

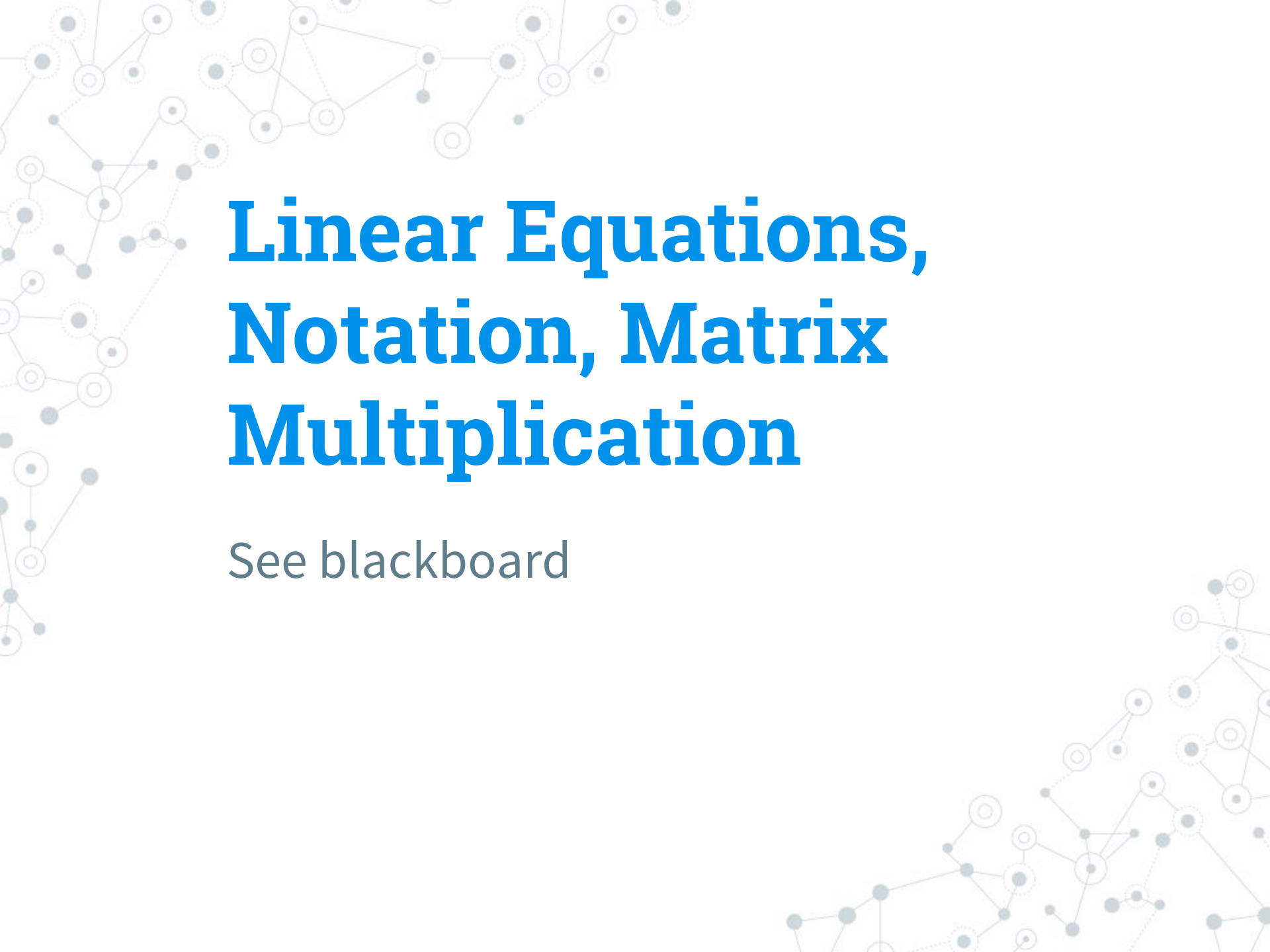


# Linear Algebra Review

Jeremy Irvin and  
Daniel Spokoyny

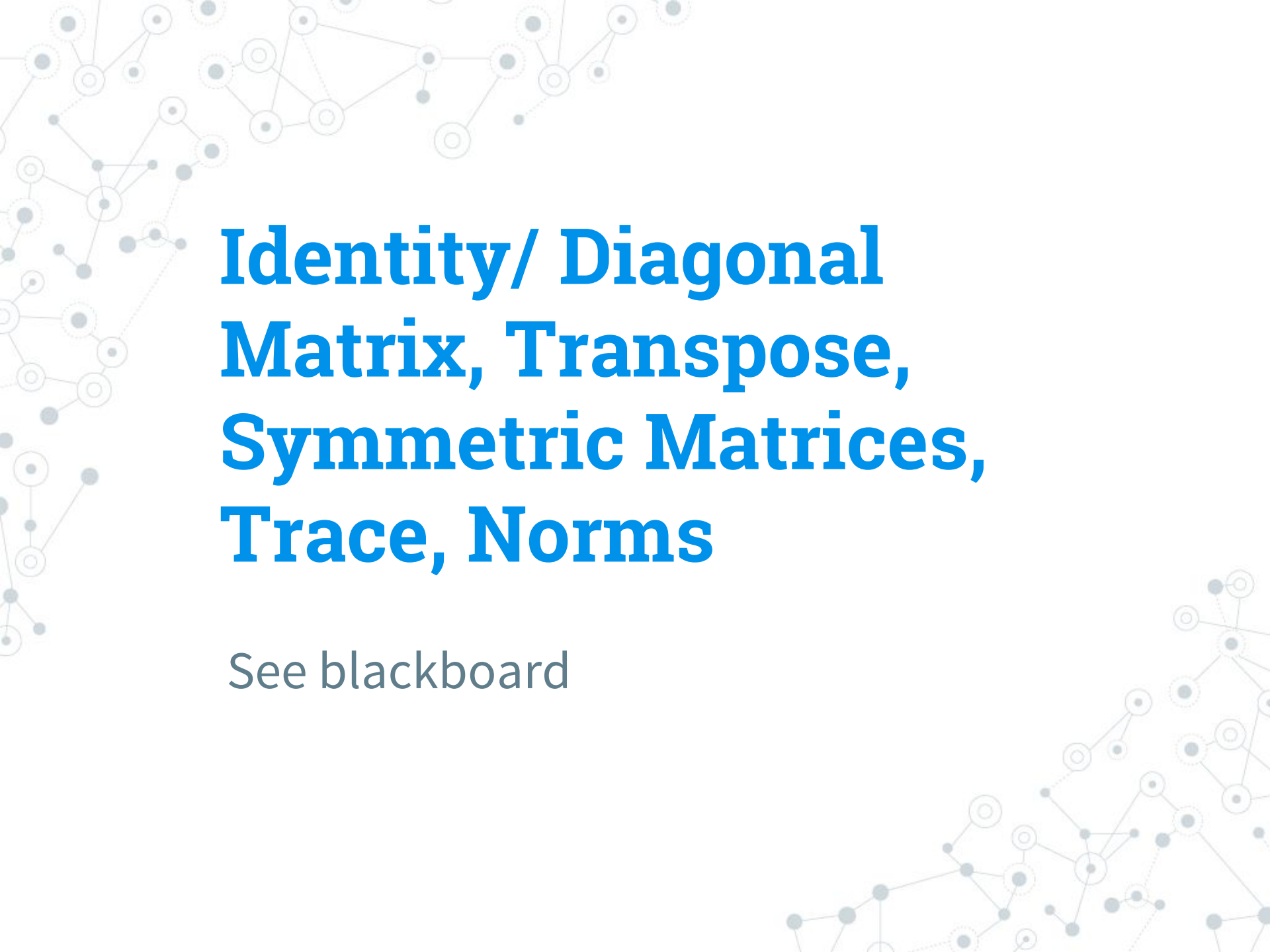
Created from Kolter's Linear  
Algebra Review for Stanford  
CS229



A decorative background featuring a network diagram with nodes and connecting lines, primarily located in the top-left and bottom-right corners. The nodes are represented by circles of varying sizes, some with concentric rings, and the lines are thin and grey.

# **Linear Equations, Notation, Matrix Multiplication**

See blackboard



# **Identity/ Diagonal Matrix, Transpose, Symmetric Matrices, Trace, Norms**

See blackboard

# Linear Independence

- A set of vectors  $\{x_1, x_2, \dots, x_n\} \subset \mathbb{R}^m$  is linearly independent if no vector can be written as a linear combination of the other vectors.

# Linear Independence

- Conversely, the set is linearly dependent if *one vector can* be written as a linear combination of the remaining vectors. i.e. if

$$x_n = \sum_{i=1}^{n-1} \alpha_i x_i$$

for some scalar values  $\alpha_1, \dots, \alpha_{n-1} \in \mathbb{R}$ .

- Example on board

# Rank

- Let  $A \in \mathbb{R}^{m \times n}$ ,  $B \in \mathbb{R}^{n \times p}$ .
  - The column rank of  $A$  is the number of linearly independent columns of  $A$ .
  - The row rank of a  $A$  is the number of linearly independent rows of  $A$ .
  - The row rank of  $A$  equals the column rank of  $A$ , and is called the rank of  $A$ .

# Properties of Rank

- Let  $A \in \mathbb{R}^{m \times n}$ ,  $B \in \mathbb{R}^{n \times p}$ .
  - $\text{rank}(A) \leq \min(m, n)$
  - If  $\text{rank}(A) = \min(m, n)$ , A has full rank.
  - $\text{rank}(A) = \text{rank}(A^T)$
  - $\text{rank}(AB) \leq \min(\text{rank}(A), \text{rank}(B))$
  - $\text{rank}(A+B) \leq \text{rank}(A) + \text{rank}(B)$

# Inverse

- The inverse of a square matrix  $A \in \mathbb{R}^{n \times n}$  is denoted by  $A^{-1}$  and is the unique matrix such that

$$A^{-1}A = I = AA^{-1}$$

- Not all matrices have inverses.
- $A$  is invertible (or non-singular) if  $A^{-1}$  exists and is called non-invertible or singular otherwise.



# Inverse

- $A$  is invertible iff it is square and has full rank.
- $(A^{-1})^{-1} = A$
- $(AB)^{-1} = B^{-1}A^{-1}$
- $(A^{-1})^T = (A^T)^{-1}$
- $A$  is non-singular iff  $Ax=b$  has a unique solution for any  $b$  ( $x=A^{-1}b$ ).

# Orthogonality and Normalization

- Recall two vectors  $x, y \in \mathbb{R}^n$  are orthogonal if

$$x^T y = 0$$

- A vector  $x \in \mathbb{R}^n$  is normalized if

$$\|x\|_2 = 1$$

# Orthogonal Matrices

- A square matrix  $U \in \mathbb{R}^{n \times n}$  is orthogonal if all of its columns are orthogonal to each other and are normalized, or equivalently,

$$U^T U = I = U U^T$$

- Orthogonal matrices are norm (or distance) preserving, ie,

$$\|Ux\|_2 = \|x\|_2$$

for any  $x \in \mathbb{R}^n$ ,  $U \in \mathbb{R}^{n \times n}$ .

# Span and Basis

- The span of a set of vectors  $\{x_1, x_2, \dots, x_n\} \subset \mathbb{R}^m$  is the set of all vectors that can be expressed as a linear combination of  $\{x_1, \dots, x_n\}$ , ie,

$$\text{span}(\{x_1, \dots, x_n\}) = \left\{ v : v = \sum_{i=1}^n \alpha_i x_i, \alpha_i \in \mathbb{R} \right\}$$

- If  $\{x_1, \dots, x_n\}$  is a set of  $n$  linear independent vectors then

$$\text{span}(\{x_1, \dots, x_n\}) = \mathbb{R}^n$$

- $\{x_1, \dots, x_n\}$  is called a basis of  $\mathbb{R}^n$ .

# Range and Nullspace

- The range of  $A \in \mathbb{R}^{m \times n}$  is the span of the columns of  $A$ , ie,

$$\mathcal{R}(A) = \{v \in \mathbb{R}^m : v = Ax, x \in \mathbb{R}^n\}$$

- The nullspace of a matrix  $A \in \mathbb{R}^{m \times n}$  is the set of all vectors that equal 0 when multiplied by  $A$ , ie,

$$\mathcal{N}(A) = \{x \in \mathbb{R}^n : Ax = 0\}$$

# Range and Nullspace

- The range of  $A^T$  and the nullspace of  $A$  are orthogonal complements in  $\mathbb{R}^n$ .
- This means they are
  - disjoint subsets which span the entire space, and
  - every vector in the range of  $A^T$  is orthogonal to every vector in the nullspace of  $A$ .

# Projection

- The projection of a vector  $y \in \mathbb{R}^m$  onto the span of  $\{x_1, \dots, x_n\}$  (with  $x_i \in \mathbb{R}^m$ ) is the vector  $v \in \text{span}(\{x_1, \dots, x_n\})$  closest to  $y$  as measured by Euclidean Distance (ie  $\|v - y\|_2$  ).
- Formally,

$$\text{Proj}(y; \{x_1, \dots, x_n\}) = \underset{v \in \text{span}(\{x_1, \dots, x_n\})}{\text{argmin}} \|y - v\|_2$$

# Projection

- Assuming  $A$  is full rank and  $n < m$ , the projection of  $y \in \mathbb{R}^m$  onto the range of  $A$  is

$$\text{Proj}(y; A) = \underset{v \in \mathcal{R}(A)}{\text{argmin}} ||y - v||_2 = A(A^T A)^{-1} A^T y$$

- When  $A$  contains a single column, ie,  $A = a \in \mathbb{R}^m$ , we see the ‘familiar’ case of a projection of a vector onto a line:

$$\text{Proj}(y; a) = \frac{aa^T}{a^T a} y$$



# Determinant

- The determinant of a square matrix is a function  $\det : \mathbb{R}^{n \times n} \rightarrow \mathbb{R}$ , denoted  $|A|$  or  $\det A$ .
- Let  $a_1, \dots, a_n \in \mathbb{R}^n$  denote the columns of  $A$ . Consider  $S$  the the restriction of  $\text{span}(\{a_1, \dots, a_n\})$  to linear combinations whose coefficients satisfy

$$0 \leq \alpha_i \leq 1, i = 1, \dots, n.$$

# Determinant

- The absolute value of the determinant of  $A$  is a measure of the ‘volume’ of  $S$ .
- For example, for  $2 \times 2$  matrices  $A$ ,  $|\det A|$  is a measure of the area enclosed by the parallelogram with edges the columns of  $A$ .

# Properties of Determinant

- $|I| = 1$ .
- If  $B$  is formed by multiplying a single row of  $A$  by a real number  $t$ , then

$$|B| = t|A|.$$

- If  $B$  is formed by switching any two rows of  $A$ , then

$$|B| = -|A|.$$

# Properties of Determinant

- $|A| = |A^T|$ .
- $|AB| = |A| |B|$
- $|A| = 0$  if and only if  $A$  is singular (non-invertible).
- If  $A$  is invertible, then  $|A^{-1}| = 1/|A|$ .

# Eigenvalues and Eigenvectors

- Given a square matrix  $A \in \mathbb{R}^{n \times n}$ , we say  $\lambda \in \mathbb{C}$  is an eigenvalue of  $A$  and  $x \in \mathbb{C}^n$  is the corresponding eigenvector if

$$Ax = \lambda x, \quad x \neq 0.$$

- Intuitively, this means multiplying  $A$  by  $x$  results in a new vector in the same direction as  $x$  but scaled by  $\lambda$ .

# Eigenvalues and Eigenvectors

- Note that if  $x$  is an eigenvector, then  $cx$  is an eigenvector for any complex  $c$ .

- We can rewrite the equation above as

$$(\lambda I - A)x = 0, \quad x \neq 0.$$

- $(\lambda I - A)x = 0$  has a non-zero solution iff  $(\lambda I - A)$  has a non-empty nullspace, which only happens if  $(\lambda I - A)$  is singular, ie,

$$|(\lambda I - A)| = 0.$$

# E-values and E-vectors Properties

- The trace of  $A$  is equal to the sum of its eigenvalues,

$$\text{tr} A = \sum_{i=1}^n \lambda_i.$$

- The determinant of  $A$  is equal to the product of its eigenvalues

$$|A| = \prod_{i=1}^n \lambda_i.$$

- The rank of  $A$  is equal to the number of non-zero eigenvalues of  $A$ .
- The eigenvalues of a diagonal matrix are just the diagonal entries.

# Diagonalization

- We can write all the eigenvector equations simultaneously as

$$AX = X\Lambda.$$

with  $X \in \mathbb{R}^{n \times n}$  the eigenvectors of  $A$  and a diagonal matrix  $\Lambda$  whose entries are the eigenvalues  $A$ , ie,

$$X \in \mathbb{R}^{n \times n} = \begin{bmatrix} | & | & \cdots & | \\ x_1 & x_2 & \cdots & x_n \\ | & | & \cdots & | \end{bmatrix}, \quad \Lambda = \text{diag}(\lambda_1, \dots, \lambda_n).$$



# Diagonalization

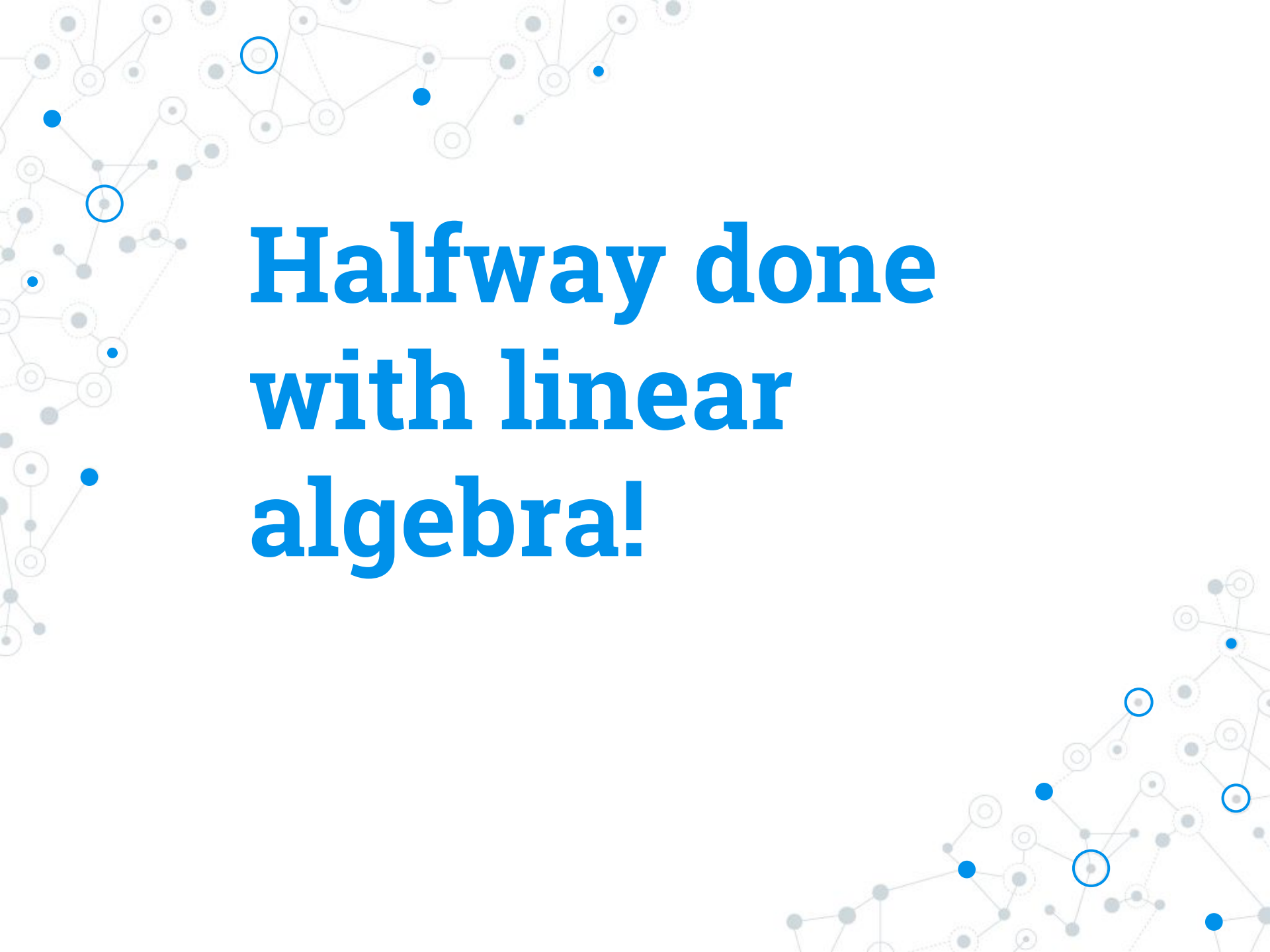
- If the eigenvectors of  $A$  are linearly independent, then  $X$  will be invertible, so

$$A = X\Lambda X^{-1}.$$

We say that  $A$  is diagonalizable.

# What Just Happened?

- Linear Equations
- Matrices / Special Matrices
- Subspaces and Projection
- Eigenvalues and Eigenvectors



**Halfway done  
with linear  
algebra!**