Some Key Multivariate Principles

James H. Steiger

Department of Psychology and Human Development
Vanderbilt University

GCM, 2010

Some Key Multivariate Principles

Introduction

In this module, we review some key ideas from multivariate analysis. These include

1. Matrix Multiplication
2. Partitioned Matrix Forms
3. Matrix Expected Value Algebra
4. Covariance Matrices and Linear Combinations
5. Multiple Regression
6. Covariance Pattern Models
7. Component Models
8. Common Factor Models and Exploratory Common Factor Analysis
9. Confirmatory Factor Analysis
10. Structural Equation Modeling and the LISREL model
11. The Extended LISREL model

I will just touch on a few key ideas here. Extensive additional material is available in the course handouts and lecture notes for Psychology 310, Psychology 312, and Psychology 319(MLRM).

Matrix Multiplication
Conformability

Conformability

- Matrix multiplication is an operation with properties quite different from its scalar counterpart.
- *order matters* in matrix multiplication.
- The product $\boldsymbol{AB}$ will exist if and only if the number of columns of $\boldsymbol{A}$ is equal to the number of rows of $\boldsymbol{B}$.

Matrix Multiplication
Importance of Order

Importance of Order

- The matrix product $\boldsymbol{AB}$ need not be the same as the matrix product $\boldsymbol{BA}$.
- Indeed, the matrix product $\boldsymbol{AB}$ might be well-defined, while the product $\boldsymbol{BA}$ might not exist.
- When we compute the product $\boldsymbol{AB}$, we say that $\boldsymbol{A}$ is *post-multiplied* by $\boldsymbol{B}$, or that $\boldsymbol{B}$ is *premultiplied* by $\boldsymbol{A}$

Matrix Multiplication
Dimension of a Matrix Product

If two or more matrices are conformable, there is a strict rule for determining the dimension of their product

Matrix Multiplication — Dimensions of a Product

- The product ${}_p\boldsymbol{A}_q\boldsymbol{B}_r$ will be of dimension $p \times r$
- More generally, the product of any number of conformable matrices will have the number of rows in the leftmost matrix, and the number of columns in the rightmost matrix.
- For example, the product ${}_p\boldsymbol{A}_q\boldsymbol{B}_r\boldsymbol{C}_s$ will be of dimensionality $p \times s$

Matrix Multiplication
Three Approaches

Three Approaches

- Matrix multiplication might well be described as the key operation in matrix algebra
- What makes matrix multiplication particularly interesting is that there are numerous lenses through which it may be viewed
- We shall examine 3 ways of "looking at" matrix algebra
- All of them rely on *matrix partitioning*, which we'll examine briefly in the next 2 slides

Matrix Multiplication
The Row by Column Approach

Any $p \times q$ matrix $\boldsymbol{A}$ may be partitioned into as a set of p rows
For example, the 2×3 matrix

$$\begin{pmatrix} 1 & 2 & 3 \\ 3 & 3 & 3 \end{pmatrix} \tag{1}$$

may be thought of as two rows, $(1 \;\; 2 \;\; 3)$ and $(3 \;\; 3 \;\; 3)$ stacked on top of each other

We have a notation for this. We write

$$\boldsymbol{A} = \begin{pmatrix} \boldsymbol{a}_1' \\ \boldsymbol{a}_2' \end{pmatrix} \tag{2}$$

In future discussions, I will refer to the “dimension of the partitioned form” which in general will be different from the dimension of the matrix. For example, the matrix $\boldsymbol{A}$ above is a 2×3 matrix, but I have expressed it as a 2×1 partitioned form by writing it as two rows.

Matrix Multiplication
The Row by Column Approach

We can also view any $p \times q$ matrix as a set of q columns, joined side-by-side

For example, for the 2×3 matrix

$$\begin{pmatrix} 1 & 2 & 3 \\ 3 & 3 & 3 \end{pmatrix} \tag{3}$$

we can write

$$\boldsymbol{A} = \begin{pmatrix} \boldsymbol{a}_1 & \boldsymbol{a}_2 & \boldsymbol{a}_3 \end{pmatrix} \tag{4}$$

where, for example,

$$\boldsymbol{a}_1 = \begin{pmatrix} 1 \\ 3 \end{pmatrix} \tag{5}$$

Matrix Multiplication
The Row by Column Approach

Suppose you wish to multiply the two matrices $\boldsymbol{A}$ and $\boldsymbol{B}$, where

$$\boldsymbol{A} = \begin{pmatrix} 2 & 7 \\ 3 & 5 \end{pmatrix}, \quad \boldsymbol{B} = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 2 & 3 \end{pmatrix} \tag{6}$$

You know that the product, $\boldsymbol{C} = \boldsymbol{A}\boldsymbol{B}$, will be a 2×3 matrix

Partition $\boldsymbol{A}$ into 2 rows, and $\boldsymbol{B}$ into 3 columns. Element $c_{i,j}$ is the scalar product of row i of $\boldsymbol{A}$ with column j of $\boldsymbol{B}$

Matrix Multiplication
The Row by Column Approach

Again suppose you wish to compute the product $\boldsymbol{C} = \boldsymbol{AB}$ using the matrices from the preceding slide.

Example

Compute $c_{1,1}$.

$$\begin{pmatrix} 2 & 7 \\ 3 & 5 \end{pmatrix} \begin{pmatrix} 1 & 2 & 1 \\ 2 & 2 & 3 \end{pmatrix} \tag{7}$$

Taking the product of the row 1 of $\boldsymbol{A}$ and column 1 of $\boldsymbol{B}$, we obtain $(2)(1) + (7)(2) = 16$

Matrix Multiplication
The Row by Column Approach

Again suppose you wish to compute the product $\boldsymbol{C} = \boldsymbol{A}\boldsymbol{B}$ using the matrices from the preceding slide.

Example

Compute $c_{2,3}$.

$$\begin{pmatrix} 2 & 7 \\ 3 & 5 \end{pmatrix} \begin{pmatrix} 1 & 2 & 1 \\ 2 & 2 & 3 \end{pmatrix} \tag{7}$$

Taking the product of the row 2 of $\boldsymbol{A}$ and column 3 of $\boldsymbol{B}$, we obtain $(3)(1) + (5)(3) = 18$

Matrix Multiplication
Linear Combination of Columns Approach

Linear Combination of Columns

- When you post-multiply a matrix $\boldsymbol{A}$ by a matrix $\boldsymbol{B}$, each column of $\boldsymbol{B}$ generates, in effect, a column of the product $\boldsymbol{AB}$
- Each column of $\boldsymbol{B}$ contains a set of linear weights
- These linear weights are applied to the columns of $\boldsymbol{A}$ to produce a single column of numbers.

Matrix Multiplication
Linear Combination of Columns Approach

Consider the product

$$\begin{pmatrix} 2 & 7 \\ 3 & 5 \end{pmatrix} \begin{pmatrix} 1 & 2 & 1 \\ 2 & 2 & 3 \end{pmatrix} \tag{8}$$

The first column of the product is produced by applying the linear weights 1 and 2 to the columns of the first matrix

The result is

$$1 \begin{pmatrix} 2 \\ 3 \end{pmatrix} + 2 \begin{pmatrix} 7 \\ 5 \end{pmatrix} = \begin{pmatrix} 16 \\ 13 \end{pmatrix} \tag{9}$$

Matrix Multiplication
Linear Combination of Columns Approach

Consider once again the product

$$\begin{pmatrix} 2 & 7 \\ 3 & 5 \end{pmatrix} \begin{pmatrix} 1 & 2 & 1 \\ 2 & 2 & 3 \end{pmatrix} \tag{10}$$

The second column of the product is produced by applying the linear weights 2 and 2 to the columns of the first matrix

The result is

$$2 \begin{pmatrix} 2 \\ 3 \end{pmatrix} + 2 \begin{pmatrix} 7 \\ 5 \end{pmatrix} = \begin{pmatrix} 18 \\ 16 \end{pmatrix} \tag{11}$$

Matrix Multiplication

Linear Combination of Rows Approach

Linear Combination of Rows

- When you pre-multiply a matrix $\boldsymbol{B}$ by a matrix $\boldsymbol{A}$, each row of $\boldsymbol{A}$ generates, in effect, a rowof the product $\boldsymbol{AB}$
- Each row of $\boldsymbol{A}$ contains a set of linear weights
- These linear weights are applied to the rows of $\boldsymbol{B}$ to produce a single row vector of numbers.

Matrix Multiplication
Linear Combination of Rows Approach

Consider the product

$$\begin{pmatrix} 2 & 7 \\ 3 & 5 \end{pmatrix} \begin{pmatrix} 1 & 2 & 1 \\ 2 & 2 & 3 \end{pmatrix} \tag{12}$$

The first row of the product is produced by applying the linear weights 2 and 7 to the rows of the second matrix The result is

$$2 \begin{pmatrix} 1 & 2 & 1 \end{pmatrix} + 7 \begin{pmatrix} 2 & 2 & 3 \end{pmatrix} = \begin{pmatrix} 16 & 18 & 23 \end{pmatrix} \tag{13}$$

Matrix Multiplication
Mathematical Properties

The following are some key properties of matrix multiplication:

Mathematical Properties of Matrix Multiplication

- *Associativity.*
$$(\boldsymbol{AB})\boldsymbol{C} = \boldsymbol{A}(\boldsymbol{BC}) \tag{14}$$
- *Not generally commutative.* That is, often $\boldsymbol{AB} \neq \boldsymbol{BA}$.
- *Distributive over addition and subtraction.*
$$\boldsymbol{C}(\boldsymbol{A}+\boldsymbol{B}) = \boldsymbol{CA} + \boldsymbol{CB} \tag{15}$$
- Assuming it is conformable, the identity matrix $\boldsymbol{I}$ functions like the number 1, that is ${}_p\boldsymbol{A}_q\boldsymbol{I}_q = \boldsymbol{A}$, and ${}_p\boldsymbol{I}_p\boldsymbol{A}_q = \boldsymbol{A}$.
- $\boldsymbol{AB} = \boldsymbol{0}$ does not necessarily imply that either $\boldsymbol{A} = \boldsymbol{0}$ or $\boldsymbol{B} = \boldsymbol{0}$.

Partitioned Matrix Forms

A matrix may be composed of submatrices, as we have already seen in our discussion of matrix multiplication.

Being able to transpose and multiply a partitioned matrix is a skill that is important for understanding the key equations of advanced multivariate techniques.

Partitioned Matrix Forms

Assuming that the matrices are partitioned properly, the rules are quite simple:

1. To transpose a partitioned matrix, treat the sub-matrices in the partition as though they were elements of a matrix, but transpose each sub-matrix. The transpose of a $p \times q$ partitioned form will be a $q \times p$ partitioned form.
2. To multiply partitioned matrices, treat the sub-matrices as though they were elements of a matrix. The product of $p \times q$ and $q \times r$ partitioned forms is a $p \times r$ partitioned form.

Transposing a Partitioned Matrix

Some examples will illustrate the above definition.

Example (Transposing a Partitioned Matrix)

Suppose $\boldsymbol{A}$ is partitioned as

$$\boldsymbol{A} = \begin{bmatrix} \boldsymbol{C} & \boldsymbol{D} \\ \boldsymbol{E} & \boldsymbol{F} \\ \boldsymbol{G} & \boldsymbol{H} \end{bmatrix} \tag{16}$$

Then

$$\boldsymbol{A}' = \begin{bmatrix} \boldsymbol{C}' & \boldsymbol{E}' & \boldsymbol{G}' \\ \boldsymbol{D}' & \boldsymbol{F}' & \boldsymbol{H}' \end{bmatrix} \tag{17}$$

Product of Two Partitioned Matrices

Example (Product of two Partitioned Matrices)

Suppose $\boldsymbol{A} = \begin{bmatrix} \boldsymbol{X} & \boldsymbol{Y} \end{bmatrix}$ and $\boldsymbol{B} = \begin{bmatrix} \boldsymbol{G} \\ \boldsymbol{H} \end{bmatrix}$. Then

$\boldsymbol{AB} = \boldsymbol{X}\boldsymbol{G} + \boldsymbol{Y}\boldsymbol{H}$.

Expected Value of a Random Vector or Matrix

The expected value of a random vector (or matrix) is a vector (or matrix) whose elements are the expected values of the individual random variables that are the elements of the random vector.

Example (Expected Value of a Random Vector)

Suppose, for example, we have two random variables x and y, and their expected values are 0 and 2, respectively. If we put these variables into a vector $\boldsymbol{\xi}$, it follows that

$$E\left(\boldsymbol{\xi}\right) = \begin{bmatrix} 0 \\ 2 \end{bmatrix} \tag{18}$$

Matrix Expected Value Algebra

If $\boldsymbol{F}$ and $\boldsymbol{G}$ are matrices of constants and $\boldsymbol{X}$ and $\boldsymbol{Y}$ are random matrices, then

1. $E(\boldsymbol{X} + \boldsymbol{Y}) = E(\boldsymbol{X}) + E(\boldsymbol{Y})$
2. $E(\boldsymbol{F}\boldsymbol{X}) = \boldsymbol{F}E(\boldsymbol{X})$
3. $E(\boldsymbol{F}\boldsymbol{X}\boldsymbol{G}) = \boldsymbol{F}E(\boldsymbol{X})\boldsymbol{G}$

Essentially these rules mean that the expectation operator distributes over addition and/or subtraction, and that the operator and its associated parentheses "pass through" matrices of constants from the left and the right until encountering the first random matrix (or vector).

Expected Value Algebra

Example (Expected Value Algebra)

As an example of expected value algebra, we reduce the following expression. Suppose the Greek letters are random vectors with zero expected value, and the matrices contain constants. Then

$$\begin{aligned} E\left(\boldsymbol{A}'\boldsymbol{B}'\boldsymbol{\eta}\boldsymbol{\xi}'\boldsymbol{C}\right) &= \boldsymbol{A}'\boldsymbol{B}'E\left(\boldsymbol{\eta}\boldsymbol{\xi}'\right)\boldsymbol{C} \\ &= \boldsymbol{A}'\boldsymbol{B}'\boldsymbol{\Sigma}_{\boldsymbol{\eta}\boldsymbol{\xi}}\boldsymbol{C} \end{aligned}$$

Variance-Covariance Matrix of a Random Vector

Given a random vector $\boldsymbol{\xi}$ with expected value $\boldsymbol{\mu}$, the variance-covariance matrix $\boldsymbol{\Sigma}_{\boldsymbol{\xi}\boldsymbol{\xi}}$ is defined as

$$\begin{aligned} \boldsymbol{\Sigma}_{\boldsymbol{\xi}\boldsymbol{\xi}} &= E(\boldsymbol{\xi} - \boldsymbol{\mu})(\boldsymbol{\xi} - \boldsymbol{\mu})' \qquad (19) \\ &= E(\boldsymbol{\xi}\boldsymbol{\xi}') - \boldsymbol{\mu}\boldsymbol{\mu}' \qquad (20) \end{aligned}$$

If $\boldsymbol{\xi}$ is a deviation score random vector, then

$$\boldsymbol{\Sigma}_{\boldsymbol{\xi}\boldsymbol{\xi}} = E(\boldsymbol{\xi}\boldsymbol{\xi}') \qquad (21)$$

Variance-Covariance Matrix of a Random Vector

Comment. Equation 21 frequently is confusing to beginners. Let's "concretize" it a bit by giving an example with just two variables. Suppose

$$\boldsymbol{\xi} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \tag{22}$$

and

$$\boldsymbol{\mu} = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} \tag{23}$$

Note that $\boldsymbol{\xi}$ contains random variables, while $\boldsymbol{\mu}$ contains constants. Computing $E(\boldsymbol{\xi}\boldsymbol{\xi}')$, we find

$$\begin{aligned} E\left(\boldsymbol{\xi}\boldsymbol{\xi}'\right) &= E\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\begin{bmatrix} x_1 & x_2 \end{bmatrix}\right) \\ &= E\left(\begin{bmatrix} x_1^2 & x_1 x_2 \\ x_2 x_1 & x_2^2 \end{bmatrix}\right) \\ &= \begin{bmatrix} E(x_1^2) & E(x_1 x_2) \\ E(x_2 x_1) & E(x_2^2) \end{bmatrix} \end{aligned} \tag{24}$$

In a similar vein, we find that

$$\begin{aligned}\boldsymbol{\mu}\boldsymbol{\mu}' &= \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} \begin{bmatrix} \mu_1 & \mu_2 \end{bmatrix} \\ &= \begin{bmatrix} \mu_1^2 & \mu_1\mu_2 \\ \mu_2\mu_1 & \mu_2^2 \end{bmatrix} \end{aligned} \tag{25}$$

Subtracting Equation 25 from Equation 24, and recalling our basic definitions for covariances and variances, we find

$$\begin{aligned} E(\boldsymbol{\xi}\boldsymbol{\xi}') - \boldsymbol{\mu}\boldsymbol{\mu}' &= \begin{bmatrix} E(x_1^2) - \mu_1^2 & E(x_1 x_2) - \mu_1\mu_2 \\ E(x_2 x_1) - \mu_2\mu_1 & E(x_2^2) - \mu_2^2 \end{bmatrix} \\ &= \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{21} & \sigma_2^2 \end{bmatrix} \end{aligned}$$

Covariance Matrix for Two Random Vectors

Given two random vectors $\boldsymbol{\xi}$ and $\boldsymbol{\eta}$, their covariance matrix $\boldsymbol{\Sigma}_{\boldsymbol{\xi}\boldsymbol{\eta}}$ is defined as

$$\begin{aligned}
\boldsymbol{\Sigma}_{\boldsymbol{\xi}\boldsymbol{\eta}} &= E\left(\boldsymbol{\xi}\boldsymbol{\eta}'\right) - E(\boldsymbol{\xi})E(\boldsymbol{\eta}') &(26)\\
&= E\left(\boldsymbol{\xi}\boldsymbol{\eta}'\right) - E(\boldsymbol{\xi})E(\boldsymbol{\eta})' &(27)
\end{aligned}$$

Key Linear Combination Results

As a generalization of results we presented in scalar algebra, we find that, for a matrix of constants $\boldsymbol{B}$, and a random vector $\boldsymbol{x}$,

$$E\left(\boldsymbol{B}'\boldsymbol{x}\right) = \boldsymbol{B}'E(\boldsymbol{x}) = \boldsymbol{B}'\boldsymbol{\mu}$$

For random vectors $\boldsymbol{x}$ and $\boldsymbol{y}$, we find

$$E\left(\boldsymbol{x}+\boldsymbol{y}\right) = E(\boldsymbol{x}) + E(\boldsymbol{y}) \tag{28}$$

Comment. The result obviously generalizes to the expected value of the difference of random vectors.

Key Linear Combination Results

Given $\boldsymbol{x}$, a random vector with p variables, having variance-covariance matrix $\boldsymbol{\Sigma}_{\boldsymbol{xx}}$. The variance-covariance matrix of any set of linear combinations $\boldsymbol{y} = \boldsymbol{B}'\boldsymbol{x}$ may be computed as

$$\boldsymbol{\Sigma}_{\boldsymbol{yy}} = \boldsymbol{B}'\boldsymbol{\Sigma}_{\boldsymbol{xx}}\boldsymbol{B} \tag{29}$$

In a similar manner, we may prove the following:

Given $\boldsymbol{x}$ and $\boldsymbol{y}$, two random vectors with p and q variables having covariance matrix $\boldsymbol{\Sigma}_{\boldsymbol{xy}}$. The covariance matrix of any two sets of linear combinations $\boldsymbol{w} = \boldsymbol{B}'\boldsymbol{x}$ and $\boldsymbol{m} = \boldsymbol{C}'\boldsymbol{y}$ may be computed as

$$\boldsymbol{\Sigma}_{\boldsymbol{wm}} = \boldsymbol{B}'\boldsymbol{\Sigma}_{\boldsymbol{xy}}\boldsymbol{C} \tag{30}$$

Multiple Regression

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\epsilon} \tag{31}$$

with $\boldsymbol{X}$ assumed fixed, and $\boldsymbol{\epsilon} \sim N(\mathbf{0}, \sigma^2 \boldsymbol{I})$.

Random Multiple Regression

In this case the model is

$$\boldsymbol{y} = \boldsymbol{B}'\boldsymbol{x} + \boldsymbol{\epsilon} \tag{32}$$

with $\boldsymbol{x}$ and $\boldsymbol{y}$ having a multivariate normal distribution, and $E(\boldsymbol{x}\boldsymbol{\epsilon}') = \mathbf{0}$.

In this case, $\boldsymbol{B}' = \boldsymbol{\Sigma}_{yx}\boldsymbol{\Sigma}_{xx}^{-1}$

Covariance Pattern Models

$$\boldsymbol{y}_i = \boldsymbol{X}_i \boldsymbol{\beta} + \boldsymbol{\epsilon}_i \tag{33}$$

with $\boldsymbol{X}_i$ assumed fixed, and $\boldsymbol{\epsilon}_i \sim N(\mathbf{0}, \Sigma_i)$

This is essentially the regression model with the errors allowed to correlate.

Component Models

$$\boldsymbol{y} = \boldsymbol{F}\boldsymbol{x} + \boldsymbol{\epsilon} \tag{34}$$

with

$$\boldsymbol{x} = \boldsymbol{B}'\boldsymbol{y} \tag{35}$$

and $E(\boldsymbol{x}\boldsymbol{\epsilon}') = \boldsymbol{0}$