LUCRAREA 9

PROGRAMAREA NELINIARĂ ÎN REZOLVAREA PROBLEMELOR DIN ENERGETICĂ. METODE DE ORDINUL 1

9.1. Aspecte generale

Metodele de ordinul unu apelează la calculul derivatelor de ordinul unu pentru determinarea direcţiei şi a pasului de deplasare. Acestea realizează un compromis între simplitate şi eficienţă, fiind cele mai utilizate în practică.

Se poate reaminti faptul că, pentru o funcţie obiectiv continuă şi derivabilă, gradientul într-un punct curent $X^{(k)}$ reprezintă vectorul derivatelor parţiale de ordinul unu în punctul respectiv. Acest vector este ortogonal la conturul lui $F(X)$ ce trece prin punctul $X^{(k)}$. Direcţia lui corespunde celei mai rapide creşteri a lui $F(X)$, ceea ce permite utilizarea ei pentru maximizare sau a direcţiei opuse, pentru minimizare.

9.2. Metoda gradientului simplu

Metoda gradientului simplu este cea mai simplă metodă de ordinul 1 pentru minimizarea unei funcții de mai multe variabile.

Metoda are la bază relaţia de recurenţă:

$$X^{(k+1)} = X^{(k)} + \lambda^{(k)} d^{(k)}\ , \quad k = 1,\ 2,\ 3, \ldots \tag{9.1}$$

unde direcţia $d^{(k)} = -\nabla F(X^{(k)})$.

Sub forma gradientului simplu, metoda s-a dovedit ineficientă din cauza deplasării în zig-zag. În schimb, metodele care au la bază gradienţii conjugaţi ne conduc la rezultate foarte bune, fiind cele mai utilizate în practică.

9.3. Metoda gradienților conjugați

În cadrul metodelor de gradienţi conjugaţi noua direcţie de deplasare poate fi determinată dacă se utilizează o combinaţie între gradientul calculat în punctul $X^{(k)}$ şi direcţia precedentă,

$$d^{(k)} = -g^{(k)} + \beta^{(k)} d^{(k-1)}, k = 1,2,... \tag{9.2}$$

$$d^{(0)} = -g^{(0)} \tag{9.3}$$

Există mai multe variante ale metodelor de gradienţi conjugaţi, acestea deosebindu-se prin modul de determinare a parametrului scalar $\beta^{(k)}$:

- metoda Fletcher-Reeves:

$$\beta^{(k)} = \frac{\left(g^{(k)}, g^{(k)}\right)}{\left(g^{(k-1)}, g^{(k-1)}\right)} \tag{9.4}$$

- metoda Polak-Ribière:

$$\beta^{(k)} = \frac{\left(g^{(k)}, g^{(k)} - g^{(k-1)}\right)}{\left(g^{(k-1)}, g^{(k-1)}\right)} \tag{9.5}$$

- metoda Hestenes-Stiefel:

$$\beta^{(k)} = \frac{\left(g^{(k)}, g^{(k)} - g^{(k-1)}\right)}{\left(g^{(k)} - g^{(k-1)}, d^{(k-1)}\right)} \tag{9.6}$$

Toate cele trei metode, calculează minimul funcţiei $F(X)$ în cel mult n iteraţii, convergenţa lor fiind bună.

9.4. Exemplu numeric

Să se determine minimul funcţiei:

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

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

Conform metodei valoarea optimă se va atinge în cel mult 2 iteraţii.

Se calculează gradientul şi hesianul funcţiei $F(X)$:

$$g = \begin{bmatrix} \dfrac{\partial F(X)}{\partial x_1} & \dfrac{\partial F(X)}{\partial x_2} \end{bmatrix}_t = [2x_1 \quad 8x_2]^t$$

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

Iteraţia 1

Se determină direcţia $d^{(0)} = -g^{(0)}$, adică:

$$d^{(0)} = -[2 \cdot 5 \quad 8 \cdot 4]^t = -[10 \quad 32]^t$$

Pentru determinarea valorii pasului de deplasare $\lambda^{(1)}$ se va utiliza relaţia (10.52):

$$\lambda^{(0)} = -\frac{[10 \quad 32] \cdot [-10 \quad -32]^t}{[-10 \quad -32]\begin{bmatrix} 2 & 0 \\ 0 & 8 \end{bmatrix}[-10 \quad -32]^t} = 0{,}1339$$

În continuare pentru determinarea noii aproximaţii se va utiliza relaţia iterativă:

$$X^{(1)} = X^{(0)} + \lambda^{(0)} d^{(0)} = [5 \quad 4]^t + 0{,}1339 \cdot [-10 \quad -32]^t = [3{,}6606 \quad -0{,}2860]^t$$

Iteraţia 2

Direcţia de deplasare $d^{(1)}$ se determină cu ajutorul relaţiei (9.2):

$$d^{(1)} = -g^{(1)} + \beta^{(1)} \cdot d^{(0)}$$

Coeficientul $\beta^{(1)}$ se calculează cu relaţia (9.4):

$$\beta^{(1)} = \frac{[7{,}3213 \quad -2{,}2879\,]\cdot[7{,}3213 \quad -2{,}2879\,]^{t}}{[10 \quad 32]\cdot[-10 \quad -32]^{t}} = 0{,}0523$$

$$d^{(1)} = -[7{,}3213 \quad -2{,}2879\,]^{t} + 0{,}0523\cdot[-10 \quad -32\,]^{t} = [-7{,}8447 \quad 0{,}6129\,]^{t}$$

Valoarea noului pas de deplasare $\lambda^{(1)}$ este:

$$\lambda^{(1)} = -\frac{[7{,}3213 \quad -2{,}2879\,][-7{,}8447 \quad 0{,}6129\,]^{t}}{[-7{,}8447 \quad 0{,}6129\,]\begin{bmatrix}2 & 0\\ 0 & 8\end{bmatrix}[-7{,}8447 \quad 0{,}6129\,]^{t}} = 0{,}4666$$

Valorile pasului de deplasare şi a direcţiei se vor introduce în relaţia iterativă obţinându-se:

$$X^{(2)} = \begin{bmatrix}3{,}6606\\ -0{,}286\end{bmatrix} + 0{,}4666\cdot\begin{bmatrix}-78447\\ 0{,}6129\end{bmatrix} = \begin{bmatrix}0\\ 0\end{bmatrix}.$$

9.5. Desfăşurarea lucrării

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

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

   folosind metoda gradientului simplu, indicându-se ca punct de pornire $X^{(0)} = [3 \quad 1]^{t}$. Rezultatele obținute vor fi comparate cu cele obținute prin folosirea funcției Matlab **fminunc**.
3. Să se determine minimul funcţiei:

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

   folosind metoda gradientului conjugat, indicându-se ca punct de pornire $X^{(0)} = [2 \quad 1]^{t}$. Rezultatele obținute vor fi comparate cu cele obținute prin folosirea funcției Matlab **fminunc**.