Disclaimer

These notes have been written with the intention of being used by people who have already taken previous analysis courses in the Tripos (beyond IA Analysis). For this reason, some of the results which have been seen already in Tripos are omitted.

1   Basic Analysis

Definition. For $A \subseteq \mathbb{R}^n$ and $x \in \mathbb{R}^n$, define

$$d(x, A) = \inf\{d(x, a) = a \in A\}.$$

Remark. If A is closed, $d(x, A) = 0$ if and only if $x \in A$. Indeed if $d(x, A) = 0$, there exists a sequence $(a_n)_{n \geq 1}$ such that $d(x, a_n) \to 0$ by the definition of infima. This is the same as $a_n \to x$, and since A is closed $x \in A$. The converse is obvious.

Proposition. The map $x \mapsto d(x, A)$ is continuous.

Proof. It is easy to show $|d(x, A) - d(y, A)| \leq d(x, y)$ and then the claim follows. $\square$

Theorem. *If $E \subseteq \mathbb{R}^n$ is compact and $f : E \to f(E)$ is continuous, then E is compact.*

Proof. Take a sequence $(f(x_n))_{n \geq 1}$ in $f(E)$. Then $(x_n)_{n \geq 1}$ has a convergent subsequence $(x_{n_j})_{j \geq 1}$, and $f(x_{n_j}) \to f(x) \in f(E)$. $\square$

Theorem (Fundamental Theorem of Algebra). *If $P(z)$ is a polynomial with coefficients in $\mathbb{C}$, then P has a root.*

Proof. Write $P(z) = \sum_{j=0}^{n} a_j z^j$ (wlog $a_n = 1$). Note

$$|P(z)| \geq |z|^n (|a_n| - |a_{n-1}||z|^{-1} - \ldots - |a_0||z|^{-n}).$$

Hence if $R = 2\left(2 + \sum_{j=0}^{n-1} |a_j|\right)$, whenever $|z| \geq R$ we have $|P(z)| > |z|^n/2$. Then $\overline{B}_R(0)$ is compact so $|P(z)|$ achieves a minimum on $B_R(0)$ at some $p_0 \in B_R(0)$. Since $|P(z)| \geq |z|^n/2 \geq R^n/2 > |a_0|$ for $|z| \geq R$, p_0 is a global minimum.

Now by considering $P(z + p_0)$, we can assume $p_0 = 0$. By considering $e^{i\theta} P(z)$ we can assume $a_0 \in \mathbb{R}$ and $a_0 > 0$. By considering $P(e^{i\theta} z)$ we can assume $P(z) = a_0 + a_m z^m + \sum_{m+1}^{n} a_j z^j$ for $a_0, a_m \in \mathbb{R}$, $a_0 > 0$ and $a_m < 0$. Then $\frac{P(\delta) - P(0)}{\delta^m} \to a_m$ as $\delta \to 0$.

So the real part of $\frac{P(\delta) - P(0)}{\delta^m}$ tends to a_m and the imaginary part tends to 0. Hence if δ is sufficiently small and positive $|P(\delta)| < |P(0)|$. $\square$

Remark. Clearly this implies every complex polynomial of degree n has exactly n roots with multiplicity (division algorithm).

2   Laplace's Equation

Definition. *Laplace's equation* is $\nabla^2 \phi = 0$.

Definition (Boundary)**.** Let (X, d) be a metric space, $E \subseteq X$.

1. The *closure* $\overline{E} = \text{Cl}(E)$ is the collection of points $x \in E$ such that there exists a sequence $(x_n)_{n \geq 1}$ in E with $x_n \to x$.

2. The *interior* $\text{int}(E)$ is the set of points $x \in E$ such that there exists $\delta > 0$ with $B(x, \delta) \subseteq E$.

3. The *boundary* ∂E is defined by $\partial E = \text{Cl}(E) \setminus \text{int}(E)$.

For the remainder of this section:

- We work in $\mathbb{R}^m$;

- Let $\Omega \neq \emptyset$ be a bounded open set;

- Let $\phi : \text{Cl}(\Omega) \to \mathbb{R}$ be a continuous function which is twice differentiable on Ω;

- Let $f : \partial\Omega \to \mathbb{R}$ be a continuous function.

We want to show that the problem $\nabla^2 \phi = 0$ on Ω, $\phi = f$ on $\partial\Omega$ has at most one solution.

Lemma. *If $\nabla^2 \phi > 0$ on Ω, then ϕ achieves its maximum on $\partial\Omega$.*

Remark. The maximum must exist since $\text{Cl}(\Omega)$ is closed and bounded.

Proof. Suppose ϕ attains its maximum at $x^* \in \Omega$. Then there exists $\delta > 0$ such that
$$(x_1^* - \delta, x^* + \delta) \times \ldots \times (x_n^* - \delta, x^* + \delta) \subseteq \Omega.$$
Define $f_j(t) = \phi(x_1^*, \ldots, x_{j-1}^*, x_j^* + t, x_{j+1}^*, \ldots, x_m^*)$. Then f_j has a maximum at 0 and f_j is twice differentiable with $f_j''(0) \leq 0$. Hence $\frac{\partial \phi(x^*)}{\partial x_j^2} \leq 0$ so $\nabla^2 \phi(x^*) \leq 0$, a contradiction. $\square$

Theorem. *If $\nabla^2 \phi = 0$ on Ω, then the maximum is achieved at $\partial\Omega$.*

Proof. Set $\phi_n(x) = \phi(x) + \frac{1}{n} \sum_{j=1}^m x_j^2$. Then $\nabla^2 \phi_n = \nabla^2 \phi + \frac{2m}{n} = \frac{2m}{n}$. Then by the previous lemma, ϕ_n has a maximum at $x_n^* \in \partial\Omega$. The boundary is closed and bounded so compact, hence we can find x^* and a subsequence $(x_{n(j)}^*)_{j \geq 1}$ such that $x_{n(j)}^* \to x^*$. Then

$$\phi(x_{n(j)}^*) + \frac{1}{n(j)} \sum_{i=1}^m (x_k^*)^2 \geq \phi(x) + \frac{1}{n(j)} \sum_{i=1}^m x_k^2 \geq \phi(x) \; \forall x \in \text{Cl}(\Omega).$$

So let $n(j) \to \infty$ and use continuity of ϕ to obtain $\phi(x^*) \geq \phi(x)$ for all $x \in \text{Cl}(\Omega)$. $\qquad\square$

Remark. Now if $\nabla^2 \phi_1 = \nabla^2 \phi_2 = 0$ on Ω and $\phi_1 = \phi_2 = f$ on $\partial\Omega$, then $\phi_1 - \phi_2 = 0$ on $\partial\Omega$ so $\phi_1 - \phi_2 \leq 0$ on $\text{Cl}(\Omega)$ by the previous theorem. Similarly $\phi_2 - \phi_1 \leq 0$ on $\text{Cl}(\omega)$ so $\phi_1 = \phi_2$.

What about existence?

Zaremba counterexample

Consider $\Omega = D \setminus \{0\} = \{x \in \mathbb{R}^2 : 0 < ||x|| < 1\}$. We'll show $\nabla^2 \phi = 0$,

$$\phi(x) = \begin{cases} 0 & ||x|| = 1 \\ 1 & x = 0 \end{cases}$$

on $\partial\Omega$ has no solution.

Indeed, let R_θ be a rotation through θ about 0. If θ is a solution, then $x \mapsto \phi(R_\theta x)$ is a also a solution. So by uniqueness $\phi = \phi \circ R_\theta$, and we can write $\phi(x) = f(||x||)$ for some f. Now either $\nabla^2 \phi$ is aready known in radial terms, or

$$\frac{\partial \phi}{\partial x} = \frac{\partial}{\partial x} f(\sqrt{x^2 + y^2}) = \frac{x}{\sqrt{x^2 + y^2}} f'(\sqrt{x^2 + y^2})$$

$$\frac{\partial^2 \phi}{\partial x^2} = \left(\frac{1}{\sqrt{x^2 + y^2}} - \frac{x^2}{(x^2 + y^2)^{3/2}} \right) f'(\sqrt{x^2 + y^2}) + \frac{x^2}{x^2 + y^2} f''(\sqrt{x^2 + y^2})$$

$$\frac{\partial^2 \phi}{\partial y^2} = \left(\frac{1}{r} - \frac{y^2}{r^3} \right) f'(r) + \frac{y^2}{r^2} f''(r)$$

$$= \nabla^2 \phi = \frac{1}{r} f'(r) + f''(r) = \frac{1}{r} \frac{\mathrm{d}}{\mathrm{d}r} (r f'(r))$$

Now we solve $0 = \frac{1}{r} \frac{\mathrm{d}}{\mathrm{d}r} (r f'(r))$. This gives $A \log r + B = f(r)$. Since ϕ is continuous on $\mathrm{Cl}(E)$, $A = 0$ (otherwise singularity at 0). So ϕ is constant, a contradiction.

3 Brouwer's Fixed Point Theorem

In one dimension: if $f : [0,1] \to [0,1]$ is continuous, there exists $x \in [0,1]$ such that $f(x) = x$ (a fixed point). Indeed consider $g(x) = f(x) - x$ so $g(0) \geq 0$, $g(1) \leq 0$, and use the intermediate value theorem.

Theorem (Brouwer's Fixed Point Theorem)**.** *If* $f : \overline{D} \to \overline{D}$ *is continuous* $(\overline{D} = \overline{D}(0,1) \subseteq \mathbb{R}^2)$ *then there exists* x^* *such that* $f(x^*) = x^*$.

Remark. This generalises to $\mathbb{R}^n$, with the proof essentially unchanged apart from one point ("Sperner's Lemma') where we have done a little simplification from the general case.

Remark. If E is homeomorphic to $\overline{D}$ then the fixed point result also holds for E. Indeed if $T : E \to \overline{D}$ is a homeomorphism, $f : E \to E$ is continuous, then $T \circ f \circ T^{-1} : \overline{D} \to \overline{D}$ is continuous. Hence there exists $x^* \in \overline{D}$ such that $T(f(T^{-1}(x^*))) = x^*$, i.e $T^{-1}(x^*)$ is a fixed point of f.

The proof of the fixed point theorem is a long line of lemmas, most of which assert that two statements are equivalent.

The first thing we'll do is show that the Fixed Point Theorem is equivalent to the "No Retraction Theorem".

Theorem (No Retraction Theorem). *There does not exist a continuous map* $g : \overline{D} \to \partial D$ *such that* $g(x) = x$ *for* $x \in \partial D$ *(called a retraction).*

Proof of equivalence. Suppose the Fixed Point Theorem is true, but there exists a retraction $g : \overline{D} \to \partial D$. Let R be a rotation through π about 0. Then $R \circ g : \overline{D} \to \overline{D}$ has no fixed point, which is a contradiction.

Now suppose the No Retraction Theorem holds, but $f : \overline{D} \to \overline{D}$ is continuous with no fixed points. Let $F : \overline{D}^2 \setminus \{(x,x) : x \in \overline{D}\} \to \overline{D} \to \partial D$ be defined so that $F(x,y)$ is the point of intersection between the (directed) line from y to x and int∂D, we see F is continuous. Now consider $g(x) = F(x, f(x))$. Since f is continuous, g is continuous, but $g(\overline{D}) \subseteq \partial D$ and $g(x) = x$ for $x \in \partial D$, a contradiction. $\qquad\square$

Now we proceed through minor lemmas with unofficial titles.

Lemma (Three Arcs Theorem). *If we consider* 3 *equal arcs of* ∂D *in polar coordinates, say* $A = \{(1,\theta) : 0 \leq \theta \leq 2\pi/3\}$, $B = \{(1,\theta) : 2\pi/3 \leq \theta \leq 4\pi/3\}$, $C = \{(1,\theta) : 4\pi/3 \leq \theta \leq 2\pi\}$, *then there does not exist* $f : \overline{D} \to \partial D$ *such that* $f(A) \subseteq A$, $f(B) \subseteq B$, $f(C) \subseteq C$.

We'll now show this is equivalent to the no retraction theorem

Proof of equivalence. Three arcs certainly proves no retraction as a special case. Conversely, if we did have a 3 arcs function f then letting R be a rotation of π, the map Rf would have no fixed points. $\qquad\square$

We now go from "three arcs" to "three sides":

Theorem. *If $\overline{\Delta}$ is an equilateral triangle with closed sides I, J, K, then there does not exist $f : \overline{\Delta} \to \partial\Delta$ continuous such that $f(I) \subseteq I$, $f(J) \subseteq J$, $f(K) \subseteq K$.*

Now we show this is equivalent to the three arcs theorem:

Proof of equivalence. Consider the homeomorphism $Tx = r(x)x$ where $r(x)$ is the distance from 0 to the point of intersection of the radius vector with $\partial\Delta$. This maps $\overline{D}$ to $\overline{\Delta}$. $\qquad\square$

Lemma (Three colours)**.** *If A, B, C closed, $A \cup B \cup C = \overline{\Delta}$, $A \supseteq I$, $B \supseteq J$, $C \supseteq K$ (I, J, K closed sides of $\overline{\Delta}$) then $A \cap B \cap C \neq \emptyset$.*

We show this is equivalent to the three sides theorem:

Proof of equivalence. If there exists $f : \overline{\Delta} \to \partial\Delta$ continuous with $f(I) \subseteq I$, $f(J) \subseteq J$, $f(K) \subseteq K$, then take $A = f^{-1}(I)$, $B = f^{-1}(J)$ and $C = f^{-1}(K)$. Then A, B, C are closed with $A \cup B \cup C = \overline{\Delta}$ and $A \cap B \cap C = f^{-1}(I \cap J \cap K) = \emptyset$, contradicting the three colour theorem.

Now suppose we could three colour with $A \cap B \cap C = \emptyset$. We embed $\overline{\Delta}$ in $\mathbb{R}^3$ as follows:
$$\overline{\Delta} = \{(x, y, z) \in \mathbb{R}^3 : x, y, z \geq 0,\ x + y + z = 1\}.$$
[Barycentric coordinates, invented by Möbius.] Consider the map
$$Tx = \frac{1}{d(x, A) + d(x, B) + d(x, C)} \left(d(x, A), d(x, B), d(x, C)\right).$$

This is well defined since $A \cap B \cap C$ and A, B, C are closed. This maps $\overline{\Delta} \to \overline{\Delta}$ and is continuous since $x \mapsto d(x, A)$ is continuous. If $x \in I$ then $x \in A$ so $d(x, A) = 0$ and $Tx = (0, y, z) \in I$, so $f(I) \subseteq I$ (similarly for J, K), contradicting three sides theorem. $\qquad\square$

Overall we have

$$\text{FPT} \iff \text{RMT} \iff 3 \text{ colouring theorem.}$$

Now we show the 3 colouring theorem via Sperner's lemma:

Theorem. *Consider an equilateral triangle Δ cut into smaller triangles by n equally spaced lines parallel to each of the sides. If we colour each vertex one of three colours R,B,G with one big side*

$$T : RRR\ldots RB$$
$$J : BBB\ldots BG$$
$$K : GGG\ldots GR.$$

Then one of the small triangles must have all its vertices differet colours.

Proof. If we look at one small triangle clockwise, we assign value θ to a side as follows

$$\theta(RB) = 1, \ \theta(BR) = -1$$
$$\theta(BG) = 1, \ \theta(GB) = -1$$
$$\theta(GR) = 1, \ \theta(RG) = -1$$
$$\theta(RR) = 0, \ \theta(GG) = 0$$
$$\theta(BB) = 0$$

and each small triangle has value ψ given by the sum of the values of its sides. Then $\psi = 0$ unless the triangle is 3 coloured. Now look at the sum of all ψ of the small triangles. The interior sides cancel (counted once in each direction), so the total sum is the sum along the sides of the large triangle, which is an odd number (changes colour 3 times) so there must be a 3 coloured small triangle. □

Finally we prove the three colour theorem:

Proof of three colour theorem. Suppose $\overline{\Delta}$ is coloured A, B, C as before. Cutting Δ into smaller triangles as above, colouring external vertices as for Sperners lemma (note $I \subseteq A$ etc and make a choice of colour when its in multiple of A, B, C). Then Sperner's lemma says there's a small triangle with 3 vertices of different colours, i,e there exist $a_n \in A$, $b_n \in B$, $c_n \in C$ such that $d(a_n, b_n), d(b_n, c_n), d(c_n, a_n) < 1/n$ (making triangles small enough). Then by compactness there exists a subsequence of (a_n) converging to some $x \in \overline{\Delta}$. This automatically implies the corresponding subsequences for (b_n) and (c_n) converge to x, so $x \in A \cap B \cap C$ since A, B, C are closed. □

Example. Suppose $\sum_{j=1}^{3} a_{ij} = 1$, $a_{ij} \geq 0$. Then the map $x \mapsto y$ with $y_i = \sum_{j=1}^{3} a_{ij}x_j$ is continuous from $\overline{\Delta} = \{(x_1, x_2, x_3) \geq 0 : x_1 + x_2 + x_3 = 1\} \to \overline{\Delta}$. Indeed $y \geq 0$ and

$$\sum_{i=1}^{3} y_i = \sum_{i=1}^{3}\sum_{j=1}^{3} a_{ij}x_j = \sum_{j=1}^{3}\sum_{i=1}^{3} a_{ij}x_j = \sum_{j=1}^{3} x_j \sum_{i=1}^{3} a_{ij} = \sum_{j=1}^{3} x_j = 1.$$

Since the map is continuous, it has a fixed point by Brouwer's Fixed Point Theorem. (Notice the parallel with invariant measures in e.g IB Markov Chains.)

4   Nash's Theorems

This section is mainly about applications of analysis in game theory and economics.

Conside a two player game with two choices for each player [this generalises to 2 players with n, m choices respectively]. Suppose person A has strategy $p = (p_1, p_2) = (p_1, 1-p_1)$, i.e choose 1 with probability p_1 and 2 with probability p_2 and player B has strategy $q = (q_1, q_2)$. The value of the game to A is

$$A(p, q) = \sum p_i a_{ij} q_j.$$

And the value of the game to B is

$$B(p, q) = \sum p_i b_{ij} q_j.$$

Where a_{ij} and b_{ij} are the gains to A and B respectively if A plays i and B plays j. [In IB Optimisation we had $a_{ij} = -b_{ij}$, i.e the games were zero sum. We will not assume the game is zero sum.]

Nash showed that there always exist p^*, q^* such that $A(p^*, q^*) \geq A(p, q^*)$ and $B(p^*, q^*) \geq B(p^*, q)$, i.e neither A nor B have any incentive to change from p^*, q^* respectively if they know each other's strategy. Such p^*, q^* are called *Nash stable points*.

Proof. We work in $\mathbb{R}^4$ on

$$\Gamma = \{(p_1, p_2, q_1, q_2) : p_1, p_2, q_1, q_2 \geq 0, \; p_1 + p_2 = q_1 + q_2 = 1\}.$$

We set $u_1(p, q) = \max\{A((1,0), q) - A((p_1, p_2), q), 0\}$ so $u_1 > 0$ means "moving p_1 towards 1 is good", $u_1 = 0$ means "moving p_1 towards 1 is not good".

Set $u_2 = \max\{A((0,1), q) - A((p_1, p_2), q), 0\}$ and define v_1, v_2 similarly for B. Set

$$F(p, q) = \left(\frac{p_1 + u_1(p, q)}{1 + u_1(p, q) + u_2(p, q)}, \frac{p_2 + u_2(p, q)}{1 + u_1(p, q) + u_2(p, 1)}, \dots \right) \in \mathbb{R}^4$$

Then F is continuous, $F(p,q) \in \Gamma$. Hence F is a map from Γ to itself, thus has a fixed point (p^*, q^*). Since u_1, u_2 cannot both be non-zero, wlog $u_2 = 0$. Then at (p^*, q^*),

$$\frac{p_1 + u_1}{1 + u_1} = p_1.$$

So either $u_1 = 0$ or $p_1 = 1$. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad\square$

Let us compute the Nash stable points for an example:

Example (Chicken). A,B drive towards each other. If both swerve they both lose 1 point. If one swerves and the other does not the swerver loses 5 points, and the non-swerver gains 10 points. If neither swerve both lose 100 points. We want to find $A(p,q)$ (the value of the game to A) and then find q such that A has no incentive to change.

$$A(p,q) = -pq - 5p(1-q) + 10(1-p)q - 100(1-p)(1-q).$$

Differentiating:

$$\frac{\partial A}{\partial p} = -q - 5(1-q) - 10q + 100(1-q).$$

If we are at an interior point $(p, q \notin \{0, 1\})$ then we must have $\frac{\partial A}{\partial p} = 0$. Substituting gives $0 = -q - 5(1-q) - 10q + 100(1-q) = 95 - 106q$, i.e $q = 95/106$. By symmetry $(95/106, 95/106)$ is a Nash point, and the only interior Nash point. There are other Nash points, namely $(1, 0)$ and $(0, 1)$.

Cooperation

Consider n parties negotiating in good faith. We have a list of assumptions:

- $\Gamma \subseteq \mathbb{R}^n$ is the set of possible choices $x = (x_1, \ldots, x_n)$, where x_j corresponds to the outcome for j;

- We make Γ closed and bounded;

- Suppose Γ is convex (indeed could choose $a \in \Gamma$ with probability p and $b \in \Gamma$ with probability $1 - p$ so $pa + (1-p)b \in \Gamma$);

- We say some $s \in \Gamma$ is the *status quo* point, i.e the result if no agreement is reached;

- Pareto optimality [secret of economics or politics is maximising $F : A \to \mathbb{R}$ makes sense but maximising $f : A \to \mathbb{R}^2$ does not]. We demand Parato optimality: $y \in \Gamma$ is *Pareto optimal* if for $x \in \Gamma$, $x_j \geq y_j$ for all j implies $x = y$;

- Independence of unrelated alternatives: if $\Gamma' \supseteq \Gamma$ and y is Nash optimal for (Γ', s), then if $y \in \Gamma$, y is Nash optimal for (Γ, s);

- Symmetry: if Γ is such that if $(x_1, \ldots, x_n) \in \Gamma$, then $(x_{\sigma(1)}, \ldots, x_{\sigma(n)}) \in \Gamma$ for all permutations $\sigma \in S_n$ and $s = (s, s, s, \ldots, s)$, then the Nash optimum is symmetric (i.e $(y, y, \ldots, y)$ for some $y \in \mathbb{R}$);

- If T is an affine map $\mathbb{R}^n \to \mathbb{R}^n$, i.e $T(x_1, \ldots, x_n) = (a_1 + \lambda_1 x_1, \ldots, a_n + \lambda_n x_n)$ with $\lambda_j > 0$, then if y is Nash optimal for (Γ, s), Ty is optimal for $(T\Gamma, Ts)$.

We briefly justify the final assumption: translation should not affect the answer. So wlog we take $s = 0$.

Theorem. *Under all these assumptions there exists a Nash optimum and it is unique.*

We first prove a preliminary lemma:

Lemma. *Suppose $s = 0$, $(1, 1, 1, \ldots, 1) \in \Gamma$ and $\prod x_i \leq 1$ for all $x \in \Gamma$, $x_i \geq 0$. Then $x + 1 + x_2 + \ldots + x_n \leq n$ for all $x \in \Gamma$.*

Remark. This is a consequence of a version of Hahn-Banach.

Proof. Suppose $x \in \Gamma$, then since $(1, 1, 1, \ldots, 1) \in \Gamma$ and by convexity of Γ, $\delta x + (1 - \delta)1 \in \Gamma$ fo all $0 \leq \delta \leq 1$. So if δ is small and positive

$$\prod_{i=1}^{n} (\delta x_i + (1 - \delta)) \leq 1.$$

By hypothesis. So

$$1 \geq \prod_{i=1}^{n} (1 + \delta(x_i - 1)) = 1 + \delta \sum_{i=1}^{n} (x_i - 1) + \delta^2 A(\delta),$$

for some $A(\delta)$ bounded as $\delta \to 0$. Thus $0 \geq \sum_{i=1}^{n} (x_i - 1) - \delta A(\delta)$ and so taking $\delta \to 0$, $\sum_{i=1}^{n} (x_i - 1) \leq 0$. $\square$

Proof of theorem. By the above lemma we have $\{x \in \Gamma : x_i \geq 0\} \subseteq \Sigma := \{x : x_i \geq 0, \sum_{i=1}^{n} x_i \leq n\}$. If we look at the problem $(\Sigma, 0)$, we need to find a point x satisfying conditions:

(a) Symmetry, i.e $x_1 = x_2 = \ldots = x_n$;

(b) Pareto optimality, so $x_1 = \ldots = x_n = 1$.

Now suppose the status quo (for the problem in Γ) is at 0. Then $f(x) = \prod_{i=1}^{n} x_i$ is continuous so will attain a maximum on Γ (by compactness of Γ), say at $y \in \Gamma$. Then for all $0 \leq x \in \Gamma$, $\prod_{i=1}^{n} (x_i/y_i) \leq 1$. So by scaling there is a y satisfying all our conditions. So there is always a solution. $\square$

5   Approximation by Polynomials

General principle - "nice functions look like polynomials". However there are limits to this view.

Example (Cauchy example)**.** Consider the function $E : bbR \to \mathbb{R}$ defined by

$$E(t) = \begin{cases} \exp(-1/t^2) & t > 0 \\ 0 & t \leq 0 \end{cases}.$$

If $t \neq 0$, then inductively $E^{(n)}$ exists and $E^{(n)} = P_n(1/t)E(t)$ for some polynomial P_n. Also inductively we have that $E^{(n)}(0)$ exists with value 0. In particular all the coefficients of a Taylor expansion around 0 are 0, so E is not analytic.

Consider interpolation: for $f \in C[a,b]$ and $x_0, x_1, \ldots, x_n \in [a,b]$ distinct, there is a unique polynomial of degree at most n such that $P(x_j) = f(x_j)$.

Uniqueness: if two polynomials of degree $\leq n$ agree on $x_0, \ldots, x_n$, they must be equal else their difference is a polynomial of degree $\leq n$ with $n+1$ roots.

Existence: define $e_j(t) = \frac{\prod_{i \neq j}(t - x_i)}{\prod_{i \neq j}(x_j - x_i)}$. Then $e_j(x_i) = 0$ for $i \neq j$ and $e_j(x_j) = 1$. Consider the polynomial $P(t) = \sum_{j=0}^{n} f(x_j) e_j(t)$. Then $P(x_k) = f(x_k)$ for all k.

Denote by $\mathcal{P}_n$ the set of polynomials of degree at most n.

Chebyshev Polynomials: there exists a polynomial $T_n \in \mathcal{P}_n$ such that $T_n(\cos\theta) = \cos(n\theta)$, and there exists $U_{n-1} \in \mathcal{P}_{n-1}$ such that $U_{n-1}(\cos\theta) = \frac{\sin(n\theta)}{\sin\theta}$.

We have

$$\cos n\theta + i \sin n\theta = (\cos\theta + i\sin\theta)^n$$

$$= \sum_{0 \le 2r \le n} (-1)^r \binom{n}{2r} \cos^{n-2r}(\theta) \sin^{2r}(\theta)$$

$$+ i \sum_{0 \le 2r+1 \le n} (-1)^r \binom{n}{2r+1} \cos^{n-(2r+1)}(\theta) \sin^{2r}(\theta) \cdot \sin\theta$$

$$= \sum_{0 \le 2r \le n} (-1)^r \binom{n}{2r} \cos^{n-2r}(\theta)(1 - \cos^2\theta)^r$$

$$+ i \sin\theta \sum_{0 \le 2r+1 \le n} (-1)^r \binom{n}{2r+1} \cos^{n-(2r+1)}(\theta)(1 - \cos^2\theta)^r$$

$$= T_n(\cos\theta) + i \sin\theta \, U_{n-1}(\cos\theta)$$

with T_n a polynomial of degree n and U_{n-1} a polynomial of degree $n-1$, $T_n(\cos\theta) = \cos n\theta$, $U_{n-1}(\cos\theta) = \frac{\sin n\theta}{\sin\theta}$.

We look at the coefficient of t^n in $T_n(t)$. Since

$$T_n(\cos\theta) = \sum_{0 \le 2r \le n} (-1)^r \binom{n}{2r} \cos^{n-2r}(\theta)(1 - \cos^2\theta)^r$$

we have $T_n(t) = \sum_{0 \le 2r \le n} (-1)^r \binom{n}{2} t^{n-2r}(1 - t^2)^r$. So the coefficient of t^n is $\sum_{0 \le 2r \le n} \binom{n}{2r} = \frac{1}{2}\left((1+1)^n + (1-1)^n\right) = 2^{n-1}$.

Is it true that functions look like polynomials over some range and in some sense?

Theorem (Weierstrass). *If $f \in C([a,b])$ and $\varepsilon > 0$, there exists a polynomial p such that $||f - p||_\infty < \varepsilon$.*

We will use Bernstein's probabilist proof.

Theorem (Bernstein). *If $f \in C([0,1])$, let $F_n(t) = \sum_{r=0}^n \binom{n}{r} f\left(\frac{r}{n}\right) t^r(1-t)^{n-r}$. Then $F_n \to f$ uniformly.*

Proof. Idea: let $X_1, X_2, \ldots, X_n$ be iid random variables with $\mathbb{P}(X_j = 1) = t$, $\mathbb{P}(X_j = 0) = 1 - t$. Then writing $\bar{X} = \frac{X_1 + \ldots + X_n}{n}$, we have

$$\mathbb{E}f(\bar{X}) = \sum_{r=0}^n \mathbb{P}\left(\bar{X} = \frac{r}{n}\right) f\left(\frac{r}{n}\right) = \sum_{r=0}^n \binom{n}{r} t^r(1-t)^{n-r} f\left(\frac{r}{n}\right).$$

We have $\mathbb{E}X_1 = t$, $\mathbb{E}X_1^2 = t$ so $\text{Var}(X_1) = t(1-t) \le \frac{1}{4}$. Hence $\text{Var}(\bar{X}) \le \frac{1}{4n}$. Since f is continuous on $[0,1]$ it is also uniformly continuous. Hence given $\varepsilon > 0$

there exists $\delta > 0$ such that $|f(x) - f(y)| < \varepsilon$ whenever $|x - y| < \delta$. So f is bounded by $||f||_\infty < \infty$. Now we have

$$
\begin{aligned}
|\mathbb{E}f(\bar{X}) - f(t)| &= |\mathbb{E}(f(\bar{X}) - f(t))| \\
&\leq |\mathbb{E}(\mathbb{1}_{|\bar{X}-t|\leq\delta}|f(\bar{X}) - f(t)|)| + |\mathbb{E}(\mathbb{1}_{|\bar{X}-t|>\delta}|f(\bar{X}) - f(t)|)| \\
&\leq \mathbb{E}(\mathbb{1}_{|\bar{X}-t|\leq\delta}\varepsilon) + \mathbb{E}(\mathbb{1}_{|\bar{X}-t|>\delta}2||f||_\infty) \\
&\leq \varepsilon + 2||f||_\infty \mathbb{P}(|\bar{X} - t| > \delta) \\
&\leq \varepsilon + 2||f||_\infty \frac{\mathrm{Var}(\bar{X})}{\delta^2} \\
&\leq \varepsilon + ||f||_\infty \frac{1}{2n\delta^2}.
\end{aligned}
$$

$\square$