NATIONAL UNIVERSITY OF SINGAPORE
MATHEMATICS SOCIETY

PAST YEAR PAPER SOLUTIONS
with credits to Teo Wei Hao, Poh Wei Shan Charlotte

MA4229 Approximation Theory
AY 2007/2008 Sem 1

SECTION A

Question 1

(a) For all $p, q \in P_n$, if $\|q\|_2 = 0$, then we have $|\langle p, q\rangle| = 0 = \|p\|_2\|q\|_2$.
Else we have $\|q\|_2 \neq 0$, and so $\left\langle p - \frac{\langle p,q\rangle}{\|q\|_2^2}q, p - \frac{\langle p,q\rangle}{\|q\|_2^2}q \right\rangle = \left\| p - \frac{\langle p,q\rangle}{\|q\|_2^2}q \right\|_2^2 \geq 0$. At the same time,

$$\begin{aligned}
\left\langle p - \frac{\langle p,q\rangle}{\|q\|_2^2}q, p - \frac{\langle p,q\rangle}{\|q\|_2^2}q \right\rangle &= \langle p,p\rangle - 2\left\langle p, \frac{\langle p,q\rangle}{\|q\|_2^2}q \right\rangle + \left\langle \frac{\langle p,q\rangle}{\|q\|_2^2}q, \frac{\langle p,q\rangle}{\|q\|_2^2}q \right\rangle \\
&= \|p\|_2^2 - 2\frac{\langle p,q\rangle}{\|q\|_2^2}\langle p,q\rangle + \frac{\langle p,q\rangle^2}{\|q\|_2^4}\|q\|^2 \\
&= \|p\|_2^2 - \frac{\langle p,q\rangle^2}{\|q\|_2^2}.
\end{aligned}$$

Thus, we have $\|p\|_2^2 \geq \frac{\langle p,q\rangle^2}{\|q\|_2^2}$, i.e. $\langle p,q\rangle^2 \leq \|p\|_2^2\|q\|_2^2$, and so $|\langle p,q\rangle| \leq \|p\|_2\|q\|_2$.

(b) We know that for the inner product given, an orthogonal basis for P_2 is the Legendre polynomials of degree up to 2, i.e. $P_0(x) = 1$, $P_1(x) = x$ and $P_2(x) = \frac{3}{2}x^2 - \frac{1}{2}$. We have,

$$\begin{aligned}
\langle 1, 1\rangle &= \int_{-1}^{1} 1\ dx = 2; \\
\langle x, x\rangle = \langle x^2, 1\rangle &= \int_{-1}^{1} x^2\ dx = \left[\frac{1}{3}x^3\right]_{-1}^{1} = \frac{2}{3}; \\
\langle x^2, x^2\rangle &= \int_{-1}^{1} x^4\ dx = \left[\frac{1}{5}x^5\right]_{-1}^{1} = \frac{2}{5}; \\
\langle 3x^2-1, 3x^2-1\rangle &= 9\langle x^2, x^2\rangle - 6\langle x^2, 1\rangle + \langle 1, 1\rangle \\
&= 9\left(\frac{2}{5}\right) - 6\left(\frac{2}{3}\right) + 2 = \frac{8}{5}.
\end{aligned}$$

Hence, after normalization, we get an orthonormal basis for P_2 to be $\left\{\frac{1}{\sqrt{2}}, \sqrt{\frac{3}{2}}x, \sqrt{\frac{5}{8}}(3x^2-1)\right\}$.

(c) We have,

$$\begin{aligned}
\langle e^x, 1\rangle &= \int_{-1}^{1} e^x\ dx = [e^x]_{-1}^{1} = e - e^{-1}; \\
\langle e^x, x\rangle &= \int_{-1}^{1} xe^x\ dx = [xe^x - e^x]_{-1}^{1} = -2e^{-1}; \\
\langle e^x, x^2\rangle &= \int_{-1}^{1} x^2e^x\ dx = \left[x^2e^x - 2xe^x + 2e^x\right]_{-1}^{1} = e + 5e^{-1}.
\end{aligned}$$

Thus, we have the least squares approximation to f to be,

$$\begin{aligned}
p_2^*(x) &= \frac{\langle e^x, 1\rangle}{\|1\|_2^2}(1) + \frac{\langle e^x, x\rangle}{\|x\|_2^2}(x) + \frac{\langle e^x, 3x^2-1\rangle}{\|3x^2-1\|_2^2}(3x^2-1) \\
&= \frac{1}{2}(e - e^{-1}) - 3e^{-1}x + \frac{5}{8}(2e + 16e^{-1})(3x^2-1) \\
&= \frac{15e + 120e^{-1}}{4}x^2 - 3e^{-1}x - \frac{3e + 42e^{-1}}{4}.
\end{aligned}$$

Question 2

(a) Notice that for all $k \in \{0, 1, \dots, n\}$, we have $\ell_k \in P_n$. Note that for $k, i \in \{0, 1, \dots, n\}$, we have,

$$\ell_k(x_i) = \begin{cases} 1 & \text{if } k = i; \\ 0 & \text{otherwise.} \end{cases}$$

Thus, for all $k \in \{0, 1, \dots, n\}$, we have $\sum_{j=0}^{n} \ell_j(x_k) = 1$.

This shows that $x_0, x_1, \cdots, x_n$ are roots for $\sum_{j=0}^{n} \ell_j(x) - 1$ and thus gives us $\sum_{j=0}^{n} \ell_j(x) - 1$ to represent a polynomial of order n with at least $n+1$ roots.

Therefore $\sum_{j=0}^{n} \ell_j(x) - 1 = 0$, i.e. $\sum_{j=0}^{n} \ell_j(x) = 1$ for all $x \in \mathbb{R}$.

(b) Let $f : [a, b] \to \mathbb{R}$ be such that $f(x) = x^{n+1} - \sum_{i=0}^{n} x_i^{n+1}\ell_i(x)$. Notice that $f \in P_{n+1}$.

For all $k \in \{0, 1, \dots, n\}$, we have $f(x_k) = x_k^{n+1} - \sum_{i=0}^{n} x_i^{n+1}\ell_i(x_k) = x_k^{n+1} - x_k^{n+1} = 0$.

Thus $\{x_0, x_1, \dots, x_n\}$ are roots of f.

Since the leading coefficient of f is 1, we conclude that $f(x) = \prod_{i=0}^{n}(x - x_i)$.

Thus $x^{n+1} = \sum_{i=0}^{n} x_i^{n+1}\ell_i(x) + \prod_{i=0}^{n}(x - x_i)$ for all $x \in \mathbb{R}$.

(c) Let $p \in P_n$ such that $p(x) = \sum_{i=0}^{n} f(x_i)\ell_i(x)$ for all $x \in \mathbb{R}$.

For all $k \in \{0, 1, \dots, n\}$, we have $f(x_k) = \sum_{i=0}^{n} f(x_i)\ell_i(x_k) = p(x_k)$, and so p is the polynomial of degree at most n that interpolates f at the $n+1$ distinct nodes $x_0, x_1, \dots, x_n$.

Let $x \in [a, b]$, and q be the polynomial of degree at most $n+1$ that interpolates f at the $n+2$ distinct nodes $x_0, x_1, \dots, x_n, x$. Using Newton form of interpolating polynomials, we will have, for every $t \in [a,b]$,

$$\begin{aligned}
p(t) &= f[x_0] + f[x_0,x_1](t-x_0) + \cdots + f[x_0,x_1,\ldots,x_n]\prod_{j=0}^{n-1}(t-x_j),\\
q(t) &= f[x_0] + f[x_0,x_1](t-x_0) + \cdots\\
&\quad + f[x_0,x_1,\ldots,x_n]\prod_{j=0}^{n-1}(t-x_j) + f[x_0,x_1,\ldots,x_n,x]\prod_{j=0}^{n}(t-x_j)\\
&= p(t) + f[x_0,x_1,\ldots,x_n,x]\prod_{j=0}^{n}(t-x_j).
\end{aligned}$$

Since q interpolates f at $\{x_0,x_1,\ldots,x_n,x\}$, we have $n+2$ roots for $f-q$.
By the Mean Value Theorem, we note that between every 2 consecutive roots of $f-q$, there exists a root of $f'-q'$. Hence, $f'-q'$ has $n+1$ roots. By recursive application of the Mean Value Theorem, $f''-q''$ has n roots; etc, and we obtain that $f^{(n+1)}-q^{(n+1)}$ has 1 root, i.e. there exists $\mu \in [a,b]$ such that,

$$0 = f^{(n+1)}(\mu) - q^{(n+1)}(\mu) = f^{(n+1)}(\mu) - (n+1)!f[x_0,x_1,\ldots,x_n,x],$$

i.e. $f[x_0,x_1,\ldots,x_n,x] = \dfrac{f^{(n+1)}(\mu)}{(n+1)!}$.
Thus for all $x \in [a,b]$, we have,

$$\begin{aligned}
|f(x)-p(x)| = |q(x)-p(x)| &= \left| f[x_0,x_1,\ldots,x_n,x]\prod_{j=0}^{n}(x-x_j)\right|\\
&= \frac{|f^{(n+1)}(\mu)|}{(n+1)!}\prod_{j=0}^{n}|x-x_j|\\
&\le \frac{\|f^{(n+1)}\|_\infty}{(n+1)!}(b-a)^{n+1}.
\end{aligned}$$

Question 3

(a) For $k \in \{1,2,3\}$, let $C_k, S_k \in T_3$ be such that $C_k(\theta) = \cos k\theta$, $S_k(\theta) = \sin k\theta$, $\theta \in [-\pi,\pi]$.
Let $C_0 \in T_3$ be the constant function $C_0(\theta) = 1$, $\theta \in [-\pi,\pi]$.
We know that an orthonormal basis for T_3 is $\left\{\frac{1}{\sqrt{2\pi}}C_0, \frac{1}{\sqrt{\pi}}C_1, \frac{1}{\sqrt{\pi}}S_1, \frac{1}{\sqrt{\pi}}C_2, \frac{1}{\sqrt{\pi}}S_2, \frac{1}{\sqrt{\pi}}C_3, \frac{1}{\sqrt{\pi}}S_3\right\}$.

We know that $C_6 : [-\pi,\pi] \to \mathbb{R}$ such that $C_6(\theta) = \cos 6\theta$ is orthogonal to T_3.
Using the identity that for all $x \in \mathbb{R}$, $\cos 3x = 4\cos^3 x - 3\cos x$, we have,

$$f(\theta) - \frac{3}{4}C_2(\theta) = \cos^3 2\theta - \frac{3}{4}\cos 2\theta = \frac{1}{4}\cos 6\theta = \frac{1}{4}C_6(\theta),$$

and so $f - \frac{3}{4}C_2$ is orthogonal to T_3.
This implies that $\frac{3}{4}C_2 \in T_3$ is the least squares approximation to f in T_3, i.e. $q_3^* = \frac{3}{4}C_2$.

(b) Using the trigonometric identity $\cos(x-y) = \cos x \cos y + \sin x \sin y$ for all $x, y \in \mathbb{R}$, we have,

$$\begin{aligned}
S_n g(\theta) &= \frac{1}{\pi}\left(\frac{a_0}{2} + \sum_{k=1}^{n}(a_k \cos k\theta + b_k \sin k\theta)\right) \\
&= \frac{1}{\pi}\left(\frac{1}{2}\int_{-\pi}^{\pi} g(\phi)\ d\phi + \sum_{k=1}^{n}\left(\left(\int_{-\pi}^{\pi} g(\phi)\cos k\phi\ d\phi\right)\cos k\theta + \left(\int_{-\pi}^{\pi} g(\phi)\sin k\phi\ d\phi\right)\sin k\theta\right)\right) \\
&= \frac{1}{\pi}\int_{-\pi}^{\pi} g(\phi)\left(\frac{1}{2} + \sum_{k=1}^{n}(\cos k\phi \cos k\theta + \sin k\phi \sin k\theta)\right)\ d\phi \\
&= \frac{1}{\pi}\int_{-\pi}^{\pi} g(\phi)\left(\frac{1}{2} + \sum_{k=1}^{n}\cos k(\theta - \phi)\right)\ d\phi \\
&= \frac{1}{\pi}\int_{-\pi}^{\pi} g(\phi) u_n(\theta - \phi)\ d\phi.
\end{aligned}$$

(c) Let $\delta \in \mathbb{R}^+$. Since $[0,\pi] \subseteq [-\pi,\pi]$, it is direct that $\omega(f,[0,\pi],\delta) \leq \omega(f,[-\pi,\pi],\delta)$.
Let $\theta_1, \theta_2 \in [-\pi,\pi]$ such that $|\theta_1 - \theta_2| \leq \delta$.
By consequence of Triangle Inequality, we have $||\theta_1| - |\theta_2|| \leq |\theta_1 - \theta_2| \leq \delta$.
Thus, we have $|f(|\theta_1|) - f(|\theta_2|)| \in \{|f(x_1) - f(x_2)| \mid |x_1 - x_2| \leq \delta, x_1, x_2 \in [0,\pi]\}$.
Since f is even, we have,

$$|f(\theta_1) - f(\theta_2)| \quad = \quad |f(|\theta_1|) - f(|\theta_2|)| \leq \omega(f,[0,\pi],\delta),$$

and so $\omega(f,[-\pi,\pi],\delta) \leq \omega(f,[0,\pi],\delta)$.
Therefore $\omega(f,[-\pi,\pi],\delta) = \omega(f,[0,\pi],\delta)$ for all $\delta \in \mathbb{R}^+$.

SECTION B

Question 4

(i) Let us use the notation $\mathcal{R}[a,b]$ to be the set of Riemann integrable functions on the interval $[a,b]$, i.e. $\alpha \in \mathcal{R}[a,b]$ iff $\displaystyle\int_a^b \alpha(x)\ dx$ exists.

Since $f \in C^2_{[a,b]}$, we have $f' \in C^1_{[a,b]}$.
Let $g \in S_1(X)$ be such that $g(x_i) = f'(x_i)$ for all $i \in \{0,1,\ldots,n\}$, i.e. g is the piecewise linear interpolation to f' on $[a,b]$ with knots X.

Since g is continuous on $[a,b]$, we have $g \in \mathcal{R}[a,b]$.
Thus by Fundamental Theorem of Calculus, we have $\displaystyle\int_a^x g(t)\ dt$ to exists for all $x \in [a,b]$.
This allows us to have a well-defined function $s : [a,b] \to \mathbb{R}$ such that,

$$s(x) \quad = \quad \int_a^x g(t)\ dt + \frac{1}{n+1}\sum_{i=0}^{n}\left(f(x_i) - \int_a^{x_i} g(t)\ dt\right).$$

By substitution, we get $\displaystyle\sum_{i=0}^{n} s(x_i) = \sum_{i=0}^{n}\int_a^{x_i} g(t)\ dt + \left(\sum_{i=0}^{n}\left(f(x_i) - \int_a^{x_i} g(t)\ dt\right)\right) = \sum_{i=0}^{n} f(x_i)$.
Also, s is differentiable, with $s'(x) = g(x)$. Since $g \in S_1(X)$, g is piecewise linear except on X, and so s is piecewise quadratic except on X, i.e. we have $s \in S_2(X)$.
Lastly, $s'(x_i) = g(x_i) = f'(x_i)$ for all $i \in \{0,1,\ldots,n\}$.
Therefore s satisfies all the conditions we wanted.

Note: The main difficulty of this problem is not to verify that the expression for s satisfy the condition (even though that is all we need to present during exams), but to obtain the expression for s. From the setting of the question, we can deduce that s is likely to be obtained by performing definite integration on g, and so we proceed in that direction. Also, to keep the proof rigorous, we will need to keep employing Fundamental Theorem of Calculus, which is covered in the prerequisite MA3110, but is not in the scope of this module.

Integrating g, we get a function s to be $s(x) = \displaystyle\int_a^x g(t)\ dt + c$, for some $c \in \mathbb{R}$.
Making use of the remaining condition, we have,

$$\begin{aligned} \sum_{i=0}^{n} f(x_i) = \sum_{i=0}^{n} s(x_i) &= c(n+1) + \sum_{i=0}^{n} \int_a^{x_i} g(t)\ dt, \\ c &= \frac{1}{n+1} \sum_{i=0}^{n} \left(f(x_i) - \int_a^{x_i} g(t)\ dt \right), \end{aligned}$$

and so we obtain our expression for s.

(ii) Let $i \in \{1, 2, \ldots, n\}$.
Since s' is linear on $[x_{i-1}, x_i]$, $s'(x_{i-1}) = f'(x_{i-1})$ and $s'(x_i) = f'(x_i)$, for $x \in [x_{i-1}, x_i]$, we have,

$$s'(x) = \frac{x_i - x}{x_i - x_{i-1}} f'(x_{i-1}) + \frac{x - x_{i-1}}{x_i - x_{i-1}} f'(x_i).$$

Since $f \in C^3_{[a,b]}$, we have f'' exists. By Mean Value Theorem, there exists $\mu_1, \mu_2 \in [x_{i-1}, x_i]$ with,

$$\begin{aligned} f'(x) - f'(x_{i-1}) &= f''(\mu_1)(x - x_{i-1}); \\ f'(x_i) - f'(x) &= f''(\mu_2)(x_i - x). \end{aligned}$$

Since $|\mu_1 - \mu_2| \le |x_i - x_{i-1}| \le \Delta$, together with the AM-GM inequality, we have,

$$\begin{aligned} \left| f'(x) - s'(x) \right| &= \left| \frac{x_i - x}{x_i - x_{i-1}} (f'(x) - f'(x_{i-1})) + \frac{x - x_{i-1}}{x_i - x_{i-1}} (f'(x) - f'(x_i)) \right| \\ &= \left| \frac{x_i - x}{x_i - x_{i-1}} (f''(\mu_1)(x - x_{i-1})) + \frac{x - x_{i-1}}{x_i - x_{i-1}} (-f''(\mu_2)(x_i - x)) \right| \\ &= \frac{1}{x_i - x_{i-1}} (x_i - x)(x - x_{i-1}) |f''(\mu_1) - f''(\mu_2)| \\ &\le \frac{1}{x_i - x_{i-1}} \left(\frac{(x_i - x) + (x - x_{i-1})}{2} \right)^2 \omega(f'', [a,b], \Delta) \\ &= \frac{x_i - x_{i-1}}{4} \omega(f'', [a,b], \Delta) \\ &\le \frac{1}{4} \omega(f'', [a,b], \Delta)\Delta, \end{aligned}$$

and therefore $\|f' - s'\|_\infty \le \dfrac{1}{4}\omega(f'', [a,b], \Delta)\Delta$.

Note: Notice that we are able to obtain a better bound than the given $\dfrac{1}{2}\omega(f'', [a,b], \Delta)\Delta$. There is nothing wrong to present a better bound solution during exam, since it still satisfies what we wanted to prove.

Question 5

(i) Let P_n be the statement "$(\Delta^n(fg))(x) = \sum_{k=0}^{n} \binom{n}{k} (\Delta^k f)(x)(\Delta^{n-k} g)(x-k)$ for all $x \in \mathbb{R}$", $n \in \mathbb{Z}_{\geq 0}$.

For all $x \in \mathbb{R}$, we have $(\Delta^0(fg))(x) = (fg)(x) = f(x)g(x) = \sum_{k=0}^{0} \binom{0}{k} (\Delta^k f)(x)(\Delta^{0-k} g)(x-k)$.

Thus P_0 is true.

Assume that P_i is true, $i \in \mathbb{Z}_{\geq 0}$.
We know that for all $k \in \mathbb{Z}_{\geq 0}$, Δ^k is linear.
Thus for all $x \in \mathbb{R}$, $k \in \mathbb{Z}_{\geq 0}$, we have $(\Delta^{k+1} f)(x) = (\Delta^k(\Delta f))(x) = (\Delta^k f)(x) - (\Delta^k f)(x-1)$.
Together with the induction hypothesis, we have,

$$
\begin{aligned}
& \sum_{k=0}^{i+1} \binom{i+1}{k} (\Delta^k f)(x)(\Delta^{i+1-k} g)(x-k) \\
= \; & \sum_{k=0}^{i+1} \left(\binom{i}{k} + \binom{i}{k-1} \right) (\Delta^k f)(x)(\Delta^{i+1-k} g)(x-k) \\
= \; & \sum_{k=0}^{i} \binom{i}{k} (\Delta^k f)(x)(\Delta^{i+1-k} g)(x-k) + \sum_{k=0}^{i} \binom{i}{k} (\Delta^{k+1} f)(x)(\Delta^{i-k} g)(x-k-1) \\
= \; & \sum_{k=0}^{i} \binom{i}{k} (\Delta^k f)(x) \Big((\Delta^{i-k} g)(x-k) - (\Delta^{i-k} g)(x-k-1) \Big) \\
& + \sum_{k=0}^{i} \binom{i}{k} \Big((\Delta^k f)(x) - (\Delta^k f)(x-1) \Big) (\Delta^{i-k} g)(x-k-1) \\
= \; & \sum_{k=0}^{i} \binom{i}{k} (\Delta^k f)(x)(\Delta^{i-k} g)(x-k) + \sum_{k=0}^{i} \binom{i}{k} (\Delta^k f)(x-1)(\Delta^{i-k} g)(x-k-1) \\
= \; & (\Delta^i(fg))(x) - (\Delta^i(fg))(x-1) \\
= \; & (\Delta^{i+1}(fg))(x),
\end{aligned}
$$

and so P_{i+1} is true.

Therefore by Mathematical Induction, we have P_n to be true for all $n \in \mathbb{Z}_{\geq 0}$.

(ii) We notice that for all $m \in \mathbb{Z}_{\geq 2}$, $x \in \mathbb{R}$, we have $x_+^m = x \cdot x_+^{m-1}$.
Let f and g be such that $f(x) = x$ and $g(x) = x_+^{m-1}$ for all $x \in \mathbb{R}$. This gives us $(fg)(x) = x_+^m$.
Now, for all $x \in \mathbb{R}$, we have $(\Delta f)(x) = (x) - (x-1) = 1$, and $(\Delta^2 f)(x) = 1 - 1 = 0$.
This implies that $(\Delta^k f)(x) = 0$ for all $k \in \mathbb{Z}_{\geq 2}$, $x \in \mathbb{R}$.
Thus from the result of (5i), we have,

$$
\begin{aligned}
\Delta^{m+1} x_+^m = (\Delta^{m+1}(fg))(x) \; &= \; \binom{m+1}{0} (x)(\Delta^{m+1} g)(x) + \binom{m+1}{1} (1)(\Delta^m g)(x-1) \\
&= \; x\big((\Delta^m g)(x) - (\Delta^m g)(x-1)\big) + (m+1)(\Delta^m g)(x-1) \\
&= \; x \Delta^m x_+^{m-1} + (m+1-x)\Delta^m (x-1)_+^{m-1}.
\end{aligned}
$$

This gives us $mM_m(x) = \frac{1}{(m-1)!} \Delta^{m+1} x_+^m = xM_{m-1}(x) + (m+1-x)M_{m-1}(x-1)$ for all $x \in \mathbb{R}$.

Question 6

(a) We established in lecture that for $n \in \mathbb{Z}^+$, the $(n+1)^{\text{th}}$ Chebyshev Polynomial T_{n+1} has an alternating set of $n+2$ points. We also know that $T_{n+1}(x)$ has a leading term $\dfrac{1}{2^n}x^{n+1}$.
This gives us $p^* : [-1,1] \to \mathbb{R}$ such that $p^*(x) = x^{n+1} - 2^n T_{n+1}(x)$ to be an element of P_n.
Also $f(x) - p^*(x) = 2^n T_{n+1}(x)$, which gives us the same alternating set for T_{n+1} to be an alternating set for $f - p^*$ consisting of $n+2$ elements.
This implies that p^* is the best uniform approximation to f on the interval $[-1,1]$.

(b) Firstly, we notice that since $[a,b]$ is compact and $f - p^*$ is continuous, by the Extreme Value Theorem, we have A_0 to be non-empty.

If $f \in P_n$, then we have $p^* - f \in P_n$.
The condition on p^* tells us that $\max\limits_{x \in A_0} \big(-(f(x) - p^*(x))^2 \big) = \max\limits_{x \in A_0} (f(x) - p^*(x))(p^*(x) - f(x)) \geq 0$, which implies that $\max\limits_{x \in A_0} \big(-(f(x) - p^*(x))^2 \big) = 0$.
This gives us $f(x) - p^*(x) = 0$ for all $x \in A_0$, and so $\|f - p^*\|_\infty = 0$, which gives us p^* to be the best uniform approximation to f (in fact, $f = p^*$).

Else, we have $f \notin P_n$. This gives us $f \neq p^*$, and so $\|f - p^*\|_\infty \neq 0$. From lecture, it is sufficient to prove that A_0 has an alternating subset for $f - p^*$ consisting of $n+2$ points. Assume on the contrary that the largest alternating subset of A_0 for $f - p^*$ consists of $k \in \mathbb{Z}^+$ points, where $k \leq n+2$.
Let $A_1 = \{x_1, \ldots, x_k\}$ be such a set, with $x_i < x_j$ for all $i, j \in \{1, \ldots, k\}$ with $i < j$.

Let $A_0^+ = \{x \in A_0 \mid f(x) - p^*(x) > 0\}$ and $A_0^- = \{x \in A_0 \mid f(x) - p^*(x) < 0\}$.
Notice that A_0^+ and A_0^- are closed sets since they are the preimages of the closed sets $\{\|f - p^*\|_\infty\}$ and $\{-\|f - p^*\|_\infty\}$ respectively under the continuous function $f - p^*$. In addition, they are disjoint.
Let $i \in \{1, \ldots, k-1\}$. WLOG let $x_i \in A_0^+$ (else, switch the role of A_0^+ and A_0^- to obtain the other case). Then $x_{i+1} \in A_0^-$.
Let $s_i^+ = \sup\{x \in A_0^+ \mid x < x_{i+1}\}$ and $s_i^- = \inf\{x \in A_0^- \mid x > x_i\}$.
Since A_1 is the largest alternating set, there should be no alternating that happens in the interval $[x_i, x_{i+1}]$, and so $s_i^+ \leq s_i^-$. Also, since A_0^+ and A_0^- are closed, we have $s_i^+ \in A_0^+$ and $s_i^- \in A_0^-$. Lastly, A_0^+ and A_0^- being disjoint gives us $s_i^+ \neq s_i^-$, i.e. $s_i^+ < s_i^-$. Let $t_i = \dfrac{s_i^+ + s_i^-}{2}$.

Again, WLOG let $x_k \in A_0^-$ (else, switch the role, similarly as above).
Let $q \in P_n$ be such that $q(x) = \displaystyle\prod_{i=1}^{k-1} (x - t_i)$.
Then for all $x \in A_0^+$, we have $q(x) < 0$, and for all $x \in A_0^-$, we have $q(x) > 0$.
This gives us $(f(x) - p^*(x))q(x) < 0$ for all $x \in A_0$, a contradiction.

Therefore there is an alternating subset of A_0 for $f - p^*$ with $n+2$ elements, and so p^* is the best uniform approximation to f.