Solution Assignment 1

1(a)

We have $H := D^2F(x)\langle u,v\rangle \in \mathbb{R}^n$ with $H_i = D^2F_i(x)\langle u,v\rangle$. As explained in class, for a scalar function F_i the second derivative D^2F_i is a quadratic form with the Hessian matrix $\left(\frac{\partial^2 F_i}{\partial x_j \partial x_k}\right)_{j,k=1,\ldots,n}$ so that $H_i = \sum_{j,k=1\ldots n} \frac{\partial^2 F_i(x)}{\partial x_j \partial x_k} u_j v_k$ and

$$|H_i| \le \sum_{j,k=1\ldots n} \left|\frac{\partial^2 F_i(x)}{\partial x_j \partial x_k}\right| |u_j|\,|v_k| \le \sum_{j,k=1\ldots n} \left|\frac{\partial^2 F_i(x)}{\partial x_j \partial x_k}\right| \|u\|_\infty \|v\|_\infty = C\,\|u\|_\infty \|v\|_\infty$$

$$\|H\|_\infty = \max_{i=1\ldots n} |H_i| \le C\,\|u\|_\infty \|v\|_\infty$$

1(b)

Newton-Kantorovich theorem from class: Let X, Y Banach spaces. Let $F\colon \Omega \to Y$ be continuously differentiable with $\Omega \subset X$ open and convex. For the initial guess $x_0 \in \Omega$ let $F'(x_0)$ be invertible. Assume that

$$\left\|F'(x_0)^{-1}F(x_0)\right\| \le \alpha \tag{1}$$

$$\left\|F'(x_0)^{-1}\left(F'(y) - F'(x)\right)\right\| \le \omega_0 \|y - x\| \qquad \text{for } x, y \in \Omega \tag{2}$$

$$h_0 := \alpha\omega_0 \le \tfrac{1}{2} \tag{3}$$

$$B := \{x \mid \|x - x_0\| \le \rho_-\} \subset \Omega \qquad \text{with } \rho_- := \left(1 - \sqrt{1 - 2h_0}\right)/\omega_0 \tag{4}$$

Then the Newton sequence x_k is well defined, $x_k \in B$ and converges to $x_* \in B$ with $f(x_*) = 0$. For $h_0 < \frac{1}{2}$ the convergence is r-quadratic, for $h_0 = \frac{1}{2}$ the convergence is r-linear.

Let $F(x) = \begin{bmatrix} x_1 - \cos(x_1 + x_2)/3 \\ x_2 - \sin(x_1 - x_2)/3 \end{bmatrix}$. Then $DF(x) = I + \frac{1}{3}\begin{bmatrix} \sin(x_1+x_2) & \sin(x_1+x_2) \\ -\cos(x_1-x_2) & \cos(x_1-x_2) \end{bmatrix}$, $D^2F_1(x) = \frac{1}{3}\cos(x_1 + x_2)\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$, $D^2F_2(x) = \frac{1}{3}\sin(x_1 - x_2)\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$. We use the ∞-vector norm and the induced norms for linear maps DF ("row sum norm") and bilinear maps D^2F (we denote all these norms by $\|\cdot\|_\infty$). Since $F \in C^2$ and Ω is convex we have $\|F'(x) - F'(y)\| \le \gamma \|x - y\|$ for $x, y \in \Omega$ with $\gamma = \max_{x\in\Omega} \left\|D^2F(x)\right\|_\infty$. From (a) we obtain $\left\|D^2F(x)\right\|_\infty = \max_{i=1\ldots n} \sum_{j,k=1\ldots n} \left|\frac{\partial^2 F_i(x)}{\partial x_j \partial x_k}\right| \le \frac{4}{3}$ since $|\sin t| \le 1$, $|\cos t| \le 1$ for any $x \in \mathbb{R}^2$. Hence $\gamma = \frac{4}{3}$. We have $F(x_0) = \begin{bmatrix} -\frac{1}{3} \\ 0 \end{bmatrix}$, $F'(x_0) = \begin{bmatrix} 1 & 0 \\ -\frac{1}{3} & \frac{4}{3} \end{bmatrix}$, $F'(x_0)^{-1} = \begin{bmatrix} 1 & 0 \\ \frac{1}{4} & \frac{3}{4} \end{bmatrix}$ and $F'(x_0)^{-1}F(x_0) = \begin{bmatrix} -1/3 \\ -1/12 \end{bmatrix}$, hence $\alpha = \frac{1}{3}$. We use $\left\|F'(x_0)^{-1}\left(F'(y) - F'(x)\right)\right\|_\infty \le \left\|F'(x_0)^{-1}\right\|_\infty \gamma \|y - x\|_\infty$ and get $\omega_0 = \left\|F'(x_0)^{-1}\right\|_\infty \gamma = \frac{4}{3}$. Therefore $h_0 := \alpha\omega_0 = \frac{4}{9} < \frac{1}{2}$ and $\rho_- = \left(1 - \sqrt{\frac{1}{9}}\right)/\left(\frac{4}{3}\right) = \frac{1}{2}$. We use the theorem with $\Omega = \mathbb{R}^2$. Then conditions (1)–(4) are satisfied, and we obtain that x_k converges r-quadratically to a solution x_* with $\|x_*\|_\infty \le \rho_- = \frac{1}{2}$, i.e., x_* is contained in the square $[-\frac{1}{2}, \frac{1}{2}] \times [-\frac{1}{2}, \frac{1}{2}]$.

1(c)

1(d)

2(a)

(A1): There is $x_* \in \Omega$ with $F(x_*) = 0$, (A2): $F'(x)$ satisfies Lipschitz condition: $\|F'(x) - F'(y)\| \le \gamma \|x - y\|$ for $x, y \in \Omega$, (A3): $F'(x_*)^{-1}$ exists.

Assume first that for k we have

$$\gamma \left\|F'(x_*)^{-1}\right\| \|x_k - x_*\| \le q < \frac{2}{3} \tag{5}$$

We redo (L2) for this assumption:

$$\left\|F'(x_k)^{-1}\right\| \le \frac{\left\|F'(x_*)^{-1}\right\|}{1 - \|I - F'(x_*)^{-1}F'(x_k)\|} \le \frac{\left\|F'(x_*)^{-1}\right\|}{1-q}$$

where we used

$$\left\|I - F'(x_*)^{-1}F'(x_k)\right\| = \left\|F'(x_*)^{-1}\left[F'(x_*) - F'(x_k)\right]\right\| \le \left\|F'(x_*)^{-1}\right\| \gamma \|x_* - x_k\| \le q.$$

Under assumption (5) we can then estimate $\|x_{k+1} - x_*\|$ in terms of $\|x_k - x_*\|$:

$$x_{k+1} - x_* = x_k - x_* - F'(x_k)^{-1}F(x_k) = F'(x_k)^{-1}\int_0^1 \left[F'(x_k) - F'\left(x_* + t(x_k - x_*)\right)\right](x_k - x_*)dt$$

$$\|x_{k+1} - x_*\| \le \left\|F'(x_k)^{-1}\right\| \int_0^1 \gamma(1-t)\|x_k - x_*\|^2 dt = \left\|F'(x_k)^{-1}\right\| \gamma \tfrac{1}{2} \|x_k - x_*\|^2 \le \frac{\left\|F'(x_*)^{-1}\right\| \gamma \|x_k - x_*\|^2}{2(1-q)}$$

$$\|x_{k+1} - x_*\| \le K \|x_k - x_*\|^2 \tag{6}$$

with $K := \left\|F'(x_*)^{-1}\right\| \gamma/(2-2q)$.

According to the problem, we initially have that (5) holds for $k = 0$. Therefore

$$K\|x_0 - x_*\| = \frac{\gamma \left\|F'(x_*)^{-1}\right\| \|x_0 - x_*\|}{2-2q} \le \frac{q}{2-2q} =: Q < \frac{\frac{2}{3}}{2 - \frac{4}{3}} = 1$$

We claim that for all $k = 0, 1, 2, \ldots$ we have (5) and

$$K\|x_k - x_*\| \le Q^{\left(2^k\right)} \tag{7}$$

We use induction: Both statements hold for $k = 0$. Assume that (5), (7) hold for k. Then $\|x_{k+1} - x_*\| \le (K\|x_k - x_*\|)\|x_k - x_*\| \le Q^{\left(2^k\right)}\|x_k - x_*\|$ which implies that (5) holds for $k+1$. Furthermore

$$K\|x_{k+1} - x_*\| \le (K\|x_k - x_*\|)^2 \le \left(Q^{\left(2^k\right)}\right)^2$$

which is (7) for $k+1$.

Now (7) implies that $\lim_{k\to\infty} x_k = x_*$. We also have (6). So both requirements for q-quadratic convergence are satisfied.

2(b)

$F'(x) = 1/(x^2+1)$, $F''(x) = -2x/(x^2+1)^2$. We have $F'(x_*)^{-1} = 1$. For the Lipschitz constant γ we use that for $F \in C^2$ we have $\gamma = \max_x \|F''(x)\|$. We can use the maximum over $x \in \mathbb{R}$: $F'''(x) = 0$ for $x = \pm 1/\sqrt{3}$, and $\left|F''(\pm 1/\sqrt{3})\right| = \frac{3}{8}\sqrt{3}$. As $|F''(x)|$ is increasing for $x < -\frac{3}{8}\sqrt{3}$ and decreasing for $x > \frac{3}{8}\sqrt{3}$ it is easy to see that $|F''(x)| \le \frac{3}{8}\sqrt{3}$ for all $x \in \mathbb{R}$.

So we obtain the condition $\gamma \left\|F'(x_0)^{-1}\right\| \|x_0 - x_*\| < \frac{2}{3}$ or $\frac{3}{8}\sqrt{3}\,|x_0| < \frac{2}{3}$ or $|x_0| < \frac{16}{9\sqrt{3}} \approx 1.0264$.

3(a)

Assume that $t_k < \rho_-$. As $f''(t) = 1 > 0$ we have $p(t) := f(t_k) + f'(t_k)(t - t_k) < f(t)$ for $t \ne t_k$ by Taylor's theorem, i.e., the tangent line is below the graph of f. We have $p(t_k) = f(t_k) > 0$, $p(\rho_-) < f(\rho_-) = 0$, so therefore the zero t_{k+1} of $p(t_k)$ must be in (t_k, ρ_-). The sequence t_k is increasing and bounded from above, hence it has a limit $t_* \le \rho_-$. Assume $t_* < \rho_-$, then taking the limit in $t_{k+1} = t_k - f(t_k)/f'(t_k)$ gives $t_* = t_* - f(t_*)/f'(t_*)$ with $f(t_*) > 0$, $f'(t_*) < 0$ which is a contradiction.

3(b)

For $h_0 = \frac{1}{2}$ we have $f(t) = \frac{1}{2}(t-1)^2$ and $t_{k+1} = t_k - \frac{\frac{1}{2}(t_k-1)^2}{t_k-1}$, i.e., $(t_{k+1}-1) = \frac{1}{2}(t_k-1)$ which is q-linear convergence to $\rho_- = 1$ with factor $\frac{1}{2}$.

Case $h_0 < \frac{1}{2}$: Consider a function $f(t)$ with zero t_*. Taylor's theorem gives $p(t) := f(t_k) + f'(t_k)(t-t_k) = f(t) - \frac{1}{2}f''(\tau_k)(t-t_k)^2$, hence $p(t_*) = f(t_k) + f'(t_k)(t_* - t_k) = 0 - \frac{1}{2}f''(\tau_k)(t_*-t_k)^2$. Dividing by $f'(t_k)$ gives

$$(t_* - t_{k+1}) = -\frac{f''(\tau_k)}{2f'(t_k)}(t_* - t_k)^2$$

We use $t_* := \rho_-$ and have for our function f that $f''(\tau_k) = 1$, $f'(t_k) = t_k - 1$. As $t_k < \rho_-$ we have $|f'(t_k)| \geq |f'(\rho_-)| = \sqrt{1-2h_0}$ and

$$|t_{k+1} - \rho_-| \leq \frac{1}{2\sqrt{1-2h_0}}\,|t_k - \rho_-|^2\,.$$

3(c)

We have $a_{k+1} = a_k(1-h_k) = a_k - d_k$ and $\tau_k - \tau_0 = d_0 + \cdots + d_{k-1} = -(a_k - a_0)$, i.e., $\tau_k = 1 - a_k$.

$$d_k = a_k h_k = a_{k-1}(1-h_{k-1})\tfrac{1}{2}\frac{h_{k-1}^2}{(1-h_{k-1})^2} = \tfrac{1}{2}\frac{(a_{k-1}h_{k-1})^2}{a_{k-1}(1-h_{k-1})} = \tfrac{1}{2}\frac{d_{k-1}^2}{a_k} = \tfrac{1}{2}\frac{d_{k-1}^2}{1-\tau_k}.$$

On the other hand t_k is defined by $t_{k+1} = t_k - \frac{\frac{1}{2}t_k^2 - t_k + h_0}{t_k - 1} = \frac{\frac{1}{2}t_k^2 - h_0}{t_k-1}$. Multiplying by $t_k - 1$ and adding $\frac{1}{2}t_{k+1}^2$ gives $\frac{1}{2}t_{k+1}^2 - t_{k+1} + h_0 = \frac{1}{2}(t_{k+1} - t_k)^2$. Therefore $t_{k+1} - t_k = -\frac{\frac{1}{2}t_k^2 - t_k + h_0}{t_k - 1} = \frac{\frac{1}{2}(t_k - t_{k-1})^2}{1-t_k}$.

We have $\tau_o = t_0 = 0$ and $\tau_1 = t_1 = h_0$. We have for $k \geq 1$ that $t_{k+1} = t_k - \frac{1}{2}(t_k - t_{k-1})^2/(1-t_k)$ and $\tau_{k+1} = \tau_k - \frac{1}{2}(\tau_k - \tau_{k-1})^2/(1-\tau_k)$. Hence we obtain by induction that $t_k = \tau_k$ for all $k \geq 0$. Therefore $t_{k+1} - t_k = \tau_{k+1} - \tau_k = d_k$.

3(d)

$$\|x_0 - x_l\| \leq \sum_{j=0}^{l-1}\|x_{j+1} - x_j\| \leq \sum_{j=0}^{l-1} d_j \overset{\text{(c)}}{=} t_l \overset{\text{(a)}}{<} \rho_- \tag{8}$$

$$\|x_k - x_l\| \leq \sum_{j=k}^{l-1}\|x_{j+1} - x_j\| \leq \sum_{j=k}^{l-1} d_j \overset{\text{(c)}}{=} t_l - t_k \tag{9}$$

Since t_k converges, it is a Cauchy sequence: We can find for every $\varepsilon > 0$ an integer k_0 such that $k, l \geq k_0$ imply $|t_l - t_k| < \varepsilon$ and hence $\|x_k - x_l\| < \varepsilon$. So x_k is also a Cauchy sequence and has a limit x_*. Taking the limit $l \to \infty$ in (8), (9) gives $\|x_0 - x_*\| \leq \rho_-$ and

$$\|x_k - x_*\| \leq \rho_- - t_k$$

We have from (b) that $\rho_- - t_k$ converges q-linearly to 0 for $h_0 = \frac{1}{2}$, and it converges q-quadratically to 0 for $h_0 < \frac{1}{2}$. By the definition of r-convergence this means that x_k converges r-linearly for $h_0 = \frac{1}{2}$, and r-quadratically for $h_0 < \frac{1}{2}$.