



在训练中显示损失值



显示损失值

开始训练, 轮数为 epoch, 采用SGD随机梯度下降优化方法

step = 0 # 记录训练步数

loss_list = [] # 用于保存loss值的列表

```
for epoch in range(train_epochs):
```

```
    for xs,ys in zip(x_data, y_data):
```

```
        _, loss=sess.run([optimizer,loss_function], feed_dict={x: xs, y: ys})
```

```
        # 显示损失值 loss
```

```
        # display_step: 控制报告的粒度
```

```
        # 例如, 如果 display_step 设为 2 , 则将每训练2个样本输出一次损失值
```

```
        # 与超参数不同, 修改 display_step 不会更改模型所学习的规律
```

```
        loss_list.append(loss)
```

```
        step=step+1
```

```
        if step % display_step == 0:
```

```
            print("Train Epoch:", '%02d' % (epoch+1), "Step: %03d" % (step), "loss=", \
                  "{:.9f}".format(loss))
```

```
b0temp=b.eval(session=sess)
```

```
w0temp=w.eval(session=sess)
```

```
plt.plot(x_data, w0temp * x_data + b0temp )# 画图
```

控制显示loss值的粒度
display_step = 10

增加一个控制显示粒度的参数



显示损失值



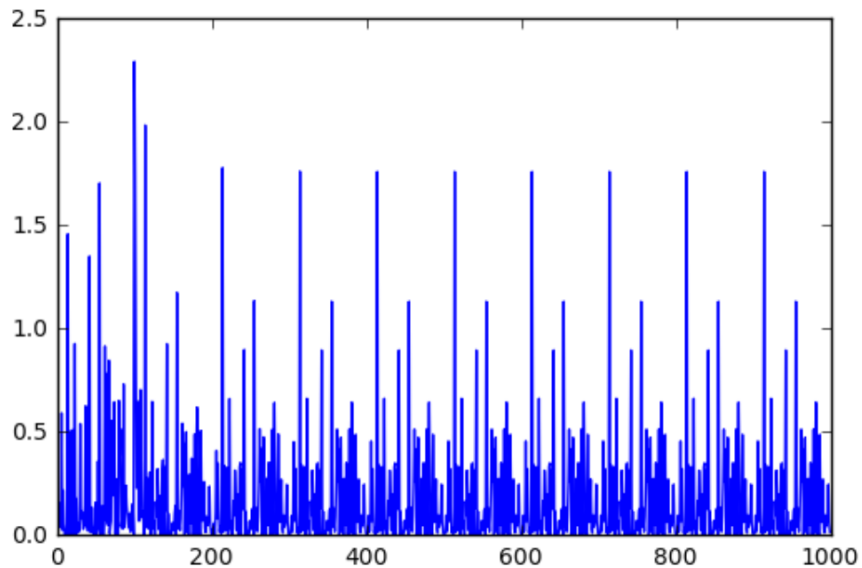
```
Train Epoch: 01 Step: 010 loss= 0.056705344
Train Epoch: 01 Step: 020 loss= 0.000246272
Train Epoch: 01 Step: 030 loss= 0.019463092
Train Epoch: 01 Step: 040 loss= 0.589754999
Train Epoch: 01 Step: 050 loss= 0.000975099
Train Epoch: 01 Step: 060 loss= 0.142678708
Train Epoch: 01 Step: 070 loss= 0.046167519
Train Epoch: 01 Step: 080 loss= 0.008698015
Train Epoch: 01 Step: 090 loss= 0.241310671
Train Epoch: 01 Step: 100 loss= 0.000510003
Train Epoch: 02 Step: 110 loss= 0.317787021
Train Epoch: 02 Step: 120 loss= 0.032408509
Train Epoch: 02 Step: 130 loss= 0.093365036
Train Epoch: 02 Step: 140 loss= 0.332121730
Train Epoch: 02 Step: 150 loss= 0.060511891
Train Epoch: 02 Step: 160 loss= 0.024091592
Train Epoch: 02 Step: 170 loss= 0.178776026
Train Epoch: 02 Step: 180 loss= 0.006450176
```



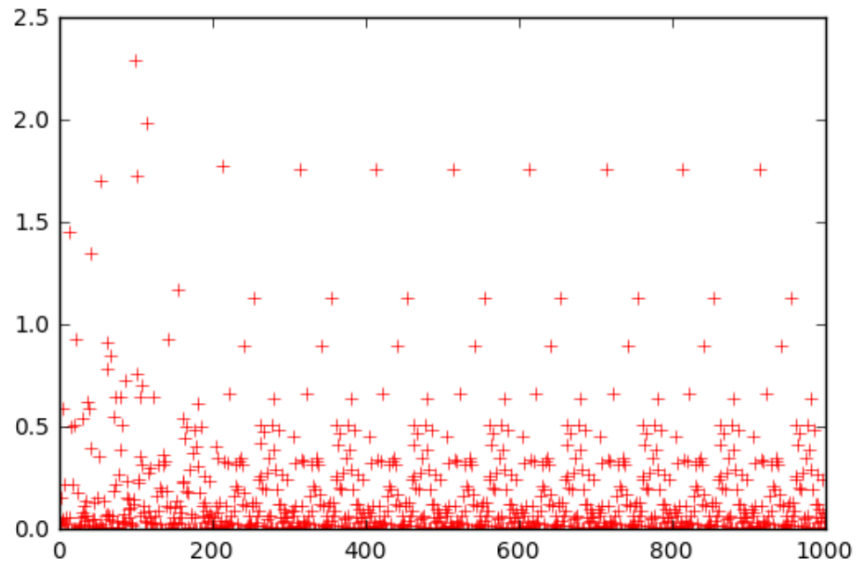
图形化显示损失值



```
plt.plot(loss_list)
```



```
plt.plot(loss_list, 'r+')
```





随机梯度下降



在梯度下降法中，**批量**指的是用于在单次迭代中计算梯度的样本总数

假定批量是指整个数据集，数据集通常包含很大样本（数万甚至数千亿），此外，数据集通常包含多个特征。因此，一个批量可能相当巨大。如果是超大批量，则单次迭代就可能要花费很长时间进行计算

随机梯度下降法（SGD） 每次迭代只使用一个样本（批量大小为 1），如果进行足够的迭代，SGD 也可以发挥作用。“随机”这一术语表示构成各个批量的一个样本都是随机选择的

小批量随机梯度下降法（小批量 SGD） 是介于全批量迭代与 SGD 之间的折衷方案。小批量通常包含 10-1000 个随机选择的样本。小批量 SGD 可以减少 SGD 中的杂乱样本数量，但仍然比全批量更高效