

A New Feasible Directions Algorithm for Effective Solution of Constrained Optimization Problems

J. Orkisz

*Technical University of Cracow, Dept. of Computational and Structural Mechanics,
ul. Warszawska 24, PL-31155 Krakow, Poland*

Abstract

Presented is a new, effective version of the feasible directions (FD) method of solution of constrained optimization problems. Considered is a minimization problem: find $\min F(\underline{x})$ with inequality constraints $g_i(\underline{x}) \leq 0$; $\underline{x} \in \mathbb{R}^n$; $i=1,2,\dots,m$.

The solution method proposed is oriented on search for both a point (FP) within a feasible region (FR) and an optimal solution within the FR. These tasks are realized by two procedures, called "standard" and "auxiliary", that involve determination of a direction vector and selection of a step size. These procedures are supposed at first either to find a FP or to show that there is no FR. Both procedures are based on purely geometrical concepts. They complete each other but they differ by a solution strategy assumed. If a FP has been found the auxiliary procedure is used again, this time in order to find the optimal point \underline{x}^* where the Kuhn-Tucker conditions are satisfied.

Denotations

g_i^I	$\Leftrightarrow g_i \leq 0$,	$i \in M^I$	- a satisfied constraint
g_i^{II}	$\Leftrightarrow g_i > 0$,	$i \in M^{II}$	- a violated constraint
g_i^{III}	$\Leftrightarrow -\epsilon < g_i \leq 0$,	$\epsilon > 0$, $i \in M^{III}$	- an active constraint i.e. $g_i \approx 0$
	∇F			- gradient of the objective function F
	∇g_i			- gradient of a constraint $g_i \leq 0$
	ϵ			- a small positive number

It is assumed that vectors dashed like $\underline{\tilde{d}}$ are normalized i.e. $|\underline{\tilde{d}}| = 1$.

1. Introduction

The problem of evaluation of residual stresses in rail road rails considered in [6,7] was finally posed in the form of the following optimization problem:

$$\text{find } \min F(\underline{x}) \quad , \quad \underline{x} = \{x_j\} \quad , \quad j=1,2,\dots,n \quad (1.1)$$

subjected to inequality constraints

$$g_i(\underline{x}) \leq 0 \quad , \quad i \in M = \{1,2,\dots,m\} . \quad (1.2)$$

Here both objective function (total complementary energy) $F(\underline{x})$ and constraints (yield conditions) are quadratic strictly convex functions. Consequently in that case any local minimum is the global one; the unique solution, if any, is obtained then.

The problem (1.1), (1.2) might be solved by the classical feasible direction method (FDM) [4,9-12,14,15] although a priori unknown initial feasible point (FP) \underline{x}_f is required then at first.

$$\underline{x}_f \Leftrightarrow g_i(\underline{x}_f) \leq 0 \quad , \quad i = 1,2,\dots,m.$$

In the present paper, however, a new original version of the FDM based on purely geometrical concepts is proposed. That version is oriented on search for both

- a)- a point (FP) within a feasible region (FR) (2); and
- b)- the optimal solution within the FR.

In the case (a) two procedures were proposed called "standard" and "auxiliary". They differ by the strategy assumed and complete each other. They are designed either to find a FP or to show that there is no FR.

The steepest descent direction ∇F for an original objective function F is, in the case

(a), replaced by a direction vector $-\sum \nabla \bar{g}_i$ representing the sum of normalized gradient vectors - ∇g_i to nonsatisfied constraints (1.2).

If a FP has been found the auxiliary procedure is used again, that time in order to find the optimal point x^* . The original objective function F is used then. Search is terminated when the Kuhn-Tucker optimality conditions are satisfied.

2. Search for a feasible point x_f

A feasible point x_f is required satisfying all constraints

$$g_i(x_f) \leq 0 \quad \text{for } i \in M.$$

Two step by step search procedures are introduced. In each step a direction vector and a step size should be determined. It is assumed that provided it is feasible, in both procedures the optimal direction vector

$$d = -\sum \nabla \bar{g}_i \quad (2.1)$$

is oriented towards currently violated constraints $g_i \leq 0$ i.e. the original gradient of the objective function ∇F is replaced then by $\sum \nabla \bar{g}_i$.

2.1 Standard procedure

The following strategy of search is assumed:

- an arbitrary starting point x_0 has to be supplied.
- The direction vector d_q is always determined by formula (2.1) since a subsequent point x_q on the step q cannot lie on any of the boundaries of the domains $\Omega_i = \{x : g_i(x) \leq 0\}$, $i \in M$.
- The step size k_q is determined as the mid value (Fig.1).

$$k_q = (k_q^+ + k_q^-) / 2 \quad (2.2)$$

between the step

$$k^+ = \sup_i k_{1q}(i) \quad , \quad k_{1q}(i) \geq 0 \quad , \quad i \in M \quad (2.3)$$

corresponding to the farthest of all points of intersections $x_{1q}(i)$ and the step

$$k^- = \inf_i k_{2q}(i) \quad , \quad k_{2q}(i) \geq 0 \quad , \quad i \in M \quad (2.4)$$

relevant to the closest of all points $x_{2q}(i)$ in the direction d_q . Here $k_{1q}(i)$ and $k_{2q}(i)$ are current step sizes measured from x_q to the intersection points of the line

$$x_{q+1} = x_q + k d_q \quad (2.5)$$

with domains Ω_i

$$g_i(x_q + k d_q) = 0 \implies k_{1q}(i) \quad \text{and} \quad k_{2q}(i) \quad (2.6)$$

Since Ω_i are strictly convex domains there are no more than two solutions for each Ω_i , namely $k_{1q}(i)$ when entering the domain and $k_{2q}(i)$ when leaving it hence $k_{2q}(i) \geq k_{1q}(i)$.

- A new point

$$x_{q+1} = x_q + k_q d_q \quad (2.7)$$

is found. Unless $k^+ = k^-$, which is one of the criteria terminating the procedure, point x_{q+1} is not located on anyone of Ω_i boundaries. That point, however, may leave constraints already satisfied by the point x_q .

In a general case the nonlinear equation (2.6) should be solved. However if constraints (1.2) are quadric

$$g_i(x) = x^t A_i x / 2 + C_i \quad , \quad A_i^t = A_i \quad (2.8)$$

the step sizes may be explicitly evaluated

$$k_{1q}(i) = (-b_i - \sqrt{b_i^2 - 4a_i c_i}) / 2a_i \quad , \quad k_{2q}(i) = (-b_i + \sqrt{b_i^2 - 4a_i c_i}) / 2a_i \quad (2.9)$$

where

$$a_i = d_q^t A_i d_q / 2 \quad , \quad b_i = d_q^t \nabla g_i(x_q) \quad , \quad c_i = g_i(x_q) \quad .$$

CRITERIA OF TERMINATION OF THE STANDARD PROCEDURE

There are three types of answers we may get from the standard procedure: I - a FP is found, II - there is no FR, III - there is no answer after a reasonable number of steps. The following termination criteria have been assumed:

- I - All constraints (1.2) are satisfied in an approximate way $g_i(x) < \epsilon$, $i \in M$.
- II - Gradient of the objective function is small enough $|\nabla F| < \epsilon$.
 - Lower and upper bounds of the step size are equal to each other, but some of constraints are still violated i.e. $k^+ = k^-$ and $\Omega^R \neq \Omega_1 \cap \Omega_2 \cap \dots \cap \Omega_m$.
 - There is no k^+ i.e. all $k_{2q}(i)$ are negative or complex.
- III - Admissible number of steps q_A is exceeded $q > q_A$.

2.2 Auxiliary procedure

The auxiliary procedure is aimed for both a search of an optimal solution within FR. Being always used for the second purpose it may be also used for the first one if the standard procedure proves to be not efficient enough (answer III).

In a case of a FP search the following strategy is assumed then:

- A starting point x_0 is supplied by the standard procedure
- Subsequent points x_q may be located either inside domains Ω_i or on one or more of their boundaries. In the first case the direction vector is given by the formula (2.1), otherwise d_q may be affected by active constraints $g_i^m(x_q) = 0$ dependently on mutual orientation of $\nabla \bar{I}$ and ∇g_i^m vectors. Possible situations distinguished will be discussed later on.

- It is assumed that all actually satisfied constraints $g_i^T \leq 0$ define a current feasible region. As opposed to the standard procedure in the auxiliary procedure search takes place only within this region its boundary including. Thus the step size is determined now as

$$k_q = \begin{cases} (k_q^- + k_q^+) / 2 & \text{if } k_q^+ \text{ exists} \\ k_q^- & \text{if } k_q^+ \text{ does not exist} \end{cases} \quad (2.10)$$

where

$$k_q^+ = \sup_i k_{1q}(i), \quad k_{1q}(i) \geq 0 \text{ and } k_{1q}(i) \leq k_q^- \quad (2.11)$$

corresponds to the farthest of all $k_{1q}(i)$ intersection points but not beyond a current feasible region (Fig.2).

- A new point x_{q+1} is found using formula (2.7) as before.

DIRECTION VECTOR FROM A POINT LOCATED ON ONE BOUNDARY

Consider a point x_q located on one boundary $g_i(x_q) = 0$. Dependently on a mutual orientation

$$\eta_i = (-\nabla \bar{I})^T \nabla g_i \quad (2.12)$$

of the normalized $\nabla \bar{I}$ and ∇g_i vectors the following situations may occur

- a) $0 < \eta_i < 1$ - direction $d_{1q} = -\nabla \bar{I}$ is feasible
- b) $\eta_i = 0$ - this is a limiting case which may be classified as feasible if g_i is convex but not strictly convex in x_q
- c) $-1 < \eta_i < 0$ - direction $-\nabla \bar{I}$ is not feasible, a special formula for d_{1q} is derived
- d) $\eta_i = -1$ - there is no directions d_{1q} both usable and feasible i.e. FR does not exist.

In the case (c)

$$d_{1q} = -\frac{1-\alpha\eta_i}{\alpha-\eta_i} \nabla \bar{I} - \nabla g_i \quad (2.13)$$

This result was based on the following assumptions (Fig.3):

- the direction vector d_{1q} is located in the plane formed by $\nabla \bar{I}$ and ∇g_i vectors;
- projections of $-\nabla \bar{I}$ and $-\nabla g_i$ vectors onto d_{1q} are equal to each other;

Here α is a scaling parameter giving a family of both usable and feasible directions bounded by $\alpha=0$ for d_{1q} perpendicular to ∇g_i and $\alpha=\infty$ for d_{1q} perpendicular to $\nabla \bar{I}$.

DIRECTION VECTOR FROM A POINT LOCATED ON MORE THEN ONE BOUNDARY

Consider a point x_q located on more than one boundary $g_i(x_q)=0$. One may distinguish the following situations (Fig.4)

- a) $0 < \eta_i < 1$ for all $i \in M^m$, then $d_{Nq} = -\nabla \bar{I}$
- b) $-1 < \eta_i < 0$ for only one i out of $i \in M^m$, then $d_{Nq} = d_{1q}$ as in (2.13)
- c) set of parameters $\lambda_j \geq 0$ may be found such that $\eta_i = -c_{ij} \lambda_j$, where $c_{ij} = (\nabla g_i)^T \nabla g_j$. Kuhn-Tucker optimality conditions are satisfied then hence FR does not exist.
- d) otherwise $-d_{Nq}$ is found by the procedure described below.

Consider at first the situation (c), which occurs when vector $-\nabla \bar{I}$ is within a cone formed by vectors ∇g_i , i.e. if $\sum \lambda_j \nabla g_j + \nabla \bar{I} = 0$, $i \in M^m$.

Hence we get simultaneous algebraic equations

$$\begin{aligned} \underline{C} \underline{\lambda} + \underline{\eta} &= 0, \quad \underline{C} = [c_{ij}], \quad c_{ij} = (\nabla g_i)^T \nabla g_j, \quad i, j \in M^m, \\ \underline{\eta} &= (\eta_j), \quad \eta_j = (\nabla g_j)^T \nabla \bar{I}, \quad \underline{\lambda} = (\lambda_j). \end{aligned} \quad (2.14)$$

The following situations may happen

- $\det(C) \neq 0 \implies \lambda = -C^{-1} \eta$
 - = if all $\lambda_i \geq 0$ Kuhn-Tucker optimality conditions are satisfied i.e. direction both feasible and usable does not exist ;
 - = otherwise both feasible and usable direction may be found using approach described in (d);
 - $\det(C) = 0$; Solution is non unique and may be obtained in the form $\lambda = I \lambda_s + \lambda_0$ using the pivotal elimination technique. Here λ_s is the vector of independent variables, I is a transformation matrix, λ_0 - a known constant vector.
- Thus a FR exists if $\lambda = I \lambda_s + \lambda_0 > 0$. Once again this is a FP search problem, although this time in the λ_s space. It may be also solved by the standard and/or auxiliary procedure adopted for linear inequalities.

SOLUTION PROCEDURE IN THE SPACE OF DIRECTIONS \underline{d}

In a situation (d) the following problem is to be solved :

find a direction \underline{d}_{Nq} satisfying the constraints

$$h_i = \underline{d}_{Nq}^T \nabla \bar{g}_i(x_q) < -g_i^{\text{III}}(x_q) \approx 0, \quad i \in M^{\text{III}}, \quad h_T = \underline{d}_{Nq}^T \nabla \bar{f} < 0. \quad (2.15)$$

Now FR, if exists, is convex though not strictly convex since it is formed by hyperplanes $h_i + g_i^{\text{III}} = 0$, and $h_T = 0$. It is also bounded by a normalization condition eg. $|\underline{d}_{Nq}| = 1$, which forms a hypersphere of the radius 1 (Fig.5).

Thus once again we face a FP search problem (now in \underline{d} space), that may be solved by the technique described earlier, with slight modifications due to linearity of constraints (2.15).

In a search process subsequent values $\underline{d}_{p+1} = \underline{d}_p + \alpha_p \underline{y}_p$ are found where α_p is a step size, and \underline{y}_p - a direction vector in the \underline{d} space. Now $\alpha_{1p(i)}$ is defined as a step size when entering to a domain $h_i^{\text{III}} < 0$ (so far violated constraint) while $\alpha_{2p(i)}$ when leaving a domain $h_i^{\text{III}} < 0$ (so far satisfied constraint). What the direction vector \underline{y}_p is concern similar situations may happen as considered before in a case of the direction vector \underline{d} in the \underline{x} space. Thus the solution procedure follows previously described patterns. The only remaining situation, not discussed earlier, takes place when there are more then one active constraints $h_i^{\text{III}} = 0, i \in M^{\text{III}}$. The following linear programming problem is to be solved then

find $\max_{\omega, \underline{y}} \omega$ subjected to

$$\begin{aligned} \omega + \underline{y}^T \nabla \bar{g}_i + h_i &\leq 0, \quad i \in M^{\text{III}}, \\ \omega / \alpha + \underline{y}^T \nabla \bar{f} &\leq 0, \quad -1 \leq y_j \leq 1, \quad \underline{y} = \{y_j\}, \quad j=1,2,\dots,n. \end{aligned} \quad (2.16)$$

Here $\nabla \bar{f} = \sum \nabla \bar{g}_i$ is gradient of a new objective function oriented towards violated constraints $h_i \leq 0, h_i^{\text{III}} < 0, i \in M^{\text{III}}$. Since only a few active constraints $h_i = 0$ may practically happen - problem (2.16), if occurs, may be efficiently solved by the standard LP approach.

Details of the auxiliary procedure as well as some other procedures involved, oriented on refinement of the described algorithm and acceleration of the solution approach are given in [7] and will not be discussed here.

It is worth to mention that an alternative way of evaluation of the direction vector \underline{d} satisfying conditions (2.15) is to apply at once the LP approach rather then to go through a search in the \underline{d} space as described before. Such approach might be logically simpler but would be less efficient.

CRITERIA OF TERMINATION OF THE AUXILIARY PROCEDURE

The following termination criteria have been assumed

- I - If all constraints are satisfied in an approximate way $g_i(x) < \epsilon, i \in M^{\text{III}}, \epsilon > 0$; a FP is found
- II - If Kuhn-Tucker optimality conditions are satisfied; there is no FR
- III - If admissible number of steps Q_A is exceeded.

3. Search for the optimal solution within a feasible region

Consider the initial problem (1.1), (1.2). The procedures described before yield a FP \underline{x}_f if a FR exists. A search should be now continued within the FR in order to find the minimum of a given objective function $F(x)$. Such a problem may be again solved by the auxiliary procedure but with slight adjustments:

- The objective functions is now $\nabla F = \underline{H}_x$ rather then $\nabla \bar{f} = \sum \nabla \bar{g}_i^{\text{III}}$
- Since the minimum of F may occur now also inside the FR the step size k_q is determined as $k_q = \inf(k_{Fq}, k_{\bar{q}})$.

Here k_{Fq} is found as the step that minimizes an unconstrained objective function F along a

direction \underline{d} : $\frac{d}{dk} F(\underline{x} + k\underline{d}) = 0 \Rightarrow k_{Fq}$, hence

$$k_{Fq} = -(\underline{d}^T \underline{H} \underline{x})_q / (\underline{d}^T \underline{H} \underline{d})_q \text{ if } F = \underline{x}^T \underline{H} \underline{x} / 2 .$$

- Termination criteria are basically unchanged, however now
 - = the case I presents only a necessary condition but not a sufficient one;
 - = the case II presents the criterion for the optimal solution point \underline{x}^* .

4. Conclusions

A new original version of the FDM has been presented. It is oriented on a search of the global minimum of a strictly convex function $F(\underline{x})$, $\underline{x} \in \mathbb{R}^n$ subjected to m strictly convex constraints $g_i(\underline{x}) \leq 0$. Both n and m are expected to be large numbers. The procedure also holds if either F or g are convex functions. This procedure is quite different from the classical FDM [14,15]. Thus e.g.: - search is based on a new, purely geometrical approach; - included is a search for a FP enabling start from an arbitrary non-FP; - better efficiency is expected due to almost complete elimination of troublesome LP approach in each step of iteration and replacing it by operationally much simpler although logically more complex algorithm.

The procedure successfully tested before is actually applied to a discrete shakedown analysis of 2D [1-3,5,13] and 3D [6,7] bodies. Discretization has been done through nonconventional hybrid element approach [8]. Since both the objective function (reduced complementary energy) and constraints (yield conditions) have been described by strictly convex quadratic functions. The proposed FD procedure seems to be particularly effective in evaluation of the global minimum.

Acknowledgments

This research has been sponsored by the DOT TSC in Cambridge Ma. The author express special thanks to Dr. P.Tong for the most valuable discussions and to Mr. Dave Jeong for the programming and computational work done.

5. References

- / 1 / BELYTSCHKO, T., "Plane Stress Shakedown Analysis by Finite Elements", *Int.J.Mech.Sci.*, 14, 619-625 (1972).
- / 2 / COHN, M.Z., MAIER, G., Editors, "Engineering Plasticity by Mathematical Programming", Part IV and V, Pergamon Press Publ., 1977.
- / 3 / CORRADI, L., ZAVELANI, A., "A Linear Programming Approach to Shakedown Analysis of Structures", *J.Comp.Meth.Appl.Mech.Engng*, 3, 37-53 (1974).
- / 4 / KIRSCH, U., "Optimum Structural Design", Mc Graw-Hill Publ., 1981.
- / 5 / MARTIN, J.B., "Plasticity: Fundamentals and General Results", The MIT Press Publ., 1975.
- / 6 / ORKISZ, J., "Progress on Elastic-Plastic Estimation of Residual Stresses in Rails", Meeting on Behavior of Materials Under Wheel/Rail Contact Conditions, TSC Pueblo, CO, April 24, 1984.
- / 7 / ORKISZ, J., "Evaluation of Residual Stresses in Railroad Rails", DOT TSC, Cambridge Mass., USA, September 1984, Progress Rpt.
- / 8 / PIAN, T.H.H., TONG, P., "Basic of Finite Element Methods for Solid Continua", *Int.J.Num.Meth.Engng*, 1, 3-28 (1969).
- / 9 / VANDERPLAATS, G.N., "A Robust Feasible Directions Algorithm for Design Synthesis", Proc. AIAA/ASME/ASCE/AHS 24th Structures, Structural Dynamics and Materials Conference, Lake Tahoe, Nevada, May 2-4, 1983.
- / 10 / VANDERPLAATS, G.N., "ADS - a FORTRAN Program for Automated Design Synthesis", Naval Postgraduate School, Monterey Ca, May 1984.
- / 11 / VANDERPLAATS, G.N., "Numerical Optimization Techniques for Engineering Design: with Applications", Mc Graw-Hill Publ., 1984.
- / 12 / WISMER, D.A., CHATTERGY, R., "Introduction to Nonlinear Optimization", North-Holland Publ., 1978.
- / 13 / ZARKA, J., CASIER, J., "Elastic-Plastic Response of a Structure to Cyclic Loading Practical Rules", *Mechanics Today Vol.6*, edited by S.Nemat-Nasser, Pergamon Press Publ., 93-198 (1981).
- / 14 / ZOUTENDIJK, G., "Methods of Feasible Directions", Elsevier, Amsterdam, 1960.
- / 15 / ZOUTENDIJK, G., "Mathematical Programming Methods", North-Holland Publ., 1976.

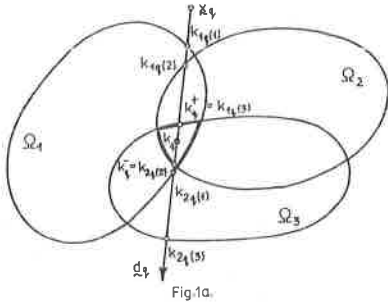


Fig. 1a.

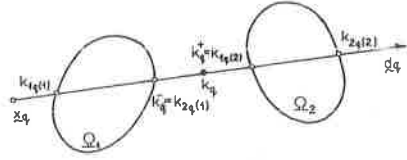


Fig. 1b.

Fig. 1 Search for a FP in the standard procedure :
 a) all domains intersect each other
 b) not all domains intersect each other

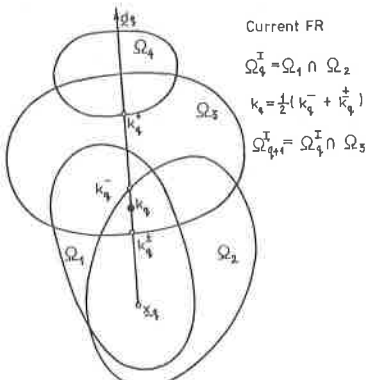


Fig. 2a.

Current FR
 $\Omega_q^I = \Omega_1 \cap \Omega_2$
 $k_q = \frac{1}{2}(k_q^- + k_q^+)$
 $\Omega_{q+1}^I = \Omega_q^I \cap \Omega_3$

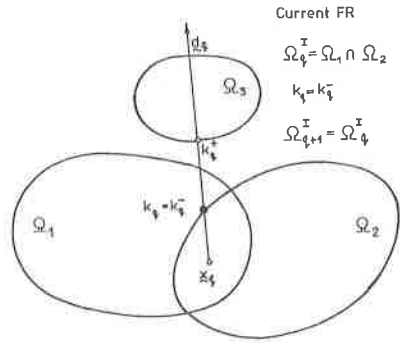


Fig. 2b.

Fig. 2 Search for a FP in the auxiliary procedure :
 a) k_q does exist
 b) k_q does not exist

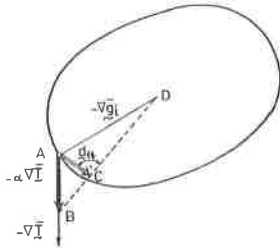


Fig. 3 Evaluation of a direction vector d

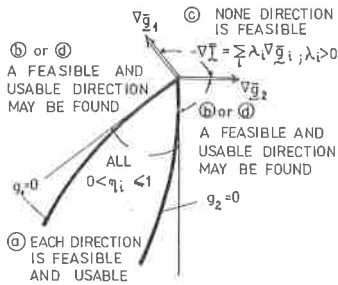


Fig. 4.

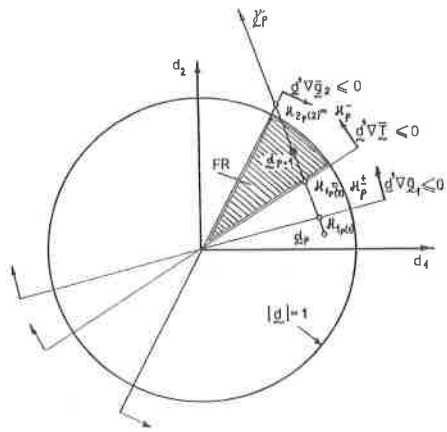


Fig. 5 Search for a FP within the FR in the d space