ISRAEL JOURNAL OF MATHEMATICS 77 (1992), 55–64

THE INFLUENCE OF VARIABLES IN PRODUCT SPACES

BY

JEAN BOURGAIN

Département de Mathématiques, IHES
35 route de Chartres, Bures-sur-Yvette, France

AND

JEFF KAHN

Department of Mathematics, Rutgers University
New Brunswick, NJ 08903, USA

AND

GIL KALAI

Landau Center for Research in Mathematical Analysis
Institute of Mathematics and Computer Science
The Hebrew University of Jerusalem, Jerusalem 91904, Israel;
and IBM Almaden Research Center

AND

YITZHAK KATZNELSON

Department of Mathematics, Stanford University
Stanford, CA 94305, USA

AND

NATHAN LINIAL*

Institute of Mathematics and Computer Science
The Hebrew University of Jerusalem, Jerusalem 91904, Israel

* Work supported in part by grants from the Binational Israel–US Science Foundation and the Israeli Academy of Science.
Received November 16, 1990 and in revised form December 5, 1991

ABSTRACT

Let X be a probability space and let $f\colon X^n \to \{0,1\}$ be a measurable map. Define the **influence of the k-th variable** on f, denoted by $I_f(k)$, as follows: For $u = (u_1, u_2, \ldots, u_{n-1}) \in X^{n-1}$ consider the set $l_k(u) = \{(u_1, u_2, \ldots, u_{k-1}, t, u_k, \ldots, u_{n-1})\colon t \in X\}$.

$$I_f(k) = \Pr(u \in X^{n-1} : f \text{ is not constant on } l_k(u)).$$

More generally, for S a subset of $[n] = \{1, \ldots, n\}$ let the influence of S on f, denoted by $I_f(S)$, be the probability that assigning values to the variables not in S at random, the value of f is undetermined.

THEOREM 1: *There is an absolute constant c_1 so that for every function $f\colon X^n \to \{0,1\}$, with $\Pr(f^{-1}(1)) = p \le \frac{1}{2}$, there is a variable k so that*

$$I_f(k) \ge c_1 p \frac{\log n}{n}.$$

THEOREM 2: *For every $f\colon X^n \to \{0,1\}$, with $\mathrm{Prob}(f = 1) = \frac{1}{2}$, and every $\epsilon > 0$, there is $S \subset [n], |S| = c_2(\epsilon)n/\log n$ so that $I_f(S) \ge 1 - \epsilon$.*

These extend previous results by Kahn, Kalai and Linial for Boolean functions, i.e., the case $X = \{0,1\}$.

1. Introduction

Let X be a probability space and let $f\colon X^n \to \{0,1\}$ be a measurable map. Define the **influence of the k-th variable** on f, denoted by $I_f(k)$, as follows: For $u = (u_1, u_2, \ldots, u_{n-1}) \in X^{n-1}$ consider the set

$$l_k(u) = \{(u_1, u_2, \ldots, u_{k-1}, t, u_k, \ldots, u_{n-1}) : t \in X\}.$$

$$I_f(k) = Pr(u \in X^{n-1} : f \text{ is not constant on } l_k(u)). \tag{1}$$

More generally, for S a subset of $[n] = \{1, \ldots, n\}$ let the influence of S on f, denoted by $I_f(S)$, be the probability that assigning values to the variables not in S at random, the value of f is undetermined. (Note that $I_f(\{k\}) = I_f(k)$.)

The purpose of this note is to supplement the papers by Kahn, Kalai and Linial [KKL, KKL′], which study the influence of variables on Boolean functions, i.e., the case $X = \{0,1\}$. The reader is referred to [BL, KKL, KKL′] for background on this problem and its relevance to extremal combinatorics and theoretical computer science.

Given X and f as above we can replace X by the unit interval $[0,1]$, and f by an appropriate function g so that the influences of f and g will be the same. Therefore, there will be no loss of generality in assuming that $X = [0,1]$.

An easy consequence of Loomis and Whitney's inequality [LW] is:

THEOREM 0: *Every function* $f\colon X^n \to \{0,1\}$ *with* $\Pr(f=1) = p \le \frac{1}{2}$ *satisfies*

$$\sum_{k=1}^{n} I_f(k) \ge p\log(\frac{1}{p}). \tag{2}$$

The following examples show that for $p > (\frac{1}{2})^n$ this inequality is sharp (up to a constant factor): If $(\frac{1}{2})^{k-1} \ge p > (\frac{1}{2})^k$ let $f = 1$ iff $p^{1/k} \ge x_i$, for $i = 1, \ldots, k$.

Theorem 0 implies that for some variable k,

$$I_f(k) \ge p\log(\frac{1}{p})\frac{1}{n}.$$

Here we improve this estimate to

THEOREM 1: *There is an absolute constant* c_1 *so that for every function*

$$f\colon X^n \to \{0,1\},$$

with $\Pr(f=1) = p \le \frac{1}{2}$, *there is a variable* k *so that*

$$I_f(k) \ge c_1 p \frac{\log n}{n}.$$

Repeated applications of Theorem 1 yields:

THEOREM 2: *For every* $f\colon X^n \to \{0,1\}$, *with* $\mathrm{Prob}(f=1) = \frac{1}{2}$, *and every* $\epsilon > 0$, *there is* $S \subset [n]$, $|S| = c_2(\epsilon)n/\log n$ *so that* $I_f(S) \ge 1-\epsilon$.

The assertions of Theorems 1 and 2 for Boolean functions (i.e., for the special case $X = \{0,1\}$) are proved in [KKL,KKL′], in response to a conjecture by Ben-Or and Linial [BL]. That Theorems 1 and 2 are asymptotically optimal for $p = \frac{1}{2}$ and $X = \{0,1\}$ is shown by the "tribes" function f from [BL]. Here, and throughout the paper, we identify elements of $\{0,1\}^n$ with subsets S of $[n]$ in the usual way. Partition $[n]$ into subsets $S_1, \ldots, S_k$ of size $\log n - \log\log n + c$

(c is an appropriate constant) and define $f(T) = 1$ iff T contains S_j for some j. Obviously, a similar function can also be realized for $X = [0,1]$.

An example which exists only in the latter case but not for $X = \{0,1\}$ is the function f which equals 1 iff $x_i \le p^{1/n}$ for every i, $1 \le i \le n$. It shows the Loomis–Whitney inequality to be tight for any $p > 0$ and also shows why the proof in [KKL, KKL′] needs to be modified to handle general probability spaces X.

2. Proofs

The proof of [KKL] relies on Beckner's hypercontractive estimate. In order to extend it to our more general case we need some additional considerations. We also sketch a variant of the proof based on another hypercontractive estimate. For simplicity we prove Theorem 1 for $p = \frac{1}{2}$, leaving the minor adjustment needed for general p to the reader.

LEMMA 1: *Given a function g: $[0,1]^n \to \{0,1\}$, there is a monotone function f: $[0,1]^n \to \{0,1\}$ such that $I_g(k) \ge I_f(k)$ for every k.*

Proof: Consider the restriction of f to the unit segment $l_k(u)$. Define $T_k(f)$ as the function which is monotone on $l_k(u)$ and satisfies $\Pr(T_k(f)^{-1}(0) \cap l_k(u)) = \Pr(f^{-1}(0) \cap l_k(u))$ for every $u \in X^{n-1}$. Note that $I_f(k) = I_{T_k(f)}(k)$ and $I_f(j) \ge I_{T_k(f)}(j)$ for $j \ne k$. Repeated applications of these operations yields in a limit a function which is fixed under all T_k, hence monotone. ∎

Remark 1: The proof of Lemma 1 is a standard combinatorial shifting argument, (see [A, Bo, F, BL]) and is also similar to the well-known Steiner symmetrization.

Remark 2: The same argument implies that $I_g(S) \ge I_f(S)$ for every S.

At this point we replace $X = [0,1]$ by the interval of integers

$$Y = \{0, 1, \dots, 2^m - 1\}$$

(with uniform probability distribution). It suffices to prove Theorem 1 with Y instead of X as long as our constants do not depend on m. It will be useful to identify Y with the discrete m-dimensional cube $\{0,1\}^m$ by the binary expansion. This allows one to express functions f: $Y \to \mathbf{R}$ in their Walsh–Fourier expansion

$$f = \sum \{\hat{f}(S) u_S : S \subset [m]\}, \tag{3}$$

where u_S is the function defined by $u_S(T) = (-1)^{|S\cap T|}$.

For a function $f\colon Y^n \to \mathbf{R}$, we write the Walsh–Fourier expansion of f in the following form:

$$f = \sum\{\hat{f}(S_1,\dots,S_n)u_{S_1,\dots,S_n} \mid S_1 \subset [m],\dots,S_n \subset [m]\}. \tag{4}$$

Here $u_{S_1,\dots S_n}(T_1,\cdots,T_n) = \prod u_{S_i}(T_i)$.

We always view Y as a probability space, and so given a function $f\colon Y \to \mathbf{R}$, its p-th norm is defined as

$$\|f\|_p = (\frac{1}{|Y|}\sum_{S\subseteq Y} |f(S)|^p)^{1/p}.$$

Parseval's identity asserts that $\|f\|_2^2 = \sum_{S\subseteq Y} \hat{f}^2(S)$. We also define

$$w(f) = \sum_{S\subset[m]} \hat{f}^2(S)|S|. \tag{5}$$

Clearly, $w(f) \geq 0$ for every function f and $w(f) = 0$ if and only if f is a constant function.

LEMMA 2: ([KKL, CG]) *For* $f\colon \{0,1\}^m \to \{0,1\}$

$$w(f) = \sum_{k=1}^{m} I_f(k). \tag{6}$$

A function f from Y to $\{0,1\}$ is monotone iff for some t, $f(i) = 0$ when $0 \leq i \leq t$ and $f(i) = 1$ when $t < i \leq 2^m - 1$. This has some implications on f's Walsh transform.

LEMMA 3: *Let* $f\colon Y \to \{0,1\}$ *be a monotone function. Then* $w(f) \leq 2$.

Proof: By definition $I_f(k)$ is 2^{-m+1} times the number of pairs v, w with $f(v) = 0, f(w) = 1$ so that v is obtained from w by flipping the k-th coordinate. (Note: here, $I_f(k)$ is the influence of a function from Y to $\{0,1\}$ regarded as a Boolean function of m variables.)

The monotonicity of f implies that

$$I_f(1) \leq \frac{1}{2^{(m-1)}}, \quad I_f(2) \leq \frac{1}{2^{(m-2)}}, \cdots, I_f(m) \leq 1$$

(in fact,

$$I_f(k) = \frac{1}{2^{(m-k)}}$$

unless $t < 2^{k-1}$ or $t > 2^m - 2^{k-1}$). Therefore $\sum_{k=1}^{m} I_f(k) \leq 2$, and by Lemma 2 this is what we need. ∎

LEMMA 4: ([KKL]) *For* $f\colon \{0,1\}^r \to \{0,1\}$, *define* $T_\epsilon(f) = \sum\{\hat{f}(S)\epsilon^{|S|}u_S : S \subset [r]\}$. *Then*

$$\|T_\epsilon f\|_2 \le \|f\|_{1+\epsilon^2}. \tag{7}$$

Proof: As shown in [KKL] this follows at once from Lemmas 1 and 2 in Beckner's paper [Be]. (We will need the case $r = mn$.) ∎

Remark: For our purposes $1+\epsilon^2$ can be replaced by any $2-\delta(\epsilon)$, so Beckner's Lemma 1 can be replaced here by an obvious estimate. ∎

Here is a quick outline of the proof of Theorem 1. We assume that f is monotone. Consider the restriction g of f to a function from Y to $\{0,1\}$ obtained by assigning values to all variables except the k-th one. $I_f(k)$ is the probability (assignments being selected at random) that g is not constant. The proof is based on two observations: First, that $w(g)$ is bounded between 0 and 2 with $w(g)=0$ if g is constant. The second observation is that if r is obtained by subtracting from g its average value, then r is bounded, and we can give an absolute upper bound for the (4/3)-norm of r. These two observations combined with Lemma 4 have consequences on the Walsh–Fourier coefficients of f which imply our theorem.

Proof of Theorem 1: Let $f\colon Y^n \to \{0,1\}$ be a function with $\Pr(f=1)=\frac{1}{2}$. We will show that for some k,

$$I_f(k) \ge c_1 \frac{\log n}{n}.$$

By Lemma 1 we may assume that f is monotone.

Let $T\colon Y \to \mathbf{R}$ be given by $T(Z) = \sum_S u_S(Z)|S|^{1/2}$, i.e. $\hat{T}(S) = |S|^{1/2}$ for all S. The convolution of T with a function $g\colon Y \to \mathbf{R}$ is denoted $T * g$, i.e., $\widehat{T*g}(S) = \hat{g}(S)|S|^{1/2}$ and

$$\|T * g\|_2^2 = \sum_{S\subset[m]} \hat{g}^2(S)|S| = w(g). \tag{8}$$

Fix an index $n \ge k \ge 1$, and define a function

$$g = g[S_1,\dots,S_{k-1},S_{k+1},\dots,S_n]\colon Y \to \{0,1\}$$

by

$$g[S_1,\dots,S_{k-1},S_{k+1},\dots,S_n](S) = f(S_1,S_2,\dots,S_{k-1},S,S_{k+1},\dots,S_n). \tag{9}$$

Define also a function $v = v[S_1, ..., S_{k-1}, S_{k+1}, ..., S_n]\colon Y \to \mathbf{R}$ by

$$v[S_1, S_2, ..., S_{k-1}, S_{k+1}, ..., S_n] = T * g[S_1, ..., S_{k-1}, S_{k+1}, ..., S_n]. \tag{10}$$

By equation (8), $\|v\|_2^2 = w(g)$, and by Lemma 3, $0 \le \|v\|_2^2 \le 2$. If g is a constant function then $\|v\|_2^2 = 0$.

Define now $W_k(S_1, S_2, ..., S_n) = u[S_1, ..., S_{k-1}, S_{k+1}, ..., S_n](S_k)$. W_k is the convolution of f with the real function T_k on Y^n given by $T_k(S_1, S_2, ..., S_n) = T(S_k)$ if $S_i = \emptyset$ for every $i \neq k$ and $T_k(S_1, S_2, ..., S_n) = 0$ otherwise. Note that $\hat{T}_k(S_1, S_2, ..., S_n) = |S_k|^{1/2}$ and therefore

$$\hat{W}_k(S_1, S_2, ..., S_n) = \hat{f}(S_1, S_2, ..., S_n)|S_k|^{\frac{1}{2}}, \tag{11}$$

and

$$\|W_k\|_2^2 = \sum (\hat{W}_k(S_1, S_2, ..., S_n))^2 = \sum \hat{f}^2(S_1, S_2, ..., S_n)|S_k|. \tag{12}$$

On the other hand,

$$\begin{aligned} \|W_k\|_2^2 &= |Y|^{-n} \sum_{S_1 \subset [m], ..., S_n \subset [m]} v[S_1, ..., S_{k-1}, S_{k+1}, ..., S_n]^2(S_k) \\ &= |Y|^{-n+1} \sum_{S_1, ..., S_{k-1}, S_{k+1}, ..., S_n} \|v[S_1, ..., S_{k-1}, S_{k+1}, ..., S_n]\|_2^2. \end{aligned} \tag{13}$$

But we saw that the value of $\|v[S_1, ..., S_{k-1}, S_{k+1}, ..., S_n]\|_2^2$ is non-negative, bounded by 2, and is equal to zero if $g[S_1, ..., S_{k-1}, S_{k+1}, ..., S_n]$ is the constant function.

Therefore we have

$$\|W_k\|_2^2 \le 2I_f(k). \tag{14}$$

Assume now that for every k,

$$I_f(k) \le c_1 \frac{\log n}{n}.$$

It follows that

$$\begin{aligned} \sum_{k=1}^{n} \|W_k\|_2^2 &= \sum_{S_1 \subset [m], ..., S_n \subset [m]} \hat{f}^2(S_1 \cdots S_n)(|S_1| + |S_2| + \cdots + |S_n|) \\ &\le 2c_1 \log n. \end{aligned} \tag{15}$$

Thus, more than half of the weight of $||f||_2^2$ is concentrated where $|S_1| + |S_2| + \cdots + |S_n| < 5c_1 \log n$.

To reach a contradiction write $R_k = \sum_{S_k \neq \emptyset} \hat{f}(S_1 \cdots S_n) u_{S_1,\dots S_n}$. Note that

$$\begin{aligned} R_k(S_1, \dots, S_{k-1}, S_k, S_{k+1}, \dots, S_n) =& f(S_1, \dots S_k, \dots, S_n) \\ &- E_{S_k} f(S_1, \dots, S_k, \dots, S_n). \end{aligned} \tag{16}$$

Here, $E_{S_k} f(S_1, \dots, S_k, \dots, S_n)$ is the average value of $f(S_1, \dots, S_k, \dots, S_n)$ over all values of S_k. Therefore $|R_k|$ is bounded (say by 2), and $R_k(S_1, \dots, S_n) = 0$ if $g[S_1, \dots, S_{k-1}, S_{k+1}, \dots, S_n]$ is a constant function.

It follows that

$$||R_k||_{4/3}^{4/3} \leq 3I_f(k). \tag{17}$$

I.e.,

$$||R_k||_{4/3}^2 \leq (3I_f(k))^{3/2}, \tag{18}$$

and by Lemma 4 for $\epsilon = \sqrt{3}/3$

$$\sum_{k=1}^{n} ||T_\epsilon R_k||_2^2 \leq \sum_{k=1}^{n} \|R_k\|_{4/3}^2 \leq c_3 (\log n)^{\frac{3}{2}} n^{-\frac{1}{2}}. \tag{19}$$

Note that

$$T_\epsilon R_k = \sum \hat{R}_k(S_1, \dots, S_n) \epsilon^{|S_1| + \cdots + |S_n|} u(S_1 \cdots S_n)$$

and that $\hat{R}_k(S_1, \dots, S_n) = 0$ or $\hat{f}(S_1, \dots, S_n)$, depending on S_k being empty or not. Therefore,

$$\begin{aligned} \sum ||T_\epsilon R_k||_2^2 =& \sum_{S_1 \subset [m], \dots, S_n \subset [m]} \hat{f}^2(S_1 \dots S_n) \mu(S_1 \dots S_n) \epsilon^{2|S_1| + \dots + 2|S_n|} \\ &\leq c_3 (\log n)^{\frac{3}{2}} n^{-\frac{1}{2}}, \end{aligned} \tag{20}$$

where $\mu(S_1 \cdots S_n) = |\{j \colon S_j \neq \emptyset\}|$.

The last relation implies that more than half the weight of $||f||_2^2$ is concentrated where $|S_1| + |S_2| + \cdots + |S_n| > c_4 \log n$ which is a contradiction if c_1 is sufficiently small. ∎

Alternative proof for Theorem 1 (sketch): Let us assume again that $X = [0,1]$ and that f is monotone. The (ordinary) Fourier expansion of f is:

$$f = \sum_{z \in \mathbf{Z}^n} \hat{f}(z) e^{2\pi i <z,x>}. \tag{21}$$

Define

$$\tilde{w}(f) = \sum \hat{f}^2(k)|k|^{\frac{1}{3}}. \tag{22}$$

Clearly $\tilde{w}(f)$ is non-negative and $\tilde{w}(f) = 0$ iff f is a constant function.

LEMMA 3′: *Let $f\colon X \to \{0,1\}$ be a monotone function, then $\tilde{w}(f) \leq c$ for some absolute constant c.*

Proof: Easy.

LEMMA 4′: *Define $P_a = \sum a^{t^{\frac{2}{3}}} e^{2\pi i t}$. Then for $a > 0$ small enough, and for every $g\colon [0,1] \to \mathbf{R}$, $||P_a * g||_2 \leq ||g||_{4/3}$.*

Proof: This follows easily by the Riesz interpolation theorem by showing that for $a > 0$ sufficiently small, $||P_a * g||_\infty \leq ||g||_2$ and $||P_a * g||_1 \leq ||g||_1$.

The proof of Theorem 1 proceeds as before: Just define

$$W_k = \sum_{z \in \mathbf{Z}^n} \hat{f}(z) z_k^{1/3} e^{2\pi i <z,x>}, \tag{23}$$

$$R_k = \sum_{z \in \mathbf{Z}^n, z_k \neq 0} \hat{f}(z) e^{2\pi i <z,x>}, \tag{24}$$

and replace the operator T_ϵ by $g \to (\bigotimes^n P_a) * g$.

Remark: In [KKL] stronger inequalities concerning the L_p-norms of the vector of influences $(I_f(1), ..., I_f(n))$ are proved, and some estimates on the absolute constants are given. Theorem 1 can be sharpened in a similar way. We omit the details.

References

[A] N. Alon, *On the density of sets of vectors*, Discrete Math. **46** (1983), 199–202.

[Be] W. Beckner, *Inequalities in Fourier analysis*, Annals of Math. **102** (1975), 159–182.

[BL] M. Ben-Or and N. Linial, *Collective coin flipping,* in *Randomness and Computation* (S. Micali ed.), Academic Press, New York, 1990, pp. 91–115. Earlier version: *Collective coin flipping, robust voting games, and minima of Banzhaf value*, Proc. 26th IEEE Symp. on the Foundation of Computer Science, 1985, 408–416.

[Bo] B. Bollobas, *Combinatorics,* Cambridge University Press, Cambridge, 1986.

[CG] B. Chor and M. Gereb-Graus, *On the influence of single participant in coin flipping schemes,* Siam J. Discrete Math. **1** (1988), 411–415.

[F] P. Frankl, *The shifting technique in extremal set theory,* in *Surveys in Combinatorics* (C. W. Whitehead ed.), Cambridge University Press, Cambridge, 1987, 81–110.

[KKL′] J. Kahn, G. Kalai and L. Linial, *The influence of variables on Boolean functions,* Proc. 29th Ann. Symp. on Foundations of Comp. Sci., Computer Society Press, 1988, pp. 68–80.

[KKL] J. Kahn, G. Kalai and L. Linial, *Collective coin flipping, the influence of variables on Boolean functions and harmonic analysis,* to appear.

[LW] L. Loomis and H. Whitney, *An inequality related to the isoperimetric inequality,* Bull. Am. Math. Soc. **55** (1949), 961–962.