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:
    \[
    A = \begin{pmatrix}
    A_{11} & A_{12} & \ldots & A_{1M} \\
    A_{21} & A_{22} & \ldots & A_{2M} \\
    \vdots & \vdots & \ddots & \vdots \\
    A_{N1} & A_{N2} & \ldots & A_{NM}
    \end{pmatrix}
    \]
  - Set of all real matrices with \(N\) rows and \(M\) columns, \(\mathbb{R}^{N \times M}\)

- Addition: \(C = A + B \Rightarrow C_{ij} = A_{ij} + B_{ij}\)

- Multiplication by scalar: \(C = cA \Rightarrow C_{ij} = cA_{ij}\)

- Transpose: \(C = A^T \Rightarrow C_{ij} = A_{ji}\)

- Trace (square matrices):
  \[
  \text{Tr } A = \sum_{i=1}^{N} A_{ii}
  \]
Scalars, Vectors and Matrices

- Matrices:
  - Matrix-vector product: $y = Ax \Rightarrow y_i = \sum_{j=1}^{M} A_{ij} x_j$
  - Matrix-matrix product: $C = AB \Rightarrow C_{ij} = \sum_{k=1}^{M} A_{ik} B_{kj}$

- Properties:
  - no commutation in general: $AB \neq BA$
  - association: $A(BC) = (AB)C$
  - distribution: $A(B + C) = AB + AC$
  - transposition: $(AB)^T = B^T A^T$
  - inversion: $A^{-1} A = AA^{-1} = I$ if $\det(A) \neq 0$
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: \( \mathbf{y}^T \mathbf{x} \)?

\[
\mathbf{C} = \mathbf{AB} \implies C_{ij} = \sum_{k=1}^{M} A_{ik} B_{kj}
\]
Scalars, Vectors and Matrices

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

- Matrices:
- Examples: $A, B \in \mathbb{R}^{N \times N}$, $x \in \mathbb{R}^N$

- How many operations to compute:
  - $Ax$
  - $AB$
  - $ABx$
  - What is $x^T ABx$?
  - What is $ABxx^T$?
Scalars, Vectors and Matrices

- Matrices:
  - Dyadic product: \( \mathbf{A} = \mathbf{x y}^T = \mathbf{x} \otimes \mathbf{y} \Rightarrow A_{ij} = x_i y_j \)
  - Determinant (square matrices only):
    \[
    \det(\mathbf{A}) = \sum_{j=1}^{N} (-1)^{i+j} A_{ij} M_{ij}(\mathbf{A})
    \]
    
    \[
    M_{ij}(\mathbf{A}) =
    \begin{pmatrix}
    A_{11} & A_{12} & \cdots & A_{1(j-1)} & A_{1(j+1)} & \cdots & A_{1N} \\
    A_{21} & A_{22} & \cdots & A_{2(j-1)} & A_{2(j+1)} & \cdots & A_{2N} \\
    \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
    A_{(i-1)1} & A_{(i-1)2} & \cdots & A_{(i-1)(j-1)} & A_{(i-1)(j+1)} & \cdots & A_{(i-1)N} \\
    A_{(i+1)1} & A_{(i+1)2} & \cdots & A_{(i+1)(j-1)} & A_{(i+1)(j+1)} & \cdots & A_{(i+1)N} \\
    \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\
    A_{N1} & A_{N2} & \cdots & A_{N(j-1)} & A_{N(j+1)} & \cdots & A_{NN}
    \end{pmatrix}
    \]
  - \( \det(c) = c \)
Scalars, Vectors and Matrices

- Matrices:
  - Determinant (square matrices only):
    \[
    \det(A) = \sum_{j=1}^{N} (-1)^{i+j} A_{ij} M_{ij}(A)
    \]
  - Properties:
    - If any row or column is zeros, \( \det(A) = 0 \)
    - If any row or column is multiplied by \( a \)
      \[
      \det(A_1^c A_2^c a A_3^c \ldots A_N^c) = a \det(A)
      \]
    - Swapping any row or column changes the sign
    - \( \det(A^T) = \det(A) \)
    - \( \det(AB) = \det(A) \det(B) \)
Scalars, Vectors and Matrices

- Matrices:
  - Example:

\[
A = \begin{pmatrix} -2 & 1 \\ 1 & -2 \end{pmatrix}
\]

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

\[
\det(A) = \sum_{j=1}^{N} (-1)^{i+j} A_{ij} M_{ij}(A)
\]
\[ A = \begin{pmatrix} -2 & 1 \\ 1 & -2 \end{pmatrix} \]

\[ \det(A) \text{ recursively takes } O(N!) \text{ but MATLAB does it in } O(N^3) \]
Scalars, Vectors and Matrices

• Matrices:
  • What are matrices?
  • They represent transformations!
  • Examples: \( \bar{y} = A\bar{x} \)

\[
\begin{pmatrix}
\frac{1}{2} & 0 \\
0 & \frac{1}{2}
\end{pmatrix}
\]

\[
y = \begin{pmatrix}
x_1/2 \\
x_2/2
\end{pmatrix}
\]
Scalars, Vectors and Matrices

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

\[
\begin{pmatrix}
0 & 1 \\
1 & 0 \\
\end{pmatrix}
\]

\[
\begin{pmatrix}
x_2 \\
x_1 \\
\end{pmatrix}
\]
Scalars, Vectors and Matrices

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

\[
\begin{pmatrix}
\cos \theta & -\sin \theta \\
\sin \theta & \cos \theta
\end{pmatrix}
\]

\[
y = \begin{pmatrix}
\cos \theta x_1 - \sin \theta x_2 \\
\sin \theta x_1 + \cos \theta x_2
\end{pmatrix}
\]
Scalars, Vectors and Matrices

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

\[
\begin{pmatrix}
1 & 1 \\
1 & 1 \\
\end{pmatrix}
\]
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: $\det(A) \neq 0$
      - A unique solution to the system of equations exists: $x = A^{-1}y$
    - What happens if a transformation is just barely unique?
      \[
      A = \begin{pmatrix}
      1 & 1 + \epsilon \\
      1 & 1
      \end{pmatrix}
      \]
Scalars, Vectors and Matrices

- Matrices:
  - Matrices are maps between vector spaces!

\[ y = Ax \]

\( x \in \mathbb{R}^M \quad \rightarrow \quad y \in \mathbb{R}^N \)

\( A \in \mathbb{R}^{N \times M} \)
Scalars, Vectors and Matrices

- Matrices:
  - Matrices are maps between vector spaces!

\[ y = Ax \]

- 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!

\[ y = Ax \]

- 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!

\[ y = Ax \]

- 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: \( A \in \mathbb{R}^{N \times M} \quad x \in \mathbb{R}^M \)
  - Induced norms:
    \[
    \|A\|_p = \max_x \frac{\|Ax\|_p}{\|x\|_p}
    \]
  - Among all vectors in \( \mathbb{R}^M \), what is the maximum “stretch” caused by the matrix \( A \)?
  - Example: let \( y = Ax \) then \( \|A\|_2 = \max_x \frac{\|y\|_2}{\|x\|_2} \)
  - What is \( \|A\|_{\infty} \)? \( \|A\|_{\infty} = \max_i \sum_{j=1}^{M} |A_{ij}| \)
  - What is \( \|A\|_1 \)? \( \|A\|_1 = \max_j \sum_{i=1}^{N} |A_{ij}| \)
Scalars, Vectors and Matrices

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

- Matrices:
  - Using matrix norms to estimate numerical error in solution of linear equations:
    - Suppose: \( Ax = b \), has exact solution: \( x = A^{-1}b \)
    - If there is a small error in \( b \), denoted \( \delta b \), how much of an error is produced in \( x \)?
      \[
      x + \delta x = A^{-1}(b + \delta b)
      \]
      \[
      \delta x = A^{-1}\delta b
      \]
    - Absolute error in \( x \):
      \[
      \| \delta x \|_p = \| A^{-1}\delta b \|_p \leq \| A^{-1} \|_p \| \delta b \|_p
      \]
    - Relative error in \( x \):
      \[
      \| b \|_p = \| Ax \|_p \leq \| A \|_p \| x \|_p \Rightarrow \| x \|_p \geq \frac{\| b \|_p}{\| A \|_p}
      \]
      \[
      \frac{\| \delta x \|_p}{\| x \|_p} \leq \| A \|_p \| A^{-1} \|_p \| \delta b \|_p \| b \|_p
      \]
Scalars, Vectors and Matrices

• Matrices:
  
  • Condition number: \( \kappa(A) = \| A \|_p \| A^{-1} \|_p \)
  
  • 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)

  • \( \log_{10} \kappa(A) \) gives the number of lost digits

  • “Ill-conditioned” means a large condition number

• Examples:
  
  • \( \kappa(I) = 1 \)

  • \( \kappa \left( \begin{pmatrix} 1 & 1 + 10^{-10} \\ 1 & 1 \end{pmatrix} \right) \approx 10^{10} \)
Scalars, Vectors and Matrices

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

- Examples:
  - Polynomial interpolation:
    \[
    y_i = \sum_{j=1}^{N} a_j x_i^{j-1}
    \]
    \[
    y = \mathbf{V} a
    \]
  - Vandermonde matrix:
    \[
    \mathbf{V} = \begin{pmatrix}
    1 & x_1 & x_1^2 & \ldots & x_1^N \\
    1 & x_2 & x_2^2 & \ldots & x_2^N \\
    \vdots & \vdots & \vdots & \ddots & \vdots \\
    1 & x_N & x_N^2 & \ldots & x_N^N
    \end{pmatrix}
    \]
    \[
    \kappa(\mathbf{V}) > N2^N, \quad N \gg 1
    \]
Scalars, Vectors and Matrices

- Matrices:
  - Condition number:
    - $Ax = b$ is ill-conditioned. What now?
    - Rescale the equations:
      $$(D_1A)x = D_1b$$
    - Rescale the unknowns:
      $$(AD_2)(D_2^{-1}x) = b$$
    - Rescale both:
      $$(D_1AD_2)(D_2^{-1}x) = D_1b$$
- $D_1$ and $D_2$ 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:

\[
A = \begin{pmatrix}
10^{10} & 1 \\
1 & 10^{-9}
\end{pmatrix}
\]

\[
\kappa(A) \approx \frac{10^{10}}{10^{-9}}
\]

\[
D = \begin{pmatrix}
10^{-10} & 0 \\
0 & 1
\end{pmatrix}, \quad \kappa(DA) \approx \frac{10^{10}}{10^{-9}}
\]

- 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: $Ax = b$
    - Solve: $(P_1AP_2)(P_2^{-1}x) = P_1b$
      - $P_1$ – left, $P_2$ – right, preconditioner
    - Perhaps the matrix $P_1AP_2$ has better properties:
      - condition number
      - structure
      - sparsity pattern