The Inverse and Implicit Function Theorems.

Proposition. Suppose X and Y are normed vector spaces and L is a linear isomorphism from X onto Y. Then
$$\frac{1}{||L^{-1}||} = \inf\{|L(x)| : x \in X \text{ and } |x| = 1\}.$$

Remark. In what follows $1/\infty = 0$ and $1/\infty = 0$.
Proof. Set $\beta = \inf\{|L(x)| : x \in X \text{ and } |x| = 1\}$.
For any $x \in X$ such that $|x| = 1$ we have
$$1 = |L^{-1}(L(x))| \leq ||L||^{-1}||\,|L(x)|$$
which implies that $1/||L^{-1}|| \leq \beta$.
For any $y \in Y$ we have that
$$|y| = |L(L^{-1}(y))| \geq \beta||L^{-1}(y)|$$
which implies that $||L^{-1}|| \leq 1/\beta$. $\square$

The Inverse Function Theorem. Suppose
(1) X and Y are Banach spaces, $a \in X$, $0 < R < \infty$, $B = \{x \in X : |x - a| \leq R\}$ and
$$f : B \to Y.$$
(2) L is a linear isomorphism from X onto Y, $||L|| < \infty$ and
$$p(x) = f(x) - [f(a) + L(x - a)] \quad \text{for } x \in B.$$
(3)
$$\alpha < \beta$$
where
$$\alpha = \mathbf{Lip}(p) \quad \text{and} \quad \beta = \inf\{|L(x)| : x \in X \text{ and } |x| = 1\}.$$

Then

(4) f^{-1} is a function and $\mathbf{Lip}(f^{-1}) \leq (\beta - \alpha)^{-1}$;

(5)
$$\{f(a) + L(h) : |h| \leq (1 - \alpha/\beta)R\} \subset \mathbf{rng}(f) \subset \{f(a) + L(h) : |h| \leq (1 + \alpha/\beta)R\};$$

(6) if f is differentiable at a with differential L then f^{-1} is differentiable at $f(a)$ with differential L^{-1}.

Remark. It is true and nontrivial that, if L is a linear isomorphism from the Banach space X onto the Banach space Y, the boundedness of L is equivalent to the boundedness of L^{-1}. Thus there is some redundancy in our hypotheses.

Proof. By virtue of the previous Proposition we have that $||L^{-1}|| = 1/\beta$. Note also that $p(a) = 0$.
Suppose x_1 and x_2 are points of B. We have
$$L(x_2 - x_1) = L(x_2 - a) - L(x_1 - a) = f(x_2) - f(x_1) - \big(p(x_2) - p(x_2)\big)$$
so
$$\begin{aligned}|x_1 - x_2| &= |L^{-1}(L(x_2 - x_1))| \\ &\leq \beta^{-1}\big(|p(x_2) - p(x_1)| + |f(x_1) - f(x_1)|\big) \\ &\leq \beta^{-1}\big(\alpha|x_2 - x_1| + |f(x_2) - f(x_1)|\big)\end{aligned}$$

which proves (4).

Next we suppose $y \in \{f(a) + L(h) : |h| \le (1 - \alpha/\beta)R\}$ and let $C : B \to Y$ be such that

$$C(x) = a + L^{-1}(y - f(a) - p(x)) \quad \text{whenever } x \in B.$$

For any $x \in B$ we have

$$\begin{aligned} |C(x) - a| &= |L^{-1}(y - f(a)) + L^{-1}(p(a) - p(x))| \\ &\le |L^{-1}(y - f(a))| + |L^{-1}(p(a) - p(x))| \\ &\le (1 - \alpha/\beta)R + \alpha/\beta|x - a| \\ &\le R \end{aligned}$$

so

$$C[B] \subset B.$$

Furthermore, for any x_1 and x_2 in B we have

$$|C(x_1) - C(x_2)| = |L^{-1}(p(x_1) - p(x_2))| \le \frac{\alpha}{\beta}|x_1 - x_2|.$$

Thus, by the Contraction Mapping Principle, C has a unique fixed point x in B. Now

$$C(x) = x \Rightarrow x = a + L^{-1}(y - f(a) - p(x)) \Rightarrow L(x - a) = y - f(a) - p(x) \Rightarrow f(x) = y;$$

thus the first inclusion in (5) is proved.

To prove the second inclusion in (5), we suppose $x \in B$, set $h = L^{-1}(f(x) - f(a))$ and note that $h = (x - a) + L^{-1}(p(x) - p(a))$. Thus

$$|h| = |(x - a) + L^{-1}(p(x) - p(a))| \le |x - a| + \frac{\alpha}{\beta}|x - a| = (1 + \frac{\alpha}{\beta})|x - a|.$$

Finally, suppose f is differentiable at a with differential L and let $\epsilon > 0$. Let $\epsilon_f = \beta\epsilon/(\beta - \alpha)$. Choose δ_f such that

$$x \in B \text{ and } |x - a| \le \delta_f \;\Rightarrow\; |f(x) - f(a) - L(x - a)| \le \epsilon_f|x - a|.$$

Let $\delta = \delta_f/(\beta - \alpha)$ and suppose $y \in \mathbf{rng}\, f$ and $|y - f(a)| \le \delta$. Set $x = f^{-1}(y)$. Then

$$|x - a| \le \mathbf{Lip}\,(f^{-1})|y - f(a)| \le \frac{1}{\beta - \alpha}|y - f(a)| \le \delta_f$$

so

$$\begin{aligned} |f^{-1}(y) - a &- L^{-1}(y - f(a)| \\ &= |L^{-1}\big(L(x - a) - f(x) - f(a)\big)| \\ &\le \beta\epsilon_f|x - a| \\ &\le \beta\epsilon_f\frac{1}{\beta - \alpha}|y - f(a)| \\ &= \epsilon|y - f(a)|; \end{aligned}$$

since we know from (5) that $f(a)$ is an interior point of $\mathbf{rng}(f)$ we conclude that f^{-1} is differentiable at $f(a)$ with differential L^{-1}. □

Corollary. Suppose

(1) X and Y are Banach spaces;

(2) $A \subset X$, $f : A \to Y$ and f is differentiable at each point of A;

(3) $a \in A$, ∂f is continuous at a and $\partial f(a)$ is a Banach space isomorphism from X onto Y.

Then there is an open subset U of A such that $a \in U$,

(4) $f[U]$ is an open subset of Y;

(5) $f|U$ is univalent;

(6) $(f|U)^{-1}$ is differentiable at each point of $f[U]$.

Proof. We let

$$\beta(x) = \inf\{|\partial f(x)(v)| : v \in X \text{ and } |v| = 1\} \quad \text{whenever } x \in A.$$

We let

$$\alpha(x, r) = \sup\{||\partial f(y) - \partial f(x)|| : y \in A \cap \mathbf{B}_r(x)\} \quad \text{whenever } x \in A \text{ and } 0 < r < \infty.$$

Then

(7) $$\alpha(x, r) \leq 2\alpha(a, |x| + r) \quad \text{whenever } x \in A \text{ and } 0 < r < \infty.$$

Indeed, if $x \in A$, $0 < r < \infty$ and $y \in \mathbf{B}_r(x)$ we have

$$||\partial f(y) - \partial f(x)|| \leq ||\partial f(y) - \partial f(a)|| + ||\partial f(x) - \partial f(a)|| \leq \alpha(a, |x| + r) + \alpha(a, |x|).$$

Moreover, if $0 < r < \infty$, $x, y \in \mathbf{B}_r(a)$, $v \in X$ and $|v| = 1$ then

$$\begin{aligned} |\partial f(y)(v)| &\geq |\partial f(x)(v)| - |\partial f(y)(v) - \partial f(x)(v)| \\ &\geq \beta(x) - ||\partial f(y) - \partial f(x)|| \\ &\geq \beta(x) - ||\partial f(y) - \partial f(a)|| - ||\partial f(x) - \partial f(a)|| \\ &\geq \beta(x) - 2\alpha(a, r) \end{aligned}$$

which implies that

$$\beta(y) \geq \beta(x) - 2\alpha(a, r);$$

thus

(8) $$|\beta(y) - \beta(x)| \leq 2\alpha(a, r) \quad \text{whenever } 0 < r < \infty \text{ and } x, y \in \mathbf{B}_r(a).$$

Since A is open we may choose $R \in (0, \infty)$ such that $\mathbf{B}_{2R}(a) \subset A$. Next we choose b such that

$$0 < b < \beta(a);$$

this is possible by because $\partial f(a)$ is a Banach space isomorphism from X onto Y. Since $\partial f(a)$ is continuous at a we may use (7)and (8) to choose R_1 such that $0 < R_1 \leq R$,

(9) $$\alpha(x, R_1) < b \leq \beta(x) \quad \text{whenever } x \in \mathbf{B}_{R_1}(a).$$

To complete the proof of the Theorem we need only apply the Inverse Function Theorem with a, R, f there replaced by $x, R_1, f|\mathbf{B}_{R_1}(x)$ for each $x \in \mathbf{B}_{R_1}(a)$. □