OUTLINE OF MATRIX ALGEBRA
by John Imbrie

1. Exhortation.

It is strongly urged that you review or learn the elements of matrix algebra. Matrix notation will be briefly explained in and used at our lectures, for only in this way can one understand the essential relationships between many mathematical models of use in geology. It is, fortunately, a very simple subject -- a language for expressing relationships between sets of numbers arranged in tables. As with any language, you must practice a bit with pencil and paper in order to feel at home in it. To this end there are included below a number of simple exercises for you to perform on your own. Remember Edith Hamilton's comment on American education: "Is hard work prominent? The world of thought can be entered in no other way."

2. Value of matrix algebra.

a. Its symbolism allows large tables of data and of derived parameters to be represented very compactly.

b. Its operations (addition and multiplication, inversion, etc.) permit complex relationships and concepts to be seen which otherwise would be obscure.

c. Most computer programs are set up to treat problems by matrix methods.

3. Definitions: Basic Matrix Symbols

A matrix is a fancy name for any complete rectangular array of numbers (or letter symbols for numbers). It may have any number of rows and columns. Specification of the number of rows and columns, always in that order, is the order of a matrix. Any number in the matrix is referred to as an element.

An entire matrix can be symbolized, as it frequently will be in these notes, by a single capital letter. Thus the matrices below (both of order 2 by 3) can be completely written down in square brackets, or defined as A and B.

$$A = \begin{bmatrix} 1 & 2 & 7 \\ 0 & 3 & -2 \end{bmatrix} \qquad B = \begin{bmatrix} 2 & 7 & 0 \\ 1 & 2 & 3 \end{bmatrix}$$

If desired, the order of a matrix may be specified thus: $A_{(2,3)}$. If it is desired to specify in the general case the order of a matrix, subscripts n, N, or m are commonly used. For example, a matrix with N rows and n columns may be written $A_{(N,n)}$.

In print it is common to put matrix symbols in bold-face type. Another system of symbols should be learned in which a lower-case letter with two subscripts, the first representing the rows and the second the columns, represents a matrix. For these row and column indices the following letters will be used: i, j, k, and p. As an example, the 2 by 3 matrix A above has N = 2 and n = 3. Using lower-case symbols, $A = a_{ij}$ where it is understood that the i index runs from 1 through 2 and the j index from 1 through 3.

Sometimes square brackets are placed around the lower case letter to indicate a matrix. Thus $\left[a_{ij}\right]$ is a matrix and a_{ij} is considered its general element. Putting actual numbers into the i and j positions we can write out the entire matrix symbolically as:

$$A = a_{ij} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix}$$

In the example on Page 2 above, the element $a_{11} = 1$; $a_{13} = 7$, etc. (refers to A on preceding page).

The general matrix A with all elements indicated will look like this:

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1j} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2j} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{i1} & a_{i2} & \cdots & a_{ij} & \cdots & a_{in} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{N1} & a_{N2} & \cdots & a_{Nj} & \cdots & a_{Nn} \end{bmatrix}$$

($i = 1, 2 \ldots i \ldots N$; $j = 1, 2 \ldots j \ldots n$).

4. Transpose of a matrix.

Given a matrix A, we may construct another matrix A' having as its first row the first column of A; its second row the second column of A, etc. A' is said to be the transpose of A. Given A of section 3 above:

$$A = \begin{bmatrix} 1 & 2 & 7 \\ 0 & 3 & -2 \end{bmatrix} \qquad A' = \begin{bmatrix} 1 & 0 \\ 2 & 3 \\ 7 & -2 \end{bmatrix}$$

Verify for yourself that the transpose of A' = A, i.e., that (A')' = A.

5. <u>Standard matrices</u>.

It is convenient to use the same symbols always for the same kind of matrix. For example, a <u>data matrix</u> is defined as a matrix representing the raw data in a problem. We will always write this as $X_{(N,n)}$ where each of the N rows represents observations made at a locality, or on a sample, individual, specimen, etc. Each of the n columns represents a set of observations on some geological <u>variable</u> over all N samples. It will normally be a <u>rectangular matrix</u>, with more rows than columns. Other standard matrix symbols will be defined as we go along.

6. <u>Vectors</u>.

A vector is defined as a matrix with one column or one row. In print it may be symbolized as a lower-case letter in bold-face type. It may also be symbolized as a lower-case letter with a single subscript such as i, j, or k.

The <u>size</u> of the vector is defined as the number of elements in it. Sometimes, to avoid confusion, a vector may be indicated by placing a tilde $\sim$ over the letter. Thus in the example of section 3 above, $\tilde{a}_j$ represents any column vector in A. For example, vector

$$\tilde{a}_2 = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$$

Note that a vector is <u>always</u> assumed to be a column vector when no transpose sign is given. Thus in the same example, the first row vector in A would be indicated as:

$$\tilde{a}_1' = \begin{bmatrix} 1 & 2 & 7 \end{bmatrix}$$

Given a factor such as $\tilde{v}' = \begin{bmatrix} 1 & 6 & 7 & 0 \end{bmatrix}$ we may transpose it as

$$\tilde{v} = \begin{bmatrix} 1 \\ 6 \\ 7 \\ 0 \end{bmatrix}$$

1\. <u>Exercise</u>: Given matrix B in section 3 above, write out the symbols and the complete vector in brackets for all the 5 vectors contained in B.

7\. <u>Matrix addition and subtraction</u>.

Any two matrices <u>of the same order</u> may be added or subtracted, by adding or subtracting the corresponding elements. Thus in section 3 above we may write C = A + B, where

$$C = \begin{bmatrix} 3 & 9 & 7 \\ 1 & 5 & 1 \end{bmatrix}$$

2\. <u>Exercise</u>: Verify this, then calculate F, where F = A-B.

8\. <u>Special kinds of matrices</u>.

Certain important types occur frequently, have important properties, and have special names or symbols. Among these are:

a. <u>The Null Matrix</u>. A matrix with all elements equal to zero. Written as 0; or $0_{(N,n)}$ if the order is not clear from the context. The null matrix plays the role of zero in matrix algebra.

3\. <u>Exercise</u>: Write out $0_{(5,10)}$

3b <u>Exercise</u>: Let A be the matrix defined in section 3; find a matrix Q such that A + Q = 0.

b. A <u>square matrix</u> has the same number of rows as columns. A <u>rectangular matrix</u> does not.

c. A <u>diagonal matrix</u> is a square matrix with zero in every position except along the <u>principal diagonal</u> (upper left to lower right). Some form of the letter D is used to indicate most diagonal matrices, as Greek Delta. A typical diagonal matrix of order 3 might be:

$$D = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 7 & 0 \\ 0 & 0 & 4 \end{bmatrix}$$

4. <u>Exercise</u>: Show that $D' = D$.

d. The <u>identity matrix</u> is a diagonal matrix with all elements equal to 1. Always symbolized by I; or I_n if the size is not clear from the context. It plays the role of unity in matrix algebra.

5. <u>Exercise</u>: Using D defined above, write out a matrix K where

$$K = I + I + D.$$

e. A symmetrical matrix is a square matrix symmetrical about the principal diagonal, i.e., with every element $a_{ij} = a_{ji}$. For example:

$$R = \begin{bmatrix} 2 & 3 & 0 & 0 \\ 3 & 7 & 4 & 1 \\ 0 & 4 & 1 & 2 \\ 0 & 1 & 2 & 3 \end{bmatrix}$$

6. <u>Exercise</u>: Show that $R' = R$.

f. A scalar is a matrix with only one row and one column, for example: $[6]$ Viewed in this rather droll way, a number is a special case of a matrix with one row and column.

Similarly, we shall see that ordinary processes of multiplication and division, as well as addition and subtraction, are special cases of more complex and fundamental matrix processes.

9. Multiplication of matrices.

Up to now you may be wondering what all the fuss has been about, for the definitions are simple and the processes of addition and subtraction are simply large-scale applications of ordinary arithmetic. In this section, however, you will catch the first glimpse of the power of matrix concepts. Press On!

a. Multiplication of a matrix by a number: Given a number N and a matrix A we may define the product of the two as another matrix (say, P) in which each element $P_{ij} = Na_{ij}$. In other words, each element of A has been multiplied by N. As an example, if C is defined as in section 7 above:

$$C = \begin{bmatrix} 3 & 9 & 7 \\ 1 & 5 & 1 \end{bmatrix} \qquad 2C = \begin{bmatrix} 6 & 18 & 14 \\ 2 & 10 & 2 \end{bmatrix} \qquad C/2 = \begin{bmatrix} 1.5 & 4.5 & 3.5 \\ 0.5 & 2.5 & 0.5 \end{bmatrix}$$

7. Exercise: Write out 3C, and 3C/2.

b. Multiplication of a vector by a vector: Two vectors of the same size may be multiplied to form a scalar product if the first member of the pair is a row vector and the second is a column vector. Given two row vectors: $\tilde{a}' = \begin{bmatrix} a_1 & a_2 & \dots & a_n \end{bmatrix}$ and $\tilde{b}' = \begin{bmatrix} b_1 & b_2 & \dots & b_n \end{bmatrix}$ for example, we may form the product $\tilde{a}'\tilde{b}$ or $\tilde{b}'\tilde{a}$. The numerical results of the two operations will be the same, as will be clear from the definition. Multiplication of a row vector by a column vector, in that order, is called the <u>scalar product</u> of the two vectors.

The multiplication $\tilde{a}'\tilde{b}$ is defined as

$$a_1b_1 \quad + \quad a_2b_2 \quad + \quad \ldots \quad + \quad a_nb_n$$

Note that the scalar product of two vectors is, oddly enough, a <u>number</u>; it may be described as the <u>sum of products</u> of corresponding elements in the vectors.

In the numerical example just below, observe the spatial arrangement of the vectors with respect to their product. If you will now, henceforth, and forevermore arrange things in this way when thinking of vector multiplication or actually performing one, you will greatly speed your comprehension of more complex operations to come.

Given $\tilde{c}' = \begin{bmatrix} 1 & 3 & -4 & 2 \end{bmatrix}$ and $\tilde{d}' = \begin{bmatrix} 2 & 0 & 1 & 3 \end{bmatrix}$ the product $\tilde{c}'\tilde{d} = 4$. This should be visualized as shown on the left; calculations given on the right.

$$\begin{array}{cc} & \begin{bmatrix} 2 \\ 0 \\ 1 \\ 3 \end{bmatrix} \\ \begin{bmatrix} 1 & 3 & -4 & 2 \end{bmatrix} & \begin{bmatrix} 4 \end{bmatrix} \end{array} \qquad (2)(1) \; + \; (0)(3) \; + \; (-4)(1) \; + \; (2)(3) \; = \; 4.$$

8. <u>Exercise</u>: Show that $\tilde{d}'\tilde{c} = \tilde{c}'\tilde{d}$.

c. "Squaring" a vector: Any vector $\tilde{c}'$ can be multiplied by its own transpose $\tilde{c}$ according to the rule just given to from the product $\tilde{c}'\tilde{c}$. This product will be the sum of the squares of the elements in the vector.

9. <u>Exercise</u>: Verify this relationship by calculating $\tilde{c}'\tilde{c}$ for the vector $\tilde{c}'$ given just above. If you carry out this difficult task carefully, you will find that $\tilde{c}'\tilde{c} = 30$.

d. Multiplication of two matrices: Two matrices can be multiplied to form a product matrix provided that the number of columns of the first is equal to the number of rows of the second. That is, given $A_{(N,n)}$ and $B_{(n,m)}$, they can be multiplied in that order to form a product matrix $C_{(N,m)}$.

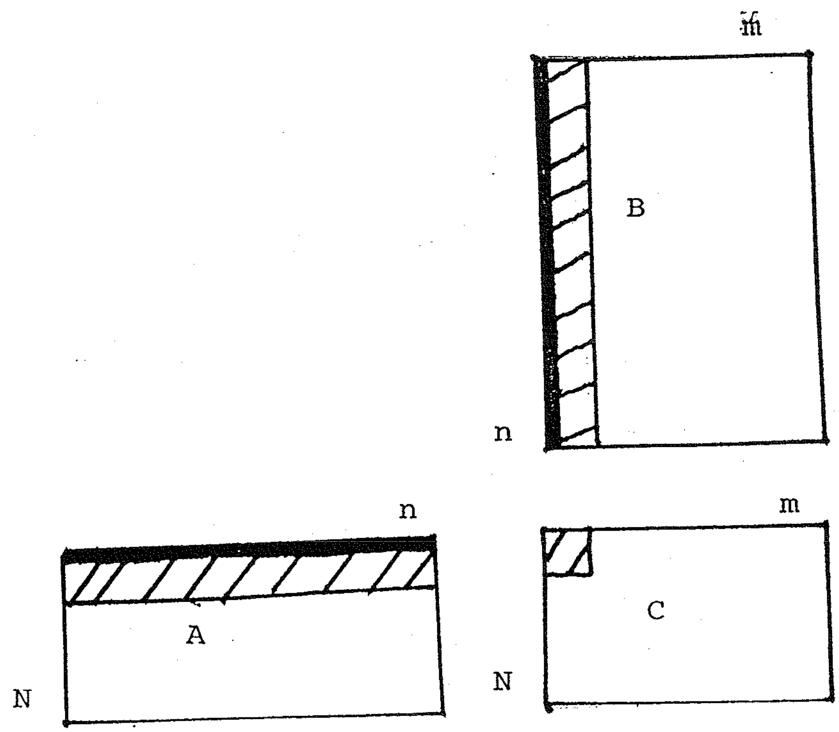
This operation can be written:

$$A_{(N,n)} \; B_{(n,m)} = C_{(N,m)}.$$

or because the requirements for the two sides to be the same size is understood, we can usually write simply:

$$A \; B = C$$

Forming the product matrix: Every element in C is a scalar product of two vectors. To form the product matrix, always arrange the matrices to be multiplied as shown below:



$$C = AB$$

To form the first element (c_{11}) in the first column of C, multiply the first column vector of B by the first row vector of A -- i.e., perform a scalar product $\tilde{a}_1'\tilde{b}_1$. This vector product is symbolized in the diagram by the shaded areas. To form each remaining element in the first column vector of C, multiply the same volumn vector of B by each of the row vectors in A in turn. Then move on to the second column of B, and do the same; and so on until the end. You will find, perhaps to your surprise, that this peculiar definition of multiplication represents many important geological, geometric, and algebraic relationships.

10. <u>Exercise</u>: Given A from section 3 above, and B defined as:

$$B = \begin{bmatrix} 2 & 1 & 0 & 3 \\ 3 & 2 & 1 & 2 \\ 6 & 1 & 0 & 4 \end{bmatrix}$$

Verify that

$$C = \begin{bmatrix} 50 & 12 & 2 & 35 \\ -3 & 4 & 3 & -2 \end{bmatrix}$$

Note that in matrix multiplication the <u>order</u> of the items multiplied is important and must not be interchanged. Thus, in the above example, the product BA does not exist.

11. <u>Exercise</u>: Try to form BA.

Definition: Given a matrix product AB, B is said to be <u>pre-multiplied</u> by A; and A is said to be <u>postmultiplied</u> by B.

e. Multiplication of more than two matrices: Given a triple product such as ABD to perform, any two adjacent matrices are multiplied first and the product of that operation is multiplied by the remaining matrix. We may either form the product AB first, and then postmultiply the result by D; or take BD and then premultiply by A. The result is identical.

12. Exercise: Construct a matrix D from the first two rows of R given under section 8e above. Taking A from Page 2 and B from Page 10 form ABD'. Do this in both alternate ways -- i.e., (AB)D' and A(BD') -- and verify that the results are the same. NOTE: If you have persevered through this exercise, congratulations! This is the most difficult problem you will have, and qualifies you for an M.Sc. (Matrix Scientist), second class. Press On! New vistas will soon open.

f. Transpose of matrix products. Given X = AB, X' = B'A'. In general, the transpose of a matrix composed of a set of matrix products is the product of the constituent matrices written in transposed form and in reverse order. For example, if X = ABC, X' = C'B'A'.

13. Exercise: Take three small matrices A, B, and C and form X where X = ABC. Then show that X' = C'B'A'.

10. Matrix products of special interest.

a. Multiplication by I. Unity can be defined as that thing which will leave another thing of the same class unchanged when multiplied with it. The identity matrix (I) has this singular property. (Note that any matrix, square or not, can be multiplied by an identity matrix of the proper size.)

14. Exercise: Make up any matrix A and verify that $AI = IA = A$.

b. Premultiplication by a diagonal matrix. Given $A_{(N,n)}$ and a diagonal matrix D with elements d_{11} d_{22} ... d_{NN} the product DA results in a matrix in which each row has been multiplied by a constant, the constants being respectively d_{11} d_{22} ... d_{NN}.

c. Postmultiplication by a diagonal matrix. Given the matrix A defined above, and a diagonal matrix D with elements d_{11} d_{22} ... d_{nn} the product AD will have each column of A multiplied by the corresponding element in D.

15. Exercise: Make up a matrix A and diagonal matrices; show that the rules just stated work out. What happens if all elements in a diagonal matrix are the same? Note that I is a special case of this.

d. Multiplication by unit vectors and matrices. A vector of size n with each element equal to 1 is called a unit vector, and is symbolized as $\tilde{j}_n$. If the size is understood, the subscript may be dropped, $\tilde{j}$.

Given any vector $\tilde{v}$, a vector multiplication of the form $\tilde{j}'\tilde{v}$ is the sum of the elements in $\tilde{v}$. Verify this for yourself.

Given any matrix A, the product $\tilde{j}'A$ results in a vector in which each element is the sum of a column of A. Verify this, then determine for yourself what $A\tilde{j}$ is.

A square matrix with each element equal to 1 is called a unit matrix and symbolized J; if the order is not clear, J_N.

16. Challenge Exercise: Given the N by n matrix C in section 9a above, calculate JC, and describe it in words. Then calculate the matrix (C - JC/N), and describe it in words.

e. Powers of a diagonal matrix. Because diagonal matrices equal their transpose (D = D'), we may write:

$$D'D = DD = D^2$$

In fact, given a diagonal matrix D, with elements d_{11} d_{22} d_{33} ... d_{nn}, we may say that D to any power ^{2} will be a diagonal matrix with non-zero elements $d_{11}^{\ 2}$ $d_{22}^{\ 2}$ $d_{33}^{\ 2}$ ... $d_{nn}^{\ 2}$

17. Exercise: Given:

$$D = \begin{bmatrix} 25 & 0 & 0 \\ 0 & 9 & 0 \\ 0 & 0 & 4 \end{bmatrix}$$

Calculate $D^{-\frac{1}{2}}$

11. The reciprocal (inverse) of a matrix.

a. The reciprocal of a number, N, is symbolized as $1/N$ or N^{-1} and defined as that quantity which satisfies the equation $N N^{-1} = 1$. For one number (zero) the reciprocal does not exist.

b. The reciprocal (= inverse) of a square matrix A may exist. It is symbolized by A^{-1} and defined as that matrix which satisfies the equation $A A^{-1} = I$. For some matrices, no inverse exists.

c. Example:

18 Given A and A^{-1} below, show that $A^{-1} A = I$ and $A^{-1} A = I$.

$$A = \begin{bmatrix} 7 & 2 \\ 10 & 3 \end{bmatrix} \qquad A^{-1} = \begin{bmatrix} 3 & -2 \\ -10 & 7 \end{bmatrix}$$

d. Calculation of an inverse is complicated for large matrices. For a 2 x 2 matrix, the formula is simple and you should learn how to use it.

$$A^{-1} = \begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix} \frac{1}{(a_{11} a_{22} - a_{12} a_{21})}$$

19 Exercise: using this formula, calculate the inverse given in c. above.

e. When does a matrix have no inverse? From the formula just given for the inverse of a 2 x 2 matrix, it is clear that if the quantity $(a_{11} a_{22} - a_{12} a_{21}) = 0$, the inverse of A does not exist. This quantity is known as the determinant of A, and symbolized as det A or as $/A/$. It has many important properties, and will be discussed later in more detail.

see 12b below

20 Exercise: given the matrices below, show that no inverse exists for them:

$$\begin{bmatrix} 0 & 7 \\ 0 & 8 \end{bmatrix} \qquad \begin{bmatrix} 3 & 3 \\ 3 & 3 \end{bmatrix} \qquad \begin{bmatrix} 1 & 4 \\ 2 & 8 \end{bmatrix} \qquad \begin{bmatrix} 1 & 0 & 10 \\ 5 & 6 & 50 \\ 2 & 1 & 20 \end{bmatrix}$$

f. Use of inverse matrices. Inverses are used in solving matrix equations. For example, given the matrix equation

$$AB = C$$

if A a square matrix, then we can solve for B by pre-multiplying both sides of the equation by A^{-1}, as follows:

$$A^{-1} A B = A^{-1} C$$

hence : $I B = A^{-1} C$

and: $B = A^{-1} C$

12. Determinant of a matrix. Every square matrix A has a number associated with it called its determinant, symbolized as det A or /A/, which is zero if and only if the matrix A has no inverse.

a. Calculation of the determinant of a 2 x 2 matrix is easy.

$$/A/ = a_{11} a_{22} - a_{12} a_{21}$$

b. Calculation for 3 x 3 is more complex:

$$/A/ = a_{11} a_{22} a_{33} + a_{21} a_{32} a_{13} + a_{12} a_{23} a_{31}$$
$$- a_{31} a_{22} a_{13} - a_{21} a_{12} a_{33} - a_{32} a_{23} a_{11}$$

c. For larger matrices, use a computer program.

d. Properties of determinants. Many useful properties occur. We have already learned their use in indicating matrices lacking inverses. Another property is their use in evaluating the tendency for numbers in a matrix to cluster symmetrically along the principal diagonal. Consider the four matrices below, each of which has the same sum of all numbers in the matrix (sum = 18), and the determinant
21 for each. As an exercise, check the ~~first two~~ values for det A and det B.

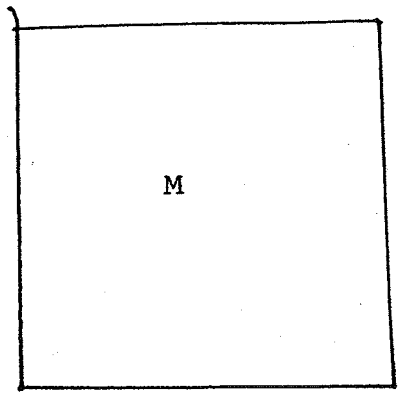
$$\underset{A}{\begin{bmatrix} 6 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 6 \end{bmatrix}} \quad \underset{B}{\begin{bmatrix} 3 & 2 & 0 \\ 2 & 4 & 2 \\ 0 & 2 & 3 \end{bmatrix}} \quad \underset{C}{\begin{bmatrix} 2 & 2 & 2 \\ 2 & 2 & 2 \\ 2 & 2 & 2 \end{bmatrix}} \quad \underset{D}{\begin{bmatrix} 0 & 0 & 6 \\ 0 & 6 & 0 \\ 6 & 0 & 0 \end{bmatrix}}$$

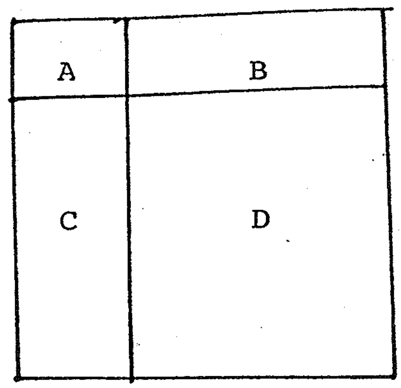
$|A| = 216 \qquad |B| = 12 \qquad |C| = 0 \qquad |D| = -216$

Other useful properties exist, and will be discussed later.

13. Trace of a matrix. Every square matrix S has a number associated with it known as the trace of S and written tr S. It is simply the sum of all elements in the principal diagonal of S. For
22 example, the matrix R in section 8e above has tr R = 13. Verify this

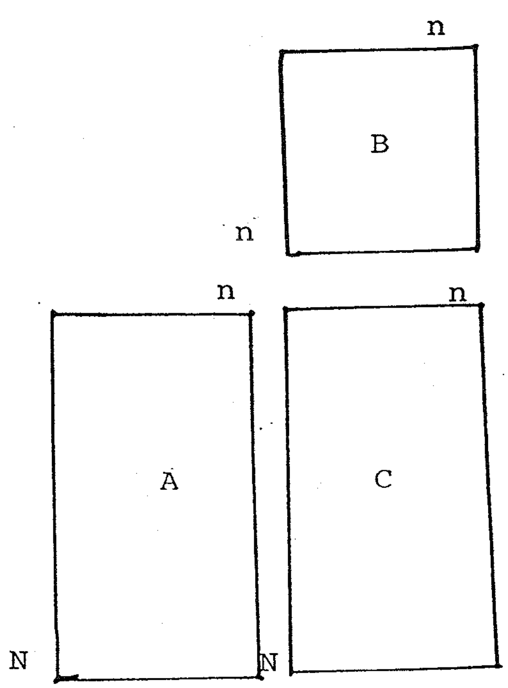
14. Partitioning of a matrix. Given any matrix M, it is sometimes convenient to partition it into two or more submatrices according to the needs of the problem. An example is diagrammed below in which M is partitioned into four submatrices A, B, C, and D.

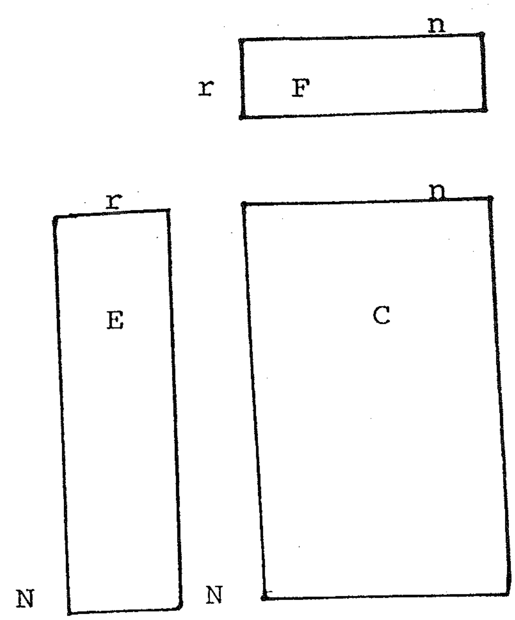




15. Rank of a matrix.

Given a matrix $C_{(N,n)}$ it is always possible to factor it into a product of two other matrices A and B, so that C = AB, provided that B is a square matrix of order n. It *may* be possible to factor C into two matrices E and F, in which the number of rows of F is *less than n*. Call this number r. Then the smallest value r can take is the *rank* of C.





23 **Exercise**: given the matrix below, find by inspection or calculation its rank.

a.
$$\begin{bmatrix} 7 & 14 & 6 \\ 4 & 8 & 7 \\ 7 & 14 & 11 \\ 13 & 26 & 19 \\ 9 & 18 & 12 \end{bmatrix}$$

$\Leftarrow$ do this last; it ain't easy

b. Do the same for the matrix below:

$$\begin{bmatrix} 1 & 2 & 3 \\ 3 & 6 & 9 \\ 13 & 26 & 39 \\ 4 & 8 & 12 \end{bmatrix}$$

c. Make up a 4 x 3 matrix of rank 3.

d. make up a 4 x 3 matrix of rank 2.

16. Vector Geometry.

Up to now we have considered vectors purely as algebraic quantities. It is useful to define the geometric properties of length and angle in such a way that, if the vectors are of size 3, these quantities correspond exactly to our normal Euclidian notions. These concepts are perfectly general, however, and extend to cases where the size of vectors is more than 3.

a. Length: The square root of the sum of the squares of the elements in a vector is defined as its length. Thus, given a vector $\tilde{x}_j$ its length (symbolized l_j) is given by

$$h_j = \sqrt{\tilde{x}'_j \tilde{x}_j}$$

b. Angle: Given two vectors $\tilde{x}_j$ and $\tilde{x}_k$, the cosine of the angle between them, $\cos\theta_{jk}$ is the scalar product divided by the product of their lengths.

$$\cos\theta_{jk} = \tilde{x}'_j \tilde{x}_k \,/\, h_j h_k$$

24. Exercise: What angle separates the two vectors

$$\tilde{b}' = \begin{bmatrix} 3 & 0 & 1 & 0 \end{bmatrix} \text{ and } \tilde{c}' = \begin{bmatrix} 0 & 7 & 0 & 3 \end{bmatrix}$$

17. Total sum of squares in a matrix

25. Exercise Write a matrix X, and calculate the square of each of its elements. Show that the sum of these squares = trX'X = trXX'.

18. Matrix formula for arithmetic mean.

Given a vector of observational data $\tilde{x}$ of size N, the mean (=average) of the observations is:

$$\bar{x} = \tilde{j}'\tilde{x}/N$$

26. Verify this for yourself with an example.

19. Matrix formula for variance.

Given a vector of observational data $\tilde{x}$ of size N, the variance of $\tilde{x}$, symbolized s_x^2 is the sum of the squared deviations from the mean. The matrix formula is:

$$s_x^2 = (\tilde{x} - J\tilde{x}/N)' \ (\tilde{x} - J\tilde{x}/N)/N$$

i.e., the squared vector $(\tilde{x} - J\tilde{x}/N)$ divided by N.
The variance is extensively used as a measured of the dispersion of values in a vector of observational data.

27. Exercise: Given $\tilde{x}' = \begin{bmatrix} 1 & 2 & 0 & 5 \end{bmatrix}$, show that $s_x^2 = 3.5$

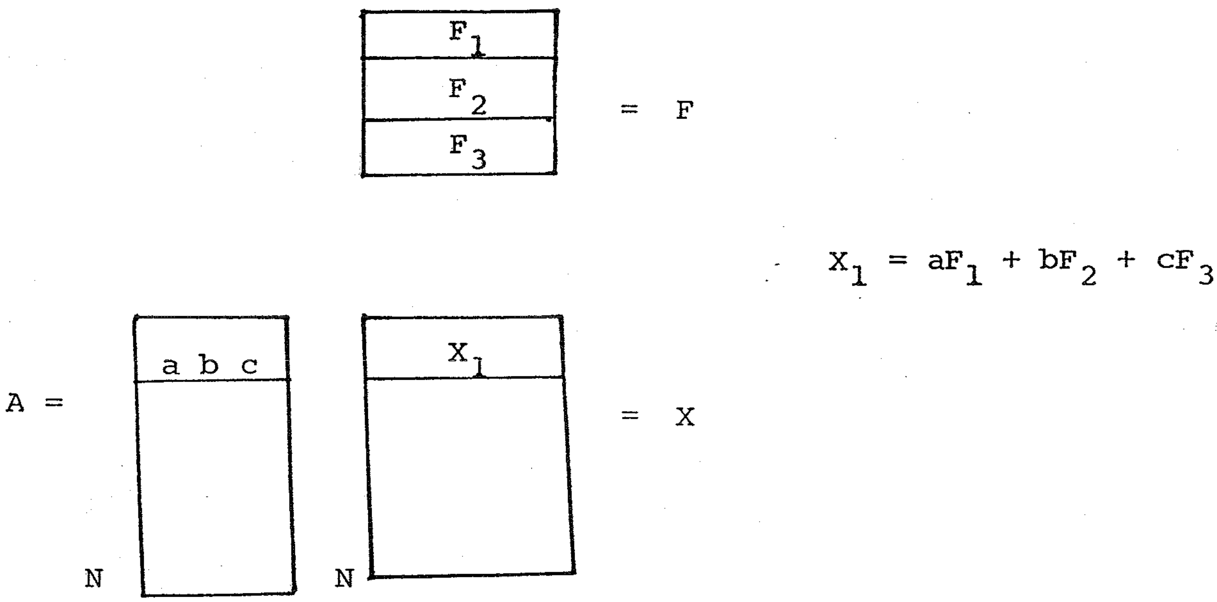
20. Product moment matrices.

Given a matrix $X_{(N,n)}$ the matrices X'X and XX' are called product moment matrices. Each is a square, symmetrical matrix. X'X is called the minor product moment, an n by n matrix with each element in the principal diagonal being the sum of squared elements in a column vector of X. XX' is an N by N major product moment with each element in the principal diagonal being the sum of squared elements in one row of X.

28. Exercise: Verify these relationships with an example.

21. MATRIX MULTIPLICATION AS A SET OF WEIGHTED COMBINATIONS OF VECTORS.

A matrix multiplication of the form X = AF may be regarded as a notation indicating how the row vectors of X are to be formed by making simple weighted combinations of the row vectors of F. Each row vector of A is the set of weights used to form the corresponding row vector of X from those in F. Stated in words, this may seem confusing. A diagram will make things clear:

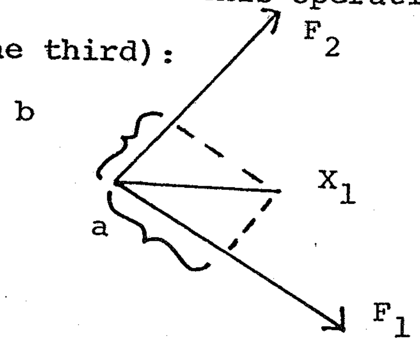


Here the data matrix X is considered to be of order N by 3; X_1 is defined as the first row vector of X; A is an N by 3 <u>matrix of weights</u>, with the vector (a b c) containing as elements the weights to be used; and F is a matrix of <u>components</u> in which the row vectors F_1, F_2, and F_3 are the three components. Then the vector X_1 is formed by adding together a parts of F_1, b parts of F_2, and c parts of F_3. As a vector equation:

$$X_1 = aF_1 + bF_2 + cF_3.$$

Other rows of X are formed according to the weights specified in other rows of A. The matrix equation X = AF is thus a set of N vector equations.

A geometric picture of this operation may be given in two dimensions (if we drop the third):



Here we form the vector X_1 by vector addition of two vectors, one being the fraction a of F_1 and the other the fraction b of F_2. If desired, the exact angles and lengths may be found using formulae in Paragraph 16 above.

Finally, if we transpose the matrix equation to $X' = F' A'$, we see that the columns of X may also be regarded as linear combinations of the columns of A. Here, the rows of F' are the <u>weights</u>, and the rows of A' are the <u>components</u> for the formation of the rows vectors in X'.

29. <u>Exercise</u>: Make up a matrix A of order 5 by 2, and a matrix F of order 2 by 2, and perform the matrix multiplication $X = A F$. Then take graph paper, and demonstrate that the geometric picture of matrix multiplication discussed above is valid for your example. Begin by plotting two component vectors F_1 and F_2, and show that each vector in X (row vectors) represent the parallelogram sums.

[redacted]

30

22. Normalized vectors.

Given a vector $\tilde{x}$ of length h, the vector $\tilde{x}/h$ is said to be normalized because its length = 1. $\tilde{x}/h \equiv \tilde{u}$

Exercise: Given $\tilde{x}' = \begin{bmatrix} 4 & 4 & 2 & 3 & 2 \end{bmatrix}$

~~calculate [illegible]~~ normalize $\tilde{x}'$ (use fractions). Call this normalized vector $\tilde{u}'$ and show that the length of $\tilde{u}$ is 1.

23. Normal matrices.

A matrix U with each column of length 1 is said to be a column-wise normal matrix. Row-wise normal matrices may also be formed.

24. Orthogonal matrices.

A matrix Q with angles between each pair of column vectors = 90° is said to be a column-wise orthogonal matrix. This is only true if $Q'Q = D$. (D is any diagonal matrix). 2

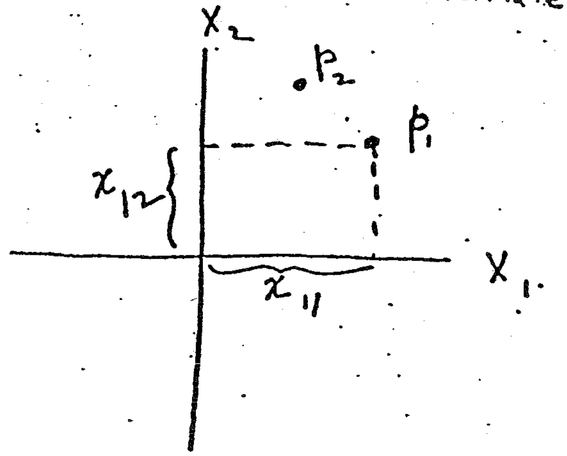
Row-wise orthogonality follows if $QQ' = D$.

25 Orthonormal matrices.

A special column-wise orthogonal matrix with column vector lengths all = 1 is a column-wise orthonormal matrix. In such a case the minor product moment = I. Row-wise orthonormal matrices also occur. How can they be recognized?

Multiplication and Coordinate Rotation.

* Given $X_{(N,2)}$, consider it to represent N points in a two-coordinate space.



$$X = \begin{array}{c} p_1 \\ p_2 \\ \vdots \\ p_N \end{array} \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \\ \vdots & \vdots \\ & \end{bmatrix}$$

* To rotate ~~coordinates~~ axes X_1, X_2 through an angle θ and get coordinates of p_1 and p_2 etc. in terms of new coordinates Y_1, Y_2, post multiply X by T where

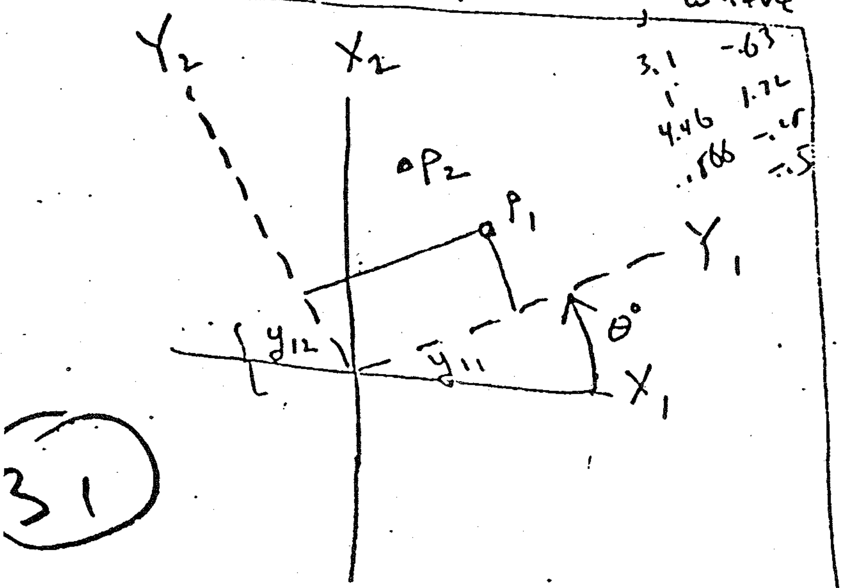
$$T = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}.$$

$YT' = X$

$X =$

Thus

$$Y = XT, \quad \text{where} \quad Y = \begin{array}{c} p_1 \\ p_2 \\ \vdots \\ p_N \end{array} \begin{bmatrix} y_{11} & y_{12} \\ & \text{etc.} \\ & \end{bmatrix}$$



3. The axes are rotated rigidly, retaining a right angle, if and only if

$$TT' = I.$$

Confirm this for the example above.

Exercise: make up and perform a rotation with $\theta = 30°$. Check graphically.

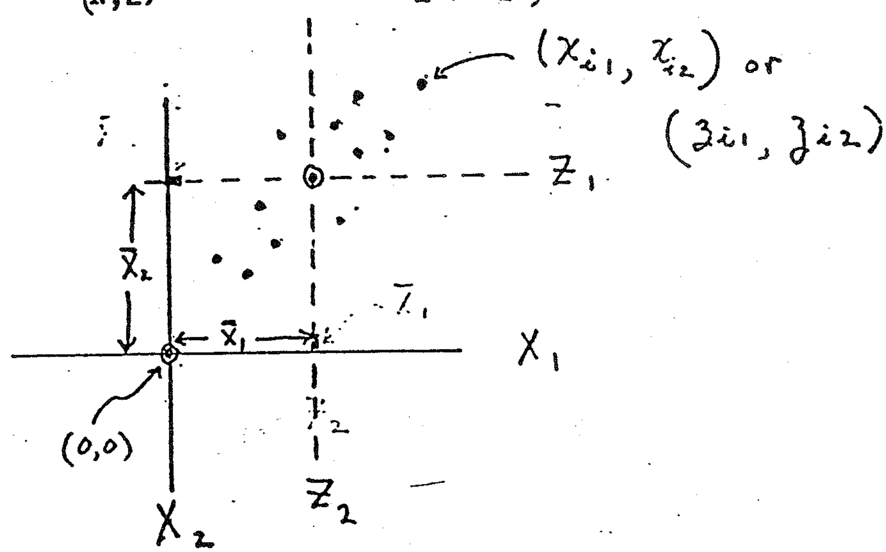
The above rotation procedure for two dimensions can be re-stated as follows: the operation $Y = X.T$ yields coordinates with respect to the new axes Y given coordinates of points with respect to initial axes X. Depending on the structure of the square transformation matrix T, the rotation may or may not distort the angular and distance relationships originally given in X. If $TT' = I$, no distortion occurs and the transformation is said to be orthogonal; it imvolves only rotation of axes.

In the two-dimensional case, there is only one parameter to be selected in an orthogonal rotation — the angle θ.

In a three-dimensional case, T is a square matrix of order three and $TT' = I$ for orthogonal transformations. In this case, two parameters must be selected. For this case and for higher dimensions, it is usual to abandon the trigonometric interpretation of the elements in the matrix and simply view eadhh element as an abstract number.

27. Coordinate translation. For many purposes it is convenient to translate the coordinate system, without rotation, to a new point of origin at the joint mean or centroid of a swarm of observations. In the figure, we wish to move the origin from the point (0,0) implied as origin of the matrix $X_{(N,2)}$ to the point$(\bar{x}_1 , \bar{x}_2)$,

$$X = \begin{array}{c c} & \begin{array}{cc} 1 & 2 \end{array} \\ \begin{array}{c} 1 \\ i \\ N \end{array} & \begin{bmatrix} x_{11} & x_{12} \\ x_{i1} & x_{i2} \\ x_{N1} & x_{N2} \end{bmatrix} \end{array}$$



$$Z = \begin{array}{c c} & \begin{array}{cc} 1 & 2 \end{array} \\ \begin{array}{c} 1 \\ i \\ N \end{array} & \begin{bmatrix} z_{11} & z_{12} \\ z_{i1} & z_{i2} \\ z_{N1} & z_{N2} \end{bmatrix} \end{array}$$

X stands for the original data; the matrix Z is defined as the matrix of mean deviates or the standardized data matrix, ~~mmmhmmh~~ representing the data referred to a new coordiate system having the joint mean as its origin. It is the algebraic representation of the translated coordinate system.

missing ; this is my attempt to reconstruct it quickly

28. Eigenvalues.

Every square matrix M of size $n \times n$, has n numbers associated with it which are called eigenvalues and generally symbolized as $\lambda_1, \lambda_2, \ldots \lambda_n$. These eigenvalues, although abstract at first definition, play an extremely important role in matrix algebra; indeed in many physical contexts they play a crucial role. What are they? By definition, the eigenvalues are a set of numbers which -- when subtracted from the principal diagonal of matrix M produce a matrix which has a determinant of zero. For convenience, let us place the n eigenvalues in the principal diagonal of a diagonal matrix, so that

$$[\lambda]_n \equiv \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \cdots \lambda_n \end{bmatrix}.$$

Thus $$|M - [\lambda]| = 0$$

Often the n eigenvalues of M are placed in the principal diagonal of a diagonal matrix and the whole matrix is symbolized as $[\lambda]$. That is,

$$[\lambda] \equiv \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 \cdots & \lambda_n \end{bmatrix}$$

(λ is lower case lambda)

2. Definition of eigenvector. Given a square matrix M of order n, and an eigenvalue λ, a column vector $\tilde{E}$ is said to be an eigenvector of M if

$$[M - \lambda I] \; \tilde{E} = \tilde{0}$$

i.e., $\tilde{E}$ is orthogonal to all row vectors in $[M - \lambda I]$

Note that $\tilde{0}$ is a null vector, i.e. a column of n zeros.

For each eigenvalue there will be a corresponding eigenvector, $\tilde{E}_1$, $\tilde{E}_2$ ... $\tilde{E}_n$

Often the n eigenvectors of M are placed as columns in a matrix of eigenvectors, E. The length of each eigenvector is arbitrary. For some applications, however, it is important to normalize each eigenvector to unit length; when this has been done, the resulting matrix of normalized eigenvectors is symbolized Λ (capital lambda). Note $\Lambda'\Lambda =$

-i.e. Λ is columwise orthogonal if M is symmetrical; if M is not symmetrical, Λ is not orthogonal.

B. Calculation of eigenvalues and eigenvectors. Numerical example.

1. Calculate eigenvalues. Given $M = \begin{bmatrix} 3 & 1 \\ 2 & 2 \end{bmatrix}$, set up the equation

$$\begin{vmatrix} (3 - \lambda) & 1 \\ 2 & (2 - \lambda) \end{vmatrix} = 0$$

and solving it for λ. Evaluating the determinant, we have

$$(3 - \lambda)(2 - \lambda) - (1)(2) = 0$$

$$6 - 5\lambda + \lambda^2 - 2 = 0$$

$$\lambda^2 - 5\lambda + 4 = 0$$

. Solving this quadratic,* we obtain the two values of $\lambda_1 = 4$ and $\lambda_2 = 1$.

Thus, $[\lambda] = \begin{bmatrix} 4 & 0 \\ 0 & 1 \end{bmatrix}$.

---

* Rembmber that you may factor or use $\lambda = \dfrac{-b \pm \sqrt{b^2 - 4ac}}{2a}$

33 **Exercise:** Verify that the values just obtained do satisfy the algebraic definition of eigenvalue. Then calculate eigenvalues of the matrix $\begin{bmatrix} 2 & 4 \\ 1 & 1 \end{bmatrix}$.   3.56   -.56

2. Calculate eigenvectors. To illustrate the procedure, we will calculate the second eigenvector of the matrix M given above — i.e., the eigenvector $\tilde{E}_2$ corresponding to $\lambda_2 = 1$.

a. The first step is to calculate the matrix $[M - \lambda I]$, using the value of λ_2:

$$[M - \lambda_2 I] = [M - (1)\, I] = \begin{bmatrix} 2 & 1 \\ 2 & 1 \end{bmatrix}$$

b. From the definition of an eigenvector, we have

$$\begin{bmatrix} 2 & 1 \\ 2 & 1 \end{bmatrix} \tilde{E}_2 = \tilde{0}$$

$\tilde{E}_2$ will be a column vector with elements e_1 and e_2; in box form:

$$\begin{bmatrix} e_1 \\ e_2 \end{bmatrix}$$

$$\begin{bmatrix} 2 & 1 \\ 2 & 1 \end{bmatrix} \quad \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

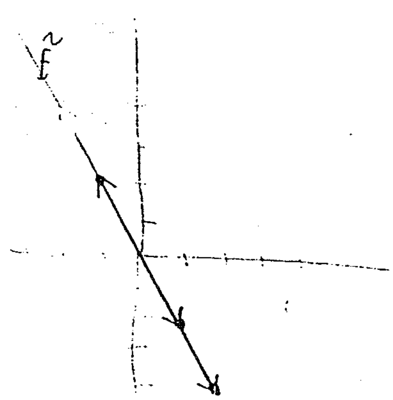
c. Expanding this matrix equation, we have two identical equations in e_1 and e_2:

$$2\, e_1 + e_2 = 0 \text{ , or}$$

$$\boxed{e_2 = -2\, e_1}$$

This equation has an infinite number of solutions, each with the value of e_2 being -2 times that of e_1. Thus the eigenvector is a direction represented by a vector of arbitrary length. Some possible $\tilde{E}_2$ vectors are:

$$\tilde{E}_2 = \begin{bmatrix} 1 \\ -2 \end{bmatrix} = \begin{bmatrix} 2 \\ -4 \end{bmatrix} = \begin{bmatrix} -1 \\ 2 \end{bmatrix} \ldots$$



d. Normalize the eigenvector, if desired, by the usual procedures. Take any solution vector for $\tilde{E}_2$, say $\begin{bmatrix} 1 \\ -2 \end{bmatrix}$ and calculate the normalized second eigenvector

$$\tilde{\Lambda}_2 = \begin{bmatrix} 1/\sqrt{5} \\ -2/\sqrt{5} \end{bmatrix}$$

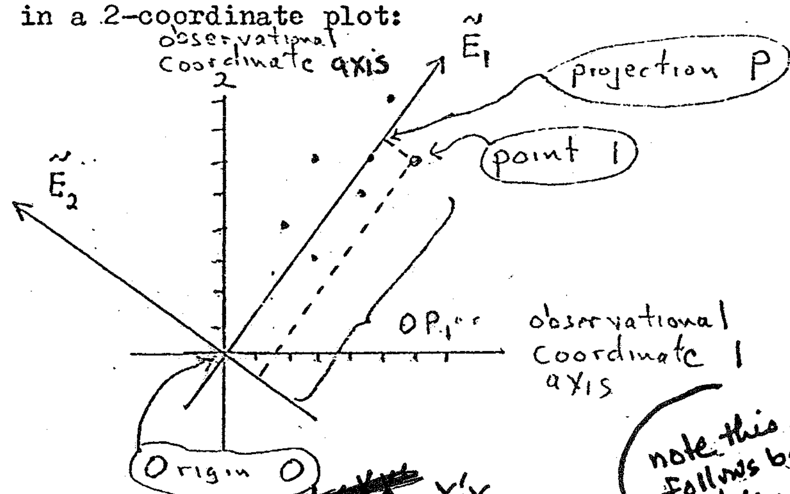
(34) **Exercise:** Verify by substitution in the basic eigenvector equation that $\tilde{\Lambda}_2$ just calculated is correct.

(35) e. Calculate $\tilde{E}_1$ and $\tilde{\Lambda}_1$ by the process just described. Then form the complete normalized eigenvector matrix Λ, and show that Λ is not columnwise orthonormal. Why is it not?

C. <u>Eigenvectors and values of the matrix $X'X$</u>. Eigenvalues and eigenvectors of certain matrices are extremely useful things. In this section we consider these quantities calculated for the <u>minor product moment</u> of a data matrix X. Given a matrix $X_{(N,2)}$, we can represent it as an elliptical swarm of N points in a 2-coordinate plot:

coordinates

$$X = \begin{array}{c} \\ 1 \\ 2 \\ \vdots \\ 3 \\ \vdots \\ N \end{array} \begin{array}{c} \begin{array}{cc} 1 & 2 \end{array} \\ \begin{bmatrix} 6 & 6 \\ 2 & 4 \\ \vdots & \vdots \\ \vdots & \vdots \\ \vdots & \vdots \\ \vdots & \vdots \end{bmatrix} \end{array}$$



Take the first ~~algebraic~~ eigenvector calculated from ~~XX~~ $X'X$ (this will be the vector of size 2, $\tilde{E}_1$) and plot it along with the N points of <u>X</u> on a 2-coordinate plot. Project each point perpendicularly on $\tilde{E}_1$ and call the projection point P, and the distance from the origin to the projection OP_1 for the first point; OP_2 for the second; etc. Now calculate the sums of the squares of these distances: this sum of squares will be the <u>maximum</u> possible for any vector drawn from the origin, and will have the value λ_1. $\tilde{E}_2$ will be perpendicular to $\tilde{E}_1$; and the sum of the squares of the projections onto $\tilde{E}_2$ will be λ_2. (If there were more columns than two in X, the second eigenvector would have the maximum possible sum of squares of any vector perpendicular to $\tilde{E}_1$; and the third the maximum for any perpendicular to $\tilde{E}_1$ and $\tilde{E}_2$, etc.)

note this follows because $X'X$ is symmetric to her 1/23/71

We may summarize the above relationships, and point out others of significance, in a series of terse statements; learn them:

1. The total "information" in X may be defined as the sum of the squared vector lengths of the N row vectors in X, and calculated as tr X'X. $= tr\ \Delta^2 = tr\ [\lambda]$

2. The eigenvector matrix E calculated from ~~X'X~~ X'X is a columwise orthogonal matrix whose column vectors $\tilde{E}_1$, $\tilde{E}_2$ etc., when plotted along with the row vectors of X in an n-coordinate plot, designate in decreasing order of importance the directions along which the variation in X is distributed.

3. The amount of variation parallel to an eigenvector is measured by the sum of the squared lengths of projections on that eigenvector: λ_1 is the sum for $\tilde{E}_1$, λ_2 for $\tilde{E}_2$, etc.

4. If there is no variation in a given direction, the correspondi ng eigenvalue is zero. If the rank of X is r, there will be r non-zero eigenvalues.

5. The total information = tr X'X = tr $[\lambda] = \lambda_1 + \lambda_2 + \cdots + \lambda_n$; i.e., the information initially recorded with respect to the axes of X is preserved in the new coordinate scheme E.

dispersion

6. The product of the non-zero eigenvalues $|[\lambda]| = |\Delta'| =$ $\lambda_1 \lambda_2 \cdots \lambda_n = /X'X/$; this quantity is related to the total volume occupied by the data points. the determinant of X'X

7. The normalized eigenvector matrix Λ is a transformation matrix for obtaining the projections of the row vectors of X on the eigenvectors: i.e., the desired projections $= X\Lambda$. Note that Λ is columnwise orthonormal. The angle between coordinate 1 and $\tilde{E}_1$ is given by the elements of Λ.

8. $\Lambda' X'X \Lambda = [\lambda]$

36

Exercise: Verify the above relationships by numerical calculation and careful plotting on graph-paper , given the following matrix: Calc. Λ & $[\lambda]$ on X'X

$$X = \begin{bmatrix} 5 & 4 \\ 1 & 6 \end{bmatrix}$$

Carry two significant decimal places. A calculator is helpful but not necessary. As a check on your work, $\lambda_1 = 68.07$.

$\Pi\ \Delta_1 \Delta_2$ = volume of ellipsoid.

29. <u>Dot notation for vectors</u>. In what follows it is convenient to have a simple notation for distinguishing row vectors from column vectors in a matrix, and to indicate particular vectors.

$$X = \begin{array}{c c} & \begin{matrix} 1 & 2 & .. & j & k & .. & n \end{matrix} \\ \begin{matrix} 1 \\ 2 \\ \vdots \\ i \\ p \\ \vdots \\ N \end{matrix} & \left[\begin{matrix} \\ \\ \\ \\ \\ \\ \\ \end{matrix} \right] \end{array}$$

Take the general notation for data matrix X given at the left; and note that the row vectors are numbered 1,2...i, p...N, indictating the first, second, i'th, p'th, and N'th vectors. The i'th vector means "any particular vector; and when i and p are both used it means "any two particular vectors. For the column vectors, they are numbered similarly 1, 2... j, k ...n.

A dot in a subscript is defined as "all the elements". Thus we indicate a desired vector as follows:

$\tilde{X}_{1.}$ = all the column elements in the first row = first row vecto in X

$\tilde{X}_{i.}$ = the i'th row vector of X

$\tilde{X}_{.j}$ = the j'th column vector of X; etc.

30. <u>Definition of a basic matrix</u>. Recall the definition of the <u>rank</u> of a matrix(p. 17). Clearly, the maximum rank a matrix can have is its ~~minimum~~ smaller dimension. For a normal data matrix X, with more rows than columns, the maximum rank possible is n. Such a matrix, with rank equal to its smaller dimension, is called a <u>basic matrix</u>. If the rank is less than the smaller dimension, it is a non-basic matrix.

31. <u>The basic structure of a matrix</u>: a fundamental law of matrix algebra, which should really turn you on. Although terribly abstract when first seen, this law represents a fantastic insight into the nature of the world. If you have comprehended what we have done so far, and work a bit to understand this, you will be rewarded with a an exhilarating sense of seeing deeply into the way things work -- or at least into the way in which we <u>think</u> about the way things work.

A. <u>Notation</u>. Let

P = a columnwise orthonormal matrix (= left basic orthonormal)
Q = another columnwise orthonormal matrix (= right basic orthonormal)
Δ = a diagonal matrix (= basic diagonal matrix)

X = a general data matrix, with no restrictions, of rank r and order N by n.

B. <u>The law</u>:

$$\boxed{X = P \Delta Q'}$$

That is, <u>any</u> matrix X may be expressed as the triple product of three matrices: a columnwise orthonormal P; a diagonal matrix Δ; and row-wise orthonormal matrix Q'. The wonder of it! Remember that each of the three matricies, P, Q, and Δ, is a <u>very</u> special <u>kind</u> of matrix. What are these matrices?

P = the normalized eigenvector matrix of XX'

Q = the normalized eigenvector matrix of X'X

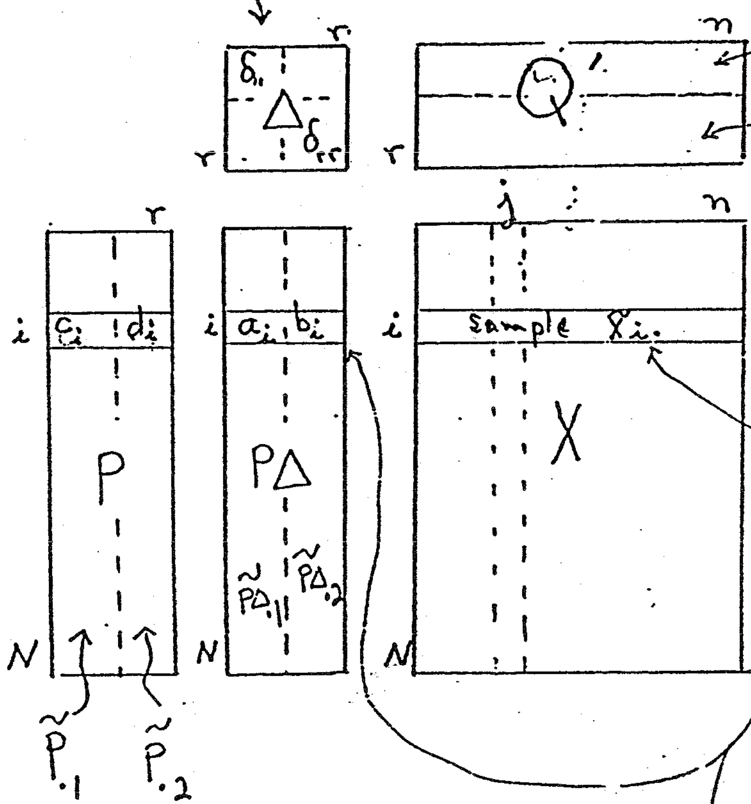
$\Delta = [\lambda]^{\frac{1}{2}}_{(r)}$ of XX' $= [\lambda]^{\frac{1}{2}}_{(r)}$ of X'X ; a matrix of order r by r formed from $[\lambda]$ by eliminating its non-zero diagonal elements and taking square

What does it all mean? In general, it means that a data matrix X usually has variables which exhibit incomplete independence and contain arbitrary scale factors. The scale factors are isolated in Δ; and the r <u>independent</u> dimensions of variation are displayed for the variables in the r columns of P, and for the samples in the r rows of Q'. Redundancy has been eliminated, scale isolated, and (hopefully) greater clarity achieved

C. <u>Analysis</u> <u>of</u> <u>the</u> <u>un-transposed</u> <u>case</u>. It will be helpful to examine the box-form of the $X = P\Delta Q'$ law [redacted].

T: Consider
st the product
$\vdots = (P\Delta)Q'$; and
in $(P\Delta) = P\Delta$

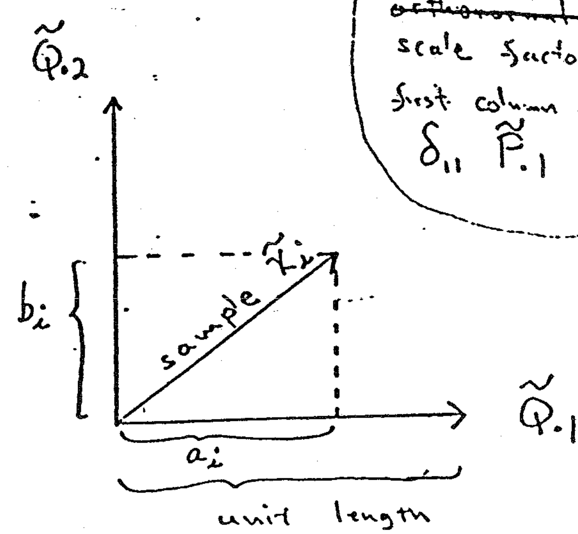
$\Delta = [\lambda]^{\frac{1}{2}}$ where $[\lambda]$ is derived from $X'X$ or XX';
$\delta_{11} = \sqrt{\lambda_{11}}$ = scale factor abstracted from $\tilde{P}_{\cdot 1}$; etc.



scale-free orthonormal <u>theoretical</u> <u>samples</u> = columns of Λ' derived from $X'X$: reference vectors for resolution of samples $\tilde{X}_{i}$

according to values given in i'th row of $P\Delta$ matrix of sample weights; columns of $P\Delta$ are theoretical orthogonal variables, ~~orthonormal but~~ with scale factors retained – first column of $P\Delta$ = $\delta_{11}\,\tilde{P}_{\cdot 1}$, etc.

and $\tilde{P}_{\cdot 2}$ are
scale-free
thonormal
<u>eoretical</u>
<u>ariables</u> =
lumns of Λ
rived from XX' =
ference vectors
r resolution of
riables $\tilde{X}_{\cdot j}$ according to weights
en in matrix $(Q\Delta)$ not
programmed here (see transposed case)



<u>Geometric Picture of $X = (P\Delta)Q'$</u>

<u>$P\Delta$</u>

min prod	maj h
$(P\Delta)'(P\Delta)$	$P\Delta(P\Delta)'$
$= \Delta^2$	$= XX'$

A numerical example, although trivially simple and artificial, ~~will~~ will help make the meaning of the basic structure law clear. The data matrix X below represents observations made at four localities arranged in increasing order of latitude. At each open-ocean station three measurements are made on density of living plankton in units of specimens per 1000 cc: forams, radiolaria, and diatoms.

$$
X = \begin{array}{l} \\ \text{Loc. 1} \\ \text{Loc. 2} \\ \text{Loc. 3} \\ \text{Loc. 4} \end{array}
\begin{array}{c} \begin{array}{ccc} \text{Forams} & \text{Rads.} & \text{Diatoms} \end{array} \\ \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 2 & 4 & 6 \\ 4 & 8 & 12 \end{bmatrix} \end{array}
$$

(1) Compute X'X , and from this $[\lambda]_{X'X} = \begin{bmatrix} 350 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$

and $\Lambda'_{X'X} = \left[\frac{1}{\sqrt{14}} \quad \frac{2}{\sqrt{14}} \quad \frac{3}{\sqrt{14}} \right] \equiv Q'$

(2) Compute XX', and from this

$$[\lambda]_{XX'} = \begin{bmatrix} 350 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \text{ and } P' \equiv \Lambda'_{XX'} = \left[\frac{1}{5} \quad \frac{2}{5} \quad \frac{2}{5} \quad \frac{4}{5} \right]$$

(3) Clearly, r = 1, and we form $\Delta^2 = [350]$ and $\Delta = [\sqrt{350}] = [5\sqrt{14}]$

(4) We now can show numerically that $X = P\Delta Q'$, and analyze the meaning.
(Exercise: verify the calculation below)
and that $\boxed{tr\, X'X = tr\, \Delta^2}$

$$
\Delta = [5\sqrt{14}] \qquad \begin{array}{c} \begin{array}{ccc} \text{For.} & \text{Rad.} & \text{Diat.} \end{array} \\ \left[\frac{1}{\sqrt{14}} \quad \frac{2}{\sqrt{14}} \quad \frac{3}{\sqrt{14}} \right] \end{array}
$$

Q' = theoretical normalized sample showing proportions 1, 2, 3; diatoms are at [illegible] of the food chain, and were [illegible] than their prey

$$
\begin{array}{l} \text{Loc 1:} \\ \text{Loc 2:} \\ \text{Loc 3:} \\ \text{Loc 4:} \end{array}
\quad P = \begin{bmatrix} \frac{1}{5} \\ \frac{2}{5} \\ \frac{2}{5} \\ \frac{4}{5} \end{bmatrix}
\qquad \underbrace{\begin{bmatrix} 1\sqrt{14} \\ 2\sqrt{14} \\ 2\sqrt{14} \\ 4\sqrt{14} \end{bmatrix}}_{P\Delta}
\qquad \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 2 & 4 & 6 \\ 4 & 8 & 12 \end{bmatrix} = X
$$

vector of weights, with scale, showing amount of theoretical sample in each observed sample

Theoretical normalized variable representing the single independent dimension of variation — in this case, fertility of oceans increasing poleward.

ok

D. Analysis of the transposed case. Given $X = P\Delta Q'$, we may write

$$X' = Q\Delta' P'$$

But since Δ is diagonal, $\Delta' = \Delta$ and the result is simply

$$\boxed{X' = Q\Delta P'}$$

7b)

Exercise: Prepare in box form an anotated analysis of the transposed case, and a geometric picture of $X' = (Q\Delta)P'$. Set this up exactly as done on the ~~previous~~ page 32.

E. Uses of the basic structure law: $X = P\Delta Q'$

1. To understand ~~thm~~ subtle and important relationships between data matrices and ~~thmmm~~ the product moment matrices derived from them, including:

a. X not being square, it can not have eigenvalues and vectors; yet it does have analagous features P, Δ, and Q, which are calculated from the product moments of X: The basic diagonal matrix Δ is obtained as the square root of the eigenvalue matrix $[\lambda]$ of X'X or XX'; the left-hand basic orthonormal P is the normalized eigenvector matrix of ~~fm~~ XX'; and the right-hand basic orthonormal Q is the eigenvector matrix of X'X.

b. The orthonormal matrix Q represents the reference vectors or principal axes within whibh the N sample vectors of X can be plotted.

c. The orthonormal matrix P represents the reference vectors or principal axes within which the n column vectors representing observational varibles in X can be plotted.

2. To prove various theorms of matrix algebra, some of which we have already simply stated. For example:

32. Augmented matrices. Sometimes it is convenient to add a column to a given matrix; it may then be said to be augmented. One common augmented matrix has a column vector of ones (a unit vector $\tilde{j}$) added on its right. Given a matrix X, a $\tilde{j}$-augmented X matrix is symbolized X^*, i.e.:

$$X^* \cong [X \quad \tilde{J}]$$

For example, if $X = \begin{bmatrix} 4 & 2 \\ 1 & 3 \\ 1 & 7 \\ 6 & 6 \end{bmatrix}$., $X^* = \begin{bmatrix} 4 & 2 & 1 \\ 1 & 3 & 1 \\ 1 & 7 & 1 \\ 6 & 6 & 1 \end{bmatrix}$

33. Quadratic forms.

(1) Pre- and post- multiplication of a matrix M by $\tilde{j}$ in the form $\tilde{j}' M \tilde{j}$ yields a scalar, which is the sum of all elements in M. For example:

$$M = \begin{bmatrix} 3 & 1 & 3 \\ 4 & 2 & 5 \\ 5 & 5 & 8 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

$$[1 \; 1 \; 1][12 \;\; 8 \;\; 16] \; [36]$$

$$\tilde{j}' M \tilde{j} = 36$$

(2) Pre-and post- multiplication of a square matrix C by a vector $\tilde{x}$ in the form $\boxed{\tilde{x}' C \tilde{x}}$ yields a scalar called a quadratic form which is a weighted sum of the elements in C. What are the weights? If the elements in C are designated by subscript symbols as $[c_{ij}] = C$; and the vector $\tilde{x}' = [x_1 \dots x_2 \dots x_i \dots]$ and xthose in its transpose $\tilde{x} = [x_1 \dots x_2 \dots x_j \dots]$, then the weight for any element c_{ij} is $(x_i)(x_j)$. Each weight is a product of two elements in $\tilde{x}$, or a square of an element in $\tilde{x}$. After multiplying each elementt in C by the appropriate weight, thh results are addde to form the final scalar value. Note that This explanation of what the weighted sum is simply follows from the basic operation of pre and post multiplioation by a vector. Although the desdription sounds complex, the results are not. Note the following example:

Note the pattern of weights in $\tilde{x}\,\tilde{x}'$: each weight is a product of two elements in $\tilde{x}$; the elements squared form the principal diagonal; and in the off-diagonal positions each possible cross-product is entered twice in symmetrical positions across the diagonal.

39 **Exercise:** Given $\tilde{x}' = \begin{bmatrix} 2 & b \end{bmatrix}$ and $C = \begin{bmatrix} 1 & 3 \\ 2 & 7 \end{bmatrix}$ evaluate the quadratic form $\tilde{x}'C\,\tilde{x}$.

(3) One important use of the quadratic form is to express compactly and to manipulate second degree equations in n variables. Consider, for example, the general second degree equation in 3 variables x_1 , x_2, and x_3:

$$\left\{a\,x_1^2 + b\,x_2^2 + c\,x_3^2 + d\,x_1x_2 + e\,x_1x_3 + f\,x_2x_3\right\} + g\,x_1 + h\,x_2 + i\,x_3 + k = 0$$

The portion in braces is the quadratic form. The remainder is a linear form (three first degree terms) and a constant. ~~By expressing the original variables in mean-deviate form (z rather than x), ...~~ ~~elimination.~~ If we arrange the coefficients in a special way in a matrix C, many useful operations are possible:

$$C = \begin{bmatrix} a & d/2 & e/2 \\ d/2 & b & f/2 \\ e/2 & f/2 & c \end{bmatrix}$$

Note that the coefficients of the square terms are placed unchanged in the principal diagonal, but that the remaining cross-product coefficents have each been divided in two and placed in symmetrical positions across the diagonal. If this is done, and we define $\tilde{x}' = \begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix}$, then $\tilde{x}'\,C\,\tilde{x}$ is in fact the matrix notation of the quadratic form written out in braces in the equation above.

40 **Exercise:** Verify that the statement in the last sentence is true. Then write down a matrix notation for the quadratic form part of a second degree equation in four variables.

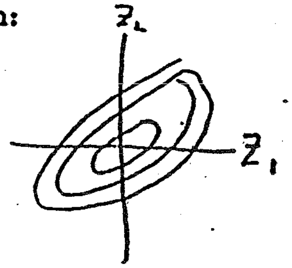
The entire equation given above can be written in an even simpler way by first transforming the x-variates to z-variates — i.e., by expressing the variables in a Z matrix of mean-deviates. This operation eliminates the linear terms, and the entire equivalent equation of a second degree equation in any number of variables becomes:

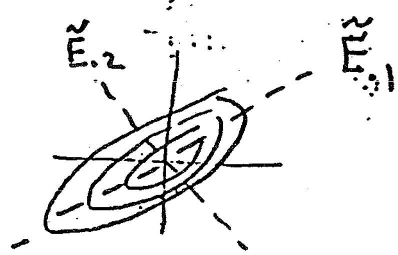
$$\boxed{\tilde{z}'\,C\,\tilde{z} + k = 0}$$

where C is an n by n square matrix of coefficients arranged as abovd.

(4) Analysis and rotation of quadric surfaces. A graph of a general second-degree equation in any number of variables is called a quadric surface of a conoid. If we take the eigenvalues and eigenvectors $[\lambda]$ and Λ of the matrix $\underline{C}$ of the quadratic form of the equation, the following useful rules hold:

a. $Z\Lambda$ rotates the surface to a new position with coordinate axes located symmetrically. Graphically, $Z\Lambda$ places the paraboloid sketched below as shown:





b. The eigenvectors are the new coordinate axes, and the new equation is in canonical form is:

$$\lambda_1 E_1^2 + \lambda_2 E_2^2 + \lambda_3 E_3^2 = k$$

where Y_1 is the first column vector of the matrix of projections $Y = Z\Lambda$, etc.

* c. If the eigenvalues are all positive, the surface is an ellipsoid; if of different sign but none are zero, a hyperboloid; if one is zero, a paraboloid.

(41) Exercise: Given the equation

$$5x^2 + 8xy + 5y^2 - 18x - 18y + 9 = 0.$$

decide what curve is represented and sketch the curve. Assume you know $\bar{x}$ and $\bar{y}$.. (Hint: all you need is the quadratic form).

$$\begin{matrix} 5 & 4 \\ 4 & 5 \end{matrix}$$

SUMMARY OF STANDARD SYMBOLS USED IN MATRIX ALGEBRA

	SYMBOL	DEFINITION	PAGE REFERENCE
NUMBERS	N	number of samples (rows) in X	3,4
	n	number of variables (columns) in X	3,4
	i, p	row index for two general rows of X	3,30
	j, k	column index for two general columns of X	30
	λ	eigenvalue	25
	r	rank of a matrix	19
	h	length of a vector	18
	$\bar{x}$	mean of vector $\tilde{x}$	19
	s^2	variance of an observed set of numbers	19
	s	standard deviation of an observed set $= \sqrt{s^2}$	36
	δ	individual value in basic diagonal matrix Δ	32
	tr M	trace of matrix M	10,18
	$/M/$	determinant of matrix M	15,16
	$\cos\theta$	cosine of angle between two vectors	18
MATRICES	X	data matrix	4,30
	Z	data matrix with columns expressed as mean deviates	25
	E	eigenvector matrix with un-normalized column vectors	25,26
	Λ	columwise ~~orth~~onormal* eigenvector matrix	26
	$[\lambda]$	matrix of eigenvalues	25,26
	P	left basic orthonormal matrix of X	31,32
	Δ	basic diagonal matrix $= [\lambda]^{\frac{1}{2}}_{(r)}$	31,32
	Q	right basic orthonormal matrix of X	31,32
	T	square transformation matrix	23-25
	U	columwise or row-wise normal matrix	22
	D	general diagonal matrix	6,13
	O	null matrix	5
	I	identity matrix	6
	J	unit matrix	13
	M^{-1}	inverse of M	14
VECTORS	$\underset{\sim}{j}$	columnwise unit vector	12
	$\underset{\sim}{X}_{.j}$	j'th column vector of X	30
	$\underset{\sim}{X}_{i.}$	i'th row vector of X	30
	$\underset{\sim}{x}$	column vector	4
	$\underset{\sim}{x}'$	row vector	4

* if M is symmetrical, Λ is orthonormal