INFORMATION SCIENCES

Volume 19
Number 2
1979

An International Journal

Abstracted in CHEMICAL ABSTRACTS; COMPUTER AND INFORMATION SYSTEMS; COMPUTING REVIEWS; ENGINEERING INDEX; LANGUAGE AND AUTOMATION; LANGUAGE AND LANGUAGE BEHAVIOR ABSTRACTS; MATHEMATICAL REVIEWS; REFERATIVNYĬ ZHURNAL MATHEMATIKA; SCIENCE ABSTRACTS; ZENTRALBLATT FÜR MATHEMATIK.

The Probability That a Matrix Has a Saddle Point

EDWARD O. THORP

Graduate School of Administration, University of California at Irvine,
Irvine, California 92717

Communicated by Frank Cannonito

ABSTRACT

Given a "randomly selected" $m\times n$ matrix, the probability that it has a saddle point is $m!n!/(m+n-1)!$.

The following question arose in correspondence between Martin Gardner and Richard Epstein. What is the probability P_n that a "randomly selected" $m\times n$ matrix has a saddle point? Recall from game theory that the i,j element a_{ij} of an $m\times n$ matrix is a *saddle point* if a_{ij} is both a minimum of its row and a maximum of its column [3]. Epstein [2] conjectures that for square matrices $P_n=n/(2^{2n-3}+n-1)$. Epstein verified the formula for $n=2,3,4$. (The formula as it appeared in [2] was misprinted with the 2 in the exponent omitted.)

The problem is solved by

THEOREM 1. *The probability $P_{m\times n}$ that a randomly selected matrix has a saddle point is $P_{m\times n}=m!n!/(m+n-1)!$ for each of the following definitions of "randomly selected."*

(1) *Let σ be a permutation of the integers $1,2,\ldots,mn$, and define the matrix $A=(a_{ij})$ by $a_{ij}=\sigma((m-1)i+j)$, $1\leqslant i\leqslant m$, $1\leqslant j\leqslant n$. Then there are $(mn)!$ permutations. We define "random selection" to mean that they are equally likely. If $f(m,n)$ is the number of permutations which produce saddle points, then $P_{m\times n}=f(m,n)/(mn)!$.*

(2) *Let the a_{ij} be independent identically distributed random variables, with a common continuous distribution (i.e., $P(a_{ij}=x)=0$ for any x).*

Proof. (1) With probability 1, all the a_{ij} are distinct in 1 and 2. Thus we may and shall assume that all a_{ij} are distinct.

 0020-0255/79/07091-05$01.75

(2) Since all saddle points have the same value, it follows that there is at most one saddle point.

(3) If all a_{ij} are distinct and $\phi: R \to R$ is order preserving (i.e., monotone increasing), then $a_{i_0 j_0}$ is a saddle point for (a_{ij}) iff $\phi(a_{i_0 j_0})$ is a saddle point for $(\phi(a_{ij}))$.

(4) For each matrix in (2), define ϕ so that it maps $\{a_{ij}\}$ onto $\{1,2,\ldots,mn\}$. Under this mapping all permutations σ of $\{1,\ldots,mn\}$ are equally probable. Thus the solution to (2) is the same as for (1).

Solution to (1). By symmetry, the number of σ producing a saddle point at $a_{i_0 j_0}$ is the same for all (i_0, j_0) pairs. Hence if $N_{1,1}$ is the number for $a_{1,1}$, then by (2) the number N for all (i,j) is $mnN_{1,1}$, and if $p_{1,1}$ is the probability of a saddle point at (1, 1), then $P_{m\times n} = mnp_{1,1}$.

Now $a_{1,1}$ is a saddle point iff $a_{i,1} < a_{1,1}$ for $2 \leqslant i \leqslant m$ and $a_{1,j} > a_{1,1}$ for $2 \leqslant j \leqslant n$. This is true iff both $a_{1,1}$ is the mth element in order of increasing size in the set S consisting of the first row and first column ($m+n-1$ elements) and also the first $m-1$ elements of the set are in the first column. Now $a_{1,1}$ is the mth element with probability $1/(m+n-1)$. There are $\binom{m+n-2}{m-1}$ equally probable ways to select the $m-1$ elements of the first column from S. Only one selects the $m-1$ smallest so the probability is $\binom{m+n-2}{m-1}^{-1}$. This is independent of the size of $a_{1,1}$, so

$$p_{1,1} = \frac{1}{m+n-1}\binom{m+n-2}{m-1}^{-1}$$

and

$$P_{m\times n} = mnp_{1,1} = \frac{m!n!}{(m+n-1)!}.$$

Note that the Epstein value $P_n = P_{n\times n}$ if and only if $n = 2,3,4$. If $n > 4$, then $P_n < P_{n\times n}$.

We need the condition in definition (2) of the theorem, that the common distribution $F(x)$ for the a_{ij} is continuous, as the example $a_{ij} \equiv 0$ for all i,j shows. This is also illustrated by an alternate proof of the theorem using definition (2). The idea was suggested by S. Karamardian.

Proof. For a given $a_{i_0 j_0}$, the conditional probability, given $a_{i_0 j_0} = x$, that $a_{i_0 j_0}$ is a saddle point is

$$P\left[\max_{i \neq i_0} a_{i j_0} \leqslant x \leqslant \min_{j \neq j_0} a_{i_0 j}\right] = [F(x^+)]^{m-1}[1-F(x)]^{n-1},$$

where $F(x)=P(a_{ij}<x)$ is the common distribution. Thus $P(a_{i_0j_0}$ is a saddle point$)=\int_{-\infty}^{\infty}[1-F(x)]^{n-1}[F(x^+)]^{m-1}\,dF(x)$. If $F(x^+)=F(x)$ for all x, i.e., if F is continuous, then this becomes $\int_{-\infty}^{\infty}[1-F(x)]^{n-1}[F(x)]^{m-1}\,dF(x)$. Again, if F is continuous (but not, in general, otherwise) this equals $B(m,n)=(m-1)!(n-1)!/(m+n-1)!$. Again, if F is continuous (but not, in general, otherwise), the probability of more than one saddle point is zero, and hence P_{mn} is mn times the probability $a_{i_0j_0}$ is a saddle point. For a proof that $\int_{-\infty}^{\infty}[1-F(x)]^{n-1}F(x)^{m-1}\,dF(x)=\int_0^1(1-y)^{n-1}y^{m-1}\,dy=B(m,n)$, use, e.g., Corollary 2.41 of [1]. The continuity of F is needed to establish that $\hat{P}(\cdot)$ is the uniform distribution on $[0,1]$.

When the distribution F has a discrete part, it appears difficult to determine the general formula for P_{mn}. Certain special cases are tractable. For example, let the a_{ij} independently be either 0 or 1 with probabilities $q=1-p$ and p respectively. A matrix of 0's and 1's has a saddle point at (i_0,j_0) iff either (A) $a_{i_0j}=1$ for all j or (B) $a_{ij_0}=0$ for all i. Let A_{i_0} be the event "$a_{i_0j}=1$ for all j," and let $A=\{\cup A_i: i=1,\ldots,m\}$. Then $P(A)=1-P(\cap_{i=1}^{m}A_i^c)$ and $P(A_i^c)=1-P(A_i)=1-p^n$, so $P(A)=1-\prod_{i=1}^{m}(1-p^n)$. Similarly if B_{j_0} is the event "$a_{ij_0}=0$ for all i" and $B=\{\cup B_j: j=1,\ldots,n\}$, then $P(B)=1-\prod_{j=1}^{n}(1-q^m)$. Finally $P_{mn}=P(A)+P(B)=2-(1-p^n)^m-(1-q^m)^n$, which proves:

THEOREM 2. *If the entries of a matrix are independently chosen to be* 0 *and* 1 *with probabilities* $q=1-p$ *and* p *respectively, then* $P_{mn}=2-(1-p^n)^m-(1-q^m)^n$.

COROLLARY 3. *The number of distinct* $m\times n$ *matrices of* 0*'s and* 1*'s with saddle points is* $2^{mn+1}-(2^n-1)^m-(2^m-1)^n$.

Proof. Set $p=\frac{1}{2}$ in Theorem 2 and multiply P_{mn} by the total number 2^{mn} of all such matrices.

Another result for $F(x)$ discrete is given in the next theorem.

THEOREM 4. *If* $m=n=2$ *and the* a_{ij} *are independent and uniformly distributed on* $x_1<x_2<\cdots<x_N$, *i.e.,* $P(a_{ij}=x_k)=1/N$, *then* $P_{2,2}=(2N^3+4N-3)/3N^3$.

Proof. We need the easily proven lemma: The set of saddle points of any matrix (b_{ij}) is a sub*matrix*.

Then we consider the following disjoint cases:

Case 1. There is exactly one saddle point. Suppose it is a_{11} and that $a_{11}=i$. Then there are three possibilities for the structure of (a_{ij}). They are, schematically,

$$\begin{pmatrix} i & >i \\ <i & \text{arb} \end{pmatrix}, \begin{pmatrix} i & >i \\ i & <i \end{pmatrix} \text{ and } \begin{pmatrix} i & i \\ <i & >i \end{pmatrix}.$$

The first has probability of

$$\sum_{i=1}^{N} \frac{1}{N} \frac{i-1}{N} \frac{N-i}{N}$$

and the last two each have probability

$$\sum_{i=1}^{N} \frac{1}{N^2} \frac{i-1}{N} \frac{N-i}{N}.$$

Summing, simplifying, and multiplying by 4 gives $2(N+2)(N-1)(N-2)/3N^3$ as the total probability of exactly one saddle point.

Case 2. There is exactly one row *of saddle points.* Schematically the three possibilities for the *first* row are

$$\begin{pmatrix} i & i \\ <i & <i \end{pmatrix}, \begin{pmatrix} i & i \\ =i & <i \end{pmatrix}, \begin{pmatrix} i & i \\ <i & =i \end{pmatrix}$$

with probabilities

$$\sum_{i=1}^{N} \frac{1}{N^2}\left(\frac{i-1}{N}\right)^2, \quad \sum_{i=1}^{N} \frac{1}{N^3} \frac{i-1}{N}, \quad \sum_{i=1}^{N} \frac{1}{N^3}\left(\frac{i-1}{N}\right).$$

Combining and doubling for two rows gives a total probability of $(N-1)(2N+5)/3N^3$.

Case 3. There is exactly one column *of saddle points.* The procedure and result is the same as in Case 2; the probability is $(N-1)(2N+5)/3N^3$.

Case 4. All 4 points are saddle points. The probability is $\Sigma_{i=1}^{N} 1/N^4 = 1/N^3$. Combining the results of all four cases gives $(2N^3+2N^2-2N+1)/3N^3 \equiv P_{2,2}(N)$.

REMARKS. $P_{2,2}(1)=1$, $P_{2,2}(2)=\frac{7}{8}$, which agrees with Theorem 2, and $\lim_{N\to\infty} P_{2,2}(N)=\frac{2}{3}$, which agrees with Theorem 1. Note too that $P_{2,2}(N)$ is monotone decreasing in N, whence $1 \geqslant P_{2,2}(N) > \frac{2}{3}$. This is consistent with the following.

CONJECTURE. If F is a distribution with a discrete part and $P_{m,n}(F)$ is the probability that (a_{ij}) has a saddle point, then $P_{m,n}(F) > P_{m,n}$ if $m,n>1$.

Note: Emmett Keeler called my attention to the earlier paper by Goldman, who proves Theorem 1, part 2, in much the same way. Part 1 follows readily. Goldman also gives the $n=2$ case of Theorem 2. Keeler also cites RAND memo 5768 (1968) by Dresher which considers nonzero sum games with two or more players. Peter Griffin pointed out an error in an earlier version.

REFERENCES

1. Leo Breiman, *Probability*, Addison-Wesley, 1968.
2. R. Epstein, *The Theory of Gambling and Statistical Logic*, Academic, 1967; revised ed., 1977, p. 37.
3. S. Karlin, *Mathematical Methods and Theory in Games, Programming, and Economics*, Addison-Wesley, 1959.

Received May 1978