nikhil_supekar_ns4486_A1_code

February 18, 2020

[36]:
```python
# Import dependencies
import random
import numpy as np
import torch
import torch.nn as nn

# You can find Alfredo's plotting code in plot_lib.py in this directory .
# Download it along with this assignment and keep it in the same directory.
from plot_lib import set_default, show_scatterplot, plot_bases

from matplotlib.pyplot import plot, title, axis

from time import perf_counter
```

[37]:
```python
# Set up your device
cuda = torch.cuda.is_available()
device = torch.device("cuda:0" if cuda else "cpu")
```

[38]:
```python
# Set up random seed to 1008. Do not change the random seed.
# Yes, these are all necessary when you run experiments!

def set_random_seed(seed):
    random.seed(seed)
    np.random.seed(seed)
    torch.manual_seed(seed)

set_random_seed(1008)
if cuda:
    torch.cuda.manual_seed(seed)
    torch.cuda.manual_seed_all(seed)
    torch.backends.cudnn.benchmark = False
    torch.backends.cudnn.deterministic = True
```

1 1. Full, slice, fill

Write a function `warm_up` that returns the 2D tensor with integers below. **Do not use any loops**.

```
1 2 1 1 1 1 2 1 1 1 1 2 1
2 2 2 2 2 2 2 2 2 2 2 2 2
1 2 1 1 1 1 2 1 1 1 1 2 1
1 2 1 3 3 1 2 1 3 3 1 2 1
1 2 1 3 3 1 2 1 3 3 1 2 1
1 2 1 1 1 1 2 1 1 1 1 2 1
2 2 2 2 2 2 2 2 2 2 2 2 2
1 2 1 1 1 1 2 1 1 1 1 2 1
1 2 1 3 3 1 2 1 3 3 1 2 1
1 2 1 3 3 1 2 1 3 3 1 2 1
1 2 1 1 1 1 2 1 1 1 1 2 1
2 2 2 2 2 2 2 2 2 2 2 2 2
1 2 1 1 1 1 2 1 1 1 1 2 1
```

Hint: Use `torch.full`, `torch.fill_`, and the slicing operator.

[39]:
```python
def warm_up():
    x = torch.full((13, 13), 1)

    x[:, 1] = x[1, :] = 2
    x[:, 6] = x[6, :] = 2
    x[:, 11] = x[11, :] = 2

    x[3:5, 3:5] = x[3:5, 8:10] = x[8:10, 3:5] = x[8:10, 8:10] = 3

    return x

# Uncomment line below once you implement this function.
print(warm_up())
```

```
tensor([[1., 2., 1., 1., 1., 1., 2., 1., 1., 1., 1., 2., 1.],
        [2., 2., 2., 2., 2., 2., 2., 2., 2., 2., 2., 2., 2.],
        [1., 2., 1., 1., 1., 1., 2., 1., 1., 1., 1., 2., 1.],
        [1., 2., 1., 3., 3., 1., 2., 1., 3., 3., 1., 2., 1.],
        [1., 2., 1., 3., 3., 1., 2., 1., 3., 3., 1., 2., 1.],
        [1., 2., 1., 1., 1., 1., 2., 1., 1., 1., 1., 2., 1.],
        [2., 2., 2., 2., 2., 2., 2., 2., 2., 2., 2., 2., 2.],
        [1., 2., 1., 1., 1., 1., 2., 1., 1., 1., 1., 2., 1.],
        [1., 2., 1., 3., 3., 1., 2., 1., 3., 3., 1., 2., 1.],
        [1., 2., 1., 3., 3., 1., 2., 1., 3., 3., 1., 2., 1.],
        [1., 2., 1., 1., 1., 1., 2., 1., 1., 1., 1., 2., 1.],
        [2., 2., 2., 2., 2., 2., 2., 2., 2., 2., 2., 2., 2.],
        [1., 2., 1., 1., 1., 1., 2., 1., 1., 1., 1., 2., 1.]])
```

2 2. To Loop or not to loop

The motivation for the following three sub-questions is to get you to think critically about how to write your deep learning code. These sorts of choices can make the difference between tractable and intractable model training.

2.1 2.1. `mul_row_loop`

Write a function `mul_row_loop`, using python loops with simple indexing but no advanced indexing/slicing, that receives a 2D tensor as input and returns a tensor of same size that is - equal to the input on the first row - 2 times the input's second row on the second row - 3 times the input's third row on the third row - etc..

For instance:

```
>>> t = torch.full((4, 8), 2.0)
>>> t
tensor([[2., 2., 2., 2., 2., 2., 2., 2.],
[2., 2., 2., 2., 2., 2., 2., 2.],
[2., 2., 2., 2., 2., 2., 2., 2.],
[2., 2., 2., 2., 2., 2., 2., 2.]])
>>> mul_row(t)
tensor([[2., 2., 2., 2., 2., 2., 2., 2.],
[4., 4., 4., 4., 4., 4., 4., 4.],
[6., 6., 6., 6., 6., 6., 6., 6.],
[8., 8., 8., 8., 8., 8., 8., 8.]])
```

[40]:
```
def mul_row_loop(input_tensor):
    rows = input_tensor.size()[0]
    cols = input_tensor.size()[1]
    output_tensor = torch.full((rows, cols), 1)

    for i in range(rows):
        for j in range(cols):
            output_tensor[i, j] = input_tensor[i, j] * (i + 1)

    return output_tensor
```

3 2.2. `mul_row_fast`

Write a second version of the same function named `mul_row_fast` which uses tensor operations and no looping.

Hint: Use broadcasting and `torch.arange`, `torch.view`, and `torch.mul`.

[41]:
```
def mul_row_fast(input_tensor):
    num_rows = input_tensor.size()[0]
    x = torch.arange(1, num_rows + 1).view(num_rows, 1)
    return input_tensor.mul(x)
```

4 2.3. `times`

Write a function `times` which takes a 2D tensor as input and returns the run times of `mul_row_loop` and `mul_row_fast` on this tensor, respectively. Use `time.perf_counter`.

Use `torch.ones` to create a 2D tensor of size (1000, 400) full of ones and run `times` on it (there should be more than two orders of magnitude difference).

[42]:
```
from time import perf_counter
def times(input_tensor):
    t1 = perf_counter()
    mul_row_loop(input_tensor)
    t2 = perf_counter()
    mul_row_fast(input_tensor)
    t3 = perf_counter()

    return t2 - t1, t3 - t2

# Uncomment lines below once you implement this function.
input_tensor = torch.ones(1000, 400)
time_1, time_2 = times(input_tensor)
print('{}, {}'.format(time_1, time_2))
```

```
5.5979539290000275, 0.00035807199992632377
```

5 3. Non-linearities

In this section, we explore similar concepts to Lab 1 and get comfortable initializing modules like nn.Linear and using non-linearities in PyTorch.

5.1 3.1. ReLU

ReLU (Rectified Linear Unit) is a non-linear activation fuction defined as:

$$y = \max(0, x)$$

Define a fully connected neural network `linear_fc_relu` which: - takes 2 dimensional data as input and passes it through linear modules (`torch.nn.Linear`) - has one hidden layer of dimension 5 - has output dimension of 2 - has ReLu as an activation function

Create a tensor with input data X of size (100, 2) using `torch.randn`.

Following the example in https://github.com/Atcold/pytorch-Deep-Learning-Minicourse/blob/master/02-space_stretching.ipynb, visualize the output of passing `X` through the neural network `linear_fc_relu`.

You can find Alfredo's plotting code in plot_lib.py in this directory. Download it along with this assignment and keep it in the same directory.

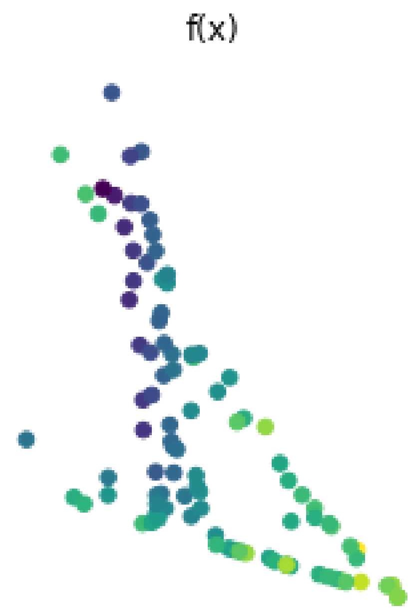
[43]:
```
# Input data
set_random_seed(1008)
X = torch.randn(100, 2)
```

[45]:
```
# create 1-layer neural networks with ReLU activation
set_random_seed(1008)
linear_fc_relu = nn.Sequential(
    nn.Linear(2, 5),
    nn.ReLU(),
```

```
    nn.Linear(5, 2)
)

linear_fc_relu.to(device)

with torch.no_grad():
    Y = linear_fc_relu(X)
show_scatterplot(X, X[:, 0], title='f(x)')
show_scatterplot(Y, X[:, 0], title='f(x)')
```



5.2 3.2. Sigmoid

The sigmoid function is another popular choice for a non-linear activation function which maps its input to values in the interval $(0, 1)$. It is formally defined as:

$$\sigma(x) = \frac{1}{1 + exp[-x]}$$

Define a new neural network `linear_fc_sigmoid` which is the same architecture as in part `3.1.` but with a sigmoid unit instead of ReLU.

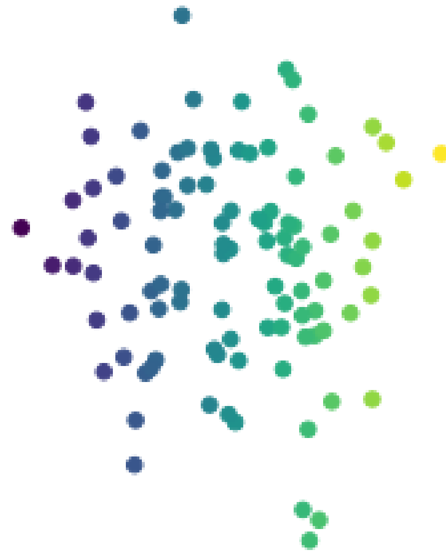
Using the same X as in part `3.1`, visualize the output of passing `X` through the neural network `linear_fc_sigmoid`.

[47]:
```
# create 1-layer neural networks with Sigmoid activation
set_random_seed(1008)
linear_fc_sigmoid = nn.Sequential(
    nn.Linear(2, 5),
    nn.Sigmoid(),
    nn.Linear(5, 2)
)

linear_fc_sigmoid.to(device)

with torch.no_grad():
    Y = linear_fc_sigmoid(X)
show_scatterplot(X, X[:, 0], title='f(x)')
show_scatterplot(Y, X[:, 0], title='f(x)')
```

f(x)



f(x)