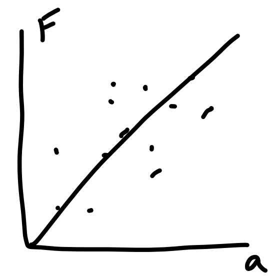
4/25/25

(substitute lecture)

Linear Regression (least squares)

$\vec{x} = (x_1, \dots, x_n)$ input data that you expect to satisfy a linear relation, i.e. $y = c_1 \vec{x}^1 + \dots + c_n \vec{x}^n$ where c_i are constants that "best fit the data and y is output



have bunch of points x_i

$$A = \begin{bmatrix} \vec{x}^1 \\ \vdots \\ \vec{x}^m \end{bmatrix} \quad m \times n \text{ matrix} \longrightarrow \text{want to find } \vec{c} = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} \text{ such that } A\vec{c} = \vec{y}$$

each data point x_i has some output y_i

We can find the best such $\vec{c}$ using orthogonal projection onto $W = \mathrm{Col}(A)$, $\vec{y} = \vec{y}_W + \vec{y}_\perp$, solve $A\vec{c} = \vec{y}_W$, then $\|\vec{y} - A\vec{c}\|$ is as small as possible

General Process

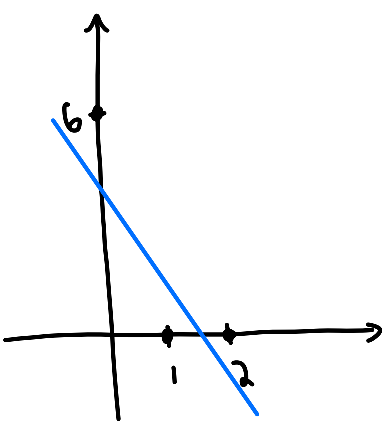
Have data points $(\vec{x}^1, y_1), (\vec{x}^2, y_2), \ldots, (\vec{x}^m, y_m)$

1) Write $A = \begin{bmatrix} \vec{x}^1 \\ \vdots \\ \vec{x}^m \end{bmatrix}$, $\vec{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix}$, $W = \text{col}(A)$

2) Solve $A^T A \vec{c} = A^T \vec{y}$ - which we know has a solution by orthogonal projection

3) The solution(s) $\vec{c} = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}$ are called least squares weights

Ex: Find the line best approximating pts. $(0,6), (1,0), (2,0)$



obviously line doesn't go through origin

↳ trick: use dummy variable $x_{n+1} = \vec{1}$

$$\Rightarrow c_1 \vec{x}_1 + \ldots + c_n \vec{x}_n + c_{n+1}(1) = \vec{y}$$

in this case: $c_1 \vec{x}_1 + c_2(1) = \vec{y}$

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 1 \\ 2 & 1 \end{bmatrix} \quad \vec{y} = \begin{bmatrix} 6 \\ 0 \\ 0 \end{bmatrix}$$

We need to solve $A^T A \vec{c} = A^T y$

$$A^T A = \begin{bmatrix} 0 & 1 & 2 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 1 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} 5 & 3 \\ 3 & 3 \end{bmatrix} \qquad A^T y = \begin{bmatrix} 0 & 1 & 2 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 6 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 6 \end{bmatrix}$$

$$\longrightarrow \begin{bmatrix} 5 & 3 \\ 3 & 3 \end{bmatrix} \vec{c} = \begin{bmatrix} 0 \\ 6 \end{bmatrix} \longrightarrow \left[\begin{array}{cc|c} 5 & 3 & 0 \\ 3 & 3 & 6 \end{array}\right] \xrightarrow{\text{RREF}} \left[\begin{array}{cc|c} 1 & 0 & -3 \\ 0 & 1 & 5 \end{array}\right]$$

$$\longrightarrow \begin{array}{l} c_1 = -3 \\ c_2 = 5 \end{array} \longrightarrow y = -3x + 5 \ !$$

Orthogonal vectors

<u>Defn</u>: A set of vectors $\{\vec{v}_1, \dots, \vec{v}_m\} \subseteq \mathbb{R}^n$ is said to be <u>orthogonal</u> $\iff$ $\vec{v}_i \cdot \vec{v}_j = \vec{0}$ whenever $i \neq j$.

Further, the set is said to be <u>orthonormal</u> if it is orthogonal and $\|\vec{v}_i\| = \sqrt{\vec{v}_i \cdot \vec{v}_i} = 1$ for all $i = 1, \dots, m$.

<u>Ex</u>: The vectors $e_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$, $e_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$, $e_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$ form an orthonormal set in $\mathbb{R}^3$

<u>Ex</u>: $\mathcal{B} = \left\{ \underbrace{\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}}_{\vec{v}_1}, \underbrace{\begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}}_{\vec{v}_2}, \underbrace{\begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}}_{\vec{v}_3} \right\}$

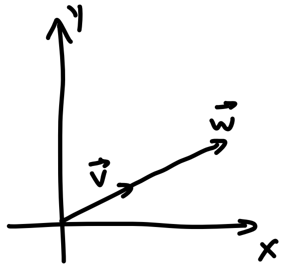
not orthonormal ($\vec{v}_1 \cdot \vec{v}_1 = 3$, etc.)

but <u>is</u> orthogonal ($\vec{v}_i \cdot \vec{v}_j = 0$ for all $i \neq j$)

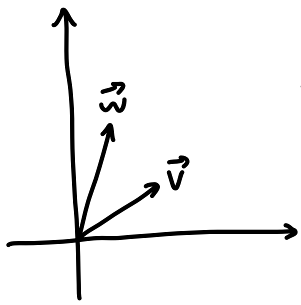
$$\mathcal{B}' = \left\{ \frac{1}{\sqrt{3}}\begin{bmatrix}1\\1\\1\end{bmatrix}, \frac{1}{\sqrt{6}}\begin{bmatrix}1\\-2\\1\end{bmatrix}, \frac{1}{\sqrt{2}}\begin{bmatrix}1\\0\\-1\end{bmatrix} \right\}$$

However, we can easily make $\mathcal{B}$ an orthonormal vector by normalizing each of its vectors, i.e. dividing by their length

Note:



$\vec{v} = (x, y)$
$\vec{w} = (x', y')$ are dependent $\iff$ $\vec{w} = \lambda \vec{v}$, $\lambda \in \mathbb{R}$



easy to see that linear independence does not equal orthogonality always

But, if orthogonal, always linearly independent

Thm

(no $v_i = \vec{0}$)

Pf. Let $\{\vec{v}_1, \ldots, \vec{v}_n\} \subseteq \mathbb{R}^n$ be an orthonal set. Suppose we have $c_1\vec{v}_1 + \ldots + c_n\vec{v}_n = 0$. WTS $c_i = 0$ for all i:

$$(c_1\vec{v}_1 + \ldots + c_i\vec{v}_i + \ldots + c_n\vec{v}_n) \cdot v_i = 0$$

$$c_1\underbrace{(v_1 \cdot v_i)}_{=0} + \ldots + c_i(v_i \cdot v_i) + c_n\underbrace{(v_n \cdot v_i)}_{=0} = 0$$

$$c_i \|\vec{v}_i\|^2 = 0 \quad , \quad \|\vec{v}_i\|^2 \neq 0$$

$$\Rightarrow c_i = 0$$

Since i was arbitrary, we are done. $\square$