Name: Matouš Dzivjak
Date: 9.11.2018

PCA

1 Zadání

`https://cw.fel.cvut.cz/wiki/courses/b33opt/cviceni/domaci_ulohy/kruznice/start`

2 Řešení

Ze zadání máme:

$$f(x) = \sum_{i=1}^{m} dist(x_i, a_i)^2$$

$$dist(x, a) = \sqrt{(a_0 - x_0)^2 + (a_1 - x_1)^2} - r$$

2.1

Mějme několik bodů $a_1, ..., a_m$ v obecné konfiguraci. Je funkce f všude diferencovatelná? Má jedno nebo více lokálních minim?

Funkce není spojite diferencovatelná protože v bodě $a_0 = x_0, a_1 = x_1$ nemá derivaci. Funkce má nekonečně mnoho lokálních minim. Např. pro jeden jediný bod bude funkce nabývat minimam v jakémkoliv bodě, pokud za poloměr dosadíme vzdálenost bodu od středu kružnice.

2.2

Diskutujte, jaký algoritmus je vhodný na minimalizaci funkce $f(x)$ a proč. Je možné, aby Gaussův-Newtonův algoritmus na naší úloze divergoval?

Gauss-Newtonova metoda konverguje pomaleji než ostatní metody. Výhodou je, že u ní není nutné počítat druhé derivace. Levenberg-Marquardtova metoda je kombinací Gauss-Newtonovo a gradientní metody - konverguje rychleji v okolí optima a zároveň zaručuje dostatečnou spolehlivost i ve větší vzdálenosti od optima. U Levenberg-Marquardtovi metody se také nestává, že by algoritmus nezkonvergoval, což se o Gauss-Newtonovo metodě říci nedá.

2.3

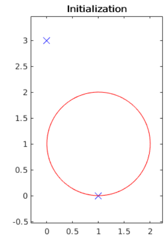
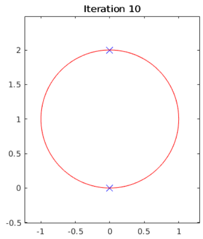
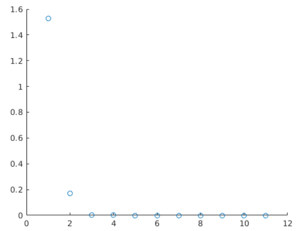
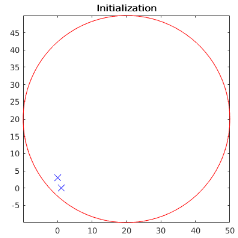
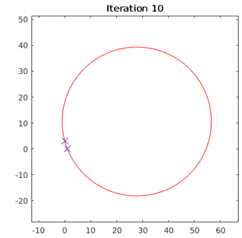
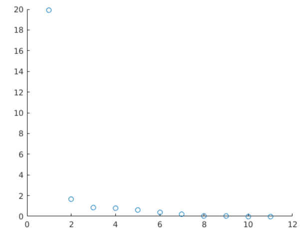
Může se zdát, že algoritmy na nelineární nejmenší čtverce bez omezení nejde použít, protože máme omezení $r \leq 0$. Vadí to? Co se stane, budeme-li toto omezení ignorovat? Můžou algoritmy konvergovat k řešení se záporným r?

V hodně specifických případech by taková možnost nejspíše mohla nastat. Nicméně záporne r by způsobilo nárůst criterionu (součet čtverců vzdáleností bodů od kružnice), protože zatím co pro kladná r se r odčítá pro záporná by se jeho velikost přičítala. Tedy ke konvergenci k zápornému číslu by dojít nemělo, což je důvod, proč jsem ve své implementaci tuto možnost neošetřoval a při testování nenastal nikdy problém.

2.4

Najděte nějakou množinu $m \leq 3$ bodů $\{a_1, a_2, ..., a_m\}$ a takovou dvojici počátečních parametrů kružnice $x_0^{(1)}$ a $x_0^{(2)}$, aby algoritmus inicializovaný těmito parametry skončil v různých lokálních minimech.

Za množinu bodů zvolíme: $\{(0, 3), (1, 0)\}$ pak pro $x_0^{(1)} = (1, 1, 1)$ dostáváme kružnici, kde každý bod leží na opačné straně kružnice, střed kružnice je na těžišti těchto dvou bodů a poloměr odpovídá jejich vzdálenosti od středu kružnice, zatím co pro $x_0^{(2)} = (20, 20, 30)$ dostáváme po konvergenci kružnici, kde oba body leží blízko sebe na kružnici a kružnice má velký poloměr

<table>
<tr><th>body a kružnice $x_0^{(n)}$</th><th>body a stav dokonvergovaný z $x_0^{(n)}$</th><th>graf f_history pro $x_0^{(n)}$</th></tr>
<tr><td>
</td><td>
</td><td></td></tr>
<tr><td>
</td><td>
</td><td></td></tr>
</table>