Mathematical Foundations for Machine Learning

Linear Algebra and Linear Regression

Instructor: Pat Virtue

Today

Linear algebra

- Primer and highlighting key aspects

Linear regression

- Calculus and optimization

Linear Algebra

Linear Algebra

Convenient and concise way to work with math involving linear operations

$$3\theta_1 + 0\theta_2 + 5\theta_3 = 36$$
$$4\theta_1 + 3\theta_2 + 2\theta_3 = 46$$
$$2\theta_1 + 2\theta_2 + 1\theta_3 = 25$$

$$V\boldsymbol{\theta} = \boldsymbol{u} \qquad V = \begin{bmatrix} 3 & 0 & 5 \\ 4 & 3 & 2 \\ 2 & 2 & 1 \end{bmatrix} \quad \boldsymbol{u} = \begin{bmatrix} 36 \\ 46 \\ 25 \end{bmatrix} \quad \boldsymbol{\theta} = \begin{bmatrix} \theta_1 \\ \theta_2 \\ \theta_3 \end{bmatrix}$$

Possible way to solve: $\boldsymbol{\theta} = V^{-1}\boldsymbol{u}$

(Actually a pretty unstable way to solve in general. More on this later.)

Matrix Multiplication

$A \in \mathbb{R}^{M \times N}$

$B \in \mathbb{R}^{M \times K}$

$C \in \mathbb{R}^{K \times N}$

$A = BC$

The entry of A in row m and column n is:

$$A_{m,n} = \sum_{k=1}^{K} B_{m,k} C_{k,n}$$

i.e. multiply the elements of row m of by the elements of column n of C, and then sum those products

Width of B must match height of C

Poll 1

What is the resulting dimension of Z, where $Z = \boldsymbol{u}\boldsymbol{v}^T$ and $\boldsymbol{u} \in \mathbb{R}^{7\times 1}$ and $\boldsymbol{v} \in \mathbb{R}^{7\times 1}$?

Select ALL that apply

$A.\ \ Z \in \mathbb{R}$

$B.\ \ Z \in \mathbb{R}^{49}$

$C.\ \ Z \in \mathbb{R}^{7\times 1}$

$D.\ \ Z \in \mathbb{R}^{1\times 7}$

$E.\ \ Z \in \mathbb{R}^{1\times 1}$

$F.\ \ Z \in \mathbb{R}^{7\times 7}$

G. I have no idea

Vectors

$\mathbf{u} \in \mathbb{R}^{M \times 1}$ Column vector of length M

$\mathbf{v} \in \mathbb{R}^{1 \times M}$ Row vector of length M

$\mathbf{z} \in \mathbb{R}^{M}$ Ambiguous. In this course, assume column vector unless stated otherwise.

$$\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_M \end{bmatrix}$$

Vectors

$\mathbf{u} \in \mathbb{R}^M$

$\mathbf{v} \in \mathbb{R}^M$ Note these are both column vectors

Multiplication of vectors of the same length

Dot product (inner product):

- $y = \mathbf{u} \cdot \mathbf{v} = \mathbf{u}^T \mathbf{v} = \sum_{i=1}^{M} u_i \, v_i \qquad y \in \mathbb{R}$
- A dot product is an inner product (inner product is a more general term)

Outer product (inner product):

- $Y = \mathbf{u} \otimes \mathbf{v} = \mathbf{u}\mathbf{v}^T \qquad Y \in \mathbb{R}^{M \times M} \qquad \mathrm{Y}_{i,j} = u_i v_j$

Vectors

$\mathbf{u} \in \mathbb{R}^M$

What happens when we do the dot product of a vector with itself?

Vectors

$\mathbf{u} \in \mathbb{R}^M$

Vector magnitude

$$|\mathbf{u}| = \sqrt{\sum_i^M u_i^2} = \left(\sum_i^M u_i^2\right)^{1/2} = (\ ?\ )^{1/2}$$

L2 norm (Euclidean norm) (More norms later)

$$\|\mathbf{u}\|_2 = \sqrt{\sum_i^M u_i^2} = \left(\sum_i^M u_i^2\right)^{1/2} = \left(\boldsymbol{u}^T \boldsymbol{u}\right)^{1/2}$$

L2 norm squared

$$\|\mathbf{u}\|_2^2 = \sum_i^M u_i^2 = \boldsymbol{u}^T \boldsymbol{u}$$

Vectors

$\mathbf{u} \in \mathbb{R}^M$

$\mathbf{v} \in \mathbb{R}^M$

$\alpha \in \mathbb{R}$

Scalar multiplication

$$\alpha\mathbf{u} = \begin{bmatrix} \alpha u_1 \\ \alpha u_2 \\ \vdots \\ \alpha u_M \end{bmatrix}$$

Vector addition

$$\mathbf{u} + \mathbf{v} = \begin{bmatrix} u_1 + v_1 \\ u_2 + v_2 \\ \vdots \\ u_M + v_M \end{bmatrix}$$

Notation alert!

Vectors

$\mathbf{u}, \mathbf{v} \in \mathbb{R}^M$

Linear combination of vectors

$$\mathbf{z} = \alpha\mathbf{u} + \beta\mathbf{v} = \begin{bmatrix} \alpha u_1 + \beta v_1 \\ \alpha u_2 + \beta v_2 \\ \vdots \\ \alpha u_M + \beta v_M \end{bmatrix} \text{ for any } \alpha, \beta \in \mathbb{R}$$

Notation alert!

Span

Set of all linear combination of a set of vectors

Given a set of vectors $\mathcal{S} = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3\}$, where $\boldsymbol{v}_i \in \mathbb{R}^M \ \forall\ i$

$$span(\mathcal{S}) = \{\, w_1\boldsymbol{v}_1 + w_2\boldsymbol{v}_2 + w_3\boldsymbol{v}_3 \mid w_i \in \mathbb{R}\,\}$$

Span of a set of vectors is an example of a vector space. Vector space is a more general term.

Span, Dimensionality, and Rank

Span, Dimensionality, and Rank

Linear Regression

Fitting a linear model

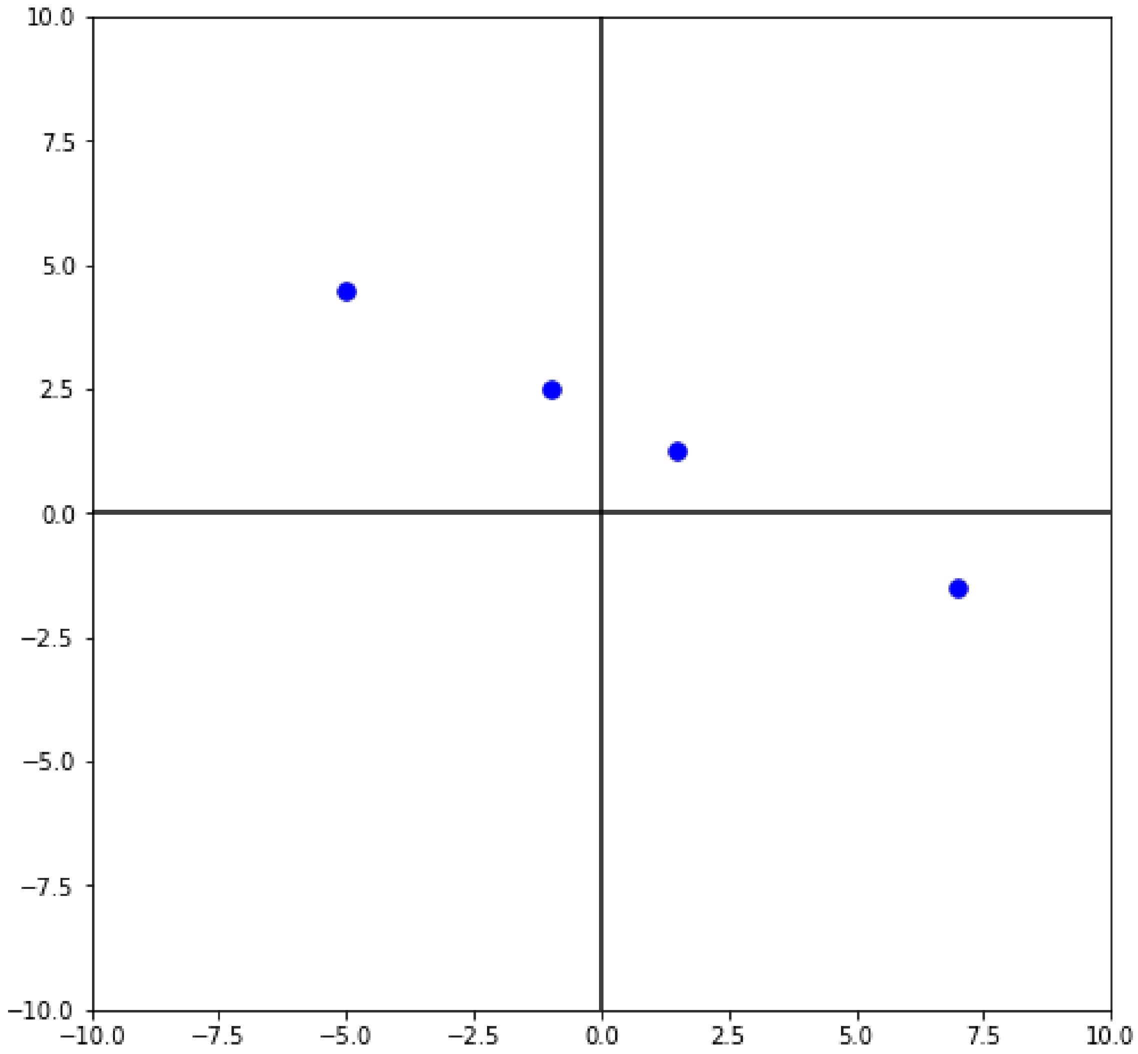
$$\mathcal{D} = \left\{\left(x^{(i)}, y^{(i)}\right)\right\}_{i=1}^{4}$$

$$= \{(-1,\ \ 2.5),$$

$$(7,\ \ -1.5),$$

$$(-5,\ \ 4.5),$$

$$(1.5,\ \ 1.25)\}$$



Jupyter Demo