Module E: Solving Systems of Linear Equations

Definition: A **linear equation** is an equation of the variables x_i of the form

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b.$$

A **solution** for a linear equation is a Euclidean vector

$$\begin{bmatrix} s_1 \\ s_2 \\ \vdots \\ s_n \end{bmatrix}$$

that satisfies

$$a_1s_1 + a_2s_2 + \cdots + a_ns_n = b$$

(that is, a Euclidean vector that can be plugged into the equation).

Definition: A **system of linear equations** (or a **linear system** for short) is a collection of one or more linear equations.

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

Its **solution set** is given by

$$\left\{ \begin{bmatrix} s_1 \\ s_2 \\ \vdots \\ s_n \end{bmatrix} \middle| \begin{bmatrix} s_1 \\ s_2 \\ \vdots \\ s_n \end{bmatrix} \text{ is a solution to all equations in the system} \right\}.$$

Definition: A linear system is **consistent** if its solution set is non-empty (that is, there exists a solution for the system). Otherwise it is **inconsistent**.

Definition: A system of m linear equations with n variables is often represented by writing its coefficients and constants in an **augmented matrix**.

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n &= b_2 \\ \vdots \quad \vdots \quad \vdots &\quad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n &= b_m \end{aligned} \qquad \left[\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{array}\right]$$

Definition: Two systems of linear equations (and their corresponding augmented matrices) are said to be **equivalent** if they have the same solution set.

For example, both of these systems share the same solution set $\left\{\begin{bmatrix}1\\1\end{bmatrix}\right\}$.

$$\begin{aligned} 3x_1 - 2x_2 &= 1 \\ x_1 + 4x_2 &= 5 \end{aligned} \qquad \begin{aligned} 3x_1 - 2x_2 &= 1 \\ 4x_1 + 2x_2 &= 6 \end{aligned}$$

Therefore these augmented matrices are equivalent, which we denote with $\sim$:

$$\left[\begin{array}{cc|c} 3 & -2 & 1 \\ 1 & 4 & 5 \end{array}\right] \sim \left[\begin{array}{cc|c} 3 & -2 & 1 \\ 4 & 2 & 6 \end{array}\right]$$

Definition: The following **row operations** produce equivalent augmented matrices:

1. Swap two rows, for example, $R_1 \leftrightarrow R_2$:

$$\left[\begin{array}{cc|c} 1 & 2 & 3 \\ 4 & 5 & 6 \end{array}\right] \sim \left[\begin{array}{cc|c} 4 & 5 & 6 \\ 1 & 2 & 3 \end{array}\right]$$

2. Multiply a row by a nonzero constant, for example, $2R_1 \to R_1$:

$$\left[\begin{array}{cc|c} 1 & 2 & 3 \\ 4 & 5 & 6 \end{array}\right] \sim \left[\begin{array}{cc|c} 2(1) & 2(2) & 2(3) \\ 4 & 5 & 6 \end{array}\right]$$

3. Add a constant multiple of one row to another row, for example, $R_2 - 4R_1 \to R_2$:

$$\left[\begin{array}{cc|c} 1 & 2 & 3 \\ 4 & 5 & 6 \end{array}\right] \sim \left[\begin{array}{cc|c} 1 & 2 & 3 \\ 4-4(1) & 5-4(2) & 6-4(3) \end{array}\right]$$

Whenever two matrices A, B are equivalent (so whenever we do any of these operations), we write $A \sim B$.

Definition: A matrix is in **reduced row echelon form** (**RREF**) if

1. The leading term (first nonzero term) of each nonzero row is a 1. Call these terms **pivots**.
2. Each pivot is to the right of every higher pivot.
3. Each term above or below a pivot is zero.
4. All rows of zeroes are at the bottom of the matrix.

Every matrix has a unique reduced row echelon form. If A is a matrix, we write $\mathrm{RREF}(A)$ for the reduced row echelon form of that matrix.

Definition: Recall that the pivots of a matrix in RREF form are the leading 1s in each non-zero row.

The pivot columns in an augmented matrix correspond to the **bound variables** in the system of equations (x_1, x_3 below). The remaining variables are called **free variables** (x_2 below).

$$\left[\begin{array}{ccc|c} 1 & 2 & 0 & 4 \\ 0 & 0 & 1 & -1 \end{array}\right]$$

To efficiently solve a system in RREF form, assign letters to the free variables, and then solve for the bound variables.

Module V: Vector Spaces

Definition: A **vector space** V is any collection of mathematical objects with associated addition $\oplus$ and scalar multiplication $\odot$ operations that satisfy the following properties. Let $\vec{\mathbf{u}}, \vec{\mathbf{v}}, \vec{\mathbf{w}}$ belong to V, and let a, b be scalar numbers.

- **Addition is associative:** $\vec{\mathbf{u}} \oplus (\vec{\mathbf{v}} \oplus \vec{\mathbf{w}}) = (\vec{\mathbf{u}} \oplus \vec{\mathbf{v}}) \oplus \vec{\mathbf{w}}$.
- **Addition is commutative:** $\vec{\mathbf{u}} \oplus \vec{\mathbf{v}} = \vec{\mathbf{v}} \oplus \vec{\mathbf{u}}$.
- **Additive identity exists:** There exists some $\vec{\mathbf{z}}$ where $\vec{\mathbf{v}} \oplus \vec{\mathbf{z}} = \vec{\mathbf{v}}$.
- **Additive inverses exist:** There exists some $-\vec{\mathbf{v}}$ where $\vec{\mathbf{v}} \oplus (-\vec{\mathbf{v}}) = \vec{\mathbf{z}}$.
- **Scalar multiplication is associative:** $a \odot (b \odot \vec{\mathbf{v}}) = (ab) \odot \vec{\mathbf{v}}$.
- **Scalar multiplication identity exists:** $1 \odot \vec{\mathbf{v}} = \vec{\mathbf{v}}$.
- **Scalar mult. distributes over vector addition:** $a \odot (\vec{\mathbf{u}} \oplus \vec{\mathbf{v}}) = a \odot \vec{\mathbf{u}} \oplus a \odot \vec{\mathbf{v}}$.
- **Scalar mult. distributes over scalar addition:** $(a+b) \odot \vec{\mathbf{v}} = a \odot \vec{\mathbf{v}} \oplus b \odot \vec{\mathbf{v}}$.

Definition: A **linear combination** of a set of vectors $\{\vec{\mathbf{v}}_1, \vec{\mathbf{v}}_2, \ldots, \vec{\mathbf{v}}_m\}$ is given by $c_1\vec{\mathbf{v}}_1 + c_2\vec{\mathbf{v}}_2 + \cdots + c_m\vec{\mathbf{v}}_m$ for any choice of scalar multiples $c_1, c_2, \ldots, c_m$.

For example, we can say $\begin{bmatrix} 3 \\ 0 \\ 5 \end{bmatrix}$ is a linear combination of the vectors $\begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$ since

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

Definition: The **span** of a set of vectors is the collection of all linear combinations of that set:

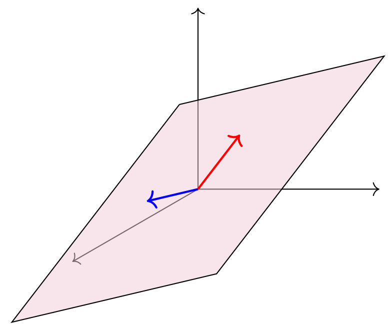
$$\operatorname{span}\{\vec{\mathbf{v}}_1, \vec{\mathbf{v}}_2, \ldots, \vec{\mathbf{v}}_m\} = \{c_1\vec{\mathbf{v}}_1 + c_2\vec{\mathbf{v}}_2 + \cdots + c_m\vec{\mathbf{v}}_m \,|\, c_i \in \mathbb{R}\}.$$

For example:

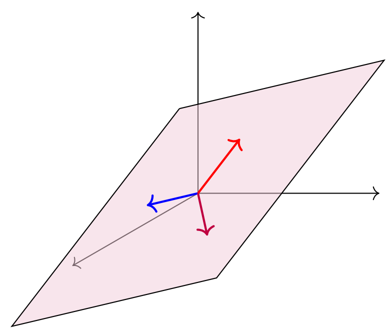
$$\operatorname{span}\left\{ \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} \right\} = \left\{ a \begin{bmatrix} 1 \\ -1 \\ 2 \end{bmatrix} + b \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} \,\middle|\, a, b \in \mathbb{R} \right\}$$

Definition: A subset of a vector space is called a **subspace** if it is a vector space on its own.

For example, the span of these two vectors forms a planar subspace inside of the larger vector space $\mathbb{R}^3$.



Definition: We say that a set of vectors is **linearly dependent** if one vector in the set belongs to the span of the others. Otherwise, we say the set is **linearly independent**.



You can think of linearly dependent sets as containing a redundant vector, in the sense that you can drop a vector out without reducing the span of the set. In the above image, all three vectors lay on the same planar subspace, but only two vectors are needed to span the plane, so the set is linearly dependent.

Definition: A **basis** is a linearly independent set that spans a vector space.

The **standard basis** of $\mathbb{R}^n$ is the set $\{\vec{\mathbf{e}}_1, \ldots, \vec{\mathbf{e}}_n\}$ where

$$\vec{\mathbf{e}}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix} \qquad \vec{\mathbf{e}}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix} \qquad \cdots \qquad \vec{\mathbf{e}}_n = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}$$

For $\mathbb{R}^3$, these are the vectors $\vec{\mathbf{e}}_1 = \hat{\imath} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \vec{\mathbf{e}}_2 = \hat{\jmath} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$, and $\vec{\mathbf{e}}_3 = \hat{k} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$.

Definition: The **dimension** of a vector space is equal to the size of any basis for the vector space.

As you'd expect, $\mathbb{R}^n$ has dimension n. For example, $\mathbb{R}^3$ has dimension 3 because any basis for $\mathbb{R}^3$ such as

$$\{\vec{\mathbf{e}}_1, \vec{\mathbf{e}}_2, \vec{\mathbf{e}}_3\} \text{ and } \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\} \text{ and } \left\{ \begin{bmatrix} 1 \\ 0 \\ -3 \end{bmatrix}, \begin{bmatrix} 2 \\ -2 \\ 1 \end{bmatrix}, \begin{bmatrix} 3 \\ -2 \\ 5 \end{bmatrix} \right\}$$

contains exactly three vectors.

Definition: A **homogeneous** system of linear equations is one of the form:

$$\begin{aligned} a_{11}x_1 + \ a_{12}x_2 + \ldots + \ a_{1n}x_n &= 0 \\ a_{21}x_1 + \ a_{22}x_2 + \ldots + \ a_{2n}x_n &= 0 \\ \vdots \qquad\quad \vdots \qquad\qquad\quad \vdots \quad &\ \ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n &= 0 \end{aligned}$$

This system is equivalent to the vector equation:

$$x_1\vec{\mathbf{v}}_1 + \cdots + x_n\vec{\mathbf{v}}_n = \vec{\mathbf{0}}$$

and the augmented matrix:

$$\left[\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & 0 \\ a_{21} & a_{22} & \cdots & a_{2n} & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & 0 \end{array}\right]$$

Module A: Algebraic Properties of Linear Maps

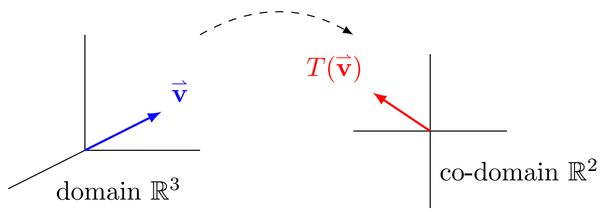
Definition: A **linear transformation** (also known as a **linear map**) is a map between vector spaces that preserves the vector space operations. More precisely, if V and W are vector spaces, a map $T : V \to W$ is called a linear transformation if

1. $T(\vec{\mathbf{v}} + \vec{\mathbf{w}}) = T(\vec{\mathbf{v}}) + T(\vec{\mathbf{w}})$ for any $\vec{\mathbf{v}}, \vec{\mathbf{w}} \in V$.
2. $T(c\vec{\mathbf{v}}) = cT(\vec{\mathbf{v}})$ for any $c \in \mathbb{R}, \vec{\mathbf{v}} \in V$.

In other words, a map is linear when vector space operations can be applied before or after the transformation without affecting the result.

Definition: Given a linear transformation $T : V \to W$, V is called the **domain** of T and W is called the **co-domain** of T.

Linear transformation $T : \mathbb{R}^3 \to \mathbb{R}^2$



Definition: Since linear transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ is determined by the standard basis $\{\vec{\mathbf{e}}_1, \dots, \vec{\mathbf{e}}_n\}$, it's convenient to store this information in the $m \times n$ **standard matrix** $[T(\vec{\mathbf{e}}_1) \cdots T(\vec{\mathbf{e}}_n)]$.

For example, let $T : \mathbb{R}^3 \to \mathbb{R}^2$ be the linear map determined by the following values for T applied to the standard basis of $\mathbb{R}^3$.

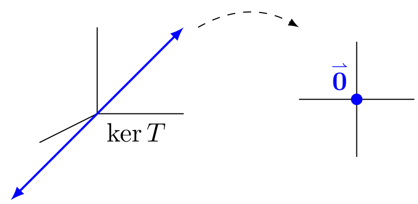
$$T(\vec{\mathbf{e}}_1) = T\left(\begin{bmatrix}1\\0\\0\end{bmatrix}\right) = \begin{bmatrix}3\\2\end{bmatrix} \qquad T(\vec{\mathbf{e}}_2) = T\left(\begin{bmatrix}0\\1\\0\end{bmatrix}\right) = \begin{bmatrix}-1\\4\end{bmatrix} \qquad T(\vec{\mathbf{e}}_3) = T\left(\begin{bmatrix}0\\0\\1\end{bmatrix}\right) = \begin{bmatrix}5\\0\end{bmatrix}$$

Then the standard matrix corresponding to T is

$$\begin{bmatrix}T(\vec{\mathbf{e}}_1) & T(\vec{\mathbf{e}}_2) & T(\vec{\mathbf{e}}_3)\end{bmatrix} = \begin{bmatrix}3 & -1 & 5\\ 2 & 4 & 0\end{bmatrix}.$$

Definition: Let $T : V \to W$ be a linear transformation. The **kernel** of T is an important subspace of V defined by

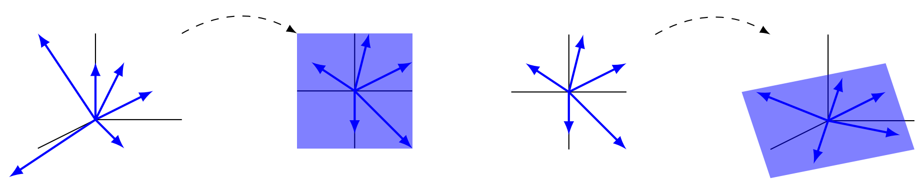
$$\ker T = \left\{ \vec{\mathbf{v}} \in V \;\middle|\; T(\vec{\mathbf{v}}) = \vec{\mathbf{z}} \right\}$$



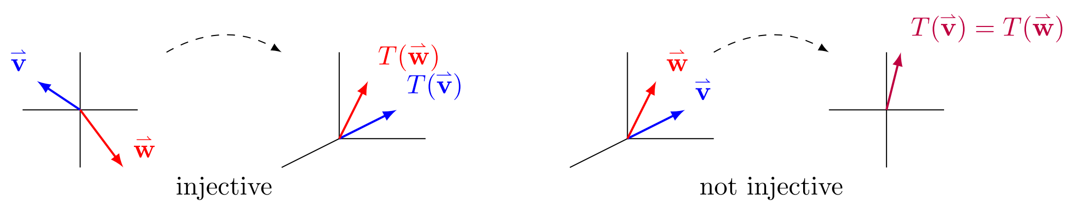
Definition: Let $T : V \to W$ be a linear transformation. The **image** of T is an important subspace of W defined by

$$\operatorname{Im} T = \left\{ \vec{\mathbf{w}} \in W \;\middle|\; \text{there is some } \vec{\mathbf{v}} \in V \text{ with } T(\vec{\mathbf{v}}) = \vec{\mathbf{w}} \right\}$$

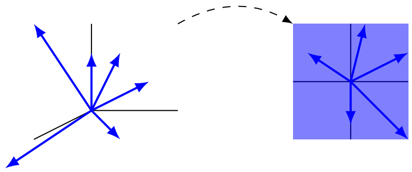
In the examples below, the left example's image is all of $\mathbb{R}^2$, but the right example's image is a planar subspace of $\mathbb{R}^3$.



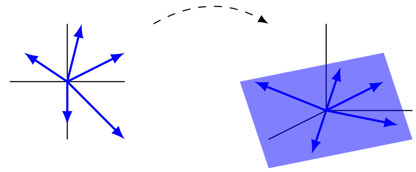
Definition: Let $T : V \to W$ be a linear transformation. T is called **injective** or **one-to-one** if T does not map two distinct vectors to the same place. More precisely, T is injective if $T(\vec{\mathbf{v}}) \neq T(\vec{\mathbf{w}})$ whenever $\vec{\mathbf{v}} \neq \vec{\mathbf{w}}$.



Definition: Let $T : V \to W$ be a linear transformation. T is called **surjective** or **onto** if every element of W is mapped to by an element of V. More precisely, for every $\vec{\mathbf{w}} \in W$, there is some $\vec{\mathbf{v}} \in V$ with $T(\vec{\mathbf{v}}) = \vec{\mathbf{w}}$.



surjective



not surjective

Module M: Understanding Matrices Algebraically

Definition: We define the **product** AB of a $m \times n$ matrix A and a $n \times k$ matrix B to be the $m \times k$ standard matrix of the composition map of the two corresponding linear functions.

For the previous activity, T was a map $\mathbb{R}^3 \to \mathbb{R}^2$, and S was a map $\mathbb{R}^2 \to \mathbb{R}^4$, so $S \circ T$ gave a map $\mathbb{R}^3 \to \mathbb{R}^4$ with a 4×3 standard matrix:

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

$$= [(S \circ T)(\vec{\mathbf{e}}_1) \quad (S \circ T)(\vec{\mathbf{e}}_2) \quad (S \circ T)(\vec{\mathbf{e}}_3)] = \begin{bmatrix} 12 & -5 & 5 \\ 5 & -3 & 4 \\ 31 & -12 & 11 \\ -12 & 5 & -5 \end{bmatrix}.$$

Definition: The identity matrix I_n (or just I when n is obvious from context) is the $n \times n$ matrix

$$I_n = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & 1 \end{bmatrix}.$$

It has a 1 on each diagonal element and a 0 in every other position.

Definition: Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be a linear map with standard matrix A.

- If T is a bijection and $\vec{\mathbf{b}}$ is any $\mathbb{R}^n$ vector, then $T(\vec{\mathbf{x}}) = A\vec{\mathbf{x}} = \vec{\mathbf{b}}$ has a unique solution.
- So we may define an **inverse map** $T^{-1} : \mathbb{R}^n \to \mathbb{R}^n$ by setting $T^{-1}(\vec{\mathbf{b}})$ to be this unique solution.
- Let A^{-1} be the standard matrix for T^{-1}. We call A^{-1} the **inverse matrix** of A, so we also say that A is **invertible**.

Module G: Geometry of Linear Maps

Definition: The **determinant** is the unique function $\det : M_{n,n} \to \mathbb{R}$ satisfying these properties:

P1: $\det(I) = 1$

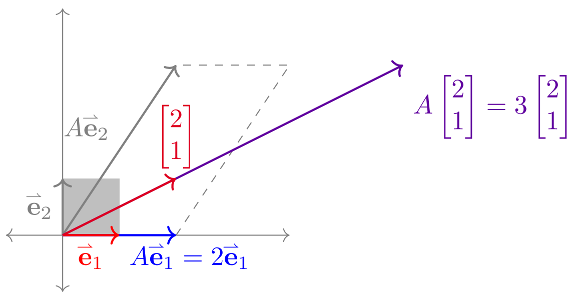
P2: $\det(A) = 0$ whenever two columns of the matrix are identical.

P3: $\det[\cdots \ c\vec{\mathbf{v}} \ \cdots] = c\det[\cdots \ \vec{\mathbf{v}} \ \cdots]$, assuming no other columns change.

P4: $\det[\cdots \ \vec{\mathbf{v}} + \vec{\mathbf{w}} \ \cdots] = \det[\cdots \ \vec{\mathbf{v}} \ \cdots] + \det[\cdots \ \vec{\mathbf{w}} \ \cdots]$, assuming no other columns change.

Note that these last two properties together can be phrased as "The determinant is linear in each column."

Definition: Let $A \in M_{n,n}$. An **eigenvector** for A is a vector $\vec{\mathbf{x}} \in \mathbb{R}^n$ such that $A\vec{\mathbf{x}}$ is parallel to $\vec{\mathbf{x}}$.



In other words, $A\vec{\mathbf{x}} = \lambda\vec{\mathbf{x}}$ for some scalar λ. If $\vec{\mathbf{x}} \neq \vec{\mathbf{0}}$, then we say $\vec{\mathbf{x}}$ is a **nontrivial eigenvector** and we call this λ an **eigenvalue** of A.

Definition: The expression $\det(A - \lambda I)$ is called **characteristic polynomial** of A.

For example, when $A = \begin{bmatrix}1 & 2\\3 & 4\end{bmatrix}$, we have

$$A - \lambda I = \begin{bmatrix}1 & 2\\3 & 4\end{bmatrix} - \begin{bmatrix}\lambda & 0\\0 & \lambda\end{bmatrix} = \begin{bmatrix}1-\lambda & 2\\3 & 4-\lambda\end{bmatrix}$$

Thus the characteristic polynomial of A is

$$\det\begin{bmatrix}1-\lambda & 2\\3 & 4-\lambda\end{bmatrix} = (1-\lambda)(4-\lambda) - (2)(3) = \lambda^2 - 5\lambda - 2$$

and its eigenvalues are the solutions to $\lambda^2 - 5\lambda - 2 = 0$.

Definition: Since the kernel of a linear map is a subspace of $\mathbb{R}^n$, and the kernel obtained from $A - \lambda I$ contains all the eigenvectors associated with λ, we call this kernel the **eigenspace** of A associated with λ.