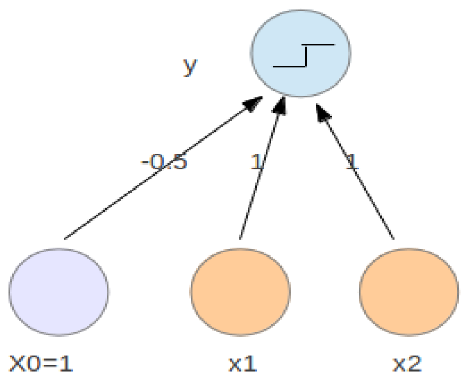
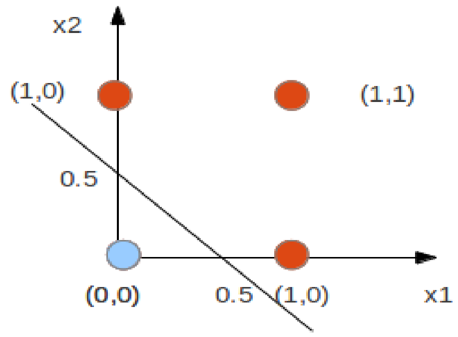
1.



x1	x2	r
0	0	0
1	0	1
0	1	1
1	1	1

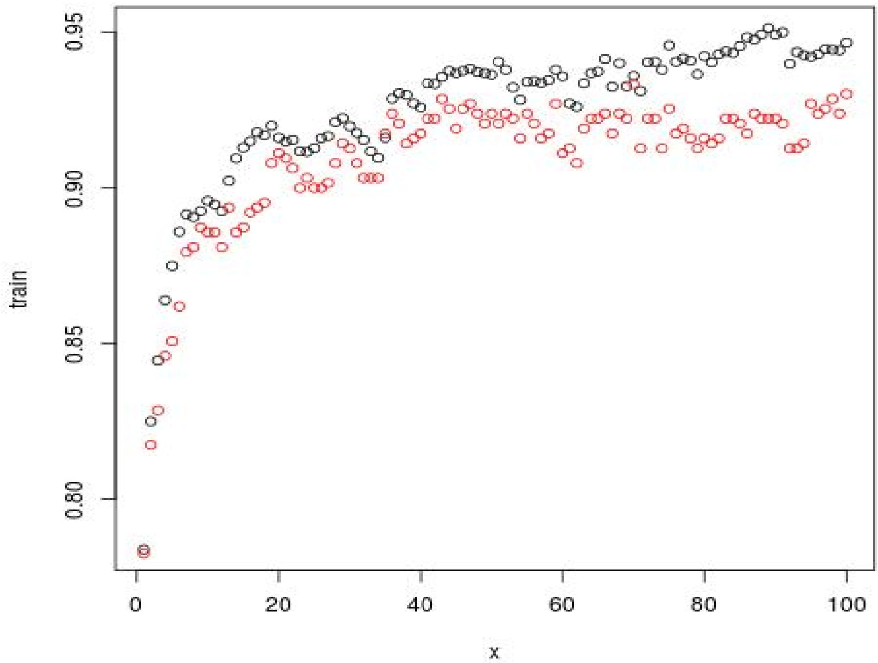


2.

(e) epoch from 1:100

black for train accuracy, red for test accuracy

with H=20, n=0.1



choose epoch=50

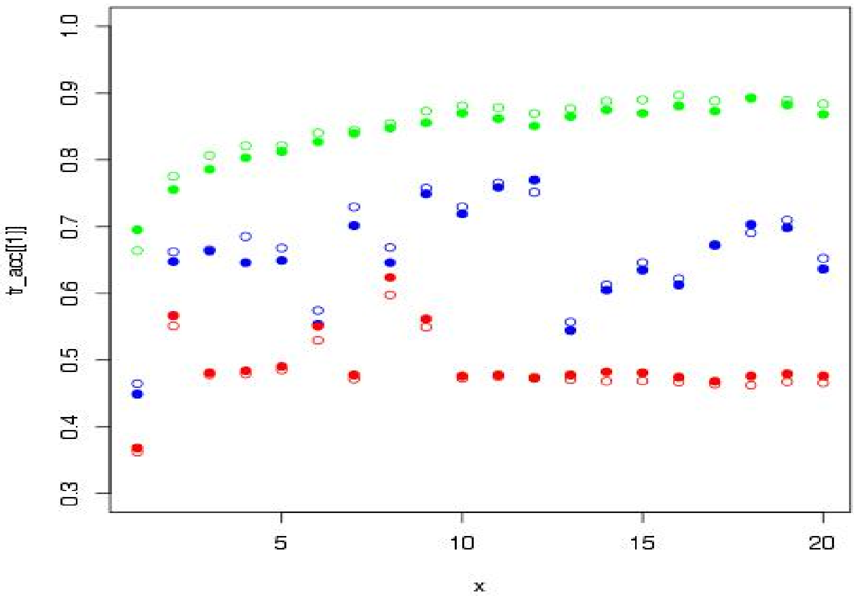
(f_n)

x is epoch from 1:50

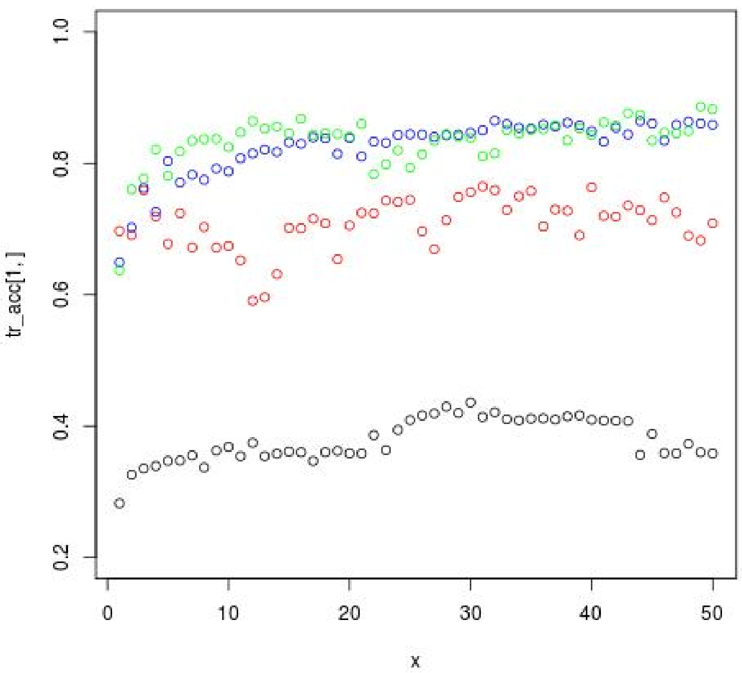
circles is training accuracy, solid points are test accuracy

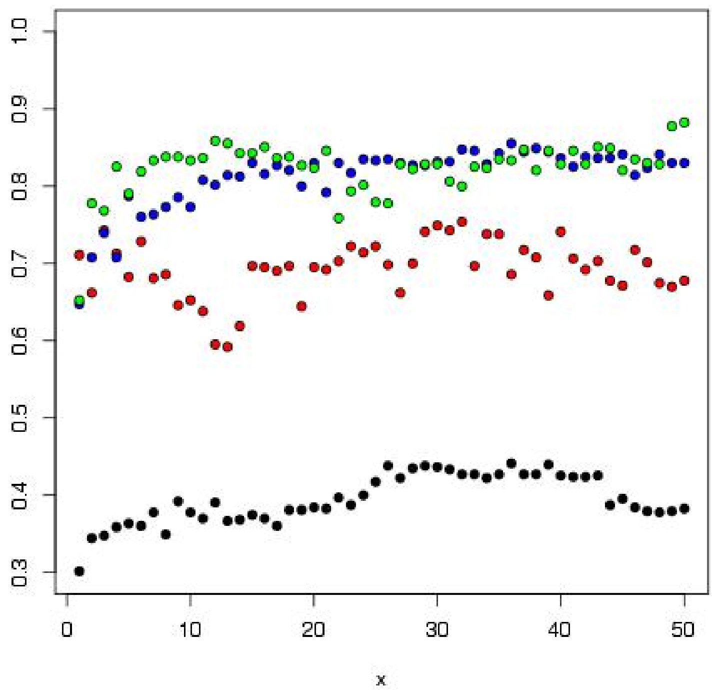
H=20

n=c(0.01, 0.1, 0.3) from top to bottom



increase n increase both accuracy and, smooth the curve
choose n=0.01
(f_h)
x is epoch from 1:50
circles is training accuracy, solid points are test accuracy
H=c(5,20,40,100) from bottom to top, green for 100, blue for 40
n=0.1





choose H=40

2.(g)

```
   [1] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 7 1 1 1 1 1 1
  [25] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 3 1 1 1 1 1
  [49] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
  [73] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 3 1 1 5 1 1 1 7 1 1
  [97] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 3
 [121] 1 5 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 7 1 1 1
 [145] 1 1 1 1 1 1 1 1 1 2 2 2 2 2 2 2 2 2 2 2 8 2 2 2
 [169] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
 [193] 2 2 2 2 2 2 2 6 2 2 2 2 2 2 9 2 2 2 2 2 2 2 2 2
 [217] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 1 2 2 2 2 2 2 2 2
 [241] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 7 8 2 2
 [265] 2 2 6 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 3 2 2 2
 [289] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 3 4
 [313] 3 3 3 3 3 3 3 3 3 3 3 4 3 3 3 4 3 3 3 3 3 3 3 3
 [337] 3 3 3 4 3 3 3 4 8 3 3 3 3 3 3 1 4 4 3 3 3 3 3 3
 [361] 3 3 3 3 3 3 3 6 3 4 3 4 3 3 3 3 3 3 3 3 3 3 3 3
 [385] 3 3 1 3 3 3 3 3 3 3 3 7 3 3 3 3 3 3 4 1 3 3 3 3
 [409] 3 3 3 3 3 3 3 3 3 3 3 3 3 6 3 3 3 3 3 3 3 3 3 3
 [433] 3 4 3 3 7 3 3 3 3 3 3 1 3 3 3 3 3 3 3 3 3 3 4 3
 [457] 3 3 3 4 7 3 3 3 4 3 3 3 3 3 3 4 4 4 4 4 4 4 4 4
 [481] 4 4 4 4 4 4 4 4 4 4 4 5 4 4 4 4 4 4 4 4 4 4 4 4
 [505] 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4
 [529] 4 4 4 1 1 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4
 [553] 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4
 [577] 5 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4 4
 [601] 5 4 4 4 4 4 5 4 6 4 4 4 4 4 4 4 4 4 4 5 4 5 5 5
 [625] 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 9 5
 [649] 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 4 5 5 5 5 5 5
 [673] 4 5 5 5 1 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5
 [697] 9 5 5 5 5 5 5 5 5 5 5 4 5 5 5 5 5 9 9 5 5 5 5 5
 [721] 5 5 5 5 5 5 6 4 5 5 5 5 9 5 5 5 5 5 5 5 5 5 5 9
 [745] 5 5 5 5 5 5 5 5 5 5 5 5 3 5 5 5 5 5 5 5 2 5 5 5
 [769] 5 5 5 9 5 5 5 5 5 6 6 6 6 4 6 6 6 6 6 6 6 6 5 6
 [793] 6 6 6 6 6 5 6 6 6 6 6 6 6 6 6 3 6 6 6 6 9 6 6 6
 [817] 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6
 [841] 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6
 [865] 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 5 6 6 6 6
 [889] 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 6 3 6 6 6 6
 [913] 6 6 6 6 6 6 3 6 6 6 6 6 6 6 6 6 6 6 3 6 6 6 5 6
 [937] 6 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 6
 [961] 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7
 [985] 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 5 7 7 7 7 7 7
[1009] 7 7 7 7 7 7 7 7 7 7 5 7 7 7 7 7 7 7 7 7 7 7 7 7
[1033] 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 3 7 7 7 7
[1057] 7 7 7 7 7 7 7 7 7 7 7 7 1 7 7 7 7 7 7 9 7 7 7 7
[1081] 7 3 7 7 7 7 7 7 8 8 8 8 8 8 8 8 8 9 8 8 7 8 8 8
[1105] 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8
[1129] 9 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8
```

```
[1153]  8  8  8  8  8  8  8  8  8  8  8  8  8  8  8  8  8  8  8  8  8  8  8  8
[1177]  8  8  8  8  8  8  8  8  8  8  8  8  8  8  8  8  8  8  8  3  8  8  8  8
[1201]  8  8  8  8  8  8  8  8  8  8  8  8  8  8  8  8  8  8  8  8  8  8  8  3
[1225]  8  8  8  8  8  8  8  8  8  8  8  8  8  8  8  8  8  8  8  8  8  8  8  9
[1249]  9  9  9  9  9  1  9  9  9  9  9  9  9  9 11  9  9  9  9  9  9  9  9  9
[1273]  9  9  9  9  9  9  9  9  9  9  9  9  9  9  9  9  9  9  9  9  9  9  9  9
[1297]  9  9  9  9  9  9  9  9  9  9  9  9  9  9  9  9  9  9  9  9  9  9  9  9
[1321]  9  9  9  9  9  9  9  9  9  9  9  9  9  9  9  9  9  9  9  9  9  9  9  9
[1345]  9  9  4  9  9  9  9  9  9  9  9  9  9  9  9  9  9  9  9  9  9  9  5  9
[1369]  9  9  9  9  9  9  9  9  9  9  9  9  9  9  9  9  9  9  9  9  8  9  9  9
[1393]  9  9  9  9  9  9  9 10 10 10 10 10 10 10 10 10 10 10 10 10  8 10 10 10
[1417] 10  4 10 10 10 10 10 10  4 10 10  1 10 10 10 10 10 10 10 10 10 10 10 10
[1441] 10 10 10 10 10 10 10 10 10 10 10 10  4 10 10 10 10 10 10 10 10  4 10 10
[1465] 10 10 10 10 10 10 10 10 10 10 10 10 10 10  7 10 10 10 10 10  4 10 10 10
[1489] 10 10 10 10 10 10  4  4 10  8 10 10 10 10 10 10 10 10  4 10 10 10 10 10
[1513] 10 10 10 10 10 10 10 10 10 10 10 10  3  4 10  2 10 10 10 10 10 10 10 10
[1537] 10 10 10 10 10 10 10 10 10 10 10 10 10 10  4 10 10 10 10 11 11 11 11 11
[1561] 11 11  1 11 11 11  1 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
[1585] 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
[1609] 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11  1 11
[1633] 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
[1657] 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
[1681] 11 11 11 11 11  3 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
[1705] 11 11 11 11 11 11 11 11 11
```

3.(a)

Minimize margin $1/2||w||^2$

Minimize soft error $C \sum \xi (t)$

s.t. $r (t) (\mathbf{w}x(t) + w0) \geq 1-\xi(t)$,

$\xi(t)>0$ for all $x(t)$

$Lp = 1/2||w||^2+ C \sum\xi(t) - \sum\alpha (t) (r(t) (w x(t) + w 0) - 1 + \xi (t)) -\sum \mu (t)\xi (t)$

minimize $w, w0, \xi (t)$

maximize $\alpha (t)$

$\partial Lp /\partial w =0$ => $w=\sum \alpha(t) r (t)x(t)$

$\partial Lp /\partial w0 =0$ => $\sum\alpha (t)r (t) = 0$

$\partial Lp /\partial\xi(t) =0$ ⇒ $C -\alpha (t) - \mu(t) = 0$

plug back to Lp maximize $\alpha(t)$

$Ld=- \sum\sum \alpha(t)\alpha(s)r(t) r(s) x (t)(t)x(t) + \sum \alpha (t)$

subject to $\sum \alpha (t)r(t) = 0$ and $0 \leq \alpha (t) \leq C$,

with KKT condition

$\alpha(t) \{r (t) (wx(t) + b) - 1 + \xi i \} = 0$

$u(t)\xi(t) = 0$

or

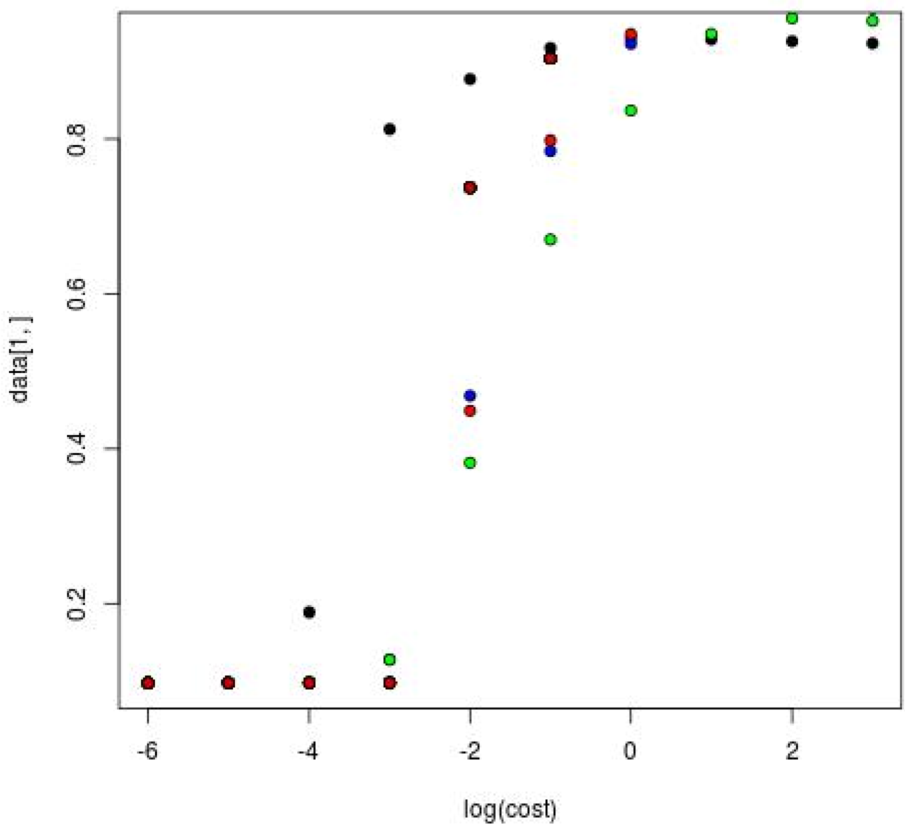
$\alpha i = 0$ $r (t) (w x(t) + w0) \geq 1$

$\alpha i = C$ $r(t) (wx(t) + w0) \leq 1$

$0 < \alpha i < C$ $r(t) (w x (t)+ w0) = 1$

3(b).

	1/1000000	1/100000	1/10000	0.0001	0.001	0.01	1	10	100	1000
linear	0.0984127	0.0984127	0.1892063	0.8120635	0.8766667	0.9168254	0.9293651	0.9279365	0.9255556	0.9228571
polynomial 2	0.0984127	0.0984127	0.0984127	0.0984127	0.4490476	0.7979365	0.9349206	0.9714286	0.9730159	0.9704762
Polynomial 3	0.0984127	0.0984127	0.0984127	0.1285714	0.4684127	0.7844444	0.9222222	0.9714286	0.9749206	0.9688889
Polynomial 4	0.0984127	0.0984127	0.0984127	0.1280952	0.3819048	0.6701587	0.8360317	0.9350794	0.9552381	0.951904
Radial 10^-6	0.0984127	0.0984127	0.0984127	0.0984127	0.7371429	0.9034921	0.9714286	0.9877778	0.9855556	0.9852381
Radial 10^-5	0.0984127	0.0984127	0.0984127	0.0984127	0.7371429	0.9034921	0.9714286	0.9877778	0.9855556	0.9852381
Radial 10^-4	0.0984127	0.0984127	0.0984127	0.0984127	0.7371429	0.9034921	0.9714286	0.9877778	0.9855556	0.9852381
Radial 10^-3	0.0984127	0.0984127	0.0984127	0.0984127	0.7371429	0.9034921	0.9714286	0.9877778	0.9855556	0.9852381
Radial 10^-2	0.0984127	0.0984127	0.0984127	0.0984127	0.7371429	0.9034921	0.9714286	0.9877778	0.9855556	0.9852381
Radial 10^-1	0.0984127	0.0984127	0.0984127	0.0984127	0.7371429	0.9034921	0.9714286	0.9877778	0.9855556	0.9852381
Radial 1	0.0984127	0.0984127	0.0984127	0.0984127	0.7371429	0.9034921	0.9714286	0.9877778	0.9855556	0.9852381
Radial 10	0.0984127	0.0984127	0.0984127	0.0984127	0.7371429	0.9034921	0.9714286	0.9877778	0.9855556	0.9852381
Radial 100	0.0984127	0.0984127	0.0984127	0.0984127	0.7371429	0.9034921	0.9714286	0.9877778	0.9855556	0.9852381
Radial 1000	0.0984127	0.0984127	0.0984127	0.0984127	0.7371429	0.9034921	0.9714286	0.9877778	0.9855556	0.9852381
Radial 10000	0.0984127	0.0984127	0.0984127	0.0984127	0.7371429	0.9034921	0.9714286	0.9877778	0.9855556	0.9852381
Radial 100000(	0.0984127	0.0984127	0.0984127	0.0984127	0.7371429	0.9034921	0.9714286	0.9877778	0.9855556	0.9852381
Radial 100000(	0.0984127	0.0984127	0.0984127	0.0984127	0.7371429	0.9034921	0.9714286	0.9877778	0.9855556	0.985238



accuracy increase with cost because when cost is low we do not pay much penety to data voilate margin

linear is sufficient for classification with cost greater than 10^-3

polynomials doesn't show better(worse) then linear

radial doesn't show better(worse) then linear

so, data are linear seperatable