2. Matrix Algebra

A linear system of m equations in n unknowns,

$$\begin{gathered} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m \end{gathered}$$

(where the a_{ij} and b_i are constants)

can be written more concisely in matrix form, as

$$\mathrm{A}\,\vec{\mathbf{x}} = \vec{\mathbf{b}}$$

where the ($m \times n$) coefficient matrix [m rows and n columns] is

$$\mathrm{A} = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{bmatrix}$$

and the column vectors (also ($n \times 1$) and ($m \times 1$) matrices respectively) are

$$\vec{\mathbf{x}} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \quad \text{and} \quad \vec{\mathbf{b}} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

Matrix operations can render the solution of a linear system much more efficient.

Sections in this Chapter

2.01 Gaussian Elimination

Example 2.01.1

In quantum mechanics, the Planck length L_P is defined in terms of three fundamental constants:

- the universal constant of gravitation, $G = 6.67 \times 10^{-11}$ N m^2 kg^{-2}
- Planck's constant, $h = 6.62 \times 10^{-34}$ J s
- the speed of light in a vacuum, $c = 2.998 \times 10^8$ m s^{-1}

The Planck length is therefore

$$L_P = k\,G^x h^y c^z$$

where k is a dimensionless constant and x, y, z are constants to be determined.
Also note that 1 N = 1 kg m s^{-2} and 1 J = 1 Nm = 1 kg m^2 s^{-2} .
Use dimensional analysis to find the values of x, y and z .

Let $\left[L_P\right]$ denote the dimensions of L_P.

Then $\left[L_P\right] = \left[k\,G^x h^y c^z\right] = \left[G\right]^x \left[h\right]^y \left[c\right]^z = \left(\text{kg}^{-1}\text{m}^3\text{s}^{-2}\right)^x \left(\text{kg m}^2\text{s}^{-1}\right)^y \left(\text{m s}^{-1}\right)^z$

Example 2.01.1 (continued)

Example 2.01.1 (continued)

Example 2.01.2

Find the solution (x, y, z, t) to the system of equations

$$\begin{aligned} x + y \phantom{{}+z+t} &= 5 \\ y + z \phantom{{}+t} &= 7 \\ 2y + z + t &= 10 \end{aligned}$$

This is an **under-determined system** of equations (fewer equations than unknowns).
A unique solution is not possible. There will be either infinitely many solutions or no solution at all.

Reduce the augmented matrix to reduced row echelon form:

Example 2.01.2 (continued)

The **rank** of a matrix is the number of leading ones in its echelon form.

If rank (A) < rank [A | **b**] , then the linear system is **inconsistent** and has no solution.

If rank (A) = rank [A | **b**] = n (the number of columns in A), then the system has a unique solution for any such vector **b** .

If rank (A) = rank [A | **b**] < n , then the system has infinitely many solutions, with a number of parameters = (n − rank (A)) = (# columns in A_r with no leading one).

Example 2.01.3

Read the solution set $(x_1, x_2, \ldots, x_n)$ from the following reduced echelon forms.

(a)

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

Example 2.01.3

(b)

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

(c)

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & a \\ 0 & 1 & 0 & b \\ 0 & 0 & 1 & c \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

Note that software exists to eliminate the tedious arithmetic of the row operations. Various procedures exist in Maple and Matlab.

A custom program, available on the course web site at "`www.engr.mun.ca/~ggeorge/9420/demos/`", allows the user to enter the coefficients of a linear system as rational numbers, allows the user to perform row operations (but will *not* suggest the appropriate operation to use) and carries out the arithmetic of the chosen row operation automatically.

2.02 Summary of Matrix Algebra

Some rules of matrix algebra are summarized here.

The **dimensions** of a matrix are (# rows $\times$ #columns) [in that order].

Addition and subtraction are defined only for matrices of the same dimensions as each other. The sum of two matrices is found by adding the corresponding entries.

Example 2.02.1

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

Scalar multiplication:
The product cA of matrix A with scalar c is obtained by multiplying every element in the matrix by c.

Example 2.02.2

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

Matrix multiplication:
The product C = AB of a $(p \times q)$ matrix A with an $(r \times s)$ matrix B is defined if and only if $q = r$. The product C has dimensions $(p \times s)$ and entries

$$c_{ij} = \sum_{k=1}^{q} a_{ik} b_{kj}$$

or c_{ij} = (i^{th} row of A) • (j^{th} column of B) [usual Cartesian dot product]

Example 2.02.3

$$\text{AB} = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 3 & 2 \end{bmatrix} \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix} =$$

Note that matrix multiplication is, in general, **not commutative**: BA $\neq$ AB.
In this example, BA is not even defined!

The **transpose** of the $(m \times n)$ matrix $\mathrm{A} = \{ a_{ij} \}$ is the $(n \times m)$ matrix $\mathrm{A}^\mathrm{T} = \{ a_{ji} \}$.
The transpose of the product AB is $(\mathrm{AB})^\mathrm{T} = \mathrm{B}^\mathrm{T}\mathrm{A}^\mathrm{T}$.

Example 2.02.4

$$\mathrm{A} = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 3 & 2 \end{bmatrix}, \quad \mathrm{B} = \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix} \quad \Rightarrow$$

A matrix is **symmetric** if and only if $\mathrm{A}^\mathrm{T} = \mathrm{A}$ (which requires $a_{ji} = a_{ij}$ for all (i, j)).
A matrix is **skew-symmetric** if and only if $\mathrm{A}^\mathrm{T} = -\mathrm{A}$.

A **square matrix** has equal numbers of rows and columns.

If a matrix is symmetric or skew-symmetric, then it must be a square matrix.

If a matrix is skew-symmetric, then it must be a square matrix whose leading diagonal elements are all zero.

Example 2.02.5

$$\mathrm{A} = \begin{bmatrix} 1 & 5 & 0 & -2 \\ 5 & 2 & -1 & 7 \\ 0 & -1 & 3 & 1 \\ -2 & 7 & 1 & 4 \end{bmatrix} \quad \text{is symmetric.}$$

$$\mathrm{B} = \begin{bmatrix} 0 & 5 & 0 & -2 \\ -5 & 0 & -1 & 7 \\ 0 & 1 & 0 & -1 \\ 2 & -7 & 1 & 0 \end{bmatrix} \quad \text{is skew-symmetric.}$$

Any square matrix may be written as the sum of a symmetric matrix and a skew-symmetric matrix.

A square matrix is **upper triangular** if all entries below the leading diagonal are zero.
A square matrix is **lower triangular** if all entries above the leading diagonal are zero.
A square matrix that is both upper and lower triangular is **diagonal**.

Example 2.02.6

$$\mathrm{A} = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 2 & \frac{1}{5} \\ 0 & 0 & 3 \end{bmatrix} \quad \text{is upper triangular.}$$

$$\mathrm{A}^{\mathrm{T}} = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 2 & 0 \\ 0 & \frac{1}{5} & 3 \end{bmatrix} \quad \text{is lower triangular.}$$

$$\mathrm{B} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix} \quad \text{is diagonal.}$$

The **trace** of a diagonal matrix is the sum of its elements. $\Rightarrow$ trace(B) = 6.

The diagonal matrix whose diagonal entries are all one is the **identity matrix** I.
Let I_n represent the $(n \times n)$ identity matrix.
$\mathrm{I}_m \mathrm{A} = \mathrm{A}\,\mathrm{I}_n = \mathrm{A}$ for all $(m \times n)$ matrices A.

If it exists, the **inverse** A^{-1} of a square matrix A is such that

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

If the inverse A^{-1} exists, then A^{-1} is unique and A is **invertible**.
If the inverse A^{-1} does not exist, then A is **singular**.

Important distinctions between matrix algebra and scalar algebra:

$ab = ba$ for all scalars a, b; but
AB = BA is true only for some special choices of matrices A, B.

$ab = 0 \Rightarrow a = 0$ and/or $b = 0$, but
AB = 0 can happen when neither A nor B is the zero matrix.

Example 2.02.7

$$\mathrm{A} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \quad \mathrm{B} = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \quad \Rightarrow \quad \mathrm{AB} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} = \mathrm{O}$$

2.03 Determinants and Inverse Matrices

The determinant of the trivial 1×1 matrix is just its sole entry:
$\det [\, a \,] = a$.

The determinant of a 2×2 matrix A is

$$\det(\mathrm{A}) = |\mathrm{A}| = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$$

For higher order $(n \times n)$ matrices $\mathrm{A} = \{ a_{ij} \}$, the determinant can be evaluated as follows:

The **minor** M_{ij} of element a_{ij} is the determinant of order $(n - 1)$ formed from matrix A by deleting the row and column through the element a_{ij}.

The **cofactor** C_{ij} of element a_{ij} is found from $\mathrm{C}_{ij} = (-1)^{i+j}\,\mathrm{M}_{ij}$

The determinant of A is the sum, along any one row or down any one column, of the product of each element with its cofactor:

$$\det(A) = \sum_{j=1}^{n} a_{ij} C_{ij} \qquad (i = \text{any one of } 1, 2, \ldots, n)$$

or

$$\det(A) = \sum_{i=1}^{n} a_{ij} C_{ij} \qquad (j = \text{any one of } 1, 2, \ldots, n)$$

If one row or column has more zero entries than the others, then one usually chooses to expand along that row or column.

The determinant of a triangular matrix is just the product of its diagonal entries.
$\det(\mathrm{I}) = 1$

<u>Example 2.03.1</u>

Evaluate the vector (cross) product of the vectors $\vec{\mathbf{a}} = \hat{\mathbf{i}} + 2\hat{\mathbf{j}} + 3\hat{\mathbf{k}}$ and $\vec{\mathbf{b}} = 2\hat{\mathbf{i}} + 4\hat{\mathbf{j}} + 3\hat{\mathbf{k}}$.

Expanding along the top row,

$\vec{\mathbf{a}} \times \vec{\mathbf{b}} =$

$$\det(AB) = \det(BA) = \det(A)\det(B)$$

$$\det(A^T) = \det(A)$$

$\det(A) = 0 \Rightarrow$ A is singular.

$$\det(A) \neq 0 \quad \Rightarrow \quad A^{-1} = \frac{\text{adj}(A)}{\det(A)}$$

where adj(A) is the adjoint matrix of A, which is the transpose of the matrix of cofactors of A. For a (2×2) matrix, the formula for the inverse follows quickly:

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \quad \Rightarrow \quad A^{-1} = \frac{1}{ad - bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix} \quad (ad \neq bc)$$

Example 2.03.2

$$A = \begin{bmatrix} 2 & 1 \\ 3 & 4 \end{bmatrix} \quad \Rightarrow \quad A^{-1} =$$

For higher order matrices, this adjoint/determinant method of obtaining the inverse matrix becomes very tedious and time-consuming. A much faster method of finding the inverse involves Gaussian elimination to transform the augmented matrix [A | I] into the augmented matrix in reduced echelon form [I | A^{-1}].

Example 2.03.3

Find the inverse of the matrix $A = \begin{bmatrix} -1 & 1 & 0 \\ 3 & 2 & 1 \\ -2 & -1 & -1 \end{bmatrix}$.

Example 2.03.3 (continued)

Example 2.03.3 (continued)

As a check on the answer,

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

Determinants may be evaluated in a similar manner:

Every row operation that subtracts a multiple of a row from another row produces a matrix whose determinant is the same as the previous matrix.

Every interchange of rows changes the sign of the determinant.

Every multiplication of a row by a constant multiplies the determinant by that constant.

Tracking the operations performed in Example 2.03.3 above (that reduced matrix A to the identity matrix I),

Operations	Net factor to date:
Multiply Row 1 by (–1):	**× (–1)**
From Row 2 subtract (3 × Row 1) and	× (–1)
to Row 3 add (2 × Row 1):	× (–1)
Multiply Row 2 by (1/5):	**× (–1/5)**
To Row 3 add (3 × Row 2):	× (–1/5)
Multiply Row 3 by (–5/2):	**× (+1/2)**
From Row 2 subtract (1/5 × Row 3):	× (+1/2)
To Row 1 add Row 2:	× (+1/2)

Therefore

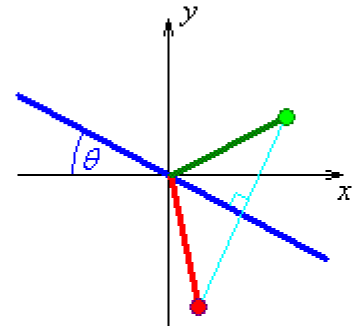
$$\det I = \frac{1}{2}\times\det A \quad \Rightarrow \quad \det A = \begin{vmatrix} -1 & 1 & 0 \\ 3 & 2 & 1 \\ -2 & -1 & -1 \end{vmatrix} = 2(\det I) = 2$$

One can also show that

$$\operatorname{adj}\left(\begin{bmatrix} -1 & 1 & 0 \\ 3 & 2 & 1 \\ -2 & -1 & -1 \end{bmatrix}\right) = \begin{bmatrix} -1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & -3 & -5 \end{bmatrix} \Rightarrow A^{-1} = \frac{\operatorname{adj}(A)}{\det(A)} = \frac{1}{2}\begin{bmatrix} -1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & -3 & -5 \end{bmatrix}$$

2.04 Eigenvalues and Eigenvectors

Example 2.04.1



In $\mathbb{R}^3$, the effect of reflection, in a vertical plane mirror through the origin that makes an angle θ with the *x-z* coordinate plane, on the values of the Cartesian coordinates (*x*, *y*, *z*), may be represented by the matrix equation

$$\vec{\mathbf{x}}_{\text{new}} = \mathrm{R}_\theta \vec{\mathbf{x}}_{\text{old}} \quad \text{or} \quad \begin{bmatrix} x_{\text{new}} \\ y_{\text{new}} \\ z_{\text{new}} \end{bmatrix} = \begin{bmatrix} +\cos 2\theta & -\sin 2\theta & 0 \\ -\sin 2\theta & -\cos 2\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_{\text{old}} \\ y_{\text{old}} \\ z_{\text{old}} \end{bmatrix}$$

The reflection matrix R_θ may be constructed from the composition of three consecutive operations:
rotate all of $\mathbb{R}^3$ about the *z* axis, so that the mirror is rotated into the *x-z* plane; then
reflect the *y* coordinate to its negative; then
rotate all of $\mathbb{R}^3$ about the *z* axis, so that the mirror is rotated back to its starting position.
With the help of some trigonometric identities, one can show that

$$\begin{bmatrix} \cos(-\theta) & -\sin(-\theta) & 0 \\ \sin(-\theta) & \cos(-\theta) & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} +\cos 2\theta & -\sin 2\theta & 0 \\ -\sin 2\theta & -\cos 2\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Obviously, any point on the mirror does not move as a result of the reflection.
Points on the mirror have coordinates (r cos θ, $-r$ sin θ, z), where r and z are any real numbers.
[Note that two free parameters are needed to describe a two-dimensional surface.]

$$\vec{\mathbf{x}} = \begin{bmatrix} r\cos\theta \\ -r\sin\theta \\ z \end{bmatrix} \quad \Rightarrow \quad \mathrm{R}_\theta \vec{\mathbf{x}} = \begin{bmatrix} \cos 2\theta & -\sin 2\theta & 0 \\ -\sin 2\theta & -\cos 2\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r\cos\theta \\ -r\sin\theta \\ z \end{bmatrix}$$

$$= \begin{bmatrix} r(\cos 2\theta\cos\theta + \sin 2\theta\sin\theta) \\ r(-\sin 2\theta\cos\theta + \cos 2\theta\sin\theta) \\ z \end{bmatrix} = \begin{bmatrix} r\cos(2\theta-\theta) \\ -r\sin(2\theta-\theta) \\ z \end{bmatrix} = \begin{bmatrix} r\cos\theta \\ -r\sin\theta \\ z \end{bmatrix} = \vec{\mathbf{x}}$$

Therefore any member of the two dimensional vector space

$$\vec{\mathbf{x}} = \left\{ r \begin{bmatrix} \cos\theta \\ -\sin\theta \\ 0 \end{bmatrix} + z \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\} \qquad (r, z \in \mathbb{R})$$

is invariant under the reflection, $\left(\mathrm{R}_\theta \vec{\mathbf{x}} = \vec{\mathbf{x}} \right)$. The basis vectors of this vector space,

$\begin{bmatrix} \cos\theta \\ -\sin\theta \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$, are the **eigenvectors** of R_θ for the **eigenvalue** +1,

(as is any non-zero combination of them).

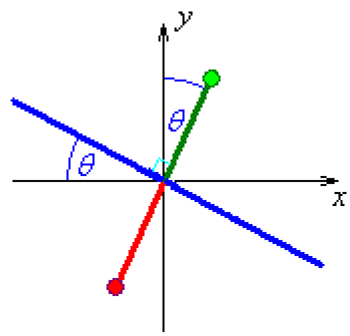
Any point on the line through the origin that is at right angles to the mirror, (r sin θ, r cos θ, 0), will be reflected to –(r sin θ, r cos θ, 0).

For these points, $\mathrm{R}_\theta \vec{\mathbf{x}} = -1\vec{\mathbf{x}}$.

The basis vector of this one-dimensional vector space,

$\begin{bmatrix} \sin\theta \\ \cos\theta \\ 0 \end{bmatrix}$, is the eigenvector of R_θ for the eigenvalue –1,

(as is any non-zero multiple of it).



The zero vector is always a solution of any matrix equation of the form $\mathrm{A}\mathbf{x} = \lambda\, \mathbf{x}$.

$\vec{\mathbf{x}} = \vec{\mathbf{0}}$ is known as the **trivial solution**.

Non-trivial solutions of $\mathrm{A}\,\mathbf{x} = \lambda\,\mathbf{x}$ are possible only for $\lambda = +1$ and for $\lambda = -1$ in this example (with $\mathrm{A} = \mathrm{R}_\theta$).

The eigenvectors for λ = +1 correspond to points on the mirror that map to themselves under the reflection operation R_θ.

The eigenvectors for λ = –1 correspond to points on the normal line that map to their own negatives under the reflection operation R_θ.

No other non-zero vectors will map to simple multiples of themselves under R_θ.

We can summarize the results by displaying the unit eigenvectors as the columns of one matrix and their corresponding eigenvalues as the matching entries in a diagonal matrix:

$$\mathrm{X} = \begin{bmatrix} \sin\theta & \cos\theta & 0 \\ \cos\theta & -\sin\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{and} \quad \Lambda = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Note that the matrix X is **orthogonal**, ($X^{-1} = X^T$ - its inverse is the same as its transpose)
[In this case, X happens to be symmetric also, so that $X^{-1} = X^T = X$.]

Also note that $X^{-1} R_\theta X = \Lambda$:

$$\begin{bmatrix} \sin\theta & \cos\theta & 0 \\ \cos\theta & -\sin\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos 2\theta & -\sin 2\theta & 0 \\ -\sin 2\theta & -\cos 2\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \sin\theta & \cos\theta & 0 \\ \cos\theta & -\sin\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Therefore the matrix X of unit eigenvectors of R_θ diagonalizes the matrix R_θ.

This is generally true of any ($n \times n$) matrix that possesses n linearly independent eigenvectors (some ($n \times n$) matrices do not).

Note that

$$A\vec{\mathbf{x}} = \lambda\vec{\mathbf{x}} \quad \Rightarrow \quad (A - \lambda I)\vec{\mathbf{x}} = \vec{\mathbf{0}}$$

The solution to this square matrix equation will be unique if and only if $\det(A - \lambda I) \neq 0$.
That unique solution is the trivial solution $\vec{\mathbf{x}} = \vec{\mathbf{0}}$.
Therefore eigenvectors can be found if and only if λ is such that $\det(A - \lambda I) = 0$.

General method to find eigenvalues and eigenvectors

$\det(A - \lambda I) = 0$ is the **characteristic equation** from which all of the eigenvalues of the matrix A can be found. For each value of λ, the corresponding eigenvectors are determined by finding the non-trivial solutions to the matrix equation $(A - \lambda I)\vec{\mathbf{x}} = \vec{\mathbf{0}}$.

Example 2.04.2

Find all eigenvalues and unit eigenvectors for the matrix $\mathrm{A} = \begin{bmatrix} -2 & 1 \\ 1 & -2 \end{bmatrix}$.

END OF CHAPTER 2