Chapter 3

1. (bias term) 偏置项 b 提供多一个参数消除无关的偏移。当你的超平面是过原点时，即成比例的假设前提下，不用 b.

2. ① $y(w) = \frac{1}{1+e^{-(w^Tx+b)}}$ 证明为非凸，

求错了 用充分条件，$y''(w)$ ~~~~ ≠ 不恒 >0 或 <0.

$\ln\frac{y}{1-y} = w^Tx+b = \ln(-1+\frac{1}{1-y})$

$-\frac{1-y}{y}\cdot\frac{y'}{(1-y)^2} = \frac{1}{y(y-1)}\cdot y' = x$

~~$y' = w^Ty(y-1) = w(w^Tx+b)(w^Tx+b-1)$~~

~~$= w\cdot w^Txx^Tw + w(2b-1)w^Tx + b(b-1)$~~

~~$y'' =$~~ $y' = xy(y-1) = x\cdot(w^Tx)^2 + x(2b-1)w^Tx + x\cdot b(b-1)$

$= 2x\cdot x^T + (2b-1)x\cdot x^T$

$= (2b+1)x\cdot x^T$

$\forall y \in \mathbb{R}^n$, $y^Tx\cdot x^Ty \geq 0 \Rightarrow x\cdot x^T$ is ~~semi-de~~ semi-positive definite.

$\Rightarrow y(w)$ is convex or concave function,

it depends on $\delta(2b+1)$. (Dirac function)

题目意思是

2. ① $y(x)=\frac{1}{1+e^{-(w^Tx+b)}}$ 证明为非凸.（下面把x当成变量，实际把w^T当变量，结果一样，但是我看错题懒得改了）

只需证明二阶偏导矩阵不为正(负)定矩阵即可.）

$$\ln\frac{y}{1-y}=w^Tx+b=\ln(-1+\frac{1}{1-y})$$

$$-\frac{1-y}{y}\cdot\frac{y'}{(1-y)^2}=\frac{y'}{y(y-1)}=w$$

$$y'=\frac{-w(e^{-(w^Tx+b)})}{(1+e^{-(w^Tx+b)})^2}$$

$y'\cdot(1+e^{-(w^Tx+b)})^2=-w\cdot e^{-(w^Tx+b)}$ 求导，设 $h(x)=e^{-(w^Tx+b)}$

$$y''(1+h(x))^2+y'\cdot 2(1+h(x))\cdot h'(x)^T=-w\,h'(x)^T$$

$$h'(x)=-w\cdot h(x)=-w\cdot(w^Tx+b)$$

$$y''=\frac{1}{(1+h(x))^2}\cdot\left(y'\cdot 2(1+h(x))+w\right)(w^T)\,h(x)$$

其实代个特值，证明不是凸函数就可以了.

② 同样，对 $\ell(\beta)=\sum_{i=1}^{m}(-y_i\beta^T\hat{x}_i+\ln(1+e^{\beta^T\cdot\hat{x}_i}))$ 对β求二阶导，

同上求法，可知~~其~~二阶偏导矩阵正定(负定)，

则可推出 $\ell(w)$ 也正定(半正定)，则为convex or concave.

6. 可用核函数，将有限维映射到高维，再求线性模型.

~~7. 码长为9，即9个分类器 f_1, ——, f_9，共4类，C_1, C_2, C_3, C_4.~~

~~求4个类的坐标，s.t. 两两之间距离和最短.~~

~~不妨设4个类距离近可能相同，~~

~~假设~~

~~假设~~ 看

~~采用海明距离，方便计算.~~

$\cancel{|c_1-c_2|+|c_1-c_3|+\cdots+|c_3-c_4| \le (|c_1|+|c_2|+|c_3|+|c_4|)\times 3}$

~~(所有不重复的距离之和，假设距离相同，即方差为0)~~

$\cancel{\Rightarrow 6\cdot|c_1-c_2| \le 9\times 4\times 3 \Rightarrow |c_1-c_2| \le}$

7. 4个类，9个码长，即9种分类器，

 22分类器有 $\frac{1}{2}\binom{4}{2}=3$ 种，

 13分类器有4种，

 最多有7种有意义的二元分类器，除非用相同的分类器

但是这样结论会重复，没有意义，所以我不知道如何分出9个码.

8. 若出错概率相当，则应尽可能遍及多的类，

 让每个类所在的分类器数量尽可能大，

 以降低干扰.

9. 假设由 $f_1,\cdots f_m$ 分类器，得到的码为 $(x_1,\cdots x_m)$

∀每个码，若在 f_i 上 ~~分配~~ ~~最值~~ 归类为 $c_1,\cdots c_N$ 概率相同

设 c_j 码为 $(c_{j,1},\cdots,c_{j,m})$，

期望距离为 $E(|x-c_j|)=E(x_1-c_{j,1})+\cdots+E(c_m-c_{j,m})$

$\cancel{E(x_i-c_{j,i})=E(E(x_i-c_{j,i})|x_i)=} \frac{k}{N}$

(假设 $c_{j,i}=1$ 时，取 k_i 个类，$c_{j,i}=-1$ 时取 $N-k_i$ 个类)

$\cancel{E(|x_i-c_{j,i}|)=E(E(|x_i-c_{j,i}|\,|x_i))} = \frac{k_i}{N}$ 当 $c_{j,i}=1$,

f_i 分类器有 k_i 个类取值 1，$N-k_i$ 个类取 -1，

$\cancel{E(|x-c_j|)=}$

$\cancel{E(|x_i-c_{j,i}|)=E(E(E(|x_i-c_{j,i}|\,|x_i)|c_{j,i}))}$

$\cancel{=\sum_i\sum_j P(|x_i-c_{j,i}|)\,|x_i-c_{j,i}|\cdot P(c_{j,i},x_i)}$

$\cancel{P(c_{j,i},x_i)=P(c_{j,i}|x_i)P(x_i)}$,

$\cancel{P(c_{j,i}|x_i)=\{}$

$$E(|x_i-c_{j,i}|)=E(E(|x_i-c_{j,i}|\ \text{if}\ c_j\in f_i))$$

$$=P(c_j\in f_i)\cdot E(|x_i-c_{j,i}|\ |\ c_j\in f_i)+\cdots$$

$$\cancel{=\frac{1}{2}\cdot\frac{N-k_i}{N}+\frac{1}{2}\cdot\frac{k_i}{N}}=\frac{k_i}{N}\cdot\frac{N-k_i}{N}+\frac{N-k_i}{N}\cdot\frac{k_i}{N}$$

$$\cancel{=\frac{1}{2}}=\frac{2k_i(N-k_i)}{N^2}$$

$E(|x-c_j|)=\sum_{i=1}^{m}\frac{2k_i(N-k_i)}{N^2}$.　与 j 无关，故期望距离相同.

f_i 的选取与期望距离相关. □

所以无需针对类别之间进行处理. □