PROJECTIVE METRICS IN DYNAMIC PROGRAMMING

BY RICHARD BELLMAN AND THOMAS A. BROWN

Communicated by E. Hewitt, April 5, 1965

It has been shown by Birkhoff [2], [3] that Hilbert's projective metric [4] may be applied to a variety of problems involving linear mappings of a function space into itself. In this note we shall point out that essentially the same metric may be applied to some nonlinear mappings which frequently arise in dynamic programming [1].

Let X be some set, and let P denote the set of all nonnegative real-valued functions which have domain X and are not identically zero. We define an extended real-valued function θ on $P \times P$ as follows:

$$\theta(f, g) = \log \left[\left(\sup_{x \in X} \frac{f(x)}{g(x)} \right) \cdot \left(\sup_{x \in X} \frac{g(x)}{f(x)} \right) \right].$$

In computing the ratios, we take $0|0$ to be 1, and $a|0$ to be ∞ if $a \neq 0$. It is easy to show that θ is an extended pseudo-metric on P. $\theta(f, g) = 0$ implies that $f = \lambda g$ for some constant $\lambda > 0$. We say that a subset P^* of P is "metric" if θ is an extended metric on P^*. That is, if for any $f, g \in P^*$, $\theta(f, g) = 0$ if and only if $f = g$.

Let L be a map of P into P. If

$$\sup_{x \in X} \frac{Lf(x)}{Lg(x)} < \sup_{x \in X} \frac{f(x)}{g(x)} \quad \text{for all } f, g \in P$$

such that $0 < \theta(f, g) < \infty$ then we say L is "ratio reducing on P." Note that if L is ratio reducing on P it follows at once that $\theta(Lf, Lg) < \theta(f, g)$ for all $f, g \in P$ such that $0 < \theta(f, g) < \infty$.

Thus L is a contraction mapping with respect to the pseudo-metric θ. Similar definitions apply on any subset of P. Many linear transformations have been shown [2], [3] to be ratio reducing (or at least ratio nonincreasing). A family $\{L_\lambda\}$ (λ ranging over some set of parameters Λ) is said to be "uniformly ratio reducing" if, given f, g,

$$\sup_{x \in X} \frac{L_\lambda(f(x))}{L_\lambda(g(x))} \leqq \sup_{x \in X} \frac{f(x)}{g(x)} - \delta_{f,g} \quad \text{for all } \lambda \in \Lambda,$$

where $\delta_{f,g} > 0$ may depend on f and g but does *not* depend on λ. Note that if Λ is a finite set then the family $\{L_\lambda\}$ is uniformly ratio reducing if each of its members is ratio reducing.

THEOREM. *If the family* $\{L_\lambda : \lambda \in \Lambda\}$ *is uniformly ratio reducing,* *then the transformation* L^1 *defined by*

$$L^1(f(x)) = \sup_{\lambda\in\Lambda} L_\lambda(f(x))$$

is ratio reducing. If in addition $L_\lambda(g(x)) > \delta_g > 0$ *for each* $g \in P$ *and all* $\lambda \in \Lambda$, *then the transformation* L^2 *defined by*

$$L^2(f(x)) = \inf_{\lambda\in\Lambda} L_\lambda(f(x))$$

is also ratio reducing.

The proof of the theorem is by straightforward computation. To illustrate the application of this theorem to dynamic programming, let us consider a class of problems referred to as "equations of type III" [1, pp. 125–129]. Suppose we are confronted with a system which may be in any one of $N+1$ states (call the states $s_0, s_1, \cdots, s_N$), and we are trying to drive the system into state s_0. At each stage, we begin by knowing a probability distribution $p = (p_0, p_1, \cdots, p_N)$, where p_i = probability that the system is in state s_i. We may either observe the system (at a cost $b > 0$), or we may perform an operation T_i on it which will alter the probability distribution in some way at a cost $a_i > 0$ $(i = 1, 2, \cdots, n)$. Then if $f(p)$ represents the expected cost of driving the system into state s_0 given that it is initially "known" to be in state s_i with probability p_i, we see that f must obey the functional equation

$$(*) \qquad f(p) = \inf \left\{ \sum_{i=1}^{N} p_i f(\hat{s}_i) + b, f(T_i p) + a_i \right\}$$

where $\hat{s}_i$ denotes the probability distribution which assigns probability 1 to state s_i.

THEOREM. *There is at most one bounded positive solution to the equation* (*).

PROOF. Let X be the set of all possible distributions over the $N+1$ possible states with the exception of $(1, 0, \cdots, 0)$. This point $(\hat{s}_0)$ is in the closure of X. Since the final operation on the system must be an observation, we see that $f(p) \geqq b$. If f is bounded, it immediately follows that $\lim_{p\to s_0} f(p) = b$. Let us restrict our attention to the metric subset P^* of P consisting of bounded f such that $\lim_{p\to \hat{s}_0} f(p) = b$.

$$L_0(f(p)) = \sum_{i=1}^{n} p_i f(\hat{s}_i) + b,$$

$$L_i(f(p)) = f(T_i p) + a_i, \qquad i = 1, 2, \cdots, n,$$

are all ratio-reducing on P^*. Thus by our Theorem above

$$L(f(p)) = \inf_{i=0,1,\dots,n} L_i(f(p))$$

is ratio-reducing on P^*. Hence, if f and g are distinct elements of P^*, then $\theta(Lf, Lg) < \theta(f, g)$, which proves there can be at most one bounded solution to $f = Lf$.

A similar method may be applied when the system may be in any one of a continuum of states. Note that in addition to proving the uniqueness of the solution (if any) to (*), the above argument shows that if $g \in P^*$, and $\{L^n g\}$ contains a uniformly convergent subsequence, then $\{L^n g\}$ converges uniformly to the solution of (*).

Bibliography

1. R. Bellman, *Dynamic programming*, Princeton Univ. Press, Princeton, N. J., 1957.

2. G. Birkhoff, *Extensions of Jentzsch's theorem*, Trans. Amer. Math. Soc. **85** (1957), 219–227.

3. ———, *Uniformly semi-primitive multiplicative processes*, Trans. Amer. Math. Soc. **104** (1962), 37–51.

4. D. Hilbert, *Neue Begründung der Bolya-Lobatschefskyschen Geometrie*, Math. Ann. **57** (1903), 137–150.

Rand Corporation, Santa Monica, California