Sheet 5 118010350

A5.1 (a) Since $f_\beta(x) = \frac{1}{2}\|x-b\|^2 + \frac{\beta}{2}(\sum_{i=1}^{n} x_i)^2$, $x \in \mathbb{R}^n$

$= \frac{1}{2}\sum_{i=1}^{n}(x_i-b_i)^2 + \frac{\beta}{2}(\sum_{i=1}^{n} x_i)^2$, $i=1,\dots,n$

Then

$$\nabla f_\beta(x) = \begin{bmatrix} (x_1-b_1)+\beta(x_1+\dots+x_n) \\ (x_2-b_2)+\beta(x_1+\dots+x_n) \\ \dots \\ (x_n-b_n)+\beta(x_1+\dots+x_n) \end{bmatrix}$$

$$\nabla^2 f_\beta(x) = \begin{bmatrix} \beta+1 & \beta & \cdots & \beta \\ \beta & \beta+1 & \cdots & \beta \\ \cdots & \cdots & \ddots & \cdots \\ \beta & \beta & \cdots & \beta+1 \end{bmatrix}$$

(b) Let $\nabla f_\beta(x) = 0$, and suppose $m = x_1+\dots+x_n$

Then we have $x_1 - b_1 + \beta m = 0$

$x_2 - b_2 + \beta m = 0$

$\dots$

$x_n - b_n + \beta m = 0$

Thus we can get

$$\left.\begin{aligned} x_1 &= b_1 - \beta m \\ x_2 &= b_2 - \beta m \\ &\dots \\ x_n &= b_n - \beta m \end{aligned}\right\} (*)$$

Since $x_1 + \dots + x_n = (b_1+\dots+b_n) - n\beta m$

$m = (b_1+\dots+b_n) - n\beta m$

$(n\beta+1)m = b_1+\dots+b_n$

$m = \frac{b_1+\dots+b_n}{n\beta+1}$

we can substitute m in equations $(*)$,

Thus we know that solution $x_1, \dots, x_n$ are unique.

So f_β has a unique stationary point x^*_β, which is

$$x^*_\beta = \begin{bmatrix} b_1 - \frac{\beta(b_1+\dots+b_n)}{n\beta+1} \\ b_2 - \frac{\beta(b_1+\dots+b_n)}{n\beta+1} \\ \dots \\ b_n - \frac{\beta(b_1+\dots+b_n)}{n\beta+1} \end{bmatrix}$$

Consider Hessian matrix $\nabla^2 f_\beta(x)$, since $\beta \geq 0$.

After applying elementary matrix to $\nabla^2 f_\beta(x)$,

we can get final matrix

$$A = \begin{bmatrix} n\beta+1 & \beta & \cdots & \beta \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}$$

Since $\det(A) > 0$,

and $\det(\nabla^2 f_\beta(x)) = \det(A)$

then $\det(\nabla^2 f_\beta(x)) > 0$.

Change $n = 1, \dots, n$, we can get all the diterminants of leading sub-matrix are positive, thus $\nabla^2 f_\beta(x)$ is positive definite, and x^*_β is a local minimizer.

(c) when $\beta \to \infty$, then $x^* = \lim_{\beta\to\infty} x^*_\beta$

Since

$$\frac{\beta(b_1+\dots+b_n)}{n\beta+1} = \frac{\beta(b_1+\dots+b_n) + \frac{b_1+\dots+b_n}{n} - \frac{b_1+\dots+b_n}{n}}{n\beta+1}$$

$$= \frac{b_1+\dots+b_n}{n} - \frac{\frac{b_1+\dots+b_n}{n}}{n\beta+1}$$

Then $\lim_{\beta\to\infty} \frac{\beta(b_1+\dots+b_n)}{n\beta+1} = \frac{b_1+\dots+b_n}{n}$

Thus,

$$x^* = \begin{bmatrix} b_1 - \frac{b_1+\dots+b_n}{n} \\ b_2 - \frac{b_1+\dots+b_n}{n} \\ \dots \\ b_n - \frac{b_1+\dots+b_n}{n} \end{bmatrix}$$

Since $\sum_{i=1}^{n} x_i^* = (b_1+\dots+b_n) - n \cdot \frac{b_1+\dots+b_n}{n} = 0$.

Then x^* satisfies the constraint $\mathbb{1}^T x^* = \sum_{i=1}^{n} x_i^* = 0$.

(d) Consider the constrained nonliear program:

$$\min_x \frac{1}{2}\|x-b\|^2$$
$$\text{s.t. } \mathbb{1}^T x = 0.$$

Define the feasible points $S = \{x : \mathbb{1}^T x = 0\}$

Since here $h(x) = x_1 + x_2 + \dots + x_n = \sum_{i=1}^{n} x_i$

Then

$$\nabla h(x) = \begin{bmatrix} 1 \\ 1 \\ \dots \\ 1 \end{bmatrix}$$, and it is the only gradient.

Thus, LICQ is generally satisfied at feasible points.

Consider the KKT-conditions for x^*.

① Main Condition: $\nabla f(x^*) + u \cdot \nabla h(x^*) = 0$.

That is

$$\begin{bmatrix} x_1^* - b_1 \\ x_2^* - b_2 \\ \dots \\ x_n^* - b_n \end{bmatrix} + u \cdot \begin{bmatrix} 1 \\ 1 \\ \dots \\ 1 \end{bmatrix} = 0$$

we can get $u = \frac{b_1+\dots+b_n}{n}$

② Dual Feasibility ③ Complementarity

No need to check.

④ Primal Feasibility: $h(x^*) = 0$

Since $\mathbb{1}^T x^* = 0$, then it is satisfied.

Then consider the second order conditions.

$\nabla^2_{xx} L(x^*, u) = \nabla^2 f(x^*) + u \cdot \nabla^2 h(x^*)$.

and

$$\nabla^2 f(x^*) = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ 0 & 0 & \cdots & 1 \end{bmatrix}, \quad \nabla^2 h(x^*) = \begin{bmatrix} 0 & \cdots & 0 \\ 0 & \cdots & 0 \\ 0 & \cdots & 0 \end{bmatrix}$$

And $C(x^*) = \{d \in \mathbb{R}^n : \nabla f(x^*)^T d = 0, \nabla h(x^*)^T d = 0\}$.

which is equivalent to the feasible points set

That is $C(x^*) = \{d \in \mathbb{R}^n : \mathbb{1}^T d = 0\}$

Then we can check for $\forall d \in C(x^*) \setminus \{0\}$, $d = \begin{bmatrix} d_1 \\ d_2 \\ \dots \\ d_n \end{bmatrix}$

$d^T \nabla^2_{xx} L(x^*, u) d = d_1^2 + d_2^2 + \dots + d_n^2 > 0$.

Then, $\nabla^2_{xx} L(x^*, u)$ is positive definite on $C(x^*) \setminus \{0\}$.

Thus, x^* is the unique local solution of this problem.

A5.2. Cosider the constrained nonlinear program:

$$\min_x f(x) = x_1^2 + x_2^2 + x_3^2 + x_1x_2 + x_2x_3 - 2x_1 - 5x_2 - 6x_3$$
$$\text{s.t. } x_1 + x_2 + x_3 \leq 1, \quad x_1 - x_2^2 = 0$$

(a). Derive the KKT-conditions.

① Main Condition: $\nabla f(x^*) + \lambda \cdot \nabla g(x^*) + u \cdot \nabla h(x^*) = 0$

$$\nabla f(x) = \begin{bmatrix} 2x_1 + x_2 - 2 \\ 2x_2 + x_1 + x_3 - 5 \\ 2x_3 + x_2 - 6 \end{bmatrix}, \quad \nabla g(x) = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \quad \nabla h(x) = \begin{bmatrix} 1 \\ -2x_2 \\ 0 \end{bmatrix}$$

② Dual Feasibility: $\lambda \geq 0$.

③ Complementarity: $\lambda \cdot g(x^*) = 0$.

④ Primal Feasibility: $g(x^*) \leq 0$, $h(x^*) = 0$.

(b) Let $x^* = [0\ 0\ 1]^T$.

Consider the KKT-conditions.

① Main Condition: $\nabla f(x^*) + \lambda \nabla g(x^*) + u \cdot \nabla h(x^*) = 0$

which is

$$\begin{bmatrix} -2 \\ -4 \\ -4 \end{bmatrix} + \lambda \cdot \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + u \cdot \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

we can get $\lambda = 4$, $u = -2$.

② Dual Feasibility: $\lambda = 4 \geq 0$.

③ Complementarity: $\lambda \cdot g(x^*) = 4 \cdot (0+0+1-1) = 0$.

④ Primal Feasibility: $g(x^*) \leq 0$, $h(x^*) = 0$.

Then consider the second order conditions.

$\nabla^2_{xx} L(x^*, \lambda, u) = \nabla^2 f(x^*) + \lambda \cdot \nabla^2 g(x^*) + u \cdot \nabla^2 h(x^*)$

and

$$\nabla^2 f(x^*) = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 2 \end{bmatrix}, \quad \nabla^2 g(x^*) = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \nabla^2 h(x^*) = \begin{bmatrix} 0 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Then $C(x^*) = \{d \in \mathbb{R}^3 : \nabla f(x^*)^T d = 0, \nabla g(x^*)^T d \leq 0, \nabla h(x^*)^T d = 0\}$.

Suppose $d = [d_1\ d_2\ d_3]^T$, then we get

$2d_1 + 4d_2 + 4d_3 = 0$, $d_1 + d_2 + d_3 \leq 0$, $d_1 = 0$.

Thus we can assume $d = [0\ t\ -t]^T$.

we can check for $\forall d \in C(x^*) \setminus \{0\}$

$$d^T \nabla^2_{xx} L(x^*, \lambda, u) d = d^T \begin{bmatrix} 2 & 1 & 0 \\ 1 & 6 & 1 \\ 0 & 1 & 2 \end{bmatrix} d$$

$= 6t^2 > 0$.

Then $\nabla^2_{xx} L(x^*, \lambda, u)$ is positive definite on $C(x^*) \setminus \{0\}$.

Thus $x^* = [0\ 0\ 1]^T$ is a strict local minimum of this problem.

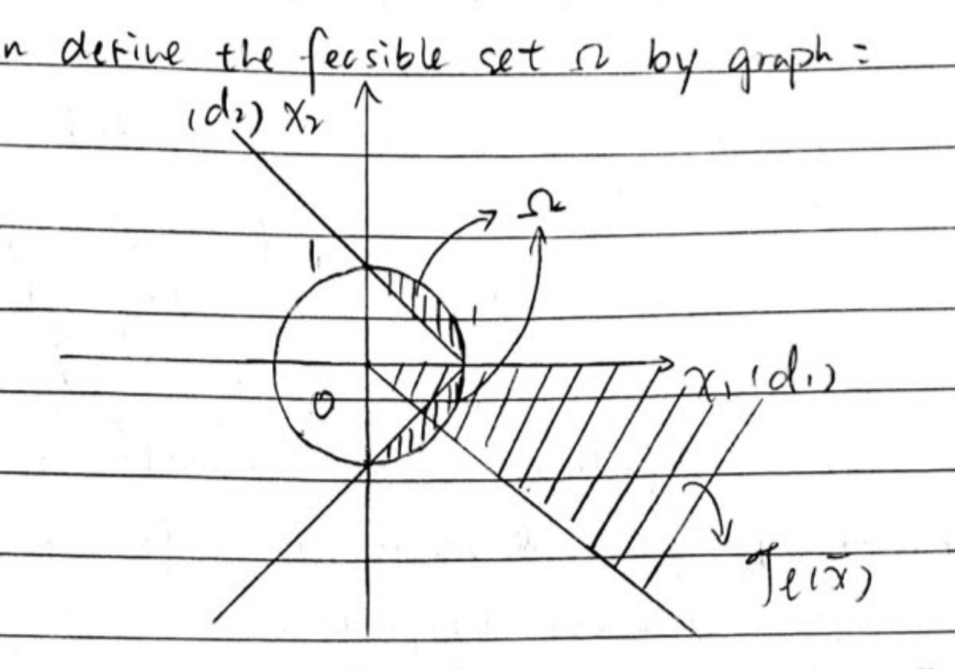
A5.3. Consider the constrained nonlinear program:

$$\min_x x_2^2 - 2x_1.$$
$$\text{s.t. } x_1^2 + x_2^2 - 1 \leq 0, \ (x_1-1)^2 - x_2^2 \leq 0.$$

And $\bar{x} = [0\ 1]^T$

(a) Since $\Omega = \{x \in \mathbb{R}^2 : g_1(x) \leq 0, g_2(x) \leq 0\}$.

We can define the feasible set Ω by graph:



(b) Since $\mathcal{A}(\bar{x}) = \{i : g_i(\bar{x}) = 0\}$, then we check $g_1(\bar{x})$ and $g_2(\bar{x})$.

$g_1(\bar{x}) = 0^2 + 1^2 - 1 = 0$, $g_2(\bar{x}) = (0-1)^2 - 1^2 = 0$. Then $\mathcal{A}(\bar{x}) = \{1, 2\}$.

Since $T_\ell(\bar{x}) = \{d : \nabla g_i(\bar{x})^T d \leq 0, \forall i \in \mathcal{A}(\bar{x})\}$.

and

$$\nabla g_1(x) = \begin{bmatrix} 2x_1 \\ 2x_2 \end{bmatrix}, \quad \nabla g_2(x) = \begin{bmatrix} 2x_1 - 2 \\ -2x_2 \end{bmatrix}$$

and

$$\nabla g_1(\bar{x}) \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad \nabla g_2(\bar{x}) = \begin{bmatrix} -2 \\ -2 \end{bmatrix}$$

Suppose $d = [d_1\ d_2]^T$, then we can get

$d_2 \leq 0$, $-2d_1 - 2d_2 \leq 0$

Thus we can assume $d = [d_1\ d_2]^T$, $d_2 \leq 0$, $d_1 \geq -d_2 \geq 0$

(C) Since $f = x_2^2 - 2x_1$ is a continuous function on $\mathbb{R}^2$.

and the feasible set Ω is a bounded, closed and nonempty set. Then by Weierstraß Theorem. f attains a global maximum and global minimum on the set Ω.

Thus, this problem has an optimal solution.

(d) Assume the LICQ holds at all feasible points other than the KKT-points.

Consider the KKT-conditions.

① Main Condition: $\nabla f(x^*) + \lambda_1 \nabla g_1(x^*) + \lambda_2 \nabla g_2(x^*) = 0$

we can get

$$\nabla f(x) = \begin{bmatrix} -2 \\ 2x_2 \end{bmatrix}, \quad \nabla g_1(x) = \begin{bmatrix} 2x_1 \\ 2x_2 \end{bmatrix}, \quad \nabla g_2(x) = \begin{bmatrix} 2x_1 - 2 \\ -2x_2 \end{bmatrix}$$

then

$$\begin{bmatrix} -2 \\ 2x_2 \end{bmatrix} + \lambda_1 \begin{bmatrix} 2x_1 \\ 2x_2 \end{bmatrix} + \lambda_2 \begin{bmatrix} 2x_1 - 2 \\ -2x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Case 1: $\lambda_1 = 0$, $\lambda_2 = 0$, impossible case

Case 2: $\lambda_1 = 0$, $\lambda_2 > 0$, then $(x_1-1)^2 - x_2^2 = 0$

Solve the equation set, impossible case

Case 3: $\lambda_1 > 0$, $\lambda_2 = 0$, then $x_1^2 + x_2^2 - 1 = 0$

solve the equation set, get $x_1 = 1$, $x_2 = 0$, $\lambda_1 = 1$, $\lambda_2 = 0$.

Case 4: $\lambda_1 > 0$, $\lambda_2 \geq 0$, then $(x_1-1)^2 - x_2^2 = 0$ and $x_1^2 + x_2^2 - 1 = 0$.

get $x_1 = 0$, $x_2 = 1$, impossible

$x_1 = 0$, $x_2 = -1$ impossible

$x_1 = 1$, $x_2 = 0$, LICQ not holds.

So we have only case 2, 3 and we can check LICQ both holds in these two cases.

② Dual Feasibility ③ Complementarity ④ Primal Feasibility

all holds for points in case 2.

Thus we get KKT points: $x_1 = 1$, $x_2 = 0$,

with $\lambda_1 = 1$, $\lambda_2 = 0$.

Then we consider second order conditions.

~~let $x^* = [0\ -1]^T$, $\lambda_1 = 0$, $\lambda_2 = 1$.~~

~~$\nabla^2_{xx} L(x^*, \lambda_1, \lambda_2) = \nabla^2 f(x^*) + \lambda_1 \nabla^2 g_1(x^*) + \lambda_2 \nabla^2 g_2(x^*)$~~

~~$\nabla^2 f(x^*) = \begin{bmatrix} 0 & 0 \\ 0 & 2 \end{bmatrix}$, $\nabla^2 g_2(\bar{x}) = \begin{bmatrix} 2 & 0 \\ 0 & -2 \end{bmatrix}$~~

~~Then $C(x^*) = \{d \in \mathbb{R}^2 : \nabla f(x^*)^T d = 0, \nabla g_2(x^*)^T d \leq 0\}$~~

~~Suppose $d = [d_1\ d_2]^T$, then we get~~

~~$-2d_1 - 2d_2 = 0$, $-2d_1 + 2d_2 \leq 0$.~~

~~Then we assume $d = [t\ -t]^T$, $t \geq 0$.~~

~~we can check for $\forall d \in C(x^*) \setminus \{0\}$,~~

~~$d^T \nabla^2_{xx} L(x^*, \lambda_1, \lambda_2) d = d^T \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix} d$~~

~~$= 2t^2 > 0$.~~

~~Then $\nabla^2_{xx} L(x^*, \lambda_1, \lambda_2)$ is positive definite on $C(x^*) \setminus \{0\}$.~~

~~Thus $x^* = [0\ -1]^T$ is a strict local minimizer~~

⇒ let $x^* = [1\ 0]^T$, $\lambda_1 = 1$, $\lambda_2 = 0$.

$\nabla^2_{xx} L(x^*, \lambda_1, \lambda_2) = \nabla^2 f(x^*) + \lambda_1 \nabla^2 g_1(x^*) + \lambda_2 \nabla^2 g_2(x^*)$

$$\nabla^2 f(x^*) = \begin{bmatrix} 0 & 0 \\ 0 & 2 \end{bmatrix}, \quad \nabla^2 g_1(x^*) = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$$

Then $C(x^*) = \{d \in \mathbb{R}^2 : \nabla f(x^*)^T d = 0, \nabla g_1(x^*)^T d \leq 0\}$

Suppose $d = [d_1\ d_2]^T$, then we get

$-2d_1 = 0$, $d_1 \leq 0$.

Then we assume $d = [0\ t]^T$. $t \in \mathbb{R}$.

we can check for $\forall d \in C(x^*) \setminus \{0\}$.

$$d^T \nabla^2_{xx} L(x^*, \lambda_1, \lambda_2) d = d^T \begin{bmatrix} 2 & 0 \\ 0 & 4 \end{bmatrix} d$$

$= 4t^2 > 0$.

Then $\nabla^2_{xx} L(x^*, \lambda_1, \lambda_2)$ is positive definite on $C(x^*) \setminus \{0\}$

Thus, $x^* = [1\ 0]^T$ is a strict local minimizer. $f = -2$

we also check f value for other endpoints.

when $x_1 = 0$, $x_2 = 1$, $f = 1$

$x_1 = 0$, $x_2 = -1$, $f = 1$

Thus, $x^* = [1\ 0]^T$ is also a strict global minimizer