设数据集$\mathcal{D}=\left\{(x_i,y_i)\right\}_{i=1}^{n}, x_i\in\mathbb{R}^d, y_i\in\{0,1\}, \theta=\{\pi,\mu_0,\Sigma_0,\mu_1,\Sigma_1\}$，则在整个数据集上有，

$$\begin{aligned}P(\mathcal{D}) &= \prod_{i=1}^{n} p(x_i,y_i;\theta)\\ &= \prod_{i=1}^{n} p(y_i;\theta)\, p(x_i|y_i;\theta)\end{aligned}$$

这里有，

$$p(y_i;\theta)=\pi^{y_i}(1-\pi)^{1-y_i}$$

$$p(x_i|y_i;\theta)=\Big(\mathcal{N}(x_i|\mu_1,\Sigma_1)\Big)^{\mathbb{1}\{y_i=1\}}\cdot\Big(\mathcal{N}(x_i|\mu_0,\Sigma_0)\Big)^{\mathbb{1}\{y_i=0\}}$$

$$\mathcal{N}(x|\mu,\Sigma)=(2\pi)^{-\frac{d}{2}}det(\Sigma)^{-\frac{1}{2}}exp\left\{-\frac{1}{2}(x-\mu)^{\top}\Sigma^{-1}(x-\mu)\right\}$$

（1）试求解$\theta=\{\pi,\mu_0,\Sigma_0,\mu_1,\Sigma_1\}$，以优化如下的目标函数，

$$\underset{\theta}{argmax}\log P(\mathcal{D})$$

（2）令

$$\begin{aligned}\log\frac{p(y_1|x;\theta)}{p(y_0|x;\theta)} &= \log\frac{p(y_1;\theta)\,p(x|y_1;\theta)}{p(y_0;\theta)\,p(x|y_0;\theta)}\\ &= \log\frac{\pi\ \cdot\mathcal{N}(x|\mu_1,\Sigma_1)}{(1-\pi)\cdot\mathcal{N}(x|\mu_0,\Sigma_0)}\\ &\doteq w^{\top}x+b \quad (where\ w\in\mathbb{R}^d, b\in\mathbb{R})\end{aligned}$$

试求解w,b，并探究w与$(\mu_1-\mu_0)$之间的关系?

注：可能会用到的公式

$$\frac{\partial\, log\, det(\Sigma)}{\partial\Sigma}=(\Sigma^{-1})^{\top}$$

$$\frac{\partial\,(x^{\top}A^{-1}x)}{\partial A}=-(A^{-1})^{\top}xx^{\top}(A^{-1})^{\top}$$

更多公式可参考[这里]