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.

1. T (F) If $A \in \mathbb{R}^{3\times 3}$ has three pivot columns, then it is possible to find invertible matrices $E_1, E_2, ..., E_p \in \mathbb{R}^{3\times 3}$ such that

$$E_p E_{p-1} \cdots E_2 E_1 A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

2. T (F) Suppose $A \in \mathbb{R}^{m \times n}$ and $B = \operatorname{RREF}(A)$. Then $\operatorname{Col}(A) = \operatorname{Col}(B)$.

3. (**T**) F Suppose $A \in \mathbb{R}^{m \times n}$ with $m, n \in \mathbb{N}$ and $m \neq n$. The spaces Nul(A) and Col(A) never share an element in common.

4. (**T**) F If $A \in \mathbb{R}^{m \times n}$, then $\operatorname{Col}(A) = \mathbb{R}^m$ if and only if $\operatorname{rank}(A) = m$

5. T (F) Suppose $A, B \in \mathbb{R}^{2 \times 2}$. If $\det(A) = 2$ and $\det(B) = 3$, then $\det(A + B) = 5$.

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. Problems that are marked "choose all that apply" may have more than one correct answer. In this case, mark all correct answers.

6. Suppose we are given a matrix $A = \begin{bmatrix} 2 & 1 & 1 \\ 4 & 5 & -2 \\ 2 & -2 & 0 \end{bmatrix}$.

Find the matrix $L \in \mathbb{R}^{3 \times 3}$ from the LU factorization of A.

| А. | $\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$ | $\begin{array}{c} 0 \\ 1 \\ -1 \end{array}$ | $\begin{pmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$ | В. | $\begin{bmatrix} 1\\ -2\\ -1 \end{bmatrix}$ | $\begin{array}{c} 0 \\ 1 \\ 1 \end{array}$ | $\begin{bmatrix} 0\\0\\1 \end{bmatrix}$ | С. | $\begin{bmatrix} 1\\ 2\\ 1 \end{bmatrix}$ | $\begin{array}{c} 0 \\ 1 \\ 1 \end{array}$ | $\begin{bmatrix} 0\\0\\1 \end{bmatrix}$ | D. | $\begin{bmatrix} 1 \\ 4 \\ 2 \end{bmatrix}$ | $\begin{array}{c} 0 \\ 1 \\ -2 \end{array}$ | $\begin{bmatrix} 0\\0\\1 \end{bmatrix}$ | E. | $\begin{bmatrix} 1\\ -2\\ -1 \end{bmatrix}$ | $0 \\ 1 \\ -1$ | $\begin{pmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$ | |
|----|---------------------------------------------|---------------------------------------------|---------------------------------------------|----|---------------------------------------------|--------------------------------------------|-----------------------------------------|----|-------------------------------------------|--------------------------------------------|-----------------------------------------|----|---------------------------------------------|---------------------------------------------|-----------------------------------------|----|---------------------------------------------|----------------|---------------------------------------------|--|
|----|---------------------------------------------|---------------------------------------------|---------------------------------------------|----|---------------------------------------------|--------------------------------------------|-----------------------------------------|----|-------------------------------------------|--------------------------------------------|-----------------------------------------|----|---------------------------------------------|---------------------------------------------|-----------------------------------------|----|---------------------------------------------|----------------|---------------------------------------------|--|

7. Suppose that $U \in \mathbb{R}^{3 \times 3}$ is the upper triangular matrix from the LU factorization of matrix

$$A = \begin{bmatrix} 2 & 1 & 1 \\ 4 & 5 & -2 \\ 2 & -2 & 0 \end{bmatrix}$$

in problem 6 above. What do you know about the product of the diagonal elements of U given by $u_{11}u_{22}u_{33}$? Choose all that apply.

A.
$$det(A) = u_{11}u_{22}u_{33}$$

B. $u_{11}u_{22}u_{33} = 0$
C. $u_{11}u_{22}u_{33} = 1$
D. $u_{11}u_{22}u_{33} = -30$
E. $u_{11}u_{22}u_{33} = 30$

8. The LU Factorization of a given 3×3 matrix is $A = \begin{bmatrix} 2 & 6 & 1 \\ 0 & 2 & 1 \\ 1 & 1 & 4 \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0.5 & -1 & 1 \end{bmatrix}}_{L} \underbrace{\begin{bmatrix} 2 & 6 & 1 \\ 0 & 2 & 1 \\ 0 & 0 & 4.5 \end{bmatrix}}_{U}.$

Use the LU factorization of ${\cal A}$ combined with forward and backward substitution to find the solution to the linear systems problem

$$\begin{bmatrix} 2 & 6 & 1 \\ 0 & 2 & 1 \\ 1 & 1 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ -9 \end{bmatrix}$$

Which of the following gives $-x_1 + x_2 + x_3$?

9. Let $A = \begin{bmatrix} 3 & 5 \\ 1 & 2 \end{bmatrix}$. Which of the following are false? Choose all that apply.

A.
$$A^{-1} = \begin{bmatrix} 2 & -5 \\ -1 & 3 \end{bmatrix}$$
. **B.** $\det(A) = -1$ **C.** $\operatorname{Nul}(A^T) \neq \emptyset$ D. $\operatorname{rank}(A) = 2$ E. $\operatorname{Col}(A) = \mathbb{R}^2$

10. Let $n \in \mathbb{N}$. Recall that the set of polynomials of degree less than or equal to n is denoted as

$$P_n = \{p(x) : p(x) = a_0 + a_1 x + \dots + a_n x^n \text{ with } a_i \in \mathbb{R} \text{ for all } i = 0, 1, \dots, n\}$$

Which of the following is false?

- A. $P_1 \subseteq P_2$
- B. P_k is a subspace of P_n for all $0 \le k \le n$
- C. P_n is a vectors space.
- D. The set of constant functions $\{p(x) : p(x) = c \text{ for } c \in \mathbb{R}\}$ is a subspace of P_n
- E. The set of linear polynomials with nonzero slope is a subspace of P_n

| 11. Let $A = \begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix}$ | $\begin{bmatrix} 0 & -2 \\ 1 & 1 \\ 1 & -t \end{bmatrix}$. Find the | set of values for which $Nul(A)$ | $) \neq \{0\}:$ | |
|----------------------------------------------------------|----------------------------------------------------------------------|----------------------------------|-----------------|----------------|
| A. $t = 3$ | B. $t = -3$ | C. $t = 3$ or $t = -3$ | D. $t \neq 3$ | E. $t \neq -3$ |

- 12. Suppose $A \in \mathbb{R}^{m \times n}$. Given a vector $\mathbf{b} \in \mathbb{R}^m$, suppose that you know:
 - I. Vectors $\mathbf{z}_1, \mathbf{z}_2 \in \mathbb{R}^n$ solve the linear system problem $A\mathbf{x} = \mathbf{0}$
 - II. Vectors $\mathbf{x}^*, \mathbf{y}^* \in \mathbb{R}^n$ solve the linear system problem $A\mathbf{x} = \mathbf{b}$.

Which of the following is NOT a solution for the linear system problem $A\mathbf{x} = \mathbf{b}$?

A. $\mathbf{x}^* + \mathbf{z}_1$ B. $\mathbf{y}^* + \mathbf{z}_2$ C. $\mathbf{x}^* + \mathbf{y}^*$ D. $3\mathbf{z}_1 + \mathbf{x}^* - 4\mathbf{z}_2$ E. $2\mathbf{x}^* - \mathbf{y}^*$

13. Let
$$B = \overbrace{\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 7 & 1 \end{bmatrix}}^{E_2} \cdot \overbrace{\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & -5 & 0 & 1 \end{bmatrix}}^{E_1} \cdot \overbrace{\begin{bmatrix} 1 & 0 & 0 & 0 \\ 3 & 1 & 0 & 0 \\ -2 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix}}^{E_1}$$
 Find B^{-1} :
A. $\begin{bmatrix} 1 & 0 & 0 & 0 \\ 3 & 1 & 0 & 0 \\ -2 & 1 & 1 & 0 \\ 1 & -5 & 7 & 1 \end{bmatrix}$ B. $\begin{bmatrix} 1 & 0 & 0 & 0 \\ -3 & 1 & 0 & 0 \\ 2 & -1 & 1 & 0 \\ -1 & 5 & -7 & 1 \end{bmatrix}$ C. $\begin{bmatrix} 1 & -3 & 2 & -1 \\ 0 & 1 & -1 & 5 \\ 0 & 0 & 1 & -7 \\ 0 & 0 & 0 & 1 \end{bmatrix}$
D. $E_3^{-1} \cdot E_2^{-1} \cdot E_1^{-1}$ E. $\begin{bmatrix} 1 & 3 & -2 & 1 \\ 0 & 1 & 1 & -5 \\ 0 & 0 & 1 & 7 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

14. Ohm's Law, governing the electrical behavior of resistors, states that the voltage access a resistor depends linearly on the current running through the resistor. Below is some current and voltage data collected from an experiment:



We can model this relationship using the linear function $v(i) = r \cdot i + b$ where the positive constant r measures the resistance value of resistor $(k\Omega)$, i represents the current through the resistor (mA), v measures voltage across resistor (V), and b is the intercept of this model with the vertical axis. Solve the least-square problem associated with this model and identify the line of best fit below:

A.
$$v = 1.0048 \cdot i + 0.0357$$

B. $v = 0.0135 \cdot i + 0.9845$
C. $v = 0.9845 \cdot i + 0.0135$
D. $v = 0.0357 \cdot i + 1.0048$
E. $v = 0.9827 \cdot i + 0.0253$

15. Suppose $\mathbf{y}, \mathbf{b} \in \mathbb{R}^3$ are given by

$$\mathbf{y} = \begin{bmatrix} -2\\1\\0 \end{bmatrix}, \qquad \qquad \mathbf{b} = \begin{bmatrix} 1\\0\\1 \end{bmatrix}$$

Let $Y = \text{Span}\{\mathbf{y}\}$ and let $Y^{\perp} = (\text{Span}\{\mathbf{y}\})^{\perp}$. Suppose that

$$\begin{aligned} \alpha \, \mathbf{y} &= \operatorname{Proj}_{Y}(\mathbf{b}) = \text{ the projection of } \mathbf{b} \text{ onto } Y, \\ \mathbf{r} &= \operatorname{Proj}_{Y^{\perp}}(\mathbf{b}) = \text{ the projection of } \mathbf{b} \text{ onto } Y^{\perp} \end{aligned}$$

Which of the following statements are true? Choose all that apply.

A.
$$\alpha \mathbf{y} = \begin{bmatrix} 0.8\\ -0.4\\ 0 \end{bmatrix}$$
 B. $\alpha \mathbf{y} = \begin{bmatrix} -1.0\\ 0.0\\ -1.0 \end{bmatrix}$ **C.** $\mathbf{r} = \begin{bmatrix} 0.2\\ 0.4\\ 1.0 \end{bmatrix}$ D. $\mathbf{r} = \begin{bmatrix} -1\\ 1\\ 1 \end{bmatrix}$ **E.** $\mathbf{y}^T \mathbf{r} = \mathbf{0}$

For Problems 16 - 20, assume that the matrix $A \in \mathbb{R}^{4 \times 6}$ is given by

$$A = \begin{bmatrix} 1 & 2 & -5 & -2 & 6 & 14 \\ 0 & 0 & -2 & -2 & 7 & 12 \\ 2 & 4 & -5 & 1 & -5 & -1 \\ 0 & 0 & 4 & 4 & -14 & -24 \end{bmatrix}$$

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

A.
$$\begin{bmatrix} 1 & 2 & -5 & -2 & 6 & 14 \\ 2 & 4 & -5 & 1 & -5 & -1 \\ 0 & 0 & -2 & -2 & 7 & 12 \\ 0 & 0 & 4 & 4 & -14 & -24 \end{bmatrix}$$
 B.
$$\begin{bmatrix} 1 & 0 & 0 & 2 & 3 & 7 \\ 0 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 0 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$
 C.
$$\begin{bmatrix} 1 & 2 & -2.5 & 0.5 & -2.5 & -0.5 \\ 0 & 0 & 1 & 1 & -3.5 & -6 \\ 0 & 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

D.
$$\begin{bmatrix} 1 & 2 & 0 & 3 & 0 & 7 \\ 0 & 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$
 E.
$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

17. Which of the following vectors in NOT a solution to $A\mathbf{x} = \mathbf{0}$?

| A. $\begin{bmatrix} -2\\1\\0\\0\\0\\0\end{bmatrix}$ B. $\begin{bmatrix} 3\\0\\1\\-1\\0\\0\end{bmatrix}$ | C. $\begin{bmatrix} 7\\0\\1\\0\\2\\-1 \end{bmatrix}$ | D. $\begin{bmatrix} -4\\2\\0\\0\\0\\0\\0\end{bmatrix} + \begin{bmatrix} 9\\0\\3\\-3\\0\\0\end{bmatrix}$ | $\mathbf{E.} \begin{bmatrix} 6\\0\\2\\-2\\0\\0\\0 \end{bmatrix} + \begin{bmatrix} -7\\0\\-1\\0\\-2\\-1 \end{bmatrix}$ |
|---------------------------------------------------------------------------------------------------------|------------------------------------------------------|---------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------|
|---------------------------------------------------------------------------------------------------------|------------------------------------------------------|---------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------|

18. Which of the following sets of vectors are linearly dependent? Choose all that apply.

A. $\{A(:,1), A(:,3), A(:5)\}$ D. $\{A(:,1), A(:,4), A(:5)\}$ E. $\{A(:,2), A(:,3), A(:6)\}$ E. $\{A(:,2), A(:,4), A(:6)\}$

19. Find dim(Nul(A)) + dim(Nul(A^T)) : A. 1 B. 2 C. 3 D. 4 E. 5 20. Which of the following must be true? Choose all that apply. A. rank $(A^T) = 3$ B. Col $(A^T) \subseteq \mathbb{R}^6$ C. $(AA^T)^{-1}$ exists D. $(A^TA)^{-1}$ exists E. Col(A) = \mathbb{R}^3

Free Response

21. (a) (4 pts) Suppose we are given vectors $\mathbf{y}, \mathbf{b} \in \mathbb{R}^m$. Let $Y = \text{Span}\{\mathbf{y}\}$ and $Y^{\perp} = \left(\text{Span}\{\mathbf{y}\}\right)^{\perp}$. Show how to construct the projections

$$\alpha \mathbf{y} = \operatorname{Proj}_{Y}(\mathbf{b})$$
 and $\mathbf{r} = \operatorname{Proj}_{Y^{\perp}}(\mathbf{b})$

by finding an explicit formula for the scalar α . Explain your assumptions and draw a diagram to support your work.

Solution: Let $\mathbf{b}, \mathbf{y} \in \mathbb{R}^m$ be linearly independent vectors. Recall that in order to project $\mathbf{r} = \mathbf{b} - \alpha \mathbf{y}$ onto Y^{\perp} , we need to choose $\alpha \in \mathbb{R}$ such \mathbf{r} is orthogonal to \mathbf{y} . to this end, we want to choose α such that

| | $\mathbf{y} \cdot \mathbf{r} = 0$ |
|---------------|-----------------------------------------------------------------------------------------------------------------------------|
| \Rightarrow | $\mathbf{y}^T \mathbf{r} = 0$ |
| \Rightarrow | $\mathbf{y}^T \left(\mathbf{b} - \alpha \mathbf{y} \right) = 0$ |
| \Rightarrow | $\mathbf{y}^T \mathbf{b} - \alpha \mathbf{y}^T \mathbf{y} = 0$ |
| \Rightarrow | $\mathbf{y}^T \mathbf{b} = \alpha \mathbf{y}^T \mathbf{y}$ |
| \Rightarrow | $\alpha = \frac{\mathbf{y}^T \mathbf{b}}{\mathbf{y}^T \mathbf{y}} = \frac{\mathbf{y} \cdot \mathbf{b}}{\ \mathbf{y}\ _2^2}$ |

(b) (6 pts) Use classical Gram-Schmidt to find an orthogonal basis for $\operatorname{Col}(A)$ for $A = \begin{bmatrix} 1 & -1 & 2 \\ 1 & 0 & -1 \\ -1 & 1 & 2 \\ 0 & 1 & 1 \end{bmatrix}$

Solution: To construct an orthogonal basis $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ for $\operatorname{Col}(A)$ we will use the Classical Gram-Schmidt Algorithm. To this end, let's initialize our basis by setting

$$\mathbf{v}_1 = A(:,1) = \begin{bmatrix} 1\\ 1\\ -1\\ 0 \end{bmatrix}$$

Now, to create our second basis vector \mathbf{v}_2 , we want to project A(:,2) onto $\left[\operatorname{Span} \{\mathbf{v}_1\}\right]^{\perp}$. To do so, we use our definition of α from part (a) above and define

$$\mathbf{v}_{2} = \left(I_{4} - \frac{\mathbf{v}_{1} \, \mathbf{v}_{1}^{T}}{\|\mathbf{v}_{1}\|_{2}^{2}}\right) A(:, 2)$$
$$= A(:, 2) - \frac{\mathbf{v}_{1} \, \mathbf{v}_{1}^{T}}{\|\mathbf{v}_{1}\|_{2}^{2}} \cdot A(:, 2)$$

$$= A(:,2) - \frac{\mathbf{v}_1^T A(:,2)}{\|\mathbf{v}_1\|_2^2} \cdot \mathbf{v}_1$$

We can calculate the dot product and norm from this equation separately as follows:

$$\mathbf{v}_1 \cdot A(:,2) = \mathbf{v}_1^T \cdot A(:,2) = \begin{bmatrix} 1 & 1 & -1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ -1 \\ 0 \end{bmatrix} = -2, \qquad \|\mathbf{v}_1\|_2^2 = 3$$

With this we have

$$\mathbf{v}_{2} = \begin{bmatrix} -1\\0\\1\\1 \end{bmatrix} + \frac{2}{3} \begin{bmatrix} 1\\1\\-1\\0 \end{bmatrix} = \begin{bmatrix} -1/3\\2/3\\1/3\\1 \end{bmatrix} = \frac{1}{3} \begin{bmatrix} -1\\2\\1\\3 \end{bmatrix}$$

We calculate the third basis vector \mathbf{v}_3 by projecting A(:,3) on $\left[\text{Span } \{\mathbf{v}_1, \mathbf{v}_2\} \right]^{\perp}$. We accomplish this projection on step at a time

$$\begin{split} \mathbf{v}_{3} &= \left(I_{4} - \frac{\mathbf{v}_{1} \, \mathbf{v}_{1}^{T}}{\|\mathbf{v}_{1}\|_{2}^{2}} - \frac{\mathbf{v}_{2} \, \mathbf{v}_{2}^{T}}{\|\mathbf{v}_{2}\|_{2}^{2}}\right) A(:,3) \\ &= A(:,3) - \frac{\mathbf{v}_{1} \, \mathbf{v}_{1}^{T}}{\|\mathbf{v}_{1}\|_{2}^{2}} \cdot A(:,3) - \frac{\mathbf{v}_{2} \, \mathbf{v}_{2}^{T}}{\|\mathbf{v}_{2}\|_{2}^{2}} \cdot A(:,3) \\ &= A(:,3) - \frac{\mathbf{v}_{1}^{T} \, A(:,3)}{\|\mathbf{v}_{1}\|_{2}^{2}} \cdot \mathbf{v}_{1} - \frac{\mathbf{v}_{2}^{T} \, A(:,3)}{\|\mathbf{v}_{2}\|_{2}^{2}} \cdot \mathbf{v}_{2} \end{split}$$

We can calculate inner products and norm from this equation separately as follows:

$$\mathbf{v}_1 \cdot A(:,3) = \mathbf{v}_1^T \cdot A(:,3) = \begin{bmatrix} 1 & 1 & -1 & 0 \end{bmatrix} \begin{bmatrix} 2 \\ -1 \\ 2 \\ 1 \end{bmatrix} = -1, \qquad \|\mathbf{v}_1\|_2^2 = 3$$

$$\mathbf{v}_2 \cdot A(:,3) = \mathbf{v}_2^T \cdot A(:,3) = \frac{1}{3} \begin{bmatrix} -1 & 2 & 1 & 3 \end{bmatrix} \begin{bmatrix} 2 \\ -1 \\ 2 \\ 1 \end{bmatrix} = \frac{1}{3}, \qquad \|\mathbf{v}_2\|_2^2 = \frac{5}{3}$$

With this we have

$$\mathbf{v}_3 = \begin{bmatrix} 2\\-1\\2\\1 \end{bmatrix} + \frac{1}{3} \begin{bmatrix} 1\\1\\-1\\0 \end{bmatrix} - \frac{1}{5} \begin{bmatrix} -1\\2\\1\\3 \end{bmatrix} = \frac{1}{5} \begin{bmatrix} 12\\-4\\8\\1 \end{bmatrix} = \begin{bmatrix} 2.4\\-0.8\\1.6\\0.2 \end{bmatrix}$$

22. (10 pts) A ride-sharing app called Super was recently released in 2009. During the first six years of it's business operations, the Super app has seen spectacular growth in its number of users:



We can model this data using a quadratic function $N(t) = a_0 + a_1 t + a_2 t^2$ for unknown positive constant a_0, a_1, a_2 . Recall from class that we can set up least-squares problem create this model.

(a) Set up the least-squares model associated with this data set. Explicitly identify the Vandermonde matrix A and the right-hand side vector **b**.

Solution:

Recall that the least squares problem is designed to fit data collected during an experiment to a particular mathematical model. In this case, we are told that our company collected seven data points $\{(t_i, N_i)\}_{i=0}^6$, where

 t_i = the number of years that the company Super has been in business for i = 0, 1, ..., 6 N_i = the number of people using the Super app during year t_i for i = 0, 1, ..., 6

We notice that the model appears to fit a quadratic model $N(t) = a_0 + a_1 t + a_2 t^2$ for unknown parameters $a_o, a_1, a_2 \in \mathbb{R}$. This model can be used to predict the internal force stored within the spring :

$$N(t_i) = a_0 + a_1 t_i + a_2 t_i^2.$$

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

$$e_i = (N_i - N(t_i)) = \left(N_i - \left(a_0 + a_1 t_i + a_2 t_i^2\right)\right)$$

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, a_2 so minimize the function

$$\sum_{i=0}^{6} e_i^2 = \left(N_i - \left(a_0 + a_1 t_i + a_2 t_i^2 \right) \right)^2 = \| \mathbf{b} - A \mathbf{x} \|_2^2 \tag{1}$$

where we define

| A = | $\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$ | $egin{array}{c} 0 \ 1 \ 2 \ 3 \ 4 \ 5 \ 6 \end{array}$ | $\begin{array}{c} 0 \\ 1 \\ 4 \\ 9 \\ 16 \\ 25 \\ 36 \end{array}$ | , $\mathbf{x} = \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix}$, $\mathbf{b} =$ | $\begin{bmatrix} 0.0 \\ 0.1 \\ 0.8 \\ 1.5 \\ 2.3 \\ 3.2 \\ 5.3 \end{bmatrix}$ | |
|-----|----------------------------------------------------------------------|--------------------------------------------------------|-------------------------------------------------------------------|-----------------------------------------------------------------------------------|-------------------------------------------------------------------------------|--|
|-----|----------------------------------------------------------------------|--------------------------------------------------------|-------------------------------------------------------------------|-----------------------------------------------------------------------------------|-------------------------------------------------------------------------------|--|

When attempting to solve this least squares problem, we want to find a vector $\mathbf{x}^* \in \mathbb{R}^3$ that solves the optimization problem

 $\min_{\mathbf{x}\in\mathbb{R}^3} \|\mathbf{b} - A\mathbf{x}\|_2^2$

This optimal \mathbf{x}^* will define parameters a_0, a_1, a_2 for the quadratic 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) Solve this least-squares problem to produce the best-fit quadratic function. If you use a calculator, be sure to explain the process you used to find your answer.

Solution: We want to solve the least-squares problem

$$\min_{\mathbf{x}\in\mathbb{R}^3} \|\mathbf{b} - A\mathbf{x}\|_2^2$$

for the Vandermonde matrix $A \in \mathbb{R}^{7 \times 3}$ and $\mathbf{b} \in \mathbb{R}^{7}$ 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 $\mathbf{x} \in \mathbb{R}^{3}$ 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

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

By Theorem 14 on p. 363, we know that $A^T A$ is invertible if and only if $\operatorname{rank}(A) = 3$. 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). The resulting linear systems problem is given by

$$\begin{bmatrix} 7 & 21 & 91 \\ 21 & 21 & 441 \\ 91 & 441 & 2275 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} 13.2 \\ 63.2 \\ 324.4 \end{bmatrix}$$

We can solve this linear-systems problem using any method we prefer to find

$$\begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} \approx \begin{bmatrix} 0.02381 \\ 0.04286 \\ 0.13333 \end{bmatrix}$$

Thus, the quadratic function of best fit is given by

$$N(t) = 0.02381 + 0.04286 \cdot t + 0.13333 \cdot t^2$$

23. Recall that for $A \in \mathbb{R}^{3 \times 3},$ the determinant of A was given by

$$\det(A) = \sum_{\pi \in S_3} \operatorname{sgn}(\pi) a_{\pi(1),1} a_{\pi(2),2} a_{\pi(3),3}$$

(a) (5 pts) List all permutations $\pi \in S_3$. In other words, list all maps $\pi : \{1, 2, 3\} \rightarrow \{1, 2, 3\}$ that are one-to-one and onto.

Solution: Consider S_3 , the permutation group on a set with three elements. We know that for each $i \in \{1, 2, ..., 6\}$, we have permutations

$$\pi_i: \{1,2,3\} \longrightarrow \{1,2,3\}$$

that are both one-to-one and onto. From our theorem above, we know that S_3 contains exactly 3! = 6 different permutations. We will label these permutations here. Consider:

$$\pi_1 := \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 2 \end{pmatrix}, \qquad \pi_2 := \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}, \qquad \pi_3 := \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix},$$
$$\pi_4 := \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}, \qquad \pi_5 := \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}, \qquad \pi_6 := \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}$$

To investigate inversions with respect to each π_i , we need to consider three different pairs:

$$(1,2)$$
 $(1,3)$ $(2,3)$

We see that $Inv(\pi_i) = \{(1,3), (2,3)\}$ are given by

| $\operatorname{Inv}\left(\pi_{1}\right)=\emptyset$ | \implies | $n(\pi_1) = 0$ |
|----------------------------------------------------|---------------|----------------|
| Inv $(\pi_2) = \{(1,3), (2,3)\}$ | \implies | $n(\pi_2) = 2$ |
| Inv $(\pi_3) = \{(1,2), (1,3)\}$ | \Rightarrow | $n(\pi_3) = 2$ |
| Inv $(\pi_4) = \{(1,2), (1,3), (2,3)\}$ | \implies | $n(\pi_4) = 3$ |
| Inv $(\pi_5) = \{(2,3)\}$ | \implies | $n(\pi_5) = 1$ |
| Inv $(\pi_6) = \{(1,2)\}$ | \implies | $n(\pi_6) = 1$ |

We can use this data to confirm that

 $sgn(\pi_1) = (-1)^{n(\pi_1)} = (-1)^0 = +1$ $sgn(\pi_2) = (-1)^{n(\pi_2)} = (-1)^2 = +1$ $sgn(\pi_3) = (-1)^{n(\pi_3)} = (-1)^2 = +1$ $sgn(\pi_4) = (-1)^{n(\pi_4)} = (-1)^3 = -1$ $sgn(\pi_5) = (-1)^{n(\pi_5)} = (-1)^1 = -1$ $sgn(\pi_6) = (-1)^{n(\pi_6)} = (-1)^1 = -1$

With this, we are ready to build the determinant function for 3×3 matrices.

(b) (5 pts) Use your work in part (a) and the determinant formula given above to prove that that the determinant of an upper triangular matrix $U \in \mathbb{R}^{3 \times 3}$ is the product of the main diagonal elements.

Solution: Using the information above, for matrix $A \in \mathbb{R}^{3 \times 3}$ we have

$$\det(A) = \sum_{\pi \in S_n} \operatorname{sgn}(\pi) \cdot a_{1\pi(1)} \cdot a_{2\pi(2)} \cdot a_{3\pi(3)}$$

$$=\sum_{i=1}^{6} \operatorname{sgn}(\pi_i) \cdot a_{1\pi_i(1)} \cdot a_{2\pi_i(2)} \cdot a_{3\pi(3)}$$

$$= a_{11} \cdot a_{22} \cdot a_{33} + a_{12} \cdot a_{23} \cdot a_{31} + a_{13} \cdot a_{21} \cdot a_{32} - a_{13} \cdot a_{22} \cdot a_{31} - a_{11} \cdot a_{23} \cdot a_{32} - a_{12} \cdot a_{21} \cdot a_{33}$$

If we assume that A is upper-triangular, we know

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{22} & a_{23} \\ 0 & 0 & a_{33} \end{bmatrix} \qquad \iff \qquad a_{ik} \begin{cases} \in \mathbb{R} & \text{if } i \le k \\ = 0 & \text{if } i > k \end{cases}$$

Thus, since all permutations $\pi_2, \pi_3, ..., \pi_6$ contain at least one inversion, we see that the determinant function greatly simplifies to

$$\det(A) = a_{11} \cdot a_{22} \cdot a_{33}$$

which is just the product of the diagonal elements. This is exactly what we wanted to show.