Linear Algebra

Linear algebra originated in methods to solve systems of linear equations. Consider these three equations in three unknowns, **x, y & z:**

$$2x - y + z = 3$$
$$-x + 3y - z = 2$$
$$x + 2y + z = 8 \quad .$$

Check that the solution is: **x = 1, y = 2, z = 3.**

If one replaces the system with the "augmented matrix" and does these row operations (r1=row#1) to put it in **reduced row echelon form (rref)**:

$$\begin{array}{cccc} x & y & z & : \\ \hline \end{array} \qquad\qquad \begin{array}{cccc} x & y & z & : \\ \hline \end{array}$$

$$\begin{pmatrix} 2 & -1 & 1 & 3 \\ -1 & 3 & -1 & 2 \\ 1 & 2 & 1 & 8 \end{pmatrix} \xrightarrow[\;(^1/_5)r2;\ r1=r1-(2)r2\;]{\;r1=r1+r2;\ \ r2=r2+r3;\ r3=r3-r1;\;} \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 3 \end{pmatrix}$$

It yields the same solution by matching the diagonal **1's** with the last column.

This system has no solution. The third row says: **0x + 0y + 0z = 1**.

$$2x - y + z = 3$$
$$-x + 3y - z = 2$$
$$x + 2y + z = 8$$

$$\begin{pmatrix} 2 & -1 & 1 & 3 \\ -1 & 3 & -1 & 2 \\ 1 & 2 & 0 & 8 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 0 & ^2/_5 & 0 \\ 0 & 1 & ^{-1}/_5 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

This system has infinitely many solutions with $x + (^2/_5)z = 0$ and $y - (^1/_5)z = 0$:

$$2x - y + z = 0$$
$$-x + 3y - z = 0$$
$$x + 2y + z = 0$$

$$\begin{pmatrix} 2 & -1 & 1 & 0 \\ -1 & 3 & -1 & 0 \\ 1 & 2 & 0 & 0 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 0 & ^2/_5 & 0 \\ 0 & 1 & ^{-1}/_5 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

or, for any number **a** , a solution is $x = (^{-2}/_5)a,\ y = (^1/_5)a,\ z = a$.

Vector and Matrix Concepts

Define an **m×n** **matrix** to be an array of numbers with **m** rows and **n** columns.
A **n×1** matrix is a **column vector.** A **1×n** matrix is a **row vector**.
Let all **m×n** matrices be denoted $\boldsymbol{R}^{m,n}$.
Let $\boldsymbol{R}^{n}$ be either all **n×1** or 1×**n.**

As examples:

$\mathbf{A} = (a_{ij})$ **For** 3×4 **A** = $\begin{pmatrix} 2 & 1 & -3 & 5 \\ -4 & 0 & -2 & 3 \\ -1 & -5 & 4 & 0 \end{pmatrix}$ then $a_{12} = 1$, $a_{23} = -2$, $a_{34} = 0$.

$\mathbf{B} = (b_{ij})$ **For** 4×2 **B** = $\begin{pmatrix} 2 & 1 \\ -4 & 0 \\ -1 & -5 \\ 3 & -2 \end{pmatrix}$ then $b_{11} = 2$, $b_{32} = -5$.

$\mathbf{X} = (x_i)$ **For** 4×1 **X** = $\begin{pmatrix} 1 \\ 3 \\ -2 \\ 4 \end{pmatrix}$. $\mathbf{Y} = (y_j)$ **For** 1×4 **Y** = $\begin{pmatrix} 3 & 2 & 1 & 5 \end{pmatrix}$.

The subscripted numbers are called the **<u>components</u>**. The **<u>transpose</u>** of a **m×n** matrix, $\mathbf{A}=(\mathbf{a_{ij}})$, is the **n×m** marix, $\mathbf{A}^{\mathbf{T}}=(\mathbf{a_{ji}})$. The **"m×n"** is the **<u>shape</u>**.

Scalar multiplication multiplies a matrix a real number, σ, $\sigma\mathbf{A} = (\sigma \cdot a_{ij})$.
Matrix addition adds matrices of the same shape: $\mathbf{A} + \mathbf{B} = (a_{ij} + b_{ij})$.
Matrix multiplication multiplies a **m×p** times a **p×n** to yield a **m×n** :

$$\mathbf{A}\bullet\mathbf{B} = \left(\sum_{k=1}^{p} a_{ik} \cdot b_{kj} \right). \quad (\text{Important: } \mathbf{A}\bullet\mathbf{B} \neq \mathbf{B}\bullet\mathbf{A})$$

<u>Laws of Matrices (0 = all zero matrix):</u>

$\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}$; $\mathbf{A} + \mathbf{0} = \mathbf{A}$; $\mathbf{A}\bullet\mathbf{0} = \mathbf{0}\bullet\mathbf{A} = \mathbf{0}$

$(\sigma\mathbf{A})\bullet\mathbf{B} = \sigma(\mathbf{A}\bullet\mathbf{B}) = \mathbf{A}\bullet(\sigma\mathbf{B})$;

$(\mathbf{A}\bullet\mathbf{B})\bullet\mathbf{C} = \mathbf{A}\bullet(\mathbf{B}\bullet\mathbf{C})$ (associative);

$\mathbf{A}\bullet(\mathbf{B} + \mathbf{C}) = \mathbf{A}\bullet\mathbf{B} + \mathbf{A}\bullet\mathbf{C}$ and $(\mathbf{B} + \mathbf{C})\bullet\mathbf{A} = \mathbf{B}\bullet\mathbf{A} + \mathbf{C}\bullet\mathbf{A}$ (distributive).

Using the above:
A + A = 2A = $\begin{pmatrix} 4 & 2 & -6 & 10 \\ -8 & 0 & -4 & 6 \\ -2 & -10 & 8 & 0 \end{pmatrix}$.

Using the above: **A•B** = $\begin{pmatrix} 18 & 7 \\ 3 & 0 \\ 14 & -21 \end{pmatrix}$ where $1,2^{th}$ entry is **7** = 2(1) + 1(0) + (-3)(-5) + 5(-2), across 1^{st} row, down 2^{nd} column.

Using the above: **A•X** = $\begin{pmatrix} 31 \\ 12 \\ -24 \end{pmatrix}$ and **Y•B** = $\begin{pmatrix} -11 & -12 \end{pmatrix}$.

Linear Independence and Bases

Given **k** vectors of shape **n**, $\mathbf{X_1, X_2, \ldots, X_k}$, a **linear combination** is a sum:

$$\sigma_1\mathbf{X_1} + \sigma_2\mathbf{X_2} + \ldots + \sigma_k\mathbf{X_k},$$

for scalars σ_j. All linear combinations form the **Spanning Space, SS{$\mathbf{X_j}$}**.

The **k** vectors are said to be **linearly dependent** if there is a linear combination of them equal to the zero vector with some of the scalars non-zero. If not linear dependent, they are called **linearly independent**. If $\mathbf{k > n}$, then the vectors must be dependent. To check for independence when $\mathbf{k \leq n}$, arrange the vectors as rows in a $\mathbf{k \times n}$ matrix and calculate **rref**. If some row in **rref** is zero, the vectors are dependent. A Spanning Space of **k** independent vectors has **dimension** **k**.

A set of **n** linearly independent vectors of shape **n** is called a **fundamental set** or **basis**. Given a basis, any vector is a **unique** linearly combination of the vectors in the basis. Given basis, $\{\mathbf{X_j}\}$, of vectors and any vector, **Y**, there are **n** unique $\{\sigma_j\}$:

$$\mathbf{Y} = \sigma_1\mathbf{X_1} + \sigma_2\mathbf{X_2} + \ldots + \sigma_n\mathbf{X_n}.$$

The **standard basis** for $\boldsymbol{R}^{\mathbf{m,n}}$ (and $\boldsymbol{R}^{\mathbf{n}}$) are those matrices with a single coordinate equal to **1** and the rest equal to **0**. $\boldsymbol{R}^{\mathbf{n}}$ has dimension **n** and $\boldsymbol{R}^{\mathbf{m,n}}$ has dimension **mn**. For any given **Y**, the $\{\sigma_j\}$ for the standard basis are its coordinates.

To find the $\{\sigma_j\}$ for some given **Y** relative to a non-standard basis $\{\mathbf{X_j}\}$, arrange the basis as column vectors and **Y** as the $\mathbf{n+1}^{\mathbf{st}}$ column in a $\mathbf{n \times n+1}$ augmented matrix. The last column of the **rref** gives the $\{\sigma_j\}$.

Linear Transformations and Matrices

Using matrix multiplication, a $\mathbf{m \times n}$ matrix, **A**, defines a function: $\mathbf{A}: \boldsymbol{R}^{\mathbf{n}} \to \boldsymbol{R}^{\mathbf{m}}$: for **X** a $\mathbf{n \times 1}$ column vector, **Y** as $\mathbf{m \times 1}$ column vector, $\mathbf{X} \to \mathbf{Y} = \mathbf{A(X)} = \mathbf{A \bullet X}$. Such a function is called a **linear transformation**. It is linear in that it satisfies:

$$\mathbf{A}(\alpha\mathbf{X} + \beta\mathbf{Y}) = \mathbf{A}\bullet(\alpha\mathbf{X} + \beta\mathbf{Y}) = \alpha\mathbf{A}\bullet\mathbf{X} + \beta\mathbf{A}\bullet\mathbf{Y} = \alpha\mathbf{A(X)} + \beta\mathbf{A(Y)}, \text{ for scalars } \alpha \text{ \& } \beta.$$

Composition of these matrix functions is the matrix which results from matrix multiplication: $(\mathbf{A} \circ \mathbf{B})(\mathbf{X}) = \mathbf{A}(\mathbf{B}(\mathbf{X})) = \mathbf{A}\bullet(\mathbf{B}\bullet\mathbf{X}) = (\mathbf{A}\bullet\mathbf{B})\bullet\mathbf{X}$.

For matrix **A**: **row space** is the spanning space of its row vectors,
column space is the spanning space of its column vectors,
null space (or **kernel**) is all vectors **X** with $\mathbf{A \bullet X = 0}$.
The row space and null space are subspaces of $\boldsymbol{R}^n$ and the column space of $\boldsymbol{R}^m$.
A basis for row space are the non-zero vectors in **rref** for **A**.
A basis for column space are the non-zero vectors in **rref** for $\mathbf{A}^T$.
A basis for null space may be constructed by solving $\mathbf{A \bullet X = 0}$ using **rref**.

Theorem: Any **m×n** matrix, $\mathbf{A}: \boldsymbol{R}^n \to \boldsymbol{R}^m$, has **domain** = $\boldsymbol{R}^n$ and
(i) The dimensions of the row space and column space are equal (= **rank(A)**);
(ii) The column space is the **range** of $\mathbf{A} \subseteq \boldsymbol{R}^m$;
(iii) For any vector **X** in $\boldsymbol{R}^n$, there are unique $\mathbf{X_r}$ in row space, $\mathbf{X_n}$ in null space with $\mathbf{X = X_n + X_r}$, and thus **n = dim(row space) + dim(null space).**
(iv) **A: row space → column space** is 1-1 and onto (**injective** and **surjective**).

Square Matrices

A square matrix is an **n×n** matrix and defines a function $\mathbf{A}: \boldsymbol{R}^n \to \boldsymbol{R}^n$. The **identity** matrix, **I**, is a square matrix with 1's on the diagonal:

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

For all appropriate **A** & **X**, **I** satisfies:
$\mathbf{I \bullet A = A = A \bullet I}$ & $\mathbf{I \bullet X = X}$ & $\mathbf{Y \bullet I = Y}$.

A square matrix, **A**, is **invertible** (1-1 and onto) if it has an inverse matrix, $\mathbf{A}^{-1}$, with the property: $\mathbf{A^{-1} \bullet A = A \bullet A^{-1} = I}$.
For example:

$$\begin{pmatrix} 2 & -1 & 1 \\ -1 & 3 & -1 \\ 1 & 2 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & {}^3/_5 & -{}^2/_5 \\ 0 & {}^1/_5 & {}^1/_5 \\ -1 & -1 & 1 \end{pmatrix}$$

RREF:

$$\left(\begin{array}{ccc|ccc} 2 & -1 & 1 & 1 & 0 & 0 \\ -1 & 3 & -1 & 0 & 1 & 0 \\ 1 & 2 & 1 & 0 & 0 & 1 \end{array}\right) \longrightarrow \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & {}^3/_5 & -{}^2/_5 \\ 0 & 1 & 0 & 0 & {}^1/_5 & {}^1/_5 \\ 0 & 0 & 1 & -1 & -1 & 1 \end{array}\right)$$

Theorem: A square **n×n** matrix, **A**, has an inverse **iff** (if and only if) A is **bijective** **iff** row vectors are independent **iff** column vectors are independent **iff** A has rank=n **iff** det(**A**) $\neq 0$ (see below for det = determinant) **iff** $\mathbf{A \bullet X = Y}$ has a unique solution for all **Y** **iff** the only solution of $\mathbf{A \bullet X = 0}$ is $\mathbf{X = 0}$.

Inverse Matrices, Determinants and Solving Systems

If one expands the definition of matrix multiplication:

A•X = Y becomes

$$a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = y_1$$
$$\ldots$$
$$a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n = y_n$$

a system of n linear equations. Suppose **Y** is known but **X** is not. To solve this system is to find $\mathbf{A}^{-1}$ since, applying $\mathbf{A}^{-1}$ to both sides we obtain: $\mathbf{X} = \mathbf{A}^{-1} \bullet \mathbf{Y}$.

Alternatively, the unique solution, **X** , is given by **Cramer's Rule**:

$$x_k = \frac{\det(A,k,Y)}{\det(A)}$$; where **(A,k,Y)** is **A** with the **Y** in kth column.

TI-83 and TI-84 Calculators Do Matrix Calculations (2^{nd}-x^{-1} = 2^{nd}-matrix)

To create or edit a matrix **A:**
2^{nd}-matrix (choose Edit, choose **A)**; Enter shape and #'s; 2^{nd}-mode=2^{nd}-quit.

To calculate determinants, **det(A):**
2^{nd}-matrix(choose Math, choose **det**); 2^{nd}-matrix(choose **A**).

To calculate transposes, $\mathbf{A}^{T}$**:**
2^{nd}-matrix(choose Math, choose **T**); 2^{nd}-matrix(choose **A**).

To calculate inverses, $\mathbf{A}^{-1}$**:**
2^{nd}-matrix(choose **A**); **"**x^{-1}**"** .

To calculate reduced row echelon form, **rref(A):**
2^{nd}-matrix(choose Math, choose **rref**); 2^{nd}-matrix(choose **A**).

To calculate **A + B (A i**s n×m & **B i**s n×m):
2^{nd}-matrix(choose **A**)**; "+";** 2^{nd}-matrix(choose **B**).

To calculate **A•B (A i**s n×p & **B i**s p×m);
2^{nd}-matrix(choose **A**)**; "×" ;**2^{nd}-matrix(choose **B**).

To solve for **C** in **A•C = D;**
Create **A**(as n×n) and **D**(as n×k); (usually k = 1)
2^{nd}-matrix(choose **A**)**; "**x^{-1}**"; "×";** 2^{nd}-matrix(choose **D**).

Metric Structure on Vectors (all vectors are column vectors in standard basis)

One may view a vector as a column matrix of numbers, or as the coordinates of a point in **n-dimensional** space, $\boldsymbol{R}^n$. Picturing an **arrow** from the origin to the point, one may view the vector as having **length** and **direction** . These are scalars defined using the **dot or inner or scalar product**. Let:

$$\mathbf{X \cdot Y = X^T \bullet Y} \text{, the inner or dot product, and, } \mathbf{X \cdot X} = \|\mathbf{X}\|^2 .$$

Call $\|\mathbf{X}\|$ the **norm** or **magnitude** or **length** of **X** . A **unit vector**, **U**, has $\|\mathbf{U}\|=1$.

For $\mathbf{X} \neq \mathbf{0}$ any vector, $\mathbf{U} = \dfrac{\mathbf{X}}{\|\mathbf{X}\|}$ is a unit vector in the same direction.

The **distance** between (the endpoints of) two vectors is $\mathbf{d(X,Y)} = \|\mathbf{X - Y}\|$.

The **angle** between two vectors, $\boldsymbol{\theta}$, is given by:

$$\mathbf{X \cdot Y} = \|\mathbf{X}\|\,\|\mathbf{Y}\| \cos(\theta) \quad \textbf{or} \quad \cos(\theta) = \frac{\mathbf{X \cdot Y}}{\|\mathbf{X}\|\,\|\mathbf{Y}\|} .$$

If $\mathbf{X \cdot Y = 0}$, **X&Y** are called **orthogonal** (or perpendicular, since $\cos(\pi/2) = 0$.)

Projection of **X** onto **Y** is: $\quad \mathrm{proj}_{\mathbf{Y}}(\mathbf{X}) = \dfrac{\mathbf{X \cdot Y}}{\|\mathbf{Y}\|^2}\,\mathbf{Y}$

Laws of Inner Products: (i) $\mathbf{X \cdot Y = Y \cdot X}$ (ii) $c(\mathbf{X \cdot Y}) = (c\mathbf{X}) \cdot \mathbf{Y}$
(iii) $\mathbf{X \cdot (Y + Z) = X \cdot Y + X \cdot Z}$

Theorems:
(i) $\|\mathbf{X + Y}\|^2 = \|\mathbf{X}\|^2 + 2\mathbf{X \cdot Y} + \|\mathbf{Y}\|^2$ **(Law of Cosines)**
(ii) $|\mathbf{X \cdot Y}| \leq \|\mathbf{X}\|\,\|\mathbf{Y}\|$ **(Cauchy-Swartz)**
(iii) $\|\mathbf{X + Y}\| \leq \|\mathbf{X}\| + \|\mathbf{Y}\|$ **(Triangle Inequality)**
(iv) if $\mathbf{X \cdot Y = 0}$ then $\|\mathbf{X + Y}\|^2 = \|\mathbf{X}\|^2 + \|\mathbf{Y}\|^2$ **(Pythagorus)**

The dot product is a scalar. In $\boldsymbol{R}^3$, one may also define the **vector or cross or outer product** of two vectors, which is another vector.

$$\mathbf{U \times V} = \begin{pmatrix} u_2v_3 - u_3v_2 \\ u_3v_1 - u_1v_3 \\ u_1v_2 - u_2v_1 \end{pmatrix} = (u_2v_3 - u_3v_2)\mathbf{i} + (u_3v_1 - u_1v_3)\mathbf{j} + (u_1v_2 - u_2v_1)\mathbf{k},$$

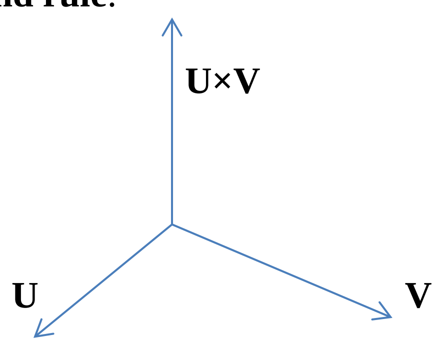
where $\mathbf{i}=(1,0,0)$, $\mathbf{j}=(0,1,0)$, $\mathbf{k}=(0,0,1)$.

Alternatively, view it as a formal determinant: $= \det \begin{pmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{pmatrix}$

Laws of Cross Products: (i) $\mathbf{U}\times\mathbf{V} = -\mathbf{V}\times\mathbf{U}$ (ii) $\mathbf{U}\times\mathbf{U} = \mathbf{0}$ (iii) $\mathbf{U}\times\mathbf{0} = \mathbf{0}$
(iv) $c(\mathbf{U}\times\mathbf{V}) = (c\mathbf{U})\times\mathbf{V}$ (v) $\mathbf{U}\times(\mathbf{V}+\mathbf{W}) = \mathbf{U}\times\mathbf{V} + \mathbf{U}\times\mathbf{W}$
(vi) $\mathbf{U}\cdot(\mathbf{V}\times\mathbf{W}) = (\mathbf{U}\times\mathbf{V})\cdot\mathbf{W}$

Theorems: (i) $\mathbf{U}\times\mathbf{V}$ is orthogonal to both **U** and **V**, and hence to their plane;
(ii) for $\mathbf{U},\mathbf{V}\neq\mathbf{0}$, $\mathbf{U}\times\mathbf{V} = \mathbf{0}$ **iff** **U** and **V** are parallel;
(iii) $||\mathbf{U}\times\mathbf{V}|| = ||\mathbf{U}||\ ||\mathbf{V}||\ \sin(\theta)$, where θ is angle between **U** and **V.**

These may be interpreted as saying that $\mathbf{U}\times\mathbf{V}$ is a vector perpendicular to the plane spanned by **U** and **V** with magnitude equal to the area of the parallelogram they form. There are two choices for the direction of $\mathbf{U}\times\mathbf{V}$, and the convention chosen here is called the **right-hand rule**:



In 2-dimensional space, two orthogonal vectors, **X&Y**, form a rectangle with area $= ||\mathbf{X}||\ ||\mathbf{Y}||$. In 3-dimensional space, three mutually orthogonal vectors, **X&Y&Z** , form a rectangular prism with volume $= ||\mathbf{X}||\ ||\mathbf{Y}||\ ||\mathbf{Z}||$. By analogy, we may define the "hyper-volume" of a "hyper-prism" formed by **n** mutually orthogonal vectors in **n**-dimensional space to be the product of their norms.

For non-orthogonal vectors, the area, volume and hyper-volume of "slanted" parallelograms, parallelepipeds and hyper- parallelepipeds by the usual procedure of cutting and pasting tri-angular pieces, or, by forming a square matrix from the vectors and taking the absolute value of the determinant. For a linear transformation defined by a matrix, **A** , the $|\det(\mathbf{A})|$ is the factor by which A expands the area, volume or hypervolume of a spanning set of vectors. This is the geometric interpretation of the **determinant**. The definition of **det(A)** for 2×2:

$$\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \ ; \quad \det(\mathbf{A}) = ad - bc\ ; \quad \mathbf{A}^{-1} = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \cdot \frac{1}{\det(\mathbf{A})}$$

In n-dimensions, the formula for **det** is messy to define and messier to calculate.

Abstract Vector Spaces and Linear Transformations

Eigenvalues and Eigenvectors

An **n×n** matrix has an **eigenvalue** λ with **eigenvector** **X** if: $\mathbf{A \cdot X = \lambda X}$. This means that the matrix (linear transformation) $\mathbf{A - \lambda I}$ is singular – and hence:

$\mathbf{\det(A - \lambda I) = 0}$ **,** which yields a polynomial of degree **n** in λ .

There will be **n** values of λ which satisfy this requirement. Those values of λ may be real or complex, and must be counted with multiplicity (Fundamental Theorem of Algebra). For a symmetric **A,** $\mathbf{A^T = A}$**,** all eigenvalues are real.

For: $\begin{pmatrix} 3 & -1 \\ 4 & -2 \end{pmatrix}$ $\lambda = 2, -1$ **X**'s are: $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ $\begin{pmatrix} 1 \\ 4 \end{pmatrix}$

For: $\begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$ $\lambda = 2, -1, -1$ **X**'s are: $\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ $\begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$ $\begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}$

For: $\begin{pmatrix} -{}^1/_2 & 1 \\ -1 & -{}^1/_2 \end{pmatrix}$ $\lambda = -{}^1/_2 \pm i$ **X**'s are: $\begin{pmatrix} 1 \\ i \end{pmatrix}$ $\begin{pmatrix} 1 \\ -i \end{pmatrix}$