An Extension of the Method of Feasible Directions

E. Blum
Universidad Nacional de Ingenieria
Lima/Perú

W. Oettli
Universität Mannheim
68 Mannheim/Germany

In this contribution we are going to discuss the extension of the method of feasible directions [1], [2], [3] to programming problems involving an infinite number of constraints. Problems of this type arise frequently in applications. We shall be working with arbitrary convex approximations instead of with linearizations, simply to emphasize the fact that the feasible direction method belongs to that class of methods where not differentiability but rather convex-likeness of the functions involved is the essential property.

Our programming problem has the following form:

$$\text{(P)} \qquad \min \{F(x) \mid x \in C, f(t,x) \le 0 \ \forall t \in T\} .$$

With $S = \{x \mid x \in C, f(t,x) \le 0 \ \forall t \in T\}$ the admissible domain of (P) we introduce for all $x \in S$ approximations $\Phi(x,\xi), \phi(t,x,\xi)$ for the functions $F(\xi), f(t,\xi)$. We assume that C is a compact convex set of some normed (metrizable) linear space, that T is a compact metric space, and that the functions $F(\xi), f(t,\xi), \Phi(x,\xi), \phi(t,x,\xi)$ are jointly continuous in all their arguments, with $\xi \in C$, $x \in S$, $t \in T$.

We shall be particularly interested in certain elements of S, henceforth denoted by $\hat{x}$, which will be limit points of our iterative procedure. Concerning these points $\hat{x} \in S$ we require in addition that the functions Φ and ϕ are "good" approximations in the sense that

$$|\Phi(\hat{x},\xi) - F(\xi)| \le o(\xi-\hat{x}), |\phi(t,\hat{x},\xi) - f(t,\xi)| \le o(\xi-\hat{x})$$

uniformly for all $t \in T$. Moreover $\Phi(\hat{x},\xi)$ and $\phi(t,\hat{x},\xi)$ have to be *convex* with regard to ξ.

For $\hat{x} \in S$ let us define the set of binding constraints

$$\hat{T} = \{t \in T \mid f(t,\hat{x}) = 0\} ,$$

and consider the following system in ξ:

$$\text{(1)} \qquad \xi \in C, \Phi(\hat{x},\xi) - F(\hat{x}) < 0, \phi(t,\hat{x},\xi) < 0 \ \forall t \in \hat{T} .$$

Under the assumptions made it is not difficult to prove the following

Lemma 1: Let ξ be a solution of (1). Then there exists $x \in [\hat{x},\xi]$ satisfying

$$x \in C, F(x) - F(\hat{x}) < 0, f(t,x) < 0 \ \forall t \in T .$$

From this lemma one obtains immediately the following necessary optimality criterion which may be considered as a generalization of Kolmogorov's criterion for best Chebyshev-approximations.

Theorem 1: If $\hat{x} \in S$ is an optimal solution for the programming problem (P), then system (1) is inconsistent.

We note that the inconsistency of (1) is also a sufficient condition for optimality, if $F(\xi)$ and $f(t,\xi)$ are convex with regard to ξ, and if Slater's assumption is satisfied: there exists $\tilde{x}\in C$ satisfying $f(t,\tilde{x}) < 0\ \forall t\in\hat{T}$.

Any limit point $\hat{x}$ of the approximation procedure to be described is a stationary point in the sense that it meets the necessary optimality condition of Theorem 1.

Let us now describe the iterative scheme. We choose a *positive* number $\alpha > 0$, and sequences $\theta_k \geq 0, \rho_k \geq 0$ such that $\theta_k \to 0, \sum_k \rho_k < +\infty$. $x^o \in S$ is arbitrary. Given $x^k \in S$ we define x^{k+1} according to the following rules: Let

$$T^k = \{t\in T \mid f(t,x^k) \geq -\alpha\} ,$$

$$H^k(\xi) = \max\{\Phi(x^k,\xi) - F(x^k), \phi(t,x^k,\xi) \colon t\in T^k\} .$$

Let $\xi^k \in C$ be such that

$$H^k(\xi^k) \leq \min\{H^k(\xi) \mid \xi\in C\} + \theta_k , \tag{2}$$

and define $x^{k+1}\in\left[x^k,\xi^k\right] \cap S$ such that

$$F(x^{k+1}) \leq \min\{F(x) \mid x\in\left[x^k,\xi^k\right] \cap S\} + \rho_k . \tag{3}$$

Obviously x^{k+1} is well defined, and is again in S. Since S is compact, the sequence $\{x^k\}$ has a cluster point $\hat{x}\in S$.

<u>Theorem 2:</u> If $\hat{x}$ is a cluster point of the sequence $\{x^k\}$, then $\hat{x}$ satisfies the necessary optimality criterion of Theorem 1.

Proof: In addition to $H^k(\xi)$ let us define the continuous functions

$$H^\infty(x,\xi) = \max\{\Phi(x,\xi) - F(x), \phi(t,x,\xi) \colon t\in T\} ,$$

$$\hat{H}(x,\xi) = \max\{\Phi(x,\xi) - F(x), \phi(t,x,\xi) \colon t\in\hat{T}\} .$$

Since C is compact we can choose a subsequence $x^{\bar{k}}$ such that

$$x^{\bar{k}} \to \hat{x},\ \xi^{\bar{k}} \to \hat{\xi}\in C .$$

From (2) follows

$$H^k(\xi^k) \leq H^k(\xi) + \theta_k\ \forall\xi\in C .$$

The continuity of $f(t,x)$ over $\hat{T}\times C$, the compactness of $\hat{T}$, and the convergence of $x^{\bar{k}}$ to $\hat{x}$ imply that $\hat{T}\subset T^{\bar{k}}$ for all sufficiently large $\bar{k}$. Also $T^{\bar{k}}\subset T$. Therefore

$$\hat{H}(x^{\bar{k}},\xi^{\bar{k}}) \leq H^\infty(x^{\bar{k}},\xi) + \theta_{\bar{k}}\ \forall\xi\in C$$

for all sufficiently large $\bar{k}$. Passing to the limit we obtain

$$\hat{H}(\hat{x},\hat{\xi}) \leq H^\infty(\hat{x},\xi)\ \forall\xi\in C . \tag{4}$$

From (3) follows

$$F(x^{k+1}) \leq F(x^k) + \rho_k ,$$

thus a fortiori

$$F(\overline{x^{k+1}}) \le F(\overline{x}^{k+1}) + \sum_{\mu=k+1}^{\infty} \rho_\mu .$$

Again by (3)

$$F(\overline{x^{k+1}}) \le F(x) + \sum_{\mu=k+1}^{\infty} \rho_\mu \quad \text{for all} \quad x \in \left[\overline{x}^k, \overline{\xi}^k\right] \quad \text{satisfying} \quad \max_{t \in T} f(t,x) \le 0 .$$

Passing to the limit we obtain by continuity

$$F(\hat{x}) \le F(x) \quad \text{for all} \quad x \in \left[\hat{x}, \hat{\xi}\right] \quad \text{satisfying} \quad \max_{t \in T} f(t,x) < 0 .$$

This means that the system

$$(5) \qquad x \in \left[\hat{x}, \hat{\xi}\right],\ F(x) - F(\hat{x}) < 0,\ f(t,x) < 0\ \forall t \in T$$

is inconsistent. Assume now that (1) has a solution. A slight variant of the proof of Lemma 1 shows that (1) has still a solution if we replace $\hat{T}$ by T. This means there exists $\xi \in C$ satisfying $H^\infty(\hat{x},\xi) < 0$. By (4) then $\hat{H}(\hat{x},\hat{\xi}) < 0$. Lemma 1 gives then the existence of x satisfying (5), a contradiction. Thus (1) is inconsistent.
q.e.d.

We may study the rate of convergence of $F(x^k)$ if we require in addition:

$$F(\xi), f(t,\xi), \Phi(x,\xi), \phi(t,x,\xi) \text{ are convex with respect to } \xi ;$$

$$\Phi(x,x) = F(x), \phi(t,x,x) = f(t,x)\ \forall x \in S ; \text{ the set}$$

$$S_o = \{x \in C \mid f(t,x) \le 0\ \forall t \in T,\ F(x) \le F(x^o)\}$$

is bounded; $\exists \tilde{x} \in C$: $f_t(\tilde{x}) < 0\ \forall t \in T$; $\theta_k = 0$ and $\rho_k = 0\ \forall k$.

We use the abbreviations

$$\tau^k = H^k(\xi^k),\ \delta^k = F(x^k) - \hat{F} ,$$

where $\hat{F}$ is the optimal value of (P). Then

$$\tau^k \le 0,\ \tau^k \to 0;\ \delta^k \ge 0,\ \delta^k \to 0 .$$

We obtain the following results.

<u>Lemma 2:</u> If there exist constants $\mu \ge 0$, $0 < m \le 1$, such that (i) $\Phi(x,\xi) - \mu|\xi - x|^2$ is convex with respect to ξ, (ii) $\phi(t,x,\xi) - \mu|\xi - x|^2 \le f(t,\xi)$, (iii) $\Phi(x,\xi) - (1-m)\mu|\xi - x|^2 \le F(\xi)$, then $\tau^k \le \rho(-\delta^k)$ for some $\rho > 0$.

<u>Lemma 3:</u> (a) If $F(\xi) \le \Phi(x,\xi) + M|\xi - x|^2$, $f(t,\xi) \le \phi(t,x,\xi) + M|\xi - x|^2$, then $\delta^{k+1} - \delta^k \le -\gamma(\tau^k)^2$ for some $\gamma > 0$. (b) If, in addition, there exists $\mu > 0$ such that $\Phi(x,\xi) - \mu|\xi - x|^2$ and $\phi(t,x,\xi) - \mu|\xi - x|^2$ are convex with regard to ξ, then $\delta^{k+1} - \delta^k \le \gamma\tau^k$ for some $\gamma > 0$.

From these follows

<u>Theorem 3:</u> If the assumptions of Lemma 2 and Lemma 3(a) hold, then $\delta^{k+1} \le (1 - \rho\delta^k)\,\delta^k$ for some $\rho > 0$. If the assumptions of Lemma 2 and Lemma 3(b) hold, then $\delta^{k+1} \le (1 - \rho)\,\delta^k$ for some $\rho > 0$.

This extends some results of Pironneau - Polak [4]. Proofs are too lengthy to be given here; they will be reproduced elsewhere.

[1] M. Frank, P. Wolfe: An algorithm for quadratic programming. *Naval Res. Logistics Quart.* 3 (1956), 95-110 .

[2] G. Zoutendijk: *Methods of Feasible Directions.* Elsevier, Amsterdam, 1960.

[3] H.P. Künzi, W. Oettli: *Nichtlineare Optimierung* (Lecture Notes in Operations Research and Mathematical Systems, 16), pp. 65-70. Springer, Berlin, 1969.

[4] O. Pironneau, E. Polak: On the rate of convergence of certain methods of centers. *Math. Programming* 2 (1972), 230-257.