Chapter 3. Matrices and Concrete Linear Maps

The Row Space, Column Space and Null Space of a Matrix

3.1 Definition: Let R be a ring. For a matrix $A \in M_{m\times n}(R)$, the **row span** of A, denoted by $\text{Row}(A)$, is the span of the rows of A, the **column span** of A, denoted by $\text{Col}(A)$, is the span of the columns of A, and the **null set** of A, is the set

$$\text{Null}(A) = \left\{x \in F^n \middle| Ax = 0\right\}.$$

When F is a field, $\text{Row}(A)$ and $\text{Col}(A)$ are also called the **row space** and **column space** of A, and we define the **rank** of A and the **nullity** of A are the dimensions

$$\text{rank}(A) = \dim\big(\text{Col}A\big) \ \text{ and } \ \text{nullity}(A) = \dim\big(\text{Null}A\big).$$

3.2 Note: For $A = (u_1, u_2, \cdots, u_n) \in M_{m\times n}(R)$ and $t \in R^n$ we have $At = \sum_{i=1}^{n} t_i u_i$, so

$$\text{Col}(A) = \left\{At \middle| t \in R^n\right\}.$$

3.3 Theorem: *Let F be a field, let $A \in M_{m\times n}(F)$ and let $b \in F^m$. If $x = p$ is a solution to the equation $Ax = b$ then*

$$\left\{x \in F^n \middle| Ax = b\right\} = p + \text{Null}(A).$$

Proof: If $Ap = b$ then for $x \in F^n$ we have

$$Ax = b \iff Ax = Ap \iff A(x-p) = 0 \iff (x-p) \in \text{Null}A \iff x \in p + \text{Null}(A).$$

3.4 Note: For $\mathcal{A} = \{u_1, u_2, \cdots, u_n\} \subseteq F^m$ and $A = (u_1, u_2, \cdots, u_n) \in M_{m\times n}(F)$,

$\mathcal{A}$ is linearly independent

$\iff$ for all $t_1, t_2, \cdots, t_n \in F$, if $\sum_{i=1}^{n} t_i u_i = 0$ then each $t_i = 0$

$\iff$ for all $t \in F^n$, if $At = 0$ then $t = 0$

$\iff \text{Null}(A) = \{0\} \iff \text{Null}(R) = \{0\}$

$\iff R$ has a pivot in every column $\iff R$ is of the form $R = \begin{pmatrix} I \\ 0 \end{pmatrix}$, and

$\mathcal{A}$ spans $F^m \iff \text{Col}(A) = F^m$

$\iff$ for every $x \in F^m$ there exists $t \in F^n$ with $At = x$

$\iff$ for every $y \in F^m$ there exists $t \in F^n$ with $Rt = y$

$\iff R$ has a pivot in every row.

3.5 Theorem: *Let F be a field, let $A = (u_1, u_2, \cdots, u_n) \in M_{m\times n}(F)$, and suppose $A \sim R$ where R is in reduced row echelon form with pivots in columns $1 \le j_1 < j_2 < \cdots < j_r \le n$. Then*

(1) the non-zero rows of R form a basis for $\text{Row}(A)$,

(2) the set $\{u_{j_1}, u_{j_2}, \cdots, u_{j_r}\}$ is a basis for $\text{Col}(A)$, and

(3) when we solve $Ax = b$ using Gauss-Jordan elimination and write the solution as $x = p + Bt$ as in Note 2.16, the columns of B form a basis for $\text{Null}(A)$.

Proof: First we prove Part (1). By Theorem 1.31, when we perform an elementary row operation on a matrix, the span of the rows is unchanged, and so we have $\text{Row}(A) = \text{Row}(R)$. The nonzero rows of R span $\text{Row}(R)$, so it suffices to show that the nonzero rows of R are linearly independent. Let $1 \leq j_1 < j_2 < \cdots < j_r$ be the indices of the pivot columns in R. Let $v_1, v_2, \cdots, v_r$ be the nonzero rows of R. Because R is in reduced row echelon form, for $1 \leq i \leq r$ and $1 \leq k \leq r$ we have $(v_i)_{j_k} = \delta_{i,k}$. It follows that $\{v_1, v_2, \cdots, v_r\}$ is linearly independent because if $\sum_{i=1}^{r} t_i v_i = 0$ with each $t_i \in F$ then for all k with $1 \leq k \leq r$ we have

$$0 = \Big(\sum_{i=1}^{r} t_i v_i\Big)_{j_k} = \sum_{i=1}^{r} t_i (v_i)_{j_k} = \sum_{i=1}^{r} t_i \delta_{i,k} = t_k.$$

To prove Part (2), let $1 \leq l_1 < l_2 < \cdots, l_{n-r} \leq n$ be the indices of the non-pivot columns. Let $v_1, v_2, \cdots, v_n \in F^m$ be the columns of R and note that we have $v_{j_i} = e_i$ for $1 \leq i \leq r$. When we use row operations to reduce A to R, the same row operations reduce $A_J = (u_{j_1}, \cdots, u_{j_r})$ to $R_J = (v_{j_1}, \cdots, v_{j_t}) = (e_1, \cdots, e_r) = \begin{pmatrix} I \\ 0 \end{pmatrix}$. This shows that $\{u_{j_1}, \cdots, u_{j_r}\}$ is linearly independent. When we use row operations to reduce A to R, the same row operations will reduce $(A|u_k)$ to $(R|v_k)$, and so the equation $Ax = u_k$ has the same solutions as the equation $Rx = v_k$. Since only the first r columns of R are nonzero, each column v_k can be written as $v_k = \sum_{i=1}^{r} (v_k)_i e_i = \sum_{i=1}^{r} (v_k)_i v_{j_i} = Rt$ where $t \in R^n$ is given by $t_J = v_k$ and $t_L = 0$. Since $Ax = u_k$ and $Rx = v_k$ have the same solutions, we also have $u_k = At = \sum_{i=1}^{r} (v_k)_i u_{j_i} \in \text{Span}\{u_{j_1}, u_{j_2}, \cdots, u_{j_r}\}$. This shows that $\text{Col}(A) = \text{Span}\{u_1, u_2, \cdots, u_n\} = \text{Span}\{u_{j_1}, \cdots, u_{j_r}\}$.

Since the solution set to the equation $Ax = b$ is the set

$$\{x \in \mathbf{R}^n | Ax = b\} = p + \text{Col}(B) = p + \text{Null}(A)$$

we must have $\text{Col}(B) = \text{Null}(A)$. Since (as in Note 2.16) we have $B_L = I$, it follows that the columns of B are linearly independent using the same argument that we used in Part (1) to show that the nonzero rows of R are linearly independent. This proves Part (3).

3.6 Corollary: *Let F be a field, let $A \in M_{m\times n}(F)$, suppose that A is row equivalent to a reduced row echelon matrix which has r pivots. Then*

$$\text{rank}(A) = \dim(\text{Row}A) = \dim(\text{Col}A) = r\ , \text{ and}$$
$$\text{nullity}(A) = \dim(\text{Null}A) = n - r.$$

3.7 Corollary: *Let F be a field, let $A \in M_{m\times n}(F)$ and suppose that A is row equivalent to a row reduced echelon matrix R.*

(1) The rows of A are linearly independent $\iff$ the columns of A span F^m $\iff$ $\text{rank}(A) = m$ $\iff$ R has a pivot in every row.

(2) The rows of A span F^n $\iff$ the columns of A are linearly independent $\iff$ $\text{rank}(A) = n$ $\iff$ R has a pivot in every column $\iff$ R is of the form $R = \begin{pmatrix} I \\ 0 \end{pmatrix}$.

(3) The rows of A form a basis for $\mathbf{R}^n$ $\iff$ the columns of A form a basis for F^m $\iff$ $\text{rank}(A) = m = n$ $\iff$ $R = I$.

Matrices and Linear Maps

3.8 Definition: Let R be a ring. A map $L : R^n \to R^m$ is called **linear** when
(1) $L(x+y) = L(x) + L(y)$ for all $x, y \in R^n$, and
(2) $L(tx) = t\,L(x)$ for all $x \in R^n$ and all $t \in R$.

3.9 Note: Given a matrix $A \in M_{m\times n}(R)$, the map $L : \mathbf{R}^n \to R^m$ given by $L(x) = Ax$ is linear.

3.10 Theorem: *Let $L : R^n \to R^m$ be linear. There exists a unique matrix $A \in M_{m\times n}(R)$ such that $L(x) = Ax$ for all $x \in R^n$, namely the matrix $A = \big(L(e_1), L(e_2), \cdots, L(e_n)\big)$.*

Proof: Let $L : R^n \to R^m$ and let $A = (u_1, u_2, \cdots, u_n) \in M_{m\times n}(R)$. If $L(x) = Ax$ for all $x \in R$ then for each index k we have $u_k = Ae_k = L(e_k)$. Conversely, suppose that $u_k = L(e_k)$ for every index k. Then for all $x \in R^n$ we have

$$L(x) = L\big(\sum_{i=1}^{n} x_i e_i\big) = \sum_{i=1}^{n} x_i\, L(e_i) = \sum_{i=1}^{n} x_i u_i = Ax\,.$$

3.11 Notation: Often, we shall not make a notational distinction between the matrix $A \in M_{m\times n}(R)$ and its corresponding linear map $A : R^n \to R^m$ given by $A(x) = Ax$. When we do wish to make a distinction, we shall use the following notation. Given a matrix $A \in M_{m\times n}(R)$ we let $L_A : R^n \to R^n$ be the linear map given by

$$L_A(x) = Ax \ \text{ for all } x \in R^n$$

and given a linear map $L : R^n \to R^m$ we let $[L]$ be the corresponding matrix given by

$$[L] = \big(L(e_1), L(e_2), \cdots, L(e_n)\big) \in M_{m\times n}(R).$$

3.12 Definition: For a linear map $L : R^n \to R^m$, the **kernel** (or the **null set** of L is the set

$$\text{Ker}(L) = \text{Null}(L) = L^{-1}(0) = \{x \in R^n | L(x) = 0\}$$

and the **image** (or the **range** of L) is the set

$$\text{Image}(L) = \text{Range}(L) = L(R^n) = \{L(x) | x \in R^n\}.$$

We also use the same terminology for a matrix $A \in M_{m\times n}(R)$ when we think of the matrix as a linear map, so when $A = [L]$ we have $\text{Ker}(L) = \text{Null}(L) = \text{Ker}(A) = \text{Null}(A)$ and $\text{Image}(L) = \text{Range}(L) = \text{Image}(A) = \text{Range}(A) = \text{Col}(A)$. When F is a field and $L : F^n \to F^m$ is linear, we define the **rank** and the **nullity** of L to be the dimensions

$$\text{rank}(L) = \dim\big(\text{Range}(L)\big) \ \text{ and } \ \text{nullity}(L) = \dim\big(\text{Null}(L)\big).$$

3.13 Theorem: *Let R be a ring and let $L : R^n \to R^m$ be a linear map. Then*
(1) L is surjective if and only if $\text{Range}(L) = F^m$, and
(2) L is injective if and only if $\text{Null}(L) = \{0\}$.

Proof: Part (1) is obvious, so we only prove Part (2). Note that since L is linear we have $L(0) = L(0\cdot 0) = 0\,L(0) = 0$ and so $0 \in \text{Null}(L)$. Suppose that L is injective. Then for $x \in R^n$ we have $x \in \text{Null}(L) \Longrightarrow L(x) = 0 \Longrightarrow L(x) = L(0) \Longrightarrow x = 0$ so $\text{Null}(L) = \{0\}$. Conversely, suppose that $\text{Null}(L) = \{0\}$. Then for $x, y \in R^n$ we have

$$L(x) = L(y) \Longrightarrow L(x-y) = 0 \Longrightarrow (x-y) \in \text{Null}(L) = \{0\} \Longrightarrow x - y = 0 \Longrightarrow x = y$$

and so L is injective.

3.14 Example: The **identity map** on R^n is the map $I : R^n \to R^n$ given by $I(x) = x$ for all $x \in R^n$, and it corresponds to the identity matrix $I \in M_n(R)$ with entries $I_{i,j} = \delta_{i,j}$. The **zero map** $O : R^n \to R^m$ given by $O(x) = 0$ for all $x \in R^n$ corresponds to the zero matrix $O \in M_{m\times n}(R)$ with entries $O_{i,j} = 0$ for all i, j.

3.15 Note: Given linear maps $L, M : R^n \to R^m$ and $K : R^m \to R^l$ and given $t \in R$, the maps $(L + M) : R^n \to R^n$, tL and $KL : R^n \to R^l$ given by $(L + M)(x) = L(x) + M(x)$, $(tL)(x) = t\,L(x)$ and $(KL)(x) = K(L(x))$ are all linear. For example, to see that KL is linear, note that for $x, y \in R^n$ and $t \in R$ we have

$$\begin{aligned}(KL)(x + y) &= K\big(L(x + y)\big) = K\big(L(x) + L(y)\big)\\ &= K\big(L(x)\big) + K\big(L(y)\big) = (KL)(x) + (KL)(y)\ \text{, and}\\ KL(tx) &= K\big(L(tx)\big) = K\big(t\,L(x)\big) = t\,K\big(L(x)\big) = t(KL)(x).\end{aligned}$$

3.16 Definition: Given $A, B \in M_{m\times n}(R)$ we define $A + B \in M_{m\times n}(R)$ to be the matrix such that $(A + B)(x) = Ax + Bx$ for all $x \in R^n$. Given $A \in M_{m\times n}(R)$ and $t \in R$, we define $tA \in M_{m\times n}(R)$ to be the matrix such that $(tA)(x) = t\,Ax$ for all $x \in R^n$. Given $A \in M_{l\times m}(R)$ and $B \in M_{m\times n}(R)$ we define $AB \in M_{l\times n}(R)$ to be the matrix such that $(AB)x = A(Bx)$ for all $x \in R^n$.

3.17 Note: From the above definitions, it follows immediately that for all matrices A, B, C of appropriate sizes and for all $s, t \in R$, we have
(1) $(A + B) + C = A + (B + C)$,
(2) $A + B = B + A$,
(3) $O + A = A = A + O$,
(4) $A + (-A) = 0$,
(5) $(AB)C = A(BC)$,
(6) $IA = A = AI$,
(7) $OA = O$ and $AO = O$,
(8) $(A + B)C = AC + BC$ and $A(B + C) = AB + AC$,
(9) $s(tA) = (st)A$,
(10) if R is commutative then $A(tB) = t(AB)$,
(11) $(s + t)A = sA + tA$ and $t(A + B) = tA + tB$, and
(12) $0A = O$, $1A = A$ and $(-1)A = -A$.
In particular, the set $M_n(R)$ is a ring under addition and multiplication of matrices.

3.18 Theorem: *For $A, B \in M_{m\times n}(R)$ and $t \in R$, the matrices $A + B$ and tA are given by $(A + B)_{i,j} = A_{i,j} + B_{i,j}$ and $(tA)_{i,j} = t\,A_{i,j}$. For $A = (u_1, u_2, \cdots, u_l)^T \in M_{l\times m}(R)$ and $B = (v_1, v_2, \cdots, v_n) \in M_{m\times n}(R)$, the matrix AB is given by*

$$(AB)_{j,k} = {v_j}^T u_k = \sum_{i=1}^{m} A_{j,i}B_{i,k}\,.$$

Proof: For $A, B \in M_{m\times n}(R)$, the k^{th} column of $(A+B)$ is equal to $(A+B)e_k = Ae_k+Be_k$ which is the sum of the k^{th} columns of A and B. It follows that $(A + B)_{j,k} = A_{j,k} + B_{j,k}$ for all j, k. Similarly for $t \in R$, the k^{th} column of tA is equal to $(tA)e_k = t\,Ae_k$ which is t times the k^{th} column of A.

Now let $A = (u_1, \cdots, u_l)^T \in M_{l\times m}(R)$ and $B = (v_1, \cdots, v_n) \in M_{m\times n}(R)$. The k^{th} column of (AB) is equal to $(AB)e_k = A(Be_k) = Av_k$, so the (j, k) entry of AB is equal to

$$(AB)_{j,k} = {v_j}^T u_k = \big(A_{j,1}, A_{j,2}, \cdots, A_{j,m}\big)$$

The Transpose and the Inverse

3.19 Definition: For a linear map $L : R^n \to R^m$ the **transpose** of L is the map $L^T : R^m \to R^n$ such that $[L^T] = [L]^T$.

3.20 Note: When R is a ring, for $A \in M_{m\times n}(R)$ we have $\text{Row}(A) = \text{Col}(A^T)$ and $\text{Col}(A) = \text{Row}(A^T)$. When F is a field, for $A \in M_{m\times n}(F)$ we have $\text{rank}(A) = \text{rank}(A^T)$ and for a linear map $L : R^n \to R^m$ we have $\text{rank}(L) = \text{rank}(L^T)$.

3.21 Definition: For linear maps $L : R^n \to R^m$ and $M : R^m \to R^n$, when $LM = I$ where $I : R^m \to R^m$ we say that L is a **left inverse** of M and that M is a **right inverse** of L, and when $LM = I$ and $ML = I$ we say that L and M are (two-sided) **inverses** of each other. When $L : R^n \to R^m$ has a (two-sided) inverse $M : R^m \to R^n$ we say that L is **invertible**. We use the same terminology for matrices $A \in M_{m\times n}(R)$ and $B \in M_{n\times m}(R)$.

3.22 Theorem: *Let R be a ring, let $A \in M_{m\times n}(R)$ and $B \in M_{n\times m}(R)$. If B is a left inverse of A and C is a right inverse of A then $B = C$. A similar result holds for linear maps $L : R^n \to R^m$ and $K, M : R^m \to R^n$.*

Proof: Suppose that $BA = I$ and that $AC = I$. Then

$$B = BI = B(AC) = (BA)C = IC = C\,.$$

3.23 Theorem: *Let R be a commutative ring.*
(1) For $A, B \in M_{m\times n}(R)$ and $t \in R$ we have

$$(A^T)^T = A\ ,\ (A+B)^T = A^T + B^T \ \text{ and }\ (tA)^T = t\,A^T.$$

A similar result holds for linear maps $L, M : \mathbf{R}^n \to R^m$.
(2) If $A \in M_{l\times m}(R)$ and $B \in M_{m\times n}(R)$ then

$$(AB)^T = B^T A^T.$$

A similar result holds for linear maps $L : R^l \to R^m$ and $M : R^m \to R^n$.
(3) For invertible matrices $A, B \in M_n(R)$ and for an invertible element $t \in R$ we have

$$(A^{-1})^{-1}\ ,\ (tA)^{-1} = \tfrac{1}{t}\,A^{-1} \ \text{ and }\ (AB)^{-1} = B^{-1}A^{-1}.$$

A similar result holds for invertible linear maps $L, M : R^n \to R^n$.

Proof: We leave the proof of Part (1) as an exercise. To prove Part (2), suppose that R is commutative and let $A \in M_{l\times m}(R)$ and $B \in M_{m\times n}(R)$. Then for all indices j, k we have

$$(AB)^T{}_{j,k} = (AB)_{k,j} = \sum_{i=1}^m A_{k,i}B_{i,j} = \sum_{i=1}^m B_{i,j}A_{k,i} = \sum_{i=1}^m B^T j, i A^T{}_{i,k} = (B^T A^T)_{j,k}.$$

To prove Part (3), let $A, B \in M_n(R)$ be invertible matrices and let $t \in R$ be an invertible element. Because $AA^{-1} = I$ and $A^{-1}A = I$, it follows that $(A^{-1})^{-1} = A$. Because $(tA)\big(\tfrac{1}{t}A\big) = \big(t\cdot\tfrac{1}{t}\big)\,AA^{-1} = 1\cdot I = I$ and similarly $\big(\tfrac{1}{t}A\big)\big(tA\big) = I$, it follows that $(tA)^{-1} = \tfrac{1}{t}\,A$. Because $(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AA^{-1} = I$ and similarly $B^{-1}A^{-1})(AB) = I$, it follows that $(AB)^{-1} = B^{-1}A^{-1}$.

3.24 Theorem: *Let F be a field and let $A \in M_{m\times n}(F)$,*
(1) A is surjective $\iff$ A has a right inverse matrix,
(2) A is injective $\iff$ A has a left inverse matrix,
(3) if A is bijective then $n = m$ and A has a (two-sided) inverse matrix, and
(4) when $n = m$, A is bijective $\iff$ A is surjective $\iff$ A is injective.

A similar result holds for a linear map $L : R^n \to R^m$.

Proof: We prove Part (1). Suppose first that A has a right inverse matrix, say $AB = I$ with $B \in M_{n\times n}(F)$. Then given $y \in F^m$ we can choose $x \in F^n$ to get

$$Ax = A(By) = (AB)y = Iy = y.$$

Thus A is surjective. Conversely, suppose that A is surjective. For each index $k \in \{1, 2, \cdots, m\}$, choose $u_k \in F^n$ so that $Au_k = e_k$, and then let $B = (u_1, u_2, \cdots, u_m) \in M_{n\times m}(F)$. Then we have

$$AB = A(u_1, u_2, \cdots, u_m) = \big(Au_1, Au_2, \cdots, Au_m\big) = (e_1, e_2, \cdots, e_m) = I\,.$$

To prove Part (2), suppose first that A has a left inverse matrix, say $BA = I$ with $B \in M_{n\times n}(F)$. Then for $x \in F^n$ we have

$$Ax = 0 \Longrightarrow B(Ax) = 0 \Longrightarrow (BA)x = 0 \Longrightarrow Ix = 0 \Longrightarrow x = 0$$

and so $\text{Null}(A) = \{0\}$. Thus A is injective. Conversely, suppose that A is injective. Then $\text{Null}(A) = \{\}$, so the columns of A are linearly independent, hence the rows of A span F^n, equivalently the columns of A^T span F^n, hence $\text{Range}(A^T) = F^n$ and so A^T is surjective. Since A^T is surjective, we can choose $C \in M_{m\times n}(F)$ so that $A^TC = I$. Let $B = C^T$ so that $A^TB^T = I$. Transpose both sides to get $BA = I^T = I$. Thus the matrix B is a left inverse of A.

Parts (3) and (4) follow easily from Parts (1) and (2) together with previous results (namely Note 3.4, Corollary 3.7 and Theorems 3.13 and 3.22).

3.25 Note: To obtain a right inverse of a given matrix $A \in M_{m\times n}(F)$ using the method described in the proof of Part (1) of the above theorem, we can find vectors $u_1, u_2, \cdots, u_m \in F^n$ such that $Au_k = e_k$ for each index k by reducing each of the augmented matrices $(A|e_k)$. Since the same row operations which are used to reduce $(A|e_1)$ to the form $(R|u_1)$, (with R in reduced echelon form) will also reduce each of the augmented matrices $(A|e_k)$ to the form $(R|e_k)$, we can solve all of the equations $Au_k = e_k$ simultaneously by reducing the matrix $(A|I) = (A|e_1, e_2, \cdots, e_m)$ to the form $(R|u_1, u_2, \cdots, u_m)$.

3.26 Example: Let $A = \begin{pmatrix} 1 & 3 & 2 \\ 2 & 4 & 1 \\ 1 & 1 & 0 \end{pmatrix} \in M_3(\mathbf{Q})$. Find A^{-1}.

Solution: We have

$$(A|I) = \left(\begin{array}{ccc|ccc} 1 & 3 & 2 & 1 & 0 & 0 \\ 2 & 4 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 \end{array}\right) \sim \left(\begin{array}{ccc|ccc} 1 & 3 & 2 & 1 & 0 & 0 \\ 0 & -2 & -3 & -2 & 1 & 0 \\ 0 & -2 & -2 & -1 & 0 & 1 \end{array}\right)$$
$$\sim \left(\begin{array}{ccc|ccc} 1 & 3 & 2 & 1 & 0 & 0 \\ 0 & 1 & \frac{3}{2} & 1 & -\frac{1}{2} & 0 \\ 0 & -2 & -2 & -1 & 0 & 1 \end{array}\right) \sim \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{1}{2} & -1 & \frac{5}{2} \\ 0 & 1 & 0 & -\frac{1}{2} & 1 & -\frac{3}{2} \\ 0 & 0 & 1 & 1 & -1 & 1 \end{array}\right)$$

and so A^{-1} is equal to the matrix whichh appears on the right of the final matrix above.