Lecture 5: Backpropagation Vector, Matrix and Tensor Derivatives

Subhransu Maji, Chuang Gan and TAs Some slides kindly provided by Fei-Fei Li, Jiajun Wu, Erik Learned-Miller Lecture 5 - 1 Sept. 19, 2023

Where we are ...

$$egin{aligned} s &= f(x;W) = Wx & ext{scores function} \ L_i &= \sum_{j
eq y_i} \max(0, s_j - s_{y_i} + 1) & ext{SVM loss} \ L &= rac{1}{N} \sum_{i=1}^N L_i + \sum_k W_k^2 & ext{data loss + regularization} \end{aligned}$$

want $\nabla_W L$

Subhransu Maji, Chuang Gan and TAs Some slides kindly provided by Fei-Fei Li, Jiajun Wu, Erik Learned-Miller Lecture 5 - 2 Sept. 19, 2023

Optimization



Subhransu Maji, Chuang Gan and TAs Some slides kindly provided by Fei-Fei Li, Jiajun Wu, Erik Learned-Miller

Lecture 5 - 3 Sept. 19, 2023

Gradient Descent

$$rac{df(x)}{dx} = \lim_{h o 0} rac{f(x+h) - f(x)}{h}$$

Numerical gradient: slow :(, approximate :(, easy to write :) **Analytic gradient**: fast :), exact :), error-prone :(

In practice: Derive analytic gradient, check your implementation with numerical gradient

Subhransu Maji, Chuang Gan and TAs Some slides kindly provided by Fei-Fei Li, Jiajun Wu, Erik Learned-Miller Lecture 5 - 4 Sept. 19, 2023

Overview of where we're going

- We want to evaluate the gradient of a Loss function L(x,W,...), with respect to the parameters (weights) of a neural network, at the "point" represented by the arguments to the function (x,W,...).
 - We are not interested in an algebraic expression for the gradient, but rather only in the evaluation of that gradient at the current value of the function arguments.

Lecture 5 - 5

Sept. 19, 2023

Consider the function

$$z(x,y) = x^2 + y^2,$$

and suppose we are interested in evaluating the gradient of this function at the point

$$(x, y) = (5, 3).$$

Evaluate the gradient:

$$\frac{\partial z}{\partial x} = 2x.$$

$$\frac{\partial z}{\partial y} = 2y.$$

The algebraic expression of the gradient is just the collection of these partials into a "vector":

The evaluation of this gradient at the point (x, y) = (5, 3) is simply

$$abla z(5,3) = \begin{bmatrix} 2 \times 5 \\ 2 \times 3 \end{bmatrix} = \begin{bmatrix} 10 \\ 6 \end{bmatrix}.$$

Do care about this

2023





Lecture 5 - 7 Sept. 19, 2023



Lecture 5 - 8 Sept. 19, 2023

Computational Graph



Subhransu Maji, Chuang Gan and TAs Some slides kindly provided by Fei-Fei Li, Jiajun Wu, Erik Learned-Miller

Lecture 5 - 9 Sept. 19, 2023

$$f(x, y, z) = (x + y)z$$

e.g. x = -2, y = 5, z = -4

Forward pass: evaluating each expression in the computational graph from the inputs to the final output (or outputs). The results of each forward step are shown in green.



Subhransu Maji, Chuang Gan and TAs Some slides kindly provided by Fei-Fei Li, Jiajun Wu, Erik Learned-Miller

Lecture 5 - 10 Sept. 19, 2023

set some inputs
x = -2; y = 5; z = -4
perform the forward pass
q = x + y # q becomes 3
f = q * z # f becomes -12

perform the backward pass (backpropagation) in reverse order: # first backprop through f = q * z dfdz = q # df/dz = q, so gradient on z becomes 3 dfdq = z # df/dq = z, so gradient on q becomes -4 # now backprop through q = x + y dfdx = 1.0 * dfdq # dq/dx = 1. And the multiplication here is the chain rule! dfdy = 1.0 * dfdq # dq/dy = 1

Subhransu Maji, Chuang Gan and TAs Some slides kindly provided by Fei-Fei Li, Jiajun Wu, Erik Learned-Miller

Lecture 5 - 11 Sept. 19, 2023

$$f(x, y, z) = (x + y)z$$

e.g. x = -2, y = 5, z = -4

Backward pass: evaluating the partial derivative of each **parameter** or **intermediate result** in the computational graph from the outputs back to the inputs.The results of each backward step are shown in red.



Goal is to calculate

$$rac{\partial f}{\partial x}, rac{\partial f}{\partial y}, rac{\partial f}{\partial z}$$

evaluated at the point

$$[x = -2, y = 5, z = -4].$$

Subhransu Maji, Chuang Gan and TAs Some slides kindly provided by Fei-Fei Li, Jiajun Wu, Erik Learned-Miller Lecture 5 - 12 Sept. 19, 2023

$$f(x, y, z) = (x + y)z$$

e.g. x = -2, y = 5, z = -4

$$q = x + y \quad \frac{\partial q}{\partial x} = 1, \frac{\partial q}{\partial y} = 1$$

$$f = qz \qquad \frac{\partial f}{\partial q} = z, \frac{\partial f}{\partial z} = q$$

$$\frac{\partial f}{\partial q} = df \quad \partial f$$

Want: $\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}$

Subhransu Maji, Chuang Gan and TAs Some slides kindly provided by Fei-Fei Li, Jiajun Wu, Erik Learned-Miller Lecture 5 - 13 Sept. 19, 2023



Lecture 5 - 14 Sept. 19, 2023

$$f(x, y, z) = (x + y)z$$

e.g. x = -2, y = 5, z = -4

$$q = x + y \quad \frac{\partial q}{\partial x} = 1, \frac{\partial q}{\partial y} = 1$$

$$f = qz \qquad \frac{\partial f}{\partial q} = z, \frac{\partial f}{\partial z} = q$$

Want: $\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}$

Lecture 5 - 15 Sept. 19, 2023

$$f(x, y, z) = (x + y)z$$

e.g. $x = -2, y = 5, z = -4$

$$q = x + y \quad \frac{\partial q}{\partial x} = 1, \frac{\partial q}{\partial y} = 1$$

$$f = qz \qquad \frac{\partial f}{\partial q} = z, \frac{\partial f}{\partial z} = q$$

Want: $\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}$

Lecture 5 - 16 Sept. 19, 2023

$$f(x, y, z) = (x + y)z$$

e.g. x = -2, y = 5, z = -4

$$q = x + y \quad \frac{\partial q}{\partial x} = 1, \frac{\partial q}{\partial y} = 1$$

$$f = qz \qquad \frac{\partial f}{\partial q} = z, \frac{\partial f}{\partial z} = q$$

$$\frac{\partial f}{\partial z}$$
Want: $\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}$

Lecture 5 - 17 Sept. 19, 2023

$$f(x, y, z) = (x + y)z$$

e.g. x = -2, y = 5, z = -4

$$q = x + y \quad \frac{\partial q}{\partial x} = 1, \frac{\partial q}{\partial y} = 1$$

$$f = qz \qquad \frac{\partial f}{\partial q} = z, \frac{\partial f}{\partial z} = q$$

$$\boxed{\begin{array}{c} x & -2 \\ y & 5 \\ z & 4 \\ \hline 3 & \end{array}}$$

$$f = \frac{1}{3}$$

$$\boxed{\begin{array}{c} \frac{\partial f}{\partial z} \\ \frac$$

Lecture 5 - 18 Sept. 19, 2023

$$f(x, y, z) = (x + y)z$$

e.g. $x = -2, y = 5, z = -4$

$$q = x + y \quad \frac{\partial q}{\partial x} = 1, \frac{\partial q}{\partial y} = 1$$

$$f = qz \qquad \frac{\partial f}{\partial q} = z, \frac{\partial f}{\partial z} = q$$

Want: $\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}$

Lecture 5 - 19 Sept. 19, 2023

$$f(x, y, z) = (x + y)z$$

e.g. $x = -2, y = 5, z = -4$

$$q = x + y \quad \frac{\partial q}{\partial x} = 1, \frac{\partial q}{\partial y} = 1$$

$$f = qz \qquad \frac{\partial f}{\partial q} = z, \frac{\partial f}{\partial z} = q$$

Want: $\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}$

Lecture 5 - 20 Sept. 19, 2023

$$f(x, y, z) = (x + y)z$$

e.g. $x = -2, y = 5, z = -4$

$$q = x + y \quad \frac{\partial q}{\partial x} = 1, \frac{\partial q}{\partial y} = 1$$

$$f = qz \qquad \frac{\partial f}{\partial q} = z, \frac{\partial f}{\partial z} = q$$

Want: $\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}$

Lecture 5 - 21 Sept. 19, 2023

$$f(x, y, z) = (x + y)z$$

e.g. $x = -2, y = 5, z = -4$

$$q = x + y \quad \frac{\partial q}{\partial x} = 1, \frac{\partial q}{\partial y} = 1$$

$$f = qz \qquad \frac{\partial f}{\partial q} = z, \frac{\partial f}{\partial z} = q$$

Want: $\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}$

$$\overset{x -2}{y - \frac{1}{4}}$$

$$\overset{y -5}{-4} \qquad (x + y + y) = \frac{1}{4}$$

$$y - \frac{1}{4} \qquad (x + y) = \frac{1}{4}$$

$$\frac{y - \frac{1}{4}}{3}$$

Chain rule: $\frac{\partial f}{\partial y}$

Lecture 5 - 22 Sept. 19, 2023

$$f(x, y, z) = (x + y)z$$

e.g. $x = -2, y = 5, z = -4$

$$q = x + y \quad \frac{\partial q}{\partial x} = 1, \frac{\partial q}{\partial y} = 1$$

$$f = qz \qquad \frac{\partial f}{\partial q} = z, \frac{\partial f}{\partial z} = q$$

Want: $\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}$

Lecture 5 - 23 Sept. 19, 2023

$$f(x, y, z) = (x + y)z$$

e.g. x = -2, y = 5, z = -4

$$q = x + y \quad \frac{\partial q}{\partial x} = 1, \frac{\partial q}{\partial y} = 1$$

$$f = qz \quad \frac{\partial f}{\partial q} = z, \frac{\partial f}{\partial z} = q$$
Want: $\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}$
Chain rule:

$$\frac{\partial f}{\partial x} = \frac{\partial f}{\partial q} \frac{\partial q}{\partial x}$$

Lecture 5 - 24 Sept. 19, 2023

```
# set some inputs
x = -2; y = 5; z = -4
# perform the forward pass
q = x + y # q becomes 3
f = q * z # f becomes -12
```

perform the backward pass (backpropagation) in reverse order: # first backprop through f = q * z dfdz = q # df/dz = q, so gradient on z becomes 3 dfdq = z # df/dq = z, so gradient on q becomes -4 # now backprop through q = x + y dfdx = 1.0 * dfdq # dq/dx = 1. And the multiplication here is the chain rule! dfdy = 1.0 * dfdq # dq/dy = 1

Subhransu Maji, Chuang Gan and TAs Some slides kindly provided by Fei-Fei Li, Jiajun Wu, Erik Learned-Miller Lecture 5 - 25 Sept. 19, 2023



Lecture 5 - 26 Sept. 19, 2023



Lecture 5 - 27 Sept. 19, 2023



Lecture 5 - 28 Sept. 19, 2023



Lecture 5 - 29 Sept. 19, 2023



Lecture 5 - 30 Sept. 19, 2023



slides kindly provided by Fei-Fei Li, Jiajun Wu, Erik Learned-Miller

$$f(w,x)=rac{1}{1+e^{-(w_0x_0+w_1x_1+w_2)}}$$

-

"sigmoid function"



Subhransu Maji, Chuang Gan and TAs Some slides kindly provided by Fei-Fei Li, Jiajun Wu, Erik Learned-Miller

Lecture 5 - 32 Sept. 19, 2023

$$f(w,x)=rac{1}{1+e^{-(w_0x_0+w_1x_1+w_2)}}$$



Subhransu Maji, Chuang Gan and TAs Some slides kindly provided by Fei-Fei Li, Jiajun Wu, Erik Learned-Miller

Lecture 5 - 33 Sept. 19, 2023

$$f(w,x)=rac{1}{1+e^{-(w_0x_0+w_1x_1+w_2)}}$$



Subhransu Maji, Chuang Gan and TAs Some slides kindly provided by Fei-Fei Li, Jiajun Wu, Erik Learned-Miller

Lecture 5 - 34 Sept. 19, 2023

$$f(w,x)=rac{1}{1+e^{-(w_0x_0+w_1x_1+w_2)}}$$



Subhransu Maji, Chuang Gan and TAs Some slides kindly provided by Fei-Fei Li, Jiajun Wu, Erik Learned-Miller

Lecture 5 - 35 Sept. 19, 2023

$$f(w,x)=rac{1}{1+e^{-(w_0x_0+w_1x_1+w_2)}}$$



Subhransu Maji, Chuang Gan and TAs Some slides kindly provided by Fei-Fei Li, Jiajun Wu, Erik Learned-Miller Lecture 5 - 36 Sept. 19, 2023

$$f(w,x)=rac{1}{1+e^{-(w_0x_0+w_1x_1+w_2)}}$$



Subhransu Maji, Chuang Gan and TAs Some slides kindly provided by Fei-Fei Li, Jiajun Wu, Erik Learned-Miller

Lecture 5 - 37 Sept. 19, 2023

$$f(w,x)=rac{1}{1+e^{-(w_0x_0+w_1x_1+w_2)}}$$



Subhransu Maji, Chuang Gan and TAs Some slides kindly provided by Fei-Fei Li, Jiajun Wu, Erik Learned-Miller

Lecture 5 - 38 Sept. 19, 2023

$$f(w,x)=rac{1}{1+e^{-(w_0x_0+w_1x_1+w_2)}}$$



Subhransu Maji, Chuang Gan and TAs Some slides kindly provided by Fei-Fei Li, Jiajun Wu, Erik Learned-Miller

Lecture 5 - 39 Sept. 19, 2023

$$f(w,x)=rac{1}{1+e^{-(w_0x_0+w_1x_1+w_2)}}$$



Subhransu Maji, Chuang Gan and TAs Some slides kindly provided by Fei-Fei Li, Jiajun Wu, Erik Learned-Miller Lecture 5 - 40 Sept. 19, 2023

$$f(w,x)=rac{1}{1+e^{-(w_0x_0+w_1x_1+w_2)}}$$



Subhransu Maji, Chuang Gan and TAs Some slides kindly provided by Fei-Fei Li, Jiajun Wu, Erik Learned-Miller

Lecture 5 - 41 Sept. 19, 2023

$$f(w,x)=rac{1}{1+e^{-(w_0x_0+w_1x_1+w_2)}}$$



Subhransu Maji, Chuang Gan and TAs Some slides kindly provided by Fei-Fei Li, Jiajun Wu, Erik Learned-Miller

Lecture 5 - 42 Sept. 19, 2023

$$f(w,x)=rac{1}{1+e^{-(w_0x_0+w_1x_1+w_2)}}$$



Subhransu Maji, Chuang Gan and TAs Some slides kindly provided by Fei-Fei Li, Jiajun Wu, Erik Learned-Miller

Lecture 5 - 43 Sept. 19, 2023



Lecture 5 - 44 Sept. 19, 2023



Lecture 5 - 45 Sept. 19, 2023

$$f(w,x) = \frac{1}{1 + e^{-(w_0 x_0 + w_1 x_1 + w_2)}} \qquad \qquad \sigma(x) = \frac{1}{1 + e^{-x}} \quad \text{sigmoid function}$$
$$\frac{d\sigma(x)}{dx} = \frac{e^{-x}}{(1 + e^{-x})^2} = \left(\frac{1 + e^{-x} - 1}{1 + e^{-x}}\right) \left(\frac{1}{1 + e^{-x}}\right) = (1 - \sigma(x))\sigma(x)$$



Lecture 5 - 46 Sept. 19, 2023

$$f(w,x) = \frac{1}{1 + e^{-(w_0 x_0 + w_1 x_1 + w_2)}} \qquad \qquad \sigma(x) = \frac{1}{1 + e^{-x}} \quad \text{sigmoid function}$$
$$\frac{d\sigma(x)}{dx} = \frac{e^{-x}}{(1 + e^{-x})^2} = \left(\frac{1 + e^{-x} - 1}{1 + e^{-x}}\right) \left(\frac{1}{1 + e^{-x}}\right) = (1 - \sigma(x))\sigma(x)$$



Lecture 5 - 47 Sept. 19, 2023

```
w = [2,-3,-3] # assume some random weights and data
x = [-1, -2]
# forward pass
dot = w[0]*x[0] + w[1]*x[1] + w[2]
f = 1.0 / (1 + math.exp(-dot)) # sigmoid function
# backward pass through the neuron (backpropagation)
ddot = (1 - f) * f # gradient on dot variable, using the sigmoid gradient derivation
dx = [w[0] * ddot, w[1] * ddot] # backprop into x
dw = [x[0] * ddot, x[1] * ddot, 1.0 * ddot] # backprop into w
# we're done! we have the gradients on the inputs to the circuit
```

Lecture 5 - 48 Sept. 19, 2023

Patterns in backward flow

add gate: gradient distributor
max gate: gradient router
mul gate: gradient... "switcher"?



Subhransu Maji, Chuang Gan and TAs Some slides kindly provided by Fei-Fei Li, Jiajun Wu, Erik Learned-Miller Lecture 5 - 49 Sept. 19, 2023

Gradients add at branches



Subhransu Maji, Chuang Gan and TAs Some slides kindly provided by Fei-Fei Li, Jiajun Wu, Erik Learned-Miller Lecture 5 - 50 Sept. 19, 2023

Implementation: forward/backward API



Graph (or Net) object. (Rough pseudo code)

| class Com | <pre>iputationalGraph(object):</pre> |
|-----------|----------------------------------------------------------------------------|
| # | |
| def f | Forward(inputs): |
| # | <pre># 1. [pass inputs to input gates]</pre> |
| # | <pre># 2. forward the computational graph:</pre> |
| f | <pre>or gate in self.graph.nodes_topologically_sorted():</pre> |
| | gate.forward() |
| r | return loss # the final gate in the graph outputs the loss |
| def b | backward(): |
| f | <pre>or gate in reversed(self.graph.nodes_topologically_sorted()):</pre> |
| | <pre>gate.backward() # little piece of backprop (chain rule applied)</pre> |
| r | <pre>return inputs_gradients</pre> |
| | |

Subhransu Maji, Chuang Gan and TAs Some slides kindly provided by Fei-Fei Li, Jiajun Wu, Erik Learned-Miller

Lecture 5 - 51 Sept. 19, 2023

Implementation: forward/backward API



(x,y,z are scalars)



Subhransu Maji, Chuang Gan and TAs Some slides kindly provided by Fei-Fei Li, Jiajun Wu, Erik Learned-Miller

Lecture 5 - 52 Sept. 19, 2023

Implementation: forward/backward API



| class Multipl | yGate(object): |
|------------------------|--------------------------------------------|
| def forwa | ard(x,y): |
| z = x | ·*γ |
| self. | <pre>x = x # must keep these around!</pre> |
| self. | y = y |
| retur | n z |
| <mark>def</mark> backw | ard(dz): |
| dx = | <pre>self.y * dz # [dz/dx * dL/dz]</pre> |
| dy = | <pre>self.x * dz # [dz/dy * dL/dz]</pre> |
| retur | n [dx, dy] |
| | |

(x,y,z are scalars)



Subhransu Maji, Chuang Gan and TAs Some slides kindly provided by Fei-Fei Li, Jiajun Wu, Erik Learned-Miller

Lecture 5 - 53 Sept. 19, 2023

Example: Torch Layers

| torch / mn | Q-Math- st ±list ra | YFek ser |
|-----------------------------|-------------------------------------------------------------------------|------------------|
| O Code C Innen 22 | The Parlingands and III Wild + Palan & Craphs | |
| description or website pro | uded. | |
| C 1 898 commits | References Definitions Definitions | |
| 0.000.0000 | present Creation Carton | |
| mits master - New pull | New Tile Find Tile HTTPS - https://github.com/torch 🚳 | Download ZIP |
| soundb Marga pull request i | 983 from Sanchimount 162 marcher :=== Laboral science 2 Male | 174 15 hours ago |
| dos | Fix tustst mutit in Margit RankingOritarion | 4 days ago |
| peneric | Improve error nessage in SpatialConvolutionMM | a day ago |
| | THNK attrining OperAP include | 2 days ago |
| rocks | Act half dependency | 14-days ago |
| glignore | tel pi to ignore build output | 4 months ago |
| Juscheckrs | [Tarth] Move test.lus to the top level | a year ago |
| travis.ymi | anal faes for lest path | 2 months ago |
| Abs.lue | Add THMN conversion of (ELU, LeekyReLU, LogSigmoid, LogSidMax, Looku | 7 days ago |
| AbsOrtherion.3us | Add THMN conversion of (ELU, LeekyReLU, LogSigmoid, LogSidMax, Looku | 7 days ago |
| Additue | for Add with multi-dim bias | 10 months ago |
| AddConstant.lue | Adding in-place AddConstant and WulConstant | 9 months ago |
| BCECriterion lue | Remove unrecessary malloca from BCEC/belon | 3 months ago |
| BatchNormalization Jus | factat/more-wast | 3 months ago |
| CAddTable lue | fixing table modules to return correct number of gradinguits | 4 months app |
| CDivTable lue | thing table modules to return correct number of gradinguts | 4 months ago |
| CMIsheLists Int | Add C implementation of SpatialBalc/Hormalization | 7 days ago |
| CMullus | on Module preserve type sharing semantics (#187); add ns Module apply | 4 months ago |
| CMullable lue | thing table mobules to return correct number of gradinguts | d months ago |
| CONTRIBUTING.md | added developing fps | 3 months ago |
| COPYRIGHT M | add copyright fla | 2 years ago |
| CSubTable Iue | fixing table modules to return correct number of gradinputs | 6 months ago |
| Clamp Iua | Use custom range in HardTanh and mask it as Clamp | 3 months ago |
| Class/NLL Oritorion lus | Add functional conversion of ClassNLLCriterion | 13 days ago |
| Concat kae | fis a hug in conditional expression | a month ago |
| Concel/Table Iue | fixing bug in Concat/Table variable length | 4 months ago |
| Container Iua | Adding applyToModulies() to m. Container, which is like :apply() but | 3 months ago |
| Copy lue | nn.Module preserve type sharing semantics (#187), add nn.Module apply | 4 months ago |
| Cosine lue | Fatge() in Cosine | a month ago |
| Cosine/Distance Iue | Donot change state variables in CosineDistance/CosineEmbeddingOnterion | 2 months ago |
| CosineEmbeddingCriterion I | Dursti change state variables in CosineDistance/CosineEmbeddingOnterion | 2 months ago |
| Criterion Jua | co.Module preserve type attarting semantics (#107), add ro.Module apply | 4 months ago |
| Criterion Table Iue | Rename argonic to lable argonic for Lue 5.2 | E-months ago |
| CrossEntropyCriterion Jus | Check for 'multipulai and 'muCriterion' in recursiveType. | E months ago |
| DepthConcet.lue | adding dexit backward to Concel, DepthConcel, Bequential | 1 months ago |
| Die/KLDivCriterion.lue | Use tensor for THMN functions even for single element outputs | 10 days ago |
| DoiProduct.lue | Add batch mode in DolProduct + unit text | 2 months ago |
| Dropout.iue | h-place draptul | 4 months ago |
| ELUlus | Add ThMW conversion of (ELU, Leeky/NaLU, LogSigmoid, LogSofMax, Looku | 7 days ago |
| ErrorMessages.lue | Give better error messages when trying to use the wrong kind of Tensor. | a year ago |
| Euclidean.lue | nn.Module preserve type sharing semantics (#167); add nn.Module apply | 4 months ago |
| Explus | Eq made lus only | 9 months ago |
| FlattenTable lue | nn:Module preserve type sharing semantics (#187), add nn:Module apply | 4 months ago |
| Orader/Reversal.lus | Act Orader/Provesal layer | 4 months ago |
| HardShrink.lus | Act functional conversion of HardShrink | 10 days ago |
| HardTarih.lua | Add functional conversion of HardTanh | 10 days ago |
| HingsEnbeddingOffanion lua | revents HingeEnteddingCiterion to support batch mode | 4 months app |
| dentry lue | Reset to previous Identity (us implementation | 2 months app |
| Index. Iue | Brightying and more efficient recircles | 2 months app |
| Jacobian lua | Add unit tests for hession (us, fix bugs detected by the tests. | 6 months app |
| Join Table Isa | on Module preserve type sharing semantics (#187), add ns Module apply | 4 months app |
| L1Centilue | Use tensor for THNN functions even for single element outputs | 10 days ago |
| 1HingeEnbeddingCriteria | Make (gpst) indy recursion. | 1 months app |
| (Penally lue | fixed LSPenalty constructor arguments | a year ago |
| LeakyReLU ius | Att THMI conversion of ELU, LaskoRaLU, LogSigmoid, LogSofMax, Looku | 7 days app |

| ogfigroit he | Add THNN conversion of (ELU, LeekyReLU, LogSigmoid, LogSidMax, Looku | 7 days |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------|-----------|
| ogSofMax.lue | Add THNN conversion of (ELU, LeakyReLU, LogSigmoid, LogSidMax, Looku | 7 days |
| cokup?able.lue | Namonize Lookupitable signature with ourn impl | 5 days |
| Bd has | Rename unpack to table unpack for Lus 5.2 | E months |
| ISE Oriterion Isa | Add SizeAverage to otheritors in the constructor | 2 months |
| Rengin Criterion Jue | modernized MarginCriterion | a year |
| ReginRankingCriterion lua | Fix batch mode in MarginRankingCriterion | 4 days |
| fer. Ive | Merge pull request #404/from vginalmaster | 2 months |
| Reamlas | Add support for negative dimension and both balkh and non-balkh input | 2 months |
| Rev. Auto | Merge pull request #400 from vgits/masker | 2 months |
| Reform Tables has | cancel unused variable and useless expression | 29 days |
| Rodule lue | Favert 'Don't re-fatten parameters if they are already fattered' | 15 hours |
| Ruf Jun | removing the requirement for providing size in nr.3/s/ | a year |
| MConstant ha | Igrow updateOradioput if self gradioput is nil | 3 months |
| Listonarios Ius | asserts in MultiOriterian and ParallelOriterion add | 2 months |
| USLabelliargeCriterion lua | Indial reverse of tords? the | 4 jears |
| Children Children Las | solometer address his | 11 months |
| | Append in Nampe no done in pace. | 10000 |
| terrorite inter | Former bern und beitfenn han Rinnelen bern en fins ellevels senner. | 11 days |
| Well Line | Euflers for PRoLU and a malementation | Logothe |
| atter be | food bridge on Palifier and was submad in backness | 1 months |
| Vinise Determine | Manage and resourced MTCC from material indianality | 20 days |
| water law | Is a hug in conditional expression | 1.000 |
| walk/Cristin he | assarts is MultiCriterian and ParabatCriterion and | 2 months |
| and of the last | Parallel entirelation. Parallel [®] able inharite Container, and tests | 2.000 |
| Traver Iva | Use UNIX line androps | 7 months |
| EADME md | do: malledos | 5 months |
| PetUlus | Add randomized lastic rectified linear unit (PReLU) | 3 months |
| MCU Iun | adds in-place RaUU and fixes a potential divide by zero in rectlight | 1 months |
| heplicate has | Replicate batchthole | 1 months |
| leshape lus | Added more informative pretty-printing. | a year |
| Infect has | Initial revenue of tandh? tree | 4 pears |
| when the first state in the state of the sta | nn.Module preserve type altering semantics (#187); add on Module apply | 4 months |
| legarita La | fixing Sequential remove corner case | 4 months |
| ignoid lue | initial reverge of tooth? two | 4 years |
| incolt4,10/lierion.lue | Add SizeAverage to otherions in the constructor | 2 months |
| iofMax.lun | Fix various unused variables in m | a year |
| in Ministration | Fix various unused variables in m | a year |
| iotPlus lue | fixed a numerical issue in the SofPlus module (It breaks for input g | 2 pears |
| URSIVITE La | initial reventp of turch? tree | 4 pears |
| iottlign ka | initial reverse of tanth? tree | 4 pears |
| pare-Jacobian lua | Fa various unused variables in m | a year |
| partel.hear lua | Using sparse implementation of zeroGradParameters for SpanseLinear | a month- |
| patal/dapti-starPooling | Added SpatialAdaptiveBlasPooling | a year |
| patal/wenger/soling ha | Spatial/versgePosting supports padding, cell mode and exclude grad dix | 29 days |
| parameter mean advantion last | Add C mpemeration of spatialbalt/Normalization | 7 days |
| and the second s | And a spectra way meaned. | 1 months |
| and an and a second | Ex total in Statisficant data | last |
| and all second size of the local | Danna in and and analysis initialization look from an Statiations | A second |
| instal Countries (Ph.) - | tade considered | 10,000 |
| and the second second | Soutial Construction District Subtraction Normalization work with bot | A months |
| and a Drame Lat | amal fa on error message | 6 months |
| and rational last sole. | Adding Fractional Naw Paoling | 3 months |
| and relationships has | Add adjustment term to SpatialFullConvolution to control the size of | 5 days |
| and all of the state of the sta | New NN classes | 2 years |
| ipatial, PPusing lua | Spatial-keragePacing divides by kIIPAH | 10 months |
| and price Provide page | SpatiaMaxPooling suggests padding and cell mode | 6 months |
| pataMarCrossing La | Add SpatiaMacDressling | 28 days |
| patial SoftMax. Los | Update SofMax to work in spatial mode | 4 months |
| patiel to the principal set | Marga branch Inn_Inst_most | 3 years |
| palathatechernals. | Spatial/Constructive, Subtractive/Hormalization work with bat | I months |
| patial.)plangingNearest I | Use UNIX line and rgs | 7 months |
| ipatia@antPatting.ka | Added more informative pretty-printing. | a year |
| pitTatrie lue | Add support for negative indices in nn.5plitTable | 7 months |
| | | |

Subhransu Maji, Chuang Gan and TAs Some slides kindly provided by Fei-Fei Li, Jiajun Wu, Erik Learned-Miller

Lecture 5 - 54 Sept. 19, 2023

Example: Torch Layers

| torch / mn | | | Q Math- 1 | aller m | YFek ar |
|-----------------------------|-----------------------------------------------------|---------------------------|-----------------|--------------------|-----------------------|
| O Code Channe 22 | Different as 10 Mil + Puls | in Graphs | | | |
| description or website pr | ovided. | | | | |
| ③ 1,839 commits | (P#branches | O Bindeau | 105 | 0.00 | All the second second |
| insche master + New yell | legant New Yor | Find file HTTPS - | https://gathati | an tarat B | Download ZP |
| sounds Merge put request | Mill from Sandhillevent Mill master 🖂 | | | alest commit 23445 | nd 15 hours ago |
| a doc | Fis batch mode in MarginRankingOrlarion | | | | 4 days ago |
| B paneric | Improve error message in SpatialConvolutionM | M | | | a day ago |
| | THNN: attl missing OpenMP include | | | | 2 days ago |
| rocks | Att half dependency | | | | 14 days ago |
| gilignore | tel pl to ignore build output | | | | 4 months ago |
| Juscheckrs | [Tanth] Move test has to the top level | | | | a year ago |
| invision (| anal fees for test path | | | | 2 months app |
| Absilue | Att THIN convention of ELU, LeakyReLU, La | offigmoid LogfietMax L | ooku | | 7 days ago |
| AbsOrtexton.lus | Att THNN convention of (ELU, Look-PaLU, Lo | glignoid LogfotMax L | ooku | | 7 days acc |
| Add Aue | fa Add with multi-dim bias | | | | 10 months app |
| AddConstant.lus | Adding in-place AddConstant and MulConstant | | | | 9 months and |
| BCECiterion has | Remove americana malloca from BCECyberl | | | | 1 months ago |
| BatchNormalization los | Schutchnorm weat | | | | 3 months and |
| CANTAIN IN | faing lattic modules to return correct number of | analfrauta | | | d martin and |
| CDivTable lue | faing table modules to return correct number of | pradhopula | | | d months and |
| Contraction of | http://www.com/com/com/com/com/com/com/com/com/com/ | - | | | T days and |
| Charles | a link in succession for derive second of the | 17. add on Martin and | | | A months and |
| Concernant | fairs table produces by starting second of p | and and the second appro | | | A marche and |
| CONTRACTOR | | franking . | | | 1 months app |
| CONTRECTINGING | attact terrenting the | | | | 1 months ago |
| COPTROPTION | and oppopt to | | | | 2 years ago |
| CoubTable Iua | flang table mobules to return correct number of | gradinputa | | | 6 months ago |
| Clampika | Use custom range in HardTanh and mask it as | Diamp | | | 3 months ago |
| Class/ALLCriterion.lus | Add functional conversion of ClassNLL Orberton | | | | 13 days ago |
| E Concatilua | fa a bug in conditional expression | | | | a month ago |
| E ConcatTable lue | faing bug in Concat/Table variable length | | | | 4 months ago |
| Container lua | Adding applyToModules() to nn Container, whi | th is like upply() but | | | 3 months ago |
| Copy lue | m.Mobile preserve type sharing semantics (P | 87), add nn.Module.apply | | | 4 months ago |
| Cosine lue | Fartger) in Cosine | | | | a month ago |
| CosineDistance lue | Donot change state variables in CosineDistant | #CosineEnbeddingCrite | rion. | | 2 months ago |
| CosineEmbeddingCriterion 1. | Do not change state variables in CosineDistant | a CosineEmbeddingCrite | rion | | 2 months ago |
| Criterion.lus | m.Mobile preserve type sharing semantics (P | 87), add rec.Module.apply | | | 4 months ago |
| CriterionTable lue | Reserve urgack to table urgack for Lue 5.2 | | | | E months ago |
| CrossEntropyOnterion Jus | Check for 'm Module' and 'm Orlierion' in rec | ansiveType. | | | I months ago |
| DepthConcat.lus | adding direct thackward to Concat, DepthConc | A. Sequential | | | 9 months ago |
| DistRLDivOnterion.lue | Use tensor for THNN functions even for single- | element outputs | | | 10 days ago |
| DoiProduct.lue | Attributch mode in DolProduct + unit text | | | | 2 months ago |
| Dropout.lus | in-place-dropool | | | | 4 months ago |
| ELUiter | Add THNN conversion of (ELU, Leeky/NeLU, Le | gSignoid LogSofMax, L | osku | | 7 days ago |
| ErofNessages.lus | One better error messages when trying to use | he wrong kind of Tensor. | | | a year ago |
| Euclidean lue | milliplais preserve type sharing semantics (# | 87), add nn Module apply | | | 4 months ago |
| Explus | Exp made lue only | | | | 9 months ago |
| FatterTable.lue | milliplais preserve type sharing semantics (# | 87), add nn Module apply | | | 4 months ago |
| OradentRevenative | Act Oracler/Reversal layer | | | | 4 months ago |
| HardShrink.lus | Add functional conversion of HardShrink | | | | 10 days ago |
| HardTanh.lua | Add functional conversion of HardTanh | | | | 10-days app |
| HepeEnbeddingCitation.lus | ments HingsEnteddingCriterion to support by | alch mode | | | d months ago |
| Martily has | Recent to previous literative lass implementation | | | | 2 months ago |
| brefers has | Simplifying and more efficient on Index | | | | 2 months and |
| Jacobian lua | Ant unit tests for hermion lus, for huge detected | by the lesits. | | | d months ago |
| D Join Table has | or Module preserve how sharing perpendice of | 1071 add on Module and | | | d months and |
| B L1Couline | Use termine for THMN functions even for visual | dement outputs | | | 10 days was |
| 1 Marsh ale Marsh | Main local link mountain | | | | 1 months and |
| 1 threads be | fact thank contains | | | | |
| a constrainty road | and a sum common whereas | | | | a year ago |

| officeroit he | Add THRW convenion of (ELU, LeakyReLU, LogRigmoid, LogRotMax, Looku | 7 days |
|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------|
| ogSofMax kun | Add THNN convenion of (ELU, LeekyPaLU, LogRigmoid, LogRidMan, Looku | 7 days |
| ookupTable lue | Harmonize Lookup/Table signature with oursn impl | 5 days |
| Bd has | Fierame unpack to table unpack for Los 5.2 | 4 months |
| 650 Ortherion Isa | Add SizeAverage to offeriors in the constructor | 2 months |
| ReginCriterion Jus | modernized MarginDifferior | a yea |
| ReginRankingCriterion Iua | Fis batch mode in MarginRankingOritarion | 4 days |
| Ren. Iva | Marga pull request #404 from vgire/master | 2 months |
| Ream Aux | Add support for regative dimension and both batch and non-batch input | 2 months |
| fin Jus | Marga pull request #404 from ogirs/master | 2 months |
| RefureTable Ise | cancel unused variable and useless expression | 29 days |
| Rodule Iue | Revet 'Don't re-fatten parameters if they are already fattered' | 15 hours |
| lui lue | removing the requirement for providing size in nn.Mul | a yea |
| MConstant lua | Ignore updateGraditput if self-graditput is nil | 3 months |
| UBOrlanter Tua | asserts in MultiCriterion and ParalleiCriterion add | 2 months |
| ML abelliarge Criterion Iua | initial reveng of torth? the | 4 years |
| Margi-Crister Iua | multimargin apports pr2 | 11 months |
| lance has | typeAa in Nampe not done in place. | 4 months |
| terrow"able has | NarrowTable | 4 months |
| converge lue | Parnove trum and tastitionin from Normalize, because they allocate memory, | 25 days |
| THE CAR | puters for Photo sate englimentation. | 1 months |
| widing has | fixed broken nr.Padiling input was returned in backprop | 5 months |
| sirviseOnlance lue | Marga pull request #252 from suggrig/hastler | 29 days |
| water isa | ta a tug in conditional expression | a month |
| waterContents to a | assets in MultOthmun and ParabelOthmon add | 2 months |
| anatient activities | Paratel systemation, Paraternable interna Container, unit testa | 1 100 |
| teen tue | Use UNIX line androps | 7 months |
| EADNE HE | doc mathemas | |
| PRECONSE | Add randomized leavy received linear unit (PTHLU) | |
| Net of the | ands in pace has a second a powerse dwale by cert in mission | 1000 |
| la de la composición de | Adda and addression | 1 1000 |
| and the second | Added more enterance programmy | 4 1999 |
| and the local | and a second of the second sec | |
| and and and and | for a first sector spectrum and a sector of the providence of the | 1 |
| imposit her | hold an ann of hold Tan | 4 1000 |
| incented Withdraw Inc. | Add Standownee to other in the constructor | 2 months |
| in White her | Fix various unused variables in no | 2.100 |
| and the last | Fix various unused variables in ne | 3 100 |
| in the last | fixed a numerical issue in the SofPlus module (I breaks for issue a | 2 1007 |
| and desired has | initial secures of teech? two | 4 1000 |
| uttigs he | initial severa of tech? test | d year |
| aut related area | Fix various unused variables in no | 3 199 |
| partel, new las | Using sparse implementation of zeroGradParameters for SpanseLinear | a month |
| patial Adaptive Man Pooling | Added SpatialAduptivelitavPooling | a yea |
| patial-wagePooling has | Spatial/wengePusing supports padding, cell mode and exclude, pad dir | 29 days |
| patallat/Momalization lua | Add C implementation of SpatiaBlatt/Normalization | 7 days |
| pataConstructionals . | Make type) tuty recursion. | 1 months |
| patie/Constitution has | Fix type() in SpatialConvolution | 3 months |
| patialCon-clutorAM lua | Fix type() in SpatiatConvolution | 3 months |
| pata/ConstatorMap.lue | Renove unused and expensive initialization logic from on SpatialConvo | I months |
| patielCrostRepLPR Las | cude consistency | 18 days |
| pataDrivieRomalizato | Spatia@createstive.Dvisive.Subtractive/Hormalization work with bat | 8 months |
| pate Dropout lue | small fix on error message | 6 months |
| patal racional das Poole | Adding Eractional Was Paoling | 3 months |
| patalFulConstation has | Add adjustment term to Spala/Ful/Convolution to control the size of | 5 days |
| gataFulCondutorMap ha | New NN classes | 3 pears |
| ipatial, PPusing lua | Spatial/versgePooling-dvides by kIIPAH | 10 months |
| patial/lasPooling.tus | SpatiaMaxPooling supports padding and cell mode | 6 months |
| pataMachpoling La | Add SpatiaMarDrynnling | 28 days |
| patial SoftMax lua | Update SoftMax to work in spatial mode | 4 months |
| patial Subfampling has | Marga branch ton_boit_seast | 3 years |
| patallubrativeRomate - | Spatial/Constructive/Delates/self-temalization work with bat- | 4 months |
| patial.)plangingNeared.1. | Use UNIX line and rgs | 7 months |
| ipetia/Zerr/Patting ka | Added more internative pretty-printing. | a yea |
| pitTable tus | Add support for negative indices in nn.SplitTable | 7 months |





Subhransu Maji, Chuang Gan and TAs Some slides kindly provided by Fei-Fei Li, Jiajun Wu, Erik Learned-Miller

Lecture 5 - 55 Sept. 19, 2023

```
local MulConstant, parent = torch.class('nn.MulConstant', 'nn.Module')
function MulConstant: init(constant scalar,ip)
 parent.__init(self)
  assert(type(constant scalar) == 'number', 'input is not scalar!')
  self.constant_scalar = constant_scalar
  -- default for inplace is false
  self.inplace = ip or false
  if (ip and type(ip) ~= 'boolean') then
     error('in-place flag must be boolean')
  end
end
function MulConstant:updateOutput(input)
 if self.inplace then
   input:mul(self.constant scalar)
   self.output = input
  else
   self.output:resizeAs(input)
   self.output:copy(input)
   self.output:mul(self.constant_scalar)
  end
 return self.output
end
function MulConstant:updateGradInput(input, gradOutput)
 if self.gradInput then
   if self.inplace then
      gradOutput:mul(self.constant_scalar)
     self.gradInput = gradOutput
      -- restore previous input value
     input:div(self.constant_scalar)
   else
      self.gradInput:resizeAs(gradOutput)
     self.gradInput:copy(gradOutput)
      self.gradInput:mul(self.constant_scalar)
   end
   return self.gradInput
  end
```

Example: Torch MulConstant

$$f(X) = aX$$

initialization

_forward()



Subhransu Maji, Chuang Gan and TAs Some slides kindly provided by Fei-Fei Li, Jiajun Wu, Erik Learned-Miller

Lecture 5 - 56 Sept. 19, 2023



Lecture 5 - 57 Sept. 19, 2023

[slides] [backprop notes] [Efficient BackProp] (optional) related: [1], [2], [3] (optional)

[slides]

handout 1: Vector, Matrix, and Tensor Derivatives handout 2: Derivatives, Backpropagation, and Vectorization Deep Learning [Nature] (optional)

[slides] tips/tricks: [1], [2] (optional)

Subhransu Maji, Chuang Gan and TAs Some slides kindly provided by Fei-Fei Li, Jiajun Wu, Erik Learned-Miller Lecture 5 - 58 Sept. 19, 2023

Vectorized operations



Subhransu Maji, Chuang Gan and TAs Some slides kindly provided by Fei-Fei Li, Jiajun Wu, Erik Learned-Miller Lecture 5 - 59 Sept. 19, 2023



Lecture 5 - 60 Sept. 19, 2023



Lecture 5 - 61 Sept. 19, 2023

Vectorized operations



Subhransu Maji, Chuang Gan and TAs Some slides kindly provided by Fei-Fei Li, Jiajun Wu, Erik Learned-Miller Lecture 5 - 62 Sept. 19, 2023

Assignment: Writing SVM/Softmax Stage your forward/backward computation!



Lecture 5 - 63

Sept. 19, 2023

Subhransu Maji, Chuang Gan and TAs Some slides kindly provided by Fei-Fei Li, Jiajun Wu, Erik Learned-Miller

Summary so far

- neural nets will be very large: no hope of writing down gradient formula by hand for all parameters
- backpropagation = recursive application of the chain rule along a computational graph to compute the gradients of all inputs/parameters/ intermediates
- implementations maintain a graph structure, where the nodes implement the forward() / backward() API.

Lecture 5 - 64 Sept. 19, 2023

- **forward**: compute result of an operation and save any intermediates needed for gradient computation in memory
- backward: apply the chain rule to compute the gradient of the loss function with respect to the inputs.

Subhransu Maji, Chuang Gan and TAs Some slides kindly provided by Fei-Fei Li, Jiajun Wu, Erik Learned-Miller