6.3 PENALTY FUNCTION METHOD & CONVEX PROGRAMS

RECALL – DUALITY GAP... $MP \geq MD$

THEOREM 57 (6.5.1)

LET $f, g_1, \ldots g_m$ ARE CONVEX FUNCTIONS WITH CONTINUOUS FIRST PARTIAL DERIVATIVES AND LET f BE COERCIVE.

IF PROGRAM

$$(P)\begin{cases} \min f(x) \\ \text{s.t. } g_1(x) \leq 0, \ldots g_m(x) \leq 0 \\ x \in \mathbb{R}^n \end{cases}$$

IS CONSISTENT THEN (DP) IS CONSISTENT AND $MP = MD$

(NO DUALITY GAP)

$(DP_k \Rightarrow \lambda_k \ldots L(x_k, \lambda_k) \Rightarrow f(x_k) \leq MD)$

NOTE: DUAL CONSISTENT AS $f(x)$ COERCIVE (AND f, g SMOOTH)

PROOF: USE $P_k(x) = f(x) + k \sum_i (g_i^+(x))^2$

COROLLARY $\Rightarrow \exists x_k$ S.T. $P_k(x_k) = \min_x P_k(x)$

& CONVERGING SUBSEQUENCE COURANT BECKMAN

$\rightarrow x^*$ SOL. OF (P).

x_k MIN. $\Rightarrow 0 = \nabla P_k(x_k) = \nabla f(x_k) + \sum_i 2k\, g_i^+(x_k) \nabla g_i(x_k)$

LET $\lambda_i^{(k)} = 2k\, g_i^+(x_k) \geq 0$

THEN $\nabla \{ L(x_k, \lambda^{(k)}) \} = 0$

ALSO $\forall \lambda$: $L(x, \lambda)$ IS CONVEX ($f(x) + \sum \lambda g_i(x)$)

FOR $\lambda^{(k)}$, x_k IS GLOBAL MINIMIZER $\Rightarrow$

$\Rightarrow \min\{L(x, \lambda^{(k)})\} > -\infty$ & $\lambda^{(k)}$ FEASIBLE

FOR DUAL.

AS (P) IS CONSISTENT & f IS COERCIVE $\Rightarrow \exists$ SOLUTION

$\Rightarrow \lim x_k = x^*$ IS ALSO A SOLUTION (BY THM BEFORE)

$f(x_k) \leq P_k(x_k) = f(x_k) + \sum k [g_i^+(x_k)]^2 \leq$

$\leq f(x_k) + \sum k 2 (g^+(x_k))^2 =$

$= f(x_k) + \sum 2k\, g_i^+(x_k)\, g_i(x_k) =$

$= f(x_k) + \sum \lambda_i^{(k)} g_i(x_k) = L(x_k, \lambda^{(k)})$

$= \min\{L(x, \lambda^{(k)})\} \leq MD$

$\Rightarrow MP = f(x^*) = \lim f(x_k) \leq MD$. [MP $\geq$ MD ALWAYS]

NOT ALL FUNCTIONS ARE COERCIVE.

$$(P)\begin{cases} \text{MIN } f(x) \\ \text{ST } g_1(x) \le 0 \ldots g_m(x) \le 0 \\ x \in \mathbb{R}^n \end{cases}$$

f NOT COERCIVE

REPLACE f BY f^{ε} COERCIVE & "CLOSE" TO f

$$f^{\varepsilon}(x) = f(x) + \varepsilon \|x\|^2$$

ε SMALL .. ALMOST $f(x)$ FOR SMALL x

$$\lim_{\|x\| \to \infty} f^{\varepsilon}(x) = +\infty$$

f IS CONVEX $\Rightarrow \exists d \in \mathbb{R}^n$ ST.

$$f(x) \ge f(0) + d^T x \quad \forall x \in \mathbb{R}^n$$

$$f^{\varepsilon}(x) = f(x) + \varepsilon \|x\|^2 \ge f(0) + d^T x + \varepsilon \|x\|^2 \ge$$

CAUCHY-SCHWARTZ: $|d^T x| \le \|d\| . \|x\|$

$$\ge f(0) - \|d\| \|x\| + \varepsilon \|x\|^2 = c_1 \|x\|^2 - c_2 \|x\| + c_3$$

$\Rightarrow f^{\varepsilon}(x)$ IS COERCIVE

$$(P)\begin{cases} \text{MIN } f(x) \\ \text{S.T. } g_1(x) \leq 0 \ \ldots \ g_m(x) \leq 0 \\ x \in \mathbb{R}^n \end{cases}$$

$\downarrow$

$$(P^\varepsilon)\begin{cases} \text{MIN } f^\varepsilon(x) \\ \text{S.T. } g_1(x) \leq 0 \ \ldots \ g_m(x) \leq 0 \\ x \in \mathbb{R}^n \end{cases}$$

$$L^\varepsilon(x,\lambda) = f(x) + \varepsilon \|x\|^2 + \sum_{i=1}^m \lambda_i g_i(x) =$$

$$= L(x,\lambda) + \varepsilon \|x\|^2$$

$$(DP^\varepsilon)\begin{cases} \text{MAX } \theta^\varepsilon(\lambda) = \inf_{x \in \mathbb{R}^n} \{L(x,\lambda) + \varepsilon \|x\|^2\} \\ \lambda \geq 0, \ \lambda \in \mathbb{R}^m \end{cases}$$

(P) CONSISTENT $\Rightarrow$ (P^ε) CONSISTENT

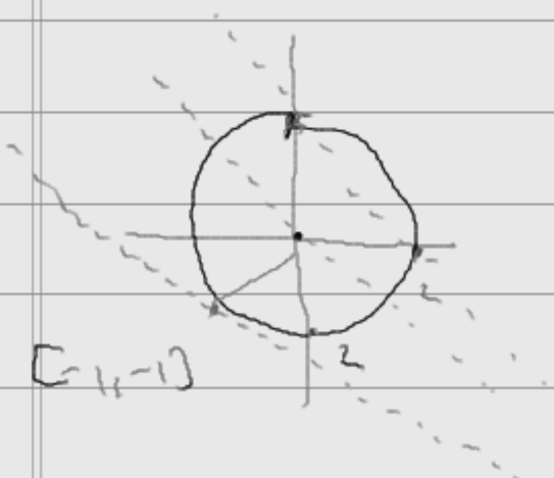
LEMMA (THM 57 - 6.3.2)

LET $f, g_1 \ldots g_m$ HAVE CONTINUOUS FIRST PARTIAL DERIVATIVES AND

(P) BE CONSISTENT. THEN FOR EVERY $\varepsilon > 0$ PROGRAMS (P^{ε}) AND (DP^{ε}) ARE CONSISTENT AND $MD^{\varepsilon} = MP^{\varepsilon}$.

EXAMPLE:

$$(P)\begin{cases} \text{MIN.} & f(x,y) = x+y \\ \text{ST.} & g(x,y) = x^2+y^2-2 \le 0 \\ & x,y \in \mathbb{R}^2 \end{cases}$$



$x^* = -1,\ y^* = -1$

AS (P) SUPER CONSISTENT

$MP = -2 = MD$

$$f^{\varepsilon}(x,y) = x+y + \varepsilon\|(x,y)\|^2 = x+y+\varepsilon(x^2+y^2)$$
$$= \varepsilon\left(x^2+\tfrac{1}{\varepsilon}x\right) + \varepsilon\left(y^2+\tfrac{1}{\varepsilon}y\right)$$
$$= \varepsilon\left(x+\tfrac{1}{2\varepsilon}\right)^2 + \varepsilon\left(y+\tfrac{1}{2\varepsilon}\right)^2 - \frac{2\varepsilon}{4\varepsilon^2}$$

CIRCLES AROUND $\left(-\frac{1}{2\varepsilon}, -\frac{1}{2\varepsilon}\right)$

MW... $\left(x+\frac{1}{2\epsilon}\right)^2 + \left(y+\frac{1}{2\epsilon}\right)^2 = 0$

$x = -\frac{1}{2\epsilon}$ BUT $x^2+y^2 \le 2$

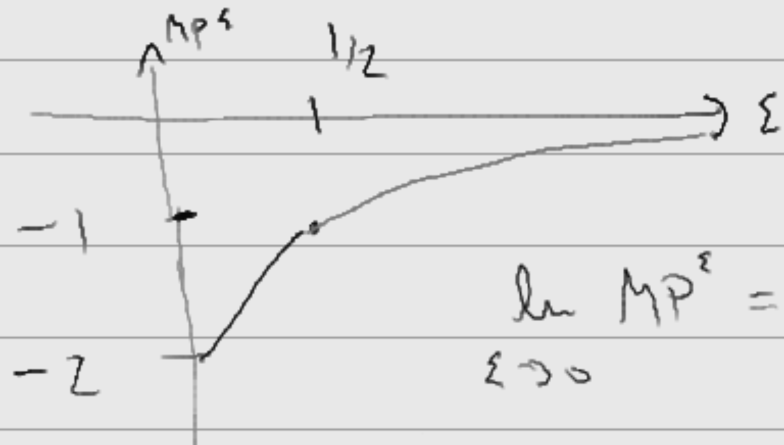
$y = -\frac{1}{2\epsilon}$ $\quad \frac{1}{4\epsilon^2} + \frac{1}{4\epsilon^2} \le 2$

$\Rightarrow \frac{1}{\epsilon^2} \le 1 \quad \therefore \quad \epsilon \ge \frac{1}{2}$

- $f^\epsilon(x,y) = -\frac{1}{2\epsilon}$ FOR $\epsilon \ge \frac{1}{2}$ $\quad MP^\epsilon = \frac{-1}{2\epsilon}$
- FOR $0 < \epsilon \le \frac{1}{2}$ SET $x = y = -1$

$$f^\epsilon(x,y) = x+y+\epsilon(x^2+y^2) =$$
$$= 2\epsilon - 2$$

SO $MP^\epsilon = 2\epsilon - 2$



$\lim_{\epsilon \to 0} MP^\epsilon = -2 = MP = MD$

So $MP^\epsilon = MD^\epsilon$ BY LEMMA

$\lim_{\epsilon \to 0} MD^\epsilon = MP$

LEMMA (6.34)

LET $f, g_1, \ldots g_m$ HAVE CONTINUOUS FIRST PARTIAL DERIVATIVES ON $\mathbb{R}^n$. IF PROGRAM (P) IS CONSISTENT AND MP $> -\infty$ THEN:

(1) THE PROGRAM (DP^ε) IS CONSISTENT FOR ALL $\varepsilon > 0$

(2) $MP = \inf\{MD^\varepsilon : \varepsilon > 0\}$

PROOF.

$\forall \varepsilon > 0 \ \forall x: \quad f(x) \leq f^{\varepsilon}(x)$

$\Rightarrow MP^{\varepsilon} \geq MP$

BY PREV. LEMMA $MP^{\varepsilon} = MD^{\varepsilon}$ AND (DP^{ε}) IS CONSISTENT.

$$MP \leq \inf \{ \underset{\substack{\shortparallel \\ \inf \{ MD^{\varepsilon} : \varepsilon > 0 \}}}{MP^{\varepsilon}} : \varepsilon > 0 \} = \inf_{\varepsilon} \{ \inf_{x} f(x) + \varepsilon \|x\|^2 : g_i(x) \leq 0 \}$$

$$= \inf_{x} \{ \inf_{\varepsilon} f(x) + \varepsilon \|x\|^2 : g_i(x) \leq 0 \} =$$

$$= \inf_{x} \{ f(x) : g_i(x) \leq 0 \} = MP.$$

$\square$

NOT TRUE THAT

$MD = \inf \{ MD^{\varepsilon} : \varepsilon > 0 \}$ ... DUALITY GAP

THEOREM 58 (6.3.5)

LET $f, g_1, \dots, g_m$ ARE CONVEX WITH CONTINUOUS FIRST PARTIAL DERIVATIVES.

IF PROGRAM

$$(P) \begin{cases} \text{MIN } f(x) \\ \text{S.T. } g_1(x) \le 0 \dots g_m(x) \le 0 \\ \text{WHERE } x \in \mathbb{R}^n \end{cases}$$

IS SUPERCONSISTENT AND $MP > -\infty$ THEN

1) (DP) IS CONSISTENT

2) $MP = MD$

3) $\exists\ \lambda^* \in \mathbb{R}^m$ A SOLUTION TO (DP)

$$(DP) \begin{cases} \max h(\lambda) = \inf \{ L(x, \lambda) : x \in \mathbb{R}^n \} \\ \text{S.T. } \lambda \ge 0 \end{cases}$$

IF x^* IS A SOLUTION TO (P) THEN

$$f(x^*) = L(x^*, \lambda^*) = h(\lambda^*)$$

MOREOVER,

$$\lambda_i^* g_i(x^*) = 0 \text{ FOR } i = 1, 2 \dots m$$

AND λ^* IS SENSITIVITY VECTOR FOR (P)