

Ausgabe: 09.06.2022

Prof. H. Jelger Risselada

Abgabe: 16.06.2022 20 Uhr

Aufgabe 0: Verständnisfragen

0 Punkte

- 1) Nennen Sie Ihnen bekannte Verfahren zur Minimierung von Funktionen der Form $f: \mathbb{R}^N \rightarrow \mathbb{R}$ mit $N > 1$. Beschreiben Sie das generelle Vorgehen bei derartigen Minimierungsaufgaben und diskutieren Sie Vor- bzw. Nachteile der jeweiligen Verfahren.

Aufgabe 1: Mehrdimensionale Minimierung

12 Punkte

Optimierung, also Minimierung oder Maximierung, stellt nicht nur in der Physik ein wichtiges Problem dar. Traditionell werden hochdimensionale Funktionen mithilfe von Gradientenmethoden optimiert. Dabei lässt sich die Konvergenzgeschwindigkeit für komplizierte Funktionsverläufe verbessern, indem die Richtung der Gradienten intelligent angepasst wird. In diese Aufgabe sollen Sie zwei Gradientenverfahren und im Laufe dessen eine eindimensionale Optimierung implementieren, um sich mit den Verfahren vertraut zu machen und sie in der Praxis zu erproben.

Im Folgenden sollen die Minima der Funktion

$$f(x, y) = \left(1 + \frac{\exp(-10(xy - 3)^2)}{x^2 + y^2} \right)^{-1} \quad (1)$$

mithilfe des SteepestDescent-(SD-) und des ConjugatedGradient-(CG-)Verfahrens bestimmt werden. Beide Verfahren benötigen eine Methode zum Minimieren einer Funktion in 1D. Hierzu soll zuerst die Intervallhalbierung implementiert werden, welche dann von den 2D-Verfahren aufgerufen werden kann.

- a) Implementieren Sie eine Methode Intervallhalbierung, welche ein lokales Minimum einer Funktion $f(\mathbf{x} + \lambda \mathbf{p})$ findet - also von einem Startpunkt $\mathbf{x} \in \mathbb{R}^2$ ausgehend in die Richtung $\mathbf{p} \in \mathbb{R}^2$. Welche Bedingung müssen die drei Stützstellen des Intervalls zu jeder Zeit erfüllen? Wie stellen Sie sicher, dass diese Bedingung bei der Wahl der ersten drei Stützstellen erfüllt ist? Testen Sie Ihre Methode an der Funktion

$$g(x, y) = x^2 + x^2 + x^2 y^2 \quad (2)$$

mit den Startwerten

$$\mathbf{x} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \mathbf{p} = \begin{pmatrix} -2 \\ 0 \end{pmatrix}, \quad \text{und} \quad \mathbf{x} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad \mathbf{p} = \begin{pmatrix} -2 \\ 0 \end{pmatrix}. \quad (3)$$

Welchen Punkt findet Ihre Intervallhalbierung jeweils als lokales Minimum?

- b) Implementieren Sie nun den SD-Algorithmus und testen ihn an der Funktion (2) mit den Startpunkten

$$\mathbf{x}_0 \in \left\{ \begin{pmatrix} 1.5 \\ 1 \end{pmatrix}, \begin{pmatrix} 100 \\ 80 \end{pmatrix}, \begin{pmatrix} 100 \\ 100 \end{pmatrix} \right\} \quad (4)$$

- c) Implementieren Sie nun den CG-Algorithmus und testen Sie ihn genauso wie SD aus Aufgabenteil b).

Vergleichen Sie die beiden Methoden bezüglich ihrer Konvergenz für die verschiedenen Startpunkte

d) Untersuchen Sie nun die anfangs genannte Funktion (1) mit den Startwerten

$$\mathbf{x}_0 \in \left\{ \begin{pmatrix} 1.5 \\ 2.3 \end{pmatrix}, \begin{pmatrix} -1.7 \\ -1.9 \end{pmatrix}, \begin{pmatrix} 0.5 \\ 0.6 \end{pmatrix} \right\} \quad (5)$$

Funktionieren die Verfahren für alle drei Startbedingungen? Welche Minima finden Sie? Vergleichen Sie wieder die beiden Methoden miteinander. Besitzt $f(x, y)$ weitere Minima, die von den drei Startpunkten aus nicht gefunden wurden? *Hinweis: Ein Plot und eine analytische Analyse der Funktion (1) können hilfreich sein.*

Aufgabe 2: Partikelschwarmoptimierung

8 Punkte

Gradientenverfahren zur Extremumssuche weisen zwei entscheidende Nachteile auf: für nicht differenzierbare Funktionen lassen sich die für das Verfahren benötigten Gradienten bestenfalls durch mehrfache Funktionsauswertungen approximieren, was rechenintensiv und ungenau ist, und für Funktionen mit mehreren Extrema können die Verfahren in lokalen Extrema stecken bleiben. Für solche Funktionen bietet es sich an, evolutionäre Algorithmen zu verwenden. In dieser Aufgabe sollen Sie eine Funktion mit vielen lokalen Minima durch Partikelschwarmoptimierung minimieren, um sich mit dem Verfahren vertraut zu machen.

Der Partikelschwarm besteht aus N Positionen \vec{r} , die in jedem Iterationsschritt mit

$$\vec{r}_{n+1} = \vec{r}_n + \vec{v}_n \quad (6)$$

$$\vec{v}_{n+1} = \omega \vec{v}_n + c_1 r_1 (\vec{r}_n^{\text{pers. best}} - \vec{r}_n) + c_2 r_2 (\vec{r}^{\text{global best}} - \vec{r}_n) \quad (7)$$

aktualisiert werden, wobei $\omega \in [0, 1]$ und $c_1, c_2 \in \mathbb{R}_+$ feste Gewichtungsfaktoren sind und $r_1, r_2 \in [0, 1]$ in jedem Schritt und für jedes Teilchen neu gezogene zufällige Skalierungsfaktoren. Hier wählen wir $\omega = 0.8$ und $c_1 = c_2 = 0.1$ und ziehen r_1 und r_2 aus einer Gleichverteilung auf dem Intervall $[0, 1]$. Die Position $\vec{r}_n^{\text{pers. best}}$ speichert die beste bisher gefundene Position des n -ten Teilchens und $\vec{r}^{\text{global best}}$ speichert die beste bisher gefundene Position aller Teilchen.

Mithilfe der Partikelschwarmoptimierung soll Sie nun das globale Minimum der Funktion

$$f(x, y) = (x - 1.9)^2 + (y - 2.1)^2 + 2 \cos(4x + 2.8) + 3 \sin(2y + 0.6) \quad (8)$$

bestimmen.

- Initialisieren Sie 10 Teilchen zufällig verteilt auf dem Intervall $\vec{r}_i \in [-5, 5]^2$ mit zufälligen Startgeschwindigkeiten $\vec{v}_i \in [-1, 1]^2$. Lassen Sie die Partikelschwarmoptimierung für mindestens 100 Iterationen laufen. *Hinweis: Wenn Ihr Algorithmus korrekt funktioniert, sollten Sie am Ende ein globales Minimum mit $f(\vec{r}^{\text{global best}}) < -4$ finden.*
- Plotten Sie in einer einzelnen Abbildung sowohl den Verlauf des besten gefundenen Funktionswerts $f(\vec{r}^{\text{global best}})$ als auch den Verlauf der individuellen besten Werte $f(\vec{r}_n^{\text{pers. best}})$ über der Anzahl der Iterationen.
- Stellen Sie den Verlauf der Partikelschwarmsuche anschaulich mithilfe mehrerer Momentaufnahmen und/oder einer Animation dar. Hierzu sollten Sie die Positionen \vec{r}_n der Teilchen (als Marker) und deren Geschwindigkeiten \vec{v}_n (als Pfeile) sowie die bisher gefundenen besten Positionen $\vec{r}_n^{\text{pers. best}}$ und $\vec{r}^{\text{global best}}$ (beispielsweise als unterschiedlich aussehende Marker) darstellen. Es ist sinnvoll, die Funktion $f(x, y)$ als Heatmap und/oder Konturplot im Hintergrund anzuzeigen, um die Bewegung des Partikelschwarms deuten zu können.

Hinweis: Sie können den Algorithmus gerne auch für weitere Werte von (ω, c_1, c_2) und andere Startwerte anwenden und mit den hier gegebenen Parametern vergleichen.