Best Rational Approximation and Strict Quasi-Convexity

MRC Technical Summary Report #1157

Ian Barrodale

December 1971

This internal report was later published in Society for Industrial Applied Mathematics (SIAM) Journal on Numerical Analysis:

March 1973, 10(1), 8–12.

https://epubs.siam.org/doi/abs/10.1137/0710002

THE UNIVERSITY OF WISCONSIN
MATHEMATICS RESEARCH CENTER

Contract No.: DA-31-124-ARO-D-462

# BEST RATIONAL APPROXIMATION AND STRICT QUASI-CONVEXITY

I. Barrodale



MRC Technical Summary Report # 1157
December 1971

Received June 15, 1971

Madison, Wisconsin 53706

ABSTRACT

If a continuous function is strictly quasi-convex on a convex set $\Gamma$, then every local minimum of the function must be a global minimum. Furthermore, every local <u>maximum</u> of the function on the interior of $\Gamma$ must also be a global minimum. Here, we prove that any minimax rational approximation problem defines a strictly quasi-convex function with the property that a best approximation (if one exists) is a minimum of that function. The same result is not true in general for best rational approximation in other norms.

# BEST RATIONAL APPROXIMATION
# AND
# STRICT QUASI-CONVEXITY

I. Barrodale

It is natural to attempt to solve nonlinear best approximation problems using numerical minimization techniques. One obvious drawback is that local optima may be encountered which are not global minima. Thus, in investigating this possibility for a given type of nonlinear approximating function, the concept of strict quasi-convexity is important.

We shall assume throughout that all functions are real-valued, continuous, and defined on subsets of Euclidean spaces. (Some of our results do not require these assumptions, but this is the normal setting for best approximation problems).

## 1. Preliminary results concerning strictly quasi-convex functions.

Definition 1: $\theta(c)$ is a strictly quasi-convex function on a convex set $\Gamma$ if and only if, for every $c_1, c_2 \in \Gamma$, $\theta(c_1) < \theta(c_2) \Rightarrow \theta(\frac{c_1 + c_2}{2}) < \theta(c_2)$.

Definition 2: $\theta(c)$ is a quasi-convex function on a convex set $\Gamma$ if and only if, for every $c_1, c_2 \in \Gamma$, $\theta(c_1) \leq \theta(c_2) \Rightarrow \theta(\frac{c_1 + c_2}{2}) \leq \theta(c_2)$.

Theorem 1: $\theta(c)$ is a quasi-convex function on a convex set $\Gamma$ if and only if $K_\alpha = \{ c \mid c \in \Gamma, \ \theta(c) \leq \alpha \}$ is a convex set for every real $\alpha$.

Sponsored by the United States Army under Contract No.: DA-31-124-ARO-D-462 and by the Department of Mathematics, University of Victoria, Victoria, B. C., Canada.

**Theorem 2:** If $\theta(c)$ is a strictly quasi-convex function on a convex set $\Gamma$ then $\theta(c)$ is a quasi-convex function on $\Gamma$.

**Theorem 3:** If $\theta(c)$ is a strictly quasi-convex function on a convex set $\Gamma$ then every local minimum is a global minimum of $\theta(c)$ on $\Gamma$.

The proofs of these theorems are given in Mangasarian [3]; the converse of Theorem 2 is not valid, and Theorem 3 does not hold in general if $\theta(c)$ is only quasi-convex.

The remainder of this section contains some new results concerning strictly quasi-convex functions.

**Theorem 4:** If $\theta(c)$ is a strictly quasi-convex function on a convex set $\Gamma$ then for every $c_1, c_2 \in \Gamma$, $\theta(c_1) < \theta(\frac{c_1 + c_2}{2}) \Rightarrow \theta(\frac{c_1 + c_2}{2}) < \theta(c_2)$.

Proof (by contradiction). Assume either (i) $\theta(\frac{c_1 + c_2}{2}) = \theta(c_2)$ or (ii) $\theta(\frac{c_1 + c_2}{2}) > \theta(c_2)$.

(i) $\theta(c_1) < \theta(\frac{c_1 + c_2}{2}) \Rightarrow \theta(c_1) < \theta(c_2)$ (since $\theta(\frac{c_1 + c_2}{2}) = \theta(c_2)$)

$\Rightarrow \theta(\frac{c_1 + c_2}{2}) < \theta(c_2)$, which contradicts (i).

(ii) Put $\theta(\frac{c_1 + c_2}{2}) + \max \{\theta(c_1), \theta(c_2)\} = 2\alpha > 0$, and define

$K_\alpha = \{c \mid c \in \Gamma, \theta(c) \le \alpha\}$.

Clearly $c_1, c_2 \in K_\alpha$ and $\frac{c_1 + c_2}{2} \notin K_\alpha$, and so $K_\alpha$ is not a convex set. Thus $\theta(c)$ is not quasi-convex on $\Gamma$, and so it is not strictly quasi-convex either. This contradicts the hypothesis, and the proof is complete.

Theorem 5: If $\theta(c)$ is a strictly quasi-convex function on a convex set $\Gamma$ then every local maximum of $\theta(c)$ on the interior of $\Gamma$ is a global minimum of $\theta(c)$ on $\Gamma$.

Proof If $c^*$ is an interior local maximum, then for some real $\delta > 0$ there exists an open ball $B_\delta(c^*) \subset \Gamma$ such that $c \in B_\delta(c^*) \Rightarrow \theta(c) \leq \theta(c^*)$. Now if $\theta(c) = \theta(c^*)$ for every $c \in B_\delta(c^*)$ it follows that $c^*$ is also a local minimum, and by Theorem 3 it is therefore a global minimum of $\theta(c)$ on $\Gamma$.

It remains to show that $\theta(c) \nless \theta(c^*)$ for any $c \in B_\delta(c^*)$. Suppose there does exist $c_1 \in B_\delta(c^*)$ such that $\theta(c_1) < \theta(c^*)$; then there also exists $c_2 \in B_\delta(c^*)$ for which $c^* = \frac{c_1 + c_2}{2}$. (Obviously, $c_2$ ($\neq c_1$) is located a distance $\|c_1 - c^*\|$ from $c^*$ on the ray through $c_1$ and $c^*$). But, by Theorem 4, $\theta(c^*) < \theta(c_2)$ and so $c^*$ is not a local maximum. This contradiction completes the proof.

We observe that both Theorem 4 and Theorem 5 may not be true if either $\theta(c)$ is not continuous or $\theta(c)$ is only quasi-convex.

Theorem 6: Let X be compact, $\Gamma$ be convex, and let $\theta(x,c)$ be defined on $X \times \Gamma$ so that,

(i) for every $x \in X$, $\theta(x,c)$ is a strictly quasi-convex function on $\Gamma$,

and (ii) for every $c \in \Gamma$, $\theta(x,c)$ is a continuous function on X.

Then $\psi(c) = \max_{x \in X} \theta(x,c)$ is strictly quasi-convex on $\Gamma$.

Proof  Assume $\psi(\frac{c_1 + c_2}{2}) = \theta(x^*, \frac{c_1 + c_2}{2})$ for some $x^* \in X$.

Now $\psi(\frac{c_1 + c_2}{2}) \geq \psi(c_2) \Rightarrow \theta(x^*, \frac{c_1 + c_2}{2}) \geq \theta(x^*, c_2)$

$\Rightarrow \quad \theta(x^*, c_1) \geq \theta(x^*, \frac{c_1 + c_2}{2})$ (by Theorem 4)

$\Rightarrow \quad \psi(c_1) \geq \psi(\frac{c_1 + c_2}{2})$ .

Thus, for every $c_1, c_2 \in \Gamma$, $\psi(\frac{c_1 + c_2}{2}) \geq \psi(c_2) \Rightarrow \psi(c_1) \geq \psi(c_2)$ and the desired result follows from the contrapositive of this implication.

Remark: In general, the sum or integral of $\theta(x,c)$ over $X$ is not a strictly quasi-convex function on $\Gamma$.

Finally, let k-dimensional Euclidean space be denoted by $E^k$, and its (strictly) positive orthant by $E^k_+$.

Theorem 7: $\phi(u,v) = \frac{|u|}{v}$ is a strictly quasi-convex function on $E^1 \times E^1_+$.

Proof $\phi(u_1,v_1) < \phi(u_2,v_2) \Rightarrow \frac{|u_1|}{v_1} < \frac{|u_2|}{v_2}$

$$\Rightarrow \frac{|u_1| + |u_2|}{v_1 + v_2} < \frac{|u_2|}{v_2}$$

$$\Rightarrow \frac{|u_1 + u_2|}{v_1 + v_2} < \frac{|u_2|}{v_2}$$

$$\Rightarrow \phi(\frac{u_1 + u_2}{2}, \frac{v_1 + v_2}{2}) < \phi(u_2,v_2) .$$

Mangasarian [ 4 ] has shown how to calculate global minima of certain nonlinear fractional programming problems. His results are a consequence of the fact that $\phi^*(u,v) = \frac{u}{v}$ is a pseudo-convex function; see also Mangasarian [5].

2. Best rational approximation.

Given a compact set X, and functions $P_1(x), P_2(x), \ldots, P_m(x)$ and $Q_1(x), Q_2(x), \ldots, Q_n(x)$ defined on X, we define a convex set

$$\Gamma = \left\{ c \,\middle|\, c = (p_1, p_2, \ldots, p_m, q_1, q_2, \ldots, q_n) \in E^{m+n}, \sum_{j=1}^{n} q_j Q_j(x) > 0 \text{ on } X \right\}.$$

Then, for a given function y(x) on X, let

$$\theta(x,c) = \left| y(x) - \sum_{i=1}^{m} p_i P_i(x) / \sum_{j=1}^{n} q_j Q_j(x) \right|$$

be defined on $X \times \Gamma$. (Notice that we can normalize the rational approximation function while preserving the convexity of $\Gamma$; for example, put $q_1 = 1$).

Our principal result concerns the general minimax rational approximation problem:

$$\min_{c \in \Gamma} \max_{x \in X} \theta(x,c). \tag{1}$$

Problem (1) includes multivariate approximation on a discrete set or on a continuum. It also includes weighted approximation using any positive weight function w(x); for we can assume in the above definition of $\theta(x,c)$ that w(x) has been absorbed within the absolute-value signs.

In addition, we have a negative result concerning best rational approximation in other norms. It is sufficient to consider the cases, in $E^1$, where $X = \{x_1, x_2, \ldots, x_N\}$ or $X = [a, b]$. For $0 < p < \infty$, the $\ell_p$ rational approximation problem is:

$$\min_{c \in \Gamma} \sum_{t=1}^{N} [\theta(x_t, c)]^p, \tag{2}$$

and the $L_p$ rational approximation problem is:

$$\min_{c \in \Gamma} \int_a^b [\theta(x, c)]^p \, dx. \tag{3}$$

3. Principal result.

Using the notation of the previous section, the following result is established for problem (1).

Theorem 8: $\psi(c) = \max_{x \in X} \theta(x, c)$ is a strictly quasi-convex function on $\Gamma$.

Proof It is convenient to put $f(x, c) = y(x) \sum_{j=1}^{n} q_j Q_j(x) - \sum_{i=1}^{m} p_i P_i(x)$ and $g(x, c) = \sum_{j=1}^{n} q_j Q_j(x)$, and to define $\phi(f(x, c), g(x, c)) = \dfrac{|f(x, c)|}{g(x, c)}$. Then if $\varepsilon(x) = \{f(x, c) \mid c \in \Gamma\}$ and $\varepsilon^+(x) = \{g(x, c) \mid c \in \Gamma\}$, Theorem 7 asserts that for any $x \in X$ the function $\phi(f(x, c), g(x, c))$ is strictly quasi-convex in $f(x,c)$ and $g(x,c)$ on the convex set $\varepsilon(x) \times \varepsilon^+(x) \subset E^1 \times E^1_+$.

Since both $f(x, c)$ and $g(x, c)$ depend linearly upon $c$, we can show easily that $\theta(x,c) = \phi(f(x, c), g(x, c))$ is a strictly quasi-convex funtion in $c$ on $\Gamma$, for every $x \in X$.

For $\theta(x, c_1) < \theta(x, c_2) \Longrightarrow \phi(f(x, c_1), g(x, c_1)) < \phi(f(x, c_2), g(x, c_2))$

$$\Longrightarrow \phi\left(\frac{f(x, c_1) + f(x, c_2)}{2}, \frac{g(x, c_1) + g(x, c_2)}{2}\right) < \phi(f(x, c_2), g(x, c_2))$$

$$\Longrightarrow \phi\left(f\left(x, \frac{c_1 + c_2}{2}\right), g\left(x, \frac{c_1 + c_2}{2}\right)\right) < \phi(f(x, c_2), g(x, c_2))$$

$$\Longrightarrow \theta\left(x, \frac{c_1 + c_2}{2}\right) < \theta(x, c_2) .$$

Finally, for $c \in \Gamma$, the continuity of $\theta(x, c)$ on $X$ follows from the continuity of $y(x)$, $P_i(x)$, and $Q_j(x)$. Thus Theorem 6 can be applied, and the result is proved.

Corollary: Any minimax rational approximation problem defines a strictly quasi-convex function with the property that a best approximation (if one exists) is a minimum of that function.

However, problems (2) and (3) do not in general lead to the minimization of a strictly quasi-convex function. For any $x \in X$ the function $[\theta(x, c)]^p$ is also strictly quasi-convex on $\Gamma$, but the remark following the proof of Theorem 6 is relevant here. For rational approximation other than in the minimax sense, local optima may exist in the interior of $\Gamma$ which are not best approximations. The following example illustrates this possibility.

Example: For $X = \{0, 0.1, 0.2, \ldots, 1.0\}$, put

$$\theta(x, c) = \left|(x^3 - 0.5) - \frac{p_1 + p_2 x + p_3 x^2}{1 + q_2 x}\right|$$

and, for $c = (p_1, p_2, p_3, 1, q_2)$, define $\psi_p(c) = \sum_{x \in X} [\theta(x, c)]^p$. Then, for $p = 1.8$, 1.9, 2.0, and 2.05, and

$$c_1 = (.88203, -.46447, .52825, 1, .22910),$$

$$c_2 = (.59354, .84698, .54733, 1, .96951),$$

we have

$$\psi_p(c_1) \neq \psi_p(c_2) \text{ and yet } \psi_p(\frac{c_1+c_2}{2}) > \max\{\psi_p(c_1), \psi_p(c_2)\} .$$

4. Remarks.

As yet, we have not experimented numerically with this direct approach to minimax rational approximation. (In Barrodale, Powell, and Roberts [1] an algorithm is provided which calculates a minimax rational approximation on a discrete set (if a best approximation exists), and its rate of convergence is quadratic under certain conditions).

The numerical example of the previous section shows that, for rational approximation, the minimax criterion sometimes leads to an easier minimization problem than do other popular norms. Nevertheless, for $p = 1$ and $p = 2$, this example is one of only six such cases where $\psi(\frac{c_1+c_2}{2}) > \max\{\psi(c_1), \psi(c_2)\}$ out of 45,000 randomly generated pairs $c_1$, $c_2$ used in test problems of various sizes and degrees. (This observation is consistent with the reported behavior of the algorithm described in Barrodale, Roberts, and Hunt [2]).

Acknowledgement. It is a pleasure to acknowledge the programming skills of Mr. Fred Crary in connection with this work.

## REFERENCES


1. I. Barrodale, M. J. D. Powell, and F. D. K. Roberts, "The differential correction algorithm for rational $\ell_\infty$ approximation", Univ. of Victoria Report #54 (1971).

2. I. Barrodale, F. D. K. Roberts, and C. R. Hunt, "Computing best $\ell_p$ approximations by functions nonlinear in one parameter", The Computer Journal, Vol. 13, 382-386 (1970).

3. O. L. Mangasarian, Nonlinear Programming, McGraw-Hill (1969).

4. O. L. Mangasarian, "Nonlinear fractional programming", Journal of the Operations Research Society of Japan, Vol. 12, 1-10 (1969).

5. O. L. Mangasarian, "Convexity, pseudo-convexity and quasi-convexity of composite functions," Cahiers du Centre d'Etudes de Recherche Operationnelle, Vol. 12, 114-122 (1970).