Lecture notes on the differential equation method

Dieter Mitsche

The idea of the differential equation method is to approximate the trajectory of a random process by solutions of (deterministic) differential equations (whose behavior is easier to understand).

Example 1: Let $c > 0$, suppose cn balls are thrown sequentially u.a.r. into one of n bins (assuming $n \to \infty$). Let $X(i)$ be the random variable counting the number of empty bins after i balls have been thrown. Clearly, $X(0) = n$. The sequence $(X(i))_i$ is Markovian, and we have

$$X(i+1) = X(i) - \delta_{E_{i+1}},$$

where $\delta_{E_{i+1}} = 1$ if the $(i+1)$-st ball is thrown in an empty bin, and 0 otherwise. We have

$$\mathbb{P}(\delta_{E_{i+1}} = 1) = \frac{X(i)}{n},$$

and hence

$$\mathbb{P}(X(i+1) - X(i) = -1) = \frac{X(i)}{n},\ \mathbb{P}(X(i+1) = X(i)) = 1 - \frac{X(i)}{n},$$

and thus

$$\mathbb{E}(X(i+1) - X(i)) = -\frac{X(i)}{n}.$$

The idea is now to convert this difference equation ($X(i)$ is always integer) into a differential equation. Let $x(t)$ model the behavior of $\frac{1}{n}X(tn)$. That is, $x(t)$ can be interpreted as the scaled number of empty bins when tn balls have been thrown; or in other words, it is also the probability of a random bin to be empty after a total of tn balls being thrown. We have $x(0) = 1$ and

$$x(t + \frac{1}{n}) - x(t) = -\delta_{E_{tn+1}}.$$

For $n \to \infty$, assuming the change of the function equal to the expected changes, we have

$$dx/dt = -x.$$

We point out that this differential equation is *only suggested*, and the steps are not independent. The differential equation can be solved by separation of variables: we have $-dx/x = dt$, and we obtain $x(t) = Ce^{-t}$ for some constant $C \in \mathbb{R}$. Since $x(0) = 1$, we obtain $C = 1$. Thus, at time t around an e^{-t}-fraction of bins are empty. Concentration can be shown using martingales: suppose T balls to be thrown in total. For $j \geq 0$, define the random variables $Z_j = \mathbb{E}(X(T) \mid \mathcal{F}_j)$, with $\mathcal{F}_j$ to be the σ-algebra of measurable events after j balls have been thrown. Then $(Z_j)_j$ is a martingale, and we have $|Z_j - Z_{j-1}| \leq 1$. Then, by Azuma's inequality, $\mathbb{P}(|X(T) - \mathbb{E}(X(T))| > \sqrt{\alpha T}) \leq 2e^{-\alpha/2}$.

Example 2: Let G_0 be the empty graph on n vertices and let G_t the graph obtained as follows: choose a vertex u of minimum degree in G_{t-1} (arbitrarily among such vertices) and choose u.a.r. a vertex v not yet adjacent to u, and set $G_t = G_{t-1} \cup \{u, v\}$. How the degrees of the vertices of G_t evolve? In order to analyze this, first define $Y_i(t)$ to be the random variable counting the number of vertices of degree i in G_t. We split the analysis into phases: we are in phase k, if the minimum

degree in G_t is k (that is, we are in phase k, if k is the smallest integer with $Y_k(t) > 0$).
We look at Phase 0 in more detail. In phase 0, u has degree 0, and vertex v has probability $\frac{Y_i(t)}{n-1}$ to have degree i, $i \geq 0$. Clearly, $Y_0(t)$ decreases by either 1 or 2, and $Y_1(t)$ increases by two (if $deg(v) = 0$), by one (if $deg(v) \geq 2$), or it stays the same (if $deg(v) = 1$). Similarly, the behavior can be obtained for $Y_i(t)$ with $i \geq 2$. More precisely, define X_i to be the indicator random variable to be 1 if v has degree i, and 0 otherwise. We have

$$Y_0(t+1) = Y_0(t) - 1 - X_0,\ Y_1(t+1) = Y_1(t) + 1 + X_0 - X_1,\ Y_i(t+1) = Y_i(t) + X_{i-1} - X_i \text{ for } i \geq 2.$$

Hence

$$\mathbb{E}(Y_i(t+1) - Y_i(t) \mid G_t) = -\mathbf{1}_{i=0} + \mathbf{1}_{i=1} + \frac{Y_{i-1}(t) - Y_i(t)}{n-1},$$

with $Y_{-1}(t) = 0$ for all t. Now put $z_i(x) = \frac{1}{n} Y_i(xn)$. Again, assuming suggested changes equal to expected changes, and taking the limit as $n \to \infty$, we get

$$z_i'(x) = -\mathbf{1}_{i=0} + \mathbf{1}_{i=1} + z_{i-1}(x) - z_i(x),$$

with $z_{-1}(t) = 0$ for all t. We have $z_0(0) = 1$, $z_i(0) = 0$ for all $i > 0$. Solving iteratively, we obtain $z_0'(x) = -1 - z_0(x)$, which can be solved by separation of variables and yields $z_0(x) = 2e^{-x} + 1$. This can be used to obtain $z_1(x) = 2xe^{-x}$, and more generally, $z_i(x) = \frac{2x^i}{i!e^x}$ for any $i \geq 1$. As before, we can use Azuma's inequality to obtain concentration. Later phases are solved similarly (although the process is not Markovian, since some edges are forbidden because the original vertex u already has some neighbors, but as long as there are not too many forbidden edges, the analysis still works out fine to obtain the first order terms).

General setup. The idea is the following: suppose we have

$$\mathbb{E}(Y(i+1) - Y(i) \mid Y(i)) = F(i/n, Y(i)/n) + o(1)$$

for some deterministic well-behaved function F and suppose that $|Y(i+1) - Y(i)|$ is never too big. Then we could expect to have $Y(tn) = y(t)n + o(n)$, where $y(t)$ is the deterministic function being the unique solution to $y'(t) = F(t, y(t))$ and $y(0) = Y(0)/n$. The motivation behind this comes from stability of differential equations that shows that functions with similar initial values and similar derivatives remain close (recall that a real function $F(x_1, x_2)$ is L-Lipschitz-continuous, if $F(x_1, x_2) - F(x_1', x_2') \leq L \max\{|x_1 - x_1'|, |x_2 - x_2'|\}$ holds). We will use the following *inequality of Gronwall*:

Theorem 0.1. *Given a continuous function $x(t)$ defined on $[0, T]$, assume that there is $L \geq 0$ such that $x(t) \leq C + L\int_0^t x(s)ds$ for $t \in [0, T)$, for some $C \in \mathbb{R}$. Then $x(t) \leq Ce^{LT}$ for $t \in [0, T]$.*

Proof. Let $y(t) = L\int_0^t x(s)ds$. We have

$$(y(t)e^{-Lt})' = (y'(t) - Ly(t))e^{-Lt} = L(x(t) - y(t))e^{-Lt} \leq LCe^{-Lt}.$$

Integration gives

$$y(t)e^{-Lt} \leq -Ce^{-Lt} + C$$

and thus $x(t) \leq C + y(t) \leq Ce^{Lt}$. □

We then have the following lemma regarding stability:

Lemma 0.1. *Let $y(t)$ and $z(t)$ two real, differentiable functions defined on $[0,T]$ for some $T \in \mathbb{R}$, with $y(0) = \hat{y}$, and let $\lambda, \delta > 0$. Suppose that for all $t \in [0,T]$, $y'(t) = F(t, y(t))$, $|z(0) - \hat{y}| \leq \lambda$, $|z'(t) - F(t, z(t))| \leq \delta$ with F being L-Lipschitz-continuous on some bounded domain $D \subseteq \mathbb{R}^2$. Let $\sigma \in [0,T)$ be such that $(t, y(t))$ and $(t, z(t))$ are in D for all $[0,\sigma]$. Then, for all $t \in [0,\sigma]$, $|z(t) - y(t)| \leq (\lambda + \delta\sigma)e^{L\sigma}$.*

Proof. We have

$$\begin{aligned} |z(t) - y(t)| \leq& |z(0) - y(0)| + \int_0^t |z'(s) - y'(s)| ds \\ \leq& \lambda + \delta t + \int_0^t |F(s, z(s)) - F(s, y(s))| ds. \end{aligned}$$

By Lipschitz-continuity of F, and since $(t, y(t))$ and $(t, z(t))$ are in D for all $t \in [0, \sigma]$, for all such t,

$$|z(t) - y(t)| \leq (\lambda + \delta\sigma) + L \int_0^t |z(s) - y(s)| ds.$$

By Gronwall's inequality, for all such t,

$$|z(t) - y(t)| \leq (\lambda + \delta\sigma)e^{L\sigma}.$$

□

We can now state the main theorem, known as *the differential equation method*:

Theorem 0.2. *Given $n \geq 1$, a bounded domain $D \subseteq \mathbb{R}^2$, a function F with $F : D \to \mathbb{R}$, and σ-algebras $\mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \ldots$, suppose that the random variables $(Y(i))_i$ are $\mathcal{F}_i$-measurable for $i \geq 0$. Suppose also that for all $i \geq 0$, the following holds whenever $(i/n, Y(i)/n) \in D$:*

1. *$|\mathbb{E}(Y(i+1) - Y(i) \mid \mathcal{F}_i) - F(i/n, Y(i)/n)| \leq \delta$ for some $\delta \geq 0$, with F being L-Lipschitz continuous for $L \in \mathbb{R}$.*
2. *$|Y(i+1) - Y(i)| \leq \beta$ for some $\beta > 0$,*
3. *$|Y(0) - \hat{y}n| \leq \lambda n$ for some $\lambda > 0$, for some $(0, \hat{y}) \in D$.*

Then there are $R = R(D, F, L) \in [1, \infty)$ and $T \in (0, \infty)$ such that for $\lambda \geq \delta \min\{T, 1/L\} + R/n$, so that with probability at least $1 - 2e^{-n\lambda^2/(8T\beta^2)}$ we have

$$\max_{0 \leq i \leq \sigma n} |Y(i) - y(i/n)n| \leq 3e^{LT}\lambda n,$$

where $y(t)$ is the unique solution to $y'(t) = F(t, y(t))$, $y(0) = \hat{y}$, and $\sigma = \sigma(\hat{y})$ is any choice of $\sigma \geq 0$ such that $(t, y(t))$ has ℓ_∞-distance at least $3e^{LT}\lambda$ from the boundary of D for all $t \in [0, \sigma)$.

Idea of proof: We will not give the proof of this theorem but give some intuition: By the first two conditions and Azuma's inequality, with high probability,

$$Y(j) - Y(0) \approx \sum_{0 \leq i < j} \mathbb{E}(Y(i+1) - Y(i) \mid \mathcal{F}_i) = \sum_{0 \leq i < j} F(\frac{i}{n}, \frac{Y(i)}{n} \pm \delta)$$

and

$$y(j/n)n - y(0)n = \sum_{0 \le i < j} (y((i+1)/n) - y(i/n))n \approx \sum_{0 \le i < j} F(i/n, y(i/n)).$$

Comparing these expressions, we can bound $|Y(i) - y(i/n)n|$ using (a discrete variant) of Gronwall's inequality, yielding the desired conclusion (taking also into account the third condition).

Remark: The theorem can be extended to a system of differential equations. Also, large one-step changes are allowed, as long as they occur with small probability.

Exercises

1. Consider the process of randomly adding edges (one after the other, one edge at each step) to an initially empty graph on n vertices subject to the condition that every vertex has degree at most 2. Analyze the number of degree 0 vertices at the t-th step of the algorithm, for t such that $\frac{n-t}{\sqrt{n}} \to \infty$.

2. Given the following randomized algorithm for obtaining a maximum independent set in a random r-regular graph $G \in \mathcal{G}_{n,r}$:
   GIVEN $V = \{v_1, \ldots, v_n\}$ VERTICES FROM A $\mathcal{G}_{n,r}$ AND r
   LET $S = \emptyset$ AND $Z = V$
   while $Z \neq \emptyset$ REPEAT
   CHOOSE U.A.R. A VERTEX $v \in Z$ AND ADD IT TO S
   EXPOSE ALL THE r EDGES FROM v TO OTHER VERTICES
   REMOVE FROM Z: v AND ALL w S.T. vw HAS BEEN EXPOSED
   end while
   OUTPUT $\alpha = |S|$

   Calculate the expected value of α and show concentration around its expected value.

3. A k-SAT formula is a conjunction of clauses, where a clause is a disjunction of exactly k Boolean variables that appear either negatively or positively. For example, given Boolean variables x_1, x_2, x_3, x_4, a 3-SAT formula with 4 clauses is $\varphi = (x_1 \vee \bar{x_2} \vee x_3) \wedge (\bar{x_1} \vee x_4 \vee x_2) \wedge (\bar{x_1} \vee \bar{x_3} \vee \bar{x_2}) \wedge (x_1 \vee x_2 \vee x_4)$. A random k-SAT formula is such that among all formulas on $x_1, \ldots, x_n$ with m clauses one chooses uniformly at random one such formula (one way to do it is to choose for each clause the following: for each of the k variables to appear in the clause, choose one uniformly at random, and then flip an unbiased coin to decide whether it appears positively or negatively; note that this might yield multiformulas with repeated variables in one clause or repeated clauses, but for $m = cn$ with $c > 0$, and k constant, there is a positive probability for this not to happen.) A k-SAT formula is *satisfiable* if there exists an (Boolean) assignment to the variables $x_1, \ldots, x_n$ such that in each clause there is at least one satisfied variable: that is, if x_i is assigned 1, all clauses with x_i appearing positively are satisfied, and if x_i is 0, all clauses with x_i appearing negatively are satsfied. (Clearly, the bigger m, the harder it is to satisfy all clauses, and one might assume that for each k there is a threshold value of $m = m(k)$ in terms of satisfiability.

   - A *pure literal* is a literal (that is, a variable or the negation of a variable) whose complement does not appear in the formula. Given a random k-SAT formula with $m = cn$, consider the following *pure literal rule elimination algorithm*:
     GIVEN $\varphi = \{c_1, \ldots, c_m\}$ AND VARIABLES $x_1, \ldots, x_n$
     while THERE IS A PURE LITERAL AVAILABLE
     CHOOSE UNIFORMLY AT RANDOM ONE OF THEM
     ASSIGN THE VARIABLE OF THE CHOSEN LITERAL ACCORDINGLY (ASSIGN IT 1 IF THE LITERAL APPEARS POSITIVELY, AND 0 IF IT APPEARS NEGATIVELY)
     REMOVE ALL CLAUSES CONTAINING THIS LITERAL
     end while
     IF CLAUSES LEFT THEN FAIL

Analyze the previous algorithm and find out the value of $m = cn$ (note that c depends on k) until which the algorithm succeeds in finding a satisfying assignment.

- (You might restrict yourself to 3-SAT). Come up with different rules that can be analyzed using differential equations and yield an improved behavior (that is, they work even for a higher clause density).

References

[1] Warnke, L. A note on Wormald's differential equation method. Preprint available at https://arxiv.org/pdf/1905.08928.pdf.