Q1 page 1

1. (a)(i)

$$Df = \begin{bmatrix} 8x_2 + (x_1 - x_2)^3 \\ 8x_1 - (x_1 - x_2)^3 \end{bmatrix}$$

$$D^2 f = \begin{bmatrix} 3(x_1 - x_2)^2 & 8 - 3(x_1 - x_2)^2 \\ 8 - 3(x_1 - x_2)^2 & 3(x_1 - x_2)^2 \end{bmatrix}$$

Let $\alpha = (x_1 - x_2)$.

Then $D^2 f = \begin{bmatrix} 3\alpha^2 & 8 - 3\alpha^2 \\ 8 - 3\alpha^2 & 3\alpha^2 \end{bmatrix}$

so $(D^2 f)^{-1} = \dfrac{1}{9\alpha^4 - (8 - 3\alpha^2)^2} \begin{bmatrix} 3\alpha^2 & 3\alpha^2 - 8 \\ 3\alpha^2 - 8 & 3\alpha^2 \end{bmatrix}$

$= \dfrac{1}{48\alpha^4 - 64} \begin{bmatrix} 3\alpha^2 & 3\alpha^2 - 8 \\ 3\alpha^2 - 8 & 3\alpha^2 \end{bmatrix}$

Newton direction

$d = -(D^2 f)^{-1} Df = \dfrac{-1}{16(3\alpha^4 - 4)} \begin{bmatrix} 3\alpha^2 & 3\alpha^2 - 8 \\ 3\alpha^2 - 8 & 3\alpha^2 \end{bmatrix} \begin{bmatrix} 8x_2 + \alpha^3 \\ 8x_1 - \alpha^3 \end{bmatrix}$

$= \dfrac{-1}{16(3\alpha^4 - 4)} \begin{bmatrix} 3\alpha^2 8x_2 + 3\alpha^5 + 24\alpha^2 x_1 - 64x_1 - 3\alpha^5 + 8\alpha^3 \\ 24\alpha^2 x_2 - 64x_2 + 3\alpha^5 - 8\alpha^3 + 24x_1\alpha^2 - 3\alpha^5 \end{bmatrix}$

$= \dfrac{-1}{2(3\alpha^2 - 4)} \begin{bmatrix} \alpha^3 + 3\alpha^2(x_1 + x_2) - 8x_1 \\ -\alpha^3 + 3\alpha^2(x_1 + x_2) - 8x_2 \end{bmatrix}$

New iterate with steplength 1:

$$\bar{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} - d = \frac{1}{2(3\alpha^2-4)} \begin{bmatrix} 2x_1(3\alpha^2-4) - \alpha^3 - 3\alpha^2(x_1+x_2) + 8x_1 \\ 2x_2(3\alpha^2-4) + \alpha^3 - 3\alpha^2(x_1+x_2) + 8x_2 \end{bmatrix}$$

$$= \frac{1}{2(3\alpha^2-4)} \begin{bmatrix} 6x_1\alpha^2 - \alpha^3 - 3\alpha^2(x_1+x_2) \\ 6x_2\alpha^2 + \alpha^3 - 3\alpha^2(x_1+x_2) \end{bmatrix}$$

$$= \frac{1}{2(3\alpha^2-4)} \begin{bmatrix} 3\alpha^2(x_1-x_2) - \alpha^3 \\ 3\alpha^2(x_2-x_1) + \alpha^3 \end{bmatrix} = \frac{\alpha^3}{3\alpha^2-4} \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

So new iterate lies on line $x_1 + x_2 = 1$ (provided $3\alpha^2 - 4 \neq 0$).

(b) If $|x_1 - x_2| < \frac{2}{\sqrt{5}}$ then $\left|\frac{2\alpha^3}{3\alpha^2-4}\right| < |\alpha|$,

so $|\bar{x}_1 - \bar{x}_2| < |x_1 - x_2|$, so get monotonic decrease, and convergence.

Also, $$\frac{|\bar{x}_1 - \bar{x}_2|}{|x_1 - x_2|^3} = \left|\frac{2\alpha^3}{3\alpha^2-4}\right| \Big/ |\alpha|^3 = \left|\frac{2}{3\alpha^2-4}\right| < \frac{5}{4},$$

so get cubic convergence.

(c) Suffices to analyze points of the form $\beta\begin{bmatrix} 1 \\ -1 \end{bmatrix}$. This has $\alpha = 2\beta$,

so $\beta\begin{bmatrix} 1 \\ -1 \end{bmatrix}$ is mapped to $\frac{8\beta^3}{12\beta^2-4}\begin{bmatrix} 1 \\ -1 \end{bmatrix} = \frac{2\beta^3}{3\beta^2-1}\begin{bmatrix} 1 \\ -1 \end{bmatrix}$

Thus $\beta \longrightarrow \dfrac{2\beta^3}{3\beta^2-1}$.

Now, $3\alpha^2 > 4 \Rightarrow$

$$|\beta| - 1 = \frac{|\alpha|^3}{3\alpha^2-4} - 1 = \frac{|\alpha|^3 - 3\alpha^2 + 4}{3\alpha^2 - 4}$$

$$= \frac{(|\alpha|-2)^2(|\alpha|+1)}{3\alpha^2-4}$$

$$\geqslant 0 \quad \text{provided } |\alpha| > \tfrac{2}{\sqrt{3}}.$$

So we always have $|\beta| \geqslant 1$.

Further, $\dfrac{2|\beta|^3}{3\beta^2-1} - |\beta| = \dfrac{|\beta| - |\beta|^3}{3\beta^2-1} < 0$ if $|\beta| > 1$.

So $|\beta|$ decreases monotonically.

Also, $$\frac{2|\beta|^3}{3\beta^2-1} - 1 = \frac{1}{3\beta^2-1}\left(2|\beta|^3 - 3\beta^2 + 1\right)$$

$$= \frac{1}{3\beta^2-1}(|\beta|-1)^2(2|\beta|+1)$$

$$\leqslant \frac{3}{2}(|\beta|-1)^2 \quad \text{since } |\beta| > 1$$

So get quadratic convergence to ± 1, depending on initial sign of β.

2 $$H^{-1} = \frac{1}{9\alpha^4 - (8-3\alpha^2)^2} \begin{bmatrix} 3\alpha^2 & 3\alpha^2 - 8 \\ 3\alpha^2 - 8 & 3\alpha^2 \end{bmatrix}$$

$$= \frac{1}{48\alpha^4 - 64} \begin{bmatrix} 3\alpha^2 & 3\alpha^2 - 8 \\ 3\alpha^2 - 8 & 3\alpha^2 \end{bmatrix}$$

$$=$$

Use $\tilde{H}^{-1} = \frac{1}{8}\begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}$

Then $\tilde{H}^{-1} Df = \frac{1}{8}\begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}\begin{bmatrix} 8x_2 + \alpha^3 \\ 8x_1 - \alpha^3 \end{bmatrix}$

$$= \frac{1}{8}\begin{bmatrix} 3\alpha^3 + 16x_2 - 8x_1 \\ -3\alpha^3 - 8x_2 + 16x_1 \end{bmatrix}$$

And $\tilde{x} = x - \tilde{H}^{-1} Df = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} - \frac{1}{8}\begin{bmatrix} 3\alpha^3 + 16x_2 - 8x_1 \\ -3\alpha^3 - 8x_2 + 16x_1 \end{bmatrix}$

$$= \begin{bmatrix} 2\alpha - \frac{3}{8}\alpha^3 \\ -2\alpha + \frac{3}{8}\alpha^3 \end{bmatrix} = \left(2\alpha - \frac{3}{8}\alpha^3\right)\begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

From $x = 2\gamma\begin{bmatrix} 1 \\ -1 \end{bmatrix}$, update is:

$$\cancel{\tilde{x} = 2\gamma\begin{bmatrix} 1 \\ -1 \end{bmatrix} - \frac{\gamma}{8}\begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}\begin{bmatrix} -1 \\ 1 \end{bmatrix}}$$

$$\tilde{x} = \left(2\gamma - \frac{3}{8}\gamma^3\right)\begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

If $|\gamma| < \frac{2}{\sqrt{3}}$ then

$$\left|2\gamma - \tfrac{3}{8}\gamma^3\right| = \left|\gamma\left(2 - \tfrac{3}{8}\gamma^2\right)\right|$$
$$\geq |\gamma|\left(2 - \tfrac{1}{2}\right)$$
$$= \tfrac{3}{2}|\gamma|$$

Thus, $|\gamma|$ is increasing by 50% at each iteration, so eventually $|\gamma| > \frac{2}{\sqrt{3}}$.

Once $|\gamma| > \frac{2}{\sqrt{3}}$ the Hessian is positive definite.

Then we are in the case from Q1, part (c), and converge quadratically to the global min.