CPT Lecture Notes 1: Linear Algebra

Levent Ülkü

August 16, 2019

Notation

$A, B, \ldots$ denote matrices.

$x, y, a, b, \ldots$ denote vectors and numbers

Vectors are always columns, but it is customary to write them as rows to save space.

Remind yourselves about simple rules of matrix algebra. For instance: transposes, inverses, the fact that AB may not equal BA even when both multiplications are allowed, etc.

What are we interested in?

Given a matrix $A_{m\times n}$ and a vector $b \in \mathbb{R}^m$, when is there a solution $x \in \mathbb{R}^n$ to $Ax = b$? When is the solution unique?

This is a system of m equations with n unknowns.

Simple observation: If $b = 0_{m\times 1}$, the system is said to be homogenous and there is at least one solution, given by $x = 0_{n\times 1}$.

If $b \neq 0$, there need not exist x such that $Ax = b$.

Definition: $y \in \mathbb{R}^m$ is a linear combination of $x_1, ..., x_n \in \mathbb{R}^m$ if there exist $\alpha_1, ..., \alpha_n \in \mathbb{R}$ such that $y = \sum_{i=1}^{n} \alpha_i x_i$.

An observation regarding matrix multiplication

Consider three matrices $B_{k\times m}, A_{m\times n}$ and $C_{k\times n}$ such that $BA = C$.

Observe that:

(1) rows of C are linear combinations of rows of A, and

(2) columns of C are linear combinations of columns of B.

These statements are in fact equivalent since

$$BA = C \Leftrightarrow A^T B^T = C^T.$$

Example:

$$\begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix}$$

$$= \begin{bmatrix} b_{11}a_{11} + b_{12}a_{21} + b_{13}a_{31} & b_{11}a_{12} + b_{12}a_{22} + b_{13}a_{32} \\ b_{21}a_{11} + b_{22}a_{21} + b_{23}a_{31} & b_{21}a_{12} + b_{22}a_{22} + b_{23}a_{32} \end{bmatrix}$$

The first row of C

$$b_{11} \begin{bmatrix} a_{11} & a_{12} \end{bmatrix} + b_{12} \begin{bmatrix} a_{21} & a_{22} \end{bmatrix} + b_{13} \begin{bmatrix} a_{31} & a_{32} \end{bmatrix}$$

is a linear combination of rows of A with entries in the first row of B as coefficients.

Elementary row operations and row reduced echelon form

Given an equation $Ax = b$, sometimes we may wish to perform a sequence of elementary row operations on A to facilitate the solution for x.

There are three elementary row operations:

(1) switching rows,

(2) multiplying a row with a nonzero constant, and

(3) replacing a row with the sum of that row and another row.

Each elementary row operation is tantamount to pre-multiplying A with a modified identity matrix.

Basically, to perform a specific elementary row operation on $A_{m\times n}$, we pre-multiply it with a matrix $E_{m\times m}$ which we get by applying the same operation to the identity matrix $I_{m\times m}$.

Such modified identity matrices are called elementary matrices .

Example 1: switch the first and third rows of A.

$$\underbrace{\begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}}_{E} \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = \begin{bmatrix} a_{31} & a_{32} & a_{33} \\ a_{21} & a_{22} & a_{23} \\ a_{11} & a_{12} & a_{13} \end{bmatrix}$$

Note that E is invertible and $E^{-1} = E$. Hence E^{-1} is also an elementary matrix.

Example 2: multiply the second row of A with a constant $c \neq 0$.

$$\underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & c & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{E} \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ ca_{21} & ca_{22} & ca_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

Note that E is invertible and that

$$E^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1/c & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

is also an elementary matrix.

Example 3: replace the first row of A with the sum of its first and the second rows.

$$\underbrace{\begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{E} \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = \begin{bmatrix} a_{11}+a_{21} & a_{12}+a_{22} & a_{13}+a_{23} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

Note that E^{-1} exists but is not elementary. However it is easily expressed as a product of other elementary matrices:

$$E^{-1} = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Theorem: Let $E_{n\times n}$ be an elementary matrix. Then:
(1) E is invertible.
(2) There exist a number k and elementary matrices $E_1, ..., E_k$ such that $E^{-1} = E_1 ... E_k$.

Proof: Skip.

Applying these three row operations sequentially to any given matrix, we can find its row reduced echelon form (RREF) .

The RREF of a matrix has the following properties:

(1) all the zero rows are at the bottom,

(2) the first nonzero entry of each nonzero row is 1,

(3) if the first nonzero entry of a row occurs on column j and if the first nonzero entry of the next row occurs on column j', then $j < j'$, and

(4) if the first nonzero entry of a row occurs on column j, then all the earlier rows have zeros on column j.

Hence RREFs have a ladder-like appearance.

The matrices

$$\begin{bmatrix} 1 & 0 & 0 & 2 \\ 0 & 0 & 1 & 13 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}, \quad I_{n \times n}$$

are RREF.

The matrix

$$\begin{bmatrix} 1 & 0 & 2 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

is not.

Theorem: Every matrix has a unique RREF.

Proof: Skip.

We will denote the RREF of A by A'. Note that given A and its RREF A', there exists a (non-unique) collection of elementary matrices E_i, $i = 1, ..., r$ such that

$$A' = E_r ... E_1 A.$$

Note:
(1) Let $E_1, ..., E_r$ be elementary matrices. $(E_r ... E_1)^{-1} = E_1^{-1} ... E_r^{-1}$ is a product of elementary matrices since each E_i^{-1} is either itself elementary, or it is the product of elementary matrices.
(2) If A is a RREF matrix, then $A' = A$. Hence $A'' = A'$.
(3) For any $b \in \mathbb{R}^m$, let $b' = E_r ... E_1 b$ where $E_1, ..., E_r$ are elementary matrices such that $A' = E_r ... E_1 A$. Then $Ax = b \Leftrightarrow A'x = b'$.

Rank

Definition: The rank of $A_{m\times n}$ denoted by rkA, is the number of nonzero rows in its RREF A'.

Note that if A is $m \times n$, then $rkA = rkA' \leq \min\{m, n\}$.

Also note, for a given $A_{m\times n}$:
(1) $rkA = m$ if and only if A' has no zero rows.
(2) $rkA = n$ if and only if all columns of A' are pivotal. (A pivotal column is one which contains the first nonzero entry of some row.)

Theorem: Let A be $m \times n$. Then

$$rkA = m \Leftrightarrow \forall b \in \mathbb{R}^m,\ \exists x \in \mathbb{R}^n \text{ such that } Ax = b.$$

In other words rank of $A_{m \times n}$ is m if and only if for each b there is **at least one** solution to the system $Ax = b$.

Proof: In one direction, suppose that $rkA = m$. Pick any $b \in \mathbb{R}^m$. We need to find some $x \in \mathbb{R}^n$ such that $Ax = b$.

Let $E_1, ..., E_r$ be elementary matrices such that $A' = E_r...E_1A$. Let $b' = E_r...E_1b \in \mathbb{R}^m$ so that $Ax = b \Leftrightarrow A'x = b'$. Now we need only find $x \in \mathbb{R}^n$ such that $A'x = b'$.

Since $rkA = m$, A' does not have a row of zeros. For every row $i = 1, ..., m$, let j_i be the column in which the first nonzero entry of row i appears. Note that $\{j_1, ..., j_m\} \subseteq \{1, ..., n\}$ (Possibly $\{j_1, ..., j_m\} \neq \{1, ..., n\}$.)

Let $x \in \mathbb{R}^n$ be such that $x_{j_i} = b'_i$ for each i and $x_k = 0$ if $k \notin \{j_1, ..., j_m\}$.

It follows that $A'x = b'$, as desired.

(Why would this argument fail if $rkA < m$?)

(Proof continued) In the other direction, suppose that for every $b \in \mathbb{R}^m$, there exists some $x \in \mathbb{R}^n$ such that $Ax = b$. We need to show that $rkA = m$. It suffices to show that A' does not have a row of zeros.

Recalling $A' = E_r ... E_1 A$, let

$$b = (E_r ... E_1)^{-1} [1]_{m \times 1} \in \mathbb{R}^m$$

and let x be such that $Ax = b$. ($[1]_{m \times 1}$ is a column of 1s.)

Then

$$A'x = [1]_{m \times 1} .$$

This implies that A' can not have a row of zeros, as desired.■

Theorem: Let A be $m \times n$. Then

$$rkA = n \Leftrightarrow \forall b \in \mathbb{R}^m, \text{ if } Ax = Ay = b, \text{ then } x = y.$$

In other words rank of $A_{m \times n}$ is n if and only if for each b there is **at most one** solution to the system $Ax = b$.

Proof: In one direction, suppose that $rkA = n$ so that $n \leq m$. Pick any $b \in \mathbb{R}^m$ and suppose that $Ax = Ay = b$. We need to show that $x = y$.

Note A' can not have a column of zeros.

Let, as usual, $E_1, ..., E_r$ be elementary matrices such that $A' = E_r ... E_1 A$. Now let $b' = E_r ... E_1 b$. It follows that $A'x = A'y = b'$.

Then:
(1) $b'_i = 0$ for all $n < i \leq m$. (Such i need not exist, i.e., possibly $n = m$.)
(2) $x_i = y_i = b'_i$ for every $1 \leq i \leq n$, i.e., $x = y$ as desired.

(What would go wrong in this argument if $rkA < n$?)

(Proof continued) In the other direction, suppose that for every $b \in \mathbb{R}^m$ if $Ax = Ay = b$, then $x = y$. We need to show that $rkA = n$. It suffices to show that all columns of A' are pivotal.

Take $b = 0 \in \mathbb{R}^m$.

The unique x that solves $Ax = 0$ is $x = 0 \in \mathbb{R}^n$.

Then the unique x that solves $A'x = 0$ is also $x = 0 \in \mathbb{R}^n$ and A' can not have a non pivotal column, as desired. ■

Corollary: Let A be $m \times n$. For each $b \in \mathbb{R}^m$, $Ax = b$ has a unique solution in $\mathbb{R}^n$ if and only if $rkA = m = n$.

Subspaces

Definition: A nonempty set $S \subseteq \mathbb{R}^m$ is a subspace of $\mathbb{R}^m$ if $\alpha x + \beta y \in S$ whenever $x, y \in S$ and $\alpha, \beta \in \mathbb{R}$.

Note:
(1) If S is a subspace of $\mathbb{R}^m$, then $0 \in S$. (Since any subspace S is by definition nonempty, there exists some $x \in S$. Then $0 = x - x \in S$ as well.)
(2) $\{0\}$ and $\mathbb{R}^m$ are subspaces of $\mathbb{R}^m$.
(3) $\mathbb{R}$ has only two subspaces.
(4) The intersection of members an arbitrary family of subspaces is a subspace. The union of subspaces need not be a subspace.

Any matrix $A_{m\times n}$ defines two subspaces.

Definition: The range of $A_{m\times n}$ is the set

$$R(A) = \{b \in \mathbb{R}^m : \exists x \in \mathbb{R}^n \text{ such that } Ax = b\}.$$

Definition: The null space of $A_{m\times n}$ is the set

$$N(A) = \{x \in \mathbb{R}^n : Ax = 0\}.$$

It is straightforward to show that $R(A)$ is a subspace of $\mathbb{R}^m$ and $N(A)$ is a subspace of $\mathbb{R}^n$.

Note:

1. $N(A) = N(A')$. (Why?)
2. $R(A) \neq R(A')$. (Why not?)
3. $R(BA) \subseteq R(B)$. (Why?)
4. $N(A) \subseteq N(BA)$. (Why?)
5. For any $A_{m\times n}$, $R(A) = \mathbb{R}^m \Leftrightarrow rkA = m$. (Why?)
6. For any $A_{m\times n}$, $N(A) = \{0\} \Leftrightarrow rkA = n$. (Why?)

Let S be a subspace of $\mathbb{R}^m$.

Definition: Vectors $a_1, ..., a_n$ in S span S if each $x \in S$ is a linear combination of $a_1, ..., a_n$. The set $\{a_1, ..., a_n\} \subset S$ spans S if its members span S.

In other words, $a_1, ..., a_n \in S$ span S if for every $x \in S$, there exists $\alpha_1, ..., \alpha_n \in \mathbb{R}$ such that $x = \sum_{i=1}^{n} \alpha_i a_i$.

For example, the columns of A span $R(A)$ and the rows of A span $R(A^T)$. (Why?)

Definition: Vectors $a_1, ..., a_n$ in $\mathbb{R}^m$ are linearly independent if for all $\alpha_1, ..., \alpha_n \in \mathbb{R}$,

$$\sum_{i=1}^{n} \alpha_i a_i = 0 \Rightarrow \alpha_1 = \cdots = \alpha_n = 0.$$

Vectors $a_1, ..., a_n$ are linearly dependent if they are not linearly independent. A set $\{a_1, ..., a_n\} \subset \mathbb{R}^m$ is linearly (in)dependent if its members are linearly (in)dependent.

Note:
(1) If $a_i = a_j$ for some $i \neq j$, then vectors $a_1, ..., a_n$ are linearly dependent. (Why?)
(2) The nonzero rows of any matrix in RREF are linearly independent. (Why?)
(3) A linearly independent set cannot contain 0.
(4) All subsets of a linearly independent set are also linearly independent.
(5) A set $\{a_1, ..., a_n\}$ in $\mathbb{R}^m$ is linearly independent if and only if $N(A_{m\times n}) = \{0_{n\times 1}\}$ where

$$A = \left[a_1 \vdots \cdots \vdots a_n\right]_{m\times n}$$

is formed by merging members of $\{a_1, ..., a_n\}$ as columns.

Lemma: $\{a_1, ..., a_n\} \subset \mathbb{R}^m$ is linearly dependent if and only if for some $k \in \{2, ..., n\}$, a_k is a linear combination of $a_1, ..., a_{k-1}$.

Proof: In one direction, if a subset of $\{a_1, ..., a_n\}$ is linearly dependent, then so is $\{a_1, ..., a_n\}$. In the other direction, suppose that $\{a_1, ..., a_n\}$ is linearly dependent, and let k be the first integer between 2 and n for which $a_1, .., a_k$ are linearly dependent. (Note that k may potentially be n.) Then $\sum_{i=1}^{k} \alpha_i a_i = 0$ for some $\alpha_1, ..., \alpha_k \in \mathbb{R}\backslash\{0\}$. Since α_k has to be different from zero by definition of k, we can write

$$a_k = \frac{-\alpha_1}{\alpha_k} a_1 + \cdots + \frac{-\alpha_{k-1}}{\alpha_k} a_{k-1}.$$

and the proof is complete. ■

Definition: Let S be a subspace of $\mathbb{R}^m$. A set $\{a_1, ..., a_n\} \subseteq S$ is a basis for S if it spans S and is linearly independent.

Note: Any two points in $\mathbb{R}^2$ not lying on the same line through the origin form a basis for $\mathbb{R}^2$. The collection of unit vectors $\{e_i\}_{i=1,...,m}$ is a basis for $\mathbb{R}^m$.

Theorem: Let S be a subspace of $\mathbb{R}^m$ with basis $\{a_1, ..., a_n\}$. If $\{b_1, ..., b_r\} \subset S$ is linearly independent, then $r \leq n$.

Proof: Let $A_{m\times n}$ be formed by merging a_i as columns:

$$A = \left[a_1 \vdots \cdots \vdots a_n\right]$$

We start by showing that $R(A) = S$. To begin pick $c \in R(A)$. There is some $x = (x_1 ..., x_n) \in \mathbb{R}^n$ -note abuse of notation: members of $\mathbb{R}^n$ are columns- such that $Ax = c$, i.e.,

$$\sum_{i=1}^{n} a_i x_i = c.$$

This means that $c \in S$ since S is a subspace and $a_i \in S$ for each i. Hence $R(A) \subseteq S$. Now pick $c \in S$. Since $\{a_1, ..., a_n\}$ is a basis for S, $c = \sum_{i=1}^{n} a_i x_i$ for some $x_1, ..., x_n \in \mathbb{R}$. In other words $Ax = c$ where $x = (x_1, ..., x_n)$ and $c \in R(A)$. Hence $S \subseteq R(A)$ as well. We conclude that $R(A) = S$.

(Proof continued) Let $\{b_1, ..., b_r\}$ be a linearly independent subset of S. Since $R(A) = S$, each $b_i \in R(A)$, i.e., there exists $y_i \in \mathbb{R}^n$ for each i such that $Ay_i = b_i$. Now let $M_{n\times r}$ be formed by merging y_i as columns and let $B_{m\times r}$ be formed by merging b_i as columns so that $AM = B$.

Suppose, towards a contradiction, that $r > n$. Then $rkM < r$ and $N(M) \neq \{0_{r\times 1}\}$. (Why?) There must, then, exist some $z = (z_1, ..., z_r) \in \mathbb{R}^r \backslash \{0\}$ such that $Mz = 0 \in \mathbb{R}^n$. Then $Bz = AMz = 0$, i.e.,

$$\left[b_1 \vdots \cdots \vdots b_r\right] z = \sum_{i=1}^{r} b_i z_i = 0_{m\times 1}$$

meaning that $\{b_1, ..., b_r\}$ is not linearly independent. This is the contradiction we needed. We conclude that $r \leq n$. ■

Corollary: Any $m+1$ vectors in $\mathbb{R}^m$ are linearly dependent.

Corollary: If $\{a_1, ..., a_n\}$ and $\{b_1, ..., b_r\}$ are two bases for S, then $n = r$.

Hence, we can talk about *the* number of elements in a basis for a subspace S.

Definition: The dimension of a subspace S, denoted $\dim S$, is the number of elements in any basis for S.

The following result is the Fundamental Theorem of Linear Algebra -part 1.

Theorem: Let A be $m \times n$. Then,

1. $\dim R(A^T) = rkA$,

2. $rkA = rkA^T$, and

3. $\dim N(A) = n - rkA$.

The proof consists of three steps:

(1) showing that the nonzero rows of A' constitute a basis for $R(A^T)$, which implies that $\dim R(A^T) = rkA$

(2) showing that $\dim R(A^T) = \dim R(A)$

(3) showing that $\dim N(A) = n - rkA$

Note that by steps (1) and (2), we will have

$$\begin{aligned} rkA &= \dim R(A^T) \\ &= \dim R(A) \\ &= rkA^T. \end{aligned}$$

Proof: Step 1... We will show that the nonzero rows of A' constitute a basis for $R(A^T)$.

To see this first note that the columns of A^T span $R(A^T)$. Hence rows of A span $R(A^T)$. Since $A' = E_r...E_1A$, rows of A' are linear combinations of rows of A and since $A = (E_r...E_1)^{-1}A'$, rows of A are linear combinations of rows of A'. Hence rows of A and rows of A' span the same subspace, $R(A^T)$.

Now drop the zero rows of A' (if they exist): the nonzero rows of A' span $R(A^T)$.

Finally note that the nonzero rows of A' are linearly independent. Thus, they form a basis for $R(A^T)$. We conclude: $\dim R(A^T) = rkA$.

(Similarly, $\dim R(A) = rkA^T$.)

(Proof continued) Step 2... We will show that $\dim R(A^T) = \dim R(A)$.

(By Step 1, The nonzero rows of $(A^T)'$ form a basis for $R(A)$, but this will not help us since we do not know how many of them there are. So we need to find another basis. Columns of A span $R(A)$ but they may not be LI.)

Let $rkA = k$ and let $\{j_1, ..., j_k\} \subseteq \{1, ..., n\}$ be the columns of A' (and of A) on which the first nonzero entries of nonzero rows occur in A'. I claim that these columns of A form a basis for $R(A)$.

Note first that if B is any matrix and if B^* is B with columns i and j interchanged, then (1) $R(B) = R(B^*)$ and (2) if $x \in N(B)$ then $x^* \in N(B^*)$ where x^* is x with the ith and jth entries interchanged.

(Proof continued) Now we can without loss of generality assume that $\{j_1, ..., j_k\} = \{1, ..., k\}$, i.e.,

$$A' = \left[\begin{array}{c} I_{k\times k} \vdots M_{k\times(n-k)} \\ \hline 0_{(m-k)\times n} \end{array}\right]$$

so that the claim we need to prove becomes: the first k columns of A form a basis for $R(A)$.

Write $A = \left[a_1 \vdots \cdots \vdots a_k \vdots a_{k+1} \vdots \cdots \vdots a_n\right]$. We need to show:

(i) $\{a_1, ..., a_k\}$ is linearly independent, and

(ii) $\{a_1, ..., a_k\}$ spans $R(A)$.

(Proof continued)

(i) Let $\alpha_1, ..., \alpha_k$ be such that $\sum_{i=1}^{k} \alpha_i a_i = 0$ and let

$$z = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_k \\ 0_{(n-k)\times 1} \end{bmatrix} \in \mathbb{R}^n.$$

Then $Az = 0 \Leftrightarrow z \in N(A) = N(A') \Leftrightarrow A'z = 0$, i.e., $\alpha_1 = \cdots = \alpha_k = 0$.

Thus $\{a_1, ..., a_k\}$ is linearly independent.

(Proof continued)

(ii) We know that $\{a_1, ..., a_n\}$ spans $R(A)$ so we only need to show that $a_{k+1}, ..., a_n$ are linear combinations of members of $\{a_1, ..., a_k\}$. Write

$$A' = \begin{bmatrix} I_{k\times k} & \vdots c^{k+1} \vdots \cdots \vdots c^n \\ 0_{(m-k)\times k} & \end{bmatrix} \text{ where } c^{k+j} = \begin{bmatrix} c_1^{k+j} \\ \vdots \\ c_m^{k+j} \end{bmatrix} \in \mathbb{R}^m$$

for all $j = 1, ..., n-k$ and let

$$x^{k+1} = \begin{bmatrix} c_1^{k+1} \\ \vdots \\ c_k^{k+1} \\ -1 \\ 0_{(n-k-1)\times 1} \end{bmatrix} \in \mathbb{R}^n.$$

Then $A'x^{k+1} = Ax^{k+1} = 0$, i.e., $a_1 c_1^{k+1} + ... + a_k c_k^{k+1} = a_{k+1}$. Similarly each a_{k+j} is a linear combination of members of $\{a_1, ..., a_k\}$, which, hence, spans $R(A)$.

(Proof continued) Step 3... To finish we need to show that $\dim N(A) = n - k$. Without loss of generality, once again let

$$A' = \left[\begin{array}{c:c} I_{k\times k} & M_{k\times(n-k)} \\ \hline 0_{(m-k)\times n} & \end{array}\right]$$

and let $x \in N(A) = N(A')$. Write

$$x = \left[\begin{array}{c} x^B_{k\times 1} \\ x^F_{(n-k)\times 1} \end{array}\right].$$

We have $A'x = 0$ implying that $x^B_{k\times 1} + Mx^F_{(n-k)\times 1} = 0$. Hence we can write x as

$$x = \left[\begin{array}{c} -M_{k\times(n-k)} \\ I_{(n-k)\times(n-k)} \end{array}\right] x^F_{(n-k)\times 1}.$$

(Proof continued)

So each $x \in N(A)$ is a linear combination of columns of

$$\begin{bmatrix} -M_{k\times(n-k)} \\ I_{(n-k)\times(n-k)} \end{bmatrix},$$

which are also linearly independent: if $\alpha \in \mathbb{R}^{n-k}$, then

$$\begin{bmatrix} -M_{k\times(n-k)} \\ I_{(n-k)\times(n-k)} \end{bmatrix} \alpha_{(n-k)\times 1} = \begin{bmatrix} -M\alpha \\ \alpha \end{bmatrix} = 0_{n\times 1}$$

which means that $\alpha = 0_{(n-k)\times 1}$. This proves that the columns of

$$\begin{bmatrix} -M_{k\times(n-k)} \\ I_{(n-k)\times(n-k)} \end{bmatrix}$$

constitute a basis of $N(A)$ and therefore that $\dim N(A) = n - k$. The proof is complete. ■

Orthogonality

If

$$A = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix},$$

then

$$\begin{aligned} R(A) &= \{(a, a) : a \in \mathbb{R}\}, \\ N(A) &= \{(0, a) : a \in \mathbb{R}\}, \\ R(A^T) &= \{(a, 0) : a \in \mathbb{R}\}, \\ N(A^T) &= \{(a, -a) : a \in \mathbb{R}\}. \end{aligned}$$

Note that $R(A)$ is orthogonal to $N(A^T)$ and $N(A)$ is orthogonal to $R(A^T)$.

Recall that if $y, z \in \mathbb{R}^m$, then $y \cdot z = y^T z = z^T y = \sum_{i=1}^m y_i z_i \in \mathbb{R}$.

Definition: Let S be a subspace of $\mathbb{R}^m$. The orthogonal complement of S is the set

$$S^\perp = \{y \in \mathbb{R}^m : y \cdot x = 0 \text{ for all } x \in S\}.$$

Note that $\{0_{m\times 1}\}^\perp = \mathbb{R}^m$ and $(\mathbb{R}^m)^\perp = \{0_{m\times 1}\}$. Also note that if S is a subspace of $\mathbb{R}^m$, then so is $S^\perp$. Furthermore,

Lemma: For any subspace S or $\mathbb{R}^m$, $S \cap S^\perp = \{0\}$.

Proof: $0 \in S$ and $0 \in S^\perp$ since both sets are subspaces. Suppose now that x is an arbitrary member of $S \cap S^\perp$. Then $x \cdot x = 0$ which is equivalent to $x = 0_{m\times 1}$. ■

We don't know yet, although it is true, that $(S^{\perp})^{\perp} = S$. To get there we need the Fundamental Theorem of Linear Algebra -part 2.

Theorem: Let A be $m \times n$. Then,

1. $R(A) = N(A^T)^{\perp}$,

2. $R(A)^{\perp} = N(A^T)$.

Proof: To prove the second equality, observe that

$$\begin{aligned} x \in R(A)^{\perp} &\Leftrightarrow x \cdot b = 0 \text{ for all } b \in R(A) \\ &\Leftrightarrow x \cdot Ay = 0 \text{ for all } y \in \mathbb{R}^n \\ &\Leftrightarrow x^T Ay = 0 \text{ for all } y \in \mathbb{R}^n \\ &\Leftrightarrow (A^T x)^T y = 0 \text{ for all } y \in \mathbb{R}^n \\ &\Leftrightarrow A^T x = 0_{n \times 1} \\ &\Leftrightarrow x \in N(A^T). \end{aligned}$$

(Proof continued) To prove the first equality, pick $b \in R(A)$ and $y \in N(A^T)$. Then there exists some $x \in \mathbb{R}^n$ such that $Ax = b$ and $b \cdot y = Ax \cdot y = x^T A^T y = 0$. So $b \in N(A^T)^\perp$.

Now pick $b \notin R(A)$. We need to show that $b \notin N(A^T)^\perp$, that is, we need to find some $x \in N(A^T)$ such that $b \cdot x \neq 0$.

Let A^* be the matrix of columns of A that form a basis for $R(A)$ and let $rkA = k$. Then A^* is $m \times k$, $rkA^* = rkA = k$, and $N(A^*) = \{0\}$.

Also note that $R(A) = R(A^*)$ and therefore $b \notin R(A^*)$ and $N(A^T) = R(A)^\perp = R(A^*)^\perp = N(A^{*T})$.

(Proof continued) Now we will show that $rk[A^*\vdots b]_{m\times(k+1)} = k+1$, which is equivalent to $N([A^*\vdots b]) = \{0_{(k+1)\times 1}\}$. To prove this claim let $y^{k+1} \in N([A^*\vdots b])$ and y^k be the vector in $\mathbb{R}^k$ consisting of the first k entries of y^{k+1}. Hence $[A^*\vdots b]y^{k+1} = 0$ and $A^*y^k + by_{k+1} = 0$. Solving,

$$b = \frac{-1}{y_{k+1}} A^* y^k,$$

unless $y_{k+1} = 0$. But this would imply that $b \in R(A^*)$ so we conclude that $y_{k+1} = 0$. Then $A^*y^k = 0$, that is $y^k = 0 \in \mathbb{R}^k$. Hence $y^{k+1} = 0 \in \mathbb{R}^{k+1}$ and $N([A^*\vdots b]) = \{0\}$.

(Proof continued) We conclude that

$$rk[A^* \vdots b] = k+1 = rk \begin{bmatrix} A^{*T} \\ b^T \end{bmatrix} \text{ and } R\left(\begin{bmatrix} A^{*T} \\ b^T \end{bmatrix} \right) = \mathbb{R}^{k+1}.$$

This means that there exists some $x \in \mathbb{R}^m$ such that

$$\begin{bmatrix} A^{*T} \\ b^T \end{bmatrix} x = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}_{(k+1)\times 1}.$$

Then $A^{*T}x = 0_{k\times 1}$ and $x \in N(A^{*T}) = N(A^T)$ and $b^T x = 1 \neq 0$. So $b \notin N(A^T)^{\perp}$. The proof is complete. ■

Corollary: For any subspace S of $\mathbb{R}^m$, $(S^\perp)^\perp = S$.

Proof: To see this let S be a subspace with basis $\{a_1, ..., a_n\}$. Let $A = [a_1 \vdots \cdots \vdots a_n]$ so that $R(A) = S$. We have

$$\begin{aligned} S^\perp &= R(A)^\perp \\ &= N(A^T) \end{aligned}$$

and

$$(S^\perp)^\perp = N(A^T)^\perp = R(A) = S$$

which finishes the proof. ■

Inverses

Theorem: Let A and C be both $n \times n$. If $AC = I$, then $rkA = rkC = n$.

Proof: It suffices to show that $\dim N(A) = \dim N(C) = 0$. Pick $x \in N(C)$. Then

$$\begin{aligned} Cx &= 0 \\ ACx &= 0 \end{aligned}$$

i.e., $x = 0_{n\times 1}$. Hence $N(C) = \{0_{n\times 1}\}$ and $\dim N(C) = 0$. Then $rkC = n$ and $R(C) = \mathbb{R}^n$. Pick $x \in N(A)$. There exists some $y \in \mathbb{R}^n$ such that $Cy = x$. Then $ACy = 0 \Leftrightarrow y = 0 \Leftrightarrow x = 0$, i.e., $N(A) = \{0_{n\times 1}\}$ and $\dim N(A) = 0$. ■

Theorem: Let A and C be both $n \times n$. $AC = I \Rightarrow CA = I$.

Proof: Note that $rkA = n$ by the result above. Then

$$\begin{aligned} A' &= I \\ &= E_r ... E_1 A \end{aligned}$$

and

$$E_r ... E_1 \underbrace{AC}_{=I} = C$$

giving us (right-multiply both hand sides by A and recall that $A' = I$) $I = CA$. This finishes the proof. ■

Note that the preceding result is actually an if and only if statement.

Definition: $A_{n\times n}$ is invertible if there exists $C_{n\times n}$ such that $AC = CA = I$.

Theorem: If $AC = CA = I$ and if $AB = BA = I$, then $C = B$.

Proof: We have $CA = I$, implying $CAB = B$, implying, $C = B$. ■

Two simple remarks:

1. If A and B are invertible, then AB is invertible as well.
2. If A is invertible, then so is A^T and $(A^T)^{-1} = \left(A^{-1}\right)^T$.