MA40050: Numerical Optimisation & Large–Scale Systems

Model Solutions to Problem Sheet 2

---

1. (a) Since x_* is a local minimiser, there exists an $r > 0$ such that

$$0 \le f(x_* + h) - f(x_*) = \nabla f(x_*) \cdot h + o(|h|), \quad \text{for all } h \in B_r(0).$$

Hence,

$$0 \le \lim_{|h|\to 0} \left(\frac{1}{|h|} \nabla f(x_*) \cdot h + \frac{o(|h|)}{|h|} \right) = \nabla f(x_*) \cdot \hat{h}$$

where $\hat{h} = h/|h|$. Since, $h \in B_r(0)$ was arbitrary, we also have $0 \le \nabla f(x_*) \cdot (-\hat{h})$, which together with the above inequality implies

$$\nabla f(x_*) \cdot \hat{h} = 0, \quad \text{for all } \hat{h} \in \mathbb{R}^N \text{ with } |\hat{h}| = 1.$$

This is equivalent to $\nabla f(x_*) = 0$.

(b) Since $f(x_* + h) \ge f(x_*)$, for all $h \in B_r(0)$, and $\nabla f(x_*) \cdot h = 0$, we have

$$0 \le f(x_* + h) - f(x_*) - \nabla f(x_*) \cdot h = \tfrac{1}{2}\, h^T \nabla^2 f(x_*) h + o(|h|^2), \quad \text{for all } h \in B_r(0),$$

and so

$$0 \le \lim_{|h|\to 0} \left(\frac{1}{2|h|^2} h^T \nabla^2 f(x_*) h + \frac{o(|h|^2)}{|h|^2} \right) = \tfrac{1}{2}\, \hat{h}^T \nabla^2 f(x_*) \hat{h}$$

Since $h \in B_r(0)$ was arbitrary and the inequality also holds for any multiple of $\hat{h}$, it follows that $\nabla^2 f(x_*) \ge 0$ (i.e. $\nabla^2 f(x_*)$ is positive semidefinite).

2. (a)

$$\begin{aligned} f(x+h) &= \tfrac{1}{2}(x+h)^T A (x+h) - b^T(x+h) + c \\ &= \tfrac{1}{2}\left[x^T A x + 2x^T A h + h^T A h\right] - b^T x - b^T h + c \\ &= f(x) + (Ax - b)^T h + \tfrac{1}{2} h^T A h + 0. \end{aligned}$$

Hence, $f \in C^2(\mathbb{R}^N)$, $\nabla f(x) = Ax - b$ and $\nabla^2 f(x) = A$.

(b) Since A is spd, it follows from the Linear Algebra handout that A^{-1} exists and thus $x_* = A^{-1}b$ is the only critical point. Since $\nabla^2 f(x) = A > 0$ it follows from Proposition 2.8 that x_* is a strict local minimum.

To show that x_* is a global minimum, let $0 \ne h \in \mathbb{R}^N$ be arbitrary. Then

$$f(x_* + h) - f(x_*) = (Ax_* - b)^T h + \tfrac{1}{2} h^T A h = \tfrac{1}{2} h^T A h > 0.$$

So x_* is a global minimiser.

3. The root is $x_* = 0.567143290409784$.

(a) Fixpoint iteration: $x^{k+1} = G(x^k) = \exp(-x^k)$. Thus

k	x^k	$\lvert x^k - x_*\rvert$	$\lvert x^k - x_*\rvert / \lvert x^{k-1} - x_*\rvert$
0	1	0.43286	
1	0.367879441171442	0.19926	0.46034
2	0.692200627555346	0.12505	0.62757
3	0.500473500563637	0.06667	0.53315
4	0.606243535085597	0.03910	0.58647
5	0.545395785975027	0.02175	0.55627
6	0.579612335503379	0.01247	0.57333
7	0.560115461361089	0.00703	0.56375
8	0.571143115080177	0.00400	0.56899
9	0.564879347391050	0.00226	0.56500
10	0.568428725029061	0.00129	0.57080

Hence, $x^k \to x_*$ q-linearly with q-factor x_*.

(b) Newton iteration: $x_{k+1} = x_k - \dfrac{x_k - \exp(-x_k)}{1 + \exp(-x_k)}$. Thus

k	x_k	$\lvert x_k - x_*\rvert$	$\lvert x_k - x_*\rvert / \lvert x_{k-1} - x_*\rvert^2$
0	1	4.32857e-1	
1	0.537882842739990	2.92604e-2	0.156166
2	0.566986991405413	1.56299e-4	0.182556
3	0.567143285989123	4.42066e-9	0.180957
4	0.567143290409784	$< 1.0e-15$	

Hence, $x^k \to x_*$ q-quadratically, much faster than the fixpoint iteration.

4. (a) $F(x) = x^k$, $k \geq 2$, is differentiable on all of $\mathbb{R}^N$ and $F'(x) = kx^{k-1} \neq 0$, for all $x \neq 0$.

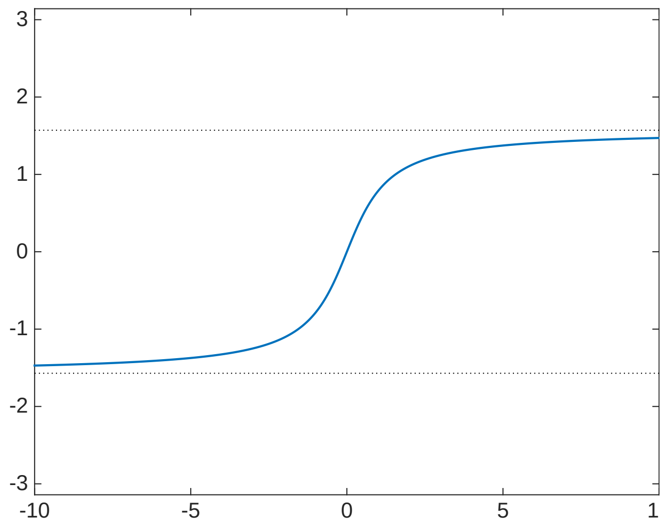
Hence, if $x_0 \neq 0$, then

$$x_{n+1} = x_n - \frac{F(x_n)}{F'(x_n)} = x_n - \frac{x_n^k}{kx_n^{k-1}} = \left(1 - \frac{1}{k}\right) x_n$$

is well-defined and converges to 0 q-linearly with q-factor $1 - \frac{1}{k}$.

(The convergence is not q-quadratic and becomes worse as k gets larger.)

(b) Choose $F(x) = \arctan(x)$:



To create a cyclic behaviour of the Newton iterates, we find an $x_0 \neq 0$ such that

$$-x_0 = x_1 = x_0 - \frac{F(x_0)}{F'(x_0)}. \qquad (1)$$

If such an x_0 exists, then $x_{2n} = x_0$ and $x_{2n+1} = -x_0$, for all $n \geq 0$.

$$(1) \quad \Leftrightarrow \quad 2x_0 F'(x_0) = F(x_0) \quad \Leftrightarrow \quad \varphi(x_0) := \frac{2x_0}{1+x_0^2} = \arctan(x_0)$$

Since $\varphi(0) = 0$ and $\arctan(0) = 0$ and

$$\varphi'(0) = \frac{2}{1+x_0^2} - \frac{4x_0}{(1+x_0^2)^2}\bigg|_{x_0=0} = 2 \quad \text{and} \quad \arctan'(0) = 1,$$

there exists an $x' > 0$ with $\varphi(x') > \arctan(x')$.

However, we also have $\lim_{x\to\infty} \varphi(x) = 0$ and $\lim_{x\to\infty} \arctan(x) = \frac{\pi}{2}$. Hence, there exists a $x'' > x'$ with $\varphi(x'') < \arctan(x'')$.

Since φ and $\arctan$ are continuous, there exists $x_0 \in (x', x'')$ where (1) holds.

5. Model code is available on the website.

   The minimum is $x_* = (1,1)^T$. Here is a table of results obtained with the model code (rounded to 5 significant figures):

n	x_n	$\lvert x_n - x_* \rvert$	$\lvert x_n - x_* \rvert / \lvert x_{n-1} - x_* \rvert^2$
0	$(-2, 2)^T$	3.1623	
1	$(-1.9925, 3.9701)^T$	4.2162	0.421614
2	$(0.96687, -7.8232)^T$	8.8232	0.49635
3	$(0.96689, 0.93488)^T$	7.3054e-2	9.3841e-4
4	$(1.0, 0.99890)^T$	1.0962e-3	0.20540
5	$(1.0, 1.0)^T$	9.5122e-10	7.9159e-4

The convergence is very erratic at first and then appears to be quadratic as predicted.

Note that $\sigma(\nabla^2 f(x_*)) = \{0.39936, 1001.6\}$ and so $\nabla^2 f(x_*) > 0$ and x_* is a strict minimiser.