

Inversio-ongelmien laskennallinen peruskurssi

Luento 2

Kevät 2012

1 Lineaarinen inversio-ongelma

Määritelmä 1.1. Yleinen (reaaliarvoinen) lineaarinen inversio-ongelma voidaan esittää muodossa

$$m = Ax + \varepsilon, \quad (1)$$

missä vektoria $x \in \mathbb{R}^N$ kutsutaan mittaukseksi, matriisia $A \in \mathbb{R}^{N \times M}$ on niin sanottu teoriamatriisi, $x \in \mathbb{R}^M$ on tuntematon ja $\varepsilon \in \mathbb{R}^N$ on mittausvirhe.

Ongelmana on löytää tarkka tai “paras mahdollinen” arvo tuntemattomalle x , kun matriisi A ja mittaus m on annettu, ja virhe ε on joko tunnettu kokonaan tai osittain (esim. joitakin virheen tilastollisia ominaisuuksia tunnetaan).

Riippuen luvuista N ja M , sekä virheen ε ja matriisin A luonteesta, seuraavat vaihtoehdot ovat mahdollisia:

1. Jos $M = N$, ja virhe ε on tunnettu (tällöin voidaan asettaa $\varepsilon = 0$), on yhtälö (1) tavallinen lineaarinen yhtälöryhmä, jossa on N tuntematonta ja $M = N$ yhtälöä. Sillä on yksikäsitteinen ratkaisu, jos ja vain jos matriisi A on kääntyvä, eli $\det(A) \neq 0$.

Jos A on singulaarinen, eli $\det(A) = 0$, on sillä nollasta eroava ydin (kernel), $\ker(A) \neq 0$. Tällöin mittaus m joko kuuluu tai ei kuulu matriisin kuva-avaruuteen (range).

Jos $m \in \text{Ran}(A)$, on yhtälöllä (1) ääretön määrä ratkaisuja. Tällöin yleinen tapa ratkaista (1) on yrittää löytää (jonkin vektorinormin mielessä) lyhyin ratkaisu $\hat{x}$.

Jos taas $m \notin \text{Ran}(A)$, ei yhtälöllä (1) ole lainkaan eksaktia ratkaisua. Tällöin ratkaisu esitetään yleisesti ns. pienimmän neliösumman mielessä, eli ratkaisua edustaa vektori $\tilde{x}$, joka minimoi funktionaalin

$$\|Ax - m\|.$$

2. Jos $N < M$ ja ε on tunnettu (jälleen voidaan asettaa $\varepsilon = 0$), os yhtälössä (1) enemmän tuntemattomia kuin yhtälöitä, eli systeemi on alideterminoitu. Tällöin se voidaan palauttaa kohdan 1 tapaukseen lisäämällä matriisiin A $M - N$ riviä nollia ja vastaavasti mittausvektoriin m $M - N$ nollaa. Näin saatu matriisi A on singulaarinen, ja ratkaisujen olemassaolo sekä määrä riippuu siitä, kuuluuko m matriisin A kuva-avaruuteen vai ei.
3. Jos $N > M$ ja ε on tunnettu ($\varepsilon = 0$), eli yhtälöitä on enemmän kuin tuntemattomia, on yhtälö (1) ylideterminoitu, paitsi jos A on degeneroitunut ($\text{rank}(A) < M$) siten, että se palautuu kohtaan 1 tai 2. Tässä tapauksessa ratkaisu on olemassa ainoastaan pienimmän neliösumman mielessä.
4. Jos virhettä ε ei tunneta, mutta sen komponenttien tiedetään olevan satunnaismuuttujia, joiden yhteisjakauma on tunnettu, kutsutaan yhtälöä (1) lineaariseksi stokastiseksi (tai tilastolliseksi) inversio-ongelmaksi. Sen ratkaisu annetaan tuntemattoman x yhteis- ja posteriori-jakaumana (joint a posteriori distribution). Käytännössä usein riittää laskea ns. MAP-estimaatti (maximum a posteriori estimate) ja a posteriori kovarianssi.

2 Moore-Penrose-inverssi

Voidaan osoittaa, että jokaista $M \times N$ -matriisia A kohti on olemassa yksikäsitteinen $N \times M$ -matriisi B , jolle pätee:

1. $ABA = A$,
2. $BAB = B$,
3. $(AB)^T = AB$,
4. $(BA)^T = BA$.

Tätä matriisia B kutsutaan pseudoinverssiksi tai Moore-Penrose-inverssiksi ja sitä merkitään symbolilla A^+ .

Huomautus 2.1. Tärkeitä erikoistapauksia:

1. Jos $A^T A$ on ei-singulaarinen (eli kääntyvä), niin

$$A^+ = (A^T A)^{-1} A^T.$$

2. Jos AA^T on ei-singulaarinen, niin

$$A^+ = A^T (AA^T)^{-1}.$$

3. Jos A on ei-singulaarinen, niin

$$A^+ = A^{-1}.$$

Tarkastellaan lineaarista yhtälöryhmää

$$m = Ax.$$

Voidaan osoittaa, että

1. Jos yhtälöryhmällä on yksi tai useampi ratkaisu, niin

$$\tilde{x} = A^+ m$$

on yhtälöryhmän miniminormiratkaisu.

2. Jos yhtälöryhmällä ei ole ratkaisua, niin

$$\tilde{x} = A^+ m$$

minimoi funktionaalin $\|Ax - m\|$, eli $\tilde{x}$ on yhtälöryhmän ratkaisu pienimmän neliösumman mielessä.

3 Matriisin normeista

Olkoon $A \in \mathbb{R}^{m \times n}$ matriisi. Matriisinormille $\|A\|$ pätee

- $\|A\| > 0$ if and only if $A \neq 0$, ja $\|A\| = 0$ jos ja vain jos $A = 0$.
- $\|\alpha A\| = |\alpha| \|A\|$ kaikilla $\alpha \in \mathbb{R}$.
- $\|A + B\| \leq \|A\| + \|B\|$ kaikille matriiseille $A, B \in \mathbb{R}^{m \times n}$.
- $\|AB\| \leq \|A\| \|B\|$ kun matriisit $A, B \in \mathbb{R}^{n \times n}$ (ts. A ja B ovat neliömatriiseja).

Jos vektorinormi avaruudessa $\mathbb{R}^n$ on annettu, niin ns. indusoitu matriisinormi määritellään kaavalla

$$\begin{aligned} \|A\| &= \max \{ \|Ax\| : x \in \mathbb{R}^n, \|x\| = 1 \} \\ &= \max \left\{ \frac{\|Ax\|}{\|x\|} : x \in \mathbb{R}^n, x \neq 0 \right\}. \end{aligned}$$

Seuraavassa on lueteltuna muutamia paljon käytettyjä matriisinormeja ($A \in \mathbb{R}^{m \times n}$):

- 1-normi (maksimi sarakesumma):

$$\|A\|_1 = \max_{1 \leq j \leq n} \sum_{i=1}^m |a_{ij}|$$

2. ∞ -normi (maksimi rivisumma):

$$\|A\|_\infty = \max_{1 \leq i \leq m} \sum_{j=1}^n |a_{ij}|$$

3. 2-normi:

$$\|A\|_2 = \sqrt{\lambda_{\max}(A^T A)},$$

missä $\lambda_{\max}(A^T A)$ on matriisin $A^T A$ suurin ominaisarvo.

4. Fröbenius-normi:

$$\|A\|_F = \sqrt{\sum_{i=1}^m \sum_{j=1}^n |a_{ij}|^2} = \sqrt{\text{tr}(A^T A)} = \sqrt{\sum_{i=1}^{\min(m,n)} \sigma_i^2},$$

missä $\text{tr}(A^T A)$ on matriisin $A^T A$ trace ja σ_i , $i = 1, 2, \dots, \min(m, n)$ ovat matriisin A singulaariarvot.

4 Matriisin ehtoluvusta (Condition number)

Olko $A \in \mathbb{R}^{n \times n}$ neliömatriisi, ja olko $Ax = m + \varepsilon$, missä ε on mitausvirhe. Jos A on kääntyvä, niin

$$x = A^{-1}(m + \varepsilon) = A^{-1}m + A^{-1}\varepsilon,$$

eli virheen osuus ratkaisussa x on $A^{-1}\varepsilon$. Nyt ratkaisun suhteellisen virheen suhde mittauksen suhteelliseen virheeseen on

$$\frac{\|A^{-1}\varepsilon\|/\|A^{-1}m\|}{\|\varepsilon\|/\|m\|} = \frac{\|A^{-1}\varepsilon\|}{\|\varepsilon\|} \left(\frac{\|A^{-1}m\|}{\|m\|} \right)^{-1},$$

ja on suhteellisen helppo havaita, että tämän suhteellisen virheen maksimi-arvo on (kun oletetaan, että $\varepsilon, m \neq 0$)

$$\|A\| \|A^{-1}\|.$$

Matriisin A ehtoluku $\kappa(A)$ määritellään kaavalla

$$\kappa(A) = \|A\| \|A^{-1}\|,$$

ja se siis antaa ratkaisun ja mittauksen suhteellisten virheiden maksimi-arvon. Mitä korkeampi ehtoluku matriisilla on, sitä herkempi yhtälö $m + \varepsilon = Ax$ on virheelle ε . Huomaa, että

$$\kappa(A) = \|A\| \|A^{-1}\| \geq \|I\| = 1.$$

Ehtoluku riippuu myös käytetystä matriisinormista. 2-normin tapauksessa voidaan osoittaa, että

$$\kappa(A) = \|A\|_2 \|A^{-1}\|_2 = \frac{\sigma_{\max}(A)}{\sigma_{\min}(A)},$$

missä $\sigma_{\max}(A)$ on matriisin A suurin ja $\sigma_{\min}(A)$ matriisin pienin singulaari-arvo.