

# Diagonalization

**Geometric Algorithms**

**Lecture 20**

# Objectives

1. Finish our discussion on the characteristic polynomial
2. Motivate diagonalization via linear dynamical systems and changes of coordinate systems
3. Describe how to diagonalize a matrix

# Keywords

multiplicity

similar matrices

diagonalizable matrices

change of basis

eigenbasis

# Recap: Characteristic Polynomial

# Recall: Determinants and Invertibility

# Recall: Determinants and Invertibility

$\det(A)$  is a value associated with the matrix  $A$

# Recall: Determinants and Invertibility

$\det(A)$  is a value associated with the matrix  $A$

**Theorem.** A matrix is invertible if and only if  $\det(A) \neq 0$

# Recall: Determinants and Invertibility

$\det(A)$  is a value associated with the matrix  $A$

**Theorem.** A matrix is invertible if and only if  $\det(A) \neq 0$

So by the Invertible Matrix Theorem:



# Recall: Determinants and Invertibility

$\det(A)$  is a value associated with the matrix  $A$

**Theorem.** A matrix is invertible if and only if  $\det(A) \neq 0$

So by the Invertible Matrix Theorem:

$$\begin{aligned} \det(A - \lambda I) = 0 &\quad \equiv \quad (A - \lambda I)\mathbf{x} = \mathbf{0} \text{ has nontrivial solutions} \\ &\quad \equiv \quad \lambda \text{ is an eigenvalue of } A \end{aligned}$$

# Recall: Determinants and Invertibility

$\det(A)$  is a value associated with the matrix  $A$

**Theorem.** A matrix is invertible if and only if  $\det(A) \neq 0$

So by the Invertible Matrix Theorem:

$$\begin{aligned} \det(A - \lambda I) = 0 &\equiv (A - \lambda I)\mathbf{x} = \mathbf{0} \text{ has nontrivial solutions} \\ \text{polynomial in } \lambda &\equiv \lambda \text{ is an eigenvalue of } A \end{aligned}$$

# How To: Finding Eigenvalues

# How To: Finding Eigenvalues

**Question.** Determine the eigenvalues of  $A$ .

# How To: Finding Eigenvalues

**Question.** Determine the eigenvalues of  $A$ .

**Solution.** Find the *roots* of the characteristic polynomial of  $A$ , which is

$$\det(A - \lambda I)$$

viewed as a *polynomial* in  $\lambda$ .

# How To: Finding Eigenvalues

**Question.** Determine the eigenvalues of  $A$ .

**Solution.** Find the *roots* of the characteristic polynomial of  $A$ , which is

$$\det(A - \lambda I)$$

viewed as a *polynomial* in  $\lambda$ .

We'll also use

*`numpy.linalg.eig(A)`*

# Example

$$A = \begin{bmatrix} 1 & -1 \\ 7 & -3 \end{bmatrix}$$

## Example: Triangular matrix

$$\begin{bmatrix} 1 & -3 & 0 & 6 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 4 \end{bmatrix}$$

The characteristic polynomial of a triangular matrix comes pre-factored:



# An Observation: Multiplicity

$$\lambda^1(\lambda - 1)^2(\lambda - 4)^1 \text{ multiplicities}$$

# An Observation: Multiplicity

$$\lambda^1(\lambda - 1)^2(\lambda - 4)^1$$

multiplicities

In the examples so far, we've seen a number appear as a root multiple times

# An Observation: Multiplicity

$$\lambda^1(\lambda - 1)^2(\lambda - 4)^1$$

multiplicities

In the examples so far, we've seen a number appear as a root multiple times

This is called the **(algebraic) multiplicity** of the root

# An Observation: Multiplicity

$$\lambda^1(\lambda - 1)^2(\lambda - 4)^1$$

multiplicities

In the examples so far, we've seen a number appear as a root multiple times

This is called the **(algebraic) multiplicity** of the root

**Is the multiplicity meaningful in this context?**

# Multiplicity and Dimension

**Theorem.** The dimension of the eigenspace of  $A$  for the eigenvalue  $\lambda$  is at most the multiplicity of  $\lambda$  in  $\det(A - \lambda I)$  (and at least 1)

The multiplicity is an upper bound on  
"how large" the eigenspace is

# Example

*Let  $A$  be a  $5 \times 5$  matrix with characteristic polynomial  $(x - 1)^3(x - 3)(x + 5)$*

*» What is  $\text{rank}(A)$ ?*

*» What is the minimum possible rank of  $A - I$ ?*

# Practice Problem

$$\begin{bmatrix} 5 & 1 \\ 4 & 2 \end{bmatrix}$$

*Determine the eigenvalues and an eigenbasis for the above matrix*

**Answer**

$$(\lambda - a)(\lambda - b) \\ = \lambda^2 - (a+b)\lambda + ab$$

$$A = \begin{bmatrix} 5 & 1 \\ 4 & 2 \end{bmatrix}$$

$$A - \lambda I = \begin{bmatrix} 5-\lambda & 1 \\ 4 & 2-\lambda \end{bmatrix}$$

$$\det(A - \lambda I) = (5-\lambda)(2-\lambda) - 4 \\ = \lambda^2 - 7\lambda + 6 = (\lambda - 1)(\lambda - 6)$$

$$A - I = \begin{bmatrix} 4 & 1 \\ 4 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 1/4 & 0 \\ 0 & 0 & 0 \end{bmatrix} \Rightarrow \vec{x} = x_2 \begin{bmatrix} -1/4 \\ 1 \end{bmatrix} \leftarrow \begin{matrix} \lambda_1 = 1 \\ \vec{v}_1 = \begin{bmatrix} -1/4 \\ 1 \end{bmatrix} \end{matrix}$$

$$A - 6I = \begin{bmatrix} -1 & 1 \\ 4 & -4 \end{bmatrix} \sim \begin{bmatrix} 1 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \Rightarrow \vec{x} = x_2 \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad \begin{matrix} \lambda_2 = 6 \\ \vec{v}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \end{matrix}$$

$$\left\{ \begin{bmatrix} -1/4 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\} \quad \text{eigenbasis}$$



# Motivating Diagonalization via Linear Dynamical Systems

(briefly)

# Recall: Eigenbasis

# Recall: Eigenbasis

**Definition.** An eigenbasis of  $H$  for the matrix  $A$  is a basis of  $H$  made up of eigenvectors of  $A$

# Recall: Eigenbasis

**Definition.** An eigenbasis of  $H$  for the matrix  $A$  is a basis of  $H$  made up of eigenvectors of  $A$

We will be almost exclusively interested of  
eigenbases of  $\mathbb{R}^n$  when  $A \in \mathbb{R}^{n \times n}$

# Recall: Eigenbasis

**Definition.** An eigenbasis of  $H$  for the matrix  $A$  is a basis of  $H$  made up of eigenvectors of  $A$

We will be almost exclusively interested of  
eigenbases of  $\mathbb{R}^n$  when  $A \in \mathbb{R}^{n \times n}$

**The Question.** When can we describe any vector in  $\mathbb{R}^n$  as a unique linear combination of eigenvectors of  $A$ ?

# Recall: Linear Dynamical Systems

$$\mathbf{v}_1 = A\mathbf{v}_0$$

$$\mathbf{v}_2 = A\mathbf{v}_1 = A^2\mathbf{v}_0$$

$$\mathbf{v}_3 = A\mathbf{v}_2 = A^3\mathbf{v}_0$$

$$\mathbf{v}_4 = A\mathbf{v}_3 = A^4\mathbf{v}_0$$

$\vdots$

A **linear dynamical system** describes a sequence of **state vectors** starting at  $\mathbf{v}_0$

# Recall: Linear Dynamical Systems

$$\mathbf{v}_1 = A\mathbf{v}_0$$

$$\mathbf{v}_2 = A\mathbf{v}_1 = A^2\mathbf{v}_0$$

$$\mathbf{v}_3 = A\mathbf{v}_2 = A^3\mathbf{v}_0$$

$$\mathbf{v}_4 = A\mathbf{v}_3 = A^4\mathbf{v}_0$$

$\vdots$

A **linear dynamical system** describes a sequence of **state vectors** starting at  $\mathbf{v}_0$

multiplying by  
 $A$  changes the  
state.

demo



# **Eigenbases and Closed-Form solutions**

# Eigenbases and Closed-Form solutions

Given  $\mathbf{v}_k = A\mathbf{v}_{k-1} = A^k\mathbf{v}_0$ , if

$$\mathbf{v}_0 = \alpha_1\mathbf{b}_1 + \alpha_2\mathbf{b}_2 + \alpha_3\mathbf{b}_3$$

# Eigenbases and Closed-Form solutions

Given  $\mathbf{v}_k = A\mathbf{v}_{k-1} = A^k\mathbf{v}_0$ , if

$$\mathbf{v}_0 = \alpha_1 \mathbf{b}_1 + \alpha_2 \mathbf{b}_2 + \alpha_3 \mathbf{b}_3$$

eigenvectors of  $A$

# Eigenbases and Closed-Form solutions

Given  $\mathbf{v}_k = A\mathbf{v}_{k-1} = A^k\mathbf{v}_0$ , if

eigenvectors of  $A$

$$\mathbf{v}_0 = \alpha_1 \mathbf{b}_1 + \alpha_2 \mathbf{b}_2 + \alpha_3 \mathbf{b}_3$$

then

$$A^k\mathbf{v}_0 = \alpha_1 \lambda_1^k \mathbf{b}_1 + \alpha_2 \lambda_2^k \mathbf{b}_2 + \alpha_3 \lambda_3^k \mathbf{b}_3$$

# Eigenbases and Closed-Form solutions

Given  $\mathbf{v}_k = A\mathbf{v}_{k-1} = A^k\mathbf{v}_0$ , if

eigenvectors of  $A$

$$\mathbf{v}_0 = \alpha_1 \mathbf{b}_1 + \alpha_2 \mathbf{b}_2 + \alpha_3 \mathbf{b}_3$$

then

eigenvalues of  $A$

$$A^k\mathbf{v}_0 = \alpha_1 \lambda_1^k \mathbf{b}_1 + \alpha_2 \lambda_2^k \mathbf{b}_2 + \alpha_3 \lambda_3^k \mathbf{b}_3$$

# Eigenbases and Closed-Form solutions

Given  $\mathbf{v}_k = A\mathbf{v}_{k-1} = A^k\mathbf{v}_0$ , if

eigenvectors of  $A$

$$\mathbf{v}_0 = \alpha_1 \mathbf{b}_1 + \alpha_2 \mathbf{b}_2 + \alpha_3 \mathbf{b}_3$$

then

eigenvalues of  $A$

$$A^k \mathbf{v}_0 = \alpha_1 \lambda_1^k \mathbf{b}_1 + \alpha_2 \lambda_2^k \mathbf{b}_2 + \alpha_3 \lambda_3^k \mathbf{b}_3$$

closed-form solution

Verify:

# Application: Eigenbases and Limiting Behavior

**Theorem.** If  $A$  has an eigenbasis with eigenvalues

$$\lambda_1 \geq \lambda_2 \dots \geq \lambda_k$$

then  $\mathbf{v}_k \sim \lambda_1^k \mathbf{u}$  for some vector  $\mathbf{u}$ .

**In the long term, the system grows exponentially in  $\lambda_1$ .**

# Application: Eigenbases and Limiting Behavior

**Theorem.** If  $A$  has an eigenbasis with eigenvalues

$$\lambda_1 \geq \lambda_2 \dots \geq \lambda_k$$

then  $\mathbf{v}_k \sim \lambda_1^k \mathbf{u}$  for some vector  $\mathbf{u}$ .

**In the long term, the system grows exponentially in  $\lambda_1$ .**



# The Takeaway

# The Takeaway

Given a basis  $\mathcal{B}$  for  $\mathbb{R}^n$ , we only need to know how  $A \in \mathbb{R}^n$  behaves on  $\mathcal{B}$ .

# The Takeaway

Given a basis  $\mathcal{B}$  for  $\mathbb{R}^n$ , we only need to know how  $A \in \mathbb{R}^n$  behaves on  $\mathcal{B}$ .

Sometimes,  $A$  behaves simply on  $\mathcal{B}$ , as in the case of eigenbases.

# The Takeaway

Given a basis  $\mathcal{B}$  for  $\mathbb{R}^n$ , we only need to know how  $A \in \mathbb{R}^n$  behaves on  $\mathcal{B}$ .

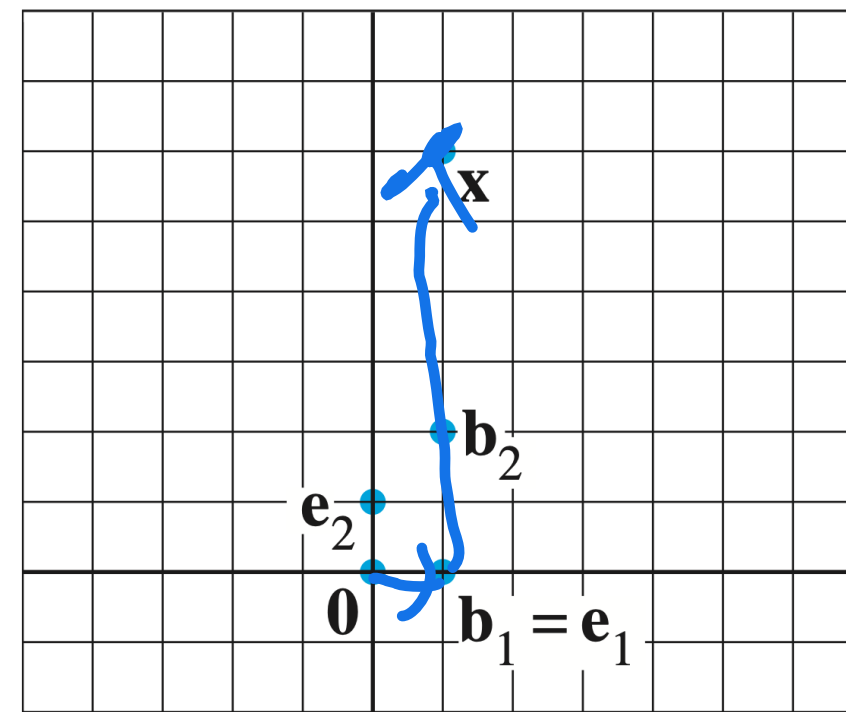
Sometimes,  $A$  behaves simply on  $\mathcal{B}$ , as in the case of eigenbases.

**What we're really doing is changing our coordinate system to expose a behavior of  $A$ .**

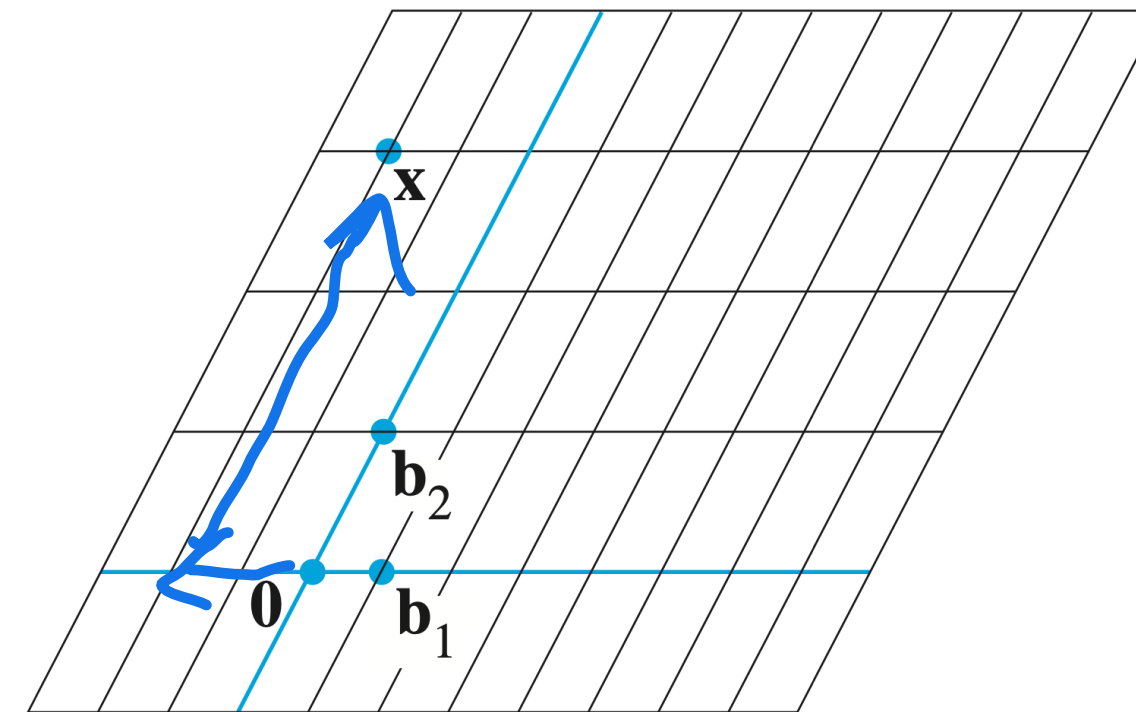
# ~~Recap:~~ Change of Basis

# Recall: Bases define Coordinate Systems

$$\vec{x} = \begin{bmatrix} 1 \\ 6 \end{bmatrix}$$



**FIGURE 1** Standard graph paper.



**FIGURE 2**  $\mathcal{B}$ -graph paper.

$$[\vec{x}]_{\mathcal{B}} = \begin{bmatrix} -2 \\ 3 \end{bmatrix}$$

$$[\vec{b}_1, \vec{b}_2 | \vec{x}]$$

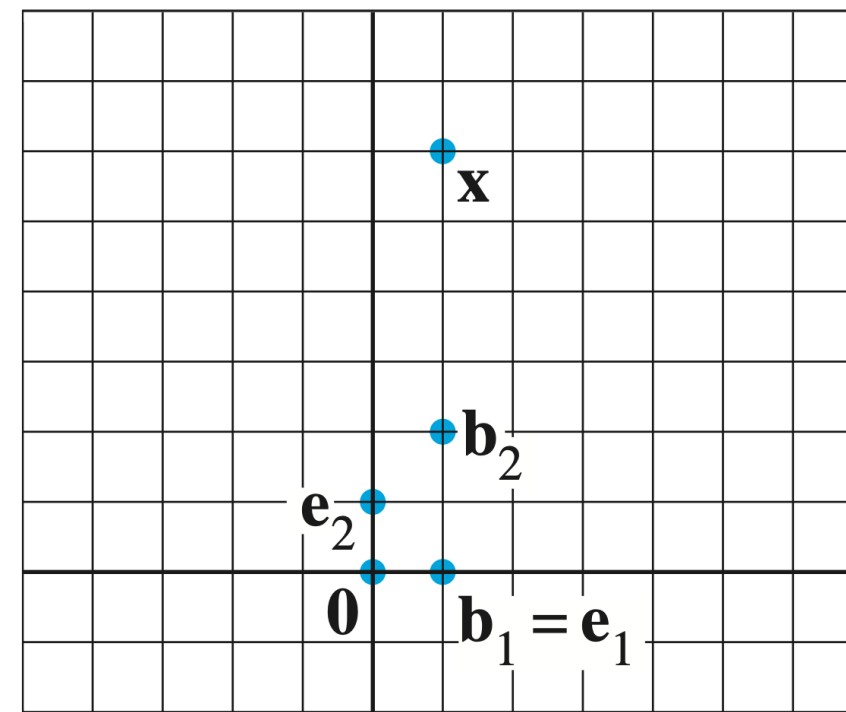
Want:  $\vec{x} = a_1 \vec{b}_1 + a_2 \vec{b}_2$

$$[\vec{b}_1, \vec{b}_2] \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = [\vec{x}]$$

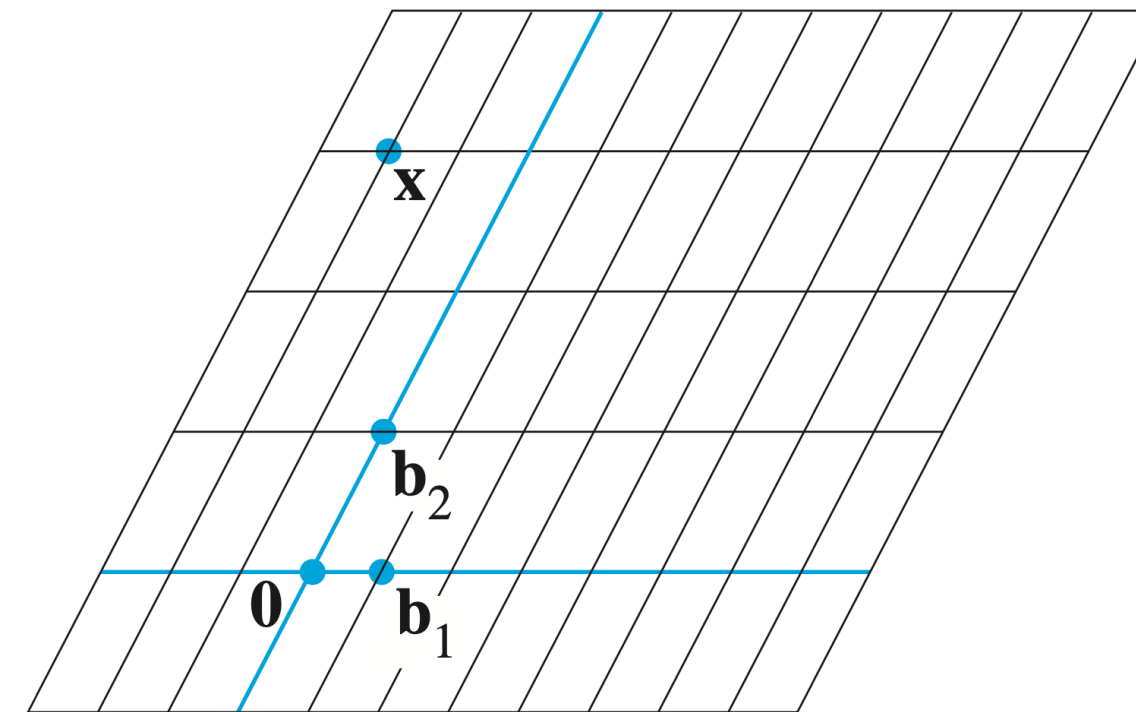
$$[\vec{x}]_{\mathcal{B}} = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = [\vec{b}_1, \vec{b}_2]^{-1} [\vec{x}]$$

↑  
COB matrix

# Recall: Bases define Coordinate Systems



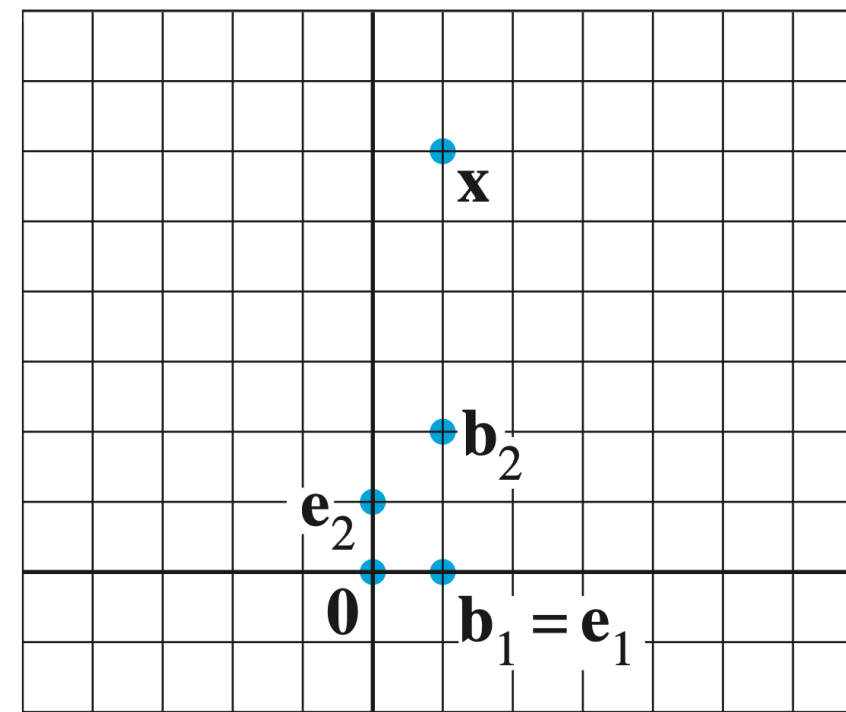
**FIGURE 1** Standard graph paper.



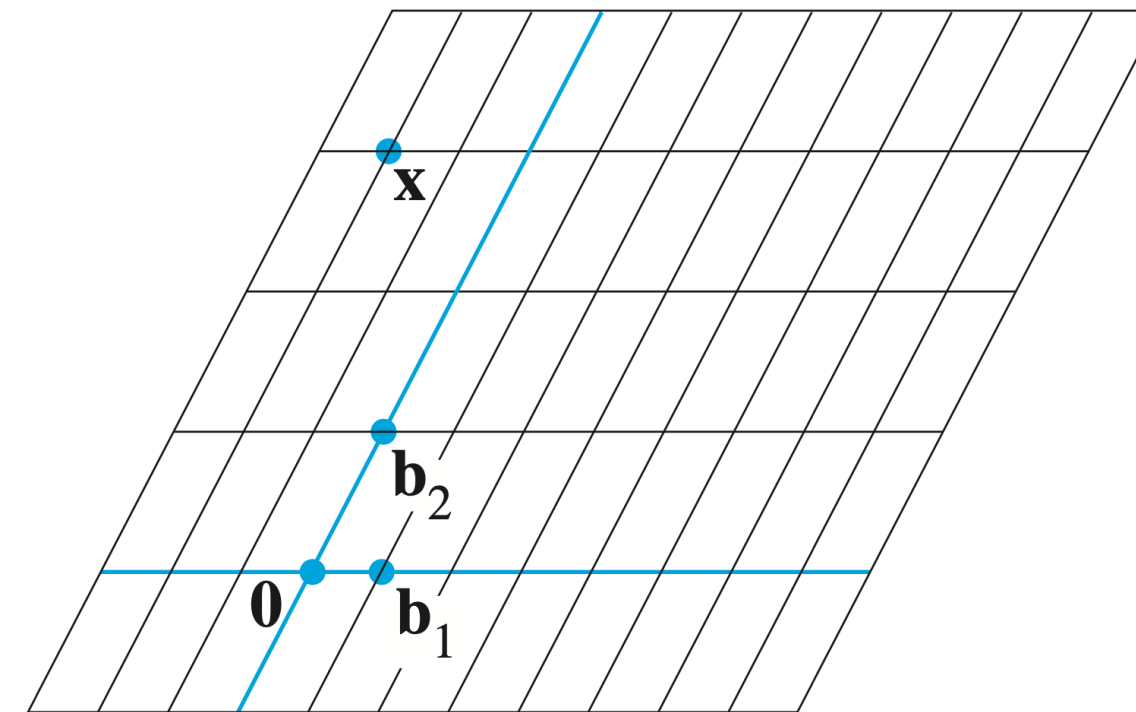
**FIGURE 2**  $\mathcal{B}$ -graph paper.

Given a basis  $\mathcal{B}$  of  $\mathbb{R}^n$ , there is **exactly one way** to write every vector as a linear combination of vectors in  $\mathcal{B}$

# Recall: Bases define Coordinate Systems



**FIGURE 1** Standard graph paper.



**FIGURE 2**  $\mathcal{B}$ -graph paper.

Given a basis  $\mathcal{B}$  of  $\mathbb{R}^n$ , there is **exactly one way** to write every vector as a linear combination of vectors in  $\mathcal{B}$

Every basis provides a way to write down *coordinates* of a vector



# Recall: Bases define Coordinate Systems

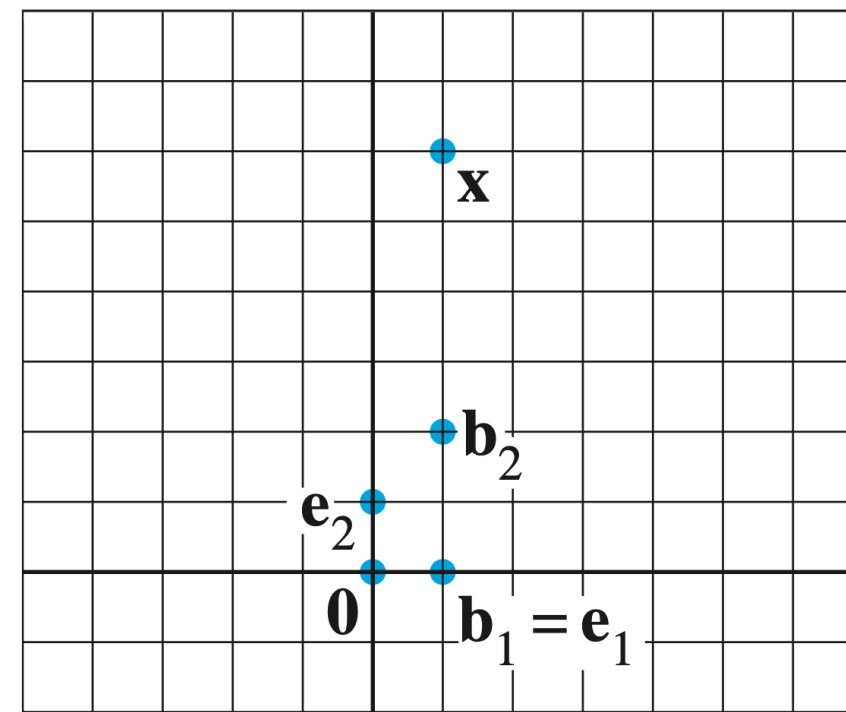


FIGURE 1 Standard graph paper.

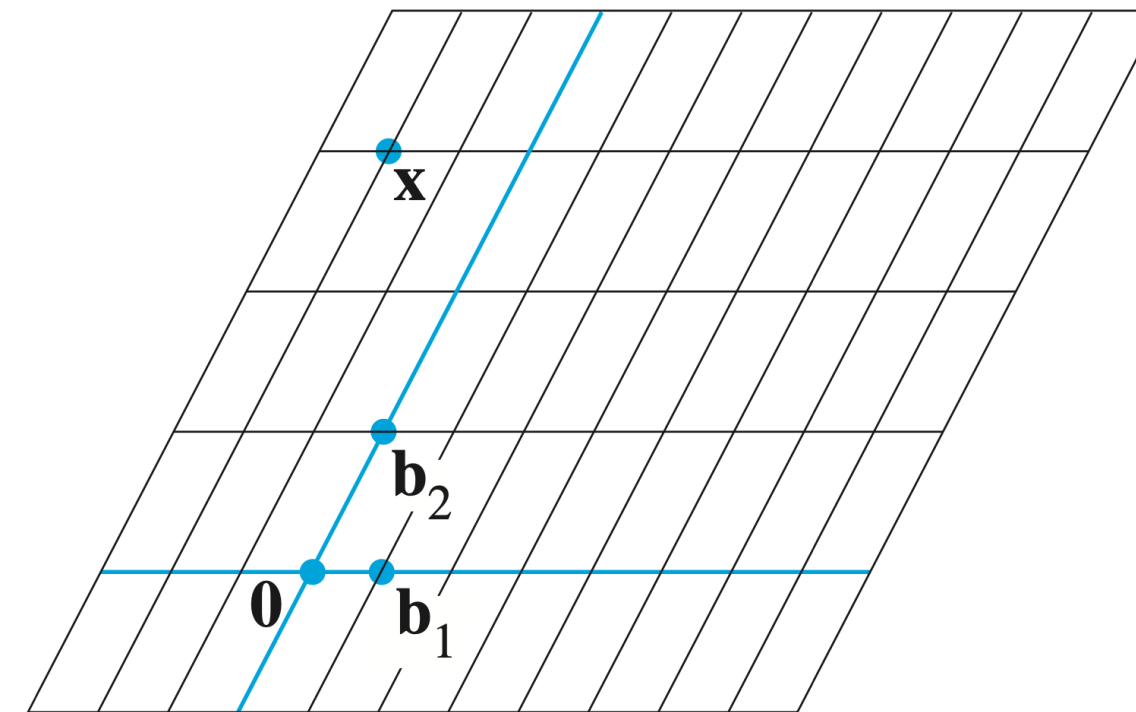


FIGURE 2  $\mathcal{B}$ -graph paper.

Given a basis  $\mathcal{B}$  of  $\mathbb{R}^n$ , there is **exactly one way** to write every vector as a linear combination of vectors in  $\mathcal{B}$

Every basis provides a way to write down *coordinates* of a vector

$\mathcal{B}$  defines a "different grid for our graph paper"

# Recall: Coordinate Vectors

# Recall: Coordinate Vectors

Let  $\mathbf{v}$  be a vector in a  $\mathbb{R}^n$  and let  $\mathcal{B} = (\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n)$  be a basis of  $\mathbb{R}^n$  where

$$\mathbf{v} = a_1\mathbf{b}_1 + a_2\mathbf{b}_2 + \dots + a_n\mathbf{b}_n$$

# Recall: Coordinate Vectors

Let  $\mathbf{v}$  be a vector in a  $\mathbb{R}^n$  and let  $\mathcal{B} = (\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n)$  be a basis of  $\mathbb{R}^n$  where

$$\mathbf{v} = a_1\mathbf{b}_1 + a_2\mathbf{b}_2 + \dots + a_n\mathbf{b}_n$$

**Definition.** The coordinate vector of  $\mathbf{v}$  relative to  $\mathcal{B}$  is

# Recall: Coordinate Vectors

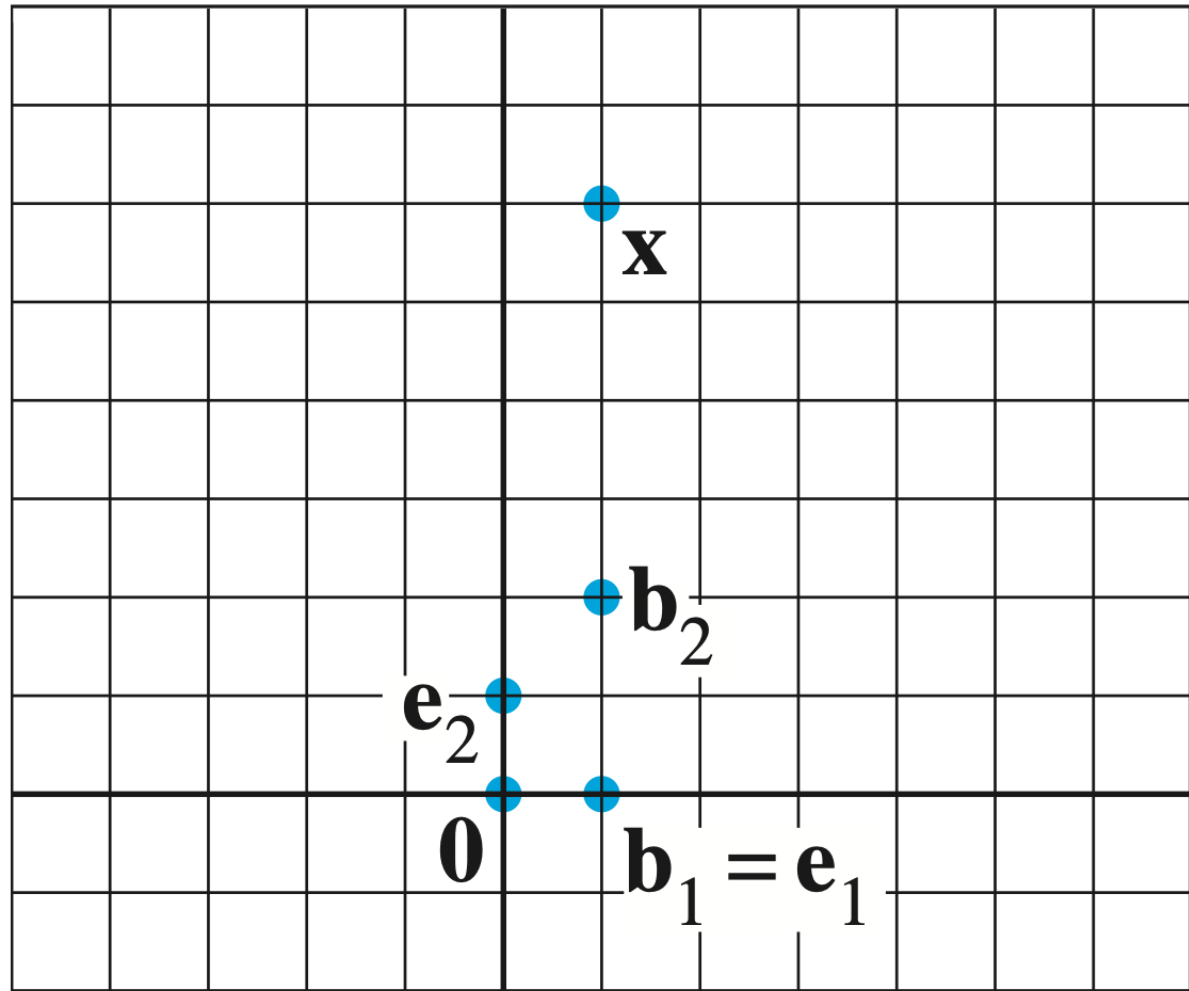
Let  $\mathbf{v}$  be a vector in a  $\mathbb{R}^n$  and let  $\mathcal{B} = (\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n)$  be a basis of  $\mathbb{R}^n$  where

$$\mathbf{v} = a_1 \mathbf{b}_1 + a_2 \mathbf{b}_2 + \dots + a_n \mathbf{b}_n$$

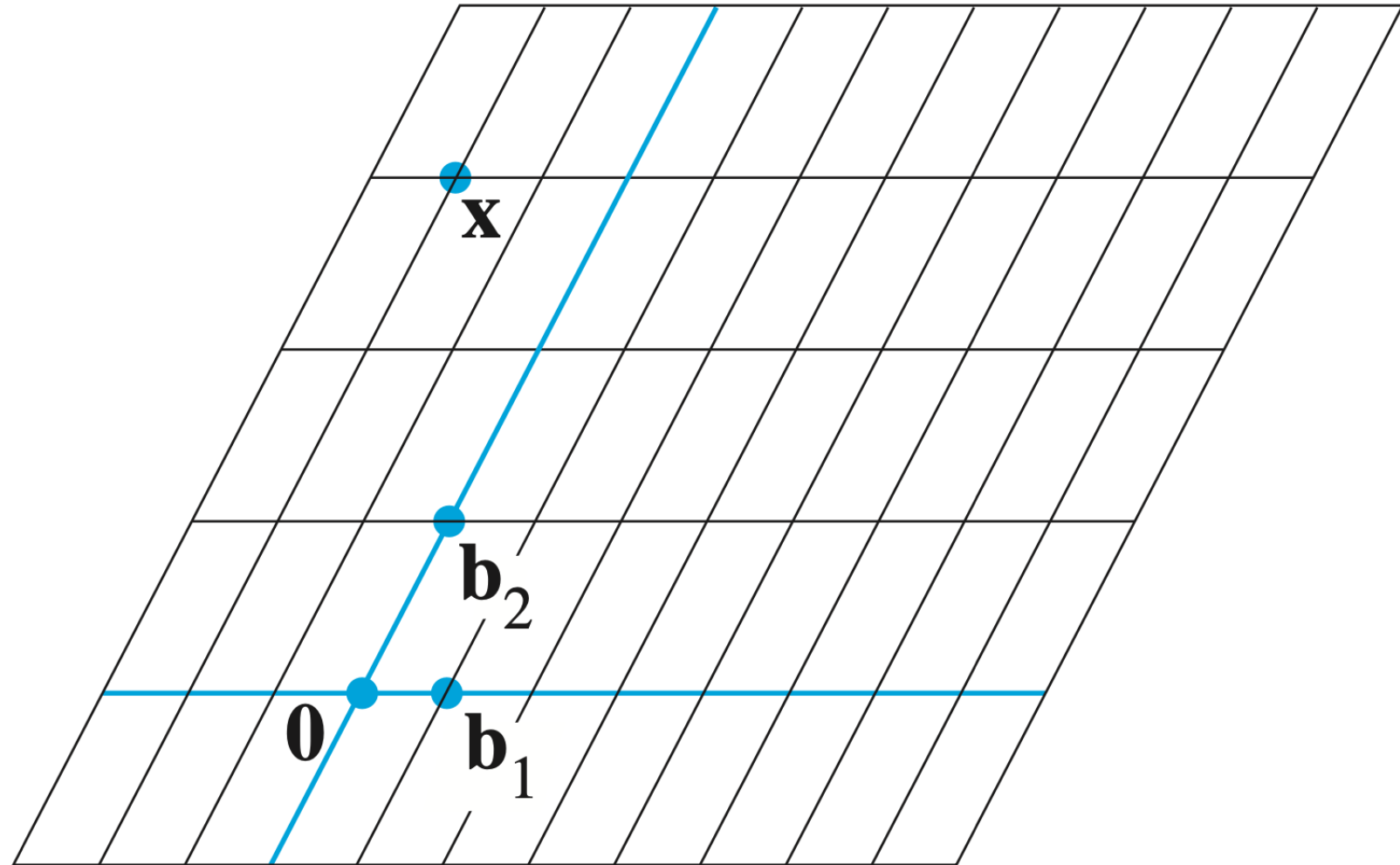
**Definition.** The coordinate vector of  $\mathbf{v}$  relative to  $\mathcal{B}$  is

$$[\mathbf{v}]_{\mathcal{B}} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}$$

# Recall: Coordinate Vectors



**FIGURE 1** Standard graph paper.



**FIGURE 2**  $\mathcal{B}$ -graph paper.

# Question (Conceptual)

We know that if a  $n \times n$  matrix  $B = [\mathbf{b}_1 \ \mathbf{b}_2 \ \dots \ \mathbf{b}_n]$  is invertible, then the columns of  $B$  form a basis  $\mathcal{B}$  of  $\mathbb{R}^n$

*What is the matrix that implements the transformation*

$$\mathbf{x} \mapsto [\mathbf{x}]_{\mathcal{B}} = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix}$$

*where  $\mathbf{x} = c_1 \mathbf{b}_1 + c_2 \mathbf{b}_2 + \dots + c_n \mathbf{b}_n$ ?*

# Change of Basis Matrix

**Theorem.** If  $\mathcal{B} = \{\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n\}$  form a basis of  $\mathbb{R}^n$ , then

$$[\mathbf{x}]_{\mathcal{B}} = [\mathbf{b}_1 \ \mathbf{b}_2 \ \dots \ \mathbf{b}_n]^{-1} \mathbf{x}$$

**Matrix inverses perform changes of bases.**

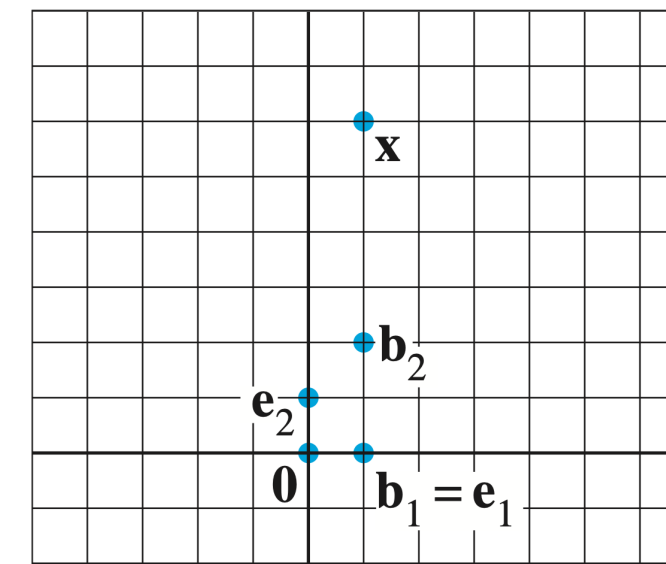


# How To: Change of Basis

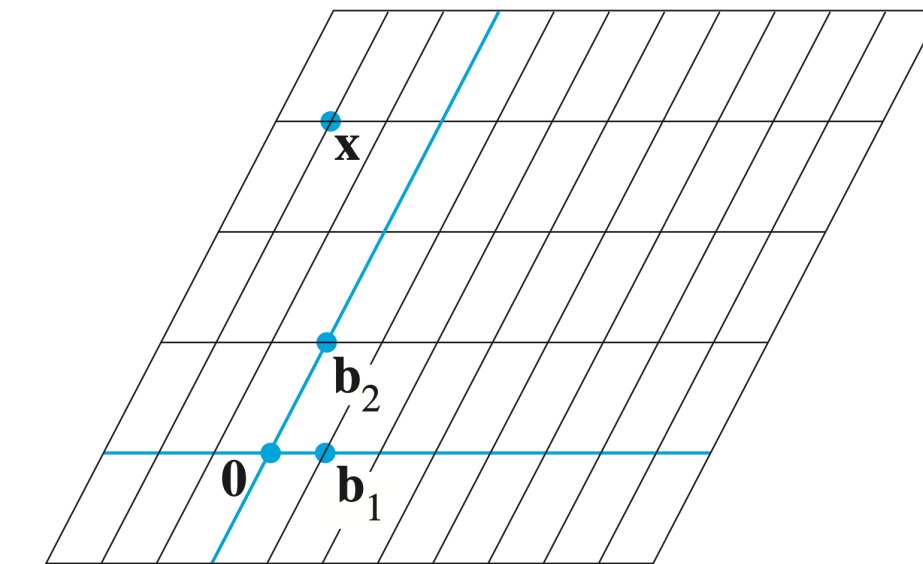
**Question.** Given a basis  $\mathcal{B} = (\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_n)$  of  $\mathbb{R}^n$ , find the matrix which implements  $\mathbf{x} \mapsto [\mathbf{x}]_{\mathcal{B}}$ .

**Solution.** Construct the matrix  $[\mathbf{b}_1 \ \mathbf{b}_2 \ \dots \ \mathbf{b}_n]^{-1}$ .

# Example



**FIGURE 1** Standard graph paper.



**FIGURE 2**  $\mathcal{B}$ -graph paper.

Write the change-of-bases matrix for the basis  $\left( \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \end{bmatrix} \right)$

$$\begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix}^{-1} = \begin{bmatrix} 1 & -1/2 \\ 0 & 1/2 \end{bmatrix}$$

$$A^{-1} = \frac{1}{\det A} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

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

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

# Diagonalization

# Diagonal Matrices

ex. 
$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -0.4 & 0 & 0 \\ 0 & 0 & 22 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

# Diagonal Matrices

ex. 
$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -0.4 & 0 & 0 \\ 0 & 0 & 22 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

**Definition.** A  $n \times n$  matrix  $A$  is **diagonal** if

$i \neq j$  ~~if and only if~~  $A_{ij} = 0$



# Diagonal Matrices

ex. 
$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -0.4 & 0 & 0 \\ 0 & 0 & 22 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

**Definition.** A  $n \times n$  matrix  $A$  is **diagonal** if

$$i \neq j \quad \text{if and only if} \quad A_{ij} = 0$$

*Only the diagonal entries can be nonzero*

# Diagonal Matrices

ex. 
$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -0.4 & 0 & 0 \\ 0 & 0 & 22 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

**Definition.** A  $n \times n$  matrix  $A$  is **diagonal** if

$$i \neq j \quad \text{if and only if} \quad A_{ij} = 0$$

*Only the diagonal entries can be nonzero*

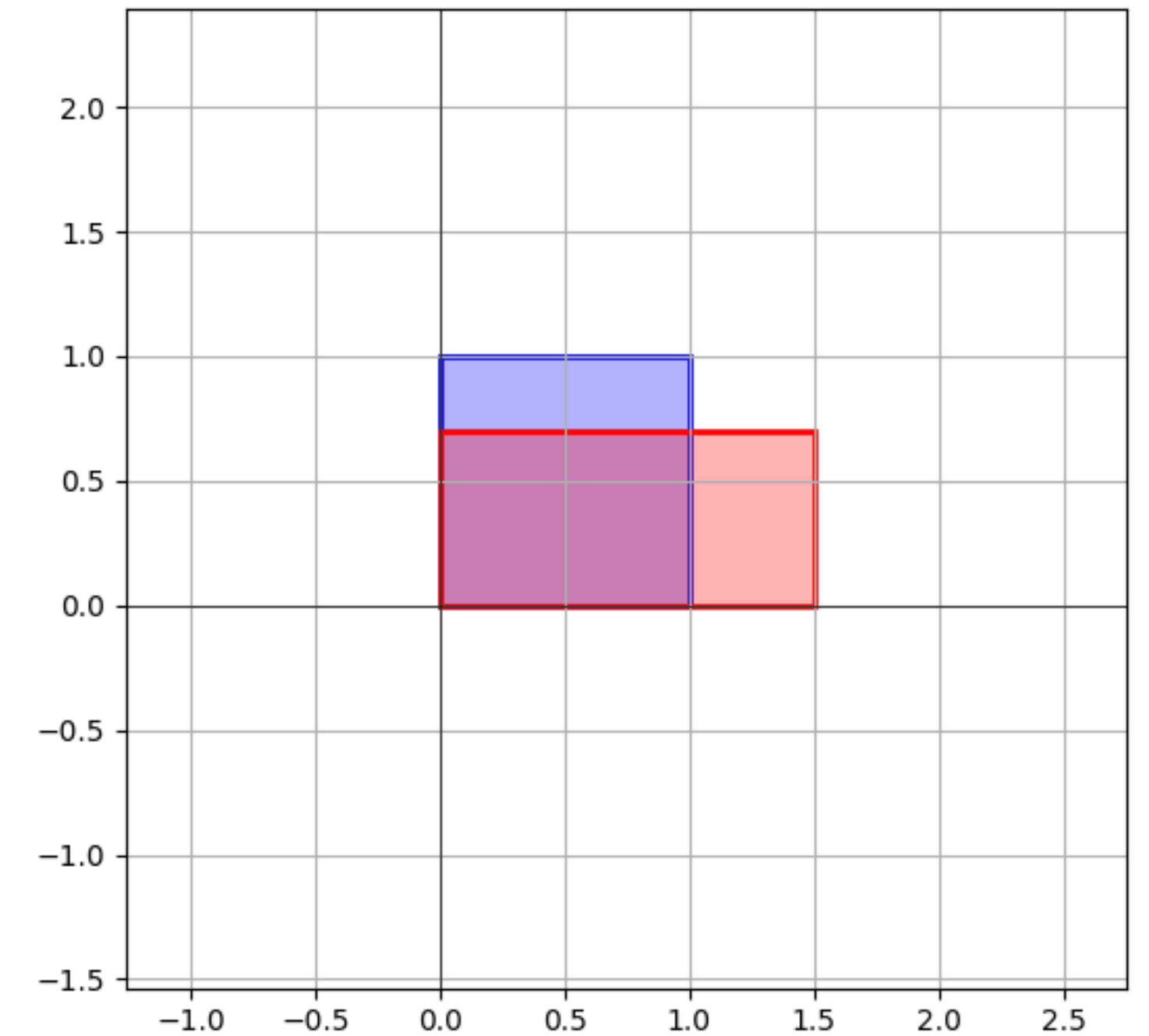
**Diagonal matrices are scaling matrices**

# Recall: Unequal Scaling

The scaling matrix *affects each component of a vector in a simple way*

The diagonal entries scale each corresponding entry

$$\begin{bmatrix} 1.5 & 0 \\ 0 & 0.7 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 1.5x \\ 0.7y \end{bmatrix}$$



$$\begin{bmatrix} 1.5 & 0 \\ 0 & 0.7 \end{bmatrix}$$



**High level question:** *unequal*  
When do matrices "behave" like scaling  
matrices "up to" change of basis?

# Scaling and Eigenvectors

# Scaling and Eigenvectors

**The idea.** Matrices behave like scaling matrices on eigenvectors.

# Scaling and Eigenvectors

**The idea.** Matrices behave like scaling matrices on eigenvectors.

$$\begin{bmatrix} 2 & 0 \\ 0 & -3 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix} (x\mathbf{e}_1 + y\mathbf{e}_2) = x2\mathbf{e}_1 + y(-3)\mathbf{e}_2$$

$$A \begin{bmatrix} x \\ y \end{bmatrix}_{\mathcal{B}} = A(x\mathbf{b}_1 + y\mathbf{b}_2) = x\lambda_1\mathbf{b}_1 + y\lambda_2\mathbf{b}_2$$

# Scaling and Eigenvectors

**The idea.** Matrices behave like scaling matrices on eigenvectors.

$$\begin{bmatrix} 2 & 0 \\ 0 & -3 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix} (x\mathbf{e}_1 + y\mathbf{e}_2) = x2\mathbf{e}_1 + y(-3)\mathbf{e}_2$$

$$A \begin{bmatrix} x \\ y \end{bmatrix}_{\mathcal{B}} = A(x\mathbf{b}_1 + y\mathbf{b}_2) = x\lambda_1\mathbf{b}_1 + y\lambda_2\mathbf{b}_2$$

**The fundamental question:**

Can we expose this behavior in  
terms of a *matrix factorization*?

# Recall: Matrix Factorization

# Recall: Matrix Factorization

A **factorization** of a matrix  $A$  is an equation which expresses  $A$  as a product of one or more matrices, e.g.,

$$A = PBP^{-1}$$



# Recall: Matrix Factorization

A **factorization** of a matrix  $A$  is an equation which expresses  $A$  as a product of one or more matrices, e.g.,

$$A = PBP^{-1}$$

Factorizations can:

- » make working with  $A$  easier
- » expose important information about  $A$

# Similar Matrices

$$A = PBP^{-1}$$

**Definition.** A matrix  $A$  is **similar** to a matrix  $B$  if there is some invertible matrix  $P$  such that  $A = PBP^{-1}$

$A$  and  $B$  are the same up to a change of basis

# Similar Matrices and Eigenvalues

**Theorem.** Similar matrices have the same eigenvalues.

Verify:

# Diagonalizable Matrices

# Diagonalizable Matrices

**Definition.** A matrix  $A$  is **diagonalizable** if it is similar to a diagonal matrix

# Diagonalizable Matrices

**Definition.** A matrix  $A$  is **diagonalizable** if it is similar to a diagonal matrix

*There is an invertible matrix  $P$  and diagonal matrix  $D$  such that  $A = PDP^{-1}$*

# Diagonalizable Matrices

**Definition.** A matrix  $A$  is **diagonalizable** if it is similar to a diagonal matrix

*There is an invertible matrix  $P$  and diagonal matrix  $D$  such that  $A = PDP^{-1}$*

**Diagonalizable matrices are the same as scaling matrices up to a change of basis**

# Important: Not all Matrices are Diagonalizable

**This is very different from the LU factorization**

We will need to figure out which matrices are diagonalizable

Question. *Is the zero matrix diagonalizable?*



# Application: Matrix Powers

**Theorem.** If  $A = PBP^{-1}$ , then  $A^k = P$ <sup>only take the power of  $B$</sup>  $B^kP^{-1}$

It may be easier to take the power of  $B$  (as in the case of diagonal matrices)

Verify:

# How To: Matrix Powers

**Question.** Given  $A$  is diagonalizable, determine  $A^k$

**Solution.** Find it's diagonalization  $PDP^{-1}$  and then compute  $PD^kP^{-1}$

*Remember that*

$$\begin{bmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{bmatrix}^k = \begin{bmatrix} a^k & 0 & 0 \\ 0 & b^k & 0 \\ 0 & 0 & c^k \end{bmatrix}$$

But how do we find the  
diagonalization...

# Diagonalization and Eigenvectors

Suppose we have a diagonalization

$$A = PDP^{-1}$$

What do we know about it?

<sup>(space)</sup>  
**Columns of  $P$  are eigenvectors**

$$A = [\mathbf{p}_1 \ \mathbf{p}_2 \ \mathbf{p}_3] \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix} [\mathbf{p}_1 \ \mathbf{p}_2 \ \mathbf{p}_3]^{-1}$$

**Verify:**

# Columns of $P$ are eigenvectors

$$A = [\mathbf{p}_1 \ \mathbf{p}_2 \ \mathbf{p}_3] \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix} [\mathbf{p}_1 \ \mathbf{p}_2 \ \mathbf{p}_3]^{-1}$$

# Columns of $P$ are eigenvectors

$$A = [\mathbf{p}_1 \ \mathbf{p}_2 \ \mathbf{p}_3] \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix} [\mathbf{p}_1 \ \mathbf{p}_2 \ \mathbf{p}_3]^{-1}$$

In fact, the columns of  $P$  form an **eigenbasis** of  $\mathbb{R}^n$  for  $A$



# Columns of $P$ are eigenvectors

$$A = [\mathbf{p}_1 \ \mathbf{p}_2 \ \mathbf{p}_3] \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix} [\mathbf{p}_1 \ \mathbf{p}_2 \ \mathbf{p}_3]^{-1}$$

In fact, the columns of  $P$  form an **eigenbasis** of  $\mathbb{R}^n$  for  $A$

And the entries of  $D$  are the **eigenvalues** associated to each eigenvector

# Columns of $P$ are eigenvectors

$$A = \overset{\text{eigenbasis}}{[p_1 \ p_2 \ p_3]} \overset{\text{eigenvalues}}{\begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix}} [p_1 \ p_2 \ p_3]^{-1}$$

In fact, the columns of  $P$  form an **eigenbasis** of  $\mathbb{R}^n$  for  $A$

And the entries of  $D$  are the **eigenvalues** associated to each eigenvector

**A diagonalization exposes a lot of information about  $A$**

# The Diagonalization Theorem

# The Diagonalization Theorem

**Theorem.** A matrix is diagonalizable if and only if it has an eigenbasis

# The Diagonalization Theorem

**Theorem.** A matrix is diagonalizable if and only if it has an eigenbasis

(we just did the hard part, if a matrix is diagonalizable then it has an **eigenbasis**)

# The Diagonalization Theorem

**Theorem.** A matrix is diagonalizable if and only if it has an eigenbasis

(we just did the hard part, if a matrix is diagonalizable then it has an **eigenbasis**)

We can use the same recipe to go in the other direction, given an eigenbasis, we can **build a diagonalization**

# Diagonalizing a Matrix

# High Level

$$A = PDP^{-1}$$



# High Level

$$A = PDP^{-1}$$

The columns of  $P$  form an eigenbasis for  $A$

# High Level

$$A = PDP^{-1}$$

The columns of  $P$  form an eigenbasis for  $A$

The diagonal of  $D$  are the eigenvalues for each column of  $P$

# High Level

$$A = PDP^{-1}$$

The columns of  $P$  form an eigenbasis for  $A$

The diagonal of  $D$  are the eigenvalues for each column of  $P$

**The matrix  $P^{-1}$  is a change of basis to this eigenbasis of  $A$**

# Step 1: Eigenvalues

$$A = \begin{bmatrix} 1 & 3 & 3 \\ -3 & -5 & 3 \\ 3 & 3 & 1 \end{bmatrix}$$

Find all the eigenvalues of  $A$

*Find the roots of  $\det(A - \lambda I)$*

*e.g.*

$$\det(A - \lambda I) = -(\lambda - 1)(\lambda + 2)^2$$

## Step 2: Eigenvectors

$$A = \begin{bmatrix} 1 & 3 & 3 \\ -3 & -5 & 3 \\ 3 & 3 & 1 \end{bmatrix}$$

$$\lambda_1 = 1$$

Find **bases** of the corresponding eigenspaces  $\lambda_2 = -2$

*e.g.*

$$\text{Nul}(A - I) = \text{span} \left\{ \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} \right\}$$

$$\text{Nul}(A + 2I) = \text{span} \left\{ \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \right\}$$

# Step 3: Construct P

$$A = \begin{bmatrix} 1 & 3 & 3 \\ -3 & -5 & 3 \\ 3 & 3 & 1 \end{bmatrix}$$

$$\lambda_1 = 1$$

$$\lambda_2 = -2$$

If there are  $n$  eigenvectors from the previous step they form an **eigenbasis**

Build the matrix with these vectors as the columns

*e.g.*

$$P = \begin{bmatrix} 1 & -1 & -1 \\ -1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$$

$$\text{Nul}(A - I) = \text{span} \left\{ \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} \right\}$$

$$\text{Nul}(A + 2I) = \text{span} \left\{ \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \right\}$$

## Step 5: Construct D

Build the matrix with eigenvalues as diagonal entries

**Note the order.** It should be the same as the order of columns of  $P$

*e.g.*

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

$$A = \begin{bmatrix} 1 & 3 & 3 \\ -3 & -5 & 3 \\ 3 & 3 & 1 \end{bmatrix}$$

$$\lambda_1 = 1$$

$$\lambda_2 = -2$$

$$P = \begin{bmatrix} 1 & -1 & -1 \\ -1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$$

## Step 6: Invert P

Find the inverse of  $P$  (we know how to do this)

$$A = \begin{bmatrix} 1 & 3 & 3 \\ -3 & -5 & 3 \\ 3 & 3 & 1 \end{bmatrix}$$

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

$$P = \begin{bmatrix} 1 & -1 & -1 \\ -1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$$



# Putting it Together

$A$

$P$

$D$

$P^{-1}$

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

# How to: Diagonalizing a Matrix

**Question.** Find a diagonalization of  $A \in \mathbb{R}^n$ , or determine that  $A$  is not diagonalizable

**Solution.**

1. Find the eigenvalues of  $A$ , and bases for their eigenspaces. If these eigenvectors don't form a basis of  $\mathbb{R}^n$ , then  $A$  is **not diagonalizable**
2. Otherwise, build a matrix  $P$  whose columns are the eigenvectors of  $A$
3. Then build a diagonal matrix  $D$  whose entries are the eigenvalues of  $A$  *in the same order*
4. Invert  $P$
5. The diagonalization of  $A$  is  $PDP^{-1}$

We know how to do every step, its  
a matter of putting it all  
together

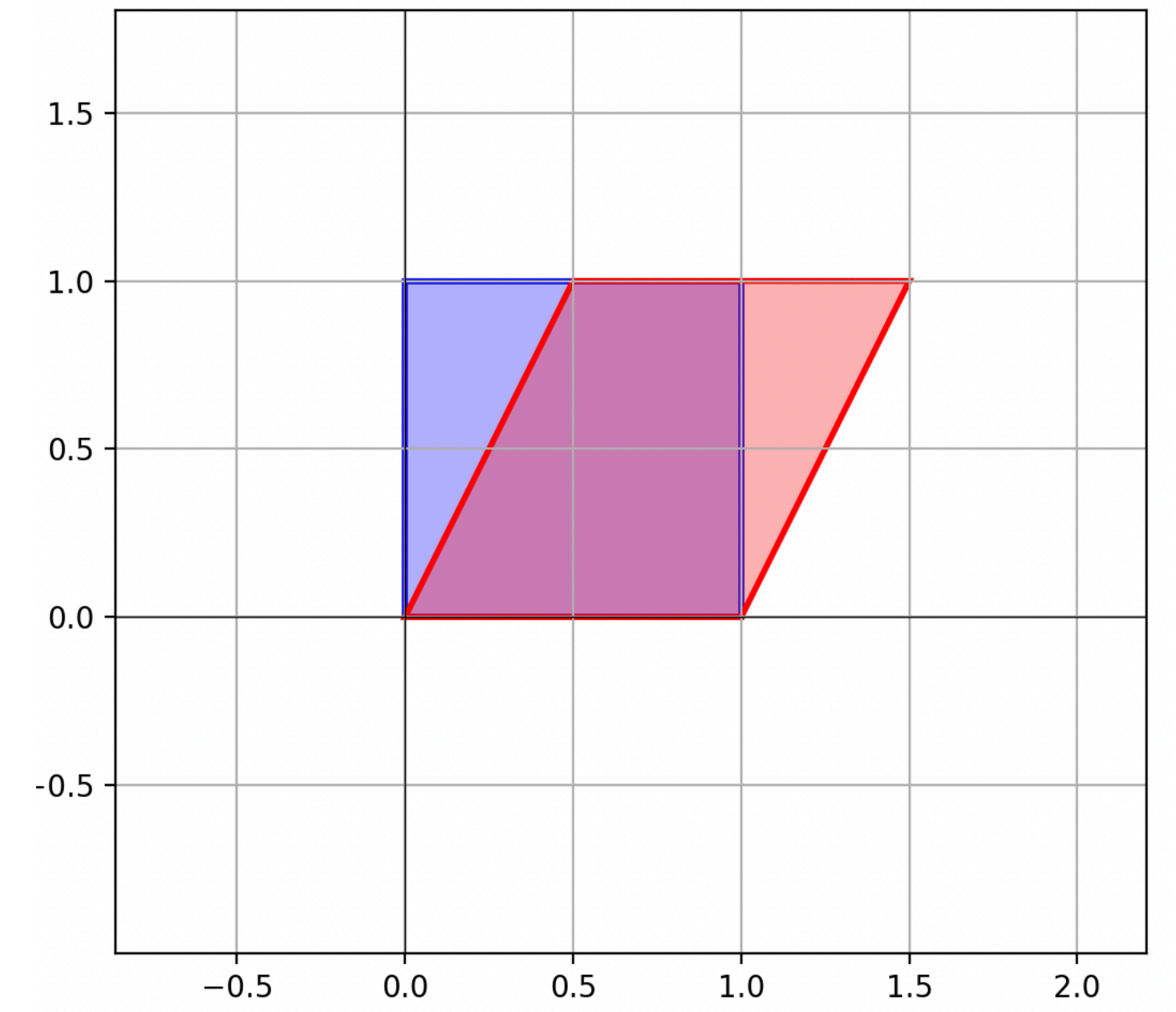
# Example of Failure: Shearing

$$A = \begin{bmatrix} 1 & 0.5 \\ 0 & 1 \end{bmatrix}$$

The shearing matrix has a single eigenvalue with an eigenspace of dimension 1

**We can't build an eigenbasis of  $\mathbb{R}^2$  for  $A$**

In other words,  $A$  is not diagonalizable



## Important case: Distinct Eigenvalues

ex. 
$$\begin{bmatrix} 1 & -3 & 4 & 2 \\ 0 & -2 & 3 & -1 \\ 0 & 0 & 10 & 5 \\ 0 & 0 & 0 & 6 \end{bmatrix}$$

**Theorem.** If an  $n \times n$  matrix has  $n$  distinct eigenvalues, then it is diagonalizable

This is because eigenvectors with distinct eigenvalues are *linearly independent*

# Example

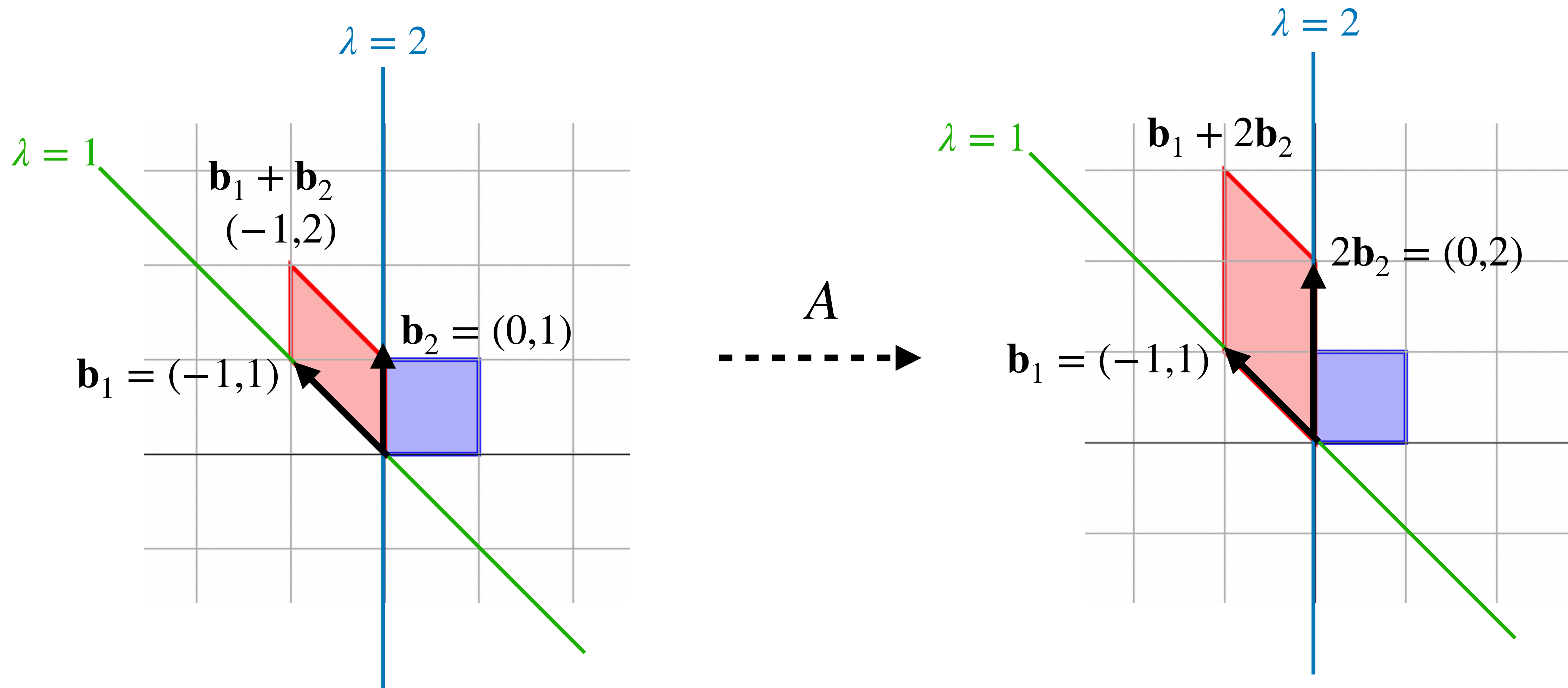
$$\begin{bmatrix} 2 & 0 \\ -1 & 1 \end{bmatrix}$$

*Find a diagonalization of the above matrix*

# The Picture

# Example (Geometric)

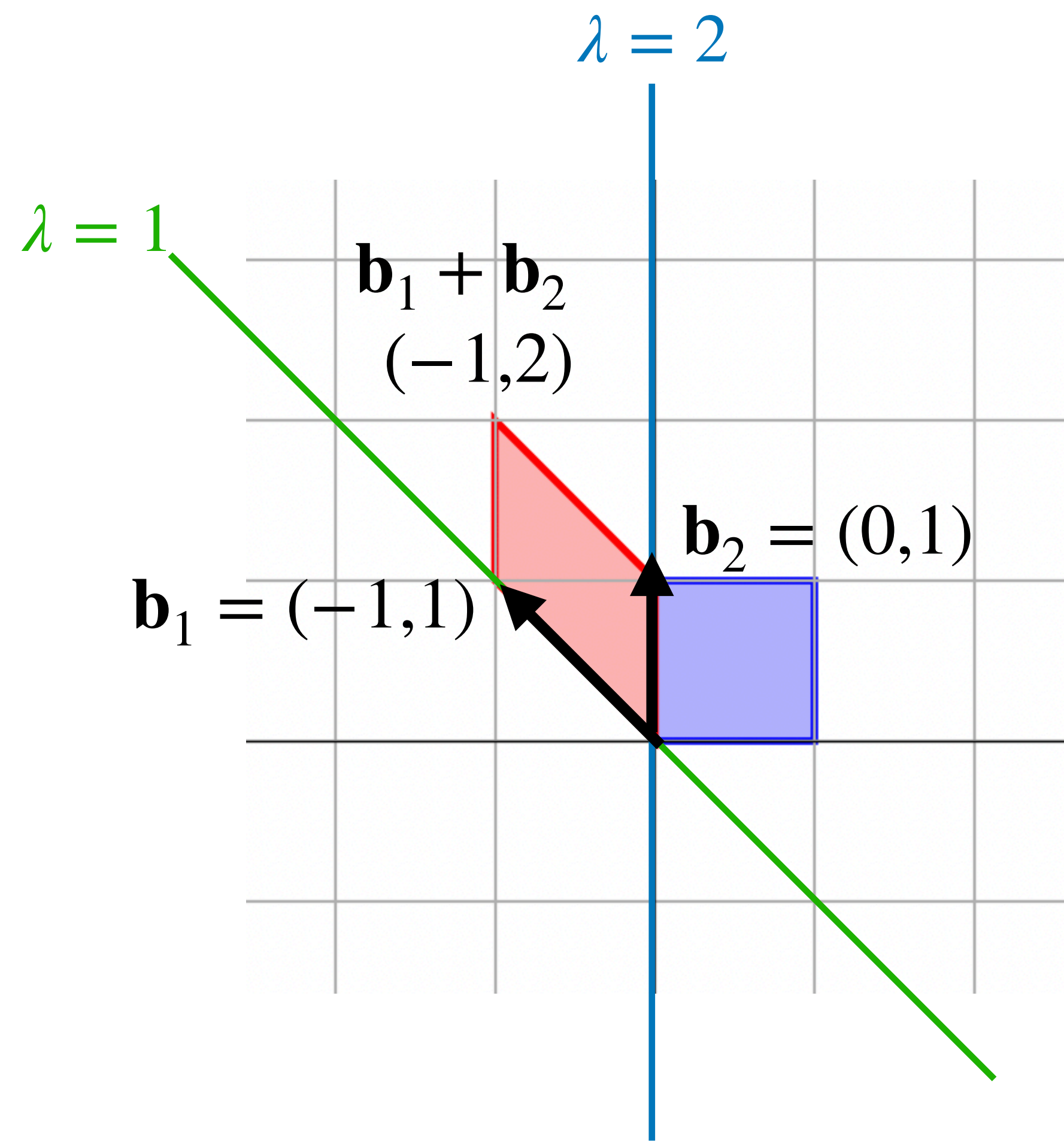
$$A = \begin{bmatrix} 2 & 0 \\ -1 & 1 \end{bmatrix}$$



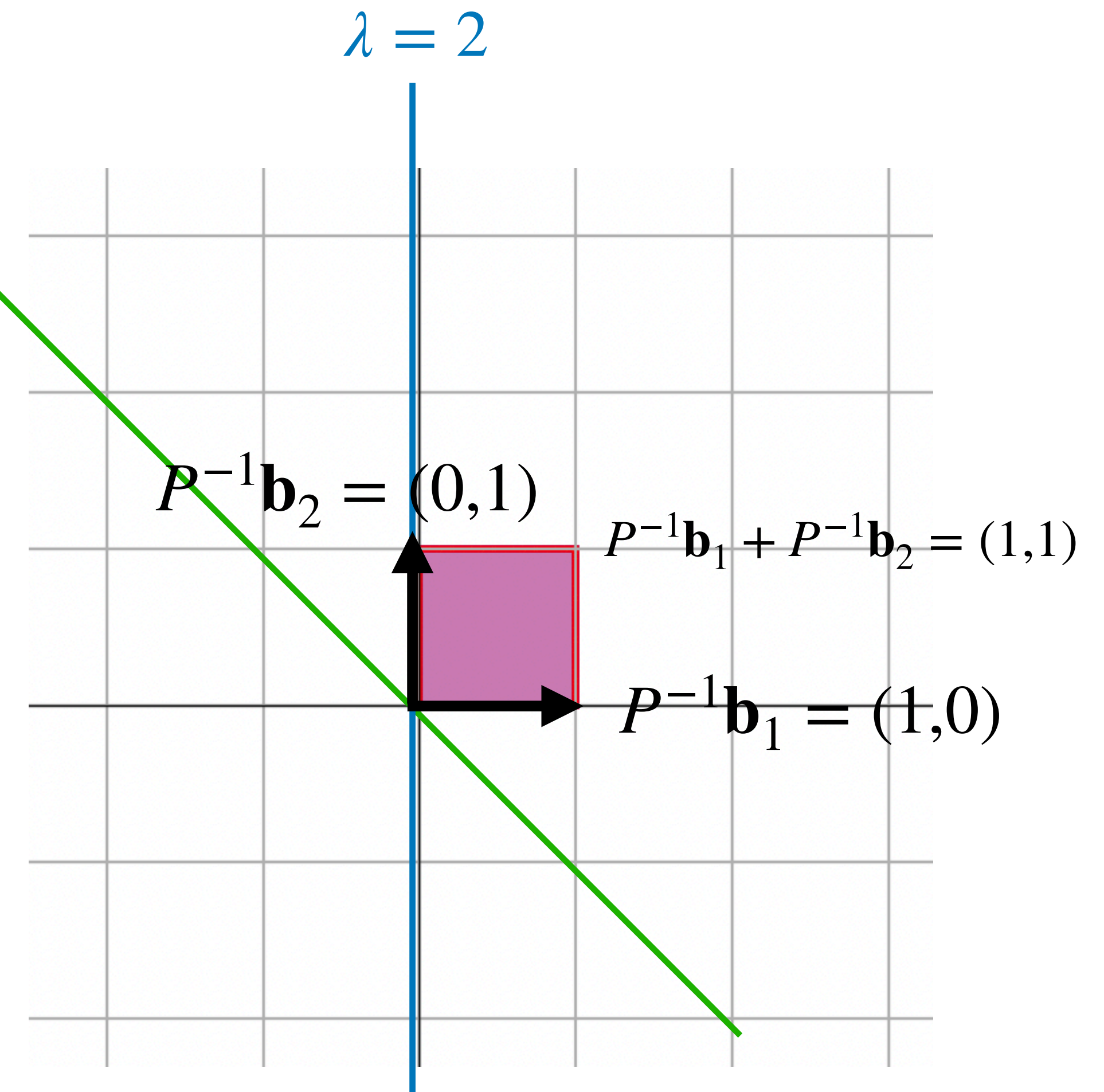


# Example (Geometric)

$$P^{-1} = \begin{bmatrix} -1 & 0 \\ 1 & 1 \end{bmatrix}^{-1}$$

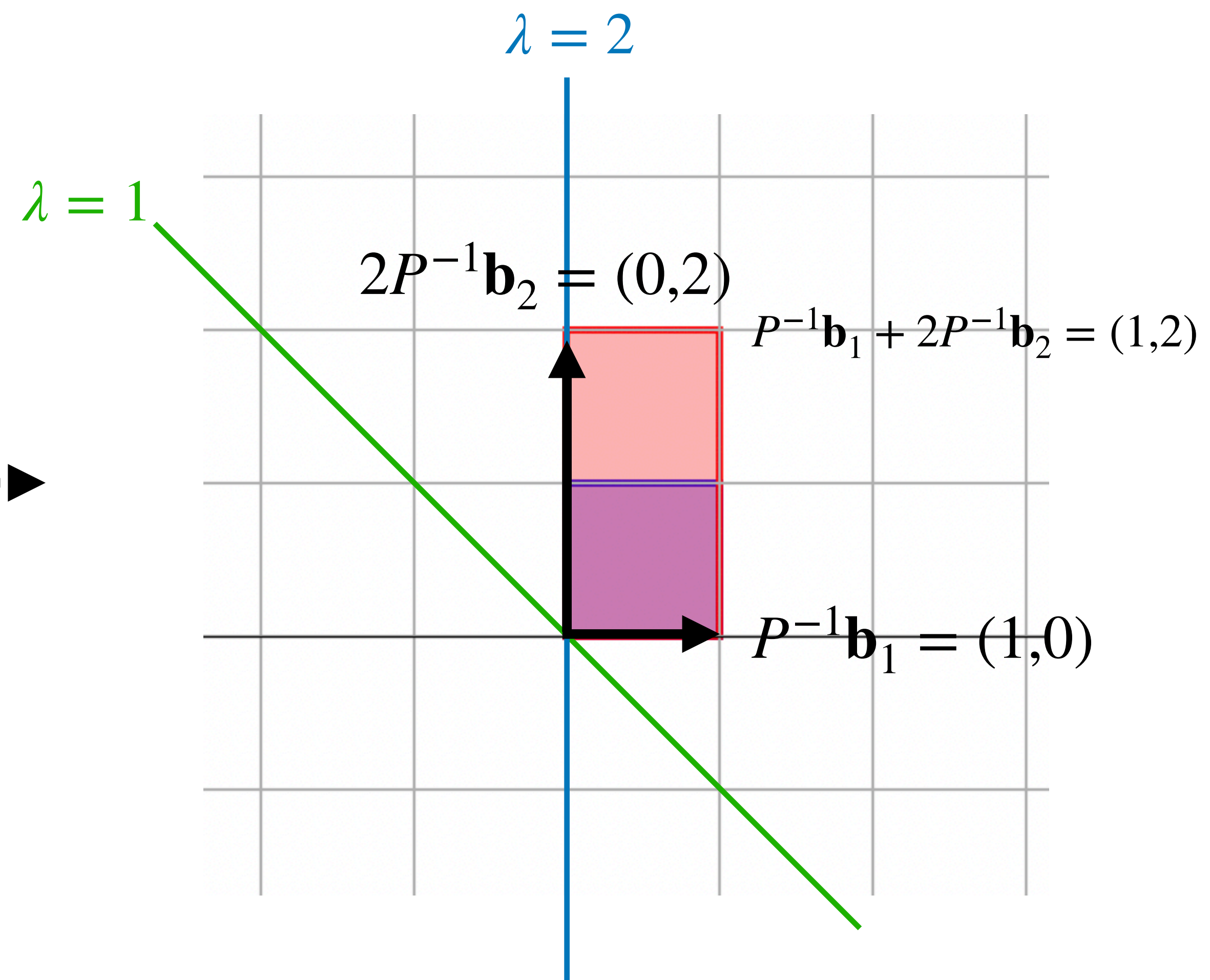
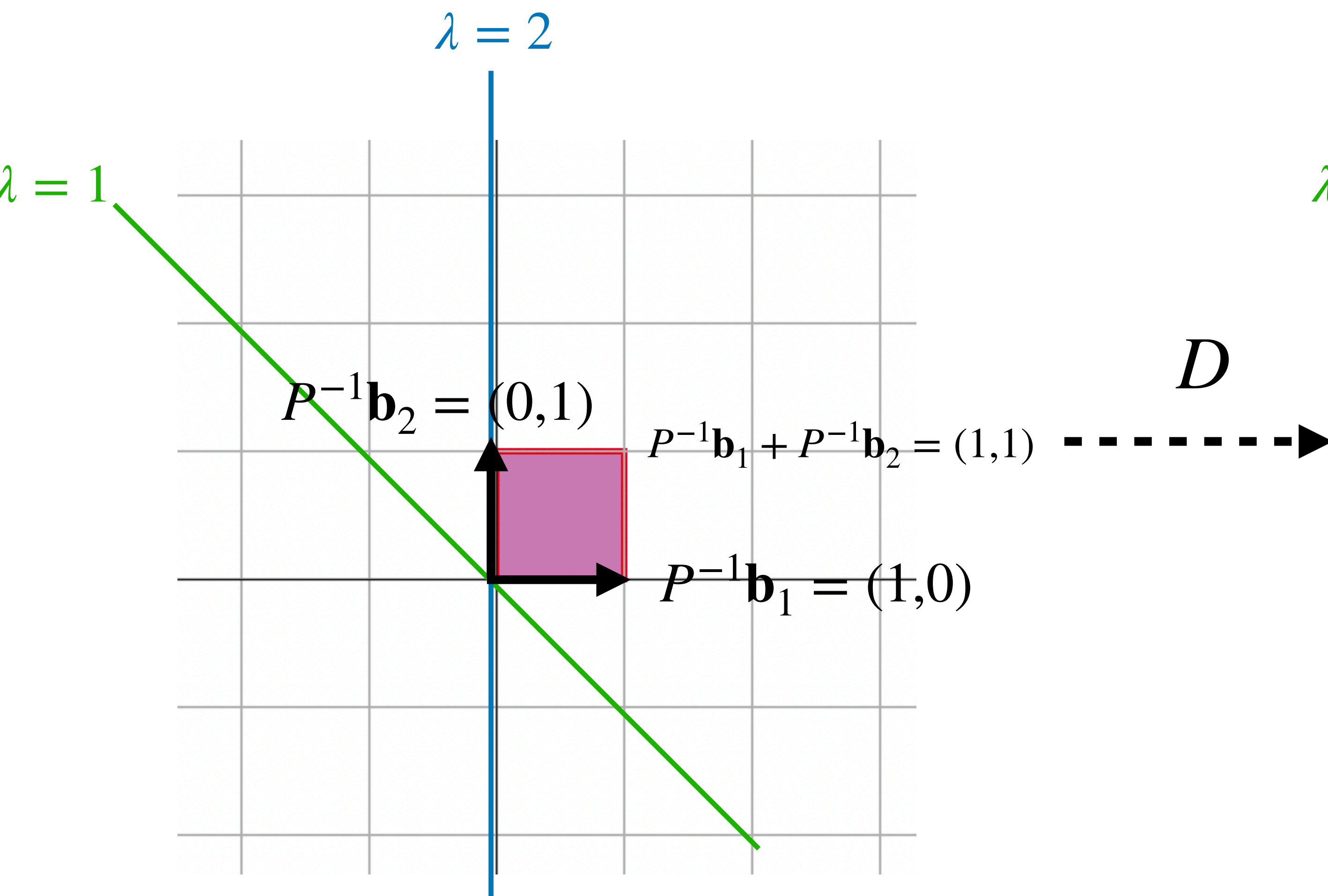


$P^{-1}$



# Example (Geometric)

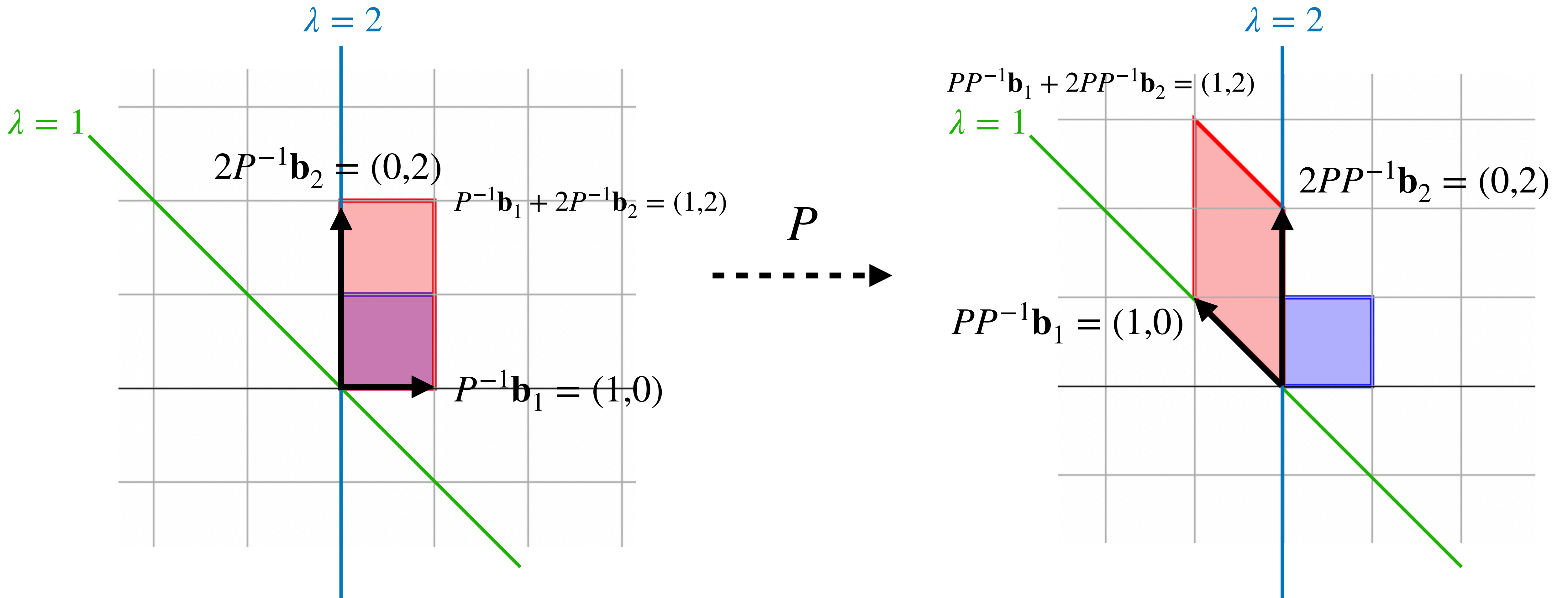
$$D = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$$





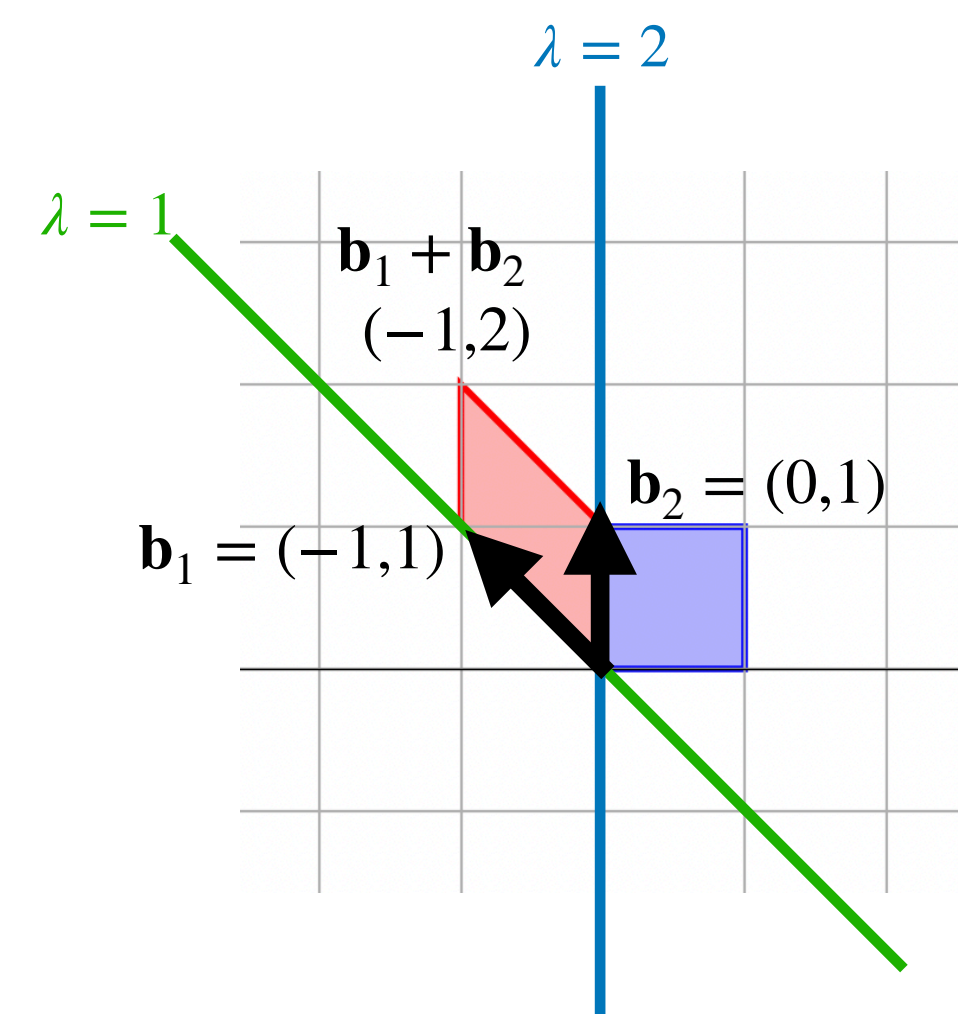
# Example (Geometric)

$$P = \begin{bmatrix} -1 & 0 \\ 1 & 1 \end{bmatrix}$$





# Example (Geometric)



$$A = PDP^{-1}$$

----->

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

