

kristiania_data_science_intro

October 11, 2022

```
[1]: import numpy as np
      from sklearn.linear_model import LinearRegression
      from sklearn.model_selection import train_test_split
      from matplotlib import pyplot as plt
```

```
[2]: n = 10
      x0 = np.arange(0, n)
      y0 = x0 + np.random.rand(n)

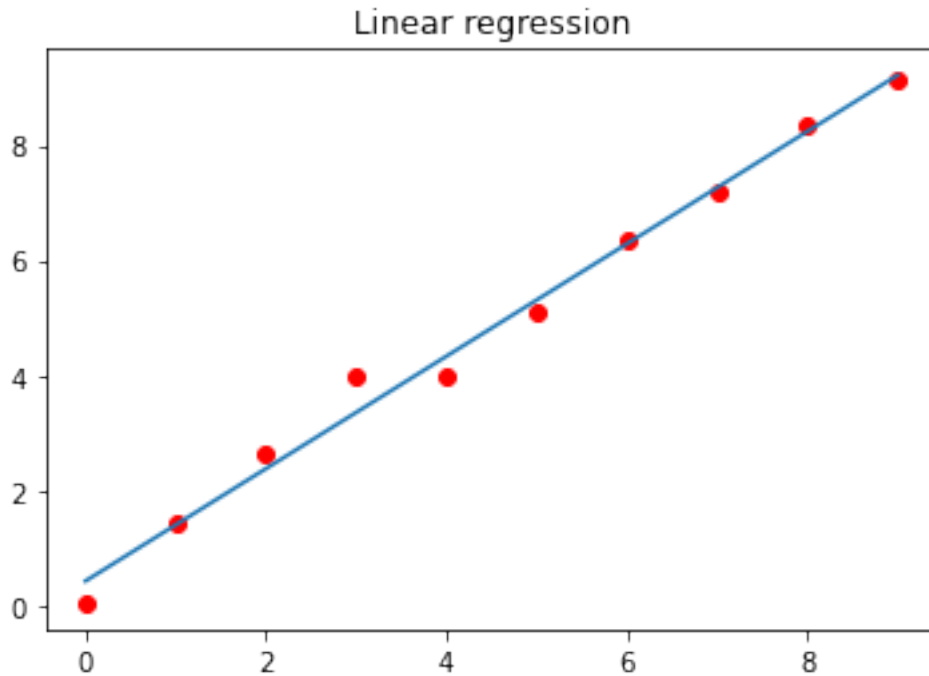
      x = np.linspace(0,n-1,1000)

      f2 = LinearRegression().fit(x0.reshape(-1, 1), y0)

      plt.figure()
      plt.plot(x0, y0, 'ro')
      plt.plot(x, f2.predict(x.reshape(-1, 1)))
      plt.title('Linear regression')

      print(f2.score(x0.reshape(-1, 1), y0))
```

0.9900119450930992

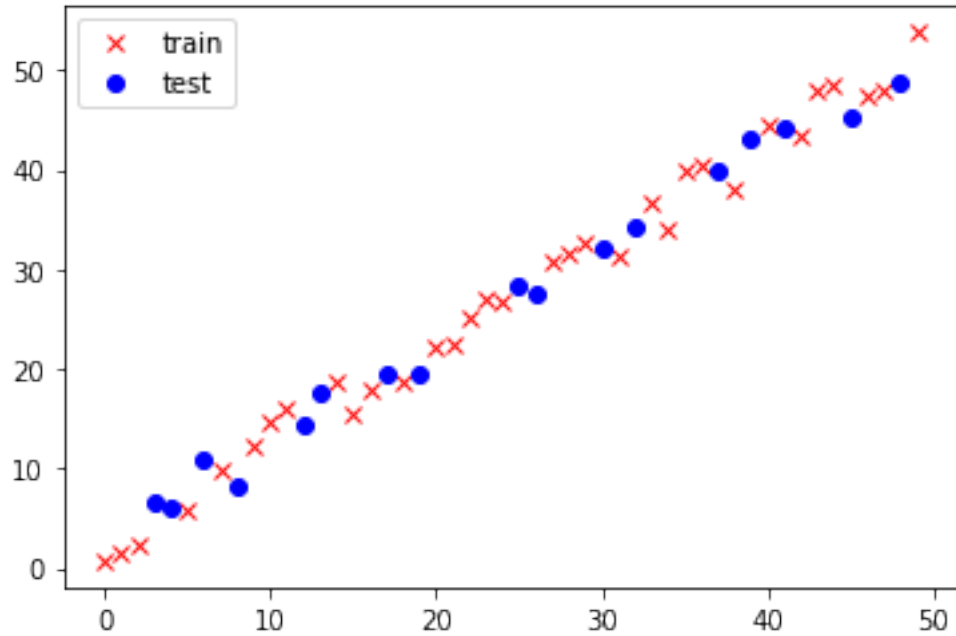


```
[3]: n = 50
x0 = np.arange(0, n)
y0 = x0 + n/10*np.random.rand(n)

x = np.linspace(0,n-1,1000)

x_train, x_test, y_train, y_test = train_test_split(x0, y0, test_size=0.33,
↳random_state=42)
plt.plot(x_train, y_train, 'rx', label='train')
plt.plot(x_test, y_test, 'bo', label='test')
plt.legend()
```

[3]: <matplotlib.legend.Legend at 0x1bd139a29d0>



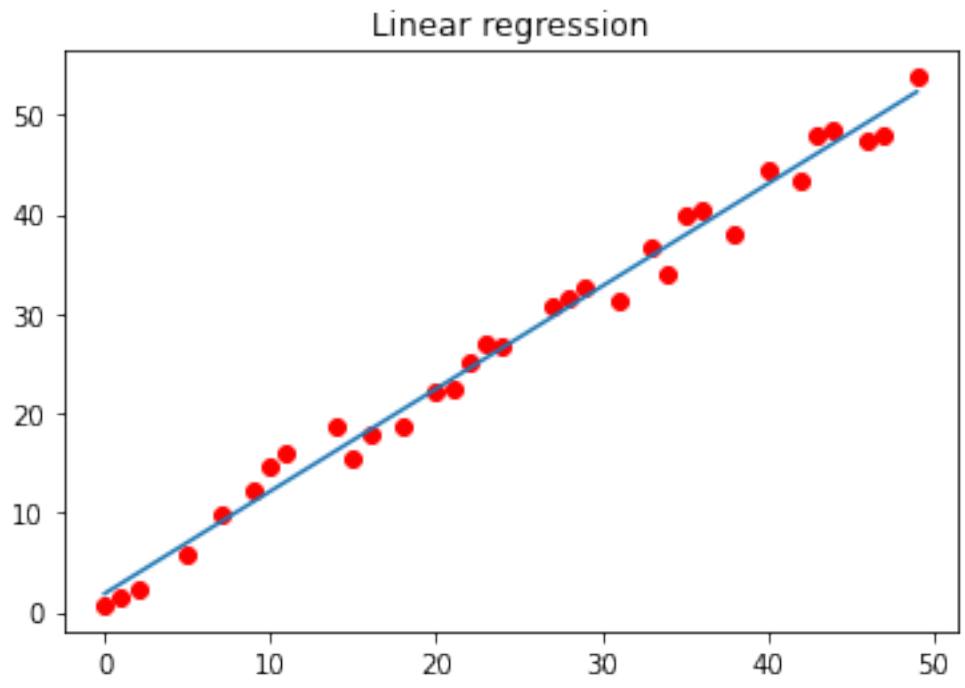
```
[4]: f2 = LinearRegression().fit(x_train.reshape(-1, 1), y_train)
```

```
plt.figure()
plt.plot(x_train, y_train, 'ro')
plt.plot(x, f2.predict(x.reshape(-1, 1)))
plt.title('Linear regression')

print(f2.score(x_train.reshape(-1, 1), y_train))
print(f2.score(x_test.reshape(-1, 1), y_test))
```

```
0.9874209563784625
```

```
0.9878412064961274
```



[]: