11.1 Inequality constrained minimization problem

$$\boxed{\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & f_i(x)\le 0 \quad i=1,\cdots,m \\ & Ax=b \end{array}}$$ Q1

$A\in R^{p\times n}$, $p<n$, $\mathrm{rank}(A)=p$ $\qquad D=\bigcap_{i=0}^{m}\mathrm{dom} f_i \bigcap_{i=1}^{p}\mathrm{dom}\, h_i$

assum the problem is strictly feasible (即 $\exists\, x\in D$ such that $Ax=b$, $f_i(x)<0$ $i=1,\cdots,m$)

$\Rightarrow$ slater's constraints qualification holds

$\Rightarrow p^*=d^*$ $\qquad$ satisfy KKT 的提是 strong duality

$\Rightarrow$ exist dual optimal $\lambda^*\in R^m$, $v^*\in R^n$, primal optimal $x^*\in R^n$ satisfies KKT

$$\boxed{\begin{array}{c} Ax^*=b\ ,\ f_i(x^*)\le 0\ ,\ i=1,\cdots,m \\ \lambda^*\ge 0 \\ \nabla f_0(x^*)+\sum_{i=1}^{m}\lambda_i^*\nabla f_i(x^*)+A^Tv^*=0 \\ \lambda_i^* f_i(x^*)=0 \quad i=1,\cdots,m \end{array}}$$ Q2

$\leftarrow \frac{\partial L(x,\lambda^*,v^*)}{\partial x}\Big|_{x=x^*}=0$

求解 optimization problem with linear equality and inequality constraints 的方法叫 interior-point method

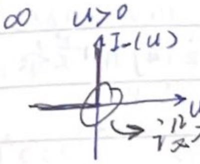
$\hookrightarrow$ a special method 是 the barrier method

11.2 logarithmic barrier function and central path

目标：将有 inequality constraints 的问题近似为仅有 equality constraints 的问题，之后可以 apply Newton's Method 求解

$\hookrightarrow$ 尝试：用 indicator function： $\quad I_-(u)=\begin{cases}0 & u\le 0\\ \infty & u>0\end{cases}$

$$\boxed{\begin{array}{ll} \text{minimize} & f_0(x)+\sum_{i=1}^{m} I_-(f_i(x)) \\ \text{subject to} & Ax=b \end{array}}$$ Q3



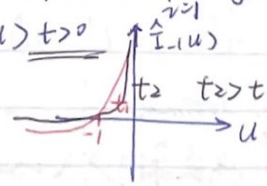
$\hookrightarrow$ 没有导数

not differentiable

$\downarrow$ 改进（用 logarithmic barrier 估计 indicator function）

logrithmic barrier: $\phi(x)=-\sum_{i=1}^{m}\log(-f_i(x))$

$\hat{I}_-(u)=-\frac{1}{t}\log(-u)$ $t>0$



$t_2>t_1$ $t\to\infty$ 时 $\hat{I}_-(u)\to I_-(u)$

$\nabla\phi(x)=\sum_{i=1}^{m}\frac{1}{-f_i(x)}\nabla f_i(x)$

$\nabla^2\phi(x)=\sum_{i=1}^{m}\frac{1}{f_i(x)^2}\nabla f_i(x)\nabla f_i(x)^T+\sum_{i=1}^{m}\frac{1}{-f_i(x)}\nabla^2 f_i(x)$

$\hookrightarrow \sum_{i=1}^{m}\hat{I}_-(f_i(x))$

$=\frac{1}{t}\phi(x)$

$$\boxed{\begin{array}{ll} \text{minimize} & f_0(x)+\frac{1}{t}\phi(x) \\ \text{subject to} & Ax=b \end{array}}$$ Q4 $\rightarrow$

$$\boxed{\begin{array}{ll} \text{minimize} & t f_0(x)+\phi(x) \\ \text{subject to} & Ax=b \end{array}}$$ Q5

for each $t>0$, Q_b has a unique solution $\underline{x^*(t)}$

central path: $\{x^*(t) \mid t>0\}$

central point: $x^*(t)$, $t>0$

有: $A(x^*(t))=b$, $f_i(x^*(t))<0$, $i=1,\dots,m$ 　$L_{Q_b}(x,\hat{v}) = tf_0(x)+\phi(x)+A^T$ $v^T(Ax-b)$ 　$x^*(t)$ minimize $L(x,\hat{v})$

exist a $\hat{v}\in R^p$ such that

central path condition:

$t\nabla f_0(x^*(t))+\nabla\phi(x)+A^T\hat{v}=0$ 　$\frac{\partial L_{Q_b}(x,\hat{v})}{\partial x}\Big|_{x=x^*(t)}=0$

$t\nabla f_0(x^*(t))+\sum_{i=1}^m \frac{1}{-f_i(x^*(t))}\nabla f_i(x^*(t))+A^T\hat{v}=0$

Lagrangian function for Q_1 is:

$L(x,\lambda^*(t),v^*(t)) = f_0(x)+\sum_{i=1}^m \lambda_i^*(t) f_i(x) + v^*(t)^T(Ax-b)$

dual function $g(\lambda^*(t),v^*(t)) = \min_x L(x,\lambda^*(t),v^*(t))$

$x^*(t)$ minimizes $L(x,\lambda^*(t),v^*(t))$ $= f_0(x^*(t))+\sum_{i=1}^m \lambda_i^*(t) f_i(x^*(t))+v^*(t)^T(Ax^*(t)-b)$

理解: 由KKT condition of Q_b, 有 $Ax^*(t)=b$ ①, $f_i(x^*(t))<0$ ②

且 $t\nabla f_0(x^*(t))+\sum_{i=1}^m \frac{1}{-f_i(x^*(t))}\nabla f_i(x^*(t))+A^T\hat{v}=0$ ③

而③的意思是 $\nabla L_{Q_b}(x,\hat{v})\big|_{x=x^*(t)}=0$

巧妙的地RP $x^*(t)$ minimizes $L_{Q_b}(x,\hat{v})$

亦即 $\Rightarrow x^*(t)$ minimizes $\frac{L_{Q_b}(x,\hat{v})}{t} = f_0(x)+\frac{1}{t}\phi(x)+\frac{\hat{v}^T}{t}(Ax-b)$

过构建 $\hat{v}$、λ^*与v^*之间的关系

使 $x^*(t)$ 同时 minimize $L_{Q_b}(x,\hat{v})$ $L_{Q_1}(x,\lambda^*,v^*)$

$\Rightarrow \frac{\partial \frac{L_{Q_b}(x,\hat{v})}{t}}{\partial x}\Big|_{x=x^*(t)} = \nabla f_0(x^*(t))+\sum_{i=1}^m \frac{1}{-tf_i(x^*(t))}\nabla f_i(x^*(t))+A^T\frac{\hat{v}}{t}=0$ 　$(t>0, f_i(x)<0)$ 　$\lambda_i^*(t)$ 　$v^*(t)$

$~~= f_0(x)+\sum_{i=1}^m \frac{1}{-tf_i(x)}~~$

如果令 $\lambda_i^*(t) = -\frac{1}{tf_i(t)}>0$ 　$v^*(t)=\frac{\hat{v}}{t}$

则有 $\nabla f_0(x^*(t))+\sum_{i=1}^m \lambda_i^*(t)\nabla f_i(x^*(t))+A^Tv^*(t)=0$ ④ 　$v^*(t)^T(Ax-b)$

② 令为 dual optimal for Q_b

联示 $L_{Q_1}(x^*(t),\lambda^*(t),v^*(t)) = f_0(x^*(t))+\sum_{i=1}^m \lambda_i^*(t) f_i(x^*(t))+$ ~~$A^Tv^*(t)$~~ ⑤

$\Rightarrow \frac{\partial L_{Q_1}(x,\lambda^*(t),v^*(t))}{\partial x}\Big|_{x=x^*(t)} = \frac{\partial ⑤}{\partial x}\Big|_{x=x^*(t)} = $ LHS of ④ $=0$

$(\lambda^*(t), v^*(t))$ are not ~~necessarily~~ dual optimal for Q_1) $\rightarrow$ 由Q_b的KKT可得 $Ax^*(t)=b$, 将该条件代入L_{Q_1}可得

$\Rightarrow x^*(t)$ minimizes $L_{Q_1}(x,\lambda^*(t),v^*(t))$ 其中 $\lambda_i^*(t)=\frac{1}{-tf_i(x^*(t))}$ 　$v^*(t)=\frac{\hat{v}}{t}$

$L_{Q_1}(x^*(t),\lambda^*(t),v^*(t)) = f_0(x^*(t)) - \frac{m}{t}$

$g(\lambda^*(t),v^*(t)) = L_{Q_1}(x^*(t),\lambda^*(t),v^*(t)) = f_0(x^*(t))+\sum_{i=1}^m \frac{1}{-tf_i(x^*(t))}f_i(x^*(t))+v^*(t)^T$ ~~$(Ax-b)$~~ $(Ax^*(t)-b)$

$f_0(x^*(t))\geq$ 　$= f_0(x^*(t))-\frac{m}{t}$

$\Rightarrow p^* \geq d^* \geq g(\lambda^*(t),v^*(t)) = L_{Q_1}(x^*(t),\lambda^*(t),v^*(t)) = f_0(x^*(t))-\frac{m}{t}$

$\Rightarrow$
———— $f_0(x^*)$
———— p^*
———— $f_0(x^*(t))-\frac{m}{t}$ 　if $t\to\infty$, $p^*\to f_0(x^*(t))$

$\Rightarrow x^*(t)$, $\lambda^*(t)$, $v^*(t)$ satisfies: 　$\lambda^*(t)$, $v^*(t)$ 并不 satisfy KKT condition

① $f_i(x^*(t))\leq 0$, $Ax^*(t)=b$

② $\lambda_i^*(t)\geq 0$

③ $\nabla f_0(x)+\sum_{i=0}^m \lambda_i^*(t) f_i(x^*(t))+A^Tv^*(t)=0$

④ $-\lambda_i^*(t) f_i(x^*(t)) = \frac{1}{t}$

But compare ①②③④ above with conventional KKT, 可发现, 仅第4条不同

conventional KKT 中 $\lambda_i f_i(x)=0$.

11.3 Barrier Method.

由于 $p^*\in(f_0(x^*(t))-\frac{m}{t}, f_0(x^*(t)))$, 可认为 if $\frac{m}{t}<\epsilon$, 那则 $f_0(x^*(t))-p^*<\epsilon$ $\Rightarrow$ can stop using different central points.

Barrier method 核心在于, 每次选择一个t值, 代入Q_b, 用 newton method 求解 Q_b, 得到 $x^*(t)$, 若 $\frac{m}{t}\geq\epsilon$, 则增大 t ($t^+=\mu t$, $\mu>1$) (使$\phi(x)$更接近 indicator function, 可得到更接近Q_1 solution的 solution), 再次代入Q_b, 用 newton 法得到新的 $x^*(t)$ 　associated with $x^*(t)$ and dual feasible pair $\lambda^*(t)$, $v^*(t)$

理解 outer iteration 与 inner iteration

令每次选择 ~~central~~ 点的过程叫 $\underline{\text{outer iteration}}$. 　outer iteration 用于缩小 duality gap $\frac{m}{t}$ by using $t^+=\mu t$, $\mu>1$

而选定t后用 newton 法求解 Q_b solution $x^*(t)$ 的 iteration (line search until $\frac{\lambda^2(x)}{2}<\epsilon$) 叫 ~~inex~~ $\underline{\text{inner iteration}}$ 　inner iteration 用于缩小 $f_0(x^*(t))$ by newton's method.

$\Rightarrow$ $f_0(x^*(t^{(k)}))$ 　$\frac{m}{t^{(k)}}$ 　$\frac{m}{t^{(k+1)}}$ 　$f_0(x^*(t^{(k+1)}))$ 　p^*

kth outer iteration 　k+1 th outer iteration

联系 newton 法需选择一个 starting point, 令 $t^{(k)}$ 次的 solution $x^*(t^{(k)})$ 做为 $k+1$ th inner iteration 的 starting point.

显然 $\mu\nearrow$ ($t^+=\mu t$), # outer iteration 减少 # of inner iteration $\uparrow$ (比较消耗)

一般来说 if $\mu\in(3,100)$, total # of newton iterations in Barrier Method

由于 $x^*(t^{(k+1)})$ 以 $x^*(t^{(k)})$ 为 starting point 用 newton 法得到 $\Rightarrow f(x^*(t^{(k+1)}))<f_0(x^*(t^{(k)}))$

$\mu\downarrow$, # of outer iteration $\nearrow$, # of inner interation $\downarrow$ 　由于 $t^{(k+1)}=\mu t^{(k)}$

The outer iteration reduces the gap by a small amount

由于 the stating point of newton method is a very good point. $\Rightarrow$ Gap $\downarrow$

if $t\to\infty$, $f(x^*(t^{\infty}))\to p^*$ 找到 p^*

但应注意：上述分析均建立在有一个 strictly feasible starting point $x^{(0)}$ 的前题下.

如何找 strictly feasible starting point？用 Phase I Method.

即找 $x^{(0)}$, such that $Ax^{(0)}=b$, $f_i(x^{(0)})<0$ for $i=1,\dots,m$ Q7

即目标: find a strictly feasible solution for $\boxed{\begin{cases} Ax=b \\ f_i(x)<0 \quad i=1,\dots,m \end{cases}}$

⇒ 联系 P25 页 feasibility problem

introduce a slack variable s

$$\boxed{\begin{aligned} &\text{minimize } s \\ &\text{subject to } f_i(x)\le s \quad i=1,\dots,m \\ &\qquad Ax=b \end{aligned}} \quad \text{Q8}$$

① if optimal value $\bar{p}^*<0$ ⇒ Q7 is strictly feasible

② if $\bar{p}^*>0$ ⇒ Q7 is infeasible

$\cancel{\max}$ $\underset{\text{over } x}{\text{minimize}}\left(\underset{\text{over } i}{\max} f_i(x)\right)>0$

③ if $\bar{p}^*=0$

(3.1) minimum is attained at x^* and $s^*=0$

⇒ Q7 is feasible but not strictly feasible

(3.2) minimum is not attained

⇒ Q7 is infeasible

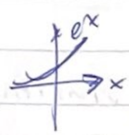
$\exists\, x^0$, such that $f_i(x^0)\le 0$, $Ax^0=b$

but $\nexists\, x'$ such that $f_i(x')<0$, $Ax'=b$

↳ feasible not strictly feasible

例. minimize s

subject to $e^x<s$



⇒ $\bar{p}^*=0$

cause for any $s>0$ can find x such that $e^x<s$

but $e^x=p^*=0$ is not attained

⇒ the original problem with constraint $f(x)=e^x\le 0$ ~~can~~ is infeasible

50