3.1

B

3.2

3.3

B

3.4

我们定义感知机的损失函数为

$$L(\hat{w})=max(0,-y_i\hat{y}_i)=-\sum_{i\in M}y_i\hat{w}^Tx_i$$

梯度表示为

$$\nabla_w L=-\sum_{i\in M}y_ix_i$$

(a)

因为$x(t)$被$w(t)$误分类,所以$w^T(t)x(t)$计算出来的y和$y(t)$符号相反，所以$y(t)w^T(t)x(t)<0$

(b)

由于$-y(t)w^T(t)x(t)$表示被误分类时的损失函数，当遇到一个误分类的点时，通过沿着梯度相反方向调整w，使损失函数变小。及

$$\begin{aligned}L(w(t))&>L(w(t+1))\\-y(t)w^T(t)x(t)&>-y(t)w^T(t+1)x(t)\\\Rightarrow y(t)w^T(t+1)x(t)&>y(t)w^T(t)x(t)\end{aligned}$$

(c)

因为$w(t+1)=w(t)+y(t)x(t)$

且$y(t)x(t)$表示梯度的相反方向，所以每次迭代都会导致损失函数$L(\hat{w})$减小,从而使参数沿着能正确分类的方向前进。

3.5

令

$$e = argmin\frac{1}{N}\sum_{i=1}^{N}(f(x_i)-y_i)^2$$

$$\begin{cases} \frac{de}{dw_1}=0 \\ \frac{de}{dw_2}=0 \\ \frac{de}{d_d}=0 \end{cases}$$

整理可得

$$w_1 = \frac{\sum x_{12}(x_{11}-\overline{x1})\sum(x_{11}(y_1-\overline{y})-\sum x_{11}(x_{12}-\overline{x2})\sum(x_{12}(y_i-\overline{y}))}{\sum(x_{11}^2-x_{11}\overline{x1})\sum(x_{12}^2-x_{12}\overline{x2})-\sum(x_{11}x_{12}-x_{12}\overline{x1})\sum(x_{11}x_{12}-x_{11}\overline{x2})}$$

$$w_2 = \frac{\sum x_{11}(x_{11}-\overline{x1})\sum(x_{12}(y_1-\overline{y})-\sum x_{12}(x_{11}-\overline{x1})\sum(x_{11}(y_i-\overline{y}))}{\sum(x_{12}^2-x_{12}\overline{x2})\sum(x_{11}^2-x_{11}\overline{x1})-\sum(x_{12}x_{11}-x_{11}\overline{x2})\sum(x_{12}x_{11}-x_{12}\overline{x1})}$$

$$b = \overline{y}-w_1\overline{x_1}-w_2\overline{x_2}$$

3.6

构建最优化问题:

$$\min_{w,b} L(w,b) = -\sum_{x_i\in M} y_i(w\cdot x+b)$$

每次迭代

$$w \leftarrow w+\eta y_i x_i$$
$$b \leftarrow b+\eta y_i$$

(1)取初值$w_0=0, b_0=0$

(2)对$x_1=(2,4)^T, y_1(w_0\cdot x_1+b_0)=0$，未能被正确分类，更新$w, b$

$$w_1 = w_0+\eta y_1 x_1 = (1,2)^T, b_1 = b_0+\eta y_1 = 1/2$$

得到线性模型

$$w_1\cdot x+b_1 = 1x^{(1)}+2x^{(2)}+1/2$$

(3)对x_1, x_2，显然$wx+b>0$，被正确分类，不更新w,b。但是对x_3，$xw+b>0$，被错误分类，更新w,b。

$$w_2 = (1,2)^T+0.5\cdot-1\cdot(0,1)^T = (1,1.5)^T, b_2 = 0.5+0.5\cdot-1 = 0$$

得到线性模型

$$w_2 x+b_2 = 1x^{(1)}+1.5x^{(2)}+0$$

(4)对x_1, x_2，显然$wx+b>0$，被正确分类，不更新w,b。但是对x_3，$xw+b>0$，被错误分类，更新w,b。

$$w_3=(1,1.5)^T+0.5\cdot -1\cdot (0,1)^T=(1,1)^T, b_3=0+0.5\cdot -1=-0.5$$

得到线性模型

$$w_3x+b_3=x^{(1)}+x^{(2)}-0.5$$

(5)对x_1, x_2，显然$wx+b>0$，被正确分类，不更新w,b。但是对x_3，$xw+b>0$，被错误分类，更新w,b。

$$w_4=(1,1)^T+0.5\cdot -1\cdot (0,1)^T=(1,0.5)^T, b_4=-0.5+0.5\cdot -1=-1$$

得到线性模型

$$w_4x+b_4=x^{(1)}+0.5x^{(2)}-1$$

到此x_1, x_2, x_3都能被正确分类，得到感知机模型为

$$f(x)=sign(x^{(1)}+0.5x^{(2)}-1)$$