Quote of Lecture 7	
Chairman Kaga: If memory serves me right ...	
	Iron Chef: (1993-1999)

In conclusion of chapter 7 we summarize the important results concerning the linear system of equations $\mathbf{Ax} = \mathbf{b}$ where $\mathbf{A} \in \mathbb{R}^{m\times n}$, $\mathbf{x} \in \mathbb{R}^{n\times 1}$, $\mathbf{b} \in \mathbb{R}^{m\times 1}$. Specifically, we begin with the following system of linear equations,

$$
\begin{aligned}
a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \cdots + a_{1n}x_n &= b_1 \\
a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \cdots + a_{2n}x_n &= b_2 \\
a_{31}x_1 + a_{32}x_2 + a_{33}x_3 + \cdots + a_{3n}x_n &= b_3 \\
\vdots \qquad &= \vdots \\
a_{m1}x_1 + a_{m2}x_2 + a_{m3}x_3 + \cdots + a_{mn}x_n &= b_m,
\end{aligned} \tag{1}
$$

which we know has the following matrix-vector representation,

$$
\mathbf{Ax} = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1m} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2m} \\ a_{31} & a_{32} & a_{33} & \dots & a_{3m} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \dots & a_{nm} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ \vdots \\ b_m \end{bmatrix} = \mathbf{b}, \tag{2}
$$

where the product of matrices is defined as $[\mathbf{AB}]_{ij} = \sum_{k=1}^{n} a_{ik}b_{kj}$. [1] We note that this system can also be defined as a **linear combination** of vectors,

$$
\mathbf{Ax} = x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + x_3\mathbf{a}_3 + \cdots x_n\mathbf{a}_n = \sum_{j=1}^{n} x_j\mathbf{a}_j = \mathbf{b}, \tag{3}
$$

where $\mathbf{a}_i \in \mathbb{R}^m$ is the i^{th} column from the **coefficient matrix A**. Given **A** and **b** we ask,

- Does there exist a solution to the linear system $\mathbf{Ax} = \mathbf{b}$?

with the understanding that there may **exist** a solution and this solution **may** be **unique**. Thus, we have three possible outcomes when trying to solve $\mathbf{Ax} = \mathbf{b}$,

(1) There exists a solution to $\mathbf{Ax} = \mathbf{b}$,
(2) There exists a solution to $\mathbf{Ax} = \mathbf{b}$ and this solution is unique,
(3) There does not exist a solution to $\mathbf{Ax} = \mathbf{b}$,

[1] The following website contains an animation of matrix multiplication, `http://www.sci.wsu.edu/math/faculty/genz/220v/lessons/kentler/FullMult/fullMatrixMultiply.html`. We should note that, visually, this is not the same way we multiply but it is equivalent. This goofy animation, `http://www.purplemath.com/modules/mtrxmult.htm` is similar to how we conduct multiplication.

which can be seen in the 1D case by quick but careful inspection of $ax = b$, $a, b \in \mathbb{R}$. For the case of coefficient data from $\mathbb{R}^{m\times n}$, we form an augmented matrix:

$$\left[\begin{array}{ccccc|c} a_{11} & a_{12} & a_{13} & \dots & a_{1m} & b_1 \\ a_{21} & a_{22} & a_{23} & \dots & a_{2m} & b_2 \\ a_{31} & a_{32} & a_{33} & \dots & a_{3m} & b_3 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \\ a_{n1} & a_{n2} & a_{n3} & \dots & a_{nm} & b_m \end{array}\right] \tag{4}$$

and apply the row-reduction algorithm.[2] to as This algorithm makes us of the following three rules for manipulating augmented matrices,

(1) row scaling: the multiplication a row by a nonzero scalar,
(2) row exchange: the exchanges of two rows,
(3) row replacement: the addition of a multiple of one row to another row,

which we know to be equivalent to algebra applied directly to the linear system of equations (1). We apply this algorithm, when taken to full completion, in a two-part process. In the so-called *forward phase* we:

(1) Begin with the leftmost nonzero column. This is a pivot column. The pivot position is at the top.
(2) Select a nonzero entry in the pivot column as a pivot. If necessary, interchange rows to move this entry into the pivot position.
(3) Use row replacement operations to create zeros in all positions below the pivot.
(4) Cover (or ignore) the row containing the pivot position and cover all rows, if any, above it. apply steps 1-3 to the submatrix that remains. Repeat the process until there are no more nonzero rows to modify.

While, in the *backward phase* we:

(1) Begin with the rightmost pivot and working upward and to the left, create zeros above each pivot. If a pivot is not 1, make it 1 by a scaling operation.

The forward phase produces a row echelon form of the input matrix. From this echelon form the backward phase produces the reduced row echelon form of the input matrix.[3] Since, row-operations do not change the solution to linear systems, getting to these echelon forms is the goal of the algorithm. For example if we take $\mathbf{A} \in \mathbb{R}^{6\times 9}$ and $\mathbf{b} \in \mathbb{R}^{6\times 1}$ and reduce it to,

$$\left[\begin{array}{ccccccccc|c} 1 & * & 0 & 0 & * & * & 0 & * & 0 & a \\ 0 & 0 & 1 & 0 & * & * & 0 & * & 0 & b \\ 0 & 0 & 0 & 1 & * & * & 0 & * & 0 & c \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & * & 0 & d \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & e \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & f \end{array}\right], \tag{5}$$

where $*$ are in general non-zero elements, then the following simplest linear system whose solution is equivalent to the solution of $\mathbf{Ax} = \mathbf{b}$, [4]

$$\begin{aligned} 1x_1 + *x_2 + *x_5 + *x_6 + *x_8 &= a \\ x_3 + *x_5 + *x_6 + *x_8 &= b \\ x_4 + *x_5 + *x_6 + *x_8 &= c \\ x_7 + *x_8 &= d \\ x_9 &= e \\ 0x_1 + 0x_2 + 0x_3 + 0x_4 + 0x_5 + 0x_6 + 0x_7 + 0x_8 + 0x_9 &= f \end{aligned}$$

[2]This algorithm is also often called Gaussian elimination in honor of its European inventor Carl Friedrich Gauss. http://en.wikipedia.org/wiki/Gaussian_elimination

[3]http://www.math.aau.dk/~ottosen/Mat2C/rralg.html

[4]The following website contains an animation of row-reduction, http://www.sci.wsu.edu/math/faculty/genz/220v/lessons/kentler/SolveAnimEch/solveAnim1.html . There are more at http://www.sci.wsu.edu/math/faculty/genz/220v/lessons/kentler/SolveAnimEch/.

We notice that if $f \neq 0$ then the final equation is inconsistent and the system has no solution. If $f = 0$ then there is a solution to the system but, since we started with more columns than rows, this solution is not unique. [5] From this we notice that not only is it important to deduce, from row-reduction, the consistency of each equation one must also compare the total number of variables to the number of pivots or the number of free variables. This difference determines the uniqueness of the solutions and is an expression of the rank-nullity theorem.

Since the calculation of the differences,

(1) $\Delta_1 =$ [number of variables] - [number of pivots]
(2) $\Delta_2 =$ [number of variables] - [number of free variables]

will be important we record the following definitions, which will allow us to calculate these numbers from row-reduced matrices. Before we recite this information we take a minute to note the logic these statements will be used for.

(1) Given some set of vectors we must determine how to make new ones from vectors from the set $\Longrightarrow$ **linear combination**.
(2) Suppose we make the set of all linear combinations, which in general contains an infinite number of vectors, we would like a way to specify those vectors necessary for the construction of this **spanning set** $\Longrightarrow$ **linear independence** .
(3) Spanning sets are examples of so-called **linear vector spaces** and all linearly independent vectors from this set/space constitutes a **basis** for this space with **dimension** equal to the number of vectors in any basis.
(4) Two vector spaces important in the study of systems involving **A** are the **null space** and **column space** of **A**. The dimension of the null space counts the number of free variables and the dimension of the column space counts the number of pivots $\Longrightarrow$ **Rank-Nullity Theorem**.

Definition 1. *Linear Combination: Let* $S = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3, \dots, \boldsymbol{v}_k\}$ *where* $\boldsymbol{v}_i \in \mathbb{R}^n$ *for* $i = 1, 2, 3, \dots, k,\ k \in \mathbb{N}$, *then we say that* $\boldsymbol{x} \in \mathbb{R}^n$ *is a linear combination of the vectors from* S *if*

$$\boldsymbol{x} = c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 + c_3\boldsymbol{v}_3 + \cdots c_k\boldsymbol{v}_k = \sum_{j=1}^{k} c_j\boldsymbol{v}_j, \tag{6}$$

where $c_j \in \mathbb{R}$.

Definition 2. *Linear Independence: Let* S *be as before then we say that* S *forms a linearly independent set if the following bidirectional implication holds,*

$$\sum_{j=1}^{k} c_j\boldsymbol{v}_j = \boldsymbol{0} \iff c_i = 0, \quad \textit{for all } i = 1, 2, 3, \dots, k. \tag{7}$$

Remark 1. *Recall that (7) is equivalent to* $\mathbf{Vc} = \boldsymbol{0}$, *where* $\mathbf{V}$ *is a matrix whose* j^{th} *column is* $\boldsymbol{v}_j$ *and* $\mathbf{c}$ *is a vector whose elements are* c_j. *If we note, without proof, that pivot columns are linearly independent then the linearly independent vectors from* S *are the pivot columns from matrix* $\mathbf{V}$ *and that these columns can be found by the row-reduction applied to* $\mathbf{V}$.

Definition 3. *Spanning Set: Let* S *be as before. Then we define the span of* S *as the set of all linear combinations of the the vectors from* S. *That is span*$\{S\}$ *is the set of all* $\boldsymbol{x}$ *defined by (6).*

Definition 4. *Linear Vector Space: A linear vector space, or just vector space for brevity, is a set of vectors* S, *which is also closed under arbitrary linear combinations of the vectors from* S.[6] *This is to say that a vector space is a the set* S *along with all linear combinations of the vectors from* S. *It follows that the span of any set of vectors is a vector space.*

[5] In other words the system has more variables than equations and from this underdetermined system one would never expect unique solutions. To have unique results one must have at least as many equations as unknowns. If there are more equations than unknowns then the system is said to be overdetermined.

[6] This definition is somewhat imprecise. There are particular algebraic rules, which must hold for the space to be a vector space. Please consult 7.9 of your text for more detailed information.

Definition 5. *Basis: Given a vector space, say $\tilde{S}$, we say that a basis for this space is the maximum collection of linearly independent vectors from $\tilde{S}$ or equivalently the minimum collection of vectors needed to span the space $\tilde{S}$.*

Definition 6. *Dimension: Given a vector space $\tilde{S}$ and a basis for this space, say $B_{\tilde{S}}$ we say that the dimension of the space is the number of vectors in this basis. That is, dim($\tilde{S}$) =dim($B_{\tilde{S}}$).*

Definition 7. *Null Space: The null space of a matrix, $Nul\{\boldsymbol{A}\}$, is the vector space defined by all solutions to the homogenous system $\boldsymbol{Ax} = \boldsymbol{0}$.*

Remark 2. *If the system $\boldsymbol{Ax} = \boldsymbol{0}$ is consistent with infinitely many solutions then the general solution will be a linear combination of vectors multiplied by free variables. These vectors form a basis for the null space and thus the null space has dimension equal to the number of free variables.*

Definition 8. *Column Space: The column space of a matrix, $col\{\boldsymbol{A}\}$, is the set of all linear combinations of the columns of $\boldsymbol{A}$.*

Remark 3. *From remark 1 we have that the pivot columns of a matrix are linearly independent and thus a basis for the column space of a matrix is its set of pivot columns. From this we conclude that the dimension of the column space of a matrix, also known as its **Rank**, is the number of pivots in the matrix.*

Theorem 1. *Rank-Nullity Theorem: Let $\boldsymbol{A} \in \mathbb{R}^{m\times n}$ then the following equality holds,*

$$Rank(\boldsymbol{A}) + dim(Nul(\boldsymbol{A})) = n, \tag{8}$$

which asserts that the number of pivots plus the number of free variables must be equal to the number of columns in $\boldsymbol{A}$.

This summarizes the major concepts from chapter 7. The following statement summarizes this material for the case where the coefficient matrix is square.

Theorem 2. ***The invertible matrix theorem**: Let $\boldsymbol{A} \in \mathbb{R}^{n\times n}$. Then the following statements are equivalent:*

(1) *$\boldsymbol{A}$ is an invertible matrix. That is $\boldsymbol{A}^{-1}$ exists*
(2) $det(\boldsymbol{A}) \neq 0$
(3) *$\boldsymbol{A}$ is row equivalent to the $n \times n$ identity matrix*
(4) *$\boldsymbol{A}$ has n-pivot positions*
(5) *The equation $\boldsymbol{Ax} = \boldsymbol{0}$ has only the trivial solution*
(6) *The columns of $\boldsymbol{A}$ form a linearly independent set*
(7) *The equation $\boldsymbol{Ax}$=$\boldsymbol{b}$ has a unique solution for each $\boldsymbol{b} \in \mathbb{R}^n$*
(8) *The columns of $\boldsymbol{A}$ span $\mathbb{R}^n$*
(9) *The columns of $\boldsymbol{A}$ for a basis for $\mathbb{R}^n$*
(10) $col(\boldsymbol{A}) = \mathbb{R}^n$
(11) $dim\{col(\boldsymbol{A})\} = n$
(12) $rank(\boldsymbol{A}) = n$
(13) $nul(\boldsymbol{A}) = \{\boldsymbol{0}\}$
(14) $dim\{nul(\boldsymbol{A})\} = 0$