5 Definiteness, Dual Spaces, and Advanced Spectral Theory

From Treil (2017).

Chapter 7

11/1: **4.1.** Using Silvester's Criterion of Positivity, check if the matrices

$$A = \begin{pmatrix} 4 & 2 & 1 \\ 2 & 3 & -1 \\ 1 & -1 & 2 \end{pmatrix} \qquad B = \begin{pmatrix} 3 & -1 & 2 \\ -1 & 4 & -2 \\ 2 & -2 & 2 \end{pmatrix}$$

are positive definite or not. Are the matrices $-A$, A^3, A^{-1}, $A+B^{-1}$, $A+B$, and $A-B$ positive definite?

Answer. $\underline{A}$: We have that

$$\det A_1 = \det \begin{pmatrix} 4 \end{pmatrix} = 4 \qquad \det A_2 = \det \begin{pmatrix} 4 & 2 \\ 2 & 3 \end{pmatrix} = 8 \qquad \det A_3 = \det \begin{pmatrix} 4 & 2 & 1 \\ 2 & 3 & -1 \\ 1 & -1 & 2 \end{pmatrix} = 5$$

Thus, since $A = A^*$ and $\det A_k > 0$ for $k = 1, 2, 3$, Silvester's Criterion of Positivity implies that A is positive definite.

$\underline{B}$: We have that

$$\det B_1 = \det \begin{pmatrix} 3 \end{pmatrix} = 3 \qquad \det B_2 = \det \begin{pmatrix} 3 & -1 \\ -1 & 4 \end{pmatrix} = 11 \qquad \det B_3 = \det \begin{pmatrix} 3 & -1 & 2 \\ -1 & 4 & -2 \\ 2 & -2 & 2 \end{pmatrix} = 2$$

Thus, since $B = B^*$ and $\det B_k > 0$ for $k = 1, 2, 3$, Silvester's Criterion of Positivity implies that B is positive definite.

$\underline{-A}$: We have that

$$\det(-A)_1 = \det \begin{pmatrix} -4 \end{pmatrix} = -4 \not> 0$$

Thus, Silvester's Criterion of Positivity implies that B is not positive definite.

$\underline{A^3}$: Since $A = A^*$, Theorem 6.2.2 implies that $A = UDU^*$ where U is unitary and $D = \operatorname{diag}\{\lambda_1, \lambda_2, \lambda_3\}$ with each λ_k real. Moreover, since A is positive definite, Theorem 7.4.1 implies that each $\lambda_k > 0$. Thus, since $A^3 = UD^3U^*$, A^3 is Hermitian, $D^3 = \operatorname{diag}\{\lambda_1^3, \lambda_2^3, \lambda_3^3\}$ where each λ_k^3 is an eigenvalue of A^3, and naturally each $\lambda_k^3 > 0$, Theorem 7.4.1 implies that A^3 is positive definite.

$\underline{A^{-1}}$: By a symmetric argument to the one used for A^3, we have that A^{-1} is positive definite.

$\underline{A+B^{-1}}$: Since A is positive definite, by definition, $(A\mathbf{x}, \mathbf{x}) > 0$ for all $\mathbf{x} \neq \mathbf{0}$. By a symmetric argument to the one used for A^{-1}, B^{-1} is positive definite. Thus, similarly, $(B^{-1}\mathbf{x}, \mathbf{x}) > 0$ for all $\mathbf{x} \neq \mathbf{0}$. It follows by combining the previous results that if $\mathbf{x} \neq \mathbf{0}$, then

$$0 < (A\mathbf{x}, \mathbf{x}) < (A\mathbf{x}, \mathbf{x}) + (B^{-1}\mathbf{x}, \mathbf{x}) = ((A + B^{-1})\mathbf{x}, \mathbf{x})$$

so $A + B^{-1}$ is positive definite.

$\underline{A+B}$: By a symmetric argument to the one used for $A+B^{-1}$, we have that $A+B$ is positive definite.

$\underline{A-B}$: We have that

$$\det(A-B)_2 = \det \begin{pmatrix} 1 & 3 \\ 3 & -1 \end{pmatrix} = -10 \not> 0$$

Thus, Silvester's Criterion of Positivity implies that $A - B$ is not positive definite. □

4.2. True or false:

a) If A is positive definite, then A^5 is positive definite.

Answer. True.
If A is positive definite, then $A = A^*$. It follows that $A = UDU^*$. Additionally, Theorem 7.4.1 implies that $\lambda_k > 0$ for all λ_k along the diagonal of D. Thus, $A^5 = UD^5U^*$ where D^5 has all positive diagonal entries because D has all positive diagonal entries. Thus, by Theorem 7.4.1 again, A^5 is positive definite. $\square$

b) If A is negative definite, then A^8 is negative definite.

Answer. False.
If A is negative definite, then as before, $A = UDU^*$ and $A^8 = UD^8U^*$. But if every entry along the diagonal of D is negative (Theorem 7.4.1), then every diagonal along $D^8 = (D^2)^4$ will be positive, so A^8 is not negative definite (it is, in fact, positive definite). $\square$

c) If A is negative definite, then A^{12} is positive definite.

Answer. True.
See the explanation to part (b). $\square$

d) If A is positive definite and B is negative semidefinite, then $A - B$ is positive definite.

Answer. True.
If A is positive definite, then $(A\mathbf{x}, \mathbf{x}) > 0$ for all $\mathbf{x} \neq 0$. Similarly, $(B\mathbf{x}, \mathbf{x}) \leq 0$ for all $\mathbf{x}$. To prove that $A - B$ is positive definite, it will suffice to show that $((A - B)\mathbf{x}, \mathbf{x}) > 0$ for all $\mathbf{x} \neq 0$. Let $\mathbf{x} \neq 0$ be arbitrary. Then

$$0 < (A\mathbf{x}, \mathbf{x}) \leq (A\mathbf{x}, \mathbf{x}) - (B\mathbf{x}, \mathbf{x}) = (A\mathbf{x} - B\mathbf{x}, \mathbf{x}) = ((A - B)\mathbf{x}, \mathbf{x})$$

as desired. $\square$

e) If A is indefinite, and B is positive definite, then $A + B$ is indefinite.

Answer. False.
Let

$$A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \qquad B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

By Theorem 7.4.1, A is indefinite and B is positive definite. However,

$$A + B = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}$$

which is positive semidefinite by Theorem 7.4.1. $\square$

4.3. Let A be a 2×2 Hermitian matrix such that $a_{1,1} > 0$, $\det A \geq 0$. Prove that A is positive semidefinite.

Answer. We have by the given constraints that A is of the form

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} \\ \bar{a}_{1,2} & a_{2,2} \end{pmatrix}$$

Additionally, we have that

$$\begin{aligned} 0 \leq \det A &= a_{1,1}a_{2,2} - a_{1,2}\bar{a}_{1,2} = a_{1,1}a_{2,2} - |a_{1,2}|^2 \\ |a_{1,2}|^2 &\leq a_{1,1}a_{2,2} \end{aligned}$$

from which it follows since $|a_{1,2}|^2 \geq 0$ that

$$0 \leq |a_{1,2}|^2 \leq a_{1,1}a_{2,2}$$

This combined with the fact that $a_{1,1} > 0$ implies that $a_{2,2} \geq 0$. Thus,

$$\operatorname{tr} A = a_{1,1} + a_{2,2} \geq a_{1,1} + 0 > 0$$

Now let λ_1, λ_2 be the eigenvalues of A. It follows from the above since $\operatorname{tr} A = \lambda_1 + \lambda_2$ that WLOG we may let $\lambda_1 > 0$. It follows that

$$\begin{aligned} 0 &\leq \det A = \lambda_1 \lambda_2 \\ 0 &\leq \lambda_2 \end{aligned}$$

Therefore, having shown that each $\lambda_k \geq 0$, Theorem 7.4.1 implies that A is positive semidefinite, as desired. □

4.4. Find a real symmetric $n \times n$ matrix such that $a_{1,1} > 0$ and $\det A_k \geq 0$ for all $k = 2, \dots, n$, but the matrix A is not positive semidefinite. Try to find an example for the minimal possible n[1].

Answer. Let

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

Then $a_{1,1} = 1 > 0$, $\det A_2 = 0 \geq 0$, and $\det A_3 = 0 \geq 0$. However, we have that its eigenvalues are $\lambda = -1, 0, 2$, so A is actually indefinite. Also, we know that this is the answer for the minimal possible n since Problem 7.4.3 proves that the conditions actually *do* imply A is positive semidefinite for $n = 2$. □

4.5. Let A be an $n \times n$ Hermitian matrix such that $\det A_k > 0$ for all $k = 1, \dots, n-1$ and $\det A \geq 0$. Prove that A is positive semidefinite.

Answer. Let $\lambda_1, \dots, \lambda_n$ be the eigenvalues of A, and let $\mu_1, \dots, \mu_{n-1}$ be the eigenvalues of A_{n-1}, both sets taken in decreasing order. By Silvester's Criterion of Positivity, the hypothesis that $\det A_k > 0$ for each $k = 1, \dots, n-1$ implies that A_{n-1} is positive definite. Thus, by Theorem 7.4.1, each $\mu_k > 0$. It follows by Corollary 7.4.4 that

$$\lambda_k \geq \mu_{n-1} > 0$$

for each $k = 1, \dots, n-1$. Thus,

$$\begin{aligned} 0 &\leq \det A = \lambda_1 \cdots \lambda_{n-1} \lambda_n \\ 0 &\leq \lambda_n \end{aligned}$$

Therefore, since each $\lambda_k \geq 0$, Theorem 7.4.1 implies that A is positive semidefinite, as desired. □

4.6. Find a real symmetric 3×3 matrix A such that $a_{1,1} > 0$, $\det A_k \geq 0$ for $k = 2, 3$, but the matrix A is not positive semidefinite.

Answer. Using the same matrix from Problem 7.4.4, we have

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

□

[1] The statement of this problem has been modified as per Chloé's instructions in the 10/28 problem session.

Chapter 8

1.1. Let $\mathbf{v}_1, \dots, \mathbf{v}_r$ be a system of vectors in X such that there exists a system $\mathbf{v}_1', \dots, \mathbf{v}_r'$ of linear functionals such that

$$\mathbf{v}_k'(\mathbf{v}_j) = \delta_{jk}$$

a) Show that the system $\mathbf{v}_1, \dots, \mathbf{v}_r$ is linearly independent.

Answer. To prove that $\mathbf{v}_1, \dots, \mathbf{v}_r$ is linearly independent, it will suffice to show that if $\alpha_1, \dots, \alpha_r \in \mathbb{F}$ make $\alpha_1\mathbf{v}_1 + \cdots + \alpha_r\mathbf{v}_r = 0$, then $\alpha_1 = \cdots = \alpha_r = 0$. Suppose that $\alpha_1, \dots, \alpha_r \in \mathbb{F}$ make

$$\alpha_1\mathbf{v}_1 + \cdots + \alpha_r\mathbf{v}_r = 0$$

It follows by linearity and the definition of the dual basis that

$$\begin{aligned} 0 &= \mathbf{v}_k'(\mathbf{0}) \\ &= \mathbf{v}_k'(\alpha_1\mathbf{v}_1 + \cdots + \alpha_r\mathbf{v}_r) \\ &= \alpha_1\mathbf{v}_k'(\mathbf{v}_1) + \cdots + \alpha_r\mathbf{v}_k'(\mathbf{v}_r) \\ &= \alpha_1 \cdot 0 + \cdots + \alpha_{k-1} \cdot 0 + \alpha_k \cdot 1 + \alpha_{k+1} \cdot 0 + \cdots + \alpha_r \cdot 0 \\ &= \alpha_k \end{aligned}$$

for each $k = 1, \dots, r$, as desired. □

b) Show that if the system $\mathbf{v}_1, \dots, \mathbf{v}_r$ is not generating, then the "biorthogonal" system $\mathbf{v}_1', \dots, \mathbf{v}_r'$ is not unique. (Hint: Probably the easiest way to prove that is to complete the system $\mathbf{v}_1, \dots, \mathbf{v}_r$ to a basis [see Proposition 2.5.4].)

Answer. By Proposition 2.5.4, we can complete the linearly independent list $\mathbf{v}_1, \dots, \mathbf{v}_r$ to a basis $\mathbf{v}_1, \dots, \mathbf{v}_n$ where $n > r$ since $\mathbf{v}_1, \dots, \mathbf{v}_r$ is not generating by hypothesis. Consider $\mathbf{v}_1', \dots, \mathbf{v}_r'$. These linear functionals' behavior on $\mathbf{v}_1, \dots, \mathbf{v}_r$ is completely defined by the given condition; however, since they act on all of X and not just $\operatorname{span}\{\mathbf{v}_1, \dots, \mathbf{v}_r\} \subsetneq X$, we can define an arbitrary linear behavior for each $\mathbf{v}_k'$ on $\operatorname{span}\{\mathbf{v}_{r+1}, \dots, \mathbf{v}_n\}$. Clearly, more than one such behavior exists (take, for example, being the zero map on that span and being the identity map on that span), so $\mathbf{v}_1', \dots, \mathbf{v}_n'$ is not unique. □

3.1. Prove that if for linear transformations $T, T_1 : X \to Y$

$$\langle T\mathbf{x}, \mathbf{y}' \rangle = \langle T_1\mathbf{x}, \mathbf{y}' \rangle$$

for all $\mathbf{x} \in X$ and for all $\mathbf{y}' \in Y'$, then $T = T_1$. (Hint: Probably one of the easiest ways of proving this is to use Lemma 8.1.3.)

Answer. Let $\mathbf{x} \in X$ be arbitrary. If $\langle T\mathbf{x}, \mathbf{y}' \rangle = \langle T_1\mathbf{x}, \mathbf{y}' \rangle$ for all $\mathbf{y}' \in Y'$, then $\mathbf{y}'(T\mathbf{x}) = \mathbf{y}'(T_1\mathbf{x})$ for all $\mathbf{x} \in X$ and for all $\mathbf{y}' \in Y'$. Thus, since every linear functional in the dual space maps the vectors $T\mathbf{x}$ and $T_1\mathbf{x}$ the same way, Lemma 8.1.3 implies that $T\mathbf{x} = T_1\mathbf{x}$. But since we let $\mathbf{x}$ be arbitrary, $T\mathbf{x} = T_1\mathbf{x}$ for all $\mathbf{x} \in X$, i.e., $T = T_1$. □

3.2. Combine the Riesz Representation Theorem (Theorem 8.2.1) with the reasoning in Section 3.1.3 above to present a coordinate-free definition of the Hermitian adjoint of an operator in an inner product space.

Answer. Let $A \in \mathcal{L}(V, W)$. We seek to define A^* as the unique element of $\mathcal{L}(W, V)$ satisfying

$$(A\mathbf{x}, \mathbf{y}) = (\mathbf{x}, A^*\mathbf{y})$$

for all $\mathbf{x} \in V$ and $\mathbf{y} \in W$. Let's begin.

Let $\mathbf{y}$ be an arbitrary element of W. We can think of $\mathbf{y}^*$ as a $1 \times \dim W$ matrix, or indeed a linear transformation $\mathbf{y}^* : W \to \mathbb{F}$. This combined with the fact that $A : V \to W$ implies that $\mathbf{y}^* A : V \to \mathbb{F}$ is a well-defined linear functional. It follows by the Riesz Representation Theorem that there exists a unique $\mathbf{z} \in V$ such that $(\mathbf{y}^* A)(\mathbf{x}) = (\mathbf{x}, \mathbf{z})$ for all $\mathbf{x} \in V$. Define $A^*\mathbf{y} := \mathbf{z}$.

Since $\mathbf{z}$ is unique by the Riesz Representation Theorem, A^* is a well-defined function for this $\mathbf{y}$. Moreover, since we let $\mathbf{y} \in W$ be arbitrary, we can define $A^*\mathbf{y}$ in the same way for *any* $\mathbf{y} \in W$. Thus, $A^* : W \to Z$ (as defined) is a well-defined function on W.

We now seek to prove that A^* is linear. Let $\mathbf{y}_1, \mathbf{y}_2 \in W$ and $\alpha_1, \alpha_2 \in \mathbb{F}$. We know that $A^*\mathbf{y}_1$ is the unique vector $\mathbf{z}_1 \in V$ such that $(\mathbf{y}_1^* A)(\mathbf{x}) = (\mathbf{x}, \mathbf{z}_1)$ for all $\mathbf{x} \in V$. We also know that $A^*\mathbf{y}_2$ is the unique vector $\mathbf{z}_2 \in V$ such that $(\mathbf{y}_2^* A)(\mathbf{x}) = (\mathbf{x}, \mathbf{z}_2)$ for all $\mathbf{x} \in V$. Lastly, we know that $A^*(\alpha_1\mathbf{y}_1 + \alpha_2\mathbf{y}_2)$ is the unique vector $\mathbf{z} \in V$ such that $[(\alpha_1\mathbf{y}_1 + \alpha_2\mathbf{y}_2)^* A](\mathbf{x}) = (\mathbf{x}, \mathbf{z})$ for all $\mathbf{x} \in V$. It follows that

$$\begin{aligned}
(\mathbf{x}, A^*(\alpha_1\mathbf{y}_1 + \alpha_2\mathbf{y}_2)) &= (\mathbf{x}, \mathbf{z}) \\
&= [(\alpha_1\mathbf{y}_1 + \alpha_2\mathbf{y}_2)^* A](\mathbf{x}) \\
&= \bar{\alpha}_1(\mathbf{y}_1^* A)(\mathbf{x}) + \bar{\alpha}_2(\mathbf{y}_2^* A)(\mathbf{x}) \\
&= \bar{\alpha}_1(\mathbf{x}, \mathbf{z}_1) + \bar{\alpha}_2(\mathbf{x}, \mathbf{z}_2) \\
&= (\mathbf{x}, \alpha_1\mathbf{z}_1 + \alpha_2\mathbf{z}_2) \\
&= (\mathbf{x}, \alpha_1 A^*\mathbf{y}_1 + \alpha_2 A^*\mathbf{y}_2)
\end{aligned}$$

for all $\mathbf{x} \in V$. Thus, by Lemma 8.1.3

$$A^*(\alpha_1\mathbf{y}_1 + \alpha_2\mathbf{y}_2) = \alpha_1 A^*\mathbf{y}_1 + \alpha_2 A^*\mathbf{y}_2$$

as desired.

We now show that A^* satisfies the desired identity: If $\mathbf{x} \in V$ and $\mathbf{y} \in W$, then we have by the definition of A^* that

$$(\mathbf{x}, A^*\mathbf{y}) = (\mathbf{y}^* A)(\mathbf{x}) = \mathbf{y}^*(A\mathbf{x}) = (A\mathbf{x}, \mathbf{y})$$

as desired.

Lastly, we prove that A^* is the unique linear map satisfying the above identity. Suppose $A^*, \tilde{A}^*$ are linear maps such that

$$(A\mathbf{x}, \mathbf{y}) = (\mathbf{x}, A^*\mathbf{y}) \qquad\qquad (A\mathbf{x}, \mathbf{y}) = (\mathbf{x}, \tilde{A}^*\mathbf{y})$$

for all $\mathbf{x} \in V$ and $\mathbf{y} \in W$. Let $\mathbf{y} \in W$ be arbitrary. Then

$$(\mathbf{x}, A^*\mathbf{y}) = (A\mathbf{x}, \mathbf{y}) = (\mathbf{x}, \tilde{A}^*\mathbf{y})$$

for all $\mathbf{x} \in V$. It follows by Lemma 8.1.3 that $A^*\mathbf{y} = \tilde{A}^*\mathbf{y}$. Furthermore, since we let $\mathbf{y}$ be arbitrary, we know that $A^*\mathbf{y} = \tilde{A}^*\mathbf{y}$ for *every* $\mathbf{y} \in W$. Therefore, $A^* = \tilde{A}^*$, so A^* is unique, as desired. ☐

3.3. Let $\mathbf{v}_1, \dots, \mathbf{v}_n$ be a basis in X and let $\mathbf{v}'_1, \dots, \mathbf{v}'_n$ be its dual basis. Let $E = \operatorname{span}\{\mathbf{v}_1, \dots, \mathbf{v}_r\}$ for $r < n$. Prove that $E^\perp = \operatorname{span}\{\mathbf{v}'_{r+1}, \dots, \mathbf{v}'_n\}$. (This problem gives a way to prove Proposition 8.3.6.)

Answer. Suppose first that $\mathbf{v}' \in E^\perp$. Then by the definition of the annihilator, $\mathbf{v}' \in X'$ and $\langle \mathbf{x}, \mathbf{v}' \rangle = 0$ for all $\mathbf{x} \in E$. It follows from the first condition that

$$\mathbf{v}' = \alpha_1\mathbf{v}'_1 + \cdots + \alpha_n\mathbf{v}'_n$$

for some $\alpha_1, \dots, \alpha_n \in \mathbb{F}$. It follows from the second condition that

$$\begin{aligned}
0 &= \langle \mathbf{v}_k, \mathbf{v}' \rangle \\
&= \alpha_1\mathbf{v}'_1(\mathbf{v}_k) + \cdots + \alpha_n\mathbf{v}'_n(\mathbf{v}_k) \\
&= \alpha_k
\end{aligned}$$

for each $k = 1, \dots, r$. Therefore,

$$\mathbf{v}' = \alpha_{r+1}\mathbf{v}'_{r+1} + \cdots + \alpha_n\mathbf{v}'_n$$

so $\mathbf{v} \in \operatorname{span}\{\mathbf{v}'_{r+1}, \dots, \mathbf{v}'_n\}$, as desired.

Now suppose that $\mathbf{v}' \in \operatorname{span}\{\mathbf{v}'_{r+1}, \dots, \mathbf{v}'_n\}$. In particular, let $\mathbf{v}' = \alpha_{r+1}\mathbf{v}'_{r+1} + \cdots + \alpha_n\mathbf{v}'_n$ for some $\alpha_{r+1}, \dots, \alpha_n \in \mathbb{F}$. To prove that $\mathbf{v}' \in E^\perp$, it will suffice to show that $\langle \mathbf{x}, \mathbf{v}' \rangle = 0$ for all $\mathbf{x} \in E$. Let $\mathbf{x}$ be an arbitrary element of E. Then by the definition of E, $\mathbf{x} = \beta_1\mathbf{v}_1 + \cdots + \beta_r\mathbf{v}_r$. It follows by the definition of $\mathbf{v}'$ and the dual basis that

$$\begin{aligned}
\langle \mathbf{x}, \mathbf{v}' \rangle &= \alpha_{r+1}\mathbf{v}'_{r+1}(\beta_1\mathbf{v}_1 + \cdots + \beta_r\mathbf{v}_r) + \cdots + \alpha_n\mathbf{v}'_n(\beta_1\mathbf{v}_1 + \cdots + \beta_r\mathbf{v}_r) \\
&= \alpha_{r+1} \cdot 0 + \cdots + \alpha_n \cdot 0 \\
&= 0
\end{aligned}$$

as desired. $\square$

Chapter 9

1.1. (Cayley-Hamilton Theorem for diagonalizable matrices). As discussed in Section 9.1, the Cayley-Hamilton theorem states that if A is a square matrix and

$$p(\lambda) = \det(A - \lambda I) = \sum_{k=0}^{n} c_k\lambda^k$$

is its characteristic polynomial, then $p(A) = \sum_{k=0}^{n} c_kA^k = \mathbf{0}$ (assuming that by definition, $A^0 = I$). Prove this theorem for the special case when A is similar to a diagonal matrix, i.e., $A = SDS^{-1}$. (Hint: If $D = \operatorname{diag}\{\lambda_1, \dots, \lambda_n\}$ and p is any polynomial, can you compute $p(D)$? What about $p(A)$?)

Answer. Suppose $A = SDS^{-1}$, and let $\lambda_1, \dots, \lambda_n$ be the eigenvalues of A. Since $D = \operatorname{diag}\{\lambda_1, \dots, \lambda_n\}$, we have by the properties of diagonal matrix exponentiation, scalar multiplication, and addition, and Exercise 4.1.10 that

$$\begin{aligned}
p(D) &= \sum_{k=0}^{n} c_kD^k \\
&= \sum_{k=0}^{n} c_k \operatorname{diag}\{\lambda_1^k, \dots, \lambda_n^k\} \\
&= \sum_{k=0}^{n} \operatorname{diag}\{c_k\lambda_1^k, \dots, c_k\lambda_n^k\} \\
&= \operatorname{diag}\left\{\sum_{k=0}^{n} c_k\lambda_1^k, \dots, \sum_{k=0}^{n} c_k\lambda_n^k\right\} \\
&= \operatorname{diag}\{p(\lambda_1), \dots, p(\lambda_n)\} \\
&= \operatorname{diag}\{0, \dots, 0\} \\
&= 0
\end{aligned}$$

It follows that

$$\begin{aligned}
p(A) &= p(SDS^{-1}) \\
&= \sum_{k=0}^{n} c_k (SDS^{-1})^k \\
&= \sum_{k=0}^{n} c_k S D^k S^{-1} \\
&= S\left[\sum_{k=0}^{n} c_k D^k\right] S^{-1} \\
&= S[p(D)]S^{-1} \\
&= S0S^{-1} \\
&= 0
\end{aligned}$$

as desired. $\square$

2.1. An operator A is called **nilpotent** if $A^k = \mathbf{0}$ for some k. Prove that if A is nilpotent, then $\sigma(A) = \{0\}$ (i.e., that 0 is the only eigenvalue of A). Can you do it without using the spectral mapping theorem?

Answer. Suppose for the sake of contradiction that $\lambda \neq 0$ for some eigenvalue λ of A. Then if $\mathbf{v}$ is a nonzero eigenvector corresponding to λ, $A\mathbf{v} = \lambda\mathbf{v}$ so $A^k\mathbf{v} = \lambda^k\mathbf{v}$. But since $\lambda^k\mathbf{v} \neq \mathbf{0}$, $A^k \neq 0$, a contradiction. $\square$