5.6 Approximations (Projections)

Learning Objectives

- terms: projection onto a line, normal equations, projection onto a subspace, relative error
- word problem on offering a company a counter-contract
Projection of a Vector Onto a Line

The vector $u - p$ is orthogonal to $v$. Also $p$ lies on the line and so is a multiple of $v$. Thus for some real number $x$:

$$(u - p) \cdot v = 0 \quad \text{and} \quad p = xv$$
The vector $u - p$ is orthogonal to $v$. Also $p$ lies on the line and so is a multiple of $v$. Thus for some real number $x$:

$$(u - p) \cdot v = 0 \quad \text{and} \quad p = xv$$

Combining these, we get

$$(u - xv) \cdot v = 0, \text{ i.e. } u \cdot v - x(v \cdot v) = 0, \text{ i.e. } x = \frac{u \cdot v}{v \cdot v}$$
Projection of a Vector Onto a Line

The vector $u - p$ is orthogonal to $v$. Also $p$ lies on the line and so is a multiple of $v$. Thus for some real number $x$:

$$(u - p) \cdot v = 0 \quad \text{and} \quad p = xv$$

Combining these, we get

$$(u - xv) \cdot v = 0, \text{ i.e. } u \cdot v - x(v \cdot v) = 0, \text{ i.e. } x = \frac{u \cdot v}{v \cdot v}$$

This result leads to:

**Definition:** The projection of a vector $u$ onto the line $\text{span}\{v\}$
Projection of a Vector Onto a Line

The vector $u - p$ is orthogonal to $v$. Also $p$ lies on the line and so is a multiple of $v$. Thus for some real number $x$:

$$(u - p) \cdot v = 0 \text{ and } p = xv$$

Combining these, we get

$$(u - xv) \cdot v = 0, \text{ i.e. } u \cdot v - x(v \cdot v) = 0, \text{ i.e. } x = \frac{u \cdot v}{v \cdot v}$$

This result leads to:

**Definition**: The projection of a vector $u$ onto the line span$\{v\}$ is

$$p = \text{Proj}_v(u) = \left(\frac{u \cdot v}{v \cdot v}\right)v$$
The vector $u - p$ is orthogonal to $v$. Also $p$ lies on the line and so is a multiple of $v$. Thus for some real number $x$:

$$(u - p) \cdot v = 0 \quad \text{and} \quad p = xv$$

Combining these, we get

$$(u - xv) \cdot v = 0, \text{ i.e. } u \cdot v - x(v \cdot v) = 0, \text{ i.e. } x = \frac{u \cdot v}{v \cdot v}$$

This result leads to:

**Definition:** The projection of a vector $u$ onto the line $\text{span}\{v\}$ is

$$p = \text{Proj}_v(u) = \left(\frac{u \cdot v}{v \cdot v}\right) v$$

It is the vector “closest” to $u$ on the line determined by $v$. 
**Example**

**Ex 1:** Find $\text{Proj}_\vec{v}(\vec{u})$ when $\vec{u} = (2, 3)$ and $\vec{v} = (2, 2)$. 

$$
\text{Proj}_\vec{v}(\vec{u}) = (\vec{u} \cdot \vec{v}) \frac{\vec{v}}{\|\vec{v}\|^2}
$$

$$
= \left( (2, 3) \cdot (2, 2) \right) \frac{(2, 2)}{(2^2 + 2^2)}
$$

$$
= (4 + 6) \frac{(2, 2)}{4}
$$

$$
= \frac{10}{4} (2, 2)
$$

$$
= \frac{5}{2} (2, 2) = (5, 5)
$$
**Example**

**Ex 1:** Find \( \text{Proj}_\vec{v}(\vec{u}) \) when \( \vec{u} = (2, 3) \) and \( \vec{v} = (2, 2) \).

\[
\text{Proj}_\vec{v}(\vec{u}) = \left( \frac{\vec{u} \cdot \vec{v}}{\vec{v} \cdot \vec{v}} \right) \vec{v}
\]

\[
= \left( \frac{(2, 3) \cdot (2, 2)}{(2, 2) \cdot (2, 2)} \right) \vec{v}
\]

\[
= \left( \frac{4 + 6}{4 + 4} \right) \vec{v}
\]

\[
= \frac{10}{8} (2, 2) = (\frac{5}{2}, \frac{5}{2})
\]
The projection of $\vec{u}$ onto the subspace $S = \text{span}\{\vec{v}_1, \vec{v}_2\}$. 
The following algorithm projects $\mathbf{u}$ onto $S = \text{span}\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$:

1. Let $A = [\mathbf{v}_1 \ \mathbf{v}_2 \ \ldots \ \mathbf{v}_k]$
2. Find $A^T A$ and $A^T u$. Solve $(A^T A)x = (A^T u)$ for $x$.
3. Final answer: $	ext{Proj}_S(\mathbf{u}) = A x$ (Or equivalently: $	ext{Proj}_S(\mathbf{u}) = x_1 \mathbf{v}_1 + x_2 \mathbf{v}_2 + \cdots + x_k \mathbf{v}_k$)
The following algorithm projects $\mathbf{u}$ onto $S = \text{span}\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$:

1. Let $A = [\mathbf{v}_1 \ \mathbf{v}_2 \ \ldots \ \mathbf{v}_k]$
Projection onto Subspaces

The following algorithm projects $\mathbf{u}$ onto $S = \text{span}\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$:

1. Let $A = [\mathbf{v}_1 \mathbf{v}_2 \ldots \mathbf{v}_k]$

2. Find $A^T A$ and $A^T \mathbf{u}$.
Projection onto Subspaces

The following algorithm projects $\mathbf{u}$ onto $S = \text{span}\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$:

1. Let $A = [\mathbf{v}_1 \ \mathbf{v}_2 \ \ldots \ \mathbf{v}_k]$

2. Find $A^T A$ and $A^T \mathbf{u}$.

   Solve $(A^T A) \mathbf{x} = (A^T \mathbf{u})$ for $\mathbf{x}$. $\leftarrow$ “the normal equations”
The following algorithm projects \( \mathbf{u} \) onto \( S = \text{span}\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\} \):

1. Let \( A = [\mathbf{v}_1 \ \mathbf{v}_2 \ \ldots \ \mathbf{v}_k] \)

2. Find \( A^T A \) and \( A^T \mathbf{u} \).
   
   Solve \( (A^T A)x = (A^T \mathbf{u}) \) for \( x \). ← “the normal equations”

3. Final answer: \( \text{Proj}_S(\mathbf{u}) = Ax \)
The following algorithm projects $\mathbf{u}$ onto $S = \text{span}\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$:

1. Let $A = [\mathbf{v}_1 \ \mathbf{v}_2 \ \ldots \ \mathbf{v}_k]$

2. Find $A^T A$ and $A^T \mathbf{u}$.
   
   Solve $(A^T A)x = (A^T \mathbf{u})$ for $x$. ← “the normal equations”

3. Final answer: $\text{Proj}_S(\mathbf{u}) = Ax$
   
   (Or equivalently: $\text{Proj}_S(\mathbf{u}) = x_1 \mathbf{v}_1 + x_2 \mathbf{v}_2 + \cdots + x_k \mathbf{v}_k$)
Example

**Ex 2:** Find $\text{Proj}_S(u)$ when $u = (2, 2, 3)$ and $S$ is the span of $v_1 = (1, -1, 1)$ and $v_2 = (2, 1, -1)$.

1. Form the matrix $A$ with $v_1$ and $v_2$ as columns: $A = \begin{bmatrix} 1 & 2 \\ -1 & 1 \\ 1 & -1 \end{bmatrix}$

2. $A^T A = \begin{bmatrix} 3 & 0 \\ 0 & 6 \end{bmatrix}, A^T u = \begin{bmatrix} 3 \\ 3 \end{bmatrix}$

The solution to the normal equations is then
Example

Ex 2: Find $\text{Proj}_S(u)$ when $u = (2, 2, 3)$ and $S$ is the span of $v_1 = (1, -1, 1)$ and $v_2 = (2, 1, -1)$.

1. Form the matrix $A$ with $v_1$ and $v_2$ as columns: $A = \begin{bmatrix} 1 & 2 \\ -1 & 1 \\ 1 & -1 \end{bmatrix}$

2. $A^T A = \begin{bmatrix} 3 & 0 \\ 0 & 6 \end{bmatrix}, A^T u = \begin{bmatrix} 3 \\ 3 \end{bmatrix}$

The solution to the normal equations is then

\[
\begin{bmatrix} 3 & 0 & 3 \\ 0 & 6 & 3 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1/2 \end{bmatrix} \rightarrow x_1 = 1, \quad x_2 = 1/2
\]
Example

**Ex 2:** Find \( \text{Proj}_S(u) \) when \( u = (2, 2, 3) \) and \( S \) is the span of \( \mathbf{v}_1 = (1, -1, 1) \) and \( \mathbf{v}_2 = (2, 1, -1) \).

1. Form the matrix \( A \) with \( \mathbf{v}_1 \) and \( \mathbf{v}_2 \) as columns: 
\[
A = \begin{bmatrix}
1 & 2 \\
-1 & 1 \\
1 & -1 \\
\end{bmatrix}
\]

2. \( A^T A = \begin{bmatrix} 3 & 0 \\ 0 & 6 \end{bmatrix}, A^T u = \begin{bmatrix} 3 \\ 3 \end{bmatrix} \)

The solution to the normal equations is then
\[
\begin{bmatrix}
3 & 0 & 3 \\
0 & 6 & 3 \\
\end{bmatrix} \rightarrow 
\begin{bmatrix}
1 & 0 & 1 \\
0 & 1 & 1/2 \\
\end{bmatrix} \rightarrow \ x_1 = 1 \\
\ x_2 = 1/2
\]

3. Final answer:
\[
\text{Proj}_S(u) = (1) \mathbf{v}_1 + (1/2) \mathbf{v}_2 = (2, -1/2, 1/2) \]
Example

Ex 2: Find $\text{Proj}_S(u)$ when $u = (2, 2, 3)$ and $S$ is the span of $v_1 = (1, -1, 1)$ and $v_2 = (2, 1, -1)$.

1. Form the matrix $A$ with $v_1$ and $v_2$ as columns: $A = \begin{bmatrix} 1 & 2 \\ -1 & 1 \\ 1 & -1 \end{bmatrix}$

2. $A^T A = \begin{bmatrix} 3 & 0 \\ 0 & 6 \end{bmatrix}$, $A^T u = \begin{bmatrix} 3 \\ 3 \end{bmatrix}$

   The solution to the normal equations is then

   $\begin{bmatrix} 3 & 0 & 3 \\ 0 & 6 & 3 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1/2 \end{bmatrix} \rightarrow x_1 = 1$

   $x_2 = 1/2$

3. Final answer:

   $\text{Proj}_S(u) = (1)v_1 + (1/2)v_2 = (2, -1/2, 1/2)$
**Example 3 from Section 5.6**

**Ex 3:** An apparel company makes sweaters, hats, and gloves at two plants. The daily production is as follows:

<table>
<thead>
<tr>
<th></th>
<th>Plant 1</th>
<th>Plant 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>sweaters</td>
<td>100</td>
<td>400</td>
</tr>
<tr>
<td>hats</td>
<td>200</td>
<td>700</td>
</tr>
<tr>
<td>gloves</td>
<td>300</td>
<td>1000</td>
</tr>
</tbody>
</table>

The Navy has approached the company with a contract proposal to make 2,000 sweaters, 6,700 hats, and 6,500 gloves per week. Can the company offer a production schedule that will meet this contract? If not, what is the best counter-contract the company can offer the Navy in getting close to its original needs?
Ex 3: An apparel company makes sweaters, hats, and gloves at two plants. The daily production is as follows:

<table>
<thead>
<tr>
<th></th>
<th>Plant 1</th>
<th>Plant 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>sweaters</td>
<td>100</td>
<td>400</td>
</tr>
<tr>
<td>hats</td>
<td>200</td>
<td>700</td>
</tr>
<tr>
<td>gloves</td>
<td>300</td>
<td>1000</td>
</tr>
</tbody>
</table>

The Navy has approached the company with a contract proposal to make 2,000 sweaters, 6,700 hats, and 6,500 gloves per week. Can the company offer a production schedule that will meet this contract? If not, what is the best counter-contract the company can offer the Navy in getting close to its original needs?

\[
P_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \quad P_2 = \begin{bmatrix} 4 \\ 7 \\ 10 \end{bmatrix}, \quad C = \begin{bmatrix} 20 \\ 67 \\ 65 \end{bmatrix}
\]

sweaters  hats  gloves
The contract cannot be met – no solution to $\mathbf{C} = c_1 \mathbf{P}_1 + c_2 \mathbf{P}_2$. 

We should run $\mathbf{P}_1$ and $\mathbf{P}_2$ for $x_1$ and $x_2$ days:

$\mathbf{C} = x_1 \mathbf{P}_1 + x_2 \mathbf{P}_2 = \frac{11}{2} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + \frac{17}{3} \begin{bmatrix} 4 \\ 7 \\ 10 \end{bmatrix} = \begin{bmatrix} 169/6 \\ 152/3 \\ 439/6 \end{bmatrix}$
The contract cannot be met – no solution to $\mathbf{C} = c_1 \mathbf{P}_1 + c_2 \mathbf{P}_2$.

Need to project $\mathbf{C}$ onto the contract space span\{\mathbf{P}_1, \mathbf{P}_2\}:
The contract cannot be met – no solution to $\mathbf{C} = c_1 \mathbf{P}_1 + c_2 \mathbf{P}_2$.

Need to project $\mathbf{C}$ onto the contract space span$\{\mathbf{P}_1, \mathbf{P}_2\}$:

$$\mathbf{A} = [\mathbf{P}_1 \ \mathbf{P}_2] = \begin{bmatrix} 1 & 4 \\ 2 & 7 \\ 3 & 10 \end{bmatrix}$$
The contract cannot be met – no solution to \( \mathbf{C} = c_1 \mathbf{P}_1 + c_2 \mathbf{P}_2 \).

Need to project \( \mathbf{C} \) onto the contract space span\( \{\mathbf{P}_1, \mathbf{P}_2\} \):

1. \[
A = [\mathbf{P}_1 \quad \mathbf{P}_2] = \begin{bmatrix} 1 & 4 \\ 2 & 7 \\ 3 & 10 \end{bmatrix}
\]

2. \[
A^T A = \begin{bmatrix} 14 & 48 \\ 48 & 165 \end{bmatrix}, \quad A^T \mathbf{C} = \begin{bmatrix} 349 \\ 1199 \end{bmatrix}.
\]
The contract cannot be met – no solution to $\mathbf{C} = c_1 \mathbf{P}_1 + c_2 \mathbf{P}_2$.

Need to project $\mathbf{C}$ onto the contract space span\{\mathbf{P}_1, \mathbf{P}_2\}:

1. $A = \begin{bmatrix} \mathbf{P}_1 & \mathbf{P}_2 \end{bmatrix} = \begin{bmatrix} 1 & 4 \\ 2 & 7 \\ 3 & 10 \end{bmatrix}$

2. $A^T A = \begin{bmatrix} 14 & 48 \\ 48 & 165 \end{bmatrix}$, $A^T \mathbf{C} = \begin{bmatrix} 349 \\ 1199 \end{bmatrix}$.

\[
\begin{bmatrix}
14 & 48 & 349 \\
48 & 165 & 1199
\end{bmatrix} \rightarrow \begin{bmatrix}
1 & 0 & 11/2 \\
0 & 1 & 17/3
\end{bmatrix} \rightarrow \begin{bmatrix}
x_1 = 5.5 \\
x_2 = 5.666...
\end{bmatrix}
\]
The contract cannot be met – no solution to \( \mathbf{C} = c_1 \mathbf{P}_1 + c_2 \mathbf{P}_2 \).

Need to project \( \mathbf{C} \) onto the contract space span\( \{ \mathbf{P}_1, \mathbf{P}_2 \} \):

1. \[
A = \begin{bmatrix}
\mathbf{P}_1 & \mathbf{P}_2
\end{bmatrix} = \begin{bmatrix}
1 & 4 \\
2 & 7 \\
3 & 10
\end{bmatrix}
\]

2. \[
A^T A = \begin{bmatrix} 14 & 48 \\ 48 & 165 \end{bmatrix}, \quad A^T \mathbf{C} = \begin{bmatrix} 349 \\ 1199 \end{bmatrix}.
\]

\[
\begin{bmatrix} 14 & 48 & 349 \\ 48 & 165 & 1199 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & 11/2 \\ 0 & 1 & 17/3 \end{bmatrix} \rightarrow \begin{array}{c} x_1 = 5.5 \\ x_2 = 5.666... \end{array}
\]

3. we should run \( \mathbf{P}_1 \) and \( \mathbf{P}_2 \) for \( x_1 \) and \( x_2 \) days:

\[
\mathbf{C} \mathbf{C} = x_1 \mathbf{P}_1 + x_2 \mathbf{P}_2 = 11/2 \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + 17/3 \begin{bmatrix} 4 \\ 7 \\ 10 \end{bmatrix} = \begin{bmatrix} 169/6 \\ 152/3 \\ 439/6 \end{bmatrix}
\]
$\mathbf{C} = (20, 67, 65)$

$\mathbf{C}\mathbf{C} = (\frac{169}{6}, \frac{152}{3}, \frac{439}{6}) \approx (28.17, 50.67, 73.17)$

Relative error $= \frac{|\mathbf{C} - \mathbf{C}\mathbf{C}|}{|\mathbf{C}|}$

$= \frac{|(20-28.17, 67-50.67, 65-73.17)|}{|(20,67,65)|}$

$= \frac{|(-8.17, 17.67, -8.17)|}{|(20,67,65)|}$

$= \frac{\sqrt{(-8.17)^2 + 17.67^2 + (-8.17)^2}}{\sqrt{20^2 + 67^2 + 65^2}}$

$= \frac{20.00}{96.47}$

$\approx 20.95\%$
5.7 Approximations (Least Squares)

Learning Objectives

- terms: fitting a line to data, least squares approximation
- learn how to:
  - find a line that best approximates a set of points in the plane
  - use the equation for the line to estimate unknown data
Example 1

**Ex 1:** In a pilot marketing analysis for a new handyman drill, the King Hardware Company distributed the drill for a month in four different cities at different prices:

<table>
<thead>
<tr>
<th>Price ((x)) dollars</th>
<th>Demand ((y)) drills</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>84</td>
</tr>
<tr>
<td>25</td>
<td>74</td>
</tr>
<tr>
<td>30</td>
<td>62</td>
</tr>
<tr>
<td>35</td>
<td>50</td>
</tr>
</tbody>
</table>

How can we estimate the demand \(y\) when the price \(x\) is 40 dollars? If we find a line \(y = mx + b\) that goes through the above points, then we can estimate \(y = m \cdot 40 + b\).
**Ex 1:** In a pilot marketing analysis for a new handyman drill, the King Hardware Company distributed the drill for a month in four different cities at different prices:

<table>
<thead>
<tr>
<th>Price (x) dollars</th>
<th>Demand (y) drills</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>84</td>
</tr>
<tr>
<td>25</td>
<td>74</td>
</tr>
<tr>
<td>30</td>
<td>62</td>
</tr>
<tr>
<td>35</td>
<td>50</td>
</tr>
</tbody>
</table>

How can we estimate the demand $y$ when the price $x$ is 40 dollars?
Example 1

Ex 1: In a pilot marketing analysis for a new handyman drill, the King Hardware Company distributed the drill for a month in four different cities at different prices:

<table>
<thead>
<tr>
<th>Price ($x$) dollars</th>
<th>Demand ($y$) drills</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>84</td>
</tr>
<tr>
<td>25</td>
<td>74</td>
</tr>
<tr>
<td>30</td>
<td>62</td>
</tr>
<tr>
<td>35</td>
<td>50</td>
</tr>
</tbody>
</table>

How can we estimate the demand $y$ when the price $x$ is 40 dollars?

If we find a line $y = mx + b$ that goes through the above points, then we can estimate $y = m \cdot 40 + b$. 
Let the linear relationship between price and demand be

\[ y = mx + b \]
Example 1

Let the linear relationship between price and demand be

\[ y = mx + b \]

If all our points lie on this line, then each point will satisfy \( y = mx + b \):

\[
84 = 20m + b \\
74 = 25m + b \\
62 = 30m + b \\
50 = 35m + b
\]
Example 1

Let the linear relationship between price and demand be

\[ y = mx + b \]

If all our points lie on this line, then each point will satisfy \( y = mx + b \):

\[
\begin{align*}
84 &= 20m + b \\
74 &= 25m + b \\
62 &= 30m + b \\
50 &= 35m + b \\
\end{align*}
\]

Now all we need do is solve these four equations for the unknowns \( m \) and \( b \).
These four equations make up a linear system for $m$ and $b$, which can be written as a matrix equation:

\[
\begin{align*}
    b + 20m &= 84 \\
    b + 25m &= 74 \\
    b + 30m &= 62 \\
    b + 35m &= 50 \\
\end{align*}
\]

\[
\begin{bmatrix}
    1 & 20 \\
    1 & 25 \\
    1 & 30 \\
    1 & 35 \\
\end{bmatrix}
\begin{bmatrix}
    b \\
    m \\
\end{bmatrix}
= 
\begin{bmatrix}
    84 \\
    74 \\
    62 \\
    50 \\
\end{bmatrix}
\]

\[\iff\]

iClicker 1: What is the solution to the above system?

(A) $b = 124$ and $m = -2$

(B) $b = 128$ and $m = -11/5$

(C) $(0, 0, 1, 0)$

(D) $(0, 0)$

(E) No solution.
These four equations make up a linear system for \( m \) and \( b \), which can be written as a matrix equation:

\[
\begin{align*}
    b + 20m &= 84 \\
    b + 25m &= 74 \\
    b + 30m &= 62 \\
    b + 35m &= 50
\end{align*}
\]

\[
\Leftrightarrow \begin{bmatrix}
1 & 20 \\
1 & 25 \\
1 & 30 \\
1 & 35
\end{bmatrix} \begin{bmatrix}
b \\
m
\end{bmatrix} = \begin{bmatrix}
84 \\
74 \\
62 \\
50
\end{bmatrix}
\]

This is \( A \mathbf{x} = \mathbf{u} \) where \( A \) is the \( 4 \times 2 \) matrix above, \( \mathbf{x} \) is the vector with entries \( b \) and \( m \) and \( \mathbf{u} \) is the 4-vector on the right-hand side.

---

**iClicker 1**: What is the solution to the above system?

(A) \( b = 124 \) and \( m = -2 \)

(B) \( b = 128 \) and \( m = -11 \) / 5

(C) \( (0, 0, 1, 0) \)

(D) \( (0, 0) \)

(E) No solution.
These four equations make up a linear system for \( m \) and \( b \), which can be written as a matrix equation:

\[
\begin{align*}
    b + 20m &= 84 \\
    b + 25m &= 74 \\
    b + 30m &= 62 \\
    b + 35m &= 50
\end{align*}
\]

\[
\begin{bmatrix}
    1 & 20 \\
    1 & 25 \\
    1 & 30 \\
    1 & 35
\end{bmatrix}
\begin{bmatrix}
    b \\
    m
\end{bmatrix}
= \begin{bmatrix}
    84 \\
    74 \\
    62 \\
    50
\end{bmatrix}
\]

This is \( Ax = u \) where \( A \) is the \( 4 \times 2 \) matrix above, \( x \) is the vector with entries \( b \) and \( m \) and \( u \) is the 4-vector on the right-hand side.

iClicker 1: What is the solution to the above system?

(A) \( b = 124 \) and \( m = -2 \)

(B) \( b = 128 \) and \( m = -11/5 \)

(C) \( (0, 0, 1, 0) \)

(D) \( (0, 0) \)

(E) No solution.
These four equations make up a linear system for $m$ and $b$, which can be written as a matrix equation:

\[
\begin{align*}
   b + 20m &= 84 \\
   b + 25m &= 74 \\
   b + 30m &= 62 \\
   b + 35m &= 50 \\
\end{align*}
\]

\[
\begin{pmatrix}
   1 & 20 \\
   1 & 25 \\
   1 & 30 \\
   1 & 35 \\
\end{pmatrix}
\begin{pmatrix}
   b \\
   m \\
\end{pmatrix}
=
\begin{pmatrix}
   84 \\
   74 \\
   62 \\
   50 \\
\end{pmatrix}
\]

This is $Ax = u$ where $A$ is the $4 \times 2$ matrix above, $x$ is the vector with entries $b$ and $m$ and $u$ is the 4-vector on the right-hand side.
Discussion

These four equations make up a linear system for $m$ and $b$, which can be written as a matrix equation:

\[
\begin{align*}
  b + 20m &= 84 \\
  b + 25m &= 74 \\
  b + 30m &= 62 \\
  b + 35m &= 50
\end{align*}
\]

\[
\begin{bmatrix}
  1 & 20 \\
  1 & 25 \\
  1 & 30 \\
  1 & 35
\end{bmatrix}
\begin{bmatrix}
  b \\
  m
\end{bmatrix}
= \begin{bmatrix}
  84 \\
  74 \\
  62 \\
  50
\end{bmatrix}
\]

This is $Ax = u$ where $A$ is the $4 \times 2$ matrix above, $x$ is the vector with entries $b$ and $m$ and $u$ is the 4-vector on the right-hand side.

- This system is inconsistent (the points are not all on a common line)!
These four equations make up a linear system for \( m \) and \( b \), which can be written as a matrix equation:

\[
\begin{align*}
    b + 20m &= 84 \\
    b + 25m &= 74 \\
    b + 30m &= 62 \\
    b + 35m &= 50
\end{align*}
\]

\[
\begin{bmatrix}
    1 & 20 \\
    1 & 25 \\
    1 & 30 \\
    1 & 35
\end{bmatrix}
\begin{bmatrix}
    b \\
    m
\end{bmatrix}
= 
\begin{bmatrix}
    84 \\
    74 \\
    62 \\
    50
\end{bmatrix}
\]

This is \( Ax = u \) where \( A \) is the \( 4 \times 2 \) matrix above, \( x \) is the vector with entries \( b \) and \( m \) and \( u \) is the 4-vector on the right-hand side.

- This system is inconsistent (the points are not all on a common line)!
- The figure shows that a straight line relationship is “almost” true.
These four equations make up a linear system for $m$ and $b$, which can be written as a matrix equation:

\[
\begin{align*}
  b + 20m &= 84 \\
  b + 25m &= 74 \\
  b + 30m &= 62 \\
  b + 35m &= 50
\end{align*}
\]

\[
\begin{pmatrix}
  1 & 20 \\
  1 & 25 \\
  1 & 30 \\
  1 & 35
\end{pmatrix}
\begin{pmatrix}
  b \\
  m
\end{pmatrix}
=
\begin{pmatrix}
  84 \\
  74 \\
  62 \\
  50
\end{pmatrix}
\]

This is $Ax = u$ where $A$ is the $4 \times 2$ matrix above, $x$ is the vector with entries $b$ and $m$ and $u$ is the 4-vector on the right-hand side.

- This system is inconsistent (the points are not all on a common line)!
- The figure shows that a straight line relationship is “almost” true.
- How can we find the best line approximation to the data?
In Section 5.6 we learned how to project a vector $u$ onto a subspace $S = \text{col}(A)$. Saying that $u$ is not in $S = \text{col}(A)$ is equivalent to saying that $Ax = u$ is inconsistent. Since the projection $p$ of $u$ onto $S = \text{col}(A)$ lies in $\text{col}(A)$, $Ax = p$ is consistent.

Recall that $p$ is as close as possible in $S$ to $u$ in the sense that $|u - p|$ is as small as possible.

e = (e_1, e_2, ..., e_n) = u - p = u - Ax$ is the "error" vector in using $u$ rather than $p$.

$|e|_2 = e_1^2 + e_2^2 + \cdots + e_n^2 = \text{sum of squares of errors}$
Remarks

Projections!
In Section 5.6 we learned how to project a vector $u$ onto a subspace $S$. 

Recall that $p$ is as close as possible in $S$ to $u$ in the sense that $|u - p|$ is as small as possible.

$e = (e_1, e_2, \ldots, e_n) = u - p = u - Ax$ is the “error” vector in using $u$ rather than $p$.

$|e|^2 = e_1^2 + e_2^2 + \cdots + e_n^2 = \text{sum of squares of errors}$
Projections!
In Section 5.6 we learned how to project a vector \( \mathbf{u} \) onto a subspace \( S \).

- Saying that \( \mathbf{u} \) is not in \( S = \text{col}(A) \) is equivalent to saying that \( A\mathbf{x} = \mathbf{u} \) is inconsistent.

\[ e = (e_1, e_2, ..., e_n) = \mathbf{u} - p = \mathbf{u} - A\mathbf{x} \]

\[ |e|_2^2 = e_1^2 + e_2^2 + \cdots + e_n^2 = \text{sum of squares of errors} \]
Projections!
In Section 5.6 we learned how to project a vector \( \mathbf{u} \) onto a subspace \( S \).

- Saying that \( \mathbf{u} \) is not in \( S = \text{col}(A) \) is equivalent to saying that \( A\mathbf{x} = \mathbf{u} \) is inconsistent.
- Since the projection \( \mathbf{p} \) of \( \mathbf{u} \) onto \( S = \text{col}(A) \) lies in \( \text{col}(A) \), \( A\mathbf{x} = \mathbf{p} \) is consistent.
Remarks

Projections!
In Section 5.6 we learned how to project a vector $\mathbf{u}$ onto a subspace $S$.

- Saying that $\mathbf{u}$ is not in $S = \text{col}(A)$ is equivalent to saying that $A\mathbf{x} = \mathbf{u}$ is inconsistent.
- Since the projection $\mathbf{p}$ of $\mathbf{u}$ onto $S = \text{col}(A)$ lies in $\text{col}(A)$, $A\mathbf{x} = \mathbf{p}$ is consistent.
- Recall that $\mathbf{p}$ is as close as possible in $S$ to $\mathbf{u}$ in the sense that $|\mathbf{u} - \mathbf{p}|$ is as small as possible.
Remarks

Projections!
In Section 5.6 we learned how to project a vector \( \mathbf{u} \) onto a subspace \( S \).

- Saying that \( \mathbf{u} \) is not in \( S = \text{col}(A) \) is equivalent to saying that \( A \mathbf{x} = \mathbf{u} \) is inconsistent.
- Since the projection \( \mathbf{p} \) of \( \mathbf{u} \) onto \( S = \text{col}(A) \) lies in \( \text{col}(A) \), \( A \mathbf{x} = \mathbf{p} \) is consistent.
- Recall that \( \mathbf{p} \) is as close as possible in \( S \) to \( \mathbf{u} \) in the sense that \( |\mathbf{u} - \mathbf{p}| \) is as small as possible.
- \( \mathbf{e} = (e_1, e_2, \ldots, e_n) = \mathbf{u} - \mathbf{p} = \mathbf{u} - A \mathbf{x} \) is the “error” vector in using \( \mathbf{u} \) rather than \( \mathbf{p} \).
Projections!

In Section 5.6 we learned how to project a vector $\mathbf{u}$ onto a subspace $S$.

- Saying that $\mathbf{u}$ is not in $S = \text{col}(A)$ is equivalent to saying that $A\mathbf{x} = \mathbf{u}$ is inconsistent.
- Since the projection $\mathbf{p}$ of $\mathbf{u}$ onto $S = \text{col}(A)$ lies in $\text{col}(A)$, $A\mathbf{x} = \mathbf{p}$ is consistent.
- Recall that $\mathbf{p}$ is as close as possible in $S$ to $\mathbf{u}$ in the sense that $|\mathbf{u} - \mathbf{p}|$ is as small as possible.
- $\mathbf{e} = (e_1, e_2, ..., e_n) = \mathbf{u} - \mathbf{p} = \mathbf{u} - A\mathbf{x}$ is the “error” vector in using $\mathbf{u}$ rather than $\mathbf{p}$.
- $|\mathbf{e}|^2 = e_1^2 + e_2^2 + \cdots + e_n^2 = \text{sum of squares of errors}$
**Least Squares Approximation**

**Definition**: The *least squares error approximation* to the inconsistent system $Ax = u$ is the solution $x$ of $Ax = p$ where $p$ is the projection of $u$ onto the column space of $A$. 

1. Let $A = \begin{bmatrix} 1 & x_1 & \cdots & x_n \end{bmatrix}$ and $u = \begin{bmatrix} y_1 & y_2 & \cdots & y_n \end{bmatrix}$.
2. Find $A^T A$ and $A^T u$, and solve normal equations $(A^T A)x = A^T u$.
3. $p = Ax$ is the approximation of $u$. The $|u - p|^2$ is the sum of squares of errors.
Least Squares Approximation

**Definition:** The **least squares error approximation** to the inconsistent system $Ax = u$ is the solution $x$ of $Ax = p$ where $p$ is the projection of $u$ onto the column space of $A$.

Given points $(x_1, y_1), (x_2, y_2), ..., (x_n, y_n)$,
Least Squares Approximation

**Definition:** The least squares error approximation to the inconsistent system $Ax = u$ is the solution $x$ of $Ax = p$ where $p$ is the projection of $u$ onto the column space of $A$.

Given points $(x_1, y_1), (x_2, y_2), ..., (x_n, y_n)$,

1. Let

   $$A = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix} \quad \text{and} \quad u = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$
Least Squares Approximation

**Definition:** The least squares error approximation to the inconsistent system $Ax = u$ is the solution $x$ of $Ax = p$ where $p$ is the projection of $u$ onto the column space of $A$.

Given points $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$,

1. Let

$$
A = \begin{bmatrix}
1 & x_1 \\
1 & x_2 \\
\vdots & \vdots \\
1 & x_n \\
\end{bmatrix}
\quad \text{and} \quad
u = \begin{bmatrix}
y_1 \\
y_2 \\
\vdots \\
y_n \\
\end{bmatrix}
$$

2. Find $A^TA$ and $A^T u$, and solve normal equations $(A^TA)x = (A^T u)$. 

$p = Ax$ is the approximation of $u$. The $|u - p|^2$ is the sum of squares of errors.
Definition: The least squares error approximation to the inconsistent system $Ax = u$ is the solution $x$ of $Ax = p$ where $p$ is the projection of $u$ onto the column space of $A$.

Given points $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$,

1. Let

$$A = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix} \quad \text{and} \quad u = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

2. Find $A^T A$ and $A^T u$, and solve normal equations $(A^T A)x = (A^T u)$.

$\Rightarrow x$ gives you the slope and intercept of best fit line: $x = \begin{bmatrix} b \\ m \end{bmatrix}$
Least Squares Approximation

**Definition:** The least squares error approximation to the inconsistent system $Ax = u$ is the solution $x$ of $Ax = p$ where $p$ is the projection of $u$ onto the column space of $A$.

Given points $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$,

1. Let

$$A = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix} \quad \text{and} \quad u = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

2. Find $A^TA$ and $A^Tu$, and solve normal equations $(A^TA)x = (A^Tu)$.

$\Rightarrow x$ gives you the slope and intercept of best fit line: $x = \begin{bmatrix} b \\ m \end{bmatrix}$

3. $p = Ax$ is the approximation of $u$. The $|u - p|^2$ is the sum of squares of errors.
**Example 1 Revisited**

**Ex 1:** Find the least squares error best fit line for the data in example 1 and plot the line with the data.

1. Let
   \[
   A = \begin{bmatrix}
   1 & 20 \\
   1 & 25 \\
   1 & 30 \\
   1 & 35
   \end{bmatrix}
   \quad \text{and} \quad
   u = \begin{bmatrix}
   84 \\
   74 \\
   62 \\
   50
   \end{bmatrix}
   .
   \]

2. Then
   \[
   A^T A = \begin{bmatrix}
   4 & 110 \\
   110 & 3150
   \end{bmatrix}
   \quad \text{and} \quad
   A^T u = \begin{bmatrix}
   270 \\
   7140
   \end{bmatrix}
   .
   \]

Solve the normal equations
   \[
   (A^T A) x = (A^T u)
   \]
Example 1 Revisited

**Ex 1**: Find the least squares error best fit line for the data in example 1 and plot the line with the data.

1. Let \( A = \begin{bmatrix} 1 & 20 \\ 1 & 25 \\ 1 & 30 \\ 1 & 35 \end{bmatrix} \) and \( u = \begin{bmatrix} 84 \\ 74 \\ 62 \\ 50 \end{bmatrix} \).

Then \( A^T A = \begin{bmatrix} 4 & 110 \\ 110 & 3150 \end{bmatrix} \) and \( A^T u = \begin{bmatrix} 270 \\ 7140 \end{bmatrix} \).

Solve the normal equations \((A^T A)x = A^T u\):

\[
\begin{bmatrix} 4 & 110 \\ 110 & 3150 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 \\ 130.2 & 1 \\ 0 & -2.28 \end{bmatrix}
\]

\( b = 130.2 \) and \( m = -2.28 \).

The best fit line is \( y = -2.28x + 130.2 \).
**Example 1 Revisited**

**Ex 1:** Find the least squares error best fit line for the data in example 1 and plot the line with the data.

1. Let \( A = \begin{bmatrix} 1 & 20 \\ 1 & 25 \\ 1 & 30 \\ 1 & 35 \end{bmatrix} \) and \( u = \begin{bmatrix} 84 \\ 74 \\ 62 \\ 50 \end{bmatrix} \).

2. Then \( A^T A = \begin{bmatrix} 4 & 110 \\ 110 & 3150 \end{bmatrix} \) and \( A^T u = \begin{bmatrix} 270 \\ 7140 \end{bmatrix} \).

The best fit line is \( y = -2.28x + 130.2 \).
**Example 1 Revisited**

**Ex 1**: Find the least squares error best fit line for the data in example 1 and plot the line with the data.

1. Let $A = \begin{bmatrix} 1 & 20 \\ 1 & 25 \\ 1 & 30 \\ 1 & 35 \end{bmatrix}$ and $u = \begin{bmatrix} 84 \\ 74 \\ 62 \\ 50 \end{bmatrix}$.

2. Then $A^T A = \begin{bmatrix} 4 & 110 \\ 110 & 3150 \end{bmatrix}$ and $A^T u = \begin{bmatrix} 270 \\ 7140 \end{bmatrix}$.

Solve the normal equations $(A^T A)x = (A^T u)$:

The best fit line is $y = -2.28x + 130.2$. 

December 12, 2017 17 / 15
Example 1 Revisited

**Ex 1**: Find the least squares error best fit line for the data in example 1 and plot the line with the data.

1. Let  \( A = \begin{bmatrix} 1 & 20 \\ 1 & 25 \\ 1 & 30 \\ 1 & 35 \end{bmatrix} \) and  \( u = \begin{bmatrix} 84 \\ 74 \\ 62 \\ 50 \end{bmatrix} \).

2. Then  \( A^T A = \begin{bmatrix} 4 & 110 \\ 110 & 3150 \end{bmatrix} \) and  \( A^T u = \begin{bmatrix} 270 \\ 7140 \end{bmatrix} \).

Solve the normal equations  \((A^T A)x = (A^T u)\):

\[
\begin{bmatrix} 4 & 110 & 270 \\ 110 & 3150 & 7140 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & 130.2 \\ 0 & 1 & -2.28 \end{bmatrix} \quad b = 130.2 \quad m = -2.28
\]
Example 1 Revisited

**Ex 1**: Find the least squares error best fit line for the data in example 1 and plot the line with the data.

1. Let \( A = \begin{bmatrix} 1 & 20 \\ 1 & 25 \\ 1 & 30 \\ 1 & 35 \end{bmatrix} \) and \( u = \begin{bmatrix} 84 \\ 74 \\ 62 \\ 50 \end{bmatrix} \).

2. Then \( A^T A = \begin{bmatrix} 4 & 110 \\ 110 & 3150 \end{bmatrix} \) and \( A^T u = \begin{bmatrix} 270 \\ 7140 \end{bmatrix} \).

Solve the normal equations \((A^T A)x = (A^T u)\):

\[
\begin{bmatrix} 4 & 110 & \text{270} \\ 110 & 3150 & \text{7140} \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & \text{130.2} \\ 0 & 1 & \text{-2.28} \end{bmatrix} \quad b = 130.2 \quad m = -2.28
\]

The best fit line is \( y = -2.28x + 130.2 \).
Example 1

3. $p = Ax = (84.6, 73.2, 61.8, 50.4)$ is the approximation of $u$. 
Example 1

3. $\mathbf{p} = A\mathbf{x} = (84.6, 73.2, 61.8, 50.4)$ is the approximation of $\mathbf{u}$.

$$|\mathbf{u} - \mathbf{p}|^2 = |(84 - 84.6, 74 - 73.2, 62 - 61.8, 50 - 50.4)|^2$$
$$= |(-0.6, 0.8, 0.2, -0.4)|^2$$
$$= (-0.6)^2 + (0.8)^2 + (0.2)^2 + (-0.4)^2 = 1.2$$
Example 1

3. \( p = Ax = (84.6, 73.2, 61.8, 50.4) \) is the approximation of \( u \).

\[
|u - p|^2 = |(84 - 84.6, 74 - 73.2, 62 - 61.8, 50 - 50.4)|^2
= |(-0.6, 0.8, 0.2, -0.4)|^2
= (-0.6)^2 + (0.8)^2 + (0.2)^2 + (-0.4)^2 = 1.2
\]

Recall: Best fit line \( y = -2.28x + 130.2 \)

The projected demand at price $40 is
Example 1

3. \( \mathbf{p} = A\mathbf{x} = (84.6, 73.2, 61.8, 50.4) \) is the approximation of \( \mathbf{u} \).

\[
|\mathbf{u} - \mathbf{p}|^2 = |(84 - 84.6, 74 - 73.2, 62 - 61.8, 50 - 50.4)|^2 \\
= |(-0.6, 0.8, 0.2, -0.4)|^2 \\
= (-0.6)^2 + (0.8)^2 + (0.2)^2 + (-0.40)^2 = 1.2
\]

Recall: Best fit line \( y = -2.28x + 130.2 \)

The projected demand at price $40 is \(-2.28(40) + 130.2 = 39\) drills.