Machine Learning

2 – Multiple Linear Regression

Avisek Gupta
Postdoctoral Fellow, IAI, TCG CREST
avisek003@gmail.com

August 16, 2022

Recap: Simple Linear Regression

Given data $\{(x_1, y_1), (x_2, y_2), ..., (x_n, y_n)\}$, $x_i, y_i \in \mathbb{R}$, we assume the data was sampled from a function:

$$y_i = f(x_i) + \epsilon_i$$

We wish to estimate this function as $\hat{f}(x_i)$ in order to predict,

$$\hat{y_i} = \hat{f}(x_i) \ .$$

In simple linear regression, we wish to estimate a line $\hat{y_i} = \hat{\beta}_0 + \hat{\beta}_1 x_i$. We can consider a choice of a *loss function* $(y_i - \hat{y_i})^2$ that can inform us how well a certain estimate of $\hat{\beta}_0, \hat{\beta}_1$ is for the given data $\{(x_1, y_1), (x_2, y_2), ..., (x_n, y_n)\}$.

Using calculus, we can arrive at the following *closed-form expressions* for $\hat{\beta}_0, \hat{\beta}_1$:

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$$

$$\hat{\beta}_1 = \frac{\sum_i x_i y_i - \bar{y} \sum_i x_i}{\sum_i x_i^2 - \bar{x} \sum_i x_i}$$

Multiple Linear Regression

Given data $\{(\mathbf{x_1}, y_1), (\mathbf{x_2}, y_2), ..., (\mathbf{x_n}, y_n)\}, \mathbf{x_i} \in \mathbb{R}^d, y_i \in \mathbb{R}$, we assume the data was sampled from a function:

$$y_i = f(\mathbf{x_i}) + \epsilon_i$$

We wish to estimate this function as $\hat{f}(\mathbf{x_i})$ in order to predict,

$$\begin{aligned} \hat{y_i} &= \hat{f}(\mathbf{x_i}) \\ &= \hat{\beta}_0 + \hat{\beta}_1 x_{i1} + \hat{\beta}_2 x_{i2} + ... + \hat{\beta}_d x_{id}. \end{aligned}$$

We consider a *loss function* $(y_i - \hat{y_i})^2$ that informs us on how well a certain estimate of $\hat{\beta}_0, \hat{\beta}_1, ...\hat{\beta}_d$ is for the given data.

How can we estimate $\hat{\beta}_0, \hat{\beta}_1, ...\hat{\beta}_d$?

- Obtain closed-form solutions of *d* variables…
- With the help of linear algebra, use *projections*

Simple Linear Regression

$$\begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix} \begin{bmatrix} w_0 \\ w_1 \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

Multiple Linear Regression

$$\begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1d} \\ 1 & x_{21} & x_{22} & \cdots & x_{2d} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{nd} \end{bmatrix} \begin{bmatrix} w_0 \\ w_1 \\ \vdots \\ w_d \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

Multiple Linear Regression

$$\begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1d} \\ 1 & x_{21} & x_{22} & \cdots & x_{2d} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{nd} \end{bmatrix} \begin{bmatrix} w_0 \\ w_1 \\ \vdots \\ w_d \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

- How can we solve a system of equations?

$$A\mathbf{x} = b$$

$$\mathbf{x} = A^{-1}b$$ (if A is a square matrix)

- If A is not a square matrix?

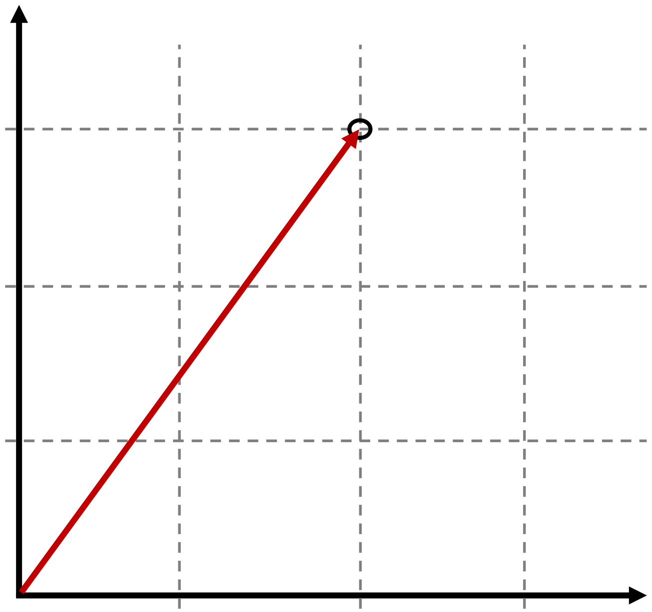
$\mathbf{x}_i$: Vectors in a space

$$\mathbf{x_i} = \{x_{i1}, x_{i2}, ..., x_{id}\}$$

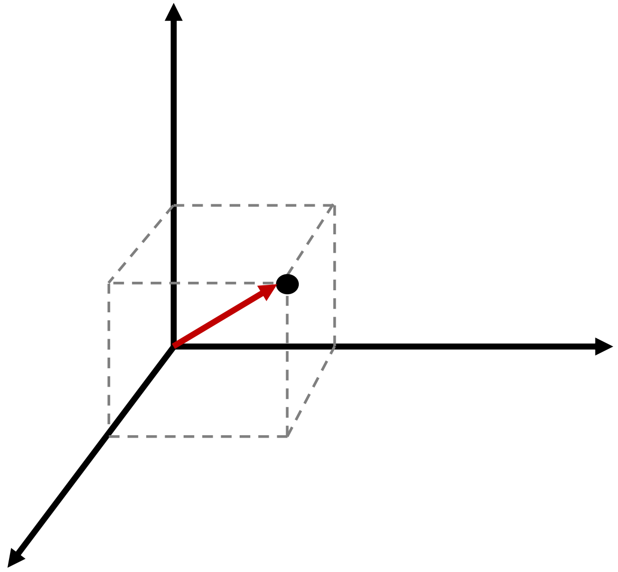
$\mathbf{x_i}$ is a point in a *d*-dimensional space

$\mathbf{x_i}$ is a vector in a *d*-dimensional space

$$\mathbf{x} = \{2, 3\}$$



$$\mathbf{x} = \{1, 1, 1\}$$

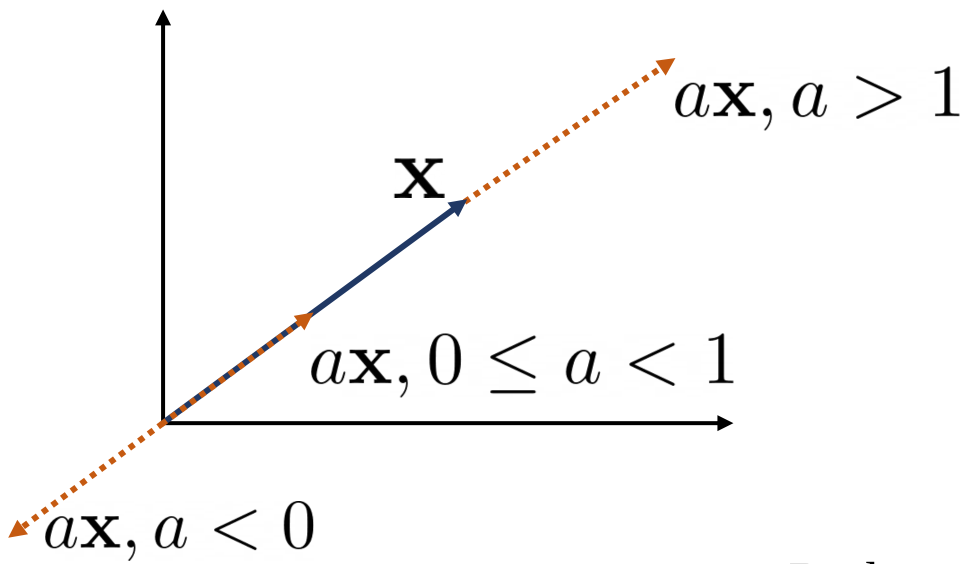


Linear Combination of vectors

- Scalar Multiple of a vector

$$a\mathbf{x},\, a \in \mathbb{R}$$

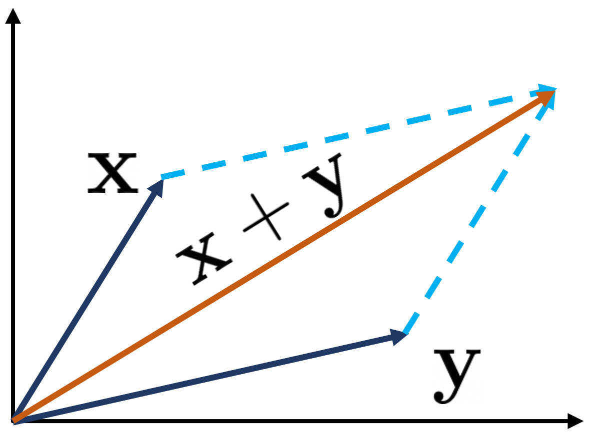
- same direction, different magnitude



- Addition of two vectors

$$\mathbf{x} + \mathbf{y}$$

- vector in the <u>plane</u> of **x** and **y**



- Both operations together?

$$a\mathbf{x} + b\mathbf{y},\;\; a, b \in \mathbb{R}$$

- Generate *<u>any</u>* vector in the plane of $\mathbf{x}$ and $\mathbf{y}$

Span of a set of vectors: Set of all linear combinations of the vectors

What is -

- The span of a single vector
- The span of two vectors
- The span of three vectors
- …

Linear Independence: For a sequence of vectors $\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_d$,

$$a_1\mathbf{x}_1 + a_2\mathbf{x}_2 + ... + a_d\mathbf{x}_d = \mathbf{0} \implies a_1 = a_2 = ... = a_d = 0$$

- What is the minimum number of *linearly independent* vectors required to span a space containing *d*-dimensional vectors?

Basis: Set of linearly independent vectors that span a vector space

Distance between two vectors:

$$d(\mathbf{x}, \mathbf{y}) = ||\mathbf{x} - \mathbf{y}|| = \sqrt{(\mathbf{x} - \mathbf{y})^T (\mathbf{x} - \mathbf{y})}$$

Norm of a vector:

$$||\mathbf{x}|| = \sqrt{\mathbf{x}^T \mathbf{x}} = \sqrt{x_1^2 + x_2^2 + ... + x_d^2}$$

Angle between vectors:

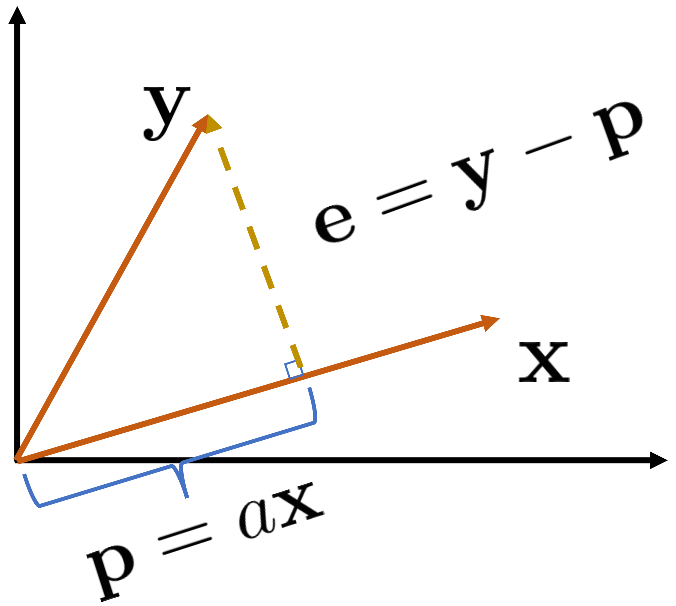
$$\cos\theta = \frac{\mathbf{x}^T \mathbf{y}}{||\mathbf{x}|| ||\mathbf{y}||}$$

Orthogonality:

$$\mathbf{x}^T \mathbf{y} = 0$$

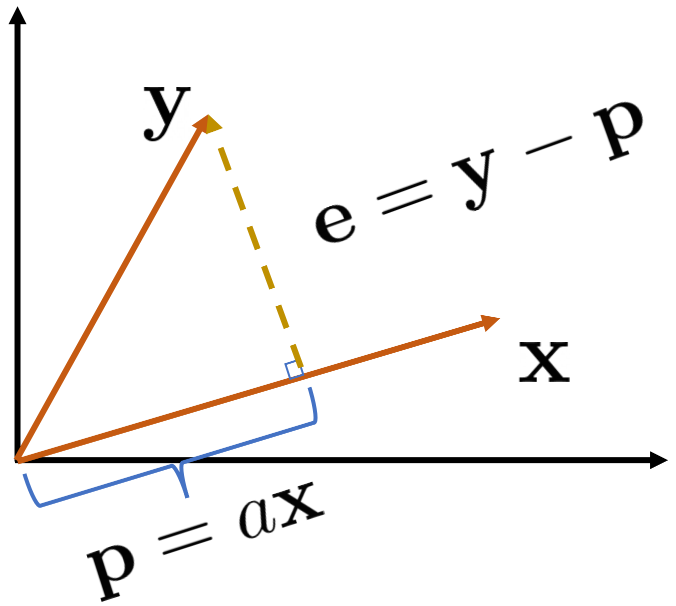
Orthogonal Projection of a vector

- We wish to find the orthogonal projection of a vector **y** on to a vector **x**.



Orthogonal Projection of a vector

- We wish to find the orthogonal projection of a vector **y** on to a vector **x**.



$$\mathbf{x}^T\mathbf{e} = 0$$

$$\implies \mathbf{x}^T(\mathbf{y} - a\mathbf{x}) = 0$$

$$\implies a = \frac{\mathbf{x}^T\mathbf{y}}{\mathbf{x}^T\mathbf{x}}$$

And so the projected vector is:

$$\mathbf{p} = \frac{\mathbf{x}^T\mathbf{y}}{\mathbf{x}^T\mathbf{x}}\mathbf{x}$$

System of equations: Row view

$$\begin{bmatrix} 1 & 3 & 2 \\ 1 & 4 & 1 \\ 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} w_0 \\ w_1 \\ w_2 \end{bmatrix} = \begin{bmatrix} 3 \\ 4 \\ 2 \end{bmatrix}$$

System of equations: Row view

$$\begin{bmatrix} 1 & 3 & 2 \\ 1 & 4 & 1 \\ 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} w_0 \\ w_1 \\ w_2 \end{bmatrix} = \begin{bmatrix} 3 \\ 4 \\ 2 \end{bmatrix}$$

Equation of a plane

System of equations: Row view

$$\begin{bmatrix} 1 & 3 & 2 \\ 1 & 4 & 1 \\ 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} w_0 \\ w_1 \\ w_2 \end{bmatrix} = \begin{bmatrix} 3 \\ 4 \\ 2 \end{bmatrix}$$

Intersection of two planes

System of equations: Row view

$$\begin{bmatrix} 1 & 3 & 2 \\ 1 & 4 & 1 \\ 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} w_0 \\ w_1 \\ w_2 \end{bmatrix} = \begin{bmatrix} 3 \\ 4 \\ 2 \end{bmatrix}$$

Intersection of three planes

System of equations: Row view

$$\begin{bmatrix} 1 & 3 & 2 \\ 1 & 4 & 1 \\ 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} w_0 \\ w_1 \\ w_2 \end{bmatrix} = \begin{bmatrix} 3 \\ 4 \\ 2 \end{bmatrix}$$

Intersection of three planes

System of equations: Column view

$$w_0 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + w_1 \begin{bmatrix} 3 \\ 4 \\ 2 \end{bmatrix} + w_2 \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix} = \begin{bmatrix} 3 \\ 4 \\ 2 \end{bmatrix}$$

System of equations: Row view

$$\begin{bmatrix} 1 & 3 & 2 \\ 1 & 4 & 1 \\ 1 & 2 & 3 \end{bmatrix} \begin{bmatrix} w_0 \\ w_1 \\ w_2 \end{bmatrix} = \begin{bmatrix} 3 \\ 4 \\ 2 \end{bmatrix}$$

Intersection of three planes

System of equations: Column view

$$w_0 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + w_1 \begin{bmatrix} 3 \\ 4 \\ 2 \end{bmatrix} + w_2 \begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix} = \begin{bmatrix} 3 \\ 4 \\ 2 \end{bmatrix}$$

Find a linear combination of the column vectors

Solving Multiple Linear Regression

$$\begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1d} \\ 1 & x_{21} & x_{22} & \cdots & x_{2d} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{nd} \end{bmatrix} \begin{bmatrix} w_0 \\ w_1 \\ \vdots \\ w_d \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

n-dimensional vectors

n-dimensional vector

Solving Multiple Linear Regression

$$\begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1d} \\ 1 & x_{21} & x_{22} & \cdots & x_{2d} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{nd} \end{bmatrix} \begin{bmatrix} w_0 \\ w_1 \\ \vdots \\ w_d \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

n-dimensional vectors

Linear combination of d vectors
$d < n$

n-dimensional vector

Solving Multiple Linear Regression

$$\begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1d} \\ 1 & x_{21} & x_{22} & \cdots & x_{2d} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{nd} \end{bmatrix} \begin{bmatrix} w_0 \\ w_1 \\ \vdots \\ w_d \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

n-dimensional vectors

Linear combination of d vectors
d < n

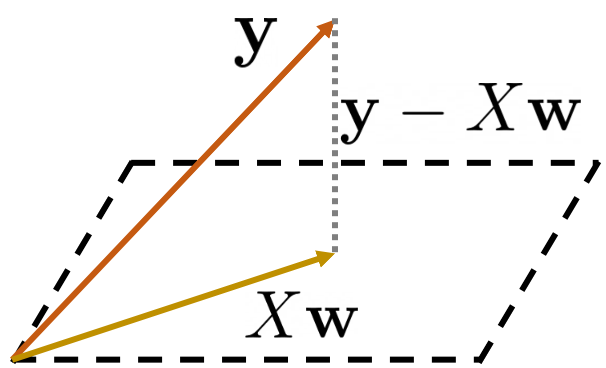
n-dimensional vector

- **y** may not lie in the span of the column vectors
- Can we project **y** on to the *subspace* spanned by the column vectors?

Solving Multiple Linear Regression

$$\begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1d} \\ 1 & x_{21} & x_{22} & \cdots & x_{2d} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{nd} \end{bmatrix} \begin{bmatrix} w_0 \\ w_1 \\ \vdots \\ w_d \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

$$X\mathbf{w} = \mathbf{y}$$

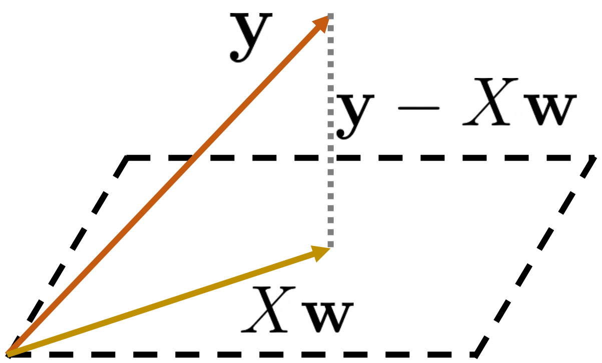


$$X^T(\mathbf{y} - X\mathbf{w}) = \mathbf{0}$$

=> $\mathbf{w} = (X^T X)^{-1} X^T \mathbf{y}$

Solving Multiple Linear Regression

$$\begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1d} \\ 1 & x_{21} & x_{22} & \cdots & x_{2d} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{nd} \end{bmatrix} \begin{bmatrix} w_0 \\ w_1 \\ \vdots \\ w_d \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$



$$X\mathbf{w} = \mathbf{y}$$

$$X^T(\mathbf{y} - X\mathbf{w}) = \mathbf{0}$$

=> $\mathbf{w} = (X^TX)^{-1}X^T\mathbf{y}$

Estimated w

=> $X\mathbf{w} = X(X^TX)^{-1}X^T\mathbf{y}$

Projection of y onto the subspace spanned by the columns of X