In [ ]:
```
%load_ext autoreload
%autoreload 2

import numpy as np
from torch import nn
import torch
from d2l import torch as d2l

import mytorch
from mytorch import nn as mynn
```

In [ ]:
```
# load data for testing
data = np.load('testerData.npz')
W, b, X, Y, dLdZ = [data[fname] for fname in data.files]

[N, num_inputs] = X.shape
num_outputs = Y.shape[1]

# converted torch versions
Xt = torch.tensor(X).float()
Wt = torch.tensor(W).float()
bt = torch.tensor(b).float()
Yt = torch.tensor(Y).float()
```

In [ ]:
```
# initialize model and fix weights to true values
my_net = mynn.Linear(num_inputs, num_outputs)
my_net.W = W
my_net.b = b.flatten()

# initialize torch model, loss, optimizer
net = nn.Linear(num_inputs, num_outputs)
net.weight = nn.Parameter(Wt.T)
net.bias = nn.Parameter(bt[:, 0])
criterion = nn.MSELoss()
optimizer = torch.optim.SGD(net.parameters(), lr=0.1, momentum=0.0)
```

Compare `forward()`

In [ ]:
```
true_out = X @ W + np.outer(np.ones(N), b)
my_out = my_net.forward(X)
torch_out = net(Xt)

print('True:\n', true_out, '\n')
print('MyTorch:\n', my_out, '\n')
print('PyTorch:\n', torch_out.data, '\n')

print('Difference:', np.linalg.norm(my_out - torch_out.data.numpy()))
```

```
True:
 [[2.45419505 3.85377348 3.86239248 4.09861525 3.42552912]
 [2.37093331 3.22945671 2.88333967 3.03220271 2.72105395]
 [1.90472749 2.88840363 2.93526692 3.4361838  3.17490315]
 [1.45392748 2.79782519 2.16759199 2.56624407 2.53054982]
 [2.04588663 3.40446758 2.76237953 3.22387254 2.86430505]]

MyTorch:
 [[2.45419505 3.85377348 3.86239248 4.09861525 3.42552912]
 [2.37093331 3.22945671 2.88333967 3.03220271 2.72105395]
 [1.90472749 2.88840363 2.93526692 3.4361838  3.17490315]
 [1.45392748 2.79782519 2.16759199 2.56624407 2.53054982]
 [2.04588663 3.40446758 2.76237953 3.22387254 2.86430505]]

PyTorch:
 tensor([[2.4542, 3.8538, 3.8624, 4.0986, 3.4255],
        [2.3709, 3.2295, 2.8833, 3.0322, 2.7211],
        [1.9047, 2.8884, 2.9353, 3.4362, 3.1749],
        [1.4539, 2.7978, 2.1676, 2.5662, 2.5305],
        [2.0459, 3.4045, 2.7624, 3.2239, 2.8643]])

Difference: 7.96027202151818e-07
```

Compare `backward` and gradients

In [ ]:
```
my_net.backward(dLdZ)
my_dLdW = my_net.dLdW
my_dLdb = my_net.dLdb

optimizer.zero_grad()
torch_loss_fn = nn.MSELoss()
torch_loss = torch_loss_fn(torch_out, Yt)
torch_loss.backward(retain_graph=True)
torch_dLdW = net.weight.grad.data
torch_dLdb = net.bias.grad.data

print('MyTorch dLdW:\n', my_dLdW, '\n')
print('PyTorch dLdW:\n', torch_dLdW.T, '\n')
print('MyTorch dLdb:\n', my_dLdb, '\n')
print('PyTorch dLdb:\n', torch_dLdb, '\n')

print('Difference in dLdW:', np.linalg.norm(my_dLdW.T - torch_dLdW.data.numpy()))
print('Difference in dLdb:', np.linalg.norm(my_dLdb.flatten() - torch_dLdb.data.num
```

```
MyTorch dLdW:
 [[ 0.12012567  0.00295367  0.04191126 -0.07267667  0.02791693]
 [ 0.03518954 -0.0200141   0.03406723 -0.02431096  0.02070968]
 [ 0.14748257 -0.01532951  0.08382155 -0.01930616 -0.0203474 ]
 [ 0.06681007 -0.03891314  0.08080953  0.01504673  0.00045012]
 [ 0.11256331 -0.01973837  0.01680606 -0.00138297  0.00995075]
 [ 0.01978376 -0.02740623  0.02936517  0.03265669 -0.00413681]
 [ 0.13485436 -0.00176794  0.03313958 -0.00091217 -0.01160413]
 [ 0.02693809 -0.02455704  0.0399628   0.00525783  0.00168999]
 [ 0.14058303 -0.02630857  0.0610937  -0.01097525  0.00104313]
 [ 0.06542919 -0.01061266 -0.00446366  0.03434501 -0.00295966]]

PyTorch dLdW:
 tensor([[ 0.1201,  0.0030,  0.0419, -0.0727,  0.0279],
        [ 0.0352, -0.0200,  0.0341, -0.0243,  0.0207],
        [ 0.1475, -0.0153,  0.0838, -0.0193, -0.0203],
        [ 0.0668, -0.0389,  0.0808,  0.0150,  0.0005],
        [ 0.1126, -0.0197,  0.0168, -0.0014,  0.0100],
        [ 0.0198, -0.0274,  0.0294,  0.0327, -0.0041],
        [ 0.1349, -0.0018,  0.0331, -0.0009, -0.0116],
        [ 0.0269, -0.0246,  0.0400,  0.0053,  0.0017],
        [ 0.1406, -0.0263,  0.0611, -0.0110,  0.0010],
        [ 0.0654, -0.0106, -0.0045,  0.0343, -0.0030]])

MyTorch dLdb:
 [ 0.17004118 -0.02935796  0.06321571 -0.02123358  0.01974512]

PyTorch dLdb:
 tensor([ 0.1700, -0.0294,  0.0632, -0.0212,  0.0197])

Difference in dLdW: 1.1085912244326638e-07
Difference in dLdb: 6.319506077808191e-08
```

Compare a single optimization step

In [ ]:

```
# my SGD step
my_optimizer = mytorch.optim.SGD(my_net, lr=0.1)
my_optimizer.step()
my_Wk = my_net.W
my_bk = my_net.b

# torch SGD step
optimizer.zero_grad()
torch_loss.backward(retain_graph=True)
optimizer.step()
torch_Wk = net.weight.data
torch_bk = net.bias.data

print('MyTorch Wk:\n', my_Wk, '\n')
print('PyTorch Wk:\n', torch_Wk.T, '\n')
print('MyTorch bk:\n', my_bk, '\n')
print('PyTorch bk:\n', torch_bk)
```

```
print('Difference in Wk:', np.linalg.norm(my_Wk - torch_Wk.data.numpy().T))
print('Difference in bk:', np.linalg.norm(my_bk.flatten() - torch_bk.data.numpy()))
```

```
MyTorch Wk:
 [[0.00225624 0.65454369 0.22786553 0.20375799 0.41525752]
 [0.06995093 0.09965917 0.54460619 0.32926168 0.65353759]
 [0.69258879 0.6152178  0.48305614 0.45607477 0.16900205]
 [0.05864521 0.35032098 0.24489474 0.8639804  0.83109932]
 [0.7373715  0.07190249 0.15495887 0.52788239 0.4534458 ]
 [0.14922424 0.40519783 0.92280482 0.5548514  0.28792973]
 [0.09215936 0.45045506 0.78606601 0.93621087 0.64919466]
 [0.28578957 0.23344379 0.51530289 0.79663504 0.54019491]
 [0.22328781 0.61469977 0.93714608 0.01456096 0.10773379]
 [0.06160243 0.89891648 0.83156413 0.43773651 0.74866049]]

PyTorch Wk:
 tensor([[0.0023, 0.6545, 0.2279, 0.2038, 0.4153],
        [0.0700, 0.0997, 0.5446, 0.3293, 0.6535],
        [0.6926, 0.6152, 0.4831, 0.4561, 0.1690],
        [0.0586, 0.3503, 0.2449, 0.8640, 0.8311],
        [0.7374, 0.0719, 0.1550, 0.5279, 0.4534],
        [0.1492, 0.4052, 0.9228, 0.5549, 0.2879],
        [0.0922, 0.4505, 0.7861, 0.9362, 0.6492],
        [0.2858, 0.2334, 0.5153, 0.7966, 0.5402],
        [0.2233, 0.6147, 0.9371, 0.0146, 0.1077],
        [0.0616, 0.8989, 0.8316, 0.4377, 0.7487]])

MyTorch bk:
 [[0.61248868 0.96033306 0.09868591 0.80567634 0.61423769]]

PyTorch bk:
 tensor([0.6125, 0.9603, 0.0987, 0.8057, 0.6142])
Difference in Wk: 1.037580056856067e-07
Difference in bk: 2.613949504884645e-08
```

In [ ]: