4.1

A

4.2

B

4.3

B

4.4

A

4.5

AC

4.6

4.7

a)

∵,

If x(t) is misclassifed by (t),H (x) and y (t) will take the inverse value(1.h(x)=1,y(t)=1 or 2.h(x)=-1,y(t)=1)

∴ < 0

b)

∵x0=1且

∴

c)

1.当h(x) = -1,y(t)=1

y(t)=1且(t)x(t) = -1时，x(t)被错误分类为-1。因为与x(t)相距较远，而t+1)=+y(t)x(t)使得t+1)更接近x(t)，从而x(t)被正确分类

2.当h(x) =1,y(t)=-1

y(t)=-1且(t)x(t) = 1时，x(t)被错误分类为+1。因为与x(t)太接近，而t+1)=+y(t)x(t)使得t+1)更远离x(t)，从而x(t)被正确分类

4.8

f(x) = sign(·x + b),其中

4.9

D

4.10

C

4.11

C

4.12

B

4.13

A

4.14

A

4.15

D

4.16

B

4.17

A

4.18

A

4.19

AB

4.20

C

4.21

a)

1.处理多重共线性问题，提高模型稳定性

2.通过引入L2正则化项，减少回归系数大小，有利于泛化

b)

1.用于特征选择，可以将某些回归系数压缩至0

2.适用于数据中纯在大量无关或冗余特征情况

c)

1.结合L1和L2正则化项，平衡特征选择和多重共线性处理

2.更适用于需要权衡两种正则华方法情况，但需调整两个超参数

4.22

不会，成本函数为凸函数。