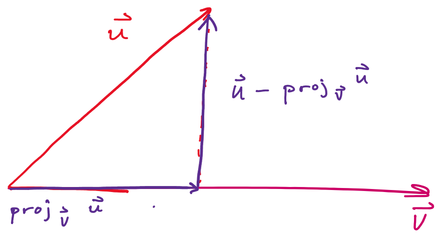
Projection Formula Derivation

November 2, 2023 9:58 PM

suspicious

Definition: Projection of $\vec{u}$ onto $\vec{v}$ $\quad \text{proj}_{\vec{v}}\vec{u} = \dfrac{\vec{u} \times \vec{v}}{\|\vec{v}\|^2}$ $(\vec{v} \neq 0)$



Derivation:

$\because \text{proj}_{\vec{v}}\vec{u} \parallel \vec{v}$, $\exists$ a scalar t s.t. $\text{proj}_{\vec{v}}\vec{u} = t\vec{v}$

we know that

$$\vec{v} \cdot (\vec{u} - \text{proj}_{\vec{v}}\vec{u}) = 0$$

and that

$$\vec{v} \cdot (\vec{u} - t\vec{v}) = 0$$

Solve for t.

$$\vec{u} \cdot \vec{v} = (\vec{v} \cdot \vec{v})t$$

$$\therefore t = \frac{\vec{u} \cdot \vec{v}}{\|v\|^2}$$

$$\because \text{proj}_{\vec{v}}\vec{u} = t\vec{v}$$

$$\therefore \text{proj}_{\vec{v}}\vec{u} = \underbrace{\frac{\vec{u} \cdot \vec{v}}{\|\vec{v}\|^2}}_{\text{scalar}}\vec{v}$$

Cross product geometry

Thursday, November 2, 2023 9:59 PM

The Cross Product

Definition: $\vec{v}_1 \times \vec{v}_2$, where $\vec{v}_1 = \begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix}$ and $\vec{v}_2 = \begin{bmatrix} x_2 \\ y_2 \\ z_2 \end{bmatrix}$

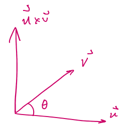
$$\vec{v}_1 \times \vec{v}_2 = \begin{bmatrix} y_1 z_2 - z_1 y_2 \\ -(x_1 z_2 - z_1 x_2) \\ x_1 y_2 - y_1 x_2 \end{bmatrix}$$

Geometric Interpretation

- $\vec{v} \times \vec{u}$ is a vector orthogonal to $\vec{v}$ and $\vec{u}$ (apply RHR)

$\|\vec{v} \times \vec{u}\|$ is the area of a parallelogram formed by $\vec{u}$ and $\vec{v}$.

Property #1: $\|\vec{u} \times \vec{v}\| = \|\vec{u}\| \, \|\vec{v}\| \sin\theta$



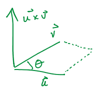
↳ Use Lagrange's Identity (you can prove Lagrange's Identity - it's just tedious)

$$\|\vec{u} \times \vec{v}\|^2 = \|\vec{u}\|^2 \|\vec{v}\|^2 - (\vec{u} \cdot \vec{v})^2$$
$$= \|\vec{u}\|^2 \|\vec{v}\|^2 - \|u\|^2 \|v\|^2 \cos^2\theta$$
$$= \|\vec{u}\|^2 \|\vec{v}\|^2 (1 - \cos^2\theta)$$
$$= \|\vec{u}\|^2 \|\vec{v}\|^2 \sin^2\theta \rightarrow \text{since}$$

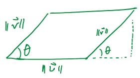
$\sin^2\theta + \cos^2\theta = 1$
$\therefore \sin^2\theta = 1 - \cos^2\theta$

$$\therefore \|\vec{u} \times \vec{v}\|^2 = \|\vec{u}\| \|\vec{v}\| \sin\theta$$

Property 2: $\|\vec{u} \times \vec{v}\|$ is equal to the area of the parallelogram formed by $\vec{u}$ and $\vec{v}$

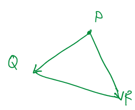


Area of parallelogram:

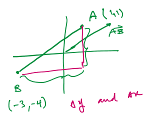


$A = \text{base} \times \text{height}$
$= \|\vec{u}\| \, \|\vec{v}\| \sin\theta$

Ex: Find $A_{\Delta PQR}$ with vertices $P(1,1,-3)$, $Q(3,5,1)$, $R(2,-1,7)$



$\vec{PQ} = \begin{bmatrix} 3-1 \\ 5-1 \\ 1+3 \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \\ 4 \end{bmatrix}$



$\begin{bmatrix} 2 \\ 4 \\ 4 \end{bmatrix} \times \begin{bmatrix} 1 \\ -2 \\ 10 \end{bmatrix}$

$2(-2) - 4$

x	y	z
2	4	4
1	-2	10

$\begin{bmatrix} 4(10) - 2(-4) \\ -2(10) + 4 \\ -4 - 4 \end{bmatrix} = \begin{bmatrix} 48 \\ -16 \\ -8 \end{bmatrix}$

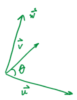
$A_{\Delta PQR} = \frac{1}{2} A_{\square}$
$= \frac{1}{2} \|\vec{PQ} \times \vec{PR}\|$
$= \frac{1}{2}(8\sqrt{41})$
$= 4\sqrt{41}$

$\vec{PR} = \begin{bmatrix} 2-1 \\ -1-1 \\ 7-(-3) \end{bmatrix} = \begin{bmatrix} 1 \\ -2 \\ 10 \end{bmatrix}$

$\vec{PQ} \times \vec{PR} = [$

$\|\vec{PQ} \times \vec{PR}\| = \sqrt{48^2 + 16^2 + 8^2}$
$= 8\sqrt{41}$

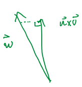
Property 3: The volume of a paralleloid is $\vec{w} \cdot (\vec{u} \cdot \vec{v})$



$V = A_{base} \times \text{height}$ (perpendicular from the base)

$A_{base} = \|\vec{u} \times \vec{v}\|$

Find height
↳ Perpendicular from the base
↳ Cross product!



Projection of $\vec{w}$ onto $\vec{u} \times \vec{v}$ is the vector representing the height of the paralelepiped

$$\left\| \left(\frac{\vec{w} \cdot (\vec{u} \times \vec{v})}{\|\vec{u} \times \vec{v}\|^2} \right) \vec{u} \times \vec{v} \right\| \text{ a vector}$$

$\underbrace{\hphantom{xxxxxxxx}}_{\text{projection formula}}$

Rearrange!

$$\left\| \frac{\vec{w} \cdot (\vec{u} \times \vec{v})}{\|\vec{u} \times \vec{v}\|} \cdot \underbrace{\frac{\vec{u} \times \vec{v}}{\|\vec{u} \times \vec{v}\|}}_{\text{unit vector}} \right\|$$

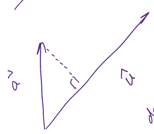
so

$$h = \left\| \frac{\vec{w} \cdot (\vec{u} \times \vec{v})}{\|\vec{u} \times \vec{v}\|} \right\|$$

$V_T = A_{base} h$
$= \cancel{\|\vec{u} \times \vec{v}\|} \left\| \frac{\vec{w} \cdot (\vec{u} \times \vec{v})}{\cancel{\|\vec{u} \times \vec{v}\|}} \right\|$

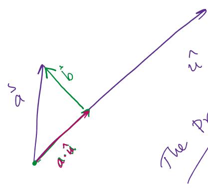
$$\therefore V_T = \underbrace{\| \vec{w} \cdot (\vec{u} \times \vec{v}) \|}_{\text{depends on how you draw your parallepdd}}$$

Dot Product

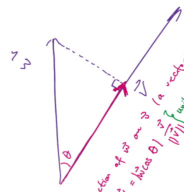


where $\hat{u}$ is a unit vector

$\vec{a} \cdot \hat{u}$ is defined as the projection of $\vec{a}$ onto $\hat{u}$
↳ the component of $\vec{a}$ in the direction of $\hat{u}$



The Projection Formula



projection of $\vec{w}$ on $\vec{v}$ (a vector)
$proj_{\vec{v}} \vec{w} = \|\vec{w}\| \cos\theta \, \underbrace{\frac{\vec{v}}{\|\vec{v}\|}}_{\text{unit vector in the direction of } \vec{v}}$

Coursera Linear Algebra

Sunday, December 10, 2023 3:11 PM

Week 1 Quiz

(2) $a - b + c + d = 0$

$\boxed{c - b = -a - d}$

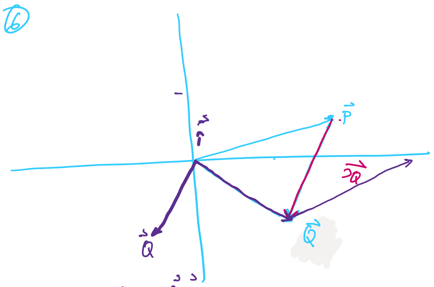
(3) $\vec{c} - 2(\vec{x} - \vec{b}) = 2\vec{b} - 2\vec{a} - \vec{c}$ solve for $\vec{x}$

$\cancel{\vec{c}} - 2\vec{x} + 2\vec{b} = \cancel{2\vec{b}} - 2\vec{a} - \vec{c}$

$-2\vec{x} = -2\vec{a} - 2\vec{c}$

$\vec{x} = \vec{a} + \vec{c}$

(6)



$\vec{PQ} = \vec{Q} - \vec{P}$

(14) $\text{proj}_{\vec{w}} \vec{v}$

projection of $\vec{v}$ onto $\vec{w}$

$(v \cdot \hat{w})(\hat{w})$

$\sqrt{2}$

$$\left(\frac{1}{\sqrt{2}}\hat{i} + \frac{1}{\sqrt{2}}\hat{j}\right) \cdot (2i + 3j)$$

$$= \frac{2}{\sqrt{2}} + \frac{3}{\sqrt{2}} = \frac{5}{\sqrt{2}}$$

$$\left(\frac{5}{\sqrt{2}}\right)\left(\frac{1}{\sqrt{2}}\hat{i} + \frac{1}{\sqrt{2}}\hat{j}\right)$$

$$\frac{5}{2}\hat{i} + \frac{5}{2}\hat{j}$$

$\vec{v} - \text{proj}_{\vec{w}} \vec{v}$

$\vec{v}$ - vector component of $\vec{v}$ in the direction of $\vec{w}$

$$\vec{v} - \frac{\vec{v} \cdot \vec{w}}{\|w\|^2}\vec{w} = \vec{v} - \left(\frac{-1}{5}\right)\vec{w}$$

$$= (i + j) - \left(\frac{-1\,i}{\sqrt{5}} + \frac{-1(-2)}{\sqrt{5}}\right)$$

$$= \frac{2i}{\sqrt{5}} + \frac{-2j}{\sqrt{5}}$$

(16) $(j + k) \times (i + j)$

$$\vec{v} \times \vec{w} = (v_2 w_3 - v_3 w_2)\hat{i} + (v_3 w_1 - v_1 w_3) j + (v_1 w_2 - v_2 w_1) k$$

a d
b e
c f

$$\begin{array}{c} \\ a \\ b \\ c \end{array} \begin{array}{c} d \\ \begin{bmatrix} j_1 + k_1 \\ j_2 + k_2 \\ j_3 + k_3 \end{bmatrix} \end{array} \times \begin{array}{c} d \\ e \\ f \end{array} \begin{bmatrix} i_1 + j_1 \\ i_2 + j_2 \\ i_3 + j_3 \end{bmatrix}$$

i j k i j
v_1 v_2 v_3 v_1 v_2
w_1 w_2 w_3 w_1 w_2
− − − + + +

down product − up product

$v_2 w_3 + v_3 w_1 + v_1 w_2 - w_1 v_2 - w_2 v_3 - w_3 v_1$

i j k j j
a b c a b
d e f d e

$$\begin{bmatrix} bf - ec \\ cd - fa \\ ae - db \end{bmatrix}$$

→ cross product

$$\Big[(j_2 + k_2)(i_3 + j_3) - (i_2 + j_2)(j_3 + k_3) \Big]$$

$$\begin{bmatrix} bf - ec \\ cd - fa \\ ae - db \end{bmatrix}$$

→ cross product

$$\begin{bmatrix} (j_2+k_2)(i_3+j_3) - (i_2+j_2)(j_3+k_3) \\ (j_3+k_3)(i_1+j_1) - (i_3+j_3)(j_1+k_1) \\ (j_1+k_1)(i_2+j_2) - (i_1+j_1)(j_2+k_2) \end{bmatrix}$$

$= j_2 i_3 + j_2 j_3 + k_2 i_3 + k_2 j_3 - i_2 j_3$

$(j+k) \times (i+j)$

i	j	k	i	j
0	1	1	0	1
1	1	0	1	1

$(0-1)i + (1+0)j + (0-1)k$

$-i + j - k$

i	j	k	i	j
0		0		
0	0	1	0	

$(1 \cdot 0) i + (0-0) j + (0-0) k$

i

i	j	k	i	j
1	1	2	1	1
2	-1	1	2	-1

$(1+2)i + (4-1)j + (-1-2)k$

$3i + 3j - 3k$

$9+9+9 = 27$

$\sqrt{27} = 3\sqrt{3}$

Understanding backpropagation

Sunday, December 24, 2023 11:34 PM

Gradient Vector

↳ vector of gradients.

for $f(x,y,z)$, $\nabla f = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \\ \frac{\partial f}{\partial z} \end{bmatrix}$

$\Rightarrow$ Imagine standing at some point $(x_0, y_0, \dots)$

$\nabla f(x_0, y_0, \dots)$ will tell you which way to ascend f the fastest.

- gradient vector are perpendicular to contour lines

Vector space!

↳ any type of mathematical object that can be multiplied by numbers and added together

Field: A field is a set of numbers, $\mathbb{F}$ with the property that if $a, b \in \mathbb{F}$, then $a+b, a-b, ab, \{a/b, b \neq 0\}$ are also in $\mathbb{F}$

$\mathbb{Q}, \mathbb{R}, \mathbb{C}$ are fields of numbers

↳ not $\mathbb{Z}$ i.e. $\frac{3}{5} \notin \mathbb{Z}$ but 3 and 5 are

---

Definition: A **vector space** consists of a set V (elements of V are called **vectors**), a field $\mathbb{F}$ (elements of $\mathbb{F}$ are called **scalars**), and two operations

- An operation called *vector addition* that takes two vectors $v, w \in V$, and produces a third vector, written $v + w \in V$.
- An operation called *scalar multiplication* that takes a scalar $c \in \mathbb{F}$ and a vector $v \in V$, and produces a new vector, written $cv \in V$.

which satisfy the following conditions (called *axioms*).

1. Associativity of vector addition: $(u + v) + w = u + (v + w)$ for all $u, v, w \in V$.
2. Existence of a zero vector: There is a vector in V, written 0 and called the **zero vector**, which has the property that $u + 0 = u$ for all $u \in V$
3. Existence of negatives: For every $u \in V$, there is a vector in V, written $-u$ and called the **negative of** u, which has the property that $u + (-u) = 0$.
4. Associativity of multiplication: $(ab)u = a(bu)$ for any $a, b \in \mathbb{F}$ and $u \in V$.
5. Distributivity: $(a + b)u = au + bu$ and $a(u + v) = au + av$ for all $a, b \in \mathbb{F}$ and $u, v \in V$.
6. Unitarity: $1u = u$ for all $u \in V$.

Backpropagation!

↳ For gradient descent we need to compute ∇C

↳ Backpropagation is a quick way to do this.

→ $\partial C/\partial w$ and $\partial C/\partial b$ tell you how quickly the cost changes as w or b change

The TWO assumptions of the cost function.

Recall $C(w,b) = \frac{1}{2n} \sum_x \| y(x) - a \|^2$

(desired output: $y(x)$; output of neural network: a)

or (with our new notation:

$$C = \frac{1}{2n} \sum_x \| y(x) - a^L(x) \|^2$$

(n: # of training examples)

Assumption #1: Our cost function can be written as an average $C = \frac{1}{n} \sum_x C_x$ over cost functions C_x for individual training examples, x.

↳ important because backpropagation computes the partial derivative for each C_x

Assumption #2: It can be written as a function of outputs from the neural network

↳ quadratic cost function satisfies this constraint

since $C_x = \frac{1}{2} \| y - a^L \|^2$ } for a single training example

$= \frac{1}{2} \sum_j (y_j - a_j^L)^2$

→ Note that the desired output, y, is a fixed parameter

Hadamard Operation $s \odot t$

↳ element wise multiplication of two vectors

i.e. $\begin{bmatrix} 1 \\ 2 \end{bmatrix} \odot \begin{bmatrix} 3 \\ 4 \end{bmatrix} = \begin{bmatrix} 3 \\ 8 \end{bmatrix}$

4 Fundamental Equations Behind Backpropagation

Backpropagation: How changing w and b changes the cost function

- we introduce an intermediate value δ_j^l ← error in the j^{th} neuron of the l^{th} layer
- Backpropagation gives a way to compute the error δ_j^l and relate this error to $\partial C/\partial b_j^l$ and $\partial C/\partial w_{jk}^l$

$$\delta_j^l \equiv \frac{\partial C}{\partial z_j^l}$$

Eq. 1: Error in output layer

$$\delta_j^L = \frac{\partial C}{\partial a_j^L} \sigma'(z_j^L)$$

We can compute $\frac{\partial C}{\partial a_j^L}$ (!)

Recall $C_j = \frac{1}{2}(y-a)^2$ } for a single training data point

$$\frac{\partial C}{\partial a} = \frac{2(y-a)(-1)}{2}$$

$$\frac{\partial C}{\partial a} = a - y$$ } easy enough to compute.

(more correct math notation: $\frac{\partial C}{\partial a_j^L} = (a_j^L - y_j^L)$)

we can also write Eq. 1 in matrix format:

$$\delta^L = \nabla_a C \odot \sigma'(z^L)$$

where $\nabla_a C$ is a vector whose components are $\frac{\partial C}{\partial a_j}$ for each activation in a layer

Eq. 2: Equation for error in terms of error in the next layer

$$\delta^l = ((w^{l+1})^T \delta^{l+1}) \odot \sigma'(z^l)$$

→ it says if you know the error of a layer ahead, you can work backwards to find the error in the current layer

- with Eq. 1 and Eq. 2 you can determine the error of every layer

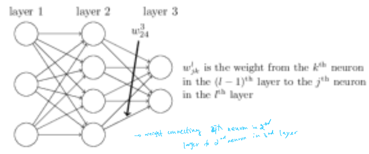
Eq. 3: $\frac{\partial C}{\partial b_j^l} = \delta_j^l$ } used for gradient descent yayyyyyy.....

Nielson's Chapter on Backpropagation: warm-up.

Fast way to compute output of neural network using matrices

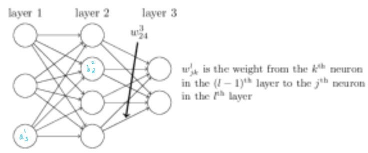
weird notation:

w_{jk}^l is a weight from the connection b/w the k^{th} neuron in the $(l-1)^{th}$ layer w/ the j^{th} neuron in the l^{th} layer



→ weight connecting 4th neuron in 2nd layer to 2nd neuron in 3rd layer

- b_j^l or a_j^l is the activation / bias for the j^{th} neuron in the l^{th} layer



we know $a = \sigma(w \cdot x + b)$, $\sigma(z) = \frac{1}{1+e^{-z}}$

so for $a_j^l = \sigma(w_{jk}^l \underbrace{a_k^{l-1}}_{\text{activation of previous layer}} + b_j^l)$ (1) } for the input layer it would be different? $a_j^2 = \sigma(w_{jk}^2 x_j + b_j^2)$

activation of previous layer ↓ output from the previous layers

"vectorization of a function"

↳ apply the same operation for multiple values

↳ helpful to store those values in a vector.

i.e. for the function $f(x) = x^2$

$$f\left(\begin{bmatrix} 2 \\ 3 \end{bmatrix}\right) = \begin{bmatrix} f(2) \\ f(3) \end{bmatrix} = \begin{bmatrix} 4 \\ 9 \end{bmatrix}$$

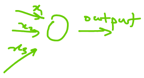
we can vectorize (1) into $a^l = \sigma(\underbrace{w^l a^{l-1} + b^l}_{z^l})$

$z^l \equiv w^l a^{l-1} + b^l$

↳ called the weighted input

Gradient Descent Revisited

Saturday, December 30, 2023 6:44 AM



For the ith neuron

$$\text{output} = \sum_i w_i x_i + b$$

Figuring old notes stuff

$$\Delta\,\text{output} = \sum_i \frac{\partial\,\text{output}}{\partial w}\Delta w_i + \frac{\partial\,\text{output}}{\partial b}\Delta b_i$$

Gradient Descent

↳ we want to minimize a cost function → quadratic cost function

$$C(w,b) = \frac{1}{2n}\sum_x \|\underbrace{y(x)}_{\text{predicted}} - \underbrace{a}_{\text{actual}}\|^2$$

why quadratic cost?
↳ making small changes to w, b won't have much change in the function, making it more difficult to find the optimal set of (w, b)

Suppose we have a ball at some arbitrary position.
If we move it in ~~s~~(w, b) in two directions
↳ move it in 1 direction v_1, and another v_2

$$\Delta C = \frac{\partial C}{\partial v_1}\Delta v_1 + \frac{\partial C}{\partial v_2}\Delta v_2$$

convenient math stuff:

let $\Delta\vec{v} \equiv (\Delta v_1, \Delta v_2)^T$
then
$\nabla C \equiv \left(\frac{\partial C}{\partial v_1}, \frac{\partial C}{v_2}\right)^T$

and
$\Delta C = \nabla C \cdot \Delta v \rightarrow \Delta C \approx \nabla C\,\Delta v$ why approximately symbol?

we can define $\Delta v = -\eta\nabla C$ where η is the learning rate
↳ guarantees ΔC will be negative
↳ so we go to a global minimum

Apply an update rule to v

$v \rightarrow v' = v - \eta\nabla C$
↳ walking to a minimum

$$\Delta C = \nabla C \cdot (-\eta\nabla C)$$
$$= -\eta\nabla C \cdot \nabla C$$
$$\Delta C = -\eta\|\nabla C\|^2$$
↳ what does this mean

FIGURE OUT

In the context of weights and biases, we want to find a combination of (w, b) that minimizes C, by moving in the steepest direction

$$w_k \rightarrow w_k' = w_k - \eta\frac{\partial C}{\partial w_k}$$
and
$$b_l \rightarrow b_l' = b_l - \eta\frac{\partial C}{\partial b_l}$$

update rule to roll down a hill

STOCHASTIC GRADIENT DESCENT

↳ gradient descent has too many things to compute (i.e. cost function for EACH neuron... this is computationally intensive)

↳ use stochastic G.D.
↳ compute ∇C_x for "a small sample of randomly chosen training inputs"
↳ use this to estimate ∇C gradients

⇒ Randomly pick out m training inputs: $\underbrace{\{X_1, X_2, X_3 \dots X_m\}}_{\text{called a mini batch}}$

Assumption: m is sufficiently large so that

$$\frac{\sum_{j=1}^{m}\nabla C_{X_j}}{m} \approx \underbrace{\frac{\sum_{j=1}^{n}\nabla C_{X_j}}{n}}_{\text{for entire training data set.}} = \nabla C$$

...imation of the gradient

$$\frac{\sum_{j=1}^{m} \nabla C_{x_j}}{m} \approx \frac{\sum_{j=1}^{n} \nabla C_{x_j}}{\underbrace{n}_{\text{for entire training data set.}}} = \nabla C$$

Our update rule using stochastic gradient descent:

$$w_k \rightarrow w_k' = w_k - \eta \nabla C \qquad \because \nabla C \approx \frac{\sum_{j=1}^{m} \nabla C_{x_j}}{m}$$

approximation of the gradient

$$\therefore \quad w_k \rightarrow w_k' = w_k - \frac{\eta}{m} \sum_{j=1}^{m} \frac{\partial C_{x_j}}{\partial w_k}$$

Brainstorming

Thursday, December 28, 2023 2:22 PM

RQ: To What Extent can mathematical models simulate intelligence.

Introduction
↳ Talk about Google Reporter calling an AI Sentient

Goal! an analysis of the capabilities of neural networks → Read a few papers on what intelligence is → philosophy

↳ what are aspects of intelligent?
What does it mean to be intelligent

] Society of Mind by Marvin Minsk → tries to explain how the mind works

↳ what mathetical model exists that can ~~be~~ intelligent
in some capability?

intelligent → depends what we call intelligent → define what it means to learn mathematically } important for analysis.

↳ Neural networks
↳ How it works
↳ How it's "intelligent"
↳ What is lacking
↳ The most advanced machine learning algorithms out there
↳ Dalle E
↳ ChatGPT / Bard / Bing
↳ Bring up claim of Google sentient AI

Math
↳ give high level explanation for non technical audience
↳ one chapter / section
can go more in depth.

Components of neural networks
↳ Neural net - perceptrons
↳ Forward propagation } probabilistic
↳ Back propagation
↳ Recurrent neural networks
↳ Biological inspiration → worm model!
↳ All under the principles ...

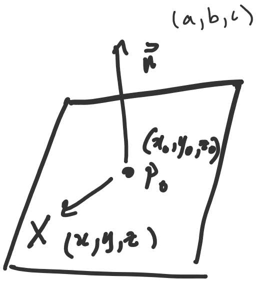
Areas for further research: Quantum Mechanics
↳ apparently there are quantum brain processes... !!
↳ recurrent neural networks
↳ is it necessary to model the human brain to create an intelligent machine?

Planes

Friday, November 10, 2023 6:17 PM

Definition of a plane:

A nonzero vector $\vec{n}$ is called normal for a plane if it is orthagonal to every vector on plane



$$\vec{P_0X} = \begin{bmatrix} x-x_0 \\ y-y_0 \\ z-z_0 \end{bmatrix}$$

Since $\vec{n} \perp \vec{P_0X}$

$\therefore (\vec{n} \cdot \vec{P_0X}) = 0$

$$\therefore \begin{bmatrix} a \\ b \\ c \end{bmatrix} \cdot \begin{bmatrix} x-x_0 \\ y-y_0 \\ z-z_0 \end{bmatrix} = 0$$

Scalar Equation of a plane: $a(x-x_0) + b(y-y_0) + c(z-z_0)$

We need 2 pieces of information to describe a plane.

↳ a point, (on a plane) and its normal vector

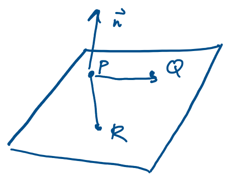
- Two planes are parallel iff their normals are parallel

Example

Find the equation of the plane that passes through the following 3 points

P(1,1,-1), Q(-1,2,-2), R(4,0,2)

ξ form of $a(x-x_0) + b(y-y_0) + c(z-z_0) = 0$



Find $\vec{PQ} \times \vec{PR}$

$$\vec{PQ} = \begin{bmatrix} -2 \\ 1 \\ -1 \end{bmatrix} \quad \vec{PR} = \begin{bmatrix} 3 \\ -1 \\ 3 \end{bmatrix}$$

Find $\vec{n}$

↳ Use cross product or dot product

Cross Product

$\vec{PQ} \times \vec{PR} =$

Matrix algebra

Saturday, November 11, 2023 6:31 AM

We'll think about matrices as algebraic objects

key objective in linear algebra

(new) Matrix Operations

zero matrix : $O_{m,n}$ } matrix w/ all 0s

Transpose : $A = [a_{ij}]$ $A^T = [a_{ji}]$ } swap rows + columns

e.g. $\begin{bmatrix} 5 & 10 \\ 9 & 7 \end{bmatrix}^T = \begin{bmatrix} 5 & 9 \\ 10 & 7 \end{bmatrix}$

If $A^T = A$, A is symmetric!

If $A^T = -A$, A is skew symmetric

Vectors as $n \times 1$ matrices

A vector is just an $n \times 1$ matrix

one column.

Matrix vector multiplication

Saturday, November 11, 2023 6:38 AM

Vector form of a Linear System of Equations

If you have m equations with n unknowns:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \cdots a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \cdots a_{2n}x_n &= b_2 \\ &\vdots \\ a_{i1}x_1 + a_{i2}x_2 + a_{i3}x_3 + \cdots a_{in}x_n &= b_i \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + a_{m3}x_3 + \cdots a_{mn}x_n &= b_m \end{aligned}$$

row by row $\Leftrightarrow$

$$\underbrace{\begin{bmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ \vdots \\ a_{i1} & a_{i2} & a_{i3} & \cdots & a_{in} \\ \vdots \\ a_{m1} & a_{m2} & a_{m3} & \cdots & a_{mn} \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} x_1 \\ \vdots \\ x_i \\ \vdots \\ x_n \end{bmatrix}}_{X} = \underbrace{\begin{bmatrix} b_1 \\ \vdots \\ b_i \\ \vdots \\ b_m \end{bmatrix}}_{\vec{b}}$$

$\Updownarrow$ col by col

$$x_1 \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{i1} \\ \vdots \\ a_{m1} \end{bmatrix} + x_2 \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{i2} \\ \vdots \\ a_{m2} \end{bmatrix} + \cdots + x_n \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{in} \\ \vdots \\ a_{mn} \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ 3 \\ \vdots \\ b_m \end{bmatrix}$$

Matrix-vector multiplication Using vector form (by columns)

$$A\vec{x} = [\vec{a}_1\ \vec{a}_2\ \vec{a}_3 \cdots \vec{a}_n]\ \vec{x} = x_1\vec{a}_1 + x_2\vec{a}_2 + x_3\vec{a}_3 \cdots x_n\vec{a}_n$$

↳ linear combination of $\vec{a}_1, \vec{a}_2, \vec{a}_3 \cdots \vec{a}_n$ (columns of the matrix)

Matrix-vector multiplication using dot product (by rows)

we can think of each row as a transposed vector.

i.e. the first row, $R_1 = [a_{11}\ a_{12}\ a_{13}\ a_{14} \cdots a_{1n}]$

Then R_1^T is a vector $= \begin{bmatrix} a_{11} \\ a_{12} \\ a_{13} \\ \vdots \\ a_{1n} \end{bmatrix}$

$$\therefore A\vec{x} = \begin{bmatrix} R_1^T \cdot \vec{x} \\ R_2^T \cdot \vec{x} \\ \vdots \\ R_m^T \cdot \vec{x} \end{bmatrix}$$

Example

$$\begin{aligned} 2x_1 + x_2 - x_3 &= 1 \\ x_1 + x_2 + x_3 &= 6 \\ -3x_1 + 2x_1 + 4x_3 &= 13 \end{aligned}$$

Matrix form: $A\vec{x}$

$$\begin{bmatrix} 1 \\ 6 \\ 13 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \begin{bmatrix} 2 & 1 & -1 \\ 1 & 1 & 1 \\ -3 & 2 & 4 \end{bmatrix}$$

Vector form

$$\underbrace{x_1 \begin{bmatrix} 2 \\ 1 \\ -3 \end{bmatrix}}_{\text{keep all } x_1 \text{ together}} + \underbrace{x_2 \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}}_{\text{keep all } x_2 \text{ together}} + \underbrace{x_3 \begin{bmatrix} -1 \\ 1 \\ 4 \end{bmatrix}}_{\text{keep all } x_3 \text{ together}}$$

Augmented Form

$$\left[\begin{array}{ccc|c} 2 & 1 & -1 & 1 \\ 1 & 1 & 1 & 6 \\ -3 & 2 & 4 & 13 \end{array}\right]$$

Inconsistent solution

→ no solution or infinite # of solutions

Thrm: A system $A\vec{x} = \vec{b}$ is consistent iff $\vec{b}$ is a linear combination of the col.A, i.e. there exists $x_1, x_2, x_3 \ldots x_n$ s.t.

$$x_1\vec{a}_1 + x_2\vec{a}_2 \cdots x_n\vec{a}_n = \vec{b}$$

Ex. Is $\left[\begin{array}{cc|c} 1 & 3 & 4 \\ 1 & -1 & 1 \\ 1 & 5 & 3 \end{array}\right]$ consistent?

$$x_1 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + x_2 \begin{bmatrix} 3 \\ -1 \\ 5 \end{bmatrix} = \begin{bmatrix} 4 \\ 1 \\ 3 \end{bmatrix}$$

OR $\vec{x} \cdot \begin{bmatrix} 1 \\ 3 \end{bmatrix} = 4$, $\vec{x} \cdot \begin{bmatrix} 1 \\ -1 \end{bmatrix} = 1$, $\vec{x} \cdot \begin{bmatrix} 1 \\ 5 \end{bmatrix} = 3$

$$\begin{bmatrix} x_1 \\ x_1 \\ x_1 \end{bmatrix} + \begin{bmatrix} 3x_2 \\ -x_2 \\ 5x_2 \end{bmatrix} = \begin{bmatrix} 4 \\ 1 \\ 3 \end{bmatrix}$$

$$\begin{aligned} x_1 + 3x_2 &= 4 \\ x_1 - x_2 &= 1 \\ x_1 + 5x_2 &= 3 \end{aligned}$$

same thing!

$$\begin{aligned} x_1 + 3x_2 &= 4 \\ x_1 - x_2 &= 1 \\ x_1 + 5x_2 &= 3 \end{aligned}$$

Ex. of determining an inconsistent solution → from Gilbert Textbook

Does $\begin{bmatrix} 1 & 3 & 5 \\ 1 & 2 & -3 \\ 2 & 5 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 4 \\ 5 \\ 8 \end{bmatrix} = b$ have any solutions

$x \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} + y \begin{bmatrix} 3 \\ 2 \\ 5 \end{bmatrix} + z \begin{bmatrix} 5 \\ -3 \\ 2 \end{bmatrix}$ represents the linear system of:

$$\begin{aligned} x + 3y + 5z &= 4 \quad (1) \\ x + 2y - 3z &= 5 \quad (2) \\ 2x + 5y + 2z &= 8 \quad (3) \end{aligned}$$

Compute (1) + (2) − (3)

$0 = 1$

↳ False, so the system does not

$x_1 + 3x_2 = 4$
$x_1 - x_2 = 1$
$x_1 + 5x_2 = 3$

same

$2x + 5y$ ---

Compute ① + ② − ⑤

$0 = 1$

↳ False, so the system does not have a solution

↳ what about in the case where x, y, z do not cancel out?

Properties of Matrix-vector multiplication

1. $A(\vec{x} + \vec{y}) = A\vec{x} + A\vec{y}$

   ↳ matrix!

2. $A(a\vec{x}) = a(A\vec{x})$
3. $(A+B)\vec{x} = A\vec{x} + B\vec{x}$

* use vector form to prove these * → 2023 EE me does not have time to prove these...

The Identity Matrix

A matrix I, s.t. $I\vec{x} = \vec{x}$

(same idea as multiplying by 1)

$$\begin{bmatrix} R_1^T \cdot x \\ R_2^T \cdot x \\ R_3^T \cdot x \\ \vdots \\ R_n^T \cdot x \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

i.e. for a 3x3 matrix

$$I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

↳ it's symmetrical!

$$\left.\begin{matrix} 1 & \underbrace{0 \quad 0 \quad 0 \quad \cdots}_{n-1 \text{ zeros}} \\ 0 \quad 1 & \underbrace{0 \quad 0 \quad 0 \quad \cdots}_{n-2 \text{ zeros}} \\ 0 \quad 0 & 1 \quad 0 \quad 0 \quad \cdots \\ \vdots \quad \vdots & \cdots \quad 0 \quad 0 \quad 0 \\ \vdots \quad \vdots & \cdots \quad 0 \quad 1 \quad 0 \\ \vdots \quad \vdots & \cdots \quad 0 \quad 0 \quad 1 \end{matrix}\right\} \text{The identity matrix}$$