.e. for

$$\mathcal{F}(a^h) = \min \qquad \text{over} \qquad \mathcal{K}_h \subset R^{\bar{m}} .$$

Since both the function $\mathcal{F}$ and the set $\mathcal{K}_h$ are convex, the problem belongs to convex programming.

In general, we distinguish two classes in convex programming:

(i) problems without constraints $(\mathcal{K}_h = R^{\bar{m}})$,

(ii) problems with constraints $(\mathcal{K}_h \subset R^{\bar{m}}, \mathcal{K}_h \neq R^{\bar{m}})$.

4.1. Problems without constraints

4.1.1 Gradient methods. The well-known approach of the steepest descent (gradient method), presented already as a method of linearization, can be employed also in a finite-dimensional space. The convergence is guaranteed by the sufficient conditions of Theorem 2.2. Moreover, we have the error estimate

$$||u_n-a^h|| \leqq q^n||u_0-a^h|| , \tag{4.1}$$

where $0 < q < 1$, provided $u_{n+1} = u_n - \alpha_n w_n$ and α_n is choosen properly - see [4] (i.e. $\alpha_n = \alpha$, $0 < \alpha < 2/M$ with the optimal value $\alpha = 2/(M+m)$ or α_n to be adjusted at any step).

Although the gradient methods are relatively simple, they converge too slowly in practice. In fact, the minimal value of q in

(4.1) being

$$q_{min} = (M-m)/(M+m) ,$$

it is very close to 1 if the matrix of the second derivatives $D^2\mathcal{F}$ (the Hessian) is ill-conditioned (i.e. if $m/M \ll 1$).

EXAMPLE 4.1 Consider $\mathcal{F}(x,y) = \frac{1}{2}(x^2/a^2 + y^2/b^2)$. The eigenvalues of $D^2\mathcal{F}$ are a^{-2}, b^{-2} and $b^2/a^2 = m/M$. If $b^2/a^2 \ll 1$, the isohypses-ellipses are very long and the method of the steepest descent requires a great number of steps to reach a sufficient accuracy. (A phenomenon of "zig-zagging".)

Consequently, we use the gradient method only as the first step of other more efficient methods.

4.1.2 Conjugate gradient methods

Here we describe only the main idea. Consider a quadratic functional

$$\mathcal{F}(v) = \tfrac{1}{2}A(v,v) - (f,v) ,$$

with A symmetric and positive definite. Assume that u_n has been found. Then we set

$$u_{n+1} = u_n + \lambda w_n , \qquad n = 0,1,\dots,\bar{m}-1 ,$$

where $\lambda \in R^1$ minimizes the function $f(\lambda) = \mathcal{F}(u_n + \lambda w_n)$ and w_n is one of the conjugate directions $\{w_0, w_1, \dots, w_{\bar{m}-1}\}$, which satisfy the conditions

$$(Aw_i, w_j) = 0 \quad \text{for} \quad i \neq j$$

and $(Aw_i, w_i) \neq 0$. We can take

$$w_0 = -G(u_0) = -Au_0 + f .$$

It is not difficult to prove that $u_{\bar{m}} = a^h$, i.e., the solution of the problem (3.5) can be obtained by a finite number of steps.

EXAMPLE 4.2 Consider again the quadratic function from EXAMPLE 4.1. It is obvious that already u_2 is the exact solution.

For a convex non-quadratic function, the algorithm can be adapted - see e.g. [2], [4]. The modified algorithm is based on the approximation of $\mathcal{F}$ by a quadratic function

$$\phi(v) = \mathcal{F}(a^h) + \tfrac{1}{2} D^2 \mathcal{F}(a^h; v-a^h, v-a^h)$$

in a neighbourhood of a^h.

4.1.3 Methods of relaxation

One of the simplest algorithms consists in fixing successively the $\bar{m} - 1$ variables of $\mathcal{F}(v_1,\dots,v_{\bar{m}})$ and relaxing only one of them to get the minimum. Thus we start with a fixed v^0 and calculate v^{n+1} step by step from the known vector v^n as follows:

v_i^{n+1} minimizes the function

$$(4.2) \quad f(t) = \mathcal{F}(v_1^{n+1},\dots,v_{i-1}^{n+1},t,v_{i+1}^n,\dots,v_{\bar{m}}^n) \quad \text{over} \quad R^1 , \quad i = 1,\dots,\bar{m} .$$

For quadratic $\mathcal{F}$ the procedure coincides with the Gauss-Seidel algorithm.

<u>*Theorem 4.1*</u> *Let*

$$\mathcal{F}(v) = \mathcal{F}_0(v) + \sum_{i=1}^{\bar{m}} \alpha_i |v_i| , \quad \alpha_i \geqq 0 ,$$

where $\mathcal{F}_0$ *is coercive, strictly convex and* C^1*-function. Then the relaxation procedure converges to the solution* a^h *of* (3.5).

For the proof see [2] .

<u>REMARK 4.1</u> Let us modify the relaxation algorithm as follows: denote the parameter minimizing $f(t)$ in (4.2) by $v_i^{n+1/2}$ and set

$$(4.3) \quad v_i^{n+1} = (1 - \omega)v_i^n + \omega v_i^{n+1/2}, \quad i=1,\dots,\bar{m} ,$$

where $0 < \omega < 2$. The algorithm with $\omega > 1$ is called the successive over-relaxation (SOR) and that with $\omega < 1$ the successive under-relaxation. The aim of introducing the parameter ω is to accelerate the convergence.

4.2 Problem with constraints

4.2.1 Relaxation with projection

Let us suppose that a convex closed set $\mathcal{K}_h \subset R^{\bar{m}}$ can be written as

$$\mathcal{K}_h = \prod_{u=1}^{\bar{m}} K_i ,$$

where $K_i = [a_i, b_i]$, $-\infty \leqq a_i \leqq b_i \leqq +\infty$, i.e. the constraints are local, prescribed separately for each coordinate. Let $\mathcal{F}$ be a quadratic functional

$$\mathcal{F}(v) = \tfrac{1}{2}A(v,v) - (f,v) .$$

Denote the parameter minimizing the function $f(t)$ in (4.2) by $v_i^{n+1/2}$ and set

$$v_i^{n+1} = P_{K_i}[(1-\omega)v_i^n + \omega v_i^{n+1/2}] , \tag{4.4}$$

where P_{K_i} is the projection onto K_i.

Theorem 4.2 Let $A(u,v)$ *be a symmetric and positive definite bilinear form,* $0 < \omega < 2$. *Then the procedure* (4.4) *converges to the solution* a^h *of* (3.5).

For the proof see [2].

REMARK 4.2 In particular, Theorem 4.2 applies to the problem without constraints, where $K_i = R^1$, $i = 1,2,\ldots,\bar{m}$.

REMARK 4.3 The algorithm appears to be the most efficient for numerous variational inequalities - cf. [2].

REMARK 4.4 The relaxation algorithm and the convergence results can be extended to a more general class of problems. Let the following decomposition hold:

$$R^{\bar{m}} = \prod_{i=1}^{N} V_i , \quad \dim V_i = m_i , \quad \sum_{i=1}^{N} m_i = \bar{m} ,$$

$$\mathcal{K}_h = \prod_{i=1}^{N} K_i , \quad K_i \subset V_i \text{ are convex and closed subsets.} \tag{4.5}$$

Then the procedure is called "relaxation by blocks" (see [2]).

REMARK 4.5 In some cases the relaxation algorithm converges even for a more general set $\mathcal{K}_h$, provided the initial element v^0 is chosen properly.

4.2.2 Methods of feasible directions

One of the first methods proposed for convex programming was the method of feasible directions (see [18],[19], [3], [4]).

Assume that

$$\mathcal{K}_h = \{v \in R^{\bar{m}} \mid f_i(v) \leq 0 , \quad i = 1,\dots,m ,$$
$$(a_j, v) = b_j, \quad j = m+1,\dots,n\}$$

where $f_i \in C^1$ are convex functions, a_j and b_j are given vectors.

Let $v_0 \in \mathcal{K}_h$ be given. We seek a ("feasible") direction $p \in R^{\bar{m}}$ such that

$$\left.\begin{array}{l} \exists \hat{\alpha} > 0, \quad v_0 + \alpha p \in \mathcal{K}_h \\ \mathcal{F}(v_0 + \alpha p) < \mathcal{F}(v_0) \end{array}\right\} \forall 0 < \alpha \leq \hat{\alpha}$$

Then we set $v_1 = v_0 + \alpha_1 p$, where α_1 minimizes the function $f(\alpha) = \mathcal{F}(v_0 + \alpha p)$ over the interval $(0,\hat{\alpha}]$; then the procedure is repeated.

In case of linear constraints

$$f_i(v) = (a_i, v) - b_i , \quad i = 1,\dots,m ,$$

the feasible direction p can be determined as follows. Let

(A) $\quad (a_i, p) \leq 0 \qquad \forall i \in \{1 \leq i \leq m \mid (a_i, v_0) = b_i\} ,$

$\quad (a_j, p) = 0 \qquad j = m+1,\dots,n ,$

(B) $\quad ||p|| \leq 1 ,$

(C) p minimizes the functional $D\mathcal{F}(v_0, p) \equiv g(p)$

over the set of all p satisfying (A) and (B).

The problem (A), (B), (C) belongs to linear programming.

Additional requirements, which guarantee the acceleration of the convergence and enable us to avoid the "zigzagging", are presented in the book [18]. (See also [19].)

4.2.3 Methods of conjugate gradients

The idea of conjugate gradients can be extended to problems with a quadratic functional $\mathcal{F}$ and linear constraints (cf. [4]).

4.2.4 Duality methods-search for a saddle point

Let us suppose that

$$\mathcal{K}_h = \{v \in R^{\bar{m}} \mid (q, \Phi(v))_n \leqq 0 \quad \forall q \in \Lambda\} \quad ,$$

where $\Lambda \subset R^n$ is a cone with the vertex 0 , and $\Phi : R^{\bar{m}} \to R^n$ is a function.

EXAMPLE. If

$$\mathcal{K}_h = \{v \in R^{\bar{m}} \mid f_i(v) \leqq 0 \ , \quad i = 1,2,\dots,n\} \quad ,$$

we can set

$$\Phi(v) = \{f_1(v), \ f_2(v), \dots, \ f_n(v)\} \quad ,$$

$$(4.6) \qquad \Lambda = \{q \in R^n \mid q_i \geqq 0 \ \forall i\} \ , \qquad (q, \Phi(v))_n = \sum_{i=1}^{n} q_i f_i(v) \ .$$

It is obvious that

$$\sup_{q \in \Lambda} \ (q, \ \Phi(v))_n = \begin{cases} +\infty & \text{for} \quad v \notin \mathcal{K}_h \ , \\ 0 & \text{for} \quad v \in \mathcal{K}_h \ . \end{cases}$$

Thus the (<u>primal</u>) problem to minimize $\mathcal{F}(v)$ over $\mathcal{K}_h$ is equivalent with:

$$(4.7) \qquad \inf_{v \in R^{\bar{m}}} \ \sup_{q \in \Lambda} \{\mathcal{F}(v) + (q, \ \Phi(v))_n\} \ .$$

Then $q \in \Lambda$ is called a <u>Lagrange multiplier</u> and

$$\mathcal{L}(v,q) = \mathcal{F}(v) + (q, \ \Phi(v))_n$$

a <u>Lagrangian</u>.

The problem

$$(4.8) \qquad \sup_{q \in \Lambda} \ \inf_{v \in R^{\bar{m}}} \mathcal{L}(v,q)$$

is called the <u>dual</u> of the problem (4.7).

If

$$\mathcal{L}(u,p) = \min_{v\in R^{\bar{m}}} \max_{q\in\Lambda} \mathcal{L}(v,q) = \max_{q\in\Lambda} \min_{v\in R^{\bar{m}}} \mathcal{L}(v,q) ,$$

the pair $\{u,p\}$ is called a <u>saddle point</u> (point of min-max).

Assume that

$$\mathcal{F}(v) = \tfrac{1}{2}A(v,v) - (f,v) ,$$

A is symmetric, positive definite on $R^{\bar{m}}$,

a saddle point exists,

$$||\Phi(u) - \Phi(v)||_n \leq C_1||u - v||_{\bar{m}} \quad \forall u,v \in R^{\bar{m}} ,$$

$f(v) \equiv (q, \Phi(v))_n$ for any fixed $q \in R^n$ is a convex lower semi-continuous function on $R^{\bar{m}}$.

One method of a search for the saddle point is the <u>algorithm of Uzawa</u> :

$p^0 \in \Lambda$ chosen , we calculate $u^0, p^1, u^1, \ldots$ such that

u^m minimizes $\mathcal{F}(v) + (p^n, \Phi(v))_n$ over $v \in R^{\bar{m}}$, $m = 0,1,\ldots,$

$p^{m+1} = P_\Lambda(p^m + \rho_m\Phi(u^m))$,

where P_Λ is the projection of R^n onto Λ and $\rho_m \in R^1$ is a properly chosen parameter.

<u>*Theorem 4.3*</u> *Under the above assumptions, it holds*

$$u^m \to u$$

provided

$$0 < \alpha_0 \leq \rho_m \leq \alpha_1 ,$$

where α_0 , α_1 *are suitable parameters and* u *minimizes* $\mathcal{F}$ *over* $\mathcal{K}_h(u \equiv a^h)$.

For the proof see [2].

<u>REMARK 4.5</u> The algorithm of Uzawa is advantageous if the projection P_Λ is easy to implement (see e.g. the EXAMPLE (4.6)).

References

[1] S. G. MICHLIN: Numerical realization of variational methods. (Russian) Nauka, Moskva, 1966.

[2] R. GLOWINSKI, J. L. LIONS, R. TREMOLIÈRES: Analyse numérique des inéquations variationelles. Dunod, Paris, 1976.

[3] J. CEA: Optimisation, théorie et algorithmes. Dunod, Paris, 1971.

[4] B. N. PŠENIČNYJ, J. M. DANILIN: Numerical methods in extremal problems. (Russian) Nauka, Moskva, 1975.

[5] G. STRANG, G. FIX: An analysis of the finite element method. Prentice Hall, New York, 1973.

[6] P. G. CIARLET: The finite element method for elliptic problems. North-Holland, Amsterdam, 1977.

[7] I. HLAVÁČEK, J. NEČAS: Mathematical theory of elastic and elasto -plastic bodies. (Czech) SNTL, Praha. English translation: Elsevier, Amsterdam (to appear).

[8] S. FUČÍK, A. KRATOCHVÍL, J. NEČAS:Kachanov's method and its application. Rev. Roum. Math. Pures et Appl. 20 (1975), 907-916.

[9] S. FUČÍK, A. KUFNER: Nonlinear differential equations. (Czech) SNTL, Praha, 1978.

[10] J. M. ORTEGA, W. C. RHEINBOLDT: Iterative solutions of nonlinear equations in several variables. Academic Press, 1970.

[11] U. MOSCO, G. STRANG: One-sided approximations and variational inequalities. Bull. Am. Math. Soc. 80 (1974), 308-312.

[12] R. S. FALK: Error estimates for the approximation of a class of variational inequalities. Math. of Comp. 28 (1974), 963-971.

[13] I. HLAVÁČEK: Dual finite element analysis for unilateral boundary value problems. Apl. mat. 22 (1977), 14-51.

[14] I. HLÁVAČEK: Dual finite element analysis for semi-coercive unilateral boundary value problems. Apl. mat. 23 (1978), 52-71.

[15] J. HASLINGER: Finite element analysis for unilateral problems with obstacles on the boundary. Apl. mat. 22 (1977), 180-188.

[16] F. BREZZI, W. W. HAGER, P. A. RAVIART: Error estimates for the finite element solution of variational inequalities. Numer. Math. 28 (1977), 431-443.

[17] F. SCARPINI, M. A. VIVALDI: Error estimates for the approximation of some unilateral problems. R.A.I.R.O. Anal. Num. 11, (1977), 197-208.

[18] G. ZOUTENDIJK:Method of feasible directions. Elsevier, Amsterdam, 1960.

[19] G. ZOUTENDIJK:Nonlinear programming. SIAM J. Control, 4 (1966), 194-210.