MATH 4332 SPRING 2008

The Weierstrass Approximation Theorem

In these notes we give a proof of the *Weierstrass Approximation Theorem* using Bernstein polynomials. This proof is different from the one that appears in Rudin and has the virtue that it is quite elementary, constructive and the proof generalizes well. The proof will follow what we did in lectures though it will be a little more polished. I will also add some comments and extra detail along the way.

THEOREM (The Weierstrass Approximation Theorem). Every continuous function on $I = [0, 1]$ can be uniformly approximated by polynomials. That is, if $f \in C^0(I)$ and $\varepsilon > 0$, then there exists a polynomial p such that

$$\rho(f, p) = \sup_{x \in I} |f(x) - p(x)| < \varepsilon.$$

We have stated the theorem for the closed interval $I = [0, 1]$ but the result holds for any closed and bounded interval $[a, b] \subset \mathbb{R}$. See the solutions to Homework 2.

Bernstein polynomials

Let $f \in C^0(I)$ and $n \geq 1$. The *nth Bernstein polynomial* $B_n(f)$ of f is the polynomial defined by

$$B_n(f)(x) = \sum_{p=0}^{n} \binom{n}{p} f(\frac{p}{n}) x^p (1-x)^{n-p}.$$

LEMMA For all $n \geq 1$ we have
(1) $B_n(cf) = cB_n(f)$, $f \in C^0(I)$, $c \in \mathbb{R}$.
(2) $B_n(f + g) = B_n(f) + B_n(g)$, $f, g \in C^0(I)$.
(3) $B_n(f) > 0$ on I if $f > 0$ on I.
(4) $B_n(1) = 1$.
(5) $B_n(t)(x) = x$ (here $f(t) = t$).
(6) $B_n(t^2)(x) = x^2 + \frac{x-x^2}{n}$ (here $f(t) = t^2$).

REMARKS
(a) (1,2) amount to the linearity of $B_n : C^0(I) \to C^0(I)$.
(b) (3) is a little stronger than what we gave in lectures. It implies that if $f > g$ then $B_n(f) > B_n(g)$ (always on I).
(c) When we replace f by an actual function the variable for f will always be t — as in $f(t)$. The Bernstein polynomial will always be a function of $x \in I$.

Date: January 29, 2008.

Proof. (1,2,3) are obvious (note for (3) that $x^p(1-x)^{n-p} > 0$ on $(0,1)$).
(4) $B_n(1)(x) = \sum_{p=0}^{n} \binom{n}{p} 1 x^p(1-x)^{n-p} = (x+(1-x))^n = 1$.
(5) We assume $n \geq 2$ — the result is easy if $n = 1$. We have

$$\begin{aligned}
B_n(t)(x) &= \sum_{p=0}^{n} \binom{n}{p} \frac{p}{n} x^p (1-x)^{n-p}, \\
&= \sum_{p=1}^{n} \frac{n!}{p!(n-p)!} \frac{p}{n} x^p (1-x)^{n-p}, \\
&= \sum_{p=1}^{n} \frac{(n-1)!}{(p-1)!(n-p)!} x^p (1-x)^{n-p}, \\
&= \sum_{p=1}^{n} \frac{(n-1)!}{(p-1)!((n-1)-(p-1))!} x x^{p-1} (1-x)^{(n-1)-(p-1)}, \\
&= x \sum_{q=0}^{n-1} \binom{n-1}{q} x^q (1-x)^{(n-1)-q}, \ (q = p-1) \\
&= x B_{n-1}(1)(x) = x.
\end{aligned}$$

(6) Again we assume $n \geq 2$, the result is easy if $n = 1$. We have

$$\begin{aligned}
B_n(t^2)(x) &= \sum_{p=0}^{n} \binom{n}{p} (\frac{p}{n})^2 x^p (1-x)^{n-p}, \\
&= \sum_{p=1}^{n} \frac{(n-1)!}{(p-1)!(n-p)!} \frac{p}{n} x^p (1-x)^{n-p}, \text{ as in (5)}, \\
&= \sum_{p=1}^{n} \frac{(n-1)!}{(p-1)!(n-p)!} (\frac{p-1}{n} + \frac{1}{n}) x^p (1-x)^{n-p}, \\
&= A + B, \text{ where} \\
A &= \sum_{p=1}^{n} \frac{(n-1)!}{(p-1)!(n-p)!} \frac{p-1}{n} x^p (1-x)^{n-p}, \\
B &= \sum_{p=1}^{n} \frac{(n-1)!}{(p-1)!(n-p)!} \frac{1}{n} x^p (1-x)^{n-p}.
\end{aligned}$$

Checking the proof of (5), we see that $B = \frac{1}{n} B_{n-1}(t)(x) = \frac{x}{n}$. It remains to evaluate A. Cancelling the factor $(p-2)$ and taking out factors $(n-2)/n$ and x^2 we have (just as in the proof of (5))

$$\begin{aligned} A &= x^2\frac{(n-1)}{n}\sum_{p=2}^{n}\frac{(n-2)!}{(p-2)!((n-2)-(p-2))!}x^{p-2}(1-x)^{(n-2)-(p-2)}, \\ &= x^2\frac{(n-1)}{n}B_{n-2}(1)(x) \\ &= x^2\frac{(n-1)}{n}. \end{aligned}$$

(Note that if $n=2$, we use $\sum_{p=2}^{2}\frac{(2-2)!}{(p-2)!((2-2)-(p-2))!}x^{p-2}(1-x)^{(2-2)-(p-2)}=1$.) Finally,

$$A+B=x^2\frac{(n-1)}{n}+\frac{x}{n}=x^2+\frac{x-x^2}{n}.$$

and so $B_n(t^2)(x)=x^2+\frac{x-x^2}{n}$. □

Proof of Weierstrass Approximation Theorem. Let $f\in C^0(I)$ and $\varepsilon>0$. Since I is compact, $f:I\to\mathbb{R}$ is uniformly continuous and so $\exists\delta>0$ such that for all $t,x\in I$ satisfying $|x-t|<\delta$ we have

$$-\varepsilon/2<f(t)-f(x)<\varepsilon/2 \text{ (ie } |f(t)-f(x)|<\varepsilon/2). \tag{1}$$

Since f is continuous on a compact set, $M=\sup_{s\in I}|f(s)|<\infty$. The next inequality follows from the triangle inequality.

$$-2M<f(t)-f(x)<2M, \quad \text{for all } t,x\in I. \tag{2}$$

Observe that the function $\frac{2M}{\delta^2}(t-x)^2$ is greater than or equal to one provided that $|t-x|\geq\delta$. It follows from (1,2) that for all $t,x\in I$ we have

$$-\varepsilon/2-\frac{2M}{\delta^2}(t-x)^2<f(t)-f(x)<\frac{2M}{\delta^2}(t-x)^2+\varepsilon/2. \tag{3}$$

Regard each term in this inequality as a function of t (so x is fixed). Noting property (3) of Bernstein polynomials we have for all $n\geq1$ the following inequality between *functions* (of x)

$$B_n(-\varepsilon/2-\frac{2M}{\delta^2}(t-x)^2)<B_n(f)-B_n(f(x))<B_n(\frac{2M}{\delta^2}(t-x)^2+\varepsilon/2).$$

(What are we doing? We fix x, set $t=\frac{p}{n}$ in (3), multiply by $\binom{n}{p}x^p(1-x)^{n-p}$ and sum from $p=0$ to $p=n$. In particular, $B_n(f(x))=f(x)B_n(1)=f(x)$, using property (1)).

Using properties (1,2,4), we have

$$\begin{aligned} B_n(\frac{2M}{\delta^2}(t-x)^2+\varepsilon/2) &= \frac{2M}{\delta^2}B_n((t-x)^2)+\varepsilon/2, \\ B_n(-\frac{2M}{\delta^2}(t-x)^2-\varepsilon/2) &= -\frac{2M}{\delta^2}B_n((t-x)^2)-\varepsilon/2. \end{aligned}$$

Hence for all $x \in I$ we have

$$-\varepsilon/2 - \frac{2M}{\delta^2} B_n((t-x)^2)(x) < B_n(f)(x) - f(x) < \frac{2M}{\delta^2} B_n((t-x)^2)(x) + \varepsilon/2. \tag{4}$$

We claim that $\exists N$ such that for $n \geq N$, $|\frac{2M}{\delta^2} B_n((t-x)^2)(x)| < \varepsilon/2$ for all $x \in I$. It follows from (4) that for $n \geq N$, $x \in I$, $|B_n(f)(x) - f(x)| < \varepsilon/2 + \varepsilon/2 = \varepsilon$ and we are done.

In order to prove the claim we evaluate $B_n((t-x)^2)$. Since $(t-x)^2 = t^2 - 2tx + x^2$, we have

$$B_n((t-x)^2) = B_n(t^2) - 2xB_n(t) + x^2 B_n(1).$$

Evaluating at x, this gives us (using (4,5,6))

$$\begin{aligned} B_n((t-x)^2)(x) &= B_n(t^2)(x) - 2xB_n(t)(x) + x^2 B_n(1)(x), \\ &= (x^2 + \frac{x - x^2}{n}) - 2xx + x^2 1, \\ &= \frac{x - x^2}{n}. \end{aligned}$$

Since $x \in [0,1]$, $0 \leq B_n((t-x)^2)(x) < 1/n$. Hence for $x \in I$,

$$0 < \frac{2M}{\delta^2} B_n((t-x)^2) + \varepsilon/2 < \frac{2M}{n\delta^2} + \varepsilon/2.$$

Now choose N so that $\frac{2M}{n\delta^2} < \varepsilon/2$, $n \geq N$. $\square$