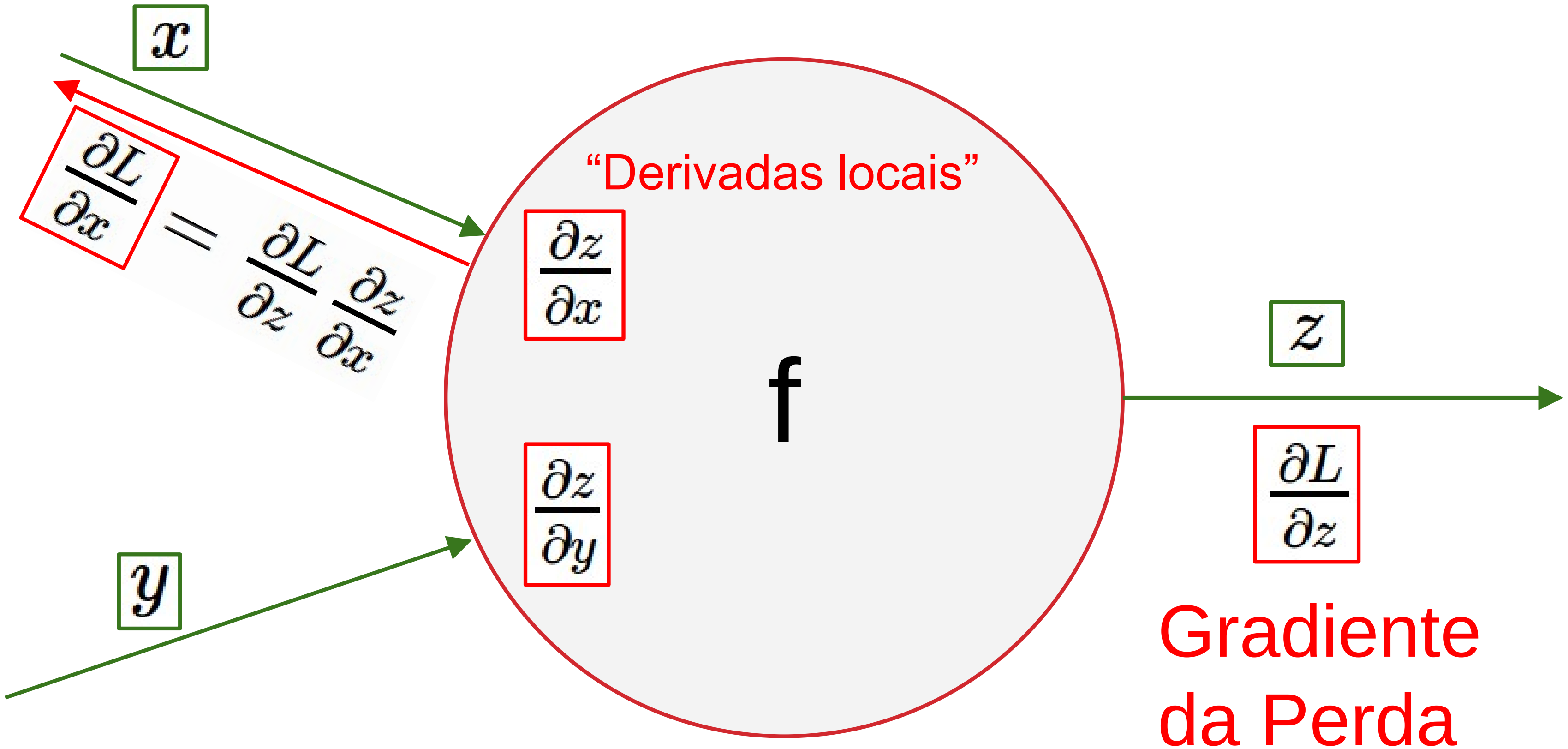
REDES NEURAIS ARTIFICIAIS

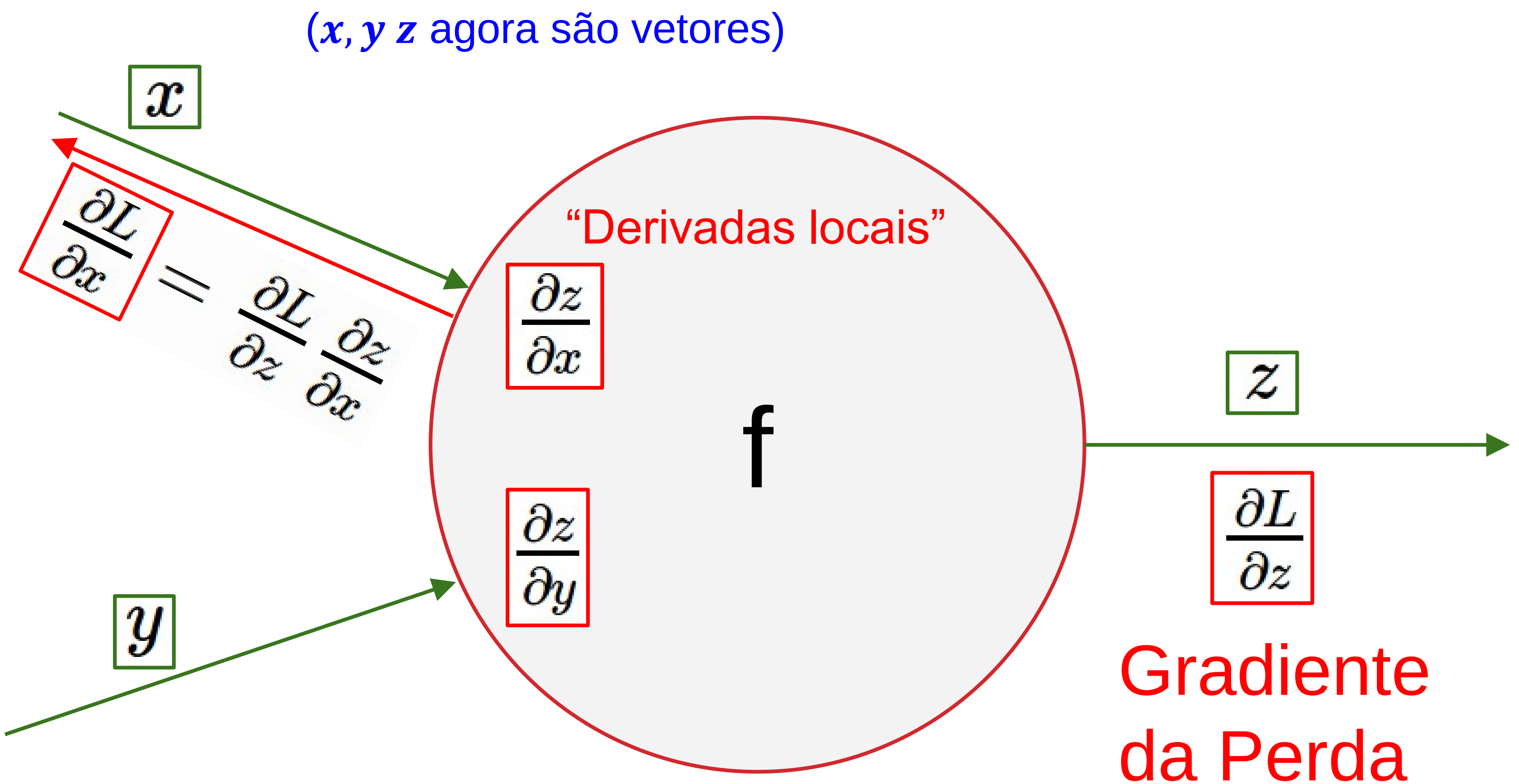
PROPAGAÇÃO RETRÓGRADA (III)

Zenilton K. G. Patrocínio Jr
zenilton@pucminas.br

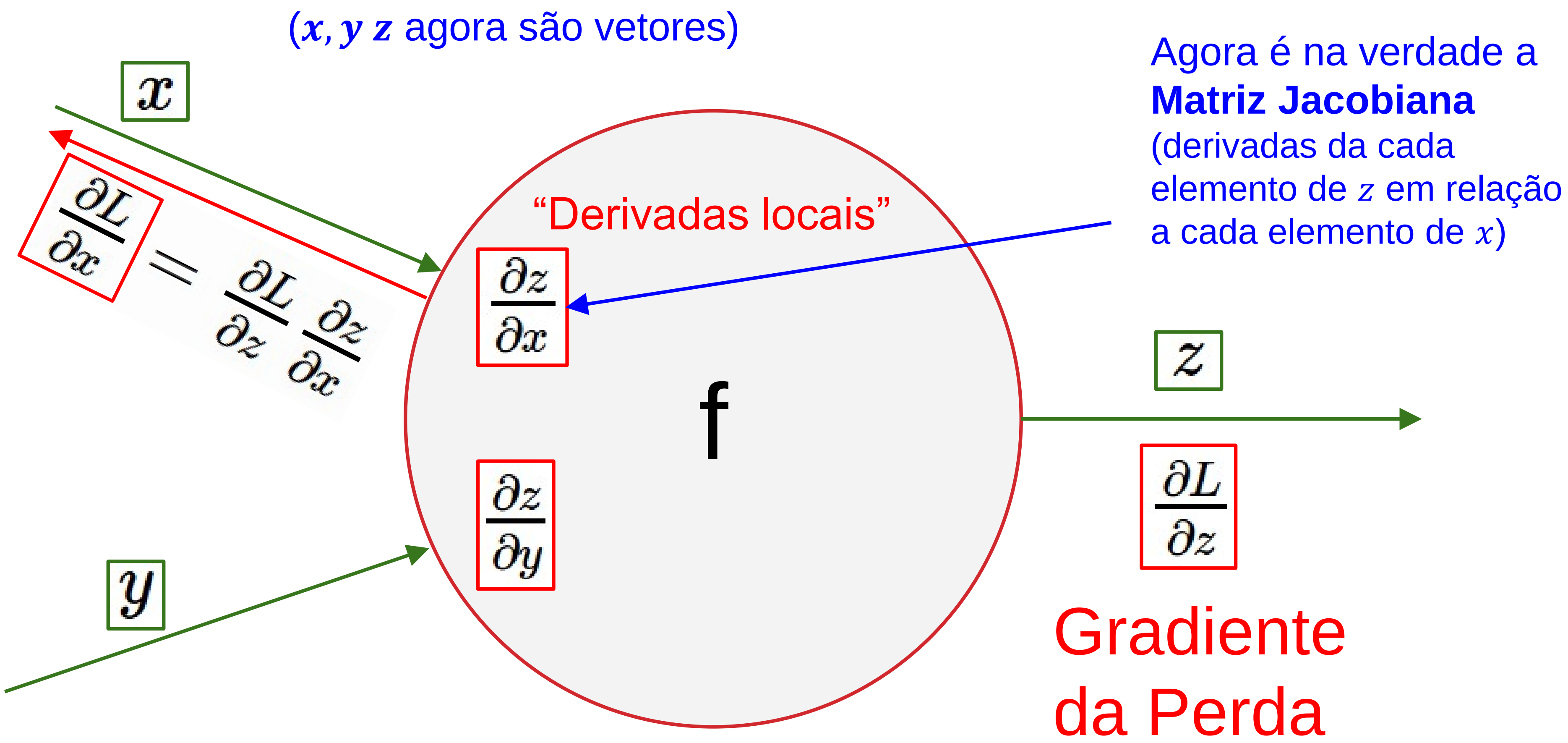
Gradientes para Dados Multidimensionais



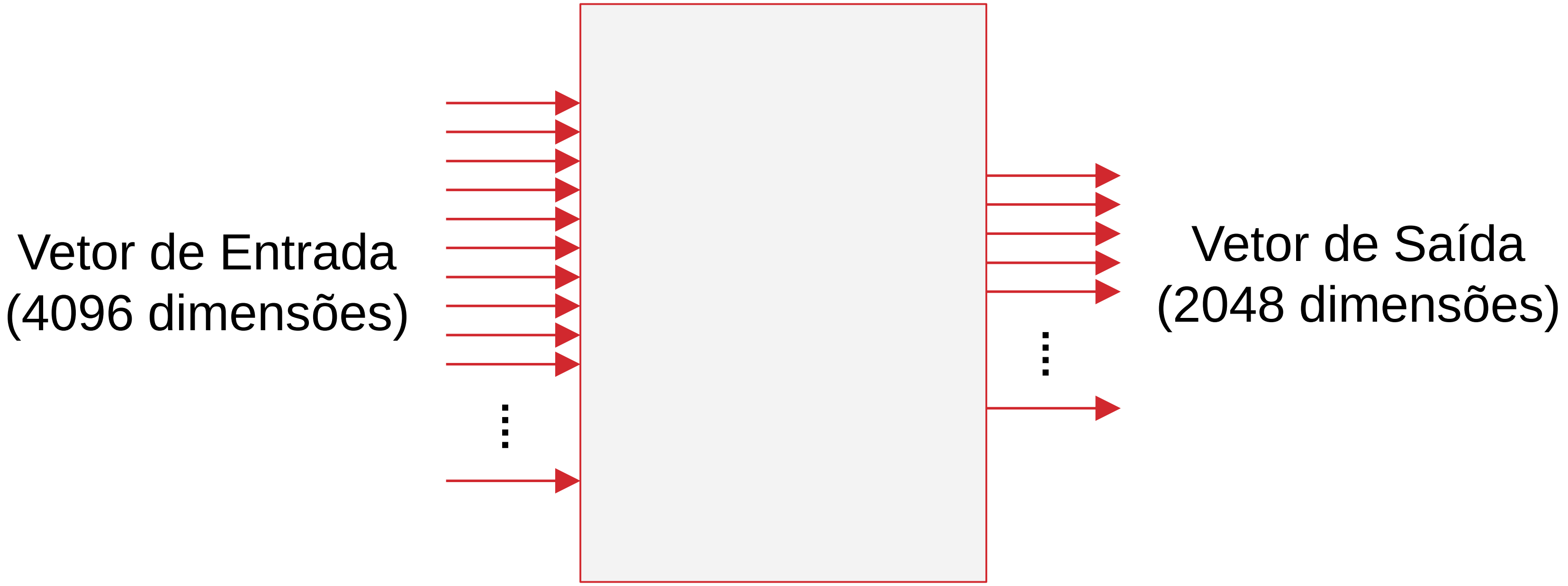
Gradientes para Dados Multidimensionais



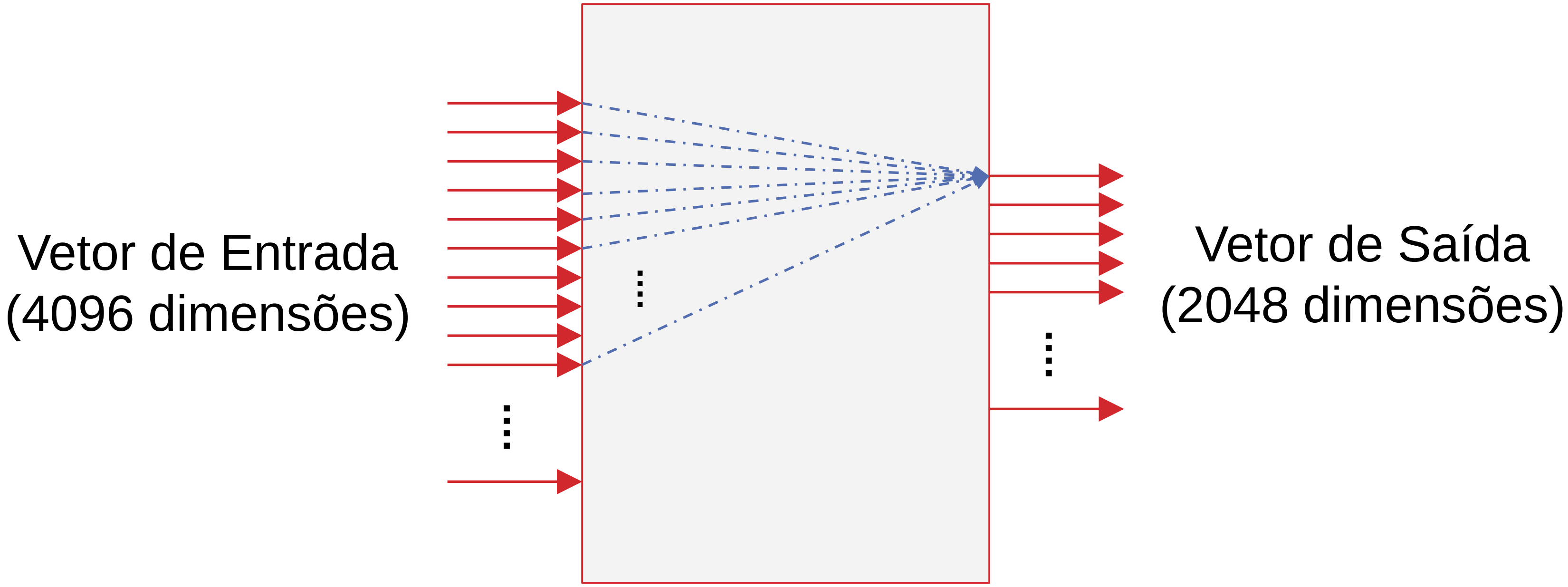
Gradientes para Dados Multidimensionais



Gradientes para Dados Multidimensionais



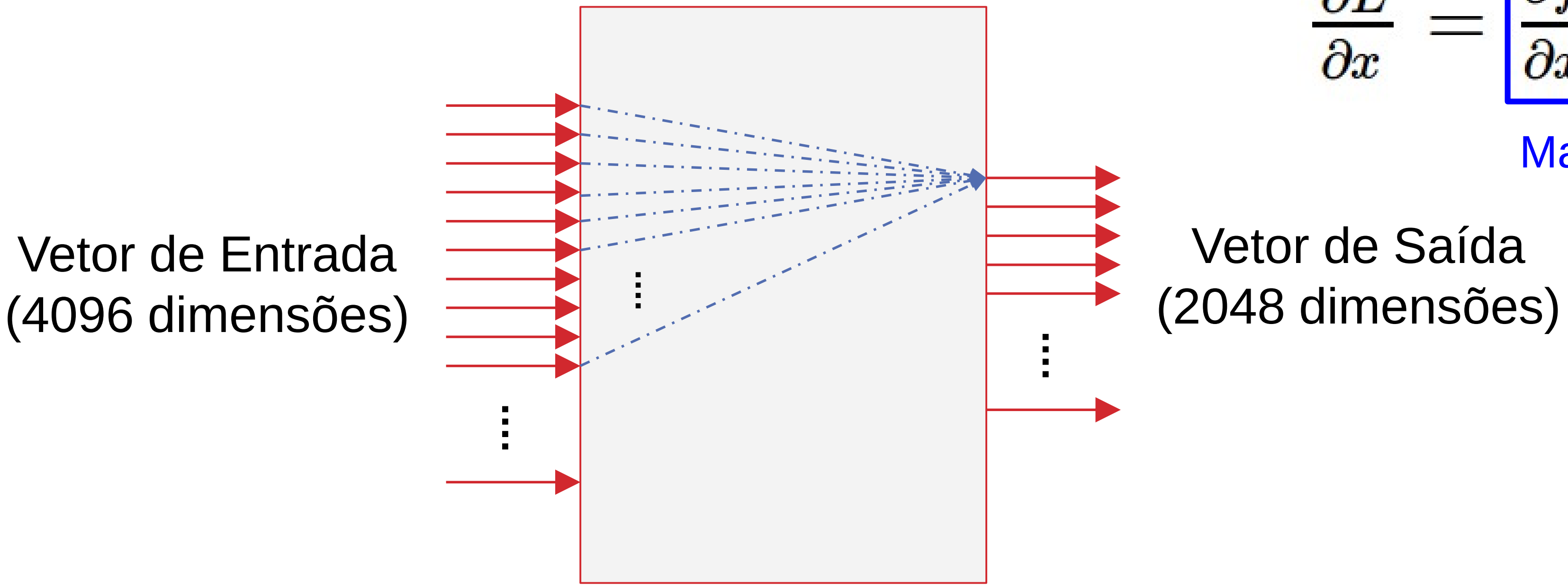
Gradientes para Dados Multidimensionais



Gradientes para Dados Multidimensionais

$$\frac{\partial L}{\partial x} = \boxed{\frac{\partial f}{\partial x}} \frac{\partial L}{\partial f}$$

Matriz Jacobiana

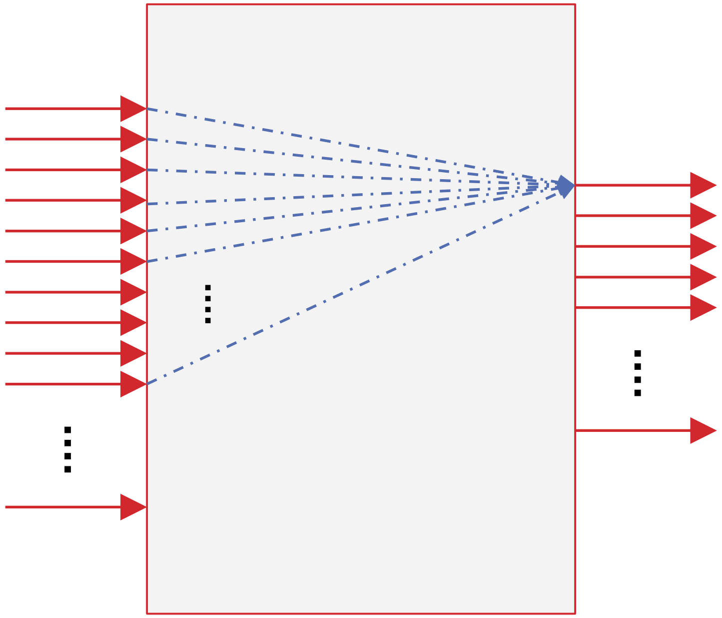


Gradientes para Dados Multidimensionais

$$\frac{\partial L}{\partial x} = \boxed{\frac{\partial f}{\partial x}} \frac{\partial L}{\partial f}$$

Matriz Jacobiana

Vetor de Entrada
(4096 dimensões)



Vetor de Saída
(2048 dimensões)

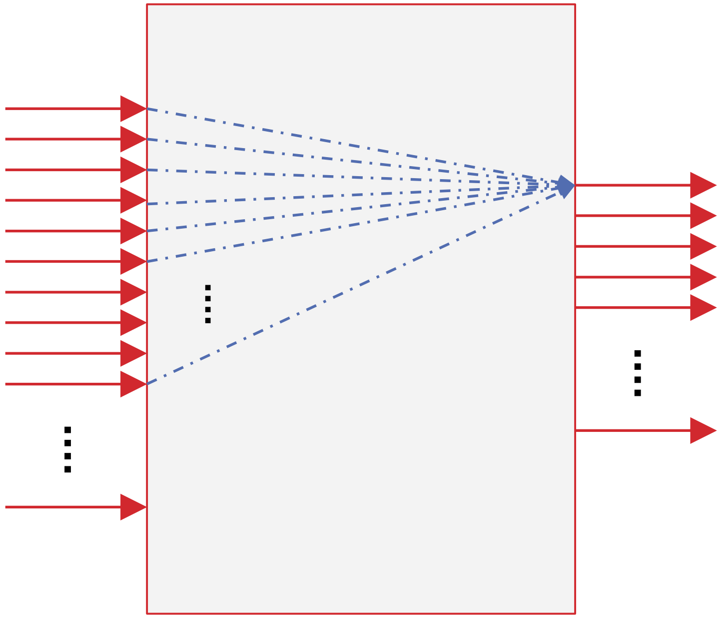
P1: qual o tamanho
da matriz Jacobiana?

Gradientes para Dados Multidimensionais

$$\frac{\partial L}{\partial x} = \boxed{\frac{\partial f}{\partial x}} \frac{\partial L}{\partial f}$$

Matriz Jacobiana

Vetor de Entrada
(4096 dimensões)



Vetor de Saída
(2048 dimensões)

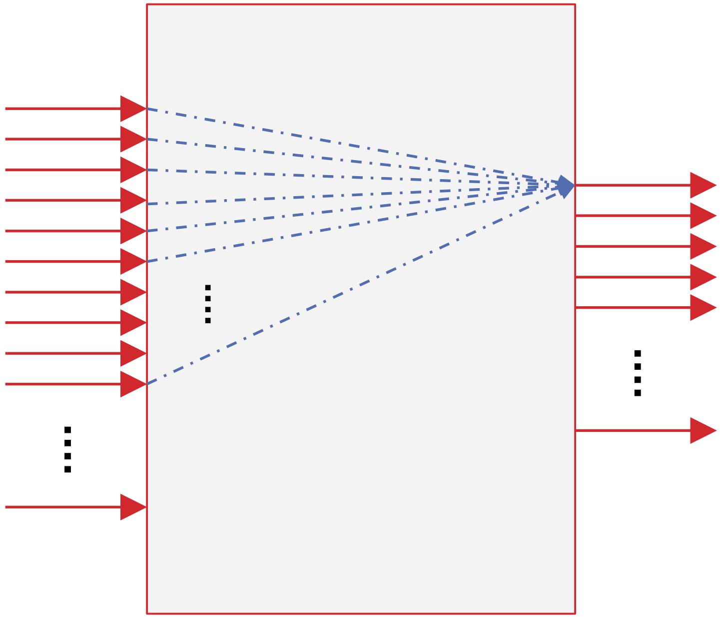
P1: qual o tamanho
da matriz Jacobiana?
$\Rightarrow$ 2048 × 4096 !

Gradientes para Dados Multidimensionais

$$\frac{\partial L}{\partial x} = \boxed{\frac{\partial f}{\partial x}} \frac{\partial L}{\partial f}$$

Matriz Jacobiana

Vetor de Entrada
(4096 dimensões)



Vetor de Saída
(2048 dimensões)

P1: qual o tamanho
da matriz Jacobiana?
$\Rightarrow$ 2048 × 4096 !

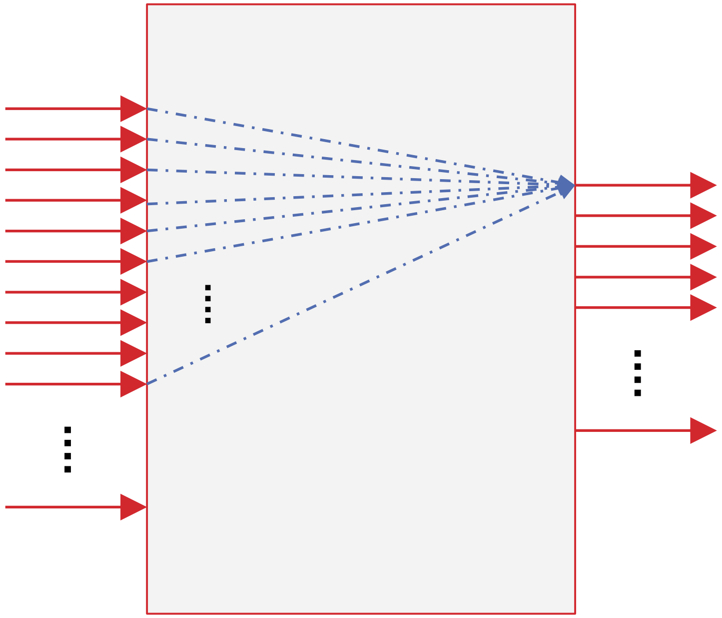
P2: qual a aparência
dessa matriz?

Gradientes para Dados Multidimensionais

$$\frac{\partial L}{\partial x} = \boxed{\frac{\partial f}{\partial x}} \frac{\partial L}{\partial f}$$

Matriz Jacobiana

Vetor de Entrada
(4096 dimensões)



Vetor de Saída
(2048 dimensões)

P1: qual o tamanho
da matriz Jacobiana?
$\Rightarrow$ 2048 × 4096 !

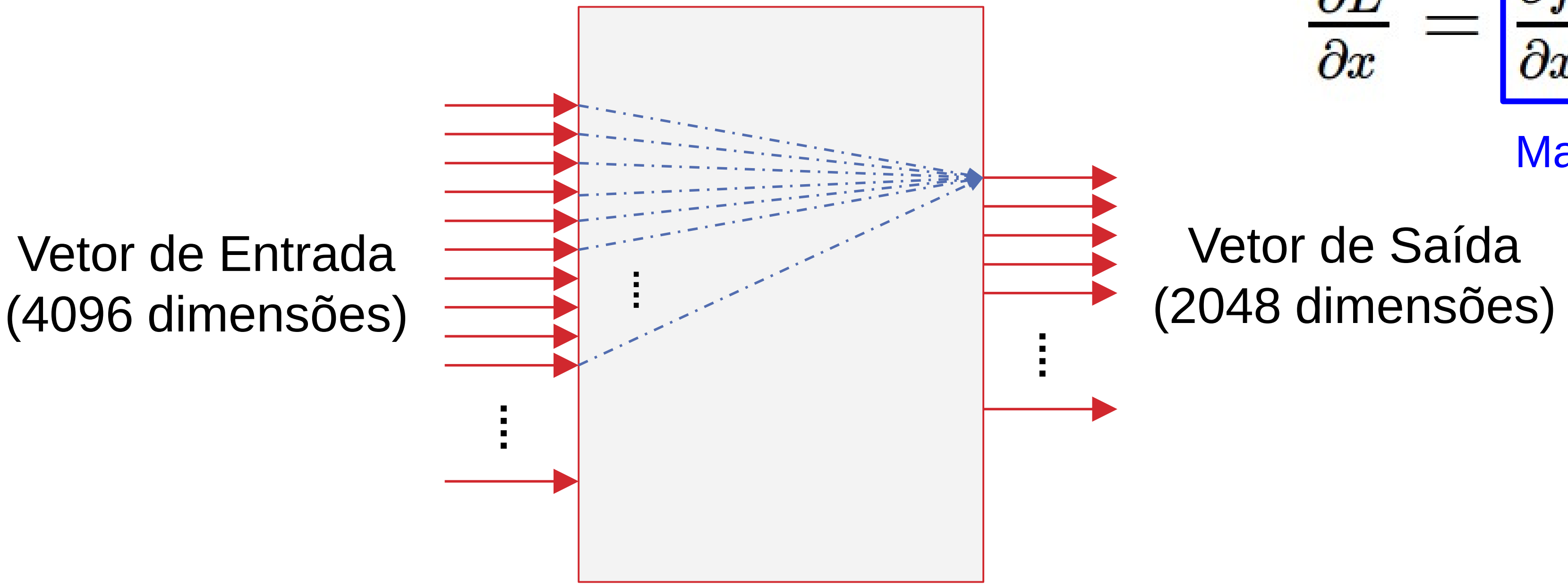
P2: qual a aparência
dessa matriz?

$$\begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_m} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_2}{\partial x_m} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_k}{\partial x_1} & \frac{\partial f_k}{\partial x_2} & \cdots & \frac{\partial f_k}{\partial x_m} \end{bmatrix}$$

Gradientes para Dados Multidimensionais

$$\frac{\partial L}{\partial x} = \boxed{\frac{\partial f}{\partial x}} \frac{\partial L}{\partial f}$$

Matriz Jacobiana



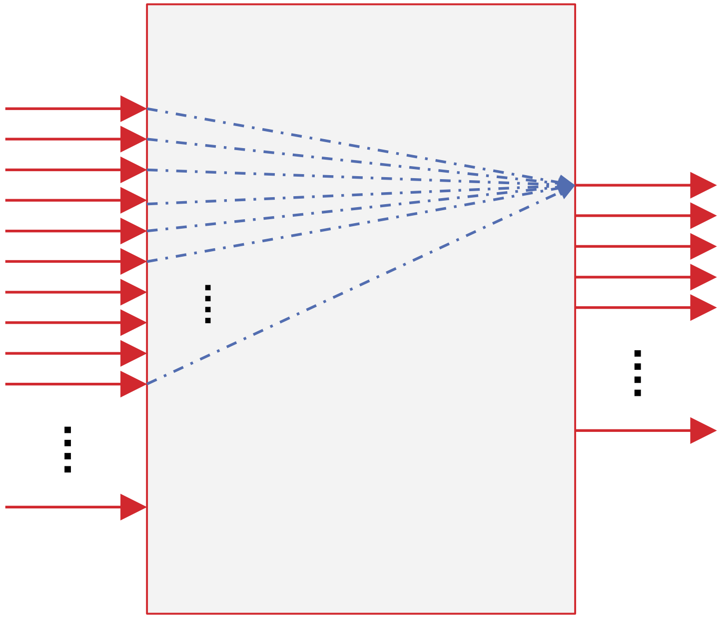
Na prática, processa-se todo um lote ou “*minibatch*” (p.ex. 100 amostras) de uma só vez

Gradientes para Dados Multidimensionais

$$\frac{\partial L}{\partial x} = \boxed{\frac{\partial f}{\partial x}} \frac{\partial L}{\partial f}$$

Matriz Jacobiana

Vetor de Entrada
(4096 dimensões)



Vetor de Saída
(2048 dimensões)

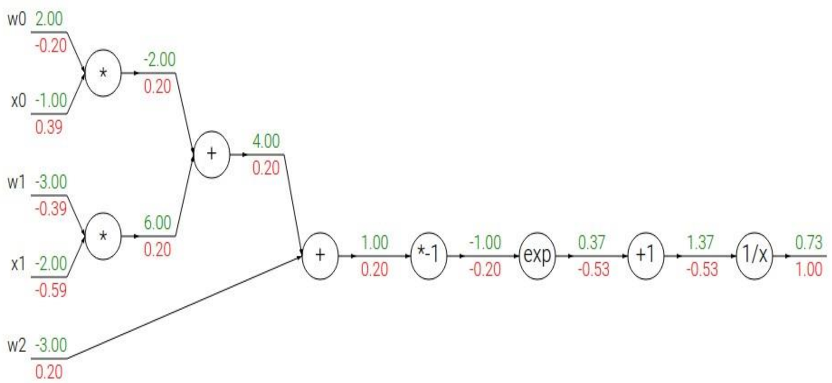
Na prática, processa-se todo um lote ou "*minibatch*" (p.ex. 100 amostras) de uma só vez

Assim, as dimensões da matriz Jacobiana desse exemplo seriam

204.800 × 409.600

≈ 83,9 bilhões de pesos :(

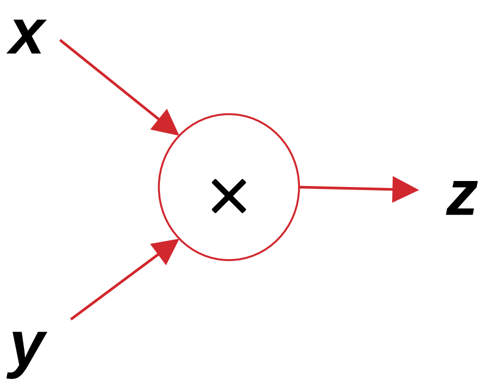
Implementação – Forward / Backward API



Para um grafo ou rede (*pseudocódigo*)

```
class ComputationalGraph(object):
    #...
    def forward(inputs):
        # 1. [pass inputs to input gates...]
        # 2. forward the computational graph:
        for gate in self.graph.nodes_topologically_sorted():
            gate.forward()
        return loss # the final gate in the graph outputs the loss
    def backward():
        for gate in reversed(self.graph.nodes_topologically_sorted()):
            gate.backward() # little piece of backprop (chain rule applied)
        return inputs_gradients
```

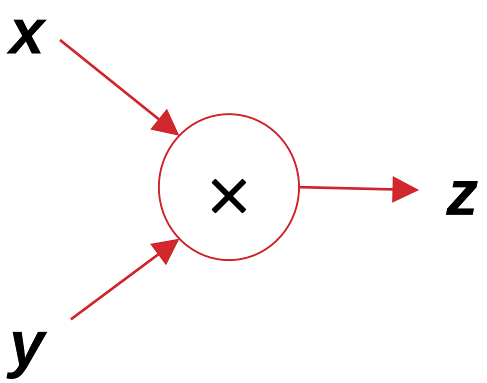
Implementação – Forward / Backward API



(x, y, z escalares)

```python
class MultiplyGate(object):
    def forward(x,y):
        z = x*y
        return z
    def backward(dz):
        # dx = ... #todo
        # dy = ... #todo
        return [dx, dy]
```

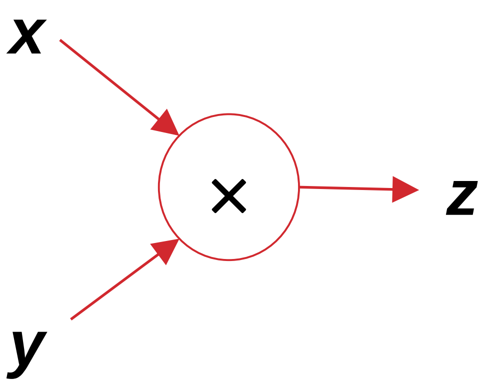
Implementação – Forward / Backward API



(x, y, z escalares)

```
class MultiplyGate(object):
    def forward(x,y):
        z = x*y
        return z
    def backward(dz):
        # dx = ... #todo
        # dy = ... #todo
        return [dx, dy]
```

Implementação – Forward / Backward API



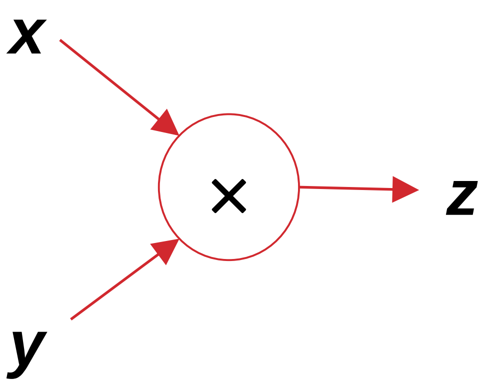
(x, y, z escalares)

```
class MultiplyGate(object):
    def forward(x,y):
        z = x*y
        return z
    def backward(dz):
        # dx = ... #todo
        # dy = ... #todo
        return [dx, dy]
```

$$\frac{\partial L}{\partial z}$$

$$\frac{\partial L}{\partial x}$$

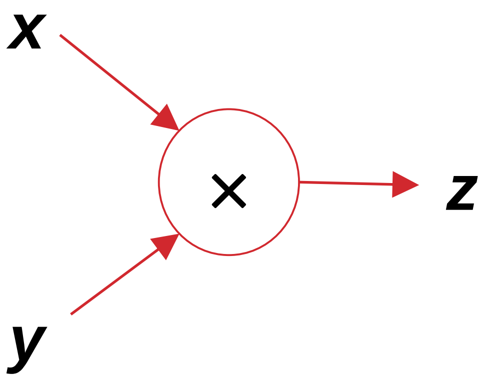
Implementação – Forward / Backward API



(x, y, z escalares)

```python
class MultiplyGate(object):
    def forward(x,y):
        z = x*y
        self.x = x # must keep these around!
        self.y = y
        return z
    def backward(dz):
        dx = self.y * dz # [dz/dx * dL/dz]
        dy = self.x * dz # [dz/dy * dL/dz]
        return [dx, dy]
```

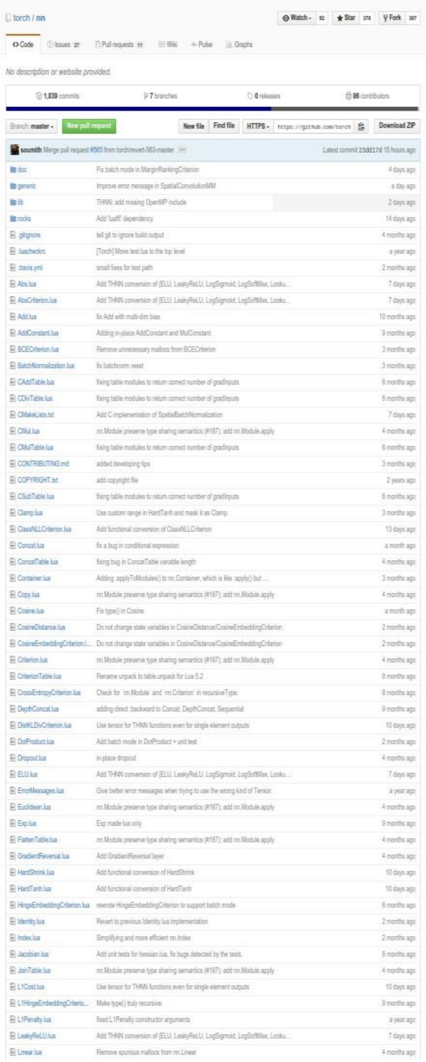
Implementação – Forward / Backward API



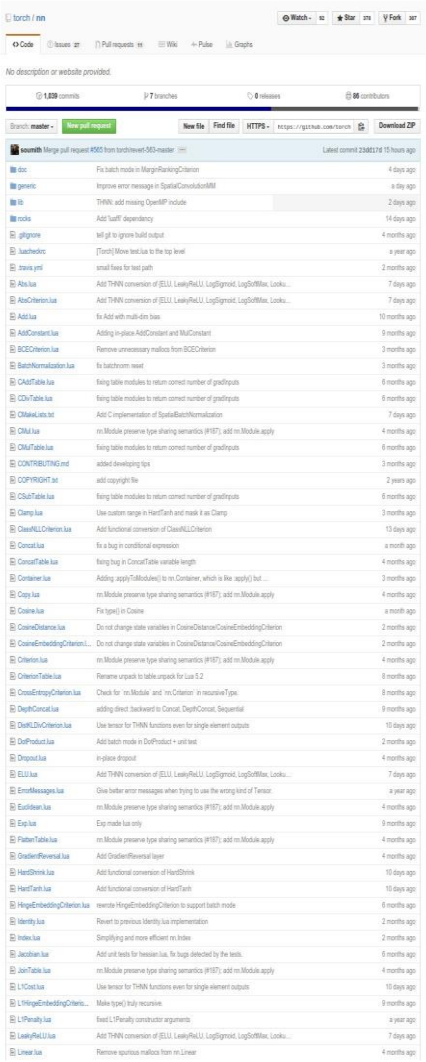
(x, y, z escalares)

```
class MultiplyGate(object):
    def forward(x,y):
        z = x*y
        self.x = x # must keep these around!
        self.y = y
        return z
    def backward(dz):
        dx = self.y * dz # [dz/dx * dL/dz]
        dy = self.x * dz # [dz/dy * dL/dz]
        return [dx, dy]
```

Biblioteca de Componentes / Camadas



Biblioteca de Componentes / Camadas



Biblioteca de Componentes / Camadas

torch / nn

Watch 92 | Star 378 | Fork 387

Code | Issues 27 | Pull requests 11 | Wiki | Pulse | Graphs

No description or website provided.

1,039 commits | 7 branches | 0 releases | 86 contributors

Branch: master | New pull request | New file | Find file | HTTPS | https://github.com/torch | Download ZIP

soumith Merge pull request #565 from torch/revert-563-master — Latest commit 23dd17d 15 hours ago

doc	Fix batch mode in MarginRankingCriterion	4 days ago
generic	Improve error message in SpatialConvolutionMM	a day ago
lib	THNN: add missing OpenMP include	2 days ago
rocks	Add 'luaffi' dependency	14 days ago
.gitignore	tell git to ignore build output	4 months ago
.luacheckrc	[Torch] Move test.lua to the top level	a year ago
.travis.yml	small fixes for test path	2 months ago
Abs.lua	Add THNN conversion of {ELU, LeakyReLU, LogSigmoid, LogSoftMax, Looku...	7 days ago
AbsCriterion.lua	Add THNN conversion of {ELU, LeakyReLU, LogSigmoid, LogSoftMax, Looku...	7 days ago
Add.lua	fix Add with multi-dim bias	10 months ago
AddConstant.lua	Adding in-place AddConstant and MulConstant	9 months ago
BCECriterion.lua	Remove unnecessary mallocs from BCECriterion	3 months ago
BatchNormalization.lua	fix batchnorm reset	3 months ago
CAddTable.lua	fixing table modules to return correct number of gradInputs	6 months ago
CDivTable.lua	fixing table modules to return correct number of gradInputs	6 months ago
CMakeLists.txt	Add C implementation of SpatialBatchNormalization	7 days ago
CMul.lua	nn.Module preserve type sharing semantics (#187); add nn.Module.apply	4 months ago
CMulTable.lua	fixing table modules to return correct number of gradInputs	6 months ago
CONTRIBUTING.md	added developing tips	3 months ago
COPYRIGHT.txt	add copyright file	2 years ago
CSubTable.lua	fixing table modules to return correct number of gradInputs	6 months ago
Clamp.lua	Use custom range in HardTanh and mask it as Clamp	3 months ago
ClassNLLCriterion.lua	Add functional conversion of ClassNLLCriterion	13 days ago
Concat.lua	fix a bug in conditional expression	a month ago
ConcatTable.lua	fixing bug in ConcatTable variable length	4 months ago
Container.lua	Adding :applyToModules() to nn.Container, which is like :apply() but ...	3 months ago
Copy.lua	nn.Module preserve type sharing semantics (#187); add nn.Module.apply	4 months ago
Cosine.lua	Fix type() in Cosine	a month ago
CosineDistance.lua	Do not change state variables in CosineDistance/CosineEmbeddingCriterion	2 months ago
CosineEmbeddingCriterion.l...	Do not change state variables in CosineDistance/CosineEmbeddingCriterion	2 months ago
Criterion.lua	nn.Module preserve type sharing semantics (#187); add nn.Module.apply	4 months ago
CriterionTable.lua	Rename unpack to table.unpack for Lua 5.2	8 months ago
CrossEntropyCriterion.lua	Check for `nn.Module` and `nn.Criterion` in recursiveType.	8 months ago
DepthConcat.lua	adding direct :backward to Concat, DepthConcat, Sequential	9 months ago
DistKLDivCriterion.lua	Use tensor for THNN functions even for single element outputs	10 days ago
DotProduct.lua	Add batch mode in DotProduct + unit test	2 months ago
Dropout.lua	in-place dropout	4 months ago
ELU.lua	Add THNN conversion of {ELU, LeakyReLU, LogSigmoid, LogSoftMax, Looku...	7 days ago
ErrorMessages.lua	Give better error messages when trying to use the wrong kind of Tensor.	a year ago
Euclidean.lua	nn.Module preserve type sharing semantics (#187); add nn.Module.apply	4 months ago
Exp.lua	Exp made lua only	9 months ago
FlattenTable.lua	nn.Module preserve type sharing semantics (#187); add nn.Module.apply	4 months ago
GradientReversal.lua	Add GradientReversal layer	4 months ago
HardShrink.lua	Add functional conversion of HardShrink	10 days ago
HardTanh.lua	Add functional conversion of HardTanh	10 days ago
HingeEmbeddingCriterion.lua	rewrote HingeEmbeddingCriterion to support batch mode	6 months ago
Identity.lua	Revert to previous Identity.lua implementation	2 months ago
Index.lua	Simplifying and more efficient nn.Index	2 months ago
Jacobian.lua	Add unit tests for hessian.lua, fix bugs detected by the tests.	6 months ago
JoinTable.lua	nn.Module preserve type sharing semantics (#187); add nn.Module.apply	4 months ago
L1Cost.lua	Use tensor for THNN functions even for single element outputs	10 days ago
L1HingeEmbeddingCriterio...	Make type() truly recursive.	9 months ago
L1Penalty.lua	fixed L1Penalty constructor arguments	a year ago
LeakyReLU.lua	Add THNN conversion of {ELU, LeakyReLU, LogSigmoid, LogSoftMax, Looku...	7 days ago
Linear.lua	Remove spurious mallocs from nn.Linear	4 months ago

LogSigmoid.lua	Add THNN conversion of {ELU, LeakyReLU, LogSigmoid, LogSoftMax, Looku...	7 days ago
LogSoftMax.lua	Add THNN conversion of {ELU, LeakyReLU, LogSigmoid, LogSoftMax, Looku...	7 days ago
LookupTable.lua	Harmonize LookupTable signature with cunn impl	5 days ago
MM.lua	Rename unpack to table.unpack for Lua 5.2	8 months ago
MSECriterion.lua	Add SizeAverage to criterions in the constructor	2 months ago
MarginCriterion.lua	modernized MarginCriterion	a year ago
MarginRankingCriterion.lua	Fix batch mode in MarginRankingCriterion	4 days ago
Max.lua	Merge pull request #464 from vgire/master	2 months ago
Mean.lua	Add support for negative dimension and both batch and non batch input...	2 months ago
Min.lua	Merge pull request #464 from vgire/master	2 months ago
MixtureTable.lua	cancel unused variable and useless expression	29 days ago
Module.lua	Revert "Don't re-flatten parameters if they are already flattened"	15 hours ago
Mul.lua	removing the requirement for providing size in nn.Mul	a year ago
MulConstant.lua	Ignore updateGradInput if self.gradInput is nil	3 months ago
MultiCriterion.lua	asserts in MultiCriterion and ParallelCriterion add	2 months ago
MultiLabelMarginCriterion.lua	initial revamp of torch7 tree	4 years ago
MultiMarginCriterion.lua	multimargin supports p=2	11 months ago
Narrow.lua	typeAs in Narrow not done in place.	6 months ago
NarrowTable.lua	NarrowTable	6 months ago
Normalize.lua	Remove bmm and baddbmm from Normalize, because they allocate memory, ...	20 days ago
PReLU.lua	Buffers for PReLU cuda implementation.	8 months ago
Padding.lua	fixed broken nn.Padding: input was returned in backprop	5 months ago
PairwiseDistance.lua	Merge pull request #532 from xwgeng/master	29 days ago
Parallel.lua	fix a bug in conditional expression	a month ago
ParallelCriterion.lua	asserts in MultiCriterion and ParallelCriterion add	2 months ago
ParallelTable.lua	Parallel optimization. ParallelTable inherits Container. unit tests	a year ago
Power.lua	Use UNIX line endings	7 months ago
README.md	doc readthedocs	5 months ago
RReLU.lua	Add randomized leaky rectified linear unit (RReLU)	3 months ago
ReLU.lua	adds in-place ReLU and fixes a potential divide-by-zero in nn.Sqrt	9 months ago
Replicate.lua	Replicate batchMode	8 months ago
Reshape.lua	Added more informative pretty-printing.	a year ago
Select.lua	initial revamp of torch7 tree	4 years ago
SelectTable.lua	nn.Module preserve type sharing semantics (#187); add nn.Module.apply	4 months ago
Sequential.lua	fixing Sequential.remove corner case	6 months ago
Sigmoid.lua	initial revamp of torch7 tree	4 years ago
SmoothL1Criterion.lua	Add SizeAverage to criterions in the constructor	2 months ago
SoftMax.lua	Fix various unused variables in nn	a year ago
SoftMin.lua	Fix various unused variables in nn	a year ago
SoftPlus.lua	fixed a numerical issue in the SoftPlus module (it breaks for input g...	2 years ago
SoftShrink.lua	initial revamp of torch7 tree	4 years ago
SoftSign.lua	initial revamp of torch7 tree	4 years ago
SparseJacobian.lua	Fix various unused variables in nn	a year ago
SparseLinear.lua	Using sparse implementation of zeroGradParameters for SparseLinear	a month ago
SpatialAdaptiveMaxPooling....	Added SpatialAdaptiveMaxPooling	a year ago
SpatialAveragePooling.lua	SpatialAveragePooling supports padding, ceil mode and exclude_pad div...	29 days ago
SpatialBatchNormalization.lua	Add C implementation of SpatialBatchNormalization	7 days ago
SpatialContrastiveNormaliz...	Make type() truly recursive.	9 months ago
SpatialConvolution.lua	Fix type() in SpatialConvolution	3 months ago
SpatialConvolutionMM.lua	Fix type() in SpatialConvolution	3 months ago
SpatialConvolutionMap.lua	Remove unused and expensive initialization logic from nn.SpatialConvo...	8 months ago
SpatialCrossMapLRN.lua	cuda consistency	18 days ago
SpatialDivisiveNormalizatio...	Spatial{Constrastive,Divisive,Subtractive}Normalization work with bat...	8 months ago
SpatialDropout.lua	small fix on error message	6 months ago
SpatialFractionalMaxPoolin...	Adding Fractional Max Pooling	3 months ago
SpatialFullConvolution.lua	Add adjustment term to SpatialFullConvolution to control the size of ...	5 days ago
SpatialFullConvolutionMap.lua	New NN classes	3 years ago
SpatialLPPooling.lua	SpatialAveragePooling divides by kW*kH	10 months ago
SpatialMaxPooling.lua	SpatialMaxPooling supports padding and ceil mode	6 months ago
SpatialMaxUnpooling.lua	Add SpatialMaxUnpooling	26 days ago
SpatialSoftMax.lua	Update SoftMax to work in spatial mode	4 months ago
SpatialSubSampling.lua	Merge branch 'nn_fast_reset'	3 years ago
SpatialSubtractiveNormaliz...	Spatial{Constrastive,Divisive,Subtractive}Normalization work with bat...	8 months ago
SpatialUpSamplingNearest.l...	Use UNIX line endings	7 months ago
SpatialZeroPadding.lua	Added more informative pretty-printing.	a year ago
SplitTable.lua	Add support for negative indices in nn.SplitTable	7 months ago