

LUCRAREA 10

PROGRAMAREA NELINIARĂ ÎN REZOLVAREA PROBLEMELOR DIN ENERGETICĂ. METODE DE ORDINUL 2

10.1. Aspecte generale

Utilizarea gradientului ca direcție de explorare este conformă cu aproximarea liniară a funcției obiectiv în dezvoltarea acesteia în serie Taylor. Totuși, o aproximare pătratică de forma:

$$F(X) = F(X^{(k)}) + (g^{(k)}, \Delta X^{(k)}) + \frac{1}{2} (d^{(k)}, H^{(k)} d^{(k)}) \quad (10.1)$$

este mai aproape de realitate. Minimul funcției se obține prin derivare în raport cu componentele lui $d^{(k)}$ și anularea derivatelor respective, de unde rezultă:

$$d^{(k)} = -H^{(k)^{-1}} g^{(k)} \quad (10.2)$$

Astfel, dacă se introduce (10.2) în relația iterativă,

$$X^{(k+1)} = X^{(k)} + \lambda^{(k)} d^{(k)}, \quad k = 1, 2, 3, \dots \quad (10.3)$$

rezultă:

$$X^{(k+1)} = X^{(k)} - \lambda^{(k)} \cdot H^{(k)^{-1}} g^{(k)} \quad (10.4)$$

unde $\lambda^{(k)}$ poate lua valori arbitrare sau optime. În cazul unor funcții obiectiv pătratice, minimul va fi atins încă din prima iterație, pentru $\lambda^{(0)} = 1$.

Aplicarea metodelor de ordinul 2, denumite și metode Newton, cere ca matricea $H^{(k)}$ să fie pozitiv definită la fiecare iterație. Pentru că aproximarea pătratică este superioară celei liniare, convergența metodei în apropierea optimului va fi sensibil mai

bună decât a metodelor de ordinul 1. În plus, disponibilitatea derivatelor de ordinul doi permite verificarea condițiilor de suficiență a optimului.

Totuși, aceste avantaje se pierd datorită volumului de calcul necesar obținerii derivatelor de ordinul doi și pentru inversarea matricei $H^{(k)}$. În plus, din cauza condiției $H^{(k)} > 0$ pot apărea situații în care metoda nu poate fi aplicată (cazurile în care $F(X)$ este liniară pe anumite porțiuni, iar $H^{(k)} = 0$).

De asemenea, pot apărea unele cazuri în care $F(X^{(k+1)}) > F(X^{(k)})$, deși $H^{(k)} > 0$, aproximația pătratică fiind valabilă doar într-o zonă foarte limitată. Pentru a evita aceste situații s-au propus diferite metode care constrâng matricea $H^{(k)}$ să fie pozitiv definită la fiecare iterație.

10.2. Metode de tip quasi-Newton

Un loc special în cadrul metodelor Newton îl ocupă așa-zisele metode quasi-Newton. Acestea din urmă au fost elaborate în scopul evitării principalelor dezavantaje ale metodelor Newton referitoare la inversarea matricei hessiene în fiecare iterație. Metodele quasi-Newton au la bază relația de recurență,

$$X^{(k+1)} = X^{(k)} - \lambda^k \cdot \tilde{H}^{(k)} g^{(k)} \quad (10.5)$$

unde matricea $\tilde{H}^{(k)}$ reprezintă o aproximație a lui $H^{(k)-1}$.

Relația (10.5) poate fi considerată ca fiind relația fundamentală care stă la baza metodelor ce utilizează derivatele. Astfel, în cazul în care matricea $\tilde{H}^{(k)}$ se identifică cu matricea unitate I_n , se obțin metodele de gradient, iar pentru cazul în care matricea $\tilde{H}^{(k)}$ se identifică cu inversa matricei hessiene $H^{(k)-1}$, se obțin metodele de tip Newton.

Cel mai performant algoritm din această categorie de metode este *algoritmul Davidson-Fletcher-Powell*. Caracteristic acestui algoritm este faptul că modificarea matricei $\tilde{H}^{(k)}$ la fiecare iterație este făcută astfel încât pentru o funcție pătratică de n variabile, în limita a n iterații, matricea $\tilde{H}^{(k)}$ să devină egală cu matricea $H^{(k)-1}$. Inițial, matricea \tilde{H} se alege ca fiind matricea unitate I_n , astfel încât prima iterație decurge după metoda gradientului. În continuare se face o modificare treptată de la direcția gradientului

la cea a metodei Newton, reușindu-se astfel combinarea avantajelor aferente celor două metode. De fapt într-o regiune îndepărtată de optim metoda gradientului este mai rapidă, în timp ce în apropierea optimului, metoda Newton este superioară.

Convergența metodei Davidon-Fletcher-Powell a fost demonstrată doar pentru funcții obiectiv pătratice, având matricea hessiană pozitiv definită. Relația de recurență care stă la baza metodei este următoarea:

$$\begin{aligned} \tilde{H}^{(k+1)} = \tilde{H}^{(k)} + A^{(k)} - B^{(k)} = \tilde{H}^{(k)} + \frac{[\Delta X^{(k)}][\Delta X^{(k)}]^t}{(\Delta X^{(k)}, \Delta g^{(k)})} - \\ - \frac{[\tilde{H}^{(k)}][\Delta g^{(k)}][\Delta g^{(k)}]^t [\tilde{H}^{(k)}]^t}{(\Delta g^{(k)}, H^{(k)} \Delta g^{(k)})} \end{aligned} \quad (10.6)$$

unde:

$$\Delta X^{(k)} = X^{(k+1)} - X^{(k)}, \quad \Delta g^{(k)} = g^{(k+1)} - g^{(k)} \quad (10.7)$$

10.3. Exemple numerice

Exemplul 1

Să se determine minimumul funcției:

$$F(X) = \min(x_1^2 + 4 \cdot x_2^2 - 4)$$

folosind metoda Newton, indicându-se ca punct de pornire $X^{(0)} = [5 \quad 4]^t$, iar $\lambda^{(0)} = 1$.

Se calculează gradientul și hesianul funcției $F(X)$:

$$g = \left[\frac{\partial F(X)}{\partial x_1} \quad \frac{\partial F(X)}{\partial x_2} \right]^t = [2x_1 \quad 8x_2]^t$$

$$H = C = \begin{bmatrix} \frac{\partial^2 F(X)}{\partial x_1^2} & \frac{\partial^2 F(X)}{\partial x_1 \partial x_2} \\ \frac{\partial^2 F(X)}{\partial x_2 \partial x_1} & \frac{\partial^2 F(X)}{\partial x_2^2} \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 8 \end{bmatrix}$$

Aplicând relația (10.4) rezultă:

$$X^{(1)} = X^{(0)} - H^{(0)^{-1}} \cdot g^{(0)} = \begin{bmatrix} 5 \\ 4 \end{bmatrix} - \begin{bmatrix} 2 & 0 \\ 0 & 8 \end{bmatrix}^{-1} \begin{bmatrix} 10 \\ 32 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Exemplul 2

1. Să se determine minimumul funcției:

$$F(X) = \min(4 \cdot (x_1 - 2)^2 + (x_2 - 3)^2)$$

folosind metoda Davidon-Fletcher-Powell, indicându-se ca punct de pornire $X^{(0)} = [6 \ 5]^t$.

Se calculează gradientul funcției obiectiv:

$$g^{(k)} = \begin{bmatrix} 8(x_1 - 2) \\ 2(x_2 - 3) \end{bmatrix}_{X=X^{(k)}}, \quad g^{(0)} = \begin{bmatrix} 32 \\ 4 \end{bmatrix}$$

Iterația 1

Se consideră $\tilde{H}^{(0)} = I_n$, astfel încât relația iterativă are următoarea formă:

$$X^{(1)} = \begin{bmatrix} x_1^{(1)} \\ x_2^{(1)} \end{bmatrix} = \begin{bmatrix} 6 \\ 5 \end{bmatrix} - \lambda^{(0)} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 32 \\ 4 \end{bmatrix} = \begin{bmatrix} 6 - 32\lambda^{(0)} \\ 5 - 4\lambda^{(0)} \end{bmatrix}$$

Pentru determinarea lungimii pasului de deplasare $\lambda^{(0)}$, se introduc valorile $x_1^{(1)}$ și $x_2^{(1)}$ în expresia funcției obiectiv F , se face derivata în raport cu $\lambda^{(0)}$, după care se anulează:

$$\frac{\partial F(x^{(1)})}{\partial \lambda^{(0)}} = 65 - 514\lambda^{(0)} = 0 \rightarrow \lambda^{(0)} = 0,1264$$

Prin urmare, valoarea noii aproximații este:

$$X^{(1)} = \begin{bmatrix} 6 \\ 5 \end{bmatrix} - 0,1264 \cdot \begin{bmatrix} 32 \\ 4 \end{bmatrix} = \begin{bmatrix} 1,955 \\ 4,494 \end{bmatrix}.$$

Iterația 2

Se calculează ΔX și Δg :

$$g^{(1)} = \begin{bmatrix} -0,36 \\ 2,988 \end{bmatrix}, \Delta X^{(1)} = \begin{bmatrix} -4,045 \\ -0,506 \end{bmatrix} \text{ și } \Delta g^{(1)} = \begin{bmatrix} -32,36 \\ -1,012 \end{bmatrix}$$

Folosind relația (10.6) se calculează matricea $\tilde{H}^{(1)}$:

$$H^{(1)} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \frac{\begin{bmatrix} -4,045 & 0 \\ -0,506 & 0 \end{bmatrix} \cdot \begin{bmatrix} -4,045 & -0,506 \\ 0 & 0 \end{bmatrix}}{\begin{bmatrix} -4,045 & -0,506 \end{bmatrix} \cdot \begin{bmatrix} -32,36 \\ -1,012 \end{bmatrix}} +$$

$$+ \frac{\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} -32,36 & 0 \\ -1,012 & 0 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}}{\begin{bmatrix} -32,36 & -1,012 \end{bmatrix} \cdot \begin{bmatrix} -32,36 \\ -1,012 \end{bmatrix}} = \begin{bmatrix} 0,1554 & -0,0147 \\ -0,0147 & 1,001 \end{bmatrix}$$

În continuare se poate calcula noua aproximație $X^{(2)}$:

$$\bullet \quad X^{(2)} = \begin{bmatrix} X_1^{(2)} \\ X_2^{(2)} \end{bmatrix} = \begin{bmatrix} 1,955 \\ 4,494 \end{bmatrix} - \lambda^{(1)} \begin{bmatrix} 0,1554 & -0,0147 \\ -0,0147 & 1,0010 \end{bmatrix} \cdot \begin{bmatrix} -0,36 \\ 4,494 \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$$

unde $\lambda^{(1)} = 0,4984$ a fost determinat în mod similar ca $\lambda^{(0)}$.

Deoarece $g^{(2)} = [0 \ 0]^t$, calculele sunt terminate, punctul $X^{(2)} = [2 \ 3]^t$ fiind soluția optimală căutată.

10.4. Desfășurarea lucrării

1. Se studiază textul lucrării.
2. Să se determine minimul funcției:

$$F(X) = \min(x_1^2 + 2 \cdot x_2^2 - 1)$$

folosind metoda Newton, indicându-se ca punct de pornire $X^{(0)} = [3 \ 1]^t$.

Rezultatele obținute vor fi comparate cu cele obținute folosind funcția Matlab **fminunc**.

3. Să se determine minimul funcției:

$$F(X) = \min(x_1^2 + 3 \cdot x_2^2 - 2)$$

folosind metoda cvasi Newton, indicându-se ca punct de pornire $X^{(0)} = [2 \ 1]^t$.
Rezultatele obținute vor fi comparate cu cele obținute folosind funcția Matlab
fminunc.