Prof. dr hab. Szymon Peszat
WMiI UJ
Łojasiewicza 6
30-348 Kraków,
e-mail napeszat@cyf-kr.edu.pl

Kraków, 12 maja 2020 r.

Recenzja rozprawy doktorskiej mgr. Szymona Majewskiego p.t. "Langevin Monte Carlo as convex optimization in the space of measures"

Rozprawa poświęcona jest problemowi próbkowania i całkowania względem miary o gęstości

$$\pi(x) := \frac{\mathrm{e}^{-U(x)}}{\int_{\mathbb{R}^d} \mathrm{e}^{-U(x)}\mathrm{d}x},$$

gdzie $U\colon \mathbb{R}^d \mapsto \mathbb{R}$ jest zadaną funkcją (w pracy wypukłą). Problem jest istotny z punktu widzenia statystyki, uczenia maszynowego, fizyki matematycznej. Poświęcone mu są liczne monografie, publikacje w renomowanych czasopismach matematycznych. Jest związany z problemem optymalnego transportu. Łączy zaawansowaną matematykę z zastosowaniami.

Klasyczne metody polegają na konstrukcji, w oparciu o na przykład algorytm Metropolisa–Hastings, łańcucha Markowa, dla którego $\pi(x)\mathrm{d}x$ jest jedyną egodyczną miarą niezmienniczą.

Punktem wyjścia do podejścia opisanego w Rozprawie jest obserwacja, pochodząca chyba od Kolmogorowa, że $\pi(x)\mathrm{d}x$ jest miarą niezmienniczą dla procesu zadanego przez równanie Langevina

$$\mathrm{d}X_t = -\nabla U(X_t)\mathrm{d}t + \sqrt{2}\mathrm{d}W_t.$$

Można dyskretyzować równanie, ale dla otrzymanego łańcucha Markowa zwykle $\pi(x)\mathrm{d}x$ nie jest miarą niezmienniczą. W pracy istotnie modyfikuje się algorytm (Unadjusted Langevin Algorythm, w skrócie ULA) oparty na dyskretyzacji Eulera–Maruyamy równania Langevina. Oczywiście rodziny rozkładów równania Langevina zbiegają do $\pi(x)\mathrm{d}x$ w bardzo mocnym sensie (zbieżność w normie wahania). Dynamika rozkładów zadana jest (deterministycznym) równaniem Fokkera–Plancka. Algorytmy prezentowane w Rozprawie oparte są na pięknym wyniku Jordana, Kinderlehrera i Otto z 1998 r., który mówi, że rozwiązaniem równania Fokkera–Plancka jest gradient flow dla pewnego funkcjonału na przestrzeni Wassersteina. Innymi słowy, rodzina rozkładów rozwiązania równania spełnia pewną zasadę wariacyjną.

Obszerna 68-stronicowa rozprawa ma charakter małej monografii. Rozprawa składa się z pięciu rozdziałów i jednego dodatku. Pierwszy rozdział to wstęp, w którym sformułowano problem oraz przedstawiono jego kontekst. W rozdziale drugim przedstawiono podstawy teoretyczne potrzebne w Rozdziałach 3, 4 i 5. Matematyka w Rozdziale 2 jest bardzo piękna i zaawansowana. To analiza nieliniowa, analiza na przestrzeni Wasersteina, teoria optymalnego transportu. Taką matematyką zajmują się obecnie Ambrossio, Otto, Viliani i inni znakomici matematycy. Rozdział zawiera potrzebne definicje. Uważam, że Pan Majewski dobrze rozumie tą piekną i trudną matematkę.

Zasadnicze wyniki uzyskane przez Pana Majewskiego przedstawione są w Rozdziałach 3,4,5. Wyniki Rozdziału 3 podają oszacowania na jądra Markowa odpowiadające jednemu krokowi w algorytmie ULA. Dowody są bardzo zaawansowane, ciekawe i trudne. Poprawiono wyniki innych autorów. Poza nielicznymi literówkami nie znalazłem żadnych błędów. W Rozdziale 4 podane są mocne rezultaty dla $U = U_1 + U_2$, w których U_i mają różną regularność. Rozprawa zawiera wyniki z pracy A. Durmus, S. Majewski, B. Miasojedow "Analysis of Langevin Monte Carlo via Convex Optimization, J. Machine Learning Research 20 (2019), 1–46.

Jestem pod wrażeniem znajomości Pana Majewskiego bardzo trudnej, nowoczesnej i ważnej matematyki. Rozprawa jest napisana bardzo starannie i poza kilkoma drobnymi literówkami nie znalazłem, żadnych innych błędów. Jestem pewien, że autor zadał sobie dużo trudu by przedstawić jak najlepiej trudną tematykę. Niestety, paradoksalnie, łatwiej było mi zrozumieć wyniki czytając pracę źródłową! Dla mnie prezentacja Pana Majewskiego, oparta o wiele analogiach była niepotrzebnie zawiła.

Drugi poważniejszy zarzut: Doktorant przedstawił do oceny bardzo ciekawą mini monografie stojącą na bardzo wysokom poziomie. Monografia zawiera wyniki wielu matematyków. Brakuje mi bardziej dokładnego opisu, które wyniki są Pana Majewskiego. Na przykład w Rozdziale 3 na stronie 27 jest napisane "The results and discussions in this chapter are largely taken from Majewski at al. (2019)". Podobnie w Rozdziale 4 na końcu strony 37. Powinno to być dokładniej opisane. Na czym polega oryginalność wyników mogłem się dowiedzieć czytając pracę zródłową!

Ostatnia sprawa, którą traktuje tylko jako niezręczność. W Rozprawie cytowana jest praca źródłowa jako S. Majewski, A. Durmus, B. Miasojedow "Analysis of Langevin Monte Carlo via Convex Optimization, J. Machine Learning Research 20 (2019), 1–46. W oryginalnej pracy i w bazie MathSciNet jest A. Durmus, S. Majewski, B. Miasojedow.

Nie mam jednak wątpliwości, że praca spełnia wymagania stawiane rozprawom doktorskim i wnioskuję o jej przyjęcie. W trakcie obrony poproszę Pana Majewskiego o określenie w sposób dokładniejszy jego wkładu.

Szymon Peszat