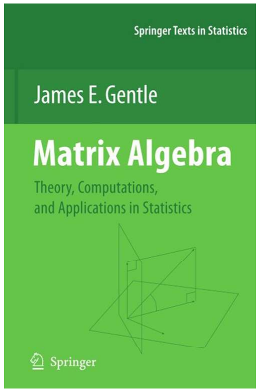
Theory of Matrices

Chapter 2. Vectors and Vector Spaces

2.1. Operations on Vectors—Proofs of Theorems



Theorem 2.1.1(A1)

Theorem 2.1.1. Properties of Vector Algebra in $\mathbb{R}^n$.
Let $x, y, z \in \mathbb{R}^n$. Then:
A1. $(x + y) + z = x + (y + z)$ (Associativity of Vector Addition)

Proof. Let $x, y, z \in \mathbb{R}^n$ be $x = [x_1, x_2, \ldots, x_n]$, $y = [y_1, y_2, \ldots, y_n]$, and $z = [z_1, z_2, \ldots, z_n]$. Then:

$$\begin{aligned}(x+y)+z &= ([x_1, x_2, \ldots, x_n] + [y_1, y_2, \ldots, y_n]) + [z_1, z_2, \ldots, z_n] \\ &= [x_1 + y_1, x_2 + y_2, \ldots x_n + y_n] + [z_1, z_2, \ldots, z_n] \\ &\quad \text{by the definition of vector addition} \\ &= [(x_1 + y_1) + z_1, (x_2 + y_2) + z_2, \ldots (x_n + y_n) + z_n] \\ &\quad \text{by the definition of vector addition} \\ &= [x_1 + (y_1 + z_1), x_2 + (y_2 + z_2), \ldots x_n + (y_n + z_n)] \\ &\quad \text{since addition in } \mathbb{R} \text{ is associative}\end{aligned}$$

Theorem 2.1.1(A1) (continued)

Theorem 2.1.1. Properties of Vector Algebra in $\mathbb{R}^n$.
Let $x, y, z \in \mathbb{R}^n$. Then:
A1. $(x + y) + z = x + (y + z)$ (Associativity of Vector Addition)

Proof (continued).

$$\begin{aligned}(x+y)+z &= [x_1 + (y_1 + z_1), x_2 + (y_2 + z_2), \ldots x_n + (y_n + z_n)] \\ &\quad \text{since addition in } \mathbb{R} \text{ is associative} \\ &= [x_1, x_2, \ldots, x_n] + [y_1 + z_1, y_2 + z_2, \ldots, y_n + z_n] \\ &\quad \text{by the definition of vector addition} \\ &= [x_1, x_2, \ldots, x_n] + ([y_1, y_2, \ldots, y_n] + [z_1, z_2, \ldots, z_n]) \\ &\quad \text{by the definition of vector addition} \\ &= x + (y + z).\end{aligned}$$

□

Theorem 2.1.2

Theorem 2.1.2. Let V_1 and V_2 be vector spaces of n-vectors. Then $V_1 \cap V_2$ is a vector space.

Proof. By our definition of "vector space," we only need to prove that $V_1 \cap V_2$ is closed under linear combinations. Let $x, y \in V_1 \cap V_2$ and $a, b \in \mathbb{R}$. Since V_1 is a vector space then it is closed under linear combinations and so $ax + by \in V_1$. Similarly, $ax + by \in V_2$. So $ax + by \in V_1 \cap V_2$. Since x and y are arbitrary elements of $V_1 \cap V_2$ and $a, b \in \mathbb{R}$ are arbitrary scalars, then $V_1 \cap V_2$ is closed under linear combinations. That is, $V_1 \cap V_2$ is a vector space. □

Theorem 2.1.3

Theorem 2.1.3. If V_1 and V_2 are vector spaces of n-vectors, then $V_1 + V_2$ is a vector space.

Proof. By our definition of "vector space," we must show that $V_1 + V_2$ is closed under linear combinations. Let $x, y \in V_1 + V_2$ and let $a, b \in \mathbb{R}$. Then $x = x_1 + x_2$ and $y = y_1 + y_2$ for some $x_1, y_1 \in V_1$ and $x_2, y_2 \in V_2$. Since V_1 and V_2 are vector spaces, then they are closed under linear combinations and so $ax_1 + by_1 \in V_1$ and $ax_2 + by_2 \in V_2$. Therefore

$$ax + by = a(x_1 + x_2) + b(y_1 + y_2) = (ax_1 + by_1) + (ax_2 + by_2) \in V_1 + V_2.$$

Since x and y are arbitrary vectors in $V_1 + V_2$ and $a, b \in \mathbb{R}$ are arbitrary scalars, then we have that $V_1 + V_2$ is closed under linear combinations. That is, $V_1 + V_2$ is a vector space. □

Theorem 2.1.4

Theorem 2.1.4. If vector spaces V_1 and V_2 are essentially disjoint then every element of $V_1 \oplus V_2$ can be written as $v_1 + v_2$, where $v_1 \in V_1$ and $v_2 \in V_2$, in a unique way.

Proof. Let V_1 and V_2 be essentially disjoint vector spaces of n-vectors; that is, $V_1 \cap V_2 = \{0\}$. Suppose some $v \in V_1 \oplus V_2$ is of the form $v = v_1 + v_2 = v_1' + v_2'$ where $v_1, v_2' \in V_1$ and $v_2, v_2' \in V_2$. Then $v_1 - v_1' = v_2' - v_2$. So $v_1 - v_1' \in V_1$ and $v_2' - v_2 \in V_2$ since V_1 and V_2 are vector space. But then $v_1 - v_1', v_2' - v_2 \in V_1 \cap V_2$ and so $v_1 - v_1' = 0$ and $v_2' - v_2 = 0$. That is, $v_1 = v_1'$ and $v_2 = v_2'$. So the representation of $v \in V_1 \oplus V_2$ as a sum of an element of V_1 and an element of V_2 is unique, as claimed. □

Theorem 2.1.5

Theorem 2.1.5. If $\{v_1, v_2, \ldots, v_k\}$ is a basis for a vector space V, then each element can be uniquely represented as a linear combination of the basis vectors.

Proof. Suppose that $x = b_1v_1 + b_2v_2 + \cdots + b_kv_k = c_1v_1 + c_2v_2 + \cdots + c_kv_k$. Then

$$0 = x - x = (b_1v_1 + b_2v_2 + \cdots + b_kv_k) - (c_1v_1 + c_2v_2 + \cdots + c_kv_k)$$

$$= (b_1 - c_1)v_1 + (b_2 - c_2)v_2 + \cdots + (b_k - c_k)v_k.$$

Since the basis consists (by definition) of a linearly independent set of vectors, then $b_1 - c_1 = b_2 - c_2 = \cdots b_k - c_k = 0$; that is, $b_1 = c_1, b_2 = c_2, \ldots, b_k = c_k$. Therefore, the representation of x is unique. Since x is an arbitrary vector in V, the claim follows. □

Theorem 2.1.6(3)

Theorem 2.1.6. Properties of Inner Products.
Let $x, y \in \mathbb{R}^n$ and let $a \in \mathbb{R}$. Then:
3. $a\langle x, y\rangle = \langle ax, y\rangle$ (Factoring of Scalar Multiplication in Inner Products).
Proof. Let $x, y \in \mathbb{R}^n$ be $x = [x_1, x_2, \ldots, x_n]$ and $y = [y_1, y_2, \ldots, y_n]$. Then

$$\begin{aligned} a\langle x, y\rangle &= a\langle [x_1, x_2, \ldots, x_n], [y_1, y_2, \ldots, y_n]\rangle \\ &= a\,(x_1y_1 + x_2y_2 + \cdots + x_ny_n) \text{ by the definition of } \langle x, y\rangle \\ &= a(x_1y_1) + a(x_2y_2) + \cdots + a(x_ny_n) \\ &\quad \text{by distribution property of multiplication over addition in } \mathbb{R} \\ &= (ax_1)y_1 + (ax_2)y_2 + \cdots + (ax_n)y_n \\ &\quad \text{by associativity for multiplication in } \mathbb{R} \\ &= \langle [ax_1, ax_2, \ldots, ax_n], [y_1, y_2, \ldots, y_n]\rangle \\ &\quad \text{by the definition of inner product} \\ &= \langle ax, y\rangle. \quad \square \end{aligned}$$

Theorem 2.1.7

Theorem 2.1.7. Schwarz Inequality.
For any $x, y \in \mathbb{R}^n$ we have $|\langle x, y\rangle| \leq \langle x, x\rangle^{1/2}\langle y, y\rangle^{1/2}$.

Proof. Let $t \in \mathbb{R}$. Then

$$\begin{aligned} 0 &\leq \langle (tx+y),(tx+y)\rangle \text{ by Theorem 2.1.6(1)} \\ &= t\langle x, tx+y\rangle + \langle y, tx+y\rangle \text{ by linearity in the 1st entry} \\ &= t(t\langle x,x\rangle + \langle x,y\rangle) + (t\langle y,x\rangle + \langle y,y\rangle) \text{ by linearity in the 2nd entry} \\ &= t^2\langle x,x\rangle + 2t\langle x,y\rangle + \langle y,y\rangle \text{ by Theorem 1.2.6(2)} \\ &= at^2 + bt + c \end{aligned}$$

where $a = \langle x,x\rangle$, $b = 2\langle x,y\rangle$, and $c = \langle y,y\rangle$. As a quadratic in t, at^2+bt+c cannot have two distinct roots or else we would have $at^2+bt+c<0$ for some t. This means that the discriminant b^2-4ac in the quadratic equation $t = (-b \pm \sqrt{b^2-4ac})/(2a)$, must be $b^2-4ac \leq 0$; that is, $(b/2)^2 \leq ac$. Hence, we have $(b/2)^2 = \langle x,y\rangle^2 \leq ac$ $= \langle x,x\rangle\langle y,y\rangle$ or $\sqrt{\langle x,y\rangle^2} = |\langle x,y\rangle| \leq \langle x,x\rangle^{1/2}\langle y,y\rangle^{1/2}$. □

Theorem 2.1.8

Theorem 2.1.8 The basis norm is indeed a norm for any basis $\{v_1, v_2, \ldots, v_k\}$ of vector space V.

Proof. Let $x = c_1v_1 + c_2v_2 + \cdots + c_kv_k$ and $y = d_1v_1 + d_2v_2 + \cdots + d_kv_k$. If $x \neq 0$ then some $c_i \neq 0$ and so $\rho(x) > 0$. Clearly $\rho(0) = 0$. So "Nonnegativity and Mapping of the Identity" is satisfied.
Next

$$\rho(ax) = \rho(a(c_1v_1+c_2v_2+\cdots+c_kv_k)) = \rho((ac_1)v_1+(ac_2)v_2+\cdots+(ac_k)v_k)$$

$$= \left\{\sum_{j=1}^{k}(ac_j)^2\right\}^{1/2} = |a|\left\{\sum_{j=1}^{k}c_j^2\right\}^{1/2} = |a|\rho(x)$$

and "Relation of Scalar Multiplication to Real Multiplication" holds.

Theorem 2.1.8 (continued 1)

Proof (continued). Finally,

$$\begin{aligned} \rho(x+y)^2 &= \rho((c_1+d_1)v_1 + (c_2+d_2)v_2 + \cdots + (c_k+d_k)v_k)^2 \\ &= \sum_{j=1}^{k}(c_j+d_j)^2 = \sum_{j=1}^{k}(c_j^2 + 2c_jd_j + d_j^2) \\ &= \sum_{j=1}^{k}c_j^2 + 2\sum_{j=1}^{k}c_jd_j + \sum_{j=1}^{k}d_j^2 \\ &\leq \sum_{j=1}^{k}c_j^2 + 2\left\{\sum_{j=1}^{k}c_j^2\right\}^{1/2}\left\{\sum_{j=1}^{k}d_j^2\right\}^{1/2} + \sum_{j=1}^{k}d_j^2 \\ &\quad \text{by Theorem 2.1.7 (Schwarz's Inequality in } \mathbb{R}^n) \\ &= \left(\left\{\sum_{j=1}^{k}c_j^2\right\}^{1/2} + \left\{\sum_{j=1}^{k}d_j^2\right\}^{1/2}\right)^2 = (\rho(x)+\rho(y))^2. \end{aligned}$$

Theorem 2.1.8 (continued 2)

Theorem 2.1.8 The basis norm is indeed a norm for any basis $\{v_1, v_2, \ldots, v_k\}$ of vector space V.

Proof (continued). ...

$$\rho(x+y)^2 = (\rho(x)+\rho(y))^2.$$

Taking square roots, $\rho(x+y) \leq \rho(x) + \rho(y)$ and so the Triangle Inequality holds. Therefore ρ is a metric on V. □

Theorem 2.1.10

Theorem 2.1.10. Every norm on (finite dimensional vector space) V is equivalent to the basis norm ρ for any given basis $\{v_1, v_2, \ldots, v_k\}$. Therefore, any two norms on V are equivalent.

Proof. Let $\|\cdot\|_a$ be any norm on vector space V and let $\{v_1, v_2, \ldots, v_k\}$ be a basis for the space. Then for some unique scalars $c_1, c_2, \ldots, c_k \in \mathbb{R}$ we have $x = c_1 v_1 + c_2 v_2 + \cdots + c_k v_k$. Then, by the Triangle Inequality and "Relation of Scalar Multiplication to Real Multiplication,"

$$\|x\|_a = \left\| \sum_{i=1}^k c_i v_i \right\|_a \leq \sum_{i=1}^k |c_i| \|v_i\|_a.$$

Now with $[|c_1|, |c_2|, \ldots, |c_k|], [\|v_1\|_a, \|v_2\|_a, \ldots, \|v_k\|_a] \in \mathbb{R}^k$ we have by the Schwarz Inequality (Theorem 2.1.7) that

$$\sum_{i=1}^k |c_i| \|v_i\|_a \leq \left\{ \sum_{i=1}^k |c_i|^2 \right\}^{1/2} \left\{ \sum_{i=1}^k \|v_i\|_a^2 \right\}^{1/2}.$$

Theorem 2.1.10 (continued 1)

Proof (continued). Hence

$$\|x\|_a \leq \left\{ \sum_{i=1}^k \|v_i\|_a^2 \right\}^{1/2} \rho(x) = \tilde{s}\rho(x) \text{ for } \tilde{s} = \left\{ \sum_{i=1}^k \|v_i\|_a^2 \right\}^{1/2}.$$

Next, let $C = \left\{ x = \sum_{i=1}^k u_i v_i \in V \,\middle|\, \sum_{i=1}^k |u_i|^2 = 1 \right\}$. Gentle states that set C is "obviously [topologically] closed" (page 20). Set C is the surface of the unit sphere in V under ρ, $C = \{x \in V \mid \rho(x) = 1\}$. We give a proof that C is a topologically closed set by showing that its complement, $V \setminus C$, is open. Let $x \in V \setminus C$ and let $\varepsilon = |1 - \rho(x)| > 0$. Then the open ball $\{v \in V \mid \rho(v - x) < \varepsilon\}$ contains no elements of C: for $y \in C$,

$$\rho(y - x) \geq \begin{cases} \rho(y) - \rho(x) \\ \rho(x) - \rho(y) \end{cases} = \begin{cases} 1 - \rho(x) \\ \rho(x) - 1 \end{cases} = \begin{cases} \varepsilon \text{ if } \rho(x) < 1 \\ \varepsilon \text{ if } \rho(x) > 1. \end{cases}$$

(Notice that the Triangle Inequality for norms implies for any $x, y \in V$ that $\|x\| = \|x - y + y\| \leq \|x - y\| + \|y\|$ or $\|x - y\| \geq \|x\| - \|y\|$.)

Theorem 2.1.10 (continued 2)

Proof (continued). Define $f : C \to \mathbb{R}$ as $f(u) = \left\| \sum_{i=1}^k u_i v_i \right\|_a$. Gentle claims that f is continuous (page 20). Let's prove this. Let $y = \sum_{i=1}^k u_i v_i \in C$ and let $\varepsilon > 0$. Set $\delta = \varepsilon$. For any $x = \sum_{i=1}^k u_i' v_i \in C$ with $\|y - x\|_a < \delta$ we have

$$\varepsilon = \delta > \begin{cases} \|y\|_a - \|x\|_a \\ \|x\|_a - \|y\|_a \end{cases} = \begin{cases} \|\sum_{i=1}^k u_i v_i\|_a - \|\sum_{i=1}^k u_i' v_i\|_a \\ \|\sum_{i=1}^k u_i' v_i\|_a - \|\sum_{i=1}^k u_i v_i\|_a \end{cases}$$

$$= \begin{cases} f(y) - f(x) \\ f(x) - f(y) \end{cases} = \begin{cases} |f(y) - f(x)| \text{ if } f(y) \geq f(x) \\ |f(x) - f(y)| \text{ if } f(y) < f(x). \end{cases}$$

That is, $|f(y) - f(x)| < \varepsilon$. So $f : C \to \mathbb{R}$ is continuous.

Theorem 2.1.10 (continued 3)

Proof (continued). By the Heine-Borel Theorem (since C is closed and bounded and V is finite dimensional), C is compact and so continuous function f attains a minimum value on C, say $f(u_*) \leq f(u)$ for all $u \in C$. Let $\tilde{r} = f(u_*) > 0$. If $x = \sum_{i=1}^k c_i v_i \neq 0$ then

$$\|x\|_a = \left\| \sum_{i=1}^k c_i v_i \right\|_a = \left\{ \sum_{j=1}^k c_j^2 \right\}^{1/2} \left\| \sum_{i=1}^k \left(\frac{c_i}{\left\{ \sum_{j=1}^k c_j^2 \right\}^{1/2}} \right) v_i \right\|_a = \rho(x) f(\tilde{c})$$

where $\tilde{c} = \sum_{i=1}^k \left(c_i \Big/ \left\{ \sum_{j=1}^k c_j^2 \right\}^{1/2} \right) v_i$, so $\tilde{c} \in C$ since

$$\rho(\tilde{c}) = \sum_{i=1}^k \left| \frac{c_i}{\left\{ \sum_{j=1}^k c_j^2 \right\}^{1/2}} \right|^2 = \frac{1}{\sum_{j=1}^k c_j^2} \sum_{i=1}^k c_i^2 = 1.$$

Theorem 2.1.10 (continued 4)

Theorem 2.1.10. Every norm on (finite dimensional vector space) V is equivalent to the basis norm ρ for any given basis $\{v_1, v_2, \ldots, v_k\}$. Therefore, any two norms on V are equivalent.

Proof (continued). Since $\tilde{r} \in C$ then $f(\tilde{c}) \geq \tilde{r}$, and so $\|x\|_a \geq \tilde{r}\rho(x)$ for all $x \in V$, $x \neq 0$. Of course $\|x\|_a \geq \tilde{r}\rho(x)$ for $x = 0$, so for all $x \in V$ we have $\tilde{r}\rho(x) \leq \|x\|_a \leq \tilde{s}\rho(x)$. That is, $\|\cdot\|_a \cong \rho(\cdot)$. Since $\cong$ is an equivalence relation for Theorem 2.1.9, we have that any two norms on V are equivalent. □

Theorem 2.1.11

Theorem 2.1.11. A set of nonzero vectors $\{v_1, v_2, \ldots, v_k\}$ in a vector space with an inner product for which $\langle v_i, v_j\rangle = 0$ for $i \neq j$ (the vectors are said to be *mutually orthogonal*) is a linearly independent set.

Proof. Let $\{v_1, v_2, \ldots, v_k\}$ be a set of mutually orthogonal nonzero vectors. ASSUME the set is not linearly independent. Then $a_1v_1 + a_2v_2 + \cdots + a_iv_i + \cdots a_kv_k = 0$ is satisfied where some coefficient is nonzero, say $a_i \neq 0$. So

$$v_i = (-a_1/a_i)v_1 + (-a_2/a_i)v_2 + \cdots + (-a_{i-1}/a_i)v_{i-1}$$
$$+(-a_{i+1}/a_i)v_{i+1} + \cdots + (-a_k/a_i)v_k.$$

But then

$$\langle v_i, v_i\rangle = \langle(-a_1/a_i)v_1 + (-a_2/a_i)v_2 + \cdots + (-a_{i-1}/a_i)v_{i-1}$$
$$+(-a_{i+1}/a_i)v_{i+1} + \cdots + (-a_k/a_i)v_k, v_i\rangle$$

Theorem 2.1.11 (continued)

Theorem 2.1.11. A set of nonzero vectors $\{v_1, v_2, \ldots, v_k\}$ in a vector space with an inner product for which $\langle v_i, v_j\rangle = 0$ for $i \neq j$ (the vectors are said to be *mutually orthogonal*) is a linearly independent set.

Proof.

$$\langle v_i, v_i\rangle = (-a_1/a_i)\langle v_1, v_i\rangle + (-a_2/a_i)\langle v_2, v_i\rangle + \cdots + (-a_{i-1}/a_i)\langle v_{i-1}, v_i\rangle$$
$$+(-a_{i+1}/a_i)\langle v_{i+1}, v_i\rangle + \cdots + (-a_k/a_i)\langle v_k, v_i\rangle = 0,$$

a CONTRADICTION to the fact that $v_i \neq 0$. So the assumption that the set is not linearly independent is false; that is, the set is linearly independent, as claimed. □