MSc: Introductory Mathematics

Matrices

A matrix is an ordered collection of numbers. For example

$$\mathbf{M} = \begin{pmatrix} 4 & -1 & 0 \\ 2 & 2 & -2 \\ 5 & -3 & 1 \end{pmatrix} \ \mathbf{P} = \begin{pmatrix} 2 & 2 & 1 \\ 4 & 6 & 8 \end{pmatrix} \ \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \ \mathbf{C} = \begin{pmatrix} 3 & 1 \\ -2 & 0 \\ 4 & 7 \\ -2 & -3 \end{pmatrix}$$

The size of the matrix is the number of rows $\times$ number of rows. Here, $\mathbf{M}$ is a 3×3 matrix, $\mathbf{P}$ is a 2×3 matrix, $\mathbf{x}$ is a 3×1 matrix (which is more commonly called a column vector) and $\mathbf{C}$ is a 4×2 matrix.

A *square* matrix is a matrix with the same number of rows and columns.

For the following discussion it is convenient to define two general 3×3 matrices

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \qquad \mathbf{B} = \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{pmatrix}$$

The extension to the discussion to different sized matrices should be obvious.

A common notation for matrices is the *subscript* notation. For example

$$\mathbf{A} = a_{ij}$$

where i refers to the row and j refers to the column.

1 Addition

In order for matrices to be added, they must be of the same size. For example, A and B can be added

$$\mathbf{A} + \mathbf{B} = \begin{pmatrix} a_{11} + b_{11} & a_{12} + b_{12} & a_{13} + b_{13} \\ a_{21} + b_{21} & a_{22} + b_{22} & a_{23} + b_{23} \\ a_{31} + b_{31} & a_{32} + b_{32} & a_{33} + b_{33} \end{pmatrix}$$

In subscript notation, this would be written

$$\mathbf{A} + \mathbf{B} = a_{ij} + b_{ij}$$

Addition is commutative, ie. $\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}$.

2 Multiplication

The product of two 3×3 matrices is defined as

$$\mathbf{AB} = \begin{pmatrix} a_{11}b_{11} + a_{12}b_{21} + a_{13}b_{31} & a_{11}b_{12} + a_{12}b_{22} + a_{13}b_{32} & a_{11}b_{13} + a_{12}b_{23} + a_{13}b_{33} \\ a_{21}b_{11} + a_{22}b_{21} + a_{23}b_{31} & a_{21}b_{12} + a_{22}b_{22} + a_{23}b_{32} & a_{21}b_{13} + a_{22}b_{23} + a_{23}b_{33} \\ a_{31}b_{11} + a_{32}b_{21} + a_{33}b_{31} & a_{31}b_{12} + a_{32}b_{22} + a_{33}b_{32} & a_{31}b_{13} + a_{32}b_{23} + a_{33}b_{33} \end{pmatrix}$$

The ij^{th} element is equal to $\sum_{k=1}^{3} a_{ik}b_{kj}$. In the subscript notation for matrices, the *summation* convention is often used. This states that if a subscript letter is repeated then it should be summed over that index. Thus, for two 3×3 matrices

$$\mathbf{AB} = a_{ik}b_{kj}$$

is a very concise way to write matrix multiplication since, by the summation convention

$$a_{ik}b_{kj} = a_{i1}b_{1j} + a_{i2}b_{2j} + a_{i3}b_{3j}$$

From the definition of multiplication, it is clear that it is necessary for the number of columns of the left matrix to be equal to the number of rows of the right matrix. For example, a 2×3 matrix can be multiplied by a 3×4 matrix giving a 2×4 matrix, but a 3×4 matrix cannot be multiplied by a 2×3 matrix.

Note that multiplication is not commutative,

$$\mathbf{AB} \neq \mathbf{BA}$$

Multiplication of more than two square matrices is associative. For example,

$$(\mathbf{AB})\mathbf{A} = \mathbf{A}(\mathbf{BA})$$

Multiplying a square matrix by a column vector yields a column vector. For example, for $\mathbf{A}$ and $\mathbf{x}$ defined above

$$\mathbf{Ax} = \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 \\ a_{31}x_1 + a_{32}x_2 + a_{33}x_3 \end{pmatrix}$$

3 Transpose

The transpose of a matrix, $\mathbf{A} = a_{ij}$ is defined as $\mathbf{A}^T = a_{ji}$. That is

$$\mathbf{A}^T = \begin{pmatrix} a_{11} & a_{21} & a_{31} \\ a_{12} & a_{22} & a_{31} \\ a_{13} & a_{23} & a_{33} \end{pmatrix}$$

Note that $(\mathbf{A}^T)^T = \mathbf{A}$.

The transpose of an $n \times m$ matrix will be an $m \times n$ matrix. For example

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

4 Diagonal Matrices

A diagonal matrix is a square matrix that has non-zero elements only along the diagonal, $a_{ij} = 0$ if $i \neq j$. If $\mathbf{A}$ is a diagonal matrix, then

$$\mathbf{A} = \begin{pmatrix} a_{11} & 0 & 0 \\ 0 & a_{22} & 0 \\ 0 & 0 & a_{33} \end{pmatrix}$$

Note that a 3×3 diagonal matrix has 3 independent elements.

5 Unit Matrix

The *unit* matrix is a diagonal matrix with 1's along the diagonal. The 3×3 unit matrix is

$$\mathbf{I} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Note that $\mathbf{AI} = \mathbf{A}$ and so the unit matrix has a similar role in linear algebra to 1 in algebra. Since we also have $\mathbf{IA} = \mathbf{A}$, multiplication by the unit matrix is commutative.

6 Symmetrical Matrices

A matrix is *symmetrical* if $a_{ij} = a_{ji}$. If A is symmetrical, then

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{12} & a_{22} & a_{23} \\ a_{13} & a_{23} & a_{33} \end{pmatrix}$$

Note that a symmetrical 3×3 matrix has 6 independent elements.

7 Skew Symmetrical Matrices

A matrix is *skew symmetrical* if $a_{ij} = -a_{ji}$. If A is skew symmetrical, then

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ -a_{12} & a_{22} & a_{21} \\ -a_{13} & -a_{21} & a_{33} \end{pmatrix}$$

Note that a skew symmetrical 3×3 matrix also has 6 independent elements.

8 Inverse

The inverse of a square matrix satisfies the relationship

$$\mathbf{A}\mathbf{A}^{-1} = \mathbf{I}$$

Multiplication by the inverse is commutative, *i.e.*

$$\mathbf{A}\mathbf{A}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$$

The inverse is very important in linear algebra because it can be used to obtain the solution of the equation

$$\mathbf{A}\mathbf{x} = \mathbf{y}$$

since multiplying both sides of the equation on the left by $\mathbf{A}^{-1}$ gives

$$\mathbf{A}^{-1}\mathbf{A}\mathbf{x} = \mathbf{I}\mathbf{x} = \mathbf{x} = \mathbf{A}^{-1}\mathbf{y}$$

Thus if we know $\mathbf{A}$ and its inverse $\mathbf{A}^{-1}$ and $\mathbf{y}$ we can solve for the unknown $\mathbf{x}$ by simple multiplication $\mathbf{A}^{-1}\mathbf{y}$. The solution of a set of linear equations therefore reduces to the problem of finding the inverse of the matrix of coefficients.

9 Determinant

The determinant is a scalar and is defined only for square matrices. The determinant of a 2×2 matrix is defined as

$$|\mathbf{A}| = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}$$

For a 3×3 the determinant is

$$\begin{aligned} |\mathbf{A}| &= \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12}\begin{vmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{vmatrix} + a_{31}\begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix} \\ &= a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31} \end{aligned}$$

The 2×2 determinants are called the cofactors and are made up of the elements of the matrix excluding the column and row of the factors. This allows the determinant to be defined iteratively, ie. the 4×4 determinant can be written in terms of the 3×3 cofactors, the 5×5 determinant in terms of the 4×4 cofactors, and so on.

The determinant is important because it can be shown that if $|\mathbf{A}| \neq 0$ then there is a unique solution to the system of equations $\mathbf{Ax} = \mathbf{y}$. If $|\mathbf{A}| = 0$ then there is either no solution or an infinity of solutions.

10 Methods of Solving $\mathbf{Ax} = \mathbf{y}$

10.1 Cramer's rule

If we think of the matrix A as a row vector of column vectors,

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = (\mathbf{a}_1\mathbf{a}_2\mathbf{a}_3) \qquad \text{where} \qquad \mathbf{a}_i = \begin{pmatrix} a_{1i} \\ a_{2i} \\ a_{3i} \end{pmatrix}$$

then we can define the determinants

$$D = |\mathbf{A}| = |\mathbf{a}_1\mathbf{a}_2\mathbf{a}_3|, \quad D_1 = |\mathbf{y}\mathbf{a}_2\mathbf{a}_3|, \quad D_2 = |\mathbf{a}_1\mathbf{y}\mathbf{a}_3|, \quad D_3 = |\mathbf{a}_1\mathbf{a}_2\mathbf{y}|$$

If $D \neq 0$, then the solution of the equation is

$$x_1 = \frac{D_1}{D} \quad x_2 = \frac{D_2}{D} \quad x_3 = \frac{D_3}{D}$$

10.2 Gauss elimination

It is probably simplest to demonstrate Gauss elimination using a simple example. Consider the system of equations

$$\begin{aligned} 2x + y + z &= 1 \\ 4x + y &= -2 \\ -2x + 2y + z &= 7 \end{aligned}$$

This can be written as the matrix equation

$$\begin{pmatrix} 2 & 1 & 1 \\ 4 & 1 & 0 \\ -2 & 2 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 \\ -2 \\ 7 \end{pmatrix}$$

Use the first equation to eliminate the x term from the second and third equations

$$\begin{pmatrix} 2 & 1 & 1 \\ 0 & -1 & -2 \\ 0 & 3 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 \\ -4 \\ 8 \end{pmatrix}$$

Use the second equation to eliminate the y term from the third equation

$$\begin{pmatrix} 2 & 1 & 1 \\ 0 & -1 & -2 \\ 0 & 0 & -4 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 \\ -2 \\ -4 \end{pmatrix}$$

This equation can be solved very straightforwardly. The third equation gives $z = 1$. Substituting this into the second equation gives $y = 2$. Substituting these into the first equation gives $x = -1$.

Thus Gauss elimination uses one equation to eliminate one variable from the other equations. One of these equations is used to eliminate another variable from the rest of the equations. This is continued until the matrix has a triangular form, which can be solved by simple back substitution. In the more sophisticated versions of the Gauss elimination algorithms, the largest element is used for each elimination step. This is called 'pivoting' and greatly decreases the sensitivity of the method to round off errors.

10.3 LU decomposition

Almost all square matrices can be written as the product $\mathbf{A} = \mathbf{LU}$ where $\mathbf{L}$ is a lower triangular matrix and $\mathbf{U}$ is an upper triangular matrix

$$\mathbf{L} = \begin{pmatrix} l_{11} & 0 & 0 \\ l_{21} & l_{22} & 0 \\ l_{31} & l_{32} & l_{33} \end{pmatrix} \qquad \mathbf{L} = \begin{pmatrix} u_{11} & u_{12} & u_{13} \\ 0 & u_{22} & u_{23} \\ 0 & 0 & u_{33} \end{pmatrix}$$

Then the equation $\mathbf{Ax} = \mathbf{y}$, can be written

$$\mathbf{Ax} = (\mathbf{LU}) = \mathbf{L}(\mathbf{Ux}) = \mathbf{Lz} = \mathbf{y} \qquad \text{where} \qquad \mathbf{Ux} = \mathbf{z}$$

Thus the problem reduces to the relatively simple one of solving the two triangular systems $\mathbf{Lz} = \mathbf{y}$ and $\mathbf{Ux} = \mathbf{z}$.

Again, it is easiest to demonstrate the method on a particular example (the same as that solved in the previous section). Start with $\mathbf{A}$ and the identity matrix $\mathbf{I}$

$$\begin{pmatrix} 2 & 1 & 1 \\ 4 & 1 & 0 \\ -2 & 2 & 1 \end{pmatrix} \qquad \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Perform Gauss elimination on the left hand matrix to obtain the $\mathbf{U}$ matrix, and put the multipliers used in the right hand matrix to obtain the $\mathbf{L}$ matrix. In this example, eliminate the 21 element of the left hand matrix by multiplying the first row by 2 and subtracting it from the second, putting the multiplier in the 21 element of the right hand matrix

$$\begin{pmatrix} 2 & 1 & 1 \\ 0 & -1 & -2 \\ -2 & 2 & 1 \end{pmatrix} \qquad \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Eliminate the 31 element of the left hand matrix by multiplying the first row by -1 and subtracting it from the third row, putting the multiplier in the 31 element of the right hand matrix

$$\begin{pmatrix} 2 & 1 & 1 \\ 0 & -1 & -2 \\ 0 & 3 & 2 \end{pmatrix} \qquad \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix}$$

Eliminate the 32 element of the left hand matrix by multiplying the second row by -3 and subtracting it from the third row, putting the multiplier in the 32 element of the right hand matrix

$$\mathbf{U} = \begin{pmatrix} 2 & 1 & 1 \\ 0 & -1 & -2 \\ 0 & 0 & -4 \end{pmatrix} \qquad \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & -3 & 1 \end{pmatrix} = \mathbf{L}$$

Now that we have $\mathbf{L}$ and $\mathbf{U}$, we can solve $\mathbf{Ax} = \mathbf{y}$ for any $\mathbf{y}$. Continuing with the example from the previous section, solve for

$$\mathbf{y} = \begin{pmatrix} 1 \\ -2 \\ 7 \end{pmatrix}$$

First solve $\mathbf{Lz} = \mathbf{y}$ for $\mathbf{z}$, using the triangular form and forward substitution

$$\begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & -3 & 1 \end{pmatrix} \mathbf{z} = \begin{pmatrix} 1 \\ -2 \\ 7 \end{pmatrix} \quad \Rightarrow \quad \mathbf{z} = \begin{pmatrix} 1 \\ -4 \\ -4 \end{pmatrix}$$

Now solve $\mathbf{Ux} = \mathbf{z}$ for $\mathbf{x}$, again using the triangular form and back substitution

$$\begin{pmatrix} 2 & 1 & 1 \\ 0 & -1 & -2 \\ 0 & 0 & -4 \end{pmatrix} \mathbf{x} = \begin{pmatrix} 1 \\ -4 \\ -4 \end{pmatrix} \quad \Rightarrow \quad \mathbf{x} = \begin{pmatrix} -1 \\ 2 \\ 1 \end{pmatrix}$$

This is the solution.

10.4 Inverse from LU decomposition

In order to obtain the general solution to $\mathbf{Ax} = \mathbf{y}$, it is necessary to find the inverse of $\mathbf{A}$, $\mathbf{A}^{-1}$. To do this, it is necessary to solve separately for the three unit vectors $\mathbf{Ax}_1 = \mathbf{u}_1$, $\mathbf{Ax}_2 = \mathbf{u}_2$, $\mathbf{Ax}_3 = \mathbf{u}_3$ where

$$\mathbf{u}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \quad \mathbf{u}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \quad \mathbf{u}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

For $\mathbf{u}_1$

$$\begin{pmatrix} 2 & 1 & 1 \\ 2 & 1 & 0 \\ -1 & -3 & 1 \end{pmatrix} \mathbf{z}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \quad \Rightarrow \quad \mathbf{z}_1 = \begin{pmatrix} 1 \\ -2 \\ -5 \end{pmatrix}$$

$$\begin{pmatrix} 2 & 1 & 1 \\ 0 & -1 & -2 \\ 0 & 0 & -4 \end{pmatrix} \mathbf{x}_1 = \begin{pmatrix} 1 \\ -2 \\ -5 \end{pmatrix} \quad \Rightarrow \quad \mathbf{x}_1 = \frac{1}{8} \begin{pmatrix} 1 \\ -4 \\ 10 \end{pmatrix}$$

For $\mathbf{u}_2$

$$\begin{pmatrix} 2 & 1 & 1 \\ 2 & 1 & 0 \\ -1 & -3 & 1 \end{pmatrix} \mathbf{z}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \quad \Rightarrow \quad \mathbf{z}_2 = \begin{pmatrix} 0 \\ 1 \\ 3 \end{pmatrix}$$

$$\begin{pmatrix} 2 & 1 & 1 \\ 0 & -1 & -2 \\ 0 & 0 & -4 \end{pmatrix} \mathbf{x}_2 = \begin{pmatrix} 0 \\ 1 \\ 3 \end{pmatrix} \quad \Rightarrow \quad \mathbf{x}_2 = \frac{1}{8} \begin{pmatrix} 1 \\ 4 \\ -6 \end{pmatrix}$$

For $\mathbf{u}_3$

$$\begin{pmatrix} 2 & 1 & 1 \\ 2 & 1 & 0 \\ -1 & -3 & 1 \end{pmatrix} \mathbf{z}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \quad \Rightarrow \quad \mathbf{z}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

$$\begin{pmatrix} 2 & 1 & 1 \\ 0 & -1 & -2 \\ 0 & 0 & -4 \end{pmatrix} \mathbf{x}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \quad \Rightarrow \quad \mathbf{x}_3 = \frac{1}{8}\begin{pmatrix} -1 \\ 4 \\ -2 \end{pmatrix}$$

Then the inverse $\mathbf{A}^{-1} = (\mathbf{x}_1\ \mathbf{x}_2\ \mathbf{x}_3)$. That is

$$\mathbf{A}^{-1} = \frac{1}{8}\begin{pmatrix} 1 & 1 & -1 \\ -4 & 4 & 4 \\ 10 & -6 & -2 \end{pmatrix}$$

This solution for $\mathbf{A}^{-1}$ can be checked by multiplication

$$\begin{aligned} \mathbf{A}\mathbf{A}^{-1} &= \frac{1}{8}\begin{pmatrix} 2 & 1 & 1 \\ 4 & 1 & 0 \\ -2 & 2 & 1 \end{pmatrix}\begin{pmatrix} 1 & 1 & -1 \\ -4 & 4 & 4 \\ 10 & -6 & -2 \end{pmatrix} \\ &= \frac{1}{8}\begin{pmatrix} 2-4+10 & 2+4-6 & -2+4-2 \\ 4-4 & 4+4 & -4+4 \\ -2-8+10 & -2+8-6 & 2+8-2 \end{pmatrix} = \mathbf{I} \end{aligned}$$

11 Eigenvalues and Eigenvectors

The equation $\mathbf{Ax} = \mathbf{y}$ can be thought of as a linear transformation from $\mathbf{x} \Rightarrow \mathbf{y}$. There are special vectors $\mathbf{e}$, for which the transformed vector is in the same direction as the original vector. That is,

$$\mathbf{Ae} = \lambda \mathbf{e}$$

where $\mathbf{e}$ is called the eigenvector (characteristic vector) and λ is called the eigenvalue (characteristic value). To determine λ and $\mathbf{e}$, rewrite the equation

$$(\mathbf{A} - \lambda \mathbf{I})\mathbf{e} = 0$$

This equation has a trivial solution $\mathbf{e} = 0$. If the determinant $|\mathbf{A} - \lambda \mathbf{I} \neq 0$ then this solution is unique. However, if $|\mathbf{A} - \lambda \mathbf{I} = 0$, then the solution is not unique and there can be other non-trivial solutions. Thus, the eigenvalues are determined from the equation

$$|\mathbf{A} - \lambda \mathbf{I}| = 0$$

In order to find the eigenvector associated with a give eigenvalue, it is necessary to solve the equation $\mathbf{Ae} = \lambda \mathbf{e}$.

For example, given the matrix

$$\mathbf{A} = \begin{pmatrix} 4 & 1 \\ 2 & 3 \end{pmatrix}$$

we must solve

$$\begin{aligned} |\mathbf{A} - \lambda \mathbf{I}| &= \begin{vmatrix} 4-\lambda & 1 \\ 2 & 3-\lambda \end{vmatrix} \\ &= (4-\lambda)(3-\lambda) - 2 \\ &= \lambda^2 - 7\lambda + 10 \\ &= (\lambda - 5)(\lambda - 2) \end{aligned}$$

Therefore, the eigenvalues for this matrix are $\lambda_1 = 5$ and $\lambda_2 = 2$.

Each eigenvalue will be associated with a different eigenvector. For $\lambda_1 = 5$

$$\begin{pmatrix} 4 & 1 \\ 2 & 3 \end{pmatrix} \begin{pmatrix} e_{1,1} \\ e_{1,2} \end{pmatrix} = 5 \begin{pmatrix} e_{1,1} \\ e_{1,2} \end{pmatrix}$$

or

$$\begin{aligned} 4e_{1,1} + e_{1,2} &= 5e_{1,1} \quad \Rightarrow \quad e_{1,1} = e_{1,2} \\ 2e_{1,1} + 3e_{1,2} &= 5e_{1,1} \quad \Rightarrow \quad 2e_{1,1} = 2e_{1,2} \end{aligned}$$

Note that the two equations give the same result. This degeneracy is the result of setting the determinant equal to zero. Thus, the eigenvector associated with the eigenvalue $\lambda_1 = 5$ is

$$\mathbf{e}_1 = \alpha \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

where α is any constant. For $\lambda_2 = 2$

$$\begin{pmatrix} 4 & 1 \\ 2 & 3 \end{pmatrix} \begin{pmatrix} e_{2,1} \\ e_{2,2} \end{pmatrix} = 2 \begin{pmatrix} e_{2,1} \\ e_{2,2} \end{pmatrix}$$

or

$$\begin{aligned} 4e_{2,1} + e_{2,2} &= 2e_{2,1} \quad \Rightarrow \quad 2e_{2,1} = -e_{2,2} \\ 2e_{2,1} + 3e_{2,2} &= 2e_{2,2} \quad \Rightarrow \quad 2e_{2,1} = -e_{2,2} \end{aligned}$$

Note that the two equations again give the same result. Thus, the eigenvector associated with the eigenvalue $\lambda_2 = 2$ is

$$\mathbf{e}_2 = \beta \begin{pmatrix} 1 \\ -2 \end{pmatrix}$$

where β is any constant.

Note that the constants can be chosen so that the eigenvectors are unit vectors (that is, vectors with unit magnitude). For this example, the unit eigenvectors are

$$\mathbf{e}_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \qquad \mathbf{e}_2 = \frac{1}{\sqrt{5}} \begin{pmatrix} 1 \\ -2 \end{pmatrix}$$

12 Singular Value Decomposition

If a matrix is singular, or near singular (ie, $\mathbf{A} \approx 0$) then the preferred method of solution is called singular value decomposition. This is based upon the theorem from linear algebra that any $n \times m$ matrix $\mathbf{A}$ can be represented as the product of three matrices

$$\mathbf{A} = \mathbf{U}\mathbf{W}\mathbf{V}^T$$

where $\mathbf{U}$ is an $n \times m$ matrix, $\mathbf{W}$ is an $n \times n$ diagonal matrix whose diagonal elements, w_{ii}, are either positive or zero and are called the singular values and $\mathbf{V}$ is an $n \times n$ matrix. Both $\mathbf{U}$ and $\mathbf{V}$ are orthonormal, which means that

$$\mathbf{U}^T\mathbf{U} = \mathbf{V}^T\mathbf{V} = \mathbf{I}$$

This decomposition can be made for any matrix. The singular values are generally arranged in decreasing order. The corresponding rows of the matrix $\mathbf{V}$ can be thought of as the vectors that define a coordinate system spanning the space defined by $\mathbf{A}$ and $\mathbf{U}$ is the transformation of $\mathbf{A}$ into the coordinates given by $\mathbf{V}$.

If $\mathbf{A}$ is a square matrix, then its inverse can be easily found from the singular value decomposition

$$\mathbf{A}^{-1} = \mathbf{V}(diag(1/w_{ii}))\mathbf{U}^T$$

where $(diag(1/w_{ii}))$ is the diagonal matrix made up of the inverses of the singular values w_{ii}.