Sistemas de ecuaciones lineales

Credit: This notes are 100% from chapter 2 of the book entitled *Linear Algebra. A Modern Introduction* by David Poole. Thomson. Australia. 2006.

Linear equation: a linear equation in the n variables $x_1, \cdots, x_n$ is an equation that can be written in the form $a_1x_1 + \cdots + a_nx_n = b$ where the coefficients $a_1, \cdots, a_n$ and the term b are all constants.

System of linear equations (SLE): a system of linear equations is a finite set of linear equations, each with the same variables. A solution of a SLE is a vector that is simultaneously a solution of each equation in the system.

Consistent and inconsistent: a SLE is called consistent if it has at least one solution. A system with no solutions is called inconsistent. A SLE with real coefficients has either (i) a unique solution (a consistent system), (ii) infinitely many solutions (a consistent system) or (iii) no solution (an inconsistent system).

Equivalent: two SLE are called equivalent if they have the same solution set. The general approach to solving a SLE is to transform the given system into an equivalent one that is easier to solve.

Coefficients and augmented matrices: there are two important matrices associated with a SLE. The coefficient matrix contains the coefficients of the variables, and the augmented matrix is the coefficient matrix augmented by an extra column containing the constant terms.

Back substitution: is the procedure which consist to solve the SLE starting from the last equation and working backward when the SLE is given in the form

$$x - y - z = 2 \quad (1)$$
$$y + 3z = 5 \quad (2)$$
$$5z = 10 \quad (3)$$

Next we turn to the general strategy for transforming a given system into an equivalent one that can be solved easily by back substitution. What follow is an example

$$x - y - z = 2 \quad (4)$$
$$3x - 3y + 2z = 16 \quad (5)$$
$$2x - y + z = 9 \quad (6)$$

We built the extended or augmented matrix, i.e. the matrix of the SLE coefficients and independent term,

$$\left[\begin{array}{ccc|c} 1 & -1 & -1 & 2 \\ 3 & -3 & 2 & 16 \\ 2 & -1 & 1 & 9 \end{array}\right] \tag{7}$$

First we eliminate x from the second and third equations: (i) we subtract the second line with 3 times the firs line: $f_2 - 3f_1$

$$\left[\begin{array}{ccc|c} 1 & -1 & -1 & 2 \\ 0 & 0 & 5 & 10 \\ 2 & -1 & 1 & 9 \end{array}\right] \tag{8}$$

(ii)next we subtract third line with 2 times the first one: $f_3 - 2f_1$

$$\left[\begin{array}{ccc|c} 1 & -1 & -1 & 2 \\ 0 & 0 & 5 & 10 \\ 0 & 1 & 3 & 5 \end{array}\right] \tag{9}$$

(iii) next we interchange the second with the third equation $f_2 \to f_3$

$$\left[\begin{array}{ccc|c} 1 & -1 & -1 & 2 \\ 0 & 1 & 3 & 5 \\ 0 & 0 & 5 & 10 \end{array}\right] \tag{10}$$

(iv) we are done. Now we write the SLE with the new coefficients and independent term. This is an equivalent SLE of the initial one and then they both share the same solution. The solution is obtained by back substitution.

$$x - y - z = 2 \tag{11}$$
$$y + 3z = 5 \tag{12}$$
$$5z = 10 \tag{13}$$

Row echelon(escalón) form: In solving a linear system, it will no always be possible to reduce the coefficient matrix to triangular form. However, we can always achieve a staircase pattern in the nonzero entries of the final matrix. A matrix is row echelon form if it satisfies the following properties

- Any rows consisting entirely of zeros are at the bottom
- In each nonzero row, the first nonzero entry (called the leading entry) is in a column to the left of any leading entries below it.

If a matrix in row echelon form is actually the augmented matrix of a linear system, the system is quite easy to solve by back substitution alone.
The row echelon form of a matrix is not unique (but they are all equivalents).

Elementary row operations: this is a procedure by which any matrix can be reduced to a matrix in row echelon form. The allowed operations, called elementary row operations, correspond to the operations that can be performed on a system of linear equations to transform it into an equivalent system. They are:

- Interchange two rows ($R_i \leftrightarrow R_j$).
- Multiply a row by a nonzero constant (kR_i).
- Add a multiple of a row to another row ($R_i + kR_j$).

Row reduction: the process of applying elementary row operations to bring a matrix into row echelon form, called row reduction, is used to reduce a matrix to echelon form.

Row equivalent: matrices A and B are row equivalent if there is a sequence of elementary row operations that converts A into B.

Theorem 2.1: Matrices A and B are row equivalent if and only if they can be reduced to the same row echelon form.
Proof: If A and B are row equivalent, then further row operations will reduce B (and therefore A) to the (same) row echelon form.
Conversely, if A and B have the same row echelon form R, then via elementary row operations, we can convert A into R and B into R. Reversing the latter sequence of operations, we can convert R into B, and therefore the sequence $A \to R \to B$ achieves the desired effect.

Gaussian elimination: when row reduction is applied to the augmented matrix of a SLE, we create an equivalent system that can be solved by back substitution. The entire process is known as Gaussian elimination.

Rank: the rank, $rank(A)$, of a matrix A is the number of nonzero rows in its row echelon form.

Theorem 2.2: The Rank Theorem Let A be the coefficient matrix of a SLE with n variables. If the system is consistent, then number of free variables $= n - rank(A)$

- If $n - rank(A) = 0$ there are no free variables and there is a unique solution.
- If $rank(A) \neq rank(A|b)$ the system is inconsistent, i.e. it has not solution. Example $\left[\begin{array}{ccc|c} 1 & -1 & 2 & 3 \\ 1 & 2 & -1 & -3 \\ 0 & 2 & -2 & 1 \end{array}\right] \to \left[\begin{array}{ccc|c} 1 & -1 & 2 & 3 \\ 0 & 1 & -1 & -2 \\ 0 & 0 & 0 & 5 \end{array}\right]$

Reduce row echelon: a modification of Gaussian elimination greatly simplifies the back substitution phase. This variant is known as Gauss-Jordan elimination, relies on reducing the augmented matrix even further. A matrix is in reduced row echelon form if:

- It is in row echelon form.
- The leading entry in each nonzero row is 1 (called a leading 1)
- Each column containing a leading 1 has zeros everywhere else.

Unlike the row echelon form, the reduced row echelon form of a matrix is unique.

Homogeneous systems: a system of linear equations is called homogeneous if the constant term in each equation is zero. A homogeneous system will have either a unique solution (namely, the zero, or trivial, solution) or infinitely many solutions.

Theorem 2.3: If $[A|0]$ is a homogeneous system of m linear equations with n variables, where $m < n$, the system has infinitely many solutions.
Proof: Since the system has at least the zero solution, it is consistent. Also rank$(A) \leq m$. Then by the rank theorem, we have

$$\text{number of free variables} = n - rank(A) \geq n - m > 0 \tag{14}$$

then, there is at least one free variable and, hence, infinitely many solutions.

The theorem 2.3 says nothing about the case where $m \geq n$. Exercise 44 asks you to give examples to show that, in this case, there can be either a unique solution of infinitely many solutions.

Theorem 2.4: A system of linear equations with augmented matrix $[A|\bar{b}]$ is consistent if and only if $\bar{b}$ is a linear combination of the column of A.

Span: If $\mathcal{S} = \{\bar{v}_1, \cdots, \bar{v}_k\}$ is a set of vectors in $\mathbb{R}^n$, the set of all linear combinations of $\bar{v}_1, \cdots, \bar{v}_k$ is called the span of $\bar{v}_1, \cdots, \bar{v}_k$ and is denoted by span$(\bar{v}_1, \cdots, \bar{v}_k)$ or span$(\mathcal{S})$=$\mathbb{R}^n$, then $\mathcal{S}$ is called a spanning set for $\mathbb{R}^n$.

Example: $\mathbb{R}^2 = span(\bar{v}_1, \bar{v}_2)$ with

$$\bar{v}_1 = \begin{bmatrix} 2 \\ -1 \end{bmatrix} \tag{15}$$

$$\bar{v}_2 = \begin{bmatrix} 1 \\ 3 \end{bmatrix} \tag{16}$$

We have to show that an arbitrary vector $\begin{bmatrix} a \\ b \end{bmatrix}$ can be written as a linear combination of $\bar{v}_1$ and $\bar{v}_2$, i.e. exist c_1 and c_2 such that $c_1 \begin{bmatrix} 2 \\ -1 \end{bmatrix} + c_2 \begin{bmatrix} 1 \\ 3 \end{bmatrix} = \begin{bmatrix} a \\ b \end{bmatrix}$ for any a and b.
Built the augmented matrix and reduce to the equivalent echelon row and use back substitution to get the following values for c_1 and c_2

$$c_1 = \frac{3a - b}{7} \tag{17}$$

$$c_2 = \frac{a + 2b}{7} \tag{18}$$

Linearly dependent (LD): A set of vectors $\bar{v}_1, \cdots, \bar{v}_k$ is linearly dependent if there are scalars $c_1, \cdots, c_k$ at least one of which is not zero, such that

$$c_1\,\bar{v}_1 + +c_k\,\bar{v}_k = \bar{0} \tag{19}$$

Any set of vectors containing the zero vector is LD, since for if $\bar{0}, \bar{v}_2, \cdots, \bar{v}_m$ are in $\mathbb{R}^m$, then we can find a nontrivial combination of the form $c_1\bar{0} + c_2\bar{v}_2 + \cdots + c_m\bar{v}_m = 0$ by setting $c_1 = 1$ and $c_2 = \cdots = c_m = 0$.

Linearly independent (LI): A set of vectors $\bar{v}_1, \cdots, \bar{v}_k$ that is not linearly dependent is called linearly independent.

Theorem 2.5: Vectors $\bar{v}_1, \cdots, \bar{v}_m$ in $\mathbb{R}^n$ are LD if and only if at least one of the vectors can be expressed as a linear combination of the others.
Proof: If one of the vectors, let as say $\bar{v}_1$ is a linear combination of the others, then there are scalars $c_2, \cdots, c_m$ such that $\bar{v}_1 = c_2\,\bar{v}_2 + \cdots + c_m\,\bar{v}_m$. Rearranging, we obtain $\bar{v}_1 - c_2\,\bar{v}_2 - \cdots - c_m\,\bar{v}_m = 0$, which implies that $\bar{v}_1, \cdots, \bar{v}_m$ are LD, since at least one of the scalar (the coefficient $c_1 = 1$ of $\bar{v}_1$) is nonzero.
Conversely, suppose that $\bar{v}_1, \cdots, \bar{v}_m$ are LD. Then there are scalars $c_1, \cdots, c_m$, not all zero, such that $c_1\,\bar{v}_1 + c_2\,\bar{v}_2 + \cdots + c_m\,\bar{v}_m = 0$. Suppose $c_1 \neq 0$. Then

$$c_1\,\bar{v}_1 = -c_2\,\bar{v}_2 - \cdots - c_m\,\bar{v}_m \tag{20}$$

and we may multiply both side by $1/c_1$ to obtain $\bar{v}_1$ as a linear combination of the others vectors:

$$\bar{v}_1 = -\frac{c_2}{c_1}\,\bar{v}_2 - \cdots - \frac{c_m}{c_1}\,\bar{v}_m \tag{21}$$

We have taken as a reference the vector $\bar{v}_1$ but this is valid for any other vector as well.

Theorem 2.6: Let $\bar{v}_1, \cdots, \bar{v}_m$ be (column) vectors in $\mathbb{R}^n$ and let A be the $n \times m$ matrix $[\bar{v}_1, \cdots, \bar{v}_m]$ with this vectors as its columns. Then $\bar{v}_1, \cdots, \bar{v}_m$ are LD if and only if the homogeneous linear system with augmented matrix $[A|\bar{0}]$ has a nontrivial solution.
Proof: $\bar{v}_1, \cdots, \bar{v}_m$ are LD if and only if there are scalars $c_1, \cdots, c_m$, not all zero, such that $c_1\bar{v}_1 + \cdots + c_m\bar{v}_m = 0$. By the theorem 2.4, this is equivalent to say that the nonzero vector $\begin{bmatrix} c_1 \\ \vdots \\ c_m \end{bmatrix}$ is a (non trivial) solution of the system whose augmented matrix is $[\bar{v}_1, \cdots, \bar{v}_m|\bar{0}]$.

Theorem 2.7: Let $\bar{v}_1, \cdots, \bar{v}_m$ be (row) vectors in $\mathbb{R}^n$ and let A be the $m \times n$ matrix $\begin{bmatrix} \bar{v}_1 \\ \vdots \\ \bar{v}_m \end{bmatrix}$ with this vectors as its rows. Then $\bar{v}_1, \cdots, \bar{v}_m$ are LD if and only if the rank$(A) < m$.
Proof: Assume that $\bar{v}_1, \cdots, \bar{v}_m$ are LD. Then, at least one of the vectors can be written as a linear combination of the others. We relabel the vectors, if necessary, so that we can write $\bar{v}_m = c_1\bar{v}_1 + \cdots + c_{m-1}\bar{v}_{m-1}$. Then the elementary row operations $R_m - c_1R_1$, $R_m - c_2R_2, \cdots, R_m - c_{m-1}R_{m-1}$ applied to A will create a zero row in row m. Thus, rank$(A) < m$.
Conversely, assume that rank$(A) < m$. Then there is some sequence of row operations that will create a zero row. A successive substitution argument analogous to that used in Example 2.25 can be used to show that $\bar{0}$ is a nontrivial linear combination of $\bar{v}_1, \cdots, \bar{v}_m$. Thus, $\bar{v}_1, \cdots, \bar{v}_m$ are LD.

Theorem 2.8: Any set of m vectors in $\mathbb{R}^n$ is LD if $m > n$.
Proof: Assume that $\bar{v}_1, \cdots, \bar{v}_m$ be (column) vectors in $\mathbb{R}^n$ and A be the $n \times m$ matrix $[\bar{v}_1, \cdots, \bar{v}_m]$ with these vectors as its columns. By Theorem 2.6 $\bar{v}_1, \cdots, \bar{v}_m$ are LD if and only if the homogeneous linear system with augmented matrix $[A|\bar{0}]$ has a nontrivial solution. But, according to Theorem 2.6, this will always be the case if A has more columns that rows; it is the case here, since number of columns m is greater than numbers of rows n.