ECE-GY 6143 Machine Learning HW 04

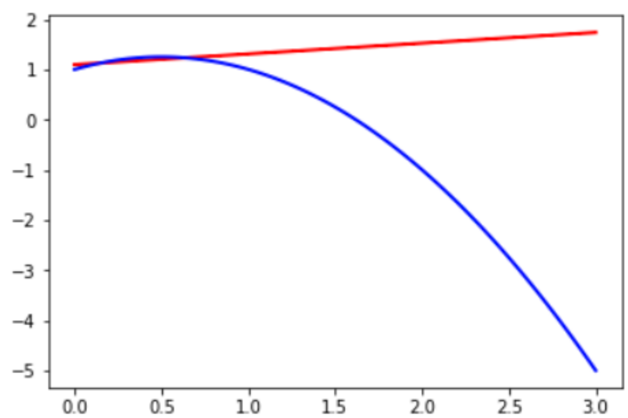
Tongda Xu

February 9, 2019

1. Question 1:

```
RSS = np.zeros(len(dtest))
for i in range(0, len(dtest)):

    j = np.arange(dtest[i])
    X_prep = np.exp(-np.dot(x.reshape(x.shape[0], 1),
        j.reshape(1, j.shape[0]))/dtest[i])

    x_train, x_test, y_train, y_test
        = np.array(model_selection.train_test_split
        (X_prep,y, test_size=0.5))

    model.fit(x_train, y_train)
    yhat = model.predict(x_test)
    RSS[i] = np.sum((yhat - y_test)**2)

optimal_d = dtest[argmin(RSS)]
```

Out[20]: [<matplotlib.lines.Line2D at 0x18794bd0860>]



2. Question 3:

   a.

   $BIAS(x) = E(f(x,\beta)) - f(x,\beta_0)$

   $= E(\frac{\Sigma\beta_0 x_i^2}{\Sigma x_i^2}x) - f(x,\beta_0)$

   $= 0$

   b.

   $Bias(x) = E(\frac{\Sigma\beta_0 x_i^2 + \Sigma\epsilon}{\Sigma x_i^2}x) - f(x,\beta_0)$

   $= \frac{x}{\Sigma x_i}\Sigma E(\epsilon)$

   $= 0$

   c.

   $Bias(x) = E(\frac{\Sigma\beta_0 x_i^2 + \Sigma 2\beta_0 x_i\epsilon + \Sigma\beta_0\epsilon^2}{\Sigma x_i^2}x) - f(x,\beta_0)$

   $= \frac{\beta_0 x}{\Sigma x_i}\Sigma E(\epsilon^2)$

   $= \frac{\beta_0 x N}{\Sigma x_i}\delta^2$

3. Question 4

   a.

   $\hat{\beta} = [(1,x)^T(1,x)]^{-1}(1,x)^T y$

   b.

   $\hat{\beta} = [(1,x)^T(1,x)]^{-1}(1,x)^T(1,x,x^2)\beta_0$

   c.

```
X = np.random.rand(10)
y = 1 + X - X**2
A = np.zeros([10, 2])
```

```
A[:,0] = 1
A[:,1] = X

beta = np.linalg.lstsq(A,y)[0]
x = x_y = np.linspace(0, 3, 100)
y_pred = beta[0] + beta[1]*x
y_true = 1 + x - x**2

plt.plot(x,y_pred,'r-',linewidth=2)
plt.plot(x,y_true,'b-',linewidth=2)
```

d.
from the Graph, $x = 3$
But there should be a prove analytically, solving β is to minimize:
$E\{(f(X) - \hat{f}(X))^2\} = \int_0^1 (1 + 2x - x^2 - \beta_0 - \beta_1 x)^2 Pr(x)dx$
Since $Pr(X)$ is a constant line, after some complex polynomial calculation, we should reach the same result

4. Question 5
   a.
   $c_i = cancer - volume, a_i = age, t1_i = is - Type - 1, t2_i = is - Type - 2$
   $Model1 : \hat{y} = (1, c)\hat{\beta}$
   $Model2 : \hat{y} = (1, c, a)\hat{\beta}$
   $Model3 : \hat{y} = (1, c, a, t1, t2)\hat{\beta}$
   b.
   $[2, 3, 5], Model3$
   c.
   $$M_1 = \begin{bmatrix} 1 & 0.7 \\ 1 & 1.3 \\ 1 & 1.6 \\ \vdots & \vdots \end{bmatrix}$$
   $$M_2 = \begin{bmatrix} 1 & 0.7 & 55 \\ 1 & 1.3 & 65 \\ 1 & 1.6 & 70 \\ \vdots & \vdots & \vdots \end{bmatrix}$$
   $$M_3 = \begin{bmatrix} 1 & 0.7 & 55 & 1 & 0 \\ 1 & 1.3 & 65 & 0 & 1 \\ 1 & 1.6 & 70 & 0 & 1 \\ \vdots & \vdots & \vdots & \vdots & \end{bmatrix}$$

d.

$Model2 : 0.72 < 0.70 + 0.05$