Exercises with Matrices

Part One – Practice with Numbers (if there is *no* answer, say so)

1. $\begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} + \begin{bmatrix} 10 & 20 \\ 30 & 40 \\ 50 & 60 \end{bmatrix}$

2. $\begin{bmatrix} 1 & 2 \end{bmatrix} \begin{bmatrix} 3 \\ 4 \end{bmatrix}$

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

4. $\begin{bmatrix} 1 & 2 \\ 10 & 20 \end{bmatrix} \begin{bmatrix} 3 & 0 \\ 4 & 1 \end{bmatrix}$

5. $\begin{bmatrix} 1 & 2 & 7 \end{bmatrix} \begin{bmatrix} 3 \\ 4 \end{bmatrix}$

6. $\begin{bmatrix} 3 \\ 4 \end{bmatrix} \begin{bmatrix} 1 & 2 & 7 \end{bmatrix}$

7. $\begin{bmatrix} 0 & 1 & 2 \\ 10 & -10 & 5 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$

8. $\begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$

   Note: The first matrix above is called a "permutation" matrix. See how it permutes the rows of the right-hand matrix? A permutation matrix is a square matrix that consists of all 0s or 1s, with a single 1 in each row and a single 1 in each column.

9. $5 \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix}$

10. $-1 \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ -5 & -6 \end{bmatrix}$

11. $\begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} - \begin{bmatrix} 10 & 20 \\ 30 & 40 \\ 50 & 60 \end{bmatrix}$

12. $\begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix}'$

13. $\begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix}^{T}$

14. $\left(\begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix}^{T} \right)^{T}$

15. Solve for x and y: $\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$

16. Solve for x and y: $\begin{bmatrix} 1 & y \\ 3 & 4 \\ x & 6 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix}$

17. $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}^{-1}$

18. $\begin{bmatrix} 1 & 2 \\ -3 & 4 \end{bmatrix}^{-1}$

19. Questions about determinants. In each case, find the determinant and indicate whether the matrix is singular. Note that, for example, $\left| \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \right| = \begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix}$.

    a. $\left| \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \right|$

    b. $\left| \begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix} \right|$

    c. $\left| \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 5 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix} \right|$

    d. $\left| \begin{bmatrix} 2 & 8 & -4 & 21 \\ 0 & 3 & 5 & 6 \\ 0 & 0 & 5 & -1 \\ 0 & 0 & 0 & -1 \end{bmatrix} \right|$

Part Two – Types of Matrices – Give the <u>most specific</u> name for each matrix, from the following choices: column

matrix, diagonal matrix, identity matrix, lower triangular matrix, permutation matrix, row matrix, square matrix, symmetric matrix, upper triangular matrix.

20. $\begin{bmatrix} a & b \\ 7 & 12 \end{bmatrix}$

21. $\begin{bmatrix} 3 \\ 4 \end{bmatrix}$

22. $\begin{bmatrix} 1 & 2 & 3 \end{bmatrix}$

23. $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

24. $\begin{bmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{bmatrix}$

25. $\begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 5 \\ 3 & 5 & 6 \end{bmatrix}$

26. $\begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$

27. $\begin{bmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{bmatrix}$

28. $\begin{bmatrix} 1 & 0 & 0 \\ 2 & 4 & 0 \\ 3 & 5 & 6 \end{bmatrix}$

Part Three – Understanding Symbolic Matrix Manipulation

In this part, think of $\mathbf{x}$, $\mathbf{y}$, and $\mathbf{X}$ as matrices containing economic data. The *sample mean* or average of n numbers is $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$. The *sample variance* of the numbers is a measure of how much they vary: $\sigma_{x^2} = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$. The *sample covariance* compares two different kinds of numbers, such as income and health, and is a measure of how much they vary together: $\sigma_{xy} = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})$. You should recognize these concepts in some of the algebraic expressions you study below.

$$\text{Let } \mathbf{i} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}, \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \mathbf{y}' = \begin{bmatrix} y_1 & y_2 & \cdots & y_n \end{bmatrix}, \text{ and } \mathbf{X} = \begin{bmatrix} x_{11} & \cdots & x_{1k} \\ x_{21} & \cdots & x_{2k} \\ \vdots & & \vdots \\ x_{n1} & \cdots & x_{nk} \end{bmatrix}.$$

For each item below, write out how the answer (usually a matrix) looks in detail, using ellipses (…) where necessary. What is:

29. $\frac{1}{n}\mathbf{ii}'$
30. $\frac{1}{n}\mathbf{ii}'\mathbf{x}$
31. $\mathbf{x} - \frac{1}{n}\mathbf{ii}'\mathbf{x}$
32. $\left[\mathbf{I} - \frac{1}{n}\mathbf{ii}'\right]\mathbf{x}$
33. $\left[\mathbf{I} - \frac{1}{n}\mathbf{ii}'\right]\mathbf{i}$
34. $\mathbf{M} \equiv \mathbf{I} - \frac{1}{n}\mathbf{ii}'$
35. $\mathbf{i}'\mathbf{x}$
36. $\mathbf{i}'(\mathbf{Mx})$
37. $\mathbf{MM}$
38. $\mathbf{x}'\mathbf{x}$
39. $(\mathbf{Mx})'(\mathbf{Mx})$
40. $\mathbf{x}'\mathbf{y}$
41. $(\mathbf{Mx})'(\mathbf{My})$
42. $\frac{1}{n-1}(\mathbf{Mx})'(\mathbf{My})$
43. $\mathbf{MX}$
44. $(\mathbf{MX})'(\mathbf{MX})$

45. $\frac{1}{n-1}(\mathbf{MX})'(\mathbf{MX})$

Part Four – Symbolic Matrix Manipulation

Simplify all expressions. Assume that matrices **A** and **B** are symmetric. Assume that inverses exist for matrices **A**, **B**, **C**, and **D**. (By the way, note that the inverse of a symmetric matrix is symmetric.)

46. $\mathbf{I}'$
47. $\mathbf{C}^{-1}\mathbf{C}$
48. $(\mathbf{C}^{-1})^{-1}$
49. $(\mathbf{X}'\mathbf{X})^{-1}(\mathbf{X}'\mathbf{X})$
50. $(\mathbf{CD})'$
51. $(\mathbf{CD})^{-1}$
52. $\mathbf{A}^{-1}\mathbf{A}'$
53. $\mathbf{C}^{-1}\mathbf{C}'$
54. $\mathbf{B}^{-1}(\mathbf{AB})'\mathbf{A}^{-1}$

Part Five – Quadratic Forms

For each of the following expressions, determine whether it is a quadratic form. If it is, write the expression in the matrix form $\mathbf{x}'\mathbf{A}\mathbf{x}$, where **x** is a vector and **A** is a symmetric matrix. Greek letters denote parameters that do not involve the x's.

55. $x_1^2 + 3x_1x_2 + 2x_2^2$
56. $x_1x_2 + x_1x_3 + x_2x_3$
57. $x_1^2 + 2x_2^2 + 7x_1$
58. x_1^2
59. x_1x_2
60. $x_1x_2x_3$
61. $x_1^2 + x_1x_2x_3$
62. $3x_1^2 + x_2^2 + 7x_3^2 + x_1x_2 - 12x_2x_3 - \frac{5}{\pi}x_1x_3$
63. $\alpha x_1^2 + \beta x_2^2 + \gamma x_3^2 + \frac{\delta}{\rho}x_1x_2 + \phi x_1x_3 + \sigma x_2x_3$

Part Six – Matrix Derivatives

If you have forgotten what derivatives are or basic formulae for derivatives, please review them. *Total versus partial derivatives*: In a total derivative, you take into account that one variable may be a function of another; for example, if $c = 2b$ and $f(b,c) = b^2 + 3c$, then $\frac{df(b,c)}{db} = 2b + 6$. In a partial derivative, you deliberately hold constant all variables but one; for example, $\frac{\partial f(b,c)}{\partial b} = 2b$ because you deliberately assume that c is being held fixed (by means which depend on the context). *Key rules to master*: Know by memory basic rules for derivatives, including eventually the chain rule. Self test: $\frac{\partial(a+bu^n)}{\partial u}$, $\frac{\partial \ln(ku)}{\partial u}$, $\frac{\partial(u^2e^{ku})}{\partial u}$, $\frac{\partial^2(u^7)}{\partial u^2}$, $\frac{\partial \exp(k(\ln u)^{107})}{\partial u}$ (answers on last page). The final one of these expressions definitely requires the chain rule.

As long as you know these basics, matrix derivatives are not too hard, especially what we do here. Suppose $\mathbf{u}$ is a vector, $\mathbf{u} = \begin{bmatrix} u_1 & \cdots & u_n \end{bmatrix}'$, and $f(\mathbf{u})$ is a function whose value depends on $u_1,\ldots,u_n$. If you need to know all n derivatives $\frac{\partial f(\mathbf{u})}{\partial u_1},\ldots,\frac{\partial f(\mathbf{u})}{\partial u_n}$, it is convenient to write all the derivatives in a single vector. The vector is written $\frac{\partial f(\mathbf{u})}{\partial \mathbf{u}}$, and is equal to $\begin{bmatrix} \frac{\partial f(\mathbf{u})}{\partial u_1} \\ \vdots \\ \frac{\partial f(\mathbf{u})}{\partial u_n} \end{bmatrix}$. The next two questions show you two useful formulae (used in chapter 1 of Hayashi) for computing such vector derivatives.

64. Let $\mathbf{a}$ be a vector of n constants $\mathbf{a} = \begin{bmatrix} a_1 & \cdots & a_n \end{bmatrix}'$, and let $\mathbf{u} = \begin{bmatrix} u_1 & \cdots & u_n \end{bmatrix}'$. First, write out $\mathbf{a}'\mathbf{u}$ in terms of the elements a_1 through a_n and u_1 through u_n (your answer should be a scalar). Second, compute the derivative $\frac{\partial \mathbf{a}'\mathbf{u}}{\partial u_i}$. Third, write the whole vector of derivatives $\begin{bmatrix} \frac{\partial \mathbf{a}'\mathbf{u}}{\partial u_1} & \cdots & \frac{\partial \mathbf{a}'\mathbf{u}}{\partial u_n} \end{bmatrix}'$. You should have just proved that $\frac{\partial(\mathbf{a}'\mathbf{u})}{\partial \mathbf{u}} = \mathbf{a}$.

65. Let $\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$, and let $\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$. Prove that $\frac{\partial(\mathbf{u}'\mathbf{A}\mathbf{u})}{\partial \mathbf{u}} = 2\mathbf{A}\mathbf{u}$ if $\mathbf{A}$ is symmetric, and prove that this formula does not work if $\mathbf{A}$ is not symmetric. (This formula in fact works for any n×n symmetric matrix $\mathbf{A}$ and n×1 vector $\mathbf{u}$.)

Above you needed to differentiate a single value ("scalar"), which we called $f(\mathbf{u})$, with respect to each element in an n×1 vector, which we called $\mathbf{u}$, giving n derivatives arranged in a vector. The same idea extends to differentiating m different values, in a vector $\mathbf{f}$, with respect to each element in the n×1 vector, $\mathbf{u}$, giving mn derivatives arranged in a matrix. Let $\mathbf{f} = \begin{bmatrix} f_1(\mathbf{u}) \\ \vdots \\ f_m(\mathbf{u}) \end{bmatrix}$ be a vector containing the m different values, which may be formulae involving $\mathbf{u}$ and hence have been written $f_1(\mathbf{u}),\ldots,f_m(\mathbf{u})$. The expression $\frac{\partial \mathbf{f}}{\partial \mathbf{u}'}$ means the following matrix of derivatives:

$\begin{bmatrix} \frac{\partial f_1(\mathbf{u})}{\partial \mathrm{u}_1} & \cdots & \frac{\partial f_1(\mathbf{u})}{\partial \mathrm{u}_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m(\mathbf{u})}{\partial \mathrm{u}_1} & \cdots & \frac{\partial f_m(\mathbf{u})}{\partial \mathrm{u}_n} \end{bmatrix}$. This matrix has m rows and n columns, i.e. one row for each entry in **f** and one column for each variable in **u**. Note that **u** is transposed at the bottom of the expression $\frac{\partial \mathbf{f}}{\partial \mathbf{u}'}$, thus emphasizing that the values $\mathrm{u}_1, \ldots, \mathrm{u}_n$ differ from left to right (not top to bottom) in the matrix. (Some authors write $\frac{\partial \mathbf{f}}{\partial \mathbf{u}}$ to mean the same matrix, but Hayashi's textbook uses $\frac{\partial \mathbf{f}}{\partial \mathbf{u}'}$ since it is clearer in this regard.)

66. Let $\mathbf{f} = \begin{bmatrix} 2\mathrm{u}_1 + 4\mathrm{u}_2 + 6\mathrm{u}_3 \\ 200 + \mathrm{u}_1^2\mathrm{u}_2^3\mathrm{u}_3^4 \end{bmatrix}$, and let $\mathbf{u} = \begin{bmatrix} \mathrm{u}_1 \\ \mathrm{u}_2 \\ \mathrm{u}_3 \end{bmatrix}$. Compute the matrix $\frac{\partial \mathbf{f}}{\partial \mathbf{u}'}$, writing out each of the derivatives in the matrix.

Finally, if you start with a single value ("scalar") $f(\mathbf{u})$ and differentiate *twice* with respect to each element in **u**, it is convenient to arrange the resulting derivatives in a matrix. The expression $\frac{\partial^2 f(\mathbf{u})}{\partial \mathbf{u} \partial \mathbf{u}'}$ means the matrix $\begin{bmatrix} \frac{\partial^2 f(\mathbf{u})}{\partial \mathrm{u}_1^2} & \cdots & \frac{\partial^2 f(\mathbf{u})}{\partial \mathrm{u}_n \partial \mathrm{u}_1} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f(\mathbf{u})}{\partial \mathrm{u}_1 \partial \mathrm{u}_n} & \cdots & \frac{\partial^2 f(\mathbf{u})}{\partial \mathrm{u}_n^2} \end{bmatrix}$. This matrix is given the name "Hessian matrix," which you will see sometimes. This matrix is symmetric as long as the second derivatives in it exist and are continuous (at the values of **u** being considered), because then "Clairaut's theorem" proves that $\frac{\partial^2 f(\mathbf{u})}{\partial \mathrm{u}_i \partial \mathrm{u}_j} = \frac{\partial^2 f(\mathbf{u})}{\partial \mathrm{u}_j \partial \mathrm{u}_i}$. The Hessian matrix is used in maximization and minimization.

67. Let $f(\mathbf{u}) = \mathrm{u}_1^2 + \mathrm{u}_2\mathrm{u}_3^4$, where $\mathbf{u} = \begin{bmatrix} \mathrm{u}_1 \\ \mathrm{u}_2 \\ \mathrm{u}_3 \end{bmatrix}$. Compute the Hessian matrix $\frac{\partial^2 f(\mathbf{u})}{\partial \mathbf{u} \partial \mathbf{u}'}$, writing out each of the derivatives in the matrix.

68. Let $\mathbf{v} = \begin{bmatrix} \mathrm{v}_1 \\ \mathrm{v}_2 \\ \mathrm{v}_3 \end{bmatrix}$, and let **A** be the matrix you computed in the previous question. Write out $\mathbf{v}'\mathbf{A}\mathbf{v}$.

Maximization and minimization problems can be solved (if first and second derivatives exist) using first and second order conditions. Suppose you are maximizing or minimizing $f(\mathbf{u})$, by choosing the values of $u_1,\ldots,u_n$. The first order condition is that each of the n derivatives in $\frac{\partial f(\mathbf{u})}{\partial \mathbf{u}}$ must be zero, i.e., $\frac{\partial f(\mathbf{u})}{\partial \mathbf{u}} = \mathbf{0}$. (Since a derivative is a slope, this just says that the slope of the function must be completely flat, in each of the n directions, at the chosen values of $u_1,\ldots,u_n$, just as there is a flat spot at the top of Mount Everest or at the bottom of the Mariana Trench). Solving these equations gives values for $u_1,\ldots,u_n$.

The second order conditions, always applied after solving the first order condition, comes in different versions. Version A. If met, this condition guarantees you have found a global* maximum (minimum). The condition is that f(**u**) is concave (convex) everywhere, i.e. that the function is dome-shaped or bowl-shaped. This is guaranteed if *for all possible values of* ***u***, the Hessian matrix $\frac{\partial^2 f(\mathbf{u})}{\partial \mathbf{u} \partial \mathbf{u}'}$ is negative (positive) semidefinite. For any $n \times n$ matrix **A** to be negative (positive) semidefinite means that $\mathbf{v}'\mathbf{A}\mathbf{v}$ is always ≤ 0 (≥ 0) for any possible $n \times 1$ vector of real numbers **v**. You don't need to know how to determine this, but if you are interested it is discussed in textbooks on quantitative methods for economics or in linear algebra texts. To summarize, to check version A of the second order condition, compute $\mathbf{A} = \frac{\partial^2 f(\mathbf{u})}{\partial \mathbf{u} \partial \mathbf{u}'}$ and then see if $\mathbf{v}'\mathbf{A}\mathbf{v} \leq 0$ (≥ 0) when you plug in all possible values of **u** and **v**.

Version B. If met, this condition guarantees you have found a local maximum (minimum), which might be the global maximum or minimum (one should also check what happens as the values in **u** become negative or approach any boundaries or discontinuities). The condition is that f(**u**) is concave (convex) immediately around the point you found using the first order condition. This is guaranteed if, for the specific point $\tilde{\mathbf{u}}$ you found, the Hessian matrix $\frac{\partial^2 f(\mathbf{u})}{\partial \mathbf{u} \partial \mathbf{u}'}$ is negative (positive) definite. For any $n \times n$ matrix A to be negative (positive) definite means that $\mathbf{v}'\mathbf{A}\mathbf{v}$ is always <0 (>0) for any possible $n \times 1$ vector of real numbers **v**, with the exception of a vector of zeros **v**=**0**. Again, you don't need to know how to determine this. To summarize, to check version B of the second order condition, compute $\mathbf{A} = \frac{\partial^2 f(\mathbf{u})}{\partial \mathbf{u} \partial \mathbf{u}'}$, plug in $\mathbf{u} = \tilde{\mathbf{u}}$ in your answer, and then see if $\mathbf{v}'\mathbf{A}\mathbf{v} < 0$ (>0) when you plug in all possible values of **v**.

If you are checking version B and $\mathbf{v}'\mathbf{A}\mathbf{v}$ is sometimes <0 and other times >0 depending on the values in **v**, then matrix **A** is called indefinite and the solution found is neither a maximum nor a minimum (it might be a "saddle point" in the function $f(\mathbf{u})$). There are further ways to check the second order condition, but these are more advanced topics.

---

* "Global" maximum means the highest maximum anywhere; "global minimum" is the lowest minimum anywhere. You could alternatively have a local maximum or minimum; for example the peak of Mount Washington is a local high point but not the highest point in the world.

69. In the previous question you wrote out $\mathbf{v}'\mathbf{A}\mathbf{v}$, where **A** was a Hessian matrix. At the point where $\begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 7 \\ 0 \end{bmatrix}$, is $\mathbf{v}'\mathbf{A}\mathbf{v}$ >0, <0, ≥0, or ≤0 for all possible values of **v**≠**0**?

70. Suppose you want to choose values of **u** that minimize the function $f(\mathbf{u}) = u_1^2 + u_2 u_3^4$. Write the first order condition and solve for the values of **u** that satisfy this condition. You have already computed the Hessian matrix for the function, so as you did in the previous question, see whether $\mathbf{v}'\mathbf{A}\mathbf{v}$ satisfies version B of the second order condition. I.e., using version B of the second order condition, can you confirm that you have found a local minimum?

Answers to Self-Test for Derivatives:

$$\frac{\partial(a + bu^n)}{\partial u} = bnu^{n-1}$$

$$\frac{\partial \ln(ku)}{\partial u} = \frac{1}{u}$$

$$\frac{\partial(u^2 e^{ku})}{\partial u} = (2u)e^{ku} + u^2(ke^{ku}) = (ku^2 + 2u)e^{ku}$$

$$\frac{\partial^2(u^7)}{\partial u^2} = \frac{\partial(7u^6)}{\partial u} = 42u^5$$

$$\frac{\partial \exp(k(\ln u)^{107})}{\partial u} = \left(\exp(k(\ln u)^{107})\right)\left(107k(\ln u)^{106}\right)\left(\frac{1}{u}\right)$$

If you had trouble only with the last question, then sometime over the first month of the semester it is a good idea to use my derivatives exercises enough to get used to the chain rule. If you had trouble with multiple questions, it is a good idea to practice just the very simple derivative rules right away.