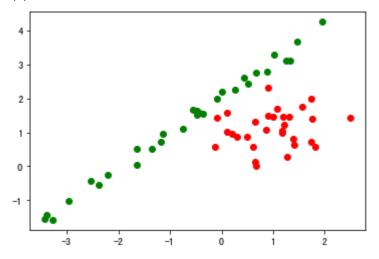
In []:			

In [4]:

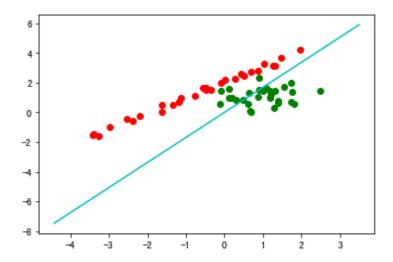
```
import sklearn
import numpy as np
def create data():
  """# 为了是数值取值为线性可分为 random_state 取值为 3 8 18 (后面的没有测试)
    18 最能体现划分过程
  from sklearn.datasets import make classification
  data, target = make_classification(n_samples=60, # 样本个数
                      n features=2, #特征个数
                      n_informative=2, # 有效特征个数
                      n_redundant=0, # 冗余特征个数(有效特征的随机组合)
                      n_repeated=0, # 重复特征个数(有效特征和冗余特征的随机组合)
                      n_classes=2, # 样本类别
                      n_clusters_per_class=1, # 簇的个数
                      random state=18)
  return data, target
#数据的可视化
def show_data(data, target):
  import matplotlib.pyplot as plt
  data label = list(set(target))
  plt.figure()
  for i in range(len(target)):
    if (target[i] == data_label[0]):
      plt.scatter(data[i][0], data[i][1], c='r')
    else:
      plt.scatter(data[i][0], data[i][1], c='g')
  plt.show()
data, target = create_data()
print("生成的数据如下:")
show data(data, target)
for i in range(len(target)):
  if target[i] == 0:
    target[i] = -1
# 分类效果展示
def show_data_line(data, target, w, b):
  import matplotlib.pyplot as plt
  import numpy as np
  data_label = list(set(target))
  plt.figure()
  for i in range(len(target)):
    if (target[i] == data_label[0]):
      plt.scatter(data[i][0], data[i][1], c='r')
    else:
      plt.scatter(data[i][0], data[i][1], c='q')
  # 为了更客观的看出图像的变化,这里我们先找出原来的图像的坐标轴的横纵坐标的最大,最小
  \# x_{min,x_{max,y_{min,y_{max}}} = 0.0,0.0,0.0,0.0
  [x_min, y_min] = np.min(data, axis=0)
  [x_max, y_max] = np.max(data, axis=0)
  XX = np.linspace(x_min - 1, x_max + 1)
```

```
YY = -(w[0] / w[1]) * XX - b / w[1]
  plt.plot(XX, YY, '-c', label='Hyperplane')
  plt.show()
# 优化目标: MIN L(w,b) = (-1)*yi * (w * xi + b) 最小
def loss_fun(data, target, w, b):
  loss = 0.0
  #划分错误的是
  loss num = 0
  for i in range(len(target)):
    loss += (-1) * target[i] * (np.dot(w, data[i].T) + b)
    if target[i] * (np.dot(w, data[i].T) + b) <= 0:
       loss num+=1
  return loss, loss num
#初始化w,b
w = np.array((0.0, 0.0))
b = 0.0
# 所有数据训练10轮
num_epochs = 1
#每10条数据为一个周期
epoch = 20
# 学习率 0.01
Ir = 0.01
while (num_epochs <= 40):
  # 随机打乱数据
  data, target = sklearn.utils.shuffle(data, target)
  # 记录已经训练过的数据的条数
  tarin data num = 0
  for i in range(len(target)):
    # 分类错误, 感知机
    if target[i] * (np.dot(w, data[i].T) + b) <= 0:
       w = w + lr * target[i] * data[i]
       b = b + lr * target[i]
    if ((i + 1) \% \text{ epoch} == 0):
       loss,loss_num = loss_fun(data, target, w, b)
       print("第{}轮训练, after第{}条数据 w:{},b:{}\n损失函数值为{} 划分错误的个数是{}: ".format(num_e
       show_data_line(data, target, w, b)
  num_epochs += 1
```

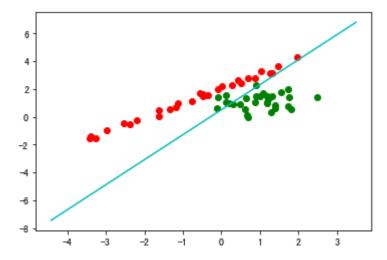
生成的数据如下:



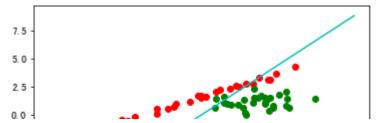
第1轮训练, after第20条数据 w:[-0.02686479 0.01588102],b:0.0 损失函数值为-1.4456892304288318 划分错误的个数是9:

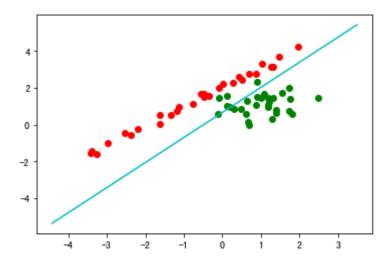


第1轮训练, after第40条数据 w:[-0.03554073 0.01975376],b:-0.01 损失函数值为-1.9045821199544153 划分错误的个数是6:

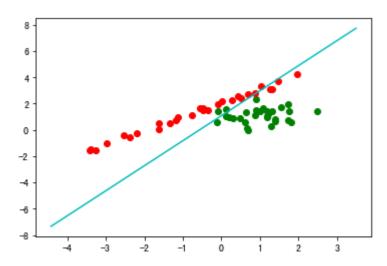


第1轮训练, after第60条数据 w:[-0.04444645 0.01994383],b:-0.02 损失函数值为-2.351550932022124 划分错误的个数是7:

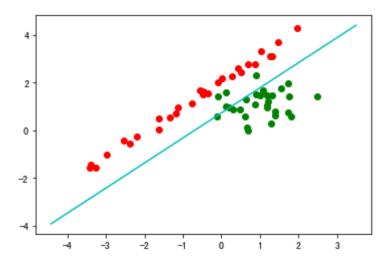




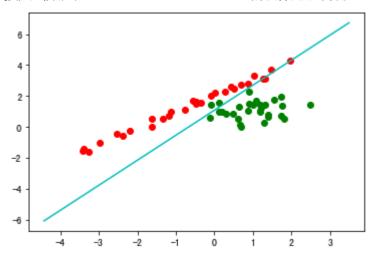
第2轮训练, after第40条数据 w:[-0.05284755 0.02775609],b:-0.03 损失函数值为-2.8217471879396405 划分错误的个数是6:



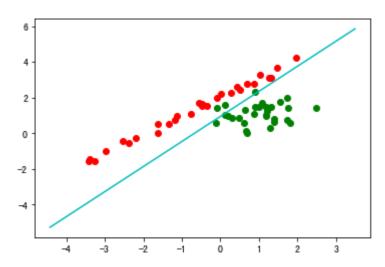
第2轮训练, after第60条数据 w:[-0.04315786 0.04117328],b:-0.03 损失函数值为-2.4220964880047955 划分错误的个数是5:



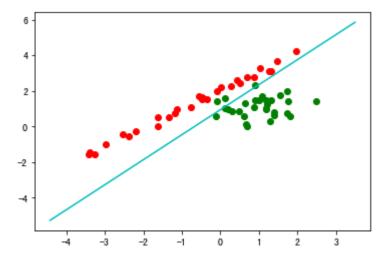
第3轮训练,after第20条数据 w:[-0.05980674 0.03697855],b:-0.04 损失函数值为-3.228742131305406 划分错误的个数是3:



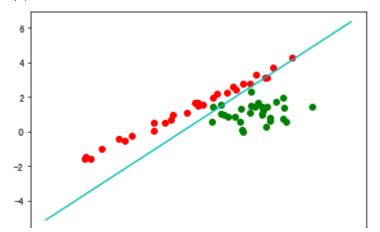
第3轮训练, after第40条数据 w:[-0.0597578 0.04251718],b:-0.04 损失函数值为-3.261524674211749 划分错误的个数是3:



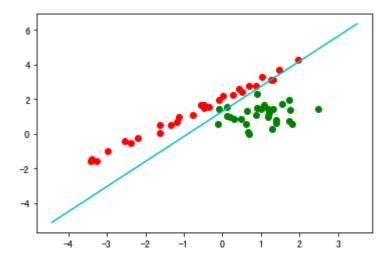
第3轮训练,after第60条数据 w:[-0.0597578 0.04251718],b:-0.04 损失函数值为-3.261524674211749 划分错误的个数是3:



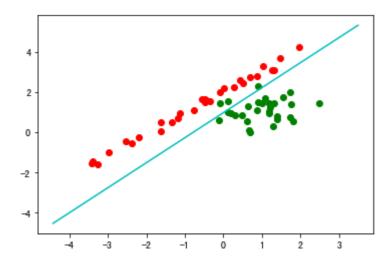
第4轮训练,after第20条数据 w:[-0.05563488 0.03845421],b:-0.05 损失函数值为-3.029314271229125 划分错误的个数是3:



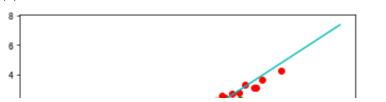
第4轮训练,after第40条数据 w:[-0.05563488 0.03845421],b:-0.05 损失函数值为-3.029314271229125 划分错误的个数是3:



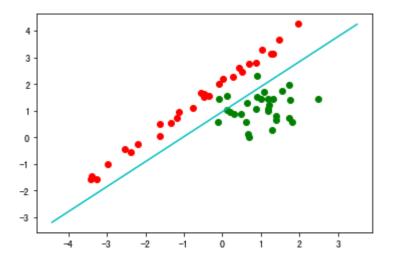
第4轮训练, after第60条数据 w:[-0.06363977 0.05112742],b:-0.05 损失函数值为-3.5106005870422243 划分错误的个数是3:



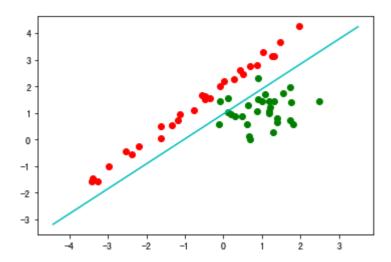
第5轮训练,after第20条数据 w:[-0.06891187 0.04086252],b:-0.06000000000000005 损失函数值为-3.709189101685324 划分错误的个数是6:



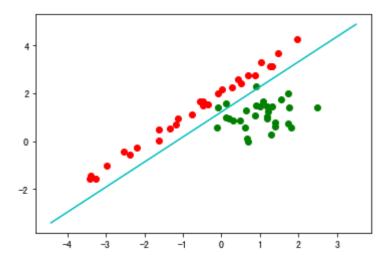
第5轮训练,after第40条数据 w:[-0.05847094 0.0623513],b:-0.06000000000000005 损失函数值为-3.3232810857065225 划分错误的个数是3:



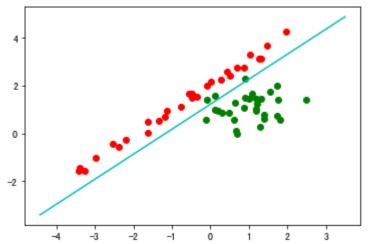
第5轮训练,after第60条数据 w:[-0.05847094 0.0623513],b:-0.060000000000000005 损失函数值为-3.3232810857065225 划分错误的个数是3:



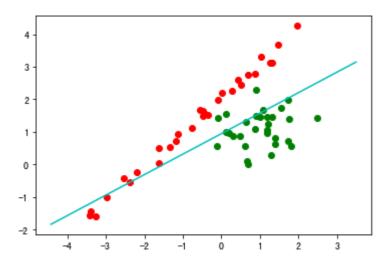
第6轮训练, after第20条数据 w:[-0.05969749 0.0569441],b:-0.07 损失函数值为-3.3502777897204115 划分错误的个数是3:



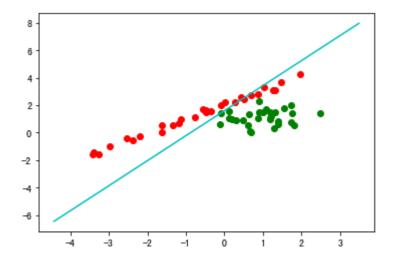
第6轮训练, after第40条数据 w:[-0.05969749 0.0569441],b:-0.07



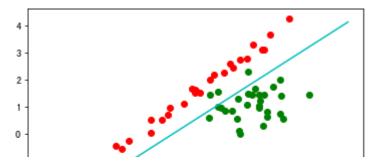
第6轮训练, after第60条数据 w:[-0.04635548 0.07374683],b:-0.07 损失函数值为-2.7893530105045667 划分错误的个数是8:



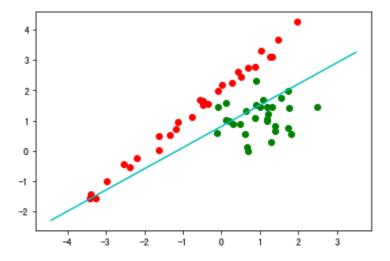
第7轮训练, after第20条数据 w:[-0.07975348 0.04377416],b:-0.07 损失函数值为-4.270367452414449 划分错误的个数是8:



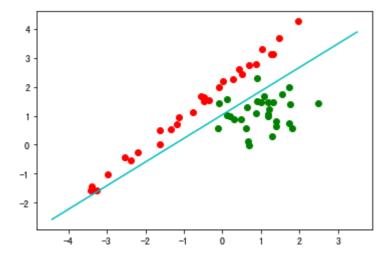
第7轮训练, after第40条数据 w:[-0.06668917 0.07324851],b:-0.07 损失函数值为-3.8039483162271783 划分错误的个数是3:



第7轮训练, after第60条数据 w:[-0.06091316 0.08693725],b:-0.07 损失函数值为-3.601917183903004 划分错误的个数是9:



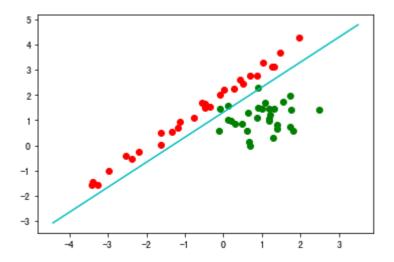
第8轮训练, after第20条数据 w:[-0.06297707 0.077247],b:-0.08 损失函数值为-3.643581035980519 划分错误的个数是3:

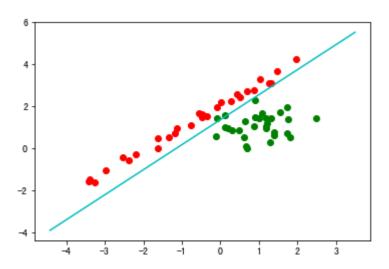


第8轮训练,after第40条数据 w:[-0.06406353 0.06710235],b:-0.09 损失函数值为-3.6334303925020923 划分错误的个数是3:

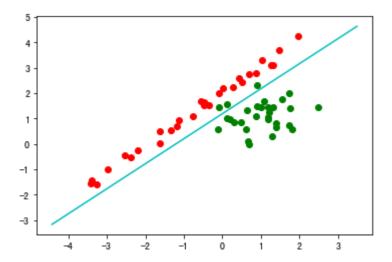
5 |

第8轮训练, after第60条数据 w:[-0.06793482 0.06844802],b:-0.09 损失函数值为-3.835760692424665 划分错误的个数是3:

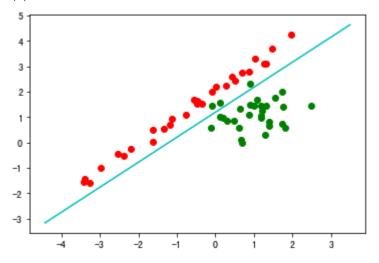




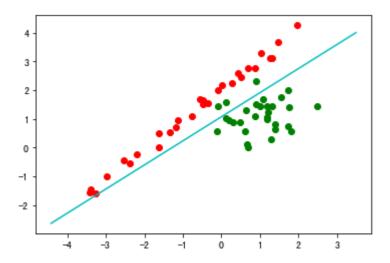
第9轮训练,after第40条数据 w:[-0.07440826 0.07561536],b:-0.09 损失函数值为-4.205369225278422 划分错误的个数是3:

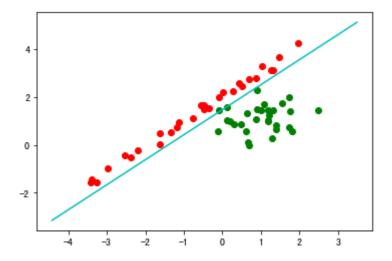


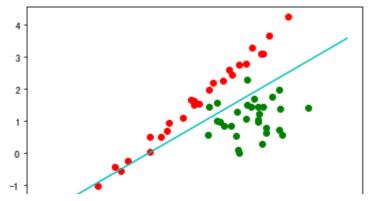
第9轮训练,after第60条数据 w:[-0.07440826 0.07561536],b:-0.09 损失函数值为-4.205369225278422 划分错误的个数是3:



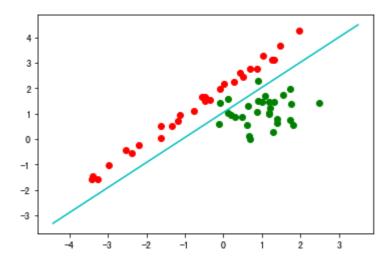
第10轮训练, after第20条数据 w:[-0.07013513 0.08374117],b:-0.09 损失函数值为-4.043175263236679 划分错误的个数是3:

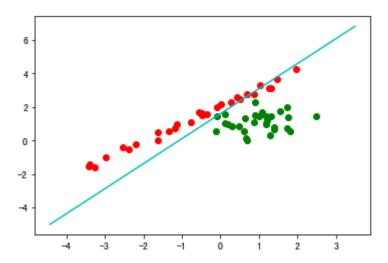


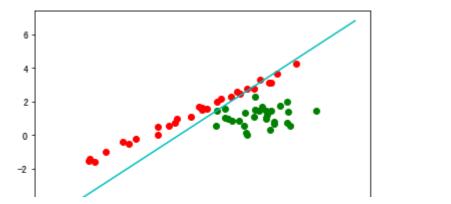


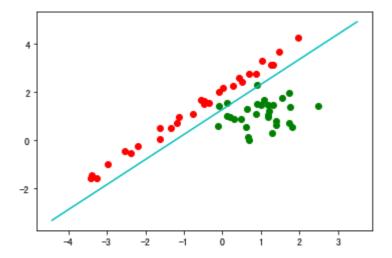


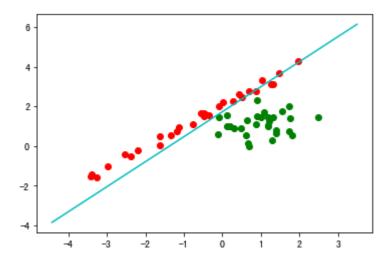
第11轮训练,after第20条数据 w:[-0.08441892 0.08566304],b:-0.09 损失函数值为-4.770349214131867 划分错误的个数是3:



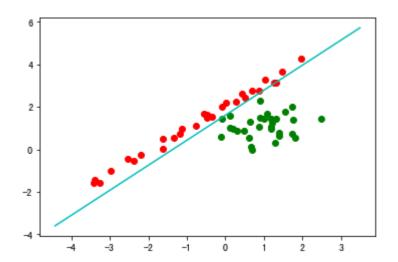


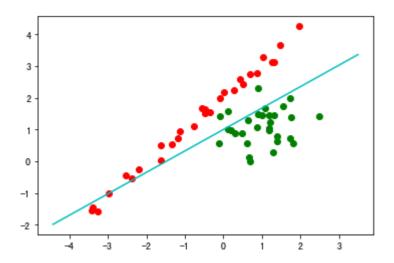


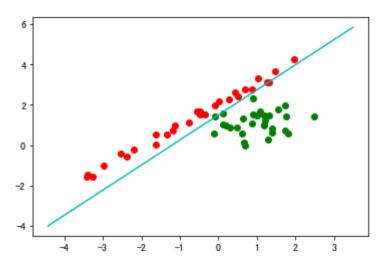


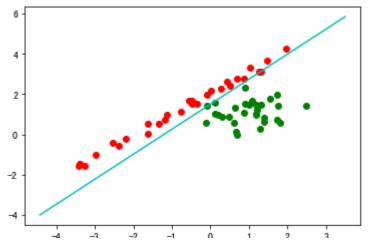


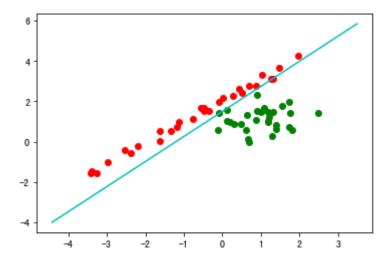


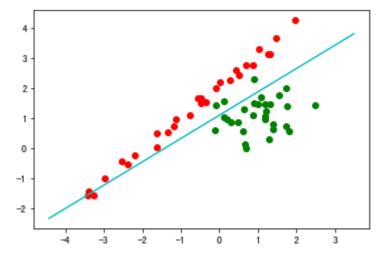


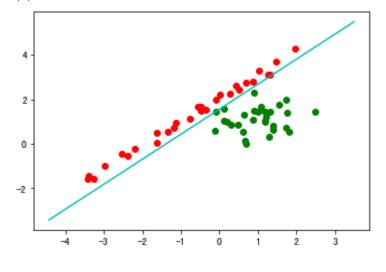


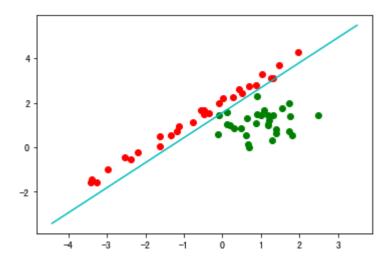


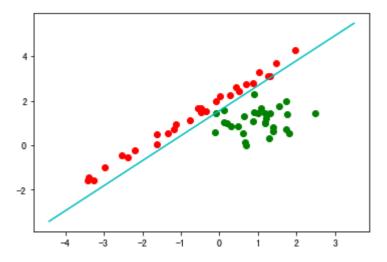


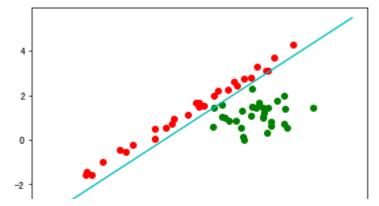


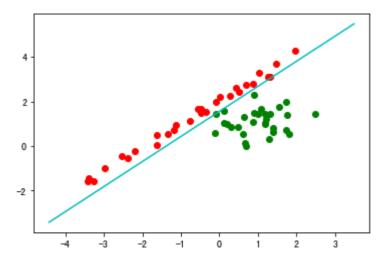


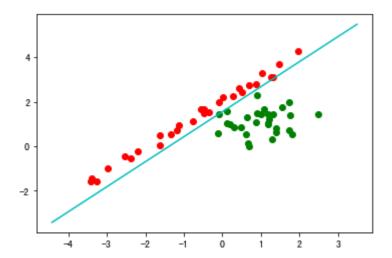




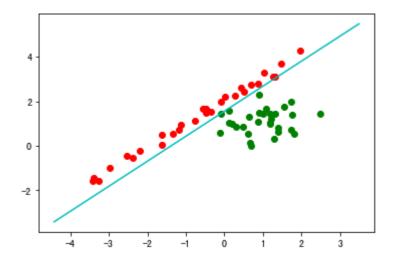


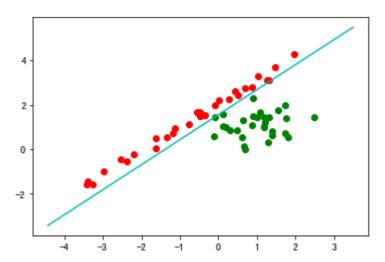


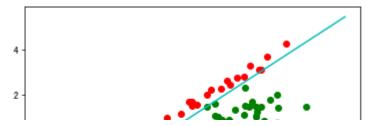


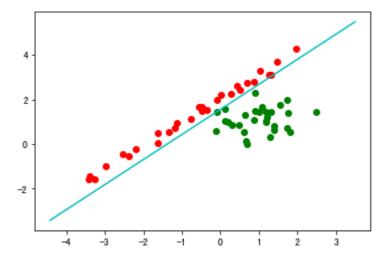


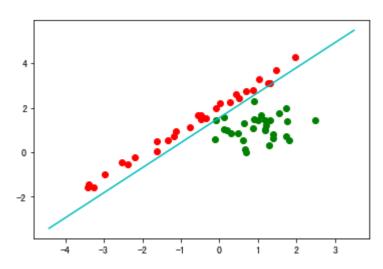


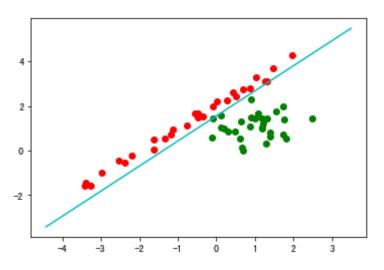


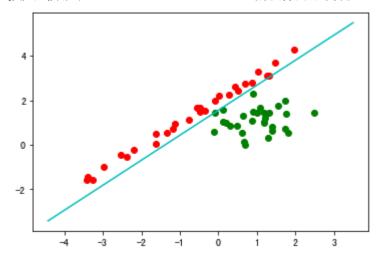


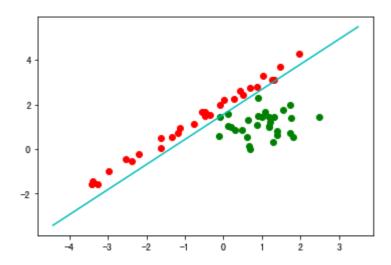


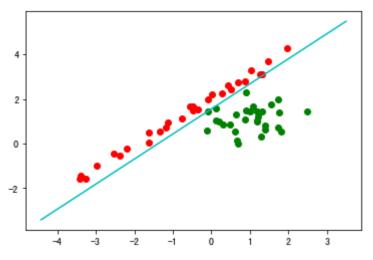


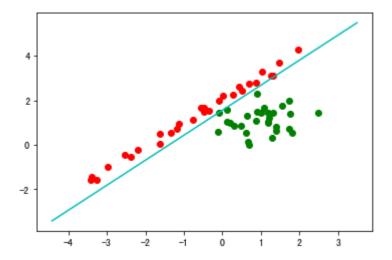


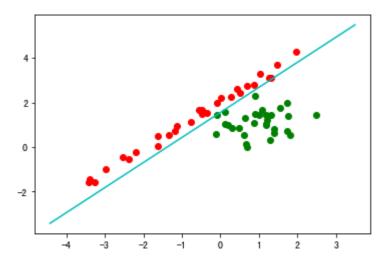


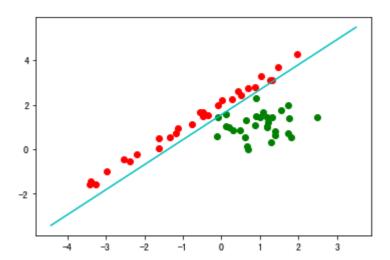


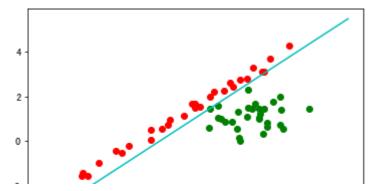


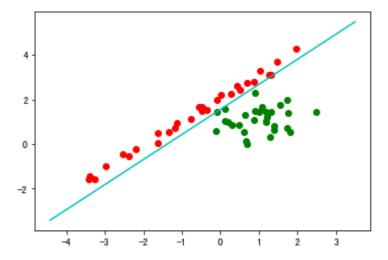


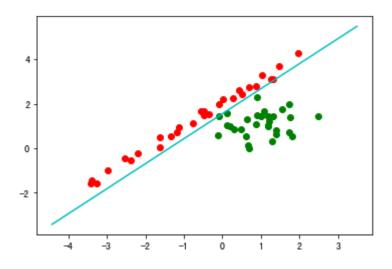


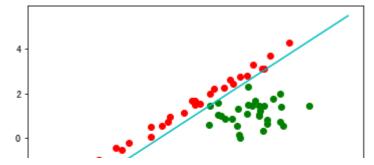


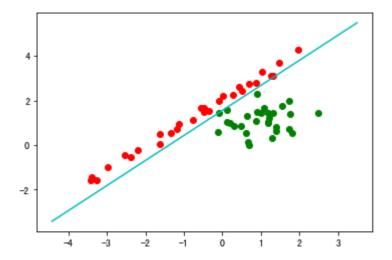


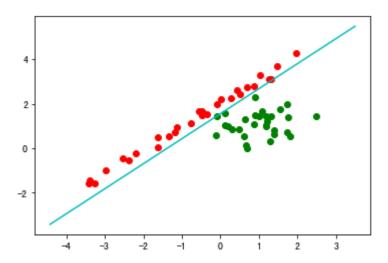






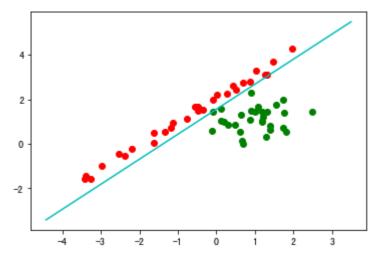


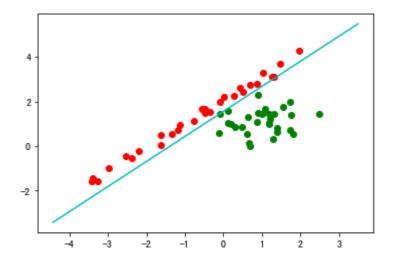


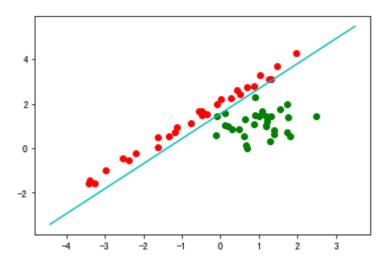


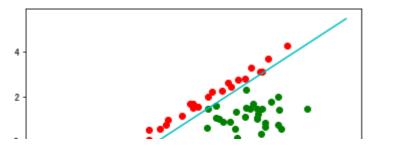
第23轮训练,after第40条数据 w:[-0.08684297 0.07723887],b:-0.119999999999999

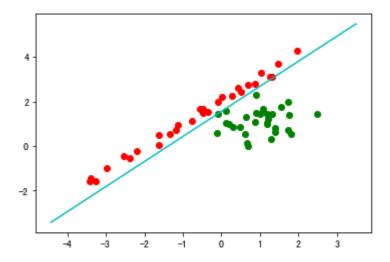
捐生函数值为-4 838092949669858 划分错误的个数型0·

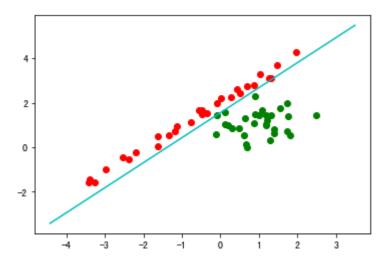


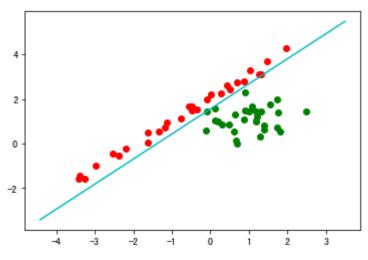


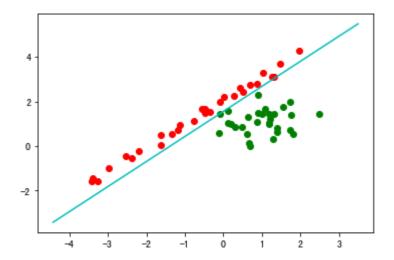


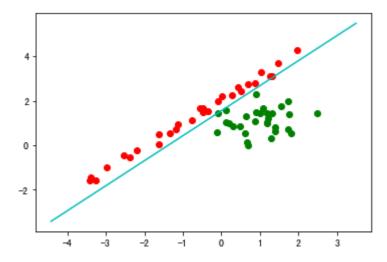




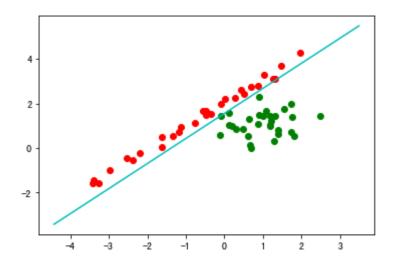


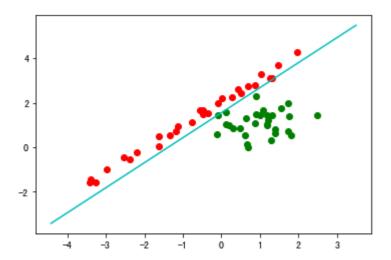


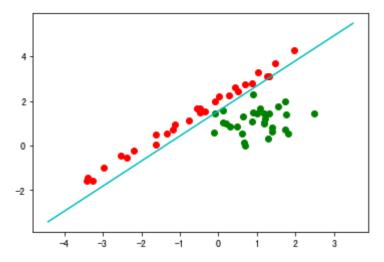


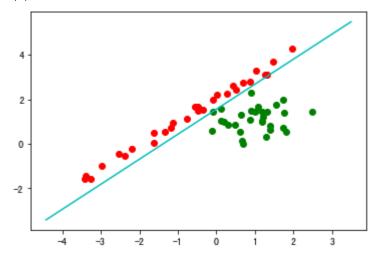


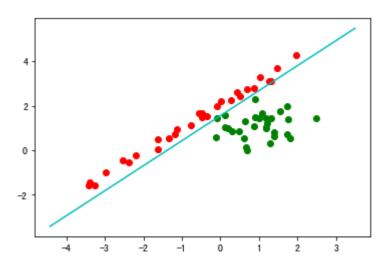


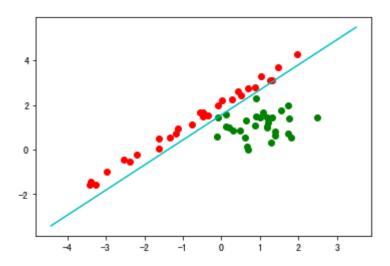




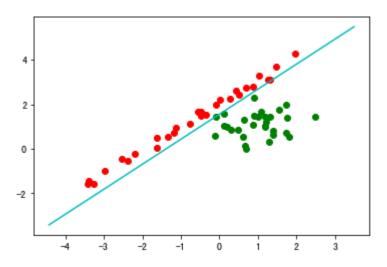


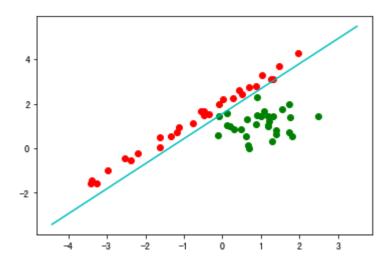


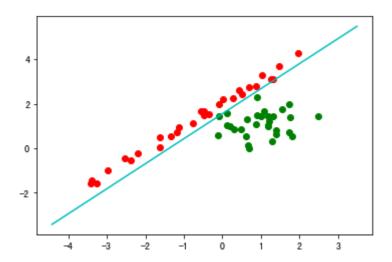


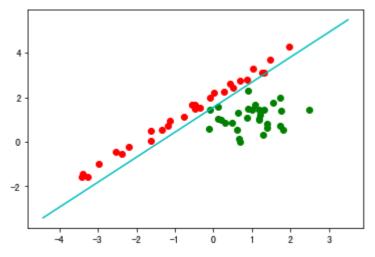


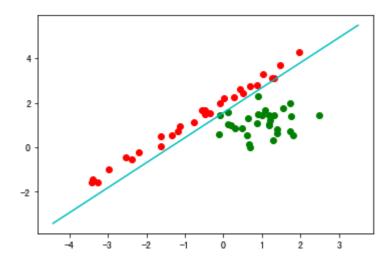
损失函数值为-4.838092949669856 划分错误的个数是0:

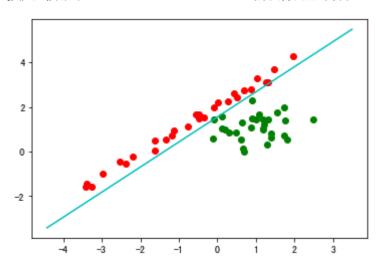












损失函数值为-4.838092949669856 划分错误的个数是0:

