COVERAGE OF A SQUARE LATTICE BY AN INCLINED RECTANGLE

DENNIS ROSEN,* *Birkbeck College, London*

Abstract

By use of a method of dissection, a formula is derived for the variance in the number of lattice points covered by a rectangle of arbitrary size, lying on a square lattice of unit spacing and inclined to the lattice at an angle of which the tangent is a rational fraction.

GEOMETRIC PROBABILITY; COVERAGE

1. Introduction

Kendall and Moran (1963) quote a formula for the variance in the coverage of a square lattice by a rectangle of arbitrary size, randomly positioned, but with its sides parallel to the two directions of the lattice. However, they note the absence of a corresponding formula for the case where the orientation of the rectangle is also random. The result presented here gives the variance when the rectangle is randomly positioned at a fixed angle to the lattice, with the tangent of the angle a rational fraction; it is obtained by dissecting the rectangle into appropriate regions for which the variances and covariances can be calculated. Although falling far short of the goal of dealing with a randomly oriented rectangle, the result is of use in tasks of automated image analysis involving mensuration of rectangular areas.

2. Areas of constant counting measure

With respect to a square lattice with points specified by integer pairs (x_i, y_i), the counting measure of a set is the number of lattice points covered. The lines passing through precisely the lattice points $(a_0 + na, b_0 + nb)$ with a, b coprime and $n = 0, \pm 1, \pm 2, \cdots$, will be called θ-lines, with $\theta = \tan^{-1}(a/b)$. The lines make an angle θ with the x-axis. The distance between consecutive points on a θ-line is $D = (a^2 + b^2)^{\frac{1}{2}}$ and the perpendicular distance between adjacent θ-lines is $U = (\cos \theta)/b = (a^2 + b^2)^{-\frac{1}{2}} = 1/D$. A parallelogram will be taken to include its left-hand edge with both ends and its lower edge with its left-hand end, but exclude the rest of the boundary.

Lemma. Every parallelogram with a pair of sides of length D, inclined to the x-axis at an angle θ, and separated by a distance U, contains exactly one lattice point.

Proof. Each such parallelogram is intersected by exactly one θ-line, on which it cuts off an intercept of length D. Hence it contains exactly one lattice point.

Thus a parallelogram with the charcteristics just stated is a set of constant measure, namely of unit measure. When such a parallelogram is a rectangle it will be referred to as a standard unit rectangle. The union of disjoint sets of constant measure has constant measure.

3. The measure of an inclined rectangle

For a rectangle of sides L_1 and L_2, inclined to the x-axis of the lattice at angle $\theta = \tan^{-1}(a/b)$, the sides may be written as $L_1 = r_1 D + s_1 U + V_1$, and $L_2 = r_2 D + s_2 U + V_2$ where D and U are defined as above, r_i, s_i are integers and $V_i < U \leqq 1$ since $D \geqq 1$. Thus the
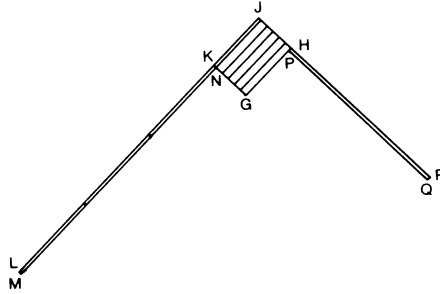
Figure 1. $LJ = L_1$, $LK = r_1 D$, $KJ = M_1 = s_1 U + V_1$: $RJ = L_2$, $RH = r_2 D$, $HJ = M_2 = s_2 U + V_2$.

rectangle is decomposed into $r_1 r_2$ squares of side D and of measure D^2, $s_1 r_2 + s_2 r_1$ standard unit rectangles and a remaining area, illustrated in Figure 1, which can be considered in terms of its constituent regions. Region I, $KLMN$ in Figure 1, consists of r_1 strips of length D and width V_2 lying end to end and parallel to side L_1 of the rectangle. Region II, $PQRH$ in Figure 1, consists of r_2 strips of length D and width V_1, lying end to end and parallel to side L_2 of the rectangle. The final part of the remaining area, $GHJK$ in Figure 1, is a small rectangle of sides $M_1 = s_1 U + V_1$ and $M_2 = s_2 U + V_2$ and can be decomposed either as s_2 strips of length M_1 and width U together with one of length M_1 and width V_2 or else as s_1 strips of length M_2 and width U together with one of length M_2 and width V_1. It is immaterial which decomposition is used, although for numerical calculation it is convenient to take that which includes the smaller number of strips. Suppose that $s_2 < s_1$ so that the first decomposition is taken, as illustrated in Figure 1. Then it is convenient to consider Region III as consisting of s_2 strips of length M_1 and width U and Region IV as the single strip of length M_1 and width V_2. The mean measures and the variances and covariances of and between the four regions can now be calculated by integrating over a standard unit rectangle .

By applying elementary methods it immediately follows that the mean measure $\mathcal{M}$ of the whole rectangle equals its area, that the variances of Regions I, II and IV are $r_1^2 DV_2(1 - DV_2)$, $r_2^2 DV_1(1 - DV_1)$ and $M_1 V_2(1 - M_1 V_2)$ respectively, the covariance, cov (I, II), between Regions I and II is zero, cov (I, IV) $= r_1 M_1 V_2(1 - DV_2)$ and cov (II, IV) $= r_2 V_1 V_2(1 - DV_1)(1 - M_1 V_2 + V_1 V_2)$.

The variance of Region III involves cov (i, j), the covariance between the ith and jth strips. The latter may be calculated by moving the rectangle so that G, the bottom left-hand corner of the first strip (Figure 1) is a lattice point and then rotating axes about G so that the new axes (X, Y) are inclined at angle θ to the old (x, y). The s_2 strips of Region III are fractions, M_1/D, of standard unit rectangles, in the ith one of which the lattice point is at (X_i, Y_i) where $X_i = y_i \sin \theta + x_i \cos \theta$, $Y_i = y_i \cos \theta - x_i \sin \theta = i/D$ and x_i, y_i are appropriate solutions of the Diophantine equation

$$\text{(1)} \qquad\qquad\qquad\qquad y_i b - x_i a = i.$$

With $E_{ij} = |X_j - X_i|$, $F_{ij} = D - E_{ij}$, and G_{ij} equal to the lesser of E_{ij} and F_{ij}, it then follows that

$$\text{(2)} \qquad
\begin{aligned}
&\text{for } M_1 \leqq E_{ij}, &&\text{cov } (i, j) = -M_1^2/D^2; \\
&\text{for } E_{ij} \leqq M_1 \leqq F_{ij}, &&\text{cov } (i, j) = (DM_1 - DG_{ij} - M_1^2)/D^2; \\
&\text{for } M_1 \geqq F_{ij}, &&\text{cov } (i, j) = -(D - M_1)^2/D^2.
\end{aligned}$$

The covariances between the strips of Region III and that of Region IV can be written as cov (i, s_2) and have values equal to the corresponding cov (i, j) values multiplied by $V_2 D$.

The covariance between Region I and each strip of Region III is zero. The covariance between Region II and each strip of Region III is $r_2 V_1(1 - DV_1)(D - M_1 - V_1)/D^2$.

Since var (m), the variance on the measure, m, of the rectangle, is given by Σ var $(l) + 2\Sigma$ cov (l, m) where $l, m = \text{I, II, III, IV}$, then on collecting terms and rearranging,

$$
\text{var}(m) = V_2(1 - DV_2)L_1^2/D + V_1(1 - DV_1)L_2^2/D
$$

(3)
$$
+ M_2\{M_1(1 - M_1 U) - UV_1(1 - DV_1)(2r_2 M_1 - 2r_2 V_1 + M_2)\}
$$

$$
+ 2\Sigma \text{ cov}(i, j) + 2\Sigma \text{ cov}(i, s_2),
$$

where cov (i, j) and cov (i, s_2) are given by (2) and the modification of (2) as noted. The case where $s_2 = 0$ is included in (3) if the last two terms are then taken as zero.

4. Comments on the results

If M_1, the remainder of L_1/D, is less than or equal to $\sin \theta$, then

(4) $\qquad 2\Sigma \text{ cov}(i, j) + 2\Sigma \text{ cov}(i, s_2) = -(s_2^2 - s_2 + 2DV_2)M_1^2/D_2$

and if $M_1 \geqq D - \sin \theta$ then

(5) $\qquad 2\Sigma \text{ cov}(i, j) + 2\Sigma \text{ cov}(i, s_2) = -(s_2^2 - s_2 + 2DV_2)(D - M_1)^2/D_2.$

Except for these cases, (3) must be evaluated after collecting solutions to (1) with $i = 0, 1, 2, \cdots, s_2$, although with some combinations of L_1 and θ, (4) and (5) will also be true with less restricted ranges of M_1, thereby simplifying the formula for var (m). The variables which may be set, that is, the size of the rectangle through L_1, L_2 and its inclination to the lattice through $\theta = \tan^{-1}(a/b)$ make var (m) vary in a highly discontinuous way.

When the rectangle is aligned with the lattice so that θ and a are zero, $b = D = u = 1$, $r_2 = s_2 = 0$ and $M_1 M_2$ reduce to V_1, V_2, (3) reduces to the formula of Kendall and Moran (1963), namely

$$
\text{var}(m) = V_2(1 - V_2)L_1^2 + V_1(1 - V_1)L_2^2 + V_1 V_2(1 - V_1)(1 - V_2).
$$

If the formula derived here is used in image analyses of digitized images then care must be taken, as previously noted (Rosen (1980)), to distinguish between lattice and grid. In the processing of digitized images, it is usual to work with pixels (picture elements) and if these are taken to lie in a lattice then the edges of the area formed by a group of pixels lie on a grid which transects the lattice.

References

KENDALL, M. AND MORAN, P. A. P. (1963) *Geometrical Probability*. Griffin, London.

ROSEN, D. (1980) On the areas and boundaries of quantized objects. *Comp. Graphics Image Proc.* **13,** 94–98.