Practice 1
Resolution of systems of linear equations: direct methods

Contents

1 Elementary row operations

Recall that there are three types of elementary row operations:

1. Row switching: a row within the matrix can be switched with another row.
2. Row multiplication: each element in a row can be multiplied by a non-zero constant.
3. Row addition: a row can be replaced by the sum of that row and a multiple of another row.

The commands of Scilab that can be used to perform these elementary operations on a matrix A are the following:

1. Row switching applied on rows i and j:

```
A ([i , j ],:) = A ([j , i ],:)
```

2. The ith row is multiplied by p:

```
A (i ,:) = p*A (i ,:)
```

2. The ith row is replaced by the sum of the ith row and p times the jth row:

```
A (i ,:) = A (i ,:) + p *A (j ,:)
```

Example 1. Let us consider the following matrix:

$$A = \begin{bmatrix} 0 & -2 & 3 & 9 \\ -4 & 6 & 0 & -4 \\ 2 & -5 & 5 & 17 \end{bmatrix}$$

We will perform, using Scilab, the following elementary operations:

a) switch rows 1 and 3,

b) multiply by $1/2$ the second row,

c) add the first row to the second one.

The sequence of used Scilab commands and the obtained outputs are:

```
A (i ,:) = p*A (i ,:)

-->A=[0 -2 3 9;-4 6 0 -4;2 -5 5 17]
 A  =

    0.  - 2.    3.    9.
  - 4.    6.    0.  - 4.
    2.  - 5.    5.    17.

-->B=A;

-->B([1,3],:)=B([3,1],:)
 B  =

    2.  - 5.    5.    17.
  - 4.    6.    0.  - 4.
    0.  - 2.    3.    9.

-->C=B;

-->C(2,:)=(1/2)*C(2,:)
 C  =

    2.  - 5.    5.    17.
  - 2.    3.    0.  - 2.
```

```
    0.  - 2.    3.    9.

-->D=C;

-->D(2,:)=D(2,:)+D(1,:)
 D  =

    2.  - 5.    5.    17.
    0.  - 2.    5.    15.
    0.  - 2.    3.    9.
```

2 Reduced row echelon form of a matrix

The **rref** function provides the reduced echelon form of whichever matrix. Let us see an example.

Example 2. Consider the system of linear equations

$$\left.\begin{array}{rrrcr} & -2y & +3z & = & 9 \\ -4x & +6y & & = & -4 \\ 2x & -5y & +5z & = & 17 \end{array}\right\}$$

whose augmented matrix is the one of Example 1. We compute now the reduced echelon form of this matrix:

```
-->A=[0 -2 3 9;-4 6 0 -4;2 -5 5 17];

-->rref(A)
 ans  =

    1.    0.    0.    1.
    0.    1.    0.    0.
    0.    0.    1.    3.
```

Therefore the unique solution of the system is:

$$x = 1, \quad y = 0, \quad z = 3.$$

3 Resolution of linear systems with the \ operator

3.1 Description of the operator \

A commonly used procedure used to solve a system of linear equations $A\vec{x} = \vec{b}$ with Scilab consists of introducing the coefficient matrix A and the vector of independent terms $\vec{b}$ and, then, writing A\b. This operator works as follows:

1. If the system has a unique solution then A\b is the solution.

2. If the system has infinitely many solutions then A\b is one of the solutions (more specifically Scilab chooses a solution with, at most, r non-zero components, where r is the rank of A).

3. If the system has no solution then Scilab computes a vector, called **least squares approximation** of the solution, that is a vector $\vec{x}'$ such that value of the norm $\|A\vec{x}' - \vec{b}\|$ is the minimum possible. Among all possible vectors $\vec{x}'$ (notice that this vector is not necessarily unique) Scilab chooses one with, at most, r non-zero components, where r is the rank of A.

Notice that, if the system is consistent, Scilab returns one of the solutions. However, if the system is not consistent, Scilab returns a vector which **is not a solution**. This means that we must be specially careful with the operator \.

Example 3. Consider the following system of linear equations:

$$\left.\begin{array}{rrrcr} & -2y & +3z & = & 9 \\ -4x & +6y & & = & -4 \\ 2x & -5y & +5z & = & 17 \end{array}\right\}$$

We will use the operator \:

```
-->A=[0 -2 3; -4 6 0; 2 -5 5]; b=[9; -4; 17];

-->x=A\b
 x  =

    1.
    0.
    3.
```

This is the unique solution of the linear system because A is a square invertible matrix (it has maximum rank):

```
-->rank(A)
 ans  =

    3.
```

Example 4. Consider the following system of linear equations:

$$\left.\begin{array}{rrrcr} x & +y & +z & = & 1 \\ x & +y & +z & = & 2 \\ 2x & +2y & +2z & = & 3 \end{array}\right\}$$

which is, evidently, inconsistent.

```
-->A=[1 1 1; 1 1 1; 2 2 2]; b=[1; 2; 3];

-->x=A\b
warning :
matrix is close to singular or badly scaled. rcond =    0.0000D+00

 x  =

    1.5
    0.
    0.
```

Scilab provides a result but, however it is not a solution of the system.

Example 5. Let us consider the system

$$\left.\begin{array}{rrrcl} x & +y & +z & = & 1 \\ x & +y & +z & = & 1 \\ 2x & +2y & +2z & = & 2 \end{array}\right\}$$

that has, evidently, infinitely many solutions.

```
-->A=[1 1 1; 1 1 1; 2 2 2]; b=[1; 1; 2];

-->x=A\b
warning :
matrix close to singular or badly scaled. rcond =    0.0000D+00

 x  =

    1.
    0.
    0.
```

In this case we have obtained a particular solution of the system. Indeed,

```
-->A*x
 ans  =

    1.
    1.
    2.
```

3.2 General solution of a system using \ and kernel

If a system has infinitely many solutions we have seen that the \ operator provides only one of the solutions of the system. However, it is possible to obtain **all** the solutions in an easy way computing the **kernel** of the coefficient matrix. First we must clarify this concept:

The **kernel** of a matrix A is the set of solutions of the homogeneous system whose coefficient matrix is A, that is, $\mathsf{A}\vec{x} = \vec{0}$. For example, the kernel of

$$\mathsf{A} = \begin{bmatrix} 1 & 2 & -3 \\ 2 & 4 & -6 \\ 5 & 10 & -15 \end{bmatrix}$$

is the set of solutions of the system

$$\begin{bmatrix} 1 & 2 & -3 \\ 2 & 4 & -6 \\ 5 & 10 & -15 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

The kernel of a matrix can be computed easily by Scilab using the **kernel** function:

```
-->A=[1 2 -3; 2 4 -6; 5 10 -15];

-->kernel(A)
 ans  =

  - 0.1195229    0.9561829
    0.8440132  - 0.0439019
    0.5228345    0.2894597
```

The kernel of the matrix is the set of linear combinations of the column vectors of the obtained matrix, that is:

$$Ker(\mathsf{A}) = \left\{ \alpha \begin{bmatrix} -0.1195229 \\ 0.8440132 \\ 0.5228345 \end{bmatrix} + \beta \begin{bmatrix} 0.9561829 \\ -0.0439019 \\ 0.2894597 \end{bmatrix} \text{ such that } \alpha, \beta \in \mathbb{R} \right\}.$$

(This can also be expressed by saying that the column vectors of the matrix form a **system of generators** of the kernel).

Now we will state a theorem that shows how to obtain the general solution of a system of linear equations from

- a particular solution and
- the kernel of the coefficient matrix.

Theorem: Let $\mathsf{A}\vec{x} = \vec{b}$ be a consistent system of linear equations and let $\vec{x}_0$ be a particular solution. Then the general solution of the system is:

$$\vec{x} = \vec{x}_0 + \lambda_1\vec{u}_1 + \lambda_2\vec{u}_2 + \ldots + \lambda_n\vec{u}_n, \qquad \lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{R},$$

where $\{\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_n\}$ is a system of generators of the kernel of A.[1]

Example 6. Consider the system of linear equations

$$\begin{bmatrix} 1 & 0 & 2 & 3 \\ 7 & 1 & 1 & 1 \\ 8 & 1 & 3 & 4 \\ 9 & 1 & 5 & 7 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ t \end{bmatrix} = \begin{bmatrix} 6 \\ 10 \\ 16 \\ 22 \end{bmatrix}.$$

We apply, with Scilab, the $\backslash$ operator in order to study the consistency of the system:

```
-->A=[1 0 2 3; 7 1 1 1 ; 8 1 3 4; 9 1 5 7]; b=[6; 10; 16; 22];

-->x=A\b
warning :
matrix is close to singular or badly scaled. rcond =    2.2204D-18

 x  =

    1.2
    0.
    0.
    1.6

-->clean(A*x-b)
 ans  =

    0.
    0.
    0.
    0.
```

From the above outputs we see that the system is consistent and that the vector $\vec{x}_0 = (1.2, 0, 0, 1.6)$ is a particular solution. Now we compute the kernel of the coefficient matrix:

```
-->kernel(A)
 ans  =
```

[1]The proof is very easy: $\vec{x}$ is a solution of the system $\Leftrightarrow \mathsf{A}\vec{x} = \vec{b} \Leftrightarrow \mathsf{A}\vec{x} - \mathsf{A}\vec{x}_0 = \vec{b} - \mathsf{A}\vec{x}_0 \Leftrightarrow \mathsf{A}(\vec{x} - \vec{x}_0) = \vec{0}$ (vecause $\mathsf{A}\vec{x}_0 = \vec{b}$) $\Leftrightarrow \vec{x} - \vec{x}_0$ belongs to the kernel of A.

```
  - 0.1490641  - 0.0418627
    0.8434185    0.5144634
    0.4510273  - 0.7061368
  - 0.2509968    0.4847121
```

This means that the system has infinitely many solutions and its general solution is:

$$\underbrace{\begin{bmatrix}1.2\\0\\0\\1.6\end{bmatrix}}_{\vec{x}_0}+\lambda_1\begin{bmatrix}-0.1490641\\0.8434185\\0.4510273\\-0.2509968\end{bmatrix}+\lambda_2\begin{bmatrix}-0.0418627\\0.5144634\\-0.7061368\\0.4847121\end{bmatrix},\qquad \lambda_1,\lambda_2\in\mathbb{R}.$$

Example 7. Consider the system of linear equations (with the same coefficient matrix as above)

$$\begin{bmatrix}1&0&2&3\\7&1&1&1\\8&1&3&4\\9&1&5&7\end{bmatrix}\begin{bmatrix}x\\y\\z\\t\end{bmatrix}=\begin{bmatrix}1\\2\\3\\5\end{bmatrix}.$$

As above, we use \ to study the consistency of the system:

```
-->b=[1; 2; 3; 5];

-->clean(A*x-b)
 ans  =

    0.
    0.
    0.
  - 1.
```

From these outputs we see that the \ operator returns a vector that **is not a solution**. This implies that **the system is inconsistent**.

Example 8. Consider the system of linear equations

$$\begin{bmatrix}1&2&3\\-1&0&1\\0&3&1\end{bmatrix}\begin{bmatrix}x\\y\\z\end{bmatrix}=\begin{bmatrix}1\\2\\3\end{bmatrix}.$$

As above, we apply first the operator \:

```
-->A=[1 2 3; -1 0 1; 0 3 1]; b=[1; 2; 3];

-->x=A\b
```

```
 x  =

  - 1.7
    0.9
    0.3

-->clean(A*x-b)
 ans  =

    0.
    0.
    0.
```

The results show that the system is consistent and that the vector $\vec{x}_0 = (-1.7, 0.9, 0.3)$ is a solution. Now we compute the kernel of the coefficient matrix:

```
-->kernel(A)
 ans  =

      []
```

This means that the kernel is trivial, that is, $Ker(\mathsf{A}) = \{\vec{0}\}$. Therefore the unique solution of the system is $\vec{x}_0$, the one obtained using $\backslash$.

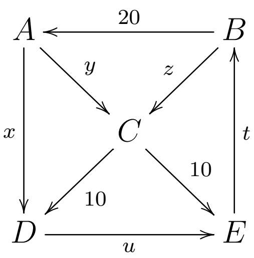
4 Application to network flows

Systems of linear equations arise when we investigate the flow of some quantity through a network. Such networks arise in science, engineering and economics. Two such examples are the pattern of traffic flow through a city and distribution of products from manufacturers to consumers through a network of wholesalers and retailers.

A network consists of a set of points, called the nodes, and directed lines connecting some or all of the nodes. The flow is indicated by a number or a variable. We observe the following basic assumptions:

- The total flow into a node is equal to the total flow out of a node.
- The total flow into the network is equal to the total flow out of the network.

Example 9. Let us consider a small closed network of pipes through which some liquid flows, as described in the following graph:



The edges represent the pipes, the arrows indicate the flows directions, and the intersections among the pipes correspond to the nodes of the graph. The weights of the edges represent the number of liters of liquid flowing per hour.

Each intersection gives rise to a linear equation:

$$\left.\begin{array}{lrcl} \text{Intersection } A\text{:} & x+y & = & 20 \\ \text{Intersection } B\text{:} & z+20 & = & t \\ \text{Intersection } C\text{:} & y+z & = & 20 \\ \text{Intersection } D\text{:} & x+10 & = & u \\ \text{Intersection } E\text{:} & u+10 & = & t \end{array}\right\},$$

The augmented matrix of this system of linear equations is

$$\begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 20 \\ 0 & 0 & 1 & -1 & 0 & -20 \\ 0 & 1 & 1 & 0 & 0 & 20 \\ 1 & 0 & 0 & 0 & -1 & -10 \\ 0 & 0 & 0 & -1 & 1 & -10 \end{bmatrix}.$$

Computing, with Scilab, the reduced row echelon form of the above matrix we obtain:

```
-->rref([1 1 0 0 0 20;0 0 1 -1 0 -20;0 1 1 0 0 20;1 0 0 0 -1 -10;0 0 0 -1 1 -10])
 ans  =

    1.    0.    0.    0.  - 1.  - 10.
    0.    1.    0.    0.    1.    30.
    0.    0.    1.    0.  - 1.  - 10.
    0.    0.    0.    1.  - 1.    10.
    0.    0.    0.    0.    0.    0.
```

Therefore, the parametric equations of the general solution of the system are: $x = -10+\lambda$, $y = 30 - \lambda$, $z = -10 + \lambda$, $t = 10 + \lambda$, $u = \lambda$, with $\lambda \in \mathbb{R}$. Since the values of the unknowns are liters of liquid we must have $x, y, z, t, u \geq 0$ and, hence, $10 \leq \lambda \leq 30$. We conclude, then, that there exist infinitely many possibilities for the distribution of flows (one for each value of λ in the interval $[10, 30]$).