Here we give a proof of the last part of the Inverse Function Theorem which says that the inverse function is continuously differentiable. This proof is non-examinable (but the statement is). For completeness we will give the whole proof including the parts given in lectures (which of course remain examinable).

Theorem (Inverse Function Theorem)**.** *Let $U \subset \mathbb{R}^n$ be an open set, let $f\colon U \to \mathbb{R}^n$ be a continuously differentiable function, and let $a \in U$. If $f'(a)$ is invertible, then there exist open neighbourhoods V of a with $V \subset U$ and W of $f(a)$ such that $f\restriction_V\colon V \to W$ is a bijection with a continuously differentiable inverse $g\colon W \to V$. Moreover, $g'(y) = \big[f'(g(y))\big]^{-1}$ for all $y \in W$.*

Proof. STEP 1: Without loss of generality, we may assume that $a = f(a) = 0$ and $f'(a) = I$, the identity map. Indeed, let $T = f'(a)$ and consider $h\colon U - a \to \mathbb{R}^n$ given by $h(x) = T^{-1}(f(x+a) - f(a))$. Then h is continuously differentiable with $h'(x) = T^{-1} \circ f'(x+a)$, $h(0) = 0$ and $h'(0) = T^{-1} \circ f'(a) = I$. If we know the result for h, then we can deduce it for f, since $f(x) = T(h(x-a)) + f(a)$.

We will now assume that $a = f(a) = 0$ and $f'(a) = I$. Since f' is continuous, we can fix $r > 0$ such that $B_r(0) \subset U$ and $\|f'(x) - I\| \leqslant \frac{1}{2}$ for all $x \in B_r(0)$.

STEP 2: For all $x, y \in B_r(0)$ we have $\|f(x) - f(y)\| \geqslant \frac{1}{2}\|x - y\|$. To see this, define $p\colon U \to \mathbb{R}^n$ by $p(x) = f(x) - x$. Then p is differentiable with $p'(x) = f'(x) - I$ for all $x \in U$, and hence $\|p'(x)\| \leqslant \frac{1}{2}$ for $x \in B_r(0)$. By the Mean Value Inequality, we have $\|p(x) - p(y)\| \leqslant \frac{1}{2}\|x - y\|$ for $x, y \in B_r(0)$. Hence by the triangle-inequality, we have

$$\frac{1}{2}\|x - y\| \geqslant \|h(x) - h(y)\| = \|x - y - (f(x) - f(y))\| \geqslant \|x - y\| - \|f(x) - f(y)\|$$

which implies the claim.

STEP 3: Set $s = r/2$. We claim that $D_s(0) \subset f\big(D_r(0)\big)$. More precisely, for each $w \in D_s(0)$ there exists a unique $x \in D_r(0)$ such that $f(x) = w$. To prove this, we fix $w \in D_s(0)$ and define $q(x) = w - p(x) = w - f(x) + x$ for $x \in B_r(0)$. [Note that $f(x) = w$ if and only if $q(x) = x$.] Since $p(0) = f(0) - 0 = 0$, it follows that

$$\|q(x)\| \leqslant \|w\| + \|p(x)\| = \|w\| + \|p(x) - p(0)\| < s + \tfrac{1}{2}\|x - 0\| \leqslant s + \tfrac{r}{2} = r\ .$$

Thus $q\big(B_r(0)\big) \subset D_r(0)$, and in particular q is a map on the non-empty complete metric space $B_r(0)$ (complete, as it is a closed subset of the complete metric space $\mathbb{R}^n$). Now, for $x, y \in B_r(0)$ we have

$$\|q(x) - q(y)\| = \|p(x) - p(y)\| \leqslant \frac{1}{2}\|x - y\|$$

and thus q is a contraction mapping. By the Contraction Mapping Theorem, q has a unique fixed point x in $B_r(0)$, and moreover $x = q(x) \in D_r(0)$ from above.

STEP 4: Set $W = D_s(0)$ and $V = f^{-1}(W) \cap D_r(0)$. Then V and W are open neighbourhoods of 0, $f\restriction_V$ is a bijection $V \to W$ whose inverse $g\colon W \to V$ is continuous.

Open balls are open and f is continuous, and so V and W are open and both contain 0 since $f(0) = 0$. By Step 3, $f\restriction_V$ is a bijection $V \to W$. We prove that the inverse $g\colon W \to V$ of $f\restriction_V$ is continuous. Indeed, given $u, v \in W$, set $x = g(u)$, $y = g(v)$ and use Step 2 to obtain

$$\|g(u) - g(v)\| = \|x - y\| \leqslant 2\|f(x) - f(y)\| = 2\|u - v\|\ .$$

Thus, g is 2-Lipschitz and, in particular, continuous. Before turning to the final step, we observe that for $x \in B_r(0)$, we have $\|f'(x) - I\| \leqslant \frac{1}{2}$, and hence

$$\|f'(x)(h)\| \geqslant \|h\| - \|f'(x)(h) - h\| \geqslant \frac{1}{2}\|h\| \ .$$

It follows that $f'(x)$ is invertible.

STEP 5: The inverse function $g\colon W \to V$ is continuously differentiable with $g'(y) = \big[f'(g(y))\big]^{-1}$ for all $y \in W$.

Fix $y \in W$. Set $x = g(y)$ and $T = f'(x)$. Then for suitable error function ε, we have

$$f(x + h) = f(x) + T(h) + \|h\|\varepsilon(h) \ .$$

Choose $\delta > 0$ such that $D_\delta(y) \subset W$. For $k \in D_\delta(0)$, define

$$h = h(k) = g(y + k) - g(y) \ .$$

Then $g(y+k) = g(y)+h = x+h$, and hence $y+k = f(x+h)$ and $k = f(x+h)-f(x)$. From above, it follows that $k = T(h)+\|h\|\varepsilon(h)$, and so $h = T^{-1}(k)-\|h\|T^{-1}\big(\varepsilon(h)\big)$. We then obtain

$$g(y + k) = g(y) + h = g(y) + T^{-1}(k) - \|h\|T^{-1}\big(\varepsilon(h)\big) \ .$$

The composite $T^{-1} \circ \varepsilon \circ h$ is 0 at 0 and continuous at 0, and in addition we have $\|h\| = \|g(y + k) - g(y)\| \leqslant 2\|k\|$ by STEP 4. Hence, we get

$$g(y + k) = g(y) + T^{-1}(k) + \mathrm{o}(\|k\|)$$

which shows that g is differentiable at y and $g'(y) = T^{-1} = \big[f'(g(y))\big]^{-1}$.

Finally, since $f' \circ g$ is continuous, and since $A \mapsto A^{-1}\colon \mathcal{G}_n \to \mathcal{M}_n$ is continuous (the entries of $A^{-1} = \frac{1}{\det A}\operatorname{adj} A$ are rational functions in the entries of A), it follows that $g'\colon W \to L(\mathbb{R}^n, \mathbb{R}^n)$ is continuous. □