4.4 Row equivalence, column equivalence and diagonal form

Given $T : U \to V$. Let $\mathcal{A}$ and $\mathcal{B}$ be basis for U and V respectively. The matrix $M_{\mathcal{A}}^{\mathcal{B}}T$ is determined by the choice of $\mathcal{A}$ and $\mathcal{B}$. We are interested in:

1. Fix $\mathcal{B}$, simplifying $M_{\mathcal{A}}^{\mathcal{B}}T$ by varying $\mathcal{A}$.
2. Fix $\mathcal{A}$, simplifying $M_{\mathcal{A}}^{\mathcal{B}}T$ by varying $\mathcal{B}$.
3. simplifying $M_{\mathcal{A}}^{\mathcal{B}}T$ by varying both $\mathcal{A}$ and $\mathcal{B}$

Consider the third case: Let $\mathcal{A} = \{u_1, ..., u_n\}$ be a basis for U such that $\{u_{k+1}, ..., u_n\}$ is a basis for $Ker(T)$ for some $k \leq n$. Define

$$v_i = T(u_j),\ \ j = 1, ..., k.$$

Then $\{v_1, ..., v_k\}$ is a basis for $Im(T)$ which is a subspace of V. If $Im(T) = V$, then $\{v_1, ..., v_k\}$ is a basis for V. In general, let

$$\mathcal{B} = \{v_1, ..., v_k, v_{k+1}, ..., v_m\}$$

be a basis for V. For each $j = 1, ..., n$, if $j \leq k$,

$$T(u_j) = a_{1j}v_1 + ... + a_{ij}v_i + ... + a_{mj}v_m = 0v_1 + ... + 1v_i + ... + 0v_m = v_i,\ \text{with } i = j.$$

Otherwise

$$T(u_j) = a_{1j}v_1 + ... + a_{ij}v_i + ... + a_{mj}v_m = 0v_1 + ... + 0v_i + ... + 0v_m = 0.$$

Hence

$$M_{\mathcal{A}}^{\mathcal{B}}T = \begin{pmatrix} I_k & 0 \\ 0 & 0 \end{pmatrix}$$

where I_k is a $k \times k$ identity matrix.

In the first case, suppose $\mathcal{B}$ is fixed and we want to change from $\mathcal{A}$ to $\mathcal{A}'$. Recall:

$$M_{\mathcal{A}'}^{\mathcal{B}}T = M_{\mathcal{A}}^{\mathcal{B}}T \cdot M_{\mathcal{A}'}^{\mathcal{A}}I_U.$$

Clearly, if $\mathcal{A} = \mathcal{A}'$, then $M_{\mathcal{A}'}^{\mathcal{A}}I_U = I_n$ is an $n \times n$ identity matrix. Suppose

$$\mathcal{A} = \{u_1, u_2, ..., u_n\}$$

is a basis for U. Let Let

$$M_{\mathcal{A}}^{\mathcal{B}}T = \begin{pmatrix} t^1 & t^2 & ... & t^n \end{pmatrix}$$

written as column vectors.

1. **Interchange vectors:** Suppose
$$\mathcal{A}' = \{u_1', u_2', ..., u_n'\} = \{u_2, u_1, ..., u_n\}.$$
Then $I_U(u_1') = u_2 = a_{12}u_2$, $I_U(u_2') = u_1 = a_{21}u_1$ and $I_U(u_j') = u_j = a_{jj}u_j$ for all $j \geq 3$. This implies
$$M_{\mathcal{A}'}^{\mathcal{A}} I_U = \begin{pmatrix} 0 & 1 & 0 & ... & 0 \\ 1 & 0 & 0 & ... & 0 \\ 0 & 0 & 1 & ... & 0 \\ ... \\ 0 & 0 & 0 & ... & 1 \end{pmatrix} \tag{4.2}$$
In other words, $M_{\mathcal{A}'}^{\mathcal{A}} I_U$ is changed by switching the column vectors. Then
$$M_{\mathcal{A}'}^{\mathcal{B}} T = M_{\mathcal{A}}^{\mathcal{B}} T \cdot M_{\mathcal{A}'}^{\mathcal{A}} I_U = \begin{pmatrix} t^2 & t^1 & ... & t^n \end{pmatrix}$$

2. **scalar multiplication:** Suppose
$$\mathcal{A}' = \{u_1', u_2', ..., u_n'\} = \{u_1, cu_2, ..., u_n\}.$$
Then
$$M_{\mathcal{A}'}^{\mathcal{A}} I_U = \begin{pmatrix} 1 & 0 & 0 & ... & 0 \\ 0 & c & 0 & ... & 0 \\ 0 & 0 & 1 & ... & 0 \\ ... \\ 0 & 0 & 0 & ... & 1 \end{pmatrix} \tag{4.3}$$
Then
$$M_{\mathcal{A}'}^{\mathcal{B}} T = M_{\mathcal{A}}^{\mathcal{B}} T \cdot M_{\mathcal{A}'}^{\mathcal{A}} I_U = \begin{pmatrix} t^1 & ct^2 & ... & t^n \end{pmatrix}$$

3. **Adding multiple of one basis element to the other:** Suppose
$$\mathcal{A}' = \{u_1', u_2', ..., u_n'\} = \{u_1, u_2 + cu_1, ..., u_n\}.$$
Then $I_U(u_2') = u_2 + cu_1 = a_{22}u_2 + a_{12}u_1$. Therefore,
$$M_{\mathcal{A}'}^{\mathcal{A}} I_U = \begin{pmatrix} 1 & c & 0 & ... & 0 \\ 0 & 1 & 0 & ... & 0 \\ 0 & 0 & 1 & ... & 0 \\ ... \\ 0 & 0 & 0 & ... & 1 \end{pmatrix} \tag{4.4}$$
Then
$$M_{\mathcal{A}'}^{\mathcal{B}} T = M_{\mathcal{A}}^{\mathcal{B}} T \cdot M_{\mathcal{A}'}^{\mathcal{A}} I_U = \begin{pmatrix} t^1 & t^2 + ct^1 & ... & t^n \end{pmatrix}$$

Suppose next that $\mathcal{A}$ is fixed and we want to change $\mathcal{B}$ to $\mathcal{B}'$. Recall,

$$M_{\mathcal{A}}^{\mathcal{B}'} = M_{\mathcal{B}}^{\mathcal{B}'} I_V \cdot M_{\mathcal{A}}^{\mathcal{B}} T.$$

Let

$$M_{\mathcal{A}}^{\mathcal{B}} T = \begin{pmatrix} t_1 \\ t_2 \\ ... \\ t_m \end{pmatrix}$$

be written as row vectors. Let $\mathcal{B} = \{v_1, ..., v_m\}$ be a basis for V. Following the same ideas, we obtain:

1. **Interchange vectors:** Suppose

$$\mathcal{B}' = \{v_1', v_2', ..., v_m'\} = \{v_2, v_1, ..., v_m\}.$$

   Then

$$M_{\mathcal{B}}^{\mathcal{B}'} I_V = \begin{pmatrix} 0 & 1 & 0 & ... & 0 \\ 1 & 0 & 0 & ... & 0 \\ 0 & 0 & 1 & ... & 0 \\ ... & & & & \\ 0 & 0 & 0 & ... & 1 \end{pmatrix} \tag{4.5}$$

$$M_{\mathcal{A}}^{\mathcal{B}'} T = M_{\mathcal{B}}^{\mathcal{B}'} I_V \cdot M_{\mathcal{A}}^{\mathcal{B}} T = \begin{pmatrix} t_2 \\ t_1 \\ ... \\ t_m \end{pmatrix}$$

2. **scalar multiplication:** Suppose

$$\mathcal{B}' = \{v_1', v_2', ..., v_m'\} = \{v_1, cv_2, ..., v_m\}.$$

   Then

$$M_{\mathcal{B}}^{\mathcal{B}'} I_V = \begin{pmatrix} 1 & 0 & 0 & ... & 0 \\ 0 & c & 0 & ... & 0 \\ 0 & 0 & 1 & ... & 0 \\ ... & & & & \\ 0 & 0 & 0 & ... & 1 \end{pmatrix} \tag{4.6}$$

$$M_{\mathcal{A}}^{\mathcal{B}'} = M_{\mathcal{B}}^{\mathcal{B}'} I_V \cdot M_{\mathcal{A}}^{\mathcal{B}} T = \begin{pmatrix} t_1 \\ ct_2 \\ ... \\ t_m \end{pmatrix}$$

3. **Adding multiple of one basis element to the other:** Suppose

$$\mathcal{B}' = \{v_1', v_2', ..., v_m'\} = \{v_2, v_2 + cv_1, ..., v_m\}.$$

Then

$$M_{\mathcal{B}}^{\mathcal{B}'} I_V = \begin{pmatrix} 1 & 0 & 0 & ... & 0 \\ c & 1 & 0 & ... & 0 \\ 0 & 0 & 1 & ... & 0 \\ ... \\ 0 & 0 & 0 & ... & 1 \end{pmatrix} \tag{4.7}$$

$$M_{\mathcal{A}}^{\mathcal{B}'} = M_{\mathcal{B}}^{\mathcal{B}'} I_V \cdot M_{\mathcal{A}}^{\mathcal{B}} T = \begin{pmatrix} t_1 \\ t_2 + ct_1 \\ ... \\ t_m \end{pmatrix}$$

Matrices of type (4.2 - 4.7) are called elementary matrices.

Let $A = (a_{ij})_{m \times n}$. We say A is **upper triangular** if $a_{ij} = 0$ whenever $i > j$. We say A is **lower triangular** if $a_{ij} = 0$ whenever $i < j$. We say A is **diagonal** if A is both lower and upper triangular, i.e. $a_{i,j} = 0$ if $i \neq j$.

We say A is **upper echelon** if A is upper triangular and if whenever $a_{ii} \neq 0$, $a_{ii} = 1$ and $a_{ij} = 0$ for all $i < j$. We say A is **lower echelon** if A is lower triangular and if whenever $a_{ii} \neq 0$, $a_{ii} = 1$ and $a_{ij} = 0$ for all $i > j$.

Theorem 17. *Let U and V be finite dimensional vector spaces and $T : U \to V$.*

1. *Fix a basis $\mathcal{B}$ for V. Then there exists a basis $\mathcal{A}'$ for U such that $M_{\mathcal{A}'}^{\mathcal{B}} T$ is lower triangular.*

2. *Fix a basis $\mathcal{A}$ for U. Then there exists a basis $\mathcal{B}'$ for V such that $M_{\mathcal{A}}^{\mathcal{B}'} T$ is upper triangular.*

3. *There exist basis $\mathcal{A}'$ and $\mathcal{B}'$ for U and V respectively such that*

$$M_{\mathcal{A}'}^{\mathcal{B}'} T = \begin{pmatrix} D_k & 0 \\ 0 & 0 \end{pmatrix},$$

where D_k is a $k \times k$ diagonal matrix.

Proof. Let $\mathcal{A} = \{u_1, ..., u_n\}$ and $\mathcal{B} = \{v_1, ..., v_n\}$ be basis for U and V respectively. Fix $\mathcal{B}$ and let

$$M_{\mathcal{A}}^{\mathcal{B}} T = (t_{ij})_{m \times n} = \begin{pmatrix} t^1 & t^2 & ... & t^n \end{pmatrix}$$

1. We will induct on n: Suppose $n = 1$, then $M_{\mathcal{A}}^{\mathcal{B}} T = (t^1)_{m \times 1}$ is a column vector, which is clearly lower triangular. In general,

• If $t_{1j} = 0$ for all $j \geq 2$, then the $(m-1) \times (n-1)$ matrix

$$M_{\mathcal{A}_1}^{\mathcal{B}_1} T = \begin{pmatrix} t_{22} & t_{23} & \ldots & t_{2n} \\ \ldots & & & \\ t_{m2} & t_{m3} & \ldots & t_{mn} \end{pmatrix}_{(m-1)\times(n-1)}$$

where $\mathcal{A}_1 = \{u_2, ..., u_n\}$ and $\mathcal{B}_1 = \{v_2, ..., v_m\}$. By induction, there exists $\mathcal{A}_1' = \{u_2', ..., u_n'\}$ (a basis for the subspace *span*$(\{u_2, ..., u_n\})$) such that $M_{\mathcal{A}_1'}^{\mathcal{B}_1}$ is lower triangular. Let $\mathcal{A}' = \{u_1, u_2', ..., u_n'\}$ be a basis for U. Then

$$M_{\mathcal{A}'}^{\mathcal{B}} T = \begin{pmatrix} t_{11} & 0 \\ \bar{t} & M_{\mathcal{A}_1'}^{\mathcal{B}_1} \end{pmatrix},$$

which is clearly lower triangular. Here $\bar{t} = (t_{21}, ..., t_{m1})^t$.

• Otherwise there exists $j \in \{1, ..., n\}$ such that $t_{1j} \neq 0$. By interchanging the column vectors t^1 and t^j (i.e. updating $\mathcal{A}$ by interchanging u_1 and u_j) we may assume $t_{11} \neq 0$. Next, we'll make $t_{1j} = 0$ for any $j = 2, .., n$ as follow: Replacing the column vector t^j with $t^j - \frac{t_{2j}}{t_{1j}} t^1$ (i.e. updating $\mathcal{A}$ by replacing u_j with $u_j - \frac{t_{2j}}{t_{1j}} u_1$). With this new basis $\mathcal{A}$, $t_{1j} = 0$. We can do this for all j, and continuously updating $\mathcal{A}$. As a result, $t_{1j} = 0$ for all $j \geq 2$. Now we are back to the previous case, and by induction, there exists $\mathcal{A}'$ such that $M_{\mathcal{A}'}^{\mathcal{B}} T$ is lower triangular.

2. Using row vectors instead of column vectors, we obtain that there exists a basis $\mathcal{B}'$ for V such that $M_{\mathcal{A}}^{\mathcal{B}'} T$ is upper triangular.

3. Apply both methods (1) and (2), one obtains (3).

□

Lemma 3. *Let U and V be finite dimensional vector spaces and $T : U \to V$.*

1. *Fix a basis $\mathcal{B}$ for V. Then there exists a basis $\mathcal{A}'$ for U such that $M_{\mathcal{A}'}^{\mathcal{B}} T$ is lower echelon.*

2. *Fix a basis $\mathcal{A}$ for U. Then there exists a basis $\mathcal{B}'$ for V such that $M_{\mathcal{A}}^{\mathcal{B}'} T$ is upper echelon.*

3. *There exist basis $\mathcal{A}'$ and $\mathcal{B}'$ for U and V respectively such that*

$$M_{\mathcal{A}'}^{\mathcal{B}'} T = \begin{pmatrix} I_k & 0 \\ 0 & 0 \end{pmatrix},$$

where I_k is a $k \times k$ identity matrix.

Lemma 4. *Let $A \in M_{m \times n}$.*

1. *There exist elementary matrices $P_1, ..., P_k$ of types (4.2 - 4.4) such that*

$$A \cdot P_1 \cdot \ldots \cdot P_k$$

is lower echelon.

2. *There exist elementary matrices $Q_1, ..., Q_l$ of types (4.5 - 4.7) such that*

$$Q_1 \cdot ... \cdot Q_l \cdot A$$

is upper echelon.

3. *There exist elementary matrices $P_1, ..., P_k$ of types (4.2 - 4.4) and elementary matrices $Q_1, ..., Q_l$ of types (4.5 - 4.7) such that*

$$Q_1 \cdot ... \cdot Q_l \cdot A \cdot P_1 \cdot ... \cdot P_k = \begin{pmatrix} I_p & 0 \\ 0 & 0 \end{pmatrix},$$

where I_p is a $p \times p$ identity matrix.

4.5 Rank of a matrix and systems of linear equations

For each $A \in M_{m\times n}$, let $a_i \in F^n$, $i = 1, ..., m$, be the row vectors of A and $a^j \in F^m$, $j = 1, ..., n$, be the column vectors of A. We have

$$R = span(\{a_1, ..., a_m\}) \subset F^n, \text{ and } C = span(\{a^1, ..., a^n\}) \subset F^m.$$

$dim(R)$ is called the row rank of A and $dim(C)$ is called the column rank of A.

Proposition 20. *Let $A \in M_{m\times n}$, then $dim(R) = dim(C)$.*

Proof. Let $\{a'_1, ..., a'_M\}$ be the maximal linearly independent subset of $\{a_1, ..., a_m\}$ (row vectors). we have, $M \le m$ and

$$R = span(\{a'_1, ..., a'_M\}) = span(\{a_1, ..., a_m\}).$$

For each $i = 1, ..., m$,

$$a_i = p_{i1}a'_1 + ... + p_{iM}a'_M = \sum_{k=1}^{M} p_{ik}a'_k.$$

This shows that, for each $j = 1, ..., n$,

$$a_{ij} = \sum_{k=1}^{M} p_{ik}a'_{kj},$$

where $P = (p_{ik})_{m\times M}$. This implies

$$a^j = \sum_{k=1}^{M} a'_{kj}p^k.$$

Thus $C = span(\{a^1, ..., a^n\}) = span(\{p^1, ..., p^M\})$ which shows that

$$dim(C) \le M = dim(R).$$

Next, let $\{a^{1'}, ..., a^{N'}\}$ be the maximal linearly independent subset of $\{a^1, ..., a^n\}$ (column vectors). Following the same techniques, we see that

$$M = dim(R) \leq N = dim(C).$$

Combining the two inequalities, we have $dim(C) = dim(R)$. $\square$

Denote by $rank(A)$ the row (or column) rank of A.

Theorem 18. *Let T be a linear map from U to V with bases $\mathcal{A}$ and $\mathcal{B}$ respectively. Then*

$$rank(M_{\mathcal{A}}^{\mathcal{B}}T) = dim(Im(T)).$$

Proof. Let $\mathcal{A} = \{u_1, ..., u_n\}$ and $\mathcal{B} = \{v_1, ..., v_m\}$. Then $M_{\mathcal{A}}^{\mathcal{B}}T = (a_{ij})_{m\times n}$, where for each $j = 1, ..., n$,

$$T(u_j) = \sum_{i=1}^{n} a_{ij}v_i.$$

and $T(u_j)^{\mathcal{B}} = a^j$ (column vector). Thus,

$$dim(Im(T)) = dim(span(\{T(u_1)^{\mathcal{B}}, ..., T(u_n)^{\mathcal{B}}\})) = dim(span(\{a^1, ..., a^n\}).$$

$\square$

Proposition 21. *An $n \times n$ matrix has rank n if and only if it has an inverse.*

Proof. Let $T : F^n \to F^n$ be the linear map associated to A. I.e. $T(x) = Ax$, for all $x \in F^n$. Let $\mathcal{B} = \{e_1, ..., e_n\}$ be the canonical basis for F^n. Then

$$A = M_{\mathcal{B}}^{\mathcal{B}}T.$$

Suppose $rank(A) = n$. $dim(Im(T)) = n$, which shows that T is an isomorphism, and hence T has an inverse T^{-1}. But

$$M_{\mathcal{B}}^{\mathcal{B}}T^{-1} = (M_{\mathcal{B}}^{\mathcal{B}}T)^{-1}.$$

Define $A^{-1} = (M_{\mathcal{B}}^{\mathcal{B}}T)^{-1}$, then

$$AA^{-1} = (M_{\mathcal{B}}^{\mathcal{B}}T) \cdot (M_{\mathcal{B}}^{\mathcal{B}}T)^{-1} = M_{\mathcal{B}}^{\mathcal{B}}TT^{-1} = I$$

Similarly, $A^{-1}A = I$. Therefore A has an inverse.

Suppose A has an inverse. Define $S(x) = A^{-1}(x)$, then

$$I = AA^{-1} = M_{\mathcal{B}}^{\mathcal{B}}T \cdot M_{\mathcal{B}}^{\mathcal{B}}S = M_{\mathcal{B}}^{\mathcal{B}}TS.$$

Similarly, $I = M_{\mathcal{B}}^{\mathcal{B}}ST$. This shows that T is an isomorphism. Hence

$$rank(A) = dim(Im(T)) = n.$$

$\square$

Solutions of linear sytems of equations: We are interested in fnding a solution of the linear system of m equations and n unknowns:

$$\begin{aligned} a_{11}x_1 + \ldots + a_{1n}x_n &= y_1 \\ &\ldots \\ a_{m1}x_1 + \ldots + a_{mn}x_n &= y_m \end{aligned} \tag{4.8}$$

In matrix form, we have $Ax = y$, where

$$A = \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ & \ldots & \\ a_{m1} & \ldots & a_{mn} \end{pmatrix}, \quad x = \begin{pmatrix} x_1 \\ \ldots \\ x_n \end{pmatrix}, \quad y = \begin{pmatrix} y_1 \\ \ldots \\ y_m \end{pmatrix}$$

Let a^j be the j^{th} column vector of A, then we can write $Ax = y$ as

$$x_1 a^1 + \ldots + x_j a^j + \ldots + x_n a^n = y.$$

$y \in span(\{a^1, \ldots, a^j, \ldots, a^n\})$ if and only if there exists a solution $x = (x_1, \ldots, x_n)^t$ such that $Ax = y$.

Let $\{a^{j_1}, \ldots, a^{j_k}\}$ be the maximal linearly independent subset of $\{a^1, \ldots, a^n\}$. WLOG, WMA $\{a^{j_1}, \ldots, a^{j_k}\} = \{a^1, \ldots, a^k\}$, $k \leq n$. Then

$$span(\{a^1, \ldots, a^j, \ldots, a^n\}) = span(\{a^1, \ldots, a^k\}).$$

This implies $a_j \in span(\{a^1, \ldots, a^k\})$ for all $j = k+1, \ldots, n$. Note that $rank(A) = k$. Next if $y \in span(\{x_1, \ldots, x_k\}$, then for any $x_{k+1}, \ldots, x_n$,

$$y - x_{k+1}a^{k+1} - \ldots - x_n a^n \in span(\{a^1, \ldots, a^k\}).$$

This implies there exist unique $x_1, \ldots, x_k \in F$ such that

$$x_1 a^1 + \ldots + x_k a^k = y - x_{k+1}a^{k+1} - \ldots - x_n a^n.$$

Therefore, $x = (x_1, \ldots, x_k, x_{k+1}, \ldots, x_n)^t$ is a solution for $Ax = y$. Note that if $k = n$, i.e. $rank(A) = n$, then the solution x is unique. Otherwise $(k < n)$ $ker(A)$ contains more than the 0 element. Let $x' \in ker(A)$ s.t. $x' \neq 0$. Then

$$A(x - x') = Ax - Ax' = y - 0 = y,$$

which shows that $x - x'$ is also a solution. Hence, we have no uniqueness.

Theorem 19. *Let A be the matrix coresponding to (4.8) such that $rank(A) = k$. Given any $y \in F^m$. The equation $Ax = y$ has a solution $x = (x_1, \ldots, x_k, x_{k+1}, \ldots, x_n)$ if y belongs to the r-dimensional subspace generated by column vectors of A. If $k = n$ the the solution x is unique. Otherwise, the components $x_{k+1}, \ldots, x_n$ can be chosen arbitrarily.*

4.6 Computation of matrix inverses

Remark 15. Let $A \in M_{n \times n}$, and $P = P_1 \cdot ... \cdot P_k$ be the product of $n \times n$ elementary matrices such that $A \cdot P$ is lower echelon. Then $rank(A) = rank(AP)$. If $rank(AP) = r < n$, then AP has exactly r linearly independent column vectors, denoted by $\{t^{j_1}, ..., t^{j_r}\}$, such that $t_{j_k, j_k} = 1$.

Theorem 20. *Suppose $A \in M_{n \times n}$ is invertible (i.e. A^{-1} exists). Let $P = P_1 \cdot ... \cdot P_k$ be the product of elementary matrices such that $A \cdot P$ is lower echelon. Then $A^{-1} = P$.*