1

7.04.2021

OPTIMIZASYON TEKNIKLERİ

Eşlenik yön yöntemleri $x^* \in \mathbb{R}^n$ f'ye bir minimum çözüm sağlıyorsa $x_0 \in \mathbb{R}^n$ başlangıç noktasını seçerek

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

algoritması $\alpha_k \in \mathbb{R}$ ve $d_k \in \mathbb{R}^n$ olmak üzere x^* optimuma ulaştırır.

$$x_1 = x_0 + \alpha_0 d_0$$

$$x_2 = x_1 + \alpha_1 d_1$$

$$= x_0 + \alpha_0 d_0 + \alpha_1 d_1$$

$$x_3 = x_2 + \alpha_2 d_2$$

$$= x_0 + \alpha_0 d_0 + \alpha_1 d_1 + \alpha_2 d_2$$

$$\vdots$$

$$x_k = x_0 + \alpha_0 d_0 + \alpha_1 d_1 + \dots + \alpha_{k-1} d_{k-1}$$ olur.

algoritma uygun $\{\alpha_i\}$ ve $\{d_i\}$ dizilerinin bulunmasıyla sürdürülür.

Tanım: Q $n \times n$ boyutlu pozitif definit simetrik matris $d_i, d_j \in \mathbb{R}^n$ olmak üzere

$$d_i' Q d_j = 0$$

ise d_i ve d_j vektörlerine Q-ortogonal vektörler denir.

Algoritmayı kuadratik fonksiyonlar için türetelim.

$$f(x) = \frac{1}{2} x' Q x - x' b$$

$$\nabla f(x) = Q x - b \quad \nabla f(x_k) = Q x_k - b = g_k$$

$$\nabla f(x^*) = Q x^* - b = 0 \Rightarrow Q(x^*) = b \text{ dir.}$$

$\boxed{2}$ Teorem: $\{d_0, d_1, \ldots, d_{k-1} \ldots d_{n-1}\}$ Q-ortogonal bir küme ise $d_0, d_1, \ldots, d_{n-1}$ vektörleri ~~ara~~ aralarında lineer bağımsızdır.

İsp. $\alpha_0 d_0 + \alpha_1 d_1 + \ldots + \alpha_k d_k + \ldots + \alpha_{n-1} d_{n-1} = 0$ eşitliği $\forall \alpha_i = 0$ için gerçekleniyorsa lineer bağımsızdır. Bu eşitlik ~~sağdan~~ soldan önce Q sonra d_k ile çarpılırsa

$$\underbrace{\alpha_0 d_k' Q d_0}_{=0} + \underbrace{\alpha_1 d_k' Q d_1}_{=0} + \ldots + \alpha_k \underbrace{d_k' Q d_k}_{\text{pozitif}} + \ldots + \underbrace{\alpha_{n-1} d_k' Q d_{n-1}}_{=0} = 0$$

Q pozitif definit olduğundan.

$$\alpha_k \underbrace{d_k' Q d_k}_{\neq 0} = 0 \qquad \alpha_k = 0 \text{ olmak zorunda.}$$

$\forall k = 0, \ldots, n-1$ için geçerli

Bütün katsayıları sıfır olur. $\Rightarrow \{d_0, d_1, \ldots, d_{n-1}\}$ kümesi lineer bağımsızdır.

n tane lineer bağımsız vektör n boyutlu uzayda bir taban yada baz oluşturur. Böylece tüm vektörler bu vektörler cinsinden ifade edilir. x^* ve x_0 $\mathbb{R}^n$'nin elemanıydı öyle ise farklarıda $\mathbb{R}^n$ nin elemanı olur. Bu durumda

$x^* - x_0$ farkı taban vektörler cinsinden yazılır.

$$x^* - x_0 = \alpha_0 d_0 + \alpha_1 d_1 + \ldots + \alpha_{n-1} d_{n-1}$$

α_i ve d_i lerin bulunmasıyla algoritmanın optimali bulması garanti ediliyor.

$$d_k' Q (x^* - x_0) = \alpha_k \underbrace{d_k' Q d_k}_{>0}$$

[3]

$$\alpha_k = \frac{d_k' Q (x^* - x_0)}{d_k' Q d_k}$$

$$d_k' Q (x^* - x_0) = d_k' Q (x^* - x_k + x_k - x_0) = d_k' Q (x^* - x_k) + d_k' Q (x_k - x_0)$$

$$x_k - x_0 = \alpha_0 d_0 + \alpha_1 d_1 + \dots + \alpha_{k-1} d_{k-1}$$

$$d_k' Q (x_k - x_0) = 0$$

$$d_k' Q (x^* - x_0) = d_k' Q (x^* - x_k)$$

Bu da başlangıç noktasına bağlı olmadan algoritmanın optimal x^* değerini içerdiğinin ispatıdır.

$$d_k' Q (x^* - x_k) = d_k' [\underbrace{\underbrace{Q(x^*)}_{b} - Q(x_k)}_{-g_k}] = -d_k' g_k$$

$$\alpha_k = \frac{-d_k' g_k}{d_k' Q d_k}$$

Demekki α'lar d'lere bağlı olacak şekilde belirleniyor.

$d_{k+1} = -g_{k+1} + \beta_k d_k$ ötelemesi yap.

Fletcher - Reeves yöntemi

Adım 1 x_0 seç $g_0 = \nabla f(x_0)$ hesapla $d_0 = -g_0$ al

Adım 2 a) α_k'yı $f(x_k + \alpha d_k)$'yı α'ya göre minimize ederek bul ve $x_{k+1} = x_k + \alpha_k d_k$ ötelemesi yap.

b) $g_{k+1} = \nabla f(x_{k+1})$ vektörünü hesapla

c) $k<n-1$ ise Adım 3'e git Değilse

$$\beta_k = \frac{g'_{k+1} g_{k+1}}{g_k' g_k}$$ olmak üzere $d_{k+1} = -g_{k+1} + \beta_k d_k$

ötelemesi yap ve a'ya dön

Adım 3 x_0 yerine x_n al ve adım1'e dön

Algoritma $g_k = 0$ ile duracaktır.

Örnek

$$\min F = x^2 + \frac{y^2}{2} - x - y$$ problemini Fletcher reeves

yöntemiyle çözelim. $x_0 = [0,0]$ başlangıç ~~yer~~ noktası

olsun.

$$\nabla f(x) = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix} = \begin{bmatrix} 2x-1 \\ y-1 \end{bmatrix}$$ $x_0 = [0,0]$ olduğundan

$$\nabla f(x_0) = \begin{bmatrix} -1 \\ -1 \end{bmatrix} \qquad g_0 = \begin{bmatrix} -1 \\ -1 \end{bmatrix}$$ $g_0 = -d_0$ olduğundan

$$d_0 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \qquad f(x_0 + \alpha d_0) = f\left[[0,0] + \alpha d_0^T\right] = f\left[[0,0] + \alpha[1,1]\right]$$

$$= f[\alpha, \alpha] \qquad \begin{matrix} x = \alpha \\ y = \alpha \end{matrix}$$

$$f(\alpha,\alpha) = \alpha^2 + \frac{\alpha^2}{2} - \alpha - \alpha$$

$$= \frac{3}{2}\alpha^2 - 2\alpha$$

$$\frac{\partial f}{\partial \alpha} = 3\alpha - 2 = 0 \qquad \alpha = 2/3$$ ilk değer olduğundan

$\alpha_0 = 2/3$ diyebiliriz

$$\frac{\partial^2 f}{\partial \alpha^2} = 3 > 0$$ bulunan α değeri minimumdur.

$\boxed{5}$

$$x_1 = x_0 + \alpha_0 d_0 = [0, 0] + \alpha_0 [1, 1] = [\alpha_0, \alpha_0] = [2/3, 2/3]$$

$$\nabla f(x_1) = \begin{bmatrix} 2 \cdot \frac{2}{3} - 1 \\ \frac{2}{3} - 1 \end{bmatrix} = \begin{bmatrix} 1/3 \\ -1/3 \end{bmatrix} = g_1$$

$$\beta_0 = \frac{g_1' \cdot g_1}{g_0' g_0} = \frac{[1/3, -1/3] \begin{bmatrix} 1/3 \\ -1/3 \end{bmatrix}}{[-1, -1] \begin{bmatrix} -1 \\ -1 \end{bmatrix}} = \frac{2/9}{2} = 1/9$$

$$d_1 = -g_1 + \beta_0 d_0 = -\begin{bmatrix} 1/3 \\ -1/3 \end{bmatrix} + \frac{1}{9} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -1/3 \\ 1/3 \end{bmatrix} + \begin{bmatrix} 1/9 \\ 1/9 \end{bmatrix}$$

$$= \begin{bmatrix} -2/9 \\ 4/9 \end{bmatrix}$$

$$f(x_1 + \alpha, d_1^T) = f\left[[2/3, 2/3] + \alpha \left[-\frac{2}{9}, 4/9\right]\right] = f\left[2/3 - \frac{2\alpha}{9}, 2/3 + \frac{4\alpha}{9}\right]$$

$$f(\alpha) = \left(\frac{2}{3} - \frac{2\alpha}{9}\right)^2 + \frac{\left(2/3 + \frac{4\alpha}{9}\right)^2}{2} - \left(\frac{2}{3} - \frac{2\alpha}{9}\right) - \left(\frac{2}{3} + \frac{4\alpha}{9}\right)$$

$$\frac{\partial \alpha}{\partial \alpha} = 2\left(\frac{2}{3} - \frac{2\alpha}{9}\right)\left(\frac{-2}{9}\right) + \cancel{2}\frac{\left(\frac{2}{3} + \frac{4\alpha}{9}\right)}{\cancel{2}} \cdot \frac{4}{9} + \frac{2}{9} - \frac{4}{9}$$

$$= \cancel{\frac{-8}{27}} + \frac{8\alpha}{81} + \cancel{\frac{8}{27}} + \frac{16\alpha}{81} + \frac{2}{9} - \frac{4}{9}$$

$$0 = \frac{24\alpha}{81} - \frac{2}{9}$$

$$\frac{24\alpha}{81} = \frac{2}{9} \Rightarrow \frac{12\alpha}{9} = 1 \qquad \alpha = \frac{9}{12} = \frac{3}{4}$$

$\boxed{6}$

$$x_2 = x_1 + \alpha_1 d_1^T = [2/3, 2/3] + \frac{3}{4}\left[\frac{-2}{9}, \frac{4}{9}\right]$$

$$= [2/3, 2/3] + \left[\frac{-6}{36}, \frac{12}{36}\right]$$

$$= [2/3, 2/3] + \left[\frac{-1}{6}, \frac{1}{3}\right]$$

$$= \left[\frac{2}{3} - \frac{1}{6}, \frac{2}{3} + \frac{1}{3}\right]$$

$$x_2 = [1/2, 1]$$

$$\nabla f(x_2) = \begin{bmatrix} 2x-1 \\ y-1 \end{bmatrix} = \begin{bmatrix} 2.\frac{1}{2} - 1 \\ 1-1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$x_2 = [1/2, 1]$ optimumdur.

$f: \mathbb{R}^2 \to \mathbb{R} = x_1^3 - x_1^2 x_2 + 2x_2^2$

fonksiyonlarının optimumlarını belirleyin.

$$\nabla f(x) = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \end{bmatrix} = \begin{bmatrix} 3x_1^2 - 2x_1x_2 \\ -x_1^2 + 4x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$3x_1^2 - 2x_1x_2 = 0$ $\quad -x_1^2 + 4x_2 = 0$

$x_1(3x_1 - 2x_2) = 0$

$\downarrow$ 0 $\quad x_1 = 0 \Rightarrow x_2 = 0$ olur. $A(0,0)$

$x_2 = \frac{3x_1}{2}$ alınırsa

$-x_1^2 + \frac{12}{2}x_1 = 0 \quad -x_1(x_1 - 6) = 0 \quad x_1 = 6$

$\boxed{x_2 = 9}$ $\quad B(6,9)$

$\boxed{7}$

(6, 9) noktasını inceleyelim

$x_1 = 6 \quad f(6, x_2) = 6^3 - 36x_2 + 2x_2^2$ x_2'ye göre 1 Tek değişkenli hale geldi

$$\frac{\partial f}{\partial x_2} = -36 + 4x_2 = 0 \quad \boxed{x_2 = 9}$$

$$\frac{\partial^2 f}{\partial x_2^2} = 4 > 0$$ (6,9) Min noktası olarak düşünülebilir mi?

yada

$x_2 = 9$'u fonksiyonda yerine koyup fonksiyonu x_1'e bağlı tek fonksiyon gibi düşünelim.

$$f(x_1, 9) = x_1^3 - 9x_1^2 + 162$$

$$\frac{\partial f}{\partial x_1} = f'(x_1) = 3x_1^2 - 18x_1 = 0 \quad 3x_1(x_1 - 6) = 0 \quad x_1 = 6$$

$$\frac{\partial^2 f}{\partial x_1^2} = 6x_1 - 18 \Big|_{x_1 = 6} = 18 > 0$$

yine (6, 9) Min noktası olarak düşünülebilir mi?

ikinci mertebe koşullar.

$$\nabla^2 f = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} \end{bmatrix} = \begin{bmatrix} 6x_1 - 2x_2 & -2x_1 \\ -2x_1 & 4 \end{bmatrix}$$

$$\underset{(0,0)}{\nabla^2 f(A)} = \begin{bmatrix} 0 & 0 \\ 0 & 4 \end{bmatrix}$$

$\boxed{8}$

$$h' \nabla^2 f(A) h = [h_1, h_2] \begin{bmatrix} 0 & 0 \\ 0 & 4 \end{bmatrix} \begin{bmatrix} h_1 \\ h_2 \end{bmatrix} = [0, 4h_2] \begin{bmatrix} h_1 \\ h_2 \end{bmatrix}$$

$= 4h_2^2 > 0$ pozitif definit olduğundan

$A(0,0)$ noktası minimum noktadır.

$$\nabla^2 f(B) = \begin{bmatrix} 18 & -12 \\ -12 & 4 \end{bmatrix}$$
$(6,9)$

$\Delta_1 = 18 > 0$

$\Delta_2 = 72 - 144 = -72$

definit değil

semer noktasıdır.

Problem: $f(x_1, x_2, x_3) = 2x_1x_2x_3 - 4x_1x_3 - 2x_2x_3 + x_1^2 + x_2^2 + x_3^2$
$- 2x_1 - 4x_2 + 4x_3$

fonksiyonun optimum çözümlerini belirleyin.

$A(0, 3, 1)$

$B(0, 1, -1)$

$C(1, 2, 0)$ * min. çöz.

$D(2, 1, 1)$

$E(2, 3, -1)$

$$\nabla^2 f = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \frac{\partial^2 f}{\partial x_1 \partial x_3} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \frac{\partial^2 f}{\partial x_2 \partial x_3} \\ \frac{\partial^2 f}{\partial x_3 \partial x_1} & \frac{\partial^2 f}{\partial x_3 \partial x_2} & \frac{\partial^2 f}{\partial x_3^2} \end{bmatrix}$$

bakılır.