

Uppgift 5: Kvasi-Newton metoder

1. a) Verifiera Sherman–Morrison formeln

$$(A + uv^T)^{-1} = A^{-1} - \frac{A^{-1}uv^T A^{-1}}{1 + v^T A^{-1}u}.$$

- b) I den kvasi-Newton metod som utnyttjar den symmetriska rang-ett-uppdateringen av approximationen D_k av Hessian-inversen gäller att

$$D_{k+1} = D_k + \frac{(p^k - D_k q^k)(p^k - D_k q^k)^T}{q^{kT}(p^k - D_k q^k)},$$

där $p^k = x^{k+1} - x^k$ och $q^k = \nabla f(x^{k+1}) - \nabla f(x^k)$.

Detta svarar naturligtvis mot en uppdatering av en approximation $H_k = D_k^{-1}$ av Hessianen. Använd Sherman–Morrison formeln för att uttrycka H_{k+1} som en rang-ett-uppdatering av H_k .

[Ledning: Den sökta rang-ett-uppdateringen uppvisar symmetri med formeln ovan.]

2. a) Låt $a \in R$ och $v \in R^n$ med $\|v\| = 1$. Betrakta rang-ett matrisen $M = avv^T$. Visa att v är en egenvektor till M . Vilket är motsvarande egenvärde?
- b) Låt $D \in R^{n \times n}$ vara symmetrisk och positivt definit. Betrakta matrisen M definierad som ovan. Antag att $a < 0$. Visa att om $|a|$ är tillräckligt stor så är matrisen $A = D + M$ inte positivt definit.
3. a) Minimera den kvadratiske funktionen

$$f(x) = x_1^2 + 4x_2^2 - 4x_1 - 8x_2$$

med kvasi-Newton metoden DFP. Använd exakt linjesökning. Starta i $x^0 = 0$ och med $D_0 = I$. [Funktionen är konvex. Minimum fås uppenbarligen för $x = (2, 1)$.]

- b) Man kan visa att om metoden DFP utnyttjar exakt linjesökning och används på ett kvadratisk problem så genererar den konjugerade sökriktningar med avseende på $H = \nabla^2 f(x)$. Verifiera detta i exemplet ovan.

4. Gör en MATLAB-implementering av Broyden-familjen med Armijo-linjesökning. Använd specialfallen DFP och BFGS på problemet att minimera

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

Starta i $x^0 = (0, 3)^T$. Illustrera metodernas uppträdande genom att plotta målfunktionsvärdena mot antalet iterationer och genom att plotta sekvenserna av iterat i R^2 . Redovisa MATLAB-koden.

5. En avsevärd nackdel med kvasi-Newton metoder är att de för storskaliga problem kräver mycket minnesutrymme, eftersom approximationen D_k av inversen till Hessianen typiskt blir tät, det vill säga innehåller många nollskilda element. För att delvis kringgå detta problem kan man beräkna kvasi-Newton riktningen med hjälp av vektorerna $p^i = x^{i+1} - x^i$ och $q^i = \nabla f(x^{i+1}) - \nabla f(x^i)$, istället för med hjälp av approximationen D_k , och samtidigt begränsa antalet vektor-par som används, typiskt till de $m \ll n$ senaste. En sådan strategi kan baseras på följande resultat.

- a) Visa att BFGS-uppdateringen kan uttryckas som

$$D_{k+1} = V_k^T D_k V_k + \varrho^k p^k p^{kT}$$

,

där $\varrho^k = 1/q^{kT} p^k$ och $V_k = I - \varrho^k q^k p^{kT}$.

- b) Visa hur man lämpligast beräknar kvasi-Newton riktningen $d^k = -D_k \nabla f(x^k)$ utifrån den initiala approximationen D_0 och vektorerna p^i och q^i , $i = 0, 1, \dots, k-1$.