

Insinöörimatematiikka: Usean muuttujan funktio 1

Mika Hirvensalo
mikhirve@utu.fi

Matematiikan ja tilastotieteen laitos
Turun yliopisto

2023

Ääriarvot

Funktion $f : \mathbb{R}^n \rightarrow \mathbb{R}$ ääriarvopisteissä $\mathbf{a}$ on $\nabla f(\mathbf{a}) = 0$.
Käsitteeseen liittyy että $\mathbf{a}$ ei ole tarkastelujoukon reunapiste.

Ääriarvot

Funktion $f : \mathbb{R}^n \rightarrow \mathbb{R}$ ääriarvopisteissä $\mathbf{a}$ on $\nabla f(\mathbf{a}) = 0$.
Käsitteeseen liittyy että $\mathbf{a}$ ei ole tarkastelujoukon reunapiste.

Kriittinen piste

Piste $\mathbf{a} \in \mathbb{R}^n$ on funktion $f : \mathbb{R}^n \rightarrow \mathbb{R}$ kriittinen piste jos $\nabla f(\mathbf{a}) = 0$. Myös tässä oletetaan, että $\mathbf{a}$ ei ole tarkastelujoukon reunapiste.

Ääriarvot

Funktion $f : \mathbb{R}^n \rightarrow \mathbb{R}$ ääriarvopisteissä $\mathbf{a}$ on $\nabla f(\mathbf{a}) = 0$.
Käsitteeseen liittyy että $\mathbf{a}$ ei ole tarkastelujoukon reunapiste.

Kriittinen piste

Piste $\mathbf{a} \in \mathbb{R}^n$ on funktion $f : \mathbb{R}^n \rightarrow \mathbb{R}$ kriittinen piste jos $\nabla f(\mathbf{a}) = 0$. Myös tässä oletetaan, että $\mathbf{a}$ ei ole tarkastelujoukon reunapiste.

Esimerkki 4.4.

(J.Lahtonen)

Vertaa:

$$f(x_0 + h) = f(x_0) + f'(x_0)h + h\epsilon_1(h)$$

Vertaa:

$$\begin{aligned}f(x_0 + h) &= f(x_0) + f'(x_0)h + h\epsilon_1(h) \\ &= f(x_0) + f'(x_0)h + \frac{1}{2}f''(x_0)h^2 + h^2\epsilon_2(h)\end{aligned}$$

Vertaa:

$$\begin{aligned}f(x_0 + h) &= f(x_0) + f'(x_0)h + h\epsilon_1(h) \\ &= f(x_0) + f'(x_0)h + \frac{1}{2}f''(x_0)h^2 + h^2\epsilon_2(h)\end{aligned}$$

$f : \mathbb{R}^n \rightarrow \mathbb{R}$

$$f(\mathbf{x}_0 + \mathbf{h}) = f(\mathbf{x}_0) + \nabla f(\mathbf{x}_0) \cdot \mathbf{h} + \|\mathbf{h}\|\epsilon_1(\mathbf{h})$$

Vertaa:

$$\begin{aligned}f(x_0 + h) &= f(x_0) + f'(x_0)h + h\epsilon_1(h) \\ &= f(x_0) + f'(x_0)h + \frac{1}{2}f''(x_0)h^2 + h^2\epsilon_2(h)\end{aligned}$$

$f : \mathbb{R}^n \rightarrow \mathbb{R}$

$$f(\mathbf{x}_0 + \mathbf{h}) = f(\mathbf{x}_0) + \nabla f(\mathbf{x}_0) \cdot \mathbf{h} + \|\mathbf{h}\|\epsilon_1(\mathbf{h})$$

Korkeamman asteen approksimaatio?

Toisen asteen approksimaatio

$$f(\mathbf{x}_0 + \mathbf{h}) = f(\mathbf{x}_0) + \nabla f(\mathbf{x}_0) \cdot \mathbf{h} + \frac{1}{2} \mathbf{h}^T H_f(\mathbf{x}_0) \mathbf{h} + \|\mathbf{h}\|^2_{\epsilon_2}(\mathbf{h})$$

Toisen asteen approksimaatio

$$f(\mathbf{x}_0 + \mathbf{h}) = f(\mathbf{x}_0) + \nabla f(\mathbf{x}_0) \cdot \mathbf{h} + \frac{1}{2} \mathbf{h}^T H_f(\mathbf{x}_0) \mathbf{h} + \|\mathbf{h}\|^2_{\epsilon_2(\mathbf{h})},$$

missä

$$H_f = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1 \partial x_1} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \frac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_n} \end{pmatrix}$$

on ns. Hessen matriisi.

Toisen asteen approksimaatio

$$f(\mathbf{x}_0 + \mathbf{h}) = f(\mathbf{x}_0) + \nabla f(\mathbf{x}_0) \cdot \mathbf{h} + \frac{1}{2} \mathbf{h}^T H_f(\mathbf{x}_0) \mathbf{h} + \|\mathbf{h}\|^2_{\epsilon_2(\mathbf{h})},$$

missä

$$H_f = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1 \partial x_1} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \frac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_n} \end{pmatrix}$$

on ns. Hessen matriisi. Joissain kirjoissa merkitään $H(\mathbf{x}_0) = \nabla^2 f(\mathbf{x}_0)$

Määritelmä

Matriisi H on

- Positiividefiniitti, jos $\mathbf{h}^T H \mathbf{h} > 0$ aina kun $\mathbf{h} \neq 0$

Määritelmä

Matriisi H on

- Positiividefiniitti, jos $\mathbf{h}^T H \mathbf{h} > 0$ aina kun $\mathbf{h} \neq 0$
- Negatiividefiniitti, jos $\mathbf{h}^T H \mathbf{h} < 0$ aina kun $\mathbf{h} \neq 0$

Määritelmä

Matriisi H on

- Positiividefiniitti, jos $\mathbf{h}^T H \mathbf{h} > 0$ aina kun $\mathbf{h} \neq 0$
- Negatiividefiniitti, jos $\mathbf{h}^T H \mathbf{h} < 0$ aina kun $\mathbf{h} \neq 0$
- Indefiniitti, jos $\mathbf{h}^T H \mathbf{h}$ saa sekä positiivisia että negatiivisia arvoja.

Määritelmä

Matriisi H on

- Positiividefiniitti, jos $\mathbf{h}^T H \mathbf{h} > 0$ aina kun $\mathbf{h} \neq 0$
- Negatiividefiniitti, jos $\mathbf{h}^T H \mathbf{h} < 0$ aina kun $\mathbf{h} \neq 0$
- Indefiniitti, jos $\mathbf{h}^T H \mathbf{h}$ saa sekä positiivisia että negatiivisia arvoja.

Termit positiivisemidefiniitti ja negatiivisemidefiniitti tarkoittavat, että $\mathbf{h}^T H \mathbf{h}$ voi olla nolla vaikka $\mathbf{h} \neq 0$, mutta H ei ole indefiniitti.

Lause

Jos funktion f kriittisessä pisteessä $\mathbf{x}_0$ Hessen matriisi $H_f(\mathbf{x}_0)$ on

- positiividefiniitti, on piste lokaali minimi

Lause

Jos funktion f kriittisessä pisteessä $\mathbf{x}_0$ Hessen matriisi $H_f(\mathbf{x}_0)$ on

- positiividefiniitti, on piste lokaali minimi
- negatiividefiniitti, on piste lokaali maksimi

Lause

Jos funktion f kriittisessä pisteessä $\mathbf{x}_0$ Hessen matriisi $H_f(\mathbf{x}_0)$ on

- positiividefiniitti, on piste lokaali minimi
- negatiividefiniitti, on piste lokaali maksimi
- indefiniitti, on piste satulapiste

Ääriarvojen laatu

Lause

Jos funktion f kriittisessä pisteessä $\mathbf{x}_0$ Hessen matriisi $H_f(\mathbf{x}_0)$ on

- positiividefiniitti, on piste lokaali minimi
- negatiividefiniitti, on piste lokaali maksimi
- indefiniitti, on piste satulapiste

Jos Hessen matriisi on semidefiniitti, ei ääriarvon laatu selviä tällä tavalla.

Todistuksen idea:

$$f(\mathbf{x}_0 + \mathbf{h}) = f(\mathbf{x}_0) + \underbrace{\nabla f(\mathbf{x}_0) \cdot \mathbf{h}}_0 + \frac{1}{2} \mathbf{h}^T H_f(\mathbf{x}_0) \mathbf{h} + \|\mathbf{h}\|^2 \epsilon_2(\mathbf{h})$$

Ääriarvojen laatu

Lause

Jos funktion f kriittisessä pisteessä $\mathbf{x}_0$ Hessen matriisi $H_f(\mathbf{x}_0)$ on

- positiividefiniitti, on piste lokaali minimi
- negatiividefiniitti, on piste lokaali maksimi
- indefiniitti, on piste satulapiste

Jos Hessen matriisi on semidefiniitti, ei ääriarvon laatu selviä tällä tavalla.

Todistuksen idea:

$$f(\mathbf{x}_0 + \mathbf{h}) = f(\mathbf{x}_0) + \underbrace{\nabla f(\mathbf{x}_0) \cdot \mathbf{h}}_0 + \frac{1}{2} \mathbf{h}^T H_f(\mathbf{x}_0) \mathbf{h} + \|\mathbf{h}\|^2 \epsilon_2(\mathbf{h}),$$

Definiittisyyden selvittäminen

Symmetrinen matriisi on diagonalisoituva

Definiittisyyden selvittäminen

Symmetrinen matriisi on diagonalisoituva
Sylvesterin kriteeri

Definiittisyyden selvittäminen

Symmetrinen matriisi on diagonalisoituva

Sylvesterin kriteeri

Esimerkkejä