MATH 3161/5165 – Term 1, 2019

1.i a) Let $q(x) = \frac{1}{2} x^T G x + d^T x + c$

The steepest descent direction at $x^{(k)}$ is given by $s_D^{(k)} = -\nabla q(x^{(k)}) = -\left(G\underline{x}^{(k)} + \underline{d}\right)$.

b) $$\nabla q(x^{(k)})^T s_D^{(k)} = -\nabla q(x^{(k)})^T \nabla q(x^{(k)})$$
$$= -\|\nabla q(x^{(k)})\|_2^2 < 0$$
$$(\because \nabla \underline{q}(x^{(k)}) \neq 0)$$

c) The line search condition is
$$\nabla q(x^{(k)} + \alpha s^{(k)})^T s_D^{(k)} = 0 \text{ at } \alpha = \alpha^{(k)}.$$

d) $\underline{x}^{(k+2)} - \underline{x}^{(k+1)} = \alpha^{(k+1)} \underline{s}_D^{(k+1)}$ and

$\underline{x}^{(k+1)} - \underline{x}^{(k)} = \alpha^{(k)} \underline{s}_D^{(k)}$, where $\alpha^{(k)}, \alpha^{(k+1)} > 0$.

$$\left(\underline{x}^{(k+2)} - x^{(k+1)}\right)^T \left(\underline{x}^{(k+1)} - \underline{x}^{(k)}\right)$$
$$= \alpha^{(k+1)} s_D^{(k+1)} \alpha^{(k)} s_D^{(k)}$$
$$= \alpha^{(k+1)} \alpha^{(k)} \left(-\nabla q(\underline{x}^{(k+1)})\right) \underline{s}_D^{(k)}$$
$$= -\alpha^{(k+1)} \alpha^{(k)} \nabla q\left(x^{(k)} + \alpha^{(k)} s_D^{(k)}\right) \underline{s}_D^{(k)}$$
$$= -\alpha^{(k+1)} \alpha^{(k)} \cdot 0 \cdot \cancel{s_D^{(k)}}$$ by the line search condition
$$= 0.$$

ii) a) Let $\nabla q(\underline{x}) = \begin{bmatrix} x_1 - x_2 + 4 \\ -x_1 + 5x_2 - 6 \end{bmatrix}$. Then, the Hessian

$$G = \nabla^2 q(\underline{x}) = \begin{bmatrix} 1 & -1 \\ -1 & 5 \end{bmatrix}$$

b) Solving $\det(\nabla^2 q(\underline{x}) - \lambda I) = 0$ gives

$$\begin{vmatrix} 1-\lambda & -1 \\ -1 & 5-\lambda \end{vmatrix} = 0 \Leftrightarrow (1-\lambda)(5-\lambda) - 1 = 0$$

$$\lambda^2 - 6\lambda + 4 = 0$$

$$\lambda = \frac{6 \pm \sqrt{36-16}}{2} = \frac{6 \pm \sqrt{20}}{2} = \frac{6 \pm 2\sqrt{5}}{2}$$

The eigenvalues are $\lambda_1 = 3+\sqrt{5}$ & $\lambda_2 = 3-\sqrt{5}$

The condition number of G is

$$\kappa = \frac{\lambda_{max}}{\lambda_{min}} = \frac{3+\sqrt{5}}{3-\sqrt{5}}$$

c) As G is positive definite, the steepest descent method generates a sequence, where for each k,

$$\|\underline{x}^{(k+1)} - \underline{x}^*\| \leq \left(\frac{\kappa - 1}{\kappa + 1}\right) \|\underline{x}^{(k)} - \underline{x}^*\| .$$

$$\|\underline{x}^{(k+1)} - \underline{x}^*\| \le \left(\frac{\frac{3+\sqrt{5}}{3-\sqrt{5}} - 1}{\frac{3+\sqrt{5}}{3-\sqrt{5}} + 1}\right) \|\underline{x}^{(k)} - \underline{x}^*\|$$

$$= \frac{2\sqrt{5}}{6} \|\underline{x}^{(k)} - \underline{x}^*\|$$

$$= \left(\sqrt{5}/3\right) \|x^{(k)} - x^*\|$$

$$\le \left(\sqrt{5}/3\right)^2 \|x^{(k-1)} - x^*\|$$

$$\vdots$$

$$\le \left(\sqrt{5}/3\right)^k \|x^{(1)} - x^*\|$$

Since $\|x^{(1)} - x^*\| \le 1$, it follows that

$$\|x^{(k+1)} - x^*\| \le \left(\sqrt{5}/3\right)^k \cdot 1$$

We find k so that $\left(\sqrt{5}/3\right)^k \le 10^{-10}$.

Taking ln each side, we get

$$k \ge \frac{-10 \ln 10}{\ln\left(\sqrt{5}/3\right)} \quad (\because \ln \sqrt{5}/3 < 0).$$

$$= 78.39.$$

The least number of iterations required to get $\|x^{(k+1)} - x^*\| \le 10^{-10}$ is 79.

ii(d) 5) Newton's direction at $x^{(1)}$ is

$$s_N^{(1)} = -G^{-1} g^{(1)} = -\begin{bmatrix} 1 & -1 \\ -1 & 5 \end{bmatrix}^{-1} \begin{bmatrix} 4 \\ -2 \end{bmatrix}$$

$$= -\tfrac{1}{4} \begin{bmatrix} 5 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 4 \\ -2 \end{bmatrix}$$

$$= -\tfrac{1}{4} \begin{bmatrix} 18 \\ 2 \end{bmatrix} = \begin{bmatrix} -9/2 \\ -1/2 \end{bmatrix}.$$

7) The next iterate is

$$\underline{x}^{(2)} = \underline{x}^{(1)} + \underline{s}_N^{(1)} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} + \begin{bmatrix} -9/2 \\ -1/2 \end{bmatrix} = \begin{bmatrix} -7/2 \\ 1/2 \end{bmatrix}.$$

As $g(\underline{x}^{(2)}) = \begin{bmatrix} -7/2 - 1/2 + 4 \\ 7/2 + 5/2 - 6 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$, & G is pos. def.,

$\underline{x}^{(2)}$ is the minimizer of $q(\underline{x})$. Hence

Newton's method terminates in one iteration.

iii) Assume that $\underline{s}^{(1)}, \dots, \underline{s}^{(n)}$ are conjugate with respect to G, where G is positive definite.

Then, $\underline{s}^{(i)T} G \underline{s}^{(j)} = 0$, for $i \neq j$

Suppose that

$\alpha_1 \underline{s}^{(1)} + \cdots + \alpha_n \underline{s}^{(n)} = \underline{0}$ for some scalars $\alpha_1, \dots \alpha_n$. Let $i \in \{1, 2, \dots n\}$

Multiplying the equality by $\underline{s}^{(i)T} G$,

$$s^{(i)T} G \left[\alpha_1 \underline{s}^{(1)} + \cdots + \alpha_n \underline{s}^{(n)}\right] = \underline{s}^{(i)T} G \underline{0} = 0$$

$$\alpha_1 s^{(i)T} G s^{(1)} + \cdots + \alpha_i s^{(i)T} G s^{(i)} + \alpha_n s^{(i)T} G s^{(n)} = 0$$

$$\Rightarrow \qquad \alpha_i \underline{s}^{(i)T} G \underline{s}^{(i)} = 0.$$

$\Rightarrow \quad \alpha_i = 0 \quad$ as $\underline{s}^{(i)T} G \underline{s}^{(i)} > 0$ & $\underline{s}^{(i)} \neq \underline{0}$..

Hence, $\{s^{(1)}, s^{(2)}, \dots, s^{(n)}\}$ is linearly independent.

b) The greatest number of iterations necessary is 5.

The Conjugate Gradient method generates linearly independent search directions.

The max no of linearly independent directions in $\mathbb{R}^5$ is 5.

2(i) (P_1) $\min\ -x_1x_2x_3$

s.t $x_1+2x_2+2x_3-72\le 0$, $x_1-50\le 0$.

a) The penalty function problem is

$$\min_{x\in\mathbb{R}^3} -x_1x_2x_3 + \mu\Big\{\left[\max\{x_1+2x_2+2x_3-72,0\}\right]^2 + \left[\max\{x_1-50,0\}\right]^2\Big\}$$

b) The gradient of the penalty function is

$$\begin{bmatrix} -x_2x_3+2\mu\max\{x_1+2x_2+2x_3-72,0\}+2\mu\max\{x_1-50,0\} \\ -x_1x_3+4\mu\max\{x_1+2x_2+2x_3-72,0\} \\ -x_1x_2+4\mu\max\{x_1+2x_2+2x_3-72\} \end{bmatrix}.$$

As $x_1(\mu)+2x_2(\mu)+2x_3(\mu)-72 = \dfrac{288}{\mu\left(1+\sqrt{1-4/\mu}\right)^2} > 0$,

$\max\{x_1(\mu)+2x_2(\mu)+2x_3(\mu)-72,0\} = \dfrac{288}{\mu\left(1+\sqrt{1-4/\mu}\right)^2}$

Also, $x_1(\mu)-50 = \dfrac{48}{1+\sqrt{1-4/\mu}} - 50$

$$= \frac{-2-50\sqrt{1-4/\mu}}{1+\sqrt{1-4/\mu}} < 0 \quad (\because \mu>4)$$

$\max\{x_1(\mu)-50,0\} = 0.$

Substituting $x(\mu)$ into the gradient, we see that

$$\begin{bmatrix} a(\mu)^2 + 2\mu \cdot \dfrac{288}{\mu\left(1+\sqrt{1-4/\mu}\right)^2} + 0 \\ 2a(\mu)^2 - 4\mu \cdot \dfrac{288}{\mu\left(1+\sqrt{1-4/\mu}\right)^2} \\ 2a(\mu)^2 + \dfrac{4\mu \cdot 288}{\mu\left(1+\sqrt{1-4/\mu}\right)^2} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

So the gradient of the penalty function vanishes at $x(\mu)$.

Hence $x(\mu)$ is a stationary point.

B) As $a(\mu) = \dfrac{24}{1+\sqrt{1-4/\mu}} \longrightarrow 12$ as $\mu \to \infty$

$$x(\mu) \longrightarrow \begin{pmatrix} 24 \\ 12 \\ 12 \end{pmatrix}, \text{ as } \mu \to \infty$$

So, $x^* = \lim\limits_{\mu \to \infty} x(\mu) = \begin{pmatrix} 24 \\ 12 \\ 12 \end{pmatrix}.$

c) Substituting x^* into the constraints of (P_1),

$$C_1(x^*) = 24 + 24 + 24 - 72 = 0 \quad \&$$

$$C_2(\underline{x}^*) = 24 - 50 < 0.$$

So, x^* is a feasible point & the constraint $C_1(x) = x_1 + 2x_2 + 2x_3 - 72$ is active at x^*.

d) The Lagrangian condition necessary conditions at x^* are

$$\nabla f(x^*) + \sum_{i \in \mathcal{A}(x^*)} \lambda_i^* \nabla C_i(x^*) = 0$$

$$C_i(x^*) = 0, \quad i \in \mathcal{A}(x^*)$$

$$\begin{bmatrix} -x_2^* x_3^* \\ -x_1^* x_3^* \\ -x_1^* x_2^* \end{bmatrix} + \lambda_1^* \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\lambda_1^* = x_2^* x_3^* = 144 > 0.$$

So, x^* is a constrained stationary point.

e) The reduced Hessian is

$W_{Z^*} = Z^{*T} \nabla^2 L(x^*, \lambda^*) Z^*$, where

$Z^{*T} \nabla c_1(x^*) = 0$ & Z^* is a matrix of dimension $n \times n-1 = 3 \times 2$ with full rank.

i.e $\underbrace{\begin{bmatrix} -2 & 1 & 0 \\ 0 & 1 & -1 \end{bmatrix}}_{2\times 3} \underbrace{\begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}}_{3\times 1} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$.

$$W_{Z^*} = \begin{bmatrix} -2 & 1 & 0 \\ 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} 0 & -x_3^* & -x_2^* \\ -x_3^* & 0 & -x_1^* \\ -x_2^* & -x_1^* & 0 \end{bmatrix} \begin{bmatrix} -2 & 0 \\ 1 & 1 \\ 0 & -1 \end{bmatrix}$$

$$= \begin{bmatrix} -2 & 1 & 0 \\ 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} 0 & -12 & -12 \\ -12 & 0 & -24 \\ -12 & -24 & 0 \end{bmatrix} \begin{bmatrix} -2 & 0 \\ 1 & 1 \\ 0 & -1 \end{bmatrix}$$

$$= \begin{bmatrix} -2 & 1 & 0 \\ 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} -12 & 0 \\ 24 & 24 \\ 0 & -24 \end{bmatrix}$$

$= \begin{bmatrix} 48 & 24 \\ 24 & 48 \end{bmatrix}$. As $\Delta_1 = 48 > 0$ & $\Delta_2 = 48^2 - 24^2 > 0$, W_{Z^*} is positive definite & $\lambda_i^* > 0$ for $i \in I \cap A(x^*)$, x^* is a strict local minimizer.

ii) (P_2) $\min\limits_{x \in \mathbb{R}^n} \ \underline{x}^T \underline{x} - 2\underline{d}^T \underline{x}$

s.t. $\underline{a}^T \underline{x} - b \leq 0.$

where $a^T d - b > 0.$

a). $\Omega = \{x \in \mathbb{R}^n : \underline{a}^T x - b \leq 0\}.$

Let $\underline{x}, \underline{y} \in \Omega$ and let $\alpha \in (0,1)$.

Then, $\underline{a}^T x - b \leq 0$ & $\underline{a}^T y - b \leq 0.$

Now, $\underline{a}^T(\alpha \underline{x} + (1-\alpha)\underline{y})$

$= \alpha a^T x + (1-\alpha) a^T y$

$\leq \alpha b + (1-\alpha) b$ $\quad (\because \alpha \in (0,1)).$

$= b.$

So, $\alpha x + (1-\alpha) y \in \Omega$. As it is true for each pair $\underline{x}, y \in \Omega$ and $\alpha \in (0,1)$, Ω is a convex set.

b). Let $f(\underline{x}) = x^T x - 2d^T \underline{x}$. Then

$\nabla f(x) = 2\underline{x} - 2\underline{d}$ and

$\nabla^2 f(\underline{x}) = 2I_{n \times n}$, where I is the identity matrix. So, $\nabla^2 f(x)$ is positive definite for all x and so $f(x)$ is convex. Hence (P_2) is a convex optimization as Ω is convex.

c). The KKT conditions at x^* are given by

$$\left.\begin{aligned} 2x^* - 2\underset{\sim}{d} + \lambda^* \underset{\sim}{a} &= 0 \\ \lambda^*(a^T x^* - b) &= 0 \\ \lambda^* &\geq 0 \end{aligned}\right\} \begin{aligned} &-(1) \\ &-(2) \\ &-(3) \end{aligned}$$

d). Solving (1) — (3),

<u>case 1</u> : $a^T x^* - b < 0$.

Then, $\lambda^* = 0$ by (2). So from (1),

$x^* = d$. This is not possible

as $a^T x^* - b = a^T d - b > 0$.

<u>Case 2</u> : $a^T x^* - b = 0$.

Then from (1), $x^* = d - \frac{\lambda^*}{2} \underset{\sim}{a}$

So, $a^T x^* - b = \underset{-}{a}^T(\underset{-}{d} - \frac{\lambda^*}{2} \underset{-}{a}) - b = 0$

$$a^T d - \frac{\lambda^*}{2} \underset{-}{a}^T a - b = 0$$

$$\lambda^* = \frac{2(a^T d - b)}{\|\underset{-}{a}\|^2} \qquad (\because a^T a = \|a\|^2 \text{ \& } a \neq 0)$$

Hence, $x^* = d - \frac{(\underset{-}{a}^T \underset{-}{d} - b)}{\|a\|^2} \underset{\sim}{a}$

As (P_2) is a convex optimization problem it follows from the KKT sufficient optimality conditions x^* is a global minimizer for (P_2).

The global minimum of (P_2) is given by

$$x^{*T}x^* - 2d^Tx^* = x^{*T}x^* - 2x^{*T}d$$
$$= x^{*T}(x^* - \cancel{2}d).$$

$$\left(\underline{d} - \frac{(a^T d - b)}{\|a\|^2} a\right)^T \left(d - \frac{(a^T d - b)}{\|a\|^2}\underline{a}\right)$$
$$- 2d^T\left(\underline{d} - \frac{a^T d - b}{\|a\|^2}\underline{a}\right)$$
$$= (\underline{d} - \gamma \underline{a})^T(\underline{d} - \gamma \underline{a}) - 2d^Td + 2\gamma d^T\underline{a}$$
$$= d^Td - \gamma d^T\underline{a} - \gamma a^T\underline{d} + \gamma^2 a^T\underline{a} - 2d^Td + 2\gamma d^T\underline{a}.$$
$$= -d^Td + \gamma^2 a^T a.$$

, where $\gamma = \frac{a^T d - b}{\|a\|^2}$

$$= -d^Td + \frac{(a^T d - b)^2}{\|a\|^4} \cdot \|a\|^2$$
$$= \frac{(a^T d - b)^2}{\|a\|^2} - \underline{d}^T\underline{d}$$

c) The Wolfe dual problem for (P_2) is

$$\max_{y,\lambda} \; y^T y - 2d^T y + \lambda (a^T y - b)$$
s.t. $\quad 2y - 2d + \lambda a = 0$

$\qquad \lambda \geq 0$

Eliminating y, we get

$$y = d - \lambda/_2 \, a.$$

So, the dual problem becomes

$$\max_{\lambda \geq 0} \; (d - \lambda/_2 a)^T (d - \lambda/_2 a) - 2d^T (d - \lambda/_2 a) + \lambda (a^T d - \lambda/_2 a^T a - b)$$

$$\max_{\lambda \geq 0} \; -\lambda^2/_4 \, a^T a + \lambda (a^T d - b) - d^T d.$$

Let $r(\lambda) = -\lambda^2/_4 \, a^T a + \lambda (a^T d - b) - d^T d$. Then,

$r'(\lambda) = -\lambda/_2 \, a^T a + a^T d - b$ & $r''(\lambda) = -\frac{a^T a}{2} < 0$

Solving $r'(\lambda) = 0$,

$$\lambda = 2(a^T d - b)/\|a\|^2$$

As $r''(\lambda) < 0$, for all $\lambda \in \mathbb{R}$ and $\lambda > 0$ ($\because a^T d - b > 0$),

$\lambda^* = 2(a^T d - b)/\|a\|^2$ is the global maximizer of (D_2).

So, $\max(D_2) = r(\lambda^*) = -\frac{(a^T d - b)^2}{\|a\|^2} + 2\frac{(a^T d - b)^2}{\|a\|^2} - d^T d$

$$= \frac{(a^T d - b)^2}{\|a\|^2} - d^T d = \min(P_2).$$

3. $\min \int_0^{t_1} (x_1^2 + u_1^2)\, dt$

s.t. $\dot{x}_1 = -x_1 + u_1$

$x_1(0) = 1, \quad x_1(t_1) = 2.$

Let (x_1^*, u_1^*) be the solution to the problem.

i) The Hamiltonian is

$$H = -(x_1^2 + u_1^2) + Z_1(-x_1 + u_1)$$

where Z_1 is a co-state variable

Satisfying $\dot{Z}_1 = -\frac{\partial H}{\partial x_1} = -(-2x_1 - z_1)$

i.e. $\dot{z}_1 = 2x_1 + z_1$

So, $\dot{z}_1^* = 2x_1^* + z_1^*$ —— ①

Maximizing H as a function of u_1,

$$\max_{u_1} -x_1^{*2} - u_1^2 - z_1^* x_1^* + z_1^* u_1$$

$$-x_1^{*2} - z_1^* x_1^* + \max_{u_1} -u_1^2 + z_1^* u_1$$

$$\text{''} \qquad + \max_{u_1} -\left(u_1 - \frac{z_1^*}{2}\right)^2 + \frac{z_1^{*2}}{4}$$

So, $u_1^* = \frac{z_1^*}{2}$ maximizes H.

Solving the state equation with $u_1 = u_1^*$

$$\dot{x}_1^* = -x_1^* + \frac{z_1^*}{2} \qquad \text{——} ②$$

Differentiating (2) w.r.t t,

$$\ddot{x}_1^* = -\dot{x}_1^* + \frac{\dot{z}_1^*}{2}$$

$$= -\dot{x}_1^* + \frac{1}{2}(2x_1^* + z_1^*) \quad \text{(by (1))}$$

$$= -\dot{x}_1^* + x_1^* + \cancel{\dot{x}_1^*} + x_1^* \quad \text{by (2).}$$

$$= 2x_1^*$$

$$\ddot{x}_1^* - 2x_1^* = 0$$

The characteristic equation is

$$\lambda^2 - 2 = 0.$$

$$\lambda = \pm\sqrt{2}.$$

So, $x_1^*(t) = Ae^{\sqrt{2}t} + Be^{-\sqrt{2}t}$, where A & B are arbitrary constants.

$$x_1^*(0) = 1 \implies A + B = 1 \quad —(3)$$

$$x_1^*(t_1) = 2 \implies Ae^{\sqrt{2}t_1} + Be^{-\sqrt{2}t_1} = 2 \quad —(4)$$

The optimal control u_1^* is given by

$$u_1^* = \frac{z_1^*}{2} = \dot{x}_1^* + x_1^*$$

$$= \sqrt{2}Ae^{\sqrt{2}t} - \sqrt{2}Be^{-\sqrt{2}t} + Ae^{\sqrt{2}t} + Be^{-\sqrt{2}t}$$

$$= (1+\sqrt{2})Ae^{\sqrt{2}t} + (1-\sqrt{2})Be^{-\sqrt{2}t}.$$

Using (3) & (4), $A = 1 - B$,

$$(1-B)e^{\sqrt{2}+1} + Be^{-\sqrt{2}+1} = 2$$

$$B = (e^{\sqrt{2}+1} - 2) / (e^{\sqrt{2}+1} - e^{-\sqrt{2}+1}).$$

& $$A = 1 - B = \frac{2 - e^{-\sqrt{2}+1}}{e^{\sqrt{2}+1} - e^{-\sqrt{2}+1}}.$$