10 November 2013

Chebyshev approximation and Helly's Theorem

Helly's Theorem states that $m \geq n+2$ convex bodies in $\mathbf{R}^n$ have non-empty intersection if any $n+1$ of them have non-empty intersection. This Theorem stated by German mathematician Helly in 1913 has many different proofs[1)] *It can be proved using just elementary mathematics (excellent topic for pupils in the school). On the other hand one of its proofs uses notion of Chech cohomology.*

In this etude I try to show application of Helly's Theorem to theory of approximation of functions. I am writing this etude inspired and based on the wonderful article of V.G.Boltiansky and N.M.Yaglom "Convex bodies" (Encyclopaedia of Elementary Mathematics. Volume 5. Geometry. Moscow 1966 (in Russian))

Helly's theorem on convex bodies have the following very interesting application to theory of approximation of continuous functions by polynomials of order n. Here we consider in detail the case when we approximate a function by lines (polynomials of order $n=1$) and briefly formulate the general case. (The idea of the proof is not very different for general case).

We consider continuous functions on the interval $[a,b]$. We define the distance d_∞ between continuous functions as

$$d(f,g) = ||f-g||_\infty = \max_{x\in[a,b]} |f(x)-g(x)| \,.$$

We say that the line $L_f = kx+b$ is a line which is the closest to the function f if for an arbitrary line l, $d(f,l) \geq d(f.L_f) = \varepsilon$:

$$\max_{x\in[a,b]} |f(x)-kx-b| = \varepsilon \text{ and for arbitrary line } y = k'x+b' \ \max_{x\in[a,b]} |f(x)-k'x-b'| \geq \varepsilon \,.$$

The following Theorem is obeyed:

Theorem 1 Let the line L_f be a closest line to the continuous function $f = f(x)$, $x \in [a,b]$. If ε is the distance between this line and the function f then there exist three points x_1, x_2, x_3, $a \leq x_1 < x_2 < x_3 \leq b$ such that the distance between function f and the line L_f at these points are $\pm\varepsilon$, and signs are alternating:

$$\begin{cases} f(x_1) - L_f(x_1) = \varepsilon \\ f(x_2) - L_f(x_2) = -\varepsilon \\ f(x_3) - L_f(x_3) = \varepsilon \end{cases} , \quad \text{or} \quad \begin{cases} f(x_1) - L_f(x_1) = -\varepsilon \\ f(x_2) - L_f(x_2) = \varepsilon \\ f(x_3) - L_f(x_3) = -\varepsilon \end{cases} , \tag{1}$$

Respectively for approximation by polynomials of order n we have:

Theorem 1* Let the $n-th$ order polynomial $P_f^{(n)}$ is the closest n-th order polynomial to the continuous function $f = f(x)$, ($x \in [a,b]$ in the $n+1$-dimensional linear space of all n-th order polynomials $P^{(n)}(x) = a_0x^n + a_1x^{n-1} + a_1x + a_0$, ($a_0, \ldots . a_n$ are arbitrary real numbers.). If ε is the distance between the parabola $P_f^{(n)}$ and the function f then in

---

1) .

the interval $[a,b]$ there exist $n+1$ points $x_1,\dots x_{n+1}$, $a\le x_1<\dots<x_{n+1}\le b$ such that the distance between function f and the parabola P_f at these points is $\pm\varepsilon$ and signs are alternating, i.e.

$$\begin{cases} f(x_1)-L_f(x_1)=\varepsilon \\ f(x_2)-L_f(x_2)=-\varepsilon \\ f(x_3)-L_f(x_3)=\varepsilon \\ \dots \\ f(x_{n+1})-L_f(x_{n+1})=(-1)^n\varepsilon \end{cases}, \quad \text{or} \quad \begin{cases} f(x_1)-L_f(x_1)=-\varepsilon \\ f(x_2)-L_f(x_2)=\varepsilon \\ f(x_3)-L_f(x_3)=-\varepsilon \\ \dots \\ f(x_{n+1})-L_f(x_{n+1})=(-1)^{n+1}\varepsilon \end{cases}.$$

Example *Chebyshev approximation and Chebyshev polynomials.*
Consider Chebyshev polynomails $\{T_k\}$,

$$T_k(x)=\frac{1}{2^k}\cos k\arccos x\,,\quad -1\le x\le 1\,,$$

$$T_1(x)=x, T_2(x)=\frac{2x^2-1}{2}\,, T_3(x)=\frac{4x^3-3x}{4}\; T_4(x)=\frac{8x^4-8x^2+1}{8}\dots,$$

$(\frac{1}{2}T_{k-1}(x)+2T_{k+1}(x)=xT_k(x).)$.

The basic property of Chebyshev polynomials is that for every natural n, the polynomial $T_n(x)$ is the polynomial which is closest to zero in the n-dimensional affine space of all polynomials of order n with leading term x^n:

$$d(T_n)=\max_{x\in[-1,1]}|T_n(x)|=\frac{1}{2^{n-1}}\le\min_{a_1,a_2,\dots,a_n}\max_{x\in[-1,1]}|x^n+a_1x^{n-1}+\dots+a_{n-1}x+a_n|\,.$$

It implies that the polynomial $P^{(n)}=x^{n+1}-T_{n+1}$ is the closest to the function $f=x^{n+1}$ in the linear space of n-th order polynomials[1)] The distance between this parabola and function $y=x^{n+1}$ is equal to $\varepsilon=\frac{1}{2^n}$. At the points $x_i=\arccos\frac{2\pi i}{n}$ the distance is $\pm\varepsilon$ and signes are alternating:

$$x_i^{n+1}-P^{(n)}(x)=-T_{n+1}(x)=-\frac{(-1)^i}{2^n}\,.$$

In this special case the nodes of Chebyshev polynomials are equidistant. The Theorem tells that the property of changing the signs is kept in the general case. This statement is very important for practical calculations of Chebyshev polynomials.

We sketch the proof of the Theorem for approximation by lines.

First of all formulate the following corollary form Helly's Theorem:

---

1) In formulation of Theorem 1^* we deal with linear space of n-th order polynomials, and regarding the basic property of Chebyshev polynomials we deal with *linear* space of n-th order polynomials with leading term x^n. E.g. polynomial $T_4(x)-x^4$ belongs to the linear space of 3-rd order polynomials, in spite of the fact that its leading term is proportional to x^2.

Corollary Let $\mathcal{M}$ be the set of parallel segments such that this set belongs to bounded domain in $\mathbf{R}^2$. Suppose that for an arbitrary three segments there exists a line which intersects these segments. Then there exists a line which intersects all the segments.

Respectively if for arbitrary $k+2$ segments there exists k-th order polynomial which intersects these segments. Then there exists k-th order polynomial which intersects all the segments.

We first sketch the proof of Theorem based on this corollary then prove the Corollary.

We will prove the Theorem for lines, i.e. for approximation by polynomials of the order $n=1$. The idea of proof is the same for an arbitrary n.

Proof

Let $L_f \colon y = kx + b$ be a closest line to the function f. Let a distance be equal to ε:

$$\max_{x\in[a,b]} |f(x) - kx - b| = \varepsilon, \quad \forall k', b', \quad \max_{x\in[a,b]} |f(x) - k'x - b'| \geq \varepsilon.$$

Pick arbitrary $\varepsilon' < \varepsilon$. Consider the set $\{d_x\}$ ($x \in [a,b]$) of segments $d_x = [a_x, b_x]$ such that points a_x, b_x have coordinates

$$a_x = (x, f(x) - \varepsilon'), b_x = (x, f(x) + \varepsilon').$$

The set of these segments is the set of vertical segments centered at the points of graph of the function f with length $2\varepsilon'$.

It follows from corollary 1 that for every $\varepsilon' < \varepsilon$ there exist three points x_1, x_2, x_3 such that *there is no a line* which intersects corresponding segments $d_{x_1}, d_{x_2}, d_{x_3}$. Indeed if for arbitrary three points x_1, x_2, x_3 there exists a line which intersects corresponding segments $d_{x_1}, d_{x_2}, d_{x_3}$ then due the corollary there exists a line L' which intersects all the segments $\{d_x\}$, i.e. the distance between line L' and a function f is less or equal to ε'. This contradicts to the fact the line L_f is the closest line.

We come to the following observation:

Observation 1 For every ε': $0 < \varepsilon' < \varepsilon$ there exist three points $x_1, x_2, x_3 \in [a,b]$ such that the distance between arbitrary line and the function f at one of these points is greater than ε'.

This observation plus continuity arguments implies the following observation:

Observation 2 There exist three points $a \leq x_1 \leq x_2 \leq x_3 \leq b$ such that the distance between arbitrary line and the function f at one of these points is greater or equal than ε.

Indeed consider the sequence $\{\varepsilon_n\}$ such that $0 < \varepsilon_n < \varepsilon$ and $\varepsilon_n \to vare$, e.g. $\varepsilon_n = vare - \frac{1}{n+M}$ (for enough big M). Choose for every ε_n points $\{x_1^{(n)}, x_2^{(n)}, x_3^{(n)}\}$ such that at these points the distance between every line including the line L_f is greater than ε_n. Due to compactness of the segment $[a,b]$ we can pick from this sequence the subsequence $\{x_1^{(n_k)}, x_2^{(n_k)}, x_3^{(n_k)}\}$ such that $\lim_{k\to\infty} x_1^{(n_k)} = x_1$, $\lim_{k\to\infty} x_2^{(n_k)} = x_2$, $\lim_{k\to\infty} x_3^{(n_k)} = x_3$. One can see that points $\{x_1, x_2, x_3\}$ are the points where for an arbitrary line L, the distance between function f and the line L is bigger that every ε_n. This we come to the the statement of Observation 2.

Now prove the Theorem using the Observation 2.

Using Observation 2 choose the points $\{x_1, x_2, x_3\}$ and show that the relations (1) are obeyed.

Consider $\Delta_i = f(x_i) - L_f(x_i)$, $i = 1, 2, 3$. We have to show that all Δ_i have the modulus ε and signes are alternating:

$$|\Delta_1| = |\Delta_2| = |\Delta_3| = \varepsilon, \quad \Delta_1\Delta_3 > 0, \quad \Delta_1\Delta 2 < 0\,, \tag{1a}$$

i.e. conditions (1) are obeyed. If these conditions are not obeyed then it is easy to show that one can always find a line L such that its distance to the function f at all points x_1, x_2 and x_3 is less than $vare$. This contradicts to Observation 2.

Suppose for example that $\Delta_1 > 0$ and $\Delta_2 > 0$. If $\Delta_3 > 0$ then one can choose $\delta > 0$ such that all the distances between function f and the line $L = L_f + \varepsilon$ at points x_1, x_2, x_3 are less than ε. If $\Delta_3 < 0$ then rotating the line L_f around the point $(x_2, L_f(x_2))$ on a smal angle we again come to the line L' such that all the distances between function f and the line L' at points x_1, x_2, x_3 are less than ε. Hence if $\Delta_1 > 0$ then $\Delta_2 < 0$.

By analogous considerations one can easy show that in all the cases when conditions (1),(1a) are not obeyed then one can choose another line L' such that the distance between the line L' and a function f at all points x_1, x_2, x_3 is less than ε. This implies the statement of Theorem.

Finally we prove the Corollary1.

Let $\mathcal{M}$ be the set of parallel segments. WLOG we may suppose that all the segments are vertical. Consider an arbitrary vertical segment d, Denote by Π_d the set of lines which intersects with the segment d. Every line which intersect this segment is non-vertical line, $y = kx + b$. We parameterise all not-vertical lines by pairs (k, b). One can see that the set of the pairs which correspond to the set of line Π_d is the convex set: If segment connects the points $(x_0, y_0), (x_0, y_0 + d)$ $(d > 0)$ then the condition that the line $kx + b$ intersect this segment is:

$$y_0 \leq kx_0 + b \leq y_0 + d$$

These conditions define the strip, the convex set in the plane (k, b). Now the Corollary follows from Helly's Theorem. ■

Example Consider approximation of the function $f = \sin x$ on the interval $[0, \pi/2]$. One can see that in this case the closest line is uniquely defined by the condition (1) of the Theorem. If L_f: $y = kx + b$ is the closest line, and the distance is equal to ε then

$$\varepsilon = (kx + b - \sin x)\big|_{x=0} = -\left(\sin x - (kx + b)\right)\big|_{x=x_0} = (kx + b - \sin x)\big|_{x=1}\,.$$

The points x_1, x_2, x_3 are: $x_1 = 0$, $x_3 = 1$ and the middle point x_2 is defined by the stationary point condition. We come to simultaneous equations

$$\begin{cases} b = \varepsilon \\ \sin x_0 - kx_0 - b = \varepsilon \\ \cos x_0 = k \\ k\frac{\pi}{2} + b - 1 = \varepsilon \end{cases}$$

Solving this system we come to $k = \frac{2}{\pi}$, $x_0 = \arccos\frac{2}{\pi}$ and

$$\varepsilon = b = \frac{1}{2}\left(\sin\arccos\frac{2}{\pi} - \frac{2}{\pi}\arccos\frac{2}{\pi}\right) = \frac{\sqrt{\pi^2 - 4} - 2\arccos\frac{2}{\pi}}{2\pi} \approx 0.10526$$

The line $y = \frac{2}{\pi}x + b$ with $b \approx 0.10526$ is the closest line to the function $f = \sin x$ on the interval $[0, \pi/2]$. The distance is equal to $\varepsilon \approx 0.10526$.