

BFGS

Nikola Ajzenhamer

Matematički fakultet,
Univerzitet u Beogradu

19. maj 2018.

Sadržaj

Uvod

BFGS metod

Algoritmi

- ▶ Broyden-Fletcher-Goldfarb-Shanno
- ▶ Jedan od Kvazi-Njutnovih metoda:
 - ▶ Zahtevaju gradijent ciljne funkcije pri svakoj iteraciji
 - ▶ Konstruišu model ciljne funkcije koji je dovoljno dobar da daje superlinearnu konvergenciju
 - ▶ S obzirom da ne zahtevaju izvode drugog reda, ponekad su brži od Njutnovih metoda

BFGS metod

- ▶ Model ciljne funkcije:

$$m_k(p) = f_k + \nabla f_k^T p + \frac{1}{2} p^T B_k p \quad (1)$$

- ▶ B_k je:
 - ▶ Dimetrična pozitivno definitna matrica dimenzija $n \times n$
 - ▶ Ažurira se pri svakoj iteraciji
 - ▶ Predstavlja aproksimaciju Hesijana

BFGS metod (2)

- ▶ Minimizator modela 1:

$$p_k = -B_k^{-1} \nabla f_k \quad (2)$$

- ▶ Koristi se kao pravac pretrage:

$$x_{k+1} = x_k + \alpha_k p_k$$

- ▶ Kasnije ćemo razmotriti određivanje koraka α_k

BFGS metod (3)

- ▶ Može se izvesti naredna formula koja se koristi za izračunavanje naredne aproksimacije B_{k+1}^{-1} ukoliko je poznata prethodna aproksimacija B_k^{-1} :

$$B_{k+1}^{-1} = (I - \rho_k s_k y_k^T) B_k^{-1} (I - \rho_k y_k s_k^T) + \rho_k s_k s_k^T \quad (3)$$

gde važi

$$s_k = x_{k+1} - x_k, \quad y_k = \nabla f_{k+1} - \nabla f_k, \quad \rho_k = \frac{1}{y_k^T s_k}$$

- ▶ Za početnu aproksimaciju može se uzeti $B_0^{-1} = I$

Određivanje koraka α_k

- ▶ Određivanje α_k se vrši tako da zadovoljava Wolfe uslove:

$$\begin{aligned} f(x_k + \alpha_k p_k) &\leq f(x_k) + c_1 \alpha_k \nabla f_k^T p_k \\ \nabla f(x_k + \alpha_k p_k)^T p_k &\geq c_2 \nabla f_k^T p_k \end{aligned} \quad (4)$$

gde važi $0 < c_1 < c_2 < 1$.

- ▶ Prvi uslov govori da α_k treba da daje *neophodan spust* ciljne funkcije f (*Armijo condition*)
- ▶ Drugi uslov eliminiše premale korake (*curvature condition*)

Određivanje koraka α_k

- ▶ Algoritam linijske pretrage za određivanje koraka vrši se u dve faze:
 1. Faza odsecanja – pronalazi se interval koji sadrži prihvatljive vrednosti
 2. Faza selekcije – pronalazi se finalni korak (algoritam Zoom)

Algoritam Linijska_pretraga

- ▶ Ulaz:
 - ▶ Funkcija $\phi(\alpha) = f(x_k + \alpha p_k)$
- ▶ Izlaz:
 - ▶ Parametar α^* koji zadovoljava Wolfe uslove 4

Algoritam Linijska_pretraga

```
Postavi  $\alpha_0 \leftarrow 0$ , odaberi  $\alpha_1 > 0$  i  $\alpha_{max}$ ;  
 $i \leftarrow 1$ ;  
repeat  
  Izračunaj  $\phi(\alpha_i)$ ;  
  if  $\phi(\alpha_i) > \phi(0) + c_1\alpha_i\phi'(\alpha_0)$  or  $[\phi(\alpha_i) \geq \phi(\alpha_{i-1})$  and  $i > 1]$   
    Postavi  $\alpha^* \leftarrow \text{Zoom}(\alpha_{i-1}, \alpha_i)$ ; Zaustavi;  
  Izračunaj  $\phi'(\alpha_i)$ ;  
  if  $|\phi'(\alpha_i)| \leq -c_2\phi'(0)$   
    Postavi  $\alpha^* \leftarrow \alpha_i$ ; Zaustavi;  
  if  $\phi'(\alpha_i) \geq 0$   
    Postavi  $\alpha^* \leftarrow \text{Zoom}(\alpha_i, \alpha_{i-1})$ ; Zaustavi;  
  Izaberi  $\alpha_{i+1} \in (\alpha_i, \alpha_{max})$ ;  
   $i \leftarrow i + 1$ ;  
end (repeat)
```

Algoritam Zoom

- ▶ Ulaz:
 - ▶ Funkcija $\phi(\alpha) = f(x_k + \alpha p_k)$
 - ▶ Vrednosti α_{lo} i α_{hi}
- ▶ Izlaz:
 - ▶ Finalni korak α^* iz intervala $[\alpha_{lo}, \alpha_{hi}]$

Algoritam Zoom

```
repeat
  Pronađi probni korak  $\alpha_j$  između  $\alpha_{lo}$  i  $\alpha_{hi}$ ;
  Izračunaj  $\phi(\alpha_j)$ ;
  if  $\phi(\alpha_j) > \phi(0) + c_1\alpha_j\phi'(\alpha_0)$  or  $\phi(\alpha_j) \geq \phi(\alpha_{lo})$ 
    Postavi  $\alpha_{hi} \leftarrow \alpha_j$ ;
  else
    Izračunaj  $\phi'(\alpha_j)$ ;
    if  $|\phi'(\alpha_j)| \leq -c_2\phi'(0)$ 
      Postavi  $\alpha^* \leftarrow \alpha_j$ ; Zaustavi;
    if  $\phi'(\alpha_j)(\alpha_{hi} - \alpha_{lo}) \geq 0$ 
      Postavi  $\alpha_{hi} \leftarrow \alpha_{lo}$ ;
    Postavi  $\alpha_{lo} \leftarrow \alpha_j$ ;
end (repeat)
```

Algoritam BFGS

- ▶ Ulaz:
 - ▶ Ciljna funkcija f
 - ▶ Gradijent funkcije ∇f
 - ▶ Početna tačka x_0
 - ▶ Parametar konvergencije ϵ
 - ▶ Inverz aproksimacije Hesijana B_0^{-1}
- ▶ Izlaz:
 - ▶ Aproksimacija minimuma x^*

Algoritam BFGS

```
 $k \leftarrow 0;$   
while  $\|\nabla f_k\| > \epsilon$   
    Izračunaj pravac pretrage  $p_k$  prema formuli 2;  
    Postavi  $\alpha_k \leftarrow \text{Linijska\_pretraga}(f, x_k, p_k);$   
    Postavi  $x_{k+1} \leftarrow x_k + \alpha_k p_k;$   
    Postavi  $s_k \leftarrow x_{k+1} - x_k;$   
    Postavi  $y_k \leftarrow \nabla f_{k+1} - \nabla f_k;$   
    Izračunaj  $B_{k+1}^{-1}$  prema formuli 3;  
     $k \leftarrow k + 1;$   
end (repeat)
```

Ocena složenosti

- ▶ Svaka iteracija se izvodi po cenu od $O(n^2)$ uz cenu izračunavanja vrednosti funkcija i gradijenata
- ▶ S obzirom da nema rešavanja sistema linearnih jednačina i matrica-matrica operacija, ne postoje ni operacije složenosti $O(n^3)$

Hvala na pažnji!

Pitanja?