HW1

October 27, 2019

[1]:
```
import pandas as pd
```

[2]:
```
dataset = pd.read_csv('dataset.csv')
```

[3]:
```
dataset
```

[3]:
```
          x      y
0     0.010  0.958
1     0.014  1.034
2     0.018  1.109
3     0.022  1.094
4     0.027  1.104
..      …    …
235   0.983 -0.975
236   0.988 -0.903
237   0.992 -1.331
238   0.996 -0.913
239   1.000 -1.019

[240 rows x 2 columns]
```

[4]:
```
%matplotlib notebook
import matplotlib.pyplot as plt
```

[5]:
```
_, ax1 = plt.subplots()
dataset.plot(x='x', y='y', ax=ax1)
```

```
<IPython.core.display.Javascript object>
```

```
<IPython.core.display.HTML object>
```

[5]:
```
<matplotlib.axes._subplots.AxesSubplot at 0x7fee9c7ca9d0>
```

In the following section we want to compare between multiple values of *alpha* for $n = 10$ and *steps* $= 150$. The best value for *alpha* around 1. we demostrate hypothesis function for `[0.1, 1, 1.5]`.

alpha = 0.1

[7]:
```
import numpy as np
import regression

_, ax2 = plt.subplots()
_, ax3 = plt.subplots()
h, mse = regression.regression(data={d[0]: d[1] for d in dataset.values}, n=10,␣
 ↪alpha=0.1, steps=150, lmb=0)

x = [d[0] for d in dataset.values]
y = [d[1] for d in dataset.values]
yh = [np.polynomial.polynomial.polyval(x, h) for x in x]

ax2.plot(x, y, 'r--', x, yh, 'b--')

ax3.plot(range(150), mse)
```

```
<IPython.core.display.Javascript object>


<IPython.core.display.HTML object>


<IPython.core.display.Javascript object>


<IPython.core.display.HTML object>
```

[7]:
```
[<matplotlib.lines.Line2D at 0x7fee5c9d4a50>]
```

alpha = 1

[9]:
```
_, ax4 = plt.subplots()
_, ax5 = plt.subplots()
h, mse = regression.regression(data={d[0]: d[1] for d in dataset.values}, n=10,␣
 ↪alpha=1, steps=150, lmb=0)

x = [d[0] for d in dataset.values]
y = [d[1] for d in dataset.values]
yh = [np.polynomial.polynomial.polyval(x, h) for x in x]

ax4.plot(x, y, 'r--', x, yh, 'b--')

ax5.plot(range(150), mse)
```

```
<IPython.core.display.Javascript object>


<IPython.core.display.HTML object>
```

```
<IPython.core.display.Javascript object>


<IPython.core.display.HTML object>
```

[9]:
```
[<matplotlib.lines.Line2D at 0x7fee5c8c2810>]
```

alpha = 1.115

[12]:
```
_, ax6 = plt.subplots()
_, ax7 = plt.subplots()
h, mse = regression.regression(data={d[0]: d[1] for d in dataset.values}, n=10,␣
 ↪alpha=1.115, steps=150, lmb=0)

x = [d[0] for d in dataset.values]
y = [d[1] for d in dataset.values]
yh = [np.polynomial.polynomial.polyval(x, h) for x in x]

ax6.plot(x, y, 'r--', x, yh, 'b--')

ax7.plot(range(150), mse)
```

```
<IPython.core.display.Javascript object>


<IPython.core.display.HTML object>


<IPython.core.display.Javascript object>


<IPython.core.display.HTML object>
```

[12]:
```
[<matplotlib.lines.Line2D at 0x7fee5c752a50>]
```

In the following section we want to compare between multiple values of *steps* for $n = 10$ and *alpha* $= 1.115$. we demostrate hypothesis function for `[100, 150, 200]`.

steps = 100

[14]:
```
_, ax8 = plt.subplots()
_, ax9 = plt.subplots()
h, mse = regression.regression(data={d[0]: d[1] for d in dataset.values}, n=10,␣
 ↪alpha=1.115, steps=100, lmb=0)

x = [d[0] for d in dataset.values]
y = [d[1] for d in dataset.values]
```

```
yh = [np.polynomial.polynomial.polyval(x, h) for x in x]

ax8.plot(x, y, 'r--', x, yh, 'b--')

ax9.plot(range(100), mse)
```

```
<IPython.core.display.Javascript object>
```

```
<IPython.core.display.HTML object>
```

```
<IPython.core.display.Javascript object>
```

```
<IPython.core.display.HTML object>
```

[14]:
```
[<matplotlib.lines.Line2D at 0x7fee5c5bbf90>]
```

steps = 150

[16]:
```
_, ax10 = plt.subplots()
_, ax11 = plt.subplots()
h, mse = regression.regression(data={d[0]: d[1] for d in dataset.values}, n=10,␣
 ↪alpha=1.115, steps=150, lmb=0)

x = [d[0] for d in dataset.values]
y = [d[1] for d in dataset.values]
yh = [np.polynomial.polynomial.polyval(x, h) for x in x]

ax10.plot(x, y, 'r--', x, yh, 'b--')

ax11.plot(range(150), mse)
```

```
<IPython.core.display.Javascript object>
```

```
<IPython.core.display.HTML object>
```

```
<IPython.core.display.Javascript object>
```

```
<IPython.core.display.HTML object>
```

[16]:
```
[<matplotlib.lines.Line2D at 0x7fee5c2dcb50>]
```

steps = 200

[18]:
```
_, ax12 = plt.subplots()
_, ax13 = plt.subplots()
h, mse = regression.regression(data={d[0]: d[1] for d in dataset.values}, n=10,␣
 ↪alpha=1.115, steps=200, lmb=0)

x = [d[0] for d in dataset.values]
y = [d[1] for d in dataset.values]
yh = [np.polynomial.polynomial.polyval(x, h) for x in x]

ax12.plot(x, y, 'r--', x, yh, 'b--')

ax13.plot(range(200), mse)
```

```
<IPython.core.display.Javascript object>


<IPython.core.display.HTML object>


<IPython.core.display.Javascript object>


<IPython.core.display.HTML object>
```

[18]:
```
[<matplotlib.lines.Line2D at 0x7fee5c186d50>]
```

In the following section we want to compare between multiple values of *degree* for *steps = 150* and *alpha = 1.115*. we demostrate hypothesis function for `[10, 15, 20]`.

degree = 10

[20]:
```
_, ax14 = plt.subplots()
_, ax15 = plt.subplots()
h, mse = regression.regression(data={d[0]: d[1] for d in dataset.values}, n=10,␣
 ↪alpha=1.115, steps=150, lmb=0)

x = [d[0] for d in dataset.values]
y = [d[1] for d in dataset.values]
yh = [np.polynomial.polynomial.polyval(x, h) for x in x]

ax14.plot(x, y, 'r--', x, yh, 'b--')

ax15.plot(range(150), mse)
```

```
<IPython.core.display.Javascript object>


<IPython.core.display.HTML object>
```

```
<IPython.core.display.Javascript object>


<IPython.core.display.HTML object>
```

[20]:
```
[<matplotlib.lines.Line2D at 0x7fee5c7e9190>]
```

degree = 12

[22]:
```
_, ax16 = plt.subplots()
_, ax17 = plt.subplots()
h, mse = regression.regression(data={d[0]: d[1] for d in dataset.values}, n=12,␣
 ↪alpha=1.115, steps=150, lmb=0)

x = [d[0] for d in dataset.values]
y = [d[1] for d in dataset.values]
yh = [np.polynomial.polynomial.polyval(x, h) for x in x]

ax16.plot(x, y, 'r--', x, yh, 'b--')

ax17.plot(range(150), mse)
```

```
<IPython.core.display.Javascript object>


<IPython.core.display.HTML object>


<IPython.core.display.Javascript object>


<IPython.core.display.HTML object>
```

[22]:
```
[<matplotlib.lines.Line2D at 0x7fee5c06d550>]
```

degree 15

[23]:
```
_, ax18 = plt.subplots()
_, ax19 = plt.subplots()
h, mse = regression.regression(data={d[0]: d[1] for d in dataset.values}, n=15,␣
 ↪alpha=1.115, steps=150, lmb=0)

x = [d[0] for d in dataset.values]
y = [d[1] for d in dataset.values]
yh = [np.polynomial.polynomial.polyval(x, h) for x in x]

ax19.plot(x, y, 'r--', x, yh, 'b--')
```

```
ax20.plot(range(150), mse)
```

```
<IPython.core.display.Javascript object>
```

```
<IPython.core.display.HTML object>
```

```
/home/linuxbrew/.linuxbrew/opt/python/lib/python3.7/site-
packages/ipykernel_launcher.py:2: RuntimeWarning: More than 20 figures have been
opened. Figures created through the pyplot interface
(`matplotlib.pyplot.figure`) are retained until explicitly closed and may
consume too much memory. (To control this warning, see the rcParam
`figure.max_open_warning`).
```

```
<IPython.core.display.Javascript object>
```

```
<IPython.core.display.HTML object>
```

```
      ␣
↪---------------------------------------------------------------------------

       NameError                                 Traceback (most recent call␣
↪last)

       <ipython-input-23-54380d1a0309> in <module>
         9 ax19.plot(x, y, 'r--', x, yh, 'b--')
        10
   ---> 11 ax20.plot(range(150), mse)


       NameError: name 'ax20' is not defined
```