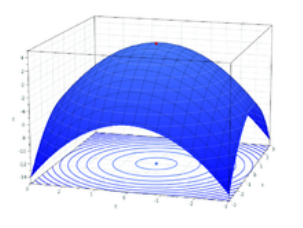
Lecture 5 - Systems of Linear Equations

COMP1046 - Maths for Computer Scientists

Dr. Ferrante Neri / Dr. Tony Bellotti

The University of Nottingham



Learning outcomes

By the end of this lecture we will have learned:

- Definition of a System of Linear Equations
- Cramer's Method
- Rouchè-Capelli Theorem
- Gaussian Elimination
- Summary of Methods

Based on Sections 3.1 to 3.3 of text book (Neri 2018).

Systems of Linear Equations

Linear equation

Definition

A *linear equation* in $\mathbb{R}$ in the variables $x_1, x_2, \ldots, x_n$ is an equation of the kind:

$$a_1x_1 + a_2x_2 + \ldots + a_nx_n = b$$

where $\forall$ index i,

- a_i is said *coefficient* of the equation,
- a_ix_i is said i^{th} *term* of the equation, and
- b is said *known term*.

Coefficients and known term are constant and known numbers in $\mathbb{R}$ while the variables are an unknown set of numbers in $\mathbb{R}$ that satisfy the equality.

Definition

Let us consider m (with $m > 1$) linear equations in the variables $x_1, x_2, \ldots, x_n$. These equations compose a *system of linear equations* indicated as:

$$\begin{cases} a_{1,1}x_1 + a_{1,2}x_2 + \ldots + a_{1,n}x_n = b_1 \\ a_{2,1}x_1 + a_{2,2}x_2 + \ldots + a_{2,n}x_n = b_2 \\ \ldots \\ a_{m,1}x_1 + a_{m,2}x_2 + \ldots + a_{m,n}x_n = b_m \end{cases} .$$

Every ordered n-tuple of real numbers substituted for $x_1, x_2, \ldots x_n$ that make the system of linear equations true is said to be a *solution*.

Systems of Linear Equations

A system can be written as a matrix equation $\mathbf{Ax} = \mathbf{b}$ where

$$\mathbf{A} = \begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} \\ \dots & \dots & \dots & \dots \\ a_{m,1} & a_{m,2} & \dots & a_{m,n} \end{pmatrix}$$

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ \dots \\ b_m \end{pmatrix}$$

Complete Matrix

The coefficient matrix $\mathbf{A}$ is said *incomplete matrix*. The matrix $\mathbf{A^c} \in \mathbb{R}_{m,n+1}$ whose first n columns are those of the matrix $\mathbf{A}$ and the $(n+1)^{th}$ column is the vector $\mathbf{b}$ is said *complete matrix*:

$$\mathbf{A^c} = (\mathbf{A}|\mathbf{b}) = \left(\begin{array}{cccc|c} a_{1,1} & a_{1,2} & \dots & a_{1,n} & b_1 \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} & b_2 \\ \dots & \dots & \dots & \dots & \dots \\ a_{m,1} & a_{m,2} & \dots & a_{m,n} & b_m \end{array}\right).$$

System of Linear Equations

Example

Consider the following system of linear equations:

$$\begin{cases} 2x - y + z = 3 \\ x + 2z = 3 \\ x - y = 1 \end{cases}.$$

The system can be re-written as $\mathbf{A}\mathbf{x} = \mathbf{b}$ where

$$\mathbf{A} = \begin{pmatrix} 2 & -1 & 1 \\ 1 & 0 & 2 \\ 1 & -1 & 0 \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} 3 \\ 3 \\ 1 \end{pmatrix}, \quad \mathbf{A}^{\mathbf{c}} = \left(\begin{array}{ccc|c} 2 & -1 & 1 & 3 \\ 1 & 0 & 2 & 3 \\ 1 & -1 & 0 & 1 \end{array}\right)$$

Exercise 1

Consider a complete matrix $\left(\begin{array}{cccc|c} 2 & 0 & -1 & 0 & 2 \\ 0 & 0 & 2 & -1 & 1 \\ 3 & 2 & 1 & 0 & 0 \end{array}\right)$.

Write as a system of linear equations in variables x_1, x_2, x_3, x_4.

Answer:

$$2x_1 - x_3 = 2$$
$$2x_3 - x_4 = 1$$
$$3x_1 + 2x_2 + x_3 = 0$$

Solving a system of linear equations: Cramer’s Method

Cramer's Theorem

Theorem

Cramer's Theorem.
Let us consider a system of n linear equations in n variables,

$$\mathbf{A}\mathbf{x} = \mathbf{b}.$$

If $\mathbf{A}$ is non-singular, there is only one solution simultaneously satisfying all the equations:

if $\det \mathbf{A} \neq 0$*, then* $\exists! \mathbf{x}$ *such that* $\mathbf{A}\mathbf{x} = \mathbf{b}$.

Definition

A system of linear equations that satisfies the hypotheses of the Cramer's Theorem is said a *Cramer system*.

Proof.

Let us consider the system $\mathbf{Ax} = \mathbf{b}$. If $\mathbf{A}$ is non-singular the matrix $\mathbf{A}$ is invertible, i.e. a matrix $\mathbf{A}^{-1}$ exists (see Lecture 3) .

Let us multiply $\mathbf{A}^{-1}$ by the equation representing the system:

$$\begin{aligned} & \mathbf{A}^{-1}(\mathbf{Ax}) = \mathbf{A}^{-1}\mathbf{b} \\ \Rightarrow \quad & \left(\mathbf{A}^{-1}\mathbf{A}\right)\mathbf{x} = \mathbf{A}^{-1}\mathbf{b} \\ \Rightarrow \quad & \mathbf{Ix} = \mathbf{A}^{-1}\mathbf{b} \\ \Rightarrow \quad & \mathbf{x} = \mathbf{A}^{-1}\mathbf{b} \end{aligned}$$

The inverse matrix $\mathbf{A}^{-1}$ is unique and thus also the vector $\mathbf{x}$ is unique, i.e. the only one solution solving the system exists. □

Hybrid Matrix

Definition

Let us consider a system of linear equations $\mathbf{Ax} = \mathbf{b}$ as defined above.

The *hybrid matrix* with respect to the i^{th} column is the matrix $\mathbf{A_i}$ obtained from $\mathbf{A}$ by substituting the i^{th} column with $\mathbf{b}$:

$$\mathbf{A_i} = \begin{pmatrix} a_{1,1} & a_{1,2} & \dots & b_1 & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & b_2 & \dots & a_{2,n} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ a_{n,1} & a_{n,2} & \dots & b_n & \dots & a_{n,n} \end{pmatrix}.$$

Equivalently if we write $\mathbf{A}$ as a vector of column vectors:

$$\mathbf{A} = \begin{pmatrix} \mathbf{a^1} & \mathbf{a^1} & \ldots & \mathbf{a^{i-1}} & \mathbf{a^i} & \mathbf{a^{i+1}} & \ldots & \mathbf{a^n} \end{pmatrix}$$

the hybrid matrix $\mathbf{A_i}$ would be

$$\mathbf{A_i} = \begin{pmatrix} \mathbf{a^1} & \mathbf{a^1} & \ldots & \mathbf{a^{i-1}} & \mathbf{b} & \mathbf{a^{i+1}} & \ldots & \mathbf{a^n} \end{pmatrix}.$$

Cramer's Method

Theorem

Cramer's Method *For a given system of linear equations* $\mathbf{Ax} = \mathbf{b}$ *with* $\mathbf{A}$ *non-singular, a generic solution* x_i *element of* $\mathbf{x}$ *can be computed as*

$$x_i = \frac{\det \mathbf{A_i}}{\det \mathbf{A}}$$

where $\mathbf{A_i}$ *is the hybrid matrix with respect to the* i^{th} *column.*

Cramer's Method

Proof.

Let us consider a system of linear equations $\mathbf{A}\mathbf{x} = \mathbf{b}$.

We can compute $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$:

$$\begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{pmatrix} = \frac{1}{\det \mathbf{A}} \begin{pmatrix} A_{1,1} & A_{2,1} & \dots & A_{n,1} \\ A_{1,2} & A_{2,2} & \dots & A_{n,2} \\ \dots & \dots & \dots & \dots \\ A_{1,n} & A_{2,n} & \dots & A_{n,n} \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ \dots \\ b_n \end{pmatrix}$$

$$\Rightarrow \begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{pmatrix} = \frac{1}{\det \mathbf{A}} \begin{pmatrix} A_{1,1}b_1 + A_{2,1}b_2 + \dots + A_{n,1}b_n \\ A_{1,2}b_1 + A_{2,2}b_2 + \dots + A_{n,2}b_n \\ \dots \\ A_{1,n}b_1 + A_{2,n}b_2 + \dots + A_{n,n}b_n \end{pmatrix}.$$

continued...

Proof.

Notice that cofactors for b_j are same as those for $a_{j,i}$.

From the I Laplace Theorem (see Lecture 3), the vector of solutions can be written as:

$$\begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{pmatrix} = \frac{1}{\det \mathbf{A}} \begin{pmatrix} \det \mathbf{A_1} \\ \det \mathbf{A_2} \\ \dots \\ \det \mathbf{A_n} \end{pmatrix}.$$

□

Example

Solve the following system by inverting the coefficient matrix:

$$\begin{cases} 2x - y + z = 3 \\ x + 2z = 3 \\ x - y = 1 \end{cases}.$$

The system can be re-written as a matrix equation $\mathbf{Ax} = \mathbf{b}$ where

$$\mathbf{A} = \begin{pmatrix} 2 & -1 & 1 \\ 1 & 0 & 2 \\ 1 & -1 & 0 \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} 3 \\ 3 \\ 1 \end{pmatrix}$$

continued...

Example

So $\det \mathbf{A} = 2 \times 2 + 1 \times -2 + 1 \times -1 = 1$ using I Laplace Theorem on row 1.

Using Cramer's Method:

$$x_1 = \frac{\det \mathbf{A_1}}{\det \mathbf{A}} = \det \begin{pmatrix} 3 & -1 & 1 \\ 3 & 0 & 2 \\ 1 & -1 & 0 \end{pmatrix} /1 = 1,$$

$$x_2 = \det \begin{pmatrix} 2 & 3 & 1 \\ 1 & 3 & 2 \\ 1 & 1 & 0 \end{pmatrix} = 0 \text{ and } x_3 = \det \begin{pmatrix} 2 & -1 & 3 \\ 1 & 0 & 3 \\ 1 & -1 & 1 \end{pmatrix} = 1.$$

Exercise 2

Solve the following system of linear equations using the inverse of the incomplete matrix.

$$2x_1 + x_2 = 1$$
$$4x_1 - x_2 = 5$$

Answer

Incomplete matrix $\mathbf{A} = \begin{pmatrix} 2 & 1 \\ 4 & -1 \end{pmatrix}$ and $\mathbf{b} = \begin{pmatrix} 1 \\ 5 \end{pmatrix}$ gives

$\mathbf{x} = \mathbf{A}^{-1}\mathbf{b} = \frac{1}{2\times-1-4\times1} \begin{pmatrix} -1 & -1 \\ -4 & 2 \end{pmatrix} \begin{pmatrix} 1 \\ 5 \end{pmatrix} = \frac{1}{-6} \begin{pmatrix} -6 \\ 6 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$

Hence $x_1 = 1$ and $x_2 = -1$.

Rouchè-Capelli Theorem

Number of solutions for systems of linear equations

Definition

A system of m linear equations in n variables is said

- *compatible* if it has at least one solution,
- *determined* if it has only one solution,
- *undetermined* if it has infinite solutions, and
- *incompatible* if it has no solutions.

Rouchè-Capelli Theorem

Theorem

Rouchè-Capelli Theorem (Kronecker-Capelli Theorem)
A system of m linear equations in n variables $\mathbf{Ax} = \mathbf{b}$ is compatible if and only if both the incomplete and complete matrices ($\mathbf{A}$ and $\mathbf{A^c}$ respectively) are characterised by the same rank $\rho_{\mathbf{A}} = \rho_{\mathbf{A^c}} = \rho$ named rank of the system.

Proof not given (it requires some further concepts about vector spaces).

Rouchè-Capelli Theorem: Cases

- If $\rho_{\mathbf{A}} < \rho_{\mathbf{A}^c}$ the system is incompatible and thus it has no solutions.
- If $\rho_{\mathbf{A}} = \rho_{\mathbf{A}^c}$ the system is compatible. Under these conditions, three cases can be identified.
  - **case 1:** If $\rho_{\mathbf{A}} = \rho_{\mathbf{A}^c} = \rho = n = m$, the system is a Cramer's system and can be solved by the Cramer's method.
  - **case 2:** If $\rho_{\mathbf{A}} = \rho_{\mathbf{A}^c} = \rho = n < m$, ρ equations of the system compose a Cramer's system (and as such has only one solution). The remaining $m - \rho$ equations are a linear combination of the other, these equations are redundant and the system has only one solution.
  - **case 3:** If $\rho_{\mathbf{A}} = \rho_{\mathbf{A}^c} = \rho \begin{cases} < n \\ \leq m \end{cases}$, the system is undetermined and has $\infty^{n-\rho}$ solutions.

Example

Let us consider the following system of linear equations:

$$\begin{cases} 3x_1 + 2x_2 + x_3 = 1 \\ x_1 - x_2 = 2 \\ 2x_1 + x_3 = 4 \end{cases} .$$

The incomplete and complete matrices associated with this system are:

$$\mathbf{A} = \begin{pmatrix} 3 & 2 & 1 \\ 1 & -1 & 0 \\ 2 & 0 & 1 \end{pmatrix} \text{and } \mathbf{A^c} = \left(\begin{array}{ccc|c} 3 & 2 & 1 & 1 \\ 1 & -1 & 0 & 2 \\ 2 & 0 & 1 & 4 \end{array}\right)$$

continued...

Example

The $\det(\mathbf{A}) = -3$. Hence, the rank $\rho_{\mathbf{A}} = 3$.
It follows that $\rho_{\mathbf{A^c}} = 3$ since a non-singular 3×3 submatrix can be extracted ($\mathbf{A}$) and a 4×4 submatrix cannot be extracted since the size of $\mathbf{A^c}$ is 3×4.
Hence, $\rho_{\mathbf{A}} = \rho_{\mathbf{A^c}} = m = n = 3$ (case 1). The system can be solved by Cramer's Method.

continued...

Rouchè-Capelli Theorem

Example

Only one solution exists and is:

$$x_1 = \frac{\det\begin{pmatrix} 1 & 2 & 1 \\ 2 & -1 & 0 \\ 4 & 0 & 1 \end{pmatrix}}{-3} = \frac{1}{3}$$

$$x_2 = \frac{\det\begin{pmatrix} 3 & 1 & 1 \\ 1 & 2 & 0 \\ 2 & 4 & 1 \end{pmatrix}}{-3} = -\frac{5}{3}$$

$$x_3 = \frac{\det\begin{pmatrix} 3 & 2 & 1 \\ 1 & -1 & 2 \\ 2 & 0 & 4 \end{pmatrix}}{-3} = \frac{10}{3}.$$

Exercise 3

Consider these three complete matrices. For each one, determine whether the corresponding system of linear equations is compatible, determined, undetermined or incompatible.

$$\mathbf{A}^c = \left(\begin{array}{cc|c} 2 & 1 & 1 \\ 4 & -1 & 5 \end{array}\right), \quad \mathbf{B}^c = \left(\begin{array}{ccc|c} 3 & 2 & 1 & 1 \\ 1 & 0 & -2 & 0 \\ 1 & 2 & 5 & 3 \end{array}\right),$$

$$\mathbf{C}^c = \left(\begin{array}{ccc|c} 3 & 0 & 1 & 1 \\ 2 & 1 & 0 & 2 \end{array}\right)$$

Exercise 3: Solution

- $\mathbf{A^c}$: For both incomplete and complete matrices, the largest square submatrix is 2×2. The first 2×2 matrix has non-zero determinant $(= -6)$, hence both $\rho_{\mathbf{A}} = \rho_{\mathbf{A^c}} = 2$ and so by Rouchè-Capelli Theorem, the system of linear equations is **compatible**. Since $\rho_{\mathbf{A^c}} = n = m = 2$, from Cramer's Theorem, there is a unique solution, hence the system is also **determined**.
- $\mathbf{B^c}$: The three rows of the incomplete matrix are linearly dependent $(\mathbf{b_1} - 2\mathbf{b_2} - \mathbf{b_3} = \mathbf{o})$, hence $\rho_{\mathbf{B}} < 3$. But the last three columns of the complete matrix form a 3×3 submatrix with non-zero determinant $(= -8)$ hence $\rho_{\mathbf{B^c}} = 3$. Since $\rho_{\mathbf{B}} < \rho_{\mathbf{B^c}}$, by Rouchè-Capelli Theorem, the system is **incompatible**.

Exercise 3: Solution

- $\mathbf{C^c}$: For both incomplete and complete matrices, the largest square submatrix is 2×2. The first 2×2 matrix has non-zero determinant $(= 3)$, hence both $\rho_{\mathbf{C}} = \rho_{\mathbf{C^c}} = 2$ and so by Rouchè-Capelli Theorem, the system of linear equations is **compatible**. Since $\rho_{\mathbf{A^c}} = 2 < n$, the system is **undetermined** (see case 3 on slide 23).

Gaussian Elimination

Direct Methods

Let us consider a Cramer's system $\mathbf{Ax} = \mathbf{b}$ where $\mathbf{A} \in \mathbb{R}_{n,n}$.

- The solution of this system can be laborious indeed as, by applying the Cramer's Theorem (matrix inversion), it would require the calculation of one determinant of a n order matrix and n^2 determinants of $n-1$ order matrices.
- The application of Cramer's Method, would require the calculation of one determinant of a n order matrix and n determinants of n order matrices.
  A determinant is the sum of $n!$ terms where each term is the result of a multiplication [Neri 2019, chapter 2].
- Hence we consider direct methods and present Gaussian Elimination in this section.

Solving by eliminating rows and substitution

Example

Consider again the system of linear equations:

$$\begin{cases} 3x_1 + 2x_2 + x_3 = 1 \\ x_1 - x_2 = 2 \\ 2x_1 + x_3 = 4 \end{cases} .$$

First: we subtract $\frac{1}{3}$ of equation 1 from equation 2 and $\frac{2}{3}$ of equation 1 from equation 3:-

$$\begin{cases} 3x_1 + 2x_2 + x_3 = 1 \\ -x_2 - \frac{2}{3}x_2 - \frac{1}{3}x_3 = 2 - \frac{1}{3} \\ -\frac{2}{3}2x_2 + x_3 - \frac{2}{3}x_3 = 4 - \frac{2}{3} \end{cases} .$$

continued...

Solving by eliminating rows and substitution

Example

Rewrite equations 2 and 3 (after subtracting):

$$\begin{cases} -\frac{5}{3}x_2 - \frac{1}{3}x_3 = \frac{5}{3} \\ -\frac{4}{3}x_2 + \frac{1}{3}x_3 = \frac{10}{3} \end{cases}.$$

Second: we subtract $\frac{4}{5}$ of first line from second:

$$\frac{1}{3}x_3 + \frac{4}{15}x_3 = \frac{10}{3} - \frac{4}{3}.$$

Solve for x_3 gives $x_3 = \frac{10}{3}$.
Substitute this back into first line to get $x_2 = -\frac{5}{3}$.
Substitute both back into equation 1 to get $x_1 = \frac{1}{3}$.

continued...

Triangular and Staircase Matrices

Example

After the two substitution steps the incomplete and complete matrices are, respectively:

$$\begin{pmatrix} 3 & 2 & 1 \\ 0 & -\frac{5}{3} & -\frac{1}{3} \\ 0 & 0 & \frac{9}{15} \end{pmatrix} \text{and} \left(\begin{array}{ccc|c} 3 & 2 & 1 & 1 \\ 0 & -\frac{5}{3} & -\frac{1}{3} & \frac{5}{3} \\ 0 & 0 & \frac{9}{15} & 2 \end{array}\right)$$

- Notice the bottom left triangles of zeroes.
- The incomplete matrix is an example of a *triangular matrix*.
- The complete matrix is an example of a *staircase matrix*.
- They make it possible to solve this problem by substitution.

Gaussian Elimination

Gaussian Elimination is an algorithm to formalize this process of elimination and substitution:

- Construct the complete matrix $\mathbf{A^c}$;
- Apply elementary row operations to obtain a staircase complete matrix and triangular incomplete matrix;
- Write down the new system of linear equations;
- Solve the n^{th} equation of the system and use the result to solve the $(n-1)^{th}$;
- Continue recursively until the first equation.

Elementary row operations

Definition

Let $\mathbf{A} \in \mathbb{R}_{m,n}$. The following operations on the matrix $\mathbf{A}$ are said *elementary row operations*:

- E1: swap of two rows $\mathbf{a_i}$ and $\mathbf{a_j}$:

$$\mathbf{a_i} \leftarrow \mathbf{a_j}$$
$$\mathbf{a_j} \leftarrow \mathbf{a_i}$$

- E2: multiplication of a row $\mathbf{a_i}$ by a non-zero scalar $\lambda \in \mathbb{R}$:

$$\mathbf{a_i} \leftarrow \lambda \mathbf{a_i}$$

- E3: substitution of a row $\mathbf{a_i}$ by the sum of the row $\mathbf{a_i}$ to another row $\mathbf{a_j}$:

$$\mathbf{a_i} \leftarrow \mathbf{a_i} + \mathbf{a_j}$$

Equivalent Matrices and Equivalent Systems

Definition

Equivalent Matrices Let us consider a matrix $\mathbf{A} \in \mathbb{R}_{m,n}$.
If we apply the elementary row operations on $\mathbf{A}$ we obtain a new matrix $\mathbf{C} \in \mathbb{R}_{m,n}$. This matrix is said *equivalent* to $\mathbf{A}$.

Definition

Equivalent Systems Let us consider two systems of linear equations in the same variables: $\mathbf{Ax} = \mathbf{b}$ and $\mathbf{Cx} = \mathbf{d}$. These two systems are *equivalent* if they have the same solutions.

Equivalent Matrices and Equivalent Systems

Theorem

Let us consider a system of m linear equations in n variables $\mathbf{A}\mathbf{x} = \mathbf{b}$. Let $\mathbf{A}^{\mathbf{c}} \in \mathbb{R}_{m,n+1}$ be the complete matrix associated with this system.

If another system of linear equations is associated with a complete matrix $\mathbf{A}'^{\mathbf{c}} \in \mathbb{R}_{m,n+1}$ equivalent to $\mathbf{A}^{\mathbf{c}}$, then the two systems are also equivalent.

Proof.

By following the definition of equivalent matrices, if $\mathbf{A'^c}$ is equivalent to $\mathbf{A^c}$, then $\mathbf{A'^c}$ can be generated from $\mathbf{A^c}$ by applying the elementary row operations.

Each operation of the complete matrix obviously has a meaning in the system of linear equations. Let us analyse the effect of the elementary row operations on the complete matrix.

- When E1 is applied, i.e. the swap of two rows, the equations of the system are swapped. This operation has no effect on the solution of the system. Thus after E1 operation the modified system is equivalent to the original one.

continued...

Proof.

- When E2 is applied, i.e. a row is multiplied by a non-null scalar λ, a scalar is multiplied to all the terms of the equation. In this case the equation $a_{i,1}x_1 + a_{i,2}x_2 + \ldots + a_{i,n}x_n = b_i$ is substituted by

$$\lambda a_{i,1}x_1 + \lambda a_{i,2}x_2 + \ldots + \lambda a_{i,n}x_n = \lambda b_i.$$

  The two equations have the same solutions and thus after E2 operation the modified systems is equivalent to the original one.

continued...

Proof.

- When E3 is applied, i.e. a row is added to another row, the equation $a_{i,1}x_1 + a_{i,2}x_2 + \ldots + a_{i,n}x_n = b_i$ is substituted by the equation *[*]*:

$$\left(a_{i,1} + a_{j,1}\right) x_1 + \left(a_{i,2} + a_{j,2}\right) x_2 + \ldots + \left(a_{i,n} + a_{j,n}\right) x_n = b_i + b_j$$

If the n-tuple $y_1, y_2, \ldots, y_n$ is the solution of the original system it is the solution of

$$a_{i,1}x_1 + a_{i,2}x_2 + \ldots + a_{i,n}x_n = b_i \text{ and}$$
$$a_{j,1}x_1 + a_{j,2}x_2 + \ldots + a_{j,n}x_n = b_j.$$

Thus, $y_1, y_2, \ldots, y_n$ also verifies equation *[*]* above.
Thus, after E3 operation the modified system is equivalent in solutions to the original one.

□

Row Vector Notation for Gaussian Elimination

Let us write an equivalent formulation of the Gaussian transformation using row vector notation.

Consider a system of linear equations in a matrix formulation: $\mathbf{Ax} = \mathbf{b}$ where $\mathbf{A} \in \mathbb{R}_{n,n}$.

Write the complete matrix $\mathbf{A^c}$ in terms of its row vectors,

$$\mathbf{A^{c(1)}} = \begin{pmatrix} \mathbf{r_1^{(1)}} \\ \mathbf{r_2^{(1)}} \\ \dots \\ \mathbf{r_n^{(1)}} \end{pmatrix}.$$

The superscript (1) emphasizes that this is step one of the algorithm.

Row Vector Notation for Gaussian Elimination

The Gaussian Elimination algorithm works for steps 2 to n.

At the generic step $(k+1)$ the Gaussian transformation formulas are

$$
\begin{aligned}
\mathbf{r_1^{(k+1)}} &= \mathbf{r_1^{(k)}} \\
\mathbf{r_2^{(k+1)}} &= \mathbf{r_2^{(k)}} \\
&\ldots \\
\mathbf{r_k^{(k+1)}} &= \mathbf{r_k^{(k)}} \\
\mathbf{r_{k+1}^{(k+1)}} &= \mathbf{r_{k+1}^{(k)}} + \left(\frac{-a_{k+1,k}^{(k)}}{a_{k,k}^{(k)}} \right) \mathbf{r_k^{(k)}} \\
\mathbf{r_{k+2}^{(k+1)}} &= \mathbf{r_{k+2}^{(k)}} + \left(\frac{-a_{k+2,k}^{(k)}}{a_{k,k}^{(k)}} \right) \mathbf{r_k^{(k)}} \\
&\ldots \\
\mathbf{r_n^{(k+1)}} &= \mathbf{r_n^{(k)}} + \left(\frac{-a_{n,k}^{(k)}}{a_{k,k}^{(k)}} \right) \mathbf{r_k^{(k)}}
\end{aligned}
$$

Row Vector Notation for Gaussian Elimination

Since all transformations involve just elementary row operations, it follows that all complete matrices given as

$$\mathbf{A}^{\mathbf{c}(\mathbf{k})} = \begin{pmatrix} \mathbf{r}_1^{(\mathbf{k})} \\ \mathbf{r}_2^{(\mathbf{k})} \\ \dots \\ \mathbf{r}_\mathbf{n}^{(\mathbf{k})} \end{pmatrix}$$

are equivalent. The theorem we proved earlier shows that they also represent equivalent systems.
Hence $\mathbf{A}^{\mathbf{c}(\mathbf{n})}$ is a staircase matrix which can easily be solved by substitution and gives the solution to the original system $\mathbf{A}\mathbf{x} = \mathbf{b}$.

Gaussian Elimination

Example

Let us apply the Gaussian elimination to solve the following system of linear equations:

$$\begin{cases} x_1 - x_2 - x_3 + x_4 = 0 \\ 2x_1 + 2x_3 = 8 \\ -x_2 - 2x_3 = -8 \\ 3x_1 - 3x_2 - 2x_3 + 4x_4 = 7 \end{cases}.$$

The associated complete matrix is

$$\mathbf{A^{c(1)}} = (\mathbf{A}|\mathbf{b}) = \left(\begin{array}{cccc|c} 1 & -1 & -1 & 1 & 0 \\ 2 & 0 & 2 & 0 & 8 \\ 0 & -1 & -2 & 0 & -8 \\ 3 & -3 & -2 & 4 & 7 \end{array}\right).$$

continued...

Example

Let us apply the Gaussian transformations to move to step (2):

$$
\begin{aligned}
\mathbf{r}_1^{(2)} &= \mathbf{r}_1^{(1)} \\
\mathbf{r}_2^{(2)} &= \mathbf{r}_2^{(1)} + \left(\frac{-a_{2,1}^{(1)}}{a_{1,1}^{(1)}}\right)\mathbf{r}_1^{(1)} = \mathbf{r}_2^{(1)} - 2\mathbf{r}_1^{(1)} \\
\mathbf{r}_3^{(2)} &= \mathbf{r}_3^{(1)} + \left(\frac{-a_{3,1}^{(1)}}{a_{1,1}^{(1)}}\right)\mathbf{r}_1^{(1)} = \mathbf{r}_3^{(1)} + 0\mathbf{r}_1^{(1)} \\
\mathbf{r}_4^{(2)} &= \mathbf{r}_4^{(1)} + \left(\frac{-a_{4,1}^{(1)}}{a_{1,1}^{(1)}}\right)\mathbf{r}_1^{(1)} = \mathbf{r}_4^{(1)} - 3\mathbf{r}_1^{(1)}
\end{aligned}
$$

thus obtaining the following complete matrix:

$$
\mathbf{A}^{\mathbf{c}(2)} = \left(\begin{array}{cccc|c} 1 & -1 & -1 & 1 & 0 \\ 0 & 2 & 4 & -2 & 8 \\ 0 & -1 & -2 & 0 & -8 \\ 0 & 0 & 1 & 1 & 7 \end{array}\right).
$$

continued...

Example

Let us apply the Gaussian transformations to move to step (3):

$$
\begin{aligned}
\mathbf{r_1^{(3)}} &= \mathbf{r_1^{(2)}} \\
\mathbf{r_2^{(3)}} &= \mathbf{r_2^{(2)}} \\
\mathbf{r_3^{(3)}} &= \mathbf{r_3^{(2)}} + \left(\frac{-a_{3,2}^{(2)}}{a_{2,2}^{(2)}}\right)\mathbf{r_2^{(2)}} = \mathbf{r_3^{(2)}} + \frac{1}{2}\mathbf{r_2^{(2)}} \\
\mathbf{r_4^{(3)}} &= \mathbf{r_4^{(2)}} + \left(\frac{-a_{4,2}^{(2)}}{a_{2,2}^{(2)}}\right)\mathbf{r_2^{(2)}} = \mathbf{r_4^{(2)}} + 0\mathbf{r_2^{(2)}}
\end{aligned}
$$

thus obtaining the following complete matrix

$$
\mathbf{A}^{\mathbf{c}(3)} = \left(\begin{array}{cccc|c}
1 & -1 & -1 & 1 & 0 \\
0 & 2 & 4 & -2 & 8 \\
0 & 0 & 0 & -1 & -4 \\
0 & 0 & 1 & 1 & 7
\end{array}\right).
$$

continued...

Example

We would need one more step to obtain a triangular matrix. However, in this case, after two steps the matrix is already triangular. It is enough to swap the third and fourth rows to obtain

$$\mathbf{A}^{\mathbf{c}(4)} = \left(\begin{array}{cccc|c} 1 & -1 & -1 & 1 & 0 \\ 0 & 2 & 4 & -2 & 8 \\ 0 & 0 & 1 & 1 & 7 \\ 0 & 0 & 0 & -1 & -4 \end{array}\right).$$

continued...

Example

Working backwards from row 4 to 1:

- The last row $(0\ 0\ 0\ -1 \mid -4)$ represents the linear equation $-x_4 = -4$, hence $x_4 = 4$.
- Then the 3rd row $(0\ 0\ 1\ 1 \mid 7)$ represents $x_3 + x_4 = 7$. Substitute $x_4 = 4$ to get $x_3 = 3$.
- The 2nd row $(0\ 2\ 4\ -2 \mid 8)$ represents $2x_2 + 4x_3 - 2x_4 = 8$. Substituting $x_3 = 3$ and $x_4 = 4$ gives $x_2 = 2$.
- The 1st row $(1\ -1\ -1\ 1 \mid 0)$ represents $x_1 - x_2 - x_3 + x_4 = 0$. Substituting $x_2 = 2$, $x_3 = 3$, $x_4 = 4$ gives $x_1 = 1$.

Therefore $x_1 = 1$, $x_2 = 2$, $x_3 = 3$, and $x_4 = 4$.

Exercise 4

Solve the following system of linear equations using the Gaussian Elimination Algorithm. Show your working using row vector notation.

$$
\begin{aligned}
2x_1 - x_2 + 3x_3 &= 1 \\
x_1 + x_2 - 2x_3 &= 4 \\
3x_1 - 2x_2 - x_3 &= 7
\end{aligned}
$$

Exercise 4: Solution

The complete matrix is $\left(\begin{array}{ccc|c} 2 & -1 & 3 & 1 \\ 1 & 1 & -2 & 4 \\ 3 & -2 & -1 & 7 \end{array}\right)$.

Then rows are

$$\begin{aligned} \mathbf{r}_1^{(1)} &= (2, -1, 3, 1) \\ \mathbf{r}_2^{(1)} &= (1, 1, -2, 4) \\ \mathbf{r}_3^{(1)} &= (3, -2, -1, 7). \end{aligned}$$

For $k = 1$,

$$\begin{aligned} \mathbf{r}_1^{(2)} &= \mathbf{r}_1^{(1)} \\ \mathbf{r}_2^{(2)} &= \mathbf{r}_2^{(1)} + \tfrac{-1}{2}\mathbf{r}_1^{(1)} = (0, \tfrac{3}{2}, \tfrac{-7}{2}, \tfrac{7}{2}) \\ \mathbf{r}_3^{(2)} &= \mathbf{r}_3^{(1)} + \tfrac{-3}{2}\mathbf{r}_1^{(1)} = (0, \tfrac{-1}{2}, \tfrac{-11}{2}, \tfrac{11}{2}). \end{aligned}$$

Exercise 4: Solution

For $k = 2$,

$$\begin{aligned}
\mathbf{r_1^{(3)}} &= \mathbf{r_1^{(2)}} \\
\mathbf{r_2^{(3)}} &= \mathbf{r_2^{(2)}} \\
\mathbf{r_3^{(3)}} &= \mathbf{r_3^{(2)}} + \tfrac{1}{3}\mathbf{r_2^{(2)}} = (0, 0, \tfrac{-20}{3}, \tfrac{20}{3})
\end{aligned}$$

which gives the complete matrix $\left(\begin{array}{ccc|c} 2 & -1 & 3 & 1 \\ 0 & \frac{3}{2} & \frac{-7}{2} & \frac{7}{2} \\ 0 & 0 & \frac{-20}{3} & \frac{20}{3} \end{array}\right)$.

Working backwards from row 3, we get $x_3 = -1$, $x_2 = 0$ and $x_1 = 2$.

Summary of Methods

Iterative Methods

- Methods such as **Jacobi's Method**, starting from an initial guess $\mathbf{x}^{(0)}$, iteratively apply some formulas to detect the solution of the system.
- For this reason, these methods are often indicated as *iterative methods*.
- Unlike direct methods that converge to the theoretical solution in a finite time, iterative methods are *approximate* since they *converge*, under some conditions.
- We will not have time to cover them in this course.

Synopsis of methods for solving linear systems

	Cramer (Rouchè-Capelli)	Direct	Iterative
Operational feature	Determinant	Manipulate matrix	Manipulate guess solution
Outcome	Exact solution	Exact solution	Approximate solution
Computational cost	Unacceptably high ($\mathcal{O}(n!)$)	High $\mathcal{O}\left(n^3\right)$	∞ to the exact solution but it can be stopped after k steps with $k \cdot \mathcal{O}\left(n^2\right)$
Practical usability	very small matrices (up to approx. 10×10)	medium matrices	large matrices (up to approx 1000×1000)
Hypothesis	No hypothesis	$a_{kk}^{(k)} \neq 0$ (solvable by pivoting)	Conditions on the eigenvalues of the matrix

Summary and next lecture

Summary

- Definition of a System of Linear Equations
- Cramer's Method
- Rouchè-Capelli Theorem
- Gaussian Elimination

The next lecture

We will learn about vector spaces.