Multivariate Data and Matrix Algebra Review

Biometry 726

Fall 2010

What is 'multivariate' data?

Data in which each sampling unit contributes more than one outcome.

Sampling unit	Multivariate outcome
Person	Duplicate serum concentration measures of a panel of cytokines (e.g. IL6, TNFα, etc.
Chick embryo heart	Number of cells in the superior and inferior atrioventricular cushions measured in six serial confocal planes
Elementary school	Third grade students' test scores
Twin pair	Age of death of each member
Cancer patient	Tumor response measured at 3 weeks, 2 months and 6 months post treatment

Multivariate data properties

What property/ies of multivariate data make commonly used statistical approaches inappropriate or impractical?

Goals of multivariate data analysis

1.
2.
3.
4.
5.

Random vectors

Because each ‘subject’ contributes multiple outcome measures to the analysis, it is convenient to organize subject i’s n_i outcomes as a column vector.

$$\mathbf{Y}_i = \begin{bmatrix} Y_{i1} \\ Y_{i2} \\ \vdots \\ Y_{i,n_i} \end{bmatrix}$$

- $\mathbf{Y}_i$’s dimension is ______________________
- $\mathbf{Y}_i$ is a random variable as are its individual elements
- The typeset depiction of a random vector uses *bold face* - $\mathbf{Y}_i$ rather than Y_i
- The handwritten depiction of a random vector is _______.

Random vectors (cont.)

Representing vectors as columns can take up a lot of space. To get around this, we often use the *transpose* operator to depict vectors. Therefore, we might write

$$\mathbf{Y}_i = (Y_{i1}, Y_{i2}, \ldots, Y_{i,n_i})'$$

where $'$ means *transpose*. Notice this representation states that $\mathbf{Y}_i$ is the transpose of a $1 \times n_i$ row vector, which makes it an $n_i \times 1$ column vector. Vectors are, by default, column vectors unless otherwise stated.

Matrix algebra - basic terminology

A rectangular array of real numbers arranged in m rows and n columns is called an $m \times n$ matrix.

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}.$$

We write $\mathbf{A} = \{a_{ij}\}$ to represent the matrix $\mathbf{A}$ whose ijth element is a_{ij}.

Matrix operations

- Addition: $\mathbf{A} + \mathbf{B} = \{a_{ij} + b_{ij}\}$ for $m \times n$ matrices $\mathbf{A}$ and $\mathbf{B}$
- Matrices $\mathbf{A}$ and $\mathbf{B}$ are *conformal* for addition (or subtraction) if the row dimensions of $\mathbf{A}$ and $\mathbf{B}$ are equal, and the column dimensions of $\mathbf{A}$ and $\mathbf{B}$ are equal.
- Matrix addition is commutative. ______________________
- Matrix addition is associative. ______________________
- Scalar multiplication: $c\mathbf{A} = \{ca_{ij}\}$

Matrix multiplication

- Matrix multiplication: For $m \times n$ matrix $\mathbf{A}$ and $n \times p$ matrix $\mathbf{B}$, the matrix product $\mathbf{AB}$ is the $m \times p$ matrix $\mathbf{C}$ where

$$c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}.$$

- Matrices $\mathbf{A}$ and $\mathbf{B}$ are conformal for the product $\mathbf{AB}$ if the column dimension of $\mathbf{A}$ equals the row dimension of $\mathbf{B}$.

Let $\mathbf{A} = \begin{bmatrix} 2 & -3 & 0 \\ 1 & 4 & 2 \end{bmatrix}$ and $\mathbf{B} = \begin{bmatrix} 3 & 4 \\ -2 & 0 \\ 1 & 2 \end{bmatrix}$. Find $\mathbf{AB}$.

Matrix multiplication (cont.)

- $\mathbf{AB}$ is the *pre-multiplication* of $\mathbf{B}$ by $\mathbf{A}$ or equivalently, the *post-multiplication* of $\mathbf{A}$ by $\mathbf{B}$.
- Matrix multiplication does not necessarily commute. That is, ________________________________.
- e.g. Let $\mathbf{A} = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$ and let $\mathbf{B} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$. Then $\mathbf{AB} =$ and $\mathbf{BA} =$
- Matrix multiplication is associative.

____

- Matrix multiplication distributes over addition.

____

Matrix transpose

- The transpose of an $m \times n$ matrix $\mathbf{A}$, denoted $\mathbf{A}'$, is the $n \times m$ matrix whose ijth element is the jith element of $\mathbf{A}$.
- More succinctly, let $\mathbf{C} = \mathbf{A}'$. Then $c_{ij} = a_{ji}$.
- $\begin{bmatrix} 2 & -3 & 0 \\ 1 & 4 & 2 \end{bmatrix}' =$
- $(\mathbf{A}')' = \mathbf{A}$
- $(\mathbf{A} + \mathbf{B})' = \mathbf{A}' + \mathbf{B}'$
- $(\mathbf{AB})' = \mathbf{B}'\mathbf{A}'$

Types of matrices

- *Square matrices* have the same number of rows and columns. The row (or column) dimension is called the *order* of the matrix.
- Note that the matrix product $\mathbf{AA}$ is defined only if $\mathbf{A}$ is square.
- If $\mathbf{A}^2 = \mathbf{AA} = \mathbf{A}$ then $\mathbf{A}$ is said to be *idempotent*.
- $\mathbf{A}$ is a *symmetric matrix* if $\mathbf{A}' = \mathbf{A}$.
- A square matrix $\mathbf{A}$ is *diagonal* if $a_{ij} = 0$ for all $i \neq j$, that is to say, if all off-diagonal elements are zero.
- The order n *identity matrix*, $\mathbf{I}_n$, is a diagonal matrix with diagonal elements equal to 1.

Triangular matrices

- Square matrix $\mathbf{A}$ is *upper-triangular* if $a_{ij} = 0$ for $i > j$.
- Square matrix $\mathbf{A}$ is *lower-triangular* if $a_{ij} = 0$ for $i < j$.

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}$$

- a_{ij} $i = j$
- a_{ij} $i < j$
- a_{ij} $i > j$

Matrix inverse definition

- An $n \times n$ matrix $\mathbf{A}$ is said to be *nonsingular* or *invertible* if there exists $n \times n$ matrix $\mathbf{B}$ such that

  $$\mathbf{AB} = \mathbf{BA} = \mathbf{I}.$$

  $\mathbf{B}$ is called the *multiplicative inverse* of $\mathbf{A}$. We write $\mathbf{B} = \mathbf{A}^{-1}$.
- A square matrix with no multiplicative inverse is said to be *singular*.

Matrix inverse - example

Demonstrate that $\mathbf{A} = \begin{bmatrix} 1 & 0 & 1 \\ 3 & 3 & 4 \\ 2 & 2 & 3 \end{bmatrix}$ is the inverse of

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

Matrix determinant

- The *determinant* of the square $n \times n$ matrix $\mathbf{A}$ is a scalar given by

$$|\mathbf{A}| = \begin{cases} a_{11} & \text{if } n = 1 \\ \sum_{j=1}^{n} a_{1j} A_{1j} & \text{if } n > 1 \end{cases}$$

- A_{1j} is called the *cofactor* of a_{1j}, and is defined as

$$A_{1j} = |\mathbf{A}_{1j}|(-1)^{1+j}$$

where $\mathbf{A}_{1j}$ is the $(n-1) \times (n-1)$ matrix obtained by deleting the first row and jth column of $\mathbf{A}$.

Determinant of a 2×2 matrix

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

Determinant of a 3×3 matrix

Find $|\mathbf{A}|$ where $\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$

Cofactor - general definition

- In general, $A_{ij} = |\mathbf{A}_{ij}|(-1)^{i+j}$ is the *cofactor* of a_{ij}.
- For $\mathbf{A} = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 5 & 3 \\ 2 & 3 & 2 \end{bmatrix}$, find A_{21}, the cofactor of a_{21}.
- It can be shown that

$$a_{i1}A_{j1} + a_{i2}A_{j2} + \ldots + a_{in}A_{jn} = \begin{cases} |\mathbf{A}| & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

Matrix adjoint

- Let $\mathbf{A}$ be an $n \times n$ matrix. We define a new matrix called the *adjoint* of $\mathbf{A}$ by

$$\text{adj } \mathbf{A} = \begin{bmatrix} A_{11} & A_{21} & \ldots & A_{n1} \\ A_{12} & A_{22} & \ldots & A_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ A_{1n} & A_{2n} & \ldots & A_{nn} \end{bmatrix}$$

- In words, the adjoint of $\mathbf{A}$ is formed by replacing each term by its cofactor, and then transposing the resulting matrix.

Matrix inverse revisited

$$\mathbf{A}\,(\text{adj }\mathbf{A}) = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix} \begin{bmatrix} A_{11} & A_{21} & \dots & A_{n1} \\ A_{12} & A_{22} & \dots & A_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ A_{1n} & A_{2n} & \dots & A_{nn} \end{bmatrix} =$$

Result: $\mathbf{A}^{-1} =$ ________________________________

Finding matrix inverse

Use the result on the previous slide to find the inverse of

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 1 \\ 3 & 3 & 4 \\ 2 & 2 & 3 \end{bmatrix}.$$

Vectors

- A column vector is an $m \times 1$ matrix.
- A row vector is a $1 \times n$ matrix.
- By default, vectors are assumed to be column vectors unless indicated otherwise.
- The number of vector elements is called its *dimension*.
- Inner product: The *inner product* or *dot product* of two m-dimensional vectors $\mathbf{x}$ and $\mathbf{y}$ is defined as

$$\mathbf{x}'\mathbf{y} = [x_1 \;\; x_2 \;\; \ldots \;\; x_m] \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix} = \sum_{i=1}^{m} x_i y_i$$

- For scalar c, $c\mathbf{x} = [cx_1 \;\; cx_2 \;\; \ldots \;\; cx_m]'$

Linear regression example

Let Y_i be the response for the ith subject with k covariates $x_{i1}, x_{i2}, \ldots, x_{ik}$. Recall that for multiple linear regression, we assume

$$Y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \ldots + \beta_k x_{ik} + \varepsilon_i.$$

The right hand side of this expression can be written as the inner product of two vectors, as follows:

Vector norm

Let $\mathbf{x}' = [x_1 \;\; x_2 \;\; \dots \;\; x_m]$.

- The *norm* (or *length* or *magnitude*) of $\mathbf{x}$ is given by:

  ____

- $\mathbf{z}$ is the vector $\mathbf{x}$ *normalized* to unit length if

$$\mathbf{z} = \frac{\mathbf{x}}{||\mathbf{x}||}.$$

  To see that $\mathbf{z}$ has unit length, note that:

Angle between two vectors

Let $\mathbf{x}$ and $\mathbf{y}$ be two m-dimensional vectors. The angle θ between the two vectors is defined such that

$$\cos(\theta) = \frac{\mathbf{x}'\mathbf{y}}{||\mathbf{x}||||\mathbf{y}||}.$$

Proof for 2-dimensional case:
(Use fact that $\cos(\theta_1 - \theta_2) = \cos(\theta_1)\cos(\theta_2) + \sin(\theta_1)\sin(\theta_2)$).

If $\mathbf{x} \perp \mathbf{y}$ then ______________________________

Orthogonal and orthonormal vectors

- The collection of equally-dimensioned vectors, $\mathbf{x}_1$, $\mathbf{x}_2$, $\ldots$, $\mathbf{x}_p$, are *orthogonal* if $\mathbf{x}_i'\mathbf{x}_j = 0$ whenever $i \neq j$.
- The collection of equally-dimensioned vectors, $\mathbf{x}_1$, $\mathbf{x}_2$, $\ldots$, $\mathbf{x}_p$, are *orthonormal* if they are orthogonal and $||\mathbf{x}_i|| = 1$ for all i. That is to say,

$$\mathbf{x}_i'\mathbf{x}_j = \begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{if } i = j \end{cases}$$

Linear dependence

- Let $\mathbf{x}_1$, $\mathbf{x}_2$, $\ldots$, $\mathbf{x}_p$ be a collection of vectors of equal dimension. We say $\mathbf{x}_1$, $\mathbf{x}_2$, $\ldots$, $\mathbf{x}_p$ are *linearly dependent* if there exist constants c_1, c_2, $\ldots$, c_p not all zero such that

$$c_1\mathbf{x}_1 + c_2\mathbf{x}_2 + \ldots + c_p\mathbf{x}_p = \mathbf{0}.$$

  Linear dependence means at least one vector in the set can be written as a linear combination of the other vectors.
- Vectors of the same dimension that are not linearly dependent are said to be *linearly independent*.

Row rank and column rank

- The *row rank* of a matrix is the maximum number of linearly independent rows.
- The *column rank* of a matrix is the maximum number of linearly independent columns.
- E.g. Find the row and column rank of

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 4 \\ 3 & 0 & 6 \\ 5 & 3 & 13 \end{bmatrix}$$

Rank (cont.)

- row rank = column rank
- If the rank of $m \times n$ matrix $\mathbf{A}$ is min(m, n), then $\mathbf{A}$ is said to be of *full rank*. Otherwise, $\mathbf{A}$ is said to be *rank deficient*.

Linear regression example - dummy variables

```
data one;                       data two;
    input y group;                  set one;
    datalines;                      if group = 1 then x1 = 1; else x1 = 0;
 0.62  2                            if group = 2 then x2 = 1; else x2 = 0;
-0.55  2                            if group = 3 then x3 = 1; else x3 = 0;
-0.50  1                        run;
 0.41  1
-0.55  3                        proc reg data = two;
 0.067 1                            model y = x1 x2 x3;
 1.27  3                        run;
-0.11  1
-0.33  2
-0.54  3
;
run;
```


Linear reg example - dummy variables (cont.)

Output

```
NOTE: Model is not full rank. Least-squares
      solutions for the parameters are not
      unique.  Some statistics will be misleading.
      A reported DF of 0 or B means that the
      estimate is biased.
NOTE: The following parameters have been set
      to 0, since the variables are a linear
      combination of other variables as shown.

         x3 =  Intercept - x1 - x2
```

Matrix trace

- Let $\mathbf{A}$ be a square $n \times n$ matrix. The *trace* of $\mathbf{A}$ is given by

$$\text{tr}(\mathbf{A}) = \sum_{i=1}^{n} a_{ii}.$$

- Trace is the sum of the diagonal elements of $\mathbf{A}$
- Properties
  - $\text{tr}(\mathbf{AB}) = \text{tr}(\mathbf{BA})$
  - $\text{tr}(\mathbf{AB}) = \text{tr}(\mathbf{B}'\mathbf{A}')$

  $\Rightarrow$ $\text{tr}(\mathbf{AB}) = \text{tr}(\mathbf{A}'\mathbf{B}')$

Orthogonal matrices

- An $n \times n$ matrix $\mathbf{A}$ is *orthogonal* if its columns, considered as vectors, form an orthonormal set.
- $\mathbf{A}$ is an *orthogonal matrix* if $\mathbf{A}'\mathbf{A} = \mathbf{I}$.
- From the previous result, we conclude that for orthogonal matrix $\mathbf{A}$, $\mathbf{A}^{-1} =$ ________________________________

Eigenvalues and eigenvectors

- Let $\mathbf{A}$ be an $n \times n$ matrix. A scalar λ is said to be an *eigenvalue* of $\mathbf{A}$ if there exists $\mathbf{x} \neq \mathbf{0}$ such that $\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$. The vector $\mathbf{x}$ is said to be an *eigenvector* of λ.
- Show that $\mathbf{x} = [2 \;\; 1]'$ is an eigenvector for $\mathbf{A} = \begin{bmatrix} 4 & -2 \\ 1 & 1 \end{bmatrix}$,

  and find the corresponding eigenvalue.

Characteristic equation

1. Recall from Slide 21, $\mathbf{A}^{-1} = \frac{1}{|\mathbf{A}|}\text{adj}\mathbf{A}$. It follows that $\mathbf{A}$ is singular if $|\mathbf{A}| = 0$.
2. Further, it can be shown that for any square matrix $\mathbf{A}$, the solution $\mathbf{x}$ to the matrix equation $\mathbf{A}\mathbf{x} = \mathbf{0}$ is non-zero only if $\mathbf{A}$ is singular.

- We can rewrite $\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$ as $\mathbf{A}\mathbf{x} - \lambda\mathbf{x} = \mathbf{0}$ or equivalently

$$(\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = \mathbf{0}.$$

  From 1 and 2 above, we know that a non-zero solution exists only if

$$|\mathbf{A} - \lambda\mathbf{I}| = 0.$$

- $|\mathbf{A} - \lambda\mathbf{I}| = 0$ is called the *characteristic equation* and is used to find the eigenvalues of a square matrix.

Finding eigenvalues and eigenvectors

Find the eigenvalues and corresponding eigenvectors for
$\mathbf{A} = \begin{bmatrix} 3 & 2 \\ 3 & -2 \end{bmatrix}$.

Quadratic forms

- Let $\mathbf{x}$ be an n-dimensional vector and let $\mathbf{A}$ be a symmetric $n \times n$ matrix. The scalar

$$\mathbf{x}'\mathbf{A}\mathbf{x}$$

  is called a *quadratic form*.
- E.g. Find the matrix associated with the quadratic form $3x_1^2 - 5x_1x_2 + x_2^2$ where $\mathbf{x} = [x_1 \;\; x_2]'$.

Positive definite matrices

A real symmetric matrix $\mathbf{A}$ is said to be

i. *Positive definite* if $\mathbf{x}'\mathbf{A}\mathbf{x} > 0$ for all nonzero $\mathbf{x}$

ii. *Negative definite* if $\mathbf{x}'\mathbf{A}\mathbf{x} < 0$ for all nonzero $\mathbf{x}$

iii. *Positive semi-definite* if $\mathbf{x}'\mathbf{A}\mathbf{x} \geq 0$ for all nonzero $\mathbf{x}$

iv. *Negative semi-definite* if $\mathbf{x}'\mathbf{A}\mathbf{x} \leq 0$ for all nonzero $\mathbf{x}$