A Survival Guide to

Linear Algebra Done Right

2015 Springer Edition

Sheldon Axler

Partial scrutiny,
Comments, Suggestions and Errata

José Renato Ramos Barbosa

2019

Departamento de Matemática
Universidade Federal do Paraná
Curitiba - Paraná - Brasil
jrrb@ufpr.br

==
==
1
==
==
EX. 20, p. **25**
On the one hand, note that each vector $u \in U$ can be written as $u = xu_1 + yu_2$ with $u_1 = (1,1,0,0)$, $u_2 = (0,0,1,1)$ and $x, y \in \mathbf{F}$. Then, since u_1 and u_2 are linearly independent (because none of them is a scalar multiple of the other), $\{u_1, u_2\}$ is a basis of U.[1] On the other hand, suppose that $w_1 = (0,1,0,0)$ and $w_2 = (0,0,1,0)$ spans W.[2] So w_1 and w_2 are linearly independent since none of them is a scalar multiple of the other. Now consider $B = \{u_1, u_2, w_1, w_2\}$. Therefore, on the one hand, the equality $(x, x, y, y) = (0, a, b, 0)$ implies that $x = y = a = b = 0$ and thus $U \cap W = \{(0,0,0,0)\}$. On the other hand, by considering the matrix whose columns are u_1, u_2, w_1 and w_2 written as column vectors, one has

$$\begin{aligned}\det\begin{pmatrix}1&0&0&0\\1&0&1&0\\0&1&0&1\\0&1&0&0\end{pmatrix} &= -\det\begin{pmatrix}1&0&0&0\\0&1&0&0\\0&1&0&1\\1&0&1&0\end{pmatrix}\\ &= \det\begin{pmatrix}1&0&0&0\\0&1&0&0\\1&0&1&0\\0&1&0&1\end{pmatrix}\\ &= 1\end{aligned}$$

and thus u_1, u_2, w_1 and w_2 are linearly independent, that is, B is a basis of $U + W$. Hence $\dim(U + W) = 4 = \dim \mathbf{F}^4$.
==
EX. 24, p. **26**
On the one hand, if $f \in U_{\mathrm{e}} \cap U_{\mathrm{o}}$, then

$$\begin{aligned}f(x) &= f(-x)\\ &= -f(x)\end{aligned}$$

for each $x \in \mathbf{R}$ and thus $f = 0$. On the other hand, if $f \in \mathbf{R}^{\mathbf{R}}$ is arbitrary, then

$$f(x) = \frac{f(x) + f(-x)}{2} + \frac{f(x) - f(-x)}{2}$$

for each $x \in \mathbf{R}$ with the first summand in U_{e} and the second summand in U_{o}.
==

[1] See CHAPTER 2.
[2] Idem.

==
==
2
==
==
EX. 1, p. **37**
On the one hand,

$$\operatorname{span}(v_1 - v_2, v_2 - v_3, v_3 - v_4, v_4) \subset \operatorname{span}(v_1, v_2, v_3, v_4)$$

since

$$\begin{aligned} v_1 - v_2 &= 1 \cdot v_1 - 1 \cdot v_2 + 0 \cdot v_3 + 0 \cdot v_4 \\ v_2 - v_3 &= 0 \cdot v_1 + 1 \cdot v_2 - 1 \cdot v_3 + 0 \cdot v_4 \\ v_3 - v_4 &= 0 \cdot v_1 + 0 \cdot v_2 + 1 \cdot v_3 - 1 \cdot v_4 \\ v_4 &= 0 \cdot v_1 + 0 \cdot v_2 + 0 \cdot v_3 + 1 \cdot v_4. \end{aligned}$$

On the other hand,

$$\operatorname{span}(v_1, v_2, v_3, v_4) \subset \operatorname{span}(v_1 - v_2, v_2 - v_3, v_3 - v_4, v_4)$$

since

$$\begin{aligned} v_1 &= 1 \cdot (v_1 - v_2) + 1 \cdot (v_2 - v_3) + 1 \cdot (v_3 - v_4) + 1 \cdot v_4 \\ v_2 &= 0 \cdot (v_1 - v_2) + 1 \cdot (v_2 - v_3) + 1 \cdot (v_3 - v_4) + 1 \cdot v_4 \\ v_3 &= 0 \cdot (v_1 - v_2) + 0 \cdot (v_2 - v_3) + 1 \cdot (v_3 - v_4) + 1 \cdot v_4 \\ v_4 &= 0 \cdot (v_1 - v_2) + 0 \cdot (v_2 - v_3) + 0 \cdot (v_3 - v_4) + 1 \cdot v_4. \end{aligned}$$

Therefore

$$\begin{aligned} \operatorname{span}(v_1 - v_2, v_2 - v_3, v_3 - v_4, v_4) &= \operatorname{span}(v_1, v_2, v_3, v_4) \\ &= V. \end{aligned}$$

==
EX. 6, p. **37**
Consider $a_i \in \mathbf{F}$, $i = 1, 2, 3, 4$, such that

$$a_1 (v_1 - v_2) + a_2 (v_2 - v_3) + a_3 (v_3 - v_4) + a_4 v_4 = 0. \tag{1}$$

So

$$a_1 v_1 + (a_2 - a_1) v_2 + (a_3 - a_2) v_3 + (a_4 - a_3) v_4 = 0.$$

Thus, due to the fact that v_1, v_2, v_3, v_4 is linearly independent,

$$\begin{aligned} a_1 &= 0, \\ a_2 - a_1 &= 0, \\ a_3 - a_2 &= 0, \\ a_4 - a_3 &= 0. \end{aligned}$$

Then

$$a_i = 0, \ i = 1, 2, 3, 4. \tag{2}$$

Therefore, since (1) $\Longrightarrow$ (2), $v_1 - v_2, v_2 - v_3, v_3 - v_4, v_4$ is linearly independent.
==
EX. 7, p. **37**
Consider $a_i \in \mathbf{F}$, $i = 1, 2, 3, \ldots, m$, such that

$$a_1 (5v_1 - 4v_2) + a_2 v_2 + a_3 v_3 + \cdots + a_m v_m = 0. \tag{3}$$

So

$$5a_1 v_1 + (a_2 - 4a_1) v_2 + a_3 v_3 + \cdots + a_m v_m = 0.$$

Thus, due to the fact that $v_1, v_2, \ldots, v_m$ is linearly independent,

$$\begin{aligned} 5a_1 &= 0, \\ a_2 - 4a_1 &= 0, \\ a_3 &= 0, \\ &\vdots \\ a_m &= 0. \end{aligned}$$

Then

$$a_i = 0, \; i = 1, 2, 3, \ldots, m. \tag{4}$$

Therefore, since (3) $\Longrightarrow$ (4), $5v_1 - 4v_2, v_2, v_3, \ldots, v_m$ is linearly independent.

==

EX. 6, p. **43**

See **EX. 1** and **EX. 6**, p. **37**.

==

EX. 8, p. **43**

First, let us show that the list of $m + n$ vectors spans V. In order to do that, consider $v \in V$. Since $V = U + W$, there exist $u \in U$ and $w \in W$ for which

$$v = u + w. \tag{5}$$

On the other hand, the bases of U and W give us scalars $\alpha_1, \ldots, \alpha_m, \beta_1, \ldots, \beta_n$ in $\mathbf{F}$ such that

$$u = \alpha_1 u_1 + \cdots + \alpha_m u_m \quad \text{and} \quad w = \beta_1 w_1 + \cdots + \beta_n w_n. \tag{6}$$

Substituting (6) in (5), it folllows that v is a linear combination of the $m + n$ vectors.
Now, let us show that the list of $m + n$ vectors is linearly independent. So, consider

$$\alpha_1 u_1 + \cdots \alpha_m u_m + \beta_1 w_1 + \cdots + \beta_n w_n = 0 \tag{7}$$

with, as before, scalars in $\mathbf{F}$. Denote u and w as in (6). Then $0 = u + w$ and, due to the fact that the sum is direct,[3]

$$u = w = 0. \tag{8}$$

Therefore, substituting (8) in (6) and observing that the list of m vectors in U and the list of n vectors in W are both linearly independent, all of the $m + n$ scalars in (7) must be equal to zero.

==

EX. 16, p. **49**

It is a direct consequence of the following generalization of **EX. 8**, p. **43**:

> If
> $$\mathcal{B}_i = \left\{ u_{i_1}, \ldots, u_{i_{n_i}} \right\}$$
> is a basis of U_i, $i = 1, \ldots, m$, then
> $$\mathcal{B} = \cup_{i=1}^{m} \mathcal{B}_i$$
> is a basis of $U = \oplus_{i=1}^{m} U_i$.

In fact, firstly, let $u \in U$. Then $u = \sum_{i=1}^{m} u_i$ with $u_i \in U_i$, $i = 1, \ldots, m$. So, since u_i is a linear combination of the elements of $\mathcal{B}_i$, $i = 1, \ldots, m$, $\mathcal{B}$ clearly spans U. Now, it remains to show that the elements of $\mathcal{B}$ are linearly independent. Hence we first write

$$c_{1_1} u_{1_1} + \cdots + c_{1_{n_1}} u_{1_{n_1}} + \cdots + c_{m_1} u_{m_1} + \cdots + c_{m_{n_m}} u_{m_{n_m}} = 0$$

with $c_{i_1} u_{i_1} + \cdots + c_{i_{n_i}} u_{i_{n_i}} \in U_i$ and $c_{i_j} \in \mathbf{F}$, $i = 1, \ldots, m$ and $j = 1, \ldots, n_i$. Then $c_{i_1} u_{i_1} + \cdots + c_{i_{n_i}} u_{i_{n_i}} = 0$, $i = 1, \ldots, m$, by 1.44, p. **23**. Therefore, since $\mathcal{B}_i$ is a basis for U_i, $i = 1, \ldots, m$, $c_{i_j} = 0$ for every i and j.

==

[3] See 1.44, p. **23**.

==
==
3
==
==
EX. 3, p. **57**
Let e_k denote the k-th vector of the standard basis of $\mathbf{F}^n$. Substitute

$$T(e_k) = (A_{1,k}, \ldots, A_{m,k})$$

in

$$T\left(\sum_{k=1}^{n} x_k e_k\right) = \sum_{k=1}^{n} x_k T(e_k).$$

==
EX. 4, p. **57**
Consider $a_1, \ldots, a_m \in \mathbf{F}$ such that

$$\sum_{j=1}^{m} a_j v_j = 0.$$

So

$$\begin{aligned}\sum_{j=1}^{m} a_j T(v_j) &= T\left(\sum_{j=1}^{m} a_j v_j\right)\\ &= T(0)\\ &= 0\end{aligned}$$

by the linearity of T and 3.11. Now use the hypothesis that $Tv_1, \ldots, Tv_m$ is linearly independent.
==
EX. 6, p. **67**
$\dim \mathbf{R}^5 = \dim \operatorname{null} T + \dim \operatorname{range} T$ by 3.22, p. **63**, which implies that

$$5 = 2(\dim \operatorname{range} T),$$

contradicting the fact that $\dim \operatorname{range} T$ is a non-negative integer.
==
EX. 9, p. **67**
Consider $a_1, \ldots, a_n \in \mathbf{F}$ such that

$$\sum_{j=1}^{n} a_j T(v_j) = 0.$$

Thus

$$T\left(\sum_{j=1}^{n} a_j v_j\right) = T(0)$$

by the linearity of T and 3.11, p. **57**. Then, since T is injective,

$$\sum_{j=1}^{n} a_j v_j = 0.$$

Now use the hypothesis that $v_1, \ldots, v_n$ is linearly independent.
==
EX. 10, p. **68**
Let $w \in \operatorname{range} T$. Thus there exist $a_1, \ldots, a_n \in \mathbf{F}$ such that

$$\begin{aligned}w &= T\left(\sum_{j=1}^{n} a_j v_j\right)\\ &= \sum_{j=1}^{n} a_j T(v_j)\end{aligned}$$

by the linearity of T.

==

EX. 14, p. **68**
$\dim \mathbf{R}^8 = \dim U + \dim \operatorname{range} T$ by 3.22, p. **63**, which implies that

$$\begin{aligned}\dim \operatorname{range} T &= 8 - 3 \\ &= 5.\end{aligned}$$

Therefore $\operatorname{range} T = \mathbf{R}^5$, which implies that T is surjective.

==

Comment, p. **83**, l. -1
We can use 3.59 to prove 3.61 provided that both $\mathbf{F}^{m,n}$ and $\mathcal{L}(V, W)$ are finite-dimensional. The former satisfies this condition by 3.40, p. **74**. Concerning the latter, use 3.22, p. **63**, with $\mathcal{M}^{-1}$ in place of T. Therefore

$$\operatorname{range} \mathcal{M}^{-1} = \mathcal{L}(V, W)$$

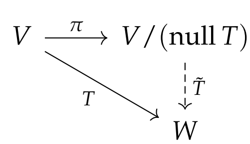
is finite-dimensional.

==

Erratum, p. **85**, ll. 1–9
$\mathcal{M}(v_k)$ should be $\mathcal{M}(Tv_k)$, four times.

==

Comment, p. **97**, 3.90
Note that $\tilde{T} \circ \pi = T$.



==

Comment, p. **98**, 3.91, Proof, (b)
$\operatorname{null} \tilde{T} = 0$ is an abuse of notation. It means $\operatorname{null} \tilde{T} = \{\operatorname{null} T\}$.

==

EX. 15, p. **100**
The condition $\varphi \neq 0$ implies that $\operatorname{range} \varphi = \mathbf{F}$, which is a one-dimensional space. Now use 3.91(d), p. **98**.

==

Erratum, p. **102**, 3.98, Proof, l. -6
F should be $\mathbf{F}$.

==

Comments, p. **112**, 3.117, Proof

- 1st paragraph, 2nd sentence
  Based on (the Proof of) 2.31, pp. **40–41**, let

  $$Tv_1, \ldots, Tv_r, r \leq n,$$

  be a basis of $\operatorname{span}(Tv_1, \ldots, Tv_n)$. Thus

  $$\operatorname{span}(Tv_1, \ldots, Tv_r) = \operatorname{span}(Tv_1, \ldots, Tv_n)$$

  and, if

  $$w = a_1 Tv_1 + \cdots + a_r Tv_r$$

  with $a_1, \ldots, a_r \in \mathbf{F}$, then, by 3.62, p. **84**,

  $$\mathcal{M}(w) = \begin{pmatrix} a_1 \\ \vdots \\ a_r \end{pmatrix}.$$

  Now, clearly,

  $$\operatorname{span}(\mathcal{M}(Tv_1), \ldots, \mathcal{M}(Tv_r)) = \operatorname{span}(\mathcal{M}(Tv_1), \ldots, \mathcal{M}(Tv_n))$$

  and

  $$\operatorname{span}(Tv_1, \ldots, Tv_r) \ni w \mapsto \mathcal{M}(w) \in \mathbf{F}^{r,1}$$

  is an isomorphism.

- 2nd paragraph, 1st sentence
  See **EX. 10**, p. **68**.

==

EX. 32, p. **115**

- (a) $\Longrightarrow$ (c) (and (b))
  By 3.69,[4] T is injective. So, by 3.16,[5] $\operatorname{null} T = \{0\}$. Then, by 3.22,[6] $\dim \operatorname{range} T = n$. Thus, by 3.117 and 3.40,[7] the column rank of $\mathcal{M}(T) = \dim \mathbf{F}^{n,1}$. Therefore, by 3.115 and 3.64,[8] $\mathcal{M}(Tu_1), \ldots, \mathcal{M}(Tu_n)$ is a basis of $\mathbf{F}^{n,1}$.

- (c) $\Longrightarrow$ (b)
  Use 2.42.[9]

- (b) $\Longrightarrow$ (a)
  By 3.115 and 3.117,[10] $\dim \operatorname{range} T = n$. So, by 3.22 and 3.16,[11] T is injective. Then, by 3.69,[12] T is invertible.

- (a) $\Longrightarrow$ (e) $\Longrightarrow$ (d) $\Longrightarrow$ (a)
  It follows from:
  * T is invertible $\Longleftrightarrow$ T' is invertible;[13]
  * $\mathcal{M}(T') = (\mathcal{M}(T))^t$ by 3.114;[14]
  * (a) $\Longrightarrow$ (c) $\Longrightarrow$ (b) $\Longrightarrow$ (a) with T' in the role of T and $\mathcal{M}(T')$ in the role of $\mathcal{M}(T)$.

==

[4] See p. **87**.
[5] See p. **61**.
[6] See p. **63**.
[7] See p. **112** and p. **74**.
[8] See p. **111** and p. **85**.
[9] See p. **46**.
[10] See pp. **111–112**.
[11] See p. **63** and p. **61**.
[12] See p. **87**.
[13] See 3.108 and 3.110, pp. **107–108**.
[14] See p. **110**.

==
==
4
==
==
Erratum, p. **123**, last paragraph, 1st sentence
"..., uses analysis its proof." should be "..., uses analysis **in** its proof.".
==
Erratum, p. **125**, 4.14, Proof, 3rd paragraph, 3rd sentence
"... So we need only show that ..." should be "... So we need only **to** show that ...".
==
EX. 2 and **EX. 3**, p. **129**
The answer to both questions is "no"![15]
==

[15]Consider, for example, the sum of $p(x) = x^2 + x$ and $q(x) = -x^2 + 1$.

==
==
5
==
==
Erratum, p. **134**, 5.6
F should be $\mathbf{F}$.
==
EX. 2, p. **138**

$$\begin{aligned} v \in \operatorname{null} S &\Longrightarrow S(v) = 0 \\ &\Longrightarrow T(S(v)) = T(0) \\ &\Longrightarrow TS(v) = 0 \\ &\Longrightarrow ST(v) = 0 \\ &\Longrightarrow S(T(v)) = 0 \\ &\Longrightarrow T(v) \in \operatorname{null} S. \end{aligned}$$

==
EX. 3, p. **139**

$$\begin{aligned} v \in \operatorname{range} S &\Longrightarrow \exists w \in V \text{ such that } v = S(w) \\ &\Longrightarrow T(v) = T(S(w)) \\ &\Longrightarrow T(v) = TS(w) \\ &\Longrightarrow T(v) = ST(w) \\ &\Longrightarrow T(v) = S(T(w)) \\ &\Longrightarrow T(v) = S(u) \text{ with } u = T(w) \\ &\Longrightarrow \exists u \in V \text{ such that } T(v) = S(u) \\ &\Longrightarrow T(v) \in \operatorname{range} S. \end{aligned}$$

==
EX. 4, p. **139**

$$\begin{aligned} v \in \sum_{i=1}^{m} U_i &\Longrightarrow v = \sum_{i=1}^{m} u_i \text{ with } u_i \in U_i,\, i = 1, \dots, m \\ &\Longrightarrow T(v) = T\left(\sum_{i=1}^{m} u_i\right) \\ &\Longrightarrow T(v) = \sum_{i=1}^{m} T(u_i) \ \text{ with } T(u_i) \in U_i,\, i = 1, \dots, m, \text{ since each } U_i \text{ is invariant under } T \\ &\Longrightarrow T(v) \in \sum_{i=1}^{m} U_i. \end{aligned}$$

==
EX. 33, p. **142**
First, since range T is invariant under T,[16] we can consider $T/(\operatorname{range} T)$.[17] So, for $v \in V$,

$$\begin{aligned} (T/(\operatorname{range} T))(v + \operatorname{range} T) &= T(v) + \operatorname{range} T \\ &= \operatorname{range} T \end{aligned}$$

since $T(v) \in \operatorname{range} T$.
==
Comment, p. **145**, 5.21, Proof, penultimate sentence
If the m operators are injective, the image of a nonzero vector under $T - \lambda_j I$ is obtained for each index j. Therefore

$$\left(\Pi_{j=1}^{n} (T - \lambda_j I)\right) v \neq 0.$$

[16]See 5.3(d), p. **132**.
[17]See 5.14, p. **137**.

==
Comment, p. **149**, Proof 1, l. 7
To be more specific, '... (see 3.69) ...' should be '... (see 5.6) ...'.[18]
==
Comment, p. **150**, Proof 2, sentence that begins with 'Unraveling'
Consider $a_2, \ldots, a_j \in \mathbf{F}$, $j = 2, \ldots, n$, such that

$$\begin{aligned} Tv_j + U &= (T/U)\big(v_j + U\big) \\ &= a_2(v_2 + U) + \cdots + a_j\big(v_j + U\big) \\ &= ((a_2 v_2) + U) + \cdots + \big((a_j v_j) + U\big) \\ &= \big(a_2 v_2 + \cdots + a_j v_j\big) + U, \end{aligned}$$

where the first equality comes from 5.14, p. **137**, and the last two equalities come from 3.86, p. **96**.
Thus, for each $j \in \{2, \ldots, n\}$, since

$$Tv_j - \big(a_2 v_2 + \cdots + a_j v_j\big) \in U,\text{[19]}$$

there exists $a_1 \in \mathbf{F}$ such that

$$Tv_j = a_1 v_1 + a_2 v_2 + \cdots + a_j v_j.$$

==
Comment - Erratum, p. **151**

- 5th paragraph ("To prove the other direction, ...")

$$\begin{aligned} Tv_1 = 0 &\Longrightarrow \text{null}\, T \neq \{0\} \\ &\Longrightarrow T \text{ is not injective (by 3.16, p. 61)} \\ &\Longrightarrow T \text{ is not invertible (by 3.69, p. 87)}, \end{aligned}$$

  which contradicts the assumption that T is invertible.

- 6th paragraph: ("Let $1 < j \leq n$, ...")
  "... T restricted to dim span$\big(v_1, \ldots, v_j\big)$ ..." should be "... T restricted to span$\big(v_1, \ldots, v_j\big)$ ...".

==
Comment, p. **152**, 5.32, Proof, last sentence
Consider 5.6, p. **134**.
==
EX. 3, p. **153**
Suppose that $T - I \neq 0$. So there is a nonzero vector $v \in V$ such that

$$\begin{aligned} w &:= (T - I)(v) \\ &\neq 0. \end{aligned}$$

Then, by 5.20, p. **144**,

$$\begin{aligned} (T + I)(w) &= (T + I)(T - I)(v) \\ &= (T^2 - I)(v) \\ &= 0 \end{aligned}$$

since $T^2 = I$. Therefore $Tw = -w$, which implies that -1 is an eigenvalue of T!
==
EX. 7, p. **153**
By 5.20, p. **144**,

$$\begin{aligned} T^2 - 9I &= (T + 3I)(T - 3I) \\ &= (T - 3I)(T + 3I). \end{aligned}$$

[18] See p. **134**.
[19] See 3.85, p. **95**.

Therefore, by 5.6, p. **134**,

$$\begin{aligned} 9 \text{ is an eigenvalue of } T^2 &\iff T^2 - 9I \text{ is not injective} \\ &\iff (T-3I) \text{ or } (T+3I) \text{ is not injective} \\ &\iff 3 \text{ or } -3 \text{ is an eigenvalue of } T. \end{aligned}$$

==

EX. 14, p. **154**
Define $T \in \mathcal{L}(\mathbf{F}^2)$ by $T(x,y) = (y,x)$. Obviously, $T^2 = I$ and the matrix of T with respect to the standard basis is

$$\mathcal{M}(T) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

==

Comment, p. **155**, sentence immediately preceding 5.37
See 5.6, p. **134**, and 3.16, p. **61**.

==

Erratum, p. **156**, 5.38, Proof, 1st sentence
'$E(\lambda, T)$' should be '$E(\lambda_j, T)$'.

==

EX. 6, p. **160**
Consider 2.39, p. **45**, and $m = \dim V$ in 5.10, p. **136**. Now let λ_i' be an eigenvalue of S with $Sv_i = \lambda_i' v_i$, $i = 1, \ldots, \dim V$. Therefore, for each basis vector v_i,

$$\begin{aligned} STv_i &= S(Tv_i) \\ &= S(\lambda_i v_i) \\ &= \lambda_i S v_i \\ &= \lambda_i \lambda_i' v_i \\ &= \lambda_i' \lambda_i v_i \\ &= \lambda_i' T v_i \\ &= T(\lambda_i' v_i) \\ &= T(Sv_i) \\ &= TSv_i. \end{aligned}$$

==

EX. 8, p. **160**
Suppose $T - 2I$ and $T - 6I$ are not invertible. Thus 2 and 6 are eigenvalues of T by 5.6, p. **134**. So, since $\dim E(8,T) = 4$, $\dim \mathbf{F}^5 = 5$ and

$$\dim E(2,T) + \dim E(6,T) + \dim E(8,T) \leq \dim \mathbf{F}^5$$

by 5.38, p. **156**, it follows that

$$E(2,T) = \{0\} \text{ or } E(6,T) = \{0\},$$

which is a contradiction.[20]

==

EX. 9, p. **161**
0 is in both eigenspaces and

$$\begin{aligned} Tv = \lambda v &\iff T^{-1}(Tv) = T^{-1}(\lambda v) \\ &\iff v = \lambda T^{-1} v \\ &\iff T^{-1} v = \lambda^{-1} v. \end{aligned}$$

==

[20] See the sentence immediately preceding 5.37, p. **155**.

==

==

6

==

==

EX. 13 and **EX. 14**, p. **176**
Let V be an arbitrary inner product space. Consider $u, v \in V$ are nonzero. Thus, by the Cauchy-Schwarz Inequality,[21]

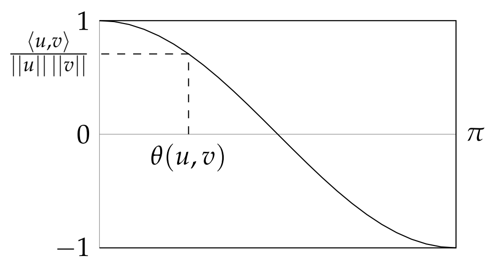
$$-||u||\,||v|| \leq \langle u, v\rangle \leq ||u||\,||v||.$$

Then, since $||u||$ and $||v||$ are positive,

$$-1 \leq \frac{\langle u, v\rangle}{||u||\,||v||} \leq 1.$$

Now let $\theta = \theta(u, v)$ be the unique number in $[0, \pi]$ (measured in radians) such that

$$\cos\theta = \frac{\langle u, v\rangle}{||u||\,||v||}.$$



Therefore $\langle u, v\rangle = ||u||\,||v||\cos\theta$ and $\theta = \arccos\frac{\langle u,v\rangle}{||u||\,||v||}$.

==

EX. 16, p. **176**
Use the Parallelogram Equality.[22]

==

EX. 17, p. **177**
If $x < 0$ and $y < 0$, then $||(x, y)|| < 0$. However, by 6.8 and 6.3,[23] a norm assigns to any $v \in V$ a non-negative real number $||v||$.

==

EX. 30, p. **179**
See 3.69, p. **87**.

==

EX. 2, p. **189**
Consider

$$||v||^2 = \sum_{j=1}^{m} |\langle v, e_j\rangle|^2$$

and

$$u = \sum_{j=1}^{m}\langle v, e_j\rangle e_j.$$

[21] Cf. 6.15, p. **172**.
[22] Cf. p. **174**.
[23] See pp. **168** and **166**.

Thus

$$\begin{aligned}\langle u, v\rangle &= \langle \sum_{j=1}^{m} \langle v, e_j\rangle e_j, v\rangle \\ &= \sum_{j=1}^{m} \langle v, e_j\rangle \langle e_j, v\rangle \\ &= \sum_{j=1}^{m} \langle v, e_j\rangle \overline{\langle v, e_j\rangle} \\ &= \sum_{j=1}^{m} |\langle v, e_j\rangle|^2 \\ &= ||v||^2 \\ &= \langle v, v\rangle.\end{aligned}$$

So $\langle u - v, v\rangle = 0$. Then, by 6.13,[24]

$$||u||^2 = ||u - v||^2 + ||v||^2.$$

Therefore, since $||u||^2 = ||v||^2$ by 6.25,[25]

$$||u - v||^2 = 0,$$

that is $u = v$, which implies that

$$v \in \text{span}(e_1, \ldots, e_m).$$

For the converse, since $e_1, \ldots, e_m$ is orthonormal, it is also linearly independent in $U = \text{span}(e_1, \ldots, e_m)$. Then $e_1, \ldots, e_m$ is an orthonormal basis of U. Therefore, if $v \in U$,

$$||v||^2 = \sum_{j=1}^{m} |\langle v, e_j\rangle|^2$$

by 6.30, p. **182**.

==

Comments, pp. **193–194**

Concerning the example (where U is a line or a plane in $\mathbf{R}^3$), since U is not necessarily a subspace of V by 6.45, it is not necessary for U to contain the zero vector. Similarly, that is why

$$U \cap U^{\perp} \supset \{0\}$$

does not necessarily hold for 6.46(d). By the way, the proof of 6.46(d) for $U \cap U^{\perp} = \emptyset$ is trivial.

==

EX. 1, p. **201**

On the one hand, from $\{v_1, \ldots, v_m\} \subset \text{span}(v_1, \ldots, v_m)$, it follows that

$$(\text{span}(v_1, \ldots, v_r))^{\perp} \subset \{v_1, \ldots, v_m\}^{\perp}$$

by 6.46(e), p. **193**. On the other hand, consider $v \in \{v_1, \ldots, v_m\}^{\perp}$ and let u be a linear combination of $v_1, \ldots, v_m$, that is, consider $c_1, \ldots, c_m \in \mathbf{F}$ such that $u = c_1 v_1 + \ldots + c_m v_m$. So

$$\begin{aligned}\langle v, u\rangle &= \overline{c_1}\langle v, v_1\rangle + \cdots + \overline{c_m}\langle v, v_m\rangle \\ &= \overline{c_1} \cdot 0 + \cdots + \overline{c_m} \cdot 0 \\ &= 0.\end{aligned}$$

Then $v \in (\text{span}(v_1, \ldots, v_m))^{\perp}$. Therefore

$$\{v_1, \ldots, v_m\}^{\perp} \subset (\text{span}(v_1, \ldots, v_m))^{\perp}.$$

==

[24]See p. **170**.

[25]See p. **180**.

EX. 4, p. **201**
For $v_1 = (1, 2, 3, -4)$ and $v_2 = (-5, 4, 3, 2)$, consider

$$u_1 = v_1 \text{ and } u_2 = v_2 - P_{\text{span}(v_1)}(v_2)$$

as in example 6.54, p. **196**. So

$$\left\{ \frac{u_1}{||u_1||}, \frac{u_2}{||u_2||} \right\}$$

is an orthonormal basis of U. On the other hand, since

$$U^\perp = \{v_1, v_2\}^\perp$$

by **EX. 1**, p. **201**, and $\dim U^\perp = 2$ by 6.50, p. **195**, if w_1, w_2 is linearly independent with

$$\langle w_i, v_j \rangle = 0,\, i, j \in \{1, 2\}\,,$$

then $\{w_1, w_2\}$ is a basis of $U^\perp$. So let $w = (a, b, c, d)$ be a vector in $\mathbf{R}^4$ for which

$$\langle w, v_j \rangle = 0,\, j = 1, 2.$$

Then, since

$$\begin{cases} a + 2b + 3c - 4d = 0 \\ -5a + 4b + 3c + 2d = 0 \end{cases}$$

and

$$\begin{cases} a = 15c - 14d \\ b = -9c + 9d \end{cases}$$

are equivalent systems,

$$\begin{aligned} w &= (15c - 14d, -9c + 9d, c, d) \\ &= c(15, -9, 1, 0) + d(-14, 9, 0, 1) \end{aligned}$$

for all real numbers c and d. Therefore, if

$$w_1 = (15, -9, 1, 0) \text{ and } w_2 = (-14, 9, 0, 1) - P_{\text{span}(w_1)}(-14, 9, 0, 1),$$

$$\left\{ \frac{w_1}{||w_1||}, \frac{w_2}{||w_2||} \right\}$$

is an orthonormal basis of $U^\perp$.
==

==
==
7
==
==
Comment, p. **204**
Consider

$$W \ni x \overset{\langle -,w \rangle}{\longmapsto} \langle x, w \rangle.$$

So $\langle -, w \rangle \in \mathcal{L}(W, \mathbf{F})$. Then

$$\varphi = \langle -, w \rangle \circ T \in \mathcal{L}(V, \mathbf{F}). \tag{9}$$

On the other hand, by 6.42,[26] let u be the vector in V such that

$$\varphi = \langle -, u \rangle. \tag{10}$$

Therefore, denoting u by T^*w, T^* is well-defined and, via (9) and (10),

$$\langle Tv, w \rangle = \langle v, T^*w \rangle.$$

==
Comment, p. **213**, 7.21
Since $T - \lambda I$ is normal, another Proof comes from the small box in the bottom left corner of p. **212**. In fact, note that

$$\text{null}\,(T - \lambda I) = \text{null}\,\left(T^* - \overline{\lambda} I\right).$$

==
EX. 2, p. **214**
Firstly, consider the following result

> *If T is invertible, then T^* is invertible and*
> $$(T^*)^{-1} = \left(T^{-1}\right)^*.$$

In fact, by 7.6, items (d) and (e), p. **206**, if

$$\begin{aligned} TT^{-1} &= I \\ &= T^{-1}T, \end{aligned}$$

then

$$\begin{aligned} \left(T^{-1}\right)^* T^* &= I \\ &= T^* \left(T^{-1}\right)^*. \end{aligned}$$

Therefore, by 5.6, p. **134**, and by 7.6, items (a), (b) and (c), p. **206**,

$$\begin{aligned} \lambda \text{ is an eigenvalue of } T &\iff T - \lambda I \text{ is not invertible} \\ &\iff (T - \lambda I)^* = T^* - \overline{\lambda} I \text{ is not invertible} \\ &\iff \overline{\lambda} \text{ is an eigenvalue of } T^*. \end{aligned}$$

==
EX. 4, p. **214**
It is a direct consequence of 7.7, p. **207**.
==
Comments, p. **220**, Proof

[26]See p. **188**.

- $\boxed{m + M \geq 1}$
  In fact, since $V \neq \{0\}$,

$$\begin{aligned} n &= \dim V \\ &\geq 1. \end{aligned}$$

  Therefore $a_0 + a_1 x + \cdots + a_n x^n$ is a nonconstant polynomial.

- $\boxed{m > 0}$
  In fact, if there is no $(T - \lambda_i I)$ factor, then $M > 0$ and the invertible operator

$$c \prod_{j=1}^{M} \left(T^2 + b_j T + c_j I \right)$$

  is the zero vector at $v \neq 0$, which is a violation of 3.69, p. **87**.

==

Comment, p. **222**, Proof, (c) $\Longrightarrow$ (a)
Suppose $\dim V = n$ and let $\mathcal{B}$ be the orthonormal basis mentioned in the 1st paragraph. Then, since $\mathcal{M}(T, \mathcal{B}) \in \mathbf{R}^{n,n}$ is diagonal,

$$\mathcal{M}(T^*, \mathcal{B}) = \mathcal{M}(T, \mathcal{B})$$

by 7.10, p. **208**. Furthermore, from the fact that

$$\begin{array}{rcccc} \mathcal{M}(-, \mathcal{B}) & : & \mathcal{L}(V) & \longrightarrow & \mathbf{R}^{n,n} \\ & & S & \longmapsto & \mathcal{M}(S, \mathcal{B}) \end{array}$$

is an isomorphism,[27] it follows that $T^* = T$.

==

EX. 6, p. **223**
Let T be a normal operator on a complex inner product space and, firstly, suppose that

$$T^* = T.$$

Furthermore, consider that λ is an eigenvalue of T. Then, by 7.21, p. **213**, $\overline{\lambda}$ is an eigenvalue of T and

$$\begin{aligned} E(\lambda, T) &= E\left(\overline{\lambda}, T^*\right) \\ &= E\left(\overline{\lambda}, T\right). \end{aligned}$$

So, by the Complex Spectral Theorem, p. **218**, T is diagonalizable. Then, by 5.41(d), p. **157**, $\lambda = \overline{\lambda}$, that is,

$$\lambda \in \mathbf{R}.$$

> THE PREVIOUS REASONING USES THE NORMALITY HYPOTHESIS. A MORE STRAIGHTFORWARD PROOF FOLLOWS FROM 7.13, P. **210**.

Now, for the converse, suppose that

$$\lambda \text{ is an eigenvalue of } T \Longrightarrow \lambda \in \mathbf{R}.$$

Hence, by the Complex Spectral Theorem, there exists an orthonormal basis $\mathcal{B}$ such that

$$D = \mathcal{M}(T, \mathcal{B})$$

is diagonal (with its (real) eigenvalues lying on the main diagonal). Then, by 7.10, p. **208**,

$$\begin{aligned} \mathcal{M}(T^*, \mathcal{B}) &= \overline{D}^t \\ &= D, \end{aligned}$$

[27] Cf. p. **83**.

which implies that

$$T^* = T.$$

==

EX. 13, p. **224**

Consider the following modifications in the Proof of page **222**:

- In the 1st paragraph, right before the last sentence, insert the sentence: $\boxed{\text{Then } T \text{ is normal.}}$;[28]
- In the 2nd paragraph, replace the word $\boxed{\text{real}}$ with the word $\boxed{\text{complex}}$;
- In the 3rd paragraph, replace:
  - the word $\boxed{\text{self-adjoint}}$ with the word $\boxed{\text{normal}}$, twice;
  - the result $\boxed{7.27}$ with the result $\boxed{5.21}$;
  - the result $\boxed{7.28(c)}$ with the result $\boxed{9.30(d)}$.

==

Comment, p. **226**, Proof, (b) $\Longrightarrow$ (c)

Let $\mathcal{B} = \{e_1, \ldots, e_n\}$. Then, since $\mathcal{M}(R, \mathcal{B}) \in \mathbf{R}^{n,n}$ is diagonal,[29]

$$\mathcal{M}(R^*, \mathcal{B}) = \mathcal{M}(R, \mathcal{B})$$

by 7.10, p. **208**. Furthermore, from the fact that

$$\begin{array}{rcccc} \mathcal{M}(-, \mathcal{B}) & : & \mathcal{L}(V) & \longrightarrow & \mathbf{F}^{n,n} \\ & & S & \longmapsto & \mathcal{M}(S, \mathcal{B}) \end{array}$$

is an isomorphism,[30] it follows that $R^* = R$, that is, R is self-adjoint.
Now, the positivity of R follows from the fact that

$$\langle Re_i, e_j \rangle = \sqrt{\lambda_i} \langle e_i, e_j \rangle$$

is nonnegative for each indices i, j.

==

EX. 4, p. **231**

Firstly, T^*T and TT^* are self-adjoint operators since, by 7.6, (c) and (e).[31]

$$\begin{aligned} (T^*T)^* &= T^* (T^*)^* \\ &= T^*T \end{aligned}$$

and

$$\begin{aligned} (TT^*)^* &= (T^*)^* T^* \\ &= TT^*. \end{aligned}$$

Now, for $v \in V$ and $w \in W$, the positivity follows from

$$\begin{aligned} \langle (T^*T) v, v \rangle &= \langle T^* (Tv), v \rangle \\ &= \langle Tv, Tv \rangle \\ &\geq 0 \end{aligned}$$

and

$$\begin{aligned} \langle (TT^*) w, w \rangle &= \langle T (T^*w), w \rangle \\ &= \langle T^*w, T^*w \rangle \\ &\geq 0. \end{aligned}$$

[28] Every self-adjoint operator is normal!

[29] For each $j \in \{1, \ldots, n\}$, the diagonal entry $\sqrt{\lambda_j}$ is the square root of a nonnegative number. Now see the last paragraph of p. **225**.

[30] Cf. p. **83**.

[31] See p. **206**.

==
EX. 10, p. **232**
By 7.42, p. **229**, (a) $\Longleftrightarrow$ (g). Then one can replace S by S^* in (a), (b), (c) and (d) of 7.42.
==
Comment, p. **233**
Concerning the sentence preceding 7.45, see **EX. 4**, p. **231**.
==
Comment, p. **236**, 7.50, Solution

- 1st sentence

$$\begin{aligned}\left\langle (z_1, z_2, z_3, z_4), T^*(w_1, w_2, w_3, w_4) \right\rangle &= \left\langle T(z_1, z_2, z_3, z_4), (w_1, w_2, w_3, w_4) \right\rangle \\ &= \left\langle (0, 3z_1, 2z_2, -3z_4), (w_1, w_2, w_3, w_4) \right\rangle \\ &= 3z_1\overline{w_2} + 2z_2\overline{w_3} - 3z_4\overline{w_4} \\ &= \left\langle (z_1, z_2, z_3, z_4), (3w_2, 2w_3, 0, -3w_4) \right\rangle\end{aligned}$$

$$\Downarrow$$

$$T^*(w_1, w_2, w_3, w_4) = (3w_2, 2w_3, 0, -3w_4)$$

$$\Downarrow$$

$$\begin{aligned}T^*T(z_1, z_2, z_3, z_4) &= T^*(0, 3z_1, 2z_2, -3z_4) \\ &= (9z_1, 4z_2, 0, 9z_4).\end{aligned}$$

- 2nd sentence
On the one hand, $\sqrt{T^*T}$ denotes the unique positive square root of the (self-adjoint) positive operator T^*T.[32] On the other hand, the operator $\sqrt{T^*T}$ presented in the solution is (self-adjoint)[33] positive since

$$\begin{aligned}\left\langle \sqrt{T^*T}(z_1, z_2, z_3, z_4), (z_1, z_2, z_3, z_4) \right\rangle &= \left\langle (3z_1, 2z_2, 0, 3z_4), (z_1, z_2, z_3, z_4) \right\rangle \\ &= 3z_1\overline{z_1} + 2z_2\overline{z_2} + 3z_4\overline{z_4} \\ &= 3|z_1|^2 + 2|z_2|^2 + 3|z_4|^2 \\ &\geq 0.\end{aligned}$$

Furthermore

$$\begin{aligned}\left(\sqrt{T^*T}\sqrt{T^*T}\right)(z_1, z_2, z_3, z_4) &= \sqrt{T^*T}\left(\sqrt{T^*T}(z_1, z_2, z_3, z_4)\right) \\ &= \sqrt{T^*T}(3z_1, 2z_2, 0, 3z_4) \\ &= (3(3z_1), 2(2z_2), 0, 3(3z_4)) \\ &= T^*T(z_1, z_2, z_3, z_4).\end{aligned}$$

==
Comment, p. **238**, 7.52, Proof
If $\mathcal{B} = \{e_1, \ldots, e_n\}$, then

$$\left(\mathcal{M}\left(\sqrt{T^*T}, \mathcal{B}\right)\right)^2 = \mathcal{M}(T^*T, \mathcal{B})$$

since $\left(\sqrt{T^*T}\right)^2 = T^*T$. Therefore, for each index j,

$$\begin{aligned}\sqrt{T^*T}e_j = \alpha_j e_j &\Longrightarrow \alpha_j^2 = \lambda_j \\ &\Longrightarrow \alpha_j = \sqrt{\lambda_j}.\end{aligned}$$

[32]See p. **233**.
[33]If $\mathcal{B}$ is the standard basis, $\mathcal{M}\left(\sqrt{T^*T}, \mathcal{B}\right)$ is diagonal with real entries!

==
EX. 1, p. **238**
In the exercise,

$$Tv = \langle v, u \rangle x. \tag{11}$$

So, on the one hand,

$$\begin{aligned} \langle v, T^* w \rangle &= \langle Tv, w \rangle \\ &= \langle \langle v, u \rangle x, w \rangle \\ &= \langle v, u \rangle \langle x, w \rangle \\ &= \langle v, \langle w, x \rangle u \rangle \end{aligned}$$

$$\Downarrow$$

$$T^* v = \langle v, x \rangle u \tag{12}$$

$$\Downarrow$$

$$\begin{aligned} (T^* T)\, v &= T^* (Tv) && (13) \\ &= T^* (\langle v, u \rangle x) && (14) \\ &= \langle v, u \rangle T^* x && (15) \\ &= \langle v, u \rangle \langle x, x \rangle u && (16) \\ &= ||x||^2 \langle v, u \rangle u. && (17) \end{aligned}$$

On the other hand, $\sqrt{T^*T}$ denotes the unique positive square root of the (self-adjoint) positive operator T^*T.[34] Furthermore, the operator $\sqrt{T^*T}$ presented in the exercise is self-adjoint,[35] positive (since

$$\begin{aligned} \left\langle \sqrt{T^*T} v, v \right\rangle &= \left\langle \frac{||x||}{||u||} \langle v, u \rangle u, v \right\rangle \\ &= \frac{||x||}{||u||} \langle v, u \rangle \langle u, v \rangle \\ &= \frac{||x||}{||u||} |\langle v, u \rangle|^2 \\ &\geq 0 \end{aligned}$$

for each $v \in V$) and a root of T^*T (since, for each $v \in V$,

$$\begin{aligned} \left(\sqrt{T^*T} \right)^2 v &= \sqrt{T^*T} \left(\sqrt{T^*T} v \right) \\ &= \sqrt{T^*T} \left(\frac{||x||}{||u||} \langle v, u \rangle u \right) \\ &= \frac{||x||}{||u||} \langle v, u \rangle \sqrt{T^*T} u \\ &= \frac{||x||}{||u||} \langle v, u \rangle \frac{||x||}{||u||} \langle u, u \rangle u \\ &= \frac{||x||^2}{||u||^2} \langle v, u \rangle ||u||^2 u \\ &= (T^*T)\, v \end{aligned}$$

by (17)).
==

[34] See p. **233**.
[35] Use (11) and (12)

EX. 3, p. **239**
By 7.45, p. **233**, there exists an isometry $\Sigma \in \mathcal{L}(V)$ such that

$$T^* = \Sigma\sqrt{(T^*)^* T^*} \tag{18}$$
$$= \Sigma\sqrt{TT^*}, \tag{19}$$

where (19) holds by 7.6(c).[36] So

$$\begin{aligned} T &= (T^*)^* \\ &= \left(\Sigma\sqrt{TT^*}\right)^* \\ &= \left(\sqrt{TT^*}\right)^* \Sigma^* \\ &= \sqrt{TT^*}\Sigma^*, \end{aligned}$$

where the first equality holds by 7.6(c), the second equality holds by (19) (as stated above), the third equality holds by 7.6(e),[37] and the last equality holds because each positive operator is self-adjoint by definition. Now use S to denote Σ^*.
==
EX. 17, p. **240**
(a)

$$\begin{aligned} \langle v, T^*w\rangle &= \langle Tv, w\rangle \\ &= \left\langle \sum_{j=1}^{n} s_j\langle v, e_j\rangle f_j, w\right\rangle \\ &= \sum_{j=1}^{n} s_j\langle v, e_j\rangle\langle f_j, w\rangle \\ &= \sum_{j=1}^{n} s_j\overline{\langle w, f_j\rangle}\langle v, e_j\rangle \\ &= \left\langle v, \sum_{j=1}^{n} \overline{s_j}\langle w, f_j\rangle e_j\right\rangle \end{aligned}$$

$$\Downarrow$$

$$T^*w = \sum_{j=1}^{n} s_j\langle w, f_j\rangle e_j$$

since, for each $j \in \{1, \ldots, n\}$, s_j is an eigenvalue of $\sqrt{T^*T}$, which is self-adjoint.[38]
(b)

$$\begin{aligned} (T^*T)\, v &= T^*(Tv) \\ &= T^*\left(\sum_{i=1}^{n} s_i\langle v, e_i\rangle f_i\right) \\ &= \sum_{i=1}^{n} s_i\langle v, e_i\rangle T^* f_i \\ &= \sum_{i=1}^{n} s_i\langle v, e_i\rangle s_i e_i \\ &= \sum_{i=1}^{n} s_i^2\langle v, e_i\rangle e_i. \end{aligned}$$

[36] See p. **206**.
[37] See p. **206**.
[38] Cf. 7.13, p. **210**.

(c) Consider $\mathcal{B} = \{e_1, \ldots, e_n\}$. Then

$$\mathcal{M}(T^*T, \mathcal{B}) = \begin{pmatrix} s_1^2 & & \text{zeros} \\ & \ddots & \\ \text{zeros} & & s_n^2 \end{pmatrix}$$

by (b). Now use 7.52, p. **238**.
(d) Firstly, note that, for each $j \in \{1, \ldots, n\}$, $s_j \neq 0$ if T^{-1} exists. In fact, by 3.43, p. **75**, if $\mathcal{B} = \{e_1, \ldots, e_n\}$ and $\mathcal{B}' = \{f_1, \ldots, f_n\}$, then

$$\begin{aligned} \mathcal{M}\left(T^{-1}, \mathcal{B}', \mathcal{B}\right) \mathcal{M}(T, \mathcal{B}, \mathcal{B}') &= \mathcal{M}\left(T^{-1}T, \mathcal{B}\right) \\ &= \mathcal{M}\left(I, \mathcal{B}\right) \end{aligned}$$

does not hold if some diagonal entry of the matrix $\mathcal{M}(T, \mathcal{B}, \mathcal{B}')$ shown on page **237**, last line, is zero.
Now, consider the operator 'T^{-1}' presented in the exercise. Then

$$\begin{aligned} \left(T^{-1}T\right) v &= T^{-1}(Tv) \\ &= T^{-1}\left(\sum_{j=1}^{n} s_j \langle v, e_j \rangle f_j\right) \\ &= \sum_{j=1}^{n} s_j \langle v, e_j \rangle T^{-1}(f_j) \\ &= \sum_{j=1}^{n} s_j \langle v, e_j \rangle \frac{\langle f_j, f_j \rangle e_j}{s_j} \\ &= \sum_{j=1}^{n} \langle v, e_j \rangle e_j \\ &= v \end{aligned}$$

and

$$\begin{aligned} \left(TT^{-1}\right) v &= T\left(T^{-1}v\right) \\ &= T\left(\sum_{j=1}^{n} \frac{\langle v, f_j \rangle e_j}{s_j}\right) \\ &= \sum_{j=1}^{n} \frac{\langle v, f_j \rangle T(e_j)}{s_j} \\ &= \sum_{j=1}^{n} \frac{\langle v, f_j \rangle s_j f_j}{s_j} \\ &= \sum_{j=1}^{n} \langle v, f_j \rangle f_j \\ &= v. \end{aligned}$$

Therefore $T^{-1}T = I$ and $TT^{-1} = I$.

==
==
8
==
==
Comment, p. **247**, Proof, 2nd paragraph
By 5.20, p. **144**, concerning the operator

$$(T-\lambda_1 I)^k (T-\lambda_2 I)^n \cdots (T-\lambda_m I)^n ,$$

the factors $(T-\lambda_1 I)^k$, $(T-\lambda_2 I)^n$, etc, commute with each other.
==
Comment, p. **248**, Proof
There is some positive integer p such that $N^p = 0$. Therefore

$$\begin{aligned}\text{null}\left(N^{\dim V}\right) &= \text{null}\left(N^p\right)\\ &= V\end{aligned}$$

by 8.2 and 8.4, pp. **242–3**.
==
EX. 4, p. **250**
Firstly, note that

$$E(\alpha, T) \cap G(\beta, T) = \{0\}\,. \tag{20}$$

In fact, consider $v \in E(\alpha, T) \cap G(\beta, T)$. So

$$\begin{aligned}0 &= (T-\beta I)^{\dim V} v\\ &= (T-\beta I)^{\dim V - 1}(T-\beta I)v\\ &= (\alpha-\beta)(T-\beta I)^{\dim V-1} v.\end{aligned}$$

Repeated use of these steps eventually give

$$0 = (\alpha-\beta)^{\dim V} v,$$

and since $\alpha \neq \beta$, we must have $v = 0$. Then (20) holds.
Now consider

$$G(\alpha, T) \cap G(\beta, T) = W.$$

By 8.21, (b), $T|_W \in \mathcal{L}(W)$. As such, if $W \neq \{0\}$, **provided that V is a complex vector space**, $T|_W$ has at least one eigenvalue with an associated eigenvector $v \in W$, which contradicts the fact that $G(\beta, T)$ contains all the eigenvectors of T corresponding to β but no other eigenvectors (by (20)). Therefore $W = \{0\}$.
==
EX. 5, p. **250**
Firstly, note that:

- $v \neq 0$ since
$$T^{m-1} v \neq 0; \tag{21}$$
- $T^m v = 0$ implies that
$$T^{m+j} v = 0,\ j = 0, 1, 2, \ldots. \tag{22}$$

Now consider

$$\sum_{j=0}^{m-1} a_{j+1} T^j v = 0 \tag{23}$$

with $a_{j+1} \in \mathbf{F}$, $j = 0, 1, \ldots, m-1$. Then

$$\begin{aligned}T^{m-1}\left(\sum_{j=0}^{m-1} a_{j+1} T^j v\right) = T^{m-1} 0 &\Longrightarrow \sum_{j=0}^{m-1} a_{j+1} T^{m-1+j} v = 0\\ &\Longrightarrow a_1 T^{m-1} v = 0\\ &\Longrightarrow a_1 = 0\end{aligned}$$

by (22) and (21). Then (23) becomes

$$\sum_{j=1}^{m-1} a_{j+1} T^j v = 0.$$

Therefore

$$\begin{aligned} T^{m-2}\left(\sum_{j=1}^{m-1} a_{j+1} T^j v\right) = T^{m-2}0 &\Longrightarrow \sum_{j=1}^{m-1} a_{j+1} T^{m-2+j} v = 0 \\ &\Longrightarrow a_2 T^{m-1} v = 0 \\ &\Longrightarrow a_2 = 0 \end{aligned}$$

by (22) and (21). Continuing with the same reasoning, we also have

$$a_3 = \cdots = a_m = 0.$$

==

EX. 7, p. **250**

Consider $\lambda \in \mathbf{F}$, $0 \neq v \in V$ and $N^p = 0$ for a positive integer p. Therefore

$$\begin{aligned} Nv = \lambda v &\Longrightarrow N^p v = \lambda^p v \\ &\Longrightarrow \lambda^p v = 0 \\ &\Longrightarrow \lambda = 0. \end{aligned}$$

==

Comment, p. **252**, Proof

By 5.20, p. **144**,

$$p(T)T = Tp(T).$$

==

Comment, p. **253**, Proof, 4th sentence

Firstly, note that $G(\lambda_j, T)$ is invariant under $T - \lambda_j I$ by (b). Therefore

$$(T - \lambda_j I)|_{G(\lambda_j,T)} \in \mathcal{L}(G(\lambda_j, T)).$$

Now, if $v \in G(\lambda_j, T)$, that is, if

$$(T - \lambda_j I)^{\dim V} v = 0,$$

then

$$\begin{aligned} \left((T - \lambda_j I)|_{G(\lambda_j,T)}\right)^{\dim V} v &= (T - \lambda_j I)^{\dim V} v \\ &= 0, \end{aligned}$$

which implies that $(T - \lambda_j I)|_{G(\lambda_j,T)}$ is nilpotent.

==

EX. 1, p. **267**

By 8.36, the characteristic polynomial of T belongs to

$$\left\{(t-3)^2(t-5)(t-8), (t-3)(t-5)^2(t-8), (t-3)(t-5)(t-8)^2\right\}.$$

Now use 8.37 and the fact that, by commutativity,

$$\begin{aligned} (T-3I)^2(T-5I)^2(T-8I)^2 &= (T-5I)(T-8I)\left[(T-3I)^2(T-5I)(T-8I)\right] \\ &= (T-3I)(T-8I)\left[(T-3I)(T-5I)^2(T-8I)\right] \\ &= (T-3I)(T-5I)\left[(T-3I)(T-5I)(T-8I)^2\right]. \end{aligned}$$

==

EX. 2, p. **267**
By 8.36, the characteristic polynomial of T belongs to

$$\left\{(t-5)^{n-i}(t-6)^{i} \;:\; i=1,\ldots,n-1\right\}.$$

Now use 8.37 and the fact that, by commutativity,

$$(T-5I)^{n-1}(T-6I)^{n-1}=(T-5I)^{i-1}(T-6I)^{n-(i+1)}\left[(T-5I)^{n-i}(T-6I)^{i}\right],\; i=1,\ldots,n-1.$$

==

==
==
9
==
==
Erratum, p. **277**, 9.4, Proof, l. -4
'Im_n' should be '$\text{Im}\,\lambda_n$'.
==
Comment, p. **279**, 9.10, Proof, 2nd sentence
Consider $p(t) = \sum_{j=0}^{m} a_j t^j$ with $a_j \in \mathbf{R}$, $j = 0, 1, \ldots, m$, and $u, v \in V$. Therefore

$$\begin{aligned}
\big(p(T)\big)_{\mathbf{C}}(u + iv) &= p(T)u + ip(T)v \\
&= \sum_{j=0}^{m} a_j T^j u + i \sum_{j=0}^{m} a_j T^j v \\
&= \sum_{j=0}^{m} a_j \left(T^j u + iT^j v\right) \\
&= \sum_{j=0}^{m} a_j \left(T_{\mathbf{C}}\right)^j (u + iv) \\
&= p(T_{\mathbf{C}})\,(u + iv).
\end{aligned}$$

==
Erratum, p. **284**, 9.23, Proof, last sentence
'8.36(a)' should be '8.36(b)'.
==
EX. 8, p. **285**
On the one hand, 5 and 7 are eigenvalues of $T_{\mathbf{C}}$ by 9.11. On the other hand, the sum of the multiplicities of all the eigenvalues of $T_{\mathbf{C}}$ equals 3 by 8.26. So there are only two possibilities:

1. 5 and 7 are the eigenvalues of $T_{\mathbf{C}}$, one of multiplicity 1 and one of multiplicity 2;
2. 5, 7 and another scalar λ are the eigenvalues of $T_{\mathbf{C}}$, each with multiplicity 1.

By 9.16, $\lambda \in \mathbf{R}$.
==
EX. 11, p. **286**
Consider $q(t) = t^2 + bt + c$ and let p denote the minimal polynomial of T. So, by 8.46, either $q = p$ or there is $s \in \mathcal{P}(\mathbf{R})$ such that $q = ps$.[39] Therefore, by 8.49,

$$\begin{aligned}
T \text{ has an eigenvalue} &\iff p \text{ has a zero} \\
&\iff b^2 - 4c \geq 0.
\end{aligned}$$

==
Comment, p. **288**, penultimate paragraph before ■

$$\mathcal{M}(T, (e_1, e_2)) = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$$

$$\Downarrow$$

$$\begin{aligned}
Te_1 &= ae_1 + be_2 \\
&= ae_1 + (-b)(-e_2), \\
Te_2 &= -be_1 + ae_2 \\
&= (-1)\left(be_1 + a\left(-e_2\right)\right)
\end{aligned}$$

$$\Downarrow$$

[39] In this case, p and s are monic polynomials of degree one.

$$\mathcal{M}(T,(e_1,-e_2)) = \begin{pmatrix} a & b \\ -b & a \end{pmatrix}.$$

==

EX. 4, p. **294**

For $u, v, x, y \in V$,

$$\begin{aligned}
\big\langle T_{\mathbf{C}}(u+iv), x+iy\big\rangle &= \langle Tu + iTv, x+iy\rangle \\
&= \langle Tu, x\rangle + \langle Tv, y\rangle + (\langle Tv, x\rangle - \langle Tu, y\rangle)i \\
&= \langle u, Tx\rangle + \langle v, Ty\rangle + (\langle v, Tx\rangle - \langle u, Ty\rangle)i \\
&= \langle u+iv, Tx+iTy\rangle \\
&= \big\langle u+iv, T_{\mathbf{C}}(x+iy)\big\rangle.
\end{aligned}$$

==

==
==
10
==
==
Comment, 10.10, p. **299**
Because the use of Determinants to obtain eigenvalues is established in the next section, **10.B**, in order to verify here that $\lambda \in \{1, 2 \pm 3i\}$ is an eigenvalue of T, we may determine a basis (v_λ) of null$(T - \lambda I)$ and verify that

$$\begin{pmatrix} 3-\lambda & -1 & -2 \\ 3 & 2-\lambda & -3 \\ 1 & 2 & -\lambda \end{pmatrix} v_\lambda = \lambda v_\lambda.$$

For example, consider $\lambda = 1$. So the reduced row echelon form of

$$\begin{pmatrix} 2 & -1 & -2 \\ 3 & 1 & -3 \\ 1 & 2 & -1 \end{pmatrix}$$

is

$$\begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

and

$$v_\lambda = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}.$$

==
EX. 4, p. **304**

$$\begin{aligned} \mathcal{M}(T, (v_1, \ldots, v_n)) &= \mathcal{M}(IT, (v_1, \ldots, v_n), (v_1, \ldots, v_n)) \\ &= \mathcal{M}(I, (u_1, \ldots, u_n), (v_1, \ldots, v_n)) \mathcal{M}(T, (v_1, \ldots, v_n), (u_1, \ldots, u_n)) \text{ (by 10.4)} \\ &= \mathcal{M}(I, (u_1, \ldots, u_n), (v_1, \ldots, v_n)) I. \end{aligned}$$

==
EX. 13, p. **305**
By 10.9, 10.13 and 10.16, we may consider the eigenvalue λ of T such that

$$-48 + 24 + \lambda = 51 + (-40) + 1.$$

==
EXERCISES 14–5, p. **305**

$$\begin{aligned} \operatorname{trace}(cT) &= \operatorname{trace} \mathcal{M}(cT) \\ &= \operatorname{trace}(c\mathcal{M}(T)) \\ &= c \operatorname{trace} \mathcal{M}(T) \\ &= c \operatorname{trace} T, \end{aligned}$$

where the first and last equalities come from 10.16, the second equality comes from 3.38, and the third equality is trivial.

$$\begin{aligned} \operatorname{trace}(ST) &= \operatorname{trace} \mathcal{M}(ST) \\ &= \operatorname{trace}(\mathcal{M}(S)\mathcal{M}(T)) \\ &= \operatorname{trace}(\mathcal{M}(T)\mathcal{M}(S)) \\ &= \operatorname{trace} \mathcal{M}(TS) \\ &= \operatorname{trace}(TS), \end{aligned}$$

where the first and last equalities come from 10.16, the second and penultimate equalities come from 3.43, and the third equality comes from 10.14.

==
Comment, 10.26, Solution, 3rd paragraph, 1st sentence, p. **310**
Here $P(j)$ represents the claim
$$T^n v_j = a_1 \cdots a_n v_j \text{ for each } j.$$
So $P(1)$ is valid since $T^{n-1}v_1 = a_1 \cdots a_{n-1}v_n$ implies that
$$\begin{aligned} T^n v_1 &= a_1 \cdots a_{n-1} T v_n \\ &= a_1 \cdots a_{n-1} a_n v_1. \end{aligned}$$
Now suppose $P(j-1)$ holds for a fixed $j \in \{2, \dots, n\}$. Therefore
$$\begin{aligned} Tv_{j-1} = a_{j-1}v_j &\implies T^n v_{j-1} = a_{j-1}T^{n-1}v_j \\ &\implies a_1 \cdots a_n v_{j-1} = a_{j-1}T^{n-1}v_j \\ &\implies a_1 \cdots a_{j-2}a_j \cdots a_n v_{j-1} = T^{n-1}v_j \\ &\implies a_1 \cdots a_n v_j = T^n v_j \\ &\implies P(j) \text{ holds.} \end{aligned}$$
==
Erratum, 10.54, Proof, 1st sentence, p. **326**
'V' should be '$\mathbf{R}^n$'.
==
EX. 1, p. **330**
T is assumed to have no real eigenvalues. So $T_{\mathbf{C}}$ has no real eigenvalues by 9.11. However, by 5.21 and 9.16, $T_{\mathbf{C}}$ has a pair of complex conjugate eigenvalues. Therefore, by 10.20, $\det T = \det T_{\mathbf{C}}$ is a product of positive real numbers, with each number expressed in the form $\lambda\bar{\lambda}$ where λ is a complex eigenvalue of $T_{\mathbf{C}}$.
==
EX. 8, p. **331**
Suppose firstly that $\lambda_1, \dots, \lambda_n$ are the eigenvalues of T (or of $T_{\mathbf{C}}$ if V is a real vector space) with each eigenvalue repeated according to its multiplicity. So
$$\begin{aligned} \overline{\det T} &= \overline{\lambda_1 \cdots \lambda_n} \\ &= \overline{\lambda_1} \cdots \overline{\lambda_n} \\ &= \det T^* \end{aligned}$$
by definition and by exercise **2** in **SECTION 7.A**.[40] Therefore
$$\begin{aligned} |\det T|^2 &= \det T\, \overline{\det T} \\ &= \det T \det T^* \\ &= \det (TT^*) \\ &= \det\left(\sqrt{TT^*}\sqrt{TT^*}\right) \\ &= \det\left(\sqrt{TT^*}\right) \det\left(\sqrt{TT^*}\right) \\ &= \left(\det\left(\sqrt{TT^*}\right)\right)^2 \end{aligned}$$
where the first equality holds because $|z|^2 = z\bar{z}$ for each complex number z, the third and fifth equalities hold by 10.44, p. **320**, and the fourth equality holds by the remarks made before 10.46, p. **322**.
==
EX. 9, p. **331**
Firstly, in 10.56, p. **327**, replace T with $T_i \in \mathcal{L}(\mathbf{R}^n)$, $i = 1, 2$. Now, by the Triangle Inequality,[41]
$$||T_1 y - T_2 y|| \leq ||\sigma(x+y) - \sigma(x) - T_2 y|| + ||T_1 y - \sigma(x+y) + \sigma(x)||$$

[40] See pages **307** and **214**.
[41] 6.18, p. **173**.

for $x \in \Omega$ and $y \in \mathbf{R}^n$ such that $x + y \in \Omega$. So

$$\lim_{y \to 0} \frac{||(T_1 - T_2)(y)||}{||y||} = 0.$$

Then, in particular, for each nonzero vector y and all sufficiently small nonzero real number t,

$$\begin{aligned} 0 &= \lim_{t \to 0} \frac{||(T_1 - T_2)(ty)||}{||ty||} \\ &= \lim_{t \to 0} \left(\frac{|t|}{|t|} \right) \left(\frac{||(T_1 - T_2)(y)||}{||y||} \right) \\ &= \frac{||(T_1 - T_2)(y)||}{||y||}. \end{aligned}$$

Hence $(T_1 - T_2)(y) = 0$ for each $y \neq 0$. Therefore, along with the fact that every linear transformation maps the zero vector to the zero vector, $T_1 - T_2 = 0$.

==

EX. 10, p. **331**

By 10.56, p. **327**, $T \in \mathcal{L}(\mathbf{R}^n)$ is differentiable at $x \in \mathbf{R}^n$ if

$$\lim_{y \to 0} \frac{||T(x+y) - Tx - Ly||}{||y||} = 0 \tag{24}$$

for an operator $L \in \mathcal{L}(\mathbf{R}^n)$. Furthermore, by **EX. 9**, p. **331**, this $L := T'(x)$ is unique. So, since

$$x, y \in \mathbf{R}^n \Longrightarrow ||T(x+y) - Tx - Ty|| = 0,$$

(24) holds and $T'(x) = T$ for each $x \in \mathbf{R}^n$.

==