In the proof of Kantorovitch's theorem we neglected to justify uniqueness. The proof given below is similar to the proof of uniqueness in Section A.4 (part 3, pages 600-601).

Now we will prove uniqueness. To prove that the solution in U_0 is unique, we will prove that if $\mathbf{y} \in U_0$ and $\mathbf{f}(\mathbf{y}) = \mathbf{0}$, then

$$|\mathbf{y} - \mathbf{a}_{i+1}| \le \frac{1}{2}|\mathbf{y} - \mathbf{a}_i|. \qquad A2.38$$

First, set

$$\mathbf{0} = \mathbf{f}(\mathbf{y}) = \mathbf{f}(\mathbf{a}_i) + [\mathbf{Df}(\mathbf{a}_i)](\mathbf{y} - \mathbf{a}_i) + \vec{\mathbf{r}}_i, \qquad A2.39$$

where $\vec{\mathbf{r}}_i$ is the remainder necessary for the second equality to be true. This gives

$$\mathbf{y} - \mathbf{a}_i = \overbrace{-[\mathbf{Df}(\mathbf{a}_i)]^{-1}\mathbf{f}(\mathbf{a}_i)}^{\substack{+\vec{\mathbf{h}}_i \text{ by earlier part} \\ \text{of proof}}} - [\mathbf{Df}(\mathbf{a}_i)]^{-1}\vec{\mathbf{r}}_i, \qquad A2.40$$

which we can rewrite as

$$\mathbf{y} - \overbrace{(\mathbf{a}_i + \vec{\mathbf{h}}_i)}^{\mathbf{a}_{i+1}} = \mathbf{y} - \mathbf{a}_{i+1} = -[\mathbf{Df}(\mathbf{a}_i)]^{-1}\vec{\mathbf{r}}_i. \qquad A2.41$$

Now we return to Equation A2.39. By Proposition A2.1, since $\mathbf{y}, \mathbf{a}_i \in U_0$, we have

$$|\vec{\mathbf{r}}_i| = \Big|\underbrace{\mathbf{f}(\mathbf{y}) - \mathbf{f}(\mathbf{a}_i)}_{\text{increment to } \mathbf{f}} - \underbrace{[\mathbf{Df}(\mathbf{a}_i)](\mathbf{y} - \mathbf{a}_i)}_{\substack{\text{linear approx.} \\ \text{of increment to } \mathbf{f}}}\Big| \le \frac{M}{2}|\mathbf{y} - \mathbf{a}_i|^2. \qquad A2.42$$

Next replace the $\vec{\mathbf{r}}_i$ in Equation A2.41 by the $\frac{M}{2}|\mathbf{y} - \mathbf{a}_i|^2$ of Equation A2.42, and take absolute values, to get

$$|\mathbf{y} - \mathbf{a}_{i+1}| \le \left|[\mathbf{Df}(\mathbf{a}_i)]^{-1}\right| \frac{M}{2}|\mathbf{y} - \mathbf{a}_i|^2. \qquad A2.43$$

Now we will prove Equation A2.38 by induction. To start the induction, note that

$$\begin{aligned}
|\mathbf{y} - \mathbf{a}_1| &\le \left|[\mathbf{Df}(\mathbf{a}_0)]^{-1}\right| \frac{M}{2}|\mathbf{y} - \mathbf{a}_0|^2 \\
&\le \left|[\mathbf{Df}(\mathbf{a}_0)]^{-1}\right| \frac{M}{2}|2\vec{\mathbf{h}}_0||\mathbf{y} - \mathbf{a}_0| \\
&\le \left|[\mathbf{Df}(\mathbf{a}_0)]^{-1}\right| M \underbrace{|\mathbf{f}(\mathbf{a}_0)|\left|[\mathbf{Df}(\mathbf{a}_0)]^{-1}\right|}_{\ge|\vec{\mathbf{h}}_0| \text{ by Eq. A2.1}} |\mathbf{y} - \mathbf{a}_0| \\
&\le \frac{1}{2}|\mathbf{y} - \mathbf{a}_0|.
\end{aligned} \qquad A2.44$$

(To get from the first to the second line of Equation A2.44, note that since $\mathbf{y} \in U_0$, it is in a ball of radius $\vec{\mathbf{h}}_0$, with center $\mathbf{a}_1$, and thus can be at most $2\vec{\mathbf{h}}_0$ away from $\mathbf{a}_0$. So we can replace one of the $|\mathbf{y} - \mathbf{a}_0|$ by $|2\vec{\mathbf{h}}_0|$. Going from the third to the fourth line uses Equation A2.3.)

Now assume by induction that

$$|\mathbf{y} - \mathbf{a}_j| \le \frac{1}{2}|\mathbf{y} - \mathbf{a}_{j-1}| \quad \text{for all } j \le i, \qquad A2.45$$

and rewrite Equation A2.43, dividing each side by $|\mathbf{y} - \mathbf{a}_i|$: The inductive hypothesis is used to replace the $|\mathbf{y} - \mathbf{a}_i|$ in the first line of Equation A2.46 by $\frac{1}{2}|\mathbf{y} - \mathbf{a}_{i-1}|$ in the second line. Lemma A2.3 justifies replacing $|[\mathbf{Df}(\mathbf{a}_i)]^{-1}|$ in the first line by $2|[\mathbf{Df}(\mathbf{a}_{i-1})]^{-1}|$ in the second line.

$$\begin{aligned}
\frac{|\mathbf{y} - \mathbf{a}_{i+1}|}{|\mathbf{y} - \mathbf{a}_i|} &\le \left|[\mathbf{Df}(\mathbf{a}_i)]^{-1}\right| \frac{M}{2}|\mathbf{y} - \mathbf{a}_i| \\
&\le 2\left|[\mathbf{Df}(\mathbf{a}_{i-1})]^{-1}\right| \frac{M}{2}\frac{|\mathbf{y} - \mathbf{a}_{i-1}|}{2} \\
&\le \dots \\
&\le \frac{1}{2}M|\mathbf{y} - \mathbf{a}_0| \left|[\mathbf{Df}(\mathbf{a}_0)]^{-1}\right| \\
&\le \frac{1}{2}M|2\vec{\mathbf{h}}_0| \left|[\mathbf{Df}(\mathbf{a}_0)]^{-1}\right| | \le |\mathbf{f}(\mathbf{a}_0)||[\mathbf{Df}(\mathbf{a}_0)]^{-1}|^2 M \\
&\le \frac{1}{2} \qquad \text{(We again use Equation A2.3.)}
\end{aligned} \qquad A2.46$$

Thus $|\mathbf{y} - \mathbf{a}_{i+1}| \le \frac{1}{2}|\mathbf{y} - \mathbf{a}_i|$. This proves that $\mathbf{y} = \lim \mathbf{a}_i$, and that $\lim \mathbf{a}_i$ is the unique solution of $\mathbf{f}(\mathbf{x}) = \mathbf{0}$ in U_0. □

Remark A2.5. If we change the hypotheses of Kantorovitch's theorem by defining a larger ball U_{-1}:

$$U_{-1} = \{\mathbf{x} | |\mathbf{x} - \mathbf{x}_0| \le 2|\vec{\mathbf{h}}_0|\},$$

and require that U_{-1} be a subset of U, and that the Lipschitz condition A2.2 holds for all $\mathbf{u}_1, \mathbf{u}_2 \in U_{-1}$, then we can strengthen the conclusion to say that the equation $\mathbf{f}(\mathbf{x}) = 0$ has a unique solution in U_{-1}, and that Newton's method with initial guess $\mathbf{x}_0$ converges to it. The proof is exactly identical. △

(The smaller the set in which one can guarantee existence, the better: it is a stronger statement to say that there exists a William Ardvark in the town of Nowhere, NY, population 523, than to say there exists a William Ardvark in New York State.

The larger the set in which one can guarantee uniqueness, the better: it is a stronger statement to say there exists a unique John W. Smith in the state of California than to say there exists a unique John W. Smith in Tinytown, CA.

There are times, such as when proving uniqueness for the inverse function theorem, that one wants the Kantorovitch theorem stated for the larger ball U_{-1}. There are other times when the function is not Lipschitz on the larger space, or is not even defined on the larger space, and the original Kantorovitch theorem is the useful one.)