

9.6 Some proofs

Proposition 9.7. *Let $T: V \rightarrow W$ be an \mathbb{F} -linear transformation. Let 0_V and 0_W be the zeros for V and W respectively. Then*

- (a) $T(0_V) = 0_W$, and
 (b) If $v \in V$ then $T(-v) = -T(v)$.

Proof.

- (a) Add $-T(0_V)$ to both sides of the following equation,

$$T(0_V) = T(0_V + 0_V) = T(0_V) + T(0_V).$$

- (b) Since $T(v) + T(-v) = T(v + (-v)) = T(0_V) = 0_W$ and

$$T(-v) + T(v) = T((-v) + v) + T(0_V) = 0_W$$

then $-T(v) = T(-v)$.

□

Proposition 9.8. *Let $T: V \rightarrow W$ be an \mathbb{F} -linear transformation. Then*

- (a) $\ker T$ is a subspace of V .
 (b) $\text{im } T$ is a subspace of W .

Proof. Let 0_V and 0_W be the zeros in V and W , respectively.

- (a) By condition (a) in the definition of linear transformation, T is a group homomorphism.

To show: (aa) If $k_1, k_2 \in \ker T$ then $k_1 + k_2 \in \ker T$.

(ab) $0_V \in \ker T$.

(ac) If $k \in \ker T$ then $-k \in \ker T$.

(ad) If $c \in \mathbb{F}$ and $k \in \ker T$ then $ck \in \ker T$.

- (aa) Assume $k_1, k_2 \in \ker T$.

Then $T(k_1) = 0_W$ and $T(k_2) = 0_W$.

By condition (a) in the definition of a linear transformation,

$$T(k_1 + k_2) = T(k_1) + T(k_2) = 0 + 0 = 0.$$

So $k_1 + k_2 \in \ker T$.

- (ab) By Proposition 9.7(a), $T(0_V) = 0_W$.

So $0_V \in \ker T$.

- (ac) Assume $k \in \ker T$.

By Proposition 9.7(b), $T(-k) = -T(k)$.

So $T(-k) = -T(k) = -0_W = 0_W$, and $-0_W = 0_W$ since $0_W + 0_W = 0_W$.

So $-k \in \ker T$.

- (ad) Assume $c \in \mathbb{F}$ and $k \in \ker T$.

Then, by the definition of linear transformation,

$$T(ck) = cT(k) = c0_W = 0_W, \quad \text{and} \quad c0_W = 0_W,$$

by adding $-c0_W$ to each side of $c0_W + c0_W = c(0_W + 0_W) = c0_W$.

So $T(ck) = 0_W$ and $ck \in \ker T$.

So $\ker T$ is a subspace of V .

(b) By condition (a) in the definition of an \mathbb{F} -linear transformation, T is a group homomorphism.

To show: (ba) If $w_1, w_2 \in \text{im } T$ then $w_1 + w_2 \in \text{im } T$.

(bb) $0_W \in \text{im } T$.

(bc) If $w \in \text{im } T$ then $-w \in \text{im } T$.

(bd) If $c \in \mathbb{F}$ and $w \in \text{im } T$ then $cw \in \text{im } T$.

(ba) Assume $w_1, w_2 \in \text{im } T$.

Then there exist $v_1, v_2 \in V$ such that $T(v_1) = w_1$ and $T(v_2) = w_2$.

By condition (a) in the definition of an \mathbb{F} -linear transformation,

$$T(v_1 + v_2) = T(v_1) + T(v_2) = w_1 + w_2.$$

So $w_1 + w_2 \in \text{im } T$.

(bb) By Proposition 9.7(a), $T(0_V) = 0_W$.

So $0_W \in \text{im } T$.

(bc) Assume $w \in \text{im } T$.

Then there exists $v \in V$ such that $T(v) = w$.

By Proposition 9.7(b), $T(-v) = -T(v) = -w$.

So $-w \in \text{im } T$.

(bd) To show: If $c \in \mathbb{F}$ and $a \in \text{im } T$ then $ca \in \text{im } T$.

Assume $c \in \mathbb{F}$ and $a \in \text{im } T$.

Then there exists $v \in V$ such that $a = T(v)$.

By the definition of an \mathbb{F} -linear transformation,

$$ca = cT(v) = T(cv).$$

So $ca \in \text{im } T$.

So $\text{im } T$ is a subspace of W .

□

Proposition 9.9. *Let $T: V \rightarrow W$ be an \mathbb{F} -linear transformation. Let 0_V be the zero in V . Then*

(a) $\ker T = (0_V)$ if and only if T is injective.

(b) $\text{im } T = W$ if and only if T is surjective.

Proof. Let 0_V and 0_W be the zeros in V and W respectively.

(a) \implies : Assume $\ker T = (0_V)$.

To show: If $T(v_1) = T(v_2)$ then $v_1 = v_2$.

Assume $T(v_1) = T(v_2)$.

Since T is an \mathbb{F} -linear transformation then

$$0_W = T(v_1) - T(v_2) = T(v_1 - v_2).$$

So $v_1 - v_2 \in \ker T$.

Since $\ker T = (0_V)$ then $v_1 - v_2 = 0_V$.

So $v_1 = v_2$.

So T is injective.

\Leftarrow : Assume T is injective

To show: (aa) $(0_V) \subseteq \ker T$.

(ab) $\ker T \subseteq (0_V)$.

(aa) Since $T(0_V) = 0_W$ then $0_V \in \ker T$.

So $(0_V) \subseteq \ker T$.

(ab) Let $k \in \ker T$.

Then $T(k) = 0_W$.

So $T(k) = T(0_V)$.

Thus, since T is injective then $k = 0_V$.

So $\ker T \subseteq (0_V)$.

So $\ker T = (0_V)$.

(b) \Rightarrow : Assume $\text{im } T = W$.

To show: If $w \in W$ then there exists $v \in V$ such that $T(v) = w$.

Assume $w \in W$.

Then $w \in \text{im } T$.

So there exists $v \in V$ such that $T(v) = w$.

So T is surjective.

\Leftarrow : Assume T is surjective.

To show: (ba) $\text{im } T \subseteq W$.

(bb) $W \subseteq \text{im } T$.

(ba) Let $x \in \text{im } T$.

Then there exists $v \in V$ such that $x = T(v)$.

By the definition of T , $T(v) \in W$.

So $x \in W$.

So $\text{im } T \subseteq W$.

(bb) Assume $x \in W$.

Since T is surjective there exists $v \in V$ such that $T(v) = x$.

So $x \in \text{im } T$.

So $W \subseteq \text{im } T$.

So $\text{im } T = W$.

□

Proposition 9.10. *Let V be an \mathbb{F} -vector space and let B be a subset of V . The following are equivalent:*

- (a) B is a basis of V .
- (b) B is a minimal element of $\{S \subseteq V \mid \text{span}_{\mathbb{F}}(S) = V\}$.
- (c) B is a maximal element of $\{L \subseteq V \mid L \text{ is linearly independent}\}$.

(In (b) and (c) the ordering is by inclusion.)

Proof.

(b) \Rightarrow (a): Let $S \subseteq V$ such that $\text{span}_{\mathbb{F}}(S) = V$.

To show: If S is minimal such that $\text{span}_{\mathbb{F}}(S) = V$ then S is a basis.

To show: If S is minimal such that $\text{span}_{\mathbb{F}}(S) = V$ then S is linearly independent.

Proof by contrapositive.

To show: If S is not linearly independent then S is not minimal such that $\text{span}_{\mathbb{F}}(S) = V$.

Assume S is not linearly independent.

To show: There exists $s \in S$ such that $\text{span}_{\mathbb{F}}(S - \{s\}) = V$.

Since S is not linearly independent then there exist $k \in \mathbb{Z}_{>0}$ and $s_1, \dots, s_k \in S$ and $c_1, \dots, c_k \in \mathbb{F}$ and $i \in \{1, \dots, k\}$ such that $c_1 s_1 + \dots + c_k s_k = 0$ and $c_i \neq 0$.

Let $s = s_i$.

Using that \mathbb{F} is a field and $c_i \neq 0$ then

$$\begin{aligned} s &= s_i = c_i^{-1}(c_1 s_1 + \dots + c_{i-1} s_{i-1} + c_{i+1} s_{i+1} + \dots + c_k s_k) \\ &= c_i^{-1} c_1 s_1 + \dots + c_i^{-1} c_{i-1} s_{i-1} + c_i^{-1} c_{i+1} s_{i+1} + \dots + c_i^{-1} c_k s_k. \end{aligned}$$

So $V = \text{span}_{\mathbb{F}}(S) = \text{span}_{\mathbb{F}}(S - \{s\})$.

So S is not minimal such that $\text{span}_{\mathbb{F}}(S) = V$.

(a) \Rightarrow (b): Proof by contrapositive.

To show: If B is not minimal element of $\{S \subseteq V \mid \text{span}_{\mathbb{F}}(S) = V\}$ then B is not a basis of V .

Assume B is not minimal element of $\{S \subseteq V \mid \text{span}_{\mathbb{F}}(S) = V\}$.

So there exists $b \in B$ such that $\text{span}_{\mathbb{F}}(B - \{b\}) \neq V$.

To show: (aa) $B \in \{S \subseteq V \mid \text{span}_{\mathbb{F}}(S) = V\}$.

(ab) If $b \in B$ then $B - \{b\} \notin \{S \subseteq V \mid \text{span}_{\mathbb{F}}(S) = V\}$.

(aa) Since $\text{span}_{\mathbb{F}}(B) = V$ then $B \in \{S \subseteq V \mid \text{span}_{\mathbb{F}}(S) = V\}$.

(ab) Assume $b \in B$.

To show: $B - \{b\} \notin \{S \subseteq V \mid \text{span}_{\mathbb{F}}(S) = V\}$.

To show: $\text{span}_{\mathbb{F}}(B - \{b\}) \neq V$.

Since $\text{span}_{\mathbb{F}}(B) = V$ then there exist $k \in \mathbb{Z}_{>0}$, $b_1, \dots, b_k \in B$ and $c_1, \dots, c_k \in \mathbb{F}$ such that $b = c_1 b_1 + \dots + c_k b_k$.

So $0 = c_1 b_1 + \dots + c_k b_k + (-1)b$.

(a) \Rightarrow (c): Assume B is a basis of V .

Since B is linearly independent then $B \in \{L \subseteq V \mid L \text{ is linearly independent}\}$.

To show: If $v \in V$ and $v \notin B$ then $B \cup \{v\}$ is not linearly independent.

Assume $v \in V$ and $v \notin B$.

Since $\text{span}_{\mathbb{F}}(B) = V$ then there exists $k \in \mathbb{Z}_{>0}$ and $b_1, \dots, b_k \in B$ and $c_1, \dots, c_k \in \mathbb{F}$ such that $v = c_1 b_1 + \dots + c_k b_k$.

So $0 = c_1 b_1 + \dots + c_k b_k + (-1)v$.

So $B \cup \{v\}$ is not linearly independent.

(c) \Rightarrow (a): Assume S is a maximal element of $\{L \subseteq V \mid L \text{ is linearly independent}\}$.

To show: $\text{span}_{\mathbb{F}}(S) = V$.

To show: $V \subseteq \text{span}_{\mathbb{F}}(S)$.

Let $v \in V$.

To show: $v \in \text{span}_{\mathbb{F}}(S)$.

Case 1: $v \in S$. Then $v \in \text{span}_{\mathbb{F}}(S)$.

Case 2: $v \notin S$.

Then $S \cup \{v\}$ is not linearly independent and S is linearly independent.

So there exist $k \in \mathbb{Z}_{>0}$ and $s_1, \dots, s_k \in S$ and $c_0, c_1, \dots, c_k \in \mathbb{F}$ such that

$$c_0 \neq 0 \quad \text{and} \quad c_0 v + c_1 s_1 + \dots + c_k s_k = 0.$$

Since \mathbb{F} is a field and $c_0 \neq 0$ then

$$v = (-c_0^{-1} c_1) s_1 + \dots + (-c_0^{-1} c_k) s_k.$$

So $v \in \text{span}_{\mathbb{F}}(S)$.

So $V \subseteq \text{span}_{\mathbb{F}}(S)$ and $V = \text{span}_{\mathbb{F}}(S)$.

So S is linearly independent and $\text{span}_{\mathbb{F}}(S) = V$.

So S is a basis of V .

□

Theorem 9.11. *Let V be an \mathbb{F} -vector space. Then*

(a) V has a basis, and

(b) Any two bases of V have the same number of elements.

Proof.

(a) The idea is to use Zorn's lemma on the set $\{L \subseteq V \mid L \text{ is linearly independent}\}$, ordered by inclusion. We will not prove Zorn's lemma, we will assume it. Zorn's lemma is equivalent to the axiom of choice. For a proof see Isaacs book [Isa. §11D].

Zorn's Lemma. *If S is a nonempty poset such that every chain in S has an upper bound then S has a maximal element.*

Let $v \in V$ such that $v \neq 0$.

Then $L = \{v\}$ is linearly independent.

So $\{L \subseteq V \mid L \text{ is linearly independent}\}$ is not empty.

To show: If $\cdots \subseteq S_{k-1} \subseteq S_k \subseteq S_{k+1} \subseteq \cdots$ chain of linearly independent subsets of V then there exists a linearly independent set S that contains all the S_k .

Assume $\cdots \subseteq S_{k-1} \subseteq S_k \subseteq S_{k+1} \subseteq \cdots$ is a chain of linearly independent subsets of V .

Let $L = \bigcup_k S_k$.

To show L is linearly independent.

Assume $\ell \in \mathbb{Z}_{>0}$ and $s_1, \dots, s_\ell \in L$.

Then there exists k such that $s_1, \dots, s_\ell \in S_k$.

Since S_k is linearly independent then if $c_1, \dots, c_\ell \in \mathbb{F}$ and $c_1 s_1 + \cdots + c_\ell s_\ell = 0$ then $c_1 = 0, c_2 = 0, \dots, c_\ell = 0$.

So L is linearly independent.

So, if $\cdots \subseteq S_{k-1} \subseteq S_k \subseteq S_{k+1} \subseteq \cdots$ chain of linearly independent subsets of V then there exists a linearly independent set B that contains all the S_k .

Thus, by Zorn's lemma, $\{L \subseteq V \mid L \text{ is linearly independent}\}$ has a maximal element B .

By Proposition 9.3, B is a basis of V .

(b) Let B and C be bases of V .

Case 1: V has a basis B with $\text{Card}(B) < \infty$.

Let $b \in B$.

Then there exists $c \in C$ such that $c \notin \text{span}_{\mathbb{F}}(B - \{b\})$.

Then $B_1 = (B - \{b\}) \cup \{c\}$ is a basis with the same cardinality as B .

Since B is finite then, by repeating this process, we can, after a finite number of steps, create a basis B' of V such that $B' \subseteq C$ and $\text{Card}(B') = \text{Card}(B)$.

Thus $\text{Card}(B) = \text{Card}(B') \leq \text{Card}(C)$.

A similar argument with C in place of B gives that $\text{Card}(B) \geq \text{Card}(C)$.

So $\text{Card}(B) = \text{Card}(C)$.

Case 2: V has an infinite basis B .

Let C be a basis of V .

Define $P_{cb} \in \mathbb{F}$ for $c \in C$ and $b \in B$ by

$$b = \sum_{c \in C} P_{cb} c, \quad \text{and let} \quad S_b = \{c \in C \mid P_{cb} \neq 0\} \quad \text{for } b \in B.$$

If $b \in B$ then S_b is a finite subset of C and

$$C = \bigcup_{b \in B} S_b, \quad \text{since } C \text{ is a minimal spanning set.}$$

So $\text{Card}(C) \leq \max\{\text{Card}(S_b) \mid b \in B\} \leq \aleph_0 \text{Card}(B)$.

A similar argument with B and C switched shows that $\text{Card}(B) \leq \aleph_0 \text{Card}(C)$.

So $\text{Card}(C) \leq \aleph_0 \text{Card}(B) = \text{Card}(B) \leq \aleph_0 \text{Card}(C) = \text{Card}(C)$.

Since $\text{Card}(C) \leq \text{Card}(B) \leq \text{Card}(C)$ then $\text{Card}(C) = \text{Card}(B)$.

□

Proposition 9.12. Let V and W and Z be \mathbb{F} -vector spaces with bases B, C and D , respectively. Let

$$f: V \rightarrow W, \quad g: V \rightarrow W, \quad h: W \rightarrow Z \quad \text{be linear transformations}$$

and let $c \in \mathbb{F}$. Then

$$(cf)_{CB} = c \cdot f_{CB}, \quad f_{CB} + g_{CB} = (f + g)_{CB} \quad \text{and} \quad (h \circ g)_{DB} = h_{DC} g_{CB}.$$

Proof. Let $b \in B$ and $c' \in C$. Taking the coefficient of c' on each side of

$$\sum_{c \in C} (\alpha f)_{CB}(c, b)c = (\alpha f)(b) = \alpha \cdot f(b) = \alpha \cdot \left(\sum_{c \in C} f_{CB}(c, b)c \right) = \sum_{c \in C} \alpha f_{CB}(c, b)c$$

gives $(\alpha f)_{CB}(c', b) = \alpha \cdot f_{CB}(c', b)$.

So $(\alpha f)_{CB} = \alpha \cdot f_{CB}$.

Let $b \in B$ and $c' \in C$. Taking the coefficient of c' on each side of

$$\begin{aligned} \sum_{c \in C} (f + g)_{CB}(c, b)c &= (f + g)(b) = f(b) + g(b) = \sum_{c \in C} (f_{CB}(c, b)c + g_{CB}(c, b)c) \\ &= \sum_{c \in C} (f_{CB}(c, b)c + g_{CB}(c, b)c) = \sum_{c \in C} (f_{CB}(c, b) + g_{CB}(c, b))c \end{aligned}$$

gives $(f_{CB} + g_{CB})(c', b) = f_{CB}(c', b) + g_{CB}(c', b)$.

So $f_{CB} + g_{CB} = (f + g)_{CB}$.

Let $b \in B$ and $d' \in D$. Taking the coefficient of d' on each side of

$$\begin{aligned} \sum_{d \in D} (h \circ g)_{DB}(d, b)d &= (h \circ g)(b) = h(g(b)) = h\left(\sum_{c \in C} g_{CB}(c, b)c \right) \\ &= \sum_{c \in C} g_{CB}(c, b)h(c) = \sum_{c \in C} \sum_{d \in D} g_{CB}(c, b)h_{DC}(d, c)d, \end{aligned}$$

gives $(h \circ g)_{DB}(d', b) = \sum_{c \in C} \sum_{d \in D} h_{DC}(d', c)g_{CB}(c, b) = (h_{DC}g_{CB})(d', b)$.

So $(h \circ g)_{DB} = (h_{DC}g_{CB})$. □

Proposition 9.13. Let $g: V \rightarrow W$ and $f: V \rightarrow V$ be \mathbb{F} -linear transformations. Let

B_1 and B_2 be bases of V , and let C_1 and C_2 be bases of W ,

and let $P_{B_1B_2}$ and $P_{C_2C_1}$ be the change of basis matrices defined as in (9.1). Then

$$g_{C_2B_2} = P_{C_2C_1}g_{C_1B_1}P_{B_1B_2} \quad \text{and} \quad f_{B_2B_2} = P_{B_1B_2}^{-1}f_{B_1B_1}P_{B_1B_2}.$$

Proof. Let $\beta, \beta' \in B_2$. Comparing coefficients of β' on each side of

$$\begin{aligned} \beta &= \sum_{b \in B_1} P_{B_1B_2}(b, \beta)b = \sum_{b \in B_1} P_{B_1B_2}(b, \beta) \sum_{\beta' \in B_2} P_{B_2B_1}(\beta', b)\beta' \\ &= \sum_{b \in B_1} \sum_{\beta' \in B_2} P_{B_2B_1}(\beta', b)P_{B_1B_2}(b, \beta)\beta' = \sum_{b \in B_1} \sum_{\beta' \in B_2} (P_{B_2B_1}P_{B_1B_2})(\beta', \beta)\beta' \end{aligned}$$

gives

$$(P_{B_2B_1}P_{B_1B_2})(\beta', \beta) = \delta_{\beta'\beta}.$$

So $P_{B_2B_1} = P_{B_1B_2}^{-1}$.

Let $\beta \in B_1$ and $c \in B_2$. Taking the coefficient of b' on each side of

$$f(c) = \sum_{c' \in B_2} f_{B_2B_2}(c', c)c' = \sum_{b' \in B_1} f_{B_2B_2}(c', c)P_{B_1B_2}(b', c')b'$$

and

$$f(c) = f\left(\sum_{b \in B_1} P_{B_1 B_2}(b, c)b\right) = \sum_{b \in B_1} P_{B_1 B_2}(b, c)f(b) = \sum_{b \in B_1} P_{B_1 B_2}(b, c) \sum_{b' \in B_1} f_{B_1 B_1}(b', b)b'$$

gives

$$(P_{B_1 B_2} f_{B_2 B_2})(\beta, b) = (f_{B_1 B_1} P_{B_1 B_2})(\beta, b).$$

So

$$P_{B_1 B_2} f_{B_2 B_2} = f_{B_1 B_1} P_{B_1 B_2} \quad \text{and thus} \quad f_{B_2 B_2} = P_{B_1 B_2}^{-1} f_{B_1 B_1} P_{B_1 B_2}.$$

Let $\gamma' \in C_2$ and $\beta \in B_2$. Taking the coefficient of γ on each side of

$$\begin{aligned} \sum_{\gamma \in C_2} g_{C_2 B_2}(\gamma, \beta)\gamma &= g(\beta) = g\left(\sum_{b \in B_1} P_{B_1 B_2}(b, \beta)b\right) = \sum_{b \in B_1} P_{B_1 B_2}(b, \beta)g(b) \\ &= \sum_{b \in B_1} P_{B_1 B_2}(b, \beta) \sum_{c \in C_1} g_{C_1 B_1}(c, b)c \\ &= \sum_{b \in B_1} P_{B_1 B_2}(b, \beta) \sum_{c \in C_1} g_{C_1 B_1}(c, b) \sum_{\gamma \in C_2} P_{C_2 C_1}(\gamma, c)\gamma \\ &= \sum_{b \in B_1, c \in C_1, \gamma \in C_2} P_{C_2 C_1}(\gamma, c)g_{C_1 B_1}(c, b)P_{B_1 B_2}(b, \beta)\gamma \\ &= \sum_{\gamma \in C_2} (P_{C_2 C_1} g_{C_1 B_1} P_{B_1 B_2})(\gamma, \beta)\gamma \end{aligned}$$

gives $g_{C_2 B_2}(\gamma', \beta) = (P_{C_2 C_1} g_{C_1 B_1} P_{B_1 B_2})(\gamma', \beta)$. So $g_{C_2 B_2} = P_{C_2 C_1} g_{C_1 B_1} P_{B_1 B_2}$. □

Proposition 9.14. *Let $P \in M_n(\mathbb{F})$. The matrix P is invertible if and only if the columns of P are linearly independent in \mathbb{F}^n .*

Proof.

\Rightarrow : Assume P is invertible. Let p_1, \dots, p_n be the columns of P .

To show: $\{p_1, \dots, p_n\}$ is linearly independent.

Assume $c_1, \dots, c_n \in \mathbb{F}$ and $c_1 p_1 + \dots + c_n p_n = 0$.

Let $c = (c_1, \dots, c_n)^t \in \mathbb{F}^n$.

Since $c_1 p_1 + \dots + c_n p_n = 0$ then $Pc = 0$.

So $c = P^{-1}Pc = P^{-1}0 = 0$.

So $c_1 = 0, \dots, c_n = 0$.

\Leftarrow : Assume the columns of P are linearly independent.

To show: There exists $Q \in M_n(\mathbb{F})$ such that $QP = 1$.

Let p_1, \dots, p_n be the columns of P .

Since $B = \{p_1, \dots, p_n\}$ is linearly independent and $\dim(\mathbb{F}^n) = n$ then B is a maximal linearly independent set.

Thus, by Theorem 9.3, B is a basis.

Let $S = \{e_1, \dots, e_n\}$ where e_i has 1 in the i th spot and 0 elsewhere.

Then $P = P_{BS}$, the change of basis matrix from S to B .

Let $Q = P_{SB}$, the change of basis matrix from B to S .

Then $QP = P_{SB}P_{BS} = P_{SS} = 1$.

So P is invertible. □