

Examensarbete: Asymmetric Bregman Forward-Backward Splitting with Projection Correction

Första Ordningens Bregmanmetoder Applicerade på Primär-Dual-Problem

Max Nilsson

December 2022

Att hitta minimum av funktioner är det huvudsakliga syftet inom optimeringslära, med applikationer inom bland annat maskininlärning. En minimeringsalgoritm kan antingen förena och generalisera en samling existerande algoritmer, eller så kan den vara nyskapande och i sin tur inspirera nya algoritmer. Detta arbete introducerar Bregman NOFOB-algoritmen som både generaliserar kända algoritmer, såsom Bregman Chambolle-Pock, men också ger upphov till en ny algoritm.

Om minimeringsproblemet är två-dimensionellt ser funktionens graf ut som en samling toppar och dalar. Dess minimum är då den punkt med djupast dal. I högre dimensioner förlorar vi förmågan att visualisera grafen och då går det inte längre att avläsa funktionens minimum med ögat. Vi måste då förlita oss på algoritmer.

Givet en initial gissning listar en minimeringsalgoritm ut en lämplig nästa punkt som kan användas som input i algoritmens nästa iteration. En *första ordningens metod* använder funktionens lokala egenskaper, såsom dess gradient, för att beräkna uppdateringssteget. I många fall räcker inte denna information, utan funktionens globala geometri måste på något sätt inkorporeras i algoritmen. Detta arbete löser detta problem med hjälp av Bregmanfunktioner. Dessa Bregmanfunktioner används sedan i uppdateringssteget och bidrar där med ett globalt perspektiv.

Vi vill i slutändan att vår algoritm ska konvergera mot en lösning. Men vilka egenskaper måste Bregmanfunktionen då ha? Detta är en både subtil och

spännande fråga! Inte bara måste Bregmanfunktionen och minimeringsfunktionen passa bra ihop, men Bregmanfunktionen måste också uppföra sig rimligt för sig självt.

Dual-teorin är en av de mest intressanta delarna av optimeringslära. För varje givet optimeringsproblem (som vi kallar *primärproblemet*) finns det ett associerat *dualproblem*. Detta dualproblem utspelar sig ofta i en annan dimension än det ursprungliga primärproblemet. Under vissa naturliga förutsättningar kan man gå från lösningar i dualproblemet till lösningar i primärproblemet. Alltså, kan vi lösa dualproblemet så kan vi lösa primärproblemet, och vice versa. Detta är önskvärt om dualproblemet är lättare att lösa än primärproblemet.

I detta arbete applicerar vi första ordningens Bregmanmetoder på problem som löser primär-dual-problem, och vi kan därmed få lösningen till både primär- och dual-problemet samtidigt. Denna Bregmanalgoritm kallar vi för Bregman NOFOB-algoritmen.

Vi kommer att undersöka ett specifikt exempel som kallas *D-optimal design* problemet. Detta problem grundar sig i statistik analys. Dualproblemet till *D-optimal design* grundar sig förvånansvärt istället i geometri: givet ett antal punkter i ett högdimensionellt rum, vilken är den ellipsoid med minst volym som samtidigt täcker alla punkter? *D-optimal design* ett exempel på ett problem som går att lösa med Bregmanmetoder. Den nya Bregman NOFOB-algoritmen ger oss dessutom en möjlighet att lösa en utvidgning av *D-optimal design* problemet i primär-dual-form, som andra algoritmer inte klarar av.