

Matrix Algebra Problems with Applications in Psychology and Multivariate Analysis

Psychology 6140

September 6, 2016

Students in Multivariate Analysis need to develop skills in working with matrices and reading matrix expressions. Some of the skills which seem to be important are:

1. Performing simple matrix calculations
2. Manipulating algebraic expressions in matrices—substituting identities, simplifying expressions, etc.
3. Formulating substantive and statistical problems in matrix terms—or, recognizing such in journal articles.
4. Understanding certain key results (theorems) of matrix algebra.
5. Recognizing these results when they are used in your text.

The following problems are directed toward developing these skills. Along the way, some matrix applications in psychology and multivariate statistics are introduced.

I. Elementary matrix expressions

I-1. Factoring a matrix into the product of two matrices is the basis of many methods in multivariate analysis. Choleski's method involves factoring a matrix into the product of two *triangular* matrices, one of which has 0s above the diagonal, the other below the diagonal. Fill in the ? entries in the following matrix equations:

Hint: Let each ? be an unknown letter; solve the equations.

(a)

$$\begin{bmatrix} 5 & 10 \\ 2 & 7 \end{bmatrix} = \begin{bmatrix} ? & 0 \\ ? & ? \end{bmatrix} \times \begin{bmatrix} 1 & ? \\ 0 & 1 \end{bmatrix}$$

(b)

$$\begin{bmatrix} 3 & -3 & 6 \\ -1 & 5 & 14 \\ -3 & 1 & 12 \end{bmatrix} = \begin{bmatrix} ? & 0 & 0 \\ ? & ? & 0 \\ ? & ? & ? \end{bmatrix} \times \begin{bmatrix} 1 & ? & ? \\ 0 & 1 & ? \\ 0 & 0 & 1 \end{bmatrix}$$

I-2. Premultiply the matrix M by each of the following matrices. Describe in words in each case what effect the premultiplier has on the rows of M .

$$M = \begin{bmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{bmatrix}$$

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad B = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \quad C = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

I-3. Let D be an $m \times n$ matrix and let $e_i = (0, 0, \dots, 1, 0, \dots, 0)'$ denote the n -dimensional vector consisting of zeros except for the i -th element, which is 1.

- (a) What is $D e_i$ (in words)?
 (b) Express the matrices A and B from Problem I.2 above in terms of e vectors.

I-4. Experimenters in color vision have found that a subject can match any spot of colored light by an additive mixture of three colored lights of fixed spectral composition (Judd, 1951). Any three colors can be used as the primary lights provided that none is a combination of the other two.

Say that an experimenter has a colorimeter that has three primary lights with values r , g , and b . She wants to specify her results in terms of a standard set of tri-stimulus values, R , G , and B . By experiment she finds that the amount of each of the standard primaries needed to match each of his primaries is

$$\begin{aligned} r &= a_{11} R + a_{12} G + a_{13} B \\ g &= a_{21} R + a_{22} G + a_{23} B \\ b &= a_{31} R + a_{32} G + a_{33} B \end{aligned}$$

Define the appropriate vectors and matrices to express the above system of equations in matrix terms.

I-5. Let $\mathbf{x} = (x_1, \dots, x_n)'$ be a vector containing the number of units purchased of each of a variety of grocery items. Let $\mathbf{y} = (y_1, \dots, y_n)'$ be a vector of unit prices, such that $y_i =$ the price/unit of item i . For example, $\mathbf{x} = (4, 3, 2)'$ and $\mathbf{y} = (.95, .25, 6.50)'$ might represent 4 dozen eggs at \$0.95 per dozen, 3 lbs. of apples at \$0.25/lb, and 2 cans of pate de fois gras at \$6.50 per can (cheap, if it's *entier*).

- (a) Formulate a matrix expression for the total (net) cost of the commodities in \mathbf{x} .
 (b) Suppose each commodity is subject to a particular rate of tax, these being given by a vector, $\mathbf{t} = (t_1, \dots, t_n)'$ so that if commodity i is taxed at 5%, $t_i = 0.05$. Formulate an expression in terms of matrices and vectors for the total cost of \mathbf{x} including taxes. [Remember, cost after tax = net cost $\times (1 + t)$.]

I-6. Cliff (1959) showed that the affective value of an adverb-adjective combination (e.g., *moderately nice*) could be predicted quite accurately by multiplying an affective value for the adjective (*nice*) by an intensity value for the adverb (*moderately*), i.e., if *nice* had an affective value of 5.4, and *moderately* an intensity value of .75, then *moderately nice* have a predicted affective value of $5.4 \times .75 = 4.05$.

Suppose one took a set of m personality-trait adjectives like *industrious*, *happy*, *carefree*, *shy*, etc. and formed all possible combinations of these with n adverbs, *seldom*, *somewhat*, *moderately*, *extremely*— $m \times n$ pairs.

If a group of people rated each combination on an evaluative scale, the average ratings could be assembled in an $m \times n$ matrix, $\mathbf{C} = \{c_{ij}\}$ = average rating of adverb i and adjective j . Let $\mathbf{v} = (v_1, \dots, v_m)'$ be a vector of intensity values for the adverbs and $\mathbf{d} = (d_1, \dots, d_n)'$ be a vector of average ratings for the adjectives by themselves. Cliff's model was:

$$c_{ij} = v_i \times d_j + k$$

where k is a constant. [Cliff's model fit the data rather well, in fact, and created considerable interest in the possibility of developing mathematical models of aspects of language.]

Hint: The number of summed terms is the number of columns in a matrix product.

- (a) Express \mathbf{C} in this model as a product of two matrices.
 (b) If this model holds, what is the rank of \mathbf{C} ?

I-7. A quadratic function of a scalar variable, x can be written

$$f(x) = b_0 + b_1 x + b_2 x^2$$

- (a) Express $f(x)$ as an algebraic expression of the vectors $\mathbf{b} = (b_0, b_1, b_2)'$ and $\mathbf{x} = (1, x, x^2)'$.

- (b) In terms of this expression, evaluate $f(2)$ for $\mathbf{b} = (b_0, b_1, b_2)'$.
 (c) Generalize to find an expression for an n -th degree polynomial of x ,

$$f(x) = b_0 + b_1 x + b_2 x^2 + \cdots + b_n x^n$$

- (d) If person i has a value of x_i , let \mathbf{x}_i be the vector $(1, x_i, x_i^2)'$. Write a matrix expression which will give the values of $f(x_i)$ as a vector for all persons in a group.

[Completing this question, you have just (re-)discovered the General Linear Model!]

I-8. Let \mathbf{j}_n be a column vector of n unities (ones). Using the following vectors,

$$\begin{aligned} \mathbf{a} &= (1, 3, 5)' & \mathbf{c} &= (2, 1, 1)' \\ \mathbf{b} &= (-5, 2, 3)' & \mathbf{x} &= (x_1, x_2, x_3)' \end{aligned}$$

find each of the following inner products with \mathbf{j}_3 . State in words the effect of taking an inner product with \mathbf{j}_n .

- (a) $\mathbf{a}' \mathbf{j}_3$
 (b) $\mathbf{b}' \mathbf{j}_3$
 (c) $\mathbf{j}_3' \mathbf{c}$
 (d) $\mathbf{x}' \mathbf{j}_3$

[Matrix notation was invented to simplify the expression of linear equations. The operation of multiplying vectors by the unit vector has a particularly simple interpretation.]

I-9. Let \mathbf{X} be a $N \times p$ matrix consisting of the scores of N people on p tests, where x_{ij} = the score of person i on test j . Define the matrix \mathbf{Z} ($N \times p$) of standard scores, $z_{ij} = (x_{ij} - \bar{x}_j) / s_j$. From the fact that the correlation between two standardized variables is $r_{ij} = (\frac{1}{N}) \sum_{i=1}^N z_{ji} z_{ik}$, show that the $p \times p$ matrix of all intercorrelations can be expressed as

$$\mathbf{R} = \left(\frac{1}{N}\right) \mathbf{Z}' \mathbf{Z}$$

[In standard scores, the correlation is simply the average cross-product. Isn't that nice?]

II. Applications: Choice, preference and graph theory

The problems in this section deal with applications of matrix algebra to representing relationships among objects by matrices and directed or undirected graphs. Two basic type of relations between objects can be represented in this way: An **incidence matrix** contains entries of 0 or 1 in cell i, j to indicate that the object in row i is related in a symmetric way to the object in column j . A **dominance** or **preference** matrix also uses binary entries, but the relation is not necessarily symmetric, as in "person i likes person j ", or "player i beats player j " in a tournament.

II-1. In a sociometric experiment, members of a group are asked which other members they like. Suppose the data are collected in a choice diagram as given in Figure 1, where an arrow going from i to j means that " i likes j ".

- (a) Convert the diagram to a matrix, \mathbf{C} , where $c_{ij} = 1$ if i likes j , and 0 otherwise. The diagonal elements, $c_{ii} = 0$.
 (b) Let \mathbf{u} be an $n \times 1$ vector with unit elements. One might say that $\mathbf{u}' \mathbf{C}$ gives scores for "popularity". Explain why.
 (c) Explain how $\mathbf{C} \mathbf{u}$ can be interpreted as scores for "generosity".
 (d) Compute the matrix $\mathbf{C}^2 = \mathbf{C} \times \mathbf{C}$ and interpret the elements of this matrix.
 (e) Try also to find an interpretation for the elements of \mathbf{C}^3 .
 (f) Is there any interpretation you can come up with for the elements of $\mathbf{C} \mathbf{C}'$ or $\mathbf{C}' \mathbf{C}$?

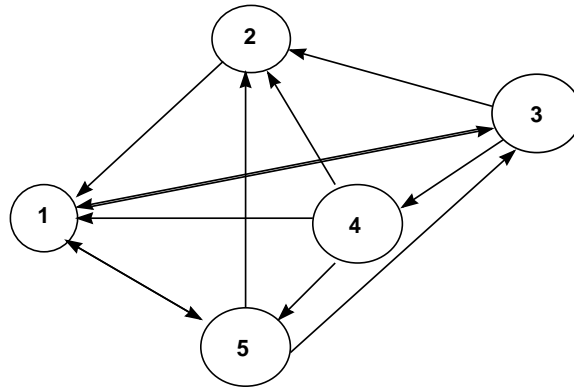


Figure 1: Directed graph for a sociometric experiment. A directed arrow points from a person to the person he/she likes.

II-2. Suppose B is a similar matrix of choices, where

$$b_{ij} = \begin{cases} 1 & i \text{ likes } j \\ -1 & i \text{ dislikes } j \end{cases}$$

Imagine a similar matrix, P , where $p_{ij} = 1$ if j believes that i likes him and $p_{ij} = -1$ if j believes that i dislikes him. The diagonal elements of B and P are equal to 0. It will be helpful for you to make up a small concrete example for yourself, with two diagrams, one for like/dislike (B) and the other for belief about like/dislike (P)

- Let $N = P'B$, so that $n_{ij} = \sum_k p_{ki} b_{kj}$. Explain why n_{ij} can be interpreted as a measure of “identification” between person i and person j .
 - Determine whether the diagonal elements of N , $n_{ii} = \sum_k p_{ki} b_{ki}$ can be interpreted as a measure of “realism” (k is realistic if his beliefs about who likes him agree with reality).
 - Determine whether the diagonal elements of $B P'$ can be interpreted as a measure of “overtness” (k is overt if his beliefs about X conform to X’s likes and dislikes).
- II-3. The expression, $u' C$, where u is a unit vector, tells how often a person is liked, so it is a measure of popularity. Suppose two persons are chosen equally often, but the first is liked by popular people and the second is liked by less popular people; then it may be reasonable to say that the first person is more popular, since he/she has more indirect choices.

Suppose now we want to combine direct and indirect choices in a measure of popularity, defined as follows: Define a “transmission” parameter, $0 < p < 1$, and assume that each direct choice contributes an amount p to a person’s popularity; an indirect choice through one intermediate person contributes an amount p^2 ; an indirect choice through two intermediates contributes an amount p^3 , etc.

- Show that the measure that takes into account all indirect choices in this way can be found from the vector $u'T$, where

$$T = pC + p^2 C^2 + p^3 C^3 + \dots + p^{n-1} C^{n-1} \quad (1)$$

- Show that the following identity holds for T defined above:

$$T(I - pC) = pC \quad (2)$$

This equation provides a way to solve for T without evaluating the entire matrix sum in equation (1).

Hint: Form a second equation by multiplying (1) by C , and subtract.

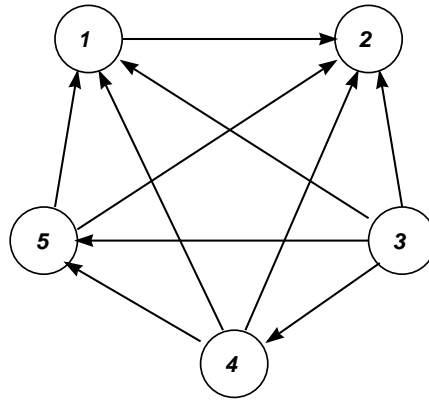


Figure 2: Choice graph for a paired comparison experiment.

- (c) Show that equation (2) above is equivalent to

$$t'(I - pC) = ps$$

where, $t' = u'T$ and $s' = u'C$. [Source: Van de Geer (1971)]

- II-4. Consider a hypothetical sociometric choice matrix for three people, as in the previous problem:

$$C = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$$

and assume that the quantity p , the relative effectiveness of transmission thru a sociometric link, has the value $1/2$. Solve for the respective elements of t , the vector of status indices.

- II-5. In a paired comparison experiment, a subject is presented with all pairs of stimuli that can be formed out of a set (e.g., political candidates), and is asked to indicate for each pair which stimulus he prefers. For n stimuli, the data can be collected in square matrix, X of order $n \times n$ where $x_{ij} = 1$ if j is preferred to i and is 0 otherwise. The diagonal elements are zero.

- Verify that if $x_{ij} = 1$ then $x_{ji} = 0$ where $i \neq j$.
- Figure 2 gives a diagram of choices, where an arrow going from i to j indicates that i is preferred to j . Convert this diagram to a choice matrix X .
- Calculate $u'X$ where u is a unit vector and interpret the elements of the resulting vector.
- Rearrange the rows and columns of X in such a way that all 1 elements are above the diagonal, so the result is a triangular matrix. Interpret the order of the rows and columns in this new matrix.
- Inspection of the diagram reveals that all choices are *transitive*, that is for any three stimuli, if i is preferred to j and j to k , then i is always preferred to k . If k were preferred to i , we would have a *circular triad*. Draw a diagram in which circular triads do occur, and convert it to the corresponding choice matrix. Calculate $u'X$ for this matrix, and compare the value for the result for a transitive matrix.
- Let $v = u'X$. Compare the values of $v'v$ as calculated from the figure with those calculated from the matrix X you constructed in part (e). What effect does the presence of circular triads have on the values of $v = u'X$?

- II-6. The problem in archeology of placing sites and artifacts in proper chronological order is called *sequence dating* or *seriation*. The same problem, of "putting in order" arises in developmental psychology and test theory, but the context in archeology is simplest.

An assumption made in archeology is that graves which are “close together” in temporal order will be more likely to have similar contents than graves further apart in time. Consider the situation in which various types of pottery, P_1, P_2, \dots, P_n are contained in graves G_1, G_2, \dots, G_m . Let \mathbf{A} be the $m \times n$ matrix with elements,

$$a_{ij} = \begin{cases} 1 & \text{if } G_i \text{ contains } P_j \\ 0 & \text{if } G_i \text{ doesn't contain } P_j \end{cases}$$

- Show that the element g_{ij} of the matrix $\mathbf{G} = \mathbf{A} \mathbf{A}'$ is equal to the number of common varieties of pottery found in *both* graves G_i and G_j .
- Show that the diagonal element, g_{ii} of \mathbf{G} gives the number of varieties of pottery in grave G_i .
- Show that the archeological assumption stated above is equivalent to the statement that *the larger the g_{ij} , the closer the graves G_i and G_j are in time.*
- If five types of pottery are contained in four graves according to the matrix \mathbf{A} below, find \mathbf{G} , and arrange the four graves in time order.

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix}$$

[Source: Williams (198?), Kendall (1969)]

- II-7. In a Markov model for short term serial recall, Murdock (1976) suggests that an item (e.g., nonsense syllables, like DAXUR) may be in any one of three states for a learner:

state IO where both item and order information have been retained (i.e., the learner remembers the elements of the item and the order in which they occurred).

state I where only item information is retained (i.e., s/he remembers it contained D, A, X, U, R but not what order they were in).

state N where neither type of information is retained.

A *Markov model* is specified by identifying the states and by specifying a “transition matrix”, which contains the probabilities that the item changes from one state to another. In Murdock’s model, the transition probabilities are shown symbolically in the matrix \mathbf{T} below.

$$\mathbf{T} = \begin{bmatrix} 1 - a & a(1 - b) & ab \\ 0 & 1 - c & c \\ 0 & 0 & 1 \end{bmatrix}$$

The parameters are defined as follows:

- a gives the probability that an item will leave the state IO.
- If it leaves that state, b is the probability that item information as well as order information will be forgotten.
- Given that the item is in state I (order information has not been retained), c is the probability that item information will be forgotten as well.

A characteristic of the Markov model is that these probabilities apply to each interval of time, Δt during the retention interval. Murdock further assumes that at the start of the retention interval, an item starts in state IO with probability p , and with probability $1 - p$ it starts in state I. The psychological rationale here might be that encoding an item at all establishes item information, and something more is required for the learner to encode serial order too. An attentive subject would never miss an item completely (Murdock assumes), so there is no probability that an item starts in state N. These starting state probabilities can be put in a vector, \mathbf{s} ,

$$\mathbf{s} = (p, 1 - p, 0)$$

- (a) Show that the probabilities that an item is in any of the states, IO, I, or N after 1 time interval, Δt is $(s'T)$.
- (b) Show that after 2 transition intervals, the probabilities are $s'T \times T = s'T^2$, and after 2 time intervals, they are $s'T^3$. [You may reason from the result of part (a), together with the diagram above, rather than multiplying the symbolic matrix entries.]
- (c) Assume parameter values of $p = 1.0$, $a = b = 0.1$, and $c = 0.5$. Find the probabilities that an item is in state IO, I, and N after 1, 2, and 3 transitions.

II-8. The matrix M below represents the performance of five people on three binary items, where $m_{ij} = 1$ indicates that item i is answered correctly by person j .

$$M = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 \end{bmatrix}$$

- (a) Calculate the product MM' . What do the diagonal elements represent? What do the off-diagonal elements represent?
- (b) Calculate the product $M'M$. What do the diagonal elements represent? What do the off-diagonal elements represent?
- II-9. In studies of perceptual identification, it is common to present stimuli (say, letters) in a background of noise and require the subject to identify the stimulus seen or heard. For n stimuli, the results are collected in an $n \times n$ matrix, X , where

$$x_{ij} = \text{Prob}(\text{subject says } j \mid \text{stimulus } i \text{ was presented})$$

In studying the relationships between stimuli in such a matrix, it is sometimes assumed that confusions are symmetric, $x_{ij} = x_{ji}$, but for some types of stimuli, the results are not symmetric.

- (a) Express the average proportion of correct responses as a function of the elements of X .
- (b) Show that an arbitrary square matrix, X can be expressed as the sum of two matrices,

$$X = S + A$$

where $S = (X + X')/2$ is symmetric, and A is a *skew-symmetric* matrix, i.e., $a_{ij} = -a_{ji}$.

- (c) Why might it be reasonable to regard A as containing “pure asymmetry” information and S as containing “pure symmetry” information?

[Psychological models for perceptual identification sometimes assume symmetry for errors, and sometimes do not. The decomposition of X allows both to be studied.]

III. Bases, orthogonality and vector spaces

III-1. On a sheet of graph paper,

- (a) Draw directed lines from the origin to represent the vectors,

$$\begin{array}{ll} \mathbf{a}' = (1, 0) & \mathbf{p}' = (4, 1) \\ \mathbf{b}' = (0, 1) & \mathbf{q}' = (2, 3) \end{array}$$

- (b) Draw on the graph directed lines which represent the vectors:

$$\begin{array}{l} \mathbf{r} = \mathbf{p} - \mathbf{q} \\ \mathbf{s} = \mathbf{q} - \mathbf{p} \\ \mathbf{t} = \mathbf{p} + \mathbf{q}/2 \end{array}$$

- (c) The vectors \mathbf{a} and \mathbf{b} form a basis for the two-dimensional vector space of the graph paper. Show that each of \mathbf{p} , \mathbf{q} , \mathbf{r} , \mathbf{s} , and \mathbf{t} can be expressed as linear combinations of \mathbf{a} , and \mathbf{b} of the form,

$$n \mathbf{a} + m \mathbf{b}$$

where n and m are real numbers.

- (d) The vectors \mathbf{p} and \mathbf{q} also form a basis for this vector space. Show that \mathbf{a} , and \mathbf{b} can be expressed as linear combinations of \mathbf{p} , and \mathbf{q} of the form, $u \mathbf{p} + v \mathbf{q}$ where u and v are real numbers.

III-2. Show that the vector $\mathbf{c} = (c_1, c_2, \dots, c_n)'$ is orthogonal to the vector $(1, 1, \dots, 1) \div \sqrt{n}$ if and only if $\sum_{i=1}^n c_i = 0$.

III-3. Consider the set of $(n - 1)$ n -dimensional vectors:

$$\begin{aligned} \mathbf{c}'_1 &= (1, -1, 0, 0, \dots, 0) \div \sqrt{2} \\ \mathbf{c}'_2 &= (1, 1, -2, 0, \dots, 0) \div \sqrt{6} \\ \mathbf{c}'_3 &= (1, 1, 1, -3, 0, \dots, 0) \div \sqrt{12} \\ &\dots \\ \mathbf{c}'_{n-1} &= (1, 1, 1, 1, \dots, 1, -(n-1)) \div \sqrt{n^2 - n} \end{aligned}$$

These vectors are referred to as *Helmert contrasts*. They are useful for comparing ANOVA factors whose levels are ordered, but not quantitative. Show that for $n = 5$ the vectors, $\mathbf{c}_1, \mathbf{c}_2, \dots, \mathbf{c}_{n-1}$:

- are all orthonormal (i.e., pairwise orthogonal and of unit length).
- are all orthogonal to the unit vector, \mathbf{j}_n .
- any other vector \mathbf{d} , which is orthogonal to the unit vector, \mathbf{j}_n . can be expressed as a linear combination of $\mathbf{c}_1, \mathbf{c}_2, \dots, \mathbf{c}_{n-1}$.

These three facts imply together that any other set of contrasts among n levels can be expressed in terms of $\mathbf{c}_1, \mathbf{c}_2, \dots, \mathbf{c}_{n-1}$. Isn't that nice?

III-4. Consider the $n \times n$ matrix

$$\mathbf{C} = \begin{bmatrix} 1-n & 1 & 1 & \dots & 1 \\ 1 & 1-n & 1 & \dots & 1 \\ \dots & \dots & \dots & \dots & \dots \\ 1 & 1 & 1 & \dots & 1-n \end{bmatrix}$$

- Show that $\text{rank}(\mathbf{C}) = n - 1$.
- Show that $\mathbf{C} \mathbf{j} = \mathbf{0}$. That is, \mathbf{j} is orthogonal to every row of \mathbf{C} , so \mathbf{C} is the orthogonal complement of \mathbf{j} .
- Let $\mathbf{x} = (x_1, x_2, \dots, x_n)'$ be the means of n treatment groups in a single factor experiment. Interpret $\mathbf{x}' \mathbf{j}$.
- Evaluate $\mathbf{C} \mathbf{x}$. Using the results of (a), (b), and (c) above, show that $\mathbf{C} \mathbf{x}$ is a vector of n contrasts among the group means, of which only $n - 1$ are independent. What do these contrasts represent?

III-5. If $\mathbf{c}_1, \mathbf{c}_2, \dots, \mathbf{c}_n$ are vectors, all orthogonal to a vector \mathbf{x} , show that any vector in the span of $\mathbf{c}_1, \mathbf{c}_2, \dots, \mathbf{c}_n$ is orthogonal to \mathbf{x} .

III-6. Show that any linear combination of a set of independent contrasts is also a contrast.

III-7. Given the three 4-dimensional vectors below, suppose we wish to find a unit-length vector, \mathbf{x} , which is orthogonal to each.

$$\begin{aligned} \mathbf{c}_1 &= (-1, -1, 0, 0)' \\ \mathbf{c}_2 &= (0, 2, -1, -1)' \\ \mathbf{c}_3 &= (3, -2, 0, -1)' \end{aligned}$$

- Show that \mathbf{x} must satisfy $\mathbf{c}'_1 \mathbf{x} = \mathbf{c}'_2 \mathbf{x} = \mathbf{c}'_3 \mathbf{x} = 0$.
- What additional equation must \mathbf{x} satisfy?
- Find \mathbf{x} .

III-8. Find two unit vectors, \mathbf{a} , and \mathbf{b} , which are each orthogonal to the vector, $\mathbf{m} = 1/\sqrt{3}(1, 1, 1)'$.

III-9. The general linear model, written in matrix notation, is given by

$$\mathbf{y} = \mathbf{X}\beta + \epsilon \quad (3)$$

where \mathbf{y} is an $n \times 1$ observation vector, \mathbf{X} is an $n \times p$ design matrix, β is an $p \times 1$ vector of unknown parameters, and ϵ is an $n \times 1$ random vector of residuals. For regression models, \mathbf{X} is typically a unit column vector followed by columns of predictor variables. For ANOVA models, \mathbf{X} is typically a unit vector followed by columns of indicator (0/1) variables.

- Consider a two-way 2×3 ANOVA design with $n = 1$ observation per cell (Elswick et al., 1991). If the model (without interaction) is given by

$$y_{ij} = \mu + \alpha_i + \beta_j + \epsilon_{ij}, \quad i = 1 : 2; j = 1 : 3 \quad (4)$$

and the vector of parameters is $\beta' = (\mu, \alpha_1, \alpha_2, \beta_1, \beta_2, \beta_3)$, find the design matrix \mathbf{X} so that the elements in y (a 6 by 1 vector) can be expressed in the form of Eqn. (3).

- By considering linear dependencies among the columns of \mathbf{X} , determine the rank of \mathbf{X} . Hazard a guess about the relation between rank and degrees of freedom.
- Find the row echelon form \mathbf{X}^* of \mathbf{X} (Try the ECHELON function in APL or SAS/IML). Does this confirm your observation of the rank of \mathbf{X} ?
- The product $\mathbf{X}^*\beta$ shows the linear combinations of the elements of β which can be estimated from the data. Find $\mathbf{X}^*\beta$ and interpret the result.

IV. Transformations, projections & quadratic forms

IV-1. Let $\mathbf{x} = (x_1, x_2)'$ be a 2-dimensional vector and consider a rotation in which the vector is rotated counter-clockwise about the origin through an angle of θ degrees, to a position, $\mathbf{x}^* = (x_1^*, x_2^*)'$, as shown in Figure 3. In scalar terms, the rotation can be expressed as

$$\begin{aligned} x_1^* &= x_1 \cos \theta - x_2 \sin \theta \\ x_2^* &= x_1 \sin \theta + x_2 \cos \theta \end{aligned}$$

- Find the matrix \mathbf{T} in the matrix expression of the above equations, $\mathbf{x}^* = \mathbf{T}\mathbf{x}$.
- Show that this matrix \mathbf{T} is orthonormal.

IV-2. Let $\mathbf{\Lambda}$ be a matrix of factor loadings of a set of tests with respect to one set of coordinate axes. If the coordinates are rotated, the matrix equation for expressing the matrix \mathbf{V} of loadings with respect to the rotated axes is $\mathbf{V} = \mathbf{\Lambda}\mathbf{T}$. Suppose there are four tests and two factors, and $\mathbf{\Lambda}$ is given by

$$\mathbf{\Lambda} = \begin{bmatrix} .52 & .30 \\ .35 & .20 \\ .40 & .69 \\ .30 & .52 \end{bmatrix}$$

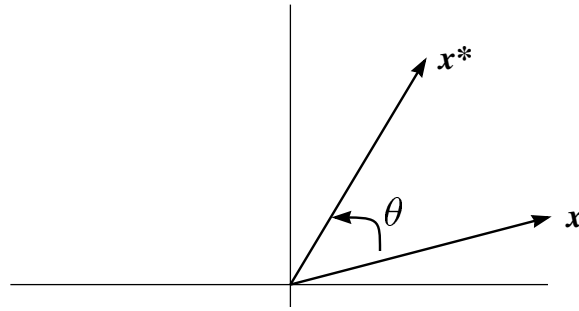


Figure 3: Rotation of a vector in 2-dimensional space

- Find a matrix T representing a transformation through a positive angle of 30° (counterclockwise).
- Evaluate the product, ΛT , rounding to two decimals. Note that some elements in $V = \Lambda T$ are (approximately) zero. The principle of *simple structure* in factor rotation seeks a rotation for which many elements of the rotated matrix become zero.

IV-3. The rotation procedure of factor analysis can be expressed as a set of linear transformations, i.e., as a matrix product. Let $V = \Lambda T$ represent such a linear transformation. The matrix Λ is a factor matrix satisfying $R = \Lambda \Lambda'$, where R is the correlation matrix V is the transformed factor matrix, and T is a matrix of coefficients which specifies the rotation.

- Find a condition on the matrix T such that the *transformed* factor matrix, V will also satisfy the factor equation.
- Show that the particular matrix T of problem I satisfies this condition, within rounding error.

IV-4. Consider a simple two-dimensional example in which three tests have these loadings on two underlying factors:

	Factor1	Factor2
Test 1	0.50	0.30
Test 2	0.30	0.20
Test 3	-0.40	0.70

Each test can be represented as a point in the plane of Factor1 (x-axis) and Factor2 (y-axis). Plot these test points, and find their transformed coordinates when the coordinate axes are rotated through a positive angle of 37° . Use $\sin 37^\circ = 0.60$, $\cos 37^\circ = 0.80$.

IV-5. Let $\mathbf{x} = (x_1, x_2)'$, and $\mathbf{A} = \begin{bmatrix} 2 & 2 \\ 4 & -2 \end{bmatrix}$, $\mathbf{B} = \begin{bmatrix} 1/2 & 0 \\ 0 & 1 \end{bmatrix}$ Expand the following quadratic forms:

- $\mathbf{x}' \mathbf{A} \mathbf{x}$
- $\mathbf{x}' \mathbf{B} \mathbf{x}$
- Describe in words the difference between the expansions in (a) and (b).

Hint: Multiply out, then repeat: {substitute a trial value for x and solve for y }.

IV-6. Describe (or plot) the set of points with coordinates (x, y) that satisfy the matrix equations:

-

$$\begin{pmatrix} x & y \end{pmatrix} \begin{bmatrix} 4 & 0 \\ 0 & 5 \end{bmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 60$$

(b)

$$\begin{pmatrix} x & y \end{pmatrix} \begin{bmatrix} 4 & 3 \\ 3 & 5 \end{bmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 60$$

IV-7. Evaluate the quadratic forms, $\mathbf{u}' \mathbf{A} \mathbf{u}$, where $\mathbf{u} = (x, y, 1)'$ and where the matrix \mathbf{A} is:

$$(a) \begin{bmatrix} 0 & 0 & -2a \\ 0 & 1 & 0 \\ -2a & 0 & 0 \end{bmatrix} \quad (b) \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 4 \end{bmatrix} \quad (c) \begin{bmatrix} a & 0 & b \\ 0 & 0 & -1 \\ b & -1 & c \end{bmatrix} \quad (d) \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & -4 \end{bmatrix}$$

What functions would be obtained by setting each of these quadratic forms equal to 0?

IV-8. The next few problems demonstrate how some of the common expressions for sums of squares can be represented by quadratic forms. Suppose n observations are represented by a vector $\mathbf{x} = (x_1, x_2, \dots, x_n)'$ and the sample variance is based on the sum of squared deviations,

$$SS = \sum_{i=1}^n (x_i - \bar{x})^2 = \sum_{i=1}^n x_i^2 - n\bar{x}^2$$

Show that

- (a) $\bar{x} = (1/n) \mathbf{x}' \mathbf{j}$.
- (b) $\sum x_i^2 = \mathbf{x}' \mathbf{x}$.
- (c) $n\bar{x}^2 = (1/n) \mathbf{x}' \mathbf{J} \mathbf{x} = (1/n) \mathbf{x}' \mathbf{j} \mathbf{j}' \mathbf{x}$, where \mathbf{J} is the $n \times n$ matrix with all elements equal to 1.
- (d) Hence, show that SS can be expressed by the quadratic form,

$$SS = \mathbf{x}' (\mathbf{I} - \mathbf{J}/n) \mathbf{x}$$

- (e) Write out explicitly the matrix $(\mathbf{I} - \mathbf{J}/n)$ for $n = 3$. [Source: Searle (1966)]

IV-9. ANOVA and regression tests are based on breaking up sums of squares into independent, additive portions attributable to various sources of variance. This problem demonstrates some of the properties of the quadratic forms associated with these sums of squares.

Consider the well-known identity for a set of test scores, $y_i, i = 1, \dots, n$:

$$\sum_i^n y_i^2 = n\bar{y}^2 + \sum_i^n (y_i - \bar{y})^2 \quad (5)$$

In the previous problem, it was shown that the sum of squared deviations term could be written as a quadratic form,

$$\mathbf{y}' (\mathbf{I} - \mathbf{J}/n) \mathbf{y} = \mathbf{y}' \mathbf{A}_2 \mathbf{y} \quad (\text{say})$$

- (a) Show that the above identity (5) can be written as

$$\mathbf{y}' \mathbf{A}_0 \mathbf{y} = \mathbf{y}' \mathbf{A}_1 \mathbf{y} + \mathbf{y}' \mathbf{A}_2 \mathbf{y}$$

where $\mathbf{A}_0, \mathbf{A}_1, \mathbf{A}_2$ are $n \times n$ symmetric. Find \mathbf{A}_0 and \mathbf{A}_1 .

- (b) Show that the following properties hold:

$$\begin{aligned} \mathbf{A}_0 &= \mathbf{A}_1 + \mathbf{A}_2 \\ \mathbf{A}_1 &= \mathbf{A}_1^2 \\ \mathbf{A}_2 &= \mathbf{A}_2^2 \\ \mathbf{A}_1 \mathbf{A}_2 &= 0 \end{aligned}$$

(c) Show that:

$$\begin{aligned} \text{tr}(\mathbf{A}_0) &= n = \text{rank}(\mathbf{A}_0) \\ \text{tr}(\mathbf{A}_1) &= 1 = \text{rank}(\mathbf{A}_1) \\ \text{tr}(\mathbf{A}_2) &= n - 1 = \text{rank}(\mathbf{A}_2) \end{aligned}$$

Matrices such as \mathbf{A}_1 and \mathbf{A}_2 with these properties are called **idempotent projection matrices**. Every sum of squares formula such as (5) in ANOVA and regression can be represented as a sum of quadratic forms with idempotent projection matrices.

(d) Evaluate the expression,

$$\frac{\mathbf{y}' \mathbf{A}_1 \mathbf{y} / 1}{\mathbf{y}' \mathbf{A}_2 \mathbf{y} / n - 1}$$

for $\mathbf{y} = (2, 3, 3, 3, 5, 8)'$. Interpret this expression and state what statistical test it is used in.

IV-10. In psychological experiments, response classes are often defined arbitrarily. As a result, an experimenter may wish to combine response classes when he analyzes data. In formal models it is useful to have a representation of this procedure of combining response classes.

Bush et al. (1954) use a *projection matrix* to combine responses. An example is the matrix,

$$\mathbf{P} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

This matrix is used as follows: Let there be three response classes initially, with probabilities, r_1 , r_2 , and r_3 , in a column vector, \mathbf{r} .

- Postmultiply \mathbf{P} by \mathbf{r} , and describe verbally the result.
- Show that the matrix \mathbf{P} satisfies $\mathbf{P}^2 = \mathbf{P}$.

IV-11. Guttman (1944) discusses a method of factoring a correlation matrix, based on the equation,

$$\mathbf{f} = \mathbf{R}\mathbf{x}(\mathbf{x}'\mathbf{R}\mathbf{x})^{-1/2} \quad (6)$$

where \mathbf{f} is one column of a factor matrix, \mathbf{F} ; \mathbf{R} is the correlation matrix; and \mathbf{x} is a column vector containing arbitrarily selected coefficients. Once a single factor has been found using this method, the same formula can be applied to the residual correlation matrix, $\mathbf{R}_1 = \mathbf{R} - \mathbf{f}\mathbf{f}'$, to find a second factor, and so forth. A factor matrix \mathbf{F} which is build up in this way will always satisfy the basic equation of factor analysis, $\mathbf{F}\mathbf{F}' = \mathbf{R}$.

Suppose that the matrix of intercorrelations on four tests is

$$\mathbf{R} = \begin{bmatrix} .64 & .56 & .48 & .40 \\ .56 & .49 & .42 & .35 \\ .48 & .42 & .36 & .30 \\ .40 & .35 & .30 & .25 \end{bmatrix}$$

The elements on the main diagonal of \mathbf{R} are not 1.0 as would be ordinarily expected, but are values called *communalities*. representing the variance of each test which is shared with the other tests.

- Apply Guttman's formula (6) to the matrix \mathbf{R} to find a single column \mathbf{f} of the factor matrix. Assume \mathbf{x} to be the vector $\mathbf{x} = (1, 1, 1, 1)'$
- Show that in this case, $\mathbf{f}\mathbf{f}' = \mathbf{R}$, or equivalently, $\mathbf{R}_1 = \mathbf{R} - \mathbf{f}\mathbf{f}' = \mathbf{0}$. In other words, the only one factor can account for all the correlations in \mathbf{R} .

[Congradualtions! You've just discovered the basis for factor analysis.]

IV-12. Resolve the column vector $\mathbf{x} = (x_1, x_2, x_3)'$ into two components, \mathbf{x}_a and \mathbf{x}_b , such that

- (a) $\mathbf{x} = \mathbf{x}_a + \mathbf{x}_b$
- (b) \mathbf{x}_a is parallel to the vector $(1, 1, 1) / \sqrt{3}$
- (c) \mathbf{x}_b is perpendicular (orthogonal) to $(1, 1, 1) / \sqrt{3}$

[And now, the Gram-Schmidt method for transforming correlated variables into orthongonal ones!]

IV-13. [Arbuckle and Friendly \(1977\)](#) consider the problem of transforming a vector of factor loadings, \mathbf{x} to a vector \mathbf{y} so that \mathbf{y} is as *smooth* as possible in the sense that the sum of squares of successive differences, $\sum_{i=1}^{n-1} (y_i - y_{i+1})^2$ is as small as possible.

- (a) Find an $(n - 1) \times n$ matrix \mathbf{P} such that

$$\mathbf{P} \mathbf{y} = \begin{bmatrix} y_1 - y_2 \\ y_2 - y_3 \\ \vdots \\ y_{n-1} - y_n \end{bmatrix} = \mathbf{d} \text{ (say)}$$

- (b) Show that the sum of squares of successive differences referred to above is given by $\mathbf{d}' \mathbf{d}$.
- (c) If \mathbf{y} is obtained as a linear combination of \mathbf{x} , by $\mathbf{y} = \mathbf{T} \mathbf{x}$, find an expression not using parentheses to express the sum of squared successive differences in terms of \mathbf{T} , \mathbf{x} , and \mathbf{P} .

V. Determinants, inverse, and rank

V-1. Find the inverse of

$$\mathbf{A} = \begin{bmatrix} 5 & 3 \\ 3 & 2 \end{bmatrix}$$

V-2. (a) Verify that the inverse of

$$\begin{bmatrix} 3 & 0 \\ 0 & 1/5 \end{bmatrix} \text{ is } \begin{bmatrix} 1/3 & 0 \\ 0 & 5 \end{bmatrix}$$

- (b) What is the inverse of

$$\begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}$$

Can you determine what the inverse of any diagonal matrix is?

V-3. If

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

show that $|\mathbf{A}| = a_{11}a_{22} - a_{21}a_{12} = a_{11}(a_{22} - a_{21}a_{11}^{-1}a_{12})$. This last expression has an important analog for partitioned matrices.

V-4. A frequent question in the testing field is whether to add another test to a battery of n tests to attempt to increase the prediction of an external criterion. [Horst \(1951\)](#) proposes a time-saving method of answering this question. He derives a formula for the validity of an additional test, k , that would be required to increase the validity of the battery by an amount a .

Let \mathbf{R} be the $(n + 1) \times (n + 1)$ matrix of intercorrelations, partitioned with test k as the last row and column,

$$\mathbf{R} = \begin{bmatrix} \mathbf{R}^* & \mathbf{r}_k \\ \mathbf{r}'_k & 1 \end{bmatrix}$$

where \mathbf{R}^* is the $n \times n$ matrix of correlations of all tests except test k , and \mathbf{r}_k is the column vector of correlations of test k with each other test.

To develop his formula, Horst needed to find the inverse of \mathbf{R} , that is a matrix \mathbf{S} , such that $\mathbf{R}\mathbf{S} = \mathbf{I}$, the identity matrix. Thus, $\mathbf{S} = \mathbf{R}^{-1}$.

(a) Partition the matrix \mathbf{S} conformably with \mathbf{R} , as

$$\mathbf{S} = \begin{bmatrix} \mathbf{S}^* & \mathbf{u} \\ \mathbf{u}' & d \end{bmatrix}$$

and write out the equation $\mathbf{R}\mathbf{S} = \mathbf{I}$ in terms of partitioned matrices.

(b) Write out the four sub-matrix equations implied by the equation in part (a).

(c) Find $\mathbf{S} = \mathbf{R}^{-1}$ by solving for \mathbf{S}^* , \mathbf{u} , and d .

V-5. If a matrix is upper (lower) triangular, then its inverse is also upper (lower) triangular. Show that this is true for the matrix \mathbf{S} below by finding its inverse.

$$\mathbf{S} = \begin{bmatrix} 1 & 2 & -1 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{bmatrix}$$

Hint: Let

$$\mathbf{S}^{-1} = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$$

and write out the 9 scalar equations implied by $\mathbf{S}\mathbf{S}^{-1} = \mathbf{I}$.

V-6. In a ratio scaling experiment four stimuli are presented in all possible pairs to the subject. For each pair, the subject is required to give a number representing the relative “strength” of the first stimulus compared to the second. For example, he/she may be asked to judge the relative brightness of two lights, or which of two tones is higher in pitch.

One hypothesis for this type of judgment is that the subject has internal subjective intensity values for each stimulus, and to make a judgment, “computes” and reports the ratio of the two values for a pair.

Let the matrix $\mathbf{R} = \{r_{ij}\}$ contain the judgments of each pair, for stimulus i relative to stimulus j , and let $\mathbf{s} = (s_1, \dots, s_n)'$ be a vector of the subject’s internal scale values for these stimuli.

(a) Express the hypothesis, $r_{ij} = s_i \div s_j$ in matrix terms.

(b) What does this hypothesis imply about the rank of \mathbf{R} if the model holds?

(c) What does this imply about the latent roots of $\mathbf{R}'\mathbf{R}$?

V-7. In the randomized block design, each of n subjects is given each of a series of p experimental treatments, resulting in a score, y_{ij} = score of subject i under treatment j . With one observation per cell, the standard analysis of variance model (without interaction) is,

$$y_{ij} = \alpha_i + \beta_j \quad (+\epsilon_{ij})$$

(a) Write $\mathbf{Y}(n \times p) = y_{ij}$ and express \mathbf{Y} in terms of α_i and β_j as a *product* of two matrices.

- (b) Tukey's "one degree of freedom for non-additivity" adds an additional term to this model, giving:

$$y_{ij} = \alpha_i + \beta_j + g \alpha_i \beta_j \quad (+\epsilon_{ij})$$

where g is one additional parameter to be estimated. Write a matrix *product* expression for \mathbf{Y} in this model. [You may not simply multiply \mathbf{Y} by the identity matrix. Use no + signs in the elements in the factors in your product.]

- (c) What is the rank of \mathbf{Y} if the model of part (a) holds?
 (d) What is the rank of \mathbf{Y} if the model of part (b) holds?

- V-8. Suppose an animal running a maze is scored under 3 experimental conditions for (a) number of wrong turns made, (b) number of pauses, and (c) time to reach the goal. Some data is given below, with scores expressed as deviations from the column means.

	Errors	Pauses	Time
Cond. 1	2	0	4
Cond. 2	1	6	10
Cond. 3	-3	-6	-14

- (a) Determine the rank of this matrix.
 (b) If time and errors have been measured, does the pause measure contribute any additional information?
 (c) If \mathbf{e} is the (column vector) number of errors, \mathbf{p} is the number of pauses, and \mathbf{t} the time taken, can you find a constant a to make the following equation true for these data?

$$\mathbf{t} = a \mathbf{e} + \frac{4}{3} \mathbf{p}$$

VI. Simultaneous equations

- VI-1. On a sheet of graph paper draw lines representing each of the three linear equations,

$$x - y = 0 \quad (1)$$

$$3x - 2y = 1 \quad (2)$$

$$4x - 2y = 3 \quad (3)$$

- (a) Show that this system of equations has no solution and interpret this fact with reference to your geometrical interpretation.
 (b) Suppose we adopt the following procedure for "solving" the system. Take the equations and solve them in pairs— (1) and (2), (1) and (3), (2) and (3). This will give values, say (x_1, y_1) , (x_2, y_2) , and (x_3, y_3) which may not agree. "Solve" the equations in this way and locate the solution points on your graph.
 (c) Now, suppose we adopt as the final solution to this system the average value,

$$(\bar{x}, \bar{y}) = \left(\frac{(x_1 + x_2 + x_3)}{3}, \frac{(y_1 + y_2 + y_3)}{3} \right)$$

Locate this point on your graph and interpret geometrically. [This is one definition of a 'best approximate solution.]

- VI-2. Show that the system of equations,

$$\begin{aligned} x &= 1 - y \\ 2x + 2y &= 2 \end{aligned}$$

is underdetermined. That is, there is no *unique* pair (x, y) which satisfy both equations.

- (a) Draw these two equations as lines on a graph and explain why there is no unique solution.
 (b) If the second equation is changed to $2x + 2y = 3$, explain why these two equations become inconsistent.

VII. Latent roots and vectors

Multiplication of a vector \mathbf{v} by a square matrix \mathbf{A} can be regarded as a mapping or transformation of \mathbf{v} into some other vector, \mathbf{v}^* in the vector space,

$$\mathbf{A} \mathbf{v} = \mathbf{v}^*$$

A *latent vector*, or *eigenvector* of the matrix \mathbf{A} is a special vector whose direction is unchanged by that transformation. That is, if \mathbf{v} is a non-zero vector such that

$$\mathbf{A} \mathbf{v} = k \mathbf{v}$$

then \mathbf{v} is an eigenvector or latent vector of \mathbf{A} , and k , the constant of proportionality, is the eigenvalue or latent root corresponding to \mathbf{v} .

VII-1. In Householder's method for determining eigenvalues of a real symmetric matrix, \mathbf{A} , the matrix is subjected to a series of orthogonal transformations of the form

$$\mathbf{P} = \mathbf{I} - 2 \mathbf{w} \mathbf{w}' \quad \text{where } \mathbf{w}' \mathbf{w} = 1$$

- Show that \mathbf{P} is orthonormal and symmetric.
- If the latent roots of \mathbf{A} are $\lambda_1, \lambda_2, \dots, \lambda_n$, what are the latent roots of $\mathbf{P} \mathbf{A} \mathbf{P}$?
- If the latent vectors of \mathbf{A} are $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$, what are the latent vectors of $\mathbf{P} \mathbf{A} \mathbf{P}$?

VII-2. The eigenvalues and eigenvectors of any square, symmetric matrix \mathbf{A} ($n \times n$) can (in principle) be found by (1) finding the n values of λ which satisfy $|\mathbf{A} - \lambda \mathbf{I}| = 0$; (2) solving $\mathbf{A} \mathbf{v} = \lambda \mathbf{v}$ for \mathbf{v} using each value of λ determined by step (1). Using this method, find the eigenvalues and eigenvectors of the following matrices:

$$(a) \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \quad (b) \begin{bmatrix} 0 & -1 \\ 1 & -1 \end{bmatrix}$$

VII-3. Show that the eigenvalues of the matrix $\begin{bmatrix} \cos \theta & \sin \theta \\ \sin \theta & -\cos \theta \end{bmatrix}$ are ± 1 .

VII-4. In the Markov model for serial learning outlined in Problem II.7 verify that the eigenvalues of the transition matrix, \mathbf{T} are

Hint: Solve the characteristic equation $|\mathbf{T} - \lambda \mathbf{I}| = 0$.

$$\lambda = \begin{cases} 1 \\ 1 - c \\ 1 - a \end{cases}$$

VII-5. A *magic square* is a matrix in which the elements of each row, column and the two main diagonals add up to the same number, this number being called the *magic number*. For the magic square, \mathbf{A} , below, show that one latent root of this magic square is its magic number. (This is true for all magic squares).

$$\mathbf{A} = \begin{pmatrix} 8 & 1 & 6 \\ 3 & 5 & 7 \\ 4 & 9 & 2 \end{pmatrix}$$

References

- Arbuckle, J. and Friendly, M. L. On rotating to smooth functions. *Psychometrika*, 42:127-140, 1977. 13
- Bush, R. R., Mosteller, F., and Thompson, G. L. A formal structure for multiple choice situations. In Thrall, C. . D., editor, *Decision Processes*. Wiley, NY, 1954. 10

- Cliff, N. Adverbs multiply adjectives. *Psychological Review*, 66:27–44, 1959. 6
- Elswick, R. K. J., Gennings, C., Chinchilli, V. M., and Dawson, K. S. A simple approach for finding estimable functions in linear models. *The American Statistician*, 45:51–53, 1991. 9a
- Guttman, L. General theory and methods of matrix factoring. *Psychometrika*, 9:1–16, 1944. 11
- Horst, P. The relation between the validity of a single test and its contribution to the predictive efficiency of a test battery. *Psychometrika*, 16:57–66, 1951. 4
- Judd, D. B. Basic correlates of the visual stimulus. In Stevens, S. S., editor, *Handbook of Experimental Psychology*. Wiley, New York, 1951. 4
- Kendall, D. G. Some problems and methods in statistical archaeology. *World Archaeology*, 1:61–, 1969. 6
- Murdock, B. B. J. Item and order information in short-term serial memory. *Journal of Experimental Psychology: General*, 105:191–216, 1976. 7
- Searle, S. R. *Matrix Algebra for the Social Sciences*. Wiley, New York, 1966. 8e
- Van de Geer, J. P. *Introduction to Multivariate Analysis for the Social Sciences*. W H. Freeman., San Francisco, 1971. 3c
- Williams, G. Mathematics in archaeology. *Mathematics Teacher*, pp. 56–58, 198? 6