In [7]:

```
import numpy as np
import scipy as sp
import matplotlib.pyplot as plt
import scipy.linalg as la
%matplotlib inline
```

data is from here: http://mlpy.sourceforge.net/docs/3.2/_downloads/iris1.csv (http://mlpy.sourceforge.net/docs/3.2/_downloads/iris1.csv)

dataset has 150 observations of three different flowers:

1. setosa
2. versicolor
3. virginica

the data taken is the sepal length and width, and the petal length and width

In [8]:

```
iris = np.loadtxt('iris.csv', delimiter=',')

X = iris[:,:-2]
target = np.int32(iris[:,-1])
```

X has the data `target` shows the exact description

Now make a *covariance* matrix and find the dominant eigenvectors:

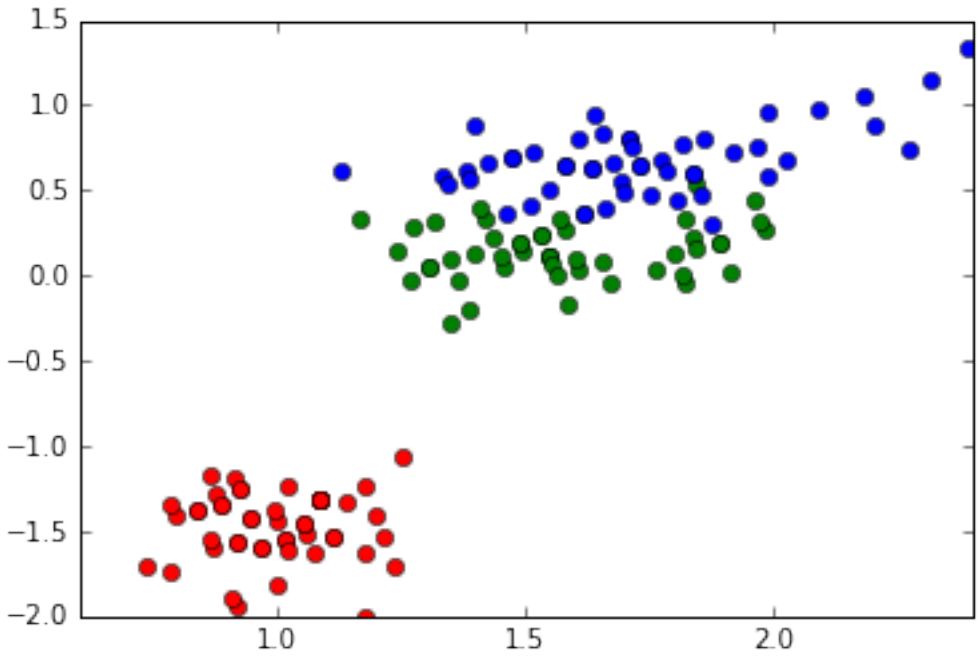
In [9]:

```
C = np.cov(X)
eigenvalues, eigenvectors = la.eigh(C)

I = np.argsort(np.abs(eigenvalues))

xi = I[-1]
yi = I[-2]
```

The dominant vectors are the *coordinates*

In [10]:

```
xcoord = eigenvectors[:,xi] * np.sqrt( eigenvalues[xi] )
ycoord = eigenvectors[:,yi] * np.sqrt( eigenvalues[yi] )

c = ['red', 'green', 'blue']
plt.figure()
plt.hold(True)
for i in range(iris.shape[0]):
    plt.plot(xcoord[i], ycoord[i], 'o', color=c[target[i]-1])
```



In [ ]:

```

```