True/False (10 points: 2 points each) For the problems below, circle T if the answer is true and circle F is the answer is false. After you've chosen your answer, mark the appropriate space on your Scantron form. Notice that letter A corresponds to true while letter B corresponds to false.

- (**T**) F If A is a  $3 \times 3$  matrix with three pivot positions, then for some  $p \in \mathbb{N}$  there exist elementary matrices  $E_1, E_2, ..., E_p \in \mathbb{R}^3$  such that  $E_p \cdots E_2 \cdot E_1 \cdot A = I_3$ .
- 2. (**T**) F Let  $m, n \in \mathbb{N}$  with m > n and suppose  $A \in \mathbb{R}^{m \times n}$  and  $\mathbf{b} \in \mathbb{R}^m$  are given. Then, a least-squares solution corresponding to matrix A and vector  $\mathbf{b}$  is a vector  $\mathbf{x}^* \in \mathbb{R}^n$  such that  $A\mathbf{x}^* \in \operatorname{Col}(A)$  and

$$\|\mathbf{b} - A\mathbf{x}^*\|_2 \le \|\mathbf{b} - A\mathbf{x}\|_2$$

for all  $\mathbf{x} \in \mathbb{R}^n$ .

T (F) Let  $A \in \mathbb{R}^{m \times n}$  and suppose  $\mathbf{x} \in \mathbb{R}^n$ . Suppose  $\mathbf{b} \in \mathbb{R}^m$  is nonzero. Suppose  $\mathbf{x}_1^*$  and  $\mathbf{x}_2^*$  are solutions to the inhomogeneous system  $A\mathbf{x} = \mathbf{b}$ . Then any linear combination  $c_1\mathbf{x}_1^* + c_2\mathbf{x}_2^*$  is a solution to the linear system  $A\mathbf{x} = \mathbf{b}$ .

4. T  $(\mathbf{F})$ 

1.

3.

Consider the linear systems problem

## $A\mathbf{x}=\mathbf{b}$

where matrix  $A \in \mathbb{R}^{m \times n}$  and vector  $\mathbf{b} \in \mathbb{R}^m$  are given and vector  $\mathbf{x} \in \mathbb{R}^n$  is unknown and desired. If this linear system is inconsistent, there may be an  $\mathbf{x} \in \mathbb{R}^n$  such that

$$\|\mathbf{b} - A\mathbf{x}\|_2 = 0.$$

5. T (**F**) If  $A \in \mathbb{R}^{3 \times 3}$ , then  $\det(5A) = 5 \det(A)$ 

Multiple Choice (60 points: 4 points each) For the problems below, circle the correct response for each question. After you've chosen your answer, mark your answer on your Scantron form.

6. Below we are given the LU Factorization of the matrix A and vector **b**:

$$A = \begin{bmatrix} 2 & -1 & 1 \\ 6 & -4 & 2 \\ 4 & 2 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 3 & 1 & 0 \\ 2 & -4 & 1 \end{bmatrix} \begin{bmatrix} 2 & -1 & 1 \\ 0 & -1 & -1 \\ 0 & 0 & -5 \end{bmatrix}, \qquad \mathbf{b} = \begin{bmatrix} 1 \\ 6 \\ 0 \end{bmatrix}$$

Use this LU Factorization to solve the linear system  $A\mathbf{x} = \mathbf{b}$  by solving the two linear systems

$$L\mathbf{y} = \mathbf{b} \qquad \text{and} \qquad U\mathbf{x} = \mathbf{y}$$
  
Then find  $\mathbf{y}^T \mathbf{x}$ :  
**A. -22** B. 22 C. -16 D. -18 E. 18

7. Let  $A \in \mathbb{R}^{m \times n}$  and set  $f(\mathbf{x}) = A\mathbf{x}$  for  $\mathbf{x} \in \mathbb{R}^n$ . Which of the following is not equivalent to  $\operatorname{Rng}(f)$ :

A. 
$$\operatorname{Col}(A)$$
 B.  $(\operatorname{Nul}(A^T))^{\perp}$  C.  $\operatorname{span}\{A(:,k)\}_{k=1}^n$ 

**D.** span{
$$A(i,:)$$
}<sup>m</sup><sub>i=1</sub> E. { $\mathbf{b} \in \mathbb{R}^m : \mathbf{b} = A\mathbf{x}$  for some  $\mathbf{x} \in \mathbb{R}^n$ }

8. Consider the following matrix:

Then

$$A = \begin{bmatrix} 1 & -1 & 0 & 0 \\ 2 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -1 & 1 & 0 & 1 \end{bmatrix}$$

Which of the following statements is true:

- A. The columns of A are linearly dependent.
- B. det(A) = -1
- C. dim  $(\operatorname{Nul}(A^T)) = 1$ .
- D. The matrix is not invertible.
- **E.**  $(A^T A)^{-1}$  exists

9. Let  $A \in \mathbb{R}^{4 \times 3}$  be defined by

$$A = \begin{bmatrix} 1 & 2 & 1 \\ 5 & -4 & 3h \\ 1 & 3 & h+5 \\ 3 & 7 & 4 \end{bmatrix}$$

For what value(s) of h is dim (Nul(A)) = 1:

A. rank(A) = 3 for all h. **B.** h = -3 C. h = -1 D. h = -1 and h = 1 E. h = 3

10. Let 
$$M = \begin{bmatrix} 1 & -1 & 2 \\ 0 & 1 & -5 \\ 0 & 0 & 1 \end{bmatrix}$$
. Then  $M^{-1}$  is given by which of the following:  
A.  $\begin{bmatrix} 1 & 1 & -2 \\ 0 & 1 & 5 \\ 0 & 0 & 1 \end{bmatrix}$  B.  $\begin{bmatrix} -1 & 1 & -2 \\ 0 & -1 & 5 \\ 0 & 0 & -1 \end{bmatrix}$  C.  $\begin{bmatrix} 1 & 1 & -3 \\ 0 & 1 & 5 \\ 0 & 0 & 1 \end{bmatrix}$  D.  $\begin{bmatrix} 1 & 1 & 3 \\ 0 & 1 & 5 \\ 0 & 0 & 1 \end{bmatrix}$  E.  $\begin{bmatrix} 1 & -1 & -2 \\ 0 & 1 & -5 \\ 0 & 0 & 1 \end{bmatrix}$ 

11. Stephen Curry plays for the Golden State Warriors as a professional basket ball player in the National Basketball Association (NBA). Suppose we analyze Curry's shooting technique. Below we see a diagram that highlights the typical trajectory of one of Mr. Curry's free-throw shots.



Our data points for this shooting trajectory  $\{(x_i, h_i)\}_{i=1}^3$  represent the observed height  $h_i$  of our basketball when the ball has moved  $x_i$  feet in the horizontal direction for i = 1, 2, 3. From our study of introductory physics, we choose to model the trajectory of the basket ball using the function

$$h(x) = a_0 + a_1 x + a_2 x^2.$$

Find the corresponding linear-systems problem to produce our desired polynomial model.

A.	$\begin{bmatrix} 1\\1\\1 \end{bmatrix}$	$\begin{array}{c} 0 \\ 8 \\ 14 \end{array}$	$\begin{bmatrix} 0\\ 64\\ 196 \end{bmatrix} \begin{bmatrix} a_0\\ a_1\\ a_2 \end{bmatrix} =$	$\begin{bmatrix} 7.25\\ 13.5\\ 10.0 \end{bmatrix}$		В.	$\begin{bmatrix} 0\\8\\14\end{bmatrix}$	$\begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} =$	$\begin{bmatrix} 7.25 \\ 13.50 \\ 10.00 \end{bmatrix}$			С.	$\begin{bmatrix} 1\\1\\1 \end{bmatrix}$	$7.25 \\ 13.50 \\ 10.00$	$\begin{array}{c} 52.5625\\ 182.2500\\ 100.0000 \end{array}$	$\begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix}$	=	$\begin{bmatrix} 0\\8\\14 \end{bmatrix}$
			D. $\begin{bmatrix} a_0 \\ a_0 \\ a_0 \end{bmatrix}$	$egin{array}{c} a_1 \ a_1 \ a_1 \end{array}$	$\begin{bmatrix} a_2\\a_2\\a_2\\a_2 \end{bmatrix} \begin{bmatrix} 0\\8\\14 \end{bmatrix}$	$=\begin{bmatrix}7\\13\\10\end{bmatrix}$	.25 3.50 ).00			E.	$\begin{bmatrix} 1\\ 1\\ 1\\ 1 \end{bmatrix}$	$\begin{array}{c} 0 \\ 8 \\ 14 \end{array}$	7.25 13.50 10.00	$\begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix}$	$= \begin{bmatrix} 0\\0\\0\end{bmatrix}$			

12. Solve the linear systems problem from Problem 11 above using any method. With your solution, determine how high the ball was in the air after traveling x = 5 feet in the horizontal direction. In other words, approximate h(5) using the solution of your linear systems problem. Round your answer to the nearest tenth (round to the nearest one digit to the RIGHT of the decimal place).

A. -23.7 ft B. 0 ft C. 12.6 ft D. 12.9 ft E. 11.9 ft

13. Suppose  $A \in \mathbb{R}^{n \times n}$ . Which of the following guarantees that  $\operatorname{Nul}(A) \neq \{\mathbf{0}\}$ ?

A. $\operatorname{rank}(A) = n$	B. $det(A) \neq 0$	C. $\operatorname{Col}(A) = \mathbb{R}^n$	D. $A^{-1}$ exists	E. $\dim($	$\left(\mathbf{Nul}\left(A^{T}\right)\right)$	) > 0
---------------------------------	--------------------	---	--------------------	------------	---	-------

For the next problem, consider the following spring-mass system



14. Suppose that you are apply masses 1, 2, and 3 to the mass-spring chain illustrated above such that

$$\mathbf{m} = \begin{bmatrix} m_1 \\ m_2 \\ m_3 \end{bmatrix} = \begin{bmatrix} 0.200 \\ 0.400 \\ 0.200 \end{bmatrix}$$

measured in kg. Assume the acceleration due to earth's gravity is  $g = 9.8m/s^2$ . Also assume that the mass of each spring is zero and that these springs satisfy the ideal version of Hooke's law. Then, which of the following gives the displacement vector

$$\mathbf{u} = \begin{bmatrix} u_1(T) \\ u_2(T) \\ u_3(T) \end{bmatrix}$$

measured in meters at t = T when the system is at equilibrium under the force of gravity on earth.

A.  $\begin{bmatrix} -0.196\\ -0.147\\ -0.196 \end{bmatrix}$ B.  $\begin{bmatrix} 0.098\\ 0.098\\ 0.049 \end{bmatrix}$ C.  $\begin{bmatrix} 0.196\\ 0.245\\ 0.196 \end{bmatrix}$ D.  $\begin{bmatrix} 0.020\\ 0.025\\ 0.020 \end{bmatrix}$ E.  $\begin{bmatrix} 0.098\\ 0.098\\ 0.098 \end{bmatrix}$ 

15. Let  $A \in \mathbb{R}^{5 \times 5}$  with det(A) = -4. Suppose

 $S_{14}(4) \cdot P_{14} \cdot D_3(1/8) \cdot P_{23} \cdot D_3(4) \cdot P_{12} \cdot B = A$ 

where we use standard notation for elementary matrices as discussed in class. Then det(B) is

A. 0	B. 2	C2	D. 8	E8

For Problems 16 - 18, assume that the matrix  $A \in \mathbb{R}^{4 \times 7}$  is given by

	[1	3	-2	0	0	0	1
Δ	2	6	-5	-2	-1	-2	1
$A \equiv$	0	0	5	10	5	10	5
	2	6	0	8	4	12	8

## 16. Find $\operatorname{RREF}(A)$ :

A.	$\begin{bmatrix} 1\\0\\0\\0\end{bmatrix}$	${3 \\ 0 \\ 0 \\ 0 \\ 0 }$	$\begin{array}{c} -2.5 \\ 1 \\ 0 \\ 0 \end{array}$	$\begin{array}{c} -1 \\ 2 \\ 0 \\ 0 \end{array}$	5 1 0 0	$\begin{array}{c} -1 \\ 2 \\ 1 \\ 0 \end{array}$	$\begin{bmatrix} 0.5 \\ 1.0 \\ 0.5 \\ 0 \end{bmatrix}$	B.	$\begin{bmatrix} 1\\0\\0\\0\end{bmatrix}$	3 0 0 0	0 1 0 0	$     \begin{array}{c}       4 \\       2 \\       0 \\       0     \end{array} $	$2 \\ 1 \\ 0 \\ 0$	$egin{array}{c} 0 \\ 0 \\ 1 \\ 0 \end{array}$	1. 0. 0. 0.	$egin{array}{c} 0 \\ 0 \\ 5 \\ 0 \end{array}$	С	$\cdot \begin{bmatrix} 1\\0\\0\\0 \end{bmatrix}$	0 0 0 0	$egin{array}{c} 0 \\ 1 \\ 0 \\ 0 \end{array}$	0 0 0 0	0 0 0 0	$egin{array}{c} 0 \\ 0 \\ 1 \\ 0 \end{array}$	$\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$
		D	$\begin{array}{c}1\\0\\0\\0\\0\end{array}$	$egin{array}{c} -3 \\ 0 \\ 0 \\ 0 \end{array}$	$\begin{array}{cccc} 0 & 4 \\ 1 & -2 \\ 0 & 0 \\ 0 & 0 \end{array}$	4 2 2 1 0 0 0 0	0 - 0 1 0	$\begin{bmatrix} -1.0 \\ 0.0 \\ 0.5 \\ 0.0 \end{bmatrix}$		]	E.	$\begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$	$egin{array}{c} 0 \\ 1 \\ 0 \\ 0 \end{array}$	$egin{array}{c} 0 \\ 0 \\ 1 \\ 0 \end{array}$	$     \begin{array}{c}       3 \\       0 \\       0 \\       0     \end{array} $	$     \begin{array}{c}       4 \\       2 \\       0 \\       0     \end{array} $	$\begin{array}{c} 2 \\ 1 \\ 0 \\ 0 \end{array}$	$ \begin{bmatrix} 1.0 \\ 0.0 \\ 0.5 \\ 0 \end{bmatrix} $						

17. How many linearly independent solutions are there to the homogeneous linear system  $A\mathbf{x} = \mathbf{0}$ :

A. 1	B. 3	<b>C.</b> 4	D. 5	E. 7

18. Which of the following is NOT a solution for the linear-systems problem  $A\mathbf{x} = \mathbf{0}$ ?

	[-3]	4	] [-2]	] [	2]	[1	1.0
	1	0	0		0	0	).0
	0	2	-1		0	0	).0
А.	0	B.  -1	C. 0	D.	0	<b>E</b> . (	).0
	0	0	1		0	0	).0
	0	0	0		1	0	).5
	0	0			-2	1	L.0

$\begin{array}{c} \text{Month } i \text{ in } 2015 \\ \text{(by Number)} \end{array}$	Cumulative Gallons $g_i$ of Gas used at end of Month	Odometer reading $r_i$ at end of month (in miles)
0 1 2 3	$0.000 \\ 41.490 \\ 98.804 \\ 143.622$	86,286.2 87,240.3 88,500.6 89,432.7

19. Below is a data set for of the gasoline usage from 1/1/2015 at 12am to 3/31/2015 at 11:59pm:

We will model this data using a piecewise function

$$R(g) = \begin{cases} R_1(g) & \text{if} \quad 0.000 < g \le 41.490\\ R_2(g) & \text{if} \quad 41.490 < g \le 98.804\\ R_3(g) & \text{if} \quad 98.804 < g \le 143.622 \end{cases}$$

where output R gives the odometer reading of the car (in miles) as a function of the number of gallons g used during the year. Each segment of the piecewise linear function given by

$$R_i(g) = m_i(g - g_i) + r_i,$$
 with  $R(g_{i+1}) = r_{i+1}$ 

Set up a  $3 \times 3$  linear-systems problem  $G \mathbf{m} = \mathbf{r}$  to find the average miles per gallon efficiency of this car during this time period. The diagonal matrix  $G \in \mathbb{R}^{3\times 3}$  and vector  $\mathbf{r} \in \mathbb{R}^3$  are constructed from our data set. The vector  $\mathbf{m}$  stores the unknown slopes  $m_i$  of the three line segments connecting these data points. Remember, for all i = 0, 1, 2, we have

$$(r_{i+1} - r_i) = m_i \cdot (g_{i+1} - g_i)$$

Which of the following gives the linear-systems problem associated with this model?



20. Find average miles per gallon (mpg) efficiency of this car during month 2 of this data set. Round your answer to the integer (one digit to the left of the decimal place).

A. 21	B. 23	C. 896	D. 0	E. 22

## Free Response

21. Let  $n, i, k \in \mathbb{N}$  such that  $1 \leq i \leq n, 1 \leq k \leq n$  and  $i \neq k$ . Suppose that  $c \in \mathbb{R}$ .

(a) Show 
$$(S_{ik}(c))^{-1} = S_{ik}(-c)$$

**Solution:** Suppose  $i, k, n \in \mathbb{N}$  such that  $1 \leq i, k \leq n$  with  $i \neq k$  and suppose  $c \in \mathbb{R}$  is a nonzero constant. Let's begin by proving that  $(S_{ik}(c))^{-1} = S_{ik}(-c)$  by multiplying  $S_{ik}(c)$  by the stated inverse to produce  $I_n$ . Consider

$$S_{ik}(c) \cdot S_{ik}(-c) = \left(I_n + c \,\mathbf{e}_i \,\mathbf{e}_k^T\right) \cdot \left(I_n - c \,\mathbf{e}_i \,\mathbf{e}_k^T\right)$$
$$= I_n - c \,\mathbf{e}_i \,\mathbf{e}_k^T + c \,\mathbf{e}_i \,\mathbf{e}_k^T \cdot I_n - c^2 \,\left(\mathbf{e}_i \,\mathbf{e}_k^T\right) \cdot \left(\mathbf{e}_i \,\mathbf{e}_k^T\right)$$
$$= I_n - c \,\mathbf{e}_i \,\mathbf{e}_k^T + c \,\mathbf{e}_i \,\mathbf{e}_k^T - c^2 \,\mathbf{e}_i \,\left(\mathbf{e}_k^T \,\mathbf{e}_i\right) \,\mathbf{e}_k^T$$

Notice that the matrix-matrix product  $\mathbf{e}_k^T \mathbf{e}_i$  results in a scalar output equivalent to the inner product  $\mathbf{e}_k \cdot \mathbf{e}_i$ . If  $j \in \mathbb{N}$  with  $1 \leq j \leq n$  we know  $\mathbf{e}_j \in \mathbb{R}^n$  is the *j*th elementary basis vector with all zero entries except the *j*th coefficient, which has value equal to one. Because of this structure and since  $i \neq k$  by assumption, we see that  $\mathbf{e}_k^T \mathbf{e}_i = \mathbf{e}_k \cdot \mathbf{e}_i = 0$ . With this we have,

$$S_{ik}(c) \cdot S_{ik}(-c) = I_n - c \mathbf{e}_i \mathbf{e}_k^T + c \mathbf{e}_i \mathbf{e}_k^T = I_n$$

We conclude that  $(S_{ik}(c))^{-1} = S_{ik}(-c).$ 

(b) Show  $\left(D_i(c)\right)^{-1} = D_i\left(\frac{1}{c}\right)$ 

Solution: Let's establish that  $\left(D_i(c)\right)^{-1} = D_i\left(\frac{1}{c}\right)$ . To this end, consider  $D_i(c) \cdot D_i\left(1/c\right) = \left(I_n + (c-1)\mathbf{e}_i\mathbf{e}_i^T\right) \cdot \left(I_n + (1/c-1)\mathbf{e}_i\mathbf{e}_i^T\right)$   $= I_n \cdot \left(I_n + (1/c-1)\mathbf{e}_i\mathbf{e}_i^T\right) + (c-1)\mathbf{e}_i\mathbf{e}_i^T \cdot \left(I_n + (1/c-1)\mathbf{e}_i\mathbf{e}_i^T\right)$   $= I_n + (1/c-1)\mathbf{e}_i\mathbf{e}_i^T + (c-1)\mathbf{e}_i\mathbf{e}_i^T + (c-1)(1/c-1)\mathbf{e}_i\mathbf{e}_i^T\mathbf{e}_i\mathbf{e}_i^T$  $= I_n + (1/c+c-2)\mathbf{e}_i\mathbf{e}_i^T + (c-1)(1/c-1)\mathbf{e}_i\mathbf{e}_i^T$ 

Using distributivity of scalar multiplication over addition, we see

$$(c-1) (1/c-1) = 2 - c - 1/c.$$

Thus, we have

$$D_i(c) \cdot D_i(1/c) = I_n + (1/c + c - 2) \mathbf{e}_i \mathbf{e}_i^T + (2 - c - 1/c) \mathbf{e}_i \mathbf{e}_i^T = I_n$$

By definition of the matrix inverse, we have  $\left(D_i(c)\right)^{-1} = D_i\left(\frac{1}{c}\right)$ 

22. (10 pts) A small bike company selling utility bicycles for daily commuting has been in business for four years. This company has recorded annual sales (in tens of thousands of dollars) as follows:



This data is plotted in the figure next to the table above. Although the data do not exactly lie on a straight line, we can create a linear model to fit this data.

a. (4 points) Set up the least squares problem to fit this data to a linear model.

**Solution:** Recall that the least squares problem is designed to fit data collected during an experiment to a particular mathematical model. In this problem, we want to fit our given data points  $(t_i, S_i)$  to a linear model

$$S(t) = a_0 + a_1 \cdot t$$

In this case we say that  $t_i$  measures the year of business and  $S_i$  is the Sales Revenue for our company. We can create a system of equations that attempts to model the collected input date  $t_i$  with an unknown model output  $S(t_i)$  as follows:

$$S(t_i) = a_0 + a_1 \cdot t_i \approx S_i$$

for i = 1, ..., 4.

The difference between the observed data and the model prediction is known as the model error in the *i*th term, given by:

$$e_i = (S_i - S(t_i)) = \left(S_i - (a_0 + a_1 t_i)\right)$$

This is a quantitative measurement of the error between our model and the collected data. For our least-squares problem, we want to minimize the total squared error for each data point. In other words, we want to find parameters  $a_0, a_1$  so minimize the function

$$\sum_{i=1}^{4} e_i^2 = \left( S_i - (a_0 + a_1 t_i) \right)^2 = \|\mathbf{b} - A\mathbf{x}\|_2^2 \tag{1}$$

where we define

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \end{bmatrix}, \qquad \mathbf{x} = \begin{bmatrix} a_0 \\ a_1 \end{bmatrix}, \qquad \mathbf{b} = \begin{bmatrix} 23 \\ 27 \\ 30 \\ 34 \end{bmatrix}$$

Then, we have translated our given data into an equivalent linear model in matrix form.

**Solution:** From here, we can solve the corresponding least-squares problem. When attempting to solve this least squares problem, we want to find a vector  $\mathbf{x}^* \in \mathbb{R}^2$  that solves the optimization problem

$$\min_{\mathbf{x}\in\mathbb{R}^2} \|\mathbf{b} - A\mathbf{x}\|_2^2$$

This optimal  $\mathbf{x}^*$  will define parameters  $a_0, a_1$  for the linear polynomial of "best fit" for this data set.

Remark:

The reason that we choose to measure total error using the sum-of-squared errors equation (1) of the individual error terms relates directly to our knowledge of multivariable calculus and to the inner product formula for the two-norm of a vector.

- We recall that when optimizing a multivariable function using calculus, we apply the second derivative test. To do so, we need to identify critical points (where the gradient is zero) and then do some analysis at those points. In general, the derivatives of power functions  $x^n$  for some  $n \in N$  are very straight forward. Thus, when we establish the squared-error  $e_i^2$ , we set ourselves up well to optimize this function using technique from calculus. While this does provide a solution mechanism, the algebra and arithmetic of optimization are prohibitively time consuming.
- Recall that for any  $\mathbf{r} \in \mathbb{R}^m$ , we can use the inner-product formula for the 2-norm to write

$$\|\mathbf{r}\|_2^2 = \mathbf{r} \cdot \mathbf{r}.$$

This is extremely helpful because the inner product operation has a geometric interpretation (the cosine formula for the inner product). Thus, we can translate a minimization problem into a geometry problem. This technique helps us to avoid the tedium of calculus based optimization. Moreover, because we can model our minimization problem using matrices, we have translated our minimization problem into a matrix analysis problem for which we have lots of powerful tools. This is the other reason we use the squares error terms to solve mathematical modeling problems.

- Students who want to pursue greater depths about this choice of squared errors: Look into the difference between a Banach Space and a Hilbert Space. This is very much related to the choice of squared errors in the least-squares problem.

b. (4 points) Solve the least squares problem using the normal equations to find the line of best fit.

Solution: We want to solve the least-squares problem

$$\min_{\mathbf{x}\in\mathbb{R}^2}\|\mathbf{b}-A\mathbf{x}\|_2^2$$

for the Vandermonde matrix  $A \in \mathbb{R}^{4 \times 2}$  and  $\mathbf{b} \in \mathbb{R}^4$  described in the problem statement. We know that the theoretic solution to this problem is to project  $\mathbf{b}$  onto the orthogonal complement of  $\operatorname{Col}(A)$ . In other words, we want to choose  $\mathbf{x} \in \mathbb{R}^2$  such that the residual vector

$$\mathbf{r} = \mathbf{b} - A\mathbf{x}$$

is orthogonal to the column space of A. Thus, we want to choose  $\mathbf{x}$  such that

 $\mathbf{r} \in \left[ \operatorname{Col}(A) \right]^{\perp} \implies \mathbf{r} \in \operatorname{Nul}(A^{T}) \quad \text{by Thm 3 p. 335}$   $\implies A^{T}\mathbf{r} = \mathbf{0} \qquad \text{by definition of Null Space}$   $\implies A^{T}(\mathbf{b} - A\mathbf{x}) = \mathbf{0} \qquad \text{by definition of } \mathbf{r}$   $\implies A^{T}\mathbf{b} - A^{T}A\mathbf{x} = \mathbf{0} \qquad \text{by distributivity of matrix mult}$   $\implies A^{T}A\mathbf{x} = A^{T}\mathbf{b}$ 

By Theorem 14 on p. 363, we know that  $A^T A$  is invertible if and only if  $\operatorname{rank}(A) = 2$ . We see that our A in this problem has three pivot columns and thus the Gram matrix  $A^T A$  is invertible. Then, to solve our least-squares problem, we will solve the normal equation associated with this matrix model given by

$$A^T A \mathbf{x} = A^T \mathbf{b}$$

where we calculate each of these matrices using the technology of our choice (i.e. a TI Calculator). We use the matrices that we generated in part (a) to find

$$A^{T}A = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \end{bmatrix} = \begin{bmatrix} 4 & 10 \\ 10 & 40 \end{bmatrix}, \qquad A^{T}\mathbf{b} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 \end{bmatrix} \begin{bmatrix} 23 \\ 27 \\ 30 \\ 34 \end{bmatrix} = \begin{bmatrix} 114 \\ 303 \end{bmatrix}$$

Now we will solve the equivalent linear system problem

$$\begin{bmatrix} 4 & 10 \\ 10 & 40 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \end{bmatrix} = \begin{bmatrix} 114 \\ 303 \end{bmatrix}$$

Using Cramer's rule for the inverse of a  $2 \times 2$  system, we see that

$$\mathbf{x}^* = \begin{bmatrix} a_0\\a_1 \end{bmatrix} = \frac{1}{10} \begin{bmatrix} 40 & -10\\-10 & 4 \end{bmatrix} \begin{bmatrix} 114\\303 \end{bmatrix} = \begin{bmatrix} 19.5\\3.6 \end{bmatrix}$$

Thus, our line of best fit that models this data is given by

$$S(t) = 19.5 + 3.6 \cdot t$$

c. (2 points) Use your linear model to estimate the sales for year five. Your solution should be include units and be in a full sentence.

Solution: Using our line of best fit

$$S(t) = 19.5 + 3.6 \cdot t$$

we can predict the sales in year t = 5 as

$$S(5) = 19.5 + 3.6 \cdot 5 = 37.5$$

which is \$375,000.

23. Recall that Cramer's formula for the inverse of a  $2 \times 2$  matrix  $A \in \mathbb{R}^2$  is given by

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}^{-1} = \frac{1}{\det(A)} \begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix}$$

where  $\det(A) = a_{11}a_{22} - a_{12}a_{21}$ .

(a) (5 pts) Using a sequence of elementary matrices, transform A into  $I_2$ . Show each matrix you use.

**Solution:** Suppose  $A \in \mathbb{R}^{2 \times 2}$  is nonsingular. Let

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

In order to produce the inverse of A, let's transform A into  $I_2$  using elementary matrices. To this end, let  $\delta = a_{11}a_{22} - a_{21}a_{12}$ . As we will see in Chapter 7, we call  $\delta$  the determinant of A and  $\delta \neq 0$  if and only if A is invertible. As we will see in the next section, since A is invertible, we know that either  $a_{11} \neq 0$  or  $a_{21} \neq 0$ . Thus, let's assume that  $a_{11} \neq 0$ . Now, we reduce A to  $I_2$ . Consider the step-by-step calculation beginning with the product

$$S_{21}\left(-a_{21}/a_{11}\right) \cdot A = \begin{bmatrix} 1 & 0\\ -a_{21}/a_{11} & 1 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12}\\ a_{21} & a_{22} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12}\\ 0 & \delta/a_{11} \end{bmatrix}$$

Now we consider

$$D_2(a_{11}/\delta) \cdot D_1(1/a_{11}) = \begin{bmatrix} 1/a_{11} & 0\\ 0 & a_{11}/\delta \end{bmatrix}$$

We can multiply our product by this diagonal matrix to find

$$\begin{bmatrix} 1/a_{11} & 0\\ 0 & a_{11}/\delta \end{bmatrix} \begin{bmatrix} a_{11} & a_{12}\\ 0 & \delta/a_{11} \end{bmatrix} = \begin{bmatrix} 1 & a_{12}/a_{11}\\ 0 & 1 \end{bmatrix}$$

Finally, multiplying this entire product on the left-hand side by we see

$$\begin{bmatrix} 1 & -a_{12}/a_{11} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & a_{12}/a_{11} \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

(b) (5 pts) Write  $A^{-1}$  as a product of the elementary matrices and confirm Cramer's Rule.

Solution: From our work in part (a) above, we see

 $A^{-1} = S_{12} \left( -a_{12}/a_{11} \right) \cdot D_2 \left( a_{11}/\delta \right) \cdot D_1 \left( 1/a_{11} \right) \cdot S_{21} \left( -a_{21}/a_{11} \right).$ 

The reduction of matrix A into  $I_2$  above gives a constructive mechanism to explicitly calculate  $A^{-1}$ . We begin by finding the product

$$\left(D_2\left(a_{11}/\delta\right) \cdot D_1\left(1/a_{11}\right)\right) \cdot S_{21}\left(-a_{21}/a_{11}\right)$$

given as

$$\begin{bmatrix} 1/a_{11} & 0\\ 0 & a_{11}/\delta \end{bmatrix} \begin{bmatrix} 1 & 0\\ -a_{21}/a_{11} & 1 \end{bmatrix} = \begin{bmatrix} 1/a_{11} & 0\\ -a_{21}/\delta & a_{11}/\delta \end{bmatrix}$$

We can calculate  $A^{-1}$  using the product

$$\begin{bmatrix} 1 & -\frac{a_{12}}{a_{11}} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{a_{11}} & 0 \\ -\frac{a_{21}}{\delta} & \frac{a_{11}}{\delta} \end{bmatrix} = \begin{bmatrix} \frac{1}{a_{11}} + \frac{a_{21}a_{12}}{a_{11}\delta} & -\frac{a_{12}}{a_{11}} \\ -\frac{a_{21}}{\delta} & \frac{a_{11}}{\delta} \end{bmatrix} = \frac{1}{\delta} \begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix}$$

This is exactly what we wanted to show.