

POSUDEK OPONENTA BAKALÁŘSKÉ PRÁCE

Název: Konvergence stochastického gradientu v úlohách strojového učení

Autor: Marie Jelínková

Shrnutí obsahu práce

Práce pojednává o algoritmech pro hledání minima účelové funkce, která má speciální tvar. Autorka demonstruje využití na učebnicovém příkladu farmáře.

V první kapitole jsou formálně zavedeny pojmy a základní tvrzení, které se v dalších kapitolách využívají v důkazech konvergence. V druhé kapitole je představen algoritmus pro případ, že gradient známe. Ve třetí kapitole je rozšíření algoritmu pro případ, že gradient musíme odhadovat. Ve čtvrté kapitole je přestavena krátká numerická studie na problému farmáře, která přibližuje chování algoritmu.

Celkové hodnocení práce

Téma práce. Téma strojového učení je velmi moderní a je postaveno na optimalizačních algoritmech, které jsou detailně rozebírány v práci. Zmínka o strojovém učení je ale jen v názvu a zadání práce, v celém textu, a hlavně pak v numerické studii, se strojové učení již nevyskytuje.

Vlastní příspěvek. Příspěvkem je rozepsání některých kroků v důkazech z citovaných článků a opravení nepřesností. Dalším příspěvkem je aplikace zkoumaných metod na příklad farmáře, aplikace teoretických poznatků na daný příklad je detailně popsána, všechny pojmy zdefinovány.

Matematická úroveň. Matematický text je přehledně strukturován do vět, lemmat a poznámek. Značení je přehledné a včas zavedeno. Matematická úroveň je dobrá, ale většinou jsou tvrzení převzata i s důkazy z literatury.

Práce se zdroji. Zdroje jsou korektně citovány, většina teoretické části práce je převzata z článku Ghandimi a kol. (2016). Vlastní příspěvek a převzatá část jsou vždy přehledně rozepsány před každým důkazem.

Formální úprava. Formální úprava je dobrá, text je přehledně a logicky rozdělen do kapitol a podkapitol, obrázky jsou popsány. Text neobsahuje žádné překlepy bránící snadnému pochopení textu.

Připomínky

1. Celkově bych klidně ocenil ještě podrobnější vysvětlení některých kroků, moje osobní preference je radši více matematického zápisu, než se pouze v krátkém textu odkázat na 4 předchozí tvrzení.
2. Strana 6, poznámka za (1.11), asi je potřeba předpokládat něco i o funkci h .
3. Strana 15, druhý řádek textu – překlep, jedná se o rovnost a ne nerovnost
4. Strana 23, nejdříve je prezentována optimální hodnota -118 000, na konci stránky je však hodnota -118 600.
5. Celá teoretická část pojednává o speciální formě účelové funkce, v numerické studii pak ale úplně vyhodíme h a uvažujeme jen velmi pěknou (konvexní) funkci f . Z názvu i zadání jsem doufal v aplikaci na složitější případy, a hlavně na algoritmy strojového učení (a reálná data).

Otázky

1. Dává v praxi nějaký smysl volit počet kroků R menší než maximální? Nemůžeme si pak jednoduše vybrat nejlepší řešení, i když zvolíme větší počet kroků? Hlavně když vidíme špatnou konvergenci algoritmu pro malé R .
2. Pokud jsem správně pochopil spojení s potenciálním strojovým učáním, pak speciální tvar účelové funkce $f + h$ reprezentuje součet ztrátové funkce a regularizace. Regularizace je

hlavním rozdílem strojového učení od konvenčních statistických postupů. Je možné využít dokázané teoretické poznatky například na LASSO (least absolute shrinkage and selection operator)? Jak by pak vypadal rozpad na f a h ? Jaké budou mít vlastnosti?

3. Na jaké jiné/další metody strojového učení lze prezentované algoritmy použít?

Závěr

Téma práce je velmi zajímavé a aktuální. Teoretická část práce je ve velké míře převzata z článku Ghandimi a kol. (2016).

Příspěvkem autorky je praktická aplikace algoritmů příklad farmáře. Autorka také velmi podrobně popisuje, proč a jak dané teoretické výsledky můžeme využít. Obecně chybí slibovaná aplikace na strojové učení a využití teoretických výsledků v plné míře.

Práci považuji za nadprůměrnou a doporučuji ji uznat jako bakalářskou práci.

Karel Kozmík

Katedra pravděpodobnosti a matematické statistiky

11.8.2022