

1 Pauli matrices

The Pauli matrices are a set of three 2×2 complex matrices that are very useful in physics. They are defined as,

$$\sigma_x = \sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z = \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

In this exercise, we will prove a number of useful properties of these matrices.

- (a) Show that σ_x , σ_y , σ_z are *Hermitian*, that is, they are equal to the conjugate transpose of themselves. (Note that here by conjugate, we mean the complex conjugate of the matrix, i.e., the matrix obtained by conjugating every entry, and transpose means switching rows and columns.)

- (b) Show that $\sigma_x^2 = \sigma_y^2 = \sigma_z^2 = -i\sigma_x\sigma_y\sigma_z = I$, where I is the 2×2 identity matrix.

- (c) The *commutator* is a matrix operation that quantifies how much two matrices fail to commute. It is defined as $[A, B] = AB - BA$. Prove the following formulas for the commutator of the Pauli matrices,

$$[\sigma_x, \sigma_y] = 2i\sigma_z, \quad [\sigma_y, \sigma_z] = 2i\sigma_x, \quad [\sigma_z, \sigma_x] = 2i\sigma_y.$$

Note: Sometimes these relations are denoted as $[\sigma_i, \sigma_j] = 2i \epsilon_{ijk} \sigma_k$.

- (d) Another matrix operation similar to the commutator is the *anti-commutator*, defined by $\{A, B\} = AB + BA$. Show that $\{\sigma_i, \sigma_j\} = 2\delta_{ij} I$ where we've used the alternative notation $\sigma_1, \sigma_2, \sigma_3$ for $\sigma_x, \sigma_y, \sigma_z$.

2 Matrix functions and series

Since we can sum and multiply matrices, we can define polynomial functions of matrices. If $p(x) = a_0 + a_1x + \cdots + a_nx^n$, we can basically plug in a matrix A for x , and define $p(A) = a_0 + a_1A + \cdots + a_nA^n$. Just as we can extend polynomials to power series, we can extend polynomial functions of matrices to matrix functions defined in terms of series. For example, we can define the matrix exponential,

$$\exp A = \sum_{n=0}^{\infty} \frac{1}{n!} A^n.$$

(We define $A^0 = I$, the identity matrix.)

- (a) Let D be a diagonal matrix with diagonal entries $(\lambda_1, \dots, \lambda_k)$. Show that $\exp D$ is a diagonal matrix with diagonal entries $(e^{\lambda_1}, \dots, e^{\lambda_k})$.

- (b) Using the results from question 1(b), show that

$$(i\sigma_y)^n = \begin{cases} I & n \equiv 0 \pmod{4} \\ i\sigma_y & n \equiv 1 \pmod{4} \\ -I & n \equiv 2 \pmod{4} \\ -i\sigma_y & n \equiv 3 \pmod{4} \end{cases}$$

(c) Using the results of part (b), show that

$$\exp(-i\theta \sigma_y) = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}$$

Multiplying a vector in the xy -plane by this matrix rotates it counterclockwise by the angle θ .

3 Elementary row operations and Gaussian elimination

3.1 Introduction to Gaussian elimination

Gaussian elimination is a process by which we start with the defining equation of the matrix inverse $AA^{-1} = I$ and perform a series of linear transformations on this equation on this equation which can be represented by matrix multiplication from the left.

$$O_N O_{N-1} \dots O_2 O_1 A A^{-1} = O_N O_{N-1} \dots O_2 O_1 I,$$

where the O_i are the linear transformation matrices. The goal is to perform linear transformations such that after all the transformations are done the matrix A on the left hand side has been transformed into the identity matrix $O_N O_{N-1} \dots O_2 O_1 A = I$. Then the equation implies that the matrix on the right hand side $O_N O_{N-1} \dots O_2 O_1$ is the inverse of A .

It's best to explain how this is done in practice by using an example, lets try to find the inverse of a matrix A (note that this only works if A has an inverse, i.e. is not singular):

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 2 & -1 & 1 \\ 1 & 2 & -1 \end{pmatrix}.$$

3.1.1 Mirror Gaussian Elimination Operations on I

We start by forming the augmented 3×6 matrix $[A \mid I]$ by adding a unit matrix to the right of A

$$\left[\begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 2 & -1 & 1 & 0 & 1 & 0 \\ 1 & 2 & -1 & 0 & 0 & 1 \end{array} \right]$$

The goal is now to perform a series of transformations on the full augmented matrix such that A is transformed into the unit matrix. At the end of the process the unit matrix on the right will have automatically been transformed into the inverse A^{-1} . (you can think of this process as applying sequential operations O_i from the proof above to both A and I at the same time)

— Column 1:

Step 1 Move the largest element of column 1 into the 1-1 position (the “pivot”): Largest element in column 1 is 2 (row 2). Swap $R1 \leftrightarrow R2$.

$$\left[\begin{array}{ccc|ccc} 2 & -1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 2 & -1 & 0 & 0 & 1 \end{array} \right]$$

Step 2 Scale the pivot to 1: $R1 \rightarrow R1 \cdot 1/2$

$$\left[\begin{array}{ccc|ccc} 1 & -1/2 & 1/2 & 0 & 1/2 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 2 & -1 & 0 & 0 & 1 \end{array} \right]$$

Step 3 Zero out the second element of the first column by subtracting a multiple of the pivot row : $R2 \rightarrow R2 - R1$

$$\left[\begin{array}{ccc|ccc} 1 & 1 & 0 & 0 & \frac{1}{2} & 0 \\ 0 & 1 & 0 & 1 & -\frac{1}{2} & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 \end{array} \right]$$

Step 4 and zero out the 3rd element: $R3 \rightarrow R3 - R1$

$$\left[\begin{array}{ccc|ccc} 1 & 1 & 0 & 0 & \frac{1}{2} & 0 \\ 0 & 1 & 0 & 1 & -\frac{1}{2} & 0 \\ 0 & 0 & 1 & 0 & -\frac{1}{2} & 1 \end{array} \right]$$

— Column 2:

Step 5 Here the pivot (2-2 element) is already the largest element in column 2. No swap needed.

Step 6 Scale: Pivot is already 1. No scaling needed.

Step 7 Eliminate: $R1 \rightarrow R1 - R2$

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -1 & 1 & 0 \\ 0 & 1 & 0 & 1 & -\frac{1}{2} & 0 \\ 0 & 0 & 1 & 0 & -\frac{1}{2} & 1 \end{array} \right]$$

(R3 already has 0 in column 2 nothing to do.)

— Column 3:

Step 8: Pivot is already 1, and both R1 and R2 already have 0 in column 3. Done.

— Result:

$$A^{-1} = \begin{pmatrix} -1 & 1 & 0 \\ 1 & -\frac{1}{2} & 0 \\ 0 & -\frac{1}{2} & 1 \end{pmatrix}$$

3.1.2 Constructing A^{-1} by Representing Row Operations as Matrices

Here we describe an alternative way to perform Gaussian elimination which is to construct the matrices O_i explicitly and multiply them. Since every row operation we do on A can be represented by a left-multiplication by an elementary matrix that represent the operation, we can determine what those matrices are and left multiply them one after another in the same order we carry them out on A to get A^{-1} . The next section explores why these matrix operations work, so feel free to come back to this after going through Sec.3.2.

Step 1: Swap $R1 \leftrightarrow R2$

$$P = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Step 2: $R1 \rightarrow R1/2$

$$S = \begin{pmatrix} \frac{1}{2} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Step 3: $R2 \rightarrow R2 - R1$

$$E_3 = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Step 4: $R_3 \rightarrow R_3 - R_1$

$$E_4 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix}$$

Step 7: $R_1 \rightarrow R_1 - R_2$

$$E_7 = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

— Since $E_7 \cdot E_4 \cdot E_3 \cdot S \cdot P \cdot A = I$, we have:

$$A^{-1} = E_7 \cdot E_4 \cdot E_3 \cdot S \cdot P$$

Multiplying right to left:

$$S \cdot P = \begin{pmatrix} 0 & \frac{1}{2} & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$E_3 \cdot S \cdot P = \begin{pmatrix} 0 & \frac{1}{2} & 0 \\ 1 & -\frac{1}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$E_4 \cdot E_3 \cdot S \cdot P = \begin{pmatrix} 0 & \frac{1}{2} & 0 \\ 1 & -\frac{1}{2} & 0 \\ 0 & -\frac{1}{2} & 1 \end{pmatrix}$$

$$E_7 \cdot E_4 \cdot E_3 \cdot S \cdot P = \begin{pmatrix} -1 & 1 & 0 \\ 1 & -\frac{1}{2} & 0 \\ 0 & -\frac{1}{2} & 1 \end{pmatrix}$$

We get the same answer as in Sec.3.1.1.

3.2 Exercises

The operations used in Gaussian elimination can be represented by matrix multiplications. In the following we will consider an $n \times n$ matrix A .

- (a) First there are row-switching transformations. Consider the matrix obtained by swapping the i -th and j -th row of the identity matrix.

$$T_{ij} = \begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 0 & \cdots & 1 & \\ & & \vdots & \ddots & \vdots & \\ & & 1 & \cdots & 0 & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix}$$

Show that $T_{ij}A$ is equal to A with the i -th and j -th row swapped. Also show that $\det T_{ij} = -1$.

- (b) Scaling the row i by a factor k can be represented by the identity with the i -th diagonal entry set to k ,

$$D_i(k) = \begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & k & & \\ & & & & 1 & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix}$$

Show that $D_i(k)A$ is equal to A with the i -th row multiplied by k , and that $\det D_i(k) = k$.

- (c) The third type of row operations used in Gaussian elimination is adding the j -th row multiplied by a scalar r to the row i . This is represented by the identity matrix with the (i, j) -th entry set to r ,

$$L_{ij}(r) = \begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & \cdots & r & \\ & & & \ddots & & \\ & & & & 1 & \\ & & & & & \ddots \\ & & & & & & 1 \end{pmatrix}$$

Again, show that $L_{ij}(r)A$ is what we expect, and $\det L_{ij}(r) = 1$.

- (d) Using the above facts, show that the scaling factors multiply to $\det A$ times $(-1)^s$ where s is the number of row swaps.