Solving systems of linear equations

1. Assemble the augmented matrix of the system.
2. By use of elementary equivalent row transforms convert the matrix to the row echelon form.
3. By the backward substitution describe *all* solutions.

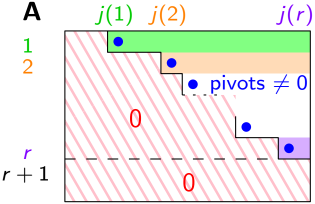
Definition: A matrix $\boldsymbol{A}$ is in the *row echelon form (REF)* if the nonzero rows are strictly ordered by the number of leading zeroes and the zero rows are below the nonzero ones.

Formally: Denote $j(i) := \min\{j : a_{ij} \neq 0\}$.
A matrix $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ is in REF iff $\exists r \in \{1, \dots, m\}$:

a) $j(1) < j(2) < \cdots < j(r)$,

b) $\forall i > r, \forall j : a_{ij} = 0$.

The first nonzero element $a_{i,j(i)}$ in the i-th row is called the *pivot*.



Naïve algorithm for Gaussian elimination

Input: A matrix $\boldsymbol{A}$
Output: A matrix $\boldsymbol{A}$ in REF
foreach *row* i **do** compute $j(i)$ `/* empty row:` $j(i) = \infty$ `*/`
sort the rows of $\boldsymbol{A}$ by $j(i)$
forever
  if $\exists i : j(i) = j(i+1) < \infty$ **then**
    `/* the` i`-th and the` $(i+1)$`-st rows are nonzero and have the same number of leading` 0`s */`
    add the $-\frac{a_{i+1,j(i)}}{a_{i,j(i)}}$-multiple of the i-th row to the $(i+1)$-st row `/* now:` $a_{i+1,j(i)} = 0$ `*/`
    update $j(i+1)$ and put the $(i+1)$-st row in place
  else
    `/* all nonzero rows have distinct numbers of leading zeroes */`
    return $\boldsymbol{A}$

Finiteness: In each loop the total number of leading zeroes grows.

Example of Gaussian elimination

$$(\boldsymbol{A}|\boldsymbol{b}) = \left(\begin{array}{ccccc|c} 1 & 4 & 3 & 2 & 1 & 1 \\ 2 & 8 & 4 & 0 & 0 & 0 \\ 0 & 0 & 3 & 6 & 9 & 5 \\ 2 & 8 & 7 & 6 & 3 & 3 \end{array}\right) \underset{\substack{\text{IV} \\ \text{III}}}{\sim} \left(\begin{array}{ccccc|c} 1 & 4 & 3 & 2 & 1 & 1 \\ 2 & 8 & 4 & 0 & 0 & 0 \\ 2 & 8 & 7 & 6 & 3 & 3 \\ 0 & 0 & 3 & 6 & 9 & 5 \end{array}\right) \overset{-2\text{I}}{\sim}$$

$$\left(\begin{array}{ccccc|c} 1 & 4 & 3 & 2 & 1 & 1 \\ 0 & 0 & -2 & -4 & -2 & -2 \\ 2 & 8 & 7 & 6 & 3 & 3 \\ 0 & 0 & 3 & 6 & 9 & 5 \end{array}\right) \overset{\text{III}}{\underset{\text{II}}{\sim}} \left(\begin{array}{ccccc|c} 1 & 4 & 3 & 2 & 1 & 1 \\ 2 & 8 & 7 & 6 & 3 & 3 \\ 0 & 0 & -2 & -4 & -2 & -2 \\ 0 & 0 & 3 & 6 & 9 & 5 \end{array}\right) \overset{-2\text{I}}{\sim}$$

$$\left(\begin{array}{ccccc|c} 1 & 4 & 3 & 2 & 1 & 1 \\ 0 & 0 & 1 & 2 & 1 & 1 \\ 0 & 0 & -2 & -4 & -2 & -2 \\ 0 & 0 & 3 & 6 & 9 & 5 \end{array}\right) \underset{+2\text{II}}{\sim} \left(\begin{array}{ccccc|c} 1 & 4 & 3 & 2 & 1 & 1 \\ 0 & 0 & 1 & 2 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 3 & 6 & 9 & 5 \end{array}\right) \underset{\substack{\text{IV} \\ \text{III}}}{\sim}$$

$$\left(\begin{array}{ccccc|c} 1 & 4 & 3 & 2 & 1 & 1 \\ 0 & 0 & 1 & 2 & 1 & 1 \\ 0 & 0 & 3 & 6 & 9 & 5 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right) \underset{-3\text{II}}{\sim} \left(\begin{array}{ccccc|c} 1 & 4 & 3 & 2 & 1 & 1 \\ 0 & 0 & 1 & 2 & 1 & 1 \\ 0 & 0 & 0 & 0 & 6 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right) = (\boldsymbol{A}'|\boldsymbol{b}')$$

Augmented matrices $(\boldsymbol{A}|\boldsymbol{b})$ and $(\boldsymbol{A}'|\boldsymbol{b}')$ correspond to systems of equations $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ and $\boldsymbol{A}'\boldsymbol{x} = \boldsymbol{b}'$ with the same sets of solutions.

Backward substitution

Notation: $(\boldsymbol{A}'|\boldsymbol{b}')$ the augmented matrix of a system in REF.

Observe: If $\boldsymbol{b}'$ contains a pivot then the system has no solution.

Definition: For a system $\boldsymbol{A}'\boldsymbol{x} = \boldsymbol{b}'$ with $\boldsymbol{A}'$ in REF, the variables matching the columns with pivots are *leading*, the other are *free*.

Theorem: For $\boldsymbol{A}'\boldsymbol{x} = \boldsymbol{b}'$ with $(\boldsymbol{A}'|\boldsymbol{b}')$ in REF and no pivot in $\boldsymbol{b}'$ *any* choice of free variables can be *uniquely* extended to a solution.

Proof: By induction on $i = r, r-1, \dots, 1$. In the i-th equation:
$0x_1 + \cdots + 0x_{j(i)-1} + a'_{i,j(i)}x_{j(i)} + a'_{i,j(i)+1}x_{j(i)+1} + \cdots + a'_{i,n}x_n = b'_i$
the values all *subsequent* leading variables $x_{j(i+1)}, \dots, x_{j(r)}$
are known by the induction hypothesis (also the values of all free variables), hence the value of $x_{j(i)}$ is uniquely given by the formula:
$x_{j(i)} = \frac{1}{a'_{i,j(i)}}(b'_i - a'_{i,j(i)+1}x_{j(i)+1} - \ldots - a'_{in}x_n)$.

Example of backward substitution

$$(\boldsymbol{A}|\boldsymbol{b}) \sim\sim (\boldsymbol{A}'|\boldsymbol{b}') = \left(\begin{array}{ccccc|c} x_1 & x_2 & x_3 & x_4 & x_5 & \\ 1 & 4 & 3 & 2 & 1 & 1 \\ & & 1 & 2 & 1 & 1 \\ & & & & 6 & 2 \\ & & & & & \end{array}\right)$$

Variables x_1, x_3 and x_5 are leading, while x_2 and x_4 are free.

For any particular values of free variables, e.g. $x_2 = -1$, $x_4 = \frac{1}{3}$, we get the unique solution of $\boldsymbol{Ax} = \boldsymbol{b}$ as follows:

The 3rd equation yields x_5, as: $6x_5 = 2 \quad \Longrightarrow \quad x_5 = \frac{1}{3}$.
The 2nd x_3: $x_3 + 2x_4 + x_5 = x_3 + 2 \cdot \frac{1}{3} + \frac{1}{3} = 1 \quad \Longrightarrow \quad x_3 = 0$.
The 1st x_1: $x_1 + 4x_2 + 3x_3 + 2x_4 + x_5 =$

$$= x_1 + 4 \cdot (-1) + 3 \cdot 0 + 2 \cdot \tfrac{1}{3} + \tfrac{1}{3} = 1 \quad \Longrightarrow \quad x_1 = 4.$$

$$\boldsymbol{x} = \left(4, -1, 0, \tfrac{1}{3}, \tfrac{1}{3}\right)^{\mathsf{T}}$$

Consequences

Corollary: Any solution can be found by the backward substitution.

Proof: Given any solution $\boldsymbol{x}$, the values of leading variables of $\boldsymbol{x}$ are uniquely determined by the free variables of $\boldsymbol{x}$.

Theorem: For any matrix $\boldsymbol{A}$ and any $\boldsymbol{A}'$ in REF s.t. $\boldsymbol{A} \sim\sim \boldsymbol{A}'$ the indices of columns of $\boldsymbol{A}'$ with pivots are determined uniquely by $\boldsymbol{A}$.

Proof: Assume for contrary that $\boldsymbol{A} \sim\sim \boldsymbol{A}' \sim\sim \boldsymbol{A}''$. Let i be the highest index where the character of variables w.r.t. $\boldsymbol{A}'$ and $\boldsymbol{A}''$ differs. Assume w.l.o.g. that x_i is leading in $\boldsymbol{A}'$ and free in $\boldsymbol{A}''$.

For an arbitrary choice of free variables of $\boldsymbol{A}'$ the system $\boldsymbol{A}'\boldsymbol{x} = \mathbf{0}$ yields a unique value of x_i.

Since x_i is free in $\boldsymbol{A}''$, we may choose the free variables for $\boldsymbol{A}''$ the same as above, but the *value of x_i differently*.

This solution of $\boldsymbol{A}''\boldsymbol{x} = \mathbf{0}$ does not solve $\boldsymbol{A}'\boldsymbol{x} = \mathbf{0}$, a contradiction.

Example for the proof argument

We show that for $\boldsymbol{A}' = \begin{pmatrix} 0 & 3 & 0 & 6 \\ 0 & 0 & 1 & 2 \end{pmatrix}$ and $\boldsymbol{A}'' = \begin{pmatrix} 2 & 3 & 0 & 6 \\ 0 & 0 & 1 & 2 \end{pmatrix}$, the systems $\boldsymbol{A}'\boldsymbol{x} = \mathbf{0}$ and $\boldsymbol{A}''\boldsymbol{x} = \mathbf{0}$ have different sets of solutions:

The variable x_4 is free in both, we may choose e.g. $x_4 = 1$.

The variable x_3 is leading in both. From the second equation $x_3 + 2 \cdot 1 = 0$ (identical in both systems) we derive that $x_3 = -2$.

The variable x_2 is leading in $\boldsymbol{A}'$. Its value $x_2 = -2$ is *uniquely* determined from the 1st equation $0x_1 + 3x_2 + 0 \cdot (-2) + 6 \cdot 1 = 0$.

The variable x_2 is free in $\boldsymbol{A}''$, so we may choose it arbitrarily. If we choose a value *distinct* from the unique one from the previous case, e.g. $x_2 = 10$, then this choice *can* be completed to a solution $\boldsymbol{x} = (-18, 10, -2, 1)^\mathsf{T}$ of $\boldsymbol{A}''\boldsymbol{x} = \mathbf{0}$.

In contrast, $\boldsymbol{x} = (-18, 10, -2, 1)^\mathsf{T}$ is not a solution of $\boldsymbol{A}'\boldsymbol{x} = \mathbf{0}$ as it violates its 1st equation.

Matrix rank

Definition: The *rank* of a matrix $\boldsymbol{A}$, denoted by $\text{rank}(\boldsymbol{A})$ is the number of pivots in any $\boldsymbol{A}'$ in REF such that $\boldsymbol{A} \sim\sim \boldsymbol{A}'$.

Theorem: A system $\boldsymbol{Ax} = \boldsymbol{b}$ has a solution if and only if the rank of the matrix $\boldsymbol{A}$ equals as the rank of the augmented matrix $(\boldsymbol{A}|\boldsymbol{b})$.

According to Wikipedia, the theorem is variously known as the:

- Frobenius theorem in the Czech Republic and in Slovakia;
- Rouché–Capelli theorem English speaking countries and in Italy;
- Rouché–Frobenius theorem in Spain and many countries in Latin America;
- Kronecker–Capelli theorem in Austria, Poland, Romania and Russia;
- Rouché–Fontené theorem in France;

while there are several other theorems named after Ferdinand Georg Frobenius (1849–1917).

Photo: Wikipedia

Matrix rank

Definition: The *rank* of a matrix $\boldsymbol{A}$, denoted by $\text{rank}(\boldsymbol{A})$ is the number of pivots in any $\boldsymbol{A}'$ in REF such that $\boldsymbol{A} \sim\sim \boldsymbol{A}'$.

Theorem: A system $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ has a solution if and only if the rank of the matrix $\boldsymbol{A}$ equals as the rank of the augmented matrix $(\boldsymbol{A}|\boldsymbol{b})$.

Proof: Choose any $(\boldsymbol{A}'|\boldsymbol{b}')$ in REF s.t. $(\boldsymbol{A}|\boldsymbol{b}) \sim\sim (\boldsymbol{A}'|\boldsymbol{b}')$.

A solution $\boldsymbol{x}$ exists $\iff$

$\iff$ $\boldsymbol{b}'$ has no pivot

$\iff$ the pivots of $\boldsymbol{A}'$ coincide with the pivots of $(\boldsymbol{A}'|\boldsymbol{b}')$

$\iff$ $\text{rank}(\boldsymbol{A}) = \text{rank}(\boldsymbol{A}|\boldsymbol{b})$,

... because the transformation $\boldsymbol{A} \sim\sim \boldsymbol{A}'$
can be performed by the same elementary
transformations as $(\boldsymbol{A}|\boldsymbol{b}) \sim\sim (\boldsymbol{A}'|\boldsymbol{b}')$.
(Just "forget" the last column.)

Questions to understand the lecture topic

- Can sorting rows in Gaussian elimination be replaced by some other computationally more efficient procedure?
- How many arithmetic operations does Gaussian elimination and backward substitution perform asymptotically?
- How fast can the number of digits in the solution grow during the backward substitution? First, try to construct a system with a matrix of order 10 with numbers $0, \dots, 10$ so that some component of the solution is of the order of billions.
- Would the proof of the theorem on the uniqueness of free and basis variables hold if instead of the homogeneous system $\boldsymbol{Ax} = \mathbf{0}$ a general system $\boldsymbol{Ax} = \boldsymbol{b}$ were used?
- Can it be deduced from the proof of the theorem on the uniqueness of free and leading variables that $\{\boldsymbol{x} : \boldsymbol{A}'\boldsymbol{x} = \mathbf{0}\} \subsetneq \{\boldsymbol{x} : \boldsymbol{A}''\boldsymbol{x} = \mathbf{0}\}$?