Fall 2008 **ARE211**

Problem Set #06- Answer key

Third Lin Algebra Problem Set

Due date: Oct 23

Problem 1

Let $\mathbf{A} = \begin{pmatrix} 3 & 3 & -1 & 0 \\ -2 & -3 & 1 & 0 \\ -6 & -2 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$ and $\mathbf{B} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & -6 \end{pmatrix}$

a) Show that $\mathbf{A}$ and $\mathbf{B}$ are nonsingular.

b) Calculate $\mathbf{A}^{-1}$, $\mathbf{B}^{-1}$, $(\mathbf{AB})^{-1}$, and $\left(\mathbf{A}^T\right)^{-1}$.

c) Solve the linear equation system $\mathbf{Ax} = \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \end{pmatrix}$ and $(\mathbf{AB})\mathbf{x} = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}$

Ans:

a) emphShow that **A** and **B** are nonsingular. A matrix is nonsingular if it's determinant is nonzero.

$$det(\mathbf{A}) = \begin{vmatrix} 3 & 3 & -1 & 0 \\ -2 & -3 & 1 & 0 \\ -6 & -2 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{vmatrix} \qquad \text{(Develop after the fouth column)}$$

$$= \begin{vmatrix} 3 & 3 & -1 \\ -2 & -3 & 1 \\ -6 & -2 & 1 \end{vmatrix} \qquad \text{(Develop after the third column)}$$

$$= (-1)\begin{vmatrix} -2 & -3 \\ -6 & -2 \end{vmatrix} + (-1)\begin{vmatrix} 3 & 3 \\ -6 & -2 \end{vmatrix} + \begin{vmatrix} 3 & 3 \\ -2 & -3 \end{vmatrix}$$

$$= -(4-18) - (-6+18) + (-9+6) = 14 - 12 - 3 = -1 \neq 0$$

Note that for a diagonal matrix the determinant is simply the product of the diagonal elements:

$$det(\mathbf{B}) = \begin{vmatrix} -1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & -6 \end{vmatrix} = 36 \neq 0$$

b) Let's first calculate the inverse of A

$$\left(\begin{array}{cccc|cccc} 3 & 3 & -1 & 0 & 1 & 0 & 0 & 0 \\ -2 & -3 & 1 & 0 & 0 & 1 & 0 & 0 \\ -6 & -2 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{array}\right)$$

Add the 3^{rd} row to the 1^{st} row and*(-1) 3^{rd} row to the 2^{nd} row:

$$\Leftrightarrow \left(\begin{array}{cccc|cccc} -3 & 1 & 0 & 0 & 1 & 0 & 1 & 0 \\ 4 & -1 & 0 & 0 & 0 & 1 & -1 & 0 \\ -6 & -2 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{array}\right)$$

Add the 2^{nd} row to the 1^{st} row:

$$\Leftrightarrow \left(\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 4 & -1 & 0 & 0 & 0 & 1 & -1 & 0 \\ -6 & -2 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{array}\right)$$

Add 6*1^{st} row to the 3^{rd} row and (-4)*1^{st} row to the 2^{nd} row:

$$\Leftrightarrow \left(\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & -1 & 0 & 0 & -4 & -3 & -1 & 0 \\ 0 & -2 & 1 & 0 & 6 & 6 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{array}\right)$$

Divide the 2^{nd} row by (-1) and add 2*2^{nd} row to the 3^{rd} row:

$$\Leftrightarrow \left(\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 4 & 3 & 1 & 0 \\ 0 & 0 & 1 & 0 & 14 & 12 & 3 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{array}\right)$$

Hence, $\mathbf{A}^{-1} = \underline{\underline{\begin{pmatrix} 1 & 1 & 0 & 0 \\ 4 & 3 & 1 & 0 \\ 14 & 12 & 3 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}}}$

Since **B** is a digonal matrix, the inverse is simply a digonal matrix where each diagonal element is the inverse of the corresponding diagonal element of **B**.

Hence, $\mathbf{B}^{-1} = \underline{\underline{\begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & \frac{1}{3} & 0 \\ 0 & 0 & 0 & -\frac{1}{6} \end{pmatrix}}}$

Using the fact that $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$ we know:

$$(\mathbf{AB})^{-1} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & \frac{1}{3} & 0 \\ 0 & 0 & 0 & -\frac{1}{6} \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 & 0 \\ 4 & 3 & 1 & 0 \\ 14 & 12 & 3 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = \underline{\underline{\begin{pmatrix} -1 & -1 & 0 & 0 \\ 2 & \frac{3}{2} & \frac{1}{2} & 0 \\ \frac{14}{3} & 4 & 1 & 0 \\ 0 & 0 & 0 & -\frac{1}{6} \end{pmatrix}}}$$

Using the fact that $(\mathbf{A}^T)^{-1} = (\mathbf{A}^{-1})^T$ we know:

$$(\mathbf{A}^T)^{-1} = \underline{\underline{\begin{pmatrix} 1 & 4 & 14 & 0 \\ 1 & 3 & 12 & 0 \\ 0 & 1 & 3 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}}}$$

c) Solving $\mathbf{Ax} = \mathbf{b}$ we know: $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$

$$\mathbf{x} = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 4 & 3 & 1 & 0 \\ 14 & 12 & 3 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \end{pmatrix} = \underline{\underline{\begin{pmatrix} 0 \\ 1 \\ 2 \\ 0 \end{pmatrix}}}$$

Solving $(\mathbf{AB})\mathbf{x} = \mathbf{b}$ we know: $\mathbf{x} = (\mathbf{AB})^{-1}\mathbf{b}$

$$\mathbf{x} = \begin{pmatrix} -1 & -1 & 0 & 0 \\ 2 & \frac{3}{2} & \frac{1}{2} & 0 \\ \frac{14}{3} & 4 & 1 & 0 \\ 0 & 0 & 0 & -\frac{1}{6} \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} = \underline{\underline{\begin{pmatrix} -2 \\ 3.5 \\ \frac{26}{3} \\ 0 \end{pmatrix}}}$$

Problem 2

Cramer's rule: Simon & Blume question 9.13 (page 196).

Ans: In this problem you were asked to solve linear equation systems of the form $\mathbf{Ax} = \mathbf{b}$ using Cramer's rule.

a) $\mathbf{A} = \begin{pmatrix} 5 & 1 \\ 2 & -1 \end{pmatrix}$ and $\mathbf{b} = \begin{pmatrix} 3 \\ 4 \end{pmatrix}$

$$x_1 = \frac{\begin{vmatrix} 3 & 1 \\ 4 & -1 \end{vmatrix}}{\begin{vmatrix} 5 & 1 \\ 2 & -1 \end{vmatrix}} = \frac{-5-2}{-5-2} = \frac{-7}{-7} = \underline{\underline{1}}$$

$$x_2 = \frac{\begin{vmatrix} 5 & 3 \\ 2 & 4 \end{vmatrix}}{\begin{vmatrix} 5 & 1 \\ 2 & -1 \end{vmatrix}} = \frac{20-6}{-5-2} = \frac{14}{-7} = \underline{\underline{-2}}$$

b) $\mathbf{B} = \begin{pmatrix} 2 & -3 & 0 \\ 4 & -6 & 1 \\ 1 & 10 & 0 \end{pmatrix}$ and $\mathbf{b} = \begin{pmatrix} 2 \\ 7 \\ 1 \end{pmatrix}$

Always develop the determinant after the third column:

$$x_1 = \frac{\begin{vmatrix} 2 & -3 & 0 \\ 7 & -6 & 1 \\ 1 & 10 & 0 \end{vmatrix}}{\begin{vmatrix} 2 & -3 & 0 \\ 4 & -6 & 1 \\ 1 & 10 & 0 \end{vmatrix}} = \frac{-(20+3)}{-(20+3)} = \frac{-23}{-23} = \underline{\underline{1}}$$

$$x_2 = \frac{\begin{vmatrix} 2 & 2 & 0 \\ 4 & 7 & 1 \\ 1 & 1 & 0 \end{vmatrix}}{\begin{vmatrix} 2 & -3 & 0 \\ 4 & -6 & 1 \\ 1 & 10 & 0 \end{vmatrix}} = \frac{2-2}{-(20+3)} = \frac{0}{-23} = \underline{\underline{0}}$$

$$x_3 = \frac{\begin{vmatrix} 2 & -3 & 2 \\ 4 & -6 & 7 \\ 1 & 10 & 1 \end{vmatrix}}{\begin{vmatrix} 2 & -3 & 0 \\ 4 & -6 & 1 \\ 1 & 10 & 0 \end{vmatrix}} = \frac{2(40+6)-7(20+3)+(-12+12)}{-(20+3)} = \frac{-69}{-23} = \underline{\underline{3}}$$

Problem 3

Show that if the matrix $\mathbf{A}$ is nonsingular and symmetric, then the matrix $\mathbf{A}^{-1}$ is also symmetric. (You can use as a fact that the left- and right inverse of the matrix $\mathbf{A}$ are the same and that the inverse is unique).

Ans: $\mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$ (Definition of the inverse)

$= \mathbf{I}^T$ (Identity matrix is symmetric)

$= \left(\mathbf{A}\mathbf{A}^{-1}\right)^T$ (Left inverse equals right inverse)

$= \left(\mathbf{A}^{-1}\right)^T \mathbf{A}^T$ (see problem 1a)

$= \left(\mathbf{A}^{-1}\right)^T \mathbf{A}$ (A is symmetric)

Since the inverse is unique we therefore know that: $\mathbf{A}^{-1} = \left(\mathbf{A}^{-1}\right)^T$, i.e., $\mathbf{A}^{-1}$ is symmetric.

Problem 4

Given an example of a 2x2 matrix $\mathbf{A}$ such that the function $f(\mathbf{x}) = \mathbf{Ax}$

a) maps the unit circle to itself.

b) maps the unit circle to a line in $\mathbb{R}^2$.

c) maps the unit circle to a single point.

d) over several iterations, maps the unit circle to an ellipse the size of Miami.

e) rotates every vector by 45 degrees (counter-clockwise) and stretches it by a factor of 2.

For each of the above, provide a sketch (or series of sketches) and an explanation.

Ans: Given an example of a 2x2 matrix $\mathbf{A}$ such that the function $f(\mathbf{x}) = \mathbf{Ax}$

a) *maps the unit circle to itself*:
The only *symmetric* matrix that maps the unit circle to itself is the identity matrix (there are also non-symmetric matrices which do so)
$$\mathbf{A}_a = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

b) *maps the unit circle to a line in* $\mathbb{R}^2$.
Any matrix with rank 1 will do so, for example:
$$\mathbf{A}_b = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$$

c) *maps the unit circle to a single point.*
The only matrix that maps the unit circle to a point is the zero matrix:
$$\mathbf{A}_c = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

d) *over several iterations, maps the unit circle to an ellipse the size of Miami.*
Any matrix with two eigenvalues bigger than one in modulus will do so. Since I am lazy, I take the example from part e).

e) *rotates every vector by 45 degrees (counter-clockwise) and stretches it by a factor of 2.*
Following the example in the lecture notes, the matrix is:
$$\mathbf{A}_e = \begin{pmatrix} \sqrt{2} & -\sqrt{2} \\ \sqrt{2} & \sqrt{2} \end{pmatrix}$$

The images are diplayed on the next page. In the first graph of figure 1, the vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ are the eigenvectors. In the second graph, the matrix has no eigenvectors, but for illustration purposes, the unit vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ and their images are displayed.

Problem 5

Compute the directional derivative of the function $f(x, y) = xy^2 + x^3y$ at the point (4,-2) in the direction $(\frac{1}{\sqrt{10}}, \frac{3}{\sqrt{10}})$.

Ans: There are 2 standard steps to calculate a directional derivative:

step 1: Parameterize the direction $\mathbf{w}$ from your given starting point $\mathbf{x}^0$:

$$\mathbf{g}(t) = \mathbf{x}^0 + \frac{t}{||\mathbf{w}||}\mathbf{w} = \begin{pmatrix} x_1^0 + \frac{t*w_1}{||\mathbf{w}||} \\ x_2^0 + \frac{t*w_2}{||\mathbf{w}||} \end{pmatrix}$$

Note: Make sure to divide by length of $\mathbf{w}$. (In this example the length was 1, i.e., $||\mathbf{w}|| = 1$)

For this problem: $\mathbf{g}(t) = \begin{pmatrix} 4 + \frac{t}{\sqrt{10}} \\ -2 + \frac{3t}{\sqrt{10}} \end{pmatrix}$

step 2: Take the chain rule for the combined function h(t) = f(g(t)) to obtain:

$$h'(t) = \nabla f(\mathbf{x}^0) \;*\; \frac{\mathbf{w}}{||\mathbf{w}||}$$

For this question $f_{x_1} = x_2^2 + 3x_1^2x_2$ and $f_{x_2} = 2x_1x_2 + x_1^3$ and hence:
$f_{x_1}(\mathbf{x}^0) = (-2)^2 + 3 * 16 * (-2) = -92$ and
$f_{x_2}(\mathbf{x}^0) = 2 * 4 * (-2) + 4^3 = 48$

$$h'(t) = \begin{pmatrix} -92 \\ 48 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{10}} \\ \frac{3}{\sqrt{10}} \end{pmatrix} = \frac{-92 + 144}{\sqrt{10}} = \underline{\underline{\frac{52}{\sqrt{10}}}}$$

Problem 6

Consider the function $f(x) = e^x$.

a) Calculate the (i) third order and (ii) fourth order Taylor approximation of $f(.)$ around the point $x = 0$.

b) Approximate the value of $f(x)$ using the two Taylor approximations of part (a) for x = 0.2 and x = 1.

c) Calculate the actual value of $f(x)$ for x = 0.2 and x = 1.

Ans: Recall the definition of a Taylor expansion (here $x_0 = 0$):

$$f(x) = f(x_0 = 0) + \sum_{k=1}^{n} \frac{f^{(k)}(x_0 = 0)(x-0)^k}{k!}$$

Note that for the exponential distribution $\forall n : \ f^{(n)}(0) = e^0 = 1$

a) Therefore the third and fourth order Taylor expansions are:

(i) $f^{3^{rd}}(x) = 1 + \frac{x^1}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!}$
$= 1 + x + \frac{x^2}{2} + \frac{x^3}{6}$

(ii) $f^{4^{th}}(x) = 1 + \frac{x^1}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!}$
$= 1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \frac{x^4}{24}$

b) $f^{3^{rd}}(0.2) = 1 + 0.2 + \frac{0.2^2}{2} + \frac{0.2^3}{6} = \underline{\underline{\frac{1893}{1550} = 1.2213}}$

$f^{3^{rd}}(1) = 1 + 1 + \frac{1}{2} + \frac{1}{6} = \underline{\underline{\frac{8}{3} = 2.667}}$

$f^{4^{th}}(0.2) = f^{3^{rd}}(0.2) + \frac{0.2^4}{24} = \underline{\underline{\frac{1893}{1550} + \frac{1}{15000} = 1.2214}}$

$f^{4^{th}}(1) = f^{3^{rd}}(1) + \frac{1^4}{24} = \frac{64}{24} + \frac{1}{24} = \underline{\underline{\frac{65}{24} = 2.7083}}$

c) $e^{0.2} = \underline{\underline{1.2214}}$ and $e^1 = \underline{\underline{2.7183}}$

Problem 7

Recall the slightly different version of Taylor's theorem from Section:

If the function f is (n+1) times continously differentiable on $I = (a, b)$ then we know that for $x_0 \in I$ and for $x = x_0 + h \in I$:

$$f(x) = f(x_0 + h) = f(x_0) + \sum_{k=1}^{n} \frac{f^{(k)}(x_0)(x-x_0)^k}{k!} + R_n(x)$$

where $R_n(x) = f^{(n+1)}(\eta)\frac{(x-x_0)^{n+1}}{(n+1)!}$ for some η between x_0 and x.

a) Show that a sufficient condition for f to a attain a strict (local) maximum on I at x_0 is that the derivatives $f^{(k)}(x_0)$ are zero for k = 1...n, and $f^{(n+1)}(x_0)$ is negative for some odd n.

b) If $f^{(k)}(x_0)$ is zero for k = 1...n-1 and $f^{(n)}(x_0)$ is non-zero, show that there exists an ϵ-neighborhood around x_0 where the absolute value of the n^{th}-order Taylor expansion is bigger than the absolute value of the remainder term $R_n(x)$.

c) Show that the n^{th}-order Taylor approximation around any point x_0 of a polynomial of degree n (i.e. a function of the form $f(x) = \sum_{k=0}^{n} \alpha_k x^k$) is the function itself.

Ans:

a) <u>*Given :*</u> $f^{(k)}(x_0) = 0$ for k = 1...n, and $f^{(n+1)}(x_0) < 0$ for some odd n.
Proof that f attains a strict local maximum at x_0:
   (1) Since f is (n+1) times *continously* differentiable and $f^{(n+1)}(x_0) < 0$ we know from problem 9 on the midterm that for $\epsilon = |f^{(n+1)}(x_0)|$
   $\exists \delta > 0$ such that $\forall x$ with $|x - x_0| < \delta$:

$$|f^{(n+1)}(x) - f^{(n+1)}(x_0)| < \epsilon = |f^{(n+1)}(x_0)|$$

   and hence $f^{(n+1)}(x) < 0$.
   (2) We are given that n is odd. Hence (n+1) is even and $(x - x_0)^{(n+1)} > 0$ for $x \neq x_0$
   (3) From (1) and (2) we know that $\forall x \neq x_0$ with $|x - x_0| < \delta : R_n(x) < 0$
   (4) We are given that $f^{(k)}(x_0)$ are zero for k = 1...n
   (5) Therefore $\forall x \neq x_0$ with $|x - x_0| < \delta$ from (1):

$$f(x) = f(x_0) + \underbrace{\sum_{k=1}^{n} \frac{f^{(k)}(x_0)(x - x_0)^k}{k!}}_{=0 \text{ by (4)}} + \underbrace{R_n(x)}_{<0 \text{ by (3)}}$$

   And hence in the δ-neighborhood around x_0: $f(x) < f(x_0)$ which is the definition for a strict local max.

b) You are asked to show Leo's version from class. i.e., for a sufficiently small ϵ-neighborhood around x_0 the first non-zero Taylor expansion will give the right sign.
Since the function is (n+1) times continously differentiable we know from the Weierstrass theorem that the continous function $|f^{(n+1)}(x_0)|$ has to attain a maximum on the compact set $[x_0 - 1; x_0 + 1]$. Let's denote this maximum by $\bar{f}$.
case I: If $\bar{f} = 0$ the remainder term $R_n(x) = 0$ for $x \in (x_0 - 1; x_0 + 1)$ and the proposition thus holds trivially as we are given $|f^{(n)}(x_0)| > 0$
case II: If $\bar{f} \neq 0$, then define $\epsilon = \min\{1, \frac{|f^{(n)}(x_0)|}{\bar{f}}\} > 0$.

(1) First, note that by the construction of ϵ above, $\forall x$ with $|x - x_0| < \epsilon$ we have $\frac{\bar{f}\ |x-x_0|}{|f^{(n)}(x_0)|} < \frac{\bar{f}\ \epsilon}{|f^{(n)}(x_0)|} \leq 1$.

(2) We therefore know that $\forall x \in (x_0 - \epsilon, x_0 + \epsilon)$:

$$\left|\frac{f^{(n)}(x_0)(x-x_0)^n}{n!}\right| > \left|\frac{f^{(n)}(x_0)(x-x_0)^n}{(n+1)!}\right| \qquad \text{(Divide by } n+1>1\text{)}$$

$$> \left|\frac{f^{(n)}(x_0)(x-x_0)^n}{(n+1)!}\right| \frac{\bar{f}*|x-x_0|}{|f^{(n)}(x_0)|} \qquad \text{(By (1))}$$

$$= \left|\bar{f} * \frac{(x-x_0)^{n+1}}{(n+1)!}\right|$$

$$\geq \left|\frac{f^{(n+1)}(\eta)(x-x_0)^{n+1}}{(n+1)!}\right| \qquad \text{(By def. of } \bar{f}\text{)}$$

$$= |R_n(x)|$$

c) From your basic calculus class you should remember that for $g(t) = \alpha t^k$ If $n \leq k$: $g^{(n)}(t) = t * (t-1) * ... * (t-n) * \alpha t^{k-n}$

If $n > k$: $g^{(n)}(t) = 0$

Hence the $(n+1)^{th}$ derivative of a function $f(x) = \sum_{k=0}^{n} \alpha_k x^k$ is equal to zero and the remainder term $R_n(x) = f^{(n+1)}(\eta)\frac{(x-x_0)^{n+1}}{(n+1)!}$ is zero. And $R_n(x) = 0$ implies that the approximation is correct.

Problem 8

Consider the function $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ with $f(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{b}$ (where $\mathbf{A}$ is a (nxn) matrix and $\mathbf{b}$ is a (nx1) column vector).

Show that f is bijective $\Leftrightarrow$ A is nonsingular.

Ans: Show that f is bijective $\Leftrightarrow$ **A** is nonsingular.

"$\Rightarrow$" Let's prove the contra-positive:

(1) Assume **A** is singular. Then we know that the rank of **A** must be less than n.

(2) Since the rank of **A** plus the dimension of the kernel has to sum to n (see lecture notes #15) the dimension of the kernel is at least 1. Hence there are infinitely many points that get mapped to **b**.

(3) From (2) we know that f can't be injective, hence it is not bijective.

"$\Leftarrow$" Let's prove the contra-positive:

Assume f is not bijective, i.e., it is either not injective or not surjective.

case 1: If f is not injective, then $\exists \mathbf{x}^1, \mathbf{x}^2$ with $\mathbf{x}^1 \neq \mathbf{x}^2$ such that $\mathbf{A}\mathbf{x}^1 + \mathbf{b} = \mathbf{A}\mathbf{x}^2 + \mathbf{b}$ which implies that $\mathbf{A}\underbrace{(\mathbf{x}^1 - \mathbf{x}^2)}_{\neq 0} = 0$

Hence, by the same argument as before the dimension of the kernel is at least 1 and thus the rank of the matrix can be at most (n-1). This implies that the matrix is singular.

case 2: If f is not surjective, then $\exists \mathbf{y} \in \mathbb{R}^n$ such that $\forall \mathbf{x} \in \mathbb{R}^n$:

$$\mathbf{A}\mathbf{x} \neq \mathbf{y} - \mathbf{b}.$$

This implies that the n column vectors of **A** do not span $\mathbb{R}^n$, and hence they can't all be linearly independent (as n linearly independent vectors always span $\mathbb{R}^n$).
However a matrix with linearly dependent column vectors is singular.

FIGURE 1. Images of the unit circle under $\mathbf{A}_a$,$\mathbf{A}_b$, and $\mathbf{A}_c$ (top graph) and under $\mathbf{A}_e$ (bootom graph)