Homework 7

December 6, 2016

* For each question, give your answer. Justify your answer using a detailed explanation.

1 Understanding kNN)

Consider the email and spam prediction problem. We obtain the following data set of binary-valued features about each email, including whether I know the author or not, whether the email is long or short, and whether it has any of several key words, along with my final decision about whether to read it ($y = +1$ for "read", $y = -1$ for "discard").

x_1 know author?	x_2 is long?	x_3 has 'research'	x_4 has 'grade'	x_5 has 'lottery'	y read?
0	0	1	1	0	-1
1	1	0	1	0	-1
0	1	1	1	1	-1
1	1	1	1	0	-1
0	1	0	0	0	-1
1	0	1	1	1	1
0	0	1	0	0	1
1	0	0	0	0	1
1	0	1	1	0	1
1	1	1	1	1	-1

Now we have new emails as following:

$$
\begin{aligned}
T_1 =& [1 \quad 1 \quad 0 \quad 0 \quad 1] \\
T_2 =& [1 \quad 0 \quad 1 \quad 0 \quad 1] \\
T_3 =& [0 \quad 1 \quad 0 \quad 1 \quad 0]
\end{aligned}
$$

(i) Use 1NN, classify T_1, T_2 and T_3.

(ii) Use 3NN, classify T_1, T_2 and T_3.

(iii) Calculate the Cosine Similarity we introduced in class, and use 1NN to classify T_1, T_2 and T_3.

2 First-order optimization methods

(1) For the univariate ℓ_0 regularization problem

$$
\min_{x \in \mathbb{R}^n} \quad \frac{1}{2}(x-z)^2 + \tau \|x\|_0,
$$

show that the optimal solution is given by

$$
x^* = \mathbb{H}_\tau(z) = \begin{cases} z, & |z| > \sqrt{2\tau}, \\ 0, & |z| \le \sqrt{2\tau}. \end{cases}
$$

Here $\mathbb{S}_\tau$ is called the hard-thresholding operator with threshold τ.

(2) For the univariate ℓ_1 regularization problem

$$
\min_{x \in \mathbb{R}^n} \quad \frac{1}{2}(x-z)^2 + \tau \|x\|_1,
$$

show that the optimal solution is given by

$$
x^* = \mathbb{S}_\tau(z) = \begin{cases} z - \tau, & z > \tau, \\ 0, & |z| \le \tau, \\ z + \tau, & z - \tau. \end{cases}
$$

Here $\mathbb{S}_\tau$ is called the soft-thresholding operator with threshold τ.

(3) Let $f(x) = \frac{1}{2}\|Ax - y\|_2^2$. Then at point x,

$$
Q(z; x) = f(x) + \nabla f(x)^T (z - x) + \frac{1}{2}\|z - x\|_2^2
$$

is a quadratic approximation for $f(z)$ at x. Now consider the ℓ_1 regularized problem $\min_x \; f(x) + \tau\|x\|_1$, at iterate x_k, one can replace the original function by a local model $Q(z; x_k) + \tau\|z\|_1$. The new iterate x_{k+1} is determined by the optimal solution of

$$x_{k+1} = \arg\min_z Q(z; x_k) + \tau\|z\|_1.$$

Find the expression for x_{k+1} (using soft-thresholding operator $\mathbb{S}_\tau$). This is the IST (iterative soft-thresholding) algorithm I introduced in the class.

(4) In the attached zipped file, you can find `demo_compressed.m, soft.m, IST.m`. In `demo_compressed.m`, a compressed sensing example is generated, and the IST algorithm (implemented in `IST.m`, the soft-thresholding operator is implemented in `soft.m`) is applied to solve the problem. The code is almost finished, leaving only a few lines (marked as "???") for you to complete.

Complete the missing lines (you can write additional lines if you need), and run `demo_compressed.m`. Show the three plots in your homework.