2 Matrices

"Do not worry about your difficulties in Mathematics. I can assure you mine are still greater."
- Albert Einstein

2.1 Matrix Algebra

Matrices are arrays of (usually) scalars arranged in rows and columns. For example:

$$\hat{\boldsymbol{A}}=\begin{bmatrix}1&2&3\\4&5&6\\7&8&9\end{bmatrix}\qquad \hat{\boldsymbol{B}}=\begin{bmatrix}2&3&4\\5&6&7\\8&9&1\end{bmatrix}$$

Matrices of equivalent dimensions are added or subtracted by adding or subtracting their corresponding individual elements:

$$\hat{\boldsymbol{A}}+\hat{\boldsymbol{B}}=\begin{bmatrix}1&2&3\\4&5&6\\7&8&9\end{bmatrix}+\begin{bmatrix}2&3&4\\5&6&7\\8&9&1\end{bmatrix}=\begin{bmatrix}1+2&2+3&3+4\\4+5&5+6&6+7\\7+8&8+9&9+1\end{bmatrix}=\begin{bmatrix}3&5&7\\9&11&13\\15&17&10\end{bmatrix}$$

$$\hat{\boldsymbol{A}}+\hat{\boldsymbol{B}}=\hat{\boldsymbol{B}}+\hat{\boldsymbol{A}}\quad \text{(commutative)}$$
$$\hat{\boldsymbol{A}}+(\hat{\boldsymbol{B}}+\hat{\boldsymbol{C}})=(\hat{\boldsymbol{A}}+\hat{\boldsymbol{B}})+\hat{\boldsymbol{C}}\quad \text{(associative)}$$
$$\hat{\boldsymbol{A}}+\hat{\boldsymbol{0}}=\hat{\boldsymbol{A}}\quad \text{(identity)}$$
$$\hat{\boldsymbol{A}}+(-\hat{\boldsymbol{A}})=\hat{\boldsymbol{0}}\quad \text{(additive~inverse)}$$

[2-1]

Normal matrix multiplication proceeds row by column:

$$\hat{\boldsymbol{A}}\,\hat{\boldsymbol{B}}=\begin{bmatrix}1&2&3\\4&5&6\\7&8&9\end{bmatrix}\times\begin{bmatrix}2&3&4\\5&6&7\\8&9&1\end{bmatrix}$$
$$=\begin{bmatrix}1\times2+2\times5+3\times8&1\times3+2\times6+3\times9&1\times4+2\times7+3\times1\\4\times2+5\times5+6\times8&4\times3+5\times6+6\times9&4\times4+5\times7+6\times1\\7\times2+8\times5+9\times8&7\times3+8\times6+9\times9&7\times4+8\times7+9\times1\end{bmatrix}$$
$$=\begin{bmatrix}36&42&63\\85&96&57\\126&150&93\end{bmatrix}$$

Note that matrix multiplication is not necessarily commutative. That is, $\hat{\boldsymbol{A}}\hat{\boldsymbol{B}}$ does not necessarily equal $\hat{\boldsymbol{B}}\hat{\boldsymbol{A}}$. Also, the matrices to not have to be square but the number of *rows* of $\hat{\boldsymbol{A}}$ must be the same as the number of *columns* in $\hat{\boldsymbol{B}}$ in order to do a normal multiplication.

Matrices may be multiplied or divided by a scalar:

$$6\times\begin{bmatrix}1 & 2\\ 3 & 4\end{bmatrix}=\begin{bmatrix}6\times1 & 6\times2\\ 6\times3 & 6\times4\end{bmatrix}=\begin{bmatrix}6 & 12\\ 18 & 24\end{bmatrix}$$

$$\begin{aligned}\alpha(\hat{\boldsymbol{A}}+\hat{\boldsymbol{B}})&=\alpha\hat{\boldsymbol{A}}+\alpha\hat{\boldsymbol{B}} \quad (\text{scalar distributive})\\ (\alpha+\beta)\hat{\boldsymbol{A}}&=\alpha\hat{\boldsymbol{A}}+\beta\hat{\boldsymbol{A}} \quad (\text{matrix distributive})\\ (\alpha\beta)\hat{\boldsymbol{A}}&=\alpha(\beta\hat{\boldsymbol{A}}) \quad (\text{associative law for multiplicat}\end{aligned} \qquad [2\text{-}2]$$

These definitions for addition and scalar multiplication are of importance in the field of vector analysis. We will return to this later.

The direct product (sometimes called the Kronecker product or tensor product) proceeds as follows:

$$\begin{aligned}\hat{\boldsymbol{A}}\otimes\hat{\boldsymbol{B}}&=\begin{bmatrix}1 & 2\\ 3 & 4\end{bmatrix}\otimes\begin{bmatrix}5 & 6\\ 7 & 8\end{bmatrix}\\ &=\begin{bmatrix}1\begin{bmatrix}5 & 6\\ 7 & 8\end{bmatrix} & 2\begin{bmatrix}5 & 6\\ 7 & 8\end{bmatrix}\\ 3\begin{bmatrix}5 & 6\\ 7 & 8\end{bmatrix} & 4\begin{bmatrix}5 & 6\\ 7 & 8\end{bmatrix}\end{bmatrix}\\ &=\begin{bmatrix}5 & 6 & 10 & 12\\ 7 & 8 & 14 & 16\\ 15 & 18 & 20 & 24\\ 21 & 24 & 28 & 32\end{bmatrix}\end{aligned} \qquad [2\text{-}3]$$

As with normal matrix multiplication, the Kronecker product is not generally commutative. This type of matrix multiplication will be of importance to us when we encounter product operators of multiple spins.

The zero matrix, referred to in equations [2-1], is (not surprisingly) full of zeros:

$$\begin{bmatrix}0 & 0 & 0 & \cdots\\ 0 & 0 & 0 & \cdots\\ 0 & 0 & 0 & \cdots\\ \vdots & \vdots & \vdots & \ddots\end{bmatrix}=\hat{\mathbf{0}} \qquad [2\text{-}4]$$

and the unit or identity matrix is:

$$\begin{bmatrix}1 & 0 & 0 & \cdots\\ 0 & 1 & 0 & \cdots\\ 0 & 0 & 1 & \cdots\\ \vdots & \vdots & \vdots & \ddots\end{bmatrix}=\hat{\mathbf{1}} \qquad [2\text{-}5]$$

with 1's on the diagonal so that:

$$\hat{A}\times\hat{0}=\hat{0} \quad \hat{0}\times\hat{A}=\hat{0}$$
$$\hat{A}\times\hat{1}=\hat{A} \quad \hat{1}\times\hat{A}=\hat{A} \qquad [2\text{-}6]$$

If $\hat{A}\hat{B}$ = $\hat{1}$ where $\hat{1}$ is the identity matrix then $\hat{A}$ and $\hat{B}$ are the *inverse* of each other. $\hat{B}$ = $\hat{A}^{-1}$ and $\hat{A}$ = $\hat{B}^{-1}$. The inverse of the product of two matrices is equal to the product of the inverses of the two matrices *in reverse order*:

$$(\hat{A}\hat{B})^{-1}=\hat{B}^{-1}\hat{A}^{-1} \qquad [2\text{-}7]$$

That this is so can be seen from the following:

$$\begin{aligned}\hat{A}\hat{B}(\hat{A}\hat{B})^{-1}&=\hat{A}\hat{B}\hat{B}^{-1}\hat{A}^{-1}\\ &=\hat{A}\hat{1}\hat{A}^{-1}\\ &=\hat{A}\hat{A}^{-1}\\ &=\hat{1}\end{aligned}$$

The transpose matrix $\hat{A}^T$ is one in which the rows and columns have been interchanged. Thus, if $\hat{A}$ is a n x m matrix and $\hat{A}^T$ is the transpose of $\hat{A}$ then $\hat{A}^T$ will be a mxn matrix.

$$\hat{A}=\begin{bmatrix}0 & 1 & 2\\ 3 & 4 & 5\\ 6 & 7 & 8\end{bmatrix} \quad \hat{A}^T=\begin{bmatrix}0 & 3 & 6\\ 1 & 4 & 7\\ 2 & 5 & 8\end{bmatrix} \qquad [2\text{-}8]$$

Obviously, the transpose of the transpose of a matrix equals the original matrix:

$$(\hat{A}^T)^T=\hat{A}$$

The transpose of the sum of two matrices is the same as the sum of their transposes:

$$(\hat{A}+\hat{B})^T=\hat{A}^T+\hat{B}^T$$
$$\text{proof:}$$
$$\text{let}\ \hat{C}=(\hat{A}+\hat{B})^T \qquad [2\text{-}9]$$
$$c_{ij}=(a_{ij}+b_{ij})^T=a_{ji}+b_{ji}$$

The transpose of the product of two matrices is equal to the *reverse* product of the transposes of the matrices:

$$(\hat{A}\hat{B})^T=\hat{B}^T\hat{A}^T \qquad [2\text{-}10]$$

If the product of $\hat{A}^T$ and $\hat{A}$ is equal to the identity matrix they are *orthogonal*:

$$\hat{\boldsymbol{A}}\hat{\boldsymbol{A}}^T = \hat{\boldsymbol{A}}^T\hat{\boldsymbol{A}} = \hat{\mathbf{1}} \qquad [2\text{-}11]$$

If a matrix is equal to its transpose it is *symmetric*:

$$\hat{\boldsymbol{A}} = \hat{\boldsymbol{A}}^T \quad (\text{symmetric matrices})$$

The sum of two symmetric matrices is another symmetric matrix:

$$(\hat{\boldsymbol{A}} + \hat{\boldsymbol{B}})^T = \hat{\boldsymbol{A}}^T + \hat{\boldsymbol{B}}^T = \hat{\boldsymbol{A}} + \hat{\boldsymbol{B}} \qquad [2\text{-}12]$$

In order for the product of two symmetric matrices to be symmetric, the two matrices *must* commute:

$$\hat{\boldsymbol{C}}^{\boldsymbol{T}} = (\hat{\boldsymbol{A}}\hat{\boldsymbol{B}})^T = \hat{\boldsymbol{B}}^T\hat{\boldsymbol{A}}^T = \hat{\boldsymbol{B}}\hat{\boldsymbol{A}} = \hat{\boldsymbol{A}}\hat{\boldsymbol{B}} = \hat{\boldsymbol{C}}$$

The conjugate matrix $\hat{\boldsymbol{A}}^*$ is one in which the elements are the complex conjugates of the original matrix elements.

$$\hat{\boldsymbol{A}} = \begin{bmatrix} 0 & 1 & 2\mathrm{i} \\ 3 & 4 & 5 \\ 6 & 7\mathrm{i} & 8 \end{bmatrix} \qquad \hat{\boldsymbol{A}}^* = \begin{bmatrix} 0 & 1 & -2\mathrm{i} \\ 3 & 4 & 5 \\ 6 & -7\mathrm{i} & 8 \end{bmatrix} \qquad [2\text{-}13]$$

The matrix $(\hat{\boldsymbol{A}}^*)^T$ is the transpose of the complex conjugate of $\hat{\boldsymbol{A}}$. If $(\hat{\boldsymbol{A}}^*)^T$ = $\hat{\boldsymbol{A}}$ it is *hermitian*. Obviously, an Hermitian matrix is square. If $(\hat{\boldsymbol{A}}^*)^T$ = $\hat{\boldsymbol{A}}^{-1}$ it is *unitary* (and square). Thus we can write:

$$\begin{gathered} (\hat{\boldsymbol{A}}^*)^T = \boldsymbol{A}^{-1} \quad (\text{unitary}) \\ (\hat{\boldsymbol{A}}^*)^T \cdot \hat{\boldsymbol{A}} = \hat{\boldsymbol{A}}^{-1} \cdot \hat{\boldsymbol{A}} \\ (\hat{\boldsymbol{A}}^*)^T \cdot \hat{\boldsymbol{A}} = \hat{\mathbf{1}} \\ \\ (\hat{\boldsymbol{A}}^*)^T = \hat{\boldsymbol{A}} \quad (\text{hermitian}) \\ (\hat{\boldsymbol{A}}^*)^T \cdot \hat{\boldsymbol{A}}^{-1} = \hat{\boldsymbol{A}} \cdot \hat{\boldsymbol{A}}^{-1} \\ (\hat{\boldsymbol{A}}^*)^T \cdot \hat{\boldsymbol{A}}^{-1} = \hat{\mathbf{1}} \end{gathered} \qquad [2\text{-}14]$$

An example of unitary matrices:

$$\hat{\boldsymbol{A}} = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \quad \hat{\boldsymbol{A}}^{-1} = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \quad (\hat{\boldsymbol{A}}^*)^T = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}$$

$$(\hat{\boldsymbol{A}}^*)^T\hat{\boldsymbol{A}} = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \times \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \mathbf{1}$$

and an example of hermitian matrices:

$$\hat{\boldsymbol{A}}=\begin{bmatrix}0 & -i\\ i & 0\end{bmatrix} \quad \hat{\boldsymbol{A}}^{-1}=\begin{bmatrix}0 & -i\\ i & 0\end{bmatrix} \quad (\hat{\boldsymbol{A}}^{*})^{T}=\begin{bmatrix}0 & -i\\ i & 0\end{bmatrix}$$

$$(\hat{\boldsymbol{A}}^{*})^{T}\hat{\boldsymbol{A}}^{-1}=\begin{bmatrix}0 & -i\\ i & 0\end{bmatrix}\times\begin{bmatrix}0 & -i\\ i & 0\end{bmatrix}=\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}=\hat{\mathbf{1}}$$

In this case you can see that $\hat{\boldsymbol{A}}$ is both hermitian and unitary.

A small amount of effort on the reader's part will reveal that the diagonal elements of the hermitian matrix must, from the definition ([2-14]), be real. Hermitian matrices play a central role in quantum mechanics and will figure prominently in our discussion of the quantum mechanics of spins.

The *trace* of a matrix is the sum of the diagonal elements:

$$Tr(\hat{\boldsymbol{A}})=\sum_{i} a_{ii}$$

$$\hat{\boldsymbol{A}}=\begin{bmatrix}1 & 2\\ 3 & 4\end{bmatrix} \quad Tr(\hat{\boldsymbol{A}})=1+4=5 \qquad [2\text{-}15]$$

and is equal to the sum of the eigenvalues of the matrix (see below). The trace of the product of two matrices is (or rather, will be) of interest. We start with 2x2 matrice, $\hat{\boldsymbol{A}}$ and $\hat{\boldsymbol{B}}$:

$$\hat{\boldsymbol{A}}=\begin{bmatrix}a_{11} & a_{12}\\ a_{21} & a_{22}\end{bmatrix} \quad \hat{\boldsymbol{B}}=\begin{bmatrix}b_{11} & b_{12}\\ b_{21} & b_{22}\end{bmatrix}$$

The product of these matrices, $\hat{\boldsymbol{C}}$ = $\hat{\boldsymbol{A}}\hat{\boldsymbol{B}}$, is:

$$\hat{\boldsymbol{C}}=\begin{bmatrix}a_{11} & a_{12}\\ a_{21} & a_{22}\end{bmatrix}\times\begin{bmatrix}b_{11} & b_{12}\\ b_{21} & b_{22}\end{bmatrix}$$

$$=\begin{bmatrix}a_{11}b_{11}+a_{12}b_{21} & a_{11}b_{12}+a_{12}b_{22}\\ a_{21}b_{11}+a_{22}b_{21} & a_{21}b_{12}+a_{22}b_{22}\end{bmatrix}$$

and the trace of **C** is:

$$tr\,\hat{\boldsymbol{C}}=a_{11}b_{11}+a_{12}b_{21}+a_{21}b_{12}+a_{22}b_{22}$$

and in general, for nxn matrices the trace is easily calculated:

$$tr(\hat{\boldsymbol{A}}\,\hat{\boldsymbol{B}})=\sum_{i}^{n}\sum_{j}^{n} a_{ij}b_{ji} \qquad [2\text{-}16]$$

Note that the subscripts i and j are reversed from a to b. What about the trace of $\hat{\boldsymbol{B}}$ times $\hat{\boldsymbol{A}}$? We know that reversing the order of multiplication does not necessarily give the same result as $\hat{\boldsymbol{A}}$ times

$\hat{\boldsymbol{B}}$ but how does this reversal affect the trace? Again we start with our 2x2 matrices and compute $\hat{\boldsymbol{D}} = \hat{\boldsymbol{B}}\hat{\boldsymbol{A}}$:

$$\hat{\boldsymbol{D}}=\begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}\times\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$
$$=\begin{bmatrix} b_{11}a_{11}+b_{12}a_{21} & b_{11}a_{12}+b_{12}a_{22} \\ b_{21}a_{11}+b_{22}a_{21} & b_{21}a_{12}+b_{22}a_{22} \end{bmatrix}$$

The trace of **D** is:

$$tr\,\hat{\boldsymbol{D}}=b_{11}a_{11}+b_{12}a_{21}+b_{21}a_{12}+b_{22}a_{22}$$

which is the same as the trace of $\hat{\boldsymbol{C}}$. In general the trace of the product of two matrices does not rely on the order of multiplication of the matrices:

$$tr(\hat{\boldsymbol{A}}\hat{\boldsymbol{B}})=\sum_{i}^{n}\sum_{j}^{n}a_{ij}b_{ji}=\sum_{i}^{n}\sum_{j}^{n}b_{ij}a_{ji}=tr(\hat{\boldsymbol{B}}\hat{\boldsymbol{A}}) \quad [2\text{-}17]$$

The *inverse* matrix is one which abides by the following definition:

$$\hat{\boldsymbol{A}}\times\hat{\boldsymbol{A}}^{-1}=\hat{\boldsymbol{A}}^{-1}\times\hat{\boldsymbol{A}}=\hat{\boldsymbol{1}} \quad [2\text{-}18]$$

That is, if a matrix $\hat{\boldsymbol{A}}$ is multiplied by its inverse, $\hat{\boldsymbol{A}}^{-1}$, the result is the unity matrix. This is of course, analogous to multiplying a simple number by its inverse to get a result of one. The inverse matrix is of use in solving matrix equations but first we need to know about determinants and adjoint matrices before we can figure out what the inverse of a matrix is.

2.2 Determinants

A determinant is a form of matrix which can be evaluated to a number or a bit more formally, the determinant of a matrix evaluates to a scalar. The procedure is to multiply each member of a row or column by an element of the matrix that is **not** in the same row and column.

Thus, for our example matrix, $\hat{\boldsymbol{A}}$, the determinant is:

$$det\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$$
$$=1\cdot5\cdot9-1\cdot8\cdot6-2\cdot4\cdot9+2\cdot7\cdot6+3\cdot4\cdot8-3\cdot7\cdot5$$
$$=45-48-72+84+96-105$$
$$=0$$

Note that there a couple of ways to denote determinants:

$$det\,\hat{\boldsymbol{A}} \quad \text{or} \quad |\hat{\boldsymbol{A}}|$$

We will use both of these as needed.

A more useful evaluation technique for the general matrix:

$$\hat{\boldsymbol{E}}=\begin{bmatrix} e_{11} & e_{12} & e_{13} \\ e_{21} & e_{22} & e_{23} \\ e_{31} & e_{32} & e_{33} \end{bmatrix} \qquad [2\text{-}19]$$

is to evaluate by taking a row (or a column) and multiplying each member of that row (column) by its *cofactor*. What is the cofactor? In our general matrix, $\hat{\boldsymbol{E}}$, let's choose the first (top) row. For e_{11} the cofactor is:

$$E_{11}=\begin{vmatrix} e_{22} & e_{23} \\ e_{32} & e_{33} \end{vmatrix}$$

where E_{11} refers to the cofactor of element e_{11} . Generally, here we will refer to cofactors in subscripted uppercase. Note that the minor sub-matrix includes all terms from the original matrix that are *not* in the same row or column as e_{11} . It is particularly simple to evaluate 3x3 matrix since at this point we evaluate the *cofactor* of the minor matrix by doing:

$$E_{11}=e_{22}\cdot e_{33}-e_{32}\cdot e_{23}$$

which is, itself a determinant calculation of a 2x2 matrix. We then multiply by e_{11} to complete the calculation:

$$e_{11}(e_{22}\cdot e_{33}-e_{32}\cdot e_{23})$$

We do the same for the other terms in the first row with one very minor complication .. for second term we multiply by -1 and +1 for the third term:

$$-e_{12}(e_{21}\cdot e_{33}-e_{31}\cdot e_{23})+e_{13}(e_{21}\cdot e_{32}-e_{31}\cdot e_{22})$$

Cofactors are signed either +1 or -1 depending on the sum of their subscripts. Thus the sign of the cofactor is determined by:

$$(-1)^{i+j}$$

Thus, cofactor E_{11} has sign +1 and cofactor E_{21} has sign -1. One can visualise this nicely for a 3x3 determinant:

$$\begin{vmatrix} + & - & + \\ - & + & - \\ + & - & + \end{vmatrix}$$

So the whole determinant evaluation looks like:

$$\begin{aligned} \det \hat{\mathbf{E}} &= \begin{vmatrix} e_{11} & e_{12} & e_{13} \\ e_{21} & e_{22} & e_{23} \\ e_{31} & e_{32} & e_{33} \end{vmatrix} \\ &= e_{11} E_{11} + e_{12} E_{12} + e_{13} E_{13} \\ &= e_{11}(e_{22} \cdot e_{33} - e_{32} \cdot e_{23}) - e_{12}(e_{21} \cdot e_{33} - e_{31} \cdot e_{23}) + e_{13}(e_{21} \cdot e_{32} - e_{31} \cdot e_{22}) \\ &= e_{11} e_{22} e_{33} - e_{11} e_{32} e_{23} - e_{12} e_{21} e_{33} + e_{12} e_{31} e_{23} + e_{13} e_{21} e_{32} - e_{13} e_{31} e_{22} \end{aligned} \qquad [2\text{-}20]$$

This is a very simplified explanation and there is much more to it than this but this is all we will need. Most of our determinant evaluations will be for 3x3 matrices. For more in-depth explanations see any competent mathematics text.

Determinants have several interesting properties that may be taken advantage of when calculating their value.

First, interchanging rows or columns in the determinant changes the sign but not the value of the scalar result. An even number of interchanges produces no sign change and an odd number of interchanges produces a sign change. We can show this with our example determinant; we will exchange rows 1 and 2 to produce a new determinant, det $\hat{\mathbf{E}}'$:

$$\begin{aligned} \det \hat{\mathbf{E}}' &= \begin{vmatrix} e_{21} & e_{22} & e_{23} \\ e_{11} & e_{12} & e_{13} \\ e_{31} & e_{32} & e_{33} \end{vmatrix} \\ &= e_{21} E'_{21} + e_{22} E'_{22} + e_{23} E'_{23} \\ &= e_{21}(e_{12} \cdot e_{33} - e_{32} \cdot e_{13}) - e_{22}(e_{11} \cdot e_{33} - e_{31} \cdot e_{13}) + e_{23}(e_{11} \cdot e_{32} - e_{31} \cdot e_{12}) \\ &= e_{21} e_{12} e_{33} - e_{21} e_{32} e_{13} - e_{22} e_{11} e_{33} + e_{22} e_{31} e_{13} + e_{23} e_{11} e_{32} - e_{23} e_{31} e_{12} \end{aligned} \qquad [2\text{-}21]$$

This is equal to but of opposite sign to the result in [2-20].

Multiplication of each of the elements in a row or column in the determinant by a scalar results in the value of the determinant being multiplied by the scalar.

$$
\begin{aligned}
\det \hat{\boldsymbol{E}}'' &= \begin{vmatrix} \lambda e_{11} & \lambda e_{12} & \lambda e_{13} \\ e_{21} & e_{22} & e_{23} \\ e_{31} & e_{32} & e_{33} \end{vmatrix} \\
&= \lambda e_{11} E''_{11} + \lambda e_{12} E''_{12} + \lambda e_{13} E''_{13} \\
&= \lambda e_{11}(e_{22} \cdot e_{33} - e_{32} \cdot e_{23}) - \lambda e_{12}(e_{21} \cdot e_{33} - \lambda e_{31} \cdot e_{23}) + e_{13}(e_{21} \cdot e_{32} - e_{31} \cdot e_{22}) \\
&= \lambda \cdot \det \boldsymbol{E}''
\end{aligned}
\qquad [2\text{-}22]
$$

If every element of a row or column is equal to zero then the value of the determinant is zero.

$$
\begin{aligned}
\det \hat{\boldsymbol{E}} &= \begin{vmatrix} 0 & 0 & 0 \\ e_{21} & e_{22} & e_{23} \\ e_{31} & e_{32} & e_{33} \end{vmatrix} \\
&= 0\,E_{11} + 0\,E_{12} + 0\,E_{13} \\
&= 0(e_{22} \cdot e_{33} - e_{32} \cdot e_{23}) - 0(e_{21} \cdot e_{33} - e_{31} \cdot e_{23}) + 0(e_{21} \cdot e_{32} - e_{31} \cdot e_{22}) \\
&= 0
\end{aligned}
\qquad [2\text{-}23]
$$

The value of the determinant of the transpose of matrix **A** is equal to the value of the determinant of matrix **A**.

$$
\begin{aligned}
\det \hat{\boldsymbol{E}}^{T} &= \begin{vmatrix} e_{11} & e_{21} & e_{31} \\ e_{12} & e_{22} & e_{32} \\ e_{13} & e_{23} & e_{33} \end{vmatrix} \\
&= e_{11} E_{11} + e_{12} E_{12} + e_{13} E_{13} \\
&= e_{11}(e_{22} \cdot e_{33} - e_{32} \cdot e_{23}) - e_{12}(e_{21} \cdot e_{33} - e_{31} \cdot e_{23}) + e_{13}(e_{21} \cdot e_{32} - e_{31} \cdot e_{22}) \\
& e_{11} e_{22} e_{33} - e_{11} e_{32} e_{23} - e_{12} e_{21} e_{33} + e_{12} e_{31} e_{23} + e_{13} e_{21} e_{32} - e_{13} e_{31} e_{22} \\
&= \det \boldsymbol{E}
\end{aligned}
\qquad [2\text{-}24]
$$

If two rows or columns are identical then the value of the determinant is zero.

$$
\begin{aligned}
\det \hat{\boldsymbol{E}}''' &= \begin{vmatrix} e_{11} & e_{12} & e_{13} \\ e_{11} & e_{12} & e_{13} \\ e_{31} & e_{32} & e_{33} \end{vmatrix} \\
&= e_{11} E'''_{11} + e_{12} E'''_{12} + e_{13} E'''_{13} \\
&= e_{11}(e_{12} \cdot e_{33} - e_{32} \cdot e_{13}) - e_{12}(e_{11} \cdot e_{33} - e_{31} \cdot e_{13}) + e_{13}(e_{11} \cdot e_{32} - e_{31} \cdot e_{12}) \\
&= e_{11} e_{12} e_{33} - e_{11} e_{32} e_{13} - e_{12} e_{11} e_{33} + e_{12} e_{31} e_{13} + e_{13} e_{11} e_{32} - e_{13} e_{31} e_{12} \\
&= 0
\end{aligned}
\qquad [2\text{-}25]
$$

If a row or column in the determinant can be obtained by a combination of the other rows or columns in the matrix then the value of the determinant is zero.

$$\begin{aligned}
det\,\hat{\boldsymbol{E}} &= \begin{vmatrix} \alpha e_{21}+\beta e_{31} & \alpha e_{22}+\beta e_{32} & \alpha e_{23}+\beta e_{33} \\ e_{21} & e_{22} & e_{23} \\ e_{31} & e_{32} & e_{33} \end{vmatrix} \\
&= (\alpha e_{21}+\beta e_{31})E_{11}+(\alpha e_{22}+\beta e_{32})E_{12}+(\alpha e_{23}+\beta e_{33})E_{13} \\
&= (\alpha e_{21}+\beta e_{31})(e_{22}\cdot e_{33}-e_{32}\cdot e_{23}) \\
&\quad -(\alpha e_{22}+\beta e_{32})(e_{21}\cdot e_{33}-e_{31}e_{23}) \\
&\quad +(\alpha e_{23}+\beta e_{33})(e_{21}\cdot e_{32}-e_{31}\cdot e_{22}) \\
&= \alpha e_{21}e_{22}e_{33}+\beta e_{31}e_{22}e_{33}-\alpha e_{21}e_{32}e_{23}-\beta e_{31}e_{32}e_{23} \\
&\quad -\alpha e_{22}e_{21}e_{33}-\beta e_{32}e_{21}e_{33}+\alpha e_{22}e_{31}e_{23}+\beta e_{32}e_{31}e_{23} \\
&\quad +\alpha e_{23}e_{21}e_{32}+\beta e_{33}e_{21}e_{32}-\alpha e_{23}e_{31}e_{22}-\beta e_{33}e_{31}e_{22} \\
&= 0
\end{aligned} \qquad [2\text{-}26]$$

Close inspection of the resulting terms in [2-26] reveals that each one has its opposite and thus the entire expression collapses to zero. This result will be of use to us in our explorations of vectors.

The determinant of a product of matrices is:

$$det(\hat{\boldsymbol{A}}\hat{\boldsymbol{B}})=det\,\hat{\boldsymbol{A}}\,det\,\hat{\boldsymbol{B}} \qquad [2\text{-}27]$$

We can show this by using a pair of 2x2 matrices:

$$\begin{aligned}
&\hat{\boldsymbol{A}}=\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \quad \hat{\boldsymbol{B}}=\begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \\
&det\,\hat{\boldsymbol{A}}=a_{11}a_{22}-a_{21}a_{12} \\
&det\,\hat{\boldsymbol{B}}=b_{11}b_{22}-b_{21}b_{12} \\
&det\,\hat{\boldsymbol{A}}\,det\,\hat{\boldsymbol{B}}=(a_{11}a_{22}-a_{21}a_{12})(b_{11}b_{22}-b_{21}b_{12}) \\
&= a_{11}a_{22}b_{11}b_{22}-a_{11}a_{22}b_{21}b_{12}-a_{21}a_{12}b_{11}b_{22}+a_{21}a_{12}b_{21}b_{12}
\end{aligned}$$

Now, using this result and multiplying $\hat{\boldsymbol{A}}$ and $\hat{\boldsymbol{B}}$:

$$\begin{aligned}
\hat{\boldsymbol{A}}\hat{\boldsymbol{B}} &= \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \times \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \\
&= \begin{bmatrix} a_{11}b_{11}+a_{12}b_{21} & a_{11}b_{12}+a_{12}b_{22} \\ a_{21}b_{11}+a_{22}b_{21} & a_{21}b_{12}+a_{22}b_{22} \end{bmatrix} \\
det(\hat{\boldsymbol{A}}\hat{\boldsymbol{B}}) &= (a_{11}b_{11}+a_{12}b_{21})(a_{21}b_{12}+a_{22}b_{22})-(a_{21}b_{11}+a_{22}b_{21})(a_{11}b_{12}+a_{12}b_{22}) \\
&= a_{11}b_{11}a_{21}b_{12}+a_{11}b_{11}a_{22}b_{22}+a_{12}b_{21}a_{21}b_{12}+a_{12}b_{21}a_{22}b_{22} \\
&\quad -a_{21}b_{11}a_{11}b_{12}-a_{21}b_{11}a_{12}b_{22}-a_{22}b_{21}a_{11}b_{12}-a_{22}b_{21}a_{12}b_{22} \\
&= a_{11}b_{11}a_{22}b_{22}+a_{12}b_{21}a_{21}b_{12}-a_{21}b_{11}a_{12}b_{22}-a_{22}b_{21}a_{11}b_{12} \\
&= det\,\hat{\boldsymbol{A}}\,det\,\hat{\boldsymbol{B}}
\end{aligned} \qquad [2\text{-}28]$$

The determinant of the identity matrix is equal to one:

$$\begin{aligned}
&det\begin{bmatrix}1&0&0&\cdots\\0&1&0&\cdots\\0&0&1&\cdots\\\vdots&\vdots&\vdots&\ddots\end{bmatrix}\\
&=1\times det\begin{bmatrix}1&0&\cdots\\0&1&\cdots\\\vdots&\vdots&\ddots\end{bmatrix}+(n-1)\times 0\times det\begin{bmatrix}1&0&\cdots\\0&1&\cdots\\\vdots&\vdots&\ddots\end{bmatrix}\\
&=1\times det\begin{bmatrix}1&0&\cdots\\0&1&\cdots\\\vdots&\vdots&\ddots\end{bmatrix}\\
&=1\times 1\times det\begin{bmatrix}1&\cdots\\0&\ddots\end{bmatrix}+(n-2)\times 0\times det\begin{bmatrix}1&\cdots\\0&\ddots\end{bmatrix}\\
&=1\times 1\times det\begin{bmatrix}1&\cdots\\0&\ddots\end{bmatrix}\\
&etc.\\
&\downarrow\\
&=\Pi_1^n(1\times 1)=1
\end{aligned}$$

The determinant of an inverse matrix is simply related to the original matrix:

$$\begin{gathered}
\hat{\mathbf{A}}^{-1}\hat{\mathbf{A}}=\hat{\mathbf{1}}\\
det\left(\hat{\mathbf{A}}^{-1}\hat{\mathbf{A}}\right)=det\,\hat{\mathbf{1}}\\
det\,\hat{\mathbf{A}}^{-1}det\,\hat{\mathbf{A}}=1\\
det\,\hat{\mathbf{A}}^{-1}=\frac{1}{det\,\hat{\mathbf{A}}}
\end{gathered}\qquad [2\text{-}29]$$

We are in a position to define the adjoint matrix now. For our general matrix, $\hat{\boldsymbol{E}}$, the adjoint is:

$$adj\,\hat{\mathbf{E}}=\begin{bmatrix}E_{11}&E_{21}&E_{31}\\E_{12}&E_{22}&E_{32}\\E_{13}&E_{23}&E_{33}\end{bmatrix}\qquad [2\text{-}30]$$

where $\hat{\boldsymbol{E}}_{ij}$ is the cofactor of element e_{ij} in $\hat{\boldsymbol{E}}$. Note that this is a *transposed* matrix of cofactors.

The determinant of the transpose of a matrix is equal to the determinant of the original matrix:

$$det\,\hat{\mathbf{A}}=det\,\hat{\mathbf{A}}^{T}\qquad [2\text{-}31]$$

We show this for a 3 x 3 matrix:

$$det\,\hat{\mathbf{A}} = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}$$
$$= a_{11}(a_{22}a_{33} - a_{32}a_{23}) - a_{12}(a_{21}a_{33} - a_{31}a_{23}) + a_{13}(a_{21}a_{32} - a_{31}a_{22})$$
$$= a_{11}a_{22}a_{33} - a_{11}a_{32}a_{23} - a_{12}a_{21}a_{33} + a_{12}a_{31}a_{23} + a_{13}a_{21}a_{32} - a_{13}a_{31}a_{22}$$

$$det\,\hat{\mathbf{A}}^T = \begin{vmatrix} a_{11} & a_{21} & a_{31} \\ a_{12} & a_{22} & a_{32} \\ a_{13} & a_{23} & a_{33} \end{vmatrix}$$
$$= a_{11}(a_{22}a_{33} - a_{23}a_{32}) - a_{21}(a_{12}a_{33} - a_{13}a_{32}) + a_{31}(a_{12}a_{23} - a_{13}a_{22})$$
$$= a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} - a_{21}a_{12}a_{33} + a_{21}a_{13}a_{32} + a_{31}a_{12}a_{23} - a_{31}a_{13}a_{22}$$

$$det\,\hat{\mathbf{A}} = det\,\hat{\mathbf{A}}^T$$

2.3 Solving Matrix Equations

Let's suppose that we have three equations:

$$6\mathrm{x} - 4\mathrm{y} + z = 3$$
$$x - 2\mathrm{y} + 5z = 1$$
$$3\mathrm{x} + 2\mathrm{y} + z = 0$$

We can represent these equations using three matrices:

$$\hat{\mathbf{A}} = \begin{bmatrix} 6 & -4 & 1 \\ 1 & -2 & +5 \\ 3 & 2 & 1 \end{bmatrix}$$
$$\boldsymbol{x} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$
$$\boldsymbol{d} = \begin{bmatrix} 3 \\ 1 \\ 0 \end{bmatrix}$$

Then we write:

$$\hat{\mathbf{A}}\boldsymbol{x} = \boldsymbol{d}$$
$$\begin{bmatrix} 6 & -4 & 1 \\ 1 & -2 & +5 \\ 3 & 2 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \\ 0 \end{bmatrix}$$

which regenerates our original equations. Now what we usually want to do is solve for x, y and z. What if we were to multiply both side of the matrix equation by the inverse of $\hat{\mathbf{A}}$?

$$\hat{\boldsymbol{A}}^{-1}\hat{\boldsymbol{A}}\,\boldsymbol{x}=\hat{\boldsymbol{A}}^{-1}\boldsymbol{d}$$

and

$$\boldsymbol{x}=\hat{\boldsymbol{A}}^{-1}\boldsymbol{d}$$

Matrix $\boldsymbol{x}$ now holds the solutions that we are looking for. The problem now is “how do we find the inverse matrix?”. We will find out how on our general matrix, $\hat{\boldsymbol{E}}$:

$$\hat{\boldsymbol{E}}\,\boldsymbol{x}=\boldsymbol{d}$$

$$\begin{bmatrix} e_{11} & e_{12} & e_{13} \\ e_{21} & e_{22} & e_{23} \\ e_{31} & e_{32} & e_{33} \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix}=\begin{bmatrix} d_1 \\ d_2 \\ d_3 \end{bmatrix}$$

In equation form this looks like:

$$\begin{gathered} e_{11}x+e_{12}y+e_{13}z=d_1 \\ e_{21}x+e_{22}y+e_{23}z=d_2 \\ e_{31}x+e_{32}y+e_{33}z=d_3 \end{gathered}$$

The inverse of matrix $\hat{\boldsymbol{E}}$ is:

$$\hat{\boldsymbol{E}}^{-1}=\frac{adj\,\boldsymbol{E}}{det\,\boldsymbol{E}} \qquad [2\text{-}32]$$

Note that since this depends on the value of det $\hat{\boldsymbol{E}}$, if det $\hat{\boldsymbol{E}}$ equals zero the inverse of $\hat{\boldsymbol{E}}$ is undefined. We can prove this (assuming det $\hat{\boldsymbol{E}}$ does not equal zero):

$$\hat{\mathbf{E}}\times\frac{adj\,\hat{\mathbf{E}}}{det\,\hat{\mathbf{E}}}=$$

$$\frac{\begin{bmatrix} e_{11} & e_{12} & e_{13} \\ e_{21} & e_{22} & e_{23} \\ e_{31} & e_{32} & e_{33} \end{bmatrix}\times\begin{bmatrix} E_{11} & E_{21} & E_{31} \\ E_{12} & E_{22} & E_{32} \\ E_{13} & E_{23} & E_{33} \end{bmatrix}}{det\,\boldsymbol{E}}$$

$$=\frac{1}{det\,\hat{\mathbf{E}}}\begin{bmatrix} e_{11}E_{11}+e_{12}E_{12}+e_{13}E_{13} & e_{11}E_{21}+e_{12}E_{22}+e_{13}E_{23} & e_{11}E_{31}+e_{12}E_{32}+e_{13}E_{33} \\ e_{21}E_{11}+e_{22}E_{12}+e_{23}E_{13} & e_{21}E_{21}+e_{22}E_{22}+e_{23}E_{23} & e_{21}E_{31}+e_{22}E_{32}+e_{23}E_{33} \\ e_{31}E_{11}+e_{32}E_{12}+e_{33}E_{13} & e_{31}E_{21}+e_{32}E_{22}+e_{33}E_{23} & e_{31}E_{31}+e_{32}E_{32}+e_{33}E_{33} \end{bmatrix}$$

$$=\frac{1}{det\,\hat{\mathbf{E}}}\begin{bmatrix} det\,\hat{\mathbf{E}} & 0 & 0 \\ 0 & det\,\hat{\mathbf{E}} & 0 \\ 0 & 0 & det\,\hat{\mathbf{E}} \end{bmatrix}=\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}=\hat{\mathbf{1}}$$

The diagonal terms are the definitions of det $\hat{\boldsymbol{E}}$ and of course det $\hat{\boldsymbol{E}}$ divided by det $\hat{\boldsymbol{E}}$ equals 1 but why are the off-diagonal terms zero? Let's look at one of them:

$$\begin{aligned}
&e_{21}E_{11}+e_{22}E_{12}+e_{23}E_{13}\\
&=e_{21}\,det\begin{bmatrix}e_{22}&e_{23}\\e_{32}&e_{33}\end{bmatrix}-e_{22}\,det\begin{bmatrix}e_{21}&e_{23}\\e_{31}&e_{33}\end{bmatrix}+e_{23}\,det\begin{bmatrix}e_{21}&e_{22}\\e_{31}&e_{32}\end{bmatrix}\\
&=e_{21}(e_{22}e_{33}-e_{32}e_{23})-e_{22}(e_{21}e_{33}-e_{31}e_{23})+e_{23}(e_{21}e_{32}-e_{31}e_{22})\\
&=e_{21}e_{22}e_{33}-e_{21}e_{32}e_{23}-e_{22}e_{21}e_{33}+e_{22}e_{31}e_{23}+e_{23}e_{21}e_{32}-e_{23}e_{31}e_{22}\\
&=0
\end{aligned}$$

We now write:

$$\begin{aligned}
\boldsymbol{x}&=\hat{\boldsymbol{E}}^{-1}\boldsymbol{d}\\
&=\frac{adj\,\hat{\boldsymbol{E}}}{det\,\hat{\boldsymbol{E}}}\boldsymbol{d}
\end{aligned}$$

Thus, if we can determine the value of det $\hat{\boldsymbol{E}}$ and what the adjoint matrix of $\hat{\boldsymbol{E}}$ is we can solve for the variables in the equations.

So, using our example at the beginning of this section:

$$\begin{aligned}
\hat{\mathbf{A}}&=\begin{bmatrix}6&-4&1\\1&-2&5\\3&2&1\end{bmatrix}\\
det\,\hat{\mathbf{A}}&=6\cdot((-2)\cdot1-2\cdot5)-(-4)\cdot(1\cdot1-3\cdot5)+1\cdot(1\cdot2-(-2)\cdot3)\\
&=-72-56+8\\
&=-120\\
adj\,\hat{\mathbf{A}}&=\begin{bmatrix}A_{11}&A_{21}&A_{31}\\A_{12}&A_{22}&A_{32}\\A_{13}&A_{23}&A_{33}\end{bmatrix}\\
&=\begin{bmatrix}-12&6&-18\\14&3&-29\\8&-24&-8\end{bmatrix}
\end{aligned}$$

and we set up the equations:

$$\mathbf{x} = \mathbf{A}^{-1}\mathbf{d}$$
$$= \frac{adj\,\hat{\mathbf{A}}}{det\,\hat{\mathbf{A}}}\mathbf{d}$$
$$= \frac{\begin{bmatrix} -12 & 6 & -18 \\ 14 & 3 & -29 \\ 8 & -24 & -8 \end{bmatrix}}{-120}\begin{bmatrix} 3 \\ 1 \\ 0 \end{bmatrix}$$
$$= \begin{bmatrix} \frac{-12}{-120} & \frac{6}{-120} & \frac{-18}{-120} \\ \frac{14}{-120} & \frac{3}{-120} & \frac{-29}{-120} \\ \frac{8}{-120} & \frac{-24}{-120} & \frac{-8}{-120} \end{bmatrix}\begin{bmatrix} 3 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{3\cdot(-12)+6\cdot1-18\cdot0}{-120} \\ \frac{14\cdot3+3\cdot1-29\cdot0}{-120} \\ \frac{8\cdot3-24\cdot1-8\cdot0}{-120} \end{bmatrix}$$
$$= \begin{bmatrix} \frac{-30}{-120} \\ \frac{45}{-120} \\ 0 \end{bmatrix}$$

and:

$$x = \frac{-30}{-120} = 0.25$$
$$y = \frac{45}{-120} = -0.375$$
$$z = 0$$

We can check our arithmetic in producing the inverse matrix using:

$$\hat{\mathbf{A}} \times \mathbf{A}^{-1} = \mathbf{1}$$
$$= \hat{\mathbf{A}} \times adj\,\hat{\mathbf{A}} \times \frac{1}{det\,\hat{\mathbf{A}}}$$
$$= \begin{bmatrix} 6 & -4 & 1 \\ 1 & -2 & +5 \\ 3 & 2 & 1 \end{bmatrix} \times \begin{bmatrix} -12 & 6 & -18 \\ 14 & 3 & -29 \\ 8 & -24 & -8 \end{bmatrix} \frac{1}{-120}$$
$$= \begin{bmatrix} -120 & 0 & 0 \\ 0 & -120 & 0 \\ 0 & 0 & -120 \end{bmatrix} \frac{1}{-120}$$
$$= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \hat{\mathbf{1}}$$

There is another way to solve for the x's by a technique known as Gaussian elimination. We use a simple set of elementary row operations that follow directly from the algebra of equations:

(I) any equation can be multiplied by a non-zero constant. This follows from the fact that doing the same thing to each side of an equation leaves each side equal to each other.
(II) any two equations can be interchanged. It does not matter what order the equation are presented to you .. any order of presentation is surely arbitrary.
(III) any equation can be replaced by the sum of itself and a multiple of another equation. This is really just rule (I) in a bit more elaborate form.

Let's use our example equations from the beginning of this section:

eq 1: $6x-4y+z=3$
eq 2: $x-2y+5z=1$
eq 3: $3x+2y+z=0$

From rule (I) we can, say, multiply equation 3 by 6:

eq 1: $6x-4y+z=3$
eq 2: $x-2y+5z=1$
eq 3 times 6: $18x+12y+6z=0$

and from rule (III) we can add it to equation 1:

eq 1 + eq 3: $24x-8y+7z=3$
eq 2: $x-2y+5z=1$
eq 3 times 6: $18x+12y+6z=0$

and so on.

How does this help us? We can use these rules to solve the equations by Gaussian elimination. The technique is to manipulate the equations using the elementary operations so that in one of the equations the coefficients of x and y are zero. The value of z can then be solved for and back-substituted to solve for x and y. Also, there is not need to write the entire equation but only the coefficients and the right-hand side of the equations in an augmented matrix:

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

We apply our elementary row operations to get the matrix into what is called echelon form in which there are all zeros below the diagonal elements starting in the upper left. Let's proceed:

$$\begin{bmatrix} 6 & -4 & 1 & 3 \\ 1 & -2 & 5 & 1 \\ 3 & 2 & 1 & 0 \end{bmatrix} \overset{r3\times 2}{\rightarrow} \begin{bmatrix} 6 & -4 & 1 & 3 \\ 1 & -2 & 5 & 1 \\ 6 & 4 & 2 & 0 \end{bmatrix} \overset{r3-r1}{\rightarrow} \begin{bmatrix} 6 & -4 & 1 & 3 \\ 1 & -2 & 5 & 1 \\ 0 & 8 & 1 & -3 \end{bmatrix}$$

Here we have multiplied row 3 by 2 and then subtracted row 1 from row 3. Next:

$$\begin{bmatrix} 6 & -4 & 1 & 3 \\ 1 & -2 & 5 & 1 \\ 0 & 8 & 1 & -3 \end{bmatrix} \overset{r2\times 6}{\rightarrow} \begin{bmatrix} 6 & -4 & 1 & 3 \\ 6 & -12 & 30 & 6 \\ 0 & 8 & 1 & -3 \end{bmatrix} \overset{r2-r1}{\rightarrow} \begin{bmatrix} 6 & -4 & 1 & 3 \\ 0 & -8 & 29 & 3 \\ 0 & 8 & 1 & -3 \end{bmatrix} \overset{r2+r3}{\rightarrow} \begin{bmatrix} 6 & -4 & 1 & 3 \\ 0 & -8 & 29 & 3 \\ 0 & 0 & 30 & 0 \end{bmatrix}$$

The final matrix is in echelon form and written out as a set of equations is:

$$\begin{matrix} 6x & -4y & z & = & 3 \\ 0x & -8y & 29z & = & 3 \\ 0x & 0y & 30z & = & 0 \end{matrix}$$

We solve for z in the last equation to get z = 0. Substituting this into equation 2 we solve for y to get y = -8/3. Then substituting these values in equation 1 we get x = -21/18.

It is quite possible that there will be either no solution or and infinite number of solutions to a set of equations. For example:

$$\begin{gathered} x+y-z=3 \\ 3x-y+3z=5 \\ x-y+2z=2 \end{gathered}$$

These equations cannot be solved using the inverse matrix method and the Gaussian elimination gives the absurd conclusion that 0 = 1. Try it for yourself.

If we alter the above set of equations to produce another example:

$$\begin{gathered} x+y-z=1 \\ 3x-y+3z=5 \\ x-y+2z=2 \end{gathered}$$

Gaussian elimination using the augmented matrix gives us:

$$\begin{bmatrix}1 & 1 & -1 & 1\\ 3 & -1 & 3 & 5\\ 1 & -1 & 2 & 2\end{bmatrix} \underset{r3-r1}{\overset{r2-3r1}{\rightarrow}} \begin{bmatrix}1 & 1 & -1 & 1\\ 0 & -4 & 6 & 2\\ 0 & -2 & 3 & 1\end{bmatrix} \overset{r3-\frac{1}{2}r1}{\rightarrow} \begin{bmatrix}1 & 1 & -1 & 1\\ 0 & -4 & 6 & 2\\ 0 & 0 & 0 & 0\end{bmatrix}$$

in which the last matrix is in echelon form. When we look at solving for z we see that *any* value of z will work:

$$z=\eta$$

where η is any number. Back-substituting, we get for x and y:

$$x=\frac{3}{2}-\frac{\eta}{2}$$
$$y=\frac{-1}{2}+\frac{3}{2}\eta$$

Thus, for this set of equations there are an infinite number of solutions.

Under some circumstances the matrix equation is said to be *homogeneous*:

$$\hat{\boldsymbol{A}}\hat{\boldsymbol{x}}=\hat{\boldsymbol{0}} \quad [2\text{-}33]$$

Obviously, we cannot use an inverse matrix, $\mathbf{A}^{-1}$, to get anything other than a solution of zero for x:

$$\hat{\boldsymbol{A}}^{-1}\hat{\boldsymbol{A}}\boldsymbol{x}=\hat{\boldsymbol{A}}^{-1}\hat{\boldsymbol{0}}$$
$$\boldsymbol{x}=\hat{\boldsymbol{A}}^{-1}\mathbf{0}=\mathbf{0}$$

This type of solution is called a *trivial* solution to the equation. There may well be *non-trivial* solutions, however. If det **A** = 0 there will be an infinite number of non-trivial solutions and if det **A** is not equal to zero the only solution is **x** = **0**.

The usefulness of this is in the search for *eigenvalues* of a matrix. Let's suppose that we have the matrix equation:

$$\hat{\boldsymbol{A}}\boldsymbol{x}=\lambda\boldsymbol{x} \quad [2\text{-}34a]$$

or:

$$\hat{\boldsymbol{A}}\boldsymbol{x}-\lambda\boldsymbol{x}=\hat{\boldsymbol{0}} \quad [2\text{-}34b]$$

or:

$$(\hat{\boldsymbol{A}}-\lambda\hat{\mathbf{1}})\boldsymbol{x}=\hat{\boldsymbol{0}} \quad [2\text{-}34c]$$

This is called the *characteristic equation*. The $\lambda's$ are the eigenvalues and x is the *eigenvector* or more properly, the right eigenvector. If:

$$\hat{\mathbf{A}}^T \boldsymbol{x} = \lambda \boldsymbol{x}$$

then:

$$\lambda \boldsymbol{x}^T = (\lambda \boldsymbol{x})^T = (\hat{\mathbf{A}}^T \boldsymbol{x})^T = \boldsymbol{x}^T \hat{\mathbf{A}}$$

where x is now referred to as the *left eigenvector*.

There will be non-trivial solutions for the characteristic equation only if:

$$det(\hat{\mathbf{A}} - \lambda \hat{\mathbf{1}}) = 0 \qquad [2\text{-}35]$$

or:

$$\begin{vmatrix} a_{11}-\lambda & a_{12} & a_{13} & \dots \\ a_{21} & a_{22}-\lambda & a_{23} & \dots \\ a_{31} & a_{32} & \ddots & \dots \\ \dots & \dots & \dots & a_{nn}-\lambda \end{vmatrix} = 0$$

which will evaluate to a polynomial known as the *characteristic polynomial*.

An example:

$$\hat{\mathbf{A}} = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 4 & 0 \\ 2 & 5 & 2 \end{bmatrix}$$

The characteristic equation is:

$$\begin{vmatrix} 2-\lambda & 1 & 0 \\ 1 & 4-\lambda & 0 \\ 2 & 5 & 2-\lambda \end{vmatrix} = 0$$

Multiplying out gives:

$$\begin{aligned} &(2-\lambda)\begin{vmatrix} 4-\lambda & 0 \\ 5 & 2-\lambda \end{vmatrix} - \begin{vmatrix} 1 & 0 \\ 2 & 2-\lambda \end{vmatrix} \\ &= (2-\lambda)(4-\lambda)(2-\lambda) - (2-\lambda) \\ &= (2-\lambda)[\lambda^2 - 6\lambda + 7] \\ &= (2-\lambda)(3+\sqrt{2}-\lambda)(3-\sqrt{2}-\lambda) = 0 \end{aligned}$$

Thus, λ can equal 2, $3+\sqrt{2}$ or $3-\sqrt{2}$ in order to satisfy the characteristic equation. Note that since we have solved a polynomial for its roots it is quite possible that those roots will be complex numbers and we must be prepared for that eventuality. A quick check of these eigenvalues is to calculate the trace of the matrix and compare it to the sum of the eigenvalues (see equation [2-15]). They should be equal:

$$
\begin{gathered}
Tr\,\hat{\mathbf{A}}=2+4+2=8 \\
\sum_{i}^{n}\lambda_i=2+(3+\sqrt{2})+(3-\sqrt{2})=8 \\
Tr\,\hat{\mathbf{A}}=\sum_{i}^{n}\lambda_i
\end{gathered}
\qquad [2\text{-}36]
$$

We defer proof of this for a bit later.

Knowing the eigenvalues, we can find possible eigenvectors from our eigenvalue equation [2-31a], starting with l=2:

$$
\begin{gathered}
\hat{\mathbf{A}}\boldsymbol{x}=\lambda\boldsymbol{x} \\
\begin{bmatrix} 2 & 1 & 0 \\ 1 & 4 & 0 \\ 2 & 5 & 2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}=2\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}=\begin{bmatrix} 2\mathrm{x}_1 \\ 2\mathrm{x}_2 \\ 2\mathrm{x}_3 \end{bmatrix}
\end{gathered}
$$

Our simultaneous equations are:

$$
\begin{gathered}
2\mathrm{x}_1+x_2=2\mathrm{x}_1 \\
x_1+4\mathrm{x}_2=2\mathrm{x}_2 \\
2\mathrm{x}_1+5\mathrm{x}_2+2\mathrm{x}_3=2\mathrm{x}_3
\end{gathered}
$$

Solution of these simultaneous equations shows that $x_1 = x_2 = 0$ A possible value for x_3 could be, say, 1. This being the case, a possible eigenvector is:

$$
\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}
$$

This is, of course, one of an infinite number of possible eigenvectors, all of which have different values of x_3. For the other values of λ :

$$\lambda = 3+\sqrt{2}$$

$$\begin{bmatrix} 2 & 1 & 0 \\ 1 & 4 & 0 \\ 2 & 5 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = (3+\sqrt{2}) \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} (3+\sqrt{2})x_1 \\ (3+\sqrt{2})x_2 \\ (3+\sqrt{2})x_3 \end{bmatrix}$$

$$2\mathrm{x}_1 + x_2 = (3+\sqrt{2})x_1$$

$$x_1 + 4\mathrm{x}_2 = (3+\sqrt{2})x_2$$

$$2\mathrm{x}_1 + 5\mathrm{x}_2 + 2\mathrm{x}_3 = (3+\sqrt{2})x_3$$

$$x_1 = x_1$$

$$x_2 = (1+\sqrt{2})x_1$$

$$x_3 = (3+2\sqrt{2})x_1$$

and a possible eigenvector is:

$$\begin{bmatrix} 1 \\ 1+\sqrt{2} \\ 3+2\sqrt{2} \end{bmatrix}$$

$$\lambda = 3-\sqrt{2}$$

$$\begin{bmatrix} 2 & 1 & 0 \\ 1 & 4 & 0 \\ 2 & 5 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = (3-\sqrt{2}) \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} (3-\sqrt{2})x_1 \\ (3-\sqrt{2})x_2 \\ (3-\sqrt{2})x_3 \end{bmatrix}$$

$$2\mathrm{x}_1 + x_2 = (3-\sqrt{2})x_1$$

$$x_1 + 4\mathrm{x}_2 = (3-\sqrt{2})x_2$$

$$2\mathrm{x}_1 + 5\mathrm{x}_2 + 2\mathrm{x}_3 = (3-\sqrt{2})x_3$$

$$x_1 = x_1$$

$$x_2 = (1-\sqrt{2})x_1$$

$$x_3 = (3-2\sqrt{2})x_1$$

and a possible eigenvector is:

$$\begin{bmatrix} 1 \\ 1-\sqrt{2} \\ 3-2\sqrt{2} \end{bmatrix}$$

These results are easily checked by doing the matrix multiplication:

$$\hat{\boldsymbol{A}}\boldsymbol{x} = \lambda \boldsymbol{x}$$

and substituting the appropriate eigenvector. Let's check using the last eigenvector:

$$\hat{\mathbf{A}}\boldsymbol{x}=\lambda\boldsymbol{x}$$
$$\begin{bmatrix}2&1&0\\1&4&0\\2&5&2\end{bmatrix}\begin{bmatrix}1\\1-\sqrt{2}\\3-2\sqrt{2}\end{bmatrix}=\begin{bmatrix}3-\sqrt{2}\\5-4\sqrt{2}\\13-9\sqrt{2}\end{bmatrix}=(3-\sqrt{2})\begin{bmatrix}1\\1-\sqrt{2}\\3-2\sqrt{2}\end{bmatrix} \quad [2\text{-}37]$$

Lovely! You can check the other two eigenvectors for yourself.

A step further is to combine the separate column eigenvectors, $\boldsymbol{x}$, into one matrix, $\boldsymbol{X}$, and to put the eigenvalues into a diagonal matrix, $\boldsymbol{\Lambda}$, so that we have a characteristic equation with all eigenvalues and eigenvectors:

$$\hat{\mathbf{A}}\boldsymbol{X}=\boldsymbol{X}\hat{\boldsymbol{\Lambda}} \quad [2\text{-}38]$$

Why put $\boldsymbol{X}$ first and $\boldsymbol{\Lambda}$ second on the right side of the equation instead of the other way around? We have constructed $\boldsymbol{X}$ from our column vectors, $\boldsymbol{x}_i$, each of which is associated with an eigenvalue, λ_i . You can see an example of this relationship in equation [2-36]. In order to maintain this association we must multiply $\boldsymbol{\Lambda}$ from the *left* by $\boldsymbol{X}$:

$$\boldsymbol{X}\hat{\boldsymbol{\Lambda}}$$
$$=\begin{bmatrix}X_{11}&X_{12}&X_{13}\\X_{21}&X_{22}&X_{23}\\X_{31}&X_{32}&X_{33}\end{bmatrix}\begin{bmatrix}\lambda_1&0&0\\0&\lambda_2&0\\0&0&\lambda_3\end{bmatrix}$$
$$=\begin{bmatrix}\lambda_1X_{11}&\lambda_2X_{12}&\lambda_3X_{13}\\\lambda_1X_{21}&\lambda_2X_{22}&\lambda_3X_{23}\\\lambda_1X_{31}&\lambda_2X_{32}&\lambda_3X_{33}\end{bmatrix}$$

Thus, λ_1 is associated with column 1 which is one of its eigenvectors. Similarly, λ_2 and λ_3 are in columns 2 and 3 respectively.

Using [2-17] and [2-38], we can now offer a proof of [2-36]. Multiplication of both sides of [2-38] by $\boldsymbol{X}^{-1}$ gives:

$$\begin{gathered}\boldsymbol{X}^{-1}\hat{\mathbf{A}}\boldsymbol{X}=\boldsymbol{X}^{-1}\boldsymbol{X}\hat{\boldsymbol{\Lambda}}\\=\boldsymbol{\Lambda}\\Tr(\boldsymbol{X}^{-1}\hat{\mathbf{A}}\boldsymbol{X})=Tr(\hat{\boldsymbol{\Lambda}})\\Tr(\boldsymbol{X}\boldsymbol{X}^{-1}\hat{\mathbf{A}})=Tr(\hat{\boldsymbol{\Lambda}})\\Tr(\hat{\mathbf{A}})=Tr(\hat{\boldsymbol{\Lambda}})\\Tr(\hat{\mathbf{A}})=\sum_i\lambda_i\end{gathered} \quad [2\text{-}39]$$

2.4 Transforms

Two n x n matrices $\hat{\boldsymbol{A}}$ and $\hat{\boldsymbol{B}}$ are said to be *similar matrices* if there exists an invertible n x n matrix, $\hat{\boldsymbol{T}}$ such that:

$$\hat{\boldsymbol{B}}=\hat{\boldsymbol{T}}^{-1}\hat{\boldsymbol{A}}\hat{\boldsymbol{T}} \qquad [2\text{-}40]$$

and the transformation from $\hat{\boldsymbol{A}}$ to $\hat{\boldsymbol{B}}$ is called a *similarity transform*. We can back-transform $\hat{\boldsymbol{B}}$ to $\hat{\boldsymbol{A}}$ again:

$$\begin{aligned}\hat{\boldsymbol{T}}\hat{\boldsymbol{B}}\hat{\boldsymbol{T}}^{-1}&=\hat{\boldsymbol{T}}\hat{\boldsymbol{T}}^{-1}\hat{\boldsymbol{A}}\hat{\boldsymbol{T}}\hat{\boldsymbol{T}}^{-1}\\&=\hat{\boldsymbol{A}}\end{aligned}$$

Note that reversing the order of the **T** matrices gives a different matrix, $\hat{\boldsymbol{C}}$, but is still, of course, a similarity transformation:

$$\hat{\boldsymbol{C}}=\hat{\boldsymbol{T}}\hat{\boldsymbol{A}}\hat{\boldsymbol{T}}^{-1}$$

Note, also, that $\hat{\boldsymbol{C}}$ is similar to $\hat{\boldsymbol{A}}$ *and* to $\hat{\boldsymbol{B}}$:

$$\begin{aligned}\hat{\boldsymbol{C}}&=\hat{\boldsymbol{T}}\hat{\boldsymbol{A}}\hat{\boldsymbol{T}}^{-1}\\ \hat{\boldsymbol{A}}&=\hat{\boldsymbol{T}}^{-1}\hat{\boldsymbol{T}}\hat{\boldsymbol{A}}\hat{\boldsymbol{T}}^{-1}\hat{\boldsymbol{T}}=\hat{\boldsymbol{T}}^{-1}\hat{\boldsymbol{C}}\hat{\boldsymbol{T}}\\ \hat{\boldsymbol{B}}&=\hat{\boldsymbol{T}}^{-1}\hat{\boldsymbol{A}}\hat{\boldsymbol{T}}=\hat{\boldsymbol{T}}^{-1}\hat{\boldsymbol{T}}^{-1}\hat{\boldsymbol{C}}\hat{\boldsymbol{T}}\hat{\boldsymbol{T}}=(\hat{\boldsymbol{T}}^{-1})^2\hat{\boldsymbol{C}}(\hat{\boldsymbol{T}})^2\end{aligned}$$

The reason that $\hat{\boldsymbol{A}}$ and $\hat{\boldsymbol{B}}$ (and $\hat{\boldsymbol{C}}$) are called similar is that they have several properties in common. First, the determinants of $\hat{\boldsymbol{A}}$ and $\hat{\boldsymbol{B}}$ are the same. We can show this quite easily using [2-27] and [2-28]:

$$\begin{aligned}\hat{\boldsymbol{B}}&=\hat{\boldsymbol{T}}^{-1}\hat{\boldsymbol{A}}\hat{\boldsymbol{T}}\\ det\,\hat{\boldsymbol{B}}&=det\,(\hat{\boldsymbol{T}}^{-1}\hat{\boldsymbol{A}}\hat{\boldsymbol{T}})\\&=det\,\hat{\boldsymbol{T}}^{-1}det\,\hat{\boldsymbol{A}}\,det\,\hat{\boldsymbol{T}}\\&=det\,\hat{\boldsymbol{A}}\end{aligned} \qquad [2\text{-}41]$$

Next, using [2-17], the trace of $\hat{\boldsymbol{A}}$ and $\hat{\boldsymbol{B}}$ are equal:

$$\begin{aligned}\hat{\boldsymbol{B}}&=\hat{\boldsymbol{T}}^{-1}\hat{\boldsymbol{A}}\hat{\boldsymbol{T}}\\ tr\,\hat{\boldsymbol{B}}&=tr\,(\hat{\boldsymbol{T}}^{-1}\hat{\boldsymbol{A}}\hat{\boldsymbol{T}})\\&=tr\,(\hat{\boldsymbol{T}}^{-1}\hat{\boldsymbol{T}}\hat{\boldsymbol{A}})\\&=tr\,(\hat{\boldsymbol{1}}\hat{\boldsymbol{A}})\\&=tr\,\hat{\boldsymbol{A}}\end{aligned} \qquad [2\text{-}42]$$

Two matrices related by a similarity transformation have the same eigenvalues. If we have matrices $\hat{\boldsymbol{A}}$ and $\hat{\boldsymbol{B}}$ related by a similarity transformation and eigenvector matrix $\hat{\boldsymbol{X}}$ then:

$$\begin{aligned}\hat{\boldsymbol{B}}&=\hat{\boldsymbol{T}}\hat{\boldsymbol{A}}\hat{\boldsymbol{T}}^{-1}\\ \hat{\boldsymbol{A}}\boldsymbol{X}&=\lambda\boldsymbol{X}\end{aligned}$$

Now, using the transformation matrix $\hat{T}$, we write:

$$\begin{aligned} \hat{B}\hat{T}X &= (\hat{T}\hat{A}\hat{T}^{-1})\hat{T}X \\ &= \hat{T}(\hat{A}X) \\ &= \hat{T}(\lambda X) \\ &= \lambda\hat{T}X \end{aligned} \quad [2\text{-}43]$$

Thus, if $\hat{X}$ is an eigenvector matrix of $\hat{A}$ then $\hat{T}\hat{X}$ is an eigenvector of $\hat{B}$ and $\hat{B}$ has the same eigenvalues as $\hat{A}$.

From a practical point of view, the order in which the (left-to-right) multiplication of the matrices is done is irrelevant. That is, for the similarity transform:

$$\hat{B} = \hat{T}^{-1}\hat{A}\hat{T}$$

if we make the definitions:

$$\hat{X} = \hat{T}^{-1}\hat{A}$$
and
$$\hat{Y} = \hat{A}\hat{T}$$

then we can write our similarity transform in two equivalent ways:

$$\hat{B} = \hat{X}\hat{T}$$
or
$$\hat{B} = \hat{T}^{-1}\hat{Y}$$

each representing a different order of multiplication of matrices $\hat{T}^{-1}$, $\hat{A}$ and $\hat{T}$.

Another type of transform is the *orthogonal* transformation which uses the transpose matrix:

$$\hat{B} = \hat{T}^{T}\hat{A}\hat{T} \quad [2\text{-}44]$$

This only works however, if:

$$\hat{T}^{T} = \hat{T}^{-1}$$

and, this being the case, all of the properties of the similar matrices discussed above apply here as well.

The last type of transformation that we shall be interested in is the *unitary* transformation:

$$\hat{B} = (\hat{T}^{*})^{T}\hat{A}\hat{T} \quad [2\text{-}45]$$

where:

$$(\hat{\boldsymbol{T}}^*)^T = \hat{\boldsymbol{T}}^{-1}$$

In other words, the matrix $\hat{\boldsymbol{T}}$ is unitary (see [2-14]). If all the elements of $\hat{\boldsymbol{T}}$ are real then it is also orthogonal. Again, there is much similarity (pun intended) between these and the similarity and orthogonal transformations. The determinant, trace and eigenvalues of the transformed matrix are invariant under the transformation.

Very closely related to the unitary transformation is the hermitian transformation:

$$\hat{\boldsymbol{B}} = (\hat{\boldsymbol{T}}^*)^T \hat{\boldsymbol{A}} \hat{\boldsymbol{T}}^{-1}$$

where now:

$$(\hat{\boldsymbol{T}}^*)^T = \hat{\boldsymbol{T}}$$

Note that the inverse matrix is used explicitly in this transformation, unlike the previous ones. As before, we can show that the trace, determinant and eigenvalues are invariant with this transformation.

We have already used a similarity transformation to advantage in equation [2-39]. We can use these ideas in our investigation of eigenvalue equations. In fact, equation [2-39] is referred to as the *diagonalization* of $\hat{\boldsymbol{A}}$ since it transforms $\hat{\boldsymbol{A}}$ into $\hat{\boldsymbol{\Lambda}}$ which is, of course, a diagonal matrix of λ values. The diagonalization problem is essentially the problem of finding the eigenvalues for a given matrix and is therefore an important one in the physical sciences.

2.5 Problems

2.6 References

1. J.S. Golan, *The Linear Algebra a Beginning Graduate Student Ought to Know*, Kluwer Academic Publishers, 2004.
2. M. O'nan, *Linear Algebra,* Harcourt, Brace Jovanovich, inc. 1976.
3. F.L. Pilar, *Elementary Quantum Chemistry*, McGraw-Hill Book Company, 1968.
4. K.A. Stroud and D. Booth, *Linear Algebra*, Industrial Press Inc., 2008.
5. G. Strang, *Linear Algebra and its Applications*, Harcourt Brace Jovanovich, 1988.