Topics in Linear Algebra

Michael Sever

1 Inner product spaces

We assume given a linear vector space V over a field F. Throughout these notes, unless explicitly stated otherwise, $F = \mathbb{R}$ or $F = \mathbb{C}$. Often, but not always, V will be finite dimensional.

A norm on V is simply a measure of the size of the elements of V, and is required to satisfy the following: For all $u, v \in V, \alpha \in F$

(1.1) $\|u\| \geq 0$, and $\|u\| = 0$ if and only if $u = 0$

(1.2) $\|\alpha u\| = |\alpha|\ \|u\|$

(1.3) $\|u + v\| \leq \|u\| + \|v\|$

Since $\|u\| = \|u - v + v\| \leq \|u - v\| + \|v\|$, an equivalent form of the triangle inequality (1.3) is $\|u\| - \|v\| \leq \|u - v\|$.

Examples: For $F = \mathbb{R}, V = \mathbb{R}^n, u = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix} \in V, u_1, u_2, \ldots, u_n \in \mathbb{R}$:

(1.4) the ℓ_1 norm $\|u\|_1 = \sum\limits_{i=1}^{n} |u_i|$

(1.5) the ℓ_2 norm $\|u\|_2 = (\sum\limits_{i=1}^{n} u_i^2)^{1/2}$

(1.6) the ℓ_∞ norm $\|u\|_\infty = \underset{i=1,2,\ldots,n}{\text{maximum}}\ |u_i|$

A linear vector space V, equipped with a norm $\|\cdot\|$, is said to be complete if every Cauchy sequence within the space converges to a limit within the space, i.e.

Let $\{v_m\}_{m=1}^{\infty}$ be a sequence of elements v_m in V, such that for any $\varepsilon > 0$ there exists N_ε such that for all $j, k > N_\varepsilon, \|v_j - v_k\| < \varepsilon$; then there exists $v \in V$ such that $\|v_m - v\| \to 0$ as $m \to \infty$.

A complete normed linear space is called a Banach space.

Example 1: $V = \mathbb{R}^n$, with any of the norms (1.4), (1.5), (1.6)

Example 2: The space of continuous functions on the interval $[0, 1]$ with the norm $\|u\| = \sup\limits_{x \in [0,1]} |u(x)|$ for $u \in C[0, 1]$.

Example 3: The space of continuous functions on the interval $(0,1)$, given the norm $\|u\| = (\int_0^1 u(x)^2 dx)^{1/2}$ is not complete, as there are sequences of continuous functions, convergent with respect to this norm, such that the limit is not continuous.

A special example of a Banach space is an inner product space, or Hilbert space. A Hilbert space H is equipped with an inner product, denoted here by $(\cdot,\cdot)$, a mapping of $H \times H$ into F. For any $p,q,r,s \in H, \alpha,\beta \in F$, the inner product satisfies

(1.7) $(p+q,r+s) = (p,r)+(p,s)+(q,r)+(q,s)$

(1.8) $(\alpha p,q) = \alpha(p,q)$

(1.9) $(p,\beta q) = \bar{\beta}(p,q)$

(1.10) $(p,q) = \overline{(q,p)}$

(1.11) $(p,p) \geq 0$, and $(p,p) = 0$ only for $p = 0$

In (1.9), (1.10), we have tacitly assumed $F = \mathbb{C}$, and $^-$ denotes complex conjugate. For $F = \mathbb{R}$, all such complex conjugate symbols, here and below, are understood omitted.

Furthermore, the norm of a Hilbert space is directly related to the inner product, and is obtained from

(1.12) $\|p\| = (p,p)^{1/2}$

Example 1: $H = \mathbb{R}^n, p = \begin{pmatrix} p_1 \\ \vdots \\ p_n \end{pmatrix}, q = \begin{pmatrix} q_1 \\ \vdots \\ q_n \end{pmatrix}$

$$(1.13) \qquad (p,q) = \sum_{i=1}^n p_i q_i \quad , \quad \|p\| = (\sum_i p_i^2)^{1/2}$$

Example 2: $H = \mathbb{C}^n, p, q$ as in example 1,

$$(1.14) \qquad (p,q) = \sum_{i=1}^n p_i \bar{q}_i \quad , \quad \|p\| = (\sum_i |p_i|^2)^{1/2}$$

Example 3: $H = \mathbb{R}^n, p, q$ as above, $\alpha_1, \alpha_2, \ldots \alpha_n$ strictly positive real numbers, we could use a weighted inner product,

$$(1.15) \qquad (p,q)_\alpha = \sum_{i=1}^n \alpha_i p_i q_i \quad , \|p\| = (\sum_i \alpha_i p_i^2)^{1/2}$$

From (1.11), it is clear that the norm obtained from (1.12) satisfies (1.1), and (1.2) follows from (1.12) as well, using (1.8), (1.9).

Theorem 1.1. *[the Schwarz inequality]: For any $p, q \in H$,*

$$|(p,q)| \leq ||p||\ ||q|| \tag{1.16}$$

Proof. For p or $q = 0$, or if $(p,q) = 0$ this is obvious, so we assume $(p,q) \neq 0$. Then for any $t \in F$, from (1.11), using (1.8 - 1.10), (1.12)

$$\begin{aligned} 0 &\leq (p + tq, p + tq) \\ &= (p,p) + t(q,p) + \bar{t}(p,q) + |t|^2(q,q) \\ &= ||p||^2 + 2Re(t(q,p)) + |t|^2||q||^2 \end{aligned} \tag{1.17}$$

This has to hold for any $t \in F$. We choose the phase of t so that $t(q,p)$ is a real negative number, and the magnitude $|t|$ to minimize the right side of (1.17). This is accomplished by the choice

$$t = -\frac{(p,q)}{|(p,q)|}\frac{||p||}{||q||}, \tag{1.18}$$

and inserting this value into (1.17) we obtain

$$\begin{aligned} 0 &\leq ||p||^2 - 2\frac{(p,q)\overline{(p,q)}}{|(p,q)|}\frac{||p||}{||q||} + ||p||^2 \\ &= 2||p||^2 - 2|(p,q)|\ ||p||/||q|| \end{aligned} \tag{1.19}$$

from which (1.16) follows easily.

Lemma 1.2. *The norm obtained from (1.12) always satisfies the triangle inequality (1.3).*

Proof. Given $p, q \in H$,

$$\begin{aligned} ||p+q||^2 &= (p+q, p+q) \\ &= ||p||^2 + (p,q) + (q,p) + ||q||^2 \\ &= ||p||^2 + 2Re(p,q) + ||q||^2 \\ &\leq ||p||^2 + 2|(p,q)| + ||q||^2 \\ &\leq ||p||^2 + 2||p||\ ||q|| + ||q||^2 \\ &= (||p|| + ||q||)^2 \end{aligned} \tag{1.20}$$

using the Schwarz inequality in the next to last step.

A subspace U within a Banach space V is a linear vector space within V, equipped with the same norm as V, and required to be complete with respect to this norm. In the case that V is also a Hilbert space, U is also a Hilbert space, equipped with the same inner product as V.

A bounded linear functional ℓ on a Banach space V is a mapping $\ell : V \to F$ satisfying, for all $p, q \in V, \alpha, \beta \in F$:

$$\ell(\alpha p + \beta q) = \alpha \ell(p) + \beta \ell(q) \tag{1.21}$$

$$|\ell(p)| \leq \|\ell\| \, \|p\| \tag{1.22}$$

for some nonnegative real constant $\|\ell\|$, independent of p in particular, called the norm of the functional ℓ.

For V of finite or countable dimension, the form of a bounded linear functional follows from (1.21). Let $\{e_i\}$ be a basis for V; then any $v \in V$ is of the form $v \equiv \sum_i \alpha_i e_i, \alpha_i \in F$, and $\ell(v) = \ell(\Sigma \alpha_i e_i) = \sum_i \alpha_i \ell(e_i) = \sum_i \alpha_i \ell_i$, with each $\ell_i = \ell(e_i)$.

Theorem 1.3. *The set of all bounded linear functionals on a given Banach space V, equipped with this norm, is itself a Banach space. It is called the dual space of V, and denoted typically by V^*.*

Proof. For $\ell, b \in V^*, \alpha \in F, p \in V$ we identify

$$(\ell + b)(p) = \ell(p) + b(p) \tag{1.23}$$

$$(\alpha \ell)(p) = \alpha \ell(p) \tag{1.24}$$

and it is obvious that V^* is a linear vector space. From (1.23),

$$\begin{aligned} |(\ell + b)(p)| &= |\ell(p) + b(p)| \\ &\leq |\ell(p)| + |b(p)| \\ &\leq (\|\ell\| + \|b\|) \, \|p\| \end{aligned}$$

so the triangle inequality (1.3) holds for the norm of V^*. It remains to prove completeness; let $\{\ell_n\}_{n=1}^{\infty}$ be any Cauchy sequence within V^*. Then for any

$p \in V, \{\ell_n(p)\}_{n=1}^{\infty}$ is a Cauchy sequence within F, as given any $\varepsilon > 0, j, k > N_\varepsilon$ (so that $\|\ell_j - \ell_k\| \leq \varepsilon$)

$$\begin{aligned} |\ell_j(p) - \ell_k(p)| &= |(\ell_j - \ell_k)(p)| \\ &\leq \|\ell_j - \ell_k\| \, \|p\| \\ &\leq \varepsilon \|p\| \end{aligned} \tag{1.25}$$

and $\|p\|$ is obviously independent of j, k. Since F is complete, with respect to 1.1 (whether F is $\mathbb{R}$ or $\mathbb{C}$), the sequence $\{\ell_n(p)\}$ converges to a limit, which we identify as $\ell(p)$. It is easily verified that $\ell : V \to F$ so determined is indeed a bounded linear functional, i.e. an element of V^* and one can take

$$\|\ell\| = \lim_{n\to\infty} \|\ell_n\|. \tag{1.26}$$

To finish the proof of completeness of V^*, we need to show that $\|\ell_n - \ell\| \to 0$ as $n \to \infty$, that is for any $\varepsilon > 0$ there exists M_ε such that $|\ell_n(u) - \ell(u)| \leq \varepsilon$ for any $n \geq M_\varepsilon$ and any $u \in V$ with $\|u\| = 1$. We choose $M_\varepsilon = N_{\varepsilon/2}$, and then $k \geq M_\varepsilon$, depending on ε and the given vector u, such that $|\ell_k(u) - \ell(u)| \leq \varepsilon/2$. Then for any $n \geq M_\varepsilon$

$$\begin{aligned} |\ell_n(u) - \ell(u)| &\leq |\ell_n(u) - \ell_k(u)| + |\ell_k(u) - \ell(u)| \\ &\leq \|\ell_n - \ell_k\| \, \|u\| + |\ell_k(u) - \ell(u)| \\ &\leq \varepsilon \end{aligned}$$

In the same manner we can find a Banach space V^{**}, the dual of V^*. Indeed, for any $p \in V, \ell \in V^*$, we can identify p as an element of V^{**} by taking

$$p(\ell) = \ell(p) \tag{1.27}$$

so

$$\begin{aligned} |p(\ell)| &= |\ell(p)| \\ &\leq \|\ell\| \, \|p\| \end{aligned} \tag{1.28}$$

so we can use the same norm $\|p\|$ for p, regarded as an element of V^{**}, as for p regarded as an element of V. Thus we have shown

Theorem 1.4. *V is a subspace of V^{**}.*

The following example, however, shows that V^{**} can be bigger than V. Let V be the space of infinite bounded sequence of real numbers, i.e. $v = \{v_n\}_{n=1}^{\infty}$, equipped with the norm

$$\|v\|_V = \underset{n=1,2,\ldots}{\text{maximum}} \; |v_n| \tag{1.29}$$

and satisfying

$$v_n \to 0 \text{ as } n \to \infty. \tag{1.30}$$

Then V^* is also a space of infinite bounded sequences of real numbers, $\ell = \{\ell_n\}_{n=1}^{\infty}$, $\ell_n \in \mathbb{R}$, identifying $\ell(v) = \sum_n \ell_n v_n$; V^* is equipped with the norm

$$\|\ell\|_{V^*} = \sum_{n=1}^{\infty} |\ell_n|$$

Similarly, V^{**} is the space of infinite bounded sequences of real numbers; for $v \in V^{**}, \ell \in V^*$, we identify $v(\ell) = \sum_n v_n \ell_n$. Thus V^{**} is equipped with the same norm $\|\cdot\|_V$ given by (1.29), but the elements of V^{**} are not required to satisfy (1.30). Thus the sequence $(1, 1, 1, \ldots)$ is an element of V^{**} but not of V.

It will be noted that these spaces V, V^*, V^{**} are not Hilbert spaces, because of the specific choice of norm; the norms chosen do not obtain from an inner product. For such infinite dimensional spaces, the specific choice of norm is quite important. (For finite dimensional spaces, it generally is not; the norms given in (1.4 - 1.6), for example, satisfy

$$\|u\|_\infty \le \|u\|_2 \le \|u\|_1 \le \sqrt{n}\|u\|_2 \le n\|u\|_\infty \tag{1.31}$$

so a Cauchy sequence with respect to any of these norms is a Cauchy sequence with respect to the others as well.)

A Hilbert space, however, is its own dual, whether finite or infinite dimensional. We first prove the following lemma:

Lemma 1.5. *Let U be a proper subspace of Hilbert space H. Then there exists a nonzero element $x \in H$, such that*

$$(x, u) = 0 \textit{ for all } u \in U, \; \textit{often written } x \perp U \tag{1.32}$$

Proof. Since U is a proper subspace of H, there exists an element $v \in H, v \notin U$. Consider

$$\inf_{u \in U} \|v-u\|^2 = \alpha \tag{1.33}$$

Clearly $\alpha \geq 0$. Furthermore α cannot be zero, as if $\alpha = 0$, there is a sequence $\{u_n\}_{n=1}^{\infty}$ in U such that $\|v-u_n\| \to 0$ as $n \to \infty$. But U is complete, so any such limit v is also within U, contradicting our specification $v \notin U$.

But for any sequence $\{\varepsilon_n\}, \varepsilon_n \downarrow 0$ as $n \to \infty$ there will be a sequence $\{u_n\} \in U$ such that

$$\|v-u_n\|^2 \leq \alpha + \varepsilon_n^2 \tag{1.34}$$

And for any fixed n, let $t \in F$, $w \in U$ such that $\|w\| = 1$, otherwise arbitrary; then $u_n + tw \in U$, and from (1.33) we still have

$$\|v-u_n-tw\|^2 \geq \alpha. \tag{1.35}$$

Expanding (1.35), using (1.34)

$$\begin{aligned} \alpha &\leq \|v-u_n-tw\|^2 \\ &= \|v-u_n\|^2 - 2Re(\bar{t}(v-u_n,w)) + |t|^2 \\ &\leq \alpha + \varepsilon_n^2 - 2Re(\bar{t}(v-u_n,w)) + |t|^2. \end{aligned} \tag{1.36}$$

For any w such that $(v-u_n,w) \neq 0$, we choose

$$t = \varepsilon_n (v-u_n,w)/|(v-u_n,w)|, \tag{1.37}$$

obtaining from (1.36)

$$0 \leq \varepsilon_n^2 - \varepsilon_n |(v-u_n,w)| \tag{1.38}$$

equivalently

$$|(v-u_n,w)| \leq \varepsilon_n \text{ for all } w \in U, \|w\| = 1. \tag{1.39}$$

Fix any $\varepsilon > 0$, and obtain N_ε such that $\varepsilon_n < \varepsilon/2$ for all $n > N_\varepsilon$. Then for any $m, n > N_\varepsilon$, we have (1.39) and $u_m \in U$ such that

$$|(v-u_m,w)| \leq \varepsilon_m, \text{ for any } w \in U, \|w\| = 1. \tag{1.40}$$

Combining (1.39) and (1.40),

$$
\begin{aligned}
|(u_n - u_m, w)| &\le \varepsilon_n + \varepsilon_m \\
&\le \varepsilon;
\end{aligned}
\tag{1.41}
$$

Since $u_n, u_m \in U$, so is $u_n - u_m$, and we can take $w = (u_n - u_m)/\|u_n - u_m\|$ in (1.41), obtaining

$$
\|u_n - u_m\| \le \varepsilon \text{ for all } n, m > N_\varepsilon \tag{1.42}
$$

Thus the sequence $\{u_j\}_{j=1}^{\infty}$ is in fact Cauchy, and as U is necessarily complete, $u_n \to z$ as $n \to \infty$, for some $z \in U$. Then for any $w \in U, \|w\| = 1$, and any n

$$
\begin{aligned}
|(v - z, w)| &\le |(v - u_n, w)| + |(u_n - z, w)| \\
&\le \varepsilon_n + \|u_n - z\|,
\end{aligned}
\tag{1.43}
$$

and the right side of (1.43) clearly approaches zero as $n \to \infty$. Since $z \in U$ and $v \notin U, v - z \neq 0$, and $x = v - z$ satisfies (1.32). We note that $x \perp U$ achieves the minimum value α in (1.33).

Theorem 1.6. *(The Riesz representation theorem for Hilbert spaces): A bounded linear functional ℓ on a Hilbert space H admits a representation as the inner product with a unique element of H, i.e. for any $\ell \in H^*$ there exists a unique $v \in H$ such that*

$$
\ell(u) = (u, v) \textit{ for all } u \in H. \tag{1.44}
$$

Proof. Set $U = \{u \in H | \ell(u) = 0\}$. It is easily verified that U is a subspace of H, and a proper subspace of H unless $\ell(u) = 0$ for all $u \in H$, in which case (1.44) holds with $v = 0$. For U a proper subspace of H, there exists $x \in H, x \perp U, x \neq 0$ by lemma 1.5. Thus $\ell(x) \neq 0$, and setting $\eta = x/\ell(x)$ we have $\ell(\eta) = 1, \eta \perp U$. Now η is unique within H, for if $\tilde{\eta}$ satisfies $\ell(\tilde{\eta}) = 1, \tilde{\eta} \perp U$, then $\eta - \tilde{\eta} \perp U$, but $\ell(\eta - \tilde{\eta}) = 0$ so $\eta - \tilde{\eta} \in U$ and thus $\eta = \tilde{\eta}$. Therefore the one-dimensional space $K = \{\alpha\eta | \alpha \in F\}$ is the perpendicular complement of u in H, i.e. $K = \{y \in H | y \perp U\}$. Now for any $u \in H$, we can write

$$
u = u_U + u_K, u_U = u - \frac{(u, \eta)\eta}{\|\eta\|^2}, u_K = \frac{(u, \eta)\eta}{\|\eta\|^2} \tag{1.45}
$$

clearly $u_K \in K$ and $u_U \perp \eta$. We claim that $u_U \in U$. If not, i.e. if $\ell(u_U) \neq 0$, then $\tilde{\eta} = u_U/\ell(u_U)$ is perpendicular to η and satisfies $\ell(\tilde{\eta}) = 1$. Thus $\ell(\eta - \tilde{\eta}) = 0$, so $\eta - \tilde{\eta} \in U$. As $\eta \perp U, \eta \perp \eta - \tilde{\eta}$. This is impossible, as $\tilde{\eta} \perp \eta$ and $(\eta, \eta) = ||\eta||^2 \neq 0$, thus proving the claim $u_U \in U$. Thus

$$\begin{aligned} \ell(u) &= \ell(u_U + u_K) \\ &= \ell(u_U) + \ell(u_K) \\ &= 0 + (u, \eta)\ell(\eta)/\|\eta\|^2 \\ &= (u, \eta)/\|\eta\|^2 \end{aligned} \tag{1.46}$$

so $v = \eta/\|\eta\|^2$ satisfies (1.44). And should another such $\tilde{v} \in H$ satisfy (1.44) as well, then for all $u \in H$

$$\begin{aligned} 0 &= \ell(u) - \ell(u) \\ &= (u, v) - (u, \tilde{v}) \\ &= (u, v - \tilde{v}) \\ &= \|v - \tilde{v}\|^2 \end{aligned} \tag{1.47}$$

taking $u = v - \tilde{v}$.

Lemma 1.7. *Let U be a proper subspace of a Hilbert space H, and let $U_\perp = \{v \in H, v \perp U\}$. Then $U_\perp$ is also a proper subspace of $H, U \cap U_\perp = \{0\}$, and $H = U \oplus U_\perp$, i.e. for any $h \in H$, there exist unique $u \in U, v \in U_\perp$ such that $h = u + v$.*

Proof. Only the last statement requires proof. Given $h \in H$, we have a bounded linear functional on U determined from $\ell(\omega) = (\omega, h)$ for all $\omega \in U$. By the Riesz theorem, there exists $u \in U$ such that $\ell(\omega) = (\omega, h) = (\omega, u)$ for all $\omega \in U$. Thus $h - u$ is perpendicular to U and so is in $U_\perp$. Suppose uniqueness fails, and $h = u + v = \tilde{u} + \tilde{v}$, with $u, \tilde{u} \in U, v, \tilde{v} \in U_\perp$. Then $u - \tilde{u} = \tilde{v} - v$, but $u - \tilde{u} \in U$ and $\tilde{v} - v \in U_\perp$.

Problems

1. Let V, V^*, V^{**} be the spaces of real sequences in the example following Theorem 1.4. Show that each is a Banach space, i.e. prove completeness. Show that V^* is the dual of V, i.e. that $\ell \in V^*$ if and only if $\|\ell\|_{V^*} < \infty$. Similarly, show $v \in V^{**}$ if and only if $\|v\|_V < \infty$.

2. Let $H = \mathbb{R}^4$, and $U = \text{span}\left(\begin{pmatrix}1\\0\\1\\0\end{pmatrix}, \begin{pmatrix}1\\-1\\-1\\1\end{pmatrix}, \begin{pmatrix}1\\1\\3\\-1\end{pmatrix}\right)$ find $U_\perp$.

3. Let $H \subset C[0,1]$, such that any $u \in H$ satisfies $u(0) = 0, \int\limits_0^1 (\frac{du}{dx})^2 dx < \infty$, with $(u, v) = \int\limits_0^1 \frac{du}{dx}\frac{dv}{dx} dx$. For each of the following linear functionals on H, find a suitable bound for $\|\ell\|$, and find $z \in H$ such that $\ell(u) = (u, z)$ for all $u \in H$

   a) $\ell_1(u) = u(1/2)$

   b) $\ell_2(u) = \int\limits_{1/2}^1 u(x)dx$

   c) $\ell_3(u) = \int\limits_0^1 \frac{u(x)dx}{x}$

   (More precisely, the functions $u \in H$ are differentiable almost everywhere in the interval $(0, 1)$ and satisfy $u(z) - u(y) = \int\limits_y^z \frac{du}{dx} dx$, for any $0 \le y < z \le 1$. You may assume that this space is indeed a Hilbert space.)

4. A bilinear form $[\ ,\ ]$ on a Hilbert space H is a mapping of $H \times H$ into F, i.e. $[p, q] \in F$ for any $p, q \in H$, required to satisfy (1.7), (1.8), (1.9), but not necessarily (1.10) or (1.11).

   a) Assume H a real Hilbert space (i.e. $F = \mathbb{R}$) and that $[p, p] = 0$ for all $p \in H$. Show $[p, q] = -[q, p]$ for all $p, q \in H$, and find a nontrivial example for $H = \mathbb{R}^2$.

   b) Assume H complex ($F = \mathbb{C}$) and $[p, p] = 0$ for all $p \in H$. Show that $[p, q] = 0$ for all $p, q \in H$.

   c) Again assume H complex, and assume that $\alpha\|p\|^2 \le [p, p] \le \beta\|p\|^2$ for some positive constants α, β. Find a constant C, depending on α, β but independent of $p, q \in H$, such that $|[p, q]| \le C\|p\|\ \|q\||$.

5. Construct an example of a linear functional on a real Hilbert space, satisfying (1.21), which is not bounded, i.e. for any constant C there exist $p \in H, \|p\| = 1$, such that $\ell(p) > C$.

6. Let V be a Banach space of finite dimension n. Show that V^{**} is isomorphic to V, and find the dimensions of V^* and V^{**}.

7. Let H be the space of problem 3, and let $U \subset H$ be a subspace of H, the elements of which satisfy $u(1) = 0$ as well. Find $U_\perp$.

8. Let U be a subspace of a Hilbert space H, and let $P : H \to U$ be the projection of H into U, i.e. for all $h \in H, v = Ph \in U$ and satisfies $(u, h) = (u, v)$ for all $u \in U$.

   a) show that P is defined, i.e. that Ph is unique

   b) show that P is onto U

   c) show that $\|P\| = 1$, i.e. $\|Ph\| \le \|h\|$ for all $h \in H$, and that there exists $h \in H$ such that $\|Ph\| = \|h\|$

   d) show that P is symmetric, i.e. that $(g, Ph) = (Pg, h)$ for all $g, h \in H$.

9. For $H = \mathbb{R}^n$, denote by $|u|_\alpha = (\sum_{i=1}^{n} |u_i|^\alpha)^{1/\alpha}$ for any $0 < \alpha < \infty$.

   a) For $u \in \mathbb{R}^n$ fixed, show that $|u|_\alpha \to \|u\|_\infty$, given by (1.6), as $\alpha \to \infty$.

   b) Show that $|\cdot|_\alpha$ is <u>not</u> a norm for $\alpha < 1$. A simple counterexample will suffice.

   c) For $1 \le \alpha < \beta < \infty$, show $|u|_\beta \le |u|_\alpha$ and find a constant C, depending on α, β, such that $|u|_\alpha \le C|u|_\beta$. For which vectors u does equality hold in each case?

10*. Let [,] be a bilinear form on a Hilbert space H, i.e. $[p, q] \in F$ for all $p, q \in H$, and [,] satisfies (1.7), (1.8), (1.9), and (1.11), but not necessarily (1.10). Furthermore $|[p, q]| \le \|p\| \, \|q\|$.

   Statement: Every bounded linear functional on H can be represented in terms of [,], i.e. for any $\ell \in H^*$ there exists a unique $u \in H$ such that $\ell(v) = [v, u]$ for all $v \in H$.

Justification: Uniqueness of u is easy, using (1.11). To show existence, we note that for any fixed $u, [v, u]$ determines a bounded linear functional on H, i.e. $b(v) = [v, u]$ for any $v \in H$. Therefore by the Riesz theorem, for any $v \in H, b(v) = (v, z) = [v, u]$ for some unique $z \in H$.

Consider the mapping $Q : H \to H$ determined by $Qu = z$ for any $u \in H$. This mapping is linear, using (1.7 - 1.9), and bounded. Therefore it suffices to show that Q is onto H. Suppose not, then $QH = \{Qu, u \in H\}$ is a <u>proper</u> subspace of H, and by lemma 1.5, there exists a nonzero element of $H, x \perp QH$, i.e. for any $u \in H, 0 = (x, Qu) = [x, u]$. Then choosing $u = x$ gives a contradiction with (1.11).

This argument is incorrect. It is true for finite - dimensional Hilbert spaces, and it is true with the stronger assumption

$(*)$ $\quad [p, p] \geq \alpha \|p\|^2$ for all $p \in H$, with some constant $\alpha > 0$. (With the assumption $(*)$ it is well-known as the Lax-Milgram theorem.)

a) Find the mistake in the argument as stated.

b) Fix the argument, given the assumption $(*)$.

11. For H of dimension n, let the ℓ_1, ℓ_2 and ℓ_∞ norms on H be as shown in (1.4), (1.5), (1.6). For $A = \{A_{ij}\}_{i,j=1}^n$ any $n \times n$ matrix, we determine the corresponding <u>induced norms</u> from

$$\|A\|_1 = \sup_{h \neq 0} \|Ah\|_1 / \|h\|_1, \|A\|_2 = \sup_{h \neq 0} \|Ah\|_2 / \|h\|_2,$$

$\|A\|_\infty = \sup_{h \neq 0} \|Ah\|_\infty / \|h\|_\infty$. Sometimes it is more convenient to work with other matrix norms, such as $|||A|||_\infty = \underset{i,j}{\text{maximum}} |A_{ij}|$ or $|||A|||_1 = \sum_{i,j=1}^n |A_{ij}|$. Show that these norms are all equivalent, i.e. that any one of them majorizes all the others, with factors depending only on n. A convenient way of stating these results is in terms of chains of inequalities of the form (1.31). Show that for any such choice of $\|A\|$, there is a positive constant α such that $\|A^m\| \leq \alpha^m \|A\|^m$ for all positive integers m. Here α may depend on n and on the specific choice for $\|A\|$, but is independent of A and m. Try to make all of these results "best possible", i.e. the inequalities can become equalities for a specific choice of the matrix A.

Show that all of these definitions of the norm of a matrix satisfy the triangle inequality

12*. Let A be a bounded linear operator on an infinite-dimensional Hilbert space H. Prove the existence of another bounded linear operator $A^\dagger : H \to H$, such that for all $x, y \in H, (Ax, y) = (x, A^\dagger y)$. Hint: Denote by $\ell_y : H \to F$ the mapping determined by $\ell_y(x) = (Ax, y)$ for all $x \in H$, any fixed y, and use the Riesz theorem.

13. Let X be a given Hilbert space, and $A : X \to X$ a given linear transformation, with $A0 = 0$ in particular.

(a) Show that if X is finite-dimensional, either there exists nonzero $x \in X$ such that $Ax = 0$, or else $A^{-1} : X \to X$ exists, i.e. for every $y \in X$ there exists a unique $z \in X$ such that $Az = y$. (This is commonly referred to as the Fredholm alternative.)

(b) Give an example showing that this is not true for infinite-dimensional X.

14. In the proof of theorem 1.3:

(a) Show V^* as defined is a linear vector space

(b) Show that the limit (1.26) exists, and that $\|\ell\|$ so obtained satisfies (1.22)

15. Let $\ell : V \to F$ be a bounded linear functional on $V, \rho_i \in V, i = 1, \ldots,$ such that $\sum\limits_{i=1}^{\infty} \|\rho_i\|$ is finite.

With $\ell\Big(\sum\limits_{i=1}^{\infty} \rho_i\Big)$ defined as $\lim\limits_{k\to\infty} \ell\Big(\sum\limits_{i=1}^{k} \rho_i\Big)$, show that $\ell\Big(\sum\limits_{i=1}^{\infty} \rho_i\Big)$ exists and is equal to $\sum\limits_{i=1}^{\infty} \ell(\rho_i)$.

16. Suppose that $|(u, v)| = \|u\| \, \|v\|$ for some $u, v \in H, u, v \neq 0$. Show that $v = \alpha u$ for some $\alpha \in F$.

17* Let $H = L^2(0, 1)$, i.e. for $u \in H,\ u : (0, 1) \to \mathbb{R}$ such that $\int\limits_0^1 u(x)^2 dx < \infty$, with the inner product $(u, v) = \int\limits_0^1 u(x)v(x)dx$. Let $Y : H \to H$ be determined from $Yu(x) = 2u(\frac{x}{2}) - u(1 - x),\ 0 < x < 1$. Find $Y^\dagger : H \to H$ such that $(Yu, v) = (u, Y^\dagger v)$ for all $u, v \in H$. Find a number K such that $\|Yu\| \leq K\|u\|$ for all $u \in H$.

18* Let $H = L^2(0,1)$ as in problem 17, and let $Z : H \to H$ be determined from $Zu(x) = u(x^2)$, $0 < x < 1$ for all $u \in H$. Prove that Z is unbounded, that for all (large) M there exists u_M such that $\|Zu_M\| \geq M\|u_M\|$.

19 Let $X \subset H$ be a proper subspace, and $P : H \to X$ determined by $(u,x) = (Pu,x)$ for all $x \in X$. Show that P is symmetric, i.e. $(Pu,v) = (u,Pv)$ for all $u,v \in H$.

2 Some basic properties of matrices

For a fixed positive integer n, let $H = \mathbb{R}^n$ or $\mathbb{C}^n$ depending on whether F is $\mathbb{R}$ or $\mathbb{C}$. A linear operator $A : H \to H$ then admits a representation as an $n \times n$ matrix, and we shall denote both the operator and the matrix by the same symbol, thus assuming a given basis for H.

Given an $n \times n$ matrix A, we have the following definitions:

(2.1) The range of $A, R(A) = \{\eta \in H | \eta = Ax \text{ for some } x \in H\}$

(2.2) the kernel of $A, \ker(A) = \{x \in H | Ax = 0\}$

Lemma 2.1. *Both $R(A)$ and ker (A) are subspaces of H.*

The proof is immediate, since A is a linear operator, but depends on the fact that H is finite-dimensional. Thus $\ker(A)$ is complete, that is closed, because A is bounded, $\|Ah\| \leq \|A\| \, \|h\|$ for any $h \in H$, for some nonnegative number $\|A\|$, the norm of A. And $R(A)$ is also complete, as it is simply span $\{Ae_i\}_{i=1}^n$, where $\{e_i\}_{i=1}^n$ is a basis for H. For example, taking the e_i as unit vectors, $(e_i)_j = \delta_{ij}$, we note that $R(A)$ is just the span of the columns of A, viewed as separate n-dimensional vectors.

Special cases are A nonsingular, i.e. A is 1-1 onto H, so $R(A) = H, \ker(A) = \{0\}$, and A identically zero, so $R(A) = \{0\}, \ker(A) = H$. For A singular but not identically zero, both $R(A)$ and $\ker(A)$ are proper subspaces of H. Alternatively, $R(A)$ is denoted by $AH = \{Ah | h \in H\}$.

The relation between $R(A)$ and $\ker(A)$ can be described easily in terms of the transpose or adjoint of A, respectively given by

(2.3) $$(A^t)_{ij} = A_{ji} \quad , \quad i,j = 1,\ldots,n$$

(2.4) $$(A^\dagger)_{ij} = \overline{A_{ji}}, i,j = 1,\ldots,n$$

For $F = \mathbb{R}$ (i.e. A real) $A^t = A^\dagger$; for $F = \mathbb{C}, A^\dagger = \overline{A^t}$. From (2.3) or (2.4) it follows easily that $(A^t)^t = (A^\dagger)^\dagger = A$.

Theorem 2.2. *For any $x, y \in H$,*

$$(A^\dagger x, y) = (x, Ay) \tag{2.5}$$

Proof. Since $(p, q) = \sum_{i=1}^{n} p_i \bar{q}_i$,

$$\begin{aligned} (x, Ay) &= \sum_{i=1}^{n} x_i \overline{(Ay)}_i = \sum_{i=1}^{n} x_i \overline{(\sum_{j=1}^{n} A_{ij} y_j)} \\ &= \sum_{i=1}^{n} \sum_{j=1}^{n} x_i \overline{A_{ij}} \bar{y}_j = \sum_{j=1}^{n} (\sum_{i=1}^{n} \overline{A_{ij}} x_i) \bar{y}_j \\ &= \sum_{j=1}^{n} (\sum_{i=1}^{n} A^\dagger_{ji} x_i) \bar{y}_j = \sum_{j=1}^{n} (A^\dagger x)_j \bar{y}_j \\ &= (A^\dagger x, y) \end{aligned} \tag{2.6}$$

In the case of F real, $A^\dagger$ becomes A^t.

Theorem 2.3.

$$R(A) \perp \ker(A^\dagger) \textit{ and } \ker(A) \perp R(A^\dagger). \tag{2.7}$$

Proof. Let $y \in R(A)$ and $z \in \ker(A^\dagger)$. Then $y = Ax$ for some $x \in H$, and

$$(z, y) = (z, Ax) = (A^\dagger z, x) = (0, x) = 0 \tag{2.8}$$

using (2.5). Similarly, for $z \in \ker(A), q \in R(A^\dagger), q = A^\dagger p$ for some $p \in H$, and

$$(z, q) = (z, A^\dagger p) = (Az, p) = (0, p) = 0. \tag{2.9}$$

In general it is not true that $R(A)$ and $\ker(A)$ are perpendicular.

These results extend very easily to non-square matrices. Let $H_n = \mathbb{R}^n$ (or $\mathbb{C}^n$) and $H_m = \mathbb{R}^m$ (or $\mathbb{C}^m$). Then a linear operator $A : H_m \to H_n$ is represented by an $n \times m$ matrix. A^t and $A^\dagger$ are $m \times n$ matrices, obtained analogously with (2.3), (2.4). In this case the respective inner products on H_m and H_n are

$$(p,x)_n = \sum_{i=1}^{n} p_i \bar{x}_i \quad , \quad (q,y)_m = \sum_{i=1}^{m} q_i \bar{y}_i \tag{2.10}$$

for $x, p \in H_m, y, q \in H_n$. Theorem 2.2 now holds in the form

$$(A^\dagger x, y)_m = (x, Ay)_n \tag{2.11}$$

(same proof as above), and Theorem 2.3 also remains valid, with $R(A) \perp \ker(A^\dagger)$ in H_n and $R(A^\dagger) \perp \ker(A)$ in H_m (again, same proof as above).

From (2.3) or (2.4), we have $(A+B)^\dagger = A^\dagger + B^\dagger$ whenever A and B are matrices of the same dimensions, and for A, B such that AB is defined,

$$(AB)^\dagger = B^\dagger A^\dagger \tag{2.12}$$

immediately from (2.4) and the definition of matrix multiplication. Alternatively, one can use (2.5), which is a more general definition of an adjoint operator, and note that $(x, ABy) = (x, A(By)) = (A^\dagger x, By) = (B^\dagger(A^\dagger x), y) = (B^\dagger A^\dagger x, y) = ((AB)^\dagger x, y)$ for all $x, y \in H$, thus identifying $B^\dagger A^\dagger$ with $(AB)^\dagger$. For any nonsingular A, taking $B = A^{-1}$ in (2.12), since $I^\dagger = I$ we find

$$(A^{-1})^\dagger = (A^\dagger)^{-1}. \tag{2.13}$$

Next we introduce two classes of $n \times n$ matrices of special structure. A <u>diagonal</u> matrix $D = \operatorname{diag}(d_1, d_2, \ldots d_n), d_i \in F, i = 1, \ldots, n$, has zero elements except on its main diagonal, i.e.

$$D_{ij} = d_i \delta_{ij} \quad 1 \le i, j \le n \tag{2.15}$$

where here and below, δ is the Kroneker delta,

$$\delta_{ii} = 1, \delta_{ij} = 0 \text{ for } i \neq j. \tag{2.16}$$

We shall also use a class of elementary matrices. The term elementary matrix is used somewhat differently in different texts, and we shall use a restricted class of such matrices here. Herein, an $n \times n$ elementary matrix, is of the form E_{kma} where $1 \le k, m \le n, k \neq m$, and $a \in F$. This matrix has diagonal elements equal to 1, its km element is equal to a, and it is 0 everywhere else,

$$(E_{kma})_{ij} = \delta_{ij} + a\delta_{ik}\delta_{jm} \quad 1 \le i, j \le n \tag{2.17}$$

From (2.17), it is easily verified that

$$\det(E_{kma}) = 1 \text{ and } (E_{kma})^{-1} = E_{km(-a)}. \tag{2.18}$$

Multiplication of a given $n \times n$ matrix by a diagonal or elementary matrix is especially simple. Given an $n \times n$ matrix A, let $r_i, i = 1, \ldots, n$, denote the i-th row of A, and let $c_i, i = 1, \ldots, n$ denote the i-th column of A. We can think of r_i as a row vector of dimension n, and c_i as a column vector of dimension n. Furthermore, the matrix A is just the r_i assembled rowwise, or the c_i assembled columnwise,

$$A = \begin{pmatrix} r_1 \\ r_2 \\ \vdots \\ r_n \end{pmatrix} = \begin{pmatrix} c_1 c_2 \ldots c_n \end{pmatrix} \tag{2.19}$$

or we can use the notation

$$A_{ij} = (r_i)_j = (c_j)_i \quad , \quad 1 \le i, j \le n. \tag{2.20}$$

Then with $D = \operatorname{diag}(d_1, d_2, \ldots, d_n)$ as above, we have

$$\begin{aligned} &(DA)_{ij} = d_i A_{ij} = d_i (r_i)_j = (d_i r_i)_j \\ &(AD)_{ij} + A_{ij} d_j = (c_j)_i d_j = (d_j c_j)_i \quad 1 \le i, j \le n \end{aligned} \tag{2.21}$$

left multiplication of A by D multiplies each row r_i by the corresponding d_i, while right multiplication of A by D multiplies each column c_j of A by the corresponding d_j.

Next we compute

$$\begin{aligned} (E_{kma}A)_{ij} &= \sum_{\ell=1}^{n} (E_{kma})_{i\ell} A_{\ell j} \\ &= A_{ij} + a \sum_{\ell=1}^{n} \delta_{ik}\delta_{\ell m} A_{\ell j} \\ &+ A_{ij} + a\delta_{ik} A_{mj} \\ &= (r_i)_j + a\delta_{ik}(r_m)_j \\ &= (r_i + a\delta_{ik} r_m)_j, \quad 1 \le i, j \le n; \end{aligned} \tag{2.22}$$

for $i \neq k$, this is just A_{ij}, but

$$(E_{kma}A)_{kj} = A_{kj} + aA_{mj}, \quad 1 \leq j \leq n. \tag{2.23}$$

Left multiplication of A by E_{kma} simply adds $a \in F$ times the m-th row of A to the k-th row of A, leaving the other rows unchanged. Similarly

$$\begin{aligned}
(AE_{kma})_{ij} &= \sum_{\ell} A_{i\ell}(E_{kma})_{\ell j} \\
&= A_{ij} + a \sum_{\ell} A_{i\ell}\delta_{\ell k}\delta_{mj} \\
&= A_{ij} + aA_{ik}\delta_{mj} \\
&= (c_j)_i + a\delta_{mj}(c_k)_i \\
&+ (c_j + a\delta_{mj}c_k)_i \quad , 1 \leq i, j \leq n.
\end{aligned} \tag{2.24}$$

For $j \neq m$, this is just A_{ij}, but

$$(AE_{kma})_{im} = A_{im} + aA_{ik}, \qquad 1 \leq i \leq n, \tag{2.25}$$

so right multiplication of A by E_{kma} adds a times the k-th column of A to the m-th column of A, leaving the other columns unchanged. In the "notation" of (2.19)

$$DA = \begin{pmatrix} d_1 r_1 \\ d_2 r_2 \\ \vdots \\ d_n r_n \end{pmatrix}, AD = \begin{pmatrix} d_1 c_1 \; d_2 c_2 \ldots d_n c_n \end{pmatrix} \tag{2.26}$$

$$E_{kma}A = \begin{pmatrix} r_1 \\ r_k + ar_m \\ r_n \end{pmatrix}, AE_{kma} = \begin{pmatrix} c_1 \ldots c_m + ac_k \ldots c_n \end{pmatrix} \tag{2.27}$$

Lemma 2.4. *For any $n \times n$ matrix A, diagonal matrix D, and elementary matrix E_{kma},*

$$\det(AD) = \det(DA) = \det A \det D \tag{2.28}$$

$$\det(AE_{kma}) = \det(E_{kma}A) = \det A \tag{2.29}$$

Proof. We use the expressions

$$\det A = \sum_{\sigma\in P_n} N_\sigma \prod_{i=1}^{n} A_{i\sigma(i)} \tag{2.30}$$

$$\det A = \sum_{\sigma=P_n} N_\sigma \prod_{i=1}^{n} A_{\sigma(i)i} \tag{2.31}$$

where P_n is the permutation group of order n, and

$$N_\sigma = \prod_{1\le j<i\le n} \frac{i-j}{\sigma(i)-\sigma(j)}. \tag{2.32}$$

Then using (2.21) and (2.30)

$$\begin{aligned}
\det(DA) &= \sum_{\sigma\in P_n} N_\sigma \prod_{i=1}^{n} (DA)_{i\sigma(i)} \\
&= \sum_{\sigma\in P_n} N_\sigma \prod_{i=1}^{n} d_i A_{i\sigma(i)} \\
&= \Big(\sum_{\sigma\in P_n} N_\sigma \prod_{i=1}^{n} A_{i\sigma(i)}\Big)\Big(\prod_{i=1}^{n} d_i\Big) \\
&= (\det A)(\det D) \qquad (2.33)
\end{aligned}$$

The other half of (2.28) is proved in the same manner.

To compute $\det(E_{kma}A)$, we use (2.22), (2.23),

$$\begin{aligned}
\det(E_{kma}A) &= \sum_{\sigma\in P_n} N_\sigma \prod_{i=1}^{n} (A_{i\sigma(i)} + a\delta_{ik}A_{m\sigma(i)}) \\
&= \sum_{\sigma\in P_n} N_\sigma (\prod_{\substack{i=1\\ i\neq k}}^{n} A_{i\sigma(i)})(A_{k\sigma(k)} + aA_{m\sigma(k)}) \\
&= \sum_{\sigma\in P_n} N_\sigma [\prod_{i=1}^{n} A_{i\sigma(i)} + a(\prod_{\substack{i=1\\ i\neq k}}^{n} A_{i\sigma(i)})A_{m\sigma(k)}] \\
&= \det A + \det \tilde{A}_k^m \qquad (2.34)
\end{aligned}$$

where $\tilde{A}_k^m$ is the matrix A except that the k-th row is replaced by the m-th. Thus the k-th and m-th rows of $\tilde{A}_k^m$ coincide, and $\det \tilde{A}_k^m = 0$, proving one of the statements (2.29). The other is proved similarly.

Next we prove a representation theorem for any $n \times n$ matrix A, in terms of diagonal and elementary matrices.

Theorem 2.5. *Any $n \times n$ matrix A may be written*

$$A = \Big(\prod_{i=1}^{M} E_{k_i m_i a_i}\Big) D \Big(\prod_{i=M+1}^{N} E_{k_i m_i a_i}\Big) \tag{2.35}$$

where D is a diagonal $n \times n$ matrix, the $E_{k_i m_i a_i}$ are elementary matrices, and $N \leq n^2$.

Remark. The representation (2.35) is not unique; there are in fact an infinite number of such representations. But from lemma 2.4, (2.29) in particular, it follows that

$$\det A = \det D \tag{2.36}$$

for any such representation.

Proof. We construct a finite sequence of matrices, $A^{(1)}, A^{(2)}, \ldots, A^{(n)}$. We take $A^{(1)} = A$, and identify $A^{(n)}$, which will be diagonal, as D. For $\ell = 2, \ldots, n$, we construct $A^{(\ell)}$ from $A^{(\ell-1)}$ by no more than $2(n-\ell)+3$ left or right multiplications by elementary matrices.

For $\ell > 1$, $A^{(\ell)}$ is partially diagonal, that is

$$A^{(\ell)}_{ij} = 0 \text{ if } i \neq j \text{ and either } i < \ell \text{ or } j < \ell. \tag{2.37}$$

Given $A^{(\ell)}$, we proceed as follows:

First, suppose that both the ℓ-th row and the ℓ-th column of $A^{(\ell)}$ vanish identically. Then we can simply take $A^{(\ell+1)} = A^{(\ell)}$ and (2.37) remains satisfied.

Next, suppose that $A^{(\ell)}_{\ell\ell} \neq 0$. Then take $B^{(\ell)} = A^{(\ell)}$ and proceed as below.

If $A^{(\ell)}_{\ell\ell} = 0$, then either $A^{(\ell)}_{\ell j} \neq 0$ for some $j > \ell$ or else $A^{(\ell)}_{i\ell} \neq 0$ for some $i > \ell$. Take $B^{(\ell)} = A^{(\ell)} E_{j\ell 1}$ in the first case and $B^{(\ell)} = E_{\ell i 1} A^{(\ell)}$ in the second. However determined, $B^{(\ell)}$ satisfies (2.37) as well and $B^{(\ell)}_{\ell\ell} \neq 0$.

Given $B^{(\ell)}$, we make the off-diagonal elements of the ℓ-th row and the ℓ-th column equal to zero by elementary operations, as in ordinary Gaussian elimination. Specifically, we find

$$A^{(\ell+1)} = \Big(\prod_{\substack{i>\ell \\ B^{(\ell)}_{i\ell} \neq 0}} E_{i\ell\beta_i}\Big) B^{(\ell)} \Big(\prod_{\substack{j>\ell \\ B^{(\ell)}_{\ell j} \neq 0}} E_{\ell j \gamma_j}\Big) \tag{2.38}$$

where $\beta_i = -B^{(\ell)}_{i\ell}/B^{(\ell)}_{\ell\ell}$ and $\gamma_j = -B^{(\ell)}_{\ell j}/B^{(\ell)}_{\ell\ell}$. The matrix $A^{(\ell+1)}$ so determined satisfies (2.37) as required.

By successive application of this algorithm, we arrive at $A^{(n)}$

$$A^{(n)} = \Big(\prod_{i=1}^{M} E_{p_i q_i z_i}\Big) A \Big(\prod_{i=M+1}^{N} E_{p_i q_i z_i}\Big); \tag{2.39}$$

from (2.37), it follows that $A^{(n)}$ is diagonal, so we can identify $A^{(n)} = D$. Now multiplying (2.39) successively by the $E^{-1}_{p_i q_i z_i}$, using (2.18), we obtain (2.35). This completes the proof.

Corollary 1. *Right inverses are also left inverses (and vice versa), i.e. if A, B are $n \times n$ matrices, $AB = I$ if and only if $BA = I$.*

Proof. Assume $AB = I$; and using (2.35)

$$A = \prod_{i=1}^{M} E_i D \prod_{i=M+1}^{N} E_i, \tag{2.40}$$

where here and below, we use the abbreviation $E_i = E_{k_i m_i a_i}$ where no ambiguity is possible. The multiplying successively by the E_i^{-1} and D^{-1} we find

$$B = \prod_{j=M+1}^{N} E^{-1}_{N+M+1-j} D^{-1} \prod_{j=1}^{M} E^{-1}_{M+1-j} \tag{2.41}$$

and one easily checks that $BA = I$. The same argument holds, assuming $BA = I$ initially.

Corollary 2. *For any $n \times n$ matrices A, B,*

$$\det(AB) = (\det A)(\det B) \tag{2.42}$$

This follows immediately from (2.35), (2.28), and (2.29).

We shall also use Theorem 2.5 to simplify a discussion of the rank of a matrix. For any matrix A, the rank of A, denoted here by $r(A)$, is simply $\dim R(A)$. We shall denote by $\gamma(A)$ the dimension of the kernel of A, i.e. $\dim\{x \in H | Ax = 0\}$. These definitions apply whether the matrix A is square or not, but we can simplify the discussion by first assuming square matrices and then extending the results.

Lemma 2.6. *Let $U \subset H, H$ of dimension n, and Z a nonsingular $n \times n$ matrix. Then*

$$\dim(ZU) = \dim U. \tag{2.43}$$

Proof. Obviously $\dim(ZU) \leq \dim U$. But if $\dim(ZU) < \dim U$, since U is itself a linear space, there exists $u \in U, u \neq 0$, such that $Zu = 0$ contradicting the fact that Z is nonsingular.

Theorem 2.7. *For any $n \times n$ matrix A, and any representation of A of the form (2.35),*

$$r(A) = \text{the number of nonzero diagonal elements of } D; \tag{2.44}$$

$$\gamma(A) = \text{the number of zero diagonal elements of } D. \tag{2.45}$$

Furthermore, $r(A)$ is the number of linearly independent rows of A and also the number of linearly independent columns of A.

As a corollary of (2.44), (2.45), we have

$$r(A) + \gamma(A) = n \tag{2.46}$$

Proof.

$$\begin{aligned} r(A) &= \dim(AH) \\ &= \dim\Big(\prod_{i=1}^{M} E_i D \prod_{i=M+1}^{N} E_i H\Big) \\ &= \dim\Big(\prod_{i=1}^{M} E_i DH\Big) \\ &= \dim(DH) \end{aligned}$$

using lemma 2.6 in the last step. Obviously $\dim(DH)$ is the number of nonzero diagonal elements of D.

Next let $U = \ker A$, and

$$V = \prod_{i=M+1}^{N} E_i U \tag{2.47}$$

From (2.35) it is clear that $AU = 0$ if and only if $DV = 0$. Clearly the dimension of V is the number of zero diagonal elements of D. And from lemma 2.6, $\dim V = \dim U = \gamma(A)$, so (2.45) holds as well.

Now from (2.22) and (2.24), or perhaps more clearly form (2.27), it follows that left or right multiplication of a matrix by an elementary matrix does not change the number of linearly independent rows or columns. Since D has $r(D) = r(A)$ linearly independent rows and columns, so does A.

Corollary 3. *For any $n \times n$ matrix $A, r(A^\dagger) = r(A^t) = r(A)$*

$$\text{and } \gamma(A^\dagger) = \gamma(A^t) = \gamma(A). \tag{2.48}$$

These results extend easily to non-square matrices. Let A be an $m \times n$ matrix, i.e. $A : H_n \to H_m$. Let $\tilde{A}$ be a square matrix of dimension maximum (m, n), obtained by adding rows or columns of zeros to A. For $n > m$, we add $n - m$ rows of zeros to the bottom of A to obtain $\tilde{A}$, which is an $n \times n$ matrix, mapping H_n to H_n. Clearly $r(A) = r(\tilde{A}), \gamma(A) = \gamma(\tilde{A})$, and $r(\tilde{A}) + \gamma(\tilde{A}) = n$ from (2.46) so $r(A) + \gamma(A) = n$. For $m > n$, we add $m - n$ columns of zeros to A to form $\tilde{A}$; $\tilde{A}$ is $m \times m$. Again $r(A) = r(\tilde{A})$, but now $\tilde{A} : H_m \to H_m$ and $\gamma(\tilde{A}) = \gamma(A) + m - n$. Now $r(\tilde{A}) + \gamma(\tilde{A}) = m$, so again (!) $r(A) + \gamma(A) = n$. Thus we have proved

Corollary 4. *For any $m \times n$ matrix A,*

$$r(A) + \gamma(A) = n. \tag{2.49}$$

We note that $r(A) \leq$ minimum (m, n).

The rank of the product of two matrices is not uniquely determined by the rank of each factor, but satisfies an important inequality.

Theorem 2.8 (Sylvester's theorem). *: Let A be an $m \times n$ matrix and B an $n \times k$ matrix, so AB is defined and is an $m \times k$ matrix. Then*

$$r(A) + r(B) - n \leq r(AB) \leq \textit{minimum } (r(A), r(B)) \tag{2.50}$$

Proof. $A : H_n \to H_m, B : H_k \to H_n$, and $(AB) : H_k \to H_m$.

$$\begin{aligned} r(AB) &= \dim(AB)H_k \\ &= \dim(A(BH_k)) \\ &= \dim(AR(B)) \end{aligned} \tag{2.51}$$

Obviously $\dim(AR(B)) \leq \dim(AH_n) = r(A)$, since $R(B) \subset H_n$. And $\dim(AR(B)) \leq \dim R(B) = r(B)$. Using (2.51) this proves the upper bound in (2.50).

To get the lower bound for $r(AB)$, we note that

$$\ker(AB) = \{x \in H_k | Bx \in \ker(A)\} \tag{2.52}$$

so that

$$\gamma(AB) = \dim\ker(AB)$$
$$= \dim(R(B) \cap \ker(A)) + \dim\ker B. \tag{2.53}$$

As $\dim(R(B) \cap \ker(A)) \leq \dim\ker(A) = \gamma(A)$, (2.53) becomes

$$\gamma(AB) \leq \gamma(A) + \gamma(B). \tag{2.54}$$

Applying (2.49) to A, B and AB, we find

$$r(A) + \gamma(A) = n, r(B) + \gamma(B) = k, r(AB) + \gamma(AB) = k; \tag{2.55}$$

combining (2.54) and (2.55) gives the lower bound in (2.50).

We continue with a brief discussion of the transformation of vectors and matrices under a change of basis. Let $\{e_i\}_{i=1}^n$ be a given basis for V^n, a linear vector space of dimension n. Let $\{e'_j\}_{j=1}^n$ be an alternative basis, which necessarily can be expressed by

$$e'_j = \sum_{i=1}^{n} T_{ij} e_i \tag{2.56}$$

with some $n \times n$ matrix T thus determined. Indeed, T is obtained by assembling columnwise the vectors e'_j; expressed with respect to the basis $\{e_i\}$. The vectors representing the e'_j must be linearly independent, so T is necessarily nonsingular.

Now a vector $v \in V$ can be represented with respect to either basis, for example

$$v = \sum_{i=1}^{n} p_i e_i = \sum_{j=1}^{n} q_j e'_j; \tag{2.57}$$

equivalently, we can say that v is represented with respect to the basis $\{e_i\}$ by p and with respect to $\{e'_j\}$ by q, with

$$p = \begin{pmatrix} p_i \\ \vdots \\ p_n \end{pmatrix}, q = \begin{pmatrix} q_1 \\ \vdots \\ q_n \end{pmatrix}. \tag{2.58}$$

The relation between p and q, that is the transformation of the vector v under this change of basis, is readily obtained from (2.56), (2.57),

$$
\begin{aligned}
v = \sum_{i=1}^{n} p_i e_i &= \sum_{j=1}^{n} q_j e'_j \\
&= \sum_{j=1}^{n} q_j \sum_{i=1}^{n} T_{ij} e_i \\
&= \sum_{i=1}^{n} (\sum_{j=1}^{n} T_{ij} q_j) e_i \qquad (2.59)
\end{aligned}
$$

so clearly

$$Tq = p, \quad q = T^{-1}p. \tag{2.60}$$

This expresses the transformation of vectors, under a change of basis. Next let A be an $n \times n$ matrix. Under this change of basis, p is transformed to $T^{-1}p$ and Ap is necessarily transformed to $T^{-1}Ap = (T^{-1}AT)(T^{-1}p)$, so if A is the representation of a given linear operator on V with respect to the basis $\{e_i\}$, the representation with respect to the basis $\{e'_j\}$ is $T^{-1}AT$. This is often called a similarity transformation of a matrix.

Finally, let $B : V^n \to V^m$ be a transformation from V^n to V^m, of dimension m. With respect to the basis $\{e_i\}_{i=1}^n$ for V^n and $\{\beta_i\}_{i=1}^m$ for V^m, B is represented by on $m \times n$ matrix, also denoted by B. Now suppose we wish to change basis in (either or) both spaces, via transformations (2.56) for V^n and a similar relation $\beta'_j = \sum\limits_{i=1}^{m} Q_{ij}\beta_i$ for V^m with some nonsingular $m \times m$ matrix Q. For any $p \in V^n, Bp \in V^m; p$ transforms to $T^{-1}p$ and Bp to $Q^{-1}Bp = Q^{-1}BT(T^{-1}p)$ under these basis changes, so B transforms to $Q^{-1}BT$. This is the generalization of a similarity transformation to such non-square matrices.

A common example of such a change of basis occurs when one wishes an orthonormal basis for a Hilbert space H^n, i.e. basis functions $\{e'_j\}_{j=1}^n$ satisfying

$$(e'_j, e'_k) = \delta_{jk}, 1 \leq j, k \leq n. \tag{2.61}$$

Given any basis $\{e_i\}_{i=1}^n$, one can easily compute such an orthonormal basis by the Gram-Schmidt procedure,

$$
\begin{aligned}
&e_1' = e_1/\|e_1\| \\
&\alpha_j = e_j - \sum_{i=1}^{j-1}(e_j, e_i')e_i',\ e_j' = \alpha_j/\|\alpha_j\|, j = 2, \ldots, n
\end{aligned}
\tag{2.62}
$$

In (2.62), none of the α_j can be zero, for this would contradict linear independence of the $\{e_i\}$.

A change of basis is often used to transform some given square matrix into an attractive structural form, e.g. diagonal. Not all matrices, however, can be so "diagonalized" by a similarity transformation, as we shall discuss in the following chapter. Furthermore, a real matrix A may have complex eigenvalues, and thus be similar to a complex diagonal matrix via a complex similarity transformation; in some applications, this is quite inconvenient.

We are thus motivated to introduce a broader class of matrices than diagonal, that of block diagonal matrices. An $n \times n$ matrix A is block diagonal, with "diagonal blocks" $b_1, \ldots, b_m$ if there exist integers α_k and blocks b_k

$$
0 = \alpha_0 < \alpha_1 < \alpha_2 < \cdots < \alpha_m = n, \quad b_k = \{\alpha_{k-1}+1, \cdots, \alpha_k\} \quad k = 1, \cdots, m \tag{2.63}
$$

such that

$$
A_{ij} = 0 \quad \text{for all } i, j, i \in b_k, \quad j \neq b_k \quad \text{for some } k, \tag{2.64}
$$

i.e. A_{ij} is zero unless i and j are in the same diagonal block.

The extreme examples of block diagonal matrices correspond to diagonal matrices, i.e. $m = n, \alpha_k = k$ for all k in (2.63), (2.64), and "full" matrices, with $m = 1$ in (2.63) and (2.64) vacuous. We shall see in chapter 3 that the Jordan canonical form of a matrix (not similar to a diagonal matrix) is in a convenient block diagonal form, and that real matrices can often be transformed to convenient block diagonal structure by real similarity transformations.

The attractiveness of block diagonal structure of a matrix is that for many purposes, the individual blocks can be treated separately and the analysis considerably simplified. the simplification results essentially from the following lemma:

Lemma 2.9. *Suppose a square matrix A is block diagonal. Then so are $A^\dagger, A^p$ for any positive integer p, and A^{-1} (if A is nonsingular) with the same diagonal blocks as for A.*

The proof follows from the basic rule for matrix multiplication, and is left as an exercise. Furthermore, if B has the same block diagonal structure as A, then so do the products AB and BA.

The block structure of matrices, square matrices in particular, is used more generally. Let A be an $n \times n$ matrix and blocks b_k as in (2.63). Then we may regard A as a matrix or assembly of its blocks $[A]_{k\ell}, 1 \leq k, \ell \leq m$, where each block $[A]_{k\ell}$ is a matrix of dimension $\dim b_k \times \dim b_\ell$

$$[A]_{k\ell} = \{A_{ij}\} \; i \in b_k, \; j \in b_\ell. \tag{2.65}$$

If two $n \times n$ matrices A, B are so partitioned, with the same blocks b_k, then matrix multiplication holds in block form

$$[AB]_{k\ell} = \sum_{j=1}^{m} [A]_{kj}[B]_{j\ell} \quad , \; 1 \leq k, \ell \leq m. \tag{2.66}$$

In this notation, a matrix A is block diagonal if and only if $[A]_{k\ell} = 0$ for $k \neq \ell$.

Let G_n denote the set of nonsingular $n \times n$ matrices containing only n nonzero elements. Obviously, such matrices will have only one nonzero element in any one row or any one column, and the determinant expansion of such a matrix will have only one nonvanishing term. Nonsingular diagonal matrices are special examples of elements of G_n.

The set G_n is obviously not a linear vector space, but it is a (noncommutative) multiplicative group. In particular

Lemma 2.10. *Let $A, B \in G_n$. Then*

$$AB, BA \in G_n \tag{2.67}$$

(2.68)
$A^\dagger \in G_n, \quad A^\dagger A \quad$ *and* $AA^\dagger \quad$ *are diagonal with strictly positive diagonal elements*

$$A^{-1} \in G_n. \tag{2.69}$$

Proof. (2.67) and (2.68) are verified immediately from the rule for matrix multiplication. In particular, for $A \in G_n$ and i, j such that $A_{ij} \neq 0$, then $(AA^\dagger)_{ii} = (A^\dagger A)_{jj} = |A_{ij}|^2$. To show (2.69), we use $A^{-1} = (A^\dagger A)^{-1} A^\dagger$, which is obviously true for any nonsingular matrix A, not just those in G_n. However, for $A \in G_n, A^\dagger A$ is diagonal, so $(A^\dagger A)^{-1}$ is also diagonal, and in G_n, and thus so in the product $(A^\dagger A)^{-1} A^\dagger$, using (2.67). Indeed, for $A \in G_n$ and i, j such that $A_{ij} \neq 0, \quad A^{-1}_{ji} = 1/A_{ij}$.

Problems

1. Let $p, q, r, s \in \mathbb{R}^3$, and let the 3×3 matrix A be given by $A = pq^t + rs^t$, i.e. $A_{ij} = p_i q_j + r_i s_j$ for $i, j = 1, 2, 3$. Prove that A is singular. Generalize this result to $\mathbb{R}^n$.

2. Let $p, q, r \in \mathbb{R}^n$ be given, and $x \in \mathbb{R}^n$ determined from $x + p(p, x) - q(q, x) = r$. What condition(s) on p, q, r are necessary so that x is uniquely determined? Find x explicitly in this case.

3. Given $A : \mathbb{R}^4 \to \mathbb{R}^3$, given by $A = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 1 & -1 & -1 & 1 \\ 1 & 1 & 3 & -1 \end{pmatrix}$; find $\ker(A), R(A)$, $\ker(A^t), R(A^t)$.

4. Given $A = \begin{pmatrix} 0 & 1 & -1 \\ 1 & 3 & 1 \\ 0 & 0 & 0 \end{pmatrix}$; find a diagonal matrix D and elementary matrices $E_1, \ldots, E_N$ such $A = \prod\limits_{i=1}^{M} E_i D \prod\limits_{i=M+1}^{N} E_i$. Also find a second example of a representation of A in this form.

5. A square matrix A is called symmetric if it is real and $A^t = A$, and, more generally, Hermitian if $A = A^\dagger$. Let H be a complex, finite-dimensional Hilbert space ($F = \mathbb{C}, H = \mathbb{C}^n$) and A an $n \times n$ matrix such that (h, Ah) is always real, i.e. $(h, Ah) \in \mathbb{R}$ for all $h \in H$. Prove that this happens if and only if A is Hermitian.

6. Find an example of a matrix A which does <u>not</u> admit a representation of the form $\prod\limits_{i=1}^{N} E_i D$, where the E_i are elementary matrices and D is diagonal.

7.* Show that if a (square) matrix A is nonsingular, then it does admit such "one-sided" representations, i.e. there exist E_i, D such that $A = \prod\limits_{i=1}^{N} E_i D$ (or $A = D \prod\limits_{i=1}^{N} E_i$).

8. Given $A = \begin{pmatrix} 0 & 1 & -1 \\ 1 & 3 & 1 \\ 0 & 0 & 0 \end{pmatrix}$, $B = \begin{pmatrix} 1 & -2 & -3 \\ 0 & 1 & 2 \\ -2 & 1 & 0 \end{pmatrix}$; find $r(AB)$ and $r(BA)$.

9. True or false: If A, B are $n \times n$ matrices, then $r(AB) = r(BA)$. Either prove this or find a counterexample.

10. Let A be an $m \times n$ matrix, and B an $n \times m$ matrix, with $m > n$. True or false: AB is always singular. Either prove this or find a counterexample.

11. True or false: If A is an $n \times n$ matrix invariant under similarity transformations, i.e. $T^{-1}AT = A$ for all nonsingular $n \times n$ T, then $A = \alpha I, \alpha \in F, I$ the $n \times n$ identity matrix. Prove or find a counterexample.

12*. Show that the space of $n \times n$ matrices, equipped with the norm $\|A\| = \sup\limits_{\substack{h \in H^n \\ h \neq 0}} \|Ah\|/\|h\|$, is a Banach space. You may choose whichever equivalent norm you wish for H^n.

13. Prove Gershgorin's theorem: Let A be an $n \times n$ matrix, and for $i = 1, \cdots, n$, let ρ_i be a closed circle in the complex plane, $\rho_i = \{z \in \mathbb{C} \Big| |z - A_{ii}| \leq \sum_{j \neq i} |A_{ij}|\}$. Show that all of the eigenvalues of A lie in the region $\bigcup\limits_{i=1}^{n} \rho_i$.

14. For vectors in H, of dimension n, let $\|\cdot\|$ denote the ℓ_2 norm given by (1.5), and let the norm of an $n \times n$ matrix be as shown in problem 12. Let $\sigma(A) = \underset{i}{\text{maximum}} |\lambda_i(A)|$, where $\lambda_i(A)$ are the eigenvalues of A_j. $\sigma(A)$ is called the <u>spectral radius</u> of A. Prove the following:

    (a) $\sigma(A) \leq \|A\|$

    (b) $\|AB\| \leq \|A\| \quad \|B\|$, for any $n \times n$ matrices A, B

    (c) $\|A\| \quad \|A^{-1}\| \geq 1$, for any nonsingular A. The product $\|A\| \quad \|A^{-1}\|$ is sometimes called the <u>condition number</u> of A, it is a measure of how close A is to being singular.

15. Prove lemma 2.9.

16. Let H be of dimension 5, spanned by $\{v, w, x, y, z\}$, and let A be a linear operator on H satisfying $Av = w - y, \quad Aw = v + 2z, \quad Ax = v - z, \quad Ay = w - x - y, \quad Az = -2x$. Find a matrix representation for A with respect to each of the following bases for H: (1) $\{v, w, x, y, z\}$; (2) $\{w, v, x, y, z\}$; (3) $\{v + w, v - w, x + y, x - y, v - x + z\}$.

17. Let $A : H_m \to H_n$ be represented by an $n \times m$ matrix, where $H_m = \mathbb{C}_m$ and $H_n = \mathbb{C}_n$, but with inner products

$$(p, x)_n = \sum_{i=1}^{n} \beta_i p_i \bar{x}_i, \;\; (q, y)_m = \sum_{i=1}^{m} \gamma_i q_i \bar{y}_i$$

where the β_i, γ_i are given positive (real) numbers. Determine a formula for $A^\dagger$ such that (2.11) holds for all $x \in H_n, y \in H_m$.

18. Let $X = \mathbb{R}^2$ with the inner product $(p, q)_x = 4p_1q_1 + p_1q_2 + p_2q_1 + p_2q_2$

    a) Show this is indeed an inner product for X.

    b) Given $A : X \to X$, find an expression for $A^\dagger$ such that $(Ap, q)_x = (p, A^\dagger q)_x$ for all $p, q \in X$.

19. Given nonzero $p, q \in \mathbb{R}^n$, and a real $n \times n$ matrix B satisfying $r \cdot Br > 0$ for all nonzero $r \in \mathbb{R}^n$ satisfying $r \cdot p = 0$. Let A be an $(n+1) \times (n+1)$ matrix with block structure

$$A = \begin{pmatrix} B & p + Bq \\ p^\dagger & p \cdot q \end{pmatrix}$$

    Prove that A is nonsingular.

20. Prove that $r(AA^\dagger) = r(A^\dagger A)$ for any matrix A, whether or not a square matrix.

3 Invariant subspaces and Jordan form of a matrix

Throughout this chapter, $F = \mathbb{C}$ for definiteness and simplicity.

Given an $n \times n$ matrix $A : H \to H, H$ of dimension n, it is natural to ask for what values of $\lambda \in \mathbb{C}$ is the matrix $A - \lambda I$ singular. Here I is the $n \times n$ identity matrix; hereafter we shall write simply $A - \lambda$ for $A - \lambda I$, etc.

The matrix $A - \lambda$ is singular if and only if λ satisfies

$$P_A(\lambda) = 0 \tag{3.1}$$

where

$$P_A(\lambda) = \det(A - \lambda) \tag{3.2}$$

is the characteristic polynomial of A. Obviously $P_A(\lambda)$ is a polynomial of order exactly n,

$$P_A(\lambda) = (-\lambda)^n + (trA)(-\lambda)^{n-1} + \cdots + \det A \tag{3.3}$$

where $trA = \sum_{i=1}^{n} A_{ii}$ is called the trace of A. Clearly $P_A(\lambda)$ can be factored,

$$P_A(\lambda) = \prod_{i=1}^{k} (\lambda_i - \lambda)^{m_i} \tag{3.4}$$

where the $\lambda_i \in \mathbb{C}$ are the distinct eigenvalues of A, satisfying $\lambda_i \neq \lambda_j$ for all $i \neq j$, and the positive integer m_i the multiplicity of λ_i. Since the λ_i are the roots of (3.1), we have from (3.3)

$$\prod_{i=1}^{k} \lambda_i^{m_i} = \det A, \sum_{i=1}^{k} m_i \lambda_i = trA, \sum_{i=1}^{k} m_i = n \tag{3.5}$$

A closely related question is that of the minimal polynomial M_A of A, the polynomial of lowest degree such that

$$M_A(A) = 0 \tag{3.6}$$

(The zero in the right side of (3.6) means the ideatically zero $n \times n$ matrix, since $M_A(A)$ is obviously an $n \times n$ matrix.)

The minimal polynomial M_A is unique up to normalization (multiplication by a constant) and divides every polynomial Q such that $Q(A) = 0$.

The polynomial M_A is of order no greater than n^2, as A has only n^2 elements; we shall see that in fact it is of order no greater than n. The polynomial M_A can also be factored, and we claim

Lemma 3.1.

$$M_A(\lambda) = \prod_{i=1}^{k} (\lambda_i - \lambda)^{\mu_i} \tag{3.7}$$

where the μ_i are positive integers and λ_i the eigenvalues of A.

Proof. Suppose $M_A(A)$ contained a factor $\xi - A$, where ξ is not one of the eigenvalues. Then $\xi - A$ is nonsingular, and since all of the factors of $M_A(A)$ commute, we can write $M_A(A) = B(\xi - A)$, where the matrix B, so determined, is just the product of the remaining factors.

As $\xi - A$ is nonsingular, $R(\xi - A) = H$, and if $M_A(A)H = 0$, then $B(\xi - A)H = BH = 0$, so B is identically zero. Thus M_A is not the minimal polynomial of A, since B is also a polynomial, of order are less.

Furthermore, every eigenvalue λ_i appearing in (3.4) must also appear in (3.7), with some positive μ_i. Suppose not, suppose (3.7) contains no factors of $(\lambda_j - A)$. Since $A - \lambda_j$ is singular, there exists a nonzero vector x_j such that $(A - \lambda_j)x_j = 0$. And

$$M_A(A)x_j = \prod_{i \neq j} (\lambda_i - A)^{\mu_i} x_j = \prod_{i \neq j} (\lambda_i - \lambda_j)^{\mu_i} x_j \neq 0,$$

since $\lambda_i \neq \lambda_j$ for all $i \neq j$. This contradicts (3.6), proving our assertion.

Both the characteristic and minimal polynomials of a matrix are invariant under similarity transformations. Let A, B be $n \times n$ matrices, with $B = T^{-1}AT$ for some nonsingular T. Then $\det(B - \lambda I) = \det(T^{-1}AT - \lambda I) = \det(T^{-1}(A - \lambda I)T) = \det(A - \lambda I)$, so $P_B(\lambda) = P_A(\lambda)$. And $0 = M_B(B) = M_B(T^{-1}AT) = T^{-1}M_B(A)T$, since M_B is a polynomial, so $M_B(A) = 0$ and $M_A(B) = 0$. Thus if each is the minimal polynomial for its respective matrix, normalized, for example, as in (3.7), they must coincide.

Suppose not, i.e., suppose that $M_A(\lambda) \neq M_B(\lambda)$. We first assert that degree $M_A(\lambda)$ = degree $M_B(\lambda)$. If, for example, degree $M_A(\lambda)$ < degree $M_B(\lambda)$, then as $M_A(B) = 0, M_B$ is not the minimal polynomial for B. With M_A and M_B normalized, as in (3.7), for example, since they are of the same degree it follows that $M_A(\lambda) - M_B(\lambda)$ is a polynomial of degree at most one less than that of $M_A(\lambda)$ or $M_B(\lambda)$. Since $M_A(A) - M_B(A) = M_A(B) = M_B(B) = 0$, it follows that if $M_A(\lambda) - M_B(\lambda)$ is not identically zero, neither M_A nor M_B is the minimal polynomial for its respective matrix.

Often the eigenvalues of A are distinct, with $k = n$ and all the $m_i = 1$ in (3.4), (3.5). In this case, there are always n linearly independent eigenvectors. Thus A is always similar to a diagonal matrix, and the characteristic and minimal polynomials coincide, i.e. $P_A(\lambda) = M_A(\lambda)$. To justify these statements, assume that the $\lambda_i, i = 1, \ldots, n$ are distinct but that the corresponding eigenvectors $\{x_i\}$ are linearly dependent. Then possibly after a reordering of the λ_i, x_i there exist $\alpha_2, \ldots, \alpha_n \in \mathbb{C}$ such that

$$x_1 = \sum_{j=2}^{n} \alpha_j x_j$$

so that

$$\begin{aligned}
\prod_{i=2}^{n} (A - \lambda_i I) x_1 &= \prod_{i=2}^{n} (A - \lambda_i I) \sum_{j=2}^{n} \alpha_j x_j \\
&= \sum_{j=2}^{n} \alpha_j \prod_{i=2}^{n} (\lambda_j - \lambda_i) x_j \\
&= 0
\end{aligned}$$

whereas the left side is

$$\prod_{i=2}^{n} (\lambda_1 - \lambda_i) x_1 \neq 0$$

by the assumption of distinct eigenvalues.

So the $\{x_i\}$ are linearly independent, and the matrix T, the i-th column of which is $x_i,\ i = 1, \ldots, n$, is nonsingular. Then one easily verifies that

$$T^{-1} A T = \mathrm{diag}(\lambda_1, \ldots, \lambda_n).$$

But a simple example shows that this is not always the case: let

$$B = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}, G = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix};$$

then $P_B(\lambda) = P_G(\lambda) = (2 - \lambda)^2, M_B(\lambda) = 2 - \lambda$ but $M_G(\lambda) = (2 - \lambda)^2$.

Let

$$U_{ij} = \ker(A-\lambda_i)^j, \text{ for } i=1,\dots,k; j=1,\dots,\mu_i. \tag{3.8}$$

The U_{ij} are subspaces of H, and from (3.8) we have

$$U_{i(j-1)} \subset U_{ij}, j>1, \text{ and } \dim U_{ij} \geq 1. \tag{3.9}$$

We give a proof of the strict inclusion in (3.9). First we show that $U_{i(\mu_i-1)} \subset U_{i\mu_i}$. If not, i.e. if these spaces coincide, set

$$V = \prod_{\substack{j=1\\ j\neq i}}^{k} (A-\lambda_j I)^{\mu_j} H.$$

Recalling M_A given by (3.7), it follows that $V \subseteq U_{i\mu_i}$ and thus $V \subseteq U_{i(\mu_i-1)}$ by assumption. But then

$$(A-\lambda_i I)^{\mu_i-1} V = (A-\lambda_i I)^{\mu_i-1} \prod_{\substack{j=1\\ j\neq i}}^{k} (A-\lambda_j I)^{\mu_j} H = 0,$$

contradicting the fact that M_A given by (3.7) is the minimal polynomial.

Thus there exists $x \in U_{i\mu_i}$, $x \notin U_{i(\mu_i-1)}$, and for any $j=1,\dots,\mu_i-2$, $y_j = (A-\lambda_i I)^j x$ satisfies $y_j \in U_{i(\mu_i-j)}$ $y_j \notin U_{i(\mu_i-j-1)}$.

Theorem 3.2. *The U_{ij} are invariant under multiplication by $A-\xi$, for any $\xi \in F$, i.e. $(A-\xi)U_{ij} \subset U_{ij}$. Furthermore, if $\xi \neq \lambda_i$ and $x \in U_{ij}, x \neq 0$, then $(A-\xi)x \neq 0$.*

Proof. For $z \in U_{ij}, y=(A-\xi)z$ satisfies

$$(A-\lambda_i)^j y = (A-\lambda_i)^j (A-\xi) z = (A-\xi)(A-\lambda_i)^j z = 0 \tag{3.10}$$

since these factors commute, so $y \in U_{ij}$.

The second statement is proved by contradiction. Suppose not, i.e. suppose there exists $x \neq 0, x \in U_{ij}$ such that $(A-\xi)x = 0$ with $\xi \neq \lambda_i$. Then

$$\begin{aligned}
0 &= (A-\lambda_i)^j x \\
&= (A-\xi+(\xi-\lambda_i))^j x \\
&= (A-\xi)^j x + j(A-\xi)^{j-1}(\xi-\lambda_i)x + \cdots + (\xi-\lambda_i)^j x \\
&= (\xi-\lambda_i)^j x
\end{aligned}$$

which is not zero, as $x \neq 0$ and $\xi \neq \lambda_i$. This contradiction proves the statement.

Theorem 3.3. *The subspaces $U_{i\mu_i}$ are linearly independent, i.e. if* $\sum_{i=1}^{k} \alpha_i x_i = 0$ *with each* $x_i \in U_{i\mu_i}, x_i \neq 0, \alpha_i \in F$, *then all of the* $\alpha_i = 0$.

Proof. First we show that any U_{i1} is linearly independent of the other $U_{j\mu_j}$. Suppose not, then there exists $x \in U_{i1}, x \neq 0$, such that

$$x = \sum_{j \neq i} \alpha_j x_j, x_j \in U_{j\mu_j}, \alpha_j \in F. \tag{3.11}$$

Denote by

$$M_i = \prod_{j \neq i} (\lambda_j - A)^{\mu_j} \tag{3.12}$$

then

$$M_i x_j = 0 \quad , \quad x_j \in U_{j\mu_j}, \tag{3.13}$$

for any $j \neq i$, using (3.8). But $x \in U_{i1}$ satisfies $(A - \lambda_i)x = 0$, and so by application of the second statement of Theorem 3.2, $M_i x \neq 0$. Thus applying M_i to both sides of (3.11) gives a contradiction, proving this result.

Now by induction, for some integer $\beta \geq 2$, suppose that $U_{i(\beta-1)}$ is linearly independent of the $U_{j\mu_j}$ but $U_{i\beta}$ is not. Then (3.11) holds for some $x \in U_{i\beta}, x \notin U_{i(\beta-1)}$. Set $y = (A - \lambda_i)x$. Since $x \notin U_{i(\beta-1)}, y \neq 0$, but as $x \in U_{i\beta}, y \in U_{i(\beta-1)}$. Applying $(A - \lambda_i)$ to both sides of (3.11), we find

$$\begin{aligned} y &= (A - \lambda_i)x \\ &= \sum_{j \neq i} \alpha_j (A - \lambda_i) x_j. \end{aligned} \tag{3.14}$$

From Theorem 3.2, $(A - \lambda_i)x_j \in U_{j\mu_j}$, so (3.14) now contradicts the induction hypothesis that $U_{i(\beta-1)}$ is linearly independent of the $U_{j\mu_j}$. This completes the proof.

Theorem 3.4.

$$\bigoplus_{i=1}^{k} U_{i\mu_i} = H \tag{3.15}$$

Proof. Denote by $X = \overset{k}{\underset{i=1}{\oplus}} U_{i\mu_i}$; clearly $X \subset H$, so it suffices to show that $\dim X = n$. We use $k-1$ applications of Sylvester's theorem, (2.54) in particular, to obtain, using (3.6), (3.7)

$$\begin{aligned} n &= \gamma(M_A(A)) \\ &= \gamma(\prod_{i=1}^{k}(\lambda_i - A)^{\mu_i}) \\ &\le \sum_{i=1}^{k} \gamma((\lambda_i - A)^{\mu_i}) \\ &= \sum_{i=1}^{k} \dim U_{i\mu_i}. \end{aligned} \tag{3.16}$$

From theorem 3.3, the spaces $U_{i\mu_i}$ are linearly independent, so

$$\dim X = \sum_{i=1}^{k} \dim U_{i\mu_i}, \tag{3.17}$$

and from (3.16), (3.17), we find $\dim X \ge n$. Thus $\dim X = n$, completing the proof.

From Theorem 3.4, it follows that if we can find a convenient representation for the matrix A on any such subspace $U_{i\mu_i}$, we will have such a representation for A on all of H. Thus our problem is reduced to finding a convenient basis for $U_{i\mu_i}$.

First we note that if $\mu_i = 1$, then A restricted to $U_{i\mu_i} = U_{i1}$ is just λ_i times the identity matrix, so any basis for U_{i1} may be used. Thus we consider the case $\mu_i > 1$.

Denote by $V_{i\nu}$, the subspaces determined by $V_{i1} = U_{i1}$,

$$V_{i\nu} = U_{i1} \cap R((A - \lambda_i)^{\nu-1}), \quad \nu = 2, \cdots, \mu_i; \tag{3.18}$$

since $(A - \lambda_i)^{\mu_i}$ appears in the minimal polynomial, $V_{i\mu_i}$ is not trivial (but $V_{i(\mu_i+1)}$ would be if defined similarly).

Clearly

$$V_{i\mu_i} \subseteq V_{i(\mu_i-1)} \subseteq \cdots \subseteq V_{i2} \subseteq V_{i1} = U_{i1};$$

we choose a basis for U_{i1} by first choosing a basis for $V_{i\mu_i}$, then adding additional elements to get a basis for $V_{i(\mu_i-1)}$, etc. Thus a set of vectors

$\{e^i_{j1}\}, j = 1, \cdots, \dim V_{i\nu}$, is a basis for $V_{i\nu}$, for any $\nu = 1, \cdots, \mu_i$. And given j, for $\nu = \nu(j)$ determined such that is $e^i_{j1} \in V_{i\nu}$ but $e^i_{j1} \notin V_{i(\nu+1)}$, that is

$$\dim V_{i(\nu+1)} < j \leq \dim V_{i\nu} \quad (\dim V_{i(\mu_i+1)} \equiv 0) \tag{3.19}$$

it follows from (3.18) that there is a nonzero vector $e^i_{j\nu} \in U_{i\mu_i}$ such that $(A - \lambda_i)^{\nu-1} e^i_{j\nu} = e^i_{j1}$. We determine a basis $\{e^i_{j\theta}\}_{\theta=1}^{\nu(j)}$ for a Jordan block Y^i_j, of dimension $\nu(j)$ obtained form (3.19), by

$$(A - \lambda_i) e^i_{j\theta} = e^i_{j(\theta-1)}, \theta = 2, \cdots, \nu(j), (A - \lambda_i) e^i_{j1} = 0, \tag{3.20}$$

alternatively written in the form

$$A \begin{pmatrix} e^i_{j1} \\ e^i_{j2} \\ \vdots \\ e^i_{j\nu} \end{pmatrix} = \lambda_i \begin{pmatrix} e^i_{j1} \\ e^i_{j2} \\ \vdots \\ e^i_{j\nu} \end{pmatrix} + \begin{pmatrix} 0 \\ e^i_{j1} \\ \vdots \\ e^i_{j(\nu-1)} \end{pmatrix}$$

or in block form, A restricted to Y^i_j with this basis, assumes the form

$$A|_{Y^i_j} = \begin{pmatrix} \lambda_i & & 1 & & 0 \\ & \ddots & & \ddots & \\ & & & & 1 \\ 0 & & & & \lambda_i \end{pmatrix} \tag{3.21}$$

which we call a Jordan block. We shall show that

$$U_{i\mu_i} = \bigoplus_{j=1}^{\dim U_{i1}} Y^i_j \tag{3.22}$$

so that $\{\{e^i_{j\theta}\}_{j=1}^{\nu(j)}\}_{j=1}^{\dim U_{i1}}$ is a basis for $U_{i\mu_i}$, and with respect to this basis, A restricted to $U_{i\mu_i}$ is just an assembly at Jordan blocks, of a form

$$A|_{U_{i\mu_i}} = \begin{pmatrix} \lambda_i & 1 & & & & & & \\ & & 1 & & & & 0 & \\ & & \lambda_i & 0 & & & & \\ & & & \lambda_i & 1 & & & \\ & & & & \ddots & 1 & & \\ & & & & & \lambda_i & 0 & \\ & 0 & & & & & \lambda_i & \\ & & & & & & & \lambda_i \end{pmatrix} \tag{3.23}$$

Next we justify (3.22). Clearly the right side of (3.22) is a subspace of $U_{i\mu_i}$ and contains U_{i1} as a subspace. Thus, suppose it contains $U_{i(\eta-1)}$ as a subspace but not $U_{i\eta}$, for some $\eta, 2 \leq \eta \leq \mu_i$. For arbitrary $x \in U_{i\eta}, y = (A-\lambda_i)^{\eta-1}x$ satisfies $y \in U_{i1}$ and $y \in R((A-\lambda_i)^{\eta-1})$, so $y \in V_{i\eta}$ from (3.18).

Then from (3.19), there exist constants τ_j such that

$$y = \sum_{j=1}^{\dim V_{i\eta}} \tau_j e^i_{j1}$$

and from (3.20), it follows that

$$x - \sum_j \tau_j e^i_{j\eta} \in U_{i(\eta-1)}$$

so that $\bigoplus_j Y^i_j$ also contains $U_{i\eta}$ as a subspace, proving our claim.

Finally extending to all of H, using theorem 3.4, we choose a basis $\{\{\{e^i_{j\theta}\}_{\theta=1}^{\dim Y^i_j}\}_{j=1}^{\dim U_{i1}}\}_{i=1}^k$, and the matrix A is put in (upper) Jordan canonical form. The diagonal elements are the eigenvalues λ_i; the off - diagonal elements are equal to zero, except that

$$\text{(3.24)} \qquad \text{if } A_{\ell\ell} = A_{(\ell+1)(\ell+1)}, \text{ then } A_{\ell(\ell+1)} \text{ may be 0 or 1.}$$

From this construction we see that the Jordan form of a matrix A is diagonal, i.e. A is similar to a diagonal matrix, if all of the μ_i appearing in (3.7) are equal to 1.

More generally, we see that the dimensions of the Jordan blocks are determined uniquely by the dimensions of the subspaces U_{ij}, which are obviously independent of the specific choice of basis. Thus the Jordan form of a matrix is unique up to possible reordering of the Jordan blocks.

Two $n \times n$ matrices A, B are similar to each other, i.e. $B = T^{-1}AT$ for some nonsingular T, if and only if their Jordan forms coincide (up to recordering of the Jordan blocks). For if $Q^{-1}BQ = J$ is in Jordan form, then $J = Q^{-1}T^{-1}ATQ = (TQ)^{-1}A(TQ)$ is also the Jordan form for A. And if $B = QJQ^{-1}$ and $A = SJS^{-1}$, then $B = Q(S^{-1}AS)Q^{-1} = (SQ^{-1})^{-1}A(SQ^{-1})$.

Finally, we return to the characteristic polynomial of A. The dimension of the matrix in (3.23) is $\dim U_{i\mu_i}$, so by inspection

$$\text{(3.25)} \qquad \det(A-\lambda) = \prod_{i=1}^{k} (\lambda_i - \lambda)^{\dim U_{i\mu_i}},$$

i.e. m_i in (3.4) is equal to $\dim U_{i\mu_i} = \sum_{j=1}^{\dim U_{i1}} \nu(j)$. As each $Y_j^i \subset U_{i\mu_i}, \dim Y_j^i = \nu(j)$, obtained from (3.19), and

$$\mu_i = \text{ maximum } \nu(j), j = 1, \ldots, \dim U_{i1}, \tag{3.26}$$

it follows from (3.26) that

$$\mu_i \leq m_i, i = 1, \ldots, k. \tag{3.27}$$

Thus from (3.4) and (3.7), $M_A(\lambda)$ divides $P_A(\lambda)$ and we have proved

Theorem 3.5 (Cayley-Hamilton theorem). *Every $n \times n$ matrix satisfies its own characteristic equation, $P_A(A) = 0$.*

In the special case that the matrix A is real, its eigenvalues are either real or occur in complex conjugate pairs. if all of the eigenvalues happen to be real, then so is the Jordan form of A, and the similarity transformation matrix T, such that $T^{-1}AT$ is in Jordan form can also be taken real. More generally, at least some of the eigenvalues are complex, and we wish to transform A to some sort of canonical form by a real similarity transformation.

The idea of this construction is very simple; instead of using the $e_{j\nu}^i$ as basis vectors, we use their real and imaginary parts. More specifically, we can place the eigenvalues in any convenient order, so we consider any complex pair in the form

$$Re\lambda_i = Re(\lambda_{i+1}) = \xi, Im(\lambda_i) = -Im(\lambda_{i+1}) = \eta \tag{3.28}$$

where ξ, η are real. The characteristic and minimal polynomials of A, factored into real factors, contain factors of

$$((\lambda - \xi)^2 + \eta^2)^{m_i}, ((\lambda - \xi)^2 + \eta^2)^{\mu_i} \tag{3.29}$$

respectively, and we have, for such values of i,

$$m_i = m_{i+1}, \mu_i = \mu_{i+1}, \dim Y_j^i = \dim Y_j^{i+1}, j = 1, \ldots, \mu_i \tag{3.30}$$

$$e_{j\nu}^{i+1} = \overline{e_{j\nu}^i}, \nu = 1, \ldots, \nu(j), j = 1, \ldots, \mu_i. \tag{3.31}$$

Instead of working with U_{ij} and $U_{i+1,j}$ separately, we consider a basis for their direct sum

$$X_j = U_{ij} \oplus U_{i+1,j} = \{x \in H | [(A - \xi)^2 + \eta^2]^j x = 0\} \tag{3.32}$$

We rewrite (3.20), taking $Re(\lambda_i) = \xi$, $Im(\lambda_i) = \eta$, $Re(e^i_{j\nu}) = p^i_{j\nu}$, $Im(e^i_{j\nu}) = q^i_{j\nu}$, with $\xi, \eta, p^i_{j\nu}, q^i_{j\nu}$ all real. Taking real and imaginary parts we obtain

$$\begin{aligned} Ap^i_{j1} &= \xi p^i_{j1} - \eta q^i_{j1}, \\ Aq^i_{j1} &= \eta p^i_{j1} + \xi q^i_{j1}; \\ Ap^i_{j\nu} &= \xi p^i_{j\nu} - \eta q^i_{j\nu} + p^i_{j(\nu-1)}, \\ Aq^i_{j\nu} &= \eta p^i_{j\nu} + \xi q^i_{j\nu} + q^i_{j(\nu-1)}; \text{ for } \nu > 1 \end{aligned} \tag{3.33}$$

instead of (3.20). Taking $\{p^i_{j\nu}, q^i_{j\nu}\}_{\nu=1}^{\dim Y^i_j}$ as a basis for $W_j = \text{span}\ \{p^i_{j\nu}, q^i_{j\nu}\}_{\nu=1}^{\dim Y^i_j}$, from (3.33) we get a real representation for A of the form

$$A|_{W_j} = \begin{pmatrix} \xi & \eta & 1 & 0 & & & & 0 & \\ -\eta & \xi & 0 & 1 & & & & & \\ & & \xi & \eta & 1 & 0 & & & \\ & & -\eta & \xi & 0 & 1 & & & \\ & 0 & & & & & \ddots & & \\ & & & & & & & \xi & \eta \\ & & & & & & & -\eta & \xi \end{pmatrix} \tag{3.34}$$

analogously with (3.21). This is the form of "Jordan block" obtainable by a real similarity transformation. As a special case, if $\xi \pm i\eta$ are simple eigenvalues of A, i.e. $\mu_i = m_i = 1$, we obtain the "diagonal block"

$$\begin{pmatrix} \xi & \eta \\ -\eta & \xi \end{pmatrix}$$

in the real canonical form of A.

Problems

1. Find the Jordan canonical form of each of the following

$$\begin{pmatrix} 1 & -1 & -1 \\ 2 & 3 & 1 \\ 0 & 0 & 0 \end{pmatrix} \quad \begin{pmatrix} 1 & 2 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 2 \end{pmatrix} \quad \begin{pmatrix} -1 & 0 & 2 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

2. Prove that there is no $n \times n$ matrix A, such that $T^{-1}AT$ is diagonal for some nonsingular T, whereas $Q^{-1}AQ$ is in Jordan canonical form, but not diagonal, for some nonsingular Q.

3. Let A be an $n \times n$ matrix, all of whose eigenvalues are equal, but such that A is not a scalar multiple of the $n \times n$ identity matrix. Prove that A is not similar to any diagonal matrix.

4. Prove that any $n \times n$ matrix A can be uniformly approximately by matrices similar to diagonal matrices, i.e. for any $\varepsilon > 0$, there exists A_ε, similar to a diagonal matrix, such that $|A_{ij} - A_{\varepsilon ij}| < \varepsilon$ for any $1 \leq i, j \leq n$. Outline: $A = T^{-1}JT$ for some J in Jordan canonical form, but presumably not diagonal. But J can be perturbed (show how) to $\tilde{J}$, such that $\tilde{J} - J$ is arbitrarily small, but the eigenvalues of $\tilde{J}$ are distinct. Thus $\tilde{J}$ is similar to a diagonal matrix. Take $A_\varepsilon = T^{-1}\tilde{J}T$; it remains to show that if $\tilde{J}$ is similar to a diagonal matrix, so is A_ε, and that $A_\varepsilon - A$ is small as described above.

   Explain why it is <u>not</u> possible to pass to the limit as $\varepsilon \to 0$ and thus show that A must be similar to a diagonal matrix.

5. True or false, for A, B $n \times n$ matrices:

   a) If A, B are both similar to diagonal matrices, then so are AB and BA.

   b) If E_{jka} is an $n \times n$ elementary matrix, as described in chapter 2, and if A is similar to a diagonal matrix, then so are AE and EA.

   c) If $AB = BA$ and A is similar to a diagonal matrix, then so are B and AB.

   d) If A is a real matrix, and is similar to a diagonal matrix D, i.e. $D = T^{-1}AT$, then T can always be taken to be a real matrix.

   e) If A is similar to a diagonal matrix and B is itself diagonal, then $A + B$ is similar to a diagonal matrix.

   f) If A is a 3×3 matrix and $M_A(\lambda) = \lambda^2$, then $r(A) = 1$, where $r(A)$ is the rank of A as in chapter 2.

   g) If $\dim U_{i1} = n$ for an $n \times n$ matrix A, with U_{ij} as given in (3.8), then $A = \lambda_i I$.

   h) If two $n \times n$ matrixes A, B, have the same characteristic and minimal polynomials,
   i.e. $P_A(\lambda) = P_B(\lambda)$ and $M_A(\lambda) = M_B(\lambda)$, then they are similar to each other ($B = T^{-1}AT$ for some nonsingular T.)

6. Show that a real 2×2 matrix can always be transformed into one of the forms
$$\begin{pmatrix} \xi & 0 \\ 0 & \eta \end{pmatrix}, \quad \begin{pmatrix} \xi & 1 \\ 0 & \xi \end{pmatrix}, \quad \text{or} \quad \begin{pmatrix} \xi & \eta \\ -\eta & \xi \end{pmatrix}$$
by a real similarity transformation, with ξ, η real. What is the corresponding statement for real 3×3 matrices?

7. Let λ_i be an eigenvalue of an $n \times n$ matrix A. Show that $\dim \ker(A - \lambda_i)^j = \dim \ker(A^\dagger - \bar{\lambda}_i)^j$ for any positive integer j, and thus that the dimensions of the invariant subspaces U_{ij} of A are the same as those of the corresponding subspaces for $A^\dagger$.

8. Use the result of problem 4 to get an independent proof of the Cayley-Hamilton theorem.

9. Suppose that $J = T^{-1}AT$ is in Jordan canonical form, but not diagonal. Describe the structure of the matrices $T^{-1}A^jT$ for j a positive integer.

10. Let A be a 4×4 matrix such that $M_A(\lambda) = (\lambda - 2)^2(\lambda + 1)$. What are the possible structures of the Jordan canonical form of A?

11. Let J be the Jordan form of an $n \times n$ matrix A. Show that $J = D + G$, where D is diagonal and G satisfies $G^\mu = 0$, where $\mu \underset{i}{=\rightarrow}$ maximum μ_i, μ_i the exponents appearing in the minimal polynomial of A, given by (3.7).

12. Find the characteristic polynomial of each of the following matrices of dimension n, for $a, a_1, \ldots, a_n \in \mathbb{C}$ given. In the second example, an induction argument on the dimension n will work.

$$\begin{pmatrix} 0 & 1 & & 0 & \\ & 0 & & 0 & 1 \\ a_1 & a_2 & \ldots & a_{n-1} & a_n \end{pmatrix} \quad \begin{pmatrix} 0 & a & & & \\ a & 0 & a & 0 & \\ & a & 0 & & \\ & 0 & & 0 & a \\ & & & a & 0 \end{pmatrix}$$

4 Some special classes of square matrices

Let f be a polynomial function of a single variable and A an $n \times n$ matrix. If $f(A) = 0$, meaning that the $n \times n$ matrix $f(A)$ vanishes identically, then we easily verify

$$f(T^{-1}AT) = T^{-1}f(A)T = 0 \quad \text{for any nonsingular } T \tag{4.1}$$

$$\bar{f}(A^{\dagger}) = (f(A))^{\dagger} = 0, \tag{4.2}$$

where $\bar{f}$ is the complex conjugate of f, i.e. the polynomial whose coefficients are the complex conjugates of those of f.

Lemma 4.1. *A square matrix A is similar to a diagonal matrix if and only if its minimal polynomial contains no repeated factors, i.e. if all of the μ_i are equal to 1 in (3.7).*

Proof. Choose T such that $J = T^{-1}AT$ is in Jordan form. Then from (4.1), J satisfies $M_A(J) = 0$. If J is diagonal, then no repeated factors are needed in (3.7) to satisfy $M_A(J) = 0$. If J is not diagonal, i.e. a Jordan block for some λ_i is not diagonal, then the corresponding block of $J - \lambda_i$ also does not vanish, and thus if $M_A(\lambda)$ contains no additional factors of $(\lambda - \lambda_i)$, neither does $M_A(J)$.

Corollary 1. *If A is similar to a diagonal matrix, then so is $A^{\dagger}$.*

Proof. Since $M_A(A) = 0, \bar{M}_A(A^{\dagger}) = 0$ from (4.2), it is easily verified that $\bar{M}_A$ is the minimal polynomial for $A^{\dagger}$. Thus if M_A has no repeated factors, neither does $\bar{M}_A$.

Remark. The adjoint property is not, in general, invariant under similarity transformations, i.e. in general

$$(T^{-1}AT)^{\dagger} \neq T^{-1}A^{\dagger}T \tag{4.3}$$

Thus in general, if $T^{-1}AT$ is diagonal, $T^{-1}A^{\dagger}T$ is not.

Corollary 2. *Let f be a polynomial with no repeated factors. If $f(A) = 0$, then A is similar to a diagonal matrix.*

Proof. Either f is itself the minimal polynomial for A, or else $M_A(\lambda)$ is obtained from $f(\lambda)$ by dividing $f(\lambda)$ by those factors $(\lambda - \alpha)$ such that α_i is not an eigenvalue of A, i.e. such that $A - \alpha_i$ is nonsingular. In either case, the minimal polynomial of A contains no respected factors.

An application is the class of idempotent matrices, $n \times n$ matrices A satisfying

$$Q^2 = Q \tag{4.4}$$

equivalently $f(Q) = Q(Q-1) = 0, f$ a polynomial without repeated factors. Here obviously either $Q = 0, Q = I$, or else (4.4) is actually the minimal polynomial for Q.

Clearly idempotent matrices are similar to diagonal matrices, and have only $0, 1$ as possible eigenvalues.

The class of idempotent matrices is invariant under similarity transformations, i.e. if Q is idempotent,

$$\begin{aligned}(T^{-1}QT)(T^{-1}QT) &= T^{-1}Q^2T \\ &= T^{-1}QT\end{aligned}$$

so $T^{-1}QT$ is also idempotent. Consequently, any matrix A which has only $0, 1$ as eigenvalues, and which is similar to a diagonal matrix, is idempotent.

An especially important class of idempotent matrices are the projection operators $Q : H \to K, K \subset H$ a closed subspace satisfying

$$(h - Qh, y) = 0 \quad \text{for all } h \in H,\ y \in K. \tag{4.5}$$

If we choose $h = Qx \in K$ in (4.5), then $y = Qx - Q^2x \in K$ is a legal choice also, and it follows immediately that Q satisfies (4.4). In addition, for any $x, y \in H, Qx$ and Qy are in K, so by repeated use of (4.5), we have $(Qx, y) = (Qx, Qy) = (x, Qy) = (x, Q^\dagger y)$ so $Q = Q^\dagger$. Such matrices are called Hermitian, and will be discussed further below.

Two $n \times n$ matrices A, B commute if $AB = BA$. This is obviously a very special property, but it occurs in a number of interesting cases.

For example, if B commutes with a diagonal matrix D, then comparing $(BD)_{ij}$ with $(DB)_{ij}, 1 \leq i, j \leq n$, we find $B_{ij}(D_{ii} - D_{jj}) = 0$ for all $1 \leq i, j \leq n$, generally a very strong restriction on the off-diagonal elements of B. For example, if the diagonal elements of D are distinct, that is $D_{ii} \neq D_{jj}$ for all $i \neq j$, and if $BD = DB, B$ is also diagonal. Furthermore, commutivity survives similarity transformations; if A, B commute, so do $T^{-1}AT$ and $T^{-1}BT$.

Theorem 4.2. *Let A, B be commuting $n \times n$ matrices. Then:*

(1) *If $U_{ij} = \ker(A - \lambda_i)^j$, then $B : U_{ij} \to U_{ij}$.*

(2) *If the eigenvalues of A are distinct, then A, B can be made diagonal simultaneously by a similarity transformation, i.e. these exists a nonsingular T such that $T^{-1}AT$ and $T^{-1}BT$ are both diagonal. Indeed, $T^{-1}BT$ is diagonal for any T such that $T^{-1}AT$ is diagonal.*

(3) If A is similar to a diagonal matrix, there exists T such that $T^{-1}AT$ is diagonal and $T^{-1}BT$ is in Jordan canonical form.

Proof. For $x \in U_{ij}, Bx$ satisfies

$$(A - \lambda_i)^j Bx = B(A - \lambda_i)^j x = 0, \tag{4.6}$$

so $Bx \in U_{ij}$. Now if the eigenvalues of A are distinct, then for λ_i any eigenvalue of A, U_{ij} is of dimension 1, so $x \in U_{ij}$ implies $Bx \in U_{ij}$ implies $Bx = \alpha x$ for some $\alpha \in F$. Thus the eigenvectors of A are also the eigenvectors of B, and there are n linearly independent eigenvectors, since the eigenvalues of A are distinct. Choosing the eigenvectors of A as a basis for H puts A, B simultaneously into diagonal form.

If A is similar to a diagonal matrix, then for any eigenvalue $\lambda_i, U_{ij} = U_{i1}$ for $j \geq 1$, but $\dim U_{i1}$ can be larger than 1. Now $B : U_{i1} \to U_{i1}$, but A, restricted to the subspace U_{i1}, is just λ_i times the identity matrix, so we learn nothing further from the commutivity property. We can choose a basis for U_{i1} which puts B, restricted to U_{i1}, in Jordan canonical form. Such a change of basis will obviously not change the from of A, on this subspace. Proceeding in this manner for each U_{i1} such that $\dim U_{i1} > 1$, we retain A in diagonal form, but can simultaneously put B in Jordan canonical form. Of course, if it is also known that B is similar to a diagonal matrix, then the Jordan form of B is necessarily diagonal, so A and B can be made diagonal simultaneously.

An $n \times n$ matrix A is said to be <u>normal</u> if it commutes with its adjoint

$$AA^\dagger = A^\dagger A. \tag{4.7}$$

Immediately from (4.7) we have $\|Ax\| = \|A^\dagger x\|$ for any vector x, as

$$\begin{aligned} \|Ax\|^2 &= (Ax, Ax) \\ (A^\dagger Ax, x) &= (AA^\dagger x, x) \\ &= (A^\dagger x, A^\dagger x) \\ &= \|A^\dagger x\|^2. \end{aligned}$$

Theorem 4.3. *A normal matrix A is similar to a diagonal matrix. Furthermore $ker(A) = \ker(A^\dagger)$, and $R(A) \perp \ker(A)$.*

Proof. Suppose not, then from Lemma 4.1 it follows that $\mu_i \geq 2$ in (3.7) for one (or more) of the eigenvalues λ_i, so U_{i1} is a proper subspace of U_{i2}. From Lemma 1.5, there exists a nonzero $x \in U_{i2}, x \perp U_{i1}$, i.e. x satisfies

$$Ax = \lambda_i x + y \quad \text{for some } y \neq 0 \quad \text{in } U_{i1} \tag{4.8}$$

$$(x, w) = 0 \quad \text{for all } w \in U_{i1}. \tag{4.9}$$

Now from (4.7) and the first statement of Theorem 4.2,

$$A^\dagger : U_{i1} \to U_{i1}. \tag{4.10}$$

Thus, from (4.8), using (4.9)

$$\begin{aligned} (Ax, y) &= \lambda_i (x, y) + \|y\|^2 \\ &= \|y\|^2 \\ &> 0, \end{aligned} \tag{4.11}$$

but

$$\begin{aligned} (Ax, y) &= (x, A^\dagger y) \\ &= 0 \end{aligned}$$

using (4.10) and (4.9), since $y \in U_{i1}$. This contradiction proves the first part of the theorem.

Suppose $x \in ker(A^\dagger)$. Then

$$\begin{aligned} 0 &= (A^\dagger x, A^\dagger x) \\ &= (AA^\dagger xx) = (A^\dagger Axx) = (Ax, Ax) \end{aligned}$$

so $x \in \ker(A)$, and vice versa. From (2.7) we now have $R(A) \perp \ker(A)$.

Corollary. *An $n \times n$ matrix A is similar to a diagonal matrix if and only if it is similar to a normal matrix.*

Proof. Since diagonal matrices are normal, this part is trivial. Thus assume $T^{-1}AT = B$, where B is normal. By Theorem 4.3, there exists nonsingular Q such that $Q^{-1}BQ = D, D$ diagonal. Thus

$$\begin{aligned} D &= Q^{-1}(T^{-1}AT)Q \\ &+ (Q^{-1}T^{-1})A(TQ) \\ &= (TQ)^{-1}A(TQ), \end{aligned}$$

so A is similar to a diagonal matrix.

Two important classes of normal matrices are unitary matrices, which are nonsingular and satisfy

$$A^{-1} = A^{\dagger}, \tag{4.12}$$

and Hermitian matrices, satisfying

$$A = A^{\dagger}. \tag{4.13}$$

For $F = \mathbb{R}$, a nonsingular matrix such that $A^t = A^{-1}$ is called orthogonal, and a matrix satisfying $A = A^t$ is called symmetric. The term self-adjoint is also sometimes used for Hermitian and/or symmetric matrices, but this term means something stronger in the context of unbounded operators an infinite dimensional Hilbert spaces. The terms symmetric and Hermitian are sometimes used interchangeably when no ambiguity is possible.

From (4.12) it follows immediately that a product of unitary matrices is unitary.

Theorem 4.4. *Let $r_i, i = 1, \ldots, n,$ and $c_i, i = 1, \ldots, n$ denote the rows and columns, respectively, of an $n \times n$ matrix A. The matrix A is unitary if and only if*

$$(r_i, r_j) = (c_i, c_j) = \delta_{ij}, \quad 1 \le i, j \le n. \tag{4.14}$$

Proof. A is unitary if and only if $AA^{\dagger} = A^{\dagger}A = I$. By inspection

$$(AA^{\dagger})_{ij} = (r_i, r_j) \quad \text{and } (A^{\dagger}A)_{ij} = \overline{(c_i, c_j)}, 1 \le i, j \le n. \tag{4.15}$$

Theorem 4.5. *Inner products of pairs of vectors and lengths of vectors are invariant under unitary transformations. Furthermore, the eigenvalues of a unitary matrix are of magnitude unity.*

Proof. Let A be unitary, and x, y any vectors in H. Then

$$\begin{aligned}(Ax, Ay) &= (A^\dagger(Ax), y)\\ &= (A^\dagger Ax, y)\\ &= (x, y) \qquad (4.16)\end{aligned}$$

using (4.12). Setting $x = y$, we have $\|Ax\| = \|x\|$ from (4.16). Conversely, if (4.16) holds for all $x, y \in H$ then A satisfies (4.12). And for x any eigenvector of A, corresponding to eigenvalue λ,then $\|x\| = \|Ax\| = \|\lambda x\| = |\lambda|\|x\|$ so $|\lambda| = 1$.

Hermitian or symmetric matrixes satisfy several special conditions.

Theorem 4.6. *A linear operator A on a complex Hilbert space H is Hermitian if and only if*

$$(Az, z) \in \mathbb{R} \quad \textit{for all } z \in H. \qquad (4.17)$$

Proof. If H is Hermitian, then for any $z \in H$

$$\begin{aligned}(z, Az) &= (A^\dagger z, z)\\ &= (Az, z)\\ &= \overline{(z, Az)} \qquad (4.18)\end{aligned}$$

so $(z, Az) \in \mathbb{R}$. Next assume (4.17) holds. For arbitrary $p, q \in H$, set $z = p + tq$ in (4.17), $t \in \mathbb{C}$ to be chosen below. Then

$$(A(p + tq), p + tq) = (Ap, p) + |t|^2(Aq, q) + t(Aq, p) + \bar{t}(Ap, q), \qquad (4.19)$$

so by hypothesis, the expression

$$t(Aq, p) + \bar{t}(Ap, q) = t(Aq, p) + \bar{t}\overline{(q, Ap)} \qquad (4.20)$$

has to be real for any choice of t. By direct computation

$$\begin{aligned}&Im(t(Aq, p) + \bar{t}\overline{(q, Ap)})\\ &\quad Im(t)Re((Aq, p) - (q, Ap))\\ &\quad + Re(t)Im((Aq, p) - (q, Ap)) \qquad (4.21)\end{aligned}$$

which is zero for any t only if $(Aq, p) = (q, Ap)$. And if $(Aq, p) = (q, Ap)$ for any $p, q \in H$, then A is Hermitian.

Remark. This theorem is true for infinite-dimensional Hilbert spaces as well, with the same proof.

Corollary. *The eigenvalues of a Hermitian matrix are real.*

Theorem 4.7. *The eigenvectors of a Hermitian matrix may be made an orthonormal basis for H.*

Proof. Normalize all the eigenvectors x_i so that $\|x_i\| = 1$. Next let x_i, x_j be eigenvectors corresponding to different eigenvalues λ_i, λ_j. Then

$$\begin{aligned}(Ax_i, x_j) &= \lambda_i(x_i, x_j) \\ &= (x_i, Ax_j) \\ &= \lambda_j(x_i, x_j)\end{aligned} \tag{4.22}$$

since λ_j is real, so $(\lambda_i - \lambda_j)(x_i, x_j) = 0$ and $(x_i, x_j) = 0$. If there are eigenvalues λ_i of multiplicity $m_i > 1$, choose any orthonormal basis for U_{i1}, for example by application of the Gram-Schmidt procedure. All of the eigenvectors are now orthonormal, and as A is similar to a diagonal matrix, they form a basis for H.

Corollary. *The matrix T such that $T^{-1}AT$ is diagonal can be taken to be unitary.*

Proof. Given the eigenvectors for A, the matrix T is just the set of eigenvectors of A, assembled columnwise. Choose the eigenvectors of A to be orthonormal, and T is unitary by application of Theorem 4.4.

Denote by

$$\|A\| = \sup_{\substack{x \in H \\ x \neq 0}} \frac{\|Ax\|}{\|x\|} \tag{4.23}$$

where $\|\cdot\|$ is the ℓ_2 norm given in (1.5), and by

$$\sigma(A) = \text{maximum } |\lambda_i| \quad \{\lambda_i\} \quad \text{the eigenvalues of } A. \tag{4.24}$$

$\sigma(A)$ is called the spectral radius of A; as x in (4.23) can be chosen any eigenvector of A, it follows that

$$\sigma(A) \leq \|A\|. \tag{4.25}$$

Theorem 4.8. *If A is Hermitian, then*

$$\|A\| = \sigma(A). \tag{4.26}$$

Proof. Let $\{x_i\}_{i=1}^n$ be the orthonormal eigenvectors of A, corresponding to eigenvalues λ_i. Then any $x \in H$ is of the form

$$x = \sum_{i=1}^n \alpha_i x_i, \quad \alpha_i \in F; \tag{4.27}$$

$$Ax = \sum_{i=1}^n \lambda_i \alpha_i x_i \tag{4.28}$$

$$\begin{aligned} \|x\|^2 &= (x, x) \\ &= \sum_{i,j} \alpha_i \bar{\alpha}_j (x_i, x_j) \\ &= \sum_i |\alpha_i|^2 \end{aligned} \tag{4.29}$$

since the $\{x_i\}$ are orthonormal, and

$$\begin{aligned} \|Ax\|^2 &= (Ax, Ax) \\ &= \sum_{i,j} \lambda_i \bar{\lambda}_j \alpha_i \bar{\alpha}_j (x_i, x_j) \\ &= \sum_i |\lambda_i|^2 |\alpha_i|^2 \end{aligned} \tag{4.30}$$

From (4.29) and (4.30) it now follows that $\|A\| \le \sigma(A)$, which together with (4.25) gives (4.26).

A matrix satisfying $A^\dagger = -A$ is called <u>anti-Hermitian</u> or <u>skew-Hermitian</u>; a real matrix such that $A^t = -A$ is also called <u>antisymmetric</u> or <u>skew-symmetric</u>.

Theorem 4.9. *A linear operator A on a complex Hilbert space H is anti-Hermitian if and only if (z, Az) is pure imaginary for all $z \in H$. The eigenvalues of an anti-Hermitian matrix are pure imaginary.*

Theorem 4.10. *The eigenvectors of an anti-Hermitian matrix may be made an orthonormal basis for H.*

Theorem 4.11. *If A is anti-Hermitian, $\sigma(A) = \|A\|$.*

If A is anti-Hermitian, iA is Hermitian, and these results follow from the corresponding results for Hermitian matrices.

Any square matrix can be (trivially) written as the sum of a Hermitian and an anti-Hermitian matrix,

$$A = \frac{1}{2}(A + A^\dagger) + \frac{1}{2}(A - A^\dagger) \tag{4.31}$$

A Hermitian matrix is positive definite (respectively positive semidefinite, negative definite, negative semidefinite) if

$$(Ax, x) > 0 \text{ for all } x \in H,\ x \neq 0. \tag{4.32}$$

(respectively $\geq 0, < 0, \leq 0$ in (4.32)). Sometimes these terms are also applied to
nonsymmetric matrices A on a real Hilbert space, in which case they refer to the symmetric part of A, as obtained from (4.31). A real matrix A is thus positive definite, for example, if and only if $\frac{1}{2}(A + A^\dagger)$ is.

Lemma 4.12. *A Hermitian matrix A is positive definite if and only if all of its eigenvalues are positive.*

Proof. Obviously positive eigenvalues are necessary to satisfy (4.32). Let $\{x_i\}_{i=1}^n$ be orthonormal eigenvectors of A, corresponding to eigenvalues λ_i. Then for any $x \in H$, from Theorem 4.7,

$$\begin{aligned} x &= \sum_{i=1}^n \alpha_i x_i \\ Ax &= \sum_{i=1}^n \lambda_i \alpha_i x_i \\ (Ax, x) &= \sum_{i,j} \lambda_i \alpha_i \bar{\alpha}_j (x_i, x_j) \\ &= \sum_i \lambda_i |\alpha_i|^2; \qquad (4.33) \end{aligned}$$

obviously if all of the $\lambda_i > 0$, then (4.32) will be satisfied.

The following theorem provides a relatively easy method of checking whether a given Hermitian matrix is indeed positive definite. Given an $n \times n$ Hermitian matrix $A : H^n \to H^n$, for $m = 1, 2, \ldots, n$, denote by

$$A^{(m)} = \{A_{ij}\}_{i,j=1}^{m} \tag{4.34}$$

i.e. the $m \times m$ matrix obtained by truncating the matrix A to its first m rows and columns. For example, $A^{(1)} = (A_{11})$ and $A^{(n)} = A$. (These truncated matrices $A^{(m)}$ have various names; the literature is not consistent on this point.)

We think of $A^{(m)}$ as mapping H^m to H^m, where H^m, of dimension m, obviously, is isomorphic to a subspace of H^n, the elements of which have zeros in their last $n - m$ components.

Theorem 4.13 (Sylvester's Theorem). *: A Hermitian matrix A is positive definite if and only if*

$$\det A^{(m)} > 0, m = 1, \ldots, n \tag{4.35}$$

Proof. First suppose that (4.35) fails, i.e. that $\det A^{(k)} \leq 0$ for some k. From (3.5), det $A^{(k)}$ is the product of the eigenvalues of $A^{(k)}$, so one or more of the eigenvalues of $A^{(k)}$ is nonpositive. As $A^{(k)}$ is also Hermitian, from lemma 4.12 it follows that $A^{(k)}$ is not positive definite, i.e. there exists nonzero $x^{(k)} \in H^k$ such that $(A^{(k)}x^{(k)}, x^{(k)}) \leq 0$. Obtain $x^{(n)} \in H^n$ from $x^{(k)}$ by placing zeros in the last $n - k$ components; then $(Ax^{(n)}, x^{(n)}) = (A^{(k)}x^{(k)}, x^{(k)}) \leq 0$, so (4.32) fails and A is not positive definite.

Therefore we assume now that (4.35) holds. In this case, we proceed by induction. Obviously the theorem holds for matrices of dimension 1; therefore we assume it is true for matrices of dimension up to and including $n - 1$, and prove it for matrices of dimension n. In particular, we may assume that $A^{(n-1)}$ is positive definite.

Now as $\det A > 0$, by another application of (3.5) and lemma 4.12, we may infer that if A is not positive definite, there are (at least) two negative real eigenvalues, denoted by λ_1, λ_2. Let x_1, x_2 be the corresponding eigenvectors, taken orthonormal, using Theorem 4.7.

Given x_1, x_2, there is a nontrivial linear combination

$$x = \alpha_1 x_1 + \alpha_2 x_2,\ \alpha_1, \alpha_2 \in F,\ |\alpha_1|^2 + |\alpha_2|^2 > 0 \tag{4.36}$$

such that the last (n-th) component of x is zero.

Now

$$\begin{aligned}(Ax,x) &= (\alpha_1\lambda_1 x_1 + \alpha_2\lambda_2 x_2, \alpha_1 x_1 + \alpha_2 x_2)\\ &= \lambda_1|\alpha_1|^2 + \lambda_2|\alpha_2|^2\\ &< 0,\end{aligned} \tag{4.37}$$

since both $\lambda_1, \lambda_2 < 0$. But let $x^{(n-1)} \in H^{n-1}$ be the vector x with that last component, equal to zero, dropped. Then

$$\begin{aligned}(Ax,x) &= (A^{(n-1)}x^{(n-1)}, x^{(n-1)})\\ &> 0\end{aligned} \tag{4.38}$$

since $A^{(n-1)}$ is known to be positive definite. This contradiction with (4.37) completes the proof.

As an application of several of the above results, we have the following.

Theorem 4.14. *A normal matrix whose eigenvalues are all real (respectively, all pure imaginary or zero) is Hermitian (respectively anti-Hermitian). For N normal, there is a unitary T such that $T^{-1}NT$ is diagonal.*

The proof of this uses the following

Lemma 4.15. *If T is unitary and A is Hermitian (or anti-Hermitian), then so is $T^{-1}AT$.*

Proof of lemma:

$$\begin{aligned}(T^{-1}AT)^\dagger &= (T^\dagger AT)^\dagger\\ &= T^\dagger A^\dagger T\\ &= T^{-1}A^\dagger T\\ &= \pm T^{-1}AT.\end{aligned} \tag{4.39}$$

for A Hermitian or anti-Hermitian, respectively.

Proof of theorem: Given a normal matrix N, set $S = \frac{1}{2}(N + N^\dagger), A = \frac{1}{2}(N - N^\dagger)$, which are Hermitian and anti-Hermitian, respectively, and satisfy

$$N = S + A. \tag{4.40}$$

Since N is normal, $NN^\dagger = N^\dagger N$, and therefore

$$SA = AS. \tag{4.41}$$

Let T_0 be a unitary matrix such that

$$D_0 = T_0^{-1} S T_0$$

is diagonal. Using lemma 4.14,

$$A_0 = T_0^{-1} A T_0 \tag{4.42}$$

is anti-Hermitian, and from (4.41)

$$D_0 A_0 = A_0 D_0. \tag{4.43}$$

From (4.43), the elements of D_0, A_0 satisfy

$$(D_{0jj} - D_{0kk}) A_{0jk} = 0, \quad 1 \leq j, k \leq n \tag{4.44}$$

so A_0 is also block diagonal, the blocks $\{b_i\}$ determined by the requirement that the diagonal elements of D_0 are equal within each block.

If the eigenvalues of D_0 happen to be distinct, then A_0 is already diagonal. Otherwise, we find unitary matrices $\{T_i\}$, one for each diagonal block b_i of A_0, such that T_i is the identity matrix outside of the block b_i,and

$$A_i = T_i^{-1} A_{i-1} T_i \tag{4.45}$$

is diagonal within the block b_i. (The matrix T_i is unitary because A_{i-1} is anti-Hermitian within the block b_i). As D_0, restricted to b_i, is a multiple of the identity matrix, it is unchanged,

$$D_0 = T_i^{-1} D_0 T_i. \tag{4.46}$$

By induction, each A_{i-1} is anti-Hermitian and thus so is A_i, from (4.45) and the unitary property of T_i.

We conclude this process with two diagonal matrices D_0 and A_m, and T obtained from

$$T = T_0 T_1 \ldots T_m, \tag{4.47}$$

such that

$$T^{-1} S T = D_0 \tag{4.48}$$

$$T^{-1}AT = A_m \tag{4.49}$$

As the product of unitary matrices, T is unitary, so D_0 is Hermitian and A_m anti-Hermitian. Therefore $\{D_{0jj}\}$ are real and $\{A_{mjj}\}$ are pure imaginary, and since the eigenvalues of $D_0 + A_m$ are the same as those of $S + A$, using (4.48), (4.49), from (4.40) we see that the D_{0jj}, A_{mjj} are just the real and imaginary parts of the eigenvalues of N.

Suppose that the eigenvalues of N are real (the pure imaginary case is proved exactly the same way). Then $A_{mjj} = 0$ for all $j = 1, \ldots, n$, and since A_m is diagonal, $A_m = 0$ and $A = 0$ from (4.49). Then N is Hermitian from (4.40).

Since T is unitary in (4.47), we have also proved the following:

Corollary. *The eigenvectors of a normal matrix N can be chosen to be orthonormal, satisfying $(x_i, x_j) = \delta_{ij}, 1 \leq i, j \leq n$; with the eigenvectors so chosen, N admits a representation of the form*

$$N = \sum_{i=1}^{n} \lambda_i x_i x_i^\dagger \tag{4.50}$$

where λ_i is the corresponding eigenvalue.

In (4.50), the eigenvalues λ_i are in general complex, and each product $x_i x_i^\dagger$ is a matrix of rank 1.

These results may be summarized as follows, giving a comparison between normal, Hermitian, and unitary matrices.

Theorem 4.16. *A normal matrix A admits a representation of the form*

$$A = T^{-1}DT, \tag{4.51}$$

with D diagonal and T unitary. A is Hermitian if and only if D is real. A is anti-Hermitian if and only if D is pure imaginary. A is unitary if and only if the diagonal elements of D are of unit magnitude, i.e. $|D_{ii}| = 1$ for $i = 1, \ldots, n$. A is idempotent and Hermitian, i.e. a projection operator onto its range if and only if each of the diagonal elements of D is either 1 or 0. Finally, it follows from (4.51) that $\sigma(A) = \|A\|$ (as determined by (4.23), (4.24)) for any normal matrix A, and not just for Hermitian matrices.

Problems

1. Consider the mixed eigenvalue problem, $Bx = \lambda Ax$, where B, A are $n \times n$ Hermitian matrices, A is nonsingular, $\lambda \in F$ is the eigenvalue and x the eigenvector. Show by example that in general the λ are not real. Prove that if A is positive definite, then λ is real.

2. What can be said about the eigenvalues, eigenvectors, and minimal polynomial of a matrix Q satisfying $Q^2 = Q^\dagger$?

3. Find an example of a real matrix Q, <u>not symmetric</u> which satisfies $Q = Q^2$.

4. Show that a real, antisymmetric matrix of odd dimension is always singular.

5. True or false:

   a) If A is Hermitian and positive definite, there exists a Hermitian, positive definite "square root", i.e. S such that $S^2 = A$.

   b) There is no square matrix A, of any (finite) dimension, which satisfies $A = I + AA^\dagger, I$ the identity matrix.

   c) There is no Hermitian matrix, of any dimension, whose minimal polynomial
   $M_A(A) = (A^2 - 1)(A^3 - A^2 - 4A + 4)$.

   d) If A is any $n \times n$ matrix and m any positive integer, then $\|A^m\| \leq \|A\|^m$, with $\|A\|$ defined in (4.23).

   e) $\sigma(A + B) \leq \sigma(A) + \sigma(B)$ for any $n \times n$ A, B, with σ defined (4.24).

   f) $\sigma(A^2) = (\sigma(A))^2$

6. For what real values of t is the matrix

$$\begin{pmatrix} 1+t & t & t \\ t & t^2 & sin\ t \\ t & -sin\ t & 1-t \end{pmatrix}$$

   positive definite?

7. Let $\{x_i\}_{i=1}^n$ be the given eigenvectors of an unknown $n \times n$ matrix X, which is known to be similar to a diagonal matrix. A vector b and Xb are also given. Under what conditions on $b, Xb, \{x_i\}$ is the matrix X uniquely determined?

8. Let A, B be two nonsingular, $n \times n$ matrices which anticommute, i.e. $AB = -BA$. Show that n is even.

9*. Let A be a nonsingular $n \times n$ matrix, with only one nonvanishing element in each row (or each column). Show that A is similar to a diagonal matrix. Show that the set of such matrices forms a multiplicative group.

10*. For A any $n \times n$ matrix, let $\|A\|$ be as given in (4.23) and $\sigma(A)$ as given in (4.24). Prove that $\lim_{m\to\infty} \|A^m\|^{1/m} = \sigma(A)$.

Hint: Consider the action of $A^m = (A - \lambda_i + \lambda_i)^m$ on the subspace U_{ij} given in (3.8) and use Theorem 3.4.

11. Let Q, S be given $n \times n$ Hermitian matrices, with Q positive definite. Prove existence of an $n \times n$ matrix A satisfying $A^\dagger Q + QA = S$.

12. Let A, B be $n \times n$ matrices with B nonsingular, and satisfying $AB = BA^\dagger$. What can be said about the eigenvalues and eigenvectors of A?

13*. Let A be a given $n \times n$ matrix, regarded as a linear map of $H \to H, H$ of dimension n. Denote by $Z = \lim_{\varepsilon\to 0}(A^\dagger A + \varepsilon)^{-1} A^\dagger$.

(a) Show that Z determines a (bounded) linear map of $H \to H$, i.e. that for any $x \in H, Zx$ is uniquely determined within H. Find Z explicitly if $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$.

(b) Show that if A is nonsingular, then $Z = A^{-1}$.

(c) Show that for any $x \in H, x - ZAx = Px$, where P is the symmetric projection onto ker $(A^\dagger A)$.

(d) Show that Z is a pseudoinverse of A in the following sense: $ZAx = x$, for all $x \in R(A^\dagger A); y - AZy \perp R(A)$ for all $y \in R(A)$.

(e) Show that if A is normal (presumably singular) then $AZx = ZAx = x$ for all $x \in R(A)$.

14. Suppose $AB = BA$ and $BC = CB$, for A, B, C given $n \times n$ matrices. Under what condition on B may it be concluded that $AC = CA$?

15. Suppose $AB = BA, A, B$ $n \times n$ matrices, where B is of a special structure as given below. In each case, what may be concluded about the structure of A?

    a) $B = E_{jma}$, an elementary matrix as described in chapter 2

    b) B is nonsingular, with only one nonvanishing element in each row or each column.

    c) B is nonsingular and block diagonal

    d) $B_{ij} = 1$ for $j = i + 1$ and 0 otherwise

16* Given a square matrix A satisfying $A^2 = AA^\dagger$, show that A is similar to a diagonal matrix.

17* For any square matrix A, show that $\|A\|^2 = \sigma(A^\dagger A)$, with $\|A\|$ the induced norm of A given in (4.23).

    Hint: Consider $\|A(x + ty)\|$ for $t \in F, \|x\| = 1$ such that $\|Ax\| = \|A\|$, and $(x, y) = 0$.

18. Find the minimal polynomial $M_A(\lambda)$ of a matrix A which is both Hermitian and unitary, but is not a multiple of the identity matrix.

19. Let A be an $n \times n$ orthogonal (real, unitary) matrix with n odd. Prove that $\det A$ is an eigenvalue of A with odd multiplicity.

20. For I the $n \times n$ identity matrix and $w \in \mathbb{R}^n$, show that the matrix $I - 2ww^t$ is orthogonal if $\|w\| = 1$ (Euclidean norm).

21* Given an $n \times n$ real matrix A, show the existence of an orthogonal matrix Q and an upper triangular matrix R ($R_{ij} = 0$ for $i > j$) such that $A = QR$. This does not determine Q, R uniquely - explain.

22. Let A be an $n \times n$ matrix with distinct eigenvalues, and B be an $n \times n$ matrix commuting with A. Show that B can be written as a polynomial function of A, of degree not exceeding $n - 1$.

23* Let A be an $n \times n$ matrix whose minimal polynomial is of degree n, and let B be such that $AB = BA$. Show that B is a polynomial function of A, of degree not exceeding $n - 1$.

24. Let A, B, C be $n \times n$ matrices, satisfying $AB = BA$ and $C = AB - B^3 + 6I$. Furthermore C is nonsingular. Show that $AC^{-1} = C^{-1}A$.

25∗. Let A be a Hermitian $n \times n$ matrix, and denote by $\lambda_i^{(m)}, i = 1, \dots, m$ the eigenvalues of the truncated matrix $A^{(m)}$ as in the proof of Sylverter's theorem (theorem 4.13). Suppose that $\{\lambda_i^{(k)}, i = 1, \dots, k\} \subset \{\lambda_i^{(m)}, i = 1, \dots, m\}$ for all $1 \le k < m \le n$. Prove that A is diagonal.

26. Given an $n \times n$ matrix A which (simultaneously) satisfies $Q_1(A) = Q_2(A) = Q_3(A) = 0$, where $Q_1(\lambda) = \lambda^2(\lambda-1)(\lambda+2)(\lambda-3)^2, Q_2(\lambda) = (\lambda-1)^2(\lambda+3)(\lambda-3), Q_3(\lambda) = \lambda(\lambda-1)(\lambda+2)^3(\lambda-3)(\lambda+3)^2$

a) Prove that A^{-1} exists and is similar to a diagonal matrix.

b) If, in addition, it is known that A is Hermitian, show that $\|A^{-1}\| \le 1$. Give an example with $\|A^{-1}\| < 1$.

c) Give an example of non-Hermitian A such that $\|A^{-1}\| > 1$.

5 Matrices depending on a parameter

Very often one encounters matrices whose elements are not simply constants, but depend on one or more variables. Here we discuss the variation of various properties of such a matrix as these variables change in value.

For the present discussion, it will suffice to consider matrices depending on a single real variable t, for $t \in L$, L some open interval in $\mathbb{R}$. Furthermore, only square matrices will be discussed, so throughout this chapter all matrices are understood as of dimension $n \times n$, for some n fixed. Thus a matrix $A(t) = \{A_{ij}(t)_{i,j=1}^n\}, t \in L$, may be considered as determining a mapping (also denoted by A) of L into the space of $n \times n$ matrices over F, denoted here by M. The space M is equipped with any one of the matrix norms as appearing in chapter 1 or 2 above.

Then for A such a matrix-valued function on L, we have the obvious definitions of continuity and differentiability.

Definition 5.1. $A : L \to M$ is continuous at $t_0 \in L$ if for all $\varepsilon > 0$, there exists a $\delta > 0$ such that

$$\|A(t_0 + t) - A(t_0)\| < \varepsilon \text{ for all } |t| < \delta \tag{5.1}$$

Definition 5.2. $A : L \to M$ is differentiable at $t_0 \in L$, with derivative $A'(t_0) = \frac{dA}{dt}(t_0) \in M$, if there exists such a matrix $A'(t_0) \in M$ satisfying

$$\|A(t_0 + t) - A(t_0) - tA'(t_0)\| \leq o(|t|) \tag{5.2}$$

(The notation $o(|t|)$ is generic, for any function of t such that $o(|t|)/t \to 0$ as $t \to 0$.)

From these definitions, we have an obvious statement

Lemma 5.3. *$A : L \to M$ is continuous (resp. differentiable) at t_0 if and only if each of its components A_{ij} is a continuous (resp. differentiable) function of t at t_0.*

Further, if A is differentiable at t_0, then

$$A'(t_0)_{ij} = \frac{d}{dt} A_{ij}(t_0) \quad i, j = 1, \ldots, n, \tag{5.3}$$

and if A, B are both matrices differentiable at t_0, then so is the product AB, and

$$(AB)'(t_0) = A'(t_0)B(t_0) + A(t_0)B'(t_0) \tag{5.4}$$

Proof. The norm of a matrix is small if and only if the magnitude of each element is, so (5.3) follows immediately from (5.2). Then (5.4) follows from (5.3) and the ordinary expression for the product of two matrices.

As $\det A$, now a mapping of $L \to F$, is just a sum of products of elements of A, we have immediately

Lemma 5.4. *If A is continuous (resp. differentiable) at t_0, then so is $\det A$.*

There is a representation for $\frac{d}{dt}(\det A)$, which is occasionally used. For $k = 1, \ldots, n$ denote by $A_{(k)}, A^{(k)}$ the matrices with elements given by

$$\begin{aligned} A_{(k)ij} &= A_{ij}, i \neq k \\ A_{(k)kj} &= \frac{d}{dt} A_{kj} \end{aligned} \tag{5.5}$$

$$\begin{aligned} A_{ij}^{(k)} &= A_{ij},\ j \neq k \\ A_{ik}^{(k)} &= \frac{d}{dt} A_{ik}, \end{aligned} \tag{5.6}$$

i.e. one takes the t-derivative of the k-th row (resp. k-th column) of A to get $A_{(k)}$ (resp. $A^{(k)}$).

Lemma 5.5. *If A is differentiable at t_0, then*

$$\frac{d(\det A)}{dt}(t_0) = \sum_{k=1}^{n} \det A_{(k)}(t_0) = \sum_{k=1}^{n} \det A^{(k)}(t_0) \tag{5.7}$$

Proof. Differentiate (2.30) and (2.31) with respect to t, using the "product rule", and collect the differentiated elements rowwise or columnwise. Then using (5.5) or (5.6), one gets (5.7).

If A is continuous at t_0, and if $A(t_0)$ is nonsingular, then $A^{-1}(t_0)$ exists, and indeed from Lemma (5.4) it follows that $A^{-1}(t)$ exists for t in some neighborhood of t_0. And from an application of Cramer's rule, for nonsingular $A(t), A^{-1}(t)$ is just a sum of products of elements of $A(t)$ divided by $\det A(t)$. Thus we obtain

Lemma 5.6. *If A is continuous (resp. differentiable) at t_0, and if $A(t_0)$ is nonsingular, then A^{-1} is also continuous (resp. differentiable) at t_0. If A^{-1} is differentiable at t_0, then*

$$(A^{-1})'(t_0) \equiv \frac{d}{dt} A^{-1}(t_0) = -A^{-1}(t_0) A'(t_0) A^{-1}(t_0) \tag{5.8}$$

Proof. We need only obtain (5.8). Differentiating $A(t)A^{-1}(t) = I$ with respect to t and using (5.4), this is immediate.

Unlike (5.7), the result (5.8) is often used.

The behavior of the eigenvalues and eigenvectors of a matrix $A(t)$ is more interesting. We first illustrate the situation by several examples.

Example 1: $A(t) = \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}$ · $A(t)$ is similar to a diagonal matrix only for $t = 0$.

Example 2: $A(t) = \begin{pmatrix} 1 & t^2 \\ t & 1 \end{pmatrix}$ · $A(t)$ is similar to a diagonal matrix for all values of t, but the eigenvalues are $1 \pm t^{3/2}$, real for $t \geq 0$ and complex for $t < 0$. Although A is a smooth function of t, the eigenvalues are only once continuously differentiable in a neighborhood of $t = 0$. The corresponding eigenvectors are $\begin{pmatrix} \pm & t^{1/2} \\ & 1 \end{pmatrix}$, and are not differentiable at $t = 0$.

Example 3: $A(t) = \begin{pmatrix} 1+t & 1 \\ 0 & 1 \end{pmatrix}$. Now A is similar to a diagonal matrix for all $t \neq 0$. The eigenvalues are 1 and $1 + t$, with corresponding eigenvectors $\binom{1}{-t}$ and $\binom{1}{0}$. As $t \to 0$, the two eigenvectors coincide, i.e. linear independence is lost.

Example 4: $A(t) = \begin{pmatrix} 1 & 1 \\ t & 1 \end{pmatrix}$. Again A is similar to a diagonal matrix for all $t \neq 0$. The eigenvalues are $1 \pm t^{1/2}$, with corresponding eigenvectors $\begin{pmatrix} 1 & \\ \pm & t^{1/2} \end{pmatrix}$. A is a smooth function of t, but the eigenvalues and eigenvectors change from real to complex, and are not differentiable, at $t = 0$.

These examples share a common feature - the eigenvalues become equal at $t = 0$. When the eigenvalues of a matrix are distinct, both the eigenvalues and eigenvectors are as smooth, with respect to t, as the matrix itself.

Theorem 5.7. *If A is continuous at t_0, then so are its eigenvalues.*

Proof. $P_{A(t)}(\lambda)$ is a polynomial of order n, whose leading coefficient is $(-1)^n$, i.e. not zero uniformly in t, and the other coefficients of which are continuous functions at t at $t = t_0$. Therefore so are the roots of $P_{A(t)}(\lambda)$, which are the eigenvalues of $A(t)$.

Corollary. *If the eigenvalues of $A(t_0)$ are distinct, and if A is continuous at t_0, then the eigenvalues of A are distinct for t is an open neighborhood of t_0.*

Thus we consider such a matrix A with distinct eigenvalues, $\lambda_i \neq \lambda_j$ for $i \neq j$. Thus $m_i = \mu_i = 1$ for all values of i and $k = n$ in (3.4), (3.7). Denote by $\{x_i\}_{i=1}^n$ the corresponding eigenvectors,

$$Ax_i = \lambda_i x_i, \ i = 1, \ldots, n \tag{5.9}$$

which are linearly independent, using theorem 3.3. We normalize the eigenvectors so that $\|x_i\| = 1$.

For each $t, t = 1, \ldots, n, x_i(t) \in H$, and $A(t) = H \to H, H$ of dimension n. We regard H as a Hilbert space, equipped with the inner product as given in (1.13) or (1.14).

Now as $\bar{\lambda}_i \neq \bar{\lambda}_j$ for $i \neq j$, the eigenvalues of $A^\dagger$ are also distinct, and thus there is a full set of eigenvectors for $A^\dagger$, also linearly independent,

$$A^\dagger y_i = \bar{\lambda}_i y_i, \ i = 1, \ldots, n. \tag{5.10}$$

Taking the adjoint of (5.10), we get

$$y_i^\dagger A = \lambda_i y_i^\dagger, \tag{5.11}$$

i.e. the $y_i^\dagger$, viewed as row vectors, are thus the left eigenvectors for A.

Lemma 5.8. *For $i \neq j, (x_i, y_j) = 0$. Furthermore $(x_i, y_i) \neq 0$ for $i = 1, \ldots, n$.*

Proof. For all $i \neq j$,

$$\begin{aligned} (Ax_i, y_j) &= \lambda_i (x_i, y_j) \\ &= (x_i, A^\dagger y_j) \\ &= \lambda_j (x_i, y_j) \end{aligned}$$

so as $\lambda_i \neq \lambda_j, (x_i, y_j) = 0$. As the $\{x_i\}$ span H, since $y_i \in H$ is not zero, it cannot be perpendicular to x_i. Typically the y_i are normalized by the condition

$$(x_i, y_j) = \delta_{ij}, i, j = 1, \ldots, n \tag{5.12}$$

Theorem 5.9. *Assume that the eigenvalues of $A(t_0)$ are distinct. If A is continuous at t_0, so are the eigenvalues and eigenvectors. If A is differentiable* *at t_0, then so are the eigenvalues and eigenvectors, and their t-derivatives are given explicitly by*

$$\lambda_i'(t_0) = (A'(t_0)x_i, y_i) \tag{5.13}$$

$$x_i'(t_0) = \sum_{j \neq i} \frac{(A'(t_0)x_i, y_j)}{\lambda_i - \lambda_j} x_j + \alpha_i x_i, \;\; i = 1, \ldots, n, \tag{5.14}$$

assuming the x_i, y_j satisfy (5.12). The constants α_i are arbitrary, determined by the normalization condition on the $\{x_i\}$.

In (5.13), (5.14) we have abbreviated $x_i(t_0)$ by x_i, etc., and as above, primes indicate t-derivatives.

Proof. For A continuous at t_0, we have already seen that the eigenvalues $\{\lambda_i\}$ are continuous at t_0, so it suffices to check the eigenvectors $\{x_i\}$. For $\varepsilon > 0$ small, satisfying in particular

$$\varepsilon < \frac{1}{3} \underset{i \neq j}{\text{minimum}} |\lambda_i(t_0) - \lambda_j(t_0)| \tag{5.15}$$

we use the continuity of A and the $\{\lambda_i\}$, of which there are only a finite number, to obtain $\delta > 0$ such that

$$\|A(t) - A(t_0)\| < \varepsilon, \tag{5.16}$$

$$|\lambda_i(t) - \lambda_i(t_0)| < \varepsilon, i = 1, \ldots, n, \text{ for all } t \in (t_0 - \delta, t_0 + \delta). \tag{5.17}$$

From (5.17) and (5.15) an application of the triangle inequality gives

$$\begin{aligned} |\lambda_i(t) - \lambda_j(t)| &> \frac{1}{3} |\lambda_i(t_0) - \lambda_j(t_0)| \\ &\geq \text{ constant } > 0 \end{aligned} \tag{5.18}$$

for all $i \neq j, t \in (t_0 - \delta, t_0 + \delta)$, i.e. the eigenvalues remain distinct, uniformly with respect to t in this interval.

Now fix t arbitrarily within the interval $(t_0 - \delta, t_0 + \delta)$; since the $\{x_i(t_0)\}_{i=1}^n$ form a basis for H, we can make an expansion

$$x_i(t) = \sum_{j=1}^n \alpha_j x_j(t_0). \tag{5.19}$$

Our claim that x_i is continuous at t_0 will now follow if we can show that the $|\alpha_j|/|\alpha_i|$ are small for $j \neq i$. For this we use (5.9), (5.16) and (5.17)

$$\begin{aligned} A(t)x_i(t) &= \lambda_i(t)x_i(t) \\ &= (\lambda_i(t_0) + O(\varepsilon))x_i(t) \\ &= (\lambda_i(t_0) + O(\varepsilon)(\alpha_i x_i(t_0) + \sum_{j\neq i} \alpha_j x_j(t_0)) \end{aligned} \tag{5.20}$$

$$\begin{aligned} A(t)x_i(t) &= A(t)(\alpha_i x_i(t_0) + \sum_{j\neq i} \alpha_j x_j(t_0)) \\ &= (A(t_0) + O(\varepsilon))(\alpha_i x_i(t_0) + \sum_{j\neq i} \alpha_j x_j(t_0)) \\ &= \lambda_i(t_0)\alpha_i x_i(t_0) + \sum_{j\neq i} \alpha_j \lambda_j(t_0) x_j(t_0) \\ &+ O(\varepsilon))(\alpha_i x_i(t_0) + \sum_{j\neq i} \alpha_j x_j(t_0)) \end{aligned} \tag{5.21}$$

Here and below, we employ the common generic notation $O(\varepsilon)$ for quantities satisfying $|O(\varepsilon)| \leq$ constant ε for $\varepsilon > 0$ small.

Comparing (5.20) and (5.21) we find

$$\begin{aligned} \sum_{j\neq i} \alpha_j(\lambda_i(t_0) - \lambda_j(t_0) + O(\varepsilon))x_j(t_0) &= O(\varepsilon)\alpha_i x_i(t_0) \\ &= O(\varepsilon), \end{aligned} \tag{5.22}$$

and for $\varepsilon > 0$ small, $j \neq i$, $\lambda_i(t_0) - \lambda_j(t_0) + O(\varepsilon)$ is bounded away from zero, and so (5.22) implies $|\alpha_j/\alpha_i| = O(\varepsilon)$ for all $j \neq i$. This proves the first part of Theorem 5.9.

Next suppose that A is differentiable at t_0. Then there exists a matrix $A'(t_0)$, such that for $|t - t_0|$ small,

$$\|A(t) - A(t_0) - (t - t_0)A'(t_0)\| \leq o(|t - t_0|). \tag{5.23}$$

Our theorem asserts the existence of $\lambda_i'(t_0) \in F, x_i'(t_0) \in H, i = 1, \ldots, n$ given by (5.13), (5.14), respectively, such that

$$|\lambda_i(t) - \lambda_i(t_0) - (t - t_0)\lambda_i'(t_0)| \leq o(|t - t_0|) \tag{5.24}$$

$$\|x_i(t) - x_i(t_0) - (t - t_0)x_i'(t_0)\| \leq o(|t - t_0|) \tag{5.25}$$

We make the expansions

$$\lambda_i(t) = \lambda_i(t_0) + (t - t_0)\gamma_i + \theta_i(t) \tag{5.26}$$

$$x_i(t) = x_i(t_0) + (t - t_0)\sum_{j=1}^{n} \alpha_j x_j(t_0) + \eta_i(t) \tag{5.27}$$

in which γ_i and the α_j are to be chosen below and θ_i, η_i thus determined. Our objective, of course, is to choose γ_i and the α_i such that $\theta_i(t)$ and $\eta_i(t)$ are $o(|t - t_0|)$, and thus γ_i is identified with $\lambda_i'(t_0)$ and $\sum \alpha_j x_j(t_0)$ with $x_i'(t_0)$ using (5.24), (5.25).

With $\lambda_i(t), x_i(t)$ given by (5.26), (5.27), and $A(t)$ satisfying (5.23), (5.9) still has to be satisfied at each value of t,

$$\begin{aligned} A(t)x_i(t) &= \lambda_i(t)x_i(t) \\ &= (\lambda_i(t_0) + (t - t_0)\gamma_i + \theta_i(t))(x_i(t_0) + (t - t_0)\sum_j \alpha_j x_j(t_0) + \eta_i(t)) \\ &= \lambda_i(t_0)x_i(t_0) + (t - t_0)[\gamma_i x_i(t_0) + \lambda_i(t_0)\sum_j \alpha_j x_j(t_0)] \\ &\quad + \theta_i(t)x_i(t) + \lambda_i(t)\eta_i(t) + O(t - t_0)^2 \end{aligned} \tag{5.28}$$

using (5.26), (5.27). Using (5.23) and (5.27),

$$\begin{aligned} A(t)x_i(t) &= (A(t_0) + (t - t_0)A'(t_0) + o(|t - t_0|)) \\ &\quad (x_i(t_0) + (t - t_0)\sum_j \alpha_j x_j(t_0) + \eta_i(t)) \\ &= A(t_0)x_i(t_0) + (t - t_0)[A'(t_0)x_i(t_0) \\ &\quad + A(t_0)\sum_j \alpha_j x_j(t_0)] + A(t)\eta_i(t) + o(|t - t_0|) \end{aligned} \tag{5.29}$$

Comparing (5.28) and (5.29), we separate the linear terms in $(t - t_0)$ from those of "higher order", and require each to be satisfied separately. Thus we choose γ_i and $\{\alpha_j\}$ so that

$$A'(t_0)x_i(t_0) + \sum_{j=1}^{n} \alpha_j(\lambda_j(t_0) - \lambda_i(t_0))x_j(t_0) = \gamma_i x_i(t_0) \tag{5.30}$$

and then θ_i and η_i must satisfy

$$(A(t) - \lambda_i(t))\eta_i(t) - \theta_i(t)x_i(t) = o(|t - t_0|). \tag{5.31}$$

Taking the inner product of (5.30) with $y_i(t_0)$ and with $y_k(t_0)$ for some $k \neq i$, we find, using (5.12)

$$\gamma_i = (A'(t_0)x_i(t_0), y_i(t_0)) \tag{5.32}$$

$$\alpha_k = (A'(t_0)x_i(t_0), y_k(t_0))/(\lambda_i(t_0) - \lambda_k(t_0)),\ k \neq i; \tag{5.33}$$

(5.13) now follows from (5.32), (5.24) and (5.26), and (5.14) from (5.33), (5.25) and (5.27), but we still have to show that $\theta_i(t)$ and $\eta_i(t)$ are $o(|t-t_0|)$.

Denote by ζ the right side of (5.31). Taking the inner product of (5.31) with $y_i(t)$ and $y_j(t), j \neq i$, using (5.12) again, we find

$$\theta_i(t) = (-\zeta, y_i(t)) \tag{5.34}$$

$$\eta_i(t) = \sum_{j \neq i} (\zeta, y_j(t))x_j(t)/(\lambda_j(t) - \lambda_i(t)) \tag{5.35}$$

Now using (5.18), $|\zeta| = o(|t - t_0|)$, and $\|x_j(t)\| = 1$, we find from (5.34), (5.35)

$$|\theta_i(t)| + \|\eta_i(t)\| \leq o(|t - t_0|) \underset{j}{\rightarrow} \text{maximum}\|y_j(t)\|, \tag{5.36}$$

so our proof is reduced to showing that the $\|y_j(t)\|$ are uniformly bounded for $|t - t_0| < \delta$. This is not completely trivial, as the $y_j(t)$ are normalized by (5.12), with each $\|x_i(t)\| = 1$.

We apply the first part of the theorem to $A^\dagger(t)$. If A is continuous, at t_0, so is $A^\dagger$, and as the eigenvalues $\lambda_i(t_0)$ of $A(t_0)$ are distinct, so are the $\bar{\lambda}_i(t_0)$, the eigenvalues of $A^\dagger(t_0)$. Therefore by the first part of the theorem, the normalized eigenvectors of $A^\dagger, y_j/\|y_j\|,\ j = 1, \ldots, n$ are continuous at t_0. If x_i and $y_i/\|y_j\|$ are continuous at t_0, then so is their product, $(x_i, y_i)/\|y_i\| = 1/\|y_i\|$. Thus the $\|y_i\|$ are continuous at t_0, and since there are only a finite number of them, taking δ smaller if necessary, they are uniformly bounded with respect to t in $(t_0 - \delta, t_0 + \delta)$.

This completes the proof.

When two or more of the eigenvalues of $A(t_0)$ are equal, differentiability of the eigenvalues or eigenvectors cannot be expected in general. Examples given above illustrate this. Furthermore, we find factors of $\lambda_i(t_0) - \lambda_j(t_0)$ in the denominator of (5.14) and (5.35), and numerous appearance of $y_j(t_0)$ normalized by (5.12). If $x_i(t) - x_j(t) \to 0$ as $t \to t_0$ with $\|x_i(t)\| = \|x_j(t)\| = 1$, then $\|y_j(t)\| \to \infty$ as $t \to t_0$.

Nevertheless, we saw in Theorem 5.7 that if A is continuous at t_0, then so are its eigenvalues, and it is obviously natural to ask if there is not some corresponding condition on the eigenvectors. Indeed, in the simple examples given above one will note that when $A(t_0)$ is not similar to a diagonal matrix but $A(t)$ is, for $t \neq t_0$, the two eigenvectors coalesce as $t \to t_0$, and the limit is indeed the one eigenvector of $A(t_0)$. This is not accidental.

The subspaces U_{ij} discussed in chapter 3 satisfy a continuity condition, even though their dimensions may change abruptly as the matrix $A(t)$ varies.

Theorem 5.10. *Let* $U_{ij}(t_0) = \ker(A(t_0) - \lambda_i(t_0))^j$, *and suppose that* $A(t), \lambda(t), x(t)$ *satisfy*

$$|\lambda(t) - \lambda_i(t_0)| \leq \varepsilon \tag{5.37}$$

$$\|A(t) - A(t_0)\| \leq \varepsilon \tag{5.38}$$

$$\|x(t)\| = 1 \tag{5.39}$$

$$x(t) \in \ker(A(t) - \lambda(t))^j, i.e. (A(t) - \lambda(t))^j x(t) = 0 \tag{5.40}$$

Then there exists $x_0 \in U_{ij}(t_0)$ *and a constant* c, *independent of* ε, *such that*

$$\|x(t) - x_0\| \leq c\varepsilon. \tag{5.41}$$

As an example, suppose that A is continuous at t_0, and that $\lambda_i(t_0)$ is of multiplicity $m_i(t_0) > 1$. Suppose further that $m_i(t_0) = \mu_i(t_0)$ in (3.4), (3.7), so there is only one Jordan block for $\lambda_i(t_0)$ and $U_{i1}(t_0)$ is of dimension 1, spanned by x_0, with $\|x_0\| = 1$. For $|t - t_0| \neq 0$ but small, there will be $m_i(t_0)$ eigenvalues of $A(t)$ close to $\lambda_i(t_0)$ by Theorem 5.7. Suppose these eigenvalues are all simple (distinct), i.e. there are m_i linearly independent corresponding eigenvectors $x(t)$, normalized by $\|x(t)\| = 1$. Then as $t \to t_0$, all of these eigenvectors coalesce, and their limit as $t \to t_0$ is x_0 (up to multiplication by a phase factor). This is the phenomenon we observed in example 3 and 4 above.

Proof. Denote by

$$W = \bigoplus_{\substack{\ell=1 \\ \ell \neq i}}^{k(t_0)} U_{\ell\mu_\ell(t_0)}(t_0), \tag{5.42}$$

$$Z = \{z \in U_{i\mu_i(t_0)}^{(t_0)} | z \perp U_{ij}(t_0)\}; \tag{5.43}$$

in (5.42), (5.43), $k(t_0)$ is the k appearing in (3.4), (3.7), etc. for the matrix $A(t_0), \mu_\ell(t_0)$ the corresponding μ_ℓ in (3.7), and

$$U_{\ell\mu_\ell(t_0)}(t_0) = \ker(A(t_0) - \lambda_\ell(t_0))^{\mu_\ell(t_0)} \tag{5.44}$$

as discussed in chapter 3. (For $j = \mu_i(t_0)$ in the statement of Theorem 5.12, the space Z is not included.) Now from Theorem 3.3, noting that $U_{i\mu_i(t_0)}(t_0) = U_{ij}(t_0) \oplus Z$ and $U_{ij}(t_0) \cap Z = \{0\}$, the spaces $W, U_{ij}(t_0)$ and Z are linearly independent, that is there are no nontrivial linear combinations of elements from these spaces equal to zero. All of these spaces are closed, finite dimensional subspaces of H. From Theorem 3.4, these spaces together span H,

$$H = W \oplus U_{ij}(t_0) \oplus Z, \tag{5.45}$$

and therefore we have

$$x(t) = w + x_0 + z, \ w \in W, \ x_0 \in U_{ij}(t_0), \ z \in Z. \tag{5.46}$$

Using $x(t)$ given by (5.46) in (5.40) with the estimates (5.37), (5.38), we obtain

$$\begin{aligned}(A(t) - \lambda(t))^j x(t) = 0 \\ &= (A(t_0) - \lambda_i(t_0) + O(\varepsilon))^j x(t) \\ &= (A(t_0) - \lambda_i(t_0))^j (w + x_0 + z) + O(\varepsilon)\|x(t)\| \\ &= (A(t_0) - \lambda_i(t_0))^j (w + z) + O(\varepsilon)\end{aligned} \tag{5.47}$$

using (5.39) and $x_0 \in U_{ij}(t_0)$ is the last step.

We claim that there exists a positive constant τ such that for any $q \in W \oplus Z$

$$\|(A(t_0) - \lambda_i(t_0))^j q\| \geq \tau \|q\|; \tag{5.48}$$

putting $q = w + z$ in (5.47) and using (5.46), will obtain (5.41) immediately. Thus it suffices to prove the claim. Suppose (5.38) fails, then there is an infinite sequence $\{q_s\}_{s=1}^{\infty}$, such that each $q_s \in W \oplus Z$, $\|q_s\| = 1$, and

$$\|(A(t) - \lambda_i(t_0))^j q_s\| \to 0 \text{ as } s \to \infty. \tag{5.49}$$

Since $W \oplus Z$ is finite dimensional, taking a subsequence of $\{q_s\}$ is necessary, we may assume that $\{q_s\}$ converges to a limit q_0, which will satisfy $\|q_0\| = 1$, $q_0 \in W \oplus Z$, as $W \oplus Z$ is closed, and by continuity

$$\|(A(t_0) - \lambda_i(t_0)^j q_0\| = 0 \tag{5.50}$$

But from (5.50), $q_0 \in U_{ij}(t_0)$, which contradicts the linear independence of $W, U_{ij}(t_0)$ and Z, thus proving the claim and the theorem.

Problems

1. Suppose the derivative of a matrix A were defined by (5.23). Show that $A'(t_0)$ thus determined, if it exists, is unique.

2. Suppose $\|x_j(t)\|$, $\|x_i(t)\| = 1$ and $\|x_i(t) - x_j(t)\| \to 0$ as $t \to t_0$. Let the y_j be normalized by (5.12). Show that $\|y_i(t)\| \to \infty$ as $t \to t_0$.

3. Find explicit expressions for the constants α_i in (5.14), assuming the α_i real and that the normalization of the x_i is via $\|x_i\| = (x_i, x_i)^{1/2} = 1$. Explain why it is no loss of generality to assume the α_i real.

4. Suppose that the eigenvalues of $A(t_0)$ are distinct and that all the eigenvalues $\lambda_i(t)$ and eigenvectors $x_i(t)$ are continuous at t_0. Show that A is continuous at t_0.

5. True or false: Let $x_i(t), x_j(t)$ be normalized eigenvectors of $A(t)$, corresponding to eigenvalues $\lambda_i(t), \lambda_j(t)$. For $t \neq t_0, \lambda_i, \lambda_j$ are simple eigenvalues. Furthermore, A is continuous at t_0. As $t \to t_0, x_i(t) - x_j(t) \to 0$. Then, necessarily, $\lambda_i(t) - \lambda_j(t) \to 0$ as $t \to t_0$, and $A(t_0)$ is <u>not</u> similar to a diagonal matrix. Either prove this result or find a counterexample.

6. Given a unitary matrix U, find a matrix $A(t)$, depending smoothly on t and unitary for all $t \in [0, 1]$, such that $A(0) = I$ and $A(1) = U$. Under what conditions on U can $A(t)$ be made real for all $t \in [0, 1]$?

6 Elementary functions of matrices

Throughout this chapter, matrices are understood as of dimension $n \times n$ for arbitrary fixed n unless stated otherwise. For definiteness, we take $F = \mathbb{C}$; the particular results for real matrices will be stated explicitly.

Elementary functions of scalar variables satisfy numerous identities and differential equations. Here we wish to define elementary functions of matrices, in such a way that many, if not all of these properties are preserved.

Our point of departure is to note that elementary functions often admit representations as infinite power series

$$f(z) = \sum_{j=0}^{\infty} a_j z^j \tag{6.1}$$

for example, converging in some region $|z| < \rho$. Then the series (6.1) converges absolutely in this region, and so do sums and products of such series, the series obtained by term by term differentiation or integration, etc. Thus many formal computations with such series are in fact justified, and numerous properties of the function $f(z)$ obtained in this manner.

We can do the same for elementary functions of matrices, determined analogously from

$$f(A) = \sum_{j=0}^{\infty} a_j A^j, \tag{6.2}$$

where A is some matrix, $a_j \in F$, and $A^0 = I$ by convention, provided the series (6.2) also converges. This question, however, has a simple answer.

Theorem 6.1. *If the series (6.1) converges for $|z| < \rho$, then the series (6.2) converges absolutely for A such that*

$$\sigma(A) < \rho. \tag{6.3}$$

Here $\sigma(A)$ is the spectral radius of A, given in (4.24). We shall prove this by a method which provides a procedure for computing $f(A)$ explicitly.

Proof. We have $A = T^{-1}JT, T$ nonsingular, J in Jordan canonical form. For any partial sum

$$
\begin{aligned}
f_m(A) &= \sum_{j=0}^{m} a_j A^j \\
&= \sum_{j=0}^{m} a_j (T^{-1}JT)^j \\
&= \sum_{j=0}^{m} a_j T^{-1} J^j T \\
&= T^{-1} f_m(J) T, \qquad (6.4)
\end{aligned}
$$

so if the partial sums $f_m(J)$ converge absolutely, so do the $f_m(A)$, and in addition we shall have established

$$f(A) = T^{-1} f(J) T \tag{6.5}$$

for the infinite series as well.

Thus it suffices to consider the $f_m(J)$, which is facilitated by the simple structure of the Jordan form J. In particular, J, is block diagonal, with each block of dimension no greater than μ, μ the largest of the exponents μ_i appearing in the minimal polynomial for A, given by (3.7). In any event, $\mu \leq n$. Now since J is block diagonal, so are all the J^j and thus the $f_m(J)$, with the same block structure, by application of lemma 2.9. Thus it suffices to prove the convergence of $f_m(J)$, restricting attention to any single Jordan block.

The matrix J, restricted to a single block, of dimension k, say, is of a very special structure

$$J = \lambda I_k + L \tag{6.6}$$

where I_k is the $k \times k$ identity matrix, L a $k \times k$ matrix satisfying

$$L_{ij} = 1 \text{ for } j = i + 1, \;\; L_{ij} = 0 \text{ otherwise} \tag{6.7}$$

and λ is one of the eigenvalues of J, or of A. In particular

$$|\lambda| \leq \sigma(A). \tag{6.8}$$

The matrix L obviously commutes with I_k; furthermore, the powers of L are easily seen to satisfy

$$(L^m)_{ij} = 1 \text{ if } j = i + m, \;\; 0 \text{ otherwise}, \tag{6.9}$$

and since L is of dimension k, with k not exceeding μ,

$$L^{\mu} = 0. \tag{6.10}$$

As the series (6.1) converges for $|z| < \rho$, it follows that

$$|a_j| \; |z|^j \le K(|z|) \quad |z| < \rho \tag{6.11}$$

where the bound K depends on $|z|$ (possibly blowing up as $|z| \to \rho$) but is independent of j.

Thus for any $N > M$, the absolute sum of the terms in $f_N(J) - f_M(J)$ is

$$\begin{aligned}
&\sum_{j=M+1}^{N} |a_j| \; \| \, (\lambda I_k + L)^j \| \\
&\qquad = \sum_{j=M+1}^{N} |a_j| \; \| \sum_{m=0}^{j} \binom{j}{m} \lambda^{j-m} L^m \| \\
&\qquad = \sum_{j=M+1}^{N} |a_j| \; \| \sum_{m=0}^{\mu-1} \binom{j}{m} \lambda^{j-m} L^m \| \\
&\qquad \le \sum_{m=0}^{\mu-1} \|L^m\| \sum_{j=M+1}^{N} |a_j| \binom{j}{m} |\lambda|^{j-m} \\
&\qquad \le \sum_{m=0}^{\mu-1} \|L^m\| \frac{1}{\sigma(A)^m m!} \sum_{j=M+1}^{N} |a_j| j^m \sigma(A)^j
\end{aligned} \tag{6.12}$$

using (6.10), (6.8), and $\binom{j}{m} = j!/(m!(j-m)!)$.

In (6.12), choose M initially sufficiently large that

$$\sigma(A) j^{(\mu-1)/j} \le (2\sigma(A) + \rho)/3 \text{ for } j > M, \tag{6.13}$$

using $\sigma(A) < \rho$. Then the second sum in (6.12) is estimated

$$
\begin{aligned}
\sum_{j=M+1}^{N} |a_j| \left(\frac{2\sigma(A) + \rho}{3} \right)^j & \\
&\leq \sum_{j=M+1}^{N} |a_j| \left(\frac{\sigma(A) + 2\rho}{3} \right)^j \left(\frac{2\sigma(A) + \rho}{\sigma(A) + 2\rho} \right)^j \\
&\leq K \left(\frac{\sigma(A) + 2\rho}{3} \right) \sum_{j=M+1}^{\infty} \left(\frac{2\sigma(A) + \rho}{\sigma(A) + 2\rho} \right)^j \\
&= K \left(\frac{\sigma(A) + 2\rho}{3} \right) \left(\frac{\sigma(A) + 2\rho}{\rho - \sigma(A)} \right) \left(\frac{2\sigma(A) + \rho}{\sigma(A) + 2\rho} \right)^{M+1} \qquad (6.14)
\end{aligned}
$$

Again using $\sigma(A) < \rho$, the right side of (6.14) is made arbitrarily small by choosing M sufficiently large. This completes the proof.

Proceeding as in obtaining (6.12), we see that the corresponding block of $f(J)$ is given explicitly by

$$\sum_{m=0}^{\mu-1} L^m \sum_{j=m}^{\infty} a_j \lambda^{j-m} \binom{j}{m} = \sum_{m=0}^{\mu-1} \frac{f^{(m)}(\lambda)}{m!} L^m \qquad (6.15)$$

which reduces to $f(\lambda)$ in the case of $\mu = 1$, i.e. this block of J in fact diagonal.

Some implications of this result are worth noting. If A is a real matrix, if $f(z)$ is real for z real and if $\sigma(A) < \rho$, then $f(A)$, defined by (6.2), is a real matrix - the eigenvalues of A do not have to be real.

Entire functions of matrices, like e^A and $\sin A$ are defined for any matrix, and satisfy relations such as

$$\sin^2 A + \cos^2 A = I, \frac{d}{dt}(e^{tA}) = Ae^{tA} \qquad (6.16)$$

as these are readily obtained from the power series expansions. But for two matrices, A, B in general

$$e^{A+B} \neq e^A e^B, \sin(A + B) \neq \sin A \cos B + \cos A \sin B \qquad (6.17)$$

because the proof of such results for scalars, based on the power series expansions for the respective functions, depends on interchanging products BA and AB. Thus equality applies in relations such as (6.17) if A, B commute.

The regularity of a function $f(A)$ so determined is described by the following lemma:

Lemma 6.2. *Suppose the series (6.1) converges for $|z| < \rho$, and that $\sigma(A) < \rho$. Then for any matrix $B, f(A + tB)$ is defined and is an infinitely differentiable function of the real parameter t in some neighborhood of $t = 0$. Furthermore,*

$$\text{(6.18)} \qquad \textit{If } AB = BA, \frac{d}{dt} f(A + tB)|_{t=0} = f'(A)B.$$

Proof. As $\sigma(A) < \rho, \sigma(A + tB) < \rho$ for $|t|$ sufficiently small, and thus formal differentiation any number of times of the power series for $f(A + tB)$ is justified. If A, B commute, the power series for $A + tB, |t|$ small, may be reassembled in the form $f(A) + tf'(A)B + O(t^2)$, and (6.18) follows.

In view of (6.10), the right side of (6.15) is just the formal Taylor series for $f(\lambda I + L)$. The generalization of this is the following lemma:

Lemma 6.3. *Suppose A, B commute, and $\sigma(A + tB) < \rho$ for all $0 \leq t \leq 1$, where the series (6.1) converges for $|z| < \rho$. Then for any integer m*

$$\text{(6.19)} \quad f(A + B) = \sum_{j=0}^{m} \frac{f^{(j)}(A)}{j!} B^j + \frac{1}{m!} \int_0^1 (1 - t)^m f^{(m+1)}(A + tB) dt B^{m+1}.$$

Proof. For any $t \in [0, 1], \sigma(A + tB) < \rho$ and $A + tB$ commutes with B, so

$$\text{(6.20)} \qquad \frac{d}{dt} f(A + tB) = f'(A + tB)B;$$

integrating (6.20) we obtain

$$\text{(6.21)} \qquad f(A + B) - f(A) = \int_0^1 f'(A + tB) dt B.$$

The higher derivatives of $f(A + tB)$ are defined by their power series, which is known to converge. Therefore we simply apply m partial integrations to (6.21) to obtain (6.19).

Other functions of matrices are defined as the inverse functions of functions defined by convergent power series. In this manner, one is often able to define such a function in a larger region than would be possible by a power series. Thus the formal power series for $\ell n(I + A)$, for example, converges for $\sigma(A) < 1$, but we shall see that in fact any nonsingular matrix has a logarithm.

Theorem 6.4. *Suppose the series (6.1) converges for $|z| < \rho$. Given a matrix Z suppose that for each eigenvalue λ_i of Z there exists η_i, such that*

$$|\eta_i| < \rho \tag{6.22}$$

$$f(\eta_i) = \lambda_i \tag{6.23}$$

and

$$f'(\eta_i) \neq 0. \tag{6.24}$$

Then $f^{-1}(Z)$ is defined, i.e. there exists a matrix X such that $f(X) = Z$, with $\sigma(X) < \rho$.

Proof. Put $Z = T^{-1}JT, J$ in Jordan form. If we can solve

$$f(Y) = J \tag{6.25}$$

for a matrix Y such that

$$\sigma(Y) < \rho,$$

then $X = T^{-1}YT$ satisfies $f(X) = Z$,

$$\begin{aligned} f(T^{-1}YT) &= T^{-1}f(Y)T \\ &= T^{-1}JT \\ &= Z. \end{aligned} \tag{6.26}$$

Of course we seek Y in the same block diagonal structure as that of J. For a block of dimension k, J is of the form (6.6), with L given by (6.7) and λ one of the eigenvalues of Z. By hypothesis, therefore, there exists η satisfying

$$|\eta| < \rho \tag{6.27}$$

$$f(\eta) = \lambda \tag{6.28}$$

$$f'(\eta) \neq 0. \tag{6.29}$$

We seek the corresponding block of the matrix Y in the form

$$\eta I_k + \sum_{j=1}^{k-1} \zeta_j L^j \tag{6.30}$$

noting that $L^k = 0$, and the lower powers of L obtained from (6.9). The ζ_j are scalars to be determined.

We wish to apply (6.19) to (6.30), identifying A in (6.19) with $\eta I_k, B$ with $\sum \zeta_j L^j$, and m in (6.19) with $k-1$. Obviously A, B commute, and the remainder term in (6.19) vanishes, because of the factor L^k. From the structure of L and (6.30), B is upper triangular with zero diagonal elements, and thus for any $t, \sigma(A+tB) = \eta < \rho$. Therefore in this case

$$\begin{aligned}
f(A+B) &= f(A) + f'(A)B + \cdots + \frac{1}{(k-1)!} f^{(k-1)}(A)B^{k-1} \\
&= f(\eta)I_k + \sum_{m=1}^{k-1} \frac{1}{m!} f^{(m)}(\eta I_k)(\sum_{j=1}^{k-1} \zeta_j L^j)^m \\
&= f(\eta)I_k + \sum_{j=1}^{k-1} L^j (f'(\eta)\zeta_j + g_j(f'(\eta), \ldots, f^{(j)}(\eta), \zeta_1 \ldots, \zeta_{j-1}))
\end{aligned} \tag{6.31}$$

where the functions g_j are obtained by expanding the sum in (6.31) and collecting the powers of L. In particular, g_j includes all the quadratic and higher order terms in the ζ_i, and therefore g_j depends on the ζ_i only for $i < j$, and g_1 vanishes identically.

We wish to solve $f(A+B) = \lambda I_k + L$; using (6.31) and comparing powers of L, using (6.28) it will suffice that

$$f'(\eta)\zeta_1 = 1, f'(\eta)\zeta_j + g_j = 0, j = 2, \ldots, k-1 \tag{6.32}$$

which is easily solved recursively using (6.29). Independently of the ζ_j, we will have $\sigma(A+B) = |\eta| < \rho$, so the matrix Y determined in the manner indeed satisfies (6.25).

The matrix $f^{-1}(Z)$ so obtained will be real if Z is a real matrix with real eigenvalues, the η_i can be taken real valued, and $f(z)$ is real for real z. Thus lnZ is real, for example, for a real matrix Z with positive (real) eigenvalues, and a real $\sin^{-1} Z$ exists for real Z with real eigenvalues and $\sigma(Z) < 1$.

Problems

1. Evaluate explicitly: $\exp\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}, \ell n\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}, \sin\begin{pmatrix}\pi/2 & 1\\ 0 & \pi/2\end{pmatrix}, \tan^{-1}\begin{pmatrix}1 & 1\\ 0 & 1\end{pmatrix}$.

2. Show that e^A is nonsingular for any matrix A.

3. Find an example of nonsingular A, B such that $\ell n(AB) \neq \ell nA + \ell nB$

4. Suppose that Z satisfies the conditions of Theorem 6.4, so that $X = f^{-1}(Z)$ is defined. Show that $f'(X)$ is nonsingular; and that $XZ = ZX$.

5. Show directly that $\sin^{-1}\begin{pmatrix}1 & 1\\ 0 & 1\end{pmatrix}$ does not exist. Hint: Any matrix X commutes with $\sin X$, so the possible structure of such a matrix is limited.

6. Find an example of a nonsingular 2×2 matrix A such that $A(\sin A) = -(\sin A)A$

7. Show that if $\sin A = 0$, then A is similar to a diagonal matrix, each of whose diagonal elements is an integer multiple of π.

8. Evaluate $ln\begin{pmatrix}2 & 1 & 0\\ 0 & 2 & 0\\ 0 & 0 & 1\end{pmatrix}$

9. Classify the matrices $A \in M_{2\times 2}(\mathbb{C})$ such that there is no $Z \in M_{2\times 2}(\mathbb{C})$ satisfying $Z^2 = A$.

10 Find ρ and the function K satisfying (6.11) for the case $a_j = j,\ j = 0, 1, 2, \ldots$.

11 Prove that A is anti-Hermitian if and only if e^A is unitary.

12* For $A, B \in M_{n\times n}(\mathbb{C})$ suppose that $e^{tA}B = Be^{tA}$ for $t \in \mathbb{R}$. Prove that $AB = BA$. In this true if $\cos(tA)B = B\cos(tA)$?

Appendix A. Roots of polynomials

Lemma A.1. *A factorization of a polynomial is unique up to rearrangement of the roots.*

Proof. Suppose that for all $z \in \mathbb{C}$

$$H(z) = \prod_{i=1}^{k}(z-r_i)^{\mu_i} = \prod_{i=1}^{\ell}(z-s_i)^{\nu_i} \tag{A.1}$$

where $r_i, s_i \in \mathbb{C}$, k, ℓ, μ_i, ν_i are positive integers, and

$$r_i \neq r_j,\ s_i \neq s_j \quad \text{for all } i \neq j. \tag{A.2}$$

The zeros of H are independent of such a factorization, so the sets $\{r_1, \ldots, r_k\}$, $\{s_1, \ldots, s_\ell\}$ coincide. From (A.2) it follows that $\ell = k$ and $s_i = r_i$, $i = 1, \ldots, k$ possibly after rearranging the s_i. So (A.1) becomes

$$H(z) = \prod_{i=1}^{k}(z-r_i)^{\nu_i} = \prod_{i=1}^{k}(z-r_i)^{\nu_i}. \tag{A.3}$$

Similarly, the zeros of the z-derivative of H are independent of the specific factorization. For any $j = 1, \ldots, k$, from the first part of (A.3)

$$\frac{d^m H(r_j)}{dz^m} = 0,\ m = 0, \ldots, \mu_j - 1 \tag{A.4}$$

$$\frac{d^{\mu_j} H(r_j)}{dz^{\mu_j}} = \mu_j! \prod_{\substack{i \neq j \\ i=1}}^{k}(r_j - r_i)^{\mu_i} \neq 0 \tag{A.5}$$

From the second part of (A.3)

$$\frac{d^m H(r_j)}{dz^m} = 0,\ m = 0, \ldots, \nu_j - 1; \frac{d^{\nu_j}(r_j)}{dz^{\nu_j}} \neq 0. \tag{A.6}$$

Comparing (A.4), (A.5), (A.6), it follows that $\nu_j = \mu_j$ for all j, completing the proof.

For $a = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}, r = \begin{pmatrix} r_1 \\ \vdots \\ r_n \end{pmatrix} \in \mathbb{C}^n$, let $f : \mathbb{C}^n \to \mathbb{C}^n$ be determined by $f(r) = a$, where for all $z \in \mathbb{C}$

$$P_a(z) = z^n + a_1 z^{n-1} + \cdots + a_{n-1} z + a_n = (z - r_1)(z - r_2) \ldots (z - r_n). \tag{A.7}$$

Clearly f is a smooth mapping, a polynomial mapping in particular. But f is not one-to-one, as $f(r)$ is unchanged by the interchange of any r_i with $r_j, j \neq i$.

Lemma A.2. *For every $K > 0$*

$$|f(r)| \leq K \quad \textit{implies } |r| \leq n \quad \textit{maximum } (1, K). \tag{A.8}$$

Proof. The components of r are just the zeros of $P_a(z)$ given by (4.7) where $a = f(r)$ so from (A.8)

$$|a| \leq K. \tag{A.9}$$

Without loss of generality assume $K \geq 1$. Then for $|z| > nK$, for any $j = 1, \ldots, n$,

$$\begin{aligned} |a_j z^{n-j}| &\leq |a||z|^{n-1}, \\ &\leq K|z|^{n-1} \\ &< |z|^n/n. \end{aligned} \tag{A.10}$$

From (A.7) and (A.10) it follows from the triangle inequality that $P_a(z)$ does not vanish, proving (A.8).

Denote by

$$D = \{r \in \mathbb{C}^n | r_i \neq r_j \quad \text{for all } i \neq j\} \tag{A.11}$$

and

$$S = f(D) \tag{A.12}$$

Lemma A.3. *S is an open subset of $\mathbb{C}^n$.*

Proof. Given $a \in S$, there exists $r \in D$ such that $f(r) = a$, equivalently from (A.7)

$$P_a(z) = 0, \quad \text{for } z \in \{r_i, \dots, r_n\}. \tag{A.13}$$

From the continuity of f, for $|\tilde{a} - a|$ sufficiently small, $|P_{\tilde{a}}(z)|$ can be made arbitrarily small for z in a sufficiently small neighborhood of any of of the r_j. Furthermore, for any $j = 1, \dots, n$ fixed

$$\begin{aligned} \frac{dP_a(r_j)}{dz} &= \prod_{\substack{i=1 \\ i \neq j}}^{n} (r_j - r_i) \\ &\neq 0, \end{aligned} \tag{A.14}$$

since $r \in D$, using (A.11). So by the implicit function theorem, there is a zero $\tilde{r}_j$ of $P_{\tilde{a}}$ approaching r_j as $\tilde{a}$ approaches a, for each j. In particular, the $\tilde{r}_j$ are distinct for $|\tilde{a} - a|$ sufficiently small.

The next lemma shows that there exists a right inverse of f, defined on the closure of S.

Lemma A.4. *Suppose that $a^m \in S, m = 1, 2, \dots,$ and that $a^m \to a \in \mathbb{C}^n$ as $m \to \infty$. Then there exists $r \in \mathbb{C}^n$ such that $f(r) = a$.*

Proof. As $a^m \in S$, there exists $r^m \in D$ such that $f(r^m) = a^m$; by lemma A2, as the a^m are bounded uniformly with respect to m, so are the r^m. By the Bolzano-Weissstrauss theorem, a subsequence of the r^m converges to r as $r \to \infty$. Now $f(r) = a$ from the continuity of f.

Theorem A.5. *In the complex field, any polynomial of degree n can be factored into n linear terms.*

Proof. We have established the existence of a right inverse of f defined on the closure of S, so it suffices to show that the open set S is dense in $\mathbb{C}^n$.

Denote by

$$E = \{r \in \mathbb{C}^n | r_{n-1} = r_n\} \tag{A.15}$$

and

$$Q = f(E) \tag{A.16}$$

From lemma A1, although E is not all of the boundary of D, Q is all of the boundary of S. Therefore either S is dense in K^n or else Q partitions $\mathbb{C}^n$ into nonintersecting regions.

There are not enough points in Q to partition $\mathbb{C}^n$; it is only here that the assumption of the complex field is used. At any point $a \in Q$, denote by T_a the tangent space to Q at a. As E is isomorphic to $\mathbb{R}^{2n-2}$, and f is a smooth map, it follows that T_a is defined for all $a \in Q$ and that

$$\dim T_a \leq 2n - 2. \tag{A.17}$$

However, if Q or a subset thereof partitions $\mathbb{C}^n$, isomorphic to $\mathbb{R}^{2n}$, we need T_a of dimension $2n - 1$. So indeed S is dense in $\mathbb{C}^n$ and the theorem is proved.

Theorem A.6. *In the complex field, the roots of a polynomial of finite degree depend continuously on the coefficients of the polynomial.*

Proof. Suppose not, then there exist $r, a = f(r) \in \mathbb{C}^n$, and a sequence $\{(r^m, a^m)\}_{m=1}^{\infty}$ satisfying

$$a^m = f(r^m), m = 1, \dots \tag{A.18}$$

$$a^m \to a \quad \text{as } m \to \infty \tag{A.19}$$

but

$$|r^m - r| \geq \alpha > 0 \tag{A.20}$$

after any rearrangement of the components of each of the r^m.

By lemma A2, as the $|a^m|$ are bounded uniformly with respect to m, so are the $|r^m|$, and a subsequence converges to some $\tilde{r} \in \mathbb{C}^n$. From (A.18), (A.19) and the continuity of $f, f(\tilde{r}) = a$. Now from lemma A1, $\tilde{r} = r$, possibly after a rearrangement of the components contradicting (A.20) and thus proving the theorem.

Appendix B: Sequences of functions and power series

Let $L \subseteq \mathbb{R}$ be an interval, and let $f_n, n = 1, \dots,$ be an infinite sequence of functions from L into $\mathbb{R}$.

Definition. The sequence $\{f_n\}_{n=1}^{\infty}$ converges <u>pointwise</u> in L if for all $\varepsilon > 0$ and all $x \in L$, there exists $N_{\varepsilon,x}$ such that

$$|f_n(x) - f_m(x)| \leq \varepsilon \quad \text{for all } m, n \geq N_{\varepsilon,x} \tag{B.1}$$

Definition. The sequence $\{f_n\}_{n=1}^{\infty}$ converges <u>uniformly</u> in L if for all $\varepsilon > 0$ there exists N_ε such that

$$|f_n(x) - f_m(x)| \leq \varepsilon \quad \text{for all } m, n \geq N_\varepsilon \quad \text{and } x \in L. \tag{B.2}$$

Throughout we denote by f the limit function of such convergent sequence.

Theorem B.1. *Assume that the sequence $\{f_n\}_{n=1}^{\infty}$ converges uniformly in L. If each of the f_n is continuous at a point $x \in L$, then the limit function is continuous at x. If each of the f_n is uniformly continuous in L, then the limit is uniformly continuous in L.*

Proof. We prove only the second statement here, as the arguments are very similar. Given $\gamma > 0$, choose $\varepsilon = \gamma/3$, find N_ε from (B.2), and fix $n \geq N_\varepsilon$. Then

$$|f(x) - f_n(x)| \leq \gamma/3 \quad \text{for all } x \in L. \tag{B.3}$$

Now find δ, depending on γ and on n, such that

$$|f_n(x) - f_n(y)| \leq \gamma/3 \quad \text{for all } |x-y| \leq \delta,\ x, y \in L \tag{B.4}$$

using the fact that f_n is uniformly continuous in L. Then using (B.3) and (B.4), for any $x, y \in L,\ |x - y| \leq \delta$,

$$\begin{aligned} |f(x) - f(y)| &\leq |f(x) - f_n(x)| + |f_n(x) - f_n(y)| + |f_n(y) - f(y)| \\ &\leq \gamma. \end{aligned}$$

Uniform convergence is needed because we have to choose n without knowing x, y. A simple example where the theorem fails, given only pointwise convergence is given by $L = (-1,\ 1)$,

$$f_n(x) = \begin{cases} e^{nx}, x < 0 \\ 1, x \geq 0 \end{cases}$$

where $\{f_n\}_{n=1}^{\infty}$ converges pointwise to the (discontinuous) Heavyside function.

Theorem B.2. *Let $(a, b) \subseteq L$ be a bounded interval, and assume that the sequence $\{f_n\}_{n=1}^{\infty}$ converges uniformly in the interval (a, b). Then*

$$\int_a^b f_n(x)dx \to \int_a^b f(x)dx \quad \text{as } n \to \infty. \tag{B.5}$$

Proof. Given any $\gamma > 0$ set $\varepsilon = \gamma/(b-a)$ and fix $n \geq N_\varepsilon$ from (B.2). (We need $b - a$ finite so $\varepsilon > 0$.) Then

$$\begin{aligned} |\int_a^b f_n(x)dx - \int_a^b f(x)dx| &\leq \int_a^b |f_n(x) - f(x)dx \\ &\int_a^b \varepsilon dx \\ &= (b-a)\varepsilon \\ &= \gamma. \end{aligned}$$

The example $a = 0, b = 1, f_n(x) = ne^{-nx}$ shows that uniform convergence is needed. The sequence $\{f_n\}_{n=1}^{\infty}$ converges pointwise to the zero function on $(0,\ 1)$, but

$$\int_0^1 f_n(x)dx = 1 - e^{-n}$$

converges to 1 as $n \to \infty$.

The example $a = 0,\ b = \infty$,

$$f_n(x) = \begin{cases} 1/n, & 0 < x < n \\ 0, & x \geq n \end{cases}$$

shows that a finite interval is needed. The sequence $\{f_n\}_{n=1}^{\infty}$ converges uniformly to zero, but for all n

$$\int_0^{\infty} f_n(x)dx = 1.$$

Theorem B.3. *Assume that each of the f_n is continuously differentiable in L, that the sequence $\{f_n\}_{n=1}^{\infty}$ converges pointwise to f in L, and that the sequence $\{f_n'\}_{n=1}^{\infty}$ converges uniformly to g in L. Then f is differentiable in L and*

$$f'(x) = g(x) \quad \textit{for all } x \in L \tag{B.6}$$

Proof. For any fixed $x, y \in L$ and any n,

$$f_n(y) - f_n(x) = \int_x^y f_n'(t)dt. \tag{B.7}$$

As $n \to \infty$, the left side of (B.7) converges to $f(y) - f(x)$, whereas by theorem B.2 the right side of (B.7) converges to $\int_x^y (g(t)dt$. So for any $x, y \in L$

$$f(y) - f(x) = \int_x^y g(t)dt \tag{B.8}$$

Now from Theorem B.1, g is continuous in L, so from (B.8) f is differentiable in L and satisfies (B.6).

The application of these results needed in chapter 6 is that where the f_n are given as a power series in x

$$f_n(x) = \sum_{j=0}^{n} a_j x^j \tag{B.9}$$

where the $a_j \in \mathbb{C}$ (or $\mathbb{R}$) are given coefficients. The formal limit

$$f(x) = \sum_{j=0}^{\infty} a_j x^j \tag{B.10}$$

obviously exists for $x = 0$; indeed, we shall show it exists for $|x| < \rho$, ρ the radius of convergence of the series (B.10), given explicitly by

$$\rho = 1/\limsup_j |a_j|^{1/j}. \tag{B.11}$$

As determined from (B.11), ρ may be zero (as in the case $a_j = j!$), a finite positive number (as in the case $a_0 = 0, a_j = (-1)^{j+1}/j$, corresponding to the Taylor series for $\log(1+x)$), or infinite (as in the case $a_j = 1/j!$, corresponding to the Taylor series for e^x).

For any $|x| < \rho$, from (B.11)

$$\begin{aligned} \limsup_j |a_j x^j| &= \limsup_j |a_j| \rho^j \left(\frac{|x|}{\rho}\right)^j \\ &= 0, \end{aligned}$$

so there is a number K, depending on $|x|$, such that

$$|a_j x^j| \le K \quad \text{for all } j,$$

and in particular (6.11) holds. For $|x| > \rho$, from (B.11)

$$\limsup_j |a_j x^j| = \infty;$$

the limit may or may not exist for $x = \pm\rho$, depending on the a_j.

Theorem B.4. *Let $L \subset (-\rho, \rho)$ be any closed interval. Then f given by (B.10) exists and is an infinitely differentable function in L. Furthermore*

$$f_n \to f \tag{B.12}$$

$$f_n^{(\ell)} \to f^{(\ell)}, \ell = 1, 2 \ldots . \tag{B.13}$$

both uniformly in L as $n \to \infty$, and for any $a, b \in L$

$$\int_a^b f_n(x) \to \int_a^b f(x)dx \quad \textit{as } n \to \infty \tag{B.14}$$

Proof. It suffices to prove that for any arbitrary fixed $\ell = 0, 1, 2, \ldots$ that the sequence $\{f_n^{(\ell)}\}$ converges uniformly in L. Then (B.12) follows from theorem B.1, setting $\ell = 0$, (B.13) follows from theorem B.3, and (B.14) from theorem B.2.

With L fixed, denote by

$$\bar{x} = \underset{x \in L}{\text{maximum}} |x|; \tag{B.15}$$

clearly

$$\bar{x} < \rho. \tag{B.16}$$

Given $\varepsilon > 0$, we seek N_ε such that (B.2) holds for the sequence $\{f_n^{(\ell)}\}_{n=1}^{\infty}$. With the f_n obtained from (B.9), this becomes

$$\Big| \sum_{j=m+1}^{n} a_j \frac{j!}{(j-\ell)!} x^{j-\ell} \Big| \le \varepsilon \tag{B.17}$$

for all $m, n \ge N_\varepsilon$ and all $x \in L$.

In fact we shall prove a statement stronger than (B.17), that for N_ε sufficiently large, depending on $\bar{x}, \rho, \ell$ and the function K satisfying (6.11),

$$\sum_{j=N_\varepsilon}^{\infty} |a_j| \frac{j!}{(j-\ell)!} \bar{x}^{j-\ell} \le \varepsilon. \tag{B.18}$$

The left side of (B.18) is equal to

$$\begin{aligned} &\frac{1}{(\frac{\bar{x}+\rho}{2})^\ell} \sum_{j=N_\varepsilon}^{\infty} |a_j| \Big(\frac{\bar{x}+\rho}{2}\Big)^j \frac{j!}{(j-\ell)!} \Big(\frac{2\bar{x}}{\bar{x}+\rho}\Big)^{j-\ell} \\ &\qquad \le \frac{K\left(\frac{\bar{x}+\rho}{2}\right)}{\left(\frac{\bar{x}+\rho}{2}\right)^\ell} \sum_{j=N_\varepsilon}^{\infty} \frac{j!}{(j-\ell)!} \theta^{j-\ell} \end{aligned} \tag{B.19}$$

using (6.11) with $|x| = (\bar{x}+\rho)/2 < \rho$ from (B.16) and denoting by

$$\theta = \frac{2\bar{x}}{\bar{x}+\rho} < 1$$

also from (B.16). The right side of (B.19) is equal to

$$
\begin{aligned}
&\frac{K(\frac{\bar{x}+\rho}{2})}{(\frac{\bar{x}+\rho}{2})^{\ell}} \sum_{k=\ell}^{\infty} \frac{(k+N_\varepsilon-\ell)!}{(k+N_\varepsilon-2\ell)!}\theta^{k+N_\varepsilon-2\ell} \\
&< \frac{K(\frac{\bar{x}+\rho}{2})}{(\frac{\bar{x}+\rho}{2})^{\ell}}(N_\varepsilon-\ell)^{\ell}\theta^{N_\varepsilon-\ell} \sum_{k=\ell}^{\infty} \frac{k!}{(k-\ell)!}\theta^{k-\ell} \\
&= \frac{K(\frac{\bar{x}+\rho}{2})}{(\frac{\bar{x}+\rho}{2})^{\ell}}(N_\varepsilon-\ell)^{\ell}\theta^{N_\varepsilon-\ell} \frac{\ell!}{(1-\theta)^{\ell+1}} \qquad \text{(B.20)}
\end{aligned}
$$

using $\theta < 1$. Now as $N_\varepsilon \to \infty$ with $\ell, \theta < 1$ fixed, the function $(N_\varepsilon-\ell)^{\ell}\theta^{N_\varepsilon-\ell}$ approaches zero, so taking N_ε sufficiently large the right side of (B.20) is made smaller than the given ε.