CONVEX PROGRAMMING IN HILBERT SPACE

BY A. A. GOLDSTEIN[1]

Communicated by V. Klee, May 1, 1964

This note gives a construction for minimizing certain twice-differentiable functions on a closed convex subset C, of a Hilbert Space, H. The algorithm assumes one can constructively "project" points onto convex sets. A related algorithm may be found in Cheney-Goldstein [1], where a constructive fixed-point theorem is employed to construct points inducing a minimum distance between two convex sets. In certain instances when such projections are not too difficult to construct, say on spheres, linear varieties, and orthants, the method can be effective. For applications to control theory, for example, see Balakrishnan [2], and Goldstein [3].

In what follows P will denote the "projection" operator for the convex set C. This operator, which is well defined and Lipschitzian, assigns to a given point in H its closest point in C (see, e.g., [1]). Take $x \in H$ and $y \in C$. Then $[x-y,\ P(x)-y] \geqq \|P(x)-y\|^2$. In the nontrivial case this inequality is a consequence of the fact that C is supported by a hyperplane through $P(x)$ with normal $x-P(x)$. Let f be a real-valued function on H and x_0 an arbitrary point of C. Let S denote the level set $\{x \in C: f(x) \leqq f(x_0)\}$, and let $\hat{S}$ be any open set containing the convex hull of S. Let $f'(x,\ \cdot) = [\nabla f(x),\ \cdot\,]$ signify the Fréchet derivative of f at x. A point z in C will be called stationary if $P(z-\rho\nabla f(z)) = z$ for all $\rho > 0$; equivalently, when f is convex the linear functional $f'(z,\ \cdot)$ achieves a minimum on C at z.

THEOREM. *Assume f is bounded below. For each $x \in \hat{S}$, h in H and for some $\rho_0 > 0$, assume that $f'(x, h)$ exists in the sense of Fréchet, $f''(x, h, h)$ exists in the sense of Gâteaux, and $|f''(x, h, h)| \leqq \|h\|^2/\rho_0$. Choose σ and ρ_k satisfying $0 < \sigma \leqq \rho_0$ and $\sigma \leqq \rho_k \leqq 2\rho_0 - \sigma$. Set $x_{k+1} = P(x_k - \rho_k \nabla f(x_k))$. Then*:

(i) *The sequence x_k belongs to S, $(x_{k+1} - x_k)$ converges to* 0, *and $f(x_k)$ converges downward to a limit L.*

(ii) *If S is compact, z is a cluster point of $\{x_k\}$, and ∇f is continuous in some neighborhood of z, then z is a stationary point. If z is unique, x_k converges to z, and z minimizes f on C.*

(iii) *If S is convex and $f''(x, h, h) \geqq \mu\|h\|^2$ for each $x \in S$, $h \in H$ and some $\mu \geqq 0$, then $L = \inf\{f(x) : x \in C\}$.*

(iv) *Assume* (iii) *with S bounded. Weak cluster points of $\{x_k\}$ minimize f on C.*

---

[1] Present address, University of Washington, Seattle. This research was supported by grant AF-AFOSR-62-348.

(v) *Assume* (iii) *with μ positive and ∇f bounded on S. Then $f(z)=L$ for some z in S, x_k converges to z, and z is unique.*

PROOF. Assume x_k belongs to S and that x_k is not stationary. Let $\nabla f(x_k)=\nabla f_k$, $x(\rho)=P(x_k-\rho\nabla f_k)$, $\delta(\rho)=x(\rho)-x_k$ and $\Delta(\rho)=f(x_k)-f(x(\rho))$. If we notice that $-\rho[\nabla f_k,\ \delta(\rho)]\geqq\|\delta(\rho)\|^2$ and invoke Taylor's theorem, we obtain $\Delta(\rho)\geqq\|\delta(\rho)\|^2\{\rho^{-1}-f''(\xi(\rho),\ \delta(\rho),\ \delta(\rho))/2\|\delta(\rho)\|^2\}$. Here $\xi(\rho)=x_k+t\delta(\rho)$ with $t\in(0,1)$. For some ρ sufficiently small and positive, $\Delta(\rho)$ is positive and continuous. Let $\hat{\rho}$ denote the least positive ρ satisfying $\Delta(\rho)=0$, if such exists. If $\hat{\rho}$ exists, $\Delta(\hat{\rho})=0$ implies that $\hat{\rho}\geqq 2\rho_0$. Thus if $\sigma\leqq\rho\leqq 2\rho_0-\sigma$, $\Delta(\rho)>0$ and $x(\rho)\in S$, whence $\Delta(\rho_k)\geqq\|x_{k+1}-x_k\|^2\sigma/4\rho_0^2$, proving (i).

The proof of (ii) being straightforward, we proceed with the proof of (iii). Suppose that $L\neq\inf\{f(x): x\in C\}$ and choose $z\in C$ such that $f(z)<L$. Then $0>f(z)-f(x_k)\geqq[\nabla f_k,\ z-x_k]$. If $\lim\inf[\nabla f_k,\ z-x_k]=\beta$ were non-negative, a contradiction would be manifest. But the inequality $[\rho_k\nabla f_k,\ z-x_{k+1}]\geqq[x_k-x_{k+1},\ z]+[x_{k+1},\ x_{k+1}-x_k]$ holds because either $x_k-\rho_k\nabla f_k-x_{k+1}$ is the normal to C at x_{k+1}, or it is 0. If the sequence x_k is bounded, clearly $\beta=0$; otherwise choose a subsequence satisfying $\|x_{k+1}\|>\|x_k\|$. Then $\beta\geqq 0$.

To prove (iv) we observe that f is lower semi-continuous on S if and only if the set $S_m=\{x\in S: f(x)\leqq m\}$ is closed in S for each m. Since f is convex and continuous, S_m is closed and convex, and is thus weakly closed. Hence f is weakly l.s.c. If x_k converges weakly to z, then $\lim\inf f(x_k)=L\geqq f(z)$.

Assume the hypotheses of (v). If $s>k$, we may write that $0>f(x_s)-f(x_k)\geqq[\nabla f_k,\ x_s-x_k]+(1/2)\mu\|x_s-x_k\|^2$, whence $\{x_s\}$ is bounded. Invoking again the supporting hyperplane at x_{k+1}, $[\rho_k\nabla f_k,\ x_s-x_k]\geqq[\rho_k\nabla f_k,\ x_{k+1}-x_k]+[x_{k+1}-x_k,\ x_{k+1}-x_s]$. Thus when k is sufficiently large $\|x_s-x_k\|<\epsilon$. There exists therefore $z\in S$ minimizing f on C, and $f(x)\geqq f(z)+[\nabla f(z),\ x-z]+(1/2)\mu\|x-z\|^2$. Since $[\nabla f(z),\ x-z]\geqq 0$, $f(x)-f(z)\geqq(1/2)\mu\|x-z\|^2$; and therefore z is unique.

REFERENCES

1. E. W. Cheney and A. A. Goldstein, *Proximity maps for convex sets*, Proc. Amer. Math. Soc. **10** (1959), 448–450.
2. A. V. Balakrishnan, *An operator theoretic formulation of a class of control problems and a steepest descent method of solution*, J. SIAM Control Ser. A **1** (1963), 109–127.
3. A. A. Goldstein, *Minimizing functionals on Hilbert space*, Computer methods in optimization problems, Academic Press, New York, 1964, pp. 159–165.

UNIVERSITY OF TEXAS