KERTAUSTA

1 Gradientti — sen tulkinta
Suunnattu derivaatta:
— normalisointi on välttämätön

2 Ketjusäännöt

3 Pistеapproksimaatiot
— linearisointi
— Taylorin polynomi

4 Rajoiteoptimointi
— Lagrangen kertoimet
→ epälineaarinen yhtälöryhmä

5 Taso- ja avaruusintegraali
— muuttujanvaihto
— koordinaatistomuunnokset

6 Massakeskipiste / Hitausmomentti
— geometrinen hahmotus

Ensimmäiset sanat:

Eräs koe (ei välttämättä meidän)

1 Olkoon
2 Määritä
3 Funktioista
4 Määritä
5 Tarkastellaan
  a) Muodosta
  b) Laske
6 Olkoon

Yhtälöryhmän ratkaisu gradienttimenetelmällä

$Ax = b$; A symmetrinen ja pos. def.

Yhtälöryhmän ratkaisu: x_*

Alkuperäinen tehtävä voidaan esittää minimoimistehtävänä:

Etsi funktion

$$\phi(x) = \frac{1}{2} x^T A x - x^T b$$

minimi!

Huomaa: $\nabla \phi(x) = Ax - b$

Merkitään: x_c on ääriarvopisteen approksimaatio

Kysymys: Kuinka arvioida menetelmän laatua (ja tarkkuutta)?

Valitaan normi: $\|v\|_A = \sqrt{v^T A v}$

$$\phi(x_c) = \frac{1}{2} x_c^T A x_c - x_c^T b \qquad (*)$$

$$= \frac{1}{2}(x_c - x_*)^T A (x_c - x_*) \qquad (**)$$

$$- \frac{1}{2} b^T A^{-1} b$$

$$= \underbrace{\frac{1}{2}\|x_c - x_*\|_A^2}_{} + \phi(x_*)$$

Jos iteraatio suppenee, niin normin neliö $\longrightarrow 0$ iteraation edetessä.

Idea: Kuljetaan aina negatiivisen gradientin suuntaan!

$(*)$ $b = A x_*$

$(**)$ $x_* = A^{-1} b$

Esim. $A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$

$$\phi(x) = \frac{1}{2} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}^T \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

$$= \frac{1}{2} (x_1^2 + x_2^2)$$

$$\Rightarrow \nabla \phi(x) = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = A \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

Palataan iteraatioon:

Uusi iteraatti: $x_+ = x_c - \mu_c g_c$,

missä $g_c = Ax_c - b$ ja μ_c on ns. askelpituus, jolla $\phi(x_c - \mu_c g_c)$ minimoituu.

Suora lasku:

$$\phi(x_c - \mu_c g_c) = \frac{1}{2} (x_c - \mu_c g_c)^T A (x_c - \mu_c g_c) - (x_c - \mu_c g_c)^T b$$

Derivointi μ_c:n suhteen:

$$\phi'(x_c - \mu_c g_c) = 0$$

$$\Longleftrightarrow \mu_c = \frac{g_c^T g_c}{g_c^T A g_c}$$

$$= \underbrace{\frac{1}{2} x_c^T A x_c - x_c^T b}_{\phi(x_c)}$$

$$- \frac{1}{2} x_c^T A \mu_c g_c - \frac{1}{2} \mu_c g_c^T A x_c \leftarrow$$

$$+ \frac{1}{2} \mu_c^2 g_c^T A g_c + \mu_c g_c^T b$$

$$= \phi(x_c) - \mu_c g_c^T \underbrace{(A x_c - b)}_{g_c} + \frac{1}{2} \mu_c^2 g_c^T A g_c$$

$$= \phi(x_c) - \frac{1}{2} \frac{(g_c^T g_c)^2}{g_c^T A g_c}, \text{ valinnalle}$$

$$\underline{\underline{\mu_c}} = \frac{g_c^T g_c}{g_c^T A g_c}$$

Kyseessä on ns. viivahaku.

Suppenemisnopeus?

$$\kappa_c = \frac{g_c^T A g_c}{g_c^T g_c} \cdot \frac{g_c^T A^{-1} g_c}{g_c^T g_c} \le \frac{\lambda_{max}(A)}{\lambda_{min}(A)}$$

$$= \kappa_2(A)$$