1

OPTIMIZASYON TEKNIKLERI 31/mart/2021

kuadratik fonksiyonlarda En hızlı düşüş algoritması

$Q \in \mathbb{R}^{n \times n}$ pozitif definit bir matris $x \in \mathbb{R}^n$ olmak üzere

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

fonksiyonuna kuadratik fonksiyon denir.

1 mertebe koşul $\nabla f(x) = Qx - b$, x^*, $\nabla f(x)$'i 0 yapan değişken

$$E(x) = \frac{1}{2} x' Q x - x' b + \frac{1}{2} x^{*\prime} Q x^*$$

$b = Q x^*$ yerine konursa

$$E(x) = \frac{1}{2} x' Q x - x' Q x^* + \frac{1}{2} x^{*\prime} Q x^*$$

$$= \frac{1}{2} x' Q x - \frac{1}{2} x' Q x^* - \frac{1}{2} x' Q x^* + \frac{1}{2} x^{*\prime} Q x^*$$

$$= \frac{1}{2} x' Q (x - x^*) - \frac{1}{2} (x - x^*)' Q x^*$$

$$= \frac{1}{2} x' Q (x - x^*) - \frac{1}{2} Q x^{*\prime} (x - x^*)$$

$$= \frac{1}{2} (x' Q (x - x^*))' - \frac{1}{2} Q x^{*\prime} (x - x^*)$$

$$= \frac{1}{2} (x - x^*)' Q x - \frac{1}{2} (x - x^*)' Q x^*$$

$$= \frac{1}{2} (x - x^*)' Q (x - x^*)$$ bulunur.

$\boxed{2}$ x^* aynı zamanda $E(x)$'in de bir köküdür.

α_i adım boyunu belirleyelim.

$x_{i+1} = x_i - \alpha_i \nabla f(x_i)$ $\nabla f(x_i) = g_i$ alalım.

Böylece en hızlı düşüş algoritması

$$x_{i+1} = x_i - \alpha_i g_i$$

$$x_{i+1} = x_i - \frac{g_i' g_i}{g_i' Q g_i} g_i$$

<u>İsp.</u>

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

koyalım.

$$f(x_i - \alpha g_i) = \frac{1}{2}(x_i - \alpha g_i)' Q (x_i - \alpha g_i) - (x_i - \alpha g_i)' b$$

α'ya göre türev alalım.

$$\frac{df}{d\alpha} = -g_i' Q (x_i - \alpha g_i) + g_i' b$$

$$= -g_i' Q x_i + (g_i' Q g_i)\alpha + g_i' b$$

$$= g_i' \underbrace{(b - Q x_i)}_{-g_i} + (g_i' Q g_i)\alpha$$

$$0 = -g_i' g_i + (g_i' Q g_i)\alpha$$

$$\boxed{\alpha_i = \frac{g_i' g_i}{g_i' Q g_i}}$$

$\boxed{3}$ $f(x,y) = 2x^2 + 2y^2 + 2xy - 4x - 6y$ başlangıç noktasını $x_0 = [1,1]^T$ alarak en hızlı düşüş algoritmasıyla yapınız. $\|\nabla f\| \leq 0.3$ olduğunda işlemi durdurunuz.

$$f(x,y) = \frac{1}{2} \underbrace{\begin{bmatrix} x & y \end{bmatrix}}_{x'} \underbrace{\begin{bmatrix} 4 & 2 \\ 2 & 2 \end{bmatrix}}_{Q} \underbrace{\begin{bmatrix} x \\ y \end{bmatrix}}_{x} - \underbrace{\begin{bmatrix} x, y \end{bmatrix}}_{x'} \underbrace{\begin{bmatrix} 4 \\ 6 \end{bmatrix}}_{b}$$

$f(x,y) = \frac{1}{2} x' Q x - x' b$ şeklinde yazıldığı için kuadratik fonksiyondur.

$$\nabla f = \begin{bmatrix} 4x + 2y - 4 \\ 4y + 2x - 6 \end{bmatrix} = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix}$$

$$\nabla^2 f = \begin{bmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial y^2} \end{bmatrix} = \begin{bmatrix} 4 & 2 \\ 2 & 4 \end{bmatrix}$$

$\nabla f(x_0) = g_0 = \begin{bmatrix} 2 \\ 0 \end{bmatrix}$ $\quad \|g_0\| = \|\nabla f(x_0)\| = \sqrt{2^2 + 0^2} = 2$

$2 \geq 0.3$ olduğu için devam.

$$\alpha_0 = \frac{g_0' g_0}{g_0' Q g_0} = \frac{[2,0]\begin{bmatrix} 2 \\ 0 \end{bmatrix}}{[2,0]\begin{bmatrix} 4 & 2 \\ 2 & 4 \end{bmatrix}\begin{bmatrix} 2 \\ 0 \end{bmatrix}} = \frac{4}{[8,4]\begin{bmatrix} 2 \\ 0 \end{bmatrix}} = \frac{4}{16}$$

$$= \frac{1}{4}$$

4

$$x_1 = x_0 - \alpha_0 g_0$$

$$x_1 = \begin{bmatrix}1\\1\end{bmatrix} - \frac{1}{4}\begin{bmatrix}2\\0\end{bmatrix} = \begin{bmatrix}1\\1\end{bmatrix} - \begin{bmatrix}1/2\\0\end{bmatrix} = \begin{bmatrix}1/2\\1\end{bmatrix}$$

$$g_1 = \begin{bmatrix}4 \cdot 1/2 + 2 \cdot 1 - 4\\ 4 \cdot 1 + 2 \cdot 1/2 - 6\end{bmatrix} = \begin{bmatrix}0\\-1\end{bmatrix} \qquad g_1 = \sqrt{0^2 + (-1)^2} = 1$$

$1 \geq 0.3$ devam.

$$\alpha_1 = \frac{g_1' g_1}{g_1' Q g_1} = \frac{[0,-1]\begin{bmatrix}0\\-1\end{bmatrix}}{[0,-1]\begin{bmatrix}4 & 2\\2 & 4\end{bmatrix}\begin{bmatrix}0\\-1\end{bmatrix}} = \frac{1}{[-2,-4]\begin{bmatrix}0\\-1\end{bmatrix}} = 1/4$$

$$x_2 = x_1 - \alpha_1 g_1$$

$$x_2 = \begin{bmatrix}1/2\\1\end{bmatrix} - \frac{1}{4}\begin{bmatrix}0\\-1\end{bmatrix} = \begin{bmatrix}1/2\\5/4\end{bmatrix}$$

$$g_2 = \begin{bmatrix}4 \cdot 1/2 + 2 \cdot 5/4 - 4\\ 4 \cdot 5/4 + 2 \cdot 1/2 - 6\end{bmatrix} = \begin{bmatrix}1/2\\0\end{bmatrix}$$

$$\|\nabla f_2\| = \|g_2\| = \sqrt{1/4 + 0^2} = 1/2 = 0.5 > 0.3 \text{ Devam.}$$

$$\alpha_2 = \frac{g_2' g_2}{g_2' Q g_2} = \frac{[1/2, 0]\begin{bmatrix}1/2\\0\end{bmatrix}}{[1/2, 0]\begin{bmatrix}4 & 2\\2 & 4\end{bmatrix}\begin{bmatrix}1/2\\0\end{bmatrix}} = \frac{1/4}{[2\ 1]\begin{bmatrix}1/2\\0\end{bmatrix}}$$

$$x_3 = x_2 - \alpha_2 g_2 \qquad = \frac{1/4}{1} = 1/4$$

$$x_3 = \begin{bmatrix}1/2\\5/4\end{bmatrix} - 1/4\begin{bmatrix}1/2\\0\end{bmatrix} = \begin{bmatrix}1/2\\5/4\end{bmatrix} - \begin{bmatrix}+1/8\\0\end{bmatrix} = \begin{bmatrix}3/8\\5/4\end{bmatrix}$$

$\boxed{5}$ $g_3 = \begin{bmatrix} 4x+2y-4 \\ 4y+2x-6 \end{bmatrix} = \begin{bmatrix} 4.3/8+2.5/4-4 \\ 4.5/4+2.3/8-6 \end{bmatrix} = \begin{bmatrix} \frac{6}{4}+\frac{10}{4}-4 \\ \frac{20}{4}+3/4-6 \end{bmatrix}$

$$= \begin{bmatrix} 0 \\ -1/4 \end{bmatrix}$$

$\|g_3\| = \|\nabla f(x_3)\| = \sqrt{0^2+(1/4)^2} = 1/4 = 0.25 < 0.3$

$x_3 = \begin{bmatrix} 3/8 \\ 5/4 \end{bmatrix}$ aranan çözümdür.

Açık ve kapalı kümeler

$X \subset \mathbb{R}^n$ olsun.

Tanım $p \in X$ ve $\forall \varepsilon \in \mathbb{R}^+$ için

$d(p,x) < \varepsilon$ eşitsizliğini sağlayan $\forall x \in X$ ise p ye X'in bir iç noktası denir.

Tanım X'in bütün noktaları iç nokta ise X'e açık küme denir.

Tanım: $\forall x \in X$ ve $\forall \varepsilon \in \mathbb{R}^+$ x'in komşuluğunun en az bir noktası X'de ise x'e kapanım noktası denir. kapanım noktalarının oluşturduğu kümeye X'in kapanışı denir.

Tanım X'in kapanışı kendisine eşit ise X'e kapalı küme denir.

6

Metrik: X vektör uzayında $\forall x,y \in X$ için tanımlanan reel değerli $d(x,y)$ fonksiyonu aşağıdaki şartları sağlarsa bir metrik belirler.

1) $\forall x,y \in X$ için $d(x,y) \geq 0$ dır.

2) $d(x,y)=0 \Leftrightarrow x=y$

3) $\forall x,y \in X$ için $d(x,y)=d(y,x)$ Simetri

4) $\forall x,y,z \in X$ için

$d(x,z) \leq d(x,y)+d(y,z)$

Tanım: üzerinde metrik tanımlanan uzaya metrik uzay denir. Ve (X,d) ile gösterilir.

Tanım konveks kümede herhangi farklı iki noktanın konveks kombinezonu olarak yazılamayan noktalara uç nokta denir.

Norm: $\| \ \|: X \to \mathbb{R}$ şeklinde tanımlanır.

$x \to \|x\|$

fonksiyondur. Fonksiyonun norm belirtmesi için aşağıdaki şartları sağlaması gerek ve yeterdir.

1) $\|x\| \geq 0$ $\|x\|=0 \Leftrightarrow x=0$

2) $\forall x,y \in X$ için $\|x+y\| \leq \|x\| + \|y\|$

3) $\forall x \in X$ $\alpha \in \mathbb{R}$ için $\|\alpha x\| = |\alpha| \|x\|$ dir

Genel olarak $\mathbb{R}^n$'de L_p normu

$\|x\|_p = \left[\sum_{i=1}^{n} |x_i|^p \right]^{1/p}$ ile ilgileneceğiz

$\boxed{7}$

$p = 1 \Rightarrow \|x\|_1 = \sum_{i=1}^{n} |x_i| = |x_1| + |x_2| + \dots + |x_n|$

$p = 2 \Rightarrow \|x\|_2 = \left[\sum_{i=1}^{n} x_i^2\right]^{1/2} = \sqrt{x_1^2 + x_2^2 + \dots + x_n^2}$

Norm, vektörün büyüklüğünü, Metrik iki vektör arasındaki uzaklığı gösterir.

<u>Eşlenik yön (Gradyant) Algoritması</u>

$x_0 \in \mathbb{R}^n$ seç $\nabla f(x_0) = g_0$ hesapla $d_0 = -g_0$ al

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

2) $\alpha_k = -\dfrac{d_k' g_k}{d_k' Q d_k}$ yı hesapla

3) $d_{k+1} = -g_{k+1} + \beta_k d_k$

4) $\beta_k = \dfrac{g'_{k+1} Q d_k}{d_k' Q d_k}$

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

$f(x,y) = x^2 - xy + y^2 - 2x - y$ problemini eşlenik yön algoritması ile yapalım. $x_0 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ seçilsin

$d_0 = -g_0$ $\qquad \nabla f = \begin{bmatrix} 2x - 2 \\ -x + 2y - 1 \end{bmatrix} = \begin{bmatrix} -2 \\ -1 \end{bmatrix} = \nabla f(x_0) = g_0$

x_0'ı koyarsak

$\boxed{8}$

$d_0 = \begin{bmatrix}2\\1\end{bmatrix}$ $\quad \alpha_0 = \dfrac{-d_0' g_0}{d_0' Q d_0} \Rightarrow \alpha_0 = \dfrac{-[2,1]\begin{bmatrix}-2\\-1\end{bmatrix}}{[2,1]\begin{bmatrix}2 & -1\\-1 & 2\end{bmatrix}\begin{bmatrix}2\\1\end{bmatrix}}$

$\nabla^2 f = \begin{bmatrix}\frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y}\\ \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial y^2}\end{bmatrix}$

$Q = \begin{bmatrix}2 & -1\\-1 & 2\end{bmatrix}$ $\quad \alpha_0 = \dfrac{5}{[3,0]\begin{bmatrix}2\\1\end{bmatrix}} = \dfrac{5}{6}$

$x_1 = x_0 + \alpha_0 d_0 = \begin{bmatrix}0\\0\end{bmatrix} + \frac{5}{6}\begin{bmatrix}2\\1\end{bmatrix} = \begin{bmatrix}\frac{10}{6}\\ \frac{5}{6}\end{bmatrix}$

$g_1 = \nabla f(x_1) = \begin{bmatrix}2x - y - 2\\ -x + 2y - 1\end{bmatrix}$

$g_1 = \begin{bmatrix}\frac{20}{6} - \frac{5}{6} - 2\\ -\frac{10}{6} + \frac{10}{6} - 1\end{bmatrix} = \begin{bmatrix}1/2\\-1\end{bmatrix}$

$\beta_0 = \dfrac{g_1' Q d_0}{d_0' Q d_0} = \dfrac{[1/2\ -1]\begin{bmatrix}2 & -1\\-1 & 2\end{bmatrix}\begin{bmatrix}2\\1\end{bmatrix}}{[2,1]\begin{bmatrix}2 & -1\\-1 & 2\end{bmatrix}\begin{bmatrix}2\\1\end{bmatrix}}$

$\beta_0 = \dfrac{[2,\ -5/2]\begin{bmatrix}2\\1\end{bmatrix}}{[3,0]\begin{bmatrix}2\\1\end{bmatrix}} = \dfrac{4 - 5/2}{6} = \dfrac{3/2}{6} = 1/4$

$d_1 = -g_1 + \beta_0 d_0$

$d_1 = \begin{bmatrix}-1/2\\1\end{bmatrix} + 1/4\begin{bmatrix}2\\1\end{bmatrix}$

$d_1 = \begin{bmatrix}0\\5/4\end{bmatrix}$

$x_2 = x_1 + \alpha_1 d_1$

$x_2 = \begin{bmatrix}\frac{10}{6}\\ \frac{5}{6}\end{bmatrix} + (-1)\begin{bmatrix}0\\ \frac{5}{4}\end{bmatrix}$

$= \begin{bmatrix}10/6\\ 5/6 - \frac{5}{4}\end{bmatrix} = \begin{bmatrix}5/3\\ .\end{bmatrix}$

$$\alpha_1 = \frac{-d_1' g_1}{d_1' Q d_1} = \frac{-\begin{bmatrix} 0, \frac{5}{4} \end{bmatrix} \begin{bmatrix} 1/2 \\ -1 \end{bmatrix}}{\begin{bmatrix} 0, 5/4 \end{bmatrix} \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} 0 \\ 5/4 \end{bmatrix}}$$

$$= \frac{+5/4}{\underbrace{\begin{bmatrix} -\frac{5}{4} & +\frac{5}{2} \end{bmatrix}}_{5/4} \begin{bmatrix} 0 \\ 5/4 \end{bmatrix}} = \frac{+5/4}{+25/8} = \frac{5}{4} \cdot \frac{8}{25} = +2/5$$

$$X_2 = X_1 + \alpha_1 d_1$$

$$X_2 = \begin{bmatrix} 10/6 \\ 5/6 \end{bmatrix} + \frac{2}{5} \begin{bmatrix} 0 \\ 5/4 \end{bmatrix} = \begin{bmatrix} 10/6 \\ \frac{5}{6} + \frac{1}{2} \end{bmatrix} = \begin{bmatrix} 10/6 \\ 8/6 \end{bmatrix}$$

$$X_2 = \begin{bmatrix} 5/3 \\ 4/3 \end{bmatrix}$$

$$\nabla f(x_2) = \begin{bmatrix} 2x - y - 2 \\ -x + 2y - 1 \end{bmatrix} = \begin{bmatrix} 2 \cdot 5/3 - 4/3 - 2 \\ -5/3 + 8/3 - 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

problem optimuma ulaştı.