CHAPTER 6

GENERALIZED INVERSE OF MATRICES AND SEMIGROUPS

[*This Chapter is based on the lectures of Professor K.S.S. Nambooripad, Director-in-Charge, Centre for Mathematical Sciences, Trivandrum Campus, Kerala, India.*]

6.0. Introduction

Linear algebra is perhaps the most important branch of mathematics from both theoretical and practical point of view. In this context checking the existence of inverses (checking the non-singularity) of matrices and linear transformations and computing the same are crucial. I am sure that you are familiar with the simple criterion for verifying non-singularity of matrices and computing the inverses. Often a problem that can be solved by matrix (or linear transformation) inversion may be relevant even when the matrix under consideration is singular. For example, the solution of a matrix equation $uA = v$ where u and v are vectors and A is a matrix or linear transformation is clearly $u = vA^{-1}$ if A is non-singular. The problem is clearly relevant even if A is singular. Several mathematicians including Moore (1935), Penrose (1955) provided constructions that can replace inversion when the matrix involved is singular. Terms like 'generalized inverses' ('ginverses' by Rao and Mitra (1971)), 'relative inverses' by R.C. Burk (in Survey of Binary Systems) and Clifford (1941), 'inverses' by Preston (1954), etc. are used in this context. Our aim here is to describe some of these constructions and also suggest some computational procedures based on the underlying semigroups. You may refer to Rao and Mitra (1971) for applications of generalized inverses.

6.1. Preliminary Definitions

I assume that you are familiar with basic linear algebra. The book Artin (1990) is a good source for these ideas (especially the first 3 chapters). However, for convenience we shall briefly recall some of them.

6.1.1. Vector spaces

Let $\mathbb{k}$ be a field. A vector space over $\mathbb{k}$ consists of a set V, a binary operation

$$+ : V \times V \to V; \quad (u, v) \mapsto u + v$$

called the vector addition and an action of $\mathbb{k}$ on V

$$\mathbb{k} \times V \to V; \quad (\alpha, v) \mapsto \alpha v.$$

These must satisfy the following conditions:

(v1) $(u + v) + w = u + (v + w)$ for all $u, v, w \in V$;
(v2) $u + v = v + u$ for all $u, v \in V$;
(v3) there exists $0 \in V$ with $0 + v = v$ for all $v \in V$;
(v4) for each $v \in V$ there exists $-v \in V$ with $v + (-v) = 0$;
(v5) $1v = v$ for all $v \in V$ where 1 denote the identity of the field $\mathbb{k}$;
(v6) $(\alpha\beta)v = \alpha(\beta v)$ for all $\alpha, \beta \in \mathbb{k}$ and $v \in V$;
(v7) $\alpha(u + v) = \alpha u + \alpha v$ for all $\alpha \in \mathbb{k}$ and $u, v \in V$; and
(v8) $(\alpha + \beta)v = \alpha v + \beta V$ for all $\alpha, \beta \in \mathbb{k}$ and $v \in V$.

When V is a vector space over $\mathbb{k}$, the elements of V are called vectors and those of $\mathbb{k}$ are called scalars.

Observe that the first four axioms above says that $(V, +)$ is an abelian group. Thus a vector space over $\mathbb{k}$ can be defined as an abelian group V (with addition + which is omitted for brevity) on which $\mathbb{k}$ acts satisfying the last four axioms above. We now look at some examples:

Example 6.1.1. Let $\mathcal{P}(A)$ be the set of all subsets of a set A and for $A_1, A_2 \in \mathcal{P}$ let

$$A_1 \oplus A_2 = A_1 \cup A_2 \setminus A_1 \cap A_2$$

denote the symmetric difference of A_1 and A_2. Then $G = (\mathcal{P}(A), \oplus)$ is an abelian group. Let $\mathbb{k} = \mathbb{Z}_2$, where $\mathbb{Z}_2 = \{0, 1\}$ denote the prime field of characteristic 2. Then G is a vector space over $\mathbb{Z}_2$ with respect to the action defined by $0A' = 0$ and $1A' = A'$ for all $A' \in G$.

Example 6.1.2. Let $\mathbb{k}^n = \{(x_1, \ldots, x_n) : x_i \in \mathbb{k},\ i = 1, 2, \ldots, n\}$ where $\mathbb{k}$ denotes some field (the cases $\mathbb{k} = \mathbb{R}$ or $\mathbb{C}$ are the more important). Define addition in $\mathbb{k}^n$ by

$$(x_1, \ldots, x_n)+(y_1, \ldots, y_n) = (x_1+y_1+\cdots+x_n+y_n) \quad \text{for all} \quad (x_1, \ldots, x_n), (y_1, \ldots, y_n) \in \mathbb{k}^n$$

and the action of $\mathbb{k}$ on $\mathbb{k}^n$ by

$$k(x_1, \ldots, x_n) = (kx_1, \ldots, kx_n).$$

With these addition and scalar multiplication, $\mathbb{k}^n$ is a vector space over $\mathbb{k}$.

Example 6.1.3. Let $\mathsf{C}^{(n)}[a, b]$ denote the set of all complex valued functions on the interval $[a, b]$ that are n times continuously differentiable. If $f, g \in \mathsf{C}^{(n)}[a, b]$ and $\alpha \in \mathbb{C}$, then define the functions $f + g$ and αf by

$$(f + g)(t) = f(t) + g(t) \quad \text{and} \quad (\alpha f)(t) = \alpha(f(t))$$

for all $t \in [a, b]$. It can be seen that the functions $f + g$ and αf are continuously differentiable n times and it can be verified that $\mathsf{C}^{(n)}[a, b]$ is a vector space over $\mathbb{C}$.

A number of other examples can be constructed by modifying the data in the last two examples.

Let V be a vector space over $\mathbb{k}$. A subset U of V is a subspace of V if U is closed with respect to vector addition and scalar multiplication or equivalently, U satisfies the following condition:

$$\alpha u + \beta v \in U \quad \text{for all} \quad u, v \in U \quad \text{and} \quad \alpha, \beta \in \mathbb{k}.$$

It is clear that a vector space V has at least two subspaces; the set $\{0\}$ and V itself. All subspaces different from these two are called proper subspaces. If $v \neq 0$, $\mathcal{L}\langle v\rangle = \{\alpha v : \alpha \in \mathbb{k}\}$ is clearly a proper subspace of V. Consequently any non-zero vector space has a rich family of subspaces. Furthermore the set of all subspaces of V is a partially ordered set $\mathcal{P}(V)$ under inclusion and is called the projective geometry coordinatized by V.

Let $\mathcal{A} = \{U_\alpha : \alpha \in \Lambda\}$ denote an arbitrary family of subspaces of V. Using the definitions and elementary set-theoretic arguments one can verify that

$$\wedge\mathcal{A} = \bigcap_{\alpha\in\Lambda} U_\alpha \tag{6.1.1}$$

is the largest subspace of V contained in every member of $\mathcal{A}$. If $W = \wedge\mathcal{A}$ then the set $\mathcal{U}$ of all subspaces of V that contain every member of $\mathcal{A}$ is non-empty since $V \in \mathcal{U}$. Hence, by the above $\wedge\mathcal{U}$ is the smallest subspace that contains every member of $\mathcal{A}$. We use the notation $\sum \mathcal{A}$ to denote the smallest subspace of V containing U_α for every $\alpha \in \Lambda$; that is,

$$\wedge \mathcal{U} = \sum \mathcal{A} \tag{6.1.2}$$

The subspace $\sum \mathcal{A}$ is called the sum of the family $\mathcal{A}$. With the operations $\cap$ (intersection) and $\sum$ (sum) defined by (**??**) and (**??**), the partially ordered set $\mathcal{P}(V)$ becomes a complete lattice. For basic definitions of lattices and related results (see Skornyakov, 1964).

When $\Lambda = \{1, 2, \ldots, n\}$ is finite, we write

$$\sum \mathcal{A} = V_1 + V_2 + \cdots + V_n.$$

We can verify that the right-hand side represents the subspace of all vectors $v \in V$ that can be written as $v = v_1 + v_2 + \cdots + v_n$ where $v_i \in V_i$ for all i. Moreover, the subspace $W = V_1 + V_2 + \cdots + V_n$ is said to the direct sum of subspaces V_i, $i = 1, 2, \ldots, n$ if every vector v in W has a unique representation in the form $v = v_1 + v_1 + \cdots + v_n$ with $v_i \in V_i$. In this case we write

$$W = V_1 \oplus V_2 \oplus \cdots \oplus V_n.$$

The reader may verify that the following criterion is necessary and sufficient for the sum $W = \sum_{i=1}^{n} V_i$ of subspaces of V to be direct:

$$V_i \cap \left(\sum_{j \neq i} V_j \right) = \{0\} \tag{6.1.3}$$

for all $i = 1, 2, \ldots, n$. In particular, if $n = 2$, the sum $V_1 + V_2$ is direct if and only if $V_1 \cap V_2 = \{0\}$. We say that the subspace V_2 is a complement of the subspace V_1 or that V_1 and V_2 are complementary subspaces if

$$V_1 \oplus V_2 = V \qquad (\textbf{??}^*)$$

(see Example 6.1.7). Every $U \in \mathcal{P}(V)$ has a complement (see Corollary **??**) so that $\mathcal{P}(V)$ is called a complemented lattice.

Finally we observe that, if $U, V, W \in \mathcal{P}(V)$ and $U \supseteq W$ then we have:

$$U + (V \cap W) = (U + V) \cap W. \tag{6.1.4}$$

This equation is called the modular law satisfied by the geometry $\mathcal{P}(V)$. Thus $\mathcal{P}(V)$ is a complete, complemented modular lattice (see Skornyakov, 1964) for a more comprehensive discussion of complemented modular lattices.

6.1.2. Linear independence and bases

If $X \subseteq V$ is any subset of the vector space V, a vector $v \in V$ which can be expressed as a finite sum of the form

$$v = \sum_{i=1}^{r} \alpha_i v_i \quad \text{where} \quad \alpha_i \in \mathbb{k},\ v_i \in X,\ \text{for all } i = 1, \dots, r$$

is called a (finite) linear combination of vectors in X. Elements $\alpha_i \in \mathbb{k}$ are called the coefficients of the linear combination v. Here we assume that the sum is written in such a way that the vectors $v_1, \dots, v_r$ are distinct. Clearly we may add terms $0w_1 + 0w_2 + \dots 0w_k$ to the linear combination v without affecting the value of v. We denote by $\mathcal{L}\langle X\rangle_{\mathbb{k}} = \mathcal{L}\langle X\rangle$ the set of all finite linear combinations of vectors in X. X is said to be linearly independent if whenever a finite linear combination of vectors in X is zero then all its coefficients are zero; that is,

$$\sum_{i=1}^{r} \alpha_i v_i = 0 \quad \text{implies} \quad \alpha_i = 0 \quad \text{for all} \quad i = 1, 2, \dots, r.$$

Notice that any subset of V that contains 0 is linearly dependent. Moreover, it is useful to observe that the empty set (considered as a subset of V) is linearly independent. In the following we will be concerned with non-empty linearly independent sets.

In view of the importance of linear independence, it is useful to look at some equivalent formulations of the definition. Suppose that $v = \sum_i^r \alpha_i v_i$, $w = \sum_j^s \beta_j w_j$ are two linear combinations. By adding 0-terms as required, we can assume that v and w are linear combinations of the same set of vectors

$$\{u_1, \dots, u_m\} = \{v_1, \dots, v_r\} \cup \{w_1, \dots, w_s\}.$$

Then we may write

$$v = \sum_{i}^{m} \alpha'_i u_i \quad \text{where} \quad \begin{cases} \alpha'_i = \alpha_j & \text{if } u_i = v_j \text{ for some } 1 \leq j \leq r; \\ \alpha'_i = 0 & \text{otherwise.} \end{cases}$$

Similarly $$w = \sum_{i}^{m} \beta'_i u_i \quad \text{where} \quad \begin{cases} \beta'_i = \beta_{r+k} & \text{if } u_i = w_{r+k} \text{ for some } 1 \leq k \leq s; \\ \alpha'_i = 0 & \text{otherwise.} \end{cases}$$

It follows that

$$v + w = \sum_{k=1}^{m} (\alpha'_k + \beta'_k) u_k.$$

Hence $v + w$ is a linear combination as specified above. Again if $k \in \Bbbk$ and $v = \sum_i^m \alpha_i u_i$ then

$$kv = \sum_i^m (k\alpha_i) u_i$$

is a linear combination of vectors in X. Therefore, by the definition above, the set $\mathcal{L}\langle X\rangle$ of all linear combinations of vectors in X is a subspace of V. On the other hand, if U is any subspace of V containing X, then the definition of subspaces shows that any finite linear combination of vectors in X belongs to U. Thus $\mathcal{L}\langle X\rangle \subseteq U$. Consequently $\mathcal{L}\langle X\rangle$ is the intersection of all subspaces of V containing X. The subspace $\mathcal{L}\langle X\rangle$ is called the linear span of X or that $\mathcal{L}\langle X\rangle$ is the subspace generated by X. Notice that every vector in $\mathcal{L}\langle X\rangle$ can be written in the form $v = \sum_{i=1}^r \alpha_i v_i$ where $v_i \in X$. However, it is possible that the linear expression representing v may not be unique. Also, it is clear that $\mathcal{L}\langle U\rangle = U$ for any subspace U of V; in particular, $\mathcal{L}\langle \mathcal{L}\langle X\rangle\rangle = \mathcal{L}\langle X\rangle$ for all $X \subseteq V$. Notice that for any family $\mathcal{A} = \{U_\alpha : \alpha \in \Lambda\}$ of subspaces of V, we have

$$\sum \mathcal{A} = \mathcal{L}\langle \bigcup_{\alpha\in\Lambda} U_\alpha \rangle$$

The reader may enjoy proving the following equivalent characterizations of linear independence.

Theorem 6.1.1. *Let V be a vector space over the field $\Bbbk$. Then the following statements are equivalent for $X \subseteq V$:*

(a) *X is linearly independent.*

(b) *Every vector in $\mathcal{L}\langle X\rangle$ is a* unique *linear combination of vectors in X.*

(c) *No proper subset of X generates $\mathcal{L}\langle X\rangle$.*

A set $X \subseteq V$ is called a generator of V over $\Bbbk$ (or simply, a generator of V if the field $\Bbbk$ is clear from the context) if $\mathcal{L}\langle X\rangle = V$. Notice that V itself is a generator of V and so, every vector space has a generator. Moreover, X is called a basis of V over $\Bbbk$ (or simply, a basis of V) if X is a linearly independent set that generates V. By the theorem above, X is a basis if and only if

1) Every $v \in V$ can be written as $v = \sum_i \alpha_i v_i$ with $\alpha_i \in \Bbbk$ and $v_i \in X$; and
2) if $\sum_i^r \alpha_i v_i = \sum_j^s \beta_j w_j$ then $r = s$ and after reordering of terms, if necessary, we have $\alpha_i = \beta_i$ and $v_i = w_i$ for $i = 1, 2, \dots, r$.

As above, it will be convenient to have some equivalent characterizations of basis. We need some terminology about partially ordered sets to formulate these equivalent definitions of linear independence. The reader may refer to some book on algebra (such a Artin 1990) or topology (Nanda and Nanda 1990) for more extensive treatment of these topics. Recall that a relation $\leq$ on a set X (that is, $\leq \subseteq X \times X$) is a partial order if it satisfies the following: For arbitrary $x, y, z \in X$ we have

1) $x \leq x$;
2) $x \leq y \leq z \Longrightarrow x \leq z$; and
3) $x \leq y$ and $y \leq x$ implies $x = y$.

When $\leq$ satisfies these conditions we shall say that X is a partially ordered set with respect to $\leq$ or briefly that X is a partially ordered set since the relation $\leq$ will usually be clear from the context. X is said to be linearly ordered (or X is a chain) if, for all $x, y \in X$, either $x \leq y$ or $y \leq x$. If $Y \subseteq X$ then Y is a partially ordered set with respect to the relation $\leq \cap (Y \times Y)$ and this partial ordered set is called a partially ordered subset of X. Given $Y \subseteq X$, $m \in X$ is an upper [lower] bound of Y if $y \leq m$ [$m \leq y$] for all $y \in Y$; m is the least upper bound [greatest lower bound] if m is the smallest [largest] among all upper bounds [lower bounds]; it is denoted by $\sup Y$ [$\inf Y$] if these exist. $m \in X$ is maximal [minimal] if $x \in X$ and $m \leq x$ [$x \leq m$] implies $m = x$.

We use the terminology introduced above in the following statement.

Theorem 6.1.2. *The following statements are equivalent for a subset $X \subseteq V$.*

(a) *X is a basis of V.*
(b) *Every vector in V is a unique linear combination of vectors in X.*
(c) *X is maximal in the partially ordered set of all linearly independent subsets of V under inclusion.*
(d) *X is minimal in the partially ordered set of all generating subsets of V.*

For many practical as well as theoretic purposes it is necessary to know whether the given vector space V has a basis satisfying some additional restrictions and to find one such basis. For example, given a linearly independent set A

and a generating set X in V we would like to find a basis B such that

$$A \subseteq B \quad \text{and} \quad B \setminus A \subseteq X.$$

To prove the existence of such a basis we need the following set-theoretic principle (see Artin, 1990, page 588):

Lemma 6.1.1 (Zorn's Lemma). Let P be a partially ordered set with respect to the relation $\leq$. If every chain in P has an upper bound then P has maximal elements.

Suppose that A be a linearly independent set and X be a generating set in V. Consider the partially ordered set $\mathcal{P}$ with respect to inclusion of all linearly independent subsets C of V satisfying the conditions

$$A \subseteq C \quad \text{and} \quad C \setminus A \subseteq X. \tag{6.1.5}$$

Clearly $A \in \mathcal{C}$ and so, $\mathcal{C} \neq \emptyset$. If $\mathcal{C} = \{C_\alpha : \alpha \in \Lambda\}$ is any chain in $\mathcal{P}$, then

$$A \subseteq C = \bigcup_{\alpha \in \Lambda} C_\alpha \quad \text{and} \quad C \setminus A \subseteq X.$$

It is easy to see that C is linearly independent and so $C \in \mathcal{P}$ is an upper bound of $\mathcal{C}$ in $\mathcal{P}$. Hence by Zorn's lemma $\mathcal{P}$ contains maximal elements. Let $B \in \mathcal{P}$ be a maximal element in $\mathcal{P}$. Then B is linearly independent. We claim that $\mathcal{L}\langle B \rangle = V$. For, if $v \in X \setminus B$ and if $v \notin \mathcal{L}\langle B \rangle$, then the set $B' = B \cup \{v\}$ is a linearly independent set satisfying (6.1.5). Since B is a proper subset of B', this contradicts the maximality of B. Therefore $X \subseteq \mathcal{L}\langle B \rangle$ and so, $V = \mathcal{L}\langle X \rangle \subseteq \mathcal{L}\langle B \rangle$. Therefore we have the following.

Theorem 6.1.3. *Let A be a linearly independent subset of a vector space V. If X is a generating set of V there exists a subset C of X such that $A \cap C = \emptyset$ and $B = A \cup C$ is a basis of V.*

The above result has several important consequences. The first statement below follows by taking $X = V$ in the theorem above. The second statement is obtained by taking, in addition, $A = \emptyset$.

Corollary 6.1.1. Let V be a vector space over $\mathbb{k}$.

A. Every linearly independent set in V can be extended to a basis.

B. Every generating set $X \subseteq V$ contains a basis.

In particular, every vector space has a basis.

Let $U \subseteq V$ be any subspace. If A is a basis of U then A is a linearly independent subset of V. By Theorem **??** (or Corollary **??** A), there is $C \subseteq V$ such that $A \cap C = \emptyset$ and $B = A \cup C$ is a basis of V. Let $W = \mathcal{L}\langle C\rangle$. Then W is a subspace of V with basis C. Also $U \cap W = \{0\}$. Since $B \subseteq U \oplus W$, we have $U \oplus W = V$. Furthermore, if V is finite dimensional, so are U and W. Then $\dim U = |A|$ and $\dim W = |C|$. Therefore

$$\dim V = |B| = |A| + |C| = \dim U + \dim W.$$

Thus we have:

Corollary 6.1.2. Let U be a subspace of a vector space V over $\Bbbk$. Then U has a complement W in V. If V is finite dimensional so are U and W. Further $\dim U + \dim W = \dim V$.

The construction of the complement of a given subspace indicated in the paragraph preceding the corollary indicates that a non-degenerate subspace U has more than one complement in V (infinitely many complements if the characteristic of the field $\Bbbk$ is 0). We shall denote the set of all complements of U by $\complement[U]$.

Suppose that B and B' are two bases of a vector space V. If $u_0 \in B$, then by Theorem **??**(d), $C_0 = B \setminus \{u_0\}$ does not generate V and so $B' \setminus \mathcal{L}\langle C_0\rangle \neq \emptyset$; let $v_0 \in B' \setminus \mathcal{L}\langle C_0\rangle$. Then $B_0 = C_0 \cup \{v_0\}$ is a linearly independent set having the same number of vectors as B. Since B is a basis, there exist $u_i \in B$ and $a_i \in \Bbbk$, $i = 0, 1, \ldots, r$ with $v_0 = \sum_{i=0}^{r} a_i u_i$. Since $v_0 \notin \mathcal{L}\langle C_0\rangle$, $a_0 \neq 0$. Hence

$$u_0 = a_0^{-1} v_0 - \sum_{i\neq 0} (a_0^{-1} a_i) u_i \in L(B_0).$$

Therefore $B_0 \subseteq \mathcal{L}\langle B\rangle$ so that $\mathcal{L}\langle B_0\rangle = \mathcal{L}\langle B\rangle = V$. Therefore B_0 is a basis of V containing the same number of vectors as B. If $C_0 \neq \emptyset$, we may choose $u_1 \in C_0 \subseteq B_0$ and proceeding as above, get a new basis B_1 in which the vectors u_0 and u_1 has been replaced by two vectors $v_0, v_1 \in B'$; B_1 contains the same number of vectors as B. If B is finite, say, $B = \{u_0, \ldots, u_{n-1}\}$, then all bases $B_0, B_1, \ldots, B_i, \ldots$ contain the same number n of vectors. In B_i we have replaced $i + 1$ vectors from B with vectors from B'. Hence B_{n-1} will contain n vectors from B'. Since B_{n-1} is a basis contained in B', it follows from Theorem **??** that $B_{n-1} = B'$. Thus B' also contains n vectors.

Theorem 6.1.4. *If a vector space V has a finite basis containing n vectors, then every basis of V has exactly n vectors.*

We say that a vector space V is finite dimensional if it has a finite basis. If V has this property, Theorem **??** shows that the number $\dim V$ of vectors in a basis is an invariant for V and is called the dimension of V. If $B = \{u_1, u_2, \ldots, u_n\}$ is a basis of V, it is clear that every vector $v \in V$ can be uniquely written as

$$v = \sum_{i=1}^{n} a_i u_i \quad \text{where} \quad a_i \in \Bbbk, \quad i = 1, 2, \ldots, n.$$

Also any set of $n + 1$ vectors in V is linearly dependent.

6.1.3. Euclidean spaces

Let V be a real or complex vector space. An inner product on V is a function $\langle\ \rangle : V \times V \to \Bbbk$ such that

(I1) For all $v \in V$, $\langle v, v\rangle \geq 0$ and $\langle v, v\rangle = 0$ if and only if $v = 0$.

(I2) For all $u, v, w \in V$ and $a, b \in \Bbbk$, $\langle au + bv, w\rangle = a\langle u, w\rangle + b\langle v, w\rangle$.

(I3) For all $u, v \in V$, $\langle u, v\rangle = \overline{\langle v, u\rangle}$ where $\overline{z}$ denote the complex conjugate of $z \in \mathbb{C}$.

The pair $(V, \langle\ \rangle)$ is called an inner product space. We shall say for brevity that V is an inner product space if the inner product $\langle\ \rangle$ is clear from the context. Axiom (3) above is called the conjugate symmetry of the inner product. Notice that for a real inner product this reduces to symmetry: $\langle u, v\rangle = \langle v, u\rangle$ for all $u, v \in V$. A symmetric inner product is often referred to as a dot product. A real inner product is always a dot product; the converse is not true.

Here we will be concerned with finite dimensional inner product spaces. The function $\langle\ \rangle : \Bbbk^n \times \Bbbk^n \to \Bbbk$ defined by

$$\langle x, y\rangle = \sum_i x_i \overline{y_i} \quad \text{for all} \quad x, y \in \Bbbk^n \tag{6.1.6}$$

is an inner product on $\Bbbk^n$. A real inner product space of finite dimension is called a Euclidean space. In particular, $\mathsf{E}_n = (\mathbb{R}^n, \langle\ \rangle)$ is a Euclidean space of dimension n. A finite dimensional complex inner product space will be denoted by H (or H_n if necessary). H is called a unitary space (or a Hilbert space (see Limaye, 1981, Page 182–184)).

Recall that a norm on a vector space V over $\mathbb{k}$ is a function $\|\cdot\| : V \to \mathbb{R}$ satisfying the following conditions (see Limaye, 1981, Page 35). For all $x, y \in V$ and $k \in \mathbb{k}$, we have:

i. $\|x\| \geq 0$ and $\|x\| = 0$ if and only if $x = 0$;
ii. $\|x + y\| \leq \|x\| + \|y\|$; and
iii. $\|kx\| = |k|\|x\|$.

The following theorem shows that any inner product space is a normed space (that is, a vector space on which a norm is given). See (Apostol, 1985, Page 48) and/or (Limaye, 1981, Page 176-182) for proofs and further details.

Theorem 6.1.5. *Let V be an inner product space with respect to the inner product $\langle\ \rangle$ and let $\|\cdot\| : V \to \mathbb{R}$ denote the map defined by*

$$\|u\| = +\sqrt{\langle u, u\rangle} \tag{6.1.7}$$

for all $u \in V$. Then for all $u \in V$, $\|u\|$ is a non-negative real number. Furthermore, we have

(a) ***(Schwartz's inequality)*** *$|\langle x, y\rangle| \leq \|x\|\|y\|$ for every $x, y \in V$ where equality holds if and only if x and y are linearly dependent.*
(b) *The map $\|\cdot\| : V \to \mathbb{R}$ is a norm on V.*
(c) ***(Parallelogram law)*** *For all $x, y \in V$,*
$$\|x + y\|^2 + \|x - y\|^2 = 2(\|x\|^2 + \|y\|^2).$$
(d) ***(Polarization identity)*** *For all $x, y \in V$,*
$$4\langle x, y\rangle = \begin{cases} \|x + y\|^2 - \|x - y\|^2 + i\|x + iy\|^2 - i\|x - iy\|^2 & \text{if } \mathbb{k} = \mathbb{C}; \text{ and} \\ \|x + y\|^2 - \|x - y\|^2 & \text{if } \mathbb{k} = \mathbb{R}. \end{cases}$$

Conversely, if $\|\cdot\| : V \to \mathbb{R}$ is a norm on V, there exists an inner product $\langle\ \rangle$ satisfying (**??**) *if and only if $\|\cdot\|$ also satisfies parallelogram law.*

In the following discussion (in this section) the spaces considered will be Euclidean (and occasionally unitary). Two vectors $x, y \in \mathsf{E}_n$ are said to be orthogonal, written $x \perp y$ if $\langle x, y\rangle = 0$. Also, for any $X \subseteq \mathsf{E}_n$, we write

$$X^{\perp} = \{y \in \mathsf{E}_n : x \perp y \quad \text{for all} \quad x \in X\}. \tag{6.1.8}$$

It is easy to see that the following hold.

Proposition 6.1.1. For any $X \subseteq \mathsf{E}_n$, the set $X^{\perp}$ defined by Equation (**??**) is a subspace of E_n. Moreover, if X is a subspace of E_n, then $X^{\perp}$ is a complement of X and $X^{\perp\perp} = X$.

For any subset X of E_n, $X^{\perp}$ is called the orthogonal subspace of X (or simply the orthogonal of X). When X is a subspace $X^{\perp}$ is a complement of X (see Equation (**??**)*) and is called the orthogonal complement of X.

Remark 6.1.1. Consider $x, y \in \mathsf{E}_n$. By Schwartz inequality,

$$\left| \frac{||x||||y||}{\langle x, y \rangle} \right| \leq 1$$

and so, there is a real number θ such that

$$\cos\theta = \frac{||x||||y||}{\langle x, y \rangle}.$$

Notice that this equation defines θ in magnitude but is indeterminate in sign. However, lines containing vectors x and y are orthogonal if and only if vectors x and y are orthogonal in the sense defined above.

Let V be a finite dimensional inner product space. A subset $X \subseteq V$ is said to be orthogonal if $0 \notin X$ and satisfies the following:

$$\langle x, y \rangle = 0 \quad \text{for all} \quad x, y \in X \quad \text{with} \quad x \neq y.$$

X is said to be normal if

$$||x|| = 1 \quad \text{for all} \quad x \in X.$$

X is ortho-normal if it is both orthogonal and normal. Any orthogonal set is linearly independent and so an orthogonal set in V contains utmost n non-zero vectors. Any singleton set containing a normal (unit) vector is clearly orthonormal. A basis which is also an orthonormal set is called an orthonormal basis. We proceed to show that V has an orthonormal bases.

Let $\dim V = n$. Suppose that $C_r = \{w_1, \ldots, w_r\}$ is an orthonormal set in V. Since C_r is linearly independent, by Theorem **??** there is $D_r = \{u_{r+1}, \ldots, u_n\}$ such that $B_r = C_r \cup D_r$ is a basis of V. Now $u_{r+1} \notin \mathcal{L}\langle C_r \rangle$ and so,

$$w'_{r+1} = u_{r+1} - \sum_{j=1}^{r} \langle u_{r+1}, w_j \rangle \neq \emptyset.$$

Therefore if $w_{r+1} = \frac{w'_{r+1}}{\|w_{r+1}\|}$, then $C_{r+1} = \{w_1, \ldots, w_{r+1}\}$ is an orthonormal set and $B_{r+1} = C_{r+1} \cup D_{r+1}$ is a basis of V. Then, by induction, $B_n = \{w_1, \ldots, w_n\}$ is an orthonormal basis. The procedure used in the induction step of the above argument is called Gram-Schmidt's orthogonalization procedure.

If v is a unit vector then $\{v\}$ is clearly an orthonormal set in V. So, starting from any non-zero unit vector we can construct an orthonormal basis of V by executing the Gram-Schmidt procedure repeatedly.

Proposition 6.1.2. Let V be an inner product space with $\dim V = n$. Given any orthonormal set $X = \{u_i : 1 \leq i \leq r\}$ in V, there exists an orthonormal basis B of V containing X. Therefore, every finite dimensional inner product space has an orthonormal basis.

Example 6.1.4. Let X be a set and let $v : X \to \Bbbk$ be a mapping. The support of v is the set

$$\operatorname{Supp}(v) = \{x \in X : v(x) \neq 0\}.$$

Let $\Bbbk^X$ denote the set all mappings v which satisfy the condition that $\operatorname{Supp}(v)$ is finite. Define addition and scalar multiplication pointwise: For $u, v \in \Bbbk^X$ and $k \in \Bbbk$ let

$$(u + v)(x) = u(x) + v(x) \quad \text{and} \quad (kv)(x) = k(v(x))$$

for all $x \in X$. Show that $\Bbbk^X$ is a vector space over $\Bbbk$. For each $x \in X$, let $e_x : X \to \Bbbk$ be the function such that

$$e_x(y) = \begin{cases} 1 & \text{if } y = x; \text{ and} \\ 0 & \text{if } y \neq x. \end{cases}$$

Prove that $E_X = \{e_x : x \in X\}$ is a basis of $\Bbbk^X$ over $\Bbbk$ and that the map $x \mapsto e_x$ is a bijection of X with $\mathcal{E}_X$. $\mathcal{E}_X$ is called the natural basis of $\Bbbk^X$. In particular, if $X = \{1, \ldots, n\}$ any n-tuple can be interpreted as a function on X to $\Bbbk$. Since X is finite, any function on X to $\Bbbk$ has finite support. Therefore the vector space $\Bbbk^n$ is a particular case of $\Bbbk^X$. The natural basis of $\Bbbk^n$ is denoted by $\mathcal{E}_n = \{e_1, e_2, \ldots, e_n\}$.

Example 6.1.5. Let $\mathbb{N} = \{0, 1, 2, \ldots\}$ denote the set of natural numbers. Consider the set

$$\Bbbk[x] = \left\{ \sum_{i=0}^{n} a_i x^i : n \in \mathbb{N}, \quad a_i \in \Bbbk \right\}$$

of all polynomials in one variable x with coefficients in $\Bbbk$ where $x^0 = 1$. If $p(x), q(x) \in \Bbbk[x]$, define

$$p(x) + q(x) = \sum_i (a_i + b_i) x^i$$

where a_i denotes the coefficient of x^i in $p(x)$ if x^i occurs in $p(x)$ and zero otherwise. Similarly b_i denotes the coefficient of x^i in $q(x)$ if the power x^i occurs in $q(x)$ and 0 otherwise. Further if $k \in \mathbb{k}$ then let

$$kp(x) = k\left(\sum_i a_i x^i\right) = \sum_i (ka_i)x^i.$$

Show that $\mathbb{k}[x]$ is a vector space over $\mathbb{k}$ and $B = \{x^n : n \in \mathbb{N}\}$ is a basis of $\mathbb{k}[x]$. Consequently $\mathbb{k}[x]$ is not finite dimensional.

Example 6.1.6. Let $\mathbb{k}^n$ be the vector space of Example 6.1.2. Let

$$\mathcal{E}_n = \{e_i \in \mathbb{k}^n : i = 1, 2, \dot{;} n\}$$

where $$e_i = (\delta_{1i}, \delta_{2i} \ldots, \delta_{ni}) \quad \text{with} \quad \delta_{ij} = \begin{cases} 0 & \text{if } i \neq j; \text{ and} \\ 1 & \text{if } i = j. \end{cases}$$

Show that $\mathcal{E}_n$ is a basis of $\mathbb{k}^n$ over $\mathbb{k}$. Consequently $\mathbb{k}^n$ is a finite dimensional vector space with $\dim \mathbb{k}^n = n$.

Example 6.1.7. Observe that the Euclidean plane can be identified with the real vector space $\mathbb{R}^2$ (see Example 6.1.6). Any non-zero proper subspace of $\mathbb{R}^2$ is a line through the origin and any two distinct lines through the origin are direct summands. Similarly, $\mathbb{R}^3$ can be identified with three dimensional Euclidean space and non-zero proper subspaces are lines and planes through the origin. If P is a plane and l is a line through the origin not lying on P, then we have $P \oplus l = \mathbb{R}^3$. Moreover if l_1, l_2, l_3 are distinct, none-coplanar lines through the origin, then

$$l_1 \oplus l_2 \oplus l_3 = \mathbb{R}^3$$

Hence the triad of lines (l_1, l_2, l_3) represent an oblique coordinate system for $\mathbb{R}^3$.

Example 6.1.8. Prove the following:

(a) $\mathcal{E}_n$ (see Example 6.1.6) is an orthonormal basis with respect to the inner product defined by Equation (**??**).

(b) Let $\mathbb{C}[x]$ be the vector space of all complex polynomials in one variable. For $p(x) = \sum_i a_i x^i, q(x) = \sum_j b_j x^j \in \mathbb{C}[x]$, define

$$\langle p(x), q(x)\rangle = \sum_i a_i \overline{b_i}.$$

Prove that this is an inner product on $\mathbb{C}[x]$ and that the set $\{x^n : n \in \mathbb{N}\}$ is an orthonormal basis of $\mathbb{C}[x]$.

(c) Let X be a set and $V = \mathbb{C}^X$ be the vector space of all complex valued functions on X with finite support (see Example 6.1.4). For $f, g \in V$ let

$$\langle f, g\rangle = \sum_{x\in X} f(x)\overline{g(x)}.$$

Show that this defines an inner product on V and that the set E_X of Example 6.1.4 is an orthonormal basis of V.

Example 6.1.9. Show that an orthonormal basis of E_2 has the form

$$\{(\cos\theta, \sin\theta), (-\sin\theta, \cos\theta)\} \quad \text{or} \quad \{(\cos\theta, -\sin\theta), (\sin\theta, \cos\theta)\}.$$

Can you get a similar description of orthonormal basis in E_3.

Example 6.1.10. Let $u \in \mathsf{E}_3$ be a unit vector. Show that

$$u = (\cos\theta\cos\phi, \cos\theta\sin\phi, \sin\theta), \quad \text{for some} \quad \theta \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right], \quad \phi \in [0, 2\pi].$$

Prove further that

$$\{(\cos\theta\cos\phi, \cos\theta\sin\phi, \sin\theta), (-\sin\theta\cos\phi, -\sin\theta\sin\phi, \cos\theta), (-\sin\phi, \cos\phi, 0)\}$$

is an orthonormal basis of E_3.

6.2. Linear Transformations

Every type of mathematical common object such as sets, groups, vector spaces, topological spaces,etc has associated with it a class mappings that enable us to compare two object of the same class. For example homomorphisms of groups do this work for groups, continuous mappings for topological spaces, etc. The maps that help to compare vector spaces over the same field are called linear transformations.

6.2.1. Linear transformations and matrices

Let V and W be vector spaces over the field $\Bbbk$. A map $f : V \to W$ is called a linear transformation if

(Lt1) $f(u + v) = f(u) + f(v)$ for all $u, v \in V$; and

(Lt2) $f(\alpha v) = \alpha f(v)$ for all $v \in V$, $\alpha \in \Bbbk$.

Observe that the first statement above shows that a linear transformation $f : V \to W$ preserves vector addition so that f is an additive homomorphism of $(V, +)$ to $(W, +)$. The second shows that f is compatible with the scalar multiplication. We can combine the two statements as follows. $f : V \to W$ is a linear transformation if and only if

$$f(au + bv) = af(u) + bf(v) \quad \text{for all} \quad a, b \in \Bbbk, \quad u, v \in V.$$

This statement says that f is a linear transformation if and only if it preserves linear combinations of two-element sets of vectors in V. This can be obviously extended to any arbitrary finite set of vectors; that is f is a linear transformation if and only if f satisfies the following condition for all $n \in \mathbb{N}$:

$$f\left(\sum_{i=1}^{n} a_i v_i\right) = \sum_{i=1}^{n} a_i f(v_i) \quad \text{for all} \quad a_i \in \Bbbk, \quad v_i \in V, \quad i = 1, \ldots, n.$$

A linear isomorphism or simply isomorphism of vector spaces is a bijective linear transformation. Notice that, if f is a linear isomorphism so is f^{-1}. Two isomorphic vector spaces are algebraically indistinguishable even though the representation of vectors in the vector spaces may be different (see Example 6.4.1 below).

We have seen that every vector space V is the linear span of a basis B of V. Similarly, we can see that every linear transformation $f : V \to W$ is the linear extension of a map $\sigma : B \to W$ where B is a basis of V. Here, by linear extension of σ we mean the map $f_\sigma : V \to W$ defined as follows:

$$f_\sigma(v) = \sum_{u \in B} a_u \sigma(u) \quad \text{for all} \quad v = \sum_{u \in B} a_u u \in V \tag{6.2.1}$$

Since every $v \in V$ is a unique finite linear combination of vectors in B, the map f_σ is well-defined. The extension is clearly unique and linear. If $f : V \to W$ is any linear transformation then f is the linear extension of $\sigma = f|B$ and the uniqueness above implies that $f = f_\sigma$.

It is possible to discuss various properties of f_σ in terms of σ. Now f_σ is not one to one if and only if there is a non-zero vector $v \in V$ with $f_\sigma(v) = 0$. If $v = \sum_B a_u u$, since $v \neq 0$, $a_u \neq 0$ for some $u \in B$. Hence $f_\sigma(v) = \sum_B a_u \sigma(u) = 0$ is a non-zero linear combination of vectors in $\sigma(B)$ and so, $\sigma(B)$ is not linearly independent. Suppose that f_σ is one-to-one. Let $\sum_B a_u \sigma(u) = 0$ be a finite linear combination of vectors in $\sigma(B)$ which is 0. If $v = \sum_B a_u u$ then by the definition of f_σ above, $f_\sigma(v) = 0$. Since f_σ is one-to-one, $v = 0$. Since B is a basis of V, it follows that $a_u = 0$ for all $u \in B$ which says that $\sigma(B)$ is linearly independent in W.

Next, let f_σ be surjective and let $w \in W$. Then there exist $v = \sum_B a_u u \in V$ with

$$w = f_\sigma(v) = \sum_B a_u f_\sigma(u)$$

which implies that $w \in \mathcal{L}\langle\sigma(B)\rangle$. Therefore $\sigma(B)$ is a generating set for W. On the other hand, let $\sigma(B)$ be generating. If $w \in W$, there is $a_u \in \Bbbk$ with $w = \sum_B a_u \sigma(u)$. If $v = \sum_B a_u u$ then $f_\sigma(v) = w$ and so, f_σ is surjective. Combining the two cases considered above, it follows that $f_\sigma : V \to W$ is a linear isomorphism if and only if $\sigma(B)$ is a basis of W.

Theorem 6.2.1. *Let V and W be vector spaces over the field $\Bbbk$ and B be a basis of V. Suppose that $\sigma : B \to W$ be a mapping. Let $f_\sigma : V \to W$ be the map defined by* (**??**)*. Then $f_\sigma : V \to W$ is the unique linear transformation which extends σ on B.*

(A) *f_σ is injective if and only if $\sigma(B)$ is linearly independent in W.*

(B) *f_σ is surjective if and only if $\sigma(B)$ is a generating set for W.*

Consequently f_σ is a linear isomorphism if and only if $\sigma(B)$ is a basis of W.

Conversely if $f : V \to W$ any linear transformation, then for any basis B of V, we have $f = f_\sigma$ where $\sigma = f|B$.

If V and W are finite dimensional and if $\dim V = \dim W = n$, then both V and W have bases B and C containing n vectors and so, there is a bijection of B onto C which extends to an isomorphism. Conversely let $\dim V = n$ and $f : V \to W$ be a linear isomorphism. Then for any basis B of V,

$$\dim V = |B| = |C| = \dim W$$

where $C = f(B)$ is a basis of W. Thus

Corollary 6.2.1. Two finite dimensional vector spaces V and W are isomorphic if and only if $\dim V = \dim W$.

The following result lists some very important properties of arbitrary linear transformations.

Proposition 6.2.1. Let $f : V \to W$ be a linear transformation. Then

$$\mathsf{N}(f) = \{u \in V : f(u) = 0\} \tag{6.2.2}$$
$$\text{and} \quad \mathsf{R}^{\mathrm{r}}(f) = \{w \in W : f(v) = w \quad \text{for some} \quad v \in V\}.$$

are subspaces of V and W respectively. If U is a complement of $\mathsf{N}(f)$ in V, then $f|U$ is a linear isomorphism of U onto $\mathsf{R}^{\mathrm{r}}(f)$. In particular, if V is finite dimensional, so are $\mathsf{N}(f)$ and $\mathsf{R}^{\mathrm{r}}(f)$ and

$$\dim V = \dim \mathsf{N}(f) + \dim \mathsf{R}^{\mathrm{r}}(f).$$

Proof 6.2.1. It is easy to verify that $\mathsf{N}(f)$ and $\mathsf{R}^{\mathrm{r}}(f)$ are subspaces of V and W respectively. Let U be a complement of $\mathsf{N}(f)$ and let $f' = f|U$. Clearly $f' : U \to \mathsf{R}^{\mathrm{r}}(f)$ is linear. Also, if $u \in U$ and $f'(u) = f(u) = 0$, then $u \in \mathsf{N}(f) \cap U$. Hence $u = 0$ which says that f' is injective. Also, for any $w \in \mathsf{R}^{\mathrm{r}}(f)$, there exists $v \in V$ with $f(v) = w$. Since $\mathsf{N}(f) \oplus U = V$, there exists unique $v_0 \in \mathsf{N}(f)$ and $v_1 \in U$ such that $v = v_0 + v_1$. Then $f(v) = f(v_0) + f(v_1) = f(v_1) = f'(v_1) = w$ which shows that f' is surjective. Therefore $f' : U \to \mathsf{R}^{\mathrm{r}}(f)$ is a linear isomorphism. If V is finite dimensional, so are $\mathsf{N}(f)$ and U. Since $\mathsf{R}^{\mathrm{r}}(f)$ is isomorphic to U, $\mathsf{R}^{\mathrm{r}}(f)$ is finite dimensional and by Corollary **??** $\dim \mathsf{R}^{\mathrm{r}}(f) = \dim U$. Hence the desired equality follows from Corollary **??**.

The integer $\dim \mathsf{R}^{\mathrm{r}}(f)$ is called the rank of f and is denoted as Rank f. Similarly, the integer $\dim \mathsf{N}(f)$ is called the nullity of f and is denoted by Nullity f. Then, by the proposition above, we have

$$\text{Rank } f + \text{Nullity } f = \dim V = n$$

for all linear transformations with $\dim f = V$.

6.2.2. The space of linear transformations and matrices

Let $\mathsf{H}_{\Bbbk}(V, W)$ denote the set of all linear transformations of an arbitrary vector space V over $\Bbbk$ to a vector space W. Since $\Bbbk$ will be clear from the context, we shall abbreviate the notation to $\mathsf{H}(V, W)$. It is useful to note that the set $\mathsf{H}(V, W)$ itself carries a vector space structure.

Proposition 6.2.2. Let V and W be vector spaces over the field $\Bbbk$. Then $\mathsf{H}(V, W)$ is a vector space with respect to addition and scalar multiplication defined as follows:

$$(f+g)(v) = f(v) + g(v); \quad \text{and} \quad (kf)(v) = k(f(v)) \tag{6.2.3}$$

for all $f, g \in \mathsf{H}(V, W)$, $k \in \Bbbk$ and $v \in V$.

In the following, unless otherwise specified, V, W, etc. will denote finite dimensional vector spaces over $\Bbbk$.

Suppose that $f : V \to W$ is a linear transformation of a vector space V to a vector space W. Let $B = \{u_i : 1 \le i \le n\}$ and $C = \{w_j : 1 \le j \le m\}$ be bases of V and W respectively so that $\dim V = n$ and $\dim W = m$. There are bijections (see Example 6.4.2)

$$\mathfrak{c}_B : B \to \mathcal{E}_n; \quad u_i \mapsto e_i \quad \text{and}$$
$$\mathfrak{c}_C : C \to \mathcal{E}_n; \quad w_j \mapsto e_j.$$

By Theorem **??**, the linear extensions of these, again denoted by $\mathfrak{c}_B$ and $\mathfrak{c}_C$ are linear isomorphisms of V onto $\Bbbk^n$ and W onto $\Bbbk^m$ respectively (see Example 6.4.2). Hence $\tilde{f} = \mathfrak{c}_B^{-1} \circ f \circ \mathfrak{c}_C$ is a linear transformation of $\Bbbk^n$ to $\Bbbk^m$. Let

$$f(u_i) = \sum_{j=1}^{m} a_{ij} w_j; \tag{6.2.4}$$

for all $i = 1, 2, \ldots, n$. The map

$$\mathfrak{m}_{B,C}(f) : e_i \to (a_{i1}, \ldots, a_{im}) : \mathcal{E}_n \to \Bbbk^m$$

is called the matrix of f in the bases B and C. We use the usual matrix notation to indicate this map:

$$\mathfrak{m}_{B,C}(f) = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{im} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \cdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{pmatrix}$$

Clearly $\mathfrak{m}_{B,C}(f)$ is an $n \times m$-matrix. Conversely if $A = (a_{ij})$ is any $n \times m$-matrix, it determines a map

$$A : e_i \mapsto (a_{i1}, a_{i2}, \cdots, a_{im})$$

of $\mathcal{E}_n$ to $\Bbbk^m$ which sends e_i to the $i-th$ row of A. The linear extension f_A of the map A is a linear transformation of $\Bbbk^n$ to $\Bbbk^m$ and

$$f = \mathfrak{c}_B \circ f_A \circ \mathfrak{c}_C^{-1} : V \to W$$

is a linear transformation such that $\mathfrak{m}_{B,C}(f) = A$. Thus the correspondence $f \mapsto \mathfrak{m}_{B,C}(f)$ is a bijection of the set of $\mathsf{H}(V, W)$ of all linear transformations of V to W with the set $\mathsf{M}at_{n\times m}$ of all $n \times m$ matrices over $\Bbbk$.

Let $\sigma_{ij} : B \to W$ be the map defined as follows:

$$\sigma_{ij}(v_k) = \begin{cases} w_j & \text{if } k = i \text{ and} \\ 0 & \text{otherwise.} \end{cases}$$

By Theorem **??** there is a unique linear transformation, $f_{ij} = f_{\sigma_{ij}} : V \to W$, the linear extension of σ_{ij}. Let

$$F = F_{B,C} = \{f_{ij} : 1 \le i \le n, 1 \le j \le m\}.$$

F is linearly independent since $\sum_{ij} d_{ij} f_{ij} = 0$ implies

$$\left(\sum_{ij} d_{ij} f_{ij}\right)(v_k) = 0 \quad \text{for all} \quad i = 1, \ldots, n.$$

Hence

$$\sum_{j} d_{kj} f_{kj}(v_k) = 0 \quad \text{for all} \quad i = 1, \ldots, n.$$

Since C is linearly independent, this implies that $d_{kj} = 0$ for all j and k. So F is linearly independent. Now let $g \in \mathsf{H}(V, W)$ with $\mathfrak{m}_{B,C}(g) = (a_{ij})$. Then by (**??**),

$$g(v_i) = \sum_j a_{ij} w_j \quad \text{for all} \quad i = 1, \ldots, n.$$

Suppose that $g' = \sum_{i,j} a_{ij} f_{ij}$. By (**??**) $g' \in \mathsf{H}(V, W)$ and for any $k = 1, \ldots, n$,

$$\begin{aligned} g'(v_k) &= \sum_{i,j} a_{ij} f_{ij}(v_k) \\ &= \sum_j a_{kj} w_j = g(v_k). \end{aligned}$$

Therefore, by the uniqueness in Theorem **??**, we conclude that $g = g'$ and so, F is a basis of $\mathsf{H}(V, W)$. It follows that $\dim \mathsf{H}(V, W) = mn$.

Again $\mathsf{M}at_{n\times m}$ is a vector space over $\Bbbk$ with addition and scalar multiplication defined by

$$\begin{aligned} (a_{ij}) + (b_{ij}) &= (a_{ij} + b_{ij}) \quad \text{for all} \quad (a_{ij}), (b_{ij}) \in \mathsf{M}at_{n\times m}; \\ k(a_{ij}) &= (ka_{ij}) \quad \text{for all} \quad (a_{ij}) \in \mathsf{M}at_{n\times m}, \quad k \in \Bbbk. \end{aligned} \tag{6.2.5}$$

Let $f, g \in \mathsf{H}(V, W)$ and $a, b \in \Bbbk$. Suppose that $\mathfrak{m}_{B,C}(f) = (a_{ij})$ and $\mathfrak{m}_{B,C}(g) = (b_{ij})$. Then for each $v_i \in B$

$$\begin{aligned}(af + bg)(v_i) &= a(f(v_i)) + b(g(v_i)) \\ &= a\left(\sum_j a_{ij}w_j\right) + b\left(\sum_j b_{ij}w_j\right) \\ &= \sum_j (aa_{ij} + bb_{ij})w_j.\end{aligned}$$

It follows from (**??**) that

$$\begin{aligned}\mathfrak{m}_{B,C}(af + bg) &= (aa_{ij} + bb_{ij}) = a(a_{ij}) + b(b_{ij}) \\ &= a\mathfrak{m}_{B,C}(f) + b\mathfrak{m}_{B,C}(g).\end{aligned}$$

Therefore the map $\mathfrak{m}_{B,C} : f \mapsto \mathfrak{m}_{B,C}(f)$ is linear. We have already noted that $\mathfrak{m}_{B,C} : \mathsf{H}(V, W) \to \mathsf{M}at_{n\times m}$ is a bijection and so, $\mathfrak{m}_{B,C}$ is a linear isomorphism.

Let U, V and W be vector spaces with bases A, B and C respectively. If $f \in \mathsf{H}(U, V)$ and $g \in \mathsf{H}(V, W)$ then by (**??**), equations defining matrices $\mathfrak{m}_{A,B}(f) = (a_{ij})$ and $\mathfrak{m}_{B,C}(g) = (b_{jk})$ are

$$f(u_i) = \sum_j a_{ij}v_j \quad \text{and} \quad g(v_j) = \sum_k b_{jk}w_k.$$

If $h = f \circ g$, equations defining $\mathfrak{m}_{A,C}(h)$ are

$$\begin{aligned}h(u_i) &= \sum_i a_{ij}g(v_j) = \sum_j \sum_k a_{ij}b_{jk}w_k \\ &= \sum_k c_{ik}w_k \quad \text{where} \quad c_{ik} = \sum_j a_{ij}b_{jk}.\end{aligned}$$

Hence $\mathfrak{m}_{A,C}(h) = (c_{ik})$. Notice that (c_{ik}) is the usual row-column product $(a_{ij})(b_{jk})$. Therefore we have

$$\mathfrak{m}_{A,C}(f \circ g) = \mathfrak{m}_{A,B}(f) \times \mathfrak{m}_{B,C}(g). \tag{6.2.6}$$

For the convenience of later reference, we summarize the discussion. Recall that the set $\mathsf{H}(V, W)$ of all linear transformations of a vector space V to W is a vector space (see Proposition **??**).

Theorem 6.2.2. *Let V and W be vector spaces over $\Bbbk$ with $\dim V = n$ and $\dim W = m$. Then $\mathsf{H}(V, W)$ is a vector space over $\Bbbk$ satisfying the following: [(1)] $\dim \mathsf{H}(V, W) = nm$.*
*$\mathsf{M}at_{n\times m}$ is a vector space over $\Bbbk$ with addition and scalar multiplication defined by (**??**).*
If B and C are bases of V and W respectively, then

$$\mathfrak{m}_{B,C} : \mathsf{H}(V, W) \to \mathsf{M}at_{n\times m}; \quad f \mapsto \mathfrak{m}_{B,C}(f)$$

is a linear isomorphism of $\mathsf{H}(V, W)$ onto $\mathsf{M}at_{n\times m}$. In particular $\dim \mathsf{M}at_{n\times m} = mn$.

Moreover, if A, B and C are bases for vector spaces U, V and W respectively then

$$\mathfrak{m}_{A,C}(f \circ g) = \mathfrak{m}_{A,B}(f) \times \mathfrak{m}_{B,C}(g)$$

for all $f \in \mathsf{H}(U, V)$ and $g \in \mathsf{H}(V, W)$.

6.2.3. The algebra of linear transformations and matrices

If V is a vector space with $\dim V = n$ over $\Bbbk$, it follows by setting $V = W$ in Theorem **??** that $\mathsf{H}(V, V)$ is a vector space of dimension n^2 over $\Bbbk$. Moreover, for $f, g \in \mathsf{H}(V, V)$ the composite $f \circ g$ (often abbreviated as fg) is again a linear transformation in $\mathsf{H}(V, V)$. This gives a multiplication in $\mathsf{H}(V, V)$ which is associative and distributive. It is easy to see that $\mathsf{H}(V, V)$ is a ring with respect to vector addition (defined by Equation (**??**)) and composition. We denote this ring by $\mathfrak{A}(V)$. Moreover, $\mathfrak{A}(V)$ is also a vector space over $\Bbbk$ and satisfies the following:

$$k(fg) = (kf)g = f(kg) \quad \text{for all} \quad k \in \Bbbk, \quad f, g \in \mathfrak{A}(V). \tag{6.2.7}$$

Rings that are vector spaces over the field $\Bbbk$ that satisfies the property above are called $\Bbbk$-algebras (or simply algebras if $\Bbbk$ is clear from the context). Thus $\mathfrak{A}(V)$ is a finite dimensional $\Bbbk$-algebra of dimension n^2. Similarly $\mathsf{M}at_{n\times n}$ is an algebra of dimension n^2 over $\Bbbk$ with respect to matrix addition and scalar multiplication defined by Equation (**??**) and usual matrix multiplication which is isomorphic to $\mathfrak{A}(\Bbbk^n)$. We shall identify $\mathsf{M}at_{n\times n}$ with $\mathfrak{A}(\Bbbk^n) = \mathsf{M}_n$.

Let B be a basis of V over $\Bbbk$. If $f \in \mathfrak{A}(V)$ then f is a linear transformation from V to itself and so, its matrix can be defined as in Equation (**??**) where we may choose $B = C$. We denote this matrix by $\mathfrak{m}_B(f)$ so that $\mathfrak{m}_B(f) = \mathfrak{m}_{B,B}(f)$.

For $f \in \mathfrak{A}_n$, we write $\mathfrak{m}(f)$ for the matrix $\mathfrak{m}_E(f)$ where $\mathcal{E}_n$ denotes the natural basis of $\mathbb{k}^n$ (see Example 6.1.6). Again by Theorem **??**, the map $\mathfrak{m}_B : \mathfrak{A}(V) \to \mathsf{M}_n$ is a linear isomorphism and by Equation (**??**), the map $\mathfrak{m}_B$ also preserves multiplication. A map $h : \mathfrak{A} \to \mathfrak{A}'$ of an algebra $\mathfrak{A}$ to $\mathfrak{A}'$ is an algebra isomorphism if it is linear isomorphism and ring isomorphism. Therefore $\mathfrak{m}_B : \mathfrak{A}(V) \to \mathsf{M}_n$ is an isomorphism of algebras.

Theorem 6.2.3. *Let V be a vector space with $\dim V = n$. Then $\mathfrak{A}(V)$ is an algebra of dimension n^2 over $\mathbb{k}$ with respect to vector addition and scalar multiplication defined by Equation* (**??**) *and composition. Similarly the set M_n of all $n \times n$ matrices over $\mathbb{k}$ is an algebra of dimension n^2 over $\mathbb{k}$ with respect to matrix addition and scalar multiplication defined by Equation* (**??**) *and usual matrix multiplication. Moreover, the map $\mathfrak{m}_B : \mathfrak{A}(V) \to \mathsf{M}_n$ is an isomorphism of algebras.*

In the following discussion, $\mathfrak{A} = \mathfrak{A}(V)$ will denote the algebra of linear transformation of a vector space V with $\dim V = n$ over $\mathbb{k}$. The theorem above says that we can identify the algebra $\mathfrak{A}(V)$ with the matrix algebra M_n by the isomorphism $\mathfrak{m}_B$. Clearly the identification depends on the given basis B. In any case, given any statement regarding linear operators in $\mathfrak{A}$ we can routinely translate the statement to an appropriate statement regarding matrices in M_n (that is, statement about square $n \times n$-matrices). The operator version is geometric as it involves affine and linear subspaces of V and certain geometric transformations whereas the matrix version is computational. Often we shall state the geometric version and leave its translation to matrices as an exercise to the reader.

6.3. Ginverses

Let U_0 be a subspace of V. Then by Corollary **??** U_0 has a complement U_1 in V so that $U_0 \oplus U_1 = V$. We shall say that (U_0, U_1) is a direct-sum decomposition of V. Then each vector $u \in V$ can be uniquely decomposed as $u = u_0 + u_1$ where $u_i \in U_i$, $i = 0, 1$. Let

$$e(U_0, U_1) : V \to V; \quad u \mapsto u_1.$$

Then $e(U_0, U_1)$ is the projection of V onto U_1 determined by the direct sum decomposition. It is easy to verify that $e(U_0, U_1)$ is a linear operator on V such that

$$\mathsf{N}(e(U_0, U_1)) = U_0 \quad \text{and} \quad \mathsf{R}^{\mathrm{r}}(e(U_0, U_1)) = U_1.$$

Recall that (see Clifford and Preston, 1961, for example) an operator f on V (that is, $f \in \mathfrak{A}(V)$) is called an idempotent if $f^2 = f$. Given a direct sum decomposition (U_0, U_1) of V, the map $e(U_0, U_1)$ defined above is an idempotent. Some useful properties equivalent to being an idempotent are the following.

Lemma 6.3.1. The following statements are equivalent for an operator $f \in \mathfrak{A}$.

(1) f is an idempotent in $\mathfrak{A}$.

(2) $f|\mathsf{R}^{\mathrm{r}}(f)$ is the identity transformation on $\mathsf{R}^{\mathrm{r}}(f)$.

(3) $\mathsf{N}(f) \oplus \mathsf{R}^{\mathrm{r}}(f) = V$ and $f = e(\mathsf{N}(f), \mathsf{R}^{\mathrm{r}}(f))$.

Thus there is a bijection between idempotents in $\mathfrak{A}$ and direct decomposition of V.

Let $f \in \mathsf{H}(V, W)$ be a linear transformation. We shall say that $g \in \mathsf{H}(W, V)$ is a generalized inverse (or just ginverse for short) of f if

$$fgf = f; \tag{6.3.1}$$

g is called an semigroup inverse, or simply an sinverse if

$$fgf = f \quad \text{and} \quad gfg = g; \quad \text{and} \tag{6.3.2}$$

and g is called a group inverse, or simply an inverse if

$$fgf = f, \quad gfg = g \quad \text{and} \quad gf = fg. \tag{6.3.3}$$

We shall denote by $\mathsf{gV}(f)$, the set of all ginverses of f and by $\mathsf{V}(f)$, the set of all sinverses of f. If $g \in \mathsf{gV}(f)$, then we have $fgf = f$. So if $h = gfg$ then

$$fhf = (fgf)gf = fgf = f; \quad \text{and}$$
$$hfh = g(fgf)gfg = g(fgf)g = gfg = h.$$

Hence $h \in \mathsf{V}(f)$. Therefore if $\mathsf{gV}(f) \neq \emptyset$ so is $\mathsf{V}(f)$ and we have

$$\mathsf{V}(f) \subseteq \mathsf{gV}(f).$$

It can be shown that an inverse of f, if exists, is unique which will be denoted by f^{-1}. Notice that if f is invertible in the usual sense (that is, if $\det f \neq 0$) then f^{-1} is the inverse of f as defined above. However, if the inverse f^{-1} exists as above, f need not be invertible

Theorem 6.3.1. *Let $f, g \in \mathfrak{A}$. Then $g \in \mathsf{gV}(f)$ if and only if there is a complement U of $\mathsf{N}(f)$ such that*

$$g|\mathsf{R}^{\mathrm{r}}(f) = (f|U)^{-1} \tag{6.3.4}$$

A ginverse g is an sinverse if and only if

$$\mathsf{N}(f) \oplus \mathsf{R}^{\mathrm{r}}(g) = \mathsf{R}^{\mathrm{r}}(f) \oplus \mathsf{N}(g) = V. \tag{6.3.5}$$

In this case we have $\mathsf{R}^{\mathrm{r}}(fg) = \mathsf{R}^{\mathrm{r}}(g)$ *and* $\mathsf{R}^{\mathrm{r}}(gf) = \mathsf{R}^{\mathrm{r}}(f)$.

Furthermore, f has an inverse if and only if

$$\mathsf{N}(f) \oplus \mathsf{R}^{\mathrm{r}}(f) = V. \tag{6.3.6}$$

Proof 6.3.1. Suppose that $g \in \mathfrak{A}$ satisfies the given condition. By Proposition **??** $f|U$ is a linear isomorphism of U onto $\mathsf{R}^{\mathrm{r}}(f)$. Hence $(f|U)^{-1}$ is well-defined. Also, if $v \in V$, then $v = v_0 + v_1$ for a unique $v_0 \in \mathsf{N}(f)$ and $v_1 \in U$. Hence $vf = (v_1)f$ and by the given condition, we have

$$(v)fgf = (((v_1)f)g)f = (v_1)f = (v)f.$$

Thus $fgf = f$ and so, $g \in \mathsf{gV}(f)$. Conversely, let $g \in \mathsf{gV}(f)$. Then $e = fg$ is an idempotent such that $\mathsf{N}(fg) = \mathsf{N}(f)$. so, by Lemma **??**, $U = \mathsf{R}^{\mathrm{r}}(fg)$ is a complement of $\mathsf{N}(f)$. Also by Proposition **??**, $f|U : U \to \mathsf{R}^{\mathrm{r}}(f)$ is a linear isomorphism. Hence if $u \in U$,

$$(u)(f|U)(g|(\mathsf{R}^{\mathrm{r}}(f))) = (u)fg = u$$

by Proposition **??**. This proves that

$$g|\mathsf{R}^{\mathrm{r}}(f) = (f|U)^{-1}$$

as desired.

Suppose that $g \in \mathsf{V}(f)$ so that $fgf = f$ and $gfg = g$. Then $e = fg$ is an idempotent with $\mathsf{N}(f) = \mathsf{N}(fg)$ and so, by Lemma **??**, $\mathsf{R}^{\mathrm{r}}(fg) \oplus \mathsf{N}(f) = V$. Since $\mathsf{R}^{\mathrm{r}}(fg) \subseteq \mathsf{R}^{\mathrm{r}}(g)$, by Proposition **??**, we have

$$\dim \mathsf{R}^{\mathrm{r}}(fg) = \operatorname{Rank} f \leq \operatorname{Rank} g \quad \text{and by symmetry,} \quad \operatorname{Rank} g \leq \operatorname{Rank} f.$$

Consequently $\mathsf{R}^{\mathrm{r}}(fg) = \mathsf{R}^{\mathrm{r}}(g)$ and so, $\mathsf{N}(f) \oplus \mathsf{R}^{\mathrm{r}}(g) = V$. Again by symmetry we have $\mathsf{R}^{\mathrm{r}}(f) \oplus \mathsf{N}(g) = V$.

Conversely, assume that g satisfies Equation (**??**). Then the complement U in Equation (**??**) is $\mathsf{R}^{\mathrm{r}}(g)$ which shows that $g \in \mathsf{gV}(f)$ and $f \in \mathsf{gV}(g)$. Therefore $g \in \mathsf{V}(f)$ as desired. If $v \in \mathsf{R}^{\mathrm{r}}(fg)$, $v = ug$ for some $u \in V$. Hence $(v)fg =$

$(u)gfg = ug = v$ and so, $\mathsf{R}^{\mathrm{r}}(g) \subseteq \mathsf{R}^{\mathrm{r}}(fg)$. Since clearly, $\mathsf{R}^{\mathrm{r}}(fg) \subseteq \mathsf{R}^{\mathrm{r}}(g)$ we have $\mathsf{R}^{\mathrm{r}}(fg) = \mathsf{R}^{\mathrm{r}}(g)$. Similarly $\mathsf{R}^{\mathrm{r}}(gf) = \mathsf{R}^{\mathrm{r}}(f)$.

Suppose that $g = f^{-1}$ so that $g \in \mathsf{V}(f)$ and $fg = gf$. By Equation (**??**),

$$V = \mathsf{N}(f) \oplus \mathsf{R}^{\mathrm{r}}(g) = \mathsf{N}(f) \oplus \mathsf{R}^{\mathrm{r}}(fg) = \mathsf{N}(f) \oplus \mathsf{R}^{\mathrm{r}}(gf) = \mathsf{N}(f) \oplus \mathsf{R}^{\mathrm{r}}(f).$$

On the other hand, suppose that $\mathsf{N}(f) \oplus \mathsf{R}^{\mathrm{r}}(f) = V$ and let $e = e(\mathsf{N}(f), \mathsf{R}^{\mathrm{r}}(f))$. Then $f = e \circ f_0$ where $f_0 = f|\mathsf{R}^{\mathrm{r}}(f)$ is, by Proposition **??**, a linear automorphism of $\mathsf{R}^{\mathrm{r}}(f)$. It is easy to check that $g = e \circ f_0^{-1}$ is the inverse of f as defined in Equation (**??**).

Proposition **??** together with the last statement of the theorem above (see Equation (**??**)) gives the following:

Corollary 6.3.1. Let U and W be subspaces of a vector space V (with $\dim V = n$). Let

$$H(U, W) = \{f \in \mathfrak{A} : \mathsf{N}(f) = U, \quad \mathsf{R}^{\mathrm{r}}(f) = W\}.$$

Then $H(U, W) \neq \emptyset$ if and only if $\dim U + \dim W = \dim V$. Moreover, $H(U, W)$ is a multiplicative subgroup of $\mathfrak{A}(V)$ if and only if $U \oplus W = V$. If this is the case, then $H(U, W)$ is the group with identity $e(U, W)$ and inverse of any $f \in H(U, W)$ is the group inverse satisfying Equation (**??**).

Notice that $H(U, W)$ is a subgroup of $\mathfrak{A}$ if and only if $e(U, W)$ is an idempotent. Here U (and/or W) can take any value subject to the restriction that (U, W) is a direct-sum decomposition (see Lemma **??**). If $U = \{0\}$ is the trivial subspace, then $W = V$ and the group $H(U, W)$ is the automorphism group of V which is denoted by $\mathsf{G}L(V)$. $\mathsf{G}L(V)$ is called the general linear group of V. The corresponding subgroup of M_n is denoted by $\mathsf{G}L_n$ which is called the general linear group of rank n. This is the group of all invertible matrices in M_n and $\mathsf{G}L_n = \mathfrak{m}_B(\mathsf{G}L(V))$ for any basis B of V (see Theorem **??**).

If G is any multiplicative subgroup of $\mathfrak{A}$, its identity is an idempotent in $\mathfrak{A}$ and hence, by Lemma **??**, has the form $e(U, W)$ where (U, W) is a direct sum decomposition of V. Hence we must have $G \subseteq H(U, W)$ and so, $H(U, W)$ is a maximal subgroup of $\mathfrak{A}$. It is easy to see that the map

$$f \in H(U, W) \mapsto f|W : H(U, W) \to \mathsf{G}L(W)$$

is a natural isomorphism of $H(U, W)$ on to $\mathsf{G}L(W)$. Thus maximal subgroups of $\mathfrak{A}$ are, up to isomorphism, general linear groups of rank less than or equal to n.

Theorem **??** shows that we may construct ginverses and sinverses as follows.

Theorem 6.3.2. *Let $f \in \mathfrak{A}$. For any complement U of $\mathsf{R}^{\mathrm{r}}(f)$ and W of $\mathsf{N}(f)$ the map*

$$h = e(U, \mathsf{R}^{\mathrm{r}}(f)) \circ (f|W)^{-1}$$

is the unique sinverse of f with $\mathsf{N}(h) = U$ and $\mathsf{R}^{\mathrm{r}}(h) = W$. Every sinverse of f is constructed in this way. Moreover, $g \in \mathsf{gV}(f)$ if and only if

$$g = h + \alpha$$

where $h \in \mathsf{V}(f)$ and $\alpha \in \mathfrak{A}$ with $\alpha(\mathsf{R}^{\mathrm{r}}(f)) \subseteq \mathsf{N}(f)$.

Proof 6.3.2. Let U and W be as given above. Suppose that $f_0 = f|W$ so that $f_0 : W \to \mathsf{R}^{\mathrm{r}}(f)$ is a linear isomorphism by Proposition **??**. It follows from Lemma **??** that $f = e(\mathsf{N}(f), W) \circ f_0$. Also

$$\begin{aligned} h &= e(U, \mathsf{R}^{\mathrm{r}}(f)) \circ (f_0)^{-1} \\ &= e(U, \mathsf{R}^{\mathrm{r}}(f)) \circ \left((f_0)^{-1} \circ f_0(f_0)^{-1}\right) = hfh; \end{aligned}$$

and
$$\begin{aligned} fhf &= e(\mathsf{N}(f), W) \circ \left(f_0(f_0)^{-1} f_0\right) \\ &= e(\mathsf{N}(f), W) \circ f_0 = f. \end{aligned}$$

Therefore $h \in \mathsf{V}(f)$. Since $f_0 : W \to \mathsf{R}^{\mathrm{r}}(f)$ is a linear isomorphism, clearly,

$$\mathsf{N}(h) = \mathsf{N}(e(U, \mathsf{R}^{\mathrm{r}}(f)) = U \quad \text{and}$$
$$\mathsf{R}^{\mathrm{r}}(h) = \mathsf{R}^{\mathrm{r}}(f_0^{-1}) = W.$$

To prove the uniqueness of h assume that $g \in \mathsf{V}(f)$ with $\mathsf{N}(g) = U$, $\mathsf{R}^{\mathrm{r}}(g) = W$. By Theorem **??**, $g|\mathsf{R}^{\mathrm{r}}(f) = h|\mathsf{R}^{\mathrm{r}}(f) = (f_0)^{-1}$. Clearly $g|U = h|U$. Since $U \oplus \mathsf{R}^{\mathrm{r}}(f) = V$, it follows that $g = h$.

Suppose that $g = h + \alpha$ where $h \in \mathsf{V}(f)$ and $\alpha : \mathsf{R}^{\mathrm{r}}(f) \to \mathsf{N}(f)$ is a linear transformation. Let $v \in V$. Since $vf \in \mathsf{R}^{\mathrm{r}}(f)$, $(vf)\alpha \in \mathsf{N}(f)$ so that $((vf)\alpha)f = 0$. Therefore

$$(v)fgf = (v)fhf - (v)f\alpha f = (v)fhf = (v)f.$$

Since this hold for all $v \in V$, $fgf = f$ and so, $g \in \mathsf{gV}(f)$. Conversely, let $g \in \mathsf{gV}(f)$. Then $h = gfg \in \mathsf{V}(f)$. Since $\mathfrak{A}$ is an algebra, $\alpha = g - h \in \mathfrak{A}$. Let $w \in \mathsf{R}^{\mathrm{r}}(f)$. Then for $v \in V$ with $vf = w$, we have

$$(w\alpha)f = (vf)(gf - hf) = (v)fgf - (v)fhf = (v)f - (v)f = 0.$$

Hence $w\alpha \in \mathsf{N}(f)$ and so, $\alpha|\mathsf{R}^{\mathrm{r}}(f) : \mathsf{R}^{\mathrm{r}}(f) \to \mathsf{N}(f)$. This completes the proof.

The theorem above shows that any ginverse can be obtained as a suitable perturbation of an sinverse. One can perturb $h \in \mathsf{V}(f)$ so that the resulting ginverse is invertible.

It is clear from Theorem **??** that ginverses and sinverses of a singular linear transformation $f \in \mathfrak{A}$ are not unique since a subspace of V may have more than one (infinitely many if the field $\mathbb{k}$ is infinite) complements. However, given any $g \in \mathsf{gV}(f)$, every $h \in \mathsf{V}(f)$ can be constructed as follows:

Theorem 6.3.3. *Let $g \in \mathsf{gV}(f)$, U be a complement of $\mathsf{R}^{\mathrm{r}}(f)$ and W be a complement of $\mathsf{N}(f)$. Then*

$$h(U, W) = e(U, \mathsf{R}^{\mathrm{r}}(f)) \circ g \circ e(\mathsf{N}(f), W)$$

is the unique sinverse of f such that $\mathsf{N}(h(U, W)) = U$ and $\mathsf{R}^{\mathrm{r}}(h(U, W)) = W$. Consequently

$$\mathsf{V}(f) = \left\{h(U, W) : U \in \complement(\mathsf{R}^{\mathrm{r}}(f)), \quad W \in \complement(\mathsf{N}(f))\right\}.$$

Proof 6.3.3. The choice of U and W implies by Lemma **??** that

$$e(\mathsf{N}(f), W) \circ f = f = f \circ e(U, \mathsf{R}^{\mathrm{r}}(f)). \qquad \text{(a)}$$

Also, let $g_0 = g|\mathsf{R}^{\mathrm{r}}(f)$. Then by Equation (**??**), $g_0 = f_0^{-1}$ where $f_0 = (f|W')$ for some $W' \in \complement(\mathsf{N}(f))$. Since f_0 is a linear isomorphism of the subspace W' onto $\mathsf{R}^{\mathrm{r}}(f)$, g_0 is a linear isomorphism of $\mathsf{R}^{\mathrm{r}}(f)$ onto the complement W' of $\mathsf{N}(f)$. Then by Lemma **??** $e(\mathsf{N}(f), W)|W' : W' \to W$ is a linear isomorphism and so,

$$g \circ e(\mathsf{N}(f), W)|\mathsf{R}^{\mathrm{r}}(f) = g_0 \circ (e(\mathsf{N}(f), W)|W')$$

is a linear isomorphism of $\mathsf{R}^{\mathrm{r}}(f)$ onto W. Since

$$h = h(U, W) = e(U, \mathsf{R}^{\mathrm{r}}(f)) \circ g \circ e(\mathsf{N}(f), W) = e(U, \mathsf{R}^{\mathrm{r}}(f)) \circ (g \circ e(\mathsf{N}(f), W))\,|\mathsf{R}^{\mathrm{r}}(f),$$

it follows that $\mathsf{N}(h) = U$ and $\mathsf{R}^{\mathrm{r}}(h) = W$. Furthermore,

$$\begin{aligned} fhf &= (fe(U, \mathsf{R}^{\mathrm{r}}(f)))\, g\, (e(\mathsf{N}(f), W) f) \\ &= fgf = f && \text{by Equation(a);} \end{aligned}$$

and

$$\begin{aligned} hfh &= e(U, \mathsf{R}^{\mathrm{r}}(f))(gfg)e(\mathsf{N}(f), W) \\ &= e(U, \mathsf{R}^{\mathrm{r}}(f))(gfg|\mathsf{R}^{\mathrm{r}}(f))e(\mathsf{N}(f), W). \end{aligned}$$

Since $g_0 = f_0^{-1}$, we have $gfg|\mathsf{R}^{\mathrm{r}}(f) = g_0 f_0 g_0 = g_0$ and so,

$$hfh = e(U, \mathsf{R}^{\mathrm{r}}(f))(g_0)e(\mathsf{N}(f), W) = h.$$

Therefore $h = h(U, W) \in \mathsf{V}(f)$ and satisfies the desired conditions. Uniqueness of h follows from Theorem **??**. The last statement is now clear.

Remark 6.3.1. Here we have considered ginverses of linear transformations $f : V \to V$ in the algebra $\mathfrak{A}(V)$ of linear transformations of a (finite dimensional) vector space V. We may extend this to linear transformations of the form $h : V \to W$ from a vector space V to W. For let U be any complement of $\mathsf{R}^{\mathrm{r}}(h) = \operatorname{Im} h$ in W. If U' is any complement of $\mathsf{N}(h)$ in V, then $h|U'$ is a linear isomorphism of U' onto $\mathsf{R}^{\mathrm{r}}(h)$. Then the linear transformation

$$g = e(U, \mathsf{R}^{\mathrm{r}}(h)) \circ (f|U')^{-1}$$

satisfies $fgf = f$. Theorem **??** can also be extended to this more general case easily: any sinverse of a linear transformation $f : V \to W$ can be constructed from a given sinverse (or ginverse) g as

$$g' = e(U, \mathsf{R}^{\mathrm{r}}(h)) \circ g \circ e(\mathsf{N}(h), U')$$

where $e(U, \mathsf{R}^{\mathrm{r}}(h)) \in \mathfrak{A}(W)$ and $e(\mathsf{N}(h), U') \in \mathfrak{A}(V)$ are idempotents.

If A is an $n \times m$-matrix, it represents a linear transformation of $V = \Bbbk^n$ to $W = \Bbbk^M$ and the matrix G of a ginverse g of A constructed as above also satisfies $AGA = A$. Consequently if the matrix $\mathfrak{m}_g = A^-$ of g is known, the matrix of any sinverse g' of the matrix $\mathfrak{m}_f = A$ of f can be computed by multiplying A^- by suitable idempotent matrices.

6.3.1. Ginverses of matrices

In the following discussion we shall be concerned with $n \times n$-matrices (matrices in M_n). As observed above in Remark **??**, all this can be extended to the more general case of $n \times m$-matrices.

Recall that each $A \in \mathsf{M}_n$ the unique linear operator on E_n that send each $e_i \in E_n$ to $\mathfrak{r}_i$, $i = 1, 2, \ldots, n$ has matrix A in the natural basis of E_n (see Example 6.1.4). Since no confusion is likely we shall use the same notation for both the linear transformation and its matrix in the natural basis of E_n. We proceed to compute the matrix of some sinverse B of A. Suppose that

$$\mathfrak{r}_i^A = \mathfrak{r}_i = (a_{i1}, a_{i2}, \ldots, a_{in}), \quad 1 \le i \le n$$

be the set of rows of A and

$$\mathfrak{c}_j^A = \mathfrak{c}_j = (a_{1j}, a_{2j}, \ldots, a_{nj}), \quad 1 \le j \le n$$

be the set of columns. Then

$$A = \begin{pmatrix} \mathfrak{r}_1 \\ \mathfrak{r}_2 \\ \cdots \\ \mathfrak{r}_n \end{pmatrix} = \begin{pmatrix} \mathfrak{c}_1 & \mathfrak{c}_2 & \ldots & \mathfrak{c}_n \end{pmatrix}$$

The linear span

$$\mathsf{R}^{\mathfrak{r}}(A) = \mathcal{L}\langle \mathfrak{r}_1, \ldots, \mathfrak{r}_n \rangle$$

of row-vectors of A is called the row-space. Clearly $\mathsf{R}^{\mathfrak{r}}(f)$ is the range (image) of the linear transformation A. Similarly the space

$$\mathsf{R}^{\mathfrak{c}}(A) = \mathcal{L}\langle \mathfrak{c}_1, \ldots, \mathfrak{c}_n \rangle$$

is called the column space of A. Now $x \in \mathsf{N}(A)$ if and only if

$$xA = (\langle x, \mathfrak{c}_1 \rangle, \langle x, \mathfrak{c}_2 \rangle, \ldots, \langle x, \mathfrak{c}_n \rangle) = 0$$

which is true if and only if x is a solution of the system of homogeneous linear equations

$$\sum_{j=1}^{n} x_j a_{ji} = \langle x, \mathfrak{c}_i \rangle = 0 \quad \text{for all} \quad i = 1, 2, \ldots, n. \tag{6.3.7}$$

It follows that the space $\mathsf{R}^{\mathfrak{c}}(A)$ is the orthogonal to $\mathsf{N}(A)$. Also $\mathsf{R}^{\mathfrak{c}}(A)$ is the row-space of A', the transpose of A and it is well-known that $\operatorname{Rank} A = \operatorname{Rank} A'$. Hence $\dim \mathsf{R}^{\mathfrak{r}}(A) = \dim \mathsf{R}^{\mathfrak{c}}(A) = \operatorname{Rank} A = t$. Since $\mathsf{R}^{\mathfrak{c}}(A) \cap \mathsf{N}(A) = \{0\}$, it follows that $\mathsf{R}^{\mathfrak{c}}(A)$ is the orthogonal complement of $\mathsf{N}(A)$. Hence A induces a linear isomorphism of $\mathsf{R}^{\mathfrak{c}}(A)$ onto $\mathsf{R}^{\mathfrak{r}}(A)$. Therefore there exists an ordered set of rows of A

$$A_{\mathfrak{r}} = \left\{ \mathfrak{r}_{i_1}, \mathfrak{r}_{i_2}, \ldots, \mathfrak{r}_{i_t} \right\} \qquad \text{with} \quad i_1 < i_2 < \cdots < i_t \tag{6.3.8}$$

that forms a basis of $\mathsf{R}^{\mathfrak{r}}(f)$. Similarly there is an ordered set of columns of A

$$A_{\mathfrak{c}} = \left\{ \mathfrak{c}_{j_1}, \mathfrak{c}_{j_2}, \ldots, \mathfrak{c}_{j_t} \right\} \qquad \text{with} \quad j_1 < j_2 < \cdots < j_t \tag{6.3.9}$$

that forms a basis of $\mathsf{R}^{\mathfrak{c}}(A)$.

We say that a matrix $D \in \mathfrak{A}$ dominates a matrix A with respect to a row-base A_{r} if

$$\mathfrak{r}_k^D = \mathfrak{r}_k^A \quad \text{if and only if} \quad k = i_l, \quad \text{for some} \quad 1 \le l \le t. \tag{6.3.10}$$

If D is non-singular, then the set of rows of D not in A_{r} spans a complement of $\mathsf{R}^{\mathrm{r}}(A)$. Conversely, given a row-base A_{r} and a basis $\mathcal{B}_U$ of a complement U of $\mathsf{R}^{\mathrm{r}}(A)$, there is a unique matrix $D = \mathfrak{m}(A_{\mathrm{r}}, \mathcal{B}_U)$ such that D dominates A and the set of rows of D not in A_{r} is $\mathcal{B}_U$.

As above let $D = (m_{ij}) = (A_{\mathrm{r}}, \mathcal{B}_U)$ be a non-singular matrix dominating A. By definition of the linear transformation A,

$$A|W; W \to \mathsf{R}^{\mathrm{r}}(A) \quad \text{where} \quad W = \mathcal{L}\langle e_{i_1}, e_{i_2}, \dots, e_{i_n}\rangle.$$

is a linear isomorphism of W onto $\mathsf{R}^{\mathrm{r}}(A)$. Hence W is a complement of $\mathsf{N}(A)$. Again by the definition of $D = \mathfrak{m}(A_{\mathrm{r}}, \mathcal{B}_U)$, $D|W = A|W$. If $M = (m_{ij}^*) = D^{-1}$, we have

$$M|\mathsf{R}^{\mathrm{r}}(A) = (D|W)^{-1} = (A|W)^{-1}.$$

By Theorem **??** M is an invertible ginverse of A.

Proposition 6.3.1. Let $A \in \mathfrak{A}_n$ and let $D = \mathfrak{m}(A_{\mathrm{r}}, \mathcal{B}_U)$ be a non-singular matrix dominating A where $\mathcal{B}_U$ is a basis of the complement U of $\mathsf{R}^{\mathrm{r}}(A)$. Then D^{-1} is a non-singular ginverse of A.

Given $A \in \mathfrak{A}$ and a complement U of $\mathsf{R}^{\mathrm{r}}(A)$, the result above shows that it is possible to calculate one ginverse M of A provided we can compute A_{r} and $\mathcal{B}_U$. Standard linear algebraic methods are available for these computations. Clearly M depends on the choice of U. If we set $U = \mathsf{R}^{\mathrm{r}}(A)^{\perp}$, the orthogonal complement of $\mathsf{R}^{\mathrm{r}}(A)$, we may choose $\mathcal{B}_U$ as the set of all linearly independent solutions of the system of equations

$$\langle x, \mathfrak{r}_{i_k}\rangle = 0 \quad \text{for all} \quad k = 1, 2, \dots, t.$$

A complete system of linearly independent solutions may be explicitly computed in terms of A using Cramer's (see Nambooripad, 2000, page 109). Consequently it is always possible to compute one (invertible) ginverse. Then Theorem **??** shows that any sinverse can be constructed by pre- and post-multiplying by suitable idempotents and by Theorem **??**, any ginverse can be computed by a suitable perturbation of an sinverse.

Again by Theorem **??** any sinverse $g \in \mathsf{V}(A)$ is uniquely determined by a complement U of $\mathsf{R}^{\mathfrak{r}}(A)$ and a complement W of $\mathsf{N}(A)$ so that

$$\mathsf{N}(g) = U \quad \text{and} \quad \mathsf{R}^{\mathfrak{r}}(g) = W.$$

We proceed to compute the matrix of g in the natural basis of E_n. Let $\mathcal{B}_U$ and $\mathcal{B}_W$ be bases of U and W respectively. Also let $D = \mathfrak{m}(A_{\mathfrak{r}}, \mathcal{B}_U)$ be a nonsingular matrix dominating A. If $M = (m^*_{ij}) = D^{-1}$ then $MD = I_n$. Since $e_i = (\delta_{i1}, \delta_{i2}, \ldots, \delta_{in})$ is the i^{th} row vector of I_n, we have

$$e_i = \sum_{\mathfrak{r}_k \in A_{\mathfrak{r}}} m_{ik}\mathfrak{r}_k + \sum_{u_j \in \mathcal{B}_U} m^*_{ij} u_j$$

for all $i = 1, 2, \ldots, n$. Since $u_i \in \mathsf{N}(g)$, application of g to either side of the above equation gives

$$(e_i)g = \sum_{\mathfrak{r}_k \in A_{\mathfrak{r}}} m^*_{ik}(\mathfrak{r}_k)g.$$

Since $g|\mathsf{R}^{\mathfrak{r}}(A)$ is an isomorphism, by Equation (**??**), for each $k = 1, 2, \ldots, t$ there exists a unique $w_k \in W$ such that

$$(\mathfrak{r}_{i_k})g = w_k = \sum_{j=1}^{n} w_{kj} e_j$$

and $\{w_k : 1 \leq k \leq t\}$ is a basis of W. Therefore

$$\begin{aligned}
(e_i)g &= \sum_{k=1}^{t} m^*_{ii_k} w_k \\
&= \sum_{k=1}^{t} m^*_{ii_k} \left(\sum_{j=1}^{n} w_{kj} e_j \right) \\
&= \sum_{j=1}^{n} \left(\sum_{k} m^*_{ii_k} w_{kj} \right) e_j
\end{aligned}$$

for all $1 \leq i \leq n$.

Thus we have proved the first part of the following. The converse part is an immediate consequence of Theorem **??** and the first part.

Theorem 6.3.4. *Let $A \in \mathsf{M}_n$ and $g \in \mathsf{V}(A)$. Suppose that $\mathsf{N}(g) = U$ and $\mathsf{R}^{\mathfrak{r}}(g) = W$. Further, let*

$$A_{\mathfrak{r}} = \{\mathfrak{r}_{i_1}, \ldots, \mathfrak{r}_{i_t}\}$$

be a row-base of A. If $\mathfrak{m}^{-1} = (m^*_{ij})$ be the inverse of the non-singular matrix $\mathfrak{m} = \mathfrak{m}(A_{\mathfrak{r}}, \mathcal{B}_U)$ that dominates A, then

$$\mathfrak{m}(g) = (h_{ik}) \quad \textit{where} \quad h_{ij} = \sum_j m^*_{i_j} w_{jk}. \tag{6.3.11}$$

is the matrix of the sinverse g of A.

Conversely, let $\mathcal{B}_U = \{u_j\}$ and $\mathcal{B}_W = \{w_i\}$ be bases of complements U and W of $\mathsf{R}^{\mathfrak{r}}(A)$ and $\mathsf{N}(A)$ respectively. If $\mathfrak{m}^{-1} = (m^*_{ij})$ is the inverse of $\mathfrak{m}(A_{\mathfrak{r}}, \mathcal{B}_U)$, the matrix formed as in Equation (**??**) *is the matrix of an sinverse of A.*

The theorem above gives a construction of an arbitrary sinverse of a matrix $A \in \mathsf{M}_n$. The construction uses the following data:

(a) A row-base $A_{\mathfrak{r}}$ of A.

(b) A basis $\mathcal{B}_U$ of a complement U of the subspace $\mathsf{R}^{\mathfrak{r}}(A)$.

(c) A basis $\mathcal{B}_W$ of a complement W of the subspace $\mathsf{N}(A)$.

Obtaining these items of data in the appropriate format may need some computations. Known computational methods are available for this purpose. The remaining computation consist mainly of inversion of a non-singular matrix for which also classical methods are available. Finally, as shown by Equation (**??**), we need to compute certain inner products which are quite straightforward.

Various special types of ginverses can be computed by choosing the complements U of $\mathsf{R}^{\mathfrak{r}}(A)$ and W of $\mathsf{N}(A)$ suitably and using the procedure described in Theorem **??**. An important particular case is obtained by choosing $U = \mathsf{R}^{\mathfrak{r}}(A)^{\perp}$ and letting W as an arbitrary complement of $\mathsf{N}(A)$ (and or choosing $W = \mathsf{N}(A)^{\perp}$ and U as arbitrary). If $U = \mathsf{R}^{\mathfrak{r}}(A)^{\perp}$, then $\mathcal{B}_U = \{u_1, \ldots, u_{n-t}\}$ may be chosen as the maximal set of linearly independent solutions

$$\sum_{i=1}^{n} a_{ik} x_k = \langle x, \mathfrak{r}_i \rangle = 0 \quad \text{for all} \quad i \in \{1, 2, \ldots, n\}. \tag{6.3.12}$$

It is possible to compute these explicitly in terms of A using the standard Cramer's rule (see Nambooripad, 2000, page 109). $W = \mathsf{N}(A)^{\perp}$ we may similarly choose

$\mathcal{B}_W = \{w_1, \ldots, w_t\}$ as the maximal set of linearly independent solution of the system of homogeneous linear equations (**??**) We will indicate the significance of these choices in the next section. In particular, the choices

$$W = \mathsf{N}(A)^{\perp} \quad \text{and} \quad U = \mathsf{R}^{\mathrm{r}}(A)^{\perp}$$

give a unique sinverse $A^{\dagger}$ which is called the Moore-Penrose inverse of A.

6.4. Systems of Linear Equations

Here we consider applications of ginverses in solving systems of linear equations. Again we shall confine ourselves to square matrices since extension to the more general case is routine.

Let $A \in \mathsf{M}_n$ and $x = (x_1, \ldots, x_n)$. Then for all $y \in \mathsf{E}_n$, the the matrix equation

$$xA = y \tag{6.4.1}$$

can be equivalently written as the system of simultaneous linear equations

$$\langle x, \mathfrak{c}_i \rangle = \sum_{j=1}^{n} a_{ji} x_j = y_i \quad \text{for all} \quad i = 1, 2, \ldots, n. \tag{6.4.2}$$

The equation above is called the system of linear equations determined by the matrix A. It has a solution $v \in \mathsf{E}_n$ if and only if $vA = y$ so that y is the image of v under the linear transformation A; that is, $y \in \mathsf{R}^{\mathrm{r}}(A)$. If this is the case, we say that the given system of equation is consistent. Notice that, for any solution v of the given system, $v + u$ is also a solution provided u is the solution of the homogeneous system $uA = 0$. Hence for any solution v of the system (6.4.1), $v + \mathsf{N}(A)$ represents a complete set of solutions of (6.4.1). Since any basis

$$\mathcal{B}_{\mathsf{N}(A)} = \{\nu_1, \nu_2, \ldots, \nu_{n-t}\}$$

of $\mathsf{N}(A)$ represents a complete linearly independent set of solutions of the homogeneous system $xA = 0$, it is easy to see that

$$\mathcal{B}_A = \{v, v + \nu_1, \ldots, v + \nu_{n-t}\}$$

represents a complete set of (affine) independent solutions of (6.4.1).

Theorem 6.4.1. *Let $A \in \mathsf{M}_n$ and let $\tilde{A}$ be a ginverse of A. If the linear equation (6.4.1) determined by A is consistent then $\hat{x} = y\tilde{A}$ is a solution and $y\tilde{A} + \mathsf{N}(A)$ represents the set of all solutions of (6.4.1).*

Proof 6.4.1. If (6.4.1) is consistent then for any $\tilde{A} \in \mathsf{gV}(A)$, there is $v \in \mathsf{E}_n$ such that $vA = y$ and we have

$$y\tilde{A}A = (vA)\tilde{A}A = (v)(A\tilde{A}A) = (v)A = y.$$

Therefore $\hat{x} = y\tilde{A}$ is a solution of (6.4.1)

Clearly for any $\tilde{A} \in \mathsf{V}(A)$ $y\tilde{A} + u$ is a solution of (6.4.1) for all $u \in \mathsf{N}(A)$. Conversely if $x_1 = \hat{x}$ is any solution, then $(x_1)A = y$ and so, $u = x_1 - x_0 \in \mathsf{N}(A)$ where $x_0 = y\tilde{A}$. Hence $x_1 = x_0 + u \in x_0 + \mathsf{N}(A)$. Therefore $x_0 + \mathsf{N}(A)$ represents the set of all solutions of (6.4.1).

The theorem above applies for all systems of linear equations both homogeneous and inhomogeneous. Next we will deal with inhomogeneous systems.

Theorem 6.4.2. *Let (6.4.1) be a system of consistent linear equations and $y \neq 0$. If $\hat{x}$ is any solution of (6.4.1), then $\hat{x} \neq 0$ and there is a $\tilde{A} \in \mathsf{V}(A)$ such that $\hat{x} = y\tilde{A}$. Consequently*

$$\hat{x} + \mathsf{N}(A) = \{y\tilde{A} : \tilde{A} \in \mathsf{V}(A)\} \tag{6.4.3}$$

is the set of all solutions of the equation (6.4.1).

Proof 6.4.2. Let x_0 be a solution of (6.4.1). Then $(x_0)A = y \neq 0$ and so, $x_0 \neq 0$. In particular $x_0 \notin \mathsf{N}(A)$. It follows that there is a complement W of $\mathsf{N}(A)$ such that $x_0 \in W$. By Theorem **??** we can find $\tilde{A} \in \mathsf{V}(A)$ such that $\mathsf{R}^{\mathrm{r}}(\tilde{A}) = W$. Since $(x_0)A = y$, by Theorem **??**, $y\tilde{A} = x_0$. The last statement is now obvious.

For $A \in \mathsf{M}_n$, let W be a complement of $\mathsf{N}(A)$. If $x_0 \in W$ then x_0 is a solution of the equation $xA = y$ determined by A where $y = (x_0)A$. Since $A|W : W \to \mathsf{R}^{\mathrm{r}}(A)$ is a linear isomorphism, it is clear that x_0 is the solution of exactly one such equation determined by A. Moreover, by Theorem **??**, every consistent equation determined by A has a unique solution in W.

Proposition 6.4.1. Let $A \in \mathsf{M}_n$ and let W be a complement of $\mathsf{R}^{\mathrm{r}}(A)$. Then every $x_0 \in W$ is a solution of a unique linear equation (6.4.1). Moreover, every consistent linear equation (6.4.1) has a unique solution in W.

We now consider some special type of sinverses that afford solutions of linear equations with various properties.

6.4.1. Ginverse for minimum norm solutions

We say that a ginverse $\tilde{A} \in \mathsf{V}(A)$ gives "minimum norm" solution of the equation $xA = y$ (equation (6.4.1))if the solution $y\tilde{A}$ has minimum norm; that is, $\tilde{A}$ satisfies the condition

$$||y\tilde{A}|| = \inf_{x \in y\tilde{A}+\mathsf{N}(A)} ||x||. \tag{6.4.4}$$

Theorem 6.4.3. *Let $A \in \mathsf{M}_n$. Then for $A^+ \in \mathsf{V}(A)$, yA^+ gives the minimum norm solution of equation (6.4.1) if and only if*

$$\mathsf{R}^{\mathrm{r}}(A^+) = \mathsf{N}(A)^{\perp}. \tag{6.4.5}$$

When A^+ satisfies (6.4.5) above, yA^+ is the unique solution of (6.4.1) with minimum norm. Moreover, if A^+ gives minimum norm of (6.4.1) for some $y \in \mathsf{R}^{\mathrm{r}}(A)$ then it gives minimum norm for all $y \in \mathsf{R}^{\mathrm{r}}(A)$.

Proof 6.4.3. Fix $y \in \mathsf{R}^{\mathrm{r}}(A)$ and let v be a solution of (6.4.1). Since $\mathsf{N}(A) \oplus W = \mathsf{E}_n$ where $W = \mathsf{N}(A)^{\perp}$ there exists unique $v_0 \in \mathsf{N}(A)$ and $v_1 \in W$ with $v = v_0 + v_1$. Then

$$vA = (v_0)A + (v_1)A = (v_1)A = y$$

so that v_1 is a solution of (6.4.1) in W. Since $\langle v_0, v_1 \rangle = 0$ we have

$$||v||^2 = ||v_0||^2 + ||v_1||^2 \geq ||v_1||^2$$

Since $v_1 \in W$, it follows by the definition of A^+ that $v_1 = yA^+$. To see that v_1 is unique, we observe that $v_0 = 0$ if v where a solution with minimum norm so that $v \in W$. It follows from Proposition **??** that $v = v_1$.

Since the condition (6.4.5) does not involve y, the last statement is obvious.

6.4.2. Ginverse for least square solutions

Another question often comes up when dealing with optimization problems is the following: If the equation $xA = y$ is not consistent, can we find an optimal solution ? That is, we wish to find a vector z such that

$$||zA - y|| = \inf_{x \in \mathsf{E}_n} ||xA - y||. \tag{6.4.6}$$

where $||z||$ denote the euclidean norm of z. A vector z with this property is called a least square solution (abbreviated as 'lss') of the inconsistent equation (6.4.1).

We say that a ginverse $\tilde{A} \in \mathsf{V}(A)$ gives a least square solution of (6.4.1) if $z = y\tilde{A}$ is a least square solution of (6.4.1).

Theorem 6.4.4. *Let $A \in \mathsf{M}_n$. Then $z \in \mathsf{E}_n$ is a least square solution of (6.4.1) if and only if there is $A_+ \in \mathsf{V}(A)$ satisfying the condition*

$$\mathsf{N}(A_+) = \mathsf{R}^{\mathrm{r}}(A)^{\perp} \tag{6.4.7}$$

such that $z = yA_+$. If $z, s \in \mathsf{E}_n$ are least square solutions of (6.4.1) if and only if $s = z + u$ for some $u \in \mathsf{N}(A)$.

Furthermore if A_+ gives least square solution of (6.4.1) for some $y \in \mathsf{R}^{\mathrm{r}}(A)$ then it gives least square solution for all $y \in \mathsf{R}^{\mathrm{r}}(A)$.

Proof 6.4.4. Suppose that $A_+ \in \mathsf{V}(A)$ satisfies (6.4.7). If $z = yA_+$ then we have

$$(zA - y)\,A_+ = 0$$

and so, $zA - y \in \mathsf{N}(A_+)$. Furthermore for any $x \in \mathsf{E}_n$

$$||xA - y|| = ||zA - y + wA||.$$

where $w = x - z$. Since $\mathsf{N}(A_+) = \mathsf{R}^{\mathrm{r}}(A)^{\perp}$, we have $\langle zA - y, wA\rangle = 0$ and so,

$$\begin{aligned} ||xA - y||^2 &= ||zA - y||^2 + ||wA||^2 \\ &\geq ||zA - y||^2. \end{aligned}$$

Therefore z satisfies Equation (**??**) and hence it is a least square solution of equation (6.4.1). If $z = yA_+$ is an lss of (6.4.1), and if $u \in \mathsf{N}(A)$ then $s = z + u$ is clearly an lss of (6.4.1).

As above let $z = yA_+$ where A_+ satisfies (is clearly an lss of (6.4.7)). Suppose that s is a lss of equation (is clearly an lss of (6.4.1)) other than z. If $u = s - z$, as above $\langle zA - y, uA\rangle = 0$ and so,

$$\begin{aligned} ||sA - y||^2 &= ||zA - y + uA||^2 \\ &= ||zA - y||^2 + ||uA||^2 \\ &= ||zA - y||^2. \end{aligned}$$

Therefore $uA = 0$ and so $u \in \mathsf{N}(A)$. Consequently $z, s \in \mathsf{E}_n$ are least square solutions of (is clearly an lss of (6.4.1)) if and only if $s = z+u$ for some $u \in \mathsf{N}(A)$.

Since equation (is clearly an lss of (6.4.1)) is inconsistent $z \notin \mathsf{N}(A)$ and so, $s = z + u \notin \mathsf{N}(A)$. Therefore there is a complement W_s of $\mathsf{N}(A)$ containing s and such that $ze(\mathsf{N}(A), W_s) = s$. By Theorem **??**

$$A_{++} = A_{+}e(\mathsf{N}(A), W_s)$$

is an sinverse of A satisfying (is clearly an lss of (6.4.7)) and such that $s = yA_{++}$.

Since the condition (is clearly an lss of (6.4.7)) does not involve y, it is clear that if yA_{+} is an lss for some y then yA_{+} is an lss for every $y \in \mathsf{E}_n$.

Corollary 6.4.1. The Moore-Penrose inverse $A^{\dagger}$ gives the least square solution of equation (is clearly an lss of (6.4.1).) with minimum norm.

Proof 6.4.5. By definition the Moore-Penrose inverse satisfies the conditions (6.4.5) and (is clearly an lss of (6.4.7)) and so $z = yA^{\dagger}$ is both a minimum norm solution as well as a least square solution of (is clearly an lss of (6.4.1)) by (Theorem **??**, Theorem **??**.)

For an extensive treatment of ginverses of metrices we refer the reader to(Rao and Mitra, 1971). Here we shall go through some of the results from their work; but the treatment will be more geometric.

Example 6.4.1. Let $\mathbb{k}_n[x]$ be the set of all polynomials of degree less than $n \in \mathbb{N}$. Prove that $\mathbb{k}_n[x]$ is a subspace of the vector space $\mathbb{k}[x]$ of Example 6.1.5 and that $B = \{1, x, \dots, x^{n-1}\}$ is a basis of $\mathbb{k}_n[x]$ over $\mathbb{k}$. Consequently $\mathbb{k}_n[x]$ is a finite dimensional vector space with $\dim \mathbb{k}_n[x] = n$. Let $\mathbb{k}^n$ denote the vector space of Example 6.1.6. Prove that the mapping $f : \mathbb{k}_n[x] \to \mathbb{k}^n$, defined by

$$f(p(x)) = (a_0, a_1, \dots, a_{n-1}) \quad \text{where} \quad p(x) = a_0 + a_1 x + \cdots + a_{n-1}x^{n-1} \in \mathbb{k}_n[x]$$

is an isomorphism of $\mathbb{k}_n[x]$ onto $\mathbb{k}^n$.

Example 6.4.2. Let B be a basis of V over $\mathbb{k}$ and let $\mathbb{k}^B$ be the vector space of Example 6.1.4. Then the map $\sigma : v \mapsto e_v$ is a bijection of B onto the basis E_B of $\mathbb{k}^B$ (see Example 6.1.4). Then by Theorem **??** $f_\sigma : V \to \mathbb{k}^B$ is an isomorphism of vector spaces, called the coefficient isomorphism determined by the basis B. If $v = \sum_B a_u u \in V$, by (**??**), we have

$$f_\sigma(v) = \sum_B a_u \sigma(u) = \sum_B a_u e_u$$

and so $$f_\sigma(v)(u) = a_u \qquad \text{for all } u \in B.$$

Thus for each $v \in V$, $f_\sigma(v)$ is the function on B which maps $u \in B$ to the coefficient of u in the linear combination for v with respect to B; we denote this isomorphism by $\mathfrak{c}_B$. In particular, if B is finite containing n elements, as noted in Example 6.1.4, functions on $B = \{u_1, \ldots, u_n\}$ to $\Bbbk$ may be identified as n-tuples. Hence, in this case, $\mathfrak{c}B(v) = (a_1, \ldots, a_n)$ if $v = \sum_{i=1}^{n} a_i u_i$.

Example 6.4.3. Let $\Bbbk[x]$ denote the vector space of all polynomials over $\Bbbk$ (see Example 6.1.5). For $p(x) = \sum_{i=0}^{r} a_i x^i \in \Bbbk[x]$, let $(p(x))D^t$ denote the t-th derivative of $p(x)$. Then the map

$$D^n : p(x) \mapsto (p(x))D^n$$

is a linear transformation $D^n : \Bbbk[x] \to \Bbbk[x]$. Show that D^n is surjective and that $\mathsf{N}(D^n) = \Bbbk[x]_n$. Deduce that D^n is not injective.

Example 6.4.4. If $n, r \in \mathbb{N}$ prove that $D^r : \Bbbk[x]_n \to \Bbbk[x]_n$ is a linear transformation which is non-zero if $r < n$. Find the matrix of D^r in the basis $\{1, x, x^2, \ldots, x^{n-1}\}$.

Example 6.4.5. Call a function $f : [a, b] \to \mathbb{C}$ smooth if f has continuous derivatives of all orders. Prove that the set $C^\infty[a, b]$ of all smooth complex-valued functions on the interval $[a, b] \subseteq \mathbb{R}$ is a vector space under point-wise addition and scalar multiplication: For $f, g \in C^\infty[a, b]$ and $\alpha \in \mathbb{C}$,

$$(f + g)(x) = f(x) + g(x) \quad \text{and} \quad (\alpha f)(x) = \alpha(f(x))$$

for all $x \in [a, b]$. If D^t denote the t-th derivative (see Example 6.4.3), show that $T(D) = \sum_{i=0}^{n} b_i D^i$ is a linear transformation in $\mathfrak{A}(C^\infty[a, b])$. Show further that the subspace $\mathsf{N}(T(D))$ is finite dimensional. Find $\dim \mathsf{N}(T(D))$. Determine also the subspace $\mathsf{R}^{\mathfrak{r}}(T(D))$.

References

Apostol, Tom M. (1985). *Mathematical Analysis*, Narosa Publishing House, New Delhi, Madras.

Artin, M. (1990). *Algebra*, Prentice-Hall of India, Private limited, M-97, Connaught Circus, New Delhi -110 001. Printed in India by special arrangement with Prentice-Hall, Inc., Englewood Cliffs, N.J., U.S.A.; Fourth Printing : December 1998.

Clifford A.H. (1941). Semigroups admitting relative inverses. *Annals of Math.*, 1037-1049.

Clifford A.H. and Preston G.B. (1961). *Algebraic theory of semigroups*, Math. surveys No.7, American Mathematical Socety, Providence, Rhode Island.

Limaye, B.V. (1981). *Functional Analysis*, Wiley Eastern, New Delhi.

Moore, R.A. (1935). General Analysis, *Memoirs of the Amer. Phil. Soc.*, **1** Chapter 3, Section 29. Amer. Phil. Soc., Philadelphia.

Nambooripad, K.S.S. (2000). *Matrix Groups*, CMS Preprint. Centre for Mathematical Sciences, Thiruvananthapuram, This is a first draft.

Nanda, Sribatsa and Nanda, Sudarsan. (1990). *General Topology*. Macmillan India Limited.

Penrose, R. (1955). A generalized inverse for matrices, *Proc. Cambridge Philos. Soc.*, **51**, 406-413.

Preston, G. B. (1954). Inverse semi-groups, *J. London Math. Soc.*, **29**, 396-403.

Rao, C. R. and Mitra, S. K. (1971). *Generalized Inverse of Matrices and its Applications*. John Wiley and Spons, New York.

Skornyakov, L. A. (1964). *Complemented Modular Lattices and Regular Rings*. Olive & Boyd, Edinburg & London. This is the translation of the book by the same author in Ruuian published in 1961(Moscow).