Institute for Operations Research and the Management Sciences

Some Bounds for the Banzhaf Index and Other Semivalues
Author(s): Ron Holzman, Ehud Lehrer, Nathan Linial
Source: *Mathematics of Operations Research*, Vol. 13, No. 2 (May, 1988), pp. 358-363
Published by: INFORMS
Stable URL: http://www.jstor.org/stable/3689830
Accessed: 22/08/2011 05:13

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at http://www.jstor.org/page/info/about/policies/terms.jsp

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.

INFORMS is collaborating with JSTOR to digitize, preserve and extend access to *Mathematics of Operations Research.*

http://www.jstor.org

MATHEMATICS OF OPERATIONS RESEARCH
Vol. 13, No. 2, May 1988
Printed in U.S.A.

SOME BOUNDS FOR THE BANZHAF INDEX AND OTHER SEMIVALUES*

RON HOLZMAN,† EHUD LEHRER‡ and NATHAN LINIAL§

The normalized Banzhaf index of a player in a monotone simple game (thought of as a voting model) is the probability for that player to swing the outcome of the vote. We bound the Euclidean norm of the vector of Banzhaf indices of simple games in terms of the number winning coalitions. The Banzhaf index is a semivalue, so we proceed to estimate norms of general semivalues.

1. Introduction. The Banzhaf index of a player in a monotone simple game counts the number of coalitions that are losing but become winning when that player joins them. When normalized appropriately, it yields the probability β_i' that player i swings the outcome in a voting model associated with the simple game v (for a detailed presentation, see Dubey and Shapley [DS], whose notation we follow).

The vector $\beta' = (\beta_i')$ of Banzhaf indices of the game is related to some interesting properties of the game. For example, its l_1 norm, $\|\beta'\|_1 = \Sigma\beta_i'$ is not necessarily 1 and its value reflects the game's responsiveness to individuals' views. A good deal of attention has been paid to $\|\beta'\|_1$ (see [H] and [DS]). For example it may be shown that among simple n-player games, $\|\beta'\|_1$ takes its maximum for the majority game where it is about $\sqrt{2n/\pi}$.

Recently interest arose in computer science in finding methods for collective coin flipping by n processors which take part in a distributed randomized computation. It is the goal to generate coin flips which are as unbiased as possible despite possible malfunctioning of some of the participating processors. This problem turns out to be equivalent to the quest of simple games with small l_∞ norm for β', $\|\beta'\|_\infty = \max \beta_i'$. Examples were given of games where exactly half of the coalitions win and $\|\beta'\|_\infty = O(\log n/n)$ (see [BL] for more on this subject).

The present article is concerned with the l_2 norm $\|\beta'\|_2 = \sqrt{\Sigma(\beta_i')^2}$. In the study of spin glass models of neural networks [AGS] one is interested in the operator associating with a vector $\mathbf{m} \in R^n$ the vector $2^{-n}\Sigma_{x \in \{-1,+1\}^n} x \cdot \operatorname{sgn}(\mathbf{m} \cdot x)$. It was conjectured that the range of this map is contained in the Euclidean n-dimensional unit ball. A. Neyman observed that this question is equivalent to asking whether $\|\beta'\|_2 \leqslant 1$ for weighted majority games. Our Corollary 1 shows that this is true for all simple games. Before we proceed to survey our results, let us remark that notions equivalent to the Banzhaf power index come up in a variety of other areas, such as threshold logic in circuit design [W] and extremal problems on hypergraphs (see [F]).

*Received December 6, 1985; revised January 5, 1987.
AMS 1980 subject classification. Primary: 90D12.
IAOR 1973 subject classification. Main: Games.
OR/MS Index 1978 subject classification. Primary: 234 Games: Cooperative.
Key words. Games, Coalitions, Banzhaf indices, Bounds for.
† Rutgers University.
‡ The Hebrew University.
§ IBM Almaden Research Center and The Hebrew University.

0364-765X/88/1302/0358$01.25
Copyright © 1988, The Institute of Management Sciences/Operations Research Society of America

Our first result (Theorem 1) states that if a simple game has ω winning coalitions and $\lambda = 2^n - \omega$ losing ones, then $\Sigma(\beta_i')^2 \leqslant 2^{-n+1}\min(\omega, \lambda)$. In particular $\|\beta'\|_2^2 \leqslant 1$ for all simple games with equality holding only for dictatorial games (Corollary 1).

The failure of the Banzhaf indices to add up to 1, which was mentioned above, is the main difference between this notion and the well known Shapley value. In fact, the operator corresponding to v the vector (β_i'), $i = 1, \ldots, n$ is a semivalue. It satisfies all of Shapley's axioms except efficiency. The class of all semivalues has been characterized in [DNW]. In §3 we seek to extend our bound on the Banzhaf index to other semivalues. This involves a bound on the Euclidean norm of the gradient of a multilinear function defined on an n-dimensional cube evaluated along the main diagonal. We also obtain (Theorem 3) a lower bound on the l_1 norm of the gradient (of a multilinear extension of a monotonic simple game), generalizing a similar bound given in [DS] for the Banzhaf case. Here the extreme elements (attaining the bound) are all unanimity rules.

We think that the techniques employed in the proofs merit interest. The one used in §3 differs from the one used in §2, and neither of them is quite straightforward.

2. The Banzhaf index. A game (N, v) consists of a set of *players*, $N = \{1, 2, \ldots, n\}$ and a function $v\colon 2^N \to R$ satisfying $v(\varnothing) = 0$. A *coalition* is a subset of N. A game (N, v) is *simple* if the range of v is $\{0, 1\}$ and $v(N) = 1$. The coalitions with $v(S) = 1$ are called *winning*, those with $v(S) = 0$ are called *losing*. A game is called *monotone* if $S \subset T$ implies $v(S) \leqslant v(T)$. Define η_i to be $\Sigma_{S \in 2^N} v(S \cup i) - v(S)$. Equivalently, $\eta_i = \Sigma_{S \in 2^N} \epsilon_i(S) v(S)$, where $\epsilon_i(S)$ equals 1 if $i \in S$ and -1 otherwise.

Let v be a monotone game. For a fixed i, there are 2^{n-1} pairs of the form $(S, S \cup \{i\})$ with $i \notin S$. Therefore, if one such pair is picked at random, the probability of a swing (i.e. $v(S) + v(S \cup \{i\}) = 1$) is $\beta_i' = \eta_i/2^{n-1}$. The β_i' are the *normalized Banzhaf indices*.

We recall now a few notions about simple games. A simple game v is a *T-unanimity game* if T is a nonempty coalition and $v(S) = 1$ iff $S \supset T$. v is *dictatorial* if it is an $\{i\}$-unanimity game for some $i \in N$. The *dual* of a game v is the game v^* defined by $v^*(S) = v(N) - v(N - S)$. Note that a dual of a simple game is a simple game.

THEOREM 1. *Let v be a simple game with n players, ω winning coalitions and λ losing coalitions (thus, $\omega + \lambda = 2^n$). Denote $\omega' = \omega/2^n$, $\lambda' = \lambda/2^n$. Then:*

$$\sum_{i=1}^{n} (\beta_i')^2 \leqslant 2\min\{\omega', \lambda'\}. \tag{1}$$

Equality holds in (1) *iff v is dictatorial or a T-unanimity game with $|T| = 2$ or the dual of the latter.*

PROOF. We consider V, the 2^n-dimensional Euclidean space, with the standard scalar product $(\cdot,\cdot)$ and norm $\|\cdot\|$ derived from it. We index the coordinates by coalitions (subsets of the set N of players), and write the elements of the space as functions from 2^N into R.

Given the game v, the vector of its Banzhaf indices β' equals $2^{-n+1}Av$, where A is an n by 2^n matrix whose rows are the vectors ϵ_i $(i = 1, \ldots, n)$ defined above. Let $u \in V$ be given by $u(S) = v(N - S)$. Then it is easily verified that $Au = -2^{n-1}\beta'$. Therefore if $f = v - u$, then $Af = 2^n\beta'$ and so

$$\|\beta'\|^2 = 2^{-2n}\|Af\|^2 = 2^{-2n}(f, A^TAf) \leqslant 2^{-2n}\Lambda\|f\|^2 \tag{2}$$

where Λ is the largest eigenvalue of A^TA. Now

$$\|f\|^2 = \sum_{S \in 2^N} [v(S) - v(N-S)]^2$$

$$= |\{S | v(S) = 1, v(N-S) = 0\}| + |\{S | v(S) = 0, v(N-S) = 1\}|.$$

Neither of the summands can exceed ω or λ, hence

$$\|f\|^2 \leqslant 2\min\{\omega, \lambda\}. \tag{3}$$

As for Λ, the rank of A^TA cannot exceed n and $\epsilon_1, \dots, \epsilon_n$ are easily seen to be eigenvectors of this matrix with eigenvalue 2^n. Therefore all the other eigenvalues are zero and so $\Lambda = 2^n$. The conclusion (1) now follows.

For equality to hold in (1), both (2) and (3) must be satisfied with equality. Let us assume that $\min\{\omega, \lambda\} = \omega$. (If this is not the case, we may argue about the dual game v^*.) Then equality in (3) means that for all $S \in 2^N$

$$(S) = 1 \Rightarrow v(N-S) = 0. \tag{4}$$

(2) holds with equality iff f belongs to the span of $\epsilon_1, \dots, \epsilon_n$. But since the ϵ_i are pairwise orthogonal and with the same norm $2^{n/2}$ that means that $f = \sum_{i=1}^n (f, \epsilon_i)\epsilon_i/2^n$. This means that for all $S \in 2^N$

$$v(S) - v(N-S) = \sum_{i \in S} \beta_i' - \sum_{i \notin S} \beta_i'. \tag{5}$$

Let us use (5) for $S = N$. This implies that $\sum_1^n \beta_i' = 1$. Now for every $S \subseteq N$

$$\left(\sum_{i \in S} \beta_i' - \sum_{i \notin S} \beta_i'\right) \in \{-1, 0, 1\} \tag{6}$$

and it follows that all $\beta_i' \geqslant 0$. Now use (6) for $S = N \setminus \{j\}$ where $\beta_j' > 0$. Since we obtain either 0 or -1 it follows that β_j' is either $\frac{1}{2}$ or 1. There are only two possibilities for β' up to change in coordinates:

(i) $\beta' = (1, 0, \dots, 0)$ which corresponds only to the dictatorial game, and

(ii) $\beta' = (\frac{1}{2}, \frac{1}{2}, 0, \dots,)$ corresponding (by the following) only to the $T = \{1, 2\}$ unanimity game.

By (5) if $S \supseteq \{1,2\}$ then $v(S) = 1$ and if $S \cap \{1,2\} = \varnothing$, then $v(S) = 0$. If $S \cap \{1,2\} = \{1\}$, then $v(S) = v(N-S)$. But then (4) implies that $v(S) = v(N-S) = 0$ holds. This shows that the game is the $\{1,2\}$-unanimity game and completes the proof of our theorem. ■

Corollary 1 (I. Ninomiya, see [W]). *For every simple game $\sum_{i=1}^n (\beta_i')^2 \leqslant 1$. This bound is attained iff the game is dictatorial.*

The proof of this result as presented in [W] is based on Bessel's inequality applied to V with the ϵ_i as a basis. Their method does not seem to imply our Theorem 1.

We do not know how good the bound given in (1) is. That is, given $0 \leqslant \omega' \leqslant 1$, how large can $\sum(\beta')^2$ be? We have a piecewise linear upper bound, namely, $2\min(\omega', 1 - \omega')$. We know that this bound is tight for $\omega' = 0$, $1/4$, $1/2$, $3/4$ and 1. But the supremum of $\|\beta'\|^2$ over all simple games with a given ω' is unknown for other values of ω'.

A related, more refined question is to find $\Theta(n, \omega)$, the maximum of $\sum_1^n (\beta_i')^2$ over all n-player simple games with ω winning coalitions. This question seems hard and we only have the following remarks to make about it, which we state without proof:

PROPOSITION 1. (i) *For* $\omega < 2^{n-1}$, *there holds* $\Theta(n, \omega + 1) > \Theta(n, \omega)$.

(ii) *If* G *is a simple n-player game with* $\omega \leqslant 2^{n-1}$ *winning coalitions and with* $\Sigma(\beta_i')^2 = \Theta(n, \omega)$, *then* G *is a weighted majority game and the* β_i' *may be taken to be the weights.* ∎

3. Other semivalues. If g is the multilinear extention of a game (see [O]) a general semivalue is obtained by integrating its gradient along the main diagonal with respect to a certain probability measure (see [DNW]). In particular, the normalized Banzhaf index for simple games is obtained using Dirac measure at $\frac{1}{2}$. This provides a generalization of the Banzhaf index to nonsimple games.

Denote $C^n = [0,1]^n$ and its set of vertices by V^n. For $0 \leqslant t \leqslant 1$ we denote the vector $(t, \ldots, t)$ by $\underline{t}$. By slightly modifying the proof of Theorem 1 we get:

THEOREM 1′. *For every multilinear function* $g\colon C^n \to [0,1]$

$$\left\|\nabla g(\underline{1/2})\right\|_2^2 \leqslant 2\min\left\{g(\underline{1/2}), 1 - g(\underline{1/2})\right\}. \quad ∎$$

COROLLARY 2. *For every multilinear function* $g\colon C^n \to [0,1]$ *and every* $0 < t < 1$

$$\|\nabla g(\underline{t})\|_2^2 \leqslant \frac{\min\{g(\underline{t}), 1 - g(\underline{t})\}}{2[\min\{t, 1 - t\}]^2}. \tag{7}$$

The proof follows by applying Theorem 1′ to $f(x_1, \ldots, x_n) = g(2tx_1, \ldots, 2tx_n)$ for $t \leqslant \frac{1}{2}$, and by using this result for $h(x_1, \ldots, x_n) = g(1 - x_1, \ldots, 1 - x_n)$. ∎

The bound in (7) deteriorates as t approaches 0 or 1. We establish now another bound that does not suffer from this shortcoming.

THEOREM 2. *For every multilinear function* $g\colon C^n \to [0,1]$ *and every* $0 < t < 1$

$$\|\nabla g(\underline{t})\|_2^2 < \frac{g(\underline{t})(1 - g(\underline{t}))}{t(1 - t)}. \tag{8}$$

Equality holds in (8) *iff* g *is of one of the following forms*: (i) $g(x_1, \ldots, x_n) = x_i$, (ii) $g(x_1, \ldots, x_n) = 1 - x_i$, (iii) $g = 0$ *or* (iv) $g = 1$.

PROOF. We fix t throughout the proof, and proceed by induction on n. For $n = 1$, $g'(t) = (g(t) - g(0))/t = (g(1) - g(t))/(1 - t)$, hence

$$[g'(t)]^2 = \frac{[g(t) - g(0)][g(1) - g(t)]}{t(1 - t)} \leqslant \frac{g(t)(1 - g(t))}{t(1 - t)}.$$

Suppose now that $g\colon C^{n+1} \to [0,1]$ and (8) is true in the n-dimensional case. We define two multilinear functions g^0 and g^1 from C^n into $[0,1]$ by:

$$g^i(x_1, \ldots, x_n) = g(x_1, \ldots, x_n, i), \qquad i = 0, 1,$$

and get the following (we suppress $\underline{t}$ from our formulae; it will be understood that all functions are evaluated at that point with the appropriate dimension):

$$g = (1-t)g^0 + tg^1,$$

$$\frac{\partial g}{\partial x_i} = (1-t)\frac{\partial g^0}{\partial x_i} + t\frac{\partial g^1}{\partial x_i}, \qquad i = 1,\dots,n,$$

$$\frac{\partial g}{\partial x_{n+1}} = g^1 - g^0.$$

Using these relations, the convexity of $\|\cdot\|_2^2$ and the induction hypothesis for g^0 and g^1, we have:

$$\begin{aligned}\|\nabla g\|_2^2 &= \|(1-t)\nabla g^0 + t\nabla g^1\|_2^2 + (g^1 - g^0)^2 \\ &\leqslant (1-t)\|\nabla g^0\|_2^2 + t\|\nabla g^1\|_2^2 + (g^1-g^0)^2 \\ &\leqslant (1-t)\frac{g^0(1-g^0)}{t(1-t)} + t\frac{g^1(1-g^1)}{t(1-t)} + (g^1-g^0)^2 = \frac{g(t)(1-g(t))}{t(1-t)}.\end{aligned}$$

This establishes (8) in the $(n+1)$-dimensional case. The proof of the statement about equality is by following the inductive process and is left to the reader. ■

COROLLARY 3. *For every multilinear function* $g\colon C^n \to [0,1]$ *and every* $0 < t < 1$

$$\|\nabla g(\underline{t})\|_2^2 \leqslant \frac{1}{4t(1-t)}. \tag{9}$$

Equality holds in (9) *iff* $t = \frac{1}{2}$ *and* g *is of the form* $g(x_1,\dots,x_n) = x_i$ *or* $g = (x_1,\dots,x_n) = 1 - x_i$.

Note that neither of the bounds (7) and (8) dominates the other. It would be interesting if one could provide a nice bound that combines the advantages of the two.

We turn to a lower bound on the l_1 norm of the gradient. This result, which is another byproduct of the inductive technique, requires monotonicity. Its proof is along the lines of the proof for Theorem 2 and is omitted.

THEOREM 3. *Let g be a multilinear function on C^n satisfying*: (i) $g(V^n) \subseteq \{0,1\}$, *and* (ii) $\underline{x} \leqslant \underline{y} \to g(\underline{x}) \leqslant g(\underline{y})$, $\forall \underline{x}, \underline{y} \in V^n$. *Then for every* $0 < t < 1$

$$\|\nabla g(\underline{t})\|_1 \geqslant \frac{g(\underline{t})\ln g(\underline{t})}{t\ln t}. \tag{10}$$

(*If* $g(\underline{t}) = 0$ *we use the convention* $0\ln 0 = 0$.) *Equality holds in* (10) *iff there exists a nonempty* $T \subseteq N$ *such that* $g(x_1,\dots,x_n) = \prod_{i\in T} x_i$, *or* $g = 0$ *or* $g = 1$.

If we consider (10) for $t = \frac{1}{2}$, we may substitute $\omega' = \omega/2^n$ for $g(\underline{1/2})$ and $\beta_i' = \eta_i/2^{n-1}$ for $(\partial g/\partial x_i)(\underline{\tfrac{1}{2}})$, thus obtaining $\sum_{i=1}^n \eta_i \geqslant \omega(n - \log_2\omega)$. (We may also state this inequality with λ, instead of ω.) By applying (10) to the dual function g^* (of the dual game) one can get also

$$\|\nabla g(t)\|_1 \geqslant \frac{(1-g(\underline{t}))\ln(1-g(\underline{t}))}{(1-t)\ln(1-t)}.$$

This is a lower bound on the total number of swings in a simple game with n players and w winning coalitions. A similar result was obtained by Dubey and Shapley [DS, p.

108], based on the results of Hart [H]. Our elementary proof thus answers negatively their question whether use of [H] was necessary.

Acknowledgements. We are much indebted to A. Neyman, who triggered off this research by conjecturing Corollary 1 and challenging us to prove it. We are also grateful to J. F. Mertens and R. Meshulam for their helpful ideas.

References

[AGS] Amit, D. J., Gutfreund, H. and Sompolinsky H. (1985). Spin-Glass Models of Neural Networks. Preprint.

[BL] Ben-Or, M. and Linial, N. (1985). Collective Coin Flipping, Robust Voting Schemes and Minima of Banzhaf Values. *26th Conf. Foundations of Computer Science* 408–416.

[DNW] Dubey, P., Neyman, A. and Weber, R. J. (1981). Value Theory without Efficiency. *Math. Oper. Res.* **6** 122–128.

[DS] ______ and Shapley, L. S. (1979). Mathematical Properties of the Banzhaf Power Index. *Math. Oper. Res.* **4** 99–131.

[F] Frankl, P. (1983). On Traces of Sets, *J. Combin. Theory Ser. A* **34** 41–45.

[H] Hart, S. (1976). A Note on the Edges of the *n*-Cube. *Discrete Math.* **14** 157–163.

[O] Owen, G. (1972). Multilinear Extensions of Games. *Management Sci.* **18** 64–79.

[W] Winder, R. O. (1971). Chow Parameters in Threshold Logic. *J. Assoc. Comput. Mach.* **18** 265–289.

HOLZMAN: RUTGERS UNIVERSITY, NEW BRUNSWICK, NEW JERSEY 08903

LEHRER: INSTITUTE OF MATHEMATICS AND COMPUTER SCIENCE, THE HEBREW UNIVERSITY, JERUSALEM, 91904, ISRAEL

LINIAL: DEPARTMENT OF COMPUTER SCIENCE, THE HEBREW UNIVERSITY, JERUSALEM 91904, ISRAEL and IBM ALMADEN RESEARCH CENTER, 650 HARRY ROAD, SAN JOSE, CALIFORNIA 95120-6099