James B. McDonald
Brigham Young University
12/2011

I. Matrix algebra

1. **Basic definitions**
2. **Matrix operations**
   a. **Definitions**
   b. **Properties**
3. **Partitioned matrices**
4. **Quadratic forms**
5. **Kronnecker products**
6. **Characteristic roots and vectors**
7. **Vector and matrix differentiation with applications to optimization problems**
8. **Problem sets**

I. Matrix algebra

1. Basic definitions:

   matrix (order of matrix)

   square matrix

   transpose of a matrix

   symmetric matrix

   diagonal matrix

   identity matrix

   null or zero matrix

   row or column vector

   determinant of a square matrix

   inverse of a square matrix

   idempotent matrix

   orthogonal matrix

   trace of a square matrix

   rank of a matrix

2. Matrix operations

   a. Definitions

      (1) scalar multiplication

      (2) addition of matrices

      (3) multiplication of matrices

(4) inverses of nonsingular matrices

b. Some properties of matrix operations. Let c denote a real number and A,B,C denote matrices. <u>The following properties are conditional on the operations being defined for the case in point</u>.

<u>(1)</u> Scalar multiplication:

$c(A + B) = cA + cB = (A + B)c$

<u>(2)</u> Addition:

$A + B = B + A$
$A + (B + C) = (A + B) + C$
$A + 0 = 0 + A = A$

<u>(3)</u> Multiplication:

*$AB \neq BA$
*$AB = AC \neq> \text{(for all cases) } B = C$
$A(BC) = (AB)C$
$A(B + C) = AB + AC$
$(B + C)A = BA + CA$
$A0 = 0A = 0$
$AI = IA = A$

<u>(4)</u> Transposes and inverses:

$(A')' = A$
$(ABC)' = C'B'A'$
$(A + B)' = A' + B'$
$(A^{-1})^{-1} = A$
$(ABC)^{-1} = C^{-1}B^{-1}A^{-1}$, if A, B, and C have inverses
*$(A + B)^{-1} \neq A^{-1} + B^{-1}$
$(A^{-1})' = (A')^{-1}$

<u>(5)</u> Trace:
Trace (ABC) = Trace (CAB) = Trace (BCA)
Trace (A + B) = Trace (A) + Trace (B)

3. Partitioned Matrices (Greene, 4th ed., pp. 33-34)

Let $A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$ be an m x n matrix

where A_{11} is m_1 x n_1

A_{22} is m_2 x n_2

A_{12} is m_1 x n_2

A_{21} is m_2 x n_1

$n_1 + n_2 = n$

$m_1 + m_2 = m$

Also let B and C denote m x n and n x q matrices which are conformably partitioned as

$$B = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}$$

and

$$C = \begin{pmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{pmatrix}$$

so that the operations to be discussed are defined.

Addition of partitioned matrices:

$$A + B = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} + \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}$$

$$= \begin{pmatrix} A_{11} + B_{11} & A_{12} + B_{12} \\ A_{21} + B_{21} & A_{22} + B_{22} \end{pmatrix}$$

*A_{ij} and B_{ij} must be of the same dimension.

Multiplication of partitioned matrices:

$$AC = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \begin{pmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{pmatrix}$$

$$= \begin{pmatrix} A_{11}C_{11} + A_{12}C_{21} & A_{11}C_{12} + A_{12}C_{22} \\ A_{21}C_{11} + A_{22}C_{21} & A_{21}C_{12} + A_{22}C_{22} \end{pmatrix}$$

*The number of columns in A_{ij} must be the same as the number of rows in B_{jk}

Inverses of partitioned matrices:

Let A be a partitioned square matrix of order m x m

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$$

where A_{11} and A_{22} are nonsingular square matrices of order m_1 and m_2, respectively ($m_1 + m_2 = m$).

The <u>determinant</u> of A can be expressed by either of the following relations:

$$A = |A_{11}|\ |A_{22} - A_{21}A_{11}^{-1}A_{12}|$$

$$= |A_{22}|\ |A_{11} - A_{12}A_{22}^{-1}A_{21}|$$

If A is also nonsingular and

$$A^{-1} = B = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}, \text{ then}$$

the partitioned inverse of A, B can be expressed as

$$B_{11} = (A_{11} - A_{12}\, A^{-1}{}_{22}\, A_{21})^{-1}$$

$$B_{12} = - B_{11}\, A_{12}\, A^{-1}{}_{22}$$

$$B_{21} = - A^{-1}{}_{22}\, A_{21}\, B_{11}$$

$$B_{22} = A^{-1}_{22} + A^{-1}_{22} A_{21} B^{-1}_{11} A_{12} A_{22}$$

and an alternative form for the B_{ij} (inverse of A) is given by

$$B_{11} = A^{-1}_{11} + A^{-1}_{11} A_{12} B^{-1}_{22} A_{21} A_{11}$$

$$B_{12} = - A^{-1}_{11} A_{12} B_{22}$$

$$B_{21} = - B_{22} A_{21} A^{-1}_{11}$$

$$B_{22} = (A_{22} - A_{21} A^{-1}_{11} A_{12})^{-1}$$

Note that if the matrix A is

Block diagonal $A = \begin{pmatrix} A_{11} & 0 \\ 0 & A_{22} \end{pmatrix}$

or Block triangular $A = \begin{pmatrix} A_{11} & 0 \\ A_{21} & A_{22} \end{pmatrix}$ or $A = \begin{pmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{pmatrix}$

then the expressions for the determinants and inverses just considered are simplified considerably.

4. Quadratic forms and their classification.

Let A denote an n x n symmetric matrix with real entries and let X denote an n x 1 column vector.

Q = X'AX is said to be a quadratic form. Note that

$$Q = X'AX = (x_1 \ldots x_n) \begin{pmatrix} a_{11} \cdots a_{1n} \\ \vdots \qquad \vdots \\ a_{n1} \cdots a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

$$= (x_1 \ldots x_n) \begin{pmatrix} \sum a_{1i} x_i \\ \vdots \\ \sum a_{ni} x_i \end{pmatrix}$$

$$\begin{aligned} = \quad & a_{11}x_1^2 + a_{12}x_1x_2 + \ldots + a_{1n}x_1x_n \\ & + a_{21}x_2x_1 + a_{22}x_2^2 + \ldots + a_{2n}x_2x_n \\ & \quad \vdots \qquad\qquad \vdots \qquad\qquad\qquad \vdots \\ & + a_{n1}x_nx_1 + a_{n2}x_nx_2 + \ldots + a_{nn}x_n^2 \end{aligned}$$

Classification of the quadratic form $Q = X'AX$:

negative definite: $Q < 0$ if $x \neq 0$

negative semidefinite: $Q \leq 0$ for all x and $Q = 0$ for some $x \neq 0$

positive definite: $Q > 0$ if $x \neq 0$

positive semidefinite: $Q \geq 0$ for all x and $Q = 0$ for some $x \neq 0$

indefinite: $Q > 0$ for some x and $Q < 0$ for some other x

<u>A necessary and sufficient condition for positive or negative definiteness (A is symmetric)</u>:

Positive Definite	**Negative Definite**
$a_{11} > 0$	$a_{11} < 0$
$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} > 0$	> 0
$\vdots$	$\vdots$
$\begin{vmatrix} a_{11} a_{12} \cdots a_{1n} \\ a_{21} a_{22} \cdots a_{2n} \\ \vdots \quad \vdots \\ a_{n1} a_{n2} \cdots a_{nn} \end{vmatrix} > 0$	(has sign of $(-1)^n$)
(all positive determinants)	(first negative, then alternate in sign)

<u>Note</u>: In the discussion on regression theory it will be useful to note that <u>if</u> X is N x K (N > K) and if $(X'X)^{-1}$ exists, <u>then</u> (X'X) is positive definite.

5. Kronnecker Products

Let $A = (a_{ij})$ m x n

$B = (b_{ij})$ p x q,

then the Kronnecker product of A and B is denoted by A $\otimes$ B and is defined by

$$A \otimes B = \begin{pmatrix} a_{11}B & a_{12}B & \dots & a_{1n}B \\ a_{21}B & a_{22}B & \dots & a_{2n}B \\ . & . & & . \\ & & \dots & \\ . & . & & . \\ a_{m1}B & a_{m2}B & \dots & a_{mn}B \end{pmatrix}_{mp \times nq}$$

Properties: (matrices are assumed to be conformable)

1. $(A \otimes C)(B \otimes D) = (AB) \otimes (CD)$
2. $(P \otimes Q)^{-1} = P^{-1} \otimes Q^{-1}$, P and Q square and nonsinguar
3. $(M \otimes N)' = M' \otimes N'$
4. $|A \otimes B| = |A|^p \, |B|^m$ A and B are square of order m x m and p x p.
5. trace $(A \otimes B)$ = trace (A) trace (B)

Example:

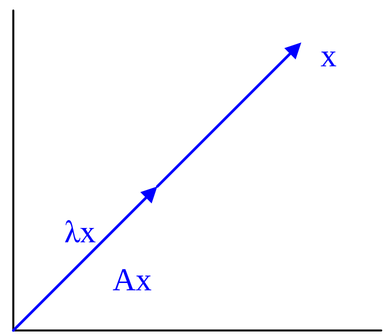
$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \otimes \begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix} = \begin{pmatrix} \begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix} & 2\begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix} \\ 3\begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix} & 4\begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix}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

$$= \begin{pmatrix} 5 & 6 & 10 & 12 \\ 7 & 8 & 14 & 16 \\ 15 & 18 & 20 & 24 \\ 21 & 24 & 28 & 32 \end{pmatrix}$$

6. Characteristic Roots and Vectors

Let A denote an n x n matrix. Consider the matrix equation:

$AX = \lambda X$



Given the nxn matrix A, it is frequently useful to determine (nx1) vectors X and corresponding scalars such that $AX = \lambda X$, i.e.

$$AX - \lambda X = (A - \lambda I)X = 0$$

The only way in which we can obtain a nontrivial solution ($X \neq 0$) to this system of equations is for $|A - \lambda I| = 0$.

The characteristic equation associated with the matrix A is defined by

$$|A - \lambda I| = 0.$$

The characteristic equation is an n^{th} degree polynomial in λ and will, by the fundamental theorem of algebra, have n solutions ($\lambda_1, \ldots, \lambda_n$) which are referred to as the characteristic (latent or eigen) roots of the matrix A. The λ_i are also referred to as eigen values and may be real or possibly imaginary. The characteristic roots will be real if A is a real symmetric matrix.

The characteristic (latent or eigen) vectors associated with A are determined by solving

$$AV_i = \lambda_i V_i \qquad i = 1,2, \ldots, n$$

where λ_i denotes the associated characteristic root and V_i is the i^{th} characteristic vector. Note: $A(cV_i) = \lambda_i(cV_i)$ and the characteristic vectors are only unique up to a scalar multiple.

Some Important Properties. Let $A = (a_{ij})$ be an n x n matrix with characteristic roots λ_i (i = 1, 2, .., n)

(1) $\lambda_1 + \lambda_2 + \ldots + \lambda_n = \text{trace}(A) = a_{11} + \ldots + a_{nn}$

(2) $\lambda_1 \lambda_2 \ldots \lambda_n = |A|$

Some additional properties are useful for symmetric matrices.

(1) A (X'AX) is positive definite <=> $\lambda_i > 0$ for all i

(2) A (X'AX) is negative definite <=> $\lambda_i < 0$ for all i

(3) If $\lambda_i \neq \lambda_j$ and A is symmetric, then $V_i' \cdot V_j = 0$

(4) If A is symmetric and if the characteristic vectors are chosen to have length one, then the matrix $V = (V_1, \ldots, V_n)$ is orthogonal and

$$AV = V\Lambda \text{ where } \Lambda = \begin{pmatrix} \lambda_1 & & 0 \\ & & \\ 0 & & \lambda_n \end{pmatrix}$$

It follows that

$$V'AV = \begin{pmatrix} \lambda_1 & & 0 \\ & & \\ 0 & & \lambda_n \end{pmatrix}$$

Note:

(1) This may be true even if not all of the characteristic roots are unique.

(2) This allows for the diagonalization of quadratic forms
$Q = X'AX = (X'VV'AVV'X)$
$= Z'\Lambda Z = \Sigma \lambda_i z_i^2$
where Z = X'V. This is a very important result.

(3) Rank (A) = the number of nonzero characteristic roots

(4) Characteristic roots of A^k are λ^k.

(5) If A is an idempotent matrix, then its characteristic roots are zero or one and

$$\text{Rank (A)} = \text{trace (A)}$$

(6) The condition number of the matrix A is defined to be

$$\left(\frac{\text{max root}}{\text{min root}}\right)^{1/2}$$

Large values of the condition number can indicate near singularity of the matrix A and is sometimes used in econometrics where A is the correlation matrix associated with the explanatory variables.

7. Vector and matrix differentiation with applications to constrained and unconstrained optimization problems.

a. Basic definitions

Let $f(X) = f(x_1, x_2, ..., x_n) = y$ be a real valued function of the vector X. The derivative of f(X) with respect to the vector X will be defined by

$$\frac{df}{dX} = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, ..., \frac{\partial f}{\partial x_n}\right)'$$

and the second derivative of f(x) with respect to X is defined by

$$\frac{d^2 f}{dX^2} = \frac{d}{dX}\left(\frac{df}{dX}\right) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1 \partial x_1} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \vdots & & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_n} \end{pmatrix}$$

which is known as the Hessian matrix. This suggests an obvious definition of the derivative of a real valued function with respect to a matrix, i.e., the derivative of a real valued function of a matrix is defined to be the matrix of derivatives of corresponding elements.

b. Unconstrained optimization

Consider the optimization problem:

$$\underset{X}{\text{maximize}}\ f(X).$$

The necessary (first order) conditions for a maximum or minimum are given by

$$\frac{df}{dX} = 0$$

The sufficient or second order condition for a maximum (minimum) is that

$$\frac{d^2 f}{dX^2}$$

is negative (positive) definite.

These results are readily obtained from a Taylor Series expansion of f(x),

$$f(x) \approx f(x_0) + \left(\frac{d'f}{dX}\right)(X- X_0) + \frac{1}{2}(X- X_0)' \frac{d^2 f}{dX^2}(X- X_0).$$

Hint : X_o is selected to satisfy the necessary condition.

c. Constrained optimization

Consider the <u>constrained optimization</u> problem defined by

$$\underset{X}{\text{maximize}}\ f(X)$$

subject to g(X) = 0.

where g(X) = 0 denotes a m x 1 vector of constraints, m < n. The solution can be obtained from the Lagrangian function

$$L(X;\lambda) = f(X) + \lambda' g(X)$$

where $\lambda' = (\lambda_1, ..., \lambda_m)$.

The <u>necessary</u> (first order) conditions for a solution to this problem are that

$$\frac{\partial L}{\partial X} = f_X + (g_X)'\lambda = 0$$

$$\frac{\partial L}{\partial \lambda} = g(X) = 0.$$

Sufficient conditions for a maximum or minimum can be stated in terms of the Hessian of the Lagrangian function (or bordered Hessian)

$$H = \frac{d^2 L}{d^2(\lambda, X)} = \begin{bmatrix} \frac{\partial^2 L}{\partial \lambda^2} & \frac{\partial^2 L}{\partial \lambda \partial X} \\ \frac{\partial^2 L}{\partial X \partial \lambda} & \frac{\partial^2 L}{\partial X^2} \end{bmatrix}$$

$$= \begin{bmatrix} 0 & g_{X'} \\ g_X & L_{XX} \end{bmatrix}$$

A sufficient condition for a <u>minimum</u> is that the determinants of border preserving principal minors of H have sign $(-1)^m$ or zero. If |H| is of sign $(-1)^n$ and the determinants of border preserving principal minors are zero or alternate in sign, then a sufficient condition for a <u>maximum</u> is satisfied.

<u>Some Useful Results</u> (Matrix Cookbook, http://matrixcookbook.com)

Let $a = (a_1, ..., a_n)'$,

$A = (a_{ij})_{n \times n}$ and $X = (x_1, ..., x_n)'$.

1. $$\frac{d(a'X)}{dX} = a$$

2. $$\frac{d(X'AX)}{dX} = 2AX \text{ (A is symmetric); } = (A + A')X \text{ otherwise}$$

3. $$\frac{d^2(X'AX)}{dX^2} = 2A \text{ (A is symmetric); } (A + A') \text{ otherwise.}$$

4. $$\frac{\partial trace(A)}{\partial A} = I$$

5. $$\frac{\partial|A|}{\partial A} = \begin{cases} |A|\left(A'\right)^{-1} & \text{if } |A| \neq 0 \\ 0 & \text{if } |A| = 0 \end{cases}$$

6. $$\frac{\partial \log|A|}{\partial A} = \left(A'\right)^{-1}$$

Hint: $|A| = \Sigma\, a_{ij}\ |C_{ij}|$

7. $$\frac{dA^{-1}(x)}{dx} = -A^{-1}\frac{dA(x)}{dx}A^{-1}$$

8. $$\frac{\partial(X'AX)}{\partial A} = XX'$$

<u>Example</u>

Consider $f(X) = (X - \mu)'B(X - \mu)$ where B is positive definite and symmetric. Determine the value of X which minimizes f(X). Check the sufficient conditions.

$$f(X) = (X - \mu)'B(X - \mu) = X'BX - 2X'B\mu + \mu'B\mu$$

$$\frac{df}{dx} = 2Bx - 2B\mu = 2B(X - \mu) = 0$$

$$\frac{d^2f}{dx^2} = 2B$$

Since $B \neq 0$, the solution is $X = \mu$. Sufficient conditions for a minimum are satisfied if B is positive definite.

8. Problems sets

1. Expand (A + B) (A - B) and (A - B) (A + B). Are these expansions the same? If not, why not?

2. Given $A = \begin{pmatrix} 1 & 0 & 3 \\ 2 & -1 & 1 \end{pmatrix}$, $B = \begin{pmatrix} 3 & 4 & 1 \\ 0 & -1 & 5 \\ 0 & 2 & -2 \end{pmatrix}$, $C = \begin{pmatrix} 2 \\ -1 \\ 4 \end{pmatrix}$

   Calculate (AB)', B'A', C'A' and (AC)'.

3. Prove that diagonal matrices of the same order are commutative in multiplication with each other.

4. Prove that

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{vmatrix} = a_{11} a_{22} \cdots a_{nn}$$

5. Show that the matrix

$$Q = \begin{bmatrix} 1/\sqrt{6} & 2/\sqrt{5} & 1/\sqrt{30} \\ -2/\sqrt{6} & 1/\sqrt{5} & -2/\sqrt{30} \\ 1/\sqrt{6} & 0 & -5/\sqrt{30} \end{bmatrix}$$

is orthogonal, i.e., $Q' = Q^{-1}$.

6. Determine whether the following quadratic forms are positive definite:

(a) $6x_1^2 + 49x_2^2 + 51x_3^2 - 82x_2x_3 + 20x_1x_3 - 4x_1x_2$

(b) $4x_1^2 + 9x_2^2 + 2x_3^2 + 8x_2x_3 + 6x_3x_1 + 6x_1x_2$

7. Prove that

$$A = \begin{pmatrix} 1/2 & 1 \\ 1/4 & 1/2 \end{pmatrix}$$

is a nonsymmetric, idempotent matrix.

8. Let X denote an N x K (N > K) matrix. Demonstrate that $B = I_N - X(X'X)^{-1}X'$ is symmetric and idempotent. You will see this matrix again.

9. Obtain the characteristic roots of the matrix

$$C = \begin{pmatrix} 5/2 & 1/2 \\ 1/2 & 5/2 \end{pmatrix}$$

Can you determine the sign of $(x_1,x_2)\ C\ (x_1,x_2)'$ for arbitrary $X = (x_1, x_2) \neq 0$? Defend your answer.

10. Using the technique of inverses of partitioned matrices, determine the inverse of

$$D = \left[\begin{array}{cc|c} 1 & 0 & 0 \\ 0 & 5 & 2 \\ \hline 0 & 2 & 1 \end{array}\right] = \begin{bmatrix} D_{11} & D_{12} \\ D_{21} & D_{22} \end{bmatrix}$$

11. Let

$$A = \begin{pmatrix} 5 & 2 \\ 2 & 1 \end{pmatrix} \text{ and } B = \begin{pmatrix} 3 & 5 \\ 1 & 2 \end{pmatrix}$$

Evaluate:

$A \otimes B$

$(A \otimes B)^{-1}$

$(A \otimes B)'$

trace $(A \otimes B)$

$|A \otimes B|$

Some additional fun exercises:

12. Let A and B be square matrices. Prove that

$$\text{trace}(AB) = \text{trace}(BA)$$

Hint: The ij^{th} element in the matrix product AB is given by $\Sigma_k a_{ik} b_{kj}$.

13. Prove:

a. $(P \otimes Q)^{-1} = P^{-1} \otimes Q^{-1}$

b. $|A \otimes B| = |A|^p |B|^m$ where A and B are respectively mxm and pxp matrices

c. $\text{Trace}(A \otimes B) = \text{trace}(A)\,\text{trace}(B)$

14. Prove that the characteristic vectors corresponding to unique (unequal) characteristic roots of a symmetric matrix A are orthogonal.

15. Prove that the characteristic roots of a real symmetric matrix are real.

16. Prove $\dfrac{\partial\, \text{trace}(A)}{\partial A} = I$.

17. a. Prove $\Sigma\lambda_i = \text{trace}(A)$ and $\Pi\lambda_i = |A|$.

b. Prove that the characteristic roots of A^k are λ^k.

c. Prove that the characteristic roots of an idempotent matrix are either zero or one.

18.. Determine an upper triangular matrix T, such that

$$T'T = \begin{pmatrix} 1 & 2 \\ 2 & 13 \end{pmatrix}$$

Any symmetric positive definite matrix (A) can be written as a product of a lower triangular matrix and its transpose (upper triangular matrix). This is referred to as the Cholesky factorization of A.

References

Green, William H., Econometric Analysis, 4th edition, New Jersey: Prentice Hall, 2000.

Hadley, G., Linear Algebra, Palo Alto: Addison-Wesley, 1961.

Magnus, Jan R. and Heinz Neudecker, Matrix Differential Calculus with Application in Statistics and Econometrics, New York: J. Wiley & Sons, 1988.

Petersen, K. B. and M. S. Pedersen, The Matrix Cookbook, http://matrixcookbook.com.

Searle, Shayler R, Matrix Algebra useful for Statistics, New York: J. Wiley and Sons, 1982.