

COMPANION NOTES AND SOLUTIONS

by Mufaro Machaya

Contents

1	Chapter 1: Matrices	2
1.1	Labs 0 and 1: Introduction and Matrix Basics	2
1.2	Lab 2: A Matrix Representation of Linear Systems	3
1.3	Lab 3: Powers, Inverses, and Special Matrices	6

1 Chapter 1: Matrices

1.1 Labs 0 and 1: Introduction and Matrix Basics

NOTES

1. Matrix addition is commutative, so addition in any order produces the same result.

$$\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A} \quad (1)$$

2. Scalar multiplication is applied on all elements of the given matrix as a regular multiplication. For example, for 3×3 matrix \mathbf{A} and scalar s :

$$s\mathbf{A} = \begin{pmatrix} sA_{11} & sA_{12} & sA_{13} \\ sA_{21} & sA_{22} & sA_{23} \\ sA_{31} & sA_{32} & sA_{33} \end{pmatrix}. \quad (2)$$

3. A transposition is essentially swapping your rows and columns (do not, however, mistake it for a rotation), so 2×3 matrix \mathbf{A} would produce 3×2 matrix \mathbf{A}^{-1} :

$$\begin{pmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \end{pmatrix}^T = \begin{pmatrix} A_{11} & A_{21} \\ A_{12} & A_{22} \\ A_{13} & A_{23} \end{pmatrix} \quad (3)$$

4. Matrix multiplication essentially computes the dot products of the rows of the first matrix with the columns of the second matrix, thus why the number of rows of the first matrix must equal the number of columns of the second. For example, for 2×3 matrix \mathbf{A} and 3×2 matrix \mathbf{B} , the resulting matrix would be

$$\begin{aligned} \mathbf{AB} &= \begin{pmatrix} \mathbf{A}_{r=1} \cdot \mathbf{B}_{c=1} & \mathbf{A}_{r=1} \cdot \mathbf{B}_{c=2} \\ \mathbf{A}_{r=2} \cdot \mathbf{B}_{c=1} & \mathbf{A}_{r=2} \cdot \mathbf{B}_{c=2} \end{pmatrix}_1 \\ &= \begin{pmatrix} A_{11}B_{11} + A_{12}B_{21} + A_{13}B_{31} & A_{11}B_{12} + A_{12}B_{22} + A_{13}B_{32} \\ A_{21}B_{11} + A_{22}B_{21} + A_{23}B_{31} & A_{21}B_{12} + A_{22}B_{22} + A_{23}B_{32} \end{pmatrix} \end{aligned} \quad (4)$$

¹ $M_{r=m}$ denotes “the vector comprised of the elements of row m in matrix \mathbf{M} ” and $M_{c=n}$ denotes “the vector comprised of the elements of column n in matrix \mathbf{M} .”

5. Matrix multiplication is not commutative, so multiplying matrices in different orders will produce different results.

$$\mathbf{AB} \neq \mathbf{BA} \quad (5)$$

6. Matrix addition and scalar multiplication are commutative with transposition, but multiplication is not.

$$(\mathbf{A} + \mathbf{B})^T = \mathbf{A}^T + \mathbf{B}^T \quad (6)$$

$$(s\mathbf{A})^T = s\mathbf{A}^T \quad (7)$$

$$(\mathbf{AB})^T \neq \mathbf{A}^T \mathbf{B}^T \quad (8)$$

7. The trace of a matrix is the sum of the elements on its main diagonal.
8. The trace is commutative with matrix addition (i.e., $\text{tr}(\mathbf{A} + \mathbf{B}) = \text{tr}(\mathbf{A}) + \text{tr}(\mathbf{B})$), but non-commutative with matrix multiplication.
9. Matrix multiplication within trace in either order will produce the same result,

$$\text{tr}(\mathbf{AB}) = \text{tr}(\mathbf{BA}) \quad (9)$$

1.2 Lab 2: A Matrix Representation of Linear Systems

EXERCISE 2(g)

Given

$$\mathbf{M} = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 0 & 3 \\ 0 & 1 & 0 \end{pmatrix},$$

we can produce $\mathbf{M} \rightarrow \mathbf{I}_3$ via:

1. Divide row 3 by 3.
2. Add multiples of -2 from row 2 to row 1.
3. Swap rows 2 and 3².

²This was originally wrong and placed first. You need to swap last because matrix multiplication is right-associative, so the rightmost operation occurs first, not last. In other words, a “correct” way to approach this would be to do the entire process in reverse when multiplying.

This effect can be produced by the following elementary matrices.

The swap matrix can be created from emulation with I_3 :

$$\mathbf{E}_{swap} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix},$$

similarly, the division matrix would be

$$\mathbf{E}_{div} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \frac{1}{3} \end{pmatrix},$$

and lastly, the additive matrix would be

$$\mathbf{E}_{add} = \begin{pmatrix} 1 & -2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Altogether, the following multiplication should hold true:

$$\mathbf{E}_{div}\mathbf{E}_{add}\mathbf{E}_{swap}\mathbf{M} = \mathbf{I}_3$$

EXERCISE 2(h)

Given

$$\begin{aligned}x + 2y &= 4 \\ 3z &= 6 \\ y &= 8,\end{aligned}$$

we can produce the following coefficient matrix \mathbf{A} and constant vector \mathbf{B} :

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 0 & 3 \\ 0 & 1 & 0 \end{pmatrix} \quad \mathbf{B} = \begin{pmatrix} 4 \\ 6 \\ 8 \end{pmatrix},$$

and given that $\mathbf{A} = \mathbf{M}$, from exercise 2(g) the solutions can be calculated by performing the RREF calculations on \mathbf{B} to produce solution vector \mathbf{S} :

$$\mathbf{E}_{div}\mathbf{E}_{add}\mathbf{E}_{swap}\mathbf{B} = \mathbf{S},$$

yielding

$$\mathbf{S} = \begin{pmatrix} -12 \\ 8 \\ 2 \end{pmatrix},$$

showing that $x = -12, y = 8, z = 2$.

EXERCISE 2(i)

This isn't solvable because the best RREF form is

$$[\mathbf{M}|\mathbf{B}] = \begin{pmatrix} 1 & 0 & -6 & 4 \\ 0 & 1 & 3 & 6 \\ 0 & 0 & 0 & 6 \end{pmatrix}$$

or, in other words, equations (2) and (3) contradict each other.

EXERCISE 2(j)

The initial state of the system is

$$\begin{pmatrix} 1 & 2 & 0 & 4 \\ 0 & 1 & 3 & 6 \\ 0 & 2 & 6 & 12 \end{pmatrix},$$

and the RREF of this system becomes

$$\begin{pmatrix} 1 & 0 & -6 & -8 \\ 0 & 1 & 3 & 6 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Again, this system is not solvable as equations 2 and 3 are equivalent, resulting in this system not containing enough information to be adequately solved, but the reduced echelon form of the coefficient matrix is the same as in exercise 2(i).

1.3 Lab 3: Powers, Inverses, and Special Matrices

NOTES

1. A matrix multiplied by its transposition will produce a gram matrix which is symmetric (does not change on transposition).

$$\mathbf{M}^T \mathbf{M} = (\mathbf{M}^T \mathbf{M})^T \neq \mathbf{M} \mathbf{M}^T = (\mathbf{M} \mathbf{M}^T)^T \quad (10)$$

2. A matrix multiplied by its inverse, in any order, will produce an identity matrix.

$$\mathbf{M} \mathbf{M}^{-1} = \mathbf{M}^{-1} \mathbf{M} = \mathbf{I}_n \quad (11)$$

3. A matrix inverted twice produces itself.

$$(\mathbf{M}^{-1})^{-1} = \mathbf{M} \quad (12)$$

4. Matrix multiplication is not commutative with inversion, however, they are commutative in alternating order.

$$(\mathbf{A} \mathbf{B})^{-1} = \mathbf{B}^{-1} \mathbf{A}^{-1} \neq (\mathbf{B} \mathbf{A})^{-1} = \mathbf{A}^{-1} \mathbf{B}^{-1} \quad (13)$$

5. Matrix inversion is commutative with transposition.

$$(\mathbf{M}^T)^{-1} = (\mathbf{M}^{-1})^T \quad (14)$$

6. A square matrix \mathbf{M} is

- (a) Symmetric when $\mathbf{M} = \mathbf{M}^T$
- (b) Diagonal when $\mathbf{M}_{ij} = 0$ if $i \neq j$
- (c) Upper triangular when $\mathbf{M}_{ij} = 0$ if $i > j$ or lower triangular when $i < j$.

7. Assuming a matrix to be square, a matrix plus its transposition is symmetric.

$$\mathbf{M} + \mathbf{M}^T = (\mathbf{M} + \mathbf{M}^T)^T \quad (15)$$

THEOREM 1. The inverse of an elementary matrix is an elementary matrix.

There are three elementary operations to be considered: scalar multiplication of a row, addition of the scalar multiplication from one row to another, and swapping rows. Let \mathbf{I}_m be an identity matrix of size $m \times m$, which can be described in the following matrix-vector notation,

$$\mathbf{I}_m = \begin{pmatrix} \vec{\mathbf{I}}_{m,r=1} \\ \vec{\mathbf{I}}_{m,r=2} \\ \vdots \\ \vec{\mathbf{I}}_{m,r=i} \\ \vdots \\ \vec{\mathbf{I}}_{m,r=m-1} \\ \vec{\mathbf{I}}_{m,r=m} \end{pmatrix},$$

where $\vec{\mathbf{I}}_{m,r=i}$ is a vector of row i in \mathbf{I}_m . This notation can be further simplified to produce generalized elementary matrices for each operation. Let row i be our of interest (the row in which we are manipulating/changing) for scalar multiplication and rows i and j be our rows of interest for scalar addition and swapping. Therefore, we can denote the elementary operations, in order, as

$$\begin{aligned} s\vec{\mathbf{I}}_{m,r=i} &\rightarrow i \\ \vec{\mathbf{I}}_{m,r=i} + s\vec{\mathbf{I}}_{m,r=j} &\rightarrow i \\ \vec{\mathbf{I}}_{m,r=i} &\leftrightarrow \vec{\mathbf{I}}_{m,r=j} \end{aligned}$$

where $X \rightarrow I_{m,r=i}$ indicates that the result of operation X is stored at row i , and $\vec{\mathbf{A}}_{r=1} \leftrightarrow \vec{\mathbf{A}}_{r=2}$ swaps the vectors of \mathbf{A} at rows 1 and 2.

Given this notation, we can write each of their corresponding elementary matrices. For scalar multiplication, this would be

$$\mathbf{E}_{mult} = \begin{pmatrix} \vdots \\ s\vec{\mathbf{I}}_{m,r=i} \\ \vdots \end{pmatrix},$$

so its inverse, naturally, would be

$$\mathbf{E}_{mult}^{-1} = \begin{pmatrix} \vdots \\ \frac{1}{s}\vec{\mathbf{I}}_{m,r=i} \\ \vdots \end{pmatrix},$$

which, similarly, is an elementary matrix of scalar multiplication by $1/s$. Likewise, scalar addition, the elementary matrix would be

$$\mathbf{E}_{add} = \begin{pmatrix} \vdots \\ \vec{\mathbf{I}}_{m,r=i} + s\vec{\mathbf{I}}_{m,r=j} \\ \vdots \end{pmatrix},$$

so its inverse, naturally, would be

$$\mathbf{E}_{add}^{-1} = \begin{pmatrix} \vdots \\ \vec{\mathbf{I}}_{m,r=j} - s\vec{\mathbf{I}}_{m,r=i} \\ \vdots \end{pmatrix},$$

which, similarly, is an elementary matrix of scalar addition by $-s$. Lastly, the inverse of any elementary matrix produced by a row swap is itself, and that can be proven with the following example:

$$\mathbf{E}_{swap} = \begin{pmatrix} \vdots \\ \vec{\mathbf{I}}_{m,r=j} \\ \vdots \\ \vec{\mathbf{I}}_{m,r=i} \\ \vdots \end{pmatrix}$$

where $i < j$. The gaussian elimination matrix would be

$$[\mathbf{E}_{swap} | \mathbf{I}_m] = \left(\begin{array}{c|c} \vdots & \vdots \\ \vec{\mathbf{I}}_{m,r=j} & \vec{\mathbf{I}}_{m,r=i} \\ \vdots & \vdots \\ \vec{\mathbf{I}}_{m,r=i} & \vec{\mathbf{I}}_{m,r=j} \\ \vdots & \vdots \end{array} \right),$$

so therefore, the corresponding inverse matrix would be

$$\left(\begin{array}{c|c} \vdots & \vdots \\ \vec{\mathbf{I}}_{m,r=i} & \vec{\mathbf{I}}_{m,r=j} \\ \vdots & \vdots \\ \vec{\mathbf{I}}_{m,r=j} & \vec{\mathbf{I}}_{m,r=i} \\ \vdots & \vdots \end{array} \right) = [\mathbf{I}_m, \mathbf{E}_{swap}],$$

thus showing that $\mathbf{E}_{swap}^{-1} = \mathbf{E}_{swap}$.