

1 The Arithmetic Derivative

"Mamma told me "you can't take derivatives on integers!!!" and I was like "hold my Leibniz rule""

Definition 1 (The Arithmetic Derivative for natural numbers).

For any $n \in \mathbb{N}_0$ the **arithmetic derivative**, denoted $(n)'$ is defined as follows:

- $(p)' = 1$, for any prime p
- $(pq)' = (p)'q + p(q)'$, for any $p, q \in \mathbb{N}$ (**Leibniz Rule**)

Corollary 1 (Elementary derivatives).

$(0)' = 0$ and $(1)' = 0$

Proof. Follows immediately from the Leibniz rule with $p = q = 1$ □

Corollary 2 (Power Rule). For any integers p and $n \geq 0$:

$$(p^n)' = np^{n-1}(p')$$

Proof. Trivial □

Corollary 3 (Arithmetic derivatives of Integers and rational numbers).

One can extend the arithmetic derivative to the integers by showing

$$(-x)' = -(x)'$$

Further, the **quotient rule** is also well defined on \mathbb{Q} :

$$\left(\frac{p}{q}\right)' = \left(\frac{(p)'q - p(q)'}{q^2}\right)$$

Proof. Ain't nobody got da time fo' dis. □

Corollary 4 (Prime factorization derivative formula). Let $\omega(x)$ be the prime omega function, indicating the number of distinct prime factors in x , and $\nu_p(x)$ be the p -adic valuation of x . Then,

$$(x)' = \sum_{\substack{p|x \\ p \text{ prime}}} \frac{v_p(x)}{p} x$$

Proof. The prime factorization of an integer $x \in \mathbb{Z}$ is given by

$$x = \prod_{i=1}^{\omega(x)} p_i^{v_{p_i}(x)}$$

it follows that

$$D(x) = \sum_{i=1}^{\omega(x)} \left[v_{p_i}(x) \left(\prod_{j=1}^{i-1} p_j^{v_{p_j}(x)} \right) p_i^{v_{p_i}(x)-1} \left(\prod_{j=i+1}^{\omega(x)} p_j^{v_{p_j}(x)} \right) \right]$$

$$=\sum_{i=1}^{\omega(x)}\frac{v_{p_i}(x)}{p_i}x=\sum_{\substack{p|x\\p\text{ prime}}}\frac{v_p(x)}{p}x$$

□

Example 1.

$$(60)'=(2^2\cdot 3\cdot 5)'=\left(\frac{2}{2}+\frac{1}{3}+\frac{1}{5}\right)\cdot 60=92,$$

$$(81)'=D(3^4)=4\cdot 3^3\cdot D(3)=4\cdot 27\cdot 1=108.$$

Hello!

$$x+3=67$$

This is really fun!!

$$x+35^2=???\\ 2x^{(35x+x^x)}$$

Greek letters :

$$\pi$$

$$\beta$$

$$\alpha$$

$$A=\pi r^2$$

$$\log_5(35)$$

fractions:

$$\frac{x}{y}$$

$$\frac{x}{y}$$

$$\frac{x}{y}$$

$$\frac{x}{1+x+x^2}$$

$$\frac{\sqrt{x^2+1}}{\sqrt{x^2-1}}$$

$$\frac{\sqrt{x^2+1}}{\sqrt{x^2-1}}$$

brackets:

$$\{a,b,c\}$$

$$\left(\frac{m_3+m_4}{x^2}\right)$$

$$\left| \frac{dx}{dy} \right|_{x=1}$$

$$\left| \frac{\delta x}{\delta y} \right|_{x=1}$$

$$\forall n \in \mathbb{Z} \quad a-1|a^n-1$$

$$\int x^{dx}-1$$

This is a table

x	1	2	3	4	5
$f(x)$	10	11	12	13	14

Exercise 1. Prove that if p : *prime* , then \sqrt{p} is irrational. More generally, prove that if $n \in \mathbb{Z}$ and $n \neq m^2$, then \sqrt{n} is irrational.

$$\begin{array}{rcl} x & \approx & 1.567 \\ 4x+3^x & = & 12 \end{array}$$

Solve the following systems of congruences:

$$\begin{cases} 12x+31y\equiv 2\pmod{127} \\ 2x+89y\equiv 23\pmod{127} \end{cases} \tag{1}$$

$$\begin{cases} x\equiv 1\pmod{3} \\ x\equiv 1\pmod{4} \\ x\equiv 1\pmod{5} \\ x\equiv 0\pmod{7} \end{cases} \tag{2}$$

Solve the following congruence polynomials:

$$x^2\equiv 29\pmod{35} \tag{3}$$

$$3x^2+6x+5\equiv 0\pmod{7} \tag{4}$$

Find y such that the following holds:

$$19^{y^{1000}}\equiv 1\pmod{20} \tag{5}$$

Assume p is prime, and $\exists a$ $1 \leq a \leq (p-1)$ then

$$\text{Show } [a], [2a], \dots, [(p-1)a] \text{ are all residue classes} \tag{6}$$

Prove that if $n \in \mathbb{Z}$ and $n > 0$, then the prime factorization of the binomial coefficient $C(n,k)$ is given by

$$\binom{n}{k} = \prod_{i=1}^n p_k^{\sum_{i=1}^n \left\lfloor \frac{n}{p_k^i} \right\rfloor - \left\lfloor \frac{n-k}{p_k^i} \right\rfloor - \left\lfloor \frac{k}{p_k^i} \right\rfloor} \quad (7)$$

Proof. This is written in L^AT_EX, so it must be true. □

Theorem 1. $P = NP$

Proof. Let $S = \{X | X \notin X\}$, and let $NP \in S$. Since $P \subseteq NP$ then $P \in P \iff P \notin P$, but clearly, we also have that $NP \in NP \iff NP \notin NP$. It follows trivially that $P \in NP \iff P \notin NP$ and $NP \in P \iff NP \notin P$. Thus we conclude that $P = NP$ \square

Corollary 5. Where is my million dollars???

$$I_n = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix}$$

$$I = (\delta_{ij})$$

$$A = U\Sigma V^*$$

Prove Mantel's Theorem, that is, prove that

$$|E| \leq \frac{|V^2|}{4} \quad (8)$$

Theorem 1 (LaPlace Expansion). *The determinant of an $n \times n$ matrix can be calculated as follows:*

$$\det(A) = \sum_{i=1}^n a_{i,j} \text{Cof}(A_{i,j})$$

where

$$\text{Cof}(A_{i,j}) = (-1)^{i+j} \det \left(\begin{matrix} & \underline{A_{i,j}} & \\ & (n-1) \times (n-1) & \end{matrix} \right)$$

and the expansion occurs along the i -th row. (Note the expansion can also be taken along the j -th column).

If A is an $n \times n$ matrix, and E_1, E_2, \dots, E_k are elementary matrices resulting from taking $[A \xrightarrow{RREF} I_n]$, then $A = E_1^{-1} E_2^{-1} \dots E_k^{-1} I_n$, and so

$$\det(A) = \det(E_1^{-1}) \dots \det(E_k^{-1})$$

Exercise 1. Let

$$A = \begin{bmatrix} a & d & 2 \\ b & e & 1 \\ c & f & 1 \end{bmatrix}, B = \begin{bmatrix} d & e & f \\ a & b & c \\ 1 & 1 & 0 \end{bmatrix}, C = \begin{bmatrix} -a & 2c + 3a & 1 \\ -b & 2e + 3b & 0 \\ -c & 2f + 3c & 1 \end{bmatrix}$$

If $\det(A) = 2$ and $\det(B) = -3$, find $\det[2A^3 B^{-1} A^T \text{adj}(3C^2)]$

Theorem 2 (Inclusion-Exclusion for Probability). *The general inclusion-exclusion formula for the union of sets A_1, A_2, \dots, A_n is completely determined by the simple formula*

$$\mathbb{P}\left(\bigcup_{i=1}^n A_i\right) = \mathbb{P}(\emptyset) - \sum_n \left(\left[\sin\left(k\pi + \frac{\pi}{2}\right) \right] \sum_{\substack{I \subset \{1, \dots, n\} \\ |I|=k}} \left(1 - \mathbb{P}\left(\overline{\bigcap_{i \in I} A_i}\right) \right) \right)$$

Proof. The proof is trivial and left as an exercise for the reader. \square

$$(\mathbf{w}^*, \xi^*, \alpha^*, \beta^*) = \underset{\mathbf{w}, \xi}{\operatorname{argmin}} \underset{\alpha, \beta}{\operatorname{argmax}} \mathcal{L}(\mathbf{w}, \xi, \alpha, \beta)$$

$$= \underset{\mathbf{w}, \xi}{\operatorname{argmin}} \underset{\alpha, \beta}{\operatorname{argmax}} \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i=1}^n \xi_i - \sum_{i=1}^n \alpha_i (y_i \mathbf{w}^T x_i - 1 + \xi_i) - \sum_{i=1}^n \beta_i \xi_i$$

Proposition 3. Let $\delta^{\text{TM}} = (\Sigma, q, \Phi, \gamma, Q_O, q_{rEjEcT}, q_{ACCePT})$ be a **Turing Machine**. you know the rest.

$$\forall u \in U, \mathbf{w} \in u\mathbb{W}, U_{\mathbf{w}} \bigcup_{u \in \mathcal{U}} uWu \bigcup_{w \in \mathcal{W}} w_u := \mathcal{O}w\mathcal{O}$$

$$A = \{u1v : u, v \in \Sigma^* \text{ and } |u|, |v| \geq 1\} \, .$$

2 Gradient Descent for Linear Regression

Suppose we have a hypothesis $h : \mathbb{R}^n \rightarrow \mathbb{R}$, $h_\theta(\mathbf{x}) = \hat{y}$ with parameters $\theta \in \mathbb{R}^n$. Recall the **Mean-Squared Loss** (MSE) metric, applied to linear regression:

$$MSE(y, h_\theta(\mathbf{x})) = \frac{1}{n} \|\mathbf{y} - h_\theta(\mathbf{x})\|_2^2 = \frac{1}{2n} \sum_{i=1}^n (y^{(i)} - \hat{y}^{(i)})^2 = \frac{1}{2n} \sum_{i=1}^n (y^{(i)} - \mathbf{w}^T \mathbf{x}^{(i)})^2$$

Then we have the general update:

$$\mathbf{w}^{k+1} = \mathbf{w}^k + \alpha \frac{\partial}{\partial \mathbf{w}} MSE(y, \hat{y})$$

Batch Gradient Descent: For $k = 0, 1, \dots$

1. For $k = 0, 1, \dots$

$$\mathbf{w}^{k+1} = \mathbf{w}^k + \alpha \frac{1}{n} \sum_{i=1}^n (y^{(i)} - \mathbf{w}^T \mathbf{x}^{(i)}) \mathbf{x}^{(i)}$$

Mini-batch Gradient Descent:

1. For $k = 0, 1, \dots$
 - (a) Split data D into T subsets D_t of sizes n_0, \dots, n_{T-1} , s.t. $\sum_t n_t = 1$.
 - (b) For each subset D_t :

$$\mathbf{w} := \mathbf{w} + \alpha \frac{1}{n_t} \sum_{i: \mathbf{x}^{(i)} \in D_t}^{n_t} (y^{(i)} - \mathbf{w}^T \mathbf{x}^{(i)}) \mathbf{x}^{(i)}$$

Stochastic Gradient Descent:

1. For $k = 0, 1, \dots$
 - (a) For $i = 1, \dots, n$:

$$\mathbf{w} := \mathbf{w} + \alpha (y^{(i)} - \mathbf{w}^T \mathbf{x}^{(i)}) \mathbf{x}^{(i)}$$

L1-norm

$$\|\mathbf{w}\|_1 = \sum_i |w_i|$$

$$\frac{\partial}{\partial \mathbf{w}} \|\mathbf{w}\|_1 = \text{sign}(\mathbf{w}) = [\text{sign}(w_1), \dots, \text{sign}(w_m)]^T$$

$$C = \sum_{w \in LDA(D)} E[idx(w)]$$

Theorem 4 (Inclusion-Exclusion in Measure Theory). *Let (X, μ) be a finite measure space. For any finite measurable sets $A_1, \dots, A_n \subseteq X$*

$$\begin{aligned}
& \mu \left(\bigcup_{i=1}^n A_i \right) \\
&= \sum_{i=1}^n \mu(A_i) - \sum_{1 \leq i < j \leq n} \mu(A_i \cap A_j) + \sum_{1 \leq i < j < k \leq n} \mu(A_i \cap A_j \cap A_k) - \dots + (-1)^{n-1} \mu(A_1 \cap \dots \cap A_n) \\
&= \mu \left(\bigcup_{i=1}^n A_i \right) = \sum_{\emptyset \neq J \subseteq \{1, \dots, n\}} (-1)^{|J|+1} \mu \left(\bigcap_{j \in J} A_j \right). \\
&= \mu \left(\bigcup_{i=1}^n A_i \right) = \sum_{k=1}^n \left((-1)^{k-1} \sum_{\substack{I \subseteq \{1, \dots, n\} \\ |I|=k}} \mu(A_I) \right), \\
&= \mu(\emptyset) - \sum_n^{k=1} \left(\left[\sin \left(k\pi + \frac{\pi}{2} \right) \right] \sum_{\substack{I \subseteq \{n, \dots, 1\} \\ |I|=k}} \left(1 - \mu \left(\overline{\bigcap_{i \in I} A_i} \right) \right) \right)
\end{aligned}$$

Proof. Yikes □

Theorem 5 (Inclusion-Exclusion). *Given n sets A_1, \dots, A_n in an universal space S , the cardinality of the union of n such sets is given by*

$$\begin{aligned}
& \left| \bigcup_{i=1}^n A_i \right| \\
&= \sum_{i=1}^n |A_i| - \sum_{1 \leq i < j \leq n} |A_i \cap A_j| + \sum_{1 \leq i < j < k \leq n} |A_i \cap A_j \cap A_k| - \dots + (-1)^{n-1} |A_1 \cap \dots \cap A_n| \\
&= \left| \bigcup_{i=1}^n A_i \right| = \sum_{\emptyset \neq J \subseteq \{1, \dots, n\}} (-1)^{|J|+1} \left| \bigcap_{j \in J} A_j \right|. \\
&= \left| \bigcup_{i=1}^n A_i \right| = \sum_{k=1}^n \left((-1)^{k-1} \sum_{\substack{I \subseteq \{1, \dots, n\} \\ |I|=k}} |A_I| \right), \\
&= |\emptyset| - \sum_n^{k=1} \left(\left[\sin \left(k\pi + \frac{\pi}{2} \right) \right] \sum_{\substack{I \subseteq \{n, \dots, 1\} \\ |I|=k}} \left(1 - \left| \overline{\bigcap_{i \in I} A_i} \right| \right) \right)
\end{aligned}$$

Proof. Yikes □

Claim 6. The Post-Correspondence Problem PCP is decidable relative to the acceptance problem A_{TM} .

Proof. We essentially want to show PCP is reducible to ATM , i.e., if ATM were decidable, then the PCP problem would also be decidable. For this purpose, we suppose A_{TM} were decidable, and use the fact that mapping reducibility is transitive. We first reduce $HALT_{TM}$ to A_{TM} . If A_{TM} were decidable, $HALT_{TM}$ would also be decidable. Let S be Turing Machine deciding A_{TM} . We construct the T.M. H for $HALT_{TM}$ as follows:

$H =$ "On input $\langle M, x \rangle$, where M is a Turing Machine and w its input:

1. Construct machine M' from M by marking all rejecting states of M as accepting.
2. Run S on $\langle M', x \rangle$, if it accepts, **accept**, if it rejects **reject**. "

Clearly, $\langle M, x \rangle \in HALT_{TM}$ if $\langle M', w \rangle \in A_{TM}$. Let this machine H also be an oracle T.M. Now, by using the transitivity of mapping reduction, if A_{TM} were decidable, then $HALT_{TM}$ would also be decidable, so we can show that if $HALT_{TM}$ were decidable, PCP would also be. Now let R be a Turing Machine that recognizes PCP , (which checks in linear time whether a proposed match is an actual solution to PCP by simply comparing the top and bottom symbols and accepting if all of them are equal and rejecting otherwise). Assume that s_1, s_2, \dots is a list of all possible strings in PCP . Construct the following enumerator E .

$E =$ "Ignore the input.

1. Repeat the following for $i = 1, 2, 3, \dots$
2. Run R for i steps on each input s_1, s_2, \dots, s_i .
3. If any computations accept, print out the corresponding s_j ."

If R accepts a particular string s , eventually it will appear on the list generated by E . Finally, we construct the decider P for PCP :

$P =$ "On input $\langle R, x \rangle$, where R is the machine described above:

1. Query the oracle for H on input $\langle E, x \rangle$.
2. If $E \in HALT_{TM}$, a solution to PCP exists, so run E until the solution is out and **accept**.
3. If $E \notin HALT_{TM}$, no solution to PCP exists, so **reject**. "

This shows that PCP is decidable relative to A_{TM}

□

Claim 7. Use the languages $A = a^m b^n c^n | m, n \geq 0$ together with example 2.36 to show that the class of context-free languages is not closed under intersection.

Proof. Note that

$$L(A \cap B) = \{w \mid w = a^m b^n c^n \wedge$$

□

$$C = \sum_{w \in D, t \text{ fidf}(w) > t} E[id_x(w)]$$

Proposed change
 \implies

$$C = \sum_{w \in LDA(D)} E[id_x(w)]$$

$$\int x^{dx} - 1$$

$$\int \frac{x^{dx} - 1}{dx} dx = \int \lim_{\Delta x \rightarrow 0} \frac{x^{\Delta x} - 1}{\Delta x} dx$$

Theorem 8 (Freshman's Dream).

Roses are red

Violets are blue

$(x+y)^n = x^n + y^n$ is true

in $\mathbb{Z}/2\mathbb{Z}$ (9)

Theorem 9 (Fubini's Theorem). Let X and Y be σ -finite measure spaces, and suppose $X \times Y$ is the given product measure. Then, if f is a $X \times Y$ (measurable) function and

$$\int_{X \times Y} |f(x, y)| d(x, y) < \infty$$

then

$$\int_X \left(\int_Y f(x, y) dy \right) dx = \int_Y \left(\int_X f(x, y) dx \right) dy = \int_{X \times Y} f(x, y) d(x, y)$$

Theorem 10 (Freshman's dream). *Let n be prime, then*

$$(x + y)^n = x^n + y^n$$

holds in $\mathbb{Z}/n\mathbb{Z}$

Proof. Let p be a prime, and note that

$$(x + y)^p = x^p + \binom{p}{1}x^{p-1}y + \binom{p}{2}x^{p-2}y^2 + \cdots + \binom{p}{p-1}x^{p-1}y + y^p$$

Let $1 \leq i \leq p-1$, and consider

$$\binom{p}{i} = \frac{p!}{i!(p-i)!} = p \frac{(p-1)!}{i!(p-i)!} \in \mathbb{N}$$

Any factor from $i!(p-i)!$ is not going to divide p .

Why is this true? We know that

$$i! = 1 * 2 * \cdots * i < p$$

$$(p-1)! = 1 * 2 * \cdots * (p-1) < p$$

which implies that none of the factors above can divide $p \implies$ any of them must divide $(p-i)!$.

$$\implies \binom{p}{i} = p * k_i \in \mathbb{N} \implies p \mid \binom{p}{i}, i = 1, \dots, (p-1)$$

$$\therefore (x + y)^p \equiv x^p + y^p \pmod{p}$$

□

$$J(\Theta) = -\frac{1}{n} \left[\sum_{i=1}^n \sum_{k=1}^K y_k^{(i)} \log h_{\theta}(x^{(i)})_k + (1 - y_k^{(i)}) \log(1 - h_{\theta}(x^{(i)})_k) \right] + \frac{\lambda}{2n} \sum_{l=1}^{L-1} \sum_{i=1}^{s_l} \sum_{j=1}^{s_l+1} (\Theta_j^l)^2$$

$$\frac{d^2y}{dx^2} = \left(\frac{dy}{dx}\right)^2$$

$$\lim_{\Delta x \rightarrow 0} \int_{-1}^0 \left(\frac{1}{x^2} + \mathbb{B}ruh \right) \Delta x$$

$$\int_a^b f(x)dx = - \int_b^a f(x)dx$$

$$\begin{aligned} \Rightarrow \lim_{N \rightarrow \infty} \sum_{i=1}^N f(x_i^*) \Delta x &= \lim_{N \rightarrow \infty} \sum_{i=1}^N f(x_i^*) \left(\frac{b-a}{N} \right) = - \lim_{N \rightarrow \infty} \sum_{i=1}^N f(x_i^*) \left(\frac{a-b}{N} \right) \\ \Rightarrow \sum_{i=1}^N f(x_i^*) \left(\frac{b-a}{N} \right) &= - \sum_{i=1}^N f(x_i^*) \left(\frac{a-b}{N} \right) := - \sum_{i=1}^N f(x_i^*) \left(\frac{b-a}{N} \right) \end{aligned}$$

Define

$$\sum_N^{n=1} a_n := -a_N - a_{N-1} - \cdots - a_1$$

Then

$$\sum_{n=1}^N a_n = - \sum_N^{n=1} a_n$$

This is well-defined as it is no more than a special case of the definite integral :

$$\sum_{k=a}^b f(k) = \int_{[a,b]} f d\mu$$

where μ is the **counting measure**.

$$\int_0^{\phi=y} \exp \left\{ \int_{-\infty}^{x=t} e^x dx \right\} d\phi = ye^t$$

Exercise 2. Let $x \in \mathbb{R}$, $N \in \mathbb{Z}$, $m \in \mathbb{N}$. Find the exact value of

$$\underbrace{\sqrt[N]{x \sqrt[N]{x \sqrt[N]{x \dots \sqrt[N]{x}}}}}_{m\text{-times}}$$

1. When $m < \infty$
2. Show that the expression above converges when $m \rightarrow \infty$ and find describe the conditions for convergence on the value of x and N .

Proof. Note that

$$\begin{aligned} \sqrt[N]{x} &= x^{\frac{1}{N}} \quad (m = 1) \\ \sqrt[N]{x \sqrt[N]{x}} &= \sqrt[N]{x * x^{\frac{1}{N}}} = x^{\frac{1}{N}} * x^{\frac{1}{N^2}} \quad (m = 2) \\ \sqrt[N]{x \sqrt[N]{x \sqrt[N]{x}}} &= \sqrt[N]{x * x^{\frac{1}{N}} * x^{\frac{1}{N^2}}} = x^{\frac{1}{N}} * x^{\frac{1}{N^2}} * x^{\frac{1}{N^3}} \quad (m = 3) \\ &\vdots \\ \sqrt[N]{x \sqrt[N]{x \sqrt[N]{x \dots \sqrt[N]{x}}}} &= \prod_{i=1}^m x^{1/N^i} \quad (m = m) \end{aligned}$$

But this is simply

$$\prod_{i=1}^m x^{1/N^i} = x^{\sum_{i=1}^m \frac{1}{N^i}}$$

When $m < \infty$, when $N \neq 1 \implies \frac{1}{N} \neq 1$

$$\sum_{i=1}^m \frac{1}{N^i} \stackrel{i=j+1}{=} \sum_{j=0}^{m-1} \frac{1}{N^{j+1}} = \frac{1}{N} \sum_{j=0}^{m-1} \frac{1}{N^j} = \frac{1}{N} \left(\frac{1 - \left(\frac{1}{N}\right)^m}{1 - \frac{1}{N}} \right)$$

Further, when $m \rightarrow \infty$, as for $|N| > 1$, $\left| \frac{1}{N} \right| < 1$

$$\lim_{m \rightarrow \infty} \sum_{i=1}^m \frac{1}{N^i} = \frac{1}{N} \sum_{i=1}^{\infty} \frac{1}{N^{i-1}} = \frac{1}{N} \sum_{i=1}^{\infty} \left(\frac{1}{N} \right)^{i-1} = \frac{1}{N} \left(\frac{1}{1 - \frac{1}{N}} \right)$$

So that

$$\underbrace{\sqrt[N]{x \sqrt[N]{x \sqrt[N]{x \dots \sqrt[N]{x}}}}_{m\text{-times}} = \prod_{i=1}^m x^{1/N^i} = x^{\sum_{i=1}^m \frac{1}{N^i}} = x^{\left\{ \frac{1}{N} \left(\frac{1 - \left(\frac{1}{N}\right)^m}{1 - \frac{1}{N}} \right) \right\}} \quad (10)$$

and

$$\underbrace{\sqrt[N]{x \sqrt[N]{x \sqrt[N]{x \dots \sqrt[N]{x}}}}_{m \rightarrow \infty} = \prod_{i=1}^{\infty} x^{1/N^i} = x^{\sum_{i=1}^{\infty} \frac{1}{N^i}} = x^{\left\{ \frac{1}{N} \left(\frac{1}{1 - \frac{1}{N}} \right) \right\}} \quad (11)$$

□

Corollary 6 (Infinite nested square root of 2).

$$\underbrace{\sqrt{2 \sqrt{2 \sqrt{2 \dots \sqrt{2}}}}}_{m \rightarrow \infty} = 2$$

Proof. Trivial y the formula above with $x = 2$, $N = 2$ □

Excuse me sir, do you have a moment to talk about our God and saviour

L^AT_EX?

$$y_{+++} = \sum_{I=1}^I \sum_{j=1}^J \sum_{k=1}^K y_{ijk}$$

Theorem 11. Let $D = \text{diag}(d_1, \dots, d_m) \in \mathbb{F}^{m \times m}$ de a real or complex-valued matrix. Then

$$e^D = \begin{pmatrix} e^{d_1} & & & & \\ & \ddots & & & \\ & & e^{d_j} & & \\ & & & \ddots & \\ & & & & e^m \end{pmatrix}$$

Proof. We have that

$$e^D = \mathbf{I}_m e^D \quad (12)$$

$$= \mathbf{I}_m \sum_{n=0}^{\infty} \frac{1}{n!} D^n \quad (13)$$

$$= \sum_{n=0}^{\infty} \frac{1}{n!} \left(\sum_j [\mathbf{I}_m]_{ij} D_{jk} \right)^n \quad (14)$$

Since

$$[\mathbf{I}_m]_{ij} D_{jk} = \begin{cases} D_{jk} & \forall \ i = j, j = k \\ 0 & \text{else} \end{cases}$$

$$\Rightarrow e^D = \sum_{n=0}^{\infty} \frac{1}{n!} \left(\sum_j D_{jj} \right)^n \quad (15)$$

$$= \sum_{n=0}^{\infty} \frac{1}{n!} \left(\sum_j d_j \right)^n \quad (16)$$

$$= e^{\sum_j d_j} \quad (17)$$

$$\therefore e^D = \mathbf{I}_m e^D \quad (18)$$

$$= \mathbf{I}_m e^{\sum_j d_j} \quad (19)$$

$$= \begin{pmatrix} e^{d_1} & & & & \\ & \ddots & & & \\ & & e^{d_j} & & \\ & & & \ddots & \\ & & & & e^m \end{pmatrix} \quad (20)$$

□