MAT 222 Linear Algebra
Week 11
Lecture Notes 2

Murat Karaçayır

Akdeniz Üniversitesi
Matematik Bölümü

15th May 2025

Iterative methods for linear systems

- We focus on $n \times n$ systems with a unique solution, i.e. $\mathbf{Ax} = \mathbf{b}$ where $\mathbf{A}$ is invertible.
- Gaussian elimination takes $O(n^3)$ operations to find the solution.
- It turns out that we can do better using an iterative approach.
- The approach we will take is a version of **fixed point iteration** technique for nonlinear equations in one variable.
- Under some conditions, given an initial guess x_0, the sequence defined by $(x_n)_{n=1}^{\infty}$, where $x_{k+1} = g(x_k)$ converges to a solution of the equation $g(x) = x$.
- This method can be adapted to linear systems of equations.

Iterative methods for linear systems

- As an example, let's take the following system:

$$x + 2y - 2z = 7$$
$$x + y + z = 2$$
$$2x + 2y + z = 1$$

- Let us write this system as follows:

$$x = -2y + 2z + 7$$
$$y = -x - z + 2$$
$$z = -2x - 2y + 1$$

- Now, let us take some initial guesses for the unknowns x, y, z and use these guesses to update their values. For example let us take $x^{(0)} = 0, y^{(0)} = 0, z^{(0)} = 0$.

$$x = -2 \cdot 0 + 2 \cdot 0 + 7 = 7$$
$$y = -0 - 0 + 2 = 2$$
$$z = -2 \cdot 0 - 2 \cdot 0 + 1 = 1$$

- So we have found $x^{(1)} = 7, y^{(1)} = 2, z^{(1)} = 1$. Let us iterate again.

Iterative methods for linear systems

- This time we obtain

$$x^{(2)} = -2\cdot 2 + 2\cdot 1 + 7 = 5$$
$$y^{(2)} = -7 - 1 + 2 = -6$$
$$z^{(2)} = -2\cdot 7 - 2\cdot 2 + 1 = -17$$

- Two more steps yield

$$x^{(3)} = -2\cdot(-6) + 2\cdot(-17) + 7 = -15$$
$$y^{(3)} = -5 - (-17) + 2 = 14$$
$$z^{(3)} = -2\cdot 5 - 2\cdot(-6) + 1 = 3$$

and

$$x^{(4)} = -2\cdot 14 + 2\cdot 3 + 7 = -15$$
$$y^{(4)} = -(-15) - 3 + 2 = 14$$
$$z^{(4)} = -2\cdot(-15) - 2\cdot(14) + 1 = 3$$

- The results did not change! That's beacuse $x = -15, y = 14, z = 3$ is the exact solution of the system.
- This method is known as Jacobi iteration.

Jacobi iteration

- This scheme can be explained in terms of matrices in a more concise way.
- The equations

$$\begin{aligned} x &= -2y + 2z + 7 \\ y &= -x - z + 2 \\ z &= -2x - 2y + 1 \end{aligned} \tag{1}$$

can be expressed as follows:

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 & -2 & 2 \\ -1 & 0 & -1 \\ -2 & -2 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} + \begin{bmatrix} 7 \\ 2 \\ 1 \end{bmatrix}$$

- Thus, letting

$$\mathbf{x} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}, \mathbf{T} = \begin{bmatrix} 0 & -2 & 2 \\ -1 & 0 & -1 \\ -2 & -2 & 0 \end{bmatrix}, \mathbf{c} = \begin{bmatrix} 7 \\ 2 \\ 1 \end{bmatrix}$$

yields the following compact expression for (1):

$$\mathbf{x} = \mathbf{T}\mathbf{x} + \mathbf{c}$$

Jacobi iteration

- The mechanism of Jacobi iteration can be explained by a simple algorithm.
  - Express the system $\mathbf{Ax} = \mathbf{b}$ as $\mathbf{x} = \mathbf{Tx} + \mathbf{c}$.
  - Starting by an initial guess $\mathbf{x}^{(0)}$, for $n = 0, 1, 2, \ldots$ compute the approximations
$$\mathbf{x}^{(n+1)} = \mathbf{Tx}^{(n)} + \mathbf{c}.$$
  - Stop when $\mathbf{x}^{(n+1)}$ and $\mathbf{x}^{(n)}$ are "close enough".
- Let us now illustrate the steps of the same example using matrices.
- First two steps are

$$\mathbf{x}^{(1)} = \mathbf{Tx}^{(0)} + \mathbf{c} = \begin{bmatrix} 0 & -2 & 2 \\ -1 & 0 & -1 \\ -2 & -2 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 7 \\ 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 7 \\ 2 \\ 1 \end{bmatrix}$$

and

$$\mathbf{x}^{(2)} = \mathbf{Tx}^{(1)} + \mathbf{c} = \begin{bmatrix} 0 & -2 & 2 \\ -1 & 0 & -1 \\ -2 & -2 & 0 \end{bmatrix} \begin{bmatrix} 7 \\ 2 \\ 1 \end{bmatrix} + \begin{bmatrix} 7 \\ 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 5 \\ -6 \\ -17 \end{bmatrix}.$$

Jacobi iteration

- The remaining steps are

$$\mathbf{x}^{(3)} = \mathbf{T}\mathbf{x}^{(2)} + \mathbf{c} = \begin{bmatrix} 0 & -2 & 2 \\ -1 & 0 & -1 \\ -2 & -2 & 0 \end{bmatrix} \begin{bmatrix} 5 \\ -6 \\ -17 \end{bmatrix} + \begin{bmatrix} 7 \\ 2 \\ 1 \end{bmatrix} = \begin{bmatrix} -15 \\ 14 \\ 3 \end{bmatrix},$$

$$\mathbf{x}^{(4)} = \mathbf{T}\mathbf{x}^{(3)} + \mathbf{c} = \begin{bmatrix} 0 & -2 & 2 \\ -1 & 0 & -1 \\ -2 & -2 & 0 \end{bmatrix} \begin{bmatrix} -15 \\ 14 \\ 3 \end{bmatrix} + \begin{bmatrix} 7 \\ 2 \\ 1 \end{bmatrix} = \begin{bmatrix} -15 \\ 14 \\ 3 \end{bmatrix}.$$

- It is an exceptional case to have $\mathbf{x}^{(4)} = \mathbf{x}^{(3)}$. In general, we must find a way to measure how close two successive approximations are.
- Since approximations are vectors, it is natural to use vector norms for this purpose.

Stopping criterion for Jacobi iteration

Definition: Let $\mathbf{x} = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$. "Uniform norm" or "Infinity norm" of the vector $\mathbf{x}$ is defined by

$$\|\mathbf{x}\|_\infty = \max\{|x_i| : i = 1, 2, \ldots, n\}.$$

- For instance, the infinity norm of the vector $\mathbf{x} = (-5, 3, 1.5)$ is

$$\|\mathbf{x}\|_\infty = \max\{|x_1|, |x_2|, |x_3|\} = \max\{5, 3, 1.5\} = 5.$$

- It is straightforward to use this norm to measure the distance between two vectors: If $\mathbf{u}$ and $\mathbf{v}$ are two vectors of the same size, then their distance (according to ∞-norm) would be

$$\|\mathbf{u} - \mathbf{v}\|_\infty = \max\{|u_i - v_i| : i = 1, 2, \ldots, n\}.$$

- For instance, the distance between $\mathbf{u} = (1.5, -4, 7)$ and $\mathbf{v} = (2.5, 0, 6.2)$ is

$$\|\mathbf{u}-\mathbf{v}\|_\infty = \max\{|u_i - v_i| : i = 1, 2, \ldots, n\} = \max\{|-1|, |-4|, |0.8|\} = 4.$$

- We will use this norm to induce a stopping criterion for the Jacobi iteration algorithm.

Stopping criterion for Jacobi iteration: An example

Example:

$$\begin{aligned} 7x_1 + x_2 - 2x_3 + x_4 &= 1 \\ x_1 + 8x_2 + x_3 &= -1 \\ -2x_1 + x_2 + 5x_3 - x_4 &= 1 \\ x_1 - x_3 + 3x_4 &= -1 \end{aligned}$$

Let us solve this system by Jacobi iteration using the tolerance $\varepsilon = 10^{-3}$.

- We first write

$$\begin{aligned} x_1 &= -\frac{1}{7}x_2 + \frac{2}{7}x_3 - \frac{1}{7}x_4 + \frac{1}{7} \\ x_2 &= -\frac{1}{8}x_1 - \frac{1}{8}x_3 - \frac{1}{8} \\ x_3 &= \frac{2}{5}x_1 - \frac{1}{5}x_2 + \frac{1}{5}x_4 + \frac{1}{5} \\ x_4 &= -\frac{1}{3}x_1 + \frac{1}{3}x_3 - \frac{1}{3} \end{aligned}$$

Stopping criterion for Jacobi iteration: An example

- In this problem we have

$$\mathbf{T} = \begin{bmatrix} 0 & -\frac{1}{7} & \frac{2}{7} & -\frac{1}{7} \\ -\frac{1}{8} & 0 & -\frac{1}{8} & 0 \\ \frac{2}{5} & -\frac{1}{5} & 0 & \frac{1}{5} \\ -\frac{1}{3} & 0 & \frac{1}{3} & 0 \end{bmatrix}, \quad \mathbf{c} = \begin{bmatrix} \frac{1}{7} \\ -\frac{1}{8} \\ \frac{1}{5} \\ -\frac{1}{3} \end{bmatrix}.$$

- Again, let us start with the guess $x_1 = x_2 = x_3 = x_4 = 0$.
- We will iterate until $\|\mathbf{x}^{(n+1)} - \mathbf{x}^{(n)}\|_\infty < 10^{-3}$ is satisfied.
- After the first step we obtain

$$\mathbf{x}^{(1)} = \mathbf{T}\mathbf{x}^{(0)} + \mathbf{c} = \begin{bmatrix} 0 & -\frac{1}{7} & \frac{2}{7} & -\frac{1}{7} \\ -\frac{1}{8} & 0 & -\frac{1}{8} & 0 \\ \frac{2}{5} & -\frac{1}{5} & 0 & \frac{1}{5} \\ -\frac{1}{3} & 0 & \frac{1}{3} & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} \frac{1}{7} \\ -\frac{1}{8} \\ \frac{1}{5} \\ -\frac{1}{3} \end{bmatrix} = \begin{bmatrix} \frac{1}{7} \\ -\frac{1}{8} \\ \frac{1}{5} \\ -\frac{1}{3} \end{bmatrix}.$$

- Since $\mathbf{x}^{(1)} - \mathbf{x}^{(0)} = (\frac{1}{7}, -\frac{1}{8}, \frac{1}{5}, -\frac{1}{3})$ we have $\|\mathbf{x}^{(1)} - \mathbf{x}^{(0)}\|_\infty = \dfrac{1}{3}$. So we don't stop at this point.

Stopping criterion for Jacobi iteration: An example

- Let us perform step 2.

$$\mathbf{x}^{(2)} = \mathbf{T}\mathbf{x}^{(1)} + \mathbf{c} = \begin{bmatrix} 0 & -\frac{1}{7} & \frac{2}{7} & -\frac{1}{7} \\ -\frac{1}{8} & 0 & -\frac{1}{8} & 0 \\ \frac{2}{5} & -\frac{1}{5} & 0 & \frac{1}{5} \\ -\frac{1}{3} & 0 & \frac{1}{3} & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{7} \\ -\frac{1}{8} \\ \frac{1}{5} \\ -\frac{1}{3} \end{bmatrix} + \begin{bmatrix} \frac{1}{7} \\ -\frac{1}{8} \\ \frac{1}{5} \\ -\frac{1}{3} \end{bmatrix} = \begin{bmatrix} 0.26547 \\ -0.16785 \\ 0.21547 \\ -0.31428 \end{bmatrix}$$

- Since $\mathbf{x}^{(2)} - \mathbf{x}^{(1)} = (0.12262, -0.04285, 0.01547, 0.01904)$, it follows that $\|\mathbf{x}^{(2)} - \mathbf{x}^{(1)}\|_\infty = 0.1226 > 10^{-3}$. So we continue to iterate.
- At the end of the next step we obtain

$$\mathbf{x}^{(3)} = \mathbf{T}\mathbf{x}^{(2)} + \mathbf{c} = \begin{bmatrix} 0 & -\frac{1}{7} & \frac{2}{7} & -\frac{1}{7} \\ -\frac{1}{8} & 0 & -\frac{1}{8} & 0 \\ \frac{2}{5} & -\frac{1}{5} & 0 & \frac{1}{5} \\ -\frac{1}{3} & 0 & \frac{1}{3} & 0 \end{bmatrix} \begin{bmatrix} 0.26547 \\ -0.16785 \\ 0.21547 \\ -0.31428 \end{bmatrix} + \begin{bmatrix} \frac{1}{7} \\ -\frac{1}{8} \\ \frac{1}{5} \\ -\frac{1}{3} \end{bmatrix} = \begin{bmatrix} 0.27329 \\ -0.18511 \\ 0.27690 \\ -0.35 \end{bmatrix}.$$

  Since $\|\mathbf{x}^{(3)} - \mathbf{x}^{(2)}\|_\infty = 0.0614$, we again move on.
- The remaining steps are given in the table on the next page.

Stopping criterion for Jacobi iteration: An example

n	$x_1^{(n)}$	$x_2^{(n)}$	$x_3^{(n)}$	$x_4^{(n)}$	$\lVert \mathbf{x}^{(n)} - \mathbf{x}^{(n-1)} \rVert_\infty$
0	0	0	0	0	—
1	$\frac{1}{7}$	$-\frac{1}{8}$	$\frac{1}{5}$	$-\frac{1}{3}$	0.3333
2	0.26547	−0.16785	0.21547	−0.31428	0.1226
3	0.27329	−0.18511	0.27690	−0.35	0.0614
4	0.29841	−0.19377	0.27634	−0.33213	0.0251
5	0.29694	−0.19684	0.29169	−0.34069	0.0153
6	0.30298	−0.19857	0.29001	−0.33508	0.0060
7	0.30195	−0.19912	0.29389	−0.33766	0.0038
8	0.30351	−0.19948	0.29307	−0.33601	0.0016
9	0.30309	−0.19957	0.29409	−0.33681	0.0010
10	0.30351	−0.19964	0.29378	−0.33633	0.0004

- It took 10 steps to achieve the desired accuracy.
- The approximate solution is computed to be $x_1 \approx 0.30351$, $x_2 \approx -0.19964$, $x_3 \approx 0.29378$, $x_4 \approx -0.33633$.

Jacobi iteration may not converge

- Jacobi iteration has a complexity of $rO(n^2)$, where r is the number of steps taken to converge. This is advantageous over Gaussian elimination especially for large n.
- On the other hand, it has a major drawback: It may not converge.
- As an example let us consider the following system:

$$\begin{aligned} 2x - y + z &= -1 \\ x + y + z &= 2 \\ -x - y + 2z &= -5 \end{aligned}$$

- The first five steps of Jacobi iteration with the initial guess $\mathbf{x}^{(0)} = (0, 0, 0)$ is illustrated in the below table. The method does not converge.

n	$x^{(n)}$	$y^{(n)}$	$z^{(n)}$
0	0	0	0
1	-0.5	2	-0.4
2	0.7	2.9	0.35
3	0.775	0.95	1.4
4	-0.725	-0.175	0.4625
5	-0.8187	2.2625	-0.85

Convergence criterion for Jacobi iteration

- Various criteria have been proposed for the convergence of Jacobi iteration.
- Let us state the following (quite strong) sufficient condition related to the coefficient matrix **A**.

Theorem: Consider the $n \times n$ linear system $\mathbf{Ax} = \mathbf{b}$ where the matrix $\mathbf{A} = [a_{ij}]$ is invertible. If **A** satisfies

$$|a_{ii}| > \sum_{j \neq i} |a_{ij}|$$

for every $i = 1, 2, \ldots, n$, then Jacobi iteration for the system $\mathbf{Ax} = \mathbf{b}$ converges to the solution of the system for any initial guess $\mathbf{x}^{(0)}$.

- In other words, if the absolute value of each of the diagonal entries of **A** is greater than the sum of the absolute values of the other entries in its row, then you can apply Jacobi iteration to the system $\mathbf{Ax} = \mathbf{b}$ with any initial guess $\mathbf{x}^{(0)}$ that you like; the system will certainly converge.
- The matrices with this property are called strictly diagonally dominant.
- This convergence criterion may give us an idea on how we should order the equations in the system $\mathbf{Ax} = \mathbf{b}$ in some cases.

Convergence criterion for Jacobi iteration: An example

Example: Let us apply Jacobi iteration to the system

$$\begin{aligned} x_1 + 8x_2 + x_3 &= -1 \\ 7x_1 + x_2 - 2x_3 + x_4 &= 1 \\ -2x_1 + x_2 + 5x_3 - x_4 &= 1 \\ x_1 - x_3 + 3x_4 &= -1 \end{aligned} .$$

- Proceeding as usual, we obtain

$$\begin{aligned} x_1 &= -8x_2 - x_3 - 1 \\ x_2 &= -7x_1 + 2x_3 - x_4 + 1 \\ x_3 &= \frac{2}{5}x_1 - \frac{1}{5}x_2 + \frac{1}{5}x_4 + \frac{1}{5} \\ x_4 &= -\frac{1}{3}x_1 + \frac{1}{3}x_3 - \frac{1}{3}. \end{aligned}$$

- As a result, we have

$$\mathbf{T} = \begin{bmatrix} 0 & -8 & -1 & 0 \\ -7 & 0 & 2 & -1 \\ \frac{2}{5} & -\frac{1}{5} & 0 & \frac{1}{5} \\ -\frac{1}{3} & 0 & \frac{1}{3} & 0 \end{bmatrix}, \mathbf{c} = \begin{bmatrix} -1 \\ 1 \\ \frac{1}{5} \\ -\frac{1}{3} \end{bmatrix}.$$

Convergence criterion for Jacobi iteration: An example

- Jacobi iteration with $\mathbf{x}^{(0)} = (0, 0, 0, 0)$ gives the following results:

n	$x_1^{(n)}$	$x_2^{(n)}$	$x_3^{(n)}$	$x_4^{(n)}$
0	0	0	0	
1	-1	1	0.2	-0.3333
2	-9.2	8.7333	-0.4667	0.0667
3	-70.4	64.4	-5.2133	2.5778

- The method does not converge.
- In this case, we could have secured convergence by ordering the equations so that the coefficient matrix **A** becomes strictly diagonally dominant.

Making the system strictly diagonally dominant

- In order to make the system strictly diagonally dominant, we can swap the positions of the first and second equations.

$$\begin{aligned} 7x_1 + x_2 - 2x_3 + x_4 &= 1 \\ x_1 + 8x_2 + x_3 &= -1 \\ -2x_1 + x_2 + 5x_3 - x_4 &= 1 \\ x_1 - x_3 + 3x_4 &= -1 \end{aligned}.$$

- Thus, we can apply Jacobi iteration as follows:

$$\begin{aligned} x_1 &= -\frac{1}{7}x_2 + \frac{2}{7}x_3 - \frac{1}{7}x_4 + \frac{1}{7} \\ x_2 &= -\frac{1}{8}x_1 - \frac{1}{8}x_3 - \frac{1}{8} \\ x_3 &= \frac{2}{5}x_1 - \frac{1}{5}x_2 + \frac{1}{5}x_4 + \frac{1}{5} \\ x_4 &= -\frac{1}{3}x_1 + \frac{1}{3}x_3 - \frac{1}{3}. \end{aligned}$$

Making the system strictly diagonally dominant

- The results are seen below.

n	$x_1^{(n)}$	$x_2^{(n)}$	$x_3^{(n)}$	$x_4^{(n)}$	$\|\mathbf{x}^{(n)} - \mathbf{x}^{(n-1)}\|_\infty$
0	0	0	0	0	—
1	$\frac{1}{7}$	$-\frac{1}{8}$	$\frac{1}{5}$	$-\frac{1}{3}$	0.3333
2	0.26547	-0.16785	0.21547	-0.31428	0.1226
3	0.27329	-0.18511	0.27690	-0.35	0.0614
4	0.29841	-0.19377	0.27634	-0.33213	0.0251
5	0.29694	-0.19684	0.29169	-0.34069	0.0153
6	0.30298	-0.19857	0.29001	-0.33508	0.0060
7	0.30195	-0.19912	0.29389	-0.33766	0.0038
8	0.30351	-0.19948	0.29307	-0.33601	0.0016
9	0.30309	-0.19957	0.29409	-0.33681	0.0010
10	0.30351	-0.19964	0.29378	-0.33633	0.0004

- The method now converges.

Gauss-Seidel iteration

- At each step of Jacobi iteration, the values computed for the unknowns $x_1, x_2, \ldots, x_n$ are not used until the next step.
- This may cause a slowdown in the convergence process.
- Let us consider the same system

$$
\begin{aligned}
7x_1 + x_2 - 2x_3 + x_4 &= 1 \\
x_1 + 8x_2 + x_3 &= -1 \\
-2x_1 + x_2 + 5x_3 - x_4 &= 1 \\
x_1 - x_3 + 3x_4 &= -1.
\end{aligned}
$$

  This time, when updating the values of the unknowns, we will use the computed values as soon as they are computed.
- Thus, we will apply the update equations

$$
\begin{aligned}
x_1 &= -\frac{1}{7}x_2 + \frac{2}{7}x_3 - \frac{1}{7}x_4 + \frac{1}{7} \\
x_2 &= -\frac{1}{8}x_1 - \frac{1}{8}x_3 - \frac{1}{8} \\
x_3 &= \frac{2}{5}x_1 - \frac{1}{5}x_2 + \frac{1}{5}x_4 + \frac{1}{5} \\
x_4 &= -\frac{1}{3}x_1 + \frac{1}{3}x_3 - \frac{1}{3}
\end{aligned}
$$

  not through a matrix multiplication but one by one.
- This method is known as Gauss-Seidel iteration.

Gauss-Seidel iteration

- For example, taking $\mathbf{x}^{(0)} = (0,0,0,0)$ as the initial guess, let us compute the next guess for x_1.

$$x_1^{(1)} = -\frac{1}{7}x_2^{(0)} + \frac{2}{7}x_3^{(0)} - \frac{1}{7}x_4^{(0)} + \frac{1}{7} = \frac{1}{7}$$

- Now we will use this "fresh" value to compute our next guess for x_2. We have

$$x_2^{(1)} = -\frac{1}{8}x_1^{(1)} - \frac{1}{8}x_3^{(0)} - \frac{1}{8} = -\frac{1}{8}\cdot\frac{1}{7} - \frac{1}{8}\cdot 0 - \frac{1}{8} = -\frac{1}{7}.$$

- The results of the first four steps thus computed are listed below.

n	$x_1^{(n)}$	$x_2^{(n)}$	$x_3^{(n)}$	$x_4^{(n)}$	$\lVert\mathbf{x}^{(n)} - \mathbf{x}^{(n-1)}\rVert$
0	0	0	0	0	—
1	$\frac{1}{7}$	$-\frac{1}{7}$	$\frac{2}{7}$	$-\frac{2}{7}$	0.2857
2	0.28571	−0.19642	0.29642	−0.32976	0.1428
3	0.30272	−0.19989	0.29511	−0.33586	0.0170
4	0.30371	−0.19985	0.29428	−0.33647	0.0009

- As seen in the table, the algorithm has reached a sensitivity of $\varepsilon_y = 10^{-3}$ in four steps. In Jacobi iteration it took ten.
- The sufficient criterion for convergence (strictly diagonally dominant coefficient matrix) is also valid for Gauss-Seidel iteration.