Computer Vision_HW4

学号：2101212808 | 姓名：郭静文

目录： 题目 | 解答 | PyTorch 验证

题目

目标函数：$f = ||\max(XW, 0) - Y||_F^2$

手动写出以下表达式，并用 PyTorch 进行验证：$\frac{\partial f}{\partial W}$、$\frac{\partial f}{\partial X}$、$\frac{\partial f}{\partial Y}$

解答

令 $Z = \max(XW, 0)$，则目标函数可化简为 $f = ||Z - Y||_F^2$。

利用矩阵的 F 范数的定义：

$$||A||_F = \sqrt{(\sum_{i=1}^{m}\sum_{j=1}^{n}|a_{i,j}|^2)} = \sqrt{(\mathrm{tr}\,(A^TA))}$$

将矩阵的 F 范数平方改写为矩阵与自身的内积的迹：

$$\begin{aligned} f &= ||Z - Y||_F^2 \\ &= tr((Z-Y)^T(Z-Y)) \\ &= tr(Z^TZ - Z^TY - Y^TZ + Y^TY) \end{aligned}$$

参考 The Matrix Cookbook [Page12-13] 给出的矩阵的迹及其微分定理公式：

$$\frac{\partial}{\partial \mathbf{X}}\mathrm{Tr}(\mathbf{XA}) = \mathbf{A}^T \quad (100)$$

$$\frac{\partial}{\partial \mathbf{X}}\mathrm{Tr}(\mathbf{X}^T\mathbf{BX}) = \mathbf{BX} + \mathbf{B}^T\mathbf{X} \quad (108)$$

$$\frac{\partial}{\partial \mathbf{X}}\mathrm{Tr}(\mathbf{X}^T\mathbf{XB}) = \mathbf{XB}^T + \mathbf{XB} \quad (113)$$

函数 f 分别对 Z、Y 求偏导，可得 $\frac{\partial f}{\partial Z}$、$\frac{\partial f}{\partial Y}$：

$$\frac{\partial f}{\partial Z} = Z + Z - Y - Y = 2(Z - Y)$$

$$\frac{\partial f}{\partial Y} = -2Z + 2Y = 2(Y - Z)$$

根据矩阵导数与微分的关系：

$$df = \mathrm{tr}\,(\frac{\partial f}{\partial Z}^T dZ)$$

进一步地，将 $\frac{\partial f}{\partial Z}$ 代入，微分 df 可以化简为：

$$\begin{aligned} df &= tr\{2(Z-Y)^T dZ\} \\ &= tr\{2(Z-Y)^T d(\max(XW,0))\} \\ &= tr\{2(Z-Y)^T \max{}'(XW,0) \odot d(XW)\} \end{aligned}$$

将 df 用 dW 表示并代入：

$$\begin{aligned} df &= tr\{2(Z-Y)^T \max{}'(XW,0) \odot XdW\} \\ &= tr\{2[(Z-Y) \odot \max{}'(XW,0)]^T XdW\} \end{aligned}$$

从而推出 $\frac{\partial f}{\partial W}$：

$$\frac{\partial f}{\partial W} = 2X^T[(Z-Y) \odot \max{}'(XW,0)]$$

将 df 用 dX 表示并代入：

$$\begin{aligned} df &= tr\{2(Z-Y)^T \max{}'(XW,0) \odot (dX)W\} \\ &= tr\{2W(Z-Y)^T \max{}'(XW,0) \odot dX\} \\ &= tr\{2W[(Z-Y) \odot \max{}'(XW,0)]^T dX\} \end{aligned}$$

从而推出 $\frac{\partial f}{\partial X}$：

$$\frac{\partial f}{\partial X} = 2(Z-Y) \odot \max{}'(XW,0)W^T$$

综上：

$$\begin{aligned} \frac{\partial f}{\partial W} &= 2X^T[(\max(XW,0)-Y) \odot \max{}'(XW,0)] \\ \frac{\partial f}{\partial X} &= 2(\max(XW,0)-Y) \odot \max{}'(XW,0)W^T \\ \frac{\partial f}{\partial Y} &= 2(Y-\max(XW,0)) \end{aligned}$$

PyTorch 验证

构造数据：

```
X = torch.randn(10, 4, requires_grad = True)
W = torch.randn(4, 4, requires_grad = True)
Y = torch.randn(10, 4, requires_grad = True)
print("X = ", X)
print("W = ", W)
print("Y = ", Y)
```

```
X =  tensor([[-1.1258, -1.1524, -0.2506, -0.4339],
        [ 0.8487,  0.6920, -0.3160, -2.1152],
        [ 0.3223, -1.2633,  0.3500,  0.3081],
        [ 0.1198,  1.2377,  1.1168, -0.2473],
        [-1.3527, -1.6959,  0.5667,  0.7935],
        [ 0.5988, -1.5551, -0.3414,  1.8530],
        [-0.2159, -0.7425,  0.5627,  0.2596],
        [-0.1740, -0.6787,  0.9383,  0.4889],
        [ 1.2032,  0.0845, -1.2001, -0.0048],
        [-0.5181, -0.3067, -1.5810,  1.7066]], requires_grad=True)
W =  tensor([[ 0.2055, -0.4503, -0.5731, -0.5554],
        [ 0.5943,  1.5419,  0.5073, -0.5910],
        [-1.3253,  0.1886, -0.0691, -0.4949],
        [-1.4959, -0.1938,  0.4455,  1.3253]], requires_grad=True)
Y =  tensor([[ 1.5091,  2.0820,  1.7067,  2.3804],
        [-1.1256, -0.3170, -1.0925, -0.0852],
        [ 0.3276, -0.7607, -1.5991,  0.0185],
        [-0.7504,  0.1854,  0.6211,  0.6382],
        [-0.0033, -0.5344,  1.1687,  0.3945],
        [ 1.9415,  0.7915, -0.0203, -0.4372],
        [-0.2188, -2.4351, -0.0729, -0.0340],
        [ 0.9625,  0.3492, -0.9215, -0.0562],
        [-0.6227, -0.4637,  1.9218, -0.4025],
        [ 0.1239,  1.1648,  0.9234,  1.3873]], requires_grad=True)
```

计算函数 f：

```
A = torch.max(torch.mm(X, W), torch.zeros_like(X)) - Y
f = torch.trace(torch.mm(A.t(), A))
print("f = ", f)
```

```
f =  tensor(99.9048, grad_fn=<TraceBackward>)
```

计算梯度：

```
f.backward()
print("W = ", W.grad)
print("X = ", X.grad)
print("Y = ", Y.grad)
```

```
W =  tensor([[ 18.2980,   2.7573,   2.3914,  -0.1974],
        [ 11.0817,   6.6428,   2.5163, -20.3225],
        [ -8.6662,   3.4506,  -1.8979,  -3.3608],
        [-21.1681,  -6.6739,  -1.0693,  27.0278]])
X =  tensor([[  1.1002,   0.0860,   5.3377,   0.2788],
        [  0.9583,  10.4633, -13.5234, -16.3639],
        [ -0.8712,  -0.9272,  -0.7764,   2.0790],
        [ -1.4504,   5.6914,   0.7613,  -0.9693],
        [ -1.2892,  -3.4714,  -1.9788,   4.8091],
        [ -4.0523,  -4.3127,  -3.6114,   9.6703],
        [ -0.7312,  -0.7782,  -0.6516,   1.7449],
        [ -0.8191,  -0.8718,  -0.7300,   1.9547],
        [  1.0350,   2.9930,  -6.6743,  -7.5333],
        [ -2.4616,  -2.4243,  -2.1164,   5.7128]])
```

```
Y =  tensor([[ 2.8885e+00,  4.1639e+00,  3.4134e+00,  3.0501e+00],
        [-1.0589e+01, -2.7045e+00, -2.1849e+00, -1.7039e-01],
        [ 6.5523e-01, -1.5214e+00, -3.1982e+00, -1.5687e+00],
        [-1.5009e+00, -3.8551e+00,  4.9843e-01,  1.2764e+00],
        [-6.6077e-03, -1.0689e+00,  1.8791e+00, -4.2604e+00],
        [ 3.8829e+00,  1.5830e+00, -4.0504e-02, -7.2968e+00],
        [-4.3767e-01, -4.8701e+00, -1.4583e-01, -1.3166e+00],
        [ 1.9250e+00,  6.9834e-01, -1.8429e+00, -1.4750e+00],
        [-5.0359e+00, -9.2744e-01,  3.8436e+00, -8.0509e-01],
        [ 2.4780e-01,  2.3296e+00, -1.7491e-01, -4.2519e+00]])
```

```
max_grad = torch.max(X.mm(W), torch.zeros_like(X.mm(W))) > 0
max_grad
```

```
tensor([[ True, False, False,  True],
        [ True,  True, False, False],
        [False, False, False,  True],
        [False,  True,  True, False],
        [False, False,  True,  True],
        [False, False, False,  True],
        [False, False, False,  True],
        [False, False, False,  True],
        [ True, False, False, False],
        [False, False,  True,  True]])
```

验证一：$\frac{\partial f}{\partial W} = 2X^T[(\max(XW, 0) - Y) \odot \max'(XW, 0)]$

```
W.grad == 2 * torch.mm(X.t(), (torch.max(X.mm(W), torch.zeros_like(X.mm(W))) - Y) * max_grad)
```

```
tensor([[True, True, True, True],
        [True, True, True, True],
        [True, True, True, True],
        [True, True, True, True]])
```

验证二：$\frac{\partial f}{\partial X} = 2(\max(XW, 0) - Y) \odot \max'(XW, 0)W^T$

```
X.grad == 2* ((torch.max(X.mm(W), torch.zeros_like(X.mm(W))) - Y) * max_grad).mm(W.t())
```

```
tensor([[True, True, True, True],
        [True, True, True, True],
        [True, True, True, True],
        [True, True, True, True],
        [True, True, True, True],
        [True, True, True, True],
        [True, True, True, True],
        [True, True, True, True],
        [True, True, True, True],
        [True, True, True, True]])
```

验证三：$\frac{\partial f}{\partial Y} = 2(Y - \max(XW, 0))$

```
Y.grad == 2 * (Y - torch.max(torch.mm(X, W), torch.zeros_like(X.mm(W))))
```

```
tensor([[True, True, True, True],
        [True, True, True, True],
        [True, True, True, True],
        [True, True, True, True],
        [True, True, True, True],
        [True, True, True, True],
        [True, True, True, True],
        [True, True, True, True],
        [True, True, True, True],
        [True, True, True, True]])
```

源码：gradient_calculation.ipynb