# 10.34: Numerical Methods Applied to Chemical Engineering

Lecture 2:
More basics of linear algebra
Matrix norms,
Condition number

# Recap

- Numerical error
- Scalars, vectors, and matrices
  - Operations
  - Properties

# Recap

- Vectors:
  - What mathematical object is the equivalent of an infinite dimensional vector?

# Scalars, Vectors and Matrices

- Vectors:
  - What mathematical object is the equivalent of an infinite dimensional vector?
    - A function.

# Scalars, Vectors and Matrices

- Matrices:
  - Ordered sets of numbers:

$$\mathbf{A} = \begin{pmatrix} A_{11} & A_{12} & \dots & A_{1M} \\ A_{21} & A_{22} & \dots & A_{2M} \\ \vdots & \vdots & \ddots & \vdots \\ A_{N1} & A_{N2} & \dots & A_{NM} \end{pmatrix}$$

  - Set of all real matrices with N rows and M columns, $\mathbb{R}^{N \times M}$
  - Addition: $\mathbf{C} = \mathbf{A} + \mathbf{B} \Rightarrow C_{ij} = A_{ij} + B_{ij}$
  - Multiplication by scalar: $\mathbf{C} = c\mathbf{A} \Rightarrow C_{ij} = cA_{ij}$
  - Transpose: $\mathbf{C} = \mathbf{A}^T \Rightarrow C_{ij} = A_{ji}$
  - Trace (square matrices):

$$\mathrm{Tr}\,\mathbf{A} = \sum_{i=1}^{N} A_{ii}$$

# Scalars, Vectors and Matrices

- Matrices:
  - Matrix-vector product: $\mathbf{y} = \mathbf{A}\mathbf{x} \Rightarrow y_i = \sum_{j=1}^{M} A_{ij} x_j$
  - Matrix-matrix product: $\mathbf{C} = \mathbf{A}\mathbf{B} \Rightarrow C_{ij} = \sum_{k=1}^{M} A_{ik} B_{kj}$
    - Properties:
      - no commutation in general: $\mathbf{AB} \neq \mathbf{BA}$
      - association: $\mathbf{A}(\mathbf{BC}) = (\mathbf{AB})\mathbf{C}$
      - distribution: $\mathbf{A}(\mathbf{B} + \mathbf{C}) = \mathbf{AB} + \mathbf{AC}$
      - transposition: $(\mathbf{AB})^T = \mathbf{B}^T \mathbf{A}^T$
      - inversion: $\mathbf{A}^{-1}\mathbf{A} = \mathbf{A}\mathbf{A}^{-1} = \mathbf{I}$ if $\det(\mathbf{A}) \neq 0$

# Scalars, Vectors and Matrices

- Matrices:
  - Matrix-matrix product:
    - Vectors are matrices too:
      - $\mathbf{x} \in \mathbb{R}^N \quad \mathbf{x} \in \mathbb{R}^{N \times 1}$
      - $\mathbf{y}^T \in \mathbb{R}^N \quad \mathbf{y}^T \in \mathbb{R}^{1 \times N}$
    - What is: $\mathbf{y}^T \mathbf{x}$ ?

$$\mathbf{C} = \mathbf{AB} \Rightarrow C_{ij} = \sum_{k=1}^{M} A_{ik} B_{kj}$$

# Scalars, Vectors and Matrices

- Matrices:
  - Matrix-matrix product:
    - Vectors are matrices too:
      - $\mathbf{x} \in \mathbb{R}^N \quad \mathbf{x} \in \mathbb{R}^{N \times 1}$
      - $\mathbf{y}^T \in \mathbb{R}^N \quad \mathbf{y}^T \in \mathbb{R}^{1 \times N}$
      - What is: $\mathbf{y}^T \mathbf{x}$ ?

# Scalars, Vectors and Matrices

- Matrices:
  - Examples: $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{N \times N}$ $\mathbf{x} \in \mathbb{R}^N$
    - How many operations to compute:
      - $\mathbf{Ax}$
      - $\mathbf{AB}$
      - $\mathbf{ABx}$
    - What is $\mathbf{x}^T\mathbf{ABx}$?
    - What is $\mathbf{ABxx}^T$?

# Scalars, Vectors and Matrices

- Matrices:
  - Dyadic product: $\mathbf{A} = \mathbf{x}\mathbf{y}^T = \mathbf{x} \otimes \mathbf{y} \Rightarrow A_{ij} = x_i y_j$
  - Determinant (square matrices only):

$$\det(\mathbf{A}) = \sum_{j=1}^{N} (-1)^{i+j} A_{ij} M_{ij}(\mathbf{A})$$

$$M_{ij}(\boldsymbol{A}) = \det \begin{pmatrix} A_{11} & A_{12} & \dots & A_{1(j-1)} & A_{1(j+1)} & \dots & A_{1N} \\ A_{21} & A_{22} & \dots & A_{2(j-1)} & A_{2(j+1)} & \dots & A_{2N} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ A_{(i-1)1} & A_{(j-1)2} & \dots & A_{(i-1)(j-1)} & A_{(i-1)(j+1)} & \dots & A_{(i-1)N} \\ A_{(i+1)1} & A_{(j+1)2} & \dots & A_{(i+1)(j-1)} & A_{(i+1)(j+1)} & \dots & A_{(i+1)N} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ A_{N1} & A_{N2} & \dots & A_{N(j-1)} & A_{N(j+1)} & \dots & A_{NN} \end{pmatrix}$$

- $\det(c) = c$

# Scalars, Vectors and Matrices

- Matrices:
  - Determinant (square matrices only):

$$\det(\mathbf{A}) = \sum_{j=1}^{N} (-1)^{i+j} A_{ij} M_{ij}(\mathbf{A})$$

  - Properties:
    - If any row or column is zeros, $\det(\mathbf{A}) = 0$
    - If any row or column is multiplied by $a$

$$\det(\mathbf{A}_1^c \; \mathbf{A}_2^c \; a\mathbf{A}_3^c \; \ldots \; \mathbf{A}_N^c) = a \det(\mathbf{A})$$

    - Swapping any row or column changes the sign
    - $\det(\mathbf{A}^T) = \det(\mathbf{A})$
    - $\det(\mathbf{AB}) = \det(\mathbf{A}) \det(\mathbf{B})$

# Scalars, Vectors and Matrices

- Matrices:
  - Example:

$$\mathbf{A} = \begin{pmatrix} -2 & 1 \\ 1 & -2 \end{pmatrix}$$

    - Calculate: $\det(\mathbf{A})$
    - How many operations to compute $\det(\mathbf{A})$ in general?

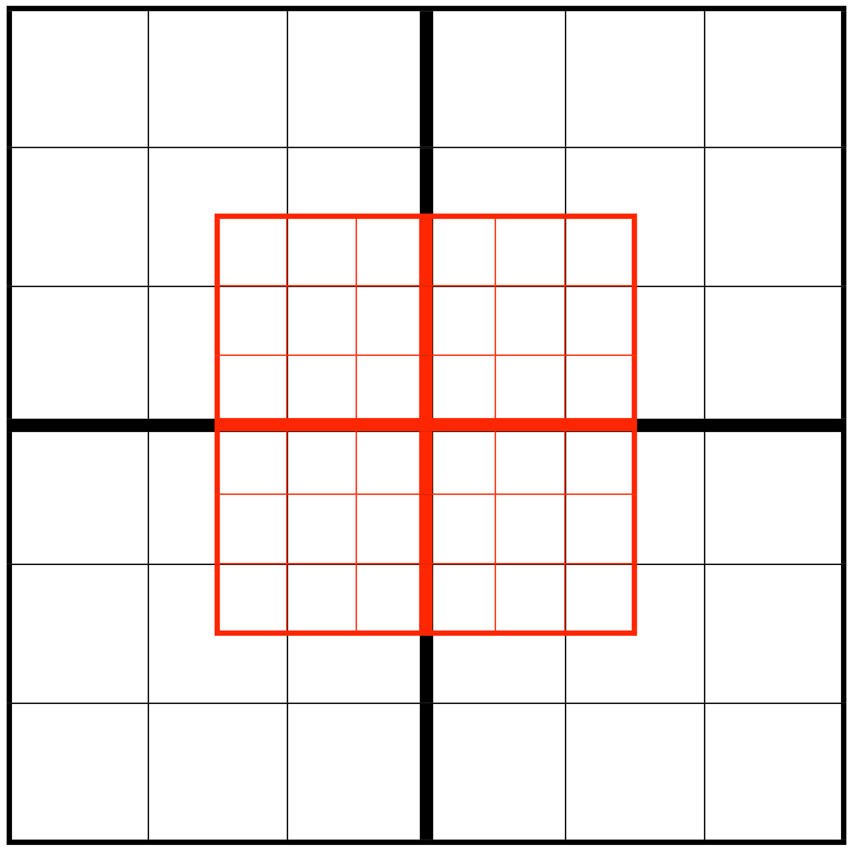
$$\det(\mathbf{A}) = \sum_{j=1}^{N} (-1)^{i+j} A_{ij} M_{ij}(\mathbf{A})$$

$$\mathbf{A} = \begin{pmatrix} -2 & 1 \\ 1 & -2 \end{pmatrix}$$

$\det(\mathbf{A})$ recursively takes $O(N!)$ but MATLAB does it in $O(N^3)$

# Scalars, Vectors and Matrices

- Matrices:
  - What are matrices?
    - They represent transformations!
    - Examples: $\bar{\mathbf{y}} = \mathbf{A}\bar{\mathbf{x}}$

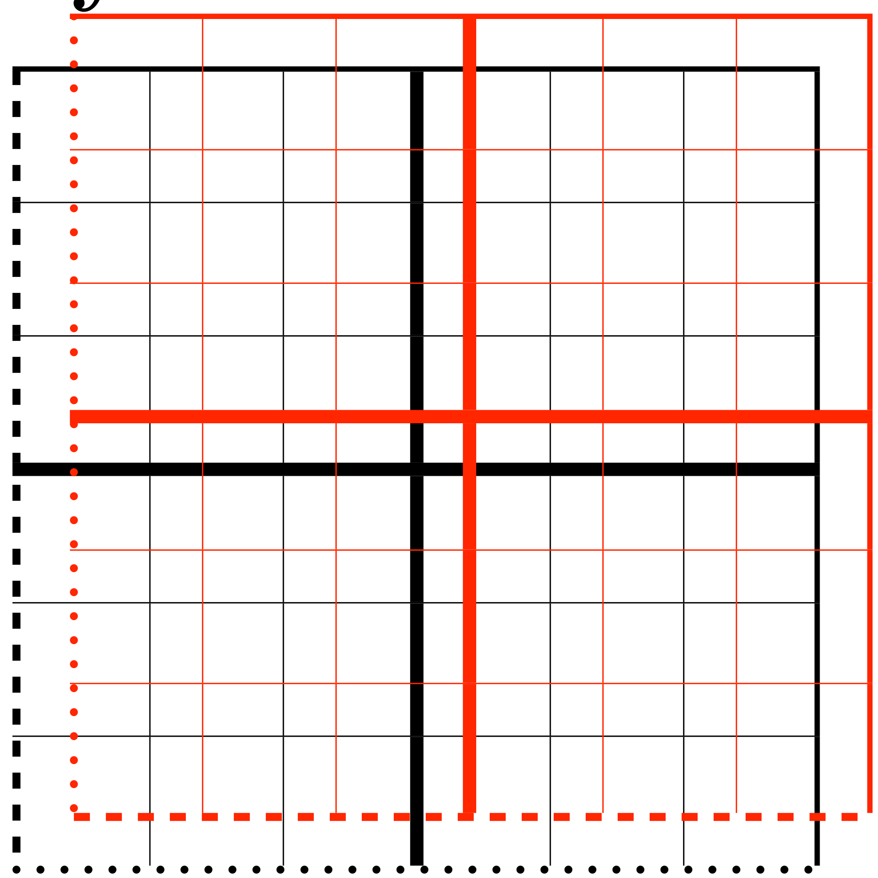


$$\mathbf{y} = \begin{pmatrix} x_1/2 \\ x_2/2 \end{pmatrix}$$

$$\begin{pmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{pmatrix}$$

# Scalars, Vectors and Matrices

- Matrices:
  - What are matrices?
    - They represent transformations!
    - Examples: $\overline{\mathbf{y}} = \mathbf{A}\overline{\mathbf{x}}$

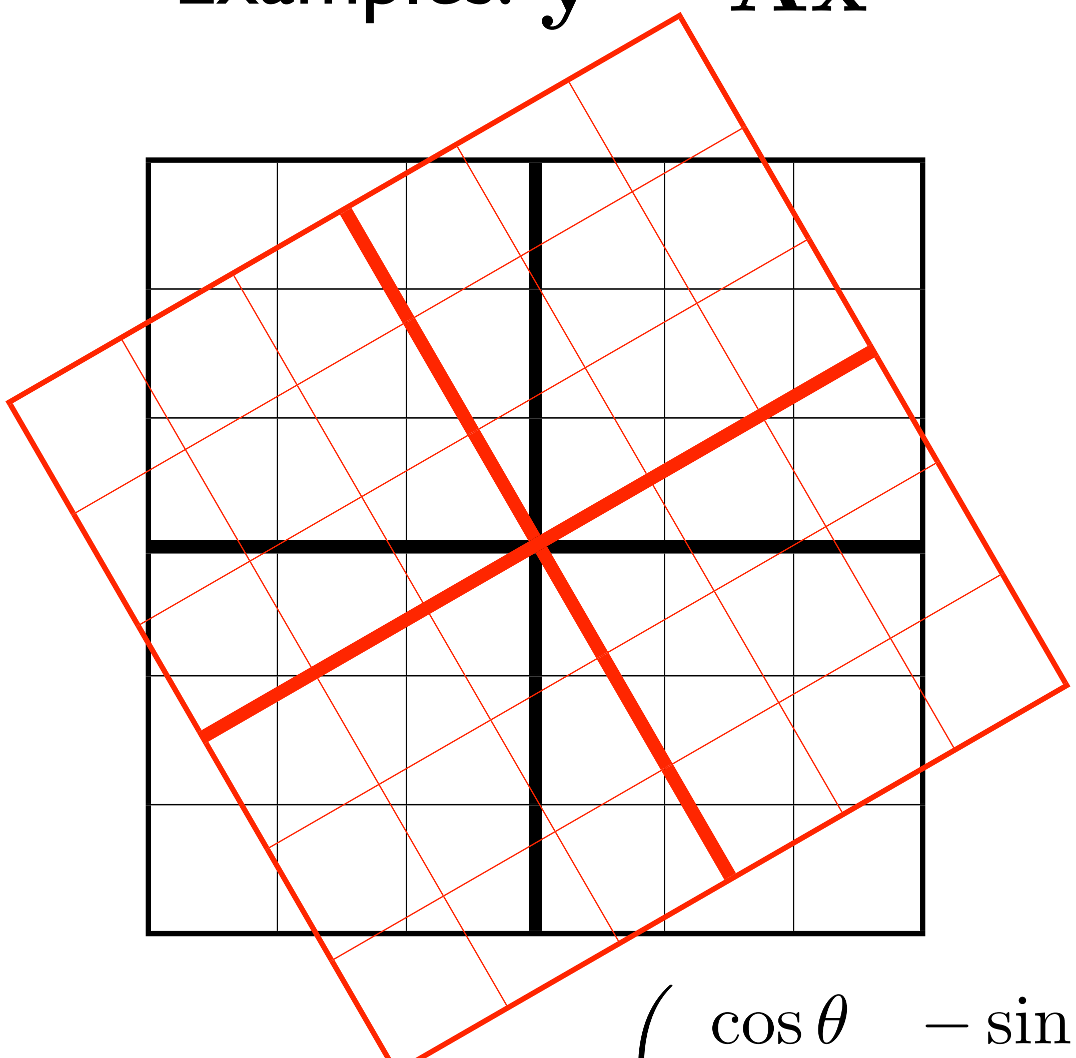


$$\mathbf{y} = \begin{pmatrix} x_2 \\ x_1 \end{pmatrix}$$

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

# Scalars, Vectors and Matrices

- Matrices:
  - What are matrices?
    - They represent transformations!
    - Examples: $\bar{\mathbf{y}} = \mathbf{A}\bar{\mathbf{x}}$

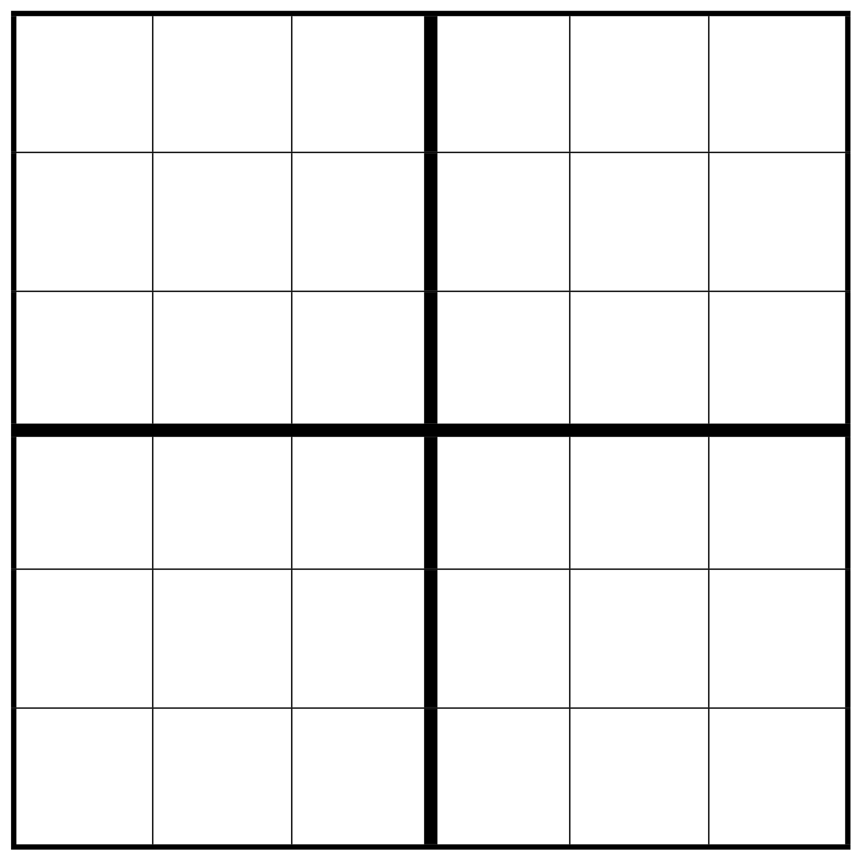


$$\mathbf{y} = \begin{pmatrix} \cos\theta x_1 - \sin\theta x_2 \\ \sin\theta x_1 + \cos\theta x_2 \end{pmatrix}$$

$$\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

# Scalars, Vectors and Matrices

- Matrices:
  - What are matrices?
    - They represent transformations!
    - Examples: $\overline{\mathbf{y}} = \mathbf{A}\overline{\mathbf{x}}$



$$\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$$

# Scalars, Vectors and Matrices

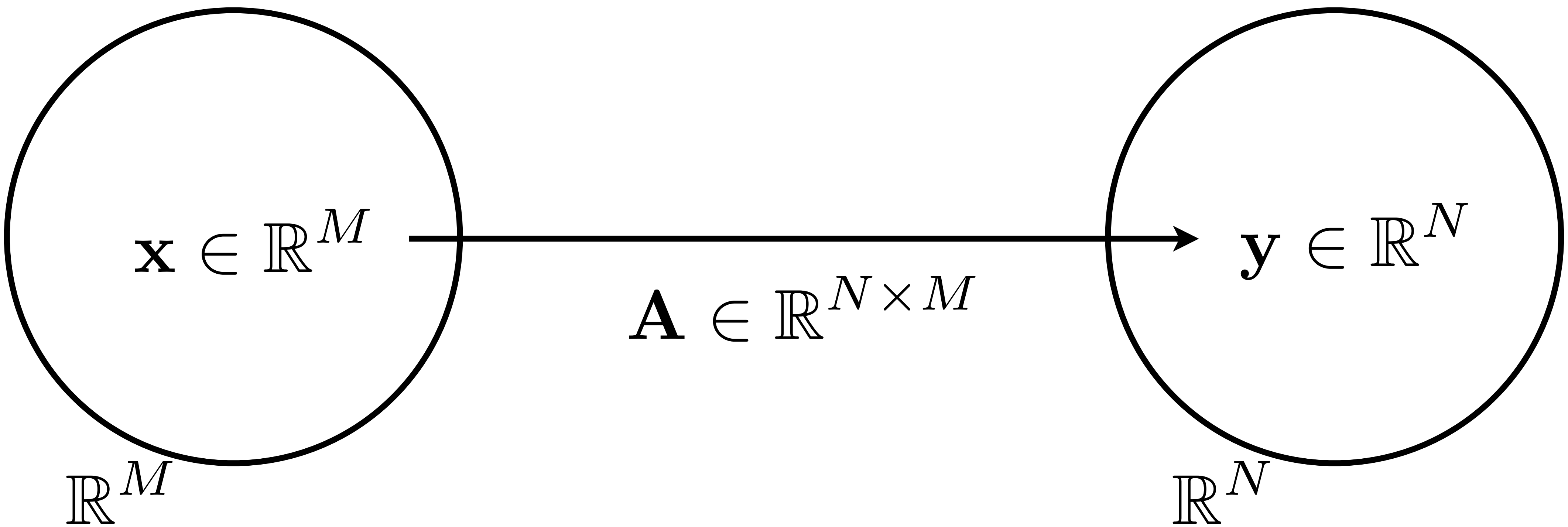
- Matrices:
  - What are matrices?
    - They represent transformations!
    - If a transformation is unique, then it can be undone.
      - The matrix is invertible: $\det(\mathbf{A}) \neq 0$
      - A unique solution to the system of equations exists: $\mathbf{x} = \mathbf{A}^{-1}\mathbf{y}$
    - What happens if a transformation is just barely unique?

$$\mathbf{A} = \begin{pmatrix} 1 & 1+\epsilon \\ 1 & 1 \end{pmatrix}$$

# Scalars, Vectors and Matrices

- Matrices:
  - Matrices are maps between vector spaces!

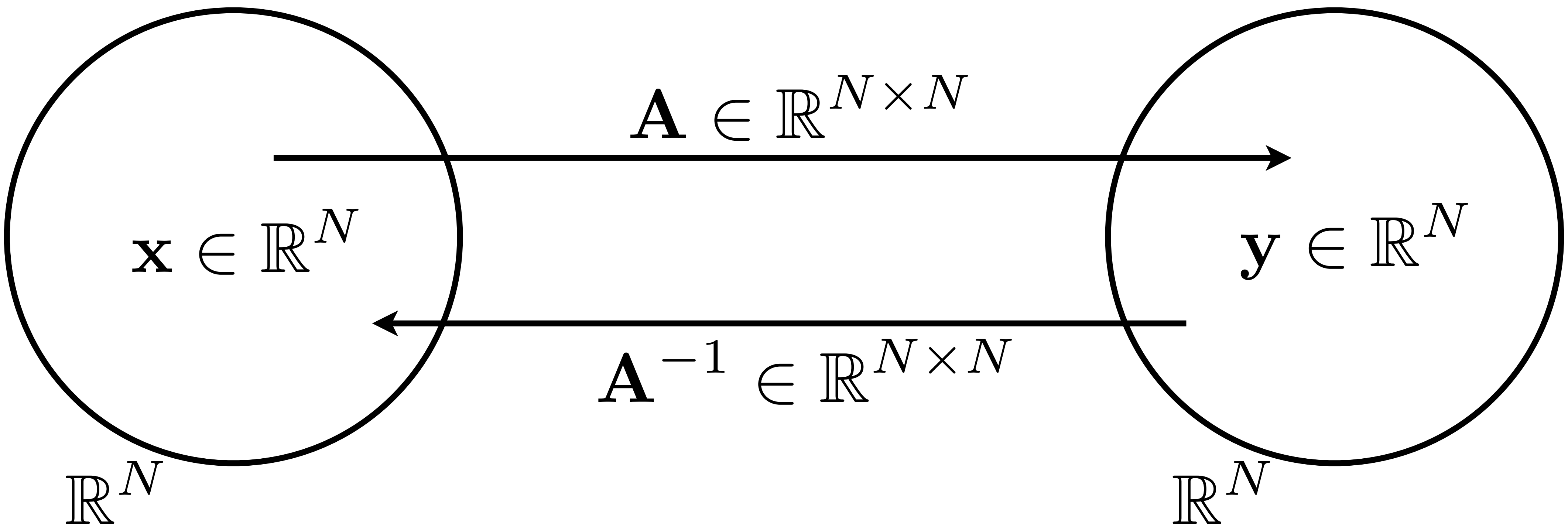
$$\mathbf{y} = \mathbf{A}\mathbf{x}$$



# Scalars, Vectors and Matrices

- Matrices:
  - Matrices are maps between vector spaces!

$$\mathbf{y} = \mathbf{A}\mathbf{x}$$

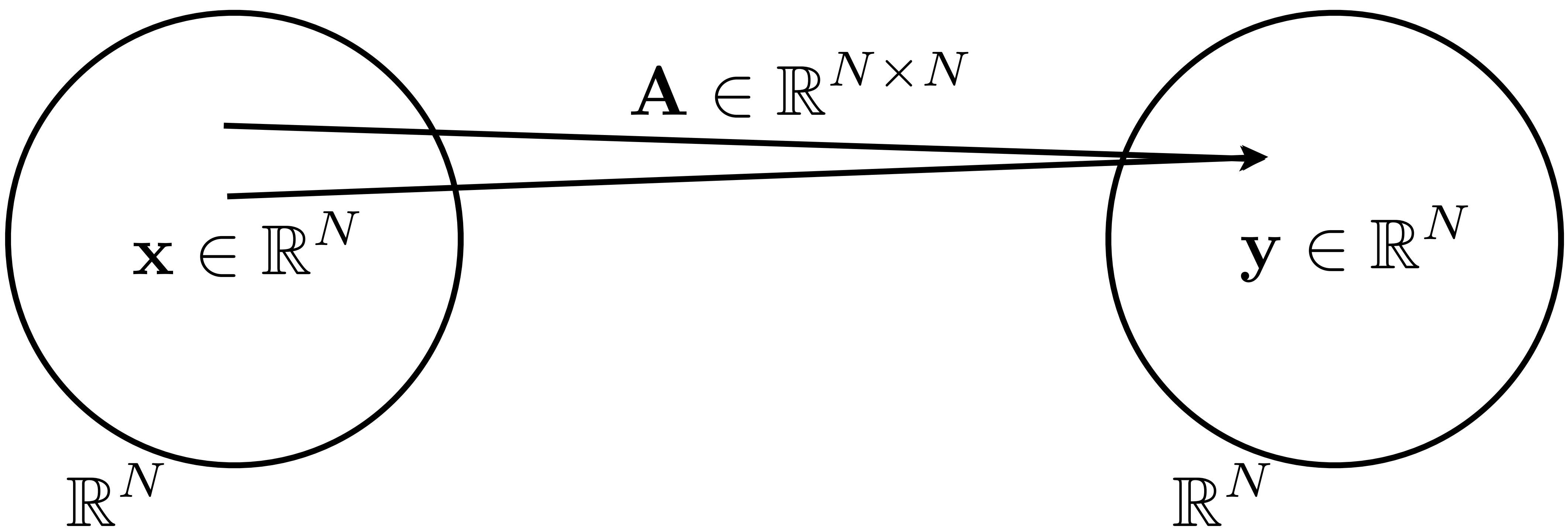



  - When a square matrix is invertible, there is a unique map back the other direction

# Scalars, Vectors and Matrices

- Matrices:
  - Matrices are maps between vector spaces!

$$\mathbf{y} = \mathbf{A}\mathbf{x}$$

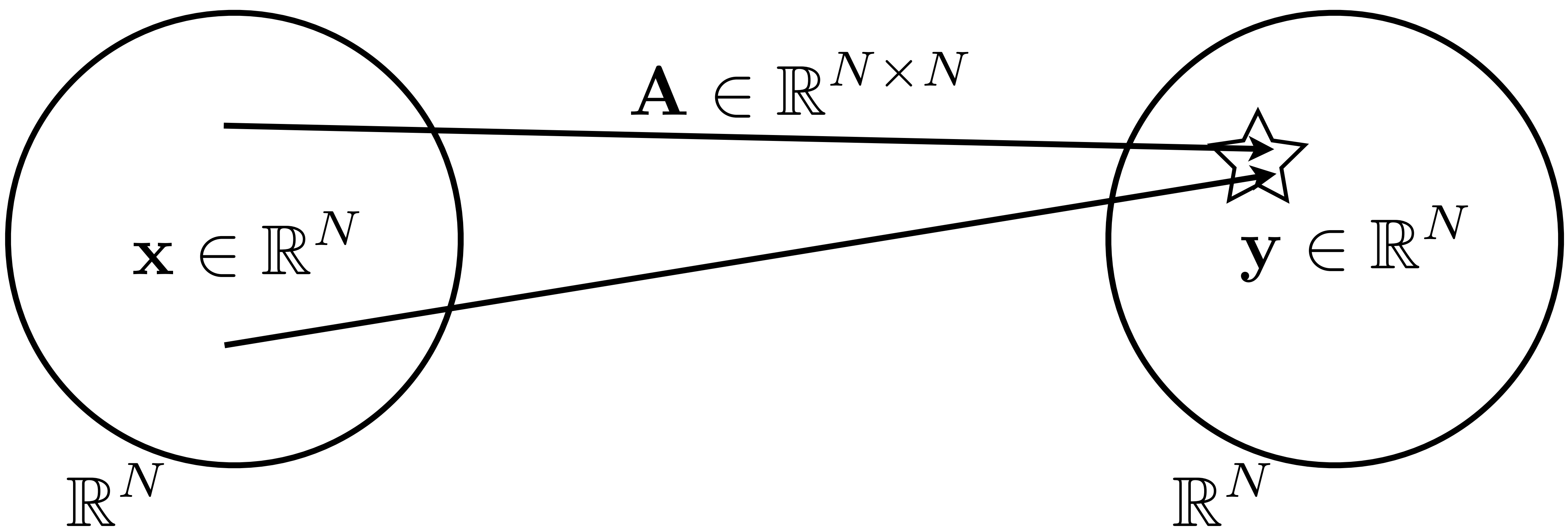



  - When a square matrix is not invertible, the map is not unique or does not cover the entire vector space.

# Scalars, Vectors and Matrices

- Matrices:
  - Matrices are maps between vector spaces!

$$\mathbf{y} = \mathbf{A}\mathbf{x}$$




  - When a square matrix is not invertible, the map is not unique or does not cover the entire vector space.

# Scalars, Vectors and Matrices

- Matrices:
  - Matrix norms: $\mathbf{A} \in \mathbb{R}^{N \times M} \quad \mathbf{x} \in \mathbb{R}^M$
    - Induced norms:

$$\|\mathbf{A}\|_p = \max_{\mathbf{x}} \frac{\|\mathbf{A}\mathbf{x}\|_p}{\|\mathbf{x}\|_p}$$

      - Among all vectors in $\mathbb{R}^M$, what is the maximum "stretch" caused by the matrix $\mathbf{A}$?
      - Example: let $\mathbf{y} = \mathbf{A}\mathbf{x}$ then $\|\mathbf{A}\|_2 = \max_{\mathbf{x}} \frac{\|\mathbf{y}\|_2}{\|\mathbf{x}\|_2}$
      - What is $\|\mathbf{A}\|_\infty$? $\|\mathbf{A}\|_\infty = \max_i \sum_{j=1}^{M} |A_{ij}|$
      - What is $\|\mathbf{A}\|_1$ ? $\|\mathbf{A}\|_1 = \max_j \sum_{i=1}^{N} |A_{ij}|$

# Scalars, Vectors and Matrices

- Matrices:
  - Matrix norms: $\mathbf{A} \in \mathbb{R}^{N \times M}$ $\mathbf{x} \in \mathbb{R}^{M}$ $\mathbf{B} \in \mathbb{R}^{M \times O}$
    - What is $\|\mathbf{A}\|_2$? $\|\mathbf{A}\|_2 = \sqrt{\max_j \lambda_j(\mathbf{A}^T\mathbf{A})}$
      - $\lambda_j(\mathbf{A}^T\mathbf{A})$ is an eigenvalue of $\mathbf{A}^T\mathbf{A}$
    - Properties:
      - $\|\mathbf{A}\|_p \geq 0, \quad \|\mathbf{A}\|_p = 0$ only if $\mathbf{A} = 0$
      - $\|c\mathbf{A}\|_p = |c|\|\mathbf{A}\|_p$
      - $\|\mathbf{A}\mathbf{x}\|_p \leq \|\mathbf{A}\|_p\|\mathbf{x}\|_p$
      - $\|\mathbf{A}\mathbf{B}\|_p \leq \|\mathbf{A}\|_p\|\mathbf{B}\|_p$
      - $\|\mathbf{A}+\mathbf{B}\|_p \leq \|\mathbf{A}\|_p + \|\mathbf{B}\|_p$

# Scalars, Vectors and Matrices

- Matrices:
  - Using matrix norms to estimate numerical error in solution of linear equations:
    - Suppose: $\mathbf{A}\mathbf{x} = \mathbf{b}$, has exact solution: $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$
    - If there is a small error in $\mathbf{b}$, denoted $\delta\mathbf{b}$, how much of an error is produced in $\mathbf{x}$?

$$\mathbf{x} + \delta\mathbf{x} = \mathbf{A}^{-1}(\mathbf{b} + \delta\mathbf{b})$$

$$\delta\mathbf{x} = \mathbf{A}^{-1}\delta\mathbf{b}$$

      - Absolute error in $\mathbf{x}$ :

$$\|\delta\mathbf{x}\|_p = \|\mathbf{A}^{-1}\delta\mathbf{b}\|_p \leq \|\mathbf{A}^{-1}\|_p\|\delta\mathbf{b}\|_p$$

      - Relative error in $\mathbf{x}$ :

$$\|\mathbf{b}\|_p = \|\mathbf{A}\mathbf{x}\|_p \leq \|\mathbf{A}\|_p\|\mathbf{x}\|_p \Rightarrow \|\mathbf{x}\|_p \geq \frac{\|\mathbf{b}\|_p}{\|\mathbf{A}\|_p}$$

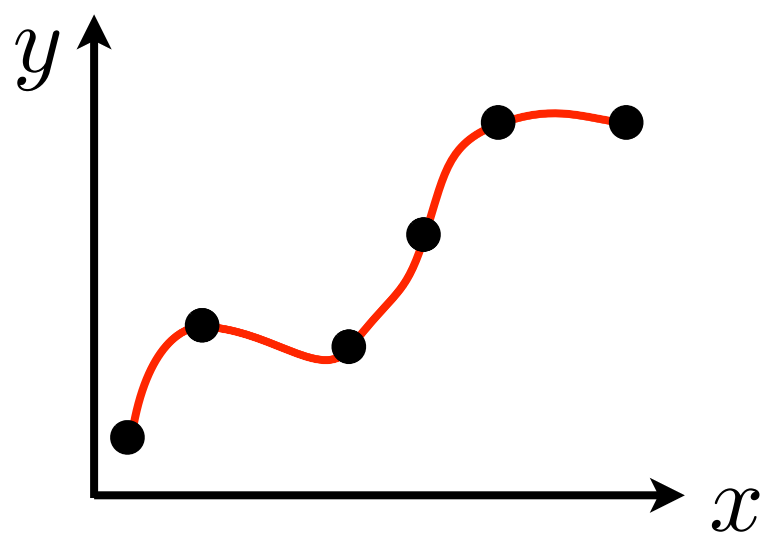
$$\frac{\|\delta\mathbf{x}\|_p}{\|\mathbf{x}\|_p} \leq \|\mathbf{A}\|_p\|\mathbf{A}^{-1}\|_p\frac{\|\delta\mathbf{b}\|_p}{\|\mathbf{b}\|_p}$$

# Scalars, Vectors and Matrices

- Matrices:
    - Condition number: $\kappa(\mathbf{A}) = \|\mathbf{A}\|_p \|\mathbf{A}^{-1}\|_p$
        - Measures how numerical error is magnified in solution of linear equations.
        - Assume a unique solution exists, can we find it?
            - (R.E. in answer) is bounded by (condition number) x (R.E. in data)
            - $\log_{10} \kappa(\mathbf{A})$ gives the number of lost digits
            - "Ill-conditioned" means a large condition number
        - Examples:
            - $\kappa(\mathbf{I}) = 1$
            - $\kappa \begin{pmatrix} 1 & 1 + 10^{-10} \\ 1 & 1 \end{pmatrix} \approx 10^{10}$

# Scalars, Vectors and Matrices

- Matrices:
    - Condition number: $\kappa(\mathbf{A}) = \|\mathbf{A}\|_p \|\mathbf{A}^{-1}\|_p$
        - Examples:
            - Polynomial interpolation:




$$y_i = \sum_{j=1}^{N} a_j x_i^{j-1}$$

$$\mathbf{y} = \mathbf{V}\mathbf{a}$$

- Vandermonde matrix:

$$\mathbf{V} = \begin{pmatrix} 1 & x_1 & x_1^2 & \dots & x_1^N \\ 1 & x_2 & x_2^2 & \dots & x_2^N \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_N & x_N^2 & \dots & x_N^N \end{pmatrix}$$

$$\kappa(\mathbf{V}) > N2^N, \quad N \gg 1$$

# Scalars, Vectors and Matrices

- Matrices:
  - Condition number:
    - $\mathbf{Ax} = \mathbf{b}$ is ill-conditioned. What now?
      - Rescale the equations:

        $$(\mathbf{D}_1\mathbf{A})\mathbf{x} = \mathbf{D}_1\mathbf{b}$$

      - Rescale the unknowns:

        $$(\mathbf{A}\mathbf{D}_2)(\mathbf{D}_2^{-1}\mathbf{x}) = \mathbf{b}$$

      - Rescale both:

        $$(\mathbf{D}_1\mathbf{A}\mathbf{D}_2)(\mathbf{D}_2^{-1}\mathbf{x}) = \mathbf{D}_1\mathbf{b}$$

        - $\mathbf{D}_1$ and $\mathbf{D}_2$ are diagonal matrices
      - An optimal rescaling exists: Braatz and Morari, SIAM J. Control and Optimization 32, 1994

# Scalars, Vectors and Matrices

- Matrices:
  - Condition number:
    - Rescaling example:

$$\mathbf{A} = \begin{pmatrix} 10^{10} & 1 \\ 1 & 10^{-9} \end{pmatrix}$$

$$\kappa(\mathbf{A}) \approx$$

$$\mathbf{D} = \begin{pmatrix} 10^{-10} & 0 \\ 0 & 1 \end{pmatrix}, \quad \kappa(\mathbf{DA}) \approx$$

- The simplest solution is to rescale rows or columns by their maximum element

# Scalars, Vectors and Matrices

- Matrices:
  - Preconditioning:
    - Change the problem so it is easier to solve!
    - Instead of solving: $\mathbf{A}\mathbf{x} = \mathbf{b}$
    - Solve: $(\mathbf{P}_1\mathbf{A}\mathbf{P}_2)(\mathbf{P}_2^{-1}\mathbf{x}) = \mathbf{P}_1\mathbf{b}$
      - $\mathbf{P}_1$ – left, $\mathbf{P}_2$ – right, preconditioner
    - Perhaps the matrix $\mathbf{P}_1\mathbf{A}\mathbf{P}_2$ has better properties:
      - condition number
      - structure
      - sparsity pattern