

## 导入Iris鸢尾花数据集

In [5]:

```
import pandas as pd

#共150条记录，分别是50条山鸢尾（Iris-setosa）、50条变色鸢尾(Iris-versicolor)、50条维吉尼亚鸢尾(Iris-virginica)
url = "https://archive.ics.uci.edu/ml/machine-learning-databases/iris/iris.data"
names = ['花萼-length', '花萼-width', '花瓣-length', '花瓣-width', 'class']
dataset = pd.read_csv(url, names=names)
dataset.head()
```

Out[5]:

	花萼-length	花萼-width	花瓣-length	花瓣-width	class
0	5.1	3.5	1.4	0.2	Iris-setosa
1	4.9	3.0	1.4	0.2	Iris-setosa
2	4.7	3.2	1.3	0.2	Iris-setosa
3	4.6	3.1	1.5	0.2	Iris-setosa
4	5.0	3.6	1.4	0.2	Iris-setosa

## 预测花瓣长度与花瓣宽度的关系

In [6]:

```
X = dataset["花瓣-length"]
Y = dataset["花瓣-width"]
X = X.reshape(len(X),1)
Y = Y.reshape(len(Y),1)
```

c:\python 3.6.5\lib\site-packages\ipykernel\_launcher.py:3: FutureWarning: reshape is deprecated and will raise in a subsequent release. Please use .values.reshape (...) instead  
This is separate from the ipykernel package so we can avoid doing imports until  
c:\python 3.6.5\lib\site-packages\ipykernel\_launcher.py:4: FutureWarning: reshape is deprecated and will raise in a subsequent release. Please use .values.reshape (...) instead  
after removing the cwd from sys.path.

## 拆分数据集成训练集、测试集

In [7]:

```
from sklearn.model_selection import train_test_split
X_train, X_test, Y_train, Y_test = train_test_split(X, Y, test_size=0.2, random_state=0)
```

## 训练线性回归模型

In [8]:

```
from sklearn.linear_model import LinearRegression
regressor = LinearRegression()
regressor = regressor.fit(X_train, Y_train)
```

## 查看训练效果

In [9]:

```
import matplotlib.pyplot as plt

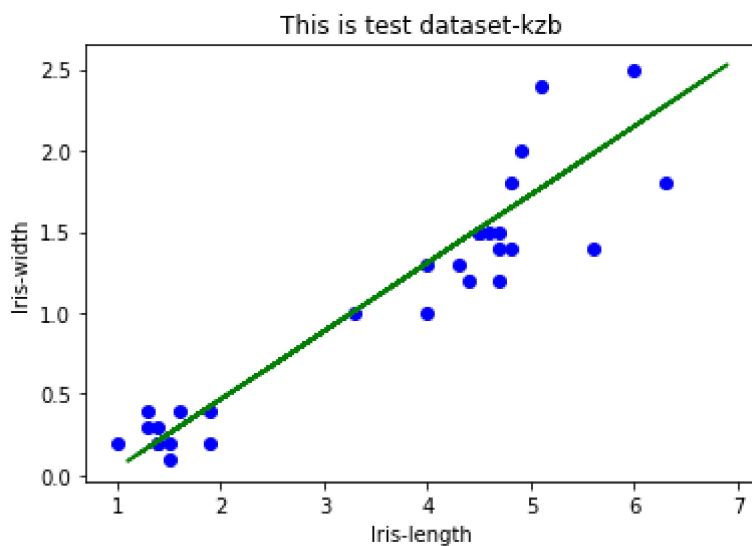
plt.scatter(X_train, Y_train, color='red')
plt.plot(X_train, regressor.predict(X_train), color='green')
plt.xlabel("Iris-length")
plt.ylabel("Iris-width")
plt.title("This is train dataset-kzb")
plt.show()
```

<Figure size 640x480 with 1 Axes>

## 用测试集数据检测训练效果

In [10]:

```
plt.scatter(X_test, Y_test, color='blue')
plt.plot(X_train, regressor.predict(X_train), color='green')
plt.xlabel("Iris-length")
plt.ylabel("Iris-width")
plt.title("This is test dataset-kzb")
plt.show()
```

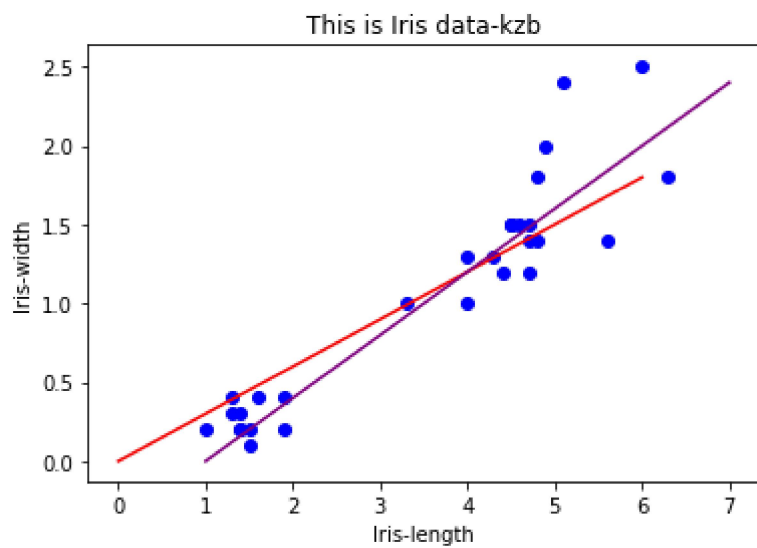


In [40]:

```
import numpy as np
a = np.linspace(0, 6, 300)

plt.scatter(X_test, Y_test, color='blue')
# plt.plot(a+1, 0.5*a, color='green')
plt.plot(a, 0.3*a, color='red')
plt.plot(a+1, 0.4*a, color='purple')

plt.xlabel("Iris-length")
plt.ylabel("Iris-width")
plt.title("This is Iris data-kzb")
plt.show()
```



In [ ]: