Chapter 1. Linear Algebra

1.1. Introduction

The solution of algebraic equations pervades science and engineering. Thermodynamics is an area rife with examples because of the ubiquitous presence of equilibrium constraints. Thermodynamic constraints are typically algebraic equations. For example, in the two-phase equilibrium of a binary system, there are three algebraic constraints defining the equilibrium state: thermal equilibrium, or equality of temperatures in the two phases, $T^I = T^{II}$; mechanical equilibrium, or equality of pressures in the two phases, $p^I = p^{II}$; and chemical equilibrium, or equality of chemical potentials of component *i*, in the two phases, $\mu_i^I = \mu_i^{II}$. Under most circumstances, these constraints are algebraic equations because there are no differential or integral operators in the equations.

In the discussion of algebraic equations, it is natural to divide the topic into the solution of linear and nonlinear equations. The mathematical framework for the methodical solution of linear algebraic equations is well-established. There are rigorous techniques for the determination of the existence and uniqueness of solutions. There are established procedures for Eigenanalysis. The discussion of the solution of non-linear algebraic equations is postponed because it is not as straight-forward and requires most of the tools that we develop in our solution of linear algebraic equations.

1.2. Linearity

We have already hinted at the importance of being able to distinguish whether an equation is linear or nonlinear since the solution technique that we adopt is different for linear and nonlinear equations. We begin with a discussion of linear operators. In mathematics, an operator is a symbol or function representing a mathematical operation. Operators that are familiar to undergraduates include exponents, logarithms, differentiation and integration. We can investigate the linearity of each of these operators by applying the following test of linearity

$$L[ax+by] = aL[x] + bL[y] \tag{1.1}$$

where $L[x]$ is a linear operator, operating on the variable x, y is another variable, and a and b are constants.

We can directly check the four operators listed above for linearity. The differential operator, $L[x(t)] = \frac{d}{dt}[x(t)]$, can be substituted into equation (1.1) to yield

$$\frac{d}{dt}[ax(t) + by(t)] = a\frac{d}{dt}[x(t)] + b\frac{d}{dt}[y(t)] \tag{1.2}$$

The differential operator is indeed linear because we know from differential calculus that a constant can be pulled out of the differential and that the differential of a sum is the sum of the differentials. Similarly, the integral operator, $L[x(t')] = \int_{t_o}^{t}[x(t')]dt'$, can be substituted into equation (1.1) to yield

$$\int_{t_o}^{t}[ax(t') + by(t')]dt' = a\int_{t_o}^{t}[x(t')]dt' + b\int_{t_o}^{t}[y(t')]dt' \tag{1.3}$$

The integral operator is indeed linear because we know from integral calculus that a constant can be pulled out of the integral and that the integral of a sum is the sum of the integrals.

The exponential operator, $L[x] = x^n$, can be substituted into equation (1.1) to yield

$$[ax + by]^n \neq ax^n + by^n \tag{1.4}$$

Equation (1.4) is not generally true. It is true for $n = 1$. However, it is not true for any other integer (positive, negative or zero). We can demonstrate directly that equation (1.4) is not true for $n = 2$ through algebraic manipulation.

$$[ax + by]^2 = a^2x^2 + 2abxy + b^2y^2 \neq ax^2 + by^2 \tag{1.5}$$

Equation (1.4) is also not true for fractional exponents, such as $n = 1/2$.

$$[ax + by]^{1/2} = \sqrt{ax + by} \neq a\sqrt{x} + b\sqrt{y} \tag{1.6}$$

Similarly, the logarithm operator, the inverse operator of the exponential operator is not linear.

$$\log_n(ax + by) \neq a\log_n(x) + b\log_n(y) \tag{1.7}$$

Without demonstration, we also state that all trigonometric functions are nonlinear.

In the solution of algebraic equations, the first step is therefore to determine if the equation is linear. An equation is linear if it does not contain any nonlinear operations. If we consider only one equation with one variable, the linear equation has the form,

$$ax = b \tag{1.8.a}$$

which we will choose to rewrite in the more general form of a function as

$$f(x) = ax - b \tag{1.8.b}$$

Examples of nonlinear algebraic equations include

$$f(x) = ax^2 + b \tag{1.9.a}$$

$$f(x) = a\sin(x) + b \tag{1.9.b}$$

$$f(x) = a\exp(x) + x + b \tag{1.9.c}$$

If there is a single nonlinear term in the equation, the entire equation is nonlinear.

In the consideration of systems of equations, if a single equation in the system is nonlinear, then the entire system must be treated as nonlinear.

1.3. Matrix Notation

A system of n algebraic equations containing m unknown variables has the form general form

$$f_i(\underline{x}) = \sum_{j=1}^{m} a_{i,j} x_j - b_i \qquad \text{for } i = 1 \text{ to } n \tag{1.10}$$

For example, a system with n = 2 algebraic equation and m = 2 variables has the form

$$\begin{aligned} f_1(x_1, x_2) &= a_{1,1}x_1 + a_{1,2}x_2 - b_1 \\ f_2(x_1, x_2) &= a_{2,1}x_1 + a_{2,2}x_2 - b_2 \end{aligned} \tag{1.11}$$

It is conventional to adopt a short-hand notation, known as matrix notation, and express equation (1.10) as

$$\underline{\underline{A}}\,\underline{x} = \underline{b} \tag{1.12}$$

where the matrix of constant coefficients, $\underline{\underline{A}}$, is

$$\underline{\underline{A}} = \begin{bmatrix} a_{1,1} & a_{1,1} & \cdots & a_{1,m} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \cdots & a_{n,m} \end{bmatrix} \tag{1.13}$$

The two lines underline in the notation $\underline{\underline{A}}$ indicate that the $\underline{\underline{A}}$ matrix is a two-dimensional matrix. We refer to $\underline{\underline{A}}$ as an *n*x*m* matrix because it contains n rows (equations) and m columns (variables). The vectors $\underline{x}$ and $\underline{b}$ are

$$\underline{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{bmatrix} \quad \text{and} \quad \underline{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} \tag{1.14}$$

The single underline in the notation $\underline{x}$ and $\underline{b}$ indicates that $\underline{x}$ and $\underline{b}$ are vectors, or one-dimensional matrices. We refer to $\underline{x}$ and $\underline{b}$ as "column vectors" of size *m*x1 and *n*x1 respectively.

The crucial thing to remember about matrix notation is that it is a convention to simplify the notation. It does not introduce any new mathematical rules. It does not add or change the rules of algebra. The solution to the set of equation can be obtained following the familiar rules of algebra, although such manipulations become cumbersome when the number of equations is large. Therefore we will shortly adopt a set of notations for matrix operations, which consist of a sequence of algebraic rules.

In equation (1.12), we present the first matrix operation, matrix multiplication, $\underline{\underline{A}}\underline{x}$. Matrices can only be multiplied if the inner indices match. In this case $\underline{\underline{A}}$ is of size *n*x*m* and $\underline{x}$ is of size *m*x1. Since the last index of $\underline{\underline{A}}$ matches the first index of $\underline{x}$, they can be multiplied. If the indices do not match, there is not matrix multiplication. For example, the matrix multiplication $\underline{x}\underline{\underline{A}}$ cannot be performed because $\underline{x}$ is of size *m*x1 and $\underline{\underline{A}}$ is of size *n*x*m* and the inner indices, 1 and *n* are not the same.

The matrix resulting from a valid matrix multiplication is of a size defined by the two outer indices of the factor matrices. Thus $\underline{\underline{A}}\underline{x}$ yields a matrix of size *n*x1, a column vector of length n. The i^{th} element of an *n*x*m* matrix, $\underline{\underline{A}}$, and an *m*x1 column vector, $\underline{x}$, is defined as

$$\sum_{j=1}^{m} a_{i,j} x_j \qquad \text{for } i = 1 \text{ to } n \tag{1.15}$$

Similarly, the i,j^{th} element of an nxm matrix, $\underline{\underline{A}}$, and an mxp column vector, $\underline{\underline{B}}$, is defined as

$$\sum_{k=1}^{m} a_{i,k} b_{k,j} \qquad \text{for } i = 1 \text{ to } n \text{ and for } j = 1 \text{ to } p \tag{1.16}$$

1.4. The Determinant and Inverse

We now consider the solution of equation (1.10) or alternatively equation (1.12). When n=1 equation and m=1 variable, we have

$$a_{1,1} x_1 = b_1 \tag{1.15}$$

This of course has the general solution, $x_1 = b_1 / a_{1,1}$. This simple problem illustrates the issue of existence of a solution. The solution only exists if $a_{1,1} \neq 0$.

We can next consider a set of linear algebraic equations with n=2 equations and m=2 variables, we have

$$\begin{aligned} a_{1,1} x_1 + a_{1,2} x_2 &= b_1 \\ a_{2,1} x_1 + a_{2,2} x_2 &= b_2 \end{aligned} \tag{1.16}$$

Through a series of algebraic manipulations, we can arrive at the solution

$$x_1 = \frac{a_{22} b_1 - a_{12} b_2}{a_{11} a_{22} - a_{21} a_{12}} \qquad \text{and} \qquad x_2 = \frac{-a_{21} b_1 + a_{11} b_2}{a_{11} a_{22} - a_{21} a_{12}} \tag{1.17}$$

Note that both x_1 and x_2 have the same denominator. This denominator is given a special name, the determinant, $\det\left(\underline{\underline{A}}\right)$.

$$\det\left(\underline{\underline{A}}_{2x2}\right) = a_{11} a_{22} - a_{21} a_{12} \tag{1.18}$$

It is clear that a solution only exists if $\det\left(\underline{\underline{A}}\right) \neq 0$. This is exactly parallel to the single equation case given above, where the determinant of the one equation case is simply $\det\left(\underline{\underline{A}}_{1x1}\right) = a_{11}$.

We can next consider a set of linear algebraic equations with n=3 equations and m=3 variables, we have

$$\begin{aligned} a_{1,1}x_1 + a_{1,2}x_2 + a_{1,3}x_3 &= b_1 \\ a_{2,1}x_1 + a_{2,2}x_2 + a_{2,3}x_3 &= b_2 \\ a_{3,1}x_1 + a_{3,2}x_2 + a_{3,3}x_3 &= b_3 \end{aligned} \tag{1.19}$$

Through a series of algebraic manipulations, we can arrive at the solution

$$\begin{aligned} x_1 &= \frac{(a_{22}a_{33} - a_{32}a_{23})b_1 + (-a_{12}a_{33} + a_{32}a_{13})b_2 + (a_{12}a_{23} - a_{22}a_{13})b_3}{\det(\underline{\underline{A}}_3)} \\ x_2 &= \frac{(-a_{21}a_{33} + a_{31}a_{23})b_1 + (a_{11}a_{33} - a_{31}a_{13})b_2 + (-a_{11}a_{23} + a_{21}a_{13})b_3}{\det(\underline{\underline{A}}_3)} \\ x_3 &= \frac{(a_{21}a_{32} - a_{31}a_{22})b_1 + (-a_{11}a_{32} + a_{31}a_{12})b_2 + (a_{11}a_{22} - a_{21}a_{12})b_3}{\det(\underline{\underline{A}}_3)} \end{aligned} \tag{1.20}$$

where

$$\det(\underline{\underline{A}}_{3x3}) = a_{11}(a_{22}a_{33} - a_{32}a_{23}) + a_{12}(a_{23}a_{31} - a_{33}a_{21}) + a_{13}(a_{21}a_{32} - a_{31}a_{22}) \tag{1.21}$$

Note that both x_1, x_2 and x_3 have the same denominator. It is clear that a solution only exists if $\det(\underline{\underline{A}}) \neq 0$.

There is no theoretical reason that we could not continue to solve systems of *n* linear algebraic equations with *n* variables for arbitrary *n*. However, practically speaking it becomes very time consuming. Note in all these cases that the determinant is strictly a function of $\underline{\underline{A}}$. The determinant is not a function of $\underline{x}$ and $\underline{b}$. At this point, we simply extrapolate the mathematical observation that **a unique solution to $\underline{\underline{A}}\underline{x} = \underline{b}$ exists only if the determinant of the matrix $\underline{\underline{A}}$ exists.**

It turns out that the solution to the 2x2 problem given in equation (1.17) and the solution to the 3x3 problem given in equation (1.20) and the solution for the general *n*x*n* problem can be expressed in matrix notation as
Through a series of algebraic manipulations, we can arrive at the solution

$$\underline{x} = \underline{\underline{A}}^{-1}\underline{b} \tag{1.22}$$

where $\underline{\underline{A}}^{-1}$ is called the inverse matrix of $\underline{\underline{A}}$. In addition to providing the solution to $\underline{\underline{A}}\underline{x} = \underline{b}$ as given in equation (1.22), the inverse all has the additional property,

$$\underline{\underline{A}}^{-1}\underline{\underline{A}} = \underline{\underline{A}}\,\underline{\underline{A}}^{-1} = \underline{\underline{I}} \tag{1.23}$$

where $\underline{\underline{I}}$ is the identity matrix, defined as

$$\underline{\underline{I}} \equiv \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix} \tag{1.24}$$

One can also derive equation (1.22) as follows

$$\begin{aligned} &\underline{\underline{A}}\,\underline{x} = \underline{b} \\ &\underline{\underline{A}}^{-1}\underline{\underline{A}}\,\underline{x} = \underline{\underline{A}}^{-1}\underline{b} \\ &\underline{\underline{I}}\,\underline{x} = \underline{\underline{A}}^{-1}\underline{b} \\ &\underline{x} = \underline{\underline{A}}^{-1}\underline{b} \end{aligned}$$

We can observe directly from the examples given above that for the small 1x1, 2x2 and 3x3 systems,

$$\underline{\underline{A}}_{1x1}^{-1} = \frac{1}{\det\left(\underline{\underline{A}}_{1x1}\right)} = \frac{1}{a_{11}} \tag{1.25.a}$$

$$\underline{\underline{A}}_{2x2}^{-1} = \frac{1}{\det\left(\underline{\underline{A}}_{2x2}\right)} \begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix} \tag{1.25.b}$$

$$\underline{\underline{A}}_{3x3}^{-1} = \frac{1}{\det\left(\underline{\underline{A}}_{3x3}\right)} \begin{bmatrix} a_{22}a_{33} - a_{32}a_{23} & -a_{12}a_{33} + a_{32}a_{13} & a_{12}a_{23} - a_{22}a_{13} \\ -a_{21}a_{33} + a_{31}a_{23} & a_{11}a_{33} - a_{31}a_{13} & -a_{11}a_{23} + a_{21}a_{13} \\ a_{21}a_{32} - a_{31}a_{22} & -a_{11}a_{32} + a_{31}a_{12} & a_{11}a_{22} - a_{21}a_{12} \end{bmatrix} \tag{1.25.c}$$

Clearly from equation (1.25), the inverse does not exist if the determinant is zero. If the inverse of a matrix $\underline{\underline{A}}$ exists, $\underline{\underline{A}}$ is said to be non-singular. If the inverse of the matrix $\underline{\underline{A}}$ does not exist, $\underline{\underline{A}}$ is said to be singular.

1.5. Elementary Row Operations

There exists a methodical procedure for generating inverses analytically. With the ubiquitous presence of computers, it is unlikely that any student will ever have any need to perform such a procedure. It is not even perfectly clear that it is essential to include such a procedure in a modern textbook. Nevertheless, since students may be called upon to generate inverses of small systems in an examination in which computers are not available, we present the procedure here.

The procedure uses three elementary row operations. The first elementary row operation is the multiplication of a row by a constant.

$$row1 = \mathrm{c} \cdot \mathrm{row}\,1 \tag{1.26}$$

The second elementary row operation is switching the order of rows. Clearly, the order that the equations are written should not influence the validity of the equations.

$$row\,2 \leftrightarrow \mathrm{row}\,1 \tag{1.27}$$

The third elementary row operation is the replacement of an equation by the linear combination of that equation with other equations. In other words, either equation in (1.28) can be replaced by

$$row\,2 = \mathrm{a} \cdot \mathrm{row}\,1 + \mathrm{b} \cdot \mathrm{row2} \tag{1.28}$$

and the resulting system of equations will still yield the same result.

These three elementary row operations provide the necessary tools to (i) determine the existence and unique of solutions, (ii) determine the inverse of $\underline{\underline{A}}$ if it exists and (iii) provide the solution to $\underline{\underline{A}}\underline{x} = \underline{b}$.

For a 2x2 matrix, the procedure for finding the inverse is given below. First, we create an augmented $\underline{\underline{A|I}}$ matrix. If $\underline{\underline{A}}$ is an *n*x*n* matrix and $\underline{\underline{I}}$ is an *n*x*n* identity matrix, then $\underline{\underline{A|I}}$ is an *n*x(2*n*) matrix. defined as

$$\underline{\underline{A|I}} = \left[\begin{array}{ccc|ccc} a_{1,1} & \cdots & a_{1,n} & 1 & 0 & 0 \\ \vdots & \ddots & \vdots & 0 & \ddots & 0 \\ a_{n,1} & \cdots & a_{n,n} & 0 & 0 & 1 \end{array}\right] \tag{1.29}$$

In general, one performs elementary row operations that convert the $\underline{\underline{A}}$ side of the augmented matrix to $\underline{\underline{I}}$. At the same time, one performs the same elementary row operations to the $\underline{\underline{I}}$ side of the augmented matrix, which converts it to the inverse of $\underline{\underline{A}}$.

We can illustrate the process for a 2x2 matrix.

$$\underline{\underline{A|I}} = \left[\begin{array}{cc|cc} a_{11} & a_{12} & 1 & 0 \\ a_{21} & a_{22} & 0 & 1 \end{array}\right]$$

(1) Put a one in the diagonal element of ROW 1.

$$ROW1 = \frac{ROW1}{a_{11}}$$

$$\left[\begin{array}{cc|cc} 1 & a_{12}/a_{11} & 1/a_{11} & 0 \\ a_{21} & a_{22} & 0 & 1 \end{array}\right]$$

(2) Put zeroes in all the entries of COLUMN 1 except ROW 1.

$$ROW2 = ROW2 - a_{21}ROW1$$

$$\left[\begin{array}{cc|cc} 1 & a_{12}/a_{11} & 1/a_{11} & 0 \\ 0 & a_{22} - a_{21}a_{12}/a_{11} & -a_{21}/a_{11} & 1 \end{array}\right]$$

(3) Put a one in the diagonal element of ROW 2.

$$ROW2 = \frac{ROW2}{a_{22} - a_{21}a_{12}/a_{11}}$$

$$\left[\begin{array}{cc|cc} 1 & a_{12}/a_{11} & 1/a_{11} & 0 \\ 0 & 1 & \dfrac{-a_{21}/a_{11}}{a_{22} - a_{21}a_{12}/a_{11}} & \dfrac{1}{a_{22} - a_{21}a_{12}/a_{11}} \end{array}\right]$$

(4) Put zeroes in all the entries of COLUMN 2 except ROW 2.

$$ROW1 = ROW1 - \frac{a_{12}}{a_{11}}ROW2$$

$$\left[\begin{array}{cc|cc} 1 & 0 & {}^{1}\!/\!{}_{a_{11}} - \frac{a_{12}}{a_{11}}\left(\frac{-{}^{a_{21}}\!/\!{}_{a_{11}}}{a_{22} - {}^{a_{21}a_{12}}\!/\!{}_{a_{11}}}\right) & -\frac{a_{12}}{a_{11}}\left(\frac{1}{a_{22} - {}^{a_{21}a_{12}}\!/\!{}_{a_{11}}}\right) \\ 0 & 1 & \frac{-{}^{a_{21}}\!/\!{}_{a_{11}}}{a_{22} - {}^{a_{21}a_{12}}\!/\!{}_{a_{11}}} & \frac{1}{a_{22} - {}^{a_{21}a_{12}}\!/\!{}_{a_{11}}} \end{array}\right]$$

which can be simplified as:

$$\left[\begin{array}{cc|cc} 1 & 0 & \frac{a_{22}}{a_{11}a_{22} - a_{21}a_{12}} & \frac{-a_{12}}{a_{11}a_{22} - a_{21}a_{12}} \\ 0 & 1 & \frac{-a_{21}}{a_{11}a_{22} - a_{21}a_{12}} & \frac{a_{11}}{a_{11}a_{22} - a_{21}a_{12}} \end{array}\right]$$

Here we have converted the matrix on the left hand side to the identity matrix. The matrix on the right hand side is now the inverse as can be seen through comparison of equation (1.25.b).

We can learn several things about the inverse from this demonstration. The most important thing is that if the determinant is zero, the inverse does not exist (because we divide by the determinant to obtain the inverse.)

Never calculate an inverse until you have first shown that the determinant is not zero.

1.6. Rank and Row Echelon Form

To determine the existence and uniqueness of the solution to $\underline{\underline{A}}\underline{x} = \underline{b}$, we must create an augmented $\underline{\underline{A|b}}$ matrix. If $\underline{\underline{A}}$ is an nxn matrix and $\underline{b}$ is an nx1 column vector, then $\underline{\underline{A|b}}$ is an nx(n+1) matrix. defined as

$$\underline{\underline{A|b}} = \left[\begin{array}{ccc|c} a_{1,1} & \dots & a_{n,1} & b_1 \\ \vdots & \ddots & \vdots & \vdots \\ a_{n,1} & \dots & a_{n,n} & b_n \end{array}\right] \tag{1.30}$$

In order to determine the existence and uniqueness of a solution to $\underline{\underline{A}}\,\underline{x} = \underline{b}$, we need to put the matrix $\underline{\underline{A}}$ and the augmented matrix $\underline{\underline{A|b}}$ into **row echelon form** (ref) (or **reduced row echelon form** (rref)) using a sequence of elementary row operations.

Row echelon form is also called upper triangular form, in which all elements below the diagonal are zero. For an arbitrary 2x2 matrix $\underline{\underline{A}}$, we have

$$\underline{\underline{A}}_{2x2} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

We can put this matrix into row echelon form with one elementary row operation, namely

$$ROW\ 2 = ROW\ 2 - \frac{a_{21}}{a_{11}} ROW\ 1$$

which yields

$$ref\left(\underline{\underline{A}}_{2x2}\right) = \begin{bmatrix} a_{11} & a_{12} \\ 0 & a_{22} - \dfrac{a_{21}}{a_{11}} a_{12} \end{bmatrix}$$

Equation (1.32) is the row echelon form of $\underline{\underline{A}}$. Reduced row echelon form simply requires dividing each row of the row echelon form by the diagonal element of that row,

$$ROW\ 1 = \frac{1}{a_{11}} ROW\ 1$$
$$ROW\ 2 = \frac{1}{a_{22} - \dfrac{a_{21}}{a_{11}} a_{12}} ROW\ 2$$

which yields

$$rref\left(\underline{\underline{A}}_{2x2}\right) = \begin{bmatrix} 1 & a_{12}/a_{11} \\ 0 & 1 \end{bmatrix}$$

Equation (1.33) is the reduced row echelon form of $\underline{\underline{A}}$ because it is in row echelon form and it has ones in the diagonal elements.

Similarly, the augmented $\underline{\underline{A|b}}$ matrix can be put into row echelon form or reduced row echelon form using precisely the same set of elementary row operations. For a 2x2 example, we have for the row echelon form

$$\underline{\underline{A|b}} = \left[\begin{array}{cc|c} a_{1,1} & a_{2,1} & b_1 \\ a_{2,1} & a_{2,2} & b_2 \end{array}\right]$$

$$ROW\ 2 = ROW\ 2 - \frac{a_{21}}{a_{11}} ROW\ 1$$

$$ref\left(\underline{\underline{A|b}}\right) = \left[\begin{array}{cc|c} a_{1,1} & a_{2,1} & b_1 \\ 0 & a_{2,2} - \frac{a_{21}}{a_{11}} a_{12} & b_2 - \frac{a_{21}}{a_{11}} b_1 \end{array}\right]$$

For a 2x2 example, we have for the reduced row echelon form

$$ROW\ 1 = \frac{1}{a_{11}} ROW\ 1$$

$$ROW\ 2 = \frac{1}{a_{22} - \frac{a_{21}}{a_{11}} a_{12}} ROW\ 2$$

$$rref\left(\underline{\underline{A|b}}\right) = \left[\begin{array}{cc|c} 1 & a_{2,1}/a_{1,1} & b_1/a_{1,1} \\ 0 & 1 & \left(b_2 - \frac{a_{21}}{a_{11}} b_1\right) \Big/ \left(a_{2,2} - \frac{a_{21}}{a_{11}} a_{12}\right) \end{array}\right] = \left[\begin{array}{cc|c} 1 & a_{2,1}/a_{1,1} & b_1/a_{1,1} \\ 0 & 1 & \frac{a_{11}b_2 - a_{21}b_1}{\det\left(\underline{\underline{A}}\right)} \end{array}\right]$$

In order to evaluate the existence and uniqueness of a solution, we also require the **rank** of a matrix. The rank of a matrix $\underline{\underline{A}}$ is the number of non-zero rows in a matrix when it is put in row echelon form. The rank of a matrix in row echelon form is the same as the rank of a matrix in reduced row echelon form.

Consider the following upper triangular matrices.

$$\underline{\underline{U}} = \begin{bmatrix} u_{11} & u_{12} & u_{13} \\ 0 & u_{22} & u_{23} \\ 0 & 0 & u_{33} \end{bmatrix} \tag{1.31}$$

The rank of this matrix is 3. The determinant of this matrix is non-zero.

If the determinant of an nxn matrix is zero, then the $rank(\underline{\underline{A}}_n)$ is less than n.

$$\underline{\underline{U}} = \begin{bmatrix} u_{11} & u_{12} & u_{13} \\ 0 & u_{22} & u_{23} \\ 0 & 0 & 0 \end{bmatrix} \tag{1.32}$$

Non-square matrices can also be put in row echelon form. Consider the augmented nx(n+1) matrix of the form:

$$\underline{\underline{U}} = \left[\begin{array}{ccc|c} u_{11} & u_{12} & u_{13} & v_1 \\ 0 & u_{22} & u_{23} & v_2 \\ 0 & 0 & u_{33} & v_3 \end{array}\right] \tag{1.33}$$

The rank of this matrix is still defined as the number of non-zero rows in the row echelon form of the matrix. The rank of the matrix shown above is 3. For augmented matrices, the non-zero element can appear on either matrix. The rank of the following matrix is still three.

$$\underline{\underline{U}} = \left[\begin{array}{ccc|c} u_{11} & u_{12} & u_{13} & v_1 \\ 0 & u_{22} & u_{23} & v_2 \\ 0 & 0 & 0 & v_3 \end{array}\right] \tag{1.34}$$

In an augmented matrix, both sides of a row must be zero for the row to be considered zero. The rank of the following matrix is two.

$$\underline{\underline{U}} = \left[\begin{array}{ccc|c} u_{11} & u_{12} & u_{13} & v_1 \\ 0 & u_{22} & u_{23} & v_2 \\ 0 & 0 & 0 & 0 \end{array}\right] \tag{1.35}$$

The rank provides the number of independent equations in the system. For example, consider the 3x3 example given below. The third equation is a linear combination of the first two equations. This matrix can be put in row echelon form using the following elementary row operations,

$$\underline{\underline{A}} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ ca_{11} + ka_{21} & ca_{12} + ka_{22} & ca_{13} + ka_{23} \end{bmatrix} \tag{1.36}$$

This matrix can be put in row echelon form using the following elementary row operations,

$$ROW\ 3 = ROW\ 3 - cROW\ 1 - kROW\ 2$$

$$ROW\ 2 = ROW\ 2 - \frac{a_{21}}{a_{11}} ROW\ 1$$

which yields

$$\left(\underline{\underline{A}}\right) = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22} - \frac{a_{21}}{a_{11}} a_{12} & a_{23} - \frac{a_{21}}{a_{11}} a_{13} \\ 0 & 0 & 0 \end{bmatrix} \tag{1.37}$$

1.7. Existence and Uniqueness of a Solution

The existence and uniqueness of a solution to $\underline{\underline{A}}\underline{x} = \underline{b}$ can be determined with either the row echelon form or the reduced row echelon form of $\underline{\underline{A}}$ and $\underline{\underline{A|b}}$ as follows. In dealing with linear equations, we only have three choices for the number of solutions. We either have 0, 1, or an infinite number of solutions.

No Solutions:

$rank(\underline{\underline{A}}) < n$ and $rank(\underline{\underline{A}}) < rank(\underline{\underline{A}}|\underline{b})$

One Solution:

$rank(\underline{\underline{A}}) = rank(\underline{\underline{A}}|\underline{b}) = n$

Infinite Solutions:

$rank(\underline{\underline{A}}) = rank(\underline{\underline{A}}|\underline{b}) < n$

When $rank(\underline{\underline{A}}|\underline{b}) > rank(\underline{\underline{A}})$, your system is over-specified. At least one of the equations is linearly dependent in the a matrix but is assigned to an inconsistent value of in the $\underline{b}$ vector. There are no solutions to your problem. When $rank(\underline{\underline{A}}|\underline{b}) = rank(\underline{\underline{A}}) = n$, you have a properly specified system with *n* equations and *n* unknown variables and you have one, unique solution. When $rank(\underline{\underline{A}}|\underline{b}) = rank(\underline{\underline{A}}) < n$, then you have less equations than unknowns. You can pick $n - rank(\underline{\underline{A}})$ unknowns arbitrarily then solve for the rest. Therefore you have an infinite number of solutions. We will work one example of each case below.

Example 1.1. One Solution to $\underline{\underline{A}}\underline{x} = \underline{b}$

Let's find

(a) the determinant of $\underline{\underline{A}}$

(b) the inverse of $\underline{\underline{A}}$

(c) the solution of $\underline{\underline{A}}\underline{x} = \underline{b}_1$

(d) the solution of $\underline{\underline{A}}\underline{x} = \underline{b}_2$

where

$$\underline{\underline{A}} = \begin{bmatrix} 2 & 1 & 3 \\ 1 & 2 & 1 \\ 1 & 1 & 1 \end{bmatrix} \qquad \underline{b}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \qquad \underline{b}_2 = \begin{bmatrix} 2 \\ 0 \\ 2 \end{bmatrix}$$

Solution:

(a) The determinant of $\underline{\underline{A}}$ is (by equation 1.21) $\det(\underline{\underline{A}}) = -1$.

(b) Because the determinant is non-zero, we know there will be an inverse. Let's find it.
STEP ONE. Write down the initial matrix augmented by the identity matrix.

$$\underline{\underline{A|I}} = \left[\begin{array}{ccc|ccc} 2 & 1 & 3 & 1 & 0 & 0 \\ 1 & 2 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 \end{array}\right]$$

STEP TWO. Using elementary row operations, convert $\underline{\underline{A}}$ into an identity matrix.

(1) Put a one in the diagonal element of ROW 1. $ROW1 = \dfrac{ROW1}{a_{11}} = \dfrac{ROW1}{2}$

$$\left[\begin{array}{ccc|ccc} 1 & 1/2 & 3/2 & 1/2 & 0 & 0 \\ 1 & 2 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 \end{array}\right]$$

(2) Put zeroes in all the entries of COLUMN 1 except ROW 1.

$$ROW2 = ROW2 - a_{21}ROW1 = ROW2 - ROW1$$
$$ROW3 = ROW3 - a_{31}ROW1 = ROW3 - ROW1$$

$$\left[\begin{array}{ccc|ccc} 1 & 1/2 & 3/2 & 1/2 & 0 & 0 \\ 0 & 3/2 & -1/2 & -1/2 & 1 & 0 \\ 0 & 1/2 & -1/2 & -1/2 & 0 & 1 \end{array}\right]$$

(3) Put a 1 in the diagonal element of ROW 2. $ROW2 = \dfrac{ROW2}{a_{22}} = \dfrac{ROW2}{3/2}$

$$\left[\begin{array}{ccc|ccc} 1 & 1/2 & 3/2 & 1/2 & 0 & 0 \\ 0 & 1 & -1/3 & -1/3 & 2/3 & 0 \\ 0 & 1/2 & -1/2 & -1/2 & 0 & 1 \end{array}\right]$$

(4) Put zeroes in all the entries of COLUMN 2 except ROW 2.

$$ROW1 = ROW1 - a_{12}ROW2 = ROW1 - 1/2 * ROW2$$
$$ROW3 = ROW3 - a_{32}ROW2 = ROW3 - 1/2 * ROW2$$

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 5/3 & 2/3 & -1/3 & 0 \\ 0 & 1 & -1/3 & -1/3 & 2/3 & 0 \\ 0 & 0 & -1/3 & -1/3 & -1/3 & 1 \end{array}\right]$$

(5) Put a 1 in the diagonal element of ROW 3. $ROW3 = \dfrac{ROW3}{a_{33}} = \dfrac{ROW3}{-1/3}$

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 5/3 & 2/3 & -1/3 & 0 \\ 0 & 1 & -1/3 & -1/3 & 2/3 & 0 \\ 0 & 0 & 1 & 1 & 1 & -3 \end{array}\right]$$

(6) Put zeroes in all the entries of COLUMN 3 except ROW 3.

$$ROW1 = ROW1 - a_{13}ROW3 = ROW1 - 5/3 * ROW3$$
$$ROW2 = ROW2 - a_{23}ROW3 = ROW2 + 1/3 * ROW3$$

$$\left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -1 & -2 & 5 \\ 0 & 1 & 0 & 0 & 1 & -1 \\ 0 & 0 & 1 & 1 & 1 & -3 \end{array}\right]$$

We have the inverse.

$$\underline{\underline{A}}^{-1} = \begin{bmatrix} -1 & -2 & 5 \\ 0 & 1 & -1 \\ 1 & 1 & -3 \end{bmatrix}$$

(c) The solution to $\underline{\underline{A}}\underline{x} = \underline{b}_1$ is $\underline{x} = \underline{\underline{A}}^{-1}\underline{b}_1$.

$$\underline{x} = \begin{bmatrix} -1 & -2 & 5 \\ 0 & 1 & -1 \\ 1 & 1 & -3 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \\ -1 \end{bmatrix}$$

(d) The solution to $\underline{\underline{A}}\underline{x} = \underline{b}_2$ is $\underline{x} = \underline{\underline{A}}^{-1}\underline{b}_2$.

$$\underline{x} = \begin{bmatrix} -1 & -2 & 5 \\ 0 & 1 & -1 \\ 1 & 1 & -3 \end{bmatrix}\begin{bmatrix} 2 \\ 0 \\ 2 \end{bmatrix} = \begin{bmatrix} 8 \\ -2 \\ -4 \end{bmatrix}$$

We see that we only need to calculate the inverse once to solve both $\underline{\underline{A}}\underline{x} = \underline{b}_1$ and $\underline{\underline{A}}\underline{x} = \underline{b}_2$. That's nice because finding the inverse is a lot harder than solving the equation once the inverse is known.

Example 1.2. No Solutions to $\underline{\underline{A}}\underline{x} = \underline{b}$

$$\underline{\underline{A}} = \begin{bmatrix} 2 & 1 & 3 \\ 1 & 2 & 1 \\ 3 & 3 & 4 \end{bmatrix} \qquad \underline{b}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

In this case, when we compute the determinant, we find that $\det(\underline{\underline{A}}) = 0$. The determinant is zero. No inverse exists. To determine if we have no solution or infinite solutions find the ranks of $\underline{\underline{A}}$ and $\underline{\underline{A|b}}$. In row echelon form, A becomes:

$$ref(\underline{\underline{A}}) = \begin{bmatrix} 2 & 1 & 3 \\ 0 & -3 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

By inspection of the row echelon form, $rank(\underline{\underline{A}}) = 2$. In row echelon form, $\underline{\underline{A \mid b}}$ becomes

$$ref\left(\underline{\underline{A|b}}\right) = \left[\begin{array}{ccc|c} 2 & 1 & 3 & 1 \\ 0 & -3 & 1 & -1 \\ 0 & 0 & 0 & -2 \end{array}\right]$$

By inspection of the row echelon form, $rank\left(\underline{\underline{A \mid b}}\right) = 3$. Since $rank\left(\underline{\underline{A \mid b}}\right) > rank\left(\underline{\underline{A}}\right)$, there are no solutions to $\underline{\underline{A}}\underline{x} = \underline{b}$.

Example 1.3. Infinite Solutions to $\underline{\underline{A}}\underline{x} = \underline{b}$

Consider the same matrix, $\underline{\underline{A}}$, as was used in the previous example 1.2. The determinant is zero and the rank is 2. Now consider a different b vector.

$$\underline{b}_2 = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}$$

In row echelon form, $\underline{\underline{A \mid b}}$ becomes:

$$ref\left(\underline{\underline{A}}|\underline{b}\right) = \left[\begin{array}{ccc|c} 2 & 1 & 3 & 1 \\ 0 & -3 & 1 & -1 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

By inspection of the row echelon form, $rank\left(\underline{\underline{A}} \mid \underline{b}\right) = 2$. Since $rank(\underline{\underline{A}}) = rank(\underline{\underline{A}}|\underline{b}) = 2 < n = 3$, there are infinite solutions.

We can find one example of the infinite solutions by following a standard procedure. First, we arbitrarily select $n - rank(\underline{\underline{A}})$ variables. In this case we can select one variable. Let's make $x_3 = 0$. Then substitute that value into the row echelon form of $\underline{\underline{A}}|\underline{b}$ and solve the resulting system of $rank(\underline{\underline{A}})$ equations.

$$ref\left(\underline{\underline{A}}|\underline{b}\right) = \left[\begin{array}{ccc|c} 2 & 1 & 3 & 1 \\ 0 & -3 & 1 & -1 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

When $x_3 = 0$

$$ref\left(\underline{\underline{A}}|\underline{b}\right) = \left[\begin{array}{ccc|c} 2 & 1 & 0 & 1 \\ 0 & -3 & 0 & -1 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

Now solve a new $\underline{\underline{A}}\underline{x} = \underline{b}$ problem where $\underline{\underline{A}}$ and $\underline{b}$ come from the non-zero parts of $ref\left(\underline{\underline{A}}|\underline{b}\right)$.

$$\begin{bmatrix} 2 & 1 \\ 0 & -3 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

This problem will always have an inverse. The solution is given by

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1/3 \\ 1/3 \end{bmatrix}$$

So one example of the infinite solutions is

$$\underline{x} = \begin{bmatrix} 1/3 \\ 1/3 \\ 0 \end{bmatrix}$$

1.8. Eigenanalysis

Eigenanalysis involves the determination of eigenvalues and eigenvectors. It is a part of linear algebra that is extremely important to scientists and engineers in a broad variety of applications. Here we first provide the mathematical framework for obtaining eigenvalues and eigenvectors. Then we provide an example.

For an *n*x*n* square matrix, there are *n* eigenvalues, though they need not all be different. If the determinant of the matrix is non-zero, all of the eigenvalues are non-zero. If the determinant of the matrix is zero, at least one of the eigenvalues is zero.

To calculate the eigenvalues, $\{\lambda\}$, for an *n*x*n* matrix, one begins by subtracting the eigenvalue from all diagonal elements.

$$\underline{\underline{A}} - \lambda \underline{\underline{I}} = \begin{bmatrix} a_{1,1} - \lambda & a_{1,1} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} - \lambda & \dots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \dots & a_{n,n} - \lambda \end{bmatrix} \tag{1.38}$$

Second, the determinant of $\underline{\underline{A}} - \lambda \underline{\underline{I}}$ is set to zero,

$$\det\left(\underline{\underline{A}} - \lambda \underline{\underline{I}}\right) = 0 \tag{1.39}$$

Third you must solve this equation for λ. Equation (1.39) is called the characteristic equation and is a polynomial in λ of order *n*. Thus, this equation has *n* roots. As with any polynomial equation, the roots may be complex. The *n* roots of equation (1.39) are the *n* eigenvalues.

Each eigenvalue has associated with it an eigenvector. Thus, if there are *n*-eigenvectors, there are also *n*-eigenvalues. The ith eigenvector, $\underline{w}_i$, for the matrix $\underline{\underline{A}}$, is obtained by solving

$$\left(\underline{\underline{A}} - \lambda_i \underline{\underline{I}}\right)\underline{w}_i = \underline{0} \tag{1.40}$$

for $\underline{w}_i$. This equation defines the eigenvectors and can be solved *n* times for all *n* eigenvalues to yield *n* eigenvectors.

For a 2x2 matrix, we have

$$\underline{\underline{A}}_{2x2} - \lambda \underline{\underline{I}}_{2x2} = \begin{bmatrix} a_{1,1} - \lambda & a_{1,2} \\ a_{2,1} & a_{2,2} - \lambda \end{bmatrix} \tag{1.41}$$

The determinant of $\underline{\underline{A}} - \lambda \underline{\underline{I}}$ is set to zero,

$$\det\left(\underline{\underline{A}}_{2x2} - \lambda \underline{\underline{I}}_{2x2}\right) = \left(a_{1,1} - \lambda\right)\left(a_{2,2} - \lambda\right) - a_{1,2}a_{2,1} = \lambda^2 - (a_{1,1} + a_{2,2})\lambda + \det\left(\underline{\underline{A}}_{2x2}\right) = 0 \tag{1.42}$$

For a 2x2 matrix, the characteristic equation is a quadratic polynomial. The two roots of the equation are given by the quadratic formula,

$$\lambda = \frac{(a_{1,1} + a_{2,2}) \pm \sqrt{(a_{1,1} + a_{2,2})^2 - 4\det\left(\underline{\underline{A}}_{2x2}\right)}}{2} \tag{1.43}$$

The eigenvectors are given by

$$\left(\underline{\underline{A}}_{2x2} - \lambda_i \underline{\underline{I}}_{2x2}\right)\underline{w}_i = \begin{bmatrix} a_{1,1} - \lambda_i & a_{1,2} \\ a_{2,1} & a_{2,2} - \lambda_i \end{bmatrix}\underline{w}_i = \underline{0} \tag{1.44}$$

This set of linear equations must be solved. The first step in solving a system of linear equation is finding the determinant. Because the eigenvalues are solutions to $\det\left(\underline{\underline{A}} - \lambda \underline{\underline{I}}\right) = 0$ (from equation (1.42)), the determinant of the matrix in equation (1.44) is always zero. Since the b vector in equation (1.44) is the zero vector, the rank of $\left(\underline{\underline{A}} - \lambda \underline{\underline{I}}\right)$ is equal to the rank of the augmented matrix $\left(\underline{\underline{A}} - \lambda \underline{\underline{I}}\right) | \underline{b}$, which is less than *n*,

$$rank\left[\left(\underline{\underline{A}} - \lambda \underline{\underline{I}}\right)\right] = rank\left[\left(\underline{\underline{A}} - \lambda \underline{\underline{I}}\right) | \underline{b}\right] < n \tag{1.45}$$

Consequently, there are always infinite solutions for each eigenvector. Another way to think of this is that eigenvectors provide directions only, but not magnitude.

For a 2x2 matrix, we can randomly set the second element of the eigenvector to an arbitrary variable, *x*. Solving the first equation in equation (1.44) yields the eigenvectors,

$$\underline{w}_i = \begin{bmatrix} -a_{1,2}x \\ \left(a_{1,1} - \lambda_i\right)x \end{bmatrix} \; for\ i = 1\ to\ n,\ for\ abitrary\ x \neq 0 \tag{1.46}$$

Typically, eigenvectors are reported as normalized vectors, where the magnitude of the vector is one. The magnitude of an arbitrary vector, $\underline{x}$, of length *n* is defined as

$$|\underline{x}| = \sqrt{\sum_{i=1}^{n} x_i^2} \tag{1.47}$$

Therefore a normalized vector, $\underline{x}'$, is given by

$$\underline{x}' = \frac{1}{|\underline{x}|}\underline{x} \tag{1.48}$$

By construction, the magnitude of this normalized vector is one. The normalized eigenvectors for the 2x2 example is then

$$\underline{w}'_i = \frac{1}{|\underline{w}_i|}\underline{w}_i = \frac{1}{\sqrt{\left(a_{1,2}{}^2 + (a_{1,1}-\lambda_i)^2\right)x^2}}\begin{bmatrix} -a_{1,2}x \\ (a_{1,1}-\lambda_i)x \end{bmatrix} \; for\; i = 1\, to\, n,\; for\; abitrary\; x \neq 0 \tag{1.49}$$

Even the normalized eigenvectors still have two equivalent expressions, which involves multiplication by -1. A normalized eigenvector multiplied by -1 is still a normalized eigenvector. For a common list of eigenvectors, we adopt the convention that the real component of the first element of each eigenvector should be positive.

Example 1.4.

Consider the 2x2 matrix, $\underline{\underline{A}} = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$. The characteristic equation is given by

$$\det\left(\underline{\underline{A}} - \lambda\underline{\underline{I}}\right) = (1-\lambda)(1-\lambda) - 4 = \lambda^2 - 2\lambda - 3 = 0$$

The eigenvalues are given by $\lambda_1 = -1$ and $\lambda_2 = 3$. From equation (1.46), the unnormalized eigenvectors are for $x = 1$,

$$\underline{w}_1 = \begin{bmatrix} -2\cdot 1 \\ (1--1)\cdot 1 \end{bmatrix} = \begin{bmatrix} -2 \\ 2 \end{bmatrix} \qquad \underline{w}_2 = \begin{bmatrix} -2\cdot 1 \\ (1-3)\cdot 1 \end{bmatrix} = \begin{bmatrix} -2 \\ -2 \end{bmatrix}$$

The normalized eigenvectors are

$$\underline{w}'_1 = \frac{1}{|\underline{w}_1|}\underline{w}_1 = \frac{1}{2\sqrt{2}}\underline{w}_1 = \begin{bmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix} \quad \underline{w}'_2 = \frac{1}{|\underline{w}_2|}\underline{w}_2 = \frac{1}{2\sqrt{2}}\underline{w}_2 = \begin{bmatrix} -\frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{bmatrix}$$

If we follow the convention that the first element should be positive in our eigenvectors, then our normalized eigenvectors are

$$\underline{w}'_1 = \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix} \qquad \underline{w}'_2 = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$$

Example 1.5. Normal Mode Analysis

Students frequently ask for an example of the physical meaning of eigenvalues and eigenvectors. Such a meaning can be directly observed in the application of normal mode analysis. Below we provide a normal mode analysis of a simple one-dimensional model of carbon dioxide in the ideal gas state.

Consider a carbon dioxide molecule modeled as three particles connected by two springs, with the carbon atom in the middle, as shown in Figure 1.1. The positions in the laboratory frame of reference are subscripted with an "0". The positions as deviations from their equilibrium positions relative to the molecule center-of-mass contain only the essential information for this problem and do not have the "0" subscript.

We model the interaction between molecules as Hookian springs. For a Hookian spring, the potential energy, U, is

$$U = \frac{k}{2}(x - x_0)^2$$

and the force, F, is

$$F = -k(x - x_0)$$

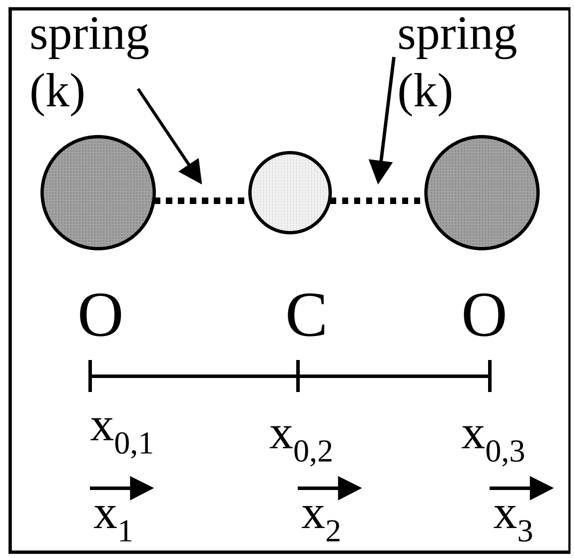

Figure 1.1. An idealized model of the carbon dioxide molecule.

where k is the spring constant (units of kg/s^2). We can write Newton's equations of motion for the three molecules:

$$m_O a_1 = F_1 = k(x_2 - x_1)$$
$$m_C a_2 = F_2 = -k(x_2 - x_1) + k(x_3 - x_2)$$
$$m_O a_3 = F_3 = -k(x_3 - x_2)$$

Knowing that the acceleration is the second derivative of the position, we can rewrite the above equations in matrix form as (first divide both side of all of the equations by the masses)

$$\frac{d^2 \underline{x}}{dt^2} = \underline{\underline{A}}\,\underline{x}$$

where

$$\underline{\underline{A}} = \begin{bmatrix} -{}^{k}\!/_{m_O} & {}^{k}\!/_{m_O} & 0 \\ {}^{k}\!/_{m_C} & -{}^{2k}\!/_{m_C} & {}^{k}\!/_{m_C} \\ 0 & {}^{k}\!/_{m_O} & -{}^{k}\!/_{m_O} \end{bmatrix} \qquad \underline{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

Today, we are not interested in solving this systems of ordinary differential equations. We are content to perform an eigenanalysis on the matrix $\underline{\underline{A}}$. First we have

$$\underline{\underline{A}} - \lambda \underline{\underline{I}} = \begin{bmatrix} -{}^{k}\!/_{m_O} - \lambda & {}^{k}\!/_{m_O} & 0 \\ {}^{k}\!/_{m_C} & -{}^{2k}\!/_{m_C} - \lambda & {}^{k}\!/_{m_C} \\ 0 & {}^{k}\!/_{m_O} & -{}^{k}\!/_{m_O} - \lambda \end{bmatrix}$$

The characteristic equation for this matrix is

$$\begin{aligned} \det\left(\underline{\underline{A}} - \lambda \underline{\underline{I}}\right) &= \left(-{}^{k}\!/_{m_O} - \lambda\right)\left(-{}^{2k}\!/_{m_C} - \lambda\right)\left(-{}^{k}\!/_{m_O} - \lambda\right) \\ &\quad - {}^{k}\!/_{m_O}\,{}^{k}\!/_{m_C}\left(-{}^{k}\!/_{m_O} - \lambda\right) - {}^{k}\!/_{m_O}\,{}^{k}\!/_{m_C}\left(-{}^{k}\!/_{m_O} - \lambda\right) = 0 \end{aligned}$$

Rearranging this cubic polynomial for λ yields

$$\det\left(\underline{\underline{A}} - \lambda \underline{\underline{I}}\right) = -\lambda\left(\lambda + \frac{k}{m_O}\right)\left(\lambda + \frac{k}{m_O}\left(1 + 2\frac{m_O}{m_C}\right)\right) = 0$$

In this form, by inspection the roots to the characteristic equation are

$$\lambda_1 = 0 \qquad \lambda_2 = -\frac{k}{m_O} \qquad \lambda_3 = -\frac{k}{m_O}\left(1 + 2\frac{m_O}{m_C}\right)$$

The eigenvectors for each of these eigenvalues are given by

$$\left(\underline{\underline{A}} - \lambda_i \underline{\underline{I}}\right)\underline{w}_i = \underline{0}$$

$$\left(\underline{\underline{A}} - \lambda_1 \underline{\underline{I}}\right)\underline{w}_1 = \begin{bmatrix} -k/m_O & k/m_O & 0 \\ k/m_C & -2k/m_C & k/m_C \\ 0 & k/m_O & -k/m_O \end{bmatrix} \underline{w}_1 = \underline{0}$$

Since the equations are not linearly independent, we can remove $w_{3,1}$ as a variable and set it equal to 1. Then our system becomes:

$$\begin{bmatrix} -k/m_O & k/m_O \\ k/m_C & -2k/m_C \end{bmatrix} \begin{bmatrix} w_{1,1} \\ w_{2,1} \end{bmatrix} = \begin{bmatrix} 0 \\ -k/m_C \end{bmatrix}$$

This 2x2 matrix has a non-zero determinant. We can solve it uniquely to yield

$$\begin{bmatrix} w_{1,1} \\ w_{2,1} \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

and the eigenvector that corresponds to $\lambda_1 = 0$ is

$$\underline{w}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

To find the second eigenvector, the eigenvector that corresponds to $\lambda_2 = -k/m_O$

$$\underline{\underline{A}} - \lambda_2 \underline{\underline{I}} = \begin{bmatrix} 0 & k/m_O & 0 \\ k/m_C & -2k/m_C + k/m_O & k/m_C \\ 0 & k/m_O & 0 \end{bmatrix}$$

As before, we remove the third equation and remove $w_{3,2}$ as a variable and set it equal to 1. Then our system becomes:

$$\begin{bmatrix} 0 & k/m_O \\ k/m_C & -2k/m_C + k/m_O \end{bmatrix} \begin{bmatrix} w_{1,2} \\ w_{2,2} \end{bmatrix} = \begin{bmatrix} 0 \\ -k/m_C \end{bmatrix}$$

With this new matrix, we can calculate that the determinant is non-zero and the solution is

$$\begin{bmatrix} w_{1,2} \\ w_{2,2} \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \end{bmatrix}$$

and the eigenvector that corresponds to $\lambda_2 = -k/m_O$ is

$$\underline{w}_2 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$$

To find the third eigenvector, the eigenvector that corresponds to $\lambda_3 = -\frac{k}{m_O}\left(1 + 2\frac{m_O}{m_C}\right)$

$$\underline{\underline{A}} - \lambda_3 \underline{\underline{I}} = \begin{bmatrix} 2k/m_C & k/m_O & 0 \\ k/m_C & k/m_O & k/m_C \\ 0 & k/m_O & 2k/m_C \end{bmatrix}$$

As before, we remove the third equation and remove $w_{3,3}$ as a variable and set it equal to 1. Then our system becomes:

$$\begin{bmatrix} 2k/m_C & k/m_O \\ k/m_C & k/m_O \end{bmatrix} \begin{bmatrix} w_{1,3} \\ w_{2,3} \end{bmatrix} = \begin{bmatrix} 0 \\ -k/m_C \end{bmatrix}$$

With this new matrix, we can calculate that the determinant is non-zero and the solution is

$$\begin{bmatrix} w_{1,3} \\ w_{2,3} \end{bmatrix} = \begin{bmatrix} 1 \\ -\dfrac{2m_O}{m_C} \end{bmatrix}$$

and the eigenvector that corresponds to $\lambda_3 = -\dfrac{k}{m_O}\left(1 + 2\dfrac{m_O}{m_C}\right)$ is

$$\underline{w}_3 = \begin{bmatrix} 1 \\ -\dfrac{2m_O}{m_C} \\ 1 \end{bmatrix}$$

So we have the three eigenvalues and the three eigenvectors. So what? What good do they do us? For a vibrating molecule, the square root of the absolute value of the eigenvalues from doing an eigenanalysis of Newton's equations of motion, as we have done, are the normal frequencies. You see that the units of the eigenvalues are 1/sec^2, so the square root has units of frequency (or inverse time).

For carbon dioxide, the three normal frequencies are:

$$\underline{\omega} = \begin{bmatrix} 0 \\ \sqrt{\dfrac{k}{m_O}} \\ \sqrt{\dfrac{k}{m_O}\left(1 + \dfrac{2m_O}{m_C}\right)} \end{bmatrix}$$

The frequency of zero is no frequency at all. It is not a vibrational mode. In fact, it is a translation of the molecule. We can see this by examining the eigenvectors. The eigenvector that corresponds to $\lambda_1 = 0$ or $\omega_1 = \sqrt{|\lambda_1|} = 0$ is

$$\underline{w}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

This is a description of the normal vibration associated with frequency of zero. It says that all atoms move the same amount in the x-direction. See Figure 1.2.

The eigenvector that corresponds to $\lambda_2 = -{}^{k}\!/\!_{m_O}$ or $\omega_2 = \sqrt{|\lambda_2|} = \sqrt{\dfrac{k}{m_O}}$ is

$$\underline{w}_2 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$$

This eigenvector describes a vibration where both the oxygen move away from the carbon equally and the carbon does not move.

The eigenvector that corresponds to $\lambda_3 = -\dfrac{k}{m_O}\left(1 + 2\dfrac{m_O}{m_C}\right)$ or $\omega_3 = \sqrt{|\lambda_3|} = \sqrt{\dfrac{k}{m_O}\left(1 + 2\dfrac{m_O}{m_C}\right)}$ is

$$\underline{w}_2 = \begin{bmatrix} 1 \\ -\dfrac{2m_O}{m_C} \\ 1 \end{bmatrix}$$

This eigenvector describes a vibration where both the oxygen move to the right and the carbon move more to the left, in such a way that there is no center of mass motion.

The normal modes of motion provide a complete, independent set of vibrations from which any other vibration is a linear combination.

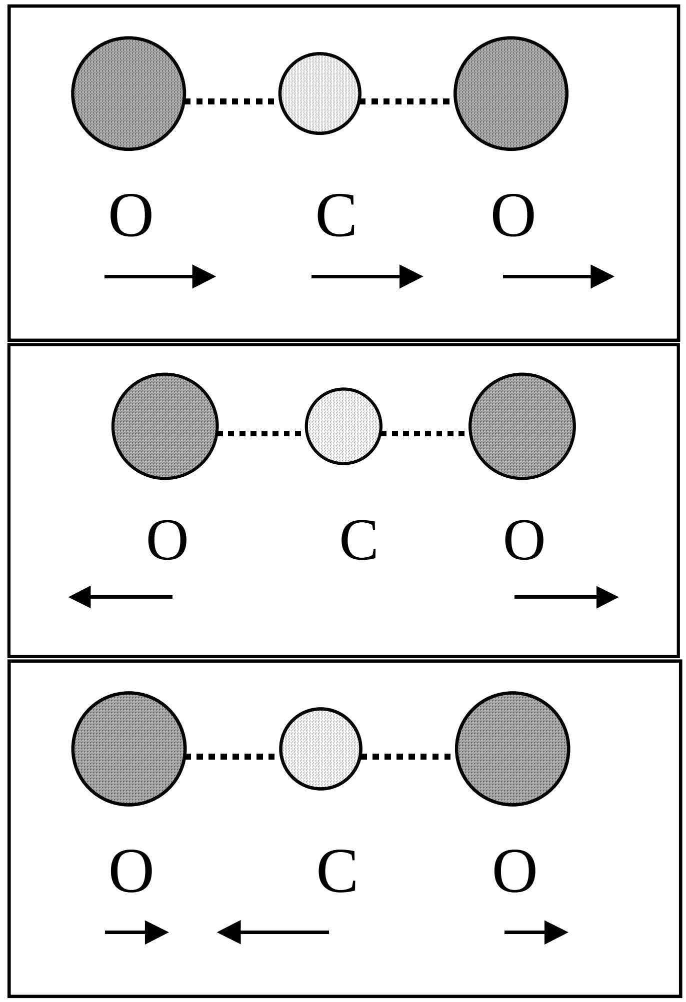

Figure 1.2. Normal modes of a one-dimensional model of carbon dioxide. The top mode is a translational mode with $\omega_1 = 0$. The middle mode is a vibrational mode with $\omega_2 = \sqrt{\dfrac{k}{m_O}}$. The bottom mode is a vibrational mode with $\omega_3 = \sqrt{\dfrac{k}{m_O}\left(1 + 2\dfrac{m_O}{m_C}\right)}$.

1.9. Summary of Logically Equivalent Statements

At this point, we have identified the most common tasks required in the solution of a system of linear algebraic equations. An example of the analytical method by which numerical values can be obtained by hand has been presented. The value in presenting these hand calculations lies in developing an understanding of the general behavior of systems of linear equations. It is unlikely that we will ever be called upon (outside of exams) to calculate eigenvalues or inverses by hand. Nevertheless, knowing what to expect from an analytical understanding better prepares us to make sense of the numerical tools and better understand why numerical tools fail. For example, we can ask a program to compute the inverse of a matrix with a determinant of zero. Depending on the software, a variety of cryptic messages may be provided when the code crashes. Showing that the determinant is zero first, allows us to understand that not all of our equations were independent. Alternatively, some software will simply return some matrix without ever notifying the user that the inverse does not exist. Again, the basic understanding provided above can go a long way in interpreting the results of mathematical software.

To this end, we can identify summarize the logically equivalent statements about an *n*x*n* matrix, $\underline{\underline{A}}$. If any one of these statements is true, all the others are true.

• If and only if $\det(\underline{\underline{A}}) \neq 0$	• If and only if $\det(\underline{\underline{A}}) = 0$
• then inverse exists	• then inverse does not exist
• then $\underline{\underline{A}}$ is non-singular	• then $\underline{\underline{A}}$ is singular
• then $rank(\underline{\underline{A}}) = n$	• then $rank(\underline{\underline{A}}) < n$
• then there are no zero rows in the row echelon form of $\underline{\underline{A}}$	• then there is at least one zero row in the row echelon form of $\underline{\underline{A}}$
• then $\underline{\underline{A}}\underline{x} = \underline{b}$ has one, unique solution	• then $\underline{\underline{A}}\underline{x} = \underline{b}$ has either no solution or infinite solutions
• all eigenvalues of $\underline{\underline{A}}$ are non-zero	• at least one eigenvalue of $\underline{\underline{A}}$ is zero

1.10. Summary of MATLAB Commands

In the table below, a summary of important linear algebra commands in MATLAB is given.

<table>
<tr><td colspan="2">Entering a matrix

<code>A=[a11,12;a21,a22]</code>

(commas separate elements in a row, semicolons separate rows)
(easiest for direct data entry)

<code>A=[a11   a12</code>
<code>a21   a22]</code>

(tabs separate elements in a row, returns separate rows)
(useful for copying data from a table in Word or Excel)</td></tr>
<tr><td>Entering a column vector

<code>b=[b1;b2;b3]</code>

(an nx1 vector)</td><td>Entering a row vector

<code>b=[b1,b2,b3]</code>

(a 1xn vector)</td></tr>
<tr><td>determinant of a matrix

<code>det(A)</code>

(scalar)</td><td>rank of a matrix

<code>rank(A)</code>

(scalar)</td></tr>
<tr><td>inverse of an nxn matrix

<code>inv(A)</code>

(nxn matrix)</td><td>transpose of an nxm matrix or an nx1 vector

<code>A=A'</code>

(mxn matrix or 1xn vector)</td></tr>
<tr><td>solution of Ax=b

<code>x=A\b</code> or <code>x=inv(A)*b</code>

(nx1 vector)</td><td>reduced row echelon form of an nxn matrix

<code>rref(A)</code>

(nxn matrix)</td></tr>
<tr><td colspan="2">eigenvalues and eigenvector of an nxn matrix

<code>[w,lambda]=eig(A)</code>

(w is an nxn matrix where each column is an eigenvector,
lambda is a nxn matrix where each diagonal element is an eigenvalue, off-diagonals are zero).</td></tr>
</table>

1.11. Problems

Problems are located on course website.