Theorem of the Alternative
Notes for Math 242, Linear Algebra, Lehigh University fall 2008

Consider a system of linear equations $A\boldsymbol{x} = \boldsymbol{b}$ that has no solutions. In the text this is called incompatible and a single equation is inconsistent if it is $0 = c$ where c is a constant not equal to 0. We will refer to a system with no solution as inconsistent. The text comments that in such a situation 'there is nothing else to do.' One thing that might be done is to look for a 'best' approximation to a solution. Generally 'best' refers to an $\boldsymbol{x}$ that is closest to being a solution under some method of measuring closeness. In Chapter 4 there is discussion of the least squares problem which is one such approach.

Another thing that we may want to do if there is no solution is to provide some certificate of this fact. This is essentially the same as asking if there is a theorem characterizing when a system of linear equations has a solution. One such result, sometimes called Fredholm's Alternative and also called the Theorem of the Alternative for linear systems is given in section 5.7. The statement there is: 'The linear system $A\boldsymbol{x} = \boldsymbol{b}$ has a solution if and only if $\boldsymbol{b}$ is orthogonal to the cokernal of A.' A second version of this which does not require the extra terminology is given after this statement. We can write this as:

Theorem of the Alternative: 'The linear system $A\boldsymbol{x} = \boldsymbol{b}$ has a solution if and only if $\boldsymbol{y}^T\boldsymbol{b} = 0$ for every column vector $\boldsymbol{y}$ such that $\boldsymbol{y}^T A = \mathbf{0}$.'

A geometric view is hidden in these statements. Consider 'only if': ($\boldsymbol{y}^T\boldsymbol{b} = 0$ for every column vector $\boldsymbol{y}$ such that $\boldsymbol{y}^T A = \mathbf{0}$) $\Rightarrow$ ($A\boldsymbol{x} = \boldsymbol{b}$ has a solution). The contrapositive of this is: Not ($A\boldsymbol{x} = \boldsymbol{b}$ has a solution) $\Rightarrow$ Not ($\boldsymbol{y}^T\boldsymbol{b} = 0$ for every column vector $\boldsymbol{y}$ such that $\boldsymbol{y}^T A = \mathbf{0}$). This is: if $A\boldsymbol{x} = \boldsymbol{b}$ is inconsistent then there exists a $\boldsymbol{y}^T$ such that $\boldsymbol{y}^T A = \mathbf{0}$ and $\boldsymbol{y}^T\boldsymbol{b} \neq 0$. A solution to $A\boldsymbol{x} = \boldsymbol{b}$ gives a way of writing $\boldsymbol{b}$ as a linear combination of the columns of A. That is, $\boldsymbol{b}$ can be written as a sum of scalar multiples of the columns of A. When there is no solution, the alternative states that there is a vector $\boldsymbol{y}$ that is orthogonal to each column of A but not orthogonal to $\boldsymbol{b}$.

For example consider $\begin{array}{ccccc} 2x_1 & + & x_2 & = & b_1 \\ x_1 & - & 3x_2 & = & b_2 \\ 3x_1 & + & 3x_2 & = & b_3 \end{array}$. This is $\begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix} x_1 + \begin{pmatrix} 1 \\ -3 \\ 3 \end{pmatrix} x_2 = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$. The column vectors $\begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix}$ and $\begin{pmatrix} 1 \\ -3 \\ 3 \end{pmatrix}$ span a plane $12x - 3y - 7z = 0$ in $\Re^3$. If $\boldsymbol{b}$ is in the plane then the system has a solution. For example it is easy to check that $\boldsymbol{b} = \begin{pmatrix} 1 \\ 11 \\ -3 \end{pmatrix}$ is in the plane, corresponding to $\begin{array}{ccccc} 2x_1 & + & x_2 & = & 1 \\ x_1 & - & 3x_2 & = & 11 \\ 3x_1 & + & 3x_2 & = & -3 \end{array}$ having a solution $x_1 = 2, x_2 = -3$.

Alternatively, $\boldsymbol{b} = \begin{pmatrix} 2 \\ 11 \\ -3 \end{pmatrix}$ is not in the plane. The equation for the plane gives us a normal vector $\begin{pmatrix} 12 \\ -3 \\ -7 \end{pmatrix}$, which is orthogonal to every vector in the plane. Corresponding to this, taking the system $\begin{array}{rcrcr} 2x_1 & + & x_2 & = & 2 \\ x_1 & - & 3x_2 & = & 11 \\ 3x_1 & + & 3x_2 & = & -3 \end{array}$ and multiplying the first equation by 12, the second by -3 and the third by -7 and combining the resulting equations yields $0x_1 + 0x_2 = 12$. Since $0 \neq 12$ the system must be inconsistent.

Restating the Theorem of the alternative again, using the contrapositive as above we have:

Theorem of the Alternative for systems of linear equations: Exactly one of the following holds:
(I) $A\boldsymbol{x} = \boldsymbol{b}$ has a solution $\boldsymbol{x}$
(II) $\boldsymbol{y}^T A = \boldsymbol{0}, \quad \boldsymbol{y}^T \boldsymbol{b} \neq 0$ has a solution $\boldsymbol{y}$.

We can already prove the Theorem of the Alternative using what we know from chapter 1.

Proof of the Theorem of the Alternative: If $\hat{\boldsymbol{x}}$ satisfies (I) and $\hat{\boldsymbol{y}}$ satisfies (II) then
$0 = \boldsymbol{0}^T \hat{\boldsymbol{x}} = (\hat{\boldsymbol{y}}^T A)\hat{\boldsymbol{x}} = \hat{\boldsymbol{y}}^T (A\hat{\boldsymbol{x}}) = \hat{\boldsymbol{y}}^T \hat{\boldsymbol{b}} \neq 0$.
So both cannot hold.

Consider a factorization $PA = LU$ where P is a permutation matrix, L is lower triangular with the diagonal entries equal to 1 and U is in row echelon form. Such a factorization can be found using Gaussian elimination. Observe that L is nonsingular. We can see this in several ways: Gaussian elimination allows pivots on each diagonal entry since they are nonzero or the determinant is 1 or forward substitution yields solutions to $L\boldsymbol{c} = \boldsymbol{b}$ for all $\boldsymbol{b}$.

To solve $A\boldsymbol{x} = \boldsymbol{b}$ look at $PA\boldsymbol{x} = P\boldsymbol{b}$ which is $L(U\boldsymbol{x}) = P\boldsymbol{b}$. Since L is nonsingular we can solve $L\boldsymbol{c} = P\boldsymbol{b}$ with $\boldsymbol{c} = L^{-1}P\boldsymbol{b}$. We then try to solve $U\boldsymbol{x} = \boldsymbol{c}$. We can solve this using back substitution unless there is a row of U which is all 0's with the corresponding entry of $\boldsymbol{c}$ not zero. That is, the system $U\boldsymbol{x} = \boldsymbol{c}$, which is equivalent to $A\boldsymbol{x} = \boldsymbol{b}$ contains an equation $0 = c_i$ for some $c_i \neq 0$ so there is no solution. If this is the i^{th} row let $\tilde{\boldsymbol{y}}^T$ be the i^{th} row of $L^{-1}P$. Then $\tilde{\boldsymbol{y}}^T A$ is the i^{th} row of $L^{-1}PA = U$ so it is $\boldsymbol{0}^T$ and $\tilde{\boldsymbol{y}}\boldsymbol{b}$ is the i^{th} entry c_i of $L^{-1}P\boldsymbol{b} = \boldsymbol{c}$ which is not zero. $\square$

We next give an example to illustrate this

Let $A = LU$ with

$$A = \begin{bmatrix} 3 & 9 & 3 & 7 & -1 & 3 \\ 1 & 3 & 6 & 4 & 9 & 4 \\ 1 & 3 & 0 & 2 & -1 & 1 \\ -2 & -6 & 6 & -2 & 8 & -1 \end{bmatrix} \qquad U = \begin{bmatrix} 1 & 3 & 0 & 2 & -1 & 1 \\ 0 & 0 & 3 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$L = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 3 & 1 & 0 & 0 \\ -2 & 2 & 1 & 0 \\ 1 & 2 & 3 & 1 \end{bmatrix} \qquad L^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -3 & 1 & 0 & 0 \\ 8 & -2 & 1 & 0 \\ -19 & 4 & -3 & 1 \end{bmatrix} \qquad P = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

Also let

$$\boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{pmatrix} \qquad \boldsymbol{b'} = \begin{pmatrix} 11 \\ 23 \\ 1 \\ 16 \end{pmatrix} \qquad \boldsymbol{b''} = \begin{pmatrix} 11 \\ 24 \\ 1 \\ 16 \end{pmatrix}.$$

Consider both $A\boldsymbol{x'} = \boldsymbol{b'}$ and $A\boldsymbol{x''} = \boldsymbol{b''}$. We have $P\boldsymbol{b'} = \begin{pmatrix} 1 \\ 11 \\ 16 \\ 23 \end{pmatrix}$ and $P\boldsymbol{b''} = \begin{pmatrix} 1 \\ 11 \\ 16 \\ 24 \end{pmatrix}$.

Solving $L\boldsymbol{c'} = P\boldsymbol{b'}$ and $L\boldsymbol{c''} = P\boldsymbol{b''}$ we get $\boldsymbol{c'} = \begin{pmatrix} 1 \\ 8 \\ 2 \\ 0 \end{pmatrix}$ and $\boldsymbol{c'} = \begin{pmatrix} 1 \\ 8 \\ 2 \\ 1 \end{pmatrix}$.

Then using back substitution to solve $U\boldsymbol{x'} = \boldsymbol{c'}$ we get for any real numbers r, s, t

$$\begin{pmatrix} x'_1 \\ x'_2 \\ x'_3 \\ x'_4 \\ x'_5 \\ x'_6 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \\ 2 \\ 0 \\ 1 \\ 0 \end{pmatrix} + \begin{pmatrix} -3 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} r + \begin{pmatrix} -2 \\ 0 \\ -1/3 \\ 1 \\ 0 \\ 0 \end{pmatrix} s + \begin{pmatrix} -3/2 \\ 0 \\ 1/3 \\ 0 \\ 0 \\ 1 \end{pmatrix} t.$$

Now $U\boldsymbol{x''} = \boldsymbol{c''}$ includes the inconsistency $0 = 1$ in the last row. So there is no solution. The 4^{th} row of $L^{-1}P$ is $\boldsymbol{y''}^T = \begin{pmatrix} 4 & 1 & -19 & -3 \end{pmatrix}$. We note that $\boldsymbol{y''}^T A = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$ and the 4^{th} entry of $\boldsymbol{y''}^T \boldsymbol{b''}$ is 1. This shows that the original system is inconsistent.