Homework 7

3.4.1 $\quad h_1 = a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + b_1 \qquad A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$ invertible

$\qquad h_2 = a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + b_2$

a) $\nabla h_1(x^*) = \begin{pmatrix} a_{11} \\ a_{12} \\ a_{13} \end{pmatrix} \qquad \nabla h_2(x^*) = \begin{pmatrix} a_{21} \\ a_{22} \\ a_{23} \end{pmatrix}$

Suppose $\nabla h_1(x^*)$. $\nabla h_2(x^*)$ are dependent. Then $\exists\ \lambda, \mu \neq 0$ s.t $\lambda \begin{pmatrix} a_{11} \\ a_{12} \\ a_{13} \end{pmatrix} + \mu \begin{pmatrix} a_{21} \\ a_{22} \\ a_{23} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$

In particular $\lambda \begin{pmatrix} a_{11} \\ a_{12} \end{pmatrix} + \mu \begin{pmatrix} a_{21} \\ a_{22} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Rightarrow rk \begin{bmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{bmatrix} < 2$

$\Rightarrow rk \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = rk \begin{bmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{bmatrix} < 2$ which contradicts that $\begin{bmatrix} a_{12} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$ is invertible.

$\therefore$ By Lemma 3.4.1 for any $w \in \mathbb{R}^3$ satisfying $\nabla h_1(x^*)^T w = 0$, $\nabla h_2(x^*)^T w = 0$ $\exists\ \bar{t} \in (0, 1]$ and differentiable functions $x(t) = (x_1(t), x_2(t), x_3(t))$ $t \in [0, 1]$ s.t $x(0) = x^*$ $x'(0) = w$ & $x(t)$ feasible $\forall\ t \in [0, \bar{t}]$.

b) $h_1(x(t)) = 0 \Rightarrow a_{11}x_1(t) + a_{12}x_2(t) + a_{13}x_3(t) + b_1 = 0$

$\quad h_2(x(t)) = 0 \Rightarrow a_{21}x_1(t) + a_{22}x_2(t) + a_{23}x_3(t) + b_2 = 0$

$$\Rightarrow \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix} \begin{pmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{pmatrix} = \begin{pmatrix} -b_1 \\ -b_2 \end{pmatrix} \Rightarrow \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}^{-1} \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix} \begin{pmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{pmatrix}$$

$$\Rightarrow \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix} + A^{-1} \begin{pmatrix} a_{13} \\ a_{23} \end{pmatrix} x_3(t) = A^{-1} \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$$

$$\begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix} = A^{-1} b - A^{-1} \begin{pmatrix} a_{13} \\ a_{23} \end{pmatrix} x_3(t) \quad //$$

c) More generally if $\exists$ m equations $h_i = \sum_{j=1}^{n} a_{ij}x_j + b_i \quad i=1,\dots,m$
and $A = \begin{bmatrix} a_{11} & \cdots & a_{1m} \\ a_{m1} & \cdots & a_{mm} \end{bmatrix}$ is invertible

then $\nabla h_i(x^*) = \begin{pmatrix} a_{i1} \\ a_{i2} \\ \vdots \\ a_{in} \end{pmatrix}$. They are LI since the top $m \times m$ minor of the matrix $\left(\nabla h_1(x^*), \dots \nabla h_m(x^*)\right)$ is invertible

and
$$I \begin{pmatrix} x_1(t) \\ \vdots \\ x_m(t) \end{pmatrix} + A^{-1} \begin{bmatrix} a_{1m+1} & \cdots & a_{1n} \\ \vdots & & \\ a_{m\,m+1} & \cdots & a_{mn} \end{bmatrix} \begin{pmatrix} x_{m+1}(t) \\ \vdots \\ x_n(t) \end{pmatrix} = A^{-1}b$$

$$\therefore \begin{pmatrix} x_1(t) \\ \vdots \\ x_m(t) \end{pmatrix} = A^{-1}b - A^{-1} \begin{bmatrix} a_{1m+1} & \cdots & a_{1n} \\ a_{m\,m+1} & \cdots & a_{mn} \end{bmatrix} \begin{pmatrix} x_{m+1}(t) \\ \vdots \\ x_n(t) \end{pmatrix}$$

Also know $x'(0) = w$ and $x(0) = x^*$.

3.4.2: NLP $\min f(x)$
s.t $g_1(x) \le 0, \dots, g_\ell(x) \le 0$
1st partials of f, g_i's are continuous on $\mathbb{R}^n$
x^* local opt of NLP. $I(x^*) = \{ i \in \{1,\dots,\ell\} : g_i(x^*) = 0 \}$

Consider the inequality system:

$$(\dagger) \quad \begin{aligned} &\nabla f(x^*)^T w < 0 \\ &\nabla g_i(x^*)^T w < 0 \quad \forall i \in I(x^*) \end{aligned}$$

Show this system has no sol^n w.

Suppose $\exists$ w satisfying $(\dagger)$. We will contradict that x^* is a local min.

Let $\phi(t) = f(x^* + tw)$. Then $\phi'(t) = \nabla f(x^* + tw)^T w \quad \forall t \in \mathbb{R}$ by Lemma 3.1.1.

$\Rightarrow \phi'(0) = \nabla f(x^*)^T w < 0$ by assumption.

Since the 1st partial of f is conts on $\mathbb{R}^n$, $\exists\ \bar{t}$ s.t $\forall$
$0 < t \le \bar{t}$ $\quad \frac{\phi(t) - \phi(0)}{t} < 0 \quad \Rightarrow \phi(t) < \phi(0)$

$\Rightarrow f(x^* + tw) < f(x^*)$.

This will contradict that x^* is a local min if we can further show that $g_i(x^* + tw) \le 0 \quad \forall \quad t \in (0, \bar{t}]$

If $i \notin I(x^*)$ then $g_i(x^*) < 0$ $\therefore$ for small t
$g_i(x^* + tw) \le 0$. So we need only worry about $i \in I(x^*)$.

If $i \in I(x^*)$, we are given $\nabla g_i(x^*)^T w < 0$

$$g_i(x^* + tw) = g_i(x^*) + \nabla g_i(x^*)^T tw + \frac{1}{2}(tw)^T H g_i(z)(tw)$$

for some $z \in [x^*, x^* + tw]$ by Taylor

Now $g_i(x^*) = 0$ since $i \in I(x^*)$

$\nabla g_i(x^*)^T tw < 0$ since $t > 0$ and $\nabla g_i(x^*)^T w < 0$ by assumption

and the second order term is small for t small.

$\therefore g_i(x^* + tw) \le 0$ for t small.

Choosing $\bar{t}$ to be small enough to make all our "t small" statements true

we get that $f(x^* + tw) < f(x^*) \quad \forall \quad 0 < t \le \bar{t}$
But since this segment $[x^*, x^* + tw]$ ~~lies in~~ intersects any ball with center x^* we get that x^* is not a local opt. Contradiction!

(b) To do this part you need a version of duality that deals with strict inequalities

Claim: Either $Ax < b$ has no sol$^{\underline{n}}$ or $y = 0$ is not the only sol$^{\underline{n}}$ of the system $y \geq 0$, $yA = 0$, $yb \leq 0$.

Apply this to your system from (a) let $p = |I(x^*)|$

$\nabla f(x^*)^T w < 0$ $\quad \nabla g_i(x^*)^T w < 0$ $\quad i \in I(x^*)$ has no sol$^{\underline{n}}$ $\iff$

$u = (u_0, u_1, \dots u_p)$ is not the only sol$^{\underline{n}}$ to

$$u \geq 0, \quad (u_0, u_1, \dots u_p) \begin{bmatrix} \nabla f(x^*) \\ \nabla g_i(x^*) \end{bmatrix} \Big\} i \in I(x^*) = 0 \qquad u \cdot 0 \leq 0$$

Extend to $u_0, u_1, \dots u_\ell$ by setting $u_i = 0$ if $i \notin I(x^*)$

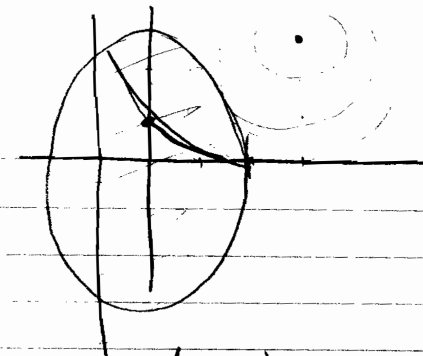
$\therefore \exists$ a nonzero sol$^{\underline{n}}$ $(u_0, u_1, \dots, u_\ell)$ to the system

$$u_0 \geq 0 \quad u_1, \dots, u_\ell \geq 0$$

$$u_0 \nabla f(x^*) + \sum_{i=1}^{\ell} u_i \nabla g_i(x^*) = 0$$

Check that $u_i g_i(x^*) = 0 \ \forall \ i = 1 \dots \ell$ since

if $g_i(x^*) \neq 0$ then $i \notin I(x^*)$ $\Rightarrow$ $u_i = 0$ by the way we set it.



3.4.3

$$\min \ (x_1-3)^2 + (x_2-3)^2 \qquad f$$

$$\text{s.t} \quad 4x_1^2 + 9x_2^2 - 36 \le 0 \qquad g_1$$

$$-x_1 - 1 \le 0 \qquad g_2$$

$$x_1^2 + 3x_2 - 3 = 0 \qquad h_1$$

$$\nabla f = \begin{pmatrix} 2(x_1-3) \\ 2(x_2-3) \end{pmatrix} \quad \nabla g_1 = \begin{pmatrix} 8x_1 \\ 18x_2 \end{pmatrix} \quad \nabla g_2 = \begin{pmatrix} -1 \\ 0 \end{pmatrix} \quad \nabla h_1 = \begin{pmatrix} 2x_1 \\ 3 \end{pmatrix}$$

(KKT)
$$2x_1 - 6 + \lambda_1 8x_1 - \lambda_2 + 2\mu_1 x_1 = 0$$
$$2x_2 - 6 + 18\lambda_1 x_2 + 3\mu_1 = 0$$
$$\left\{ \begin{array}{ll} \lambda_1 \ge 0 & \lambda_2 \ge 0 \\ \lambda_1 g_1 = 0 & \lambda_2 g_2 = 0 \end{array} \right\}$$

Case 1: $\lambda_1 = \lambda_2 = 0$

(KKT) $\Rightarrow \quad 2x_1 - 6 + 2\mu_1 x_1 = 0 \quad \Rightarrow 2x_1(\mu_1 + 1) = 6$

$\Rightarrow \boxed{x_1 = \frac{3}{1+\mu_1}}$

$2x_2 - 6 + 3\mu_1 = 0 \quad \Rightarrow \boxed{x_2 = \frac{6-3\mu_1}{2}}$

$h_1 = 0 \quad \Rightarrow \left(\frac{3}{1+\mu_1}\right)^2 + 3\left(\frac{6-3\mu_1}{2}\right) = 3$

$\Rightarrow \frac{9}{(1+\mu_1)^2} = 3\left(\frac{\cancel{2} - 6 + 3\mu_1}{2}\right) = \frac{3}{2}(3\mu_1 - 4)$

$\Rightarrow 6 = (3\mu_1 - 4)(1+\mu_1)^2 = (3\mu_1 - 4)(1 + \mu_1^2 + 2\mu_1)$

$= 3\cancel{\mu_1} + 3\mu_1^3 + 6\mu_1^2 - 4 - 4\mu_1^2 - 8\cancel{\mu_1}$

$6 = 3\mu_1^3 + 2\mu_1^2 - 5\mu_1 - 4$

$\Rightarrow 3\mu_1^3 + 2\mu_1^2 - 5\mu_1 - 10 = 0$ has only one real root $\boxed{\mu_1 = 1.6238}$

$\therefore \boxed{x_1 = \frac{3}{2.6238} = 1.1433 \qquad x_2 = 0.5642}$ Satisfies $g_1 \le 0 \quad g_2 \le 0$

Case 2: $\lambda_1 = 0, \quad g_2 = 0$

$g_2 = 0 \quad \Rightarrow \quad -x_1 - 1 = 0 \quad \Rightarrow \boxed{x_1 = -1}$

KKT① $\Rightarrow \quad 2x_1 - 6 - \lambda_2 + 2\mu_1 x_1 = 0 \quad \Rightarrow -2 - 6 - \lambda_2 - 2\mu_1 = 0$

$\Rightarrow -8 - \lambda_2 - 2\mu_1 = 0 \quad \Rightarrow \boxed{\lambda_2 = -8 - 2\mu_1}$

not ok so discard

$\lambda_2 = -8 - \frac{28}{9} = -\frac{100}{9}$

$h_1 = 0 \Rightarrow 3x_2 = 3 - x_1^2 = 3 - 1 = 2 \qquad \boxed{x_2 = 2/3}$

KKT ② $\Rightarrow 2\left(\frac{2}{3}\right) - 6 + 3\mu_1 = 0 \quad \Rightarrow 3\mu_1 = 6 - \frac{4}{3} = \frac{14}{3} \quad \therefore \boxed{\mu_1 = \frac{14}{9}}$

Case 3 $\lambda_2 = 0 \quad g_1 = 0$

$g_1 = 0 \Rightarrow \boxed{4x_1^2 + 9x_2^2 = 36}$ (i) $\quad h_1 = 0 \Rightarrow \boxed{x_1^2 + 3x_2 - 3 = 0}$ (ii)

(KKT) $\Rightarrow$ $2x_1 - 6 + 8\lambda_1 x_1 + 2\mu_1 x_1 = 0$

$2x_2 - 6 + 18\lambda_1 x_2 + 3\mu_1 = 0$

(ii) $\Rightarrow x_1^2 = 3 - 3x_2$ substituting in (i) we get

$4(3-3x_2) + 9x_2^2 = 36 \Rightarrow 12 - 12x_2 + 9x_2^2 - 36 = 0$

$\Rightarrow 9x_2^2 - 12x_2 - 24 = 0 \Rightarrow 3x_2^2 - 4x_2 - 8 = 0$

$x_2 = \dfrac{4 \pm \sqrt{16 + 96}}{6} = \dfrac{4 \pm \sqrt{112}}{6} = \boxed{2.4305 \text{ or } -1.0971}$

$x_1^2 = 3 - 3(2.4305) = -4.2915$ (not possible)

$x_1^2 = 3 - 3(-1.0971) \Rightarrow x_1 = 2.50824$

$\therefore$ One possibility is $x_1 = 2.50824 \quad x_2 = -1.0971$

Now substituting this into KKT gives

$\lambda_1 = -0.2395 \quad \mu_1 = 1.15436$ (discard since $\lambda_1 < 0$)

Case 4 $g_1 = 0 \quad g_2 = 0$

$g_2 = 0 \Rightarrow \boxed{x_1 = -1} \quad g_1 = 0 \Rightarrow 4x_1^2 + 9x_2^2 - 36 = 0$

$\Rightarrow 4 + 9x_2^2 - 36 = 0 \rightarrow 9x_2^2 = 32$

$x_2^2 = \dfrac{32}{9} \quad \boxed{x_2 = \pm 1.885618}$

Check if $h_1 = 0$

$1 + 3(1.885618) - 3 = -2 + 3(1.885618) \neq 0$

$\therefore$ This sol^n is not feasible

$1 + 3(-1.885618) - 3 \neq 0 \quad \therefore$ both sol^{n}s not feasible.

$\therefore$ Only $x_1 = 1.1433 \quad x_2 = 0.5642 \quad \mu_1 = 1.6238 \quad \lambda_1 = \lambda_2 = 0$ is a possible local opt.

$Hf = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$ $Hg_1 = \begin{bmatrix} 8 & 0 \\ 0 & 18 \end{bmatrix}$ $Hg_2 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$ $Hh = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}$

$\therefore Hf + \lambda_1 Hg_1 + \lambda_2 Hg_2 + \mu_1 Hh = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} + 1.6238 \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}$

$= \begin{bmatrix} 5.2476 & 0 \\ 0 & 2 \end{bmatrix} \succ 0$ since diagonal entries are positive.

So this sol^n is a local opt.