Appendix A
Vector Spaces

This appendix reviews some of the basic definitions and properties of vector spaces. It is presumed that, with the possible exception of Theorem A.14, all of the material presented here is familiar to the reader.

Definition A.1. A set $\mathscr{M}$ is a *vector space* if, for any $x, y, z \in \mathscr{M}$ and scalars α, β, operations of vector addition and scalar multiplication are defined such that:

(1) $(x+y)+z = x+(y+z)$.
(2) $x+y = y+x$.
(3) There exists a vector $0 \in \mathscr{M}$ such that $x+0 = x = 0+x$ for any $x \in \mathscr{M}$.
(4) For any $x \in \mathscr{M}$, there exists $y \equiv -x$ such that $x+y = 0 = y+x$.
(5) $\alpha(x+y) = \alpha x + \alpha y$.
(6) $(\alpha+\beta)x = \alpha x + \beta x$.
(7) $(\alpha\beta)x = \alpha(\beta x)$.
(8) There exists a scalar ξ such that $\xi x = x$. (Typically, $\xi = 1$.)

In nearly all of our applications, we assume $\mathscr{M} \subset \mathbf{R}^n$.

Definition A.2. Let $\mathscr{M}$ be a vector space, and let $\mathscr{N}$ be a set with $\mathscr{N} \subset \mathscr{M}$. $\mathscr{N}$ is a *subspace* of $\mathscr{M}$ if and only if $\mathscr{N}$ is a vector space.

Vectors in $\mathbf{R}^n$ will be considered as $n \times 1$ matrices. The 0 vector referred to in Definition A.1 is just an $n \times 1$ matrix of zeros. Think of vectors in three dimensions as $(x,y,z)'$, where w' denotes the *transpose* of a matrix w. The subspace consisting of the z axis is

$$\left\{ \begin{pmatrix} 0 \\ 0 \\ z \end{pmatrix} \middle| z \in \mathbf{R} \right\}.$$

R. Christensen, *Plane Answers to Complex Questions: The Theory of Linear Models*, Springer Texts in Statistics, DOI 10.1007/978-1-4419-9816-3,

The subspace consisting of the x, y plane is

$$\left\{ \begin{pmatrix} x \\ y \\ 0 \end{pmatrix} \middle| \, x, y \in \mathbf{R} \right\}.$$

The subspace consisting of the plane that is perpendicular to the line $x = y$ in the x, y plane is

$$\left\{ \begin{pmatrix} x \\ -x \\ z \end{pmatrix} \middle| \, x, z \in \mathbf{R} \right\}.$$

Theorem A.3. Let $\mathscr{M}$ be a vector space, and let $\mathscr{N}$ be a nonempty subset of $\mathscr{M}$. If $\mathscr{N}$ is closed under vector addition and scalar multiplication, then $\mathscr{N}$ is a subspace of $\mathscr{M}$.

Theorem A.4. Let $\mathscr{M}$ be a vector space, and let $x_1, \ldots, x_r$ be in $\mathscr{M}$. The set of all linear combinations of $x_1, \ldots, x_r$, i.e., $\{v \mid v = \alpha_1 x_1 + \cdots + \alpha_r x_r, \alpha_i \in \mathbf{R}\}$, is a subspace of $\mathscr{M}$.

Definition A.5. The set of all linear combinations of $x_1, \ldots, x_r$ is called the *space spanned by* $x_1, \ldots, x_r$. If $\mathscr{N}$ is a subspace of $\mathscr{M}$, and $\mathscr{N}$ equals the space spanned by $x_1, \ldots, x_r$, then $\{x_1, \ldots, x_r\}$ is called a *spanning set* for $\mathscr{N}$.

For example, the space spanned by the vectors

$$x_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad x_2 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

consists of all vectors of the form $(a, b, b)'$, where a and b are any real numbers.

Let A be an $n \times p$ matrix. Each column of A is a vector in $\mathbf{R}^n$. The space spanned by the columns of A is called the *column space* of A and written $C(A)$. (Some people refer to $C(A)$ as the *range space* of A and write it $R(A)$.) If B is an $n \times r$ matrix, then $C(A, B)$ is the space spanned by the $p + r$ columns of A and B.

Definition A.6. Let $x_1, \ldots, x_r$ be vectors in $\mathscr{M}$. If there exist scalars $\alpha_1, \ldots, \alpha_r$ not all zero so that $\sum \alpha_i x_i = 0$, then $x_1, \ldots, x_r$ are *linearly dependent*. If such α_is do not exist, $x_1, \ldots, x_r$ are *linearly independent*.

Definition A.7. If $\mathscr{N}$ is a subspace of $\mathscr{M}$ and if $\{x_1, \ldots, x_r\}$ is a linearly independent spanning set for $\mathscr{N}$, then $\{x_1, \ldots, x_r\}$ is called a *basis* for $\mathscr{N}$.

Theorem A.8. If $\mathscr{N}$ is a subspace of $\mathscr{M}$, all bases for $\mathscr{N}$ have the same number of vectors.

Theorem A.9. If $v_1, \ldots, v_r$ is a basis for $\mathscr{N}$, and $x \in \mathscr{N}$, then the characterization $x = \sum_{i=1}^r \alpha_i v_i$ is unique.

PROOF. Suppose $x = \sum_{i=1}^r \alpha_i v_i$ and $x = \sum_{i=1}^r \beta_i v_i$. Then $0 = \sum_{i=1}^r (\alpha_i - \beta_i) v_i$. Since the vectors v_i are linearly independent, $\alpha_i - \beta_i = 0$ for all i. $\square$

Definition A.10. The *rank* of a subspace $\mathscr{N}$ is the number of elements in a basis for $\mathscr{N}$. The rank is written $r(\mathscr{N})$. If A is a matrix, the rank of $C(A)$ is called the rank of A and is written $r(A)$.

The vectors

$$x_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad x_2 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad x_3 = \begin{pmatrix} 2 \\ 3 \\ 3 \end{pmatrix}$$

are linearly dependent because $0 = 3x_1 - x_2 - x_3$. Any two of x_1, x_2, x_3 form a basis for the space of vectors with the form $(a, b, b)'$. This space has rank 2.

Definition A.11. The (Euclidean) *inner product* between two vectors x and y in $\mathbf{R}^n$ is $x'y$. Two vectors x and y are *orthogonal* (written $x \perp y$) if $x'y = 0$. Two subspaces $\mathscr{N}_1$ and $\mathscr{N}_2$ are orthogonal if $x \in \mathscr{N}_1$ and $y \in \mathscr{N}_2$ imply that $x'y = 0$. $\{x_1, \ldots, x_r\}$ is an *orthogonal basis* for a space $\mathscr{N}$ if $\{x_1, \ldots, x_r\}$ is a basis for $\mathscr{N}$ and for $i \neq j$, $x_i'x_j = 0$. $\{x_1, \ldots, x_r\}$ is an *orthonormal basis* for $\mathscr{N}$ if $\{x_1, \ldots, x_r\}$ is an orthogonal basis and $x_i'x_i = 1$ for $i = 1, \ldots, r$. *The terms orthogonal and perpendicular are used interchangeably*. The *length* of a vector x is $\|x\| \equiv \sqrt{x'x}$. The *distance* between two vectors x and y is the length of their difference, i.e., $\|x - y\|$.

The lengths of the vectors given earlier are

$$\|x_1\| = \sqrt{1^2 + 1^2 + 1^2} = \sqrt{3}, \quad \|x_2\| = 1, \quad \|x_3\| = \sqrt{22} \doteq 4.7.$$

Also, if $x = (2, 1)'$, its length is $\|x\| = \sqrt{2^2 + 1^2} = \sqrt{5}$. If $y = (3, 2)'$, the distance between x and y is the length of $x - y = (2, 1)' - (3, 2)' = (-1, -1)'$, which is $\|x - y\| = \sqrt{(-1)^2 + (-1)^2} = \sqrt{2}$.

Just prior to Section B.4 and in Sections 2.7 and 6.3 we discuss more general versions of the concepts of inner product and length. In particular, a more general version of Definition A.11 is given in Subsection 6.3.5. The remaining results and definitions in this appendix are easily extended to general inner products.

Our emphasis on orthogonality and our need to find orthogonal projection matrices make both the following theorem and its proof fundamental tools in linear model theory:

Theorem A.12. *The Gram–Schmidt Theorem.*
Let $\mathscr{N}$ be a space with basis $\{x_1, \ldots, x_r\}$. There exists an orthonormal basis for $\mathscr{N}$, say $\{y_1, \ldots, y_r\}$, with y_s in the space spanned by $x_1, \ldots, x_s$, $s = 1, \ldots, r$.

PROOF. Define the y_is inductively:

$$\begin{aligned} y_1 &= x_1 \Big/ \sqrt{x_1' x_1}\,, \\ w_s &= x_s - \sum_{i=1}^{s-1} (x_s' y_i) y_i\,, \\ y_s &= w_s \Big/ \sqrt{w_s' w_s}\,. \end{aligned}$$

See Exercise A.1. $\square$

The vectors

$$x_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad x_2 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

are a basis for the space of vectors with the form $(a, b, b)'$. To orthonormalize this basis, take $y_1 = x_1/\sqrt{3}$. Then take

$$w_2 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} - \frac{1}{\sqrt{3}} \begin{pmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{pmatrix} = \begin{pmatrix} 2/3 \\ -1/3 \\ -1/3 \end{pmatrix}.$$

Finally, normalize w_2 to give

$$y_2 = w_2 \Big/ \sqrt{6/9} = (2/\sqrt{6}, -1/\sqrt{6}, -1/\sqrt{6})'.$$

Note that another orthonormal basis for this space consists of the vectors

$$z_1 = \begin{pmatrix} 0 \\ 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix}, \quad z_2 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}.$$

Definition A.13. For $\mathscr{N}$ a subspace of $\mathscr{M}$, let $\mathscr{N}_{\mathscr{M}}^{\perp} \equiv \{y \in \mathscr{M} | y \perp \mathscr{N}\}$. $\mathscr{N}_{\mathscr{M}}^{\perp}$ is called the *orthogonal complement* of $\mathscr{N}$ with respect to $\mathscr{M}$. If $\mathscr{M}$ is taken as $\mathbf{R}^n$, then $\mathscr{N}^{\perp} \equiv \mathscr{N}_{\mathscr{M}}^{\perp}$ is simply referred to as the orthogonal complement of $\mathscr{N}$.

Theorem A.14. Let $\mathscr{M}$ be a vector space, and let $\mathscr{N}$ be a subspace of $\mathscr{M}$. The orthogonal complement of $\mathscr{N}$ with respect to $\mathscr{M}$ is a subspace of $\mathscr{M}$; and if $x \in \mathscr{M}$, x can be written uniquely as $x = x_0 + x_1$ with $x_0 \in \mathscr{N}$ and $x_1 \in \mathscr{N}_{\mathscr{M}}^{\perp}$. The ranks of these spaces satisfy the relation $r(\mathscr{M}) = r(\mathscr{N}) + r(\mathscr{N}_{\mathscr{M}}^{\perp})$.

For example, let $\mathscr{M} = \mathbf{R}^3$ and let $\mathscr{N}$ be the space of vectors with the form $(a,b,b)'$. It is not difficult to see that the orthogonal complement of $\mathscr{N}$ consists of vectors of the form $(0,c,-c)'$. Any vector $(x,y,z)'$ can be written uniquely as

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x \\ (y+z)/2 \\ (y+z)/2 \end{pmatrix} + \begin{pmatrix} 0 \\ (y-z)/2 \\ -(y-z)/2 \end{pmatrix}.$$

The space of vectors with form $(a,b,b)'$ has rank 2, and the space $(0,c,-c)'$ has rank 1.

For additional examples, let

$$X_0 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \quad \text{and} \quad X = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix}.$$

In this case,

$$C(X_0)^\perp = C\left(\begin{bmatrix} -1 & 1 \\ 0 & -2 \\ 1 & 1 \end{bmatrix}\right), \quad C(X_0)^\perp_{C(X)} = C\left(\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}\right),$$

and

$$C(X)^\perp = C\left(\begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}\right).$$

PROOF OF THEOREM A.14. It is easily seen that $\mathscr{N}^\perp_{\mathscr{M}}$ is a subspace by checking Theorem A.3. Let $r(\mathscr{M}) = n$ and $r(\mathscr{N}) = r$. Let $v_1, \ldots, v_r$ be a basis for $\mathscr{N}$ and extend this with $w_1, \ldots, w_{n-r}$ to a basis for $\mathscr{M}$. Apply Gram–Schmidt to get $v_1^*, \ldots, v_r^*$, $w_1^*, \ldots, w_{n-r}^*$ an orthonormal basis for $\mathscr{M}$ with $v_1^*, \ldots, v_r^*$ an orthonormal basis for $\mathscr{N}$.

If $x \in \mathscr{M}$, then

$$x = \sum_{i=1}^{r} \alpha_i v_i^* + \sum_{j=1}^{n-r} \beta_j w_j^*.$$

Let $x_0 = \sum_{i=1}^{r} \alpha_i v_i^*$ and $x_1 = \sum_{j=1}^{n-r} \beta_j w_j^*$. Then $x_0 \in \mathscr{N}$, $x_1 \in \mathscr{N}^\perp_{\mathscr{M}}$, and $x = x_0 + x_1$.

To establish the uniqueness of the representation and the rank relationship, we need to establish that $\{w_1^*, \ldots, w_{n-r}^*\}$ is a basis for $\mathscr{N}^\perp_{\mathscr{M}}$. Since, by construction, the w_j^*s are linearly independent and $w_j^* \in \mathscr{N}^\perp_{\mathscr{M}}$, $j = 1, \ldots, n-r$, it suffices to show that $\{w_1^*, \ldots, w_{n-r}^*\}$ is a spanning set for $\mathscr{N}^\perp_{\mathscr{M}}$. If $x \in \mathscr{N}^\perp_{\mathscr{M}}$, write

$$x = \sum_{i=1}^{r} \alpha_i v_i^* + \sum_{j=1}^{n-r} \beta_j w_j^*.$$

However, since $x \in \mathscr{N}^\perp_{\mathscr{M}}$ and $v_k^* \in \mathscr{N}$ for $k = 1, \ldots, r$,

$$
\begin{aligned}
0 = x'v_k^* &= \left(\sum_{i=1}^{r} \alpha_i v_i^* + \sum_{j=1}^{n-r} \beta_j w_j^*\right)' v_k^* \\
&= \sum_{i=1}^{r} \alpha_i v_i^{*\prime} v_k^* + \sum_{j=1}^{n-r} \beta_j w_j^{*\prime} v_k^* \\
&= \alpha_k v_k^{*\prime} v_k^* = \alpha_k
\end{aligned}
$$

for $k = 1, \ldots, r$. Thus $x = \sum_{j=1}^{n-r} \beta_j w_j^*$, implying that $\{w_1^*, \ldots, w_{n-r}^*\}$ is a spanning set and a basis for $\mathscr{N}_{\mathscr{M}}^{\perp}$.

To establish uniqueness, let $x = y_0 + y_1$ with $y_0 \in \mathscr{N}$ and $y_1 \in \mathscr{N}_{\mathscr{M}}^{\perp}$. Then $y_0 = \sum_{i=1}^{r} \gamma_i v_i^*$ and $y_1 = \sum_{j=1}^{n-r} \delta_j w_j^*$; so $x = \sum_{i=1}^{r} \gamma_i v_i^* + \sum_{j=1}^{n-r} \delta_j w_j^*$. By the uniqueness of the representation of x under any basis, $\gamma_i = \alpha_i$ and $\beta_j = \delta_j$ for all i and j; thus $x_0 = y_0$ and $x_1 = y_1$.

Since a basis has been found for each of $\mathscr{M}$, $\mathscr{N}$, and $\mathscr{N}_{\mathscr{M}}^{\perp}$, we have $r(\mathscr{M}) = n$, $r(\mathscr{N}) = r$, and $r(\mathscr{N}_{\mathscr{M}}^{\perp}) = n - r$. Thus, $r(\mathscr{M}) = r(\mathscr{N}) + r(\mathscr{N}_{\mathscr{M}}^{\perp})$. □

Definition A.15. Let $\mathscr{N}_1$ and $\mathscr{N}_2$ be vector subspaces. Then the sum of $\mathscr{N}_1$ and $\mathscr{N}_2$ is $\mathscr{N}_1 + \mathscr{N}_2 = \{x | x = x_1 + x_2, x_1 \in \mathscr{N}_1, x_2 \in \mathscr{N}_2\}$.

Theorem A.16. $\mathscr{N}_1 + \mathscr{N}_2$ is a vector space and $C(A, B) = C(A) + C(B)$.

Exercises

Exercise A.1 Give a detailed proof of the Gram–Schmidt theorem.

Questions A.2 through A.13 involve the following matrices:

$$
A = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix}, \quad
B = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad
D = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 2 & 5 \\ 0 & 0 \end{bmatrix}, \quad
E = \begin{bmatrix} 1 & 2 \\ 1 & 2 \\ 2 & 7 \\ 0 & 0 \end{bmatrix},
$$

$$
F = \begin{bmatrix} 1 & 5 & 6 \\ 1 & 5 & 6 \\ 0 & 7 & 2 \\ 0 & 0 & 9 \end{bmatrix}, \quad
G = \begin{bmatrix} 1 & 0 & 5 & 2 \\ 1 & 0 & 5 & 2 \\ 2 & 5 & 7 & 9 \\ 0 & 0 & 0 & 3 \end{bmatrix}, \quad
H = \begin{bmatrix} 1 & 0 & 2 & 2 & 6 \\ 1 & 0 & 2 & 2 & 6 \\ 7 & 9 & 3 & 9 & -1 \\ 0 & 0 & 0 & 3 & -7 \end{bmatrix},
$$

$$K = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}, \quad L = \begin{bmatrix} 2 & 0 & 0 \\ 2 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}, \quad N = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}.$$

Exercise A.2 Is the space spanned by the columns of A the same as the space spanned by the columns of B? How about the spaces spanned by the columns of K, L, F, D, and G?

Exercise A.3 Give a matrix whose column space contains $C(A)$.

Exercise A.4 Give two matrices whose column spaces contain $C(B)$.

Exercise A.5 Which of the following equalities are valid: $C(A) = C(A,D)$, $C(D) = C(A,B)$, $C(A,N) = C(A)$, $C(N) = C(A)$, $C(A) = C(F)$, $C(A) = C(G)$, $C(A) = C(H)$, $C(A) = C(D)$?

Exercise A.6 Which of the following matrices have linearly independent columns: A, B, D, N, F, H, G?

Exercise A.7 Give a basis for the space spanned by the columns of each of the following matrices: A, B, D, N, F, H, G.

Exercise A.8 Give the ranks of A, B, D, E, F, G, H, K, L, N.

Exercise A.9 Which of the following matrices have columns that are mutually orthogonal: B, A, D?

Exercise A.10 Give an orthogonal basis for the space spanned by the columns of each of the following matrices: A, D, N, K, H, G.

Exercise A.11 Find $C(A)^{\perp}$ and $C(B)^{\perp}$ (with respect to $\mathbf{R}^4$).

Exercise A.12 Find two linearly independent vectors in the orthogonal complement of $C(D)$ (with respect to $\mathbf{R}^4$).

Exercise A.13 Find a vector in the orthogonal complement of $C(D)$ with respect to $C(A)$.

Exercise A.14 Find an orthogonal basis for the space spanned by the columns of

$$X = \begin{bmatrix} 1 & 1 & 4 \\ 1 & 2 & 1 \\ 1 & 3 & 0 \\ 1 & 4 & 0 \\ 1 & 5 & 1 \\ 1 & 6 & 4 \end{bmatrix}.$$

Exercise A.15 For X as above, find two linearly independent vectors in the orthogonal complement of $C(X)$ (with respect to $\mathbf{R}^6$).

Exercise A.16 Let X be an $n \times p$ matrix. Prove or disprove the following statement: Every vector in $\mathbf{R}^n$ is in either $C(X)$ or $C(X)^\perp$ or both.

Exercise A.17 For any matrix A, prove that $C(A)$ and the null space of A' are orthogonal complements. Note: The null space is defined in Definition B.11.