Course 12: 07.01.2021

3.7 Linear systems of equations

As usual, throughout this section K will be a field. When needed, we use superior indices to denote vectors in K^n and inferior indices to denote their components. For instance, $x^0 = (x_1^0, \dots, x_n^0) \in K^n$.

Consider a *linear system* of m equations with n unknowns $x_1, \dots, x_n$:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \dots\dots\dots\dots\dots\dots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{cases} \tag{S}$$

where the coefficients are $a_{ij}, b_i \in K$ $(i = 1, \dots, m,\ j = 1, \dots, n)$.

Denote $A = (a_{ij}) \in M_{mn}(K)$. The matrix A is called the *matrix of the system* (S).

Also denote $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ and $b = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$.

Then the system (S) can be written as:

$$A \cdot x = b \tag{S}$$

Furthermore, denote

$$\bar{A} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{pmatrix}.$$

The matrix $\bar{A}$ is called the *extended* (or *augmented*) *matrix of the system* (S).

We know that there exists a bijective correspondence between K-linear maps and matrices. Thus, since $A \in M_{mn}(K)$, there exists $f_A \in Hom_K(K^n, K^m)$ such that $[f_A]_{EE'} = A$, where E and E' are the canonical bases in K^n and K^m respectively.

Denoting $x = (x_1, \dots, x_n) \in K^n$ and $b = (b_1, \dots, b_m) \in K^m$, it follows that

$$[f_A(x)]_{E'} = [f_A]_{EE'} \cdot [x]_E = A \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} = [b]_E \,.$$

It follows that $f_A(x) = b$. Thus, the system (S) can be written as:

$$f_A(x) = b \tag{S}$$

Remark 3.7.1 (1) Thus, for a linear system of equations we have three equivalent forms, namely: the classical one with coefficients and unknowns, the one using matrices and the one using the corresponding linear map.

(2) We have denoted by x and b first column-matrices and then row-matrices to get nicer results, without using any transposed matrices.

If $b = 0$, then the system (S) is called *homogeneous* and it is denoted by (S_0).

Definition 3.7.2 An element $x^0 \in M_{n1}(K)$ $(x^0 \in K^n)$ is called a:
(1) *(particular) solution* of (S) if $A \cdot x^0 = b$ (or equivalently $f_A(x^0) = b$).
(2) *(particular) solution* of (S_0) if $A \cdot x^0 = 0$ (or equivalently $f_A(x^0) = 0$).

Denote the sets of solutions of (S) and (S_0) by

$$S = \{x^0 \in M_{n1}(K) \mid A \cdot x^0 = b\} \quad \text{or} \quad S = \{x^0 \in K^n \mid f_A(x^0) = b\}\,,$$

$$S_0 = \{x^0 \in M_{n1}(K) \mid A \cdot x^0 = 0\} \quad \text{or} \quad S_0 = \{x^0 \in K^n \mid f_A(x^0) = 0\}\,.$$

Theorem 3.7.3 *The set S_0 of solutions of the homogeneous linear system of equations (S_0) is a subspace of the canonical vector space K^n over K and*

$$\dim S_0 = n - \operatorname{rank}(A)\,.$$

Proof. Since

$$S_0 = \{x^0 \in K^n \mid f_A(x^0) = 0\} = \operatorname{Ker} f_A$$

and the kernel of a linear map is always a subspace of the domain vector space, it follows that $S_0 \leq K^n$. Now by the first dimension formula, it follows that

$$\dim S_0 = \dim(\operatorname{Ker} f_A) = \dim K^n - \dim(\operatorname{Im} f_A) = n - \operatorname{rank}(f_A) = n - \operatorname{rank}(A)\,.$$

□

Theorem 3.7.4 *If $x^1 \in S$ is a particular solution of the system (S), then*

$$S = x^1 + S_0 = \{x^1 + x^0 \mid x^0 \in S_0\}\,.$$

Proof. Since $x^1 \in S$, we have $Ax^1 = b$. We are going to prove the requested equality by double inclusion.

First, let $x^2 \in S$. Then

$$Ax^2 = b \Longrightarrow Ax^2 = Ax^1 \Longrightarrow A(x^2 - x^1) = 0 \Longrightarrow x^2 - x^1 \in S_0 \Longrightarrow x^2 \in x^1 + S_0\,.$$

Conversely, let $x^2 \in x^1 + S_0$. Then there exists $x^0 \in S_0$ such that $x^2 = x^1 + x^0$. It follows that

$$Ax^2 = A(x^1 + x^0) = Ax^1 + Ax^0 = b + 0 = b$$

and consequently $x^2 \in S$.

Therefore, $S = x^1 + S_0$. □

Remark 3.7.5 By Theorem 3.7.4, the general solution of the system (S) can be obtained by knowing the general solution of the homogeneous system (S_0) and a particular solution of (S).

In the sequel, we are going to see when a linear system of equations has a solution.

Definition 3.7.6 The system (S) is called *compatible* (or *consistent*) if $S \neq \emptyset$, that is, it has at least one solution.

A compatible system (S) is called *determinate* if $|S| = 1$, that is, it has exactly one solution.

Remark 3.7.7 (1) The system (S) is compatible if and only if $\exists x^0 \in K^n$ such that $f_A(x^0) = b$ if and only if $b \in \operatorname{Im} f_A$.

(2) The system (S_0) is compatible if and only if $\exists x^0 \in K^n$ such that $f_A(x^0) = 0$ if and only if $0 \in \operatorname{Im} f_A$. But the last condition always holds, since $\operatorname{Im} f_A$ is a subspace of K^m. Hence any homogeneous linear system of equations is compatible, having at least the zero (trivial) solution.

Theorem 3.7.8 *The system (S_0) has a non-zero solution if and only if* $\operatorname{rank}(A) < n$*.*

Proof. By Theorem 3.7.3, we have $S_0 = \operatorname{Ker} f_A \neq \{0\} \Longleftrightarrow \dim S_0 \neq 0 \Longleftrightarrow n - \operatorname{rank}(A) \neq 0 \Longleftrightarrow \operatorname{rank}(A) < n$. □

Corollary 3.7.9 *Let $A \in M_n(K)$. Then* $S_0 = \{0\} \Longleftrightarrow \operatorname{rank}(A) = n \Longleftrightarrow \det(A) \neq 0$.

Definition 3.7.10 If $A \in M_n(K)$ and $\det(A) \neq 0$, then the system (S) is called a *Cramer system*.

Theorem 3.7.11 *A Cramer system has a unique solution.*

Proof. The matrix of a Cramer system is an invertible matrix $A \in M_n(K)$. Then we deduce that $x = A^{-1}b$ is the unique solution. □

Corollary 3.7.12 *A homogeneous Cramer system has only the zero solution.*

Let us now give two classical compatibility theorems.

Theorem 3.7.13 (Kronecker-Capelli) *The system (S) is compatible if and only if* $\mathrm{rank}(\bar{A}) = \mathrm{rank}(A)$.

Proof. Let $(e_1, \ldots, e_n)$ be the canonical basis of the canonical vector space K^n over K and denote by $a^1, \ldots, a^n$ the columns of the matrix A. Then we have

$$(S) \text{ is compatible } \iff \exists x^0 \in K^n : \ f_A(x^0) = b \iff b \in \mathrm{Im} f_A \iff$$
$$\iff b \in f_A(\langle e_1, \ldots, e_n \rangle) \iff b \in \langle f_A(e_1), \ldots, f_A(e_n) \rangle \iff b \in \langle a^1, \ldots, a^n \rangle \iff$$
$$\iff \langle a^1, \ldots, a^n, b \rangle = \langle a^1, \ldots, a^n \rangle \iff \dim\langle a^1, \ldots, a^n, b \rangle = \dim\langle a^1, \ldots, a^n \rangle \iff$$
$$\iff \mathrm{rank}(\bar{A}) = \mathrm{rank}(A)\,.$$

□

Definition 3.7.14 A minor d_p of the matrix A is called a *principal determinant* if $d_p \neq 0$ and d_p has the order $\mathrm{rank}(A)$.

We call *characteristic determinants associated to a principal determinant d_p of A* the minors of the extended matrix $\bar{A}$ obtained by completing the matrix of d_p with a column containing the corresponding constants b_i and a row containing the corresponding elements of a row of $\bar{A}$.

Remark 3.7.15 Notice that if the principal determinant d_p has order r, then the characteristic determinants associated to d_p have order $r + 1$.

Now we give without proof the second compatibility theorem.

Theorem 3.7.16 (Rouché) *The system (S) is compatible if and only if all the characteristic determinants associated to a principal determinant are zero.*

3.8 Gauss method

In this section we present a very useful practical method to solve linear systems of equations, called the *Gauss method.*

In the sequel, suppose that $m \leq n$, that is, we talk about systems with less equations than unknowns. In fact, this is the interesting case.

The Gauss method consists of the following steps:

1. *Write the extended matrix $\bar{A}$ of the system (S).*

2. *Apply elementary operations on rows for $\bar{A}$ to get to an echelon form A'.*

3. *Use the Kronecker-Capelli Theorem to decide if the system is compatible or not.*

4. *If compatible, write and solve the system corresponding to the echelon form, starting with the last equation.*

Remark 3.8.1 (1) Actually, the Gauss method simulates working with equations. When we apply an elementary operation on the rows of $\bar{A}$, say multiply a row by a scalar and add it to another row, in fact we multiply an equation by a scalar and add it to another equation. That is why it is important to apply elementary operations only on rows, in order not to interchange the order of the unknowns.

(2) The initial system and the system corresponding to the echelon form are equivalent, that is, they have the same solutions. The great advantage is that the last system can be easily solved, starting with the last equation.

(3) The Gauss method includes checking compatibility, done by the Kronecker-Capelli Theorem.

(4) If the system is compatible, we have a principal determinant of order $r = rank(\bar{A}) = rank(A)$ and it is possible to continue the procedure on the matrix A' to get to a diagonal form having r elements on the principal diagonal and all the other elements zero. Then, when writing the equivalent system, in fact we directly get the solution. This completion of the Gauss method is called the *Gauss-Jordan method.*

Example 3.8.2 (a) Consider the system

$$\begin{cases} x+y-z=2 \\ 3x+2y-2z=6 \\ -x+y+z=0 \end{cases}$$

with real coefficients. Then its extended matrix is

$$\bar{A}=\begin{pmatrix} 1 & 1 & -1 & 2 \\ 3 & 2 & -2 & 6 \\ -1 & 1 & 1 & 0 \end{pmatrix} \sim \begin{pmatrix} 1 & 1 & -1 & 2 \\ 0 & -1 & 1 & 0 \\ 0 & 2 & 0 & 2 \end{pmatrix} \sim \begin{pmatrix} 1 & 1 & -1 & 2 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & 2 & 2 \end{pmatrix}.$$

Since $\operatorname{rank}(\bar{A})=3=\operatorname{rank}(A)$, the system is determinate compatible. The equivalent system is

$$\begin{cases} x+y-z=2 \\ \quad -y+z=0 \\ \qquad 2z=2. \end{cases}$$

We immediately get the solution $x=2$, $y=1$, $z=1$.

We could have got to the same solution by continuing with the Gauss-Jordan method. Indeed,

$$\bar{A}\sim\begin{pmatrix} 1 & 1 & -1 & 2 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & 2 & 2 \end{pmatrix} \sim \begin{pmatrix} 1 & 1 & -1 & 2 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix} \sim \begin{pmatrix} 1 & 1 & 0 & 3 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix},$$

whence we immediately read the solution $x=2$, $y=1$, $z=1$.

(b) Consider the system

$$\begin{cases} x+y+z=0 \\ x+4y+10z=3 \\ 2x+3y+5z=1 \end{cases}$$

with real coefficients. Then its extended matrix is

$$\bar{A}=\begin{pmatrix} 1 & 1 & 1 & 0 \\ 1 & 4 & 10 & 3 \\ 2 & 3 & 5 & 1 \end{pmatrix} \sim \begin{pmatrix} 1 & 1 & 1 & 0 \\ 0 & 3 & 9 & 3 \\ 0 & 1 & 3 & 1 \end{pmatrix} \sim \begin{pmatrix} 1 & 1 & 1 & 0 \\ 0 & 1 & 3 & 1 \\ 0 & 1 & 3 & 1 \end{pmatrix} \sim \begin{pmatrix} 1 & 1 & 1 & 0 \\ 0 & 1 & 3 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Since $\operatorname{rank}(\bar{A})=2=\operatorname{rank}(A)$, the system is non-determinate compatible. The equivalent system is

$$\begin{cases} x+y+z=0 \\ \quad y+3z=1. \end{cases}$$

Then x and y are principal unknowns and z is a secondary unknown. We immediately get the solution

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

We could have got to the same solution by continuing with the Gauss-Jordan method. Indeed,

$$\bar{A}\sim\begin{pmatrix} 1 & 1 & 1 & 0 \\ 0 & 1 & 3 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & -2 & -1 \\ 0 & 1 & 3 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

The equivalent system is

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

whence we get the solution

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

Extra: Simple authentication scheme

Let us consider the following simple authentication scheme. We denote by E the canonical basis of the canonical vector space $\mathbb{Z}_2^n$ over $\mathbb{Z}_2$.

- The password is a vector $v = (x_1, \dots, x_n) \in \mathbb{Z}_2^n$.
- As a challenge, Computer sends a random vector $u = (u_1, \dots, u_n) \in \mathbb{Z}_2^n$.
- As the response, Human sends back the dot-product vector $u \cdot v = u_1x_1 + \cdots + u_nx_n \in \mathbb{Z}_2$.
- The challenge-response interaction is repeated until Computer is convinced that Human knows password v.

Eve eavesdrops and learns m pairs (a_1, b_1), ..., (a_m, b_m) such that each b_i is the correct response to challenge a_i. For every $i \in \{1, \dots, m\}$, denote $a_i = (a_{i1}, \dots, a_{in})$.

Then the password $v = (x_1, \dots, x_n)$ is a solution of the linear system of equations:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \dots\dots\dots\dots\dots\dots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m \end{cases} \tag{S}$$

Once the rank of the matrix of the system reaches n, the solution is unique, and Eve can use the Gauss method to find it, obtaining the password.

Reference: P.N. Klein, Coding the Matrix. Linear Algebra through Applications to Computer Science*, Newtonian Press, 2013.*