5.10 $\min\limits_{x\in\mathbb{R}^2} \ x_1$

s.t. $16-(x_1-4)^2-x_2^2\geq 0$ (a)

$x_1^2+(x_2-2)^2-4=0$ (b)

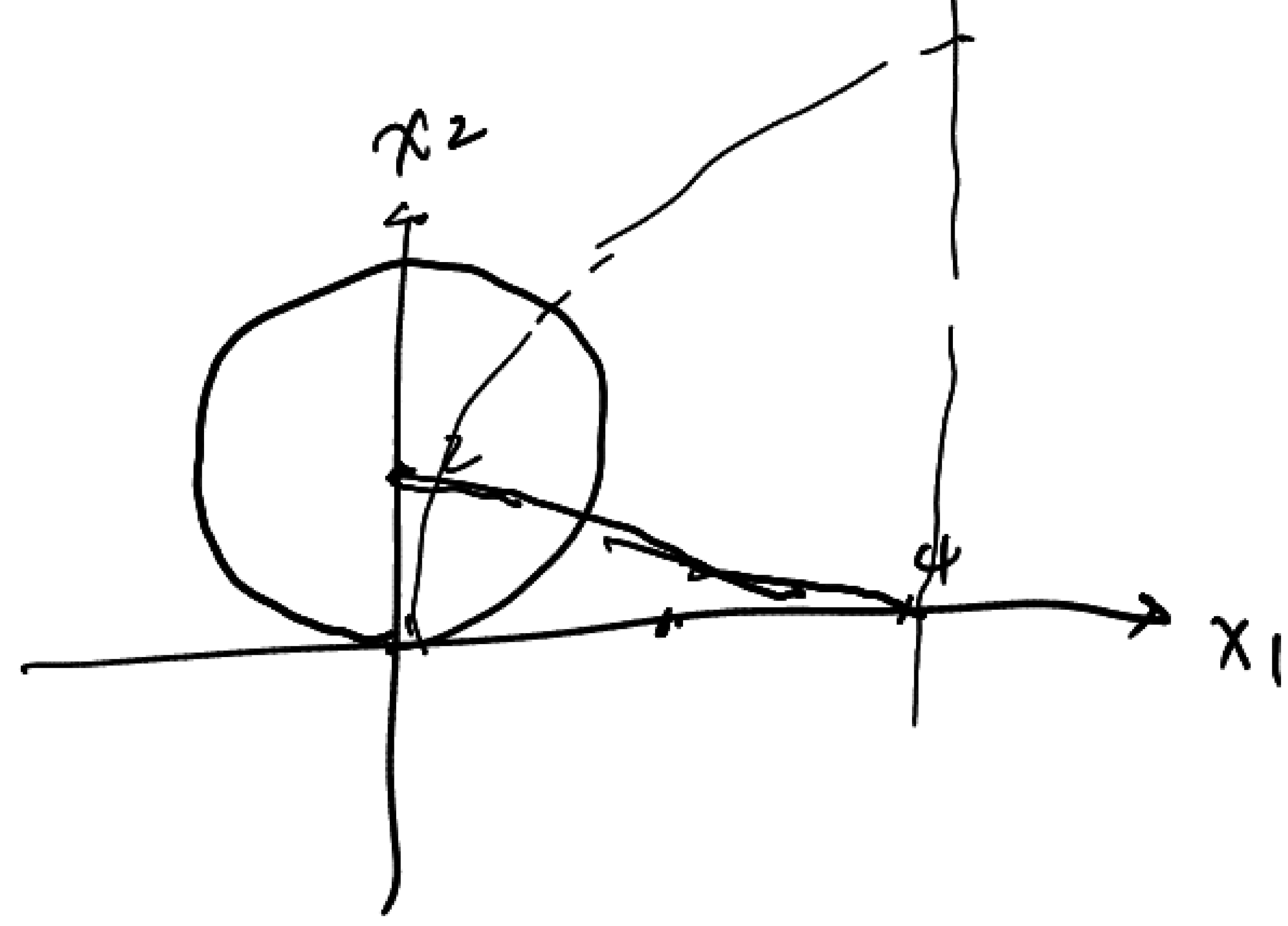
求KKT点，它们是否对应局部极小，鞍点，全局极小？

解．记 $c_1(x)=(x_1-4)^2+x_2^2-16$

$c_2(x)=x_1^2+(x_2-2)^2-4$

$\nabla c_1(x)=2((x_1-4),\ x_2)^T$ $\nabla c_2(x)=2(x_1,(x_2-2))^T$

$\nabla c_1(x)$ 和 $\nabla c_2(x)$ 线性相关 $\Leftrightarrow$ $x_1+2x_2=4$ (c) 但满足(b)(c)时 (x_1',x_2') 使 (a)不平等，不在积极集中 且 $\nabla c_1(x)\neq 0$ $\nabla c_2(x)\neq 0$ 故LICQ满足



考虑KKT条件

$$\mathcal{L}(x,\lambda)=x_1+\lambda_1((x_1-4)^2+x_2^2-16)+\lambda_2(x_1^2+(x_2-2)^2-4)$$

$$\begin{cases}1+\lambda_1(2(x_1-4))+\lambda_2\cdot 2x_1=0\\ \lambda_1\cdot 2x_2+\lambda_2\cdot(2(x_2-2))=0\\ x_1^2+(x_2-2)^2-4=0\\ \lambda_1\geq 0\\ \lambda_1((x_1-4)^2+x_2^2-16)=0\\ (x_1-4)^2+x_2^2\leq 16\end{cases}$$

解得 $(x^*,\lambda^*)=(0,0,\frac{1}{8},0)$, $(2,2,0,-\frac{1}{4})$, $(\frac{8}{5},\frac{16}{5},\frac{3}{40},-\frac{1}{5})$

$$\nabla_{xx}^2\mathcal{L}(x,\lambda)=\begin{bmatrix}2\lambda_1+2\lambda_2 & 0\\ 0 & 2\lambda_1+2\lambda_2\end{bmatrix}$$ 对 $(0,0,\frac{1}{8},0)$ 为正定阵，其余为负定

又 满足(a)(b)时 x 均有 $x_1\geq 0$

故 $(0,0,\frac{1}{8},0)$ 为全局极小 $(2,2,0,-\frac{1}{4})$ 及 $(\frac{8}{5},\frac{16}{5},\frac{3}{40},-\frac{1}{5})$ 不是极小也不是鞍点

5-16 SVM
$$\min_{x\in\mathbb{R}^n,\xi} \frac{1}{2}\|x\|_2^2+\mu\sum_{i=1}^m \xi_i$$

$$\text{s.t.}\quad b_i a_i^T x \geq 1-\xi_i \quad i=1,\dots m$$
$$\xi_i \geq 0,\ i=1,\dots m$$

其中 $\mu>0$ $b_i\in\mathbb{R}$, $a_i\in\mathbb{R}^n$, $i=1,\dots m$ 已知，求对偶问题

解 考虑 $\mathcal{L}(x,\xi,\lambda,t)=\frac{1}{2}\|x\|_2^2+\mu\sum_{i=1}^m \xi_i+\sum_{i=1}^m \lambda_i(-\xi_i)+\sum_{i=1}^m t_i(1-\xi_i-b_i a_i^T x)$

其中 $\lambda_i\geq 0$, $i=1,\dots m$, $t_i\geq 0$, $i=1,\dots m$

$$\min_{x\in\mathbb{R}^n,\xi\in\mathbb{R}^m}\mathcal{L}(x,\xi,\lambda,t)=\min_{x,\xi}\frac{1}{2}\|x\|_2^2+\sum_{i=1}^m(\mu-\lambda_i-t_i)\xi_i+\sum_{i=1}^m t_i$$
$$-\sum_{i=1}^m t_i b_i a_i^T x=\begin{cases}-\frac{1}{2}\sum_{i=1}^m\sum_{j=1}^m t_i t_j b_i b_j a_i^T a_j+\sum_{i=1}^m t_i & \text{if } \lambda_i+t_i=\mu,\ i=1,\dots m\\ -\infty & \text{else}\end{cases}$$

对偶问题是
$$\max_{\substack{\lambda\in\mathbb{R}^n\\ t\in\mathbb{R}^n}} -\frac{1}{2}\sum_{i=1}^m\sum_{j=1}^m t_i t_j b_i b_j a_i^T a_j+\sum_{i=1}^m t_i$$
$$\text{s.t.}\quad \lambda\geq 0$$
$$t\geq 0$$
$$\lambda_i+t_i=\mu \quad i=1,\dots m$$

等价于
$$\max_{t\in\mathbb{R}^n} -\frac{1}{2}\sum_{i=1}^m\sum_{j=1}^m t_i t_j b_i b_j a_i^T a_j+\sum_{i=1}^m t_i$$
$$\text{s.t.}\quad 0\leq t_i\leq\mu \quad i=1,\dots m$$

#