ROB 501 Exam-I Solutions

30 October 2018

Problem 1:

(a) False. Ordinary Induction and Strong Induction are logically equivalent.

(b) True. We prove the contrapositive and assume that n is even. Hence, there exists a natural number k such that $n = 2k$, and thus $n = 4k^2 = 2 \times (2k^2)$. Because $2k^2$ is a natural number, we have that n^2 is even. Hence, n not odd implies n^2 not odd.

(c) True. $p \implies q \Leftrightarrow \neg (p \wedge (\neg q)) \Leftrightarrow (\neg p) \vee q$

(d) False. We know the truth table for $a \implies b$ is

a	b	$a \implies b$
1	1	1
1	0	0
0	1	1
0	0	1

Hence, substituting $a = \neg q$ and $b = p$ we have

p	q	$\neg q$	$\neg q \implies p$
1	1	0	1
1	0	1	1
0	1	0	1
0	0	1	0

Problem 2:

(a) False. $\text{trace}(AA^\top) = \sum_{i=1}^{n} \sum_{j=1}^{m} \left([A]_{ij}\right)^2$

(b) True. The ij-element of $A^\top B$ equals the i-th row of $A^\top$ times the j-th column of B. Hence, $[A^\top B]_{ij} = (A_i)^\top B_j$.

(c) True. For $x = [x_1, x_2, \ldots, x_m]^\top$, we have that $Ax = A_1 x_1 + A_2 x_2 + \cdots + A_m x_m$ Because the coefficients x_i are arbitrary, we generate the span of the columns of A.

(d) False. $Ax = A_1 x_1 + A_2 x_2 + \cdots + A_n x_n = x_1 A_1 + x_2 A_2 + \cdots + x_n A_n$

Remark: Yes, parts (c) and (d) were different ways of asking the same question: do you recognize that

$$Ax = A_1 x_1 + A_2 x_2 + \cdots + A_m x_m = \text{ linear combination of columns of } A \text{ ?}$$

If we looked at $x^\top A$, we'd have a linear combination of the rows of A.

Problem 3:

(a) False. Let $\mathcal{X} = \mathbb{R}$, $S_1 = \{1, 2\}$ and $S_2 = \{1\}$. Then $S_1 \not\subset S_2$ and yet $\text{span}\{S_1\} = \text{span}\{S_2\} = \mathbb{R}$.

(b) True. From lecture, for every subspace $M \subset \mathcal{X}$, we have that $M \oplus M^\perp = \mathcal{X}$. From one of the assigned exercises (or by a very quick calculation), we have that $S_1^\perp = (\text{span}\{S_1\})^\perp$. Hence,

$$\mathcal{X} = \text{span}\{S_1\} \oplus (\text{span}\{S_1\})^\perp = \text{span}\{S_1\} \oplus S_1^\perp$$

(c) True. From HW.

(d) False. Let $S_1 = \mathcal{X}$ and $S_2 = \{0\}$ (the zero subspace). Then, $S_1^\perp \cap S_2^\perp = \{0\}$ while

$$\left[\text{span}\{S_1\} \cap \text{span}\{S_2\}\right]^\perp = \left[S_1 \cap S_2\right]^\perp = \left[\{0\}\right]^\perp = \mathcal{X}$$

Problem 4:

(a) False. Recall the example at the end of the Sept. 27 lecture:
Let $\mathcal{X} = \mathbb{R}^2$, $M = \text{span}\left\{\begin{bmatrix}-1\\1\end{bmatrix}\right\}$, $x = \begin{bmatrix}1\\1\end{bmatrix}$, and let $||\bullet||$ be the 1-norm: $\left|\left|\begin{bmatrix}x_1\\x_2\end{bmatrix}\right|\right| = |x_1| + |x_2|$.
Now let $m^1 = \begin{bmatrix}0\\0\end{bmatrix} \in M$, and $m^2 = \begin{bmatrix}1\\-1\end{bmatrix} \in M$. Then $||x - m^1|| = 2$ and $||x - m^2|| = 2$.
Next, we see that $\forall m = \begin{bmatrix}\alpha\\-\alpha\end{bmatrix} \in M : ||x - m|| = |1 - \alpha| + |1 + \alpha| \geq |1 - \alpha + 1 + \alpha| = 2$.
Therefore, both m^1 and m^2 satisfy: $||x - m^i|| \leq ||x - m||, \forall m \in M$, which contradicts the problem statement.

(b) True. For $||\bullet||$ to satisfy the norm axioms, we need $||x||^2 > 0, \forall x \neq 0$. This implies that $\forall x \neq 0, x^\top P x > 0$, i.e. P is positive definite. Since positive definite matrices have all positive eigenvalues, the statement must be true.

(c) False. We use the Schur complement theorem to show this.

Let $M = \begin{bmatrix}A & B\\B^\top & C\end{bmatrix}$, where $A = 6$, $B = \begin{bmatrix}2 & 4\end{bmatrix}$, and $C = \begin{bmatrix}3 & 1\\1 & 1\end{bmatrix}$. Then $C^{-1} = \frac{1}{2}\begin{bmatrix}1 & -1\\-1 & 3\end{bmatrix}$, and:

$$\begin{aligned}
A - B^\top C^{-1} B &= 6 - \frac{1}{2}\begin{bmatrix}2 & 4\end{bmatrix}\begin{bmatrix}3 & 1\\1 & 1\end{bmatrix}\begin{bmatrix}2\\4\end{bmatrix}\\
&= 6 - \begin{bmatrix}1 & 2\end{bmatrix}\begin{bmatrix}10\\6\end{bmatrix} = 6 - 22 = -16\\
&< 0
\end{aligned}$$

Because the Schur complement of C in M is not positive definite (we computed that it is a negative number), M is not positive definite.

(d) True. Since $\{v^1, \ldots, v^n\}$ is a basis, we know that $\text{span}\{v^1, \ldots, v^n\} = \mathbb{R}$.

Therefore $x \in \text{span}\{v^1, \ldots, v^n\}$, i.e. $x = \alpha^1 v^1 + \cdots + \alpha^n v^n$, for some $\alpha^i \in \mathbb{R}$. Substituting this into the right side of $< x, x >$, we get:

$$\begin{aligned} < x, x > &= < x, \alpha^1 v^1 + \cdots + \alpha^n v^n > \\ &= \alpha^1 < x, v^1 > + \cdots + \alpha^n < x, v^n > \\ < x, x > &= 0 \end{aligned}$$

Therefore $x = 0$ by the properties of the inner product.

Problem 5:

(a) False. The i-th column of the change of basis matrix is $P_i = \left[u^i\right]_{\{\bar{u}\}} = \frac{1}{i} e^i$.

(b) False. Let $A_3 = \begin{bmatrix} a_{31} \\ a_{32} \\ \vdots \\ a_{3n} \end{bmatrix}$. Then, we have that

$$L(u^3) = a_{31} u^1 + a_{32} u^2 + \ldots + a_{3n} u^n = a_{31} \bar{u}^1 + \frac{1}{2} a_{32} \bar{u}^2 + \ldots \frac{1}{n} a_{3n} \bar{u}^n$$

where the first equality is from the definition of a matrix representation and the second equality uses the fact that $u^i = \frac{1}{i} \bar{u}^i$. Hence,

$$[L(u^3)]_{\{\bar{u}\}} = \left[a_{31} \bar{u}^1 + \frac{1}{2} a_{32} \bar{u}^2 + \ldots \frac{1}{n} a_{3n} \bar{u}^n\right]_{\{\bar{u}\}} = \begin{bmatrix} a_{31} \\ \frac{1}{2} a_{32} \\ \vdots \\ \frac{1}{n} a_{3n} \end{bmatrix} \neq \frac{1}{3} A_3$$

(c) True. $[L(\bar{u}^3 + \bar{u}^4)]_{\{u\}} = [L(3u^3 + 4u^4)]_{\{u\}} = 3[L(u^3)]_{\{u\}} + 4[L(u^4)]_{\{u\}} = 3A_3 + 4A_4$

(d) True. Because u^1 and u^2 are linearly independent, for all $\alpha \neq 0$, $u^1 \neq \alpha u^2$ and $u^2 \neq \alpha u^1$. Hence, by the definition of a strict norm, $||u^1 + u^2|| \neq ||u^1|| + ||u^2||$. The triangle inequality gives $||u^1 + u^2|| \leq ||u^1|| + ||u^2||$, thus, we deduce that

$$||u^1 + u^2|| < ||u^1|| + ||u^2||$$

must hold.

Problem 6: (a) We use the Gram-Schmidt process to construct an orthogonal basis.

Start by computing the inner products $< u_i, u_j >$ for later use. We can use the fact that the integral of an odd function from -1 to 1 is 0 to immediately get:

$$< 1, t > \; = \; < 1, \sin(\pi t) > \; = \; < t, t^2 > \; = \; < t^2, \sin(\pi t) >= 0$$

From the hint we have:

$$< t, \sin(\pi t) >= \frac{2}{\pi}, \qquad < \sin(\pi t), \sin(\pi t) >= 1$$

Then we can evaluate the integrals of polynomials to get:

$$< 1, 1 >= \int_{-1}^{1} 1 dt = 2, \qquad < 1, t^2 > \; = \; < t, t >= \int_{-1}^{1} t^2 dt = \frac{2}{3}$$

We can now construct the basis:

$$\begin{aligned}
u_1 &= 1 && = 1 \\
u_2 &= t - \frac{< 1, t >}{< 1, 1 >} 1 && = t \\
u_3 &= t^2 - \frac{< 1, t^2 >}{< 1, 1 >} 1 - \frac{< t, t^2 >}{< t, t >} t && = t^2 - \frac{1}{3} \\
u_4 &= \sin(\pi t) - \frac{< 1, \sin(\pi t) >}{< 1, 1 >} 1 - \frac{< t, \sin(\pi t) >}{< t, t >} t - \frac{< t^2 - \frac{1}{3}, \sin(\pi t) >}{< t^2 - \frac{1}{3}, t^2 - \frac{1}{3} >} t^2 && = \sin(\pi t) - \frac{3}{\pi} t
\end{aligned}$$

Where we used the following result:

$$< t^2 - \frac{1}{3}, \sin(\pi t) >=< t^2, \sin(\pi t) > + \frac{1}{3} < 1, \sin(\pi) >= 0$$

Grading Notes: None to give. Most of you nailed this problem.

(b) We find the nearest point to x in M ($m^* = \arg \min_{m \in M} ||x - m||$) using the Projection Theorem.

If we let $m^* = \alpha_1 t + \alpha_2 t^2$, the normal equations for this problem are given by:

$$\begin{aligned}
\begin{bmatrix} < t, t > & < t, t^2 > \\ < t, t^2 > & < t^2, t^2 > \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} &= \begin{bmatrix} < \sin(\pi t), t > \\ < \sin(\pi t), t^2 > \end{bmatrix} \\
\begin{bmatrix} \frac{2}{3} & 0 \\ 0 & \frac{2}{5} \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} &= \begin{bmatrix} \frac{2}{\pi} \\ 0 \end{bmatrix} \\
\begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} &= \begin{bmatrix} \frac{3}{\pi} \\ 0 \end{bmatrix}
\end{aligned}$$

And the nearest point is $m^* = \frac{3}{\pi}t$. The distance is then:

$$\begin{aligned}
d_0 &= ||\sin(\pi t) - \frac{3}{\pi}t|| \\
&= \left[\int_{-1}^{1} \left(\sin(\pi t) - \frac{3}{\pi}t\right)^2 dt\right]^{\frac{1}{2}} = \left[\int_{-1}^{1} \sin(\pi t)^2 - \frac{6}{\pi}t\sin(\pi t) + \frac{9}{\pi^2}t^2 dt\right]^{\frac{1}{2}} = \left[1 - \frac{12}{\pi^2} + \frac{6}{\pi^2}\right]^{\frac{1}{2}} \\
&= \sqrt{1 - \frac{6}{\pi^2}}
\end{aligned}$$

You could also write this as

$$\begin{aligned}
d_0^2 &= < \sin(\pi t) - \frac{3}{\pi}t,\ \sin(\pi t) - \frac{3}{\pi}t > \\
&= < \sin(\pi t),\ \sin(\pi t) > - \frac{6}{\pi} < \sin(\pi t),\ t > + \left(\frac{3}{\pi}\right)^2 < t,\ t > \\
&= 1 - \frac{12}{\pi^2} + \frac{6}{\pi^2} \\
&= 1 - \frac{6}{\pi^2}
\end{aligned}$$

This looks less scary because there are no integral signs!

Grading Notes

- If you applied the Projection Theorem or the Normal Equations and computed

  $$m^* = \text{argmin}_{x \in M} d(x, M)$$

  correctly, you earned 4 points.

- If you reported $m^*(t)$ as the answer, **even when the hint said the answer is a number and not a function of time**, there was not much more to give you.

- If you used $m^*(t)$ to compute a number, I tired to give points. For example, if you knew the distance was the norm of the approximation error, $||x - m^*||$, but messed up the calculation, you got 2 more points. If you made a major conceptual error on distance: such as reporting the α coeff for t as a distance or the inner product of x and $m^*(t)$ as the distance, or the norm of $m^*(t)$ as the distance, but, **at least it was a number**, you earned 1 point!

Problem 7:
The set of linear combinations of $\{u_b, u_r\}$ is a subspace of $\mathcal{X}$, so we use the normal equations to find α^*:

$$\begin{bmatrix} < u_b, u_b > & < u_b, u_r > \\ < u_b, u_r > & < u_r, u_r > \end{bmatrix} \begin{bmatrix} \alpha_b^* \\ \alpha_r^* \end{bmatrix} = \begin{bmatrix} < x, u_b > \\ < x, u_r > \end{bmatrix}$$

$$\begin{bmatrix} 1 & 1/2 \\ 1/2 & 1 \end{bmatrix} \begin{bmatrix} \alpha_b^* \\ \alpha_r^* \end{bmatrix} = \begin{bmatrix} 5/2 \\ 3/2 \end{bmatrix}$$

$$\begin{bmatrix} \alpha_b^* \\ \alpha_r^* \end{bmatrix} = \frac{4}{3} \begin{bmatrix} 1 & -1/2 \\ -1/2 & 1 \end{bmatrix} \begin{bmatrix} 5/2 \\ 3/2 \end{bmatrix}$$

$$\begin{bmatrix} \alpha_b^* \\ \alpha_r^* \end{bmatrix} = \begin{bmatrix} 7/3 \\ 1/3 \end{bmatrix}$$

(a) The answer is unique, since the Gram matrix: $\begin{bmatrix} < u_b, u_b > & < u_b, u_r > \\ < u_b, u_r > & < u_r, u_r > \end{bmatrix}$ is full rank and invertible.

Note 1: the projection theorem states only that the resultant color is unique, not that the specific linear combination is unique. That is, the projection theorem tells us only that $m^* = \alpha_b^* u_b + \alpha_r^* u_r$ is unique, but doesn't rule out the possibility of a different $\alpha^{**} = \begin{bmatrix} \alpha_b^{**} \\ \alpha_r^{**} \end{bmatrix}$ that satisfies $m^* = \alpha_b^{**} u_b + \alpha_r^{**} u_r$.
If u_b and u_r are linearly *dependent*, then another linear combination that minimizes the distance could be given by:

$$\alpha^{**} = \begin{bmatrix} \alpha_b^* \\ \alpha_r^* \end{bmatrix} + \begin{bmatrix} \alpha_b^0 \\ \alpha_r^0 \end{bmatrix}$$

Where α_b^0 and α_r^0 satisfy: $\alpha_b^0 u_b + \alpha_r^0 u_r = 0$. If u_b and u_r are linearly *independent*, then α_b^0 and α_r^0 must be 0, and the combination is unique.

Note 2: The Gram matrix being invertible implies linear independence of the basis vectors.
This can be seen as follows:

$$\begin{aligned} \alpha_b^0 u_b + \alpha_r^0 u_r = 0 \implies & \begin{matrix} < u_b, \alpha_b^0 u_b + \alpha_r^0 u_r >= 0 \\ < u_b, \alpha_b^0 u_b + \alpha_r^0 u_r >= 0 \end{matrix} \\ \implies & \begin{bmatrix} < u_b, u_b > & < u_b, u_r > \\ < u_b, u_r > & < u_r, u_r > \end{bmatrix} \begin{bmatrix} \alpha_b^0 \\ \alpha_r^0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \\ \implies & \begin{bmatrix} \alpha_b^0 \\ \alpha_r^0 \end{bmatrix} = \begin{bmatrix} < u_b, u_b > & < u_b, u_r > \\ < u_b, u_r > & < u_r, u_r > \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \end{aligned}$$

Where we assume an invertible Gram matrix in the last step.

Grading Notes

- If you used the invertibility of the Gram matrix to justify uniqueness you earned full points.
- If you stated that linear independence of the basis vectors leads to uniqueness, and saw from the figure that the basis vectors are linearly independent, you earned full points.
- If you stated that uniqueness follows if the basis vectors are linearly independent, but don't show linear independence you lost one point.

- If you use the projection theorem but don't reason about uniqueness of the representation, you lost one point.
- If you justified your answer, but your reasoning was unclear, you earned two points.
- If you didn't justify your answer, you earned no points.

(b) $\alpha^* = \begin{bmatrix} 7/3 \\ 1/3 \end{bmatrix}$

Grading Notes

- If you made calculation errors, you lost one point
- If you made minor errors in your method, such as using the wrong G matrix, you lost one point.
- If your method was mostly on the right track, but you missed some parts, you lost two points.
- If your method was mostly unclear, you lost eight points.

Problem 8:

- P: each of the sets $\{v^1, v^2\}$, $\{v^2, v^3\}$, and $\{v^3, v^1\}$ is linearly independent.
- Q: the set $\{v^1, v^2, v^3\}$ is linearly independent.

(a) $P \implies Q$ is **F** and here is why. Consider the two-dimensional vector space $(\mathbb{R}^2, \mathbb{R})$ and let $v^1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $v^2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, and $v^3 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$. Then each of the sets $\{v^1, v^2\}$, $\{v^2, v^3\}$, and $\{v^3, v^1\}$ is linearly independent, but because the vector space $(\mathbb{R}^2, \mathbb{R})$ is two-dimensional, any set with three vectors has to be linearly dependent.

Grading Notes

- If you answered **F** and gave a counter example, you got 7 points.
- If you answered **F** and sketched out a plausible argument about why it should be false, but did not make it concrete with a specific choice of vectors, you probably got 3 or 4 points, depending on just how plausible your argument was.
- If you answered **T** and applied the definition of linear independence or dependence in some clear form, then you got 3 points.
- If you answered **T** and attempted an argument like this: a general set is linear independent if and only if all finite subsets are linearly independent, you earned 2.0 points. The big problem here is that you forgot that any set is a subset of itself. Hence, if the set is finite, as in our case, the general definition is vacuous. You have to go back to the definition of a finite collection of vectors being linearly independent, which those attempting this approach did not do.

(b) $Q \implies P$ is **T** and here is why. **Claim** Any nonempty subset of a linearly independent set is linearly independent. ■

Proof Let S be a linearly independent set and $\emptyset \neq T \subset S$. We show the contrapositive: If T is linearly dependent, then there exists a linear combination of elements of T for which at least one of the coefficients is non-zero and yet the sum is zero,

$$\alpha_1 v^1 + \alpha_2 v^2 + \cdots + \alpha_k v^k = 0.$$

Because each of the $v^i \in T \subset S$, this proves that S is not linearly independent, showing the result. ■

Because each of the sets in P is a subset of the linearly independent set in Q, the result is proven. ■

Alternative Solution In lecture, an arbitrary set was **defined** to be linearly independent if every finite subset is linearly independent. In a definition, the if is really an if, and only if. Hence, $Q \implies P$ because the sets in P are finite subsets of the set in Q.

Grading Notes

- If you answered **T** and did the proof by a general set is linear independent if and only if all finite subsets are linearly independent, you earned 8 points. See above. Why can you use this proof for part (b) and not for part (a)? Because to prove part (b), you did not need to use ALL finite subsets of $\{v^1, v^2, v^3\}$, just the three given in the problem statement. What are all the (nonempty) subsets? $\{v^1, v^2, v^3\}$ itself and $\{v^1, v^2\}$, $\{v^2, v^3\}$, $\{v^1, v^3\}$, $\{v^1\}$, $\{v^2\}$, and $\{v^3\}$.

- Many of you wrote that if $\{v^1, v^2\}$ is linearly dependent then $\exists\ a \neq 0$ such that $v^2 = av^1$. This is false in general, for example, if $v^1 = 0$ and $v^2 \neq 0$. I did not take off points if you were otherwise 100% on a right track. But, please note your error. What is true? If $\{v^1, v^2\}$ is linearly dependent, then $\exists\ a \neq 0$ such that either $v^2 = av^1$ OR $v^1 = av^2$. I figured in the heat of battle, you were doing pretty well. Now, just a few of you wrote, if $\{v^1, v^2\}$ is linearly dependent, then, WLOG, we can assume $\exists\ a \neq 0$ such that $v^2 = av^1$; this is spot on because one can always swap the labeling of v^1 & v^2 to make the statement true.

- Some of you need to review the difference between a proof by contradiction and proving the contra-positive. Once again, in the heat of battle, if you said one and used the other, I pretended not to notice.

- If you answered **F**....I did my best to find a point or two. Almost no one did this.

Problem 9:

Remark: When grading the A+ problem, I apply a higher degree of rigor. To earn the full five points, you need to justify every claim. If you have points taken off and a friend did not for the "same proof", it is very likely that your friend proved more of the key points in their proof than you did in yours.

Proof 0: 5 points (From your solutions!) We prove the general case by induction:

We order the e-values and e-vectors by $\lambda_1, \lambda_2, \ldots$ and $v^1, v^2, \ldots$. For $k \geq 1$, let $P(k)$ be that $\lambda_1, \ldots, \lambda_k$ distinct implies that $\{v^1, \ldots, v^k\}$ is linearly independent.

Base Case: k=1. Because v^1 is an e-vector, it is non-zero and thus the set $\{v^1\}$ is linearly independent.

Induction Step: $P(k) \implies P(k+1)$. We suppose $\lambda_1, \ldots, \lambda_k$ are distinct, $\{v^1, \ldots, v^k\}$ is linearly independent, and that $\{v^1, \ldots, v^{k+1}\}$ is linearly dependent. Hence, it must be the case that $v^{k+1} \in \text{span}\{v^1, \ldots, v^k\}$. Because v^{k+1} is an e-vector and hence non-zero, there must exist $\alpha_1, \ldots, \alpha_k$ not all zero such that

$$v^{k+1} = \alpha_1 v^1 + \cdots + \alpha_k v^k.$$

Hence,

$$\lambda_{k+1} v^{k+1} = L(v^{k+1}) = L(\alpha_1 v^1 + \cdots + \alpha_k v^k) = \alpha_1 \lambda_1 v^1 + \cdots + \alpha_k \lambda_k v^k,$$

which yields

$$0 = \alpha_1(\lambda_1 - \lambda_{k+1}) v^1 + \cdots + \alpha_k(\lambda_k - \lambda_{k+1}) v^k.$$

Because $\{v^1, \ldots, v^k\}$ is linearly independent and because at least one of the α_i is non-zero, we have that for some $1 \leq i \leq k$, $\lambda_i = \lambda_{k+1}$, proving the existence of a repeated e-value. ■

Proof 1: 5 points

We prove the contrapositive, that is, if $\{v^1, v^2, v^3\}$ is linearly dependent, then e-values are repeated. We use the notation that for two functions $f : \mathcal{X} \to \mathcal{X}$ and $g : \mathcal{X} \to \mathcal{X}$, their composition is $f \circ g(x) := f(g(x))$. We note that if f and g are linear, then so is $f \circ g$ because

$$f \circ g(\alpha x + \beta y) := f(g(\alpha x + \beta y)) = f(\alpha g(x) + \beta g(y)) = \alpha f(g(x)) + \beta f(g(y)) = \alpha f \circ g(x) + \beta f \circ g(y)$$

Claim 1 W.L.O.G., we can assume there exist α and β such that $\alpha v^1 + \beta v^2 + v^3 = 0$

Pf. Because the set $\{v^1, v^2, v^3\}$ is linearly dependent, there exist c_1, c_2, $c_3 \in \mathbb{C}$ not all zero such that $c_1 v^1 + c_2 v^2 + c_3 v^3 = 0$. If necessary, relabel the vectors such that $c_3 \neq 0$. Then we can define $\alpha := \frac{c_1}{c_3}$ and $\beta := \frac{c_2}{c_3}$.

Claim 2: Define $f : \mathcal{X} \to \mathcal{X}$ by $f(x) = L(x) - \lambda_1 x$ and $g : \mathcal{X} \to \mathcal{X}$ by $g(x) = L(x) - \lambda_2 x$. Then f and g are linear operators and $f \circ g = g \circ f$.

Pf. The linearity of both f and g is obvious and thus it is not proved. For the rest,

$$\begin{aligned} f \circ g(x) &:= f(g(x)) \\ &= f(L(x) - \lambda_2 x) \\ &= L(L(x) - \lambda_2 x) - \lambda_1 (L(x) - \lambda_2 x) \\ &= L \circ L(x) - \lambda_2 L(x) - \lambda_1 L(x) - \lambda_1 \lambda_2 x \end{aligned}$$

and

$$\begin{aligned} g \circ f(x) &:= g(f(x)) \\ &= g(L(x) - \lambda_1 x) \\ &= L(L(x) - \lambda_1 x) - \lambda_2 (L(x) - \lambda_1 x) \\ &= L \circ L(x) - \lambda_1 L(x) - \lambda_2 L(x) - \lambda_2 \lambda_1 x. \end{aligned}$$

Hence, $f \circ g = g \circ f$ and the claim is shown.

Claim 3: $f \circ g(v^2) = 0$, $g \circ f(v^1) = 0$, and $f \circ g(v^3) = (\lambda_3 - \lambda_1)(\lambda_3 - \lambda_2) v^3$

Pf. Direct substitution.

Claim 4: There is a repeated e-value

Pf. We use the linearity of $f \circ g$ and the results of Claims 1 through 3 to deduce that

$$\begin{aligned} 0 &= f \circ g(0) \\ &= f \circ g(\alpha v^1 + \beta v^2 + v^3) \\ &= \alpha f \circ g(v^1) + \beta f \circ g(v^2) + f \circ g(v^3) \\ &= 0 + 0 + (\lambda_3 - \lambda_1)(\lambda_3 - \lambda_2) v^3. \end{aligned}$$

Because $v^3 \neq 0$, we deduce that $(\lambda_3 - \lambda_1)(\lambda_3 - \lambda_2) = 0$, and thus either $(\lambda_3 - \lambda_1) = 0$ or $(\lambda_3 - \lambda_2) = 0$.

∎

Proof 2: We assume that $\mathcal{X}$ is finite dimensional. (3 points)

Because $\mathcal{X}$ is finite dimensional, we can choose a basis $\{u\} := \{u^1, \ldots, u^n\}$ for $\mathcal{X}$, and in that basis, compute a matrix representation of L, which we denote by A.

Claim 1: λ is an e-value of L if, and only if, λ is an e-value of A. Moreover,

$$L(v) = \lambda v \Leftrightarrow A[v]_{\{u\}} = \lambda[v]_{\{u\}}$$

Pf. By definition, A is a matrix representation of L if, and only if,

$$\forall\, x \in \mathcal{X}, \quad [L(x)]_{\{u\}} = A[x]_{\{u\}}.$$

We first assume that λ is an e-value of L. Then $\exists(v \in \mathcal{X}, v \neq 0)$ such that $L(v) = \lambda v$. Taking $x = v$ and using v is an e-vector of L, we have that

$$\begin{aligned}
[L(v)]_{\{u\}} &= A[v]_{\{u\}} \\
[\lambda v]_{\{u\}} &= A[v]_{\{u\}} \\
\lambda[v]_{\{u\}} &= A[v]_{\{u\}}
\end{aligned}$$

and thus we conclude that $\tilde{v} := [v]_{\{u\}} \in \mathbb{C}^n$ is an e-vector of the matrix A once we note that $\tilde{v} \neq 0$ if, and only if, $v \neq 0$.

To show the other direction, we assume that λ is an e-value of A. Hence, $\exists(\tilde{v} \in \mathbb{R}^n, \tilde{v} \neq 0)$ such that $A\tilde{v} = \lambda\tilde{v}$. We define $v \in \mathcal{X}$ by $[v]_{\{u\}} = \tilde{v}$, that is, $v := \tilde{v}_1 u^1 + \ldots + \tilde{v}_n u^n$. Then, using properties of representations and the fact that $A\tilde{v} = \lambda\tilde{v}$ we have that

$$\begin{aligned}
0 = &[L(v)]_{\{u\}} - A[v]_{\{u\}} \\
&[L(v)]_{\{u\}} - \lambda[v]_{\{u\}} \\
&[L(v) - \lambda v]_{\{u\}}
\end{aligned}$$

and thus $L(v) - \lambda v = 0$.

Claim 2: If λ_1, λ_2, and λ_3 are *distinct* e-values of L, then the corresponding set of e-vectors $\{v^1, v^2, v^3\}$ is *linearly independent.*

Pf. For $1 \leq i \leq 3$, define $\tilde{v}^i := [v^i]_{\{u\}}$. By Claim 1, $\{\tilde{v}^1, \tilde{v}^2, \tilde{v}^3\}$ are e-vectors of A. From our work in lecture, $\{\tilde{v}^1, \tilde{v}^2, \tilde{v}^3\}$ is linearly independent if, and only if, $\{v^1, v^2, v^3\}$ is linearly independent. Hence, from lecture, $\{v^1, v^2, v^3\}$ linearly independent implies the e-values are distinct.

■

Proof 3: (5 points)

We use results from lecture in a clever way.

Let $M := \text{span}\{v^1, v^2, v^3\}$ and note that for each $x \in M$, $L(x) \in M$ because

$$L(\alpha_1 v^1 + \alpha_2 v^2 + \alpha_3 v^3) = \alpha_1 \lambda_1 v^1 + \alpha_2 \lambda_2 v^2 + \alpha_3 \lambda_3 v^3.$$

Hence, $T : M \to M$ by $T(x) = L(x)$ is a well defined linear operator on the finite-dimensional vector space $(M, \mathbb{C})$, and we have that, for $1 \leq i \leq 3$, $T(v^i) = L(v^i) = \lambda_i v^i$ and therefore λ_i are e-values of T. We can now re-do the proof for a finite-dimensional space given above (i.e,m Proof 2) and deduce that if $\lambda_1, \lambda_2, \lambda_3$ are distinct, then $\{v^1, v^2, v^3\}$ is linearly independent.

This proof uses the fact that M is finite dimensional and does not assume that $(\mathcal{X}, \mathbb{C})$ is finite dimensional. Hence, it is worth five (5) points! ■

Proof 4: (5 points)

We prove the contrapositive.

Claim 3 Suppose $\{v^1, v^2, v^3\}$ is linearly dependent. W.L.O.G., we can assume there exist α and β **not both zero** such that $v^1 = \alpha v^2 + \beta v^3$

Proof: Same as Claim 1.

We are given that $L(v^1) = \lambda_1 v^1$. Hence,

$$\lambda_1(\alpha v^2 + \beta v^3) = \lambda_1 v^1 = L(v^1) = L(\alpha v^2 + \beta v^3) = \alpha \lambda_2 v^2 + \beta \lambda_3 v^3$$

and therefore, subtracting the far left side from the far right side, we have

$$0 = \alpha(\lambda_2 - \lambda_1) v^2 + \beta(\lambda_3 - \lambda_1) v^3$$

Case 1: If $\{v^2, v^3\}$ is linearly independent, then $\alpha(\lambda_2 - \lambda_1) = 0$ and $\beta(\lambda_3 - \lambda_1) = 0$. Because at least one of α and β are non-zero, we deduce that either $\lambda_2 = \lambda_1$ or $\lambda_3 = \lambda_1$ and we have proven the result.

Case 2 By Claim 3, we have $v^1 \in \text{span}\{v^2, v^3\}$. If $\{v^2, v^3\}$ is linearly dependent, because $\{v^1, v^2, v^3\}$ are e-vectors and hence are non-zero, there must $\exists\ \gamma_1 \neq 0$ and $\gamma_2 \neq 0$ such that $v^1 = \gamma_1 v^3$, and $v^2 = \gamma_2 v^3$. We then deduce that

$$\begin{aligned} 0 = \lambda_1 \gamma_1 v^3 = \lambda_1 v^1 = L(v^1) = L(\gamma_1 v^3) = \gamma_1 \lambda_3 v^3, \\ 0 = \lambda_2 \gamma_2 v^3 = \lambda_2 v^2 = L(v^2) = L(\gamma_2 v^3) = \gamma_2 \lambda_3 v^3. \end{aligned}$$

Because $\gamma_1 v^1 \neq 0$ and $\gamma_2 v^3 \neq 0$, we deduce that $\lambda_1 = \lambda_2$ and $\lambda_1 = \lambda_3$ and hence, $\lambda_1 = \lambda_2 = \lambda_3$.

Hence, $\{v^1, v^2, v^3\}$ is linearly dependent implies the e-values are repeated, proving the contrapositive. ■