Proof of Weiersrtass-Stone Theorem in general case:

Any algebra $\mathcal{A}$ of continuous functions on a compact metric space X that separates points and contains constants is dense in $C(X)$.

Step 1: Since the absolute value can be approximated by polynomials (proved in class) If $f \in \mathcal{A}$, then $|f| \in \overline{\mathcal{A}}$, the closure of $\mathcal{A}$.

Step 2: If $f, g \in \mathcal{A}$, then $\max(f, g)$ and $\min(f, g)$ are in $\overline{\mathcal{A}}$.

This follows immediately from Step 1, using the formulae

$$\max(f,g) = \frac{f+g}{2} + \frac{|f-g|}{2}$$
$$\min(f,g) = \frac{f+g}{2} - \frac{|f-g|}{2},$$

Step 3: Given any continuous function $f : X \to \mathbb{R}$, $x \in X$ and $\varepsilon > 0$ there exists a function $g_x \in \overline{\mathcal{A}}$ such that $g_x(x) = f(x)$ and $g_x(t) > f(t) - \varepsilon$ for all $t \in X$.

By the separation property end inclusion of constants in $\mathcal{A}$ for any $y \in X$ there exists $h_y \in \mathcal{A}$ such that $h_y(x) = f(x)$ and $h_y(y) = f(y)$. Since h_y is continuous, there exists an open neighborhood $I_y \ni y$ such that $h_y(t) > f(t) - \varepsilon$ for all $t \in I_y$. Since X is compact the cover of X by open sets I_y contains a finite subcover $I_{y_1}, I_{y_2}, \ldots, I_{y_n}$. Any $x \in X$ lies in some I_{y_k}. Take $g_x = \max(h_{y_1}, \ldots, h_{y_n})$. Then $g_x \in \overline{\mathcal{A}}$ with the required property.

Step 4: Given any continuous function $f : X \to \mathbb{R}$, $x \in X$ and $\varepsilon > 0$ there exists a function $h \in \overline{\mathcal{A}}$ such that $f(t) - \varepsilon < h(t) < f(t) + \varepsilon$ for all $t \in X$.

Now look at the collection $\{g_x\}$ constructed in Step 3. Since each function is continuous, by for each $x \in X$ there exists an open neighborhood $J_x \ni x$ on which $g_x(t) < f(t) + \varepsilon$. Use compactness and choose a finite subcover $J_{x_1}, J_{x_2}, \ldots, J_{x_m}$. Take $h = \min(g_{x_1}, \ldots, g_{x_m})$. Then $h \in \overline{\mathcal{A}}$ with the required property, and since also $h(t) > f(t) - \varepsilon$ for all $t \in X$, we have $\|f - h\| < \varepsilon$, proving the result.