torch.optim中包含了SGD，Adagrad，RMSprop，Adam，LBFGS

**torch.optim.Adam(params,lr)**

params为模型参数

lr为学习率

返回一个优化器对象

**反向传播优化**

设optimizer是optim.Adam的实例，则可用optimizer.zero\_grad()清空梯度，用optimizer.step()更新参数

for input, target in dataset:

optimizer.zero\_grad()*#清空梯度*output = model(input)  
loss = loss\_fn(output, target)

loss.backward()

optimizer.step()*#更新参数*

**torch.optim.LBFGS(params)**

设optimizer是optim.LBFGS的实例，则可用optimizer.zero\_grad()清空梯度，用optimizer.step(closure)更新参数