Indian J. pure appl. Math., 33(9) : 1373-1377, September 2002
© Printed in India.

THE UNEXPECTED LINEAR ALGEBRA OF EXPECTATION

J. M. HORGAN* AND J. WALDRON**

**School of Computer Applications, Dublin City University, Dublin, Ireland*
***Department of Computer Science, Trinity College, Dublin, Ireland*

(*Received* 3 *May* 2001; *accepted* 7 *December* 2001)

Horgan *et al.* proposed a coefficient to measure the instruction count of two compilers of Java code. This paper investigates the linear algebra of the coefficient.

Key Words : Compilers; Contingency Measures

1. BACKGROUND

Suppose $x = (x_j)$ and $y = (y_j)$ are variables (j = 1, 2, ..., n): in our motivating example they describe the instruction count of compilers or programs P_x and P_y. As a measure of the similarity or otherwise of the compilers or programs we first form the $n \times 2$ matrix $a = (a_{ij}) = (x \mid y)$ whose columns are given by x and y, and then compare it with a new matrix $J(a)$, where we write

$$\left(\sum_{ij} a_{ij} \right) J(a) = ((x_j + y_j) \Sigma_k x_k \quad (x_j + x_y) \Sigma_k y_k), \qquad \ldots (1)$$

the $n \times 2$ matrix whose columns are multiples of the sum of the two columns of $a = (a_{ij})$ by the sum of the column elements. Now as a measure of the statistical 'independence' of the rows and columns of the original matrix a we write

$$\| a - J(a) \|^2 = \sum_{ij} (a_{ij} - J(a)_{ij})^2, \qquad \ldots (2)$$

This could be thought of as a "normalised mean square contingency measure": we have used it in connection with the comparison of the performance of two compilers[1]. As a measure of the association between a and $J(a)$ we consider instead the expression

$$\sum_{ij} \frac{(a_{ij} - J(a)_{ij})^2}{J(a)_{ij}}, \qquad \ldots (3)$$

and make some observations.

2. PRELIMINARIES

Our first observation is that this is subject to an inequality :

Theorem 1 — *If* $x=(x_j)$ *and* $y=(y_j)$ *are n-tuples of positive numbers and* $a=(x\,|\,y)\in \mathbf{R}^{n\times 2}$ *then*

$$\sum_{ij}\frac{(a_{ij}-J(a)_{ij})^2}{J(a)_{ij}}\leq \sum_j a_{ij}. \qquad \text{... (4)}$$

PROOF : We compute

$$\left(\sum_j x_j+\sum_j y_j\right)\left(\sum_i \frac{\left(x_i-\frac{(x_i+y_i)\,\Sigma_j x_j}{\Sigma_j x_j+\Sigma_j y_j}\right)^2}{\frac{(x_i+y_i)\,\Sigma_j x_j}{\Sigma_j x_j+\Sigma_j y_j}}+\sum_i \frac{\left(y_i-\frac{(x_i+y_i)\,\Sigma_j y_j}{\Sigma_j x_j+\Sigma_j y_j}\right)^2}{\frac{(x_i+y_i)\,\Sigma_j y_j}{\Sigma_j x_j+\Sigma_j y_j}}\right)$$

$$=\sum_i \frac{(\Sigma_j x_j+\Sigma_j y_j)\,x_i-(x_i+y_i)\,\Sigma_j x_j)^2}{(x_i+y_i)\,\Sigma_j x_j}+\sum_i \frac{(\Sigma_j x_j+\Sigma_j y_j)\,y_i-(x_i+y_i)\,\Sigma_j y_j)^2}{(x_i+y_i)\,\Sigma_j y_j}$$

$$=\sum_i \frac{((\Sigma_j y_j)\,x_i-y_i\,\Sigma_j x_j)^2}{(x_i+y_i)\,\Sigma_j x_j}+\sum_i \frac{((\Sigma_j y_j)\,x_i-x_i\,\Sigma_j y_j)^2}{(x_i+y_i)\,\Sigma_j y_j}$$

$$=\left(\frac{1}{\Sigma_j x_j}+\frac{1}{\Sigma_j y_j}\right)\sum_i \frac{((\Sigma_j y_j)\,x_i-y_i\,\Sigma_j y_j)^2}{(x_i+y_i)}$$

$$=\frac{\Sigma_j x_j+\Sigma_j y_j}{(\Sigma_j x_j)\,(\Sigma_j y_j)}\sum_i \frac{((\Sigma_j y_j)\,x_i-y_i\,\Sigma_j x_j)^2}{(x_i+y_i)},$$

which is $\leq(\Sigma_j x_j+\Sigma_j y_j)^2$: for we claim

$$\sum_i\left(\left(\sum_j y_j\right)^2 x_i^2+\left(\sum_j x_j\right)^2 y_i^2\right)\prod_{k\neq i}(x_k+y_k)$$

$$\leq\left(\sum_j x_j\right)\left(\sum_j y_j\right)\left(\sum_j x_j+\sum_j y_j\right)\prod_k (x_k+y_k).$$

Indeed we count $n\cdot 2n^2\cdot 2^{n-1}=n^3\,2^n$ terms on the left and $n\cdot n\cdot 2n\cdot 2^n=n^3\,2^{n+1}$ terms on the right, and can pick off a match on the right for each term on the left •

3. LINEAR ALGEBRA OF THE DISSIMILARITY COEFFICIENT

In this note we want to look at some of the curious linear algebra hiding behind the construction. If $a = (a_{ij}) \in A^{m \times n}$ is a general matrix over a ring A we shall write

$$col_k(a) = \begin{pmatrix} a_{1k} \\ a_{2k} \\ \cdots \\ a_{mk} \end{pmatrix} \in A^{m \times 1}, \quad row_j(a) = (a_{j1}\, a_{j2} \ldots a_{jn}) \in A^{1 \times n} \qquad \ldots (5)$$

for the kth column and jth row of the matrix a, so that

$$\sigma(a) = \sum_{jk} a_{jk} = \upsilon \sum_{k} col_k(a) = \sum_{j} row_j(a)\, u \in A \qquad \ldots (6)$$

is the total sum of the entries, where

$$u = \begin{pmatrix} 1 \\ 1 \\ \cdots \\ 1 \end{pmatrix} \in A^{n \times 1} \text{ and } \upsilon = (1\ 1 \ldots 1) \in A^{1 \times m} \qquad \ldots (7)$$

Definition 1 — If $a \in A^{m \times n}$ and $b \in A^{p \times q}$ are arbitrary set

$$K(a, b) = \sum_{jk} col_k(a)\, row_j(b) = au\, \upsilon b \in A^{m \times q}. \qquad \ldots (8)$$

When $p = m$ and $q = n$ we write

$$\sigma(a)\ J(a) = \upsilon a u J(a) = K(a, a) \in A^{m \times n} \qquad \ldots (9)$$

Naturally we will be most interested in the commutative rings $A = \boldsymbol{R}$ and $A = \boldsymbol{C}$ of real or of complex numbers. The matrix $K(a, b)$ is evidently a bilinear function of a and b, but in general far from invertible, being always of rank ≤ 1 :

Theorem 2 — *If $a \in A^{m \times n}$ is commutative then*

$$J(J(a)) = J(a) \in A^{m \times n} \qquad \ldots (10)$$

PROOF : Compute

$$\sigma K(a, b) = \sigma(a)\, \sigma(b), \qquad \ldots (11)$$

$$col_i K(a, b) = \sum_{kj} col_k(a)\, b_{ji} \text{ and } row_i K(a, b) = \sum_{kj} a_{ik}\, row_j(b),$$

while of course

$$col_i(ab) = \sum_{k} col_k(a)\, b_{ki} \text{ and } row_i(ab) = \sum_{j} a_{ij}\, row_j(b).$$

Provided A is commutative this gives in particular

$$K(a, K(b, c)) = \sigma(b)\, K(a, c) = K(K(a, b), c)$$

and hence

$$K(K(a, b), K(c, d)) = \sigma(b)\, \sigma(c)\, K(a, d).$$

Taking $a = b = c = d$ then gives (10) •

When $a = (x \mid y)$ and $b = (z \mid w)$ consist of two columns we get

$$K(a, b) = K(x \mid y, z \mid w) = (x + y)\, (\sigma(z) \quad \sigma(w)). \qquad \ldots (12)$$

If instead of matrices we look at linear operators then the product K takes the form

$$K : L(X, Y) \times L(Z, W) \to L(Z, Y)$$

given by

$$K(a, b) = w'\, b \odot ax : y \to w'(by)\, ax \qquad \ldots (13)$$

with $x \in X$ and $w' \in L(W, \boldsymbol{K})$, where K is either the real field $\boldsymbol{R}$ or the complexes $\boldsymbol{C}$. Now

$$(w'\, ax)\, J(a) = K(a, a) \qquad \ldots (14)$$

with $w' \in L(Y, \boldsymbol{K})$.

Theorem 3 — *If K is given by (13) and J by (14) with $x = y = z = w$ then, provided $w'\, ax \neq 0$.*

$$\sigma(a) = w'\, ax \text{ is an eigenvalue of } K(a, 1) = w' \odot ax \qquad \ldots (15)$$

and $$\sigma(a) = w'\, ax \text{ is a co-eigenvalue of } K(1, a) = w'\, a \odot x. \qquad \ldots (16)$$

PROOF : ax is the eigenvector for $w' \odot ax$, and $w'\, a$ the co-eigenvector for $w'\, a \odot x$ •

The mapping J is "positive", in the sense that if $A = \boldsymbol{R}$ and if all the entries of $a \in A^{m \times n}$ are positive then the same is true of $J(a)$:

Theorem 4 — *If $a \in A^{m \times n}$ there is implication*

$$a_{ij} \geq 0 \Rightarrow J(a)_{ij} \geq 0. \qquad \ldots (17)$$

PROOF : Observe that

$$a_{ij} \geq 0, b_{ij} \geq 0 \Rightarrow K(a, b)_{ij} \geq 0 \bullet \qquad \ldots (18)$$

When A is topological, or more generally normed[2], then J is continuous, and can be differentiated :

Theorem 5 — *The differential $d_a J$ of J at $a \in A^{m \times n}$ is given by*

$$\sigma(a)\, d_a J(h) = K(a, h) + K(h, a) - \sigma(h)\, J(a). \qquad \ldots (19)$$

PROOF : Compute •

In the case of bounded operators on normed spaces we are interested in the distance

$$\sigma(a) \| a - J(a) \| = \| \sigma(a)\, a - K(a, a \|$$

$$= \sup_{y,\, y' \leq 1} | (w' ax)(y' ay) - (y' ax)(w' ay) |. \qquad \ldots (20)$$

The quantity (3) of the matrix case is more difficult to identify in a linear context.

REFERENCES

1. J. M. Horgan, J. Power and J. Waldron, *Measurement and analysis of Runtime Profiling Data for Java Programs, Technical* Report NUIM-CS-TR-2001-04, Department of Computer Science, NUI Maynooth, February 2001.
2. R. Harte, *Invertibility and singularity for bounded linear operators*, Dekker New York 1988.