MATHEMATISCHES INSTITUT
PROF. DR. FLORIAN JARRE

Heinrich Heine Universität Düsseldorf

17.11.2023

Einführung in die Optimierung – 6. Übungsblatt

Programmieraufgabe 3

Man implementiere das Verfahren des steilsten Abstiegs und die Momentum Methode jeweils in "Reinform" (wie in der Vorlesung) und mit gewichteten Mitteln (wie unten beschrieben) in Matlab/Octave oder Python oder R für den nachfolgenden Spezialfall.
Zu minimieren sei die konvexe quadratische Funktion $f : \mathbb{R}^n \to \mathbb{R}$ mit $f(x) \equiv \frac{1}{2} x^T D x$ mit einer Diagonalmatrix $D \succ 0$. Startpunkt sei stets der Vektor $(1, \ldots, 1)^T$.
Zu gegebener Konditionszahl $\texttt{KD} \geq 1$ sei D dabei gegeben durch die Diagonaleinträge
$D_{i,i} =$ `(1+(i-1)*(KD-1)/(n-1))` für $1 \leq i \leq n$.
Der stochastische Gradient an der Stelle `x` sei gegeben durch
`g = D*x + (epsgrad/sqrt(n)) * randn(n,1)`, wobei `epsgrad=0.1` gewählt werde und `randn` einen 0-1-normal verteilten (Pseudo-)Zufallsvektor erzeuge.
Für die feste Schrittweite α wähle man $\alpha = 1/\texttt{KD}$ in allen Verfahren. Für β wähle man den Wert $(1 - \sqrt{1/\texttt{KD}})^2$. In allen Fällen führe man `itmax` Iterationen aus, wobei `itmax` ein Eingabeparameter ist.
Die gewichteten Mittel $\bar{x}^k$ seien aus den Iterierten x^k der "Reinform" dabei jeweils wie folgt erzeugt: Initialisiere $w := 0$ und $\bar{x}^0 := 0$. Für $k \geq 0$ setze $\bar{x}^{k+1} := \bar{x}^k + k^3 * x^{k+1}$ sowie $w := w + k^3$.
Ganz zum Schluss setze $\bar{x}^{k+1} := \frac{1}{w} \bar{x}^{k+1}$.

Man überlege die Antwort auf folgende Fragen: (Braucht nicht abgegeben zu werden.)
Besitzt obiges D wirklich die Konditionszahl `KD`?
Man begründe: Nach Abschluss der Berechnung ist $\bar{x}^{k+1}$ eine Konvexkombination der Iterierten $x^1, x^2, \ldots, x^{k+1}$, d.h.

$$\bar{x}^{k+1} = \sum_{i=1}^{k+1} \lambda_i x^i \quad \text{mit} \quad \lambda_i \geq 0 \ \forall i \quad \text{und} \quad \sum_{i=1}^{k+1} \lambda_i = 1.$$

Ferner sind die Gewichte λ_i streng monoton wachsend, d.h. spätere Iterierte werden stärker gewichtet.

- Testen Sie das Programm für $n = 100$, $\texttt{KD} = 100$ sowie $\texttt{itmax} = 100$ und $\texttt{itmax} = 1000$ und vergleichen Sie die Ergebnsisse.
  Wie sehen die Ergebnisse aus wenn `epsgrad=0` gesetzt wird?

- Schicken Sie die erzeugten Dateien in der üblichen Form bis zum 15.12.2023 um 12:00 Uhr an die E-Mail Adresse `opt0@uni-duesseldorf.de`