LINEAR ALGEBRA

LECTURE 7: EIGENDECOMPOSITION

goorm

KAIST AI
Graduate School of AI

DAVIAN
Data and Visual Analytics Lab

Lecture Overview

- Elements in linear algebra
- Linear system
- Linear combination, vector equation,
  Four views of matrix multiplication
- Linear independence, span, and subspace
- Linear transformation
- Least squares
- Eigendecomposition
- Advanced eigendecomposition
- Singular value decomposition

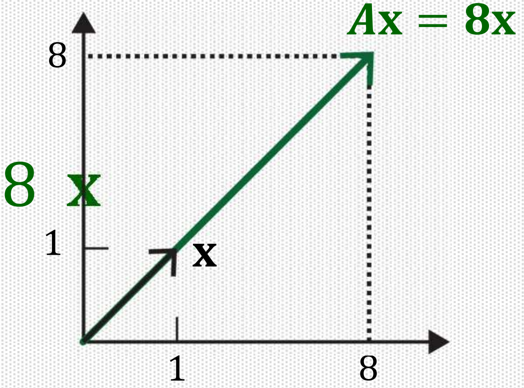
Eigenvectors and Eigenvalues

- **Definition**: An **eigenvector** of a **square** matrix $A \in \mathbb{R}^{n \times n}$ is a **nonzero** vector $\mathbf{x} \in \mathbb{R}^n$ such that $A\mathbf{x} = \lambda\mathbf{x}$ for some scalar λ
  In this case, λ is called an **eigenvalue** of A, and
  such an $\mathbf{x}$ is called an ***eigenvector corresponding to λ***.

Transformation Perspective

- Consider a linear transformation $T(\mathbf{x}) = A\mathbf{x}$.
- If $\mathbf{x}$ is an eigenvector, then $T(\mathbf{x}) = A\mathbf{x} = \lambda\mathbf{x}$, which means the output vector has **the same direction** as $\mathbf{x}$, but the length is scaled by a factor of λ.
- **Example:** For $A = \begin{bmatrix} 2 & 6 \\ 5 & 3 \end{bmatrix}$, an eigenvector is $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ since

$$T(\mathbf{x}) = A\mathbf{x} = \underbrace{\begin{bmatrix} 2 & 6 \\ 5 & 3 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} 1 \\ 1 \end{bmatrix}}_{\mathbf{x}} \underbrace{=}_{=} \begin{bmatrix} 8 \\ 8 \end{bmatrix} = \underbrace{8}_{8} \underbrace{\begin{bmatrix} 1 \\ 1 \end{bmatrix}}_{\mathbf{x}}$$



Computational Advantage

- Which computation is faster between $\begin{bmatrix} 2 & 6 \\ 5 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $8 \begin{bmatrix} 1 \\ 1 \end{bmatrix}$?

Eigenvectors and Eigenvalues

- The equation $A\mathbf{x} = \lambda\mathbf{x}$ can be re-written as

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

- λ is an eigenvalue of an $n \times n$ matrix A if and only if this equation has a **nontrivial** solution (since $\mathbf{x}$ should be a nonzero vector).

Eigenvectors and Eigenvalues

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

- The set of *all* solutions of the above equation is the **null space** of the matrix $(A - \lambda I)$, which we call the **eigenspace** of A **corresponding to λ**.
- The eigenspace consists of the zero vector and all the eigenvectors corresponding to λ, satisfying the above equation.

Example: Eigenvalues and Eigenvectors

- **Example**: Show that 8 is an eigenvalue of a matrix
  $A = \begin{bmatrix} 2 & 6 \\ 5 & 3 \end{bmatrix}$ and find the corresponding eigenvectors.

- **Solution**: The scalar 8 is an eigenvalue of A if and only if
  the equation $(A - 8I)\mathbf{x} = \mathbf{0}$ has a nontrivial solution:

$$(A - 8I)\mathbf{x} = \begin{bmatrix} -6 & 6 \\ 5 & -5 \end{bmatrix}\mathbf{x} = \mathbf{0}$$

- The solution is $\mathbf{x} = c\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ for any nonzero scalar c,
  which is Span $\left\{\begin{bmatrix} 1 \\ 1 \end{bmatrix}\right\}$.

Example: Eigenvalues and Eigenvectors

- In the previous example, -3 is also an eigenvalue:

$$(A+3I)\mathbf{x} = \begin{bmatrix} 5 & 6 \\ 5 & 6 \end{bmatrix}\mathbf{x} = \mathbf{0}$$

- The solution is $\mathbf{x} = c\begin{bmatrix} 1 \\ -5/6 \end{bmatrix}$ for any nonzero scalar c, which is Span $\left\{\begin{bmatrix} 1 \\ -5/6 \end{bmatrix}\right\}$.

Characteristic Equation

- How can we find the eigenvalues such as 8 and -3?
- If $(A - \lambda I)\mathbf{x} = \mathbf{0}$ has a nontrivial solution, then the columns of $(A - \lambda I)$ should be noninvertible.
- If it is invertible, $\mathbf{x}$ cannot be a nonzero vector since
$$(A - \lambda I)^{-1}(A - \lambda I)\mathbf{x} = (A - \lambda I)^{-1}\mathbf{0} \Rightarrow \mathbf{x} = \mathbf{0}$$
- Thus, we can obtain eigenvalues by solving
$$\det(A - \lambda I) = 0$$
called a **characteristic equation**.
- Also, the solution is not unique, and thus $A - \lambda I$ has linearly dependent columns.

Example: Characteristic Equation

- In the previous example, $A = \begin{bmatrix} 2 & 6 \\ 5 & 3 \end{bmatrix}$ is originally invertible since

$$\det(A) = \det \begin{bmatrix} 2 & 6 \\ 5 & 3 \end{bmatrix} = 6 - 30 = -24 \neq 0.$$

- By solving the characteristic equation, we want to find λ that makes $A - \lambda I$ non-invertible:

$$\begin{aligned} \det(A - \lambda I) &= \det \begin{bmatrix} 2-\lambda & 6 \\ 5 & 3-\lambda \end{bmatrix} \\ &= (2-\lambda)(3-\lambda) - 30 \\ &= -\lambda^2 - 5\lambda - 25 = (8-\lambda)(-3-\lambda) = 0 \end{aligned}$$

$\lambda = -3$ or 8

Null Space

- **Definition**: The **null space** of a matrix $A \in \mathbb{R}^{m \times n}$ is the set of all solutions of a homogeneous linear system, $A\mathbf{x} = \mathbf{0}$.
  We denote the null space of A as Nul A.

- For $A = \begin{bmatrix} \mathbf{a}_1^T \\ \mathbf{a}_2^T \\ \vdots \\ \mathbf{a}_m^T \end{bmatrix}$, $\mathbf{x}$ should satisfy the following:

$$\mathbf{a}_1^T\mathbf{x} = 0, \mathbf{a}_2^T\mathbf{x} = 0, \dots, \mathbf{a}_m^T\mathbf{x} = 0$$

- That is, $\mathbf{x}$ should be orthogonal to every row vector in A.

Null Space is a Subspace

- **Theorem**: The null space of a matrix $A \in \mathbb{R}^{m \times n}$ is a subspace of $\mathbb{R}^n$. In other words, the set of all the solutions of a system $A\mathbf{x} = \mathbf{0}$ is a subspace of $\mathbb{R}^n$.

- **Note**: An eigenspace thus have a set of basis vectors with a particular dimension.

Orthogonal Complement

- If a vector $\mathbf{z}$ is orthogonal to every vector in a subspace W of $\mathbb{R}^n$, then $\mathbf{z}$ is said to be **orthogonal to** W.
- The set of all vectors $\mathbf{z}$ that are orthogonal to W is called the **orthogonal complement** of W and is denoted by $W^\perp$ (and read as "W perpendicular" or simply "W perp").
- A vector $\mathbf{x} \in \mathbb{R}^n$ is in $W^\perp$ if and only if $\mathbf{x}$ is orthogonal to every vector in a set that spans W.
- $W^\perp$ is a subspace of $\mathbb{R}^n$.
- Nul $A = (\mathrm{Row}\, A)^\perp$.
- Likewise, Nul $A^T = (\mathrm{Col}\, A)^\perp$.

Fundamental Subspaces Given by A

- Nul $A = (\text{Row } A)^{\perp}$.
- Nul $A^T = (\text{Col } A)^{\perp}$.

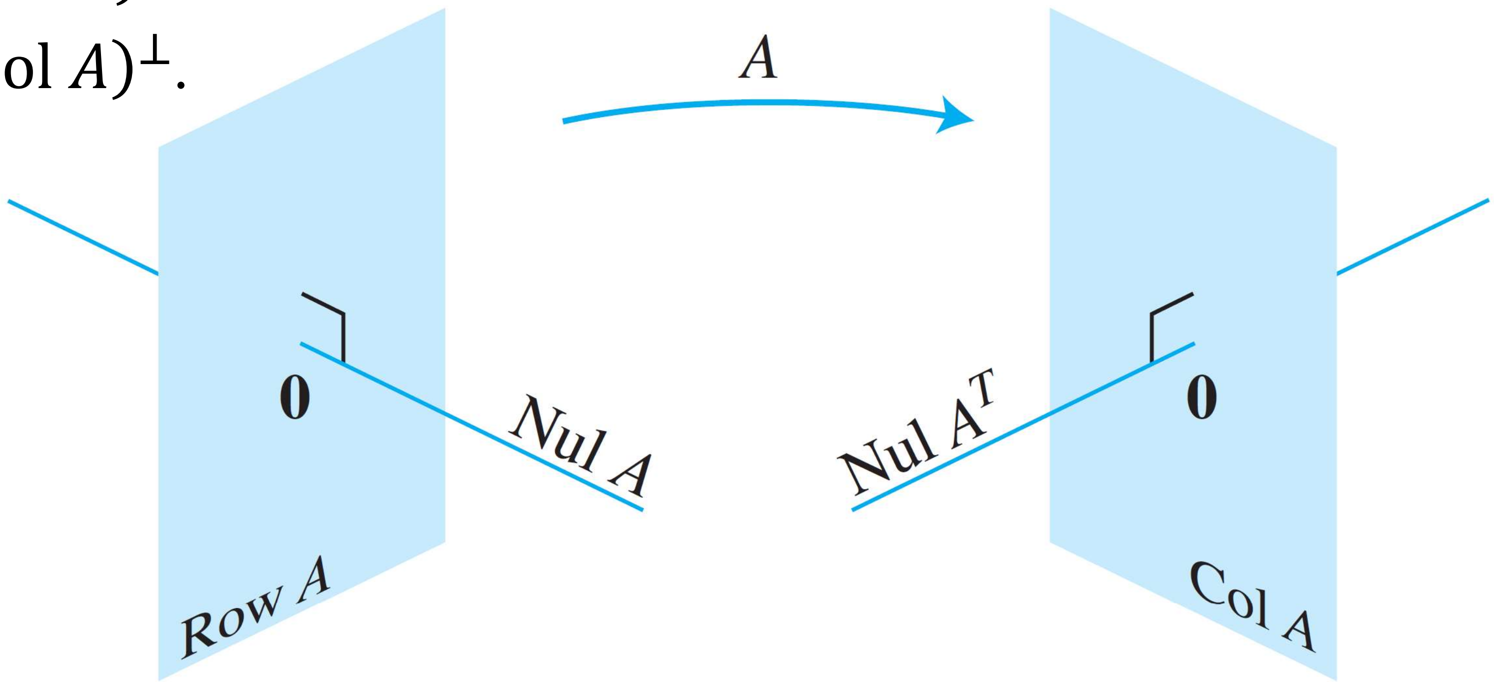
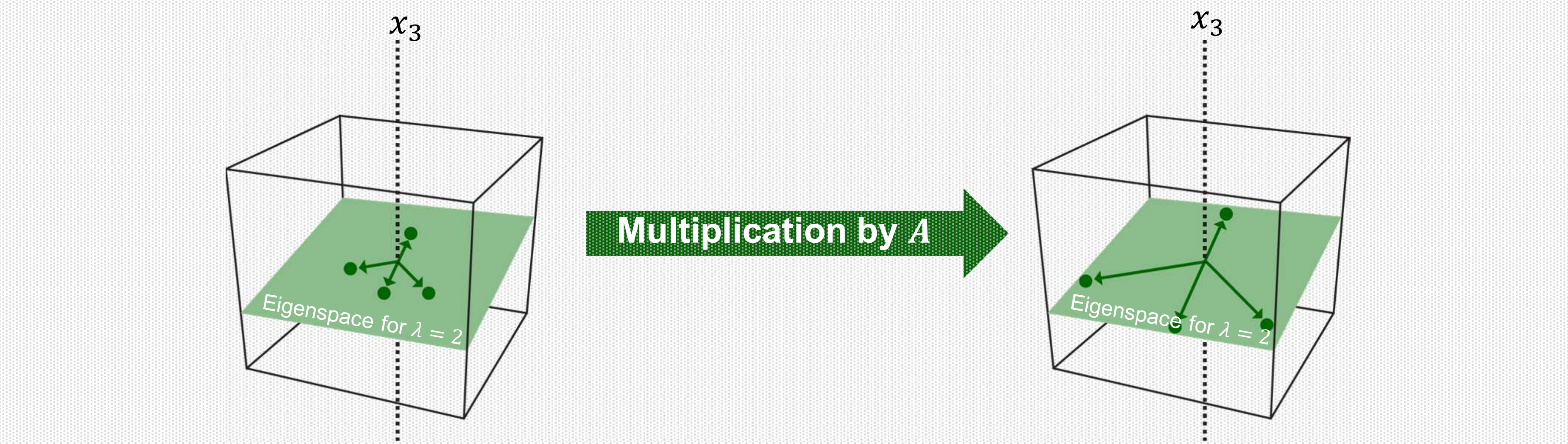


FIGURE 8 The fundamental subspaces determined by an $m \times n$ matrix A.

Eigenspace

- Note that the dimension of the eigenspace (corresponding to a particular λ) can be larger than one. In this case, any vector in the eigenspace satisfies

$$T(\mathbf{x}) = A\mathbf{x} = \lambda\mathbf{x}$$



A acts as a dilation on the eigenspace

Finding All Eigenvalues and Eigenvectors

- In summary, we can find all the possible eigenvalues and eigenvectors, as follows.
- First, find all the eigenvalue by solving the **characteristic equation**:

$$\det(A - \lambda I) = 0$$

- Second, for each eigenvalue λ, solve for $(A - \lambda I)\mathbf{x} = \mathbf{0}$ and obtain the set of basis vectors of the corresponding eigenspace.

Diagonalization

- We want to change a given square matrix $A \in \mathbb{R}^{n \times n}$ into a diagonal matrix via the following form:
$$D = V^{-1}AV$$
where $P \in \mathbb{R}^{n \times n}$ is an invertible matrix and $D \in \mathbb{R}^{n \times n}$ is a diagonal matrix. This is called a diagonalization of A.

- It is not always possible to diagonalize A. For A to be diagonalizable, an invertible V should exist such that $V^{-1}AV$ becomes a diagonal matrix.

Finding V and D

- How can we find an invertible P and the resulting diagonal matrix $D = V^{-1}AV$?
- $D = V^{-1}AV \Longrightarrow VD = AV$
- Let us represent the following:
- $V = [\mathbf{v}_1 \quad \mathbf{v}_2 \quad \cdots \quad \mathbf{v}_n]$ where $\mathbf{v}_i$'s are column vectors of V
- $D = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{bmatrix}$

Finding V and D

- $AV = A[\mathbf{v}_1 \quad \mathbf{v}_2 \quad \cdots \quad \mathbf{v}_n] = [A\mathbf{v}_1 \quad A\mathbf{v}_2 \quad \cdots \quad A\mathbf{v}_n]$
- $VD = [\mathbf{v}_1 \quad \mathbf{v}_2 \quad \cdots \quad \mathbf{v}_n]\begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{bmatrix}$

  $= [\lambda_1\mathbf{v}_1 \quad \lambda_2\mathbf{v}_2 \quad \cdots \quad \lambda_n\mathbf{v}_n]$
- $VD = AV \Leftrightarrow [A\mathbf{v}_1 \quad A\mathbf{v}_2 \quad \cdots \quad A\mathbf{v}_n] = [\lambda_1\mathbf{v}_1 \quad \lambda_2\mathbf{v}_2 \quad \cdots \quad \lambda_n\mathbf{v}_n]$

Finding V and D

- Equating columns, we obtain

$$A\mathbf{v}_1 = \lambda_1\mathbf{v}_1,\ A\mathbf{v}_2 = \lambda_2\mathbf{v}_2, \ldots,\ A\mathbf{v}_n = \lambda_n\mathbf{v}_n$$

- Thus, $\mathbf{v}_1$, $\mathbf{v}_2$, …, $\mathbf{v}_n$ should be eigenvectors and λ_1, λ_2, …, λ_n should be eigenvalues.

- Then, for $VD = AV \Longrightarrow D = V^{-1}AV$ to be true, V should invertible.

- In this case, the resulting diagonal matrix D has eigenvalues as diagonal entries.

Diagonalizable Matrix

- For V to be invertible,
  V should be a square matrix in $\mathbb{R}^{n \times n}$, and
  V should have n linearly independent columns.

- Recall columns of V are eigenvectors.
  Hence, A should have n linearly independent eigenvectors.

- It is not always the case, but if it is, A is diagonalizable.

Eigendecomposition

- If A is diagonalizable, we can write $D = V^{-1}AV$.

- We can also write $A = VDV^{-1}$.
  which we call eigendecomposition of A.

- A being diagonalizable is equivalent to
  A having eigendecomposition.

Linear Transformation via Eigendecomposition

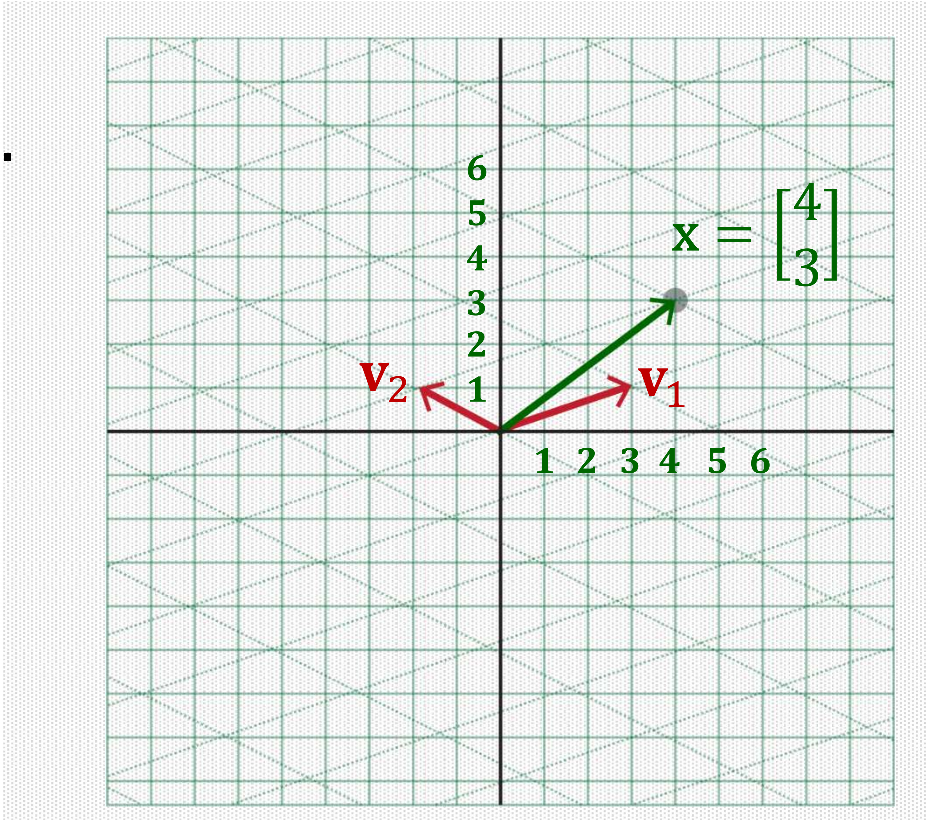
- Suppose A is diagonalizable, thus having eigendecompositio n

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

- Consider the linear transformation $T(\mathbf{x}) = A\mathbf{x}$.
- $T(\mathbf{x}) = A\mathbf{x} = VDV^{-1}\mathbf{x} = V\left(D(V^{-1}\mathbf{x})\right)$.

Change of Basis

- Suppose $A\mathbf{v}_1 = -1\mathbf{v}_1$ and $A\mathbf{v}_2 = 2\mathbf{v}_2$.
- $T(\mathbf{x}) = A\mathbf{x} = VDV^{-1}\mathbf{x} = V\left(D(V^{-1}\mathbf{x})\right)$
- Let $\mathbf{y} = V^{-1}\mathbf{x}$. Then,
$$V\mathbf{y} = \mathbf{x}$$
- $\mathbf{y}$ is a new coordinate of $\mathbf{x}$ with respect to a new basis of eigenvectors $\{\mathbf{v}_1, \mathbf{v}_2\}$.



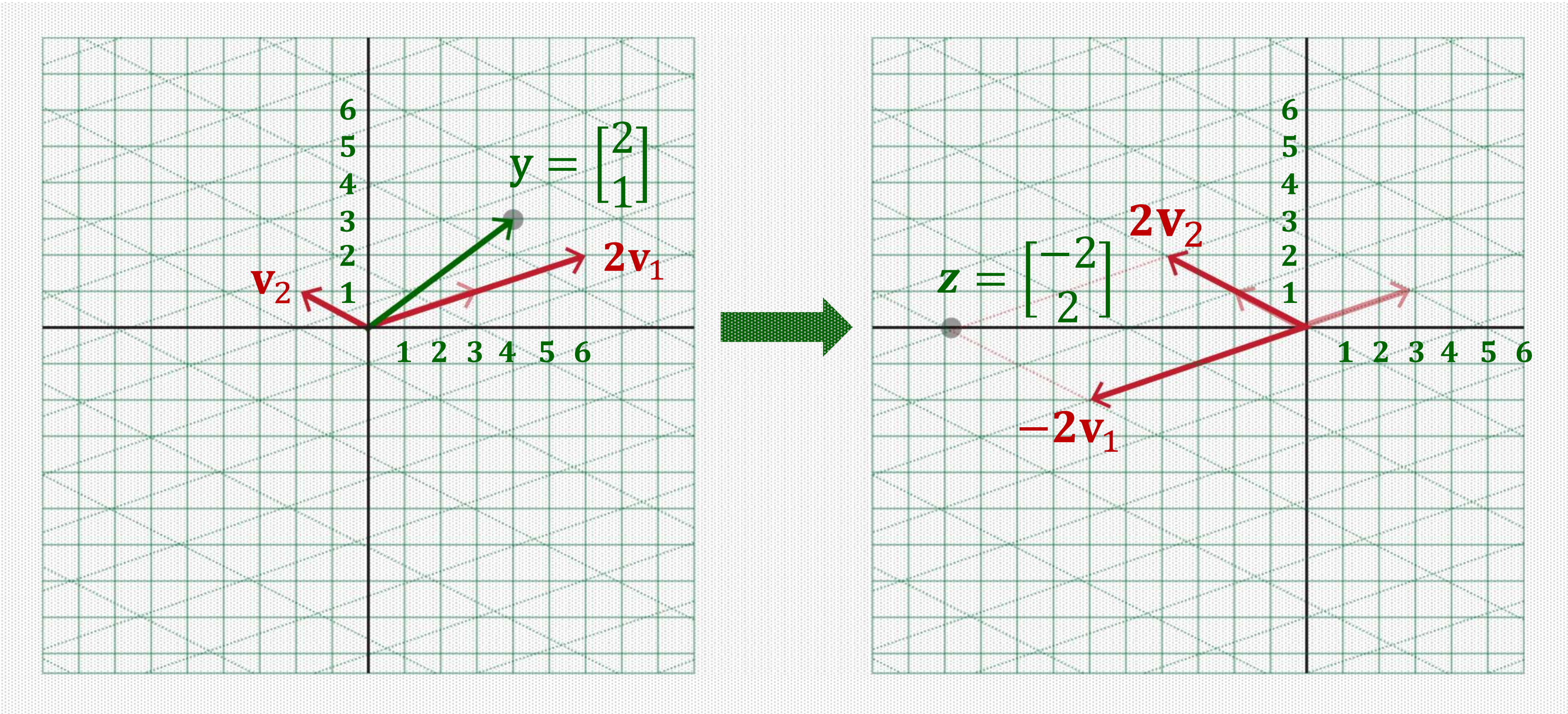
$$\mathbf{x} = \begin{bmatrix}4\\3\end{bmatrix} = 4\begin{bmatrix}1\\0\end{bmatrix} + 3\begin{bmatrix}0\\1\end{bmatrix} = \begin{bmatrix}1 & 0\\0 & 1\end{bmatrix}\begin{bmatrix}4\\3\end{bmatrix} = V\mathbf{y} = [\mathbf{v}_1 \quad \mathbf{v}_2]\begin{bmatrix}y_1\\y_2\end{bmatrix} = 2\mathbf{v}_1 + 1\mathbf{v}_2 \Rightarrow \mathbf{y} = \begin{bmatrix}2\\1\end{bmatrix}$$

Element-wise Scaling

- $T(\mathbf{x}) = V\left(D(P^{-1}\mathbf{x})\right) = V(D\mathbf{y})$
- Let $z = D\mathbf{y}$. This computation is a simple element-wise scaling of $\mathbf{y}$.
- **Example**: Suppose $D = \begin{bmatrix} -1 & 0 \\ 0 & 2 \end{bmatrix}$. Then

$$z = D\mathbf{y} = \begin{bmatrix} -1 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} (-1) \times 2 \\ 2 \times 1 \end{bmatrix} = \begin{bmatrix} -2 \\ 2 \end{bmatrix}$$

Dimension-wise Scaling



Back to Original Basis

- $T(\mathbf{x}) = V(D\mathbf{y}) = Vz$
- z is still a coordinate based on the new basis $\{\mathbf{v}_1, \mathbf{v}_2\}$.
- Vz converts z to another coordinates based on the original standard basis.
- That is, Vz is a linear combination of $\mathbf{v}_1$ and $\mathbf{v}_2$ using the coefficient vector z.
- That is,

$$Vz = [\mathbf{v}_1 \quad \mathbf{v}_2]\begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = \mathbf{v}_1 z_1 + \mathbf{v}_2 z_2$$

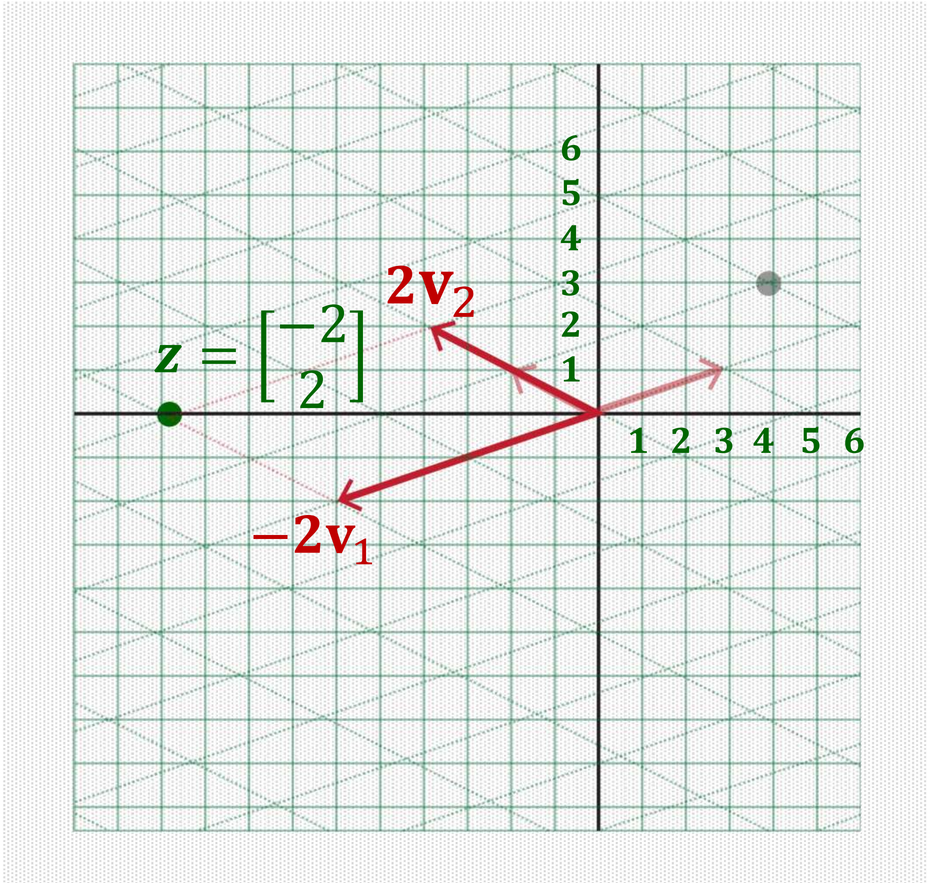
Back to Original Basis

- $T(\mathbf{x}) = V\color{red}{z} = [\mathbf{v}_1 \quad \mathbf{v}_2]\begin{bmatrix}-2\\2\end{bmatrix}$

$$= -2\mathbf{v}_1 + 2\mathbf{v}_2$$

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

$$= \begin{bmatrix}-10\\0\end{bmatrix}$$



Overview of Transformation using Eigendecomposition

$$
\begin{tikzcd}
x \arrow[r, "\boldsymbol{A}"] \arrow[d, "\boldsymbol{V^{-1}}"'] & T(x) = Ax = V\big(D(V^{-1}x)\big) \\
y \arrow[r, "\boldsymbol{D}"] & z \arrow[u, "\boldsymbol{V}"']
\end{tikzcd}
$$

Linear Transformation via A^k

- Now, consider recursive transformation $A \times A \times \cdots \times A\mathbf{x} = A^k \mathbf{x}$.
- If A is diagonalizable, A has eigendecomposition

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

- $A^k = (VDV^{-1})(VDV^{-1})\cdots(VDV^{-1}) = VD^kV^{-1}$
- D^k is simply computed as

$$D^k = \begin{bmatrix} \lambda_1^k & 0 & \cdots & 0 \\ 0 & \lambda_2^k & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n^k \end{bmatrix}$$

Linear Transformation via A^k

- $A^k\mathbf{x} = VD^kV^{-1}\mathbf{x}$ can be computed in the similar manner to the previous example.
- It is much faster to compute $V\left(D^k(V^{-1}\mathbf{x})\right)$ than to compute $A^k\mathbf{x}$.