（1）对于一个最小二乘问题，令J代表其 Jacobian 矩阵，是一个$m\times n$的矩阵，$m\ge n$，证明：（a）J是列满秩的充要条件是J^TJ非奇异，(b)J是列满秩的充要条件是J^TJ正定。(10 分)

(a) "$\Rightarrow$" 当J列满秩时，假设J^TJ奇异，则$\exists x\neq 0$ 使得 $(J^TJ)x=0$

注意到 $0=x^TJ^TJx=\|Jx\|^2\Rightarrow Jx=0$，但J列满秩，若$Jx=0$只能$x=0$，矛盾，故假设不成立，原命题得证

"$\Leftarrow$" 当J^TJ非奇异时，若J非列满秩，则$\exists x\neq 0$ 使得 $Jx=0$，但这样一来 $(J^TJ)x=J^T(Jx)=0$，这与$|J^TJ|\neq 0$矛盾，故假设不成立，原命题得证

(b) 注意到 $\forall x\in\mathbb{R}^n$，$x^T(J^TJ)x=\|Jx\|^2\ge 0$，

J列满秩 $\Leftrightarrow$ $Jx=0$只有零解 $\Leftrightarrow$ $\forall x\neq 0$，$x^T(J^TJ)x=\|Jx\|^2>0$

$\Leftrightarrow$ J^TJ正定

（2）对于一个最小二乘问题，假设每个剩余函数r_j和其梯度均是 L-Lipschitz 连续的，

$$\left\|r_j(x)-r_j(\tilde{x})\right\|\le L\|x-\tilde{x}\|,\qquad \left\|\nabla r_j(x)-\nabla r_j(\tilde{x})\right\|\le L\|x-\tilde{x}\|$$

对于$j=1,2,\dots,m;\ x,\tilde{x}\in D$. 假设$r_j$在$D$上是有界的，既存在$M$使得$\left|r_j(x)\right|\le M$. 请找到雅各比矩阵$J$的 Lipschitz 常数，和目标函数梯度$\nabla f(x)$的 Lipschitz 常数. （14 分）

设 $J(x)=[\nabla r_1(x),\dots\nabla r_m(x)]^T\in\mathbb{R}^{m\times n}$，$r(x)=[r_1(x),\dots,r_m(x)]^T\in\mathbb{R}^n$

$\forall x,\tilde{x}\in D$，$\|J(x)-J(\tilde{x})\|^2=(J(x)-J(\tilde{x}))^T(J(x)-J(\tilde{x}))$

$$=J(x)^TJ(x)+J(\tilde{x})^TJ(\tilde{x})-2J(x)^TJ(\tilde{x})$$

$$=\sum_{i=1}^m(\nabla r_i(x)^T\nabla r_i(x))+\sum_{i=1}^m(\nabla r_i(\tilde{x})^T\nabla r_i(\tilde{x}))-\sum_{i=1}^m 2\nabla r_i(x)^T\nabla r_i(\tilde{x})$$

$$=\sum_{i=1}^m\|\nabla r_i(x)-\nabla r_i(\tilde{x})\|^2\le L^2m\|x-\tilde{x}\|^2$$

由于 $\|r_i(x)-r_j(\tilde{x})\|\le L\|x-\tilde{x}\|\Rightarrow$ 当$\|x-\tilde{x}\|$充分小时，$\frac{\|r_i(x)-r_i(\tilde{x})\|}{\|x-\tilde{x}\|}\to\|\nabla r_i(x)\|\le L$

$$\|\nabla f(x)-\nabla f(\tilde{x})\|=\left\|\sum_{i=1}^m\left(r_i(x)\nabla r_i(x)-r_i(\tilde{x})\nabla r_i(\tilde{x})\right)\right\|\le\sum_{i=1}^m\|r_i(x)\nabla r_i(x)-r_i(\tilde{x})\nabla r_i(\tilde{x})\|$$

$$=\sum_{i=1}^m\left\|(r_i(x)-r_i(\tilde{x}))\nabla r_i(x)+r_i(\tilde{x})(\nabla r_i(x)-\nabla r_i(\tilde{x}))\right\|$$

$$\le\sum_{i=1}^m\|\nabla r_i(x)\|\cdot L\|x-\tilde{x}\|+ML\|x-\tilde{x}\|$$

$$\le(mL^2+mML)\|x-\tilde{x}\|$$

因此J的 Lipschitz 常数为 $\sqrt{m}L$，$\nabla f(x)$的 Lipschitz 常数为 mL^2+mML

（3）对于带约束的优化问题$\min_{x\in\Omega} f(x)$，如果f是凸函数并且可行域Ω是凸集，请证明该问题的局部解必然是全局解，并且全局解的集合是凸集。（10 分）

(1) $\forall \varepsilon>0$，设x^*是$f(x)$在x^*的一个邻域$O(x^*,\varepsilon)$内的局部解，从而

$$f(x^*)\le f(x),\qquad \forall x\in O(x^*,\varepsilon)$$

$\forall x\in\Omega$，取$0<\lambda<\frac{\varepsilon}{|x-x^*|}$，作$x'=\lambda x+(1-\lambda)x^*$，则$|x'-x^*|=\lambda|x-x^*|<\varepsilon$

故$x'\in O(x^*,\varepsilon)$，由于Ω是凸集，易见$x'\in\Omega$，又因为f为凸函数，故

$$f(x^*)\le f(x')\le \lambda f(x)+(1-\lambda)f(x^*)\le \lambda f(x)+(1-\lambda)f(x)=f(x)$$

由x在Ω的任意性知

$$f(x^*)\le f(x),\qquad \forall x\in\Omega$$

这说明x^*也为问题的全局解

(2) 记全局解的集合为$S=\{x^*: f(x^*)\le f(x),\ x\in\Omega\}$

① 当$|S|=0$或$|S|=1$时，显然S为凸集

② 当$|S|\ge 2$时，任取$x_1^*, x_2^*\in S$，考虑$x_3^*=\lambda x_1^*+(1-\lambda)x_2^*$，$(\forall\lambda\in[0,1])$

$$f(x_3^*)\le \lambda f(x_1^*)+(1-\lambda)f(x_2^*)$$
$$\le \lambda f(x)+(1-\lambda)f(x)=f(x),\qquad \forall x\in\Omega$$

因此$x_3^*\in S$，由x_1^*, x_2^*的任意性知S为凸集

（4）考虑一个半空间$H=\{x\in R^n|a^{\mathrm T}x+\alpha\ge 0\}$ 其中$a\in R^n$ 并且$\alpha\in R$ 是给定的。请在H中找到一个x，使其具有最小的欧式范数。将上述问题表达为一个带约束的优化问题，然后采用约束优化理论的方法求解。（10 分）

$$\min \|x\|_2^2 \qquad \text{s.t.}\quad a^{T}x+\alpha\ge 0$$

构造$L(x,\lambda)=\|x\|_2^2-\lambda(a^Tx+\alpha)$，显然约束满足LICQ，故

$$\begin{cases}\nabla L_x=2x-\lambda a=0\\ \nabla L_\lambda=a^Tx+\alpha=0\\ \lambda\cdot(a^Tx+\alpha)=0\\ \lambda\ge 0\end{cases}$$

$\Rightarrow$ ① $\lambda=0$时，$x^*=0$，$x^*=0$处有最小的Euclidean范数，$\|x^*\|=0$

② $\lambda>0$时，解得 $x^*=\frac{-\alpha\cdot a}{a^Ta}$，$\lambda^*=\frac{-2\alpha}{a^Ta}$

由于$\lambda>0$，因此必然$\alpha<0$

可行方向集$F(x^*)=\{d: a^Td>0\}$。$\forall d\in F(x^*)$，

$$g(x^*)^Td=2x^{*T}d=-\frac{2\alpha(a^Td)}{a^Ta}>0$$

故$x^*=\frac{-\alpha\cdot a}{a^Ta}$是最优解，此时$\|x^*\|_2=\frac{-\alpha}{a^Ta}$

（5）在函数$y=\frac{1}{5}(x-1)^2$上找到离$(x,y)=(1,2)$欧氏距离最近的点。我们可以把上述问题转变为一个带约束的优化问题：

$$\min f(x,y)=(x-1)^2+(y-2)^2$$
$$s.t.\quad (x-1)^2=5y$$

(a) 找到所有的 KKT 点
(b) 哪个点才是问题的解？
(c) 我们可以直接把约束条件代入目标函数，消去变量x，从而获得一个无约束优化问题；请证明该无约束优化问题的解不是原始问题的解。（想一想为什么？）（14 分）

(a) $L(x,\lambda)=(x-1)^2+(y-2)^2+\lambda((x-1)^2-5y)$

$$\begin{cases}\nabla L_x = 2(x-1)+\lambda(2(x-1)) = (2+2\lambda)(x-1)=0\\ \nabla L_y = 2(y-2)-5\lambda=0\\ \lambda^*((x-1)^2-5y)=0\\ \nabla L_\lambda = (x-1)^2-5y=0\end{cases} \Rightarrow \begin{cases}x=1\\ y=0\\ \lambda=-\frac{4}{5}\end{cases}$$

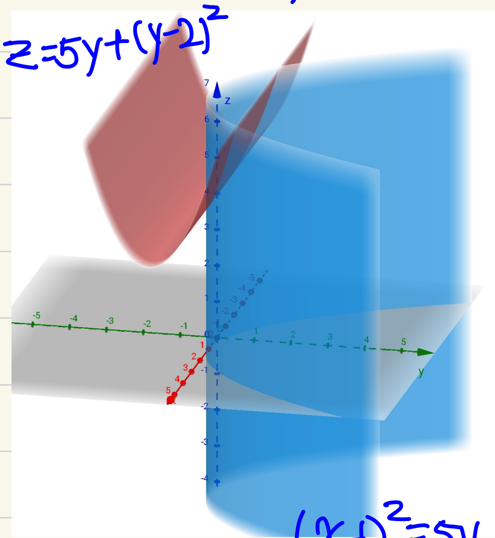
故全部KKT点为 $((x^*,y^*),\lambda^*)=((1,0),-\frac{4}{5})$

(b) $\nabla^2 L=\begin{pmatrix}2+2\lambda & 0\\ 0 & 2\end{pmatrix}=\begin{pmatrix}\frac{2}{5} & \\ & 2\end{pmatrix}$ 正定 $\Rightarrow (x^*,y^*)=(1,0)$ 就是问题的最优解

此时 $f(x^*,y^*)=4$

(c) 若直接代入，则 $f(x,y)=5y+(y-2)^2 \Rightarrow \frac{\partial f}{\partial y}=5+2(y-2)=0 \Rightarrow y=-\frac{1}{2}$

但 $(x-1)^2=5y=-\frac{5}{2}$ 无解！无约束问题的解根本不存在，不是原始问题的解



因为原本的约束隐含 $y=\frac{1}{5}(x-1)^2\geq 0$，代换后的无约束优化问题解得的 $y^*<0$，不满足这一约束

（6）将下列优化问题转化为无约束优化问题直接进行求解，同时用外点罚函数法求解该约束优化问题，（14 分）

$$\min f(x) = x_1 + x_2$$
$$s.t.\quad x_2 - x_1^2 = 0$$

无约束优化问题：$f(x_1,x_2) = x_1 + x_2 = x_1 + x_1^2 \Rightarrow \frac{\partial f}{\partial x_1} = 1 + 2x_1 = 0 \Rightarrow \begin{cases} x_1 = -\frac{1}{2} \\ x_2 = \frac{1}{4} \end{cases}$

$\Rightarrow$ 最优解为 $(x_1^*, x_2^*) = (-1/2, 1/4)$，此时 $f(x_1^*, x_2^*) = -1/4$

外点罚函数：构造罚函数 $P(x,\sigma) = x_1 + x_2 + \frac{1}{2}\sigma(x_2 - x_1^2)^2$

$$\begin{cases} \partial P/\partial x_1 = 1 + \frac{1}{2}\sigma \cdot 2(x_2 - x_1^2)(-2x_1) = 1 - 2\sigma(x_2 - x_1^2)x_1 = 0 \\ \partial P/\partial x_2 = 1 + \frac{1}{2}\sigma \cdot 2(x_2 - x_1^2)\cdot 1 = 1 + \sigma(x_2 - x_1^2) = 0 \end{cases} \Rightarrow \begin{cases} x_1 = -\frac{1}{2} \\ x_2 = \frac{1}{4} - \frac{1}{\sigma} \end{cases}$$

$\Rightarrow$ 最优解为 $(x_1^*(\sigma), x_2^*(\sigma)) = (-1/2,\ 1/4 - 1/\sigma) \to (-1/2,\ 1/4)$，$\sigma \to \infty$

此时 $f(x_1^*, x_2^*) = -1/4$

（7）用对数障碍函数法求解如下优化问题：（14 分）

$$\min f(x) = x_1^2 + 2x_2^2$$
$$s.t.\quad x_1 + x_2 - 1 \geq 0$$

构造对数障碍函数 $B_L(x,\mu) = x_1^2 + 2x_2^2 - \mu\ln(x_1 + x_2 - 1)$

$$\begin{cases} \partial B_L/\partial x_1 = 2x_1 - \mu\cdot\frac{1}{x_1 + x_2 - 1} = 0 \\ \partial B_L/\partial x_2 = 4x_2 - \mu\cdot\frac{1}{x_1 + x_2 - 1} = 0 \end{cases} \Rightarrow \begin{cases} x_1 = \frac{1 \pm \sqrt{1+3\mu}}{3} \\ x_2 = \frac{1 \pm \sqrt{1+3\mu}}{6} \end{cases}$$

又 $x_1 + x_2 \geq 1 \Rightarrow$ 取 $\begin{cases} x_1 = \frac{1 + \sqrt{1+3\mu}}{3} \\ x_2 = \frac{1 + \sqrt{1+3\mu}}{6} \end{cases} \Rightarrow$ 最优解为 $(x_1^*, x_2^*) = \left(\frac{1+\sqrt{1+3\mu}}{3}, \frac{1+\sqrt{1+3\mu}}{6}\right)$

$\to \left(\frac{2}{3}, \frac{1}{3}\right)$，$\mu \to 0$

此时 $f(x_1^*, x_2^*) = 2/3$

（8）采用增广拉格朗日函数方法求解

$$\min f(x) = 2x_1^2 + x_2^2 - 2x_1x_2$$
$$s.t.\quad x_1 + x_2 - 1 = 0$$

增广拉格朗日函数方法中的第一步$\lambda^1 = 1$，σ恒定取 2.（14 分）

构造增广 Lagrangian 函数：$\Phi(x,\lambda^k) = 2x_1^2 + x_2^2 - 2x_1x_2 - \lambda^k(x_1 + x_2 - 1) + \frac{1}{2}\cdot 2(x_1 + x_2 - 1)^2$

$= 2x_1^2 + x_2^2 - 2x_1x_2 - \lambda^k(x_1 + x_2 - 1) + (x_1 + x_2 - 1)^2$

$$\begin{cases} \partial\Phi/\partial x_1 = 4x_1 - 2x_2 - \lambda^k + 2(x_1 + x_2 - 1) = 0 \\ \partial\Phi/\partial x_2 = 2x_2 - 2x_1 - \lambda^k + 2(x_1 + x_2 - 1) = 0 \end{cases} \Rightarrow \begin{cases} x_1^* = \frac{1}{6}(\lambda^k + 2) \\ x_2^* = \frac{1}{4}(\lambda^k + 2) \end{cases}$$

其中 $\lambda^{k+1} = \lambda^k - 2([x_1^{(k)}]^* + [x_2^{(k)}]^* - 1) = \lambda^k - 2\left(\frac{5}{12}(\lambda^k + 2) - 1\right) = \frac{1}{6}\lambda^k + \frac{1}{3}$

解不动点方程 $x = \frac{1}{6}x + \frac{1}{3} \Rightarrow x = \frac{2}{5}$，故 $\lambda^{k+1} - \frac{2}{5} = \frac{1}{6}(\lambda^k - \frac{2}{5}) \Rightarrow \lambda^k = \left(\frac{1}{6}\right)^k \cdot \frac{3}{5} + \frac{2}{5}$

当 $k \to \infty$ 时，$\lambda^k \to \frac{2}{5}$，那么 $(x_1^*, x_2^*) \to \left(\frac{2}{5}, \frac{3}{5}\right)$，此时 $f(x_1^*, x_2^*) = 1/5$