General Idea:
- There are no solutions to $\mathbf{Ax} = \mathbf{b}$, but we can get an $\mathbf{Ax}$ to be as close as possible to $\mathbf{b}$.
- How?
  - Let's first visualize this:

Key Ideas
- $\mathbf{b} = \mathbf{b}'' + \mathbf{b}^\perp$
- $\mathbf{Ax} \neq \mathbf{b}$ aka $\mathbf{b} \not\in \text{Span}(\mathbf{A})$ aka no linear combinations of $\mathbf{A}$ gets you $\mathbf{b}$.
- $\mathbf{b}'' \in \text{A}^\perp$. There exists some vector $\mathbf{x}$ where $\mathbf{Ax} = \mathbf{b}''$
- We denote $\mathbf{x}^*$ as the vector in subspace $\text{A}$ that will be the closest to $\mathbf{b}$.
  - The closest vector in $\mathbf{A}$ to $\mathbf{b}$ has to be $\mathbf{b}''$ (w/r to $\mathbf{A}$).
  - $\mathbf{x}^* = \mathbf{b}'' = \text{proj}_A(\mathbf{b})$
  - To minimize $\|\mathbf{b} - \mathbf{Ax}\|$, $\mathbf{Ax} = \text{proj}_A(\mathbf{b})$

**Thm**: Let $\mathbf{A}$ be an $n \times m$ matrix and $\mathbf{b} \in \mathbb{R}^n$. The closest vector to be on $\text{im}(\mathbf{A})$ is $\mathbf{x}^* = \text{proj}_{\text{im}(\mathbf{A})}(\mathbf{b})$

**Thm**: $\mathbf{x}^* = \mathbf{b}''$ where $\mathbf{x}^* \in \text{im}(\mathbf{A})$, $\mathbf{b}^\perp \in \text{im}(\mathbf{A})^\perp$, $\mathbf{b} = \mathbf{x}^* + \mathbf{b}^\perp$

Approx Solution to $\mathbf{Ax} = \mathbf{b}$ (aka find $\text{proj}_A(\mathbf{b}) = \mathbf{x}''$)

\[
\mathbf{b} = \mathbf{b}'' + \mathbf{b}^\perp
\]
\[
\mathbf{b} = \mathbf{Ax}^* + \mathbf{b}^\perp
\]
\[
\mathbf{Ax}^* - \mathbf{b} = \mathbf{b}^\perp
\]

\[
\mathbf{A}^T(\mathbf{Ax}^* - \mathbf{b}) = \mathbf{A}^T\mathbf{b}^\perp
\]

\[
\mathbf{A}^T\mathbf{Ax}^* - \mathbf{A}^T\mathbf{b} = 0
\]

**IMPORTANT**

**Thm**: The approx solutions to $\mathbf{Ax} = \mathbf{b}$ are exactly the solutions $\mathbf{A}^T\mathbf{Ax}^* = \mathbf{A}^T\mathbf{b}$.
Approximate Solutions to $Ax = b$

The problem...

Given $n \times n$ matrix $A$, for which $b \in \mathbb{R}^2$ does $Ax = b$ have a solution?

The quick answer is when $b \in \text{Im}(A)$. However, if you very slightly move this vector $b$, then it will lie outside the image causing there to be no solution.

Instead of finding $x$ such that $Ax = b$, find $x$ such that $Ax$ is as close to $b$ as possible (this is finding the approximate solutions).

How?

To find the approximate solutions, solve $A^T A x = A^T b$.

Where...

$A$ = given
$A^T$ = transpose of $A$
$b$ = given
$x$ = what you are trying to find

then...

$$[A^T A] x = [A^T b]$$

solve for $x$

$$[A^T A : A^T b] \xrightarrow{\text{row reduce}} [I : M]$$

$$[M] = [x], \text{ the unique approximate solution}$$

Warning! This assumes $A^T A$ is invertible (in general you may not get a unique approximate solution, but the process is the same).
VIDEO SERIES 3 GRAPH FITTING
Graph fitting is the generalization of this problem:
Given some data set \((x_1, y_1), \ldots, (x_n, y_n)\) \(\in \mathbb{R}^2\) and functions \(f_1, \ldots, f_n : \mathbb{R} \to \mathbb{R}\) find a function \(f : \mathbb{R} \to \mathbb{R}\) such that \(f(x_i) = y_i\) for all \(x_i\) in this data set.

I have a bench of data points & I want to find a function that goes through those data points.

Problem: Consider the following data set

<table>
<thead>
<tr>
<th>(x)</th>
<th>(y)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
</tr>
</tbody>
</table>

Find a quadratic polynomial \(f(x) = Ax^2 + Bx + C\) (i.e. find \(A, B, C \in \mathbb{R}\)) such that \(f(x) = y\) for all \(x\) in this data set.

Solution idea:
Plug in data points to \(f(x) = Ax^2 + Bx + C\) to obtain algebraic relations between \(A, B,\) and \(C\).

Data point: \(0 \quad 0\)
Equation: \(A \cdot 0 + B \cdot 0 + C = f(0) = 0\)

Data point: \(1 \quad 0\)
Equation: \(A \cdot 1 + B \cdot 1 + C = f(1) = 0\)

Data point: \(2 \quad 1\)
Equation: \(A \cdot 2 + B \cdot 2 + C = f(2) = 1\)

To find \(f\), we have to solve this system:

\[
\begin{bmatrix}
0 & 0 & 1 \\
1 & 1 & 1 \\
4 & 2 & 1 \\
\end{bmatrix}
\begin{bmatrix}
A \\
B \\
C \\
\end{bmatrix}
= \begin{bmatrix}
0 \\
0 \\
1 \\
\end{bmatrix}
\]

After the augmenting & row reducing we get:

\[
\begin{bmatrix}
1 & 0 & 0 & \frac{1}{2} \\
0 & 1 & 0 & -\frac{1}{2} \\
0 & 0 & 1 & 0 \\
\end{bmatrix}
\]

Answer is \(f(x) = \frac{1}{2}x^2 - \frac{1}{2}x = \frac{1}{2}(x)(x-1)\)
Problem
Consider the following data set

<table>
<thead>
<tr>
<th>x</th>
<th>y</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
</tr>
</tbody>
</table>

Find a quadratic poly.

\[ f(x) = Ax^2 + Bx + C \]

i.e. find \(A, B, C\) such that
\[ f(0) = 0 \]
\[ f(1) = 0 \]
\[ f(2) = 0 \]
\[ f(3) = 1 \]

for all \((x, y)\) in the data set.

Solution
Plug in data points to obtain linear relations:

\[ A \cdot 0 + B \cdot 0 + C = f(0) = 0 \]
\[ A \cdot 1 + B \cdot 1 + C = f(1) = 0 \]
\[ A \cdot 2^2 + B \cdot 2 + C = f(2) = 0 \]
\[ A \cdot 3^2 + B \cdot 3 + C = f(3) = 1 \]

No quad poly has 3 roots except \(f(c) = 0\)

System Solution

\[
\begin{bmatrix}
0 & 0 & 1 \\
1 & 1 & 1 \\
2 & 2 & 1 \\
3 & 3 & 1 \\
\end{bmatrix} \begin{bmatrix}
A \\
B \\
C \\
\end{bmatrix} = \begin{bmatrix}
0 \\
0 \\
0 \\
1 \\
\end{bmatrix}
\]

To solve, now reduce augmented matrix:

\[
\begin{bmatrix}
0 & 0 & 1 \\
1 & 1 & 1 \\
2 & 2 & 1 \\
3 & 3 & 1 \\
\end{bmatrix} \rightarrow \begin{bmatrix}
1 & 1 & 0 \\
0 & 1 & 0 \\
0 & 0 & 1 \\
0 & 0 & 0 \\
\end{bmatrix}
\]

No solution

How to fix this lack of solution

- Find a quad. poly that best models the data rather than directly passing through it.
- Find a quad. poly such that the distance between the experimental outputs & modeled outputs is minimized.

\[
\begin{bmatrix}
f(c_0) \\
f(c_1) \\
f(c_2) \\
f(c_3) \\
\end{bmatrix}
\]
Modified Problem:
Find a quadratic polynomial such that the distance between

\[ f(0) \quad f(1) \quad f(2) \]

is minimized.

\[ \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \quad \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \quad \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \]

Solve \( A^T A x = A^T b \)

\[
\begin{bmatrix}
0 & 0 & 1 \\
1 & 1 & 1 \\
2 & 2 & 1 \\
3 & 3 & 1 \\
\end{bmatrix}
\begin{bmatrix}
x_0 \\
x_1 \\
x_2 \\
x_3 \\
\end{bmatrix}
= 
\begin{bmatrix}
f(0) \\
f(1) \\
f(2) \\
f(3) \\
\end{bmatrix}
\]

We’re solving \( A^T A x = A^T b \)

\[
A^T A = 
\begin{bmatrix}
98 & 36 & 14 \\
36 & 14 & 6 \\
14 & 6 & 4 \\
\end{bmatrix}
\]

Remark: \( A^T A \) is symmetric across diagonal.

\[
A^T b = 
\begin{bmatrix}
0 & 1 & 4 & 97 \\
0 & 1 & 2 & 3 \\
1 & 1 & 1 & 1 \\
\end{bmatrix}
\begin{bmatrix}
0 \\
0 \\
0 \\
\end{bmatrix}
= 
\begin{bmatrix}
9 \\
5 \\
1 \\
\end{bmatrix}
\]

Solve \( A^T A x = A^T b \)

Form augmented matrix

\[
\begin{bmatrix}
98 & 36 & 14 & 9 \\
36 & 14 & 6 & 5 \\
14 & 6 & 4 & 1 \\
\end{bmatrix}
\begin{bmatrix}
1 & 0 & 0 & 1/4 \\
0 & 1 & 0 & -9/20 \\
0 & 0 & 1 & 1/20 \\
\end{bmatrix}
\]

\[
\begin{bmatrix}
98 & 36 & 14 & 9 \\
36 & 14 & 6 & 5 \\
14 & 6 & 4 & 1 \\
\end{bmatrix}
\begin{bmatrix}
1/4 \\
-9/20 \\
1/20 \\
\end{bmatrix}
\]

\[ f(x) = 4 \frac{x^2 - 9}{20} x + \frac{1}{20} \]
Summary

General Problem

Given a data set
\[(x_1, y_1), \ldots, (x_n, y_n) \in \mathbb{R}^2\]
and functions
\[f_1, \ldots, f_m : \mathbb{R} \to \mathbb{R}\]
find a function \(f : \mathbb{R} \to \mathbb{R}\) such that
\[f = A_1 f_1 + \ldots + A_m f_m\]

General Solution

1. Form matrix \(A = \begin{bmatrix} f(x_1) & f(x_2) & \cdots \\ f(x_2) & f(x_2) & \cdots \\ \vdots & \vdots & \ddots & \vdots \\ f(x_n) & f(x_n) & \cdots & f(x_n) \end{bmatrix}\)

2. Solve \(A x = b\) approximately

Which is: solve \(A^T A x = A^T b\)

\[x = A_1 \begin{bmatrix} \vdots \\ A_m \end{bmatrix}\]

How do you solve Graph Fitting problems

1. Plug in the data sets to get a system of linear equations
2. Solve this system: the rows of the resultant vector are the subsequent variable constants of your function.
3. If the system has no solution then you find the least squared solution

2. Explanation: if the function you plugged into was \(A x^2 + B x + C = y\) & the matrix you got was \(A = \begin{bmatrix} 1 & 1 & 1 \\ 2 & 2 & 2 \\ 3 & 3 & 3 \end{bmatrix}\) then the function is \(2x^2 + 5x + 7 = y\)
When given a data set, we can find a function that most closely describes the values that we observe. We can do this by minimizing the orthogonal distance between actual and theoretical values. Theoretical values are predicted based on the function we are looking for.

This is analogous to solving for an approximate solution to $Ax = b$,

$$A = \begin{bmatrix} f_1(x_1) & \cdots & f_m(x_1) \\ \vdots & \ddots & \vdots \\ f_1(x_n) & \cdots & f_m(x_n) \end{bmatrix}, \quad b = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} \quad \text{and} \quad x \text{ contains the coefficients to the equation we are solving for}.$$
Computing the Orthogonal Complement

Overview:
- Definition: Let $U$ be a subspace of $\mathbb{R}^n$, the orthogonal complement of $V$ is the set $V^\perp = \{ x \in \mathbb{R}^n : x \cdot v = 0 \text{ for all } v \in U \}$.
  - $V^\perp$ is subspace
  - $V^\perp = \text{orthogonal complement of } V$

- Theorem: $\ker(A^T) = V^\perp$, where $A = [v_1, v_2, \ldots, v_m]$ and $V = \text{span}(v_1, \ldots, v_m)$.

Possible solution format:
$G$: Let $V = \text{span}(v_1, v_2, \ldots, v_m)$. Compute $V^\perp$.

Solution:
Let $A = [v_1, v_2, \ldots, v_m]$, compute $\ker(A^T)$. Then $V^\perp = \ker(A^T)$.

Attachment: Note For 3/13:

Computing the Orthogonal Complement

$\Delta$: Let $U$ be a subspace of $\mathbb{R}^n$. The orthogonal complement of $V$ is the set $V^\perp = \{ x \in \mathbb{R}^n : x \cdot v = 0 \text{ for all } v \in U \}$.

- $V^\perp$ is subspace
- $V^\perp = \text{orthogonal complement of } V$

To compute lots of dot products at the same time, organize them as a single matrix multiplication.

Consider $A = [v_1, v_2, \ldots, v_m]$. Then $A^T x = \text{span}(v_1, \ldots, v_m)$.

Ex. Problem: Let $V = \text{span}(\begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix}, \begin{bmatrix} 4 \\ 8 \\ 6 \end{bmatrix})$. Compute $V^\perp$.

Submatrix $A = \begin{bmatrix} 1 & 4 \\ 3 & 8 \\ 2 & 6 \end{bmatrix}$, compute $\ker(A^T)$.

$A^T = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 2 & 3 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ and $\ker(A^T) = V^\perp$.

Unpack: $x_1 + 2x_2 = 0 \Rightarrow x_2 = \frac{x_1}{2}$

$y + 2x_2 = 0 \Rightarrow y = -x_1$

$z = 0$

$\Rightarrow V^\perp = \ker(A^T) = \{ \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} : 2x_2 = x_3 \} = \text{span}(\begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix})$