Linköpings universitet
Matematiska institutionen
Optimeringslära

Matematisk optimering

24 november 2020

Uppgift 4: Konjugerade riktningar och gradienter

1. Betrakta problemet att minimera

$$f(x) = x_1^2 + 2x_1x_2 + 2x_2^2 + x_3^2 - x_2x_3 + x_1 + 3x_2 - x_3.$$

   a) Tillämpa *Gram–Schmidt proceduren* på de tre enhetsvektorerna i R^3 för att beräkna tre konjugerade riktningar för detta problem.

   Verifiera att de tre riktningarna blev konjugerade.

   b) Lös problemet genom att utgående från origo göra minimeringar längs det tre konjugerade riktningarna.

   Verifiera optimalitet för den funna punkten

2. Låt matrisen $Q \in R^{n\times n}$ vara symmetrisk. Visa att egenvektorer som svarar mot olika egenvärden är Q-konjugerade.

3. Visa att en *Gram–Schmidt ortogonalisering* av kolumnerna i en matris $A \in R^{n\times n}$ kan användas för att göra en *QR-faktorisering* av A, det vill säga för att konstruera en ortonormalmatris $Q \in R^{n\times n}$ ($Q^{-1} = Q^{\mathrm{T}}$) och en övertriangulär matris $R \in R^{n\times n}$ sådana att $A = QR$.

4. Finn minimum till funktionen $f(x) = 2x_1^2 + x_2^2$ med hjälp av konjugerade gradientmetoden från startpunkten $x_0 = (3/2,\ 2)^{\mathrm{T}}$. Verifiera att de genererade sökriktningarna är konjugerade och verifiera optimalitet.

5. Gör en MATLAB-implementering av konjugerade gradientmetoden *Polak–Ribière* på problemet att minimera

$$f(x_1, x_2) = (x_1 - 2)^4 + (x_1 - 2x_2)^2.$$

   Starta i $x^0 = (0, 3)^{\mathrm{T}}$. Illustrera metodens uppträdande genom att plotta målfunktionsvärdet mot antalet iterationer och genom att plotta sekvensen av iterat i R^2. Redovisa MATLAB-koden.

   [Observera att konjugerade gradient-metoden kräver exakt linjesökning. Utnyttja en färdig MATLAB-funktion för detta eller implementera själva en enkel exakt linjesökning, till exempel baserad på ekvidistant sökning kombinerad med några iterationer av Newtons metod. Minimum antas i punkten $x^* = (2, 1)$ med $f(2, 1) = 0$. Utnyttja detta för att verifiera resultatet.]

6. a) Betrakta den vanliga konjugerade gradientmetoden tillämpad på problemet

$$\min_{x \in R^n} f(x) = \frac{1}{2} x^{\mathrm{T}} Q x - b^{\mathrm{T}} x,$$

där $Q \in R^{n \times n}$ är symmetrisk och positivt definit. Visa att varje iterationspunkt x^k minimerar f över den linjära mångfald som definieras av x^0 och det underrum som spänns upp av vektorerna $g^0, Qg^0, Q^2 g^0, \ldots, Q^{k-1} g^0$, där $g^0 = Qx^0 - b$.
[Ledning: Använd till exempel induktion med avseende på k.]

a) Betrakta det numeriska exemplet

$$Q = \begin{pmatrix} 2 & -1 & 0 & 0 \\ -1 & 2 & -1 & 0 \\ 0 & -1 & 2 & -1 \\ 0 & 0 & -1 & 2 \end{pmatrix}$$

och $b = (1, 0, -2, -\sqrt{5})^{\mathrm{T}}$. Låt $x^0 = (0, 0, 0, 0)^{\mathrm{T}}$. Avgör hur många iterationer som konjugerade gradientmetoden skulle behövas för att finna optimum, genom att studera antalet linjärt oberoende vektorer i sekvensen $g^0, Qg^0, Q^2 g^0, Q^3 g^0$.

7. Betrakta problemet

$$\min_{x \in R^n} f(x) = \frac{1}{2} x^{\mathrm{T}} Q x - b^{\mathrm{T}} x,$$

där $Q \in R^{n \times n}$ är symmetrisk och positivt definit. Låt punkterna $x', x'' \in R^n$ vara olika och låt riktningen $d \in R^n$ vara nollskild och icke-parallell med $x'' - x'$. Antag att punkterna x^1 och x^2 fås efter exakta minimeringar av f längs riktningen d från punkterna x' respektive x''. Visa att riktningen $x^2 - x^1$ är Q-konjugerad mot d.

8. a) Angrip problemet att minimera $f(x) = \frac{1}{2} x^{\mathrm{T}} Q x - b^{\mathrm{T}} x$, där

$$Q = \begin{pmatrix} 1001 & 1 & 1 & 1 & 1 \\ 1 & 999 & 1 & 1 & 1 \\ 1 & 1 & 101 & 1 & 1 \\ 1 & 1 & 1 & 99 & 1 \\ 1 & 1 & 1 & 1 & 10 \end{pmatrix}$$

och $b = (1, 1, 1, 1, 1)^{\mathrm{T}}$, med konjugerade gradientmetoden med analytisk linjesökning. Starta i origo. Plotta hur $f(x^k) - f(x^*)$ och $\|\nabla f(x^k)\| = \|Qx^k - b\|$ förändras med iterationerna. (Använd logaritmisk skala.)

b) Angrip problemet med den prekonditionerade konjugerade gradientmetoden, med

$$H = \begin{pmatrix} \frac{1}{1000} & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{1000} & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{100} & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{100} & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{10} \end{pmatrix},$$

och studera samma storheter som ovan.

c) Förklara skillnaden mellan resultaten i deluppgifterna ovan genom att studera egenvärdesstrukturerna för Q och HQ.