MAT 429 Notes

Max Chien

Fall 2025

Contents

Introduction

Chapter 1

Preliminaries

Definition 1.1

$C^m(\mathbb{R}^n)$ denotes the class of functions $f : \mathbb{R}^n \to \mathbb{R}$ which have m continuous derivatives on $\mathbb{R}^n$, and moreover are *bounded.* Accompanying this, a norm is defined by

$$\|f\|_{C^m(\mathbb{R}^n)} = \max_{|\alpha|\le m} \sup_{\mathbb{R}^n} |\partial^\alpha f|$$

Hence $(x \mapsto x) \notin C^m(\mathbb{R})$ for any m. Also, the norm is equivalent to taking the sum of such suprema. The problem at hand, then, is to consider a function $f : E \to \mathbb{R}$, where $E \subseteq \mathbb{R}^n$ is compact, and ask how we might extend it to a $C^m(\mathbb{R}^n)$ function on all of $\mathbb{R}^n$. Moreover, supposing that such a function exists, we ask for bounds on the $C^m(\mathbb{R}^n)$ norm of such an extension. We also seek formulas for the extension in terms of f, and for bounds on th ederivatives at points close to E.

We will also consider the particular case where E is a finite set. Using bump functions, the first question is trivial; we can easily create extensions that are in $C^\infty(\mathbb{R}^n)$. But this is not a particularly reasonable or useful choice for interpolation or extrapolation. Instead, we will attempt to find extensions which minimize the $C^m(\mathbb{R}^n)$ norm over solutions to the problem, up to some constant factor which depends only on m, n.

For the finite case, we will also attempt to find algorithms which can search for these minimizing solutions, and which minimize compute time and memory resources.

Theorem 1.1

Let $E \subseteq \mathbb{R}^2$ be finite with $|E| \geq 6$, and consider $C^2(\mathbb{R}^2)$. Let $f : E \to \mathbb{R}$. For any 6 distinct points $G \subseteq E$, let

$$g(G) = \inf_{\substack{F \in C^m(\mathbb{R}^n) \\ F|_G = f|_G}} \|F\|_{C^m(\mathbb{R}^n)}$$

Then there is a constant C, not depend on f, such that

$$\max_{\substack{G \subseteq E \\ |G| = 6}} g(G) \leq C \inf_{\substack{F \in C^m(\mathbb{R}^n) \\ F|_E = f|_E}} \|F\|_{C^m(\mathbb{R}^n)}$$

The above is not true for 5 instead of 6.

Another problem, which leads to the Whitney extension theorem, asks if and how we may extend functions which have not just their values but also their derivatives prescribed on an initial set.

To illustrate why the n dimensional case is significantly harder than the 1 dimensional case, suppose E consists of many points on the x axis, and a single point which is offset in the y direction by $+\varepsilon$. Without this point, we have no information about y direction derivatives at other points nearby on the x axis. However, if we do have that point, we can use it to extrapolate bounds on the y derivative at all of the other points. Nevertheless, our algorithm needs to understand how to identify that this data must be obtained from that point.

Similarly, let P be a third degree polynomial in x, y, and let E be its zero set, with a single point lying just off of the curve. Once again we can use this point to estimate the gradient of our function nearby on E.

For another example, consider $C^4(\mathbb{R}^3)$, and let G be an algebraic surface $\{x : P = 0\}$, and consider a curve in G $\{x : P = Q = 0\}$. If all the points in E lie in or near G, and moreover there are points which are close to our curve, the interpolation algorithm should identify these patterns just from the input data E.

We can also make this useful for experiments by considering data which is prescribed with errors. Concretely, consider $E \subseteq \mathbb{R}^n$ finite for m, n fixed, and for $x \in E$ suppose we are given $f(x)$ and $\sigma(x) > 0$. Then we want to find functions $F \in C^m(\mathbb{R}^n)$ such that

$$|F(x) - f(x)| < \sigma(x) \qquad \forall x \in E$$

and such that $\|F\|_{C^m(\mathbb{R}^n)}$ is as small as possible up to a constant factor C. There exist algorithms which can compute a solution for $|E| = N$ in $O(N \log N)$ time, print queries in $O(\log N)$ time, and use $O(N)$ memory.

Example 1.1

Homework: We are given points $x_1, x_2, \ldots, x_N \in \mathbb{R}$ and are working in $C^2(\mathbb{R})$, with initial data $f(x_k) = y_k$. We want an algorithm which computes extensions $F \in C^2(\mathbb{R})$

such that $\sup|F''|$ is optimal up to a universal constant C, in the sense that if $\tilde{F}$ is also an extension, then $\sup|F''| \le C \sup\left|\tilde{F}''\right|$. Similarly we can look for computations of extension such that $\max_{k=0,1,2} \sup\left|F^{(k)}\right|$ is optimal to a constant. The algorithm should compute F with $O(N \log N)$ overhead, and computation of specific values $F(x)$ should occur in $O(\log N)$ time.

If we add error bars $\sigma(x_k) > 0$ to the above problem, it becomes an open problem, although it can be solved if a $1+\varepsilon$ factor is introduced and the constants allowed to depend on ε.

To extend this problem to arbitrary metric spaces (X, d), with $F : X \to \mathbb{R}^D$, we replace smooth functions with Lipschitz functions. Supposing for each $x \in X$ we are given a convex set $K(x) \subseteq \mathbb{R}^D$, we want to calculate F such that $F(x) \in K(x)$ for all x, and moreover the Lipschitz constant for F is as small as possible, up to a constant. This may be solved using a similar strategy as the previous 6-point method, with 2^D test points.

1.1 Well Separated Pairs Decomposition

Let $E \subseteq \mathbb{R}^n$ and $|E| = N$. Given $f : E \to \mathbb{R}$, compute the Lipschitz constant

$$L(f) = \max_{\substack{x,y \in E \\ x \neq y}} \frac{|f(x) - f(y)|}{|x - y|}$$

Clearly the constant may be computed in $O(N^2)$ time by using direct comparison of each pair. If we need to compute $L(f)$ merely within 1% error, we can do so in $O(N \log N)$ operations using **well separated pairs decomposition**.

Suppose we have to sets E', E'', and

$$\operatorname{dist}(E', E'') > A(\operatorname{diam}(E') + \operatorname{diam}(E''))$$

with A large, say 10^3. Then we may save a number of operations by computing

$$\max_{\substack{x' \in E' \\ x'' \in E''}} \frac{|f(x') - f(x'')|}{|x' - x''|}$$

The distances between the points may be bounded below easily by the separation assumption. We can also bound the numerator by computing the maximum and minimum values of f over points E', E''. The first step takes $O(1)$ operations, and the second takes $O(N)$ operations. To produce these sets, we will need more operations, but not $O(N^2)$.

Theorem 1.2: Well Separated Pairs Decomposition

Let $E \subseteq \mathbb{R}^n$ and $|E| = N$. Let $\mathcal{E} = E \times E \setminus$ (diagonal). Then $\mathcal{E}$ may be partitioned into $\nu_{\max}$ rectangular sets $E'_\nu \times E''_\nu, \nu = 1, \ldots, \nu_{\max}$ such that:

- For each ν, $\operatorname{dist}(E'_\nu, E''_\nu) \geq A(\operatorname{diam}(E'_\nu) + \operatorname{diam}(E''_\nu))$.
- $\nu_{\max} \leq CN$, where C depends on A and n.

Proof. Assume that $A \geq 10$ and $A \in \mathbb{N}$, say. We proceed using dyadic cubes. We'll use the convention that cubes are Cartesian products of half open intervals; that is, we begin with the unit cube (say) $[0,1)^n$ and subdivide further. Note that the problem is clearly invariant under scale, so we can assume $E \subseteq [0,1)^n$.

Choose any two dyadic cubes of side length $2^{-(k+A)}$ such that the distance between their centers is between $2^{-(k+1)}$ and 2^{-k}, and each contains a point in E. Call them Q'_ν, Q''_ν. Then set $E'_\nu = Q'_\nu \cap E, E''_\nu = Q''_\nu \cap E$, and take all such ν. This construction suffices as a partition, because any two points lie in exactly one such pair of dyadic cubes.

Let $\mathcal{Q}$ be the collection of all such (Q'_ν, Q''_ν), so that $\nu_{\max} = |\mathcal{Q}|$. Consider the subclass $\hat{\mathcal{Q}} \subseteq \mathcal{Q}$ which consists of those pairs (Q'_ν, Q''_ν) where Q'_ν, Q''_ν have sidelength $2^{-(k+A)}$, and they lie in the same dyadic cube of sidelength $2^{-(k-A)}$. In general the cube of sidelength $2^{-(k-A)}$ is much larger than the distance between the two cubes, so that nearly all pairs in $\mathcal{Q}$ should lie in $\hat{\mathcal{Q}}$.

For pairs which do not, Q''_ν lies in the margin of width $2^{-(k-2)}$ around the $2^{-(k-A)}$ dyadic square containing Q'_ν. If we shift the entire dyadic grid by $1/3$ in one of the dimensions, and repeat the procedure, we will capture some of the pairs in this margin. Repeating in each dimension, and each subset of the possible dimensions (so 2^n of them), we will capture all of the problematic pairs. We may overcount the pairs, and we also may create new bad pairs, but the point is that if one shows that the set of good pairs is bounded CN for any one of these shifts, then the entire set of pairs is boudned by $2^n CN$. So it suffices to consider the good pairs.

Consider the tree of dyadic cubes beginning with $[0,1)^n$ and including all children which have at least one point of E. For any children which have at least two points of E, we include their children as well, and so on. Note that any path will terminate once a cube is reached which contains exactly one point of E. Call any node which has at least two children a "branching node." For instance, for any pair $(Q'_\nu, Q''_\nu) \in \hat{\mathcal{Q}}$, there is a branching node at most $2A$ generations above the pair (since at worst, the $2^{-(k-A)}$ cube containing them both will be a branching node). The association from (Q'_ν, Q''_ν) to their first mutual ancestor may be many-to-one, but it is bounded-to-one, since the ancestor has a bounded number of descendant pairs at most $2A$ generations deep. Thus we just need to count the number of branching nodes in order to count $\hat{\mathcal{Q}}$.

Note that the number of branching nodes in a tree with N leaves is at most $N - 1$. $\square$

Having chosen a well separated pairs decomposition, we can pick a representative point

$(x'_\nu, x''_\nu) \in E'_\nu \times E''_\nu$ for each ν. Then (say $A = 10^6$), $L(f)$ is approximated by

$$\max_\nu \frac{|f(x'_\nu) - f(x''_\nu)|}{|x'_\nu - x''_\nu|} \leq \max_{\substack{x,y \in E \\ x \neq y}} \frac{|f(x) - f(y)|}{|x - y|} \leq (1 + 10^{-3}) \max_\nu \frac{|f(x'_\nu) - f(x''_\nu)|}{|x'_\nu - x''_\nu|}$$

The first inequality is clear. To prove the second, observe the following:

Proposition 1.3

For $A = 10^6$, if $|f(x'_\nu) - f(x''_\nu)| \leq |x'_\nu - x''_\nu|$ for all representatives, then it follows that for any $x, y \in E$, $|f(x) - f(y)| \leq (1 + 10^{-3})|x - y|$.

Proof. Suppose not. Then $|f(x) - f(y)| > (1 + 10^{-3})|x - y|$ for some $x, y \in E$. Pick such x, y where $|x - y|$ is minimzed among pairs of points satisfying this. In particular $x \neq y$, so $(x, y) \in E'_\nu \times E''_\nu$, and we take the representative pair (x'_ν, x''_ν). We have

$$\begin{aligned} |x - x'_\nu| &\leq \operatorname{diam}(E'_\nu) \\ |y - x''_\nu| &\leq \operatorname{diam}(E''_\nu) \\ \implies |x - y| &\geq \operatorname{dist}(E'_\nu, E''_\nu) \geq 10^6(\operatorname{diam}(E'_\nu) + \operatorname{diam}(E''_\nu)) \geq |x - x'_\nu| + |y - x''_\nu| \end{aligned}$$

In particular, since (x, y) were the minimal counterexample, (x, x'_ν) and (y, x''_ν) are not counterexamples. Hence

$$\begin{aligned} |f(x) - f(x'_\nu)| &\geq (1 + 10^{-3})|x - x'_\nu| \\ |f(y) - f(x''_\nu)| &\geq (1 + 10^{-3})|y - x''_\nu| \end{aligned}$$

Also by assumption,

$$|f(x'_\nu) - f(x''_\nu)| \leq |x'_\nu - x''_\nu|$$

so

$$|f(x) - f(y)| \leq |x'_\nu - x''_\nu| + (1 + 10^{-3})(|x - x'_\nu| + |y - x''_\nu|) \leq (1 + 10^{-3})|x - y|$$

But this contradicts the assumption that (x, y) is a counterexample. $\square$

1.2 Whitney Extension Theorem

Definition 1.2

Let $f \in C^m(\mathbb{R}^n)$, and let $x \in \mathbb{R}^n$. The **jet** of f at x i the function defined by

$$J_x F(y) = \sum_{|\alpha| \leq m} \frac{1}{\alpha!} \partial^\alpha f(x)(y - x)^\alpha$$

Suppose we are given $E \subseteq \mathbb{R}^n$ compact, and for each $x \in E$ we are given an m degree polynomial P^x on $\mathbb{R}^n$. We want to know what conditions allow us to find a function

$f \in C^m(\mathbb{R}^n)$ such that $J_x(f) = P^x$ on E. This problem gives rise to the Whitney extension theorem.

Clearly a necessary condition is that the derivatives

$$\left\{ \left| \partial_y^\alpha P^x \Big|_{y=x} \right| : x \in E, |\alpha| \leq m \right\}$$

are uniformly bounded, since any $C^m(\mathbb{R}^n)$ function must satisfy this. Of course, any individual P^x has uniformly bounded derivatives since E is compact, the point is that the family is uniformly bounded over $x \in E$.

Definition 1.3

Let $f : E \to \mathbb{R}$ be continuous and $E \subseteq \mathbb{R}^n$ compact. Then $\omega : \mathbb{R} \to \mathbb{R}$ is said to be a **modulus of continuity** if for all x, y,

$$|f(x) - f(y)| \leq \omega(|x - y|)$$

Then for $|\alpha|, |\beta| \leq m$, and $f \in C^m(\mathbb{R}^n)$, by Taylor's Theorem,

$$\left| \partial^\beta f(y) - \sum_{|\alpha| \leq m - |\beta|} \frac{1}{\alpha!} \partial^{\alpha+\beta} f(x)(y-x)^\alpha \right| \leq \omega(|x-y|)|y-x|^{m-|\beta|}$$

The same must be true, then, for the P^x, except that the modulus of continuity may be replaced with a uniform bound M, since the derivatives are bounded. Lastly, we have a kind of compatibility condition on the imposed derivatives:

$$\frac{\left| \partial^\beta P^y(y) - \sum_{|\alpha| \leq m-|\beta|} \frac{1}{\alpha!} \partial^{\alpha+\beta} P^x(x)(y-x)^\alpha \right|}{|y-x|^{m-|\beta|}} \xrightarrow[x,y \in E]{|y-x| \to 0} 0$$

Theorem 1.4: Whitney Extension Theorem

Suppose $E \subseteq \mathbb{R}^n$ compact, and let $\{P^x : x \in E\}$ be a family of m degree polynomials on $\mathbb{R}^n$. Then there exists $F \in C^m(\mathbb{R}^n)$ such that $J_x f = P^x$ for all $x \in E$ if and only if there exists $M < \infty$ such that

1. $\left| \partial^\beta P^x(x) \right| \leq M$ for all $x \in E, |\beta| \leq m$,
2. $\left| \partial^\beta \left(P^y - P^x\right)(y) \right| \leq M|y-x|^{m-|\beta|}$,
3. $\dfrac{\left| \partial^\beta \left(P^y - P^x\right)(y) \right|}{|y-x|^{m-|\beta|}} \xrightarrow[x,y \in E]{|y-x| \to 0} 0.$

Moreover, if M is a constant such that $1, 2, 3$ hold, then F may be chosen such that $\|f\|_{C^m(\mathbb{R}^n)} \leq CM$, where C depends only on m, n.

Proof. Let Q^0 be a cube containing E. For any cube Q, let Q^* be the cube with the same center as Q but 5 times (say) the sidelength. We say that Q is "good" if $Q^* \cap E = \varnothing$. Clearly Q^0 is bad. Then we bisect Q^0 into 2^n subcubes. For each subcube, if the subcube is good, we call it a "Whitney cube" derived from Q^0 (note that good cubes are defined without reference to Q^0, but Whitney cubes must be obtained from this process). If it is bad, we bisect it and consider the children.

Let $x \notin E$. Then there is a neighborhood around x contained in $Q^0 \setminus E$, so the chain of cubes containing x will eventually terminate at a Whitney cube. On the other hand, if $x \in E$ then clearly this chain never terminates. So the Whitney cubes form an almost pairwise disjoint family of cubes whose union is precisely $Q^0 \setminus E$.

Consider any Whitney cube Q, and let Q^+ be the dyadic cube which is its direct ancestor (we know it has one since Q^0 is not a Whitney cube). Then $Q^* \cap E = \varnothing$ since Q is good, but $(Q^+)^* \cap E \neq \varnothing$ because otherwise Q^+ would be a Whitney cube. So $\operatorname{dist}(Q, E) \sim_n \operatorname{len}(E)$.

Moreover, if Q, Q' are two Whitney cubes with $\overline{Q} \cap \overline{Q'} \neq \varnothing$, then their sidelengths differ by at most a factor of 2. Indeed, suppose not. Then taking Q to have the longer sidelenght, it must have at least 4 times the sidelength of Q'. Then $10Q'$ lies within $5Q = Q^*$, which does not intersect E, so the parent of Q' should have been chosen as a Whitney cube. Also, any point x may lie in at most $A_{m,n}$ slightly dilated Whitney cubes $(1.01)Q$.

Now, we attempt to construct F. For any point $x \in E$, we are forced to define $F(x) = P^x(x)$. On the other hand, let x_Q, Q be a point in E and cube such that $\operatorname{dist}(x_Q, Q) \leq CS_Q$, where S_Q is the sidelength. Since Q is close to E, it is reasonable to choose to define $F_Q = P^{x_Q}$ on Q. However, this can give discontinuities when cubes touch each other. Therefore, we instead define F_Q on $(1.01)Q$, and use a partition of unity to patch these together.

Let Q_0 be a reference cube with sidelength 1, and define θ_0 to be a C^∞ function that is supported on $(1.01)Q_0$ and equal to 1 on Q_0. For any Q, define $\tilde{\theta}_Q = \theta_0(x)\left(\frac{x - z_Q}{S_Q}\right)$ where z_Q is the center of Q. Let C_α be defined for any multiindex such that

$$|\partial^\alpha \theta_0(x)| \leq C_\alpha$$

Then

$$\left|\partial^\alpha \tilde{\theta}_Q(x)\right| \leq C_\alpha S_Q^{-|\alpha|}$$

Then define

$$\theta_Q(x) = \begin{cases} \frac{\tilde{\theta}_Q(x)}{\sum_{Q'} \tilde{\theta}_{Q'}(x)}, & x \in E^c \\ 0, & x \in E \end{cases}$$

There are no issues of convergence in the denominator since only finitely many terms are supported. Also, we have

$$\left|\partial^\alpha \sum_{Q'} \tilde{\theta}_{Q'}(x)\right| \leq C_\alpha A_{m,n} 2^{|\alpha|} S_Q^{-|\alpha|}$$

for $x \in (1.01)Q$. So for such x, the derivatives $|\alpha| \leq m$ of the numerator and denominator are bounded and the denominator is at least 1, so

$$|\partial^\alpha \theta_Q(x)| \leq A'_{m,n} S_Q^{-|\alpha|}$$

We now return to constructing F. As before, for $x \in E$, we have $F(x) = P^x(x)$. For $x \notin E$, we instead use the following: let x_Q, Q be a point in E and cube as chosen above. Then we define

$$F(x) = \begin{cases} P^x(x), & x \in E \\ \sum_Q \theta_Q(x) P^{x_Q}(x), & x \notin E \end{cases}$$

It now remains to verify that F is C^m and that its norm is bounded. The fact that this is true on E follows from the assumptions. Fix some cube $\hat{Q}$. Then

$$F(x) = P^{x_{\hat{Q}}} + \sum_Q \theta_Q \cdot (P^{x_Q} - P^{x_{\hat{Q}}})$$

So using the first and second assumptions,

$$\begin{aligned} |\partial^\alpha F(x)| &\leq |\partial^\alpha P^{x_{\hat{Q}}}(x)| + \left| \sum_Q \sum_{|\beta| \leq |\alpha|} \operatorname{coeff}(\alpha, \beta) \partial^\beta \theta_Q(x) \partial^{\alpha - \beta} (P^{x_Q} - P^{x_{\hat{Q}}})(x) \right| \\ &\leq |\partial^\alpha P^{x_{\hat{Q}}}(x)| + A_{m,n} \sum_{|\beta| \leq |\alpha|} |\operatorname{coeff}(\alpha, \beta)| A'_{m,n} S_Q^{-|\beta|} CM |x - y|^{m - (|\alpha| - |\beta|)} \end{aligned}$$

For any $Q, \hat{Q}$ where θ_Q is nonzero on $\hat{Q}$, then the distance between x_Q and $x_{\hat{Q}}$ is comparable to the sidelength. So we get the new bound

$$|\partial^\alpha F(x)| \leq M + A''_{m,n} M S_Q^{m - |\alpha|} \leq C_{m,n} M$$

$\square$

Definitions