## The Method of Least Squares

**Theorem 9.6.1 - The Approximation Theorem**

Let $\mathbb{W}$ be a finite dimensional subspace of an inner product space $\mathbb{V}$. If $\vec{v} \in \mathbb{V}$, then

$$\|\vec{v} - \vec{w}\| > \|\vec{v} - \text{proj}_{\mathbb{W}}(\vec{v})\|$$

for all $\vec{w} \in \mathbb{W}$, $\vec{w} \neq \text{proj}_{\mathbb{W}}(\vec{v})$.

### Proof

We have

$$\vec{v} - \vec{w} = \vec{v} - \text{proj}_{\mathbb{W}}(\vec{v}) + \text{proj}_{\mathbb{W}}(\vec{v}) - \vec{w}$$

Observe that

$$\vec{v} - \text{proj}_{\mathbb{W}}(\vec{v}) = \text{perp}_{\mathbb{W}}(\vec{v})$$

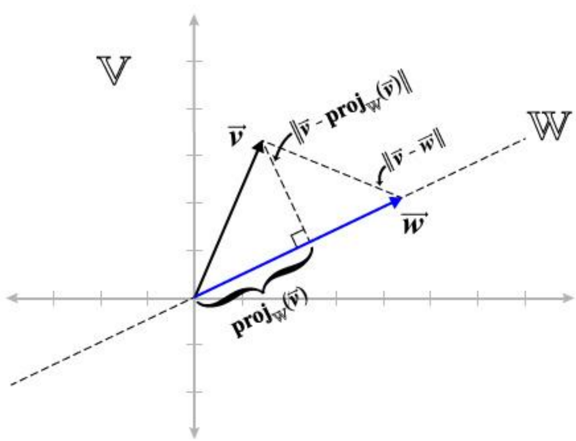
So, $\langle \vec{v} - \text{proj}_{\mathbb{W}}(\vec{v}), \text{proj}_{\mathbb{W}}(\vec{v}) - \vec{w} \rangle = 0$.
Hence, by Theorem 9.2.2 we get

$$\begin{aligned} \|\vec{v} - \vec{w}\|^2 &= \|(\vec{v} - \text{proj}_{\mathbb{W}}(\vec{v})) + (\text{proj}_{\mathbb{W}}(\vec{v}) - \vec{w})\|^2 \\ &= \|\vec{v} - \text{proj}_{\mathbb{W}}(\vec{v})\|^2 + \|\text{proj}_{\mathbb{W}}(\vec{v}) - \vec{w}\|^2 \\ &> \|\vec{v} - \text{proj}_{\mathbb{W}}(\vec{v})\|^2 \end{aligned}$$

□

## The Method of Least Squares

The Approximation Theorem states that the vector in $\mathbb{W}$ that is closest to the vector $\vec{v}$ in $\mathbb{V}$ is the projection of $\vec{v}$ onto $\mathbb{W}$. This is demonstrated geometrically in the picture below:



# The Method of Least Squares

## A Real World Application

Assume a scientist wants to relate two quantities $x$ and $y$, and by theoretical reasons, they know that they are related by an equation of the form, say,

$$y = a_0 + a_1 x + a_2 x^2$$

The scientist wants to determine values of the coefficients $a_0, a_1, a_2$ so that this will best model the real world.

To do this, the scientist performs some experiments involving the two quantities. Every time they do the experiment they will get quantities $y_i$ and $x_i$. Putting this into their equation, they get a linear equation

$$y_i = a_0 + a_1 x_i + a_2 x_i^2$$

A system of linear equations which has more equations than unknowns is said to be overdetermined.

# The Method of Least Squares

## A Real World Application

Lets begin by representing the system in the form $X\vec{a} = \vec{y}$ where $X$ is the coefficient matrix, $\vec{y}$ is the vector of $y$-values and $\vec{a}$ is the vector of unknowns.

We want to find the vector $\vec{a}$ such that $\|X\vec{a} - \vec{y}\|$ is as small as possible. By the Approximation Theorem, this is $\text{proj}_{\text{Col}(X)}(\vec{y})$. Hence, we want to find $\vec{a}$ such that

$$X\vec{a} = \text{proj}_{\text{Col}(X)}(\vec{y})$$

Subtracting both sides of the equation above from $\vec{y}$, we get

$$\vec{y} - X\vec{a} = \vec{y} - \text{proj}_{\text{Col}(X)}(\vec{y}) = \text{perp}_{\text{Col}(X)}(\vec{y}) \in (\text{Col}(X))^{\perp} = \text{Null}(X^T)$$

So,

$$X^T(\vec{y} - X\vec{a}) = \vec{0}$$

Rearranging we get

$$X^T X\vec{a} = X^T\vec{y}$$

We call this system the normal system.

**Note:** The normal system need not have a unique solution. If it does have infinitely many solutions, then each of the solutions will minimize $\|X\vec{a} - \vec{y}\|$.

## The Method of Least Squares

### Example

Determine the vector $\vec{x}$ that is closest to being a solution for the system

$$\begin{aligned} 3x_1 - x_2 &= 4 \\ x_1 + 2x_2 &= 0 \\ 2x_1 + x_2 &= 1 \end{aligned}$$

**Solution**

Let $A = \begin{bmatrix} 3 & -1 \\ 1 & 2 \\ 2 & 1 \end{bmatrix}$, $\vec{y} = \begin{bmatrix} 4 \\ 0 \\ 1 \end{bmatrix}$, and $\vec{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$. We want to find the vector $\vec{x}$ that minimizes $\|A\vec{x} - \vec{y}\|$.

So, we solve the normal system

$$A^T A\vec{x} = A^T\vec{y}$$

We find $A^TA = \begin{bmatrix} 14 & 1 \\ 1 & 6 \end{bmatrix}$ which is invertible, so

$$\vec{x} = (A^TA)^{-1}A^T\vec{y} = \frac{1}{83}\begin{bmatrix} 6 & -1 \\ -1 & 14 \end{bmatrix}\begin{bmatrix} 14 \\ -3 \end{bmatrix} = \begin{bmatrix} 87/83 \\ -56/83 \end{bmatrix}$$

Thus, $x_1 = \frac{87}{83}$ and $x_2 = \frac{-56}{83}$.

We also get $3x_1 - x_2 = 3.82$, $x_1 + 2x_2 = -0.3$, $2x_1 + x_2 = 1.42$.

## The Method of Least Squares

### Example

Determine the vector $\vec{x}$ that minimizes $\|A\vec{x} - \vec{b}\|$ for the system

$$\begin{aligned} 2x_1 + x_2 &= -5 \\ -2x_1 + x_2 &= 8 \\ 2x_1 + 3x_2 &= 1 \end{aligned}$$

**Solution**

Let $A = \begin{bmatrix} 2 & 1 \\ -2 & 1 \\ 2 & 3 \end{bmatrix}$, $\vec{b} = \begin{bmatrix} -5 \\ 8 \\ 1 \end{bmatrix}$, and $\vec{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$. Then, the normal system gives

$$\begin{aligned} A^TA\vec{x} &= A^T\vec{b} \\ \vec{x} &= (A^TA)^{-1}A^T\vec{b} \\ &= \frac{1}{96}\begin{bmatrix} 11 & -6 \\ -6 & 12 \end{bmatrix}\begin{bmatrix} -24 \\ 6 \end{bmatrix} \\ &= \begin{bmatrix} -25/8 \\ 9/4 \end{bmatrix} \end{aligned}$$

Thus, $x_1 = \frac{-25}{8}$ and $x_2 = \frac{9}{4}$.

We also get $2x_1 + x_2 = -4$, $-2x_1 + x_2 = 8.5$, $2x_1 + 3x_2 = 0.5$

## The Method of Least Squares

### Example

Determine the vector $\vec{x}$ that minimizes $\|A\vec{x} - \vec{b}\|$ for the system

$$\begin{aligned} 3x_1 - 2x_2 &= 1 \\ x_1 + 4x_2 &= 1 \\ 2x_1 + x_2 &= -2 \end{aligned}$$

**Solution**

We have $A = \begin{bmatrix} 3 & -2 \\ 1 & 4 \\ 2 & 1 \end{bmatrix}$, $\vec{b} = \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix}$, and $\vec{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$. Then, solving the normal system we get

$$\vec{x} = (A^TA)^{-1}A^T\vec{b} = \begin{bmatrix} 1/14 & 0 \\ 0 & 1/21 \end{bmatrix} \begin{bmatrix} 3 & 1 & 2 \\ -2 & 4 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ -2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

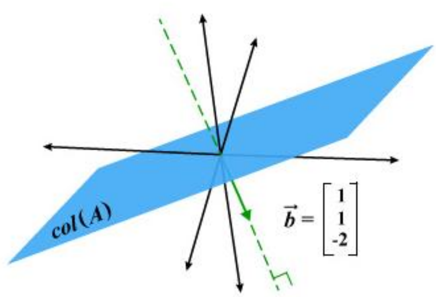
## The Method of Least Squares

### Example

Determine the vector $\vec{x}$ that minimizes $\|A\vec{x} - \vec{b}\|$ for the system

$$\begin{aligned} 3x_1 - 2x_2 &= 1 \\ x_1 + 4x_2 &= 1 \\ 2x_1 + x_2 &= -2 \end{aligned}$$

**Solution**




The method we have derived in this lecture for finding the best approximation to a solution of an overdetermined system is called the method of least squares.