

Quiz on Linear Equations with Answers

(a) State the defining equations for linear transformations.

(i) $L(u + v) = L(u) + L(v)$, \forall vectors u and v .

(ii) $L(Av) = AL(v)$, \forall vectors v and \forall numbers A .

or combine these two equations as:

$$L(Au + Bv) = AL(u) + BL(v), \forall \text{ vectors } v \text{ and } w \text{ and } \forall \text{ numbers } A \text{ and } B.$$

(b) State and prove the theorem on homogeneous linear equations.

This is Theorem II.3.1. There are two statements of this theorem in the text.

Proof by calculation: Given $L(v) = 0$ and $L(w) = 0$ and $u = Av + Bw$.

Then $L(Av + Bw) = AL(v) + BL(w) = A \times 0 + B \times 0 = 0$. Thus $L(u) = 0$, $\forall A$ and B . \checkmark

(c) State the algorithm implied by the theorem on homogeneous linear equations. From text, immediately after Theorem II.3.1:

Step 0. Observe/Check that the equation is indeed a homogeneous linear equation

Step 1. Find the simplest solutions.

Step 2. Take all their linear combinations.

(d) State the theorem on non-homogeneous linear equations. This is Theorem II.4.3.

(e) Given $\ddot{x} - x = 3$.

(i) Using the Proposition on Recognizing linear transformation, prove that the relevant transformation is a linear transformation. State which parts of the Proposition on Recognizing linear transformation you used.

Set $L(x) = \ddot{x} - x = 3$. Review Example II.2.5, use either scheme.

(ii) Solve the equation. Indicate the line where the theorem on homogeneous linear equations is used. Indicate the line where the theorem on non-homogeneous linear equations is used.

Similar to Example II.4.8 and Example II.3.7.

For a simple solution to the non-homogeneous linear equation, try $x_c = C$, where C is an unknown, to-be-found constant. Then $L(C) = 0 - C = 3$, $\implies x_c = C = -3$. The associated homogeneous equation is $L(u) = \ddot{u} - u = 0$. Try $u = e^{rt}$, where r is an

2

unknown, to-be-found constant. Plug in, solve for $r = \pm 1$. Hence $u_1 = e^t$ and $u_2 = e^{-t}$ are two solutions of the associated homogeneous equation.

Using the theorem on homogeneous linear equations: $u = Au_1 + Bu_2 = Ae^t + Be^{-t}, \forall A$ and B .

Using the theorem on non-homogeneous linear equations: $x = -3 + Ae^t + Be^{-t}, \forall A$ and B .

Exercise II.6.4

Is zero an eigenvalue, why or why not?

Answer: For $\lambda = 0$, the main eigenvalue equation simplifies to $L(X(x)) = -X''(x) = 0$. Integrating this equation twice, yields: $X(x) = Ax + B$, $\forall A$ and B .

Plugging in the boundary conditions, $X(0) = 0 = X(\pi)$ will yield:

$$0 = X(0) = B; \text{ this simplifies the function to } X(x) = Ax.$$

$$\text{Then } 0 = X(\pi) = A\pi \implies A = 0 \implies X(x) = 0.$$

But eigenfunctions cannot be zero, hence $X(x) = 0$ is not an eigenfunction. Without an eigenfunction, $\lambda = 0$ cannot be an eigenvalue.

Prove that there are no negative eigenvalues.

For $\lambda < 0$, we note: $-\lambda > 0$. The main eigenvalue equation is a homogeneous linear equation. The standard simple guess is $X(x) = e^{rx}$. Plugging this in and then solving for r , yields: $r^2 = -\lambda$, and hence $r = \pm\sqrt{-\lambda}$. Thus $X_1 = e^{\sqrt{-\lambda}x}$ and $X_2 = e^{-\sqrt{-\lambda}x}$ are two solutions. Then the Theorem on Homogeneous Linear Equations implies that $X(x) = Ae^{\sqrt{-\lambda}x} + Be^{-\sqrt{-\lambda}x}$, $\forall A$ and B are many solutions. Plugging in the boundary conditions, $X(0) = 0 = X(\pi)$ will yield:

$$0 = X(0) = A + B \implies A = -B.$$

$$0 = X(\pi) = Ae^{\sqrt{-\lambda}\pi} + Be^{-\sqrt{-\lambda}\pi} = B(-e^{\sqrt{-\lambda}\pi} + e^{-\sqrt{-\lambda}\pi}).$$

Since e^t is a strictly monotone increasing function, $e^t \neq e^{-t}$, except when $t = 0$. Hence $e^{\sqrt{-\lambda}\pi} \neq e^{-\sqrt{-\lambda}\pi}$, $\forall \lambda < 0$. Hence $B = 0$.

Thus $A = 0 = B$, and hence $X(x) = 0$. But eigenfunctions cannot be zero, hence $X(x) = 0$ is not an eigenfunction. Without an eigenfunction, all possible negative λ 's cannot be eigenvalues.

Solutions to some Exercises

Exercise III.2.14. Shortcuts: Use a big zero instead of writing a block of many zeros.

Observe that $C = 10B$, then $C^2 = 100B^2$ and $C^3 = 1000C^3$, etc.

$$\boxed{B^6 = O = C^6.}$$

Exercise III.2.17. An answer is supplied by Proposition III.5.9.

Exercise III.2.31. Note: The last matrix is a 1×1 -mx; not a 2×1 -mx. One tip-off is the plus sign at the end of the top line. The other tip-off is that multiplying $(x, y, z)S \begin{pmatrix} x \\ y \\ z \end{pmatrix}$ produces a 1×1 -mx.

Exercise III.2.32. An answer is supplied in Ch. 7 Sec.2 by Observation 2.3 (turn page) and by the first 4 lines of Corollary 7.2.3 (turn page).

Exercise III.3.7. (c) By part (b), $N = PMP^{-1}$, is the product of invertible matrices, hence it is also invertible. (General product Rule for Inverses.) Then using the formula for the inverse of the product of invertible matrices, $((ABC)^{-1} = C^{-1}B^{-1}A^{-1})$,

$$N = PMP^{-1} \implies N^{-1} = (P^{-1})^{-1}M^{-1}P^{-1} = PM^{-1}P^{-1}$$

Similarly, given $M = P^{-1}NP$, using the formula for the inverse of the product of invertible matrices, will provide $M^{-1} = P^{-1}N^{-1}P$.

Exercise III.2.36. The eigenvectors associated with eigenvalue one, for the matrix are $x \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \forall x$. The useful, steady-state vector, for a stochastic matrix, is the one that is a

probability vector. Here that is the vector: $\begin{pmatrix} .5 \\ .5 \\ 0 \end{pmatrix}$.

Exercise III.2.37. The purpose of this exercise is to find “which” matrices commute with the diagonal matrix: $D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}$. In math, the word “which” means “all those which”.

The set of matrices, which commute with $D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}$ is the set of *all* 3×3 -diagonal matrices. That is, if $MD = DM$, then M is a 3×3 -diagonal matrix.

(iii) A general rule is that, if $D = \begin{pmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{pmatrix}$, then $D^n = \begin{pmatrix} a^n & 0 & 0 \\ 0 & b^n & 0 \\ 0 & 0 & c^n \end{pmatrix}, \forall n \in \mathbb{Z}^+$.

For $n = 0$, one uses the definition: $M^0 = I_n, \forall n \times n$ -matrices. This formula is also valid for $n = -1$.

Exercise III.2.37. The matrices, which commute with $D_2 = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}$ are the matrices, $\begin{pmatrix} a & b & 0 \\ c & d & 0 \\ 0 & 0 & e \end{pmatrix}, \forall a,b,c,d,e$.

Exercise III.3.7. (c) Given invertible matrices, M and P , such that $M = P^{-1}NP$, prove that N is an invertible matrix.

Multiply $M = P^{-1}NP$ on the left by P and on the right by P^{-1} , this yields: $P^{-1}MP = PP^{-1}NPP^{-1} = N$. Thus N is the product of three invertible matrices. Hence it is invertible.

Exercise III.4.9. The stochastic matrix is $\begin{pmatrix} .5 & .7 \\ .5 & .3 \end{pmatrix}$. Tomorrow’s probability vector is $\begin{pmatrix} .56 \\ .44 \end{pmatrix}$. Yesterday’s probability vector was $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$.

Exercise III.4.12. After doing the matrix multiplication, both off diagonal terms are:

$$-b(\sin^2 \theta - \cos^2 \theta) + (c - a) \sin \theta \cos \theta = 0.$$

Either (i) use the double angle formulas for $\sin 2\theta$ and $\cos 2\theta$, and then solve for $\tan 2\theta$, or (ii) divide by $\sin \theta \cos \theta$, convert into a quadratic equation in $\tan \theta$, set $x = \tan \theta$, solve for x .

When $a = c$, off diagonal terms simplify to $-b(\sin^2 \theta - \cos^2 \theta) = 0$, which implies that $\theta = \frac{\pi}{4}$.

Exercise III.4.24 and 17. See the proof of Proposition VI.3.2.

Rule III.6.4 (i) To prove $(A^{-1})^T = (A^T)^{-1}$

Remark. We will use the defining equation for inverse matrices, which says that two matrices, M and N are inverse matrices, when $MN = I = NM$. Or verbally, *two matrices are inverse matrices, when their product (both ways) is the identity matrix.*

We will use the Product Rule for transposes.

Remark. There are two basic ways to prove that two matrices are inverse matrices; the following two proofs are good examples:

Proof. #1. We start with the given information that matrix A is an invertible matrix; we translate this fact into an equation, namely the defining equations:

$$AA^{-1} = I = A^{-1}A$$

We will now manipulate these equations, in order to obtain the defining equations for $(A^{-1})^T$ and (A^T) to be invertible matrices:

$$I = I^T = (A^{-1}A)^T = A^T (A^{-1})^T$$

Similarly:

$$I = I^T = (AA^{-1})^T = (A^{-1})^T A^T.$$

Thus the product (both ways) of $(A^{-1})^T$ and (A^T) is the identity matrix; hence

$$\boxed{(A^{-1})^T = (A^T)^{-1}.$$

Proof. #2. The equation, $(A^{-1})^T = (A^T)^{-1}$ means that $(A^{-1})^T$ and (A^T) are inverse matrices. To prove this, we check that their product (both ways) is the identity matrix:

$$(A^{-1})^T \times (A^T) = (A^{-1}A)^T = I^T = I$$

$$(A^T) \times (A^{-1})^T = (A^{-1}A)^T = I^T = I$$

hence $\boxed{(A^{-1})^T = (A^T)^{-1}.$

Exercise III.6.14 (b) and key part of proof of Theorem III.7.5. Given $P = M(M^T M)^{-1}M^T$, with $M^T M$, an invertible matrix; but M not an invertible matrix. Prove that $P^T = P$.

Proof.

$$P^T = [M \times (M^T M)^{-1} \times M^T]^T$$

What is this? This is the transpose of the product of matrices. So we use the Product Rule (III.6.4(f)) for Transposes. Thus:

$$P^T = [M^T]^T \times [(M^T M)^{-1}]^T \times M^T = M \times [(M^T M)^T]^{-1} \times M^T$$

$$(M^T M)^T = M^T [M^T]^T = M^T M$$

We have also used the rules $[M^T]^T = M$ and $(A^{-1})^T = (A^T)^{-1}$.

Exercise III.6A.8. Let S be an invertible symmetric matrix. Check that S^{-1} is also symmetric.

Proof.

$$(S^{-1})^T = (S^T)^{-1} = S^{-1},$$

(using the “inverse” rule for transposes.) Thus S^{-1} is also symmetric.

Exercise III.6A.11. Given unknown 7×7 matrices, S, D and P , where P is an invertible matrix and S is a symmetric matrix, such that $S = P^T D P$. Let $\mathbf{v} \in \mathbb{R}^7$ be a coordinate vector. Let $\mathbf{w} = P\mathbf{v}$. Given $(z) = \mathbf{v}^T S \mathbf{v}$.

Show that $(z) = \mathbf{w}^T D \mathbf{w}$.

Proof.

$$(z) = \mathbf{v}^T S \mathbf{v} = \mathbf{v}^T P^T D P \mathbf{v} = \mathbf{w}^T D \mathbf{w}$$

since $\mathbf{v}^T P^T = (P\mathbf{v})^T = \mathbf{w}$.

Answers to some exercises for Ch. IV.

Exercise IV.3A.2. $|E| \leq J_2 = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$. $|M^{-1}E|\mathbf{1} \leq |M^{-1}| \times |E|\mathbf{1} \leq \frac{1}{49} \binom{22}{18} < \mathbf{1}$.
Therefore, $M + E$ is an invertible matrix.

Exercise IV.3A.3. $|M^{-1}E|\mathbf{1} \leq |M^{-1}| \times |E|\mathbf{1} \leq \binom{2.2}{1.8} > \mathbf{1}$. Therefore, the lemma provides *no* info on the invertibility of the matrix, $M + E$.

Remember, E is unknown, hence $M^{-1}E$ is unknown and cannot be calculated. Only, bounds on E are known. Therefore, need to use $|M^{-1}E|\mathbf{1} \leq |M^{-1}| \times |E|\mathbf{1}$.

Using $E = \frac{1}{10} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$, and calculating $|M^{-1}E|\mathbf{1} = \left| M^{-1} \frac{1}{10} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \right| \mathbf{1}$ will produce a *wrong* answer.

Exercise IV.3B.2. $d = 0.4002 < 1$. Hence $M + E$ is invertible. $\|\Delta v\|_\infty \leq 1.6681$.

Exercise IV.3B.3. $d = 0.1 < 1$. Hence $M + E$ is invertible.

Part (i) According to Theorem IV.3B.1, $\Delta v = 0$; according to Corollary IV.3B.2, $\|\Delta v\|_\infty \leq \frac{1}{9}$. For this matrix and vector, the error bound provided by the theorem is much lower than the one provided by the corollary.

Part (ii) According to both Theorem IV.3B.1 and Corollary IV.3B.2, $\|\Delta v\|_\infty \leq \frac{1}{9}$. For this matrix and vector, the two error bounds are the same.

Exercise IV.3B.5. (Not Superposition)

Part (a). $|\Delta v_a| \leq \begin{pmatrix} .004 \\ .006 \end{pmatrix}$.

Part (b). $d = 0.012 < 1$. Hence $M + E$ is invertible. $\|\Delta v_b\|_\infty \leq \frac{.066}{1-0.12} = 0.0668$.

Part (c). $\|\Delta v_c\|_\infty \leq 0.0668 + \frac{.006}{1-0.12} = 0.07287$.

Part (d). $\|\Delta v_a\|_\infty + \|\Delta v_b\|_\infty \leq 0.006 + 0.0668 = 0.0724 < 0.07287$, the bound for $\|\Delta v_c\|_\infty$.

Thus the error bound, due to the two errors, is larger than the sum of the two error bounds due to the individual errors. This is to be expected, since the total error due to two individual errors is often larger than the sum. Superposition does *not* occur for errors or their bounds.

Exercise IV.3B.6. (Iterative improvement)

Part (b). $|\Delta v_{6b}| \leq \frac{1}{1000} \left[\binom{44}{66} + .0668 \binom{8}{12} \right] = \begin{pmatrix} .04454 \\ .0668 \end{pmatrix}$. This provides a significant reduction in the bound on the first coordinate but no reduction in the bound on the second coordinate (in comparison with the results of Exercise IV.3B.5 part (b)).

Part (c) $|\Delta v_{6c}| \leq \frac{1}{1000} \left[\binom{44}{66} + .07287 \binom{8}{12} + \binom{4}{6} \right] = \begin{pmatrix} .04858 \\ .07287444 \end{pmatrix}$. Again, this provides a significant reduction in the bound on the first coordinate but no reduction in the bound on the second coordinate (in comparison with the results of Exercise IV.3B.5 part (c)). Actually, the bound on the second coordinate is slightly higher; but since both pairs of bounds are valid, we may combine them to obtain: $|\Delta v_{6c}| \leq \begin{pmatrix} .04858 \\ .07287 \end{pmatrix}$.